International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 11 | November - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

A Comparative Study of Neural Networks for Emergency Vehicle Detection in Traffic

Shreya Palande, Rhea Coutinho
Fr. Conceicao Rodrigues College of Engineering

Abstract:

This study explores the application of Neural Networks for efficient emergency vehicle identification. Utilizing a dataset of emergency and non-emergency vehicle images, the model demonstrates high accuracy in classification. A noteworthy use case involves real-time integration with traffic systems, where the identification of an emergency vehicle triggers an automatic change of traffic lights to green, facilitating swift and unimpeded passage.

I. INTRODUCTION

The implementation of Convolutional Neural Network (CNN) models for emergency vehicle detection in video content stands as a pivotal technological advancement with multifaceted advantages. Primarily, it plays a crucial role in bolstering the safety and efficiency of emergency response services. Rapid and accurate identification of emergency vehicles like ambulances, fire trucks, and police cars in video streams facilitates quicker response times, especially in life-threatening situations. Recognizing these vehicles promptly and affording them priority on the road significantly improves the operational effectiveness of emergency services, potentially leading to lives saved and reduced response times to critical incidents.

The adoption of CNN-based models for emergency vehicle detection extends its impact to overall road traffic safety. Identifying emergency vehicles in video feeds enables nearby vehicles to be alerted, allowing drivers to yield the right of way and create space for the emergency vehicle. This not only reduces the likelihood of accidents but also minimizes traffic congestion, helping prevent gridlock in urban areas. Moreover, it plays a pivotal role in upholding traffic regulations, discouraging the unauthorized impersonation of emergency vehicles for unjust advantages on the road. The environmental benefits are also substantial, as optimizing the movement of emergency vehicles through real-time video analysis can curtail fuel consumption and emissions, resulting in a decreased carbon footprint and contributing to a cleaner and more eco-friendly environment. The swifter response times facilitated by effective detection and traffic management further reduce the need for emergency vehicles to idle or take less efficient routes, mitigating their overall environmental impact.

The integration of CNN models for emergency vehicle detection in video content is indispensable for ensuring the safety, productivity, and overall well-being of the public and emergency services. This application of advanced computer vision technology exemplifies how AI-driven systems can profoundly and positively influence real-world scenarios, ultimately benefiting society at large.

II. RELATED WORKS

The study led by Hari Vignesh K, Musharraf U, and Balaji A [1] aims to develop an emergency vehicle detection system using CNN Xception. It employs a diverse dataset for training, covering scenarios like heavy traffic and various lighting conditions, enabling real-time detection with bounding boxes and vehicle tracking across frames. Despite concerns about overfitting, the system demonstrates the ability to generalize, albeit with slightly lower validation accuracy. The system's efficacy relies on a comprehensive dataset, and insufficient coverage may impact detection accuracy. It assumes high-quality video feeds for accurate detection, and lower quality may affect performance. Real-time performance is crucial for emergency response, with potential delays based on computational resources and video complexity. While promising for heavy traffic scenarios, addressing these limitations is vital for optimal performance and accuracy.

The paper "Detection of Active Emergency Vehicles using Per-Frame CNNs and Output Smoothing" [2] introduces an efficient EV detection system with pre-processing, inference, and post-processing phases. The pre-processing involves a tracking system generating 3D coordinates, projecting them onto images, and resizing to 224x224 pixels. The inference module, based on ResNet-18, identifies active EVs, and the post-processing includes a downstream smoother with an average latency of under 10 milliseconds. The EV classifier uses ResNet-18 with pre-trained weights, focal loss, and data augmentation, resulting in a significant 5.73% improvement in the max-F1 score. Techniques like mining data, event labeling, and iterative model retraining address false positives, emphasizing the model's effectiveness in mitigating errors. The downstream smoother enhances accuracy, demonstrated by true positives and challenging true negatives in the results.

