Conventional Machine Learning Approach for Waste Classification

Kharittha Jangsamsi*
Computer Engineering Department, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
kharittha.jan@kmutt.ac.th

ABSTRACT

Waste management is a complex and challenging process, especially waste classification to sort waste by categories. The paper aims to overcome these challenges by proposing a waste classification approach that uses various feature extraction algorithms along with a support vector machine (SVM). The purpose is to identify the most effective feature for building a classification model, even with a low number of samples and high intra-class variance. SVM was used for classification while Fourier descriptors (FDs), histogram of oriented gradients (HOG), and local binary pattern (LBP) were used for feature extraction. The dataset used in this paper was obtained from Kaggle.com and Google.com with different types of vision problems. The experimental results showed that classification with LBP feature extraction achieves the highest accuracy. This accuracy is higher than the experiments with other feature extractions.

CCS CONCEPTS

• **Computing methodologies**; • **Machine learning**;

KEYWORDS

waste classification, feature extraction, Fourier descriptors, histogram of oriented gradients, local binary pattern

ACM Reference Format:
Kharittha Jangsamsi. 2023. Conventional Machine Learning Approach for Waste Classification. In *2023 6th Artificial Intelligence and Cloud Computing Conference (AICCC) (AICCC 2023), December 16–18, 2023, Kyoto, Japan.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3639592.3639594

1 INTRODUCTION

Based on data from the World Bank, there has been a significant surge in waste production in recent years. In 2020, the global generation of solid waste amounted to approximately 2.24 billion tonnes, and it is projected to reach 3.88 billion tonnes by 2050 [1, 2]. Due to the volume of waste, sorting the waste at the beginning of the waste management process will increase the amount of recyclable materials and lessen the chance that other materials will contaminate the environment. If the waste is separated and recycled using the most modern technological breakthroughs, it gains value and becomes a useful entity. Nonetheless, the existing recycling procedures demand manual sorting at recycling facilities and rely on a complex system of extensive filters to distinguish specific objects. Consequently, waste classification has gained considerable attention among researchers and holds great promise as an application of computer vision within the industry.

Computers can now analyze and extract information from visual input owing to the branch of study known as computer vision. The two most popular uses of computer vision are object detection and image classification. Image classification is the process of predicting the class or category of an object. Both object detection and image classification have seen substantial use of machine learning and deep learning techniques. Image classification-based machine learning exploits the potential of algorithms to learn previously unknown information from a dataset containing both structured and unstructured patterns, while deep learning, a type of machine learning that uses numerous hidden layers within a model, is the most commonly used approach in the field today.

2 RELATED WORKS

The research community has recently given image-based waste classification systems a lot of attention. Over the past few years, several intriguing solutions using a variety of approaches and technology have been put forth. This section reviews the existing waste classification systems that operate in varied environments utilizing conventional machine learning and deep learning techniques. The advantages and limitations of each approach are described to support our study.

2.1 Waste classification systems based on conventional machine learning techniques

There are some researchers who concentrate on extracting features to be carried out the unique features of a waste object that can describe the characteristics of the object.

In 2020, K. Ahmad et al. [3] introduced an approach to classify waste called double fusion that combines multiple deep learning models using feature and score-level fusion methods. There are three steps: feature extraction, classification, and fusion. Deep features were extracted using several deep models, followed by Support Vector Machine (SVM) based classification. While six different fusion methods were employed and compared, including two feature-level fusion schemes, i.e., discriminant correlation analysis, and simple concatenation of deep features and four late fusion methods, i.e., particle swarm optimization, genetic modeling of deep features, induced ordered weighted averaging, and a baseline method. The goal is to leverage the diverse and complementary features extracted by different models to enhance waste classification.

