Eurasian J. Sci. Eng., 2024, 10(2), 165-180
https://doi.org/10.23918/eajse.v10i2p12

Published by Tishk International University
Available at: https://eajse.tiu.edu.iq/submit/

RESEARCH ARTICLE Open Access

A Deep Learning-Based Approach to Garbage Detection in Urban Centers

Slvar Abdulazeez Arif [1]*, and Abubakar M. Ashir [2]

[1]Department of Computer Education, Faculty of Education, Tishk International University, Erbil, Iraq.
[2]Department of Computer Engineering, Faculty of Engineering, Tishk International University, Erbil, Iraq.

Article History
Received: 05.06.2024
Revised: 29.07.2024
Accepted: 27.08.2024
Published: 29.08.2024
Communicated by: Dr. Orhan Tug
**Email address:*
slvar.abdulazeez@tiu.edu.iq
**Corresponding Author*

Abstract:
The proposed method is an innovative way that address environmental concerns through the integration of artificial intelligence (AI) and image recognition technology. The application leverages the powerful Residual Network 101(ResNet101) model, trained on a dataset of over 3000 images, to accurately classify and analyze environmental issues depicted in user-submitted images. By utilizing the user's location data, the proposed method enables users to report environmental problems effectively, such as pollution, waste accumulation, deforestation, and more. For testing and practical applicability of the proposed approach, a mobile applicaton is developed to provide a user-friendly interface that allows users to capture images of environmental issues and submit them with relevant details. The images are then processed by the ResNet101 model, which employs deep learning techniques to classify and provide insights into the severity of the reported problems. The model achieved a low training loss of 0.0553 and a high training accuracy of 0.9809, indicating that learned well from the training data. The validation set's metrics, with a validation loss of 1.7773 and a validation accuracy of 0.6566, that show the model's performance on unseen data was relatively lower compared to the training set. However, the model's ability to generalize to new data is demonstrated by achieving a test loss of 0.8396 and a test accuracy of 0.8602 abstracts.

Keywords: *Garbage Detection: Deep Learning Computer Vision: ResNet, AI.*

1. Introduction

Waste collection is a big challenge that requires innovative ways to address it. One of the ways is developing a mobile application that can help locate waste in public places. The need for such an application is growing as a result of the increase in waste production, especially in metropolitan areas caused by population growth over the last few decades. Waste production has also significantly increased in urban areas largely due to the rapid growth of consumable products by the public. For instance, in a city like Sulaymaniyah-Kurdistan Region Government (KRG), researchs have shown that 1,200 tons of waste were produced daily in the city center in 2016, averaging about 2,400 tonnes produced daily across the larger Sulaymaniyah governorate [1]. A simple calculation reveals that the yearly waste production in the city exceeded one million tons, which projects to about 31,000 tons of waste daily at a rate of 1.4 kg of waste per capita for the entire Iraq in the same period, implying a yearly waste production of approximately 11 million tons [2].

One major challenge in the process of waste collecting and cleaning is waste identification. With the mushrooming residential places in urban areas, it is becoming increasingly difficult for waste-collecting agencies to identify and locate household waste disposal points. It wastes time for the cleaning agencies to track down all waste collection points and may require a large number of staff, which translates to more costs by the agency, eventually passing down to the general public to shoulder the burden. Additionally, waste collection in the traditional way does not guarantee that all waste can be collected and done on time.

Wastes that has not been collected on time and disposed of properly cause a lot of health issues. Such as waste releases toxic substances into the environment, polluting waterways, the air around, and even the surrounding land. This changes the abiotic factors of an ecosystem, adversely affecting the plant and animal habitats and causing drastic lifestyle changes for flora and fauna in the system. For instance, plastic waste is known to be one of the leading causes of animal deaths. Besides releasing harmful substances into the environment. Wastes also attract rodents, roaches, flies, and other disease-causing animals and insects. This can be very harmful to those people, who are living around waste collection points. Aside from the potential dangers to their health, the people who reside in nearby buildings will be able to smell the waste as if it were in their homes [2].

Economical growth dangers brought about by improper waste management abound. Potential foreign investors may decline to invest in a city or country for fear of health risks to employees, which may come in the form of high health insurance premiums and other health-related costs. Lack of investment slows down the economic growth of a city and may hamper the overall development. Therefore, investment is crucial for developing countries like Iraq to spur economic growth that leads to overall development. Economic growth and development are possible and can be achieved much more easily by cities that have a well-managed waste management system, where investors do not have to worry about their health and that of their employees.

The world population is projected to be more than 9.5 billion people by 2050, and more than 80% of them will live in cities, up from 55% currently. If waste management systems are not properly developed by then, ultimately, people will have no choice ,but they can construct their houses out of junk since the quantity of waste accessible will have risen and will continue to climb at a rapid pace. However, the development of an application that can make cleaning quicker, easier, and ideal to approach, people would not have to resort to such desperate measures.

Since the application can be operational on a mobile gadget, people can easily share information with cleaning agencies without wasting time and money. Therefore no waste would be left uncollected at any time the agencies go to any particular area in the city, as they would know exactly where the waste collection points are, which ones may contain more waste than others, and such other data. This makes the process of reporting waste more widespread. Because of this, it would now be much simpler for individuals to participate in the protection of the environment without being forced to depend on the government [3]. This research therefore aims to develop an application that would be used by common citizens to report waste collection points anywhere in the country, and the following are the specific research objectives:

1. To design, develop, and implement a deep-learning application that can effectively detect and classify waste in urban areas.
2. To assess and evaluate the effectiveness, reliability, and accuracy of the developed application in a real-world city, taking into account all natural conditions, such as different weather patterns.

2. Literature Review

Researchers from Bangladesh University of Engineering and Technology proposed a system for autonomous mobile trash collectors [3]. The system uses a mechanism called the collecting mechanism to transport trash to a container [3]. Previously, this process was implemented using an ultrasonic sensor and object detector to find objects within a range of 30 centimeters [3]. When an object is detected, the robot captures an image for trash identification, which is then processed by a deep learning algorithm, specifically a Convolutional Neural Network (CNN), to identify items such as papers, bottles, and cans [3]. The model achieved an accuracy of 98% during training and 96% during testing [3].

