Traffic Congestion Prediction with a Self-Collected Dataset Using Image-Based Analysis

F. A. Moustafa[1], H. M. Nasr[2], M. Nashaat[3], M. Hossam[4], A. Hossam[5]

[1 2 3 4 5]Bachelor of Artificial Intelligence in School of Information Technology and Computer Science (ITCS), Nile University, Egypt.

F.ahmed2270@nu.edu.eg, H.Nasr2270@nu.edu.eg, M.Nashaat2265@nu.edu.eg
M.Hossam2201@nu.edu.eg, A.Hossam2232@nu.edu.eg

Abstract - This research introduces a novel dataset for traffic congestion prediction, collected from two critical locations on the 26th of July Corridor in Egypt: Juhayna Square and the MUST University Pedestrian Bridge. To ensure comprehensive data, images were captured at various times of the day, capturing diverse traffic conditions. Initially consisting of 370 images labeled as high, medium, or low traffic, the dataset was augmented to enhance generalizability due to its small size, and then split into training, validation, and test sets. We evaluated the dataset using several deep learning models, including VGG16, InceptionV3, MobileNet, ResNet, DenseNet and EfficientNetB0. Ensemble learning was applied to optimize predictions, resulting in promising accuracy. Despite challenges in data collection, our dataset contributes valuable insights into real-time traffic management and prediction systems.

Keywords - Traffic Congestion, Augmentation

I. INTRODUCTION

Predicting traffic congestion has led to an increase in research in recent years, particularly in the field of Artificial Intelligence, Deep Learning and Machine Learning. This field of study has grown significantly over the past few decades due to the emergence of huge data from stationary sensors or probe vehicle data and the creation of new AI algorithms as M. Akhtar and S. Moridpour [2] stated.

So far, traffic congestion prediction has been receiving much attention in the context of civil engineering as well as information technology. Traffic congestion prediction plays an important role in route guidance and traffic management. We formulate it as a binary classification problem S. Yang [4]. A vital problem faced by urban areas, traffic congestion impacts wealth, climate, and air pollution in cities. Sustainable transportation systems (STSs) play a crucial role in traffic congestion prediction for adopting transportation networks to improve the efficiency and capacity of traffic management M. Anjaneyulu, M. Kubendiran [5].

Traffic congestion is a significant challenge that affects urban areas worldwide, with both direct and indirect impacts on economies, public health, and the environment. In Egypt, where urban expansion and vehicle usage are rapidly increasing, there is a noticeable lack of localized datasets for traffic congestion analysis. This gap hinders the ability of authorities to monitor and manage road conditions effectively, which is critical for supporting the country's ongoing developments in the transportation sector.

Effective traffic monitoring is vital for urban planning, road infrastructure optimization, and reducing congestion-related issues. As highlighted by Ali et al. [3], traffic congestion in Pakistan causes daily losses of Pak Rs. 1 million due to opportunity costs and fuel consumption. Egypt is experiencing similar economic effects, compounded by individual-level impacts such as time loss during peak hours, mental stress, and the contribution to global warming through increased emissions. These challenges underscore the urgent need for a comprehensive traffic congestion monitoring system supported by robust datasets.

This research addresses the gap by presenting a novel, self-collected dataset specifically designed for traffic congestion prediction in Egypt. Images were captured at two critical locations on the 26th of July Corridor, Juhayna Square and the MUST University Pedestrian Bridge, cross different times of the day to ensure diversity in traffic conditions. The dataset initially had 370 images labeled as High Traffic, Medium traffic, or Low Traffic, was preprocessed and augmented to enhance its utility and size reaching 2240 image with noticeable varieties and was tested using advanced machine learning models, including VGG16, InceptionV3, MobileNet, ResNet, DenseNet and EfficientNetB0. Ensemble learning techniques further validated its suitability for predicting traffic congestion with promising results.

This paper aims to provide a benchmark dataset to aid authorities and researchers in developing smarter traffic monitoring systems, ultimately contributing to better urban mobility and reduced environmental impacts.

II. STUDY AREA

The 26th of July Corridor - is one of Egypt's most critical roadways, serving as the most suitable link between Giza and Cairo. This road plays a vital role in connecting residential, educational, and commercial zones and is the most heavily used route for commuting between the two cities. Due to its strategic importance and heavy daily traffic, it frequently experiences significant congestion, making it a prime area for studying urban traffic patterns.

