Kent Academic Repository

Ejiofor, Uche, Wang, Lijuan and Wang, Dayang (2025) ***Plastic identification using NIR spectral analysis and KNN model with correlation feature selection.*** **In: 2025 IEEE International Instrumentation and Measurement Technology Conference. . IEEE ISBN 979-8-3315-0500-4.**

Downloaded from

https://kar.kent.ac.uk/108950/ The University of Kent's Academic Repository KAR

The version of record is available from

https://doi.org/10.1109/I2MTC62753.2025.11079091

This document version

Author's Accepted Manuscript

DOI for this version

https://doi.org/10.22024/UniKent/01.02.108950.3453401

Licence for this version

CC BY (Attribution)

Additional information

For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in ***Title of Journal*** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

Plastic Identification Using NIR Spectral Analysis and KNN Model with Correlation Feature Selection

Uche Ejiofor
School of Engineering
University of Kent
Canterbury, United Kingdom
uue2@kent.ac.uk

Lijuan Wang
School of Engineering
University of Kent
Canterbury, United Kingdom
l.wang@kent.ac.uk

Dayang Wang
School of Engineering
University of Kent
Canterbury, United Kingdom
dw523@kent.ac.uk

Abstract*—The increasing demand for high-purity post-consumer plastics for recycling highlights the importance of efficient material identification. Near-Infrared (NIR) spectroscopy is widely used to capture unique spectral fingerprints of materials for identification purposes. However, developing efficient and lightweight machine learning models with NIR spectral data for accurate and real-time classification of plastic materials is still in high demand. This paper presents a new hybrid approach combining the k-nearest neighbours (KNN) model with the filter-based correlation feature selection (CFS) technique for plastic identification. Experiments were conducted to test six key recyclable plastic polymers consisting of Polyethylene Terephthalate (PET), High-Density Polyethylene (HDPE), Low-Density Polyethylene (LDPE), Polyvinyl Chloride (PVC), Polypropylene (PP), and Polystyrene (PS). Feature selection techniques - CFS and principal component analysis (PCA) are applied to extract the essential discriminative features of each type of plastic. With the extracted features, KNN and support vector machine (SVM) models are developed respectively for plastic identification. Experimental results demonstrate that the CFS-KNN model achieves a success rate of 100% with a computational time of 40ms under laboratory conditions, outperforming the PCA-SVM, PCA-KNN and CFS-SVM models in terms of accuracy and computational efficiency.

Keywords— Plastic Polymers Recycling, Near-Infrared (NIR), Correlation Feature Selection (CFS), Principal Component Analysis (PCA), K-nearest Neighbors (KNN), Support Vector Machine (SVM)

I. INTRODUCTION

Solid waste management is a growing global issue, with over 2 billion tons of municipal solid waste generated annually, projected to reach 3.4 billion tons by 2050 [1] [2]. Only 13.5% of waste is recycled [3], highlighting the inefficiency of the current recycling systems. Plastic solid waste poses a significant environmental concern as it is non-biodegradable, leading to severe harm to both land and water ecosystems. Sorting methods in material recovery facilities, whether manual or automated, are costly, slow, and unreliable in identifying valuable recyclables like PET, HDPE, and PP. Using inferior raw materials in recycling poses risks and inefficiencies [4] [5]. Developing an efficient identification system tailored for recycling facilities is essential for delivering high-quality secondary plastics, enabling effective recycling, and addressing the global plastic waste crisis.

A range of sensing technologies, including Raman [6] and NIR spectroscopy [7], hyperspectral imaging [8], computer vision [9], and ultrasonic sensors [10] have been applied for material identification. However, Raman spectroscopy struggles with poor efficiency and alloy detection. Computer vision relies on image features, limiting its effectiveness for plastic polymer identification. Ultrasonic sensors face challenges in detecting plastic polymers due to negligible differences in the acoustic impedance of polymers. Hyperspectral imaging and NIR spectroscopy offer spectral information for analysing molecular structures, making them a versatile, efficient tool for identifying diverse materials, including plastic polymers [11].

