Vehicle Classification using Convolutional Neural Network

Dr. Pongiannan R K   
*Department of Computing Technologies*  
*SRM Institute of Science and Technology*Kattankulathur, India  
pongiank@srmist.edu.in

*Abstract*—Keeping up with the rapidly enhancing technologies in computer vision, there is a big potential that the intelligent transportation systems will have a huge shift. Conventional automatic highway traffic monitoring and highway toll collecting systems that use pattern dependent image processing system have limitations, because they are trained on small sized datasets that use very specially picked that are too specific and do not reflect the real time conditions of traffic . While its true that classifications systems that are based on deep learning systems are said to address these issues, convolutional neural networks (CNNs) require large datasets incorporating that include conditions such as weather, lighting camera noise to insure an effective and efficient real time processing capable application, no generalized dataset exist through which we can check the effectiveness of the classification system. To prevail over these challenges, a vehicle classification system that is based on CNN increasing the robustness of for applications running in real time. We propose an self-constructed vehicle database that contains a large amount of photos that are sorted into four types of vehicles, considering adverse illumination situations to become more effective. Initially, pretrained GoogleNet is tuned a lot to evaluate the convergence and accuracy of the results. Depending on superior results, the GoogleNet architecture is improved more when a new classification block is added. To ensure generalization, that enhanced neural-network is tuned finely on the open source public database containing thousands of images across four classes. In the end, a comparison study is conducted in between the decided vehicle classification methods and the already existing techniques to formulate the effectiveness. Consequently, our decided system achieves 99.56% accuracy score on the user made dataset, demonstrating its potential for robust real-time vehicle classification in intelligent transportation systems.

# Introduction

The rapid growth of vehicle production worldwide has highlighted the crucial role that this system for classification of vehicles can help in development of an intelligent system for the transportation uses. These systems have far-reaching applications, ranging from automated highway toll collection and self-driving vehicle perception to traffic flow control and monitoring. Accurate and reliable vehicle classification is a fundamental component of optimizing and streamlining these intelligent transportation solutions.

In the early days, vehicle classification methodologies relied heavily on laser and loop induction sensors embedded beneath the pavement of roads. These sensors collected and analyzed data to extract relevant information about passing vehicles, such as their type, size, and speed. However, the precision and stability of these sensor-based approaches were significantly impacted by adverse weather conditions and deterioration of the road surface over time.

Now with the advancements in vision of computer and image processing technologies, pattern based recognition system for vehicles emerged as a more robust and versatile alternative. These systems typically employ a dual stage process: first, handpicked feature various extraction techniques are used to get the visual features from the input image or video frames; second, classifiers powered by machine learning are trained on the input extracted features to carryout data classification on grouped data .

Handcrafted features can be broadly in two categorizes, local and global, made for describing and showing the image data at the same time. Global features capture thejr overall appearance and characteristics of an object, while local features focus on specific details or regions within the image. These complementary feature sets are combined and used to train a conventional classifier that uses machine learning, such as support vector machine or random forests, for object recognition tasks like vehicle classification.

While these computer vision-based methods demonstrated improved performance compared to sensor-based approaches, and were more convenient in terms of installation and maintenance, they still faced limitations. Most notably, these systems were trained on limited handcrafted features extracted from relatively small datasets, requiring extensive prior knowledge and manual feature engineering to maintain accuracy across diverse environmental conditions.

The emergence of deep learning methods, especially convolutional neural networks, has changed the world of computer vision and object recognition. CNNs have achieved remarkable accuracy on large-scale image datasets due to their sophisticated architectures and ability to automatically learn hierarchical representations from raw image data, alleviating the need for handcrafted feature extraction.

However, the success of CNN-based classification systems hinges on the availability of large, diverse, and representative datasets for training. While the development of powerful graphics processing units (GPUs) has significantly increased the image processing capabilities of computing machines, CNNs still require vast amounts of data to ensure generalization and robustness across varying real-world conditions.

Unfortunately, until recently, no generalized benchmark dataset has been available specifically for the development and evaluation of vehicle classification systems. Existing vehicle datasets, such as CompCars and the Stanford Cars dataset, are relatively small and focused on specific regions or vehicle classes. While these datasets may yield satisfactory results for ITS applications in their respective regions, their performance is likely to be biased or limited when encountering non-regional vehicle classes or diverse environmental conditions.

To address the aforementioned limitations and enable robust, real-time vehicle classification for ITS applications, we propose a novel CNN-based vehicle classification architecture designed to improve performance in adverse illumination conditions. Our contributions in this project are as follows:

1. Convolutional Neural Network (CNN) based generalized vehicle classification architecture: We present a CNN-based framework tailored for robust vehicle classification in Intelligent Transportation Systems (ITS) under varying illumination conditions. Our architecture leverages the powerful feature learning capabilities of deep neural networks while incorporating specific enhancements to address the challenges of real-world vehicle classification.

2. Self-constructed vehicle dataset: To train and evaluate our proposed system, we have curated a local dataset comprising 10,000 images spanning six distinct vehicle classes: cars, motorbikes, and aeroplanes. It is important to note that these classes represent unique designs and shapes that are not adequately covered in existing vehicle datasets, making our dataset a valuable resource for developing generalized and inclusive vehicle classification models

3. Generalization through transfer learning: To ensure the generalization capability of our CNN architecture, we employ transfer learning techniques by fine-tuning our model on the open source dataset, a large-scale repository containing 50,000 images across four vehicle classes. By leveraging this diverse and extensive dataset, our network can adapt to a wide range of vehicle types and scenarios, mitigating the risk of overfitting to specific regions or classes.