Juhayna Square - is a major intersection along the 26th of July Corridor. It acts as a hub for vehicular and pedestrian activity, connecting numerous residential areas, businesses, and recreational sites. Additionally, it is located next to **Mall of Arabia**, one of the most visited destinations in Egypt, attracting thousands of visitors daily. The square also houses **Nile University**, a prominent educational institution, and serves as a public transportation stop for many commuters. Additionally, it is surrounded by various companies and other important establishments, making it a hotspot for traffic flow and congestion, especially during peak hours.

MUST University Pedestrian Bridge – is another critical location on the corridor, offers unique traffic variability. This area is home to several educational institutions including **MUST University** and two of Egypt's largest hospitals **Souad Kafafi** and **Dar Al Fouad**, making it a significant point for both vehicular and pedestrian movement. The presence of a public transportation stop contributes to its bustling activity throughout the day. These factors make it an ideal site for capturing diverse traffic conditions and understanding time-dependent variations.

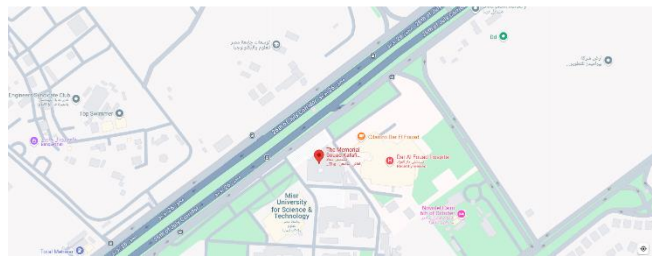
These two locations were selected for them:

- **Critical Nature**: Representing major congestion hotspots.
- **Diversity**: Providing a variety of traffic patterns due to their mixed-use for educational, medical, and commercial purposes.
- **Relevance**: Offering actionable insights for traffic management and optimization.

The data collection process was challenging. Capturing high-quality images during periods of heavy traffic, accounting for lighting changes throughout the day, and navigating the dense activity in these locations required careful planning and execution.

Figure 1: Map showing the location of Juhayna Square on the 26th of July Corridor, Egypt.

Figure 2: Map showing the location of MUST University Pedestrian Bridge on the 26th of July Corridor, Egypt.



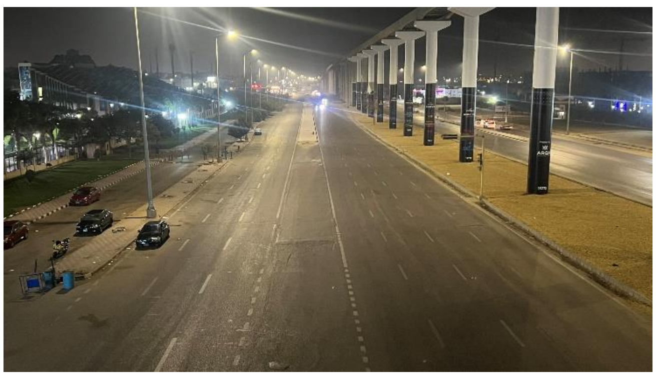
III. DATASET DESCRIPTION

Collection Process - Images were captured several times throughout the day to reflect diverse traffic patterns and levels:

- **Morning Rush Hours (8:00 AM - 9:30 AM):** There was high traffic due to individuals heading to workplaces, schools, universities, hospitals, governmental departments, and bus stops. This period exhibited significant traffic jams and high flow rates.
- **Late Morning (11:00 AM - 12:30 PM):** Traffic levels were partially flowing, leaning towards above-average congestion as most people were occupied at their destinations.
- **Afternoon/Evening Rush (2:30 PM - 5:00 PM):** The highest traffic period of the day, with individuals returning home, heading to bus stops, or moving between various destinations.
- **Night (9:00 PM - 11:00 PM):** Above-average traffic persisted during this period, influenced by evening activities and leisure.
- **Late Night (1:30 AM - 3:00 AM):** Traffic levels began to decrease as roads started clearing, but still captured some activity, ensuring the dataset included diverse conditions.
- **Early Morning (4:30 AM - 6:00 AM):** The least congested time of the day, as most people were asleep, and traffic was minimal.