Different machine learning models have been employed along with NIR technology to enhance the accuracy of material identification. In line with that initiative, Tehrani and Karbasia [12] proposed a three-layer Artificial Neural Networks (ANN) model with four hidden neurons to analyse the hyperspectral data captured from a NIR hyperspectral camera to discriminate only three electronic waste plastics – Polycarbonate (PC), Acrylonitrile Butadiene Styrene (ABS), and High Impact Polystyrene (HIPS) with almost 99% accuracy. Maliks and Kadik [13] developed a plastic classification system using NIR spectroscopy (900-1700 nm) and a simplified Convolutional Neural Networks (CNN) model with PCA achieving 98.4% accuracy in identifying only four plastic polymers: PET, HDPE, LDPE, and PP. Zhu et al. [14] utilized NIR reflectance spectroscopy combined with an SVM model and PCA to classify six municipal and industrial plastics polymers: PP, PS, polyethylene (PE), polymethyl methacrylate (PMMA), ABS, and PET achieving an accuracy of 97.5%. Moirogiorgou et al. [15] evaluated SVM and KNN models using PCA-extracted NIR features to separate plastic materials including PET, HDPE, PP, PE-film from non-plastic materials such as aluminium, and paper. The study found KNN to be faster and achieved an accuracy of 91.34% compared to SVM with 63.97%. However, the research was limited to binary classification, segregating plastics from other materials. Kumar et al. [16] combined PCA with the Spectral Angle Mapper (SAM) algorithm to process NIR data, achieving 97.5% in accuracy for identifying five common plastic polymers: PET, HDPE, PS, PP, and PVC. Cesetti et al. [17] utilized a hyperspectral camera in the NIR range (1100–2500 nm) with the Indices-Definition model and Correlation Matrix feature selection to classify waste materials, including plastics, wood, stone, laminate, and glass. While introducing filter-based feature selection, the study [17] did not specify the types of plastics tested nor report the model's accuracy. Overall, in the existing research, deep learning models such as CNN achieves high accuracy but are computationally expensive, while classical models such as SVM and KNN combining PCA extracted features achieves comparable accuracy only for the classification of limited types of municipal plastic materials. Though PCA is commonly applied for feature selection, the comparison with other techniques, like filter-based methods, for the application of plastic identification remains underexplored.

This paper proposes a novel hybrid approach combining the KNN model with the filter-based CFS technique for the multi-class identification of plastic materials. Experimental tests were conducted to identify six key recyclable plastic polymers consisting of PET, HDPE, LDPE, PVC, PP and PS. Feature selection techniques - CFS and PCA are applied to extract the relevant features of each type of plastic. Accordingly, KNN and SVM models are developed to utilise the extracted features for plastic identification. Performance comparisons among hybrid models PCA-SVM, PCA-KNN, CFS-SVM and CFS-KNN were undertaken to evaluate the accuracy and computational time of these models.

II. METHODOLOGY

A. *Plastic Identification Strategy*

The proposed plastic identification strategy shown in Fig.1 consists of a sensing unit, a data processing unit, and an identification unit. The sensing unit incorporates the NIR reflectance measurement setup, which includes key components such as a light source, a spectrometer, and an optical reflectance probe. The spectra of materials captured in the sensing unit are pre-processed in the processing unit. The chemometric techniques including standard normal variate (SNV), polynomial fitting and Savitzky–Golay are used for pre-processing data to remove additive noise and align data classes to similar spectral responses. The CFS feature selection technique is adopted to extract useful features to facilitate classification. The machine learning models, such as KNN, implemented in the identification unit classifies materials into different polymer types.

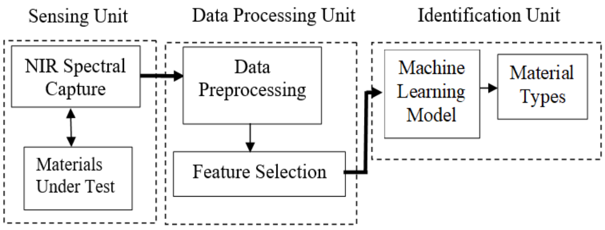


Fig. 1. Principle of plastic identification strategy

B. *Near-Infrared Sensing Principle*

This study employs the NIR spectroscopic sensing operated within the wavelength range of 900nm to 1700nm. The interaction of NIR light with polymer samples induces molecular vibrations, primarily comprising overtones and combination bands of hydrogen bonds (C-H, O-H, and N-H), which are used to identify different plastic polymers [18]. Among the NIR spectroscopy sampling techniques - transmission, attenuated total reflection, specular reflection, and diffuse reflectance, diffuse reflectance is most suitable for non-contact acquisition of the spectral signatures of plastic materials [19]. The operating principle of the NIR diffuse reflectance spectroscopy is based on the Beer-Lambert Law [18] which stipulates that the intensity of reflected NIR radiation depends on the amount of light absorbed by a material which is directly related to the light path length and the molecular concentration of the sample. The Beer-Lambert Law is described as:

$$A = \log_{10}\left(\frac{I_o}{I_1}\right) = \varepsilon l c \quad (1)$$

where, A is absorbance, I_0 is intensity of incident light, I_1 is intensity of transmitted light, l is the thickness of sample or light path length, c is concentration of chemical compound in the sample, while ε is the molar absorptivity of material.

