Comparison of Pigment Classification Algorithms on Non-Flat Surfaces using Hyperspectral Imaging

Dipendra J. Mandal[1], Marius Pedersen[1], Sony George[1], and Clotilde Boust[2]

[1]Department of Computer Science, Norwegian University of Science and Technology (NTNU), Norway

[2]Center for Research and Restoration of Museums of France (C2RMF), France

Abstract

Cultural heritage objects, such as paintings, provide valuable insights into the history and culture of human societies. Preserving these objects is of utmost importance, and developing new technologies for their analysis and conservation is crucial. Hyperspectral imaging is a technology with a wide range of applications in cultural heritage, including documentation, material identification, visualization and pigment classification. Pigment classification is crucial for conservators and curators in preserving works of art and acquiring valuable insights into the historical and cultural contexts associated with their origin. Various supervised algorithms, including machine learning, are used to classify pigments based on their spectral signatures. Since many artists employ impasto techniques in their artworks that produce a relief on the surface, i.e., transforming it from a flat object to a 2.5D or 3D, this further makes the classification task difficult. To our knowledge, no previous research has been conducted on pigment classification using hyperspectral imaging concerning an elevated surface. Therefore, this paper compares different spectral classification techniques that employ deterministic and stochastic methods, their hybrid combinations, and machine learning models for an elevated mockup to determine whether such topographical variation affects classification accuracy. In cultural heritage, the lack of adequate data is also a significant challenge for using machine learning, particularly in domains where data collection is expensive, time-consuming, or impractical. Data augmentation can help mitigate this challenge by generating new samples similar to the original. We also analyzed the impact of data augmentation techniques on the effectiveness of machine learning models for cultural heritage applications.

Keywords – *Pigment Classification, Impasto, Supervised Classification, Machine Learning, Hyperspectral Imaging*

1 Introduction

In Cultural Heritage (CH), paintings are an essential tangible component that provides valuable insights into our history, social norms, and beliefs. Therefore, the preservation and restoration of paintings are crucial and poses numerous challenges, including the removal of dirt and old varnish without damaging the paint layer [1] and the selection of appropriate materials for retouching [2]. To address these problems, it is essential to accurately identify the pigment used by the artist in an artwork. Scientific analysis sometimes require the physical samples, however, due to the nature of CH objects, it is not recommended to take samples from the object which in fact destroy the object even at a microscale and so very often, it is not permitted. Consequently, non-invasive or non-contact imaging techniques [3, 4, 5, 6] are necessary.

Hyperspectral Imaging (HSI) is a technology that has gained increasing attention in recent years due to its wide range of applications in various fields, including remote sensing [7], agriculture [8], medical sciences [9], forensics [10], biomedical engineering [11], and CH [12]. In addition in CH, spectral system have also been used for measuring special appearance properties such as materials that are goniochromatic [13] or by integrating 3D information with spectral data [14]. An important aspect of CH is pigment classification [15], and HSI can facilitate it by using spectral information about pigments and different classification algorithms. In this work we focus on pigment classification using hyperspectral imaging. Artwork is not confined to two-dimensional canvases or boards, which means they are not always flat. The addition of relief, which introduces 2.5D or 3D to artwork, is also an

important consideration [16, 17]. Several factors can contribute to this third dimension [18, 19, 20]; For example, morphological textures of brushstrokes on the painted surface [21], a thick layer of pigments applied by many renowned artists to their artwork for creating depth(impasto technique) [22, 23, 24, 25]. The geometry of a relief raised from a brush painting, impasto techniques, or any other factors may affect how light interacts with the surface, affecting the spectral signature captured by the hyperspectral sensor for a given pixel. Considering the importance of pigment identification or classification of an artwork, the influence of such factors must be explored.

Most of the research conducted so far has explored the effectiveness of various traditional supervised algorithms and machine learning models for pigment classification using HSI [26, 27]. Pigment classification has been done for single pigments or mixtures of pigments. In the latter case, unmixing is required to identify the different pigments. When the materials are known pigment mapping is used to find the right combination of the materials and their relative concentrations [28]. This can be done by using the optical propeties of the material and a mixing model. Several models for this exist, for example based on Kubelka-Munk theory [29, 28, 30, 31]. Recently, Fukomoto et al. [32, 33] used an encoder-decoder neural network to estimate pigment concentrations. In our work we deal with a single pigment, not mixtures of pigments. The existing studies have primarily focused on flat surfaces and, to date, no research has investigated the same for artwork with an elevation. This paper aims to investigate how surface elevation in artworks affects the accuracy of pigment classification using HSI, with the underlying hypothesis that surface elevation impacts this accuracy. In CH, the lack of sufficient training datasets is also a considerable challenge for using machine learning, particularly in domains where collecting data is expensive, time-consuming, or impractical [34, 35]. Data augmentation is a technique that can help mitigate this challenge by generating new samples similar to the original data. Therefore, in this

paper, we have also compared and analyzed the impact of data augmentation techniques on the effectiveness of machine learning algorithms for pigment classification. This study is primarily concerned with answering the following research questions:

- To what extent does elevation of a surface in artworks affect the accuracy of pigment classification using hyperspectral imaging?
- What is the influence of data augmentation techniques on the efficacy of machine learning models for material classification?

The rest of the paper is organized as follows. Section 2 briefly reviews the classification algorithms used in the CH field, focusing on pigment classification. Section 3 provides a brief overview of the classification algorithms used in this study. Section 4 describes the materials and methods used in the present study. Section 5 presents the results of the experiments and discusses the findings in detail. Finally, Section 6 concludes the paper and highlights directions for future research.

2 Background

Over the last few years, HSI, a non-invasive technique, has been widely employed for pigment classification in artwork [36, 5, 26]. It has resulted in significant advances in the study of spectral signatures and matching, broadening the scope of HSI technology in the CH domain. Molecules in the materials have unique vibration frequencies, which can be detected by analyzing how they absorb or reflect light at specific wavelengths. These characteristics of materials are known as spectral signatures and help to identify and distinguish pigment based on how it interacts with electromagnetic radiation. Many spectral matching algo-

rithms have evolved in hyperspectral image processing, ranging from traditional clustering techniques to more recent automated matching models. The approaches used for matching spectra can be classified as deterministic, i.e., based on geometrical and physical aspects, or stochastic, which is based on the distributions [37]. These algorithms are essential for accurate and efficient pigment identification and analysis, making them a critical component of any HSI workflow. This section will briefly overview the classification algorithms employed in CH, specifically for pigment classification.

The Euclidean Distance (ED) metric is widely used to measure spectral similarity in HSI, and it works well when a data set has distinct or isolated clusters [38, 39, 40]. Mandal et al. [27] implemented ED and other supervised classification techniques to classify pigments on a flat surface and observed that classification accuracy declines for some pigments with similar spectral characteristics.

Deborah et al. [5] explored the application of HSI in mapping the pigments of Edvard Munch's painting The Scream. They used two methods for spectral image classification, namely Spectral Angle Mapper (SAM) and Spectral Correlation Mapper (SCM). They observed that SCM performed better than SAM as it accounted for both positive and negative correlations between the spectra. Adjusting the threshold value for SAM reduced false detection, but it also decreased the accuracy of pigment classification, which varied depending on the type of pigment. They further suggested that different classes of pigments would require the use of distinct threshold values. In [41], the authors used the SAM technique to classify traditional Chinese pigments and recommended a similar suggestion of using different thresholds. The SAM algorithm measures the angle between two vectors, independent of the vector length, and therefore, insensitive to gain. As a result, this algorithm does not account for magnitude shifts in the spectrum. Please refer to Osmar et al. [42] for more information.

George and Hardeberg [43] demonstrate the usefulness of HSI to separate inks using SAM and SID as classification algorithms. They found that the SID algorithm performs better than SAM in cases where two distinct inks were overlaid. However, they also mentioned misclassification might arise from noise and non-uniformity in spectral signatures resulting from ink-paper blends. Mishra et al. [44] used HSI to evaluate hybrid spectral similarity measures to classify paper samples used in forensic investigations. The findings indicate that the hybrid similarity measures of SIDSCM demonstrate better classification accuracy than conventional spectral similarity measures. Furthermore, the classification accuracies of SIDSCM and SIDSAM are comparable.

The authors [45] discuss the use of SAM and machine learning (ML) models to classify mineral pigments used in ancient Chinese paintings using HSI. The results show that, for similar colors and spectra, SAM is unable to classify; however, combining it with a decision tree can effectively improve the accuracy. The authors of [46] discuss the effectiveness of HSI technology in archaeological research for identifying and classifying materials in ancient tombs. They found that combining HSI data with Principal Component Analysis (PCA) transformation and SVM classification was an effective method for accurately classifying and identifying materials. The SVM classification based on feature bands improved classification accuracy and reduced data processing time. Kleynhans et al. [47] discuss using reflectance HSI and Neural Networks (NNs) to create labeled pigment maps of paintings. The authors report that a one-dimensional convolutional neural network (1D-CNN) model could accurately label the pigments in most of the paintings studied. However, this finding highlights the importance of having comprehensive training data for the model to perform well, the need for further studies to expand the training dataset, and the possibility of augmenting existing training datasets to develop a more robust solution.

Lie et al. [48] explore the potential of using NNs to analyze HSI data in the CH field. They

present a thorough overview of the different applications and constraints of NNs models. The findings indicate that NNs offer a promising alternative to conventional statistical and multivariate analysis techniques for pigment identification and classification. The authors in [49] present a method for identifying pure pigments in CH using a combination of CNNs and SCM. The HSI data, collected within the range of 400nm to 720nm and at a resolution of 10nm, was pre-processed by smoothing and computing the first derivative before being fed to the network. The study emphasizes the significance of employing deep learning NNs for this application and the requirement for a comprehensive training dataset. A recent study by Mandal et al. [27] investigated the performance of various traditional supervised algorithms, their hybrid combinations, and ML models for pigment classification on flat surfaces. A research gap exists as the efficacy of these algorithms on non-flat objects has not been explored by any of the authors. Thus, further investigation is necessary for this area.

3 Classification Algorithms

This section presents the fundamental theory and mathematical expressions for the classification algorithms used in the study.

3.1 Euclidean Distance

Euclidean Distance (ED) is a distance metric that measures the distance between two points in an N-dimensional space [50]. It is calculated as the square root of the sum of the squared differences between the corresponding elements of the two points. In spectral analysis, the ED can compare the similarity between two spectra by measuring the difference between their respective pixel or spectral values. The formula for the Euclidean distance between

the image spectrum t_i and a reference spectrum r_i, each with n elements, is defined using Equation (1).

$$ED = \sqrt{\sum_{i=1}^{nb} (t_i - r_i)^2}, \tag{1}$$

where, nb is the number of spectral bands.

3.2 Spectral Angle Mapper

Spectral Angle Mapper (SAM), introduced by Boardman in 1992 [51], is a method for measuring the spectral similarity between two spectra, i.e., a test (also referred to as a target) and reference spectrum. This technique treats the spectra as N-dimensional vectors in space, where N is the number of spectral bands and calculates the arccosine angles between them. The spectral angle, α, between the two spectra is computed using Equation (2). A smaller angle indicates a better match between the spectra.

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} {t_i}^2}\sqrt{\sum_{i=1}^{nb} {r_i}^2}}\right), \tag{2}$$

where t_i represents the image spectrum, r_i denotes the reference spectrum, and nb is the total number of bands.

3.3 Spectral Correlation Mapper

Spectral Correlation Mapper (SCM) is one of several algorithms used in spectral similarity analysis for classification and feature extraction. It measures the Pearson correlation coefficient between two spectra by standardizing the data and centering them around the mean of the test and reference spectra. The result is then expressed as an angle using the arccosine function. This algorithm excludes negative correlation and retains the shading effect minimization characteristics similar to SAM, resulting in better classification results [42, 52]. SCM is computed using Equation (3).

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{nb}\left(t_i - \bar{t}_i\right)\left(r_i - \bar{r}_i\right)}{\sqrt{\sum_{i=1}^{nb}\left(t_i - \bar{t}_i\right)^2 \sum_{i=1}^{nb}\left(r_i - \bar{r}_i\right)^2}}\right) \tag{3}$$

where α is the arccosine of the spectral correlation measure in radians, t_i and $\bar{t}_i$ are the image spectrum and its sample mean, similarly r_i and $\bar{r}_i$ are the reference spectrum and its sample mean; and nb is the total number of bands.

3.4 Spectral Information Divergence

In terms of spectral similarity, Spectral Information Divergence (SID) measures the dissimilarity between two spectra by comparing their spectral information content. It is based on the concept of Kullback-Leibler (KL) divergence, a measure of the difference between two probability distributions. In SID, the spectral information content of each pixel is modeled as a probability distribution, and the divergence between the two distributions is calculated. If two pixels have similar spectral information, their probability distributions will be similar, and the SID value will be low, and vice versa. The probability distribution of the

test and reference spectra is expressed as Equation (4) and Equation (5), respectively [53].

$$p_i = \frac{t_i}{\sum_{i=1}^{nb} t_i} \tag{4}$$

$$q_i = \frac{r_i}{\sum_{i=1}^{nb} r_i} \tag{5}$$

where, t_i is the image spectrum, r_i is the reference spectrum, and nb is the total number of bands. Using these two probability distributions, SID can be calculated with Equation (6).

$$SID = \sum_{i=1}^{nb} p_i \log\left(\frac{p_i}{q_i}\right) + \sum_{i=1}^{nb} q_i \log\left(\frac{q_i}{p_i}\right) \tag{6}$$

3.5 Spectral Similarity Scale

The Spectral Similarity Scale (SSS) provides a quantitative measure of the similarity between two spectra. This algorithm uses the Euclidean distance metric for magnitude and correlation to compare the shape of the spectra. This method combines both calculations, giving each equal weighting [54]. An SSS value of 0 indicates that the two spectra are identical, while a value of 1 indicates that the two spectra are entirely dissimilar. SSS is computed using Equation (7).

$$SSS = \sqrt{(d_e)^2 + (\hat{r})^2} \tag{7}$$

where, d_e is the Euclidean distance between two spectra and is computed using Equa-

tion (8) and its value ranges from 0 to 1 due to the factor $1/nb$.

$$d_e = \sqrt{\frac{1}{nb}\sum_{i=1}^{nb}(t_i - r_i)^2} \tag{8}$$

Equation (9) computes the value for $\hat{r}$, where r is the correlation coefficient between the two spectra and is computed using Equation (10).

$$\hat{r} = (1 - r^2) \tag{9}$$

$$r^2 = \left(\frac{\sum_{i=1}^{nb}(t_i - \bar{t}_i)(r_i - \bar{r}_i)}{\sqrt{\sum_{i=1}^{nb}(t_i - \bar{t}_i)^2 \sum_{i=1}^{nb}(r_i - \bar{r}_i)^2}} \right)^2 \tag{10}$$

3.6 Spectral information divergence-spectral angle mapper

The Spectral Information Divergence Spectral Angle Mapper (SIDSAM) is a hybrid approach that incorporates quantitative and qualitative matching measures. It utilizes the SID algorithm to assess the dissimilarity between two spectra and the SAM algorithm to evaluate their geometric similarity. This hybrid computation enhances the comparability of similar spectra and makes dissimilar spectra more distinctive, thus improving spectral discriminability. SIDSAM is computed by multiplying the SID by the tangent or sine function of the SAM, which calculates the perpendicular distance between the test and reference vectors. Both measures yield similar results, as reported in previous studies [55]. This hybrid computation SIDSAM can be calculated using either Equation (11) or (12).

$$SIDSAM = SID * \tan(SAM) \tag{11}$$

$$SIDSAM = SID * \sin(SAM) \tag{12}$$

where, SID and SAM are calculated using Equations (6) and (2), respectively.

3.7 Spectral information divergence-spectral correlation mapper

The Spectral Information Divergence Spectral Correlation Mapper (SIDSCM) is another hybrid approach that combines qualitative and quantitative metrics to increase spectral discriminability. It combines SID and SCM algorithms, similar to SIDSAM, where SID measures the difference between two spectra and SCM determines the Pearson correlation coefficient. To integrate the two measures, the product of SID and either the tangent or sine function of the correlation coefficient between two spectra is used [56]. The resultant method may be calculated using either Equation (13) or (14).

$$SIDSCM = SID * \tan(SCM) \tag{13}$$

$$SIDSCM = SID * \sin(SCM) \tag{14}$$

where SID and SCM can be computed using Equations (6) and (3) respectively.

