



Intelligent Vehicle Driver Face and Conscious Recognition

Hiba Ali Ahmed^{1*}, Muayad Sadik Croock¹, Mohammed A. Noaman Al-hayanni²

¹ Department of Control and System Engineering, University of Technology-Iraq, Baghdad 10066, Iraq

² Department of Electrical Engineering, University of Technology-Iraq, Baghdad 10066, Iraq

Corresponding Author Email: cse.21.15@grad.uotechnology.edu.iq

Copyright: ©2023 IETA. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ria.370612>

ABSTRACT

Received: 21 August 2023

Revised: 12 October 2023

Accepted: 21 November 2023

Available online: 27 December 2023

Keywords:

classification, CNN, dataset, emotional recognition, person face recognition

The car manufacturing industry faces pressing issues of vehicle theft and driver conscious related accidents. This study introduces AI-powered computer applications to tackle these challenges, aiming to enhance security and safety in the automotive sector. The study developing two distinct models—one for driver identification via facial recognition prior to ignition, and another for continuous driver state monitoring during travel—this research aims to bolster vehicle security and enhance driver safety. Two carefully curated data sets consisting of images of four individuals were used to train and validate the models, one for facial recognition and the other for conscious and unconscious driver detection. The models achieved accuracy rates exceeding 99%, and cross-validation confirmed their reliability, with consistent performance showing accuracy ranging from 95% to 100%. The study underscores the potential of AI to revolutionize vehicle security and driver safety mechanisms. The implementation of these models promises to significantly curtail the incidence of car theft and the risk of accidents cause by driver un conscious, heralding a new era of ethical and advanced automotive technologies.

1. INTRODUCTION

Due to rapid advancements in computational capabilities and the availability of modern sensing, processing, and visualization tools, computers are gaining higher levels of intelligence. They can now seamlessly engage with humans by utilizing cameras and microphones to perceive and understand people's actions, and subsequently respond in a personable manner. This capability has been showcased in various research endeavors and real-world business applications [1]. One of the primary techniques facilitating this seamless interaction between humans and computers is face detection [2, 3]. Face detection serves as the fundamental building block for a wide range of facial analysis algorithms, encompassing tasks like aligning faces, creating facial models, adjusting lighting conditions on faces, recognizing faces, authenticating identities, tracking head poses, identifying and tracking facial emotions, as well as determining gender and age [4]. Additionally, a number of computer vision-based methods have been developed for the non-intrusive, real-time identification of driver sleep stages using a variety of visual cues and observable face features [5, 6]. It is well recognized that a person's level of alertness and weariness can be inferred from an observable pattern of eye, head, and facial expression movement. Indicators of a person who is extremely exhausted and drowsy include eye closure, head movement, a drooping jaw, an asymmetrical brow, and eyelid movement. A remote camera is typically put on the dashboard of the car to use these visual signals. It analyzes the driver's physical conditions and determines whether or not they are drowsy by using a variety

of facial traits that have been extracted from the driver [7, 8].

Driver identification and conscious detection are two critical challenges in the automotive industry. Vehicle theft and the conscious of driver are major safety concerns, and existing solutions are often limited in their effectiveness.

This study aims to address these challenges by developing AI-powered computer applications for driver identification and detect the driver conscious. These applications are designed to improve vehicle security and safety by preventing unauthorized vehicle use through facial recognition-based driver identification and reducing the risk induced accidents cause by driver un conscious by detecting un conscious drivers and alerting them to the need to take a break.

The proposed system consists of two deep learning models, one to recognize the face of the authorized driver and the other to monitor the state of conscious of the authorized driver to drive the car. The two models are trained and validated using two carefully curated data sets of images of four individuals. Accuracy and loss ratios are used to calculate efficiency. For the training and validation sets, the accuracy of the recommended model for face recognition is 99.63% and 99.48%, respectively. While the proposed model of conscious performs well, training and validation scores of 99.54% and 99.37%, respectively, were determined.

2. RELATED WORK

The study of this field has been considered by numerous researchers. In this section, we divided the related studies into

two parts for the sake of simplicity. The recognition works for people's faces as well as emotional face gestures.

2.1 Person face recognition

The rapid growth of computer science and information technology has resulted in an increase in the use of Artificial Intelligence (AI) in a wide range of applications. Deep Neural Networks (DNN) have proven effective in tasks such as speech recognition, image identification, and character recognition [9-13]. In particular, DNN-based facial recognition has received considerable attention in recent years. Several studies have investigated the use of DNNs for facial recognition, including the development of a novel linked mapping technique based on Convolutional Neural Networks (CNN) to recognize low-resolution facial images [11-14], and the use of standard face databases to demonstrate that DNNs can successfully extract facial characteristics using a range of image processing approaches [15]. To improve the accuracy of facial recognition, various techniques have been integrated, including Principal Component Analysis (PCA) and Support Vector Machine (SVM) [16, 17]. For instance, one study utilized CNN to extract facial features before applying PCA to condense the size of the extracted features, followed by a combined Bayesian method for facial recognition [18, 19]. The results indicated that the hybrid approach improved recognition precision. However, DNN-based face recognition algorithms are often built on large original datasets, which can be difficult to obtain [15, 20, 21]. While larger datasets can increase model accuracy and the network's capacity for generalization, labeling the data is a laborious and time-consuming process. Consequently, developing DNN-based face recognition methods on a limited original dataset is an interesting issue [21].

2.2 Emotional recognition

Computer vision is a promising method for detecting driver conscious primarily due to its non-intrusive nature, which does not require drivers to wear any special equipment that might cause discomfort or distraction. It utilizes cameras to observe and analyze facial cues and behaviors that are indicative of conscious, such as eyelid closure and head nodding. This method leverages existing camera hardware in many modern vehicles, making it practical and cost-effective [22].

In comparison to other methods, such as physiological signal monitoring (which might involve measuring heart rate variability, brain waves, or skin conductance), computer vision-based systems are less invasive and easier to implement at scale [23]. While physiological methods may provide direct measures of a driver's state, they require specialized sensors and can be impractical for everyday use in a typical driving environment.