*kharittha.jan@mail.kmutt.ac.th

AICCC 2023, December 16–18, 2023, Kyoto, Japan

ACM ISBN 979-8-4007-1622-5/23/12
https://doi.org/10.1145/3639592.3639594

Their results demonstrated that the fusion of multiple deep models outperforms the individual models by exploiting their learning capabilities. However, the performances of the models vary across different waste categories due to the lower number of samples in the class and high intra-class variations. Next, A. P. Puspaningrum et al. [4] proposed waste classification using the Support Vector Machine (SVM) for classification and Scale Invariant Feature Transform - Principal Component Analysis (SIFT-PCA) for feature extraction. The SIFT algorithm was a suitable method for detecting and describing waste image keypoints and descriptors while the PCA algorithm was used to reduce the dimensionality of the extracted features. This is important to accurately recognize waste objects and reduce the dimensionality of the extracted feature data. The experimental results indicated that waste classification using SIFT feature extraction outperformed the SIFT-PCA combined feature extraction method in terms of accuracy. According to the PCA algorithm reduced the important components of the SIFT feature vectors used in the classification process. S. K. Behera et al. [5] then represented a deep learning model to improve waste classification ability, along with the use of Support Vector Machine (SVM). They used You Only Look Once version 3 (YOLOv3) for object detection and feature extraction. Their experimental results indicated that the improvement in terms of accuracy and memory utilization.

2.2 Waste classification systems based on deep learning techniques

In recent years, deep learning techniques have been widely used in image classification and have proved very effective at waste classification as follows.

In 2019, CompostNet model [6] was presented for waste classification. The waste materials were categorized using a convolutional neural network, with an emphasis on compostable, recyclable, and landfill materials. Two different approaches were designed: one using a custom model and the other using transfer learning with a pre-trained image classification model, i.e., MobileNet. The experimental results showed that the transfer learning approach had better success. This paper does not explore the use of other deep learning models for waste classification to compare performance with CompostNet. Next, A. Vo et al. [7] proposed deep neural networks called DNN-TC for waste classification. Their model is an improvement of ResNext to improve the predictive performance of waste classification. Their dataset belongs to three different classes: Organic, Inorganic, and Medical wastes. The experimental results showed that their model improved the accuracy of waste classification, particularly in distinguishing between Organic and Inorganic classes. However, the model did not significantly improve the classification accuracy for the Medical class compared to ResNext101. In 2020, Z. Yang et al. [8] proposed WasNet system, which includes a lightweight neural network for waste classification, an intelligent trash can, a waste recognition application, and a visualization and decision support platform. Their model referred to how high-precision neural networks are currently designed. To get the optimum model for their dataset, the experiment was continuously changed for the number of convolutional layers, as well as the depth and width of the model. The system was tested on an extended dataset and showed reliable performance. The authors concerned the challenge of multi-label waste classification, where waste items may belong to multiple categories. In 2021, a fully automated waste management system [9] was developed for reducing the risk of health issues for municipal workers and preventing the spread of transmissible diseases. The authors utilized Convolutional Neural Networks (CNN) for waste classification and employed an Internet of Things (IoT) system to integrate the waste management process with hardware devices, such as servo arms controlled by Arduino, for automated segregation. Their system achieved accurate waste classification into organic and recyclable categories with high accuracy. Next, several convolutional neural networks (CNN), specifically AlexNet, DenseNet121, and SqueezeNet were implemented for waste classification by A. E. B. Alawi et al. [10]. The models were trained on waste images to learn the features and patterns associated with different waste categories, enabling accurate classification. The performance of the models was evaluated based on accuracy, with DenseNet121 achieving the highest accuracy. The authors suggest expanding the number of images and using different datasets to improve performance. Then, GarbageNet [11] was developed. This framework utilizes transfer learning techniques, feature mixup module, memory pool, and metric-based classifier to enhance the model's capability and achieve noise-robust features. The experimental results demonstrated that GarbageNet outperforms other visual recognition models in terms of accuracy and robustness. After that, R. Faria et al. [12] introduced a method to automatically classify waste into four categories: organic waste, glass waste, metal waste, and plastic waste. Several convolutional neural networks (CNN), including 3-layer CNN, VGG16, VGG19, Inception-V3, and ResNet50, were implemented for training on the dataset. Among them, VGG16 achieved the highest accuracy outperforming other models. Their method does not concern testing on a larger scale or in different environmental contexts, which could affect the reliability and applicability of the results. Finally, the use of EfficientNet-B0 and MobileNets-V1 was represented to classify waste accurately by W. Mulim et al. [13] and I. F. Nurahmadan et al. [14], respectively. For EfficientNet-B0 model, the authors also modified EfficientNet-B0 model with additional layers, such as 2D pooling, batch normalization, and dropout including tuning hyperparameters to improve the classification performance. The datasets used for waste classification were pre-processed by scaling them to 224x224 and adding an augmentation layer before inputting them into the model. The model encountered difficulty with recyclable images, especially in terms of recall capability, indicating the need for further analysis and dataset variety to address this issue. For MobileNets-V1, the authors used to classify waste for three categories: Non-Recyclable, Organic, and Recyclable. The MobileNets-V1 model showed excellent performance in waste classification, indicating its effectiveness in accurately classifying waste types. In 2023, M. Polchan et al. [15] compared five models including MobileNetV2, InceptionV3, ResNet34, VGG16, and CNN to develop a mobile application for waste classification. The waste was classified into four categories: Wet waste, General waste, Recyclables, and Hazardous waste. The MobileNetV2 was found to have the highest validation results. The authors suggest using data augmentation for the learning of the VGG16 and InceptionV3 models to address data overfitting issues. Recently, A. Pandey et al. [16] investigated waste classification using Convolutional Neural

