Convolutional Neural Networks for Driver Behaviour Prediction: A Comparative Analysis of AlexNet, VGGNet and ResNet

Orlando Iparraguirre-Villanueva[1], Melquiades Melgarejo-Graciano[2]
[1,2] Facultad de Ingeniería, Universidad Tecnológica del Perú, Chimbote, Perú

Abstract*– Understanding driver activity in real time is complex, yet it is important for in-vehicle systems that aim to reduce car crashes. This work addressed the problem by relying on state-of-the-art methods, specifically, evaluating three convolutional neural network (CNN) models, such as: AlexNet, VGGNet and ResNet with the aim of predicting real-time driver activity and behaviour during driving. For the development, a dataset consisting of 10,751 images was used (The dataset was obtained from the Kaggle platform). To achieve good results, there are multiple factors, such as the volume of the dataset, the quality of the data, the application of optimisation techniques, among others. The findings of the proposal showed the VGGNet model to be the most efficient model for efficiently classifying and predicting driver's behaviour, achieving an accuracy rate of over 98%. It is closely followed by the AlexNet model, with an accuracy rate of 98%, a very significant result for this type of task, this model also obtained an F1-Score of 86%, and a count rate of 75%. However, the metrics obtained by the ResNet model are much lower than the compassion of the other models, it only achieved an accuracy rate of 35.28%, which indicates that it has limitations in identifying features and predicting driver behaviour. Finally, it is concluded that the VGGNet model slightly outperforms the AlexNet model, which shows that both models are efficient for this type of task.

Keywords-- convolutional neural networks; traffic accidents; prediction; driver behavior; automobiles.

I. Introduction

In recent years, road crashes have intensified significantly, creating a serious problem that affects millions of people around the world. Annually, around 1.35 million people lose their lives on the roads, and between 20 and 50 million are seriously injured, often resulting in permanent disabilities [1] [2]. This problem is further intensified in resource-poor countries, where the death rate is higher, and road crashes are among the leading causes of death [3] [4].

For example, in Korea, motor vehicle crashes generated costs of more than USD 23 billion, equivalent to 1.3% of its GDP [5]. In Saudi Arabia, car crashes were the second leading cause of death in 2019 [6], and in Pakistan, 15-16,000 deaths are recorded per year [7]. In Poland, in 2021 alone, more than 22 000 car accidents resulted in more than 2245 deaths and 26415 injuries [8].

The World Health Organization (WHO), a body that seeks to promote action plans to improve road safety, has predicted that by 2024, motor vehicle crashes could become the second leading cause of death and disability worldwide [9]. This problem is becoming more and more profound and affects the most vulnerable populations, such as pedestrians and pregnant women, among others [10], [11]. The UN, another body responsible for setting sustainable development goals, has set a target to reduce deaths and injuries related to motor vehicle crashes by half, although progress has been modest [12]. While it is true that the main drivers of car accidents are human error, this can be reduced thanks to technological progress, specifically with computer vision, which gives us the tools to develop intelligent solutions for vehicles. For example, in countries such as the USA and Switzerland, they are responsible for more than 50% of car accidents [13]. Anticipating such events is very important, and with artificial intelligence (AI) models we can predict and understand drivers' actions and behavior in real time [14].

In the last decade, AI has become a key tool in technological developments, specifically in classification models, such as deep CNNs, among others [15]. These technologies offer great potential to mitigate automobile accidents, through intelligent solutions embedded in vehicles [16]. This process has been made possible by the rapid development of computer technology, which has allowed AI to be integrated into tasks ranging from simple to complex decisions [17] [18]. AI not only seeks to replicate human intelligence, but is designed to enhance it, providing more efficient and effective tools for tackling real-world problems [19]. Its ability to process large volumes of data and find patterns makes it indispensable for decision making, especially in scenarios where time and accuracy are critical [20], and its contribution to automotive manufacturing is increasingly significant, and this can be seen in companies such as General Motors, Tesla, Waymo, BMW, Nissan, Toyota, Ford, among others. Soon, AI is expected to continue to evolve, endowing machines with human-like cognitive capabilities, which would extend their reach and usefulness [21]. This technology is already providing highly efficient innovative solutions to complex challenges such as autonomous driving, data analytics and in other fields [22]. One example is the use of CNN, a type of deep learning that is particularly effective for image recognition and processing [23].

This paper aims to compare the performance of three CNN models in driver behaviour detection. AlexNet, VGGNet and ResNet architectures were evaluated to determine which of them offers better performance in this task. In the initial section, the context of the problem addressed in this research is presented. In the second section, the literature review related to the prediction of car driver behaviour is presented. In the third section, the methodology used in the development of the case is discussed in detail. Then, in the fourth section, the results obtained are presented. Finally, in the last two concluding sections, the results are analyzed and discussed and the conclusions drawn from the research are presented.

II. Literature review

For example, in [24] they proposed a deep convolutional model that employs a multiple granularity strategy by recalibrating and fusing features to improve driver fatigue detection. The model achieved 89.42% accuracy. Similarly, [25] used a combination of ResNet50, DenseNet201 and InceptionV3 models together with XGBoost for driver behavior anomaly detection. The accuracy and precision metrics obtained showed an average value of 99.52%. Similarly, the authors of [26] detected the driver's emergency braking intention by using different CNN architectures, specifically ShallowNet, DeepNet and EEGNet. Among the models evaluated, EEGNet proved to be the most effective, predicting emergency braking 200 milliseconds in advance and with an accuracy of more than 80%. In addition, [27] aimed to classify driving behavior into defensive, normal and aggressive driving categories. A model combining stacked LSTM with a 4-layer CNN (CNN-2), optimized to capture patterns in the driving data, was employed. System performance achieved an accuracy of approximately 98%.

On the other hand, in [28] they sought to categorize driving behaviors into five different types: typical, aggressive, distracted, drowsy and intoxicated driving, using CNN. The CNN model achieved 99% accuracy in identifying these behaviors. Also, in [29] they focused on detecting driving behavior by using the Dilated Lighthead R-CNN (DL-RCNN) method. This method achieved a classification accuracy of 93.2%. Similarly, the authors of [30] proposed a driving behavior recognition algorithm that combines an attention mechanism with a lightweight network. This improvement of the YOLO-V4 model, integrated with MobileNet-V3, achieved an accuracy of 96.49%. In [31] they developed a system to detect drowsiness and distraction in drivers using a CNN-LSTM model. Results showed 93.61% accuracy.

