Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

Enhancing Vehicle Classification Accuracy: A Convolutional Neural Network (CNN) Based Model

[1]Rashid Iqbal, [2]Dr. Abdul Basit
[1]MS Scholar, Department of Computer Science and IT, University of Balochistan, Quetta-Pakistan.
[2]Assistant Professor, Department of Computer Science and IT, University of Balochistan, Quetta-Pakistan.
dso.rashid@gmail.com, drabasit@um.uob.edu.pk

DOI: 10.5281/zenodo.11074270

ABSTRACT

Using the convolutional neural network (CNN) fine-tuned method, this article introduces a vehicle categorization system. The system's goal is to properly categorize popular vehicle types in the domestic market, which will help with traffic control, monitoring, and traffic accident prevention. The efficacy of VGG-16 and Inception V3 architectures is demonstrated by their evaluation of a real-world dataset consisting of 2000 photos of vehicles. While VGG-16 attains an accuracy of 99.11%, Inception V3 reaches an accuracy of 96.43%. In terms of overall accuracy, VGG-16 outperforms Inception V3, highlighting the importance of architectural decisions in achieving accurate vehicle classification. The suggested technique significantly improves computer vision applications in the domain of vehicle classification, making valuable contributions to traffic management and accident prevention efforts.

Keywords: Vehicle Detection, Transfer Learning, Fine-tuned, CNN, Inception V3, VGG-16.

Cite as: Rashid Iqbal, & Abdul Basit. (2024). Enhancing Vehicle Classification Accuracy: A Convolutional Neural Network (CNN) Based Model. *LC International Journal of STEM, 5(1)*, 1–11. https://doi.org/10.5281/zenodo.11074270

INTRODUCTION

Daily occurrences of traffic offenses are steadily rising due to the exponential increase in the number of automobiles in emerging countries. Because of this, commuters, pedestrians, and drivers all face increased risks of traffic jams and accidents. To solve these problems, a reliable traffic monitoring system is required. The Intelligent Transportation System (ITS) relies on vehicle classification for road traffic monitoring, which is crucial for efficient transportation planning, control, and the creation of autonomous vehicles. Vehicle classification is essential for monitoring traffic flow, identifying risky driving behaviors, and detecting signal infractions in unregulated road environments with poor adherence to traffic norms. In addition, it helps with forecasting transportation needs, increasing road safety, and figuring out whether new road infrastructure is needed based on the kinds and amounts of vehicles that use the roads. Scholarly literature has offered a variety of approaches to vehicle classification.

The use of deep learning techniques, such as Convolutional Neural Networks (CNNs), for vehicle categorization has grown in popularity in the past several years. We used a convolutional neural network (CNN) with tweaked VGG-16 and Inception V3 architectures to classify cars. A great deal of progress has been achieved because of this study. First, we claim that the best models for vehicle classifications in unstructured contexts are models of fine-tuned convolutional neural networks, like Inception V3 and VGG-16. In order to make the dataset even more diverse, we have also included technologies that enhance information and modify images. To train their models, Inception V3 and VGG-16 make use of the specialized vehicle dataset. Finally, the results show that optimizing VGG-16 and Inception V3 CNNs makes them the best at vehicle classification. Part 2 details the Literature Review.

As stated in Section 3, the Research Methodology is detailed. Section 4 presents the results and analysis of the experiments, and Section 5 wraps up the paper.

LITERATURE REVIEW

Maity, S., Saha, D., Singh, P. K., and Sarkar, R. (2024) developed the JUIVCDv1 dataset for Indian vehicle classification. They evaluated eight pre-trained CNN models on the dataset, with Xception, InceptionV3, and DenseNet121 achieving the highest accuracy scores of 0.94, 0.93, and 0.92 respectively. The authors further improved classification performance by combining these models using ensemble techniques. The ensemble models, based on majority voting, weighted average, and sum rule, achieved accuracy scores of 0.95, 0.94, and 0.94 respectively. This research highlights the importance of tailored datasets and effective model ensembling for accurate vehicle classification.

