European Journal of Remote Sensing

ISSN: 2279-7254 (Online) Journal homepage: www.tandfonline.com/journals/tejr20

Spectral image data aggregation for multisource data augmentation

R. I. Luca, A. Baicoianu & I. C. Plajer

To cite this article: R. I. Luca, A. Baicoianu & I. C. Plajer (2025) Spectral image data aggregation for multisource data augmentation, European Journal of Remote Sensing, 58:1, 2492295, DOI: 10.1080/22797254.2025.2492295

To link to this article: https://doi.org/10.1080/22797254.2025.2492295

Published online: 18 Apr 2025.

EUROPEAN JOURNAL OF REMOTE SENSING
2025, VOL. 58, NO. 1, 2492295
https://doi.org/10.1080/22797254.2025.2492295

Spectral image data aggregation for multisource data augmentation

R. I. Luca, A. Baicoianu and I. C. Plajer

Faculty of Mathematics and Computer Science, Transilvania University of Braşov, Braşov, Romania

ABSTRACT

Multispectral and hyperspectral images are increasingly popular in different research fields, such as remote sensing, astronomical imaging, or precision agriculture. However, the amount of free data available to perform machine learning tasks is relatively small. Moreover, artificial intelligence models developed in the area of spectral imaging require input images with a fixed spectral signature, expecting the data to have the same number of spectral bands or the same spectral resolution. This requirement significantly reduces the number of usable sources that can be used for a given model. The scope of this study is to introduce a methodology for spectral image data aggregation, in order to allow machine learning models to be trained and/or used on data from a larger number of sources, thus providing better generalization. For this purpose, we propose different interpolation techniques, in order to make multisource spectral data compatible with each other. The interpolation outcomes are evaluated through various approaches. This includes direct assessments using surface plots and metrics such as a Custom Mean Squared Error and the Normalized Difference Vegetation Index. Additionally, indirect evaluation is done by estimating their impact on machine learning model training, particularly for semantic segmentation.

ARTICLE HISTORY
Received 15 August 2024
Revised 18 December 2024
Accepted 6 April 2025

KEYWORDS
Spectral images; interpolation; data aggregation; multisource data; neural networks

Introduction

The ongoing progress in remote sensing sensors helps us to better understand various phenomena around us (Brezini & Deville, 2023). Each type of material possesses different spectral characteristics that alter how the light is absorbed. This leads to unique spectral fingerprints. Unlike RGB images that only have three spectral bands, multispectral (MS) and hyperspectral (HS) sensors are capable of capturing data from tens or even hundreds of different wavelengths. As a result, each pixel in the image holds a wealth of information, including spectral data related to the chemical composition of the objects (Plajer, B˘aicoianu, & Majercsik, 2023).

These HS sensors can be found on spaceborne platforms like Hyperion (Middleton et al., 2013), the Precursore Iperspettrale della Missione Applicativa (PRISMA) (Loizzo et al., 2018), or the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) (Pancorbo et al., 2023), or on aircraft equipped with sensors such as the Airborne Prism Experiment (APEX) (Schaepman et al., 2015). The information given by the high spectral resolution sensors is extremely valuable for enabling accurate classifications and detection of pure materials. Therefore, MS and HS images can be very useful in numerous domains like agriculture (Tuzzi et al., 2023), coastal areas monitoring (Guillaume et al., 2020), mineral detection (Benhalouche et al., 2022), and military applications (Kartikaningsih et al., 2021).

Although the interest in spectral images is growing and the public availability of more and more satellite data increases there are serious limitations in using them for machine learning tasks, as the different satellites and other equipment produce images with different spectral signatures and spatial resolutions. Additionally, in this context, labeled data for training and testing the models is sparse, which further restricts the datasets that can be practically used. A common approach to expand the training data for an ML model is data augmentation (Gao et al., 2024). While this technique is highly beneficial in many scenarios, it is less effective for very small datasets, as it generates only a limited amount of new information. Additionally, data augmentation does not enable the model to generalize its inference to data with spectral signatures that differ from those of the training samples. Therefore, we can say that is exceptionally difficult to generalize models for the available data without a preprocessing step. Such a step should enable the aggregation of the data based on their characteristics, leading to a uniform and consistent dataset that can be used in any machine learning model. Furthermore, it could allow a model trained on a certain dataset to infer spectral data from other sources or sensors. However, research in this area remains limited, and a standardized solution has yet to be established. Additionally, it is crucial to analyze the available data

CONTACT A. Baicoianu a.baicoianu@unitbv.ro Faculty of Mathematics and Computer Science, Transilvania University of Braşov, B-dul Eroilor 29, Braşov 500036, Romania

and objectively assess whether aggregation is feasible for a specific task, such as segmentation.

Semantic segmentation of spectral images is an important operation in the area of remote sensing. It plays a fundamental role in tasks such as land cover classification, monitoring environmental changes, and overseeing military spaces by labeling each pixel to its respective category. Lately, this task has gained significant attention in the sector of machine learning research. Traditional approaches require human assistance to identify and design the features that will then be used for image segmentation. In contrast, machine learning-based methods can automatically perform this task in an end-to-end manner by training specific neural networks (Liang & Wang, 2022; Ma et al., 2023; Saxena & Raman, 2020). These networks can range from shallow ones, such as Fully Connected Neural Networks (FCNN) (Chen et al., 2014), to more complex and deep ones, often with better results in semantic segmentation, such as Convolutional Neural Networks (CNN) CNN) (Benhida et al., 2022).

In recent years a series of approaches to spectral data aggregation were done by means of deep learning super-resolution, using CNNs Zheng et al. (2020), model-informed networks Li et al. (2024) and auto-encoders Gao et al. (2023); Li et al. (2023).

Although machine learning-based aggregation of spectral images is highly valuable and yields good results, it also involves significant complexity and requires substantial computational resources. Additionally, these methods are dependent on the availability of training data.

In this paper, we introduce a simple and cost-efficient methodology to integrate HS data from different sources into a single set, aiming to create a larger collection of images that can be used in different machine-learning algorithms. The process of data aggregation has always been a complex task and still represents a challenge. It has been extensively researched in machine learning and big data, aiming to optimize and improve the overall data quality. Furthermore, this operation is considered to be an integral component of the applications developed in the present (Divya & Manish, 2020; Zhang, Xie et al., 2018).

Our approach differs from synchronization methods such as that of (Nalepa et al., 2020), which reduce the number of spectral bands and apply transfer learning specifically for semantic segmentation. By contrast, our method focuses on enriching spectral information through the generation of additional bands, enabling broader applicability across tasks like classification, reconstruction, or colorization. While semantic segmentation is used as an example in our study, our approach is task-agnostic and addresses the more general challenge of improving model generalization for spectral data with unseen signatures.

Moreover, in this study we analyze the available datasets, identify the similarities and differences, choose a reference dataset, and execute an interpolation operation over the spectral bands, in order to aggregate all images into a single dataset. Furthermore, to demonstrate this approach's utility, correctness, and accuracy, we also introduce two neural networks specialized in semantic segmentation, which are trained and tested on the aggregated dataset. The networks are also tested on images from other sources, which were not used for the training. Since both FCNNs and CNNs are employed as backbones in the field of spectral image classification and semantic segmentation, we will evaluate our results on both types.

To provide a comprehensive understanding of our study, the manuscript is structured as follows: the Introduction Section reviews the relevant literature on spectral image data aggregation and multisource data augmentation, providing the foundation for our approach. In the Materials and Methods Section, we describe the datasets used, the data aggregation techniques applied, and the validation methodologies employed. The Results and Discussions Section details two main approaches for validating the proposed method: visual inspection and quantitative measures, as well as analyzing the impact of the methodology on segmentation accuracy using neural networks. Finally, the Conclusions Section summarizes the main findings and outlines potential directions for future research.

Materials and methods

To combine spectral data from various sources, our initial step involved a comprehensive analysis of the publicly available HS datasets, trying to understand their characteristics and selecting those that we considered appropriate for our study. Then, we aggregated the data using various interpolation methods to create a unitary dataset.

Two key methods were used to validate the suggested methodology for spectral integration of the data. Through direct comparison of the interpolated spectral image with the original using some of the metrics outlined in Section 3, and indirectly through an examination of the segmentation accuracy of the networks trained on the combined datasets.

Dataset description

The following public HS datasets, with various spectral signatures, were used for the experiments. Firstly we selected Pavia University [dataset] (Graña et al., 2021), Kennedy Space Center (KSC)[dataset] (Graña et al., 2021), Botswana [dataset] (Graña et al., 2021), and Indian Pines [dataset] (Graña et al., 2021), which

also come with labels for classification. These were chosen because they share some of the classes, enabling the testing of the interpolation results on classification tasks with neural networks. In order to further analyze the accuracy of the different interpolation methods, we additionally used the well-known Columbia Imaging and Vision Laboratory (CAVE) [dataset] (Yasuma et al., 2008) and the Universidad de Granada (UGR) dataset] (Eckhard et al., 2015).

Given the difficulty in finding several HS images with the same labeled classes, we decided to combine the existing labels into two main categories that can be easily found in multiple scenes: *Vegetation* and *Non-Vegetation*. For each of the datasets, an illustration of the original labels and the fused ones is provided in the following.

Pavia university

The Pavia University image was obtained in 2001 by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor. It was captured during a flight campaign over Pavia, Northern Italy. The uncorrected data consists of 610×610pixels and 115 spectral bands between 430 and 860 nm, with a spectral resolution of 4 nm. However, some samples did not contain useful information, so they were removed. This resulted in a corrected dataset of 610×340pixels with 103 spectral bands and a spatial resolution of 1.3 m per pixel. The image is divided into nine ground truth classes. Figure 1 illustrates the results for these classes as well as for the two ones considered in our study.

This dataset is frequently explored in the literature for tasks related to image classification using different types of CNNs. An example of this is the Multiscale Spectral-Spatial Convolutional Neural Network presented in (Xu et al., 2021).

Kennedy space center (KSC)

The KSC image, captured in 1996 by the NASA Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor, provides a view over the Kennedy Space Center, Florida. It contains a wide wavelength range of 224 spectral bands between 400 and 2500 nm, with a spectral resolution of 10 nm. Among these, 48 water absorption and low signal-to -noise ratio (SNR) bands were removed, resulting in 176 bands available for analysis. The image, with a resolution of $512 \times 614$and a spatial resolution of 18 m per pixel, originally contained 13 ground truth classes. The ground truth for these classes, as well, as for the two classes considered by our study, are presented in Figure 2.

KSC is well-known for its applications in classification tasks (Fang et al., 2022) but also in the field of band selection based on different criteria. The study presented in (Jia et al., 2012) uses this dataset to demonstrate the accuracy of extracting discriminative properties from HS images.

Botswana

The Botswana image was collected in 2001 by the Hyperion sensor on EO-1 over the Okavango Delta, Botswana. The original uncorrected data consists of 1476×256pixels, each with a spatial resolution of 30 m and 242 spectral bands, spanning from 400 to 2500 nm, with a sampling interval of 10 nm. After the removal of uncalibrated and noisy bands, 145 bands remain for further use. The original 14 labels and the two merged labels are presented in Figure 3.

Apart from classification tasks, Botswana served as an essential tool for experimental tasks in the field of hyperspectral unmixing Botswana_unmixing.

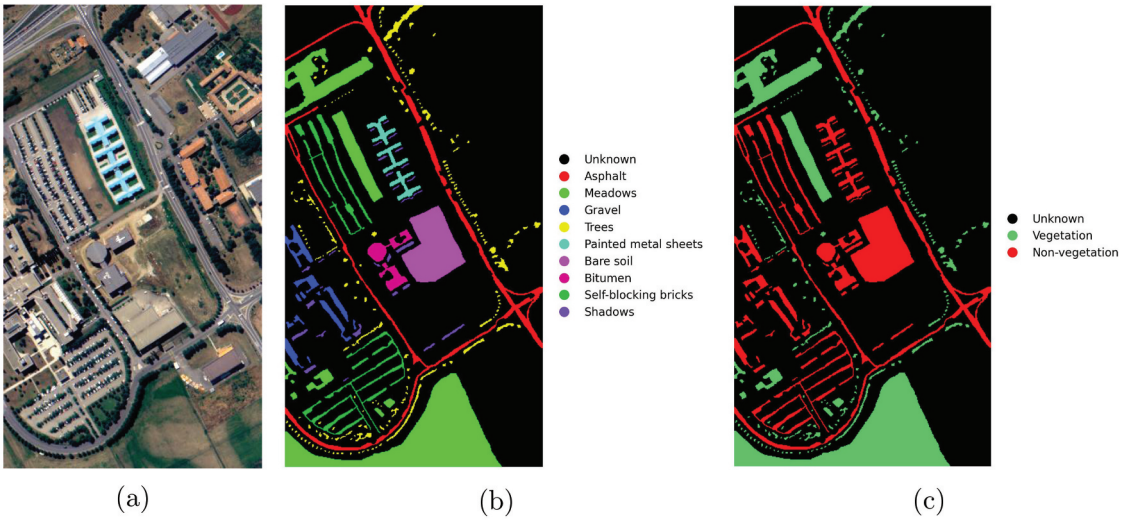


Figure 1. Pavia university: (a) visualization using 3 bands; (b) original ground truth; (c) processed ground truth (*black* for unknown; *green* for vegetation; *red* for non-vegetation).

