ISSN: 2182-2069 / E-ISSN: 2182-2077

Evaluation of Performance of Different Optimizers of Convolutional Neural Network in the Classification of Images of Urban Domestic Solid Waste

Vicknesh Kumar Yogamadhavan[1], and Dr. Giriraj Mannayee[2*]

[1]Research Scholar, School of Mechanical Engineering, Vellore Institute of Technology, Vellore Campus, Tiruvalam Rd, Katpadi, Vellore, Tamil Nadu, India.
vickneshkumar.y2016@vitstudent.ac.in, https://orcid.org/0000-0002-7505-8173

[2*]Professor, Department of Design and Automation, School of Mechanical Engineering, Vellore Institute of Technology, Vellore Campus, Tiruvalam Rd, Katpadi, Vellore, Tamil Nadu, India.
m.giriraj@vit.ac.in, https://orcid.org/0000-0002-8726-2972

Received: February 24, 2024; Revised: April 20, 2024; Accepted: May 30, 2024; Published: August 30, 2024

Abstract

The inevitable urbanization in the modern world with its exponential population generates highly voluminous domestic waste. Disposal of this waste has become a challenging task. The process of solid waste management is a very tedious one and requires large manpower and poses serious health hazards. Hence, the implementation of automation in the process of collection and segregation of domestic solid waste has become mandatory. In our work, we have tried to identify and classify domestic urban solid waste in their real background using convolutional neural network (CNN), a genre of deep learning. 1892 photos of commonly littered street wastes were taken with their real background in the presence of natural sunlight. The photos of the wastes were grouped into 22 classes and labeled accordingly. These images were trained through transfer learning in the various pre-trained neural networks such as AlexNet, ResNet-18, Places365-GoogLeNet, SqueezeNet, GoogLeNet, ResNet-50, ShuffleNet, MobileNet-v2, NasNet-Mobile, Inception-v3, and ResNet-101. The performance of the different optimizers sgdm, adam, and rmsprop was evaluated in each of these networks for the different initial learn rates. It was found that the overall performance of the optimizers was similar, where 0.001 was the initial learn rate achieved maximum validation accuracy in most of the convolutional neural network pre-trained models. Among all the networks MobileNet-v2 achieved maximum validation accuracy and was able to predict and classify a maximum of 17 classes of waste. Footwear and wrappers were easily identified by most of the neural networks. Cigarette butts, dry flowers, fabric waste, vegetable waste, and wooden waste were never able to be classified by any of the chosen networks.

Keywords: Initial Learn Rate, Optimizer, Training, Validation, Trial Time, Confusion Matrix.

Journal of Internet Services and Information Security (JISIS), volume: 14, number: 3 (August), pp. 42-77.
DOI: *10.58346/JISIS.2024.I3.004*

*Corresponding author: Professor, Department of Design and Automation, School of Mechanical Engineering, Vellore Institute of Technology, Vellore Campus, Tiruvalam Rd, Katpadi, Vellore, Tamil Nadu, India.

1 Introduction

The rate of production in fast-growing countries like India necessitates advanced technology to handle the equally voluminous urban domestic solid waste produced daily. It is time to switch over from conventional waste collection to technology-based waste collection and management. Researches show that the vision system (Aarthi & Rishma, 2023) is the primary requirement for automated waste management systems like robots. The challenging task for the vision system of the autonomous robot is the identification and classification of domestic solid waste from the real and natural environment. The deep learning technique has a high scope in sorting out solid waste through pre-trained networks of convolutional neural networks. Municipal solid waste management with the aid of CNN provides relatively better accuracy (Mookkaiah et al., 2022) when compared to other existing methods. Lot of research works have been carried out on the classification of images in various fields like medical science, construction, waste management, natural disasters, agriculture, entomology, industrial management, biology, traffic management, space research etc., with the help of various pretrained convolutional networks such as AlexNet, GoogLeNet, ResNet, VGG16, VGG19, SqueezeNet, Inceptionv3, Densenet201, Resnet18, ResNet34, Resnet50, Resnet101, DenseNet169, and their modified versions. In most of the research works clear images of the objects of interest with preferable backgrounds have been used for the training of the dataset in the selected network models. Moreover, in the management of municipal solid waste, different kinds of solid waste have been grouped into their corresponding common 6 major classes such as cardboard, glass, metal, plastic, paper, and other garbage pictures (Melinte et al., 2020) and used as datasets for the training and further classification. In our research work, photos of various categories of domestic solid waste have been taken from the street with the real background in the natural environment with actual sunlight. All the captured images have been sorted into 46 various classes of solid waste and kept in different folders that form the dataset for training and validation. The folders have been labeled with the respective waste names. In our research work, we have used the following 11 pretrained neural networks to evaluate the performance of the different optimizers sgdm, adam, and rmsprop with the help of transfer learning technique in the prediction and classification of urban domestic solid waste. The pretrained convolutional networks used are AlexNet, ResNet-18, Places365-GoogLeNet, SqueezeNet, GoogLeNet, ResNet-50, ShuffleNet, MobileNet-v2, NasNet-Mobile, Inception-v3, and ResNet-101. During the experimentation process, after conducting a few training trials through transfer learning in various neural networks, we faced some problems in the generation of the confusion matrix. The recall values and predicted values in each cell of the confusion matrix did not appear. Then by trial and error, the number of waste classes was reduced to 22 by considering the important wastes. Now the confusion matrix was able to show the values of both recalling and prediction in each cell of it.

2 Related Work

A study of previous research works shows that various CNN models have been used by researchers in various fields, for the classification of images. AlexNet, GoogLeNet, and ResNet have been used (Kurtulmuş, 2021) for the identification of sunflower seeds. Raghu et al., (2020) have used the following pretrained networks - AlexNet, VGG16, VGG19, SqueezeNet, GoogLeNet, Inceptionv3, Densenet201, Resnet18, Resnet50, and Resnet101 for their work on the classification of the type of multi-class seizures based on EEG. GoogLeNet and Inceptionv3 have produced the highest classification accuracy of 82.85% and 88.30% respectively. In their paper, (Zhang et al., 2021 “a”) have concluded that the accuracy of classification of the waste is higher when using the CNN model DenseNet169. This model, according to them, based on transfer learning is more accurate when compared to the other CNN models AlexNet, GooleNet and VGG. Chen et al., (2022) have used the pretrained network models AlexNet and Resnet 18, for the reduction of traffic accidents by monitoring the internal state of the drivers based on 30 channel EEG signals. Rehman et al., (2020) have used AlexNet, GoogLeNet and VGG for the classification of brain tumors, in which the highest accuracy is shown by the network model VGG. Inception V3, ResNet50, VggA, and VGG16 have been used (Wang et al., 2020) for the identification of pest images. In a research on the classification of skin cancer, (Dorj et al., 2018) have used the pretrained CNN model AlexNet to extract the image features. In the research work carried out on the classification of the crop pest, (Thenmozhi & Srinivasulu Reddy, 2019) have used the pretrained CNN models of AlexNet, GoogLeNet, ResNet and VGGNet for evaluating the performance of their proposed model. Rahman et al., (2022) have used AlexNet, ResNet34, and VGG16 for the classification of indigestible waste such as cardboard, glass, metal, paper, plastic, and trash to build an intelligent waste management system. The authors found that ResNet34 performs considerably better when compared to other pre-trained network models. Kim et al., (2020) have used the ResNet model for the training of waste categorization. A garbage detecting system has been developed (Abinandan et al., 2022) with the help of pretrained CNN models like VGG-16 and Inception V3. In the research work on the development of noise-robust CNN for the classification of the images, (Momeny et al., 2021) have used GoogLeNet, ResNet, VGG-Net-Medium, and VGG-Net-Slow to evaluate the performance of their proposed CNN. In the research work of the positioning of the endotracheal tube and classification of X-ray images the (Lakhani, 2017) has used the pretrained CNN networks AlexNet and GoogLeNet. In the research work of classification of images of clinical melanoma and atypical nevi, authors (Brinker et al., 2019) have used the ResNet-50. Inception-v3 has been used (Bard et al., 2019) in their research work on the classification of images for robotic plastering. Fu & Aldrich, (2019) have used AlexNet, VGG16, and ResNet for the estimation of the froth grades from the images captured from industries in which AlexNet shows better performance. The CNN pretrained model Inceptionv3 has been used (Ramírez et al., 2020) in the detection and recognition of street dumpsters. A DenseNet169 based visual recognition model has been generated in the research work on unattended gauging of the composition of construction waste, (Chen et al., 2021; Muangnak et al., 2021) have investigated the CNN models of ResNet-152 and ResNet-50 to develop a suitable classification model for the classification of wastes. EfficientNet-B2 has been used (Majchrowska et al., 2022) for the classification of detected wastes into seven categories. ShuffleNet v2 has been used to develop a lightweight waste classification model (Chen et al., 2022; Sun & Gu, 2022) have used the CNN models GoogLeNet, ResNet, and VGG16 for verifying the quality and efficiency of their dataset built up of construction waste materials such as brick, concrete, metal, stone, and wood. VGG-16 has been used (Iwashita et al., 2022) for the development of a CNN model in the work of prediction of soil particle size deposited during the natural disaster of a Tsunami. Madappa et al., (2020) have concluded that ResNet and VGG CNN models perform better than the customized CNN model in the classification of images of the waste for aiding the automated waste-picking robotic arm. The performance of DenseNet-201 has been evaluated (Chen et al., 2022), in the prediction of environmental microorganisms. In the research carried out on the performance analysis among the CNN networks AlexNet, GoogLeNet, and ResNet 50 in the classification of images, the (Sharma et al., 2018) have found that GoogLeNet and ResNet-50 performed better than AlexNet. In the research work carried out by the authors (Kadhim & Abed, 2020) on the classification of satellite images with the pre-trained neural networks AlexNet, GoogLeNet, ResNet-50, and VGG19, the ResNet 50 shows a promising result in the feature extraction when compared to that of other networks. Jain et al., (2022) have claimed that the CNN model EfficientNet-b0 performed effectively, in their work on a smart garbage classification system. According to the (Unal et al., 2022) the CNN model SqueezeNet has performed well in the identification of species of coffee beans when compared to that of other networks Inception v3, VGG16, and VGG19. In the research work on the recognition of COVID-19 from the noisy X-ray images with the modified pretrained networks EfficientNetb0, GoogLeNet, MobileNetv2, ResNet18, ResNet50, ShuffleNet, and SqueezeNet for enhancing the robustness against the impulse noise, the authors (Akbarimajd et al., 2022) have concluded that ResNet50 shows better accuracy on the classification. In the research work on the development of a system for the automatic classification of MRI brain images the (El Boustani et al., 2020) have concluded that the performance of the optimizer RMSprop is better than the other optimizers Adam and SGD in terms of accuracy, the loss and the time of execution.

3 Methodology

The software module used for the classification of images of domestic urban solid waste is the Deep Network Designer module of MATLAB. The transfer learning technique was used to classify the captured solid waste images to improve the accuracy of the classification of the waste. The laptop Acer Nitro AN515-57 was used for the experiments. The processor of the laptop has 11th Gen Intel(R) Core (TM) i5-11400H @ 2.70GHz 2.69 GHz. The installed RAM capacity is 16.0 GB (15.8 GB usable). The system type is a 64-bit operating system with an x64-based processor. The laptop is inbuilt with the GPU NVIDIA GeForce RTX3050 with 4GB memory.

The experiments were conducted in 11 various pre-trained convolutional networks. The AlexNet has a total of 25 layers with 24 connections and 8 layers deep. The input size of the image is 227 x 227. The ResNet-18 has a total of 71 layers with 78 connections and 18 layers deep. The input size of the image is 224 x 224. The Places365-GoogLeNet has a total of 144 layers with 170 connections and 22 layers deep. The input size of the image is 224 x 224. The SqueezeNet has a total of 68 layers with 75 connections and 18 layers deep. The input size of the image is 227 x 227. GoogLeNet has a total of 144 layers with 170 connections and 22 layers deep. The input size of the image is 224 x 224. The ResNet-50 has a total of 177 layers with 192 connections and 50 layers deep. The input size of the image is 224 x 224. The ShuffleNet has a total of 172 layers with 187 connections and 50 layers deep. The input size of the image is 224 x 224. The MobileNet-v2 has a total of 154 layers with 163 connections and 53 layers deep. The input size of the image is 224 x 224. The NasNet-Mobile has a total of 913 layers with 1072 connections. The input size of the image is 224 x 224. The Inception-v3 has a total of 315 layers with 349 connections and 48 layers deep. The input size of the image is 299 x 299. The ResNet-101 has a total of 347 layers with 379 connections and 101 layers deep. The input size of the image is 224 x 224.

The (Chen et al., 2021) have constructed the data set using 2000 solid waste images. Among them 500 images have been taken from the network resources and the other 1500 solid waste images have been captured on a black background panel by the camera. The (Mao et al., 2021) have used six different categories of waste cardboard, glass, metal, paper, plastic, and trash to evaluate the performance of the CNN to employ a vision-based robot that can automatically classify the wastes for recycling and to reduce the labor requirement. In another waste classification problem, the (Toğaçar et al., 2020) have used the recycling of wastes such as cardboard, cloth, glass, metal, paper, plastics, and organic wastes such as organic foods to generate the dataset. In their research work, the (Altikat et al., 2022) have used pictures of glass, organic waste, paper, and plastic wastes captured from the natural environment for the classification. Ruiz et al., (2019) have used RGB images of waste of the following six classes cardboard, general trash, glass, metal, paper, and plastic for generating the dataset which was taken by placing the objects over a white postcard under sunlight and room light. Three classes of recyclable waste glass-metal, paper, and plastic along with the biodegradable, and non-recyclable wastes have been used

(Nguyen et al., 2022) to generate the image dataset with the view to evaluate the performance of deep learning in the development of the green smart environment. Islam & Alam, (2022) have proposed to develop a new deep learning-based approach to identify and classify five types of wastes such as Glass, Kitchen, Metal, Paper, and Plastic.

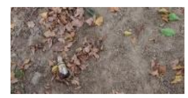
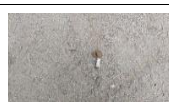
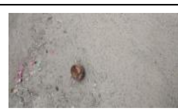
The images used in our research were taken randomly over a month along 6 streets in the area of Iyyapanthangal, Chennai, Tamil Nādu, India, using the LG-G8 ThinQ with a 128GB Memory mobile. The pixels of the rear-facing camera was 16MP. It has a front and back integrated camera. The recording resolution was 3840 x 2160 (4K). The frame rate was 60 frames per second. The most commonly found waste in these photographs were the wrappers of biscuits, chocolates, and other food items. The least found were the glass bottles. Next to the wrappers, the most identified wastes were plastic waste, cardboard and tetra packs, plastic bottles, face masks, and paper cups. The images that were used for the classification were with actual backgrounds. The item of interest in the images was fully set with a real environment background. The photos were also captured in real sunlight with no artificial lighting. The pictures were not cropped for the item of interest. The image augmentation was not performed to evaluate the performance of the various pre-trained networks in the classification of images of various domestic solid wastes that were surrounded by natural scenarios. The sample photos of each waste class are given in Table 1. In our research work the image dataset has been formed from the total number of captured images of 1892 and was grouped into 22 classes. The various classes of solid wastes used in our research work were bottle caps, cardboards and tetra packs, cigarette butts, coconut waste, dairy packets, dry flowers, dry leaves, dry twigs and branches, fabric waste, face masks, footwear, fruit waste, glass bottles, matchboxes, paper cups, paper waste, plastic bottles, plastic waste, rope and thread, vegetable waste, wooden waste, and wrappers.

