IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 24 April 2023, accepted 9 June 2023, date of publication 20 June 2023, date of current version 28 June 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3288008

RESEARCH ARTICLE

Facial Feature-Based Drowsiness Detection With Multi-Scale Convolutional Neural Network

V. VIJAYPRIYA AND MOHAN UMA
Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur 603203, India
Corresponding author: Mohan Uma (umam@srmist.edu.in)

ABSTRACT Recently, the upsurge in accidents is caused due to driver drowsiness arises due to lack of sleep, fatigue and other health factors. The drowsiness leads to mortality, loss of properties and serious health conditions. Hence, it is necessary to prevent accidents caused by drowsiness in drivers. At present, the automated model is effective for detection drowsiness and recognition. In this research paper, developed a MCNN (Multi-Scale Convolutional Neural Network) framework for the classification of drowsiness. Initially, the YAWDD dataset and NTHU-DDD dataset is utilized for acquiring video sequences about driving. The acquired video is converted into frames for the keyframe extraction and selection. With the Dlib library face recognition is utilized for the localization of facial points in the extracted frames. The extracted image frames are pre-processed with the Cross Guided Bilateral Filtering followed by the feature extraction with hybrid dual-tree complex wavelet transforms with Walsh-Hadamard transform for the extraction of the feature vector in the image frame blocks. The feature points are optimized with the Flamingo search algorithm (FSA) integrated with the deep learning model Multiscale convolutional neural network (MCNN). With the proposed method, MCNN based FSA model feature extraction drowsy and non-drowsy are classified. The proposed simulation results illustrated the MCNN with FSA model attains an accuracy value of around 98.38% for YAWDD dataset and NTHU-DDD model exhibits an accuracy value of 98.26%. The performance of the proposed MCNN model exhibits approximately 6% higher accuracy than the conventional state-of-art methods.

INDEX TERMS Accuracy, drowsiness, deep learning, feature extraction, optimization, pre-processing.

I. INTRODUCTION

In recent years, road accidents cause huge fatalities and injuries due to the motorization surge with an increase in population and penetration of vehicle based on length and breadth [1]. Globally injuries from road accidents are the leading death cause in between the age group of around 15 – 49 years. In the year 2020, road accidents in India reach 1.3lakhs accidents and causes injuries to more than 3.4 lakh people. However, the accident number decreases fatalities in 2020 compared with 2019. The reduction of fatalities due to road accidents is better management of traffic, implementation of the new Motor Vehicle Act and COVID-19 lockdown. In the year 2020, the reduction of road accident counts by 18.46% with decrease in the average injury rate of 22.84% and the total death rate is observed as 12.84% [2]. Table 1 presented the road accident instances between the year 2019–2020.

TABLE 1. Road accidents recorded in 2019-2020.

Parameter	2019	2020	% Change
Number of Injury	4,51,361	3,48,279	-22.8
Number of Deaths	1,51,113	1,31,714	-12.8
Number of Accidents	4,49,002	3,66,138	-18.5
Severity of Accident (Number of person death/ 100 accidents)	33.7	36.0	

In the case of State and Union Territories (UTs) road accident claim the injuries rate of 3,48,279 and death rate of 1,31,714 lives. In the year 2020, the road accident count is reduced by 18.5% on average compared with the year 2019. On the whole, the count of deaths in road accidents are

The associate editor coordinating the review of this manuscript and approving it for publication was Rajeeb Dey.

reduced to 12.8% from 22.8% as in Table 1. In Tamilnadu state government accounts for the total accident of 45,484 rates of 12.4% in the year 2020. The number of accidents in the Uttar Pradesh state accounts count of 19,149 with 14.5% [3]. The number of road accidents category in Road for 2020, in National Highways is observed as 1,37,191 accidents presented in Table 2. The total accidents on the road are reduced by 15.1% with the previous year. The National Highway person killed in 2019 is observed as 53,872 and for 2020 it is observed as 47,984. The overall accident percentage growth rate is minimized by 10.9%. In 2020, the State Highway reported a minimal number of accidents and fatalities as in Table 2. In the National Highway, Tamilnadu accident accounts 15,269 and Uttar Pradesh count 13,695 and Maharashtra 3,528. The state-wise activities on the National Highways are estimated between 2019 – 2020 as in Table 2.

TABLE 2. Total highways road accidents recorded in 2019 – 2020.

	2019 Record		2020 Record	
Category of Road	Accidents	Death	Accidents	Death
National Highways	1,37,191	53,872	1,16,496	47,984
State Highways	1,08,976	38,472	90,755	33,148
Other Roads	2,02,835	58,769	1,58,887	50,582
All India	4,49,002	1,51,113	3,66,138	1,31,714

Recently, accidents are increases drastically due to several factors that drowsiness of drivers is the leading cause of the accident [4]. The drowsiness of drivers is caused due to low heart rate, alcohol intake, sleep deprivation, depression, stress, improper medication and anxiety [5]. The drowsiness causes frequent yawning, eye blinking, heavy eyelids and abnormalities. Additionally, cardiac abnormalities and lane position are also common characteristics of drowsiness. Even though some drivers neglect those symptoms and keep driving those leads to accidents [6]. Accident due to drowsiness not only affects the drivers, exhibits severe consequences in the families. Conventionally, drowsiness is measured based on driving behaviours, biological and vehicle measures [7]. Video processing base driving pattern estimation comprises the examination of body postures, eye-opening and blinking, yawning detection and detection of head gaze movement. With an effective video processing technique, it is easy to evaluate the drivers face and eye movement through which it can be evaluated to examine the driving condition of drivers [8]. In those system, video or images are captured with the use of a camera to evaluate the variation in the driver behaviour.

The behaviours-based drowsiness detection system uses the image or video processing approach [9]. Video or image processing system evaluate the rate of eyelid open and closure duration and movement of head, eyelid blinking and yawning rate [10]. In biological measure process medical signals is utilized are Electrocardiogram (ECG), Electrooculogram (EOG), Electroencephalogram (EEG) and temperature [11]. In those methods, the physiological activities need to be monitored and reported continuously. Finally, with the vehicle-based measurement comprises of the calculation of the lane deviation and the measurement of vehicle pressure while using breaks and measurement of steering angle deviation [12].

Conventionally, facial feature detection is operative for the driver drowsiness detection [13]. The comprehensive technique comprises of the simple methods for the driver drowsiness detection. As drowsiness are state lies between wakefulness and sleep it can be evaluated based on the reaction, behaviour changes, reduction of reflex and difficulty in placing frontal position of the head in the vision field [14]. According to this scenario, behaviour-based drowsiness detection will be effective in the form of video and biomedical signals. As the behaviour model is effective indicator for the detection of drowsiness related to the video features [15]. Over a decade Machine Learning (ML) based technique are implemented for drowsiness detection tasks to perform classification [16]. The limitation associated with the ML (Machine Learning) requires massive range of datasets for predicting and evaluating the performance of the anomaly's detection [17].

Additionally, deep learning (DL) model exhibits the noticeable progress in the classification-based detection model [18]. However, the DL is the sub-field technology of ML it exhibits promising advancement in the simplification of the raw data. The process of DL comprises feature extraction, feature selection and denoising to achieve effective classification results [19]. The DL technique utilized for the classification in the field of DSNs (deep stacking networks), LSTM (Long Short-Term Memory), CNNs (convolutional neural networks), AN (Auto-encoder), RNNs (Recurrent Neural networks), etc [20]. To increase the classification accuracy the DL approach is implemented for accurate classification performance.

This paper concentrated on the design of an effective drowsy detection model with the use of artificial intelligence. The present research uses the YAWDD and NTHU-DDD datasets for the classification of Drowsy and Non-drowsy. The specific contribution of the present work is

1. Initially, the YAWDD and NTHU-DDD dataset is acquired as the video sequence. The acquired video is converted into frame.
2. The frames are utilized for face localization with the Dlib face recognition system for landmark estimation.
3. Through the processed face recognition system pre-processing is performed with Cross- Guided Bilateral Filtering. With the processed image feature

extraction is performed with hybrid dual-tree complex wavelet transforms with Walsh-Hadamard for the conversion of feature vector into image frames.