Figure 3: Showing an image of peak traffic during the morning rush hours at Juhayna Square before preprocessing and augmentation.

Figure 4: Showing an image of the least traffic during the early morning hours at MUST University Pedestrian Bridge before preprocessing and augmentation.



Images were taken using smartphones with high-quality cameras, ensuring clarity and detail. The elevated positions from which images were captured contributed to the overall visibility and quality of the dataset.

Dataset Composition - The dataset was divided into three traffic categories: High Traffic, Medium Traffic, and Low Traffic.
Initially, the dataset comprised 370 images distributed as follows:

- High Traffic: 99 images
- Medium Traffic: 100 images
- Low Traffic: 171 images

To address the imbalance in class distribution, augmentation techniques were applied, increasing the dataset to a total of 2,240 images with near-equal class representation:

- High Traffic: 735 images
- Medium Traffic: 742 images
- Low Traffic: 763 images

The dataset was then split into training, validation, and testing subsets as follows:

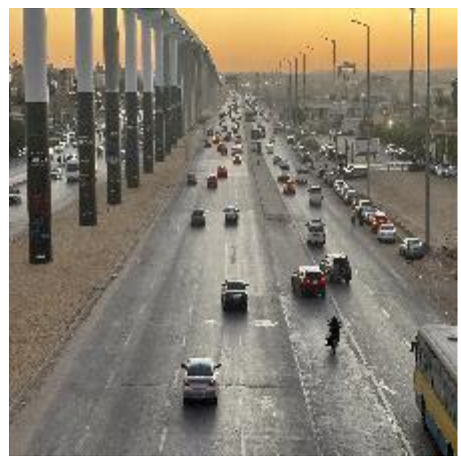
1. **Training and Validation Set (1,870 images):**
    - High Traffic: 615 images
    - Medium Traffic: 625 images
    - Low Traffic: 630 images
    - Training Set: 1,496 images
    - Validation Set: 374 images
2. **Testing Set (370 images):**
    - High Traffic: 120 images
    - Medium Traffic: 117 images
    - Low Traffic: 133 images

Day and night images were approximately equally distributed, with a slightly higher proportion of night-time images.

Data Preprocessing and Augmentation

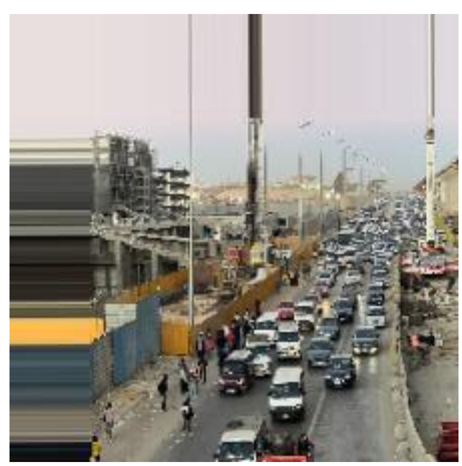
- **Preprocessing:** Images were resized to a uniform input size of 224x224 pixels, normalized, and converted to RGB format to standardize input across all models.

Figure 5: Showing an image of medium traffic at MUST University Pedestrian Bridge after preprocessing.



- **Augmentation:** Techniques applied included horizontal flipping, horizontal and vertical shifting, rotation, brightness adjustment, and zooming. Augmentation was specifically tailored to balance class distributions, ensuring approximately equal representation for all three traffic levels.

Figure 6: Showing an image of high traffic at Juhayna Square after augmentation.



Challenges in Data Collection

- **Traffic Conditions:** Capturing images during peak traffic periods was challenging due to congestion and environmental factors, such as weather and lighting variations.
- **Safety and Perception Issues:** Photographing roads and public spaces occasionally raises suspicions from pedestrians or authorities. Some individuals questioned the purpose of capturing road images, leading to potential misunderstandings.
- **Class Imbalance:** The initial dataset had a significant imbalance, with the Low Traffic class being overrepresented. This was mitigated using augmentation techniques to ensure fair model training.

VI. MODELS AND METHODOLOGY

In this study, we leverage state-of-the-art convolutional neural networks (CNNs) to classify traffic congestion images. CNNs have shown impressive results in object classification and recognition tasks, particularly in scenarios where traditional feature extraction methods struggle due to high occlusion, scale variation, or lighting conditions. Moreover, some traditional methods require laborious annotations for each object, which can be time-consuming and error-prone. CNNs, on the other hand, eliminate these issues by automatically learning relevant features directly from raw images, making them an ideal choice for this task.

