Vision Based Floating Garbage Classification using SIFT

Swati Shilaskar  
E&TC Dept.,  
VIT, Pune , India.  
swati.shilaskar@vit.edu

Rushikesh Mali  
E&TC Dept.,  
VIT, Pune , India.  
rushikesh.mali20@vit.edu

Shripad Bhatlawande   
E&TC Dept.,  
VIT, Pune , India.  
 shripad.bhatlawande@vit.edu

Sahil Parekh  
E&TC Dept.,  
VIT, Pune , India.  
sahil.parekh20@vit.edu

Sarthak Bhake  
E&TC Dept.,  
VIT, Pune , India.  
sarthak.bhake20@vit.edu

.

***Abstract*—** ***Pollution caused due to unregulated dumping of waste in water bodies has been a problem for some time now. Water Pollution contributes to a lot of environmental as well as social hazards and poses a risk to the public health and well-being of humans. Major Steps are being taken to limit this type of pollution but it’s far from being completely tackled. This study proposes a vision-based approach to identifying different types of water pollutants. Classification-based machine learning algorithms like KNN, Random Forest, XGBoost, SVM, etc. are used on a custom-made dataset containing 1500 images of waste. Feature extraction of all the images is done using the SIFT feature extraction algorithm which is then fed to the classifiers. Machine learning techniques like Dimensionality reduction, K fold and Grid Search Cross Validation are used to improve the accuracy and optimize model performance. The proposed model boasts a highest accuracy of 95 percent with a ROC-AUC score of 0.97. The result of the proposed study is an efficient classification of floating waste in different target classes.***

***Keywords— SIFT, Kmeans Clustering, OpenCV, Floating waste classification, Machine Learning***

1. INTRODUCTION

Pollution has been affecting the earth, ever since industrialization and globalization too advent. Pollution is caused in all types of natural places like land, water bodies and air, to save our mother earth and its natural resources it is essential to take preventive measures which would curb pollution. The yearly pollution created humans across all forms of pollutions is 1.3 billion tons which is equivalent to the weight of equivalent to the weight of 6.5 million whales. Recently water and air pollution rates have been on the rise and have seen incredible number of wastes being exposed to them. Plastic pollution in water bodies raised a lot of concerns when plastic pollution in the Pacific Ocean formed a patch which was thrice the size of France [1]. This event was named as the Great Pacific Garbage Patch.

Water pollution is one of the primal concerns of environmentalists because this type of pollution is not only harmful for humans but is also very harmful aquatic life. Floating wastes on the surface of water bodies, blocks sunlight to enter beneath the ocean which is very harmful for creatures living beneath. If the concentration of wastes is more, it also blocks oxygen which is fatal. Water pollution is a host to a lot of toxic and hazardous pollutants which are

directly being dumped into the water body without any filtering and processing. Plastic Wastes, Toxic Chemicals, Domestic and Industrial Sewage, and oil pollution are some of the types of pollutants which contribute to most wastes in water bodies. Toxic wastes/chemicals are the most harmful types of pollutants, the possess a threat to the groundwater ecosystem which is still a source of drinking water for billions of people across the earth. Another major concerning type of pollutant is plastic and single used plastic materials like bottles, straws etc. which are used in abundance by humans. Sadly, more than 8.5 million tons weight of plastic trash, or 19 billion pounds, are thought to enter our seas each year. This plastic is largely derived from single-use containers like bottles and plastic bags as well as other single-use items like straws and disposable cups and plates. According to one estimate, there will be more plastic in the oceans by weight than aquatic animals by the year 2050 [2].

Recycling is essential for a society that is sustainable since it reduces the quantity of garbage. Recycling facilities must currently manually sift rubbish and employ several big filters to separate more definite things as part of the recycling process. As a result, waste categorization is another industry-relevant application of computer vision that has lately gained a lot of academic attention. Machine learning and Deep learning might increase the effectiveness of processing facilities by classifying garbage. This will have advantageous economic consequences in addition to favorable environmental effects.

