**Journal of Theoretical and Applied Information Technology**
28th February 2023. Vol.101. No 4


ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# AN EFFECTIVE HYBRID FEATURES FOR DRIVER FATIGUE DETECTION USING CONVOLUTIONAL NEURAL NETWORK

**V.VIJAYPRIYA[1], M. UMA[2]**

[1]Research Scholar, Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur, India.

[2]Associate Professor, Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur, India.

E-mail: [1]vv8380@srmist.edu.in, [2]umam@srmist.edu.in

**ABSTRACT**

Drowsy driving is one of the leading causes of road accidents worldwide. Drowsiness is caused by a lack of sleep and irregular sleeping patterns. People who are drowsy may have difficulty driving and cause major accidents, as evidenced by the high number of injuries and deaths worldwide. As a result, detecting driver drowsiness is critical in order to save many lives. The proposed model made use of non-intrusive methods related to fatigue states. The proposed system employs a multi-task convolutional neural network to predict driver fatigue states based on the percentage of eye closure (PERCLOS) and yawning detection in a video stream. Faces are identified using facial landmarks in the extracted images from the video streams, and the trained model predicts whether the person's face in the image is fatigued with closed eyes and yawning or non-fatigued with open eyes, talking, and a closed mouth. In comparison to the existing model, our proposed model has been tested on the D3S publicly available datasets and has produced better classification models for predicting fatigue states with 97% accuracy using deep learning models.

**Keywords:** *Convolutional Neural Network, Driver Drowsiness, Facial Landmark, Percentage Of Eye Closure*

## 1. INTRODUCTION

According to the statistical data from the Ministry of Road Transport and Highways, India ranks first in road accident deaths compared to other countries [1]. Every year in India, more than 150,000 people die due to road accidents caused by cars, taxis, and vans. Every year, drivers fall asleep on the road, leading to thousands of deaths and severe injuries. According to the Transport Commissionerate and State Transport Authority Report for 2020, there were 40725 accidents, 7287 deaths, and 45286 injuries in Tamilnadu. In 2020, Tamil Nadu reported the highest number of road accidents. The accidents caused by cars, Jeeps, and taxis/tempos account for 25.89% of all road accidents that occur all over the world [2]. In 2019, the AAA Foundation for Traffic Safety's annual Culture Index study found 96% of drivers find drowsy driving a dangerous driving behaviour; 29% of drowsy drivers are caught by police; and 24% of drivers feel tired and find it hard to keep their eyes open at least once in the past 30 days [3]. The major causes of road accidents in India are distracted driving, intoxicated driving, speeding, not wearing a seatbelt, rain or wet roads, potholes, bad road conditions, breaking traffic rules, multi-tasking while driving, and tailgating [4]. Road accidents can occur for a variety of reasons,

including human error, road conditions, or vehicle conditions [5]. When comparing the various road accident reasons, driver fatigue and drowsiness are the major reasons for the occurrence of road accidents due to human error. Texting, talking, viewing Google Maps, distraction, and drowsiness are some of the activities caused by a person while driving. It may lead to devastating accidents for the person driving, the passengers travelling with them, or other road users [6]. A driver assistance technique can be used to sense the drowsiness of the driver, and precautionary notification can be used to prevent accidents caused by drowsiness [7]. Some of the drowsiness symptoms include frequent blinking, dropping eyelids, rubbing eyes, nodding heads, and yawning frequently. In all types of cars, intrusive or non-intrusive methods are used to modify the driver's drowsiness.

The intrusive method is based on physiological methods like an electrocardiogram (ECG), an electroencephalogram (EEG), a photoplethysmogram (PPG), and an

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

electrooculogram (EOG). The intrusive method uses several electrodes and sensors that are fitted to the driver's body. It may cause the driver discomfort while driving. The non-intrusive is based on two methods: the behavioural method (like blinking of eyes, head pose, yawning, and closing of eyes) and the vehicle method (like controlling wheel movements, acceleration, lane position, and speed) [8]. As compared with existing methodology, the behavioural faults of people driving lead to 90% of road accidents. In 2030, the roads' higher traffic injuries could cause several deaths around the world [9]. It is comfortable for the driver to detect drowsiness using a non-intrusive method like the behavioural method.

