UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

FACULTAD DE CIENCIAS

Advanced blind decomposition methods for end-members and abundance estimation in medical imaging

a dissertation submitted for the degree
of
Doctor of Engineering Sciences

Present:
M.Eng. Inés Alejandro Cruz Guerrero
Advisor:
Ph.D. Daniel Ulises Campos Delgado

San Luis Potosí, S.L.P., September 2023

“Con fundamento en los artículos 21 y 27 de la Ley Federal del Derecho de Autor y como titular de los derechos moral y patrimonial de la obra titulada “Métodos avanzados de descomposición ciega para la estimación de perfiles y abundancias en imagenología médica”, Tesis para obtener el grado de Doctor en Ciencias de la Ingeniería, otorgo de manera gratuita y permanente a la Facultad de Ciencias de la Universidad Autónoma de San Luis Potosí, la autorización para que difunda la obra en cualquier medio, incluido el electrónico, y la divulgue entre sus usuarios, profesores, estudiantes o terceras personas, sin que pueda percibir por tal divulgación una contraprestación”

Fecha de presentación: _____

Fecha de autorización: _____

Nombre y Firma de AUTORES

_____
M.I.E. Inés Alejandro Cruz Guerrero
Estudiante

_____
Dr. Daniel Ulises Campos Delgado
Asesor

Abstract

Hyperspectral imaging (HSI) capture a wide range of spectral bands across the electromagnetic spectrum, including both the visible range and beyond human perception. These images contain valuable information about the scenes captured by the optical sensors. By using this information in conjunction with classification algorithms, it is possible to determine the material or substance present in each pixel of the image. One of the main benefits of this technology lies in its versatility, as it can be used as a visual assistance tool in various areas, from industrial applications to the medical field. It is precisely in this medical domain where HSI has been applied for the classification and identification of biological tissues affected by certain pathologies, showing promising results in characterizing their spatial-spectral properties.

Despite the numerous advantages offered by HSI technology, the task of identifying pathologies through spatial-spectral information is not straightforward. This is due to the variability among the samples and the lack of distinctive spectral separability between healthy and diseased tissues. Moreover, the large volume of spectral information can lead to redundancies, as increasing the number of spectral bands does not always result in improved accuracy. Furthermore, the design, evaluation, and optimization of classification methods by HSI present a computational challenge, particularly due to the high dimensionality of the data. Furthermore, there is limited availability of HSI databases in the medical field and an even more restrictive scarcity of labeled databases in this area.

This dissertation work aims to exploit the characteristics of hyperspectral images to develop unmixing and classification algorithms, in order to provide precise localization of different components present in hyperspectral images. To achieve this goal, spectral unmixing methodologies were developed, considering spatial coherence and nonlinear interactions (multi-linear mixing model) among the components in the scene of interest. Additionally, hybrid classification methods were generated, combining unmixing algorithms with machine learning for hyperspectral image evaluation, to reduce computational costs and avoid overfitting. A new data calibration method was also proposed to reduce the variability in the information. In addition, state-of-the-art image processing methods were explored and adapted for hyperspectral applications.

The results of this work showed that the proposed methods allow an accurate classification of different classes of interest, outperforming state-of-the-art methods in most of the evaluated metrics. Additionally, classification maps can be generated with a higher level of agreement with the initial segmentations produced by clinical experts. Furthermore, the proposed methods reduce training and inference times, opening up the feasibility of implementing these in real-time applications.

Resumen

Las imágenes hiperespectrales (HSI) capturan una amplia gama de bandas espectrales a lo largo del espectro electromagnético, incluyendo tanto el rango visible como más allá de la percepción humana. Estas imágenes contienen información valiosa sobre las escenas capturadas por los sensores ópticos. Mediante el uso de esta información en conjunto con algoritmos de clasificación, es posible determinar el material o sustancia presente en cada píxel de la imagen. Uno de los principales beneficios de esta tecnología radica en su versatilidad, ya que puede utilizarse como herramienta de asistencia visual en diversas áreas, desde aplicaciones industriales hasta el campo médico. Precisamente en este ámbito médico es donde se ha aplicado la HSI para la clasificación e identificación de tejidos biológicos afectados por ciertas patologías, mostrando resultados prometedores en la caracterización de sus propiedades espaciales y espectrales.

A pesar de las numerosas ventajas que ofrece la tecnología HSI, la tarea de identificar patologías a través de información espacio-espectral no es sencilla. Esto se debe a la variabilidad entre las muestras y a la falta de separabilidad espectral distintiva entre tejidos sanos y enfermos. Además, el gran volumen de información espectral puede llevar a redundancias, ya que aumentar el número de bandas espectrales no siempre resulta en una mayor precisión. Aunado a esto, el diseño, la evaluación y la optimización de los métodos de clasificación mediante HSI presentan un desafío computacional, particularmente debido a la alta dimensionalidad de los datos. Asimismo, existe una disponibilidad limitada de bases de datos HSI en el campo médico y una escasez aún más restrictiva de bases de datos etiquetadas en esta área.

El objetivo de este trabajo de tesis es aprovechar las características de las HSI para desarrollar algoritmos de desmezcla y clasificación, con el fin de proporcionar una localización precisa de los diferentes componentes presentes en las mismas. Para lograr este objetivo, se desarrollaron metodologías de desmezcla espectral, considerando coherencia espacial e interacciones no lineales (modelo de mezcla multilínea) entre los componentes en la escena de interés. Además, se generaron métodos de clasificación híbridos, combinando algoritmos de desmezcla con aprendizaje automático para la evaluación de HSI, con el fin de reducir los costos computacionales y evitar el sobreajuste. Asimismo, se propuso un nuevo método de calibración de datos para reducir la variabilidad en la información. Además, se exploraron y adaptaron métodos de procesamiento de imágenes de vanguardia para aplicaciones hiperespectrales.

Los resultados de este trabajo mostraron que los métodos propuestos permiten una clasificación precisa de diferentes clases de interés, superando a los métodos de vanguardia en la mayoría de las métricas evaluadas. Además, se pueden generar mapas de clasificación con un mayor nivel de coincidencia con segmentaciones iniciales producidas por expertos clínicos. De igual forma, los métodos propuestos reducen los tiempos de entrenamiento e inferencia, abriendo la posibilidad de implementarlos en aplicaciones en tiempo real.

Dedication

To my family, for being by my side with every step I've taken in my pursuit of becoming a better person and professional.

Likewise, I dedicate this thesis to my beloved grandparents: Tomasa, Emilia, and Inés; who, unfortunately, passed away before this thesis dissertation.

Acknowledgements

First and foremost, I want to express my heartfelt gratitude to my family. Your unwavering support, valuable advice, and constant example of humility and sacrifice have been essential on my journey. Thank you for being my constant source of encouragement and, above all, for believing in me.

Furthermore, I would like to extend my gratitude to my thesis advisor, Dr. Daniel Ulises Campos Delgado, for his invaluable guidance, support, and expertise throughout my doctoral journey. His mentorship has not only allowed me to acquire knowledge in various aspects of research but also to learn about the successful execution of research projects. Additionally, his support in collaborative work with different research groups has provided me with essential tools for the successful completion of this work and for future steps in my professional career.

Furthermore, I would like to extend special thanks to Dr. Himar Fabelo, Dr. Samuel Ortega, and Prof. Gustavo Callico from the Institute of Applied Microelectronics (IUMA) at the University of Las Palmas de Gran Canaria for their invaluable support and collaboration throughout my doctoral journey, as well as during my research stay in their laboratory. This collaboration has proven to be highly fruitful for both parties and has greatly enriched my experience. My gratitude also extends to Raquel León, Laura Quintana, María Castro, Antonio Rodríguez, and Abián Hernández for their invaluable assistance during my stay.

On the other hand, I wish to extend a special gratitude to my team colleagues: Nicolás Mendoza, Erick Zavala, Liliana Granados, and Diego Polanco. Their involvement in the various projects undertaken has been invaluable. Without their invaluable collaboration, this thesis would not have reached the high level of quality it has achieved. Likewise, I would like to extend a very special thank you to Jimena García for advising me, spending time with me, and offering her support during the final stage of this thesis.

Lastly, I would like to express my heartfelt gratitude to the professors who guided and supported me throughout this journey. My appreciation extends to the National Council for the Humanities, Sciences and Technologies (Consejo Nacional de Humanidades, Ciencias y Tecnologías, CONAHCYT) for their invaluable funding of this project through a National Scholarship for the Doctoral program (Grant No. 865747). Furthermore, I acknowledge and am thankful to the Consortium of Mexican Universities (Consorcio de Universidades Mexicanas, CUMEX) and the Ibero-American University Postgraduate Association (Asociación Universitaria Iberoamericana de Postgrado, AUIP) for their generous support in granting me a joint CUMEX-AUIP scholarship, which enabled me to carry out an enriching research stay. I would also like to thank the San Luis Potosi Council for Science and Technology (Consejo Potosino de Ciencia y Tecnología, COPOCYT) and the coordination of the postgraduate program in Electronic Engineering for their essential funding that allowed me to participate in various national and international conferences, thus enhancing my academic development.

Finally, I want to express my gratitude to the University Autonomus of San Luis Potosí (Universidad Autónoma de San Luis Potosís, UASLP) that has been my academic home, which has challenged and encouraged me in equal measure, and has provided me the opportunity to achieve the PhD degree that today I obstented.

Contents

List of Tables

Table

List of Figures

Figure

List of Acronyms

Acronyms

1D-DNN 1-Dimensional Deep Neural Network.

2D-CNN 2-Dimensional Convolutional Neural Network.

ABCDE Asymmetry of the mole, Border irregularity, Color uniformity, Diameter, and Evolving size.

ABS Acrylonitrile Butadiene Styrene.

AD All Data.

ADMM Alternating Direction Method of Multipliers.

AE Autoencoder.

ANC Abundances Non-negativity Constraint.

ASC Abundance Sum-to-one Constraint.

BCC Basal Cell Carcinoma.

BG Background.

BIL Band-Interleaved-by-Line.

BIP Band-Interleaved-by-Pixel.

BLU Blind Linear Unmixing.

BSNHU Blind Sparse Nonlinear Hyperspectral Unmixing.

BSQ Band Sequential.

CCD Charge-Coupled Devices.

CCDA Cyclic Coordinate Descent Algorithm.

CDA Coordinate Descent Algorithm.

CEI Comité de Ética en la Investigación.

CEIC Comité Ético de Investigación Clínica.

CMOS Complementary Metal-Oxide Semiconductors.

CNN convolutional Neural Network.

CNS Central Nervous System.

CQO Constrained Quadratic Optimization.

CT Computed Tomography.

DBN Deep Belief Network.

DD Data Division.

DL Deep Learning.

DLF Deep Learning Framework.

EBEAE Extended Blind End-member and Abundance Estimation.

EBEAE-SC Extended Blind End-member and Abundance Extraction with Spatial Coherence.

ESC End-members Sum-of-one Constraint.

FN False Negative.

FP False Positive.

GB Glioblastoma.

GBM Generalized Bilinear Model.

GLNMF Graph-Regularized $L_{1/2}$ Nonnegative Matrix Factorization.

gtvMBO Graph Total Variation by Merriman-Bence-Osher.

HS Hyperspectral.

HSI Hyperspectral Imaging.

HT Hypervascularized Tissue.

KNN K-Nearest Neighbors.

LCTF Liquid Crystal Tunable Filter.

LEDs Light-Emitting Diodes.

LLE Local Linear Embedding.

LMM Linear Mixing Model.

LU Linear Unmixing.

MCC Matthews Correlation Coefficient.

ML Machine Learning.

MMM Multilinear Mixing Model.

MMMNSU Multilinear Mixing Model for Nonlinear Spectral Unmixing.

MRI Magnetic Resonance Imaging.

MSE Mean Squared Error.

MSI Multispectral Imaging.

NEBEAE Nonlinear Extended Blind End-member and Abundance Estimation.

NIR Near Infrared.

NLP Natural Language Processing.

NMF-QMV Nonnegative Matrix Factorization-Quadratic Minimum Volume.

NMM Nonlinear Mixing Models.

NMSC Nonmelanoma Skin Cancer.

NN Neural Network.

NT Normal Tissue.

PC Proposed Calibration.

PCA Principal Component Analysis.

PET Positron Emission Tomography.

PETG Polyethylene Terephthalate Glycol.

PISINMF Preserving the Intrinsic Structure Invariant Nonnegative Matrix Factorization.

PLA Polylactic Acid.

PPI Pixel Purity Index.

PPNM Polynomial Post-Nonlinear Model.

PSNR Peak Signal-to-Noise Ratio.

QTH Quartz Tungsten Halogen.

RBF Radial Basis Function.

ReLU Rectified Linear Unit.

ResNet Residual Networks.

RF Random Forest.

RGB Red, Green, and Blue.

RNN Recurrent Neural Network.

SAM Spectral Angle Mapper.

SC Standard Calibration.

SCC Squamous Cell Carcinoma.

SD standard deviation.

SeCoDe Sparsity-Enhanced Convolutional Decomposition.

SID Spectral Information Divergence.

SLHSI Structured Light and Hyperspectral Imager.

SLV7 Spectral Library Version 7.

SNR Signal-to-Noise Ratio.

SU Spectral Unmixing.

SVM Support Vector Machine.

SWIR Short-Wave Infrared.

t-SNE t-Distributed Stochastic Neighbor Embedding.

TN True Negative.

TP True Positive.

TT Tumor Tissue.

TV Total Variation.

UASLP Universidad Autónoma de San Luis Potosí.

UNSUBMMM Unsupervised Nonlinear Spectral Unmixing Based on MMM.

VCA Vertex Component Analysis.

ViT Vision Transformers.

VNIR Visual and Near Infrared.

WHO World Health Organization.

Chapter 1

Introduction

1.1 Overview of Hyperspectral Imaging

Cancer has been a global health problem, since according to the World Health Organization (WHO) data, in 2018 there were 18.1 million new cases and 9.8 million deaths worldwide [7]. Cancer is characterized by uncontrolled cell division in specific body regions that can later spread to other areas and encompasses various pathologies [8, 9]. Although symptoms can initially be mistaken for other conditions, physical examinations, blood tests, imaging tests, and biopsies can help to identify the presence of the disease [10]. However, a definitive diagnosis usually requires a histological examination to classify the lesion, which is necessary to determine the most appropriate treatment based on the type of cancer, its stage, and its extent [10, 11].

Medical technology is a vital part of assessing, diagnosing, and treating cancer patients. These technologies provide vital information that enables a detailed analysis using processing techniques or the ability to seek a second opinion through telemedicine technologies [12]. Among these technologies, medical imaging is one of the most informative and essential in clinical practice and research [13, 14]. Examples of medical imaging technologies include X-rays, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) scans. These technologies use various forms of radiation, such as magnetic fields or X-rays, to create detailed images of the internal parts of the body. These images enable doctors to identify and measure different tissue types, thus detecting diseases in their early stages, sometimes even before they cause symptoms [13, 15].

Although imaging techniques have advanced considerably in the detection and treatment of various tumors, it remains challenging to differentiate between healthy and diseased tissues using only the information provided by digital images [15]. Furthermore, these analyses are conducted by clinical experts, making them prone to interpretation variations and fatigue. As a result, assistive medical technologies have been de-

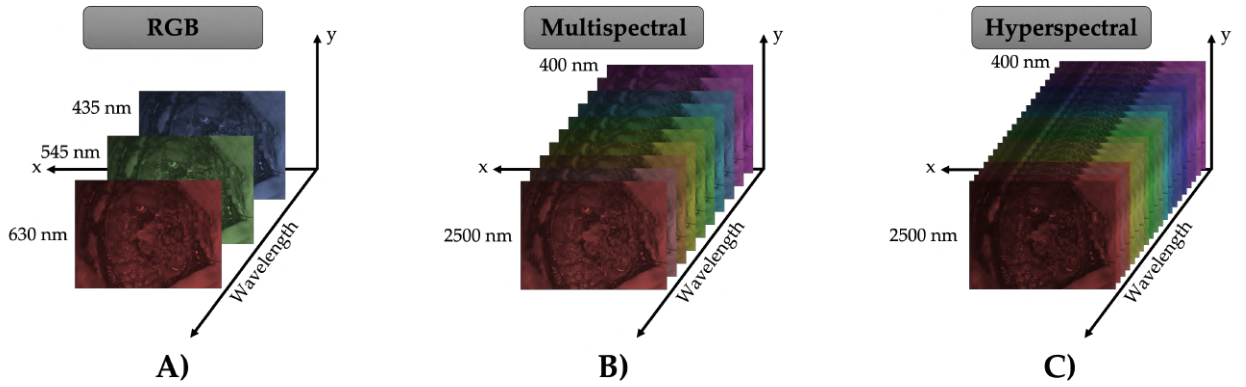


Figure 1.1: Exemplification of different image capture modalities: A) RGB, B) Multispectral, and C) Hyperspectral.

veloped that are based on medical image processing and classification techniques [15, 16]. However, despite the assistance provided by these technologies, it is still not possible to accurately distinguish among different types of tissue and composites. Consequently, there has been growing interest in multispectral (Figure 1.1 B) and hyperspectral (HS) images (Figure 1.1 C) in recent years. These technologies can identify various components using the optical response of a sample, enabling to distinguish of different types of tissues and composites with greater accuracy [17].

Hyperspectral imaging (HSI) is a relatively new type of imaging modality that originated in the field of remote sensing and has found many applications in various industrial and research areas in recent years [18]. Some of the areas that have benefitted from HSI are food quality inspection [19], pharmaceutical product quality control [20], marine ecosystem monitoring [21], soil pollution monitoring [22], petrochemical industry [23], and defense and security [24]. HS images are composed of spatial and spectral information, forming a three-dimensional matrix known as a hypercube or HS cube [25, 26]. Each spatial pixel corresponds to a vector of intensity values that describe the response at hundreds of different spectral wavelengths or bands, forming a continuous spectrum called a spectral signature. Unlike standard digital color cameras that only capture red, green, and blue (RGB) images with three wavelengths (see Figure 1.1), HS cameras can cover wide spectral ranges using different sensor types such as visual and near infrared (VNIR) between 400 and 1000 nm, near infrared (NIR) from 900 to 1700 nm, or short-wave infrared (SWIR) from 900 to 2500 nm (Figure 1.2) [27]. The spectral signature allows the differentiation, at pixel level, of the materials presented in the captured scene based on their chemical composition and optical properties [28].

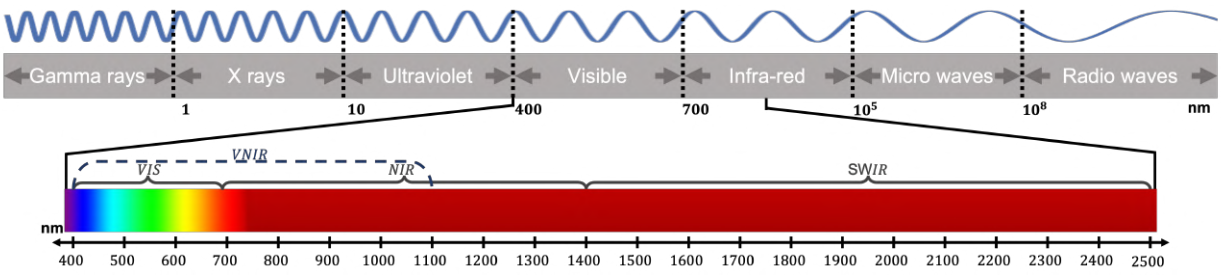


Figure 1.2: Electromagnetic spectrum with the most commonly used divisions.

HSI techniques have made significant advances in various fields of study, including the medical field, because of their noninvasive, nonionizing, nonlabeled, and contactless nature. In the medical field, HSI has been utilized to measure changes in deoxyhemoglobin and oxyhemoglobin using isosbestic points of the hemoglobin absorption spectral at the oxygen-sensitive wavelength [29]. Additionally, HSI systems have been used to identify skin lesions by analyzing the spectral properties of the skin affected by chromophores [30]. In another application, HSI was used for organ quality assessment during perfusion, where it predicted the tissue water index in the kidneys, allowing the analysis of tissue-related damage during ex vivo preservation [31]. Furthermore, Alzheimer's disease has been investigated by analyzing amyloid-beta protein in the retina in combination with HSI systems without contrast agents [32]. All of these applications are made possible because the optical characteristics of tissues, such as reflection, diffraction, and scattering, change as pathology progresses [26].

1.2 HSI in Cancer Detection

The applications of HSI in the healthcare environment have evolved significantly, to the point that in recent years this technology has been investigated and employed to identify various types of tumor tissues [33]. Consequently, this section provides a brief overview of how HSI is used to detect and evaluate various types of cancer.

1.2.1 Skin Cancer Applications

The incidence of skin cancer has been increasing around the world, with malignant melanoma and nonmelanoma skin cancer (NMSC) being the most common types. NMSC includes basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and other less prevalent types, while melanoma ranks 17th among the most frequent cancers worldwide with 325,000 new cases in 2020, and NMSC ranks 5th with 1,200,000 new cases in 2020 [34]. BCC is the most frequent type of skin cancer, accounting for 80-85% of all cases, followed by SCC with 15-20% [35].

Traditionally, skin cancer is detected through visual inspection by the naked eye and dermoscopy, which evaluates the shape and color of the lesion using the ABCDE (Asymmetry of the mole, Border irregular-

ity, Color uniformity, Diameter, and Evolving size) protocol [36]. In recent years, various imaging techniques have been employed in clinical practice, including confocal microscopy, polarized imaging, three-dimensional topography, thermal imaging, multispectral imaging (MSI), and HSI [30]. These latter imaging modes can capture information beyond the visible range related to different chromophores [37]. MSI technology has been widely used to identify skin cancer, with commercial systems such as MelaFind developed specifically for melanoma detection [38]. This device uses a snapshot imaging system to capture skin lesions. In 2015, Neittaanmäki-Perttu et al. presented a prototype HSI system to define the contour of lentigo maligna and its progression to melanoma [39]. The HS images were analyzed using a linear mixture model to identify the spectral responses of malignant lesions and normal skin, producing abundance maps that delineated the lesion borders with a match of 94.7%. In 2019, the same system was used in a pilot study to delineate BCC, achieving an accuracy of 75% [40]. In 2021, the system was tested to distinguish between BCC and melanoma, and a convolutional neural network (CNN) classifier was employed to identify 26 pigmented lesions, achieving a sensitivity of 100% and a specificity of 90% [40].

Zherdeva et al. proposed an HSI system that uses the optical density of hemoglobin and melanin to discriminate among different types of skin cancer [41]. The authors evaluated 45 skin lesions and achieved a sensitivity and specificity of 84% and 87%, respectively. Hosking et al. employed a Melanoma Advanced Imaging Dermatoscope (mAID) that covers 21 distinct wavelengths in the spectral range of 350-950 nm to classify nevus and melanoma lesions [42]. The study collected 70 HS images of skin lesions, performing a classification between nevus and melanoma with sensitivity and specificity of 100% and 36%, respectively.

Fabelo et al. developed a system that captured HS images with 125 spectral bands in the 450-950 nm range to evaluate 49 HS images of skin lesions corresponding to 36 patients [43]. Later, Leon et al. proposed a methodology that combined unsupervised and supervised techniques for the automated classification of pigmented skin lesions [44]. In this study, excellent differentiation of benign and malignant skin lesions was achieved with 87.5% sensitivity and 100% specificity.

In 2021, Courtenay et al. reported the use of a pushbroom HS camera to discriminate between healthy and non-healthy skin [45]. The acquisition system was composed of an HS camera with two 60 W halogen light lamps on each side and captured 270 spectral bands in the 398-995 nm range. Statistical tests were performed on a total of 60 patients, including 41 BCC and 19 SCC diagnosed cases. In a later study, a combination of a CNN and a final activation layer of a support vector machine was proposed to classify the same dataset, achieving up to 90% overall accuracy [46]. Furthermore, the SICSURFIS system has recently been proposed as a hand-held HSI tool for complex skin surfaces [46]. In this work, a total of 42 skin lesions were evaluated using only 33 of the thousands of spectral bands captured by the system in a spectral range of 475 to 975 nm. The classification and delineation method employed a CNN that utilized spectral, spatial, and skin-surface models, resulting in an 87% sensitivity and 93% specificity for the different tissues evaluated.

1.2.2 Brain Cancer Applications

A considerable cause of morbidity and mortality worldwide is represented by primary brain and central nervous system cancers. The treatment consists of biopsy or aggressive surgical resection with postoperative radiation and chemotherapy [47]. However, to achieve successful resection that increases the probability of survival, it is necessary to precisely delineate the boundaries between tumor and normal tissue [48]. During surgery, different tools are often used to act as surgical guidance, such as intraoperative image-guided stereotactic neuronavigation, intraoperative MRI, or fluorescent tumor markers like 5-aminolevulinic acid [49].

An emerging intraoperative guidance tool is HSI, which has been used to monitor brain oxygenation and hemodynamics in animals [50]. Recently, Fabelo et al. designed an intraoperative HS system to identify human cancer tissue during in vivo brain surgery [51]. The system employed two pushbroom HS cameras, an illumination source, and a scanning platform. The VNIR spectral range between 400 and 1000 nm was covered by one HS camera, capable of acquiring 1004 pixels with 826 spectral bands. Another HS camera imaged the NIR range between 900 and 1700 nm, capturing 172 spectral bands and 320 spatial pixels. The illumination system was based on a quartz tungsten halogen (QTH) lamp of 150 W with a bandwidth emission in the range 400 to 2200 nm. A scanning platform was used to provide the necessary spatial scanning to build the HS cubes. An HS human brain database was obtained from 22 patients employing this system [52].

In fact, the database obtained from 22 patients by Fabelo et al. [52] has been utilized in several works for the classification and delimitation of brain cancer. A hybrid approach that combined supervised and unsupervised machine learning (ML) methods was developed to perform a spatio-spectral classification. In the supervised stage, the SVM algorithm was used for pixel-wise classification, and the resulting map was spatially smoothed by K-Nearest Neighbors (KNN) filtering on a representative band of the HS cube. The output from this stage was combined with the unsupervised stage, which utilized a hierarchical K-Means strategy to obtain a segmentation map. The results showed that the approach accurately distinguished between normal tissue, tumor tissue, blood vessels, and background, with an overall accuracy close to 100%. Separately from traditional ML methods, deep learning (DL) approaches were also proposed to identify glioblastoma (GB) tumors. The proposed framework employed a 2D fully CNN to identify the parenchymal region, which corresponds to the principal surgical zone of the brain, and a 2D-CNN to identify blood vessels. The HS cube was then classified by a 1-dimensional deep neural network (1D-DNN) generating a classification map with the four tissue classes. The blood vessels and parenchyma maps were combined into the 1D-DNN classification map. This framework was able to identify GB tumors with a general mean accuracy of 80%. Another research used a blind linear unmixing (BLU) method to identify GBs as a low computational time-cost alternative [5]. This method was compared with a supervised SVM strategy, which required a high training time, achieving similar classification results, but with a speed-up factor of $429\times$.

Recently, a method combining multiple ML models was proposed to use spectral and spatial information to identify GBs [53]. The main strength of the framework is the joint implementation of 1D-DNN and 2D-CNN architectures, which yield spectral and spectral-spatial HSI feature extraction and classification, respectively. These structures, in conjunction with fusion and optimization, based on edge-preserving filtering and background class estimation by a fully CNN, enabled the proposed method to achieve an overall accuracy of 96.69% for four-classes classification and 96.34% for GB identification.

An HS acquisition system was presented by Urbanos et al. to acquire and process HS images during the surgical environment [54]. The system employs a snapshot HS camera that is capable of capturing 25 bands in the range of 655 to 975 nm. The illumination system uses a 150 W halogen lamp connected to two fiber optic cables to remove thermal exposure. In this study, an HS database was generated, consisting of over 50 images of various pathologies and labeled into five different categories: normal tissue, tumor, dura mater, venous, and arterial blood vessels. Finally, 13 images with advanced stages of GB (grades III and IV) were used to train and assess the SVM, Random Forest (RF), and CNN classifiers, resulting in an overall accuracy between 60% and 95%.

In addition, HSI systems have been used to monitor the differences in oxygenated and deoxygenated hemoglobin levels that occur in the brain during neurosurgical procedures [55]. A camera capable of capturing 25 bands in the spectrum between 675 and 975 nm was used for this application. Another HSI system was used to identify postoperative cerebral hyperperfusion syndrome by monitoring intraoperative changes in brain surface hemodynamics using the spectral range between 400 and 800 nm [56].

1.2.3 Gastrointestinal Cancer Applications

Gastrointestinal cancer includes the stomach, liver, esophagus, pancreas, and colorectum; this cancer modality represents 26% of the global incidence and 35% of mortality in 2018 [57]. Endoscopic tools such as gastroscopy, colonoscopy, and wireless capsule endoscopy using RGB cameras are commonly employed to detect gastrointestinal cancers and their abnormalities [58]. To increase versatility, HSI has been incorporated into endoscopy to explore the optical properties of the tissue. In 2018, a system called ICL SLHSI (Structured Light and Hyperspectral Imager) was developed by Lin et al., which employed a pushbroom HS camera that covers the spectral range 400-1000 nm with 270 spectral bands [59]. Similarly, a study proposed a technique to distinguish early esophageal cancer lesions using standard and HSI endoscopy with a spectral range of 350-800 nm [60]. Yoon et al. reported a line-scanning HSI endoscopy that was able to capture 100 spectral bands that spanned wavelengths of 680 to 730 nm. The HSI system was used to improve polyp discrimination for detection and resection in seven patients undergoing routine colonoscopy screening, and the KNN algorithm was used as a classifier to discriminate between patients with and without polyps [61]. Furthermore, Köhler et al. developed an HSI laparoscope capable of capturing 100 spectral bands ranging in wavelengths from 500 to 1000 nm, and the platform was tested with resected human tissue [62].

A pushbroom NIR HS camera with 256 bands and a spectral range of $1000 - 2350$ nm was employed by Sato et al. to image 12 ex vivo gastrointestinal stromal tumors [63]. During this the study, the SVM algorithm was utilized to predict normal and lesion regions, resulting in a performance of more than 73% in all evaluated metrics. In other studies, the TIVITA Tissue System (Diaspective Vision GmbH, Am Salzhaff, Germany), a commercial HSI system capable of capturing images with a spatial dimension of 640×480 pixels and a spectral range of 500 to 1000 nm, was evaluated to determine the resection margin during colorectal surgery in 24 patients [64]. Additionally, the same tool was used in another study to detect colorectal carcinoma in 54 patients, achieving a sensitivity of 86% and a specificity of 95% through the use of a neural network (NN) to classify tumor and healthy mucosa in colorectal carcinoma [65].

1.2.4 Head and Neck Cancer Applications

The oral cavity, nasopharynx, pharynx, and larynx are all included in head and neck cancer, which is often diagnosed through patient self-identification by presenting symptoms like voice changes or cranial nerve palsies [66]. HSI applications for this type of cancer are still in their early stages of development. An optical biopsy method using HSI and CNN was proposed by Halicek et al. for ex-vivo head and neck cancer [67]. The acquisition structure utilized a CRI Maestro imaging system (Perkin Elmer Inc., Waltham, Massachusetts) comprising a xenon white-light illumination lamp, a liquid crystal tunable filter, and a 16-bit charge-coupled devices (CCD) camera capable of capturing 91 bands with wavelengths ranging from 450 to 900 nm. Recently, Eggert et al. conducted a prospective clinical observational study to classify healthy and tumor tissue of laryngeal, hypopharyngeal, and oropharyngeal mucosa [68]. The HSI system used was capable of capturing 30 spectral bands ranging from 390 to 680 nm. In this work, 98 patients were examined in vivo due to suspicious lesions of the mucosal membrane before surgery. DL methods were employed to achieve an average accuracy of 81%, a sensitivity of 83%, and a specificity of 79%.

1.2.5 Histological Samples in Cancer Applications

Several diseases can be identified by examining histological samples using digital pathology. Microscopy is used to digitize samples and capture partial or complete images at higher magnifications [69]. HSI has been employed in various studies to analyze histological samples using microscopy [25].

A process to attain high-quality HS images using a pushbroom HS microscope was proposed by Ortega et al. [70]. The HS camera captured 826 spectral bands and 1004 spatial pixels in the spectral range of 400 to 1000 nm. The camera was coupled to a conventional light microscope to capture 83 HS images for $5\times$ and $10\times$ magnifications. These images were obtained from 13 pathology slides of human brain tissue affected by a grade IV GB tumor resected during surgery [71]. Three supervised classification algorithms, SVM, NNs, and RF, were utilized to classify the HS images captured at $5\times$ magnification. The system achieved competitive results in distinguishing between normal and tumor tissue with an accuracy of more than 80%

[72]. In a recent study, the pushbroom HS microscope system was modified to overcome the limitation of the effective spectral range. A new database of 527 HS images was collected, including 337 non-tumor brain samples and 190 GB samples [25]. The system was also employed to differentiate between normal and tumor breast cancer cells using 112 HS images captured from histologic samples of human patients at $20\times$ magnification. A DL network was employed, achieving an area under the curve of over 0.89 for all experiments.

An HS microscopic imaging system was developed by Ma et al., covering a spectral range of 460 to 750 nm with 87 spectral bands, employing a SnapScan HS camera [73]. A total of 15 histologic slides of larynx and hypopharynx from 15 head and neck cancer patients were collected at $40\times$ magnification. The author proposed a nuclei segmentation strategy based on the analysis of the main components. Spectral-based SVM and patch-based CNN were used for nuclei classification. The average accuracy of the spectral-based SVM classification was 68%, and of CNN was 82%. Finally, a system was presented by Souza et al. for acquiring HS images using a conventional microscope and a liquid crystal tunable filter (LCTF) [74]. The light in various wavelengths from 400 to 720 nm is filtered using this system based on light polarization. The system was tested on a $H\&E$-stained slide of rat skin treated with ALA-mediated photodynamic therapy. Four different algorithms (KNN, SVM, and RF) were employed, resulting in an accuracy between 96% and 98%.

1.3 Problem Statement

The potential and versatility of HSI in the medical field have been demonstrated in recent years, as shown in previous sections. Despite these advances, there are some limitations to this technology that must be taken into account. One of them is that direct identification of tumor or injured tissue regions is not directly possible in the raw HS images. Because of this, advanced processing is necessary to identify the spectral signatures of the different elements that make up a sample, and at the same time classify the regions that have them. These processes, as can be seen in the previous cases, are diverse with different advantages and disadvantages of each particular algorithm and application. In addition, this heterogeneous condition involves high difficulty in analyzing the data and extracting meaningful information with a high computational cost for its study. Finally, capture conditions can affect data quality and certain materials can be complex to identify, such as those with low reflectance or transparent materials.

In the literature, there are standard algorithms that allow to perform HS analysis in multiple applications in different research fields [75]. Among these algorithms, decomposition or unmixing methods are versatile techniques widely used in various research fields to analyze HS data. These algorithms allow estimating the characteristic spectral signatures of the different components present in a sample, also called profiles or end-members, and their fractional contributions known as abundance or concentration at each pixel [76, 77]. In the state-of-the-art, different approaches have been proposed to perform spectral decomposition, methods

such as those based on artificial intelligence (SVM, NNs, among others), spectral divergence information, spectral angle mapping, and mathematical linear spectral unmixing (SU) techniques [77]; these methods have emerged as several important processing options [78]. Nevertheless, identifying characteristic spectral signatures of a particular tissue, such as the regions with the highest concentration of them, without prior knowledge is not a simple task, even with training information. In this sense, linear SU methods let blind or unsupervised estimation of both end-members and abundance maps without prior training information [79].

Although linear unmixing (LU) techniques have been shown to be efficient in classifying various types of tissues, they are susceptible to noise from different sources and light scattering from adjacent pixels [80]. Still, these techniques have positioned themselves as a powerful tool for converting spectral information into data that may be related to the physical abundances of materials on the surface [80, 81]. But this condition is strictly valid only for the situation where the end-members are arranged at discrete locations on the surface, condition rarely met in practice. Moreover, in many research fields these material mixtures are closely related to each other, which is very difficult to consider for simple linear models [81, 82]. However, in the case of nonlinear spectral decomposition, it is possible to distinguish between microscopically mixed materials and even to take into account events that present multiple scattering effects, in addition to showing greater robustness to the presence of noise [80, 81]. Furthermore, taking into account spatial coherence properties, corresponding to the spectral signatures surrounding the analyzed area allows a much more accurate and robust estimation of abrupt changes [83].

We believe that this research will provide the foundation for developing various SU algorithms based on linear and nonlinear mixture models for the classification and identification of basic components in tissue samples affected by different pathologies. These methods would enable rapid and accurate analysis of the data using unsupervised, semisupervised, and supervised approaches. Additionally, we aim to explore the possibility of combining these methods with other decomposition or classification techniques, thereby enhancing the accuracy of the estimations.

1.4 Research Objectives

The overall objective of this dissertation project is to propose advanced analysis methods that allow the identification of different components present in HS biomedical images. This work is based on the unmixing mathematical framework of the Extended Blind End-member and Abundance Estimation (EBEAE) algorithm [84]. These methods involved the addition of spatial coherence information and the extension of the linear model to a nonlinear one. Additionally, hybrid methodologies that combined unmixing algorithms with ML techniques were explored and proposed. Furthermore, possible sources of variability in the HS data that could have affected classification processes were analyzed.

The following specific aims were established for the period September 2019 - August 2023 to meet the general objectives:

(1) Study, analyze and understand the mathematical framework of the EBEAE algorithm.

(2) Evaluate the EBEAE method on different types of HS images.

(3) Develop a blind estimation methodology that includes spatial coherence information in its formulation.

(4) Evaluate the unmixing algorithm with spatial coherence with different HS databases.

(5) Develop, study, and understand the nonlinear extended blind abundance and end-member estimation (NEBEAE) method.

(6) Evaluate the NEBEAE algorithm on experimental HS images.

(7) Explore ML algorithms for joint classification methodologies with spectral decomposition methods.

(8) Reduce variability in the classification and identification of components by a new calibration method.

The specific aim 1 will be achieved by reviewing in the literature the EBEAE algorithm, which was proposed in [84], and its previous versions. In this stage, the mathematical analysis will be carried out to deduce each component present in the synthesis framework, such as the hyperparameters. Subsequently, aim 2 will consist of studying the behavior of the decomposition algorithm in different types of HS images, both synthetic and experimental. This stage will provide a better understanding of the hyperparameters and the implications of modifying each of them. In addition, the experiments will allow interpretation of the results and identification of possible limitations in the evaluations.

The work done in specific aims 1 and 2 will serve as the basis for the specific aim 3, in which various types of spatial coherence techniques in the literature will be analyzed. In addition, decomposition algorithms that implement spatial coherence information in the estimation will be studied. Once this aim is achieved, a spatial coherence version of EBEAE will be proposed. The specific aim 4 of this study is to evaluate the performance of the proposed algorithm on both experimental and synthetic HS images and to compare it with the methods discussed in the literature.

Subsequently, with the experience obtained in the previous objectives and examining in detail the state-of-the-art, a version of nonlinear unmixing will be proposed. This algorithm will be evaluated in different databases, as in the specific aim 2, which will fulfill the specific aims 5 and 6. After that, in specific aim 7 it will seek to improve the results generated by means of hybrid classification algorithms, which combine spectral decomposition and ML techniques. Finally, in specific aim 8 will seek to improve the estimations made by a new preprocessing proposal and a new DL approach using attention methods.

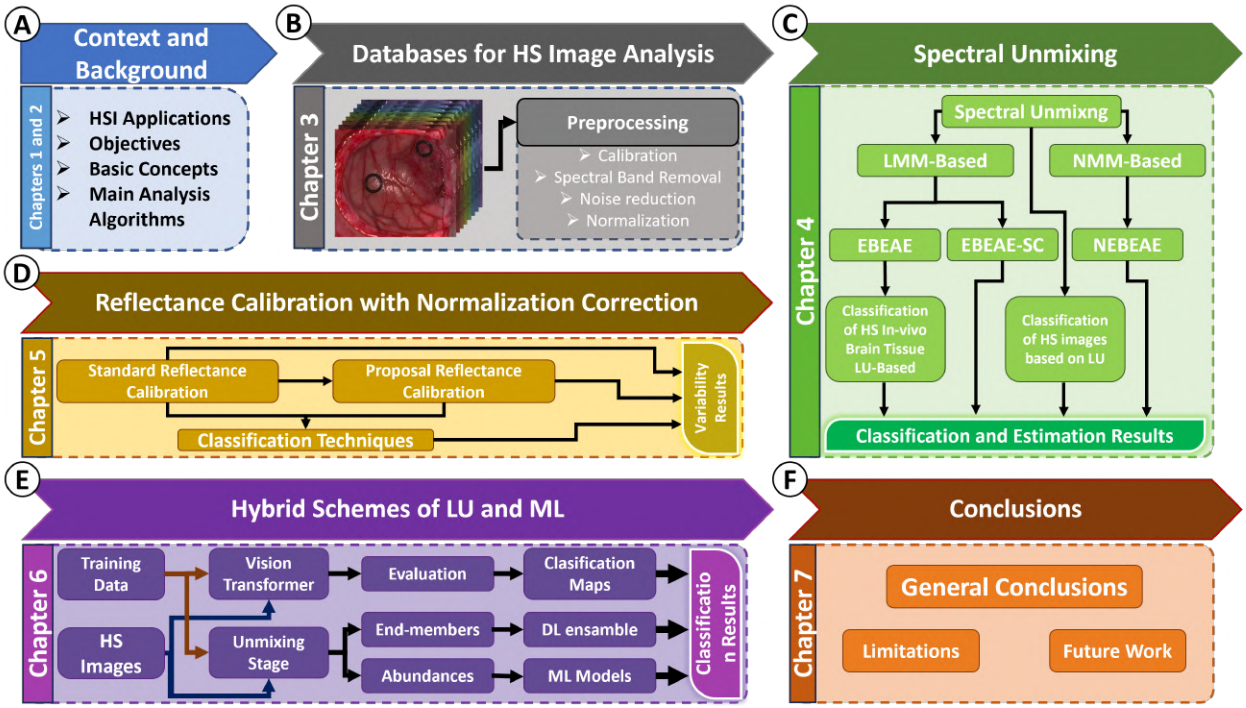


Figure 1.3: General thesis organization diagram.

1.5 Thesis Organization and Scientific Contributions

This section describes the organization of this dissertation document, which has been divided into six chapters covering the general objectives described previously. The structure and general content of this dissertation are depicted in Figure 1.3, providing an overview of its distribution. This diagram also illustrates the journey undertaken during the doctoral period. Initially, the focus was solely on unmixing algorithms, but the challenges encountered during the project's development led to the creation of new strategies, presented in subsequent chapters. Additionally, it highlights the main contributions resulting from achieving the proposed specific objectives and the challenges overcome. It also discusses the main contributions made to achieve the set goals.

Initially, Chapter 1 provides an overview of the main argument, the specific aims, and structure of this document. The general view of HSI in medical applications emerges as part of the following book chapter:

- I. A. Cruz-Guerrero, R. Leon, A. R. Mejia-Rodriguez, D. U. Campos-Delgado, S. Ortega, H. Fabelo, and G. M. Callico, "Hyperspectral Imaging for Cancer Applications," in Diagnosis and Treatment of Cancer using Thermal Therapies, CRC Press, 2023, pp. 81–101, doi: 10.1201/9781003342663.

Subsequently, each chapter aims to address the proposed objectives and presents the contributions developed based on them.

Chapter 2: Methodological Background

Contains detailed information about different concepts, topics, and definitions used throughout this document. In addition, this chapter includes subjects not so extensively detailed in the first one.

Chapter 3: Databases for Hyperspectral Image Analysis: Description and Preprocessing

In this chapter, the HS databases employed in this research work are presented, along with their properties, acquisition platform, and preprocessing.

Chapter 4: Spectral Unmixing

This chapter presents the theoretical background of the mixing models found in the literature, which form the basis for the different SU methods developed during this dissertation. It covers both linear and nonlinear algorithms. The objective of this chapter is to provide an overview of the proposed SU methods, explore various approaches to achieve classification based on these methods, and develop classification proposals based on LU. This chapter materialized in four scientific contributions, whose references are provided below:

- I. A. Cruz-Guerrero, R. Leon, D. U. Campos-Delgado, S. Ortega, H. Fabelo, and G. M. Callico, "Classification of hyperspectral in vivo brain tissue based on linear unmixing," Appl. Sci., vol. 10, no. 16, Aug. 2020, doi: 10.3390/app10165686.
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejia-Rodriguez, A. J. Jo, S. Ortega, H. Fabelo, and G. M. Callico, "Multi and Hyperspectral Image Unmixing with Spatial Coherence by Extended Blind End-member and Abundance Extraction," J. Franklin Inst., 2023.
- D. U. Campos-Delgado, I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, A. R. Mejía-Rodríguez, S. Ortega, H. Fabelo, and G. M. Callico, "Nonlinear extended blind end-member and abundance extraction for hyperspectral images," Signal Processing, vol. 201, p. 108718, 2022, https://doi.org/10.1016/j.sigpro.2022.108718.

Chapter 5: Reflectance Calibration with Normalization Correction In this chapter, a proposal for spectral calibration is introduced with the aim of reducing spectral variability, followed by its respective evaluation in both a controlled scenario and an experimental case. This proposal was presented at the Euromicro Conference on Digital System Design, while the experimental evaluation was presented at the National Congress of Biomedical Engineering, whose references are provided below:

- I. A. Cruz-Guerrero, R. Leon, L. Granados-Castro, H. Fabelo, S. Ortega, D. U. Campos-Delgado, and G. M. Callico, "Reflectance Calibration with Normalization Correction in Hyperspectral Imaging," in 2022 25th Euromicro Conference on Digital System Design (DSD), 2022, pp. 855-862, https://doi.org/10.1109/DSD57027.2022.00120.

- I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, and D. U. Campos-Delgado, "Glioblastoma Classification in Hyperspectral Images by Reflectance Calibration with Normalization Correction and Nonlinear Unmixing," in Congreso Nacional de Ingeniería Biomédica, Cham: Springer International Publishing 2022, pp. 393–402, https://doi.org/10.1007/978-3-031-18256-3_43.

Chapter 6: Advanced Hyperspectral Brain Tissue Classification with Hybrid Schemes of Linear Unmixing and Artificial Intelligence

This chapter addresses the methodologies developed regarding hybrid approaches, which combine SU techniques with ML algorithms for classification and identification of components. Additionally, a proof of concept based on vision transformers for spectral information classification is evaluated. The results and proposals in these topics are reflected in three scientific contributions, whose references are listed below:

- I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejia-Rodriguez, H. Fabelo, S. Ortega, and G. M. Callico, "A hybrid approach to the hyperspectral classification of in vivo brain tissue: linear unmixing with spatial coherence and machine learning," in Artificial Intelligence in Cancer Diagnosis and Prognosis, Volume 3: Brain and prostate cancer, IOP Publishing, 2022, https://doi.org/10.1088/978-0-7503-3603-1ch9.
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejia-Rodriguez, R. Leon, S. Ortega, H. Fabelo, R. Camacho, M. Plata, and G. M. Callico, "Hybrid Brain Tumor Classification Scheme of Histopathology Hyperspectral Images Using Linear Unmixing and Deep Learning," Submitted in Healthcare Technology Letters, 2023, Manuscript ID: HTL-2023-08-0027.
- I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, D. U. Campos-Delgado, H. Fabelo, S. Ortega, and G. M. Callico, "Classification of Brain Tissues in Hyperspectral Images Using Vision Transformers," in 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), 2023, pp. 1–4, doi: 10.1109/ISBI53787.2023.10230806.

Chapter 6: Conclusions

This last chapter summarizes the main findings of this dissertation, as well as the limitations and areas of future development and application of the methods proposed in this research work.

Chapter 2

Methodological Background

The field of HSI has undergone remarkable expansion and advancement in recent years, driven by the emergence of novel technologies and cutting-edge techniques designed to tackle a diverse array of challenges. To truly appreciate the significance of the research unveiled in this dissertation, it becomes imperative to cultivate a comprehensive grasp of the contemporary landscape within the HSI domain. In essence, the overarching objective of this chapter is to lay down a sturdy bedrock of knowledge, one that will serve as a vantage point for the research expedition that unfolds in the ensuing chapters. By delving into the fundamental principles, methodologies, and innovations that underlie HSI, it aims to provide readers with the essential context required to grasp the implications of the research endeavors documented within this dissertation.

2.1 HSI Platform

As described in Chapter 1, HSI is a technology used to capture and analyze the full spectrum of light reflected or emitted by an object or scene [26]. A typical HS image is a 3-dimensional array, where the first two dimensions correspond to the spatial x and y coordinates of the image, and the third dimension corresponds to the wavelength of light [25, 26]. This means that each pixel of an HS image represents a collection of measurements taken at different wavelengths (such as in the visible, infrared, or ultraviolet regions of the electromagnetic spectrum); this information is the unique pattern of light that is reflected, emitted, or absorbed by a substance or object and can be used to identify it or determine its properties [25].

In general, HSI has several advantages over other types of imaging techniques, which include:

- High spectral resolution: HSI can capture light across a wide range of wavelengths and can be used to detect materials or features that are not visible to the naked eye, which allows the identification of hidden features and the characterization of a wide range of materials [33, 85]. This property of

HSI is useful in a variety of applications, such as mineral exploration, environmental monitoring, and medical imaging [85].

- Chemical analysis: HSI can be used to determine the chemical composition of a wide range of materials, such as rocks, soils, and biological tissue [86].
- Non-destructive testing: HSI is a non-destructive testing method, meaning that it does not damage the object being imaged, which is relevant in applications such as art conservation and archaeological research [33, 85].
- High spatial resolution: With the advancements of technology, HSI sensors can also provide high spatial resolution, allowing them to detect small objects on a big image and a high level of detail [85, 87].
- Combination with other sensors: HS data can be easily integrated with data from other sensors, such as lidar and radar, to provide a more complete understanding of the objects being imaged [33, 85].

Despite the many advantages HSI offers, there are some limitations such as:

- Data size: The large amount of data collected by an HSI system can be difficult to store and transmit [33, 85].
- Spatial resolution: The spatial resolution of an HS image is usually lower than that of a traditional image, especially when a high spectral resolution is required or a quick capture is desired [87].
- Signal-to-noise ratio: In some cases, the signal-to-noise ratio of an HS image may be too low to obtain useful data, particularly in cases where the scene being imaged is low in contrast or variability [87, 88].
- Interference: Some environmental factors can introduce noise and interfere with the image acquisition process, which can reduce the accuracy of the data [85, 86].
- Object occlusion: Certain items or substances may be hidden by others, making it difficult to obtain precise information about them. [87].
- Spectral analysis: Sometimes, the accurate spectral analysis might be limited by the complexity of the target and the quality of a known prior spectral library [33, 86].
- High computational requirements: HS data require a lot of computational power and time to extract information [85].

The selection of HSI instrumentation can have a significant impact on the limitations of the technique, since there are always trade-offs between advantages and drawbacks. In the following subsection, the instrumentation required within HSI and the main characteristics of each element will be studied.

2.1.1 HSI Instrumentation

In HSI, instrumentation is a crucial element for reliable, efficient, and high-quality spectral data acquisition. Usually, an HSI platform consists of an HS camera, a light source, a computer with the acquisition software, and, in some cases, a motorized mobile station, which depends on the scanning mode employed by the HS camera [86, 89]. This section provides an overview of the key components of an HSI system.

2.1.1.1 HS Camera

The HS camera is the main component of the acquisition system, which consists of two structures: spectrographs or spectrometers and a detector or array of photosensitive detectors [90]. Spectrographs allow the dispersal of incident polychromatic light into beams with specific wavelengths, with three types of devices [86]: monochromator, optical bandpass filter, and single-shot imager. These scattering devices focus the narrow wavelength light toward each of the detectors. In this sense, the photosensors most used in HSI are CCD and complementary metal-oxide semiconductors (CMOS) [88]. The principal difference between these two sensors lies in the transmission scheme of the incoming signals. On the one hand, CCD sensors focus on measuring the luminous intensity, transferring the resulting multisensor signal to a digital/analog converter. On the other hand, CMOS sensors incorporate the photodetector and the digital/analog converter together; thus, the information from each sensor is independent of the rest. Because of this difference, CMOS sensors are faster in measuring and capturing photons, but these sensors are susceptible to the presence of noise and are mostly affected by dark currents [88, 90]. This situation is compensated by CCD sensors, since digitizing the signals outside the photodiode allows for the inclusion of components with different characteristics that mitigate noise, dark current, and acquisition speed. In addition, CCDs have better sensitivity for visible and VNIR wavelengths, whereas CMOS have higher efficiency in the infrared range.

In general, HSI cameras are classified according to the scanning method used to generate hypercubes, with four main types of scanning: whiskbroom, pushbroom, focal plane, and snapshot [85, 86, 88]. Whiskbroom or point-scanning cameras are characterized by capturing the spectral information of one pixel at a time (Figure 2.1A); this means that to scan a particular region, it is necessary to have a mobile station that travels through the scanning area of the camera at each location in dimensions X and Y, where the acquired hypercube is stored in a band-interleaved-by-pixel (BIP) format [86, 88]. Due to this, whiskbroom cameras require considerable time to acquire an image, so spatial resolution is often limited. Nevertheless, the main strength of these cameras is their high spectral resolution, which permits one to capture a large amount of information. Like in the previous case, the pushbroom or line-scanning cameras (Figure 2.1B) acquire complete spectra of several pixels continuously, that is, the area of interest is scanned line by line until the entire image is captured, storing the hypercube in a band-interleaved-by-line (BIL) format [85, 88]. This scanning mode requires a mobile station to traverse the scan line through the area, however, motion artifacts may occur. The pushbroom camera provides high spatial and spectral resolution, and for this reason is the

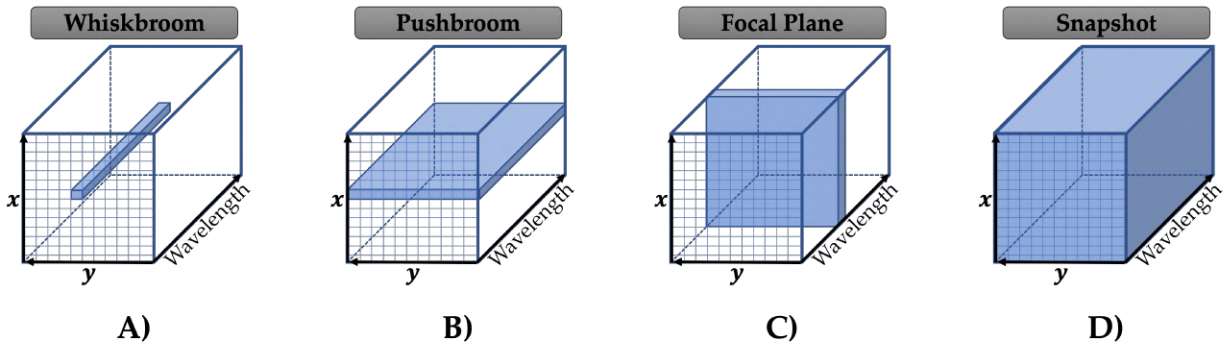


Figure 2.1: Hyperspectral imaging scanning techniques: A) whiskbroom, B) pushbroom, C) focal plane, and D) snapshot.

most used today.

Focal plane and snapshot cameras (Figures 2.1C and 2.1D) allow scanning of complete areas, capturing spatial and spectral information together [85, 88]. Focal plane cameras acquire a 2D monochromatic image at a given wavelength, i.e., each spectrum is captured independently until completion of the hypercube, using a band sequential format (BSQ) as a form of storage [86, 88]. The main advantage of these cameras is that they can capture a single wavelength or several wavelengths by selecting them, but they are susceptible to the presence of motion artifacts. As in the previous case, snapshot cameras acquire spatial and spectral information simultaneously, but unlike focal plane cameras, this type of camera produces the hypercube in a single shot, which results in a reduction of the capture time. However, snapshot cameras can acquire a limited number of spectral bands, so the spectral resolution is lower than with other camera types.

2.1.1.2 Light Source

In this sense, the light source is another crucial component of the HSI acquisition system, as light is the medium that provides information about the objects under study. Currently, halogen lamps are the most widely used because of their broad-spectrum, which is continuous, soft, and without sharp peaks [86, 90]. However, this type of illumination has certain disadvantages, such as a temperature rise in the sample, short lifetime, and spectral peak change as a result of variations in temperature, voltage, and time of use. On the other hand, light-emitting diodes (LEDs) have started to be used as light sources due to their long lifetime, fast response, compact size, low power consumption, and low heat generation. LEDs can produce broad and short spectra in the ultraviolet, visible, and infrared regions. However, they are not very efficient in dissipating heat, which reduces their lifetime and affects their spectrum. Finally, lasers are light sources with narrow bandwidth, linear direction, and are used mainly in fluorescence and photoluminescence applications [86, 90].

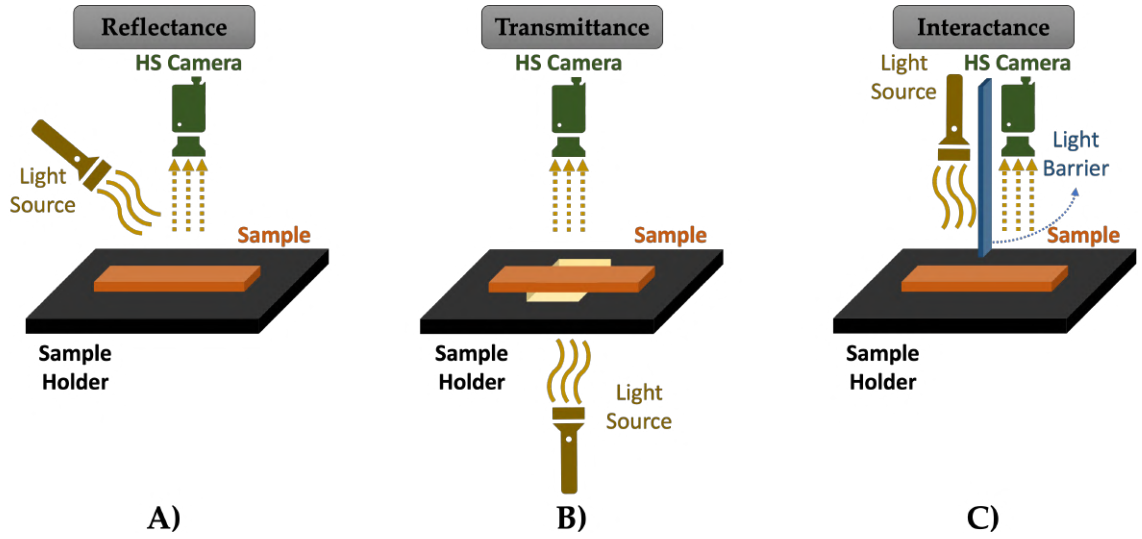


Figure 2.2: Most common hyperspectral acquisition modes: A) reflectance, B) transmittance, and C) interactance.

2.1.2 HSI Acquisition Methodologies

The most common acquisition modes or methodologies are reflectance, transmittance, and interactance, which configurations are depicted in Figure 2.2. These acquisition modes are fundamental in the interaction studies of electromagnetic radiation with objects and are often used to provide different kinds of information about a material [86, 89, 90].

Reflectance is the amount of light reflected by a surface across the electromagnetic spectrum [86, 89, 90]. It is measured in percent, with 100% reflectance equal to the total amount of light incident on the surface, while 0% indicates that the surface reflects no light. This information allows to create of spectral signatures, which are unique patterns of reflectance that can be used to identify different materials by measuring their properties, and tracking their changes over time. This acquisition mode has a preprocessing step linked to the reflectance information to eliminate the effects of temperature and illumination changes, as well as the aging of the light source [85, 86, 89, 90]. This preprocessing is known as reflectance calibration, so that the raw image ($I_0 \in \mathbb{R}^{X\times Y\times L}$) is modified based on a dark ($I_D \in \mathbb{R}^{X\times Y\times L}$) and a white ($I_W \in \mathbb{R}^{X\times Y\times L}$) reference images. The reference I_D is captured by closing the camera shutter, while I_W is captured from a highly reflective and uniform white surface. These images are related by the following equation

$$I_R = 100 \times \frac{I_0 - I_D}{I_W - I_D}, \tag{2.1}$$

where $I_R \in \mathbb{R}^{X\times Y\times L}$ is the resultant image, with spatial dimensions $X \times Y$ and L number of spectral bands [86, 90].

Transmittance is the measurement of the amount of light that passes through a material at different

wavelengths. It is measured in percent, with 100% transmittance equal to the total amount of light passing through the surface and 0% indicates that light cannot pass through the sample under analysis [85]. This information can measure the optical properties of translucent materials. As in the case of the reflectance mode, transmittance also has a preprocessing stage to adjust the spectral information and eliminate unwanted effects. This preprocessing is known as transmittance calibration, using the same images as in the case of reflectance calibration (I_0, I_D, and I_W) [86]. The transmittance is obtained through

$$I_T = -\log_{10}\left(\frac{I_0 - I_D}{I_W - I_D}\right), \tag{2.2}$$

where $I_T \in \mathbb{R}^{X \times Y \times L}$ is the transmittance image.

Finally, interactance is the measurement of the amount of light interacting with a surface, including both reflectance and transmittance. In this mode, the light sources and the optical detector are on the same side of the sample, but the capture field of the detector is not illuminated [86], as shown in Figure 2.2C. The sensor captures the light as it passes through the sample and reaches the surface, reducing the adverse effects caused by the thickness of the sample and those caused by specular reflection, thus obtaining both external and internal information. This information measures the optical properties of opaque or semitransparent materials [86].

2.2 HSI Analysis Algorithms

HS image processing has emerged as a significant research field to extract valuable information from captured objects and scenes. This processing involves the application of various algorithms, which play a crucial role in information extraction, material classification, change detection, object segmentation, and identification of chemical compounds present in a scene.

This section presents the most common algorithms for HS image processing. Additionally, the analysis methodologies to tackle specific tasks, such as supervised, unsupervised, and semi-supervised learning, will be discussed. Each of these learning approaches and processing methods comes with its own set of advantages and challenges, and the selection depends on the specific characteristics of the problem at hand and the available data.

2.2.1 Analysis Methodologies

In general, the models used to analyze HS images employ diverse strategies to perform tasks such as classification, segmentation, compound identification, and even regression. Among the different analysis approaches, supervised, unsupervised, and semi-supervised learning stand out [91, 92].

2.2.1.1 Supervised Approach

Supervised learning is a strategy in which a model is trained using pairs of input/output data, where the outputs can be discrete (classification) or in real numbers (regression) [91, 92]. The goal of the supervised methodology is for the model to learn to correctly map input data to the expected outputs. The principle of this approach is to provide the model with a significant amount of training data, where each input data is associated with a specific output value. During training, the models adjust their parameters to minimize the difference between the predicted outputs and the expected outputs. Finally, once the model has been trained, it can be used to make predictions on unpaired input data by employing the learned patterns to infer the corresponding outputs. In this way, the supervised method allows one to gain knowledge and identify patterns from previously labeled examples [91, 92]. It is also important to note that the performance of supervised learning is highly dependent on the quality and representativeness of the training data. The more varied and comprehensive the training data set is, the better the model will be in making precise and applicable predictions.

2.2.1.2 Unsupervised Approach

Unsupervised methodology is a type of ML in which the models learn from data without any information about the corresponding outputs. Unlike the supervised approach, the unsupervised method does not have explicit information to provide to the model, whose main goal is to discover patterns, clusters, hidden structures, or intrinsic relationships in the data without any external support or prior knowledge of the labels or classes to which they belong. However, interpreting the results can be more challenging and subjective since there is no clear reference to evaluate the quality of the predictions [91, 92]. Furthermore, the effectiveness of the unsupervised approach, as in the supervised case, depends on the quality of the data and the appropriate choice of algorithms and parameters.