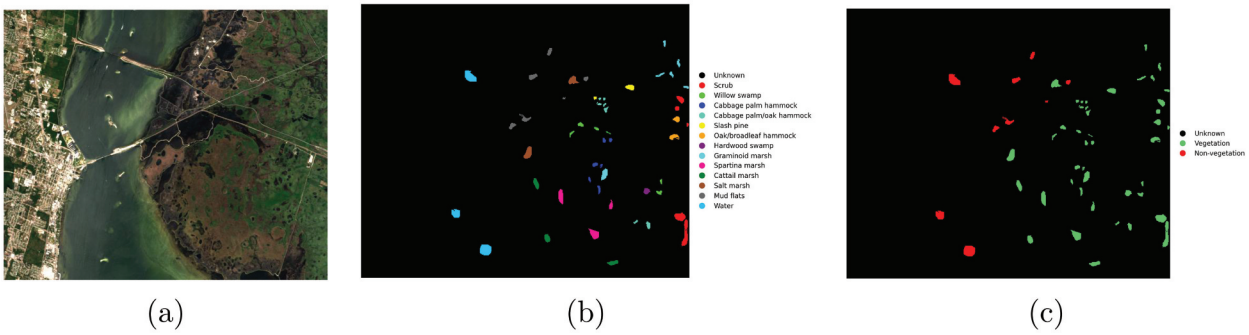


Figure 2. KSC image: (a) visualization using 3 bands; (b) original ground truth; (c) processed ground truth (*black* for unknown; *green* for vegetation; *red* for non-vegetation).

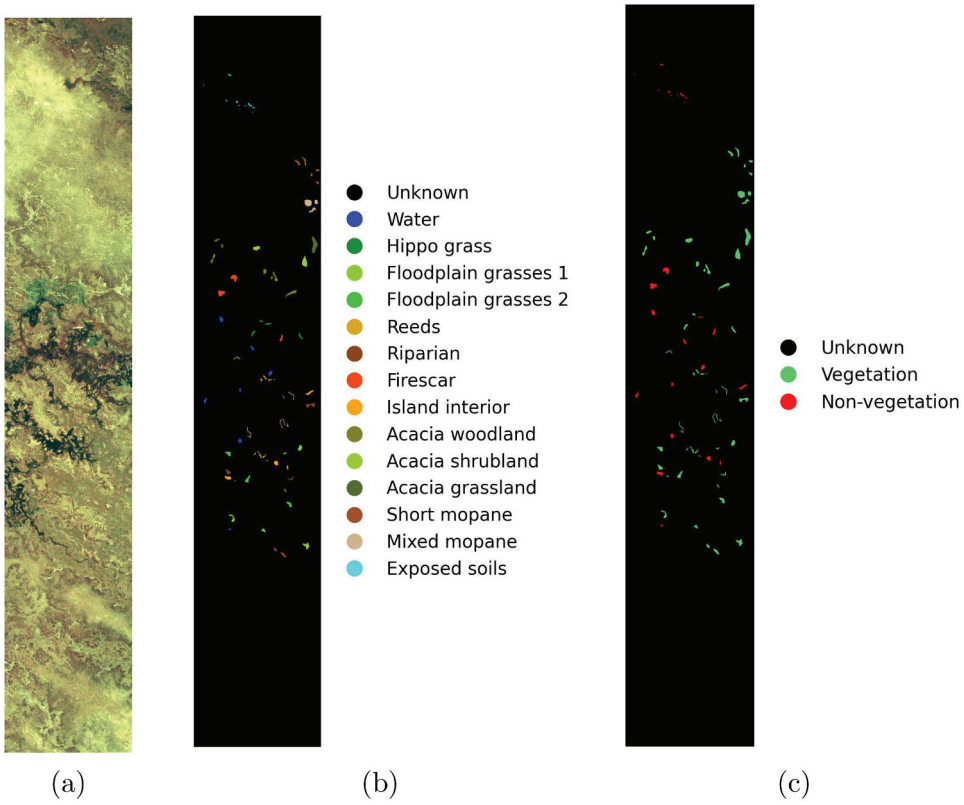


Figure 3. Botswana image: (a) visualization using 3 bands; (b) original ground truth; (c) processed ground truth (*black* for unknown; *green* for vegetation; *red* for non-vegetation).

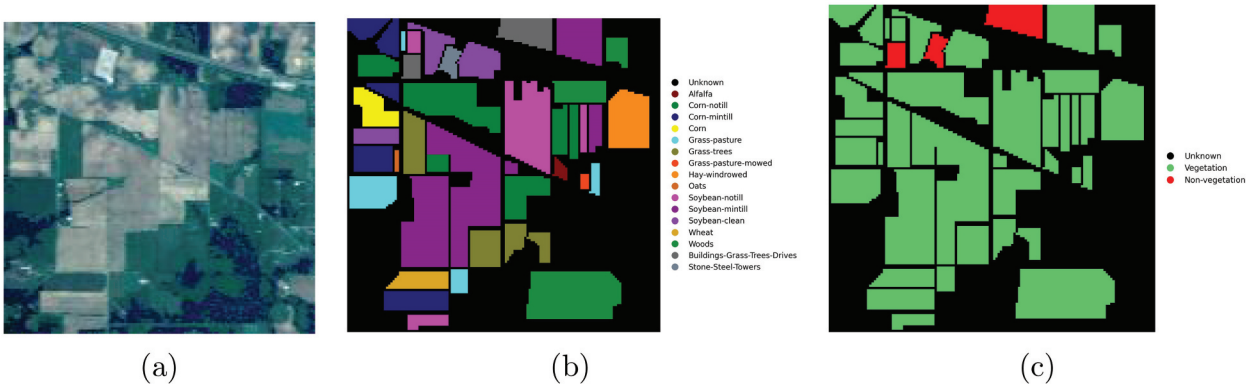


Figure 4. Indian Pines image: (a) visualization using 3 channels; (b) original ground truth; (c) processed ground truth (*black* for unknown; *green* for vegetation; *red* for non-vegetation).

Indian Pines

The Indian Pines dataset was gathered in 1992 by the AVIRIS sensor over the Indian Pines test site in northwestern Indiana. The uncorrected data consists of 145 × 145pixels, with a spatial resolution of 20 m. The 224 spectral bands range from 400 to 2500 nm, having a spectral resolution of 10 nm. After removing the bands that cover the water absorption region, 200 bands remained in the corrected dataset. Initially, there were 16 ground truth classes. Figure 4 shows these classes and the merged two classes considered.

The Indian Pines dataset represents a valuable hyperspectral image used across a broad spectrum of applications. These include classification tasks using Fast 3D CNNs, as presented in (Ahmad et al., 2022). It also finds extensive use in the field of super-resolution classification (Zhang, Zhang et al., 2018) and band selection studies (Wang et al., 2022).

CAVE

The CAVE dataset consists of 32 spectral scenes, each of 512 × 512pixels, captured using a cooled Charge-Coupled Device (CCD) camera. There are 31 spectral bands, spanning from 400 nm to 700 nm, with a spectral resolution of 10 nm. All scenes were captured indoors under controlled lighting conditions. There are no segmentation labels provided.

This dataset is appropriate for applications in the domain of RGB visualization (Plajer, Băicoianu, & Majercsik, 2023), using various artificial intelligence methodologies. Additionally, it is used for the aggregation of spectral channels (Lu et al., 2021) to improve image quality.

UGR

The UGR dataset contains 14 outdoor urban HS scenes, each captured using a V-EOS HS camera by Photon. The images have a resolution of 1000 × 900pixels. There are 61 spectral bands between 400 nm and 1000 nm, with a spectral resolution of 10 nm. Similarly, no segmentation labels are provided.

This spectral image is used in various scenarios, particularly for the evaluation of machine learning models that are specialized in detecting anomalies, as referenced in (Nguyen et al., 2019).

Other datasets

Hyperview. The Hyperview dataset [dataset] (Nalepa et al., 2024), collected in March 2021, was obtained using hyperspectral imagers mounted on a Piper PA-31 Navajo aircraft. This dataset comprises 2886 hyperspectral patches of varying sizes, each with a spectral range from 400 to 2500 nm. The spectral resolution is 3.26 nm for the [400, 1000] range and 5.45 nm for the [930, 2500] range, resulting in 288 bands for SWIR and 186 for VNIR.

Given its high spectral resolution, similar to the Pavia University dataset which has a spectral resolution of 4 nm, the Hyperview dataset can serve as a reference for interpolation processes. However, this study will primarily focus on the Pavia University dataset, while acknowledging that the Hyperview dataset holds potential for future research and improvements in the proposed methodology.

CloudSEN12. The CloudSEN12 dataset [dataset] (Aybar et al., 2022) is obtained from Sentinel-2 multispectral imagery, including 13 spectral bands within the [443, 2190] nm range. It consists of 49,400 image patches and is extensively used for cloud and cloud shadow detection tasks. The dataset is divided into four classes (Clear, Thick Cloud, Thin Cloud, and Cloud Shadow) and is primarily employed for training and testing cloud detection algorithms. Given this, we opted not to include the CloudSEN12 dataset in our study, as its primary focus is on cloud detection rather than hyperspectral image classification in vegetation and non-vegetation areas. The presence of clouds in the images would introduce noise into the classification process, and since our neural networks were not trained to detect clouds, the results would not be accurate. However, we consider improving our research in the future by incorporating multiple datasets in the training process for the neural networks and therefore including CloudSEN12 in the testing phase.

Onera satellite change detection dataset (OSCD). The Onera Satellite Change Detection Dataset (OSCD) [dataset] (Daudt, 2018) is specifically designed for tasks involving change detection in satellite imagery. It includes 24 pairs of multispectral images captured by the Sentinel-2 satellites from various global locations between 2015 and 2018. These images span 13 spectral bands within the wavelength range of 443 to 2190 nm. The primary application of this dataset is for the training and testing of change detection algorithms. However, to enhance our study and demonstrate the extensibility of our methodology, we incorporated two specific images from the dataset, one from Norcia and the other from Chongqing. These selected images are depicted in Figure 5.

Interpolation approach

In practice, the spectral data aggregation concept is associated with generating integrated data from multiple MS/HS sources (Pal et al., 2019). However, directly merging different datasets into a single one may not produce a usable dataset for a machine-learning task. Traditional neural networks require the input data to be in a consistent format, with a fixed number of channels, but gathering a substantial amount of such data for spectral images can be challenging. The

(a) (b)

Figure 5. OSCD images: (a) Norcia; (b) Chongqing.

proposed approach aims to show that a preprocessing step applied to the samples in the datasets, which will then be used in the neural network, can address this issue. This step intends to transform the data to a predetermined format with a predetermined number of channels and wavelengths considered, based on the available datasets using the concept of interpolation.

Different interpolation techniques are used in the area of digital image processing (Fadnavis, 2014) because because they can improve the quality of the output in different scenarios. Interpolation can be used to determine the unknown value of a sample from known values present in a dataset. Moreover, using interpolation for HS images may not only generate values for different channels but may also increase the quantity of the information contained in each pixel.

We needed to establish a reference for the interpolation to properly aggregate or fuse the selected spectral datasets. Most of the chosen datasets have a spectral resolution of 10 nm, except for the Pavia University scene, which has a spectral resolution of 4 nm. As this image, has the best resolution in terms of the wavelength spectrum, it is the most appropriate as a reference for the interpolation. In this way, by increasing the number of wavelengths considered for each image, we expect to improve the quantity of the information contained in each pixel. Additionally, given that the maximum wavelength for the CAVE dataset is 700 nm, we decided to further limit the wavelengths to a maximum of 690 nm. This decision was made because interpolating values for the interval [700, 860] nm for images in this dataset would have produced unrealistic results.

Considering Pavia University as the reference dataset, for each wavelength of any other image, the value of each Pavia University channel is interpolated from the values of the adjacent channels of the processed image. We used four different interpolation methods on the five selected images and compared their results. The interpolation techniques used are classical methods, well-known in the literature for applications in different domains, including image processing, and will be presented subsequently. Although the selected methods may not be custom ones, applying these techniques in this study's novel domain allows us to effectively aggregate spectral images.

Linear interpolation

Linear interpolation is a mathematical method that generates new data points within the boundaries of a set of known discrete data using linear functions. The formula used to calculate the interpolated value y for a new point x is given by:

$$y = y_1 + (x - x_1)\frac{(y_2 - y_1)}{(x_2 - x_1)} \tag{1}$$

where (x_1, y_1) and (x_2, y_2), are the known data points.

This method has been used since antiquity for finding missing values in tables, and it is successfully used in various domains such as medical science (Kumar & Verma, 2022), simulations (Chunlin et al., 2014), electronics (Yang et al., 2015) and image processing (Zhu et al., 2015).

Quadratic interpolation

Quadratic interpolation assumes that the points follow a parabolic curve modeled by a quadratic equation with the following general form: $y = ax^2 + bx + c$. Having three known data points (x_0, y_0), (x_1, y_1) and (x_2, y_2), the following quadratic equation is used to estimate the value of a new point:

$$y = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x) \tag{2}$$

where $L_0(x), L_1(x), L_2(x)$ are the Lagrange basis polynomials.

This method is widely used in data analysis and curve fitting, but it also has applicability in the image resampling field (Dodgson, 1997).

Cubic spline interpolation

Cubic spline interpolation, a more complex procedure, is a mathematical method used to create a curve that connects data points with a degree of three, using piecewise third-degree polynomials. It is a special case for Spline interpolation, which successfully avoids Runge's phenomenon (Jiabo & Yanxiang, 2019).

