Solid waste classification using pyramid scene parsing network segmentation and combined features

Khadijah, Sukmawati Nur Endah, Retno Kusumaningrum, Rismiyati, Priyo Sidik Sasongko, Iffa Zainan Nisa
Department of Informatics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

Article Info

ABSTRACT

Solid waste problem become a serious issue for the countries around the world since the amount of generated solid waste increase annually. As an effort to reduce and reuse of solid waste, a classification of solid waste image is needed to support automatic waste sorting. In the image classification task, image segmentation and feature extraction play important roles. This research applies recent deep leaning-based segmentation, namely pyramid scene parsing network (PSPNet). We also use various combination of image feature extraction (color, texture, and shape) to search for the best combination of features. As a comparison, we also perform experiment without using segmentation to see the effect of PSPNet. Then, support vector machine (SVM) is applied in the end as classification algorithm. Based on the result of experiment, it can be concluded that generally applying segmentation provide better source for feature extraction, especially in color and shape feature, hence increase the accuracy of classifier. It is also observed that the most important feature in this problem is color feature. However, the accuracy of classifier increase if additional features are introduced. The highest accuracy of 76.49% is achieved when PSPNet segmentation is applied and all combination of features are used.

Corresponding Author:

Khadijah
Department of Informatics
Faculty of Science and Mathematics
Universitas Diponegoro
Prof. H. Soedarto St., SH. Tembalang Semarang 50275, Indonesia
Email: khadijah@live.undip.ac.id

1. INTRODUCTION

Solid waste problem become a serious issue for the countries around the world since the amount of generated solid waste increase annually. In 2016 the total generation of solid waste by the world's cities was up to 2.01 billion tonnes. It was equal to 0.74 kilogram of solid waste generated by a person in a day. This number is estimated to increase by 70% or up to 3.40 billion tonnes of solid waste in 2050. Population growth and urbanization are the most siginificant factors that trigger the increase in the amount of waste. Poor management of waste may create serious problem related to health, safety, and environment. Therefore, proper waste management strategy is needed to minimize such negative impacts [1].

One of the effort to reduce the number of solid waste is by improving the waste reusability. Waste sorting plays significant role to support the waste reusability [2]. Since the number of waste is great and the awarness of people in waste sorting is still low, an automatic waste sorting is needed. The starting point to

produce an automatic waste sorting is by building a classification model that may recognize the type of waste image.

　　　　Some previous studies have been investigated the use of machine learning algorithm to classify or recognize waste image. Mustaffa *et al.* [3] and Torres-gracia *et al.* [4] classified waste image into three classes using conventional machine learning algorithm and were able to achieve good accuracies, but their experiment used only 20 samples in each class. Therefore, the generalization of the resulting model for classifying the variety of real waste image could not be assured. Adedeji and Wang [5] and Costa *et al.* [6] classified waste image using deep learning model and more number of samples, but they used the capture of waste image directly without any segmentation. On the other hand, segmentation is one of the most important part in the image preprocessing. Poor segmentation result may degrade the performance of the subsequent processes, such as feature extraction and classification [7].

　　　　Segmentation is the process of partitioning an image into different several disjoint subset [8], for example partitioning an image into background and foreground. Segmentation can also be used to extract region of interest of an image [7]. State of the art of segmentation methods are kind of deep learning algorithm with special architecture, such as encoder-decoder which have better performance than the conventional one (such as thresholding method) [9]. Pyramid scene parsing netwrok (PSPNet) is a kind of deep learning network that can be used for semantic image segmentation. PSPNet successfully outperformes other deep learning based segmentation in some large benchmark dataset, such as fully convolutional network (FCN), DeepLab, deep parsing network (DPN) and Laplacian pyramid reconstruction and refinement (LRR). PSPNet is able to achieve better segmentation result because it considers global context of the image and uses pyramid pooling module to obtain different region based context of an image [10].

　　　　In addition, feature extraction of the image must be determined properly to achieve good classification result [11], [12]. Feature extraction is aimed to extract relevant subset of features from an image and to reduce the large dimension of image to the lower dimensional set of image features [13]. Color, texture, and shape are the most visual features extracted from an image. Colormoments is one of the simplest color feature compared to the other, such as color histogram, color coherence vector, and color correlogram [14]. Color moment is also proven to be effective and efficient for extracting color features of an image [15]. In addition texture feature is also important to extract the relationship from neighboring pixel. Gray level co-occurence matrix (GLCM) is one of the popular texture-based feature extraction that has been successfully applied in many computer vision problem [16]-[18]. The other important image feature to describe the object of an image is shape feature. Some morpohlogical features, such as area, perimeter, major and minor axis, centroid-x, centroid-y, roundness, rectangularity, eccentricity and elongation, can be used as shape descriptor [19]. In addition, Hu moment is also important to extract shape features. Hu moment is region-based method that uses second and third order central moments and constructs seven invariant moments whose values are not affected when the image is translated, rotated, or scaled [20].