3.8 Jeffries Matusita-Spectral Angle Mapper

The Jeffries Matusita (JM) distance is a statistical metric considering the covariance of two spectral vectors. The SAM method computes the angle between two spectral vectors to determine their spectral similarity. Jeffries Matusita-Spectral Angle Mapper (JMSAM) is

calculated by first calculating the JM distance and then converting it to an angle with the inverse cosine method. The angle obtained is then compared to a threshold value to assess whether the two spectra belong to the same class. It can increase spectral classification accuracy by considering the spectral similarity and statistical distance [57]. It can be computed using either Equation (15) or (16).

$$JMSAM = JMD * \tan(SAM) \tag{15}$$

$$JMSAM = JMD * \sin(SAM) \tag{16}$$

where, JMD is JM distance and is computed using Equation (17)

$$JMD = 2\left(1 - e^{-B}\right) \tag{17}$$

Here B is the Bhattacharyya distance and is computed using Equation (18).

$$B = \frac{1}{8}\left(\mu_t - \mu_r\right)^T \left[\frac{\sigma_t + \sigma_r}{2}\right]^{-1} \left(\mu_t - \mu_r\right) + \frac{1}{2} ln \left[\frac{\left|\frac{\sigma_t + \sigma_r}{2}\right|}{\sqrt{|\sigma_t||\sigma_r|}}\right], \tag{18}$$

where, μ_t and μ_r are the mean of the test and reference spectra, respectively; σ_t and σ_r are the covariance of the test and reference spectra, respectively.

3.9 Support Vector Machine

Support Vector Machine (SVM) is a ML algorithm used for classification and regression analysis [58]. SVM classification aims to find a hyperplane that separates the data into two classes with maximum margin. It can handle non-linearly separable data using a kernel trick

that maps it into a higher-dimensional space. This algorithm involves data preprocessing, such as normalization, to ensure that the features are on the same scale. It selects the most relevant features for the classification task. After that, it trains the model by finding the optimal hyperplane that maximizes the margin using a cost function. The cost function penalizes misclassified data points and encourages the SVM to find the hyperplane that separates the data with the largest margin. Finally, the model is tested on a validation set or test data set to evaluate its performance [59, 60].

3.10 1D-Convolutional Neural Network

A NN is a machine learning algorithm inspired by the structure and function of the human brain. The basic building block of a neural network is the artificial neuron, which takes inputs and applies a transformation to produce an output [61]. The architecture of a NN can vary widely depending on the task and the data being used. A 1D Convolutional Neural Network (CNN) is a type of NN commonly used for processing one-dimensional data such as time series, audio signals, and text data [62, 63, 64]. It consists of one or more convolutional layers, a pooling layer, and fully connected layers. The convolutional layer applies convolution operations to the input sequence using a set of learnable filters or kernels. This generates a set of feature maps representing the convolutional layer's output. The pooling layer is typically used to reduce the dimensionality of feature maps while maintaining the most important information. It applies an aggregation function such as max or average pooling to extract the most relevant features from each feature map. This helps reduce the number of parameters in the model and prevent overfitting. The fully connected layer takes the output of the pooling layer, flattens it into a one-dimensional vector, and passes it through a set of fully connected neurons. The output of the fully connected layer is often fed into a softmax function to generate class probabilities [65].

Overall, the 1D-CNN architecture is designed to extract and learn discriminative features from one-dimensional data sequences, making it suitable for various applications. Figure 1 illustrates the general architecture of a 1D CNN for use with HSI data.

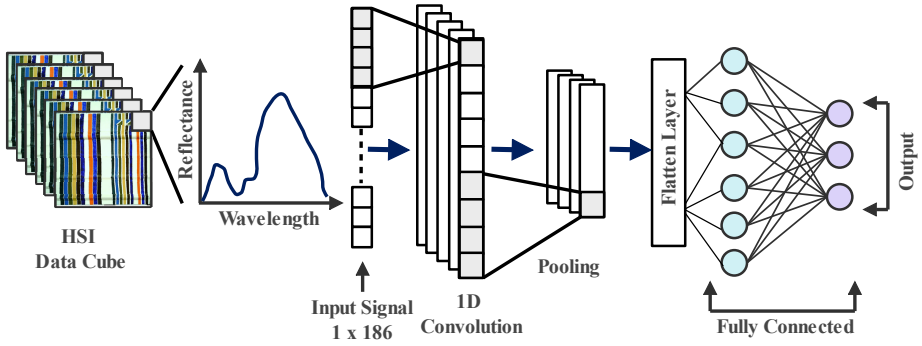


Figure 1: The architecture of 1D-CNN; typically comprises three fundamental layers: convolutional, pooling, and fully connected layer. Input from the HSI datacube (pixel value across wavelength) is fed to the convolutional layers, which apply a sliding window over given input data to perform feature extraction. The pooling layers reduce the size of the extracted features, and the fully connected layers classify the input based on the features obtained from the previous layers.

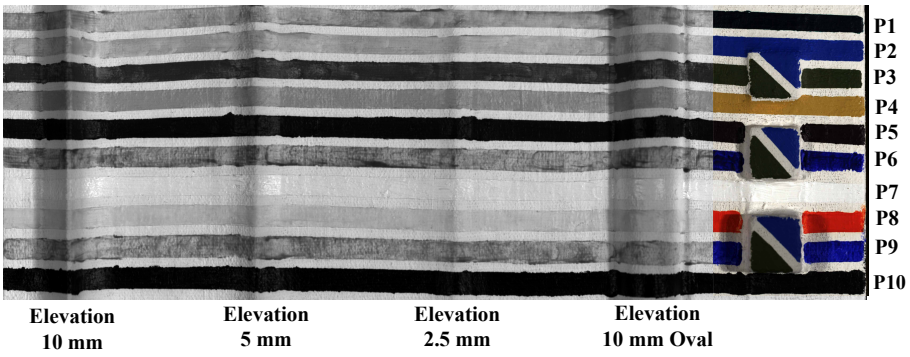
4 Materials and Methods

This section describes the mockup used in the study and the HSI acquisition laboratory setup. Additionally, we will describe the steps used to process the HSI data for the classification task.

4.1 Mockup

As shown in Figure 2a, a pigment mockup was prepared and used in a laboratory environment. The mockup's base was printed using 3D printing. It consisted of different elevation levels, including a flat surface and regions raised to different levels 2.5mm, 5mm, and 10mm from the base. A linen fabric was glued to the surface of the base. Three layers of white

gesso were applied evenly across the entire surface of the canvas. The surface was carefully sanded between each layer using sandpaper to create a smooth and even coat. Figure 2b illustrates the grayscale depth map of the pigment mockup to visualize the relative distances of different elevations from the camera's perspective.



(a)

(b)

Figure 2: The pigment mockup used in the study. (a) The mockup consisted of ten pigments labelled P1 to P10. These pigments were applied to a surface that included a flat region and three different elevation levels, namely 2.5mm, 5mm, and 10mm. For better visualization of the elevation, the left part of the image is shown in grayscale, captured at 998 nm in the IR region, while the right side is a colour image produced using bands at 640 nm, 551 nm, and 458 nm. (b) A grayscale depth map of the mockup, this map is inverted so black represent the foreground or the regions that has highest elevation (10mm height).

The selection of pigments for the research work was a crucial step to ensure that the mockup

accurately reflected the properties of pigments commonly used in historical artworks. The pigments were carefully chosen on the basis of their spectral characteristics and frequent appearance in CH research articles. The selection process also involved consultations with CH experts. Furthermore, web-based research was conducted on 164 known paintings from different centuries to determine the final pigments selection. We selected the pigments that were used most frequently in those paintings. This approach ensured that the pigments used in the mockup represented those used in real paintings, making the research findings more applicable to real-world scenarios. The final selection of pigments included Veridian (V), Cerulean Blue (CB), Green Earth (GE), Yellow Ochre Light (YOL), Blue Cobalt (BC), Ultramarine Blue Deep (UBD), Lead White Hue (LWH), Genuine Vermilion (GV), Burnt Umber (BU), and Ivory Black (IB). Each pigment was applied to the mockup surface with a width of 6mm, and a 3mm gap between each pigment. Each pigment was applied to the mockup surface with a width of 6mm and a 3mm gap between each pigment. We tried to achieve a uniform thickness for all pigments. Additionally, the pigments utilized were mostly opaque, effectively concealing the underlying substrate. The tubes were purchased from Zecchi, a supplier of art materials [66]. Safflower oil was used for the whites, while linseed oil was used for all other pigment tubes.

4.2 HSI Acquisition Setup

Hyperspectral images were acquired in a laboratory using HySpex VNIR-1800, a line scanner camera developed by Norsk Elektro Optikk [67], and a translation stage setup, the schematic illustration of the HSI system used in this study is illustrated in Figure 3. The detector of the HySpex camera consists of an actively cooled and stabilized complementary metal-oxide-semiconductor (CMOS). It has a spectral range of 400 to 1000 nm with 186 spectral bands having a spectral resolution of 3.26 nm. A computer with HySpex-GROUND

software controlled the HSI acquisition system. This software synchronizes the scanning speed for the integration time set by the user. We have used a close-range 30cm cylindrical lens for the acquisition. This allows to capture 1800 spatial pixels across a line with a field of view of approximately 86mm.

During the experiment, we positioned the Spectralon® [68], a ColorChecker [69], and a pigment mockup on a movable part of the translation stage, as shown in Figure 3. These were placed at the same horizontal level and perpendicular to the camera's focal axis [70]. The Spectralon®, a multi-step reference target with four adjacent panels with reflectance levels of 99%, 50%, 25%, and 12% was used to calculate the normalized reflectance at the pixel level. To verify the obtained spectral data, we used a ColorChecker.

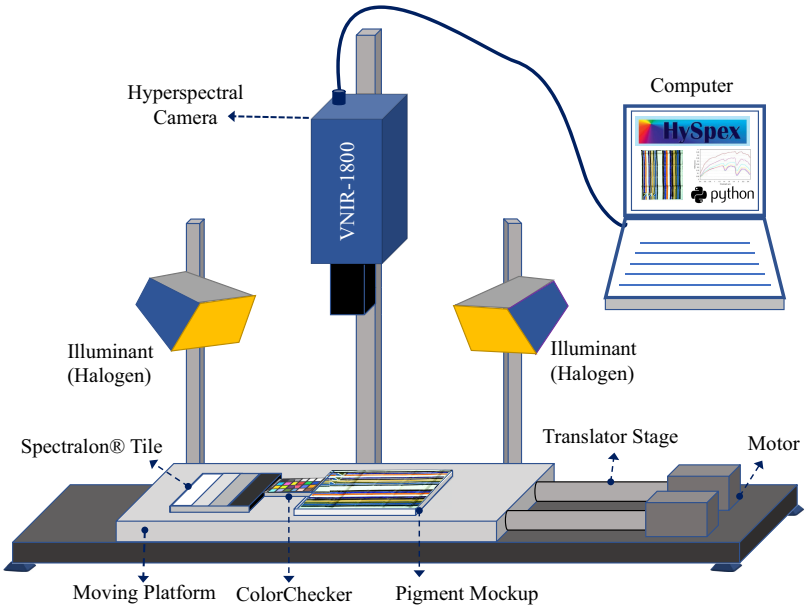


Figure 3: Schematic illustration of the HSI system used in the experiment. A 500 Watt halogen-based floodlight was used for illumination, and the illumination geometry was set to 45°-0°-45°, where 0° is the camera angle to normal.

4.3 Data Processing

The raw hyperspectral data are preprocessed for dark current factor, sensor correction, and radiometric calibration using HySpex RAD software provided by the camera manufacturer. The preprocessed data (converted to sensor-level absolute radiance value) are then converted to normalized reflectance using the known reflectance value of the reference target used in the experiment. The reference target surface might have some variation in the pixel value, so we averaged the values from 100 pixels for each line scan and calculated the reference target radiance value. Due to the shorter distance between the sensor and the object, we assumed that the path radiance effect was negligible. The spectral data were then cropped to exclude the ColorChecker and the reference target. The data processing steps were computed using the open-source software Python 3.9 [71]. Equation 19 provides the mathematical formulation used for conversion. Further data processing steps for supervised and ML-based classifications are discussed in the following sections.

$$R_{Obj}(\lambda) = R_{\mathrm{Ref\_t}}(\lambda)\frac{r_{Obj}(\lambda)}{r_{\mathrm{Ref\_t}}(\lambda)}, \tag{19}$$

where $R_{Obj}(\lambda)$ is the reflectance of an object, $R_{\mathrm{Ref\_t}}(\lambda)$ is the reflectance of reference target, $r_{Obj}(\lambda)$ and $r_{\mathrm{Ref\_t}}$ are the absolute sensor radiance values for the object and the reference target, respectively.

4.3.1 Data Processing Steps for Supervised Classification Algorithms

To conduct supervised classification, reference or ground truth spectra are required to compare similarity. For this purpose, we selected a flat region with dimensions of around 10×10 pixels to establish a spectral library. We then saved the mean spectra for each pigment based on these regions, ensuring that the number of pixels was consistent. The

plot for the spectrum of ten pigments and a substrate is included in the Appendix D. Our approach to constructing this library involved considering three different elevations for each pigment, as well as calculating an average spectrum that accounted for both elevated and flat surfaces. Data processing steps are shown with a block diagram in Figure 4.

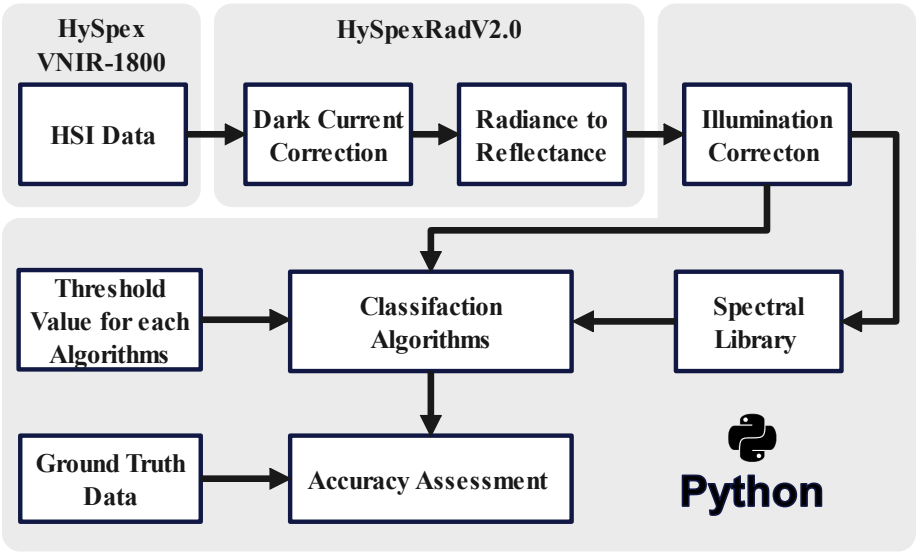


Figure 4: Workflow diagram illustrating the data processing steps for supervised classification algorithms used in the present study.

Selecting a threshold value is an important step in spectral matching [5, 72, 41], which involves identifying specific spectral ranges as belonging to one of several given classes. To achieve optimal classification results, the threshold value must balance minimizing misclassification rates and maximizing the number of correctly classified pigments. Mandal et al. [27] employed an empirical approach to determine an optimal threshold value, where the authors chose a small section of the HSI dataset from a mockup and extracted the reference spectrum by averaging 11 $\times$ 11 pixels from a flat region. They tested a range of values, computed the classification accuracy for various algorithms, and evaluated their accuracy using a confusion matrix. The threshold value used in our study was directly taken from

this research and is mentioned in Table 1

Table 1: The selected threshold value for eight different classification algorithms [27].

Algorithms	Threshold Value
ED	0.9
SAM	0.1
SCM	0.8
SID	0.03
SSS	1.1
SIDSAM	0.003
SIDSCM	0.005
JMSAM	0.09

4.3.2 Data Processing Steps for ML-based Classification Algorithms

Before feeding the normalized reflectance HSI data to the ML model, the data were labeled for distinct classes using a label encoder. We used one hot encoder for our dataset, where each class has one hot value (1), and the rest are cold (0). The data was then split into training and testing sets using an 80-20 split and was further standardized. Subsequently, the model was built and implemented. The training dataset was used to train the model by updating the weights and biases of neurons with each epoch until a considerably low Mean Square Error (MSE) and high accuracy were achieved. Once the model was trained, the test dataset was used to validate its performance. The overall workflow is illustrated in a block diagram in Figure 5.