The proposed model is one such approach in tackling the modern problem of water pollution by using computer vision and machine learning. The proposed work is capable to classify three types of wastes namely plastic bottles, plastic bags, and biomedical waste. A custom-made dataset consisting of 1500 images for used for model training. The model accurately classifies the type of waste based on the input image given.

The remaining paper is structured as follows, section 2 illustrates about the related work that has been previously done in this space. Followed by this Section 3 elaborates about the methodology of the proposed approach along with the results obtained. This is followed by Section 5 which concludes this manuscript.

1. RELATED WORK

Zhuang Kang et al. have proposed a smart garbage classification system based on deep learning which consists of a solar powered garbage bin that classifies the litter into biodegradable and non-biodegradable categories. The model used for the system is ResNet-34 which was further modified in three aspects in its network structure. The model gives an outstanding final accuracy of 99.6% [3].

The authors have put forward a garbage classification system established on deep learning principles. The system makes use of two deep learning models viz. ResNet-50 which is a 50 layer deep neural network and YOLO an extensively used object detection system [4]. The models were tested on a custom dataset which gave accuracies of 98.5% for ResNet-50 and 97.6 for YOLO respectively. A trash classification system based on machine learning models which would aid for recyclability status. The model aims to classify the images into one of the five categories viz. plastic, metal, paper, glass, cardboard. A dataset of 400-500 images per class was used for model training. The system makes use of various models including KNN, SVM, XgBoost, Random Forest and CNN. Out of all the models CNN gives the most accuracy of 89.81% [5].

Farzana Shaikh et al. have put forward a waste profiling system backed by a deep learning approach to classify the garbage into biodegradable and non-biodegradable classes. The Inception-v3 model was trained on about 2700 images. The trained model was deployed using a flask web server and an android application was built to capture images of garbage and send it to the web server for further classification. The system was able to classify images with an accuracy of 83.30% [6]. Anh H. Vo et al. have designed model to classify trash based on deep learning. The model uses a refined version of the ResNext model i.e., DNN-TC. The model was trained on two datasets, VN-trash dataset which includes three classes of trash namely organic, inorganic, biomedical and the Trashnet dataset. The model gave accuracies of 94% and 98% for Trashnet and VN-trash respectively [7].

Gary Thung et. al performed classification of waste across 6 different classes namely glass, paper, metal, plastic, cardboard, and trash. The author proposed an approach using SVM and CNN for classification by using SIFT as a feature extractor. Although the CNN wasn’t trained on the data due to lack of correct hyperparameters the accuracy achieved through SVM was 63%. The reason for low accuracy and precision/recall score was lack of a good data source [8].

A waste bin was suggested by Wesley et al. [9]. Although the method is highly reliant on the environment, there are still some restrictions on its classification. The research on trash categorization systems is advanced, but there is still room for improvement in terms of speed and accuracy. Additionally, there aren't many deep learning-based models on the classification of garbage/trash. Deep learning technology is now utilized extensively in picture categorization and has produced some amazing results. Talking about hardware solutions to waste classification, a lot of robots and automated bots are being researched about in the academia. A robot system was put out by Kano et al. [10] which can/should offer decentralized control method for indoor garbage/trash pickup and collection. However, it requires a lot of time and cooperation from the robot’s end.

Bridging the gap between AI and hardware was a model proposed by M. Swathi et al. [11] where the smart dustbin can read the type of waste present in it and create a real time dataset for the neural network. The output of the neural network is then used to connect all sorts of peripheral devices and various sensors to create an end-to-end system. For majority of the systems to train a model on waste classification a dataset named TrashNet is used a lot. This dataset has six different classes with roughly 500-600 images per class. Utilizing this dataset for by building sophisticated deep learning models on top of it can lead to a good accuracy and highly optimized model performance. One such study proposed to categorize garbage on the Trash net dataset, Aral et al. [12] used transfer learning models derived from several well-known CNN models for image classification, including Densenet121, DenseNet169, InceptionResnetV2, MobileNet, and Xception. The authors employed 70% of the Trashnet dataset for training, 13% for validation, and 17% for testing in their studies. Additionally, 8 and 224 x 224 were chosen as the batch and input size, respectively. The trial findings showed that a DenseNet121 transfer learning model archived the greatest accuracy, which produces at 95% score.