Given $n+1$ data points (x_i, y_i) for $i = \overline{0,n}$, the cubic spline interpolation creates a set of cubic polynomials $S(x) = \{C_i(x), x_{i-1} \le x \le x_i | i \in \overline{0,n}\}$ where

$$C_i(x) = a_i + b_i x + c_i x^2 + d_i x^3, d_i \neq 0 \qquad (3)$$

is a cubic function defined on the interval $[x_{i-1}, x_i]$, $i = \overline{1,n}$. The coefficients for the cubic spline $S(x)$ a_i, b_i, c_i, d_i are calculated for each interval $[x_i, x_{i+1}]$ by solving the set of equations:

$$\begin{array}{ll} C_i(x_{i-1}) = y_{i-1}, & i = \overline{1,n} \\ C_i(x_i) = y_i, & i = \overline{1,n} \\ C_i^{'}(x_i) = C_{i+1}^{'}(x_i), & i = \overline{1,n-1} \\ C_i^{''}(x_i) = C_{i+1}^{''}(x_i), & i = \overline{1,n-1} \end{array} \qquad (4)$$

This method is commonly employed in computer graphics, animations, robots (Ogniewski, 2019), and image magnifying (Chu et al., 2011).

Piecewise cubic hermite interpolating polynomial

Piecewise cubic Hermite interpolating polynomial (PCHIP), is a piecewise polynomial function with a degree of three. It maintains the original shape of the data and, unlike cubic splines, aims to match only the first-order derivatives at the data points with those of the preceding and following intervals (Rabbath & Corriveau, 2019). We used the scientific computing Python library SciPy (https://scipy.org/) for the implementation.

Results and discussions

We used two main approaches to validate the proposed method and compare the results of the interpolations. The first one is visual inspection and some quantitative measures, and the second one is analyzing the impact of the methodology on the accuracy of the segmentation using neural networks. All tests are available in the source code repository, which can be accessed at https://github.com/RobertaLuca/spectral-data-aggregation.git.

Quality assessment by different measurement

The first way of assessing the outcome of the interpolation methods of the spectral pixels was done visually by 2D and 3D comparative plots and is described below.

Plots

The interpolation results for an image were compared both with the reference and among themselves. This was done by the generation of two distinct types of graphical representations: 2D plots of given pixels and 3D surface plots of the entire image. In this way, we can identify which interpolation result models better the shape of the original sample. Moreover, it provides early indications if certain interpolation types are unsuitable for a specific dataset. However, this information alone cannot validate or invalidate the interpolation because its goal is not only to preserve the initial information but also to augment the quantity of information contained.

For the 2D plots, we selected a random pixel and represented its spectral signature graphically. Each plot contains the pixel from the original image as well as all the interpolated versions. Results on pixels from each dataset are illustrated in Figure 6 (CAVE), 7 (UGR), 8 (Indian Pines),9 (KSC), 10 (Botswana) 11 (OSCD Norcia), and 12 (OSCD Chongqing). The coordinates of the pixel in the interpolated images are the same as in the original one, as the interpolation was performed only over the spectral channels.

The 3D plots offer a comprehensive visualization of the pixel surface distribution across the entire image, including multiple spectral bands. By leveraging the 3D plotting capabilities of the Matplotlib Python library, we stacked multiple semi-transparent surfaces (one for each spectral band) along the Z-axis to highlight the variation of intensity values across the spectral bands. Similar to the 2D plots, the images include both the reference dataset and the results obtained from various interpolation methods. The primary objective of displaying both the reference and interpolated images is to demonstrate that the overall visualization remains consistent, thereby confirming that the interpolation process does not alter the original image. Some plots can be seen in Figure 13 (one image of CAVE dataset), 14 (one image of UGR dataset), 15 (Indian Pines), 16 (KSC), 17 (Botswana), 18 (OSCD Norcia), and 19 (OSCD Chongqing).

By analyzing the graphical representations, we can affirm that all selected interpolation techniques provide satisfactory results. They predominantly follow the original shape of the pixel, but there are also different changes in intensities. The tests performed on the neural networks will provide a better perception, enabling us to estimate the impact of these alterations.

Surface differences

There are a lot of well-known metrics in the academic literature used for assessing the quality of HS

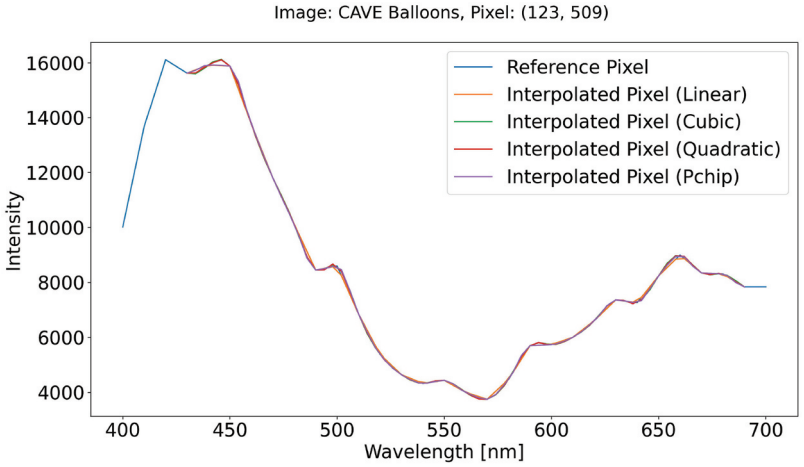


Figure 6. Reference and interpolated pixel for CAVE balloons.

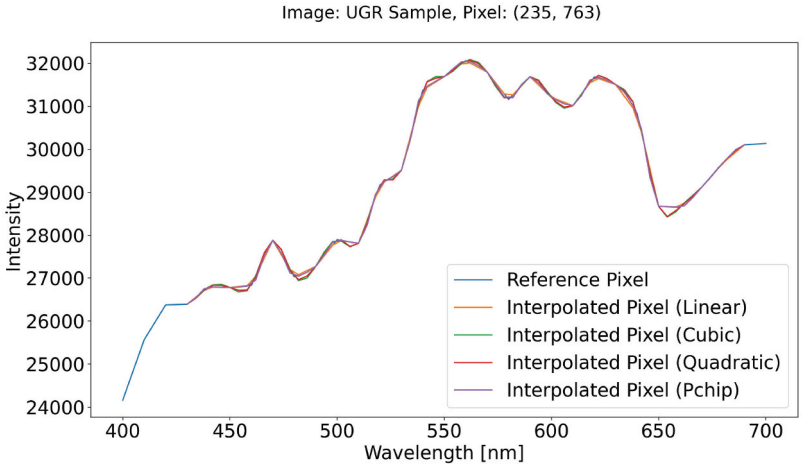


Figure 7. Reference and interpolated pixel for UGR.

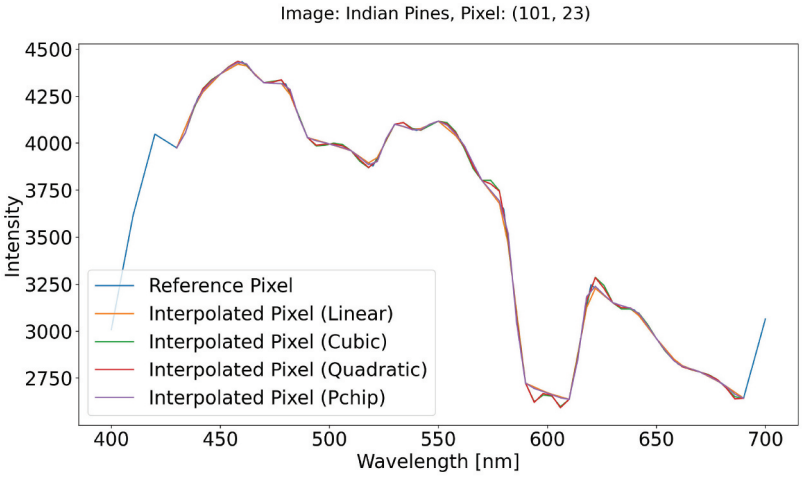


Figure 8. Reference and interpolated pixel for Indian Pines.

images. These include but are not limited to, Spectral Discrepancy, Correlation Coefficient, Mean Square Error (MSE), Root Mean Square Error (RMSE), Spectral Angle Mapper (SAM), Standard Deviation, Peak Signal-to-Noise Ratio (PSNR), and Image Fidelity (Christophe et al., 2004; Gavade & Rajpurohit, 2016). Nevertheless, these metrics are extremely valuable only if the compared images have the same number of spectral bands.

In the context of this study, we increased the number of spectral bands, starting with a spectral resolution of 10

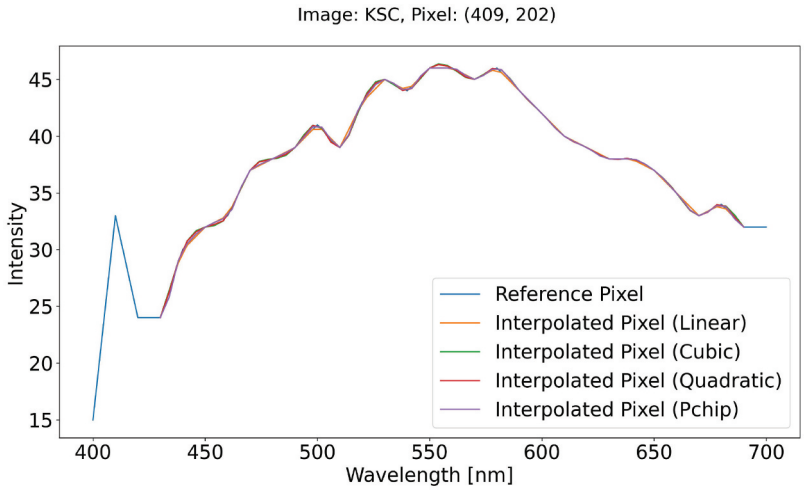


Figure 9. Reference and interpolated pixel for KSC.

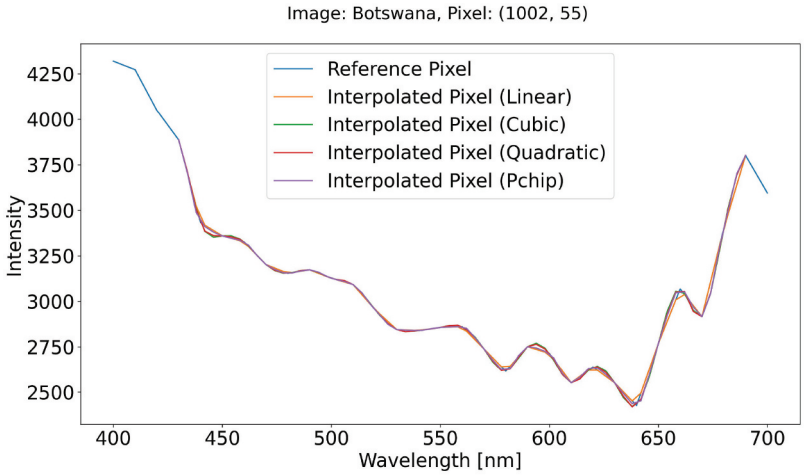


Figure 10. Reference and interpolated pixel for Botswana.

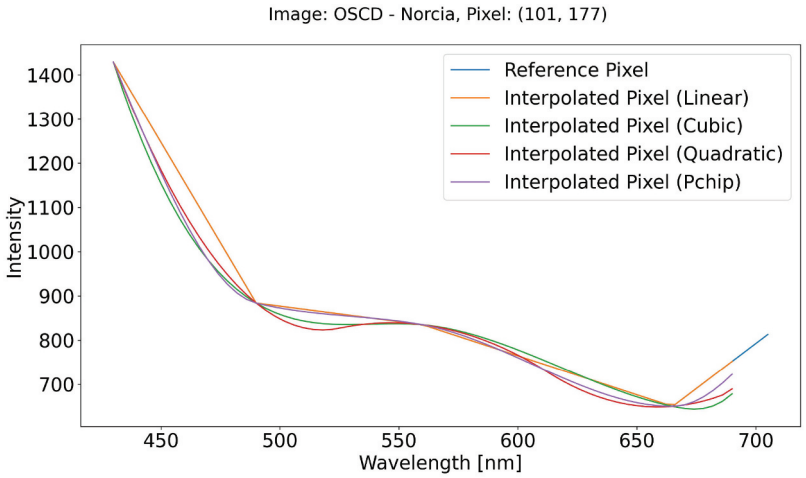


Figure 11. Reference and interpolated pixel for OSCD Norcia.

nm and interpolating the intensities for the missing wavelengths to have a spectral resolution of 4 nm. Given these circumstances, we opted to use as one of the validation methods the results provided by computing the surfaces under the 2D plots using the Trapezoidal Rule (Fornberg, 2021) and comparing the average differences between the reference and the interpolations. This approach offers a more precise comparison than the visual one offered by the plots, as it enables a quantitative assessment. This criterion allows us to

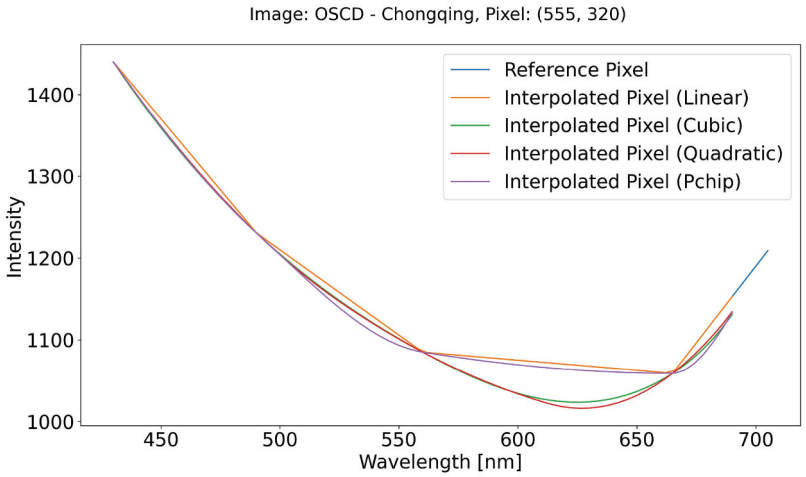


Figure 12. Reference and interpolated pixel for OSCD Chongqing.