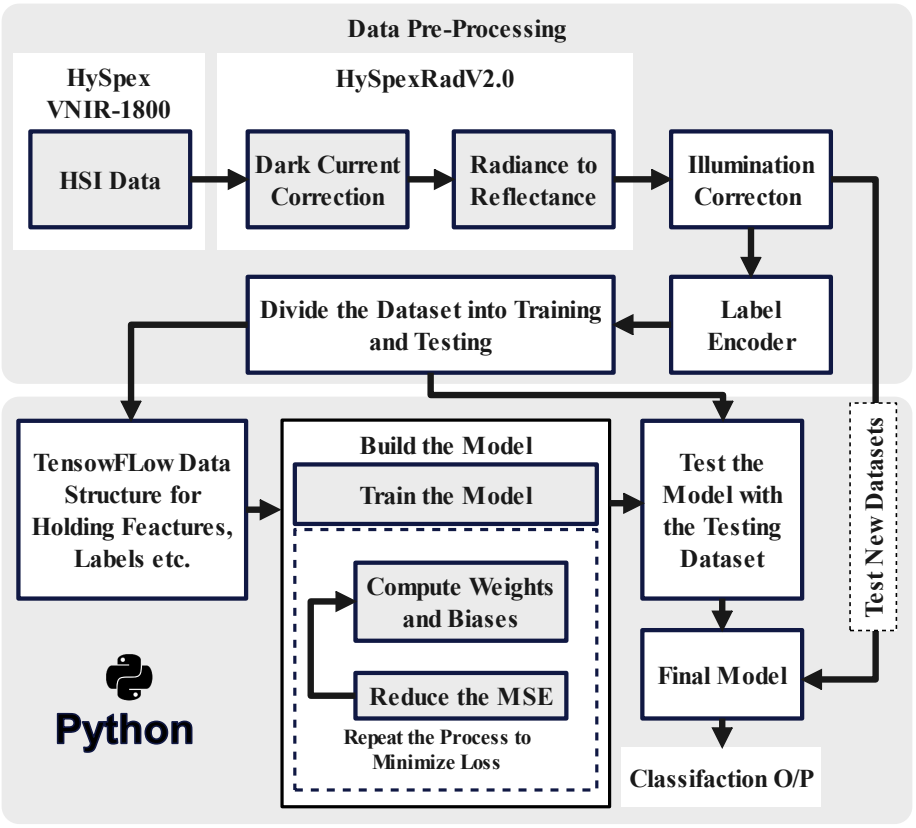


Figure 5: Workflow diagram illustrating the data processing steps for ML.

Hyperparameter tuning is crucial in building robust and better generalized SVM models [73]. We tuned our SVM model for three key hyperparameters: kernel function, regularization (C), and gamma [59]. The kernel function transforms the input space into a higher-dimensional feature space, allowing the SVM to find a hyperplane (decision boundary) that can separate the classes. Several kernel functions are available, including linear, polynomial, and radial basis functions (RBF). The C parameter in SVM introduces a penalty for each misclassified data point. A smaller value of C results in a low penalty for misclassifications, leading to a decision boundary with a larger margin but more misclassifications. On the other hand, a larger value of C results in a higher penalty for misclassifications, leading

to a decision boundary with a smaller margin and fewer misclassifications. The gamma parameter of RBF controls the distance of influence of a single training point. Low gamma values indicate a large similarity radius, resulting in more points being grouped together. For high gamma values, the points must be very close to each other to be considered in the same class. Therefore, models with very large gamma values tend to overfit. If the gamma is large, the effect of C becomes negligible. GridSearch cross-validation was used to optimize the hyperparameters of the SVM model [74]. This involved generating and testing the model for every possible combination of algorithm parameters specified in a grid. Table 2 shows the details of the hyperparameters.

Table 2: SVM hyperparameters, tuning range, and the optimal value selected for classification.

Hyperparameter	Range Used	Optimum Value Selected
Kernel	'Polynomial', 'RBF', 'Sigmoid', 'Linear'	RBF
C	0.1, 1, 10, 100, 1000	100
Gamma	1, 0.1, 0.01, 0.001,0.0001	1
k-fold	5	5

To optimize the hyperparameters of the 1D-CNN model, we used KerasTuner [75].The tuning process involved adjusting the number of convolutional layers, filter size, dropout rate, dense layer filter size, learning rate, and epoch. The resulting optimized model is illustrated in Figure 6, along with the specific hyperparameters used. The Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss function was used in the training process.

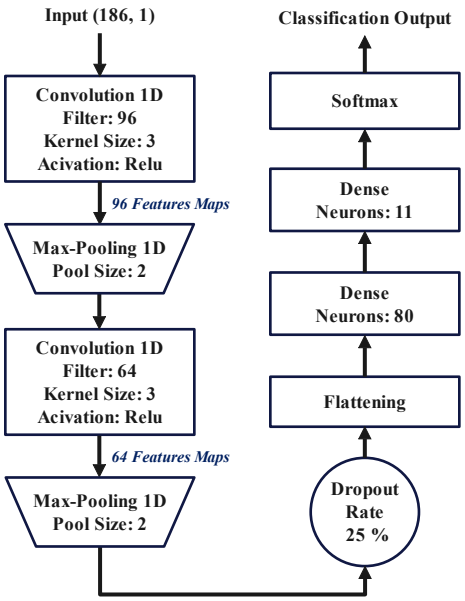


Figure 6: The architecture of a tuned 1D CNN model with optimized hyperparameters.

4.4 Data Augmentation

In ML, data augmentation is a method employed to expand the size of the training dataset by implementing various transformations on the available training data samples. The fundamental idea behind data augmentation is that alterations made to the labeled data should not modify the semantic interpretation of the labels [76, 77]. McFee et al. [78] suggested using deformation techniques that preserve the semantics of audio signals, improving the model's accuracy for the music classification task. Bjerrum et al. [79] used data augmentation techniques on spectral data to employ deep learning algorithms to predict drug composition in tablets using near-infrared regions. The results showed that data augmentation improves overall performance.

The data augmentation technique should be chosen based on the specific characteristics of the analyzed signal. Some techniques might be more appropriate than others, depending on

the context. In our study, we augmented the datasets by introducing four attributes to the spectrum: offset, multiplication, Gaussian noise [80], and speckle noise [81]. An example of the implementation of these attributes is shown in Figure 7. Offset was varied ±(0.0001 to 0.1 with a step size of 0.001) times the standard deviation of the training set. Multiplication was done with 1±(0.0001 to 0.1 with a step size of 0.001) times the standard deviation of the training set, and the two different noises, Gaussian distributed additive noise and speckle which is a multiplicative noise, were added ten times with variation of 0.00001 and .000001 respectively. Using augmentation techniques, we produced two different training datasets. The first training dataset was generated by considering only a single spectrum from the spectral library, whereas for the second training dataset, we augmented each spectrum within the training dataset.

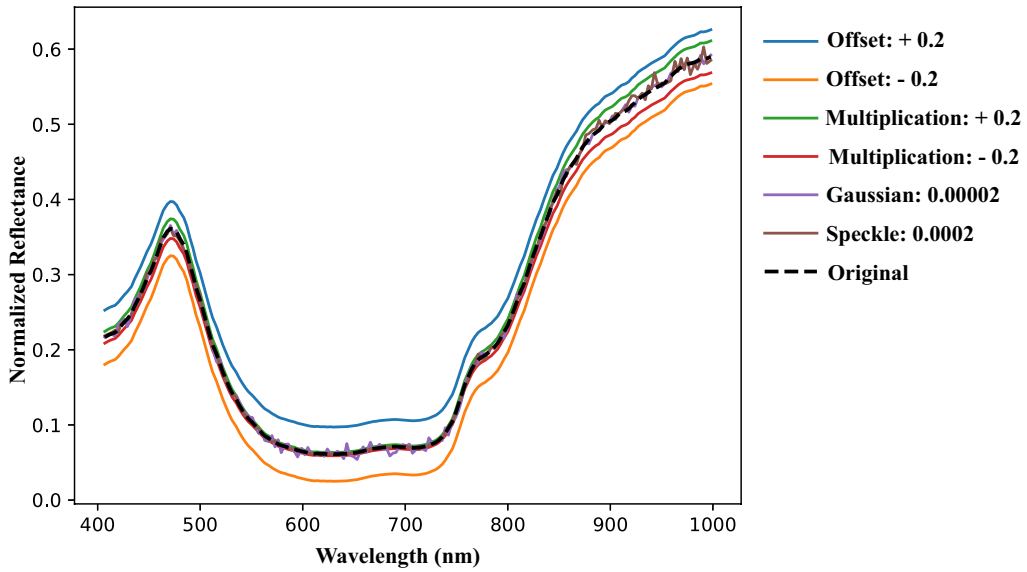


Figure 7: An example of data augmentation; is created with an addition of offset and multiplication to data by a factor of 0.2; Gaussian noise and speckle noise are added to the data with variations of 0.00002 and 0.0002, respectively.

4.5 Accuracy Assessment

Accuracy assessment is an important step in evaluating the performance of classification algorithms. The most common and widely accepted method to express classification accuracy is confusion matrices. It helps to visualize the cross-tabulation of classified pigments; the matrix's main diagonal represents the correctly classified values, while the other elements indicate how many pixels in one category are incorrectly classified into other categories. For additional information on the confusion matrix, we refer to the work of Congalton [82]. For each algorithm, we calculate the accuracy for the predefined region illustrated in Figure 8.

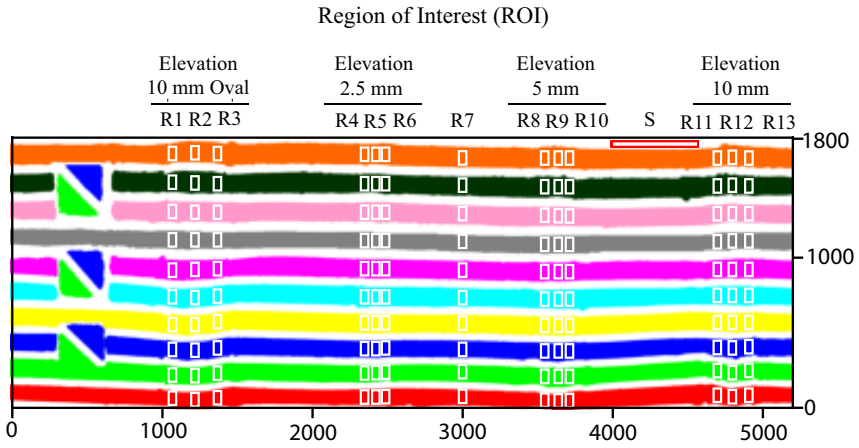


Figure 8: The ground truth image of the mockup showing regions of interest (ROIs) used for accuracy assessment corresponds to different elevation labels. R7 represents the flat region, while S represents the region for a substrate.

5 Results and Discussion

This section presents the classification results obtained, along with an evaluation of the overall accuracy of the algorithms used, considering various reference spectra (ground truth). Furthermore, the outcomes of the SVM model employing data augmentation will also be elaborated.

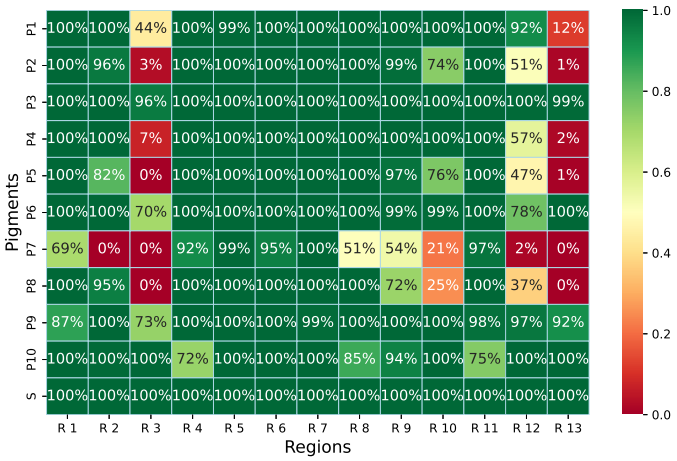
	Flat Region										Elevation : 2.5 mm										Elevation : 5 mm										Elevation : 10 mm									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
ED	1	0.9	1	1	1	0.1	0.7	1	0.3	0.8	0.8	0.7	1	0.8	0.9	0.3	0.3	0.8	0.6	0.3	0.5	0.4	0.8	0.4	0.8	0.2	0.2	0.5	0.4	0.2	0.5	0.3	0.5	0.4	0.5	0.2	0.1	0.3	0.2	0.5
	77.29%										**66.24%**										**44.48%**										**35.55%**									
SAM	1	1	1	1	1	0.2	1	1	0.6	1	0.9	1	0.9	1	0.9	0.7	0.9	1	0.9	0.5	0.9	1	0.8	1	0.7	0.6	0.6	1	0.9	0.3	1	1	0.9	1	0.9	0.9	0.5	1	0.7	0.5
	88.05%										**87.92%**										**77.96%**										**84.65%**									
SCM	1	1	1	1	1	0.2	1	1	0.7	0.9	1	1	1	1	0.9	0.7	0.9	1	0.9	0.5	1	1	1	1	0.6	0.5	0.6	1	0.9	0.3	1	1	1	1	0.9	0.8	0.5	1	0.8	0.4
	88.08%										**89.89%**										**79.33%**										**83.91%**									
SID	0.6	1	0.7	1	1	0.2	1	1	0.6	1	0.3	0.9	0.5	1	0.9	0.7	1	1	0.9	0.6	0.1	0.9	0.6	1	0.7	0.5	0.6	1	0.9	0.4	0.1	1	0.7	1	0.9	0.8	0.5	1	0.8	0.6
	80.86%										**76.77%**										**66.73%**										**72.87%**									
SID-SAM	0.7	1	0.6	1	1	0.2	1	1	0.6	1	0.3	0.9	0.4	1	0.9	0.7	0.9	1	0.9	0.6	0.1	0.9	0.5	1	0.7	0.5	0.6	1	0.9	0.3	0.1	1	0.6	1	0.9	0.8	0.5	1	0.8	0.6
	79.98%										**76.09%**										**66.01%**										**72.22%**									
SID-SCM	0.7	1	0.5	1	1	0.2	1	1	0.6	0.9	0.4	1	0.4	1	0.7	0.7	0.9	1	0.9	0.3	0.2	1	0.6	1	0.4	0.5	0.6	1	0.9	0.1	0.2	1	0.6	1	0.7	0.8	0.5	1	0.8	0.2
	79.86%										**74.16%**										**62.59%**										**67.68%**									
JMSAM	1	0.6	1	1	1	0.1	0.7	1	0.3	0.8	0.9	0.6	1	1	0.6	0.4	0.4	1	0.7	0.3	0.7	0.4	1	1	0.4	0.3	0.5	1	0.7	0.2	0.8	0.5	0.9	1	0.3	0.6	0.3	0.9	0.7	0.3
	75.38%										**68.84%**										**61.39%**										**63.20%**									
SSS	1	0.9	1	1	1	0.1	0.8	1	0.3	1	0.9	0.7	1	0.9	1	0.4	0.4	1	0.6	0.5	0.8	0.4	1	0.5	0.8	0.2	0.3	0.5	0.4	0.3	0.5	0.3	0.8	0.4	1	0.2	0.1	0.4	0.3	0.6
	80.03%										**74.05%**										**52.30%**										**44.86%**									
SVM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	0.9	1	1	0.9	1	0.4	0.7	1	0.9	0.7	0.5	1	0.5	0.5	0.9	0.3	0.5	1	0.9
	99.89%										**98.60%**										**88.23%**										**67.88%**									
1D-CNN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	0.9	1	1	1	1	1	1	0.6	1	1	0.8	1	1	1	1	1	1	0.4	1	1	0.9
	99.91%										**98.35%**										**93.48%**										**93.38%**									

Figure 9: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; the reference spectrum from the flat area was used for supervised and ML models; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

Figure 9 illustrates the classification accuracy of ten different algorithms, including two machine learning models, for each pigment on both elevated surfaces and a flat surface. Algorithms 1D-CNN and SVM, perform better than all eight supervised algorithms, with the least accurate being algorithm ED. SAM and SCM performed better than the other algorithms after the machine learning models. On the other hand algorithms with hybrid approaches did not performed well overall. The images showing the classification results for each algorithm are attached in Appendix A.

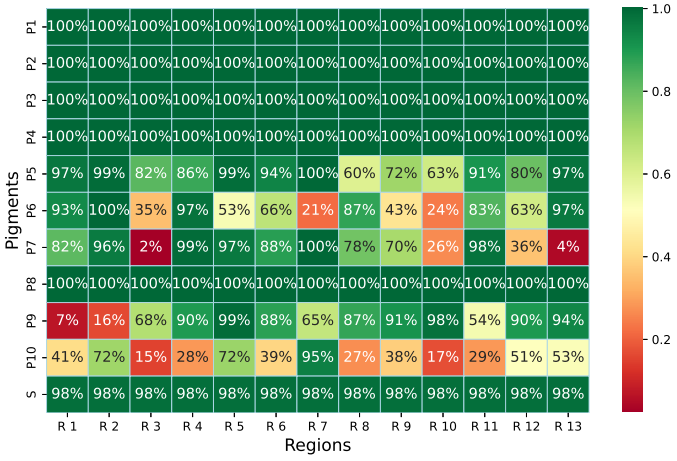
Although there are variations in accuracy among the different algorithms, we can discern a pattern in the obtained classification results; the classification accuracy decreases with increasing elevation. An interesting observation is that SVM has lower accuracy for a region with an elevation of 10mm than most other algorithms for that same region. Figure 10a also presents the confusion matrix for SVM, revealing that the accuracy is particularly low for regions R3, R13, and R10, all located on one particular side of the elevation in the mockup (Figure 8). This side of the mockup has a shadow (Figure 2a), with the lowest accuracy being on the shadow side of the highest elevated regions (Figure 2b), which could be a significant factor in misclassifying these regions. Abed [31] stated in his work that surface changes become problematic when it comes to material aspects, which our results confirm.