Computer vision methods also tend to be preferred over behavioral measures, such as steering patterns or lane-keeping ability, because they can detect driver conscious before it affects driving performance, thereby providing an earlier warning system. Furthermore, advancements in artificial intelligence and machine learning have significantly improved the accuracy and reliability of computer vision techniques, enabling real-time processing and analysis that can match or even surpass other methods in both convenience and performance. For instance, the authors [24] presented a Driver Drowsiness Detection system employing four deep learning

models to identify driver drowsiness based on RGB videos. This system encompasses four major categories of characteristics: facial expressions, behavioral traits, and hand and head movements, all to detect signs of tiredness and sleepiness.

Another system, called DriCare, developed by Deng and Wu [25], utilizes video frames to detect indicators of driver fatigue, such as blinking, yawning, and the duration of eye closure, without requiring any physical attachments to the driver. To enhance tracking accuracy, the authors devised a new facial tracking algorithm due to the limitations of existing algorithms.

Maior et al. [26] introduced a cost-effective real-time system aimed at reducing the risk of human errors and preventing accidents. This system aims to enhance safety. Additionally, Zandi et al. [27] incorporated nonintrusive eye-tracking data to detect driver drowsiness. In a simulated driving experiment, data from 53 volunteers' eye movements was collected and analyzed using RF and non-linear SVM algorithms to classify their alertness levels. Region of Interest (ROI) images were employed to assess the condition of the eyes and mouth [28].

3. PROPOSED METHODOLOGY

The terms "conscious" and "no conscious" cases within the context of this study refer to the driver's state of alertness, with "conscious" denoting a state of wakefulness and active responsiveness, and "no conscious" indicating a reduced responsiveness, potentially symptomatic of sleep or drowsiness. These terms are critical as they underpin the system's capability to assess a driver's readiness to operate a vehicle safely.

The methodology distinguishes between conscious and unconscious states by analyzing the status of the person's eyes, whether they are open or closed. To accomplish this, we employ facial images of the four individuals, including both closed and open-eye instances, as the input dataset for the deep learning model. This approach allows us to make a definitive determination regarding their conscious status.

The basic steps for each stage in this methodology are described in the following sections and the overall proposed methodology is shown in Figure 1.

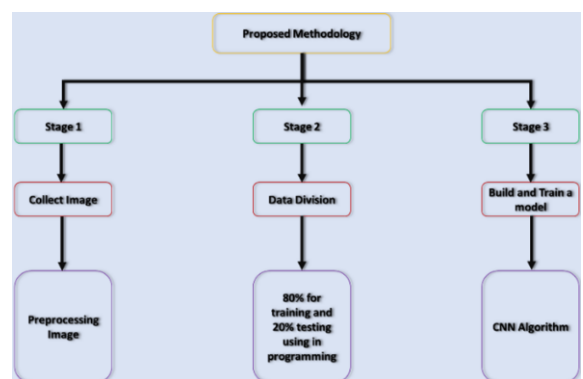


Figure 1. Overall proposed methodology

3.1 Dataset collection

The images collected for use in our proposed system include two types of recognition. First, we carry out 2000 photos for

four persons (Ali, Hiba, Lara, and Yaraa) in order to satisfy the face recognition stage. Each person (class) has 500 images (400 photos for the training set, which performs 80%, and 100 photos for the validation set, which performs 20%), as shown in Figure 2(a). Second, 1600 photos run on the same classified persons to determine the classification for conscious and no conscious cases. As shown in Figure 2(b), each class has 800 photos (600 photos for the training set, which performs 80%, and 200 photos for the validation set, which performs 20%).

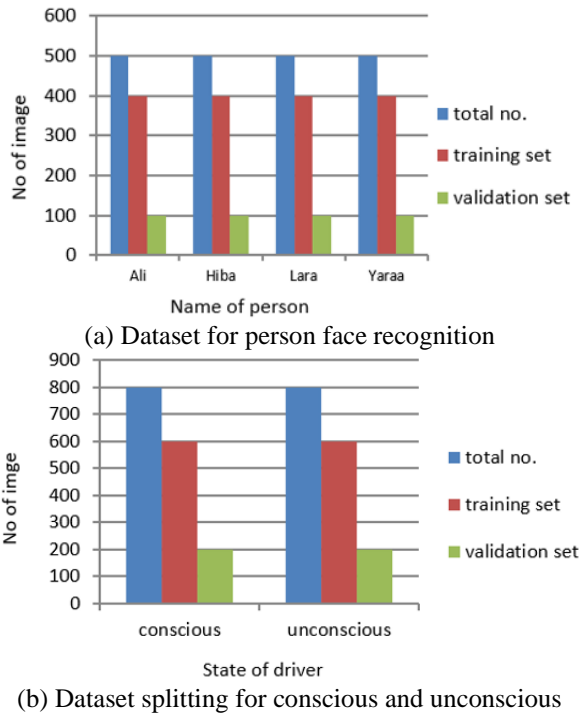


Figure 2. Dataset allocation for each stage of the proposed classification CNN model

3.2 Pre-processing stage

Mostly, image pre-processing techniques are applied before many machine learning and deep learning processes. Image pre-processing includes many popular methods, such as reducing calculations and dimensions, thus the performance can be improved. The dataset normalized the various sizes of the original photos to 224×224 pixels. Before data splitting, the data shuffled to ensure that training includes both sorted and unsorted data from the dataset as well as narrow data. The intensity pixel value range is changed by normalizing the image's pixel values. The complicated calculation cause by using the image's channel (0-255), which normalizes the image's pixel range to be in this range (0-1). Data augmentation techniques are employed with Keras augmentations to enhance both the quantity and quality of training samples, while also mitigating overfitting issues. They also increase the generality and efficacy of the model [29]. The augmentations are set up before the models are trained. the training samples expanded dataset by three by using three augmentation approaches (rotation, horizontal flip, and zoom). The majority of the newly formed training set is composed of the initial training images that have been enhanced using the selected augmentation strategies. The following are descriptions of each augmentation [30, 31]:

(1) Rotation: A 20-degree rotation is applied to the image to aid in its classification.

