

## Research Article

# A New Rule-Based Classification Method Using Shape-Based Properties of Spectral Curves

Songuel Polat <sup>1,2</sup>, Alain Tremeau <sup>2</sup>, and Frank Boochs <sup>1</sup>

<sup>1</sup>*i3mainz, Institute for Spatial Information and Surveying Technology, Mainz University of Applied Sciences, 55128 Mainz, Germany*

<sup>2</sup>*Hubert Curien Laboratory, University Jean Monnet, 42100 Saint-Etienne, France*

Correspondence should be addressed to Songuel Polat; [songuel.polat@gmx.de](mailto:songuel.polat@gmx.de)

Received 10 August 2021; Revised 15 September 2021; Accepted 15 January 2022; Published 24 February 2022

Academic Editor: Arnaud Cuisset

Copyright © 2022 Songuel Polat et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to its high spatial and spectral information content, hyperspectral imaging opens up new possibilities for a better understanding of data and scenes in a wide variety of applications. An essential part of this process of understanding is the classification part. However, the high spatial and spectral resolution also leads to enormous amounts of data. The effective handling and use of such datasets for classification requires processing steps (dimensionality reduction through feature selection or feature extraction) that are not always goal-oriented. In this article, a new general classification approach is presented that uses the geometric shape of spectral signatures instead of purely statistical methods. In contrast to classical classification approaches (e.g., SVM, KNN), not only are reflectance values taken into account, but also parameters such as curvature points, curvature values, and the curvature behavior of spectral signatures are used to develop shape-describing rules in order to use them for classification by a rule-based procedure with IF-THEN queries. The flexibility and efficiency of the methodology are demonstrated on datasets from two different application domains and lead to convincing results with good performance.

## 1. Introduction

Optical technology developments are extending the possibilities to better understand the world and its resources. Starting with images consisting of three color channels covering the visual electromagnetic spectrum, the developments since the late 1960s opened up the possibility of using spectral properties for identification of materials by using multispectral images with tens of channels. Especially with the developments in the last two decades, another enormous step forward was made and the low spectral resolution of multispectral images was overcome. With several hundred narrow channels, hyperspectral imaging (HSI) opens up completely new possibilities for analysis in a wide variety of application fields [1]. As examples, [2–4] use HSI to maintain and increase crop yields in precision agriculture. The evaluation of food quality and safety through the use of HSI is part of the research of [5–7]. Other applications of HSI can be found in medicine to diagnose

diseases or to monitor wound healing [8–10], in the art market to verify the authenticity of artworks [11–13], and in forensics to analyze crime scenes [14–16].

The key technology behind those applications is the HSI with detailed spectral and spatial information, which makes the HSI a powerful information source for advanced classification methods like k-nearest neighbor, support vector machines, random forests, neural networks, and deep learning approaches. A comparison of these mentioned classification methods is conducted in [17] and shows that there is no classifier that consistently provides the best performance and that the quality of the classification result mainly depends on factors such as the availability of training samples, processing requirements, tuning parameters, and speed of the algorithm. Another aspect to be considered in the context of mentioned classifiers is the Curse of Dimensionality. If high dimensional HSI are directly used as input, the classification accuracy decreases, while the computational effort of the model tends to increase

exponentially. To avoid this problem, a dimensionality reduction is essential [18]. The reduction of dimensions implies that algorithms automatically have to extract a set of characteristic spectral values, which have to represent the deciding features of the objects in question, from the entire course of a spectral signature. The automatic determination of the optimal number of relevant features [19] or the time required for example in the case of graph-based methods [20] is among the main problems. The amount of advanced existing studies (e.g., [19–27]) proves that there is no existing satisfactory, robust, and reliable methodology and that this topic is one of the main open issues in spectral imaging. As a consequence, this recommended step leads to loss of information in the spectral space, because the selected bands do not give an accurate description of the original spectral signature.

Other classification approaches based on indices [28–30] also arbitrarily select a subset of spectral values. One of the most popular is the Normalized Difference Vegetation Index (NDVI) from the field of remote sensing [31], which allows the classification of vegetation covered areas on the Earth's surface by combining only a few bands of NIR and Red wavelengths. Another example of such an index is the Normalized Difference Plastic Index (NDPI), which combines shortwave infrared bands for the classification of plastic materials in urban areas [32]. By limiting to a small number of spectral bands in a restricted spectral range, broad classifications such as separation between plants, water, or urbanized areas can be made, but finer separation (e.g., between plant species or plastic materials) would be difficult due to the reduced number of spectral bands and contradicts the use of HSI, which offers the enormous advantage of high-resolution spectral information.

Both band selection methods and indices consider only a small set of characteristic values, while the shape of the spectra is not taken into account. A consideration of the shape of spectral curves as important information source can be found in [33–35].

The work of [34] consists of a classification approach that fully utilizes the shape of spectral curves by using a code to parameterize the spectral curve shape. The research of [35] deals with the definition of analysis rules based on spectral features like band position, band depth, band width, and band asymmetry. These key parameters are used to describe the absorption features of spectra. Disadvantage of both works is the time required to develop the description parameters of the curve shapes. Both approaches use tables to store the parameters and a subsequent matching process between the parameters of the reference data and the spectra that need to be classified, which also makes the classification process a time-consuming task. The advantage of considering shapes, especially in combination with high-resolution spectral data, is clear: spectra express the mixed reflectivity of the elements that make up an object (e.g., molecules, pigments, cell structure, and water content), which is why each individual component has only a proportional influence on a spectrum. Changes in the composition of the elements then mainly change the mixture of all spectral contributions, resulting in local or regional variations and changing the shape of a spectrum.

In this article we follow the shape-based works of [33–35] and propose a new rule-based classification method using shape-based properties of spectral curves like curvature points, curvature values, curvature direction, and spectral values.

Particular attention should be paid to the following points:

- (i) The formulation of the rules should not become complex
- (ii) The classification process should not be time-consuming

Keeping these points in mind, an approach was developed that allows establishing rules, using any kind of logical elements, in a straightforward manner. The establishment of rules implies existing knowledge. A prior analysis of the data and the acquisition of knowledge enable a better understanding of the data and allow structuring and simplifying a problem. Research papers from the field of remote sensing that use the advantage of knowledge can be found in [36–41]. The experimental part for the evaluation of the method was performed on two different datasets. A detailed description of the method and the used datasets are part of the following section.

## 2. Materials and Methods

*2.1. Spectral Data Set Acquisition.* The hyperspectral systems used in this work are two pushbroom cameras from Specim Ltd. (Oulu, Finland). The Specim FX10 captures the spectral signature from 400 nm to 1000 nm (233 bands), while the Specim FX17 captures the spectral signature from 900 nm to 1700 nm (229 bands). All spectral images acquired by these cameras are radiometrically normalized by using dark reference images for dark-current (closed shutter) and a white reference image to reduce the influence of the intensity variability. For the white calibration a 99% reflectance tile was used. As shown in Figure 1, a set of different HSI consisting of two different object types are used to demonstrate the proposed method.

The first dataset was captured with the Specim FX17 and shows an image of seven classes of different plastic types. It has a resolution of  $661 \times 500$  pixels and 229 bands with a spectral range from 900 nm to 1700 nm. The second scene was captured with the Specim FX10 and the Specim FX17 and shows ten classes of different plant types. The HSI of both cameras were combined and have a resolution of  $1220 \times 640$  pixels and 462 bands covering the spectral range from 400 nm to 1700 nm.

The HSI with different plastic types is used to demonstrate the basic functionality of the approach and gets more attention due to existing ground truth (the datasets can be found on <https://doi.org/10.5281/zenodo.5068201>). As part of a waste sorting application, the demonstration also covers experiments with real waste consisting of plastics and electronic waste like printed circuit boards (PCB). The plant-based dataset, on the other hand, is used to demonstrate the potential and flexibility of the approach regarding different kind of applications.

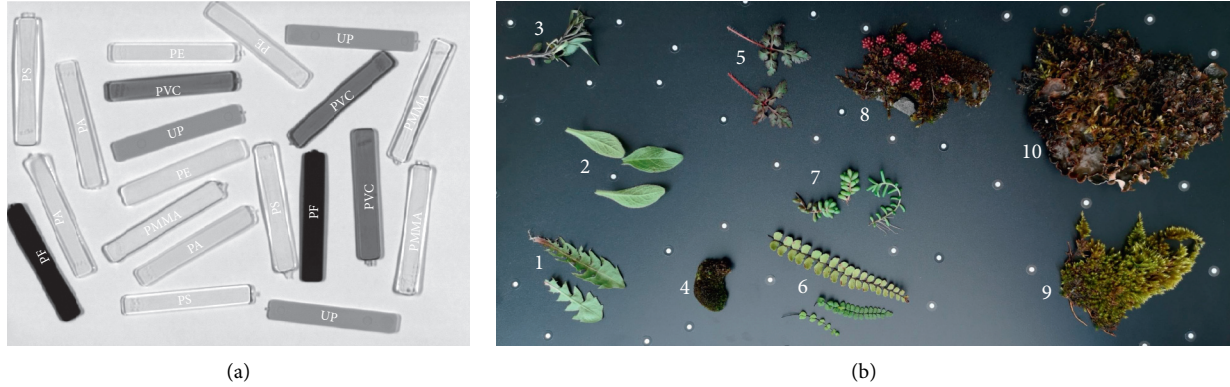


FIGURE 1: Images of the captured scenes. (a) Seven classes of different plastics (plastic samples were kindly provided by PlasticsEurope Deutschland e.V. (<https://www.plasticseurope.org/de>)) consisting of phenolic resin (PF), polyamide (PA), polyvinylchloride (PVC), polyethylene (PE), polymethylmethacrylate (PMMA), polyester resins (UP), and polystyrene (PS); (b) ten classes of different vegetation types consisting of *Taraxacum* (1), *Inula conyza* (2), *Campanula rotundifolia* (3), moss (4), *Geranium robertianum* (5), *Asplenium trichomanes* (6), green sedum (7), moss and red sedum (8), moss (9), moss and lichen (10).

