

Article

Eye-Blink Event Detection Using a Neural-Network-Trained Frame Segment for Woman Drivers in Saudi Arabia

Muna S. Al-Razgan ^{1,*}, Issema Alruwaly ² and Yasser A. Ali ^{3,*}

¹ Software Engineering Department, College of Computer & Information Sciences, King Saud University, P.O. Box 22452, Riyadh 11495, Saudi Arabia

² Sociology Department, College of Humanities and Social Sciences, King Saud University, P.O. Box 2456, Riyadh 11451, Saudi Arabia; aalrowily@ksu.edu.sa

³ Department of Information Systems, College of Computer and Information Sciences, King Saud University, P.O. Box 51178, Riyadh 11543, Saudi Arabia

* Correspondence: malrazgan@ksu.edu.sa (M.S.A.-R.); yasali@ksu.edu.sa (Y.A.A.)

Abstract: Women have been allowed to drive in Saudi Arabia since 2018, revoking a 30-year ban that also adhered to the traffic rules provided in the country. Conventional drivers are often monitored for safe driving by monitoring their facial reactions, eye blinks, and expressions. As driving experience and vehicle handling features have been less exposed to novice women drivers in Saudi Arabia, technical assistance and physical observations are mandatory. Such observations are sensed as images/video frames for computer-based analyses. Precise computer vision processes are employed for detecting and classifying events using image processing. The identified events are unique to novice women drivers in Saudi Arabia, assisting with their vehicle usage. This article introduces the Event Detection using Segmented Frame (ED-SF) method to improve the abnormal Eye-Blink Detection (EBD) of women drivers. The eye region is segmented using variation pixel extraction in this process. The pixel extraction process requires textural variation identified from different frames. The condition is that the frames are to be continuous in the event detection. This method employs a convolution neural network with two hidden layer processes. In the first layer, continuous and discrete frame differentiations are identified. The second layer is responsible for segmenting the eye region, devouring the textural variation. The variations and discrete frames are used for training the neural network to prevent segment errors in the extraction process. Therefore, the frame segment changes are used for Identifying the expressions through different inputs across different texture luminosities. This method applies to less-experienced and road-safety-knowledge-lacking woman drivers who have initiated their driving journey in Saudi-Arabia-like countries. Thus the proposed method improves the EBD accuracy by 9.5% compared to Hybrid Convolutional Neural Networks (HCNN), Long Short-Term Neural Networks (HCNN + LSTM), Two-Stream Spatial-Temporal Graph Convolutional Networks (2S-STGCN), and the Customized Driving Fatigue Detection Method CDFDM.

Keywords: CNN; eye blink detection; feature extraction; woman driving



Citation: Al-Razgan, M.S.; Alruwaly, I.; Ali, Y.A. Eye-Blink Event Detection Using a Neural-Network-Trained Frame Segment for Woman Drivers in Saudi Arabia. *Electronics* **2023**, *12*, 2699. <https://doi.org/10.3390/electronics12122699>

Academic Editor: Beiwen Li

Received: 6 May 2023

Revised: 7 June 2023

Accepted: 13 June 2023

Published: 16 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

There are few women drivers in Saudi Arabia. Financial barriers and high driving class fees present difficulties in getting a license for women in Saudi Arabia. Driving class fee rates are high for women when compared to driving classes for men [1,2]. Proper driving skills and knowledge are a must for women drivers to get a license for driving a vehicle in Saudi Arabia. Women drivers require the proper training to drive vehicles [3]. Saudi Arabia was the world's last country with a low rate of women drivers. Today, the various opportunities and possibilities available are increasing in Saudi Arabia for women drivers [4].

Learning the safety and driving rules is important for every woman driver in Saudi Arabia. Driving classes and sessions are provided to women drivers to learn the skills and safety rules for the driving process [5,6]. High and low driving classes and practices are provided to women drivers in Saudi Arabia. Safety rules such as backups, cautions, precautions, and the actual meaning of signs are provided via these driving classes [7,8]. The proper road safety skills and knowledge are also provided to the women drivers, reducing the accident rates in Saudi Arabia [9].

Eye-blink detection is a method used to detect a person's eye-blinking range. The eye-blink detection method is used in the monitoring system for women drivers in Saudi Arabia [10,11]. The eye-blink detection method identifies important evidence for detecting a driver's driving behavior during their driving. A driver's normal blinking (NB) rate is monitored via wireless sensors at both roadsides and in vehicles [12,13]. The eye-blink detection method maximizes the safety of travelers and minimizes the overall accident rate [14].