(2) Horizontal Flipping: This geometric enhancement flips the image along the horizontal axis.

(3) The CIFAR-10 and ImageNet datasets have shown its benefits, and it is more prevalent than vertical flip.

(4) Zoom: To better show details, we magnified the images.

3.3 Proposed system

This work, use two deep-learning CNN models for four persons. The first model is for face recognition, and the other is for conscious and unconscious recognition Figure 3(a) demonstrates the block diagram, used to implement the classification process for person identities. face images for four persons (Ali, Hiba, Lara, and Yaraa) used as the input data set for the CNN model. This model divides the images into four categories for the persons mentioned before, thus allow to recognize the person that are looking for. Figure 3(b) shows the emotional-based conscious and unconscious recognition model, used for completing the classification process for conscious and unconscious people. The proposed models lead to identifying the driver (one of the four authorized people) and then recognizing if this driver is conscious or unconscious.

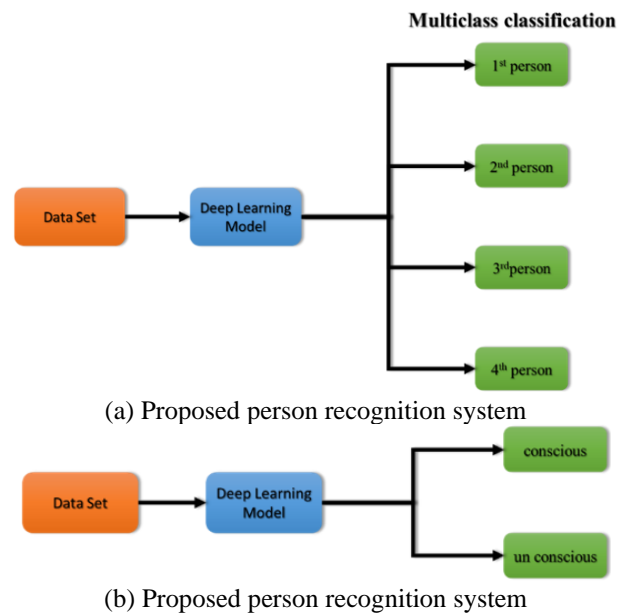


Figure 3. Block diagrams for the proposed systems

4. THE PROPOSED CNN MODEL

The proposed backbone CNN model, used for driver and emotional recognition, is illustrated in Figure 4, which lists the CNN's 15 layers. We designed CNN's layers as follows:

(1) Starting from the first layer which is performed as the input layer. It carried out images from the pre-processing stage described in the pre-processing stage section.

(2) Three stages of convolution layers, each stage consists of convolution and rectified linear unit (Relu), which are the activation function, max-pooling, and dropout ranges (from 25% to 50%) layers.

(3) One fully connected layer is implemented.

(4) The dropout layer with a probability ratio equal to 50% occurs before the final layer (sigmoid layer). The dropout layer uses four classes of face images in the CNN model for face recognition, while the softmax layer uses two classes of face images in the CNN model of emotional recognition.

The differences between both proposed CNN models in terms of sophisticated description of each layer including

output shape and parameters are illustrated in Table 1, and Table 2.

Table 1. An overview of CNN's layer of face recognition

Layer (Type)	Output Shape	Param #
image_input (InputLayer)	[(None, 224, 224, 3)]	0
layer_1 (Conv2D)	(None, 224, 224, 32)	896
layer_2 (Conv2D)	(None, 224, 224, 64)	18496
layer_3 (MaxPooling2D)	(None, 112, 112, 64)	0
layer_4 (Conv2D)	(None, 112, 112, 64)	36928
layer_5 (MaxPooling2D)	(None, 56, 56, 64)	0
dropout_1 (Dropout)	(None, 56, 56, 64)	0
layer_6 (Conv2D)	(None, 56, 56, 128)	73856
layer_7 (MaxPooling2D)	(None, 28, 28, 128)	0
dropout_2 (Dropout)	(None, 28, 28, 128)	0
fc_1 (Flatten)	(None, 100352)	0
layer_8 (Dense)	(None, 64)	6422592
dropout_3 (Dropout)	(None, 64)	0
pridictions (Dense)	(None, 4)	260

Table 2. An overview of CNN's layers of emotional recognition

Layer (Type)	Output Shape	Param #
image_input (InputLayer)	[(None, 224, 224, 3)]	0
layer_1 (Conv2D)	(None, 224, 224, 32)	896
layer_2 (Conv2D)	(None, 224, 224, 64)	18496
layer_3 (MaxPooling2D)	(None, 112, 112, 64)	0
layer_4 (Conv2D)	(None, 112, 112, 64)	36928
layer_5 (MaxPooling2D)	(None, 56, 56, 64)	0
dropout_1 (Dropout)	(None, 56, 56, 64)	0
layer_6 (Conv2D)	(None, 56, 56, 128)	73856
layer_7 (MaxPooling2D)	(None, 28, 28, 128)	0
dropout_2 (Dropout)	(None, 28, 28, 128)	0
fc_1 (Flatten)	(None, 100352)	0
layer_8 (Dense)	(None, 64)	0
dropout_3 (Dropout)	(None, 64)	0
pridictions (Dense)	(None, 2)	130

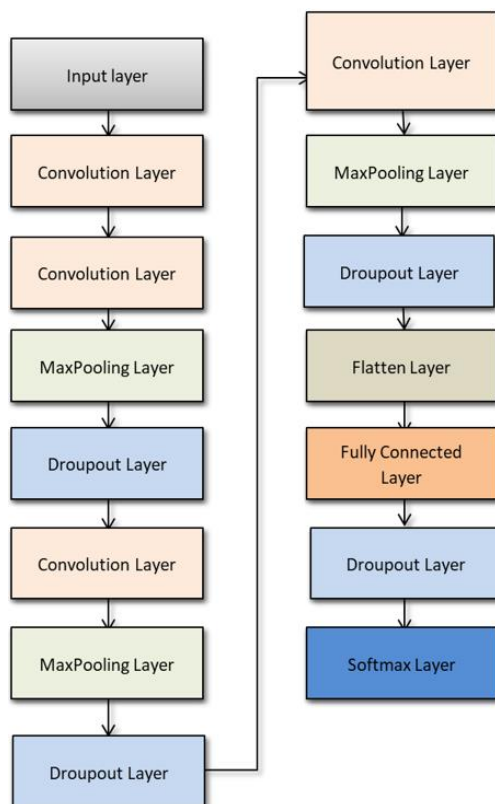


Figure 4. CNN model proposed

The proposed CNN model's architecture above, comprised of 15 layers, has been meticulously designed to facilitate efficient feature extraction and classification for both driver identification and emotional state recognition. The rationale for the chosen configuration is rooted in the model's ability to progressively abstract and interpret complex features from facial images, which necessitates a deliberate layering strategy that combines convolutional layers, ReLU activations, pooling, and dropout layers.

