2024 6th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)
Military Institute of Science and Technology (MIST), Dhaka-1216, Bangladesh

Garbage Classification using a Transfer Learning with Parameter Tuning

Nazmul Haque, Rizoan Toufiq, Md. Zahirul Islam, Md Al Amin Tokder Shoukhin
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
Emails: nazmulruetcse18@gmail.com, rizoantoufiq@gmail.com, km.zahir61@gmail.com, alamintokdercse@gmail.com

***Abstract*—To better manage waste resources, minimize environmental pollution, and simplify garbage sorting, this paper presents a method for garbage classification and recognition using a fusion network based on transfer learning. We begin by balancing the dataset through data augmentation and then integrate a pre-trained model with a customized CNN model. Three Convolutional Blocks with varying filter sizes (32, 64, and 128) and Mish activation functions are employed. The flattened output of the last MaxPool2D layer is fed into three fully connected layers (1024, 512, and 12 neurons) with Softmax activation for multi-class prediction. Evaluation of the model reveals DenseNet169 achieved the highest accuracy of 99.58%. Our approach yields an efficient garbage classification model, requiring less training time and parameters. Mish activation function demonstrates superior performance compared to ReLU. Previous models were more complex, requiring extensive training time and parameters. This study underscores the importance of automated garbage classification and presents an advanced solution.**

***Index Terms*—Garbage Classification, Transfer Learning Fusion (TLF), Convolutional Block (CB).**

I. Introduction

Garbage in the environment refers to waste or rejected components that are considered no longer useful. Garbage has significantly harmed the natural circumstance.Therefore, differentiation and reuse of garbage is very significant to make a green and sustainable development society [1]. The global production of urban solid waste totals 2.01 billion tons annually, this is a threat to the ecological environment. garbage generation will be grown by 70% ,if the recent status persist. Therefore, automated garbage detection system is very urgent. Nevertheless, the entire process of recycling incurs significant hidden expenses, stemming from the sorting, categorization, and treatment of recycled materials. [2]. But Deep Learning in computer vision, especially for tasks like image classification and object detection, has proven effective and straightforward. CNNs, a vital part of deep models, have significantly advanced image classification by capturing features and making robust assumptions about image nature [3].
Several research endeavors have employed a variety of models in their investigations. For example Shanshan Meng and Wei-Ta Chu [4], Harshita Dooja Poojary et al [5] incorporated Support Vector Machines and Artificial Neural Networks (ANN) in their research. Specific studies utilized Convolutional Neural Network (CNN) models referenced by [4]–[10]. Meanwhile, Tanya Gupta et al. [11] and Mohammad Kamrul Hasan et al. [12] applied the Inception model. Mohammad Kamrul Hasan et al. [12] employed a combination of DenseNet169, InceptionNet-v4, and MobileNet. Conversely, Nikita Garg and Sunanda Das [13] chose SqueezeNet, VGG-19, and GoogLeNet. Shanshan Meng and Wei-Ta Chu [4] reported a 95.35% accuracy in garbage classification using the RestNet50 model.Additional models, such as HOG+SVM and CNN, exhibited impressive performance. Capturing an image of an object against a pristine backdrop is impractical. Harshita Dooja Poojary et al. [5] attained a 97.0% accuracy employing CNN. Other models, like ANN and TL, demonstrated robust and commendable performance. The inability to highlight a specific region remains a limitation. Yuchen Wang et al. [6] achieved accuracy of 90.88% using CNN and weight pruning.A significant drawback of this study is the absence of loss recovery techniques, such as fine-tuning. Li Cao and Wei Xiang [7] attained a 93.20% accuracy by employing transfer learning with the Inception-v3 model. The image size and quality is not satisfiable and relatively low accuracy. Sabitabrata Bhattacharya et al. [8] employed CNN to achieve a 95.0% accuracy in garbage classification. The primary drawbacks include the inability to highlight crucial features of an image and a comparatively low level of performance. Ms. Ishita Joshi et al. [9] achieved 92.5% accuracy using CNN.Comparatively low performance and difficulty in emphasizing a specific region persists as a limitation. Dong Wang and Zhongsheng Wang [10] introduced a garbage detection approach with 96.29% accuracy using CNN. Tanya Gupta et al. [11] employed InceptionNet model, with achieving classification rates of 98.15%. Mohammad Kamrul Hasan et al. [12] utilized DenseNet169, achieving an accuracy of 97%. Additional models, including DenseNet121, InceptionNet-v4, and MobileNet, exhibited strong and commendable performance. A limitation of their work lies in poor interpretability. Nikita Garg and Sunanda Das [13] obtained an accuracy of 98.79% utilizing SqueezeNet. Alternative models, such as VGG-19 and GoogLeNet, showcased strong and good performance. However, a significant limitation lies in poor generalization stemming from a limited amount of data. Cumulatively, these investigations provide valuable perspectives and approaches for the identification and classification of garbage, making substantial contributions to progress in environmental management. Literature review briefly explain in Table I

