Educational Administration: Theory and Practice
2024, 30(1) 3527 - 3533
ISSN: 2148-2403
https://kuey.net/
Research Article

Deep Learning Model Enabled for Distracted Driver Behavior Detection

Kirti Singh[1*], Dr. Amit Singhal[2]

[1*]Research Scholar Email- ks4may@gmail.com
[2]Professor, Head Of Department, MONAD University, Hapur U.P. Email- hod.cs@monad.edu.in

Citation: Kirti Singh, Dr. Amit Singhal (2024), Deep Learning Model Enabled for Distracted Driver Behavior Detection, *Educational Administration: Theory and Practice*, *30(1)* 3527 - 3533
Doi: 10.53555/kuey.v30i1.7327

ARTICLE INFO **ABSTRACT**

This research presents a novel method for detecting distracted driver behaviour by utilising Deep Learning (DL) models. The proposed architecture comprises three deep learning models: AlexNet, VGGNet and ResNet. These models utilise a state farm dataset that includes data on one safe driving class and four risky behaviours namely Talking on the phone, Texting on the phone, Turning and Other activities. Distracted driving remains a major contributor to road accidents worldwide. However automated detection technologies show great potential in improving road safety. The experimental results indicate that AlexNet and VGGNet achieved accuracy rates of 98% as well as 98.4% respectively showcasing their efficacy in detecting distracted driving behaviours. This study highlights the significance of choosing deep learning architectures that are customised for certain objectives. Our proposed models have been compared to existing state-of-the-art works and the study confirms that our models perform competitively. The results add to the progress of automated systems designed to decrease accidents resulting from distracted driving highlighting the potential of deep learning in enhancing road safety.

Keywords:. *Distracted Driver; AlexNet; VGGNet; Deep Learning; ResNet.*

1. Introduction

The majority of collisions and fatalities have occurred as a result of drivers being inattentive or distracted driving. 90% of all accidents globally are caused by human mistakes which is the primary cause of most traffic accidents. WHO reported in 2020 that automobile accidents resulted in around 1.35 million deaths annually. Every year reckless driving results in one million deaths and over 50 million injuries worldwide [1]. Any activity that takes the focus away from driving a vehicle while driving is considered as distracted driving like eating, drinking, conversing with passengers, texting, adjusting the radio or using navigation. Both drivers and passengers may have grave safety issues when distracted driving occurs. Using a mobile phone is one of several activities that divert a driver's focus from the road and it is known to be a high-level distraction.
The study aims to automatically identify the causes of driver distraction by using the DL model. Although DL models have demonstrated impressive performance in image identification, they come with a cost in terms of computation time and data [6]. Nonetheless, methods like quantization, pruning and transfer learning can lessen DL's computing requirements [7]. Furthermore a great deal of research has looked into how to make DL more successful and efficient in a variety of fields [8]. DL can enhance representation and performance by mimicking the human brain's capacity to extract significant features which can lead to better decision-making across a wide range of applications [9].
There are numerous effective deep networks for classification such as AlexNet, ResNet and VGGNet. We present a new and effective approach to driver behavior classification in this paper. The driver behaviors are separated into five classes. The remaining sections are structured as follows: Section 2 presents the literature review. The framework of DL model is proposed in Section 3. Section 4 presents the experiment design and results. Section 5 concludes with future work.