Continuing the work done using deep learning algorithms, a few other models were also researched where different experiments were run using different hyperparameters. One experiment was run for the categorization of waste, Bircanoglu et al. [13] created the RecycleNet light-weight convolutional neural network model. RecycleNet decreased the time complexity by lowering the number of parameters from seven million to three million, although only achieving 81% accuracy for the Trashnet dataset with 70% of pictures for training, 13% for validation, and 17% for testing. RecycleNet is a simple concept for many systems that limit hardware devices as a result.

A lot of work followed these experiments by proposing more newer and advanced models for waste classification built upon the trashnet dataset. For automatically classifying garbage, Ruiz et al. [16] tested the usage of numerous CNN models, including VGG, Inception, and ResNet. 80% of the Trashnet dataset was utilised for training in this study, 10% for validation, and the remaining 10% for testing. With an average accuracy of 88.66% for the Trashnet dataset, the ResNet-based architecture produced the greatest performance results. In addition to the approaches created for trash classification that the study has just described, several well-known CNN models for image classification, including ResNext [14], ImageNet [15], VGG [17], ResNet [18], and DenseNet [19], may also be utilized to categorize garbage.

Table 1 depicts the comparison of the papers which were the most relevant to our proposed system. The table contains comparison of the papers on this basis of the dataset used, algorithms implemented and the performance of their model. The comparison table helps in finding research gaps and thus helps in figuring the novelty of the model/system that is being proposed. The research gaps and the novelty of the proposed model is illustrated following the table.

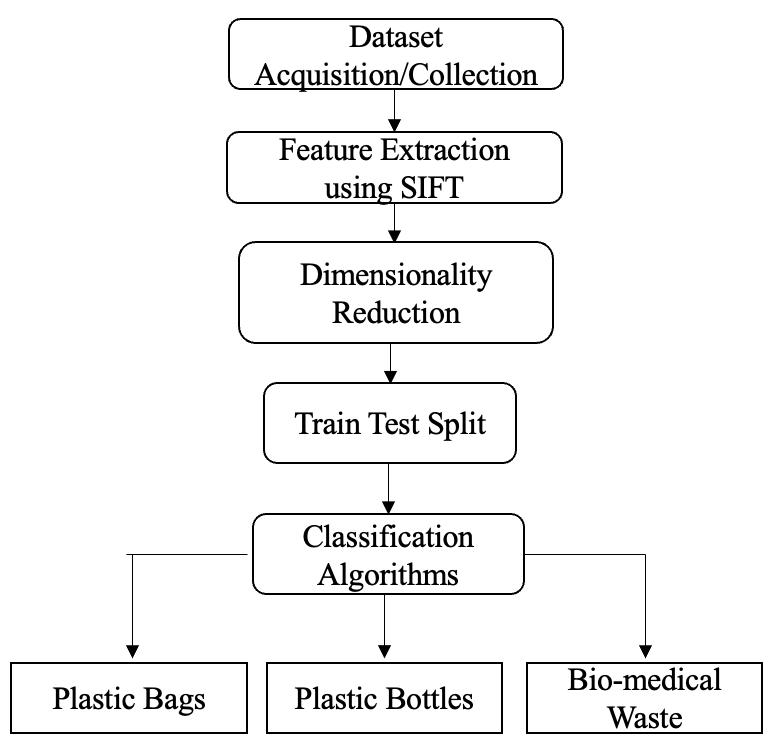
TABLE I. Comparison of dataset and performance of different papers discussed.