Image processing technology is used to identify the important features and patterns from the given images [15]. The image processing technique is used for the women driver's Eye Blink Detection (EBD) process in Saudi Arabia [16]. The image processing technique detects the exact blinking range of these women drivers, which the Saudi Arabia government uses for its safety management systems [17]. Image processing reduces the eye-blink detection process's latency and error ratio, improving the system's efficiency range [18]. Several methods, such as Convolution Neural Networks (CNN), the attention-based multi-modal fusion approach, the Head-up-display (HUD) technique, the Situation awareness global assessment technique (SAGAT), and long short-term memory (LSTM), are widely applied to perform the blink detection process. However, while these existing methods achieve the maximum accuracy, they have a latency and error rate. The research issues are addressed with the help of two-layer convolution neural networks. The contributions of this article are listed as follows:

- Designing an event detection method for identifying the eye blinks of women drivers in Saudi Arabia to assist in safe driving
- Designing an Event Detection using the Segmented Frame (ED-SF) method, which uses a two-layer convolution neural network for frame differentiation and sequence detection in order to reduce the variation errors in the event detection
- Performing an experimental analysis using the Niqab dataset to prove the consistency of the proposed method
- Performing a comparative analysis using specific metrics and methods for external verification.

2. Related Works

This section presents a discussion of the works related to the proposed concept. The works from the previous authors, with their key focus areas and findings, are tabulated and theoretically expressed in this section. First, Table 1 summarizes the references [19–25] for the same.

Yamabe et al. [26] proposed a new comfortable awakening method for sleeping drivers. The proposed method is commonly used for automated driving systems. The proposed method detects the exact drowsiness range of drivers during autonomous driving. A human-machine interface (HMI) is implemented in the vehicles, which provides the necessary interaction services to the drivers. HMI reduces these drowsiness levels (2.3%) and ensures the safety of travelers.

Guo et al. [27] presented a Hybrid Convolutional Neural Network (HCNN)-based drowsiness detection method. The long short-term memory (LSTM) algorithm is also used in the detection method, which detects the relevant data for further processes. The LSTM and CNN algorithms decrease the drowsiness detection process's time and energy consumption ratio. The experimental results show that the presented CNN-based method maximizes the detection process's accuracy (84.85%).

Table 1. Summary of references [19–25].

Author	Title	Key Areas	Methods Used	Findings
Jordan et al. [19]	Deep learning (DL)-based EBD method for edge computing systems.	The main aim of the method is to identify the driver's drowsiness ratio during driving.	The convolutional neural network (CNN) model is used here to detect the eye blinking range of the person.	Increases the accuracy of the EBD process.
Mou et al. [20]	An isotropic self-supervised learning (IsoSSL) model for driver drowsiness detection process.	IsoSSL detects the exact drowsiness range of drivers via videos and images.	An attention-based multi-modal fusion method is implemented to identify the facial features of the drivers.	The IsoSSL model reduces the accident range on roadsides.
Bai et al. [21]	A Two-Stream Spatial-Temporal Graph Convolutional Network (2S-STGCN) for drowsiness detection.	The STGCN identifies the facial expression ratio of the drivers.	A feature extraction method is used to extract the important data for detection.	Minimizes the error ratio in drowsiness detection.
Li et al. [22]	A new EBD method for non-driving-related tasks (NDRT) in automated vehicles.	The goal is to reduce accidents and ensure the users' safety.	Head-up-display (HUD) technique is used to predict the eye-blinking range of the drivers.	Increases the accuracy of the EBD process.
Liang et al. [23]	A new technique for the eye-tracking investigation process.	The proposed technique investigates the effects of pre-takeover request (TOR) for the tracking process.	Situation awareness global assessment technique (SAGAT) is used here to analyze the actual behavioral patterns of the drivers.	Enhances the accuracy and efficiency range of the eye-tracking system.
Akrout et al. [24]	A novel approach for driver fatigue detection process.	The proposed approach identifies drivers' drowsiness, fatigue, and yawning levels during driving.	Visual characteristics analysis is used to produce optimal data for the detection process.	Maximizes the accuracy of fatigue detection.
Zeng et al. [25]	A Customized Driving Fatigue Detection Method (CDFDM).	The developed method identifies the fatigue level of drivers from the given images.	A long short-term memory (LSTM) algorithm is used here to detect the important datasets from the database.	Increases the overall accuracy in the fatigue recognition process.