　　　　In this research, we propose the application of PSPNet as segmentation to provide good source for feature extraction. We also use various combination of image feature extraction (color, texture, and shape) to search fo the best combination of features for solid waste image classification. As a comparison, we also perform experiment without using segmentation to see the effect of PSPNet. Then, support vector machine (SVM) is applied at the end as a classifier. SVM is a binary classification algorithm proposed by Cortes and Vapnik [21] which works by finding the optimal hyperplane to maximize the separation between binary class data. SVM has been successfully applied in various classification problem and proven to better than other popular classification algorithm, such as artificial neural network (ANN) [22], [23], Naïve Bayes classifier dan random forest [24].

2.　RESEARCH METHODOLOGY

　　　　Figure 1 shows the stages of process in this research. First, the image dataset is segmented by using PSPNet segmentation, then the process is continued by feature extraction, classification, and evaluation. As a comparison, to examine the effect of PSPNet, we also perform experiment without using segmentation, hence the process of segmentation in Figure 1 is skipped.
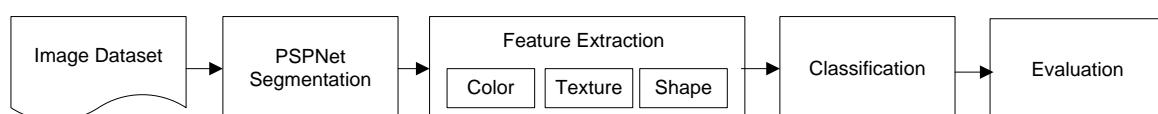

Figure 1. Research methodology

2.1. Dataset

Public trash image dataset from Trashnet are used in this research as source data for conducting experiments. Trashnet dataset was collected by Yang and Thung [25]. This dataset contain 2,527 trash images of 224×224 pixels which is grouped into six classes: glass (501), paper (594), cardboard (403), plastic (482), metal (410), and trash (137). A sample image from each class can be seen in Figure 2 [25].
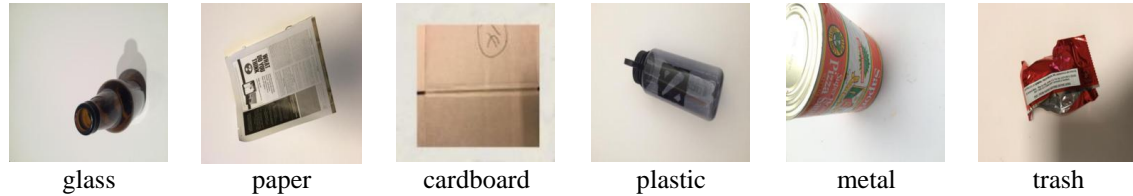

glass   paper   cardboard   plastic   metal   trash

Figure 2. Sample image from each class of Trashnet dataset [25]

2.2. PSPNet for image segmentation

PSPNet is performed to generate segmented binary image, then the bounding box of segmented image are calculated and image is cropped so that only the main object remain. PSPNet is a kind of deep learning network for semantic image segmentation. PSPNet outperformed FCN based segmentation because PSPNet consider global context of the image and uses pyramid pooling module to obtain different region based context of an image. The architecture of PSPNet is shown in Figure 3 [10].

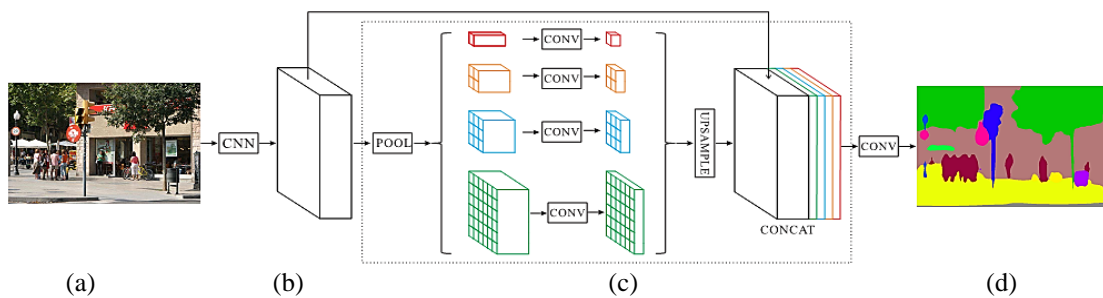(a)   (b)   (c)   (d)

Figure 3. The Architecture of PSPNet [10]: (a) input image, (b) feature map, (c) pyramid pooling module, and (d) final prediction

First, an input image (Figure 3(a)) is fed into convolutional neural networks (CNN) with dilated network strategy which genarate feature map (Figure 3(b)) with size 1/8 of original input image. Then, the feature map is forwarded to the pyramid pooling module (Figure (c)) which generate the concatenated feature map in the end of the module. In the last step (Figure (d)) convolution layer is applied on the concatenated feature map to generate the final prediction of each pixel in an image. There are four operations in the pyramid pooling module as [10], [26]:

− Sub region average pooling

Each feature map is pooled over different sub-region to obtain different context reprsentation in each sub-region. In the first level (red), the global average pooling is performed in each feature map. The result is a single bin output for each fature map. In the second level (orange), third level (blue), and fourth level (green), each feature map is divided into 2×2, 3×3 and 6×6 sub-region, respectively, then each sub-region is pooled by average pooling.

− Convolution

The 1x1 convolution is performed at each level to reduces the size of feature map at each level into 1/N of the original one (black) where N is the level size of pyramid.