4. The extracted features are optimized with the FSA for the features and deep learning model is utilized for the classification and detection of drowsiness in 2-classes of Drowsy and Non-drowsy.
5. The classification is implemented in Python software and the analysis is evaluated comparatively with the machine learning classifier along with the pre-trained deep learning model. The comparative analysis observed that the machine learning model and deep learning performance is higher for both YAWDD and NTHU-DDD dataset. With YAWDD proposed MCNN model achieves an accuracy value of 98.38% and NTHU-DDD provides an accuracy value of 98.26%.

This research is systematized as follows: Section II provides the related works of the drowsiness detection system and overall research methodology adopted and presented in section III. The simulation results for YAWDD and NTHU-DDD datasets is obtainable in section IV and the overall conclusion is presented in section V.

II. RELATED WORK

Over the decades, a video-based feature detection model is utilized for tracking in different applications. With the Personal Computer Vision Strategy, the accuracy of the performance is improved through the webcams for facial feature detection and tracking. The existing literature on the detection of drowsiness is presented as follows:

Arefnezhad et al. [21] developed a non-obtrusive drowsiness detection system with a deep learning neural network based on vehicle measurement. The proposed deep learning model integrates the RNN (Recurrent Neural network) and CNN (Convolutional Neural network). The examination is performed with the consideration of network inputs such as wheel velocity, lateral deviation, acceleration, wheel angle, yaw rate, and road centreline. The drowsiness level is measured with the consideration of LSTM (Long-short term memory) and GRU (Gated recurrent unit) layers are used as RNN within the deep learning network. Therefore, the proposed model evaluates around 44 sessions of the experimental data with the fixed driving simulator. The simulation results demonstrated the classification accuracy of the deep networks is higher compared with the conventional classifier such as k-nearest neighbours and support vector machine. The proposed model achieves a higher accuracy of 96.0%. Bekhouche et al. [22] constructed a transfer learning-based framework model for the deep feature extraction from the image of the driver faces. The model uses the DCNN (deep convolutional neural network) with the pre-trained model with the facial recognition dataset. The developed mode uses the sliding temporal window frame sample with the 9 frames. The features extracted are converted in the observation matrix with the aggregation of the temporal features for the construction of raw feature vectors. The examination is conducted based on the NTHU-DDD (NTHU Drowsy Driver Detection) video dataset. The simulation results demonstrated that the proposed model exhibits outperformance for the other art methods with an accuracy of 97.45%.

Cui et al. [23] developed an interpretable Convolutional Neural Network (CNN) with an examination of EEG features to detect drowsiness in drivers. The model performance is based on the local region estimation in the input signal evaluated with the GAP (Global Average Pooling) with CAM (Class Activation Map) method to perform classification. The simulation output analysis expressed that the enhanced model achieves an accuracy of around 73.22% for the 2-class cross-subject in the classification. It is stated that compared with the EEG signal-based method visualization technique is effective. Additionally, the research expressed that CNN is an effective tool for the examination of the mental state of the drivers. Moujahid et al. [10] constructed a PML (Pyramid-Multi Level) framework for face detection, representing the feature extraction and classification. The constructed framework model is evaluated for the NTHUDDD (NTH Drowsy Driver Detection). The analysis expressed that with examination of face descriptors the proposed method achieves significant performance for drowsiness detection through deep Convolutional Neural Networks. Zhang et al. [24] established a model to compute the drowsiness level of the drivers measured with the time cumulative effect (TCE). The evaluation is conducted with the dataset of 27 drivers evaluated with the eye-tracking system for the driving simulation. Through Karolinska Sleepiness Scale (KSS) records the perception of the drivers about drowsiness. The drowsiness level is increased with time and the accumulation of data is achieved with the MOL (Mixed-effect ordered logit) model. Henceforth, the experimental results demonstrated that with an increase in drowsiness level the standard deviation of lateral position and PERCLOS (percentage of driver eyelid closure) also increased rapidly. The MOL-TCE model provides an accuracy of 62.84% while the MGOR-TCE accuracy is measured as 61.04% and the non-TCE MOL model accuracy is measured as 52.47%.

Quddus et al. [12] proposed model to detect the eye movement associated with drowsiness without the use of an eye-tracking system. The proposed model uses the RNN (Recurrent Neural network) model integrated with LSTM (Long Short-Term Memory) class is utilized for the eye movement detection. The model uses the 2 LSTM classes with the 1-D LSTM (R-LSTM) and C-LSTM (convolutional LSTM) for the extraction with the patch size of 48×48 for the 38 subjects. The simulation results demonstrated that R-LSTM achieves an accuracy of 82% and C-LSTM model achieves the accuracy range of 95 – 97%. Through examination, it can be stated that the LSTM technique is effective for the detection of the drowsiness level. Pandey and Muppalaneni [25] developed a LSTM model for the computer vision model for the examination of realistic drowsiness detection in the video dataset. A model is performed with the two characteristics along with spatiotemporal characteristics for

Model-A and Model-B. The Model-A uses the YOLOv3 with LSTM to retrieve the temporal characteristics and Model-B uses the CNN Inception V3 model for the evaluation of each condition. With the developed model the overfitting problems in the frames are discarded with the use of TransGAN's augmentation. Model -B uses the CNN based Inception V3 model for drowsiness detection. Model-A achieves the accuracy of 86% and Model-B achieves an accuracy of 97.5%. Through comparative examination, CNN based Inception V3 is effective compared with the YOLO V3 integrated with LSTM.

Chinara [26] constructed a model for the drowsiness detection (DD) with the EEG based signal with the utilization of the Wavelet packet transform (WPT) along with the consideration of the time-domain factors. The WPT is employed with the Virtual Driving Driver for drowsiness detection. The experimental analysis expressed that SVDD model achieves an accuracy of 94.45% and 84.3% respectively. Therefore, the experimental output results demonstrated that the enhanced model attains significant accuracy for the stated physiological dataset for the time-domain feature selection method with consideration of 66 features. Lee and An. [27] adopts CNN model with the four-feature based LST approach for the detection of drowsiness with the consideration of feature-based model with six different levels. The model uses the CNN integrated with LSTM measured with erratically assigned auditory stimuli. Therefore, the EEG data is labelled in three different classes such as drowsiness, sleep and awake with the input vector size of 8 sec. The LSTM model achieves an accuracy of 86% for the window length of 1 sec and the kappa index is measured as 0.77 for the 4 sec window length. With the application of binary classification model, the LSTM-CNN model achieves the F-score of 0.95 with an average accuracy of 85.6% and kappa index 0.77. Shen et al. [28] utilized the EEG signal for the classification of drowsiness in the person to perform feature extraction and classification. The proposed model is stated as the MSSA (Multi-source signal alignment) with the multi-dimensional feature classification. The examination is based on the covariance matrices-based computation for the multi-dimensional feature extraction and classification with the tensor network (TN). The evaluation is considered for the EEG dataset for drowsiness detection. The analysis stated that classification accuracy is achieved as 86.7% which is 3.71% higher than the state-of-art methods.

Phan et al. [29] proposed an IoT based deep learning mode for the detection of drowsiness in drivers using LSTM, VGG16, InceptionV3, and Dense Net. The examination is based on the utilization of the transfer learning technique to detect the drowsiness of driver under different conditions. The model accounts for the time-varying factors for the detection of fatigue with the warning messages with the detection of sounds through the Jetson Nano monitoring system. The simulation analysis stated that the proposed deep learning model achieves a higher accuracy of 98%. Bekhouche et al. [22] constructed a framework with a transfer learning process to extract the feature for the detection of driver's face. The model uses the facial recognition dataset with the pre-trained deep convolutional network model for the extraction of deep learning features. The model uses the sliding temporal window with the observation matrix for the aggregation of the temporal features in the raw vector. With the use of NTHU-DDD (NTHU Drowsy Driver Detection) video dataset performance of the binary classifier is related with the conventional art methods. Fouad [30] evaluated the EEG signal with the consideration of 12 subjects to evaluate the recorded data. The data were labelled based on the consideration of labelled and drowsy epochs. The labelled data about brain signals are pre-processed based on the relevant feature extraction based on consideration of KNN (K-Nearest Neighbour), Random Forest, Support vector machine and Naive Bayes. The accuracy of classification is observed as 100% with minimal loss in the driver's data.