The contribution of our proposed deep learning-based

979-8-3503-8577-9/24/ $31.00 ©2024 IEEE

TABLE I

SUMMARY OF LITERATURE REVIEW

Article	Accuracy	Precision	Specificity	Published Year
ResNet50 [4]	95.35	-	-	2020
CNN [6]	90.88	-	-	2022
CNN [7]	93.20	-	-	2020
CNN [8]	95.00	-	-	2023
CNN [5]	97.00	-	-	2022
CNN [9]	92.50	-	-	2023
CNN [10]	96.29	-	-	2021
InceptionNet [11]	98.15	-	-	2022
DenseNet169 [12]	97.00	-	-	2022
SqueezeNet [13]	98.79	99.11	97.16	2022

model is integrating a pre-trained model with a customized CNN model to create an improved transfer learning model.

Our paper is organized as follows: Section II provides an overview of the dataset and Sections III explain our research approach, data preprocessing, Proposed Transfer Learning (TL) Architecture, and Justification of Our Procedural Architecture. In Section IV discuss about performance evaluation metrics, Experimental Setup, Performance analysis for training and validation data, and Performance analysis for test data. section V contain considerations of potential result validity threats. Finally, Section VI summarizes findings, and in Section VII, we suggest future research directions, appendix and followed by references.

II. DATASET DESCRIPTION

The summary of our "garbage" dataset is conveniently presented in Table II, which can be easily accessed on Kaggle [14]. Furthermore, Figure 1 provides a visual representation of the dataset's distribution across various classes.

TABLE II

SUMMARY OF THE DATASET

No of Images	**Format**	**No of Classes**	**Source**
15515	JPG	12	kaggle.com

TABLE III

DATA DISTRIBUTION

Classes	No Of Images	Classes	No Of Images
battery	945	metal	769
biological	985	paper	1050
Brown-glass	607	plastic	865
cardboard	891	shoes	1977
clothes	5325	trash	697
green-glass	629	white-glass	777

The dataset exhibits an imbalance in its distribution of data. It is important to note that all images obtained through web scraping are the intellectual property of their respective original photographers or owners.

III. METHODOLOGY

A. Sequential workflow of the proposed methodology

The garbage dataset was obtained from kaggle.com [14], but it exhibited an imbalance, as illustrated in Table II. To address this issue, we applied Data Augmentation techniques. Subsequently, we conducted preprocessing operations outlined in Section III-B. The preprocessed dataset was split into training (70%), validation (10%), and test (20%) sets. We then integrated a Customized CNN model with a pretrained model. The training data was utilized to train the customized transfer learning model, while the validation data was used to validate it. Finally, we evaluated the model's performance using the test data. The workflow of our research is illustrated in Figure 2.

B. Data Preprocessing

In preprocessing, images undergo cleaning, resizing, and normalization for consistency and reduced noise. The dataset is divided into training (70%), validation (10%), and test (20%) data. Data Augmentation is applied to address the imbalanced class distribution, ensuring effective training and model assessment. Data augmentation benefits by creating extra training samples, making the dataset larger. Diverse variations help the model understand new data and handle input variations more effectively. Conversely, disadvantages include the risk of the model becoming too specific to augmented data and increased computing demands.

C. Architecture of Proposed Transfer Learning (TL)

Our approach begins by inputting images into pre-trained models, followed by reshaping tensors for fine-tuning. Three convolution blocks (CBs) with 32, 64 and 128 filters each are utilized, incorporating varying kernel sizes (5x5, 3x3, and 1x1) and three 'BatchNormalization' layers. Each CB concludes with a MaxPooling2D layer, employing Mish activation functions on all convolution layers to address the vanishing gradient problem. The output from the last max-pooling layer is flattened and traverses three fully connected layers (1024, 512, and 12 neurons) with Mish activation functions. The final layer uses a softmax activation function to predict class probabilities. Selected pre-trained models, such as DenseNet, ResNet, and Xception, are strategically trained using diverse architecture types. The architectural overview is shown in Figure 3.