− Upsampling

Upsampling is performed by using bilinear interpolation to make each feature map have the equal size as the original one (black).

− Concatenation

The original feature map (black) and all upsampled feature map from the first to fourth level are concatenated and the result is forwarded to convolutional layer for prediction.
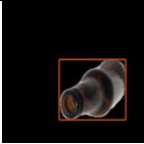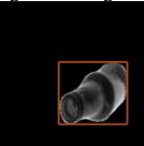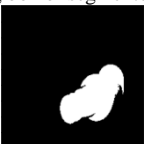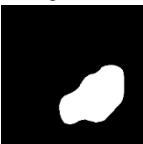
The process of segmentation using PSPNet consists of training and testing. Dataset is divided into 70% of training data, 15% of validation data, and 15% of testing data. Training is performed using some combination of hyperparameter: learning rate (0.001, 0.0001, and 0.00001) and batch (5 and 10) in 50 epoch. After training using such combination of hyperparameters, six models of image segmentation are obtained, then testing data is used to evaluate and select the best model. Dice coeffect (DC) is used to evaluate the results of segmentation as shown in (1) where A and B is the image regions being compared [27].

$$DC = \frac{2|A \cap B|}{|A| + |B|} \tag{1}$$

2.3. Feature extraction

Feature extraction is aimed to extract relevant subset of features from an image [13]. This research uses three kinds of image features, namely color features extracted by using color moments, texture features extracted by using gray level co-occurence matrix (GLCM) and shape features. Experiments are run using one or combination of such features to obtain the best classification result. Table 1 shows the comparison of source image for each feature extraction method. When using PSPNet segmentation, original red, green, and blue (RGB) image is segmented resulting the segmented binary image. Then, for extraction of color and texture features, the image is cropped around the bounding box using OpenCV library, findcontour. When the segmentation is skipped, before shape features are extracted, each image is converted into binary image by using inverse binary thresholding (value of threshold = 128).

Table 1. The comparison of input image for feature extraction between PSPNet segmentation and without segmentation

Feature	Using PSPNet segmentation	Without segmentaion
Color	Cropped bounding box of segmented RGB image	Original RGB image
Texture	Cropped bounding box of segmented grayscale image	Transformation from original RGB to grayscale image
Shape	Segmented binary image	Transformation from original RGB to binary image using inverse binary tresholding

2.3.1. Colormoments

Color feature is visual feature that can be used to discriminate or recognize visual information. If the color distribution of an image is interpreted as a probability distribution, then color moments can be used to characterize the color distribution [28]. Three color moments (mean, standar deviation, and skewness) are extracted for every image channel, therefore there are 9 numerical values extracted for an image in RGB color space. Mean E_i is the average of pixel values as shown in (2), standard deviation σ_i is the variation of pixel values as shown in (3), and skewness s_i is the degree of asymmetry in the color distribution in an image channel as shown in (4). N is total number of pixel in each channel and p_{ij} is the j-th pixel value in channel i [15].

$$E_i = \frac{1}{N} \sum_{j=1}^{N} p_{ij} \tag{2}$$

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^{N} (p_{ij} - E_i)^2 \right)} \tag{3}$$

$$s_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^{N} (p_{ij} - E_i)^3 \right)} \tag{4}$$

2.3.2. Gray level co-occurence matrix (GLCM)

GLCM is a method for extracting texture features of an image. First, co-occurence matrix P is created. P is a square matrix whose size is equal to the number of gray intensity value of an image. Each element p_{ij} in the matrix is the number of occurence (frequency) of two neigboring pixel in specific orientation where the gray intensity value of the first pixel is equal to i and gray intensity value of the second pixel is equal to j [29]. Neighboring pixel can be selected based on specified spatial orientation. For example when the orientation is 0^0, then the neighbor of a pixel is a pixel that is on the right side. The resulting GLCM matrix can be obtained by making P as symmetrical matrix (adding matrix P with its transpose) and then normalizing the value of each element into [0, 1]. Some metrics can be calculated based on the resulting GLCM matrix, they are contrast, angular second moment (ASM), energy, homogenity, correlation, and dissimilarity. The detail formula for each metric can be referred at [30]. In this research, we construct GLCM matrix in various spatial orientation $(0^0, 45^0, 90^0,$ and $135^0)$.

2.3.3. Shape

Shape is also prominet feature to discirminate an image to another. This research extracts shape descriptors of an image from morphological features and Hu Invariant Moment. Some morpohlogical features extracted are area, perimeter, major and minor axis, centroid-x, centroid-y, roundness, rectangularity, eccentricity, elongation, dispersion I, dispersion IR, convexity, and solidity [19]. The illustration of such morphological shape features can be seen in Figure 4.