Savas Burcu Kir and Yasar Becerikli proposed a multi-task convolutional neural network model for detecting driver fatigue using both eye and mouth features and classified fatigue states into three classes: very tired, not tired, and less tired, which produced 98% accuracy using two datasets, YawDD and NthuDDD [10]. Ji, Yingyu, et al. proposed a multi-task cascaded convolutional neural network for detecting driver fatigue using a separate network structure for eye and mouth state recognition networks and classified fatigue states using a fatigue judgement model, which produced 98% accuracy using different types of datasets [11]. Gupta, Isha, et al. proposed a non-intrusive method using a support vector machine (SVM) classifier for detecting a person's face or non-face and used sleepiness and yawning features for detecting drowsiness using a particular threshold value for checking the driver's drowsiness state and produced 90% accuracy using the D3S dataset [12]. Saurav, Sumeet, et al. used an in-depth learning feature for detecting yawning, which used two networks: a bi-directional LSTM network and a convolutional neural network for predicting drowsiness from various video streams using three datasets: YawDD, NthuDDD, and the FER+ dataset, all of which predicted drowsiness with 95% accuracy [13]. Bai, Jing, et al. used both spatial and temporal graph-based convolutional networks to detect driver drowsiness using four features such as blinking, closing eyes, yawning, and talking collected from two datasets, NthuDDD and YawDD, and produced 93% accuracy [14]. Omidyeganeh, Mona, et al. used an embedded smart camera to detect yawning features using two methods: the Viola-Jones algorithm and back projection theory, and produced 75% accuracy using the YawDD dataset [15]. Mbouna, R.O. et al. collected a real-time dataset using visual features like eye and head movements from real-time driving conditions, extracted images from the real-time videos using different subjects, and used a single camera for detecting drowsiness alert and non-alert classifiers using SVM classifiers [16]. Gu, Wang Huan, et al. used a multi-task convolutional neural network for detecting a person's face, eye, and mouth states, and multi-scale pooling (MSP-Net) was used to extract eye and mouth state regions for detecting driver fatigue from 21 volunteers [17]. Markoni H. and Guo J.M. used a publicly available drowsy ACCV driver dataset to detect drowsiness using a neural network called a hybrid convolutional neural network and long-term memory [18]. Hu, Jie, et al. analyse a person's abnormal driving behaviour by analysing the person's normalised behaviour using three characteristics: fatigue/drunkness, recklessness, and mobile phone use while driving, and following three driving techniques like a driver's driving, vehicle, and environmental conditions for assessing abnormal driving behaviour [19]. Zhao, Zuopeng, et al. developed a multitask cascading convolutional network architecture for detecting faces and facial feature points. EM-CNN was used to detect PERCLOS and mouth opening degree (POM)-based features for detecting fatigue and produced 93% accuracy [20]. Fatima, Bahjat, et al. developed micro-sleep patterns for detecting driver fatigue states and detecting faces using the Viola-Jones algorithm [21]. A person's driving video was collected by a mobile phone from six different genders, and the fatigue state was classified using two methods: support vector machines and Adaboost, which produced 96% accuracy. A. Dasgupta et al. used a smartphone-based drowsiness detection system. The first stage detects PERCLOS features through images captured by the front smartphone camera. The second stage contains a microphone for alerting the driver. The third stage contains a touch response for the driver to detect drowsiness. The Invedrifac dataset is a publicly available dataset that was used and collected with 1200 images at 640*480 at 30 frames per second using an SVM classifier and yielded 93.33% accuracy [22]. Dua, Mohit, et al. proposed four deep learning models: Alex Net for environmental changes; VGG Face Net for extracting facial features; Flow ImageNet for extracting behavioural features; and Resnet for extracting hand gestures. An ensemble algorithm was used for detecting drowsiness and produced 85% accuracy [23]. Dricare was proposed by Deng, Wanghua, and Ruoxue Wu for detecting driver fatigue states using behavioural methods like yawning, blinking, and duration of eye closure. They used MC-KCF and MTCNN algorithms to

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

track the driver's face in each video frame with different threshold values for each behavioural feature and produced 92% accuracy for two classification methods: drowsy and awake [24].
The proposed work makes a significant contribution by converting the D3S video dataset into images for predicting fatigue states. RGB images are used in the experiment. Manual annotation for two class labels, fatigue and non-fatigue, has been completed. To predict fatigue states, features such as yawning, eye opening, and eye closing are extracted from the same subject and classified and predicted using deep learning models such as the Multi-Task Convolutional Neural Network. Shape landmarks in a specific position of behavioural features are detected, and the hyper-tuning method is used to improve the accuracy of the Multi-Task Convolutional Neural Network. Finally, the model's performance for each fatigue state is assessed using the confusion matrix, precision, recall, F1score, and accuracy results.