Generally, the performance will be poor on new data sets (Ahmad et al., 2021). This is the overfitting problem. To avoid this problem the dataset has been divided into training sets and test sets. Various network models have been generated using the training dataset. The accuracy of the models has been verified using the test dataset. In the research work on identifying and classifying acute lymphoblastic leukemia, the (Das & Meher, 2021) have concluded that the use of 70% of the dataset as a training set and 30% of the dataset as the test set has achieved the most excellent accuracy. In our classification work experiments, 70% of the Images have been used to generate datasets for training, and 30% of the images have been used to generate datasets for validation.

Our experiments were conducted in various pretrained convolutional networks of matlab for evaluating the performances of the different optimizers Sgdm (Stochastic Gradient Descent with Momentum), adam (Adaptive Moment), and rmsprop (Root Mean Square Propagation). One experiment and three trials were conducted for each optimizer in each neural network. Likewise, 33 experiments and 99 trials were also conducted.

In the training of the 11 pretrained network models on the given dataset of images, the mini-batch size was taken as 128, and the maximum number of iterations was 300. The hardware source used was a single GPU. The learning rate schedule was fixed constant. The maximum number of epochs was 30 and the iterations per epoch were 10. In the sgdm optimizer, the momentum was set to 0.9. In the adam optimizer, the gradient decay factor was set as 0.9 and the squared gradient decay factor was set as 0.999. In the rmsprop optimizer, the Squared Gradient Decay Factor was set as 0.9. The model's performance may depend on the learning rate value we choose. The model may show poor performance (Lin et al., 2023) with a large learning rate value, whereas the time taken for training a model may be more with a very less learning rate value. Until now the practice is to set the learning rate at random, which has an impact on the computing sources. In our experiments, the network models have been trained with the InitialLearnRates 0.001, 0.01, and 0.1 with sgdm optimizer, with the InitialLearnRates 0.0001, 0.001, and 0.01 with adam optimizer, and with the InitialLearnRates 0.0001, 0.001, and 0.01 with rmsprop optimizer. While acquiring the images from the image dataset the total observations obtained from the input of 1892 images were 1325. The highest number of observations was acquired for the wrappers with a quantity of 305. The lowest number of observations was acquired for the glass bottles with a quantity of 12.

Table 1: Sample Photos of Urban Street Solid Waste

Waste class number	**Name of the waste**	**Sample photos**	**Waste class number**	**Name of the waste**	**Sample photos**
1	Bottle cap		12	Fruit waste	
2	Cardboard and tetra pack		13	Glass bottle	
3	Cigarette butt		14	Matchbox	
4	Coconut waste		15	Paper cup	
5	Dairy packet		16	Paper waste	
6	Dry flowers		17	Plastic bottle	
7	Dry leaves		18	Plastic waste	
8	Dry twigs & branches		19	Rope and thread	
9	Fabric waste		20	Vegetable waste	
10	Face mask		21	Wooden waste	
11	Footwear		22	Wrapper	

4 Training and Validation Results

- **AlexNet**

In the AlexNet network, the maximum training accuracy achieved for the adam optimizer was high whereas the maximum validation accuracy was achieved by sgdm optimizer.

In the sgdm optimizer, the training accuracy was the maximum for the InitialLearnRate of 0.001. The total time taken for the maximum iteration of 300 was 33 minutes and 34 seconds. The training accuracy achieved was 96.875% and its corresponding validation accuracy was 23.2804% and it was the maximum among the 3 trials. The training accuracy was the minimum for InitialLearnRate of 0.01. The time taken was 5 minutes 53 seconds. The training accuracy achieved was 17.9688%. The validation accuracy was 1.7637%. The training accuracy and validation accuracy of InitialLearnRate 0.1 were 27.3438% and 23.1041%. The elapsed trial time was 29 minutes and 53 seconds.

In the adam optimizer, the highest training accuracy was achieved for the InitialLearnRate of 0.0001. The training accuracy was 97.6563% with the corresponding validation accuracy of 22.2222%. The time taken to complete the trial was 52 minutes and 5 seconds. The maximum validation accuracy achieved was 23.1041%. It was the same for both the InitialLearnRates 0.001 and 0.01 and the corresponding training accuracies were 23.4375% and 22.6563%. The elapsed trial time was 51 minutes and 43 seconds, and 51 minutes and 12 seconds.

Simultaneously, in the rmsprop optimizer, the maximum training accuracy was achieved for the InitialLearnRate of 0.0001. The training accuracy was 89.0625% and the corresponding validation accuracy was 14.2857%. The elapsed trial time was 46 minutes and 12 seconds. The maximum validation accuracy achieved was 23.1041% and it was the same for both the InitialLearnRates 0.001 and 0.1 and the corresponding training accuracies were 21.875% and 17.1875%. The elapsed trial time was 50 minutes and 29 seconds, and 45 minutes and 57 seconds.

Table 2 shows only the values of maximum training and validation accuracies achieved over the training of AlexNet with different InitialLearnRate and optimizers.

Table 2: Training Results of AlexNet for Maximum Training and Validation Accuracies

Optimizer	InitialLearnRate	Elapsed Time	Maximum Training Accuracy (%)	Corresponding Validation Accuracy (%)	Maximum Validation Accuracy (%)	Corresponding Training Accuracy (%)
sgdm	0.001	33 minutes 34 seconds	96.875		23.2804	
adam	0.0001	52 minutes 5 seconds	97.6563	22.2222		
adam	0.001	51 minutes 43 seconds			23.1041	23.4375
adam	0.01	51 minutes 12 seconds			23.1041	22.6563
rmsprop	0.0001	46 minutes 12 seconds	89.0625	14.2857		
rmsprop	0.001	50 minutes 29 seconds			23.1041	21.875
rmsprop	0.01	45 minutes 57 seconds			23.1041	17.1875

- **SqueezeNet**

In the SqueezeNet network, the maximum training accuracy and validation accuracy achieved for the rmsprop optimizer was high.

In the sgdm optimizer, the training accuracies were the maximum and the same is seen for the initial learning rate of 0.001 and 0.1 also. The value of training accuracy was 18.75%. The elapsed trial time for both trials was 9 minutes 33 seconds and 9 minutes and 37 seconds. The validation accuracy was also the same for both trials. Its value was 23.1041%. The training accuracy was zero for the InitialLearnRate of 0.01 whereas the validation accuracy was 1.7637%. The training loss and validation loss were NaN. The time consumed was 1 minute 54 seconds.

At the same time, in the adam optimizer, the highest training accuracy and validation accuracy were achieved with an InitialLearnRate of 0.001. The corresponding training accuracy was 32.8125% whereas the validation accuracy was 23.9859%. The time taken to run the trial was 10 minutes and 4 seconds. Simultaneously, the minimum training accuracy achieved was 18.75% for the InitialLearnRate of 0.01 with an elapsed trial time of 10 minutes and 5 seconds. The corresponding validation accuracy was 23.1041%. The training accuracy and validation accuracy of InitialLearnRate 0.0001 were 32.0313% and 22.575%. The elapsed trial time was 9 minutes and 51 seconds.

In the experiment with rmsprop optimizer, the maximum training and validation accuracy were achieved for the InitialLearnRate of 0.0001 whose value was 76.5625%. The experiment consumed 9 minutes and 43 seconds, and the corresponding validation accuracy was 25.9259%. At the same time, the minimum training accuracy was achieved for the InitialLearnRate of 0.01 with the value of 18.75%. The corresponding validation accuracy of this optimizer was 23.1041%. It took 17 minutes and 17 seconds to complete the trial. The training accuracy and validation accuracy of InitialLearnRate 0.001 were 32.0313% and 22.7513%. The elapsed trial time was 9 minutes and 46 seconds.

Only the values of maximum training and validation accuracies achieved over the training of SqueezeNet with different InitialLearnRate and optimizers are depicted in Table 3.

Table 3: Training Result of SqueezeNet for Maximum Training and Validation Accuracies

Optimizer	**InitialLearnRate**	**Elapsed Time**	**Maximum Training Accuracy (%)**	**Corresponding Maximum Validation Accuracy (%)**
sgdm	0.001	9 minutes 33 seconds	18.75	23.1041
sgdm	0.1	9 minutes 37 seconds	18.75	23.1041
adam	0.001	10 minutes 4 seconds	32.8125	23.9859
rmsprop	0.0001	9 minutes 43 seconds	76.5625	25.9259

- **ShuffleNet**

In the ShuffleNet network, the maximum training accuracy value was achieved for sgdm optimizer whereas the maximum validation accuracy was achieved for the rmsprop optimizer.

The maximum training accuracy was achieved in sgdm optimizer, for the InitialLearnRate of 0.001 and for the elapsed trial time of 22 minutes 45 seconds and the value was 98.4375%. The corresponding validation accuracy was 19.9295%. The maximum validation accuracy value was 22.9277% which was achieved for the InitialLearnRate of 0.01. The corresponding training accuracy value for this InitialLearnRate of 0.01 was 96.875%, and the trial time lapse was 22 minutes and 52 seconds. The minimum training accuracy value of 96.0938% was achieved for the InitialLearnRate of 0.1 with the lapsed trial time of 22 minutes and 57 seconds. The subsequent validation accuracy value was 17.284%.

In the adam optimizer, the maximum training accuracy values were the same for both the InitialLearnRates 0.0001 and 0.001 with the value 96.875%. The elapsed trial time for the InitialLearnRate of 0.0001 and 0.001 were 22 minutes and 55 seconds and 22 minutes and 45 seconds. The maximum validation accuracy was achieved for the InitialLearnRate of 0.001 and the corresponding value was 23.8095%. At the same time, the validation accuracy for the InitialLearnRate of 0.0001 was 20.4586%. The minimum value of training accuracy was achieved for the InitialLearnRate of 0.01 with the trial time lapse of 34 minutes and 34 seconds. 93.75% and 19.9295% were the training accuracy and corresponding validation accuracy values.

In a similar manner, in the rmsprop optimizer, the maximum training accuracy values achieved were the same for both the InitialLearnRates 0.0001 and 0.001 and the value was 96.0938%. The trial time lapse for the InitialLearnRate of 0.0001 and 0.001 were 26 minutes and 3 seconds and 25 minutes and 31 seconds. The maximum validation accuracy was achieved for the InitialLearnRate of 0.001 with the corresponding value of 24.6914%. The validation accuracy for the InitialLearnRate of 0.0001 was 20.6349% while the minimum value of training accuracy was achieved for the InitialLearnRate of 0.01 for the trial time lapse of 25 minutes and 37 seconds. The training accuracy and corresponding validation accuracy values were 86.7188% and 24.515%.

Only the values of maximum training and validation accuracies achieved over the training of ShuffleNet with different InitialLearnRate and optimizers are brought out clearly in Table 4.

Table 4: Training Result of ShuffleNet for Maximum Training and Validation Accuracies

Optimizer	**InitialLearnRate**	**Elapsed Time**	**Maximum Training Accuracy (%)**	**Corresponding Validation Accuracy (%)**	**Maximum Validation accuracy (%)**	**Corresponding Training Accuracy (%)**
sgdm	0.001	22 minutes 45 seconds	98.4375	19.9295		
sgdm	0.01	22 minutes 52 seconds			22.9277	96.875
adam	0.0001	22 minutes 55 seconds	96.875	20.4586		
adam	0.001	22 minutes 45 seconds			23.8095	96.875
rmsprop	0.0001	26 minutes 3 seconds	96.0938	20.6349		
rmsprop	0.001	25 minutes 31 seconds			24.6914	96.0938

- **Places365-GoogLeNet**

In the Places365-GoogLeNet network, the maximum training accuracy was achieved for the adam optimizer. The maximum validation accuracy was achieved for the sgdm optimizer.

In the sgdm optimizer, the training accuracy was at its maximum for the initial learning rate of 0.01. The total time taken for the maximum iteration of 300 was 28 minutes and 23 seconds with the resultant accuracy of 79.6875% whereas the corresponding validation accuracy was 27.6896% and it was the maximum. The training accuracy was zero for InitialLearnRate of 0.1. The time taken was 1 minute and 53 seconds with the corresponding validation accuracy of 1.7637%. The training loss and validation loss were NaN. The training accuracy and validation accuracy of InitialLearnRate 0.001 were 72.6563% and 25.0441%. The elapsed trial time was 28 minutes and 30 seconds.

In the adam optimizer, the highest training accuracy was achieved for the InitialLearnRate of 0.0001. The training accuracy achieved was 93.75%. The corresponding validation accuracy was 25.9259% consuming 29 minutes and 32 seconds trial time. In adam optimizer, the validation accuracy for both the InitialLearnRates 0.0001 and 0.001 was the same and maximum with the value 25.9259%. At the same time, the training accuracy of the InitialLearnRate 0.001 was 37.5% and the time elapsed for the trial was 28 minutes and 40 seconds. The minimum validation accuracy was 23.1041% achieved for the InitialLearnRate 0.01 with the trial time lapse of 28 minutes and 33 seconds. The corresponding training accuracy value was 25.7813%.

In the rmsprop optimizer, the InitialLearnRate of 0.0001 achieved the maximum training accuracy and the value was 78.9063%. The time taken to complete the trial was 28 minutes and 57 seconds. The corresponding validation accuracy was 19.7531%. The validation accuracy for the InitialLearnRate of 0.001 and 0.01 was the same and it was also the maximum. The value of the validation accuracy was 23.1041%. The corresponding training accuracies were 26.5625% and 17.1875% and the time taken for the trial were 29 minutes and 6 seconds and 49 minutes and 38 seconds.

Table 5 delineates only the values of maximum training and validation accuracies achieved over the training of Places365-GoogLeNet with different InitialLearnRate and optimizers.

Table 5: Training Result of Places365-GoogLeNet for Maximum Training and Validation Accuracies

Optimizer	**InitialLearnRate**	**Elapsed Time**	**Maximum Training Accuracy (%)**	**Corresponding Validation Accuracy (%)**	**Maximum Validation accuracy (%)**	**Corresponding Training Accuracy (%)**
sgdm	0.01	28 minutes 23 seconds	79.6875		27.6896	
adam	0.0001	29 minutes 32 seconds	93.75		25.9259	
adam	0.001	28 minutes 40 seconds			25.9259	37.5
rmsprop	0.0001	28 minutes 57 seconds	78.9063	19.7531		
rmsprop	0.001	29 minutes 6 seconds			23.1041	26.5625
rmsprop	0.01	49 minutes 38 seconds			23.1041	17.1875

- **GoogLeNet**

In the GoogLeNet network, the maximum training accuracy and validation accuracy were achieved for the sgdm optimizer. But the maximum training accuracy for the adam and rmsprop optimizers was the same.