Some examples of common unsupervised techniques include clustering and dimensionality reduction methods [92]. Clustering algorithms are employed to group data into sets or clusters based on their similarity, to allow data within the same cluster to be as similar as possible to each other, while data between clusters should be significantly different. On the other hand, dimensionality reduction is used to represent data in a lower-dimensional space while preserving characteristic information. This new representation allows better visualization and understanding of the data, as well as reducing computational complexity in subsequent analysis after reduction [91, 92].

2.2.1.3 Semi-supervised Approach

The semi-supervised method combines the capability of the unsupervised approach with the use of supervised information to improve accuracy [93]. In this strategy, labeled data (typically a small amount) is used alongside unlabeled data, where the latter usually represents a larger proportion for training [91, 92].

The main idea behind the semi-supervised approach is to leverage the limited information from the labeled data to enhance the performance of the model. However, unlike supervised learning, the semi-supervised method uses labeled in conjunction with unlabeled data to learn patterns and underlying structures, establishing the assignment of classes to the unlabeled information [91, 92, 93]. In other words, the objective is to use the labeled data to guide the model in the classification or prediction of unlabeled data.

The semi-supervised strategy is a highly useful tool in scenarios where obtaining labeled data is challenging or costly. However, similar to previous approaches, the performance of this method is based on the quality and representativeness of the data. Additionally, finding an appropriate balance between labeled and unlabeled data is important to achieve good performance, as an imbalance can introduce bias in model generalization [91, 92, 93].

2.2.2 Spectral Information Divergence

Spectral Information Divergence (SID), together with Spectral Angle Mapper, is one of the most common methods in HS image analysis. This method assesses the difference between the spectra of two pixels in an image [26, 33]. In simple terms, the more prominent the divergence, the larger the difference between the spectra and vice versa. The SID is calculated based on the probability distribution of two spectral vectors: one corresponding to the pixel under analysis $\{t_i\}$ and the second corresponding to a reference pattern $\{r_i\}$. The probability distribution of the reference spectra $\{q_i\}$ is calculated as:

$$q_i = \frac{r_i}{\sum_{i=1}^{c} r_i} \quad \forall i \in \{1, ..., c\}, \tag{2.3}$$

while the distribution of the test spectra $\{p_i\}$ is calculated using the following equation:

$$p_i = \frac{t_i}{\sum_{i=1}^{c} t_i} \quad \forall i \in \{1, ..., c\}. \tag{2.4}$$

Once the probability distributions $\{p_i\}$ and $\{q_i\}$ are calculated, it is possible to estimate the value of the SID using the equation:

$$SID = \sum_{i=1}^{c} p_i \log \frac{p_i}{q_i} + \sum_{i=1}^{c} q_i \log \frac{q_i}{p_i}, \tag{2.5}$$

where c represents the number of spectral channels or bands [94, 95].

2.2.3 Spectral Angle Mapper

The Spectral Angle Mapper (SAM) is a widely used algorithm in the field of HS image analysis. This method measures the spectral similarity between pixels in an HS image and a reference spectra by calculating the angle between both spectra [26, 33]. Typically, this evaluation is performed by comparing each pixel in the image with a known reference spectra, based on the assumption that similar materials have similar

spectral signatures [33, 96]. SAM determines the similarity using the following equation:

$$\alpha = \cos^{-1} \frac{\sum_{i=1}^{c} t_i r_i}{\sqrt{\sum_{i=1}^{c} t_i^2} \sqrt{\sum_{i=1}^{C} r_i^2}}, \tag{2.6}$$

where α is the angle between the spectral response $\{t_i\}$ (the spectrum under analysis) and $\{r_i\}$ (the reference spectrum). In this case, if the result of this equation is small, the pixel is considered to belong to the same class as the reference spectrum [96]. On the contrary, if the angle is large, the pixel is classified as belonging to a different class.

2.2.4 Spectral Unmixing

Spectral unmixing (SU) methods are pixel analysis algorithms that decompose the spectral information of a pixel under the assumption that it is a mixture of elemental spectra or end-members, along with a set of fractional abundances indicating the proportion of each end-member. There are two main approaches to evaluate unmixing algorithms: supervised and unsupervised [26, 84]. In supervised unmixing, prior knowledge about the basic materials composing the mixture is available, such as their reflectance patterns or spectral signatures, or their abundance maps. On the other hand, unsupervised or blind unmixing aims to identify end-members and their abundances solely based on measured data, without any prior information [26, 84]. This topic is explored in detail in Chapter 4, where the theoretical foundations and proposed approaches are discussed.

2.2.5 Machine Learning

Machine learning (ML) is a set of techniques and algorithms within the field of artificial intelligence that allow learning and decision making directly from data, rather than learning through mathematical inference or explicit programming [91, 92]. ML algorithms are based on identifying patterns and relationships in training data, to carry out tasks such as classification, regression, clustering, and anomaly detection [97, 98]. As a result, these algorithms can improve their performance as they are provided with more data and feedback from users and are capable of identifying these patterns in unlabeled data and even in unforeseen situations.

In the literature, ML algorithms have been widely used for HS image classification with very good results [99, 100]. Some of these algorithms include:

- Random Forest (RF),
- Support Vector Machines (SVM),
- Fully Connected Neural Network (NN),
- Convolutional Neural Network (CNN),

- Autoencoders (AE),
- Deep Belief Network (DBN),
- Recurrent Neural Network (RNN),
- Residual Network (ResNet),
- Vision Transformers (ViT).

Each of these algorithms employs a different theoretical approach and a distinct functional approach to produce information classification, hence they have advantages and disadvantages specific to each strategy [97, 98, 99, 100]. Because of this, the results vary depending on the application. The methods listed above are briefly discussed in the Appendix B, mentioning some applications of these strategies in the field of HSI.

Chapter 3

Databases for Hyperspectral Image Analysis: Description and Preprocessing

HS images, characterized by their numerous spectral bands, often reaching into the hundreds or even thousands per pixel, present a rich source of information for a wide range of applications. Yet, their high-dimensional nature brings forth a unique set of challenges when it comes to their analysis. Among these challenges, one stands out prominently: the need for extensive and diverse databases to support the development, training, and evaluation of the various analysis methods commonly employed in HS image research.

These datasets are the lifeblood of research in HSI. They serve as the foundation upon which innovative techniques are honed, algorithms are fine-tuned, and novel approaches are rigorously tested. However, the demand for these datasets has been steadily growing and their availability remains limited. This chapter provides an overview of the synthetic and experimental HS image databases used in this dissertation work and explores their characteristics, including the number of bands, spatial resolution, and types of scenes that they cover. Additionally, it delves into the preprocessing stages typically employed to prepare HS images for analysis.

3.1 Synthetic Databases

The synthetic databases consist of three datasets, the first of which corresponds to a VNIR image, which was generated with three components ($N = 3$) over a spatial domain of 120×120 pixels. The pixel spectral response ranged from 450 to 950 nm and consisted of 129 spectral bands [84]. It should be clarified that in this HS image, a linear mixing model (LMM) was followed to generate the dataset. Figure 3.1 displays the visualization of the end-members and abundances for both synthetic images.

In addition, two additional synthetic images were created based on two different types of HS signatures and a nonlinear mixing model (NMM): (i) Spectral Library Version 7 (SLV7) [101], and (ii) in-vivo human brain tissue in the VNIR range [84]. Both synthetic images were generated using four end-members ($N = 4$)

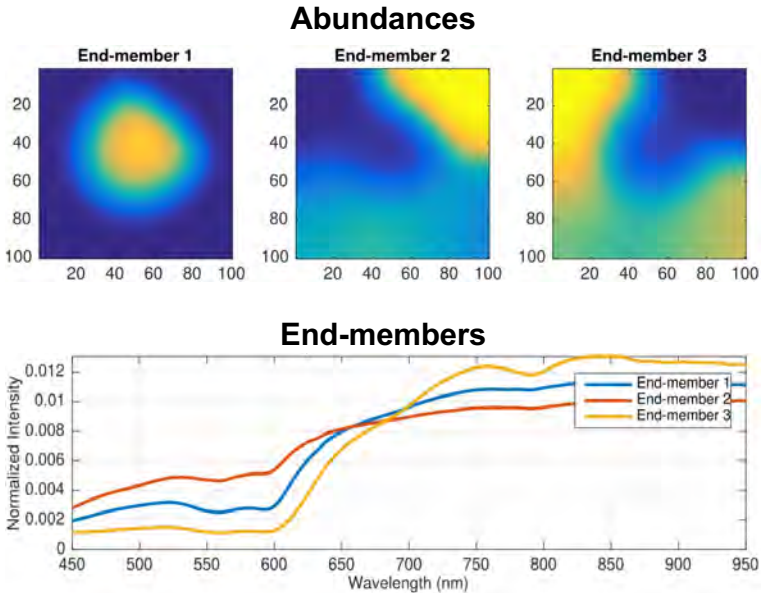


Figure 3.1: Components of end-members and abundances of VNIR synthetic image.

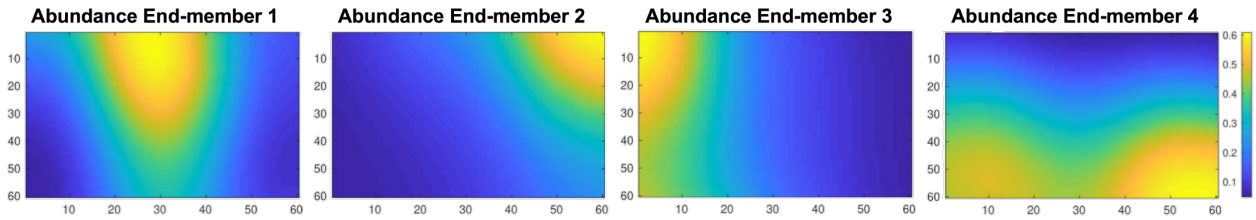


Figure 3.2: Mappings showing the ground-truth abundances of the four end-members during the synthetic nonlinear evaluation.

and a multilinear mixing model (MMM), which will be discussed in the next chapter. These synthetic images shared similar spatial characteristics, with a dimension of 60×60 (i.e., a total of 3600 spatial measurements) and the corresponding abundance maps (in Figure 3.2). For the SLV7 data set, each end-member had 470 spectral channels ranging from 0.21 to 2.69 μm [101]. On the other hand, the VNIR end-members consisted of 128 spectral channels spanning from 450 to 950 nm [84]. Figure 3.3 visually illustrates the end-members for both synthetic datasets.

The main advantage of these databases lies in the availability of ground truth for both abundances and end-members. These reference data enable a precise comparison of the various unmixing methods under study. Their primary objective is to create a controlled environment that facilitates the evaluation of results generated by algorithms based on both LMM and NMM.

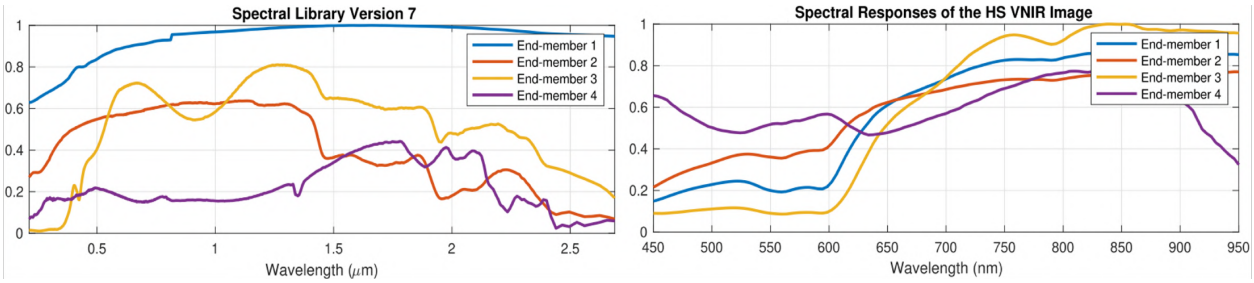


Figure 3.3: Profiles illustrating the spectral characteristics of the end-members during the nonlinear synthetic evaluation.

3.2 Real Remote Sensing Datasets

Typically, different algorithms for HS unmixing and classification are evaluated using remote sensing data. In this dissertation work, the HS images utilized were Jasper Ridge, Samson, Cuprite, Urban, and Pavia University. The Jasper Ridge, Samson, and Cuprite datasets consist of VNIR-HSI images captured by the NASA AVIRIS system (available at *https://rslab.ut.ac.ir/data* and *http://lesun.weebly.com/hyperspectral-data-set.html*). Meanwhile, information regarding the infrastructure or the specific regions where the Urban image was acquired is currently unavailable. Lastly, the HS image of Pavia University was acquired during a flight campaign over Pavia, located in northern Italy, using the ROSIS sensor.

The Jasper Ridge database provides spectral information from a region within the Jasper Ridge Biological Reserve in California, USA. This high-resolution HS image covers a diverse landscape and contains four main components, representing water, soil, road, and trees as four end-members. The database consists of 100×100 pixels and 198 spectral bands, spanning the spectral range from 380 to 2500 nm. Similarly, the Samson database is widely used in HSI research. Captured in the Samson area in California, an agricultural region with various types of crops, it includes three components water, soil, and trees. This database has a spatial dimension of 95×95 and 156 spectral bands that range from 401 to 889 nm. These datasets are valuable resources, much like the synthetic datasets used in this work.

The Cuprite database refers to an HS image captured in Cuprite, Nevada, United States [102]. This HS image was obtained using a VNIR spectrometer. The dataset consists of 188 spectral channels in the wavelength range 1.94 to $2.48\mu m$, after removing some noisy and water absorption channels. The spatial domain contains a region of 250×190 pixels, and there are 12 target end-members or components to be identified: “alunite", “andradite", “buddingtonite", “dumortierite", “kaolinite1", “kaolinite2", “muscovite", “montmorillonite", “nontronite", “pyrope", “sphene", and “chalcedony". For this dataset, ground-truth information is available only for the end-members and not for the abundance maps.

The Urban database comprises an HS image obtained from an undisclosed urban area [103]. The image covers a spatial area of 2×2 square meters, corresponding to a size of 307×307 pixels. It contains 210

spectral channels spanning the spectral range from 400 to 2500 nm. After excluding channels affected by dense water vapor and atmospheric effects, a total of 162 spectral channels are available. One notable advantage of this dataset is the availability of three versions of ground-truth information for land cover (end-members and abundances) with four, five, and six end-members respectively, representing “asphalt", “grass", “tree", “roof", “metal", and “dirt".

The final database for HS remote sensing is the Pavia University scene, which refers to an image captured in the city of Pavia, Italy. This dataset was obtained using a reflective optical system imaging spectrometer over the city and consists of 103 spectral bands ranging from 430 to 860 nm [104]. The HS image covers an area of 610×340 pixels with a spatial resolution of 1.3 m^2 per pixel. The scene contains nine ground-truth end-members: ‘Asphalt”, “meadows”, “gravel”, “trees”, “painted metal sheets”, “bare soil”, “bitumen”, ”self-blocking bricks” and “shadows”. Nevertheless, there is no ground-truth available for the abundance maps, meaning that the exact composition of the materials in the scene is not provided.

Real remote sensing databases, similarly to the case of synthetic databases, are used to evaluate the algorithms studied in the subsequent chapters of this study. However, unlike synthetic databases, the real remote sensing databases described in this section correspond to real-world scenarios. In these databases, the information has been pre-processed, and the ground truth for each HS image has been reliably established, meaning that accurate reference data are available to assess the performance of the algorithms.

3.3 Experimental Databases

3.3.1 Plastics Hyperspectral Database

The HS image dataset for plastic materials was obtained using two HS cameras. The first camera, operating in the VNIR range, captured spectral information between 400 and 1000 nm. The second camera, which captures the NIR range, covered wavelengths from 900 to 1700 nm. To illuminate the samples, a 150 W QTH lamp with a broad emission spectrum from 400 to 2200 nm was employed [105]. To avoid excessive heat, the QTH lamp was connected to a cold light emitter via an optical fiber. Both HS cameras and the cold light emitter were mounted on a scanning platform, enabling the push-broom technique to generate complete HS cubes. The working distance between the camera lens and the samples was maintained at 40 cm. It should be clarified that in this dissertation work only the data from the VNIR camera were used. As a result, more emphasis is placed on this dataset [105]. However, the data from the NIR camera are available for further analysis.

For the analysis, ten squares of synthetic materials made from various plastic polymers (labeled S1 to S10) and four different colors (white, black, red, and magenta) were utilized. The test-bench comprised three main materials: polylactic acid (PLA), acrylonitrile butadiene styrene (ABS), and polyethylene terephthalate glycol (PETG). These materials have been extensively studied in the field of HS characterization [105]. The materials used in this database are shown in Figure 3.4. The left portion presents a standard diffuse re-

flectance material called Spectralon White, which has a reflectance of 99% and serves as the white reference tile, denoted $R_W(x, y, \lambda)$. The right section includes PLA (S2, S5, and S6), ABS (S1, S8, S9, and S10), and PETG (S3, S4, and S7) materials, representing the colors white (S1 and S4), magenta (S2, S8, and S10), red (S3, S6, and S9), and black (S5 and S7).

Figure 3.4: Materials of the test-bench that conform to the HS plastic database. On the right side, there is the standard diffuse reflectance material, Spectralon White, with a reflectance of 99%, which serves as the white reference tile denoted as $R_W(x, y, \lambda)$. The left part of the image showcases ten plastic squares made from three materials: polylactic acid (S2, S5, and S6), acrylonitrile butadiene styrene (S1, S8, S9, and S10), and polyethylene terephthalate glycol (S3, S4, and S7). These squares come in different colors and contribute to the diversity of the database.

The HS sample of the test bench was imaged by simultaneously capturing the white reference and plastic squares, as shown in Fig. 3.4. The result was a raw image ($I_{raw}(x, y, \lambda)$) consisting of 440×1004 spatial pixels with 826 spectral bands ranging from 400 to 1000 nm. The dark reference ($R_D(x, y, \lambda)$) was captured across the entire area of interest with the shutter closed, while the white reference ($R_W(x, y, \lambda)$) was created by averaging the rows corresponding to the spectral bands of the reflecting surface in Fig. 3.4. This process resulted in a vector of 1004 pixels with 826 spectral bands. Since the push-broom camera captures the data line by line, the spectral response remains relatively consistent along the columns. By averaging the values, the variation is reduced and the reference spectral signatures are smoothed.

The hyperspectral plastic database was created to provide a reference point in a controlled experimental context. This is because the materials used in this database have been widely employed in studies of spectral variability and hyperspectral resolution, thanks to their well-known spectral signatures and minimal presence of geometric aberrations [105]. These characteristics make these materials ideal for investigating and comprehending variations in spectral information with a high degree of precision and control.

3.3.2 Overview Brain Tumor

A tumor is an abnormality in tissues that causes an unusual increase in volume anywhere in the body. Specifically, it is an accumulation of cells that grow uncontrollably [106, 107]. This cellular growth can originate in various ways: if it comes from multiple cells (polyclonal), it is called hyperplasi; conversely, when it arises from a single cell (monoclonal), it is referred to as neoplasm [108]. The distinction between these terms is related to the nature of the cellular growth process. The classification of tumors is based on their ability to infiltrate surrounding tissues [107]. In general terms, tumors are divided into two main categories: benign and malignant [109, 110]. Benign tumors tend to grow locally without invading nearby tissues or spreading to other parts of the body. On the other hand, malignant tumors, also known as cancers, have the ability to infiltrate and damage surrounding tissues and can spread through the bloodstream or lymphatic system to other areas of the body.

Neuronal tumors are a group of neoplasms that garner significant attention in research due to their complexity and the diversity of types and subtypes that can manifest in the central nervous system (CNS) and peripheral nervous system [111]. These tumors have a considerable impact on brain and nervous system function, making it crucial to study them in order to comprehend their pathophysiology and develop effective diagnostic and treatment strategies [109]. The importance of addressing these tumors is reflected in incidence statistics. In 2020, brain and CNS tumors ranked twelfth on the list of most common cancers in terms of mortality, with approximately 308,102 new cases worldwide and associated with 251,329 deaths across all ages and genders [112, 113]. In the young population, particularly those under 35 years old, these tumors represented the second leading cause of mortality, while in children under 14 years old, these tumors were the second most common cause of both diagnosis and mortality [113, 114]. In general, brain tumors constitute over 90% of all cases of CNS cancer, and they are associated with high rates of mortality and morbidity [112, 114].

Brain tumors can be categorized into two main groups: primary tumors and secondary tumors (metastatic tumors). Primary tumors originate within the brain itself, while secondary tumors originate in other parts of the body and then spread to the brain through metastasis. Primary tumors are further divided into low-grade and high-grade tumors based on their level of malignancy [115, 116]. Low-grade tumors encompass grades I and II, whereas high-grade tumors correspond to grades III and IV. Grade IV GB is the most common type, accounting for approximately 50% of all primary brain tumors [109, 115]. It is also the most deadly, with an average survival time of 16 months and a 5-year survival rate of 5.5%. GB is characterized by rapid growth and infiltration into underlying tissues, consisting of a heterogeneous mixture of poorly differentiated astrocytic tumor cells with pleomorphism, necrosis, vascular proliferation, and frequent mitosis [107].

Currently, the most common treatment for primary brain tumors is surgical resection. Early and complete removal of the tumor significantly improves overall survival rates, increasing the 5-year survival rate

to 50% [116]. To achieve maximum resection, neurosurgeons must accurately identify the boundaries of the tumor and the infiltrated tissue during surgery using image-guided techniques[117]. Additionally, neurosurgeons must avoid damaging normal brain tissue, as this can lead to neurological deficits in patients and negatively impact their quality of life [118]. Hence, HSI proves to be an important tool, allowing the identification of regions affected by tumor tissue without compromising the analyzed samples, being its feasibility for use as a surgical guide in clinical procedures [119].

Due to the aforementioned reasons, this dissertation investigates two HS image databases: histopathological samples and in-vivo brain tissue. These databases are described in detail below, with the aim of developing algorithms that enable the precise identification of brain tumor tissue, both at the cellular and macroscopic levels. Each database includes labeled information that is used to create training and validation sets. The main objective of these databases is to validate the analysis of complex tissues, such as brain tumor tissue, using HS images. Additionally, the study of these databases intends to expand the knowledge in this specific area with the expectation of standardizing algorithms to facilitate the development of practical applications in the clinical field.

3.3.2.1 Histopathology Hyperspectral Database

The histopathology database consists of a set of HS images acquired from various histological samples of human brain tissue, as described in [1, 2]. These samples were processed and analyzed by the Pathological Anatomy Service of the University Hospital Doctor Negrín in Las Palmas de Gran Canaria, Spain, with the ethical approval of the Clinical Research Ethics Committee (CEIC/CEI). The definitive diagnosis of the samples was determined by pathologists from the aforementioned institution according to the WHO classification of central nervous system tumors [120]. Once the diagnosis of GB was confirmed, a clinical expert made macroscopic annotations on the slides using a marker, delineating non-tumor areas in blue, while tumor areas were marked in red (as shown in Figure 3.5A). After delimiting both tissue classes in the samples, the regions were divided into different HS images.

HS images were obtained using an Olympus BX-53 microscope (Olympus, Tokyo, Japan) equipped with a Hyperspec VNIR A-Series scanning HS camera from HeadWall Photonics (Fitchburg, MA, USA). The camera was set to a $20\times$ magnification, resulting in HS images with dimensions of 375×299 μm. This imaging system is described in detail in [72]. The spectral range of the images spans 400 to 1000 nm, with a resolution of 2.8 nm. The captured HS images have a resolution of 1004×800 pixels and include 826 spectral channels. Figure 3.5A provides an example of HS images for the two classes analyzed in this study: *non-tumor* and *tumor*.

This database was derived from 13 tissue samples corresponding to various patients with GB tumors (grade IV). Based on the pathologist's comments, different regions of interest within the sample were selected, from which up to 10 HS images (as described in [2]) were acquired per region, resulting in 494 HS

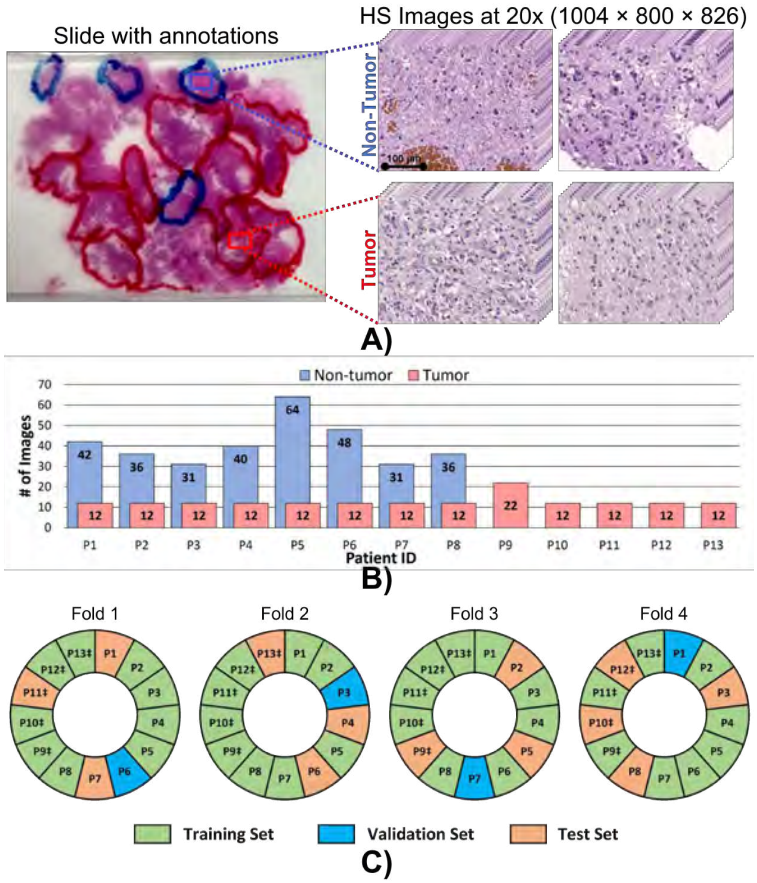


Figure 3.5: Database summary: A) An illustration of a pathological sample displaying annotations made by clinical experts, delineating regions of tumor (indicated by a red line) and non-tumor (indicated by a blue line) tissues. This sample was used for diagnosis, and HS images of both non-tumor and tumor regions were captured for classification purposes, utilizing a $20\times$ magnification. B) Total number of HS images obtained from each patient sample, categorized into tumor and non-tumor images. C) The data distribution for each patient in the proposed folds, including patients with solely tumor data (indicated by ‡).

images. Of these HS images, 328 corresponded to the non-tumor class and 166 were classified as tumor. The tissue distribution on the slide for each patient is shown in Figure 3.5B. As can be seen, the number of images marked as non-tumor varies from patient to patient. For some patients, only tumor class HS images were available. This distribution of data causes three main problems: (i) limited number of patients and samples; (ii) not all patients include both classes (non-tumor and tumor), only 8 patients present both histologies; (iii) the dataset is unbalanced between both classes due to the availability of more non-tumor images. For this reason, it was decided to spread the data over four folds to perform the evaluation and classification. These folds gather three different sets, each consisting of an independent subset of patients for training, validation, and testing. Three of these folds consist of 9 training samples, 1 validation sample, and 3 test samples, and the remaining fold contains 8 training samples, 1 for validation, and 4 for testing.

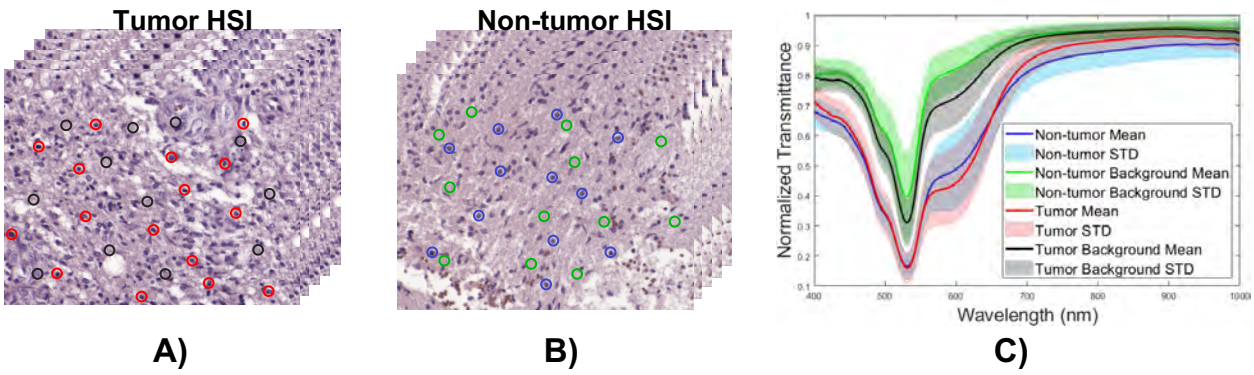


Figure 3.6: Example of the spectral information in the pathological HS images: A) and B) HS image with tumor and non-tumor tissue, respectively (the examples of selected regions are labeled with a circular marker to analyze their spectral information); and C) mean average spectral signatures and STDs: tumor cells (red), non-tumor cells (blue), tumor background tissue (black), and non-tumor background tissue (green).

The proposed distribution of the four folds is shown in Figure 3.5C. The random assignment of samples within each fold ensured that each one was presented once to the test subgroup, and patients with both types of labeled data (non-tumor and tumor) were included in the validation subgroup. Therefore, at least 3 of the patients with pure tumor information were assigned to the training subgroup. This organization of the four folds was used in previous works with the same database [1, 2].

Subsequent to the setting up of the database and folds, the data were visualized to assess the complexity of the classification problem. This challenge was solved by selecting different regions corresponding to cell nuclei and their external structures in images of both classes and different patients. An example of this visualization is shown in figure 3.6. After the data from different regions have been acquired (selected by the red, blue, green, and black circles in figures 3.6A and 3.6B), the spectral signatures of the selected regions were analyzed using mean and standard deviation (STD) for both classes. In Figure 3.6C, it can be observed that there is a great similarity between the spectral signatures of the non-tumor and tumor classes. This similarity can be measured using the correlation coefficient between the mean spectral responses of each class. Therefore, with the applied H&E staining process, a high correlation (correlation > 0.98) is observed that is very close to perfect registration. In other words, the spectral responses tend to be very similar in both cases, as the same reagents were used for both classes. This means that it is not possible to distinguish between both types of tissue simply by analyzing their spectral signatures, so more complex classification schemes are needed.

The study of this database was conducted with the aim of streamlining and enhancing the analysis of histological samples. This is because histopathological studies typically consume a considerable amount of time, and performing this task in large quantities can lead to human errors. The primary focus of this database is its practical application, as it seeks to simplify and expedite work in the field of histopathology.

This is crucial to ensure accuracy and efficiency in sample analysis, which, in turn, can have a positive impact on clinical decision-making and medical diagnostics.

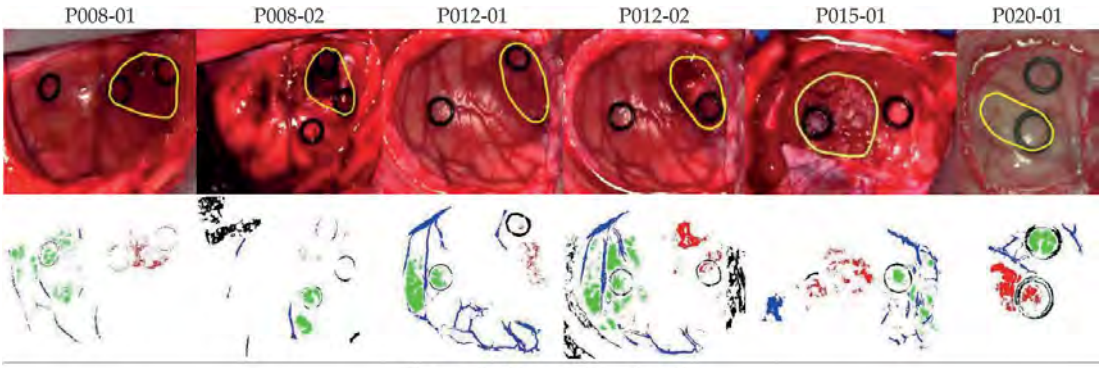
3.3.2.2 In-vivo Human Brain HS Dataset

The database of in-vivo human brain images was obtained using a custom-built intraoperative HS acquisition system developed in [121], as part of the HELICoiD European project (#618080) [122]. The system included an illumination system based on a 150 W quartz tungsten halogen QTH lamp with broadband emission between 400 and 2200 nm; a pushbroom HS camera in the VNIR spectral range of 400 to 1000 nm (Hyperspec® VNIR A-Series, Headwall Photonics Inc., Fitchburg, MA, USA); and a scanning platform to provide the necessary movement for pushbroom scanning, capable of covering an effective area of 230 mm^2. The generated HS cubes have a spatial resolution of 128.7 μm, a spectral resolution of 2-3 nm, and 826 spectral bands with a sampling of 0.73 nm.

The HS database used is reported in [52] and gathered 26 pictures from 16 adult patients. During clinical practice at the University Hospital Doctor Negrin at Las Palmas de Gran Canaria (Spain), patients underwent craniotomies to remove intraaxial brain tumors or other forms of brain surgery. Grade IV GB tumors were discovered in eight separate individuals and 11 HS pictures of exposed tumor tissue were recorded. The remaining patients had different tumors or other illnesses that required a craniotomy to expose the surface of the brain. All participant subjects provided their written informed consent, and the University Hospital's CEIC/CEI (*Comité Ético de Investigación Clínica-Comité de Ética en la Investigación*) Doctor Negrin, approved the study protocol and consent procedures.

The methodology outlined in [121] had to be followed to obtain HS pictures during surgical procedures. The operating surgeon initially pinpointed the general position of the tumor (if applicable) and the normal brain during the craniotomy and removal of the dura. On the surface of the brain, where the existence of the tumor and non-tumor was determined on the basis of preoperative imaging data, the surgeons inserted sterilized rubber ring markers. Once the tumor markers were found, the operator took an HS image while the clinical expert performed a biopsy on the tissue inside the tumor markers and sent the sample to a pathologist. The surgeon used a histopathological diagnosis to establish the kind and grading of tumor as well as to confirm its presence or absence.

After capturing the HS images, a specified group of pixels was tagged using a semi-automatic method based on the SAM algorithm published in [52]. Once the clinical procedure was completed, the surgeon used the semiautomatic tool to select small groups of very reliable pixels to build the ground-truth map for each acquired HS image. Subsequently, the SAM was calculated for the entire image with regard to the previously chosen pixel. The other pixels with the most similar spectral characteristics to the selected one can be found using a threshold manually set by experts. Tumor tissue (TT), normal tissue (NT), hypervascularized tissue (HT) (mainly blood vessels), and background (BG) were the four classes used to label pixels. The



Size (height × width × bands)					
460×549×826	480×553×826	443×497×826	445×498×826	376×494×826	378×330×826
#Labeled Pixels					
5477	10,769	15,753	24,464	8082	9635

Figure 3.7: Six HS test datasets of synthetic RGB images with the tumor area enclosed in yellow (first row) and ground-truth maps (second row). Green, red, blue, and black, respectively, are used to label the NT, TT, HT, and BG classes. Meanwhile, white pixels in the images represent unlabeled data [4].

background class covered additional elements or components that were included in the surgical scenario but were not related to the tumor resection technique, such as skull bone, dura, skin, or surgical supplies.

Due to poor image conditions, only four of the eight patients who were initially affected by GB tumors in this database had their GB tumor pixels labeled. The remaining GB tumor HS images were added to the database, but no tumor samples were taken into account. Six HS cubes in total (P008-01, P008-02, P012-01, P012-02, P015-01, and P020-01) were used as test databases and labeled with the four examined classes (NT, TT, HT, and BG). The datasets studied are depicted in Figure 3.7, where the tumor areas are enclosed by a yellow line in the synthetic RGB images, and the ground-truth maps of each HS image are displayed below.

The main objective of this study on the in-vivo human brain tissue database is the development of standardized applications that facilitate differentiation between healthy and diseased tissue during surgical procedures. This is of utmost importance since, in general, visually identifying tumor tissue is a complex task, and histopathological studies require a significant amount of time. Furthermore, due to ethical considerations, it is not feasible to analyze all tissues, both healthy and tumorous. In this context, the use of HSI allows for the identification of specific compounds without jeopardizing the examined areas. Therefore, the development of algorithms for the study and assessment of the mentioned database aims to enable practical applications once the methods are standardized and robust technology is established in this field.

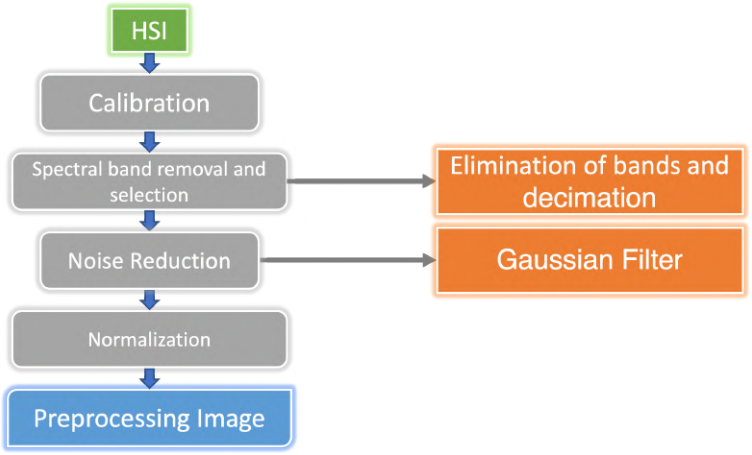


Figure 3.8: Blocks diagram of preprocessing chan.

3.4 Preprocessing Chain

The preprocessing of the databases was carried out following the HS image preprocessing pipeline described in Figure 3.8 and proposed in [4] and [123]. This process consists of four steps: image calibration, spectral band removal and selection, noise reduction, and spectral signature normalization.

The first step was the calibration of the HS image, which was performed to smooth out the raw spectral signatures and compensate for the nonlinear camera response. This step was carried out using two reference images. The first reference image (I_W) was acquired on a white surface at the same location as the clinical procedure, under the same lighting conditions. The second reference image (I_D) was obtained by capturing an image with the shutter closed, generating a dark reference image. The preprocessed HS image was then computed via a normalization step after these two images were acquired. This process is carried out based on the acquisition mode used, with reflectance and transmittance being the most commonly employed modes. Therefore, Equations (2.1) and (2.2) were used to obtain the calibrated HS images.

Subsequently, the spectral bands in the low- and high-frequency ranges were removed because of their low signal-to-noise ratio (SNR) caused by the poor performance of the CCD sensor in these sections of the spectrum. Next, a 1:5 decimation method was used to down-scale the spectral signature samples to eliminate redundant spectral values. The experimental databases of Plastic and In Vivo Human Brain were consequently condensed to 128 spectral bands. The data was then filtered through a spectral domain Gaussian filter to smooth them out. The final stage of the preprocessing pipeline comprised spectral normalization to take into account fluctuations in radiation intensity caused by irregular surfaces. Two methods were used to normalize the spectral signatures: the first involved scaling the values of the spectral signatures between zero and one, with the minimum value being zero and the maximum value being one, and the second involved normalizing the values of each spectral signature to have a sum of one.

Regarding the histopathology database, a preprocessing involving four steps was applied. First, the radiance images were transformed into normalized transmittance using a reference white image captured from a blank area of the pathological slide, and a dark image with the shutter closed. Subsequently, in the second step, neighboring spectral bands were averaged to reduce redundant information resulting from a high correlation between adjacent spectral bands. This averaging step led to a reduction in the number of spectral bands from 826 to 275. The third stage involved filtering the transmittance data using a Gaussian filter. Finally, the spectral signatures were normalized from zero to one, as described previously.

In this dissertation work, two different schemes were implemented to evaluate the various proposed methods: intra-patient and inter-patient. In the intra-patient approach, the labeled data from each image was used to generate its analysis. On the other hand, in the inter-patient case, all the labeled data from the HS database were considered, except for the labeled pixels of the HS image to be classified. Furthermore, it is important to clarify that preprocessing was applied only to the experimental databases. Synthetic and remote sensing data were appropriately generated (synthetic data) or had undergone preprocessing by the authors (remote sensing data).

Chapter 4

Spectral Unmixing

HS images, as discussed in previous chapters, have emerged as a new visualization tool in diverse research and industrial fields. However, this tool presents a significant challenge that can greatly influence the expected results. This challenge is the presence of mixed pixels resulting from a combination of different materials [124, 125, 126]. Moreover, if the spatial resolution of the HS camera is not sufficient to identify objects on the surface in a heterogeneous manner, it is inevitable that a pixel in the HS image contains the spectral information of more than one object [124, 126]. As a result, SU techniques have emerged, which allow the analysis and identification of the spectral information of pure materials also known as end-members and their respective fractional proportions or abundances in the analyzed scene [124, 127]. Due to these properties, SU has become one of the most studied topics in HSI [126].

In general, SU models are based on a mixture representation that describes the interactions among various pure elements present in a scene and the process of combination among these end-members [79, 128]. In the literature on the subject, two main categories of mixtures that are frequently encountered in HS images are identified: LMM and NMM [79, 128, 129]. These categories are defined based on the type of interaction that optical information undergoes when captured by the sensor. These mixture models serve as a fundamental foundation for estimating unmixing methods, which are formulated as an inversion problem for both LMM and NMM. This process involves recovering the spectral signatures of pure elements from the observed spectral signatures in the image.

In this chapter, as described in the diagram of Figure 4.1, the fundamentals of LMM and NMM are explored. This is done to establish the context of the origin of unmixing models. Subsequently, the unmixing approach developed by the UASLP research team, titled "extended blind end-member and abundance extraction" (EBEAE) [84], is reviewed. This method serves as the basis for two proposed unmixing approaches, which will be discussed in detail in this chapter. Following this section, the principles of image classification

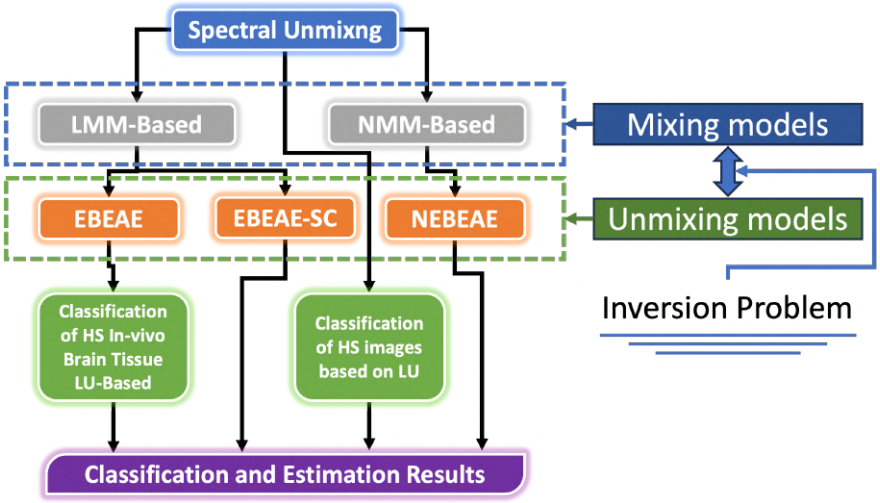


Figure 4.1: General diagram of chapter content

using unmixing algorithms are described, in order to understand how information is classified using these methods. Once the foundations are laid, two contributions related to unmixing proposals are described. The first of these is an extension of EBEAE that incorporates spatial coherence information, while the second is a nonlinear version of the same algorithm. Additionally, two strategies for tissue classification based on EBEAE are presented.

4.1 Linear Mixing Model

The LMM, as its name suggests, is a model that describes a linear combination of the spectral information from the end-members [79]. In this mixing approach, it is assumed that the reflected light from the scene reaching the sensor has only interacted with single elements or components [79, 129]. In other words, each incident ray of light interacts with unique materials; this type of interaction is exemplified in Figure 4.2. Due to its simplicity and interpretability, the LMM has become the most studied in the literature, demonstrating good results in multiple application fields.

Specifically, the interactions produced by the LMM generate convex linear combinations of the end-member information. This leads to two commonly applied constraints in unmixing algorithms to ensure a physical interpretation [128, 129]. The first is abundances non-negativity constraint (ANC). Since the fractional contributions are the result of the light reflected by the scene, this information must be positive [128]. The second is abundance sum-to-one constraint (ASC), which means that during unmixing, the abundances are a fractional contribution of the end-members, and their sum at each location must be one, excluding unknown components or scaling factors [79, 128, 129]. In addition to these constraints, in the literature often employs the end-members sum-of-one constraint (ESC), which aims to delimit the search

space for estimation and also reduce variability [84, 129, 130]. These constraints allow the SU algorithms to have a physical interpretation.

The LMM is based on the assumption that there are K spatial measurements of L-dimensional vectors

$$\mathbf{z}_k \in \mathbb{R}^L \quad \forall k \in [1, K], \tag{4.1}$$

where $\mathbf{z}_k \geq 0$. Generally, to maintain consistency with the physical interpretation and the commonly considered abundance constraints, all measurements $\mathcal{Z} = \{\mathbf{z}_1, ..., \mathbf{z}_K\}$ are normalized to sum up to one, such that

$$\mathbf{y}_k \triangleq \frac{1}{\mathbf{1}_L^\top \mathbf{z}_k} \mathbf{z}_k, \tag{4.2}$$

where $\mathbf{1}_L$ is an L-dimensional vector filled with ones and $(\cdot)^\top$ denotes the transpose operation. Once the above is established, it is possible to define the LMM for the k-th measurement in $\mathcal{Y} = \{\mathbf{y}_1, ..., \mathbf{y}_K\}$ by

$$\mathbf{y}_k = \sum_{n=1}^{N} \alpha_{k,n} \mathbf{p}_n + \mathbf{v}_k = \underbrace{\begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_N \end{bmatrix}}_{\mathbf{P} \in \mathbb{R}^{L \times N}} \underbrace{\begin{bmatrix} \alpha_{k,1} \\ \vdots \\ \alpha_{k,N} \end{bmatrix}}_{\boldsymbol{\alpha}_k \in \mathbb{R}^N} + \mathbf{v}_k = \mathbf{P}\, \boldsymbol{\alpha}_k + \mathbf{v}_k \quad \forall k \in \{1, \dots, K\}, \tag{4.3}$$

where $\mathbf{p}_n \in \mathbb{R}^L$ is the n-th end-member, $\alpha_{k,n} \geq 0$ its abundance in the k-th pixel, N represents the LMM order ($2 \leq N < L$), and $\mathbf{v}_k \in \mathbb{R}^L$ denotes a noise or uncertainty vector ($\mathbf{v}_k \sim \mathcal{N}(\mathbf{0})$) [79, 129, 131]. The scaled measurements, end-members, abundances, and noise vectors can be represented with a matrix notation $(\mathbf{Y}, \mathbf{P}, \mathbf{A}, \mathbf{V})$, respectively, so the LMM in Equation (4.3) can be written as:

$$\mathbf{Y} = \mathbf{P}\mathbf{A} + \mathbf{V} \tag{4.4}$$

where

$$\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_K] \in \mathbb{R}^{L \times K} \tag{4.5}$$

$$\mathbf{A} = [\boldsymbol{\alpha}_1 \dots \boldsymbol{\alpha}_K] \in \mathbb{R}^{N \times K} \tag{4.6}$$

$$\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_K] \in \mathbb{R}^{L \times K}. \tag{4.7}$$

Given the LMM presented in Equation 4.3 (see Figure 4.2), the goal of SU is to estimate the inverse of this mixing process in order to deduce the desired quantities based on the measured spectra [82]. SU typically involves three main stages: determining the number of end-members present, extracting the spectral characteristics of these end-members, and estimating the abundance of the end-members in each pixel [125]. The determination of the number of end-members is a complex problem, for which multiple solutions have emerged. These solutions not only estimate the number of components in the mixture, but also calculate their spectral information based on the measurements. Examples of such solutions include virtual dimensionality and HS signal identification by minimum error (HysSime) [132].

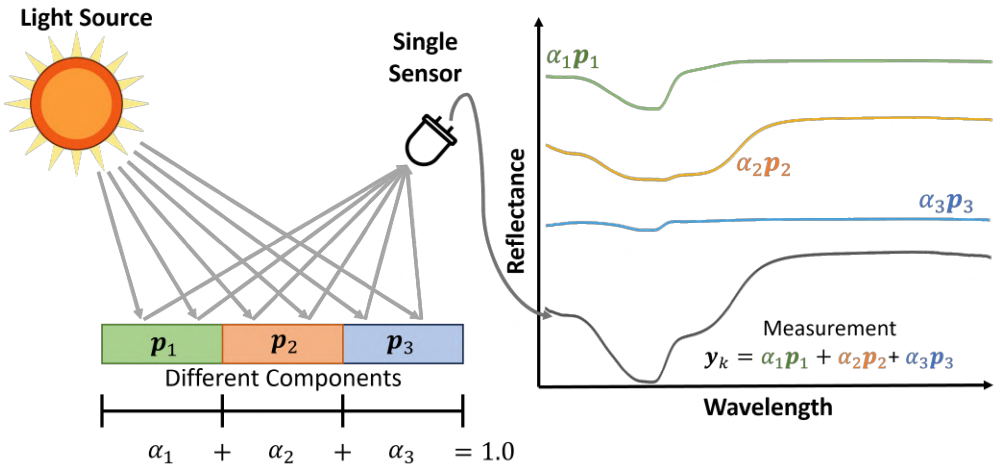


Figure 4.2: Diagram of the LMM, where it can be observed that the measurement captured by a sensor, corresponding to the information of a pixel, is a weighted average of the reflectances of the materials present in the pixel.

Regarding the estimation of end-members, the focus is solely on the unsupervised identification by assuming that the measurements contain pure spectral signatures corresponding to the components and assuming a certain number of end-members [79]. On the other hand, the estimation of abundances takes a supervised approach, where characteristic end-members are provided to estimate their contributions [79]. However, LMM also allow for the joint estimation of end-members and abundances in an unsupervised or semi-supervised manner [79]. The former is known as BLU, which enables the simultaneous estimation of end-members and abundances without previous knowledge [79]. The semi-supervised approach, in contrats, takes pre-known end-members (with certain similarity or affinity to those present in the data to be analyzed) as input and modifies them based on the measurements, acquiring the characteristic end-members and abundances from the database [79].

Furthermore, algorithms based on the LMM are usually classified into four categories based on the proposed approach for estimation: geometric, sparse, statistical, and DL models [129, 125]. On one hand, geometric unmixing algorithms are based on the assumption that there are pure or nearly pure end-members within the measurements, acting as vertices in a geometric space. Within this space, a convex simplex is generated that encloses the rest of the spectral data. Geometric unmixing methods are often employed to estimate pure end-members present in the measurements [77, 125, 129]. Some examples of these methods include Pixel Purity Index (PPI) [133], N-FINDR [134], Iterative Error Analysis [135], Vertex Component Analysis (VCA) [136], and the simplex growing algorithm [137]. Nevertheless, geometric algorithms face difficulties in extracting end-members from data with a high degree of mixing, as they lack access to pure spectral signatures [125, 129].

On the other hand, sparse unmixing methods emerged due to the availability of spectral libraries con-

taining various materials. The goal is to estimate the abundances of end-members using the prior information provided by these libraries [125]. It is assumed that each measurement combines linearly with a few spectra from the library. As a result, there will often be a higher number of end-members than pure elements within the HS image, leading to a scarcity of information in some abundances [129]. This is why these methods are referred to as sparse unmixing. However, this category of unmixing has the disadvantage of requiring spectral libraries composed of different spectra of interest, which is not always feasible. Additionally, there are multiple factors that can cause differences between the spectral library and the actual characteristics [125, 129]. To address these problems, a technique called sparse coding has been proposed in the literature [77]. It aims to learn the end-members directly from the measurements, creating the spectral library from the measured data. This approach eliminates the need for libraries with prior information on the characteristic components and avoids the issues of variability related to the different conditions under which the libraries and data were acquired [125]. This allows for the joint and unsupervised estimation of the characteristic end-members and their abundances, which is known as blind estimation.

In turn, statistical algorithms identify the end-members and abundances simultaneously using prior knowledge obtained from the statistical properties of the measurements [77]. This strategy has the advantage of not being conditioned by the initial conditions of the process, which can lead to a better solution in certain cases [125, 129]. Among the most popular statistical algorithms are Independent Component Analysis [138] and Bayesian approaches [139, 140].

Finally, DL methods have emerged in recent years and have shown their potential in SU tasks [125, 141]. In this regard, methods such as autoencoders, generative NNs, CNNs, and even fully connected networks have been widely used to process HS data for unmixing purposes. However, there are still some drawbacks with these types of techniques. One of the main challenges is that these algorithms require a large amount of training data. Additionally, they have a high number of parameters, and due to the properties of HS images, they often require significant computational resources [125, 141].

4.2 Nonlinear Mixing Models

Although LMM-based SU methods have shown good performance, they can present low performance when the interactions between different elements in the scene do not conform to linear mixing [82, 142]. Examples of such interactions occur when scenes exhibit geometric structures, including multiple shadows, reflections, absorptions, diffractions, and the presence of different incident illuminations, which can lead to changes in spectral properties [142, 143, 144]. Additionally, LMM typically considers only the reflected light information and not the transmitted one. This limitation can overlook intimate and multilayer relations in the analyzed scene (see Figure 4.3), where the incoming light ray may interact multiple times with different components, resulting in nonlinear relationships that are difficult to model using LMM [82, 128, 142, 143, 144]. To address this condition, various types of NMM have been proposed to better

understand how different end-members in a given scene interact with each other. Moreover, scenarios that can be represented by NMM are more common in nature compared to LMM cases [142, 143, 144].

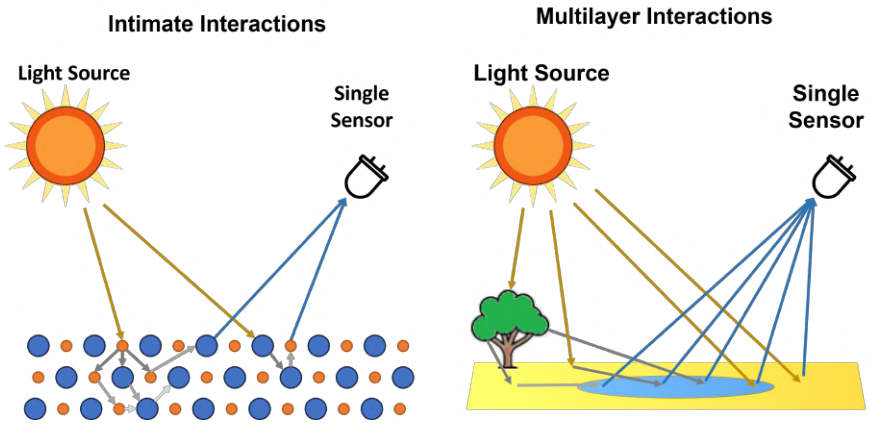


Figure 4.3: Diagrams of the NMM, where it can be observed that the measurement captured by a sensor, corresponding to the information of a pixel, is a weighted average of the reflectances of the materials present in the pixel.

Since the beginning of the past decade, various teams have proposed different NMM that consider the nonlinear phenomena caused by the physical interactions of light in the area under analysis. Among these models, the most representative ones are: the Fan et al. model [145], the Generalized Bilinear Model (GBM) [146], the Polynomial Post-Nonlinear Model (PPNM) [147], and more recently, the MMM [128]. These models build upon the linear perspective and include the nonlinear formulation of the physical interactions considered for each case. In particular, GBM and PPNM aim to scale the bilinear interactions with the linear abundances, assuming that the probability of interacting with two different materials should be proportional to their presence in the pixel [128, 130]. Likewise, it has been suggested that bilinear models can be expanded to higher-order terms or using higher-order polynomials. However, the literature argues that these expansions do not significantly contribute to the spectral signal and do not justify the added complexity [128]. MMM is also an extension of bilinear models that attempts to consider all possible interactions through a nonlinear interaction level [128].

Similar to the LMM, in this case, K spatial measurements are considered, expressed as real positive L-dimensional vectors $\mathbf{z}_k \in \mathbb{R}^L$ with $k \in \mathcal{K} \triangleq \{1, ..., K\}$. Each measurement $\mathbf{z}_k$ can be analyzed under the aforementioned nonlinear frameworks, but unlike LMM, NMM approaches do not strictly require initial scaling of the measurements. Additionally, in these frameworks, the contribution variables $\beta_{k,n}$ are considered as abundances, resulting in the following representations:

a) Fan et al. model [145]:

$$\mathbf{z}_k = \sum_{n=1}^{N} \beta_{k,n}\mathbf{p}_n + \sum_{n=1}^{N-1}\sum_{m=n+1}^{N} \beta_{k,n}\beta_{k,m}\mathbf{p}_n \odot \mathbf{p}_m + \mathbf{v}_k. \tag{4.8}$$

b) GBM [146]:

$$\mathbf{z}_k = \sum_{n=1}^{N} \beta_{k,n}\mathbf{p}_n + \sum_{n=1}^{N-1}\sum_{m=n+1}^{N} \gamma_{k,n,m}\beta_{k,n}\beta_{k,m}\mathbf{p}_n \odot \mathbf{p}_m + \mathbf{v}_k$$

where $\gamma_{k,n,m} \in [0, 1]$ denotes the degree of interaction between the n-th and m-th end-members at k-th measurement.

c) PPNM [147]:

$$\mathbf{z}_k = \sum_{n=1}^{N} \beta_{k,n}\mathbf{p}_n + \chi_k \left(\sum_{n=1}^{N} \beta_{k,n}\mathbf{p}_n\right) \odot \left(\sum_{m=1}^{N} \beta_{k,m}\mathbf{p}_m\right) + \mathbf{v}_k$$

where $\chi_k \in (-0.3, 0.3)$ refers to the strength of nonlinear interaction in the k-th measurement.

d) MMM [128]:

$$\mathbf{z}_k = (1 - d_k)\sum_{n=1}^{N} \beta_{k,n}\mathbf{p}_n + d_k \sum_{n=1}^{N} \beta_{k,n}\mathbf{z}_k \odot \mathbf{p}_n + \mathbf{v}_k \tag{4.9}$$

where $d_k \in (-\infty, 1]$ defines the nonlinear interaction level. As described in [128], if $d_k \in [0, 1]$, this factor will imply a decrease in reflectance with respect to the LMM, and conversely $d_k < 0$ an increment [128].

4.3 Extended blind end-member and abundance extraction

The approach known as extended blind end-member and abundance extraction, introduced in the work by Campos et al. [84] as EBEAE, provides a method to estimate end-members and their corresponding abundances in non-negative datasets using a LMM. Moreover, the BLU procedure implemented by EBEAE is regulated by hyperparameters that regulate the similarity between end-members and the entropy of the abundance, thus influencing the final results.

EBEAE is based on an LMM, as described in Section 4.1, by considering to the constraints of ANC, ASC, ESC, and measurement scaling. Based on the above, the synthesis problem of EBEAE is defined as the following constrained optimization process:

$$\min_{\{\boldsymbol{\alpha}_k\},\mathbf{P}} J(\{\boldsymbol{\alpha}_k\}, \mathbf{P}), \tag{4.10}$$

where

$$J(\{\boldsymbol{\alpha}_k\}, \mathbf{P}) = \frac{1}{2K} \sum_{k=1}^{K} \frac{||\mathbf{y}_k - \mathbf{P}\boldsymbol{\alpha}_k||^2}{||\mathbf{y}_k||^2} - \frac{\mu \lambda_{min}(\mathbf{P}^\top \mathbf{P})}{2K} \sum_{k=1}^{K} ||\boldsymbol{\alpha}_k||^2 \tag{4.11}$$

$$+ \frac{\rho}{2\vartheta} \sum_{n=1}^{N-1} \sum_{j=n+1}^{N} ||\mathbf{p}_n - \mathbf{p}_j||^2, \tag{4.12}$$

$\lambda_{min}(.)$ represents the minimum eigenvalue of the argument matrix, $\mu \in [0, 1)$ (*entropy weight*) and $\rho \geq 0$ (*similarity weight*) are hyperparameters, and

$$\vartheta = \begin{cases} (N-1) + \cdots + 1 & N \geq 3, \\ 1 & N = 2, \end{cases} \tag{4.13}$$

with restrictions $\mathbf{1}_N^\top \boldsymbol{\alpha}_k = 1$, $\mathbf{1}_L^\top \mathbf{p}_n = 1$ and $\boldsymbol{\alpha}_k$, $\mathbf{p}_n \geq 0$. Therefore, the hyperparameters in EBEAE are defined as (N, μ, ρ). To solve the optimization problem, an alternating least squares approach [148] is used to overcome the nonlinear dependence of the end-members $\{\mathbf{p}_n\}_{n=1}^N$ and the abundances $\{\boldsymbol{\alpha}_k\}_{k=1}^K$ in Equation (4.12), until a convergence condition is met or a maximum number of iterations is reached [84]. In this formulation, the end-members $\{\mathbf{p}_n\}_{n=1}^N$ identify N characteristic or representative components to reproduce all pixels in the measurements set $\mathcal{Y}$.

4.4 Classification of Hyperspectral Image Based on Linear Unmixing

The classification of HS data based on SU methods has been extensively explored in the literature, primarily focusing on its constituent elements, namely end-members and abundances. In this regard, both end-members and abundances allow for the identification and classification of different materials present in a scene or sample. Specifically, there are two types of approaches used for HS image classification based on SU estimates: abundance-based and end-member-based methods [149, 150, 151].

Abundance-based classification is a commonly used approach in HS image analysis. This strategy involves identifying the abundance of the material or component that has the highest incidence at each spatial location and assigning the label of that material to that location [84, 149]. In other words, the dominant class is determined based on the highest incidence in each pixel. It is important to note that classification using this approach can be affected by the amount of noise present in the image, the contribution level of each component to the mixture, and the accuracy of the SU algorithms [149, 150]. Due to its characteristics, abundance-based classification is a useful technique in HS data analysis, especially when evaluating HS information of materials with well-defined spectral signatures and significant spectral differences [150].

In contrast, end-member-based classification utilizes the characteristic spectral information of each known material to create a set of reference spectral signatures or ground-truth. This dataset is then used to compare the spectral signatures of each pixel using similarity metrics [151]. Subsequently, the class label

is assigned based on the best match with the materials within the ground-truth [151]. It is important to consider that the accuracy of this type of classification, similar to abundance-based, depends heavily on the quality of the reference spectral signatures, spectral variability in the image, the level of noise present, and the consideration of spectral mixtures [150, 151].

In recent years, classification approaches that integrate SU methods with various types of classifiers have increased, effectively incorporating the inherent characteristics of each algorithm into the application of HS images. These methodologies are commonly referred to as hybrid classification approaches as they merge abundance and end-member estimation with classification algorithms, such as ML methods, resulting in more precise classification outcomes. In these methodologies, relevant spatial features are extracted alongside with the spectral information using SU methods. These extracted patterns, comprising end-members and abundances, are then utilized in various classification algorithms to learn pattern characteristics and assign classes [152, 153]. The objective of these approaches is to address the challenges faced by each individual technique through their combined implementation [152, 153, 154]. Such proposals offer two significant advantages. Firstly, they reduce the dimensionality of input data for classification techniques [152, 153]. Secondly, by leveraging this dimensional reduction, the proposed classifiers can adopt simpler architectures, thereby facilitating their implementation on computing platforms and reducing execution time. Furthermore, this joint framework has the potential to enhance classification results by mitigating the risk of overfitting [152, 153]. In this dissertation work, different classification approaches based on SU are proposed that rely on ML.