In the next paper [32], they presented the D-CRNN architecture based on a deep neural network combining CNN and RNN to identify driving styles, seeking to highlight the need for diverse data to improve the generalization of the models. Similarly, in work [33], they trained Hybrid CNN, VGG16 and VGG19 models to detect real-time driver distractions in different situations. The results indicated that the VGG19 model obtained more than 98% accuracy among the trained models, followed closely by the Hybrid CNN model with 98% accuracy. Also, in [34] they proposed an in-vehicle sensor-based system obtained with CAN-BUS, this tool provides reliable measurements of driver behavior, such as acceleration, RPM, speed, pedal value and throttle position, among others. They used a hybrid deep learning model, which combines CNN and long-term memory (LSTM). The results highlight the potential of these models in the real world, greatly improving safety by providing information on driver behavior. In addition, this work contributes to the development of intelligent transport systems by providing a deeper understanding of driver behavior in various driving situations.

Along the same lines, in [35] they analyzed the LSTM and CNN architectures. They used three variables: vehicle progress, vehicle speed and vehicle acceleration, to predict the vehicle speed for the next 6 seconds. As a result, they determined that both LSTM and CNN architecture achieve significant performance and the complexity of the models influence the results, despite how complex capturing driver behavior can be. In the paper [36], the author proposed a deep CNN-based architecture to address the problem of classifying driver behaviors in real time. As input for the project, they used information captured through in-vehicle cameras. The results showed that the proposed project can classify nine different behaviors with an accuracy of 0.7397, and five behaviors with an accuracy rate of 0.8166. Also, the findings show that the proposal is computationally efficient, since the inferences are 15 ms, allowing it to satisfy the constraints of smart cars. Also, in the paper [37] they proposed a CNN and transformer-based driver attention prediction method, where they seek to address the need for interpretability of prediction models. The findings validate the effectiveness of the proposed method (ACT-Net), confirming that CNN architectures are effective for this type of task.

Finally, in [38] they presented an improved method based on YOLOv8 to detect both distracted behavior of drivers and their emotions. The approach achieved an accuracy of 0.814 for the detection of distracted driving and 0.733 for the identification of driver emotions.

III. Methodology

In this section, we present the methodology that provides the theoretical basis for the development of the case study, which is organized in two phases. First, a detailed description of the architecture (AlexNet, VGGNet and ResNet) that will be used to carry out the predictions related to driver behavior is presented. In the second and last part, an analysis of the dataset is performed.

A. *Convolutional Neural Networks*

CNNs have been widely used, showing great predictive power in image classification [39]. Their main advantage lies in their self-learning ability to extract features through convolution filters, which allows them to generate invariant representations of data in two and three dimensions [40]. Recent advances in CNN have significantly improved accuracy in image, text, speech processing and natural language recognition applications [41].

In CNNs, activation functions are key mathematical equations that capture non-linear connections, essential for processing information in a similar way to the brain [42], [43]. Clustering layers simplify data by reducing its dimensionality by grouping the most relevant inputs together, thus facilitating downstream processing [44] [45].. Fully connected layers, also called dense layers, combine the learned features to generate results such as classifications [46], [47]. To avoid overfitting, techniques are applied to regulate the complexity of the model and optimize its performance with new data [48], [49]. Finally, transfer learning reuses previously acquired knowledge in one task to improve effectiveness in another, maximizing the efficiency of the model [50].

B. AlexNet

AlexNet is a deep convolutional neural network architecture, proposed by Krizhevsky et al. [51], widely used for image classification. This network consists of five convolutional layers followed by three fully connected layers, interspersed with clustering and activation layers [52]. Recognized as one of the most influential CNNs, AlexNet has been used in several applications [53] and gained notoriety by winning first place in the large-scale visual recognition challenge [54]. Fig. 1 shows how this model works.

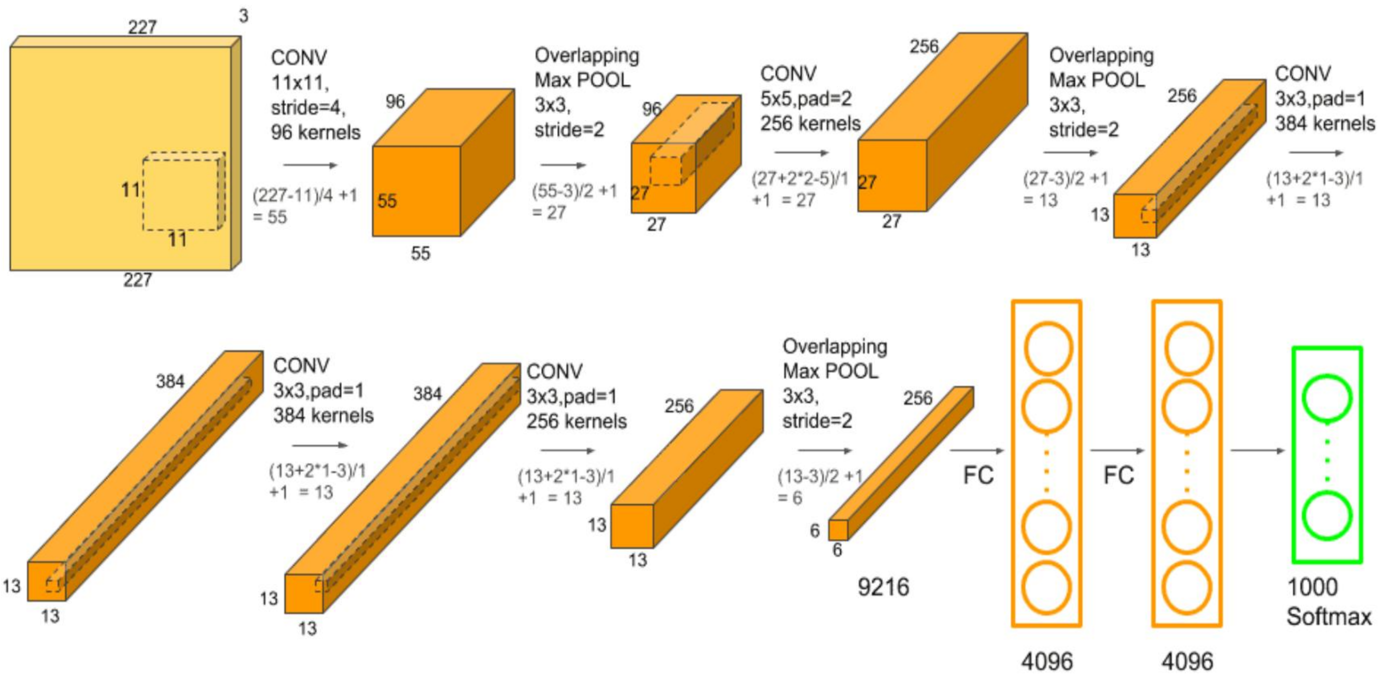

Fig. 1 AlexNet Architecture [63]

C. VGGNet

VGGNet, created by the University of Oxford, managed to stand out at ILSVRC-2014, achieving first place in the localisation task and second place in the classification task [55]. This neural network is composed of five main blocks, each consisting of one or more convolutional layers followed by a clustering layer [56].