The CNN models chosen for this study are VGG16, InceptionV3, MobileNet, ResNet, EfficientNetB0, and DenseNet. These models were selected based on their demonstrated success in various image classification applications. Additionally, we employed transfer learning using pre-trained models, which have been trained on large datasets like ImageNet, allowing them to generalize well to our dataset.

As reported by C. Lin et al. [6], "the accuracy of MobileNetV2 with attention module on the test set exceeds that of several advanced CNN models such as VGG16, GoogLeNet, ResNet50, EfficientNetB0, DenseNet121." This highlights the efficacy of MobileNetV2, especially when used with attention mechanisms, in producing accurate results while maintaining computational efficiency. MobileNet is known for its lightweight architecture, making it suitable for real-time applications without sacrificing performance.

1. **VGG16:** is a deep architecture that is known for its simplicity and effectiveness in various image classification tasks. Despite its relatively large size compared to newer models, VGG16's performance remains competitive, making it a solid baseline in transfer learning applications.

2. **InceptionV3:** is selected for its ability to achieve high performance while keeping the model size smaller compared to other deep CNNs. Despite its compact design, it performs exceptionally well in image classification tasks, making it an excellent choice for transfer learning applications [7].

3. **MobileNet:** is chosen for its efficient use of parameters and computational resources. As mentioned by C. Lin et al. [6], MobileNetV2 with an attention module outperforms several advanced CNN models, making it an ideal choice for traffic congestion prediction where real-time performance is crucial.

4. **ResNet:** with its residual connections, is effective at training very deep networks and helps mitigate the vanishing gradient problem. It has been key architecture in achieving state-of-the-art performance on several image classification benchmarks.

5. **EfficientNetB0:** is known for its efficient scaling, balancing depth, width, and resolution. It achieves superior performance with fewer parameters, making it ideal for tasks like traffic congestion prediction, where computational efficiency is important.

6. **DenseNet:** is selected for its unique architecture where each layer is connected to every other layer, allowing for efficient feature reuse. This results in fewer parameters while maintaining high performance. DenseNet has shown state-of-the-art results in various image classification tasks, making it a strong contender for this study.

Data Preprocessing Pipeline - To ensure consistency in the input data and prepare it for the models, we applied a series of preprocessing techniques. All images were resized to 224x224 pixels, normalized to a range of 0-1, and converted to RGB format. These steps helped standardize the input, making it suitable for the models.

We also employed various data augmentation techniques to increase the diversity of the dataset. These included horizontal flipping, horizontal shifting, vertical shifting, rotation, brightness adjustments, and zooming. This augmented dataset helped balance the class distribution and improved the models' generalization capabilities. The augmentation process resulted in a final dataset of 2240 images, evenly distributed across the three traffic categories: high, medium, and low traffic.

Training and Validation Workflow - The dataset was split into 60% for training, 20% for validation, and 20% for testing. The training set consisted of 1496 images, while the validation and test sets contained 374 and 370 images, respectively. The class distribution was balanced through augmentation, ensuring that the models would not be biased toward the majority class.

Ensemble Learning Approach - To enhance the performance of individual models, we employed an ensemble learning strategy. By combining the predictions from multiple models, we were able to leverage the strengths of each model, improving the overall prediction accuracy. Ensemble learning is particularly useful when individual models exhibit complementary strengths and weaknesses. In this study, ensemble learning helped reduce errors, providing more reliable and robust predictions for traffic congestion.

Evaluation Metrics - We evaluated the performance of the models using several metrics, including accuracy, precision, recall, and F1-score. These metrics were chosen to assess not only the overall classification accuracy but also the model's ability to correctly classify high-traffic, medium-traffic, and low-traffic images. F1-score was used to balance the trade-off between precision and recall, ensuring that the models performed well across all classes.