|  |  |  |  |
| --- | --- | --- | --- |
| References | Dataset | Algorithm | Performance |
| [4] H. S. Pandita et al.  2022 | Custom Dataset with 400-500 images for each class | ResNet-50 | 98.50% Test Accuracy |
| [8] Gary Thung et al.  2016 | Own dataset with total of 2400 images | SVM with 70/30 Train Test Split | 63% Test Accuracy |
| R. A. Aral Et al.  2018 | TrashNet, Adam, Adadelta | DenseNet121, DenseNet169 | 95% Test Accuracy |
| [16] V. Ruiz Et al.  2019 | TrashNet with 2527 images | CNN Based ResNet | 88.66% Test Accuracy |

Majority of the papers reviewed/surveyed had limitations catering two mainly two aspects, lack of a good dataset and less accuracy and precision scores with machine learning algorithms. Almost all the authors resorted to deep learning techniques and algorithms to improve the accuracy and optimize the model performance. This study presents a machine learning based model backed with a good data source which classifies between three types of wastes with an accuracy of 95%, without the use of any deep learning algorithms.

1. METHODOLOGY
2. *Block Diagram*

The outline/flow of the proposed research work can be understood by referring to Fig 1. The figure contains the block diagram of the proposed model, right from data collection to model’s performance evaluation. Feature extraction using SIFT, Dimensionality reduction and k means were the major techniques performed*.* The model uses 5 different classification algorithms which are discussed in further detail in the following sections.



**Fig. 1.** Outline/Flow of the Proposed work

1. *Data Collection and Pre-processing*

The data used for the study was custom made by downloading selective images from the internet. The data consists of 3 distinct classes with approximately 500 images in each class, taking it to a total of 1500 sample images. The 3 distinct classes consist of different types of wastes in water bodies viz. Plastic Bottles, Plastic Bags and Biomedical Waste. After the collection of primary images, preprocessing of images was undertaken by resizing them and gray scaling to nullify the RGB channels. Fig 2 shows a sample image of the plastic bottle class before and after gray scaling.

A picture containing outdoor object

Description automatically generated

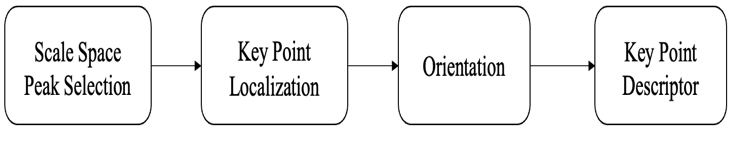


**Fig. 2.** Sample Images from the 3 considered classes

1. *Feature extraction using SIFT.*

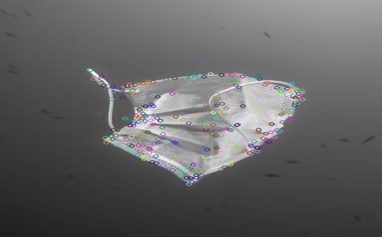
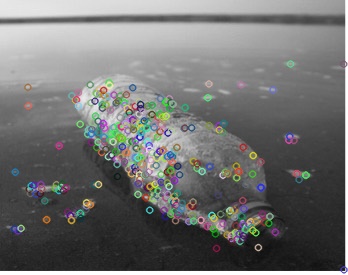
The feature extraction technique used for this study is Scale-invariant feature transform popularly known as SIFT. This technique is mainly used to detect interest points on specific input images. It also supports the identification of localized features. SIFT can carry out feature identification independently of the image's viewpoint, depth, and scale, in contrast to other feature extraction techniques, which depend on these factors. In the proposed work, SIFT is applied to each image of each class. The output of the feature extractor is a pandas data frame which is then appended to each other to form a combined data frame of features.

Fig 3 shows the sequence of steps followed in SIFT feature extraction technique.