4.5 Extended Blind End-member and Abundance Extraction with Spatial Coherence

In the field of BLU, EBEAE methodology provides a comprehensive framework for solving the unmixing problem. However, the existing EBEAE approach does not consider spatial coherence in its formulation, which limits its ability to handle different types of noise in the HS images. To address this limitation, an extension of the EBEAE methodology that incorporates spatial coherence using Total Variation (TV) theory is proposed. This enhanced methodology is referred to as extended blind end-member and abundance extraction with spatial coherence (EBEAE-SC).

EBEAE-SC formulates the BLU problem with a generic noise model and introduces a TV regularization component in the synthesis problem. This TV regularization effectively compensates and attenuates the effects of various types of noise, including Gaussian, shot, and impulse noise. By considering internal abundances for each calculated end-member, the TV regularization enhances the robustness of the algorithm. In this section, the mathematical formulation of EBEAE-SC is introduced, and the estimation scheme is described by applying constrained quadratic optimization (CQO), coordinate descent algorithm (CDA), and the split Bregman formulation [155].

In EBEAE-SC, three main variables are estimated: end-members, their nominal abundances, and in-

ternal abundances. Nominal abundances are estimated locally, treating each measurement independently, while a global approach is adopted for the overall end-members using the complete dataset. For internal abundances, a global approach is once again employed to capture spatial correlations. A notable advantage of EBEAE-SC over other state-of-the-art BLU methods is the joint estimation of end-members, nominal abundances, and noise-free internal abundances using the Bregman split [155]. This approach enhances the algorithm's robustness without affecting the accuracy of the end-members. The Bregman split has demonstrated superior performance in TV-based noise removal and artifact correction compared to other strategies in the literature.

4.5.1 Model Formulation

Similar to EBEAE, EBEAE-SC is based on the LMM described in Section 4.1, adhering to the constraints of measurement normalization ($\mathbf{1}_L^\top \mathbf{y}_k = 1 \quad \forall k \in [1, K]$), ANC ($\mathbf{1}_N^\top \boldsymbol{\alpha}_k = 1$), and ASC ($\mathbf{1}_N^\top \boldsymbol{\alpha}_k = 1$). Furthermore, the end-members in EBEAE-SC are normalized to sum up to one ($\mathbf{1}_L^\top \mathbf{p}_n = 1 \quad \forall n \in [1, N]$), following the aforementioned constraints. In the original BLU problem, the estimated abundances $\{\boldsymbol{\alpha}_k\}_{k=1}^K$ can be influenced by noise from different sources as well as light scattering from neighboring pixels. To address this issue, a new set of variables called internal abundances $\{\mathbf{w}_k\}_{k=1}^K$ is introduced in this formulation. These internal abundances $\{\mathbf{w}_k\}_{k=1}^K$ incorporate spatial coherence in both the vertical and horizontal components using the TV theory. The goal is for the internal abundances to remain close to the nominal abundances $\{\boldsymbol{\alpha}_k\}_{k=1}^K$, while also satisfying positivity and sum-to-one conditions.

The formulation of EBEAE-SC deviates from the approach presented in [84] by incorporating a modification of the energy functional in Equation (4.12) to include a TV component. The problem can be described as follows:

$$\min_{\mathbf{A},\mathbf{P},\mathbf{W}} \frac{1}{2K} \sum_{k=1}^{K} \frac{\|\mathbf{y}_k - \mathbf{P}\boldsymbol{\alpha}_k\|^2}{\|\mathbf{y}_k\|^2} + \frac{\rho}{2\vartheta} \sum_{n=1}^{N-1} \sum_{j=n+1}^{N} \|\mathbf{p}_n - \mathbf{p}_j\|^2 + \frac{\lambda}{2K} \sum_{k=1}^{K} \|\mathbf{w}_\mathbf{k} - \boldsymbol{\alpha}_k\|^2 + \frac{\tau}{K} HTV(\mathbf{W}), \tag{4.14}$$

where $\|\cdot\|$ stands for the Euclidean norm, and ρ, λ, and τ are hyper-parameters that control the strength of the regularization process. The normalization variable ϑ in Equation (4.14) takes the value described in Equation (4.13). In addition, matrix $\mathbf{W} = [\mathbf{w}_1, ..., \mathbf{w}_K] \in \mathbb{R}^{N \times K}$ gathers the internal abundances $\{\mathbf{w}_k\}_{k=1}^K$ in all pixels. The $HTV(.)$ component in Equation (4.14) is a TV regularization that adds SC in the problem formulation [156], which is defined as:

$$HTV(\mathbf{W}) = \sum_{n=1}^{N} \|\tilde{\mathbf{w}}_n \mathbf{D}_x^\top\|_1 + \sum_{n=1}^{N} \|\tilde{\mathbf{w}}_n \mathbf{D}_y^\top\|_1, \tag{4.15}$$

where the L_1 or Manhattan norm is denoted by $\|\cdot\|_1$. The vector $\tilde{\mathbf{w}}_n \in \mathbb{R}^{1 \times K}$ corresponds to the n-th row in the matrix $\mathbf{W}$ (internal abundances of all pixels for n-th end-member), where n takes values from 1 to N.

The matrices $\mathbf{D}_x \in \mathbb{R}^{K\times K}$ and $\mathbf{D}_y \in \mathbb{R}^{K\times K}$ represent 2-D horizontal and forward vertical finite difference operators applied to the internal abundances. To maintain the original dimensions, zero adjustments are made at the borders. In Equation (4.14), the parameter $\rho > 0$ controls the similarity among the resulting end-members, $\lambda > 0$ regulates the difference between the nominal and internal abundances at each spatial position, and $\tau > 0$ determines the strength of the SC during the estimation process.

Hence, the optimization problem stated in Equation (4.14) can be divided into three estimation sub-problems: (i) end-members $\mathbf{P}$, (ii) nominal abundances $\mathbf{A}$, and (iii) internal abundances $\mathbf{W}$, using a CDA. To solve these sub-problems, a CQO approach is employed, where at each step of the CDA, two sub-problems are held constant while the remaining one is optimized, following the methodology outlined in [84].

In this proposed approach, particular attention is given to the estimation of the internal abundances $\mathbf{W}$, which represents a significant difference compared to the original formulation [84]. By incorporating the $HTV(\cdot)$ term in the cost function of Equation (4.14), the internal abundances are able to capture the spatial variations between neighboring pixels using TV theory. Consequently, the resulting concentration maps exhibit reduced granularity and a higher correlation with their neighboring pixels. The estimation of the end-members follows the same approach as in the original EBEAE formulation [84]. The subsequent sections provide detailed descriptions of the proposed solutions for the three sub-problems.

4.5.2 End-member Estimation

In this particular sub-problem, the abundance matrices $\mathbf{A}$ and $\mathbf{W}$ are maintained as known constants during the estimation process. To achieve this goal, the expression in Equation (4.14) is modified to focus solely on the end-member matrix $\mathbf{P}$ [84]:

$$\min_{\mathbf{P}\geq 0, \mathbf{P}^\top \mathbf{1}_L = \mathbf{1}_N} \frac{1}{2K}\sum_{k=1}^{K} \frac{\|\mathbf{y}_k - \mathbf{P}\boldsymbol{\alpha_k}\|^2}{\|\mathbf{y}_k\|^2} + \frac{\rho}{2\vartheta}\sum_{n=1}^{N-1}\sum_{j=n+1}^{N} \|\mathbf{p}_n - \mathbf{p}_j\|^2. \tag{4.16}$$

To meet the constraints of the above problem, a Lagrange multiplier vector $\boldsymbol{\chi} \in \mathbb{R}^L$ is added to consider the equality constraint to obtain the stationary equations:

$$\mathbf{P}\left(\mathbf{A}\mathbf{T}\mathbf{A}^\top + \frac{\rho}{\vartheta}\mathbf{O}\right) - \mathbf{Y}\mathbf{T}\mathbf{A}^\top + \mathbf{1}_L\boldsymbol{\chi}^\top = \mathbf{0} \tag{4.17}$$

$$\mathbf{P}^\top \mathbf{1}_L - \mathbf{1}_N = \mathbf{0} \tag{4.18}$$

with $\mathbf{T} \triangleq \frac{1}{K}\text{diag}([\frac{1}{\|\mathbf{y}_1\|^2}, \ldots, \frac{1}{\|\mathbf{y}_K\|^2}])$, $\text{diag}(\cdot)$ denotes the resulting diagonal matrix, and $\mathbf{O} \triangleq N\mathbf{I}_N - \mathbf{1}_N\mathbf{1}_N^\top$. From Equations (4.17) and (4.18), a closed-solution for $\mathbf{P}$ is reached as follows:

$$\mathbf{P} = (\mathbf{I}_L - \frac{1}{L}\mathbf{1}_L\mathbf{1}_L^\top)\mathbf{Y}\mathbf{T}\mathbf{A}^\top(\mathbf{A}\mathbf{T}\mathbf{A}^\top + \frac{\rho}{\vartheta}\mathbf{O})^{-1} + \frac{1}{L}\mathbf{1}_L\mathbf{1}_N^\top. \tag{4.19}$$

After computing (4.19), a linear rectification is applied to all elements in $\mathbf{P}$, and each column is normalized to sum-to-one.

4.5.3 Abundance Estimation

In the second sub-problem, the matrices $\mathbf{W}$ and $\mathbf{P}$ are considered known and remain fixed during the estimation process. To address the estimation of the abundance vector at the k-th measurement, Equation (4.14) can be reformulated, and by introducing the notation $\hat{y}_k = \|\mathbf{y}_k\|^2$, the estimation process can be expressed as follows:

$$\min_{\boldsymbol{\alpha_k}\geq 0, \boldsymbol{\alpha}_k^\intercal \mathbf{1}_N = 1} \frac{1}{2}\|\mathbf{y}_k - \mathbf{P}\boldsymbol{\alpha}_k\|^2 + \frac{\lambda \hat{y}_k}{2}\|\mathbf{w}_k - \boldsymbol{\alpha}_k\|^2 \qquad \forall k \in [1, \ldots, K]. \tag{4.20}$$

Similar to the previous sub-problems, a Lagrange multiplier is introduced to enforce the constraints in Equation (4.20), resulting in a system of linear equations. The closed-form solution for this system is given by:

$$\boldsymbol{\alpha}_k = \boldsymbol{\Theta}\left(\mathbf{P}^\top \mathbf{y_k} + \lambda \hat{y}_k \mathbf{w_k} - \frac{\mathbf{y_k}^\top \mathbf{P}\boldsymbol{\Theta}\mathbf{1}_N + \lambda \hat{y}_k \mathbf{w_k}\boldsymbol{\Theta}\mathbf{1}_N - 1.0}{\mathbf{1}_N^\top \boldsymbol{\Theta}\mathbf{1}_N}\mathbf{1}_N\right), \tag{4.21}$$

with $\boldsymbol{\Theta} = (\mathbf{P}^\top\mathbf{P} + \lambda \hat{y}_k \mathbf{I}_N)^{-1}$. In Equation (4.21), a large value of λ will induce a similarity in the nominal $\{\boldsymbol{\alpha}_k\}_{k=1}^K$ and internal $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_K] \in \mathbb{R}^{N\times K}$ abundances, and if $\lambda = 0$, the induced SC is omitted.

4.5.4 Internal Abundance Estimation

In this final step, similar to previous iterations, the matrices $\mathbf{P}$ and $\mathbf{A}$ are considered known and unchanged. The objective is to estimate the n-th row $\tilde{\mathbf{w}}_n$ in $\mathbf{W}$ by rewriting Equation (4.14) in vector form. In this step, it is assumed $\lambda > 0$, so the optimization problem can be formulated as follows:

$$\min_{\tilde{\mathbf{w}}_n} \lambda||\tilde{\mathbf{w}}_n - \tilde{\mathbf{a}}_n||^2 + \tau\|\tilde{\mathbf{w}}_n\mathbf{D}_x^\top\|_1 + \tau\|\tilde{\mathbf{w}}_n\mathbf{D}_y^\top\|_1 \qquad \forall n \in [1, \ldots, N], \tag{4.22}$$

where $\tilde{\mathbf{a}}_n \in \mathbb{R}^{1\times K}$ represents the n-th row in the abundance matrix $\mathbf{A}$. The estimation problem described in Equation (4.22) poses a challenge as it involves a high-dimensional, non-differentiable optimization due to the non-separability of $\tilde{\mathbf{w}}_n$. To address this issue, new variables $\mathbf{R}_n = \tilde{\mathbf{w}}_n\mathbf{D}_x^\top$ and $\mathbf{Q}_n = \tilde{\mathbf{w}}_n\mathbf{D}_y^\top$ are introduced to restructure the optimization problem. Consequently, the original problem can be reformulated as an unconstrained optimization problem:

$$\min_{\tilde{\mathbf{w}}_n, \mathbf{R}_n, \mathbf{Q}_n} \frac{\lambda}{2}\|\tilde{\mathbf{w}}_n - \tilde{\mathbf{a}}_n\|^2 + \tau\|\mathbf{R}_n\|_1 + \tau\|\mathbf{Q}_n\|_1 + \frac{\nu}{2}\|\mathbf{R}_n - \tilde{\mathbf{w}}_n\mathbf{D}_x^\top\|^2 + \frac{\nu}{2}\|\mathbf{Q}_n - \tilde{\mathbf{w}}_n\mathbf{D}_y^\top\|^2, \tag{4.23}$$

where parameter ν serves as a regularization term, ensuring equal contributions from both the horizontal and vertical components. This balance is achieved by considering the 2D finite difference operator of TV theory. Additionally, parameter τ regulates the influence of the substituted variables, controlling their impact on the overall optimization process.

The optimization problem presented in Equation (4.23) involves three variables ($\tilde{\mathbf{w}}_n$, $\mathbf{R}_n$, $\mathbf{Q}_n$), but it can be decomposed into three separate sub-problems using the split Bregman approach [155, 157]. This

decomposition allows us to address each sub-problem individually. To achieve this decomposition, Bregman variables $\mathbf{B}_{1,n}$ and $\mathbf{B}_{2,n}$ are introduced into Equation (4.23), resulting in the following modified form:

$$\min_{\tilde{\mathbf{w}}_n,\mathbf{R}_n,\mathbf{Q}_n,\mathbf{B}_{1,n},\mathbf{B}_{2,n}} \frac{\lambda}{2}\|\tilde{\mathbf{w}}_n - \tilde{\mathbf{a}}_n\|^2 + \tau\|\mathbf{R}_n\|_1 + \tau\|\mathbf{Q}_n\|_1 + \frac{\nu}{2}\|\mathbf{R}_n - \tilde{\mathbf{w}}_n\mathbf{D}_x^\top - \mathbf{B}_{1,n}\|^2 + \frac{\nu}{2}\|\mathbf{Q}_n - \tilde{\mathbf{w}}_n\mathbf{D}_y^\top - \mathbf{B}_{2,n}\|^2. \tag{4.24}$$

By examining Equation (4.24), three distinct sub-problems can be identified:

- P1: $\min_{\mathbf{R}_n} \|\mathbf{R}_n\|_1 + \frac{\nu}{2\tau}\|\mathbf{R}_n - \tilde{\mathbf{w}}_n\mathbf{D}_x^\top - \mathbf{B}_{1,n}\|^2$,
- P2: $\min_{\mathbf{Q}_n} \|\mathbf{Q}_n\|_1 + \frac{\nu}{2\tau}\|\mathbf{Q}_n - \tilde{\mathbf{w}}_n\mathbf{D}_y^\top - \mathbf{B}_{2,n}\|^2$,
- P3: $\min_{\tilde{\mathbf{w}}_n} \lambda\|\tilde{\mathbf{w}}_n - \tilde{\mathbf{a}}_n\|^2 + \nu\|\mathbf{R}_n - \tilde{\mathbf{w}}_n\mathbf{D}_x^\top - \mathbf{B}_{1,n}\|^2 + \nu\|\mathbf{Q}_n - \tilde{\mathbf{w}}_n\mathbf{D}_y^\top - \mathbf{B}_{2,n}\|^2$.

Sub-problems P1 and P2 can be effectively addressed by employing the soft thresholding operation [155, 158] as follows:

$$\text{SoftTh}(\mathbf{U}, c) = \text{sign}(\mathbf{U}) \times \max\left(0, |\mathbf{U}| - \frac{c}{2}\right), \tag{4.25}$$

where the absolute value $|\cdot|$ and $\text{sign}(\cdot)$ functions are applied component-wise. In contrast, P3 sub-problem can be formulated as a least-squares problem, and its solution is derived from the following system of linear equations:

$$(\lambda\mathbf{I} + \nu\nabla)\tilde{\mathbf{w}}_n = \lambda\tilde{\mathbf{a}}_n + \nu\mathbf{D}_x^\top(\mathbf{R}_n - \mathbf{B}_{1,n}) + \nu\mathbf{D}_y^\top(\mathbf{Q}_n - \mathbf{B}_{2,n}), \tag{4.26}$$

with $\nabla \triangleq \mathbf{D}_x^\top\mathbf{D}_x + \mathbf{D}_y^\top\mathbf{D}_y$. The Bregman variables $\mathbf{B}_{1,n}$ and $\mathbf{B}_{2,n}$ are iteratively estimated to generate an internal loop over the index j. This observation involves solving sub-problems P1, P2, and P3 to compute $(\mathbf{R}_n^j, \mathbf{Q}_n^j, \tilde{\mathbf{w}}_n^j)$ and updating the variables, as described in [156, 158]:

$$\begin{aligned}\mathbf{B}_{1,n}^{j+1} &= \mathbf{B}_{1,n}^j + \mathbf{D}_x\tilde{\mathbf{w}}_n^{j+1} - \mathbf{R}_n^{j+1},\\ \mathbf{B}_{2,n}^{j+1} &= \mathbf{B}_{2,n}^j + \mathbf{D}_y\tilde{\mathbf{w}}_n^{j+1} - \mathbf{Q}_n^{j+1}.\end{aligned} \tag{4.27}$$

Once all the rows $\{\tilde{\mathbf{w}}_n\}_{n=1}^N$ in the internal abundances matrix $\mathbf{W}$ have been computed, a linear rectification procedure is performed to ensure that each element remains positive. Furthermore, each column $\{\mathbf{w}_k\}_{k=1}^K$ is normalized to ensure that the sum of its elements is equal to one.

4.5.5 Implementation and Analysis

The pseudocode for the implementation of EBEAE-SC is outlined in Algorithm 1, along with several noteworthy observations. Initially, the convergence of the CDA will be assessed based on the percentage change in the estimation error $J^i \triangleq \|\mathbf{Y} - \mathbf{P}^i\mathbf{A}^i\|_F$ at the i-th iteration, denoted as:

$$\frac{|J^i - J^{i-1}|}{J^{i-1}} < \epsilon, \tag{4.28}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, and $\epsilon > 0$ is a convergence threshold. Additionally, a maximum number of iterations is defined as a termination criterion.

Regarding the optimization process in Equation (4.23), the approach employed is based on the alternating direction method of multipliers (ADMM), where sub-problems P3, P1, and P2 are solved sequentially. The Bregman variables are updated using Equation (4.27) at the end of each iteration [159]. To estimate the rows of the internal abundance matrix $\{\tilde{\mathbf{w}}_n^i\}_{n=1}^N$ at the i-th iteration, an inner iterative loop is defined [157]. The stopping condition for this loop is based on the percentage change between the $(j-1)$-th and j-th iterations, calculated as $L_n^{i,j} \triangleq \|\tilde{\mathbf{w}}_n^{i,j} - \tilde{\mathbf{w}}_n^{i,j-1}\|$, with a convergence threshold of $\epsilon_{int} > 0$ and a limit on the maximum number of iterations. As explained in [157], since the data-fitting term in Equation (4.24) is quadratic, the split Bregman formulation guarantees convergence.

The iterative procedures involved in EBEAE-SC are summarized in Algorithm 1. If the hyperparameters are not specified, the original values from EBEAE are used [84]. Algorithm 1 primarily focuses on solving the subproblem of estimating internal abundances, where the innermost while loop is dedicated to smoothing the abundance map corresponding to each end-member. It is evident that the internal abundance estimation subproblem is more complex compared to the other two subproblems. This can be observed by analyzing the computational complexity of Algorithm 1 using Big $\mathcal{O}$ notation [160]. Specifically, the complexity of estimating end-members and their abundances is $\mathcal{O}(L \cdot K)$ and $\mathcal{O}(K \cdot N^4)$, respectively. Moreover, the complexity of estimating internal abundances is $\mathcal{O}(maxiter \cdot N \cdot K^2)$, which is higher than the other subproblems due to the large value of K compared to N. Therefore, considering the outer while loop in Algorithm 1, the overall complexity of EBEAE-SC is $\mathcal{O}\left(maxiter \cdot K \cdot (maxiter \cdot N \cdot K + L + N^4)\right)$, with the main contribution coming from the internal abundance subproblem. However, in our approach, K is significantly larger than L, and L is larger than the model order N ($K \gg L \gg N$). As a result, the complexity of EBEAE-SC can be reduced to $\mathcal{O}(maxiter^2 \cdot N \cdot K^2)$. In contrast, the original EBEAE approach has a complexity of $\mathcal{O}\left(maxiter \cdot K \cdot (L + N^4)\right)$, which is directly proportional to the number of maximum iterations and the number of measurements K. On the other hand, EBEAE-SC exhibits a quadratic complexity growth with respect to the same parameters.

Furthermore, the quadratic optimizations in Equations (4.16), (4.20), and Bregman's division [157] will exhibit convergence by analyzing the original formulation in Equation (4.14) at each stage of the CDA. As a result, the global convergence of the iterative CDA scheme presented in Algorithm 1 is guaranteed. The details regarding this point are explored extensively in the following section. However, it is important to note that the selection of the initial end-members matrix $\mathbf{P}^0$ and regularization variables $(\rho, \lambda, \tau, \nu)$ will significantly impact the optimal global solution. Finally, the internal abundance matrix $\mathbf{W}$ is chosen as the abundance matrix obtained from the EBEAE-SC method because it yields more uniform abundance maps with reduced granularity. Nonetheless, matrix $\mathbf{A}$ also produces acceptable results; however, it exhibits a higher level of granularity.

Algorithm 1 EBEAE-SC

Input: Matrix of scaled measurements $\mathcal{Y}$, number of total measurements K, number of end-members N, initial end-members matrix $\mathbf{P}^0$, maximum number of iterations $maxiter$, convergence threshold ϵ, and hyperparameters ρ, λ, ν, τ.

Output: End-member matrix $\mathbf{P}$, abundance matrix $\mathbf{A}$ and internal abundance matrix $\mathbf{W}$.

Normalize measurements to sum-to-one

$$\mathbf{y}_k = \frac{1}{\mathbf{1}_L^\top \mathbf{z}_k}\mathbf{z}_k$$

Initialize matrix $\mathbf{W}$ to zero, the end-members matrix by $\mathbf{P}^0$, and set $i = 0$.

if Estimation with spatial coherence is desired (i.e. $\lambda > 0$, $\nu > 0$, and $\tau > 0$) **then**

  while Stopping criterion in Equation (4.28) is not satisfied or $i < maxiter$ **do**

    Estimate the columns $\boldsymbol{\alpha}_k^i$ of the abundance matrix $\mathbf{A}^i$ per each spatial measurement $k \in \{1, \ldots, K\}$ by (4.21).

    Estimate the end-members matrix $\mathbf{P}^i$ over the whole dataset by Equation (4.19).

    for $n = 1$ until N **do**

      Set $j = 1$

      while $\frac{|L_n^{i,j} - L_n^{i,j-1}|}{L_n^{i,j-1}} < \epsilon_{int}$ or $j < maxiter$ **do**

        Estimate internal abundance row $\tilde{\mathbf{w}}_n^{i,j}$ by Equation (4.26),

        Calculate P1 and P2 sub-problems,

        Update $\mathbf{B}_{1,n}^{j+1}$ and $\mathbf{B}_{2,n}^{j+1}$ by Equation (4.27),

        Set $j = j + 1$.

      end while

    end for

    Set $i = i + 1$.

  end while

else

  Evaluate the original EBEAE formulation in [84].

end if

4.5.6 Synthetic, Remote Sensing and Experimental Results, and Discussion

This section shows the application of the proposed EBEAE-SC method on a synthetic VNIR image and two other widely used HS images in the remote sensing literature [161, 162]. To quantify the accuracy in the proposed unmixing process, the resulting end-members and their abundances at all spatial locations are collected in sets $\mathcal{P}$ and $\mathcal{A}$, respectively. The labeled ground-truth sets are defined as $\bar{\mathcal{A}}$ and $\bar{\mathcal{P}}$. Subsequently, estimation errors are calculated as follows:

$$E_p = \frac{1}{\text{card}(\overline{\mathcal{P}}) + \text{card}(\mathcal{P})} \min_{\forall \overline{\mathbf{p}} \in \overline{\mathcal{P}}, \mathbf{p} \in \mathcal{P}} \|\overline{\mathbf{p}} - \mathbf{p}\|, \tag{4.29}$$

$$E_a = \frac{1}{\text{card}(\overline{\mathcal{A}}) + \text{card}(\mathcal{A})} \min_{\forall \bar{\mathbf{a}} \in \overline{\mathcal{A}}, \mathbf{a} \in \mathcal{A}} \|\bar{\mathbf{a}} - \mathbf{a}\|, \tag{4.30}$$

where $\text{card}(\cdot)$ denotes the cardinality of a set [163]. The metrics described above are used to evaluate the performance of unmixing algorithms.

Five state-of-the-art BLU methods that incorporate spatial information were considered for comparison with the EBEAE-SC results. These methods are: (i) graph-regularized $L_{1/2}$ nonnegative matrix factorization (GLNMF) [164], (ii) preserving the intrinsic structure invariant nonnegative matrix factorization (PISINMF) [165], (iii) nonnegative matrix factorization-quadratic minimum volume (NMF-QMV) [166], (iv) graph total variation regularized NMF solved by Merriman-Bence-Osher (gtvMBO) [162], and (v) sparsity-enhanced convolutional decomposition (SeCoDe) [167]. Different mathematical principles have been employed by all these algorithms, and previous studies [162, 165, 167, 168] have compared them with multiple unmixing methods. Furthermore, an evaluation with an in-vivo human brain HS image was conducted for EBEAE-SC. The scripts used to generate the results presented in the figures and tables can be found in the following GitHub repository: https://github.com/Alex-CruzG/EBEAE-SC.

In order to assess the effectiveness of the proposed method under adverse conditions, two types of noise, namely Gaussian and shot noise, are taken into consideration. To generate a measurement incorporating both types of noise, denoted as $\mathbf{y}_k$, from a noise-less observation $\mathbf{y}_k^0$, the subsequent transformation is applied:

$$\mathbf{y}_k = \mathbf{y}_k^0 + \mathbf{n}_k + \mathbf{m}_k \qquad \forall k \in \{1, \ldots, K\}, \tag{4.31}$$

where $\mathbf{n}_k \in \mathbb{R}^L$ and $\mathbf{m}_k \in \mathbb{R}^L$ are associated with the Gaussian and shot noise components, respectively. The Gaussian noise vector $\mathbf{n}_k$ is characterized by a zero mean and a STD of σ_k^{SNR}, represented by $\mathcal{E}\{\mathbf{n}_k\} = \mathbf{0}_L$ and $\mathcal{E}\{\mathbf{n}_k^\top \mathbf{n}_k\} = (\sigma_k^{SNR})^2 \mathbf{I}_L$, respectively. Here, $\mathcal{E}\{\cdot\}$ denotes the expected value. On the other hand, the shot noise component $\mathbf{m}_k$ is modeled as a signal-dependent term, which can be approximated using a continuous distribution according to [169, 170]:

$$\mathcal{E}\{\mathbf{m}_k\} = \mathbf{0} \quad \& \quad \mathcal{E}\{\mathbf{m}_k^\top \mathbf{m}_k\} = (\sigma_k^{PSNR})^2 \text{diag}(\mathbf{y}_k^0). \tag{4.32}$$

Therefore, the variances of the components within $\mathbf{m}_k$ are determined by the amplitudes in $\mathbf{y}_k^0$, which correspond to positive values. In this model, the mean value of the shot noise component is derived from $\mathbf{y}_k^0$ in Equation (4.31) and is therefore excluded from the noise distribution. These definitions incorporate σ_k^{SNR} and σ_k^{PSNR} as specified by the desired signal-to-noise ratio (SNR) and peak signal-to-noise ratio ($PSNR$) within the datasets:

$$\sigma_k^{SNR} = \sqrt{\frac{\frac{1}{L-1}\|\mathbf{y}_k^0\|^2}{10^{\frac{SNR}{10}}}}, \qquad \sigma_k^{PSNR} = \sqrt{\frac{\max_{l\in\{1,\ldots,L\}}(\mathbf{y}_k^0)_l^2}{10^{\frac{PSNR}{10}}}}, \tag{4.33}$$

where $(\cdot)_l$ represents the l-th component within the vector. To assess the performance of EBEAE-SC and the state-of-the-art methods, evaluations were conducted under different noise scenarios. The noise levels were set to the same values of the SNR and the $PSNR$, specifically $(SNR, PSNR) \in \{(20, 20), (17.5, 17.5), (15, 15), (12.5, 12.5), (10, 10), (7.5, 7.5), (5, 5)\}$. These conditions impose high demands on most BLU algorithms. A Monte Carlo evaluation was conducted for each combination of SNR and $PSNR$ values. This evaluation involved calculating the estimation errors E_p in Equation (4.29) and E_a in Equation (4.30) by performing 50 noise realizations for both the synthetic and real datasets.

The hyperparameters of the examined methods were determined using the *fminsearch* function in Matlab within an intermediate noise scenario ($SNR = 12.5$ dB and $PSNR = 12.5$ dB). Hence, the values with the lowest overall estimation error were selected. The hyperparameters that were chosen are presented in Table 4.1, along with the estimation errors for abundances and end-members at the highest noise level during the Monte Carlo evaluation. Consequently, EBEAE-SC demonstrated a significant reduction of 11% in abundance estimation errors compared to the best-performing state-of-the-art BLU methods. Regarding end-members, PISINMF and EBEAE-SC exhibited similar error levels, with the former achieving the lowest error. In terms of execution time, EBEAE-SC showed an increase compared to gtvMBO, which is the fastest algorithm in this scenario. However, compared to other state-of-the-art BLU methods, EBEAE-SC ranked second, improving the execution time of SeCoDe by 67%. In summary, EBEAE-SC achieved the second shortest computational time among synthetic VNIR dataset, while exhibiting the best performance in terms of abundance errors.

The subplots in Figure 4.4 show the abundance maps for a single realization at the noise level of $SNR = 12.5$ dB and $PSNR = 12.5$ dB in the VNIR scenario. It is observed that, in general, the state-of-the-art BLU methods yielded higher errors. Furthermore, in areas where high and low concentrations transitioned, all methods exhibited a noticeable level of noise-induced granularity. Nevertheless, the abundance maps generated by EBEAE-SC demonstrated greater homogeneity compared to the other methods and exhibited better consistency with ground-truth data.

Boxplots depicting the abundance and end-member errors from the Monte Carlo test are illustrated in Figure 4.5 for the synthetic VNIR-HSI datasets. As the noise levels ($SNR, PSNR$) increased, all the methods demonstrated a corresponding increase in errors (E_p, E_a). Regarding abundance errors, EBEAE-SC consistently delivered the most accurate results. In terms of end-member errors, EBEAE-SC reached the lowest errors in the majority of the noise scenarios, although SeCoDe, PISINMF, GLNMF, and EBEAE-SC displayed notable similarity in their end-member errors across all noise levels. In summary, based on the evaluations conducted using the synthetic VNIR dataset, EBEAE-SC exhibited the most consistent results

Table 4.1: Optimal hyperparameters estimated for the Monte Carlo evaluation, resulting in: (i) estimation errors and their STD for abundances and end-members at a noise level of $SNR = 5$ dB and $PSNR = 5$ dB, and (ii) computational time and its STD for the VNIR synthetic dataset. The method with the lowest value in each category is highlighted in bold font.

	VNIR			
Method	**Hyperparameters**	E_a (mean±STD)	$E_p \times 10^{-4}$ (mean±STD)	**Computational Time (seconds)** (mean±STD)
GLNMF	$\mu = -0.0015$, $\lambda = -0.6463$	4.712±0.282	24.493±5.151	91.972±0.268
PISINMF	$\mu = 0.01$, $\alpha_0 = 0.001$, $\delta = 45$, $\tau = 25$, window size = 5×5	4.703±0.283	**23.440±4.956**	8.178±0.057
NMF-QMF	$\beta \in [10^{-5}, 10^{5}]$	5.040±0.266	39.230±6.189	15.391±0.969
gtvMBO	$\lambda = 4.333$, $\rho = 0$, $\gamma = 10.331$	4.901±0.209	34.817±6.927	**0.493±0.059**
SeCoDe	$\alpha = 5.103 \times 10^{-6}$, $\beta = 3.102 \times 10^{-4}$, $\gamma = 0.205$	4.706±0.284	24.285±5.055	4.182±0.157
EBEAE-SC	$\rho = 0.085$, $\lambda = 0.347$, $\mu = 1.237 \times 10^{-7}$, $\nu = 1.3 \times 10^{-3}$, $\tau = 0.097$	**4.201±0.252**	23.460±5.052	1.375±0.219

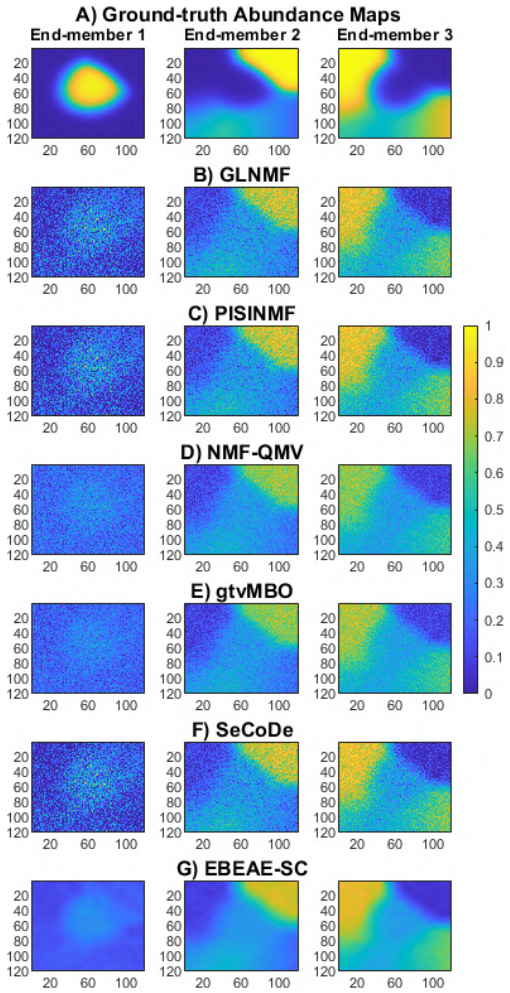


Figure 4.4: The Monte Carlo estimation results for VNIR synthetic dataset ($N = 3$, $SNR = 12.5$ dB, and $PSNR = 12.5$ dB) are illustrated in seven panels: A) Ground-truth, B) GLNMF, C) PISINMF, D) NMF-QMV, E) gtvMBO, F) SeCoDe, and G) EBEAE-SC, representing one realization of the abundance maps.

in terms of end-member and abundance errors, as well as computational time, when compared to state-of-the-art BLU algorithms.

4.5.6.1 Remote Sensing Datasets

Jasper Ridge and Samson HS images are widely used as experimental images to characterize unmixing algorithms in studies such as [161, 162, 171]. Similar to the synthetic datasets, the noise levels mentioned above were added to the real data. Additionally, a Monte Carlo test comprising 50 realizations was carried out and the hyperparameters of each methodology were fine-tuned using the *fminsearch* function in Matlab within an intermediate noise condition ($SNR = 12.5$ dB and $PSNR = 12.5$ dB), selecting the values that

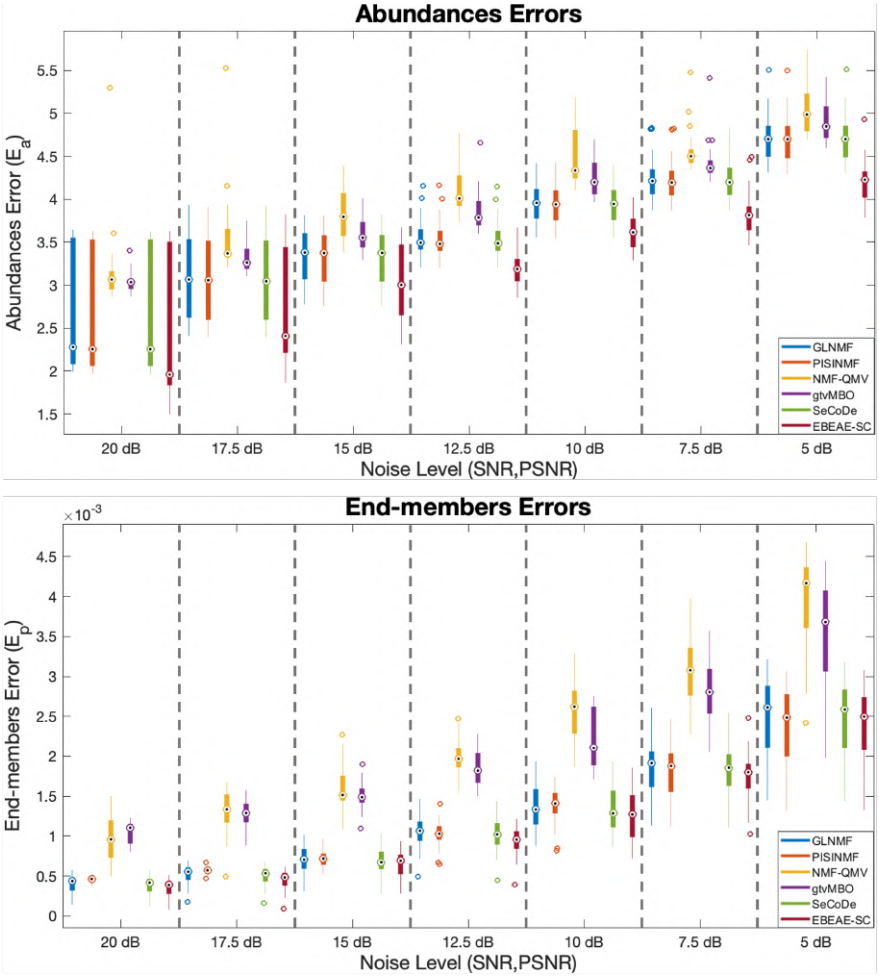


Figure 4.5: The boxplots of the Monte Carlo test for estimation errors of abundances (top plot) and end-members (bottom plot) in VNIR-HSI synthetic datasets ($N = 3$) across various SNR and $PSNR$ values.

resulted in the lowest overall estimation error. For the Jasper and Samson HS images, the VCA and FCLSU algorithms were employed as the initialization methods for end-members and abundances, respectively, following the same approach as with the synthetic dataset. The hyperparameters selected for both datasets are presented in Table 4.2. The execution time results for both HSIs demonstrate that gtvMBO achieved the shortest processing time, with EBEAE-SC closely following. On the other hand, GLNMF exhibited the longest duration to generate the estimates.

Table 4.3 presents the results of the Monte Carlo test for the mean estimation errors E_a and E_p in Jasper Ridge HS image. As the noise intensity increased (indicated by lower pairs of $(SNR, PSNR)$), the estimation errors generally increased for all state-of-the-art BLU methods, which is expected. However, our proposed method, EBEAE-SC, consistently reduced the estimation errors for both abundances and end-members in both low and high noise scenarios, outperforming all state-of-the-art BLU methods. In fact,

Table 4.2: Optimal hyperparameters estimated for the Monte Carlo test and mean computational time and its STD for the Jasper Ridge and Samson HS images. The method with the lowest value is highlighted in bold font.

Method	Hyperparameters	Computational Time (seconds) (mean±STD)
	Jasper Ridge HSI	
GLNMF	$\mu = 2.579 \times 10^{-5}$ and $\lambda = -0.2462$	65.478±2.578
PISINMF	$\mu = 0.001$, $\alpha_0 = 1$, $\delta = 45$, $\tau = 20$, window size = 5×5	2.015±0.238
NMF-QMF	$\beta \in [10^{-5}, 10^{5}]$	4.283±0.042
gtvMBO	$\lambda = 5.6 \times 10^{-3}$, $\rho = 7.365 \times 10^{-9}$, $\gamma = 16.582$	**0.257±0.004**
SeCoDe	$\alpha = -0.021$, $\beta = 6.324 \times 10^{-4}$, $\gamma = 0.826$	4.512±1.513
EBEAE-SC	$\rho = 0.225$, $\lambda = 9 \times 10^{-4}$, $\mu = 1.081 \times 10^{-4}$, $\nu = 1.59 \times 10^{-10}$, $\tau = 2.392 \times 10^{-6}$	0.789±0.128
	Samson HSI	
GLNMF	$\mu = 1.163 \times 10^{-4}$, $\lambda = -8.037 \times 10^{-2}$	43.401±1.361
PISINMF	$\mu = 1 \times 10^{-4}$, $\alpha_0 = 0.001$, $\delta = 10$, $\tau = 35$, window size = 5×5	8.178±0.057
NMF-QMF	$\beta \in [10^{-5}, 10^{5}]$	15.391±0.969
gtvMBO	$\lambda = 19.833$, $\rho = 9.806 \times 10^{-9}$, $\gamma = 6.586$	**0.493±0.059**
SeCoDe	$\alpha = -0.033$, $\beta = 5.218 \times 10^{-4}$, $\gamma = 3.168$	4.182±0.157
EBEAE-SC	$\rho = 0.766$, $\lambda = 4.5 \times 10^{-3}$, $\mu = 1.455 \times 10^{-6}$, $\nu = 3.533 \times 10^{-8}$, $\tau = 0.014$	1.375±0.219

Table 4.3: Mean estimation errors for abundances and end-members in the Monte Carlo test for various SNR and $PSNR$ values in the Jasper Ridge HS image. The lowest value for each pair of SNR and $PSNR$ is indicated in bold font, while results without statistically significant difference from EBEAE-SC (determined by ANOVA with $p > 0.05$) are highlighted in blue.

	Method											
Noise level	GLNMF		PISINMF		NMF-QMV		gtvMBO		SeCoDe		EBEAE-SC	
$(SNR, PSNR)$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$
20	0.259	1.448	0.249	1.868	1.311	4.702	0.796	4.349	0.597	16.085	**0.183**	**0.946**
17.5	0.357	2.072	0.346	2.474	1.497	6.854	0.942	6.009	0.552	15.524	**0.274**	**1.580**
15	0.474	3.159	0.461	3.514	1.693	10.419	1.111	9.154	0.570	15.659	**0.384**	**2.756**
12.5	0.604	3.869	0.592	4.777	1.837	14.752	1.249	11.886	0.803	17.639	**0.530**	**3.456**
10	0.757	5.484	0.743	6.269	1.978	20.044	1.367	13.550	0.839	17.030	**0.688**	**4.926**
7.5	1.034	7.426	1.021	8.256	2.226	26.435	1.626	15.881	1.106	18.365	**0.982**	**6.679**
5	1.168	7.677	1.155	10.190	2.453	31.597	1.861	20.190	1.198	18.022	**1.117**	**6.585**

EBEAE-SC exhibited similar results to GLNMF and PISINMF in most cases, while also achieving shorter execution times (see Table 4.2). Figure 4.6 illustrates a comparison of the abundance maps generated by the state-of-the-art BLU methods for a specific noise level of $SNR = 12.5$ dB and $PSNR = 12.5$ dB. It can be observed that GLNMF, PISINMF, SeCoDe, and EBEAE-SC produced estimations that closely resemble the ground-truth. Although there are subtle differences among the four methods, EBEAE-SC accurately estimated regions with very low concentrations and transition zones. Furthermore, our proposed method reduced the granularity present in the abundance maps compared to the state-of-the-art BLU methods.

Table 4.4 displays the results of the Monte Carlo test for the mean errors E_a and E_p in the Samson HS image. Similar to Jasper Ridge, EBEAE-SC exhibited estimates that were closer to the ground-truth for both abundances and end-members. However, EBEAE-SC did not demonstrate significant differences in end-member errors compared to GLNMF, PISINMF, and SeCoDe. Nonetheless, this proposed method

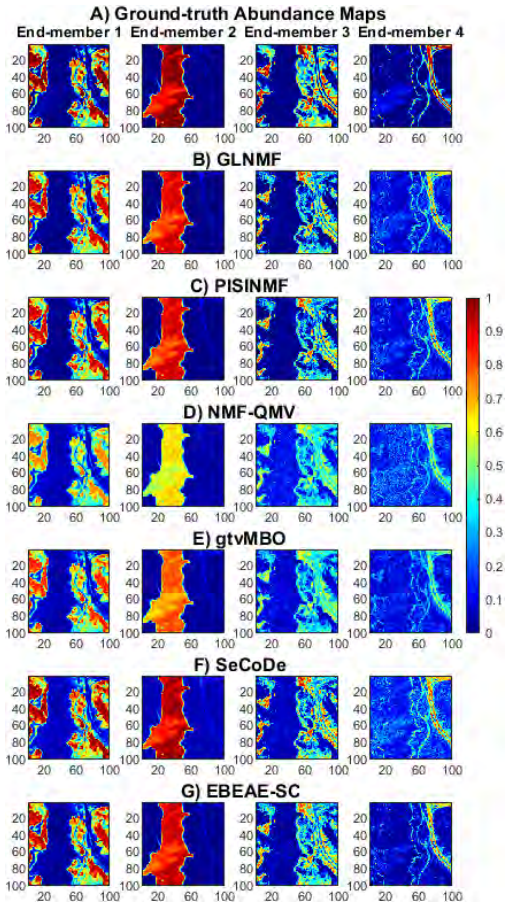


Figure 4.6: One realization of the Monte Carlo estimation results (abundance maps) in Jasper Ridge HS image ($N = 4$, $SNR = 12.5$ dB and $PSNR = 12.5$ dB): A) Ground-truth, B) GLNMF, C) PISINMF, D) NMF-QMV, E) gtvMBO, F) SeCoDe, and G) EBEAE-SC.

achieved shorter execution times (see Table 4.2). Figure 4.7 illustrates the abundance maps obtained from the state-of-the-art BLU methods for an intermediate noise realization ($SNR = 12.5$ dB, $PSNR = 12.5$ dB). Qualitatively, it is evident that EBEAE-SC closely resembles the ground-truth, with differences primarily observed in regions with extreme abundances. Specifically, areas with low abundance in the ground-truth appear slightly brighter in the EBEAE-SC estimates, indicating a slight increase in concentration. Conversely, areas with high concentration in the ground-truth exhibited lower concentration in EBEAE-SC estimates. Notably, the abundance maps generated by the state-of-the-art BLU methods showed more pronounced effects and significant granularity.

In certain scenarios, EBEAE-SC may have a potential drawback related to oversmoothing when dealing with low noise in synthetic and remote sensing HS images [162]. This phenomenon arises from the utilization of TV theory in computing internal abundances (see Section 4.5.4), which can result in smoothing and uniformization effects [162, 172]. Consequently, if the TV component is applied without modifying the hy-

Table 4.4: Mean estimation errors for abundances and end-members in the Samson HS image during the Monte Carlo test for various SNR and $PSNR$ values. The lowest value for each pair of SNR and $PSNR$ is indicated in bold font, while results without statistically significant difference from EBEAE-SC (determined by ANOVA with $p > 0.05$) are highlighted in blue.

	Method											
Noise level	GLNMF		PISINMF		NMF-QMV		gtvMBO		SeCoDe		EBEAE-SC	
$(SNR, PSNR)$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$	E_a	E_p $\times 10^{-4}$
20	2.604	4.029	2.582	4.561	3.157	9.715	3.043	10.509	2.583	3.868	**2.363**	**3.463**
17.5	3.076	5.152	3.058	5.732	3.544	13.096	3.325	12.855	3.056	5.004	**2.761**	**4.496**
15	3.342	6.905	3.324	7.151	3.823	16.024	3.602	15.033	3.325	6.699	**3.056**	**6.392**
12.5	3.547	10.451	3.535	10.287	4.086	19.867	3.863	18.559	3.532	10.251	**3.194**	**9.493**
10	3.947	13.524	3.933	13.782	4.498	25.466	4.237	22.090	3.934	13.230	**3.617**	**12.444**
7.5	4.250	18.779	4.236	18.036	4.552	30.781	4.426	28.179	4.241	18.426	**3.833**	**17.748**
5	4.712	24.493	4.703	**23.440**	5.040	39.230	4.901	34.817	4.706	24.285	**4.201**	23.460

perparameters (λ, τ, ν) in low noise conditions, the discrepancies between the EBEAE-SC estimates and the ground-truths may increase, particularly when analyzing images with significant intensity variations. Nevertheless, it is important to note that EBEAE-SC still demonstrates competitive performance across various noise levels by appropriately selecting hyperparameters in low noise scenarios. Moreover, in cases with higher noise, EBEAE-SC consistently produces estimates with lower errors compared to state-of-the-art BLU methods, while maintaining relatively low computational time (refer to Tables 4.1 and 4.2).

4.5.6.2 VNIR Biomedical Application

For the final evaluation of EBEAE-SC, a VNIR image of a patient diagnosed with a grade IV glioblastoma brain tumor was analyzed. This particular image, denoted as HS P012-01, is part of the in-vivo human brain database discussed in Subsection 3.3.2.2 of this dissertation work. The tumor area in the HS image was delineated with a yellow line, as depicted in Figure 4.8A, which illustrates the corresponding synthetic RGB image.

In the analysis of the VNIR image, the initial step involved defining characteristic end-members by extracting the labeled pixels from the gold standard map for the four classes under study. Using the mean spectral signatures $\{\mathbf{p}_{NT}, \mathbf{p}_{TT}, \mathbf{p}_{BV}, \mathbf{p}_{BG}\}$, an end-members matrix $\mathbf{P}$ was constructed, as these labeled pixels represent pure spectral signatures of each class. To generate abundance maps, only the algorithms described by Equations (4.21) and (4.23) were used, with hyperparameter values of $\lambda = 6 \times 10^{-5}$, $\tau = 0.1$, and $\nu = 1 \times 10^{-4}$.

The resulting abundances and end-members obtained by EBEAE-SC for this application are shown in Figures 4.8B) and 4.8C), respectively. Furthermore, a classification map indicating the labels NT, TT, BV, or BG was generated based on the maximum abundance per pixel, as shown in Figure 4.8E). These results are consistent with the gold standard depicted in Figure 4.8D) and the manual segmentation performed by

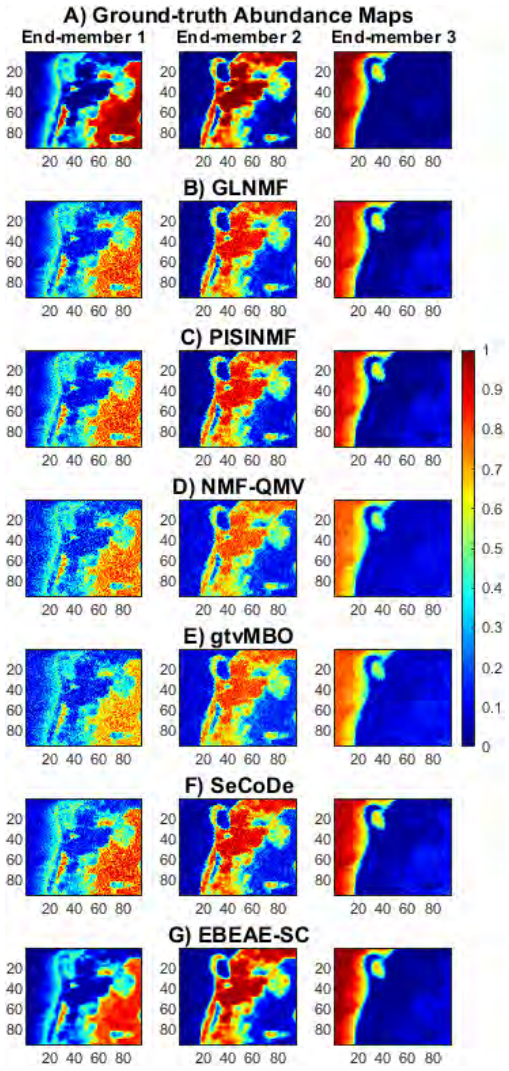


Figure 4.7: One realization of the Monte Carlo estimation results (abundance maps) for Samson HS image ($N = 3$, $SNR = 12.5$ dB and $PSNR = 12.5$ dB): A) Ground-truth, B) GLNMF, C) PISINMF, D) NMF-QMV, E) gtvMBO, F) SeCoDe, and G) and E) EBEAE-SC.

the clinical expert shown in Figure 4.8A). Additionally, the accuracy of the estimates compared to the pixels labeled in the gold standard was 97.4%, which confirms their high similarity with the findings reported in [5, 51]. One advantage of using EBEAE-SC over existing methodologies in the literature is that it eliminates the need for a smoothing stage after abundance estimation [5, 51]. Moreover, EBEAE-SC opens up possibilities for reducing processing stages without sacrificing the ability to identify characteristic elements in such applications.

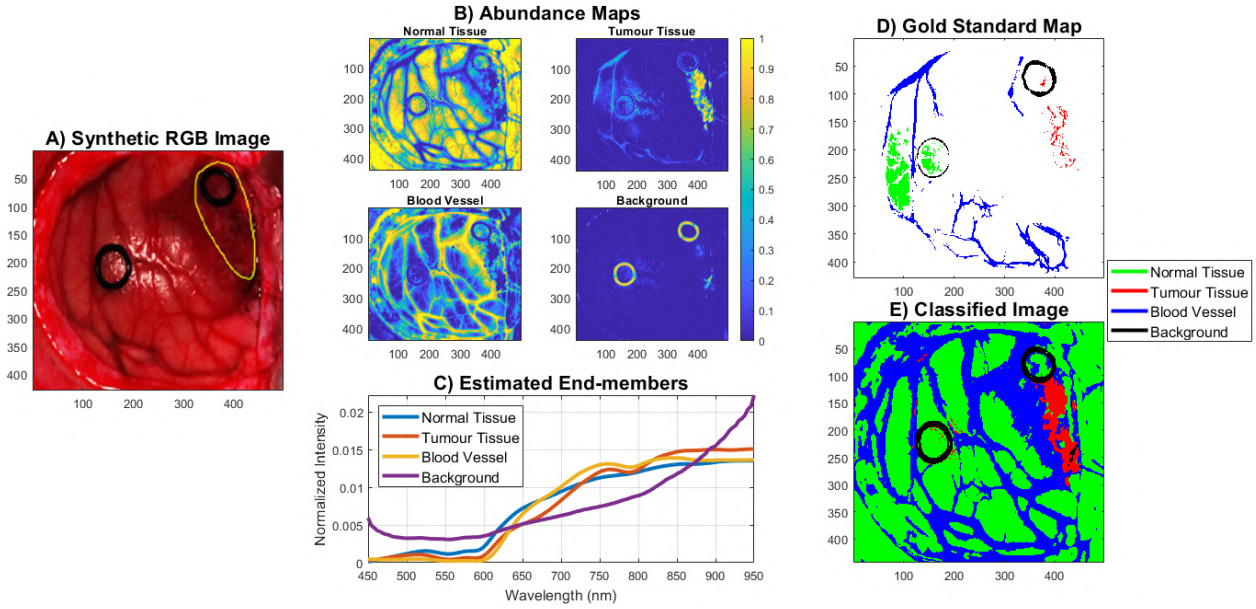


Figure 4.8: Results of the VNIR application using EBEAE-SC: A) A synthetic RGB image highlighting the tumor area with a yellow line, B) Abundance maps representing four classes: Normal Tissue (NT), Tumor Tissue (TT), Blood Vessel (BV), and Background (BG), with values ranging from 0 to 1, C) Estimated end-members for each class, D) Gold standard map, and E) Classified image. The NT, TT, BV, and BG classes are visually represented by green, red, blue, and black colors, respectively. White pixels in the images indicate non-labeled data.

4.5.6.3 Qualitative Analysis

The results in Figures 4.4, 4.6, 4.7, and 4.8 for synthetic and remote sensing HS images demonstrate that EBEAE-SC effectively reduces the granularity present in the concentration maps derived from the internal abundances matrix $\mathbf{W}$, while preserving a significant level of fine detail that is not discernible in state-of-the-art BLU methods. Moreover, EBEAE-SC maintains sharp edges, a characteristic often compromised by denoising algorithms employing TV components in their formulation. This improved uniformity and detail preservation in the internal abundance maps result in estimations that closely align with the ground-truth, outperforming the state-of-the-art BLU methods. Additionally, the inclusion of the TV component slightly enhances the end-member estimation, which proves advantageous in BLU tasks. Consequently, leveraging prior knowledge of the end-members and employing EBEAE-SC for abundance calculation yields remarkably similar results to the ground-truth dataset, effectively eliminating the majority of granularity. Furthermore, the Bregman split technique prevents excessive smoothing by controlling the relationship between the hyperparameters μ and τ.

Nonetheless, the primary hurdle faced by EBEAE-SC lies in the calibration of its hyperparameters to achieve well-defined transition regions and prevent oversmoothing, as discussed earlier in the text. The

hyperparameters of EBEAE-SC, as mentioned in a previous section, affect various aspects of the unmixing process. For instance, hyperparameter ρ modifies the similarity between end-members, where a value close to 1 produces similar morphology among end-members, while a value close to 0 allows for different shapes. The impact of this parameter is evident in the implementation results presented in preceding sections. On the other hand, hyperparameters λ, ν, and τ control the smoothness and uniformity of the internal abundance maps. Hence, hyperparameter λ regulates the similarity between the matrices $\mathbf{A}$ and $\mathbf{W}$; when λ approaches 1, both matrices become highly similar, resulting in minimal impact from the TV smoothing. Consequently, small values of λ were employed in the experiments. Lastly, the hyperparameters ν and τ regulate the variability and smoothness of the internal abundance maps. These hyperparameters mitigate granularity by assigning weight to the finite difference operators (ν) and controlling the effect of the Bregman variables (τ).

4.5.6.4 Convergence Analysis

In order to demonstrate the convergence characteristics of EBEAE-SC, a detailed analysis of the estimation errors was performed at each iteration using both synthetic (VNIR) and remote sensing (Samson HS image) data sets. These evaluations were carried out using identical hyperparameters, as discussed in Subsections 4.5.6 and 4.5.6.1, with a fixed SNR and $PSNR$ noise levels of 5 dB, and a maximum iteration limit of 1000. The convergence plots, illustrated in Figure 4.9, clearly indicate a consistent pattern of decreasing estimation errors with each iteration. Eventually, the errors stabilize and converge to a fixed value, as the number of iterations increases.

Furthermore, upon examination of the energy functional of EBEAE-SC presented in Equation (4.14), it can be observed that most of its components exhibit quadratic characteristics, rendering them convex and differentiable functions. However, when addressing the subproblem associated with estimating the internal abundances $\mathbf{W}$ in (4.22), the cost function involves the utilization of both L_2 and L_1 norms. The latter is effectively resolved through the Bregman split technique, which has been extensively studied in the literature and proven to converge without imposing constraints or exact updates, as demonstrated in [157]. The aforementioned discussion, combined with our experimental evaluations, provides evidence of the convergence properties exhibited by EBEAE-SC, even in scenarios characterized by high levels of noise.

4.6 Nonlinear Extended Blind End-member and Abundance Extraction for Hyperspectral Images

In the field of nonlinear optical interaction analysis, the MMM provides a comprehensive framework for unmixing. However, the research on unsupervised unmixing strategies based on this model is limited, and the work by Wei et al. [176] is one of the few contributions in this area. Furthermore, the develop-

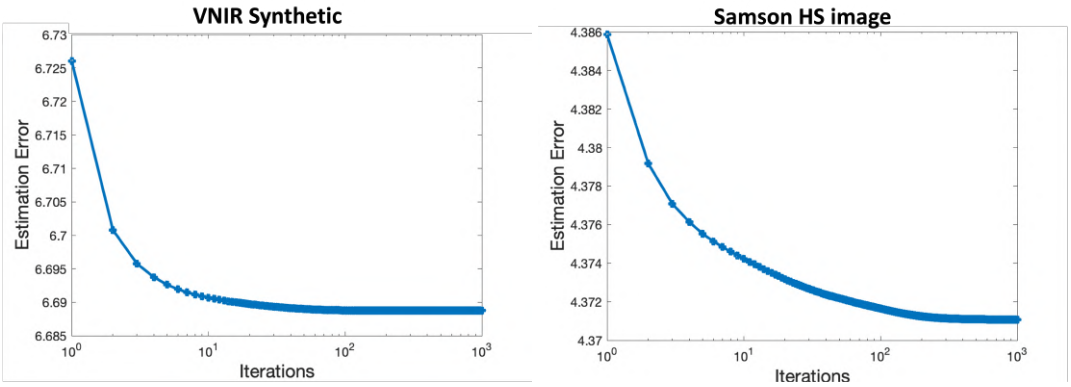


Figure 4.9: Convergence analysis of EBEAE-SC, which was performed on two datasets: VNIR synthetic (shown in the left plot) and Samson HS image (shown in the right plot) datasets. Both datasets were evaluated at a noise level of 5 dB for both SNR and PSNR.

ment of new blind unmixing methods based on the MMM requires considerations of precision, robustness, and computational efficiency as essential performance requirements. To address these concerns, a novel methodology unmixing proposal is suggested based on the optimization structure of EBEAE. This novel methodology is known as NEBEAE.

An advantage of the original EBEAE formulation was the normalization of measurements, which restricted the search space and reduced estimation variability. In NEBEAE, a similar normalization condition is suggested to constrain the abundances. The synthesis optimization objective function of NEBEAE comprises three components: estimation error, abundance entropy, and end-member similarity. Due to the nonlinear relationship between decision variables, a cyclic CDA (CCDA) is adopted to derive an iterative scheme [173].

4.6.1 Problem Formulation

Similarly to the case of EBEAE-SC, NEBEAE draws inspiration from both EBEAE and LMM to formulate its approach. In this case, the normalization conditions in the formulation are described next:

$$\langle \mathbf{1}_L, \mathbf{z}_k \rangle = \mathbf{1}_L^\top \mathbf{z}_k = 1, \quad \forall k \in \mathcal{K} \triangleq \{1, ..., K\}, \tag{4.34}$$

$$\langle \mathbf{1}_L, \mathbf{p}_n \rangle = \mathbf{1}_L^\top \mathbf{p}_n = 1 \quad \forall n \in \mathcal{N} \triangleq \{1, ..., N\}, \tag{4.35}$$

$$\langle \mathbf{1}_N, \boldsymbol{\alpha}_k \rangle = \sum_{n=1}^{N} \alpha_{k,n} = 1, \quad \forall k \in \mathcal{K}, \tag{4.36}$$

with $\langle \mathbf{x}, \mathbf{y} \rangle$ representing the inner product $\mathbf{x}^\top \mathbf{y}$, $\mathbf{1}_L$ denoting an L-dimensional vector of ones, and $\alpha_{k,n}$ representing the normalized abundances $\alpha_{k,n} \triangleq \beta_{k,n} / \mathbf{1}_L^\top \mathbf{z}_k$. With these restrictions, it is possible to achieve an equivalence of the LMM in Equation (4.3) based on nonlinear variables.