The VGGNet architecture includes a total of sixteen convolutional layers and three fully connected layers, interspersed with clustering layers. One of its key features is the use of small 3X3 convolutional kernels, which, when expanded, allow the depth of the network to be increased and more detailed image features to be captured. This architecture has a simple design and is easy to implement in different applications [57]. Although it may seem simple, this model marked a significant advance in the field of image classification and continues to be a reference in the development of deep learning models [58].

D. ResNet

The CNN ResNet is based on an innovative idea, which stacks residual blocks on top of each other to form a network. Its design follows two key principles: first, all layers retain the same number of filters regardless of their feature map size, and second, when the feature map size is halved, the number of filters is doubled to maintain a balanced computational load [59], [60]. ResNet to further optimize its performance, uses a design called bottleneck. This approach helps to reduce the number of parameters in the model, which speeds up training without compromising the quality of learning [61]. One of the main advantages of this CNN is that it allows us to build much deeper networks, allowing us to solve the degradation problem that is used to limit traditional neural networks [62].

E. DataSet

To predict driver behaviour, a dataset of 10751 images, extracted from the Kaggle platform, was used. These images were grouped into five categories representing different actions performed by drivers. Safe driving, this represents the ideal scenario, where the driver keeps full attention on the road and drives safely, respecting the rules and in control of the vehicle. Other activities, this category includes a variety of behaviours that do not fit into the other defined classes. It can range from actions such as eating or drinking, to interacting with passengers, which could distract the driver. Talking_on_the_phone: This category captures moments when the driver uses a mobile device to make phone calls, a common distraction that can significantly affect concentration. Texting_phone: focuses specifically on the action of texting while driving, a highly dangerous practice requiring hand-eye coordination that distracts the driver from the main task. Turning: This category focuses on turning manoeuvres, which require special attention to direction indicators and the road environment.

Each category represents a specific action performed by the driver, allowing the model to learn to distinguish patterns of behavior that may influence driving safety. These categories were selected based on their relevance for identifying common driving distractions, such as using a phone or performing other tasks while driving. Fig. 2 provides a visual representation of the dataset and the case study development process. The images of the dataset show examples of the different categories, making it easier to understand the variability and complexity of the data. It also shows the class stages of the process from exploration and analysis of the data to training the models and obtaining the final predictions.

Within each category, there is likely to be a wide variety of images due to factors such as lighting conditions, camera angles, driver clothing and other elements of the environment. Also, the number of images in each category may not be the same, which could lead to an imbalance of classes. This could affect the performance of the models, as it could favor the prediction of the most represented classes. The model development and training process is presented in Fig. 2.

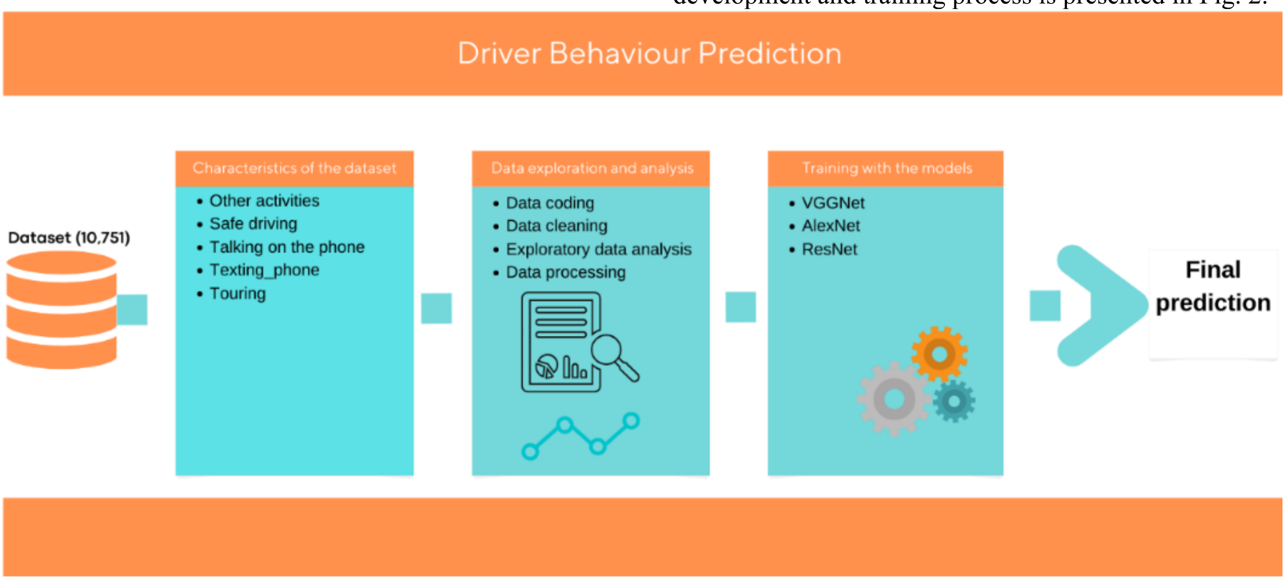

Fig. 2 Case development process

1. Data Cleaning and Processing: In this phase, the organization and cleaning of the images in the dataset was carried out. First, the images were classified into five classes: "other activities", "safe driving", "talking on the phone", "texting" and "turning". Subsequently, images not readable by the system, which could affect the quality of the model, were removed. Next, random samples were generated for each class, allowing an initial visualization of the dataset. Finally, dataframes were created from classified images, assigning labels corresponding to each class to facilitate their use in subsequent stages of the model. It was also determined that the images needed to be resized (to 240x240 pixels) to adapt them to the input model. For this, image generators were implemented to automate the re-scaling and extraction of images from the dataframes, preparing them for model training. Also, file paths were mapped to labels. One-hot coding of categories was performed. Custom generators were then developed, such as real-time data incrementation, random rotations, horizontal and vertical displacements. The memory was optimized to allow batch loading and cache management. Next, pipeline configuration of real-time processing, on-the-fly normalization and efficient buffer management were performed. Next, validation and final quality control were performed, including load testing, output format validation, label consistency checking, buffer parameter adjustments, load speed optimization, resource usage monitoring.