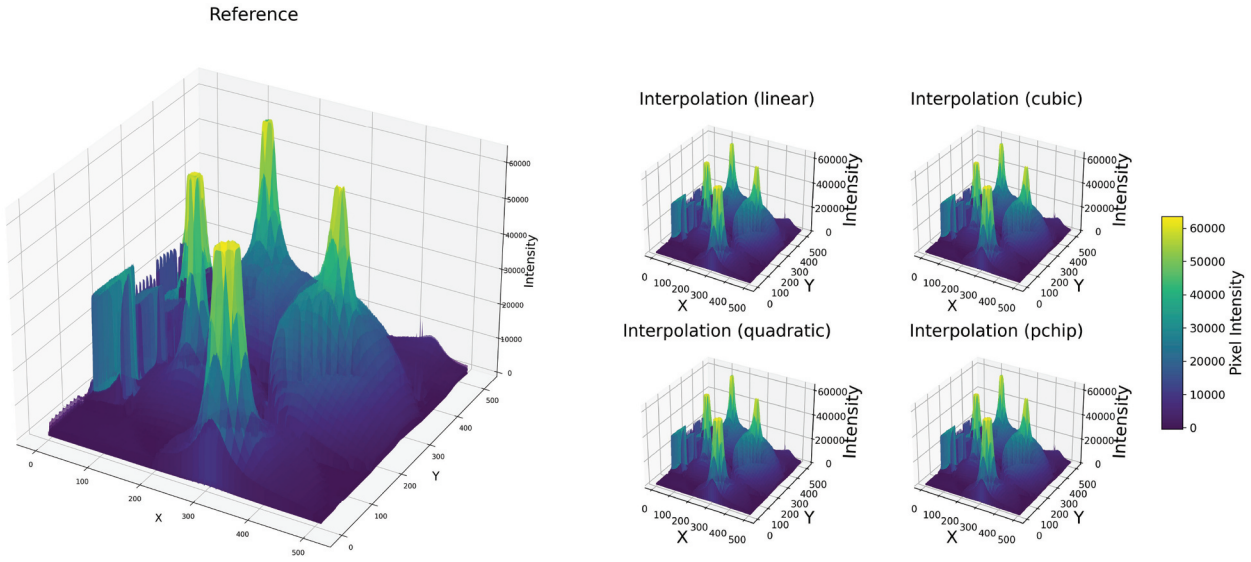


Figure 13. Pixels surface for CAVE balloons.

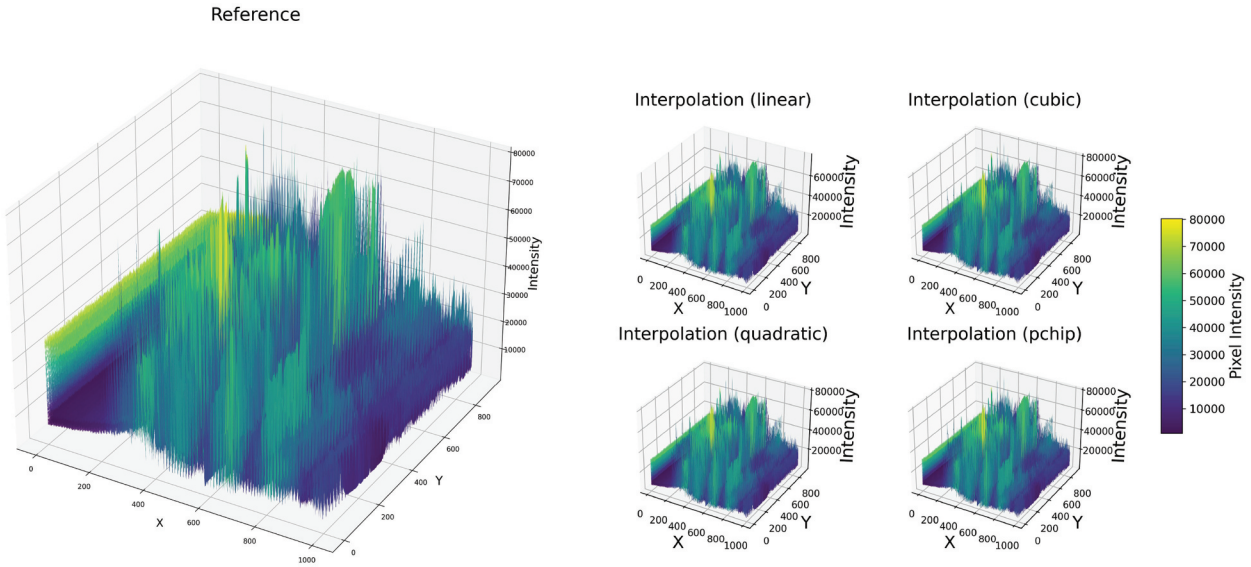


Figure 14. Pixels surface for UGR.

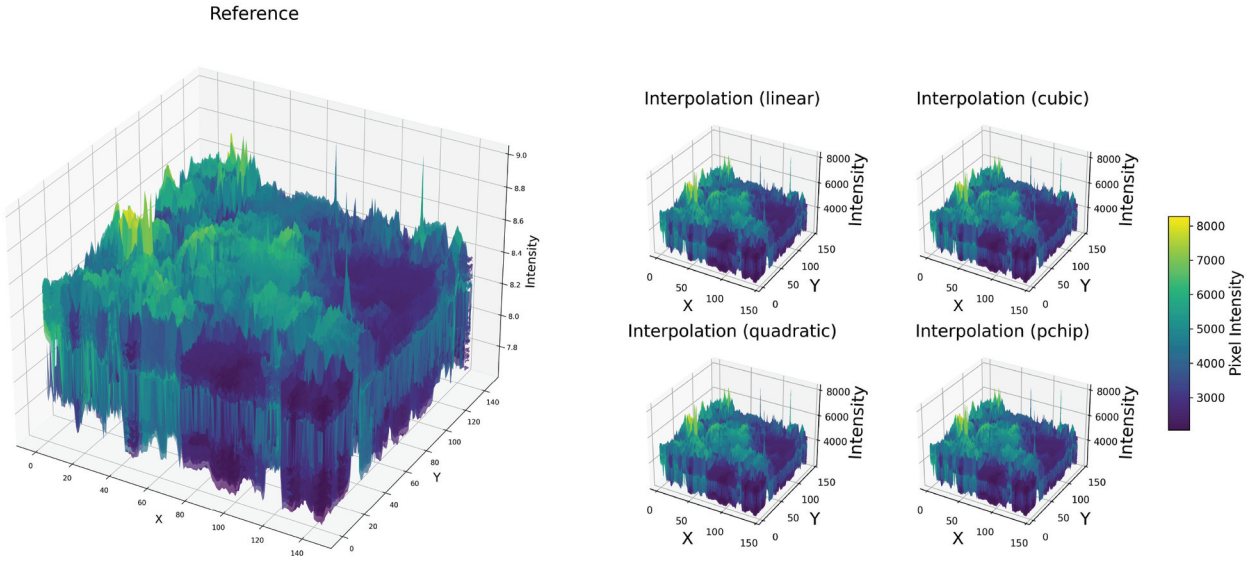


Figure 15. Pixels surface for Indian Pines.

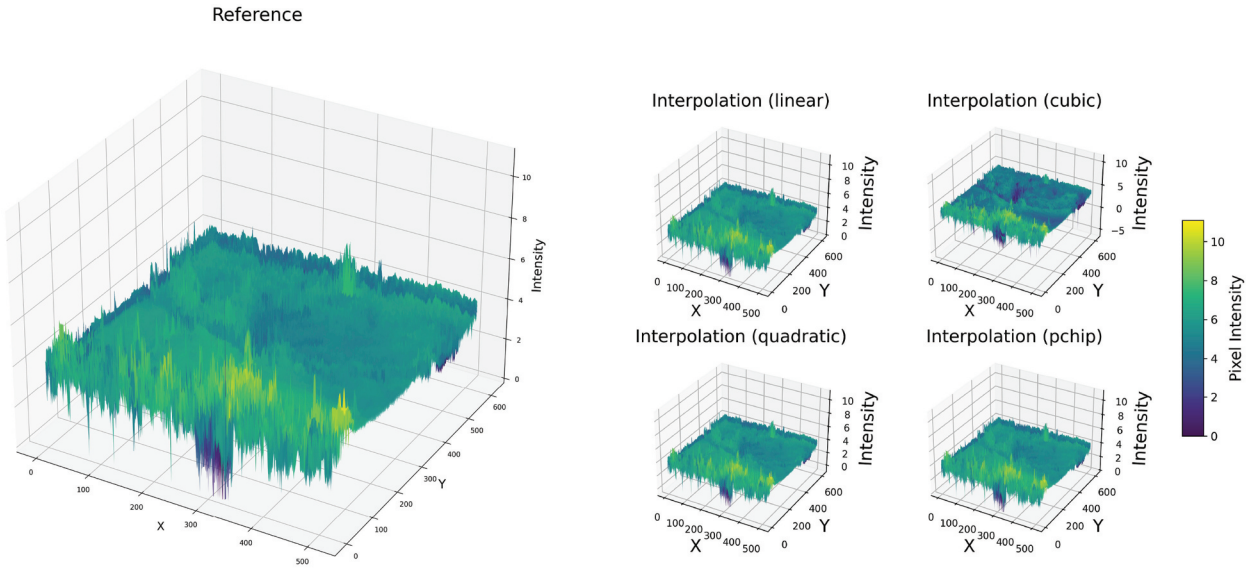


Figure 16. Pixels surface for KSC.

identify if specific methods are not appropriate to be used on different datasets, as indicated by exceptionally high or low values compared to the rest of the results. Table 1 presents a summary of the obtained results for these surfaces.

The results for the interpolations across all datasets have a high degree of similarity, with no prominent outliers, suggesting that all proposed methods are valid. Nevertheless, depending on the dataset, we can choose the most appropriate interpolation technique depending on the task that we further want to solve.

Custom mean squared error metric

In order to consistently evaluate and compare the different interpolation methods, we searched for an appropriate metric. Except for the plot-based measures, we could not find in literature such a measure, and therefore, we tried to adapt one of the most used metrics for the evaluation of results, i.e. the Mean Squared Error (MSE). The MSE usually compares two signals/functions with the same data points, yielding the degree of similarity between them. In our case, it could compare two HS pixels that have the same spectral signature. Alternatively, MSE could be used in the case of interpolation if a ground truth exists to compare with, which, in our case, does not exist.

In the case of interpolating HS images, we want to compare an interpolated pixel with the original one and evaluate how well the resulting interpolated HS pixel respects the shape of the original one; the MSE formula cannot be directly used. If we assume that the interpolated pixel faithfully replicates the shape of the original one, then interpolating back on the source wavelengths should produce a result very similar to the source pixel.

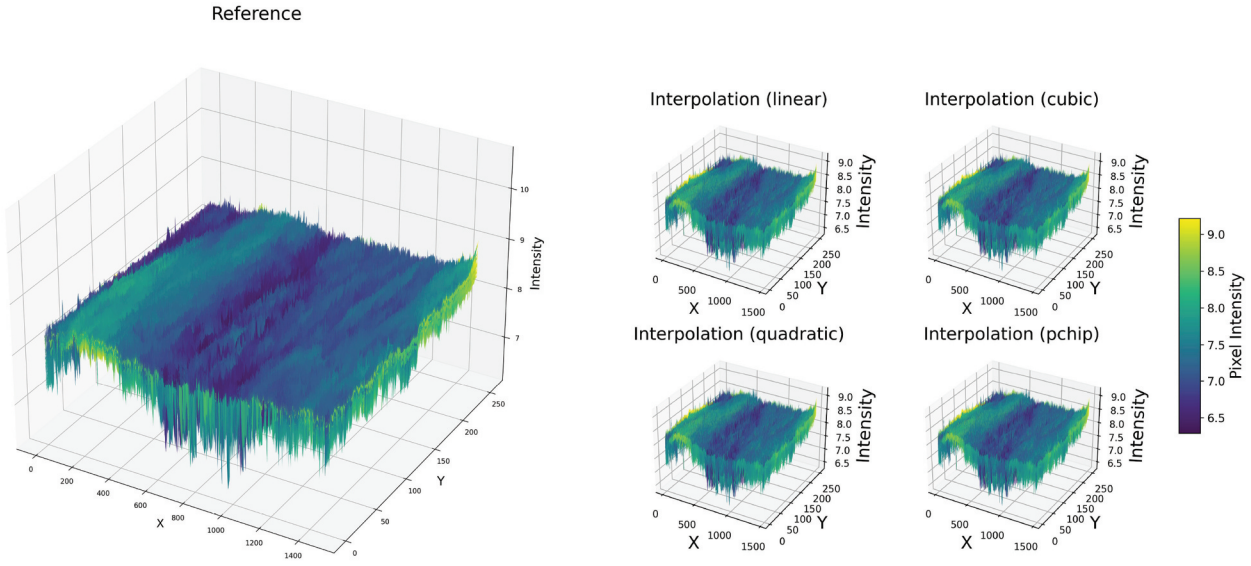


Figure 17. Pixels surface for Botswana.