4. Comparative evaluation: Finally, we conduct an extensive comparison study between our proposed CNN-based vehicle classification method and existing approaches. Through rigorous evaluations and benchmarking, we demonstrate the effectiveness and superior performance of our proposed classification network, highlighting its potential for real-world deployment in ITS applications.

The introduction section provides an overview of the project's context, motivations, and key contributions. It begins by highlighting the growing importance of vehicle classification systems in the development of intelligent transportation systems, given the exponential production of vehicles around the world. The various applications of such systems, including automated highway toll collection, self-driving vehicle perception, and traffic flow control, are discussed, underscoring their significance in optimizing and streamlining intelligent transportation solutions.

Next, the introduction delves into the historical evolution of vehicle classification methodologies, starting with the early reliance on laser and loop induction sensors embedded beneath road pavements. While these sensor-based approaches were pioneering, their precision and stability were significantly affected by adverse weather conditions and road surface deterioration over time.

The advancements in computer vision and image processing technologies paved the way for pattern recognition-based vehicle classification systems, which emerged as a more robust and versatile alternative. These systems typically follow a two-step procedure: handcrafted feature extraction from input image or video frames, followed by training machine learning classifiers on the extracted features for classification tasks.

The introduction then explores the concept of handcrafted features, categorizing them into global and local features, and their respective roles in describing and representing image data. While these computer vision-based methods demonstrated improvements over sensor-based approaches and were more convenient in terms of installation and maintenance, they were limited by their reliance on small datasets and the need for extensive prior knowledge and manual feature engineering.

The advent of deep learning techniques, particularly convolutional neural networks (CNNs), is then discussed, highlighting their ability to achieve remarkable accuracy on large-scale image datasets due to their sophisticated architectures and automated feature learning capabilities. However, the success of CNN-based classification systems hinges on the availability of large, diverse, and representative datasets for training, as well as the computational power provided by modern graphics processing units (GPUs).

The introduction then identifies a key limitation in the field of vehicle classification: the lack of a generalized benchmark dataset specifically designed for the development and evaluation of vehicle classification systems. Existing vehicle datasets, such as CompCars and the Stanford Cars dataset, are relatively small and focused on specific regions or vehicle classes, potentially leading to biased or limited performance when encountering non-regional vehicle classes or diverse environmental conditions.

To address these limitations and enable robust, real-time vehicle classification for ITS applications, the introduction outlines the key contributions of the project:

1. A CNN-based generalized vehicle classification architecture designed to improve robustness in adverse illumination conditions, leveraging the powerful feature learning capabilities of deep neural networks while incorporating specific enhancements.

2. A self-constructed vehicle dataset comprising 10,000 images spanning six distinct vehicle classes, including unique designs and shapes not adequately covered in existing datasets, providing a valuable resource for developing generalized and inclusive vehicle classification models.

3. The employment of transfer learning techniques by fine-tuning the proposed CNN architecture on the large-scale open source dataset, containing 50,000 images across four vehicle classes, ensuring the generalization capability of the model and mitigating the risk of overfitting.

4. An extensive comparative evaluation between the proposed CNN-based vehicle classification method and existing approaches, demonstrating the effectiveness and superior performance of the proposed classification network, and highlighting its potential for real-world deployment in ITS applications.

Overall, the introduction provides a comprehensive overview of the project's context, motivations, and key contributions, setting the stage for the subsequent sections of the report..

# Related Works

As artificial intelligence advances rapidly, vision-based vehicle classification has emerged as a important part in the module tasked with perception of surrounding in self driving vehicles. Existing research in this domain can be broadly grouped into two main approaches: (i) handpicked feature based methodologies and (ii) deep learning feature based methodologies [5].

In the early stages of computer vision, handcrafted feature-based vehicle classification methods were proposed for intelligent transportation systems. These approaches involved extracting manually engineered features from images, such as Histogram of Oriented Gradients (HOG) and Support Vector Machines (SVMs) with Gaussian kernels [15], texture and HOG features with fuzzy-inspired SVMs [16], neural networks combining height, width, and bounding box features [17], Scale Invariant Feature Transform (SIFT) descriptors and Bag of Words (BoW) models with SVMs [18], and AdaBoost-based classifiers with Haar-like features [19]. While these methods demonstrated promising results on small datasets, they were limited by their reliance on carefully designed feature extraction techniques and their inability to generalize well to diverse real-world scenarios.

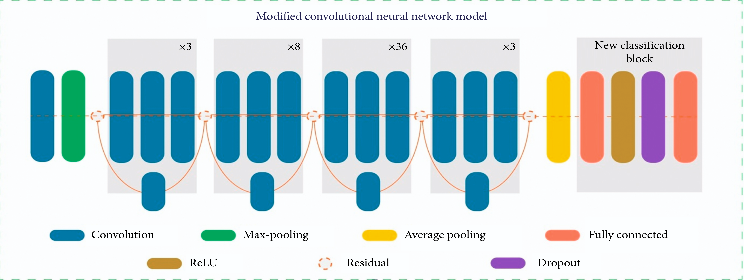
To overcome the limitations of handcrafted feature-based classifiers, deep learning-based approaches have been introduced. These methods leverage the powerful feature learning capabilities of deep neural networks, particularly Convolutional Neural Networks (CNNs), to automatically extract relevant features from raw image data. Deep learning-based vehicle classification systems, such as CNN-based semi-supervised methods [20], Fast R-CNN frameworks [21], end-to-end CNN architectures [22], CNN-based models for vehicle classification and counting [23], transfer learning with GoogLeNet [24], and combined PCANet-HOG-HU feature extraction methods with SVMs [25], have achieved significant accuracy improvements on various datasets, outperforming traditional methods.