The following steps guide the organization of the proposed paper: Step 2 lists a proposed convolutional neural network algorithm for predicting fatigue. In Step 3, the D3S datasets were used to generate a variety of experimental results for predicting fatigue. The conclusion is discussed in Step 4.

## 2. THE PROPOSED METHOD

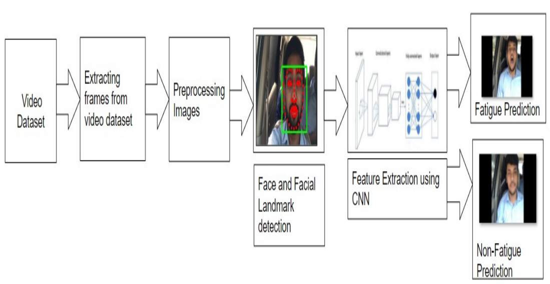



*Figure 1. Proposed Fatigue State Prediction Methodology*

In the suggested method, as indicated in Figure 1, the video-based driver drowsiness detection method is applied in the deep learning method on the D3S datasets. In the proposed model, D3S video streams of various individuals are collected with tiredness features, and images are extracted from the video streams at 15 frames per second. The D3S videos are divided into frames based on facial and behavioural characteristics. Images are manually annotated and separated into training, validation, and testing models. The experimental outcomes are evaluated and trained using fatigue images extracted from D3S video datasets. As a behavioural strategy, the non-intrusive method is suggested for forecasting drowsiness in individuals; it utilises facial motions such as yawning, eye-opening, and eye-closing. Our suggested model predicts the various variables and extracts the features using an advanced neural network, making it ideal for comparative drowsiness prediction. It involves numerous pre-processing procedures.

### 2.1 D3S Dataset using Convolutional Neural Network Algorithm

**Input:** D3S videos are extracted into images. Training and Testing based on Dlib Facial landmark position and labels.
**Output:**

1. Loading the D3S Training and Testing Dataset images
2. Initialize the CNN parameters
   Batch size at 32, Epochs at 30, Drop rate at 30% and Learning rate at 0.003
3. Adding Convolution and Hidden layer
   For convolution layer (n, 2n)
   For Filter Size (3)
   do
   a. The First convolution layers take the input (224x224x3) as Image size with 3 as its filter size and 2x2 Max pooling size.
   b. Four convolution layers with 32, 64,64 and 128 with ReLU activation function.
   c. In Hidden layers 64 and 512 number of neurons is added with ReLU activation function.
   d. Drop out is set to 30% to prevent over fitting
   End
4. Finally Fully Connected layer sigmoid function is used and produced two class outputs Fatigue and Non fatigue
5. CNN Model Performance Measures
End

Our proposed method has Seven steps:

**Step 1: Extract videos from the D3S dataset**
In the first step, the videos are extracted from the D3S dataset in mp4 format. It contains four subjects with different facial features, like yawning, PERCLOS, and eye blinking.

**Step 2: Identifying Fatigue states Images from D3S Dataset Video Frames**
In the second step, the D3S video dataset cannot train the neural network, so the images are extracted from the D3S video dataset at a resolution

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of 1920 x 1080 pixels at a time frame of 30 per second.

**Step 3: Pre-processing**
In the third step, the images are resized from 1920x1080 pixels to 640x480 pixels. The face is detected using the Dlib face detector.
**Step 4: Obtaining Coordinates for Facial Landmarks from Images**
In the fourth step, the Dlib library is used to extract facial landmark coordinates from each image and to map the x-coordinates and y-coordinates of the facial structure of the face around 68 points and detect each facial feature with shape estimation methods.
**Step 5: Extraction of Features**
The images are fed into the CNN network for detection of the eye and mouth features based on the 68 x-y facial landmark coordinates.
**Step 6: Algorithm Training**
The facial landmark coordinates are extracted from images and fed as input to the network based on the CNN classifier. The images are split into training at 70% and testing at 30% from the D3S dataset images.
**Step 7: Model Extraction**
Finally, the model is fed into the CNN layer for classification. It is classified as fatigue or non-fatigue based on the facial landmark.