2. Literature Review

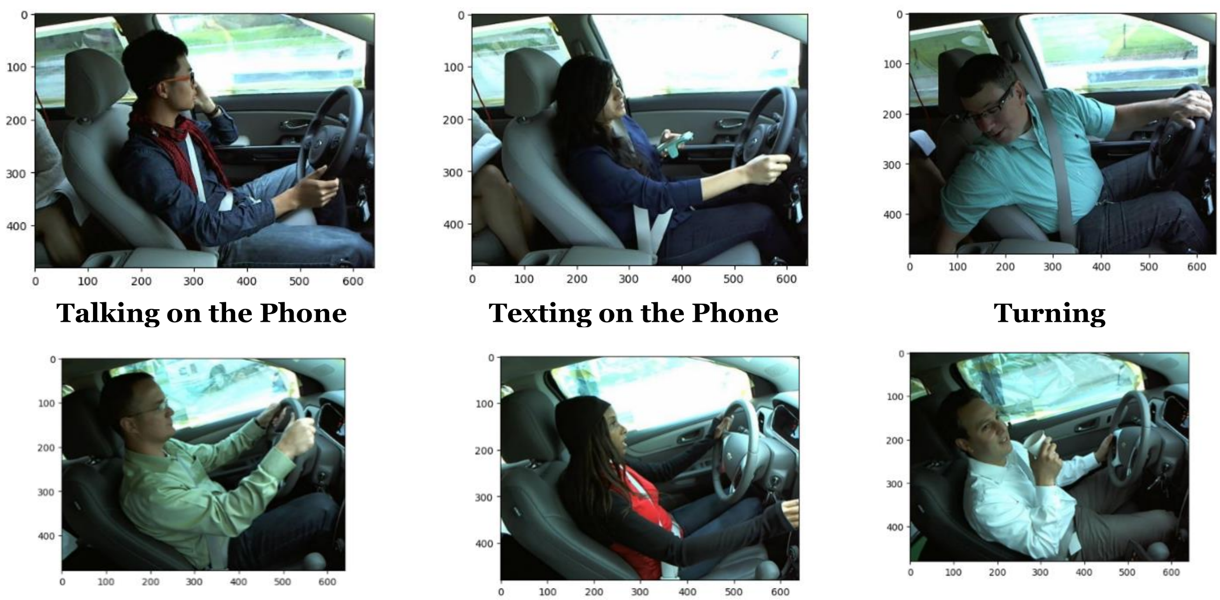
One of the hottest ML techniques in artificial intelligence is DL. DL is being widely applied in automatic driving, NLP and image identification. It performed well as evidenced. Ding et al. (2023) presented a lightweight ensemble model designed for real-time distracted driving detection achieving a notable accuracy rate of 99%. The study leverages the capabilities of pre-trained networks such as Inception, Xception, DenseNet and MobileNet tailored to address the computational and storage constraints of edge devices. The aim is to combine these lightweight models to achieve high accuracy, low latency and reduced parameter set [10].

Nawaf et al. (2023) introduced a novel architecture called U2-net an extension of U-net convolutional neural network which is tailored for detecting distracted drivers. Recognizing the increasing number of traffic accidents due to distracted driving globally the study aims to contribute the development of intelligent vehicles and safer roads by monitoring driver behaviors [11]. Apurva et al. (2023) delves into the intricate challenge of detecting cognitive distractions among drivers where drivers appear focused but are mentally disengaged. Unlike distractions cognitive distractions like fatigue or conversations are harder to identify as they don't necessarily divert a driver's eyes from the road. The study emphasizes the importance of continuously monitoring driver states to preemptively address potential hazards [12].
Hasan et al. (2024) addressed the challenge of recognizing distracted driving activities in real-world scenarios emphasizing the critical role of ensuring safety for both drivers and pedestrians. Their study acknowledges the limitations of traditional computer vision methods which often demand extensive annotated data and may lack scalability as well as generalization [13]. Ping et al. (2023) introduced an innovative distracted driving behavior recognition model that addressed the limitations of traditional DL methods particularly CNN-based approaches. They emphasized the importance of causal reasoning and robustness in identifying the distracted driving behaviors [14].
Abeer. A. Aljohani (2023) presented an advanced framework for detecting driver distractions by combining artificial deep learning and machine learning models with genetic algorithms (GA). Their study offered a hybrid approach by merging the strengths of various models to achieve high accuracy in classifying distracted driving behaviours and recognizing the driver's commands [15]. Gawtham et al. (2023) addressed the critical issue of distracted driving by introducing a lightweight DL approach tailored for resource-constrained devices like Raspberry Pi and mobile phones. Their study leverages the SqueezeNet 1.1 architecture by employing transfer learning to adapt it for distracted driving detection with high efficiency [16].