The maximum training accuracy was achieved for the InitialLearnRate of 0.01 in sgdm optimizer, with the elapsed trial time of 38 minutes and 9 seconds achieving the accuracy value of 96.875%. The corresponding validation accuracy achieved was 30.3351%. But the validation accuracy was the maximum for the training accuracy of 90.625% with the elapsed trial time of 1 hour 4 minutes and 8 seconds and for the InitialLearnRate of 0.001. The validation accuracy value for the InitialLearnRate of 0.001 was 32.2751%. At the same time, the minimum training accuracy value was 5.4688% for the InitialLearnRate of 0.1 with the corresponding validation accuracy of 1.7637%. The elapsed trial time was 2 minutes and 3 seconds. The training loss and validation loss were NaN

Using adam optimizer, the maximum training accuracy was achieved for the InitialLearnRate of 0.0001 and the value was 94.5313% with the elapsed trial time of 29 minutes and 23 seconds. The corresponding validation accuracy value was 29.6296% and it was the maximum. The training accuracy and validation accuracy values were the same for both the InitialLearnRate of 0.001 and 0.01. The corresponding values were 26.5625% and 23.1041%. The respective elapsed trial times were 29 minutes 35 seconds and 29 minutes and 19 seconds.

In rmsprop optimizer, the maximum training and validation accuracies were achieved for the InitialLearnRate of 0.0001 with a trial time lapse of 30 minutes and 1 second. The respective training and validation accuracy values were 94.5313% and 31.3933%. Meanwhile, the minimum training accuracy was achieved for the InitialLearnRate of 0.01, consuming a lapsed trial time of 29 minutes and 16 seconds. The respective training and validation accuracy values were 17.1875% and 7.5838%. However, the training and validation accuracy values for the InitialLearnRate of 0.001 for the elapsed trial time of 29 minutes and 22 seconds were 23.4375% and 23.1041%.

In Table 6, only the values of maximum training and validation accuracies achieved over the training of GoogLeNet with different InitialLearnRate and optimizers are portrayed.

Table 6: Training Result of GoogLeNet for Maximum Training and Validation Accuracies

Optimizer	**InitialLearnRate**	**Elapsed Time**	**Maximum Training Accuracy (%)**	**Corresponding Validation Accuracy (%)**	**Maximum Validation accuracy (%)**	**Corresponding Training Accuracy (%)**
sgdm	0.01	38 minutes 9 seconds	96.875	30.3351		
sgdm	0.001	1 hour 4 minutes 8 seconds			32.2751	90.625
adam	0.0001	29 minutes 23 seconds	94.5313	29.6296		
rmsprop	0.0001	30 minutes 1 second	94.5313	31.3933		

- **ResNet-18**

In the ResNet-18 network, the maximum training accuracy achieved for the sgdm optimizer was high. But the maximum validation accuracy was achieved by adam optimizer.

When sgdm optimizer was used, the training accuracy was the maximum for the InitialLearnRate of 0.001. The total time taken for the maximum iteration of 300 was 26 minutes and 5 seconds. The training accuracy achieved was 97.6563%. But the corresponding validation accuracy was 22.9277%. The training accuracy was the minimum for InitialLearnRate of 0.1. The time consumption was 11 minutes and 47 seconds. The training accuracy achieved for this was 17.1875%. The subsequent validation accuracy was 23.1041%. The value of maximum validation accuracy was 24.8677% achieved for the InitialLearnRate of 0.01 and for the elapsed trial time of 11 minutes and 51 seconds. The corresponding training accuracy value was 96.0938%.

At the same time, in the adam optimizer, the highest training accuracy was achieved for the InitialLearnRate of 0.001. The training accuracy was 96.875% with the corresponding validation accuracy of 28.3951%. The time taken to complete the trial was 12 minutes and 25 seconds. In adam optimizer, the highest validation accuracy was achieved for the highest training accuracy. The training accuracy, validation accuracy and trial time lapse for the InitialLearnRates 0.0001 and 0.01 were as follows: 95.3125%, 23.4568%, 12 minutes and 18 seconds, and 21.875%, 19.224%, 12 minutes and 24 seconds.

In the rmsprop optimizer, the maximum training accuracy achieved for the InitialLearnRate of 0.0001 and 0.001 was the same and the maximum training accuracy was 95.3125%. The validation accuracy for the InitialLearnRate of 0.0001 was 24.6914% consuming 28 minutes and 11 seconds whereas the validation accuracy for the InitialLearnRate of 0.001 was 27.6896% and it was the maximum whose time consumption was 31 minutes and 19 seconds. The training accuracy, validation accuracy, and trial time lapse for the InitialLearnRate 0.01 were as follows: 17.1875%, 19.0476%, and 22 minutes and 53 seconds.

Only the values of maximum training and validation accuracies achieved over the training of ResNet-18 with different InitialLearnRate and optimizers are shown in Table 7.

Table 7: Training Result of ResNet-18 for Maximum Training and Validation Accuracies

Optimizer	InitialLearnRate	Elapsed Time	Maximum Training Accuracy (%)	Corresponding Validation Accuracy (%)	Maximum Validation accuracy (%)	Corresponding Training Accuracy (%)
sgdm	0.001	26 minutes 5 seconds	97.6563	22.9277		
sgdm	0.01	11 minutes 51 seconds			24.8677	96.0938
adam	0.001	12 minutes 25 seconds	96.875		28.3951	
rmsprop	0.0001	28 minutes 11 seconds	95.3125	24.6914		
rmsprop	0.001	31 minutes 19 seconds	95.3125		27.6896	

- **MobileNet-v2**

The maximum training accuracy value was achieved in the MobileNet-v2 network for sgdm optimizer. But the maximum validation accuracy was achieved for the adam optimizer.

In sgdm optimizer, the maximum training and validation accuracy values were achieved for the InitialLearnRate of 0.01 for an elapsed trial time of 38 minutes and 20 seconds. The respective training and validation accuracy values were 96.0938% and 28.9242%. The training accuracy and validation accuracy of InitialLearnRate 0.001 were 95.3125% and 23.2804%. The time consumed was 39 minutes and 41 seconds. And the training accuracy and validation accuracy of InitialLearnRate 0.1 were 34.375% and 19.9295% with the time consumption of 38 minutes and 16 seconds.

As per the results of adam optimizer, the maximum training accuracy values were the same for the InitialLearnRates 0.0001 and 0.001 and the value was 95.3125%. The maximum validation accuracy with the value of 33.157% was achieved for the InitialLearnRate of 0.001 for an elapsed trial time of 40 minutes and 37 seconds. At the same time, the value of validation accuracy achieved for the InitialLearnRate of 0.0001 for an elapsed trial time of 41 minutes and 36 seconds was 24.515%. The training accuracy and validation accuracy shown by InitialLearnRate 0.01 were 28.125% and 19.7531% and the time taken was 40 minutes and 30 seconds.

In rmsprop optimizer, the maximum training accuracy was achieved for the InitialLearnRate of 0.0001 for an elapsed trial time of 45 minutes and 55 seconds. The corresponding training and validation accuracy values were 94.5313% and 26.1023%. The maximum validation accuracy was achieved for the InitialLearnRate of 0.001 for an elapsed trial time of 45 minutes and 29 seconds. The validation accuracy value was 26.9841% and the corresponding training accuracy value was 92.9688%. The training accuracy and validation accuracy of InitialLearnRate 0.01 were 28.125% and 23.6332%. The elapsed trial time was 46 minutes and 9 seconds.

Table 8 brings out only the values of maximum training and validation accuracies achieved over the training of MobileNet-v2 with different InitialLearnRate and optimizers.

Table 8: Training Result of mobileNet-v2 for Maximum Training and Validation Accuracies

Optimizer	InitialLearnRate	Elapsed Time	Maximum Training Accuracy (%)	Corresponding Validation Accuracy (%)	Maximum Validation accuracy (%)	Corresponding Training Accuracy (%)
sgdm	0.01	38 minutes 20 seconds	96.0938		28.9242	
adam	0.0001	41 minutes 36 seconds	95.3125	24.515		
adam	0.001	40 minutes 37 seconds			33.157	95.3125
rmsprop	0.0001	45 minutes 55 seconds	94.5313	26.1023		
rmsprop	0.001	45 minutes 29 seconds			26.9841	92.9688

- **NasNet-Mobile**

As in MobileNet-v2 network, in NasNet-Mobile network also the maximum training accuracy value was achieved for sgdm optimizer, and the maximum validation accuracy was achieved for the adam optimizer.

In sgdm optimizer, the maximum training accuracy was achieved for the InitialLearnRate of 0.001 for an elapsed trial time of 2 hours 9 minutes and 7 seconds. The corresponding training accuracy and validation accuracy values were 99.2188% and 23.8095%. The training accuracy values were the same for both the InitialLearnRates, 0.01 and 0.1. The training accuracy value was 98.4375%. The time taken for each experiment was 1 hour 11 minutes and 52 seconds and 1 hour 11 minutes and 18 seconds. The result showed that the maximum validation accuracy was achieved for the InitialLearnRate of 0.01. This value was 26.2787%, whereas, the validation accuracy achieved for the InitialLearnRate of 0.1 was 25.3968%.

In adam optimizer, the maximum training and validation accuracy values 97.6563% and 29.6296% were achieved for the InitialLearnRate of 0.001. The corresponding elapsed trial time was 1 hour 22 minutes and 4 seconds. The training accuracy and validation accuracy of InitialLearnRate 0.0001 were 96.875% and 25.2205%. It took 1 hour 21 minutes, and 34 seconds to complete the experiment. At the same time, the training accuracy and validation accuracy of InitialLearnRate 0.01 were 73.4375% and 11.4638% with the trial time lapse of 1 hour 22 minutes and 22 seconds.

In the rmsprop optimizer method, the maximum training accuracy value of 98.4375% was achieved for the InitialLearnRate of 0.0001 for an elapsed trial time of 1 hour 17 minutes, and 27 seconds. The corresponding validation accuracy value was 26.9841%. At the same time, the maximum validation accuracy value of 27.3369% was achieved for the InitialLearnRate of 0.001 for an elapsed trial time of 1 hour 17 minutes and 40 seconds. The corresponding training accuracy value was 96.0938%. The training accuracy and validation accuracy of InitialLearnRate 0.01 were 20.3125% and 23.1041% taking 1 hour 18 minutes 0 seconds to complete the trial.

Only the values of maximum training and validation accuracies achieved over the training of NasNet-Mobile with different InitialLearnRate and optimizers are depicted in Table 9.

Table 9: Training Result of NasNet-Mobile for Maximum Training and Validation Accuracies

Optimizer	**InitialLearnRate**	**Elapsed Time**	**Maximum Training Accuracy (%)**	**Corresponding Maximum Validation Accuracy (%)**
sgdm	0.01	1 hour 6 minutes 6 seconds	98.4375	28.9242
adam	0.0001	1 hour 28 minutes 29 seconds	98.4375	28.3951
rmsprop	0.0001	1 hour 1 minute 8 seconds	97.6563	30.8642

- **ResNet-50**

In the ResNet-50 network, the maximum training accuracy values of both sgdm and adam optimizers were the same. The maximum validation accuracy was achieved for the rmsprop optimizer.

In sgdm optimizer, the maximum training accuracy values were the same for the InitialLearnRates 0.001 and 0.01. The training accuracy value was 98.4375%. But the validation accuracies achieved were different. The InitialLearnRate of 0.01 achieved the maximum validation accuracy value of 28.9242% for the elapsed trial time of 1 hour 6 minutes and 6 seconds. The validation accuracy value of 25.0441% was achieved for the InitialLearnRate of 0.001 for the elapsed trial time of 1 hour 20 minutes and 50 seconds. The minimum training accuracy of value 42.9688% was achieved for the InitialLearnRate 0.1 for the elapsed trial time of 21 hours 21 minutes 21 seconds, one of the longest runs. The corresponding validation accuracy value was 22.9277%.

In adam optimizer, the maximum training accuracy value of 98.4375% was achieved for the InitialLearnRate of 0.0001. The validation accuracy was 28.3951% and it was the maximum while the trial time taken was 1 hour 28 minutes 29 seconds. The second maximum training accuracy value of 95.3125% was achieved for the InitialLearnRate of 0.001. The corresponding elapsed trial time was 55 minutes 46 seconds and the validation accuracy was 24.6914%. At the same time, the minimum training accuracy value of 33.5938% was achieved for the InitialLearnRate of 0.01. The validation accuracy was 23.1041% with the elapsed trial time of 55 minutes and 46 seconds.

In rmsprop optimizer also, the maximum training and validation accuracy value were achieved for the InitialLearnRate of 0.0001. The respective training accuracy and validation accuracy values were 97.6563% and 30.8642% consuming 1 hour 1 minute and 8 seconds. With the adam optimizer, the second maximum training accuracy value of 96.0938% was achieved for the InitialLearnRate of 0.001 and trial time consumed was 1 hour 0 minute 34 seconds. The corresponding validation accuracy value was 23.1041%. The minimum training accuracy value was achieved for the InitialLearnRate of 0.01, for the elapsed trial time of 1 hour 0 minutes, and 14 seconds. The corresponding training and validation accuracy values were 28.9063% and 22.7513%.

Only the values of maximum training and validation accuracies achieved over the training of ResNet-50 with different InitialLearnRate and optimizers are seen in Table 10.

Table 10: Training Result of ResNet-50 for Maximum Training and Validation Accuracies

Optimizer	InitialLearnRate	Elapsed Time	Maximum Training Accuracy (%)	Corresponding Validation Accuracy (%)	Maximum Validation accuracy (%)	Corresponding Training Accuracy (%)
sgdm	0.001	2 hours 9 minutes 7 seconds	99.2188	23.8095		
sgdm	0.01	1 hour 11 minutes 52 seconds			26.2787	98.4375
adam	0.001	1 hour 22 minutes 4 seconds	97.6563		29.6296	
rmsprop	0.0001	1 hour 17 minute 27 seconds	98.4375	26.9841		
rmsprop	0.001	1 hour 17 minutes 40 seconds			27.3369	96.0938

- **Inception-v3**

In the Inception-v3 network, the maximum training accuracy value was achieved for adam optimizer. The maximum validation accuracy was also achieved for the adam optimizer.

In sgdm optimizer, the maximum training and validation accuracy were achieved for the InitialLearnRate of 0.01 for the elapsed trial time of 1 hour 32 minutes, and 13 seconds. The training and validation accuracy values were 98.4375% and 29.806%. The training accuracy and validation accuracy of InitialLearnRate 0.001 were 97.6563% and 24.515% with the elapsed trial time of 2 hours 9 minutes and 1 second. At the same time, the training accuracy and validation accuracy of InitialLearnRate 0.1 were 53.125% and 23.2804% taking 10 hours 18 minutes, and 35 seconds to complete the experiment.

In adam optimizer, the maximum training accuracy of value 99.2188% was achieved for the InitialLearnRate of 0.0001, for an elapsed trial time of 1 hour 25 minutes, and 59 seconds. The corresponding validation accuracy was 28.5714%. The maximum value of validation accuracy 31.3933% was achieved by the InitialLearnRate of 0.001 consuming the trial time of 1 hour 25 minutes and 28 seconds. The corresponding value of training accuracy was 94.5313%. Simultaneously, the training accuracy and validation accuracy of InitialLearnRate 0.01 were 36.7188% and 23.6332%. It took the trial time of 1 hour 25 minutes and 39 seconds.