## 3. METHOD

### 3.1 Driver Drowsiness Dataset (D3S)

The Driver Drowsiness Dataset (D3S) consists of four participants in order to detect driver fatigue conditions. The data collection includes both males and females of various ages. Both video data and extracted video frames are included in the dataset [9]. In the proposed method, each image from the D3S video dataset is manually annotated as either fatigued or unfatigued. The video depicts many facial traits, including eyes open, eyes closed, and yawning, as well as happy and neutral activities. Videos are 1920 x 1080 pixels at 30 frames per second and are in mp4 format without audio. Figure 2 depicts various subject images from the driver drowsiness dataset (D3S).

*Figure 2. Images of Various Subjects from the Driver Drowsiness Dataset (D3S)*

### 3.2 Detection of Face and Facial Landmarks

In the proposed method, faces are extracted from the training and testing images using the Dlib face detector and then converted into grayscale images. The face is extracted into four coordinates: top, left, right, and bottom, and a rectangular bounding box is created. Then, 68 facial landmarks are used to extract eye and mouth regions from the D3S dataset, and the Dlib shape predictor is used to locate and track the facial landmarks on the detected face in training and testing images. The recognition of 68 facial landmarks on images is highly effective under different lighting conditions [25]. Figure 3 depicts the face and facial landmarks using D3S dataset images. 36 to 45 facial points correspond to the eyes, while 48 to 60 facial points correspond to the mouth. After facial landmarks are extracted, they are fed into CNN's feature extraction.

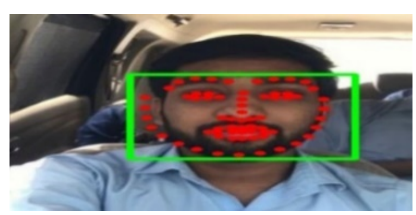


*Figure 3. Detecting Face and Facial Landmarks Using the D3S Dataset*

### 3.3 Feature Extraction

In the proposed method, a convolutional neural network model is used for feature extraction, and an image of size 222*222*3 is used as the input image. The model then learns the weights and biases of the D3S image datasets and classifies them into two classes: fatigue and non-fatigue.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The first convolution layer of a 3*3 filter with a stride1 receives 222*222*3 training images as input, and the second, third, and fourth layers contain the maximum pooling layer of a 2*2 filter with a stride2. The output of the first layer is sent to the second convolution layer, which has a size of 109*109*64, a 3*3 filter, and a stride1 value of 1. The third convolutional layer is comprised of a 52*52*64 matrix and a 3*3 filter with a stride1 parameter. The size of the fourth convolutional layer is 24*24*128 and it contains a 3*3 filter with the stride1 parameter. As shown in Figure 4, these fully connected layers generate two output layers for predicting fatigue state using two classes: class1 as fatigue and class2 as non-fatigue.

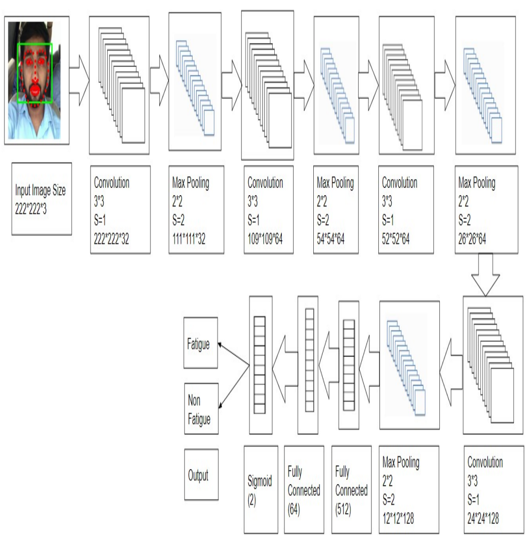



*Figure 4. Multitask Convolutional Neural Network Model on the D3S Dataset*

## 4. RESULTS AND DISCUSSION

The proposed experimental result was implemented using Python 3.6 on the Google Colab platform and an Intel Core i7 processor with a 3.00 GHz CPU and 16 GB of RAM. The CNN model was developed in the proposed method to predict fatigue states using 4628 images from the D3S dataset. To avoid overfitting, data augmentation methods such as rescaling images from 0 to 255, shearing the range by 20%, zooming the image by 20%, splitting the dataset into validation datasets by 10%, and flipping the dataset images horizontally were used. Figure 5 depicts the D3S dataset after data augmentation, which is divided into three categories: the training dataset (2916 images), the validation dataset (324 images), and the testing dataset (1388 images).

