MJSAT

Malaysian Journal of Science and Advanced Technology

journal homepage: https://mjsat.com.my/

Dry Waste Classification Using Quadratic Support Vector Machine for Intelligent Waste Management System

Ahmad Fahim Naqib Ahmad Faisal*[1], Jabbar Al-Fattah Yahaya[1]

[1] *Dept. Electric, Faculty of Electrical and Electronics Engineering, Universiti Tun Hussein Onn Malaysia. Batu Pahat, Malaysia.*

KEYWORDS

Dry Waste
Quadratic SVM
Cubic SVM
Fine KNN
Weighted KNN

ARTICLE HISTORY

Received 26 July 2021
Received in revised form 23 September 2021
Accepted 26 September 2021
Available online 27 September 2021

ABSTRACT

There has been a serious increment in solid waste in the past decades due to rapid urbanization and industrialization. Therefore, it becomes a big issue and challenges which need to have a great concern, as accumulation of solid waste would result in environmental pollution. Recycling is a method which has been prominent in order to deal with the problems, as it is assumed to be economically and environmentally beneficial. It is important to have a wide number of intelligent waste management system and several methods to overcome this challenge. This journal explores the application of image processing techniques in recyclable variety type of dry waste. An automated vision-based recognition system is modelled on image analysis which involves image acquisition, feature extraction, and classification. In this study, an intelligent waste material classification system is proposed to extract features from each dry waste image. The Quadratic Support Vector Machine, Cubic Support Vector Machine, Fine K-Nearest Neighbor, and Weighted K-Nearest Neighbor were used to classify the waste into different type such as bottle, tin, crumble, and flat waste sample. A Quadratic Support Vector Machine (QSVM) classifier led to promising results with accuracy of training, 89.7%.

1. INTRODUCTION

Solid wastes refer here to all non-liquid wastes. In general, this does not include excreta, although sometimes nappies and the feces of young children may be mixed with solid waste. Solid waste can create significant health problems and a very unpleasant living environment if not disposed of safely and appropriately. If not correctly disposed of, waster may provide breeding sites for insect vectors, pests, snakes, and vermin (rats) that can increase the likelihood of disease transmission. It may also pollute water sources and the environment. Therefore, solid wastes are known as the most critical problems of our time. There is no place in the corner of the earth that currently immune from the municipal solid waste. The statistics show that, since the beginning of last decade, an increasing sharply number of solid waste production is correlated with the gross domestic product (GDP). High GDP tends to produce large quantity of solid waste, From an updated report shown by the world bank has reported that there are about 4 billion tons of waste generated every year globally which urban area is one of the main contributors to the huge numbers and the waste is estimated to be up until 70% by 2025. In the next 25 years, number of wastes accumulated will be rapidly increased in underdeveloped nations due to accelerated pace of urbanization and industrialization [1]-[2].

The population growth rate has continued of 2.4% per year since 1994 based on Department of Statistic Malaysia in 2012. The higher the growing number of people with higher consumption rates, the higher the amount of waste generation. There is a correlation between the income rate and urbanization. As the disposable income and living standard increase, the consumption of products and services, as well as the quantity of waste produces rise correspondingly. With an incrementing number of industries in the urban area, thus, solid waste management becomes a critical concern and challenging for municipal authorities worldwide since the drastically amount of waste generated. Solid waste management confronts more complex problems in the developing nations due to limited door-to-door collection, inefficient treatment, and inadequate disposal facilities [3]. Majority of the solid waste is comprised

*Corresponding author:
E-mail address: Ahmad Fahim Naqib Ahmad Faisal < fahimnaqib97@gmail.com>.

of waste that mainly found in public which consists of paper, plastics, and glass waste material.

This research work embarks on the objectives of finding the suitable image processing method for feature extraction of bottle, tin, crumble, and flat waste sample. Furthermore, to investigate the suitable features used for classification to increase performance. Following objective is to find optimum classification method between quadratic support vector machine, cubic support vector machine, fine K nearest neighbor and weighted K nearest neighbor.

After the introduction, the details of image processing discussed in Section 2. Then Section 3 will focus on the result of training and testing system. Finally, Section 4 is the conclusion of this project.