**2.2. Methodology.** Spectral signatures offer the ability to distinguish between materials and are the result of reflected light from the surface, which is captured within a broad electromagnetic spectrum. The reason for different signatures is a combination of molecules in materials and the morphological structure. Different molecules result in different spectral signatures. The morphological structure, on the other hand, is important because of the resulting light path that is created by reflection, absorption, transmission, and deflection from the different components of an object. That is why plants show different spectra when the cell structure changes, e.g., due to stress or aging. Both differences in molecules and structure result in different shaped spectral curves. The idea of our work is to describe the spectral curve shape by a combination of spectral values and shape-based parameters, to use this knowledge for the formation of rules.

Changes in the material composition of objects lead to local or regional changes in the course of spectral signatures. These changes inevitably lead to changes in the curvature behavior, what makes the curvature  $\kappa$  to a significant parameter for the shape description and the modelling of spectral changes.

Mathematically, the curvature is the change of a curve that occurs when the curve is traversed and can be expressed in parametric form for each point  $\mathbf{P}(x(t), y(t))$  using equation (1), where points refer to derivatives.

$$\kappa = \frac{|\dot{x}\ddot{y} - \dot{y}\ddot{x}|}{(\dot{x}^2 + \dot{y}^2)^{3/2}}. \quad (1)$$

While the curvature of a straight line is zero everywhere and the curvature of a circle is equal at all points, the curvature for all other curves changes from point to point and indicates how strongly the curve at a point  $\mathbf{P}$  deviates from a straight line. Thus, for the description of spectral shapes, we use the following properties of curvature:

- (i) The dimension of the current rate of change of the direction of a point moving on the curve: the greater the curvature, the greater the dimension of change

- (ii) The behavior of the curvature: if the curvature value is positive, it is called a convex curve, and in the case of a negative curvature value, it is considered a concave curve.

As shown in Figure 2, extreme values of the second derivative are used to select significant parameters. The combination of these parameters with selected spectral values allows a precise description of the shape of spectral curves using a few selective spectral bands. The base for the calculation of the curvatures is preprocessed spectral curves. The preprocessing consists of a smoothing and a subsequent step of Continuum Removal. A schematic representation of all essential steps can be found in Figure 3.

Continuum removal is a normalization procedure which allows a better quantification of absorption peaks after removing the overall concave shape of spectral curves [42, 43] and is illustrated in Figure 4. Due to the particularly highlighted absorption bands, rules based on curvature values can be developed much more efficiently. An example for continuum removed spectra and the calculated curvatures is shown in Figure 5. Depicted are continuum removed spectra for two different plastic types (PE, PS). In addition to the course of the spectral signature, the calculated curvature values are shown as positive and negative vertical lines in green and red color. The longer the line, the stronger the curvature at the respective band.

The curve behavior is represented by red and magenta colored dots. These points are maximum and minimum points and are automatically determined by the local maxima and minima of the second derivative. It helps to distinguish between concave and convex curve behavior, providing an important source of information for describing the shape. For the subsequent formation of rules, mainly the red lines are used, since these reflect both the concave and convex behavior and a significant change in curvature. All other curvature values (green lines) are not considered in the rule formation. To ensure that only significant changes in the curve are captured, a threshold is set for the selection of the relevant curvatures (red lines). The setting of the threshold

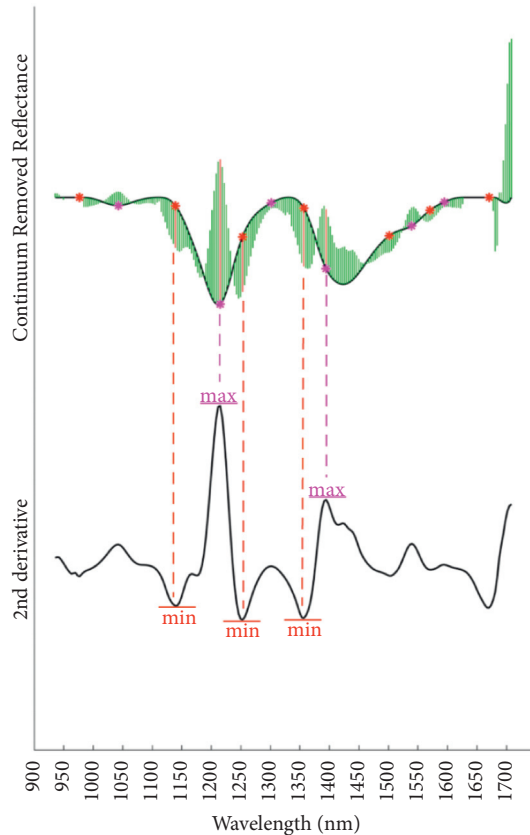


FIGURE 2: Use of minimum and maximum positions of 2nd derivative to select significant parameters for shape description. Parameters are the location (min/max), the curvature values at these locations (red lines), and the direction of curvature value (up for convex and down for concave behavior).

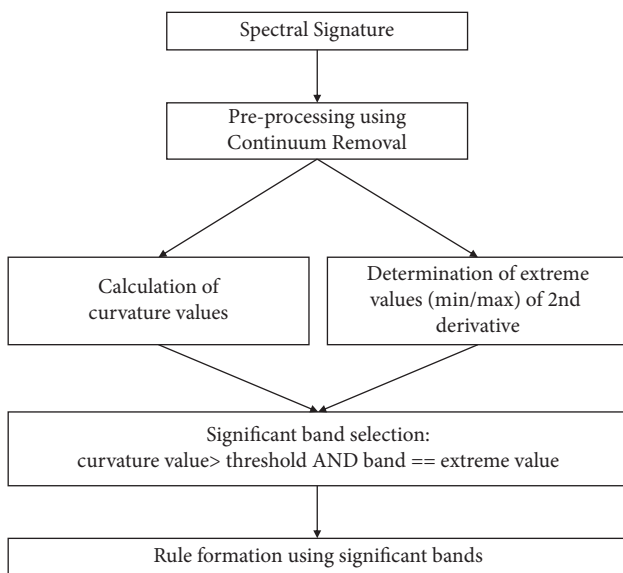


FIGURE 3: Schematic representation of proposed method.

mainly depends on the curve shape. The smaller the threshold the more detailed the description of the curves. However, for highly variant spectra (Figure 6) higher

threshold is sufficient. The examples in Figure 5 show the result of selected red lines for a threshold of 0.1. Building on these extracted parameters regarding the spectra for each material, a collection of conditions is formulated and used for the classification. It should be mentioned here that, in addition to the used local shaped values, other rules can also be integrated. For example, rules are describing global shape effects (e.g., the expressivity of the green peak for plants) or adding relations of averaged reflectivity in different regions as indices like NDVI do.

One of the main advantages of rule-based classifier is the simplicity. Once the knowledge on which rules are based has been worked out, conditions can be easily set up. Further advantages are the performance, the ability to handle redundant and irrelevant attributes, and the flexible extensibility of rule sets [44]. Acquiring knowledge can seem effortful, but in view of the resulting advantages, it should be seen as a clear benefit which allows structuring and simplifying problems by using expert knowledge. A comparison of deep learning and a knowledge-based method can be found in [45] and show that a rule-based method can even be better as machine learning-based methods. For better illustration of rule formation, we refer to the spectral curves of PE and PS from Figure 5 and express them as shown in Table 1.

The parameter CV stands for the curvature value at a specific spectral band and a positive or negative threshold is used to distinguish between convex and concave curve behavior. All conditions with a negative threshold will capture the downward red lines (concave behavior), while all conditions with a positive threshold will capture upward red lines (convex behavior). As already mentioned, for rule formation mainly the bands with high curvature are taken into account (red lines). Nevertheless, this does not mean that all bands are really necessary for the rule formation. By analyzing the data in advance, it is possible to achieve a clean classification even with a smaller number of selected bands. Therefore, only six conditions are defined for sample PS in Table 1, while in Figure 5 a total of seven red lines are present. The continuum removed spectra and rules for all other existing plastic types in the used dataset are listed in the Appendix (Figures 7–15). The corresponding spectral signatures are also shown in Figure 6. As an additional condition the continuum removed reflectance value (CRRV) could be used, if it is not possible to sufficiently distinguish the shape. For instance, to separate PF-black from all other available plastic types.

### 3. Results and Discussion

Once the rules are established, the next step is to apply them. This involves a pixel-by-pixel processing of the corresponding dataset and a check of the rule conditions for each individual pixel. If a condition applies, this pixel will be assigned to the appropriate class.