ReLU activation functions are particularly employed for their proven efficacy in mitigating the vanishing gradient problem and facilitating faster convergence during training compared to other activation functions such as sigmoid or tanh, thereby optimizing the network's performance for the high-dimensional data characteristic of image recognition tasks.

4.1 Convolution layer

The convolution layer, which comprises shared weights and local ties, is the main part of the CNN model. Its goals are to comprehend how to represent entered features. It has a number of feature maps. To extract the local properties at the various positions of the preceding layer, it is employed the similarity of the neuron properties in various regions. In the preceding layer, individual neurons facilitate the extraction of characteristics within the same feature map region. To extract features from the input image, we utilized a convolution filter (kernel) with three distinct widths (33, 55, and 77) that were

$$D2 = D1 \quad (4)$$

overlapped both horizontally and vertically. Stride is two steps and padding is one pixel, correspondingly. the convolution layer actions are shown in Figure 5. A non-linear activation function called Relu is then activated as a result of the result in Figure 6 clarifies Relu's work [32].

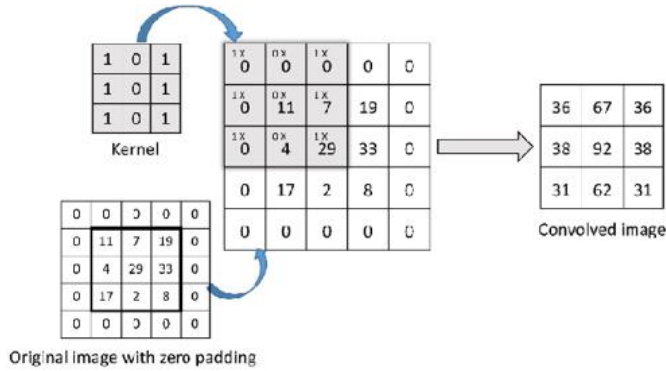


Figure 5. The example for operations of convolution layer (where the input_shape: 3×3, padding=1, kernel_size= 3×3, Stride = 2, output_shape: 2×2)

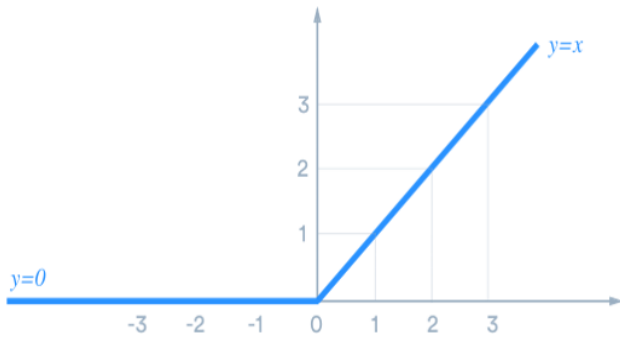


Figure 6. ReLU activation function [26]

The Eq. (1) of Relu is [33]:

$$F(X) = \text{MAX}(0, X) \quad (1)$$

where, $F(X) = X$ if x : positive and $F(X) = 0$ if x : negative.

4.2 Max pooling layer

The convolution layer (often many convolutions and pool layers rounds) minimizes the amount of information in each feature it has collected while maintaining the most crucial data. The CNN model usually uses large photographs; thus, we must limit the number of parameters in our photos. All employed max-pooling layers are (2, 2) except the stride, which can move both vertically and horizontally. To do this, the image is divided into smaller squares (2×2 is the preferred method), and each of them is passed over the complete image with a special string (2×2). Following that, the highest value is selected from the matrix of four numbers [34]. Figure 7 illustrates the Max pooling layer procedure.

In Eq. (2) - Eq. (4) are used for Max pooling [35]:

$$W2 = \frac{W1 - F}{S + 1} \quad (2)$$

$$H2 = \frac{H1 - F}{S + 1} \quad (3)$$

where, F is the spatial extent of the filter, S is the stride, $D1$ is the depth of the convolution layer, $D2$ is the depth-convolution layer, $W1$ is the width of the convolution layer, $H1$ is the height of the convolution layer, $W2$ is the maximum layer width, and $H2$ is the maximum layer height.

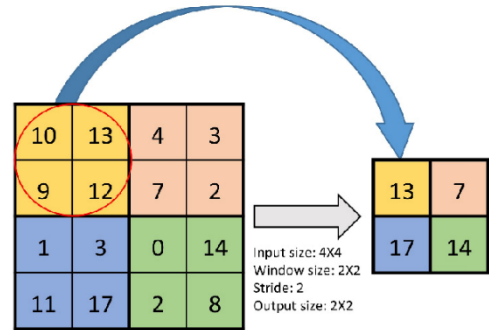


Figure 7. The operation of the max-pooling layer to one block

4.3 Dropout layer

In this layer, a substantial portion of activations (nodes) is intentionally dropped at random, leading to a noticeable reduction in training time and mitigating overfitting concerns. In our suggested approach, the dropout probability for the four Dropout layers succeeding the Max pooling layers ranged between 25% and 50%. The dropout ratio for the ultimate layer, subsequent to the fully connected layer, is set at 30% [36]. Figure 8 illustrates the presentation of dropout layers.