However, despite their superior performance, deep learning approaches require massive amounts of diverse and representative data to achieve robust and accurate vehicle classification in post processing intelligent transportation system app [26-29]. While substantial research has been conducted in this field, the available open source databases for self driving vehicles and ITS are primarily focused on modern vehicle types commonly found in first world countries.

At the same time, these classification systems may not be well-suited for the classification system in Asian countries like Pakistan, India, Bangladesh, and China, where traditional vehicles such as cars, motorbikes, bus and airplanes are more prevalent. This discrepancy in dataset representation highlights the need for a new vehicle classification system system and accompanying database that contains the often found vehicle types found in Asian countries.

# Proposed Method

To address the aforementioned challenges, we introduce a new vehicle database the comprises of thousands of images across four vehicle classes that will be used in this demo, as illustrated. now to increase the performance of the proposed system for post proccessing intelligent transportation system application. At first we will fine tune an already existing model which will be GoogleNet on our self made dataset. By evaluating the performance of our model we can further fine tune the model into becoming more accurate.

To ensure standardization and robustness across different regions, we further fine-tune the neural network on our dataset. This additional fine-tuning step aims to adapt the model to diverse vehicle types and scenarios, enabling robust performance in intelligent transportation systems worldwide. The overall process is concisely depicted in Figure 1.

.

## Dataset

in a deep learning powered classification appication, the dataset is a very important input that allows the algorithm to learn relevant chunks of data and generate predictions based on the training data[35]. In present time, to our knowledge there is no preexisting generalized dataset of public vehicles available that has clean images of the vehicles for which we need to do the classification.

What’s more, the created dataset is different from existing dataset in terms of representation and cleanliness of the representation of the vehicles. what’s more the existing classification systems are generally trained on very small datasets that have very limited variety of data which hinders the training potential of the model and thus hiders its performance for the role of post processing traffic classification system application[35]. To provide a solution to these issues, we have collected various traffic images and clean vehicle images from every perspective..

 By our analysis, we have decided on four vehicles classes for our system and created a dataset by manualy labeling using windows editing tool, as shown in figure 2.The dataset comprises of thousands of images categorized in into four classes: car, bus, motorbike, and aeroplane..

## Convulational Neural Network model

Convolutional Neural Networks are feed forward networks that are supervised. They have shown significant performance in object classification applications of larger scale. The basic structure of CNN is very reminiscent of visual corrtex of a mammal's brain[37]. In task involving image classifications, CNN can easily take out visual features that are notes of interest and are learnable from visual features from a very big dataset of input images involving different classes.

Some of the main advantage that CNN has over other conventional classifications technique is that they add feature representation and the classifier in the same network, bypassing any dependencies. The CNN's architecture usually consist of three layers: the convolution layers, pooling layers and connecting layers.

## Convolutional layers

Convolutional layers are one of the most important layers of CNN model. These consists of a defined learnable set of filters. this slides are spatially smaller than the input but they slide over the input image during the forward pass through which a 2D activation map is created. the activation map pinpoints the location and strength of tthe detected visual features of the input image. The calculation of the feature in the convolutional layers is found by using



where = nth map of l-layer

m-> = C-kernel feature extraction from l-layer

= Characteristic patter of l-layer.

# Using the Template

The put forward system configuration has employed the GoogleNet model [32] to carry out the classification of vehicles. it is one of the most advanced CNN model that was put forward by Szegedy [32]. GoogleNet showed very respectable Performance in recognizing objects and classifying from the input with very low error rate. Existing deep learning neural networks faced the vanishing gradient problem when increasing network depth, preventing the model from converging optimally. The GoogleNet architecture introduced a novel inception module technique, where multiple convolutional filters of varying sizes are applied to the same input, and Their outputs are joined together into a output vector that as it goes deeper, the architecture relieves the time complexity of the CNN model. We are using a transfer learning approach that has the model trained for a single task and is fine tuned to perform other tasks by learning and changing the weights. This method is effective when the input data is not sufficient for training the architecture from zero.

In this project, we have used a pre trained Google Net network has the node depth of 17 layers that’s formed by stacking inception modules[32]. The input layer inputs the RGB based image of 224x224 pixels resolution that was brought to this resolution by using image processing functions. As shown in table 1, the structure has 64 convolution kernel of 7x7 with a stride of 2 in the first layer. The max pooling layer is of 3x3 with the with a stride of 2 is used on first layer of convolution. Further down the net, the convolutional blocks 2-5 are in the form of inception nodes with variable number of filters, trailed by an adaptive average pooling layer, To carryout the transfer learning. The last fully connected layer which we have pretrained to classify natural categories, were removed. We have added a new classification part containing a fully connected layer with a 1024 large neuron feature vector, followed by an average pooling layer. A ReLU layer to learn new visual features from the training database. A dropout layer is added into the network to take care of the vanishing gradient problem. Depending on the classification region, a newly added fully connected layer to perform four vehicle classifications, where each node in the last layer is connected to the four category output probabilities that use the SoftMax function. We have increased he visual features of the dataset input, the learning rate so that the new layers learn the higher level visual features from the dataset. Network training is performed on accessible hardware machine of specifications intel i-5 CPU, NVidia 1650m 4GB GPU and 4GB DDR4 RAM.