Networks (CNNs) and compared it with Support Vector Machine (SVM) for categorizing three categories: Plastic, Paper, and Metal. The authors explored the use of two popular CNN architectures, i.e., VGG16 and FastNet-34. Their study compared the performance of CNNs with SVM for waste classification. The experimental results showed that SVM had higher accuracy compared to CNN for three waste categories, but CNN had potential for higher accuracy once hyperparameter optimization challenges are overcome.

According to the literature reviews, we found that Support Vector Machine (SVM) is one of the popular classification algorithms that is known to perform well to classify data [17] and obtains higher accuracy compared to convolutional neuron networks (CNNs) for waste classification. Therefore, waste classification is carried out using the SVM algorithm with various feature extraction algorithms in this paper. We aim to identify the most effective feature for building a classification model, even with a low number of samples and high intra-class variance. We also focus on categorizing waste to determine whether it falls into the bottle, can, or snack package category, with the primary goal of enhancing waste management process.

3 PROPOSED METHOD

An overview of our proposed method is shown in Figure 1. There are several processes including data collection, feature extraction, model training, and model evaluation. Details of each process are described as follows.

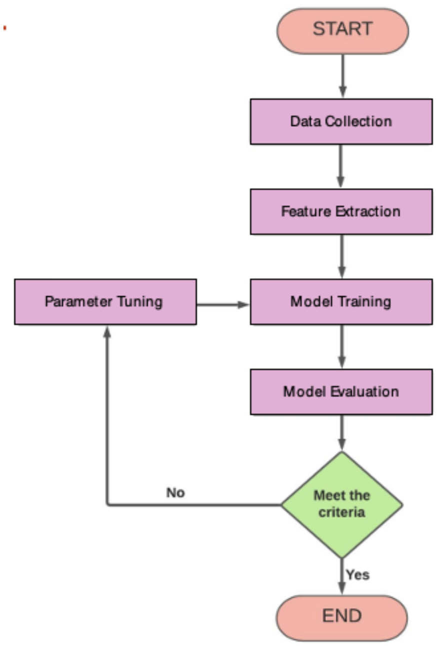


Figure 1: An overview of our proposed method

3.1 Data Collection

This paper uses waste images as dataset obtained from Kaggle.com and Google.com with various types of vision problems such as illumination conditions, insufficient data, and object size and position. The dataset is in JPG format and contains 165 waste images of various sizes. The dataset has been categorized into three categories, namely bottle, can, and snack package. Each image has been labeled and named according to the name of its category followed by a number. Each object in the dataset from Kaggle.com is captured on a white background and each object in the dataset from Google.com is captured on a complex background. Figure 2 displays several examples of waste images from each category. The dataset is divided into a training set comprising 85% of the data and a test set comprising 15% of the data for each category.

