# Design of an End-to-End Deep Learning System for the Spectral Band Selection and the Classification of Near-infrared Spectral Signatures.

Ing. Karen Andrea Fonseca Estupiñan

A thesis presented in fulfillment of the requirements for the degree of master electronic engineer

Advisor:

*PhD.* Henry Arguello Fuentes.

Co-directors:

*PhD.* Hans Yecid Garcia Arenas

*PhD.* Jorge Luis Bacca Quintero

Universidad Industrial de Santander

Facultad de ingenierías físicomecánicas.

Escuela de ingenierías eléctrica, electrónica y de telecomunicaciones.

2024

## Dedication

*I would like to dedicate this project to each person who has made my personal and academic growth possible throughout my time in the master's program. My director, my co-directors, my family, my fiancé and my friends.*

**Acknowledgments**

To PhD Henry Arguello, my degree project director for his support and wisdom.

To PhD Jorge and PhD Hans, for their guidance, dedication and constant teaching during my professional and academic development.

To my parents, without whom my professional development would not be possible.

To my fiancé, who has supported me day by day to achieve my goals.

This project was supported by the Vicerrectoría de Investigacion y Extensión UIS, project code 3924 under the name of ”Sistema óptico-computacional multiespectral de bajo costo en el infrarrojo cercano para la estimación de propiedades físico-químicas de granos secos de cacao mediante aprendizaje profundo.”

## Table of Contents

## List of Figures

## List of Tables

## Resumen

**Título:** Diseño de un sistema de aprendizaje profundo de extremo a extremo para la selección de bandas espectrales y la clasificación de firmas espectrales en el infrarojo cercano.

**Autora:** Karen Andrea Fonseca Estupiñan

**Palabras clave:** Imágenes espectrales, Firmas espectrales, Sistemas ópticos, Selección de bandas

**Description:** Los datos espectrales proporcionan información específica de los materiales en una amplia gama de longitudes de onda electromagnéticas. Sin embargo, la adquisición de estos datos presenta retos como la redundancia de datos, largos tiempos de adquisición y gran capacidad de almacenamiento. Para hacer frente a estos retos, se introduce la selección de bandas como una estrategia que se centra en utilizar sólo las bandas más relevantes para preservar la información espectral para una tarea específica. Los métodos más avanzados se centran en encontrar las bandas a partir de datos adquiridos previamente, independientemente del sistema óptico y del modelo de clasificación. Sin embargo, algunos métodos de aprendizaje profundo, como extremo a extremo, permiten diseñar sistemas ópticos y aprender los parámetros de la red en conjunto. En este trabajo, modelamos la selección de banda óptica como una capa entrenable acoplada a una red de clasificación cuyos parámetros se aprenden en un marco extremo a extremo. Para garantizar un sistema físicamente implementable, se propusieron dos términos de regularización en el entrenamiento para promover la binarización y el número de bandas seleccionadas, ya que se necesita proporcionar las condiciones para diseñar el elemento físico. También, se implementó un sistema de escaneo de punto, que es la versión emulada de la selección de bandas para firmas espectrales de 15 materiales diferentes. El método propuesto proporciona un mejor rendimiento que los métodos de selección de bandas más avanzados para tres conjuntos de datos espectrales públicos diferentes a pesar de capturar 478 longitudes de onda en una implementación óptica, el mismo rendimiento se logró con datos reducidos, utilizando sólo 4 bandas y en otro sistema se redujo al 2 % la cantidad de bandas obteniendo resultados comparables.

## Abstract

**Title:** Design of an end-to-end deep learning system for the spectral band selection and the classification of near-infrared spectral signatures.

**Author:** Karen Andrea Fonseca Estupiñan

**Keywords:** Spectral Imaging, Spectral Signatures, Optical systems, Band selection

**Description:** Spectral data provides material-specific information across a broad range of electromagnetic wavelengths by capturing numerous spectral bands. However, acquiring such a large amount of data presents challenges such as data redundancy, long acquisition times, and significant storage capacity. To address these challenges, band selection is introduced as a strategy that focuses on using only the most relevant bands to preserve spectral information for a specific task. State-of-the-art methods focus on finding the most significant bands from previously acquired data, regardless of the optical system and classification model. However, some deep learning methods, such as end-to-end frameworks, allow the design of optical systems and the learning of classification network parameters. In this paper, the optical band selection was modeled as a trainable layer coupled with a classification network whose parameters are learned in an end-to-end framework. To guarantee a physically implementable system, two regularization terms were proposed in the training step to promote binarization and also the number of selected bands, as we need to provide the conditions to design the physical element through which the light passes. Also, an optical whiskbroom system was implemented, which is the emulated version of band selection for spectral signatures of 15 different materials. The proposed method provides better performance than state-of-the-art band selection methods for three different public spectral datasets under the same conditions and for the acquired data despite capturing 478 wavelengths in an optical implementation, the equivalent performance can be achieved with reduced data, using only 4 bands and in another system the number of bands was reduced to 2 %, obtaining comparable results.

## Introduction

Hyperspectral imaging (HSI) captures information across a range of wavelengths in the electromagnetic spectrum. Spectral signatures are created when a beam of light strikes a sample. Depending on the material, certain wavelengths are reflected, emitted, or absorbed, and the response of each wavelength contributes to the creation of a spectral signature. Each material has a spectral response that provides essential information for its identification Bacca et al. (2023). Different spectrum ranges, such as visible (VIS) or near-infrared (NIR), provide specific spectral responses for each material, making them ideal for highlighting features related to material composition Salamati et al. (2009). HSI usually employs optical systems such as line-scanning or point-scanning Garini et al. (2006). The spectral signatures correspond to the spectral response of a spatial point and are discretized as a vector in which each position is known as a spectral band. A spectral band represents a specific segment of wavelengths; acquiring more bands means either expanding the range of the electromagnetic spectrum or providing more spectral resolution in a given range. This means that more data is required to be collected with the optical system to increase the number of spectral bands. Nevertheless, the use of these HSI can be challenging because of their acquisition time, storage limitations, and data redundancy, as highlighted in Sun and Du (2019). One particular concern for computer vision tasks is data redundancy, as training with large datasets may not adequately capture all the necessary distribution of the database for the network to effectively perform the intended task, as shown in Audebert et al. (2019).

A strategy to reduce data redundancy is band selection, which focuses on only using the

most significant bands or wavelengths to preserve the most relevant information. Band selection addresses these challenges by reducing the number of acquired hyperspectral bands without compromising the application performance of computer vision tasks. Many computer vision tasks, such as classification, only require specific bands where material signatures can be distinguished Roy et al. (2019). This approach offers the opportunity to design an optical system that captures the essential hyperspectral bands, significantly reducing costs and acquisition time by focusing on the relevant spectral ranges Sun and Du (2019). Usually, band selection is applied as a post-processing method, offering computational efficiency advantages, as full spectral data analysis can produce high computational costs. Several approaches have been proposed in the field of band selection, including ranking-based, clustering-based, sparsity-based, scheme-based, and embedding learning methods, where principal component analysis (PCA) and clustering are among the most popular techniques Wang et al. (2018). These methods aim to identify the most representative bands on a broad scale, and subsequently validate the bands in a task, such as classification. However, these methods do not train for a specific task, and their solution may not always be the optimal solution for a particular application since the best bands for classification may not be the same for reconstruction. Although some of these methods include classification tasks, they often do not optimize the two tasks, leading to satisfactory overall outcomes but do not guarantee the best performance for each task individually. On the other hand, methods based on embedding learning often incorporate terms from classification techniques into optimization algorithms for band selection; however, these techniques often use the classification score as a value that gives weight to the ranking criteria for further selection Sun and Du (2019), which implies an improvement in the selection criterion,

while the classification task is not prioritized. Therefore, the current state-of-the-art is based on the critical information needed for the overall representation of the data, rather than focusing on training with band selection criteria aligned with the specific computer vision task at hand. In addition, it is still necessary to capture all wavelengths during data acquisition to select the essential features of each band, rather than minimizing the number of captures in the optical acquisition process. Several deep learning methods for hyperspectral band selection have been proposed to deal with band selection using trainable parameters, including reinforcement learning Mou et al. (2022), unsupervised learning, and optimal clustering Wang et al. (2018). Nonetheless, end-to-end (E2E) approaches provide an optimization framework for simultaneous learning of the codification for optical acquisition and the computer vision task Arguello et al. (2023). E2E systems require using a proper optical coding system and training a neural network for a given task with suitable loss and regularization functions Bacca et al. (2021). However, these techniques often include a component for feature removal rather than focusing on training with band selection criteria Audebert et al. (2019), this implies that, despite potential improvements in computational efficiency, it is still necessary to capture all wavelengths during data acquisition. This effort is time-consuming and expensive, as the acquisition requires specialized equipment to select essential features from each band. This approach focuses on selecting essential features from each band rather than minimizing the number of captures in the optical acquisition process, an approach that has never been used in designing optical band selection. Some recent works have implemented some spectral dimension reduction. For example, in Fang et al. (2024), an inertial-free acousto-optic tunable filter allows for a non-uniform sampling strategy, or in Cull et al. (2010), a snapshot system captures a spatial

dimension and transforms into another matrix representation cube using spectral priors, allowing the transformation of a spectral cube to one of a lower dimension. However, none of the above has been aimed at the physical implementation of an optical system specialized in band selection.

Therefore, this project introduces a joint deep-learning approach for optical band selection and classification and proposes an E2E framework composed of regularization terms in the training loss function to encourage network convergence for the classification task, leading to state-of-the-art results. E2E learning frameworks improve the band selection process by integrating it into the pipeline itself, enabling the discovery of relevant features and integrated optimization. This approach simplifies development and potentially improves performance across multiple datasets, increasing flexibility for applications requiring band selection. Also, this research presents the selection and acquisition of the most relevant spectral bands for the task of material classification, through the optical implementation of band selection that allows classification in different materials. In addition, this thesis provides two datasets with 6 and 15 different classes respectively captured with 2048 and 478 spectral bands.

The contributions of this thesis are as follows:

- Two material databases acquired at the HDSP optical laboratory of visible and near-infrared spectral signatures.
- The fundamental integration of the spectral band selection mathematical operator, based on optical acquisition in the deep learning framework.
- A joint training algorithm, in which band selection becomes a customized layer that focuses

the neural network parameters on selecting the best bands for the classification task.

- The incorporation of regularization terms to guarantee a physically implementable system, which combines the trainable band selection with the global joint classification task, enhances the overall deep learning framework.

- The optical emulation of band selection for material classification.

The remaining sections of this thesis are organized as follows: Problem statement, research question, hypothesis, objectives, state-of-the-art, related work, model sensing, end-to-end band selection, optical implementation, databases, network architectures, simulations and results, conclusions and discussion, and ending with expected results.

## 1. Problem statement

Spectral imaging aims to capture spectral information over the electromagnetic spectrum. While the visible (VIS) range is widely studied, the near-infrared (NIR) spectrum enables more efficient identification and classification of materials because the NIR spectrum reveals specific material characteristics that the VIS can not. Although spectral deep learning classification is very accurate, its performance is limited by the significant volume of data. Remarking that the acquisition of this spectral data is associated with high costs because of the difficulty of creating a sensor that can capture the NIR spectral range. Therefore, selecting the most relevant spectral ranges or bands containing the information that determines material properties is essential for the classification task. Traditional methods usually perform the task and data processing separately, but integrating both objectives can focus on a more specific problem area. Other methods such as E2E seek to train two tasks together, the acquisition and a computer vision task. This method can join the band selection acquisition and the classification task. Therefore, this proposed project aims to address the problem of integrating the spectral band selection and classification task using the NIR spectral range into an E2E approach.

## 2. Research question

Assuming that the spectral bands with the best information for the classification task can be selected. What is the design of an E2E training algorithm that allows the acquisition of a low number of spectral bands and achieves high performance in a classification task for specific materials

compared with a state-of-the-art band selection approach?

## 3. Hypothesis

It is possible to design an E2E deep learning-based algorithm that can jointly select the optical band acquisition and learn classification task. Therefore, the method will be able to achieve high performance by selecting the best spectral bands for classification in the NIR spectrum range.

# 4. Objectives

## General objective

- To design an E2E deep learning-based system for the spectral band selection and classification tasks using near-infrared spectral signatures.

## Specific objectives

- To acquire a database of spectral signatures in the near-infrared spectrum of materials to be classified with a commercial NIR spectrometer available in the HDSP laboratory.
- To model in a differentiable manner the optical acquisition system for spectral signatures in the near-infrared spectrum to be included as a layer in the deep neural network.
- To design an E2E deep learning algorithm that learns the band selection structure specialized for the classification task using the dataset acquired in the near-infrared spectrum.
- To evaluate the designed optical band selection and the classification deep learning model using the dataset acquired in the near-infrared spectrum and emulate the band selection optical system.

# 5. State-of-the-art

## 5.1. Spectral imaging

Spectral imaging captures and processes information along spatial and spectral dimensions, which can be represented as a 3D spectral data cube. Imaging spectroscopy, also called hyperspectral imaging (HSI), is captured by a technique that combines conventional digital imaging and spectroscopy in a single system Bacca et al. (2022). This study works on the interaction of light with the observed materials, measuring the amount of light emitted, reflected, or transmitted from a given object, or target Paoletti et al. (2019). HSI sensors (spectrometers) typically operate in the visible and sun-reflected infrared spectrum region. However, unlike broadband sensing systems that subsample the available spectral information, narrow-band HSI systems are capable of producing, for each captured target, a distinctive spectral signature composed of reflectance measurements in hundreds of different wavelength channels Goetz et al. (1985). Providing three-dimensional (3D) images, called spectral data cubes that capture spatial information across the electromagnetic spectrum Kamruzzaman et al. (2011) as shown in Fig. 1 this information can be processed as the whole data cube, as spectral signatures (spectral dimension only in one space pixel) or spacial slice (spacial dimension only in one spectral range).