Chen et al. [31] constructed an approach for feature selection and classification with the deep learning based ConvNets (convolutional neural network) for the detection of drowsiness in drivers through multi-channel EEG signal. The developed model comprises the 12-layer deep ConvNet model for the learning and extraction of features in raw EEG data with the use of 5 convolutional layer, 3 max pooling and 1 pooling layer those are optimized for the classification. The experimental analysis stated that the proposed model attains an accuracy of 97.02% $\pm$ 0.0177, precision of 96.74% $\pm$ 0.0347, sensitivity of 97.76% $\pm$ 0.0168, specificity of 96.22% $\pm$ 0.0426, and mean f-measure of 97.19% $\pm$ 0.0157, respectively. Rajamohana et al. [32] constructed a hybrid model for the Convolutional Neural Network (CNN) integrated with the Bidirectional Long-Term Dependencies (BiLSTM) for the detection of drowsiness in drivers. Initially, with the proposed model facial images are tracked in the blink of eyes with consideration of the main phases. The phase of drivers is estimated and identified with the Identified drivers face with the use of Web camera. While in second phase, the features are extracted based on Euclidean distance based on continuous monitoring. Kamath and Renuka [33] reviewed 167 studies with consideration of systematic kind of literatures for the estimation of visualization for the identification of gaps, possibilities and future perspectives in research. Quddus et al [12] developed an appropriate eye-tracking system for the detection of drowsiness detection RNN (Recurrent Neural Network) and LSTM (Long Short Term Memory) model. With the employed 1-D LSTM based model 2-D images are processed directly with the extraction of 38 subjects in the 48 $\times$ 48 pixels. The experimental analysis stated that the proposed model with electroencephalogram (EEG) signals achieves an accuracy of 82% R-LSTM and C-LSTM achieved an accuracy value of 95%–97%. Jabbar et al. [34] proposed a model for the real-time drowsiness detection with deep learning model, The analysis is based on the utilization of Android application with the recognition of facial landmarks with an accuracy value of more than 80%. Pandey and Muppalaneni. [25] designed a model based on the spatio temporal feature estimation is based on the computer vision technique. The features are processed with the YOLOv3 detector with the TransGAN augmentation model. The spatial information is extracted based on the Inception V3 with a minimal accuracy level of 86% with the model – B accuracy of 97.5%. Vijayan and Pushpalatha [35] classified the anomalies with the consideration of deep learning model with the detection. Through the incorporation of the autoencoder based Long Short Memory for the detection of abnormal points based on the intervals. The evaluation is based on the consideration of relational feature with consideration of time series data in the intervals. The anomalies explain ability is increased for the detection of an anomaly.

He et al. [36] comparatively examined the feature extraction with the RootSIFT process with consideration of Scale Invariant Feature Transform (SIFT). The examination is based up-on the consideration of RootSOFT and SIFT descriptors for the feature extraction with consideration of features such as translation, rotation, scale. The model RootSIFT accuracy is achieved at the 93.55% for the extraction of drowsy detection. Escorcia-Gutierrez et al. [37] proposed an Automated Deep Learning-Enabled Brain Signal Classification for Epileptic Seizure Detection (ADLBSC-ESD). With the Proposed ADLBSC-ESD model the feature in EEG signal is selected with the Improved Teaching and Learning-Enabled Optimization (ITLBO). The Deep Belief Network (DBN) uses the EEG based signal classification model with the hyperparameters based optimal tuning with the Swallow Swarm Optimization Algorithm (SSA). With the proposed ADLBSC-ESD scheme the brain signal is classified based on the consideration of different metrices. The experimental analysis states that ADLBSC-ESD scheme achieves an accuracy of 0.8316 for the multiple classes.

Chinara [26] developed a model for drowsiness detection with the use of Wavelet Packet Transform (WPT). The analysis is based on the consideration of time-domain features in the EEG channels. The feature selection is performed with the detection of drowsiness. The simulation results demonstrated that Simulated Virtual Driving Driver (SVDD) achieves an accuracy of 94.45%. Khare and Bajaj [38] developed a Tunable Q wavelet transform (TQWT) decomposes the EEG signal. The model comprises of the optimization of the EEG features with O-TQWT with an overall accuracy value of 96.14%. The overall summary of the literature is tabulated in Table 3.

TABLE 3. Existing literature are summarised.

Reference	Method	Dataset	Accuracy
[21]	Long-short-term-memory (LSTM) with CNN and RNN	44 experimental data	96 %
[22]	Transfer learning	NTHU-DDD	97.45%
[23]	Global Average Pooling (GAP) with Class Activation Map (CAM)	EEG Signal	73.22%
[10]	Pyramid-Multi Level (PML)	NTHU-DDD	86.57%
[24]	Time cumulative effect (TCE) to measure the drowsiness scale	Eyelid closure scale	62.84%
[12]	Convolutional LSTM	Eye movement detection	96%
[25]	Model A - YOLOv3 with LSTM Mode B - CNN Inception V3 model	Driving Video	Model A – 86% Model B – 97.5%
[26]	Wavelet packet transform (WPT)	Virtual Driving Driver	94.45%
[27]	CNN integrated with LSTM	EEG data	85.6%
[28]	Multi-source signal alignment (MSSA)	EEG data	86.7%
[29]	LSTM, VGG16, InceptionV3	IoT data	98%
[22]	Transfer learning	NTHU-DDD	77%
[30]	CNN	EEG	93%
[31]	ConvNets	EEG Signal	97.02 % ± 0.0177
[12]	Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM)	Eye Tracking Data	95%–97%
[25]	YOLOv3 detector TransGAN	Spatio Temporal	97.5%
[37]	Automated Deep Learning-Enabled Brain Signal Classification for Epileptic Seizure Detection (ADLBSC-ESD)	EEG	83%
[26]	Wavelet Packet Transform (WPT)	EEG	94.45%
[38]	Tunable Q wavelet transform (TQWT)	EEG	96.14%

III. PROPOSED METHODOLOGY

The process involved in the drowsiness detection in drivers through the utilization of a drowsiness dataset. The complete process is presented in Figure 1 for the detection of drowsiness in drivers to prevent accidents. Initially, the video is converted into frames based on key frame extraction and selection process. With the key frame selection process facial features are recognized through the use of Dlib library. Secondly, pre-processing is performed for the elimination of noises followed by the optimization based Multiscale Convolutional neural network (MCNN). Finally, with the training and testing images the data was analyzed.

The steps adopted for the detection of drowsiness are explained as follows:

A. DATASET

To train the model, this research selects the two datasets such as YAWDD- Yawning Detection Dataset) [39] and NTHU-DDD (NTHU Drowsy Driver Detection) dataset [40]. The YAWDD model comprises of the video recorded through the in-car cameras and driver's dash. The YAWDD dataset is utilized for yawning detection and face or mouth tracking

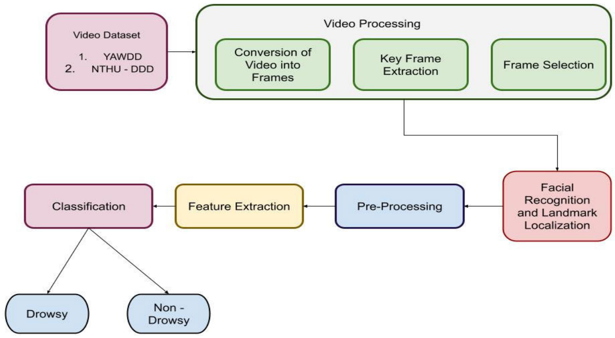

FIGURE 1. Proposed method for drowsiness detection.

through variations of illumination. The NTHU-DDD dataset is generated in NTHU Computer Vision Lab with 36 subjects under different scenarios with an entire duration of 9 and half hours. The attributes for the YAWDD dataset are presented in Table 4.

TABLE 4. Attributes of various drowsiness dataset.

