Deep Learning-Based Vehicle Accident and Traffic Image Classification

N. V. M. B. VITHANAWASAM and T. KOKUL
Department of Computer Science, University of Jaffna, Jaffna, Sri Lanka.
madushkavithanawasam148@gmail.com

Introduction

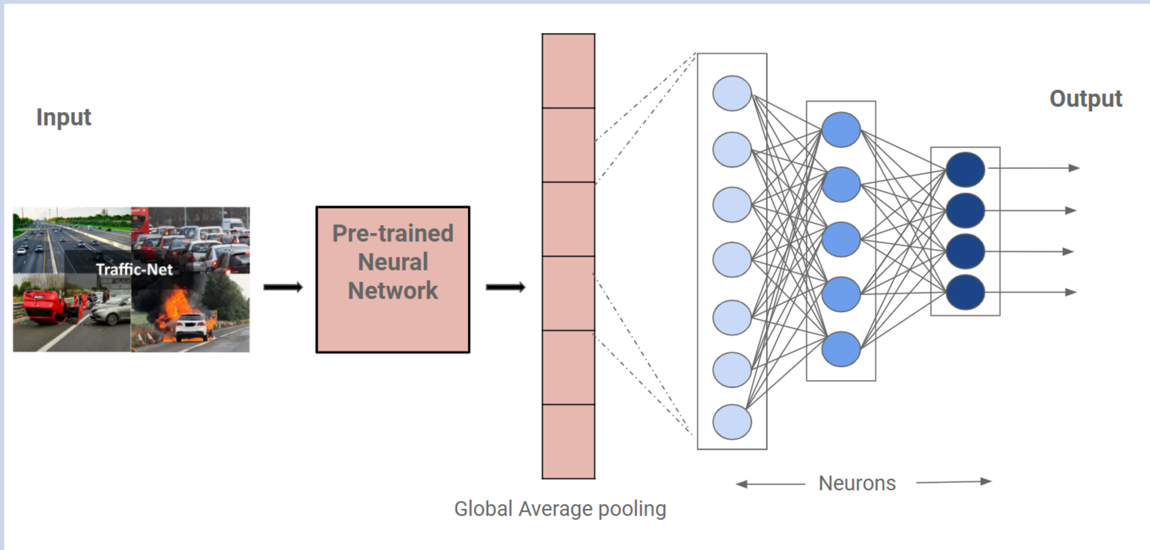
- Road traffic accident cases are a big local and global issue. Because of that, globally, every year the lives of approximately 1.19 million people are cut short due to a road traffic crash. Between 20 and 50 million more people suffer non-fatal injuries, with many incurring a disability.
- In this situation, improving traffic surveillance and accident detection systems using deep learning could help reduce both human and economic losses.
- An automated traffic and accident classification system not only enhances the traffic monitoring but also improves the decision-making process for emergency responding, ensures to supply medical support at the correct time, and increases the safety for the accident victims.
- In this particular instance, we explore the use of deep learning techniques for vehicle accident and traffic image classification, comparing multiple models to identify the best-performing architecture.
- **Objective**: To compare the performance of VGG16, ResNet152, and DenseNet 121 models for vehicle accident and traffic image classification.
- **Aim**: Use deep learning techniques to improve the classification of traffic incidents.
- **Scope**: Enhance the accuracy of accident classifications and enhance the efficiency of emergency support systems.

Dataset

- Source: Traffic-Net Dataset. [5]
- 4,400 traffic images in 'jpeg' format in four classes. (Accident, Fire, Sparse Traffic, Dense Traffic)
- Training/Test Split: 900 images for training and 200 for testing in each category.

Methodology

- Initially, we performed the dataset preparation. The training set is split into training and validation subsets with an 80-20 split.
- Image augmentation techniques are used to prevent overfitting, such as rescaling pixel values, random rotation, shifts, shearing, zooming, horizontal flips, and nearest neighbor filling.



Overview of the model

- The ResNet152, DenseNet121, and VGG16 pre-trained on ImageNet are used separately. Those are used as the backbones of the model for image feature extraction. (Transfer learning, channel attention mechanism)
- A final output layer with 4 neurons and a SoftMax activation function is used for multi-class classification.
- With the use of the Adam optimizer and loss function sparse categorical cross-entropy, we handle the multiclass classification.
- For optimizing training, we used several callbacks, like early stopping and model checkpoints. Early stopping will prevent overfitting by halting training. It halts when performance on the validation stops improving. Model checkpoints were used to save the, best model with the best validation accuracy during training.
- Evaluation of the model performed based on test dataset with test accuracy, confusion matrix, and classification report. (Recall, Precision, F1-score)

Experimental Setup

- The experiments were done separately on,
  VGG16| ResNet152| DenseNet121
- Optimizer: Adam |Batch Size:32 | No of Epochs:80
- Dropout:0.3-0.5
- Loss Function: Categorical Cross Entropy
- Activation Function: ReLU, Softmax
- Early Stopping: Enabled for improved efficiency
- Language: Python | Framework: Keras-Tensorflow
- IDE: Google Colab

Results

The performance of the proposed approach is assessed using classification accuracy. The accuracy expressed mathematically as follows,

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Model	Accuracy
ResNet152	0.927
DenseNet121	0.862
VGG16	0.909

Base model, in all the pre-trained models are frozen and we unfroze the last few layers. The last convolutional blocks were fine-tuned with training data samples. The best classification accuracies of the pre-trained models are compared above table .

Image Type	Precision	Recall	F1-Score	Accuracy
Fire	0.95	0.88	0.91	0.927
Sparse-traffic	0.93	0.94	0.94	
Dense-traffic	0.96	0.96	0.96	
Accident	0.87	0.93	0.93	

As shown in the above first table, ResNet152 achieves the best accuracy compared to the other two models. Also, the second table shows the precision, recall, F1-score, and accuracy metrics for ResNet152.

Discussion & Conclusion

- In this study, when we use a model without callbacks or early stopping, we can miss the models with good validation performance.
- In the ResNet152 model without the channel attention mechanism, we achieved a validation accuracy of 0.901 and a test accuracy of 0.87.But That test result shows a little bit low accuracy compared to the model which includes the channel attention mechanism.
- Channel attention mechanisms make the model more accurate and reliable when making predictions in image classification tasks. Test accuracy shows how the model reacts to unseen data.
- In summary, the developed ResNet152 model with channel attention is an effective solution for traffic event classification. The model shows high performance compared to the other architectures.

References

[1] B. Kumeda, Z. Fengli, A. Oluwasanmi, F. Owusu, M. Assefa and T. Amenu, "Vehicle Accident and Traffic Classification Using Deep Convolutional Neural Networks," 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing, Chengdu, China, 2019, pp. 323-328, doi: 10.1109/ICCWAMTIP47768.2019.9067530.

[2] Fengyun Cao, Sijing Chen, Jin Zhong, Yikai Gao, "Traffic Condition Classification Model Based on Traffic-Net", Computational Intelligence and Neuroscience, vol. 2023, Article ID 7812276, 13 pages, 2023. https://doi.org/10.1155/2023/7812276.

[3] M. Zheng et al., "Traffic Accident's Severity Prediction: A Deep-Learning Approach-Based CNN Network," in IEEE Access, vol. 7, pp. 39897-39910, 2019, doi: 10.1109/ACCESS.2019.2903319.

[4] Kanakala, R., & Reddy, K. (2023). Modelling a deep network using CNN and RNN for accident classification. Measurement: Sensors, 100794.

[5]Olafenwa Moses, Traffic-Net Dataset https://github.com/OlafenwaMoses/Traffic-Net .