Using ATR-FTIR Spectra and Convolutional Neural Networks for Characterizing Mixed Plastic Waste

Shengli Jiang,[1] Zhuo Xu,[2] Medhavi Kamran,[2] Stas Zinchik,[2] Sidike Paheding,[3] Armando G. McDonald,[4] Ezra Bar-Ziv,[2] Victor M. Zavala[1]

[1] Department of Chemical and Biological Engineering, University of Wisconsin-Madison
[2] Department of Mechanical Engineering, Michigan Technological University
[3] Department of Applied Computing, Michigan Technological University
[4] Department of Forest, Rangeland and Fire Sciences, University of Idaho

Abstract

We present a convolutional neural network (CNN) framework for classifying different types of plastic materials that are commonly found in mixed plastic waste (MPW) streams. The CNN framework uses experimental ATR-FTIR (attenuated total reflection-Fourier transform infrared spectroscopy) spectra to classify ten different plastic types. An important aspect of this type of spectral data is that it can be collected in real-time; as such, this approach provides an avenue for enabling the high-throughput characterization of MPW. The proposed CNN architecture (which we call PlasticNet) uses a Gramian angular representation of the spectra. We show that this 2-dimensional (2D) matrix representation highlights correlations between different frequencies (wavenumber) and leads to significant improvements in classification accuracy, compared to the direct use of spectra (a 1D vector representation). We also demonstrate that PlasticNet can reach an overall classification accuracy of over 87% and can classify certain plastics with 100% accuracy. Our framework also uses saliency maps to analyze spectral features that are most informative.

Keywords

machine learning; plastic waste; IR spectra; classification; real-time

1. Introduction

Plastics are essential materials that are used in a wide range of applications such as food packaging, construction, transportation, health care, and electronics. Since 1856 (when the first plastic celluloid was invented), the plastics industry has grown rapidly not only in terms of volume, but also in terms of the variety of materials produced. This rapid expansion has resulted in a massive environmental footprint; to give some perspective, in 2015, nearly 381 million tons of mixed plastic waste (MPW) were produced, this is more than the total weight of humans on earth (316 million tons). Notably, only 20% of all plastics produced were recycled (Ritchie and Roser, 2018); this recycling rate is notably low compared to that of other materials (e.g., aluminum has a recycling rate of nearly 100%). Most MPW end up in landfills and incinerators; landfills are unsustainable, especially when land availability is constrained (Abdel-Shafy and Mansour, 2018). MPW incineration reduces the need for landfills, but this process can release hazardous substances into the atmosphere (Hopewell et al., 2009).

MPW recycling is essential for mitigating the environmental impact of plastics, but this practice faces many obstacles (Schlesinger, 2013). Most of the recycled plastic is reprocessed into downgraded products (of a lower value); for instance, plastics used for food packaging are often converted into cheaper building materials such as plastic lumber (Awoyera and Adesina, 2020). In other words, recycled plastic products are less valuable and thus there are limited incentives to produce them. Another key factor that hinders plastic recycling is our limited ability to effectively characterize and sort MPW streams (which can be quite complex) (Milios et al., 2018). Traditionally, plastic components in MPW can only be partially identified based on techniques such as coding, density differences, and froth-flotation (Gundupalli et al., 2017). These

technologies are easy to implement but are low-throughput and have several other limitations (Zhu et al., 2019); for example, density separation in water can effectively separate polypropylene (PP) and polyethylene (PE) from polyvinyl chloride (PVC), polyethylene-terephthalate (PET), and polystyrene (PS); however, PVC cannot be removed from PET in this manner because their density ranges overlap (Hopewell et al., 2009). Automated sorting with high-throughput, high-accuracy, and low-labor is necessary for effective MPW management.

Recent innovations in recycling technology include increasingly reliable detection instruments and improved materials identification algorithms; these have improved the accuracy and productivity of automated sorting. Methods such as spectroscopy, hyperspectral imaging (HSI), ultrasonic techniques, X-ray diffraction (XRD), thermal imaging or infrared imaging, combined with machine learning (ML) algorithms, have been successful in accurately identifying plastics that are commonly found in MPW (da Silva and Wiebeck, 2020; Karlsson et al., 2016; Siddiqui et al., 2008; Signoret et al., 2020, 2019; Wu et al., 2015). Michele et al. analyzed four different spectroscopic methods with various machine learning (ML) algorithms, such as k-nearest neighbors (KNN), linear discriminant analysis (LDA), and support vector machines (SVM), to identify marine plastic debris and consumer plastic (Michel et al., 2020). Among the four spectroscopic methods, the attenuated total reflection-Fourier transform infrared spectroscopy (ATR-FTIR) technique performed best, with an accuracy of 89-98%. Da Silva et al. developed a method to identify nine different types of plastics, including polyamide (PA) and polycarbonate (PC), based on µFTIR hyperspectral imaging and ML (Da Silva et al., 2020). Roh et al. used laser-induced breakdown spectroscopy with an algorithm-based radial basis function neural network to identify black plastics, including PP, PS, and acrylonitrile-butadiene-styrene (ABS), and achieved an accuracy

3

of over 95% (Roh et al., 2018). Wu et al. proposed an automated sorting system using near-infrared spectroscopy to identify waste from electronic and electrical equipment (WEEE) (Allen et al., 1999). Gundupalli et al. used a thermal imaging camera integrated with a robotic manipulator to classify recyclable materials in MSW and achieved a 90% accuracy (Gundupalli et al., 2017). While these results are highly encouraging, these methods are slow and low-throughput and are not tailored to real-time sorting (they rely on manual sample collection).

ATR-FTIR can analyze plastic components found in MPW in real-time; as such, one can envision the development of fast, online ML techniques that can analyze ATR-FTIR spectra to characterize MPW streams. Recently, ML methods such as convolutional neural networks (CNNs) have been used to analyze spectral data (Ng et al., 2019). A key advantage of CNNs over other ML methods is their ability to automatically extract and organize discriminative features directly from raw data (without the need to pre-compute hand-crafted features). The training of powerful CNN models can be facilitated by the availability of advanced computing hardware (e.g., GPUs) and of vast data streams found in online systems. The integration of online ATR-FTIR and CNNs thus provides a potential avenue to sort plastic waste with high accuracy and throughput in real-time.

In this work, we propose a computational framework to characterize plastic components of MPW by analyzing ATR-FTIR spectra using CNNs. Experimental data was obtained by preparing small sheets of plastics of different shapes and used ATR-FTIR to scan sheets for 10 different types; this data collection approach mimics how rigid waste plastics are found in online processing of MPW streams. The proposed framework uses CNNs to analyze the spectra and sort/classify plastic components. The spectra collected can be represented as 1D vectors and analyzed by using 1D

CNNs (Chen et al., 2019). The 1D CNN extracts features of a spectrum by convolving it with different filters. A limitation of this approach, however, is that it might fail to capture correlations across frequencies (wavenumbers which may compromise the prediction accuracy). To deal with this issue, we present a new data representation that captures signal correlation information; specifically, we represent a spectrum as Gramian angular fields (GAFs). GAFs are matrices (2D data objects) that can be analyzed using 2D CNNs (Wang and Oates, 2015) and these data objects can better capture spectral correlations. A problem with this approach, however, is that the training of 2D CNNs is significantly more computationally expensive than that of 1D CNNs. To ameliorate this issue, we use a Piecewise Aggregate Approximation (PAA) approach to reduce the dimension of the input GAF matrices (Keogh and Pazzani, 2000). This framework also uses saliency analysis (Sundararajan et al., 2017) to understand the most important features of spectra that can help identify different plastic components. We demonstrate that this CNN framework (which we call PlasticNet) can reach an overall classification accuracy of over 87% and can classify certain plastics with 100% accuracy. The conjunction of ATR-FTIR and CNN creates a powerful, low-cost, and rapid method for analyzing the composition of plastic waste and enables future recycling and reproduction of high-quality plastics.