**Fig. 3.** Sequence of steps in SIFT feature extraction technique

Fig 4 shows the key points extracted from SIFT feature extractor for sample images from all the 3 classes viz plastic bag, plastic bottle and biomedical waste.



**Fig. 4.** Key points Extracted from SIFT

1. *Dimensionality reduction*

Dimensionality reduction is a technique in machine learning used to reduce the dimensions (columns) of a particular file. In this study, the output obtained from the SIFT feature extraction technique were huge data frames. After combining the data from all 3 classes the final data frame had a shape of 962917 rows × 128 columns. To reduce the 128 dimensions/columns, k means clustering technique is used which reduces the dimensions to just 8 columns. Finalizing the k in this K means the method is the most essential step. This value is achieved by performing the elbow method which is a graph with the number of clusters on the x-axis and the within-cluster sum of squares (WCSS) on the y-axis. The corresponding value of K where the graph molds into an elbow-like shape is chosen.

Fig 5 illustrates the elbow graph obtained in this proposed study. According to the graph, the corresponding value of K is equal to 7.

Chart, line chart

Description automatically generated

**Fig. 5.** Elbow method graph (K vs WCSS)

1. *Classification Algorithms and Performance Metrics*

Depending upon the type of waste in the input image the image will be classified into 3 classes viz plastic bottles, plastic bags, and biomedical waste. For classification purposes, almost all algorithms have been used to do a comparative analysis between all of them. All algorithms used in the study are listed below with a brief description. Decision Trees builds a tree-like model by recursively splitting the data into subsets based on the most important features, allowing for easy interpretation and visualization of the decision-making process. K Nearest Neighbors is used for predicting the classification or value of a new data point by identifying the k nearest training data points in the feature space and assigning the majority class or average value among the k neighbors.

Random Forest creates a collection of decision trees and uses their collective output to make predictions. It reduces overfitting by randomly selecting subsets of features and observations for each tree in the forest. Support Vector Machine constructs a hyperplane or a set of them in a high-dimensional space to separate different classes. XgBoos*t* is a gradient-boosting algorithm that is designed to minimize prediction errors by combining several weak models in a weighted manner. These were the 5 different types of algorithms used for classification. These algorithms were judged and ranked based on some performance metrics which are discussed below.

The specific metrics chosen for evaluating the performance of a classifier are Accuracy, F1 Score, ROC AUC score and confusion matrix. These are statistically proven metrics to evaluate the performance of a classifier.

Precision is the measure to evaluate the number of times the model got the correct prediction in correspondence to the training data. Whereas Recall is a metric which indicates the ratio of positives which are predicted to the total number of positive labels. But when dealing with multi class classification the true metric to evaluate the performance of the classifier is F1 score. F1 score is the harmonic mean between precision and recall and can wave all the tradeoffs offered by precision and recall. The formula for F1 score is (2 X Precision X Recall)/ (Precision + Recall). Figure 6 illustrates the Accuracy, F1 and ROC AUC scores of the algorithms.

Chart, bar chart

Description automatically generated

**Fig. 6.** F1 and ROC AUC scores of the algorithms

The dataset containing images for training were resized to 512 x 512 dimension and then grayscale to nullify the RGB channels. As part of an essential step with working with in machine learning is feature extraction, which is performed by using the SIFT technique for key point localization and description. The descriptor output which was received from SIFT was then clustered using K means clustering algorithm. After this an efficient technique named K fold was used to train the algorithms on the final dataset. K fold uses leave one out method, in our case k is equal to 5, so out of 5 subsets 4 are used for training while the one left is used for testing. Total iterations are equal to the number k, and we consider the mean accuracy of all the iterations. Table 2 depicts F1 score, ROC AUC Score and the mean accuracies for all algorithms along with their standard deviation.