Researchers examined the efficacy of three CNN models (Mask Region-based Convolutional Neural Network (Mask R-CNN), Segmenting Objects by Locations (SOLO), and You Only Look Once version 6 (YOLOv6)) to detect the amount of trash accumulation on roadsides using vehicle-mounted cameras [4]. Additionally, trash volumes, visual assessments Optimized Vector Transformation Algorithm (OVTA), and photographs (amounting to about 3,800 in total) for 84 roads were collected and evaluated [4]. The outcomes for the Mask Region-based Convolutional Neural Network (R-CNN) model were 95% recall, 92% precision, and 87% accuracy during the training phase, while in the validation phase, they were 91% recall, 84% precision, and 77% accuracy [4]. The SOLO model achieved 78% recall, 82% precision, and 66% accuracy in training, with validation results of 81% recall, 60% precision, and 52% accuracy [4]. The YOLOv6 model's training results were 75% recall, 80% precision, and 65% accuracy, while validation results were 70% recall, 79% precision, and 59% accuracy [4].

Developers developed a system that can differentiate between dry and wet waste solely based on a taken picture of the waste, The system used machine learning and the TensorFlow platform to detect and classify the contents of waste images taken from an Android device into Bio-degradable and Non-Bio-degradable waste, as well as the amount of banned plastic content. Then, the results have been sent back to the Android device for analysis. They used a dataset of 2700 images, and the system used the version of the Inception-V3 algorithm which is referred to as V3 with Batch Normalization (BN) auxiliary includes a fully connected layer in the auxiliary classifier which has been normalized [5].

Up to 2700 pictures were used to train the model, and the accuracy of the system's ability of classification was 83.3% in 1.3 seconds. Three parameters were used for testing the model, and the parameters were color, brightness, and pureness [5].

Investigators presented a system with an automatic ability for waste detection, which can recognize and categorize different kinds of waste, for instance, trash, glass, plastic, metal, cardboard, and paper. In the training, the total taken time was 17 hours and 3 minutes [6].

For the output, Real-Time Object Detection with Region Proposal Networks (Faster RCNN) was chosen for creating mother-class and bounding box [6] .

For the input of the Convolutional Neural Network (CNN), the real image was trimmed. For creating the final classification of the image, the child class was the out of the CNN, and it was combined with the mother class. Then the output of the CNN was used for predicting the object boundaries before model detection. For analyzing six groups of waste materials, the researcher applied the Faster R-CNN [6].

The author employed the Faster R-CNN to identify six categories of waste material, after 17 hours 91% of accuracy rate has been achieved with 200000 steps of epochs for the neural network. By this as a result almost 8.05 seconds took as processing time per prediction [6].

Developers present a real-time garbage detection system based on the Convolutional Neural Network (CNN) model by Closed-Circuit Television (CCTV) cameras. The improved CNN model utilized a dataset of 8 classes of trash consisting of more than 2100 images, with each class containing 300 images. These classes were masks, tissue papers, shoppers, boxes, automobile parts, pampers, bottles, and juice boxes. Videos of each class of trash from different angles were made, and each frame was extracted, to avoid the need to use data augmentation. The performance of the models was measured by accuracy and average weight loss, the original model achieved 97.2% accuracy and an average weight loss of 0.0930, while the improved model had 99.6% accuracy and an average weight loss of 0.6928. Detection time accuracy and average mean precision (mAP) were used to compare the performance of both models [7].

Researchers described a system for garbage detection and classification into organic waste, non-organic waste, and non-waste. The suggested system uses the Haar-Cascade technique to detect the presence of trash, the Gray-Level Co-Occurrence Matrix (GLCM) [8] to get the characteristics of textures that can be acquired from the statistics of gray intensity values in the image, and the Histogram of Oriented Gradient (HOG) which is a feature description used in computer vision to calculate the orientation of a gradient in a localized image. Before that, the object must be preprocessed by cropping and resizing to 32x32 pixels; after that, it will use Support Vector Machines (SVM) for classifying objects. The system was tested with 5-fold cross-validation, resulting in 82.7% accuracy for offline testing, and 63.5% accuracy for online testing with a camera tilted down to -40o, and minimum distance of 80 cm, and a maximum distance of 200 cm for detection. A Kalman-based tracker was applied to enhance the object detection stability. The results of the detection test on the waste image object yielded 93.76% detection accuracy [9].

The researchers used (Visual Geometry Group) VGG16, a convolutional neural network with thirteen layers for feature extraction from a dataset for multi-subcategory or multi-category classification. The neural network was built using Keras and TensorFlow. The features are extracted in each layer, becoming more generalizable, and after thirteen layers, the features are passed through the softmax function to determine the label of the input image [10].

A two-stage Waste Recognition-Retrieval algorithm (W2R) has been used by the researchers for classifying the waste; the Recognition Model (RegM) consists of the algorithm that can be recognized into 13 sub-groups. The sub-groups have been divided into 4 groups by the Recognition-Retrieval Model (RevM). Also, for the comparison, the one-stage Classification Model (ClfM) was trained. Both of them were installed on an automatic sorting machine; The RegM had an average accuracy of 93.80% ± 1.71, while the RevM had an average accuracy of 94.71% ± 1.69, which was better than the ClfM-VGG's accuracy of 69.66% ± 3.43 [10].

The Hybrid Transfer Learning method and a Generative Patch-based GAN, Generative Adversarial Network (GP-GAN) to blend images have been used by the researchers to describe a system for waste object detection and classification by “Waste Object Detection and Classification” [11].

For improving the quantity of the dataset via creating collages of multiple input pictures, three methods have been proposed by the researchers. theses methods have been putting pictures randomly, putting pictures at four quadrants and capturing the place for putting the images.

The Faster R-CNN network was well-tuned through a Region Proposal Network (RPN) that can distribute full-image features with having the detection network that allows proposals that can cost-free region.