Classifying vehicles is an essential component of Intelligent Transportation Systems' capacity to monitor traffic. There have been a number of methods investigated in the literature that aim to solve the problems with vehicle classification. Within the framework of Intelligent Transportation Systems, Khairi et al. (2023) explored deep convolutional neural networks (DCNNs) as a possible method for granular vehicle categorization. Considerations such as lighting impacts, variations in vehicle size and design, weather circumstances, and similarities within and between vehicle classes were explicitly examined in their study. Three separate datasets were used to assess the efficacy of various DCNN models: ResNet-50, VGG-19, Inception-v3, and MobileNet-v2. These datasets were BMW-10, Stanford Cars, and PAK Cars. Because of their more intricate architecture, ResNet-50 and VGG-19 fared better than MobileNet-v2, the smallest model.

To improve vehicle recognition and type classification in traffic video surveillance, Ennehar, B. C., and Samra, B. (2023) proposed a unique CNN framework known as the master-slave convolutional deep architecture. This architecture consists of two networks and aims to reduce the search area and processing time by activating the slave network only after the master network detects cars. By combining deep and shallow neural networks, the framework facilitates effective information sharing between the networks. Experiments conducted on 3200 images showed that vehicles could be accurately detected (92% TP, 95% TN) and classified with an average accuracy of 93.38%. Of particular note is the fact that vehicles attained the maximum categorization rate of 98.63%.

A hybrid model for vehicle categorization, proposed by Alghamdi, A. S., Saeed, A., Kamran, M., Mursi, K. T., and Almukadi, W. S. (2023), combines genetic algorithms with a pre-trained Convolutional Neural Network (CNN). The model utilizes the Stanford car dataset and successfully classifies eight different types of vehicles. By employing deep feature fusion and evolutionary feature selection, a Cubic Support Vector Machine (SVM) achieved an impressive accuracy rate of 99.7 percent. This hybrid approach demonstrates the potential application of the model in various fields, such as security, traffic monitoring, and autonomous car technology.

A model for vehicle classification in intelligent transportation systems was suggested by Muhib, R. B., Ahmad, I. S., and Boufama, B. (2023). This model incorporates transfer learning and data augmentation techniques. Built on ResNet-50, the model achieves successful classification of various types of vehicles with an impressive accuracy rate of 90.07%. Further improvement is observed by integrating additional categorization blocks, resulting in enhanced performance. The proposed model surpasses baseline approaches and pre-trained deep learning systems such as VGG-16 and AlexNet, which achieved an accuracy of 82.5% according to the provided comparison. The implementation of smart vehicle counters and number plate identification significantly enhances the efficiency of the transportation system.

Within the context of an Intelligent Transportation System (ITS), Avianto, D., Harjoko, A., and Afiahayati (2022) tackled the classification problem of differentiating across linked vehicle models and manufacturers. A multi-task learning-based, all-encompassing classifier based on CNNs was proposed by them. Car types and manufactures were identified using features extracted from images using the VGG-16 architecture, which were subsequently split into several branches. The suggested method accomplished the remarkable feat of 98.73% accuracy for automobile model designation and 97.69% accuracy for car make authentication on the InaV-Dash dataset. Based on the results, the proposed approach is superior to the status quo when it comes to differentiating between identical vehicle models and manufactures.

In their study, Hasan, M. M., Wang, Z., Hussain, M. A. I., and Fatima, K. (2021) presented a model that leverages deep learning and transfer learning techniques to classify native automobiles in Bangladesh. Utilizing a ResNet-50 residual network as the backbone, augmented with additional classification blocks, the model achieved an impressive 98.00% accuracy in vehicle classification. The researchers collected a dataset of 10,440 photos representing 13 distinct vehicle categories commonly found in Bangladesh. Notably, the proposed model outperformed both traditional methods and pre-trained artificial intelligence models like VGG-16 and AlexNet. Performance evaluation metrics, including recall, precision, and F1-Score, were employed to assess the model's effectiveness.

Using a Transfer Learning (TL) approach, Baghdadi, S., and Aboutabit, N. (2021) tackled the problem of vehicle view categorization. The researchers employed a pre-trained Convolutional Neural Network (CNN) model called AlexNet to classify photos into 1000 object categories. This AlexNet model had been trained on a large dataset consisting of more than a million photographs. Encouraging results were obtained by leveraging the learned information of AlexNet to distinguish between different vehicle perspectives. Two studies were conducted, one focusing on TL using the AlexNet model, while the other involved combining fully connected layers with a Support Vector Machine (SVM) classifier. The results demonstrated that the improved accuracy rates observed in the second experiment were attributed to the efficacy of the Transfer Learning method.