Traditionally, spectral images can be acquired through a spectrometer, which measures the light intensity of light waves in a specific wavelength range of the electromagnetic spectrum. There are different traditional techniques for capturing the spectral data cube, in which small areas of the scene are scanned until all the information from a spatial and spectral data cube can be obtained.

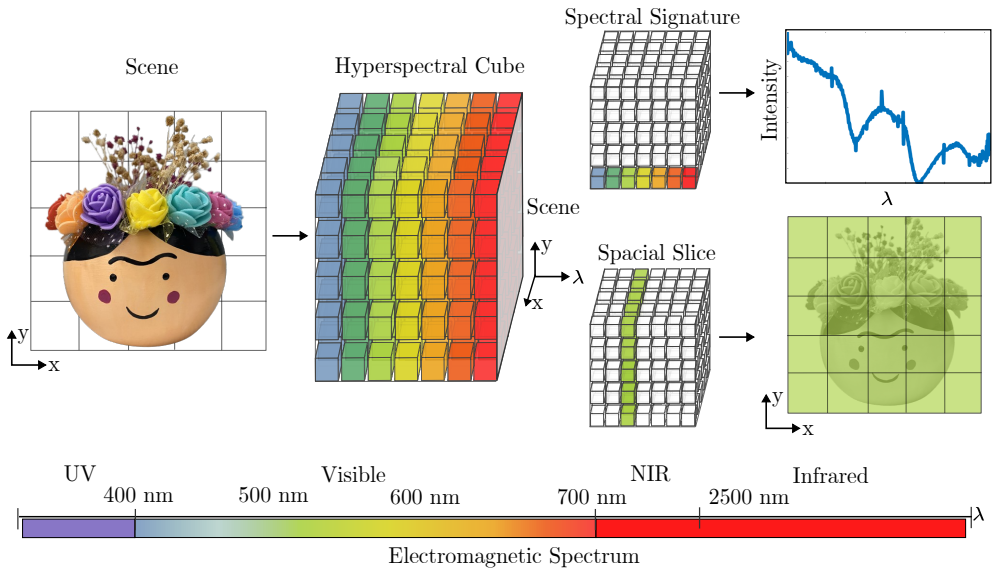



*Figure 1.* Splitting of the hyperspectral cube

Depending on the way of scanning, Fig. 2 shows that they are classified as pixel acquisition Green et al. (1998), linear acquisition Aiazzi et al. (2006), and spectral filtering acquisition Gupta (2011). The first is based on the implementation of a set of rotating mirrors that individually focus on each spatial position of the scene. The second simultaneously acquires a slice of spatial and spectral information. Finally, the third acquires each spectral band as a grayscale image with complete spatial information.

The spectrometers consist of optimized elements that operate within specific spectral ranges, thereby influencing their respective costs. Spectrometers designed for the near-infrared range tend to be approximately six times more expensive than those intended for the visible range once they go over 1100nm. This significant cost difference is attributed to the necessary complete transformation of the sensor's physical composition in near-infrared spectrometers compared to their counterparts.

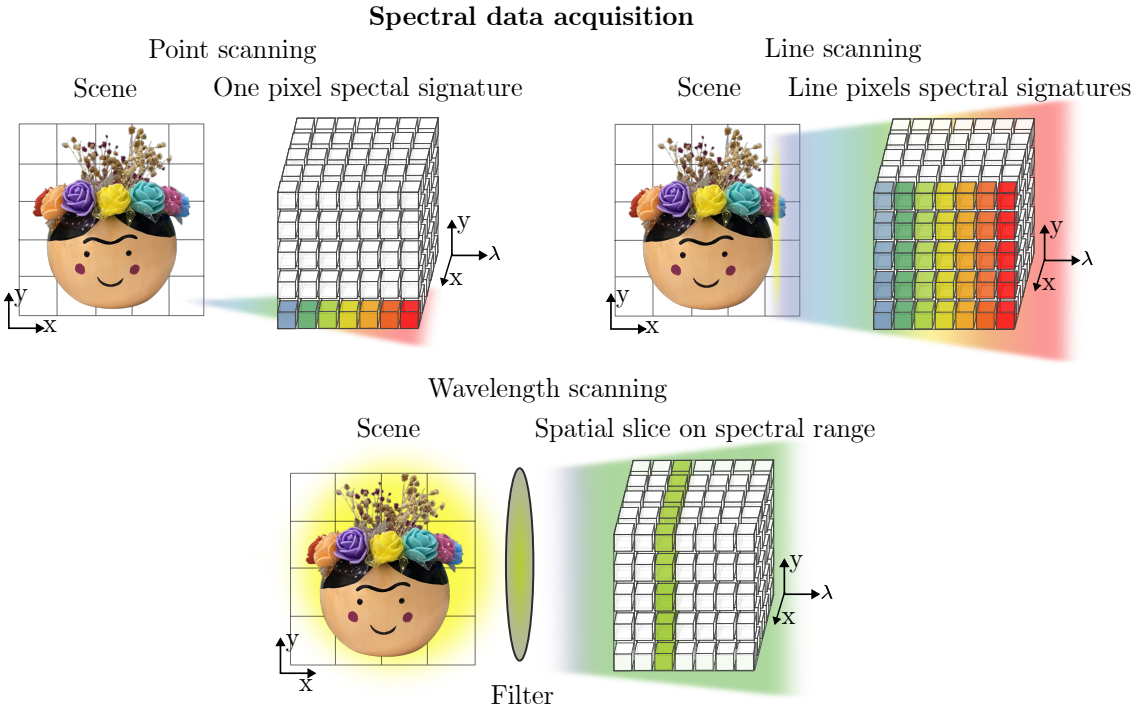



*Figure 2.* Traditional spectral image acquisition architectures.

### 5.2. Near infrared range spectrum

In the visible spectrum, significant advances have been developed exploring wavelengths from 400 nm to 700 nm since most optical systems are developed in this range. If move to the near-infrared spectrum (NIR), i.e., wavelengths from 900nm to 2500nm, this range provides greater information than data in the visible spectrum Bacca et al. (2022). However, it can be more expensive because different photodetectors must be used to create the specific sensor Hakkel et al. (2022). Some fields of interest and research areas in which hyperspectral imaging on the NIR electromagnetic spectrum is applied are pharmaceuticals Otsuka (2006), and medical applicationsSakudo (2016), among others. The reflectance light in the NIR spectrum provides a particular

spectral response ideal to highlight features related to the object composition. The NIR has been used in remote sensing Van Den Broek et al. (1996), biomedicine Tao and Farokhzad (2022), food analysis Zareef et al. (2020); Williams et al. (2009), surveillance Elihos et al. (2018), and drug detection Rodionova and Pomerantsev (2010). In Fig. 3, the visible spectrum and NIR spectrum exhibit a distinction in terms of their characteristics. While the visible spectrum only emphasizes colors, it does not effectively distinguish between materials. On the other hand, NIR images reflect additional features through variations in image intensities. This can be observed in their representation in spectral signatures in Fig. 4, where NIR shows more information than VIS with the same number of bands.

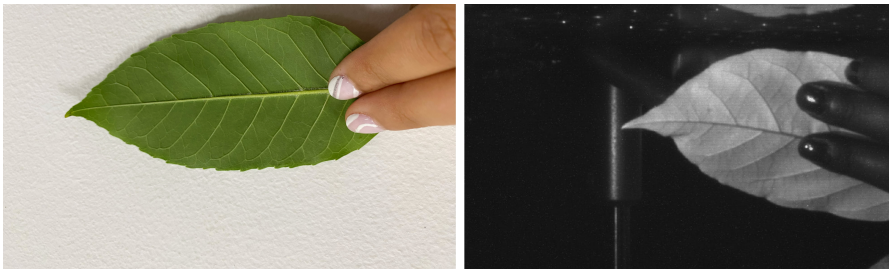


*Figure 3.* Differences between infrared and visible.

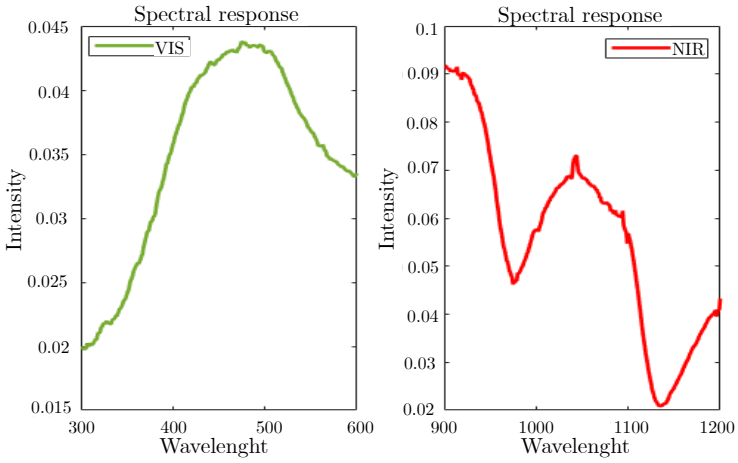



*Figure 4.* Differences between infrared and visible in spectral signatures.

### 5.3. Band Selection

Band selection aims to select the relevant spectral bands and reduce acquisition time, storage, and complexity problem Sun and Du (2019). Large amounts of data can lead to instability of the convergence of the models designed for classification, known as the curse of dimensionality Audebert et al. (2019). In Minet et al. (2010), a hyperspectral band selection approach in the VIS/NIR spectrum maximizes the contrast between the background and the target by adequately optimizing the positions and line widths of a limited number of filters.

Six hyperspectral band selection methods based on different techniques were discussed in Sun and Du (2019): In ranking-based is used a predefined band-prioritization criterion is. In searching-based, an optimization problem is employed of a given criterion function and search for the optimal solution bands. Clustering-based method groups the original bands into clusters and selects the representative bands from each one. Also, sparsity-based methods use sparse representation or regression to discover underlying structures within data. Hybrid scheme-based methods implement several schemes to select appropriate bands. Moreover, embedding learning-based methods incorporate it into the optimization of application models such as classification and target detection.

State-of-the-art (SOTA) also includes Reinforcement Learning Mou et al. (2022), the main focus of this article revolves around cluster-based band selection. An optimal clustering framework, capable of obtaining the best clustering result for a specific objective function within reasonable constraints, is introduced. Also in Wang et al. (2018) propose a Top-Rank Cut (TRC) method

utilizing Optimal Clustering (OC), and Density-Peak-Based Clustering (FDPC). Other approaches involve combining Normalized Cut (NC) with Information Entropy (IE) and Maximum-Variance Principal Component Analysis (MVPCA) and combinations between them. Although these methods aim to preserve spectral information, only Mou et al. (2022) have in consideration the application, about the others, they are not designed to select characteristic information to realize a specific task such as classification.

### 5.4. End-to-end model

End-to-end (E2E) consists of an optimization framework for jointly learning a spectral classification network's optical patterns and parameters Arguello et al. (2023). E2E systems require four foundations: a suitable optical system, the modeling of optical acquisition codes, the performance of a computational imaging task, and the training of a neural network with appropriate loss and regularization functions Bacca et al. (2021). Optimized coding offers new data-driven optical designs that outperform conventional non-data-driven approaches Arguello et al. (2023). A disadvantage in all deep learning models is their dependence on gradient descent, which could lead to slow convergence or poor local optima Glasmachers (2017).

The optimization problem defined to perform end-to-end model is represented in Fig. 5 and modeled as:

$$\{\phi^*, \theta^*\} = \underset{\phi,\ \theta}{\arg\min} \sum_k \mathscr{L}_{\text{cross}} \left( \mathscr{N}_\theta \left( \mathscr{M}_\phi \left( \mathbf{x}_k \right) \right), \mathbf{d}_k \right), \tag{1}$$

with a set of $k$ spectral signatures, $\mathbf{x}_k$, and its respective labels $\mathbf{d}_k$, the sensing matrix band selection

$\mathscr{M}_{\phi}(\cdot)$ as band selection operator, and the non-linear classification operator $\mathscr{N}_{\theta}(\cdot)$.

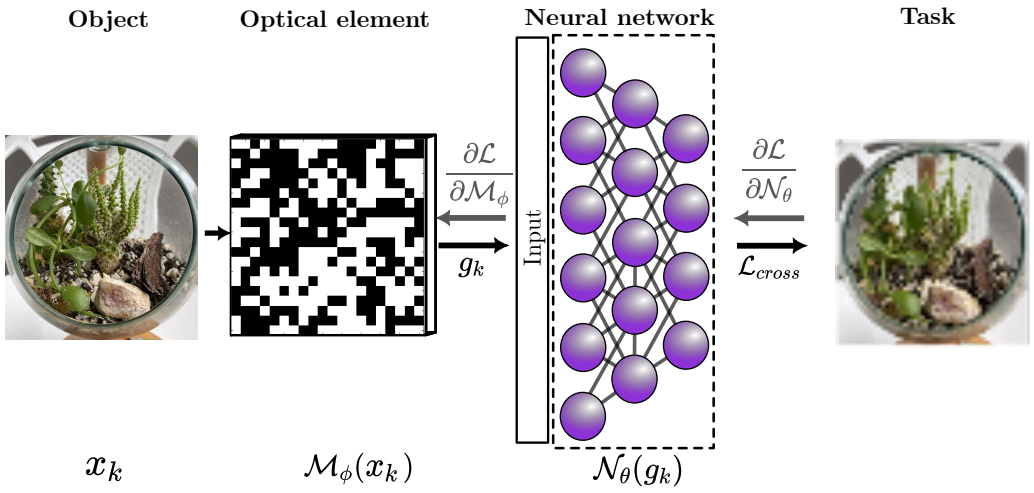



*Figure 5.* Pipeline of training of the end-to-end deep learning model.