*3.1. Classification Results for Plastic Samples.* Applying the developed rules in the Appendix to the dataset consisting of plastic samples in Figure 1, results shown in Figure 16 were

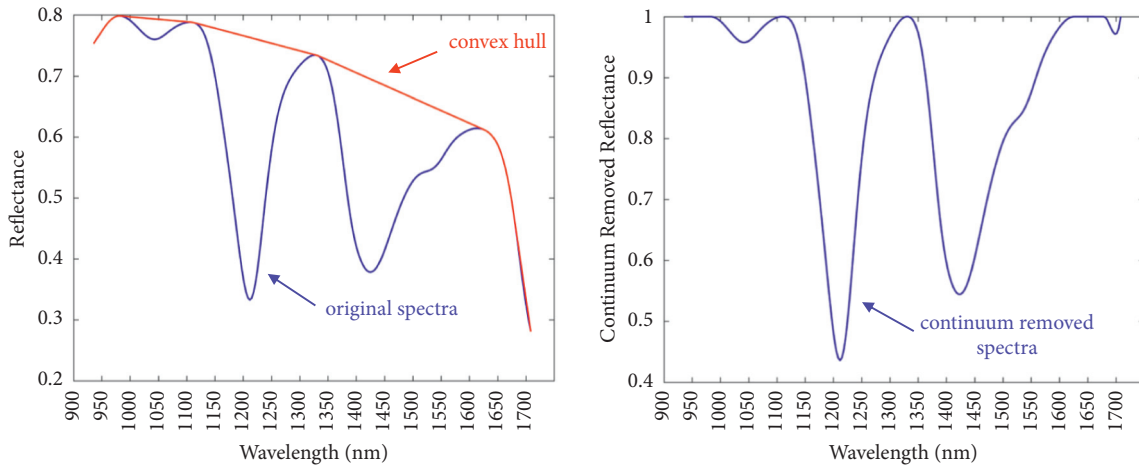


FIGURE 4: The use of continuum removal for quantification of absorption peaks by connecting the maxima of spectra with straight lines (left), setting the convex hull to 100%, and subtracting the original spectra (right).

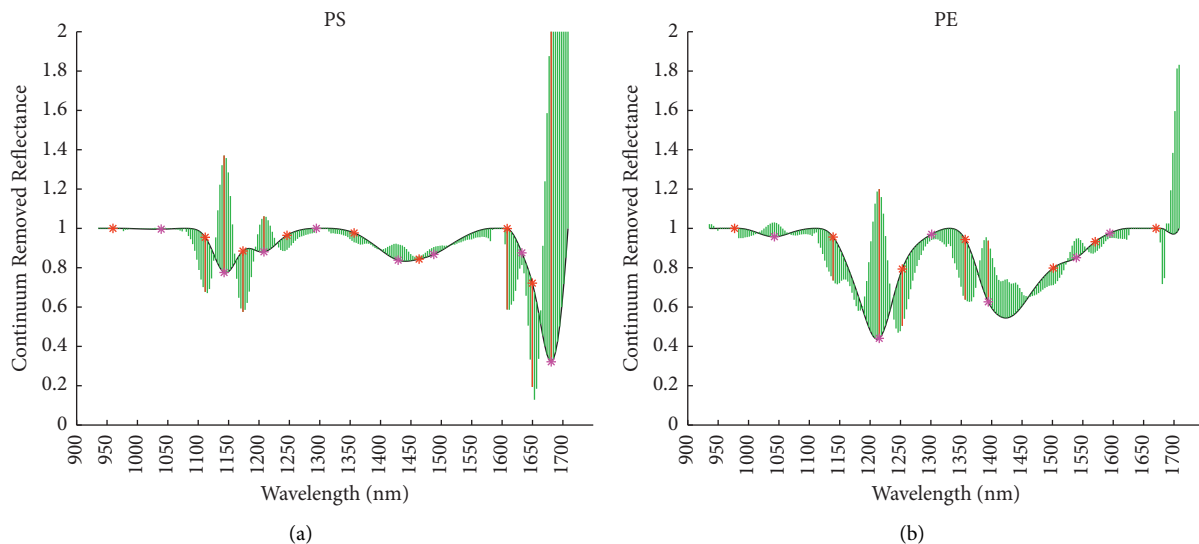


FIGURE 5: Continuum Removed Reflectance spectra (black line), calculated curvature values (green lines), maximum points with concave behavior (red dots), and minimum points with convex behavior (magenta dots) for (a) polystyrene (PS); (b) polyethylene (PE).

obtained. The result of the classification shows that the individual samples were classified correctly. A closer look at the border areas of the samples shows that a shadow effect occurs and that this effect can influence the quality of the classification. This shadow effect occurs especially with samples (e.g., PS) that do not lie in a planar position.

Another point to consider here is the classification of PF samples. The continuum removed reflectance spectrum illustrated in the Appendix does not correspond to the real spectrum of the material PF. The PF samples used here are composed of black colorants. A well-known problem is the strong absorption of black colorants like carbon black. Due to the strong absorption of light from the UV to the NIR there is no reflected light that can be detected by the sensor and thus no spectral information that can be used for a classification [46, 47]. The rule-based approach presented here offers the advantage that even in the absence of

reflectance rules can be established based on the existing limited information, which permit such a classification. In the case of these samples, it means all black plastics will be classified as PF-black. It must also be noted in this example that the samples used have clean, homogeneous surfaces and the resulting spectra have a correspondingly high degree of shape similarity. However, considering applications in the field of waste sorting, it is more common to work with dirty and damaged materials.

Therefore, for the evaluation of the methodology, a dataset ( $1253 \times 578$  pixels and 229 bands) with real waste consisting of plastic parts (objects 1–19) and circuit boards (objects 20–27) was processed additionally. The classification result for the real waste dataset is shown in Figure 17. The objects in the dataset were chosen randomly from a collection of different plastic parts. Therefore, it is not surprising that certain materials (e.g., PMMA, PVC, PE, and



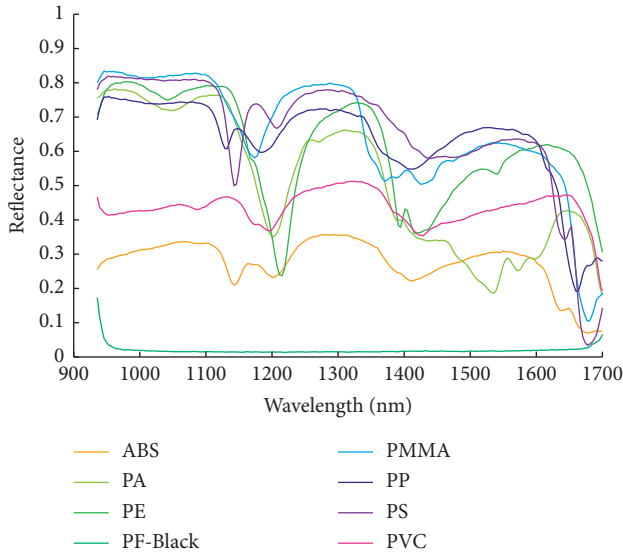


FIGURE 6: Reflectance spectra of available plastics. Each category of plastic material corresponds with a specific spectral shape; likewise the shapes of Continuum Removed Reflectance Spectra are different. This means the rule defined for a category is specific to this category and there is no possible confusion as the shapes are distinct enough. So, such rules can be used only if spectral curve shapes are all sufficiently distinct.

TABLE 1: Shape-based rule for PS and PE.

<b>IF</b> $CV_{1108} < -0.1$ <b>AND</b>	<b>IF</b> $CV_{1139} < -0.1$ <b>AND</b>
$CV_{1174} < -0.1$ <b>AND</b>	$CV_{1253} < -0.1$ <b>AND</b>
$CV_{1608} < -0.1$ <b>AND</b>	$CV_{1357} < -0.1$ <b>AND</b>
$CV_{1143} > +0.1$ <b>AND</b>	$CV_{1215} > +0.1$ <b>AND</b>
$CV_{1204} > +0.1$ <b>AND</b>	$CV_{1394} > +0.1$ <b>AND</b>
$CV_{1677} > +0.1$	<b>THEN class</b> $\rightarrow$ <i>PE</i>
<b>THEN class</b> $\rightarrow$ <i>PS</i>	

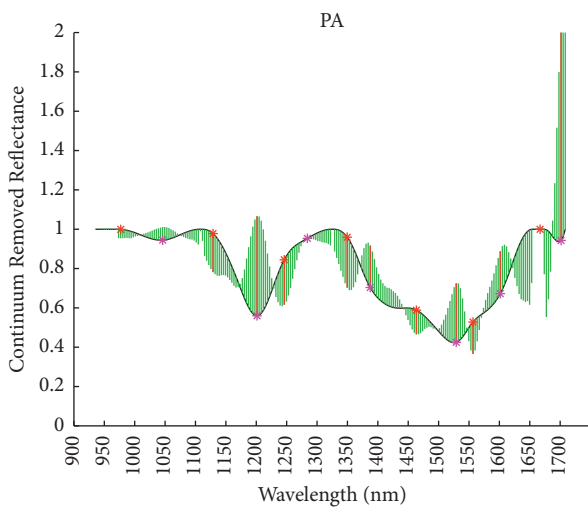


FIGURE 7: Continuum removed reflectance spectra and shape-based rule for PA.

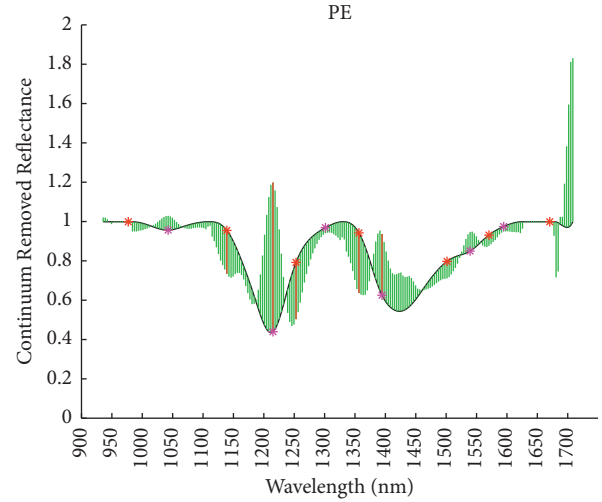


FIGURE 8: Continuum removed reflectance spectra and shape-based rule for PE.