Butt, M. A., Khattak, A. M., Shafique, S., Hayat, B., Abid, S., Kim, K. I., and Adnan, A. (2021) implemented a vehicle classification system that achieved remarkable results on their unique dataset. The system demonstrated impressive performance, achieving a recall of 99.56%, precision of 99.65%, and accuracy of 99.68%. Remarkably, even in the presence of poor illumination, the system accurately classified automobiles, as evidenced by these metrics. In summary, the suggested method proves to be more efficient for instant vehicle categorization in modern transportation systems, and the comparative analysis revealed its superiority over existing alternatives.

Bin Che Mansor, M. A. H., Kamal, N. A. M., Baharom, M. H. B., and Zainol, M. A. B. (2021) investigated the task of emergency vehicle type classification through the utilization of a Convolutional Neural Network (CNN) approach. The study addressed the challenge of road congestion, which

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

hampers the prompt response of emergency vehicles. With the aim of optimizing efficiency, the proposed method focused on identifying and categorizing emergency vehicles operating on public roadways. By employing a custom configuration of convolutional layers and adjusting filter sizes in the pre-trained VGG-16 model, the system achieved an impressive accuracy rate of approximately 95%. This outcome unequivocally demonstrates the system's ability to accurately classify emergency vehicles, underscoring the potential of CNNs to enhance intelligent transportation systems.

In their article "Improving Vehicle Classification and Detection with Deep Neural Networks," Mohamed, A. K., Ibrahim, A. K., Akram, A., Ibrahim, A., and Gamal, M. (2020) introduced a deep-learning method for detecting and classifying vehicles. Improving the performance of the popular object identification model YOLOv4, with a focus on car recognition, was the primary goal of the researchers. The improved model's ability to correctly recognise autos was validated in field trials carried out at different sites in Egypt. With ResNet50 and VGG-16 as primary classification frameworks, the authors trained their models on various datasets, leading to notable enhancements in classification accuracy and Mean Average Precision (MAP). The study's major contribution to the field of vehicle classification and detection is highly advantageous for applications like ADAS.

Meng, W., and Tia, M. (2020) explored the topic of "Unmanned Aerial Vehicle Classification and Detection Based on Deep Transfer Learning" in their study. The article discusses the security risks associated with unauthorized unmanned aerial vehicle (UAV) operations and proposes the utilization of transfer learning techniques to enhance UAV image recognition and designation. The authors evaluated three classification models, namely Inception V3, ResNet 101, and VGG-16, along with two detection mechanisms, Faster RCNN and SSD, using transfer learning. Particularly noteworthy is the Inception V3 model, which achieved an impressive recall rate of 96.48% while exhibiting substantial improvements in accuracy and precision.

Tabassum, Ullah, Al-Nur, and Shatabda (2020) developed a CNN-based transfer learning approach for native vehicle classification on Bangladeshi roads. They utilized the YOLO framework and a dataset of 9000 annotated images containing 15 unique vehicles. The proposed method achieved an impressive 73% IoU accuracy in detecting native vehicles, demonstrating its potential for enhancing the traffic management system in developing countries. This research highlights the effectiveness of CNNs and transfer learning in addressing the challenges of vehicle identification in unique traffic scenarios.

RESEARCH METHODOLOGY

One form of deep neural network that is frequently used for visual image processing in deep learning is the convolutional neural network (CNN). Using an image as input, a convolutional neural network (CNN) can distinguish between various regions of the image and assign a value to each. Because of its great accuracy, CNN is used for picture classification analysis. A fully connected layer, with all neurons coupled and data processed, is the result of the CNN model's hierarchical structure, which generates a funnel-shaped network. Recent years have seen remarkable progress in AI when it comes to computer vision tasks, particularly with convolutional neural networks (CNNs) and other deep learning approaches helping to bridge the gap between human and machine skills.