Now, to establish the MMM in Equation (4.9), it is necessary to restructure the restriction in the normalized abundances $\{\boldsymbol{\alpha}_k\}_{k\in\mathcal{K}}$ in Equation (4.36) to account for nonlinear interactions, based on the assumption of Equations (4.35) and (4.34), as follows:

$$(1-d_k)\sum_{n=1}^{N}\alpha_{k,n}+d_k\sum_{n=1}^{N}\alpha_{k,n}\langle\mathbf{z}_k,\mathbf{p}_n\rangle=1 \tag{4.37}$$

$$\underbrace{\Big[(1-d_k)\mathbf{1}+d_k\mathbf{P}^\top\mathbf{z}_k\Big]^\top}_{\boldsymbol{\delta}_k^\top}\boldsymbol{\alpha}_k=\boldsymbol{\delta}_k^\top\boldsymbol{\alpha}_k=1. \tag{4.38}$$

In a vector format, the MMM in Equation (4.9) can be written as:

$$\mathbf{y}_k=(1-d_k)\mathbf{P}\boldsymbol{\alpha}_k+d_k\left(\mathbf{P}\boldsymbol{\alpha}_k\right)\odot\mathbf{z}_k+\mathbf{w}_k\quad\forall k\in\mathcal{K}. \tag{4.39}$$

Therefore, the matrix $\mathbf{P}$ representing the end-members is shared among all measurements, while the normalized abundances $\{\boldsymbol{\alpha}_k\}$ and the levels of nonlinear interaction $\{d_k\}$ are unique to each observation.

The problem formulation of NEBEAE differs from the approach presented in [84] by considering a MMM for each individual measurement. The NEBEAE synthesis problem can be represented as follows:

$$\begin{aligned}\min_{\{\boldsymbol{\alpha}_k\},\mathbf{P},\{d_k\}}\ &\underbrace{\frac{1}{2K}\sum_{k=1}^{K}\frac{\|\mathbf{y}_k-(1-d_k)\mathbf{P}\boldsymbol{\alpha}_k-d_k\left(\mathbf{P}\boldsymbol{\alpha}_k\right)\odot\mathbf{z}_k\|^2}{\|\mathbf{y}_k\|^2}}_{(a)}\\ &\underbrace{-\frac{\mu}{2K}\sum_{k=1}^{K}\|\boldsymbol{\alpha}_k\|^2}_{(b)}+\underbrace{\frac{\rho}{2\vartheta}\sum_{n=1}^{N-1}\sum_{j=n+1}^{N}\|\mathbf{p}_n-\mathbf{p}_j\|^2}_{(c)},\end{aligned} \tag{4.40}$$

such that:

$$\boldsymbol{\alpha}_k\geq 0\quad\&\quad\boldsymbol{\delta}_k^\top\boldsymbol{\alpha}_k=1\quad\forall k\in\mathcal{K}, \tag{4.41}$$

$$d_k\in(-\infty,1], \tag{4.42}$$

$$\mathbf{p}_n\geq 0\quad\&\quad\mathbf{1}^\top\mathbf{p}_n=1\quad\forall n\in\mathcal{N}, \tag{4.43}$$

where $\mu\geq 0$ and $\rho\geq 0$ are hyperparameters in NEBEAE, and the normalization variable ϑ takes the value described in Equation (4.13).

In Equation (4.40), the component denoted by (a) represents the fidelity error as a percentage in the estimation made by the MMM. This error accumulates across the entire dataset. On the other hand, the component denoted by (b) acts as a regularization term, taking into account the Euclidean norm of the abundance vectors with a normalization weight. Finally, the component denoted by (c) quantifies the Euclidean distance between pairs of end-members, serving as a similarity metric among them. Since the objective function in Equation (4.40) involves joint minimization, the fidelity error in (a) and the similarity among

resulting end-members in (c) are reduced. However, due to the negative sign in (b), the abundance vectors tend to increase their values. Nevertheless, as the abundances are positive and constrained by the equality condition in Equation (4.41), and the Euclidean norm is squared in each term, the overall effect is to promote larger abundances towards their upper limit and diminish smaller abundances towards zero. Consequently, this pattern can be interpreted as an increase in the entropy of the abundances. The hyperparameter μ is thus associated with the abundance entropy, while ρ relates to the similarity among end-members.

Furthermore, the synthesis scheme described in Equation (4.40) poses an ill-posed inverse problem, where the end-members are computed across the entire dataset, and the abundances and nonlinear interaction levels vary spatially. Therefore, the regularization term in (b) plays a crucial role in stabilizing this optimization process and enhancing the robustness of the estimation scheme. In summary, the overall synthesis framework presented in Equation (4.40), inspired by [84], aims to minimize estimation errors while controlling the entropy of normalized abundances and promoting similarity among resulting end-members.

To address the optimization problem presented in Equation (4.40), three sub-problems are formulated for estimation using a CCDA [173]: (i) normalized abundances $\{\boldsymbol{\alpha}_k\}_{k\in\mathcal{K}}$, (ii) nonlinear interaction levels $\{d_k\}_{k\in\mathcal{K}}$, and (iii) end-members matrix $\mathbf{P}$. To solve these sub-problems, a CQO approach is adopted within each stage of the CCDA. This approach involves fixing the variables of two sub-problems while optimizing the remaining one [84].

4.6.2 Abundance Estimation

In this initial sub-problem, the nonlinear interaction levels $d_{k\,k\in\mathcal{K}}$ and the end-member matrix $\mathbf{P}$ are kept fixed during the optimization process. To achieve this goal, the cost function in Equation (4.40) is rewritten and each individual k-th measurement is analyzed individually from the others as:

$$\min_{\boldsymbol{\alpha}_k} \frac{1}{2}\frac{\|\mathbf{y}_k - \boldsymbol{\Lambda}_k\boldsymbol{\alpha}_k\|^2}{\|\mathbf{y}_k\|^2} - \frac{\mu}{2}\|\boldsymbol{\alpha}_k\|^2, \tag{4.44}$$

such that the condition in Equation (4.41) is satisfied, where:

$$\boldsymbol{\Lambda}_k \triangleq \mathbf{P} \odot \left[(1-d_k)\mathbf{1}_{L\times N} + d_k\mathbf{z}_k\mathbf{1}_N^\top\right] \in \mathbb{R}^{L\times N}. \tag{4.45}$$

The cost function is augmented by incorporating a Lagrange multiplier $\sigma_k > 0$ to account for the equality condition in Equation (4.41). The resulting expression is as follows:

$$J_1(\boldsymbol{\alpha}_k, \sigma_k) \triangleq \frac{\|\mathbf{y}_k\|^2 - 2\boldsymbol{\alpha}_k^\top\boldsymbol{\Lambda}_k^\top\mathbf{y}_k + \boldsymbol{\alpha}_k^\top\boldsymbol{\Lambda}_k^\top\boldsymbol{\Lambda}_k\boldsymbol{\alpha}_k}{2\|\mathbf{y}_k\|^2} - \frac{\mu}{2}\boldsymbol{\alpha}_k^\top\boldsymbol{\alpha}_k + \sigma_k\left(\boldsymbol{\delta}_k^\top\boldsymbol{\alpha}_k - 1\right). \tag{4.46}$$

The optimization problem stated in Equation (4.44) can be solved by deriving a closed-form solution based on the stationary conditions:

$$\frac{\partial J_1}{\partial \boldsymbol{\alpha}_k} = 0 \quad \& \quad \frac{\partial J_1}{\partial \sigma_k} = 0 \tag{4.47}$$

and redefining the weight hyperparameter $\mu = \frac{\lambda_{min}(\mathbf{\Lambda}_k^\top \mathbf{\Lambda}_k)}{\|\mathbf{y}_k\|^2}\tilde{\mu}$ with $\tilde{\mu} \in [0, 1)$, such that:

$$\boldsymbol{\alpha}_k = \mathbf{\Theta}_k^{-1}\left[\mathbf{\Lambda}_k \mathbf{y}_k - \frac{\boldsymbol{\delta}_k^\top \mathbf{\Theta}_k^{-1}\mathbf{\Lambda}_k^\top \mathbf{y}_k - 1}{\boldsymbol{\delta}_k^\top \mathbf{\Theta}_k^{-1}\boldsymbol{\delta}_k}\boldsymbol{\delta}_k\right], \tag{4.48}$$

where $\mathbf{\Theta}_k \triangleq \mathbf{\Lambda}_k^\top \mathbf{\Lambda}_k - \tilde{\mu}\,\lambda_{min}(\mathbf{\Lambda}_k^\top \mathbf{\Lambda}_k)\,\mathbf{I} \in \mathbb{R}^{N\times N}$. The calculation of Equation (4.48) involves inverting matrix $\mathbf{\Theta}_k$, which is an $N \times N$ matrix determined by the order of the MMM in Equation (4.9). However, the order N is significantly smaller compared to the size of each measurement L or the cardinality of the dataset $K = \text{card}(\mathcal{Z})$. Therefore, the computational complexity of this step is not expected to be restrictive. Additionally, the construction of matrix $\mathbf{\Theta}_k$ based on Equation (4.45) primarily depends on the linear independence of the end-members in $\mathbf{P}$. It is worth noting that this condition is continuously monitored through the computation of $\lambda_{min}(\mathbf{\Lambda}_k^\top \mathbf{\Lambda}_k)$. In the event of encountering a singularity condition, the CCDA process is halted as the resulting end-members in $\mathbf{P}$ would lack a physical interpretation.

4.6.3 Estimation of Nonlinear Interaction Level

In the case of the second sub-problem, the absence of interaction between the measurements in the cost function allows for the independent estimation of each nonlinear interaction level. Therefore, the optimization problem for this estimation process can be expressed as follows:

$$\min_{d_k \in (-\infty, 1]} \frac{1}{2}\|\mathbf{y}_k - (1-d_k)\mathbf{P}\boldsymbol{\alpha}_k - d_k\,(\mathbf{P}\boldsymbol{\alpha}_k) \odot \mathbf{z}_k\|^2. \tag{4.49}$$

Following several mathematical derivations, the cost function can be formulated as:

$$J_2(d_k) \triangleq \frac{1}{2}\|\mathbf{y}_k - \mathbf{P}\boldsymbol{\alpha}_k\|^2 + d_k\,(\mathbf{y}_k - \mathbf{P}\boldsymbol{\alpha}_k)^\top\,(\mathbf{P}\boldsymbol{\alpha}_k - [\mathbf{P}\boldsymbol{\alpha}_k] \odot \mathbf{z}_k) + \frac{d_k^2}{2}\|\mathbf{P}\boldsymbol{\alpha}_k - [\mathbf{P}\boldsymbol{\alpha}_k] \odot \mathbf{z}_k\|^2. \tag{4.50}$$

Next, by taking into account the optimality condition $\partial J_2/\partial d_k = 0$ and considering the feasible interval, the optimal solution is written as follows:

$$d_k = \Pi\left(-\frac{(\mathbf{y}_k - \mathbf{P}\boldsymbol{\alpha}_k)^\top\,(\mathbf{P}\boldsymbol{\alpha}_k - [\mathbf{P}\boldsymbol{\alpha}_k] \odot \mathbf{z}_k)}{\|\mathbf{P}\boldsymbol{\alpha}_k - [\mathbf{P}\boldsymbol{\alpha}_k] \odot \mathbf{z}_k\|^2}\right), \tag{4.51}$$

where $\Pi(\cdot)$ is a projection function:

$$\Pi(x) = \begin{cases} x & x \in (-\infty, 1], \\ 1 & x > 1. \end{cases} \tag{4.52}$$

4.6.4 End-member Estimation

In this final sub-problem, the normalized abundances $\{\boldsymbol{\alpha}_k\}_{k\in\mathcal{K}}$ and nonlinear interaction levels $\{d_k\}_{k\in\mathcal{K}}$ remain constant throughout the estimation process. Departing from Equation (4.40), the resultant optimization problem becomes:

$$\min_{\mathbf{P}} \frac{1}{2K}\sum_{k=1}^{K}\frac{\|\mathbf{y}_k - (1-d_k)\mathbf{P}\boldsymbol{\alpha}_k - d_k\,(\mathbf{P}\boldsymbol{\alpha}_k) \odot \mathbf{z}_k\|^2}{\|\mathbf{y}_k\|^2} + \frac{\rho}{2\vartheta}\sum_{n=1}^{N-1}\sum_{j=n+1}^{N}\|\mathbf{p}_n - \mathbf{p}_j\|^2. \tag{4.53}$$

The updated objective function for this estimation procedure can be expressed as follows:

$$J_3(\mathbf{P}) \triangleq \frac{1}{2K}\sum_{k=1}^{K}\frac{\|\mathbf{y}_k\|^2 - 2\mathbf{y}_k^\top\mathbf{M}_k\mathbf{P}\boldsymbol{\alpha}_k + \boldsymbol{\alpha}_k^\top\mathbf{P}^\top\mathbf{M}_k^\top\mathbf{M}_k\mathbf{P}\boldsymbol{\alpha}_k}{\|\mathbf{y}_k\|^2} + \frac{\rho}{2\vartheta}\text{Tr}(\mathbf{POP}^\top), \tag{4.54}$$

where

$$\mathbf{M}_k \triangleq (1-d_k)\mathbf{I} + d_k\text{diag}(\mathbf{z}_k) \in \mathbb{R}^{L\times L} \tag{4.55}$$

$$\mathbf{O} \triangleq N\mathbf{I} - \mathbf{1}_N\mathbf{1}_N^\top \in \mathbb{R}^{N\times N}. \tag{4.56}$$

The derivative of J_3 with respect to the end-member matrix $\mathbf{P}$ can be calculated straightforwardly by:

$$\frac{\partial J_3}{\partial \mathbf{P}} = \frac{1}{K}\sum_{k=1}^{K}\left\{-\frac{\mathbf{M}_k^\top\mathbf{y}_k\boldsymbol{\alpha}_k^\top}{\|\mathbf{y}_k\|^2} + \frac{\mathbf{M}_k^\top\mathbf{M}_k\mathbf{P}\boldsymbol{\alpha}_k\boldsymbol{\alpha}_k^\top}{\|\mathbf{y}_k\|^2}\right\} + \frac{\rho}{2\vartheta}\mathbf{PO} \in \mathbb{R}^{L\times N}. \tag{4.57}$$

Due to the nature of the equation mentioned earlier, finding a closed-form solution for the optimal end-members matrix $\mathbf{P}$ is not feasible. Therefore, an alternative approach is proposed, which involves a gradient descent scheme with an optimal linear search to define the step size [173]:

$$\mathbf{P}^{l+1} = \mathbf{P}^l - \gamma_l\boldsymbol{\Gamma}^l \qquad l \geq 0, \tag{4.58}$$

where

$$\gamma_l = \arg\min_{\gamma\geq 0} J_3\left(\mathbf{P}^l - \gamma\boldsymbol{\Gamma}^l\right) \tag{4.59}$$

and

$$\boldsymbol{\Gamma}^l \triangleq \left.\frac{\partial J_3}{\partial \mathbf{P}}\right|_{\mathbf{P}=\mathbf{P}^l}. \tag{4.60}$$

After performing a straightforward derivation and considering the quadratic relationship on γ in Equation (4.6.4), the optimal step in Equation (4.58) is determined by setting:

$$\frac{\partial J_3(\mathbf{P}^l - \gamma\boldsymbol{\Gamma}^l)}{\partial\gamma} = 0 \tag{4.61}$$

which produces:

$$\gamma^l = \max\left(0, \frac{\gamma_{num}^l}{\gamma_{den}^l}\right) \tag{4.62}$$

such that:

$$\gamma_{num}^l = \frac{1}{K}\sum_{k=1}^{K}\frac{\boldsymbol{\alpha}_k^\top(\boldsymbol{\Gamma}^l)^\top\mathbf{M}_k^\top\mathbf{M}_k(\mathbf{P}^l\boldsymbol{\alpha}_k - \mathbf{y}_k)}{\|\mathbf{y}_k\|^2} + \frac{\rho}{2\vartheta}\text{Tr}(\boldsymbol{\Gamma}^l\mathbf{O}[\mathbf{P}^l]^\top + \mathbf{P}^l\mathbf{O}[\boldsymbol{\Gamma}^l]^\top)$$

$$\gamma_{den}^l = \frac{1}{K}\sum_{k=1}^{K}\frac{\boldsymbol{\alpha}_k^\top(\boldsymbol{\Gamma}^l)^\top\mathbf{M}_k^\top\mathbf{M}_k\boldsymbol{\Gamma}^l\boldsymbol{\alpha}_k}{\|\mathbf{y}_k\|^2} + \frac{\rho}{\vartheta}\text{Tr}(\boldsymbol{\Gamma}^l\mathbf{O}[\boldsymbol{\Gamma}^l]^\top). \tag{4.63}$$

4.6.5 Implementation

The implementation of NEBEAE using the formulation in Equation (4.40) requires an initial step of end-members matrix $\mathbf{P^0}$ initialization during the CCDA. To achieve this, four widely used schemes from the literature are proposed: N-FINDR, VCA, SVMAX, and SISAL [136, 174, 175, 176]. However, the main computational complexity in NEBEAE arises from estimating the end-members matrix $\mathbf{P}$ shared across the measurements dataset $\mathcal{Z}$. To accelerate this process, similar to [84], we perform this estimation on reduced datasets. To achieve this, we randomly select a subset of indices $\mathcal{I} \subset \mathcal{K} = \{1, \ldots, K\}$ such that the cardinality of $\mathcal{I}$ is less than K, and form the reduced datasets:

$$\begin{aligned} \hat{\mathcal{Z}} &= \{\mathbf{z}_k | k \in \mathcal{I}\} \subset \mathcal{Z}, \\ \hat{\mathcal{Y}} &= \{\mathbf{y}_k | k \in \mathcal{I}\} \subset \mathcal{Y}. \end{aligned} \tag{4.64}$$

The assessment of convergence in the CCDA for the initial iterative scheme will be determined by examining the percentage variation in the estimation error at the l-th iteration:

$$J_l = \sum_{k \in \mathcal{I}} \|\mathbf{y}_k - (1 - d_k^l)\mathbf{P}^l \boldsymbol{\alpha}_k^l - d_k^l \left(\mathbf{P}^l \boldsymbol{\alpha}_k^l\right) \odot \mathbf{z}_k\|^2 \quad l \geq 0. \tag{4.65}$$

Additionally, a maximum number of iterations $l_{max} > 0$ is considered as a termination criterion, ensuring the following convergence property is achieved:

$$\frac{|J_l - J_{l-1}|}{J_{l-1}} < \epsilon \quad \text{OR} \quad l > l_{max}, \tag{4.66}$$

where $\epsilon > 0$ is the convergence threshold. After obtaining the estimated end-members matrix $\mathbf{P}$, the next step is to compute the normalized abundances $\{\boldsymbol{\alpha}_k\}$ and nonlinear interaction levels $\{d_k\}$ for the remaining indexes in $\mathcal{K} \backslash \mathcal{I}$. To achieve this, a new CCDA is employed, incorporating the nonlinear interaction between $\{\boldsymbol{\alpha}_k\}$ and $\{d_k\}$ within the framework of the MMM. A similar estimation error to Equation (4.65) is considered:

$$J_h = \sum_{k \in \mathcal{K} \backslash \mathcal{I}} \|\mathbf{y}_k - (1 - d_k^h)\mathbf{P} \boldsymbol{\alpha}_k^h - d_k^h \left(\mathbf{P} \boldsymbol{\alpha}_k^h\right) \odot \mathbf{z}_k\|^2. \tag{4.67}$$

Upon reaching convergence of the second iteration scheme, the reconstructed measurements $\{\hat{\mathbf{z}}_k\}_{k \in \mathcal{K}}$ can be estimated as follows:

$$\hat{\mathbf{z}}_k = (1 - d_k)\mathbf{P}(\omega_k \, \boldsymbol{\alpha}_k) + d_k \left[\mathbf{P}(\omega_k \, \boldsymbol{\alpha}_k)\right] \odot \mathbf{z}_k \quad \forall k \in \mathcal{K}. \tag{4.68}$$

As mentioned previously, the solution to Equation (4.40) considers an ill-posed inverse problem, so the proposed regularization terms in abundances and end-members look to overcome the ambiguity in the formulation. Furthermore, at each stage of the CCDA, all the terms in Equations (4.44), (4.49) and (4.53) involve the Euclidean norm, so they induce convex optimization problems. So, overall convergence is guaranteed at each stage, but not necessarily to the global optimum, due to the nonlinear dependence of the

Algorithm 2 NEBEAE

Input: The sets of original measurements and its scaled version $\mathcal{Z}$ and $\mathcal{Y}$, the number of end-members N, the initial end-members matrix $\mathbf{P}^0$, the maximum number of iterations l_{max}, the convergence threshold ϵ, and the hyperparameters ρ and $\tilde{\mu}$.

Output: End-member matrix $\mathbf{P}$, normalized abundances $\{\boldsymbol{\alpha}_k\}_{k\in\mathcal{K}}$ and nonlinear interaction level $\{d_k\}_{k\in\mathcal{K}}$.

Normalize measurements $\{\mathbf{z}_k\}_{k\in\mathcal{K}}$ to sum-to-one

$$\mathbf{y}_k = \frac{1}{\mathbf{1}^\top \mathbf{z}_k}\mathbf{z}_k \quad k \in \mathcal{K}.$$

Initialize the end-members matrix by $\mathbf{P}^0$ and set the nonlinear interaction levels to zero, i.e. $d_k = 0$ $\forall k \in \mathcal{K}$.
Select randomly the indexes set $\mathcal{I} \subset \mathcal{K}$, construct the reduced datasets by Equation (4.64), and set $l = 0$.
while Convergence condition in Equation (4.66) is not satisfied. **do**
    Estimate the normalized abundances α_k^l for all $k \in \mathcal{I}$ by Equation (4.48).
    Estimate the nonlinear interaction levels d_k^l for all $k \in \mathcal{I}$ by Equation (4.51).
    Update the gradient descent estimation of end-members matrix $\mathbf{P}^l$ over the reduced datasets $\hat{\mathcal{Y}}$ and $\hat{\mathcal{Z}}$ by Equation (4.58).
    Compute the estimation error in Equation (4.65).
end while
Set $h = 0$.
while Convergence condition in Equation (4.66) with $J_l = J_h$ is not satisfied. **do**
    Estimate the normalized abundances α_k^h for all $k \in \mathcal{K}\backslash\mathcal{I}$ by Equation (4.48).
    Estimate the nonlinear interaction levels d_k^h for all $k \in \mathcal{K}\backslash\mathcal{I}$ by Equation (4.51).
    Compute the estimation error in Equation (4.67).
end while
Compute the estimated measurements $\{\hat{\mathbf{z}}_k\}_{k\in\mathcal{K}}$ by Equation (4.68).
Compute the nonlinear interaction levels $\{d_k\}_{k\in\mathcal{K}}$ with the estimated end-members and abundances, and original measurements $\{\mathbf{z}_k\}_{k\in\mathcal{K}}$ by Equation (4.69).

variables $\{\boldsymbol{\alpha}_k\}, \mathbf{P}, \{d_k\}$ in the error fidelity term in Equation (4.40)-(a). Hence, the optimal solution will be local and depend largely on the selection of the initial end-members matrix $\mathbf{P^0}$.

Lastly, as a result of the normalized requirement specified in Equation (4.34), the calculated nonlinear interaction levels in Equation (4.51) lack the appropriate physical interpretation. Therefore, these parameters need to be computed in the final step after undergoing re-normalization using the following procedure:

$$d_k = \Pi\left(-\frac{(\mathbf{z}_k - \mathbf{P}(\omega_k\boldsymbol{\alpha}_k))^\top (\mathbf{P}(\omega_k\boldsymbol{\alpha}_k) - [\mathbf{P}(\omega_k\boldsymbol{\alpha}_k)] \odot \mathbf{z}_k)}{\|\mathbf{P}(\omega_k\boldsymbol{\alpha}_k) - [\mathbf{P}(\omega_k\boldsymbol{\alpha}_k)] \odot \mathbf{z}_k\|^2}\right) \quad \forall k \in \mathcal{K}. \tag{4.69}$$

Algorithm 2 describes the implementation of NEBEAE and provides a high-level summary of the entire discussed process.

4.6.6 Validation and Discussion

In this section, the validation of the blind nonlinear unmixing under a MMM model is conducted using two types of synthetic HS images, as well as experimental datasets from remote sensing literature and a biomedical imaging application similar to the test in the last Subsection. The performance of NEBEAE is compared to two state-of-the-art unmixing methods based on MMM: (i) Multilinear Mixing Model for Nonlinear Spectral Unmixing (MMMNSU) with end-members initialized by VCA [128], and (ii) Unsupervised Nonlinear Spectral Unmixing Based on MMM (UNSUBMMM) [144]. Additionally, for the experimental datasets, the blind sparse nonlinear HS unmixing (BSNHU) method, which utilizes an ℓ_q-regularizer and a bilinear model [177], is considered for comparison. A comparison is made between the blind nonlinear unmixing and BSNHU, as they address different NMMs than the MMM in NEBEAE, and incorporate both abundance nonnegativity and abundance sum constraints jointly. In order to measure the accuracy of the estimation process, the estimated end-member sets and their corresponding abundances are defined for each method as $\hat{\mathcal{A}}$ and $\hat{\mathcal{P}}$, respectively. The ground-truth sets were denoted as $\mathcal{A}$ and $\mathcal{P}$. The performance metrics using these sets are calculated as follows:

$$E_z = \frac{\|\mathbf{Z} - \hat{\mathbf{Z}}\|_F}{\|\mathbf{Z}\|_F}, \tag{4.70}$$

$$E_p = \min_{\forall \hat{\mathbf{p}} \in \hat{\mathcal{P}}, \mathbf{p} \in \mathcal{P}} \frac{\|\hat{\mathbf{p}} - \mathbf{p}\|}{\|\mathbf{p}\|}, \tag{4.71}$$

$$E_{SAM} = \min_{\forall \hat{\mathbf{p}} \in \hat{\mathcal{P}}, \mathbf{p} \in \mathcal{P}} \cos^{-1}\left(\frac{\hat{\mathbf{p}}^\top \mathbf{p}}{\|\hat{\mathbf{p}}\|\|\mathbf{p}\|}\right), \tag{4.72}$$

$$E_a = \min_{\forall \hat{\mathbf{a}} \in \hat{\mathcal{A}}, \mathbf{a} \in \mathcal{A}} \frac{\|\hat{\mathbf{a}} - \mathbf{a}\|}{\|\mathbf{a}\|}, \tag{4.73}$$

where $\mathbf{Z} = [\mathbf{z}_1 \ \ldots \ \mathbf{z}_K]$ represents the matrix measurements, while $\hat{\mathbf{Z}} = [\hat{\mathbf{z}}_1; \ldots; \hat{\mathbf{z}}_K]$ represents the estimated ones obtained by each blind nonlinear unmixing method. The tests were performed using Matlab© 2018a software on a MacBook Pro equipped with a 2.3 GHz Intel Core i5 dual-core processor and 16 GB of RAM. The scripts used to generate the results presented in the Figures and Tables can be found at the following GitHub repository: *https://github.com/Alex-CruzG/NEBEAE*.

The convergence threshold for NEBEAE was set at $\epsilon = 0.001$, and the maximum number of iterations was limited to 50 to balance convergence evaluation and computational complexity. The similarity weight ρ was chosen to be $\rho = 0.1$, considering a middle point within the range of $[0.01, 1.0]$ as suggested in [84]. For the entropy weight $\tilde{\mu}$, it was restricted to the interval $[0, 1)$, and in our evaluation, we selected $\tilde{\mu} = 0.1$. This choice of $\tilde{\mu}$ ensures a balance between using abundances for classification (when $\tilde{\mu} > 0.5$) and minimizing the fidelity error in Equation (4.40) (when $\tilde{\mu} = 0$). Furthermore, as noted in [84], a small value of $\tilde{\mu}$ improves noise robustness. Lastly, the initialization of the end-members matrix in NEBEAE can be performed using various methods such as VCA, N-FINDR, SVMAX, and SISAL [136, 174, 175, 176].

4.6.6.1 Synthetic Data Evaluation

The synthetic HS images used in this evaluation were generated by the MMM model, described in Section 3.1, corresponding to: (i) Spectral Library Version 7 (SLV7) [101], and (ii) VNIR of in-vivo human brain tissue [84]. It is important to clarify that in all synthetic images, the level of nonlinear interaction $\{d_k\}_{k=1}^{K}$ was set to 0.5, that is, an intermediate value in the range $[0, 1]$, to induce dominance of the nonlinear reflectance term in the MMM. Additionally, to accelerate the estimation of the end-members, a reduced dataset is assumed, reducing the cardinality to card$(\mathcal{I}) = 2700$ measurements.

To evaluate the performance of the proposed method in challenging scenarios, two types of noise commonly encountered in these images were introduced: Gaussian and shot noise. To generate a noisy measurement $\mathbf{y}_k$ from a noiseless observation $\mathbf{y}_k^0$, the following transformation is applied:

$$\mathbf{y}_k = \mathbf{y}_k^0 + \mathbf{n}_k + \mathbf{m}_k \cdot \sqrt{\mathbf{y}_k^0} \qquad \forall k \in \mathcal{K}, \tag{4.74}$$

where the vectors $\mathbf{n}_k \in \mathbb{R}^L$ and $\mathbf{m}_k \in \mathbb{R}^L$ correspond to the Gaussian and shot noise components, respectively. In our approach, the Gaussian noise vector $\mathbf{n}_k$ is characterized by having a zero mean and a STD of σ_k^{SNR}. On the other hand, the shot noise component $\mathbf{m}_k$ is also defined by a Gaussian distribution with a mean of zero, but with a STD of σ_k^{PSNR}. The values of σ_k^{SNR} and σ_k^{PSNR} were determined based on the desired SNR and $PSNR$ specified in the dataset:

$$\sigma_k^{SNR} = \sqrt{\frac{\frac{1}{L-1}\|\bar{\mathbf{y}}^0\|^2}{10^{\frac{SNR}{10}}}}, \qquad \sigma_k^{PSNR} = \sqrt{\frac{\max_{l\in[1,L]}(\bar{\mathbf{y}}^0)_l^2}{10^{\frac{PSNR}{10}}}}, \tag{4.75}$$

where $(\cdot)l$ represents the l-th component of the vector, and $\bar{\mathbf{y}}^0$ denotes the mean vector in the dataset $\mathcal{Y}$. To assess the performance of NEBEAE and the state-of-the-art methods, they were tested under challenging conditions of varying noise levels $(SNR, PSNR) \in \{40, 35, 30, 25, 20\}$ dB. These noise levels were particularly demanding for most blind nonlinear unmixing algorithms. For each combination of SNR and $PSNR$ values, a Monte Carlo evaluation was conducted, generating 50 noise realizations for the synthetic datasets. The performance metrics, including (E_y, E_p, E_{SAM}, E_a) in Equations (4.70)-(4.73), were calculated for each realization.

In the initial evaluation, the four proposals for the initial end-members matrix $\mathbf{P}^0$ in NEBEAE were considered: VCA, N-FINDR, SVMAX, and SISAL. This evaluation was conducted under various noise levels. The results of this quantitative analysis are presented in Tables 4.5 and 4.6. To assess the statistical significance of the numerical results, an ANOVA evaluation was performed. Based on the results in Table 4.5, SISAL demonstrated the best performance across most of the metrics for the SLV7 dataset. Similarly, in Table 4.6, SISAL outperformed the other methods in the VNIR dataset, although N-FINDR and VCA achieved the best results for certain noise conditions and metrics such as E_a, E_p, and E_{SAM}. Interestingly, for N-FINDR and VCA, the values of these metrics did not change significantly for noise conditions above

20 dB. Therefore, SISAL will primarily be adopted for the initialization process in NEBEAE, while VCA and N-FINDR will be considered as alternative options in some datasets.

Table 4.5: Evaluation of estimation performance using the synthetic Monte Carlo approach with different pairs of SNR/PSNR and initial end-members matrixes $\mathbf{P}^0$ for SLV7 synthetic database. The results are presented as the mean value $\pm$ STD, at the same form, the lowest error and computational time are highlighted in bold-face for each SNR/PSNR pair. It is worth noting that all values in the table show statistical differences based on ANOVA analysis with a significance level of $p < 0.05$.

		Initial End-members Matrix $\mathbf{P}^0$			
Metric	(SNR,PSNR)	N-FINDR	VCA	SVMAX	SISAL
Output Estimation Error E_z	20	$0.0990 \pm 9.7882\times10^{-5}$	$0.0992 \pm 7.2496\times10^{-05}$	$0.0992 \pm 5.9183\times10-5$	$\mathbf{0.0988 \pm 5.1793}\times10^{-5}$
	25	$0.0562 \pm 3.7470\times10^{-5}$	$0.0562 \pm 4.1662\times10^{-05}$	$0.0562 \pm 3.2716\times10^{-5}$	$\mathbf{0.0558 \pm 2.7228}\times10^{-5}$
	30	$0.0320 \pm 2.2002\times10^{-5}$	$0.0320 \pm 3.3105\times10^{-05}$	$0.0320 \pm 2.2955\times10^{-5}$	$\mathbf{0.0314 \pm 3.3431}\times10^{-5}$
	35	$0.0186 \pm 2.2638\times10^{-5}$	$0.0186 \pm 4.8501\times10^{-05}$	$0.0186 \pm 2.7026\times10^{-5}$	$\mathbf{0.0178 \pm 6.8842}\times10^{-5}$
	40	$0.0115 \pm 3.7539\times10^{-5}$	$0.0115 \pm 9.4781\times10^{-05}$	$0.0114 \pm 2.7655\times10^{-5}$	$\mathbf{0.0104 \pm 2.4093}\times10^{-5}$
Abundances Estimation Error E_a	20	0.3492 ± 0.0321	0.3442 ± 0.0317	$\mathbf{0.3422 \pm 0.0166}$	0.3597 ± 0.0158
	25	0.3567 ± 0.0108	0.3667 ± 0.0087	0.3662 ± 0.0053	$\mathbf{0.3340 \pm 0.0182}$
	30	0.3861 ± 0.0042	0.3829 ± 0.0053	0.3887 ± 0.0022	$\mathbf{0.3261 \pm 0.0149}$
	35	0.3757 ± 0.0037	0.3749 ± 0.0084	0.3777 ± 0.0034	$\mathbf{0.3281 \pm 0.0129}$
	40	0.3731 ± 0.0030	0.3731 ± 0.0098	0.3733 ± 0.0025	$\mathbf{0.3293 \pm 0.0121}$
End-members Estimation Error E_p	20	0.0444 ± 0.0100	0.0479 ± 0.0099	0.0449 ± 0.0039	$\mathbf{0.0358 \pm 0.0061}$
	25	0.0421 ± 0.0024	0.0464 ± 0.0032	0.0444 ± 0.0012	$\mathbf{0.0370 \pm 0.0053}$
	30	0.0463 ± 0.0006	0.0476 ± 0.0012	0.0460 ± 0.0002	$\mathbf{0.0387 \pm 0.0054}$
	35	0.0472 ± 0.0003	0.0478 ± 0.0010	0.0470 ± 0.0003	$\mathbf{0.0390 \pm 0.0049}$
	40	0.0474 ± 0.0002	0.0477 ± 0.0009	0.0473 ± 0.0002	$\mathbf{0.0390 \pm 0.0015}$
End-members Spectral Angle Mapper E_{SAM}	20	0.0442 ± 0.0098	0.0477 ± 0.0099	0.0448 ± 0.0039	$\mathbf{0.0356 \pm 0.0061}$
	25	0.0419 ± 0.0023	0.0462 ± 0.0031	0.0442 ± 0.0012	$\mathbf{0.0368 \pm 0.0052}$
	30	0.0461 ± 0.0006	0.0474 ± 0.0012	0.0457 ± 0.0002	$\mathbf{0.0385 \pm 0.0054}$
	35	0.0470 ± 0.0003	0.0476 ± 0.0010	0.0468 ± 0.0003	$\mathbf{0.0388 \pm 0.0049}$
	40	0.0472 ± 0.0003	0.0475 ± 0.0009	0.0471 ± 0.0002	$\mathbf{0.0388 \pm 0.0015}$
Computational Time (seconds)	20	21.0742 ± 4.2993	13.7782 ± 5.1759	13.0834 ± 2.3485	$\mathbf{5.9417 \pm 0.3974}$
	25	18.0637 ± 1.6096	20.4937 ± 4.0539	22.2896 ± 3.0965	$\mathbf{6.0775 \pm 0.7183}$
	30	34.2252 ± 2.9930	34.0510 ± 3.6826	38.0171 ± 5.0760	$\mathbf{5.7977 \pm 0.3236}$
	35	44.8495 ± 2.1743	43.6023 ± 5.6222	46.9741 ± 4.7431	$\mathbf{5.0665 \pm 0.0516}$
	40	56.5148 ± 1.6697	55.6090 ± 7.3566	57.2215 ± 1.6614	$\mathbf{5.1774 \pm 0.5516}$

In the second evaluation, the performance of NEBEAE, MMMNSU, and UNSUBMMM under different noise levels is presented in Tables 4.7 and 4.8. BSNHU is not compared since the synthetic datasets were generated by MMM, making the comparison unfair. ANOVA evaluation was conducted to assess the statistical significance of the synthetic results, revealing statistical differences for all values in Table 4.7. NEBEAE consistently outperformed MMMNSU and UNSUBMMM in terms of the lowest estimation errors (E_z, E_a, E_p, E_{SAM}) and computational times, as shown in Table 4.7. Figure 4.10 displays the estimated abundance maps for SLV7 at an SNR and PSNR of 20 dB. Visually, NEBEAE provides a more consistent estimate compared to the ground-truths depicted in Figure 3.2, which aligns with the quantitative analysis in Table 4.7. Lastly, the three blind nonlinear unmixing methods accurately estimated the nonlinear interaction levels, as seen in the last column of Figure 4.10.

Table 4.8 presents the results for synthetic VNIR HS images, where statistical differences are consistently observed. NEBEAE achieved the best performance in terms of output estimation error E_z and computational time compared to all other methods. MMMNSU obtained the lowest abundance estimation errors E_a for some noise conditions, while UNSUBMMM performed the best in terms of end-members

Table 4.6: Evaluation of estimation performance using the synthetic Monte Carlo approach with different pairs of SNR/PSNR and initial end-members matrixes $\mathbf{P}^0$ for VNIR synthetic database. The results are presented as the mean value $\pm$ STD, at the same form, the lowest error and computational time are highlighted in bold-face for each SNR/PSNR pair. It is worth noting that all values in the table show statistical differences based on ANOVA analysis with a significance level of $p < 0.05$.

		Initial End-members Matrix $\mathbf{P}^0$			
Metric	(SNR,PSNR)	N-FINDR	VCA	SVMAX	SISAL
Output Estimation Error E_z	20	0.0983 ± 0.0001	$0.0983 \pm 9.1205\times10^{-5}$	$0.0982 \pm 9.9727\times10^{-5}$	$\mathbf{0.0979 \pm 8.5406}\times10^{-5}$
	25	$0.0556 \pm 7.8924\times10^{-5}$	$0.0556 \pm 7.4253\times10^{-5}$	$0.0555 \pm 6.7405\times10^{-5}$	$\mathbf{0.0552 \pm 6.0191}\times10^{-5}$
	30	$0.0315 \pm 5.4413\times10^{-5}$	$0.0314 \pm 8.5194\times10^{-5}$	$0.0314 \pm 4.6024\times10^{-5}$	$\mathbf{0.0311 \pm 3.3697}\times10^{-5}$
	35	0.0180 ± 0.0001	0.0179 ± 0.0001	$0.0179 \pm 7.5787\times10^{-5}$	$\mathbf{0.0175 \pm 1.8572}\times10^{-5}$
	40	0.0105 ± 0.0001	0.0106 ± 0.0004	$0.0104 \pm 7.0883\times10^{-5}$	$\mathbf{0.0098 \pm 1.2410}\times10^{-5}$
Abundances Estimation Error E_a	20	0.3952 ± 0.0426	0.3892 ± 0.0643	$\mathbf{0.3890 \pm 0.0590}$	0.5226 ± 0.0130
	25	$\mathbf{0.3482 \pm 0.0526}$	0.3705 ± 0.0657	0.3610 ± 0.0473	0.5069 ± 0.0243
	30	$\mathbf{0.3262 \pm 0.0602}$	0.3554 ± 0.0430	0.3493 ± 0.0286	0.4906 ± 0.0167
	35	$\mathbf{0.3349 \pm 0.0377}$	0.3669 ± 0.0290	0.3601 ± 0.0251	0.4002 ± 0.0319
	40	0.3589 ± 0.0217	0.3778 ± 0.0427	0.3711 ± 0.0170	$\mathbf{0.3110 \pm 0.0186}$
End-members Estimation Error E_p	20	0.0589 ± 0.01096	0.0301 ± 0.0122	$\mathbf{0.0282 \pm 0.01017}$	0.0708 ± 0.0135
	25	0.0298 ± 0.00792	$\mathbf{0.0233 \pm 0.0067}$	0.0244 ± 0.0051	0.0414 ± 0.0156
	30	0.0222 ± 0.0089	0.0245 ± 0.010	0.0227 ± 0.0034	0.0255 ± 0.0119
	35	0.0212 ± 0.0021	0.0226 ± 0.0023	0.0228 ± 0.0014	$\mathbf{0.0163 \pm 0.0078}$
	40	0.0227 ± 0.0012	0.0239 ± 0.0033	0.0232 ± 0.0010	$\mathbf{0.0125 \pm 0.0025}$
End-members Spectral Angle Mapper E_{SAM}	20	0.0571 ± 0.0096	0.0293 ± 0.0116	$\mathbf{0.0272 \pm 0.0093}$	0.0675 ± 0.0105
	25	0.0291 ± 0.0078	$\mathbf{0.0226 \pm 0.0064}$	0.0235 ± 0.0049	0.0399 ± 0.0146
	30	0.0215 ± 0.0083	0.0235 ± 0.0094	0.0219 ± 0.0032	0.0243 ± 0.0108
	35	0.0206 ± 0.0019	0.0219 ± 0.0021	0.0220 ± 0.0013	$\mathbf{0.0154 \pm 0.0071}$
	40	0.0221 ± 0.0011	0.0232 ± 0.0030	0.0225 ± 0.0009	$\mathbf{0.0123 \pm 0.0023}$
Computational Time (seconds)	20	1.4050 ± 0.1609	0.7728 ± 0.2059	0.7329 ± 0.1374	$\mathbf{0.4609 \pm 0.0221}$
	25	1.3362 ± 0.1396	1.0004 ± 0.1638	0.9805 ± 0.1701	$\mathbf{0.4550 \pm 0.0194}$
	30	1.6323 ± 0.2236	1.5186 ± 0.5014	1.3581 ± 0.1234	$\mathbf{0.4445 \pm 0.0152}$
	35	2.0461 ± 0.1597	1.8679 ± 0.3594	1.7558 ± 0.1884	$\mathbf{0.4810 \pm 0.0763}$
	40	2.0149 ± 0.1847	2.2304 ± 0.3023	2.0884 ± 0.1306	$\mathbf{0.5929 \pm 0.0916}$

metrics E_p and E_{SAM}. However, overall, NEBEAE demonstrated superior performance across most tested scenarios. This synthetic evaluation emphasizes the enhanced accuracy of NEBEAE in handling challenging noise conditions and end-members with different physical interpretations.

4.6.7 Real Remote Sensing and Experimental Datasets

This section compares NEBEAE with MMMNSU and UNSUBMMM, but also includes BSNHU for other types of NMMs. For BSNHU, there is no single value per pixel for quantifying nonlinear optical interactions, so this parameter cannot be extrapolated for comparison with NEBEAE, MMMNSU, and UNSUBMMM. However, if there is a ground-truth for both variables, you can compare the estimated end members and their abundance. In this section, the different methods of blind nonlinear SU are evaluated on three remote sensing datasets described in Section 3.2: Cuprite, Urban, and Pavia University Scene. Finally, the proposal is evaluated on two VNIR images of the in-vivo human brain, described in Section 3.3.2.2.

Table 4.7: Performance evaluation of estimation in a synthetic Monte Carlo environment, presented as the mean deviation ± STD, applying nonlinear unmixing methodologies on the SLV7 dataset with different pairs of SNR and PSNR. The lowest error and computational time are highlighted in bold in this table, and all the results in the table were statistically significant with p < 0.05 according to ANOVA analysis.

		Method		
Metric	(SNR,PSNR)	NEBEAE	MMMNSU	UNSUBMMM
Output Estimation Error E_z	20	$\mathbf{0.0992 \pm 7.5409}\times 10^{-5}$	0.1018 ± 0.0012	0.0999 ± 0.0003
	25	$\mathbf{0.0562 \pm 4.4945}\times 10^{-5}$	0.0617 ± 0.0024	0.0577 ± 0.0005
	30	$\mathbf{0.0320 \pm 2.7767}\times 10^{-5}$	0.0415 ± 0.0032	0.0345 ± 0.0004
	35	$\mathbf{0.0179 \pm 6.4545}\times 10^{-5}$	0.0314 ± 0.0032	0.0226 ± 0.0013
	40	$\mathbf{0.0104 \pm 6.2130}\times 10^{-5}$	0.0269 ± 0.0055	0.0171 ± 0.0010
Abundances Estimation Error E_a	20	$\mathbf{0.3457 \pm 0.0215}$	0.3649 ± 0.0547	0.3820 ± 0.0552
	25	$\mathbf{0.3665 \pm 0.0083}$	0.3895 ± 0.0456	0.3966 ± 0.0357
	30	$\mathbf{0.3845 \pm 0.0052}$	0.3951 ± 0.0423	0.4065 ± 0.0290
	35	$\mathbf{0.3327 \pm 0.0167}$	0.3964 ± 0.0235	0.4114 ± 0.0416
	40	$\mathbf{0.3323 \pm 0.0101}$	0.4311 ± 0.0559	0.4123 ± 0.0275
End-members Estimation Error E_p	20	$\mathbf{0.0466 \pm 0.0090}$	0.0574 ± 0.0109	0.0519 ± 0.0126
	25	$\mathbf{0.0457 \pm 0.0019}$	0.0602 ± 0.0059	0.0504 ± 0.0058
	30	$\mathbf{0.0469 \pm 0.0011}$	0.0638 ± 0.0085	0.0513 ± 0.0032
	35	$\mathbf{0.0386 \pm 0.0048}$	0.0630 ± 0.0049	0.0522 ± 0.0050
	40	$\mathbf{0.0395 \pm 0.0018}$	0.0653 ± 0.0062	0.0525 ± 0.0028
End-members Spectral Angle Mapper E_{SAM}	20	$\mathbf{0.0464 \pm 0.0089}$	0.0573 ± 0.0107	0.0517 ± 0.0124
	25	$\mathbf{0.0455 \pm 0.0019}$	0.0600 ± 0.0059	0.0503 ± 0.0056
	30	$\mathbf{0.0466 \pm 0.0011}$	0.0636 ± 0.0084	0.0511 ± 0.0032
	35	$\mathbf{0.0384 \pm 0.0047}$	0.0629 ± 0.0049	0.0520 ± 0.0049
	40	$\mathbf{0.0393 \pm 0.0018}$	0.0651 ± 0.0061	0.0524 ± 0.0028
Computational Time (seconds)	20	$\mathbf{9.1165 \pm 1.6853}$	15.2092 ± 1.2149	10.0364 ± 1.4549
	25	$\mathbf{13.5596 \pm 1.9506}$	14.8229 ± 0.9186	16.7374 ± 3.4365
	30	24.4548 ± 3.1559	$\mathbf{14.8204 \pm 0.9733}$	78.2267 ± 54.6767
	35	24.0440 ± 0.0460	$\mathbf{14.1591 \pm 0.7397}$	145.7370 ± 9.3060
	40	24.0922 ± 0.1341	$\mathbf{14.6954 \pm 0.8340}$	147.5184 ± 0.9670

4.6.7.1 Cuprite Dataset

Cuprite is one of the most widely used benchmark data sets for evaluating SU methods. For this dataset, there are ground-truths just for the end-members, but not for the abundance maps. Therefore, the estimation errors for NEBEAE, MMMNSU, UNSUBMMM, and BSNHU will be evaluated by Equations (4.70) and (4.71); and at the same time by the computational time. The hyperparameters in NEBEAE were roughly the same as those used to evaluate synthetic data, only the entropy weight was increased to account for the noise in the experimental data set, so it was defined $\tilde{\mu} = 0.2$. In this scenario, VCA initialization performed better compared to SISAL and N-FINDR.

As illustrated in Table 4.9, NEBEAE shows clear advantages over MMMNSU, UNSUBMMM, and BSNHU. NEBEAE achieved the lowest estimation errors in measurements (E_z), end-members (E_p), and SAM (E_{SAM}). Additionally, NEBEAE exhibited the shortest computational time compared to the other

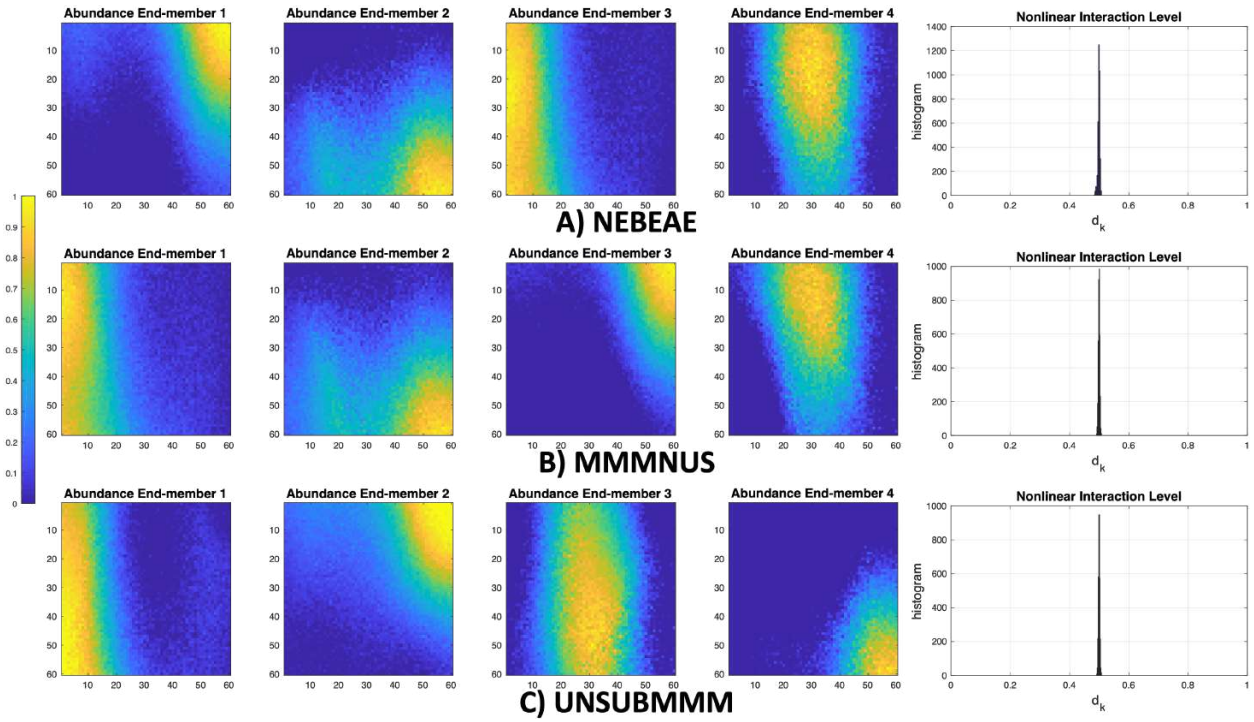


Figure 4.10: Abundance maps estimated for the **SLV7** dataset, along with the resulting nonlinear interaction levels for an SNR and PSNR to 20 dB: A) NEBEAE: 0.4993 ± 0.0033, B) MMMNSU: 0.4997 ± 0.0025, and C) UNSUBMMM: 0.4999 ± 0.0025.

methods. Figure 4.11 shows the nonlinear interaction level maps (d_k) for NEBEAE, MMMNSU, and UNSUBMMM. As mentioned earlier, BSNHU does not provide this parameter. Nonlinear interaction maps reveal negative values, indicating an increase in reflectance compared to LMM as described in [128]. Although the resulting maps of the three unmixing techniques generally exhibit consistency in the separation of the morphological regions, NEBEAE achieves better resolution, as shown in the approximated RGB image in Figure 4.11A.

4.6.7.2 Urban Dataset

In conjunction with the previous evaluation, blind nonlinear unmixing methods were evaluated in the Urban database, described in Section 3.2. Table 4.10 presents the estimation performance of NEBEAE, MMMNSU, UNSUBMMM, and BSNHU for three different conditions regarding the end-members. To improve estimation performance, slight adjustments were made to the hyperparameters in NEBEAE compared to synthetic evaluation, setting $\rho = 1.0$ and $\tilde{\mu} = 0.25$. The VCA initialization yielded the best results for NEBEAE on this dataset compared to SISAL and N-FINDR. NEBEAE demonstrated a significant reduction in computational time compared to MMMNSU, UNSUBMMM, and BSNHU. The output estimation error E_z was the lowest for UNSUBMMM with four end-members, while NEBEAE achieved smaller errors with

Table 4.8: Performance evaluation of estimation in a synthetic Monte Carlo environment, presented as the mean deviation ± STD, applying nonlinear unmixing methodologies on the VNIR dataset with different pairs of SNR and PSNR. The lowest error and computational time are highlighted in bold in this table, and all the results in the table were statistically significant with p < 0.05 according to ANOVA analysis.

		Method		
Metric	(SNR,PSNR)	NEBEAE	MMMNSU	UNSUBMMM
Output Estimation Error E_z	20	**0.0979 ± 1.1126**$\times 10^{-4}$	0.0989 ± 0.0007	0.0986 ± 0.0007
	25	**0.0553 ± 5.7677**$\times 10^{-5}$	0.0563 ± 0.0008	0.0563 ± 0.0011
	30	**0.0311 ± 3.0159**$\times 10^{-5}$	0.0325 ± 0.0004	0.0325 ± 0.0017
	35	**0.0175 ± 2.3667**$\times 10^{-5}$	0.0195 ± 0.0005	0.0191 ± 0.0011
	40	**0.0098 ± 1.5885**$\times 10^{-5}$	0.0152 ± 0.0056	0.0132 ± 0.0036
Abundances Estimation Error E_a	20	0.5228 ± 0.0156	**0.3939 ± 0.0676**	0.4215 ± 0.0720
	25	0.5043 ± 0.0231	**0.3917 ± 0.0640**	0.4534 ± 0.0812
	30	0.4851 ± 0.0202	**0.3670 ± 0.0505**	0.4206 ± 0.0569
	35	0.4075 ± 0.0278	**0.3750 ± 0.0304**	0.4222 ± 0.0593
	40	**0.3169 ± 0.0195**	0.4428 ± 0.0467	0.4399 ± 0.0557
End-members Estimation Error E_p	20	0.0724 ± 0.0132	0.0357 ± 0.0134	**0.0345 ± 0.0130**
	25	0.0369 ± 0.0118	0.0291 ± 0.0110	**0.0253 ± 0.0095**
	30	**0.0212 ± 0.0093**	0.0261 ± 0.0087	0.0234 ± 0.0070
	35	**0.0148 ± 0.0069**	0.0255 ± 0.0020	0.0242 ± 0.0080
	40	**0.0121 ± 0.0028**	0.0259 ± 0.0030	0.0244 ± 0.0050
End-members Spectral Angle Mapper E_{SAM}	20	0.0693 ± 0.0110	0.0348 ± 0.0126	**0.0334 ± 0.0119**
	25	0.0360 ± 0.0105	0.0282 ± 0.0106	**0.0245 ± 0.0091**
	30	**0.0206 ± 0.0088**	0.0253 ± 0.0081	0.0226 ± 0.0065
	35	**0.0140 ± 0.0063**	0.0248 ± 0.0019	0.0233 ± 0.0071
	40	**0.0119 ± 0.0027**	0.0252 ± 0.0027	0.0236 ± 0.0046
Computational Time (seconds)	20	**0.6262 ± 0.0261**	17.0775 ± 0.7811	2.3985 ± 1.0793
	25	**0.6266 ± 0.0401**	17.3067 ± 0.6933	5.2546 ± 8.0038
	30	**0.6156 ± 0.0066**	17.1421 ± 0.8298	8.0615 ± 10.4945
	35	**0.6458 ± 0.0790**	17.2223 ± 0.6115	8.3488 ± 7.1535
	40	**0.7647 ± 0.1279**	18.7007 ± 0.9869	13.675 ± 12.3457

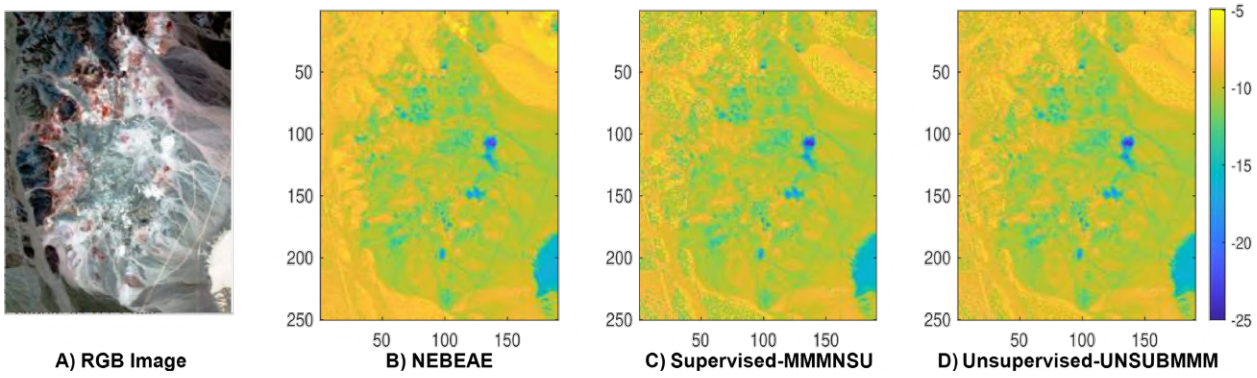


Figure 4.11: Approximate RGB image and estimated levels of nonlinear interaction for the Cuprite dataset: (a) RGB image, (b) NEBEAE, (c) MMMNSU, and (d) UNSUBMMM.

five and six end-members. In terms of abundance estimation error E_a, NEBEAE performed the best with

Table 4.9: Estimation errors for blind nonlinear unmixing methods in the Cuprite dataset: NEBEAE, MMMNSU, UNSUBMMM, and BSNHU. The lowest error and computational time are highlighted in bold.

	Parameter			
Method	E_z	E_p	E_{SAM}	Computational Time (seconds)
NEBEAE	**0.0103**	**0.0532**	**0.0532**	**257.0032**
MMMNSU	0.1614	0.0630	0.0629	617.9150
UNSUBMMM	0.0776	0.0566	0.0564	1151.5394
BSNHU	0.0233	0.0608	0.0608	125164.7874

four and six end-members, and for five end-members, NEBEAE was within the 20% range of MMMNSU, which had the lowest error. BSNHU achieved the best end-member estimation error E_p, and only UNSUBMMM outperformed NEBEAE in terms of SAM E_{SAM} for six end-members. In conclusion, NEBEAE demonstrated a significant improvement by greatly reducing computational time while maintaining a balanced level of estimation performance. Figure 4.12 displays the approximated RGB image and the estimated levels of nonlinear interaction for NEBEAE, MMMNSU and UNSUBMMM with four end-members. The results demonstrate a high level of consistency in this parameter among the three methods. The presence of negative values in Figure 4.12 indicates areas with enhanced reflectance compared to the LMM, which aligns with findings reported in [178].

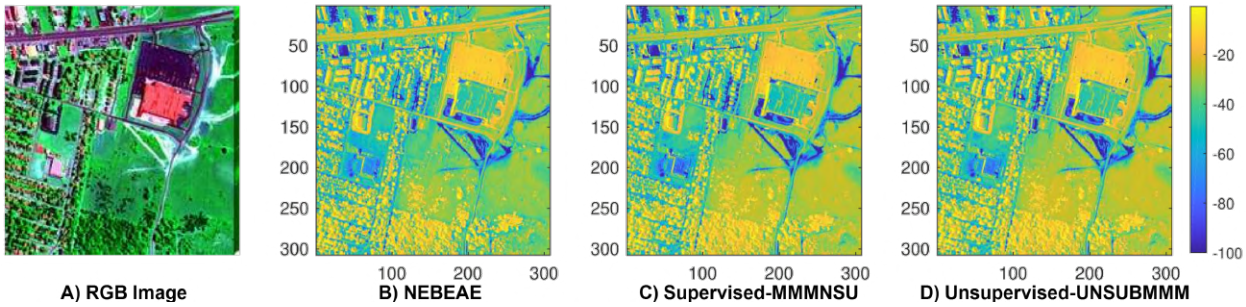


Figure 4.12: Approximate RGB image and estimated levels of nonlinear interaction for the Urban dataset: (a) RGB image, (b) NEBEAE, (c) MMMNSU, and (d) UNSUBMMM.

Table 4.10: Estimation errors for blind nonlinear unmixing methods in the Urban dataset: NEBEAE, MMMNSU, UNSUBMMM, and BSNHU. The lowest error and computational time are highlighted in bold.

	Number of End-members											
	Output Estimation Error E_z			Abundance Estimation Error E_a			End-member Estimation Error E_p			Computational Time (seconds)		
Method	4	5	6	4	5	6	4	5	6	4	5	6
NEBEAE	0.0576	**0.0449**	**0.0467**	**0.2373**	0.3314	**0.4346**	0.1306	0.2116	0.1653	**64.7214**	**70.6075**	**95.7871**
MMMNSU	0.0987	0.0832	0.1166	0.5897	**0.2757**	0.7312	0.0849	0.1148	0.1318	306.0314	349.4668	355.7806
UNSUBMMM	**0.0493**	0.0455	0.0731	0.3937	0.5476	0.5859	0.1165	0.2098	0.1242	1645.6323	1730.4609	1748.7644
BSNHU	0.1842	0.2224	0.2339	0.5018	0.7285	0.7690	**0.0452**	**0.1059**	**0.1073**	2854.1665	4031.3013	4597.8123

Table 4.11: Estimation errors for blind nonlinear unmixing methods in the Pavia University Scene dataset: NEBEAE, MMMNSU, UNSUBMMM, and BSNHU. The lowest error and computational time are highlighted in bold.

	Parameter			
Method	E_z	E_p	E_{SAM}	Computational time
NEBEAE	**0.0406**	**0.0121**	**0.0118**	**248.6713**
MMMNSU	0.0863	0.2117	0.1911	2381.4301
UNSUBMMM	0.0445	0.1642	0.1120	2896.0229
BSNHU	0.0596	0.0489	0.0481	277829.7414

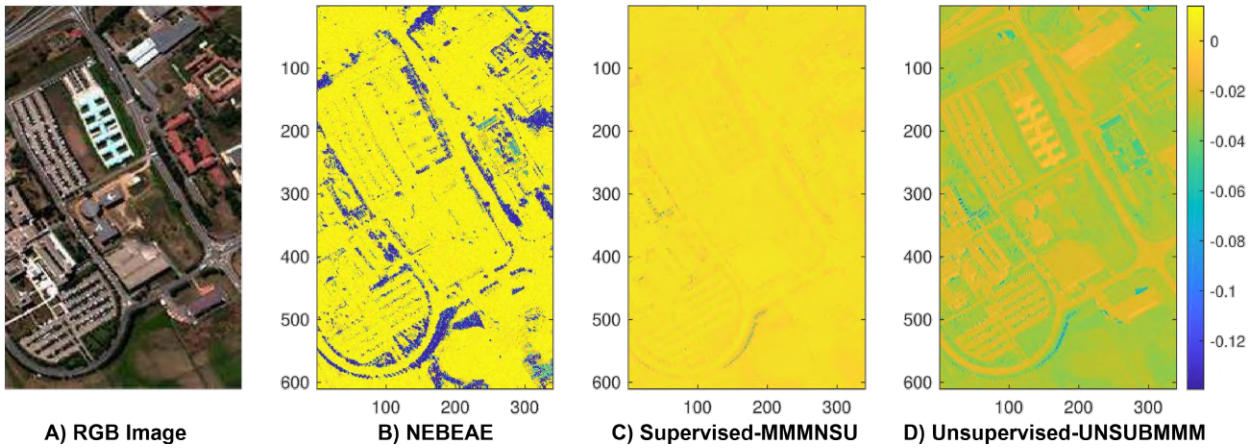


Figure 4.13: Approximate RGB image and estimated levels of nonlinear interaction for the Pavia University Scene dataset: A) RGB image, B) NEBEAE, C) MMMNSU, and D) UNSUBMMM.