Appendix B contains the confusion matrix for regions three and thirteen, illustrating the misclassification of pigments for SVM. The SAM algorithm determines the angle between two vectors, irrespective of their length, and thus its classification accuracy is less affected by any changes in the spectrum's magnitude. On the other hand, the SCM algorithm eliminates negative correlations while preserving the SAM characteristics. Therefore, these algorithms provide greater accuracy in the shadow area, as evidenced by the confusion matrix depicted in Figure 10b and Figure 10c for SAM and SCM, respectively.



(a)

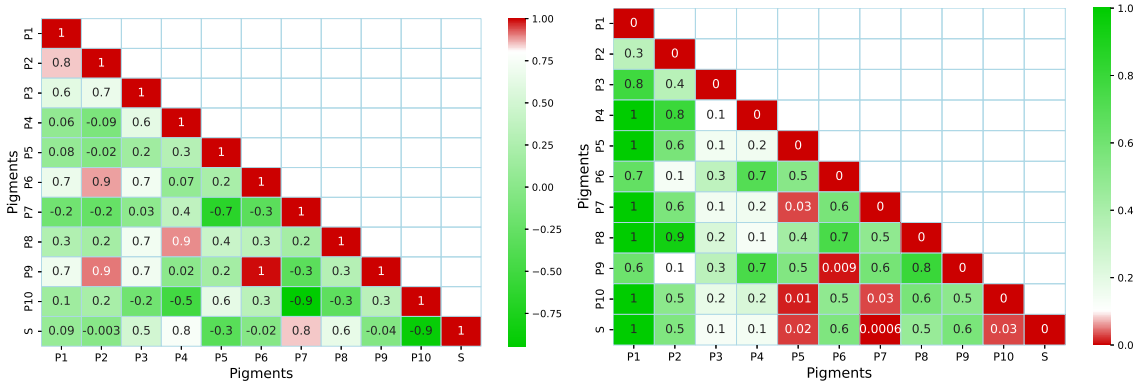
(b)



(c)

Figure 10: Classification accuracy for each pigment across all Regions of Interest (ROIs); For algorithms, (a): SVM, (b): SAM, and (c): SCM.

The accuracy of all supervised algorithms is lower for pigment P6 and P9 in flat regions, but this is only due to the exact location chosen for accuracy assessment (R7), which includes most of the unclassified areas. This trend is not necessarily representative of all flat regions on average; also, elevated surfaces do not exhibit the same level of lower accuracy as flat regions. The classification results are presented in the Appendix A. However, the ED, JMSAM, and SSS algorithms have lower accuracy for pigments P6 and P9 due to their similar spectra, which are also discussed by authors in [27]. This similarity is evident in Figure 11a, where the Pearson correlation coefficient between these two pigments is almost 1. Similarly, Figure 11b also indicates a very low dissimilarity measure between these two pigments.

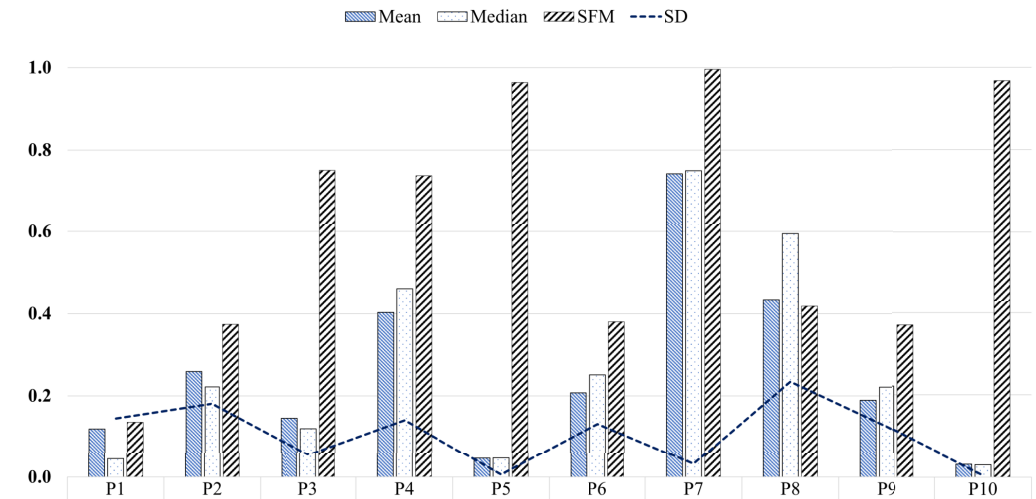


(a) (b)

Figure 11: The matrices display the correlation and dissimilarity between the pigments; (a) Pearson's correlation coefficient [83], a coefficient of 1 indicates a high correlation, while a value of zero represents no correlation; (b) SID calculated for dissimilarity, where a value of 0 implies a high degree of similarity between the spectra, while 1 indicates maximum dissimilarity.

The accuracy of the algorithms SID and its hybrid combination with SAM and SCM for pigment P1 has decreased as shown in Figure 9, and this decrease was further observed with an increase in elevation. It is important to note that not all pigments are affected

in the same way. In the case of P1, the pixels were either classified as P1 or remained unclassified, as indicated by the confusion matrix for SID in the Appendix C. Mandal et al. [27] suggested that changing the threshold value could improve classification accuracy, but this approach could lead to misclassification of other pigments, which is generally undesirable in CH. The stochastic algorithm, SID, depends on the probability distributions of spectra, and alterations in data distribution can influence the overall entropy value. Moreover, the changes in the dataset can produce a varying impact on the entropy value for normal distribution and skewed distribution. In other words, if both distributions are shifted equally, the symmetric distribution would experience less change in entropy than the skewed distribution, owing to its higher predictability and lower uncertainty compared to the skewed distribution.



	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Mean	0.1192	0.2594	0.1454	0.4033	0.0462	0.2074	0.7422	0.4332	0.1894	0.0319
Median	0.0451	0.2220	0.1200	0.4601	0.0470	0.2511	0.7498	0.5944	0.2214	0.0306
SFM	0.136	0.374	0.751	0.738	0.964	0.38	0.996	0.418	0.37	0.968
SD	0.1448	0.1810	0.0552	0.1403	0.0063	0.1313	0.0322	0.2344	0.1243	0.0044

Figure 12: Statistical measures of mean, median, standard deviation (SD), and Spectral flatness measure (SFM), computed for all ten pigments.

As illustrated in Figure 12, pigment P1 exhibits a notable difference between its mean and

median values, resulting in a skewed distribution. Additionally, the standard deviation of this distribution is greater than its mean, indicating that the data points are more widely dispersed. Given the same amount of shift in datasets (identical in absolute terms), this shift may have a more significant impact on the dataset with a larger spread than the one with data points being more tightly clustered around the mean, leading to a larger relative entropy between the two datasets. This is likely why some pigments are affected more than others. This also explains why we need to set different threshold values for different pigments.

The accuracy of Pigment P10 is higher for all the algorithms used on a flat surface, but it decreases significantly for supervised algorithms with elevation changes. On the other hand, ML models, SVM, and 1D-CNN show consistent classification accuracy for P10, regardless of elevation. When the reference spectrum is obtained from a flat region, it is likely to have minimal variation with test data sets within the same region, resulting in a higher number of correctly classified pixels. However, if the reference spectrum is taken from an elevated region, changes in reflectance values at higher elevations could cause higher variation and lead to more unclassified pixels. Based on this, one can hypothesize that using a reference spectrum taken from an elevated region is more likely to result in correct pixel classification on the same elevated surface region.

Low spectral intensity makes distinguishing between different land cover classes or features in an image difficult, leading to misclassification, especially for classes with similar low reflectance values [84]. We observed that, in most cases, the P10 and P5 pigments were misclassified as each other or as the substrate, or they remained unclassified in most of the algorithms used. The Spectral Flatness Measure (SFM), also known as Wiener Entropy, is a metric that can be used to quantify the degree of flatness or peakiness of a spectrum by computing the geometric mean ratio to the arithmetic mean of the power spectrum [85].

We computed the SFM for the pigments, as shown in Figure 12. The SFM values for P7, P10, and P5 were higher, indicating a flat spectrum with nearly identical reflectance values across different wavelengths. This can make it difficult for a classifier algorithm to differentiate between different classes, resulting in lower classification accuracy.

Using a reference target with a surface height equivalent to the flat surface of the mockup, the normalized reflectance value was calculated for the HSI datacube. Ideally, the reflectance value of an elevated surface should be higher than that of a flat surface. However, shadows caused by the surface elevation lead to a decrease in the obtained reflectance value as the surface height increases. Generally, the substrate spectrum has a lower value than the P7 reference spectrum (the reference spectrum for ten pigments used and the substrate is provided in the Appendix D). With an increase in surface elevation, the distance between the P7 and its reference spectrum increases, while the distance between the P7 and substrate reference spectrum decreases. As a result, most regions for P7 are misclassified as substrate.

In Figure 13, we can see the overall classification accuracy obtained by averaging all pigments for four different surface elevations: flat, 2.5mm, 5mm, and 10mm, using all ten algorithms in the study. The reference spectrum used for building the spectral library and training the ML models was taken from the flat surface. For almost all algorithms, the classification accuracy for an elevated surface is lower compared to the flat region. Surfaces with a 2.5mm elevation have accuracy similar to or less than the flat surface, followed by 5mm and 10mm, respectively. This pattern is consistent for ED, SSS, SVM, and 1D-CNN algorithms. However, for the other six algorithms, we see that the accuracy at an elevation of 10mm is slightly greater than that of the surface with a 5mm elevation. The increase in overall classification accuracy can be attributed to the higher accuracy obtained for pigments P5 and P10. This higher accuracy for P5 and P10 might be due to the shadow effect;

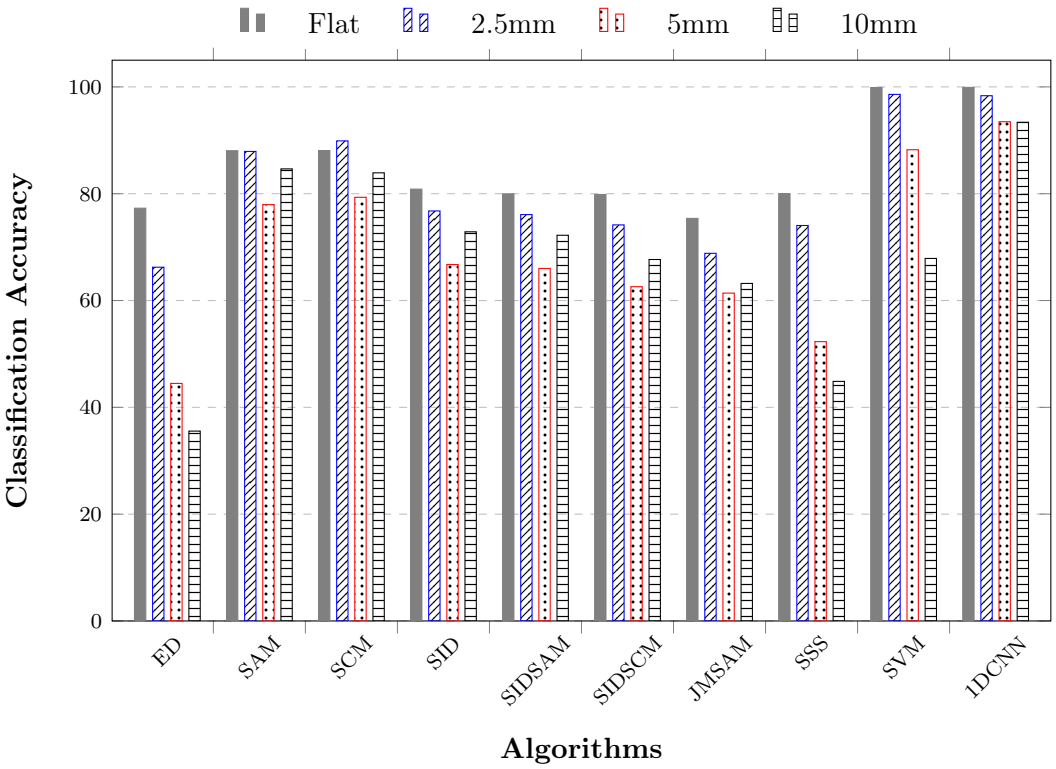


Figure 13: Overall classification accuracy for four different surface elevations (flat, 2.5mm, 5mm, and 10mm) using reference spectra from the flat surface.

decreased reflectance value in the shadow region might have reduced the distance between the reference and measured pixels.

Earlier in this paper, we hypothesized that using a reference spectrum from an elevated surface would result in more accurate pixel classification for that same elevated surface region. To test this hypothesis, we built spectral libraries and training datasets using the reference spectrum from each elevated surface, i.e., 2.5mm, 5mm, and 10mm and then computed the classification accuracy. In addition, we built spectral libraries and training datasets that represent an average spectrum of flat and varying elevation surfaces. To obtain these spectra, we averaged the pixel values over a spatial region taken as in a straight line

(for example, in Figure 8, it is a pixels values of a line from right to left) for each band. The result from classifaction accuracy is summarized in Figure 14.

	Ref. Spectrum: Flat				Ref. Spectrum: 2.5mm				Ref. Spectrum: 5mm				Ref. Spectrum: 10mm				Ref. Spectrum: Avg			
	F	E1	E2	E3	F	E1	E2	E3	F	E1	E2	E3	F	E1	E2	E3	F	E1	E2	E3
ED	0.77	0.66	0.44	0.36	0.79	0.75	0.51	0.42	0.71	0.7	0.48	0.42	0.53	0.58	0.53	0.44	0.82	0.7	0.48	0.41
	55.89%				**61.58%**				**57.45%**				**51.92%**				**60.26%**			
SAM	0.88	0.88	0.78	0.85	0.85	0.88	0.81	0.84	1	0.97	0.92	0.93	0.99	0.93	0.9	0.92	1	0.98	0.93	0.92
	84.65%				**84.52%**				**95.31%**				**93.50%**				**95.75%**			
SCM	0.88	0.9	0.79	0.84	0.86	0.87	0.78	0.82	0.91	0.95	0.91	0.92	0.98	0.93	0.89	0.91	0.99	0.96	0.88	0.91
	85.30%				**83.24%**				**92.43%**				**92.68%**				**93.59%**			
SID	0.81	0.77	0.67	0.73	0.75	0.81	0.73	0.72	0.85	0.88	0.8	0.82	0.84	0.83	0.8	0.86	0.95	0.92	0.83	0.88
	74.31%				**75.36%**				**83.88%**				**83.31%**				**89.44%**			
SID-SAM	0.8	0.76	0.66	0.72	0.75	0.8	0.72	0.71	0.83	0.88	0.79	0.82	0.79	0.81	0.78	0.84	0.94	0.92	0.83	0.87
	73.58%				**74.28%**				**82.83%**				**80.63%**				**89.08%**			
SID-SCM	0.8	0.74	0.63	0.68	0.74	0.77	0.66	0.68	0.74	0.84	0.8	0.8	0.74	0.79	0.78	0.83	0.89	0.89	0.79	0.86
	71.07%				**71.15%**				**79.72%**				**78.50%**				**85.71%**			
JMSAM	0.75	0.69	0.61	0.63	0.76	0.74	0.63	0.54	0.67	0.69	0.66	0.59	0.73	0.69	0.67	0.68	0.85	0.72	0.66	0.68
	67.20%				**66.58%**				**65.19%**				**69.14%**				**72.76%**			
SSS	0.8	0.74	0.52	0.45	0.83	0.84	0.59	0.57	0.73	0.76	0.59	0.49	0.7	0.67	0.62	0.58	0.91	0.78	0.57	0.56
	62.81%				**70.86%**				**64.26%**				**64.43%**				**70.78%**			
SVM	1	0.99	0.88	0.68	0.99	1	0.97	0.76	0.99	1	1	0.93	0.95	0.97	0.97	0.99	1	1	0.99	1
	88.65%				**93.12%**				**97.89%**				**97.24%**				**99.70%**			
1D-CNN	1	0.98	0.93	0.93	1	1	0.98	0.97	1	1	0.99	0.99	1	0.99	0.99	1	1	1	0.99	1
	96.28%				**98.62%**				**99.42%**				**99.35%**				**99.49%**			

Figure 14: Classifaction accuracy obtained using ten different algorithms for flat (F), 2.5 mm elevation (E1), 5 mm elevation(E2), and 10mm eleavation (E3), each computed for different spectral library and training datasets, built using reference spectrum from the flat region, three different elevated region and an average of all these.

Figure 14 shows that classification accuracy for the flat region (F) is higher when using reference spectra from the flat surface than at different elevations. Similarly, when using reference spectra from a 2.5mm elevated region, the classification accuracy is higher for regions with 2.5mm elevation (E1) for all algorithms. However, for regions with 5mm elevation (E2), the classification accuracy is almost identical to E1 and very close to other regions. In contrast, for regions with 10mm elevation (E3), the classification accuracy

is greater than other surfaces, mainly for stochastic algorithms (SID, SIDSAM, SIDSCM), SVM, and 1D-CNN. However, it is the lowest for ED and SSS. The accuracy of SAM, SCM, and JMSAM is similar to E1 and E2 elevations but lower than the flat surface. Using average reference spectra improved accuracy for almost all flat and elevated regions compared to the accuracy obtained when using four different conditions of reference spectra. A detailed result displaying the classification accuracy for individual pigments at each elevation and for different reference spectra conditions can be found in the Appendix E.