The RPN is trained to generate high-quality region proposals, which are then used by Fast R-CNN for detection. The fine-tuned Faster R-CNN returned good object detection results with a learning rate of 0.0002 [11].

The designers have developed a deep learning-based small object detection and classification system for garbage waste management (DLSODC-GWM). The system consists of two processes: object detection and classification. An arithmetic optimization algorithm (AOA) with an improved RefineD (IRD) model is utilized for the detection of objects; while a functional link neural network (FLNN) technique is employed for waste classification. The dataset used comprises 393 images for cardboard, 491 images for glass, 400 images for metal, 584 images for paper, 472 images for plastic, and 127 images for trash. With 2000 epochs, the AlexNet, VGG16, and ResNet50 models resulted in accuracies of 52.50%, 73.10%, and 74.70%, respectively, while the Multi-Level Hierarchical Convolutional Neural Network (MLH-CNN) technique achieved a near-optimal accuracy of 92.60%. However, the DLSODC-GWM technique outperformed the other models with an accuracy of 98.61% [12].

This study assessed five object recognition models for the purpose of trash detection (EfficientDet-D1, Single Shot MultiBox Detector (SSD) ResNet-50 V1, Faster R-CNN ResNet-101 V1, CenterNet ResNet-101 V1, and YOLOv5M) [13], [14].

The models were chosen as they were successful in their respective families in terms of inference time and accuracy and were suitable for embedded systems. A Tesla V100 Graphics Processing Unit (GPU) was utilized in order to compare the efficiency of the distinct models, with a standard of 500 epochs. The input image size was set to 640 x 640 for EffcientDet D1, RetinaNet-50, and Faster R-CNN ResNet101 experiments and 512x512 for CenterNet ResNet101 experiments. The models were evaluated on 25 unseen trash photos, and YOLOv5M was the most successful, accurately detecting the most garbage. Other models like EfficientDet-D1, Faster R-CNN ResNet-101, and SSD ResNet-50 did similarly, but with fewer predictions of the amount of trash. The images used for testing included real-world scenarios of trash near garbage bins and on roads. All models were tested on the images, and the differences in accuracy were clear [12].

Researchers at the beginning proposed two new benchmark datasets, detect-waste, and classify-waste, to merge collections from open-source datasets with unified annotations covering all possible waste categories, namely: bio, glass, metal and plastic, non-recyclable, other, paper, and unknown. Two CNN models were evaluated with state-of-the-art models using precision, recall, F1-score, and accuracy metrics. The CNN1 model achieved better performance for automatic solid waste detection with 94% accuracy. The CNN1 model was trained with different learning rates and epochs with the RMSprop and Adam optimizers. The RMSprop optimizer yielded the highest accuracy of 90%, precision of 0.86 and 0.92 in class-0 and 1, respectively, recall of 0.88 and 0.91 in class-0 and 1, respectively, and F1-score of 0.87 and 0.91 in class-0 and 1 respectively. The Adam optimizer achieved the highest accuracy of 94%, precision of 0.91 and 0.96 in class-0 and 1, respectively, recall of 0.94 and 0.94 in class-0 and 1, respectively, and F1-score of 0.93 and 0.95 in class-0 and 1 respectively [16].

Researchers introduced two new benchmark datasets, detect-waste, and classify-waste, which are merged collections of existing datasets with unified annotations covering all possible waste categories, such as bio, glass, metal, and plastic, non-recyclable, other, paper, and unknown. The detect-waste dataset contains over 28,000 images and over 40,000 objects with a unified bounding box annotation and a single-label litter. The classify-waste dataset was used to train the classifier in a semi-supervised fashion, using around 55k unlabeled images from the Open Litter Map. Additionally, a two-stage detector for litter localization and classification is presented, using the Efficient Set-D2 and EfficientNet-B2 models. The proposed approach achieves up to 70% of average precision in waste detection and around 75% of classification accuracy on the test dataset [17].

The study proposed an automatic garbage detection system based on deep learning and narrowband Internet of Things. The system uses an improved You Only Look Once version 2 (YOLOv2) model for decoration garbage detection and recognition. The model is optimized and accelerated by lightweight processing, such as target box dimension clustering and classification network pre-training, and is ported to an embedded monitor terminal. Narrowband Internet of Things can communicate among thousands of monitor terminals and background server centers. Experiments show that compared with the traditional system, the system has lower cost and better performance with a precision rate of 89.1%, a recall rate of 87.9%, and a frame rate of 42 f/s, which can effectively save manpower and material resources [18].

The authors purposed to distinguish images into two distinct categories, and they have attempted to utilize Convolutional Neural Networks (CNNs). CNNs use convolutional layers to discern the features from the images, followed by fully connected layers to label the images. They utilized TFlearn, a modular and transparent deep learning library built on top of Tensorflow, for training and prediction. They configured Tensorflow to train the network on a Tesla K40c, which has a memory of 12GB. They used CUDA Toolkit version 7.5 and cuDNN v5.1 internally. Given an image as input for prediction, the model binds two probabilities to it (since they have two categories). Each of these probabilities is a measure of the affinity of that image towards both categories. The accuracy is calculated according to a formula. The accuracy increases gradually after every epoch and reaches a peak of 83% to 85%. They get intermittent peaks up to 87%. This intermittent peak in accuracy is advantageous to them as the neural network saves the model only when it reaches a new maximum accuracy. Each epoch takes approximately 14 seconds [20].

Researchers present GarbNet, a convolutional neural network (CNN) model designed to classify garbage. The model is trained on the newly developed Garbage In Images (GINI) dataset and achieves a mean accuracy of 87.69%. Optimizations have been proposed to reduce the memory usage by 87.9% and the prediction time by 96.8%, while maintaining the accuracy of the model. The paper also introduces a new annotated GINI dataset and uses a pre-trained AlexNet model to initialize the weights of GarbNet. The model is trained on Caffe using Nvidia TitanX and tested using 5-fold stratified cross-validation. Additionally, the convolutional layers of the network are converted into fully convolutional layers to reduce the time for feedforward computation, by allowing the entire image to be processed in a single pass. The experiments conducted on the GINI dataset show a significant reduction in both the memory and time requirements, with no reduction in accuracy [21].