3.2 Feature Extraction

Feature extraction is a process that extracts a set of unique features of a waste object to increase the effectiveness and efficiency of waste classification [18]. In this paper, we use three algorithms to extract and describe feature points of a waste object, i.e., Fourier Descriptors (FDs), Histogram of Oriented Gradients (HOG), and Local Binary Pattern (LBP). Note that, the images are resized to 512 by 512 pixels before being extracted feature points.

3.2.1 Fourier Descriptors (FDs). Before using Fourier descriptors (FDs), we first perform Canny edge detection to find edges in each waste image. Then, the Fourier transform is applied to the sample points of these edges, resulting in a shape descriptor. Note that FDs are invariant to location, scale, rotation, and starting point.

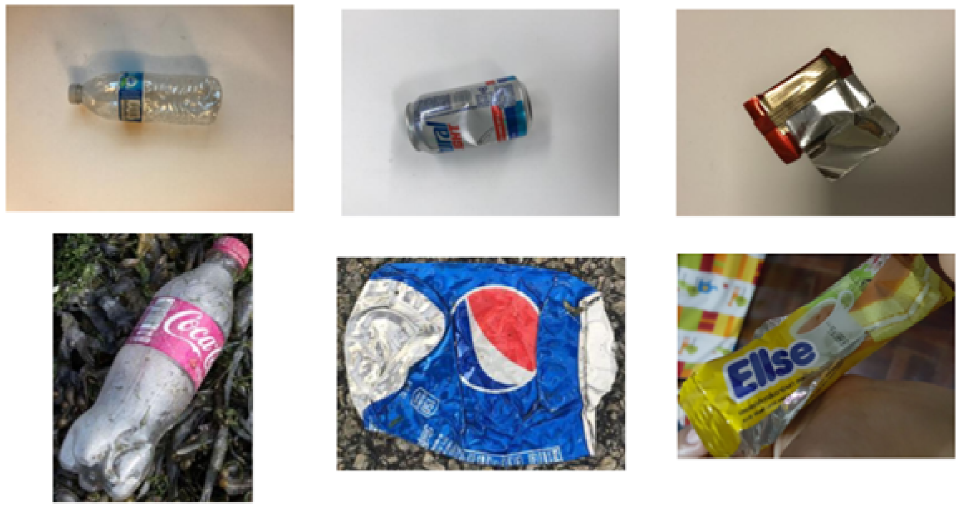
3.2.2 Histogram of Oriented Gradients (HOG). The Histogram of Oriented Gradients (HOG) is a feature descriptor. It counts the instances of gradient orientation within a given region of the waste image. The HOG description emphasizes the structure or shape of the waste object. Since it computes features based on the magnitude and direction of the gradient, it is superior to other edge descriptors. It creates histogram for the areas of the waste image based on the magnitude and direction of the gradient. Figure 3 illustrates the examples of HOG generated from each category.

3.2.3 Local Binary Pattern (LBP). Local Binary Pattern (LBP) serves as a local descriptor based on the pixel neighborhood around a given pixel. This neighborhood is defined by P neighbors within a radius of R. LBP is known for its strong ability to identify all potential edges in the image. Figure 4 shows the examples of LBP generated from each category.

3.3 Model Training

Support vector machine (SVM) is learning algorithm and supervised learning model that examine data for regression and classification. Finding a hyperplane in an N-dimensional space (N is the number of features) that categorizes the feature points clearly is the goal of the support vector machine method. The following are some common techniques for performing multi-classification using SVM: One versus One (OvO), and One versus Rest (OvR). Since the runtime is substantially less and the results are often identical, OvR classification is typically used in practice [17].

Sets of features extracted using several algorithms from the previous process are followed by SVM to classify waste of each category.



bottle can snack package

Figure 2: Examples of waste images from each category (1st and 2nd rows are the examples from Kaggle.com and Google.com, respectively.)