YAWDD			
Scenario	**Number of Videos**	**Situations**	
Camera under front mirror	322	1. Normal Driving 2. Talking or Singing 3. Yawning	
Camera in drivers dash	29	1. Driving Silently 2. Taking or Singing 3. Yawning	
NTHU-DDD			
		Drowsy	Non-Drowsy
Bare Face	90	yawning, nodding, slow blink rate)	talking, laughing, looking at both sides
Glasses			
Night_ Bare Face			
Night Glasses			
Sunglasses			

B. KEY FRAME EXTRACTION AND SELECTION

The selected dataset video sequences are converted into frames. After the conversion of the video into frames key frames are extracted and selected. Those key frames provide the overall summary of the frames in the video sequences. As the key frames are selected in the frames the facial landmarks such as eyes and mouths are identified.

C. PREPROCESSING

Pre-processing is essential step for the accurate estimation of results to perform the learning process. This paper uses Cross Guided Bilateral filtering and Hybrid dual tree complex wavelet transformation.

D. FEATURE EXTRACTION

With the feature extraction model the raw data is converted into numerical features to preserve the information about the original frames. With the utilization of hybrid dual-tree complex wavelet transforms with Walsh-Hadamard transform the features are extracted in the images. The extracted features are optimized and perform the classification to detect the drowsiness in the drivers.

E. OPTIMIZED FSA MCNN

The pre-processed feature images are optimized with the Flamingo search algorithm and MCNN for the detection of drowsy and non-drowsy. The optimal point features for the detection of drowsiness in drivers are estimated with the FSA applied over the MCNN for the classification of Drowsy and Non-drowsy. The FSA comprises of three features with the MCNN based classification. The MCNN architecture model comprises the convolution, max-pooling and classification layers.

1) PREPROCESSING

Pre-processing is a key step in image processing as it eliminates unwanted distortion or improves the features in the image for further processing. The frames extracted from the dataset are utilized for the key frame selection. Through the key frame selection facial landmarks are estimated. The pre-processing is performed with the cross guided bilateral filter.

a: CROSS GUIDED BILATERAL FILTER

The bilateral filter is an extension of the linear smoothing model computed with the image weights w_p with the spatial weight of w_s for the original image E stated as in equation (1)

$$F(x) = \frac{\sum_{t \in S_m} w_s(\|t\|)\, w_p(E(x) - E(x+t)\, E(x+t))}{\sum_{t \in S_m} w_s(\|t\|)\, w_p(E(x) - E(x+t))} \tag{1}$$

where, for the square window [−m, m] × [−m, m] it is measured as S_m for the images. The distance between the centre point of image S_m to the decrease in function is represented as $\|t\|$. The function intensity decreases for the w_p. The cross guided bilateral filtering computes similar pixels in the frames to locate the neighbourhood distance. The cross guided bilateral filtering guidance is represented as w_g as in equation (2). The spatial weight similarity indicators are not related to the photometric weights with the guided image label as the face, mouth, node, and so on. The reduced cost function $F(x)$ for the pre-processing is achieved with the cross-guided bilateral filter.

$$\sum_{t \in S_m} w_s(\|t\|)\, w_g(G(x) - G(x+t)\, \emptyset((F(x) - E(x+t)^2) \tag{2}$$

The spatial weights in the images are represented as w_s, w_g is denoted as the guide weight for the image, and the similarity between each pixel features of images are denoted as $q_t = w_s w_g$ and $\emptyset$ represents the noises in the images. The pre-processed image after the pre-processing is presented as in equation (3)

$$F_{k+1}(x) = \frac{\sum_{t \in S_m} q_t \emptyset^{;} (F_k(x) - E(x+t))^2 E(x+t)}{\sum_{t \in S_m} q_t \emptyset^{;} (F_k(x) - E(x+t))^2} \tag{3}$$

where, $q_t = w_s(\|t\|) w_g(G(x) - G(x+t))$

The noise elimination in the frames is performed with the cross guided bilateral filter with the $w_g = 1$ with the iteration value of $w_p = 1$ and $w_s = 1$. The pre-processing with the cross guided bilateral filtering process provides the original image weights equal to the adequate noise value of Ø.

2) FEATURE EXTRACTION

In this extraction process, the features in the images are considered with the hybrid dual-tree complex wavelet transforms with the Walsh-Hadamard transform. To construct an effective model with minimal effort feature extraction is considered the effective approach to reducing the redundant data. Through the effective feature extraction model the process of Machine learning or Deep Learning process is improved. The filtered or pre-processed image in the previous stage is feature extracted with the hybrid DTCWT (dual-tree complex wavelet transform) integrated with Walsh-Hadamard transform. With the hybrid DTCWT model, the complex signal transform is calculated with the separate DWT decomposition (tree A and tree B). The filtering is applied over the frames on the images then the first DWT generates the real coefficient and the second DWT generates the imaginary coefficients.Through the scaling function image signal is decomposed as $f(x, y)$ with the complex wavelet function represented in equation (4)

$$\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y}) = \sum_{l \in Z} \boldsymbol{A}_{j,l}, \boldsymbol{\emptyset}_{j,l}(\boldsymbol{x},\boldsymbol{y}) + \sum_{k \in \alpha} \sum_{j=1}^{j_0} \sum_{l \in Z^2} \boldsymbol{D}_{j.i}^{k} \boldsymbol{\varphi}_{j,l}^{k}(\boldsymbol{x},\boldsymbol{y}) \tag{4}$$

where in the above equation (4) the image decomposition level is denoted as j_0, and the wavelet scaling coefficient is represented as $A_{j,l}$ and $D_{j.i}^{k}$. the scaling function is denoted as $\varphi_{j,l}^{k}(x, y)$.

The feature extraction process comprises the two-dimensional Walsh-Hadamard Transform with the entities as either +1 or −1 those pixels rows are mutually orthogonal. The Walsh Hadamard model comprises two different rows with matching entities in one column and mismatched entities in another column. The Hadamard matrix for the image is given in equation (5)

$$HH^T = nI_n \tag{5}$$

In the above equation (5), the H^T represents the transpose of the identity matrix of $n \times n$ for the matrix H. The Hadamard matrix order is denoted as the $n = 1, 2orn == 0(mod4)$. Based on this the minimal Hadamard kernel is computed as the $H_1^1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and the higher order of the hybrid DTCWT Hadamard kernel is stated as the $H_{2n}^1 = \begin{bmatrix} H_n^1 & H_n^1 \\ H_n^1 & -H_n^1 \end{bmatrix}$. The extracted features with the hybrid DTCWT Walsh Hadamard transformation are utilized for the feature training and testing process of the optimized MCNN.

3) OPTIMIZATION OF FEATURES WITH THE FLAMINGO SEARCH ALGORITHM (FSA)

Optimization is widely used in different applications due to higher applicability with minimal parameters and higher capability of global search. Flamingos are migratory birds those are acquired food from algae, clams, small shrimps, larvae, and small worms. The physical appearance of the flamingo is presented in Figure 2.

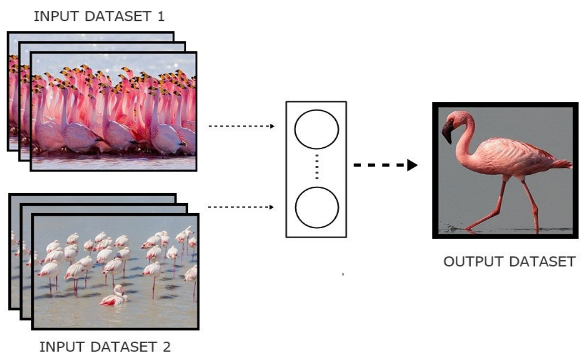

FIGURE 2. **Extraction of flamingo migration behaviour.**

The main characteristic of the FSA model is to compute the optimal feature for the effective classification process. To derive the optimal features FSA comprises the foraging and migratory behaviour. The characteristics of the flamingo are inhabitant in the area of plenty of food available and after foraging it moves to some other place with the migration model. Based on the characteristics of the flamingo the effective features are examined in the images. In Figure 3 the process flow in the FSA model is presented. The characteristics of the flamingo considered for the optimization of the features are listed as follows:

1. For local communication flamingos sings to each other about the availability of food.
2. The flamingo population does not aware of the present search area food availability. Instead, to identify the higher food area communicates with each other.
3. The change in position is computed based on the flamingo's behaviour with two behaviour such as foraging and migration. As the foraging behaviour comprises the two characteristics such as flamingo foot movement and foraging behaviour.