Smoother threshold T	% change of precision	% change of recall	% change of F1 score
0%	0	0	0
30%	10.30	-0.68	5.07
50%	9.87	-0.51	4.87
70%	6.80	-20.0	-7.02

Figure 1. Per-actor results of the output smoother [2]

Kaushik S, Abhishek Raman, and Dr. Rajeswara Rao K.V.S [3] develop a model for emergency vehicle detection, utilizing a dataset of 400 annotated images. The ResNet50 architecture forms the core, with Stochastic Gradient Descent and mini-batch processing for training. Bounding boxes and masks aid in discerning and classifying emergency vehicles. The architectural framework incorporates the Fast RCNN and the Region Proposal Network (RPN). The ResNet backbone with ReLU activation significantly contributes to overall accuracy. The model achieves an impressive 81% accuracy in object detection and an outstanding 92% accuracy in instance segmentation, showcasing its proficiency.

	True positive	True negative	False positive	False negative
Object Detection	74	7	8	11
Instance Segmentation	83	9	6	2

Figure 2. Tabular representation of accuracy [3]

The paper "A New Hybrid Architecture for Real-Time Detection of Emergency Vehicles" [4] presents a two-stage hybrid algorithm for emergency vehicle detection that combines machine learning and image processing. Operating at five frames per second using ResNet and the YOLO algorithm, the model crops the region of interest, applying an OCR module to detect "Ambulance" text, and can flip mirrored images for improved recognition. Upon identifying "AMBULANCE," control information is transmitted to the traffic controller. The approach addresses challenges in existing traffic management, stressing the need for adaptive systems and highlighting deficiencies leading to accidents and delays for emergency vehicles. The hybrid architecture aims to minimize the search space, enhance real-time detection, and improve traffic control by integrating OCR and a centroid-based method for ambulance detection. Despite challenges like fixed signals, neglect of emergency vehicles, detection issues, and limited datasets, the hybrid model holds promise in advancing traffic management and prioritizing emergency response.

The researchers Shuvendu Roy and Md. Sakif Rahman [5] utilized the COCO dataset, known for its comprehensive object detection capabilities, encompassing 330,000 diverse images. A pre-trained model with 80 classes for object detection, including various vehicles, is employed. Emergency vehicle classification adopts a binary approach, distinguishing between emergency and regular vehicles. The regular vehicle class is trained on the Stanford University Cars dataset, while emergency vehicle images are sourced from Google and Wikipedia, with 90% used for training and 10% for testing. Data augmentation techniques enhance model robustness. The baseline, a 2-layer CNN, undergoes experiments with different hyper-parameters, utilizing the Adam optimization algorithm with a learning rate of 0.001. Findings suggest the optimal performance of the simple model with an image dimension of 64 × 64, facing challenges in effectively capturing patterns in larger images due to constrained parameters and architecture.

III. DATASET USED

In the research dataset employed, a total of 1646 images depicting both emergency and non-emergency vehicles were meticulously curated. The dataset exhibits a balanced distribution, comprising 965 images categorized as class 0, representing non-emergency vehicles, and 681 images designated as class 1, signifying emergency vehicles.

To ensure a comprehensive evaluation, the dataset was meticulously partitioned into training and testing subsets, employing an 80-20 split ratio. This strategic division was executed with meticulous care to uphold the original class distribution, thereby preserving the inherent balance between emergency (class 1) and non-emergency (class 0) vehicle samples in both the training and testing sets. This approach aims to foster a representative and unbiased assessment of the model's performance across diverse scenarios.

Figure 3. Dataset used for detection

The Xception structure incorporates depth wise separable convolutions, a unique method that divides the convolution process into depthwise and pointwise convolutions. This division enhances computational efficiency while preserving expressive power, making Xception particularly beneficial in situations with limited computational resources. The model's efficiency extends to its adept use of parameters and computations, contributing to its effectiveness. Furthermore, it facilitates an expanded receptive field, enabling the capture of long-range dependencies in input data for enhanced pattern recognition. Frequently employed in transfer learning, Xception utilizes pre-training on extensive datasets such as ImageNet, followed by fine-tuning on smaller datasets. Apart from image classification, Xception's versatility spans various computer vision tasks, and its strong performance has been verified in benchmarks. It is conveniently accessible in the Keras library, providing access to pre-trained models or the development of customized models based on its architecture.