C. *Feature Selection*

Feature selection algorithms are categorized into filter, wrapper, and embedded approaches [20]. The filter-based approach does not overfit and is not computationally expensive compared to the wrapper and embedded approaches. This research explores the filter-based feature selection approach - Correlation Feature Selection (CFS) and compares it with PCA.

1) Correlation Feature Selection (CFS): CFS is considered a scaled version of covariance due to the implementation of normalization to select features. CFS computes correlation coefficients to determine relationships among features and their relevance to the target variable. This helps maximize relevance while minimizing redundancy between features. The correlation coefficients are calculated using Pearson's correlation coefficient. The feature-class correlation r_{cf} and feature-feature correlation r_{ff} computed using Pearson's correlation coefficient are described as:

$$r_{cf} = \frac{\sum_{i=1}^{n}(f_i-\bar{f})\,(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(f_i-\bar{f})^2\,\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (2)$$

$$r_{ff} = \frac{\sum_{i,j=1}^{n}(f_i-\bar{f}_i)\,(f_j-\bar{f}_j)}{\sqrt{\sum_{i=1}^{n}(f_i-\bar{f}_i)^2\,\sum_{j=1}^{n}(f_j-\bar{f}_j)}^2} \quad (3)$$

where f_i and f_j are feature values, y is target values, while $\bar{f}$ and $\bar{y}$ are the mean values of the feature and target values respectively. The CFS uses heuristic search to determine the optimal relevance-to-redundancy ratio before terminating the feature selection process. The heuristic evaluation function calculates the merit of a subset of features that are highly correlated with the target and minimally correlated with other features. The CFS heuristic merit function is represented as:

$$Merit_s = \frac{k\overline{r_{cf}}}{\sqrt{k+k(k-1)\overline{r_{ff}}}} \quad (4)$$

where, k represents the chosen numbers of features subset, $\overline{r_{cf}}$ is the average feature-class correlation, while $\overline{r_{ff}}$ denotes the average feature-feature correlation.

2) Principal Component Analysis (PCA): PCA is a dimension reduction technique used to decompose a dataset to new orthogonal components that explain the maximum variance in the dataset. PCA uses unsupervised linear technique to transform data into a new set of variables called principal components that capture maximum variance in the original data. The principal components retain the features that have the most impact on the target variable. The central idea behind PCA is to determine the principal components that retain maximum variation in data, thereby reducing the dimensionality of a dataset. The principal components of PCA are derived by solving the eigenvalue equation for the covariance matrix:

$$\sum v = \lambda v \quad (5)$$

where λ represents the eigenvalue describing explained variance by principal components, and v is the eigenvector that determines the direction of principal components.

D. *Machine Learning Models*

1) k-Nearest Neighbours (KNN): KNN is one of the simplest and most popular non-parametric machine-learning algorithms used for classification tasks involving small datasets. KNN uses the proximity of new data among data groups to make classification. KNN classifies data samples according to the majority of class in the dataset which is closest to the data point using the K nearest neighbours algorithm [21]. The neighbours of an unknown sample are the samples having the lowest Euclidean norms. The Euclidean norm is estimated by calculating the Euclidean distance using equation below:

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \tag{6}$$

where the Euclidean distance d is calculated between two points (p, q) in Euclidean n-space. q_i and p_i are the Euclidean vectors between two points.

2)Support Vector Machine (SVM): SVM algorithm aims to draw a line/ hyperplane between different data classes. The central idea behind SVM is to find the ideal hyperplane separating the support vectors with maximum margin. The support vectors which is a key term in SVM are data points in each group of data closest to the line. The concept of kernels is introduced when SVM constructs hyperplanes for non-linearly separable complex data [22]. Once the ideal line is drawn, it can then be used to classify new data. The basic equation for the hyperplane is described as:

$$\overrightarrow{W_0}x_1 + \overrightarrow{W_1}x_2 + b = 0 \tag{7}$$

where x_1 and x_2 are the data inputs, $\overrightarrow{W_0}$ and $\overrightarrow{W_1}$ are the weight support vectors, and b is the bias. The margin that separates data sets is defined as:

$$M = \frac{2}{\|W\|} \tag{8}$$

where $\|W\|$ is the normalized weight vectors ($\overrightarrow{W_0}$ and $\overrightarrow{W_1}$). SVM constrained optimization problem SVM (O) is then solved to get maximum margin M using the equation described as:

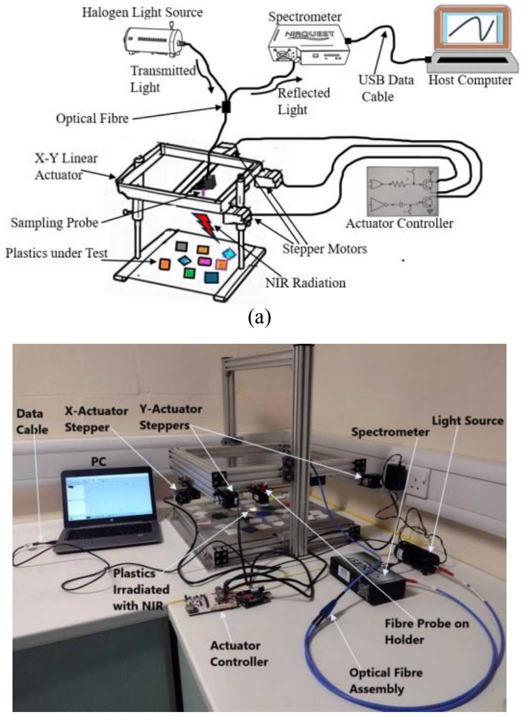
$$SVM\ (O) = min_{w,b}\frac{1}{2}\|W\|^2 \tag{9}$$

where $\|W\|$ is being minimised with respect to w and b to ensure maximum margin for a robust classier.

III. EXPERIMENTATION

A. *Experimental Setup*

The proposed material identification system illustrated in Fig. 2 (a) and (b) uses an Ocean Optics NIRQuest512-1.7 spectrometer to capture spectral data of materials. The fibre optic modular spectrometer is based on high sensitive Indium Gallium Arsenide (InGaAs) array detector with 512 pixels. The NIRQ512 spectrometer has an integration time of 1ms to 120s for high-performance spectroscopic measurements in the 900 - 1700nm wavelength region with a spectral resolution of approximately 2 nm. The NIR spectrometer and accessories including a halogen light source and optical reflection probe are setup in the diffuse reflectance mode to capture spectral data. An X-Y Linear actuator based test rig is developed to move the NIR probe to different positions to effectively sample various materials on the test area of 44cm by 44cm. The NIR probe can be moved in step size of 0.8cm/step on both on the X-axis and Y-axis by the programmed X-Y linear actuator.



(a)

(b)

Fig. 2. (a) Schematic of the experimental setup; (b) Test rig.

B. *Materials*

The representative samples used for the study consists of different household plastic samples collected randomly from various household waste bins. In total 244 samples of PET, HDPE, PVC, LDPE, PP and PS were collected. The samples distribution is tabulated in Table I. The samples consist of different clear/ transparent, white and coloured materials. However, samples of dark colours that contains carbon black were not used due to non-detection in the NIR region.

TABLE I. MATERIAL SAMPLES

Sample Set of 6 Common Municipal Plastic Polymers						
Type	**PET**	**HDPE**	**PVC**	**LDPE**	**PP**	**PS**
No. of Samples	47	48	40	30	49	30

The plastic samples were prepared by removing contamination labels, cutting and flattening the materials to varying shapes with dimensions ranging from 10 cm^2 to 40 cm^2 to ensure more samples can fit in the detection platform of the experimental setup. During experimental tests, a stable environmental conditions are maintained. Ambient fluorescent light was used, and temperature varied between 18 to 25°C for all experiments. The test surface where samples are placed for spectral data capture was polished aluminium. The probe-to-sample distance is very critical variable for the experiment. So an approximate of 5mm probe-to-sample distance was maintained throughout the experiments to ensure less noise in captured spectra and optimise spectral data acquisition time.

C. *Spectral Data Acquisition*

Before spectral data acquisition, the NIR spectrometer was calibrated accordingly. During measurements, the NIR spectrometer was used to scan each material twice to account for heterogeneity and the average of the two spectra recorded as the reflectance spectrum for that sample. Fig. 3 shows the captured NIR spectra of random polymer samples representing each of the six polymers types being classified. The different spectra has a unique spectral signature used to identify each plastic material.

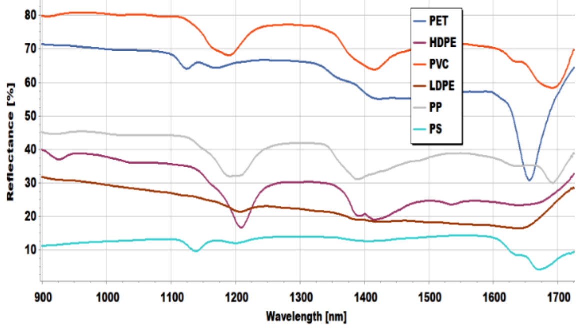


Fig. 3. Raw Spectra of six polymer samples.