2. Experimental Data Collection and Preparation

The dataset studied included ATR-FTIR spectra for 10 different, commercially-available plastic materials (see **Figure 1**). These include thermoplastic polymers, natural, and synthetic rubber that are common in the MPW. Specifically, these were ABS, acrylic (AC), PE, PET, polybutadiene (BR), polycarbonate (PC), polyisoprene (PI), PS, PP, and PVC. The spectra were collected using a Thermo Scientific, Nicolet-iS5 FTIR spectrometer equipped with an attenuated total reflectance

(ATR) accessory (ZnSe crystal, iD5), taken with 64 scans with 4 cm^{-1} resolution between 2000 to 4000 cm^{-1}. Spectral data was collected using Omnic v9.8 software, and then extracted using TQ analyst EZ software (Thermo Nicolet) and compiled for analysis.

The plastics purchased consist of different shapes; round-shaped beads were cut into less than 1 mm thickness and converted into flattened thin sheets (10 mm ×10 mm). For each plastic sample, 50 spectra were measured as the training set for the ML algorithms, and 20 spectra were used as the testing set. To obscure the spectra, the sample was not kept in close contact with the crystal and each spectrum was taken with only one scan. The background was repeated after every measurement with 64 scans. For each plastic, 70 online measurements of were obtained. Each spectrum had 4150 data points, where each point represents the intensity at a given wavenumber (cm^{-1}). Each spectrum is encoded in a vector in $\mathbb{R}^{4150}$. For the 10 types of plastics, a total of 700 IR spectra were obtained. For the analysis, all spectra were normalized to be in the range [0, 1]:

$$\hat{x} = \frac{x - \min{(X)}}{\max(X) - \min{(X)}},$$

where $x \in \mathbb{R}^{4150}$ is the original vector (a raw spectrum), $\hat{x} \in \mathbb{R}^{4150}$ is the normalized vector (normalized spectrum), and $X \in \mathbb{R}^{700 \times 4150}$ is a matrix obtained by stacking all raw spectra. By stacking all the normalized vectors, we obtain the normalized spectra matrix $\hat{X} \in \mathbb{R}^{700 \times 4150}$, which is randomly partitioned into a training set and a test set. The training set is the dataset used in the learning process to fit the parameters of the ML models. The test set is a dataset that is independent of the training set and is used to examine the performance (accuracy) of the ML model. A total of 30% of the elements of the training set were randomly selected and used as the validation set for tuning the ML architecture and preventing overfitting (Ng and Ng, 1997). For validation, we use a five-fold cross-validation approach; here, the original dataset is randomly split into five

subsets of equal size. Among the five subsets, a subset was retained as the test set, and the remaining four subsets were used as training data. The cross-validation process was repeated five times, with each of the five subsets used exactly once as test data. A schematic of the five-fold cross-validation process is shown in **Figure 2**. The training, validation, and test set consist of 392, 168, and 140 spectra, respectively. A stratification was implemented to ensure that each fold represents all strata of the data. That is, in each fold, each plastic type accounts for 10% of the data in the training and test sets. The final reported accuracy is the average of all accuracies of the five folds. The model is robust and generalizable if the test sets of each fold have similar accuracy.

The types of plastic (labels) that need to be predicted by the ML models are one-hot encoded. Specifically, each label can be represented by a vector of size 10 (only one entry in the vector is 1 and all other entries are 0). This vector representation is necessary to calculate the loss of categorical cross-entropy in our ML models.

3. Computational Framework

The proposed framework includes a CNN architecture, that we called PlasticNet; this architecture can process spectra as vectors (1D data objects); as such, PlasticNet can operate as a 1D CNN. The framework also includes a Gamian angular transformation method that transforms the spectra vectors into GAF matrices (2D objects); as such, PlasticNet can also operate as a 2D CNN. The framework also includes saliency analysis techniques, which are useful tools that allow us to understand features that the CNN might be searching for in the spectra in classifying plastic types.

3.1. 1D CNN

IR spectra vectors can be analyzed directly with a 1D CNN; 1D CNNs are widely in applications such as electrocardiography (Kiranyaz et al., 2016), near-infrared spectroscopy (Chen et al., 2019), and optimal control (Jiang and Zavala, 2021). The architecture of the proposed 1D CNN is shown in **Figure 3**. 1D CNNs extract and summarize features from spectra using convolution and pooling operations. In our architecture, each convolution filter is a vector of size three. The output of a convolution operation is a single scalar value that marks the presence (high value) or absence (low value) of the pattern the filter is trying to identify or highlight. A single convolution operation maps a given vector to another vector of the same dimension after a nonlinear transformation (i.e., rectified linear unit). In the architecture used, a set of these filters is referred as a convolutional layer. Convolutional operations greatly increase the amount of information that needs to be proposed; therefore, it is necessary to summarize such information. In our architecture, we use a max-pooling layer to reduce dimensionality. A max-pooling operation takes a subset of a given vector, in this case a part of size two, and reduces it to a single value by extracting only the maximum value. This greatly reduces the dimensionality of the vectors created by the convolutional layer and distills the important information extracted by the convolutional filters.

An IR vector of size 4150 is fed into the 1D CNN, which we call PlasticNet (1D). This architecture contains four convolutional layers, two max-pooling layers, and three fully-connected layers. This simple architecture achieves high accuracy and facilitates fast training. The convolutional layer has 64 filters of size 3 and the max-pooling layer has filters of size 2. Each of the fully-connected layers has 64 nodes and the activation functions between layers are rectified linear units (ReLUs). Between each of two fully-connected layers is a dropout layer with a dropout ratio of 0.2 to prevent

overfitting. The output layer uses a SoftMax activation function to perform classification. The output for plastic classification is a vector of dimension 10, corresponding to the probability of the IR spectra being from a specific type of plastic. The loss function coupled with the SoftMax activation function is the categorical cross-entropy. In the proposed CNN architecture, convolutional layers and max-pooling layers are performed recursively. The idea behind this recursion is to extract information at both local and global scales while condensing it so that simple classification can be performed, and the corresponding plastic types can be predicted. A recent review on fundamentals of CNNs can be found in Jiang and Zavala, 2021.

3.2. Gramian Angular Fields

Although the vector representation of IR spectra already carries rich information, the correlation between different frequencies is not explicitly encoded in the vector representation and this is difficult to extract using convolution operations. Recently, Gramian Angular fields (GAF) have been used to encode time-series objects into matrices that capture correlation structures and that are processed using 2D CNNs; this data transformation approach has been shown to improve classification accuracy (Wang and Oates, 2015). Our hypothesis was that a similar principle can be applied to IR spectra to improve prediction accuracy. A GAF represents vectors in a polar coordinate system and converts these angles into symmetric matrices using various operations. There are two types of GAFs: Gramian Angular Summation fields (GASF) and Gramian Angular Difference fields (GADF). Each element of GASF and GADF is the cosine of the sum and the sine of the difference of the angles, respectively. The first step in constructing the GAF matrix is to normalize the spectral data to a value between 0 and 1. After normalization, the second step is to

represent the normalized vector $\hat{x}$ in a polar coordinate system by using the following transformations:

$$\phi_i = \arccos(\hat{x}), i = 1, \dots, 4150$$

$$r_i = \frac{i}{4150}, i = 1, \dots, 4150,$$

where i is the index of the vector entry, $\phi \in \mathbb{R}^{4150}$ is the angle vector, and $r \in \mathbb{R}^{4150}$ is the radius vector. Finally, the GASF and GADF matrices are obtained as:

$$GASF = \cos(\phi_i + \phi_j) = \hat{x}^T \hat{x} - \sqrt{I - \hat{x}^2}^T \sqrt{I - \hat{x}^2}$$

$$GADF = \sin(\phi_i - \phi_j) = \sqrt{I - \hat{x}^2}^T \hat{x} - \hat{x}^T \sqrt{I - \hat{x}^2}$$

where $I = [1, \dots, 1]$ is a unit row vector of size 4150.