2. Exploratory Data Analysis: The exploration data analysis was conducted comprehensively, examining multiple dimensions and characteristics of the dataset to ensure a thorough understanding of its structure and quality. The Safe Driving class represented 20.49% of the total samples, basically consisting of: correct driver position, and attention to the road. The Texting Phone class represents 20.49% of the total samples, with predominant characteristics such as: use of mobile phone, visual distraction. The Talking Phone class, representing 20.17% of samples, is predominantly characterized by: talking on the phone, divided attention. The Other activities class, represented by 19.13% of the total sample, is characterized by images of various distractions, varied behaviors. Finally, the Turing class, comprising 19.13% of the total dataset, is characterized by turning the steering wheel, changing direction. The technical dimensions of the images were defined as follows: original resolution - variable (800x600 to 1920x1080). Standardized resolution: 240x240 pixels. Aspect ratio: 1:1 after standardization. Colour model: RGB. Depth: 24 bits. Value range: 0-25 per channel. Regarding sharpness metrics, average gradient value: 42.3. Average sharpness index: 0.87. RMS contrast: 0.45. Other factors considered were environmental conditions, such as: vehicle interior visible: 100%. Window reflections: 15%. Variable shadows:30%. Fig. 3 shows images of the data set.

Fig. 3 images of the dataset

The training process and model validation were key steps in our work on driver behaviour detection. To carry out this process, a structured and detailed approach was used that included three CNN architectures: AlexNet, VGGNet and ResNet. These architectures were selected for their unique characteristics and their ability to address this type of classification task.

This process began with stratified partitioning of the dataset, allocating 75% for training, 15% for validation and 15% for testing. This technique ensured that each class of behaviour was proportionally represented in the three subsets, ensuring a fair and balanced assessment of model performance. During the training process, a robust pre-processing flow was implemented that included image normalization and real-time data augmentation techniques. Images were normalized to have values in the range of 0 to 1, with a mean of 0.5 and a standard deviation of 0.5. Also, data enhancements incorporating +-15-degree rotations were applied, which helped to increase the variability in the training samples. These techniques are essential to improve the models' ability to generalize and handle different conditions in the images.

The hyperparameters were configured specifically for each architecture, considering its own characteristics. The Adam optimizer with a learning rate of 0.001 was used for AlexNet. For VGGNet a rate of 0.01, and a rate of 0.0001 for ResNet. The batch size was set to 64 samples for all models, seeking a balance between computational efficiency and training stability. During the training process, regulation strategies were implemented that included Dropout in the fully connected layers with a rate of 0.5, and a batch normalization after each convolutional layer. These techniques were necessary to prevent overfitting and improve the generalizability of the models. Training monitoring was performed using a comprehensive system that records multiple metrics in real time, including loss and accuracy in both training and validation. Early stopping was implemented with 20 epochs to optimize training time and prevent overfitting. Fig. 4 shows a comparison of the performance of the three CNN models during training over the 20 epochs. The X-axis represents the number of epochs, and the Y-axis represents the loss.

Analyzing the blue line representing the AlexNet Model. This model starts with a relatively low loss of 0.42. It shows a rapid increase up to Epoch 2 with 0.81. Along the way it shows some fluctuations, especially in epoch 5. It stabilizes after Epoch 7 at values close to 0.95. From then on it maintained a constant performance in the epoch. The orange, black line represents the training of the VGGNet model, this model starts with the lowest loss at 0.32. It shows a gradual increase until epoch 4, then it experiences more variability than AlexNet. Also, several peaks can be seen, especially in epochs 9 and 11, from epoch 12 onwards it shows similar values to AlexNet. The black line represents the performance of the ResNet model. This model starts with the lowest loss, maintains the lowest values of loss, shows remarkable stability throughout the training, stays in a range between 0.30 and 0.37, and does not present significant improvements after Epoch 3.

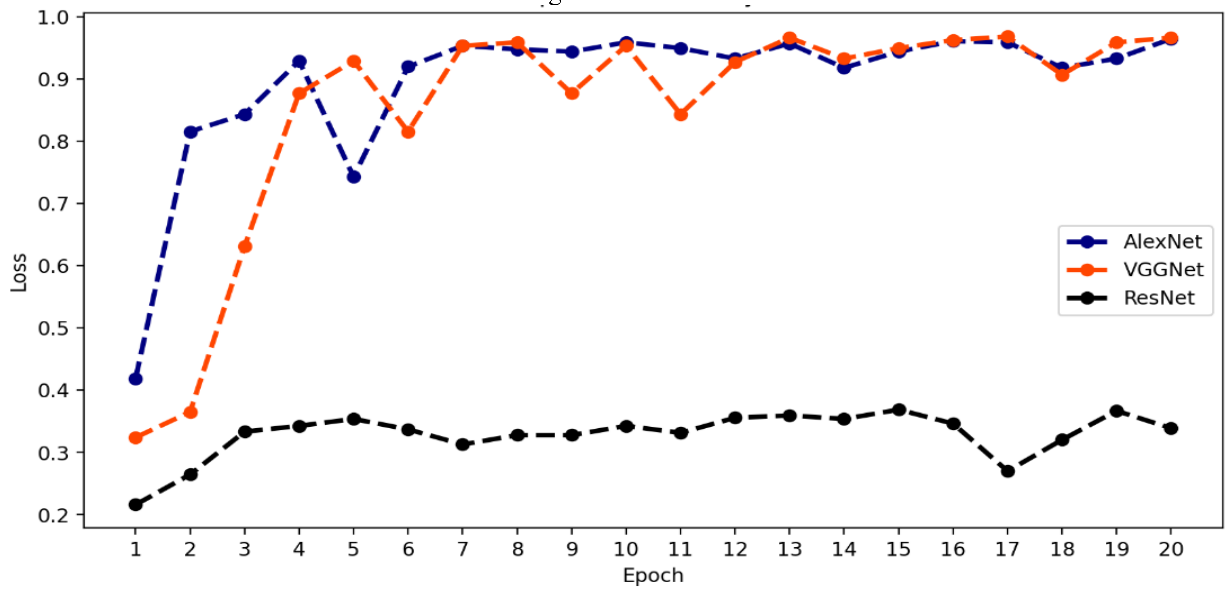

Fig. 4 Epoch numbers and Accuracy of all models

Finally, Fig. 4 shows that the ResNet model maintains more stable behavior with less loss during training, as evidenced by the accuracy results in Table 1.