In the rmsprop optimizer method, the maximum training and validation accuracy values were achieved for the InitialLearnRate of 0.0001 for an elapsed trial time of 2 hours 4 minutes, and 36 seconds. The value of maximum training and validation accuracy were 96.875% and 31.2169%. At the same time, the training accuracy and validation accuracy of InitialLearnRate 0.001 were 75.7813% and 21.6931%, taking 1 hour 38 minutes and 42 seconds to complete the experiment. But the training accuracy and validation accuracy varied for InitialLearnRate 0.01 with the respective values 31.25% and 23.1041%, and the elapsed trial time was 1 hour 37 minutes and 26 seconds.

Only the values of maximum training and validation accuracies achieved over the training of Inception-v3 with different InitialLearnRate and optimizers are found in Table 11.

Table 11: Training Result of Inception-v3 for Maximum Training and Validation Accuracies

Optimizer	InitialLearnRate	Elapsed Time	Maximum Training Accuracy (%)	Corresponding Validation Accuracy (%)	Maximum Validation accuracy (%)	Corresponding Training Accuracy (%)
sgdm	0.01	1 hour 32 minutes 13 seconds	98.4375	29.806		
adam	0.0001	1 hour 25 minutes 59 seconds	99.2188	28.5714		
adam	0.001	1 hour 25 minutes 28 seconds			31.3933	94.5313
rmsprop	0.0001	2 hours 4 minutes 36 seconds	96.875	31.2169		

- **ResNet-101**

In the ResNet-101 network, the maximum training accuracy value was achieved for sgdm optimizer. But the maximum validation accuracy was achieved for the rmsprop optimizer.

In sgdm optimizer, the value of maximum training accuracy of 99.2188% was achieved for the InitialLearnRate of 0.001 for an elapsed trial time of 13 hours 27 minutes, and 39 seconds. The corresponding validation accuracy was 25.9259%. The maximum validation accuracy with the value 28.5714% was achieved for the InitialLearnRate of 0.01 for an elapsed trial time of 4 hours 21 minutes and 57 seconds. The respective training accuracy value was 98.4375%. Subsequently, the training accuracy and validation accuracy of InitialLearnRate 0.1 were 32.8125% and 23.2804%. The elapsed trial time was 6 hours 42 minutes and 30 seconds.

In the adam optimizer method, the maximum training accuracy value was the same for both the InitialLearnRates 0.0001 and 0.001 and the value of the maximum training accuracy was 98.4375%. However, the maximum validation accuracy was achieved for the InitialLearnRate of 0.0001 and the value was 28.9242%. The trial time taken for the InitialLearnRates 0.0001 and 0.001 was 9 hours 9 minutes and 20 seconds and 9 hours 35 minutes and 9 seconds. At the same time, the validation accuracy of the InitialLearnRate 0.001 was 22.2222%. For the elapsed trial time of 9 hours 5 minutes and 35 seconds, InitialLearnRate 0.01 showed the training accuracy and validation accuracy of 31.25% and 22.3986%.

In rmsprop optimizer method, the maximum training accuracy value of 98.4375% was achieved for the InitialLearnRate of 0.001 for an elapsed trial time of 7 hours 1 minute, and 57 seconds. The corresponding validation accuracy value was 21.164%. The value of maximum validation accuracy of 29.4533% was achieved for the InitialLearnRate of 0.0001 for an elapsed trial time of 10 hours and 22 minutes. Comparatively, the time consumed is more than the other trials. The respective training accuracy value was 96.875%. The training accuracy and validation accuracy of InitialLearnRate 0.01 were 31.25% and 22.9277%. The time consumed was 9 hours 14 minutes and 41 seconds.

Table 12 shows only the values of maximum training and validation accuracies achieved over the training of ResNet-101 with different InitialLearnRate and optimizers.

Table 12: Training Result of ResNet-101 for Maximum Training and Validation Accuracies

Optimizer	**InitialLearnRate**	**Elapsed Time**	**Maximum Training Accuracy (%)**	**Corresponding Validation Accuracy (%)**	**Maximum Validation accuracy (%)**	**Corresponding Training Accuracy (%)**
sgdm	0.001	13 hours 27 minutes 39 seconds	99.2188	25.9259		
sgdm	0.01	4 hours 21 minutes 57 seconds			28.5714	98.4375
adam	0.0001	9 hours 9 minutes 20 seconds	98.4375		28.9242	
adam	0.001	9 hours 35 minutes 9 seconds	98.4375	22.2222		
rmsprop	0.001	7 hours 1 minute 57 seconds	98.4375	21.164		
rmsprop	0.0001	10 hours 22 minutes			29.4533	96.875

5 Results of the Confusion Matrix

In Matlab, the confusion matrix is the representation of the total number of observations by each cell of it. The row of the confusion matrix generally represents the true class, and the column represents the predicted class. The diagonal cell presents the observations classified correctly (Zhang et al., 2021 “b”) and the off-diagonal cells exhibit the observations classified incorrectly. A row-normalized row of each true class displays the summary of the percentage of correctly and incorrectly classified observations. Similarly, a column-normalized column of each predicted class displays the summary of the percentage of correctly and incorrectly classified observations. The evaluation of recalling i.e., the class-wise true positive rate is carried out along each row. The evaluation of precision, i.e., the class-wise positive predicted value is carried out along each column. The precision in the identification and classification of wastes should be higher (Li et al., 2022). The common metrics used to verify a classification model's performance are accuracy, precision, recall, and F1 score (Rachapudi & Lavanya Devi, 2021).

The prediction and recall percentage of the confusion matrices corresponding to the maximum validation accuracy achieved during the training under each optimizer are compared in the following part for the various pretrained neural networks. The prediction percentage value equal to and above 30% has been considered for evaluating the performance of the optimizers in the classification of the various classes of solid waste. Table 13 shows the diagonal cell values of the confusion matrices corresponding to the maximum validation accuracy values of AlexNet and SqueezeNet. Table 14 shows the diagonal cell values of the confusion matrices corresponding to the maximum validation accuracy values of ShuffleNet, Places365-GoogLeNet, GoogLeNet, and ResNet-18. Table 15 shows the diagonal cell values of the confusion matrices corresponding to the maximum validation accuracy values of MobileNet-v2, NasNet-Mobile, ResNet-50, Inception-v3, and ResNet-101.

Table 13: Diagonal Cell Values of the Confusion Matrix of AlexNet and SqueezeNet for the Trial with Maximum Validation Accuracy

			AlexNet			**SqueezeNet**		
Class No.	**Solid Waste Category**	**Number of Photos Given as Input**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**
1	Bottle caps	35	1	1	-	-	-	3
2	Cardboard and tetra packs	183	12	14	40	-	-	1
3	Cigarette butts	22	-	-	-	-	-	1
4	Coconut waste	46	3	3	4	-	-	3
5	Dairy packets	67	-	-	4	-	-	2
6	Dry flowers	22	-	-	-	-	-	-
7	Dry leaves	92	3	1	1	-	-	3
8	Dry twigs & branches	38	-	-	-	-	-	-
9	Fabric waste	38	2	-	-	-	-	-
10	Face masks	107	7	6	7	-	-	5
11	Footwear	60	5	4	3	-	-	6
12	Fruit waste	47	1	-	1	-	-	4
13	Glass bottles	17	-	-	-	-	-	-
14	Matchboxes	27	-	-	-	-	-	-
15	Paper cups	105	7	15	10	-	-	10
16	Paper waste	66	-	-	-	-	-	1
17	Plastic bottles	144	8	11	1	-	-	1
18	Plastic waste	261	16	7	9	-	46	15
19	Rope and thread	24	-	-	-	-	-	-
20	Vegetable waste	29	1	-	-	-	-	-
21	Wooden waste	26	-	-	-	-	-	-
22	Wrappers	436	66	64	1	131	90	92

Table 14: Diagonal Cell Values of the Confusion Matrix of ShuffleNet, Places365-GoogLeNet, GoogLeNet and ResNet-18 for the Trial with Maximum Validation Accuracy

			ShuffleNet			**Places365-GoogLeNet**			**GoogLeNet**			**ResNet-18**		
Class No.	**Solid Waste Category**	**Number of Photos Given as Input**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**
1	Bottle caps	35	-	-	2	1	1	-	7	5	6	1	2	1
2	Cardboard and tetra packs	183	10	11	7	9	11	27	7	12	8	8	11	12
3	Cigarette butts	22	-	-	-	-	-	-	1	1	-	-	-	-
4	Coconut waste	46	2	1	4	-	1	2	1	1	6	-	-	-
5	Dairy packets	67	2	2	3	1	3	2	5	3	2	1	6	3
6	Dry flowers	22	-	-	-	-	-	-	-	-	-	-	-	-
7	Dry leaves	92	4	4	3	3	2	-	8	6	7	8	10	9
8	Dry twigs & branches	38	1	-	-	-	-	1	-	-	-	-	-	1
9	Fabric waste	38	-	-	-	-	-	-	-	1	1	-	-	1
10	Face masks	107	4	4	10	7	6	9	14	14	19	7	13	11
11	Footwear	60	3	3	3	4	3	7	12	7	7	4	2	2
12	Fruit waste	47	-	-	-	3	1	1	2	3	4	1	2	3
13	Glass bottles	17	-	-	-	-	-	1	-	-	-	-	-	-
14	Matchboxes	27	-	-	-	-	-	-	-	-	-	-	1	-
15	Paper cups	105	7	4	13	14	10	10	17	11	15	8	14	10
16	Paper waste	66	-	-	-	-	1	-	-	1	1	-	-	-
17	Plastic bottles	144	6	9	2	5	7	5	12	12	5	5	10	12
18	Plastic waste	261	9	13	15	15	24	31	14	12	-	16	14	16
19	Rope and thread	24	-	-	1	-	-	-	1	1	1	-	-	-
20	Vegetable waste	29	-	-	-	-	-	-	-	-	-	-	-	-
21	Wooden waste	26	-	-	-	-	-	-	-	-	-	-	-	-
22	Wrappers	436	82	84	77	95	77	16	82	78	96	82	76	76

Table 15: Diagonal Cell Values of the Confusion Matrix of MobileNet-v2, NasNet-Mobile, ResNet-50, Inception-v3 and ResNet-101 for the Trial with Maximum Validation Accuracy

			MobileNet-v2			NasNet-Mobile			ResNet-50			Inception-v3			ResNet-101		
Class No.	**Solid Waste Category**	**Number of Photos Given as Input**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**	**sgdm**	**adam**	**rmsprop**
1	Bottle caps	35	-	2	4	1	4	3	2	3	6	-	4	-	-	-	1
2	Cardboard and tetra packs	183	13	14	3	7	10	15	12	9	15	13	9	13	8	7	11
3	Cigarette butts	22	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
4	Coconut waste	46	-	7	6	-	1	3	-	2	2	1	4	1	2	1	-
5	Dairy packets	67	1	3	1	7	5	3	3	4	2	4	5	5	4	3	2
6	Dry flowers	22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	Dry leaves	92	6	7	4	7	4	6	4	5	6	8	-	11	6	8	4
8	Dry twigs & branches	38	-	-	1	-	-	1	-	-	-	-	-	-	-	-	-
9	Fabric waste	38	-	-	-	-	-	1	-	-	-	-	1	-	-	-	1
10	Face masks	107	9	14	9	8	14	11	15	14	12	14	9	13	8	10	12
11	Footwear	60	4	6	8	3	5	3	3	1	2	7	5	8	5	3	6
12	Fruit waste	47	2	1	3	3	1	1	2	2	4	1	-	1	1	3	2
13	Glass bottles	17	1	-	-	-	-	-	-	-	-	-	2	-	-	-	-
14	Matchboxes	27	1	-	-	-	-	-	-	-	1	-	-	-	-	1	1
15	Paper cups	105	14	15	12	13	11	14	12	12	13	15	13	13	15	14	14
16	Paper waste	66	2	2	3	-	-	-	-	-	1	-	4	-	-	-	1
17	Plastic bottles	144	12	12	7	5	8	10	12	7	9	11	16	10	14	15	11
18	Plastic waste	261	13	16	14	15	14	14	13	15	16	11	16	17	18	17	11
19	Rope and thread	24	-	1	-	-	2	-	-	-	1	-	1	-	-	-	-
20	Vegetable waste	29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21	Wooden waste	26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22	Wrappers	436	86	88	78	80	89	69	86	87	85	85	89	85	81	82	90

- **AlexNet**

In the AlexNet network, the optimizers sgdm, and adam had similarly predicted the different classes of wastes. sgdm, and adam optimizers had predicted the wrappers mostly. The rmsprop optimizer showed a good prediction of Cardboard and tetra packs. Fig. 1 shows the confusion matrix corresponding to the maximum validation accuracy of 23.2804% achieved during the training with sgdm optimizer when compared to the training with other optimizers.

The prediction percentage of the waste classes such as footwear, and wrappers for the sgdm optimizer was 35.7%, and 38.4%. The recall percentage of the above waste classes was 27.8% and 50.4%.

The prediction percentage of the waste classes such as coconut waste, footwear, and wrappers for the adam optimizer was 60%, 30.8%, and 38.6% respectively. The recall percentage of the above waste classes was as follows: 21.4%, 22.2%, and 48.9%.

The prediction percentage of the waste classes such as cardboards and tetra packs, coconut waste, footwear, and wrappers for the rmsprop optimizer was 12.1%, 50%, 33.3%, and 100% respectively. The recall percentage of the above waste classes was 72.7%, 28.6%, 16.7%, and 0.8% respectively.

The classes cigarette butts, dry flowers, dry twigs and branches, glass bottles, matchboxes, paper waste, rope and thread, and wooden waste were never classified by the AlexNet network. 8 out of 22 classes were not classified.

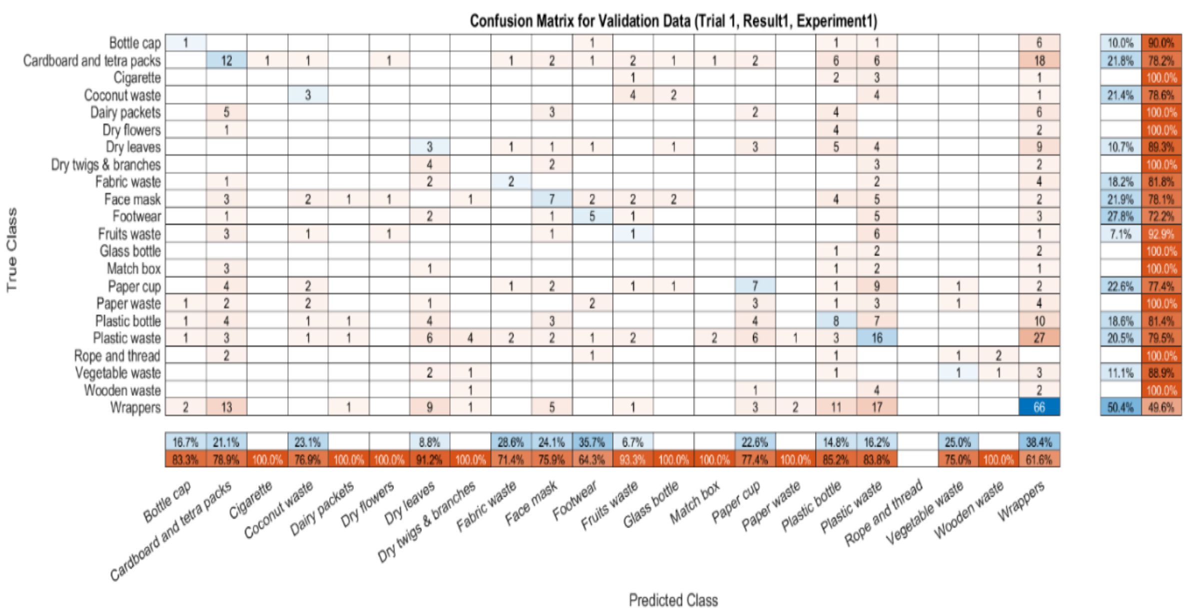


Figure 1: Confusion Matrix of AlexNet

- **SqueezeNet**

In the SqueezeNet network, the optimizer rmsprop predicted the different classes of wastes successfully when compared to other types of optimizers. The optimizer sgdm was able to classify only wrappers. The optimizer adam classified only plastic waste and wrappers. The confusion matrix corresponding to the maximum validation accuracy of 25.9259% achieved during the training with rmsprop optimizer when compared to the training with other optimizers, is shown in Fig.2.