D. Justification of Our Procedural Architecture

In CNN, batch normalization tackles difference of data distribution for each layer during training. This accelerates training, achieving convergence quickly, and improves overall performance. Three convolution layers with different kernel sizes (5x5, 3x3, and 1x1) with less filtering (32, 64, and 128) are used to create a lightweight model. This explains why our model performs well with less parameters and requires less training time compare to previous research work.

Fig. 1. Sample images for each classes

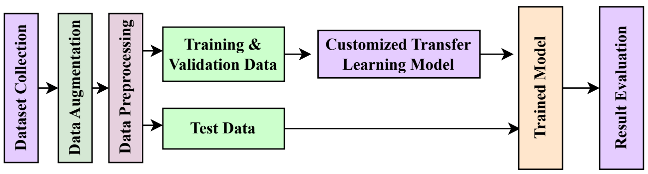

Fig. 2. Sequential workflow of the proposed methodology

IV. Result And Performance Analysis

A. Performance Evaluation Metrics

Various metrics, including accuracy, precision, recall (sensitivity), f1-score, and specificity were employed to evaluate model effectiveness. The mathematical expressions for these metrics are presented below [15].

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

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

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

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

$$Specificity = \frac{TN}{TN + FP} \tag{5}$$

Accuracy represents overall correctness, precision measures the model's ability to avoid false positives, recall assesses its ability to capture true positives, F1-score combines precision and recall, and specificity gauges the ability to avoid false negatives.

B. Experimental setup

Architectural operations were conducted on Kaggle using a GPU P100 and a 2-core Intel Xeon CPU (690 ms/step). Input images measured (224,224,3), and the dataset was split into separate training, validation, and testing folders in a 0.70-0.10-0.20 ratio. No holdout set was used. Models pass through 50 epochs (batch size 16) with the Adam optimizer (lr: 0.001, loss: categorical cross-entropy). Early stopping utilized Reduce on Plateau (patience: 25).

C. Performance analysis for training And validation dataset

Figure [4-10] illustrates how the training and validation accuracy's, as well as the training and validation losses, evolve with increasing epochs for the corresponding model. We observe that the DenseNet-based model outperforms the ResNet and Xception-based models. Specifically, DenseNet169 exhibits the highest training and validation accuracies among all models, as shown in Figure 5. In the ResNet-based model, both training and validation accuracy's experience a slight decrease compared to the DenseNet model. Conversely, in the Xception-based model, both training and validation accuracies show a slight increase compared to the ResNet model.

D. Performance analysis for test dataset

TABLE IV
Performance Analysis on test data

Algorithm	Accuracy %	Precission%	Recall%	F1-score%
DN121_Tuned	99.31	99.31	99.31	99.31
DN121	99.38	99.38	99.38	99.38
DN169_Tuned	99.58	99.58	99.58	99.58
DN169	99.47	99.47	99.47	99.95
DN201_Tuned	99.56	99.56	99.56	99.56
DN201	99.55	99.55	99.55	99.55
RN50_Tuned	98.81	98.81	98.81	98.81
RN50	99.25	99.25	99.25	99.25
RN101_Tuned	98.81	98.81	98.81	98.81
RN101	99.03	99.03	99.03	99.03
RN152_Tuned	98.97	98.97	98.97	98.97
RN152	99.39	99.39	99.39	99.39
XN_Tuned	98.91	98.91	98.91	98.91
XN	98.57	98.57	98.57	98.57

The performance outcomes from our TLF architectures has been displayed by Table IV. In our table, different model has been represented by each row, and we have been evaluated their performance by various metrics such as accuracy, precision, recall, and F1-score, each indicating different aspects of model performance. DenseNet (DN), RestNet (RN), and XceptionNet (XN) have been employed in our study. For instance, the highest accuracy of 99.58% has been achieved by "DN169_Tuned". This means correct predictions have been