4.6.7.3 Pavia University Scene Dataset

The final HS image from remote sensing is the scene of Pavia University. This database is also described in Section 3.2, where the main characteristics are discussed. The hyperparameters used in NEBEAE for this experiment were identical to those used in the synthetic evaluation, with $\rho = 0.1$ and $\tilde{\mu} = 0.1$. The initialization of end-members was performed using N-FINDR, which yielded optimal results. Table 4.11 presents the results of the nonlinear unmixing methods NEBEAE, MMMNSU, UNSUBMMM, and BSNHU in terms of quantitative indexes. Once again, NEBEAE shows superior performance in all metrics (E_z, E_p, E_{SAM}) and exhibits the shortest computational time. The second best algorithm in terms of error metrics was BSNHU, although its computational time was significantly longer compared to MMMNSU and UNSUBMMM. Figure 4.13 visually represents the approximated RGB image and the estimated nonlinear interaction levels for NEBEAE, MMMNSU, and UNSUBMMM. Notably, clear distinctions in the estimation of the nonlinear interaction level by NEBEAE are observed, particularly in certain unlabeled regions present in the ground-truth data.

4.6.7.4 Hyperspectral VNIR Images of In-vivo Human Brain Surgery Scenes

The final evaluation of the blind nonlinear unmixing methods was conducted on two in-vivo brain tissue HS images described in Subsection 3.3.2.2, specifically images P15-01 and P20-01. The main objective of this evaluation was to assess the performance of the nonlinear proposal on experimental data for the identification of the four classes of interest: NT, TT, HT, and BG. The BG class encompasses various materials that can be found in a surgical scene, including skull, bone, dura mater, skin, or other tools. This application is focused on evaluating the nonlinear properties within the experimental image.

Since the advantages of NEBEAE over various state-of-the-art blind nonlinear unmixing algorithms (MMMNUS, UNSUBMMM, and BSNHU) were described in the previous examples, in this final scenario, only the proposed algorithm is presented. However, to identify the differences in linear and nolinear unmixing methods, for comparison, EBEAE is considered. Therefore, the labeled dataset was first used for the identification of the end-members using NEBEAE and EBEAE in the two previously mentioned images, where the approximate RGB images are illustrated in the first column of Figure 4.14. The experiments were conducted following an intra-patient validation process.

In EBEAE, six end-members were identified in the labeled dataset: one for NT and TT, two for HT, and four for BG. However, increasing the number of end-members in the NT and TT classes led to a singularity condition during the abundance extraction step [84]. To improve accuracy, the number of end-members was increased in NEBEAE: two for NT, TT, and HT, and four for BG. The abundances were then calculated for the entire image. During the end-member identification step in NEBEAE, the estimated nonlinear interaction levels consistently decreased to an average value of -40 for all cases. This indicates that the reflectance by the LMM dominates, thus confirming the purity of the labeled pixels. The hyperparameters of both EBEAE and NEBEAE were set to $\rho = 0.1$ and $\tilde{\mu} = 0.2$, i.e., the same values as used in the Cuprite dataset. Concentration maps were constructed by adding the total abundance of the end-members for each class, and each pixel was labeled as (NT, TT, HT, BG) based on the maximum abundance value per class. Table 4.12 presents the quantitative results comparing EBEAE and NEBEAE, showing an overall improvement in estimation accuracy with NEBEAE, despite an increase in computational time due to the nonlinear unmixing technique. Figure 4.14 shows the RGB images, ground-truth maps, classified images by EBEAE and NEBEAE, and nonlinear interaction maps for both HS images (P015-01 and P020-01). In P015-01, the tumor region is better defined by NEBEAE, as well as the NT inside the black rubber rings in P020-01. The nonlinear interaction maps in Figure 4.14 reveal a decrease in reflectance ($d_k > 0$) compared to the LMM in certain BG regions and areas with high light reflections, and an increase in some hypervascularized zones ($d_k < 0$). The last row of Table 4.12 shows a significant difference in classification performance between EBEAE and NEBEAE. So, in this scenario, the NMM provides an improved performance in the classification compared to the LMM.

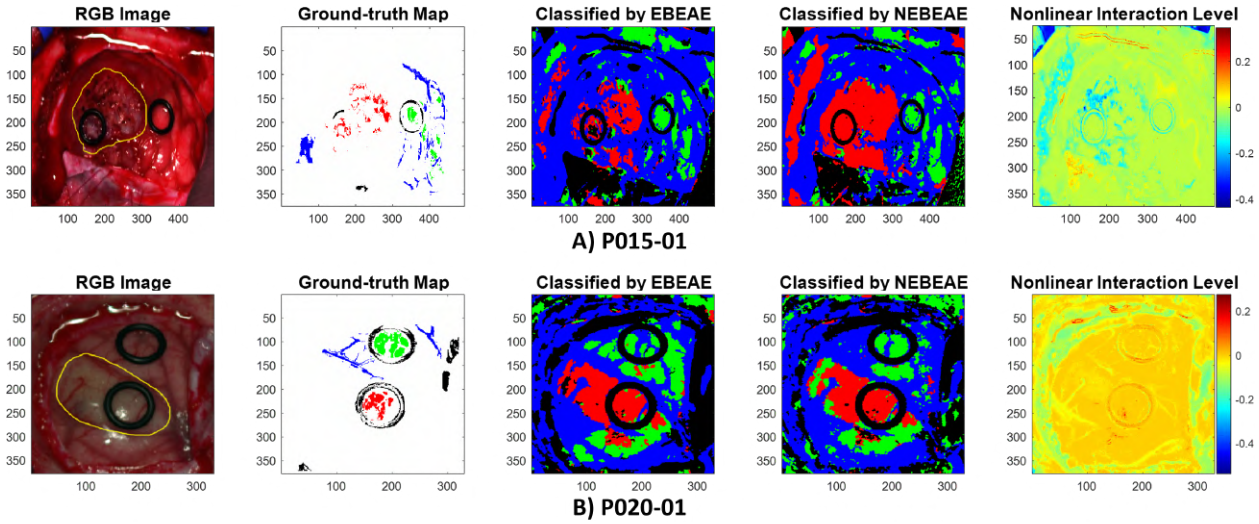


Figure 4.14: Classification results by EBEAE and NEBEAE: A) P015-01 and B) P020-01. The first column shows the synthetic RGB images (tumor region delineated by a yellow line), followed by the ground-truth background map (green - normal tissue, red - tumor tissue, blue - hypervascularized tissue, black - background). The third and fourth columns display the classified images using EBEAE and NEBEAE (same color coding). Finally, the fifth column presents the nonlinear interaction level, obtained from the evaluation with EBEAE.

Table 4.12: Quantitative assessment of estimation errors in the VNIR hyperspectral images of in-vivo human brain surgery scenes, utilizing NEBEAE and EBEAE for classification.

	Output Estimation Error E_z		Classification Precision in Ground-truth Pixels		Computational Time (seconds)		Total Error Percentage in Classified Images NEBEAE vs EBEAE	
VNIR Image	P015-01	P020-01	P015-01	P020-01	P015-01	P020-01	P015-01	P020-01
NEBEAE	**0.0971**	**0.0849**	**0.9893**	**0.9665**	50.9256	48.9586	0.2320	0.3638
EBEAE	0.1172	0.0972	0.9756	0.9311	**9.7039**	**6.0889**		

4.7 Classification of Hyperspectral In-vivo Brain Tissue Based on Linear Unmixing

In existing literature, previous studies have utilized SU approaches such as PCA and ICA for cancer detection using HSI. Other methods like MCR, NNMF, and EBEAE have been employed to identify cancer cells in various biomedical imaging applications. Moreover, ML techniques including supervised and unsupervised algorithms have been investigated for brain tissue classification. In this section, we propose two methods based on BLU using the original EBEAE approach for the classification of in-vivo brain HS images. The main goal of these methods is to achieve precise and efficient classification results for HSI, specifically for intraoperative brain tissue HS images discussed in Subsection 3.3.2.2.

These two strategies adopt a semi-supervised approach to perform the classification, starting with a small set of labeled pixels as ground-truth. These pixels belong to four general classes: NT, TT, HT, and BG. The BG class encompasses a wide range of spectral signatures, including distinctive flat spectral sig-

natures from rubber marker rings and other materials [4, 121, 123]. To mitigate variability in the unmixing process, an initial segmentation of the rubber marker rings based on their flat spectral signatures is conducted, particularly in the lower frequency bands [4]. This step aims to prevent misclassifications caused by the presence of rubber, moisture, and light scattering. The energy is calculated from the initial twenty spectral signatures and raised to the power of $3/2$ to enhance the magnitude differences. The resulting image is then normalized to grayscale tones. Finally, the Otsu method is applied to the grayscale image to segment the regions associated with the rubber marker rings [179]. Subsequently, two proposed classification methodologies based on BLU were explored, differing primarily in the assumption that an initial segmentation of the binary classes BG versus non-BG (NT, TT, and HT) could improve overall accuracy without significant computational overhead.

4.7.1 Method A

The first methodology comprises three primary stages after the segmentation of the rubber marker rings, as illustrated in Figure 4.15. In the initial stage, the distinctive end-members of the four studied classes (NT, TT, HT, and BG) were estimated using the EBEAE algorithm. This estimation process leverages the labeled pixels from the ground-truth datasets of each image as training data. In this case, method A uses labeled data from all HS images except for the HS image under analysis to train the algorithm.

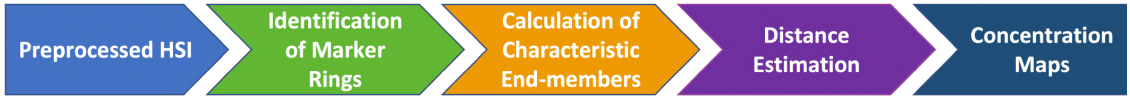


Figure 4.15: Schematic representation illustrating the different steps involved in method A [5].

The number of representative end-members for each class is determined based on the variability of spectral signatures in the dataset [123], following the approach described in [180]. In our study, we selected two representative end-members for NT ($N_{NT} = 2$), two for TT ($N_{TT} = 2$), one for HT ($N_{HT} = 1$), and three for BG ($N_{BG} = 3$) in all HS images. As for the hyperparameters of the EBEAE algorithm, they were manually chosen to improve the classification results, taking into account the recommendations in [84]. The entropy weight parameter (μ) was set to zero to minimize estimation errors, while the similarity weight parameter (ρ) was selected as 0.3 for NT ($\rho_{NT} = 0.3$), 0.2 for TT ($\rho_{TT} = 0.2$), and 0.01 for BG ($\rho_{BG} = 0.01$) to account for dataset variability. After this estimation process, the sets of representative end-members $\{\mathcal{P}_{NT}, \mathcal{P}_{TT}, \mathcal{P}_{HT}, \mathcal{P}_{BG}\}$ for NT, TT, HT, and BG classes were obtained, respectively. It should be noted that $\text{card}(\mathcal{P}_{NT}) = 2$, $\text{card}(\mathcal{P}_{TT}) = 2$, $\text{card}(\mathcal{P}_{HT}) = 1$, and $\text{card}(\mathcal{P}_{BG}) = 3$, as mentioned earlier.

In the next stage, the distance between each pixel $\mathbf{y}_k$ (where k represents the pixel index) in the HS

image and the corresponding end-member sets was calculated using the concept of point-to-set distance:

$$d(\mathbf{y}_k, \mathcal{P}_i) = \inf_{\mathbf{p}\in\mathcal{P}_i} d(\mathbf{y}_k, \mathbf{p}) \quad i \in \{NT, TT, HT, BG\},\ k \in \{1, \ldots, K\} \tag{4.76}$$

where $d(\cdot, \cdot)$ represents a distance or metric. In this study, five distances were evaluated:

$$d_1(\mathbf{x}, \mathbf{y}) = \sum_{l=1}^{L} |(\mathbf{x})_l - (\mathbf{y})_l| \qquad \text{Manhattan metric} \tag{4.77}$$

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x}-\mathbf{y})^\top(\mathbf{x}-\mathbf{y})} \qquad \text{Euclidean metric} \tag{4.78}$$

$$d_C(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^\top\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \qquad \text{Correlation metric} \tag{4.79}$$

$$d_M(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x}-\mathbf{y})^\top\mathbf{Q}(\mathbf{x}-\mathbf{y})} \qquad \text{Mahalanobis metric} \tag{4.80}$$

$$d_{SAM}(\mathbf{x}, \mathbf{y}) = \cos^{-1}\left(\frac{\mathbf{x}^\top\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\right) \qquad \text{SAM metric} \tag{4.81}$$

where $\mathbf{Q} \in \mathbb{R}^{L\times L}$ represents the covariance matrix of the dataset. In the final step, the classification of the k-th pixel $\mathbf{y}_k$ is determined by calculating the minimum distance between the pixel and the four available classes:

$$c(\mathbf{y}_k) = \arg \min_{i\in\{NT,TT,HT,BG\}} d(\mathbf{y}_k, \mathcal{P}_i). \tag{4.82}$$

4.7.2 Method B

In the case of the second approach (refer to Figure 4.16), the main focus was on achieving a precise estimate of the BG class, which exhibits the most significant spectral variation [4, 121]. Initially, the computation of representative end-members for each class follows the same steps as in method A. Subsequently, the collection of all these estimated end-members is defined as follows:

$$\mathcal{P} = \mathcal{P}_{NT} \cup \mathcal{P}_{TT} \cup \mathcal{P}_{HT} \cup \mathcal{P}_{BG}.$$

Subsequently, for each pixel $\mathbf{y}_k$ (where k ranges from 1 to the total number of pixels) in the HS image, the distance to all $M = \text{card}(\mathcal{P})$ estimated characteristic end-members in $\mathcal{P}$ is calculated. This process generates M images, each representing the distance to a specific characteristic end-member. To further analyze these images, a segmentation step is performed using the K-means algorithm, which separates the pixels into four distinct groups [181].

Subsequently, using the non-BG labeled pixels (NT, TT, and HT) in each HS image, the regions corresponding to the BG class are identified based on the positions of the spectral signatures of the non-BG class within the HS image. In this final step, a binary image is created specifically for the BG class, which is then combined with the regions containing the rubber ring markers to construct the comprehensive BG image for that particular patient. Pixels labeled BG are not considered in subsequent stages of the process. Moving

on to the next phase of method B, the distance between the k-th pixel $\mathbf{y}_k$ located outside the BG mask in the HS image is computed with respect to the sets of end-members $\mathcal{P}_{NT}, \mathcal{P}_{TT}, \mathcal{P}_{HT}$ using Equation (4.76). Finally, in the last stage, the classification of the k-th pixel $\mathbf{y}_k$ is determined based on the minimum distance calculated in Equation (4.82), considering only three classes $\{NT, TT, HT\}$.

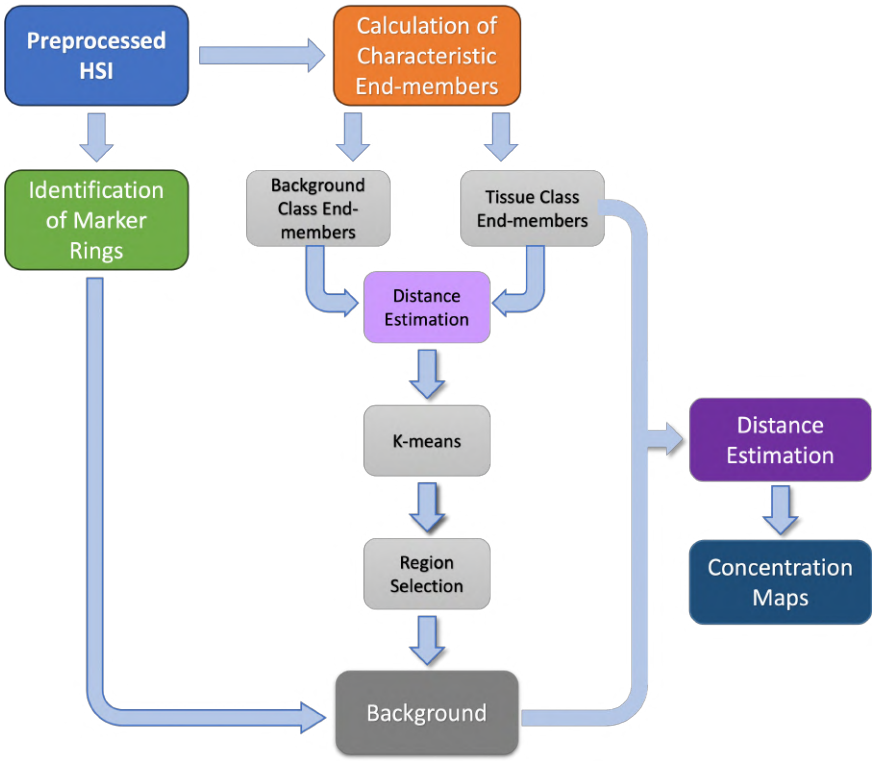


Figure 4.16: Schematic representation of the sequential steps involved in method B [5].

4.7.3 Results and Discussion

In this section, the performance of the proposed BLU-based classification methodologies by EBEAE is presented, with respect to an SVM-based approach from the literature [51]. The effect of different metrics in method A is first evaluated, and then the comparison results between method A, method B, and the SVM-based scheme are shown.

4.7.3.1 Metrics Evaluation in Distance to End-members Sets

Method A relies on the identification of rubber rings markers, which the surgeon places to delineate the tumor and healthy tissues, and then extracting the characteristic end-members by excluding these markers information. Evaluation of this methodology involved utilizing the six test HS images shown in Figure 3.5 (P008-01, P008-02, P012-01, P012-02, P015-01 and P020-01), which database is described in Section

3.3.2.2. Various distance metrics were employed for the evaluation, including Manhattan metric, Euclidean metric, correlation metric, Mahalanobis metric and SAM metric, as defined in Equations (4.77)-(4.81).

The average classification performance results for each metric are presented in Figure 4.17. This figure illustrates the accuracy, sensitivity, and specificity outcomes. Across all metrics, the accuracy values are comparable. Similarly, the sensitivity performance for the NT and BG classes is consistent across all metrics. However, notable performance differences are observed for the TT and HT classes. The Manhattan metric demonstrates the highest sensitivity in the HT class, achieving a value of approximately 70%, while the correlation metric exhibited the lowest performance around 47%. On the contrary, in the TT class, the correlation metric delivers the best sensitivity results at 50%, while the Manhattan metric performs the poorest at approximately 32%. With respect to specificity, consistent results were observed across all classes. The NT class consistently achieves specificity above 70%, while the TT, HT, and BG classes have specificity above 90%. Figure 4.17 also presents the results of the MCC. It can be observed that most metrics yielded similar performance, with the exception of the correlation metric in the HT class, which exhibits a slight decrease of around 7% compared to the best performing metric. However, it should be noted that the correlation metric performs exceptionally well in the TT class, exceeding other metrics. Therefore, on average, the correlation metric demonstrates the best overall results, particularly in the TT class, which is of great significance given the nature of the clinical application.

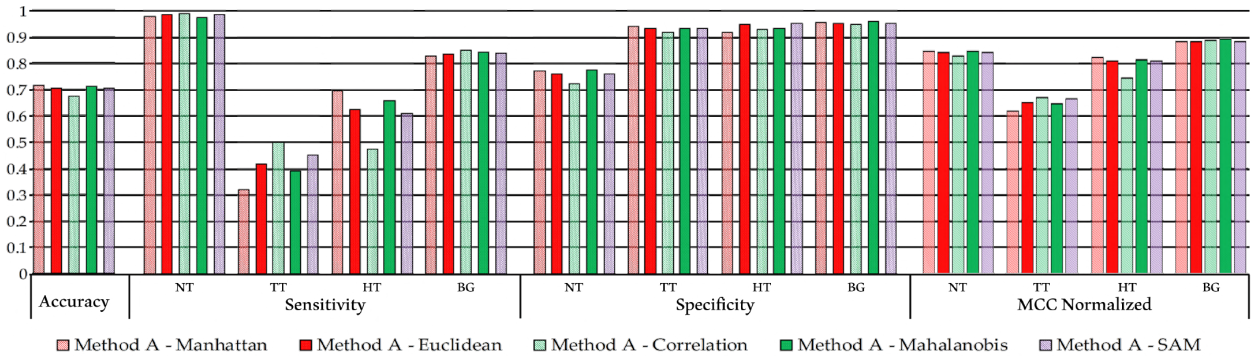


Figure 4.17: The average classification performance of method A utilizing different metrics (4.77)-(4.81): accuracy, sensitivity, specificity, and MCC normalized.

The classification maps generated using each metric for method A are illustrated in Figure 4.18. Figure 4.18A displays the synthetic RGB images where the tumor area is indicated by a yellow line. Figures 4.18B to 4.18F show the classification results obtained with all metrics in Equations (4.77)-(4.81). The classification results demonstrate consistent performance across all metrics, as indicated by the accuracy evaluation in Figure 4.17. However, when it comes to defining tumor areas in HS images, the correlation metric provided better results. This is evident in images P012-02 and P015-01, where the correlation metric successfully identified tumor areas with greater precision compared to other metrics. For example, in the

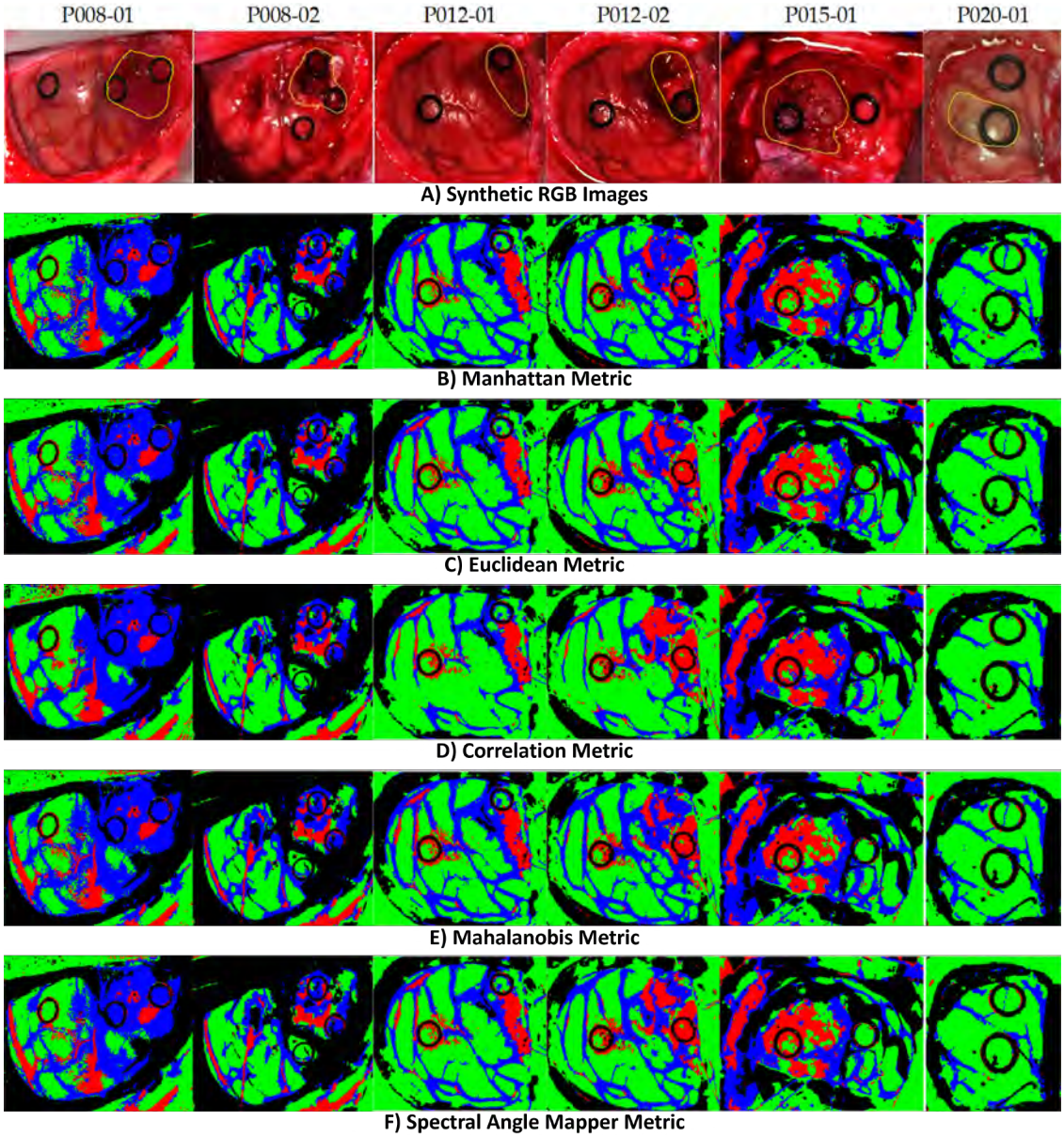


Figure 4.18: Classification maps generated for method A using the metrics defined in Equations (4.77)-(4.81): A) Synthetic RGB images with the tumor area identified by a yellow line, B) Method A with Manhattan metric, C) Method A with Euclidean metric, D) Method A with correlation metric, E) Method A with Mahalanobis metric, F) Method A with SAM metric.

case of image P012-02, the Manhattan metric (Figure 4.18B) misclassified some pixels in the HT class within the tumor area. On the contrary, the correlation metric (Figure 4.18D) accurately identified the tumor area and exhibited a more homogeneous classification compared to other distances. It should be noted that for P020-01, none of the metrics was able to distinguish the TT class. Synthetic RGB image reveals that the marked tumor area presented a similar colorization to the NT class, which aligns with the findings in [4].

As mentioned in the aforementioned study, this misclassification of the TT class may be attributed to the limited representation of inter-patient spectral variability in the database.

4.7.3.2 Comparison Results

In method B, it was determined through our analysis that the correlation metric was the most suitable choice for enhancing accuracy in extracting the BG class. Additionally, the Mahalanobis metric was selected to improve the performance in tissue classification. Consequently, a comparison was made between method A with the correlation metric and method B with the correlation/Mahalanobis metrics against the SVM-based approach. The evaluation was carried out using six test HS images (P008-01, P008-02, P012-01, P012-02, P015-01 and P020-01) in a leave-one-patient-out cross-validation setup. On average, method A yielded an overall accuracy of 67.2$\pm$11.5%, while method B achieved 76.1$\pm$12.4%. These results were lower than the accuracy obtained by the SVM-based approach, which averaged at 79.2$\pm$15.6%. However, as shown in Figure 4.19, the use of alternative metrics for each class demonstrated some improvements in classification performance.

The sensitivity results for the TT class were significantly improved with the proposed approaches compared to the SVM-based method, which is the most notable outcome for this specific application (Figure 4.19A). Method A and B achieved a median sensitivity of 47.8% and 31.3%, respectively, representing an increase of 26.2% and 9.7% compared to the sensitivity obtained with the SVM approach (21.6%). Conversely, the median sensitivity for the NT class remained relatively constant across all three approaches, with values exceeding 97% and reaching 99.7% in method A. However, both methods A and B experienced a decrease in median sensitivity for the HT class, resulting in sensitivity values of 46.5% and 18%, respectively, compared to the performance of SVM (92. 9%). Nevertheless, in this specific application, the accurate identification and differentiation of the NT and TT classes was more important than the identification of hypervascularized tissue, which could be visually distinguished or identified using image processing algorithms based on morphological characteristics of blood vessels.

In terms of the specificity results (Figure 4.19B), there were minor variations among the three proposed methods, except for the NT and TT classes. Method A exhibited a median specificity value of 70.4% for the NT class, while method B and the SVM scheme achieved values of 88.8% and 87.3%, respectively. The specificity of the TT class decreased slightly in the two proposed methods, with higher interquartile ranges (IQR) compared to the SVM-based approach, but with median values above 92%.

Regarding the F1-score results, Figure 4.19C shows that methods A and B improved the median values for the TT class, reaching 30.7% and 31%, respectively, compared to the result obtained by the SVM-based approach (25.7%). For the NT and BG classes, method B achieved the highest median results of 90.7% and 98.3%, respectively. On the other hand, the SVM-based approach obtained the best result in the HT class (91.5%), which is consistent with the results of the sensitivity metric.

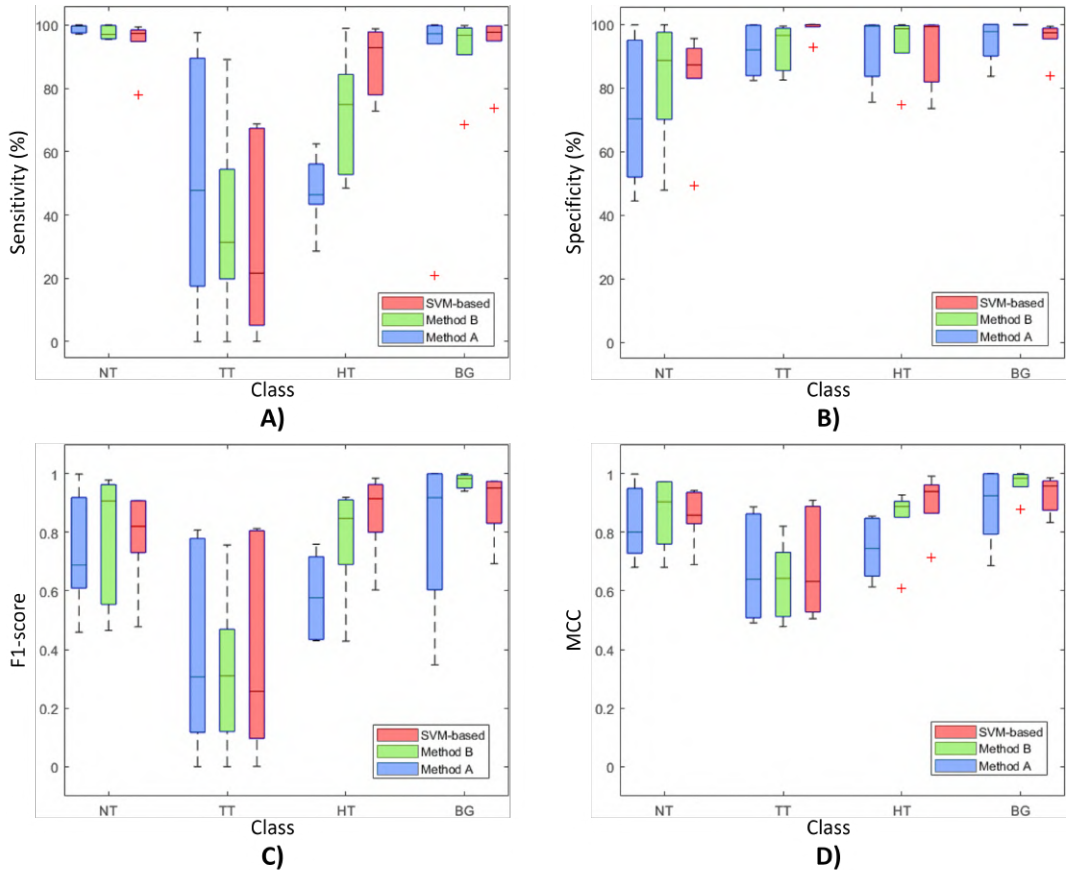


Figure 4.19: Box plots of classification performance of method A, method B and SVM-based approach using leave-one-patient-out cross-validation: A) Sensitivity, B) specificity, C) F1-score, and D) MCC (normalized).

Finally, the normalized MCC results are depicted in Figure 4.19D, that takes into account the imbalanced dataset. These results reveal that the median values for the TT class were quite similar (approximately 66%) among the three methods. Method B exhibited the highest median MCC value for the NT class at 90.4%. In contrast, methods A and B showed a reduction in the median MCC value for the HT class, approximately 19% and 5% respectively, compared to the SVM-based approach.

The classification maps shown in Figure 4.20 provide a qualitative evaluation of the entire HS images, including non-labeled pixels. Figure 4.20A presents the synthetic RGB images with the tumor area delineated by a yellow line. Figures 4.20B, C, and D display the classification maps obtained using the SVM-based approach, method A, and method B, respectively. These results demonstrate that the proposed methods improved the labeling of pixels within the tumor area compared to the SVM-based approach. However, the proposed methodologies exhibited more false positives in non-tumor areas. Regarding the other tissue classes, the qualitative results were generally similar, except for the BG class, where method B demonstrated accurate identification of the parenchymal area (exposed brain surface) in images P008-01, P008-02,

and P012-02. Furthermore, these results indicate that the lower classification performance observed in the quantitative results for the HT class in method A (see Figure 4.19) is attributed to misclassifications between the BG and HT classes in images P008-02 and P015-01, where the main blood vessels were identified as background. This phenomenon did not occur in method B, where the hypervascularized areas were overall well identified.

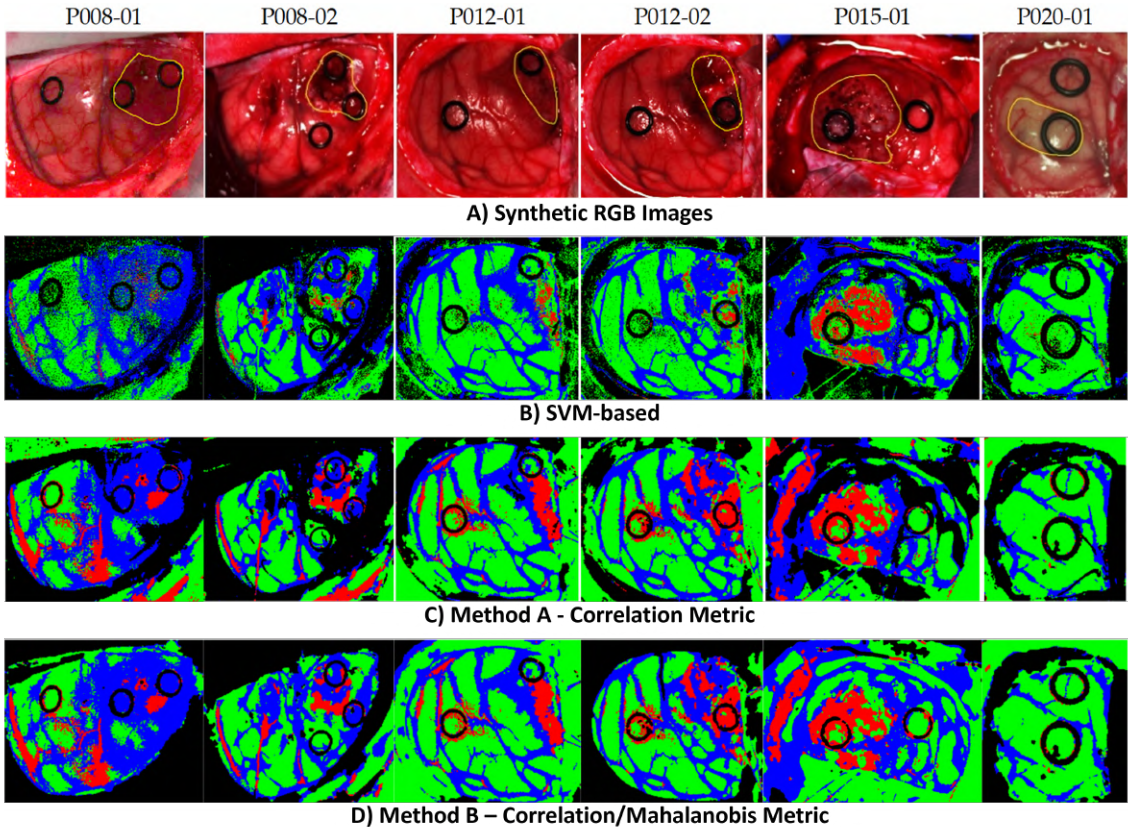


Figure 4.20: Classification maps for method A, method B, and SVM-based approach: A) Synthetic RGB images with the tumor area identified by a yellow line, B) SVM-based approach, C) Method A with correlation metric, and D) Method B with correlation/Mahalanobis metrics.

In terms of comparing the execution time among the three methods, Figure 4.21 presents the average time required for the six test HS images. To facilitate the comparison, a logarithmic scale was utilized. The execution time encompassed both the training and classification processes for the SVM-based approach, as well as the complete execution of the two proposed methods illustrated in Figures 4.15 and 4.16. These results were obtained using MATLAB® on a system equipped with an Intel i7-4790K processor at 4.00 GHz and 8 GB of RAM.

The SVM-based approach needed nearly four hours to train and classify a single HS image. Conversely, the proposed methods based on the EBEAE algorithm required an average of approximately 30 seconds to complete the training and classification tasks for the datasets. Notably, these methods achieved comparable

accuracy results as discussed in the previous sections. In summary, methods A and B offered significant speedup factors of approximately 459x and 429x, respectively, compared to the SVM-based approach.

Figure 4.21: Execution time for method A and B, and SVM-based approach (representation in logarithmic scale).

Chapter 5

Reflectance Calibration with Normalization Correction in Hyperspectral Imaging

The calibration stage establishes an amplitude adjustment for the spectral signatures stated in percentages with the intention of compensating for negative effects during acquisition [182]. This step enables the standardization of spectral signatures, which ensures that the morphology of the spectral data obtained from the same object or component by various HS cameras is the same in both cases. The standard calibration (SC) process, carried out in most high-speed imaging studies, is carried out through the linear transformation of Equation (2.1) into the raw unprocessed image ($I_R \in \mathbb{R}^{X \times Y \times L}$). Equation (2.1) uses an inverse model viewpoint to account for nonuniform gain and is derived from the spectral response of the sensors as stated in [183]. As a result, this SC procedure offers comparable spectral signatures for acquisitions conducted with various HS cameras and circumstances using the same components.

Nonetheless, the HSI stage still has problems, despite the benefits of the SC method described in the literature (Equation (2.1)). Therefore, the effects generated by optical filters used in the HS camera sensors are not compensated entirely by the conventional transformation. Optical filters are employed as bandpass components to reject undesirable data and restrict the acquisition to a particular wavelength band [184]. However, the fundamental characteristics of bandpass filters, in conjunction with electronics, can result in some parasitic effects, such as crosstalk, leakage, and harmonics, which cause sensors to measure light from undesirable wavelengths, leading to slight shape changes and an increase in the amplitude of spectral signatures [184, 185]. Harmonic effects, which mainly alter the amplitude of the spectral signatures and have different impacts depending on the material being studied, are still taken into account in measurements [184]. Because HS cameras produce discretized spectral responses for each pixel, the effects of harmonics are typically undetectable [184]. Nonetheless, it becomes feasible to assess the impact by contrasting the measurements under capture conditions identical to those of a white reference sample.

In this context, this chapter proposes a modification to the standard reflectance calibration, as described

in Equation (2.1). This modification aims to address the significant impact caused by unwanted alterations in the amplitude of spectral signatures. The proposed calibration (PC) incorporates a normalization correction, ensuring a resulting reflectance range between zero and one hundred. This adjustment eliminates the need for subsequent normalization steps. Moreover, the PC method reduces variability and preserves the morphology of spectral signatures, as demonstrated in the initial tests for the classification of different compounds. In addition to this, a second experimental evaluation of the PC was conducted in conjunction with the developed NEBEAE algorithm. This was done to assess the impact of the PC on data containing nonlinear components.

5.1 Proposal Reflectance Calibration

The proposed method departs from the raw HS image (I_R), and the dark (I_D) and white (I_W) references, which are used in the standard reflectance calibration. The proposed reflectance calibration is motivated by (2.1), but defines two auxiliary images:

$$\alpha(x, y, \lambda) = I_{raw}(x, y, \lambda) - I_D(x, y, \lambda) \tag{5.1}$$

$$\beta(x, y, \lambda) = I_W(x, y, \lambda) - I_D(x, y, \lambda) - \min_{x,y,\lambda} \alpha(x, y, \lambda), \tag{5.2}$$

where $\alpha(x, y, \lambda) \in \mathbb{R}^{X \times Y \times L}$ and $\beta(x, y, \lambda) \in \mathbb{R}^{X \times Y \times L}$ are the numerator and denominator of (2.1). To ensure that the minimum values are equal to or greater than zero, an additional component is introduced in the auxiliary image $\beta(x, y, \lambda)$ by subtracting an offset from the spectral signals in the denominator of Equation (2.1). This adjustment guarantees the non-negativity of the minimum values. Subsequently, the same operation is applied to $\alpha(x, y, \lambda)$ to correct for any negative values that may arise from undesired effects during the acquisition stage. It is important to note that in an ideal scenario, the global minimum value of $\alpha(x, y, \lambda)$ should already be greater than or equal to zero. Therefore, even in cases where negative values are present, the smallest values of the spectral signatures are assigned as zero.

After defining $\alpha(x, y, \lambda)$ and $\beta(x, y, \lambda)$, we proceed with the normalization process by:

$$\hat{\beta}(x, y, \lambda) \triangleq \frac{\beta(x, y, \lambda)}{\max_{x,y}(\beta(x, y, \lambda))} \in [0, 1], \tag{5.3}$$

when applying this operation, the lower and upper values in $\hat{\beta}(x, y, \lambda)$ are constrained to zero and one, respectively. Consequently, the purpose of $\hat{\beta}(x, y, \lambda)$ is to assess the nonuniform gain across the analyzed wavelengths in a normalized and positive manner. Likewise, $\alpha(x, y, \lambda)$ undergoes a similar adjustment:

$$\hat{\alpha}(x, y, \lambda) \triangleq \frac{\alpha(x, y, \lambda) - \min_{x,y,\lambda} \alpha(x, y, \lambda)}{\hat{\beta}(x, y, \lambda)}, \tag{5.4}$$

where the offset is calculated using the general minimum of $\alpha(x, y, \lambda)$, limiting the minimum values that can be used without constraining the maximum values ($\hat{\alpha}(x, y, \lambda) \in [0, \infty)$). The PC formula is defined as:

$$\hat{I}(x, y, \lambda) = 100 \times \frac{\hat{\alpha}(x, y, \lambda)}{\max_{x,y}(\hat{\alpha}(x, y, \lambda))} \in [0, 100], \tag{5.5}$$

this guarantees values in the [0,100] range and establishes the upper limit of the spectral signatures within the image $\hat{I}(x, y, \lambda)$. The proposed methodology addresses the non-uniform gain of the HS sensor, as shown in Equation (2.1). Hence, in this new approach represented by Equation (5.5), the resulting values are now constrained. This novel perspective effectively reduces spectral variability and improves the accuracy of classification results, as will be demonstrated in the subsequent analysis.

5.1.1 Evaluation of Calibration

This section presents the results of the comparison between the SC and the PC. The evaluation of these two approaches was performed on an HS image in Figure 3.4, and the acquisition process, image characteristics, and results discussion are detailed in the subsequent subsections. All evaluations were carried out using MATLAB®2018a on a computer equipped with a 4.2 GHz Intel Core i7 quad-core processor and 16 GB of RAM.

The evaluation of both the SC method (Equation (2.1)) and the PC method (Equation (5.5)) was carried out using the HS plastic database (see Figure 3.4). The results of this evaluation are presented in Figures 5.1 and 5.2, where each row represents the spectral information of a specific square. The first and second columns illustrate the spectral signatures obtained using the SC and PC methods, respectively. It can be observed that the SC method yields values above 100%, with scenarios such as S1 showing almost all spectral signatures exceeding this threshold. On the contrary, the PC method produces results within the expected range of [0, 100], showing less variability in spectral signatures compared to the SC method. To validate these results, the means and STD of both methods were analyzed in the third column. The values of the PC method were found to remain below 100%, exceeded for the squares S5 and S7. In contrast, the means and STD of the SC method exceed 100% in most cases, except for the aforementioned squares. The STD values demonstrate a reduction in variability for the PC method in all squares compared to SC.

Furthermore, the fourth column displays the mean and STD of both methods after applying the preprocessing stage (described in the section involving normalization from zero to one). The results show similar mean and STD values between the two methods, with PC performing better for squares S5 and S7. This improvement can be attributed to the elimination of the end bands, which helps reduce variability in the spectral signatures. It is worth noting that there is significant variability at the spectral extremes due to the response of the HS sensors at those particular wavelengths. Thus, the preprocessing step effectively minimizes differences between the two methods. However, it should be mentioned that normalization can also remove distinctive features of the information that are relevant for classification purposes. In cases where no undesired effects are present during the capture process, the normalization step may not be necessary.

Apart from the aforementioned comparison, two classifications were conducted on the HS image using both calibration methods, without considering the preprocessing stage. The evaluations involved the utilization of the standard K-means algorithm (Ahmed et al., 2020), with two different values of K ($K \in \{4, 9\}$).

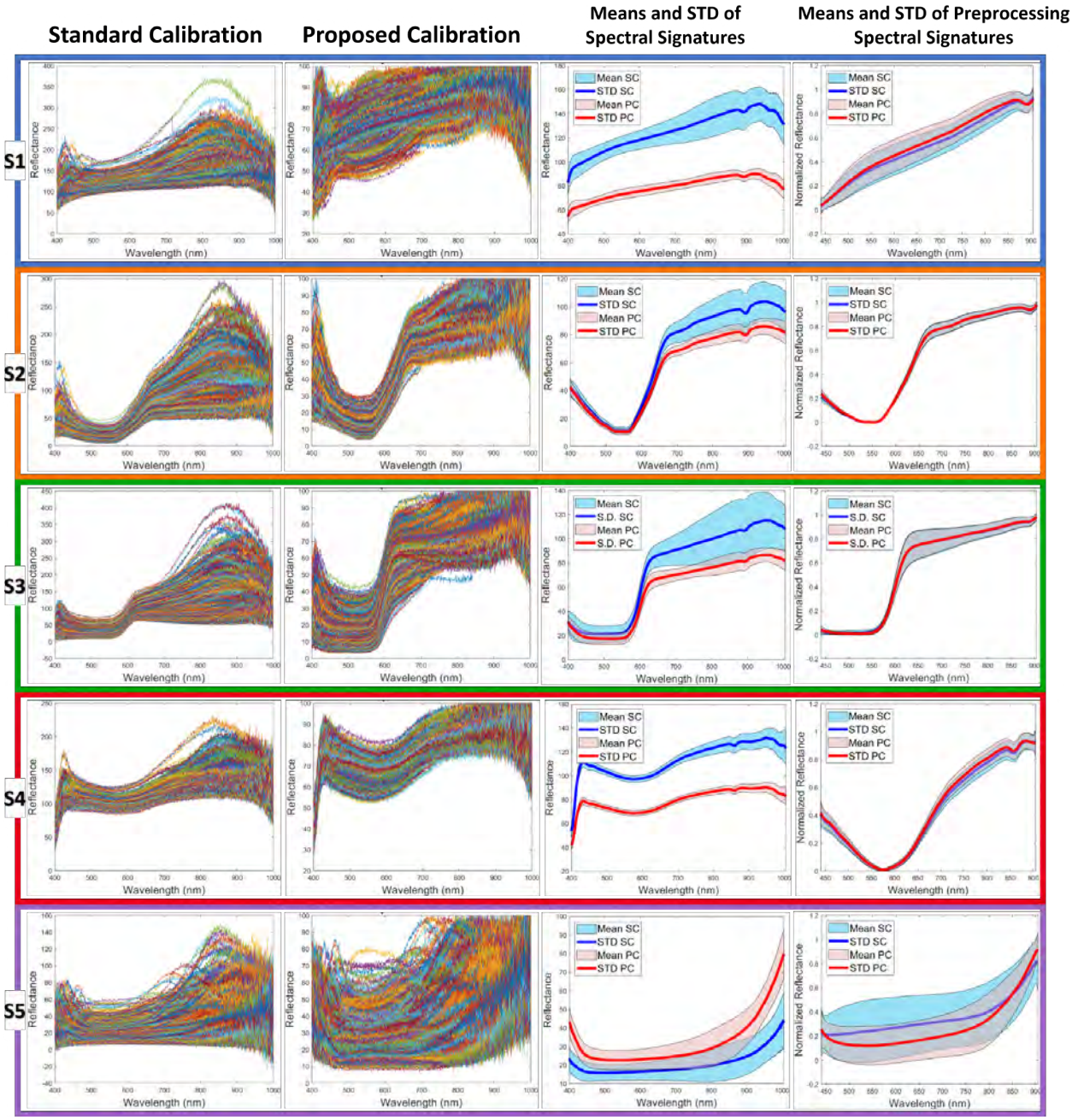


Figure 5.1: Results of calibration of the HS image, where each row represents the spectral information of each square from S1 to S5. The first column shows the results of the standard calibration (SC), while the second column presents the results of the proposed calibration (PC). The third column reports the mean and STD of both approaches. Finally, the fourth column also shows the mean and STD of both methods after performing the preprocessing chain.

The Euclidean distance metric was used for the $K = 4$ case, while the L_1 distance metric was used for the $K = 9$ case. The first classification generated a color-based segmentation, while the second classification

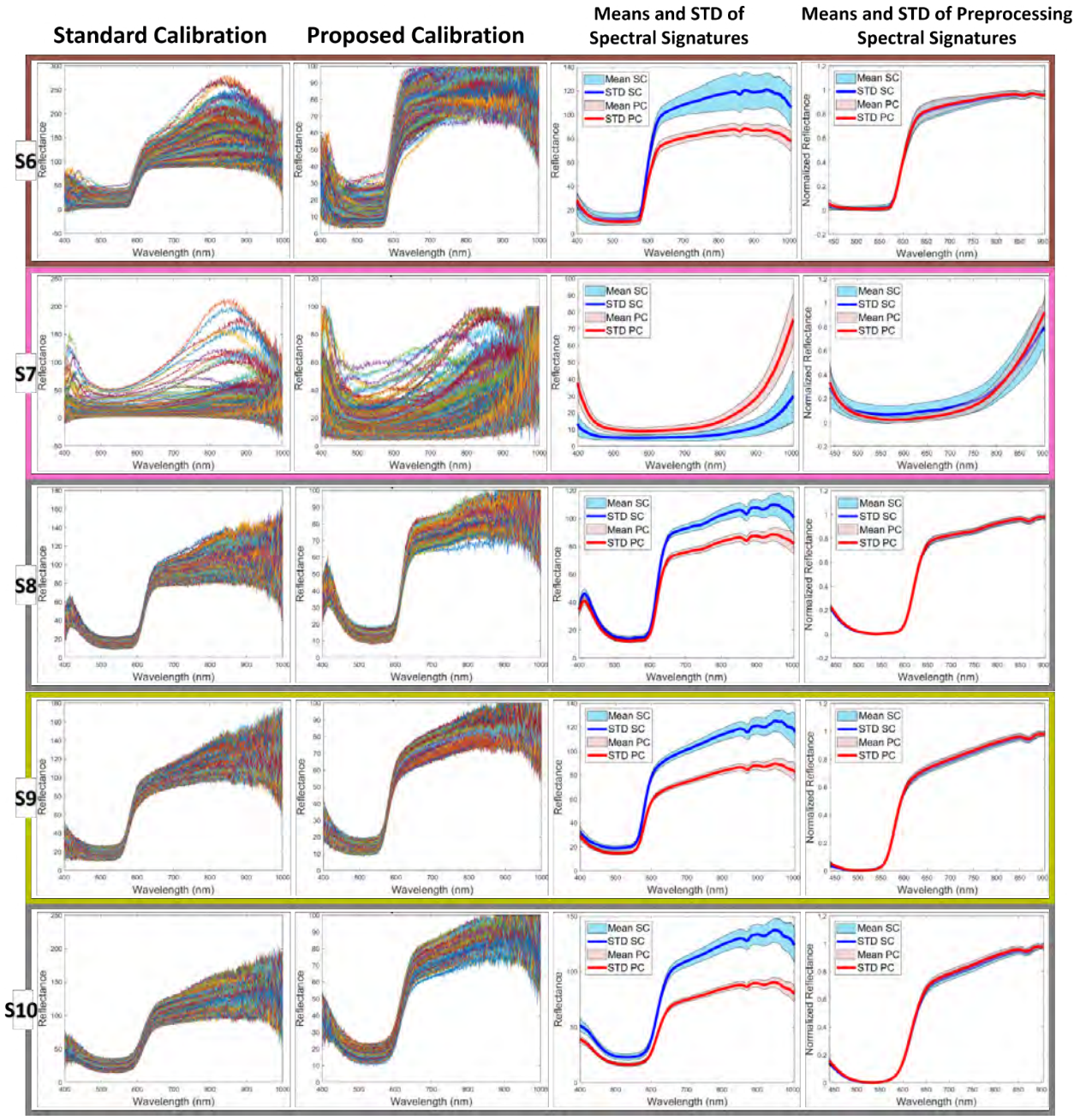


Figure 5.2: Results of calibration of the HS image, where each row represents the spectral information of each square from S6 to S10. The first column shows the results of the standard calibration (SC), while the second column presents the results of the proposed calibration (PC). The third column reports the mean and STD of both approaches. Finally, the fourth column also shows the mean and STD of both methods after performing the preprocessing chain.

focused on a per-square classification. This is to generally consider the differences between the clusters, being more specific for the second case, where the differences are directly taken into account in a linear

manner in the color identification.

The classification results for the color-based segmentation with $K = 4$ are presented in Figure 5.3. It can be seen that the results obtained using the SC method exhibited numerous errors in most squares, leading to an accuracy of 79% (Figure 5.3A). Conversely, the segmentation of the HS image with the PC method demonstrated fewer errors, primarily in squares S3 and S5 (Figure 5.3B). Consequently, the second classification achieved a significantly improved accuracy of 97% compared to the corresponding results obtained using the SC method.

Figure 5.3: Classification results per color. A) standard calibration and B) proposed calibration.

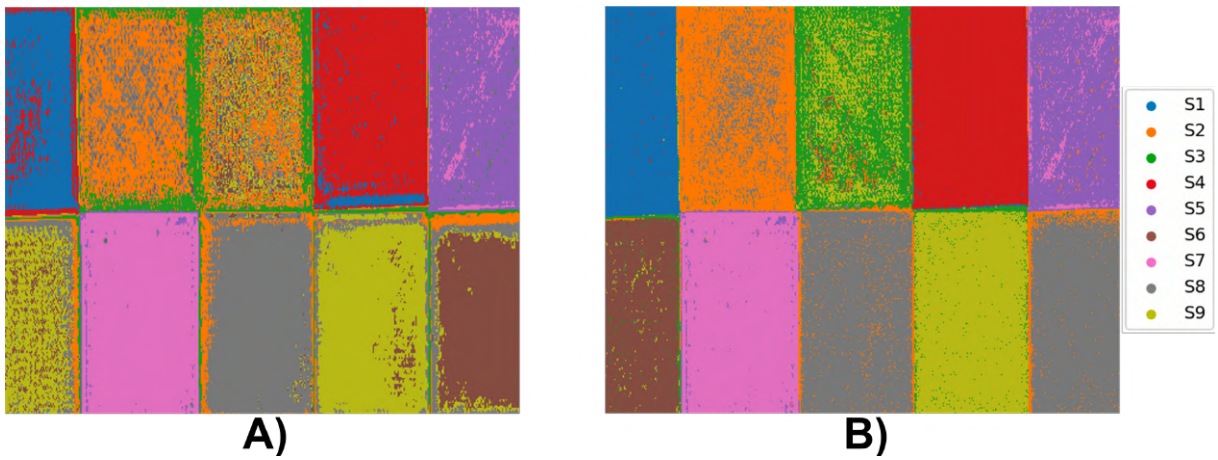


Figure 5.4: Classification results per square. A) standard calibration and B) proposed calibration.

The classification results per square are illustrated in Figure 5.4. Overall, there is an increase in classification errors compared to the color-based segmentation case. This trend can be attributed to the similarity between the spectral signatures and the presence of certain texture characteristics in the squares of the test-

bench.

Figure 5.4A displays the classification results using the SC method, revealing a higher number of classification errors among the labels assigned for different squares, particularly in S2, S3, S6, and S10. Notably, S10 is entirely misclassified as S5, and most of S9 squares are erroneously labeled as S5. In contrast, Figure 5.4B presents the classification results obtained with the PC method, demonstrating a decrease in classification errors compared to Figure 5.4A. The squares with the highest classification accuracy were S2, S3, and S5. As a result, S10 and S8 were correctly classified with the same label due to their identical material and color. In general, the SC classification achieved an accuracy of 64%, while the PC classification achieved 88%. Therefore, in this second evaluation, the PC method once again exhibited a significant improvement over the SC method.

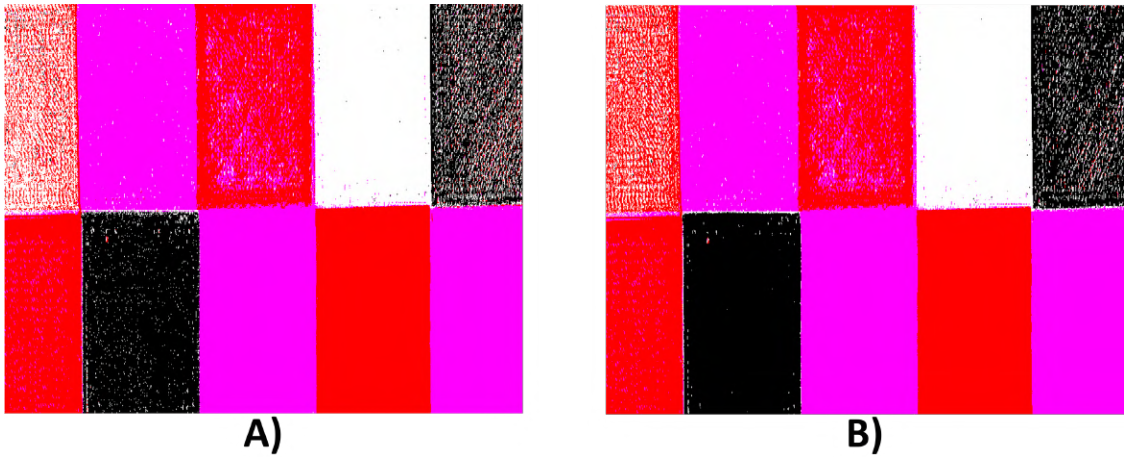


Figure 5.5: Classification results per color with preprocessing stage. A) standard calibration and B) proposed calibration.

The classification tests conducted on the calibrated data, incorporating the preprocessing stage, yielded highly similar outcomes for both SC and PC spectral signatures. The results of the color classification are displayed in Figure 5.5, where most of the errors were concentrated in squares S1 and S5, with a higher noise level observed in the case of PC for these specific squares. Conversely, the remaining squares exhibited lower noise levels when using PC compared to SC. These observations were further supported by the overall classification accuracy analysis, with PC achieving 90% accuracy compared to 89% for SC. This pattern was also evident in the per-square classification, as depicted in Figure 5.6, where PC demonstrated improved accuracy at 69% compared to the 64% accuracy achieved by SC.

In addition to the aforementioned tests, a classification evaluation was performed with both calibration methods, incorporating the preprocessing step but excluding the zero-to-one normalization. The purpose was to assess whether the normalization step was solely responsible for the similarity in the shape of the spectral signatures. Qualitatively, the results of this evaluation closely resembled those presented in Figures 5.3 and

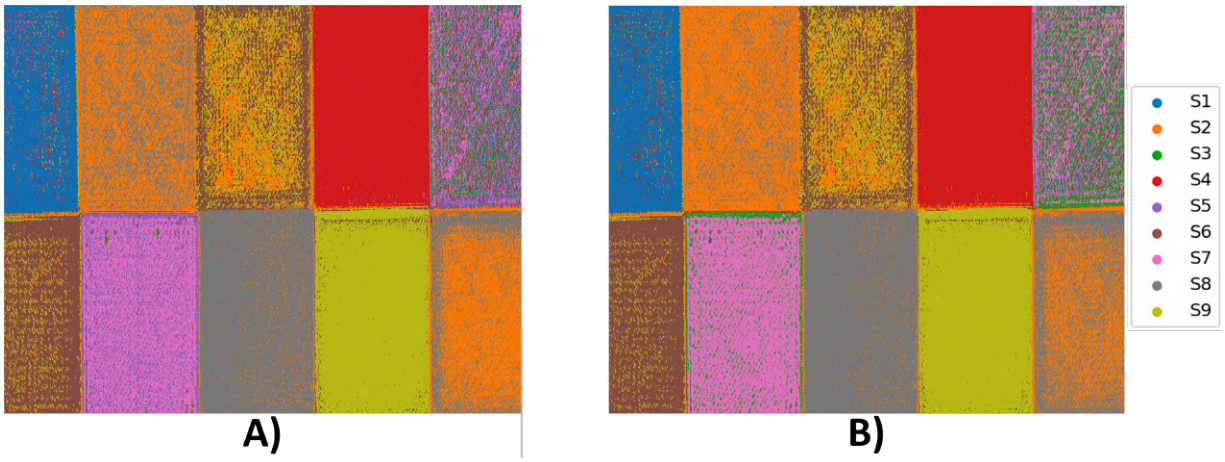


Figure 5.6: Classification results per square with preprocessing stage. A) standard calibration and B) proposed calibration.

5.4 for both SC and PC. However, when examining quantitative accuracy data, SC achieved 79% accuracy in color classification, while PC achieved 95%. Similarly, for square classification, SC yielded 64% accuracy, whereas PC generated 87%. Therefore, there was a slight decrease in precision for PC compared to the classification results obtained solely with the calibration step, while SC maintained consistent performance in both tests.

To conduct a more in-depth analysis of the evaluation results, a dimensionality reduction technique was employed on the calibrated spectral signatures. Local linear embedding algorithm (LLE) [186] was utilized to transform the data into three-dimensional points. LLE is a dimensionality reduction method that aims to preserve the intrinsic geometric properties of the original database by considering the relationships between neighboring points. In this evaluation, LLE was performed with a neighborhood size of five and nearest neighbors were estimated using the KDTree algorithm [187].

The selection of LLE for dimensionality reduction was based on a comparison with other techniques such as principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) [188]. When evaluating the data with both calibration methods, PCA and t-SNE exhibited qualitatively inferior class separation in the resulting three-dimensional space. In contrast, LLE proved effective in reducing the high-dimensional space by capturing nonlinear structures while preserving the underlying geometric characteristics of the original data [188].