The process in FSA is described as follows:

The FSA comprises the different features for the evaluation through which optimal features in the images are derived. As stated, flamingo comprises two behaviour such as foraging and migration which are explained as follows:

a: FORAGING BEHAVIOUR

FEATURE 1: COMMUNICATIVE BEHAVIOUR

The food source is not known to the flamingos, they need to spread the information about location and changes in position. In theory, the flamingos are not aware of the area of major food sources with the global optimal point. The global optimal evaluation is based on the position and availability of food sources for the estimation to derive the optimal solution in search space. This work considers the food source in j^{th} dimension of x_{bj}.

FEATURE 2: BEAK SCANNING BEHAVIOUR

The beak scanning flamingo's behaviour is evaluated based on the i^{th} position of the flamingo in the j^{th} dimensional population as x_{ij} for every flamingo's those influence on the foraging behaviour. However, the foraging behaviour is subjected to the error related to the information transmission in small probability value. The maxima foraging behaviour based on maximum distance is stated as $|G1 \times x_{bj} + \varepsilon_2 \times x_{ij}|$, with the random number of -1 or 1. The maximal distance is increased in the flamingo's beak scan for the random number G1 with the standard normal distribution. The scanning behaviour is evaluated based on the normal distribution-based variation curve with the beak scanning range of $G2 \times |G1 \times xbj + \varepsilon 2 \times xij|$, where G2 is stated as the random number of the standard distribution.

FEATURE 3: BIPEDAL MOBILE BEHAVIOUR

The flamingo's behaviour model movement is shown in Figure 4. The flamingos are scanning and forage towards the beaks and claws towards the food in abundant range for the complete population, The location of the food is stated as x_{bj} with the distance accepted as $\varepsilon 1 \times x_{bj}$, the random number $\varepsilon 1$ is denoted as -1 or 1 those increases based on the search range of the foraging in flamingos. The iteration of flamingos in the foraging distance is computed using equation (6)

$$b_{ij}^{t} = \varepsilon_1 \times xb_j^t + G_2 \times \left| G_1 \times xb_j^t + \varepsilon_2 \times x_{ij}^t \right| \tag{6}$$

The location update of the flamingo foraging behaviour is presented in equation (7)

$$x_{ij}^{t+1} = x_{ij}^{t} + \varepsilon_1 \times xb_j^t + G_2 \times \left| G_1 \times xb_j^t + \varepsilon_2 \times x_{ij}^t \right| / G \tag{7}$$

In the above equation (7) x_{ij}^{t+1} denotes the i^{th} flamingo and j^{th} population dimension for the iteration $t + 1$. The position of flamingos is stated as x_{ij}^{t} for the total iteration count of t. The best fitness value for the flamingos in the population with the t iteration is represented as xb_j^t with the chi-square distribution with the degree of freedom.

b: MIGRATION BEHAVIOUR

The flamingo population migrates from the next area with maximal food availability. The migration behaviour of the flamingo's population is given in equation (8)

$$x_{ij}^{t+1} = x_{ij}^{t} + \omega \times \left(xb_j^t - x_{ij}^t \right) \tag{8}$$

In the above equation (8), the x_{ij}^{t} provides the position of flamingos for the t iterations. The best fitness function for the population is stated as $\omega = \mathrm{N}\ (0, \mathrm{n})$ with the Gaussian random number value with a degree of freedom value for the increases in search space for the migration process.

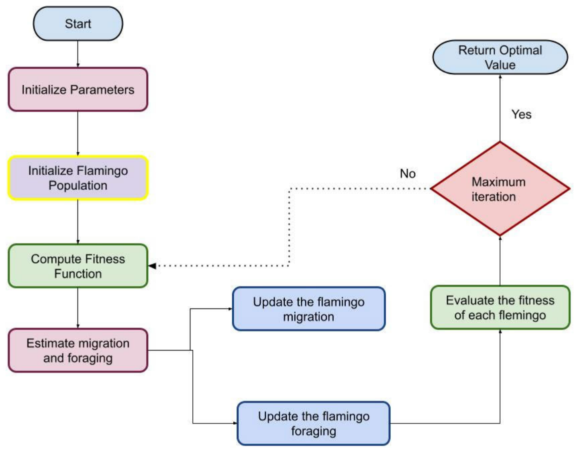

FIGURE 3. Flow chart of FSA.

4) MULTISCALE CONVOLUTIONAL NEURAL NETWORK (MCNN)

Multi-scale Convolutional Neural Networks (MCNN) is effective pattern recognition task for the computation of minimal variation is consider as the pre-processing input for the feature extraction. The drowsiness model uses the MCNN architecture comprised of the deep neural network for the convolutional and pooling layers. The MCNN model comprises the architecture with the convolution and pooling layer for the feature extraction. Through the fully connected layers, final classification is performed in the stacked top. The MCNN comprises the convolutional layer with the remapping of the single kernels. The Multi-Scale in the inception module is performed for the extraction of the relevant features with the multiple features. The MCNN model comprises the convolutional layer processing I pixel-wise to achieve a fully connected with the convolutional filter of 1×1 for the estimation of the density map as illustrated in Figure 4.

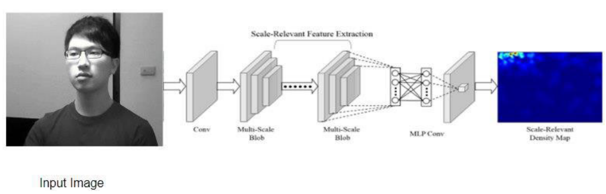

FIGURE 4. Process carried out in MCNN.

The complete architecture model employed for the drowsiness detection and classification is performed with the optimized feature extraction model MCNN as illustrated in Figure 5.

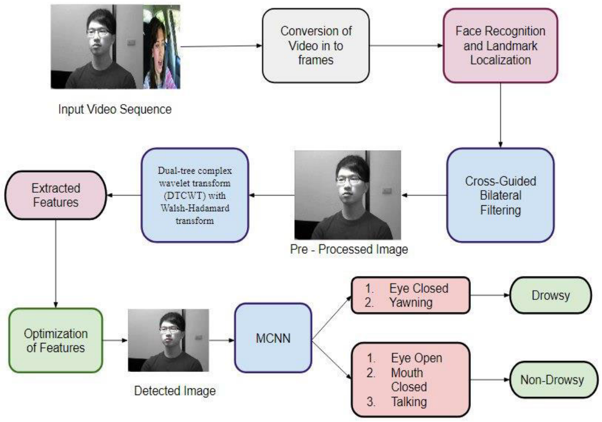

FIGURE 5. Enhanced architecture of MCNN with FSA.

a: CONVOLUTION LAYER

The convolution layer in the network is denoted as C_i with the parametrized map number of N with the maps M_{ij} for the kernel of $K_x \times K_Y$ with the connection with the previous layer is represented as L_{i-1}. The resulting model comprises the squashing function value of $\varphi(x) = 1.7159 \tanh(x)$ for the connected convolution map.

b: MAX_POOLING LAYER

The convolutional layer with the subsampling layer is involved in the reduction of map size with the introduction of invariance for the shifted inputs. The max-pooling output layer provides the maximal activation function for the overlapping region with the size of $K_x \times K_Y$ for the subsampling layer. The resulting pooling is provided to the output map through the squashing function φ.

c: CLASSIFICATION LAYER

The pooling layer is elected based on the output maps with the convolutional layer for the 1 pixel in the down sampling for the image resulting in the 1D vector. The classical feed forward model is connected to the layer and performs the classification. The present work uses the neurons count of 6 with the SoftMax activation function for the estimation of output probabilities.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The proposed approach is simulated with the Python software with Intel Core I5 Processor with the 64-bit operating system. The drowsiness detection module is designed and implemented with the use of Dlib library face recognition system model. The distribution of the dataset for the proposed model and simulation analysis of the results are presented.