3. Methodology

3.1 Introduction

The study aims to develop an application for identifying and reporting waste with a ResNet 101 to classify the waste. This section therefore, explains the process of developing the application and the techniques involved in the development process. The section also explains the model used to develop the application, complete with expected reporting and classifying accuracy level.

The ResNet101 deep learning network is specifically used for categorizing waste and has been trained on a big dataset of waste for consistency and accuracy in reporting and for the ability to classify waste correctly. As shown in figure 1, in building an application, the first step is collecting a dataset and then organizing it, after that, interpreting the images.

The model in this research is primarily designed for classifying waste, so will be trained for it after the data presentation is done. The model will be arranged to process the images which will be stated by the mobile application, after the model reaches high accurancy. For developing the mobile application, flutter has been used for achieving some of the applications requeriments.

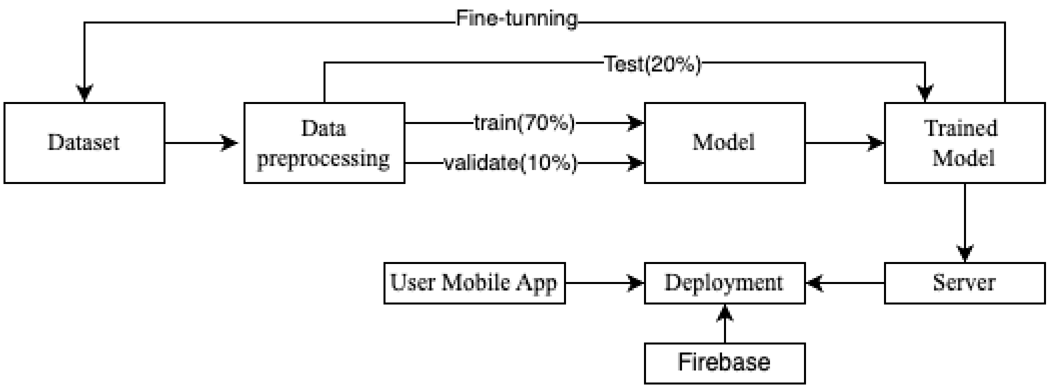

Figure 1: The full flowchart.

3.2 Data Collection

It is necessary to find and collect a large amount of data which should be suitable for training the model. This process is achieved by getting the data from websites such as Kaggle, Garbage in Image (GINI) [17] dataset, and other resources over the internet. This collected dataset consists of over 2500 images of various locations, which are either clean or dirty. The images are meant to be utilized in a model for training. The assortment of clean and dirty images from various places will give the model an extensive range of data to learn from, thus allowing it to identify clean and dirty surfaces in varying settings. Additionally, data has to be cleared from any low-resolution blurry images, and ensure that the images are appropriate to train the model. Then, three separate folders will be created: a training folder, a validation folder, and a testing folder, as shown in fig. 2 below.

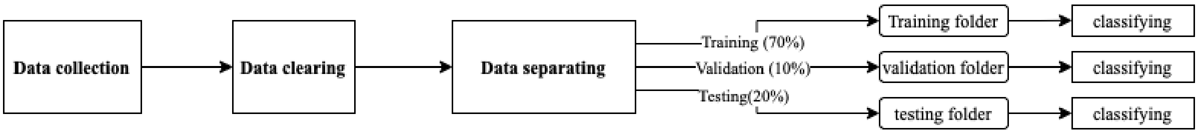

Figure 2: A flowchart of the data preparation.

Each of the folders will contain 4 levels of waste images depending on the amount of waste displayed. Examples of the dataset can be seen in the table below.

Table 1: An example of the data

LEVEL-1	LEVEL-2	LEVEL-3	LEVEL-4
			
			

3.3 Data Preprocessing

The processing started with image pre-treatment because it guarantees work efficiency and accuracy for the waste classifying model. In this process, the images were sized to unified dimensions and amount, regulating the clarity, color range, and contrast. Then, they were augmented by Python with the Pillow library for applying transformations like rotating, cropping, flipping, and resizing to a desired image. Further image augmentation using the Imaging library was applied to enhance and or minimize blurring, sharpening, noise, and color jittering. In Figure 3, an example of Image augmentation is shown.

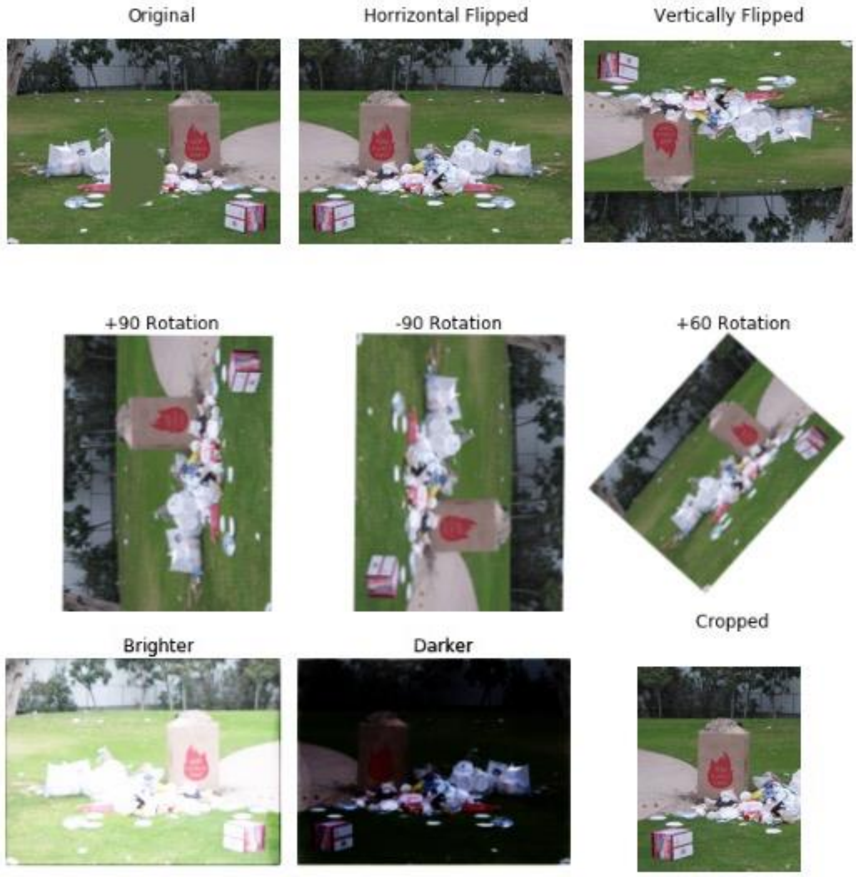

Figure 3: Image augmentation example.

