Trash Classification using Deep Learning Models

Enes Faruk Keskin
Turkish Aeronautical Association University
Department of Computer Engineering
Ankara, Türkiye
enesfarukkeskin@gmail.com

Oğulcan İşleyen
Turkish Aeronautical Association University
Department of Computer Engineering
Ankara, Türkiye
isleyenogulcan@outlook.com

Hilmi Demirhan
University of North Carolina Wilmington
Congdon School of Supply Chain, Business Analytics, and Information Systems
Cameron School of Business
Wilmington, North Carolina, USA
demirhanh@uncw.edu

Şeref Recep Keskin
Logiwa WMS
Senior Machine Learning Engineer
Chicago, USA
seref.keskin@logiwa.com

***Abstract*—The rapid growth of population and urbanization, combined with transformative technological advancements arising from the industrial revolution, have resulted in a divergence in living standards and consumption patterns. Consequently, a substantial increase in waste generation occurred compared with earlier times. This has given rise to severe air, water, and soil pollution, posing a grave threat to all forms of life. Moreover, it has accelerated the depletion of natural resources and has exacerbated the challenges associated with climate change. Given these circumstances, effective classification of recyclable waste is a crucial objective for humanity. To accomplish this, the utilization of Deep Learning models has proven highly beneficial. In order to achieve this objective, we conducted a thorough investigation where we evaluated prominent Deep Learning models to determine the most effective approach. The DenseNet121, MobileNet, ResNet50, and Xception architectures were employed in this study using the TrashNet dataset. After conducting the experiments, we found that the DenseNet121 model, fine-tuned specifically for this task, yielded the best results with a test accuracy rate of 95%.**

***Index Terms*—Trash Classification, Image Recognition, Deep learning, Convolutional Neural Networks**

I. INTRODUCTION

The combination of population growth, urbanization, and technological advancements resulting from the industrial revolution has led to a significant increase in waste generation compared to previous times. This has caused a differentiation in living standards and consumption patterns. However, this surge in waste production has also led to detrimental effects on the environment, including air, water, and soil pollution. These environmental issues pose a threat to all living organisms and contribute to the depletion of natural resources and to climate change.

The fact that waste is not used in the recycling and recovery processes causes serious economic and energy losses. The method used today in the selection process of recycled materials is based on the decomposition of these materials manually. This separation is sometimes performed by separating the waste bins of the recycled materials, and sometimes the garbage is separated one by one by the people in charge. This situation does not ensure that we are at the desired point in saving recycled materials. Great economic loss is experienced because of people's unconscious behavior or carelessness.

- By recycling 1 ton of waste paper, we can prevent the cutting down of 17 trees and save up to 12,400 m3 of greenhouse gas emissions. Additionally, recycling paper helps save 2.4 m3 of waste storage space.
- Metal and plastic recovery can lead to significant energy savings, with up to 95% energy reduction compared to new production methods.
- Recycling 1 ton of glass can save approximately 100 liters of oil, contributing to conservation efforts and reducing the reliance on fossil fuel resources.
- Waste glass can be recycled and transformed back into glass products. Similarly, plastics have the potential to be repurposed into various materials, including fibers and filling materials. Additionally, waste metals can be recycled and used to produce new metal products.
- Organic waste can be composted to produce nutrient-rich compost, then it can be used to improve soil fertility and productivity. By utilizing compost obtained from organic waste, we can improve the health and productivity of our soils.

In light of this information, in our country and in the world, it is necessary to use a technology that is beyond the traditional methods to prevent waste by detecting recyclable materials, to use resources efficiently, reduce the amount of waste generated, and establish more effective waste collection systems.

Setting up a model definition or a machine learning system using traditional machine learning techniques is a time-consuming task. These techniques often require expert assistance to preprocess raw data before they can be effectively utilized. However, deep learning has made significant advancements in addressing this issue. Deep neural networks have the capability to directly process raw data, enabling the learning process to be performed on the original, unprocessed data.

979-8-3503-2759-5/23/$31.00 ©2023 IEEE
DOI 10.1109/CSCE60160.2023.00055

Deep Learning is a category of machine learning that employs artificial neural networks, drawing inspiration from the functioning of our brains [1]. It enables computational models to understand and extract meaningful representations from complex datasets by incorporating multiple layers of processing. By utilizing a rewriting algorithm, Deep Learning allows machines to adjust the internal parameters of each layer's representation based on the information derived from preceding layers.