Li et al. [28] introduced a facial-feature-based driver fatigue detection approach. Both recurrent gate unit (GRU) and integrated facial features data are used for the detection. A multi-task convolutional neural network (MTCNN) model is used to extract the comprehensive facial features of the drivers. The MTCNN model predicts the exact fatigue level of the drivers, providing optimal security services to its users. Thus, the method detects fatigue with a 97.47% accuracy.

Wijnands et al. [29] developed a real-time monitoring system for driver drowsiness detection on a mobile platform. A three-dimensional (3D) neural network is used in the system, which gathers the drivers' facial features and patterns using wireless sensors. Both the spatial and temporal features are combined and produce optimal datasets for monitoring systems. The developed monitoring system improves the performance and efficiency (98.03%) of the driving system's range.

Cui et al. [30] proposed a deep learning (DL)-based real-time detection method for the driver fatigue recognition process. The DL technique detects the exact fatigue cause and condition range of the driver based on facial expressions and features. The DL technique

calculates the fatigue level of drivers and provides the necessary safety services to reduce the fatigue range. The proposed method achieves a high accuracy (2.1% improvement) in the detection process, enhancing the driving systems' effectiveness.

R. Ghosh et al. [31] utilized K-Nearest Neighboring and LSTM networks to identify the EBD from the EEG signal. The EEG signal is analyzed using a 0.5 s sliding window that eliminates the artifacts. The peak value and amplitude-related features are extracted and classified using the classifier. The classifier identifies the eye blinks with a 97.4% accuracy.

Egambaram et al. [32] applied a Deep Learning Model to identify and detect driver drowsiness. The EEG signal is collected and processed with the help of the BLINKER algorithm, which predicts the eye blink features. Then, a deep learning classifier is used to identify the drowsiness with a 94.91% accuracy.

Event detection relies on driver fatigue and drowsiness reflected in the eyes of these drivers. Computer-vision-based techniques extract the heterogeneous features for analyzing the time-sequence-based events. Such a process requires pre-classified sequences for identifying the missing and terminated sequences for precise error detection. To strengthen this concept, this article introduced a segmented frame method for differentiating sequences over various observation intervals. Such intervals are pre-classified for preventing the variation errors that are reflected in event detection.

3. Event Detection Using Segmented Frame Method

The ED-SF method was designed to improve the precise monitoring of women drivers' abnormal eye blink recognition, based on discrete and continuous observation inputs. The required inputs are observed from women drivers in Saudi Arabia (i.e.), and their facial expressions, eye blinks, and expressions at the time of a driving journey are observed at different time intervals. The main objective of this face image processing method is to identify the eye blink region and reduce the segment error when analyzing the variations and discrete frames in the textural feature extraction process. The challenging task in this manuscript is the extraction and augmentation of a discrete frame detected sequence of the novice women drivers. The eye blink instances are stored as records for the previously identified event-occurring instances. The proposed method's step-by-step work is illustrated in Figure 1.