The prediction percentage of the waste class wrappers for the sgdm optimizer was 23.1%. The recall percentage of the above waste classes was 100%.

The prediction percentage of the waste classes such as plastic waste, and wrappers for the adam optimizer was as follows: 13.8%, and 38.5%. The recall percentage of the above waste classes was 59% and 68.7%.

The prediction percentage of the waste classes such as coconut waste, face masks, footwear, paper cups, and wrappers for the rmsprop optimizer was as follows: 33.3%, 35.7%, 40%, 30.3%, and 38.8%. The recall percentage of the above waste classes was as follows: 21.4%, 15.6%, 33.3%, 32.3%, and 70.2%.

The classes dry flowers, dry twigs and branches, fabric waste, glass bottles, matchboxes, rope and thread, vegetable waste, and wooden waste were never classified by the SqueezeNet network. 8 out of 22 classes were not classified.

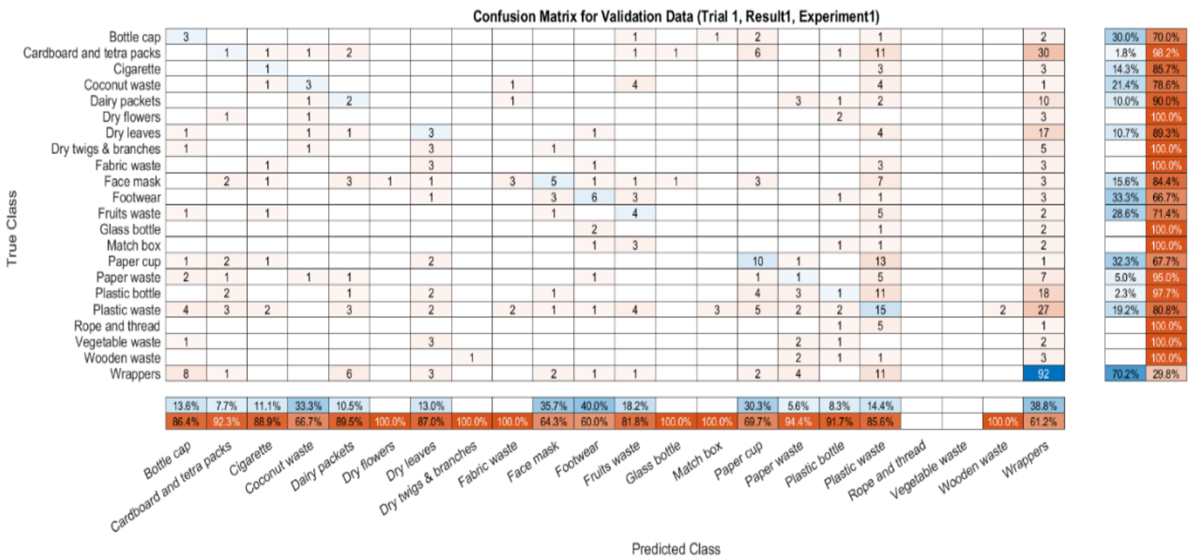


Figure 2: Confusion Matrix of SqueezeNet

- **ShuffleNet**

In the ShuffleNet network, the optimizers sgdm, adam, and rmsprop similarly predicted the different classes of wastes. In Fig. 3, the confusion matrix corresponding to the maximum validation accuracy of 24.6914% achieved during the training with rmsprop optimizer when compared to the training with other optimizers, is shown.

In sgdm optimizer, the prediction percentage of the classes of wastes such as coconut waste, footwear, and wrappers was 66.7%, 37.5% and 36.9% respectively. The corresponding recall percentages were 14.3%, 16.7% and 62.6%.

In adam optimizer, the prediction percentage of the classes of wastes such as footwear and wrappers was 42.9% and 38%. The corresponding recall percentages were 16.7%, and 64.1%.

In rmsprop optimizer, the prediction percentage of the classes of wastes such as coconut waste, face masks, footwear, paper cups, rope and thread, and wrappers was 44.4%, 38.5%, 30%, 36.1%, 50%, and

36% respectively. The corresponding recall percentages were 28.6%, 31.2%, 16.7%, 41.9%, 14.3%, and 58.8%.

The classes cigarette butts, dry flowers, fabric waste, fruit waste, glass bottles, matchboxes, paper waste, vegetable waste, and wooden waste were never classified by the ShuffleNet network. 9 out of 22 classes were not classified.

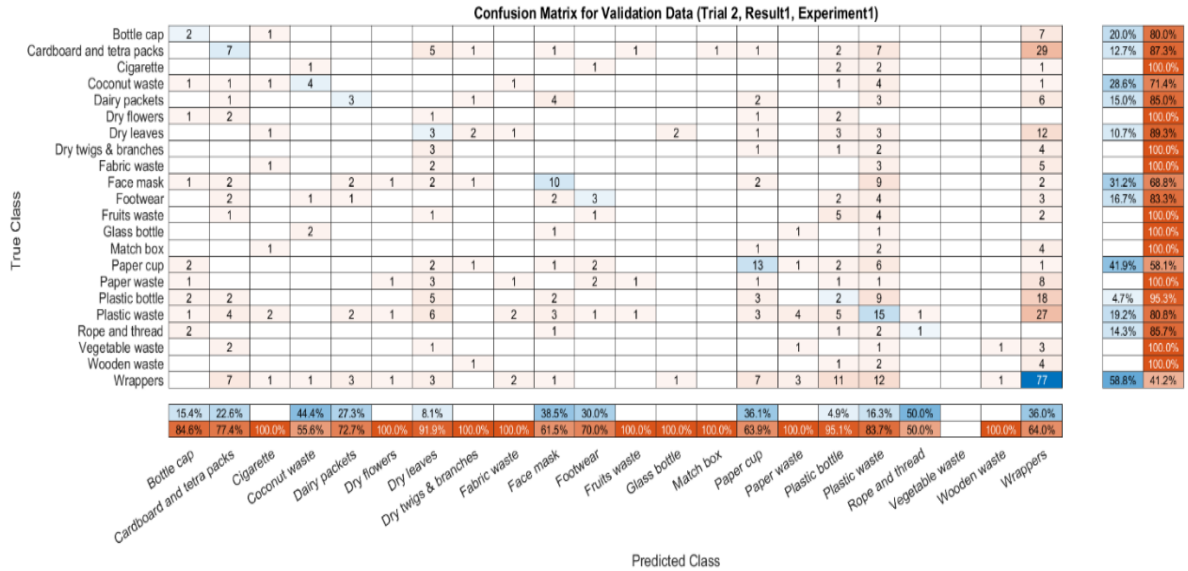


Figure 3: Confusion Matrix of ShuffleNet

- **Places365-GoogLeNet**

In the Places365-GoogLeNet network also, the optimizers sgdm, adam, and rmsprop had similarly predicted the different classes of wastes. The optimizer rmsprop showed poor classification for wrappers when compared to the other two optimizers. The confusion matrix corresponding to the maximum validation accuracy of 27.6896% achieved during the training with sgdm optimizer when compared to the training with other optimizers, is brought out in Fig. 4.

The prediction percentage of the waste classes such as bottle caps, face masks, footwear, fruit waste, paper cups, and wrappers for the sgdm optimizer was 100%, 50%, 40%, 33.3%, 45.2%, and 36.1% respectively. The recall percentages of the above waste classes were as follows: 10%, 21.9%, 22.2%, 21.4%, 42.2%, and 72.5%.

The prediction percentage of the waste classes such as dairy packets, footwear, paper cups, and wrappers for the adam optimizer was 33.3%, 37.5%, 58.8%, and 40.5% respectively. The recall percentage of the above waste classes was as follows: 15%, 16.7%, 32.3%, and 58.8%.

The prediction percentage of the waste classes such as cardboard and tetra packs, dry twigs and branches, footwear, paper cups, and wrappers for the rmsprop optimizer was as follows: 26%, 33.3%, 25%, 45.5%, and 64%. The recall percentage of the above waste classes was 49.1%, 9.1%, 38.9%, 32.3%, and 12.2% respectively.

The classes cigarette butts, dry flowers, fabric waste, matchboxes, rope and thread, vegetable waste, and wooden waste were never classified by the Places365-GoogLeNet network. 7 out of 22 classes were not classified.

- **GoogLeNet**

In GoogLeNet network also, sgdm, adam, and rmsprop optimizers had predicted the wrappers correctly. The prediction percentage and recall percentage of many of the waste classes were considerably better for these optimizers in GoogLeNet. Fig. 5 tells the confusion matrix corresponding to the maximum validation accuracy of 32.2751% achieved during the training with sgdm optimizer when compared to the training with other optimizers.

The prediction percentage for the waste classes such as bottle caps, dairy packets, face masks, footwear, paper cups, plastic bottles, and wrappers for the sgdm optimizer was as follows: 46.7%, 33.3%, 33.3%, 70.6%, 40.5%, 37.5%, and 43.6%. The recall percentage of the above waste classes was 70%, 25%, 43.8%, 66.7%, 54.8%, 27.9%, and 62.6% respectively.

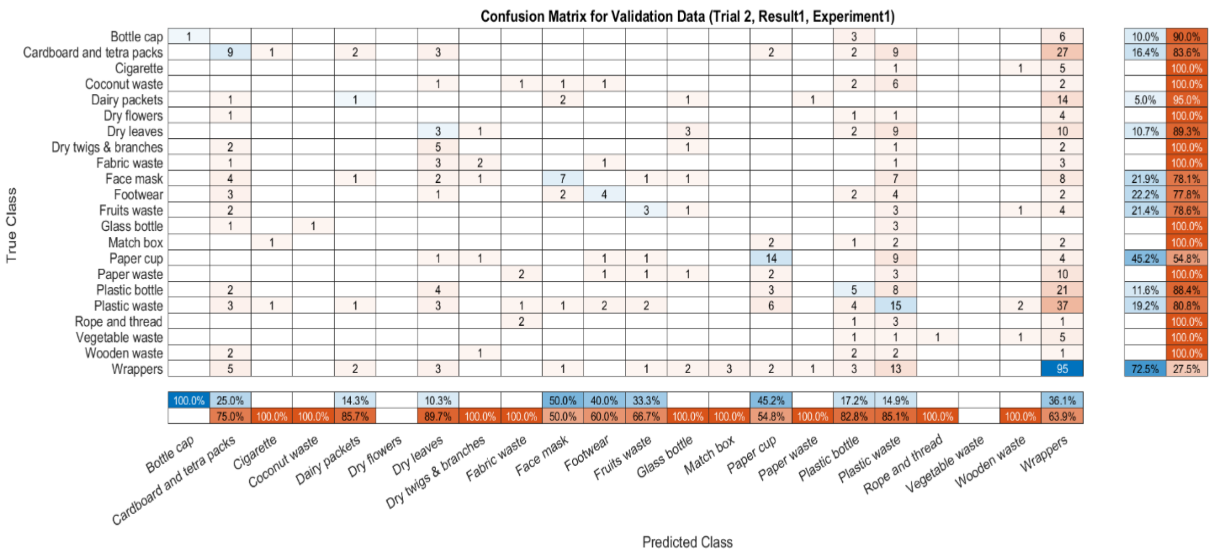


Figure 4: Confusion Matrix of Places365-GoogLeNet

The prediction percentage of the waste classes such as bottle caps, coconut waste, fabric waste, face masks, footwear, fruit waste, rope and thread, and wrappers for the adam optimizer was 35.7%, 33.3%, 100%, 33.3%, 87.5%, 30%, 33.3%, and 44.6% respectively. The recall percentage of the above waste classes was as follows: 50%, 7.1%, 9.1%, 43.8%, 38.9%, 21.4%, 14.3%, and 59.5%.

The prediction percentage of the waste classes such as bottle caps, coconut waste, footwear, fruit waste, rope and thread, and wrappers for the rmsprop optimizer was as follows: 42.9%, 40%, 63.6%, 50%, 33.3%, and 41.6%. The recall percentage of the above waste classes was 60%, 42.9%, 38.9%, 28.6%, 14.3%, and 73.3% respectively.

The classes dry flowers, dry twigs and branches, glass bottles, matchboxes, vegetable waste, and wooden waste were never classified by the GoogLeNet network. 6 out of 22 classes were not classified.

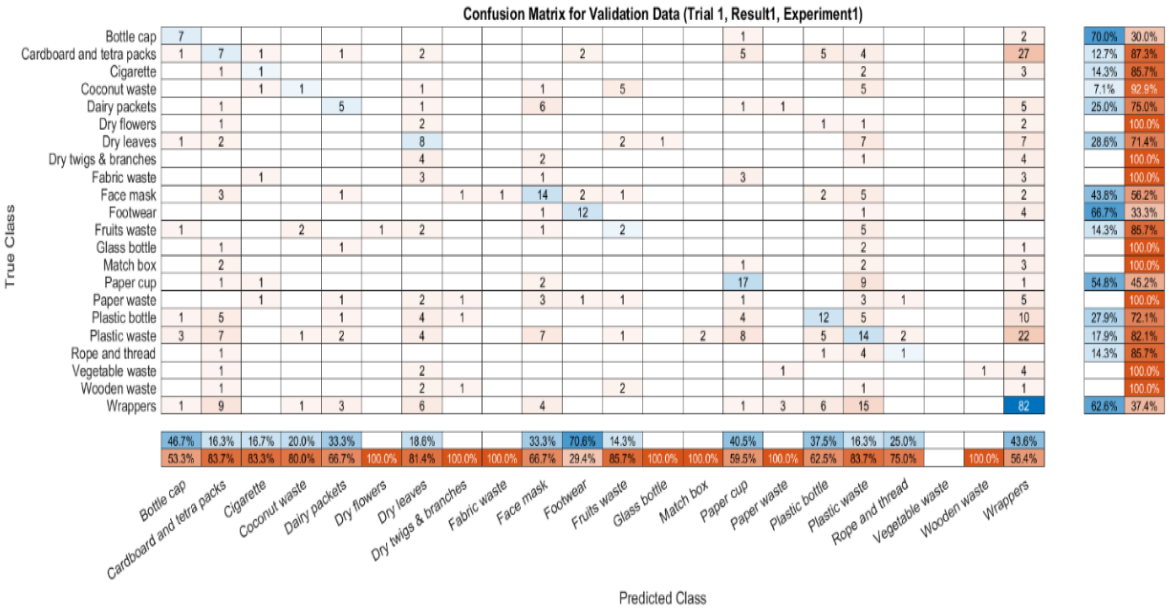


Figure 5: Confusion Matrix of GoogLeNet

- **ResNet-18**

The optimizers sgdm, adam and rmsprop optimizers had similarly predicted the different classes of wastes in ResNet-18 network. The most predicted class of waste was wrappers like that of the other networks.