In this study, the cars were detected using two CNN-tuned algorithms for vehicle classification, namely VGG-16 and Inception V3. The whole schematic of the proposed system is shown in Figure 1.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

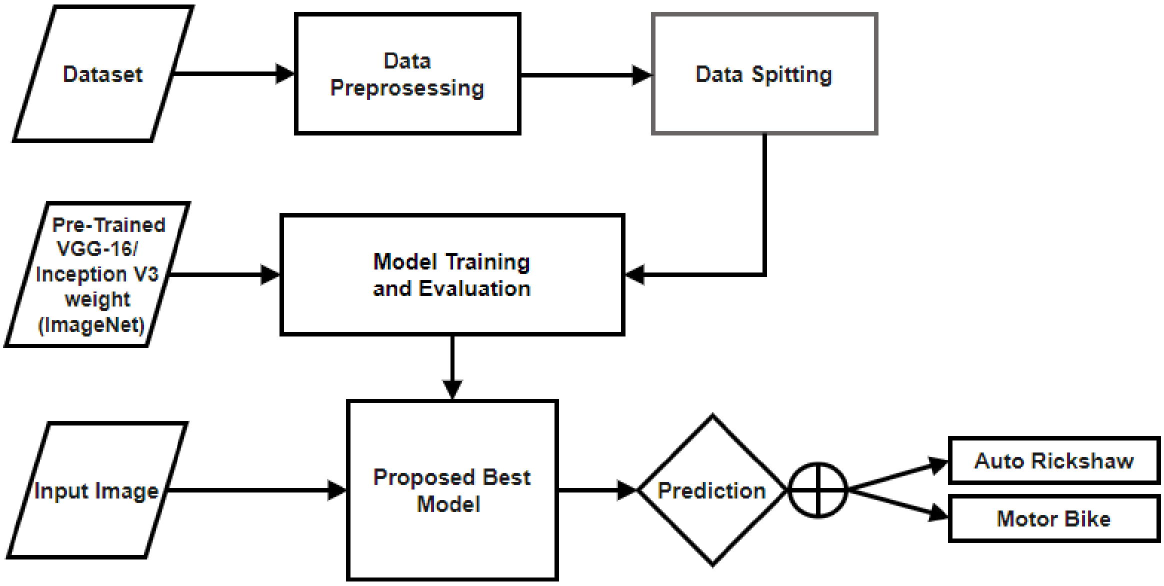

Figure 1: Vehicle-Classification Model

Our project involves the development of a vehicle classification model that makes use of a pre-processed bespoke vehicle dataset. To make the model more versatile and applicable to a wider range of situations, the pre-processing processes include using several data augmentation strategies. The following methods are included in this category: shearing, cropping, rotating, and flipping.

Data splitting occurs after pre-processing and involves randomly dividing the dataset in half, creating two sets: one for training and one for validation, with an 80/20 split. A suggested VGG-16 and Inception V3 model is developed using the training set, with pre-trained weights from the ImageNet dataset incorporated. Using this transfer learning method, the model may make use of the insights gained from studying a diverse array of objects in the ImageNet dataset.

A well-trained model that can categorize cars is obtained when the assessment and training phases are finished. When you feed a picture into the model, it can guess what kind of car it is. The trained model can use the representations gained during training to provide accurate predictions using the input image. The input image of the vehicle may be quickly and accurately classified using this prediction procedure.

Data Collection

We drew on the custom automobiles dataset for this study. There are two thousand pictures in the collection. The photos have been sorted into these groups. Both Auto Rickshaw and Motorbike fall under this category. The sizes of the photos vary. A total of 1,600 photos have been set aside for training, while 400 photographs have been reserved for validation. The sizes of the photos are the same.

Preprocessing

In this step, we picked out high-quality photos that included the specified target classes and removed any inaccurate or dirty data. The photos were cropped so that the model could train on the data with as little unnecessary information as possible. Additional picture adjustments were made to match the precise dimensions given by the model.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

There is a high probability of overfitting when the dataset used to train a deep learning model is inadequate. To improve the dataset and make sure it worked with the model, we used data augmentation techniques like Gaussian blur, flipping, rotation, and Gaussian noise. We were able to quadruple the size of our dataset by adding more photographs via data augmentation. Each training cycle involved retrieving a new batch of photos and applying data augmentation at random to each one.