Result for average classification accuracy using ten algorithms and five reference spectra conditions (i.e., from the regions F, E1, E2, E3, and the average of these regions) is shown in Figure 15. When the reference spectrum was taken from the flat region, the classification accuracy was lower, or comparable (in the case of E1), for all algorithms compared to the accuracy obtained when the reference spectra were taken from regions E2, E3, or the average spectrum. The classification accuracy for most algorithms was almost the same when considering reference spectra from E2 or E3. Notably, using an average reference spectrum improved classification accuracy for almost all algorithms. The 1D-CNN algorithm had the highest classification accuracy among all the algorithms used, with slightly lower accuracy when the reference spectrum was taken from the flat region and almost similar results for all other reference spectra conditions.The classified images produced by the SVM and 1D-CNN algorithms using an average reference spectrum are provided in the Appendix F.

Data augmentation was performed to create additional training datasets for SVM, as shown in the flow diagram illustrated in Figure 16. Before computing classification accuracy, Hyperparameter tuning was performed using these augmented datasets. Finally, the SVM model was executed with the optimal hyperparameters of a polynomial kernel function and

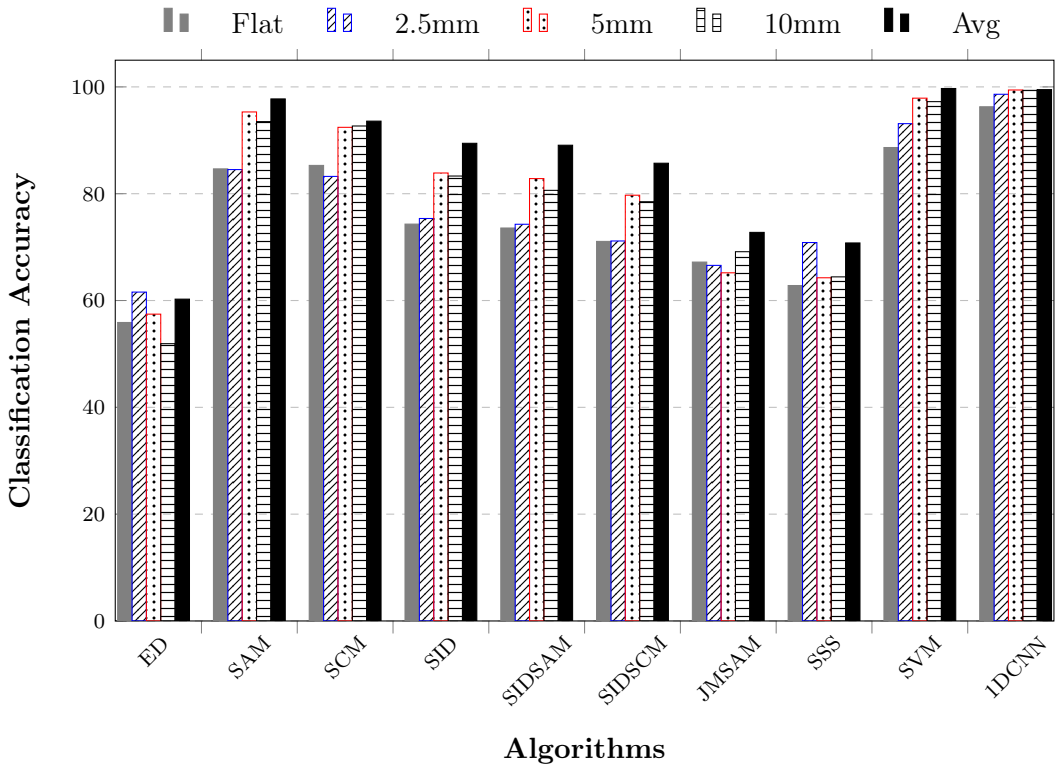


Figure 15: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 2.5 mm.

a regularization value of 0.1. Figure 17 shows the classification accuracy obtained for each pigment across all regions using the augmented training dataset where training data set from a flat region was selected. We observed that overall classification accuracy was higher than the SVM without data augmentation (Figure 10a). We also augmented the data from a single spectrum taken from the flat region first and then from the average spectrum; results for classification accuracy are included in Appendix G. Figure 18 shows the overall classification accuracy for each condition for flat and three different elevations.

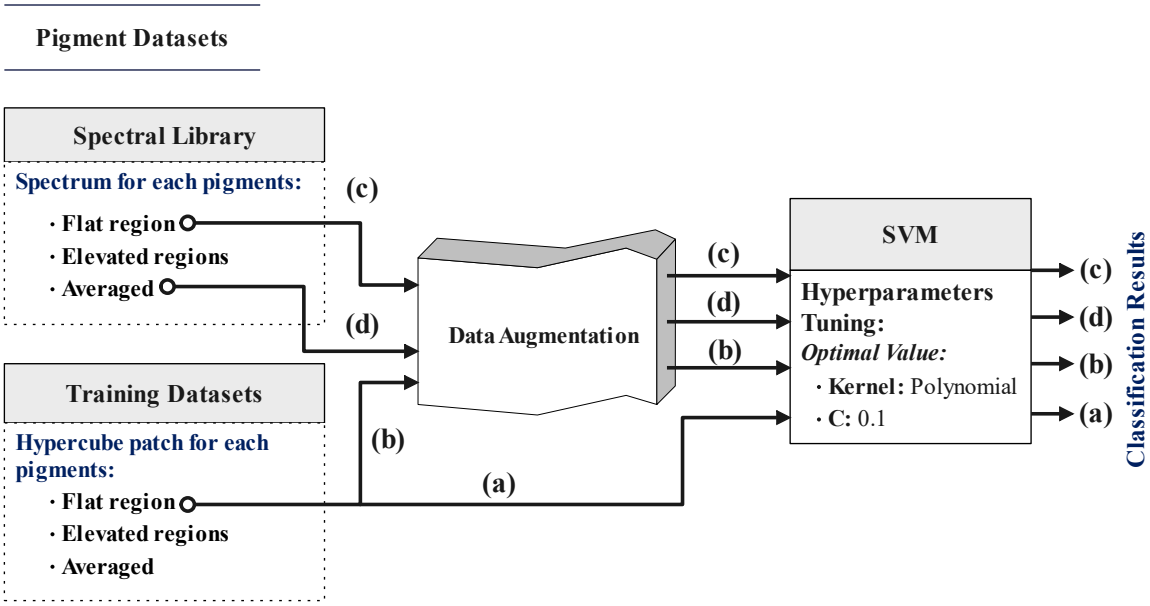


Figure 16: A workflow diagram illustrating the generation of training datasets using data augmentation for SVM; classification results from a, b, c and d are illustrated in Figure 19.

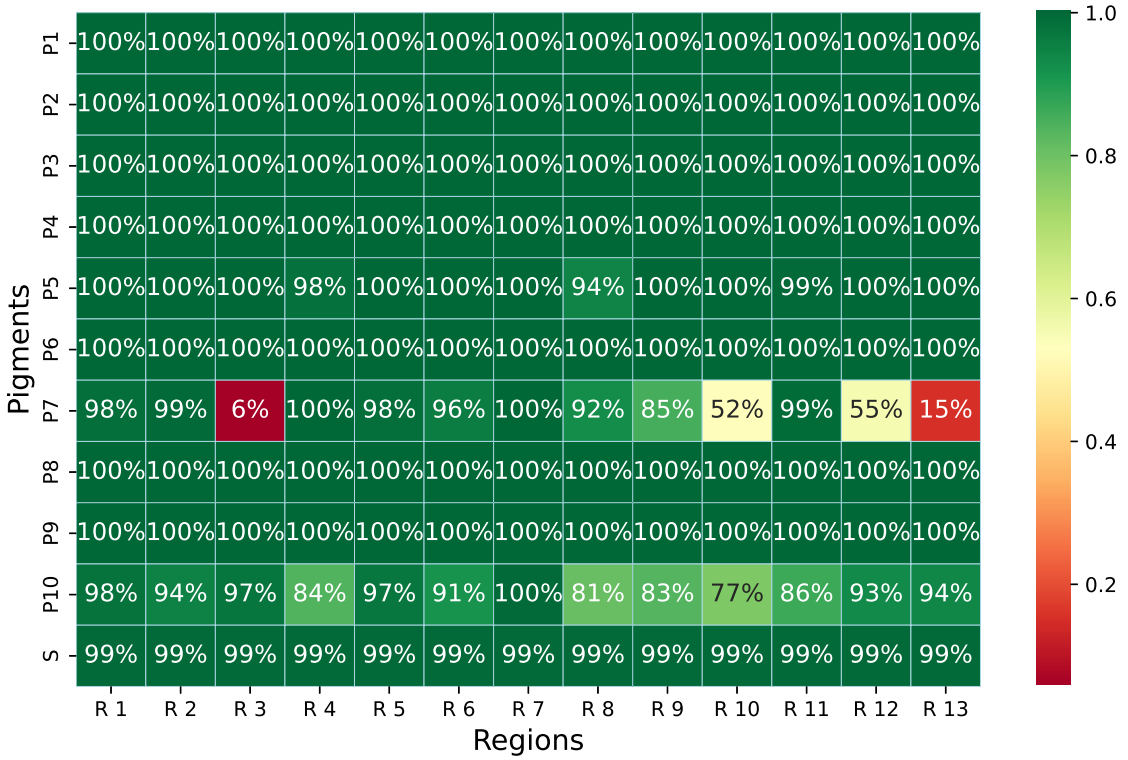


Figure 17: Classification accuracy for each pigment across all Regions of Interest (ROIs) for SVM using augmented dataset.

Classification accuracy for an augmented data set is higher when compared with all other three conditions, i.e., without data augmentation, data augmentation considering a single spectrum from a flat region and considering data augmentation considering a single average spectrum. However, using a single spectrum for augmentation yielded lower performance, except for an elevated region of 10 mm, where the accuracy was significantly improved compared to the non-augmented condition. Comparable accuracy was achieved using data augmentation with a flat region and augmenting a single averaged spectrum. As illustrated in Figure 9, SAM and SCM are the algorithms which performed better after ML models. Figure 18 shows that data augmentation using a single spectrum from a flat region is still

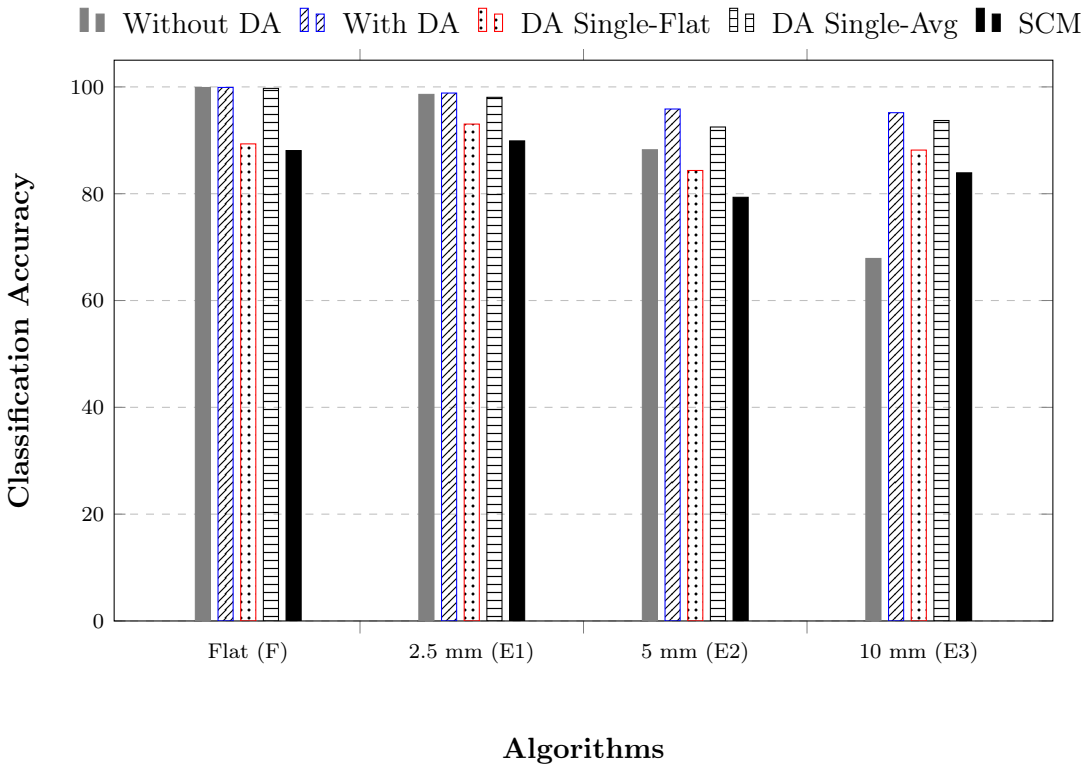


Figure 18: Overall classification accuracy for SVM at four different surface elevations (flat, 2.5 mm, 5 mm, and 10 mm) using different conditions of data augmentation, in comparison with SCM, employing reference spectra obtained from the flat region.

better than SCM, highlighting that even with the single spectrum available, one can obtain better classification accuracy than supervised-based algorithms. The results suggest that data augmentation can improve classification accuracy, particularly when multiple spectra are augmented or when an average spectrum is used for augmentation. The classified images for each of these conditions are presented in Figure 19.

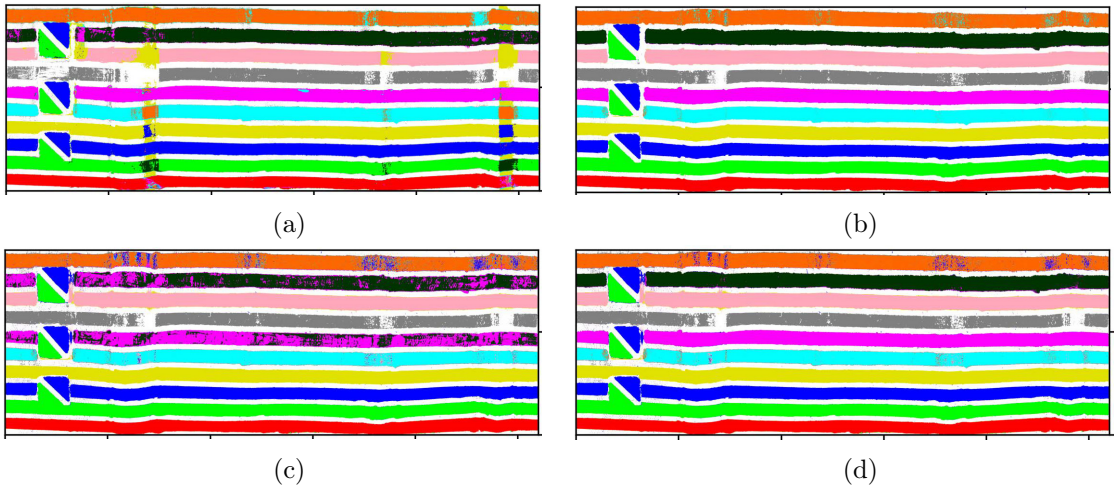


Figure 19: The classification results for SVM; (a): without data augmentation, (b): with data augmentation, (c): with data augmentation using single spectrum from flat region, and (d):with data augmentation using single spectrum from averaged region.

6 Conclusion

This study investigated the performance of different supervised algorithms and machine learning models for pigment classification using HSI on an elevated mockup. We have observed that the elevation itself does not significantly impact the classification accuracy; however, the elevation can result in the formation of shadows, which can have a significant effect on the classification accuracy of the algorithms used and varies for different algorithms. It was also observed that the choice of reference spectra plays a significant role in the accuracy of pigment classification. An average reference spectrum from different elevated regions yields better results than individual spectra. Among the ten algorithms tested, the 1D-CNN algorithm showed the highest classification accuracy, followed by SVM, SAM and SCM. Furthermore, results also indicated that data augmentation could significantly improve classification accuracy, particularly when multiple spectra are augmented or when

an average spectrum is used for augmentation. This study provides valuable insights for analysing paintings in the CH domain. It could be beneficial in selecting appropriate classification algorithms when artworks have elevation or data that have shadows. In addition to paintings, it could also be useful to other museum low elevations objects, such as glazed plates [86] and bas-reliefs [87]. In future work, removing shadows using image-processing techniques could be explored as a possible direction to further improve the accuracy of pigment classification. Future work will also include mixtures of pigments as well as the influence of the thickness of the pigment.

References

[1] Dusan Stulik, David Miller, Herant Khanjian, Narayan Khandekar, Janice Carlson, Richard Wolbers, and W Christian Petersen. *Solvent gels for the cleaning of works of art: the residue question*. Getty Publications, 2004.