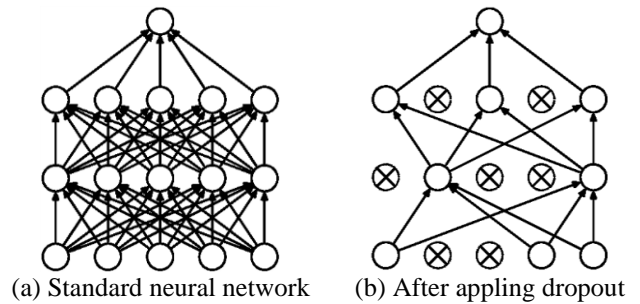


Figure 8. Example of the dropout layer [37]

4.4 Fully connected layers

A typical neural network's layers are organized similarly to how neurons are organized. As seen in Figure 9, each node in a wholly linked layer is directly connected to every other node in the levels above and below it. The first layer in the preceding frame is the vector that joins the nodes in the pooling layer and the layer is fully connected. Some variables require more time to practice for. With a 30% possibility, the Dropout approach is employed to reduce the number of nodes and links before the Softmax layer. A combination of different sigmoid activation functions results in Softmax. The sigmoid function generates results between 0 and 1. These can be thought of as the likelihood connected to a group of data items. Although Softmax can be used to tackle issues needing several classes, sigmoid functions are often utilized for binary classification. The function, as stated in the explanation in Eq.

(5) [38], returns the likelihood for each data item across all groups:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^k e^{z_k}} \text{ for } j = 1, \dots, k. \quad (5)$$

where, K represents the number of classes, resulting in an output Z equal to 1. When designing a network or model for multiple classifications, the output layer can feature the same number of neurons as classes within the target. Our investigation revealed a three-class scenario (Armenia, Belarus, and Hungary) through the use of Softmax activation, as depicted in Figure 9. Finally, the cross-entropy loss function was employed to compute the classification loss and predict the label for the input image. Eq. (6) [39] outlines the cross-entropy equation as follows:

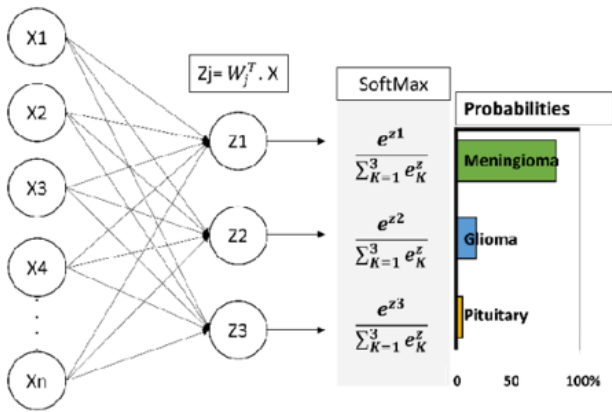
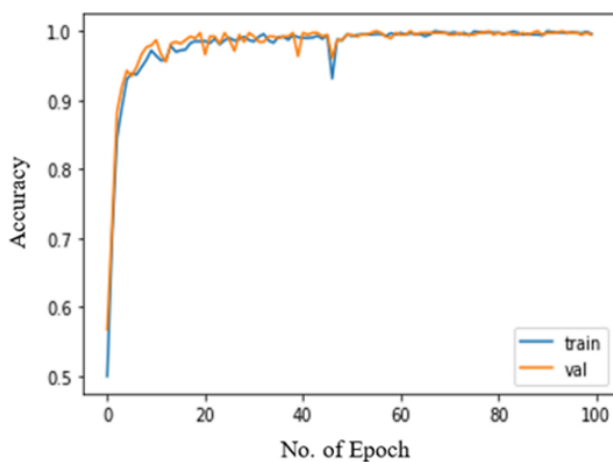


Figure 9. Softmax activation function example

$$L(Y, \hat{Y}) = - \sum_{k=0}^n Yk * \log(\hat{Y}k) \quad (6)$$

In this equation, \hat{Y} represents the predicted output layer ratio, while y denotes the binary indicator signifying whether the classification is accurate (1) or not (0). The variable n corresponds to the number of classes or labels (i.e., the number of output nodes).



(a) Training and validation accuracy

4.5 Optimization algorithm

As part of the training of the CNN model, a crucial aspect involves the iterative adjustment of the network's layer parameters. The optimizer plays a pivotal role in the training of deep CNN models, as it works to minimize the loss function. For this purpose, the Adam optimizer was employed, proving its effectiveness for learning. Notably, this optimizer operates with minimal memory consumption and achieves rapid computation [40].

5. RESULTS AND DISCUSSION

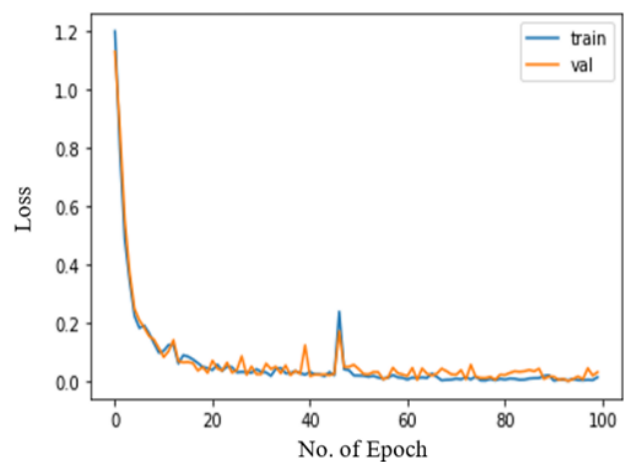
In this paper, the proposed CNN models implemented in Python using Anaconda Python 3.7, Keras (the Tensor Flow backend), and Adam as the optimizer. the algorithms carried out on a laptop operating Windows 10; the CPU is an Intel R Core (TM) i7-6820 HQ CPU that has 4 physical cores and 8 logical cores, and the RAM is 8 GB. One essential key to evaluating the accuracy of the recognition algorithms is to find the accuracy ratio. The accuracy ratio is calculated as follows [41]:

$$A_{cc} = \frac{N_{cc}}{T_s} \times 100\% \quad (7)$$

where, A_{cc} is the accuracy of classification, N_{cc} is the number of correctly classified images, and T_s is the total number of samples. In addition, we used categorical cross-entropy as a loss function, and we chose a learning rate of 0.0001.