In addition, Hu moment is also performed to extract shape descriptor of an image. Hu moment is region-based method that uses second and third order central moments and constructs 7 invariant moments. The value of invariant moment features are not affected when the image is translated, rotated, or scaled. The detail oh Hu moment can be referred in [20].
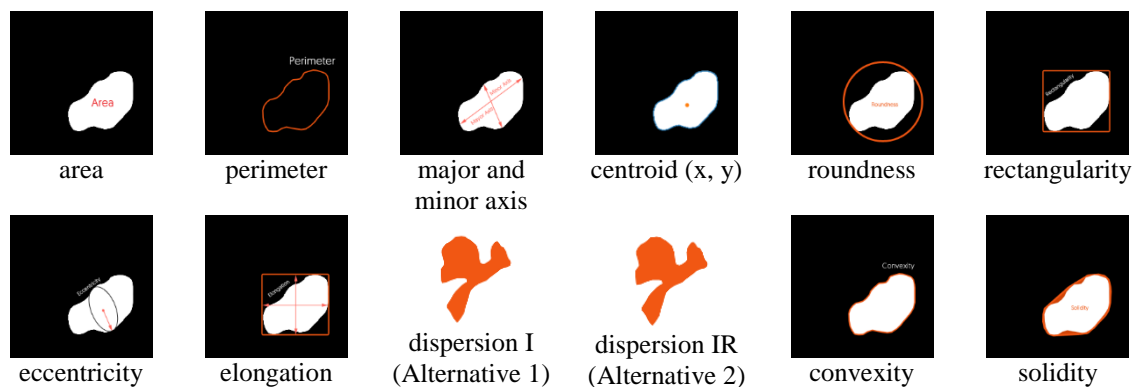

| area | perimeter | major and minor axis | centroid (x, y) | roundness | rectangularity |

| eccentricity | elongation | dispersion I (Alternative 1) | dispersion IR (Alternative 2) | convexity | solidity |

Figure 4. The illustration of shape descriptors from morphological features

2.4. Classification

Classification consists of training and testing. Training is used to build the classifier model, while testing is used to evelute the performance of the model as illustrated in Figure 5. First the dataset is splitted by using 10-fold cross validation. This research applies SVM as classification algorithm.
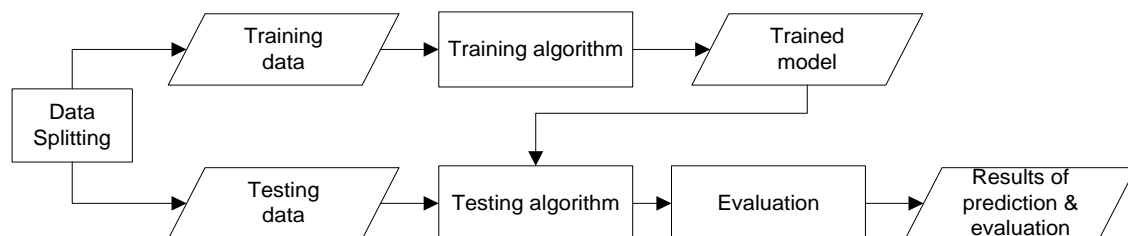

Figure 5. Traning and testing in classification

SVM training algorithm works by finding the optimal hyperplane that maximize the separation between binary class data. The closest training data to the optimal hyperplane that defined the optimal margin

are called support vectors [21]. When the data are non-linearly separable, non-linear mapping $\phi(\mathbf{x}_i)$ is applied to transform the original data into higher dimension [31]. Let the $(\mathbf{x}_i, y_i)_{i=1}^N$ where $\mathbf{x}_i \in \mathbf{R}^d$ is input training data, y_i is targeted data and N is the number of training data, SVM find the solution by solving the following optimization problem as show in (5) where $\mathbf{w}$ is weight vector and C is error penalty. Such optimization problem can be solved using Lagrangian formulation. The training data $\mathbf{x}_i$ is normalized into [0, 1] before they are inputted to the SVM and y_i is set into -1 or 1 [32].

$$\min_{\mathbf{w},b,\xi} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^l \xi_i \tag{5}$$

$$\text{subject to} \quad y_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq 1 - \xi_i,$$

$$\xi_i \geq 0, i = 1, \dots, l$$

In order to reduce the computational cost when working with nonlinear data, kernel tricks can be used to substitute the dot product between transform data tuples as (6). Some popular kernel function can be used, such as polynomial and radial basis function (RBF) as shown in (7) and (8), respectively [33].

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i).\phi(\mathbf{x}_j) \tag{6}$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \left|\left|\mathbf{x}_i - \mathbf{x}_j\right|\right|^2\right) \tag{7}$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma(\mathbf{x}_i^T\mathbf{x}_j) + r)^d \tag{8}$$

Once the optimization problem solved, the optimal hyperplane and the support vectors are obtained. Then, the output $out(\mathbf{x}_t)$ of a new test sample $\mathbf{x}_t$ can be determined by using (9) where $\mathbf{x}_i$ are support vector, y_i is class label of i-th support vector, l is the number of support verctors, α_i is Lagrange multipliers, and b is bias [32]. This research applies the one-versus-rest strategy to handle the multiclass classification problem, because Trashnet dataset consist of 6 classes.

$$out(\mathbf{x}_t) = \text{sgn}\left(\sum_{i=1}^l y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}_t) + b\right) \tag{9}$$

2.5. Evaluation

Evaluation is performed to evaluate the resulting classification model. In this research, evaluation of classification model is measured in term of accuracy. Accuracy shows the ratio between the correctly predicted data and the total number of data [34].