A. EXPERIMENTAL SETUP

For our experimental analysis YAWDD Dataset and NTHUDDD Dataset are considered. The training and testing of the dataset for the classification are presented in Table 5. The YAWDD and NTHU-DDD dataset training and testing are conducted based on two classes such as Drowsy and Non-drowsy. The dataset distribution is in the ratio of 80% for the training and 20% for the testing.

TABLE 5. Distribution of various training and testing dataset.

Dataset	Total	Classes	Training	Testing
YAWDD Dataset	10314	Drowsy	5594	728
		Non-Drowsy	3639	353
NTHU-DDD	10,155	Drowsy	3886	275
		Non-Drowsy	5316	678

B. DATA PREPROCESSING

The proposed method uses the cross guided bilateral filtering process for the elimination of redundant information. The pre-processing was performed for the extracted frames in the video sequences with the identification of facial landmarks. At first, the images are resized 640×480 to 224×224. In the next stage, the images are converted in the form of RGB to BGR scale based on consideration of mean value (103.939, 116.779, 123.68) as shown in Figure 6. Every pixel in the image is subtracted from the RGB means value for all pixels.

(a) (b)

FIGURE 6. (a) Preprocessed image (b) Detected image for classification with MCNN.

Figure 6 (a) provides the pre-processed image for the extracted frames in the video sequences and Figure 6 (b) provides the features extracted based detected image for the drowsiness classification is presented.

C. RESULTS FOR THE YAWDD DATASET

Figure 7 provides the loss estimated for the training and testing process. For the YAWDD dataset, the epoch is set as 100. The YAWDD dataset loss estimation for the training and testing is below 0.05 for the epoch above 5. This implies that the proposed model achieves minimal loss for the training and testing of the YAWDD dataset.

Table 6 provides the performance metrices of the proposed model for the different parameters for the classification of

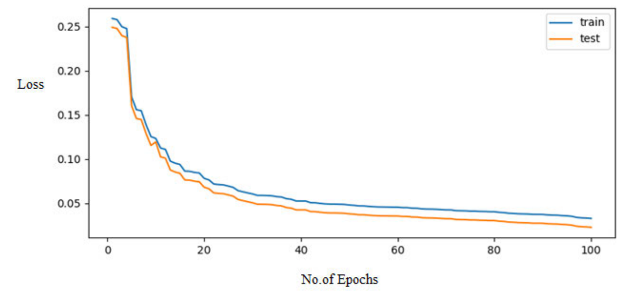

FIGURE 7. Estimation of loss for YAWDD.

drowsy and non-drowsy. The proposed model is implemented with ML (machine learning) and DL (deep learning) model. Therefore, the machine learning classifiers considered for the analysis of KNN (K-Nearest Neighbour), NB (Naive Bayes), RF (Random Forest), AdaBoost, and SVM (Support Vector Machine). The deep learning model considered for the comparative analysis are pre-trained model such as Alex Net, Res Net, VGG Net and proposed MCNN with FSA.

TABLE 6. Performance measured for the proposed YAWDD dataset.

Methods	Accuracy	Precision	Recall	F Measure	Kappa Score	FPR	FNR
Proposed MCNN with FSA	98.38	97.06667	97.84946	97.45649	96.26380275	0.013784	0.021505
Alex Net	95.3	92.95393	92.2043	92.5776	89.13803057	0.032581	0.077957
Res Net	95.56	93.71585	92.2043	92.95393	89.70830035	0.028822	0.077957
VGG Net	95.9	94.2623	92.74194	93.49593	90.49996955	0.026316	0.072581
SVM	93.33	86.93467	93.01075	89.87013	84.91041489	0.065163	0.069892
Random Forest	90.34	81.66259	89.78495	85.53137	78.30752419	0.093985	0.102151
KNN	87.69	77.27273	86.82796	81.77215	72.52930012	0.119048	0.13172
Naive Bayes	83.42	70.41284	82.52688	75.9901	63.44786272	0.161654	0.174731
AdaBoost	79.91	65.25612	78.76344	71.37637	56.1143087	0.195489	0.212366

The comparative analysis of the results represent that the DL (deep learning) model exhibits a higher performance compared with the machine learning classifier models. In comparative analysis of the deep learning model, the proposed MCNN with FSA model achieves a higher accuracy value of 98.3% for the classification.

D. RESULTS FOR NTHU-DDD DATASET

With the NTHU-DDD dataset, the loss function is computed for the training and testing dataset. The evaluation of the NTHU-DDD dataset loss value is minimal than 0.05, but it slightly higher than the YAWDD dataset as presented in Figure 8.

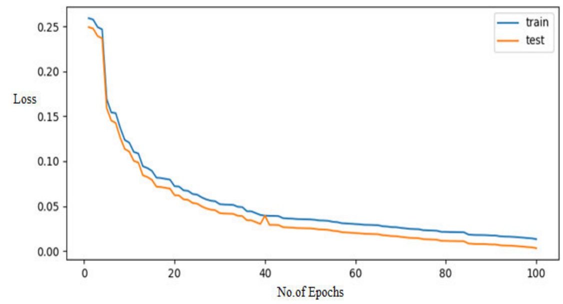

FIGURE 8. Estimation of loss for NTHU-DDD.

In Table 7 comparative examination of the proposed MCNN with FSA model performance is stated. The comparison is performed based on the consideration of the Machine Learning and Deep Learning model for the classification.

TABLE 7. Comparison of performance for NTHU-DDD dataset.

Methods	Accuracy	Precision	Recall	F measure	Kappa Score	FPR	FNR
Proposed	98.26	99.45205	98.10811	98.77551	95.77598912	0.013559	0.018919
Alex Net	96.71	98.75691	96.62162	97.6776	92.069323	0.030508	0.033784
Res Net	96.14	98.74652	95.81081	97.25652	90.72559869	0.030508	0.041892
VGG Net	95.56	98.19444	95.54054	96.84932	89.31313131	0.044068	0.044595
SVM	93.72	97.46835	93.64865	95.52033	85.03353689	0.061017	0.063514
Random Forest	91.98	96.86163	91.75676	94.24011	81.07631814	0.074576	0.082432
KNN	89.86	96.35036	89.18919	92.63158	76.43008187	0.084746	0.108108
Naive Bayes	87.92	95.82712	86.89189	91.14103	72.31465531	0.094915	0.131081
AdaBoost	85.8	95.26718	84.32432	89.46237	67.9306529	0.105085	0.156757

The experimental analysis expressed that the proposed MCNN with FSA model achieves a higher accuracy value of 98.2% which is significantly higher than the conventional machine learning model. Figure 9 – 15 provides the comparative illustration of the proposed MCNN with FSA with the conventional pre-trained and machine learning model.

The comparative analysis expressed that the proposed MCNN model with FSA attains the higher accuracy value of around 98.93% which stated significantly way up than

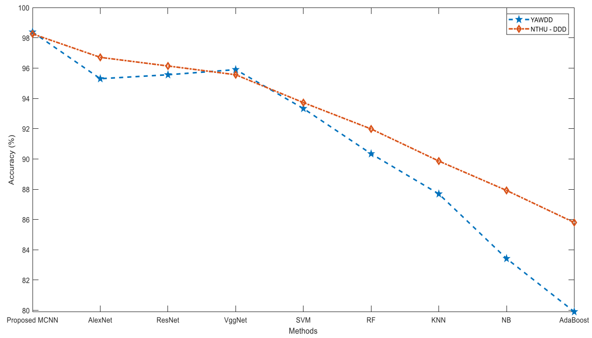

FIGURE 9. Illustrate the comparison of accuracy.

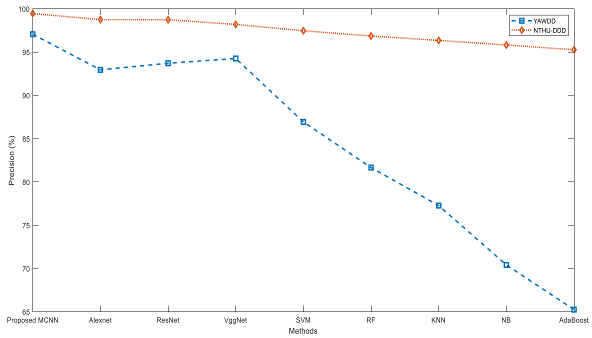

FIGURE 10. Stated the comparison of precision.