Proposed Convolution Neural Network (CNN) Architecture

VGG-16

Zhao (2022) discussed the VGG16 model, which is a deep convolutional neural network (CNN) comprising 16 layers. The initial 13 layers consist of convolutional layers with the ReLU activation function. These layers are accompanied by five Max pooling operations, which serve to reduce the spatial information size and enhance computational efficiency. Following the convolutional layers, there are three fully connected layers responsible for flattening the high-dimensional features within the data. The output from these layers is then fed into the Softmax function, which normalizes the values and converts them into a probability distribution.VGG-16 Model's overall architecture is shown in Figure 2.

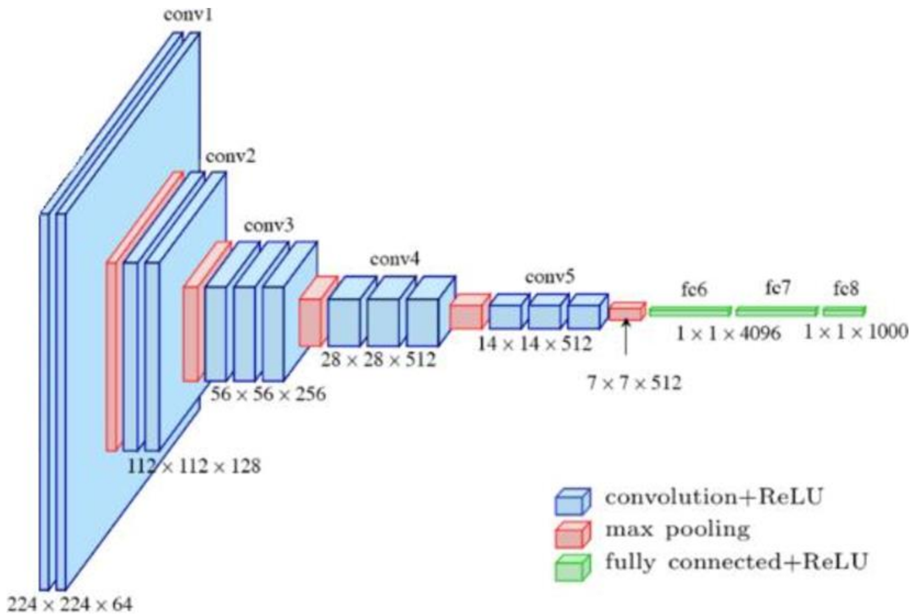

Figure 2: VGG-16 Model Architecture

Inception V3

A pre-trained model for Inception-V3 was proposed by Szegedy, Vanhoucke, Ioffe, Shlens, and Wojna (2016). This particular model, developed by a prominent hardware specialist in the field, consists of over 20 million parameters. The building blocks of the model incorporate symmetrical and asymmetrical components, featuring multiple layers of fully connected, concatenated, average, or max pooling data, as well as dropouts and convolutional layers. Batch normalization is often applied to the input of the activation layer in this model. For classification purposes, the Soft-max algorithm is utilized. Figure 3 provides a schematic representation of the Inception-V3 model.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

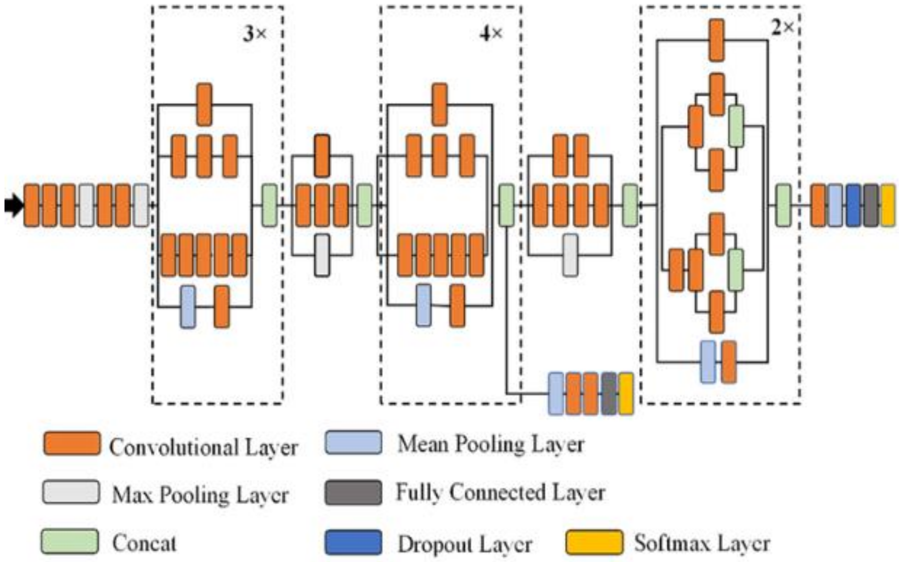

Figure 3: Inception V3 Model Architecture.