IV. RESULTS

In this section, results are presented as a product of training, validation and testing of the AlexNet, VGGNet and ResNet models for driver prediction using the 10751 images and five classes dataset. The results obtained were significantly different between the three architectures. VGGNet demonstrated superior performance, reaching an accuracy of 98.3% after epoch 20, with an approximate training time of 4.5 hours. AlexNet achieved an accuracy of 98% with a recall of 75% and an F1 score of 86% after Epoch 20. On the other hand, the performance of the ResNet Model was lower, reaching an accuracy of 35.28% after the second epoch.

Table 1 Validation and testing results

Model	Accuracy	Recall	F1-Score
VGGNet	98.3%	75%	86%
AlexNet	98%	75%	86%
ResNet	35.28	30%	32%

After validation and testing, the results obtained showed that the VGGNet model showed remarkable stability in recognizing driver behavior patterns during training. It also demonstrated an outstanding adaptive capability, specifically in efficiently classifying driver behaviors. For example, talking on the phone and texting while driving. The AlexNet model, on the other hand, demonstrated a good balance between computational efficiency and accuracy in this task. However, the ResNet model failed to efficiently classify the necessary patterns, performing significantly worse.

A detailed analysis of the classification errors allowed us to identify key areas for improving model performance, such as handling extreme lighting conditions or unusual driver postures. This information is fundamental to refine the models and optimize data augmentation strategies, thus strengthening their robustness. Cross-validation with K=5 confirmed the robustness of the results, showing minimal variation in performance between different partitions of the dataset. These findings show that CNN models, especially VGGNet, are efficient in recognizing new patterns, even in complicated lighting and other situations.

These results are significant, considering the inherent complexity of the task of classifying driver behaviors and variations in imaging conditions. The success of VGGNet shows us that its deep but well-organized architecture is particularly well suited to identifying subtle features that differentiate driver actions.

V. DISCUSSION

This section presents the conclusions of our research, contrasting them with the findings obtained in the studies reviewed during the literature review stage. This analysis allowed us to identify significant and relevant contributions in both technical and practical terms, highlighting the potential impact of these findings on improving road safety.

In this research it was shown that the VGGNet architecture was the best overachiever in the classification of driver behaviours, achieving a remarkable accuracy of 98.3%. This performance indicates that the balanced VGGNet architecture is effective in identifying key hierarchical features and patterns associated with distracted driving. Its deep learning design allows for more detailed and accurate feature extraction, which contributes to its effectiveness in this specific type of task. The stability demonstrated during training and the robustness to variations in input data reinforce its potential for applications in real-world environments. The study in [26], used EEGNet to predict emergency braking with an accuracy of over 80%, which, while a lower value, focuses on a more specific task, such as anticipation of braking. Similarly, the study in [27], which combined LSTM with CNN to classify defensive, normal and aggressive driving, achieved an accuracy of 98%, very close to that obtained with VGGNet, suggesting that this hybrid approach is equally effective in classifying driving behaviors. These studies confirm that CNN architectures are highly effective in predicting driver behavior. Which differs with [24] that employs a multiple granularity strategy to detect driver fatigue achieved an accuracy of 89.42%, suggesting that VGGNet is more efficient in predicting overall driver behavior.

AlexNet, on the other hand, demonstrated high efficiency with 98% accuracy, suggesting that simpler architectures can be very efficient for these specific tasks. This finding has significant implications for practical implementation, especially in systems with limited computational resources or time requirements. However, the study by [25], which combined advanced models such as ResNet50, DenseNet201 and InceptionV3 together with XGBoost, reported a higher average accuracy, reaching 99.52%, demonstrating that combinations of models can outperform individual networks such as VGGNet.

ResNet's low performance of 35.28% raises interesting questions about the relationship between architecture complexity and efficiency in specific classification tasks. This result contradicts the general perception that deeper architectures necessarily lead to better performance, suggesting that the nature of the driver behavior classification problem may not require the additional complexity offered by the ResNet model.

The optimization strategies that were implemented, particularly real-time data augmentation and batch normalization, proved to be very important for training success. The effectiveness of these techniques indicates that controlled variability in training data and numerical stability are critical factors in the development of driver monitoring systems. From a practical perspective, our results are especially encouraging. The relatively short training time and moderate computational requirements indicate the feasibility of implementation in real systems. The ability of the models to discriminate between different types of distractions with high accuracy opens significant possibilities for the development of more effective accident prevention systems. However, it is important to recognize the limitations and challenges that exist. The controlled conditions under which the data were collected may not fully reflect the variability present in real driving situations. In addition, practical implementation will need to address important privacy considerations and ethical aspects of continuous driver monitoring.

Overall, a review of the literature review indicates that our approach represents a significant advance, both in terms of accuracy and computational efficiency. Nevertheless, we identify several promising areas for future improvements, including the exploration of hybrid architecture, the incorporation of temporal information, and the development of methods for adaptation to different environmental conditions.

VI. CONCLUSIONS

After completing the training process and comparing AlexNet, VGGNet and ResNet models for kidney stone detection, the following conclusions were drawn.

The findings show that VGGNet excelled in terms of accuracy and performance in predicting driver behavior, achieving an accuracy of 98.30%. This outperformed AlexNet, which achieved an accuracy of 98%, with a recall of 75% and an F1 score of 86%. On the other hand, ResNet showed significantly lower performance, with an accuracy of 35.28%, a recall of 30% and an F1 score of 32%. In comparison, VGGNet ranks as the top choice for predicting driver's behavior. Accurate driver behavior prediction is crucial to improving road safety by identifying dangerous behavior and potentially avoiding accidents, which can increase road safety and reduce costs associated with vehicle incidents.

Looking ahead to improving prediction of driver behavior, it is recommended to continue using VGGNet, as it has proven to be the most accurate and reliable model. It would also be useful to explore combinations with other models that can help to further improve performance, especially in contexts where greater generalization is needed. In addition, the implementation of these models in real-time monitoring systems could help to identify dangerous driver behaviors quickly, allowing preventive measures to be taken before an accident occurs. Finally, it is important to test the model with more varied data and in different driving conditions to ensure its effectiveness and applicability in real-world situations.