2. IMAGE PROCESSING SYSTEM DESIGN

2.1 Flowchart of the system

Figure 1 shows the process of the system divided by two modules, training, and testing. The proposed dry waste classification system by using vision inspection method consists of several module: image acquisition, image processing, feature extraction, classification and finally determined the decision.

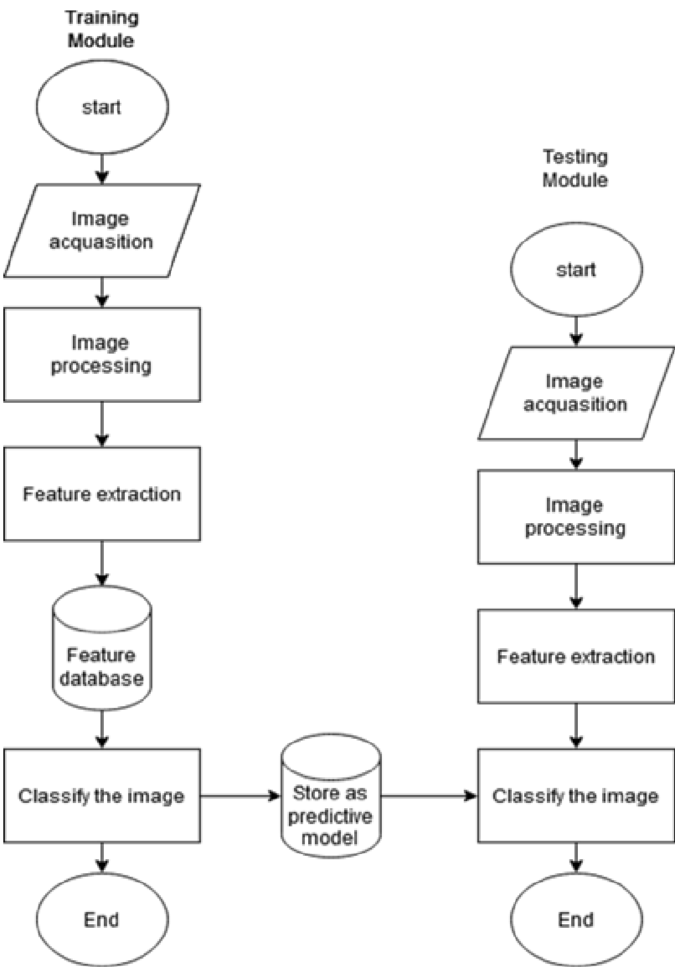

Fig. 1. Conceptual framework for training and testing stage proposed by the system.

2.2 Image Acquisition

The first step for this project is image acquisition. For the image acquisition, the image obtain are manually captured. The images need to capture with dark background. All images are manually captured and entered the simulation.

2.3 Feature Extraction

2.3.1. Grey level co-occurrence matrix

Another statistical technique that contemplates with spatial relationship of pixels is the grey level co-occurrence matrix (GLCM) [4]. GLCM works by calculating how often a pixel with the intensity value i occurs in a specific relationship to a pixel with the value j. By default, the spatial relationship is defined as the pixel of interest and the pixel to its immediate right (horizontally adjacent) [5]. Each element *(i, j)* in the resultant is simply the sum of the number times the pixel with value i occurred in the specified spatial relationship to a pixel with value j in the input image. The number of grey levels is very significant as it can be used to determine the size of GLCM. Therefore, parameters such as number of grey levels and the scaling intensity values need to be controlled. The grey level co-occurrence matrix can reveal certain properties about the spatial distribution of the grey levels in the texture image. For example, if most of the entries in the GLCM are concentrated along the diagonal, the texture is coarse with respect to the specified offset. In MATLAB R2020b software, GLCM have been programmed to derive several statistical measures such as contrast, correlation, energy, and homogeneity [6].