## 6. Related work

Several methods have been explored in the band selection field. This section will introduce some of the most pertinent state-of-the-art techniques.

### 6.1. Post-processing band selection methods

Conventional band selection methods traditionally identify the most relevant features for specific datasets. To address this challenge, Wang et al. (2018) introduces an optimal clustering framework aimed at identifying the most suitable clustering structure within hyperspectral imaging (HSI) data. A comparative analysis is conducted between various methods, including Optimal Clustering (OC), Density-Peak-Based Clustering (FDPC), Maximum-Variance Principal Component Analysis (MVPCA), and Information Entropy (IE). To facilitate the selection of representative bands within the established clustering structure, a ranking strategy for clustering is proposed. This strategy involves the fusion of Normalized Cut (NC) and Top-Rank Cut (TRC), which improves the selection of bands within the clustering structure. However, some combinations are very sensitive to noises Wang et al. (2018).

On the other hand, to achieve feature extraction, the clustering methods are widely employed, as demonstrated in Huang et al. (2022), where a tensor-based subspace clustering model is introduced, leveraging prior information about the spatial dependencies of pixels and the spectral correlations of bands. The method also utilizes heterogeneous regularization with an efficient algorithm based on the alternating direction method of multipliers (ADMM) to optimize the model. This method uses prior information about the spatial dependencies of the pixels and the spectral

correlations of the bands. However, although it introduces a novel approach to selecting bands that effectively preserve the spatial structure of the HSI, tensor-based models are more time-consuming than other methods. Furthermore, in Fu et al. (2022), dimensionality reduction focuses on addressing spatial noise, and a novel approach based on neighborhood grouping normalized matched filters for band selection is proposed, however, the computation time increases exponentially with the size of the data, which for large databases could lead to inefficiency.

### 6.2. Deep learning band selection methods

In computer vision, deep learning has become a pioneering technique that has opened the way for advances in hyperspectral band selection. A current approach is to employ unsupervised learning to discriminate optimal band configurations. For example, in Mou et al. (2021), the authors treat unsupervised hyperspectral band selection as a Markov decision process, transforming it into a problem that can be solved by deep reinforcement learning algorithms. For this purpose, a deep reinforcement learning model is trained using a Q-network. Another innovative contribution is presented in Liu et al. (2022), where an unsupervised band selection framework capable of selecting a subset of bands is introduced. This framework takes advantage of a masked convolutional autoencoder that allows simultaneous consideration of band redundancy and information content. In addition, Chang et al. (2023) proposes an unsupervised approach based on rate-distortion principles for hyperspectral image classification, which adds to the diversity of unsupervised band selection techniques. Adopting a different approach to deep learning, Li et al. (2023) explores filter selection learning within a large set of wide-band filters. They present a network designed to optimize the encoded filter array pattern along with image demosaicing and hyperspectral image

retrieval. To complete the process, a spatio-spectral network prior is employed to transform the demosaicked multispectral images into their final shapes. However, deep learning approaches for hyperspectral band selection typically focus solely on the selection task and do not concurrently address computer vision tasks.

## 7. Sensing Model

To model the selection of bands as optical layers within the deep learning framework, it is necessary to represent the optical acquisition system, as illustrated in Figure 6.

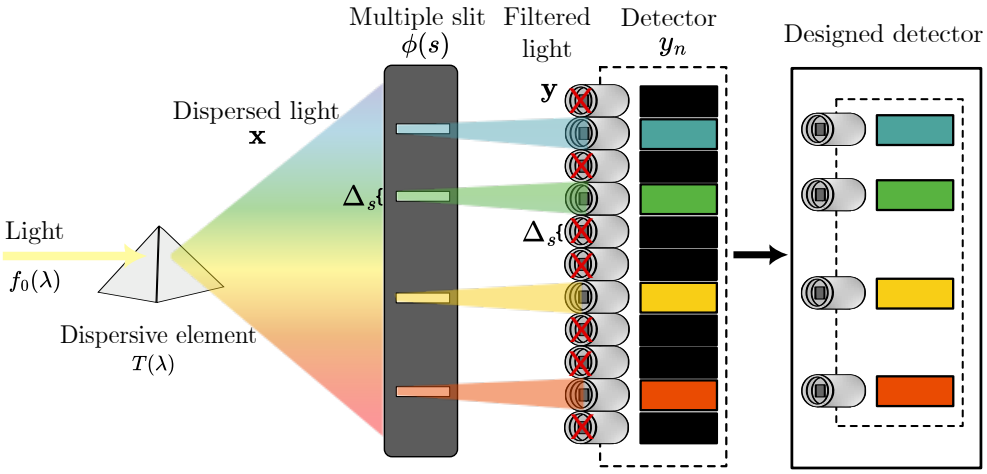



*Figure 6.* Proposed optical sensing model. The light $f_0(\lambda)$ passes through a dispersive element obtaining the dispersed spectral signature $x(s)$. Then the dispersed light goes through a slit which restricts the light passing to a discrete sensor to obtain $y_n$. By selecting specific wavelength bands, can be designed an optical system that measures only the relevant ranges using fewer sensors than the original detector.

The point incident light is denoted as $f_0(\lambda)$, where $\lambda$ represents the wavelength. The transmittance of each wavelength $\lambda$ of the dispersive element is initially characterized by $T(\lambda)$. Therefore, the dispersed spectral signature $x(s)$ at a fixed distance is expressed as

$$x(s) = T(\lambda) f_0(\lambda) * \delta(s - S(\lambda)), \tag{2}$$

where $\delta(s - S(\lambda))$ represents the point spread function of the system over the spatial dimension s, and $S(\lambda) = \alpha(\lambda)(\lambda - \lambda_c)$ is the dispersion induced by the dispersive element along

the $x$-axis centered at the wavelength $\lambda_c$ with a dispersion coefficient $\alpha(\lambda)$, and $*$ symbolizes the convolution operation. The band selection model is given by a slit array $\phi(s)$, using a digital micromirror device (DMD), this could be emulated. Such that the response after the coded selector $y(s)$ is

$$y(s) = \phi(s) \odot x(s). \tag{3}$$

Assuming that the pixel size of the slits and the sensor are the same $\Delta_s$, each spatial position of the sensor array can be modeled by $y_n$ represented as

$$y_n = \int_{(1-n)\Delta_s}^{n\Delta_s} y(s) rect\left(\frac{s}{\Delta_s} - n\right) ds, \tag{4}$$

where $n = 1, \cdots, L$ represents the number of spectral bands, $rect()$ is the function that represents the discretization caused by the sensor where it limits the light reaching the sensor over the spatial dimension. Note that some $y_n$ are zero since the slits block the lights, therefore, these bands are not selected, resulting in an optical band selection system that can be physically implemented to reject some wavelengths.

## 8. End-to-end Band Selection

To formulate the optimization problem based on the modeling of the previous section, the proposed method is presented in Fig. 7. The modeling the band selection operator $\mathscr{M}_\phi$ is defined

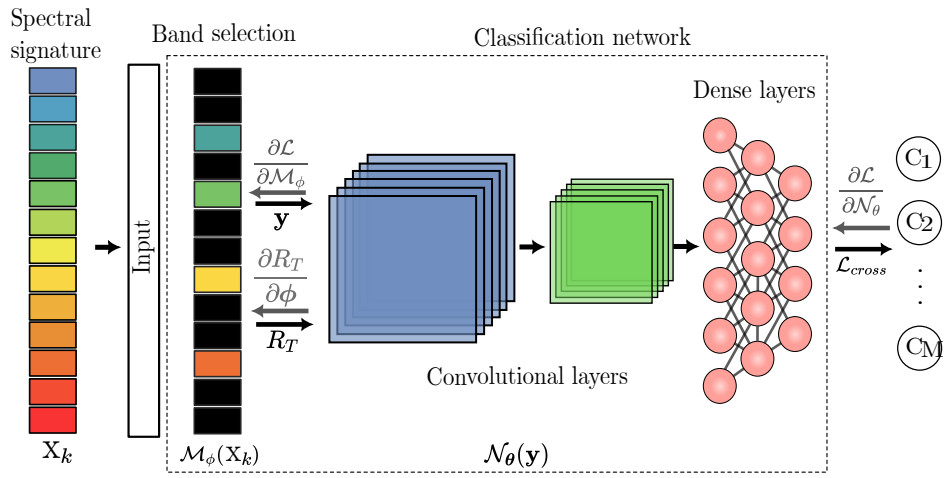



*Figure 7.* The computational pipeline of the proposed method. The element-wise product $\odot$ is performed between $\mathbf{x}_k$ and $\phi$ obtaining only the information of the selected bands. The training is performed over the $\mathscr{L}_{cross}$, including also the regularizer $R_T$ over the binary weight $\phi$.

as

$$\mathbf{y} = \mathscr{M}_{\boldsymbol{\phi}}(\mathbf{x}) = \boldsymbol{\phi} \odot \mathbf{x}, \quad \textit{where} \quad \|\mathbf{y}\|_0 = N \ll L, \tag{5}$$

with $\mathbf{x} \in \mathbb{R}^L$ as the full bands spectral signature for $L$ spectral bands and $\mathbf{y}$ is the selected bands of $\mathbf{x}$ given by the binary weight $\phi \in \{0,1\}^L$, with $N$ the number of selected bands or one values over the band selection layer. The proposed method consists of jointly designing $\phi$ and the weights $\theta$ of a classification network $\mathscr{N}_\theta(\cdot)$, taking into account a training label dataset $\{\mathbf{x}_k, \mathbf{c}_k\}_{k=1}^K$ by solving

the following optimization problem:

$$\{\phi^*, \theta^*\} = \underset{\phi,\ \theta}{\arg\min} \sum_{k=1}^{K} \mathscr{L}\left(\mathscr{N}_\theta\left(\mathscr{M}_\phi\left(\mathbf{x}_k\right)\right), \mathbf{c}_k\right) + \mu R_T(\phi), \tag{6}$$

where $\mu$ is a trade-off hyper-parameter between the cross entropy loss function $\mathscr{L}$ Nesterov (1983), and the proposed regularizer, $R_T(\cdot)$ which assists in the characterization of the model layer by converting it into binary values and ensuring that a specific number of desired bands are selected. Observe that $R_T(\cdot)$ operates only over $\phi$. Compared to equation 1, in this method, regularization terms are added for band selection, in charge of binarization and choosing the specific number of bands desired.

### 8.1. Binary regularization term

This method uses a regularization term that promotes binary values. This approach is intended to facilitate future physical implementation. When designing a sensor, it is necessary to consider whether a spectral band should be captured. The proposed binary regularization is defined as follows

$$R_{Binary}(\phi) = \left(\sum_{n=1}^{L} \beta(\phi_n^2)^\alpha (1-\phi_n)^2 + (1-\beta)(\phi_n^2)^\alpha\right). \tag{7}$$

The hyperparameters $\beta$ and $\alpha$ are essential for controlling the number of resulting bands, denoted as $N$ during the training step. In terms of binarization control, the parameter $\beta$ manages the weighting of the zeros, influenced by the structure of $R_{Binary}(\phi)$, as illustrated in Fig. 8. Note that can be inferred that a lower $\beta$ value corresponds to a greater tendency to select fewer bands. During the training process, an initial value of $\beta$ is chosen, and it is incrementally increased by

$$\beta_e = \beta_{e-1} + \frac{\tau}{50}, \tag{8}$$

where $\tau$ represents the factor by which $\beta$ increases in each epoch $e$. To control the rate of increase of $\beta$, we introduce the parameter $p_e$, which specifies after how many epochs $\beta$ should increase in value. This process continues until $\beta$ reaches a value of 1, marking the equilibrium point for the binarization process. The initial values for $\beta$, $\tau$, and $p_e$ vary for each dataset.

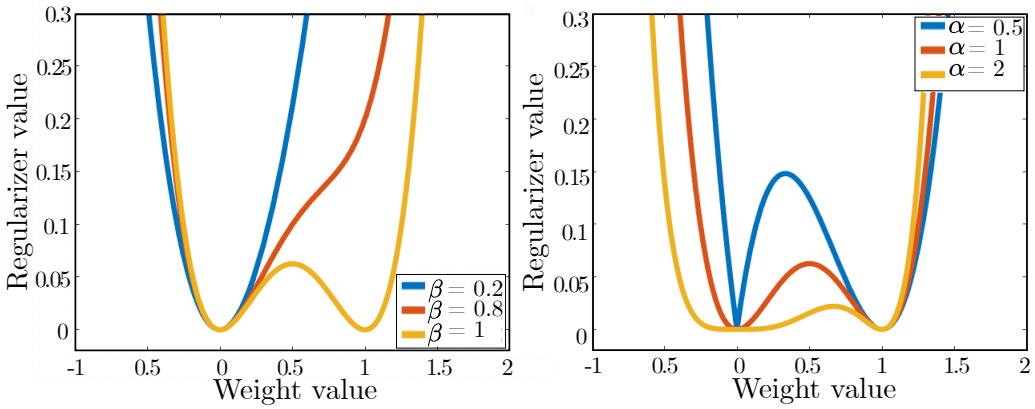



*Figure 8.* Behavior $R_B(\phi)$ depending on $\beta$ and $\alpha$. Note that the $\beta$ value affects the promotion of zeros in the binarization terms, while $\alpha$ promotes the one values.