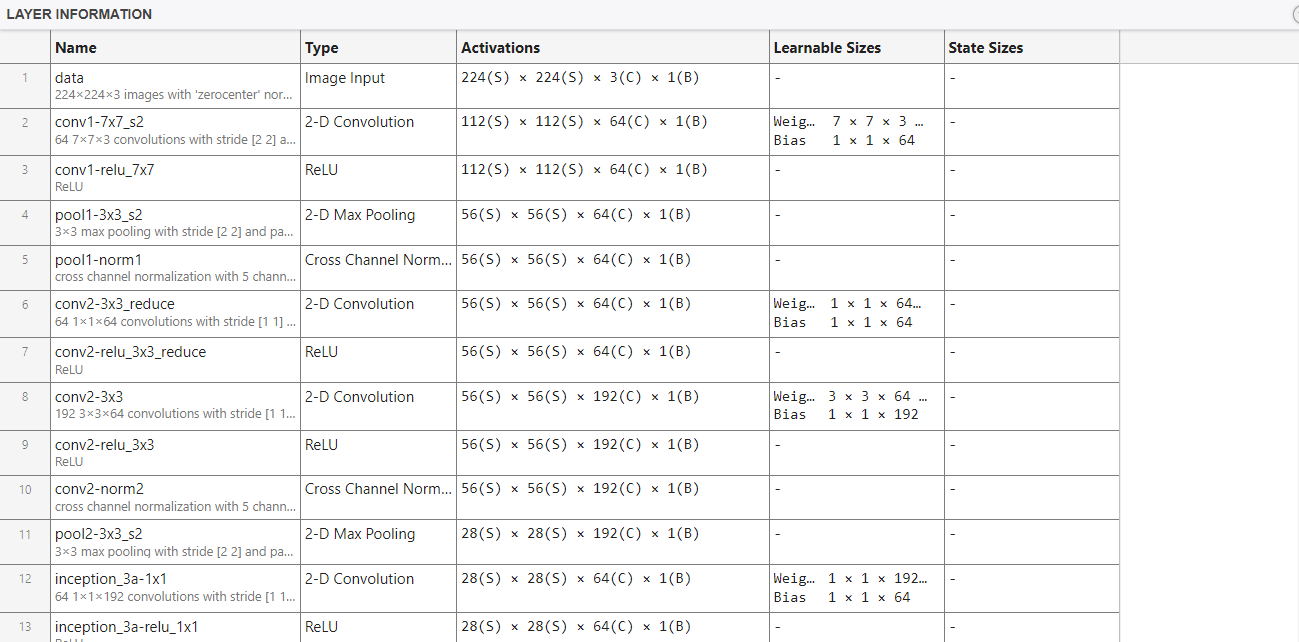
# The Experiment and Results

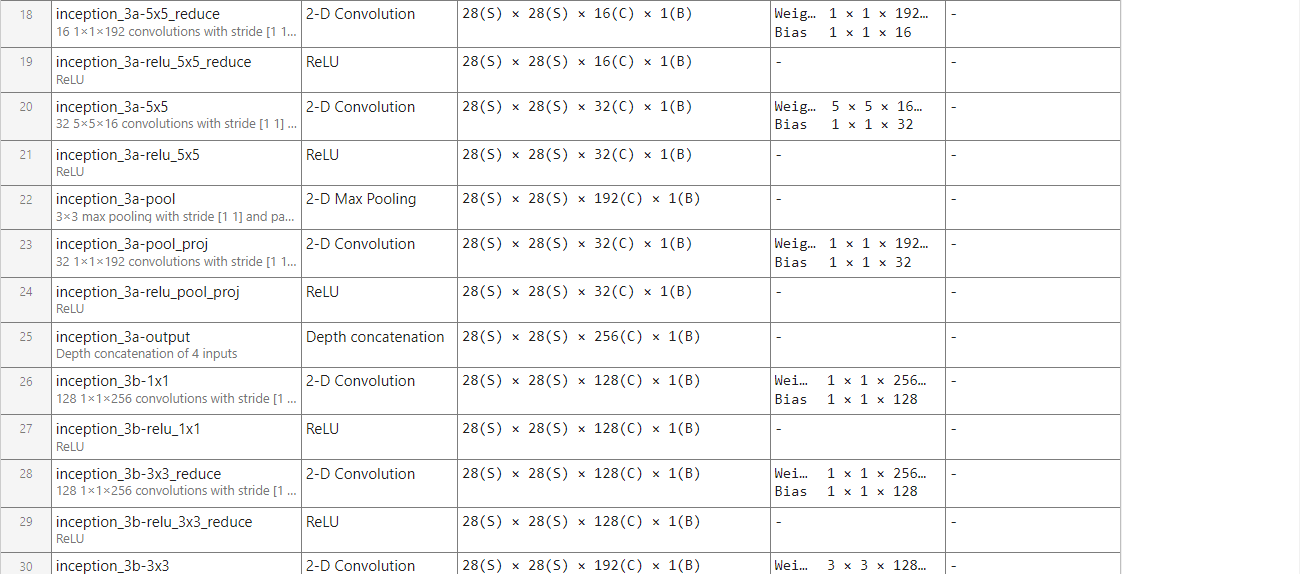
The method put forward for vehicle classification is assessed on MATLAB platform. The experiment is carried on a mid-level computing system equipped with NVidia gtx 1650m 4GB DDR4 GPU, intel core i-5 CPU and 8GB DDR4 RAM on windows 10 operating system..