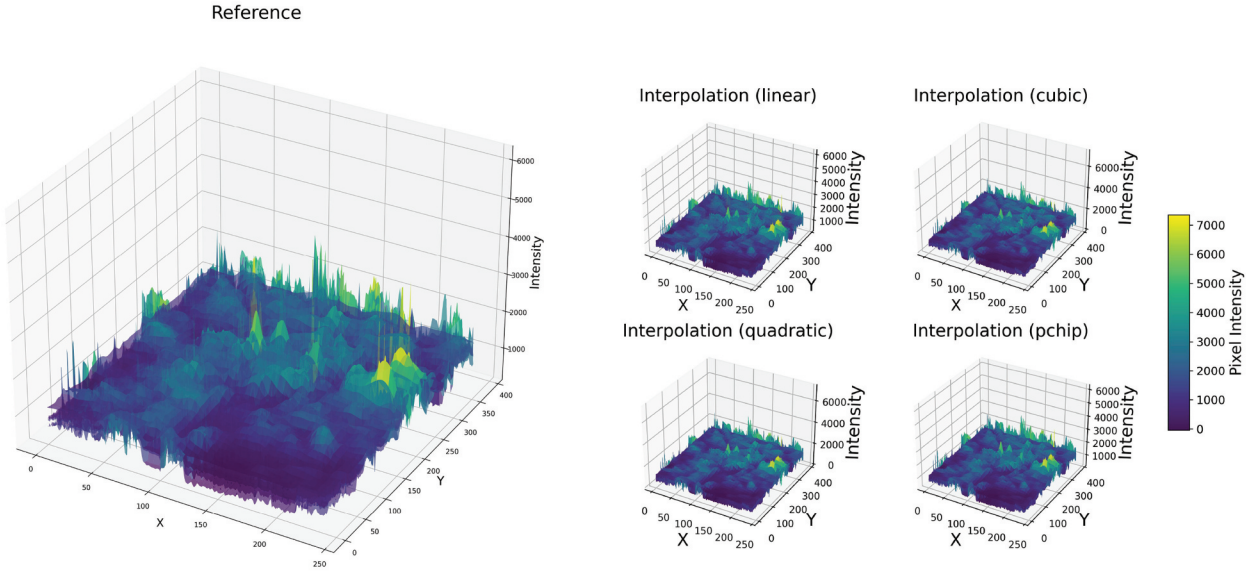


Figure 18. Pixels surface for OSCD Norcia.

As a consequence, we elaborated a Custom Mean Squared Error (CMSE) measure, which, for each MS pixel, calculates the MSE between this pixel and the result of its interpolation onto the desired wavelengths and then back onto the original ones. We could express this formula by:

$$CMSE(p) = \frac{(p - I_b(I_f(p))^2}{nw} \tag{5}$$

in which $I_f(p)$ represents the *forward interpolation*, i.e. the interpolation of pixel p on the new wavelengths and $I_b(p)$ represents the *backward* interpolation of p, i.e. the interpolation onto the source wavelength, and nw the number of wavelengths.

The results of this measure are represented in Table 2. It can be observed that, for the linear interpolation, the CMSE is very large, while for the quadratic and cubic, it is relatively small. In our opinion, this is an expected result, as linear interpolation is usually the most rough approximation for a signal. This observation underlines the results observed in the plot presented in Figure 20, which presents a magnified section from the 2D plot of a pixel of the KSC dataset.

Normalized difference vegetation index

In the context of spectral imaging processing, the Normalized Difference Vegetation Index (NDVI) is a pivotal indicator for precision agriculture, providing farmers with crucial data on crop health and resource allocation (Plajer, Băicoianu, Majercsik, & Ivanovici, 2023). The red (RED) and near-infrared (NIR) spectral bands of the spectral image are used to calculate

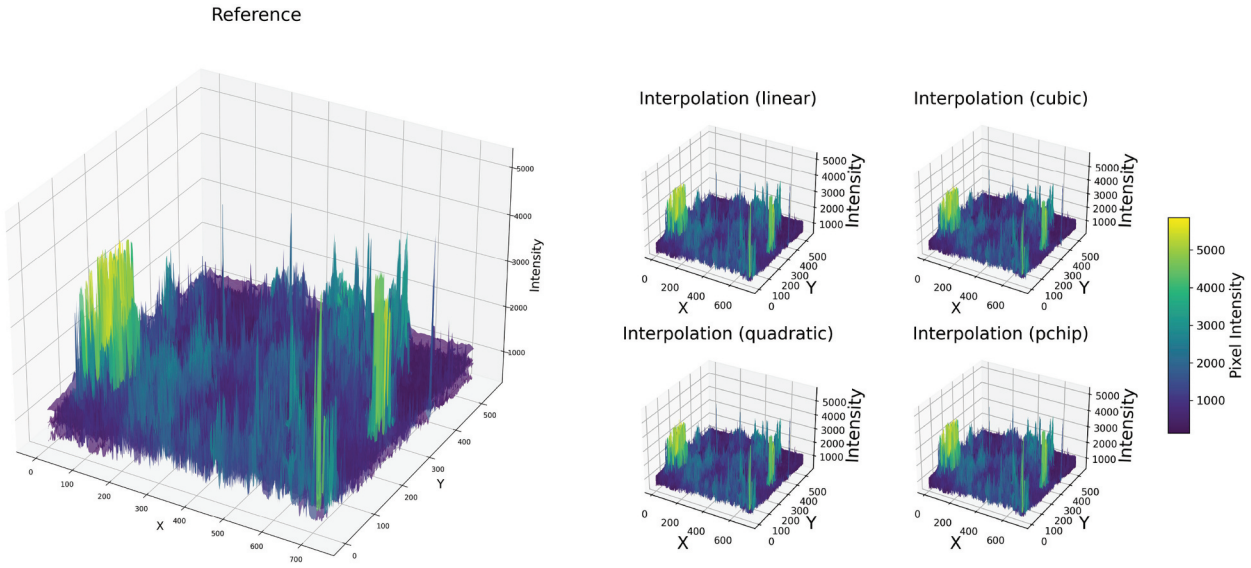


Figure 19. Pixels surface for OSCD Chongqing.

Table 1. Surface average differences.

Dataset	Interpolation	Surface Average Difference
CAVE Balloons	Linear	3.60
	Cubic	3.56
	Quadratic	3.60
	PCHIP	3.56
UGR	Linear	3.90
	Cubic	3.89
	Quadratic	3.88
	PCHIP	3.90
Indian Pines	Linear	1.89
	Cubic	1.88
	Quadratic	1.95
	PCHIP	2.13
KSC	Linear	6.01
	Cubic	7.32
	Quadratic	6.53
	PCHIP	5.54
Botswana	Linear	5.90
	Cubic	5.80
	Quadratic	5.79
	PCHIP	5.87
OSCD Norcia	Linear	74.54
	Cubic	79.26
	Quadratic	77.14
	PCHIP	76.70
OSCD Chongqing	Linear	137.46
	Cubic	134.77
	Quadratic	135.39
	PCHIP	138.79

Table 2. CMSE: reference - interpolated interpolation.

Dataset	Interpolation	MSE
CAVE Balloons	Linear	3675.01
	Cubic	7.75
	Quadratic	5.27
	PCHIP	341.31
UGR	Linear	7800.99
	Cubic	23.12
	Quadratic	15.08
	PCHIP	897.71
Indian Pines	Linear	554.50
	Cubic	1.65
	Quadratic	1.09
	PCHIP	78.59
KSC	Linear	96.17
	Cubic	0.58
	Quadratic	0.37
	PCHIP	8.76
Botswana	Linear	311.14
	Cubic	0.84
	Quadratic	0.57
	PCHIP	35.75
OSCD Norcia	Linear	24.48
	Cubic	0.000035
	Quadratic	≈ 0
	PCHIP	0.04
OSCD Chongqing	Linear	13.83
	Cubic	0.000012
	Quadratic	≈ 0
	PCHIP	0.02

the NDVI, which is a vegetation index. The general formula is given by Eq. 6.

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{6}$$

The NDVI values range within the interval $[-1, 1]$. They can be interpreted as follows:

- $(0, 0.33]$ for bare soil with little to no vegetation cover;
- $(0.33, 0.66]$ for unhealthy or sparse vegetation;
- $(0.66, 1]$ for dense and healthy vegetation;
- $[-1, 0]$ for water or inanimate things.

Our objective was to investigate the influence of interpolation on the calculation of the NDVI. From the conducted tests, as shown in Table 3, all the methods appear equally suitable. This finding is significant as it suggests that the most straightforward and quickest method, linear interpolation, can be effectively employed in this context.

Quality assessment by the accuracy of the semantic segmentation

In spectral imaging, semantic segmentation plays a crucial role by enabling pixel-wise labeling of the image. In this context increasing the number

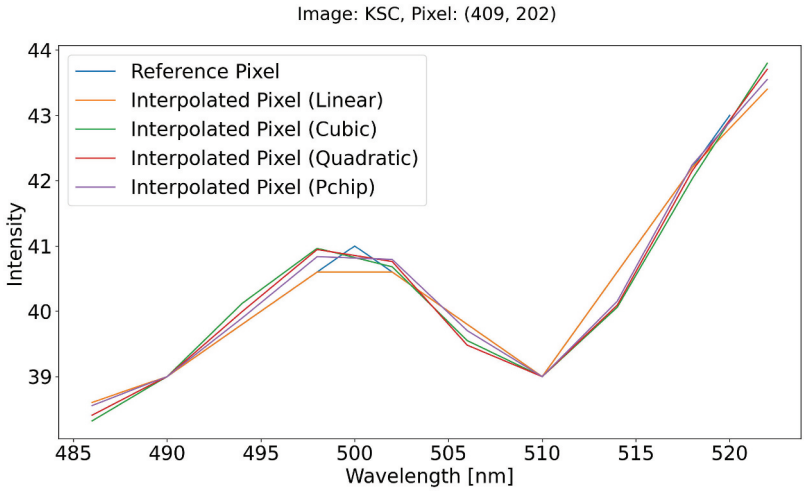


Figure 20. Magnified plot for a selected pixel from the KSC dataset.

Table 3. MSE: NDVI reference – NDVI interpolation.

Dataset	Interpolation	MSE for NDVI
Indian Pines	Linear	0.00014
	Cubic	0.00024
	Quadratic	0.00020
	PCHIP	0.00019
KSC	Linear	0.00024
	Cubic	0.00042
	Quadratic	0.00038
	PCHIP	0.00038
Botswana	Linear	0.00075
	Cubic	0.00155
	Quadratic	0.00139
	PCHIP	0.00117
OSCD Norcia	Linear	0.000033
	Cubic	0.00040
	Quadratic	0.00039
	PCHIP	0.00002
OSCD Chongqing	Linear	0.000024
	Cubic	0.00095
	Quadratic	0.00093
	PCHIP	0.000049

of samples of a dataset could be beneficial for the training of neural network models. This goal could be achieved by using interpolation for fusing several datasets, as proposed in section 2. In this subsection, we evaluate and compare the proposed aggregation method in the context of training neural networks for semantic segmentation of HS images.

For this purpose, two different neural networks were implemented. The first one is a simple Fully Connected Neural Network (FCNN), while the second architecture, U-Net, a Convolutional Neural Network (CNN), is known for providing good accuracy for remote sensing image semantic segmentation tasks.

In practical applications, FCNNs are used in a wide range of domains including object recognition. On the other hand, U-Net architectures provide robust solutions for a variety of tasks outperforming traditional methods and even human experts in certain cases including semantic segmentation tasks in remote sensing areas (Bao et al., 2021; Liu & Zhao, 2023; Xiang et al., 2022) and precision agriculture tasks (Wang et al., 2023), among many others (Hu et al., 2021; Sun et al., 2022; Suri et al., 2023).

Both architectures were trained to classify spectral pixels into two categories: vegetation and non-vegetation. The choice of these two classes was guided by two key considerations. First, our primary focus is on the application of satellite spectral imaging in agriculture, which necessitates an emphasis on vegetation. Second, the diverse class definitions in the datasets under consideration required unification into a set of common classes. Given the significant differences between the original class definitions, the only viable approach was to consolidate them into the chosen categories.

Furthermore, we want to emphasize, that our study does not focus on the best architecture for pixel classification into vegetation/non-vegetation, but on demonstrating the viability of the interpolation approach for spectral data aggregation for segmentation purposes. This section concentrates thus on this aspect, by comparing the results obtained by the two selected architectures with and without integrating different spectral data into one dataset.