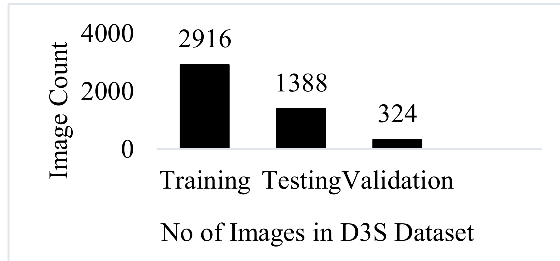



*Figure 5. Data Visualization of D3S Dataset*

The D3S dataset is divided into three sections: 60% training, 10% validation, and 30% testing. The training, validation, and testing dataset is divided into two folders, "fatigue" and "non-fatigue." The D3S training dataset is used to train an in-depth learning model based on the multi-task neural network of a conventional model, and the D3S testing dataset is used to evaluate the model's performance. Figure 6 shows a few instances of training images.

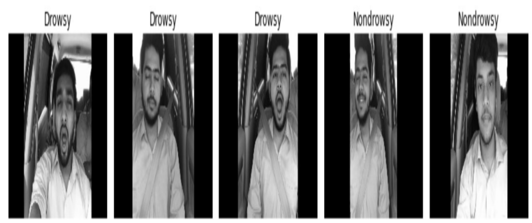



*Figure 6. Sample Training Images from D3S Dataset*

A validation dataset is used to fine-tune the parameters for the classifier. In the proposed model, training data is fed into a neural network using a deep learning framework. The Training D3S dataset contains a batch size of 32 and has a learning rate of 0.003. While fitting the D3S dataset to the deep learning model, it splits the dataset into two parts for every epoch. It trains and validates the model based on its loss and accuracy. 30 epochs are used to train all the training and validation D3S datasets once. The Adam optimization algorithm is used as an optimizer with a binary cross entropy loss function, and the loss in the validation dataset is monitored using a callback function. In the proposed method, 91 steps per 30 epochs are calculated by dividing 2916 training sample images by 32 batch sizes. The D3S dataset training and validation losses are depicted in Figure 7. The D3S dataset training and validation

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

accuracy are depicted in Figure 8. The hyperparameter optimization for building the CNN-based model is depicted in Table 1.

*Table1.Hyperparameters optimization used in CNN Model for D3S Dataset*

| | |
|---|---|
| Batch Size | 32 |
| Number of Epochs | 30 |
| Optimizer | Adam |
| Loss Function | Binary cross-entropy |
| Steps per epoch | 91 |
| Learning Rate | 0.0003 |

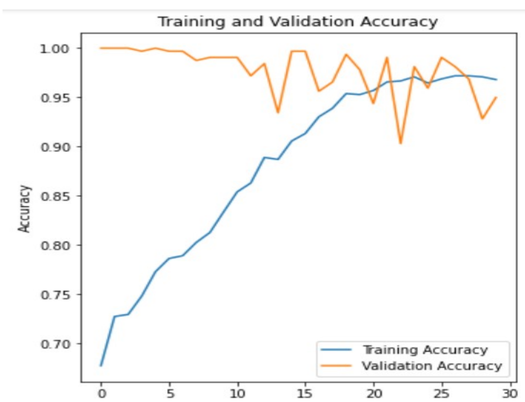



*Figure 7. D3S dataset accuracy for Training and Validation Accuracy*

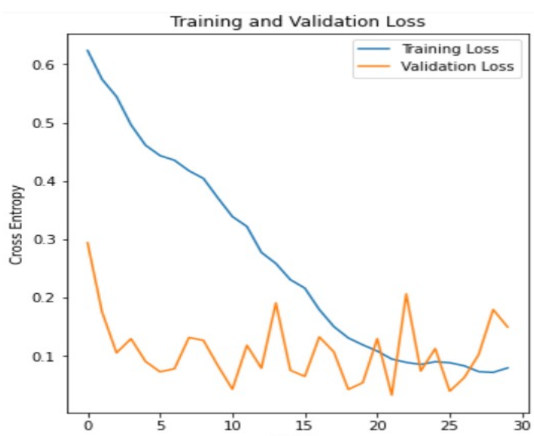



*Figure 8. D3S dataset loss cross entropy for Training and Validation Loss*

CNN's pre-trained models serve as the foundation for the proposed methodology. For both feature extraction and classification, a pretrained model is used. For classification, we have four layers in our model: convolution, max pooling, dense, and sigmoid, which are depicted in Table 2. The CNN architecture in the proposed method consists of four convolution processes with an image input size of 222x222x3, where 3 is the RGB color for the input layer and 222 and 222 are the input layer's height and width. A 3x3 Gaussian filter is used to multiply the image input size. The number of filters for each convolution layer is 9. It includes ReLU and Max-pooling as activation functions in the subsampling process. The Max-pooling function finds the maximum value to produce an extracted image by using a 2*2 matrix for each pixel and an extracted image of various sizes. A neural network uses the input vector of each pixel value as its input. To reduce overfitting issues during convolutional layers, 30% dropout is used. The output of four convolutional layers is flattened, and two dense layers are added to the hidden layer to identify the hidden features with 512 and 64.A sigmoid layer is used to generate the classification output.