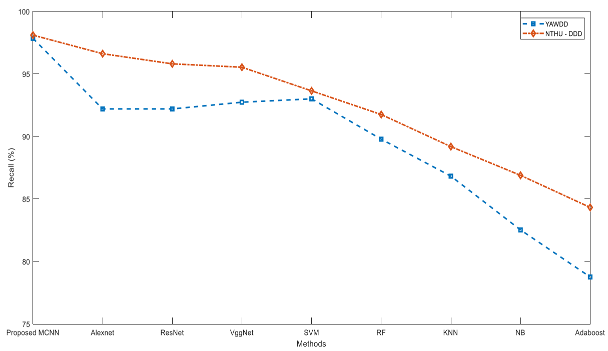

FIGURE 11. Define the comparison of recall.

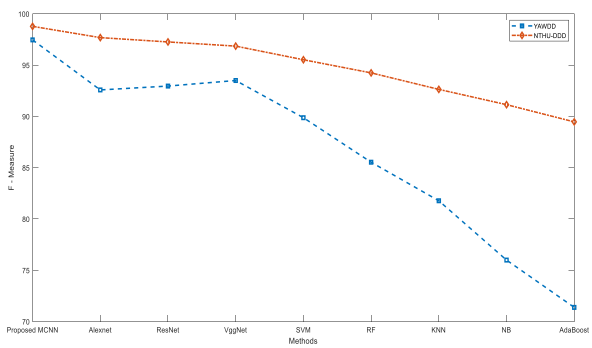

FIGURE 12. Represented the comparison of F-measure.

the conventional machine learning and pre-trained models. In machine learning model the SVM classifier achieves the higher accuracy value of 93.33%. Through the comparative

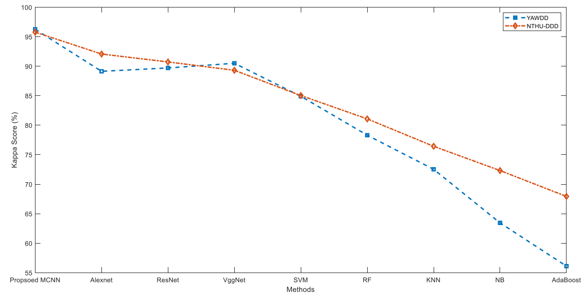

FIGURE 13. Specified the comparison of kappa score.

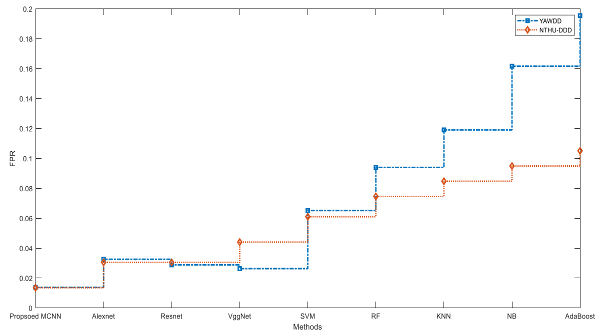

FIGURE 14. Simplified the comparison of FPR.

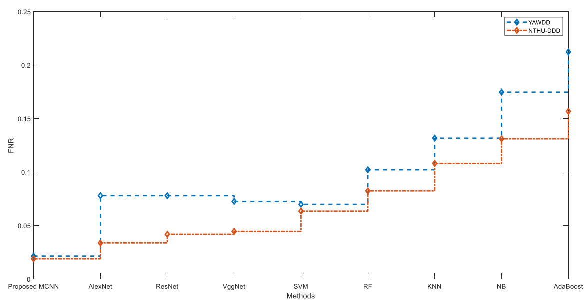

FIGURE 15. Denoted the comparison of FNR.

analysis, it can be stated that deep learning performance is higher than the machine learning classifiers. In case of deep learning model, the proposed MCNN with FSA model exhibits improved performance in terms of the Kappa Score, precision, F-measure, accuracy and recall. Also, the proposed MCNN model achieves reduced FPR and FNR value. With the YAWDD dataset, the proposed NTHU-DDD dataset achieves the higher F-measure, recall and precision value. With the NTHU-DDD model, an accuracy is observed as 98.26%B while the YAWDD model achieves the accuracy of 98.38%. The FPR value for the YAWDD is measured as 0.013 and NTHU-DD achieves an FPR of 0.013. Similarly, for the FNR value YAWDD dataset exhibits value of 0.021 and NTHU-DDD exhibits an FNR value of 0.0189. Through comparative analysis, it is concluded that DL (deep learning) based approach is effective for detecting the drowsiness. The proposed MCNN with FSA model achieves enhanced performance compared with the existing methods for the classification.

V. CONCLUSION

Drowsiness is a major cause of accidents that may lead to mortality and severe injury to the person. This paper presented an effective model for the classification of Drowsy and Non-drowsy in drivers to prevent accidents. The developed model performed the conversion of video sequences into frames followed by the location of landmarks in the face. To increase the processing accuracy pre-processing is performed with the filtering process. The feature extraction is performed with the DTCWT model. The extracted features are optimized with the flamingo search optimization model and classification is performed with the MCNN. The simulation results demonstrated that the proposed MCNN with FSA model achieves a higher accuracy value of 98.38% for the YAWDD and NTHU-DDD achieves an accuracy of 98.26%. Through analysis, it is concluded that the proposed research is effective for defining the classification of Drowsy and Non-drowsy in drivers.

