Journal of Computer Technology and Software | Vol. 1, No. 1, 2024

Advancing Soil Microplastics Detection: Insights from Hyperspectral Imaging Technology

Cheng Chen, Jiawei Zhang, Jie Guo

George Washington University,George Washington University,George Washington University
Cc1199@gmail.com, zjw6677@gmail.com,jjg666@gmail.com

Abstract: Addressing the rising concern of microplastic pollution in soil, this study proposes employing hyperspectral imaging technology for efficient detection. Utilizing supervised classification algorithms, including Support Vector Machine (SVM), Mahalanobis Distance (MD), and Maximum Likelihood (ML), microplastic pollutants in soil are directly identified and classified. Experiments conducted within a wavelength range of 400-1000 nm reveal SVM as the most suitable algorithm, achieving an average identification accuracy of 84% for white polyethylene (PE) microplastics within the 1-5 mm particle size range.

Keywords: Microplastics; Marine pollution; Hyperspectral; Support vector machine.

1. Introduction

The widespread pollution of microplastics in soil has becomes a significant concern. However, to date, research on microplastics in terrestrial ecosystems has not received much attention [1,2]. The presence of microplastics has been detected in soil samples collected near an industrial area in Sydney, with concentrations ranging from 300 mg/kg to 67,500 mg/kg [3]. Long-term use of plastic films, organic fertilizers, sewage irrigation, and sludge application in agricultural practices are the primary pathways for microplastics to enter the soil environment [1,4]. Among these, the use of significant amounts of sewage sludge as a raw material in agricultural composting processes cannot completely remove microplastics through processes such as lime stabilization, anaerobic digestion, and thermal drying [5], making sewage sludge an important pathway for microplastics to enter the soil environment. Based on the quantity of microplastics found in sludge and assuming limited sludge application rates (5 tons of dry sludge per hectare every 3 years), it has been estimated that 1.67 to 408 thousand tons of plastic particles enter agricultural soil in Germany annually through sludge fertilization [6]. Additionally, microplastics can also enter non-agricultural soil environments through improper waste disposal, flooding events, and atmospheric deposition.

Polymer materials are highly stable and resistant to degradation in the soil environment. Even though the small-sized plastic fragments that are difficult to observe visually may not draw immediate attention, they can have persistent effects on the soil for several years or even decades [7], posing challenges for their management. Microplastics, with their stable chemical properties and hydrophobic surfaces, can provide a growth environment for microorganisms and promote biofilm formation [8], potentially affecting the ecological functions of soil microorganisms. Microplastics can also introduce pollutants and toxic additives into the soil environment, which can negatively impact the health of organisms and disrupt biodiversity in soil ecosystems [9-11]. Smaller microplastic particles in the soil can be ingested by some protists [12] and earthworms, causing intestinal damage and affecting their growth [13,14]. Moreover, these microplastics can potentially enter higher trophic levels through the food chain [2].

In this study, soil samples containing plastic were collected from agricultural areas near the Quzhou West Exit of Quzhou City, Zhejiang Province, China, with a total mass of 3 kg. The samples were divided into two groups. One group was used for microplastic identification after extraction using a saturated NaCl solution, while the other group was used to establish a method for identifying soil microplastics using hyperspectral imaging technology. Five types of materials, including fresh leaves, wilted leaves, rocks, branches, and extracted microplastics, were collected to simulate different categories of microplastics in the soil environment. Hyperspectral images were obtained to establish the best hyperspectral image for microplastic detection by comparing the results of three supervised classification methods. The visual distribution of microplastics in the soil was also determined.

2. Experimental Sample Preparation

White microplastics (1-5 mm), black microplastics (1-5 mm), rocks, wilted leaves, fresh leaves, and dry branches were mixed and randomly distributed on the surface of dry soil. Each type of sample was replicated 10 times, and hyperspectral images of the samples were subsequently captured.

To validate the stability of the hyperspectral microplastic detection model established in this study, the experiments were repeated five times independently. The soil microplastic samples were scanned using a hyperspectral imaging system to obtain their image and spectral information.

3. Supervised Classification Models

3.1. ML Model (Maximum Likelihood)

In image classification models, the Maximum Likelihood (ML) model is a parameter estimation method. It assumes that the observed data is independently drawn from a probability distribution. The goal of the ML model is to estimate the parameters of the model by maximizing the likelihood function of the observed data. The likelihood function measures the probability of observing the data under different parameter values. In the case of image classification models, we assume that the observed data is generated from a probabilistic model, and the parameters of that model need to be determined. By defining an appropriate probability model and its parameters, we can compute the probability of observing the data given the parameters. The ML model estimates the parameters by finding the values that maximize the likelihood function. We use optimization algorithms such as gradient descent to search the parameter space and find the maximum likelihood estimate. Through the estimation of the ML model, we can obtain optimal parameter values that best describe the relationship between image features and classes. With these parameters, we can perform image classification and prediction by assigning new image data to the most likely class.