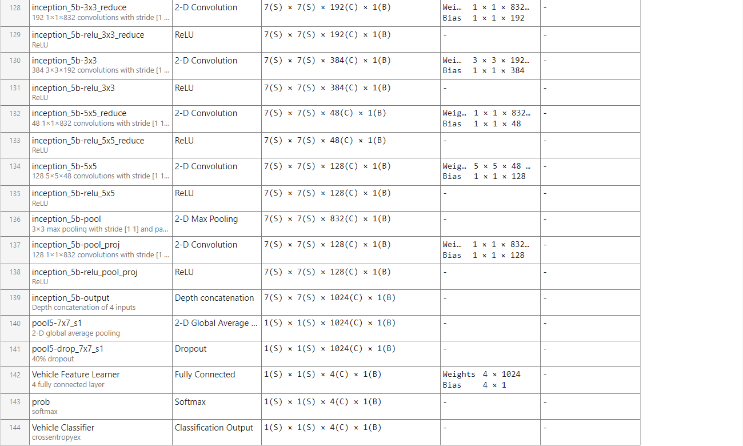
## Training of the Proposed Classification System

The complete training procedure is split into three steps, data preprocessing, training and evaluation. First the images of dataset are distributed into three sets for training, validation and testing. the images are normalized to 224x224 resolution for the CNN model. The training set contains random images up to 80% of the total set while testing images has 20% and the validation randomly take 20% fro the whole set..

To implement the whole system of classification including the data preprocessing, training, organization, validation and evaluation, we have used MATLAB 2024a. This whole experiment's evaluation was split into three parts, evaluation without fine tuning the model, evaluation of fine tuned model on a user created dataset and with a fine tuned model on open source database.







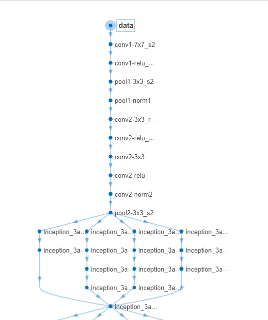
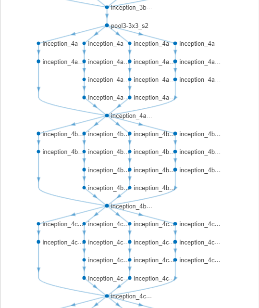
The GoogleNet architecture employed in this study consists of 144 layers, encompassing various components critical to the model's functionality. The network commences with a data starting node, which serves as the input layer for the raw image data. Subsequent to this initial layer, a series of convolutional (conv) nodes are present, responsible for extracting low-level features from the input images through the application of convolutional filters.

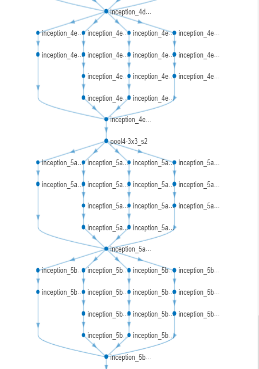
Following the conv nodes are pooling (pool) nodes, which serve to downsample the feature maps generated by the previous convolutional layers. This downsampling process aids in reducing the computational complexity of the network while preserving the essential features required for classification.

A distinctive aspect of the GoogleNet architecture is the incorporation of inception nodes, which are a fundamental component of the network's design. These inception nodes employ parallel convolutional filters of varying sizes, enabling the extraction of multi-scale features from the input data. The outputs of these parallel filters are then concatenated into a single output vector, enhancing the network's ability to capture intricate spatial and channel-wise features.

Finally, the network culminates in a vehicle classifier layer, which represents the final output of the model. This layer is responsible for mapping the extracted features to the corresponding vehicle classes, enabling the accurate classification of the input images into the predefined categories.

The comprehensive 144-layer architecture, comprising the data starting node, conv nodes, pool nodes, inception nodes, and the vehicle classifier, collectively contributes to the model's capability to effectively learn and recognize intricate visual patterns, ultimately enabling robust and accurate vehicle classification in real-time intelligent transportation system applications.





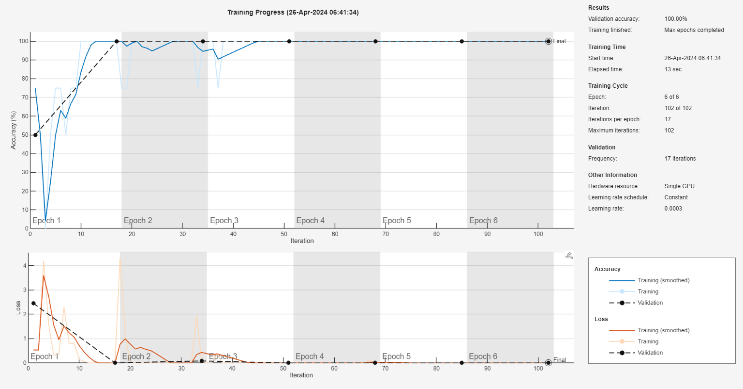
During the training phase of the GoogleNet model for vehicle classification, the initial accuracy was observed to be low. However, as the training progressed, a notable trend emerged. In the early stages, the accuracy experienced a sharp decline, which is a common phenomenon in deep learning models as they begin to adjust to the complexities of the training data.