REFERENCES

[1] Y. K. Arora and S. Kumar, "Statistical approach to predict road accidents in India," in *Computing in Engineering and Technology*. Singapore: Springer, 2020, pp. 189–196.
[2] C. Panda, A. K. Dash, and D. P. Dash, "Assessment of risk factors of road traffic accidents: A panel model analysis of several states in India," *Vis., J. Bus. Perspective*, vol. 2022, Aug. 2022, Art. no. 097226292211132.
[3] R. B. Rajasekaran, S. Rajasekaran, and R. Vaishya, "The role of social advocacy in reducing road traffic accidents in India," *J. Clin. Orthopaedics Trauma*, vol. 12, no. 1, pp. 2–3, Jan. 2021.
[4] S. D. Khan, L. Alarabi, and S. Basalamah, "A unified deep learning framework of multi-scale detectors for geo-spatial object detection in high-resolution satellite images," *Arabian J. Sci. Eng.*, vol. 47, no. 8, pp. 9489–9504, Aug. 2022.
[5] Y. Albadawi, M. Takruri, and M. Awad, "A review of recent developments in driver drowsiness detection systems," *Sensors*, vol. 22, no. 5, p. 2069, Mar. 2022.
[6] M. K. Hussein, T. M. Salman, A. H. Miry, and M. A. Subhi, "Driver drowsiness detection techniques: A survey," in *Proc. 1st Babylon Int. Conf. Inf. Technol. Sci. (BICITS)*, Apr. 2021, pp. 45–51.
[7] I. Stancin, M. Cifrek, and A. Jovic, "A review of EEG signal features and their application in driver drowsiness detection systems," *Sensors*, vol. 21, no. 11, p. 3786, May 2021.
[8] M. Dua, R. Singla, S. Raj, and A. Jangra, "Deep CNN models-based ensemble approach to driver drowsiness detection," *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3155–3168, Apr. 2021.
[9] E. Magán, M. P. Sesmero, J. M. Alonso-Weber, and A. Sanchis, "Driver drowsiness detection by applying deep learning techniques to sequences of images," *Appl. Sci.*, vol. 12, no. 3, p. 1145, Jan. 2022.
[10] A. Moujahid, F. Dornaika, I. Arganda-Carreras, and J. Reta, "Efficient and compact face descriptor for driver drowsiness detection," *Expert Syst. Appl.*, vol. 168, Apr. 2021, Art. no. 114334.
[11] S. Chaabene, B. Bouaziz, A. Boudaya, A. Hökelmann, A. Ammar, and L. Chaari, "Convolutional neural network for drowsiness detection using EEG signals," *Sensors*, vol. 21, no. 5, p. 1734, Mar. 2021.
[12] A. Quddus, A. Shahidi Zandi, L. Prest, and F. J. E. Comeau, "Using long short term memory and convolutional neural networks for driver drowsiness detection," *Accident Anal. Prevention*, vol. 156, Jun. 2021, Art. no. 106107.
[13] T. Arakawa, "Trends and future prospects of the drowsiness detection and estimation technology," *Sensors*, vol. 21, no. 23, p. 7921, Nov. 2021.
[14] M. M. Hasan, C. N. Watling, and G. S. Larue, "Physiological signal-based drowsiness detection using machine learning: Singular and hybrid signal approaches," *J. Saf. Res.*, vol. 80, pp. 215–225, Feb. 2022.
[15] G. Li and W.-Y. Chung, "Electroencephalogram-based approaches for driver drowsiness detection and management: A review," *Sensors*, vol. 22, no. 3, p. 1100, Jan. 2022.
[16] S. Bakheet and A. Al-Hamadi, "A framework for instantaneous driver drowsiness detection based on improved HOG features and Naïve Bayesian classification," *Brain Sci.*, vol. 11, no. 2, p. 240, Feb. 2021.
[17] M. Ahmed, S. Masood, M. Ahmad, and A. A. A. El-Latif, "Intelligent driver drowsiness detection for traffic safety based on multi CNN deep model and facial subsampling," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 19743–19752, Oct. 2022.
[18] V. Kumar and S. Sharma, "Driver drowsiness detection using modified deep learning architecture," *Evol. Intell.*, vol. 2022, pp. 1–10, Jun. 2022.
[19] H. Varun Chand and J. Karthikeyan, "CNN based driver drowsiness detection system using emotion analysis," *Intell. Autom. Soft Comput.*, vol. 31, no. 2, pp. 717–728, 2022.
[20] V. P. Balam, V. U. Sameer, and S. Chinara, "Automated classification system for drowsiness detection using convolutional neural network and electroencephalogram," *IET Intell. Transp. Syst.*, vol. 15, no. 4, pp. 514–524, Apr. 2021.
[21] S. Arefnezhad, S. Samiee, A. Eichberger, M. Frühwirth, C. Kaufmann, and E. Klotz, "Applying deep neural networks for multi-level classification of driver drowsiness using vehicle-based measures," *Expert Syst. Appl.*, vol. 162, Dec. 2020, Art. no. 113778.
[22] S. E. Bekhouche, Y. Ruichek, and F. Dornaika, "Driver drowsiness detection in video sequences using hybrid selection of deep features," *Knowl.-Based Syst.*, vol. 252, Sep. 2022, Art. no. 109436.
[23] J. Cui, Z. Lan, Y. Liu, R. Li, F. Li, O. Sourina, and W. Müller-Wittig, "A compact and interpretable convolutional neural network for cross-subject driver drowsiness detection from single-channel EEG," *Methods*, vol. 202, pp. 173–184, Jun. 2022.
[24] X. Zhang, X. Wang, X. Yang, C. Xu, X. Zhu, and J. Wei, "Driver drowsiness detection using mixed-effect ordered logit model considering time cumulative effect," *Analytic Methods Accident Res.*, vol. 26, Jun. 2020, Art. no. 100114.
[25] N. N. Pandey and N. B. Muppalaneni, "Dumodds: Dual modeling approach for drowsiness detection based on spatial and spatio-temporal features," *Eng. Appl. Artif. Intell.*, vol. 119, Mar. 2023, Art. no. 105759.
[26] S. Chinara, "Automatic classification methods for detecting drowsiness using wavelet packet transform extracted time-domain features from single-channel EEG signal," *J. Neurosci. Methods*, vol. 347, Jan. 2021, Art. no. 108927.
[27] C. Lee and J. An, "LSTM-CNN model of drowsiness detection from multiple consciousness states acquired by EEG," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 119032.
[28] M. Shen, B. Zou, X. Li, Y. Zheng, L. Li, and L. Zhang, "Multi-source signal alignment and efficient multi-dimensional feature classification in the application of EEG-based subject-independent drowsiness detection," *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 103023.
[29] A.-C. Phan, T.-N. Trieu, and T.-C. Phan, "Driver drowsiness detection and smart alerting using deep learning and IoT," *Internet Things*, vol. 22, Jul. 2023, Art. no. 100705.
[30] I. A. Fouad, "A robust and efficient EEG-based drowsiness detection system using different machine learning algorithms," *Ain Shams Eng. J.*, vol. 14, no. 3, Apr. 2023, Art. no. 101895.
[31] J. Chen, S. Wang, E. He, H. Wang, and L. Wang, "Recognizing drowsiness in young men during real driving based on electroencephalography using an end-to-end deep learning approach," *Biomed. Signal Process. Control*, vol. 69, Aug. 2021, Art. no. 102792.
[32] S. P. Rajamohana, E. G. Radhika, S. Priya, and S. Sangeetha, "Driver drowsiness detection system using hybrid approach of convolutional neural network and bidirectional long short term memory (CNN_BILSTM)," *Mater. Today, Proc.*, vol. 45, pp. 2897–2901, Jan. 2021.
[33] V. Kamath and A. Renuka, "Deep learning based object detection for resource constrained devices: Systematic review, future trends and challenges ahead," *Neurocomputing*, vol. 531, pp. 34–60, Apr. 2023.
[34] R. Jabbar, K. Al-Khalifa, M. Kharbeche, W. Alhajyaseen, M. Jafari, and S. Jiang, "Real-time driver drowsiness detection for Android application using deep neural networks techniques," *Proc. Comput. Sci.*, vol. 130, pp. 400–407, Jan. 2018.
[35] V. Vijayan and P. Kp, "A comparative analysis of RootSIFT and SIFT methods for drowsy features extraction," *Proc. Comput. Sci.*, vol. 171, pp. 436–445, Jan. 2020.

[36] J. He, W. Choi, Y. Yang, J. Lu, X. Wu, and K. Peng, ''Detection of driver drowsiness using wearable devices: A feasibility study of the proximity sensor,'' *Appl. Ergonom.*, vol. 65, pp. 473–480, Nov. 2017.

[37] J. Escorcia-Gutierrez, K. Beleño, J. Jimenez-Cabas, M. Elhoseny, M. D. Alshehri, and M. M. Selim, ''An automated deep learning enabled brain signal classification for epileptic seizure detection on complex measurement systems,'' *Measurement*, vol. 196, Jun. 2022, Art. no. 111226.

[38] S. K. Khare and V. Bajaj, ''Optimized tunable Q wavelet transform based drowsiness detection from electroencephalogram signals,'' *IRBM*, vol. 43, no. 1, pp. 13–21, Feb. 2022.

[39] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, and B. Hariri, ''YawDD: A yawning detection dataset,'' in *Proc. 5th ACM Multimedia Syst. Conf.*, Mar. 2014, pp. 24–28.

[40] C.-H. Weng, Y.-H. Lai, and S.-H. Lai, ''Driver drowsiness detection via a hierarchical temporal deep belief network,'' in *Proc. Int. Workshops Comput. Vis.*, vol. 13. Cham, Switzerland: Springer, 2017, pp. 117–133.

V. VIJAYPRIYA received the M.E. degree in computer science and engineering from the SRM Valliammai Engineering College, Chennai, India. For two years, she was a Teaching Associate with the SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India, where she is currently a Research Scholar with the Department of Computer Science and Engineering. She is the author of three journal publications, five international conferences, and two patents published. Her current research interests include the Internet of Things, image processing, machine learning, and computer vision.

MOHAN UMA received the B.Sc. and M.C.A. degrees from Bharathiyar University, Trichy, India, the M.Tech. degree in computer science from SRM University, Chennai, India, and the Ph.D. degree in brain–computer interface from Bharathiyar University, Coimbatore, India. She has more than 20 years of teaching and research experience. Currently, she is an Associate Professor with the Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India. She has published more than 40 papers in national and international journals and conferences. She has published six patents. Her current research interests include brain–computer interface, EEG, P300, data mining, recommendation systems, machine learning, web services, data analytics, and image processing.