Fully connected neural network

Assuming that the input data was preprocessed and the number of wavelengths present in all the images that will be used is 66, the input layer of the neural network will correspondingly contain 66 neurons. The network architecture includes four fully connected hidden layers and an output layer with two neurons for the two classification categories: vegetation and non-vegetation. To leverage the computational efficiency associated with powers of 2, the hidden layers contain 128, 256, 512, and 256 neurons, respectively. The Rectified Linear Unit (ReLU) is the activation function used for the hidden layers, while the Softmax function is used for the output layer to

convert the output into probabilities. Figure 21 provides a schematic representation of the model architecture.

The network was trained on the Pavia University image and various other dataset combinations that are detailed in Section 3.2.3. The pixels were combined into a single set, which was then shuffled and divided into training and testing subsets. After several tests, it was determined that 150 epochs were sufficient for the classification task. We used the Cross-Entropy Loss as a loss function, and we selected Adam as the optimizer, with a learning rate of 0.0001. After shuffling all the pixels, the dataset was split into training (80%) and testing (20%) subsets and random batches of 2048 pixels were used for training.

U-Net

An U-Net model is characterized by its U-shape, which contains a contracting path, or encoder, to capture the context and a symmetric expanding path, or decoder, for precise localization. In the implemented architecture, both the encoder and decoder consist of five 3D convolutional layers, followed by a Batch Normalization layer and a ReLU activation function. The last layer in the decoder path is a convolutional layer with the number of filters equal to the number of classes, which is two, and it uses a Softmax activation function. The output of the encoder's last layer and the decoder's last layer are concatenated, facilitating better localization because the high-resolution features from the encoder are combined with the upsampled output of the decoder. We used 3D convolutional layers and image patches of 10×10pixels with 66 channels for each pixel, setting the input shape to (10, 10, 66, 1). The filters were selected based on the computational benefits of operations with powers of two. Therefore, for the contracting path, they were set to 64, 64, 128, 256, and 256 respectively. We used 64, 128, 256, 512, and 256 for the expanding path, respectively.

The network was trained on 10×10 patches from Pavia University and various other combinations of patches from different datasets that are presented in detail in Subsection 3.2.3. The patches combined from all the datasets were split into training (80%), validation (10%), and testing (10%) subsets. Depending on the selected datasets, the number of epochs varies because the training is stopped when a monitored metric, the loss in our case, has stopped improving. As an optimizer, we used Adam with a learning rate of 0.001.

To expand the dataset used in the training process, we used the data augmentation technique (Dewantara et al., 2020). For each patch in the training dataset, we generated three more images using a horizontal flip, a vertical flip, and a rotation of the patch with a random angle between −180 and 180 degrees.

Training datasets

In order to train and test the results on these networks, four of the exposed datasets can be used, namely: Pavia University, KSC, Botswana, and Indian Pines, because they also contain a ground truth image with the associated labels. These labels were merged into only two main classes: Vegetation and Non-Vegetation.

We tested the neural networks on different training scenarios using both unprocessed datasets and combinations of original and interpolated sets. Table 4 presents the most representative results obtained for the semantic segmentation using a fully connected neural network, while Table 5 displays a part of the results for the U-Net architecture.

Moreover, Table 4 presents the results of the training process using both shuffled and non-shuffled pixels. The results indicate that in the non-shuffled scenario, overfitting occurred because the training accuracy was higher than in the shuffled case (exceeding 93% for all training scenarios), but the testing results were worse. While the dataset included in the training pixels showed high accuracy, the accuracy for new datasets was very low. This demonstrates that, for the neural network architecture and datasets used, overfitting is mitigated when the pixels are shuffled during the training process.

Since the OSCD dataset does not include labels for classification operations, we were unable to calculate the accuracy and other specific metrics necessary for evaluating the performance of the neural networks. Nonetheless, we proceeded with the classification task using the Fully Connected Neural Network

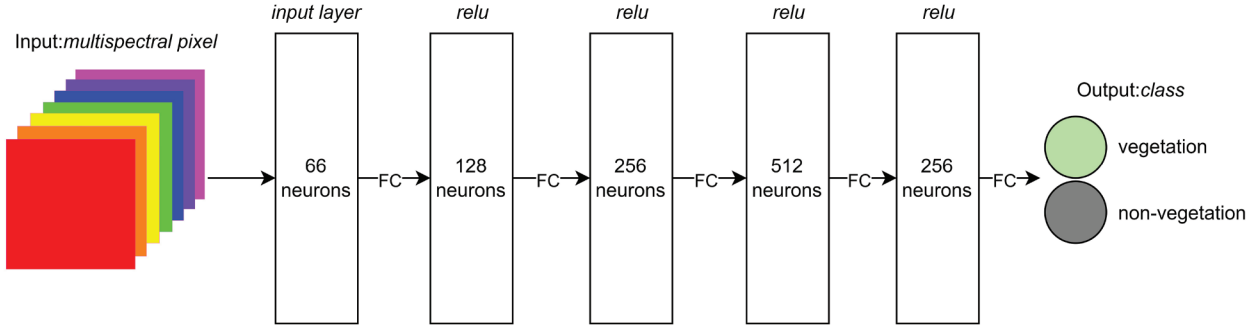


Figure 21. Network model.

Table 4. FCNN training and testing results.

Training Dataset	Training Accuracy (shuffled/non-shuffled)	Testing Dataset (Interpolation)	Testing Accuracy (shuffled/non-shuffled)
Pavia University	71.5/94.15%	KSC (Linear)	72.56/57.59%
		KSC (Cubic)	72.37/58.34%
		K SC (Quadratic)	72.37/58.49%
		KSC (PCHIP)	72.56/58.28%
		Botswana (Linear)	53.17/25.46%
		Botswana (Cubic)	53.48/25.46%
		Botswana (Quadratic)	53.17/25.46%
		Botswana (PCHIP)	53.17/25.46%
		Indian Pines (Linear)	90.75/7.27%
		Indian Pines (Cubic)	90.62/7.62%
		Indian Pines (Quadratic)	90.61/7.49%
		Indian Pines (PCHIP)	90.64/7.40%
Pavia University and Botswana (Linear)	88/93.38%	KSC (Linear)	72.56/60.74%
		KSC (Cubic)	72.40/61.14%
		KSC (Quadratic)	72.54/61.12%
		KSC (PCHIP)	72.56/60.81%
		Botswana (Linear)	90.02/97.04%
		Botswana (Cubic)	89.93/96.86%
		Botswana (Quadratic)	89.93/96.95%
		Botswana (PCHIP)	89.99/96.98%
		Indian Pines (Linear)	94.98/7.94%
		Indian Pines (Cubic)	94.84/8.10%
		Indian Pines (Quadratic)	94.89/8.13%
		Indian Pines (PCHIP)	94.97/8.10%
Pavia University and KSC (Cubic)	89.27/93.28%	KSC (Linear)	97.08/96.49%
		KSC (Cubic)	97.12/96.49%
		KSC (Quadratic)	97.12/96.49%
		KSC (PCHIP)	97.08/96.49%
		Botswana (Linear)	77.46/25.46%
		Botswana (Cubic)	77.77/25.46%
		Botswana (Quadratic)	77.83/25.46%
		Botswana (PCHIP)	78.14/25.46%
		Indian Pines (Linear)	92.97/4.68%
		Indian Pines (Cubic)	92.46/4.69%
		Indian Pines (Quadratic)	92.54/4.69%
		Indian Pines (PCHIP)	92.87/4.69%

Table 5. U-Net training and testing results.

Training Dataset	Training Accuracy	Testing Dataset (Interpolation)	Testing Accuracy
Pavia University (82 epochs)	89%	KSC (Linear)	72.08
		KSC (Cubic)	
		KSC (Quadratic)	
		KSC (PCHIP)	
		Botswana (Linear)	91.70
		Botswana (Cubic)	88.98
		Botswana (Quadratic)	90.52
		Botswana (PCHIP)	90.36
		Indian Pines (Linear)	69.41
		Indian Pines (Cubic)	71.19
		Indian Pines (Quadratic)	71.44
		Indian Pines (PCHIP)	71.13
Pavia University and Indian Pines (Linear) (95 epochs)	89%	KSC (Linear)	72.08
		KSC (Cubic)	
		KSC (Quadratic)	
		KSC (PCHIP)	
		Botswana (Linear)	86.11
		Botswana (Cubic)	86.10
		Botswana (Quadratic)	86.11
		Botswana (PCHIP)	86.11
		Indian Pines (Linear)	94.96
		Indian Pines (Cubic)	94.95
		Indian Pines (Quadratic)	94.96
		Indian Pines (PCHIP)	94.96
Pavia University and KSC (Cubic) (89 epochs)	86%	KSC (Linear)	72.08
		KSC (Cubic)	
		KSC (Quadratic)	
		KSC (PCHIP)	
		Botswana (Linear)	92.43
		Botswana (Cubic)	92.44
		Botswana (Quadratic)	92.43
		Botswana (PCHIP)	92.43
		Indian Pines (Linear)	80.04
		Indian Pines (Cubic)	78.79
		Indian Pines (Quadratic)	78.70
		Indian Pines (PCHIP)	78.83

(trained on Pavia University and Botswana linear interpolated) and generated the corresponding images. These images depict vegetation pixels in green and non-vegetation pixels in black. The results for the Norcia cubic interpolated and Chongqing quadratic interpolated images are presented in Figure 22.

We have also experimented with evaluating the FCNN using the whole Pavia University image, including all 103 spectral bands. This choice was chosen in order to take advantage of the extra data found in the supplementary spectral bands, which may be quite valuable in the case of spectral imaging. As a result, we carried out the interpolations this time, taking into account the wavelengths between 430 and 838 nm.

At first, the outcomes were less than encouraging, indicating a training accuracy of only 33%. This underperformance can be attributed to the specific architecture chosen for the neural network. Considering that we've nearly doubled the number of spectral bands, and consequently increased the neurons in the input layer, we recognized the potential benefit of also doubling the neurons in certain hidden layers. Our final tests were done on an architecture that includes four hidden layers with 256, 512, 256, and 128 neurons, respectively. In this setup, we observed noticeable improvements during the testing phase. We trained the neural network on the Pavia University dataset for 150 epochs and then tested the classification on the KSC, Botswana, and Indian Pines datasets. The results are presented in Table 6. We noticed significant progress with the Botswana dataset, where the accuracy nearly doubled compared to the previous results. For KSC, the results were roughly the same, while for Indian Pines they improved by 5%, resulting in a total of 95% correct classifications.

Given the diversity across the datasets, incorporating a wide array of both vegetation and non-vegetation associated pixels, we regard the outcomes as promising. Notably, we observed improvements in the outcomes when the training incorporated a mix of unprocessed pixels from the Pavia University dataset and results from the interpolations. Moreover, in all the cases, the results from the testing phase are better than the training ones for at least one of the three tested datasets. For the others, the accuracy is similar to the

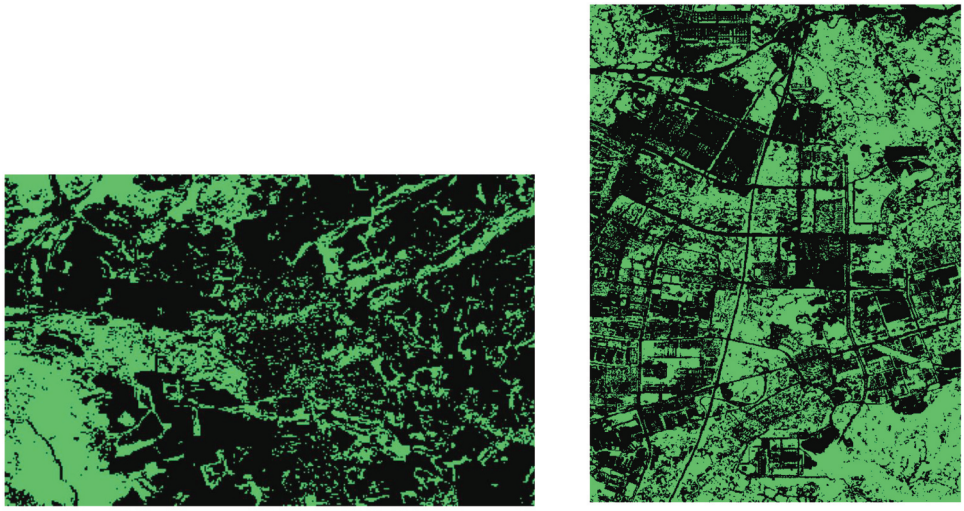


Figure 22. OSCD classification results: (a) Norcia (Cubic); (b) Chongqing (Quadratic).

training phase's. The lower results are caused by the differences in the dataset labels. Sometimes, the vegetation is associated with the color green, while in other cases, it is correlated with dry vegetation revealing shades of brown. In such scenarios, we have observed that combining pixels from two different images significantly improves the quality of the results.

Furthermore, to obtain a comprehensive view of the statistical results of our neural networks, dedicated machine learning metrics were employed. Figure 23 presents the confusion matrices for selected testing datasets for the FCNN in both shuffled and non-shuffled scenarios. These matrices corroborate the accuracy results, highlighting poorer performance in the non-shuffled scenario and indicating the occurrence of overfitting.

We conducted the same statistical analysis for the U-Net architecture, with the results for the confusion matrices available in Figure 24.

Additionally, for the same selected tests in terms of training pixels and testing datasets, Table 7. provides insights into the precision of the results along with the F1 Score values.

