Analysis of multi-class convolution neural network on garbage classification system

Vikas K R
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
vikaskr8055@gmail.com

Ullas P
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
ullasputtaraju@gmail.com

Sri Sai Teja M S
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
srisaitejms2002@gmail.com

Kumar Swamy S
Professor Head & Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
Skswamy@gat.ac.in

Ashwini Kodipalli
Associate Professor & Head,
Department of Artificial Intelligence and Data Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
dr.ashwini.k@gat.ac.in

Trupthi Rao
Assistant Professor, Department of Artificial Intelligence and Data Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
trupthirao@gat.ac.in

***Abstract*— The buildup of solid waste in metropolitan areas is a major worry that, if not effectively handled, might lead to environmental contamination and be dangerous to human health. To manage a range of waste products, it's crucial to have an advanced/intelligent waste management system. The act of separating garbage into its many components is one of the most crucial parts of waste management, and it is often carried out manually by hand-picking. To simplify this, we suggest a customized architecture for multiclass Convolutional Neural Networks (CNNs). The design includes elements that are individual-specific, enabling the model to adjust to the traits of each class. Our CNN architecture outperforms conventional CNN models like ResNet50 and VGG16 in multiclass classification problems by utilizing this personalized approach. The effectiveness and resilience of our suggested architecture are demonstrated by experimental results on our dataset, opening the path for increased classification accuracy of 71.15%.**

Keywords— Multi-class convolution neural network, Garbage classification, Waste Separation.

I. Introduction

According to the World Bank's What a Waste 2.0[1] assessment an estimated 33 percent of the annual 2,011 billion tons of municipal solid waste produced worldwide is not managed in a manner that is environmentally friendly. The globe produces 0.74 kilograms of trash per person per day on average, however there is a wide range, from 0.11 to 4.54 kilograms. Despite making up only 16% of the world's population, high-income countries create garbage around 34%, or 683 million tons of garbage.

By 2050, it is predicted that there will be 3.40 billion tons of waste produced worldwide, which is more than double the population growth during that period. Income level and waste production are frequently strongly correlated. The average amount of trash created worldwide each day is 0.74 kilograms, however there is a wide range, ranging from 0.11 to 4.54 kilograms. Despite making up only 16% of the world's population, high-income countries create around 34%, or 683 million tons, of the garbage. According to the central pollution related annual report for 2020–21. By 2050, it is anticipated that high-income countries' daily waste production per person would increase by 19%; now, more than half of trash is thrown out in the open; action must be taken right away since the trajectory of trash increase will have a negative impact on the environment, human health, and economic progress.

According to the Central Pollution Control Board of Delhi's annual report for 2020–21 [2], the nation generates 160038.9TPD of solid waste, of which 95.4% is collected and only 50% is treated; the remaining 18.4% is dumped, and the remaining 31.7% of the total waste generated goes unnoticed. In recent years in some countries have begun to investigate recycling tactics in a new sort of cyclical economy for sustainable growth that enhances environmental quality. The most common method for classifying garbage is manual sorting since it is currently the most precise one. Sadly, it takes a lot of time and requires experienced operators, which severely limits the rate of categorization of waste. Therefore, an automated trash categorization solution is urgently required to meet this expanding problem and has therefore emerged as a global research hotspot. Convolutional neural network (CNN)-based object identification algorithms have progressed quickly and are now widely employed in the environmental protection industry because of improvements in computer technology, such as the graphics processing unit (GPU), in recent years. Although the area of trash classification has attained good performance, it still confronts difficulties due to the wide range of categories, variable states, and complicated textures, which might result in a major performance drop. It is resolved in this work.

The remainder of this paper is organized as follows: The associated works are evaluated in Section II, with their benefits and drawbacks being highlighted. The recommended tactic is presented in Section III. In Section IV, all the experimental data, conclusions, and analyses are provided. The last section provides a summary of the article.