After spectral data acquisition, a dataset with a dimension of "244 x 512" was developed for model training. Most of the features (512 wavelength) are redundant and does not contribute to the discrimination of samples. So feature selection was adopted to reduce data dimension and facilitate material identification. For the development and evaluation of machine learning models, the collected dataset (244) was divided into Dataset-1 and Dataset-2. Dataset-1 consists of 192 samples that were used to train and evaluate models. The Dataset-2 contains 52 samples which were kept aside and used to validate the models.

IV. RESULTS AND DISCUSSIONS

A. Data Pre-processing

NIR spectral data is normally influenced with undesired scatter effects such as baseline shift and non-linearities. However, by applying suitable pre-processing on data, these effects can be eliminated prior to multivariate data analysis. The captured raw spectral data of different PET samples is shown in Fig. 4. Different colours represent different PET samples with different reflectance intensities due to differences in the physical attributes of materials.

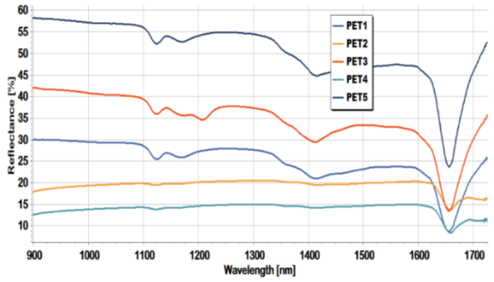


Fig. 4. Un-Processed spectra of selected PET samples.

For this work, SNV normalization was used to correct any intensity variations. Second order derivative (polynomial fitting) was employed for baseline correction to eliminate the effects of illumination intensity fluctuations and viewing geometry, and Savitzky–Golay smoothing was used for noise suppression. The final pre-processed spectra of selected samples of PET shown in Fig. 5 have similar reflectance intensity unlike the raw PET spectra in Fig. 4. Experiments have demonstrated that Savitzky–Golay smoothing with Standard Normal variate (SNV) preprocessing are superior to other pre-processing functions [23]. The SNV transformation is applied to each spectrum by centering the spectrum X around the mean $\bar{X}$ and scaling using standard deviation σ_X. The SNV transformation is calculated as:

$$X_{SNV} = \frac{X-\bar{X}}{\sigma_X} \tag{10}$$

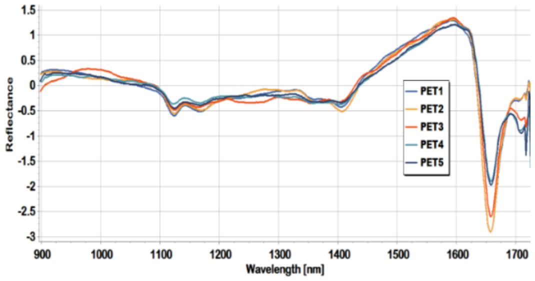
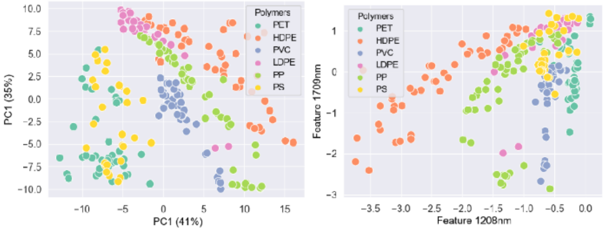


Fig. 5. Pre-Processed spectra of selected PET samples.

B. Feature Selection

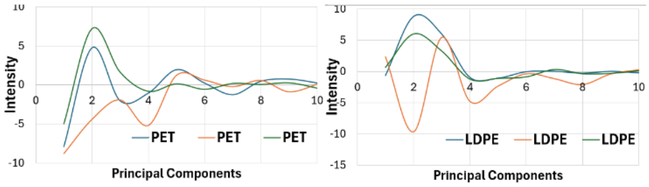
PCA and CFS feature selection techniques were implemented, and the effectiveness of each technique was analyzed. Fig. 6 (a) visualizes the classification capability of PCA using two principal components that capture 76% of variance in the data. Similarly, Fig. 6 (b) shows that CFS can reasonably classify data using only two key features.