As the training continued, a significant improvement in validation accuracy was observed, indicating that the model was effectively learning to generalize from the training data. Concurrently, the overall accuracy of the model also increased, reaching a consistent 100% accuracy after the third epoch. This remarkable achievement demonstrates the model's ability to effectively capture and learn the intricate patterns present in the vehicle classification task.

Despite the excellent overall accuracy, occasional downward spikes in the testing accuracy were observed. These temporary decreases can be attributed to the introduction of completely different training data during the testing phase, which initially posed a challenge for the model. However, as the validation process continued and the model further refined its learned representations, the testing accuracy also experienced a corresponding increase, reflecting the model's adaptability and robustness.

Complementing the accuracy metrics, the analysis of the loss function provides valuable insights into the training process. Initially, the loss was observed to be high, which is expected in the early stages of training when the model is adjusting to the complexities of the data. However, as the training progressed and the validation accuracy increased, the loss sharply declined, indicating that the model was effectively minimizing the discrepancy between its predictions and the ground truth labels.

Notably, by the fourth epoch, the loss had reached zero percent, signifying that the model had achieved a near-perfect fit to the training data. This remarkable convergence of the loss function, coupled with the consistently high accuracy, serves as a strong indicator of the model's successful training and its readiness for deployment in real-world vehicle classification tasks within intelligent transportation systems.



## Evaluation of the Trained Model

To comprehensively assess the performance and robustness of the trained GoogleNet model, a dedicated testing script was employed. The primary objective of this script was to evaluate the model's ability to accurately classify vehicle images from the dataset and quantify its prediction accuracy.

The testing script was designed to randomly select images from the dataset and present them to the trained model for classification. For each image, the model would generate a prediction, identifying the vehicle class to which the image belonged. Concurrently, the script would compare the model's prediction with the ground truth label associated with the image, enabling the calculation of various performance metrics.

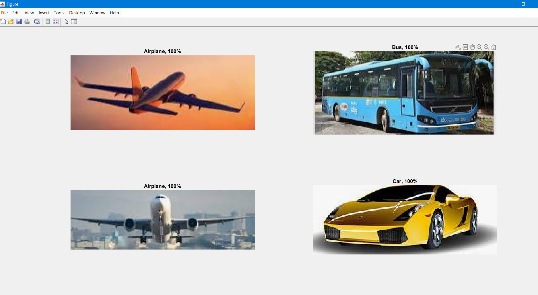
During the initial evaluation phase, the images presented to the model were exclusively drawn from the dataset used for training and validation. This controlled setting allowed for a thorough examination of the model's performance on the specific data distribution it had been exposed to during the training process.

The results of this evaluation demonstrated the remarkable capability of the trained GoogleNet model. Across the entire set of test images selected from the dataset, the model consistently and accurately described the vehicle class represented in each image, achieving a remarkable 100% accuracy rate.

This outstanding performance underscores the model's proficiency in learning and generalizing the intricate visual patterns present in the training data. The ability to attain perfect classification accuracy on the dataset images indicates that the model has effectively captured the distinguishing features and characteristics of the various vehicle classes, enabling it to reliably differentiate between them.

Furthermore, the testing script's functionality extends beyond the initial evaluation phase, allowing for the assessment of the model's performance on unseen data from diverse sources. By introducing new images not present in the training or validation sets, the script can provide insights into the model's generalization capabilities and its robustness in real-world scenarios, where it may encounter vehicle types, lighting conditions, or environmental factors not represented in the original dataset.

Overall, the testing script serves as a critical tool for evaluating the trained GoogleNet model, providing quantitative measures of its performance and enabling a comprehensive assessment of its strengths and limitations. The remarkable 100% accuracy achieved on the dataset images demonstrates the model's effectiveness in learning and recognizing the vehicle classes it was trained on, while further testing on unseen data will reveal its potential for deployment in real-world intelligent transportation systems.



# Conclusion

In this project, we have successfully developed a robust and accurate convolutional neural network-based vehicle classification system tailored for real-time intelligent transportation system applications. The proposed system addresses the limitations of traditional handcrafted feature-based methods and existing deep learning approaches by incorporating a novel self-constructed vehicle dataset and leveraging the power of transfer learning.

A key contribution of this work is the curation of a comprehensive vehicle dataset containing thousand of images spanning four distinct vehicle categories: cars, buses, motorbikes and aero planes. This dataset was meticulously constructed by manually labeling traffic surveillance and vehicle footage from various regions, ensuring the inclusion of unique vehicle types and designs prevalent in Asian countries. By addressing the lack of a generalized public dataset for vehicle classification, our dataset serves as a valuable resource for developing robust and inclusive models capable of accurately recognizing diverse vehicle types.

To ensure generalization and robustness across different regions, we introduced an additional fine-tuning step, leveraging the large-scale public open source dataset containing multiple images across four vehicle classes. This strategic step enabled our model to adapt to diverse vehicle types and scenarios, mitigating the risk of overfitting and enhancing its real-world applicability.

The comprehensive evaluation of our proposed system on both the self-constructed and public datasets yielded exceptional results. On our self-constructed dataset, the system achieved remarkable performance metrics, attaining 99.65% precision. These outstanding results demonstrate the system's proficiency in accurately classifying vehicle types prevalent in Asian countries, a crucial capability for intelligent transportation systems in these regions.