II. Existing work

Today, several automated methods for classifying waste have been suggested. These methods may be divided into three categories: mechanical methods (MAs), Internet of Things methods (ITAs), and artificial intelligence methods (AIAs). To successfully replace human garbage sorting, an MA uses microprocessors, external sensors, and mechanical gearbox in an automatic garbage categorization system. However, because to low-accuracy recognition, this method does not produce the intended trash categorization result [3]– [5]. A unique ITA technique for automated trash categorization was put out to solve this issue. A more automated and accurate waste categorization system was created using the ITA and a cloud server, however owing to the high price and intricate design of the system, it is challenging to deploy and maintain [6]– [9]. Artificial intelligence (AI)-based AIAs for waste categorization are extremely accurate, adaptable, and durable. To perform trash detection and classification tasks, a prediction model trained on garbage data might be used. However, the current classification algorithms do not meet the requirements of trash classification systems and run on high-performance servers or PCs [10]– [13]. This study suggests a brand-new deep learning-based embedded Linux system for intelligent garbage categorization. The system combines an integrated Linux platform for real-time processing with a convolutional neural network (CNN) model for picture categorization. [14]. Gives An overview of garbage detection and categorization approaches utilising deep learning algorithms is given in this survey study. It explores several methods for classifying garbage, including CNNs, and shows both their benefits and drawbacks.[15]. The study mainly concentrated on garbage identification utilising aerial hyperspectral data, which may not be easily transferrable to other circumstances or image-based methods for waste categorization [16]. Due to the benchmark datasets and the proposed garbage image recognition method's focus on domestic waste, its application to other waste categorization domains may be constrained [17]. The deep learning framework CGBNet, created exclusively for compost categorization, is introduced in this research. It may not address other waste types or management issues because it only classifies compost [18]. An artificial neural network-based method for predicting the production of municipal solid trash is presented in this paper. It does not particularly address trash classification [19]. The research focuses on utilising machine learning approaches to categorise rubbish according to its suitability for recycling. It offers insightful information on recycling but skips over other waste management or deep learning topics [20]. The idea of very deep convolutional networks (VGGNet) for extensive image recognition is introduced in this fundamental paper. Despite not being specifically for garbage categorization, it offers a basis for deep learning techniques applied in this area [21].

The previously mentioned references have various limitations, even if they offer insightful contributions to the subject of intelligent waste categorization systems employing deep learning. By considering several data modalities and expanding the evaluation to more complex and varied circumstances, our study addresses these constraints.

III. PROPOSED WORK

A. Dataset and Data Collection

The details of the dataset development are discussed in this section.

We developed the dataset, which contains 6 garbage categories and a total of 2521 images, of which 2019 images were used for training and 502 for testing, to promote the application of deep learning in the field of environmental protection and improve the performance of the object recognition algorithm for garbage images. The dataset's trash has the following characteristics: fixed shape, variable shape, constant texture, variable texture, and various scales. We collected each picture and gave it a labelled and the fig 1-6 are the sample images available in the dataset.

Fig. 1. Cardboard

Fig. 2. Glass

Fig. 3. Metal

Fig. 4. Paper

Fig. 5. Plastic

Fig. 6. Trash

Due to the lesser size of each class, data augmentation techniques were used on each image. Rescale, Shear_range of (0.2), Zoom_range (0.4), and Horizantal Flip were some of these methods. Fig. 7 is a presentation of the precise information abotut dataset and Table. 1 is the amount of images in each category of waste. And number of images for training and testing.