Finally, this work not only validates the technical feasibility of automated detection of distracted driving behavior but also establishes a solid basis for the development of safer and more efficient systems in the future of mobility. The results obtained, while promising, also indicate the need for continued research and development to address remaining challenges and maximize the potential impact on road safety.

REFERENCES

[1] "Road traffic injuries." Accessed: Aug. 15, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

[2] S. A. M. Abd El Fatah, "Multifarious experiences of women drivers post-car accidents: An urban community experience," *Womens Stud Int Forum*, vol. 85, p. 102449, Mar. 2021, doi: 10.1016/J.WSIF.2021.102449.

[3] P. Laxmanrao Sarode and S. Damodar Suryawanshi, "Analysis of vehicle accidents in a parameter of quantitative study and evaluation of safety parameters," *Mater Today Proc*, vol. 72, pp. 802–809, Jan. 2023, doi: 10.1016/J.MATPR.2022.09.052.

[4] M. Eltemasi and H. Behtooiey, "Predicting the type of road accidents based on air temperature in Iran: A case study of roads in Qazvin province," *Heliyon*, vol. 10, no. 7, p. e28536, Apr. 2024, doi: 10.1016/J.HELIYON.2024.E28536.

[5] T. Y. Kim, J. H. Kim, B. Goo, and B. K. Seo, "Attitudes and experiences of traditional Korean medicine practitioners in cases of traffic accidents: A qualitative study," *Prev Med Rep*, vol. 44, p. 102767, Aug. 2024, doi: 10.1016/J.PMEDR.2024.102767.

[6] M. Alhomoud, E. AlSaleh, and B. Alzaher, "Car accidents and risky driving behaviors among young drivers from the Eastern Province, Saudi Arabia," *Traffic Inj Prev*, vol. 23, no. 8, pp. 471–477, Jan. 2022, doi: 10.1080/15389588.2022.2113392.

[7] K. Aati, M. Houda, S. Alotaibi, A. M. Khan, N. Alselami, and O. Benjeddou, "Analysis of Road Traffic Accidents in Dense Cities: Geotech Transport and ArcGIS," *Transportation Engineering*, vol. 16, p. 100256, Jun. 2024, doi: 10.1016/J.TRENG.2024.100256.

[8] D. Bruski, L. Pachocki, A. Sciegaj, and W. Witkowski, "Speed estimation of a car at impact with a W-beam guardrail using numerical simulations and machine learning," *Advances in Engineering Software*, vol. 184, p. 103502, Oct. 2023, doi: 10.1016/J.ADVENGSOFT.2023.103502.

[9] T. Kenarangi, F. Rahmani, A. Yazdani, G. D. Ahmadi, M. Lotfi, and T. A. Khalaj, "Comparison of GAP, R-GAP, and new trauma score (NTS) systems in predicting mortality of traffic accidents that injure hospitals at Mashhad University of medical sciences," *Heliyon*, vol. 10, no. 16, p. e36004, Aug. 2024, doi: 10.1016/J.HELIYON.2024.E36004.

[10] J. Wang, Z. Li, F. ying, D. Zou, and Y. Chen, "Reconstruction of a real-world car-to-pedestrian collision using geomatics techniques and numerical simulations," *J Forensic Leg Med*, vol. 91, p. 102433, Oct. 2022, doi: 10.1016/J.JFLM.2022.102433.

[11] I. Azouz *et al.*, "Traumatic uterine rupture: A rare complication of vehicle accidents," *Radiol Case Rep*, vol. 19, no. 5, pp. 1994–1997, May 2024, doi: 10.1016/J.RADCR.2024.02.020.

[12] I. Konkor, "Examining the relationship between transportation mode and the experience of road traffic accident in the upper west region of Ghana," *Case Stud Transp Policy*, vol. 9, no. 2, pp. 715–722, Jun. 2021, doi: 10.1016/J.CSTP.2021.03.009.

[13] L. Brühwiler, C. Fu, H. Huang, L. Longhi, and R. Weibel, "Predicting individuals' car accident risk by trajectory, driving events, and geographical context," *Comput Environ Urban Syst*, vol. 93, p. 101760, Apr. 2022, doi: 10.1016/J.COMPENVURBSYS.2022.101760.

[14] L. Yang, C. Zhao, C. Lu, L. Wei, and J. Gong, "Lateral and Longitudinal Driving Behavior Prediction Based on Improved Deep Belief Network," *Sensors (Basel)*, vol. 21, no. 24, Dec. 2021, doi: 10.3390/S21248498.

[15] "Assessing the synergistic effects of artificial intelligence on pollutant and carbon emission mitigation in China," *Energy Econ*, vol. 138, p. 107829, Oct. 2024, doi: 10.1016/J.ENECO.2024.107829.

[16] A. M. Geske, D. M. Herold, and S. Kummer, "Artificial intelligence as a driver of efficiency in air passenger transport: A systematic literature review and future research avenues," *Journal of the Air Transport Research Society*, vol. 3, p. 100030, Dec. 2024, doi: 10.1016/J.JATRS.2024.100030.

[17] B. Yang, "Research on Vehicle Detection and Recognition Technology Based on Artificial Intelligence," *Microprocess Microsyst*, p. 104937, Sep. 2023, doi: 10.1016/J.MICPRO.2023.104937.

[18] K. Rana and N. Khatri, "Automotive intelligence: Unleashing the potential of AI beyond advance driver assisting system, a comprehensive review," *Computers and Electrical Engineering*, vol. 117, p. 109237, Jul. 2024, doi: 10.1016/J.COMPELECENG.2024.109237.

[19] J. Gu, Z. Wu, Y. Song, and A. C. Nicolescu, "A win-win relationship? New evidence on artificial intelligence and new energy vehicles," *Energy Econ*, vol. 134, p. 107613, Jun. 2024, doi: 10.1016/J.ENECO.2024.107613.

[20] K. Inouye, A. Petrosyan, L. Moskalensky, and F. G. Thankam, "Artificial intelligence in therapeutic management of hyperlipidemic ocular pathology," *Exp Eye Res*, vol. 245, p. 109954, Aug. 2024, doi: 10.1016/J.EXER.2024.109954.

[21] F. Bao, F. Gao, and W. Li, "Application of IoT voice devices based on artificial intelligence data mining in motion training feature recognition," *Measurement: Sensors*, vol. 34, p. 101260, Aug. 2024, doi: 10.1016/J.MEASEN.2024.101260.