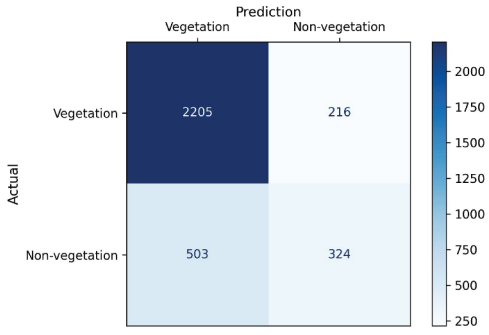
Conclusions

This paper illustrates a preprocessing approach for the spectral data aggregation task. After performing the interpolation of the spectral pixels with respect to a chosen reference wavelength spectrum and validating the results using three distinct methods, we concluded that the results obtained are encouraging and demonstrate the validity of the proposed solution.

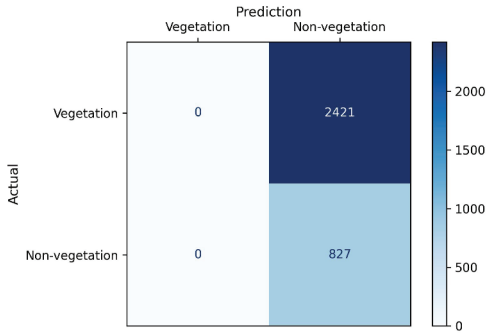
This study was developed on six different datasets. However, the selection of these datasets posed a serious challenge because of the variety of characteristics of the spectral images and the labels provided for them, making it hard to decide if they can or cannot be fused. This still remains a problem and represents a potential area for further enhancement of our methodology. Each HS dataset has a specific description that can provide useful information and insights for gathering a collection of sets that can be fused. Therefore, developing an algorithm that analyzes these descriptions and determines whether the data can be aggregated would be the next step in improving our solution. This would also allow us to test and interpret the results on larger datasets.

Table 6. FCNN training and testing results using 103 spectral bands.

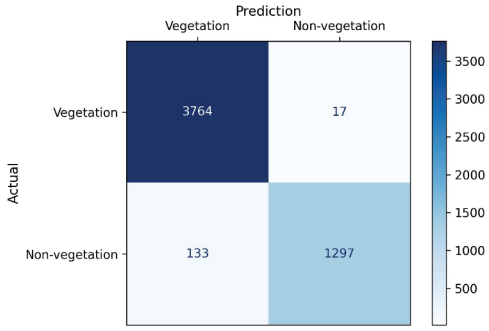
Training Dataset	Training Accuracy	Testing Dataset (Interpolation)	Testing Accuracy
Pavia University	66.26%	KSC (Linear)	72.56
		KSC (Cubic)	
		KSC (Quadratic)	
		KSC (PCHIP)	
		Botswana (Linear)	82.85
		Botswana (Cubic)	
		Botswana (Quadratic)	
		Botswana (PCHIP)	
		Indian Pines (Linear)	95.33
		Indian Pines (Cubic)	
		Indian Pines (Quadratic)	
		Indian Pines (PCHIP)	



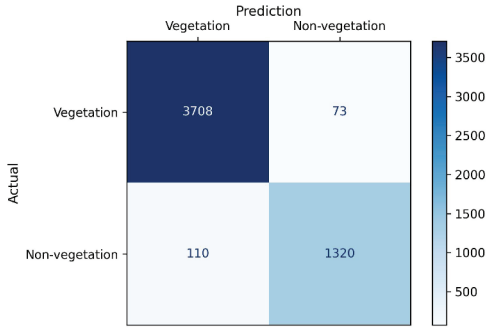
(a) Train: Pavia University and KSC Cubic (shuffle), Test: Botswana Quadratic



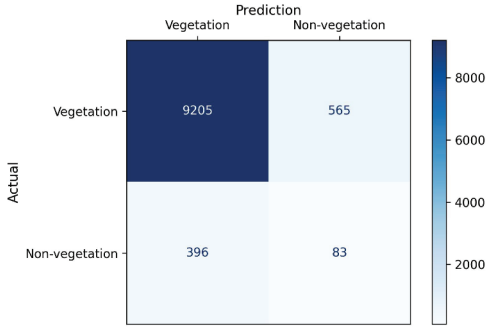
(b) Train: Pavia University and KSC Cubic (no shuffle), Test: Botswana Quadratic



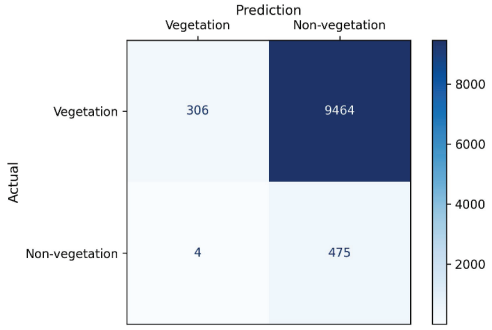
(c) Train: Pavia University and KSC Cubic (shuffle), Test: KSC Quadratic



(d) Train: Pavia University and KSC Cubic (no shuffle), Test: KSC Quadratic

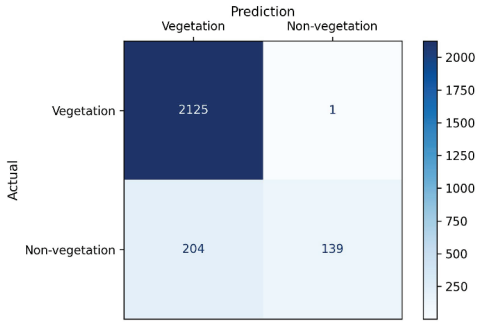


(e) Train: Pavia University (shuffle), Test: Indian Pines Cubic

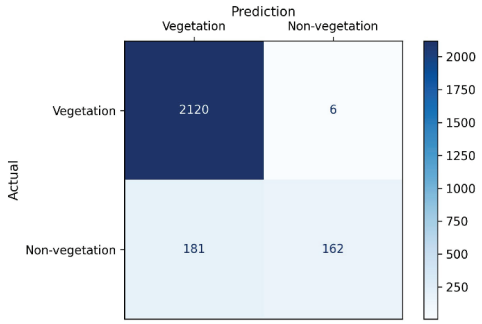


(f) Train: Pavia University (no shuffle), Test: Indian Pines Cubic

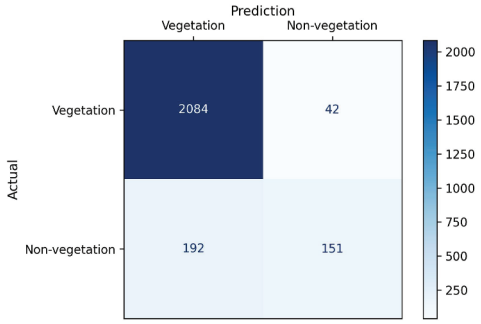
Figure 23. Confusion matrices for FCNN.



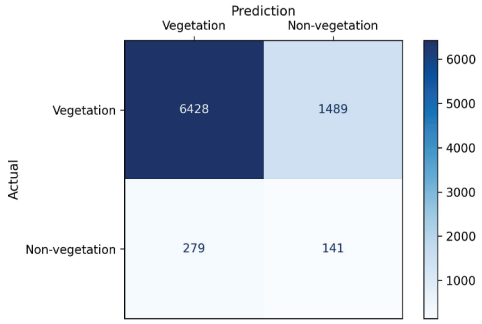
(a) Train: Pavia University, Test: Botswana Linear



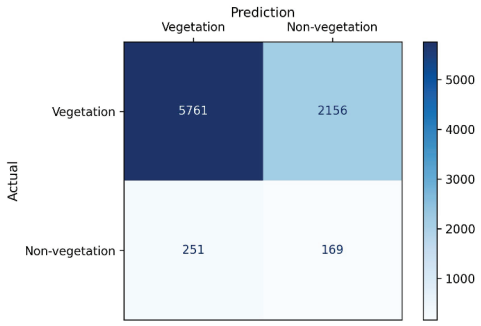
(b) Train: Pavia University and KSC Cubic, Test: Botswana Cubic



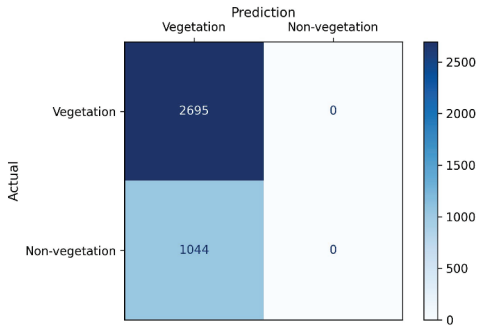
(c) Train: Pavia University, Test: Botswana Quadratic



(d) Train: Pavia University and KSC Cubic, Test: Indian Pines Cubic



(e) Train: Pavia University, Test: Indian Pines Pchip



(f) Train: Pavia University and Indian Pines Linear, Test: KSC Linear

Figure 24. Confusion matrices for UNet.

Table 7. Statistical results for the neural networks (Precision and F1 score).

Training Dataset	Testing Dataset (Interpolation)	Neural Network	Shuffle	Precision	F1 Score
Pavia University and KSC Cubic	Botswana (Quadratic)	FCNN	Yes	0.6	0.47
Pavia University and KSC Cubic	Botswana (Quadratic)	FCNN	No	0.26	0.40
Pavia University and KSC Cubic	KSC (Quadratic)	FCNN	Yes	0.98	0.94
Pavia University and KSC Cubic	KSC (Quadratic)	FCNN	No	0.94	0.93
Pavia University	Indian Pines (Cubic)	FCNN	Yes	0.12	0.14
Pavia University	Indian Pines (Cubic)	FCNN	No	0.04	0.09
Pavia University	Botswana (Linear)	U-Net	N/A	0.98	0.26
Pavia University and KSC Cubic	Botswana (Cubic)	U-Net	N/A	0.98	0.26
Pavia University	Botswana (Quadratic)	U-Net	N/A	0.96	0.25
Pavia University and KSC Cubic	Indian Pines (Cubic)	U-Net	N/A	0.70	0.56
Pavia University	Indian Pines (PCHIP)	U-Net	N/A	0.62	0.48
Pavia University and Indian Pines Linear	KSC (Linear)	U-Net	N/A	0.72	0.34

Nonetheless, the exposed results based on the four manually selected datasets, which were used in the context of validation using neural networks specialized in semantic segmentation, showed to be promising, even though the merged labels could not be as accurate as desired and the segmentation networks were relatively simple ones. Therefore, the proposed aggregation technique proves to be efficient, robust, and adaptable, showing potential to be used in different applications and to gain in accuracy, when provided with better labels or used with more sophisticated architectures.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Funded by the European Union. The AI4AGRI project entitled "Romanian Excellence Center on Artificial Intelligence on Earth Observation Data for Agriculture" received funding from the European Union's Horizon Europe research and innovation program under grant agreement no. [101079136]. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

References

Ahmad, M., Khan, A. M., Mazzara, M., Distefano, S., Ali, M., & Sarfraz, M. S. (2022). A fast and compact 3-d cnn for hyperspectral image classi_cation. *IEEE Geoscience & Remote Sensing Letters*, *19*, 1–5. https://doi.org/10.1109/LGRS.2020.3043710

Aybar, C., Ysuhuaylas, L., Loja, J., Gonzales, K., Herrera, F., Bautista, L., Yali, R., Flores, A., Diaz, L., Cuenca, N., & Espinoza, W. (2022). Cloudsen12, a global dataset for semantic understanding of cloud and cloud shadow in sentinel-2. *Scientific Data*, *9*(1). https://doi.org/10.1038/s41597-022-01878-2

Bao, Y., Liu, W., Gao, O., Lin, Z., & Hu, Q. (2021). E-unet++: A semantic segmentation method for remote sensing images. *2021 IEEE 4th Imcec*, *4*, 1858–1862 doi:10.1109/IMCEC51613.2021.9482266.

Benhalouche, F. Z., Benabbou, O., Karoui, M. S., Kebir, L. W., Bennia, A., & Deville, Y. (2022). Minerals detection and mapping in the southwestern Algeria garadjebilet region with a multistage informed nmf-based unmixing approach using prisma remote sensing hyperspectral data. In *IGARSS 2022* (pp. 6422–6425). Institute of Electrical and Electronics Engineers.

Benhida, H., Souadi, M., & El Ansari, M. (2022). Convolutional neural network for automated colorectal polyp semantic segmentation on colonoscopy frames. In *2022 9th international conference on wireless networks and mobile communications (WINCOM)* (pp. 1–5). Institute of Electrical and Electronics Engineers.

Brezini, S. E., & Deville, Y. (2023). Hyperspectral and multispectral image fusion with automated extraction of image-based endmember bundles and sparsity-based unmixing to deal with spectral variability. *Sensors (Switzerland)*, *23*(4), 2341. https://doi.org/10.3390/s23042341

Chen, Y., Lin, Z., Zhao, X., Wang, G., & Gu, Y. (2014, 06). Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, *7*(6), 2094–2107. https://doi.org/10.1109/JSTARS.2014.2329330

Christophe, E., Leger, D., & Mailhes, C. (2004, 10). Comparison and evaluation of quality criteria for hyperspectral imagery. In *SPIE conference proceedings* (p. 5668). https://doi.org/10.1117/12.587107

Chu, Y., Li, S., Wang, M., & Che, Y. (2011). Resarch of image magnifying algorithm based on cubic spline interpolation. *Emeit 2011*, *7*, 3391–3394 doi:10.1109/EMEIT.2011.6023064.