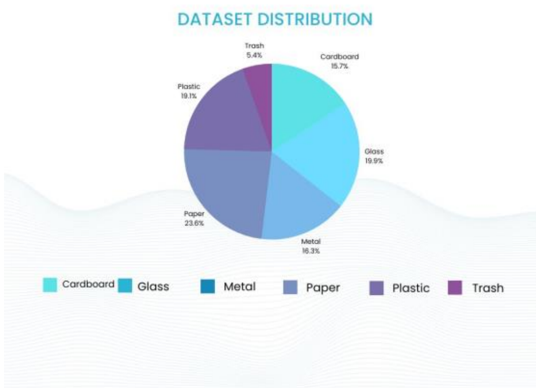

Fig.7.Dataset distribution

Table 1 . Number of images in the training and testing of each category of waste

Class	Training	Testing
Cardboard	322	75
Glass	400	101
Metal	328	82
Paper	475	119
Plastic	385	97
Trash	109	28

B. Model

In this part, in Fig. 8 we describe our model architecture in depth and go over the training process.

Deep learning models have significantly advanced picture classification tasks in recent years, revolutionizing the area of computer vision. Convolutional Neural Networks (CNNs) have become an effective method for removing significant characteristics from pictures and obtaining cutting-edge results in a variety of visual identification applications. In this study, we provide a brand-new CNN architecture for precise and effective picture categorization into multiple classes.

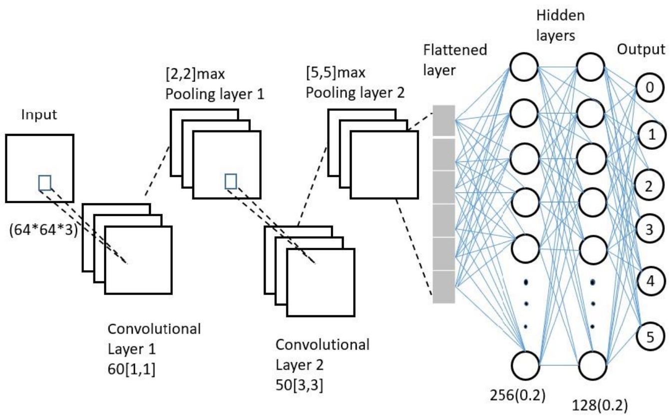

Fig 8. The architecture of the proposed method

Input Layer: pictures having the shape (64, 64, 3) are supplied to the input layer; these pictures have a width and height of 64 pixels and three RGB color channels. For jobs involving picture categorization, this input size is common.

Convolutional Layers: 60 filters with a size of (1, 1) are applied by the first convolutional layer (Conv2D). These filters assist in removing minute details from the input photos. Then Rectified Linear Unit (ReLU) activation function, which adds non-linearity to the network and enhances its capacity to recognize complex patterns. Following the first convolutional layer comes a MaxPooling2D layer with a pool size of (2, 2). It conducts a down sampling technique, cutting the feature maps' spatial dimensions in half while keeping the most important information.

In the second convolutional layer, 50 filters with a (3, 3) pixel size are applied. In comparison to the data recorded by the first layer, these filters aid in the extraction of higher-level features that are more complicated and abstract. After the second convolutional layer MaxPooling2D layer with a pool size of (5, 5). In doing so, the spatial dimensions are further reduced, and the most important data from the feature maps are captured.

Flattening Layer: The preceding layer's 2D feature maps are converted into 1D vectors using the flattening layer. To link the convolutional layers to the fully linked layers, this flattening process is required.

Fully Connected Layers: A first dense layer with 256 units is present. By collecting abstract representations of the input pictures, it performs the function of a high-level feature extractor.

After the initial dense layer, a Dropout layer with a rate of 0.2 is added to prevent overfitting. Dropout forces the model to acquire more robust and generalized characteristics by randomly deactivating certain neurons during training.

The learnt characteristics are further refined in the second dense layer, which has 128 units.

The second dense layer is followed by a Dropout layer with a rate of 0.2, which adds more regularization.

Output Layer: Six units make up the ultimate dense layer, the same number of waste classification groups. Each unit stands for a certain type of garbage.

The output layer is given the SoftMax activation function, which results in a probability distribution over the classes. As a result, the model can give each waste category a probability value, indicating how likely it is that the input image belongs to that class.