The resulting $GASF, GADF \in \mathbb{R}^{4150 \times 4150}$ matrices are dense and large, but can be reduced using the Piecewise Aggregation Approximation (PAA) technique (Keogh and Pazzani, 2000). In this study, we also compared the effect of the magnitude of matrix reduction on the results. That is, we compared matrices with the shape of 50×50, 100×100, 150×150, 200×200, and 250×250. The conversion of spectra to GASF and GADF matrices is illustrated in **Figure 4.** Here, the matrices are represented as grayscale images.

3.3. 2D CNN

2D CNNs are typically used to classify images, which are multi-channel matrices (tensors). 2D CNNs are commonly used, for instance, to classify RGB images (each channel is a color channel). In our approach, we use a two-channel, data representation that embeds the GASF and GADF matrices as channels. Depending on the scale of reduction, the size of the input varies from

$50 \times 50 \times 2$ to $250 \times 250 \times 2$. The 2D convolution operation extracts meaningful patterns from GASF and GADF matrices. In our architecture, each 2D convolution filter is a matrix of shape 3×3. The output of a 2D convolution operation also indicates the presence or absence of the pattern that the filter is searching for. A 2D max-pooling operation to reduce the dimension of the convolved matrices was also used. A 2D max-pooling operation takes a subregion of 2×2 and reduces it to a single value by taking the maximum value.

The two-channel GASF/GADF object is fed into a 2D CNN, which we refer to as PlasticNet (2D). PlasticNet (2D) contains four 2D convolutional layers, two 2D max-pooling layers, and three fully connected layers (**Figure 3**). The 2D convolutional layer has 64 filters of size 3×3 and the 2D max-pooling layer has filters of size 2×2. The settings for the fully-connected layers, activation functions between layers, dropout ratio, final layer activation function, and the loss function are the same as those used in PlasticNet (1D).

3.4. Saliency Analysis

Saliency maps are a powerful tool used for highlighting features in the input data that are considered relevant to the predictions of the CNN model. In our case, these techniques try to highlight aspects in a given input data object that the CNN is searching for. Among all saliency map methods, an integrated gradient (IG) was used that has the most theoretical completeness (Adebayo et al., 2018; Sundararajan et al., 2017). For the PlasticNet (2D) case, let $\mathcal{V} \in \mathbb{R}^{200 \times 200 \times 2}$ be the input and θ be the parameter vector, the CNN can be written as a large and complicated equation $F(\mathcal{V}; \theta): \mathbb{R}^{200 \times 200 \times 2} \mapsto \mathbb{R}^{10}$, where the output is the classification probability. The loss function is then $L\big(F(\mathcal{V}; \theta)\big): \mathbb{R}^{10} \mapsto \mathbb{R}$. The saliency map $S \in \mathbb{R}^{200 \times 200 \times 2}$ calculated by the IG as:

$$S = abs\left((\mathcal{V} - \bar{\mathcal{V}}) \cdot \int_0^1 \frac{\partial L(F(\bar{\mathcal{V}} + \beta(\mathcal{V} - \bar{\mathcal{V}}); \theta))}{\partial \mathcal{V}} d\beta \right)$$

where $\bar{\mathcal{V}} \in \mathbb{R}^{200 \times 200 \times 2}$ is a baseline input that represents the absence of a feature in the input $\mathcal{V}$. Typically, $\bar{\mathcal{V}}$ only contains zero values. Saliency maps on the GASF/GADF will not only illustrate the signals that are significant at a given frequency (cm⁻¹) but will also highlight important relationships between the signals. Saliency maps can help us understand why CNNs can accurately classify plastic types.

4. Results and Discussion

Classification results for PlasticNet (1D) and (2D) are presented in **Figure 5**, along with comparisons of different input sizes. The results reveal that PlasticNet (2D) has a higher accuracy when the input size is larger than 100×100, compared to PlasticNet (1D) on raw IR spectra (77.7%). Specifically, PlasticNet (2D) with an input size of 200×200 increases the accuracy of the PlasticNet (1D) by 12.4%; this confirms that correlation information in spectra is important for classification. The classification accuracy of PlasticNet (2D) improves as the input matrix size increases until reaching a size of 200×200. This suggests that larger input matrices may contain richer information, which is important for classification. The accuracy of the input size of 250×250 has a slightly lower accuracy (86.9%) than the one of 200×200 (87.3%). This indicates that the input matrix with a size of 200×200 contains sufficient information and continuing to increase the matrix size can lead to overfitting. **Table 1** provides a comparison of the overall accuracy obtained with all CNN architectures explored.

12

Table 1. Overall classification accuracies found with different CNN architectures.

1D	2D (50×50)	2D (100×100)	2D (150×150)	2D (200×200)	2D (250×250)
77.77 %	78.14 %	84.29 %	85.57 %	87.29%	86.86%

We obtain further insight into classification accuracies obtained for different plastic types by using confusion matrices. Each row of the confusion matrix represents instances of the predicted class and each column represents instances of the true class. The entries along the diagonal lines are where the instances are correctly classified. The confusion matrix for PlasticNet (2D) with an input size of 200×200 (**Figure 6**) indicates that plastic types are correctly predicted 87.3% of the time. The confusion matrix also indicates that PC has the lowest classification accuracy among the 10 plastics, with 9% of PC classified as Acrylic and 6% as PB, respectively. It was also found that PE, PET, and PI have classification accuracies that were close to 100%. These results indicate that certain plastics can be more easily classified than others (their spectra have more unique features). This information can be useful in identifying strategies to target specific types of plastics (e.g., by tuning IR equipment).

To validate the effectiveness of the proposed CNN models, we compared the average accuracies over 5-fold cross-validation of 1D CNN and 2D CNNs with four commonly used ML classifiers, including Radial Basis Function (RBF) based Support Vector Machine (RBF-SVM) (Vapnik, 1998), Random Forest (RF), k-Nearest Neighbors (kNN) (DICKERSON et al., 1992), Gaussian Process Classifier (GPC) (KI Williams, 2006). SVM is a learning method that was designed to find optimal decision boundaries between classes. The use of the RBF function in SVM allows for

mapping patterns nonlinearly into a high-dimensional feature space (Schölkopf et al., 1997), and it introduces a kernel parameter (γ) in addition to penalty parameter (i.e., C) in linear SVM. In our experiments, These are two parameters (γ and C) were selected from a wide range of values, i.e., $\gamma \in \{10^{-2}, 10^{-1}, \ 10^{0}, \ 10^{1}, \ 10^{2}\}$ and $C \ \in \{10^{-2}, 10^{-1}, \ 10^{0}, \ 10^{1}, \ 10^{2}, 10^{3}\}$, using grid search approach (Staelin, 2003) that performed on the training data. RF is an ensemble ML method that utilizes predictions from many randomized decision trees and it is found to be well suited to high-dimensional data modeling. There are two parameters in RF that need to be tuned to optimized the model performance, they are the number of trees to be grown in the run (*ntree*) and the number of features used in each split (*mtry*). We set *ntree* = 500 and *mtry* is set to the square root of the number of features as recommended by many studies (Immitzer et al., 2012; Sidike et al., 2019). KNN is another popular ML algorithm, which involves the measurement of k-nearest neighbors of a test sample and it is the class label that is decided on a majority vote. The number of neighbors in KNN is fixed to 5 in the experiments. GCP can be modeled based on a GP prior and the latent function. The default parameters used in GCP, as specified in the *Scikit-learn* (Pedregosa et al., 2011) ML library. **Table 2** provides a comparison of the overall accuracy of these ML algorithms. It can be observed that RBF-SVM yields the best accuracy, while kNN and GPC demonstrate similar performance but lower accuracy than RBF-SVM and RF.