3. Materials and Proposed Work

3.1. Materials

The dataset utilized in this work is Driver Behaviour Detection (DBD) which is sourced from Kaggle dataset website. The dataset focuses on detecting various behaviors exhibited by car drivers during driving. Driver behaviour detection is a specialized field aimed at identifying and analyzing the actions of drivers while they are behind the wheel. The dataset comprises five distinct classes to classify these behaviors: 1. Safe driving, 2. Talking on the phone, 3. Texting on the phone, 4. Turning and 5. Other activities. Some sample images are given in Figure 1.



Talking on the Phone **Texting on the Phone** **Turning**

Safe Driving **Other Activities** **Other Activities**

Figure 1. Five Distracted Driving Behaviors.

3.1.1. Data Splitting

The dataset has been divided into three subsets such as training, testing and validation purposes.
• Training Data: 75% of the dataset with a total of 8066 samples.
• Testing Data: 15% of the dataset with a total of 2148 samples.
• Validation Data: 5% of the dataset with a total of 537 samples.

3.1.2. Class Distribution

The number of samples for each class is as follows

- Safe Driving: 2203 samples
- Talking on the Phone: 2169 samples
- Texting on the Phone: 2203 samples
- Turning: 2057 samples
- Other Activities: 2119 samples

3.1.3. Data Frame Creation

The data has been organized into data frames for ease of use. Each data frame includes a label column to tag or classify the respective behavior. This format facilitates the utilization of the 'flow from dataframe' method which simplifies the data feeding into deep learning models. This structured approach for the data organization and splitting ensures a balanced distribution across classes as well as provides an ample data for training, testing and validation. We aim to create robust and accurate driver behaviour detection models.

3.2. The Proposed System

3.2.1. CNN

CNN sometimes referred to as ConvNet is a popular DL technique used in image processing applications. The method distinguishes between images by learning distinct weights and biases that correspond to different items contained in an input image. The algorithm uses the height, width and number of channels of an input image to create a matrix of pixel values.

The network is made up of several layers the first of which is the convolutional layer (CL) which receives input images. The CL is made up of a collection of convolutional kernels that cover a tiny portion of the image which are convolved across the whole thing to identify different characteristics.

3.2.2. AlexNet

Previously CNN was primarily meant to be used for jobs involving hand digit recognition because it didn't scale well to all image classes. To improve CNN's potential for learning the AlexNet model was released in 2012. It demonstrated strong performance on tasks including picture identification and classification. The CNN became more sophisticated with the introduction of several parameter optimization techniques by the AlexNet model. CNN was made suitable for a variety of image classification tasks by increasing the number of feature extraction steps from five (in LeNet) to seven (in AlexNet). But when the model's depth increases issues such as overfitting and vanishing gradient descents appear. To lessen overfitting as well as overlapping techniques like subsampling and local response normalization are applied alongside the large-sized filters (11 9 11 or 5 9 5) are added to the first few layers.

3.2.3. VGGNet

The architecture from Visual Geometry Group, Oxford (VGG Net) [5] demonstrates the improvement over AlexNet by substituting several 3×3 kernels-sized filters for huge kernel-sized filters. It demonstrates that numerous stacked smaller size kernels perform better than one larger size kernel with a given receptive field. In this design partial pooling is used for some of the intermediate convolution layers and 1×1 convolution filters are used. Three Fully Connected layers the first two with 4096 channels and the third with 1000 channels which follow a stack of convolutional layers to complete 1000-class ILSVRC classification.

3.2.4. ResNet

ResNet introduced the idea of residual learning to address the computational complexity of the problem that plagued the offered earlier models. ResNet has less computational cost which despite being a CNN with 152 layers making it 8 times deeper than VGGNet and 20 times deeper than AlexNet respectively. The introduction of residual blocks which allows network to address the issue of vanishing gradients in very deep networks is the main novelty of the ResNet architecture. The vanishing gradient problem which occurs when gradients decrease as they backpropagate through multiple layers and impede learning makes training CNNs more difficult as they become deeper.