Overall, this architecture blends fully connected layers for high-level feature processing and classification with convolutional layers for feature extraction. The model can learn discriminative features, decrease overfitting, and enhance generalization with the use of ReLU activation functions, max pooling, and dropout layers. This model can be taught to categorize fresh garbage photographs reliably and automatically by training on a huge dataset of labelled waste images.

ResNet50: It is a deep residual network with exemplary performance in picture classification tasks. To solve the vanishing gradient issue, it has 50 layers and incorporates skip connections, sometimes referred to as residual connections. The model can successfully transport gradients through the network thanks to the skip connections, which makes it possible to train deeper designs.

With ResNet50, the pre-trained model serves as a feature extractor when implementing transfer learning. The model's first layers, which oversee low-level feature extraction, are frozen and maintained the same. To meet the job at hand, the later layers—which oversee high-level feature extraction and classification—are adjusted or replaced with new layers. The target dataset is then used to train these additional layers.

VGG16: It is Another popular pre-trained model for image categorization It is renowned for its ease of use and potency. Small 3x3 filters are used in the 16 convolutional layers that make up VGG16, which are followed by fully linked layers for classification.

When the target dataset and the dataset on which it was trained have comparable properties, such as similar item categories or visual patterns, VGG16 is very helpful. The early layers of VGG16's learnt representations capture general visual cues that are helpful for a variety of picture classification applications. However, VGG16 has more parameters than ResNet50, which might make training it more expensive computationally.

C. Equations

Rectified Linear Unit (ReLU): To introduce non-linearity, the ReLU activation function is used.

$$ReLU(x) = max(0, x)$$

If the input value is positive, the ReLU function returns the value; if it is negative, it returns zero. It facilitates the introduction of nonlinearity and solves the vanishing gradient problem.

SoftMax: For multi-class classification problems, the SoftMax activation function is frequently utilised in the output layer of a neural network. The resulting logits are transformed into a probability distribution across the classes.

$$SoftMax(x_i) = exp(x_i) / sum(exp(x_j))$$

Each input value is exponentiated before being divided by the total of all the exponentiated values across all classes. The values are guaranteed to be between 0 and 1, total to 1, and indicate the class probabilities.

Adam optimizer: Adam (Adaptive Moment Estimation) is a popular optimisation approach for neural network training. It combines the advantages of AdaGrad and RMSProp, two additional optimisation techniques. Based on the first and second moments of the gradients, the method modifies the learning rate for each network parameter..

$$m\_t = beta1 * m\_t\text{-}1 + (1 - beta1) * g\_t$$

$$v\_t = beta2 * v\_t\text{-}1 + (1 - beta2) * g\_t\text{^}2$$

$$m\_t\_hat = m\_t / (1 - beta1\text{^}t)$$

$$v\_t\_hat = v\_t / (1 - beta2\text{^}t)$$

$$w\_t{+}1{=}w\_t\text{-}learning\_rate{*}m\_t\_hat/(sqrt(v\_t\_hat){+}epsilon)$$

The first and second moment estimations of the gradients are represented by m and v. The current time step is indicated by the exponential decay rates for beta1 and beta2 are. The gradients are shown by g. w stands for the weights. The learning rate is represented by learning_rate, and the tiny constant epsilon prevents division by zero.

Sparse Categorical Cross Entropy: A popular loss function for multi-class classification issues when the classes are mutually exclusive is sparse categorical cross entropy. It calculates the difference between the actual class labels and the expected class probabilities.

Loss = -sum (y_true * log(y_pred))

In this case, y-pred stands for anticipated class probabilities, while y_true represents the true class labels (as integers). The loss is calculated by summing the true class label times the logarithm of the expected probability for all classes. The loss is transformed into a minimization issue using the negative sign.

IV. EXPERIMENTS AND RESULTS

This section compares the results of our custom-built model employing convolutional neural network (CNN) on our specific problem domain and dataset with different transfer learning techniques, namely ResNet50 and VGG16.