The confusion matrix corresponding to the maximum validation accuracy of 28.3951% achieved during the training with adam optimizer when compared to the training with other optimizers, is delineated in Fig. 6.

The prediction percentage of the waste classes such as bottle caps, footwear, paper cups, and wrappers for the sgdm optimizer was as follows: 33.3%, 40%, 32%, and 39.2%. The recall percentage of the above waste classes was 10%, 22.2%, 25.8%, and 62.6% respectively.

The prediction percentage of the waste classes such as bottle caps, dairy packets, face masks, matchboxes, paper cups, and wrappers for the adam optimizer was 66.7%, 35.3%, 28.3%, 33.3%, 42.4%, and 49% respectively. Also, the recall percentage of the above waste classes was 20%, 30%, 40.6%, 12.5%, 45.2%, and 58% respectively.

The prediction percentage of the waste classes such as bottle caps, dairy packets, fabric waste, paper cups, and wrappers for the rmsprop optimizer was as follows: 33.3%, 30%, 33.3%, 47.6%, and 46.3%. The recall percentage of the above waste classes was 10%, 15%, 9.1%, 32.3%, and 58% respectively.

The classes cigarette butts, coconut waste, dry flowers, glass bottles, paper waste, rope and thread, vegetable waste and wooden waste were never classified by the ResNet-18 network. 8 out of 22 classes were not classified.

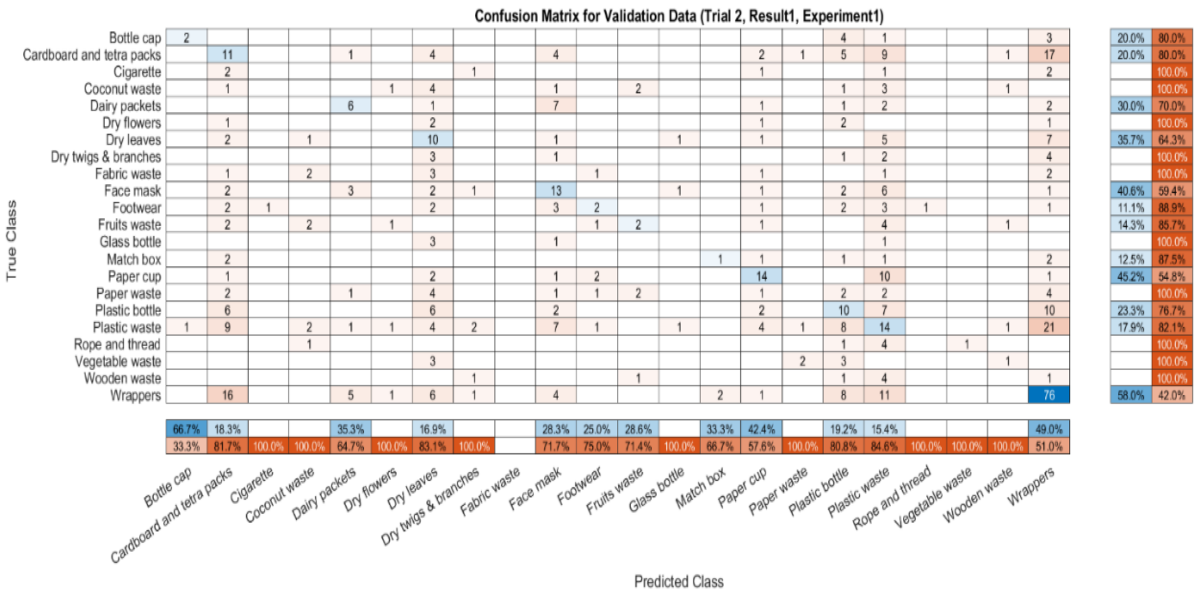


Figure 6: Confusion Matrix of ResNet-18

- **MobileNet-v2**

The optimizers sgdm, adam, and rmsprop had similarly predicted the different classes of wastes in the MobileNet-v2 network also. As in other networks, the most predicted class of waste by the MobileNet-v2 network also was wrappers. The confusion matrix corresponding to the maximum validation accuracy of 33.157% achieved during the training with adam optimizer when compared to the training with other optimizers, is depicted in Fig. 7.

In sgdm optimizer, the prediction percentage of the classes of wastes such as footwear, matchboxes, paper cups, paper waste, and wrappers was 44.4%, 33.3%, 32.6%, 40%, and 40.4% respectively. The corresponding recall percentages were 22.2%, 12.5%, 45.2%, 10%, and 65.6%.

In adam optimizer, the prediction percentage of the classes of wastes such as bottle caps, coconut waste, face masks, footwear, paper cups, rope and thread, and wrappers was 50%, 100%, 46.7%, 60%, 33.3%, 100%, and 46.8% respectively. The corresponding recall percentages were 20%, 50%, 43.8%, 33.3%, 48.4%, 14.3%, and 67.2%.

In rmsprop optimizer, the prediction percentage of the classes of wastes such as bottle caps, coconut waste, face masks, footwear, paper cups, and wrappers was 40%, 42.9%, 42.9%, 50%, 50%, and 41.5% respectively.

The corresponding recall percentages were 40%, 42.9%, 28.1%, 44.4%, 38.7%, and 59.5%.

The classes cigarette butts, dry flowers, fabric waste, vegetable waste, and wooden waste were never classified by the MobileNet-v2 network. 5 out of 22 classes were not classified.

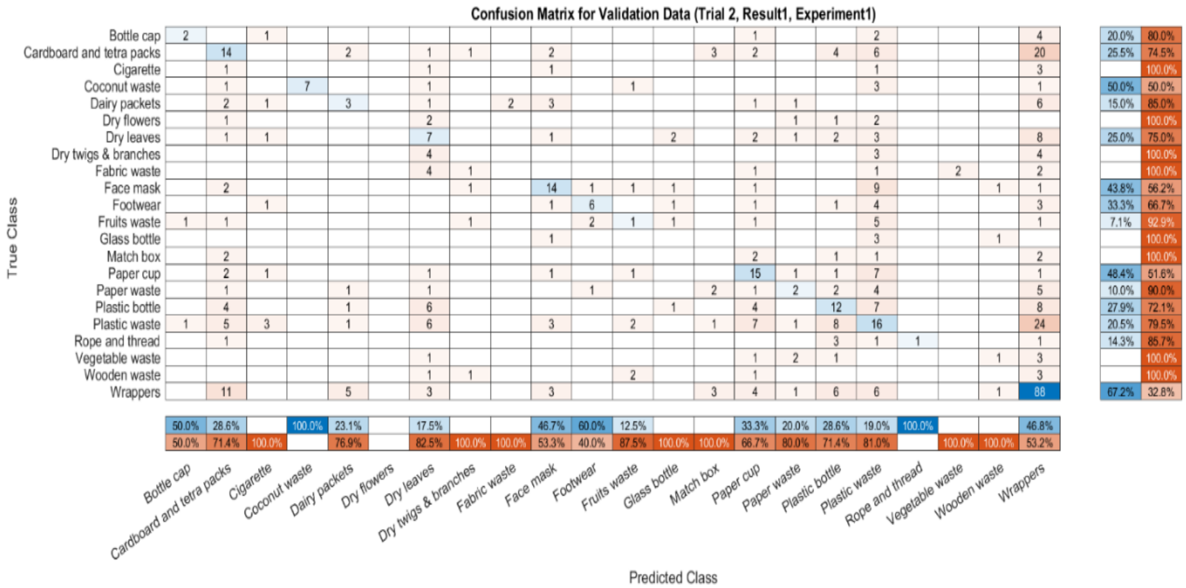


Figure 7: Confusion Matrix of MobileNet-v2

- **NasNet-Mobile**

The optimizers sgdm, adam, and rmsprop had similarly predicted the different classes of wastes in NasNet-Mobile network also. The optimizer rmsprop predicted more waste classes when compared to other optimizers. The most predicted class of waste was wrappers like that of the other networks. Fig. 8 portrays the confusion matrix corresponding to the maximum validation accuracy of 29.6296% achieved during the training with adam optimizer when compared to the training with other optimizers.

In this network, in sgdm optimizer, the prediction percentage of the classes of wastes such as bottle caps, dairy packets, paper cups, and wrappers was 33.3%, 29.2%, 29.5%, and 39.8% respectively. The corresponding recall percentages were 10%, 35%, 41.9%, and 61.1%.

In the NasNet-Mobile network, in adam optimizer, the prediction percentage of the classes of wastes such as bottle caps, coconut waste, footwear, paper cups, rope and thread, and wrappers was 36.4%, 33.3%, 45.5%, 31.4%, 66.7%, and 48.9% respectively. The corresponding recall percentages were 40%, 7.1%, 27.8%, 35.5%, 28.6%, and 67.9%.

In rmsprop optimizer, the prediction percentage of the classes of wastes such as bottle caps, cigarette butts, coconut waste, dairy packets, face masks, footwear, paper cups, and wrappers were 37.5%, 100%, 42.9%, 30%, 39.3%, 50%, 35%, and 46% respectively. The corresponding recall percentages were 30%, 14.3%, 21.4%, 15%, 34.4%, 16.7%, 45.2%, and 52.7%.

The classes dry flowers, glass bottles, matchboxes, paper waste, vegetable waste, and wooden waste were never classified by the NasNet-Mobile. 6 out of 22 classes were not classified.

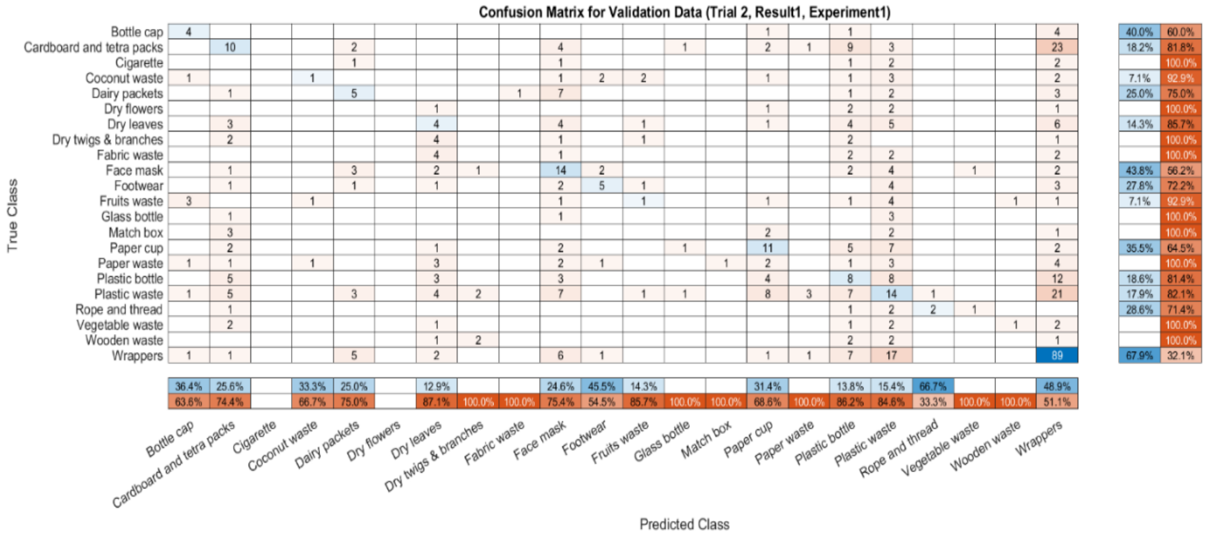


Figure 8: Confusion Matrix of NasNet-Mobile

- **ResNet-50**

The optimizers sgdm, adam, and rmsprop have similarly predicted the different classes of wastes in ResNet-50 also. The highest predicted class of waste was wrappers like that of the other network models. The confusion matrix corresponding to the maximum validation accuracy of 30.8642% achieved during the training with rmsprop optimizer when compared to the training with other optimizers, is given in Fig. 9.

In sgdm optimizer, the prediction percentage of the classes of wastes such as bottle caps, face masks, fruit waste, paper cups, and wrappers was 50%, 51.7%, 33.3%, 36.4%, and 42% respectively. The corresponding recall percentages were 20%, 46.9%, 14.3%, 38.7%, and 65.6%.

In adam optimizer, the prediction percentage of the classes of wastes such as bottle caps, coconut waste, face masks, paper cups, and wrappers was 50%, 40%, 40%, 44.4%, and 41.6% respectively. The corresponding recall percentages were 30%, 14.3%, 43.8%, 38.7%, and 66.4%.

In the ResNet-50 network, in rmsprop optimizer, the prediction percentage of the classes of wastes such as bottle caps, coconut waste, face masks, footwear, fruit waste, matchboxes, paper cups, paper waste, rope and thread and wrappers was 66.7%, 40%, 52.2%, 40%, 44.4%, 100%, 36.1%, 33.3%, 100%, and 42.3% respectively. The corresponding recall percentages were 60%, 14.3%, 37.5%, 11.1%, 28.6%, 12.5%, 41.9%, 5%, 14.3%, and 64.9%.

The classes cigarette butts, dry flowers, dry twigs and branches, fabric waste, glass bottles, vegetable waste, and wooden waste were never classified by the ResNet-50 network. 7 out of 22 classes were not classified.

- **Inception-v3**

In the network Inception-v3 also, the optimizers sgdm, adam, and rmsprop had similarly predicted the different classes of wastes. The optimizer adam had predicted more classes of waste when compared to other optimizers. The highest predicted class of waste was wrappers as done by the other networks. The confusion matrix corresponding to the maximum validation accuracy of 31.3933% achieved during the training with adam optimizer when compared to the training with other optimizers is putforth in Fig. 10.

In sgdm optimizer, the prediction percentage of the classes of wastes such as dairy packets, face masks, footwear, and wrappers was 40%, 42.4%, 70%, and 45.7% respectively. The corresponding recall percentages were 20%, 43.8%, 38.9%, and 64.9%.