For contrast, it calculates the variation of localization point in the grey level co-occurrence matrix, as shown in Eq. (1). For correlation, it measures the occurrence probability of joint in specified pairs of pixels, as shown in Eq. (2). For energy, also widely known as uniformity or the angular second moment, providing addition of squared pixels in the GLCM, as shown in Eq. (3). Lastly, homogeneity is to test the proximity of element distribution in the GLCM to the GLCM diagonal, as shown in Eq. (4). Consider a matrix denote as B (2,4), which is calculated from matrix A (2,4) [7]. Initially, GLCM implement 3 x 3 neighborhood. For example, at value 9 in matrix A, the maximum value in the neighborhood is subtract with minimum value in the neighborhood.

$$\sum_{i,j} |i-j|^2 p(i,j) \tag{1}$$

$$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \tag{2}$$

$$\sum_{i,j} p(i,j)^2 \tag{3}$$

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{4}$$

2.3.2. Ratio of grey level

Grey level is fundamental in study of image processing. The grey level or grey value indicates the brightness of a pixel

[8]. The maximum grey value depends on the depth of an image. For example, 8-bit-deep image contain levels up to 255, which they can take any value in the range. However, binary image can only take either value 0 or 255 [9]. Table 1 shows the summary of grey level. The program has been setup to calculate ratio of grey level (L) and ratio of grey level (H). Here, ratio of grey level (H) is denoted as $40 < x \leq 110$, and ratio of grey level (L) is denote as $181 \leq x \leq 255$. Table 1 Grey level and its respective colour.

Table 1. Grey level and its respective colour.

Grey Level	Colour
0	Black
0 < x < 255	Grey
255	White

2.3.3. *Entropy*

In image processing, entropy is a statistical measure of randomness that can be used to characterize the texture of the input image [10]. The higher the value of entropy will result as the more detailed information of an image. Entropy is a measure image information content, which is interpreted as the average uncertainty of information source [11]. A vector with relatively “low” entropy is a vector with relatively low information content, such as it might be [0 1 0 1 1 1 0]. A vector with relatively “high” entropy is a vector with relatively high information content such as it might be [0 242 124 222 149 13]. Hence, it is very important to have higher entropy to have precise segmented image after image post-processing method so that it can classify accordingly to its own type of groups. Entropy algorithm is shown in Eq. (5):

$$\mu_i = \sum_{i,j=0}^{N-1} p_{i,j}(-\ln p_{i,j}) \qquad (5)$$

2.3.4. *Standard Deviation*

It is the most common method used to calculate variability or diversity in statistic. In image processing, it shows how much discrepancy, or “dispersion” exists from the mean value. Total standard deviation used as a more accurate expression of the statistical distribution of each class. Liu et al. [12] stated that a standard deviation filter measures the statistical distribution of each class and locates this value to the center pixel in the output map. It also can be implemented in calculating corner sharpening as it has capability in measuring variability, and due to grey level is varied at the corner of image by larger value [13]. Standard deviation filter can be very advantages for radar images. Formally, interpretation of radar images quite hard because of back scatter (return of the pulse sent by the radar). This is due to a lot of “noise”. Therefore, some pattern can be recognizable when using standard deviation filter. Mathematically standard deviation can be represent as shown in Equation (6):

$$f'(x,y) = \sqrt{\frac{1}{mn-1}\sum(g(r,c) - \frac{1}{mn-1}\sum_{(r,c)\in W} g(r,c))^2} \qquad (6)$$

2.4 *Classification*

The classification process Support Vector Machine (SVM) in the experiment is carried out by using Classifier Learner Application in MATLAB. There are about six classification models under category SVM, which are linear SVM, quadratic SVM, cubic SVM, fine gaussian SVM, medium gaussian SVM and coarse gaussian SVM. Experiment using the 5-fold cross-validation is applied to evaluate the prediction accuracy of the model, which optimum number of k-fold applied is 5. If the number of k-fold increase, it will result in lower accuracy. In contrast, the accuracy will increase but it does not protect from the overfitting data [14]. Cost matrix applied in this experiment are using default setting for misclassification costs.

The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. The sample k is a user-defined constant at the classification stage, and a non-labelled vector (a query or test point) is classified by allocating the tag that is the one most common in the k training datasets closest to that query point [15]. The Euclidian distance is a popular distance metric for sustained variables, for which a different metric may be used, such as a differentiation metric (or Hamming distance) for discrete variables such as text classification. In the context of micro-array data on gene expression, K-Nearest Neighbor (K-NN) for example was used as metric for coefficients such as Pearson and Spearman [16]. Often, if a distance metric is learned with special algorithms, such as Large Margin Nearest Neighbor and Neighborhood components, the grading accuracy of K-NN can significantly be enhanced [17].