IV. METHODOLOGY

This research endeavor employs a dual-model approach for the detection of emergency vehicles, harnessing the capabilities of both the Xception model and an Artificial Neural Network (ANN). This methodological choice is rooted in the distinct strengths and architectures of these models, with the overarching goal of achieving robust and accurate detection performance. The Xception model, renowned for its depth and efficiency in feature extraction, is juxtaposed with the versatile and adaptable architecture of the ANN.

The first step is curating a diverse dataset encompassing images of both emergency and non-emergency vehicles followed by data standardization. Uniformly resize images to a predetermined resolution, and normalize pixel values to a standardized scale. Moreover, applying data augmentation techniques, such as rotation and zooming, to enhance the dataset's richness and promote model generalization. Further,sSegregate the dataset into training and testing subsets, preserving the balance of emergency and non-emergency vehicle instances in each subset.

Implement suitable deep learning architecture for image classification. This research implements transfer learning by utilizing a pre-trained Xception model as a feature extractor. The model is fine-tuned by incorporating additional layers and finally trained after specifying the optimizer, loss function and metrics. Alternatively, this research also includes a custom Artificial Neural Network (ANN) architecture tailored to the specifics of emergency vehicle detection. Train the selected models on the training dataset, monitoring pertinent training metrics like accuracy and loss. Finally, evaluate the trained models on the testing dataset, scrutinizing performance metrics such as accuracy, precision, recall, and F1 score.

V. IMPLEMENTATION

Xception:

The Xception model, characterized by its extreme inception, excels through a groundbreaking architecture. It leverages separable convolutions for parameter efficiency, employs a bottleneck structure to reduce dimensionality effectively, and incorporates residual connections for streamlined gradient flow. The model's entry and middle flows capture intricate features, while global average pooling ensures compact yet informative representations, collectively contributing to its exceptional performance in various computer vision tasks.

The pre-trained Xception model, initially trained on the ImageNet dataset, serves as a feature extractor due to its adeptness at capturing intricate hierarchical features within images. The model is compiled with a binary cross entropy loss function, suitable for binary classification tasks, and the RMSprop optimizer with a learning rate of 0.0001. The metric chosen for evaluation is accuracy (acc). The images are preprocessed using Keras image preprocessing module which resizes it to the specified target size (350, 350, 3). The third dimension (3) corresponds to the number of color channels (RGB). Further, the loaded image is converted into a NumPy array, and then normalizes the pixel values to be suitable for feeding into a deep learning model. The Xception model gives training accuracy of 99.05% and validation accuracy of 96.97%. This model is used for classification in test data, it gives accuracy of 95%

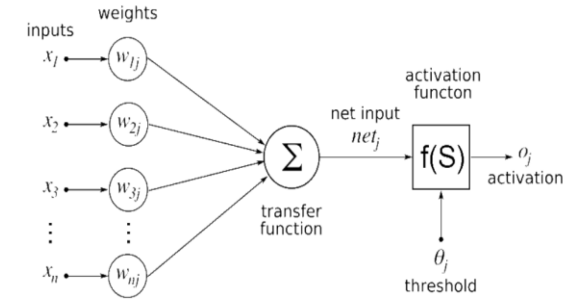
International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 11 | November - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

```
Accuracy: 0.95
Confusion Matrix:
[[192   6]
 [ 10 122]]
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.97      0.96       198
           1       0.95      0.92      0.94       132

    accuracy                           0.95       330
   macro avg       0.95      0.95      0.95       330
weighted avg       0.95      0.95      0.95       330
```