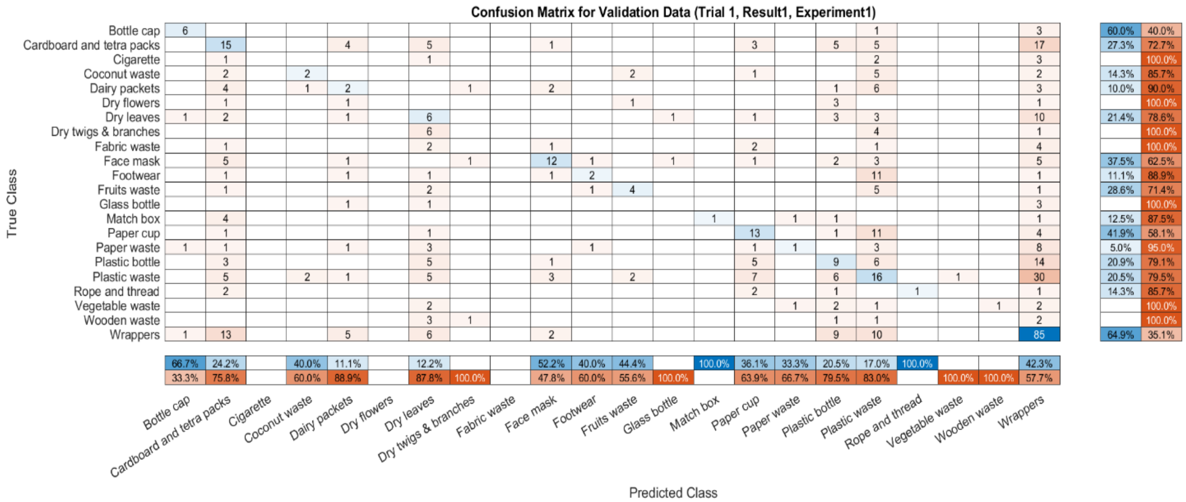


Figure 9: Confusion Matrix of ResNet-50

In adam optimizer, the prediction percentage of the classes of wastes such as bottle caps, dairy packets, face masks, footwear, paper cups, rope and thread, wrappers were 57.1%, 45.5%, 47.4%, 41.7%, 46.4%, 50%, and 43.6% respectively. The corresponding recall percentages were 40%, 25%, 28.1%, 27.8%, 41.9%, 14.3%, and 67.9%.

In rmsprop optimizer, the prediction percentage of the classes of wastes such as coconut waste, dairy packets, face masks, footwear, and wrappers were 33.3%, 50%, 40.6%, 88.9%, and 45.7% respectively. The corresponding recall percentages were 7.1%, 25%, 40.6%, 44.4%, and 64.9%.

The classes cigarette butts, dry flowers, dry twigs & branches, matchboxes, vegetable waste, and wooden waste were never classified by the Inception-v3 network. 6 out of 22 classes were not classified.

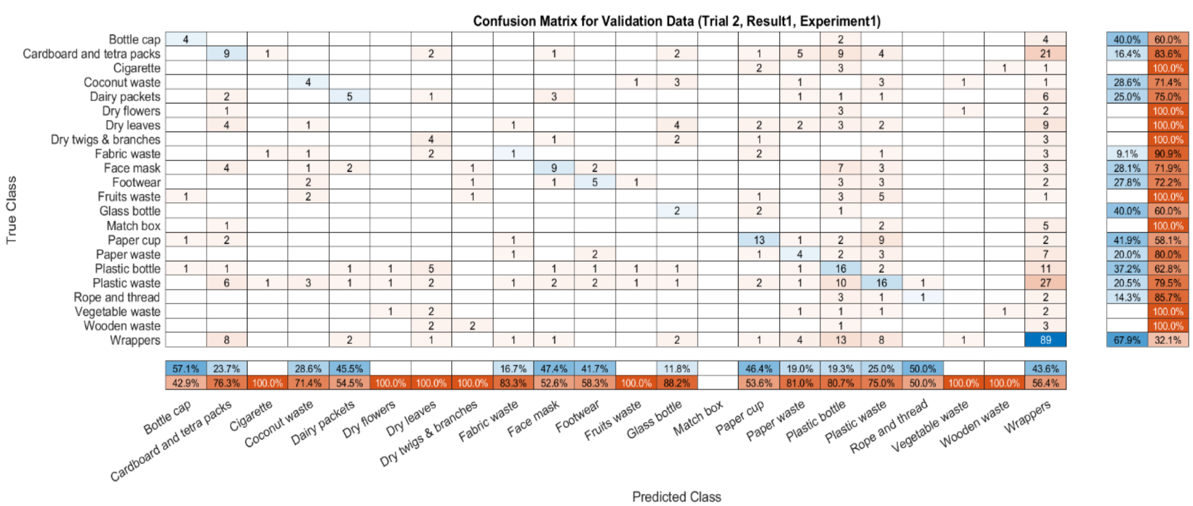


Figure 10: Confusion Matrix of Inception-v3

- **ResNet-101**

In the ResNet-101 network also, the optimizers sgdm, adam, and rmsprop had similarly predicted the different classes of wastes. Like that of the other networks, the most predicted class of waste was wrappers. The confusion matrix corresponding to the maximum validation accuracy of 29.4533% achieved during the training with the rmsprop optimizer when compared to the training with other optimizers, is revealed in Fig. 11.

In sgdm optimizer, the prediction percentage of the classes of wastes such as coconut waste, footwear, paper cups, and wrappers was 40%, 55.6%, 40.5%, and 45.3% respectively. The corresponding recall percentages were 14.3%, 27.8%, 48.4%, and 61.8%.

In adam optimizer, the prediction percentage of the classes of wastes such as footwear, fruit waste, paper cup, and wrappers was 60%, 33.3%, 31.1%, and 43.9% respectively. The corresponding recall percentages were 16.7%, 21.4%, 45.2%, and 62.6%.

In rmsprop optimizer, the prediction percentage of the classes of wastes such as face mask, footwear, and wrappers was 36.4%, 60%, and 47.1% respectively. The corresponding recall percentages were 37.5%, 33.3%, and 68.7%.

The classes cigarette butts, dry flowers, dry twigs and branches, glass bottles, rope and thread, vegetable waste, and wooden waste were never classified by the ResNet-101 network. 7 out of 22 classes were not classified.

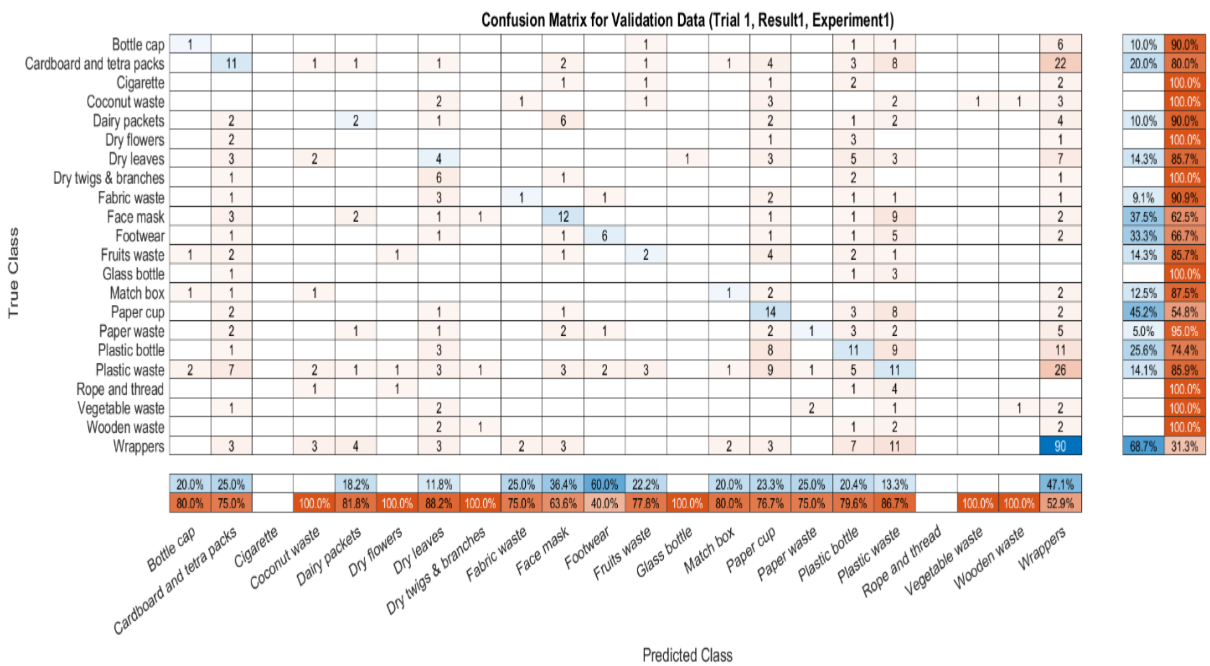


Figure 11: Confusion Matrix of ResNet-101

6 Discussion

Table 16 shows the percentage of maximum validation accuracy achieved by each CNN pretrained model. Each CNN pretrained model underwent three trials of training for each optimizer. For sgdm optimizer, trials were conducted for the InitialLearnRates 0.001, 0.01, and 0.1. For adam optimizer, the trials were conducted for the InitialLearnRates 0.0001, 0.001, and 0.01. For rmsprop optimizer, trials were conducted for the InitialLearnRates 0.0001, 0.001, and 0.01. MobileNet-v2 network, with the adam optimizer, achieved the highest validation accuracy for InitialLearnRate of 0.001 with a reasonable elapsed time of 40 minutes and 37 seconds. The highest validation accuracy achieved was 33.157%. The validation accuracy of GoogLeNet comes next to the MobileNet- v2. The elapsed trial time was more than that of MobileNet-v2. NasNet-Mobile and ResNet-101 showed similar performance and the values of validation accuracy were 29.6296% and 29.4533%. The elapsed trial time of ResNet-101 was 10 hours 22 minutes and very high when compared to that of other CNN models. AlexNet shows the least validation accuracy of value 23.2804%.

It shows that the InitialLearnRate 0.001 achieves better results when compared to the other values of InitialLearnRate. 7 CNN models have achieved the maximum validation accuracy with an InitialLearnRate of 0.001. When the optimizers are considered, all the optimizers perform equally. The optimizer sgdm shows good results with GoogLeNet, Places365-GoogLeNet, and AlexNet while the optimizer adam shows good results with MobileNet-v2, Inception-v3, NasNet-Mobile, and ResNet-18. The optimizer rmsprop shows a good result with ResNet-50, ResNet-101, SqueezeNet, and ShuffleNet. The maximum trial time taken was 10 hours and 22 minutes to train the network ResNet-101. The least trial time taken was 9 minutes and 43 seconds to train the SqueezeNet network.

Among all the networks ShuffleNet shows low performance with respect to the classification of 22 classes of given solid wastes. It could classify only 13 out of 22 classes of waste. Next to ShuffleNet network, AlexNet, SqueezeNet, and ResNet-18 were able to classify 14 classes of waste.

Then Places365-GoogLeNet, ResNet-50, and ResNet-101 were able to classify 15 classes of waste. Next to them, GoogLeNet, NasNet-Mobile, and Inception-v3 were able to classify 16 classes of waste. Finally, at the top, MobileNet-v2 was able to classify 17 out of 22 classes of waste. The wastes Cigarette, Dry flowers, Fabric waste, Vegetable waste, and Wooden waste were not able to be classified by any of the 11 selected network models.

Bottle caps were better predicted by GoogLeNet and ResNet-18 when compared to other networks. Footwear and wrappers were classified widely by all the networks with reasonably good prediction accuracy. Coconut waste was able to be predicted by AlexNet, SqueezeNet, ShuffleNet, GoogLeNet, MobileNet-v2, NasNet-Mobile, ResNet-50, and ResNet-101. Places365-GoogLeNet, GoogLeNet, ResNet-18, NasNet-Mobile, and Inception-v3 predicted dairy packets. Dry twigs and branches were able to be predicted by only Places365-GoogLeNet. Fabric waste was predicted by GoogLeNet, and ResNet-18. Face masks were able to be predicted by SqueezeNet, ShuffleNet, Places365-GoogLeNet, GoogLeNet, MobileNet-v2, NasNet-Mobile, ResNet-50, Inception-v3, and ResNet-101. Fruit waste could be predicted by Places365-GoogLeNet, GoogLeNet, ResNet-50, and ResNet-101. Matchboxes were able to be predicted by ResNet-18, MobileNet-v2, and ResNet-50. Paper cups were predicted by SqueezeNet, ShuffleNet, Places365-GoogLeNet, GoogLeNet, ResNet-18, MobileNet-v2, NasNet-Mobile, ResNet-50, Inception-v3, and ResNet-101. Paper waste was able to be predicted by only two networks MobileNet-v2, and ResNet-50. Plastic bottles were able to be predicted only by GoogLeNet. Rope and thread were predicted by ShuffleNet, GoogLeNet, MobileNet-v2, NasNet-Mobile, ResNet-50, and Inception-v3.

Table 16: Maximum Validation Accuracy of different Networks Arranged in Descending Order

CNN Model	Optimizer	InitialLearnRate	Trial Elapsed Time	Validation Accuracy (%)
MobileNet-v2	adam	0.001	40 minutes 37 seconds	33.157
GoogLeNet	sgdm	0.001	1 hour 4 minutes 8 seconds	32.2751
Inception-v3	adam	0.001	1 hour 25 minutes 28 seconds	31.3933
ResNet-50	rmsprop	0.0001	1 hour 1 minute 8 seconds	30.8642
NasNet-Mobile	adam	0.001	1 hour 22 minutes 4 seconds	29.6296
ResNet-101	rmsprop	0.0001	10 hours 22 minutes	29.4533
ResNet-18	adam	0.001	12 minutes 25 seconds	28.3951
Places365-GoogLeNet	sgdm	0.01	28 minutes 23 seconds	27.6896
SqueezeNet	rmsprop	0.0001	9 minutes 43 seconds	25.9259
ShuffleNet	rmsprop	0.001	25 minutes 31 seconds	24.6914
AlexNet	sgdm	0.001	33 minutes 34 seconds	23.2804

7 Conclusions

In a token attempt to implement automation in the process of waste collection and segregation, the present research work was conducted. In our work, we have trained our image dataset of different varieties of solid waste which were taken with their real background in the sunlight in 11 pre-trained neural networks. We have trained all 11 CNN models with the three types of optimizers sgdm, adam, and rmsprop for different InitialLearnRate. 70% of images were used as training datasets and 30% of images were used as validation datasets. It was found that with the optimizer sgdm the maximum validation accuracies were achieved for GoogLeNet, Places365-GoogLeNet, and AlexNet. At the same time, with the optimizer adam, the maximum validation accuracies were achieved for MobileNet-v2, Inception-v3, NasNet-Mobile, and ResNet-18. Finally, with the optimizer rmsprop, ResNet-50, ResNet-101, SqueezeNet, and ShuffleNet have achieved the maximum validation accuracies. The highest validation accuracy value of 33.157% was achieved by the MobileNet-v2 and the least validation accuracy value of 23.2804% was achieved by AlexNet. Among the 22 classes of waste, footwear and wrappers were classified by most of the networks. MobileNet-v2 performed well in terms of the classification of domestic solid waste. It was able to classify 17 classes of waste.

The network ShuffleNet showed the lowest performance in terms of the classification of domestic solid waste. It was able to classify only 13 classes of waste. The classes of solid wastes such as cigarette butts, dry flowers, fabric waste, vegetable waste, and wooden waste were never classified by any of the 11 selected pre-trained networks. The problem of incorrect classification, non-classification, and lower validation accuracy can be rectified and improved by increasing the number of images in each of the classes. In our work, we have used images of waste in their real and natural background without following any separate image augmentation procedure. So, each waste of interest is surrounded by other classes of waste and other disturbing factors from the environment like buildings and structures, vehicles, trees, plants, grass, debris, etc. In most research works, the images of waste of interest in artificial, clear, and plain backgrounds have been used without real-world environmental disturbances. As a challenge, we have addressed this case in our work. Our future work is to create our network which can show better identification rate using the images of similar kinds of solid waste taken in real backgrounds, which will be a tiny, innovative, upward step in the process of automated waste segregation which in turn will encourage more such research works and initiate new waste management systems in densely populated places with large waste generation.