[22] H. Li, H. Xu, Y. Li, and X. Li, "Application of artificial intelligence (AI)-enhanced biochemical sensing in molecular diagnosis and imaging analysis: Advancing and challenges," *TrAC Trends in Analytical Chemistry*, vol. 174, p. 117700, May 2024, doi: 10.1016/J.TRAC.2024.117700.

[23] N. Chakrabarty and A. Mahajan, "Imaging Analytics using Artificial Intelligence in Oncology: A Comprehensive Review," *Clin Oncol*, vol. 36, no. 8, pp. 498–513, Aug. 2024, doi: 10.1016/J.CLON.2023.09.013.

[24] R. Huang, Y. Wang, Z. Li, Z. Lei, and Y. Xu, "RF-DCM: Multi-Granularity Deep Convolutional Model Based on Feature Recalibration and Fusion for Driver Fatigue Detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 630–640, Jan. 2022, doi: 10.1109/TITS.2020.3017513.

[25] T. J. Chengula, J. Mwakalonge, G. Comert, M. Sulle, S. Siuhi, and E. Osei, "Enhancing advanced driver assistance systems through explainable artificial intelligence for driver anomaly detection," *Machine Learning with Applications*, vol. 17, p. 100580, Sep. 2024, doi: 10.1016/J.MLWA.2024.100580.

[26] X. Liang, Y. Liu, Y. Yu, and Z. Zhou, "Using Data Augmentation to Improve EEG-Based Prediction of Emergency Braking Intention," *Proceedings - 2022 Chinese Automation Congress, CAC 2022*, vol. 2022-January, pp. 5015–5019, 2022, doi: 10.1109/CAC57257.2022.10055049.

[27] S. K. Kwon, J. H. Seo, J. Y. Yun, and K. D. Kim, "Driving Behavior Classification and Sharing System Using CNN-LSTM Approaches and V2X Communication," *Applied Sciences 2021, Vol. 11, Page 10420*, vol. 11, no. 21, p. 10420, Nov. 2021, doi: 10.3390/APP112110420.