TABLE II. Classification Results before tuning.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Accuracy | F1 Score | ROC AUC Score |
| DT | 77.44 +/- 3.75 | 78 | 83 |
| KNN | 87.60 +/- 2.04 | 91 | 97 |
| RF | 87.22 +/- 2.70 | 87 | 97 |
| SVM | 77.54 +/- 2.40 | 77 | 91 |
| XgBoost | 86.45 +/- 2.02 | 88 | 96 |

*\*Note – All values are in percentage.*

However, to optimize the model performance cross validation using Grid Search Cross Validation was performed on Random Forest and K nearest neighbors’ algorithm. Both the algorithms showed improvement in accuracy with KNN giving the nest accuracy of 95%. Hyperparameters like number of decision tress and the maximum depth of those tress were tuned during the hyperparameter tuning process for random forest. The hyperparameters tuned for KNN were the distance metrics like Euclidean (L2 Norm) and Manhattan (L1 Norm) as well as the value of k which varied from 3 to 9 inclusive. Table 3 depicts the accuracies achieved after implementing the hyperparameter tuning for random forest.

TABLE III. Hyperparameter tuning results in % for Random Forest

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Max Depth/No. of DTs | 300 | 400 | 500 | 600 | 700 |
| 11 | 87.51 | 86.84 | 87.03 | 87.51 | 87.61 |
| 13 | 87.71 | 87.13 | 87.13 | 87.22 | 87.71 |
| 15 | 87.90 | **88.29** | 87.80 | 87.32 | 87.61 |
| 17 | 87.90 | 88.00 | 87.80 | 88.00 | 88.00 |
| 19 | 88.00 | 87.32 | 88.19 | 87.90 | 88.19 |

*\*Note – All values are in percentage.*

The value 88.29 which is highlighted in bold in Table 3 is the best accuracy achieved for random forest. After hyperparameter tuning on the random forest algorithm using the grid search cross validation technique the best set of hyperparameters found were (15,400) which are maximum depth of a tree and number of decision tress respectively.

The same approach of K fold + Grid search cross validation was used to tune the hyperparameters of KNN algorithm which has initially given the best accuracy of 87.6 before tuning. The parameters tuned for KNN were the distance metrics and the value of k. Algorithm 1 describes the process hyperparameter tuning for KNN. The algorithm involves initialization of the parameters grid and using grid search cross validation on top of the parameters grid declared.

**Algorithm 1:** Grid Search Cross Validation for KNN

1. **Start.**
2. Initialize a parameter grid of distance metrics and number of neighbors hyperparameters.
3. Create object of the KNN classifier.
4. Create object of Grid Search Cross Validation and pass the initialized KNN object, the parameter grid, and the value of k for k-fold Cross Validation.
5. Train the model with specified parameters.
6. Evaluate performance of the trained model.
7. Procure the best parameters.
8. **End.**

After implementation of the steps mentioned in algorithm 1, a huge jump in accuracy was noticed for KNN. The accuracy rose from 87% to 95%, thus resulting in the best accuracy achieved for the proposed model. Table 4 depicts the accuracies for each combination of the hyperparameter for KNN.

TABLE IV. Hyperparameter tuning results in % for KNN

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Distance Metrics/Value of K | 3 | 5 | 7 | 9 |
| Euclidean | 94.26 | **95.07** | 93.61 | 94.00 |
| Manhattan | 94.26 | 94.88 | 93.71 | 93.74 |

*\*Note – All values are in percentage.*

CONCLUSION AND FUTURE SCOPE

The research work mainly focuses on tackling the ever-increasing environmental issue of waster pollution. The study helps in accurately identifying and classifying the most hazardous pollutants like plastic bottles, plastic bags, and biomedical waste. The developed model is light weight system utilizing the SIFT feature extractor which can be used and deployed into any real-world device/application which aims to identify waste floating on water bodies. Machine learning algorithms are used to greater efficiency to yield respectable accuracy levels for the dataset used. Future work can be focused adding more classes i.e., for classification, so that more and more types of wastes get identified and put under the process of recycling/dumping.