The main objective of Deep Learning is to explore and comprehend the intricate patterns and relationships present in large datasets. This is achieved by iteratively modifying the internal parameters to improve the accuracy and efficiency of the model's predictions. By leveraging the hierarchical nature of neural networks, Deep Learning models can learn and extract increasingly abstract features and representations from the data.

In this study, we utilized datasets containing images of individual objects set against a clear white background. Employing Deep Learning techniques, we aimed to classify these images into six distinct categories representing different types of rubbish. By leveraging the power of image-based predictions, we can accurately determine the category to which an object belongs based solely on its image representation [2], [3].

Although there have been numerous image classification projects utilizing neural networks in recent years, there is a dearth of research specifically focusing on Rubbish classification using neural networks. Therefore, our objective was to contribute to the field of recycling by employing computer vision and Deep Learning techniques to improve the process of sorting out recyclable materials. In our previous work, we used the same dataset and applied DenseNet121, Xception, MobileNet, ResNet50 models [4]. In this study, we applied various data augmentation techniques and explored different deep learning models. Additionally, we experimented with different numbers of epochs and made modifications to the hyperparameters of the deep learning models, which differed from our previous research.

The dataset utilized in our study was obtained from research conducted by students at Stanford University [2].

II. LITERATURE REVIEW

In 2016, Gary Thung and Mindy Yang conducted a project on trash sorting [5], which served as the foundation for our own project. Thung and Yang created the TrashNet dataset, which we utilized in our project. Both projects shared a common goal: to classify single pieces of garbage or recyclables into six different categories that includes glass, plastic, cardboard, metal, paper, and trash materials.

Thung and Yang used Support Vector Machine and Convolutional Neural Networks (CNN) techniques separately and compared their results. They developed the eleven-layer CNN architecture, adjusting the number of filters in certain convolutional layers to address computational limitations. Additionally, they utilized methodologies including learning rate decay, Kaiming weight initialization and L2 regularization.

However, their results with the CNN approach were not satisfactory. While the SVM approach attained a testing accuracy of 63%, the CNN method only managed to reach a testing accuracy of 22%. Thung and Yang attributed this lower performance to suboptimal hyperparameters, which hindered the ability of the CNN to learn effectively.

Another paper by Olugboja Adedeji and Zenghui Wang [6] introduces an intelligent waste material classification system. The system utilizes a pretrained 50-layer residual network (ResNet-50) Convolutional Neural Network (CNN) model as a feature extractor, along with Support Vector Machine (SVM) for waste classification into different categories such as glass, metal, paper, and plastic.

The proposed system was evaluated using the trash image dataset developed by Gary Thung and Mindy Yang. It achieved an impressive accuracy of 87% on this dataset. By employing this waste material classification system, the separation process of waste can be made faster and more intelligent, reducing or eliminating the need for human involvement.

In 2020, Maoguo Shi, Qinyue Gu, and Yujie He conducted a project [7] where they focused on utilizing Convolutional Neural Networks (CNN) to achieve their objectives. They explored several well-known CNN architectures at the early stages of their research. Eventually, they arrived at a modified version of the AlexNet architecture by removing two layers. They conducted experiments using this modified architecture and incorporated techniques such as dropout, data augmentation, and learning rate decay.

In their study, Shi, Gu, and He experimented with two classifiers, Softmax and SVM, as the final layer of their model structure. They achieved 79.94% as their highest test accuracy using a combination of with partial data augmentation and the SVM classifier.

A. What will be our originality/contribution/difference?

The all papers/projects are done by using old convolutional neural network (CNN) models. For example, ResNet152 and Inception V3. While some of the papers were focused on modifying the structure of established architectures and conducting training from the ground up. Some of them focused on transfer learning (pre-trained) models. In the end of day, they all got the result (test accuracy) from their models. At this point, our aim is to exceed the success of these academic papers. We will use up-to-date pretrained image classification models. For example , ResNet50, MobileNet and up-to-date versions of DenseNet. Then, we will apply the our fine tuning on this pretrained models. In this way, we will fine tuning on a general purpose pretrained model and make it serviceable for our own project. In this manner, we aim to achieve more success from other works.

III. DEEP LEARNING MODELS

In this section we explain the details of transformer architectures, deep learning models used.

A. MobileNet

The MobileNet model incorporates depth-wise separable convolutions, initially utilized in the Inception model, to reduce the computational overhead in the early layers. The architecture of MobileNet revolves around depth-wise separable convolutions, except for the first layer, which employs full convolution. Following the full convolution, MobileNet applies depth-wise separable convolution. This design allows high accuracy rates to be achieved while using a minimal number of hyperparameters. MobileNet has proven to be a valuable model owing to its ability to be trained faster and with fewer resources. [8].