A. *Training Accuracy and Loss*

The training loss and accuracy graphs for three models—ResNet50, VGG16, and our own custom model are shown in Fig. 9-11 in this section. These graphs show the models' development as learners and can be used to evaluate how well they Train. The training loss graph shows how the loss function changes during training. It demonstrates how the models get better at reducing mistakes over time. Reduced loss and enhanced convergence are shown by lower values on the y-axis.

The training accuracy graph displays the accuracy of the models on the training data over time. The capacity of the models to produce accurate predictions as they learn from the training data is shown by the higher numbers on the y-axis.

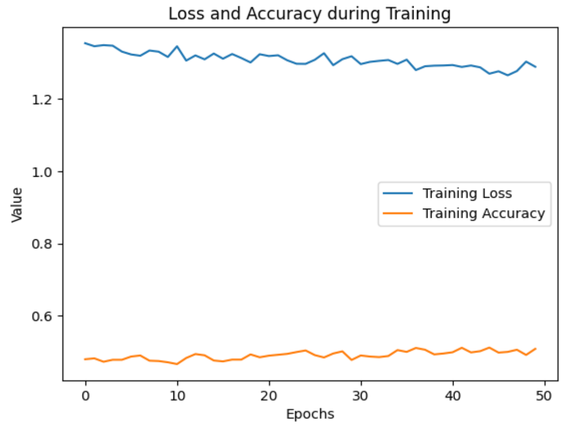

Fig. 9. Training and accuracy of ResNet50

Fig. 10. Training and accuracy of VGG16

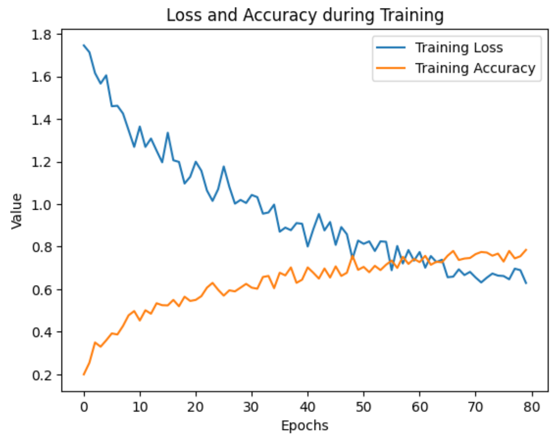

Fig. 11. Training and accuracy Custom model

Three models—ResNet50, VGG16, and our own custom model—training loss and accuracy graphs are shown in above Figures respectively. The loss and accuracy values are represented on the y-axis, while the number of training iterations or epochs is represented on the x-axis. The training loss is represented by the blue line, which demonstrates how the loss values of the models decrease during the training cycles. The objective is to spot a declining trend, which denotes enhanced convergence and model performance. The training accuracy is shown by the orange line, which throughout the process shows the accuracy values of the models on the training data. Our goal is to spot a rising trend that shows how the models get better at making predictions as the training goes on. We can learn more about the training behaviours of models by contrasting the training loss and accuracy graphs of those models. These graphs give a complete picture of the learning ability of the models and emphasise variances in convergence, overfitting, or underfitting tendencies.

B. *Testing Accuracy and Confusion_matrix*

The accuracy and loss values of three distinct image classification models—ResNet50, VGG16, and a custom model—are shown in Table 2. The same dataset is used to measure these models' accuracy in classifying images.