Evaluating performance using performance matrix:

After the testing and training phases were over, we evaluated the two models' performance based on accuracy, f1-score, recall, and precision. The following formulas were utilized:

$$\text{Accuracy} = \frac{tp+tn}{tp+tn+fp+fn} \quad (1)$$

$$\text{Precision} = \frac{tp}{tp+fp} \quad (2)$$

$$\text{Recall} = \frac{\text{tp}}{\text{tp+fn}} \quad (3)$$

$$f1 - score = 2 * \frac{(P*R)}{(P+R)} \quad (4)$$

EXPERIMENTAL RESULT ANALYSIS

Data preprocessing, training, and assessment and testing are the three primary steps in training the suggested vehicle classification model. Data preparation entails splitting the data into training and validation sets after cropping photos to extract specific classes. On top of that, we've standardized the picture dimensions at 320 × 320. The training set is made up of an 80:20 random split, while the validation set is made up of 20% randomly selected photos from the training set. Using a custom dataset, the tests on the dataset compare the Inception V3 model to the fine-tuned VGG-16 model and assess the performance of both models with data augmentation.

In order to incorporate our trainable layer, we have frozen the foundation layer for all transfer learning designs. We have set all the trainable layers to False. Preserving the learned representations of the pre-trained model is achieved by setting the layers as non-trainable. After that, more layers are built, starting with a flattened layer and continuing with two thick layers that are activated by ReLU. Class probabilities for vehicle classification are generated by the last dense layer with SoftMax activation, which distinguishes between two distinct vehicles. We've chosen 64 as the Batch size, executed our code for 80 epochs, and set the learning rate to 0.001. For learning and backpropagation, we employed the Stochastic Gradient Descent technique, and categorical cross-entropy was utilized to measure the loss function.

Table 1: Accuracy Shown by Transfer Learning Models on Training Set

Model	Training Accuracy	Validation Accuracy
VGG-16	100	99.11
Inception V3	92.25	96.43