### 8.2. Ranking regularization term

This method also further proposes incorporating a regularization term that selects only a specific number of bands. This choice is intended to ensure that a precise number of sensors can be used for data acquisition. To obtain a specific number of bands, the following regularization terms were proposed:

$$R_{bands}(\phi) = \rho \left( N - \sum_{n=1}^{L} \phi_n \right)^2 . \tag{9}$$

Observe that $R_{bands}(\cdot)$ operates only over $\phi$; where $\rho$ represents the weight given to the $R_{bands}(\phi)$ regularization term. Thus, the proposed regularizer aims to achieve a specific number of selected bands with binary values, where ones denote the selected bands.

### 8.3. Training regularization

Finally, the final regularization term for the whole training given is represented as

$$R_T(\phi) = \underbrace{\left( \sum_{n=1}^{L} \beta(\phi_n^2)^{\alpha}(1-\phi_n)^2 + (1-\beta)(\phi_n^2)^{\alpha} \right)}_{R_{binary}} + \underbrace{\rho \left( N - \sum_{n=1}^{L} \phi_n \right)^2}_{R_{bands}} \tag{10}$$

Notice that the regularizers $R_{binary}(\phi)$ and $R_{bands}(\phi)$ are minimized when the entries of $\phi$ are binary and $\|\phi\|_1 = N$, respectively.

Once equation 6 is formulated, it is solved using a gradient descent algorithm such as ADAM Kingma and Ba (2014). The proposed method is designed to adapt to the characteristics of the databases provided. Each database has a unique behavior. The method aims to progressively reduce the importance of less important bands by decreasing the value of some and increasing the value of others. Throughout the process, bands can be discarded or re-selected at each step until the most appropriate bands are identified. Once the value of the bands has been determined, the parameters of the regularized are set to a value that no longer affects the selection process but ensures the binarization of the selected bands.

## 9. Optical implementation

### 9.1. Spectrometer optical system

A Spectrometer acquisition system was developed to implement the spectral signature acquisition optical system shown in Fig. 9 and simulate the band selection process. A schematic of the optical system is shown in Fig.10.

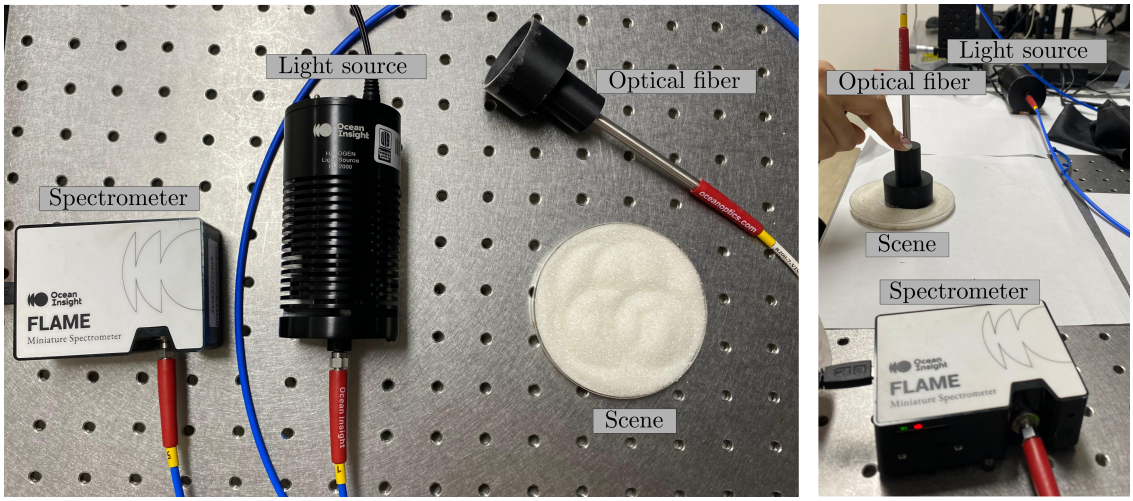



*Figure 9.* Optical implementation of spectrometer spectral signature acquisition system

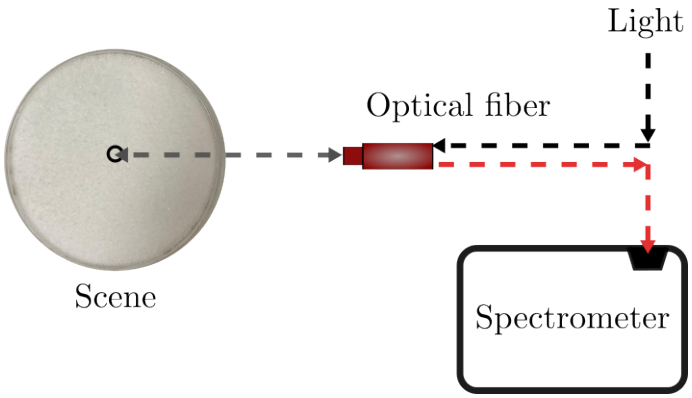



*Figure 10.* Schematic of the optical system implemented for acquiring spectral signatures with a spectrometer.

The implemented system consists of a spectrometer that captures the spectral signature of a specific point in space. The light travels through one of the two paths of the optical fiber until it reaches the scene. When the light arrives and reflects with the sample through the other path, the spectral response of that spatial point can be obtained, thus getting the spectral signature of different materials. The spectrometer spectral range is between 400-1100 [nm]. The optical elements used are a Flame miniature spectrometer by Ocean Insight, an optical fiber by Ocean Insight, and a halogen light source HL-2000 by Ocean Insight.

### 9.2. Point scanning Whiskbroom optical system

A Whiskbroom-type spectral acquisition system was developed to implement the spectral signature acquisition optical system shown in Fig. 11 and emulate the band selection process. A schematic of the optical system is shown in Fig.12.

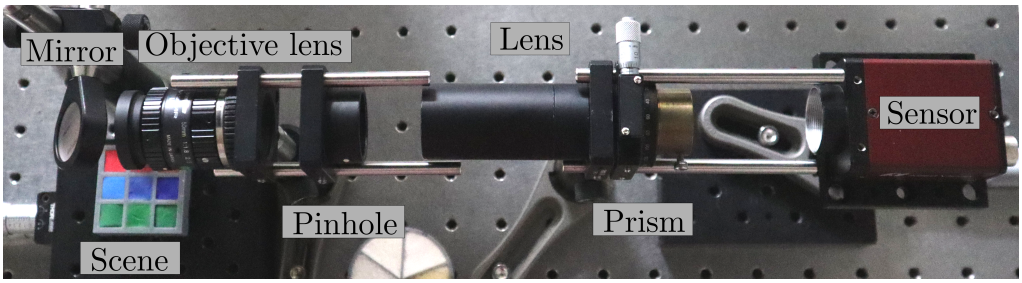



*Figure 11.* Optical implementation of a Whiskbroom-type spectral signature acquisition system.

The implemented system consists of a mirror that directs the light from the scene to the objective lens. A mirror was used because of the nature of the scenes to be captured. Subsequently, an objective lens focuses the incident light from the scene to an image plane with a pinhole, which serves as a narrow aperture for spatially filtering the scene. The light then travels from the pinhole to the sensor through a lens with a focal length of f=50 [mm] and passes through a dispersing

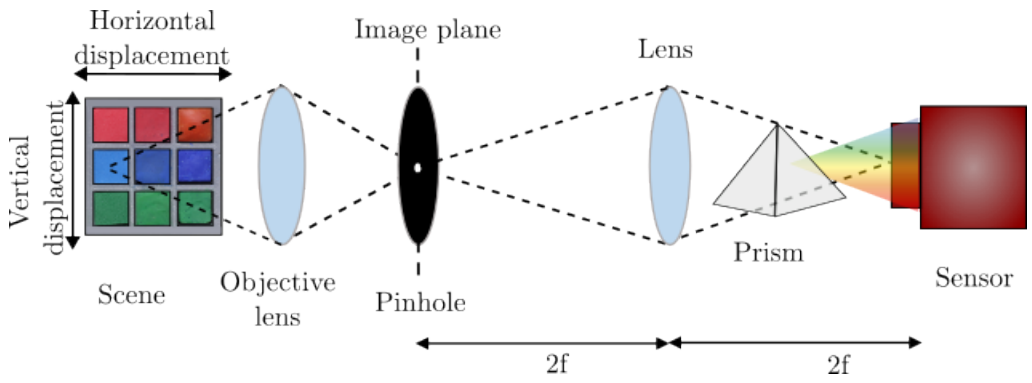



*Figure 12.* Schematic of the optical system implemented for acquiring spectral signatures of different materials.

element, in this case, a prism. The prism disperses the light beam into its spectral components. Finally, the 2D sensor, acting as a camera, captures the spectral information of the scene. Notably, the spectral information is spatially dispersed along the sensor due to the nature of light.

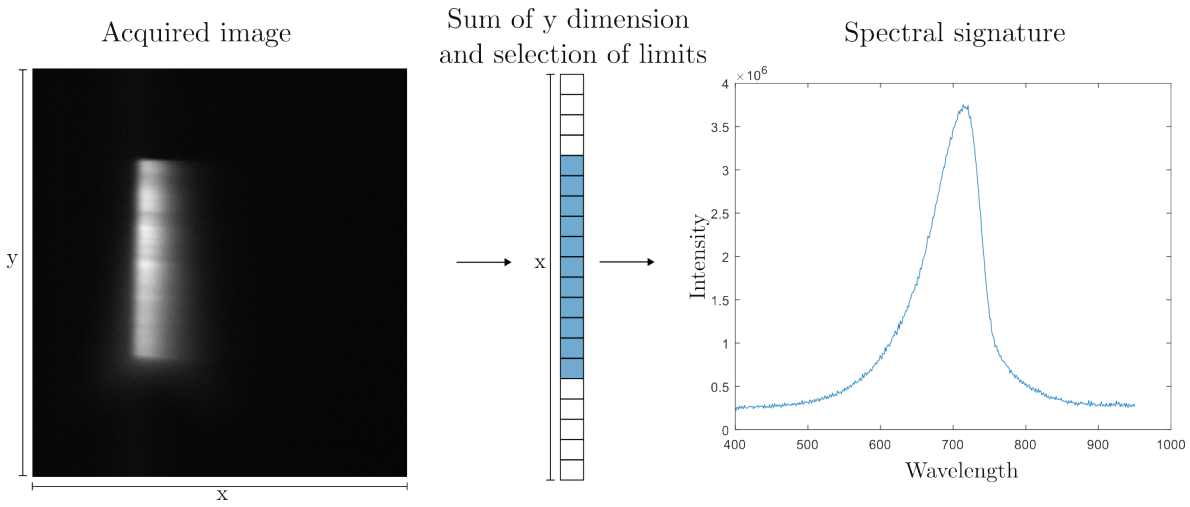



*Figure 13.* Process to acquire the spectral signature. The image is summed over the dimension to capture the spectral dispersion, then specific wavelengths are selected and the spectral signature is plotted.

Note that the distance between the prism and the sensor was selected by looking for the dispersion of the greatest number of spectral bands between 400-950 [nm], without significantly

affecting the luminous efficiency of the system. Since the system uses a 2D sensor to acquire all the spectral information, the first step was to characterize the dispersion angle of the prism to identify the pixels corresponding to each wavelength. To obtain this information, the optical system was illuminated with a monochromator, which allowed us to obtain in a controlled way the spectral bands from 400 [nm] to 950 [nm] using a 10 [nm] step. Once the characterization was obtained and the spatial positions for each row for spectral dispersion were known, the scenes were positioned as input to the optical system, moving it to take multiple signatures at different positions. Then, post-processing was performed to identify the range of pixels in which the information related to the dispersion was acquired according to the characterization as shown in Fig. 13, and based on this, each spectral signature was assembled according to the intensity at each point. The optical elements used are a stingray camera by Allied Vision, a prism CRM1P by Thorlabs, a pinhole S30K by Thorlabs, an objective lens of f=50mm by Computar VisWir Lite, and a mirror PF10-03-P01 by Thorlabs.

## 10. Datasets

In this thesis, tests were performed at different stages with a total of 5 datasets, 3 public databases found in the state of the art, and two acquired in the HDSP research laboratory.

### 10.1. Public datasets

In this article, we use the Indian Pines, Salinas, and Pavia University to train the method and compare it with the state-of-the-art, these three datasets are shown in Fig. 14.

**10.1.1. Indian Pines.** The Indian Pines dataset was captured by the AVIRIS sensor over the Indian Pines test site, located in northwestern Indiana. The image dimensions are 145 $\times$ 145 pixels, and it includes 224 spectral reflectance bands covering wavelengths from 400 to 2,500 nm. The Indian Pines scene primarily consists of two-thirds devoted to agriculture and one-third covered by forests and other forms of natural evergreen vegetation. We decided to reduce the number of spectral bands to 200 (corrected version) by excluding the bands that fall within the water absorption region. The field information available for this dataset is classified into 16 different classes, with the potential for overlap between these categories.

**10.1.2. Salinas.** The Salinas database was acquired using a 224-band AVIRIS sensor, which offers high spatial resolution with 3.7-meter pixels, covering a 512 $\times$ 217 pixels grid over the scenic Salinas Valley in California. As with the Indian Pines scene, we have removed 20 bands associated with water absorption. The Salinas dataset is classified into 16 different classes and includes a variety of features, such as vegetable fields, patches of bare soil, and extensive areas of vineyards.