3. RESULTS AND ANALYSIS

This research is performed in two main scenario. The first scenario is performed by using PSPNet segmentation, while the second scenario skip the process of segmentation. In each scenario, single or combination feature extraction of color (colormoments), texture (GLCM) and shape (morphological features and Hu Invariant Moments) are experimented to search for the best image features that well describe the trash image in order to reach the best classification results. GLCM feature extarction method is performed in various spatial orientation: 0^0, 45^0, 90^0, and 135^0. Then in the classification, SVM training algorithm is performed using some combination of parameters, namely kernel function (RBF and polynomial) and error penalty C (1 or 100). Therefore, for each feature extraction in a scenario, classification with SVM is performed four times using different combination of kernel function and error penalty C. In the last section the results of the first scenario and the second scenario are compared.

3.1. The first scenario

In this scenario segmentation is performed in the first stage by using PSPNet. In order to obtain the best model of PSPNet, this reseach try some combination of hyperparameter: learning rate (0.001, 0.0001, and 0.00001) and batch (5 and 10). An experiment for each combination of hyperparameter is performed in 50 epoch.

Based on Figure 6(a), it is shown that the learning rate of 0.0001 gives the best results than other values. It can be explained that when the learning rate is too small, the progress of network learning is very slow, then the result is lower. Conversely, when the learning rate is too high the progress of network learning

may diverge, then the network is failed to achieve the best result. Figure 6(b) shows that the batch value of 5 is able to reach better performance than the batch value of 10. It can be explained that in this case the stochastic nature of using lower number of mini batch may lead to find the optimum solution. Therefore, the segmentation in the rest of experiment are performed by using the best segmentation model trained by those combination of parameter. The results of segmentation using PSPNet for sample images in Figure 2 can be seen in Figure 7.
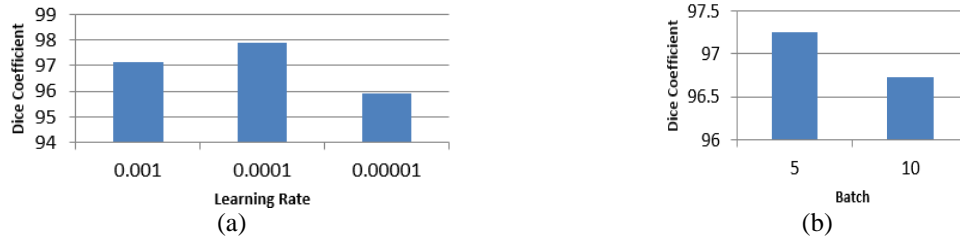

(a)

(b)

Figure 6. Experiment results of PSPNet segmentation on testing dataset: (a) the average result of all variation of batch values in each learning rate and (b) the average results of all variation of learning rate values in each batch

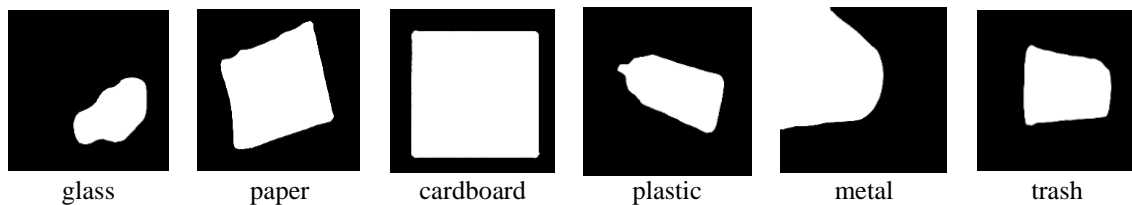glass   paper   cardboard   plastic   metal   trash

Figure 7. The sample results of segmentation using PSPNet

The result of experiment in this scenario for various features with the best accuracy in each SVM kernel (RBF and polynomial) can be seen in Figure 8. It is shown that RBF kernel is better than polynomial kernel in most of experiments, but when more combination of features are used, the polynomial kernel are better than RBF kernel. The highest accuracy of 76.49% in this scenario is achieved by polynomial kernel with $C = 1$ when using combination features of color, GLCM 135 and shape. While the highest accuracy of RBF kernel in this scenario is 74.55% when using $C = 100$ and the same combination features of color, GLCM 135 and shape. Therefore, it can be concluded that when segmentation is used, the performance of classification increase as the more combination of features are used. The use of more combination of features give the more representative feature sets of an image, therefore the accuracy of classification increase. However, the most important feature is the color feature. When the color feature is removed, the accuracy of classifier decrease.

3.2. The second scenario

The second scenario is performed without segmentation in the preprocessing. The result of experiment in this scenario for various features with the best accuracy in each SVM kernel (RBF and polynomial) can be seen in Figure 9. It is shown that RBF kernel is better than polynomial kernel in most of experiments. However, the highest accuracy of 74.83% in this scenario is achieved by polynomial kernel with $C = 100$ when using combination features of color and GLCM 90. While the best result of RBF kernel in this scenario is 74.55% when using $C = 100$ and combination features of color and GLCM 135. Therefore, it can be concluded that when segmentation is not used, the best combination of feature that well describe the trash image is combination of color and GLCM. When, the shape features are added, the performance of classification decrease. To extract shape features in this scenario, a conventional thresholding operation is applied to transform a RGB image into binary image, thefore the resulting binary image is not good enough as source for extracting shape features.