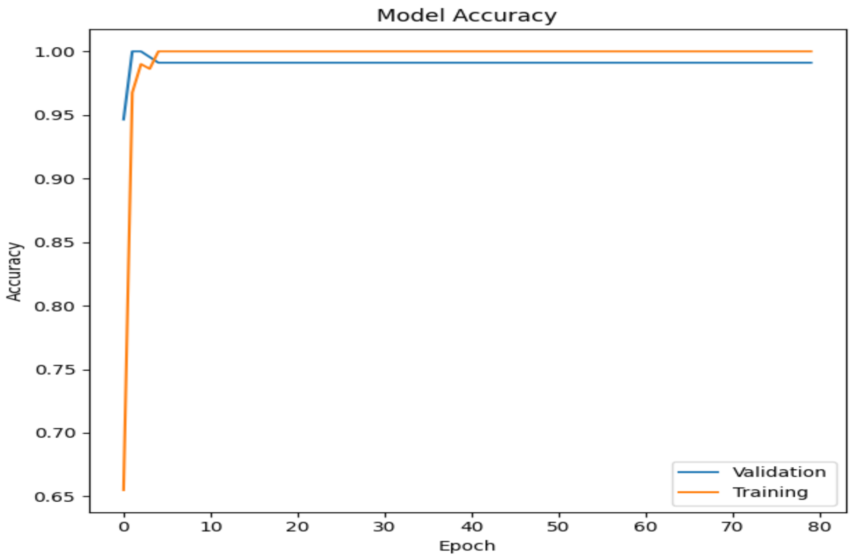

Figure 4: VGG-16 Training and Validation Accuracy curve

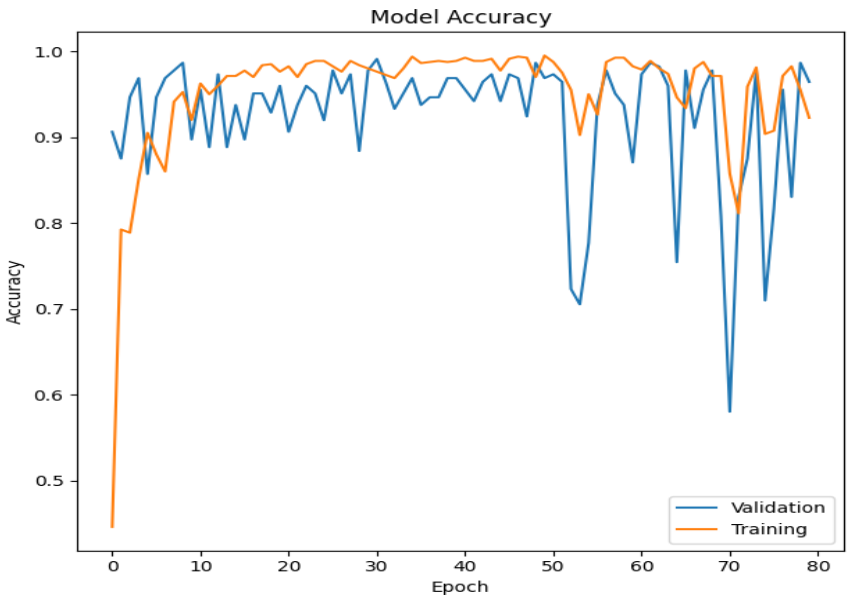

Figure 5: Inception V3 Training and Validation Accuracy Curve

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

The training and validation accuracy graphs are shown in Figures 4 and 5. Table 1 demonstrates that the VGG-16 architecture achieved an accuracy of 99.11% during validation and 100% throughout training. Additionally, Inception V3 attains a validation accuracy of 96.43% and a training accuracy of 92.25%.

Table 2 displays the results for two different transfer learning models: recall, accuracy, F1-Score, and precision. Four types of information form the basis of the classification report. You may find the formulas for Precision, Accuracy, F1-Score, and Recall here:

Table 2: Accuracy shown by Transfer Learning Models on Training set

Report	VGG-16	Inception V3
Validation Accuracy	99.11	96.43
Precision	99.50	97.00
Recall	99.50	97.00
F1-Score	99.50	97.00

Figures 6 and 7 shows the confusion matrix of Inception V3 and VGG-16 models, respectively.

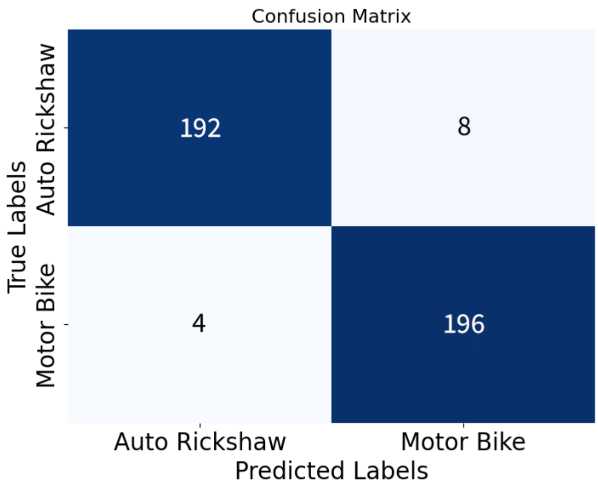

Figure 6: Confusion Matrix of VGG-16 Model

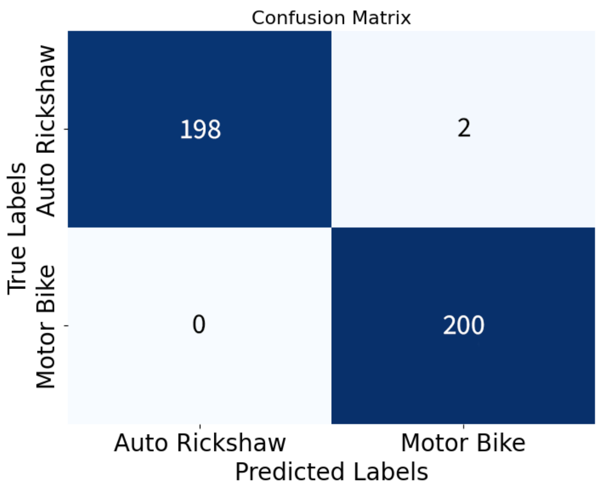

Figure 7: Confusion Matrix of Inception V3 Model

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

Lastly, computer vision techniques were used to assess the categorization model. As shown by the experimental findings, the suggested method was tested using a wide range of input photos, some of which would not have been part of the training dataset. The results of the Auto Rickshaw and Motor Bike predictions are shown in Figure 8.