To make sure the model is able to quickly and exactly analyze the images and give accurate predictions. These steps are important.The model will be more strong and capable to process a diversity of input images. By properly preprocessing the images.

3.4 ResNet Convolutional Neural Network (CNN)

The CNN model is an advanced type of deep-learning neural network that is commonly used for the identification and classification of images. As it has several coats, each one of them is responsible for different parts of the process of image analysis. The first layer is the curve layer, which extracts features from the image by going through a sequence of filters. The next layer is the grouping layer, which decreases the quality of the data by sampling the inputted image. Finally, the focussed layar collects the features from the previous two layers and classifies the image. This is broadly used in image processing, computer vision, speech recognition, and natural language. So, its ability to learn from data and generate exact results made it the standard choice in many areas. The accuracy overpasses the old models and is the classic for a lot of machine learning tasks. Resnet is a convolutional neural network architecture that is used for image classification problems. Because it is aimed to capture the shades of the image and it is contents, so that is why it is well-matched for classifying the waste. for making it useful for classifying the waste, it can be used for identifying objects in images along with classifying them into particular groups. For instance, classifying the images according to the total of the shown wastes.

For image classification and recognition Microsoft Research (MSR) in 2015 developed a CNN architecture which is ResNet101. Which is based on the ResNet architecture and has resulted in effective applications in different image datasets. For instance CIFAR-10, ImageNet, and Common Objects in Context (COCO), achieving top performance. To help the network learn more complex features, the core idea behind the ResNet architecture is to practice short connections. An identity mapping is added to meet this, as it allows the network to exclude specific layers. It helps to minimize the vanishing gradient, which is a common problem in deep learning networks. ResNet employs various layers for the purpose of identification and extraction of features from the input data. These layers include batch normalization layers, rectified linear unit (ReLU) layers, and convolutional layers. The expected output is then generated when convolutional layers' output goes through completely connected layer(s). The ResNet architecture is widely used for classifying images, detecting objects, and other related computer tasks [20].

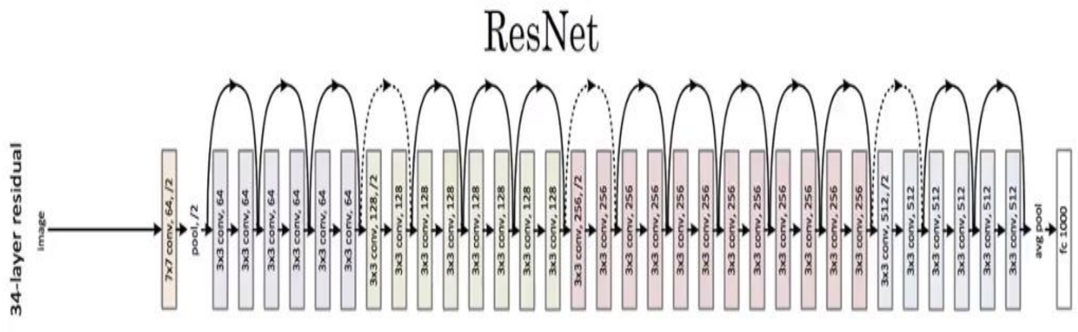

Figure 4: The 34-layer ResNet Architecture.

ResNet101 uses skip connections because of there ability to let the network to learn from more layers without compromising accuracy. They are like shortcuts that enable the network to learn intricate connections between different parts of an image. The act of learning from more layers improves network accuracy which leads to better identification of complex features in an image. Since ResNet101 is a deeper network with more layers, skip connections are especially useful here than in a typical convolutional neural network.

One of the reasons for choosing this model is due to its ability to handle large datasets quickly and accurately. It is able to process multiple images in a short amount of time, making it suitable for tackling large datasets. Additionally, Resnet is able to handle different types of data, such as images from varied sources, various types of formats and sizes. Hence most suitable for many different data sources guaranteeing the classification of wastes accurately. It's also noteworthy that it is open-source, therefore, customizable by users based on their needs, and this makes it easier to optimize network performance and accuracy.

The model will undergo training and should be able to categorize waste using 70% of the data used for training whereas 10% will be used for validation. The model will follow the supervised learning approach and should be able to adjust parameters to match the labeled training data. To check for the model's accuracy, tests will be conducted on validation data during the training. If the model achieves a score of at least 90% then the model will be ready to be used in a production environment.

3.5 Model training

The ResNet101model will be used for classifying waste based on 4 levels according to the amount of the waste as shown in Fig. 5. Different separate sets of images will be tested in order to ascertain the model's accuracy. It is expected that after training, assessment, and evaluation, the model can be used to identify different types of waste in the future. This can be helpful for waste management, as it can be used to correctly categorize and dispose of different types of waste.