The results of the dimensional reduction analysis are presented in Figure 5.7. Figures 5.7A and 5.7B show the dimensional reductions obtained using SC and PC methods, respectively. Additionally, Figures 5.7C and 5.7D display the dimensional reductions after applying the preprocessing step, as described earlier. In the plots showing the SC dimensional reductions, the data points tend to cluster in a central region with significant overlap between different squares. Conversely, the PC dimensional reductions exhibit improved

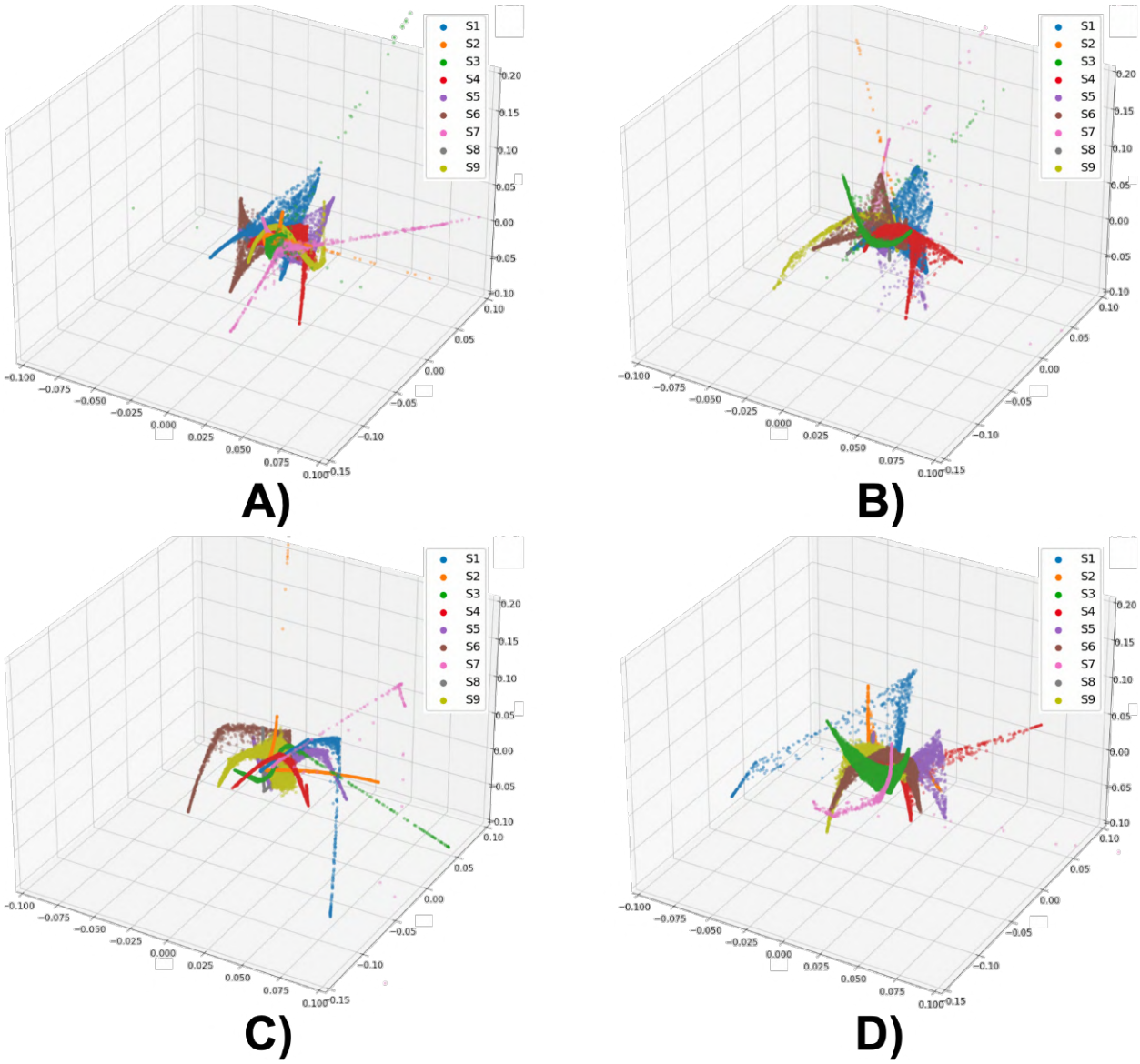


Figure 5.7: The dimensional reduction of spectral signatures was performed using the following approaches: A) standard calibration, B) proposed calibration, C) standard calibration with preprocessing stage, and D) proposed calibration with preprocessing stage.

separability for each square, indicating the effectiveness of the PC calibration method in enhancing the distinguishability of spectral signatures.

Analyzing the dimensional reductions with the preprocessing step, the PC method maintains a consistent shape similar to Figure 5.7B, preserving the proportions in the three-dimensional space and reducing the presence of outliers for each square. Similarly, the reduction in SC dimensions with preprocessing shows a reduction in outliers and sharper curves compared to Figure 5.7A. However, when the zero-to-one nor-

malization is applied, similar behavior for both SC and PC methods is observed, causing the data points to concentrate and overlap in a central area. This behavior aligns with the observations in the fourth column of Figures 5.1 and 5.2. When classification is performed with the preprocessing step, the results demonstrate a high degree of similarity. In summary, particularly for the PC method, where the normalization step is not necessary due to the limitations in the formulation, the application of this transformation modifies the spectral signatures, causing a loss of distinguishing characteristics. This, in turn, hinders the accurate identification of data from each square and consequently leads to less precise classification outcomes.

5.2 Glioblastoma Classification in Hyperspectral Images by Reflectance Calibration with Normalization Correction and Nonlinear Unmixing

In addition to the previous evaluations, comparison tests were carried out regarding the calibration methodologies, followed by a nonlinear unmixing stage using NEBEAE. To carry out this comparison, the methodology described in Figure 5.8 was followed, consisting of five stages: HS image raw input (Figure 5.8A), preprocessing (Figure 5.8B), semi-supervised dataset (Figure 5.8C), nonlinear unmixing (Figure 5.8D), and classification results (Figure 5.8E).

The data used in this evaluation, similar to the previous assessment, was the in-vivo human brain tissue database described in Section 3.3.2.2. Specifically, four images were employed from the database: P008-01, P012-01, P015-01, and P020-01. In this test, the preprocessing chain described in Section 3.4 was applied in general, except for the normalization step. In the calibration stage, Equation (2.1) was used, along with the PC method described in Section 5. After the pre-processing stage, pathology-labeled pixels were used as training data in a semi-supervised perspective, employing intra-patient and inter-patient approaches.

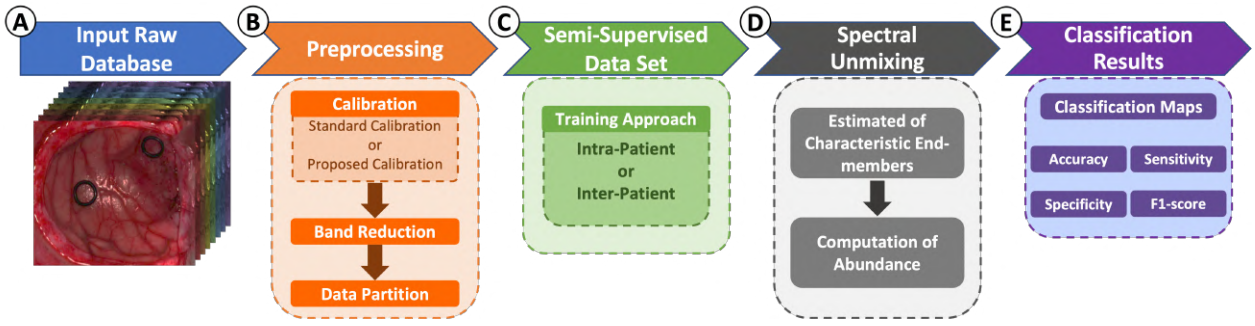


Figure 5.8: General scheme of the methodology for evaluating standard and proposed calibration in the classification of HS images.

Following the preprocessing stage, a feature extraction step was conducted as part of the classification process. This step involved separating the HS data into its constituent end-members, along with their corresponding abundances. The initial phase involved identifying the characteristic end-members for each class in the HS images using the labeled pixels from the database. These labeled pixels were divided into

four distinct sets, representing the four classes, and analyzed from both the intra-patient and inter-patient perspectives. Subsequently, various numbers of end-members were manually selected per class, as depicted in Table 5.1, ensuring consistency across both classification perspectives. These selections were intended to minimize the estimation error during data reconstruction. Hence, when considering a single end-member for a particular class, its spectral signature was determined by averaging the spectral signatures of the labeled pixels within that class. In all other cases, NEBEAE was applied to estimate the final set of end-members for each labeled pixel set. During this stage, specific hyperparameters were set for NEBEAE: a similarity coefficient of $\rho = 0.01$ and an entropy weight of $\lambda = 0$. Additionally, the VCA technique was used to initialize the end-members matrix [136]. Once the characteristic end-members were obtained, they were utilized to estimate the corresponding abundance maps using NEBEAE with $\rho = 0$ and $\lambda = 0.2$, while maintaining fixed the end-members matrix. This estimation process was performed on the HS image targeted for classification. Once the abundance maps were calculated, the classification maps were generated based on the maximum abundance criterion. In the case of having multiple end-members for a class, the abundances of the final members belonging to this class were summed, and then the pixel label was based on the maximum contribution of abundance sums. Next, the results were evaluated considering the pixels labeled by the clinical expert as the ground-truth, using the following metrics: accuracy, sensitivity, specificity, and F1-score.

Table 5.1: End-members per class in each HS image.

	End-members Per Patient			
Class	P008-01	P012-01	P015-01	P020-01
NT	1	1	1	2
TT	2	1	1	2
HT	2	1	1	2
BG	4	4	4	4

5.2.1 Calibration Results

As an initial evaluation, the labeled end-members in the database were compared after the processing stage, contrasting both calibration methods and separating them by class. Figure 5.9 presents this comparison, where except for BG, each calibration provided a different mean spectral signature. Furthermore, when analyzing the variability among the end-members, the PC reduces the variability, especially in TT and HT. This property can be observed by a larger red shaded area compared to the green area, which corresponds to the STDs.

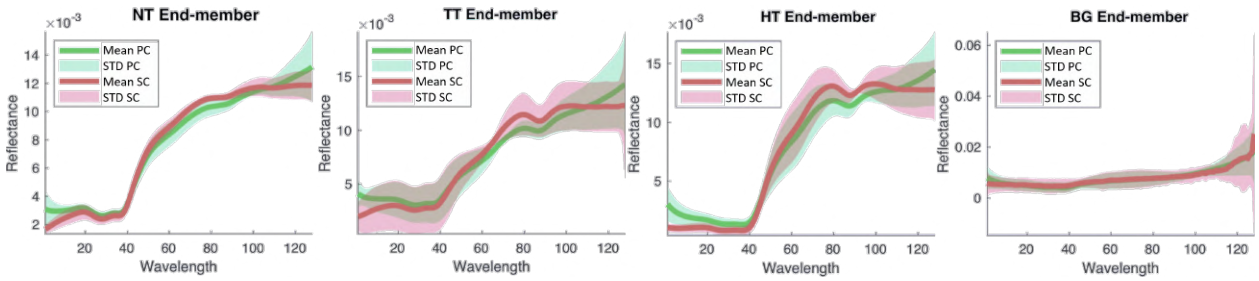


Figure 5.9: Mean and STD of the labeled end-members of each class in the database.

5.2.2 Classification Results

The resulting classification maps are presented in Figures 5.10 and 5.12. It can be observed that there is a reduction in false positives for the TT class, both in inter- and intra-patient scenarios, when using the PC approach. Additionally, this approach also presented better identification of the rubber ring markers. An example of this phenomenon is shown in Figure 5.12A for the intra-patient strategy, as well as in Figure 5.12D. Regardless of the calibration method, the labeled maps are more consistent in the intra-patient approach, as the labeled information of each patient is used for their own estimation. In contrast, in the inter-patient case, there are multiple classification errors, especially in the NT and TT classes. However, there is a better agreement with the ground-truth compared to the PC in both classification approaches.

On the other hand, when studying the quantitative results presented in Figures 5.11 and 5.13, it is possible to verify the previous observations. As seen in the classification maps of Figures 5.10 and 5.12, there is a great similarity in the results of both calibrations, which is corroborated by the performance metrics. However, it is important to clarify that in the intra-patient case, the SC produces better results, but they are quite similar to each other. On the other hand, in the inter-patient approach, the best results are obtained using the PC with a significant advantage in most of the studied metrics. These results are important because in a more realistic scenario (such as in the inter-patient case), the PC in Section 5.10 improves the performance of classification methods and spectral analysis.

These results demonstrate a significant improvement when using the PC, which can be attributed to the fact that it preserves the nonlinear relationships of the analyzed HS image and adjusts the spectral signatures to uniform conditions. In contrast, the SC, as the pixel values are not constrained, may introduce changes in the reflectance signatures, including their shape. Despite the improved performance of the PC, challenges still persist in terms of calibration, particularly in the geometric aspect. This phenomenon is because the white reference used is a flat structure, in contrast to the three-dimensional structure of the scene being analyzed.

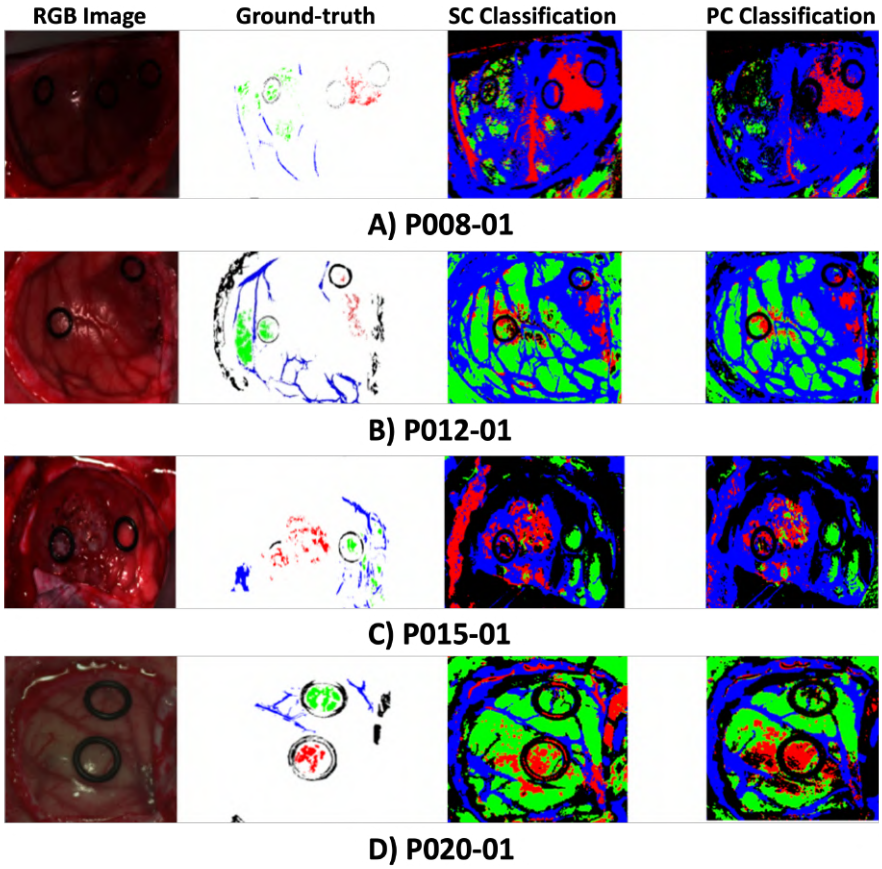


Figure 5.10: Inter-patient classification performance results using HS images: A) P008-01, B) P012-01, C) P015-01, and D) P020-01. The synthetic RGB image serves as a reference in the first column, while the ground-truth maps are shown in the second column. The SC labels maps are displayed in the third column, and the PC results are presented in the last column. The colormap used is as follows: NT (green), TT (red), HT (blue), and BG (black), with non-labeled data represented by white pixels.

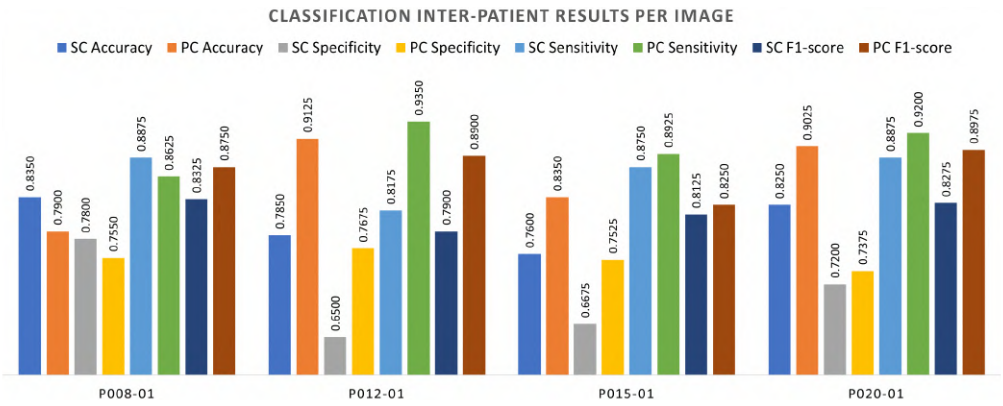


Figure 5.11: Classification performance metrics for each HS image within the inter-patient approach.

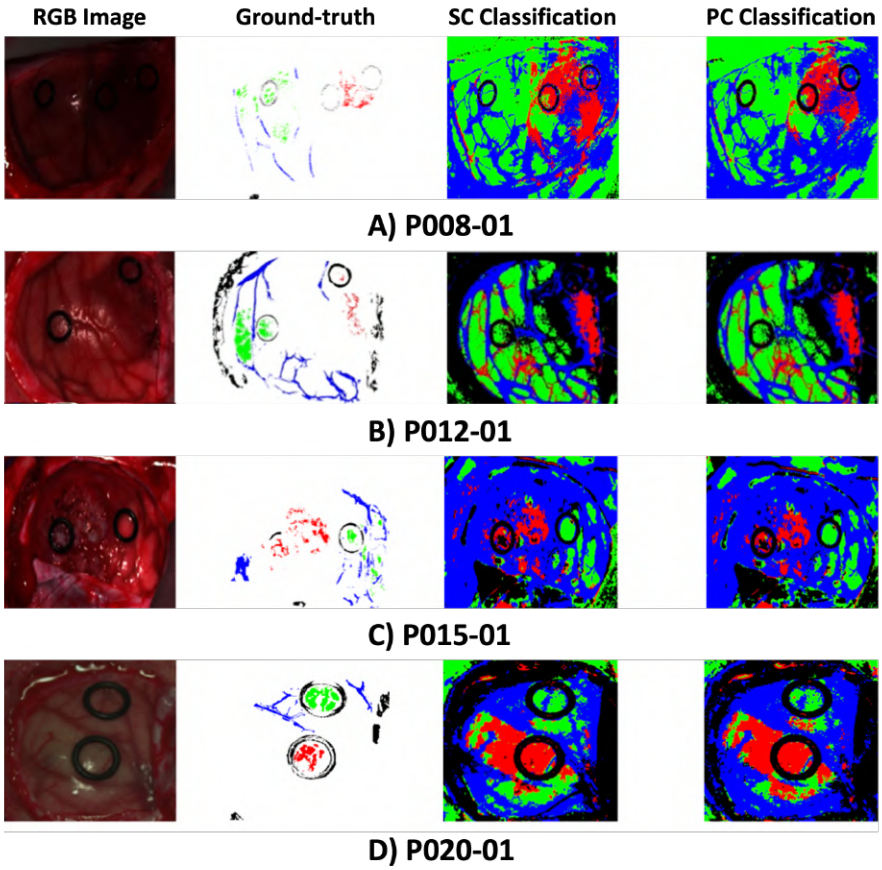


Figure 5.12: Intra-patient classification performance results using HS images: A) P008-01, B) P012-01, C) P015-01, and D) P020-01. The synthetic RGB image serves as a reference in the first column, while the ground-truth maps are shown in the second column. The SC labels maps are displayed in the third column, and the PC results are presented in the last column. The colormap used is as follows: NT (green), TT (red), HT (blue), and BG (black), with non-labeled data represented by white pixels.

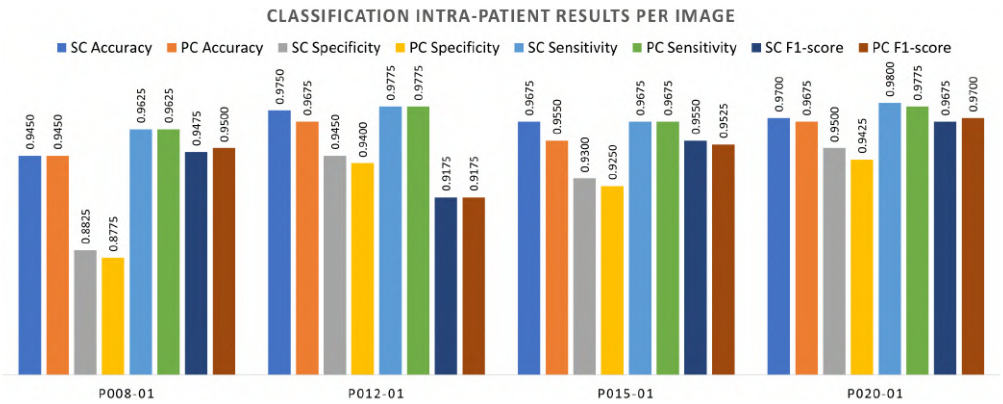


Figure 5.13: Classification performance metrics for each HS image within the intra-patient approach.

Chapter 6

Advancing Hyperspectral Brain Tissue Classification with Hybrid Schemes of Linear Unmixing and Artificial Intelligence

Unmixing methods have gained recognition in the literature as a valuable tool for HS image processing. However, in recent years, the advancement of computational architectures has paved the way for the integration of DL methods into computer vision tasks, including HS image processing [189, 190]. The combination of DL and HSI has been successfully employed for feature extraction and tissue sample classification in various studies [1, 2, 191, 192, 193, 194]. These examples exemplify the wide range of applications and opportunities for leveraging DL techniques to effectively analyze different biological samples. It is important to highlight that this line of research is relatively new [189], given that the utilization of HSI in the biomedical field is still in its experimental stages. Despite the advancements achieved by DL, these algorithms face certain challenges, such as the need for substantial amounts of training data, computational complexity, and intricate hyperparameter tuning [190].

Numerous research teams have recognized the potential challenges in DL architectures and have dedicated their efforts to addressing these issues. They have identified that the combined utilization of DL and SU can offer potential solutions for improving HS image classification performance and overcoming certain limitations [152, 153, 154]. These methodologies propose a two-step approach: first, the unmixing of HS images into abundance maps and end-members, and second, the utilization of these outputs as feature inputs for DL architectures. This integration presents two significant advantages. Firstly, it reduces the input dimensionality for DL models, enhancing computational efficiency . Secondly, it enables the development of simple DL classifiers, making them more accessible for implementation on high-performance computing platforms and reducing overall execution time. Moreover, this combined framework demonstrates the potential to improve classification accuracy by mitigating the risk of overfitting [152, 153].

This chapter introduces two methodologies that address the combination of SU and DL approaches, also known as hybrid methods. The first methodology focuses on estimating abundances, which are then

used as features for classification by various ML models. In contrast, the second approach takes a different perspective by utilizing the end-members obtained through unmixing as features for classification using an ensemble of NNs. Furthermore, the chapter concludes with a concept test involving a vision transformer classifier applied to HS images, aiming to leverage the benefits of attention-based algorithms.

6.1 A Hybrid Approach to the Hyperspectral Classification of In-vivo Brain Tissue: Linear Unmixing with Spatial Coherence and Machine Learning

The hybrid classification methodology proposed in this Section comprises two mainly stages. The initial phase involves a dimensionality reduction step using a linear unmixing (LU) approach, which aims to capture the essential features of the HS image. Subsequently, in the second stage, a classifier is employed, leveraging well-established techniques in ML such as NN, SVM, and RF. The overall process of the hybrid classification methodology is illustrated in Figure 6.1, with the LU stage performed first, followed by the application of the ML classification approach. This methodology was designed for the processing and classification of the in-vivo brain tissue database described in Section 3.3.2.2, aiming to identify the four specific classes (NT, TT, HT, and GB) within this database.

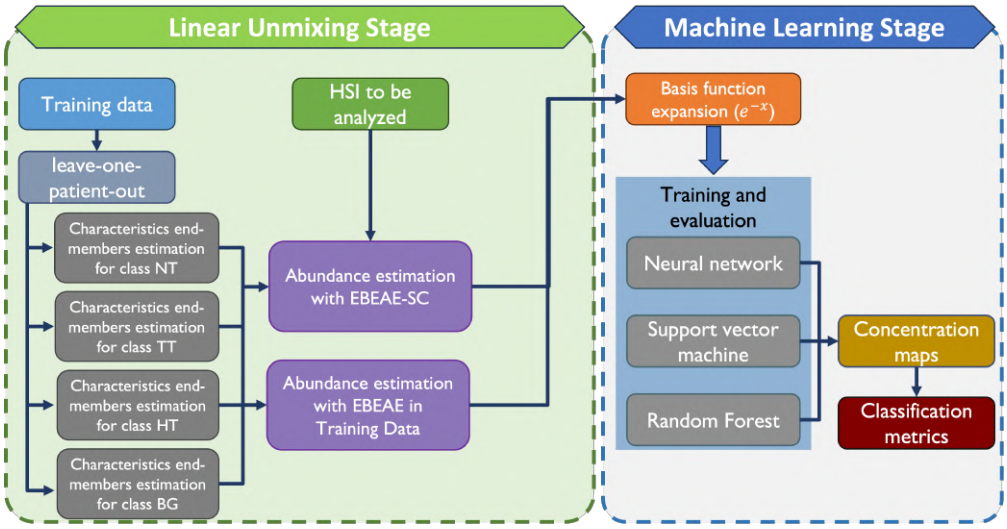


Figure 6.1: Diagram illustrating the block structure of the hybrid classification methodology.

The LU stage of the proposed methodology takes as input the HS image to be analyzed and the expert-labeled data. Once this information is defined, a leave-one-patient-out cross-validation methodology or inter-patient approach is implemented. The training dataset is then divided into four subsets representing the classes of interest. These subsets are subjected to LU using the EBEAE-SC algorithm, specifically the standard formulation without spatial coherence ($\lambda = 0$, $\tau = 0$, and $\nu = 0$), i.e., the EBEAE algorithm [84]. The goal is to obtain representative characteristic end-members ($\mathcal{P}_{NT}$, $\mathcal{P}_{TT}$, $\mathcal{P}_{HT}$, and $\mathcal{P}_{BG}$). Spatial

coherence is not utilized in this stage, as not all labeled pixels in the training dataset incorporate spatial information from the image.

The estimation of representative end-members is based on [84], defining two end-members for the NT class (card($\mathcal{P}_{NT}$) = 2), three for TT (card($\mathcal{P}_{TT}$) = 3), two for HT (card($\mathcal{P}_{HT}$) = 2), and three for GB (card($\mathcal{P}_{BG}$) = 3) for all HS images. Hyperparameters for EBEAE-SC were set manually to ensure the best representation of each class, while maintaining suitable dissimilarity among the end-members of the interclass, as stated in [33]. The selection of the similarity weight ρ varies for each class to account for the variability of the data. For NT, a weight of 0.05 ($\rho_{NT} = 0.05$) is used, while for TT, a weight of 0.1 ($\rho_{TT} = 0.1$) is chosen. HT is assigned a weight of 0 ($\rho_{HT} = 0$), and GB has a weight of 0.01 ($\rho_{GB} = 0.01$). The resulting end-members are then combined into the set $\mathcal{P}$, following this approach:

$$\mathcal{P} = \mathcal{P}_{NT} \cup \mathcal{P}_{TT} \cup \mathcal{P}_{HT} \cup \mathcal{P}_{GB}. \tag{6.1}$$

After obtaining the set of ten characteristic end-members $\mathcal{P}$ (card($\mathcal{P}$) = 10), the next step is to estimate the abundances in the training data and the HS image using these end-members. This estimation process generates the set of features that will be fed into the classifier. For the training data, the abundances are estimated using EBEAE-SC without considering spatial coherence ($\lambda = 0$, $\tau = 0$ and $\nu = 0$), and these abundances serve as the training data for the different ML models. The reason for this is that the data used to train the model only includes individual pixels, without any information about the neighboring pixels surrounding it. This can result in problems when trying to consider spatial coherence in the estimation. To address each HS image, the EBEAE-SC algorithm is used to reduce variability and granularity in the resulting abundance maps. The hyperparameters for EBEAE-SC were empirically selected to achieve smooth abundances without abrupt changes. An exhaustive search is performed for the hyperparameters (λ, τ and ν) in the range of $[1 \times 10^{-8}, 1.0]$ to minimize the error, and the optimal values are determined to be $\lambda = 1 \times 10^{-5}$, $\tau = 1 \times 10^{-4}$, and $\nu = 0.1$ for all the HS images.

In the ML phase, depicted in Figure 6.1, the first step involves a procedure called base function expansion. In DL, this step entails transforming the feature space to achieve a more distinct separation in the database, thereby enhancing the classifier's performance. Various tests were conducted to determine the most suitable expansion of the base function. These tests included squared, polynomial, sinusoidal, and exponential patterns. The exponential function with a negative argument (e^{-X}) was found to give the best performance and significantly improve the accuracy of the hybrid classifier.

The objective of the ML stage is to categorize the extracted features into four classes (NT, TT, HT, and GB). To achieve this objective, the suitability of NN, SVM, and RF classifiers was investigated, as mentioned above. Initially, the NN architecture consisted of four layers: an input layer with 10 neurons, a second layer with 50 neurons, a third layer with 100 neurons, and an output layer with four neurons. Sigmoid activation functions were employed in the intermediate layers, while a softmax activation function was used in the output layer for the multiclass problem. In addition, batch normalization with a momentum of 0.8

was applied to the training data. The Adam optimization algorithm with a learning rate of 0.0002 was used and each batch contained 64 elements. The training process consisted of 50 epochs. These parameter values were selected through extensive evaluations to optimize classification performance. Different combinations of epoch numbers, batch sizes, and layer configurations (ranging from one to ten layers) were examined, along with varying the number of neurons within each layer (from 10 to 500).

To determine the properties of the SVM classifier, various experiments were conducted using different kernels such as linear, polynomial, and RBF. The optimal configuration for the SVM classifier was achieved by employing a linear kernel along with the Crammer-Singer strategy for multiclass classification. Similar to the NN, the SVM hyperparameters were carefully selected to achieve optimal performance on the training dataset. The regularization hyperparameter, denoted as C, was tested with values of $C \in \{10, 50, 100, 200\}$; nevertheless, no substantial enhancements in training accuracy were observed.

Similar to the previous classifiers, the RF configuration was modified during the training stage to optimize the classification performance for the target dataset. In order to achieve this objective, several parameters were adjusted, including the number of trees (ranging from 10 to 500), the minimum number of samples necessary to create a split (10, 20, 30, or 40), the minimum number of samples in a leaf (1, 2, or 3), and the criterion for determining the maximum number of features considered for the best split ('log2' or 'sqrt'). As a result, the RF model was constructed with 200 trees, utilizing the Gini function as the quality criterion for generating splits. The minimum number of samples per leaf was set to two, the minimum number of samples required to generate a split was set to 20, and the 'log2' function was employed to determine the optimal split.

6.1.1 Experimental Results and Discussion

The performance of the proposed hybrid classification methodologies is evaluated in this section with respect to the state-of-the-art methods described in [5], namely the SVM-based approach and the LU-based scheme (methodology B with correlation metrics and Mahalanobis distance). All implementations were performed using the Python programming language with the use of libraries such as NumPy, SciPy, matplotlib, Scikit Learn, and TensorFlow. The experiments were carried out on a computer equipped with an Intel® CoreTM i7 quad-core processor operating at 4.2 GHz and 16 GB of RAM. The evaluation involved testing six HS images as described in Subsection 3.3.2.2 (P008-01, P008-02, P012-01, P012-02, P015-01, and P020-01) using leave-one-patient-out cross-validation, and the classification metrics were calculated based on the pixels labeled by clinical experts. The evaluation of the classification results for both the proposed methodologies and the state-of-the-art methods used the following metrics: precision, sensitivity, specificity, F1-score, and normalized MCC.

6.1.1.1 Evaluation of the Hybrid Classification Methodology

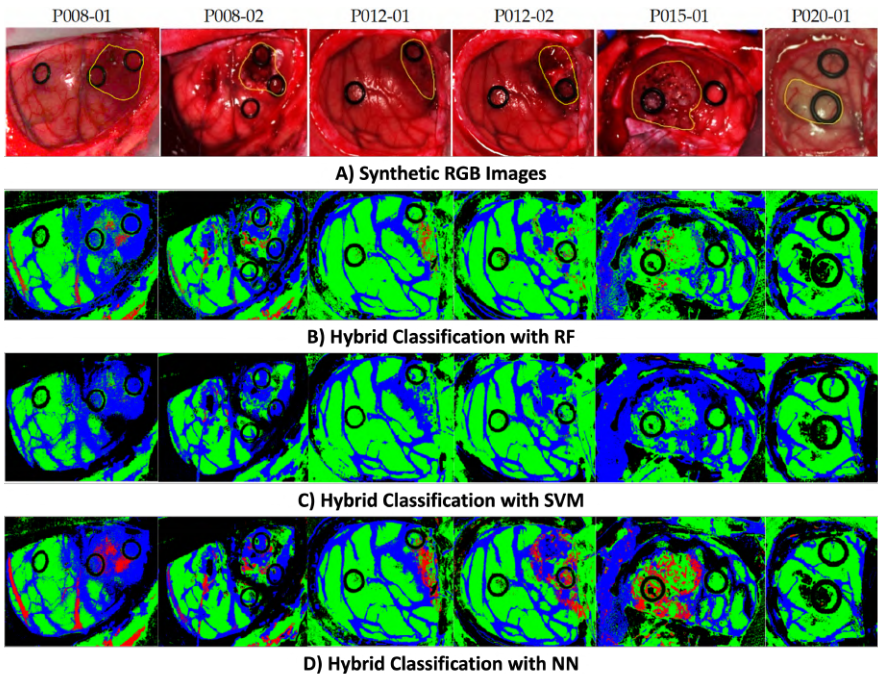


Figure 6.2: Comparison of the classification maps generated by the hybrid classification methodology, where the NT, TT, HT, and BG classes are represented by green, red, blue, and black regions, respectively. The figures include: A) Synthetic RGB images with the tumor area highlighted by a yellow line, B) Classification map produced by the RF hybrid method, C) Classification map produced by the SVM hybrid method, and D) Classification map produced by the NN hybrid method.

The hybrid classification methodology utilizing ML tools produced labeled maps, as illustrated in Figure 6.2. Figure 6.2A show the synthetic RGB image with the manual segmentation performed by a clinical expert, highlighting the approximate tumor-affected area with a yellow line. Figures 6.2B - 6.2D display the labeled maps resulting from the hybrid classification method employing RF, SVM, and NN, respectively. In these maps, the NT class is represented by green regions, the TT class by red, the HT class by blue, and the BG class by black. The labeled maps demonstrate similarities primarily in the NT and BG classes. However, errors in classification are observed, particularly in the TT class. Among the hybrid approaches, the NN method exhibited greater accuracy in identifying the tumor regions as delineated by the expert, while the SVM method achieved lower performance in identifying this class.

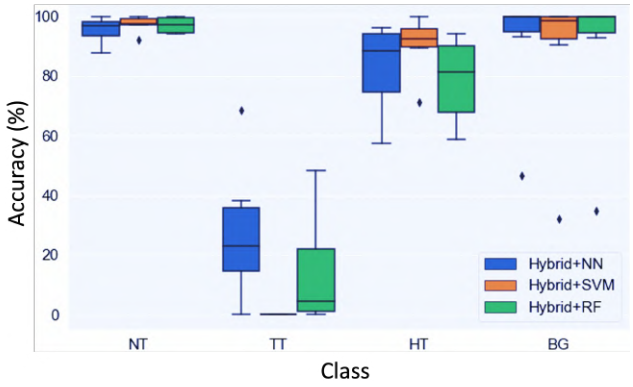


Figure 6.3: Class-wise overall accuracy generated by the hybrid methodology employing NN, SVM, and RF models.

The accuracy per class was analyzed using the labeled dataset, and the findings are visually represented in boxplot diagrams shown in Figure 6.3. The results demonstrate that the hybrid classification methodology effectively identifies the NT and BG classes, while also achieving good accuracy in the HT class, with median accuracies exceeding 80%. However, the performance in identifying the TT class is notably low, particularly in the SVM hybrid method, where the correct identification of tumor tissue pixels is practically non-existent. On the other hand, the NN approach enhances the estimation of the TT class, yielding a higher median accuracy of over 20%, although these values varied widely between 0% and 70%.

6.1.1.2 Comparison with Other Related Works

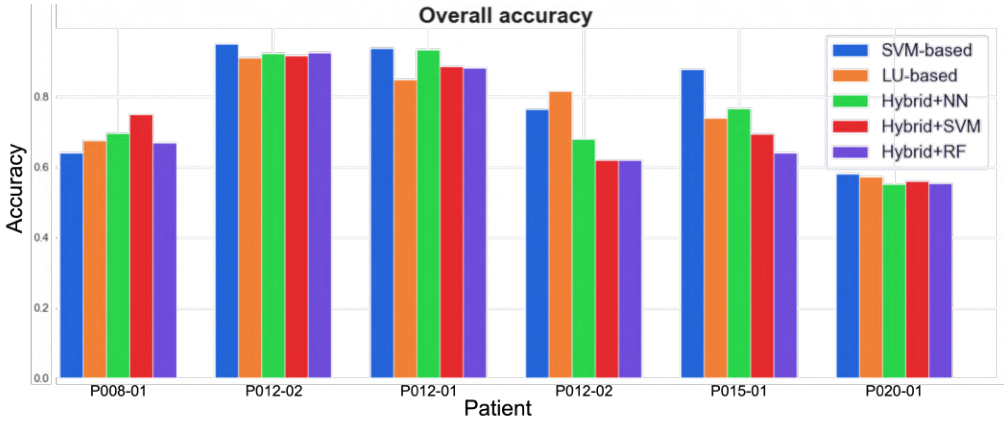


Figure 6.4: Comparison of the overall accuracy achieved per HS image using the hybrid classification methodology and its comparison to the accuracy obtained by the NN, SVM, and RF methods, along with the SVM- and LU-based approaches.

The results obtained from the SVM-based and LU-based schemes in [5] were utilized to compare with the proposed hybrid classification methodology. Figure 6.4 presents the overall accuracy results per HS image for each method. Generally, a similar performance was observed among the evaluated methods. The SVM-based approach presented slightly higher accuracy in half of the HS images, with an average accuracy of 79.2± 14%, while the LU-based method achieved an accuracy of 76.0±11%. On the other hand, analyzing the overall accuracy of the hybrid classification methodology, a similar trend was observed: 75.9±13% for NN, 73.8±13% for SVM, and 71.5±14% for RF. These results may appear somewhat counterintuitive when compared to the accuracy per class, primarily due to the varying number of labeled pixels per class. Specifically, the TT class had the lowest number of labeled pixels. In particular, the NN hybrid method demonstrated accuracy results that were nearly equivalent to the LU-based approach, with a difference of approximately 3% from the SVM-based approach.

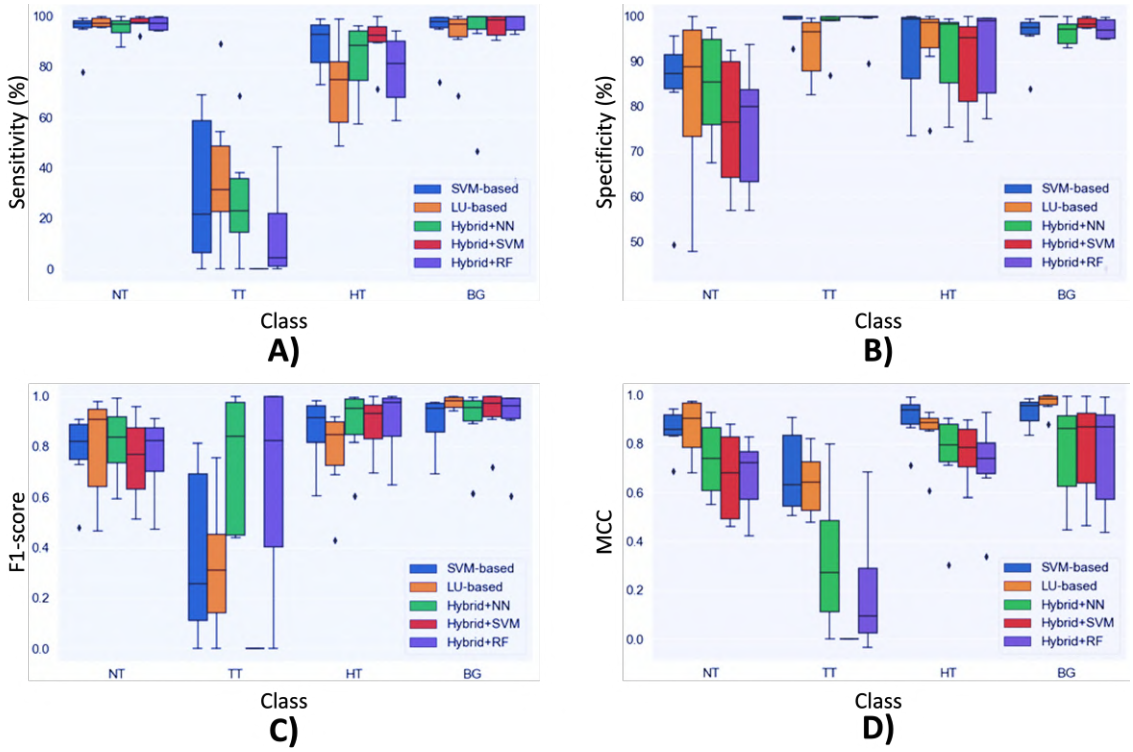


Figure 6.5: Comparison of the classification performance achieved by the NN, SVM, and RF hybrid methods, as well as the SVM-based and LU-based approaches, utilizing leave-one-patient-out cross-validation. The evaluation metrics presented are: A) sensitivity, B) specificity, C) F1-score, and D) normalized Matthews correlation coefficient.

The classification metrics in Equations (11) to (15) were presented in Figure 6.5, considering the labeled pixels. Sensitivity analysis (Figure 6.5A) revealed a similar trend for the NT and BG classes across all five

methodologies, with median values above 95%. However, greater variability was observed in the TT and HT classes. Among the classifiers, the LU-based scheme exhibited the highest sensitivity for the TT class, with a median of 31.2%, followed by the NN hybrid method (22.3%) performing better than the SVM-based approach (21.6%). The RF and SVM hybrid methods had sensitivities below 5%. In the HT class, the SVM-based, NN, and SVM hybrid methods displayed sensitivities of approximately 90%, with the latter exhibiting the least variability. On the other hand, the LU-based and RF hybrid methods had medians lower than 82%.

The specificity information produced by the classifiers is depicted in Figure 6.5B using boxplots. Regarding the TT, HT, and BG classes, all methodologies presented similar behavior with medians exceeding 95%. However, the LU-based scheme displays slightly higher variability in the TT class with respect to the comparison methods. In contrast, for the NT class, there is a high variability, where the LU-based scheme exhibiting the highest specificity with a median of 88.8%. It is followed by the SVM-based, NN, RF, and SVM hybrid methods with specificities of 87.3%, 84.4%, 79.9%, and 76.6%, respectively.

Figure 6.5C illustrates the F1-score results, which present a similar behavior among the five methods in the NT, HT, and BG classes, akin to the specificity results. However, the most significant outcome is observed in the TT class. Both the NN and RF hybrid methods achieved F1-scores of 84.0% and 82.4% respectively, representing a substantial increase compared to the SVM-based and LU-based schemes with F1-scores of 21.6% and 31.3% respectively. This improvement is particularly relevant in applications where accurately identifying regions belonging to the TT and NT classes is crucial. The MCC index results are presented in Figure 6.5D, where a similar level of variability is shown among the comparison methods. However, in contrast to the previous metrics, both the SVM-based and LU-based schemes outperformed the hybrid methodologies, exhibiting approximately 5% higher medians in the NT, HT, and BG classes. For the TT class, the SVM-based and LU-based schemes achieved medians of 63.1% and 64.2% respectively. Conversely, the hybrid methods showed a significant decrease in performance, with medians below 28% across all cases. The NN hybrid method demonstrated a moderately acceptable range of results. These findings align with the overall trend observed in Figure 6.3, where the accuracy in estimating the TT class was low. Nonetheless, the NN hybrid method exhibited the best performance specifically for this class. Figure 6.6 presents the classification maps obtained from the state-of-the-art methodologies under comparison. Specifically, Figures 6.6B to 6.6D represent the NN hybrid, LU-based, and SVM-based methods, respectively. In this case, the results by the NN hybrid method are highlighted, given its superior performance both qualitatively and quantitatively when compared to the SVM and RF variants. Overall, the classification maps generated by the three methods exhibited high similarity. However, the NN hybrid method produced maps with a reduced level of granularity compared to the SVM-based results, though not reaching the same level of uniformity as observed in the LU-based results.

It should be noted that the NN hybrid method achieved lower classification errors in the TT class compared to the other methods. Nonetheless, certain classification errors were observed between the BG

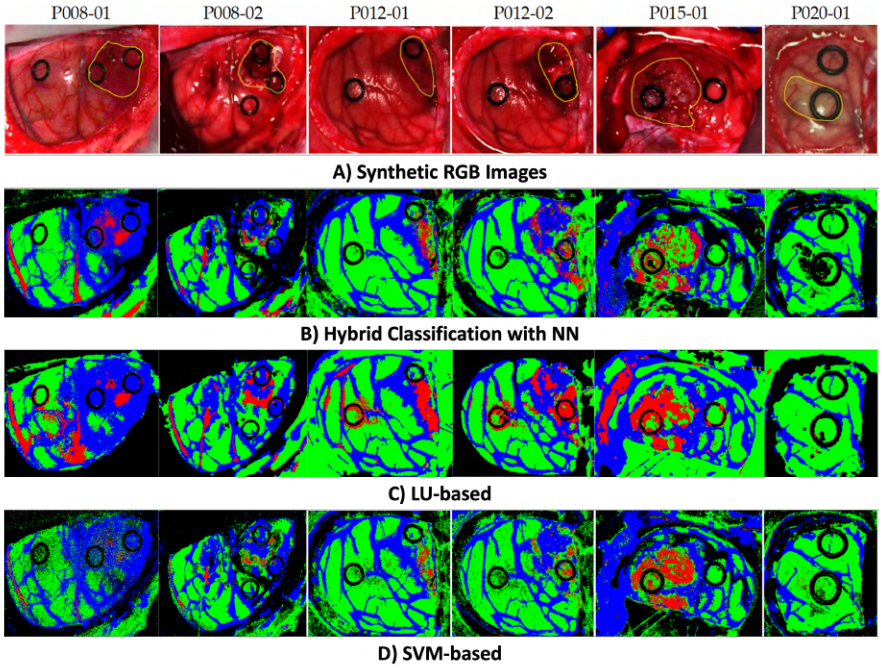


Figure 6.6: A comparison of classification maps produced by three different methods: the NN hybrid method, the SVM-based approach, and the LU-based approach. The regions labeled in green, red, blue, and black represent the NT, TT, HT, and BG classes, respectively. The images displayed include: A) synthetic RGB images with the tumor area outlined in yellow, B) the NN hybrid method, C) the LU-based approach, and D) the SVM-based approach.

and TT classes (P020-01), HT class (P015-01), and NT class (P008-01 and P008-02). The LU-based scheme showed less vulnerability to this situation. However, it does exhibit classification errors in the HT class by misclassifying regions belonging to NT and TT as HT. Moreover, the SVM-based and NN hybrid methods also presented this type of errors, albeit to a lesser extent, with a greater prevalence in subjects P008-01 and P015-01.

6.2 Hybrid Brain Tumor Classification Scheme of Histopathology Hyperspectral Images Using Linear Unmixing and Deep Learning

The second proposed hybrid approach is depicted in the diagram of Figure 6.7. It consists of seven steps and is specifically designed for the histopathology database described in Subsection 3.3.2.1. The methodology begins by preprocessing and partitioning the raw HS image (Figure 6.7A and B), as discussed earlier. Then, the preprocessed HS image undergoes the SU stage (Figure 6.7C), where abundance maps and characteristic end-members are derived. This stage is described in detail in the subsequent section. Notably, the SU stage enables the identification of regions on the histological slide that are devoid of tissue and therefore irrelevant for tissue classification. These regions, along with their corresponding abundance maps

and end-members, were eliminated (Figure 6.7D). The remaining end-members are then concatenated into a single vector (Figure 6.7E). These resulting elements serve as the input for training, validating, and testing both level-zero and level-one NNs using a transfer learning approach (Figure 6.7F), which is thoroughly explained in Section 6.2.2. Finally, the performance of the level-one NN is evaluated using classification metrics (Figure 6.7G).

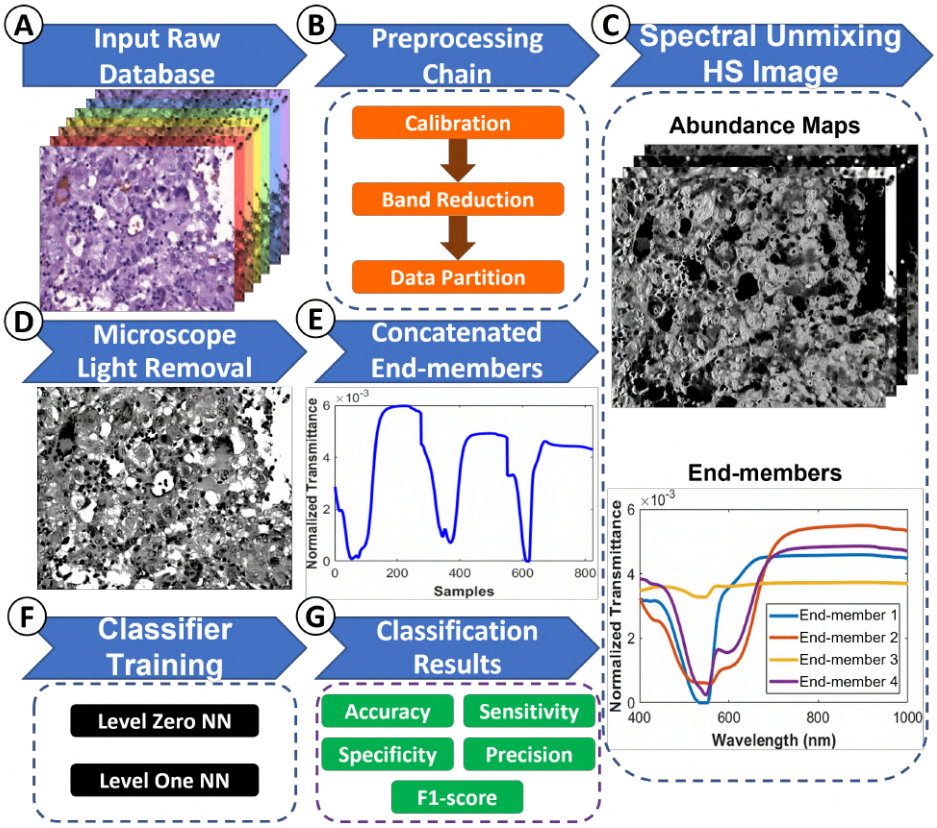


Figure 6.7: The general schematic of the hybrid scheme proposed for the processing and classification of histopatology HS images.

6.2.1 Spectral Unmixing

The SU stage serves as an intermediary step between the preprocessing phase and the DL classifier. Its purpose is to estimate abundance maps and characteristic end-members from the HS image. The end-members represent distinct components in the HS image, while the abundance maps indicate the proportion of each component present at each pixel. To perform the SU stage, the EBEAE algorithm was employed. The initialization method for the end-members was the N-FINDR scheme [134]. The hyperparameters of EBEAE were set as follows: a similarity weight (ρ) of 1, an entropy weight (μ) of 0, an error threshold (ϵ) of 1×10^{-6}, and a maximum number of iterations ($maxiter$) of 10. The LU process assumed the presence

of four characteristic components ($N = 4$). These hyperparameters were manually tuned following the guidelines outlined in [5] and [84].

6.2.1.1 End-members Identification

To address the issue of histological samples containing areas without tissue and direct transmission of light from the microscope, two possible approaches can be considered. The first approach involves segmenting the HS image to extract the regions with direct light transmission. Alternatively, the second approach involves performing SU on the entire HS image, followed by the removal of the end-member representing regions with a flat spectral signature. In this study, the second approach to simplify the preprocessing stage was adopted. This phenomenon is evident in Figure 6.7C, particularly in the spectral signature of end-member 3, which exhibits uniform information across nearly all bands due to the sum-to-one restriction imposed by the EBEAE algorithm. Considering that HS images typically contain 275 bands, this end-member can be identified by:

$$\mathbf{p}^* = \underset{\forall \tilde{\mathbf{p}} \in \mathcal{P}}{\arg\min} \left\| \tilde{\mathbf{p}} - \frac{1}{275}\mathbf{1} \right\|, \tag{6.2}$$

where $\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_N\}$ represents the set of estimated end-members by EBEAE. Consequently, the end-member $\mathbf{p}^*$ corresponds to the spectral signature of regions within the HS image that lack cellular tissue information. As a consequence, this particular end-member is discarded during the DL classification process. Furthermore, an ordering stage is implemented for the estimated end-members within each HS image. This step involves selecting a reference HS image randomly for each fold and employing its end-member order as a guide for the remaining images, using the Euclidean distance as a measure of similarity. Subsequently, the ordered end-members are concatenated to form an augmented vector that contains the spectral information, as illustrated in Figure 6.7E. Consequently, each HS image, with dimensions of $1004 \times 800 \times 275$, is represented by a single vector of dimension $3 \times 275 = 825$. This process of feature extraction is based on the assumption that the concatenated spectral components from tumorous and non-tumorous tissue samples contain distinct information that is relevant for classification. Another key consideration in choosing only end-members and discarding the abundance maps is that only one global label is available for each image. Consequently, the spatial properties of each class are unknown, posing challenges in their utilization. Using only end-members yielded excellent error metrics with low complexity. It is worth noting that in our initial experiments, classifiers trained with abundance maps produced significantly lower results compared to models trained with end-members.

6.2.1.2 Data Augmentation

The size of the database directly influences the outcome of the classification process. In general, enlarging the database can enhance the performance of a NN up to a certain threshold, beyond which the improvements become negligible. By incorporating a larger database, the NN can access a broader range

of examples, facilitating the learning of robust features and promoting better generalization to unseen data [195, 196]. Nonetheless, the quality of the data is as crucial as its quantity. A sizable database with low-quality data may not yield superior results compared to a smaller database comprising high-quality data. Hence, it is essential to ensure that the database is both representative and accurately labeled to avoid biases and errors during the classification process [196]. The SU methodology yielded 494 feature vectors, each corresponding to an HS image from the database under study. These vectors were utilized for training, validating, and testing the DL classifier. Figure 3.5 illustrates the distribution of the four folds for the learning process, with approximately 310 HS images allocated for training (Fold 1: 325, Fold 2: 327, Fold 3: 263, and Fold 4: 325), 50 for validation (Fold 1: 60, Fold 2: 43, Fold 3: 43, and Fold 4: 53), and 120 for testing (Fold 1: 109, Fold 2: 124, Fold 3: 146, and Fold 4: 115). However, the limited size of this database poses a challenge for effective learning by the DL classifier. To address this limitation, data augmentation was employed to augment the database, ensuring a more successful learning process. To achieve this, a reformulation of the SU stage was performed. The spatial pixels of each HS image were randomly divided into ten subsets, and SU was applied to each subset to extract their characteristic spectral signatures. These signatures were then concatenated, resulting in feature vectors of dimension 825. Consequently, for each HS image, ten feature vectors were generated to characterize it. Furthermore, up to 20 random subsets were evaluated for each HS image, but no improvements were observed during the initial classification tests. As a result, the data augmentation step produced a total of 4940 vectors, representing the entire database, through the concatenation of spectral signatures. These augmented feature vectors were subsequently employed for training, validating, and testing the DL classifier. This data augmentation step proved instrumental in providing more consistent training with reduced variation when classifying HS images.

6.2.2 Deep Learning Classifier

To enhance the robustness and accuracy of the proposed methodology, a DL classification stage was implemented using an ensemble of NNs. This approach aimed to decrease variability and improve overall performance through transfer learning [197, 198]. The choice of NNs for the ensemble was based on their adaptability and robustness in learning characteristic patterns of each class in the training data, as compared to SVM and RF, as observed in a previous study [199]. NNs are outstanding at capturing complex nonlinear relationships between input data and output labels, making them well-suited for the classification of histological brain samples. The ensemble strategy employed in our methodology is reminiscent of the RF approach, where multiple decision trees are combined using a voting scheme to mitigate the impact of overfitting.

To construct the ensemble, three identical level-zero NNs were combined. This choice was based on experimental findings that demonstrated stable classification performance with three level-zero NNs, as increasing the number of NNs did not yield significant improvements. The architecture of each level-zero

NN, illustrated in Figure 6.8, and it consisted of an input layer with 825 features and four hidden layers with 1000, 200, 100, and 20 neurons, respectively. This hierarchical structure facilitated effective learning within the network. The determination of the number of layers and neurons per layer was also conducted experimentally, prioritizing simplicity while maximizing performance. Each internal layer in the level-zero NNs utilized the LeakyReLu activation function with a slope of 0.1 for negative values. Additionally, L_2 regularization of 0.0002 and a dropout stage of 50% were included to mitigate overfitting during the learning process. The output layer employed a sigmoid activation function, as the classification task was binary. The binary crossentropy cost function was employed for weight adjustment [200]. The uniform variance scaling method was utilized for weight initialization in all layers of both the level-zero NNs and the ensemble, aiming to reduce result variability [201]. During training, a minibatch size of 100 elements and a learning rate of 0.001 were utilized. The number of epochs was set to 500 for the level-zero NNs and 100 for the level-one NN. As described earlier, the input data for the level-zero NNs consisted of the training sets from each fold in Figure 3.5C, incorporating only the concatenated end-members of each HS image, resulting in a dimension of 825. Once the NNs were trained, they were saved and integrated into the ensemble network or level-one network.

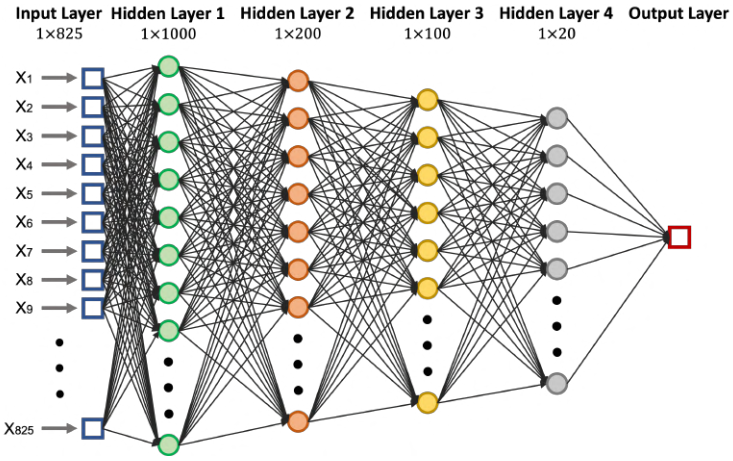


Figure 6.8: Architecture for the proposed structure of the level-zero NNs, consisting of three neural networks.

The level-one network, also known as the NNs ensemble, utilizes the three pre-trained level-zero networks in a transfer learning approach, where the weights of the level-zero NNs are frozen during the training of the level-one network. Figure 6.9 illustrates the general architecture of the ensemble, where the outputs of the level-zero NNs are discarded after the four hidden layers (illustrated in Figure 6.8), following the transfer learning methodology. Consequently, the outputs of the three level-zero networks form a vector of 60 input elements in the ensemble. This vector is then processed by a NN consisting of three hidden layers with 300, 200, and 32 neurons, respectively. Each hidden layer utilizes the LeakyReLu activation function

with a slope of 0.1 for negative values. L_2 regularization of 0.0002 and a dropout stage of 50% are applied to prevent overfitting. Similar to the level-zero networks, the binary crossentropy cost function is used to adjust the weights of the level-one network. To optimize performance, the architecture of the level-one NNs is kept simple. The initial hidden layer is composed of 300 neurons to enhance representability and interconnection across layers, allowing effective integration of data from the three level-zero NNs. This selection ensures optimal information flow throughout the network. The output layer employs a sigmoid activation function, producing a bounded output between zero and one, which is rounded to generate the final classification result. The ensemble outputs a binary value, where zero represents non-tumor class and one represents tumor. This output is verified against the labels provided by clinical experts.

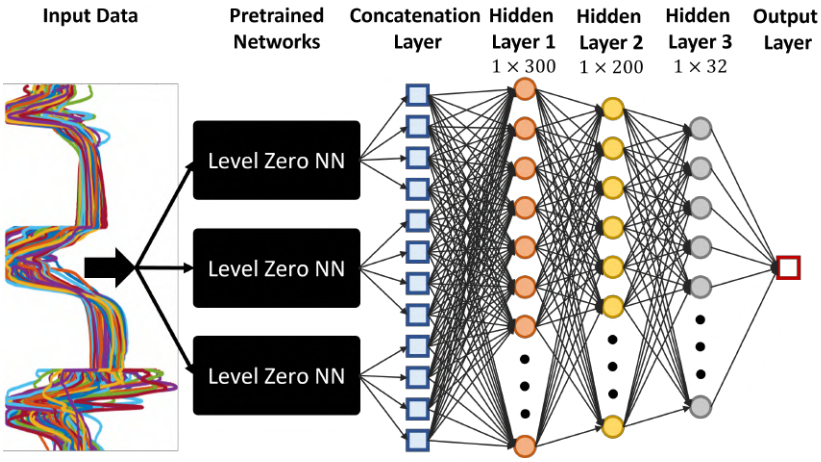


Figure 6.9: Architecture of the ensemble network, also referred to as the level-one NN.

6.2.3 Results

The results obtained by the proposed methodology are described in this section. The evaluation involved training independent classifiers for each fold in Figure 3.5C. The preprocessing and SU stages were conducted in MATLAB®, and the resulting data per patient was saved for future use. The DL classifier, on the other hand, was trained and evaluated in Python using TensorFlow modules [202]. The entire process was executed on a computer equipped with a 4.2 GHz Intel Core i7 quad-core processor and 16 GB of RAM.

For the training of the level-zero NNs, two different strategies were employed for comparison purposes. The first approach involved randomly dividing the training data into three subsets, known as *training with data division* (DD), where each subset was used for training a level-zero NN. In the second approach, all available data (AD) was used to train each NN. The training of the level-one NN was performed using the assigned data for each fold. Both strategies yielded promising results, with an average accuracy of 91% for DD training and 94% for AD. However, it is important to note that using the training data from Figure 3.5C for both the level-zero and level-one NNs could lead to overfitting, particularly in the level-zero networks

trained by AD. Therefore, the evaluation of the NNs was conducted using a combination of validation and test data to mitigate this potential condition.

6.2.3.1 Validation Results

The results obtained from the validation sets in Figure 3.5C for each fold in both training approaches are summarized in Table 6.1. Overall, both proposals exhibited a similar behavior, but the NNs trained with DD show slightly better performance across five metrics: (i) accuracy at 91%, (ii) sensitivity at 72%, (iii) specificity at 95%, (iv) precision at 80%, and (v) F1-score at 75%. On the other hand, the NNs trained with AD achieved slightly lower scores in all metrics, albeit by a small margin. In terms of STD, the NNs trained with DD presented lower variability in all metrics, with the most significant difference observed in terms of specificity. When analyzing the validation results individually, it is observed that Fold 1 performed relatively poorer in both cases, with 74% accuracy for NNs trained with DD and 57% accuracy for NNs trained with AD. In contrast, the best results were obtained in Fold 2, where both approaches reached 100% accuracy across all metrics. Fold 3 also presented high performance, with NNs trained with DD achieving 99% accuracy and NNs trained with AD 97% accuracy. This pattern suggests that Folds 2 and 3 are better suited for accurate detection of both classes, while Folds 1 and 4 reached lower classification performance.

Table 6.1: Quantitative evaluation of the validation set in each fold using the hybrid classification scheme, with the best results highlighted in bold. Two training approaches, namely DD and AD, were employed in the experiment.