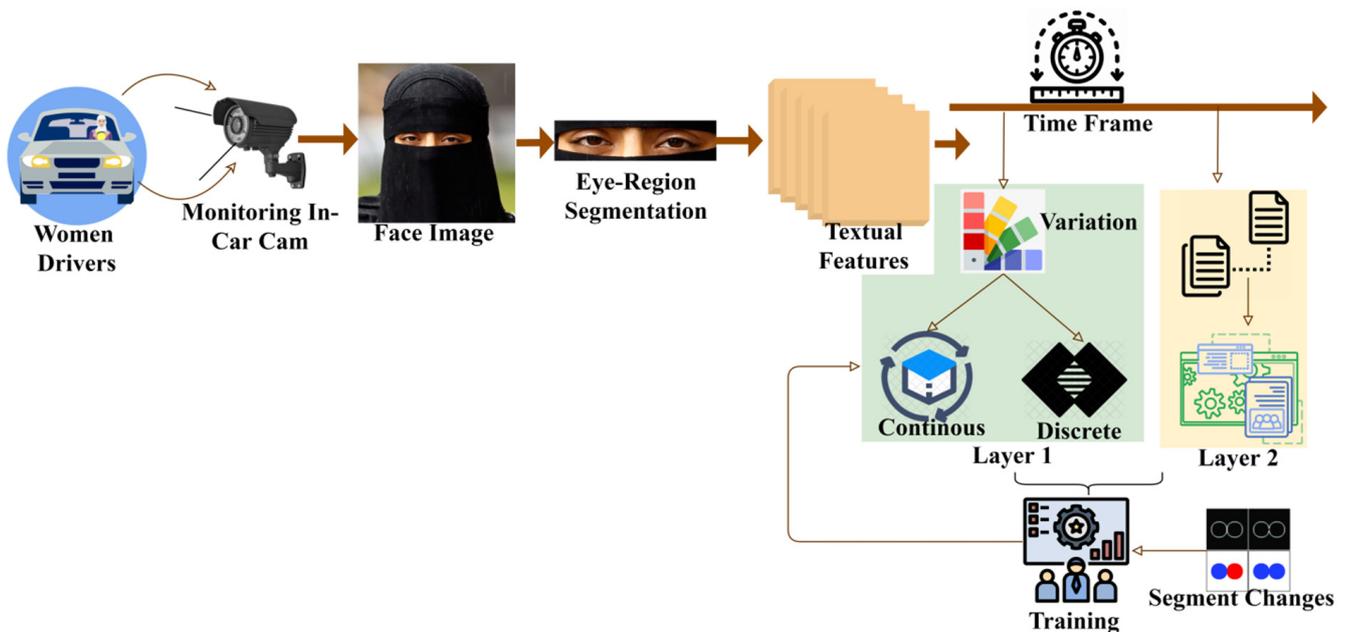


Figure 1. Proposed Method's Working Process.

The facial expressions and eye blinking of novice women drivers are observed through monitoring systems such as wearable sensors, mirrors, and in-car cams, etc., which are placed in vehicles for technical assistance and the physical observation of women drivers in Saudi Arabia. These observations are sensed as images/video frames for computer-based analyses, the identified event-occurring region is segmented, and then variation is identified through a textural extraction from different frames. Such observations are classified as discrete and continuous frames. In discrete frame observation, the identified events occur in the eye region. This region is segmented using variation pixel extraction, in which the extraction process identifies textural variations from different instances for the time/day. In continuous frame observation, the EBD of novice women drivers is said to be continuous for the time and day. This method employs a CNN with two hidden-layer processes for the frame differentiation and textural variation identification. The detection and extraction process prevents the chance of eye blinking by causing segment errors. The segment errors are identified as eye blink recognition. The proposed eye blink event detection addresses such segment errors in face image processing through pixel extraction using a convolutional neural network.

First, novice women drivers' facial expressions during driving instances are observed. Let $Eye_B(f)$ mean the sequence of the eye-blink detected frames that are observed in a different time interval. In this face image processing, the $F_{img}(f)$ of novice women drivers during different instances, Δ , is evaluated as:

$$F_{img}(\Delta) = Eye_B(\Delta) - Seg^e \times Eye_{rs}(\Delta) \quad (1)$$

In Equation (1), the variable Seg^e represents the eye region segmentation errors and the objective of this error minimization is defined as $\forall Eye_B(\Delta) \in F_{img}(\Delta)$ for detecting and classifying the identified events using face image processing. The image/video frames observed from novice women drivers are classified into two instances, such as discrete (f_d) and continuous (f_c). Therefore, the condition $\Delta = f_d + f_c$ is such that a discrete frame is detected between two continuous frames or vice versa. The identified event occurrence is unique for novice women drivers in Saudi Arabia. If the variable δ means the number of identified event occurrences, then $f_d = (\delta \times \Delta) - f_c$ is the discrete frame detected instances to be segmented from the face image. Let $f(f_d)$ and $f(f_c)$ represent the frames of $Eye_B(\Delta)$ that are observed with δ and Seg^e and detected in all the discrete frame detected instances, such that:

$$f(f_d) = \frac{Seg^e}{F_{img}(\Delta)} T \vee Seg^e \neq 0 \quad (2)$$

Additionally,

$$f(f_c) = \delta \cdot T \times F_{img}(\Delta) \vee Seg^e = 0 \quad (3)$$

The above Equations (2) and (3) represent the frames observed from the novice women driver face images, such that the observed frames are identified for segmenting the eye region based on extracting the textural features using $Eye_B(\Delta)$. T is represented as the first event detection. Now, based on the frames of the eye, parts are detected and segmented using Equation (2); therefore, the above Equation (1) is re-written as:

$$F_{img}(\Delta) = \begin{cases} f(f_d) = \delta \cdot T \times F_{img}(\Delta), & \text{if } Seg^e = 0 \\ f(f_c) - f(f_d) = \delta \cdot T \times F_{img}(\Delta) - \frac{Seg^e}{F_{img}(\Delta)} T, & \forall Seg^e \neq 0 \end{cases} \quad (4)$$