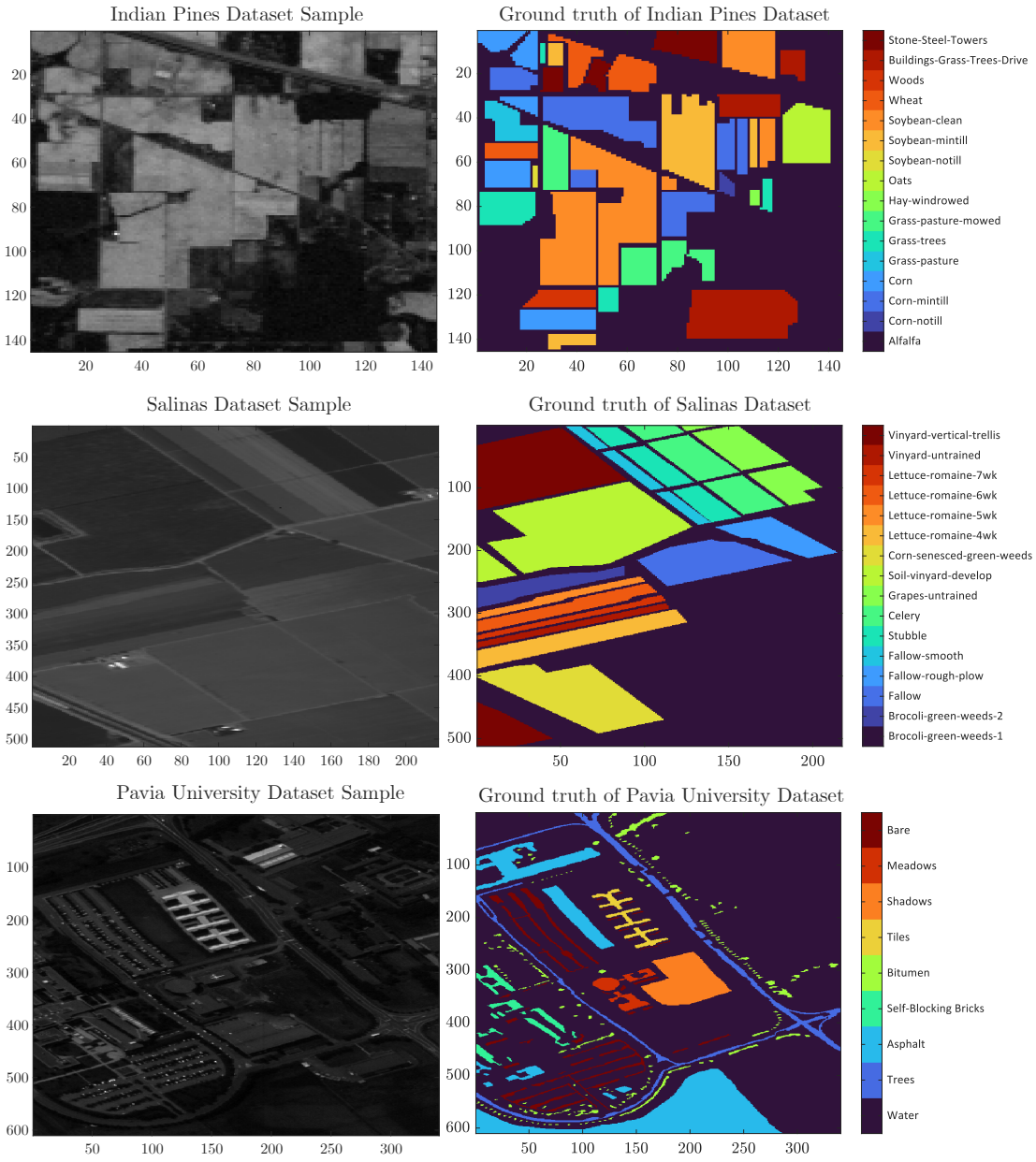



*Figure 14.* The visualization of the datasets includes representations of the Indian Pines, Salinas, and Pavia University datasets, each accompanied by its corresponding ground truth. False colors represent each class in the datasets.

**10.1.3. Pavia University.** The Pavia University dataset was captured with the ROSIS sensor during an aerial survey over the northern region of Pavia (Italy). The Pavia Univer-

sity dataset includes 103 spectral bands. In terms of spatial dimensions, Pavia University covers $610 \times 610$ pixels. This dataset is divided into 9 distinct classes.

### 10.2. Acquired datasets

**10.2.1. Spectrometer optical system.** The optical system implemented shown in Fig. 9 was used to acquire the spectral signatures of 6 different materials, using six scenes that can be seen in Fig. 15. For each scene, specifically for each of the materials, a total of 22 spectral signatures were made each one, with a total of 2048 spectral signatures.


*Figure 15.* Scenes of several white materials acquired by the spectrometer implemented optical system.

**10.2.2. Whiskbroom optical system.** The optical system implemented shown in Fig. 11 was used to acquire the spectral signatures of 15 different materials, using two scenes that can be seen in Fig. 16 as scene A and scene B. For each scene, specifically for each of the materials, a total of 35 captures of spectral signatures were made each one, with a total of 525

spectral signatures [1].

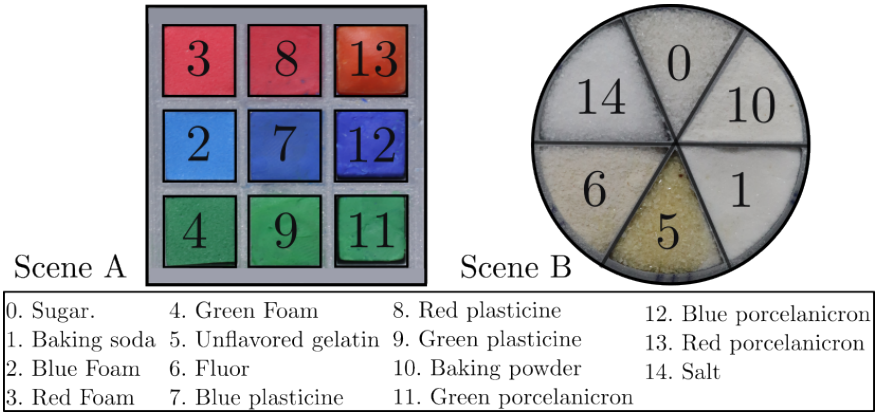



*Figure 16.* Scenes of various materials and colors acquired by the implemented optical system.

***10.2.2.1. Scene A.*** In this scene, three types of materials were included: foam, clay, and porcelain, arranged from left to right in the column, each presented in three different colors: red, blue, and green. The main reason to use this type of scene is to verify that the spectral band selection is not affected by the similarity of colors among the materials.

***10.2.2.2. Scene B.*** For scene B, were looked for different highly reflective materials with similar color tones, as depicted in Figure 16. These similar white materials include sugar, salt, baking powder, baking soda, unflavored gelatin, and flour. The main reason for acquiring the spectral signature of these materials is to evaluate the performance of the proposed method for classification tasks in scenes with similar color tones but different materials.

---

[1] Dataset available in *https://github.com/karena2/Spectral-Dataset15.git*

## 11. Network architecture

For the evaluation of the method, three different architectures were used. The classification architectures used are described in Table 1 and Table 2, which specify the type of layer, stride, and activation for each layer.

### 11.1. Public datasets

The test for comparison with other methods was conducted to ensure that the method's robustness is not dependent on architecture.

| **CNN model** | | | | **Dense model 1** | | |
|---|---|---|---|---|---|---|
| **Layer** | **Kernel** | **Stride** | **Activation** | **Layer** | **Kernel** | **Activation** |
| Input | - | - | - | Input | - | - |
| Band selector | - | - | - | Flatten | - | - |
| Conv2D | 32 | [2,1] | ReLu | Dense | 128 | ReLu |
| Conv2D | 32 | [2,1] | ReLu | Dense | 256 | ReLu |
| Conv2D | 64 | [2,1] | ReLu | Dense | 512 | ReLu |
| Conv2D | 128 | [2,1] | ReLu | Dense | num classes | Softmax |
| Conv2D | 256 | [1,1] | ReLu | | | |
| Conv2D | 512 | [1,1] | ReLu | | | |
| Flatten | - | - | - | | | |
| Dense | 128 | - | ReLu | | | |
| Dense | 256 | - | ReLu | | | |
| Dense | 512 | - | ReLu | | | |
| Dense | num classes | - | Softmax | | | |

*Table 1.* Deep learning CNN model architecture. The CNN model consists of some convolutional neural networks (CNN) followed by four dense layers to facilitate the classification task. The kernel size increases with each subsequent convolutional layer. The CNN model was used in the proposed method to select the bands. The Dense model was used as an additional test method for the bands selected by all methods.

### 11.2. Acquired datasets

To evaluate the proposed method in acquired datasets, the deep learning architecture model described in Table 2 was used. The model uses five dense layers with batch normalization and dropout to prevent overfitting.

| **Dense model 2** | | |
|---|---|---|
| **Layer** | **Kernel** | **Activation** |
| Input | - | - |
| Flatten | - | - |
| Dense | 256 | ReLu |
| Batch normalization | - | - |
| Dropout | - | - |
| Dense | 128 | ReLu |
| Batch normalization | - | - |
| Dropout | - | - |
| Dense | 64 | ReLu |
| Batch normalization | - | - |
| Dense | 32 | ReLu |
| Batch normalization | - | - |
| Dense | Num classes | Softmax |

*Table 2.* Deep learning dense model architecture used for acquired datasets.

## 12. Simulations and results

This section will present the results of different experiments and databases used, using in general a batch size of 32 and a learning rate of 1e-3 for public datasets and 1e-3 for acquired datasets.

### 12.1. Public datasets

Each dataset is partitioned into training, validation, and testing sets for these experiments in proportions of 70 %, 10 %, and 20 %. The database was divided in proportion to its classes, i.e. each percentage keeps the proportion of each class respectively, for each training the model was run for 200 epochs. The proposed optical band selection system is trained using each spatial point as an independent measurement. To determine the appropriate hyperparameter selection multiple iterations for the three databases were performed, varying the value of the $\beta$ parameter to assess its impact on the training results, this $\beta$ parameter controls the number of resulting bands and manages the weighting of the zeros. Figure 17 illustrates the influence of this parameter on the binarization process.

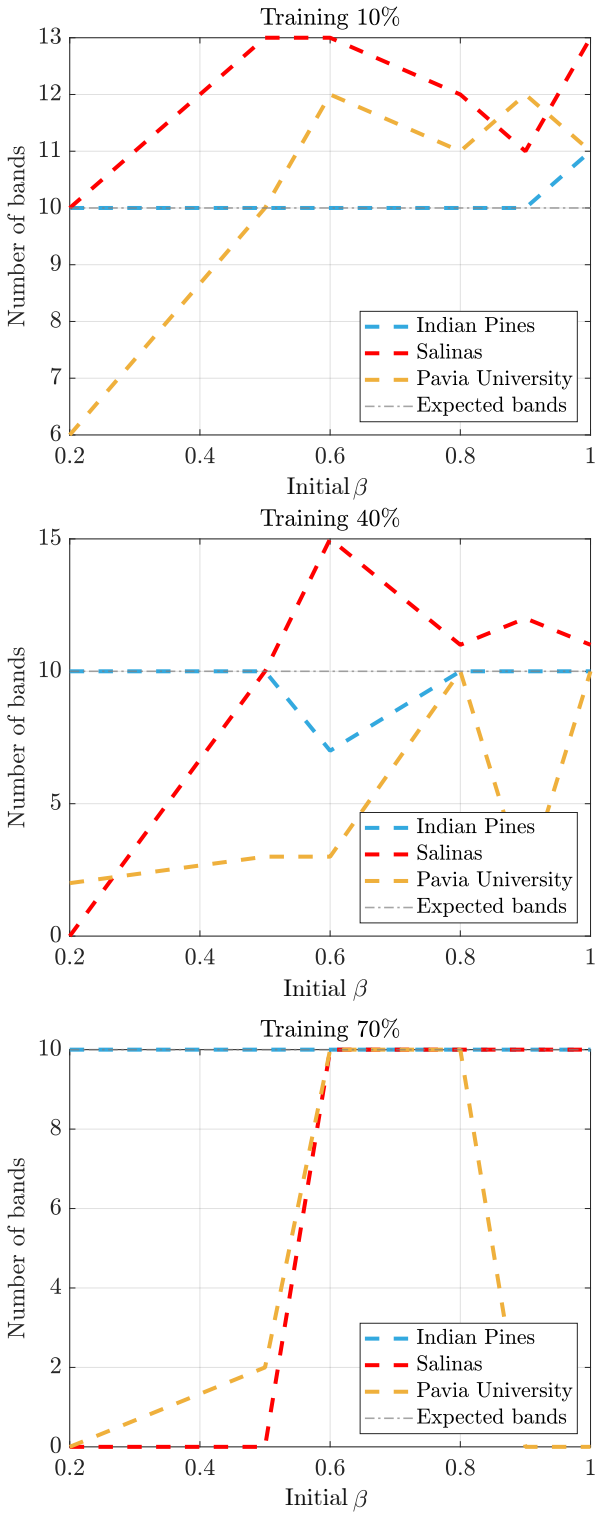



*Figure 17.* Effect of $\beta$ over selected bands in each dataset. The figure illustrates the impact of the $\beta$ parameter on the selection of 10 bands, which is dependent on the input data provided to the neural network.

A critical factor in inducing sparsity in the model weights is the $\beta$ parameter. A higher $\beta$ encourages more weights to approach zero, resulting in the selection of fewer bands with significant information. In large datasets, such as the Indian Pines database, a small $\beta$ is preferred because it allows for a gradual and smoother selection process during training. However, for smaller datasets such as Salinas and the University of Pavia, a $\beta$ that is too small can prolong the process of weight convergence to zero. This may exceed the desired number of selected bands. In these cases, a higher $\beta$ is beneficial because it speeds up the band selection process, ensuring more efficient convergence to the desired sparsity level. Table 3 displays the corresponding parameters used to achieve network convergence for each dataset using the Overall Accuracy (OA) metric.[1].

| **Dataset** | $\beta$ | $\alpha$ | $\mu$ | $\tau$ | $\rho$ | $p_e$ |
|---|---|---|---|---|---|---|
| Indian Pines | 0.1 | 1 | 1x10-6 | 3 | 10 | 4 |
| Salinas | 0.8 | 1 | 1x10-8 | 2 | 10 | 2 |
| Pavia University | 1 | 1 | 1x10-8 | 3 | 10 | 4 |

*Table 3.* Parameters selected in the training model for each dataset. $\beta$ and $\alpha$ ensure the binarization, $\mu$ is the weight given to the regularization term, $\tau$ promotes the increase of $\beta$ each $p_e$ epoch during the training until reach 1.