*Table 2. Convolutional Neural Network Summary Model on D3S Dataset*

| Layer (Type) | Output Shape | Param # |
|---|---|---|
| Conv2d | (None,222,222,32) | 896 |
| Max_pooling2d | (None,111,111,32) | 0 |
| Conv2d_1 | (None,109,109,64) | 18496 |
| Max_pooling2d_1 | (None,54,54,64) | 0 |
| Conv2d_2 | (None,52,52,64) | 36928 |
| Max_pooling2d_2 | (None,26,26,64) | 0 |
| Conv2d_3 | (None,24,24,128) | 73856 |
| Max_pooling2d_3 | (None,12,12,128) | 0 |
| Dropout | (None,12,12,128) | 0 |
| Flatten | (None,18432) | 0 |
| Dense | (None,512) | 9437696 |
| Dense_1 | (None,64) | 32832 |
| Dense_2 | (None,1) | 65 |
| Activation | (None,1) | 0 |
| Total Params: | 9,600,769 | |
| Trainable Params: | 9,600,769 | |
| Non-Trainable Params: | 0 | |

A confusion matrix is used to predict the performance level of an organized model based on the number of correct and incorrect predictions. In the proposed model, 723 correctly predicted that the subject is fatigued, while 595 correctly predicted that the subject is not fatigued, resulting in 50 incorrect predictions of fatigue and 20 incorrect predictions of non-fatigue The confusion matrix for predicting fatigue states is depicted in Table 3. The testing accuracy for the proposed D3S dataset is 95%.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

| True Positive | True Negative | Accuracy | Specificity | Precision | F1 Score | Recall |
|---|---|---|---|---|---|---|
| 723 | 595 | 97% | 92% | 97% | 94% | 93% |

| Confusion Matrix | Predicted Fatigue | Predicted Non-Fatigue |
|---|---|---|
| Actual Fatigue | 723 | 20 |
| Actual Non-Fatigue | 50 | 595 |

*Table 3. Testing of D3S Dataset Confusion Matrix*

In the proposed method, various performance-based evaluation metrics like precision, recall, specificity, F1score, and accuracy are calculated based on the following equations from equation (1) to equation (5)

Recall is known for finding correct positive results from the sum of true and false positive predictions.

$$Recall = (X)/(X + Y1) \quad (1)$$

The number of negative samples predicted by the classifier is referred to as specificity.

$$Specificity = \frac{Y}{Y+X1} \quad (2)$$

Precision is defined as the ratio of true positive results to the sum of true positive predictions.

$$Precision = \frac{X}{X+X1} \quad (3)$$

An accuracy is known as a "true prediction," which determines both positive and negative values out of the total number of samples used for testing the CNN model.

$$Accuracy = \frac{X+Y}{X+Y+X1+Y1} \quad (4)$$

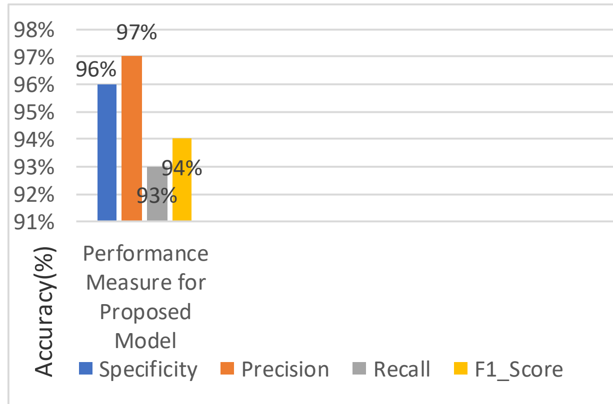
Where X= True Positive, Y=True Negative, X1=False Positive, Y1=False Negative

F1 Score is defined to evaluate a test's accuracy and also measure the harmonic mean between recall and precision.

$$F1\ Score = \frac{2*(Precision*Recall)}{Precision+Recall} \quad (5)$$

Experimental results based on different performance measures for the proposed model are depicted in Table 4 and Figure 9.