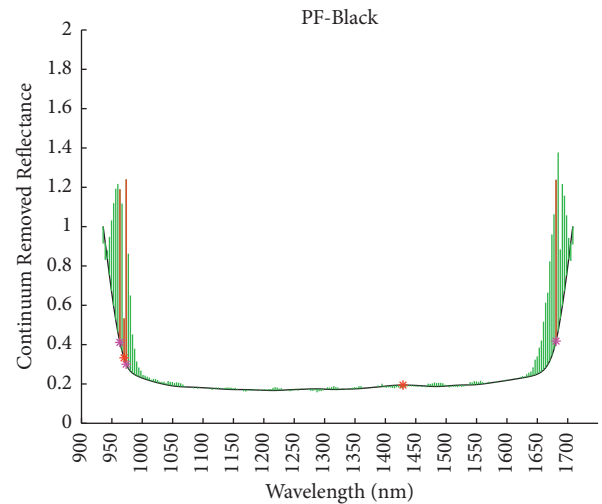


FIGURE 9: Continuum removed reflectance spectra and shape-based rule for PF-black.

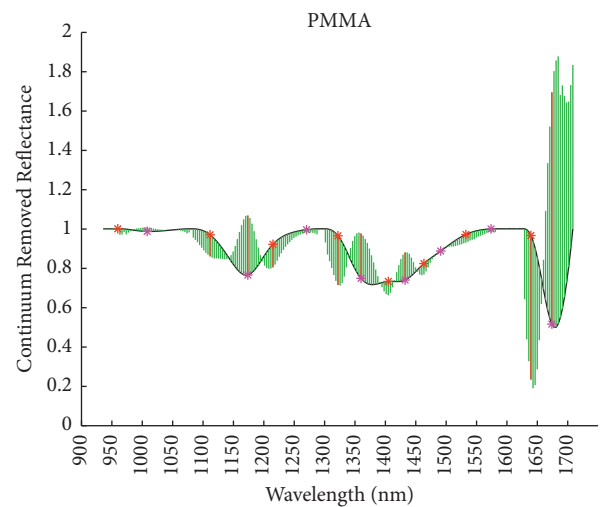


FIGURE 10: Continuum removed reflectance spectra and shape-based rule for PMMA.

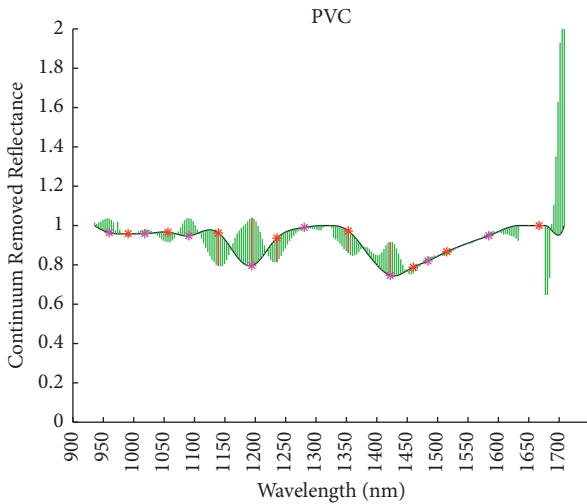


FIGURE 11: Continuum removed reflectance spectra and shape-based rule for PVC.

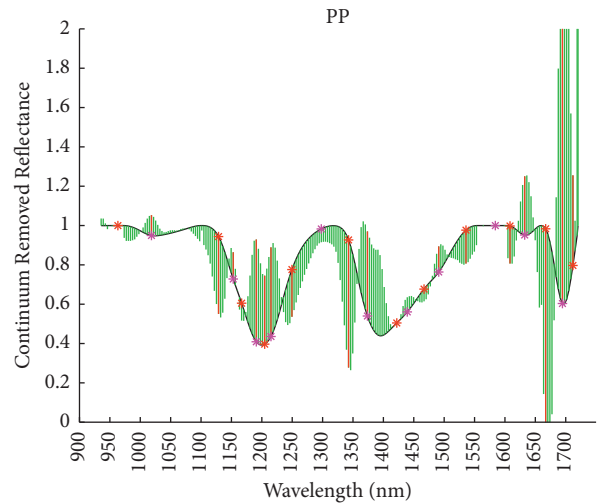


FIGURE 14: Continuum removed reflectance spectra and shape-based rule for PP.

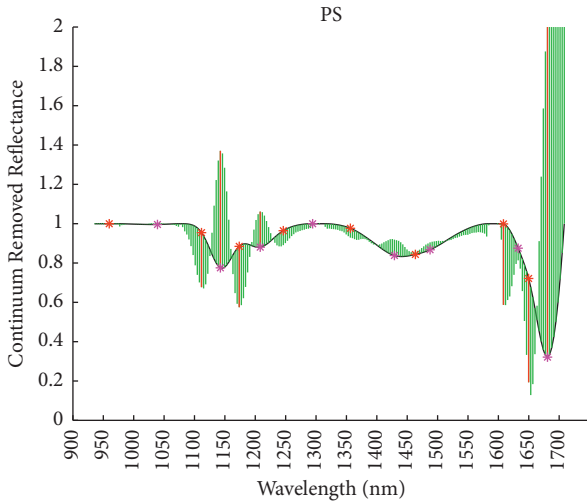


FIGURE 12: Continuum removed reflectance spectra and shape-based rule for PS.

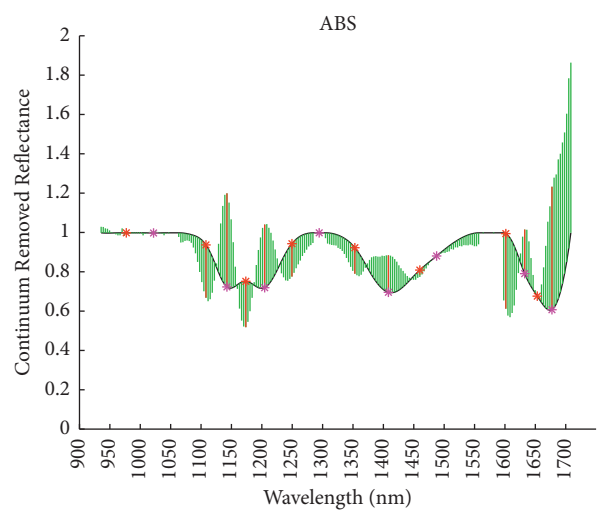


FIGURE 15: Continuum removed reflectance spectra and shape-based rule for ABS.

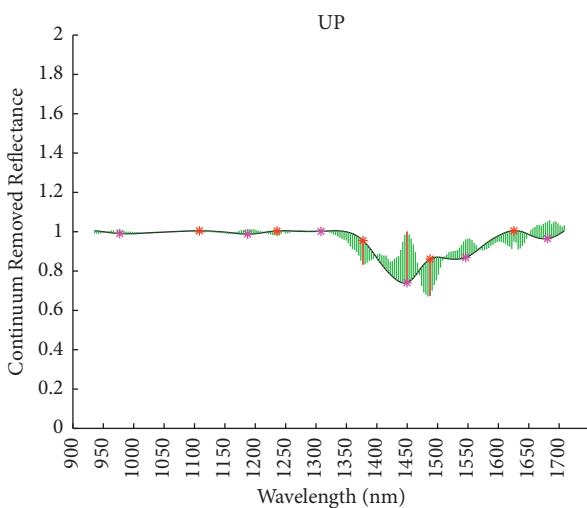


FIGURE 13: Continuum removed reflectance spectra and shape-based rule for UP.

UP) are not present. In addition to the plastic types mentioned so far, polypropylene (PP) and acrylonitrile butadiene styrene (ABS) were identified on the basis of spectra from [48, 49] and formulated as rules. Furthermore, an additional rule was established for the classification of printed circuit boards. A look at the classified plastic parts shows that the nonhomogeneity of the surfaces is partly reflected in the results in form of unclassified or misclassified pixels. Nevertheless, it can be stated that the classification was generally successful by using the developed ten rules. The black plastic parts (objects 1 and 17 in Figure 17) that are missing were not modelled in this example, because the used background also consists of black plastic and a separation in this special case proves to be difficult. Also, since the spectra of object number 5 in Figure 17 could not be assigned to any material, no rule for classification was established.

An important step in the recycling process is the separation of plastics and electronic waste. In particular, the high variation of different material compositions from

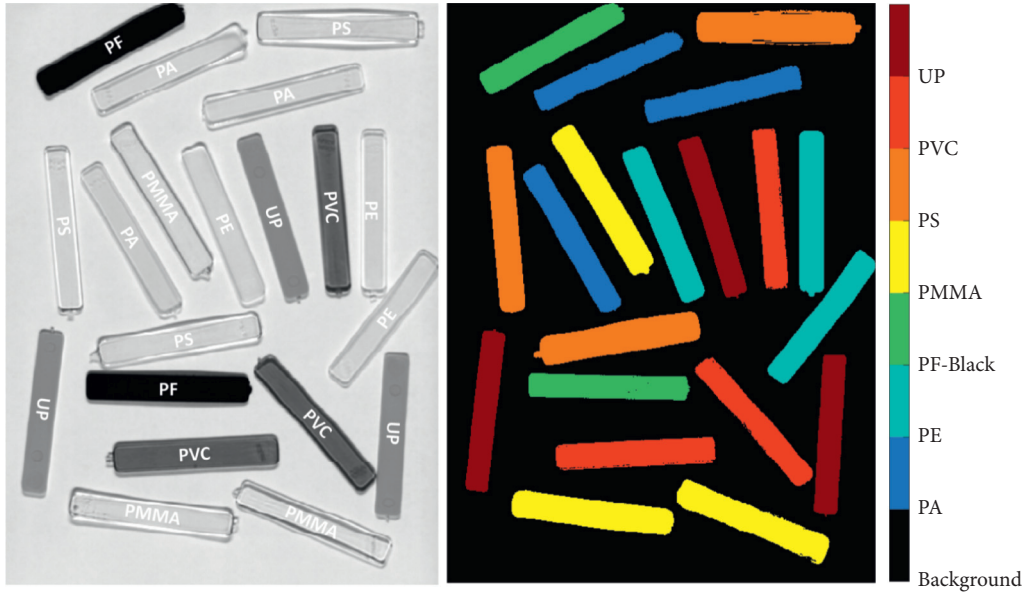


FIGURE 16: Classification result using developed rules.