Table 2. Overall accuracies obtained with other ML algorithms.

RBF-SVM	RF	k-NN	GPC
86.14 %	72.57 %	65.00 %	63.29 %

A comparison between the CNN-based methods and other ML algorithms is shown in **Figure 5**. The accuracy of PlasticNet (2D) is slightly higher (~1%) than that of RBF-SVM when the input size is larger than 200 × 200. This indicates that RBF-SVM is comparable to CNN-based methods;

however, SVMs provide limited information on features that drive predictions and offer limited flexibility to capture different representations for IR data. The results obtained with SVM confirm that there appears to be enough separation (differences) in the spectra that can be exploited to classify different types of plastic materials. However, the accuracy of all methods saturates at 87%, which suggests that the dataset itself contains significant errors that neither the CNN-based nor the SVM methods can explain.

To understand exactly what the CNNs have learned from the spectra, we used saliency maps to find the most important regions for classification. We used the results for PlasticNet (2D) with an input size of 200×200, since this has the highest accuracy. **Figure 8** shows the average (a) GASF, (b) GADF, (c) saliency map, (d) spectrum and its important regions of PE. The average saliency map for each plastic was studied because each spectrum has some subtle differences, and the common significant patterns were of interest. The darker regions in **Figure 8c** are the most important ones. Specifically, the horizontal bands near 2900 cm^{-1} and vertical bands 2400 cm^{-1} were dark, which indicates the importance of the signal at these frequencies. **Figure 8d**, shows the significant signal locations (shaded regions) and the raw spectrum. The bands between 2800 and 2900 cm^{-1} were of importance. This region provides characteristic IR bands for PE. A similar trend is observed for other plastics, such as PC and ABS shown in Error! Reference source not found. and **Figure 9**. Saliency analysis shows that the regions of interest for PlasticNet (2D) are exactly the most physically informative regions. This confirms intuition that might be exploited by humans to compare different spectra.

5. Conclusions

A convolutional neural network (CNN) framework for classifying different types of plastic materials that are commonly found in MPW based on ATR-FTIR spectra was developed. An important aspect of this type of spectral data is that it can be collected in real-time; as such, this approach provides an avenue for the high-throughput characterization of MPW. The proposed CNN framework (which we call PlasticNet) uses a Gramian angular representation of the IR spectra and we show that this approach reaches overall classification accuracies of 87%. Moreover, it has been found that certain plastics can be classified with 100% accuracy. As part of future work, we aim to test the proposed framework using high-throughput data collected in an online system and to account for other sources of complexity and noise arising in MPW systems (e.g., presence of pigments).
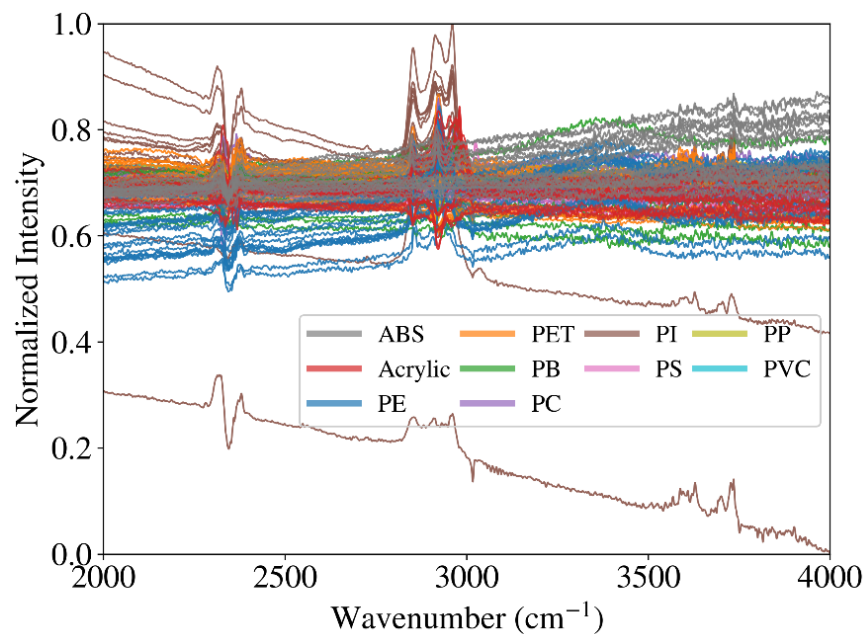
Figure 1: Normalized infrared spectral intensities of various plastic materials. Each spectrum is a vector of length 4150. The resulting spectra contain significant noise and systematic errors.
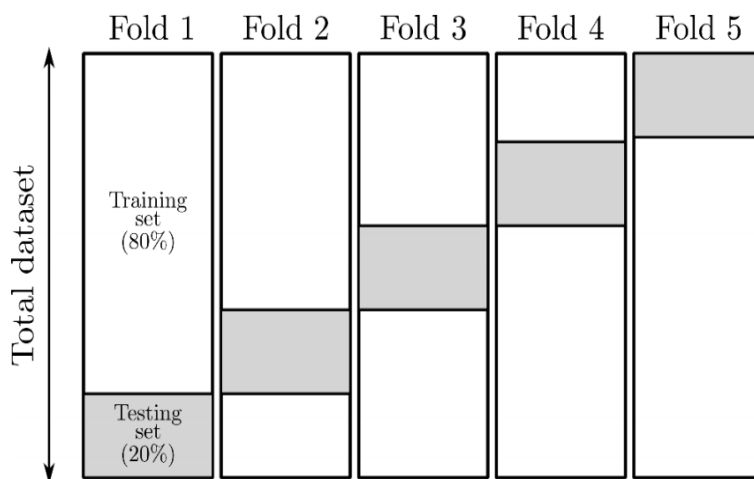