In Equation (4), the expanded face image processing and frames' detected sequence of $f_d \in T$ to be identified in the current instance addresses the first event detection in a discrete frame using textural feature extraction. The segmentation process is illustrated in Figure 2.

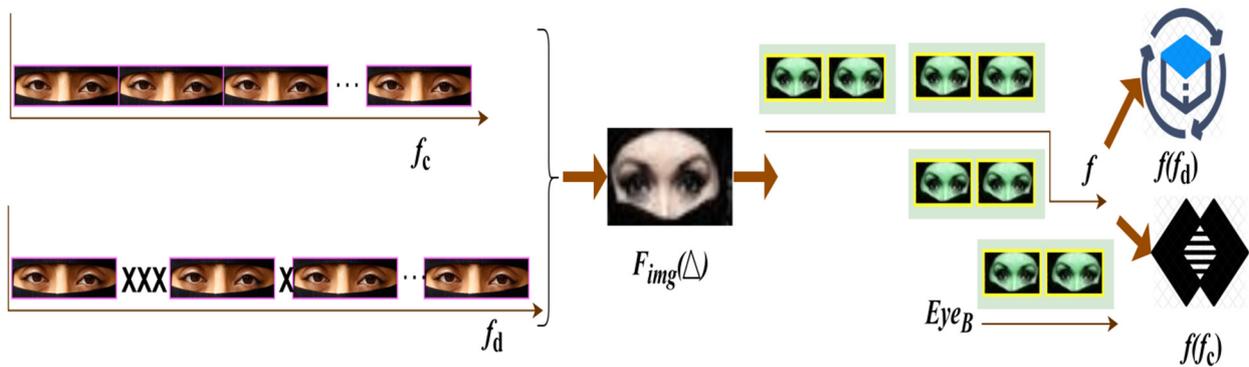


Figure 2. Segmentation Process.

The input image generates f_c and f_d , which are assimilated for the segment detection. In the segment detection, the face, eyes, and blink are extracted for improved accuracy. Based on $\Delta = f_d + f_c$, the missing sequences are pursued by either $f(f_d)$ or $f(f_c)$, for which consecutive frames are utilized. Therefore, the segmentation relies on Δ extracted from $F_{img}(f)$ (Refer to Figure 2). This eye-region segmentation identifies Seg^e based on variation detection, using a convolutional neural network and the ED-SF method. The variation pixel extraction and textural feature variation identified from the different frames using the face image are processed through the CNN. For this process, the different time frame sequences of $\delta \in f_d$ are expressed as:

$$\delta(f_d) = \left(1 - \frac{f_c(f \cdot T)}{\delta}\right) f_{d-1} + \sum_{i=1}^T \left(\frac{\left(1 - \frac{f_c}{Seg^e}\right)^{i-1} f_{d-1}}{(f \cdot T)} \right) \tag{5}$$

Equation (5) computes the previous sequence of event detection in a discrete frame at a different time interval f_{d-1} and the previous sequence of event detection in continuous frames at a different time interval $(f_c)^{i-1}$. The condition $\delta(f_d) = \left(1 - \frac{f_c(f \cdot T)}{\delta}\right) f_{d-1} + \sum_{i=1}^T \left(\frac{\left(1 - \frac{f_c}{Seg^e}\right)^{i-1} f_{d-1}}{(f \cdot T)} \right)$ is computed for identifying the eye blinking of novice women drivers driving in Saudi Arabia. Based on the frame detection, $F_{img}(\Delta) = f(f_c) - f(f_d)[1 - \delta(f_d)]$ is the final output for the $Seg^e \neq 0$ condition. The textural feature variation identifies (tex_{f_c}) and (tex_{f_d}) in both the continuous and discrete frames for the identified event detection in the first input image processing, which is computed as:

$$tex_{f_c} \simeq \frac{f(f_c) \times T}{\sum_{n \in T} [\delta \cdot f_c + Eye_B(\Delta)]_n} \tag{6}$$

$$tex_{f_d} \simeq \frac{f(f_c) * T + f(f_d) \times T}{\sum_{n \in T} (\delta \cdot f_c)_n \{ [1 - \delta(f_d)] \times f(f_c) \}_n} \tag{7}$$

where the variable n represents the number of event detections using frames and the above Equations (6) and (7) are used to identify the textural feature variations in the frames observed in the discrete and continuous sequences, which are saved as the knowledge base for future reference. In this first novice women drivers face image processing, tex_{f_c} , tex_{f_d} , $f(f_c)$, and $f(f_d)$ are the serving inputs for the CNN. Figure 3 presents the feature variation identification.

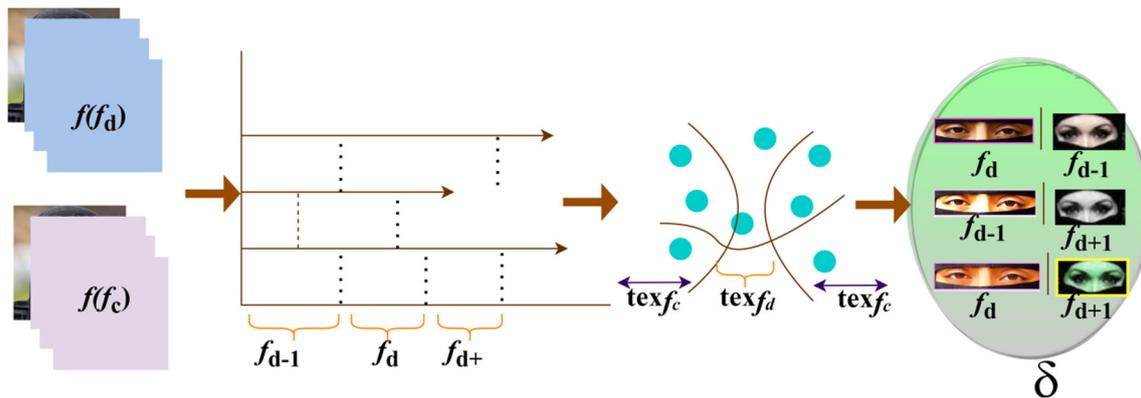


Figure 3. Feature Variation Identification.

The variation detection process relies on f_{d-1} and f_d until Δ reaches any f . If Δ is achieved, then tex_{f_c} and tex_{f_d} are segregated for identifying variations. The discrete combinations (f_d, f_{d-1}) , (f_{d-1}) , and (f_{d-1}, f_{d+1}) are separated from f_d and f_{d+1} for precision retention. Such sequences are used for extracting f_d and f_c to further identify $\Delta \in F_{img}$ (Refer to Figure 3). The consecutive processing of frames for eye-region segmentation helps to identify the segment error and variation error in the pixel and textural feature extraction between the continuous and discrete sequences. This CNN process is discussed in the following section.

4. Neural Network Process for Event Detection

In the pixel extraction process, the convolutional neural network is used to identify the correctness of the discrete–continuous textural features and detect the correctness of the discrete and continuous textural features and segment errors in the face image processing. As this process relies on already stored observations of the $Eye_B(\Delta)$, the maximum detection precision is achievable. The number of pixels may vary for all novice women drivers, though the knowledge base helps to classify and detect the eye blink segment for both instances at different time intervals. In particular, the CNN process has two hidden-layer processes: frame differentiation and eye-region segmentation. In the frame difference estimation, the discrete and continuous frames are detected to improve the stored observation of $Eye_B(\Delta)$. Instead, in the eye-region segmentation, a different textural variation is observed in $Eye_B(\Delta)$, which is used to augment the condition $F_{img}(\Delta)$ along with a better extraction process and the detection of segment errors. As per the process, the serving inputs for the discrete and continuous frame differentiation are $Eye_B(\Delta)$ and T . The computation of $Eye_B(\Delta) \in T$ is extracted under the discrete and continuous frames’ observed sequences, depending upon the occurrence of the identified event detection.