5.1 Person face recognition

The accuracy and loss for training and validation sets of the proposed face recognition model (driver recognition) are illustrated in Figure 10. From Figure 10(a), the accuracy ratio is reaching saturation at 60 epochs in both the training and validation processes. Moreover, Figure 10(b) shows that the losses occurred immediately after 60 epochs. Losses have varied throughout time; the curve begins to degrade after 48 epochs, at which point it stabilizes. The proposed model is effective despite the dataset for each class only having 500 images. The accuracy for the training and validation sets are 99.63% and 99.48%, respectively.



(b) Training and validation loss

Figure 10. Accuracy and loss in training and validation at various epochs

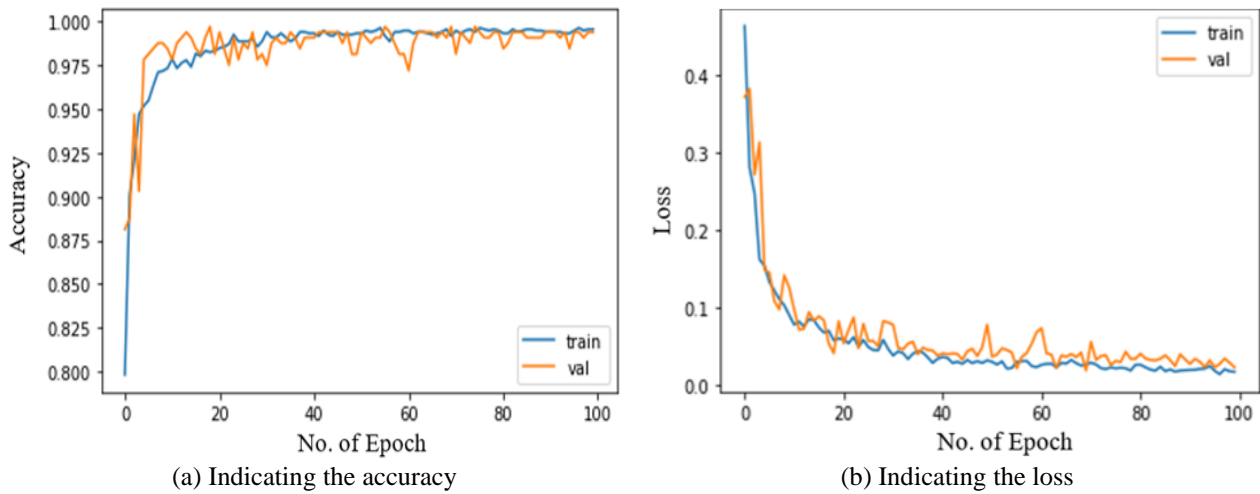


Figure 11. Displays the accuracy and loss of training and validation at various epochs

5.2 Emotional recognition

The suggested model's accuracy over the course of the training and validation stages is shown in Figure 11(a). Figure 11(b) contrasts this with loss that varies over time, starts falling after 67 epochs, stabilizes, and then decelerates for the next 100 epochs. Despite only having 800 image samples for each class in the dataset, the suggested model is successful, with accuracy scores for the training and validation sets of 99.54% and 99.37%, respectively.

5.3 Cross validation

In this section, used cross-validation to examine a new set of images that we had not seen before during training. The goal was to ensure that the proposed models worked properly. This assessment consisted of two stages: first recognizing the individual and then determining their emotional state. After that tested the face recognition model with 120 photos (30 images per person), yielding an average accuracy of 99.1%, as shown in Table 3.

Table 3. The effectiveness of the proposed CNN model for face recognition

Person Name	Total Test Images	No. of Correct Recognition Image	No. of Incorrect Recognition Image	Image Recognition Percentage
Ali	30	29	1	96.6%
Hiba	30	30	0	100%
Lara	30	30	0	100%
Yara	30	30	0	100%
Total	120	119	1	99.15

Table 4. Comparison of face recognition accuracy

Method	Accuracy%
ANN	80.3%
PCA+ANN	91.0%
PCA+SVM	97.4%
Wavelet+SVM	98.0%
Wavelet+PCA+SVM	98.1%
Proposed method (CNN)	99.15%

In order to further substantiate the efficacy of proposed approach, we carried out a comparative evaluation involving several other face recognition methodologies documented in the existing literature. These encompass ANN (artificial neural network), PCA+ANN, PCA+SVM, Wavelet+SVM, and Wavelet+PCA+SVM. The outcomes, detailing the respective

accuracies in face recognition, have been detailed in Table 4. Notably, it's observed that methods employing SVM tend to exhibit superior performance in face recognition accuracy compared to those based on ANN. However, proposed approach, which combines CNN with an augmented face dataset, consistently achieves even higher levels of accuracy in contrast to the other methods. This underscores the proposition that the integration of an augmented face dataset can yield substantial performance improvements in face recognition by capitalizing on the diverse set of features available.

Moreover, a test phase was conducted using a total of 60 images, evenly divided between the categories of conscious and unconscious models (30 images per case), which resulted in an impressive accuracy rate of 98.3%, as shown in Table 5.

Table 5. The effectiveness of the proposed CNN model for conscious recognition

State of Recognition Image	Total Test Images	No. of Correct Recognition Image	No. of Incorrect Recognition Image	Image Recognition Percentage
Conscious	30	29	1	96.6
Nonconscious	30	30	0	100%
Total	60	59	1	98.3

Table 6. Comparison of face recognition method accuracy

Method	Accuracy%
Deep-CNN-based ensemble	85%
RF and non-linear SVM	RF: 88.37% to 91.18% SVM: 77.1% to 82.62%
Multiple CNN- kernelized correlation filters method	92%
Multilayer perceptron, RF, and SVM	94.9%
Dlib's Haar Cascade model	98.0%
The proposed model by CNN	98.3%

A comparison of our proposed classification method with various other classification approaches documented in the literature is presented in Table 6.