Figure 2: Schematic illustration of 5-fold cross-validation procedure used to train and test models. The training-to-testing split is 4:1. Within the training set, we randomly select 30% of the data as the validation set to tune the parameters of the model.
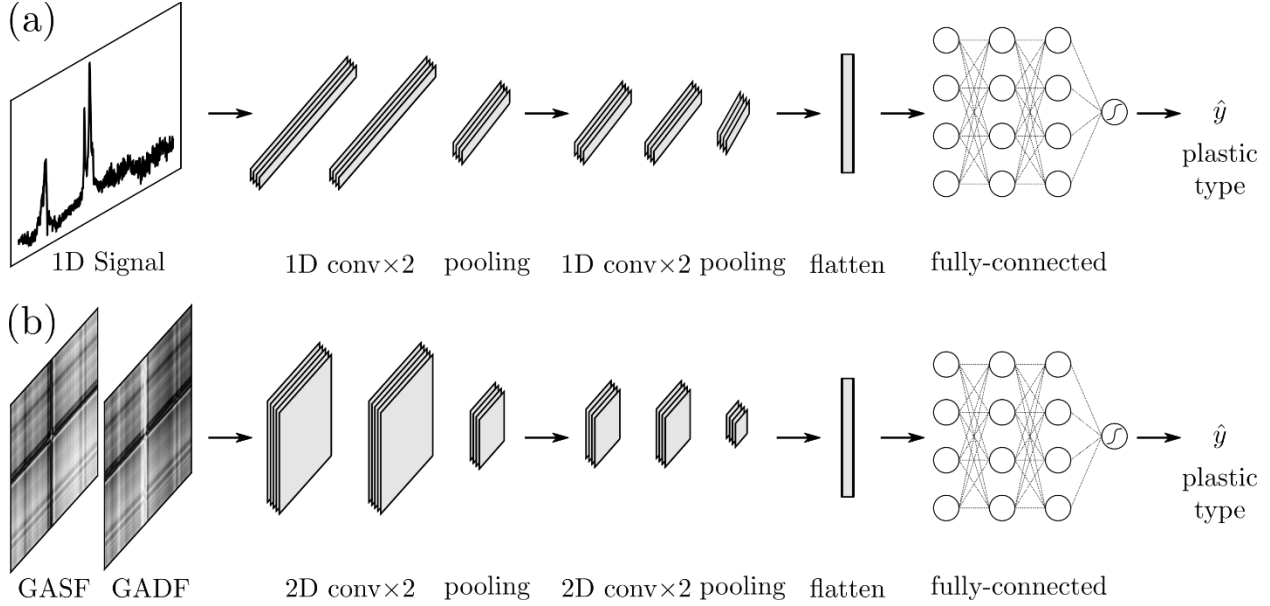
Figure 3: Architectures of (a) PlasticNet (1D) and (b) PlasticNet (2D). PlasticNet (1D) inputs a vector of 4150 and outputs the predicted plastic type. It contains 4 1D convolutional layers (each has 64 filters of dim 3), 2 1D max-pooling layers (each has a pooling window size of 2), a flatten layer, and 3 fully-connected layers (each has 64 units and a dropout ratio of 0.2). The activation functions between the layers are ReLUs. The final output activation function is softmax. PlasticNet (2D) inputs a GASF and a GADF matrix. The input size varies from 50×50×2 to 250×250×2. It has 4 2D convolutional layers (each has 64 filters of 3×3), 2 2D max-pooling layers (each has a pooling window size of 2×2). The flatten, fully-connected layers and activation function setups are the same as the ones of PlasticNet (1D).
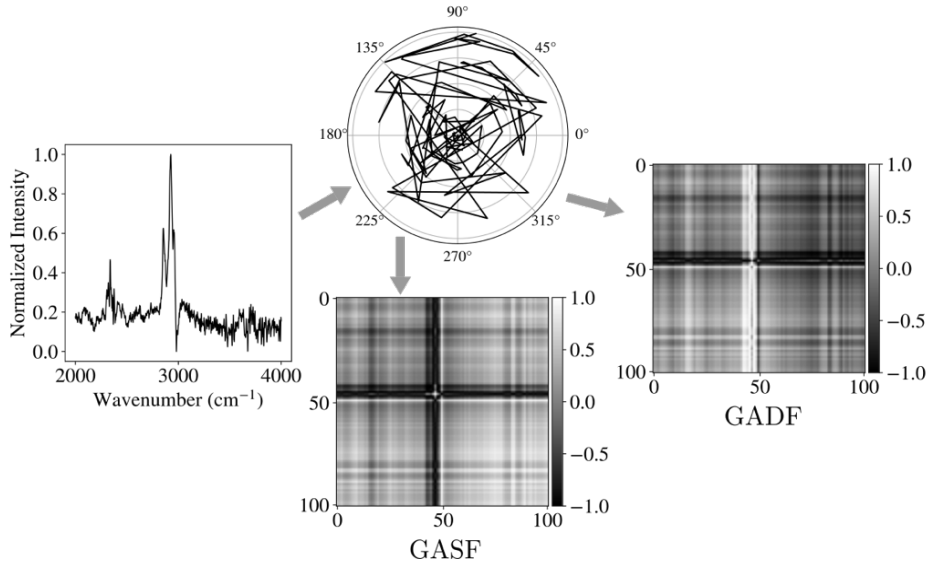