FIGURE 17: Classification result for real waste (plastics and printed circuit boards) using developed rules.



which PCBs are made makes separation difficult. However, the example shown here also demonstrates that prior analysis of data and the use of rules based on acquired knowledge can lead to a more efficient recycling process. Due to the already mentioned high variation of PCBs and the composition of different materials (e.g., board, conductors, solder joints, resistors, and capacitors), a holistic assessment of a PCB is difficult to implement.

Nevertheless, it is evident in this example that the board has been identified in all cases and, as expected, only the areas with conductive tracks and metallic or electronic components have not been assigned and consequently ended up in a category with black plastics and the background. A significant difference to the samples shown in Figure 16 is the heterogeneity of the surfaces. Differences in depth, shadow effects on the surface, and dirt lead to a very high variability of spectral signatures. As an example, a region of object number 12 in Figure 17 was selected. In this area, differences in depth, shadow effects, and soiling can be found. A representation of the high variability in form of spectral shifts is shown in Figure 18. The shape-based classification result for object number 13, however, shows that a satisfactory result was achieved despite the high variability. The reason for the robustness of the proposed method is that even if the spectral values change and result in spectral shifts due to certain factors like dirt or shadow, the geometric shape remains stable.

The processing time using MATLAB on a machine with an Intel(R) Core(TM) i7-10700 CPU @ 2.90 GHz and 16 GB RAM was 136 seconds for the real waste dataset and 44 seconds for the plastic dataset from Figure 16.

In order to compare the quality of the proposed method to standard methods, we decided to use one of the most commonly used pixel-based classifiers. Therefore, the dataset shown in Figure 16 (plastic samples) and the dataset shown in Figure 17 (real waste samples) were processed using a supervised multiclass SVM classifier (C-SVC, One-Versus-All), which is based on the selection of a training set to obtain a model and the subsequent application of this model for the prediction of classes. The dataset with the real waste was limited to the plastic objects. As already described in [17], the quality of the results strongly depends on the chosen training samples and parameterization. For this reason, three different kernels, namely, Radial Basis Function (RBF), linear, and polynomial, with different parameters have been tested on the datasets. Due to the multicategory classification, a classification threshold is used, which is the maximum distance to the hypersurface for conducted classification to a specific class. The parameters for the kernels and the classification threshold were obtained by testing and checking the results. Because of the homogeneity of the plastic samples, one material sample of each material class was manually labelled and used as training set. A description of the datasets can be found in Table 2. The best results for the plastic dataset using different kernels and classification thresholds are shown in Figure 19.

The accuracy of each classification method compared is reported in Table 3. In principle, the results of this homogeneous and optimal dataset are comparable for most of

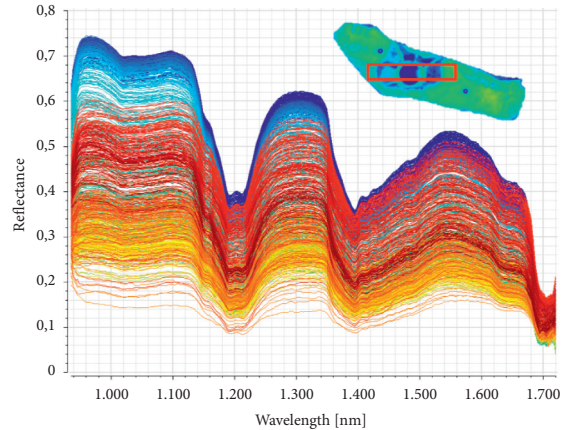


FIGURE 18: Spectral signature variability for object 12 of real waste dataset (spectra of all pixels inside the red box).

TABLE 2: Number of samples and training samples for each class.

	Class	Number of samples	Number of training samples
Plastic dataset	PA	16276	3783
	PE	16212	3659
	PF-Black	11157	4930
	PMMA	15779	4071
	PS	16877	4123
	PVC	16645	4869
	UP	16098	4613
Real waste dataset	PA	13343	2114
	PS	81238	10444
	PP	116620	11936
	ABS	11703	1638

classification metrics used. Nevertheless, significant differences can be seen in the resulting images, especially at the edge of objects, in form of unclassified or misclassified pixels. As already mentioned, this is due to a shadow effect that locally occurs in some areas, which impacts mainly the quality of the SVM methods. While the rule-based approach also classifies these edge areas and some shadow parts as corresponding material, the SVM results in misclassifications, especially using linear and polynomial kernels.

Compared to the plastic dataset, the real waste dataset has a typical situation consisting of objects with a high degree of variety. This not only increases the effort required to train the model, but also makes the training process more difficult. Small subsets of the individual samples were manually selected as training set. This strategy was considered to be appropriate due to the strong differences between the plastic samples. Thus, when selecting the subsets, special care was taken to ensure that particularly critical areas, such as shadow areas caused by depth, dirty areas, and damaged surfaces, were also covered in the training set. The best resulting class-labels are presented in Figure 20 and show also the effect of different classification thresholds. The corresponding numerical values are listed in Table 4.



FIGURE 19: Classification results using the (a) proposed shape-based method; SVM with (b) RBF kernel, C-value of 1, and classification threshold of 0.5; (c) linear kernel, C-value of 1, and classification threshold of 0.3; (d) polynomial kernel, C-value of 1, and classification threshold of 0.3.

In principle, also for this dataset it can be stated that, despite the challenging objects of the real waste dataset, a high degree of accuracy can be achieved for both methods. The best SVM result was obtained with an RBF kernel consisting of a C-Value of 0.2 and a classification threshold of 0.3. In comparison, the rule-based method produces

slightly better values for the different metrics. Also noticeable is the more homogeneous representation of objects in the resulting class-labels in Figure 20. The reason for this is again the robustness against spectral shifts. While SVM is based on pure reflectance values, the shape-based approach only needs to consider the geometric shape of spectral

TABLE 3: Classification accuracies for plastic dataset.

Metrics	Rule-based method	SVM (a)	SVM (b)	SVM (c)
Overall accuracy	0.9694	0.9647	0.9641	0.9641
Precision	0.9604	0.9731	0.9403	0.9403
Sensitivity	0.9551	0.9272	0.9623	0.9623
False positive Rate	0.0077	0.0114	0.0076	0.0076
F1-score	0.9564	0.9487	0.9503	0.9503
Kappa	0.8603	0.8384	0.8358	0.8359

(a) Solver: C-SVC; Classification type: one-versus-all; kernel: RBF; C-value: 1; threshold: 0.5. (b) Solver: C-SVC; classification type: one-versus-all; kernel: linear; C-value: 1; threshold: 0.3. (c) Solver: C-SVC; classification type: one-versus-all; kernel: polynomial; C-value: 1; threshold: 0.3.

signatures. One of the advantages of the shape-based approach is that, theoretically, only one spectrum per material is required for the rule formation. When using an SVM, it is necessary to ensure that, in case of high variability datasets, all data representing the spectral variations of one material class are included in the training set. Fulfilling this requirement can be a challenging task.

**3.2. Classification Results for Plants.** Compared to the dataset consisting of plastics and electronic waste, spectral signatures of plants often show a very similar curve shape. This fact generally complicates the classification process.

To illustrate the ability of classifying even in such difficult cases and the flexibility of the methodology presented in this work, rules were developed to distinguish between plant species based on the spectral signatures shown in Figure 21, which are coherent with spectral signatures provided in [50]. Due to the similarity of the shape, in this particular case, rules have been formulated using mainly Continuum Removed Reflectance Values (CRRV) as well as curvature parameters. This involved using the entire spectrum from 500 nm to 1700 nm to work out fine differences regarding the water content, the cell structure, and the pigments of used plants. In this context, it must be taken into account that the spectral properties of a plant species can also vary, as spectral signatures will be affected by factors such as species, variety, age, internal cell structure, environmental conditions, chemical composition, and nutrient content [51]. An example for such a variation is given in Figure 22 for moss. Although not implemented in this work, one possible way to deal with such heterogeneous spectra could be the addition of rules based on texture features or other local image features [52–54].

A total of six rules were developed to separate the ten samples in groups of moss, lichen, red sedum, green sedum, *Geranium robertianum*, and green leaves. The results in Figure 23 show that, in principle, a classification between different plant species is possible. For example, it can be clearly seen that red sedum is distinctly different from the other plant species due to its coloring. Green sedum also differs based on the cell structure, whereby the properties of this plant seem also partially reflected in moss, which can be relatively well distinguished from the other plants due to the spectral signatures in the NIR and the low reflectance around

550 nm. Furthermore, it was possible to classify the lichens in object 10.

The very strong spectral similarity between the leaves does not allow a clean separation between the individual species. Only *Geranium robertianum* differs due to its low reflectance in the range around 550 nm. In this context, it is important to note that the spectral signatures shown in Figure 21 only reflect the spectral signature of one pixel and that the recognizable differences, which refer to reflectance values or shape, are not consistently present due to the high variability within a plant species. Nevertheless, it can be stated also for this example that a rule-based approach is a flexible method and, depending on prior analytical efforts, has the potential to provide useful results.