3.3. Comparison of the first scenario and the second scenario

Figure 10 shows the comparison between the first scenario and the second scenario. Based on Figure 10, it is shown that the first scenario (using PSPNet segmentation) is better than the second scenario (without segmentation) in most of experiment. The second scenario outperforms the first scenario only in 5 from 19 experiments. Therefore, it can be concluded that generally applying PSPNet segmentation provide better source for feature extraction, especially in color and shape feature, hence increase the performance of classification.
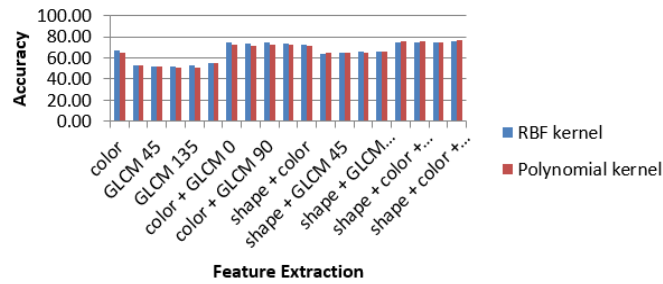
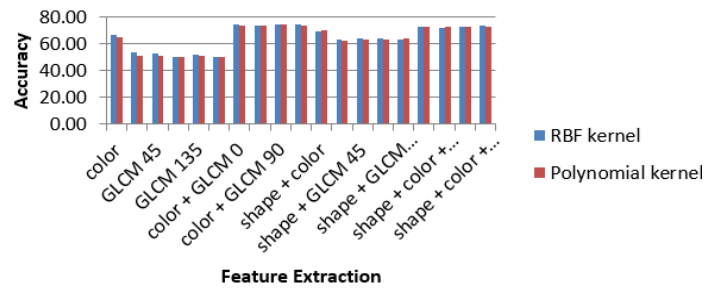Figure 8. Result of experiments in the first scenario

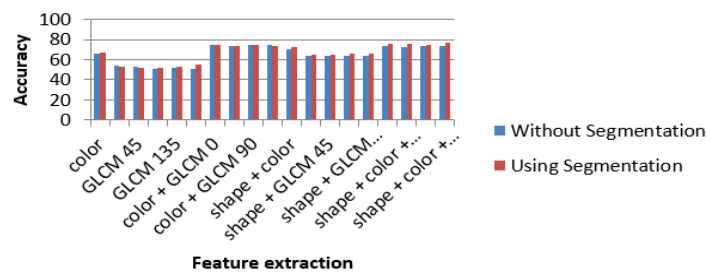Figure 9. Result of experiments in the second scenario

Figure 10. Comparison between the first scenario (using segmentation) and the second scenario (without segmentation)

It is also observed that the most important feature in this problem is color feature. When using single feature, color feature provide the highest result compared to GLCM (texture) and shape feature, both in the first and the second scenario. However, the accuracy increase if additional features are introduced. In the first scenario better results are achieved when using all combination of features, while in the second scenario better results are achieved when using only color and texture features. Therefore, it can be concluded that when segmentation is applied by using PSPNet, the segmented binary image provide better source for shape feature extraction. Conversely, when the binary image is only obtained by using inverse binary thresholding, the result is not good enough for shape feature extraction. Hence, the accuracy of classification decrease when shape feature is added in the second scenario. From all combination of parameters conducted in this research, the highest accuracy of 76.49% is achieved when using PSPNet segmentation and all combination of features (color, texture, and shape).

The results of this research show that the combination of features are able to increase the performance of the resulting model than when using the individual feature, but they are still not enough to uniquely characterize each class of solid waste image. The more representative additional features are still required to improve the performance of classifier. The tuning of parameter of classification algorithm also need to be explored to obtain better classification results.

4. CONCLUSION

In this research we apply PSPNet as segmentation and combination of image feature extraction (color, texture, and shape) to classify the solid waste image. As a comparison, to see the effect of PSPNet segmentation, we also perform experiment without using segmentation. Based on the result of experiment, it

can be concluded that generally applying segmentation provide better source for feature extraction, especially in color and shape feature, hence increase the accuracy of classification. It is also observed that the most important feature in this problem is color feature, both when the segmentation is applied or not. However, the accuracy of classifier increase if additional features are introduced. When segmenation is not used, better result is achieved when using only color and texture features, while when segmentation is applied the highest accuracy of 76.49% is achieved when using all combination of features.

ACKNOWLEDGEMENTS

REFERENCES

[1] S. Kaza, L. Yao, P. B. Tata, and F. V. Woerden, *What A Waste 2.0 A Global Snapshot of Solid Waste Management to 2050,* Washington: World Bank Publications, 2018.

[2] F. A. Rahman, "Reduce, Reuse, Recycle: Alternatives for Waste Management," NM, United State: Las Cruces, NM: NM State University, Cooperative Extension Service, 2014, pp. 1-4. [Online]. Available: https://aces.nmsu.edu/pubs/_g/G314.pdf

[3] M. R. Mustaffa *et al.,* "Automated recyclable waste classification using multiple shape-based properties and quadratic discriminant," *Int. J. Innovative Technol. and Exploring Eng. (IJITEE),* vol. 8, pp. 270-274, Jun. 2019, [Online]. Available: https://www.ijitee.org/wp-content/uploads/papers/v8i8s/H10450688S19.pdf

[4] A. T. García, O. R. Aragón, O. L. Gandara, F. S. García, and L. E. G. Jiménez, "Intelligent waste separator," *Computacion y Sistemas*, vol. 19, no. 3, pp. 487-500, 2015, doi: 10.13053/CyS-19-3-2254.