Figure 4: Conversion from 1D signal to GASF and GADF matrices. The 1D signal is first mapped to the polar coordinate system and finally converted to GASF and GADF matrices. Encoding the 1D signal into GAF matrices captures the relationship between the signal intensity at different wavenumbers.
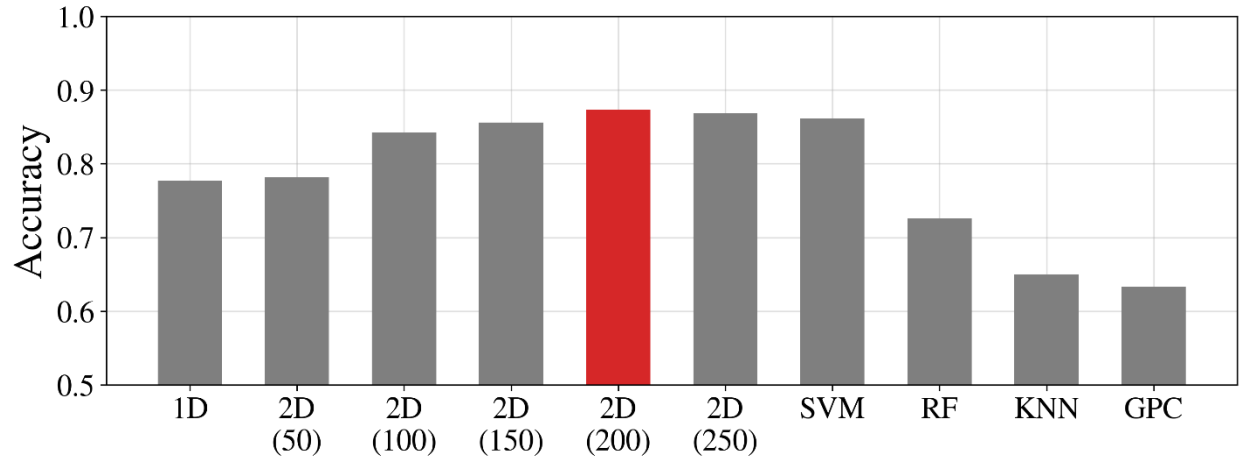
Figure 5: Comparison of the accuracy of CNN-based methods and other ML algorithms. PlasticNet (2D) with an input size of 200×200×2 has the highest accuracy of 87.29%. SVM with RBF kernels has a comparable accuracy of 86.14%. The accuracy of PlasticNet (2D) is always higher than that of PlasticNet (1D), indicating that the conversion from the original 1D signal to 2D GAF matrices captures more information. The accuracy of PlasticNet (2D) increases as the input matrix increases, indicating that a larger input matrix contains more information.
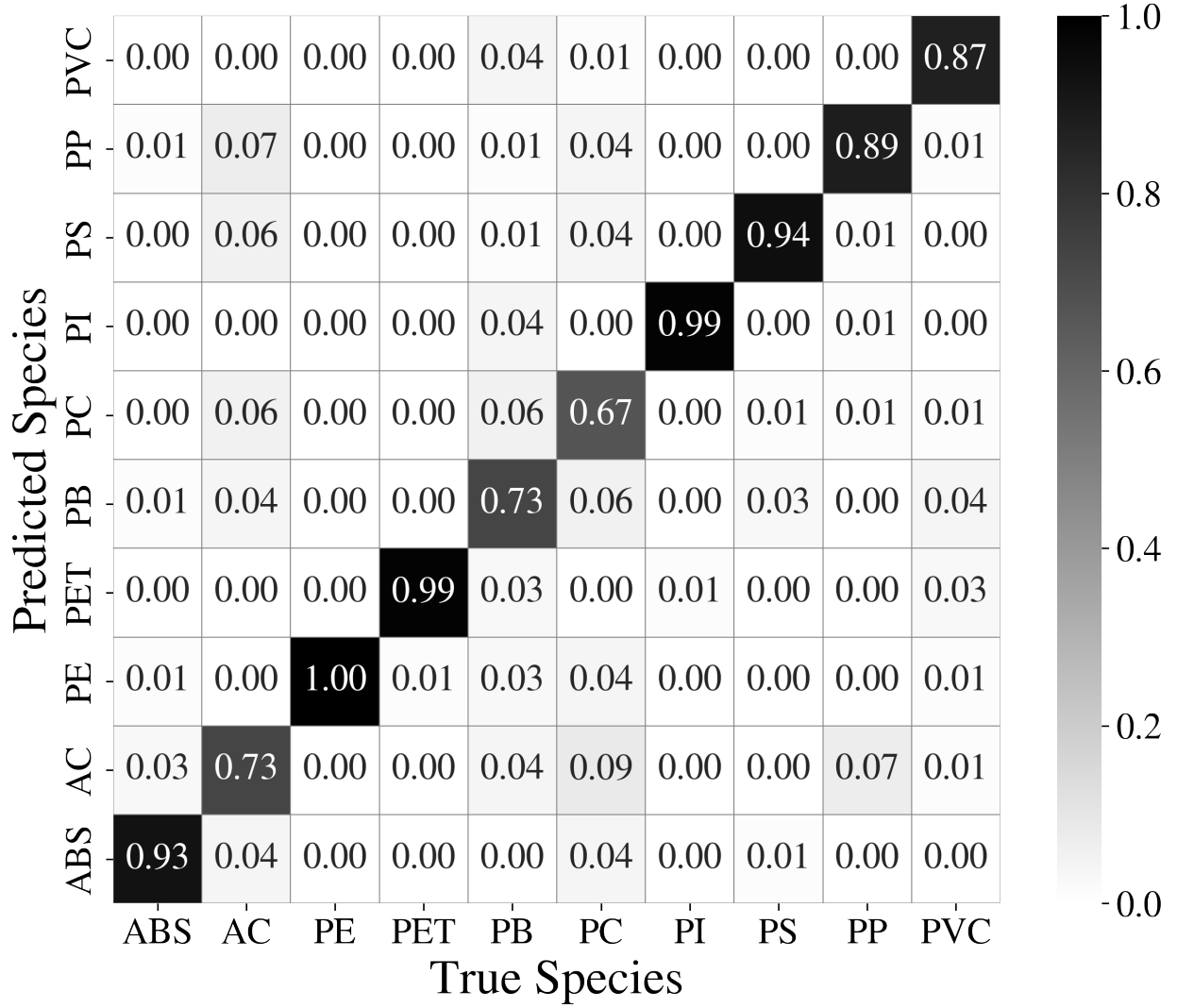
Figure 6: Confusion matrix for PlasticNet (2D) with an input size 200×200×2. The overall accuracy is 87.3%. Each column represents a true plastic species, and each row represents a model predicted plastic species. The entries along the diagonal are where the plastic species are correctly classified. Many diagonal entries are close to one, indicating that the PlasticNet (2D) has excellent classification accuracy. However, some plastic types cannot be classified with high accuracy (e.g., PC and AC).
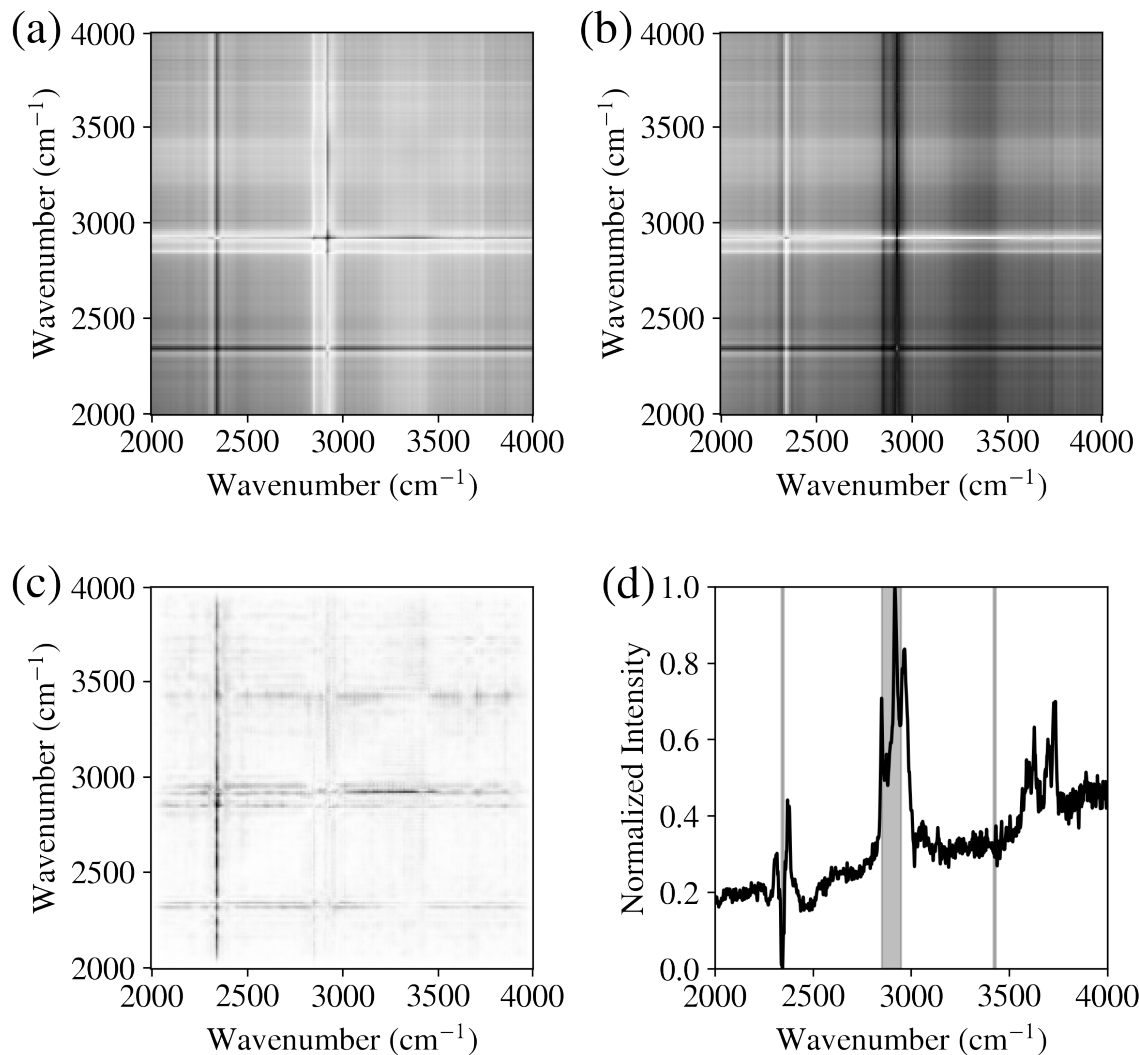
Figure 7: Saliency analysis for PE. The average (a) GASF and (b) GADF matrices of size 200 ×
200, where darker colors represent larger values. (c) The average saliency map of size 200 × 200.
The darker regions are the most important regions for classification. (d) The average IR spectrum
and the most important signals, shaded in gray. The most important region includes the bands
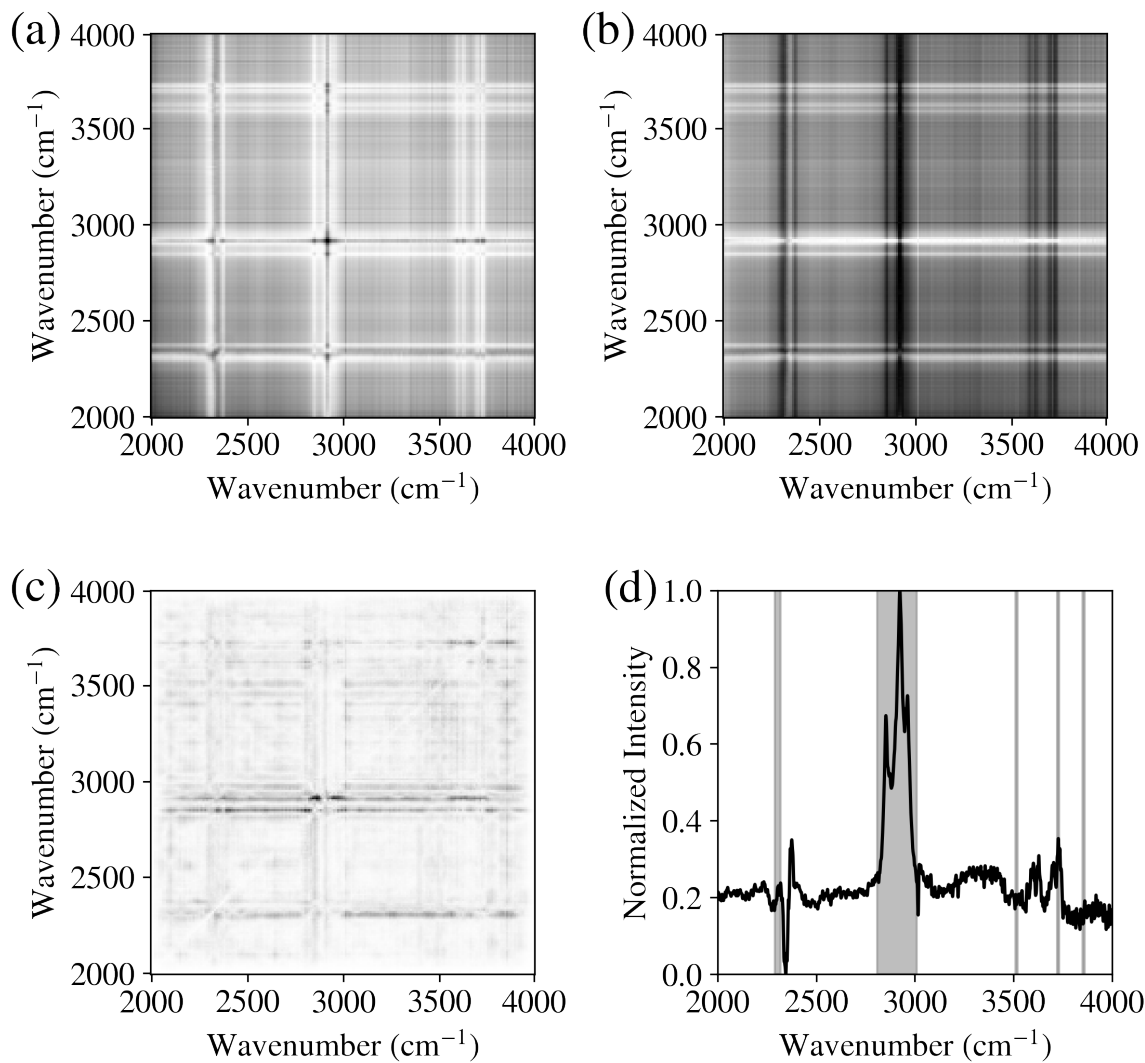between 2800-2900 cm⁻¹, which are the characteristic IR peaks of the PC.