*Table 4. Experimental Results using proposed model*
*Figure 9. Performance based different Measures for proposed model*




Experimental results of the proposed D3S dataset-based deep learning model are compared with existing methods based on accuracy and classifier. The D3S dataset is compared with the existing classifiers depicted in Table 5. Compared to existing methods, our proposed model using CNN with different layers produced 97% accuracy.

*Table 5. Performance comparison of Fatigue state classification on D3S Dataset*

| Reference No | Dataset | Classifier | Accuracy |
|---|---|---|---|
| Gupta, Isha, et.al [12] | D3S Dataset | Support Vector Machine | 90% |
| Proposed Method | D3S Dataset | Convolutional Neural Network | 97% |

## 5.CONCLUSION

The proposed deep learning-based multi-task convolutional neural network model is predicted with different features of a subject like yawning, eye-opening, eye-closing, and talking from the publicly available D3S dataset to detect fatigue states. The proposed model is based on behavioural features, and facial landmarks are used to extract deep features from the driver's face. When compared to existing models, the proposed model uses driver videos to extract images, and manual annotation s done to predict driver fatigue states. Compared to the existing machine learning classifier, deep learning produced 97% accuracy with various performance-based evaluation metrics like recall, precision, sensitivity, and specificity. Subjects wearing spectacles may be considered in future work for detecting fatigue states.

**REFERENCES:**

[1] Priya VV, Uma M. “Comparison of Multimodal Features for Detecting Driver Drowsiness using different dataset”, In2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) 2021 Jun 3 (pp. 1429-1437). IEEE. doi: 10.1109/ICOEI51242.2021.9452887.

[2] T. Johnson, 2019 Traffic Safety Culture Index Accessed: August 23, 2022, [Online]. Available: https://aaafoundation.org/2019-traffic- safety-culture-index.

[3] AAA Foundation for Traffic Safety Accessed: December 7, 2022, [Online]. Available: https://aaafoundation.org/2019-traffic-safety-culture-index.

[4] Lin CT, Chuang CH, Hung YC, Fang CN, Wu D, Wang YK. “A driving performance forecasting system based on brain dynamic state analysis using 4-D convolutional neural networks”, IEEE transactions on cybernetics. 2020 Aug 20;51(10):4959-67 doi: 10.1109/TCYB.2020.3010805

[5] Murugan, S., Selvaraj, J. and Sahayadhas, A., 2020. “Detection and analysis: driver state with electrocardiogram (ECG)”. *Physical and engineering sciences in medicine*, *43*(2), pp.525-537. doi: https://doi.org/10.1007/s13246-020-00853-8

[6] Koesdwiady, Arief, et al. "Recent trends in driver safety monitoring systems: State of the art and challenges." IEEE transactions on vehicular technology 66.6 (2016): 4550-4563. doi: 10.1109/TVT.2016.2631604.

[7] Fujiwara K, Abe E, Kamata K, Nakayama C, Suzuki Y, Yamakawa T, Hiraoka T, Kano M, Sumi Y, Masuda F, Matsuo M. “Heart rate variability-based driver drowsiness detection and its validation with EEG”. IEEE Transactions on Biomedical Engineering. 2018 Nov 2;66(6):1769-78. doi: 10.1109/TBME.2018.2879346.

[8] Jo, Jaeik, Sung Joo Lee, Kang Ryoung Park, Ig-Jae Kim, and Jaihie Kim. "Detecting driver drowsiness using feature-level fusion and user-specific classification." *Expert Systems with Applications* 41, no. 4 (2014): 1139-1152 doi: https://doi.org/10.1016/j.eswa.2013.07.108.

[9] Lin, Chin-Teng, Chun-Hsiang Chuang, Yu-Chia Hung, Chieh-Ning Fang, Dongrui Wu, and Yu-Kai Wang. "A driving performance forecasting system based on brain dynamic state analysis using 4-D convolutional neural networks." *IEEE transactions on cybernetics* 51, no. 10 (2020): 4959-4967. doi: 10.1109/TCYB.2020.3010805

[10] Savaş BK, Becerikli Y. “Real time driver fatigue detection system based on multi-task ConNN”. Ieee Access. 2020 Jan 3; 8:12491-8. doi:10.1109/ACCESS.2020.2963960.