(a) (b)

Fig. 6. Two-Dimensional plot of extracted features by (a) PCA and (b) CFS

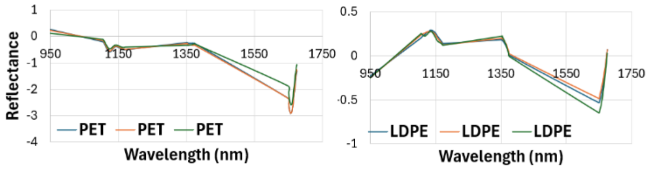
To further assess the discrimination ability of the extracted features of PCA and CFS, the features are plotted graphically to visualise any discrepancy among data classes. Fig. 7 shows the PCA graphical plot of PET and LDPE using 10 principal components that capture 99% of data variance. However, the PCA-transformed data for PET and LDPE reveals sinusoidal waveforms with phase shift. The phase shift in PCA features may affect classification accuracy for similar polymers.



(a) (b)

Fig. 7. Extracted feature spectra of (a) PET and (b) LDPE using PCA

For CFS, Fig. 8 shows that the selected features are non-sinusoidal waves with minimal phase shifts and greater uniformity among data classes compared to PCA. However, unlike PCA, CFS clearly shows in Fig. 8 that the selected features of CFS have potential for efficient classification of plastic materials. The spectral pattern of LDPE is quite distinct from that of PET for discrimination purposes.



(a) (b)

Fig. 8. Selected feature spectra of (a) PET and (b) LDPE using CFS

It can also be observed in Fig. 8 that CFS features have dominant absorbance peaks for plastic discrimination within NIR wavelengths of 1150-1250 nm, 1350-1450 nm, and 1650-1750 nm, which is associated with the overtones and combination bands of C-H, O-H, and N-H bonds [24].

C. *Evaluation of hybrid models*

1) PCA-SVM Model: For the PCA-SVM model, the PCA algorithm utilised 10 principal components to realise a reduced dataset with 99% variance. Grid Search was used to tune the critical hyperparameters (gamma and C) in the Support Vector Machine (SVM) model to realise the best model. The SVM model utilised "rbf" kernel, gamma ="0.013", C=10 to achieve best accuracy of 63.46%. PCA-SVM model performed below expectation when tested for generalization capability to predict materials that are not yet seen by the model. The average computational time achieved was 80ms. Classifcation report and confusion matrix of the models are presented below for analysis of results.

TABLE II. CLASSIFCATION REPORT OF PCA-SVM MODEL

Samples	**Total samples**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
PET	11	67	55	60
HDPE	10	70	70	70
PVC	8	89	100	94
LDPE	7	0	0	0
PP	10	60	90	72
PS	6	38	50	43
Average Accuracy				64
PCA-SVM Average computational time				80ms

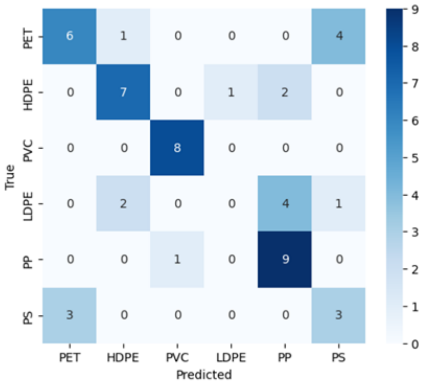


Fig. 9. Confusion Matrix for PCA -SVM model

From the classification report in Table II, majority of the error is associated with the identification of LDPE. No single LDPE was identified as shown in the Confusion Matrix of Fig. *9*. Rather, 2 HDPE, 4 PP and 1 PS samples were wrongly identified as LDPE. This error is attributed to the adoption of the PCA feature extraction technique, which transforms data into a new set of variables called principal components. As already analysed in section IV.B, the principal components of the LDPE polymers may be dominated with significant phase shifts, leading to the poor classification accuracy. Table II shows that PS identification result is also poor with 50% recall. From the confusion matrix in Fig. 9, out of the 6 PS samples, only 3 were correctly identified, while the remaining 3 were identified as PET. PET with a precision of 67% means that only 6 samples were correctly identified while 3 PS were wrongly predicted as PET.

2) PCA-KNN Model: PCA feature reduction approach was applied to KNN model and prediction accuracy of 53.85% was achieved with an average computational time of 61ms. Looking closely at the classifcation report of PCA-KNN model in Table III, the model was not able to detect any LDPE samples. Observing the confusion metrix in Fig. 10, out of the 7 LDPE, 6 samples were wrongly detected as PP, while 1 LDPE was misclassified as PS. The result for PS is also poor, with a recall of 33%, indicating that only 2 out of 6 PS samples were correctly identified. PP also achieved a low precision of 36% with only 5 PP samples correctly identified, while 1 PS, 6 LDPE, and 2 HDPE samples were wrongly predicted as PP.