4. Discussion and Experimental Results

4.1. AlexNet

The DL model is designed for real-time distracted driving detection. It uses convolutional layers to extract spatial features from 240x240 RGB images progressing from capturing edges to complex patterns. Batch normalization stabilizes activations while max pooling reduces feature map dimensions. A flattened layer prepares data for fully connected layers which is classified into five distraction types. Dropout layers prevent overfitting with 24.7 million parameters. Its complexity captures detailed patterns but requires significant computational resources. Understanding its layered architecture highlights the potential for improving road safety through future work which involves optimizing, interpreting and deploying the model effectively. We can plot the output of the model for each epoch. For this matter we can extract the training loss, valid loss and accuracy. Then we plot both of them to find out the learning path.

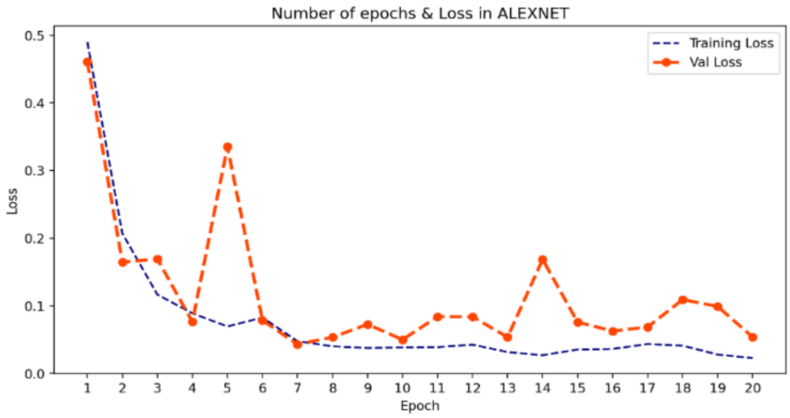


Figure 2. Number of epochs & Loss plot of Suggested AlexNet

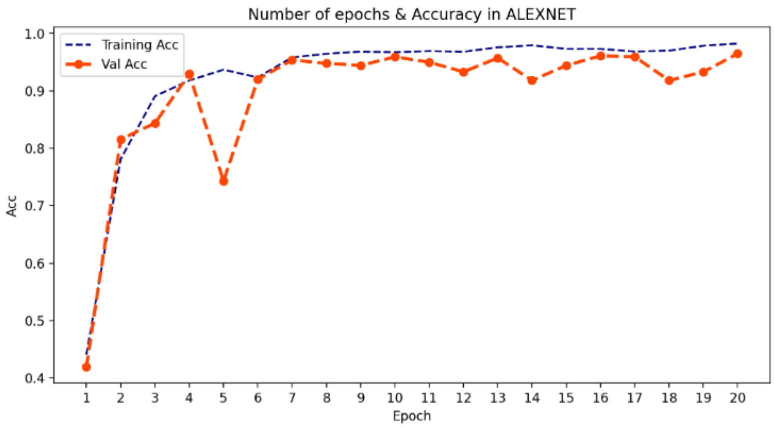


Figure 3. Number of epochs & Accuracy plot of Suggested AlexNet

4.2. VGGNet

The VGG architecture is the basis of ground-breaking object recognition models. Developed as a deep neural network VGGNet surpasses baselines on many tasks and datasets beyond ImageNet. Moreover it is still one of the most popular image recognition architectures. This is designed for distracted driving detection with a series of convolutional layers extracting features from 240x240 RGB images. It starts with two 64-filter convolutional layers followed by batch normalization and max pooling. The model then deepens with 128 and 256-filter layers each set with batch normalization and max pooling. This is followed by a set of 512 filter layers and another max pooling. A flattened layer prepares for dense layers with two 4096-neuron dense layers and dropout to prevent overfitting. The final layer is classified into five distraction types with 194.3 million parameters. It is complex but demands significant computational resources.