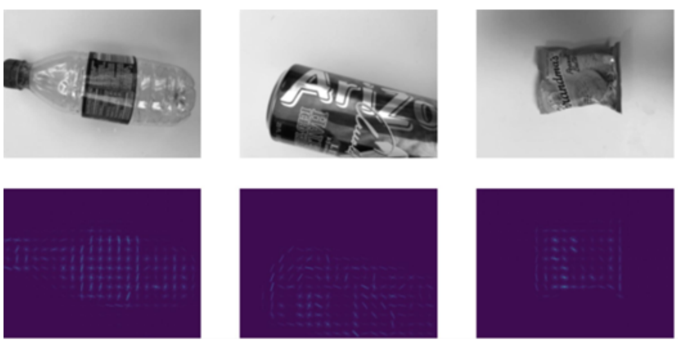


Figure 3: Examples of HOG generated from each category, i.e., bottle, can, and snack package, respectively

3.4 Model Evaluation

The assessment of classification performance involves the computation of accuracy. Accuracy is a method used to measure the total number of correct predictions made by a model and can be quantified using the formulas (1) - (4) below.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$precission = \frac{TP}{TP + FP} \tag{2}$$

$$recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = 2 * \frac{precision * recall}{precision + recall} \tag{4}$$

where TP or True Positive is correct results, TN or True Negative is correct absent results, FP or False Positive is unexpected results, and FN or False Negative is False results. Precision is the ratio of True Positives to all positive predictions, while recall is the measure of the model correctly identifying True Positives.

4 EXPERIMENTAL RESULTS

In this paper, the training process is divided into three sets, consisting of the use of each feature extraction algorithm followed by SVM. The SVM parameters, namely C and gamma (γ) parameters, are 1 and 0.0001, respectively. The experimental results can be seen in Table 1 below which are presented in terms of precision, recall, and F1 score for each category. Table 2 shows the average accuracy of each feature extraction algorithm followed by SVM.

From the experimental results, it is clear that the use of only one LBP can extract image features and use them for classification. It is not good for using the descriptors of HOG since HOG was not able to extract the features of the images well enough for use in classification, which caused the results of drop as well, and on the part of FDS, there were no good results. This may be because we have to sample the same points from each image, which causes some information to be lost.

5 CONCLUSION

The paper proposes an approach to waste classification based on conventional machine learning techniques. The proposed method combines feature extraction to obtain unique features of waste objects and a Support Vector Machine (SVM) for waste classification. Experimental results have shown that Local Binary Pattern (LBP) followed by SVM can classify the waste images into three different categories and outperforms the other methods. LBP is the effective feature for building a classification model with a low number of samples and high intra-class variance. Future work should acquire a new dataset with a larger number of waste categories and different complexity levels to achieve good accuracy and should consider

Figure 4: Examples of LBP generated from each category, i.e., bottle, can, and snack package, respectively

Table 1: Results in terms of precision, recall, and F1 score for each category

category	feature extraction algorithm	precision	recall	F1 score
bottle	FDs	0.36	0.33	0.34
	HOG	0.56	0.60	0.58
	LBP	0.83	0.67	0.74
can	FDs	0.31	0.24	0.27
	HOG	0.53	0.53	0.53
	LBP	0.81	0.76	0.79
snack package	FDs	0.38	0.61	0.54
	HOG	0.76	0.72	0.74
	LBP	0.77	0.94	0.85

Table 2: Average accuracy of each feature extraction algorithm followed by SVM

feature extraction algorithm	average accuracy
FDs	0.40
HOG	0.62
LBP	0.80

using other classifiers to handle multi-class classification compared with SVM.

REFERENCES

[1] Khan N. Sami, Zian Md M. A. Amin, and Raini Hassan. 2020. Waste management using machine learning and deep learning algorithms. Int. J. Perceptive Cogn. Comput. 6, 2 (December 2020), 97–106.

[2] Sana Shahab, Mohd Anjum, and M. S. Umar. 2022. Deep learning applications in solid waste management: A deep literature review. Int. J. Adv. Comput. Sci. Appl. 13, 3, 381–395.

[3] Kashif Ahmad, Khalil Khan, and Ala A. Fuqaha. 2020. Intelligent fusion of deep features for improved waste classification. IEEE Access. 8 (May 2020), 2169–3536.