Acknowledgments

I sincerely thank my guide, Dr. M. Giriraj, for his continued support in bringing out this research paper. I thank my mother, Dr. A. Brintha Devi, for her invaluable efforts in taking the entire set of photographs of different classes of waste littered all along the streets that are used in this research work. I also thank my mother for her motivation and language support. I also thank my family members for their constant support.

References

[1] Aarthi, R., & Rishma, G. (2023). A vision based approach to localize waste objects and geometric features exaction for robotic manipulation. *Procedia Computer Science, 218,* 1342-1352. https://doi.org/10.1016/j.procs.2023.01.113

[2] Abinandan, V., Thilagavathi, G., Gokulram, M., & Sankarraj Aditya, M. (2022). Garbage Segregator Using Deep Learning. *In Smart Data Intelligence: Proceedings of ICSMDI,* 233-247. https://doi.org/10.1007/978-981-19-3311-0_20

[3] Ahmad, M. B., Soomro, U., Muqeet, M., & Ahmed, Z. (2021). Adsorption of Indigo Carmine dye onto the surface-modified adsorbent prepared from municipal waste and simulation using deep neural network. *Journal of Hazardous Materials, 408,* 124433. https://doi.org/10.1016/j.jhazmat.2020.124433

[4] Akbarimajd, A., Hoertel, N., Hussain, M. A., Neshat, A. A., Marhamati, M., Bakhtoor, M., & Momeny, M. (2022). Learning-to-augment incorporated noise-robust deep CNN for detection of COVID-19 in noisy X-ray images. *Journal of Computational Science, 63,* 101763. https://doi.org/10.1016/j.jocs.2022.101763

[5] Altikat, A., Gulbe, A., & Altikat, S. (2022). Intelligent solid waste classification using deep convolutional neural networks. *International Journal of Environmental Science and Technology*, *19*(3), 1285–1292. https://doi.org/10.1007/s13762-021-03179-4

[6] Bard, J., Bidgoli, A., & Chi, W.W. (2019). Image Classification for Robotic Plastering with Convolutional Neural Network. *In Robotic Fabrication in Architecture, Art and Design,* 3–15. https://doi.org/10.1007/978-3-319-92294-2_1

[7] Brinker, T.J., Hekler, A., Enk, A.H., Berking, C., Haferkamp, S., Hauschild, A., Weichenthal, M., Klode, J., Schadendorf, D., Holland-Letz, T., Von Kalle, C., Fröhling, S., Schilling, B., & Utikal, J.S. (2019). Deep neural networks are superior to dermatologists in melanoma image classification. *European Journal of Cancer*, *119*, 11–17. https://doi.org/10.1016/j.ejca.2019.05.023

[8] Chen, J., Lu, W., & Xue, F. (2021). "Looking beneath the surface": A visual-physical feature hybrid approach for unattended gauging of construction waste composition. *Journal of environmental management*, *286*, 112233. https://doi.org/10.1016/j.jenvman.2021.112233

[9] Chen, J., Wang, H., Wang, S., He, E., Zhang, T., & Wang, L. (2022). Convolutional neural network with transfer learning approach for detection of unfavorable driving state using phase coherence image. *Expert Systems with Applications*, *187,* 116016. https://doi.org/10.1016/j.eswa.2021.116016

[10] Chen, W.C., Liu, P.Y., Lai, C.C., & Lin, Y.H. (2022). Identification of environmental microorganism using optimally fine-tuned convolutional neural network. *Environmental Research*, *206*, 112610. https://doi.org/10.1016/j.envres.2021.112610

[11] Chen, Y., Sun, J., Bi, S., Meng, C., & Guo, F. (2021). Multi-objective solid waste classification and identification model based on transfer learning method. *Journal of Material Cycles and Waste Management*, *23*(6), 2179–2191. https://doi.org/10.1007/s10163-021-01283-8

[12] Chen, Z., Yang, J., Chen, L., & Jiao, H. (2022). Garbage classification system based on improved ShuffleNet v2. *Resources, Conservation and Recycling*, *178*, 106090. https://doi.org/10.1016/j.resconrec.2021.106090

[13] Das, P.K., & Meher, S. (2021). An efficient deep Convolutional Neural Network based detection and classification of Acute Lymphoblastic Leukemia. *Expert Systems with Applications*, *183*. https://doi.org/10.1016/j.eswa.2021.115311

[14] Dorj, U.O., Lee, K.K., Choi, J.Y., & Lee, M. (2018). The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications*, *77*(8), 9909–9924. https://doi.org/10.1007/s11042-018-5714-1

[15] El Boustani, A., Aatila, M., El Bachari, E., & El Oirrak, A. (2020). MRI brain images classification using convolutional neural networks. *In Advanced Intelligent Systems for Sustainable Development (AI2SD'2019) Volume 4-Advanced Intelligent Systems for Applied Computing Sciences*, 308-320. https://doi.org/10.1007/978-3-030-36674-2_32

[16] Fu, Y., & Aldrich, C. (2019). Flotation froth image recognition with convolutional neural networks. *Minerals Engineering*, *132*, 183-190. https://doi.org/10.1016/j.mineng.2018.12.011

[17] Islam, S.M.Y., & Alam, Md. G.R. (2022). *Computer Vision-based Waste Detection and Classification for Garbage Management and Recycling*, 143–164. https://doi.org/10.1007/978-981-19-2445-3_10

[18] Iwashita, M., Otsuka, Y., & Katoh, M. (2022). Predicting the particle size distribution of fine-grained and sandy soils using deep learning for classifying recovered soils separated from tsunami deposits. *Journal of Material Cycles and Waste Management*, *24*(4), 1304-1316. https://doi.org/10.1007/s10163-022-01404-x

[19] Jain, A., Khetriwal, V., Daga, H., & Tripathy, B.K. (2022). Smart Garbage Classification. *In International Conference on Innovations in Intelligent Computing and Communications*, 113-124. https://doi.org/10.1007/978-3-031-23233-6_8

[20] Kadhim, M.A., & Abed, M.H. (2020). Convolutional neural network for satellite image classification. *Intelligent Information and Database Systems: Recent Developments 11*, 165-178. https://doi.org/10.1007/978-3-030-14132-5_13

[21] Kim, J.G., Jang, S.C., Kang, I.S., Lee, D.J., Lee, J.W., & Park, H.S. (2020). A study on object recognition using deep learning for optimizing categorization of radioactive waste. *Progress in Nuclear Energy*, *130*, 103528. https://doi.org/10.1016/j.pnucene.2020.103528

[22] Kurtulmuş, F. (2021). Identification of sunflower seeds with deep convolutional neural networks. *Journal of Food Measurement and Characterization*, *15*(2), 1024-1033. https://doi.org/10.1007/s11694-020-00707-7

[23] Lakhani, P. (2017). Deep Convolutional Neural Networks for Endotracheal Tube Position and X-ray Image Classification: Challenges and Opportunities. *Journal of Digital Imaging*, *30*(4), 460–468. https://doi.org/10.1007/s10278-017-9980-7

[24] Li, J., Chen, J., Sheng, B., Li, P., Yang, P., Feng, D.D., & Qi, J. (2022). Automatic Detection and Classification System of Domestic Waste via Multimodel Cascaded Convolutional Neural Network. *IEEE Transactions on Industrial Informatics*, *18*(1), 163–173. https://doi.org/10.1109/TII.2021.3085669

[25] Lin, K., Zhao, Y., Wang, L., Shi, W., Cui, F., & Zhou, T. (2023). MSWNet: A visual deep machine learning method adopting transfer learning based upon ResNet 50 for municipal solid waste sorting. *Frontiers of Environmental Science and Engineering*, *17*(6). https://doi.org/10.1007/s11783-023-1677-1

[26] Madappa, E.A., Joshi, A.A., Karthik, P.K., Shashank, E., & Veeresh, J. (2020). Dexterous Trashbot. *Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough: Latest Trends in AI*, 87-98. https://doi.org/10.1007/978-3-030-38445-6_7

[27] Majchrowska, S., Mikołajczyk, A., Ferlin, M., Klawikowska, Z., Plantykow, M.A., Kwasigroch, A., & Majek, K. (2022). Deep learning-based waste detection in natural and urban environments. *Waste Management*, *138*, 274–284.

[28] Mao, W.L., Chen, W.C., Wang, C.T., & Lin, Y.H. (2021). Recycling waste classification using optimized convolutional neural network. *Resources, Conservation and Recycling*, *164*, 105132. https://doi.org/10.1016/j.resconrec.2020.105132

[29] Melinte, D.O., Travediu, A.M., & Dumitriu, D.N. (2020). Deep convolutional neural networks object detector for real-time waste identification. *Applied Sciences*, *10*(20), 7301. https://doi.org/10.3390/app10207301

[30] Momeny, M., Latif, A.M., Sarram, M.A., Sheikhpour, R., & Zhang, Y.D. (2021). A noise robust convolutional neural network for image classification. *Results in Engineering*, *10*, 100225. https://doi.org/10.1016/j.rineng.2021.100225

[31] Mookkaiah, S.S., Thangavelu, G., Hebbar, R., Haldar, N., & Singh, H. (2022). Design and development of smart Internet of Things–based solid waste management system using computer vision. *Environmental Science and Pollution Research*, *29*(43), 64871-64885. https://doi.org/10.1007/s11356-022-20428-2

[32] Muangnak, N., Sirawattananon, C., Oza, M.G., & Sukthanapirat, R. (2021). Comparative study considering garbage classification using in-depth learning techniques. *In International Conference on Computing and Information Technology*, 169-180. https://doi.org/10.1007/978-3-030-79757-7_17

[33] Nguyen, T., Ngoc, L.C., Son, T.N., Minh, D.H., & Dinh, T.H.P. (2022). Deep Learning for Green Smart Environment. *In Computational Intelligence Techniques for Green Smart Cities*, 271-293. https://doi.org/10.1007/978-3-030-96429-0_13

[34] Rachapudi, V., & Lavanya Devi, G. (2021). Improved convolutional neural network based histopathological image classification. *Evolutionary Intelligence*, *14*(3), 1337–1343. https://doi.org/10.1007/s12065-020-00367-y

[35] Raghu, S., Sriraam, N., Temel, Y., Rao, S.V., & Kubben, P.L. (2020). EEG based multi-class seizure type classification using convolutional neural network and transfer learning. *Neural Networks*, *124*, 202–212. https://doi.org/10.1016/j.neunet.2020.01.017

[36] Rahman, M.W., Islam, R., Hasan, A., Bithi, N.I., Hasan, M.M., & Rahman, M.M. (2022). Intelligent waste management system using deep learning with IoT. *Journal of King Saud University - Computer and Information Sciences*, *34*(5), 2072–2087. https://doi.org/10.1016/j.jksuci.2020.08.016

[37] Ramírez, I., Cuesta-Infante, A., Pantrigo, J.J., Montemayor, A.S., Moreno, J.L., Alonso, V., Anguita, G., & Palombarani, L. (2020). Convolutional neural networks for computer vision-based detection and recognition of dumpsters. *Neural Computing and Applications*, *32*(17), 13203–13211. https://doi.org/10.1007/s00521-018-3390-8

[38] Rehman, A., Naz, S., Razzak, M.I., Akram, F., & Imran, M. (2020). A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning. *Circuits, Systems, and Signal Processing*, *39*(2), 757–775. https://doi.org/10.1007/s00034-019-01246-3

[39] Ruiz, V., Sánchez, Á., Vélez, J.F., & Raducanu, B. (2019). Automatic image-based waste classification. *In From Bioinspired Systems and Biomedical Applications to Machine Learning: 8th International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2019*, 422-431. https://doi.org/10.1007/978-3-030-19651-6_41

[40] Sharma, N., Jain, V., & Mishra, A. (2018). An Analysis of Convolutional Neural Networks for Image Classification. *Procedia Computer Science*, *132*, 377–384. https://doi.org/10.1016/j.procs.2018.05.198

[41] Sun, Y., & Gu, Z. (2022). Using computer vision to recognize construction material: A Trustworthy Dataset Perspective. *Resources, Conservation and Recycling*, *183*, 106362. https://doi.org/10.1016/j.resconrec.2022.106362

[42] Thenmozhi, K., & Srinivasulu Reddy, U. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture*, *164*, 104906. https://doi.org/10.1016/j.compag.2019.104906

[43] Toğaçar, M., Ergen, B., & Cömert, Z. (2020). Waste classification using AutoEncoder network with integrated feature selection method in convolutional neural network models. *Measurement: Journal of the International Measurement Confederation*, *153*. https://doi.org/10.1016/j.measurement.2019.107459

[44] Unal, Y., Taspinar, Y.S., Cinar, I., Kursun, R., & Koklu, M. (2022). Application of Pre-Trained Deep Convolutional Neural Networks for Coffee Beans Species Detection. *Food Analytical Methods*, *15*(12), 3232–3243. https://doi.org/10.1007/s12161-022-02362-8

[45] Wang, J., Li, Y., Feng, H., Ren, L., Du, X., & Wu, J. (2020). Common pests image recognition based on deep convolutional neural network. *Computers and Electronics in Agriculture*, *179*. https://doi.org/10.1016/j.compag.2020.105834

[46] Zhang, Q., Yang, Q., Zhang, X., Bao, Q., Su, J., & Liu, X. (2021). Waste image classification based on transfer learning and convolutional neural network. *Waste Management*, *135*, 150–157. https://doi.org/10.1016/j.wasman.2021.08.038

[47] Zhang, Q., Zhang, X., Mu, X., Wang, Z., Tian, R., Wang, X., & Liu, X. (2021). Recyclable waste image recognition based on deep learning. *In Resources, Conservation and Recycling* (Vol. 171). Elsevier B.V. https://doi.org/10.1016/j.resconrec.2021.105636

Authors Profile

Y. Vicknesh Kumar, author of this research paper, hailing from Virudhunagar district, Tamilnadu State, India, was awarded his Bachelor's degree in Mechanical Engineering in Mohamed Sathak Engineering College, Kilakarai, Ramanathapuram district, affiliated to Madurai Kamaraj University. He did his post-graduation in Computer Integrated Manufacturing in Sona College of Technology, Salem, affiliated to Anna University. In the field of industry, he rendered his service at Amalgamated Engineers Private Limited, Chennai, as a Sales Engineer for Saint Gobain Abrasives and at TAFE Madurai, Vadippatti Plant, as a paint shop supervisor under a contractor. In the field of teaching, he served at V.S.B. Engineering College, Karudayampalayam, Karur as a lecturer and at Velammal Engineering College, Chennai as Assistant Professor II. He has been working as a lecturer at the University of Technology and Applied Sciences, Ibra, Sultanate of Oman since 2012. He has guided many innovative student projects in areas pertinent to the native environmental conditions. His interest is mainly aligned with the development of new mechanisms for innovative projects in the field of Mechanical Engineering. Currently, he is pursuing his Ph.D. at Vellore Institute of Technology, Vellore under the guidance of Dr.M. Giriraj, Professor, Department of Design and Automation, School of Mechanical Engineering.

Dr.M. Giriraj is a seasoned professor at Vellore Institute of Technology, holding a PhD from Anna University. With over two decades of experience spanning both industry and academia, his expertise lies in the realm of secure data enhancement within manufacturing execution systems, utilizing cloud paradigms. His innovative contributions have been instrumental in advancing technological solutions tailored for industrial applications, making significant strides in optimizing efficiency and security within manufacturing processes.