When the class allocation is skewed, there is a drawback in the fundamental "majority voting" classification [18]. That is, examples of a more regular class tend to dominate the new example prediction because they are common among k neighbors, because they are numerous [19]. One way to overcome this problem is by assessing the distance between the test point and each of its closest neighbors. Each of the nearest k points is multiplied by a weight ($w = 35$) proportional to the opposite distance of that point to the test point by a class (or value, in regression problems). Cost matrix applied in this experiment are using default setting for misclassification costs [20].

3. Results And Discussion

The experiments were carried out in two stages: training phase and testing phase. In training phase, the segmentation images from the image processing step will be used for feature extraction phase where features from the dry waste are extracted. In testing phase, the dataset is not labelled so that the classification will be carried out based on training database. Table 2 shows the total of dry waste samples for training and testing such as bottle, crumble, flat and tin.

Table 2. Dataset of dry waste images

Types of dry wastes	Number of training samples	Number of testing samples
Bottle	24	3
Crumble	34	3
Flat	44	2
Tin	39	2
Total	**141**	**10**

3.1 Training Results

Since each classifier performs differently for different data. Table 3 shows the training results obtained by 4 classifiers used in this experiment. Table 4 shows the average classification accuracy for each classifier. From the experiment that has been conducted, all the classifiers were training to become predictive models for validation purpose. Table 4 shows the accuracy of the training module of the system after 10 times the system being test which produce different result for Quadratic SVM, Cubic SVM, Fine KNN and Weighted KNN then the Table 4 indicates the average of accuracy result after 10 times the system test for training module.

Table 3. Training classification accuracy results.

No. of training	Quadratic SVM	Cubic SVM	Fine KNN	Weighted KNN
1	90.8	91.5	86.5	84.4
2	90.8	88.7	88.7	82.3
3	90.1	88.7	85.1	80.9
4	87.9	87.9	87.2	83.7
5	90.1	90.1	85.8	84.4
6	87.9	85.8	85.8	82.3
7	90.1	87.9	87.2	85.8
8	87.9	87.2	85.1	80.9
9	90.8	87.9	86.5	83.0
10	90.8	90.1	85.1	83.0

Table 4. Average training accuracy.

Classifier	Average training accuracy (%)
Quadratic SVM	89.7
Cubic SVM	88.6
Fine KNN	86.3
Weighted KNN	83.1

To summarize the previous works on classifying various type of dry waste using many techniques as feature extractor and classifier. For example, metal, paper, glass plastic, bottle, tin can, cardboard, trash, e-waste etc. The suggested system is benchmarked with the previous works as shown in Table 5. As compared to previous works, most of the previous works are evaluated with database less than the proposed system. As for comparisons in terms of classification accuracy, the proposed system yields about 89.7% which is good but still lower in number of dry waste samples compared to some of previous works. However, this still proved that the suggested system that implements features and Q-SVM classifier manage to identify 4 types of groups for dry waste quite effectively.

Table 5. Previous work on dry waste recognition system.

References	No of class (No of samples)	Classifiers	Accuracy (%)
(Saeed, Mustaffa et al., 2019) [11]	3(60)	1. Fine Tree 2. Bagged Trees 3. Linear Discriminant 4. Quad Discriminant 5. Fine KNN 6. Weighted KNN 7. L-SVM 8. Q-SVM 9. FG-SVM	1. 85.7 2. 88.1 3. 73.8 4. 92.9 5. 78.6 6. 81 7. 92.9 8. 90.5 9. 83.3
(Liu et al., 2018) [12]	5(4300)	SVM	83.38
(Rajamanikam et al., 2019) [13]	2(134)	ANN	91.91
(Yang et al., 2016) [14]	6(2400)	1. SVM 2. CNN	1. 63 2. 22
(Costa et al., 2018) [15]	4(1600)	1. SVM 2. VGG16 3. KNN 4. Random Forest 5. AlexNet	1. 90 2. 93 3. 88 4. 85 5. 91
**The proposed vision-based system	4(141)	1. Q-SVM 2. C-SVM 3. Fine KNN 4. Weighted KNN	1. 89.7 2. 88.6 3. 86.3 4. 83.1

3.2 Testing Results

From previous training experiment, predictive models are saved up based on the best output produce by each of the classifier at specific number of training session. Ten images been running in testing process, all the images are correctly classified. Figures 2, 3, 4 and 5 are the examples of each type of waste that already classify through testing process.