Table.2. Accuracy and Loss Table

MODEL	ACCURACY	LOSS
ResNet50	0.4462	1.4873
VGG16	0.6992	1.3818
Custom-Built	0.7171	1.0192

The ResNet50 model obtained a greater test loss of 1.4873 and a significantly lower test accuracy of 44.62%. This suggests that, in contrast to the other models included in this study, the model had difficulty correctly classifying pictures. To increase its performance on a particular dataset The ResNet50 architecture may need more fine-tuning or changes because of its well-known deep and complicated structure. The VGG16 model demonstrated a higher test accuracy of 69.92% and a test loss of 1.3818. This indicates that VGG16 performed better than ResNet50, but there is still room for improvement. VGG16 is a well-known and widely used architecture, but its performance may vary depending on the dataset and problem domain. Fine-tuning or hyperparameter optimization techniques can potentially enhance its accuracy further. The custom-built model showcased the test accuracy of 71.71% highest among the evaluated models. It achieved a test loss of 1.0192, indicating a relatively lower error rate compared to ResNet50 and VGG16. The custom-built model's superior performance suggests that it was able to capture and leverage the dataset's specific characteristics effectively. This result highlights the potential advantages of designing a model tailored to the problem domain.

Confusion Matrix Representation:

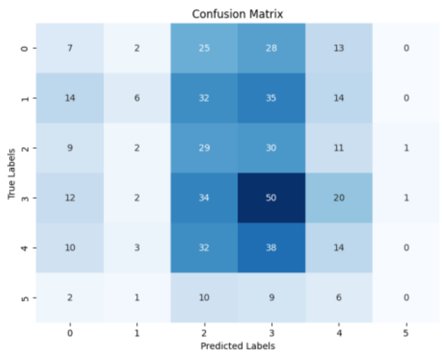

Fig. 12. Confusion matrix of ResNet50

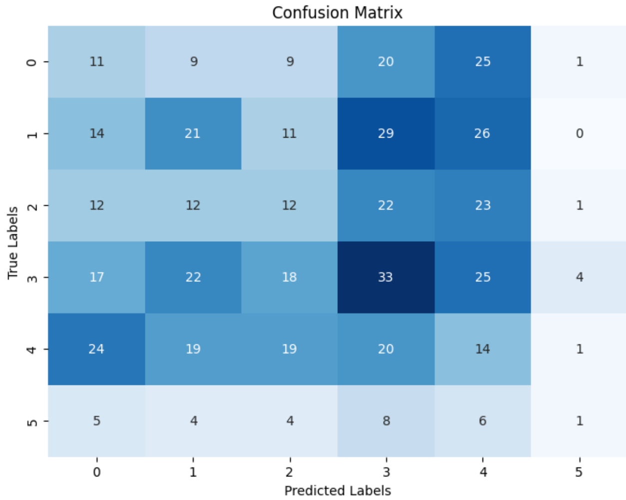

Fig. 13. Confusion matrix of VGG16

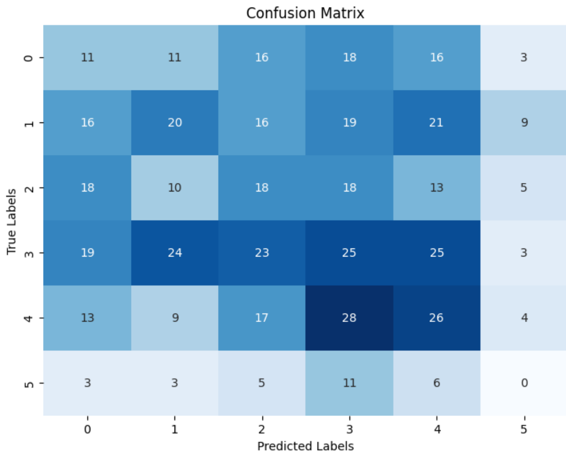

Fig. 14. Confusion matrix of Custom-Model

In Figures 12, 13, and 14, we present the confusion matrices for ResNet50, VGG16, and our custom-built model, respectively. Each matrix provides a visual representation of the model's performance in classifying the test data. The matrix elements correspond to the number of instances classified correctly (true positives and true negatives) or incorrectly (false positives and false negatives) for each class.