TABLE III. CLASSIFCATION REPORT OF PCA-KNN MODEL

Samples	**Total samples**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
PET	11	67	55	60
HDPE	10	88	70	78
PVC	8	62	100	76
LDPE	7	0	0	0
PP	10	36	50	42
PS	6	29	33	31
Average Accuracy				54
PCA-SVM Average computational time				61ms

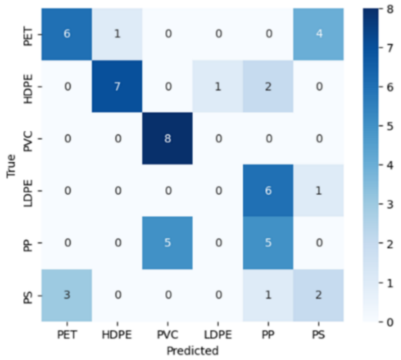


Fig. 10. Confusion Matrix for PCA -KNN model

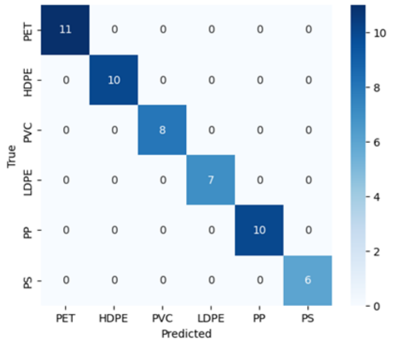


Fig. 11. Confusion Matrix for CFS -SVM model

3) CFS-SVM Model: The CFS approach was also applied to SVM to evaluate the performance of hybrid model. The confusion matrix in Fig. 11 shows the perfect prediction power of the CFS-SVM model which achieves a identification accuracy of 100%, with average computational time of 45ms.

4) CFS-KNN Model: Implementing the CFS-KNN hybrid model for deployment purposes, 40 best features were selected by the CFS. The KNN critical hyperparameter "k" was set to 1, and the hybrid CFS-KNN model achieved an exceptional identification accuracy of 100% with an average computational time of 40ms which is less than other models. The results solidifies the fact that the filter-based CFS is a robust feature selection technique for simplifying machine learning models and enhancing prediction accuracy.

V. CONCLUSION

With NIR spectroscopic sensing technique, this paper presents a new hybrid approach combining the KNN model with the correlation feature selection (CFS) technique for multi-class segregation of plastic materials. Experiments have been conducted to test six municipal polymers (PET, HDPE, PVC, LDPE, PP and PS). Results have suggested that the widely used PCA-SVM model achieves accuracy of 63.46% for identification of new and unknown polymers that have not been seen by the model, while the PCA-KNN model produces accuracy of 53.85%. By applying the selected features of CFS, the performance of SVM and KNN has been improved to 100% accuracy. This paper has demonstrated that the filter-based CFS feature selection algorithm provides more clarity for the identification and discrimination of different plastic materials. In terms of computational time, the CFS-KNN achieves 40ms which is less than the models of PCA-SVM, PCA-KNN and CFS-SVM. This high-performing, cost-effective CFS-KNN model paired with NIR spectroscopy offers a reliable solution for identifying valuable plastics in municipal waste, ensuring high-quality secondary raw materials for recycling. Future work will focus on identifying aged or damaged materials, as well as those contaminated with labels or food residues, using the developed models to provide a more practical solution.

ACKNOWLEDGMENT

The authors wish to acknowledge the UK Engineering and Physical Sciences Research Council (EP/W026228/1) for providing financial support for this research.

REFERENCES

[1] T. M. Letcher, Plastic Waste and Recycling: Environmental Impact, Societal Issues, Prevention, and Solutions, London: Academic Print-Elsevier, 2020.

[2] W. B. Group, "Global Waste to Grow by 70 Percent by 2050 Unless Urgent Action is Taken: World Bank Report," 20 September 2018. [Online]. Available: https://www.worldbank.org/en/news/press-release/2018/09/20/global-waste-to-grow-by-70-percent-by-2050-unless-urgent-action-is-taken-world-bank-report. [Accessed 12 October 2024].

[3] W. B. Group, "What a Waste: An Updated Look into the Future of Solid Waste Management," 20 September 2018. [Online]. Available: https://www.worldbank.org/en/news/immersive-story/2018/09/20/what-a-waste-an-updated-look-into-the-future-of-solid-waste-management. [Accessed 12 October 2024].

[4] L. Coltro, B. F. Gasparino and G. C. Queiroz, "Plastic material recycling: the importance of the correct identification," Polimeros, vol. 18, no. 2, pp. 119-125, 2018.