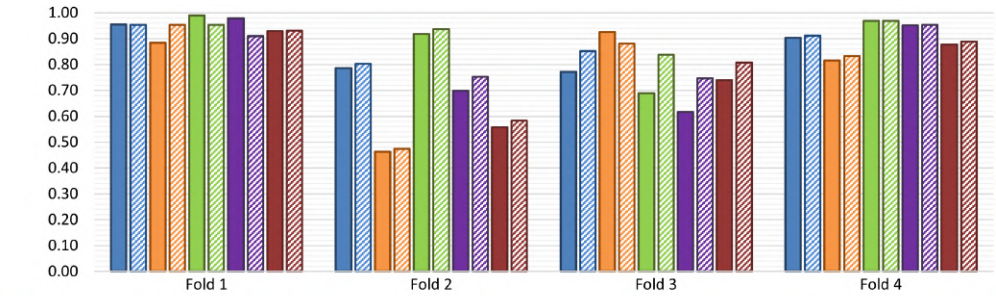
	NNs with DD					NNs with AD				
Fold	**Acc.**	**Sen.**	**Spe.**	**Prec.**	**F1**	**Acc.**	**Sen.**	**Spe.**	**Prec.**	**F1**
Fold 1	**0.74**	**0.29**	**0.85**	**0.33**	**0.31**	0.57	0.25	0.64	0.15	0.19
Fold 2	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Fold 3	**0.99**	**1.00**	**0.98**	**0.96**	**0.98**	0.97	**1.00**	0.95	0.89	0.94
Fold 4	**0.89**	**0.58**	0.98	**0.90**	**0.70**	0.87	0.47	**0.99**	**0.90**	0.62
Avg.	**0.91**	**0.72**	**0.95**	**0.80**	**0.75**	0.85	0.68	0.90	0.74	0.69
STD.	**0.12**	**0.35**	**0.07**	**0.31**	**0.32**	0.20	0.38	0.17	0.39	0.37

6.2.3.2 Test Results

The classification results for the test set for each fold in Figure 3.5C are presented in Figure 6.10. The results for NNs with DD and AD are compared in terms of various metrics. Initially, it is observed that the results obtained with both approaches are highly similar across all metrics. This similarity is further confirmed when examining the average metrics across all folds. On average, the accuracy between both approaches differed by only 3%, with NNs trained with DD achieving an accuracy of $85\% \pm 9\%$ (mean $\pm$ STD) and NNs trained with AD of $88\% \pm 7\%$.

The similarity between the approaches extends to the sensitivity and F1-score metrics as well. The NNs with DD obtained an average sensitivity of $77\% \pm 21\%$ and an average F1-score of $79\% \pm 21\%$,

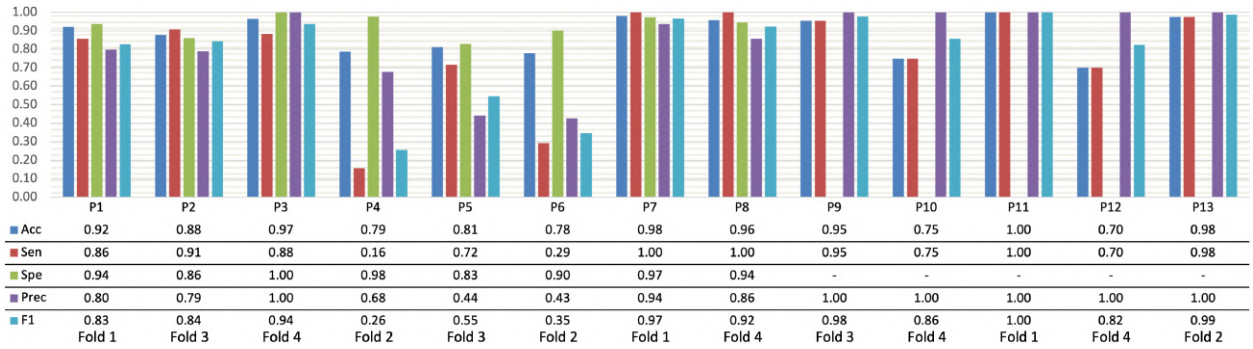
whereas the NNs trained with AD achieved an average sensitivity of $78\% \pm 17\%$ and an average F1-score of $80\% \pm 16\%$. Similarly, there is only a 3% difference in the precision metric, with NNs trained with DD achieving $81\% \pm 18\%$ and NNs trained with AD $84\% \pm 11\%$. The specificity metric also showed a 3% difference between the two approaches, with NNs with DD achieving $89\% \pm 14\%$ and NNs trained with AD $92\% \pm 6\%$. Overall, the NNs trained with AD demonstrated slightly superior performance in the proposed metrics for the test sets, as well as during the training stage.



	Fold 1	Fold 2	Fold 3	Fold 4
Acc. (DD)	0.96	0.79	0.77	0.90
Acc. (AD)	0.95	0.80	0.85	0.91
Sen. (DD)	0.88	0.46	0.93	0.81
Sen. (AD)	0.95	0.48	0.88	0.83
Spe. (DD)	0.99	0.92	0.69	0.97
Spe. (AD)	0.95	0.94	0.84	0.97
Prec. (DD)	0.98	0.70	0.62	0.95
Prec. (AD)	0.91	0.75	0.75	0.95
F1 (DD)	0.93	0.56	0.74	0.88
F1 (AD)	0.93	0.58	0.81	0.89

Figure 6.10: Quantitative results of the test set in each fold employing the hybrid classification scheme with two distinct training approaches: DD and AD.

Finally, Figure 6.11 illustrates the classification results for each patient evaluated using the NNs trained with AD. It can be observed that only four patients (P4, P6, P10, and P12) achieved an accuracy lower than 80%, while two patients (P2 and P5) obtained accuracies between 80% and 90%. The remaining patients obtained a performance greater than 90% in terms of accuracy. Similar trends are observed in other performance metrics. Upon analyzing the cases with lower performance, it was found that patients P4, P6, P10, and P12 were classified using the models trained with the data from Folds 2 and 4. However, these models demonstrated good performance in the other patients, with Fold 2 being the one with the lowest performance across all metrics. It is important to note that specificity can only be calculated for patients with both types of classes (P1 to P8), resulting in a sensitivity of 100% for patients P9 to P13. The overall high performance observed in the performance metrics suggests that our proposed methodology holds potential as a computer-aided detection system for identifying images with tumor presence.



	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13
Acc	0.92	0.88	0.97	0.79	0.81	0.78	0.98	0.96	0.95	0.75	1.00	0.70	0.98
Sen	0.86	0.91	0.88	0.16	0.72	0.29	1.00	1.00	0.95	0.75	1.00	0.70	0.98
Spe	0.94	0.86	1.00	0.98	0.83	0.90	0.97	0.94	-	-	-	-	-
Prec	0.80	0.79	1.00	0.68	0.44	0.43	0.94	0.86	1.00	1.00	1.00	1.00	1.00
F1	0.83	0.84	0.94	0.26	0.55	0.35	0.97	0.92	0.98	0.86	1.00	0.82	0.99
	Fold 1	Fold 3	Fold 4	Fold 2	Fold 3	Fold 2	Fold 1	Fold 4	Fold 3	Fold 4	Fold 1	Fold 4	Fold 2

Figure 6.11: Classification results per patient in the test set using the NN trained with AD. Patients P9 to P13 have only tumor class samples, making it impossible to calculate the specificity metric for these cases.

Table 6.2 presents the comparison of the NNs approach with AD to two state-of-the-art methodologies using the same database and artificial intelligence techniques [1, 2]. The first method is a SVM-based approach trained with a superpixel scheme [1], specifically utilizing results at the image level. The second method is a patch-based CNN [2]. Both algorithms were trained using the same database, except for patient P6, where an annotation error was discovered during the implementation of the CNN. These errors were rectified in both the SVM work with superpixels [1] and this dissertation work. Consequently, to ensure a fair comparison, patient P6 was excluded from the calculation of the overall performance metrics. The hybrid classification scheme achieved an average accuracy that was 3% higher than the CNN-based approach and 6% higher than the SVM with superpixels, while also exhibiting a reduced STD. Similar trends were observed in the specificity metric, with the proposed approach achieving $82\%\pm33\%$ compared to $78\%\pm34\%$ and $77\% \pm 16\%$ for the SVM-based and CNN-based approaches, respectively. Regarding sensitivity results, the SVM-based approach reached the best performance with $91\% \pm 22\%$, followed by the CNN-based approach with $88\% \pm 13\%$, and finally the proposed methodology with $83\% \pm 24\%$. The SVM-based approach consistently excelled in sensitivity as it was optimized with respect to this metric [1]. However, the proposed methodology in this study maintained a balanced performance between sensitivity and specificity at an overall level.

In addition to the performance results, the hybrid classification scheme offers significant advantages in terms of reduced data size and computational time. This trend is evident when examining the input data requirements for each methodology, as shown in Table 6.3. The CNN-based approach utilizes 49,565 image patches with dimensions of 87×87 pixels and 275 spectral bands, requiring approximately 768 GB of memory. On the other hand, the SVM-based approach uses 426,260 superpixels with 275 bands, which required approximately 0.87 GB of memory space. In contrast, the proposed hybrid classification scheme dramatically reduced the amount of data needed for classification. It only required 4,940 vectors, each with 875 elements, totaling 22.9 MB of memory. This represents a reduction of $\times 38$ compared to the SVM-based scheme in [1]. Consequently, the proposed approach significantly reduces memory usage and computational

Table 6.2: Evaluation and comparison of outcomes among the hybrid classification scheme, a superpixel SVM-based approach [1], and a CNN-based methodology [2], with emphasis on bold-highlighting the top-performing results.

	Hybrid Classification Scheme			**SVM-Based Approach [1]**			**CNN-Based Approach [2]**		
Patient	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**
P1	0.92	0.86	0.94	**1.00**	**1.00**	**1.00**	0.93	0.91	0.96
P2	0.88	0.91	0.86	**0.94**	0.83	**1.00**	0.89	**0.99**	0.83
P3	0.97	0.88	**1.00**	**1.00**	**1.00**	**1.00**	0.85	0.91	0.80
P4	**0.79**	0.16	**0.98**	**0.79**	**1.00**	0.73	0.57	0.57	0.58
P5	**0.81**	0.72	**0.83**	0.68	**1.00**	0.63	0.69	0.81	0.64
P6*	0.78	0.29	0.90	**0.90**	**0.50**	**1.00**	-	-	-
P7	**0.98**	**1.00**	**0.97**	0.35	**1.00**	0.10	0.66	0.71	0.63
P8	**0.96**	**1.00**	0.94	**0.96**	0.83	**1.00**	**0.96**	0.96	0.96
P9	0.95	0.95	NA	0.95	0.95	NA	**0.99**	**0.99**	NA
P10	0.75	0.75	NA	0.25	0.25	NA	**0.89**	**0.89**	NA
P11	**1.00**	**1.00**	NA	**1.00**	**1.00**	NA	0.92	0.92	NA
P12	0.70	0.70	NA	**1.00**	**1.00**	NA	0.92	0.92	NA
P13	0.98	0.98	NA	**1.00**	**1.00**	NA	0.99	0.99	NA
Avg.	**0.89**	0.83	**0.82**	0.83	**0.91**	0.78	0.86	0.88	0.77
STD.	**0.10**	0.24	0.33	0.27	0.22	0.34	0.14	**0.13**	**0.16**

Metrics marked as NA indicate that they are not applicable due to the absence of both classes in these patients.
* Data for the CNN-based approach is not available, and therefore, not included in the calculation of the mean (Avg) and STD.

Table 6.3: Comparative analysis of training duration, inference time, and storage requirements among the hybrid classification scheme and two prior methods: a superpixel SVM-based approach [1] and a CNN-based methodology [2].

Method	Training Time (minutes)	Inference Time per image (seconds)	Storage size
CNN-based Approach [2]	NA	72	768 GB
SVM-based Approach [1]	≈ 240	1440	0.870 GB
Hybrid Classification Scheme	AD proposal 7 DD proposal 4	18	0.023 GB

NA = data not available for the CNN-based approach.

overhead for data classification.

This advantage is further demonstrated in the training time for each methodology, as shown in Table 6.3. The SVM-based approach needed 4 hours for training, whereas our AD and DD proposals for zero-level NNs took approximately 7 and 4 minutes, respectively. Notably, the CNN-based approach [2] required more powerful processing equipment, as the training time exceeds one day. Regarding inference times (Table 6.3), the proposed hybrid classification scheme achieved the lowest value compared to the CNN-based approach, taking only 25% of the inference time of the DL scheme. Similarly, compared to the SVM-based method, the difference is even more significant, with the hybrid scheme requiring only 1.25% of the time. These improvements are attributed to the lower complexity of the proposed hybrid classification scheme. In fact, the optimization of the NNs ensemble involved a total of 3,229,925 trainable parameters. Among these parameters, 1,048,320 correspond to each level-zero NN, while the level-one NN has only

84,965 parameters.

Additionally, in analyzing the evaluation of HS images, the primary complexity lies in the SU stage, which has a complexity of $\mathcal{O}\left(maxiter \cdot K \cdot (L + N^4)\right)$, where L represents the number of spectral bands and K is the number of pixels to be analyzed in the HS image. This stage can be considered the bottleneck in the hybrid scheme. Similarly, the main limitation of the SVM-based approach is the bottleneck caused by the superpixel generation process. Although the superpixel algorithm can be highly parallelized, the current implementation in [1] is iterative, resulting in long inference times. Furthermore, the original paper that used this algorithm for HS histological data mentions an overestimation of the number of superpixels per image, which further increased the computational time required for superpixel generation. Nonetheless, both the proposed hybrid scheme and the comparison methods are amenable to parallelization, which would significantly reduce training and inference times.

6.2.3.3 Performance Evaluation

The results obtained from the hybrid scheme demonstrate the successful classification of HS images based solely on their spectral characteristics, without apparent overfitting. Moreover, the proposed approach presented significant improvements compared to CNN- and SVM-based methods [1, 2], which utilize both spatial and spectral information. It should be noted, however, that there might exist a CNN architecture capable of achieving similar results for this specific application. Nevertheless, the design of such a CNN model, fine-tuning of hyperparameters, and the learning process would pose significant challenges when applied to the studied database, especially when compared to the simplicity and effectiveness of the proposed hybrid classification scheme.

Although the hybrid classification scheme has shown excellent results, it does have some limitations. One major limitation is the small number of patients in the database and the class imbalance between tumor and non-tumor samples, as only eight patients had data from both classes. Although data augmentation partially compensates for this limitation, there is a limit to the improvement in classification performance by dividing spectral signatures into subsets. Moreover, the histological variability among GB-type tumors needs to be considered. To address these limitations, it is recommended to increase the number of patients in future studies.

Another limitation is the computational time required for the SU stage. The LU process for the entire database takes approximately 12 hours and 19 minutes, translating to about one and a half minutes per HS image. However, by performing a random downsample of one-tenth of the information in each HS image, consistent classification with good accuracy can still be achieved, reducing the computational time to just 2 hours and 24 minutes for the SU stage, equivalent to approximately 18 seconds per HS image. Further research is needed to determine the minimum sample reduction ratio that allows for good classification results. Overall, these limitations highlight areas for improvement and future research to enhance the hybrid

classification scheme's performance and efficiency.

6.2.3.4 Qualitative Analysis

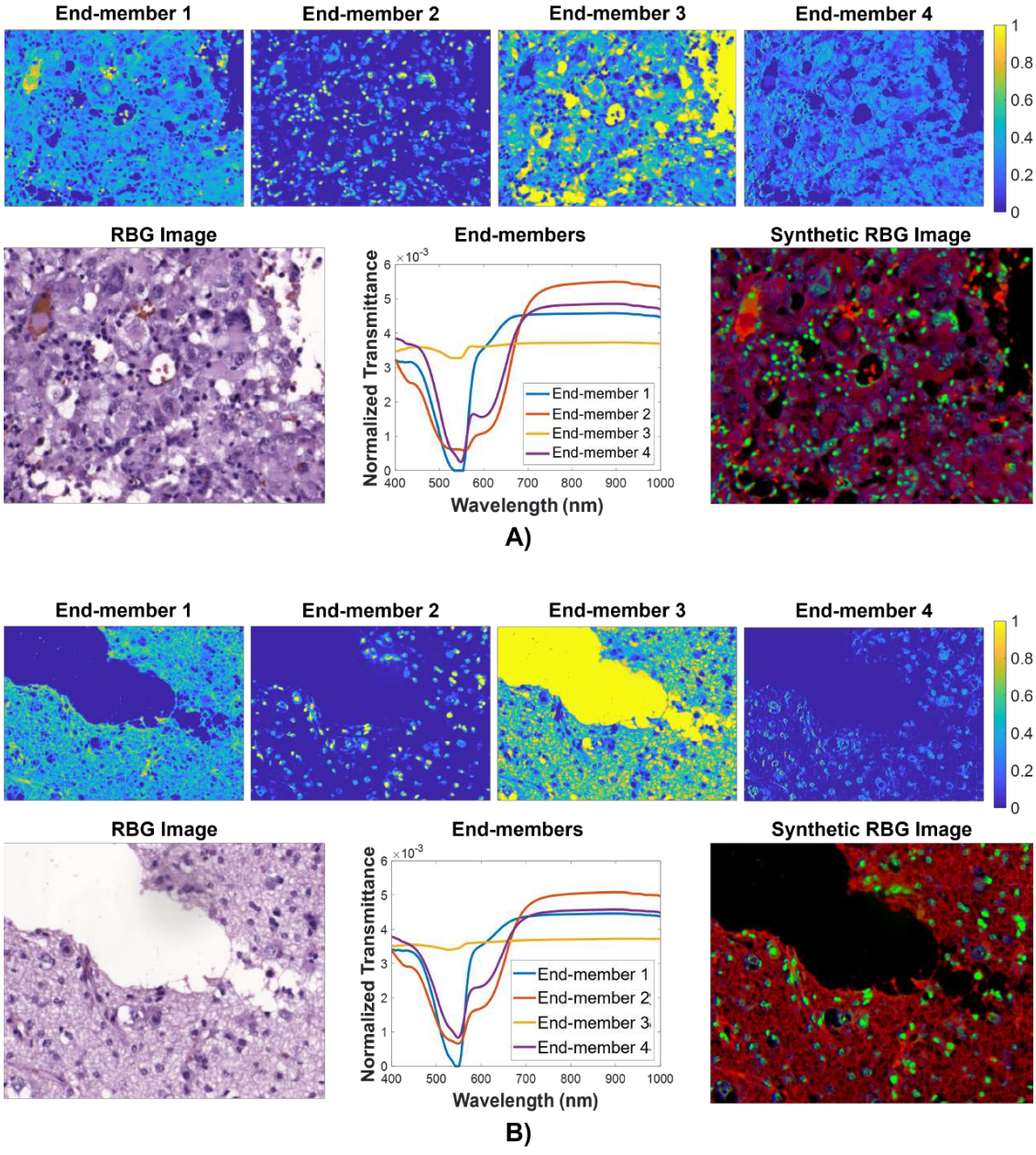


Figure 6.12: Qualitative outcomes of two HS images labeled as A) Tumor and B) Non-tumor are depicted. The figures illustrate the original RGB image (left image in the bottom row), the end-members (center of the bottom row), and the abundance maps estimated by SU (top row). Additionally, a synthetic RGB image (right image in the bottom row) is generated using the abundance maps of end-members 1, 2, and 4, which are assigned to the respective RGB channels.

While the hybrid classification methodology successfully identified tumor tissue in histological samples, it has a limitation regarding spatial classification, as it does not provide pixel-by-pixel annotation. However, the SU stage, as discussed in Section 6.2.1, allows the identification of characteristic spectral signatures and their concentrations through abundance maps. Therefore, after the classification process, the abundance maps can be utilized to identify spatial structures within the HS image. This property is demonstrated in Figure 6.12, which presents the RGB image, end-members, abundance maps, and a synthetic RGB image generated using the abundance maps corresponding to specific end-members. In this case, end-member 3, representing the direct transmission of microscope light, is not considered for the synthetic RGB image. Figure 6.12 provides two examples of correctly classified HS images, with Figure 6.12A representing tumor tissue and Figure 6.12B representing the non-tumor class.

A detailed analysis of the HS images revealed significant features, as indicated by the results of the SU stage and the original/synthetic RGB images. Firstly, the RGB images show distinguishable characteristics between the tumor and non-tumor classes. The tumor tissue image showed higher color intensity and a coarse texture, whereas the non-tumor tissue image displayed lower color intensity and a smoother texture. Secondly, the shapes of the end-members are quite similar between both classes, with slight spectral variations. End-member 1 is the most similar for both classes, hence the necessity for three distinct end-members in the training process to provide discriminative features during the classification stage. The abundance maps illustrate the concentration zones of each end-member, with the second end-member capturing structures related to cell nuclei, while end-members 1 and 4 highlight external structures such as blood vessels, gray matter, and white matter. Additionally, the abundance map of end-member 4 in the non-tumor case (Figure 6.12B) exhibited lower concentrations compared to the tumor condition (Figure 6.12A). These abundance maps offer insights into the spatial distribution of different components, enabling the identification of relevant patterns and features not readily visible in RGB images. Consequently, these patterns and abundance maps can be utilized in classification tasks. However, it is important to note that these observations are specific to the studied example. Nonetheless, synthetic RGB images provide an accurate representation of cellular structures, and in combination with the abundance maps from the SU stage, they can be valuable in future studies for pixel-wise classification.

6.3 Classification of Brain Tissues in Hyperspectral Images Using Vision Transformers

The ViT architecture, originally described in [6], was utilized in this proof of concept. A brief description of the ViT architecture will follow. The ViT primarily comprises position encoding embedding, linear mapping, transformer encoder, learnable layer, and normalization layer. The transformation of HS images $\mathbf{x} \in \mathbb{R}^{H \times W \times C}$ into flattened 2D patches $\mathbf{x}_p \in \mathbb{R}^{N \times (P^2 \cdot C)}$ occurs, where H and W represent the spatial dimensions, C signifies the number of bands, P denotes the size of each 2D patch, and $N = H \cdot W/P^2$ represents the number of patches. Furthermore, a latent constant vector of size D is employed uniformly

across all layers of the transformer. This vector is leveraged to map the flattened patches to dimensions D using a trainable linear projection along with the patch embeddings, thereby ensuring consistent patch order, regardless of resolution. Subsequently, positional encoding is applied to the patch embeddings, enabling the inclusion of spatial information in the classification process. Additionally, a learned class token is appended to the patch embeddings with positional encoding, providing global representation. The transformed input is then passed through the transformer encoder, and classification is performed through regression using a multi-layer perceptron in the header.

In this proof of concept, both training and evaluation were conducted at the pixel level instead of patch level. This decision was made as dividing the data into patches did not yield significant improvements. Moreover, increasing the number of patches to two or more resulted in smoothed classification maps with a loss of important details. The general diagram of the architecture and the input data are illustrated in Figure 6.13, where, unlike the original architecture in [6] (shown in Section 8), where intensity values are sequentially and orderly inputted to the transformer, in this proof of concept, the input data are the spectral values of each pixel. This allows for a comparison between spectral intensities, which, due to the structure of the transformer, enables a contextual comparison between spectral signature values.

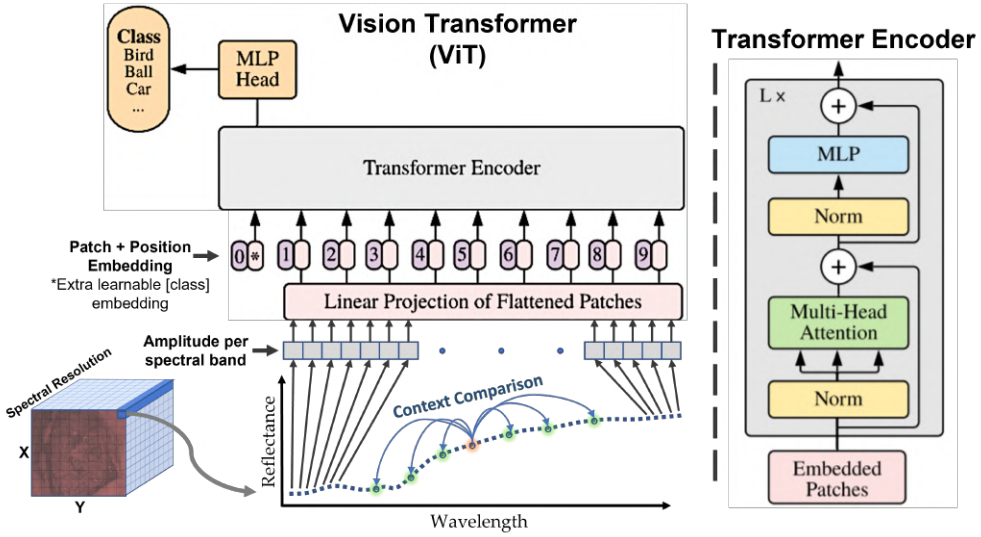


Figure 6.13: Diagram of the vision transformer architecture for spectral signature input.

The quantitative and qualitative results of ViT on the experimental brain cancer HS image database, described in Section 3.3.2.2, are presented in this section. The images P008-01, P012-01, and P020-01 were specifically analyzed. ViT's results were compared with labeled pixels using standard metrics such as accuracy, sensitivity, precision, and F1-score. The evaluation was conducted on the PyTorch platform using a computer equipped with an Intel Core i7 quad-core 4.2 GHz processor and 16 GB of RAM.

As a proof-of-concept, ViT was trained using both intra and inter-patient approaches. In the intra-patient

perspective, with 20% of the labeled data per class and the remaining as a validation set. In the inter-patient approach, where the labeled data from all patients served as the validation and training set without considering the information of the analyzed image. For this approach, 40% and 60% of the data were allocated to each set. The training percentages were selected in intervals of 5% to ensure sufficient representative information for correct classification of each class and to maintain stable accuracy performance during training. In both approaches, the test set consisted of pixels labeled by the clinical expert within the image of interest.

The intra-patient training stage was set up with 200 epochs for each image, using the Adam optimizer, a minibatch size of 32, and a learning rate of 0.001 with a decay factor of 0.9 after every tenth of the total number of epochs. Similarly, the inter-patient training stage was configured with 20 epochs for each image, a minibatch size of 128, the Adam optimizer, and a learning rate of 0.0005 with the same decay factor as in the intra-patient case. For comparison, this work also used the methodology proposed in [3], which includes a DL framework (DLF) with an inter-patient approach.

Table 6.4: Performance metrics of the ViT model in both the intra-patient and inter-patient approaches, as well as the DLF methodology [3]. The best result for each image in the inter-patient approach is emphasized in bold.

Image ID	Accuracy per class				Overall	Precision	Sensibility	F1 score
	NT	TT	HT	BG	Accuracy			
ViT in Intra-patient Approach								
P008-01	0.997	0.938	0.969	1.000	0.978	0.977	0.977	0.977
P012-01	0.992	0.995	0.999	0.994	0.996	0.997	0.997	0.997
P015-01	0.985	0.999	1.000	0.992	0.997	0.997	0.997	0.997
ViT in Inter-patient Approach								
P008-01	**0.881**	**0.997**	0.551	**1.000**	**0.840**	**0.860**	**0.841**	**0.839**
P012-01	0.985	**0.961**	0.927	**0.998**	**0.953**	**0.961**	**0.953**	**0.955**
P015-01	0.780	0.519	**0.922**	**1.000**	0.805	0.881	0.805	0.812
DLF Methodology [3] (Inter-patient Approach)								
P008-01	0.761	0.257	**0.612**	0.998	0.639	0.707	0.640	0.639
P012-01	**0.993**	0.698	**0.947**	0.987	0.951	0.955	0.951	0.950
P015-01	**0.989**	**0.916**	0.684	0.945	**0.811**	**0.911**	**0.811**	**0.833**

The results presented in Table 6.4 demonstrate that in the intra-patient case, despite the presence of unbalanced data, the ViT model achieves accuracy values above 0.90 for each class, with an overall accuracy of 0.97. These high accuracy results, along with the precision, sensitivity, and F1-values, indicate that the classification model is robust and highly efficient. On the other hand, in the inter-patient scheme, the overall accuracy of ViT is similar to the DLF methodology [3], except for the P008-01 image, where an improvement of up to 20

A closer analysis of the metrics reveals that the DLF methodology is clearly affected by the class imbalance, resulting in lower accuracy values for the TT and HT classes (classes with fewer labeled pixels). In contrast, the intra-patient approach shows a higher average F1-score (87%) and performs better in classifying the data compared to the DLF methodology (81%), which is more influenced by the class imbalance.

These findings indicate that the ViT model can be a promising and effective choice for the classification of brain cancer HS images.

The resulting classification maps in Figure 6.14 demonstrate that ViT correctly identified the rubber rings (background) in all cases, even though they were not fully considered in the ground-truth. Similarly, the consistent structures are identified for the remaining classes, aligning with what is shown in the RGB reference image. Notably, the regions classified as tumor closely follow the approximate borders delimited by the clinical expert.

The main difference among the studied strategies lies in the background area. The ViT in the inter-patient approach exhibits larger background areas compared to the intra-patient one, while the DLF methodology produces a larger division that distinguishes between the parenchymal region and the background. Despite these differences, all three approaches generate similar classification maps, as depicted in Figure 6.14, with the highest similarity observed between the inter-patient approach and the DLF methodology.

The ViT in the inter-patient approach shows a higher number of coincidences with the ground-truth, particularly in the HT class. However, the DLF methodology generates more uniform classification maps with fewer errors among the NT, TT, and HT classes, albeit with some loss of detail in the HT class. When comparing the DLF methodology with the inter-patient approach, the labeled tumor region in the latter case is closer to the area defined by the expert.

The results demonstrated the viability of ViTs as a proof-of-concept by achieving classification accuracies above 93% for all classes in the intra-patient approach, even with a limited amount of training data (20% of labeled pixels). In the inter-patient approach, the performance results exceeded 80% for all classes, with clear delineation among different tissue classes, surpassing the average performance of the DLF methodology. However, the inter-patient approach exhibited lower performance in certain classes, indicating a need to modify the training characteristics to address this issue.

Notably, the inter-patient training process was more challenging, requiring approximately three hours per image compared to approximately 15 minutes in the intra-patient case. The evaluation time for each image was around 7 minutes. Unfortunately, there were no available computational times for comparison with the DLF methodology [3]. These findings suggest that models utilizing transform-type architectures hold promise for tissue classification tasks and surgical guidance applications.

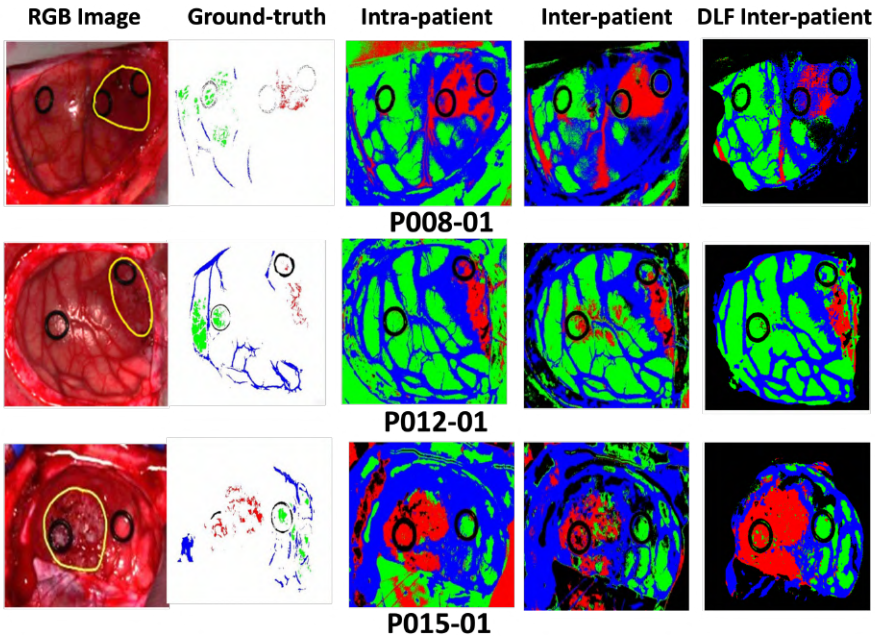


Figure 6.14: Qualitative classification results for each image are displayed in the following manner: the first column showcases a synthesized RGB image with the tumor region highlighted by a yellow boundary, while the second column exhibits the ground-truth image containing pixels labeled by the clinical expert (green for non-tumor, red for tumor tissue, blue for healthy tissue, and black for background). The remaining three columns present the classification maps generated by ViT using the intra-patient and inter-patient approaches, as well as the DLF methodology [3].

Chapter 7

Conclusions

This dissertation presents a comprehensive study of proposals for classification and SU in HS images. These methodologies include several formulations and novel techniques, which were evaluated on synthetic, real, and experimental HS images from different biomedical and remote sensing contexts. The research focuses primarily on reflectance calibration, blind linear and nonlinear unmixing, and hybrid classification methods. The results demonstrate the effectiveness of the proposed techniques in preserving spectral signatures, improving classification accuracy, and reducing computational time. This is presented in a condensed form in Table 7.1, which summarizes the main conclusions obtained in the development of each chapter, and their contributions are analyzed in detail below.

In the first instance, in Section 4.5, a methodology called Extended Blind End-member and Abundance Estimation with Spatial Coherence (EBEAE-SC) was introduced as a BLU algorithm applicable to biomedical imaging and remote sensing. EBEAE-SC showcased its remarkable ability to perform quantitative analysis with minimal prior information. It enabled the estimation of concentration maps with spatial coherence by incorporating internal abundances within the optimization process. The flexibility of its hyperparameters $(N, \rho, \lambda, \nu, \tau)$ allowed the methodology to adapt effectively to different applications. Similar to the original formulation EBEAE, the selection of hyperparameters (N, ρ) relied on prior knowledge about the dataset, such as the number of components or end-members present and the expected spectral similarities. On the other hand, the hyperparameters (λ, ν, τ) influenced the balance between nominal and internal abundances and effectively reduced the impact of noise in the measurements using TV theory. This fine-tuning approach of hyperparameters ensured a better fit to the ground-truth information while preserving crucial structures and edges in the abundance maps, leading to substantial improvements in estimation accuracy, particularly in high-noise scenarios compared to cutting-edge BLU methods. Moreover, EBEAE-SC exhibited a lower computational cost than other methodologies examined in test scenarios, resulting in an excellent quality-

Table 7.1: Overall summary of the main conclusions of each chapter presented in the dissertation.

Chapter	Title Chapter	Essential conclusions of the dissertation
1	Introduction	• Demonstrate the utility of HSI in the medical field, showcasing the ability of these techniques to detect various pathologies and tumors.
2	Methodological Background	• Expone the most prominent HS image analysis strategies in the state of the art.
3	Databases for HS Image Analysis: Description and Preprocessing	• The characteristics of the databases used in the thesis, along with the preprocessing required to standardize spectral signatures, are described.
4	Spectral Unmixing	• Theoretical foundations of LMM and NMM are established, providing a solid basis for understanding these approaches. • The theoretical framework of EBEAE is detailed, and the fundamental principles of HS data classification through unmixing techniques are presented. • Formulations for the proposed models are developed, yielding noteworthy results in terms of accuracy and computational efficiency. • Results obtained through EBEAE-SC have demonstrated a significant reduction in estimation granularity, leading to a decrease in overall estimation error and superior performance compared to traditional methods. •The formulation of NEBEAE has shown its capacity to outperform previously reported non-linear methods in the literature, achieving improved performance in both global terms and processing time. • The LU-based classification proposal has proven its ability to generate competitive results when compared to ML methods, all while consuming fewer computational resources. • The calibration proposal is evaluated in an experimental scenario in conjunction with NEBEAE, showing an increase in the utilized metrics.
6	Advancing HS Brain Tissue Classification with Hybrid Schemes of LU and Artificial Intelligence	• An innovative calibration proposal for hyperspectral data is introduced, effectively reducing the adverse effects of the standard method and resulting in a significant performance boost in controlled and experimental classification scenarios.
7	Advancing HS Brain Tissue Classification with Hybrid Schemes of LU and Artificial Intelligence	•Two proposals were developed that combine LU with various ML techniques. These proposals use abundance maps and end-members as features introduced into the classifier. The results obtained show similarities to methodologies previously published in the literature while improving the processing time and memory efficiency required to store the features. • Attention-based techniques for HS image classification were explored, and the results demonstrated superior performance in a proof-of-concept compared to reference methods.

time processing relationship, making it a highly practical and advantageous choice for various applications.

Following the development of the EBEAE-SC method, this dissertation introduces a nonlinear variant of EBEAE (Section 4.6), which is based on a multilinear mixing model. This new approach exhibits robust performance when dealing with various types and levels of noise, all while maintaining a low computational burden. The validity of this formulation is illustrated through its application to synthetic datasets and comparison with advanced remote sensing methods, showcasing its superior computational efficiency with improvements ranging from 1.6 to 1117 times faster than the comparison methods. When applied to VNIR HS images for intraoperative brain tumor detection, the proposed method exceeds its linear counterpart and effectively highlights crucial tissue characteristics. NEBEAE significantly reduces the overall estimation error and improves accuracy by an average of 2% in classification tests. Additionally, the nonlinear interaction maps provide valuable information on important tissue features, particularly in hypervascularized tissue

areas.

Subsequently, in Section 4.7, this work introduces two novel methods that use the EBEAE algorithm to classify in vivo brain tissue through intraoperative HS images. These innovative approaches outperform conventional supervised SVM-based techniques in terms of computational efficiency while providing comparable or even improved classification performance. Moreover, both methods demonstrate reduced variability in classification results and offer significantly simpler processes compared to other existing strategies, resulting in a remarkable training speed improvement of more than 420 times compared to the benchmark method used. A noteworthy advantage of these methods is their enhanced precision in identifying the BG class, particularly evident in Method B, which also shows improved sensitivity and F1-score for detecting the TT class compared to the methods proposed in the existing literature. Combining these overall advantages with the immense potential for parallelization in the proposed techniques and the utilization of snapshot HS cameras (capable of capturing both spectral and spatial information in one shot), the possibility of real-time classification results during clinical procedures becomes achievable. This advancement has the potential to significantly improve surgical outcomes, leading to improved patient results and overall quality of life.

In Chapter 5, a new formulation for reflectance calibration in HS images was introduced, which compensates for unwanted effects in the image acquisition stage. The results of the developed reflectance calibration method demonstrated its effectiveness in preserving the shape of the spectral signatures, limiting their amplitudes (in the range of 0 to 100%), and improving the separability among the different classes present in the HS image. These properties are especially crucial in medical applications, where an accurate identification of various classes is sought. This application was illustrated in Section 5.2, where the impact of calibration on the classification of medical HS images was analyzed using a nonlinear unmixing algorithm. The results of the PC showed a reduction in the variability of spectral signatures within the same class, which resulted in a 22% improvement in the initial classification tests a 6% improvement in the inter-patient tests. This variability reduction was reflected in the performance of the inter-patient strategy, which simulates real-world scenarios where labeled information from new patients is not available. The application reflects the ability of the proposed formulation to uniformly adjust spectral signatures despite physical variations present at the time of capture.

In Chapter 6, two new hybrid classification methods were introduced to address the tissue classification problem in brain tissue HS images. These methods combine EBEAE and EBEAE-SC with popular ML approaches, such as NNs, SVMs, and RFs. The primary goal of the first proposed method (Section 6.1) is to classify tissues in vivo intraoperative brain tissue HS images. It leverages linear unmixing to generate abundance maps, which serve as feature vectors for ML classifiers. The extended basis function used in this approach allows for better differentiation between the classes of interest, thus improving the overall classification performance. Remarkably, the hybrid method employing NNs achieved the most promising results ($75.9 \pm 13\%$ average accuracy), outperforming other state-of-the-art methods. It demonstrated superior performance in reducing classification errors for specific tissue classes, mitigating granularity issues present

in SVM-based approaches, and overcoming oversmoothing challenges associated with LU-based schemes. Moreover, the proposed hybrid approach retains a higher level of detail in the resulting classification maps. The flexibility of this hybrid approach is a significant advantage, enabling adaptation to various scenarios by adjusting control parameters or even considering different classifiers to optimize the results. However, it is essential to note that handling multiple control parameters may require careful tuning to achieve the best outcomes.

Meanwhile, the second hybrid classification proposal focused on the identification of GB tumor-affected tissue in histopathological HS images (Section 6.2). The methodology involved using SU as a feature extraction by the estimated end-members, which were then input into an ensemble of deep NNs for binary classification. The results showcased the effectiveness of the proposed method compared to two state-of-the-art approaches [1, 2]. The proposal achieved an average accuracy and specificity of 89% and 82%, respectively, with reduced variance. However, the comparison methods outperformed the proposed one in terms of sensitivity (average of 83%). Furthermore, this approach significantly reduced the computational cost during classification, leading to shorter learning and inference times without compromising performance. These specific results open possibilities for the proposed approach to be applied in clinical settings to identify regions in brain histological samples for further analysis by a pathologist.

The dissertation concludes with an exploration of the application of ViTs to classify HS images from in vivo brain tissue (Section 6.3). ViTs demonstrate promising results in intra-patient classification with limited training data, achieving high accuracy (above 93%) across all classes. However, inter-patient classification also shows relatively good performance (above 80%), albeit with some classes performing less effectively. This issue can be addressed by adjusting the training process, which involves a higher computational cost. For each scenario, parameter optimization takes approximately three hours, while intra-patient training only requires about 15 minutes per patient. The evaluation time for both approaches is around 7 minutes per assessment. These findings indicate that transformer-based architectures like ViTs could be suitable for tissue classification tasks and surgical guidance applications.

In conclusion, this dissertation provides a comprehensive examination of various methodologies for classification and spectral unmixing in HS images. These methodologies encompass a wide array of novel techniques and formulations, which were rigorously evaluated across synthetic, real, and experimental HS images sourced from different biomedical and remote sensing contexts. Overall, this research contributes significantly to the field of HS image analysis, offering practical insights into the applicability of these methodologies across various domains and holding the potential to improve surgical outcomes and patient care. Moreover, while successfully attaining the primary objective and specific goals set forth, this research is at a preliminary stage of exploration, generating novel avenues for further investigation. Nevertheless, it is important to acknowledge and address several limitations inherent to this project. The subsequent section outlines these identified limitations and offers potential directions for future development.

7.1 Limitations

Despite the promising results of the different methodologies presented in this work, each proposal has certain limitations. One general limitation is the relatively small size of some databases used for evaluation, which affects the generalization of the classification results. Additionally, the databases have another limitation, namely, the imbalanced distribution of classes within the labeled dataset, which can lead to an overfitting problem in classes with a larger number of data points.

In particular, each method developed in this work has specific limitations. Initially, the calibration with normalization correction achieves the intended goal of reducing variability. However, there are still open issues associated with this proposal. The main limitation arises from the use of a flat surface as the white reference. In contrast, the surfaces to be analyzed often have textures and a certain level of volume, resulting in an inconsistency between the white reference and the surface being analyzed.

On the other hand, both proposed unmixing methods EBEAE-SC and NEBEAE share similar limitations, with the main one being the proper selection of their hyperparameters. This limitation extends to hybrid methods, where significant computational time is required for this stage, potentially hindering real-time applications. Additionally, hybrid classification schemes may not be suitable for all types of brain tumors since these algorithms were designed for a specific case and histological and spectral variability among different tumor types must be considered. Finally, the proof-of-concept using the transformer, despite showing promising results, is limited by the use of the standard ViT architecture, which restricts its performance on HS data.

Regarding the databases, they have limitations that are directly related to their acquisition process. For example, in the acquisition of the in vivo brain tissue database and within the context of limitations associated with the PC method, challenges arise related to the capture environment. These challenges include variations in ambient lighting, degradation of the illumination source, and differences in depth within the area of interest. These factors can have adverse effects on the interactions of captured light. Additionally, in the surgical environment, there are often accumulations of blood, gauze, medical instruments, and other materials that can affect the spectral information and the analysis within the image. While, in the case of histopathology image databases, a physical challenge arises when attempting to increase the level of magnification for HS image capture. This challenge is linked to the characteristics of the objectives used in optical microscopes, as their aperture decreases as magnification increases. In other words, as magnification increases, the effective diameter of the objective decreases. This limitation in image capture restricts the permissible bandwidth in the images and, in some cases, can affect the entire spectrum captured by HS cameras. Additionally, it can potentially lead to optical aberrations due to the properties of the lenses used in the objectives.

In clinical practice, there are multiple technical limitations, such as the characteristics of acquisition platforms, camera properties, capture time, computing equipment for processing, and algorithmic perfor-

mance issues. The latter problem has been successfully addressed in this dissertation with good results from the developed proposals. However, it is important to note that technical limitations still require considerable technological advancement. Additionally, more research and clinical trials are needed to obtain approval for a HS surgical guidance platform by various health agencies. It is also worth mentioning that clinical experts can often be resistant to changes in standard strategies. Despite these challenges, approaches to clinical personnel have been fruitful and well-received. In the case of histopathological applications, there are fewer technical limitations as the variables that can affect image capture are more controlled. However, in both the histopathological and GB detection cases, it is essential to conduct more comprehensive studies to validate both the acquisition platform and the proposed classification algorithms.

7.2 Future Work

To address the limitations and expand the contributions made in this work, several lines for future research are identified:

- Expand the size of the databases used, along with extending the labeling of experimental HS images, to improve classification performance and enable the identification of different tissue types.
- Explore different optimization methods and hyperparameters tuning techniques to enhance the adaptability and performance of the algorithms.
- Investigate the use of morphological postprocessing methods and spatial filtering algorithms to reduce misclassifications.
- Propose approaches for algorithm parallelization using hardware accelerators such as GPUs or FPGAs, to reduce computational time and enable real-time performance.
- Develop new variants of unmixing and ML algorithms for feature extraction and pixel-wise classification in HS images.
- Research image processing techniques for depth analysis and geometric perception alterations to reduce variability when evaluating standard reflectance calibration.
- Implement the developed classification methods into APIs for surgical guidance devices.
- Rigorous and comprehensive clinical studies will be conducted to validate both the acquisition platform and the proposed classification algorithms.
- Additional research will be conducted to address technical limitations and enhance clinical applicability.
- Close collaboration with clinical professionals will continue to understand their needs and concerns.
- Examine probabilistic algorithms to overcome the physical challenge in the analysis of histological samples using HSI.

Furthermore, various cutting-edge techniques in the field of image processing will be explored to extrapolate and adapt them to HS image analysis. Among the techniques under consideration are attention mechanisms, graph architectures, deep residual learning, adaptive processes, and genetic algorithms, which have the potential to significantly improve the performance of HS processing. The proposed future work aims to overcome the identified limitations and establish a solid foundation for future lines of research and applications.

Chapter A

Scientific Contribution

The present appendix aims to provide a detailed list of the work carried out during the doctoral period, including: articles published in indexed journals, book chapters, and both national and international conference papers. Additionally, it will present the works in which participation was taken throughout this doctoral training process. Each of these works represents a fundamental pillar in the construction of knowledge and the skills acquired during these years. Furthermore, this section offers readers an overview of the achievements and the diversity of topics addressed on the journey towards obtaining the doctoral degree.

A.1 Articles

- R. Leon, S. Ortega, I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. Szolna, J. F. Piñeiro, C. Espino, A. J. O'Shanahan, M. Hernandez, D. Carrera, S. Bisshopp, C. Sosa, F. J. Balea-Fernandez, J. Morera, B. Clavo, G. M. Callico, and H. Fabelo, "Hyperspectral Imaging Benchmark based on Machine Learning for Intraoperative Brain Tumour Detection," Submitted in NPJ Precision Oncology, 2023, Manuscript ID: NPJPRECISIONONCOLOGY-03695-T, doi: 10.21203/rs.3.rs-2956240/v1
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejia-Rodriguez, R. Leon, S. Ortega, H. Fabelo, R. Camacho, M. Plata, and G. M. Callico, "Hybrid Brain Tumor Classification Scheme of Histopathology Hyperspectral Images Using Linear Unmixing and Deep Learning," Submitted in Healthcare Technology Letters, 2023, Manuscript ID: HTL-2023-08-0027.
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejia-Rodriguez, A. J. Jo, S. Ortega, H. Fabelo, and G. M. Callico, "Multi and Hyperspectral Image Unmixing with Spatial Coherence by Extended Blind End-member and Abundance Extraction," J. Franklin Inst., 2023.

- D. U. Campos-Delgado, I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, A. R. Mejía-Rodríguez, S. Ortega, H. Fabelo, and G. M. Callico, "Nonlinear extended blind end-member and abundance extraction for hyperspectral images," Signal Processing, vol. 201, p. 108718, 2022, https://doi.org/10.1016/j.sigpro.2022.108718.
- I. A. Cruz-Guerrero, R. Leon, D. U. Campos-Delgado, S. Ortega, H. Fabelo, and G. M. Callico, "Classification of hyperspectral in vivo brain tissue based on linear unmixing," Appl. Sci., vol. 10, no. 16, Aug. 2020, doi: 10.3390/app10165686.

A.2 Book Chapters

- I. A. Cruz-Guerrero, R. Leon, A. R. Mejia-Rodriguez, D. U. Campos-Delgado, S. Ortega, H. Fabelo, and G. M. Callico, "Hyperspectral Imaging for Cancer Applications," in Diagnosis and Treatment of Cancer using Thermal Therapies, CRC Press, 2023, pp. 81–101, doi: 10.1201/9781003342663.
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejia-Rodriguez, H. Fabelo, S. Ortega, and G. M. Callico, "A hybrid approach to the hyperspectral classification of in vivo brain tissue: linear unmixing with spatial coherence and machine learning," in Artificial Intelligence in Cancer Diagnosis and Prognosis, Volume 3: Brain and prostate cancer, IOP Publishing, 2022, https://doi.org/10.1088/978-0-7503-3603-1ch9.

A.3 International Conferences

- J. N. Mendoza Chavarría, I. A. Cruz-Guerrero, D. U. Campos-Delgado, F. Fabelo, S. Ortega, and G. M. Callico, "Ensemble of Artificial Intelligence Classifiers for In-vivo Identification of Glioblastoma Tumours Using Hyperspectral Images," in IEEE Engineering in Medicine and Biology Conference in Latin America (IEEE EMBS R9), submitted August 14th, 2023.
- I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, D. U. Campos-Delgado, H. Fabelo, S. Ortega, and G. M. Callico, "Classification of Brain Tissues in Hyperspectral Images Using Vision Transformers," in 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), 2023, pp. 1–4, doi: 10.1109/ISBI53787.2023.10230806.
- I. A. Cruz-Guerrero, R. Leon, L. Granados-Castro, H. Fabelo, S. Ortega, D. U. Campos-Delgado, and G. M. Callico, "Reflectance Calibration with Normalization Correction in Hyperspectral Imaging," in 2022 25th Euromicro Conference on Digital System Design (DSD), 2022, pp. 855-862, https://doi.org/10.1109/DSD57027.2022.00120.
- L. Granados-Castro, O. Gutiérrez-Navarro, I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, E. R. Zavala-Sánchez, and D. U. Campos-Delgado, "Estimation of deoxygenated and oxygenated hemo-

globin by multispectral blind linear unmixing," in 2022 25th Euromicro Conference on Digital System Design (DSD), 2022, pp. 849–854, https://doi.org/10.1109/DSD57027.2022.00119.

- J. N. Mendoza-Chavarría, E. R. Zavala-Sánchez, L. Granados-Castro, I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, and D. U. Campos-Delgado, "Glioblastoma Classification in Hyperspectral Images by Nonlinear Unmixing," in 2022 25th Euromicro Conference on Digital System Design (DSD), 2022, pp. 843–848, https://doi.org/10.1109/DSD57027.2022.00118.
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, and A. R. Mejía-Rodríguez, "Extended Blind End-member and Abundance Estimation with Spatial Total Variation for Hyperspectral Imaging," in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 1957–1960, https://doi.org/10.1109/EMBC46164.2021.9629708.

A.4 National Conferences

- D. A. Martínez-Arboleyda, I. A. Cruz-Guerrero, and D. U. Campos-Delgado "Tumor Tissue Classification in Hyperspectral Histopathology Images through Individual and Ensemble of Machine Learning Algorithms," in Congreso Nacional de Ingeniería Biomédica 2023, 2023, Accepted.
- I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, and D. U. Campos-Delgado, "Glioblastoma Classification in Hyperspectral Images by Reflectance Calibration with Normalization Correction and Nonlinear Unmixing," in Congreso Nacional de Ingeniería Biomédica, Cham: Springer International Publishing 2022, pp. 393–402, https://doi.org/10.1007/978-3-031-18256-3_43.
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejía-Rodriguez, S. Ortega, H. Fabelo, and G. M. Callico, "Enfoqué Híbrido para la Clasificación de Muestras Histológicas para la Detección de Tumores Cerebrales: Descomposición lineal y Redes Neuronales," in Memorias del Congreso Nacional de Ingeniería Biomédica, 2021, vol. 8, no. 1, pp. 97–100.
- J. N. Mendoza-Chavarría, I. A. Cruz-Guerrero, A. R. Mejia-Rodriguez, and D. U. Campos-Delgado, "Algoritmo de descomposición ciega basado en el modelo de mezcla multi-lineal," in Memorias del Congreso Nacional de Ingeniería Biomédica, 2021, vol. 8, no. 1, pp. 106–109.
- I. A. Cruz-Guerrero, D. U. Campos-Delgado, and A. R. Mejía-Rodríguez, "Estimación Ciega de Perfiles y Abundancias con Variación Total para Imágenes Hiperespectrales de Uso Clínico," in Memorias del Congreso Nacional de Ingeniería Biomédica, 2020, vol. 7, no. 1, pp. 548–555.

Chapter B

Machine Learning Methods

B.1 Random Forests

Random Forest (RF) is a supervised classification and regression algorithm, originally developed by Breiman, that adopts a learning approach using random subsets of training data [203]. The algorithm involves training multiple decision trees simultaneously to contribute collectively to the classification process, helping to reduce data variance and improve accuracy. The classification is achieved through a voting mechanism, where each decision tree casts a vote, and the final class label is assigned based on the highest number of votes [97, 204].

RF employs an ensemble strategy that includes two main methods: boosting and bagging. Boosting is a sequential process in which each subsequent model attempts to correct the errors of the previous ones in a particular sequence [203]. In contrast, bagging involves training each decision tree on different subsets of data, meaning that each tree observes distinctive data during its training phase. When the results of all trees in the forest are balanced, errors are balanced, leading to improved stability, accuracy, and reduced variance. Consequently, bagging is used more frequently in practice [203].

The hyperparameters of the RF classifier are mainly inherited from the decision trees, which include the maximum depth of the tree, the minimum number of samples required for node splitting, the minimum number of samples for a final node, and the maximum number of final nodes or leaves [204]. Additionally, the number of trees in the forest, the quality criteria (typically the Gini criterion), and the maximum number of features considered to generate a split are also defined as hyperparameters for the RF classifier [203].

RF has proven to be an effective strategy for classifying high-dimensional data, such as HS images and remote sensing data. RF combines the strength of multiple classifiers, resulting in improved accuracy and stability [205]. Compared to other classification algorithms, such as AdaBoost and SVM, RF has presented faster computation times and comparable or even superior performance in various remote sensing

applications [205].

B.2 Support Vector Machines

The Support Vector Machine (SVM) algorithm, developed by Vapnik in the late 1970s, is widely utilized for classification tasks, particularly in image classification. The SVM serves as a linear binary classifier that separates classes using an optimal hyperplane (Figure B.1) [97, 203]. During training, the SVM algorithm utilizes the available data to compute support vectors, which are generated from the data points closest to the decision boundary. These support vectors determine the optimal decision boundary, minimizing classification errors and maximizing separation [203]. While the basic SVM algorithm is designed for linear binary classification problems, real-world scenarios often require nonlinear class divisions. To address this, several variants of SVM kernels have been developed. Nonlinear kernels, such as sigmoid, radial basis function (RBF), and polynomial kernels enable improved separation of overlapping data. However, defining the hyperparameters for these nonlinear kernels can be challenging [97].

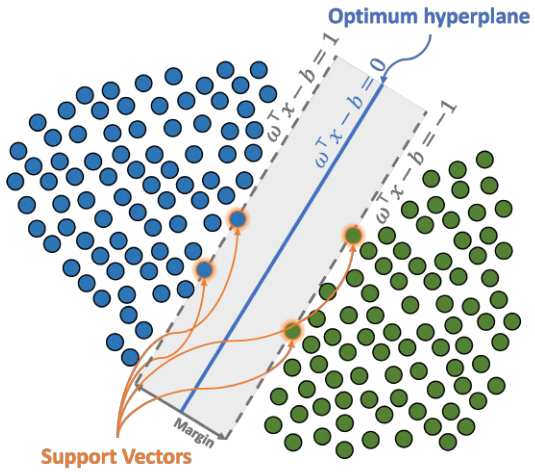


Figure B.1: Example of linearly separable data classified by an SVM.

Extensions of SVM have also been proposed for multiclass classification. For example, the "one-versus-the-rest" approach involves a single optimization process and requires fewer support vectors compared to multiple stages of binary separation [97]. On the other hand, the Crammer-Singer method solves a dual optimization problem with a set of constraints, allowing for multiple subproblems. Due to its characteristics, the Crammer-Singer method is slightly more sensitive, but achieves a more accurate classification [206].

In the classification of HS images, SVM has gained prominence due to its strong theoretical foundation, good generalization capability, low sensitivity to the curse of dimensionality, and the ability to find global classification solutions [207, 208, 209]. SVMs have shown superior performance in terms of accuracy, computational time, and stability of parameter settings when compared to other algorithms such as RBF,

NNs, and KNN classifiers [207]. Experimental results have demonstrated the effectiveness of SVMs in tasks such as spectral signature classification and tissue differentiation [207, 209]. These findings highlight the versatility and potential of SVMs in a wide range of classification tasks, particularly in the analysis of HS data.

B.3 Fully Connected Neural Networks

Neural Networks (NNs) are classification techniques that draw inspiration from biological neurons, with multilayer NN being the architecture most commonly used. An NN consists of interconnected or fully connected layers, each composed of neurons or perceptrons. The perceptron serves as the fundamental unit of an NN [97, 204]. Within a perceptron, the bias term and weight components act as adjustable parameters. While a perceptron functions as a linear classifier, the multilayer structure of interconnected perceptrons enables the separation of complex nonlinear patterns within the feature space [210]. In a supervised approach, the bias and weight components of the overall NN are modified to reproduce the desired outputs [210].

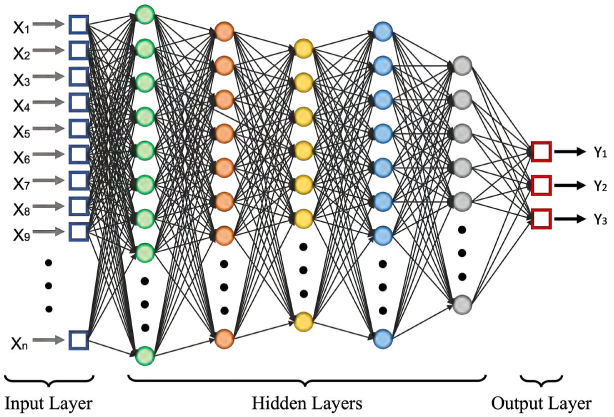


Figure B.2: Basic diagram of a fully connected neural network.

The basic structure of a NN consists of three layers: an input layer, a hidden layer, and an output layer. These layers are interconnected, as depicted in Figure B.2. The training of an NN involves two stages: forward propagation and backpropagation. In the forward propagation stage, the input data is passed through the network, flowing through the internal layers until it reaches the output. At the output layer, the estimation error is evaluated using a cost function [97, 204]. After calculating the estimation error, the second stage, known as backpropagation, begins. In this stage, the weights and biases of each layer are updated by iteratively computing the gradient of the cost function concerning each parameter [97, 204]. The entire process of forward and backward propagation is referred to as a training epoch. Once an epoch is completed, the process is repeated until a convergence condition is met [97, 204].

In addition, there is a set of parameters known as hyperparameters that play a critical role in controlling

the speed and quality of the learning process [97, 211]. These hyperparameters need to be defined prior to initiating the learning process and are further fine-tuned to achieve the optimal model for a given dataset. In the field of HS classification, several common hyperparameters are encountered. Next, some of them are described:

- Hidden layers and neurons: The number of hidden layers and neurons in each layer determines the depth and complexity of the NN architecture, affecting its performance [212].
- Activation functions: These functions define how the inputs to a neuron are transformed into outputs for the subsequent layer. Popular activation functions include sigmoid, rectified linear unit (ReLU), and softmax functions [213].
- Loss functions: These functions, such as mean squared error (MSE) and cross-entropy loss, quantify the discrepancy between the predicted output values and the ground-truth labels [214].
- Learning rate: The learning rate determines the step size at which the NN adjusts its parameters during the optimization process [215].
- Number of epochs and batch size: Epochs refer to the number of times the network processes the training data, while batch size specifies the number of input subsets used before updating the network parameters [215].

The application of a NN in the HS classification requires careful consideration of these parameters. Due to this property, NNs are typically designed based on the specific characteristics of the data to be classified [210, 216]. By optimizing these hyperparameters to suit the database and the task at hand, the NN can effectively leverage the inherent patterns within the HS data, leading to improved classification performance [210].

B.4 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a type of ML model used to process and analyze visual information, such as images and videos [97]. The basic architecture of a CNN consists of two well-defined main stages: feature extraction and classification [93, 100, 189, 217], as shown in Figure B.3. The feature extraction stage is typically located at the network input and is composed of filters that reduce the size of the input data while retaining the characteristics of the relevant information [93, 98, 189]. This spatial reduction helps capture invariances and enhances the network's ability to generalize to unknown data. The feature extraction stage involves a convolutional step, in which filters are applied to the input data, allowing the extraction of relevant patterns [93, 99, 189, 217]. These filters can detect features such as edges, textures, and spatially/temporally shaped characteristics. On the other hand, the classification stage assigns input data to a specific label and is composed of fully connected layers [93, 99, 189, 217]. Together, these two stages enable the classification of structures within different types of data, primarily images and videos.

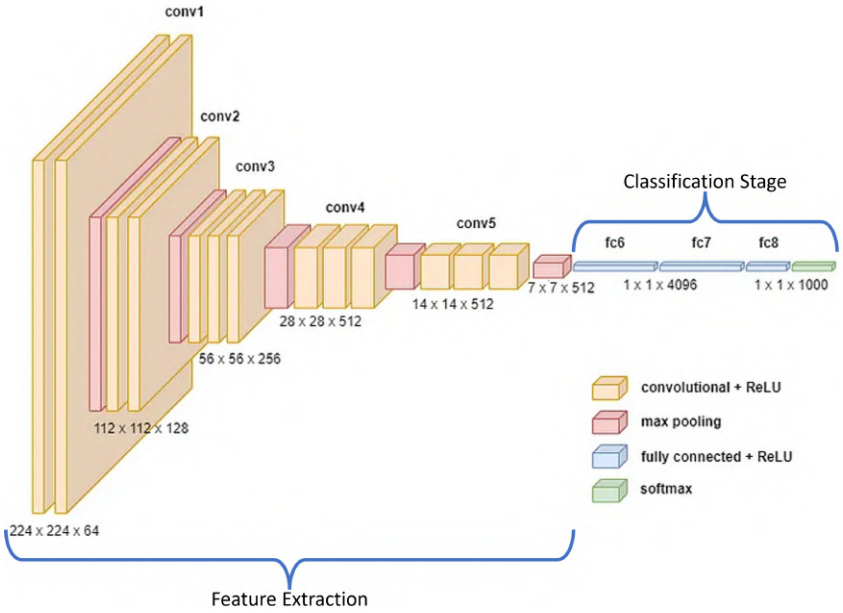


Figure B.3: Exemplification of a convolutional neural network.

In the literature, there are three main categories of CNN models used for the classification of HS data. These models differ according to the type of analysis they perform: spectral features, spatial features, and spectral-spatial features [93, 100, 189]. However, it is also common to employ only the feature extraction stage or the convolutional stage to capture patterns and subsequently use another classifier. CNN for spectral analysis uses the spectral information of each pixel in the HS image, achieved through convolutions with 1D kernels [93, 100, 189]. This approach allows obtaining the main spectral features of the image, resulting in per-pixel classification at the output. On the other hand, spatial analysis CNNs use convolutional stages with 2D kernels to extract structural features from each image throughout the spectrum [93, 100, 189]. In addition to this, authors often employ dimensional reduction methods to decrease the complexity of the algorithms without compromising the characteristic information within the data. Finally, the spectral-spatial CNN models consider both the spectral and spatial information of the HS cube [93, 100, 189]. Several strategies have been proposed in the state of the art, suggesting different architectures for conducting spectral-spatial analysis. However, the direct strategy to implement this type of architecture is through the use of 3D kernels in the convolutional stages, where the 3D filters are capable of extracting spatial-spectral feature volumes [93, 189].