[2] Shawn Digney-Peer, Karen Thomas, Roy Perry, Joyce Townsend, and Stephen Gritt. The imitative retouching of easel paintings. In *Conservation of easel paintings*, pages 626–653. Routledge, 2020.

[3] Raju Shrestha and Jon Yngve Hardeberg. Evaluation and comparison of multispectral imaging systems. In *Color and Imaging Conference*, volume 2014, pages 107–112. Society for Imaging Science and Technology, 2014.

[4] Costanza Cucci and Andrea Casini. Hyperspectral imaging for artworks investigation. In *Data handling in science and technology*, volume 32, pages 583–604. Elsevier, 2019.

[5] Hilda Deborah, Sony George, and Jon Yngve Hardeberg. Pigment mapping of the Scream (1893) based on hyperspectral imaging. In *Image and Signal Processing: 6th*

International Conference, ICISP 2014, Cherbourg, France, June 30–July 2, 2014. Proceedings 6, pages 247–256. Springer, 2014.

[6] Koen Janssens, Geert Van der Snickt, Frederik Vanmeert, Stijn Legrand, Gert Nuyts, Matthias Alfeld, Letizia Monico, Willemien Anaf, Wout De Nolf, Marc Vermeulen, et al. Non-invasive and non-destructive examination of artistic pigments, paints, and paintings by means of X-ray methods. In *Analytical Chemistry for Cultural Heritage*, pages 77–128. Springer, 2017.

[7] Freek D. van der Meer, Harald M.A. van der Werff, Frank J.A. van Ruitenbeek, Chris A. Hecker, Wim H. Bakker, Marleen F. Noomen, Mark van der Meijde, E. John M. Carranza, J. Boudewijn de Smeth, and Tsehaie Woldai. Multi- and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14(1):112–128, 2012.

[8] Laura M Dale, André Thewis, Christelle Boudry, Ioan Rotar, Pierre Dardenne, Vincent Baeten, and Juan A Fernández Pierna. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Applied Spectroscopy Reviews*, 48(2):142–159, 2013.

[9] Guolan Lu and Baowei Fei. Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1):1 – 24, 2014.

[10] Edelman, Gerda and Gaston, Edurne and van Leeuwen, Ton and Cullen, P.J and Aalders, M.C.G. Hyperspectral imaging for non-contact analysis of forensic traces. *Forensic science international*, 223:28–39, 10 2012.

[11] Qingli Li, Xiaofu He, Yiting Wang, Hongying Liu, Dongrong Xu, and Fangmin Guo. Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of Biomedical Optics*, 18(10):1 – 29, 2013.

[12] Christian Fischer and Ioanna Kakoulli. Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications. *Studies in Conservation*, 51(sup1):3–16, 2006.

[13] Yoko Arteaga, JY Hardeberg, and C Boust. Hdr multispectral imaging-based brdf measurement using flexible robot arm system. In *Proceedings of the 30th Color and Imaging Conference, Scottsdale. Society for Imaging Science and Technology*, 2022.

[14] Camille Simon Chane, Alamin Mansouri, Franck S Marzani, and Frank Boochs. Integration of 3d and multispectral data for cultural heritage applications: Survey and perspectives. *Image and Vision Computing*, 31(1):91–102, 2013.

[15] W Stanley Taft and James W Mayer. The science of paintings. *Measurement Science and Technology*, 12(5):653–653, 2001.

[16] Peter Paul Rubens. *The Virgin as the Woman of the Apocalypse*. J. Paul Getty Museum. about 1623–1624.

[17] Joyce H Townsend. Painting techniques and materials of Turner and other British artists 1775-1875. In *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*, pages 176–185, 1995.

[18] Sylvana Barrett and Dusan C Stulik. An integrated approach for the study of painting techniques. In *Historical painting techniques, materials, and studio practice: preprints*

of a symposium, University of Leiden, the Netherlands, 26-29 June 1995, pages 6–11, 1995.

[19] Eddie Sinclair. The polychromy of Exeter and Salisbury Cathedrals: a preliminary comparison. In *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*, pages 105–110. Getty Conservation Institute Malibu, 1995.

[20] William Baxter, Jeremy Wendt, and Ming C Lin. IMPaSTo: A realistic, interactive model for paint. In *Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering*, pages 45–148, 2004.

[21] Yunfei Fu, Hongchuan Yu, Chih-Kuo Yeh, Jianjun Zhang, and Tong-Yee Lee. High relief from brush painting. *IEEE transactions on visualization and computer graphics*, 25(9):2763–2776, 2018.

[22] J Salvant Plisson, Laurence de Viguerie, L Tahroucht, M Menu, and G Ducouret. Rheology of white paints: How Van Gogh achieved his famous impasto. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 458:134–141, 2014.

[23] Karin Groen. Investigation of the use of the binding medium by Rembrandt. *Zeitschrift fur Kunsttechnologie und Konservierung*, 2(2):208–211, 1997.

[24] W.S Elkhuizen, T.W.J Dore-Callewaert, Emilien Leonhardt, Abbie Vandivere, Y Song, S.C Pont, Jo M.P Geraedts, and J Dik. Comparison of three 3D scanning techniques for paintings, as applied to Vermeer's 'Girl with a Pearl Earring'. *Heritage science*, 7(1):1–22, 2019.

[25] Victor Gonzalez, Marine Cotte, Gilles Wallez, Annelies van Loon, Wout De Nolf, Myriam Eveno, Katrien Keune, Petria Noble, and Joris Dik. Unraveling the Composition

of Rembrandt's Impasto through the Identification of Unusual Plumbonacrite by Multimodal X-ray Diffraction Analysis. *Angewandte Chemie*, 131(17):5675–5678, 2019.

[26] Dipendra J Mandal, Sony George, Marius Pedersen, and Clotilde Boust. Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging. In *Color and Imaging Conference*, volume 2021, pages 334–346. Society for Imaging Science and Technology, 2021.

[27] Dipendra J. Mandal, Marius Pedersen, Sony George, Hilda Deborah, and Clotilde Boust. An Experiment-based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging. *Journal of Imaging Science and Technology*, 67(3):030403–1–030403–1, 2023.

[28] Yonghui Zhao, Roy S Berns, Lawrence A Taplin, and James Coddington. An investigation of multispectral imaging for the mapping of pigments in paintings. In *Computer image analysis in the study of art*, volume 6810, pages 65–73. SPIE, 2008.

[29] Roy S Berns and Francisco H. Imai. The use of multi-channel visible spectrum imaging for pigment identification. In *Proc. 13th triennal ICOM-CC meeting*, pages 217–222, 2002.

[30] Roy S Berns and Francisco H Imai. Pigment identification of artist materials via multispectral imaging. In *Color and Imaging Conference*, volume 2001, pages 85–90. Society for Imaging Science and Technology, 2001.

[31] Farhad Moghareh Abed. *Pigment identification of paintings based on Kubelka-Munk theory and spectral images.* Rochester Institute of Technology, 2014.

[32] Kensuke Fukumoto, Norimichi Tsumura, and Roy Berns. Estimating concentrations of pigments using encoder-decoder type of neural network. In *Color and Imaging*

Conference, volume 2019, pages 149–152. Society for Imaging Science and Technology, 2019.

[33] Kensuke Fukumoto, Norimichi Tsumura, and Roy Berns. Estimating pigment concentrations from spectral images using an encoder decoder neural network. *Journal of Imaging Science & Technology*, 64(3), 2020.

[34] Marco Fiorucci, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, and Stuart James. Machine learning for cultural heritage: A survey. *Pattern Recognition Letters*, 133:102–108, 2020.

[35] Stanislav Smirnov and Alma Eguizabal. Deep learning for object detection in fine-art paintings. In *2018 Metrology for Archaeology and Cultural Heritage (MetroArchaeo)*, pages 45–49. IEEE, 2018.

[36] Costas Balas, George Epitropou, Athanasios Tsapras, and Nicos Hadjinicolaou. Hyperspectral imaging and spectral classification for pigment identification and mapping in paintings by El Greco and his workshop. *Multimedia Tools and Applications*, 77(8):9737–9751, 2018.

[37] Sennaraj Vishnu, Rama Rao Nidamanuri, and R Bremananth. Spectral material mapping using hyperspectral imagery: a review of spectral matching and library search methods. *Geocarto international*, 28(2):171–190, 2013.

[38] Nirmal Keshava. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Transactions on Geoscience and remote sensing*, 42(7):1552–1565, 2004.

[39] Jianfeng Li, D Brynn Hibbert, Steven Fuller, Julie Cattle, and Christopher Pang Way. Comparison of spectra using a Bayesian approach. An argument using oil spills as an example. *Analytical chemistry*, 77(2):639–644, 2005.

[40] Daniele Cerra, Jakub Bieniarz, Janja Avbelj, Rupert Müller, and Peter Reinartz. Spectral Matching through Data Compression. volume Volume, pages 1–4, 01 2011.

[41] GONG Mengting and FENG Pingli. Preliminary study on the application of hyperspectral imaging in the classification of and identification Chinese traditional pigments classification——a case study of spectral angle mapper. *Sciences of conservation and archaeology*, (4):76–83, 2014.

[42] O Abilio De Carvalho and Paulo Roberto Meneses. Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM). In *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18*, volume 9. JPL Publication Pasadena, CA, 2000.

[43] Sony George and Jon Yngve Hardeberg. Ink classification and visualisation of historical manuscripts: Application of hyperspectral imaging. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pages 1131–1135, 2015.

[44] Dilish Deepthi, Binu Melit Devassy, Sony George, Peter Nussbaum, and Tessamma Thomas. Classification of forensic hyperspectral paper data using hybrid spectral similarity algorithms. *Journal of Chemometrics*, 36(1):e3387, 2022.

[45] Cheng Fan, Pengchang Zhang, Shuang Wang, and Bingliang Hu. A study on classification of mineral pigments based on spectral angle mapper and decision tree. In *Tenth International Conference on Digital Image Processing (ICDIP 2018)*, volume 10806, pages 1639–1643. SPIE, 2018.

[46] Mingyan Gu, Shuqiang Lyu, Miaole Hou, Sheng Ma, Zhenhua Gao, Shuzhang Bai, and Pingping Zhou. Classification and recognition of tomb information in hyperspectral image. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives*, 42(3):411–416, 2018.

[47] Tania Kleynhans, Catherine M Schmidt Patterson, Kathryn A Dooley, David W Messinger, and John K Delaney. An alternative approach to mapping pigments in paintings with hyperspectral reflectance image cubes using artificial intelligence. *Heritage Science*, 8(1):1–16, 2020.

[48] Lingxi Liu, Tsveta Miteva, Giovanni Delnevo, Silvia Mirri, Philippe Walter, Laurence de Viguerie, and Emeline Pouyet. Neural Networks for Hyperspectral Imaging of Historical Paintings: A Practical Review. *Sensors*, 23(5):2419, 2023.

[49] Ailin Chen, Rui Jesus, and Marcia Vilarigues. Convolutional Neural Network-Based Pure Paint Pigment Identification Using Hyperspectral Images. In *ACM Multimedia Asia*, pages 1–7. Article No.: 64, 2021.

[50] John Clifford Gower. Properties of Euclidean and non-Euclidean distance matrices. *Linear algebra and its applications*, 67:81–97, 1985.

[51] F.A. Kruse, A.B. Lefkoff, J.W. Boardman, K.B. Heidebrecht, A.T. Shapiro, P.J. Barloon, and A.F.H. Goetz. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2):145–163, 1993. Airbone Imaging Spectrometry.

[52] Osmar de Carvalho Júnior, Renato Guimarães, Alan Gillespie, Nilton Silva, and Roberto Gomes. A New Approach to Change Vector Analysis Using Distance and

Similarity Measures. *Remote Sensing, vol. 3, issue 11, pp. 2473-2493*, 3:2473–2493, 11 2011.

[53] Chein-I Chang. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory*, 46(5):1927–1932, 2000.

[54] John P Kerekes, Adam P Cisz, and Rulon E Simmons. A comparative evaluation of spectral quality metrics for hyperspectral imagery. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, volume 5806, pages 469–480. SPIE, 2005.

[55] Yingzi Du, Chein-I Chang, Hsuan Ren, Chein-Chi Chang, James O Jensen, and Francis M D'Amico. New hyperspectral discrimination measure for spectral characterization. *Optical engineering*, 43(8):1777–1786, 2004.

[56] M Naresh Kumar, MVR Seshasai, KS Vara Prasad, V Kamala, KV Ramana, RS Dwivedi, and PS Roy. A new hybrid spectral similarity measure for discrimination among Vigna species. *International journal of remote sensing*, 32(14):4041–4053, 2011.

[57] S. Padma and S. Sanjeevi. Jeffries matusita-spectral angle mapper (JM-SAM) spectral matching for species level mapping at Bhitarkanika, Muthupet and Pichavaram mangroves. volume 40, page 1403 – 1411, 2014.

[58] Lipo Wang. *Support vector machines: theory and applications*, volume 177. Springer Science & Business Media, 2005.

[59] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, 1992.

[60] William S Noble. What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567, 2006.

[61] Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, and Humaira Arshad. State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11):e00938, 2018.

[62] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[63] Shiqi Yu, Sen Jia, and Chunyan Xu. Convolutional neural networks for hyperspectral image classification. *Neurocomputing*, 219:88–98, 2017.

[64] Danfeng Hong, Lianru Gao, Jing Yao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7):5966–5978, 2021.

[65] Wei Hu, Yangyu Huang, Wei Li, Fan Zhang, and Hengchao Li. Deep Convolutional Neural Networks for Hyperspectral Image Classification. *Journal of Sensors*, 2015:1–12, 07 2015.

[66] ZECCHI, Soluzioni per artisti, materiale per restauro . http://www.zecchi.it/. Accessed: 01 May 2023.

[67] Norsk Elektro Optikk. http://www.hyspex.no/. Accessed: 01 May 2023.

[68] Spectralon multi-step targets. https://www.labsphere.com/product/spectralon-reflectance-targets/. Accessed: 1 May 2023.

[69] X-Rite ColorChecker Classic. https://www.xrite.com/. Accessed: 01 May 2023.

[70] David Saunders and John Cupitt. Image Processing at the National Gallery: The VASARI Project. 14:72–85, 1993.

[71] Python Software Foundation. Python Language Reference, Version 3.9. https://docs.python.org/3/reference/, 2021. [Accessed 05 December 2022].

[72] Li Tan et al. A study on the application of SAM classification algorithm in seal of calligraphy and painting based on hyperspectral technology. In *2016 4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, pages 415–418. IEEE, 2016.

[73] Hilde JP Weerts, Andreas C Mueller, and Joaquin Vanschoren. Importance of tuning hyperparameters of machine learning algorithms. *arXiv preprint arXiv:2007.07588*, 2020.

[74] Andreas C Müller and Sarah Guido. *Introduction to machine learning with Python: a guide for data scientists*. " O'Reilly Media, Inc.", 2016.

[75] Tom O'Malley, Elie Bursztein, James Long, François Chollet, Haifeng Jin, Luca Invernizzi, et al. KerasTuner. https://github.com/keras-team/keras-tuner, 2019. Accessed: 20 January 2023.

[76] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

[77] Qingsong Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, and Huan Xu. Time series data augmentation for deep learning: A survey. *arXiv preprint arXiv:2002.12478*, 2020.

[78] Brian McFee, Eric J Humphrey, and Juan Pablo Bello. A software framework for musical data augmentation. In *ISMIR*, volume 2015, pages 248–254. Citeseer, 2015.

[79] Esben Jannik Bjerrum, Mads Glahder, and Thomas Skov. Data augmentation of spectral data for convolutional neural network (CNN) based deep chemometrics. *arXiv preprint arXiv:1710.01927*, 2017.

[80] Wilbur B Davenport, William L Root, et al. *An introduction to the theory of random signals and noise*, volume 159. McGraw-Hill New York, 1958.

[81] Ajay Kumar Boyat and Brijendra Kumar Joshi. A review paper: noise models in digital image processing. *arXiv preprint arXiv:1505.03489*, 2015.

[82] Russell G. Congalton. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1):35–46, 1991.

[83] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. Pearson correlation coefficient. In *Noise reduction in speech processing*, pages 1–4. Springer Berlin Heidelberg, 2009.

[84] Luyan Ji, Peng Gong, Xiurui Geng, and Yongchao Zhao. Improving the accuracy of the water surface cover type in the 30 m FROM-GLC product. *Remote Sensing*, 7(10):13507–13527, 2015.