[5] S. Serranti and G. Bonifazi, "Post-Consumer Polyolefins (PP-PE) Recognition by Combined Spectroscopic Sensing Techniques," The Open Waste Management Journal, pp. 35-45, 2016.

[6] W. Musu, A. Tsuchida, H. Kawazumi and N. Oka, "Application of PCA-SVM and ANN Techniques for Plastic Identification by Raman Spectroscopy," in 1st International Conference on Cybernetics and Intelligent System (ICORIS), Denpasar, Indonesia, 2019.

[7] D. Scott, "A Two-Colour Near-Infrared Sensor for Sorting Recycled Plastic Waste," Measurement Science and Technology, vol. 6, pp. 156-159, 1995.

[8] G . Bonifazi, F. D. Maio, F. Potenza and S. Serranti, "FT-IR Analysis and Hyperspectral Imaging Applied to Postconsumer Plastics Packaging Characterization and Sorting," in IEEE Sensors Journal, vol. 16, no. 10, pp. 3428-3434, 2016.

[9] E. Scavino, D. A. Wahab, A. Hussain, H. Basri and M. M. Mustafa, "Application of automated image analysis to the identification and extraction of recyclable plastic bottles," Applied Physics & Engineering, vol. 10, no. 6, pp. 794-799, 2009, ISSN 1862-1775.

[10] K. Ohtani and M. Baba, "A Simple Identification Method for Object Shapes and Materials Using an Ultrasonic Sensor Array," in IEEE Instrumentation and Measurement Technology Conference Proceedings, Sorrento, Italy, 2006.

[11] H. W. Siesler, Y. Ozaki, S. Kawata and H. M. Heise, Near-infrared spectroscopy: principles, instruments, applications (Eds.), Chicago, USA: Wiley. corn. , 2008.

[12] A. Tehrani and H. Karbasi, "A novel integration of hyper-spectral imaging and neural networks to process waste electrical and electronic plastics," in 2017 IEEE Conference on Technologies for Sustainability (SusTech), Phoenix, AZ, USA, 2017.

[13] R. Maliks and R. Kadikis, "Multispectral data classification with deep CNN for plastic bottle sorting," in 2021 6th International Conference on Mechanical Engineering and Robotics Research (ICMERR), Krakow, Poland, 2021, pp. 58-65.

[14] S. Zhu and e. al., "Plastic Solid Waste identification system based on Near Infrared Spectroscopy in combination with support vector machine," Advanced Industrial and Engineering Polymer Research, pp. 77-81, 2019.

[15] K. Moirogiorgou, F. Raptopoulos, George Livanos, S. Orfanoudakis, M. Papadogiorgaki and M. Zervakis, "Intelligent robotic system for urban waste recycling," in Imaging Systems and Techniques, IEEE, 2023.

[16] L. M. Kumar, B. Pavan, P. V. Kalyan, N. S. Paul, R. S. Prakruth and T. Chinnu, "Design of an embedded based control system for efficient sorting of waste plastics using Near Infrared Spectroscopy," in IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2014.

[17] M. Cesetti, L. Placentino, G. Pinese, M. Libralato and P. Nicolosi, "Near infrared technology for material identification and selection," in 2015, Fotonica AEIT Italian Conference on Photonics Technologies, Turin, Italy, pp. 1-4, 2015.

[18] D. M. Adams, R. Miller and B. Stace, Laboratory Methods in Infrared Spectroscopy, London: Heyden & Son, 1972.

[19] M. R. Derrick, D. Stulik and J. M. Landry, Infrared Spectroscopy in Conservation Science, Los Angeles: The Getty Conservation Institute, 1999.

[20] W. Zhang, L. C. Kasun, Q. JieWang, Y. Zheng and Z. Lin, "A Review of Machine Learning for Near-Infrared Spectroscopy," Sensors , vol. 22, no. 24, 2022.

[21] C. D and M. D.L, "Alternative k-nearest neighbour rules in supervised pattern," Analytica Chimica Acta, vol. 136, no. 153, pp. 15-27, 1982.

[22] L. Wei-Meng, Python Machine Learning, Indianapolis , John Wiley & Sons Inc., 2019.

[23] Y. P. Jiao.; Z. C. Jiao; X.S. Chen; S.M. Fei, "Preprocessing methods for near-infrared spectrum calibration". J. Chemometrics. 2020, 34, e3306.

[24] D. Vigren, "Identification of environmental plastic samples from a caribbean beach survey by near-infrared spectroscopy with the micronir and chemometric analysis," 2015.