B. Xception

Xception, also known as Extreme Inception, is a deep learning model recommended by the Google research team. It draws inspiration from the Inception architecture but introduces a novel approach to feature extraction using depth-wise separable convolutions. By leveraging depth-wise separable convolutions, Xception enhances the learning capacity of the network while reducing the computational complexity. It employs a deep architecture composed of independently processed depth-wise separable convolution blocks, thereby enabling more effective feature extraction. The Xception model serves as a powerful tool for tasks such as complex image classification, object recognition, and various computer vision problems, thereby achieving high accuracy rates. Moreover, it accelerates the training process and utilizes resources more efficiently owing to the reduced number of hyperparameters and computational requirements. Xception is widely recognized as a robust deep learning model for a range of computer vision applications. [9].

C. Densely Connected Convolutional Networks

Densely Connected Convolutional Networks (DenseNet) are deep learning architectures that introduce dense connections between the layers. As proposed by Huang et al., DenseNet aims addressing the vanishing gradient problem and by promoting a better information transmission throughout the network. In DenseNet, each layer receives direct inputs from all the preceding layers, resulting in densely connected feature maps. This design fosters feature reuse and facilitates gradient propagation, thereby enabling deeper architectures with fewer parameters. DenseNet consists of multiple dense blocks, where each block consists of a series of convolutional layers, followed by concatenation of feature maps. By exploiting dense connections, DenseNet were able to achieve state-of-the-art performance on different computer vision tasks while being more parameter efficient. It also encourages feature extraction from multiple abstraction levels, leading to richer representations and improved model accuracy. DenseNet has gained significant attention and is widely used in the deep learning community owing to its effectiveness and interoperability. [10].

D. ResNet50

ResNet50, a Residual Network with 50 layers, is a deep learning architecture that revolutionizes image classification tasks. Introduced by He et al., ResNet50 addresses the challenge of training very deep neural networks by introducing residual connections. The key innovation of ResNet50 lies in its residual blocks, which enable the network to learn residual mappings rather than explicitly trying to fit the desired underlying mapping. These residual connections allow for the direct flow of information from one layer to another, mitigating the vanishing gradient problem and enabling the training of significantly deeper networks. ResNet50 consists of multiple residual blocks, where each block contains a set of convolutional layers along with skip connections that add the original input to the block output. This helps to preserve valuable information during the forward pass. With its deep architecture and skip connections, ResNet50 achieves remarkable accuracy in image classification tasks, even surpassing the human-level performance in some cases. It has become a popular choice for various computer vision tasks and serves as a foundation for many state-of-the-art deep learning models. [11].

IV. EXPERIMENTAL RESULTS

This section provides a comprehensive overview of the experiments conducted, including the methodologies employed, the results obtained, and how the data preprocessing implemented. Additionally, the dataset used in the study is described, along with the methods employed to obtain vector representations of the dataset.

The experiments took place on a computer powered by the macOS operating system, housing an M1 processor with 16GB of RAM.

A. Dataset

Dataset consists of photographs taken by students at Stanford using a white banner in the background [3]. 6 of the photos taken from the dataset are shown in the figure below. One photograph of each class and a total of 6 photographs were chosen to represent 6 classes. The dataset contains images of recycled objects as six classes. These are: glass, paper, cardboard, plastic, metal and trash.

Currently there are 2527 data (images) in the dataset:

- 501 glass
- 594 papers
- 403 cartons
- 482 plastic
- 410 metals
- 137 trash

In this study, 70% of the total images were employed for training, 17% were allocated for testing, and the remaining 13% were designated for validation.

B. Results

In this study we used Google Colab, Keras library with TensorFlow. We run experiments using Google GPU services.