## 4. Discussion

Results of a rule-based classification approach were presented, based on the shape of spectral signatures. The rules are established in a supervised way and are based not only on spectral values, but especially on parameters that describe the geometric form of a spectral signature. These parameters include the automatic determined curvature points (i.e., specific spectral values obtained through the 2nd derivative), curvature values, and curvature behavior. The effectiveness of this method was demonstrated with different datasets from completely different fields of application. In particular, the separation of materials with significant geometric differences in the course of the curve leads to convincing results, which is reflected in the Overall Accuracy (OA) of the plastic dataset with 96.94% and the real waste dataset with 98.42%.

An essential advantage over classical classification approaches is the possibility of describing the course of a spectrum in detail on the basis of a few selective parameters. Classic approaches do not require a previous analysis of the data and offer the advantage of automated processing based on the actual data. The prerequisite for this is, on the one hand, the availability of a large number of training data, which is not always readily available and could be a very time-consuming task (e.g., annotation), and preprocessing in the form of a dimensional reduction. This reduction is a selection of informative bands and often leads to a loss of information within the spectral signatures, because it is based on pure statistical reduction. With the method presented here, the finest changes in the course of the curve can be identified and modelled by using theoretically only one

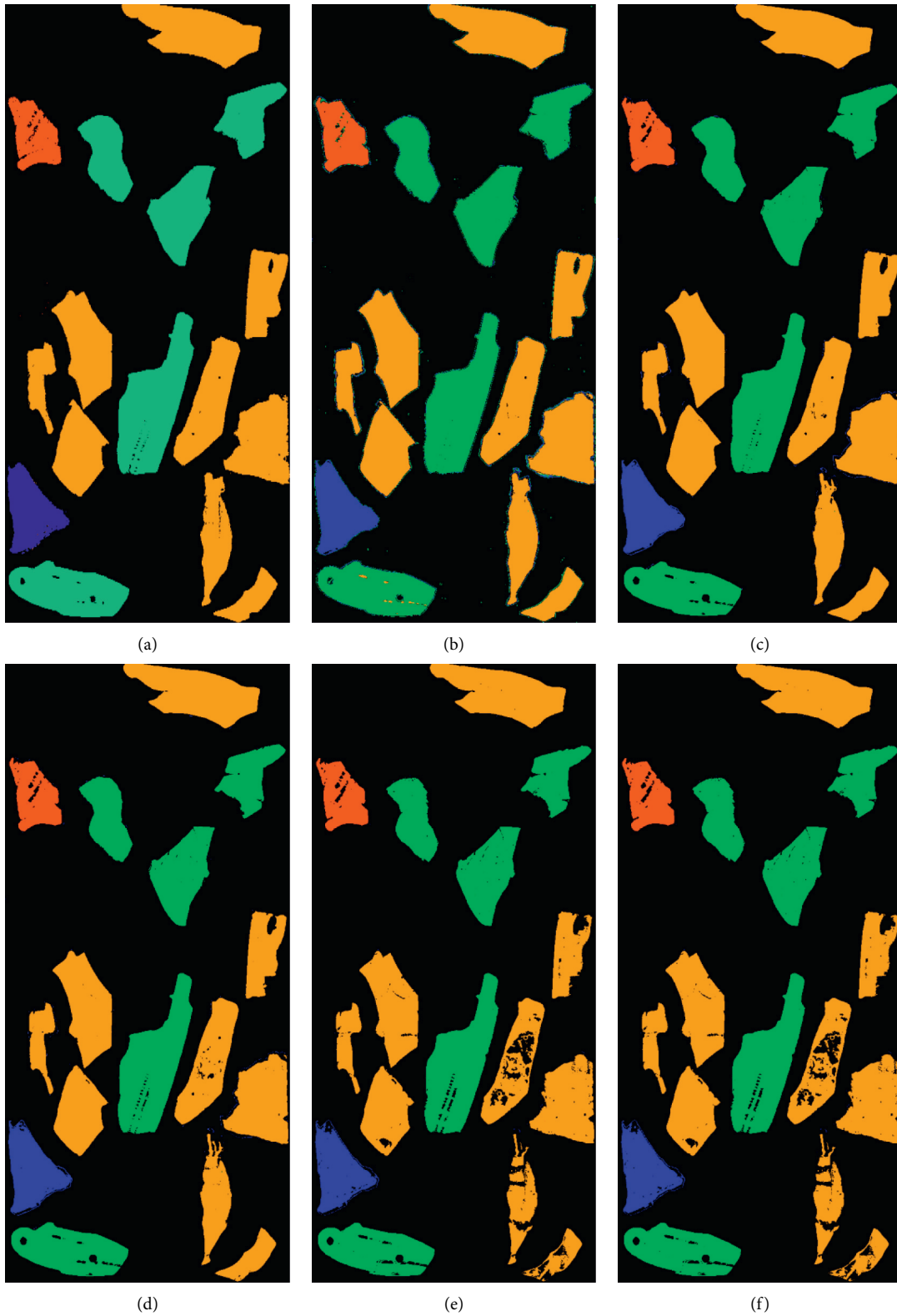


FIGURE 20: Classification results using (a) proposed shape-based method; SVM with (b) RBF kernel, C-value of 0.2, and classification threshold of 0.2; (c) RBF kernel, C-value of 0.2, and classification threshold of 0.3; (d) RBF kernel, C-value of 0.2, and classification threshold of 0.4; (e) RBF kernel, C-value of 0.2, and classification threshold of 0.5; (f) RBF kernel, C-value of 0.2, and classification threshold of 0.6.

TABLE 4: Classification results for real waste dataset.

Metrics	Rule-based method	SVM (a)	SVM (b)	SVM (c)	SVM (d)	SVM (e)
Overall accuracy	0.9842	0.9770	0.9826	0.9805	0.9760	0.9586
Precision	0.9842	0.9398	0.9698	0.9807	0.9856	0.9854
Sensitivity	0.9498	0.9477	0.9434	0.9353	0.9208	0.8831
False positive Rate	0.0075	0.0093	0.0098	0.0119	0.0152	0.0267
F1-score	0.9661	0.9396	0.9540	0.9556	0.9504	0.9290
Kappa	0.9506	0.9283	0.9456	0.9392	0.9249	0.8707

(a) Solver: C-SVC; classification type: one-versus-all; kernel: RBF; C-value: 0.2; threshold: 0.2. (b) Solver: C-SVC; classification type: one-versus-all; kernel: RBF; C-value: 0.2; threshold: 0.3. (c) Solver: C-SVC; classification type: one-versus-all; kernel: RBF; C-value: 0.2; threshold: 0.4. (d) Solver: C-SVC; classification type: one-versus-all; kernel: RBF; C-value: 0.2; threshold: 0.5. (e) Solver: C-SVC; classification type: one-versus-all; kernel: RBF; C-value: 0.2; threshold: 0.6.

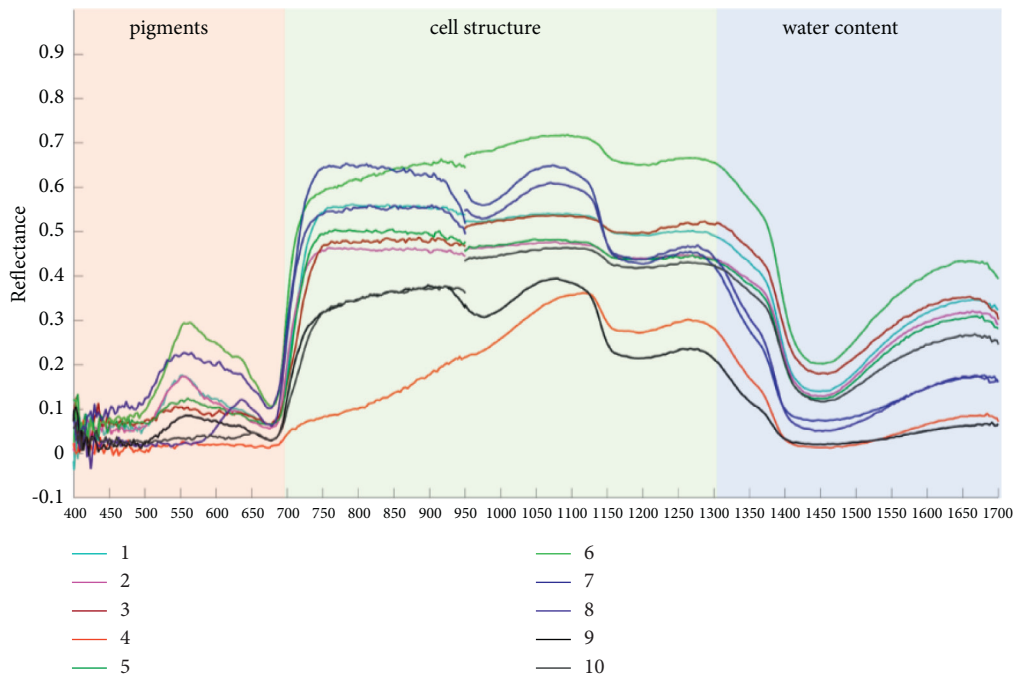


FIGURE 21: Spectra of measured plants. Two different cameras were used to acquire the spectra (Specim FX10 and Specim FX17). The discontinuities in the range of 900 nm and the greater noise in the VIS range can be explained by the differences in spectral sensitivity of the cameras. In particular, the range below 500 nm exhibits strong noise effects and was not considered in the rule formation. *Taraxacum* (1), *Inula conyza* (2), *Campanula rotundifolia* (3), moss (4), *Geranium robertianum* (5), *Asplenium trichomanes* (6), green sedum (7), red sedum (8), moss (9), and lichen (10).

spectrum instead of a large amount of training data. The prerequisite for this is a previous analysis of the spectral signatures and the establishment of the rules based on expert knowledge. This process may be considered time-consuming, but considering the parameters used, it does not require a particularly high investment of time.