B.5 Auto Encoders

Autoencoder is another widely used ML technique for HS data analysis, whose main idea is dimensionality reduction and feature extraction in an unsupervised manner [93, 217]. The main characteristic of

these networks is their ability to transform input data from one domain to another, generating compressed, expanded, or even equal-dimensional outputs to minimize the difference between the original data and the reconstructions [93, 218]. This transformation is achieved through two main components: the encoder and the decoder, which are connected by a bottleneck layer, generating the traditional architecture of this type of network, as shown in Figure B.4 [93, 217, 218]. In addition, autoencoders are also used for data generation and anomaly detection tasks, but due to their characteristics, they are not typically used directly for classification [93].

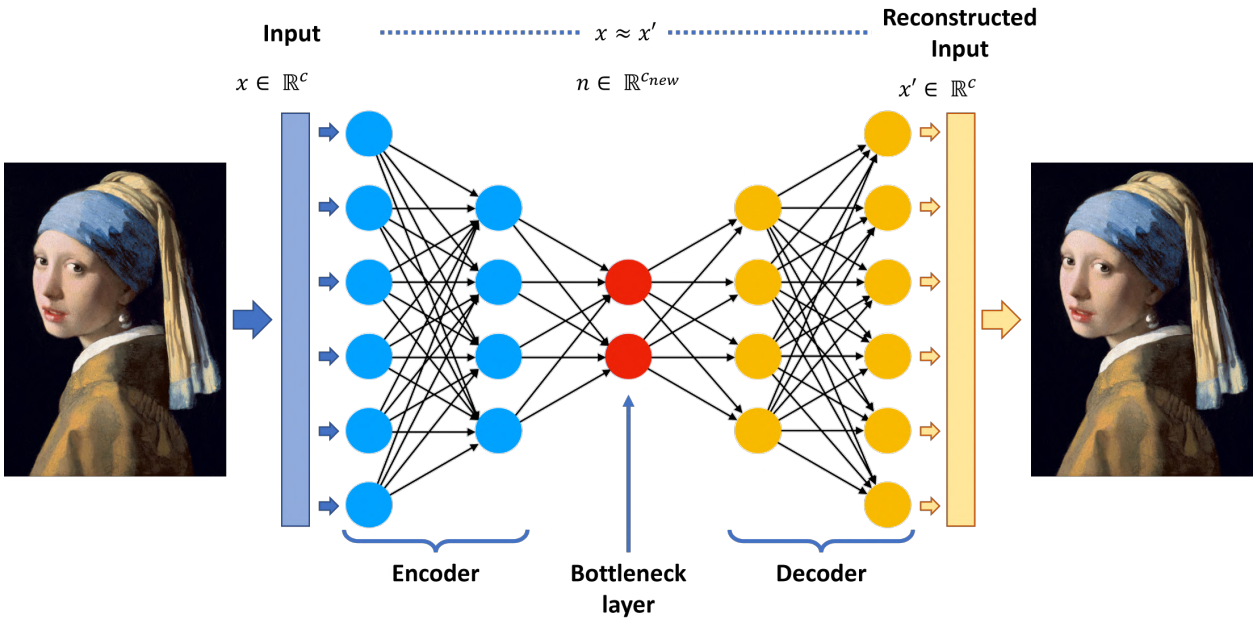


Figure B.4: Traditional representation of an autoencoder, composed of two main parts: an encoder and a decoder.

B.6 Deep Belief Networks

Deep Belief Network (DBN) is a ML model used in HS processing, as it can automatically learn relevant features and improve performance in HS image analysis tasks [219, 220]. The DBN is made up of multiple layers of processing units, where the hidden layers form a structure of restricted Boltzmann machines (RBM) [219, 220]. Additionally, the DBN is trained in an unsupervised manner, which means that it does not require prior labels or annotations [219, 220]. This is especially useful in the case of HS data, where obtaining a fully labeled training set can be costly and labor intensive [219, 220].

B.7 Recurrent Neural Networks

Recurrent Neural Networks (RNN) are another variant of NNs commonly used in the evaluation of HS information. Regardless, RNNs have recurrent connections in their architecture, which allow them to store information from previous data, generating an internal memory [93, 189]. This structure makes RNN a model capable of learning sequences with an internal memory state to produce an association between incoming and previous data [189, 221]. This hidden state is updated as the network receives new data and propagates through the recurrent layers, allowing the network to learn patterns along the sequence. This memory feature enables RNNs to predict future events based on previous data [93, 189, 221]. In this case, HS data can be treated as if they were video sequences, considering spectral bands as frames and applying an RNN to model the dependencies between bands [189, 221]. This allows for the identification of spectral features within the information of each pixel.

B.8 Residual Networks

Residual Networks (ResNet) are a type of ML architecture that stands out for incorporating residual connections, which allow the flow of information between the layers of the network [222]. In the context of HS images, ResNets are employed to tackle the challenges posed by the high data dimensionality and complex spectral features [223]. These networks are capable of mitigating the performance degradation problem that can occur in deep networks. Furthermore, ResNet architectures are easier to optimize compared to other architectures due to their smaller number of parameters [217, 223].

B.9 Vision Transformers

Vision Transformers (ViTs) are a type of ML model used for image recognition and other computer vision tasks. ViTs are based on the transformer architecture (see Figure B.5), which in turn relies on a procedure called attention that focuses on the context in sequential data. Initially, transformers were proposed as methods for natural language processing (NLP) tasks such as automatic translation, language modeling, and speech recognition [224]. The Transformer architecture was originally developed in [225] and is based on the encoder-decoder model, which transforms a given sequence of elements into another sequence. The main motivation behind Transformers was to enable parallel processing of words in a sentence, which was not possible in models like RNNs that process words one by one.

ViTs take inspiration from the original transformer architecture and apply it to the field of computer vision, as described by Dosovitskiy et al. in [6]. In this novel model, a standard transformer is directly applied by dividing the image into patches, without focusing on individual pixels. The Transformer then encodes the sequence of embeddings for these patches, transforming them into attention tokens to learn the relationships between different patches of the image. This information is utilized for tasks such as image

classification or other applications.

The advantages of ViTs have led to their integration in HSI analysis, surpassing CNNs in terms of classification metrics, processing time, and computational cost [225, 226]. Promising results have been demonstrated in various studies. For example, Bazi et al. [224] proposed the use of multihead attention mechanisms, while Gao et al. [227] developed a new architecture called Transformer-iN-Transformer for region-based analysis in HSI. Another approach by Hu et al. [228] introduced a computationally efficient unsupervised contrastive learning model. Nonetheless, it is important to note that, while transformers have shown excellent performance, most evaluations have been performed using synthetic high-resolution images, and limited studies have explored the use of transformers with experimental data.

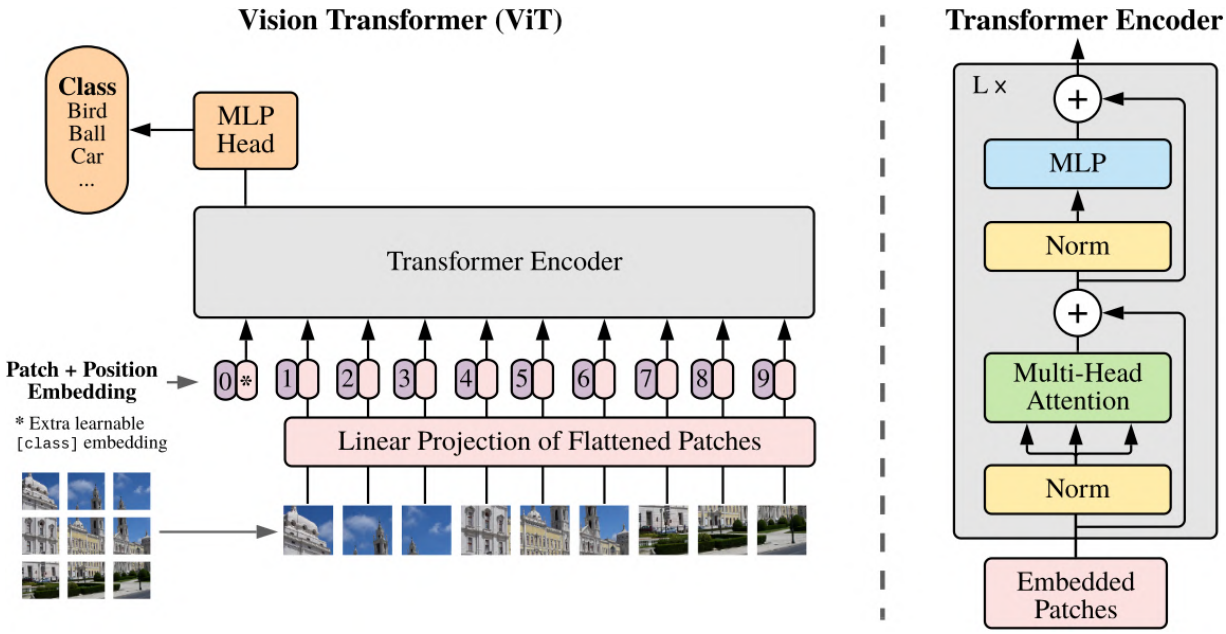


Figure B.5: The Vision Transformer architecture [6].

Chapter C

Performance Metrics

The performance of the methods studied in this dissertation will be compared to real datasets labeled by clinical experts. Additionally, to evaluate the effectiveness of classifiers in scenarios with unbalanced data, several metrics have been selected for quantitative analysis [229, 230]. These metrics include accuracy, sensitivity, specificity, F1-score, and the Matthews correlation coefficient (MCC) [229, 230, 231]. These metrics are based on the following variables to represent the possible outcomes in a binary classification problem:

- True positive (TP): instances correctly classified as positive.
- False positive (FP): instances wrongly classified as positive.
- True negative (TN): instances correctly classified as negative.
- False negative (FN): instances wrongly classified as negative.

The aforementioned classification metrics are calculated as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{C.1}$$

$$\text{sensitivity} = \frac{TP}{TP + FN} \tag{C.2}$$

$$\text{specificity} = \frac{TN}{FP + TN} \tag{C.3}$$

$$\text{F1-score} = \frac{2 \times TP}{2 \times TP + FN + FP} \tag{C.4}$$

$$\text{MCC} = \frac{1}{2}\left(1 + \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}\right) \tag{C.5}$$

In classification tasks, accuracy, sensitivity, and specificity are commonly used metrics to assess precision in terms of positive and negative true values [229, 230]. On the other hand, the F1-score quantifies the precision of a test, where a higher value indicates higher precision and sensitivity [229, 230, 231]. The MCC is employed as a measure of classification quality, considering both true and false positives along with negatives. The MCC is considered a balanced metric that can be applied even when classes have significantly different sizes. Essentially, the MCC (normalized) serves as a correlation coefficient between the observed and predicted binary classifications, producing a value ranging from 0 to 1 [229, 230, 231]. A coefficient of 1 signifies a perfect prediction, 0.5 implies that there is no improvement over a random prediction, and 0 indicates complete disagreement between prediction and observation [231].

Bibliography

[1] S. Ortega, H. Fabelo, M. Halicek, R. Camacho, M. de la Luz Plaza, G. M. Callicó, and B. Fei, "Hyperspectral superpixel-wise glioblastoma tumor detection in histological samples," Appl. Sci., vol. 10, no. 13, 2020.

[2] S. Ortega, M. Halicek, H. Fabelo, R. Camacho, M. d. l. L. Plaza, F. Godtliebsen, G. M Callicó, and B. Fei, "Hyperspectral imaging for the detection of glioblastoma tumor cells in h&e slides using convolutional neural networks," Sensors, vol. 20, no. 7, p. 1911, 2020.

[3] H. Fabelo, et. al, "Deep learning-based framework for in vivo identification of glioblastoma tumor using hyperspectral images of human brain," Sensors, vol. 19, no. 4, p. 920, 2019.

[4] B. Martinez, R. Leon, H. Fabelo, S. Ortega, J. F. Piñeiro, A. Szolna, M. Hernandez, C. Espino, A. J. O'shanahan, D. Carrera, S. Bisshopp, C. Sosa, M. Marquez, R. Camacho, M. de la Luz Plaza, J. Morera, and G. M. Callico, "Most relevant spectral bands identification for brain cancer detection using hyperspectral imaging," Sensors (Switzerland), vol. 19, no. 24, 2019.

[5] I. A. Cruz-Guerrero, R. Leon, D. U. Campos-Delgado, S. Ortega, H. Fabelo, and G. M. Callico, "Classification of Hyperspectral In Vivo Brain Tissue Based on Linear Unmixing," Appl. Sci., vol. 10, no. 16, p. 5686, 2020.

[6] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," arXiv, 2020.

[7] S. McGuire, "World cancer report 2018. geneva, switzerland: World health organization, international agency for research on cancer, who press, 2018," 2018.

[8] L. Alvarenga, "Real academia española y asociación de academias de la lengua española: Diccionario panhispánico de dudas," Realidad: Revista de Ciencias Sociales y Humanidades, vol. 23, no. 108, pp. 331–332, 2017.

[9] H. K. Matthews, C. Bertoli, and R. A. de Bruin, "Cell cycle control in cancer," Nature Reviews Molecular Cell Biology, vol. 23, no. 1, pp. 74–88, 2022.

[10] B. P. Sedeño and P. C. L. Jiménez, "Generalidades en oncología radioterápica (i): radiobiologia," Biocáncer, vol. 1, pp. 1–21, 2004.

[11] J. Arias, Generalidades médico-quirúrgicas. Editorial Tebar, 2001.

[12] I. H. Khan and M. Javaid, "Big data applications in medical field: A literature review," Journal of Industrial Integration and Management, vol. 6, no. 01, pp. 53–69, 2021.

[13] N. Goel, A. Yadav, and B. M. Singh, "Medical image processing: a review," 2016 Second International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH), pp. 57–62, 2016.

[14] F. A. Fellner, "Introducing cinematic rendering: a novel technique for post-processing medical imaging data," Journal of Biomedical Science and Engineering, vol. 9, no. 3, pp. 720–726, 2016.

[15] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," Annual review of biomedical engineering, vol. 19, p. 221, 2017.

[16] M. Bakator and D. Radosav, "Deep learning and medical diagnosis: A review of literature," Multimodal Technologies and Interaction, vol. 2, no. 3, p. 47, 2018.

[17] S. Bagheri, "Hyperspectral remote sensing of nearshore water quality," 2017.

[18] Y. Gu, T. Liu, G. Gao, G. Ren, Y. Ma, J. Chanussot, and X. Jia, "Multimodal hyperspectral remote sensing: an overview and perspective," Sci. China Inf. Sci. 2021 642, vol. 64, pp. 1–24, jan 2021.

[19] R. Vejarano, R. Siche, and W. Tesfaye, "Evaluation of biological contaminants in foods by hyperspectral imaging: A review," Int. J. Food Prop., vol. 20, pp. 1–34, jun 2017.

[20] M. Al Ktash, M. Stefanakis, B. Boldrini, E. Ostertag, and M. Brecht, "Characterization of Pharmaceutical Tablets Using UV Hyperspectral Imaging as a Rapid In-Line Analysis Tool," Sensors, vol. 21, p. 4436, jun 2021.

[21] J. C. Montes-Herrera, E. Cimoli, V. Cummings, N. Hill, A. Lucieer, and V. Lucieer, "Underwater Hyperspectral Imaging (UHI): A Review of Systems and Applications for Proximal Seafloor Ecosystem Studies," Remote Sens., vol. 13, p. 3451, aug 2021.

[22] G. Lassalle, S. Fabre, A. Credoz, R. Hédacq, D. Dubucq, and A. Elger, "Mapping leaf metal content over industrial brownfields using airborne hyperspectral imaging and optimized vegetation indices," Sci. Rep., vol. 11, p. 2, dec 2021.

[23] T. Arnold, M. De Biasio, T. Bereczki, M. Baumgart, and A. Horn, "Development of inspection system for the detection and analysis of solid particles and oil droplets in process water of the petrochemical industry using hyperspectral imaging and fluorescence imaging," in Algorithms, Technol. Appl. Multispectral Hyperspectral Imaging XXVIII (D. W. Messinger and M. Velez-Reyes, eds.), p. 37, SPIE, may 2022.

[24] M. Shimoni, R. Haelterman, and C. Perneel, "Hypersectral imaging for military and security applications: Combining Myriad processing and sensing techniques," IEEE Geosci. Remote Sens. Mag., vol. 7, pp. 101–117, jun 2019.

[25] S. Ortega, M. Halicek, H. Fabelo, G. M. Callico, and B. Fei, "Hyperspectral and multispectral imaging in digital and computational pathology: a systematic review [Invited]," Biomed. Opt. Express, vol. 11, no. 6, p. 3195, 2020.

[26] B. Fei, Hyperspectral imaging in medical applications, vol. 32. Elsevier, 2020.

[27] M. Govender, K. Chetty, and H. Bulcock, "A review of hyperspectral remote sensing and its application in vegetation and water resource studies," Water {SA}, vol. 33, no. 2, 2009.

[28] M. Kamruzzaman and D.-W. Sun, "Introduction to Hyperspectral Imaging Technology," in Comput. Vis. Technol. Food Qual. Eval., pp. 111–139, Elsevier, jan 2016.

[29] A. Holmer, F. Tetschke, J. Marotz, H. Malberg, W. Markgraf, C. Thiele, and A. Kulcke, "Oxygenation and perfusion monitoring with a hyperspectral camera system for chemical based tissue analysis of skin and organs," Physiol. Meas., vol. 37, p. 2064, oct 2016.

[30] L. Rey-Barroso, S. Peña-Gutiérrez, C. Yáñez, F. J. Burgos-Fernández, M. Vilaseca, and S. Royo, "Optical Technologies for the Improvement of Skin Cancer Diagnosis: A Review," Sensors, vol. 21, p. 252, jan 2021.

[31] W. Markgraf, J. Lilienthal, P. Feistel, C. Thiele, and H. Malberg, "Algorithm for Mapping Kidney Tissue Water Content during Normothermic Machine Perfusion Using Hyperspectral Imaging," Algorithms, vol. 13, p. 289, nov 2020.

[32] J. K. H. Lim, Q.-X. Li, T. Ryan, P. Bedggood, A. Metha, A. J. Vingrys, B. V. Bui, and C. T. O. Nguyen, "Retinal hyperspectral imaging in the 5xFAD mouse model of Alzheimer's disease," Sci. Rep., vol. 11, p. 6387, dec 2021.

[33] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," J. Biomed. Opt., vol. 19, no. 1, p. 010901, 2014.

[34] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA. Cancer J. Clin., vol. 71, pp. 209–249, may 2021.

[35] M. Zambrano-Román, J. R. Padilla-Gutiérrez, Y. Valle, J. F. Muñoz-Valle, and E. Valdés-Alvarado, "Non-Melanoma Skin Cancer: A Genetic Update and Future Perspectives," Cancers (Basel)., vol. 14, p. 2371, may 2022.

[36] H. Tsao, J. M. Olazagasti, K. M. Cordoro, J. D. Brewer, S. C. Taylor, J. S. Bordeaux, M. M. Chren, A. J. Sober, C. Tegeler, R. Bhushan, and W. S. Begolka, "Early detection of melanoma: Reviewing the ABCDEs American Academy of Dermatology Ad Hoc Task Force for the ABCDEs of Melanoma," J. Am. Acad. Dermatol., vol. 72, no. 4, pp. 717–723, 2015.

[37] R. Jolivot, Y. Benezeth, and F. Marzani, "Skin parameter map retrieval from a dedicated multispectral imaging system applied to dermatology/cosmetology," Int. J. Biomed. Imaging, vol. 2013, 2013.

[38] M. Elbaum, A. W. Kopf, H. S. Rabinovitz, R. G. Langley, H. Kamino, M. C. Mihm, A. J. Sober, G. L. Peck, A. Bogdan, D. Gutkowicz-Krusin, M. Greenebaum, S. Keem, M. Oliviero, and S. Wang, "Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: A feasibility study," J. Am. Acad. Dermatol., vol. 44, no. 2, pp. 207–218, 2001.

[39] N. Neittaanmäki-Perttu, M. Grönroos, L. Jeskanen, I. Pölönen, A. Ranki, O. Saksela, and E. Snellman, "Delineating margins of lentigo maligna using a hyperspectral imaging system," Acta Derm. Venereol., vol. 95, no. 5, pp. 549–552, 2015.

[40] M. Salmivuori, N. Neittaanmäki, I. Pölönen, L. Jeskanen, E. Snellman, and M. Grönroos, "Hyperspectral imaging system in the delineation of Ill-defined basal cell carcinomas: a pilot study," J. Eur. Acad. Dermatology Venereol., vol. 33, pp. 71–78, jan 2019.

[41] L. A. Zherdeva, I. A. Bratchenko, O. O. Myakinin, A. A. Moryatov, S. V. Kozlov, and V. P. Zakharov, "In vivo hyperspectral imaging and differentiation of skin cancer," in Optics in Health Care and Biomedical Optics VII, vol. 10024, pp. 658–665, SPIE, 2016.

[42] A. M. Hosking, B. J. Coakley, D. Chang, F. Talebi-Liasi, S. Lish, S. W. Lee, A. M. Zong, I. Moore, J. Browning, S. L. Jacques, J. G. Krueger, K. M. Kelly, K. G. Linden, and D. S. Gareau, "Hyperspectral imaging in automated digital dermoscopy screening for melanoma," Lasers Surg. Med., vol. 51, pp. 214–222, mar 2019.

[43] H. Fabelo, G. Carretero, P. Almeida, A. Garcia, J. A. Hernandez, F. Godtliebsen, V. Melian, B. Martinez, P. Beltran, S. Ortega, M. Marrero, G. M. Callico, R. Sarmiento, and I. Castano, "Dermatologic Hyperspectral Imaging System for Skin Cancer Diagnosis Assistance," 2019 34th Conf. Des. Circuits Integr. Syst. DCIS 2019, pp. 5–10, 2019.

[44] R. Leon, B. Martinez-Vega, H. Fabelo, S. Ortega, V. Melian, I. Castaño, G. Carretero, P. Almeida, A. Garcia, E. Quevedo, J. A. Hernandez, B. Clavo, and G. M. Callico, "Non-invasive skin cancer diagnosis using hyperspectral imaging for in-situ clinical support," J. Clin. Med., vol. 9, no. 6, 2020.

[45] L. A. Courtenay, D. González-Aguilera, S. Lagüela, S. del Pozo, C. Ruiz-Mendez, I. Barbero-García, C. Román-Curto, C. Román-Curto, J. Cañueto, and J. Cañueto, "Hyperspectral imaging and robust statistics in non-melanoma skin cancer analysis," Biomed. Opt. Express, Vol. 12, Issue 8, pp. 5107-5127, vol. 12, pp. 5107–5127, aug 2021.

[46] L. A. Courtenay, D. González-Aguilera, S. Lagüela, S. Del Pozo, C. Ruiz, I. Barbero-García, C. Román-Curto, J. Cañueto, C. Santos-Durán, and M. E. Cardeñoso-álvarez, "Deep Convolutional Neural Support Vector Machines for the Classification of Basal Cell Carcinoma Hyperspectral Signatures," J. Clin. Med. 2022, Vol. 11, Page 2315, vol. 11, p. 2315, apr 2022.

[47] A. P. Patel, J. L. Fisher, E. Nichols, F. Abd-Allah, J. Abdela, A. Abdelalim, H. N. Abraha, D. Agius, F. Alahdab, and T. Alam, "Global, regional, and national burden of brain and other CNS cancer, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016," Lancet Neurol., vol. 18, pp. 376–393, apr 2019.

[48] R. S. D'Amico, Z. K. Englander, P. Canoll, and J. N. Bruce, "Extent of Resection in Glioma–A Review of the Cutting Edge," jul 2017.

[49] E. Belykh, N. L. Martirosyan, K. Yagmurlu, E. J. Miller, J. M. Eschbacher, M. Izadyyazdanabadi, L. A. Bardonova, V. A. Byvaltsev, P. Nakaji, and M. C. Preul, "Intraoperative Fluorescence Imaging for Personalized Brain Tumor Resection: Current State and Future Directions," oct 2016.

[50] L. Giannoni, F. Lange, and I. Tachtsidis, "Investigation of the quantification of hemoglobin and cytochrome-c-oxidase in the exposed cortex with near-infrared hyperspectral imaging: a simulation study," https://doi.org/10.1117/1.JBO.25.4.046001, vol. 25, p. 046001, apr 2020.

[51] H. Fabelo, S. Ortega, E. Casselden, J. Loh, H. Bulstrode, A. Zolnourian, P. Grundy, G. M. Callico, D. Bulters, and R. Sarmiento, "SVM optimization for brain tumor identification using infrared spectroscopic samples," Sensors (Switzerland), vol. 18, no. 12, 2018.

[52] H. Fabelo, S. Ortega, A. Szolna, D. Bulters, J. F. Pineiro, S. Kabwama, A. J-O'Shanahan, H. Bulstrode, S. Bisshopp, B. R. Kiran, D. Ravi, R. Lazcano, D. Madronal, C. Sosa, C. Espino, M. Marquez, M. De La Luz Plaza, R. Camacho, D. Carrera, M. Hernandez, G. M. Callico, J. Morera Molina, B. Stanciulescu, G. Z. Yang, R. Salvador Perea, E. Juarez, C. Sanz, and R. Sarmiento, "In-Vivo Hyperspectral Human Brain Image Database for Brain Cancer Detection," IEEE Access, vol. 7, pp. 39098–39116, 2019.

[53] Q. Hao, Y. Pei, R. Zhou, B. Sun, J. Sun, S. Li, and X. Kang, "Fusing Multiple Deep Models for in Vivo Human Brain Hyperspectral Image Classification to Identify Glioblastoma Tumor," IEEE Trans. Instrum. Meas., vol. 70, 2021.

[54] G. Urbanos, A. Martín, G. Vázquez, M. Villanueva, M. Villa, L. Jimenez-Roldan, M. Chavarrías, A. Lagares, E. Juárez, and C. Sanz, "Supervised Machine Learning Methods and Hyperspectral Imaging Techniques Jointly Applied for Brain Cancer Classification," Sensors 2021, Vol. 21, Page 3827, vol. 21, p. 3827, may 2021.

[55] C. Caredda, L. Mahieu-Williame, R. Sablong, M. Sdika, J. Guyotat, and B. Montcel, "Optimal Spectral Combination of a Hyperspectral Camera for Intraoperative Hemodynamic and Metabolic Brain Mapping," Appl. Sci. 2020, Vol. 10, Page 5158, vol. 10, p. 5158, jul 2020.

[56] K. Iwaki, S. Takagishi, K. Arimura, M. Murata, T. Chiba, A. Nishimura, N. Ren, and K. Iihara, "A Novel Hyperspectral Imaging System for Intraoperative Prediction of Cerebral Hyperperfusion Syndrome after Superficial Temporal Artery-Middle Cerebral Artery Anastomosis in Patients with Moyamoya Disease," Cerebrovasc. Dis., vol. 50, pp. 208–215, apr 2021.

[57] M. Arnold, C. C. Abnet, R. E. Neale, J. Vignat, E. L. Giovannucci, K. A. McGlynn, and F. Bray, "Global Burden of 5 Major Types of Gastrointestinal Cancer," Gastroenterology, vol. 159, pp. 335–349.e15, jul 2020.

[58] W. Du, N. Rao, D. Liu, H. Jiang, C. Luo, Z. Li, T. Gan, and B. Zeng, "Review on the Applications of Deep Learning in the Analysis of Gastrointestinal Endoscopy Images," IEEE Access, vol. 7, pp. 142053–142069, 2019.

[59] J. Lin, N. T. Clancy, J. Qi, Y. Hu, T. Tatla, D. Stoyanov, L. Maier-Hein, and D. S. Elson, "Dual-modality endoscopic probe for tissue surface shape reconstruction and hyperspectral imaging enabled by deep neural networks," Med. Image Anal., vol. 48, pp. 162–176, aug 2018.

[60] I. C. Wu, H. Y. Syu, C. P. Jen, M. Y. Lu, Y. T. Chen, M. T. Wu, C. T. Kuo, Y. Y. Tsai, and H. C. Wang, "Early identification of esophageal squamous neoplasm by hyperspectral endoscopic imaging," Sci. Reports 2018 81, vol. 8, pp. 1–10, sep 2018.

[61] J. Yoon, J. Joseph, D. J. Waterhouse, C. Borzy, K. Siemens, S. Diamond, V. L. Tsikitis, and S. E. Bohndiek, "First experience in clinical application of hyperspectral endoscopy for evaluation of colonic polyps," J. Biophotonics, vol. 14, p. e202100078, sep 2021.

[62] H. Köhler, A. Kulcke, M. Maktabi, Y. Moulla, B. Jansen-Winkeln, M. Barberio, M. Diana, I. Gockel, T. Neumuth, and C. Chalopin, "Laparoscopic system for simultaneous high-resolution video and rapid hyperspectral imaging in the visible and near-infrared spectral range," https://doi.org/10.1117/1.JBO.25.8.086004, vol. 25, p. 086004, aug 2020.

[63] D. Sato, T. Takamatsu, M. Umezawa, Y. Kitagawa, K. Maeda, N. Hosokawa, K. Okubo, M. Kamimura, T. Kadota, T. Akimoto, T. Kinoshita, T. Yano, T. Kuwata, H. Ikematsu, H. Takemura, H. Yokota, and K. Soga, "Distinction of surgically resected gastrointestinal stromal tumor by near-infrared hyperspectral imaging," Sci. Reports 2020 101, vol. 10, pp. 1–9, dec 2020.

[64] B. Jansen-Winkeln, N. Holfert, H. Köhler, Y. Moulla, J. P. Takoh, S. M. Rabe, M. Mehdorn, M. Barberio, C. Chalopin, T. Neumuth, and I. Gockel, "Determination of the transection margin during colorectal resection with hyperspectral imaging (HSI)," Int. J. Colorectal Dis., vol. 34, pp. 731–739, apr 2019.

[65] B. Jansen-winkeln, M. Barberio, C. Chalopin, K. Schierle, M. Diana, H. Köhler, I. Gockel, and M. Maktabi, "Feedforward Artificial Neural Network-Based Colorectal Cancer Detection Using Hyperspectral Imaging: A Step towards Automatic Optical Biopsy," Cancers 2021, Vol. 13, Page 967, vol. 13, p. 967, feb 2021.

[66] D. E. Johnson, B. Burtness, C. R. Leemans, V. W. Y. Lui, J. E. Bauman, and J. R. Grandis, "Head and neck squamous cell carcinoma," Nat. Rev. Dis. Prim. 2020 61, vol. 6, pp. 1–22, nov 2020.

[67] M. Halicek, J. V. Little, X. Wang, A. Y. Chen, and B. Fei, "Optical biopsy of head and neck cancer using hyperspectral imaging and convolutional neural networks," https://doi.org/10.1117/1.JBO.24.3.036007, vol. 24, p. 036007, mar 2019.

[68] D. Eggert, M. Bengs, S. Westermann, N. Gessert, A. O. Gerstner, N. A. Mueller, J. Bewarder, A. Schlaefer, C. Betz, and W. Laffers, "In vivo detection of head and neck tumors by hyperspectral imaging combined with deep learning methods," J. Biophotonics, vol. 15, p. e202100167, mar 2022.

[69] C. Klein, Q. Zeng, F. Arbaretaz, E. Devêvre, J. Calderaro, N. Lomenie, and M. C. Maiuri, "Artificial intelligence for solid tumour diagnosis in digital pathology," Br. J. Pharmacol., vol. 178, pp. 4291–4315, nov 2021.

[70] S. Ortega, R. Guerra, M. DIaz, H. Fabelo, S. Lopez, G. M. Callico, and R. Sarmiento, "Hyperspectral Push-Broom Microscope Development and Characterization," IEEE Access, vol. 7, pp. 122473–122491, 2019.

[71] S. Ortega, G. Callico, M. Plaza, R. Camacho, H. Fabelo, and R. Sarmiento, "Hyperspectral database of pathological in-vitro human brain samples to detect carcinogenic tissues," in Proc. - Int. Symp. Biomed. Imaging, vol. 2016-June, 2016.

[72] S. Ortega, H. Fabelo, R. Camacho, M. de la Luz Plaza, G. M. Callicó, and R. Sarmiento, "Detecting brain tumor in pathological slides using hyperspectral imaging," Biomed. Opt. Express, vol. 9, no. 2, p. 818, 2018.

[73] L. Ma, M. Halicek, X. Zhou, J. D. Dormer, and B. Fei, "Hyperspectral Microscopic Imaging for Automatic Detection of Head and Neck Squamous Cell Carcinoma Using Histologic Image and Machine Learning," Proc. SPIE–the Int. Soc. Opt. Eng., vol. 11320, p. 31, mar 2020.

[74] M. M. Souza, F. A. Carvalho, E. F. V. Sverzut, M. B. Requena, M. R. Garcia, and S. Pratavieira, "Hyperspectral Imaging System for Tissue Classification in H&E-Stained Histological Slides," in 2021 Sbfot. Int. Opt. Photonics Conf. (sbfot. IOPC), pp. 1–4, IEEE, may 2021.

[75] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," Ieee Access, vol. 6, pp. 14118–14129, 2018.

[76] C.-I. Chang, Hyperspectral data processing: algorithm design and analysis. John Wiley & Sons, 2013.

[77] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 5, no. 2, pp. 354–379, 2012.

[78] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," Journal of biomedical optics, vol. 19, no. 1, p. 010901, 2014.

[79] N. Dobigeon, Y. Altmann, N. Brun, and S. Moussaoui, Linear and Nonlinear Unmixing in Hyperspectral Imaging, vol. 30. Elsevier, 2016.

[80] J. M. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in Image and Signal Processing for Remote Sensing XV, vol. 7477, p. 74770I, International Society for Optics and Photonics, 2009.

[81] N. Keshava and J. F. Mustard, "Spectral unmixing," IEEE signal processing magazine, vol. 19, no. 1, pp. 44–57, 2002.

[82] N. Dobigeon, J. Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," IEEE Signal Process. Mag., vol. 31, no. 1, pp. 82–94, 2014.

[83] L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," IEEE Transactions on Image Processing, vol. 25, no. 8, pp. 3890–3905, 2016.

[84] D. U. Campos-Delgado, O. Gutierrez-Navarro, J. J. Rico-Jimenez, E. Duran-Sierra, H. Fabelo, S. Ortega, G. Callico, and J. A. Jo, "Extended Blind End-Member and Abundance Extraction for Biomedical Imaging Applications," IEEE Access, vol. 7, pp. 178539–178552, 2019.

[85] G. ElMasry and D. W. Sun, "Principles of Hyperspectral Imaging Technology," Hyperspectral Imaging Food Qual. Anal. Control, pp. 3–43, 2010.

[86] J. Ma, D. W. Sun, H. Pu, J. H. Cheng, and Q. Wei, "Advanced Techniques for Hyperspectral Imaging in the Food Industry: Principles and Recent Applications," Annu. Rev. Food Sci. Technol., vol. 10, pp. 197–220, 2019.

[87] J. Jia, Y. Wang, J. Chen, R. Guo, R. Shu, and J. Wang, "Status and application of advanced airborne hyperspectral imaging technology: A review," Infrared Phys. Technol., vol. 104, no. November 2019, p. 103115, 2020.

[88] T. Adão, J. Hruška, L. Pádua, J. Bessa, E. Peres, R. Morais, and J. J. Sousa, "Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry," Remote Sens., vol. 9, no. 11, 2017.

[89] G. ElMasry, M. Kamruzzaman, D. W. Sun, and P. Allen, "Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review," Crit. Rev. Food Sci. Nutr., vol. 52, no. 11, pp. 999–1023, 2012.

[90] X. Fu and J. Chen, "A Review of Hyperspectral Imaging for Chicken Meat Safety and Quality Evaluation: Application, Hardware, and Software," mar 2019.

[91] B. Mahesh, "Machine learning algorithms-a review," International Journal of Science and Research (IJSR).[Internet], vol. 9, pp. 381–386, 2020.

[92] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, "Secure and robust machine learning for healthcare: A survey," IEEE Reviews in Biomedical Engineering, vol. 14, pp. 156–180, 2020.

[93] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," ISPRS J. Photogramm. Remote Sens., vol. 158, no. November, pp. 279–317, 2019.

[94] E. Zhang, X. Zhang, S. Yang, and S. Wang, "Improving hyperspectral image classification using spectral information divergence," IEEE Geoscience and Remote Sensing Letters, vol. 11, pp. 249–253, 2014.

[95] F. Palsson, J. Sigurdsson, J. R. Sveinsson, and M. O. Ulfarsson, "Neural network hyperspectral unmixing with spectral information divergence objective," in 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), vol. 2017-July, pp. 755–758, Institute of Electrical and Electronics Engineers Inc., 12 2017.

[96] S. Addamani, "Spectral angle mapper algorithm for remote sensing image classification," IJISET-International Journal of Innovative Science, Engineering & Technology, vol. 1, 2014.

[97] C. M. Bishop, Pattern recognition and machine learning. Springer-Verlag, 2006.

[98] J. Waring, C. Lindvall, and R. Umeton, "Automated machine learning: Review of the state-of-the-art and opportunities for healthcare," Artificial intelligence in medicine, vol. 104, p. 101822, 2020.

[99] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable ai: A review of machine learning interpretability methods," Entropy, vol. 23, no. 1, p. 18, 2020.

[100] W. Lv and X. Wang, "Overview of hyperspectral image classification," Journal of Sensors, vol. 2020, 2020.

[101] R. N. Clark, G. A. Swayze, R. Wise, K. E. Livo, T. M. Hoefen, R. F. Kokaly, and S. J. Sutley, "Usgs digital spectral library splib06a," US Geological Survey, Open File Report, vol. 3395, 2003.

[102] R. G. Resmini, M. E. Kappus, W. S. Aldrich, J. C. Harsanyi, and M. Anderson, "Mineral mapping with hyperspectral digital imagery collection experiment (hydice) sensor data at cuprite, nevada, u.s.a.," International Journal of Remote Sensing, vol. 18, no. 7, pp. 1553–1570, 1997.

[103] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 6, pp. 2014–2039, 2011.

[104] S. Ren, S. Wan, X. Gu, X. Shu, and H. Xu, "Hyperspectral image classification based on robust discriminative extraction of multiple spectral-spatial features," International Journal of Remote Sensing, vol. 40, no. 15, pp. 5812–5834, 2019.

[105] R. Leon, H. Fabelo, S. Ortega, J. F. Piñeiro, A. Szolna, M. Hernandez, C. Espino, A. J. O'Shanahan, D. Carrera, S. Bisshopp, C. Sosa, M. Marquez, J. Morera, B. Clavo, and G. M. Callico, "VNIR–NIR hyperspectral imaging fusion targeting intraoperative brain cancer detection," Sci. Rep., vol. 11, no. 1, pp. 1–12, 2021.

[106] D. Lin, L. Shen, M. Luo, K. Zhang, J. Li, Q. Yang, F. Zhu, D. Zhou, S. Zheng, Y. Chen, et al., "Circulating tumor cells: Biology and clinical significance," Signal transduction and targeted therapy, vol. 6, no. 1, p. 404, 2021.

[107] A. Vollmann-Zwerenz, V. Leidgens, G. Feliciello, C. A. Klein, and P. Hau, "Tumor cell invasion in glioblastoma," International journal of molecular sciences, vol. 21, no. 6, p. 1932, 2020.

[108] B. Sipos and G. Klöppel, "Glucagon cell hyperplasia and neoplasia: a recently recognized endocrine receptor disease," Endocrine-Related Cancer, vol. 30, no. 8, 2023.

[109] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, "Brain tumor detection and classification using machine learning: a comprehensive survey," Complex & intelligent systems, pp. 1–23, 2021.

[110] A. R. Khan, S. Khan, M. Harouni, R. Abbasi, S. Iqbal, and Z. Mehmood, "Brain tumor segmentation using k-means clustering and deep learning with synthetic data augmentation for classification," Microscopy Research and Technique, vol. 84, no. 7, pp. 1389–1399, 2021.

[111] E. F. Rodriguez, R. Jones, D. Miller, and F. J. Rodriguez, "Neurogenic tumors of the mediastinum," Seminars in diagnostic pathology, vol. 37, no. 4, pp. 179–186, 2020.

[112] J. Ferlay, M. Colombet, I. Soerjomataram, D. M. Parkin, M. Piñeros, A. Znaor, and F. Bray, "Cancer statistics for the year 2020: An overview," International journal of cancer, vol. 149, no. 4, pp. 778–789, 2021.

[113] R. L. Siegel, K. D. Miller, H. E. Fuchs, A. Jemal, et al., "Cancer statistics, 2021," Ca Cancer J Clin, vol. 71, no. 1, pp. 7–33, 2021.

[114] A. P. Patel, J. L. Fisher, E. Nichols, F. Abd-Allah, J. Abdela, A. Abdelalim, H. N. Abraha, D. Agius, F. Alahdab, T. Alam, et al., "Global, regional, and national burden of brain and other cns cancer, 1990–2016: a systematic analysis for the global burden of disease study 2016," The Lancet Neurology, vol. 18, no. 4, pp. 376–393, 2019.

[115] D. N. Louis, A. Perry, P. Wesseling, D. J. Brat, I. A. Cree, D. Figarella-Branger, C. Hawkins, H. Ng, S. M. Pfister, G. Reifenberger, et al., "The 2021 who classification of tumors of the central nervous system: a summary," Neuro-oncology, vol. 23, no. 8, pp. 1231–1251, 2021.

[116] A. M. Di Giacomo, M. J. Mair, M. Ceccarelli, A. Anichini, R. Ibrahim, M. Weller, M. Lahn, A. M. Eggermont, B. Fox, and M. Maio, "Immunotherapy for brain metastases and primary brain tumors," European Journal of Cancer, vol. 179, pp. 113–120, 2023.

[117] F. J. Voskuil, J. Vonk, B. van der Vegt, S. Kruijff, V. Ntziachristos, P. J. van der Zaag, M. J. Witjes, and G. M. van Dam, "Intraoperative imaging in pathology-assisted surgery," Nature biomedical engineering, vol. 6, no. 5, pp. 503–514, 2022.

[118] N. Verburg and P. C. de Witt Hamer, "State-of-the-art imaging for glioma surgery," Neurosurgical review, vol. 44, no. 3, pp. 1331–1343, 2021.

[119] D. M. Müller, P. A. Robe, H. Ardon, F. Barkhof, L. Bello, M. S. Berger, W. Bouwknegt, W. A. Van den Brink, M. C. Nibali, R. S. Eijgelaar, et al., "On the cutting edge of glioblastoma surgery: where neurosurgeons agree and disagree on surgical decisions," Journal of Neurosurgery, vol. 136, no. 1, pp. 45–55, 2021.

[120] D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary," Acta Neuropathol., vol. 131, no. 6, pp. 803–820, 2016.

[121] H. Fabelo, S. Ortega, R. Lazcano, D. Madroñal, G. M. Callicó, E. Juárez, R. Salvador, D. Bulters, H. Bulstrode, and A. Szolna, "An Intraoperative Visualization System Using Hyperspectral Imaging to Aid in Brain Tumor Delineation," Sensors, vol. 18, p. 430, feb 2018.

[122] H. Fabelo, G. M. Callico, A. Vega, M. Aleman, A. de Pablo, L. Sanjuan, and R. Sarmiento, "HELICoiD Demonstrator for Intraoperative Brain Cancer Detection using Hyperspectral Images," Jornadas Comput. Empotrada (JCE 15), 2015.

[123] H. Fabelo, S. Ortega, D. Ravi, B. R. Kiran, C. Sosa, D. Bulters, G. M. Callicó, H. Bulstrode, A. Szolna, J. F. Piñeiro, et al., "Spatio-spectral classification of hyperspectral images for brain cancer detection during surgical operations," PloS one, vol. 13, no. 3, p. e0193721, 2018.

[124] C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," Remote Sensing of Environment, vol. 149, pp. 70–87, 2014.

[125] X. R. Feng, H. C. Li, R. Wang, Q. Du, X. Jia, and A. Plaza, "Hyperspectral Unmixing Based on Nonnegative Matrix Factorization: A Comprehensive Review," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 15, no. Mv, pp. 4414–4436, 2022.

[126] L. Drumetz, J. Chanussot, and C. Jutten, "Variability of the endmembers in spectral unmixing: Recent advances," Work. Hyperspectral Image Signal Process. Evol. Remote Sens., vol. 0, 2016.

[127] H. Han, G. Wang, M. Wang, J. Miao, S. Guo, L. Chen, M. Zhang, and K. Guo, "Hyperspectral Unmixing Via Nonconvex Sparse and Low-Rank Constraint," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 13, pp. 5704–5718, 2020.

[128] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 1, pp. 240–251, 2016.

[129] J. Wei and X. Wang, "An overview on linear unmixing of hyperspectral data," Mathematical Problems in Engineering, vol. 2020, 2020.

[130] D. U. Campos-Delgado, I. A. Cruz-Guerrero, J. N. Mendoza-Chavarría, A. R. Mejía-Rodríguez, S. Ortega, H. Fabelo, and G. M. Callico, "Nonlinear extended blind end-member and abundance extraction for hyperspectral images," Signal Processing, vol. 201, p. 108718, 2022.

[131] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," IEEE Transactions on Geoscience and Remote Sensing, vol. 54, pp. 240–251, 2016.

[132] S. Sánchez and A. Plaza, "Fast determination of the number of endmembers for real-time hyperspectral unmixing on gpus," Journal of real-time image processing, vol. 9, pp. 397–405, 2014.

[133] J. W. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," in JPL, Summaries of the 4th Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop, 1993.

[134] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in Imaging Spectrometry V, vol. 3753, pp. 266–275, International Society for Optics and Photonics, 1999.

[135] R. Neville, "Automatic endmember extraction from hyperspectral data for mineral exploration," in International Airborne Remote Sensing Conference and Exhibition, 4 th/21 st Canadian Symposium on Remote Sensing, Ottawa, Canada, 1999.

[136] J. M. Nascimento and J. M. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," IEEE transactions on Geoscience and Remote Sensing, vol. 43, no. 4, pp. 898–910, 2005.

[137] C.-I. Chang, H.-C. Li, C.-C. Wu, and M. Song, "Recursive geometric simplex growing analysis for finding endmembers in hyperspectral imagery," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 1, pp. 296–308, 2016.

[138] N. Wang, B. Du, L. Zhang, and L. Zhang, "An abundance characteristic-based independent component analysis for hyperspectral unmixing," IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 1, pp. 416–428, 2014.

[139] J. S. Bhatt, M. V. Joshi, and M. S. Raval, "A data-driven stochastic approach for unmixing hyperspectral imagery," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 6, pp. 1936–1946, 2014.

[140] K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "A novel hierarchical bayesian approach for sparse semisupervised hyperspectral unmixing," IEEE Transactions on Signal Processing, vol. 60, no. 2, pp. 585–599, 2011.

[141] U. B. Gewali, S. T. Monteiro, and E. Saber, "Machine learning based hyperspectral image analysis: a survey," arXiv preprint arXiv:1802.08701, 2018.

[142] B. Yang, B. Wang, and Z. Wu, Unsupervised nonlinear hyperspectral unmixing based on bilinear mixture models via geometric projection and constrained nonnegative matrix factorization, vol. 10. MDPI, 2018.

[143] H. Zhao and X. Zhao, "Nonlinear unmixing of minerals based on the log and continuum removal model," European Journal of Remote Sensing, vol. 52, pp. 277–293, 2019.

[144] Q. Wei, M. Chen, J. Y. Tourneret, and S. Godsill, "Unsupervised nonlinear spectral unmixing based on a multilinear mixing model," IEEE Transactions on Geoscience and Remote Sensing, vol. 55, pp. 4534–4544, 2017.

[145] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," International Journal of Remote Sensing, vol. 30, no. 11, pp. 2951–2962, 2009.

[146] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 11, pp. 4153–4162, 2011.

[147] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," IEEE Transactions on Image Processing, vol. 21, no. 6, pp. 3017–3025, 2012.

[148] O. Gutierrez-Navarro, D. U. Campos-Delgado, E. Arce-Santana, M. O. Mendez, and J. A. Jo, "Blind end-member and abundance extraction for multispectral fluorescence lifetime imaging microscopy data," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 2, pp. 606–617, 2014.

[149] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, et al., "Recent advances in techniques for hyperspectral image processing," Remote sensing of environment, vol. 113, pp. S110–S122, 2009.

[150] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," IEEE Geoscience and remote sensing magazine, vol. 1, no. 2, pp. 6–36, 2013.

[151] A. Villa, J. Chanussot, J. A. Benediktsson, and C. Jutten, "Spectral unmixing for the classification of hyperspectral images at a finer spatial resolution," IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 3, pp. 521–533, 2010.

[152] B. Fang, Y. Bai, and Y. Li, "Combining spectral unmixing and 3d/2d dense networks with early-exiting strategy for hyperspectral image classification," Remote Sens., vol. 12, no. 5, p. 779, 2020.

[153] A. J. X. Guo and F. Zhu, "Improving deep hyperspectral image classification performance with spectral unmixing," Signal Processing, vol. 183, p. 107949, 2021.

[154] F. I. Alam, J. Zhou, L. Tong, A. W.-C. Liew, and Y. Gao, "Combining unmixing and deep feature learning for hyperspectral image classification," in 2017 Int. Conf. Digit. Image Comput. Tech. Appl., pp. 1–8, IEEE, 2017.

[155] X. Liu and L. Huang, "Split Bregman iteration algorithm for total bounded variation regularization based image deblurring," J. Math. Anal. Appl., vol. 372, no. 2, pp. 486–495, 2010.

[156] H. K. Aggarwal and A. Majumdar, "Hyperspectral Image Denoising Using Spatio-Spectral Total Variation," IEEE Geosci. Remote Sens. Lett., vol. 13, no. 3, pp. 442–446, 2016.

[157] H. Nien and J. A. Fessler, "A convergence proof of the split bregman method for regularized least-squares problems," 2014.

[158] H. K. Aggarwal and A. Majumdar, "Hyperspectral unmixing in the presence of mixed noise using joint-sparsity and total variation," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 9, pp. 4257–4266, 2016.

[159] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Foundations and Trends in Machine Learning, vol. 3, no. 1, pp. 1–122, 2011.

[160] I. Chivers, J. Sleightholme, I. Chivers, and J. Sleightholme, "An introduction to algorithms and the big o notation," Introduction to Programming with Fortran: With Coverage of Fortran 90, 95, 2003, 2008 and 77, pp. 359–364, 2015.

[161] F. Xiong, Y. Qian, J. Zhou, and Y. Y. Tang, "Hyperspectral Unmixing via Total Variation Regularized Nonnegative Tensor Factorization," IEEE Trans. Geosci. Remote Sens., vol. 57, no. 4, pp. 2341–2357, 2019.

[162] J. Qin, H. Lee, J. T. Chi, L. Drumetz, J. Chanussot, Y. Lou, and A. L. Bertozzi, "Blind Hyperspectral Unmixing Based on Graph Total Variation Regularization," IEEE Trans. Geosci. Remote Sens., vol. 59, no. 4, pp. 3338 – 3351, 2021.

[163] A. Conci and C. Kubrusly, "Distance between sets-a survey," arXiv preprint arXiv:1808.02574, 2018.

[164] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse NMF for hyperspectral unmixing," IEEE Trans. Geosci. Remote Sens., vol. 51, no. 5, pp. 2815–2826, 2013.

[165] Y. Shao, J. Lan, Y. Zhang, and J. Zou, "Spectral unmixing of hyperspectral remote sensing imagery via preserving the intrinsic structure invariant," Sensors, vol. 18, no. 10, p. 3528, 2018.

[166] L. Zhuang, C. H. Lin, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Regularization parameter selection in minimum volume hyperspectral unmixing," IEEE Trans. Geosci. Remote Sens., vol. 57, no. 12, pp. 9858–9877, 2019.

[167] J. Yao, D. Hong, L. Xu, D. Meng, J. Chanussot, and Z. Xu, "Sparsity-enhanced convolutional decomposition: A novel tensor-based paradigm for blind hyperspectral unmixing," IEEE Transactions on Geoscience and Remote Sensing, vol. (Early Access), pp. 1–14, 2021.

[168] L. Dong, X. Lu, G. Liu, and Y. Yuan, "A novel nmf guided for hyperspectral unmixing from incomplete and noisy data," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–15, 2021.

[169] V. Brajovic, "Shot noise suppression in image sensors," in Visual Information Processing XIII (Z. ur Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds.), vol. 5438, pp. 193 – 197, International Society for Optics and Photonics, SPIE, 2004.

[170] S. W. Hasinoff, Photon, Poisson Noise. Boston, MA: Springer US, 2014.

[171] X.-R. Feng, H.-C. Li, J. Li, Q. Du, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using sparsity-constrained deep nonnegative matrix factorization with total variation," IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 10, pp. 6245–6257, 2018.

[172] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," Remote Sensing, vol. 10, no. 3, p. 482, 2018.

[173] D. G. Luenberger and Y. Ye, Linear and nonlinear programming. Springer, 3rd ed., 2008.

[174] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 3, pp. 650–663, 2004.

[175] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 11, pp. 4177–4193, 2011.

[176] J. M. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," in 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp. 1–4, 2009.

[177] J. Sigurdsson, M. O. Ulfarsson, and J. R. Sveinsson, "Blind sparse nonlinear hyperspectral unmixing using an ℓ_q penalty," IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 12, pp. 1907–1911, 2018.

[178] M. Li, F. Zhu, A. J. X. Guo, and J. Chen, "A graph regularized multilinear mixing model for nonlinear hyperspectral unmixing," Remote Sensing, vol. 11, no. 19, 2019.

[179] N. Otsu, "A threshold selection method from gray-level histograms," IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1, pp. 62–66, 1979.

[180] O. Gutierrez-Navarro, D. U. Campos-Delgado, E. R. Arce-Santana, K. C. Maitland, S. Cheng, J. Jabbour, B. Malik, R. Cuenca, and J. A. Jo, "Estimation of the number of fluorescent end-members for quantitative analysis of multispectral flim data," Optics Express, vol. 22, no. 10, pp. 12255–12272, 2014.

[181] A. K. Jain, "Data clustering: 50 years beyond k-means," Pattern recognition letters, vol. 31, no. 8, pp. 651–666, 2010.

[182] P. Geladi, J. Burger, and T. Lestander, "Hyperspectral imaging: calibration problems and solutions," Chemometrics and intelligent laboratory systems, vol. 72, no. 2, pp. 209–217, 2004.

[183] N. Gat, "Imaging spectroscopy using tunable filters: a review," Wavelet Applications VII, vol. 4056, pp. 50–64, 2000.

[184] J. Pichette, T. Goossens, K. Vunckx, and A. Lambrechts, "Hyperspectral calibration method For CMOS-based hyperspectral sensors," in Photonic Instrum. Eng. IV, vol. 10110, p. 101100H, SPIE, feb 2017.

[185] J.-B. Thomas, P.-J. Lapray, P. Gouton, and C. Clerc, "Spectral characterization of a prototype sfa camera for joint visible and nir acquisition," Sensors, vol. 16, no. 7, p. 993, 2016.

[186] B. Ghojogh, A. Ghodsi, F. Karray, and M. Crowley, "Locally linear embedding and its variants: Tutorial and survey," arXiv preprint arXiv:2011.10925, 2020.

[187] Y. Chen, L. Zhou, Y. Tang, J. P. Singh, N. Bouguila, C. Wang, H. Wang, and J. Du, "Fast neighbor search by using revised kd tree," Information Sciences, vol. 472, pp. 145–162, 2019.

[188] F. Anowar, S. Sadaoui, and B. Selim, "Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne)," Computer Science Review, vol. 40, p. 100378, 2021.

[189] A. Signoroni, M. Savardi, A. Baronio, and S. Benini, "Deep learning meets hyperspectral image analysis: A multidisciplinary review," Journal of Imaging, vol. 5, p. 52, 2019.

[190] Y. Zhang and Q. Ni, "Recent advances in quantum machine learning," Quantum Engineering, vol. 2, no. 1, p. e34, 2020.

[191] X. Zhang, C. Liu, T. Li, and Y. Zhou, "The whole slide breast histopathology image detection based on a fused model and heatmaps," Biomedical Signal Processing and Control, vol. 82, p. 104532, 2023.

[192] M. Maktabi, Y. Wichmann, H. Köhler, H. Ahle, D. Lorenz, M. Bange, S. Braun, I. Gockel, C. Chalopin, and R. Thieme, "Tumor cell identification and classification in esophageal adenocarcinoma specimens by hyperspectral imaging," Scientific Reports, vol. 12, no. 1, pp. 1–14, 2022.

[193] X. Xie, Y. Wang, and Q. Li, "Sr: Self-supervised spectral regression for hyperspectral histopathology image classification," in International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 46–55, Springer, 2022.

[194] T. Collins, V. Bencteux, S. Benedicenti, V. Moretti, M. T. Mita, V. Barbieri, F. Rubichi, A. Altamura, G. Giaracuni, J. Marescaux, et al., "Automatic optical biopsy for colorectal cancer using hyperspectral imaging and artificial neural networks," Surgical Endoscopy, pp. 1–11, 2022.

[195] A. Vabalas, E. Gowen, E. Poliakoff, and A. J. Casson, "Machine learning algorithm validation with a limited sample size," PloS one, vol. 14, no. 11, p. e0224365, 2019.

[196] J. G. A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," Computers and electronics in agriculture, vol. 153, pp. 46–53, 2018.

[197] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, "Chapter 12 - ensemble learning," in Data Mining (Fourth Edition) (I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, eds.), pp. 479–501, Morgan Kaufmann, fourth edition ed., 2017.

[198] T. Hastie, R. Tibshirani, and J. Friedman, Model Inference and Averaging. New York, NY: Springer New York, 2009.

[199] I. A. Cruz-Guerrero, D. U. Campos-Delgado, A. R. Mejia-Rodriguez, H. Fabelo, S. Ortega, and G. M. Callico, "A hybrid approach to the hyperspectral classification of in vivo brain tissue: linear unmixing with spatial coherence and machine learning," in Artificial Intelligence in Cancer Diagnosis and Prognosis, Volume 3: Brain and prostate cancer, IOP Publishing, 2022.

[200] Y. Ho and S. Wookey, "The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling," IEEE Access, vol. 8, pp. 4806–4813, 2019.

[201] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," arXiv, p. 1502.01852, 2015.

[202] A. Géron, Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.", 2019.

[203] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 6308–6325, 2020.

[204] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern classification. johnwiley & sons," Inc.,, 2001.

[205] J. Xia, N. Falco, J. A. Benediktsson, P. Du, and J. Chanussot, "Hyperspectral image classification with rotation random forest via kpca," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 4, pp. 1601–1609, 2017.

[206] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," Journal of machine learning research, vol. 2, no. Dec, pp. 265–292, 2001.

[207] A. Kaul and S. Raina, "Support vector machine versus convolutional neural network for hyperspectral image classification: A systematic review," Concurrency and Computation: Practice and Experience, vol. 34, no. 15, p. e6945, 2022.

[208] G. Mercier and M. Lennon, "Support vector machines for hyperspectral image classification with spectral-based kernels," in IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477), vol. 1, pp. 288–290, IEEE, 2003.

[209] S. Li, H. Wu, D. Wan, and J. Zhu, "An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine," Knowledge-Based Systems, vol. 24, no. 1, pp. 40–48, 2011.

[210] L. Liu, T. Miteva, G. Delnevo, S. Mirri, P. Walter, L. de Viguerie, and E. Pouyet, "Neural networks for hyperspectral imaging of historical paintings: A practical review," Sensors, vol. 23, no. 5, p. 2419, 2023.

[211] G. I. Diaz, A. Fokoue-Nkoutche, G. Nannicini, and H. Samulowitz, "An effective algorithm for hyperparameter optimization of neural networks," IBM Journal of Research and Development, vol. 61, no. 4/5, pp. 9–1, 2017.

[212] M. Adil, R. Ullah, S. Noor, and N. Gohar, "Effect of number of neurons and layers in an artificial neural network for generalized concrete mix design," Neural Computing and Applications, pp. 1–9, 2022.

[213] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized mlp architectures of neural networks," International Journal of Artificial Intelligence and Expert Systems, vol. 1, no. 4, pp. 111–122, 2011.

[214] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," IEEE Transactions on computational imaging, vol. 3, no. 1, pp. 47–57, 2016.

[215] F. He, T. Liu, and D. Tao, "Control batch size and learning rate to generalize well: Theoretical and empirical evidence," Advances in neural information processing systems, vol. 32, 2019.

[216] K. Golhani, S. K. Balasundram, G. Vadamalai, and B. Pradhan, "A review of neural networks in plant disease detection using hyperspectral data," Information Processing in Agriculture, vol. 5, no. 3, pp. 354–371, 2018.

[217] A. Ozdemir and K. Polat, "Deep Learning Applications for Hyperspectral Imaging: A Systematic Review," J. Inst. Electron. Comput., vol. 2, no. 1, pp. 39–56, 2020.

[218] J. Zabalza, J. Ren, J. Zheng, H. Zhao, C. Qing, Z. Yang, P. Du, and S. Marshall, "Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging," Neurocomputing, vol. 185, pp. 1–10, 2016.

[219] C. Li, Y. Wang, X. Zhang, H. Gao, Y. Yang, and J. Wang, "Deep belief network for spectral–spatial classification of hyperspectral remote sensor data," Sensors, vol. 19, no. 1, p. 204, 2019.

[220] T. Li, J. Zhang, and Y. Zhang, "Classification of hyperspectral image based on deep belief networks," in 2014 IEEE international conference on image processing (ICIP), pp. 5132–5136, IEEE, 2014.

[221] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 7, pp. 3639–3655, 2017.

[222] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral–spatial residual network for hyperspectral image classification: A 3-d deep learning framework," IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 2, pp. 847–858, 2017.

[223] Z. Zhong, J. Li, L. Ma, H. Jiang, and H. Zhao, "Deep residual networks for hyperspectral image classification," in 2017 IEEE international geoscience and remote sensing symposium (IGARSS), pp. 1824–1827, IEEE, 2017.

[224] Y. Bazi, L. Bashmal, M. M. Al Rahhal, R. A. Dayil, and N. A. Ajlan, "Vision transformers for remote sensing image classification," Remote Sens., vol. 13, no. 3, pp. 1–20, 2021.

[225] V. Ashish, et. al, "Attention is all you need," Advances in Neural Information Processing Systems, vol. 2017-Decem, pp. 5999–6009, 2017.

[226] X. Yifan, et. al, "Transformers in computational visual media: A survey," Computational Visual Media, vol. 8, pp. 33–62, 3 2022.

[227] L. Gao, Z. Wang, L. Zhuang, H. Yu, B. Zhang, and J. Chanussot, "Using low-rank representation of abundance maps and nonnegative tensor factorization for hyperspectral nonlinear unmixing," IEEE Trans. Geosci. Remote Sens., vol. 60, pp. 1–17, 2021.

[228] X. Hu, et. al, "Contrastive learning based on transformer for hyperspectral image classification," Appl. Sci., vol. 11, no. 18, 2021.

[229] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: an overview," arXiv preprint arXiv:2008.05756, 2020.

[230] A. Tharwat, "Classification assessment methods," Applied Computing and Informatics, vol. 17, no. 1, pp. 168–192, 2020.

[231] S. Boughorbel, F. Jarray, and M. El-Anbari, "Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric," PLoS One, vol. 12, p. e0177678, jun 2017.