[5] O. Adedeji and Z. Wang, "Intelligent waste classification using deep learning convolutional neural network," Procedia Manuf., vol. 35, pp. 607-612, 2019, doi: 10.1016/j.promfg.2019.05.086.

[6] B. S. Costa *at al.,* "Artificial intelligence in automated sorting in trash recycling," in *Conference: XV Encontro Nacional de Inteligência Artificial e Computacional*, 2018, pp. 198-205.

[7] S. Matta, "Review: various image segmentation techniques," *Int. J. Comput. Sci. and Inf. Technol.,* vol. 5, no. 6, pp. 7536-7539, 2014. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid= 82107504F029CD97E611CD026FC5834F?doi=10.1.1.657.6661&rep=rep1&type=pdf

[8] D. Libouga, L. Gwet, M. Otesteanu, I. O. Libouga, L. Bitjoka, and G. D. Popa, "A review on image segmentation techniques and performance measures," *Int. J. Comput. and Inf. Eng.,* vol. 12, no. 12, pp. 1107-1117, 2018, doi: 10.5281/zenodo.2579976.

[9] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: a survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, doi: 10.1109/TPAMI.2021.3059968.

[10] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, and S. G. Limited, "Pyramid scene parsing network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6230-6239, doi: 10.1109/CVPR.2017.660.

[11] I. Guyon, M. Nikravesh, S. Gunn, and L. A. Zadeh, *Feature Extraction Foundations and Applications*, Springer, 2006, doi: 10.1007/978-3-540-35488-8.

[12] B. Remeseiro and V. Bolon-canedo, "A review of feature selection methods in medical applications," *Comput. in Biol. and Medicine*, vol. 112, no. 103375, pp. 1-9, July 2019, doi: 10.1016/j.compbiomed.2019.103375

[13] G. Kumar, "A detailed review of feature extraction in image processing systems," in *2014 Fourth International Conference on Advanced Computing & Communication Technologies*, 2014, pp. 5-12, 2014, doi: 10.1109/ACCT.2014.74.

[14] D. Tian, "A review on image feature extraction and representation techniques," *Int. J. Multimedia and Ubiquoitous Eng.*, vol. 8, no.4, pp. 385-395, July 2013.

[15] A. J. Afifi and W. M. Ashour, "Image retrieval based on content using color feature," *International Scholarly Research Notices*, vol. 2012, pp. 1-11, 2012, doi: 10.5402/2012/248285.

[16] Z. Lan and Y. Liu, "Study on multi-scale window determination for GLCM texture description in high-resolution remote sensing image geo-analysis supported by GIS and domain knowledge," *International Journal Geo-Information*, vol. 7, no. 5, pp. 1-24, 2018, doi: 10.3390/ijgi7050175.

[17] P. Mohanaiah, P. Sathyanarayana, and L. Gurukumar, "Image texture feature extraction using GLCM approach," *Int. J. Scientific and Res. Publications*, vol. 3, no. 5, pp. 1-5, 2013. [Online]. Available: http://www.ijsrp.org/research-paper-0513/ijsrp-p1750.pdf

[18] A. Zotin, Y. Hamad, K. Simonov, and M. Kurako, "Lung boundary detection for chest X-ray images classification based on GLCM and probabilistic neural networks," *Procedia Computer Science*, vol. 159, pp. 1439-1448, 2019, doi: 10.1016/j.procs.2019.09.314.

[19] F. Y. Manik, Y. Herdiyeni, and E. N. Herliyana, "Leaf morphological feature extraction of digital image Anthocephalus Cadamba," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 14, no. 2, pp. 630-637, 2016, doi: 10.12928/TELKOMNIKA.v14i2.2675.

[20] Z. Wu, S. Jiang, X. Zhou, Y. Wang, and Y. Zuo, "Application of image retrieval based on convolutional neural networks and Hu invariant moment algorithm in computer telecommunications," *Comput. Commun.,* vol. 150, no. December 2019, pp. 729-738, 2020, doi: 10.1016/j.comcom.2019.11.053.

[21] C. Cortes and V. N. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995, doi: 10.1007/BF00994018.

[22] M. R. Phangtriastu, J. Harefa, and D. F. Tanoto, "Comparison between Neural network and support vector machine in optical character recognition," *Procedia Comput. Sci.,* vol. 116, pp. 351-357, 2017, doi: 10.1016/j.procs.2017.10.061.

[23] H. Bisgin *et al.,* "Comparing SVM and ANN based machine learning methods for species identification of food contaminating beetles," *Scientific Reports*, vol. 8, no. 6532, pp. 1-12, 2018, doi: 10.1038/s41598-018-24926-7.

[24] N. Kranjˇciˊc, D. Medak, R. Župan, and M. Rezo, "Machine learning methods for classification of the green infrastructure in city areas," *Int. J. Geo-Inf*, vol. 8, no. 463, pp. 1-14, 2019, doi: 10.3390/ijgi8100463.