Fig. 3. Architectural view of the proposed methodology

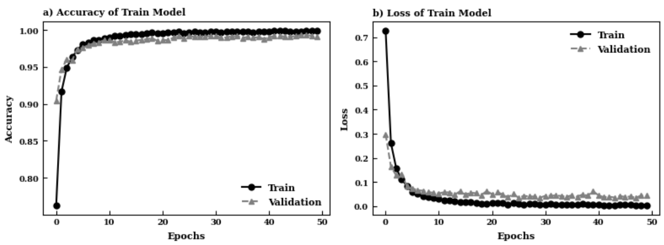

Fig. 4. Accuracy & loss curve for DenseNet121

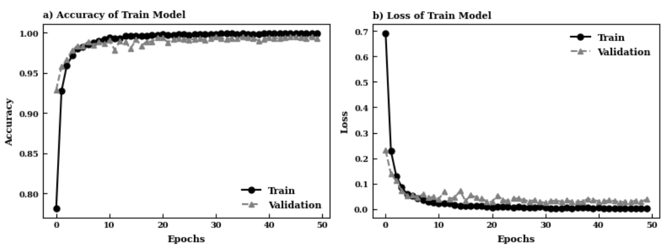

Fig. 5. Accuracy & loss curve for DenseNet169

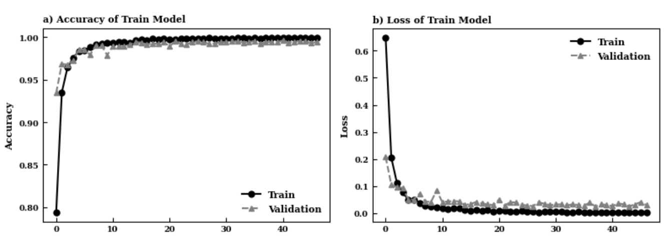

Fig. 6. Accuracy & loss curve for DenseNet201

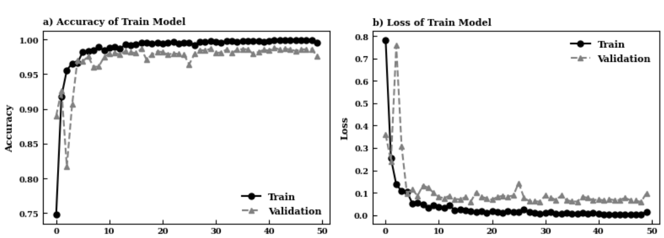

Fig. 7. Accuracy & loss curve for ResNet50

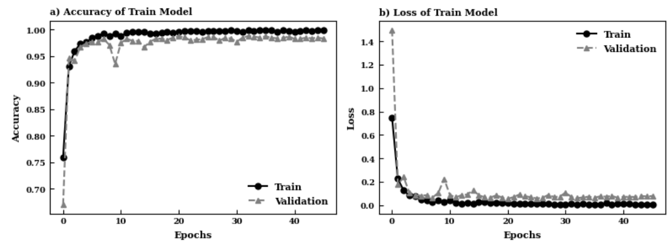

Fig. 8. Accuracy & loss curve for ResNet101

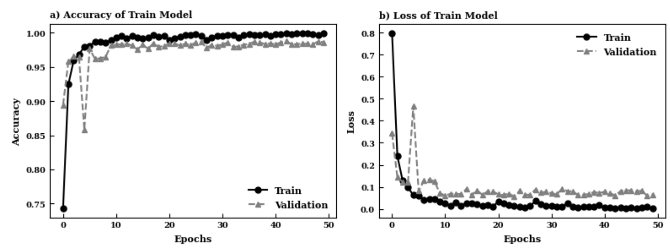

Fig. 9. Accuracy & loss curve for ResNet152

made for about 99.58% of the cases it has encountered. On the other hand, the lowest accuracy of 98.81% has been achieved by RN101_Tuned and . DN121_Tuned has achieved the lowest accuracy of 99.31% among all the DenseNet models, although all these models have provided higher performance. But better performance has been provided by DN169 models compared to DN121 and DN201 model. Here, it is demonstrated that the performance of RestNet models has been lower than other models. However, the performance of the XceptionNet model has increased slightly compared to RestNet models. We demonstrate that certain scenarios with the pretained model outperform the strategy we have suggested. Yet in this investigation, our suggested approach yielded the greatest

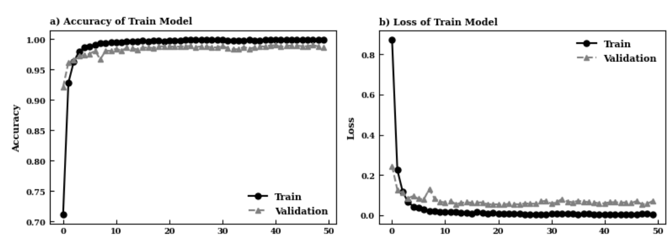

Fig. 10. Accuracy & loss curve for XceptionNet

results. Overall, a concise summary of how different models have performed across various evaluation criteria is provided by this table.