In the pixel extraction processes, if the frames are to be continuous in the event detection, then $Eye_B(\Delta)$ and the time frames are classified independently, with two hidden-layer processes. These two hidden-layer processes are performed for the discrete and continuous frames, after which, the CNN is used to update the previous event detection observation. The two hidden-layer outputs, $(H1_1$ to $H1_T)$ and $(H2_1$ to $H2_T)$, are computed as shown in Equations (8) and (9).

$$\left. \begin{aligned}
 H1_1 &= f_{c_1} \\
 H1_2 &= 2f_{c_2} - 2(f_d)_2 - f(f_c)_1 \\
 H1_3 &= 3f_{c_3} - 3(f_d)_3 - f(f_c)_3 \\
 &\vdots \\
 H1_T &= \delta f_{c_T} - \delta(f_d)_T - f(f_c)_{T-1}
 \end{aligned} \right\} \text{Layer 1 output} \tag{8}$$

$$\left. \begin{aligned}
 H2_1 &= f_{c_1} \\
 H2_2 &= 2(f_d) + f(f_c)_1 \\
 H2_3 &= 3(f_d) + f(f_c)_2 - f(f_d)_1 \\
 &\vdots \\
 H2_T &= \delta(f_d) + f(f_c)_{T-1} - f(f_d)_{T-2}
 \end{aligned} \right\} \text{Layer 2 output} \tag{9}$$

The two hidden-layer processes generate frame differentiation and eye-region segmentation outputs. The pixel extraction is processed using a convolutional neural network based on the time frame for the identified event occurrence detection. The condition $T \in H1$ must not be equal to $T \in H2$, which is the textural variation identification condition. If an identified event is observed from the continuous frames in the first image process, then $H2$ is processed for the eye-region segmentation. In this process, the time frames are divided as per the segment changes, and then $\delta(f_d) + f(f_c)_{T-1} - f(f_d)_{T-2}$ is the pixel extraction sequence of the discrete frames. The continuous and discrete frames are identified using the differentiation process of the CNN. In this differentiation process, the comparison of $H1_T$ and $H2_T$ is computed such that $f_c = \{f_c \cup f(f_c)\}$ and $f_d = \{H1_T \cup H2_T \cap f(f_d)\}$ are extracted independently for the precise recognition of the eye blinks of novice women drivers in Saudi Arabia. The eye-region segmentation with the previous knowledge base used for achieving the first layer, from which the textural variations are identified from the sequence, is grouped for verification. After the differentiation process, the frames are discrete, and then the textural features are compared with T based on $f(f_d)$ in the identified event-occurring sequence. Here, the serving inputs are $f(f_d)$ and f_d for the segment changes verification, and these inputs serve as the training for all the time-frame-classified sequences under f_d . The first hidden-layer process is illustrated in Figure 4.

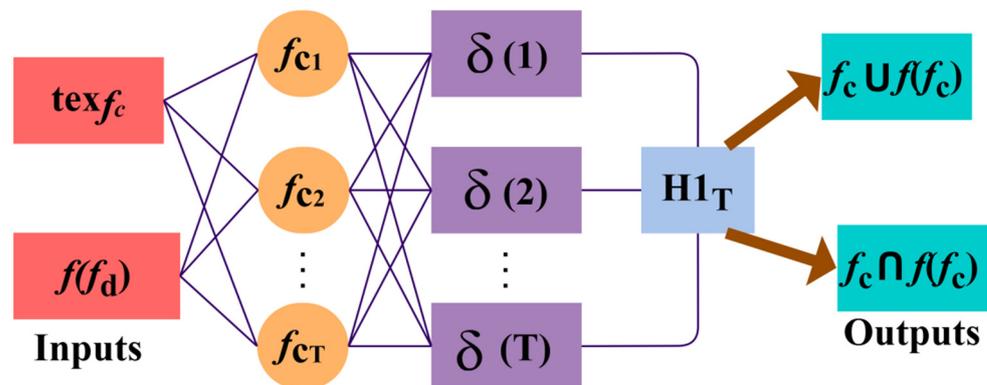


Figure 4. First hidden-layer process.