V. RESULTS AND DISCUSSION

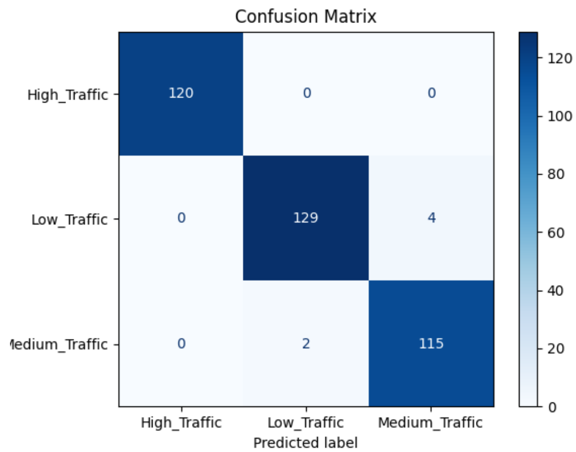
In this section, we present the performance metrics of the models used in the traffic congestion prediction task, comparing their accuracy, precision, recall, and F1-score. Additionally, we discuss key insights, challenges, and the significance of the results. Visualizations, including confusion matrices, bar plots, and training-validation curves, are included to illustrate the performance.

Performance Evaluation

Table 1 summarizes the test and training accuracies of the five models.

Model	Test Accuracy	Training Accuracy
MobileNet	96.39%	99.80%
DenseNet	95.95%	99.85%
ResNet50	85.14%	87.40%
VGG16	98.37%	98.79%
InceptionV3	98.65%	92.00%

Figure 7: Confusion matrix of the VGG16 model shows detailed classification performance across all three traffic classes: High, Medium, and Low Traffic.



Model Comparison - MobileNet achieved impressive test accuracy (96.39%) with minimal computational overhead, validating its efficiency in real-world applications. However, InceptionV3 outperformed MobileNet in test accuracy (98.65%) while maintaining comparable computational efficiency.

DenseNet also showed competitive performance, with a test accuracy of 95.95%, but required slightly more training time due to its dense connections. On the other hand, ResNet50 displayed the lowest test accuracy (85.14%) and struggled with Medium Traffic classification, as shown in its confusion matrix.

VGG16 demonstrated excellent classification performance, with a test accuracy of 98.37%. However, its high computational cost, along with ResNet50, resulted in prolonged training times.

Figure 8: Bar plot comparing the test and train accuracy of MobileNet.

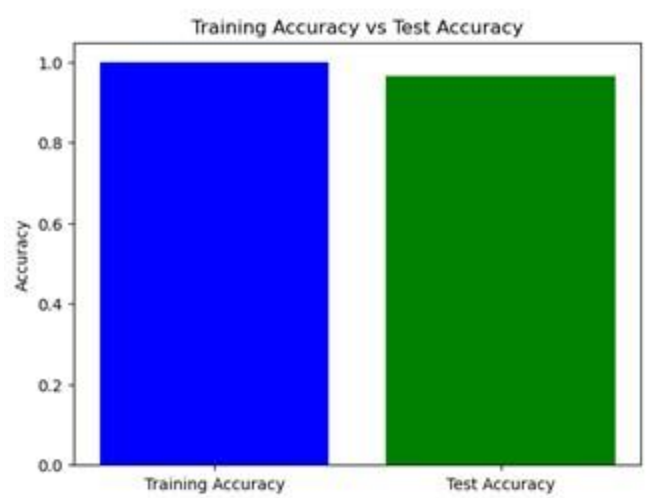
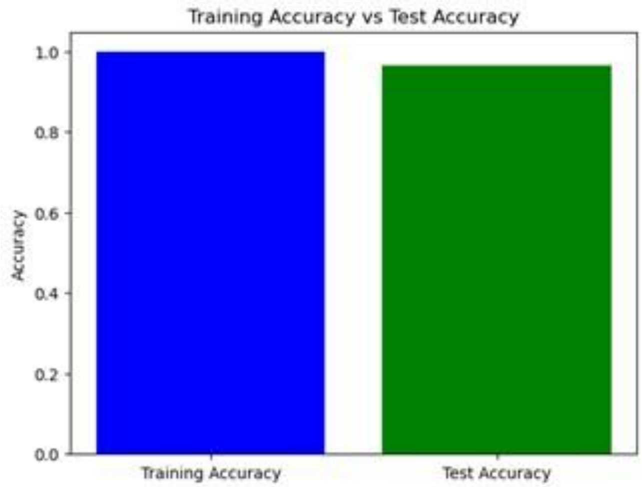


Figure 9: Bar plot comparing the test and train accuracy of InceptionV3.