[11] Ji Y, Wang S, Zhao Y, Wei J, Lu Y. “Fatigue state detection based on multi-index fusion and state recognition network”. IEEE Access. 2019 May 16;7: 64136-47. doi: 10.1109/ACCESS.2019.2917382

[12] Gupta I, Garg N, Aggarwal A, Nepalia N, Verma B. “Real-time driver's drowsiness monitoring based on dynamically varying threshold”. In 2018 Eleventh International Conference on Contemporary Computing (IC3) 2018 Aug 2 (pp. 1-6). IEEE. doi: 10.1109/IC3.2018.8530651.

[13] Saurav S, Mathur S, Sang I, Prasad SS, Singh S. “Yawn detection for driver’s drowsiness prediction using bi-directional LSTM with CNN features”. In International Conference on Intelligent Human Computer Interaction 2019 Dec 12 (pp. 189-200). Springer, Cham. doi: https://doi.org/10.1007/978-3-030-44689-5_17.

[14] Bai J, Yu W, Xiao Z, Havyarimana V, Regan AC, Jiang H, Jiao L. “Two-stream spatial-temporal graph convolutional networks for driver drowsiness detection”. IEEE Transactions on Cybernetics. 2021 Oct 4.doi: 10.1109/TCYB.2021.3110813

[15] Omidyeganeh M, Shirmohammadi S, Abtahi S, Khurshid A, Farhan M, Scharcanski J, Hariri B, Laroche D, Martel L. “Yawning detection using embedded smart cameras”. IEEE Transactions on Instrumentation and Measurement. 2016 Jan 7;65(3):570-82doi: 10.1109/TIM.2015.2507378.

[16] Mbouna RO, Kong SG, Chun MG. “Visual analysis of eye state and head pose for driver alertness monitoring”. IEEE transactions on intelligent transportation systems. 2013 May 22;14(3):1462-9. doi: 10.1109/TITS.2013.2262098.

[17] Gu WH, Zhu Y, Chen XD, He LF, Zheng BB. “Hierarchical CNN-based real-time fatigue detection system by visual-based technologies using MSP model”. IET Image Processing.

2018 Dec;12(12):2319-29. doi: https://doi.org/10.1049/iet-ipr.2018.5245

[18] Guo JM, Markoni H. “Driver drowsiness detection using hybrid convolutional neural network and long short-term memory”. Multimedia tools and applications. 2019 Oct;78(20):29059-87 doi: https://doi.org/10.1007/s11042-018-6378-6.

[19] Hu J, Xu L, He X, Meng W. “Abnormal driving detection based on normalized driving behavior”. IEEE Transactions on Vehicular Technology. 2017 Jan 27;66(8):6645-52. doi: 10.1109/TVT.2017.2660497.

[20] Zhao Z, Zhou N, Zhang L, Yan H, Xu Y, Zhang Z. “Driver fatigue detection based on convolutional neural networks using EM-CNN”. Computational intelligence and neuroscience. 2020 Nov 18;2020. doi: https://doi.org/10.1155/2020/7251280.

[21] Fatima B, Shahid AR, Ziauddin S, Safi AA, Ramzan H. “Driver fatigue detection using viola jones and principal component analysis”. Applied Artificial Intelligence. 2020 May 11;34(6):456-83. doi: https://doi.org/10.1080/08839514.2020.1723875.

[22] Dasgupta A, Rahman D, Routray A. “A smartphone-based drowsiness detection and warning system for automotive drivers”. IEEE transactions on intelligent transportation systems. 2018 Dec 28;20(11):4045-54. doi: 10.1109/TITS.2018.2879609.

[23] Dua M, Singla R, Raj S, Jangra A. “Deep CNN models-based ensemble approach to driver drowsiness detection”. Neural Computing and Applications. 2021 Apr;33(8):3155-68 doi: https://doi.org/10.1007/s00521-020-05209-7.

[24] Deng W, Wu R. “Real-time driver-drowsiness detection system using facial features”. Ieee Access. 2019 Aug 21; 7:118727-38. doi: 10.1109/ACCESS.2019.2936663.

[25] Liu, Hongzhe, Weicheng Zheng, Cheng Xu, Teng Liu, and Min Zuo. "Facial Landmark Detection Using Generative Adversarial Network Combined with Autoencoder for Occlusion." *Mathematical Problems in Engineering* 2020 (2020). doi: https://doi.org/10.1155/2020/4589260

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195