Figure 8: Saliency analysis for PC. The average (a) GASF and (b) GADF matrices of size 200 × 200. (c) The average saliency map of size 200 × 200. (d) The average IR spectrum and the most important signals, shaded in gray.
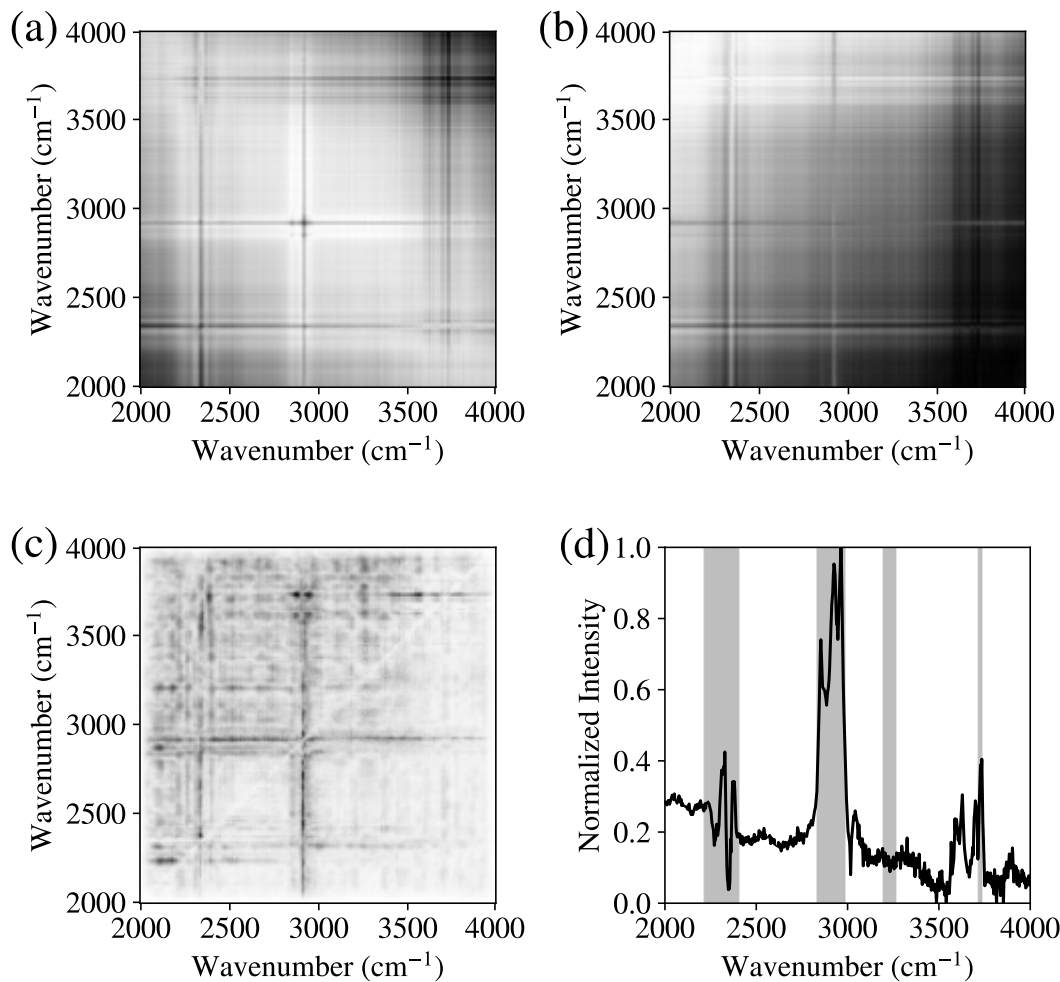
Figure 9: Saliency analysis for ABS. The average (a) GASF and (b) GADF matrices of size 200 × 200. (c) The average saliency map of size 200 × 200. (d) The average IR spectrum and the most important signals, shaded in gray.

References

Abdel-Shafy, H.I., Mansour, M.S.M., 2018. Solid waste issue: Sources, composition, disposal, recycling, and valorization. Egypt. J. Pet. https://doi.org/10.1016/j.ejpe.2018.07.003

Adebayo, J., Gilmer, J., Muelly, M., Goodfellow, I., Hardt, M., Kim, B., 2018. Sanity Checks for Saliency Maps. Adv. Neural Inf. Process. Syst. 2018-December, 9505–9515.