[25] M. Yang and G. Thung, "Classification of Trash for Recyclability Status," Stanford, 2016. [Online]. Available: http://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf

[26] S.-H. Tsang, "Review: PSPNet-Winner in ILSVRC 2016 (Semantic Segmentation/Scene Parsing)," Towards Data Science, 15 December, 2018. Accessed: 28-Oct-2020. [Online]. Available: https://towardsdatascience.com/review-pspnet-winner-in-ilsvrc-2016-semantic-segmentation-scene-parsing-e089e5df177d.

[27] M.-P. Hosseini, M. R. N.-Zadeh, D. Pompili, and H. S.-Zadeh "Statistical validation of automatic methods for hippocampus segmentation in MR images of epileptic patients," in *Int. Conf. of the IEEE Eng. in Medicine and Biol. Soc.,* Aug. 2014, doi: 10.1109/embc.2014.6944675.

[28] S. M. Singh and K. Hemachandran, "Content-based image retrieval using color moment and gabor texture feature," *IJCSI Int. J. Comput. Sci. Issues*, vol. 9, no. 5, pp. 299-309, 2012.

[29] Ş. Öztürk and B. Akdemir, "Application of feature extraction and classification methods for histopathological image using GLCM, LBP, LBGLCM, GLRLM and SFTA," *Procedia Comput. Sci.*, vol. 132, pp. 40-46, 2018, doi: 10.1016/j.procs.2018.05.057.

[30] P. S. S. Kumar and V. S. Dharun, "Extraction of texture features using GLCM and shape features using connected regions," *Int. J. Eng. and Technol. (IJET)*, vol. 8, no. 6, pp. 2926-2930, 2017, doi: 10.21817/ijet/2016/v8i6/160806254.

[31] J. Han and M. Kamber, *Data Mining Concepts and Techniques Second Edition*, Morgan Kaufmann Publisher, 2006. [Online]. Available: https://mitmecsept.files.wordpress.com/2017/04/data-mining-concepts-and-techniques-2nd-edition-impressao.pdf

[32] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.,* vol. 2, no. 3, pp. 1-27, 2011, doi: 10.1145/1961189.1961199.

[33] A. Karatzoglou, D. Meyer, and K. Hornik, "Support Vector Machines in R," *Journal of Statistical Softwer*, vol. 15, no. 9, pp. 1-28, 2006, doi: 10.18637/jss.v015.i09.

[34] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, vol. 45, no. 4, pp. 427-437, 2019, doi: 10.1016/j.ipm.2009.03.002.

BIOGRAPHIES OF AUTHORS

Khadijah obtained the Bachelor of Informatics Engineering (S.Kom) from the Universitas Diponegoro, Indonesia in 2011 and the Master of Computer Science (M.Cs) from the Universitas Gadjah Mada, Indonesia in 2014. She has been a lecturer with the Department of Informatics, Universitas Diponegoro since 2014. Her main research interests are artificial intelligence and machine learning. She is also a member of international profesional association, such as IEEE.

Sukmawati Nur Endah obtained the Bachelor of Science (S.Si) in Mathematics from the Universitas Diponegoro, Indonesia in 2001 and the Master of Computer Science (M.Kom) from the Universitas Indonesia, Indonesia in 2009. She is currently working as a lecturer with the Department of Informatics, Universitas Diponegoro. Since 2019, she has been the Head of Laboratorium of Intelligent System in the Department of Informatics, Universitas Diponegoro. Her main research interests are artificial intelligence and machine learning. She is also the member of APTIKOM, the Indonesian professional organization in computer science.

Retno Kusumaningrum obtained the Bachelor of Science (S.Si) in Matematics from the Universitas Diponegoro, Indonesia in 2003 and the Master of Computer Science (M.Kom) and the Doctoral Degree (Dr.) in Computer Science from the Universitas Indonesia, Indonesia in 2010 and 2014, respectively. She is currently working as a lecturer and the Head of Department of Informatics, Universitas Diponegoro. Her main research interests are natural language processing and machine learning. She is also a member of International profesional association, such as IEEE and member of some Indonesian professional organization, such as INACL, INAPR and APTIKOM.

Rismiyati obtained the Bachelor of Engineering (B.Eng) from the Nanyang Technological University, Singapore in 2007 and the Master of Computer Science (M.Cs) from the Universitas Gadjah Mada, Indonesia in 2016. She is currently working as a lecturer with the Department of Informatics, Universitas Diponegoro, Indonesia. Her main research interests are machine learning and computer vision. She is also a member of International profesional association, such as IEEE.

Priyo Sidik Sasosongko obtained the Bachelor of Science (S.Si) in Mathematics from the Universitas Diponegoro, Indonesia in 1996 and the Master of Computer Science (M.Cs) from the Universitas Gadjah Mada, Indonesia in 2006. He is currently working as a lecturer with the Department of Informatics, Universitas Diponegoro, Indonesia. His main research interests are machine learning and artificial intelligence.

Iffa Zainan Nisa obtained the Bachelor of Informatics (S.Kom) from the Department of Informatics, Universitas Diponegoro, Indonesia. She has been a research assistant with the Laboratorium of Intelligent System in Department of Informatics, Universitas Diponegoro since 2020. Her main research interests are machine learning and computer vision.