The ML model uses the maximum likelihood estimation method in image classification to estimate model parameters by maximizing the likelihood function of the observed data. This approach helps us find parameter values that best describe the relationship between image features and classes, enabling accurate image classification and prediction.

Fig. 1 Soil samples collected from farmland near the highway

3.2. MD Model (Mahalanobis distance)

In image classification models, the MD (Mahalanobis Distance) model is a classification method based on the Mahalanobis distance measure. This model classifies samples by calculating the Mahalanobis distance between them. The Mahalanobis distance is a distance measure that takes into account the correlation between features. In image classification, we represent image samples as vectors with multidimensional features, where each feature represents a certain attribute or description of the image. The basic idea of the MD model is to measure the similarity between samples by calculating the Mahalanobis distance between them, considering the correlation between features. The Mahalanobis distance reflects the differences between samples more accurately in the distance measure. For a given image sample, the MD model first calculates the Mahalanobis distance between the sample and samples from each class. Then, based on the magnitude of the Mahalanobis distance, the sample is assigned to the class it is most similar to. When computing the Mahalanobis distance, it is necessary to estimate the covariance matrix of each class's samples, which reflects the correlation information between the sample features. By using the covariance matrix, the Mahalanobis distance can be calculated more accurately, thereby improving the accuracy of classification.

The MD model in the image classification model uses the Mahalanobis distance to measure the similarity between samples. By calculating the Mahalanobis distance between the sample and samples from different classes, the sample is assigned to the class it is most similar to. This approach considers the correlation between features and can improve the accuracy of image classification.

3.3. SVM Model (Support Vector Machine)

In image classification models, the Support Vector Machine (SVM) model is a commonly used classification algorithm. It achieves classification of image samples by constructing an optimal hyperplane. The goal of the SVM model is to find a hyperplane with the largest margin that separates image samples of different classes. The hyperplane is a (d-1)-dimensional decision boundary, where d is the feature dimension of the image samples. The SVM model first represents image samples as feature vectors, where each feature represents a certain attribute or description of the image. By solving an optimization problem, the parameters of the hyperplane are determined to maximize the projection distance of training samples onto the hyperplane. During the training process, the SVM model divides the image samples into two classes and defines the hyperplane based on support vectors. Support vectors are the training sample points closest to the hyperplane, and they play a crucial role in defining the hyperplane. For new image samples, the SVM model can classify and predict them based on their position relative to the hyperplane. Samples on one side of the hyperplane are assigned to one class, while samples on the other side are assigned to the other class. The SVM model has advantages such as the ability to handle high-dimensional data, good generalization capability, and sensitivity to a small number of support vectors. It is widely used in image classification tasks and can effectively perform image classification and prediction.

The SVM model in the image classification model achieves the classification of image samples by constructing an optimal hyperplane. It relies on the concept of support vectors and determines the parameters of the hyperplane by maximizing the margin, enabling accurate classification and prediction of new image samples.

4. Evaluation Metrics

ENVI (The Environment for Visualizing Images) software is a professional hyperspectral image data processing software. The analysis of hyperspectral data in this experiment was performed using ENVI 4.6 software.

This study aimed to validate the stability of the hyperspectral classification model in identifying and classifying microplastics in soil environments through multiple repeated experiments. The detection capability of the model was represented by calculating the average precision (P) and recall (R) values for each experimental group. Precision (P) is an important criterion for evaluating whether the hyperspectral model can correctly identify microplastics in soil samples. It is the percentage of true positive samples (TPS) in all positive samples, including both true positive samples (TPS) and false positive samples (FPS). Recall (R) is used to evaluate the model's ability to identify true positive microplastics from environmental samples. It represents the proportion of positive samples among all samples. Recall (R) is the score of true positive samples (TPS) in the sum of false negative samples (FNS) and true positive samples (TPS). The formulas are as follows:

$$P = TPS / (TPS + FPS) \tag{1}$$

$$R = TPS / (TPS + FNS) \tag{2}$$

5. Results and Discussion

Three supervised classification methods, MD, ML, and SVM, were used to classify and identify PE microplastics in the soil surface and other interfering substances based on hyperspectral images. The performance of the three supervised classification models in identifying microplastics in the soil environment was evaluated by calculating the precision (P) and recall (R) values for six materials (white PE, black PE, rocks, withered leaves, fresh leaves, dry branches). The classification results (P and R values) for PE microplastics with particle sizes ranging from 1-5 mm and four other interfering substances are listed in Table 1. To test the accuracy and stability of the three classification models, five simulated samples were prepared in this experiment. Hyperspectral images of the samples were collected, and the three classification models were used for classification and identification. The average values and standard deviations (SD) of P and R from the five experimental results were calculated and presented in Table 1. Due to the similarity between the soil spectral curve and the spectral curves of rocks and withered leaves, some soil backgrounds in certain samples were incorrectly identified as rocks and withered leaves. Although the areas of these misclassified regions were small, their number was significant. Therefore, the precision (P) values for these rocks and withered leaves could not be accurately calculated. In this experiment, "NaN" was used to indicate these difficult-to-count invalid values.

Comparing the identification results of the three classification methods in Table 5.1, it can be observed that the SVM algorithm exhibited the best performance in identifying PE microplastics in the soil. The recall (R) values for white and black PE microplastics reached 94% and 92%, respectively. This indicates that the SVM algorithm achieved a correct identification rate of over 90% for white and black PE microplastics in complex soil environments.

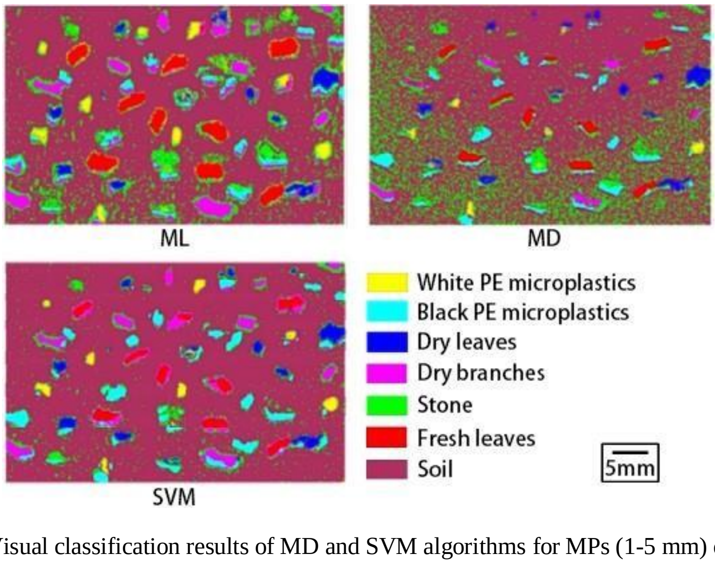

Fig. 2 Visual classification results of MD and SVM algorithms for MPs (1-5 mm) detection

From Fig.2, it can be observed that the hyperspectral images processed using the SVM algorithm exhibit smoother and clearer backgrounds compared to the MD and ML algorithms. The SVM algorithm also shows fewer interfering substances around the classified and recognized materials, and it performs well in identifying microplastics of both colors. This indicates that the SVM algorithm can effectively filter out noise information from hyperspectral images, resulting in processed hyperspectral data with a higher signal-to-noise ratio, which is beneficial for the identification and classification of various substances. Comparing the precision and recall values of the three algorithms in Table 1, it is evident that the SVM algorithm demonstrates superior recognition performance for various materials in hyperspectral images.

Tab. 1 Classification results of soils samples covered with MPs (1-5 mm) by using MD、ML and SVM algorithms.

	White PE microplastics		Black PE microplastics		Stone		Dry leaves		Fresh leaves		Dry branches	
	P	R	P	R	P	R	P	R	P	R	P	R
ML												
S1	100	100	63	100	NaN	NaN	100	90	100	100	53	100
S2	100	80	63	100	NaN	NaN	77	100	100	100	100	100
S3	71	100	77	100	100	50	100	60	91	100	83	100
S4	69	90	57	80	43	30	NaN	NaN	100	100	77	100
S5	75	90	33	20	100	50	50	80	100	100	100	100
Avg	83	92	58	80	81	42	82	83	98	100	83	100
SD(n=5)	16	8	16	35	33	12	25	17	4	0	19	0
MD												
S1	89	80	50	100	NaN	NaN	82	90	100	100	100	80
S2	100	50	63	100	100	40	53	100	100	100	100	80
S3	100	50	56	100	100	50	NaN	NaN	100	100	100	100
S4	82	90	56	100	67	20	NaN	NaN	100	100	63	100
S5	56	100	60	90	100	60	NaN	NaN	100	100	100	100
Avg	81	84	57	98	92	43	67	95	100	100	93	92
SD(n=5)	16	21	5	4	17	17	21	7	0	0	17	11
SVM												
S1	100	100	50	100	100	30	83	100	100	100	83	100
S2	100	90	71	100	100	20	71	100	100	100	100	100
S3	71	100	63	100	100	40	86	60	100	100	90	100
S4	82	90	54	70	50	40	75	90	100	100	100	100
S5	64	90	50	90	100	30	90	90	100	100	100	100
Avg	84	94	58	92	90	32	81	88	100	100	95	100
SD(n=5)	16	5	9	13	22	8	7	16	0	0	8	0