Error Analysis - ResNet50 faced challenges in correctly classifying Medium Traffic, which is evident from its classification report. This could be attributed to its inability to capture intermediate-level traffic features effectively.

Figure 10: Low-level features extracted by ResNet highlight its limitations in distinguishing between Medium and Low Traffic, contributing to its lower accuracy.

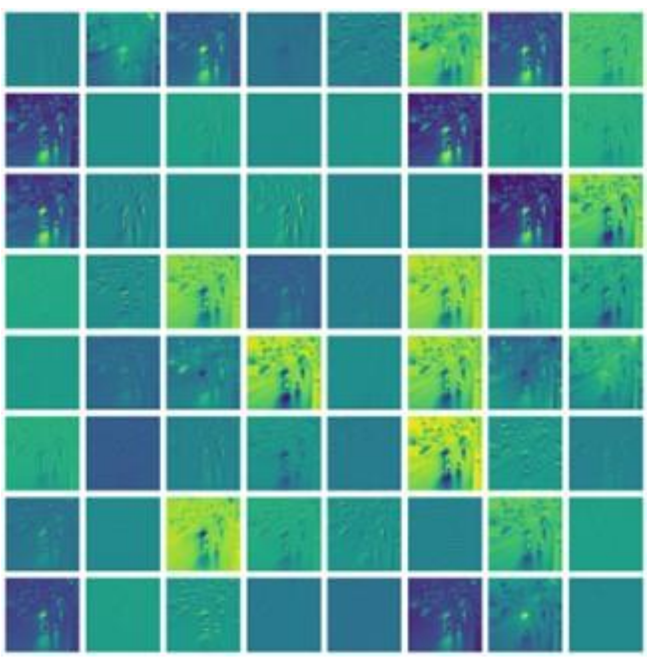
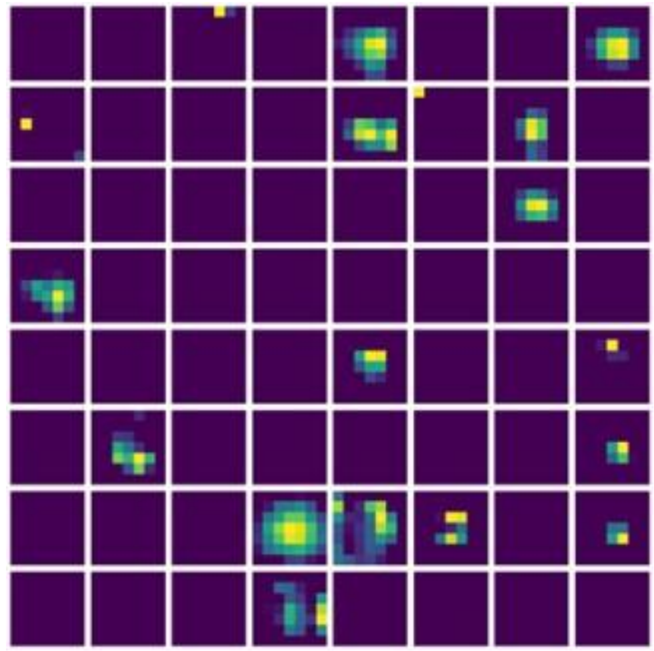


Figure 11: High-level features of ResNet show its strength in High Traffic classification but difficulty in Medium Traffic scenarios.



Training and Validation Performance - The models showed initial overfitting, which was mitigated using techniques like data augmentation, hyperparameter tuning, and regularization methods (e.g., dropout, L2 regularization, early stopping).

Figure 12: Training and validation accuracy and loss plots for VGG16 illustrate its convergence and minimal overfitting.