Figure 4. Results of Xception model for emergency vehicle identification

Artificial Neural Network (ANN):
Artificial Neural Networks (ANNs) can be effectively used for emergency vehicle detection by leveraging their ability to learn complex patterns and representations from images. Artificial Neural Networks are made up of layers and layers of connected input units and output units called neurons. First, initialize weights and biases with random values. These parameters are the learnable components of the neural network. An activation function is applied to the output of each neuron to introduce non-linearity. This paper applies the sigmoid function for implementing binary classification. Forward and backward propagation methods are defined. The main purpose of forward propagation is to make predictions based on the input data. The role of backward propagation is to adjust the weights and biases of the network to minimize the difference between predicted and actual outputs. It is the learning phase where the network learns from its mistakes.The proposed model gives a training accuracy of 83.13% and testing accuracy of 68.78%.

Figure 5. Working of Artificial Neural



Networks (ANNs)

```
# test accuracy
test_model(WB, x_test, y_test, layers)
```

```
Error:  0.20278901657410606      Accuracy:  0.6878787878787879
```

```
plot_graphs(history, "accuracy")
plot_graphs(history, "error")
```

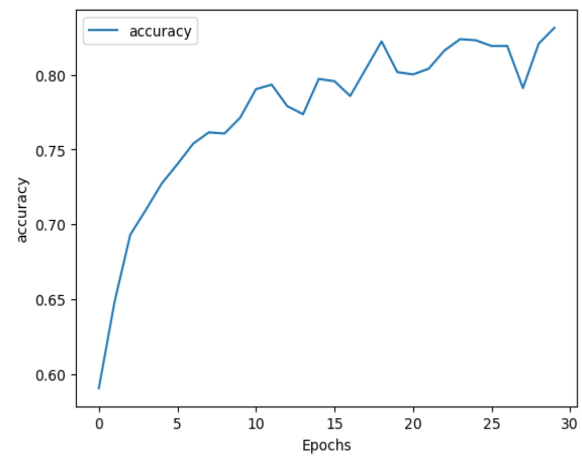


Figure 6. Results of Artificial Neural Network (ANN) model for emergency vehicle identification

VI. CONCLUSION

In summary, the Xception model outperformed the Artificial Neural Network (ANN) model, demonstrating superior accuracy in our research study.

Furthermore, the superior accuracy achieved by the Xception model can be attributed to its advanced architecture, which leverages separable convolutions, bottleneck structures, and residual connections. These design elements enhance the model's ability to capture intricate patterns and representations within the data, contributing to its improved performance compared to the more conventional Artificial Neural Network (ANN) model.

VII. REFERENCE

[1] Hari Vignesh K, Musharraf U , Balaji A , "Emergency Vehicle Detection using Deep Learning" , IJARIIE-ISSN(O)-2395-4396 Vol-9 Issue-2 2023

[2] Meng Fan, Craig Bidstrup, Zhaoen Su, Jason Owens, Gary Yang, Nemanja Djuric, "Detection of Active Emergency Vehicles using Per-Frame CNNs and Output Smoothing" , 2022

[3] Kaushik S, Abhishek Raman, Dr. Rajeswara Rao K.V.S , "Leveraging Computer Vision for Emergency Vehicle Detection - Implementation and Analysis" 11th ICCCNT 2020

[4] Eshwar Prithvi Jonnadula, Pabitra Mohan Khilar, "A New Hybrid Architecture for Real-Time Detection of Emergency Vehicles" , 2020

[5] Shuvendu Roy, Md. Sakif Rahman, "Emergency Vehicle Detection on Heavy Traffic Road from CCTV Footage Using Deep Convolutional Neural Network" , International Conference on Electrical, Computer and Communication Engineering, 2019

[6] François Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions" , IEEE Conference on Computer Vision and Pattern Recognition , 2017

[7] K Sanjay, V Madhavan, Tariq Muhammad, K Shankar, " Emergency Vehicle Detection in Heavy Traffic using Deep ConvNet2D and ComputerVision" , Proceedings of the International Conference on Innovative Computing & Communication, 2022