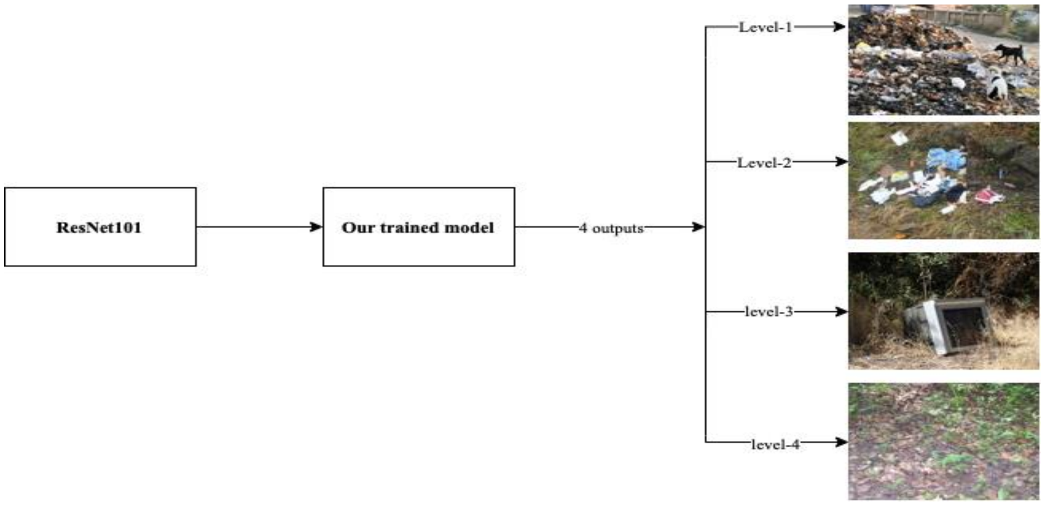

Figure 5: Trained Model Outputs.

3.6 Model Finetuning

After training, the model will undergo finetuning, a process of carefully adjusting parameters in order to optimize machine learning known as fine-tuning. It is performed to enhance the model’s performance on specific datasets. Fine-tuning adjusts the hyperparameters of a model for a better fit of data and improved accuracy. The whole process comprises adjusting parameters such as the regularization of techniques, the learning rates, the activation functions, and the number of layers and neurons. For the best possible decisions on what parameters to adjust and how best to optimize the model, it is crucial to understand the data and the model itself. The performance of the model is best evaluated to check for the improvement of the model’s overall performance after all the hyperparameters are been adjusted.

4. Experimental Results and Discussion

This section explores the results, analysis, and discussions within the technovironment after the completion of the training phase and successful integration of the various components. It involves an in-depth examination of the program's operations, along with an analysis of the output results achieved post-training.

4.1 Data Preprocessing

Google Collab was used in the experimental setup for enhanced performance during the training period. Several libraries were used in this process, and they include tensorflow, numpy, keras, and sklearn. The following steps were adhered to when installing the experimental:

- Writing the code to set the dataset.
- Writing the code for the training model, according to the ResNet101 in the original paper [21].
- Next, Train the model with the dataset from the training folder.
- Saving the trained model.

Table 2 indicates the performance of the computer that was trained on.

Table 2: Experimental Setup Specification

Item	Specification
RAM	83.48 GB
Graphics card	NVIDIA A100 Tensor Core GPU 40GB
Platform	Google Collab

4.2 Training Process

In the model training step, we utilized the ResNet101 model. ResNet101 is a deep convolutional neural network known for its excellent performance in image classification tasks, making it a suitable choice for our project. Our training dataset consisted of 3,103 images, and the training process took approximately 12 hours to complete. We trained the model with a batch size of 16 and ran it for 100 epochs. To monitor the model's performance, we employed a validation split of 0.1, which allocated a portion of the training data for validation during the training process. Once the model training was finished, we saved the trained model, which occupied a storage size of 7.94 GB. This saved model could later be used for predictions and further analysis.

The validation set is a separate set of data that is used to evaluate and validate our model's performance during training. This validation process helps us gather information that can assist in adjusting the hyperparameters of the model. As mentioned earlier, during each epoch of training, the model is trained using the training set and simultaneously validated using the validation set. In the training phase, the model classifies the input data in the training set, calculates the loss, and adjusts the model's weights accordingly. In the next epoch, the model repeats this process by classifying the same inputs again. The provided figure illustrates the results of the model validation.

During the training process, we incorporated a validation split using the split_valid parameter with a value of 0.1. This means that a portion of the training data, specifically 20%, was set aside as the validation set. The remaining 80% of the data was used for training the model. By utilizing this validation split, we were able to monitor and assess the model's performance on unseen data throughout the training iterations.

4.3 Training Results

The data set used to train the model is called the training set, the features in this data provide a good ground for training the model because the model learns during each epoch.

The main objective is for effective training of the model to be able to accurately predict information in any new or unseen data.

Figure 5 shows The model's training progress is depicted in Figure 6, showing the achieved results for loss and accuracy.

The training loss reached 0.0553, and the training accuracy reached 0.9809. On the other hand, the validation set showed a validation loss of 1.7773 and a validation accuracy of 0.6566.

These metrics provide insights into the model's performance during the training process and its ability to generalize to new data.

The model achieved a test loss of 0.8396 and a test accuracy of 0.8602, indicating its ability to accurately classify the test data with a relatively low loss value.

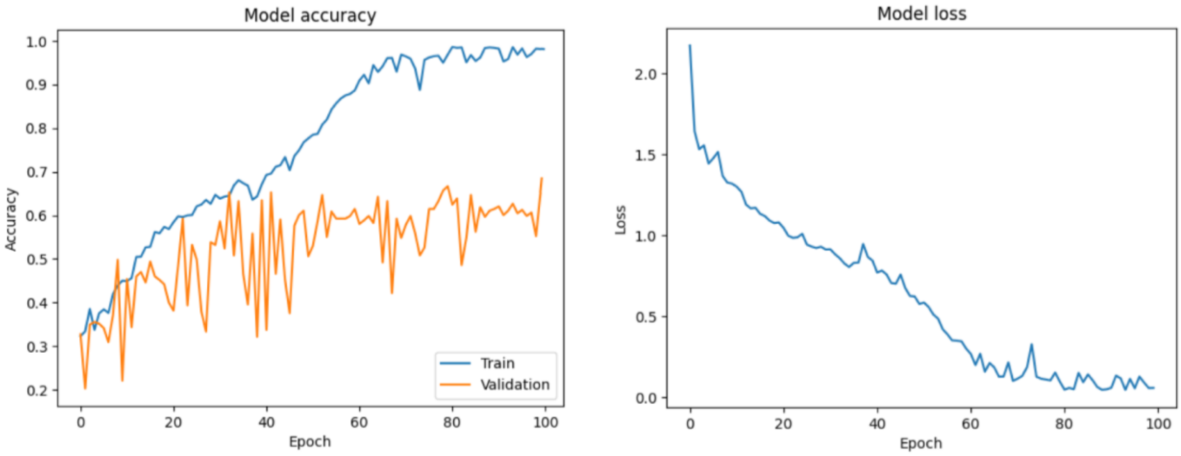

Figure 6: Training Results.