Allen, V., Kalivas, J.H., Rodriguez, R.G., 1999. Post-Consumer Plastic Identification Using Raman Spectroscopy. Appl. Spectrosc. 53, 672–681. https://doi.org/10.1366/0003702991947324

Awoyera, P.O., Adesina, A., 2020. Plastic wastes to construction products: Status, limitations and future perspective. Case Stud. Constr. Mater. 12, e00330. https://doi.org/10.1016/j.cscm.2020.e00330

Chen, X., Chai, Q., Lin, N., Li, X., Wang, W., 2019. 1D convolutional neural network for the discrimination of aristolochic acids and their analogues based on near-infrared spectroscopy. Anal. Methods 11, 5118–5125. https://doi.org/10.1039/c9ay01531k

da Silva, D.J., Wiebeck, H., 2020. Current options for characterizing, sorting, and recycling polymeric waste. Prog. Rubber, Plast. Recycl. Technol. 36, 284–303. https://doi.org/10.1177/1477760620918603

Da Silva, V.H., Murphy, F., Amigo, J.M., Stedmon, C., Strand, J., 2020. Classification and Quantification of Microplastics (<100 μm) Using a Focal Plane Array-Fourier Transform Infrared Imaging System and Machine Learning. Anal. Chem. 92, 13724–13733. https://doi.org/10.1021/acs.analchem.0c01324

DICKERSON, M.T., DRYSDALE, R.L.S., SACK, J.-R., 1992. SIMPLE ALGORITHMS FOR ENUMERATING INTERPOINT DISTANCES AND FINDING k NEAREST NEIGHBORS. Int. J. Comput. Geom. Appl. 02, 221–239. https://doi.org/10.1142/s0218195992000147

Gundupalli, S.P., Hait, S., Thakur, A., 2017. A review on automated sorting of source-separated municipal solid waste for recycling. Waste Manag. https://doi.org/10.1016/j.wasman.2016.09.015

Hopewell, J., Dvorak, R., Kosior, E., 2009. Plastics recycling: challenges and opportunities. Philos. Trans. R. Soc. B Biol. Sci. 364, 2115–2126.

Immitzer, M., Atzberger, C., Koukal, T., 2012. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. Remote Sens. 4, 2661–2693. https://doi.org/10.3390/rs4092661

Jiang, S., Zavala, V.M., 2021. Convolutional Neural Nets: Foundations, Computations, and New Applications.

Karlsson, T.M., Grahn, H., van Bavel, B., Geladi, P., 2016. Hyperspectral Imaging and Data Analysis for Detecting and Determining Plastic Contamination in Seawater Filtrates. J. Near Infrared Spectrosc. 24, 141–149. https://doi.org/10.1255/jnirs.1212

Keogh, E.J., Pazzani, M.J., 2000. Scaling up dynamic time warping for datamining applications, in: Proceeding of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery (ACM), New York, New York, USA, pp. 285–289. https://doi.org/10.1145/347090.347153

KI Williams, C., 2006. Gaussian processes formachine learning. Taylor & Francis Group.

Kiranyaz, S., Ince, T., Gabbouj, M., 2016. Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. IEEE Trans. Biomed. Eng. 63, 664–675. https://doi.org/10.1109/TBME.2015.2468589

Michel, A.P.M., Morrison, A.E., Preston, V.L., Marx, C.T., Colson, B.C., White, H.K., 2020. Rapid Identification of Marine Plastic Debris via Spectroscopic Techniques and Machine Learning Classifiers. Environ. Sci. Technol. 54, 10630–10637. https://doi.org/10.1021/acs.est.0c02099

Milios, L., Holm Christensen, L., McKinnon, D., Christensen, C., Rasch, M.K., Hallstrøm Eriksen, M., 2018. Plastic recycling in the Nordics: A value chain market analysis. Waste Manag. 76, 180–189. https://doi.org/10.1016/j.wasman.2018.03.034

Ng, A.Y., Ng, A.Y., 1997. Preventing "Overfitting" of Cross-Validation Data. Proc. FOURTEENTH Int. Conf. Mach. Learn. 245--253.

Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., McBratney, A.B., 2019. Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. Geoderma 352, 251–267. https://doi.org/10.1016/j.geoderma.2019.06.016

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Ritchie, H., Roser, M., 2018. Plastic Pollution [WWW Document]. Our World Data. URL https://ourworldindata.org/plastic-pollution (accessed 4.26.21).

Roh, S.B., Park, S.B., Oh, S.K., Park, E.K., Choi, W.Z., 2018. Development of intelligent sorting system realized with the aid of laser-induced breakdown spectroscopy and hybrid preprocessing algorithm-based radial basis function neural networks for recycling black plastic wastes. J. Mater. Cycles Waste Manag. 20, 1934–1949. https://doi.org/10.1007/s10163-018-0701-1

Schlesinger, M.E., 2013. Aluminum Recycling, Second Edition. Taylor \& Francis.

Schölkopf, B., Sung, K.K., Burges, C.J.C., Girosi, F., Niyogi, P., Poggio, T., Vapnik, V., 1997. Comparing support vector machines with gaussian kernels to radial basis function classifiers. IEEE Trans. Signal Process. 45, 2758–2765. https://doi.org/10.1109/78.650102

Siddiqui, M.N., Gondal, M.A., Redhwi, H.H., 2008. Identification of different type of polymers in plastics waste. J. Environ. Sci. Heal. - Part A Toxic/Hazardous Subst. Environ. Eng. 43, 1303–1310. https://doi.org/10.1080/10934520802177946

Sidike, P., Sagan, V., Maimaitijiang, M., Maimaitiyiming, M., Shakoor, N., Burken, J., Mockler, T., Fritschi, F.B., 2019. dPEN: deep Progressively Expanded Network for mapping heterogeneous agricultural landscape using WorldView-3 satellite imagery. Remote Sens. Environ. 221, 756–772. https://doi.org/10.1016/j.rse.2018.11.031

Signoret, C., Caro-Bretelle, A.S., Lopez-Cuesta, J.M., Ienny, P., Perrin, D., 2020. Alterations of plastics spectra in MIR and the potential impacts on identification towards recycling. Resour. Conserv. Recycl. 161, 104980. https://doi.org/10.1016/j.resconrec.2020.104980

Signoret, C., Caro-Bretelle, A.S., Lopez-Cuesta, J.M., Ienny, P., Perrin, D., 2019. MIR spectral characterization of plastic to enable discrimination in an industrial recycling context: I. Specific case of styrenic polymers. Waste Manag. 95, 513–525. https://doi.org/10.1016/j.wasman.2019.05.050

Staelin, C., 2003. Parameter selection for support vector machines. Hewlett-Packard Company, Tech. Rep. HPL-2002-354R1 1.

Sundararajan, M., Taly, A., Yan, Q., 2017. Axiomatic attribution for deep networks. arXiv.

Vapnik, V., 1998. Statistical Learning Theory. Wiley.

Wang, Z., Oates, T., 2015. Encoding time series as images for visual inspection and classification using tiled convolutional neural networks, in: AAAI Workshop - Technical Report. AI Access Foundation, pp. 40–46.

Wu, Q., Yu, F., Okabe, Y., Saito, K., Kobayashi, S., 2015. Acoustic emission detection and position identification of transverse cracks in carbon fiber–reinforced plastic laminates by using a novel optical fiber ultrasonic sensing system. Struct. Heal. Monit. An Int. J. 14, 205–213. https://doi.org/10.1177/1475921714560074

Zhu, S., Chen, H., Wang, M., Guo, X., Lei, Y., Jin, G., 2019. Plastic solid waste identification system based on near infrared spectroscopy in combination with support vector machine. Adv. Ind. Eng. Polym. Res. 2, 77–81. https://doi.org/10.1016/j.aiepr.2019.04.001