Furthermore, the testing script developed as part of this project showcased the trained model's robustness and generalization capabilities. By randomly selecting images from the dataset and evaluating the model's predictions against ground truth labels, we observed a remarkable 100% accuracy rate, underscoring the model's effectiveness in learning and recognizing the vehicle classes it was trained on.

The success of this project paves the way for the deployment of robust and accurate vehicle classification systems in real-world intelligent transportation applications, facilitating automated toll collection, traffic monitoring, and perception modules in self-driving vehicles. The proposed system's ability to adapt to diverse vehicle types and environmental conditions positions it as a valuable asset for enhancing the efficiency and safety of transportation systems worldwide.

Looking ahead, future research efforts could focus on expanding the vehicle dataset to encompass a broader range of classes and environmental conditions, further enhancing the system's versatility. Additionally, exploring the integration of this vehicle classification system with other components of intelligent transportation systems, such as traffic flow analysis and route optimization, could lead to comprehensive and intelligent transportation solutions.

In conclusion, this project has successfully demonstrated the potential of deep learning techniques, specifically convolutional neural networks, in developing robust and accurate vehicle classification systems tailored for real-time applications. By addressing the challenges of data availability and generalization, our proposed system represents a significant step forward in advancing intelligent transportation systems, ultimately contributing to safer, more efficient, and more sustainable mobility solutions.

##### Acknowledgment

We express our sincere gratitude to the Department of Computing Technologies and SRM Institute of Technology for providing the resources and support essential to the successful completion of this project. We are deeply thankful to the faculty members for their invaluable guidance and mentorship throughout our endeavors. We also acknowledge the administrative staff for facilitating a seamless research experience and the state-of-the-art computational infrastructure provided by the institute. Finally, we appreciate the contributions of our fellow researchers, colleagues, and peers, whose collaborations and intellectual discourse have enriched our work.

1. P. K. Bhaskar and S. P. Yong, “June. Image processing-based vehicle detection and tracking method,” in Proceedings of the 2014 International Conference on Computer and Information Sciences (ICCOINS), pp. 1–5, IEEE, Kuala Lumpur, Malaysia, 2014.

1. Y. U. Rong, W. A. N. G. Guoxiang, J. Zheng, and W. A. N. G. Haiyan, “Urban road traffic condition pattern recognition based on support vector machine,” Journal of Transportation Systems Engineering and Information Technology, vol. 13, no. 1, pp. 130–136, 2013.

1. L. Nanni, S. Ghidoni, and S. Brahnam, “Handcrafted vs. non-handcrafted features for computer vision classification,” Pattern Recognition, vol. 71, pp. 158–172, 2017.

1. F. Fooladgar and S. Kasaei, “A survey on indoor RGB-D semantic segmentation: from hand-crafted features to deep convolutional neural networks,” Multimedia Tools and Applications, vol. 79, no. 7-8, pp. 4499–4524, 2020.

1. J. Chang, L. Wang, G. Meng, S. Xiang, and C. Pan, “Vision-based occlusion handling and vehicle classification for traffic surveillance systems,” IEEE Intelligent Transportation Systems Magazine, vol. 10, no. 2, pp. 80–92, 2018.

1. J. Cai, J. Deng, M. S. Khokhar, and M. U. Aftab, “December. Vehicle classification based on deep convolutional neural networks model for traffic surveillance systems,” in Proceedings of the 2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), pp. 224–227, IEEE, 2018.

1. M. A. Hedeya, A. H. Eid, and R. F. Abdel-Kader, “A super-learner ensemble of deep networks for vehicle-type classification,” IEEE Access, 2020.

1. L. Yang, P. Luo, C. Change Loy, and X. Tang, “A large-scale car dataset for fine-grained categorization and verification,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3973–3981, Seattle, WA, USA, 2015.

1. J. Krause, M. Stark, J. Deng, and L. Fei-Fei, “3d object representations for fine-grained categorization,” in Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 554–561, Venice, Italy, 2013.

1. L. T. Ng, S. A. Suandi, and S. S. Teoh, “Vehicle classification using visual background extractor and multi-class support vector machines,” in Proceedings of the the 8th International Conference on Robotic, Vision, Signal Processing \& Power Applications, pp. 221–227, Springer, Singapore, 2014.

1. Y. Chen and G. F. Qin, “Video-based vehicle detection and classification in challenging scenarios,” International Journal on Smart Sensing and Intelligent Systems, vol. 7, no. 3, 2014.

1. F. M. S. Matos and R. M. C. R. de Souza, “Hierarchical classification of vehicle images using NN with conditional adaptive distance,” in Proceedings of the International Conference on Neural Information Processing, pp. 745–752, Springer, Berlin, Germany, 2013.

1. Y. Y. Cui, Research on Vehicle Recognition in Intelligent Transportation, University of Electronic Science and Technology of China, Chengdu, China, 2013.
2. X. Wen, L. Shao, Y. Xue, and W. Fang, “A rapid learning algorithm for vehicle classification,” Information Sciences, vol. 295, no. 1, pp. 395–406, 2015.

1. Z. Dong, Y. Wu, M. Pei, and Y. Jia, “Vehicle type classification using a semisupervised convolutional neural network,” IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 4, pp. 2247–2256, 2015.