An automated generation of the rules would also be conceivable, since the essentially used parameters such as curvature value, curvature point, and curvature behavior are determined by simple mathematical methods. A disadvantage that can arise with the automatic formulation of rules is that the number of rules increases with the number of categories. This can lead to complex rules that conflict with each other. Modelling the rules based on expert knowledge can avoid this by understanding the interaction of the factors, which makes it possible to formulate the simplest concept that gives the best results. Basically, when dealing

with spectral data, it can be stated that the number of categories that can be classified from spectral bands is small enough that a limited number of rules can be used. The analysis of spectral signatures in advance has the further advantage that there is no need to use the entire spectrum for classification. If significant shape differences are detected within a limited spectral range, it is sufficient to consider only this range in order to establish the rules based on it, which leads to an increase in performance.

While spectral values can vary strongly depending on different factors, differences in the shape of spectral signatures are only due to differences in material composition. This fact is the reason for the promising results and confirm our expectations of the proposed analytical method, which is robust to spectral variations and allows an unambiguous description of the spectra based on the basic idea of using the shape of the individual fingerprint of different materials or objects.



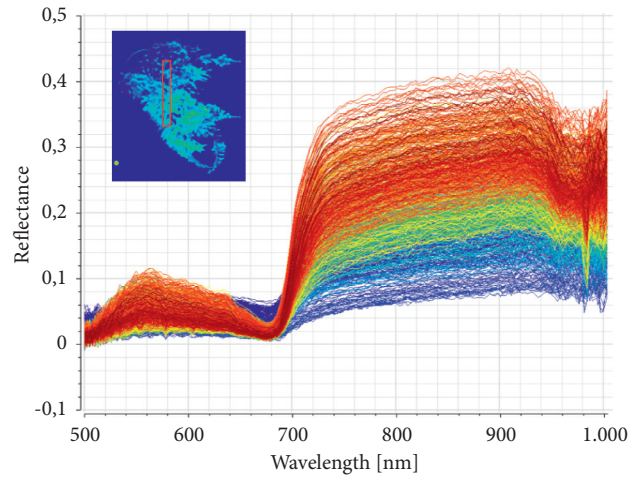


FIGURE 22: Spectral variations in VIS for one type of moss regarding all pixels in the red box.

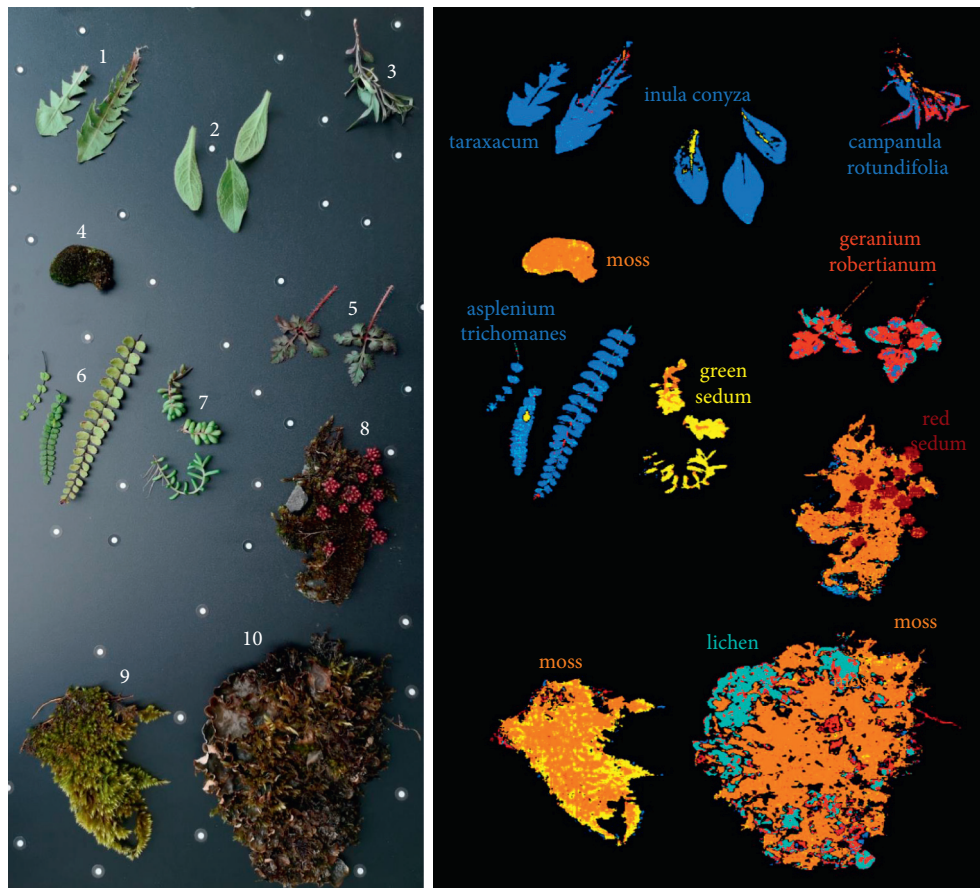


FIGURE 23: Classification result for plants using VIS and NIR (500 nm–1700 nm).

Furthermore, the processing time must be mentioned. The pixel-by-pixel processing and checking for conditions using MATLAB on a machine with an Intel(R) Core(TM) i7-10700 CPU @ 2.90 GHz and 16 GB RAM takes 113.42 seconds for the plant dataset (1220 × 640 pixels and 462 bands) shown in Figure 23. In comparison, the shape-based

classification in [34] for a dataset of 512 × 512 pixels and 6 bands takes about 9 minutes (computer configuration: CPU 2.93 GHz and installed memory 4.00 GB). Another comparison with a research using a classic classification method like SVM also shows a better performance. The processing time for a comparable dataset of 1168 × 696 pixels and 520

bands takes 2980.21 seconds using 30 bands and 9381.37 seconds using all bands of the dataset (computer configuration: Intel Core i7-6800k 3.40-GHZ CPU and installed memory 64 GB) [55].

## 5. Conclusions

Considering the shape of spectral signatures and describing them by curvature parameters and spectral values proves to be a convincing supervised method for classifying diverse groups of materials from different fields of application. Further evaluation on publicly available datasets, e.g., from the field of remote sensing, also with respect to other classification methods, would be of interest. A possible disadvantage of this methodology could be the necessary prior analysis of the data. This circumstance could be eliminated by automation. Especially in the case of significantly different spectra, like the plastic samples, automation is certainly feasible and would bring the benefit of an unsupervised method.

## Data Availability

The HSI datasets used to support the findings of this study have been deposited in the Zenodo repository (<https://doi.org/10.5281/zenodo.5068201>).

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was funded by the European Union from the European Regional Development Fund and the State of Rhineland-Palatinate. The authors thank Pellenc ST for supplying the real waste materials.

## References

- [1] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: a review," *IEEE Access*, vol. 6, pp. 14118–14129, 2018.
- [2] E. Honkavaara, M. A. Eskelinen, I. Polonen et al., "Remote sensing of 3-D geometry and surface moisture of a peat production area using hyperspectral frame cameras in visible to short-wave infrared spectral ranges onboard a small unmanned airborne vehicle (UAV)," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 9, pp. 5440–5454, 2016.
- [3] P. Vermeulen, P. Flémal, O. Pigeon, P. Dardenne, J. F. Pierna, and V. Baeten, "Assessment of pesticide coating on cereal seeds by near infrared hyperspectral imaging," *Journal of Spectral Imaging*, vol. 6, 2017.
- [4] R. Roscher, J. Behmann, A.-K. Mahlein, J. Dupuis, H. Kuhlmann, and L. Plümer, "Detection of disease symptoms on hyperspectral 3D plant models," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. III-7, pp. 89–96, 2016.
- [5] M. Al-Sarayreh, M. M. Reis, W. Qi Yan, and R. Klette, "Detection of red-meat adulteration by deep spectral-spatial features in hyperspectral images," *Journal of Imaging*, vol. 4, no. 5, p. 63, 2018.
- [6] I. Pineda, M. D. Nur Alam, and O. Gwun, "Calyx and stem discrimination for apple quality control using hyperspectral imaging," in *Communications in Computer and Information Science (Technology Trends-CITT)*, M. Botto-Tobar, G. Pizarro, M. Zúñiga-Prieto et al., Eds., vol. 895, pp. 274–287, Springer, Berlin, Germany, 2018.
- [7] Q. Wang and Y. He, "Rapid and nondestructive classification of Cantonese sausage degree using hyperspectral images," *Applied Sciences*, vol. 9, no. 5, p. 822, 2019.
- [8] M. Halicek, H. Fabelo, S. Ortega, G. M. Callico, and B. Fei, "In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer," *Cancers*, vol. 11, no. 6, 2019.
- [9] H. Ding and R. C. Chang, "Hyperspectral imaging with burn contour extraction for burn wound depth assessment," *Journal of Engineering and Science in Medical Diagnostics and Therapy*, vol. 1, 2018.
- [10] M. Wahabzada, M. Besser, M. Khosravani et al., "Monitoring wound healing in a 3D wound model by hyperspectral imaging and efficient clustering," *PLoS One*, vol. 12, no. 12, Article ID e0186425, 2017.
- [11] A. Polak, T. Kelman, P. Murray et al., "Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication," *Journal of Cultural Heritage*, vol. 26, pp. 1–11, 2017.
- [12] F. Daniel, A. Mounier, J. Pérez-Arategui et al., "Hyperspectral imaging applied to the analysis of Goya paintings in the Museum of Zaragoza (Spain)," *Microchemical Journal*, vol. 126, pp. 113–120, 2016.
- [13] C. Cucci, J. K. Delaney, and M. Picollo, "Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts," *Accounts of Chemical Research*, vol. 49, no. 10, pp. 2070–2079, 2016.
- [14] A. Majda, R. Wietecha-Posłuszny, A. Mendys, A. Wójtowicz, and B. Lydzba-Kopczyńska, "Hyperspectral imaging and multivariate analysis in the dried blood spots investigations," *Applied Physics A*, vol. 124, 2018.
- [15] S. Cadd, B. Li, P. Beveridge, W. O'Hare, and M. Islam, "Age determination of blood-stained fingerprints using visible wavelength reflectance hyperspectral imaging," *Journal of Imaging*, vol. 4, 2018.
- [16] G. J. Edelman, E. Gaston, T. G. van Leeuwen, P. J. Cullen, and M. C. Aalders, "Hyperspectral imaging for non-contact analysis of forensic traces," *Forensic Science International*, vol. 223, pp. 28–39, 2012.
- [17] P. Ghamisi, J. Plaza, Y. Chen, J. Li, and A. J. Plaza, "Advanced spectral classifiers for hyperspectral images: a review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 8–32, 2017.
- [18] J. Ren, R. Wang, G. Liu, R. Feng, Y. Wang, and W. Wu, "Partitioned relief-F method for dimensionality reduction of hyperspectral images," *Remote Sensing*, vol. 12, no. 7, p. 1104, 2020.
- [19] X. Luo, Z. Shen, R. Xue, and H. Wan, "Unsupervised band selection method based on importance-assisted column subset selection," *IEEE Access*, vol. 7, pp. 517–527, 2019.
- [20] F. Luo, L. Zhang, B. Du, and L. Zhang, "Dimensionality reduction with enhanced hybrid-graph discriminant learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 8, pp. 5336–5353, 2020.
- [21] S. Li, Q. Hao, X. Kang, and J. A. Benediktsson, "Gaussian pyramid based multiscale feature fusion for hyperspectral