[4] Adita P. Puspaningrum, Sukmawati N. Endah, Priyo S. Sasongko, Retno Kusumaningrum, Khadijah, Rismiyati, and Ferda Ernawan. 2020. Waste classification using support vector machine with SIFT-PCA feature extraction. In Proceedings of 4th International Conference on Informatics and Computational Sciences (ICICoS). IEEE, Semarang, Indonesia.

[5] Saswati K. Behera, Aouthithiye Barathwaj SR Y, Vasundhara L, Saisudha G, Haariharan N C, and Sathya Priya J. 2020. AI based waste classifier with thermo-rapid composting. In Proceedings of International Conference on Power, Energy, Control and Transmission Systems (ICPECTS). IEEE, Chennai, India.

[6] Sarah Frost, Bryan Tor, Rakshit Agrawal, and Angus G. Forbes. 2019. CompostNet: an image classifier for meal waste. In Proceedings of IEEE Global Humanitarian Technology Conference (GHTC). IEEE, Seattle, WA, USA.

[7] Anh H. Vo, Le H. Son, Minh T. Vo, and Tuong Le. 2019. A novel framework for trash classification using deep transfer learning. IEEE Access. 7 (December 2019), 178631–178639.

[8] Zhihu Yang and Dan Li. 2020. WasNet: a neural network-based garbage collection management system. IEEE Access. 8 (June 2020), 103984–103993.

[9] Mallikarjuna Gowda C. P, Sachin Yadav, Aditi Shanmugam, Hima V, and Niharika Suresh. 2021. Waste classification and segregation: machine learning and IOT approach. In Proceedings of 2nd International Conference on Intelligent Engineering and Management (ICIEM). IEEE, London, United Kingdom.

[10] Abdulfattah E. Ba Alawi, Ahmed Y. A. Saeed, Fatima Almashhor, Reem Al-Shathely, and Ahmed N. Hassan. 2021. Solid waste classification using deep learning techniques. In Proceedings of International Congress of Advanced Technology and Engineering (ICOTEN). IEEE, Taiz, Yemen.

[11] Jianfei Yang, Zhaoyang Zeng, Kai Wang, Han Zou, and Lihua Xie. 2021. GarbageNet: a unified learning framework for robust garbage classification. IEEE TAI. 2, 4 (August 2021), 372–380.

[12] Rushnan Faria, Fahmida Ahmed, Annesha Das, and Ashim Dey. 2021. Classification of organic and solid waste using deep convolutional neural networks. In Proceedings of 9th Region 10 Humanitarian Technology Conference (R10-HTC). IEEE, Bangalore, India.

[13] William Mulim, Muhammad F. Revikasha, Rivandi, and Novita Hanafiah. 2021. Waste classification using EfficientNet-B0. In Proceedings of 1st International

Conference on Computer Science and Artificial Intelligence (ICCSAI). IEEE, Jakarta, Indonesia.

[14] Irzan F. Nurahmadan, Roihan M. Arjuna, Herlambang D. Prasetyo, Pandu A. Hogantara, Ika N. Isnainiyah, and Rio Wirawan. 2021. A mobile based waste classification using MobileNets-V1 architecture. In Proceedings of International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS). IEEE, Jakarta, Indonesia.

[15] Mujalin Polchan, Ampai Pukao, Tapana Cheunban, and Somnuek Sinthupuan. 2023. Waste Management System using Waste Classification on Mobile Application. In Proceedings of Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON). IEEE, Phuket, Thailand.

[16] Ashish Pandey, Harshit Jain, Harshit Raj, and Prince Gupta. 2023. Identification and classification of waste using cnn in waste management. In Proceedings of 8th International Conference for Convergence in Technology (I2CT). IEEE, Lonavla, India.

[17] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: machine learning in python. J Mach Learn Res. 12, 85, 2825–2830.

[18] M. Yuvaraju and K. Sheela S. Rani. 2015. Feature extraction of real-time image using Sift algorithm. International Journal of Research in Electrical & Electronics Engineering. 3, 4 (October-December 2015), 1-7.