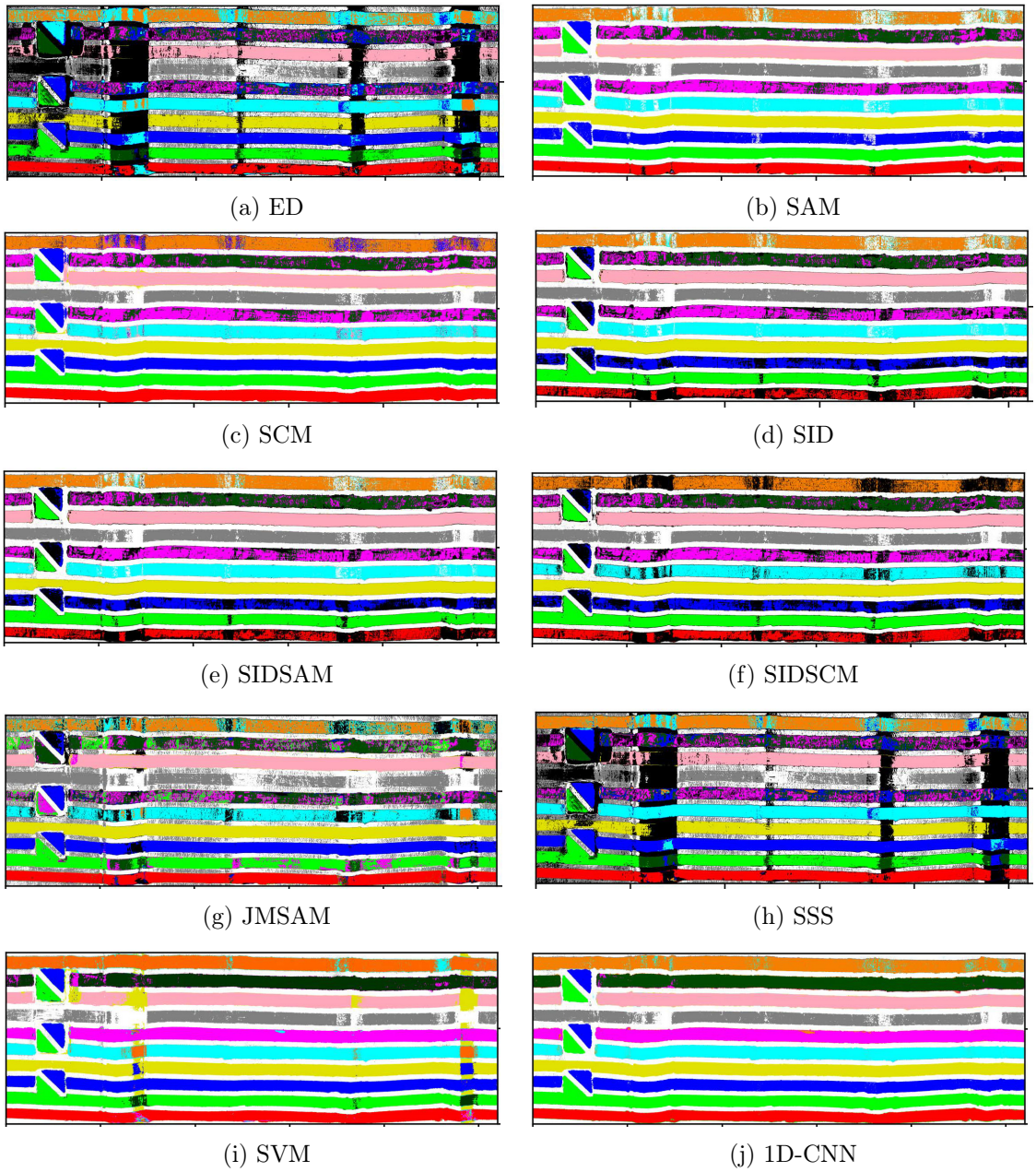
[85] Augustine Gray and John Markel. A spectral-flatness measure for studying the autocorrelation method of linear prediction of speech analysis. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 22(3):207–217, 1974.

[86] Anne Bouquillon, Gaia Ligovich, and Gauthier Roisine. De la couleur du bestiaire «esmaillé» de palissy. *Technè. La science au service de l'histoire de l'art et de la préservation des biens culturels*, (47):51–60, 2019.

[87] Louvre Museum. Relief of Seti I and Hathor. https://collections.louvre.fr/ark:/53355/cl010009693, 2021. [Accessed 05 August 2023].

Appendix

A Classification results.



(a) ED (b) SAM

(c) SCM (d) SID

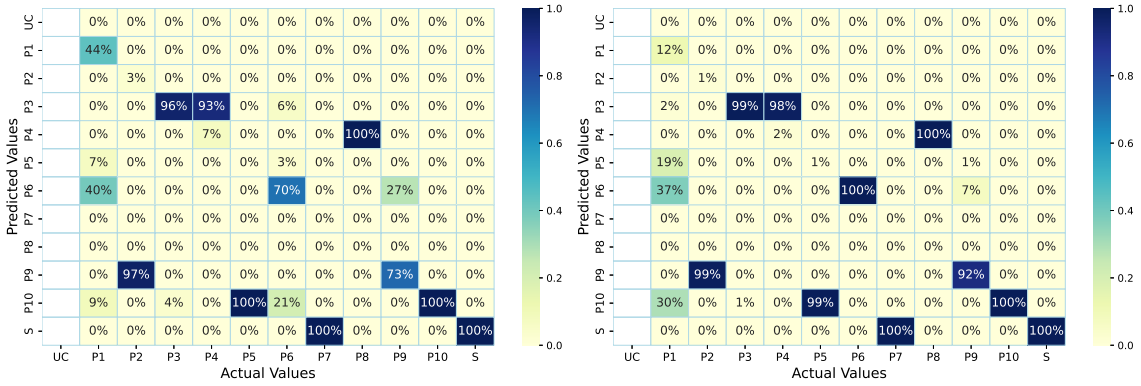
(e) SIDSAM (f) SIDSCM

(g) JMSAM (h) SSS

(i) SVM (j) 1D-CNN

Figure A.1: The classification results from ten algorithms using the reference spectrum from the flat region.

B Confusion matrix for regions R3 and R13, illustrating the misclassification of pigments for SVM.



(a) (b)

Figure B.1: Confusion matrix for SVM; (a): ROI : 3 , (b): ROI : 13.

C Confusion matrix illustrating classification accuracy obtained on a flat surface using SID Algorithm.

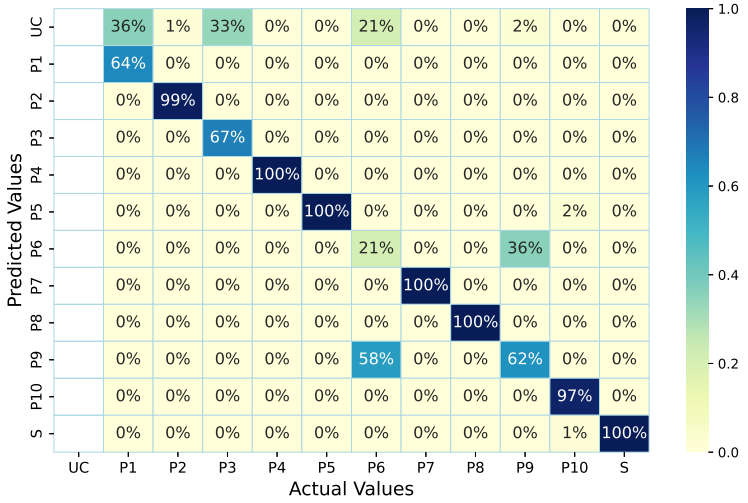


Figure C.1: Confusion matrix obtained by utilizing the SID Algorithm on a flat surface for ten pigments and substrate.

D Normalized reflectance spectrum of 10 pigments and substrate

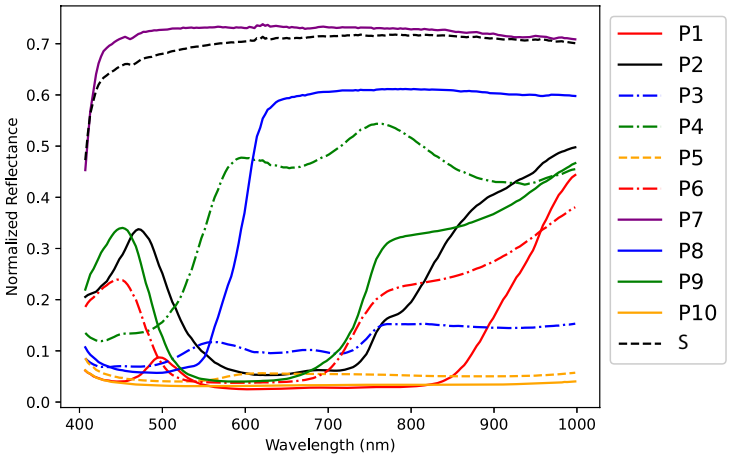


Figure D.1: Spectrum for ten pigments (P1 to P10) and substrate (S) measured at the flat surface by taking an average of 10 $\times$ 10 pixels.

E Accuracy of Pigment Classification at Various Elevations Using Different Reference Spectra Conditions.

	Flat Region										Elevation : 2.5 mm										Elevation : 5 mm										Elevation : 10 mm									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
ED	1	1	1	1	0.9	0	0.7	1	0.4	0.9	0.9	0.9	1	0.9	0.8	0.3	0.6	0.8	0.7	0.5	0.7	0.6	0.9	0.4	0.7	0.3	0.1	0.5	0.6	0.3	0.6	0.5	0.6	0.4	0.5	0.1	0.2	0.3	0.4	0.6
	78.74%										**74.82%**										**51.14%**										**41.61%**									
SAM	0.9	1	1	1	1	0.1	1	1	0.6	1	1	1	1	1	1	0.5	1	1	0.9	0.6	1	1	1	1	0.9	0.4	0.6	1	0.9	0.4	1	1	1	1	1	0.6	0.4	1	0.7	0.6
	85.34%										**88.31%**										**80.59%**										**83.84%**									
SCM	1	1	1	1	1	0.1	1	1	0.6	1	1	1	1	1	0.9	0.5	0.9	1	0.9	0.5	1	1	1	1	0.7	0.3	0.6	1	0.9	0.3	1	1	1	1	0.9	0.6	0.4	1	0.8	0.5
	86.25%										**87.38%**										**77.59%**										**81.73%**									
SID	0	1	0.8	1	1	0.1	1	1	0.6	1	0.2	1	0.9	1	1	0.5	1	1	0.9	0.7	0.3	1	0.8	1	0.9	0.4	0.6	1	0.9	0.4	0.1	0.9	0.7	1	1	0.6	0.5	1	0.8	0.7
	75.04%										**80.95%**										**72.99%**										**72.47%**									
SID-SAM	0	1	0.8	1	1	0.1	1	1	0.6	1	0.2	1	0.9	1	1	0.5	1	1	0.9	0.6	0.3	1	0.7	1	0.9	0.4	0.6	1	0.9	0.4	0.1	0.8	0.7	1	1	0.6	0.5	1	0.8	0.7
	74.60%										**80.11%**										**71.58%**										**70.81%**									
SID-SCM	0	1	0.8	1	1	0.1	1	1	0.6	0.9	0.3	1	0.9	1	0.9	0.5	0.9	1	0.9	0.4	0.4	1	0.7	1	0.5	0.4	0.6	1	0.9	0.2	0.4	0.8	0.6	1	0.8	0.6	0.4	1	0.8	0.3
	73.69%										**76.97%**										**66.13%**										**67.81%**									
JMSAM	1	0.9	0.9	1	1	0	0.6	1	0.3	0.9	0.8	0.7	1	1	0.7	0.2	0.9	1	0.6	0.4	0.7	0.4	1	1	0.5	0.2	0.5	1	0.6	0.3	0.9	0.4	0.8	1	0.4	0.1	0.4	0.5	0.6	0.3
	75.60%										**73.79%**										**62.62%**										**54.31%**									
SSS	1	1	1	1	1	0	1	1	0.4	1	1	1	1	1	1	0.3	0.8	1	0.8	0.6	0.8	0.7	1	0.5	0.8	0.3	0.2	0.5	0.6	0.4	0.9	0.5	1	0.4	1	0.1	0.3	0.4	0.5	0.7
	83.38%										**84.17%**										**58.99%**										**56.91%**									
SVM	1	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	0.8	1	1	1	1	0.8	1	0.5	0.6	1	0.4	0.6	0.7	1
	99.30%										**99.83%**										**96.98%**										**76.35%**									
1D-CNN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	0.9	1	1	1	1	1	1	1	1	1	0.7	1	1	1
	99.98%										**99.55%**										**97.74%**										**97.22%**									

Figure E.1: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; spectral library and training datasets built using reference spectrum from an elevated region of 2.5 mm; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

	Flat Region										Elevation : 2.5 mm										Elevation : 5 mm										Elevation : 10 mm									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
ED	1	1	1	1	0	0.6	0	1	0.5	1	0.9	1	0.9	0.9	0.2	0.4	0.5	0.8	0.3	0.9	0.7	0.6	0.7	0.3	0.3	0.2	0.3	0.5	0.3	0.9	0.7	0.5	0.4	0.4	0.2	0.4	0.2	0.3	0.2	0.9
	70.58%										**69.83%**										**47.81%**										**41.56%**									
SAM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1	1	0.8	1	1	1	1	0.9	1	0.6	1	1	0.7	1	1	1	1	1	1	0.5	1	1	0.9
	99.57%										**97.03%**										**92.07%**										**92.55%**									
SCM	1	1	1	1	0.9	1	1	1	1	0.3	1	1	1	1	0.9	1	1	1	1	0.7	1	1	1	1	0.9	1	0.7	1	1	0.5	1	1	1	1	0.9	1	0.5	1	1	0.8
	91.23%										**95.24%**										**91.15%**										**92.08%**									
SID	0.3	0.6	0.8	1	1	0.9	1	1	0.9	0.9	0.5	0.8	0.9	1	1	0.9	1	1	0.9	0.8	0.4	0.8	0.8	1	0.9	0.8	0.7	1	0.9	0.7	0.8	0.6	0.8	1	1	0.9	0.5	1	0.9	0.9
	84.55%										**88.43%**										**80.41%**										**82.14%**									
SID-SAM	0.4	0.7	0.8	1	1	0.8	1	1	0.9	0.7	0.5	0.8	0.9	1	0.9	0.9	1	1	0.9	0.8	0.4	0.8	0.8	1	0.9	0.8	0.7	1	0.9	0.7	0.8	0.5	0.7	1	1	0.9	0.5	1	0.9	0.9
	82.64%										**87.82%**										**79.36%**										**81.51%**									
SID-SCM	0.8	1	0.8	1	0.1	0.8	1	1	1	0.1	0.7	1	0.9	1	0.6	0.9	1	1	0.9	0.5	0.6	1	0.7	1	0.9	0.7	0.7	1	0.9	0.5	0.9	0.7	0.7	1	0.8	0.9	0.5	1	0.9	0.6
	74.13%										**84.29%**										**80.09%**										**80.36%**									
JMSAM	1	0.9	0.9	1	0	0.9	0	1	0.8	0.1	0.9	0.8	1	1	0.3	0.5	0.6	1	0.4	0.4	0.7	0.4	1	1	0.4	0.7	0.6	1	0.4	0.5	1	0.4	0.7	1	0.2	0.7	0.4	0.7	0.5	0.3
	66.85%										**68.74%**										**66.34%**										**58.83%**									
SSS	1	1	1	1	0	0.7	0	1	0.6	1	1	1	1	1	0.4	0.4	0.5	1	0.4	0.9	0.8	0.7	1	0.5	0.6	0.2	0.3	0.5	0.3	0.8	0.8	0.5	0.7	0.4	0.3	0.5	0.2	0.4	0.2	0.9
	73.03%										**75.66%**										**58.84%**										**49.49%**									
SVM	1	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8	0.8	1	0.7	1	1	1
	99.08%										**99.85%**										**99.70%**										**92.93%**									
1D-CNN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	99.98%										**99.56%**										**98.81%**										**99.34%**									

Figure E.2: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; spectral library and training datasets built using reference spectrum from an elevated region of 5 mm; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