6. CONCLUSION

In conclusion, this study presented two deep-learning models utilizing Convolutional Neural Networks (CNNs) to achieve driver face recognition and conscious recognition. These models aimed not only to identify the driver but also to discern their level of consciousness by analyzing their emotional state. The application of these models could potentially result in appropriate actions like halting the vehicle or contacting emergency services as needed.

To facilitate the creation of these models, two separate datasets were gathered and prepared. These datasets contained facial images from four different individuals, which were then employed for training the models and evaluating their performance through validation. The incorporation of diverse data for each individual ensured a thorough training process, leading to remarkably high accuracy rates of over 99% for both tasks.

Additionally, the models' effectiveness was validated through cross-validation experiments. These experiments entailed assessing the models' performance on images not used during training, producing impressive accuracy rates ranging from 95% to 100%. This outcome solidified the strength and dependability of the proposed deep-learning models.

As is common with research endeavors, there are possibilities for future exploration and improvement. One potential avenue for future research involves expanding the dataset to include a more extensive and diverse range of individuals, facial expressions, and emotional states. This broader dataset could enhance the models' ability to generalize to a wider population and increase their relevance in real-world scenarios.

Furthermore, in practical applications, the integration of these deep-learning models into existing driver assistance systems could substantially elevate road safety by providing real-time monitoring of driver alertness, thereby enabling proactive measures such as automatic vehicle control to prevent accidents due to un conscious

REFERENCES

[1] Kumar, A., Kaur, A., Kumar, M. (2019). Face detection techniques: A review. *Artificial Intelligence Review*, 52: 927-948. <https://doi.org/10.1007/s10462-018-9650-2>

[2] Sabah Sabah, S.H., Croock, M.S. (2020). Software engineering based fault tolerance method for wireless sensor network. *Iraqi Journal of Computers,*

Communications, Control and Systems Engineering, 20(4): 21-28. <https://doi.org/10.33103/uot.ijccce.20.4.3>

[3] Abdul Ameer, H.R., Hasan, H.M. (2020). Enhanced MQTT protocol by smart gateway. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, 20(1): 53-67. <https://doi.org/10.33103/uot.ijccce.20.1.6>

[4] Al-Khazraji, H., Nasser, A.R., Khilil, S. (2022). An intelligent demand forecasting model using a hybrid of metaheuristic optimization and deep learning algorithm for predicting concrete block production. *IAES International Journal of Artificial Intelligence*, 11(2): 649. <https://doi.org/10.11591/ijai.v11.i2.pp649-657>

[5] Dwivedi, K., Biswaranjan, K., Sethi, A. (2014). Drowsy driver detection using representation learning. In *IEEE International Advance Computing Conference (IACC)*, Gurgaon, India, pp. 995-999. <https://doi.org/10.1109/IAdCC.2014.6779459>

[6] Nafea, S., Hamza, E.K. (2020). Path loss optimization in WIMAX network using genetic algorithm. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, 20(1): 24-30. <https://doi.org/10.33103/uot.ijccce.20.1.3>

[7] Ali, M.O., Abou-Loukh, S.J., Al-Dujaili, A.Q., Alkhayyat, A.H., Abdulkareem, A.I., Ibraheem, I.K., Humaidi, A.J., Al-Qassar, A.A., Azar, A.T. (2022). Radial basis function neural networks-based short term electric power load forecasting for super high voltage power grid. *Journal of Engineering Science and Technology*, 17(1): 361-378.

[8] Albawi, S., Mohammed, T.A., Al-Zawi, S. (2017). Understanding of a convolutional neural network. In *International Conference on Engineering and Technology (ICET)*, Antalya, Turkey, pp. 1-6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>

[9] Hashim, S.A., Jawad, M.M., Wheedd, B. (2020). Study of energy management in wireless visual sensor networks. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, 20(1): 68-75. <https://doi.org/10.33103/uot.ijccce.20.1.7>

[10] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61: 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>

[11] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, 234: 11-26. <https://doi.org/10.1016/j.neucom.2016.12.038>

[12] Ptucha, R., Such, F.P., Pillai, S., Brockler, F., Singh, V., Hutkowski, P. (2019). Intelligent character recognition using fully convolutional neural networks. *Pattern recognition*, 88: 604-613. <https://doi.org/10.1016/j.patcog.2018.12.017>