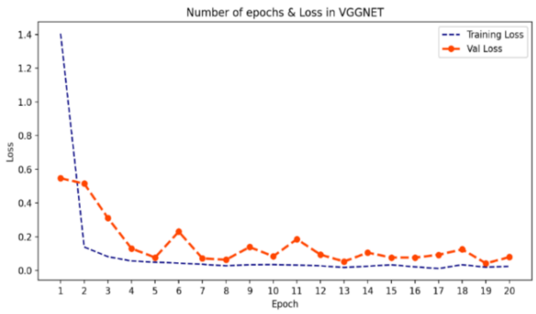


Figure 4. Number of epochs & Loss plot of Suggested VGGNet

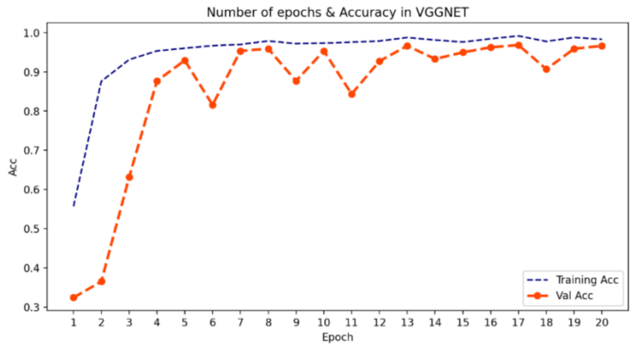


Figure 5. Number of epochs & Accuracy plot of Suggested VGGNet

4.3. ResNet

This results in training a very deep neural network without the problems caused by vanishing/exploding gradients. ResNet is widely used in image classification tasks and has inspired many other CNN architectures that have been developed since its introduction. ResNet is designed for distracted driving detection with a series of convolutional layers starting from 64 filters repeating with batch normalization and max pooling. It progresses through 128 and 256-filter layers maintaining the same pattern and ends with 512-filter layers. After feature extraction two 4096-neuron dense layers with dropout are used for classification into five distraction types. With 160.8 million parameters it's a balanced and slightly more compact model compared to others.

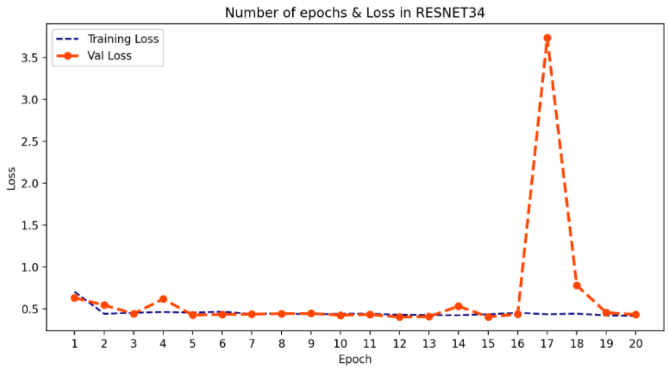


Figure 6. Number of epochs & Loss plot of Suggested ResNet

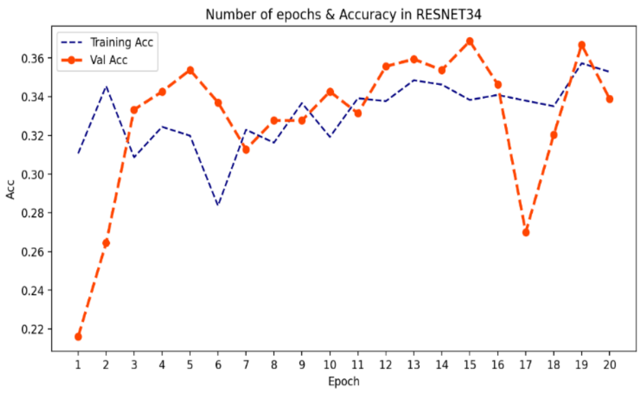


Figure 7. Number of epochs & Accuracy plot of Suggested ResNet

4.4. Model Comparison of Accuracy and Loss

Comparing the behaviour of the model outputs can be done by plotting all of the findings together in a single plot. Figure 7 illustrates that the results from AlexNet and VGGNet are satisfactory whereas the ResNet model doesn't fit well. This underscores the importance of using multiple-model architectures.