4.4 Confusion Matrix

The Confusion Matrix is an important tool in machine learning for assessing the performance of models in multilabel classification problems, where individual instances can have several labels. In such cases, the matrix gives a complete summary of the model's predictions compared to the actual labels. The summary makes it easier to understand the model for individual labels. Such errors include false positives and/or false negatives. The confusion matrix summary also, when analyzed properly, may reveal the strengths and weaknesses of the model and identify misclassified classes, which leads to informed decision-making regarding the model's performance improvement. It provides the basis for matrics evaluation calculations such as precision, recall, and F1-score, which assess the overall effectiveness of the model in handling multilabel classification.

Table 3: Result of the multilabel confusion matrix function.

	TP	FP	TN	FN
1	485	12	25	93
2	404	31	26	154
3	462	22	10	121
4	408	21	26	160

Table 4: Evaluation Metrics for Multilabel Classification Model.

	Percentage (%)
Accuracy	0.8585
Precision	0.8620
Recall	0.8568
F1 score	0.8582

4.5 Testing Samples

This section presents results of the mahual testing displayed in tables together with the predicted outcomes and their corresponding confidence rates.

Table 5: Predicted outcomes and corresponding confidence

Image	Label	Predict Result	Confidence rate
	1	1	99.99%
	1	1	89.67%
	1	1	99.68%
	1	1	98.99%

4.6 Discussion

Invaluable insights into the system's potential and performance can be obtained from the Technovironment application's training and evaluation results. Employing ResNet101 is one notable strength of the application, known for its high performance in image classification tasks. The application used more than 3000 environmental images to identify and categorize different levels of environment cleanliness, achieving exceptional accuracy in the process: 0.9809 training accuracy and 0.8602 test accuracy, demonstrating the model's ability to generalize the learned features for accurate classification of unseen environmental images. The integration of Flask as the web framework is another strength, enabling smooth communication between users and the server, making it easier and more efficient to report environmental issues.

However, the model is not without flaws. Despite the high test accuracy, the model exhibited a validation loss of 1.7773, which poses potential challenges in generalizing to unseen data. This issue may impact the model's accuracy in real-life situations. To address this, further fine-tuning and analysis of the hyperparameters are necessary to improve the model's generalization capabilities. Additionally, the training and evaluation dataset may have inherent flaws, such as biases or limitations. Any further analysis of the model would benefit from a well-represented dataset encompassing various real-world scenarios, which could enhance the model's strength and accuracy. The model's performance might also be improved by augmenting the data with a wider range of environmental conditions and issues.

A comparison with related works highlights that while many environmental monitoring systems rely on traditional machine learning models like Support Vector Machine (SVM) or decision trees, the use of deep learning architectures, such as ResNet101, in this study provides a significant improvement in accuracy and robustness. For instance, a recent study using SVM achieved an accuracy of 0.75, significantly lower than the 0.8602 test accuracy achieved by the ResNet101 model in this work. However, similar deep learning models have reported comparable performance, indicating that while the current approach is effective, there is room for further optimization and exploration of more advanced architectures [22],[22].

Moreover, the model's database functionalities, currently supported by Firebase for secure storage and retrieval of accepted reports, have room for expansion. This could be achieved by enhancing the application's effectiveness in monitoring and addressing environmental issues through integrating real-time mapping features, connecting to pollution detectors, and building algorithms for cleaning paths. Such improvements would provide users and administration staff with up-to-date information on environmental conditions, enabling timely and targeted interventions.

Invaluable insights into the system's potential and performance can be obtained from the Technovironment application training and evaluation results. Employing the ResNet101 is one such notable strength of the application. The model is known for its high performance in the tasks of classifying images. The application used more than 3000 environmental image datasets to identify and categorize the different levels of environment cleanliness, achieving exceptional accuracy in the process: 0.9809 training accuracy and 0.8602 for test accuracy, demonstrating that the model is able to successfully generalize the learned features for accurate classification of unseen environmental images. The other notable strength is the integration of Flask as the web framework. This enabled smooth communication between the users and the server, making it easier and more efficient to report environmental issues.

The model was not without flaws, it had some limitations that must be put into consideration. Despite the high test accuracy, the model had a 1.7773 validation loss, posing possible future challenges in generalizing unseen data, which has the potential to impact the accuracy of the model in real-life situations. To rectify this issue, more fine-tuning and analysis should be performed on the hyperparameters to improve the generalization capabilities of the model. Furthermore, the training and evaluation dataset may have inherent flaws, such as biases and/or limitations. Any further analysis of the model may require a well-represented dataset that encompasses different real-world scenarios, which may enhance the model's strength and accuracy. The model's performance can also be enhanced by augmented data that incorporates a wide range of environmental conditions and issues.

In addition, the model still has room for expanding its database functionalities and capabilitities in addition to the Firebase that provides secure storage and retrieval of accepted reports. This may be achieved in the form of enhancing the application's effectiveness in monitoring and addressing environment issues through integrating real-time mapping features, connecting to pollution detectors, and building algorithms for cleaning paths. Such improvements enable users and administration staff to have up-to-date information on environmental conditions for a timely and targeted intervention.

5. Conclusion

In summary, Technovironment is an original premier application that effectively uses deep learning power and image recognition technologies in conjunction with a thoroughly trained ResNet101 model on an all-inclusive dataset of over 3000 images to successfully deal with environmental issues. It has a user-friendly interface that allows users to easily capture and submit images that include accurate location details, which eases the process of reporting a wide range of environmental concerns. The application guarantees fast communication between users and the server facilitated by Flask. This allows for accuracy in the classification and prediction of submitted reports that is leading to efficient data transfer and storage security in the Firebase.