Chunlin, L., Min, K., Xingsheng, W., Xiuqing, F., & Jun, Y. (2014). Study on the finite element simulation of complex surface turning based on piecewise linear interpolation. In *2014 ISDEA* (pp. 976–979). Institute of Electrical and Electronics Engineers (IEEE).

Daudt, R. C. (2018). *Onera satellite change detection dataset*. https://rcdaudt.github.io/oscd/

Dewantara, D. S., Hidayat, R., Susanto, H., & Arymurthy, A. M. (2020). Cnn with multi stage image data augmentation methods for Indonesia rare and protected orchids classification. In *2020 ICOSICA* (pp. 1–5).

Divya, G., & Manish, T. I. (2020). Machine learning techniques and frameworks for hetero- geneous data fusion in big data analytics. In *2020 4th ICECA* (pp. 1568–1574). Institute of Electrical and Electronics Engineers (IEEE).

Dodgson, N. (1997). Quadratic interpolation for image resampling. *IEEE Transactions on Image Processing*, *6* (9), 1322–1326. https://doi.org/10.1109/83.623195

Eckhard, J., Eckhard, T., Valero, E. M., Nieves, J. L., & Contreras, E. G. (2015). Outdoor scene reflectance measurements using a Bragg-grating-based hyperspectral imager. *Applied Optics*, *54*(13), D15–D24. http://www.ugr.es/local/coloring/UGRhyperspectralimagedatabase.html

Fadnavis, S. (2014). Image interpolation techniques in digital image processing: An overview. *IJERA*, *4* 10 (Part–1) , 2248–962270 doi :2248-962270.

Fang, B., Liu, Y., Zhang, H., & He, J. (2022). Hyperspectral image classification based on 3D asymmetric inception network with data fusion transfer learning. *Remote Sensing*, *14*(7), 1711. https://doi.org/10.3390/rs14071711

Fornberg, B. (2021). Improving the accuracy of the trapezoidal rule. *SIAM Review*, *63*(1), 167–180. https://doi.org/10.1137/18M1229353

Gao, L., Li, J., Zheng, K., & Jia, X. (2023). Enhanced autoencoders with attention-embedded degradation learning for unsupervised hyperspectral image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–17 doi:10.1109/TGRS.2023.3267890.

Gao, S., Jiang, X., Zhang, Y., Liu, X., Xiong, Q., & Cai, Z. (2024). Superpixelwise pca based data augmentation for hyperspectral image classification. *Multimedia Tools and Applications* 83 , 1–21 doi:https://doi.org/10.1007/s11042-024-18667-1.

Gavade, A., & Rajpurohit, V. (2016, 01). *Multi spectral super resolution and image quality assessment comparative analysis*.

Graña, M., Veganzons, M., & Ayerdi, B. (2021). *Hyperspectral remote sensing scenes*. https://www.ehu.eus/ccwintco/index.php/HyperspectralRemote Sensing Scenes

Guillaume, M., Minghelli, A., Deville, Y., Chami, M., Juste, L., Lenot, X., Lafrance, B., Jay, S., Briottet, X., & Serfaty, V. (2020). Mapping benthic habitats by extending non-negative matrix factorization to address the water column and seabed adjacency effects. *Remote Sensing*, *12*(13), 2072. https://doi.org/10.3390/rs12132072

Hu, H., Peng, X., Yang, Q., Li, G., Wang, X., Li, G., Sun, J., & Yuan, K. (2021). A geometry information enhanced unet for tumor segmentation. In *2021 IEEE ICMIPE* (pp. 1–5). Institute of Electrical and Electronics Engineers (IEEE).

Jia, S., Shen, L., & Deng, L. (2012). Band selection-based gabor wavelet feature extraction for hyperspectral imagery classification. *2012 WHISPERS* (pp. 1–4). Institute of Electrical and Electronics Engineers (IEEE).

Jiabo, S., & Yanxiang, T. (2019, 01). Research on Runge phenomenon. *Advances in Applied Mathematics*, *8*, 1500–1510. doi:10.12677/AAM.2019.88175

Kartikaningsih, W., Triharianto, R. H., & Gunawan, D. (2021). Identification of military objects around Indonesia using public domain satellite images. *2021 ICACSIS* (pp. 1–7). Institute of Electrical and Electronics Engineers (IEEE).

Kumar, A., & Verma, V. (2022). Modeling human genome (3d) using linear & nearest inter- polation. *2022 8th Icaccs*, *1*, 751–755 doi:10.1109/ICACCS54159.2022.9785329.

Li, J., Zheng, K., Gao, L., Ni, L., Huang, M., & Chanussot, J. (2024). Model-informed multi- stage unsupervised network for hyperspectral image super-resolution. *IEEE Transactions on Geoscience & Remote Sensing*, *62*, 1–17. https://doi.org/10.1109/TGRS.2024.3391014

Li, J., Zheng, K., Li, Z., Gao, L., & Jia, X. (2023). X-shaped interactive autoencoders with cross-modality mutual learning for unsupervised hyperspectral image super-resolution. In *IEEE transactions on geoscience and remote sensing*. doi:10.1109/TGRS.2023.3300043

Liang, Z., & Wang, X. (2022). Semantic segmentation network with band-location adaptive selection mechanism for multispectral remote sensing images. *IGARSS 2022* (pp. 3488–3491). Institute of Electrical and Electronics Engineers (IEEE).

Liu, S., & Zhao, Y. (2023). A unet-like hybrid transformer for efficient semantic segmentation of remote sensing images. *2023 5th GRSM* (pp. 149–154). Institute of Electrical and Electronics Engineers (IEEE).

Loizzo, R., Guarini, R., Longo, F., Scopa, T., Formaro, R., Facchinetti, C., & Varacalli, G. (2018). Prisma: The Italian hyperspectral mission. *IGARSS 2018* (pp. 175–178). Institute of Electrical and Electronics Engineers (IEEE).

Lu, Y., Yang, J., & Xiao, L. (2021). Multi-supervised recursive-cnn for hyperspectral and multispectral image fusion. In *2021 IGARSS* (pp. 2440–2443).

Ma, L., Li, S., Zhou, Z., Yao, Y., & Du, Q. (2023). Semantic segmentation network for classification of hyperspectral images with small size samples. *IEEE GRSL*, *20*, 1–5 doi:10.1109/LGRS.2023.3314466.

Middleton, E. M., Ungar, S. G., Mandl, D. J., Ong, L., Frye, S. W., Campbell, P. E., Landis, D.R., Young, J.P., and Pollack, N. H. (2013). The earth observing one (eo-1) satellite mission: Over a decade in space. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, *6*(2), 243–256. https://doi.org/10.1109/JSTARS.2013.2249496

Nalepa, J., Myller, M., & Kawulok, M. (2020). Transfer learning for segmenting dimensionally reduced hyperspectral images. *IEEE Geoscience & Remote Sensing Letters*, *17*(7), 1228–1232.

Nalepa, J., Tulczyjew, L., Le Saux, B., Longépé, N., Ruszczak, B., Wijata, A. M., Smykala, K., Myller, M., Kawulok, M., Kuzu, R.S., & Albrecht, F. (2024). Estimating soil parameters from hyperspectral images: A benchmark dataset and the outcome of the hyperview challenge. *IEEE Geoscience and Remote Sensing Magazine*, *12*(3), 35–63. https://doi.org/10.1109/MGRS.2024.3394040

Nguyen, Q. P., Wai, L., Divakaran, D. M., Low, K., & Chan, M. (2019, 06). Gee: A gradient-based explainable variational autoencoder for network anomaly detection.

Ogniewski, J. (2019). Spline interpolation in real-time applications using three control points.

Pal, M. K., Rasmussen, T. M., & Abdolmaleki, M. (2019). Multiple multi-spectral remote sensing data fusion and integration for geological mapping. *2019 10th WHISPERS* (pp. 1–5). Institute of Electrical and Electronics Engineers (IEEE).

Pancorbo, J., Quemada, M., & Roberts, D. A. (2023). Drought impact on cropland use mon- itored with aviris imagery in central valley, california. *Science of the Total Environment*, *859*, 160198. https://doi.org/10.1016/j.scitotenv.2022.160198

Plajer, I. C., Băicoianu, A., & Majercsik, L. (2023). AI-Based visualization of remotely- sensed spectral images. In *2023 ISSCS* (pp. 1–4). doi:10.1109/ISSCS58449.2023.10190908. Institute of Electrical and Electronics Engineers (IEEE).

Plajer, I. C., Băicoianu, A., Majercsik, L., & Ivanovici, M. (2023). Ndvi computation from hyperspectral images. *13th WHISPERS*. Institute of Electrical and Electronics Engineers (IEEE).

Rabbath, C., & Corriveau, D. (2019). A comparison of piecewise cubic hermite interpolating polynomials, cubic splines and piecewise linear functions for the approximation of projectile aerodynamics. *Defence Technology*, *15*(5), 741–757. (SI: 2019 International Symp. Ballistics). https://doi.org/10.1016/j.dt.2019.07.016

Saxena, N., & Raman, B. (2020). Semantic segmentation of multispectral images using res-seg-net model. *2020 IEEE ICSC* (pp. 154–157). Institute of Electrical and Electronics Engineers (IEEE).

Schaepman, M. E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., Schneider, F.D., Laurent, V., Popp, C., Seidel, F.C., & Itten, K. I. (2015). Advanced radiometry measurements and earth science applications with the airborne prism experiment (apex). *Remote Sensing of Environment*, *158*, 207–219. https://doi.org/10.1016/j.rse.2014.11.014

Sun, S., Zheng, Y., Li, G., Guo, Z., Bie, Z., & Ma, J. (2022). Missing data reconstruction method of distribution network based on res-at-unet. *2022 CICED* (pp. 508–512). Institute of Electrical and Electronics Engineers (IEEE).

Suri, J. S., Bhagawati, M., Agarwal, S., Paul, S., Pandey, A., Gupta, S. K., Saba, L., Paraskevas, K.I., Khanna, N.N., Laird, J.R., & Naidu, S. (2023). Unet deep learning architecture for segmentation of vascular and non-vascular images: A microscopic look at unet components buffered with pruning, explainable artificial intelligence, and bias. *Institute of Electrical and Electronics Engineers Access*, *11*, 595–645. https://doi.org/10.1109/ACCESS.2022.3232561

Tuzzi, L., Busi, I., Garzonio, R., Cotrozzi, L., Risoli, S., Quaratiello, G., Colombo, R., Cogliati, S., & Sironi, L. (2023). Detection of fusarium head blight of wheat from hyperspectral images. In *2023 IEEE inter- national workshop on metrology for agriculture and forestry (Metroagrifor)* (pp. 516–520). Institute of Electrical and Electronics Engineers (IEEE).

Wang, Y., Gu, L., Jiang, T., & Gao, F. (2023). Mde-unet: A multitask deformable unet combined enhancement network for farmland boundary segmentation. *IEEE Geoscience & Remote Sensing Letters*, *20*, 1–5 doi:10.1109/LGRS.2023.3252048.

Wang, Y., Zhu, Q., Ma, H., & Yu, H. (2022). A hybrid gray wolf optimizer for hyperspectral image band selection. *IEEE TGRS*, *60*, 1–13 doi:10.1109/TGRS.2022.3167888.

Xiang, S., Xie, Q., & Wang, M. (2022). Semantic segmentation for remote sensing images based on adaptive feature selection network. *IEEE Geoscience & Remote Sensing Letters*, *19*, 1–5. https://doi.org/10.1109/LGRS.2021.3049125

Xu, Z., Yu, H., Zheng, K., Gao, L., & Song, M. (2021). A novel classification framework for hyperspectral image classification based on multiscale spectral-spatial convolutional network. *2021 WHISPERS* (pp. 1–5). Institute of Electrical and Electronics Engineers (IEEE).

Yang, W., Geng, H., Xiao, S., & Yang, G. (2015). Linear interpolation model predictive control of large wind turbines for blade asymmetric fatigue loads mitigation. In *EPE'15 ECCE-Europe* (pp. 1–7). Institute of Electrical and Electronics Engineers (IEEE).

Yasuma, F., Mitsunaga, T., Iso, D., & Nayar, S. (2008). *Cave Dataset: Multispectral Image Database*. https://www.cs.columbia.edu/CAVE/databases/multispectral/

Zhang, L., Xie, Y., Xidao, L., & Zhang, X. (2018). Multi-source heterogeneous data fusion. In *2018 ICAIBD* (pp. 47–51). Institute of Electrical and Electronics Engineers (IEEE).

Zhang, Y., Zhang, D., & Wang, T. (2018). Super-resolution classification of hyperspectral images with a small training set using semi-supervised learning. In *2018 WHISPERS* (pp. 1–5). Institute of Electrical and Electronics Engineers (IEEE).

Zheng, K., Gao, L., Liao, W., Hong, D., Zhang, B., Cui, X., & Chanussot, J. (2020). Coupled convolutional neural network with adaptive response function learning for unsupervised hyperspectral super resolution. *IEEE Transactions on Geoscience & Remote Sensing*, *59*(3), 2487–2502. https://doi.org/10.1109/TGRS.2020.3006534

Zhu, S., Zeng, B., Liu, G., Zeng, L., Fang, L., & Gabbouj, M. (2015). Image interpolation based on non-local geometric similarities. In *2015 IEEE ICME* (pp. 1–6). Institute of Electrical and Electronics Engineers (IEEE).