In the training of the band selection, when going through a binarization process, two phases are observed, the first is the selection, in which the values oscillate, looking for the indicated number of bands, once it finds this point, it binarizes radically, so a drop in accuracy can be observed in Fig. 18 since the network was trained for an input of size L of the data. However, it then recovers with these new input values.

---

[1] Codes available in *https://github.com/karena2/Deep_Jointly_Optical_BandSelecion.git*

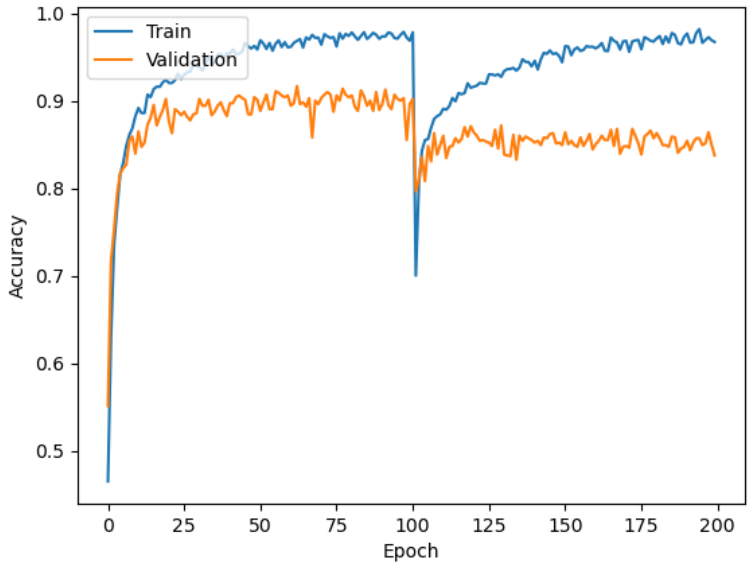



*Figure 18.* Example of band selection accuracy training

To compare the method with the state-of-the-art, the provided bands reported in each article were used. For Wang et al. (2018), the implementation was carried out with the repository code provided by the state-of-the-art articles themselves. Since many of them showed results for only 10 bands, to perform a fair comparison, this criterion was established within the proposed method. For comparison purposes, the same network for the classification task was used(Table 1). The accuracy results from the neural network are presented in Table 4, with each method run with a parameter of 100 epochs. This table summarizes the results from the same seed for each method. All methods were tested under identical conditions, utilizing the same deep-learning network. As a parameter to achieve, the classification result using all bands was used, since it contains all the information from the database. The ” – ” symbolizes that the method does not report the bands for that dataset.

To ensure the repeatability of the proposed method, this thesis offers the selected bands for both the state-of-the-art approach and the proposed method. For selected bands of Indian Pines,

| Results for 10 selected bands | | | | | | |
|---|---|---|---|---|---|---|
| | CNN Model | | | Dense Model | | |
| **Methods** | IP | SAL | PU | IP | SAL | PU |
| Full bands | 0.922 | 0.968 | 0.952 | 0.909 | 0.956 | 0.963 |
| TRC OC FDPC Wang et al. (2018) | 0.844 | 0.932 | 0.848 | 0.848 | 0.926 | 0.852 |
| NC OC MVPCA Wang et al. (2018) | 0.862 | 0.930 | 0.854 | 0.823 | 0.936 | 0.856 |
| NC OC IE Wang et al. (2018) | 0.843 | 0.929 | 0.851 | 0.854 | 0.929 | 0.850 |
| RL Mou et al. (2021) | 0.792 | - | - | 0.811 | - | - |
| RDFBSS Chang et al. (2023) | 0.779 | 0.916 | 0.904 | 0.791 | 0.922 | 0.904 |
| TSC Huang et al. (2022) | 0.856 | - | 0.920 | 0.849 | - | 0.920 |
| ASPS-MN Wang et al. (2019) | - | - | 0.907 | - | - | 0.907 |
| Proposed | **0,866** | **0,941** | **0,928** | **0,857** | **0,940** | **0,928** |

*Table 4.* Comparison of state-of-the-art methods with classification accuracy over 10 trials, the metric used is Overall Accuracy (OA). The database was used with the selected bands by each method, followed by the training to evaluate their performance.

Salinas, and Pavia University, please refer to Table 5. Since the proposed method selects the best bands based on the data distribution, choosing and discarding bands according to the classification task, allows for a better focus on each database. This adaptability not only facilitates its future physical implementation but also enhances its performance in a single specific task, such as classification. This is reflected in the results obtained for both architectures, where in both cases the proposed method outperforms the other state-of-the-art methods. However, increasing the selected bands may decrease the gap between the improvement of one method and the other Fonseca et al. (2023).

| **Method** | **Indian Pines** |
|---|---|
| TRC-OC-FDPCWang et al. (2018) | 8,28,43,50,67,107,118,128,141,173 |
| NC-OC-MVPCAWang et al. (2018) | 17,29,42,48,57,89,117,161,166,176 |
| NC-OC-IEWang et al. (2018) | 17,29,42,48,54,89,117,160,166,176 |
| RLMou et al. (2021) | 65,56,35,194,55,118,61,58,80,185 |
| RDFBSSChang et al. (2023) | 19,29,80,146,58,45,77,98,81,79 |
| TSCHuang et al. (2022) | 15,25,51,72,98,106,123,137,157,167 |
| Proposed | 21,35,69,101,117,133,149,150,165,181 |
| **Method** | **Salinas** |
| TRC-OC-FDPCWang et al. (2018) | 7,11,20,24,32,38,41,46,55,68 |
| NC-OC-MVPCAWang et al. (2018) | 8,15,25,32,45,58,93,119,125,135 |
| NC-OC-IEWang et al. (2018) | 8,15,25,34,45,58,93,120,125,135 |
| RDFBSSChang et al. (2023) | 106,149,11,224,167,7,41,59,151,104 |
| Proposed | 9,25,37,41,57,67,75,81,97,137 |
| **Method** | **Pavia University** |
| TRC-OC-FDPCWang et al. (2018) | 3,8,15,19,29,33,36,48,53, 61 |
| NC-OC-MVPCAWang et al. (2018) | 4,10,17,20,26,36,39,49,58,63 |
| NC-OC-IEWang et al. (2018) | 4,10,17,22,30,36,43,49,58,63 |
| RDFBSSChang et al. (2023) | 72,2,11,1,35,3,20,68,62,26 |
| TSCHuang et al. (2022) | 12,21,29,38,48,58,66,83,92,101 |
| ASPS-MNWang et al. (2019) | 14,24,28,41,54,67,72,87,92, 97 |
| Proposed | 27,35,51,79,83,85,87,91,93,95 |

*Table 5.* The selected indices representing band for all methods with Indian Pines (corrected version), Salinas (corrected version), and Pavia University datasets.

For each dataset, were plotted the respective selected band to examine the more relevant wavelengths in each case, as shown in Fig. 19. This figure shows the two best methods. Note that some bands are repeated in several methods, but can observed that although there are some patterns, can be obtained comparable performance from several combinations. For the Indian Pines dataset, the bands are distributed across the spectrum, with the main focus being on the peaks. For example, around 1500 [nm] both methods share a band, but the other bands focus on different parts of the spectrum. This thesis shows the standardized version to ensure the proposed method worked correctly as the data passed through the deep learning model. For the Salinas dataset, note that the main band is between 400 [nm] and 1500 [nm]. The proposed method finds another important band above 1600 [nm], which can make a difference in difficult cases. On the other hand, the University of Pavia concentrates on higher bands above 750 [nm], close to the near-infrared range.

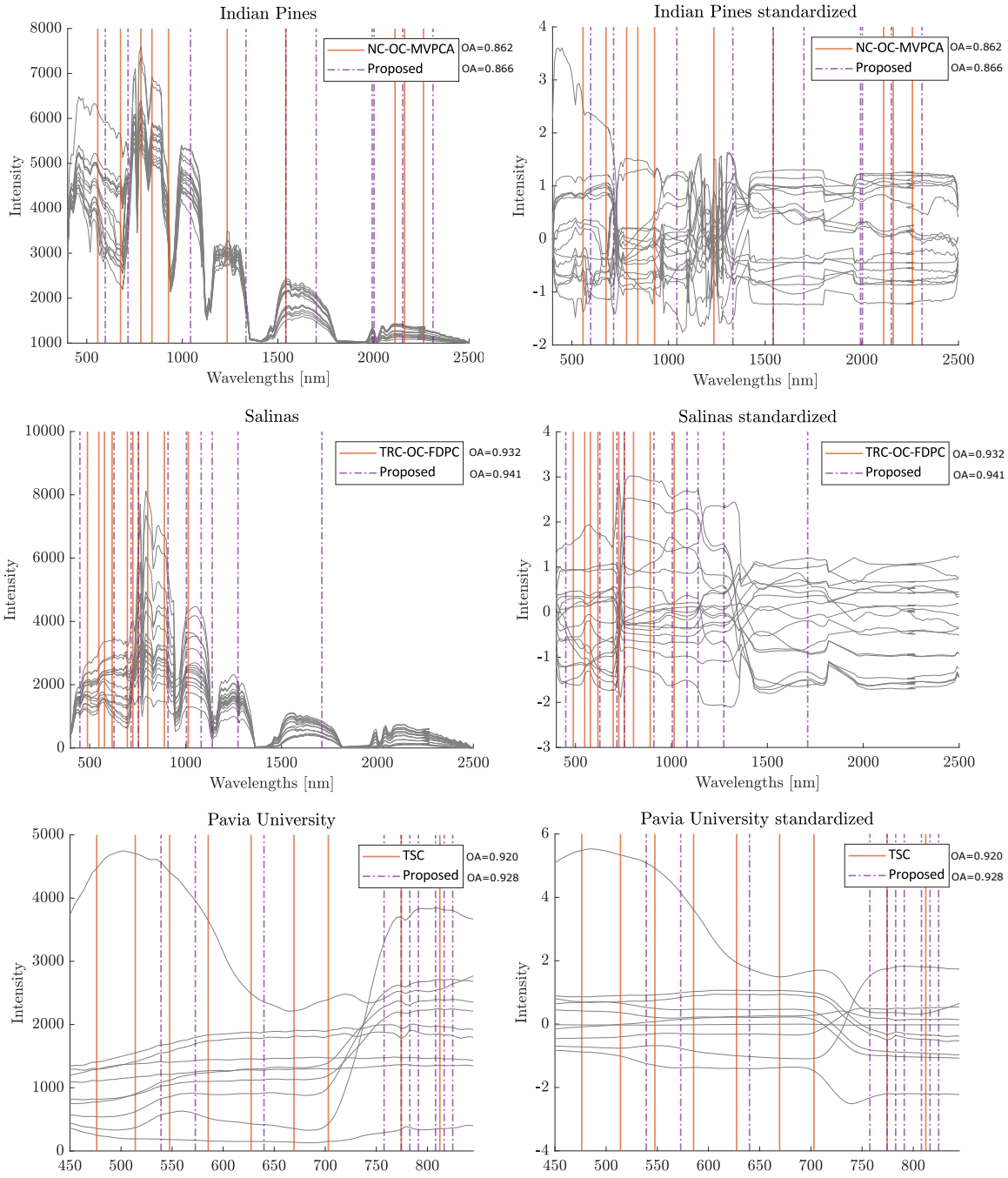



*Figure 19.* Selected bands for Indian Pines, Salinas, and Pavia University datasets. The illustration displays the bands selected by the proposed method alongside those of the second-best method for each dataset. The gray lines are the mean for each class for each dataset. The raw and standardized datasets are presented to facilitate the observation of critical spectral ranges.

To illustrate the importance of the bands, werwanalyze some spectral bands for each dataset, for Indian Pines in Fig. 20, Salinas in Fig. 21, and in Pavia University in Fig. 22. In the case of the Indian Pines, some commonly selected bands were analyzed, such as 117, as shown in Fig. 20, this band shows high intensities in several classes of the database, different from bands 35 and 133, which although they mainly highlight the background classes, their intensities are different. On the other hand, band 200, taken as a sample of a non-selected band, note that it does not provide significant information about the classes, this one in particular is mainly noise. On the other hand, in the Salinas database shown in Fig. 21, bands 41, 67, and 68 were selected by the best methods, showing a clear difference in the intensities of each material. Note that the difference between 67 and 68 is not easily highlighted visually, so it can be concluded that each method selects the bands that highlight spectral ranges that allow one to find the regions that differentiate the materials. Finally, the Pavia University database shown in Fig. 22 illustrates bands 10 and 35 as some of the most selected bands, which, for example, allows identifying tile materials easily. Each of these best methods contains at least one band in a higher spectral range that complements this information as bands 83 and 101.