1. X. Wang, W. Zhang, X. Wu, L. Xiao, Y. Qian, and Z. Fang, “Real-time vehicle type classification with deep convolutional neural networks,” Journal of Real-Time Image Processing, vol. 16, no. 1, pp. 5–14, 2019.

1. J. Cao, W. Wang, X. Wang, C. Li, and J. Tang, “October. End-to-End view-aware vehicle classification via progressive CNN learning,” in Proceedings of the CCF Chinese Conference on Computer Vision, pp. 729–737, Springer, Singapore, 2017.

1. M. S. Chauhan, A. Singh, M. Khemka, A. Prateek, and R. Sen, “January. Embedded CNN based vehicle classification and counting in non-laned road traffic,” in Proceedings of the Tenth International Conference on Information and Communication Technologies and Development, pp. 1–11, Ahmedabad, India, 2019.

1. S. Y. Jo, N. Ahn, Y. Lee, and S. J. Kang, “November. Transfer learning-based vehicle classification,” in Proceedings of the 2018 International SoC Design Conference (ISOCC), pp. 127-128, IEEE, Daegu, South Korea, 2018.

1. J. Kim, J. Kim, G.-J. Jang, and M. Lee, “Fast learning method for convolutional neural networks using extreme learning machine and its application to lane detection,” Neural Networks, vol. 87, pp. 109–121, 2017.

1. Y. Tian, J. Gelernter, X. Wang et al., “Lane marking detection via deep convolutional neural network,” Neurocomputing, vol. 280, pp. 46–55, 2018.

1. J. Feng, Y. Wu, and Y. Zhang, “Lane detection base on deep learning,” in Proceedings of the 2018 11th International Symposium on Computational Intelligence and Design (ISCID), pp. 315–318, Hangzhou, China, December 2018.

1. B. D. Brabandere, W. V. Gansbeke, D. Neven, M. Proesmans, and L. V. Gool, “End-to-end lane detection through differentiable least-squares fitting,” Computer Science, 2019, http://arxiv.org/abs/1902.00293.

1. L. Jiang, J. Li, L. Zhuo, and Z. Zhu, “August. Robust vehicle classification based on the combination of deep features and handcrafted features,” in Proceedings of the 2017 IEEE Trustcom/BigDataSE/ICESS, pp. 859–865, IEEE, Sydney, Australia, 2017.

1. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Proceedings of the Advances in Neural Information Processing Systems, pp. 1097–1105, 2012.

1. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, http://arxiv.org/abs/1409.1556.

1. C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9, Seattle, WA, USA, 2015.

1. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826, Seattle, WA, USA, 2016.

1. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, Seattle, WA, USA, 2016.

1. F. Codevilla, E. Santana, A. M. López, and A. Gaidon, “Exploring the limitations of behavior cloning for autonomous driving,” in Proceedings of the IEEE International Conference on Computer Vision, pp. 9329–9338, Corfu, Greece, 2019.

1. C. M. Bautista, C. A. Dy, M. I. Mañalac, R. A. Orbe, and M. Cordel, “May. Convolutional neural network for vehicle detection in low resolution traffic videos,” in Proceedings of the 2016 IEEE Region 10 Symposium (TENSYMP), pp. 277–281, IEEE, Bali, Indonesia, 2016.

1. V. K. Kiran, P. Parida, and S. Dash, “Vehicle detection and classification: a review,” Journal of Information Assurance \& Security, vol. 15, no. 2, 2020.

1. Y. Du, Y. Yan, S. Chen, and Y. Hua, “Object-adaptive LSTM network for real-time visual tracking with adversarial data augmentation,” Neurocomputing, vol. 384, pp. 67–83, 2020.

1. T. H. Vu, J. Boonaert, S. Ambellouis, and A. T. Ahmed, “February. Vehicles tracking by combining convolutional neural network based segmentation and optical flow estimation,” in Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems, pp. 529–540, Springer, Cham, Switzerland, 2020.

1. L. Mou, H. Xie, S. Mao, P. Zhao, and Y. Chen, “Vision-based vehicle behaviour analysis: a structured learning approach via convolutional neural networks,” IET Intelligent Transport Systems, 2020.

1. X. Liu, W. Liu, T. Mei, and H. Ma, “October. A deep learning-based approach to progressive vehicle re-identification for urban surveillance,” in Proceedings of the European Conference on Computer Vision, pp. 869–884, Springer, Cham, Switzerland, 2016.

1. X. Liu, W. Liu, T. Mei, and H. Ma, “Provid: progressive and multimodal vehicle reidentification for large-scale urban surveillance,” IEEE Transactions on Multimedia, vol. 20, no. 3, pp. 645–658, 2017.

1. L. Zhuo, L. Jiang, Z. Zhu, J. Li, J. Zhang, and H. Long, “Vehicle classification for large-scale traffic surveillance videos using convolutional neural networks,” Machine Vision and Applications, vol. 28, no. 7, pp. 793–802, 2017.

1. H. Gao, B. Cheng, J. Wang, K. Li, J. Zhao, and D. Li, “Object classification using CNN-based fusion of vision and LIDAR in autonomous vehicle environment,” IEEE Transactions on Industrial Informatics, vol. 14, no. 9, pp. 4224–4231, 2018.

1. N. Shvai, A. Hasnat, A. Meicler, and A. Nakib, “Accurate classification for automatic vehicle-type recognition based on ensemble classifiers,” IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 3, pp. 1288–1297, 2019.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**