Figure 8: Prediction Outcomes of Different Types of Vehicles

CONCLUSION

Improving the intelligent transport system is the goal of this study, which suggests a CNN-based vehicle classification. In our proposed model, we utilized VGG-16 and Inception V3 for vehicle detection. With a training and validation accuracy of 100% and 99.11% on the architecture, VGG-16 did quite well. Furthermore, future plans include exploring the integration of real-time data from IoT devices and cloud-based platforms to enhance the scalability and centralized management of the proposed vehicle classification system.

ACKNOWLEDGMENT

We express our gratitude to the Higher Education Commission of Pakistan (HEC) for its assistance in enabling Rashid Iqbal to finish his research work.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-01 | March-2024

REFERENCES

Maity, S., Saha, D., Singh, P. K., & Sarkar, R. (2024). JUIVCDv1: development of a still-image based dataset for indian vehicle classification. *Multimedia Tools and Applications*, 1-28.

Khairi, D. ul, Ayaz, F., Saeed, N., Ahsan, K., & Ali, S. Z. (2023). Analysis of deep convolutional neural network models for the fine-grained classification of vehicles. *Future Transportation, 3*(1), 133-149.

Ennehar, B. C., & Samra, B. (2023). Master-Slave Convolutional Deep Architecture for Vehicle Identification and Type Classification. *Traitement du Signal, 40*(2).

Alghamdi, A. S., Saeed, A., Kamran, M., Mursi, K. T., & Almukadi, W. S. (2023). Vehicle classification using deep feature fusion and genetic algorithms. *Electronics, 12*(2), 280.

Muhib, R. B., Ahmad, I. S., & Boufama, B. (2023). Deep Learning-Based Vehicle Classification. In *Proceedings of the Future Technologies Conference* (pp. 244-258).

Avianto, D., Harjoko, A., & Afiahayati. (2022). CNN-Based Classification for Highly Similar Vehicle Model Using Multi-Task Learning. *Journal of Imaging, 8*(11), 293.

Hasan, M. M., Wang, Z., Hussain, M. A. I., & Fatima, K. (2021). Bangladeshi native vehicle classification based on transfer learning with deep convolutional neural network. Sensors, 21(22), 7545.

Baghdadi, S., & Aboutabit, N. (2021). Transfer Learning for classifying front and rear views of vehicles. Journal of Physics: Conference Series, 1743, 012007.

Butt, M. A., Khattak, A. M., Shafique, S., Hayat, B., Abid, S., Kim, K. I., ... & Adnan, A. (2021). Convolutional neural network based vehicle classification in adverse illuminous conditions for intelligent transportation systems. *Complexity*, *2021*, 1-11.

Bin Che Mansor, M. A. H., Kamal, N. A. M., Baharom, M. H. B., & Zainol, M. A. B. (2021). Emergency vehicle type classification using convolutional neural network. In 2021 IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS) (pp. 126-129).

Mohamed, A. K., Ibrahim, A. K., Akram, A., Ibrahim, A., & Gamal, M. (2020). Improving Vehicle Classification and Detection with Deep Neural Networks. In 2020 International Symposium on Advanced Electrical and Communication Technologies (ISAECT) (pp. 1-5).

Meng, W., & Tia, M. (2020). Unmanned aerial vehicle classification and detection based on deep transfer learning. In 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) (pp. 280-285).

Tabassum, S., Ullah, M. S., Al-Nur, N. H., & Shatabda, S. (2020, June). Native vehicles classification on bangladeshi roads using cnn with transfer learning. In *2020 IEEE Region 10 Symposium (TENSYMP)* (pp. 40-43). IEEE.

Zhao, Z. (2022, December). Skin cancer classification based on convolutional neural networks and vision transformers. In *Journal of Physics: Conference Series* (Vol. 2405, No. 1, p. 012037). IOP Publishing.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).