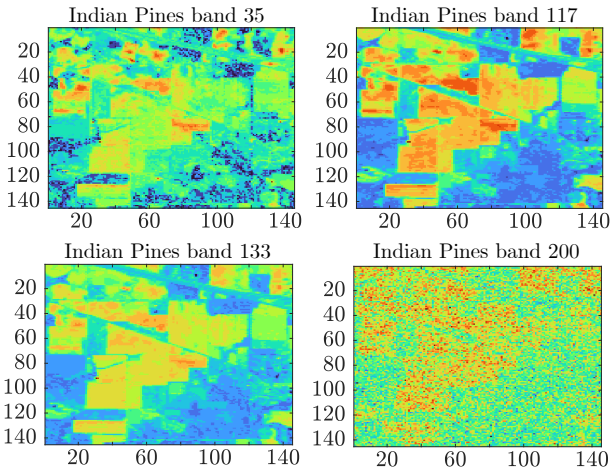



*Figure 20.* Comparison of bands in the Indian Pines dataset. Band 117, selected by multiple methods, band 133, the sole selection of the proposed proposed method, and band 200, not selected by any method.

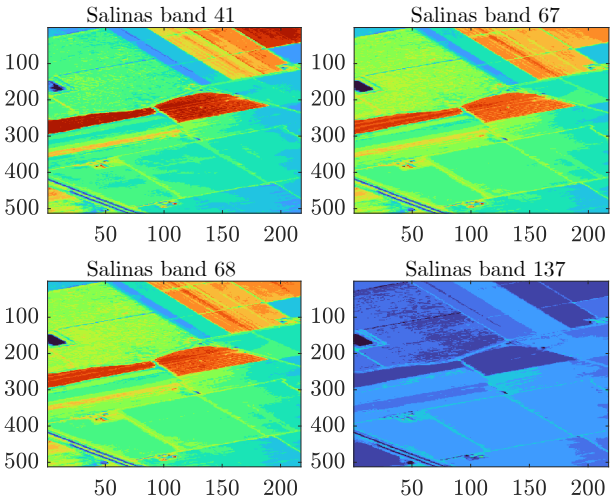



*Figure 21.* Comparison of bands in the Salinas dataset. Band 41 was selected as one of the most repetitive bands. Bands 68 and 67, very close to each other, were selected by different methods. Band 137 was the most distant band selected by any method.

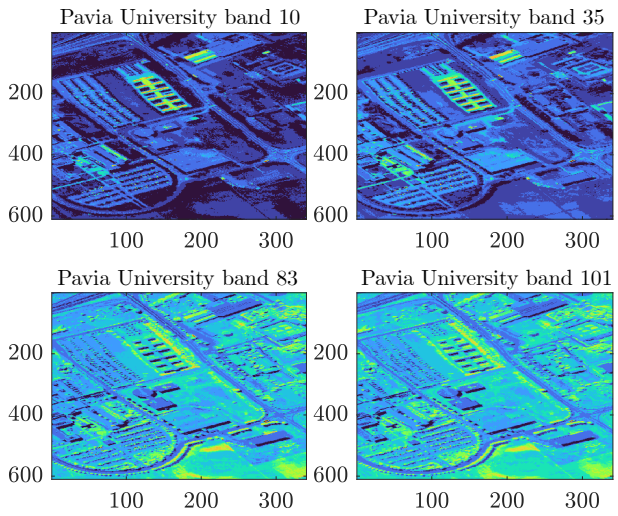



*Figure 22.* Comparison of bands in Pavia University dataset. Bands 10 and 35 were selected frequently. Band 83 was selected by the two best methods. Band 101, selected as the second-best method, did not appear in any other method's selection.

### 12.2. Cross-validation by dataset fold.

To test the robustness of the proposed method, a validation was performed between 5 different datasets. The test was conducted by defining different trials of partitions on the databases. The results show that the proposed method has a satisfactory generalization for the different data as shown in Table 6.

| **Dataset** | **Overall Accuracy** |
|---|---|
| Indian Pines | $0.862 \pm 0.010$ |
| Salinas | $0.940 \pm 0.001$ |
| Pavia University | $0.930 \pm 0.0008$ |

*Table 6.* Database cross-validation results. Selected bands proved on different dataset distributions.

## 12.3. Acquired datasets

**12.3.1. Spectrometer optical system.** The experiments were conducted as follows: first, were created the database from the acquired data, assigning a class number to each material. Once the database was obtained, was run the band selection algorithm to obtain different numbers of bands. To verify the validity of the selected bands were used the classification neural network shown in Table 2 and 5 trials with different seeds to ensure its robustness. For the experiments, several numbers of selected bands, such as 3,4,5,12,13,15, and 40 to evaluate the performance of different band combinations. As shown in Figure 23, with 30 bands, less than 2%, the proposed method can achieve a performance that is compared to the full bands by only 4 points under the baseline.

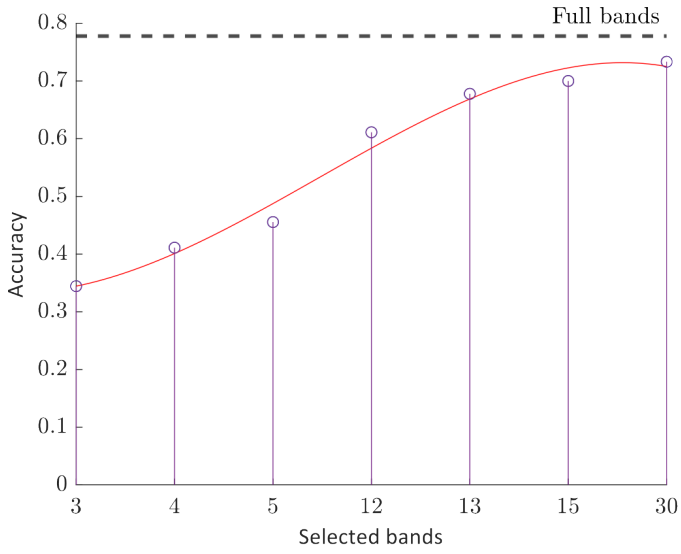



*Figure 23.* Number of bands vs. accuracy obtained in classifying the acquired materials with the spectrometer optical system. The results with full bands are made with 2048 bands

Since optical acquisition systems depend on the electromagnetic spectrum range for their

design, since optical elements are designed for a specific range, the wavelength value selected by the algorithm in Fig. 24, where the 6 average signatures of each class and the 15 selected bands. The most relevant ranges for these elements are located from approximately 500 [nm] to 700 [nm].

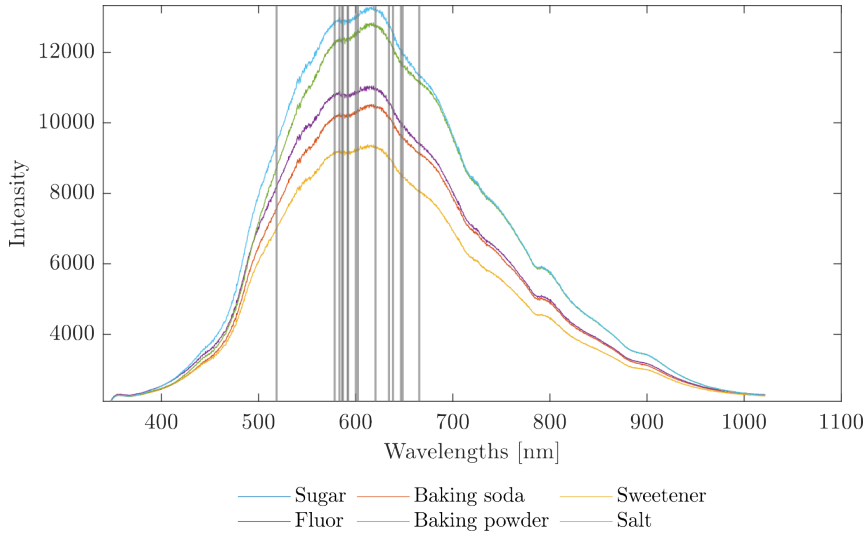



*Figure 24.* Selected bands in the electromagnetic spectrum. The figure shows the average signatures of the 6 classes acquired materials with the spectrometer optical system, labeled in different colors at the bottom. The bands that were selected by the algorithm are shown in gray.

The matrices in Fig 25 show that even with all the bands, there are materials that get confused, such as salt and flour. Likewise, some materials can be easily distinguished from 13 bands in the evaluation, such as baking powder.

**12.3.2. Whiskbroom optical system.** The experiments were conducted as follows: first, the database from the acquired data was created, assigning a class number to each material. Once the database was obtained, were ran the band selection algorithm to obtain different numbers of bands. To verify the validity of the selected bands, the classification neural network shown in Table 2 was used and performed 5 trials with different seeds to ensure its robustness. For the

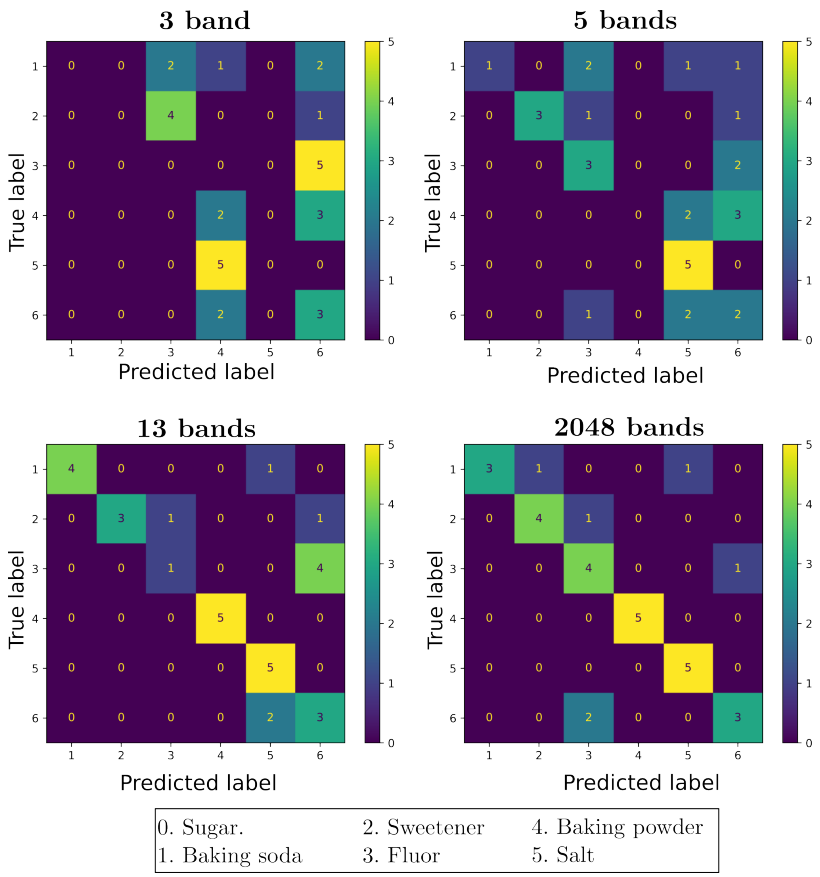



*Figure 25.* The confusion matrices of different numbers of selected bands with different classification performances to verify the materials with the highest level of confusion in the neural network model of the data acquired materials with the spectrometer optical system.

.

experiments, were used several numbers of selected bands, such as 1, 2, 3, 4, 7, 8, and 18, to evaluate the performance of different band combinations. As shown in Figure 26, with only 4 bands, the proposed method can achieve the same performance as the obtained with all bands.

This indicates that by capturing only these 4 bands, the proposed method can obtain results comparable to those obtained by capturing a wide range of the electromagnetic spectrum. Reducing the number of detected bands can optimize the optical system, as sensors can be designed to detect only the most important ranges of material properties which improve acquisition times and

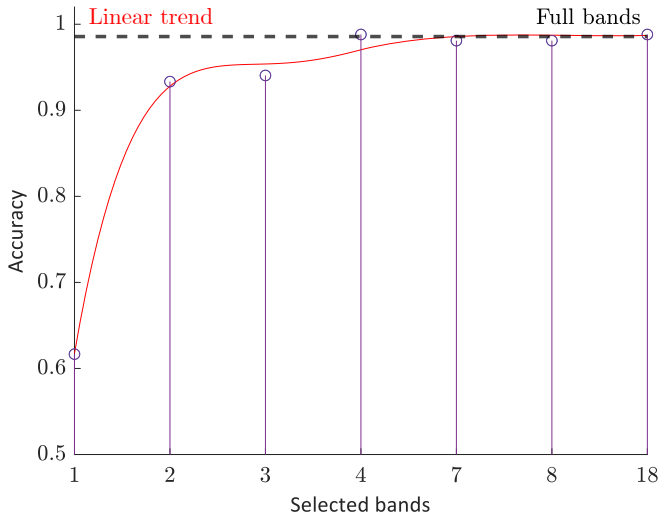



*Figure 26.* Number of bands vs. accuracy obtained in classifying the acquired materials. Note that after 4 bands, the task performance is the same as using all acquired bands. The results with full bands are made with 525 bands

simplifies optical acquisition systems.