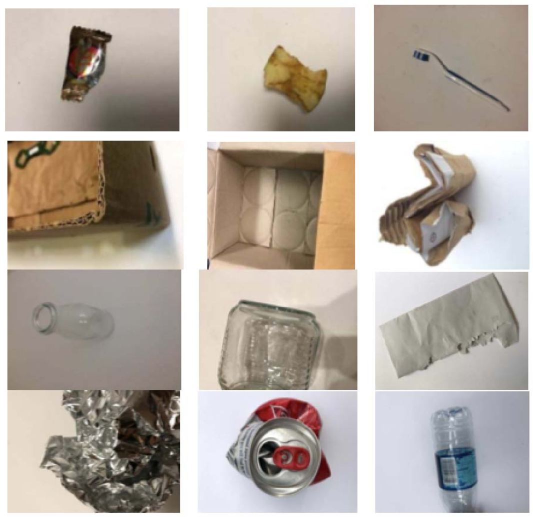

Fig. 1. Example Images [11]

We conducted fine-tuning experiments on several models by adjusting the pretrained model's weights using the ImageNet dataset. Following an initial step of pre-training the ImageNet trained weights that contains the model. Fine-tuning was carried out utilizing stochastic gradient descent alongside a Nesterov momentum of 0.9, while the learning rate of 0.0001 was applied specifically to DenseNet121. This was because we needed to achieve our aim. During the training process of all models, we employed basic data augmentation techniques, including vertical and horizontal flips, as well as 15° rotations. We utilized the Adam optimizer, with a learning established at 0.001.

During the training experiments on the DenseNet121 models, a batch size of 8 was utilized. For the remaining training experiments, a batch size of 32 was employed. Additionally, for all of our deep learning models (DenseNet121, Xception, MobileNet, ResNet50), the input size was set to 224 × 224.

For our experiments, we trained the models from scratch using dedicated training data, and incorporated validation during the training process. The weights obtained from this training were used to conduct the experiments. The details of our experiments along with the corresponding results are presented in Figure 2. This allowed us to showcase the performance and outcomes of our models in a clear and concise manner, thereby providing valuable insights into the effectiveness and capabilities of our approach.

Model	Epoch	Test Accuracy
ResNet50	100	76 %
MobileNet	100	89 %
Xception	100	83 %
DenseNet121	10 + 100	95 %

Fig. 2. Fine Tuned Models Results

When we are working on the training process of all the models, we applied straightforward data augmentation techniques. The motivation behind the implementation of data augmentation was the limited availability of images in the dataset. While we observed some instances of overfitting owing to data augmentation in certain experiments, overall, data augmentation proved to be highly effective in enhancing the performance of the models. It helped diversify the training samples, enabling the models to learn robust features and generalize better to unseen data. The careful application of data augmentation plays a important role in improving the overall performance and mitigating the challenges posed by the limited dataset size.

The three algorithms that achieved the highest success were MobileNet, Xception, and DenseNet121.

MobileNet model was applied to 100 epochs. We achieved 89% test accuracy. When we are working on the analysis of the result, we realized that the validation test results were stable. It was realy close to our aim (93%) but, we could not reach the 93by using the MobileNet model. We tried some other fine-tuning and we tried to performed training after pre-training with some epochs; however, this did not work.

The accuracy graph for the fine-tuned MobileNet model is shown in Figure 3.

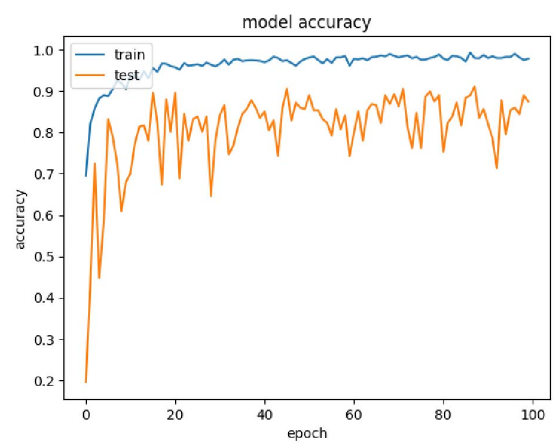

Fig. 3. Fine-tuned MobilNet model's Test Accuracy: 94%

The Xception model was trained for 100 epochs, resulting in a test accuracy of 83%. Upon analyzing the accuracy results, we observed that the validation accuracy continued to improve beyond the 70th epoch. Based on this evaluation, we decided to extend the training phase by an additional 100 epochs in the subsequent experiment. However, despite the extended training, we were unable to achieve the desired results.

The accuracy graph for the fine-tuned Xception model is shown in Figure 4.