Fig. 2. Testing result for image of bottle.

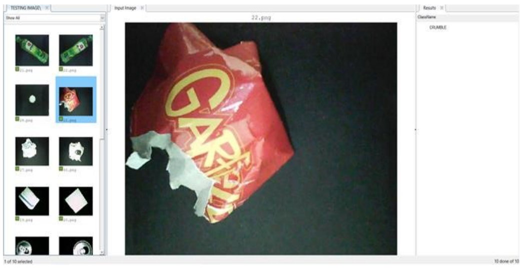

Fig. 3. Testing result for image of crumble.

Fig. 4. Testing result for image of flat garbage.

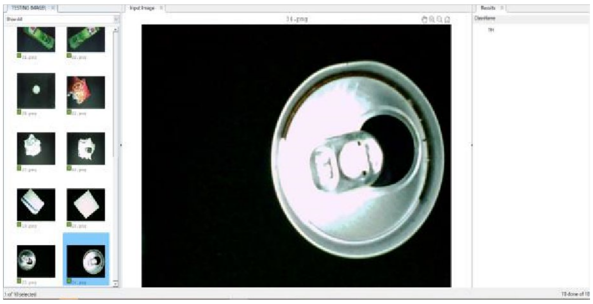

Fig. 5. Testing result for image of tin.

4. CONCLUSION

In this research study, types of dry waste can be determined by the image processing. Different types of dry waste such as bottle, crumble, flat surface, and tin success to recognize. Image processing approaches such as edge detection, image dilating, image filling and image smoothing were able to be applied in order, to differentiate all those types of dry wastes into 4 different classes.

White pixel plot also played important role in separated all those data in the beginning so that a clear picture of each class can be differentiate earlier. Image processing approaches such as edge detection, image dilating, image filling and image smoothing were able to be applied in order, to differentiate all those types of dry wastes into 4 different classes.

All the samples are classified by using Quadratic SVM, Cubic SVM, Fine KNN and Weighted KNN. Average classification training accuracy for each classifier are 89.7%, 88.6%, 86.3%, 83.1% respectively then the best predictive training models is quadratic support vector machine (QSVM) with 89.7% accuracy. Since QSVM is the best classifier for this system, the testing phase of the system will be use QSVM to achieve higher accuracy prediction for classification of dry waste for bottle, crumble, flat and tin. In future the accuracy of the system still can be improve after adding more the number of dry wastes.

ACKNOWLEDGEMENT

The authors express their sincere gratitude to the Faculty of Electrical and Electronics Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), for providing a conducive research platform.

REFERENCES

[1] Sarkodie, S.A., Owusu, P.A. Impact of COVID-19 pandemic on waste management. Environ Dev Sustain 23, 7951–7960 (2021). https://doi.org/10.1007/s10668-020-00956-y

[2] Adedeji, O., & Wang, Z. (2019, August 14). Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network.

[3] Gundupalli, S. P., Hait, S., & Thakur, A. (2017). A review on automated sorting of source-separated municipal solid waste for recycling. Waste Manag, 60, 56-74. doi:10.1016/j.wasman.2016.09.015

[4] Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., & Mahfood, B. (2020). Artificial intelligence applications in solid waste management: A systematic research review. Waste Management, 109, 231–246. https://doi.org/10.1016/j.wasman.2020.04.

[5] Gundupalli Paulraj, S., Hait, S., & Thakur, A. (2016). Automated Municipal Solid Waste Sorting for Recycling Using a Mobile Manipulator. Volume 5A: 40th Mechanisms and Robotics Conference. Published. https://doi.org/10.1115/detc2016-59842.

[6] Vicente, F. J. M., Álvarez-Sánchez, J. R., López, D. F. L. P., Moreo, T. J., & Adeli, H. (2019). From Bioinspired Systems and Biomedical Applications to Machine Learning: 8th International Work-Conference on the Interplay Between Natural and . . . II (Lecture Notes in Computer Science, 11487) (1st ed. 2019 ed.). Springer.