[28] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural

networks," *Expert Syst Appl*, vol. 149, Jul. 2020, doi: 10.1016/J.ESWA.2020.113240.
[29] M. Lu, Y. Hu, and X. Lu, "Dilated Light-Head R-CNN using tri-center loss for driving behavior recognition," *Image Vis Comput*, vol. 90, p. 103800, Oct. 2019, doi: 10.1016/J.IMAVIS.2019.08.004.
[30] L. Wang, W. Yao, C. Chen, and H. Yang, "Driving Behavior Recognition Algorithm Combining Attention Mechanism and Lightweight Network," *Entropy 2022, Vol. 24, Page 984*, vol. 24, no. 7, p. 984, Jul. 2022, doi: 10.3390/E24070984.
[31] "(PDF) A CNN-LSTM APPROACH FOR ACCURATE DROWSINESS AND DISTRACTION DETECTION IN DRIVERS." Accessed: Aug. 26, 2024. [Online]. Available: https://www.researchgate.net/publication/383227527_A_CNN-LSTM_APPROACH_FOR_ACCURATE_DROWSINESS_AND_DISTRACTION_DETECTION_IN_DRIVERS
[32] S. Moosavi, P. D. Mahajan, S. Parthasarathy, C. Saunders-Chukwu, and R. Ramnath, "Driving Style Representation in Convolutional Recurrent Neural Network Model of Driver Identification," Feb. 2021, Accessed: Jan. 09, 2025. [Online]. Available: https://arxiv.org/abs/2102.05843v1
[33] A. A. Sheikh and I. Z. Khan, "Enhancing Road Safety: Real-Time Detection of Driver Distraction through Convolutional Neural Networks," May 2024, Accessed: Jan. 09, 2025. [Online]. Available: https://arxiv.org/abs/2405.17788v1
[34] H. M. Gheni and L. A. Abdul-Rahaim, "An Efficient Deep Learning Model Based on Driver Behaviour Detection Within CAN-BUS Signals," *Revue d'Intelligence Artificielle*, vol. 38, no. 1, pp. 53–62, Feb. 2024, doi: 10.18280/RIA.380106.
[35] G. Lucente, M. S. Maarssoe, I. Kahl, and J. Schindler, "Deep Learning Algorithms for Longitudinal Driving Behavior Prediction: A Comparative Analysis of Convolutional Neural Network and Long–Short-Term Memory Models," *SAE International Journal of Connected and Automated Vehicles*, vol. 7, no. 4, pp. 389–404, Jun. 2024, doi: 10.4271/12-07-04-0025.
[36] C. Zhang, R. Li, W. Kim, D. Yoon, and P. Patras, "Driver Behavior Recognition via Interwoven Deep Convolutional Neural Nets with Multi-stream Inputs," *IEEE Access*, vol. 8, pp. 191138–191151, Nov. 2018, doi: 10.1109/ACCESS.2020.3032344.
[37] C. Gou, Y. Zhou, and D. Li, "Driver attention prediction based on convolution and transformers," *Journal of Supercomputing*, vol. 78, no. 6, pp. 8268–8284, Apr. 2022, doi: 10.1007/S11227-021-04151-2/METRICS.
[38] B. Ma *et al.*, "Distracted Driving Behavior and Driver's Emotion Detection Based on Improved YOLOv8 With Attention Mechanism," *IEEE Access*, vol. 12, pp. 37983–37994, 2024, doi: 10.1109/ACCESS.2024.3374726.
[39] Y. Zhang, J. Zhang, and W. Zhou, "Research on Image Classification Improvement Based on Convolutional Neural Networks with Mixed Training," *Proceedings of 2022 IEEE 4th International Conference on Civil Aviation Safety and Information Technology, ICCASIT 2022*, pp. 7–10, 2022, doi: 10.1109/ICCASIT55263.2022.9986643.
[40] Z. Ajra, B. Xu, G. Dray, J. Montmain, and S. Perrey, "Mental arithmetic task classification with convolutional neural network based on spectral-temporal features from EEG," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2022-July, pp. 52–55, 2022, doi: 10.1109/EMBC48229.2022.9870887.
[41] D. M. Gonzalez-Arriaga *et al.*, "Design of a Software Platform to Generate Convolutional Neural Networks for the Parametric Identification of a Cartesian Robot," *IEEE Access*, vol. 11, pp. 63371–63387, 2023, doi: 10.1109/ACCESS.2023.3289078.
[42] O. Iparraguirre-Villanueva *et al.*, "Classification of Tweets Related to Natural Disasters Using Machine Learning Algorithms," *International Journal of Interactive Mobile Technologies*, vol. 17, no. 14, pp. 144–162, Jan. 2023, doi: 10.3991/ijim.v17i14.39907.
[43] Y. Wang, Y. Yu, T. Zhang, K. Song, Y. Wang, and S. Gao, "Improved dendritic learning: Activation function analysis," *Inf Sci (N Y)*, vol. 679, Sep. 2024, doi: 10.1016/J.INS.2024.121034.
[44] L. Nahhas, M. Albahar, A. Alammari, and A. Jurcut, "Android Malware Detection Using ResNet-50 Stacking," *Computers, Materials & Continua*, vol. 74, no. 2, pp. 3997–4014, Oct. 2022, doi: 10.32604/CMC.2023.028316.
[45] O. Dahmane, M. Khelifi, M. Beladgham, and I. Kadri, "Pneumonia detection based on transfer learning and a combination of VGG19 and a CNN built from scratch," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 3, pp. 1469–1480, Dec. 2021, doi: 10.11591/IJEECS.V24.I3.PP1469-1480.
[46] O. Iparraguirre-Villanueva, J. Gonzales-Huaman, J. Machuca-Solano, and J. Ruiz-Alvarado, "Improving industrial security device detection with convolutional neural networks," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, no. 3, pp. 1935–1943, Jun. 2024, doi: 10.11591/ijeecs.v34.i3.pp1935-1943.
[47] M. Hussain, C. Cheng, R. Xu, and M. Afzal, "CNN-Fusion: An effective and lightweight phishing detection method based on multi-variant ConvNet," *Inf Sci (N Y)*, vol. 631, pp. 328–345, Jun. 2023, doi: 10.1016/J.INS.2023.02.039.
[48] D. Kojić and M. Yasui, "Why matrix-decomposition of overlapped light spectroscopic features leads to overfitting in certain cases ?," *Chemometrics and Intelligent Laboratory Systems*, vol. 239, p. 104859, Aug. 2023, doi: 10.1016/J.CHEMOLAB.2023.104859.
[49] F. Martínez, F. José, C. Bogotá, D. C. Colombia, H. Montiel, and C. Fernando Martínez, "A Machine Learning Model for the Diagnosis of Coffee Diseases," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, pp. 968–974, Summer 2022, doi: 10.14569/IJACSA.2022.01304110.
[50] O. Iparraguirre-Villanueva *et al.*, "Text prediction recurrent neural networks using long short-term memory-dropout," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 3, pp. 1758–1768, Mar. 2023, doi: 10.11591/ijeecs.v29.i3.pp1758-1768.
[51] H. Amanollah, A. Asghari, M. Mashayekhi, and S. M. Zahrai, "Damage detection of structures based on wavelet analysis using improved AlexNet," *Structures*, vol. 56, p. 105019, Oct. 2023, doi: 10.1016/J.ISTRUC.2023.105019.
[52] H. Eldem, E. Ülker, and O. Y. Işıklı, "Alexnet architecture variations with transfer learning for classification of wound images," *Engineering Science and Technology, an International Journal*, vol. 45, p. 101490, Sep. 2023, doi: 10.1016/J.JESTCH.2023.101490.
[53] D. Fan, H. Liang, C. Qu, J. Ma, and R. Hasani, "A novel approach for early gastric cancer detection using a hybrid of alexnet, extreme learning machine, and adjusted gorilla troops optimization," *Biomed Signal Process Control*, vol. 93, p. 106126, Jul. 2024, doi: 10.1016/J.BSPC.2024.106126.
[54] X. Luo, W. Wen, J. Wang, S. Xu, Y. Gao, and J. Huang, "Health classification of Meibomian gland images using keratography 5M based on AlexNet model," *Comput Methods Programs Biomed*, vol. 219, p. 106742, Jun. 2022, doi: 10.1016/J.CMPB.2022.106742.
[55] W. Fan, J. Fang, and G. Huang, "An improved image texture synthesis based on algorithm convolution neural network," *Physical Communication*, vol. 66, p. 102395, Oct. 2024, doi: 10.1016/J.PHYCOM.2024.102395.
[56] E. Prasetyo, N. Suciati, and C. Fatichah, "Multi-level residual network VGGNet for fish species classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 5286–5295, Sep. 2022, doi: 10.1016/J.JKSUCI.2021.05.015.
[57] Y. Yan, Y. Yang, S. Fang, M. Gao, and Y. Chen, "MUS Model: A Deep Learning-Based Architecture for IoT Intrusion Detection," *Computers, Materials and Continua*, vol. 80, no. 1, pp. 875–896, Jul. 2024, doi: 10.32604/CMC.2024.051685.
[58] I. Khan, S. S. Sohail, D. Ø. Madsen, and B. K. Khare, "Deep transfer learning for fine-grained maize leaf disease classification," *J Agric Food Res*, vol. 16, p. 101148, Jun. 2024, doi: 10.1016/J.JAFR.2024.101148.
[59] J. Li, R. Gong, and G. Wang, "Enhancing fitness action recognition with ResNet-TransFit: Integrating IoT and deep learning techniques for real-time monitoring," *Alexandria Engineering Journal*, vol. 109, pp. 89–101, Dec. 2024, doi: 10.1016/J.AEJ.2024.07.068.
[60] F. Ben Nasr Barber and A. Elloumi Oueslati, "Human exons and introns classification using pre-trained Resnet-50 and GoogleNet models and 13-layers CNN model," *Journal of Genetic Engineering and Biotechnology*, vol. 22, no. 1, p. 100359, Mar. 2024, doi: 10.1016/J.JGEB.2024.100359.
[61] W. Al-Khater and S. Al-Madeed, "Using 3D-VGG-16 and 3D-Resnet-18 deep learning models and FABEMD techniques in the detection of

malware," *Alexandria Engineering Journal*, vol. 89, pp. 39–52, Feb. 2024, doi: 10.1016/J.AEJ.2023.12.061.

[62] S. Yu, Z. Zhang, S. Wang, X. Huang, and Q. Lei, "A performance-based hybrid deep learning model for predicting TBM advance rate using Attention-ResNet-LSTM," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 16, no. 1, pp. 65–80, Jan. 2024, doi: 10.1016/J.JRMGE.2023.06.010.

[63] "Understanding AlexNet | LearnOpenCV #." Accessed: Nov. 13, 2024. [Online]. Available: https://learnopencv.com/understanding-alexnet/