[13] Zeng, N., Zhang, H., Song, B., Liu, W., Li, Y., Dobaie,

- A. M. (2018). Facial expression recognition via learning deep sparse autoencoders. *Neurocomputing*, 273: 643-649. <https://doi.org/10.1016/j.neucom.2017.08.043>
- [14] Zangeneh, E., Rahmati, M., Mohsenzadeh, Y. (2020). Low resolution face recognition using a two-branch deep convolutional neural network architecture. *Expert Systems with Applications*, 139: 112854. <https://doi.org/10.1016/j.eswa.2019.112854>
- [15] Zhang, Z., Li, J., Zhu, R. (2015). Deep neural network for face recognition based on sparse autoencoder. In 2015 8th International Congress on Image and Signal Processing (CISP), Shenyang, China, pp. 594-598. <https://doi.org/10.1109/CISP.2015.7407948>
- [16] Lin, G., Shen, W. (2018). Research on convolutional neural network based on improved Relu piecewise activation function. *Procedia Computer Science*, 131: 977-984. <https://doi.org/10.1016/j.procs.2018.04.239>
- [17] Mohammed, A.A., Noaman, M.A., Azzawi, H.M. (2022). A combining two KSVM classifiers based on True pixel values and discrete wavelet transform for MRI-based brain tumor detection and classification. *Engineering and Technology Journal*, 40(02): 322-333. <http://doi.org/10.30684/etj.v40i2.2180>
- [18] Zhao, F., Li, J., Zhang, L., Li, Z., Na, S.G. (2020). Multi-view face recognition using deep neural networks. *Future Generation Computer Systems*, 111: 375-380. <https://doi.org/10.1016/j.future.2020.05.002>
- [19] Gumus, E., Kilic, N., Sertbas, A., Ucan, O.N. (2010). Evaluation of face recognition techniques using PCA, wavelets and SVM. *Expert Systems with Applications*, 37(9): 6404-6408. <https://doi.org/10.1016/j.eswa.2010.02.079>
- [20] Guo, G., Zhang, N. (2019). A survey on deep learning based face recognition. *Computer Vision and Image Understanding*, 189: 102805. <https://doi.org/10.1016/j.cviu.2019.102805>
- [21] Sun, J., Meng, F. (2016). Face recognition based on deep neural network and weighted fusion of face features. *Journal of Computer Applications*, 36(2): 437-443. <https://doi.org/10.11772/j.issn.1001-9081.2016.02.0437>
- [22] Bhatti, U.A., Yu, Z., Yuan, L., Nawaz, S.A., Aamir, M., Bhatti, M.A. (2022). A robust remote sensing image watermarking algorithm based on region-specific SURF. In *Proceedings of International Conference on Information Technology and Applications: ICITA 2021*, Singapore, pp. 75-85. https://doi.org/10.1007/978-981-16-7618-5_7
- [23] Budak, U., Bajaj, V., Akbulut, Y., Atila, O., Sengur, A. (2019). An effective hybrid model for EEG-based drowsiness detection. *IEEE Sensors Journal*, 19(17): 7624-7631. <https://doi.org/10.1109/JSEN.2019.2917850>
- [24] Dua, M., Shakshi, Singla, R., Raj, S., Jangra, A. (2021). Deep CNN models-based ensemble approach to driver drowsiness detection. *Neural Computing and Applications*, 33: 3155-3168. <https://doi.org/10.1007/s00521-020-05209-7>
- [25] Deng, W., Wu, R. (2019). Real-time driver-drowsiness detection system using facial features. *IEEE Access*, 7: 118727-118738. <https://doi.org/10.1109/ACCESS.2019.2936663>
- [26] Maior, C.B.S., das Chagas Moura, M.J., Santana, J.M.M., Lins, I.D. (2020). Real-time classification for autonomous drowsiness detection using eye aspect ratio. *Expert Systems with Applications*, 158: 113505. <https://doi.org/10.1016/j.eswa.2020.113505>
- [27] Zandi, A.S., Quddus, A., Prest, L., Comeau, F.J. (2019). Non-intrusive detection of drowsy driving based on eye tracking data. *Transportation Research Record*, 2673(6): 247-257. <https://doi.org/10.1177/0361198119847985>
- [28] Bajaj, S., Panchal, L., Patil, S., Sanas, K., Bhatt, H., Dhakane, S. (2023). A real-time driver drowsiness detection using OpenCV, DLib. In: *ICT Analysis and Applications*, pp. 639-649. https://doi.org/10.1007/978-981-19-5224-1_64
- [29] Asperti, A., Mastronardo, C. (2017). The effectiveness of data augmentation for detection of gastrointestinal diseases from endoscopic images. In *Proceedings of the 5th International Conference on Bioimaging, BIOIMAGING 2018*, Funchal, Madeira – Portugal, pp. 1-7. <https://doi.org/10.48550/arXiv.1712.03689>
- [30] Yang, S., Xiao, W., Zhang, M., Guo, S., Zhao, J., Shen, F. (2022). Image data augmentation for deep learning: A survey. *arXiv preprint arXiv:2204.08610*. <https://doi.org/10.48550/arXiv.2204.08610>
- [31] Li, W., Chen, C., Zhang, M., Li, H., Du, Q. (2018). Data augmentation for hyperspectral image classification with deep CNN. *IEEE Geoscience and Remote Sensing Letters*, 16(4): 593-597. <https://doi.org/10.1109/LGRS.2018.2878773>
- [32] Al-Saffar, A.A., Tao, H., Talab, M.A. (2017). Review of deep convolution neural network in image classification. In *2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, Jakarta, Indonesia, pp. 26-31. <https://doi.org/10.1109/ICRAMET.2017.8253139>
- [33] Sultan, H.H., Salem, N.M., Al-Atabany, W. (2019). Multi-classification of brain tumor images using deep neural network. *IEEE Access*, 7: 69215-69225. <https://doi.org/10.1109/ACCESS.2019.2919122>
- [34] Paul, J.S., Plassard, A.J., Landman, B.A., Fabbri, D. (2017). Deep learning for brain tumor classification. In *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, Orlando, Florida, United States, pp. 253-268. <https://doi.org/10.1117/12.2254195>
- [35] Jmour, N., Zayen, S., Abdelkrim, A. (2018). Convolutional neural networks for image classification. In *2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET)*, Hammamet, Tunisia, pp. 397-402. <https://doi.org/10.1109/ASET.2018.8379889>
- [36] Park, S., Kwak, N. (2017). Analysis on the dropout effect in convolutional neural networks. In *13th Asian Conference on Computer Vision*, Taipei, Taiwan, pp. 189-204. https://doi.org/10.1007/978-3-319-54184-6_12
- [37] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1): 1929-1958.
- [38] Sharma, S., Sharma, S., Athaiya, A. (2017). Activation functions in neural networks. *International Journal of Engineering Applied Sciences and Technology*, 4(12): 310-316.
- [39] Iqbal, S., Ghani, M. U., Saba, T., Rehman, A. (2018). Brain tumor segmentation in multi - spectral MRI using convolutional neural networks (CNN). *Microscopy Research and Technique*, 81(4): 419-427. <https://doi.org/10.1002/jemt.22994>

[40] Metz, L., Maheswaranathan, N., Nixon, J., Freeman, D., Sohl-Dickstein, J. (2019). Understanding and correcting pathologies in the training of learned optimizers. In International Conference on Machine Learning, California, USA, pp. 4556-4565.

[41] Hidayah, M.R., Akhlis, I., Sugiharti, E. (2017). Recognition number of the vehicle plate using Otsu method and K-nearest neighbour classification. Scientific Journal of Informatics, 4(1): 66-75. <https://doi.org/10.15294/sji.v4i1.9503>