REFERENCES

1. Britannica, T. Information Architects of Encyclopaedia (2023, April 25). pollution. Encyclopedia Britannica. <https://www.britannica.com/facts/pollution-environment>
2. Petruzzello, M.. "Plastic Disaster: How Your Bags, Bottles, and Body Wash Pollute the Oceans." Encyclopedia Britannica, July 17, 2017.
3. Z. Kang, J. Yang, G. Li and Z. Zhang, "An Automatic Garbage Classification System Based on Deep Learning," in IEEE Access, vol. 8, pp. 140019-140029, 2020, doi: 10.1109/ACCESS.2020.3010496.
4. H. S. Pandita, V. Vaidya, M. Doifode, P. Bhavthankar and A. Venkatesh, "Garbage Classification using Machine Learning to Aid Recycling," 2022 3rd International Conference for Emerging Technology (INCET), Belgaum, India, 2022, pp. 1-6, doi: 10.1109/INCET54531.2022.9824215.
5. Satvilkar, Mandar. "Image based trash classification using machine learning algorithms for recyclability status." PhD diss., Dublin, National College of Ireland, 2018.
6. F. Shaikh, N. Kazi, F. Khan and Z. Thakur, "Waste Profiling and Analysis using Machine Learning," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp. 488-492, doi: 10.1109/ICIRCA48905.2020.9183035.
7. A. H. Vo, L. Hoang Son, M. T. Vo and T. Le, "A Novel Framework for Trash Classification Using Deep Transfer Learning," in IEEE Access, vol. 7, pp. 178631-178639, 2019, doi: 10.1109/ACCESS.2019.2959033.
8. Gary Thung and Mindy Yang, “Classification of Trash for Recyclability Status,” CS 229, Stanford University, 2016
9. W. Pereira, S. Parulekar, S. Phaltankar and V. Kamble, "Smart Bin (Waste Segregation and Optimisation)," 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 2019, pp. 274-279, doi: 10.1109/AICAI.2019.8701350.
10. T. Kano, E. Naito, and T. Aoshima, ‘‘Decentralized control for swarm robots that can effectively execute spatially distributed tasks,’’ Artif. Life, vol. 26, pp. 243–260, Apr. 2020, doi: 10.1162/artl\_a\_00317
11. A. Chandramohan, J. Mendonca, N. R. Shankar, N. U. Baheti, N. K. Krishnan and M. S. Suma,” Automated Waste Segregator”, 2014 Texas Instruments India Edu- cators’ Conference (TIIEC), Bangalore, 2014, pp. 1-6. doi: 10.1109/TIIEC.2014.009
12. R. A. Aral, S. R. Keskin, M. Kaya, and M. Haciomeroglu, ‘‘Classification of trashnet dataset based on deep learning models,’’ in Proc. BigData, Dec. 2018, pp. 2058–2062.
13. C. Bircanoglu, M. Atay, F. Beser, O. Genc, and M. A. Kizrak, ‘‘RecycleNet: Intelligent waste sorting using deep neural networks,’’ in Proc. INISTA, 2018, pp. 1–7.
14. A. Krizhevsky, I. Sutskever, and G. S. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ Commun. ACM, vol. 60, no. 6, pp. 84–90, 2017.
15. S. Xie, R. B. Girshick, P. Dollár, Z. Tu, and K. He, ‘‘Aggregated Residual Transformations for Deep Neural Networks,’’ in Proc. CVPR, 2017, pp. 5987–5995.
16. V. Ruiz, Á. Sánchez, J. F. Vélez, and B. Raducanu, ‘‘Automatic imagebased waste classification,’’ in Proc. IWINAC, vol. 2, 2019, pp. 422–431.
17. K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ in 2014, arXiv:1409.1556.
18. K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in Proc. CVPR, 2016, pp. 770–778.
19. G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, ‘‘Densely connected convolutional networks,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 2261–2269