- image classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 9, pp. 3312–3324, 2018.
- [22] D. Xiaohui, L. Huapeng, L. Yong, Y. Ji, and Z. Shuqing, “Comparison of swarm intelligence algorithms for optimized band selection of hyperspectral remote sensing image,” *Open Geosciences*, vol. 12, no. 1, pp. 425–442, 2020.
- [23] Y. Wang, L. Wang, H. Xie, and C.-I. Chang, “Fusion of various band selection methods for hyperspectral imagery,” *Remote Sensing*, vol. 11, no. 18, p. 2125, 2019.
- [24] Q. Wang, F. Zhang, and X. Li, “Optimal clustering framework for hyperspectral band selection,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. 1–13, 2018.
- [25] R. M. Torres, P. W. T. Yuen, C. Yuan, J. Piper, C. McCullough, and P. Godfree, “Spatial spectral band selection for enhanced hyperspectral remote sensing classification applications,” *Journal of Imaging*, vol. 6, no. 9, p. 87, 2020.
- [26] W. Sun and Q. Du, “Hyperspectral band selection: a review,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 2, pp. 118–139, 2019.
- [27] F. Luo, L. Zhang, and X. Zhou, “Sparse-adaptive hypergraph discriminant analysis for hyperspectral image classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 6, pp. 1082–1086, 2019.
- [28] A.-K. Mahlein, T. Rumpf, P. Welke et al., “Development of spectral indices for detecting and identifying plant diseases,” *Remote Sensing of Environment*, vol. 128, pp. 21–30, 2013.
- [29] R. Meng, Z. Lv, J. Yan et al., “Development of spectral disease indices for southern corn rust detection and severity classification,” *Remote Sensing*, vol. 12, no. 19, p. 3233, 2020.
- [30] A. Huete, C. Justice, and H. Liu, “Development of vegetation and soil indices for MODIS-EOS,” *Remote Sensing of Environment*, vol. 49, no. 3, pp. 224–234, 1994.
- [31] J. Rouse, R. H. Haas, D. Deering, J. A. Schell, and J. Harlan, “Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. [great plains corridor],” 1973.
- [32] X. Guo and P. Li, “Mapping plastic materials in an urban area: development of the normalized difference plastic index using WorldView-3 superspectral data,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 214–226, 2020.
- [33] F. Boochs, G. Kupfer, K. Dockter, and W. Kühbauch, “Shape of the red edge as vitality indicator for plants,” *International Journal of Remote Sensing*, vol. 11, no. 10, pp. 1741–1753, 1990.
- [34] Y. Chen, Q. Wang, Y. Wang, S.-B. Duan, M. Xu, and Z.-L. Li, “A spectral signature shape-based algorithm for landsat image classification,” *ISPRS International Journal of Geo-Information*, vol. 5, no. 9, p. 154, 2016.
- [35] F. A. Kruse, “Spectral-feature-based analysis of reflectance and emission spectral libraries and imaging spectrometer data,” *SPIE Proceedings, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery*, vol. 18, pp. 83901F–83910F, 2012.
- [36] S. Andrés, D. Arvor, I. Mougnot, T. Libourel, and L. Durieux, “Ontology-based classification of remote sensing images using spectral rules,” *Computers & Geosciences*, vol. 102, pp. 158–166, 2017.
- [37] T. Berhane, C. Lane, Q. Wu et al., “Decision-tree, rule-based, and random forest classification of high-resolution multi-spectral imagery for wetland mapping and inventory,” *Remote Sensing*, vol. 10, no. 4, p. 580, 2018.
- [38] W. Cui, M. Yao, Y. Hao et al., “Knowledge and geo-object based graph convolutional network for remote sensing semantic segmentation,” *Sensors*, vol. 21, no. 11, p. 3848, 2021.
- [39] G. Ghazaryan, O. Dubovyk, F. Löw et al., “A rule-based approach for crop identification using multi-temporal and multi-sensor phenological metrics,” *European Journal of Remote Sensing*, vol. 51, no. 1, pp. 511–524, 2018.
- [40] P. F. Houhoulis and W. Michener, “Detecting wetland change: a rule-based approach using NWI and SPOT-XS data,” *Photogrammetric Engineering & Remote Sensing*, vol. 66, pp. 205–211, 2000.
- [41] S. Liu and Q. Shi, “Multitask deep learning with spectral knowledge for hyperspectral image classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 12, pp. 2110–2114, 2020.
- [42] R. N. Clark, G. A. Swayze, K. E. Livo et al., “Imaging spectroscopy: earth and planetary remote sensing with the USGS Tetracorder and expert systems,” *Journal of Geophysical Research: Planets*, vol. 108, p. E12, 2003.
- [43] R. N. Clark and T. L. Roush, “Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications,” *Journal of Geophysical Research: Solid Earth*, vol. 89, no. B7, pp. 6329–6340, 1984.
- [44] P.-N. Tan, M. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining*, Pearson, New York, NY, USA, 2019.
- [45] J.-J. Ponciano, M. Roetner, A. Reiterer, and F. Boochs, “Object semantic segmentation in point clouds-comparison of a deep learning and a knowledge-based method,” *ISPRS International Journal of Geo-Information*, vol. 10, no. 4, p. 256, 2021.
- [46] R. Dvorak, E. Kosior, and L. Moody, “Development of NIR detectable black plastic packaging,” 2011, [http://www.wrap.org.uk/sites/files/wrap/Recyclability\\_of\\_Black\\_Plastic\\_Summary.pdf](http://www.wrap.org.uk/sites/files/wrap/Recyclability_of_Black_Plastic_Summary.pdf).
- [47] O. Rozenstein, E. Puckrin, and J. Adamowski, “Development of a new approach based on midwave infrared spectroscopy for post-consumer black plastic waste sorting in the recycling industry,” *Waste Management*, vol. 68, pp. 38–44, 2017.
- [48] S. Zhu, H. Chen, M. Wang, X. Guo, Y. Lei, and G. Jin, “Plastic solid waste identification system based on near infrared spectroscopy in combination with support vector machine,” *Advanced Industrial and Engineering Polymer Research*, vol. 2, no. 2, pp. 77–81, 2019.
- [49] X. Wu, J. Li, L. Yao, and Z. Xu, “Auto-sorting commonly recovered plastics from waste household appliances and electronics using near-infrared spectroscopy,” *Journal of Cleaner Production*, vol. 246, Article ID 118732, 2020.
- [50] N. Y. Kiang, J. Siefert, Govindjee, and R. E. Blankenship, “Spectral signatures of photosynthesis. I. Review of Earth organisms,” *Astrobiology*, vol. 7, no. 1, pp. 222–251, 2007.
- [51] D. de Carvalho Lopes and A. J. Steidle Neto, “Classification and authentication of plants by chemometric analysis of spectral data,” *Vibrational Spectroscopy for Plant Varieties and Cultivars Characterization*, Elsevier, vol. 80, pp. 105–125, , Amsterdam, Netherlands, 2018.
- [52] T. Beghin, J. S. Cope, P. Remagnino, and S. Barman, “Shape and texture based plant leaf classification,” *Advanced Concepts for Intelligent Vision Systems-ACIVS (Lecture Notes in Computer Science)*, Springer, vol. 6475, pp. 345–353, , Berlin, Germany, 2010.
- [53] P. Salve, P. Yannawar, and M. Sardesai, “Multimodal plant recognition through hybrid feature fusion technique using imaging and non-imaging hyper-spectral data,” *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 1, pp. 1361–1369, 2018.
- [54] A. C. Siravenha and S. R. Carvalho, “Plant classification from leaf textures,” in *Proceedings of the 2016 International Conference on Digital Image Computing: Techniques and*

*Applications (DICTA)*, pp. 1–8, Gold Coast, Australia, November 2016.

- [55] Q. Wang, F. Zhang, and X. Li, “Hyperspectral band selection via optimal neighborhood reconstruction,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 12, pp. 8465–8476, 2020.