V. Threats To Validity

Our study addresses dataset imbalance using augmentation but has limitations. Future work should explore new data balancing techniques. Deploying the model on mobile devices is complex and may impact performance due to resizing and normalization operations.

VI. Conclusion

In this paper, we focus on garbage image classification based on Transfer Learning Fusion Network. If we continue with the existing garbage management system, it will harm our environment. That's why efficient garbage management is vital for sustainability but faces challenges in manual sorting. Our goal is to develop a model that can effectively classify the garbage automatically. To achieve our goal try to integrating pre-trained model with customized CNN model. DN169_Tuned provide highest accuracy (99.58%) among all of these model. Our model performed well aligning with our initial expectations in the research, highlights the crucial role of automated garbage classification, providing a superior solution and advancing environmental practices.

VII. FUTURE SCOPE

In future using a federated learning technique this model can be integrate into mobile devices. The goal is to build a automated system for sorting garbage in rural areas, making it easier for people, and speeding up the development of smart waste identification.

References

[1] M. Tanveer, S. A. R. Khan, M. Umar, Z. Yu, M. J. Sajid, and I. U. Haq, "Waste management and green technology: Future trends in circular economy leading towards environmental sustainability," *Environmental Science and Pollution Research*, vol. 29, no. 53, pp. 80 161–80 178, 2022.

[2] S. Kaza, L. Yao, P. Bhada-Tata, and F. Van Woerden, *What a waste 2.0: a global snapshot of solid waste management to 2050*. World Bank Publications, 2018.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[4] S. Meng and W.-T. Chu, "A study of garbage classification with convolutional neural networks," in *2020 Indo–Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN)*. IEEE, 2020, pp. 152–157.

[5] H. D. Poojary, N. M. Shinde, A. K. Singh, V. Kothuru, S. B. Kshirsagar, and H. Agarwal, "Classification of garbage for robtic system using deep learning techniques," in *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, 2022, pp. 1345–1348.

[6] Y. Wang, G. Zhao, and X. Li, "A weight pruned convolutional neural network for garbage image classification," in *2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 2022, pp. 1–6.

[7] L. Cao and W. Xiang, "Application of convolutional neural network based on transfer learning for garbage classification," in *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*. IEEE, 2020, pp. 1032–1036.

[8] S. Bhattacharya, K. B. Sai, S. Harikrishnan, H. Peera, G. Jyothi *et al.*, "Automated garbage classification using deep learning," in *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*. IEEE, 2023, pp. 404–410.

[9] I. Joshi, P. Dev, and G. Geetha, "Garbage classification using deep learning," in *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*. IEEE, 2023, pp. 809–814.

[10] D. Wang and Z. Wang, "Improved method of garbage classification based on deep learning," in *2021 International Conference on Computer Network, Electronic and Automation (ICCNEA)*. IEEE, 2021, pp. 126–132.

[11] T. Gupta, R. Joshi, D. Mukhopadhyay, K. Sachdeva, N. Jain, D. Virmani, and L. Garcia-Hernandez, "A deep learning approach based hardware solution to categorise garbage in environment," *Complex & Intelligent Systems*, pp. 1–24, 2022.

[12] M. K. Hasan, M. A. Khan, G. F. Issa, A. Atta, A. S. Akram, and M. Hassan, "Smart waste management and classification system for smart cities using deep learning," in *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*. IEEE, 2022, pp. 1–7.

[13] N. Garg and S. Das, "An analysis on intelligent garbage detection using cnn architecture," in *2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*. IEEE, 2022, pp. 1–6.

[14] "Garbage dataset," https://www.kaggle.com/datasets/mostafaabla/garbage-classification, (Accessed on 12/10/2023).

[15] N. Haque, A. H. Efat, S. M. Hasan, N. Jannat, M. Oishe, and M. Mitu, "Revolutionizing pest detection for sustainable agriculture: A transfer learning fusion network with attention-triplet and multi-layer ensemble," in *2023 26th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2023, pp. 1–6.