To improve the performance of our custom model, we conducted a series of experiments involving different optimizers and varying dropout ratios. This was done as our model consistently outperformed established architectures such as ResNet and VGG16. We wanted to explore if fine-tuning the optimization process and adjusting the dropout ratio could further enhance our model's accuracy. By employing optimizers like Adam, SGD, ADAGRAD, ADADELTA and RMSprop with varying dropout ratio, we were able to observe their impact on training convergence and overall performance. Additionally, we adjusted the dropout ratio, which controls the regularization strength and prevents overfitting, to find an optimal value. To visualize the results, we plotted the corresponding accuracy readings on a bar graph in Fig. 15-16, allowing us to compare the different optimizer-dropout combinations and identify the most effective configuration for our custom model.

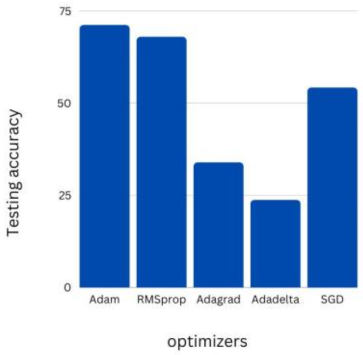

Fig.15.accuracy of Custom-model for different optimizer

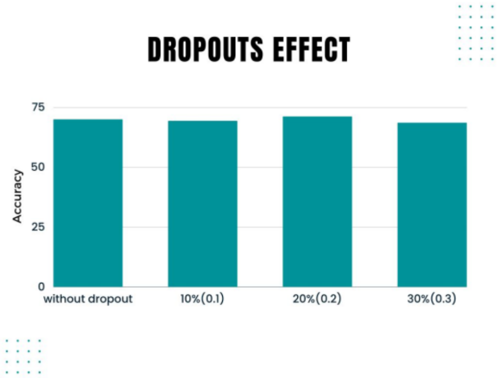

Fig.16. effect of optimizers on accuracy of model

V. Conclusion

In conclusion, the custom-built model outperformed ResNet50 and VGG16 in terms of test accuracy. The variation in performance underscores the importance of selecting the appropriate model architecture and it for the specific dataset and problem domain. We suggest a waste categorization system that uses deep learning to distinguish various waste components. This technology may be used to categorise garbage automatically, lowering the need for human involvement and decreasing pollution and illness. We obtained an accuracy of 71% from the results when tested against the garbage dataset. Using our approach, the garbage will be separated more quickly and intelligently, maybe even with less input from humans. The accuracy of the system may be increased by including additional images in the dataset. By changing some of the current settings, we will eventually enhance our system so that it can classify more trash items.

References

[1] Kaza, Silpa; Yao, Lisa C.; Bhada-Tata, Perinaz; Van Woerden, Frank. 2018. What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. Urban Development;. © Washington, DC: World Bank. http://hdl.handle.net/10986/30317 License: CC BY 3.0 IGO.

[2] CENTRAL POLLUTION CONTROL BOARD DELHI, ANNUAL REPORT 2020-21, Website: www.cpcb.nic.in.

[3] J. Hui, ‘‘Design and implementation of intelligent garbage sorting system based on RFID,’’ J. Anhui Inst. Electron. Inf. Technol., vol. 17, no. 4, pp. 10–13, 2018.

[4] T. Zhang, M. Li, L. Li, and W. Luo, ‘‘Design of garbage classification system based on RFID,’’ J. Phys., Conf. Ser., vol. 1744, no. 2, Feb. 2021, Art. no. 022111.

[5] C. Zhou, ‘‘Design of intelligent sorting trash dustbin based on STM32,’’ in Proc. E3S Web Conf., 2020, Art. no. 04032\

[6] [6] P. Pan, J. Lai, G. Chen, J. Li, M. Zhou, and H. Ren, ‘‘An intelligent garbage bin based on NB-IOT research mode,’’ in Proc. IEEE Int. Conf.Saf. Produce Informatization (IICSPI), Dec. 2018, pp. 113–117.