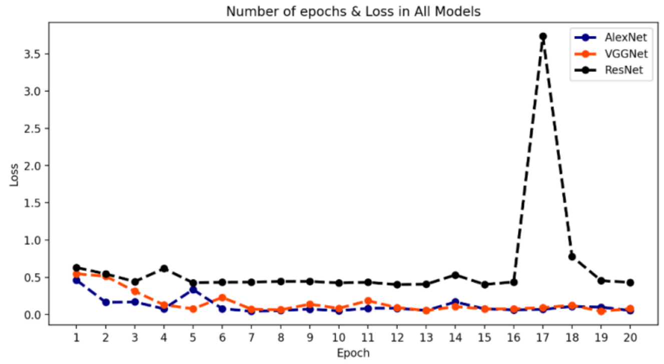


Figure 8. Comparison of the Number of epochs & Loss plot

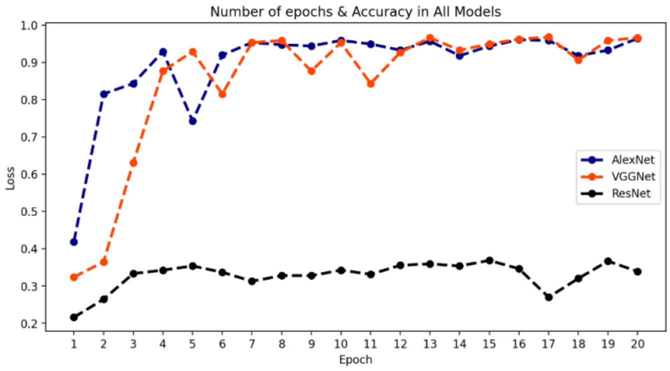


Figure 9. Comparison of the Number of epochs & Accuracy plot

Table 1. Comparison with State-of-the-Art Works

Reference	Dataset	Model	Accuracy
Proposed Model 1	DBD	AlexNet	98
Proposed Model 2	DBD	VGGNet	98.4
Proposed Model 3	DBD	ResNet	33
Janet et.al. (2020) [2]	SFDDD	CNN	97
Qin, B et al. (2021) [3]	MI-AUC	D-HCNN	95.59
Baheti et al. (2020) [4]	MI-AUC	SqueezeNet	93.21

Table 1 provides a comparative analysis of the proposed models with existing state-of-the-art works in distracted driving detection. The proposed Model 1 based on AlexNet achieves an accuracy of 98% matching the performance of Janet et.al.'s CNN on the SFDDD dataset. Model 2 built on VGGNet slightly improves the accuracy to 98.4% outperforming other models like Qin, B et al.'s D-HCNN on the MI-AUC dataset with an accuracy of 95.59%.

However, there is a notable drop in accuracy for Model 3 based on ResNet achieving only 33%. This stark difference in performance compared to the other proposed models and existing works like SqueezeNet by Baheti et al. at 93.21% suggests that ResNet may not be well-suited for this specific distracted driving detection task.

Overall, the proposed models based on ALEXNet and VGGNet demonstrate competitive performance with VGGNet showing a slight edge in accuracy. The significant disparity in ResNet's performance highlights the importance of selecting appropriate model architectures tailored to the specific task at hand.