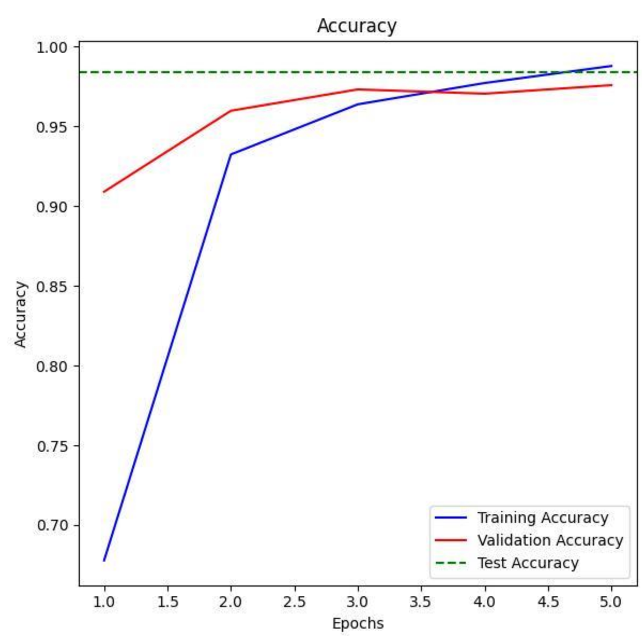


Figure 13: MobileNet training and validation curves demonstrate its stable learning process and consistent accuracy across epochs.

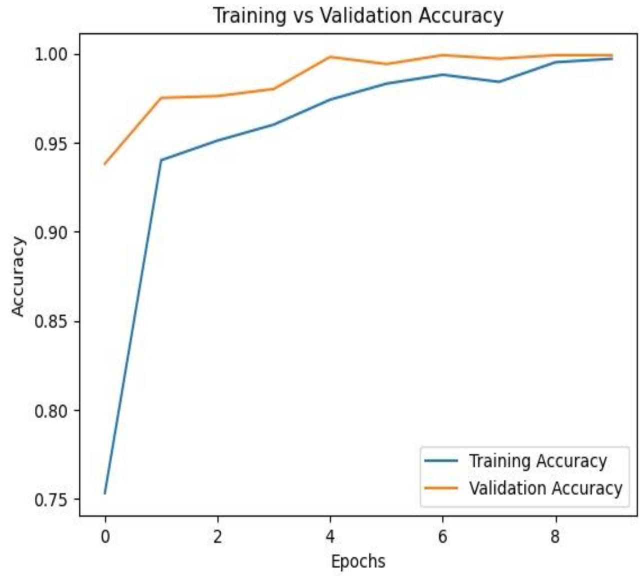


Figure 14: Training and validation accuracy plots for ResNet show slower convergence and higher loss, indicating room for improvement.

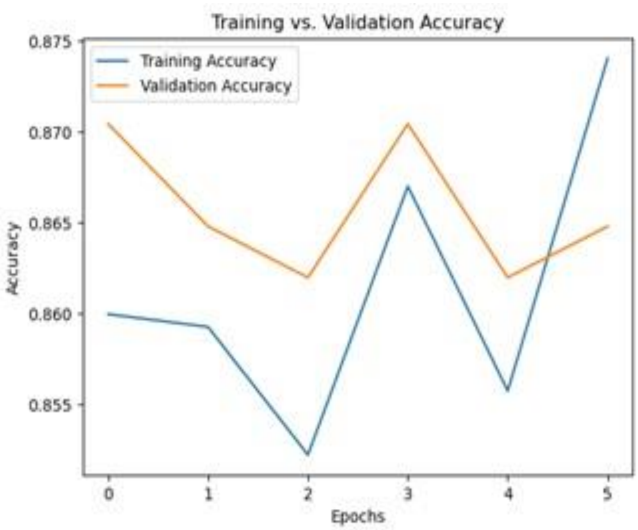
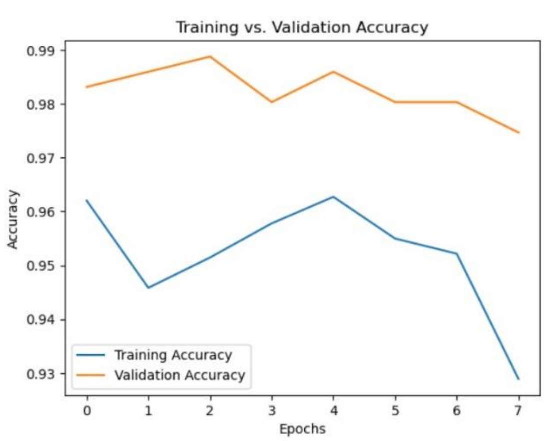


Figure 15: Training and validation accuracy for InceptionV3 show consistent improvement with minimal deviation between training and validation performance.