[7] Y. Wang, Y. Xu, B. Zhang, J. Zhang, and X. Su, ‘‘The design and imple mentation of the smart trash can based on the Internet of Things,’’ J. Phys., Conf. Ser., vol. 1550, May 2020, Art. no. 022003.

[8] Oralhan, Z., Oralhan, B., & Yiğit, Y. (2017). Smart city application: Internet of things (IoT) technologies based smart waste collection using data mining approach and ant colony optimization. Internet Things, 14(4), 5.

[9] Lu, Zhongzhi, and Na Xu, ‘‘Application strategies of waste sorting facil ities based on Internet of Things,’’ in Innovative Computing. Singapore:Springer, 2020, pp. 1291–1296.

[10] Z. Kang, J. Yang, G. Li, and Z. Zhang, ‘‘An automatic garbage classification system based on deep learning,’’ IEEE Access, vol. 8, pp. 140019–140029, 2020.

[11] O. Adedeji and Z. Wang, ‘‘Intelligent waste classification system using deep learning convolutional neural network,’’ Proc. Manuf., vol. 35,pp. 607–612, Jan. 2019.

[12] H. Wang, ‘‘Garbage recognition and classification system based on con volutional neural network VGG16,’’ in Proc. 3rd Int. Conf. Adv. Electron.Mater., Comput. Softw. Eng. (AEMCSE), Apr. 2020, pp. 252–255.

[13] K. Yan, W. Si, J. Hang, H. Zhou, and Q. Zhu, ‘‘Multi-label garbage image classification based on deep learning,’’ in Proc. 19th Int. Symp. Distrib. Comput. Appl. Bus. Eng. Sci. (DCABES), Oct. 2020, pp. 150–153.

[14] B. Fu, S. Li, J. Wei, Q. Li, Q. Wang and J. Tu, "A Novel Intelligent Garbage Classification System Based on Deep Learning and an Embedded Linux System," in IEEE Access, vol. 9, pp. 131134-131146, 2021, doi: 10.1109/ACCESS.2021.3114496.

[15] H. Abdu and M. H. Mohd Noor, "A Survey on Waste Detection and Classification Using Deep Learning," in IEEE Access, vol. 10, pp. 128151-128165, 2022, doi: 10.1109/ACCESS.2022.3226682.

[16] D. Zeng, S. Zhang, F. Chen and Y. Wang, "Multi-Scale CNN Based Garbage Detection of Airborne Hyperspectral Data," in IEEE Access, vol. 7, pp. 104514-104527, 2019, doi: 10.1109/ACCESS.2019.2932117.

[17] Z. Wu et al., "New Benchmark for Household Garbage Image Recognition," in Tsinghua Science and Technology, vol. 27, no. 5, pp. 793-803, October 2022, doi: 10.26599/TST.2021.9010072.

[18] S. Gangopadhyay and A. Zhai, "CGBNet: A Deep Learning Framework for Compost Classification," in IEEE Access, vol. 10, pp. 90068-90078, 2022, doi: 10.1109/ACCESS.2022.3201099.

[19] Kumar, S., Gaur, A., Kamal, N., Pathak, M., Shrinivas, K., & Singh, P. (2020, April). Artificial Neural Network Based Optimum Scheduling and Management of Forecasting Municipal Solid Waste Generation–Case Study: Greater Noida in Uttar Pradesh (India). In *Journal of Physics: Conference Series* (Vol. 1478, No. 1, p. 012033). IOP Publishing.

[20] Yang, M. and Thung, G., 2016. Classification of trash for recyclability status. CS229 project report, 2016(1), p.3.

[21] TY - JOUR, AU - Simonyan, Karen, AU - Zisserman, Andrew, PY - 2014/09/04, SP - ,T1 - Very Deep Convolutional Networks for Large-Scale Image Recognition,JO - arXiv 1409.1556,ER-