[7] Bobulski J., Kubanek M. (2019) Waste Classification System Using Image Processing and Convolutional Neural Networks. In: Rojas I., Joya G., Catala A. (eds) Advances in Computational Intelligence. IWANN 2019. Lecture Notes in Computer Science, vol 11507. Springer, Cham. https://doi.org/10.1007/978-3-030-20518-8_30.

[8] Jayson, M., Hiremath, S., & H.R., L. (2018). SmartBin-Automatic waste segregation and collection. 2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC). Published. https://doi.org/10.1109/icaecc.2018.8479531.

[9] Gundupalli, S. P., Hait, S., & Thakur, A. (2017). Multi-material classification of dry recyclables from municipal solid waste based on thermal imaging. Waste Management, 70, 13–21. https://doi.org/10.1016/j.wasman.2017.09.019

[10] N. S. Gupta, V. Deepthi, M. Kunnath, P. S. Rejeth, T. S. Badsha and B. C. Nikhil, (2018) "Automatic Waste Segregation," Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, pp. 1688-1692, doi: 10.1109/ICCONS.2018.8663148.

[11] Saeed, Mustafa, Sheikh, Jumani, & Mirjat. (2019). Ensemble Bagged Tree Based Classification for Reducing Non-Technical Losses in Multan Electric Power Company of Pakistan. Electronics, 8(8), 860. https://doi.org/10.3390/electronics8080860

[12] Liu, Y., Wen, K., Gao, Q., Gao, X., & Nie, F. (2018). SVM based multi-label learning with missing labels for image annotation. Pattern Recognition, 78, 307–317. https://doi.org/10.1016/j.patcog.2018.01.022

[13] Baskaran, D., Sinharoy, A., Pakshirajan, K., & Rajamanickam, R. (2020). Gas-phase trichloroethylene removal by Rhodococcus opacus using an airlift bioreactor and its modeling by artificial neural network. Chemosphere, 247, 125806. https://doi.org/10.1016/j.chemosphere .2019. 125806

[14] Yang, M., & Thung, G. (2016). Classification of trash for recyclability status. CS229 Project Report , 2016.

[15] Costa, B. S., Bernardes, A. C. S., Pereira, J. V. A., Zampa, V. H., Pereira, V. A., Matos, G. F., Soares, E. A., Soares, C. L., & Silva, A. F. (2018). Artificial Intelligence in Automated Sorting in Trash Recycling. Anais Do XV Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2018). Published. https://doi.org/10.5753/eniac.2018.4416

[16] Bircanoglu, C., Atay, M., Beser, F., Genc, O., & Kizrak, M. A. (2018). RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks. 2018 Innovations in Intelligent Systems and Applications (INISTA). Published. https://doi.org/10.1109/inista.2018.8466276

[17] Meng, X., Tan, X., Wang, Y., Wen, Z., Tao, Y., & Qian, Y. (2019). Investigation on decision-making mechanism of residents' household solid waste classification and recycling behaviors. Resources, Conservation and Recycling, 140, 224–234. https://doi.org/10.1016/j.resconrec.2018.09.021

[18] Qin, L. W., Ahmad, M., Ali, I., Mumtaz, R., Zaidi, S. M. H., Alshamrani, S. S., Raza, M. A., & Tahir, M. (2021). Precision Measurement for Industry 4.0 Standards towards Solid Waste Classification through Enhanced Imaging Sensors and Deep Learning Model. Wireless Communications and Mobile Computing, 2021, 1–10. https://doi.org/10.1155/2021/9963999

[19] You, H., Ma, Z., Tang, Y., Wang, Y., Yan, J., Ni, M., Cen, K., & Huang, Q. (2017). Comparison of ANN (MLP), ANFIS, SVM, and RF models for the online classification of heating value of burning municipal solid waste in circulating fluidized bed incinerators. Waste Management, 68, 186–197. https://doi.org/10.1016/j.wasman.2017.03.044

[20] 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. Advanced Industrial and Engineering Polymer Research, 2(2), 77–81. https://doi.org/10.1016 /j.aiepr.2019.04.001