6. Author's Contribution

The first author created and trained the model, wrote the introduction, literature review, conclusion and discussion parts. The second author was responsible for problem formulation, conceptualization, and writing the methodology part.

7. Conflict of Interest

There is no conflict of interest in this paper.

8. Acknowledgment

The authors thank two students (Yahya Adnan Ghadhban and Musatafa Dilshad Khalid) from TIU – Computer Engineering department for helping with collecting the data. And a special thank to Mr. David W. Waswa a lecturer at TIU – Education Faculty for proofreading the manuscript.

References

[1] Salimi I, Bayu Dewantara BS, Wibowo IK. Visual-based trash detection and classification system for Smart Trash Bin Robot. 2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC). 2018. https://doi.org/10.1109/KCIC.2018.8628499

[2] Hama AR, Tahir TA, Ali BJ. A study on solid waste generation, composition and management in Sulaimania City, Kurdistan Region, Iraq. IOP Conference Series: Earth and Environmental Science. 2021;779(1):012049. https://doi.org/10.1088/1755-1315/779/1/012049

[3] Hossain S, Debnath B, Anika A, Junaed-Al-Hossain M, Biswas S, Shahnaz C. Autonomous Trash collector based on object detection using Deep Neural Network. TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON). 2019. https://doi.org/10.1109/TENCON.2019.8929270

[4] Conley G, Zinn SC, Hanson T, McDonald K, Beck N, Wen H. Using a deep learning model to quantify trash accumulation for cleaner urban stormwater. Computers, Environment and Urban Systems. 2022;93:101752. https://doi.org/10.1109/TENCON.2019.8929270

[5] Shaikh F, Kazi N, Khan F, Thakur Z. Waste profiling and analysis using machine learning. 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA). 2020. https://doi.org/10.1109/ICIRCA48905.2020.9183035

[6] MITRA A. Detection of Waste Materials Using Deep Learning and Image Processing. MSc Dissertation, Dept. of Comput. Sci., CALIFORNIA STATE UNIVERSITY SAN MARCOS, US. 2020. http://hdl.handle.net/20.500.12680/hq37vs75c

[7] Raza SM, Hassan SA, Shin S. Real-Time Trash Detection for Modern Societies using CCTV to Identifying Trash by utilizing Deep Convolutional Neural Network. https://doi.org/10.48550/arXiv.2109.09611

[8] Ashir AM, Ibrahim S, Abdulghani M, Ibrahim AA, Anwar MS. Diabetic Retinopathy Detection Using Local Extrema Quantized Haralick Features with Long Short-Term Memory Network. Int J Biomed Imaging. 2021;2021. https://doi.org/10.1155/2021/6618666

[9] Zhang S, Chen Y, Yang Z, Gong H. Computer Vision based two-stage waste recognition-retrieval algorithm for waste classification. Resources, Conservation and Recycling. 2021;169:105543. https://doi.org/10.1016/j.resconrec.2021.105543
[10] Kulkarni H, Raman N. Waste Object Detection and Classification. rep, Dept. of Comput. Sci. Stanford University.
[11] Alsubaei FS, Al-Wesabi FN, Hilal AM. Deep learning-based small object detection and classification model for Garbage Waste Management in Smart Cities and IOT Environment. Applied Sciences. 2022;12[5]:2281. https://doi.org/10.3390/app12052281
[12] Patel D, Patel F, Patel S, Patel N, Shah D, Patel V. Garbage detection using advanced object detection techniques. 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS). 2021. https://doi.org/10.1109/ICAIS50930.2021.9395916
[13] Ashir AM. A Transfer-Learning-Based Approach for Emergency Vehicle Detection. Eurasian Journal of Science and Engineering. 2022;8(1):75-86. https://doi.org/10.23918/eajse.v8i1p75
[14] Ashir AM. A Real-Time Automatic Kurdistan Numberplate Recognition System. Eurasian Journal of Science and Engineering. https://doi.org/10.23918/eajse.v8i1p126
[15] Verma V, Gupta D, Gupta S, Uppal M, Anand D, Ortega-Mansilla A, et al. A deep learning-based intelligent garbage detection system using an unmanned aerial vehicle. Symmetry. 2022;14[5]:960. https://doi.org/10.3390/sym14050960
[16] Majchrowska S, Mikołajczyk A, Ferlin M, Klawikowska Z, Plantykow MA, Kwasigroch A, et al. Deep learning-based waste detection in natural and Urban Environments. Waste Management. 2022;138:274–284. https://doi.org/10.1016/j.wasman.2021.12.001
[17] Liu Y, Ge Z, Lv G, Wang S. Research on automatic garbage detection system based on deep learning and narrowband internet of things. Journal of Physics: Conference Series. https://doi.org/10.1088/1742-6596/1069/1/012032
[18] Singh S, Mehta KS, Bhattacharya N, Prasad J, Lakshmi SK, Subramaniam KV, et al. Identifying uncollected garbage in urban areas using crowdsourcing and machine learning. 2017 IEEE Region 10 Symposium (TENSYMP). 2017. https://doi.org/10.1109/TENCONSpring.2017.8070078
[19] Mittal G, Yagnik KB, Garg M, Krishnan NC. SpotGarbage: smartphone app to detect garbage using deep learning. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016 Sep;940–954. https://dl.acm.org/doi/abs/10.1145/2971648.2971731
[20] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015:770-778. https://doi.org/10.1109/CVPR.2016.90
[21] Jozdani SE, Johnson BA, Chen D. Comparing Deep Neural Networks, Ensemble Classifiers, and Support Vector Machine Algorithms for Object-Based Urban Land Use/Land Cover Classification. Remote Sensing. 2019 Jul;11[14]:1-24. https://doi.org/10.3390/rs11141713
[22] Gusain HS, Negi HS, Dhamija S, Mishra V, Snehmani. Development of Avalanche Information System Using Remote Sensing and GIS Technology in the Indian Karakoram Himalaya. Current Science. 2019;117(1):104. https://www.jstor.org/stable/27138209.