We performed our last experiment using Densenet121 model. We conducted the training utilizing 100 epochs. We reached the 89% accuracy.. We decided to do pretraining with 10 epochs. Contrary to the Xception model, our test accuracy

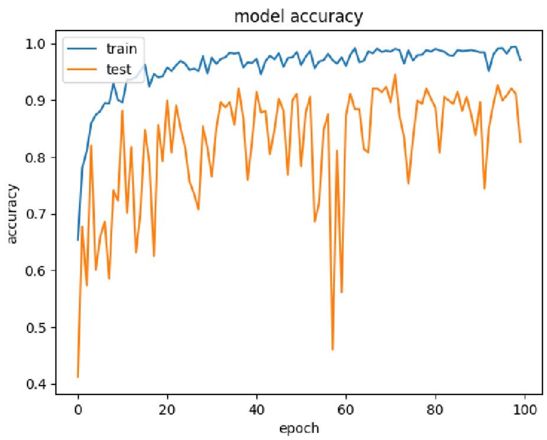

Fig. 4. Fine-tuned Xception model's test accuracy: 94%

increased to 95% when this operation was performed. As a result, we achieved 95% test accuracy with DenseNet121 10 (pretraining) + 100 epochs. We achieved the best test results via fine-tuning the models.

Fine-tuned DenseNet121 model were shown in Figure 5.

Fig. 5. Fine-tuned DenseNet121 model's Test Accuracy: 94%

Fig. 6. Comparison of Deep Learning Models

V. Discussions and Conclusion

This study involved experimentation using well-known CNN models. The results of these experiments demonstrate varying levels of test accuracy for the deep learning models used. Initially, our project plan encompassed the use of three distinct models. However, these models do not meet the desired level of anticipated accuracy. Consequently, we incorporated two additional models, namely, Xception and DenseNet121, into our research framework.

The performances of the various models were evaluated in terms of test accuracy during the course of the study. The ResNet50 model achieved a test accuracy of 76% after 100 epochs, which, although reasonable, did not meet the desired level of accuracy. After 100 epochs, the MobileNet model attained the highest test accuracy of 89%, which still did not satisfy our objectives. Consequently, we introduce two additional models for our experiments. Subsequently, the Xception model exhibited an improved performance with a test accuracy of 83% after 100 epochs. However, the DenseNet121 model outperformed the others, reaching a remarkable test accuracy of 95% after 100 epochs. The detailed outcomes of the fine-tuned models are shown in Figure 2 for reference.

The fine-tuned Densenet-121 model was the most successful in terms of the test accuracy. During the optimization process, it was observed that the Adam optimizer consistently yielded higher test accuracy than the other options. Consequently, the Adam optimizer was utilized across all training sessions, ensuring consistent and improved performance throughout the experiments.

The results obtained from the experiments highlight the potential of deep learning algorithms for classifying recyclable waste. Through a series of experiments on known deep-learning models, we explored their suitability for this purpose. Nevertheless, it should be noted that when applied to real-time systems, the accuracy of deep learning models for classifying recyclable waste may be compromised owing to factors such as limited data availability and challenges associated with inconsistent backgrounds in real-world images.

References

[1] J. Brownlee, "What is deep learning?" 2016. [Online]. Available: https://machinelearningmastery.com/what-is-deep-learning/

[2] M. Yang and G. Thung, "Classification of trash for recyclability status," *arXiv preprint*, 2016.

[3] G. Thung, "Trashnet," https://github.com/garythung/trashnet, 2016.

[4] R. A. Aral, R. Keskin, M. Kaya, and M. Hacıömeroğlu, "Classification of trashnet dataset based on deep learning models," in *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 2058–2062.

[5] G. Thung and M. Yang, "Classification of trash for recyclability status," 2016.

[6] O. Adedeji and Z. Wang, "Intelligent waste classification system using deep learning convolutional neural network," *Procedia Manufacturing*, 2019.

[7] M. Shi, Q. Gu, and Y. He, "Trash classification using convolutional neural networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 4138–4147, 2020.

[8] A. G. Howard, M. Zhu, B. Chen, and H. Adam, "Mobilenets: Applying mobilenets to real-world applications," *arXiv preprint arXiv:1708.06134*, 2017. [Online]. Available: https://arxiv.org/abs/1708.06134

[9] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258. [Online]. Available: http://openaccess.thecvf.com/content_cvpr_2017/html/Chollet_Xception_Deep_Learning_CVPR_2017_paper.html

[10] G. Huang, Z. Liu, K. Q. Weinberger, and L. van der Maaten, "Densenet: A resource-efficient deep convolutional neural network for visual recognition," *CoRR*, vol. abs/1707.07870, 2017. [Online]. Available: http://arxiv.org/abs/1707.07870

[11] B. Xu, Y. Wang, and W. Liu, "Residual networks and image classification," *CoRR*, vol. abs/1603.08029, 2016. [Online]. Available: http://arxiv.org/abs/1603.08029