5. Conclusion

This study presented a complete method for detecting distracted driver behaviour using deep learning models including AlexNet, VGGNet and ResNet. Given the growing concerns about distracted driving causing accidents and deaths on a global scale the use of automated detection systems is essential for improving road safety. The experimental findings obtained showed that both AlexNet and VGGNet achieved highly accurate results on the DBD dataset with accuracy rates of 98% as well as 98.4% respectively. These findings not only confirm the effectiveness of deep learning models in detecting distracted driving behaviours but also emphasise the possibility of utilising cutting-edge architectures. Nevertheless the ResNet model's notable decrease in performance resulting in an average 33% accuracy brings into question its appropriateness for this particular task. Choosing the appropriate model structure that is specifically designed for the dataset along with task needs is crucial. Our comparison investigation demonstrates the superior performance of our suggested models compared to regular CNNs and other sophisticated architectures like as D-HCNN and SqueezeNet. Potential future research in this field may involve investigating alternative deep learning structures, integrating more advanced preprocessing methods or even merging multiple models to enhance accuracy and resilience. In addition the immediate application and integration of these models in automobiles or intelligent gadgets could facilitate practical uses thereby preventing fatalities by decreasing occurrences of driver distraction. Ultimately this study adds to the expanding collection of studies focused on utilising deep learning to improve road safety. By implementing automated systems to identify distracted driving behaviours this work takes a step in making a substantial stride in promoting road safety and reducing avoidable accidents as well as fatalities.

Reference

1. Sahoo, G. K., Das, S. K., & Singh, P. (2023). A deep learning-based distracted driving detection solution implemented on embedded system. Multimedia Tools and Applications, 82(8), 11697-11720.
2. Janet, B., & Reddy, U. S. (2020, August). Real time detection of driver distraction using CNN. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 185-191). IEEE.
3. Qin, B., Qian, J., Xin, Y., Liu, B., & Dong, Y. (2021). Distracted driver detection based on a CNN with decreasing filter size. *IEEE transactions on intelligent transportation systems*, *23*(7), 6922-6933.
4. Baheti, B., Talbar, S., & Gajre, S. (2020). Towards computationally efficient and realtime distracted driver detection with mobilevgg network. *IEEE Transactions on Intelligent Vehicles*, *5*(4), 565-574.
5. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
6. Abdullah, B. W., & Ahmed, H. M. (2022). A Convolutional Neural Network for Detecting COVID-19 from Chest X-ray Images. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, *22*(3).
7. Abdulsahib, M. G., & Abdulmunim, M. E. (2022). Convolutional Recurrent Neural Networks for Text Lecture Summarization. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, *22*(2).
8. Obaida, T. H., Hassan, N. F., & Jamil, A. S. (2022). Comparative of Viola-Jones and YOLO v3 for Face Detection in Real time. *vol*, *22*, 63-72.
9. Kannan, K. S., & Parimyndhan, V. (2023, March). Modelling and Estimating of VaR Through the GARCH Model. In *International Conference on Advanced Engineering, Technology and Applications* (pp. 324-334). Cham: Springer Nature Switzerland.
10. Chen, D., Wang, Z., Wang, J., Shi, L., Zhang, M., & Zhou, Y. (2023). Detection of distracted driving via edge artificial intelligence. *Computers and Electrical Engineering*, *111*, 108951.
11. Alsrehin, N. O., Gupta, M., Alsmadi, I., & Alrababah, S. A. (2023). U2-Net: A Very-Deep Convolutional Neural Network for Detecting Distracted Drivers. *Applied Sciences*, *13*(21), 11898.
12. Misra, A., Samuel, S., Cao, S., & Shariatmadari, K. (2023). Detection of driver cognitive distraction using machine learning methods. *IEEE Access*, *11*, 18000-18012.
13. Hasan, M. Z., Chen, J., Wang, J., Rahman, M. S., Joshi, A., Velipasalar, S., ... & Sarkar, S. (2024). Vision-language models can identify distracted driver behavior from naturalistic videos. *IEEE Transactions on Intelligent Transportation Systems*.
14. Ping, P., Huang, C., Ding, W., Liu, Y., Chiyomi, M., & Kazuya, T. (2023). Distracted driving detection based on the fusion of deep learning and causal reasoning. *Information Fusion*, *89*, 121-142.
15. Aljohani, A. A. (2023). Real-time driver distraction recognition: A hybrid genetic deep network based approach. *Alexandria Engineering Journal*, *66*, 377-389.
16. Sahoo, G. K., Das, S. K., & Singh, P. (2023). A deep learning-based distracted driving detection solution implemented on embedded system. *Multimedia Tools and Applications*, *82*(8), 11697-11720.