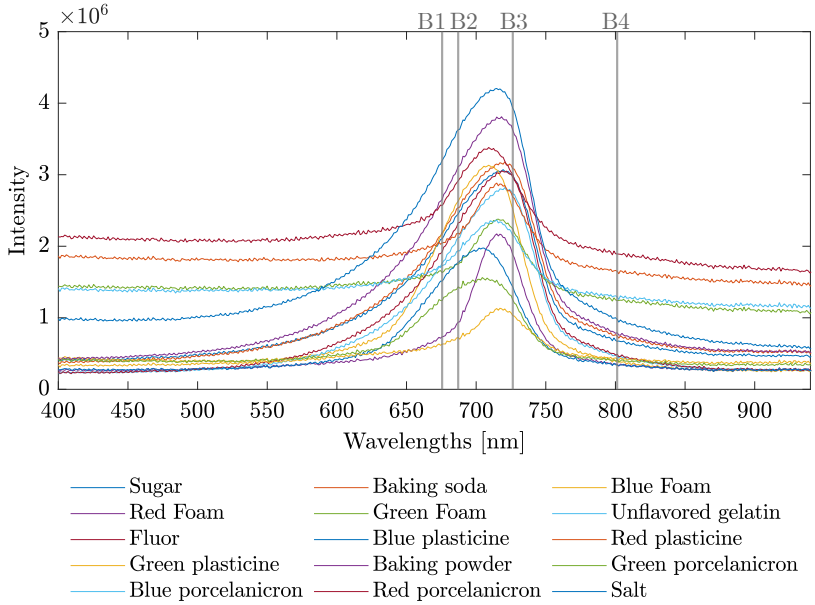
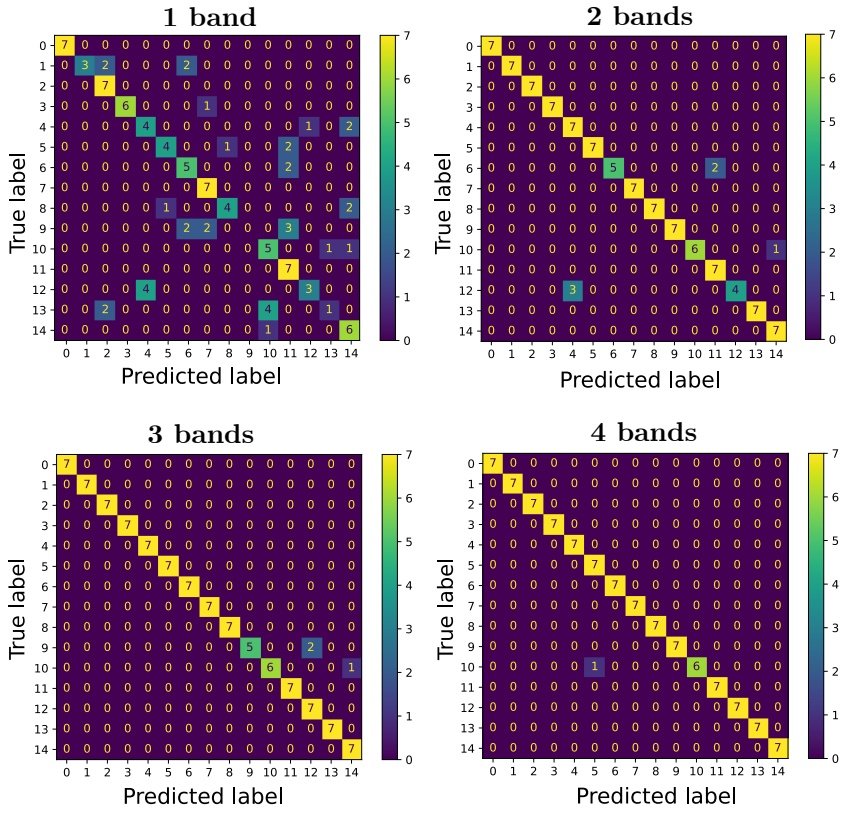



*Figure 27.* Selected bands in the electromagnetic spectrum. The figure shows the average signatures of the 15 classes labeled in different colors at the bottom. The bands that were selected by the algorithm are shown in gray.

Since optical acquisition systems depend on the electromagnetic spectrum range for their design, since optical elements are designed for a specific range, It was examined the wavelength value selected by the algorithm in Fig. 27, where the 15 average signatures of each class and the 4 selected bands. The most relevant ranges for these elements are located from approximately 650 [nm] to 800 [nm], visible (VIS) range reaching the near-infrared (NIR) spectrum.




| | | | |
|---|---|---|---|
| 0. Sugar. | 4. Green Foam | 8. Red plasticine | 12. Blue porcelanicron |
| 1. Baking soda | 5. Unflavored gelatin | 9. Green plasticine | 13. Red porcelanicron |
| 2. Blue Foam | 6. Fluor | 10. Baking powder | 14. Salt |
| 3. Red Foam | 7. Blue plasticine | 11. Green porcelanicron | |

*Figure 28.* The confusion matrices of different numbers of selected bands with different classification performances to verify the materials with the highest level of confusion in the neural network model.

Figure 28 demonstrates that the model can confuse materials that differ in color and mate-

rial when using only 1 band. However, materials with only one intensity, such as sugar, can still be accurately identified. The matrices in Fig 28 also illustrate some anticipated confusions with 2 and 3 bands, such as between baking powder and salt, both of which are white materials, or between different materials with similar colors, such as green foam and blue porcelanicron. The final confusion matrix with 4 bands only misclassified two materials: unflavored gelatin and baking powder, these two materials are visually similar, and also have similar spectrum responses.

## 13. Conclusion and Discussion

Spectral band selection is a critical method for extracting essential information from images or spectral signatures. The proposed approach integrates band selection and classification within the trainable layers of the network, seeking optimal convergence that serves both goals for spectral signatures. Results show that the proposed method outperforms other state-of-the-art methods on three different datasets and two different architectures, illustrating its robustness. In addition, it shows comparable results with full band acquisition in two acquired spectral signature databases. However, it requires fine-tuning their respective parameters, which depend on the underlying data structure. Nevertheless, the results demonstrate the efficacy of optical implementation and emulation for several materials. These results offer the opportunity to reduce data redundancy and simplify the complexity of optical systems, which can mean less expensive equipment and acquisition time.

## 14. Expected results

| Expected result | Indicator | State |
|---|---|---|
| Research accepted in conference | Letter of acceptance by the conference. | Complete, work accepted in the Optical Imaging Congress on COSI (Computational Optical Sensing and Imaging) under the title of ”Joint Deep Learning Optical Band Selection and Classification Method for Spectral Data”.<br><br>Additional:<br>Work submitted to EUSIPCO 2024, under the name .ºptical Spectral Band Selection System for Material Classification” |
| Submission of research article to journal. | Articles submitted | Complete, work submitted to the Applied Optics journal, under the submission ID 523199, and the name ”Deep Jointly Optical Spectral Band Selection and Classification Learning” |
| Project research book | Book submitted to UIS library | In progress |

*Table 7.* Expected Products of the project

## Bibliographic References

Aiazzi, B., Alparone, L., Barducci, A., Baronti, S., Marcoionni, P., Pippi, I., and Selva, M. (2006). Noise modelling and estimation of hyperspectral data from airborne imaging spectrometers. *Annals of Geophysics*, 49(1).

Arguello, H., Bacca, J., Kariyawasam, H., Vargas, E., Marquez, M., Hettiarachchi, R., Garcia, H., Herath, K., Haputhanthri, U., Ahluwalia, B. S., et al. (2023). Deep optical coding design in computational imaging: a data-driven framework. *IEEE Signal Processing Magazine*, 40(2):75–88.

Audebert, N., Le Saux, B., and Lefèvre, S. (2019). Deep learning for classification of hyperspectral data: A comparative review. *IEEE geoscience and remote sensing magazine*, 7(2):159–173.

Bacca, J., Gelvez-Barrera, T., and Arguello, H. (2021). Deep coded aperture design: An end-to-end approach for computational imaging tasks. *IEEE Transactions on Computational Imaging*, 7:1148–1160.

Bacca, J., Hernandez-Rojas, A., and Arguello, H. (2022). Deep coding patterns design for compressive near-infrared spectral classification. In *2022 30th European Signal Processing Conference (EUSIPCO)*, pages 548–552. IEEE.

Bacca, J., Martinez, E., and Arguello, H. (2023). Computational spectral imaging: a contemporary overview. *JOSA A*, 40(4):C115–C125.

Chang, C.-I., Kuo, Y.-M., and Hu, P. F. (2023). Unsupervised rate distortion function-based band subset selection for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*.

Cull, C. F., Choi, K., Brady, D. J., and Oliver, T. (2010). Identification of fluorescent beads using a coded aperture snapshot spectral imager. *Applied optics*, 49(10):B59–B70.

Elihos, A., Alkan, B., Balci, B., and Artan, Y. (2018). Comparison of image classification and object detection for passenger seat belt violation detection using nir & rgb surveillance camera images. In *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6. IEEE.

Fang, J., Huang, K., Qin, R., Liang, Y., Wu, E., Yan, M., and Zeng, H. (2024). Wide-field mid-infrared hyperspectral imaging beyond video rate. *Nature Communications*, 15(1):1811.

Fonseca, K., Garcia, H., da Silva, F., Arguello, H., and Bacca, J. (2023). Joint deep learning optical band selection and classification method for spectral data. In *Optica Imaging Congress (3D, COSI, DH, FLatOptics, IS, pcAOP)*, page CTh2A.4. Optica Publishing Group.

Fu, H., Zhang, A., Sun, G., Ren, J., Jia, X., Pan, Z., and Ma, H. (2022). A novel band selection and spatial noise reduction method for hyperspectral image classification. *IEEE transactions on geoscience and remote sensing*, 60:1–13.

Garini, Y., Young, I. T., and McNamara, G. (2006). Spectral imaging: principles and applications.

*Cytometry Part A: The Journal of the International Society for Analytical Cytology*, 69(8):735–747.

Glasmachers, T. (2017). Limits of end-to-end learning. In *Asian conference on machine learning*, pages 17–32. PMLR.

Goetz, A. F., Vane, G., Solomon, J. E., and Rock, B. N. (1985). Imaging spectrometry for earth remote sensing. *science*, 228(4704):1147–1153.

Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., Faust, J. A., Pavri, B. E., Chovit, C. J., Solis, M., et al. (1998). Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris). *Remote sensing of environment*, 65(3):227–248.

Gupta, N. (2011). Development of staring hyperspectral imagers. In *2011 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–8. IEEE.

Hakkel, K. D., Petruzzella, M., Ou, F., van Klinken, A., Pagliano, F., Liu, T., van Veldhoven, R. P., and Fiore, A. (2022). Integrated near-infrared spectral sensing. *Nature communications*, 13(1):1–8.

Huang, S., Zhang, H., Xue, J., and Pižurica, A. (2022). Heterogeneous regularization-based tensor subspace clustering for hyperspectral band selection. *IEEE Transactions on Neural Networks and Learning Systems*.

Kamruzzaman, M., ElMasry, G., Sun, D.-W., and Allen, P. (2011). Application of nir hyperspectral imaging for discrimination of lamb muscles. *Journal of food engineering*, 104(3):332–340.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Li, K., Dai, D., and Van Gool, L. (2023). Jointly learning band selection and filter array design for hyperspectral imaging. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 6384–6394.

Liu, Y., Li, X., Hua, Z., Xia, C., and Zhao, L. (2022). A band selection method with masked convolutional autoencoder for hyperspectral image. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5.

Minet, J., Taboury, J., Péalat, M., Roux, N., Lonnoy, J., and Ferrec, Y. (2010). Adaptive band selection snapshot multispectral imaging in the vis/nir domain. In *Electro-Optical Remote Sensing, Photonic Technologies, and Applications IV*, volume 7835, pages 294–303. SPIE.

Mou, L., Saha, S., Hua, Y., Bovolo, F., Bruzzone, L., and Zhu, X. X. (2021). Deep reinforcement learning for band selection in hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14.

Mou, L., Saha, S., Hua, Y., Bovolo, F., Bruzzone, L., and Zhu, X. X. (2022). Deep reinforcement learning for band selection in hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14.

Nesterov, Y. (1983). A method for solving the convex programming problem with convergence rate o$(1/k^2)$. *Proceedings of the USSR Academy of Sciences*, 269:543–547.

Otsuka, M. (2006). Near-infrared spectroscopy application to the pharmaceutical industry. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*, pages 1–15.

Paoletti, M., Haut, J., Plaza, J., and Plaza, A. (2019). Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158:279–317.

Rodionova, O. Y. and Pomerantsev, A. (2010). Nir-based approach to counterfeit-drug detection. *TrAC Trends in Analytical Chemistry*, 29(8):795–803.

Roy, S. K., Krishna, G., Dubey, S. R., and Chaudhuri, B. B. (2019). Hybridsn: Exploring 3-d–2-d cnn feature hierarchy for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 17(2):277–281.

Sakudo, A. (2016). Near-infrared spectroscopy for medical applications: Current status and future perspectives. *Clinica Chimica Acta*, 455:181–188.

Salamati, N., Fredembach, C., and Süsstrunk, S. (2009). Material classification using color and nir images. In *Color and Imaging Conference*, pages 216–222. Society for Imaging Science and Technology.

Sun, W. and Du, Q. (2019). Hyperspectral band selection: A review. *IEEE Geoscience and Remote Sensing Magazine*, 7(2):118–139.

Tao, W. and Farokhzad, O. C. (2022). Theranostic nanomedicine in the nir-ii window: Classification, fabrication, and biomedical applications. *Chemical Reviews*, 122(6):5405–5407.

Van Den Broek, W., Derks, E., Van De Ven, E., Wienke, D., Geladi, P., and Buydens, L. (1996). Plastic identification by remote sensing spectroscopic nir imaging using kernel partial least squares (kpls). *Chemometrics and Intelligent Laboratory Systems*, 35(2):187–197.

Wang, Q., Li, Q., and Li, X. (2019). Hyperspectral band selection via adaptive subspace partition strategy. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(12):4940–4950.

Wang, Q., Zhang, F., and Li, X. (2018). Optimal clustering framework for hyperspectral band selection. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10):5910–5922.

Williams, P., Geladi, P., Fox, G., and Manley, M. (2009). Maize kernel hardness classification by near infrared (nir) hyperspectral imaging and multivariate data analysis. *Analytica Chimica Acta*, 653(2):121–130.

Zareef, M., Chen, Q., Hassan, M. M., Arslan, M., Hashim, M. M., Ahmad, W., Kutsanedzie, F. Y., and Agyekum, A. A. (2020). An overview on the applications of typical non-linear algorithms coupled with nir spectroscopy in food analysis. *Food Engineering Reviews*, 12(2):173–190.