Sn: Parallel sample n,NaN: Invalid value

6. Conclusion

In conclusion, this study explores the efficacy of hyperspectral imaging technology in directly identifying microplastics on soil surfaces, obviating the need for their extraction. With particle sizes ranging from 1-5 mm, the study calculates average spectra of materials by extracting spectral regions of interest from hyperspectral images. Employing three supervised classification models, the SVM algorithm demonstrates superior recognition performance, achieving an average precision of 84% and a recall of 94% for white PE microplastics, and 58% precision with a 92% recall for black PE microplastics. These findings underscore the significant potential of hyperspectral imaging technology in detecting microplastics across complex soil surfaces

Acknowledgments

This work was partially supported by Zhejiang Public Welfare Project (Grant No. LGN21F010001), Quzhou Science and Technology Project (Grant No. 2022K94)

References

[1] Liu H, Yang X, Liu G, et al. Response of soil dissolved organic matter to microplastic addition in Chinese loess soil [J]. Chemosphere, 2017, 185: 907-917.

[2] Rillig M C. Microplastic in terrestrial ecosystems and the soil? [J]. Environmental Science & Technology, 2012, 46(12): 6453-6454.

[3] Fuller S, Gautam A. A procedure for measuring microplastics using pressurized fluid extraction [J]. Environmental Science & Technology, 2016, 50(11): 5774-5780.

[4] Nizzetto L, Futter M, Langaas S. Are agricultural soils dumps for microplastics of urban origin? [J]. Environmental Science & Technology, 2016, 50(20): 10777-10779.

[5] Mahon A M, O'Connell B, Healy M G, et al. Microplastics in sewage sludge: Effects of treatment [J]. Environmental Science & Technology, 2017, 51(2): 810-818.

[6] Bläsing M, Amelung W. Plas tics in soil: Analytical methods and possible sources [J]. Science of The Total Environment, 2018, 612: 422-435.

[7] Briassoulis D, Babou E, Hiskakis M, et al. Analysis of long-term degradation behaviour of polyethylene mulching films with pro-oxidants under real cultivation and soil burial conditions [J]. Environmental Science and Pollution Research, 2015, 22(4): 2584-2598.

[8] Harrison J P, Schratzberger M, Sapp M, et al. Rapid bacterial colonization of low-density polyethylene microplastics in coastal sediment microcosms [J]. BMC microbiology, 2014, 14: 232.

[9] Browne M A, Niven S J, Galloway T S, et al. Microplastic moves pollutants and additives to worms, reducing functions linked to health and biodiversity [J]. Current Biology, 2013, 23: 2388-2392.

[10] Koelmans A A, Besseling E, Foekema E M. Leaching of plastic additives to marine organisms [J]. Environmental Pollution, 2014, 187: 49-54.

[11] Kwon J, Chang S, Hong S H, et al. Microplastics as a vector of hydrophobic contaminants: Importance of hydrophobic additives[J]. Integrated Environmental Assessment and Management, 2017, 13(3): 494-499.

[12] Rillig M C, Bonkowski M. Microplastic and soil protists: A call for research [J]. Environmental Pollution, 2018, 241: 1128-1131.

[13] Lwanga E H, Gertsen H, Gooren H, et al. Microplastics in the terrestrial ecosystem: Implications for Lumbricus terrestris (Oligochaeta, Lumbricidae) [J]. Environmental Science & Technology, 2016, 50: 2685-2691.

[14] Cao D, Wang X, Luo X, et al. Effects of polystyrene microplastics on the fitness of earthworms in an agricultural soil [J]. IOP Conference Series: Earth and Environmental Science, 2017, 61(1): 12148.