	Flat Region										Elevation : 2.5 mm										Elevation : 5 mm										Elevation : 10 mm									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
ED	1	0.8	0.9	0.2	0.3	0.5	0	0	0.5	1	0.9	0.7	1	0.6	0.6	0.4	0.2	0.5	0.3	0.6	0.7	0.7	0.8	0.6	0.6	0.2	0.5	0.6	0.3	0.4	0.5	0.7	0.7	0.3	0.4	0.4	0.2	0.3	0.2	0.7
	52.54%										**57.61%**										**53.10%**										**44.44%**									
SAM	1	1	1	1	1	1	1	1	0.9	1	1	1	0.9	1	1	1	1	1	0.7	0.8	1	1	0.8	1	0.9	1	0.9	1	0.7	0.7	1	1	0.9	1	1	1	0.7	1	0.7	0.9
	99.08%										**93.20%**										**90.04%**										**91.67%**									
SCM	1	1	1	1	1	1	1	1	0.9	0.9	1	1	1	1	1	1	1	1	0.7	0.7	1	1	1	1	0.9	1	0.9	1	0.6	0.5	1	1	1	1	1	1	0.8	1	0.6	0.7
	97.52%										**92.98%**										**89.11%**										**91.12%**									
SID	0.2	1	0.7	1	1	0.6	1	1	0.9	1	0.5	1	0.5	1	1	0.9	1	1	0.7	0.8	0.4	1	0.6	1	0.9	0.7	1	1	0.7	0.7	0.6	1	0.7	1	1	0.9	0.8	1	0.8	0.9
	83.67%										**83.32%**										**79.95%**										**86.31%**									
SID-SAM	0.2	1	0.6	1	1	0.5	1	1	0.8	0.8	0.5	1	0.4	1	1	0.9	1	1	0.6	0.8	0.4	1	0.5	1	0.9	0.6	0.9	1	0.7	0.7	0.5	1	0.6	1	1	0.9	0.8	1	0.7	0.9
	78.88%										**81.36%**										**78.34%**										**83.92%**									
SID-SCM	0.6	1	0.5	1	0.9	0.4	1	1	0.9	0.1	0.6	1	0.4	1	0.9	0.8	1	1	0.6	0.5	0.7	1	0.6	1	0.9	0.6	0.9	1	0.7	0.4	0.7	1	0.7	1	1	0.9	0.8	1	0.7	0.6
	74.12%										**78.73%**										**77.98%**										**83.15%**									
JMSAM	1	0.9	1	1	0.7	0.9	0	1	0.7	0.2	1	0.8	0.9	1	0.7	0.5	0.4	1	0.3	0.3	0.8	0.9	0.8	1	0.6	0.4	0.6	1	0.3	0.3	0.9	0.9	1	1	0.5	0.6	0.5	0.9	0.3	0.2
	72.79%										**68.69%**										**66.79%**										**68.29%**									
SSS	1	1	1	0.6	1	0.6	0	0.3	0.6	1	1	0.7	1	0.6	1	0.5	0.4	0.5	0.3	0.7	0.9	0.7	0.9	0.7	0.8	0.2	0.5	0.6	0.3	0.5	0.6	0.8	1	0.4	1	0.4	0.3	0.4	0.2	0.8
	70.44%										**67.06%**										**62.01%**										**58.22%**									
SVM	1	1	1	1	1	1	0.6	1	1	1	1	1	1	1	1	1	1	0.8	1	1	1	1	1	0.9	1	1	0.9	0.9	1	1	1	1	1	1	1	1	1	1	1	1
	95.48%										**97.26%**										**96.94%**										**99.26%**									
1D-CNN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	99.97%										**99.40%**										**98.51%**										**99.52%**									

Figure E.3: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; spectral library and training datasets built using reference spectrum from an elevated region of 10 mm; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

	Flat Region										Elevation : 2.5 mm										Elevation : 5 mm										Elevation : 10 mm									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
ED	1	1	1	1	0.4	0.5	1	1	0.5	0.9	1	1	1	0.9	0.7	0.4	0.3	0.8	0.3	0.6	0.7	0.7	0.9	0.4	0.6	0.2	0.1	0.5	0.3	0.4	0.6	0.5	0.6	0.4	0.4	0.4	0.2	0.3	0.2	0.6
	82.31%										**70.02%**										**47.58%**										**41.12%**									
SAM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1	1	0.9	1	0.7	1	1	0.8	1	1	1	1	1	1	0.5	1	1	0.9
	99.84%										**97.85%**										**92.98%**										**93.34%**									
SCM	1	1	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	1	0.7	1	1	1	1	0.7	1	0.6	1	1	0.5	1	1	1	1	0.9	1	0.5	1	1	0.7
	99.37%										**95.82%**										**88.27%**										**90.90%**									
SID	0.7	1	0.9	1	1	0.8	1	1	1	1	0.5	1	1	1	1	0.9	1	1	0.9	0.9	0.3	0.9	1	1	0.9	0.7	0.7	1	1	0.8	0.6	1	1	1	1	0.9	0.5	1	1	0.9
	94.54%										**92.33%**										**83.26%**										**87.63%**									
SID-SAM	0.8	1	0.9	1	1	0.8	1	1	1	1	0.5	1	1	1	1	0.9	1	1	0.9	0.9	0.3	0.9	1	1	0.9	0.7	0.7	1	0.9	0.8	0.6	1	0.9	1	1	0.9	0.5	1	1	0.9
	94.01%										**92.04%**										**82.87%**										**87.38%**									
SID-SCM	1	1	0.9	1	1	0.7	1	1	1	0.3	0.7	1	1	1	0.9	0.9	1	1	0.9	0.6	0.5	1	0.9	1	0.7	0.7	0.7	1	0.9	0.4	0.7	1	1	1	0.9	0.9	0.5	1	0.9	0.7
	89.32%										**89.12%**										**78.63%**										**85.76%**									
JMSAM	1	0.9	1	1	0.6	0.9	1	1	0.6	0.5	1	0.8	1	1	0.6	0.6	0.5	1	0.3	0.4	0.7	0.5	1	1	0.5	0.7	0.5	1	0.3	0.4	0.9	0.7	1	1	0.5	0.7	0.3	0.9	0.4	0.3
	85.00%										**71.60%**										**66.47%**										**67.98%**									
SSS	1	1	1	1	1	0.6	1	1	0.6	1	1	1	1	1	1	0.5	0.4	1	0.3	0.8	0.8	0.7	1	0.5	0.8	0.2	0.2	0.5	0.3	0.6	0.7	0.5	1	0.4	1	0.4	0.2	0.4	0.2	0.8
	91.47%										**78.45%**										**57.31%**										**55.87%**									
SVM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	99.97%										**99.71%**										**99.36%**										**99.74%**									
1D-CNN	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	99.94%										**99.54%**										**98.88%**										**99.59%**									

Figure E.4: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; using average reference spectra for supervised and ML models; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

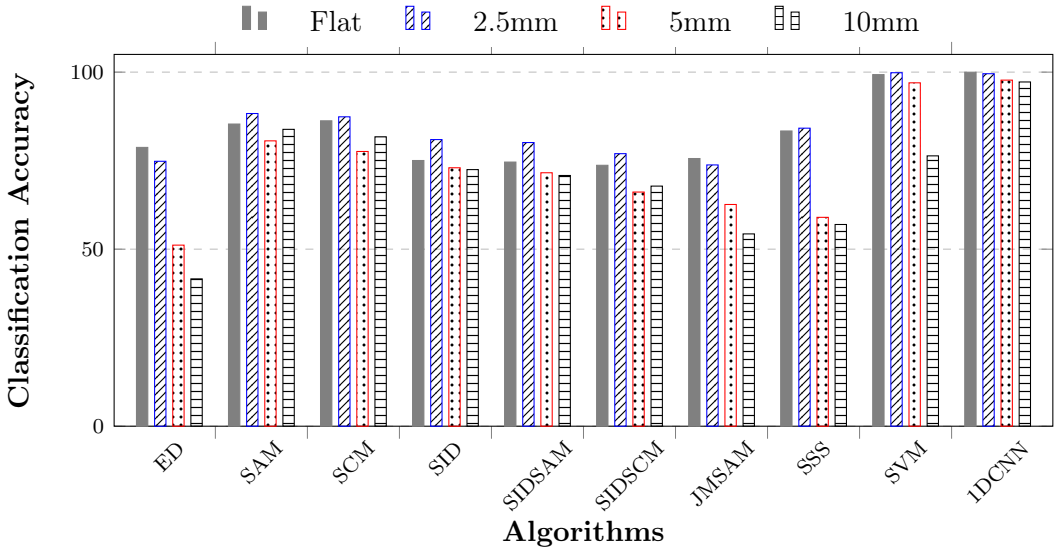


Figure E.5: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 2.5 mm.

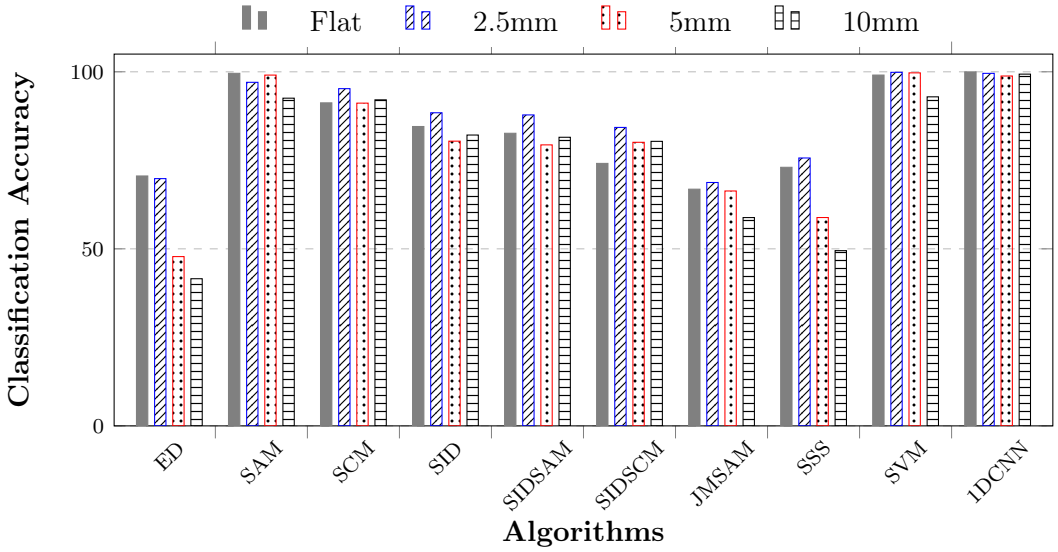


Figure E.6: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 5 mm.

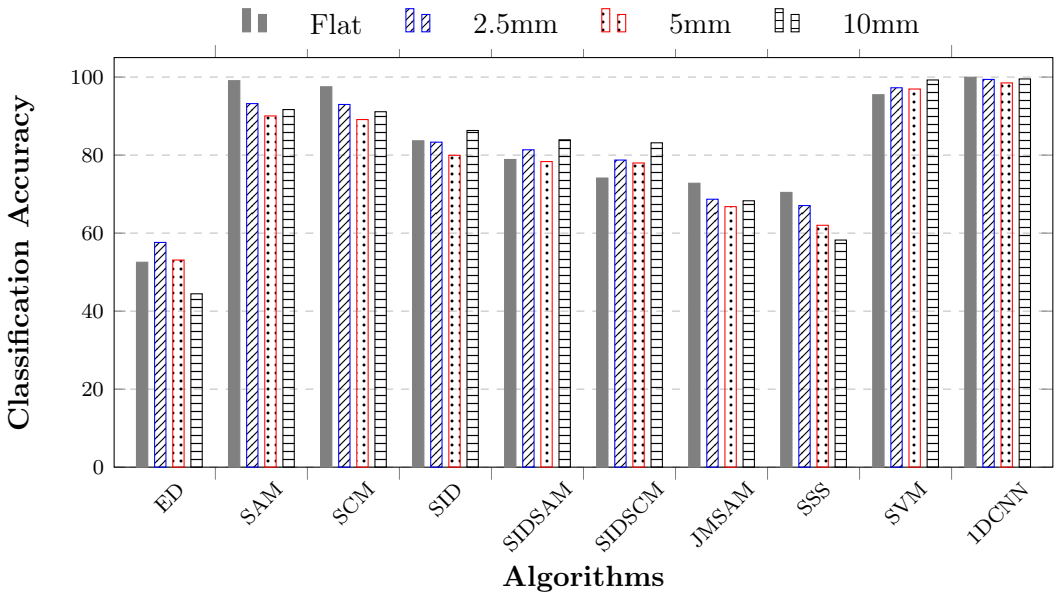


Figure E.7: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 10 mm.

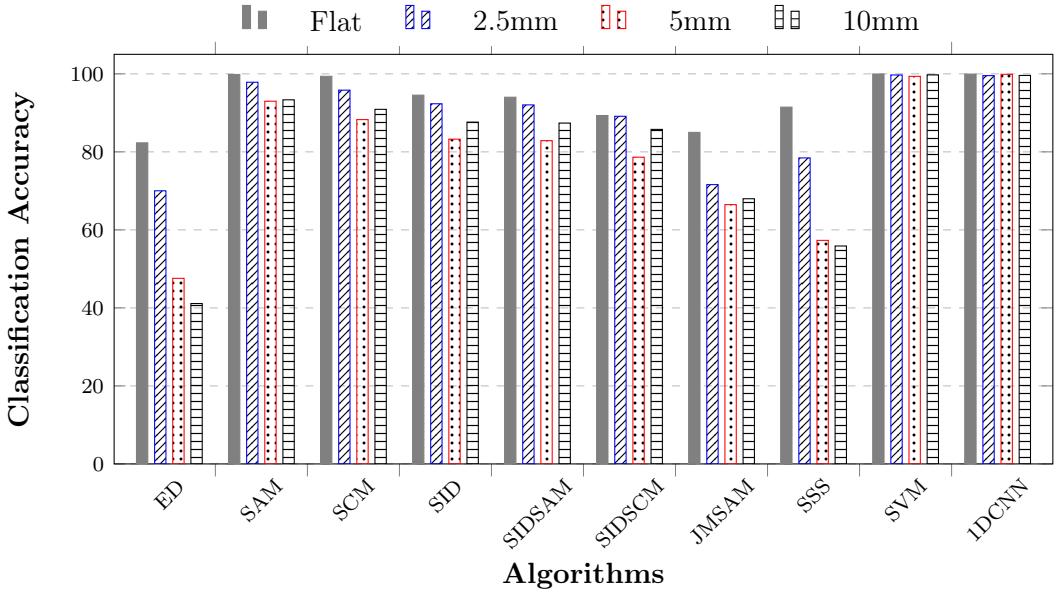
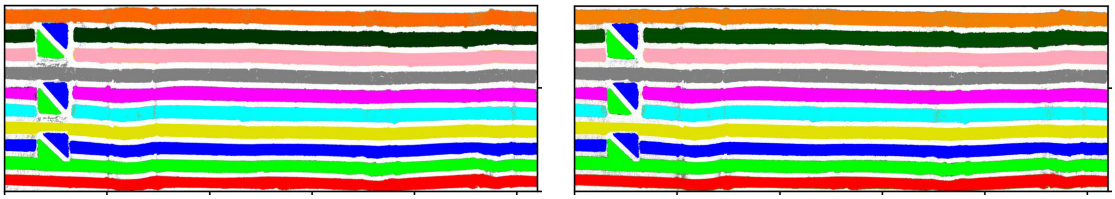


Figure E.8: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained using average reference spectra; the averages were computed by considering a pixels values of a line drawn from right to left of the mockup for each pigment.

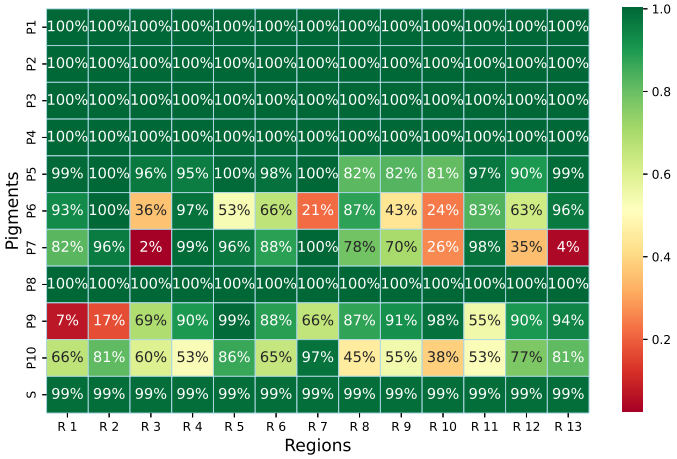
F Classification results for SVM and 1D-CNN.



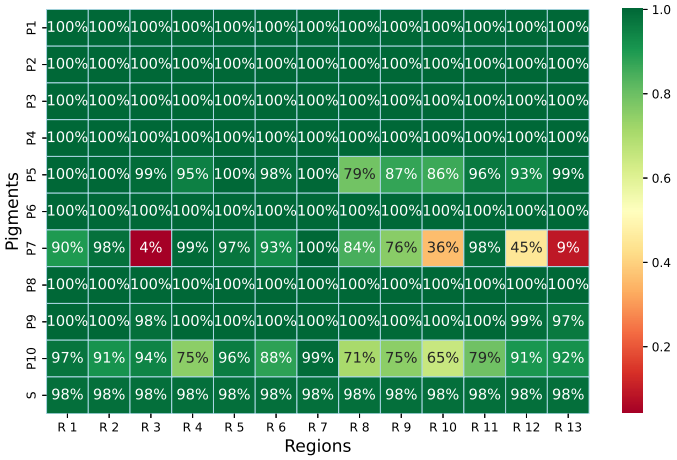
(a) SVM (b) 1D-CNN

Figure F.1: The classification results for SVM and 1D-CNN using the averaged reference spectrum.

G Classification Accuracy for SVM using Data Augmentation



(a)



(b)

Figure G.1: Classification accuracy for each pigment across all Regions of Interest (ROIs) for SVM, (a): With data augmentation using single spectrum from flat region, and (b): With data augmentation using averaged single spectrum.