Challenges encountered during the project included:

- **Overfitting:** Initially, all models suffered from overfitting. This was mitigated through techniques like data augmentation, hyperparameter adjustments, and regularization.
- **Class Imbalance:** The imbalance in the dataset was addressed by targeted augmentation to ensure equal representation of all traffic classes.
- **Computational Costs:** VGG16 and ResNet50 had longer training times due to their computational complexity.
- **Data Collection Constraints:** Capturing high-traffic images posed challenges due to security concerns and potential conflicts with authorities.

Insights:

1. InceptionV3 and MobileNet demonstrated their suitability for traffic congestion prediction, balancing high accuracy with computational efficiency.
2. Data augmentation and preprocessing played a critical role in achieving consistent performance across models.
3. ResNet50's difficulty with Medium Traffic classification highlights the need for further optimization or alternative architectures for balanced traffic scenarios.

The presented results demonstrate the effectiveness of CNN-based transfer learning models for traffic congestion prediction, with InceptionV3 and VGG16 emerging as top performers. The insights gained from this research can guide future applications and improvements in traffic management systems.

IV. CONCLUSION AND FUTURE WORK

This research demonstrated the effectiveness of using Convolutional Neural Networks (CNNs) for traffic congestion prediction through image-based analysis. By leveraging a self-collected dataset of traffic images from key locations in Egypt, we were able to classify traffic levels into high, medium, and low congestion with significant accuracy. Among the models evaluated, VGG16 and InceptionV3 exhibited the highest performance, achieving test accuracies of 98.37% and 98.65%, respectively. MobileNet also performed exceptionally well, achieving a test accuracy of 96.39%, while DenseNet proved robust with 95.95%. ResNet50 showed moderate results, with a test accuracy of 85.14%, likely due to its sensitivity to class imbalances and computational demands.

The success of this study was also attributed to extensive data preprocessing and augmentation techniques, which resolved challenges such as class imbalance and overfitting. Despite these efforts, challenges like the computational expense of certain models and difficulties during data collection in high-traffic periods highlight areas for future improvement.

For future work, this research could be expanded to integrate traffic congestion prediction into a **Speed Management System** for roads in Egypt. By dynamically adjusting road speed limits based on real-time traffic congestion levels, such a system could help optimize traffic flow, reduce accidents, and improve overall road safety. Further enhancements could also involve expanding the dataset to include additional locations and incorporating real-time video feeds to refine congestion prediction accuracy.

This study sets a foundation for leveraging deep learning models in traffic management systems, contributing toward smarter, data-driven urban planning initiatives in Egypt and beyond.

IIV. REFERENCES

[1] Y. Gao et al., *A novel image-based convolutional neural network approach for traffic congestion estimation*, Expert Systems with Applications, vol. 180, p. 115037, Oct. 2021.

[2] M. Akhtar and S. Moridpour, *A Review of Traffic Congestion Prediction Using Artificial Intelligence*, Department of Civil and Infrastructure Engineering, RMIT University, Melbourne, VIC 3000, Australia Jan. 2021.

[3] Ali M. S. et al., *Estimation of Traffic Congestion Cost-A Case Study of a Major Arterial in Karachi*, Fourth International Symposium on Infrastructure Engineering in Developing Countries, IEDC 2013.

[4] S. Yang, *On feature selection for traffic congestion prediction*, Transportation Research Part C: Emerging Technologies, Volume 26, Pages 160-169, Jan. 2013.

[5] M. Anjaneyulu, M. Kubendiran, *Short-Term Traffic Congestion Prediction Using Hybrid Deep Learning Technique*, School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology (VIT), Vellore 632014, India, Dec. 2022.

[6] C. Len et al., *MobileNetV2 with Spatial Attention module for traffic congestion recognition in surveillance images*, Expert Systems with Applications Volume 255, Part D, 1 Dec. 2024.

[7] S. B. Kelengate, R. Rismiyati, *Transfer learning with VGG16 and InceptionV3 for traffic sign classification*, 12TH INTERNATIONAL SEMINAR ON NEW PARADIGM AND INNOVATION ON NATURAL SCIENCES AND ITS APPLICATIONS (12TH ISNPINSA): Contribution of Science and Technology in the Changing World, Volume 3165, Issue 1, Semarang, Indonesia, Oct. 2022.