In this process, the frame differentiation is first identified for a precise eye-blink segmentation. In this differentiation process, the discrete and continuous frame differentiation is identified for the segment changes identification with f_d and $Eye_B(\Delta)$. If f_d and $Eye_B(\Delta)$ denote the time frame sequences, (i.e.,), if $tex_{f_c} < tex_{f_d}$, then the decision matrix value is 1. Instead, if $tex_{f_c} > tex_{f_d}$ is true, then the decision matrix value is inverted; therefore, the new sequence of the continuous frame is further extracted for its features, and the segment changes are identified under two hidden-layer processes, where $T \in f_c$. From the above condition, if $tex_{f_c} > tex_{f_d}$ is marked as “1”, then $tex_{f_c} < tex_{f_d}$ is marked as “0”. Based on the CNN output, if the identified event occurrence of 0 is detected, then discreteness is achieved. Instead, if the identified event occurrence of 1 is detected, then the consequence is achieved. Hence, the differentiation of the discrete and continuous frames initiates the two hidden-layer processes, as per the above equation. Now, the segment changes are identified, and the training will be performed by the CNN process for updating the knowledge base, as per the CNN process of eye blink recognition using the proposed ED-SF method. In this sequential process, first, the event identified f_d is used for the pixel

extraction process based on $f_c \times f(f_c)$ and $f_c \times f(f_d)$, which are the final decision matrix representations, respectively. In the extraction process, the time frames serve as the inputs for the variation identification to classify the events with the extracted textural features from f_d . In the eye-region segmentation retrieval process, the condition $(H2_T, f_d)$ is fetched for verifying the texture luminosity and identifying the segment errors. If f_c occurs, then tex_{f_c} is computed as in Equation (6). Instead, if f_d occurs, the discrete frames are grouped for a continuous sequence. Here, the two hidden-layer processes and event occurrences, which are different for all novice women drivers, are analyzed using the pixel extraction process. Therefore, based on the segment changes and training, the knowledge base is updated with the current eye-region segmentation process and computed as:

$$\left. \begin{aligned} H1_1 &= f_{d_1} \\ H1_2 &= 2f_{d_2} + f(f_d)_1 + Seg_1^e \\ H1_3 &= 3f_{d_3} + f(f_d)_2 + Seg_2^e \\ &\vdots \\ H1_T &= T \cdot f_{d_T} + f(f_d)_{T-1} + Seg_T^e \end{aligned} \right\}, \text{ for the discrete frames} \tag{10}$$

$$\left. \begin{aligned} H1_1 &= 0 \\ H1_2 &= f(f_c)_1 + 2f_{c_1} - Seg_1^e \\ H1_3 &= f(f_c)_2 + 3f_{c_2} - Seg_1^e \\ &\vdots \\ H1_T &= f(f_c)_{T-1} + T \cdot f_{c_T} - Seg_T^e \end{aligned} \right\}, \text{ for the continuous frames} \tag{11}$$

The above Equations (10) and (11) compute the identified event occurrences in both the discrete and continuous frames and are analyzed for novice women drivers. The segment changes are also identified for training. The continuous and discrete frame differentiation achieves $H1_1 = 0$, as the previous eye-region segmentation used the case of a discrete frame sequence, and the update for f_c is zero. Therefore, this is not considered in the pixel extraction process. In both instances, the segment error increases, before which, the knowledge base is updated, as in Equation (12).

$$\left. \begin{aligned} KU_1 &= \frac{1(f_{d1})+f_{c-1}}{Seg_1^e} \\ KU_2 &= \frac{2(f_{d2})+f_{c-2}}{Seg_2^e} - f(f_{d-1}) \\ KU_3 &= \frac{3(f_{d3})+f_{c-3}}{Seg_3^e} - f(f_{d-2}) \\ &\vdots \\ KU_T &= \frac{T(f_{dT})+f_{c-T}}{Seg_T^e} - f(f_{d-T-1}) \end{aligned} \right\} \tag{12}$$

The stored knowledge base update is processed at the end of all the discrete frame sequences or before the start of the next continuous frame sequence. In Figure 5, the layer two process for the training is illustrated.

The two outputs, $H1_T$ and $H2_T$, are cross-validated for f_c and $f_d \in T$ until Δ . In this process, based on f_c and f_d , the δ and Seg^e occurrences are extracted. The extraction process is tuned to generate f_c other than f_d from the layer one analysis. Both layers are jointly idolized for knowledge updates; $f_c \in H1_T$ and $H2_T \in f_d$ occur between f_d and f_{d+2} . If $f_c \in H2_T$, then the T extraction (identifying) sequence is performed. Contrarily, if f_d is the output, then Seg^e is trained until Δ and the layer one process (as in Figure 4) generates a precise $H1_T$ (Figure 5). From the knowledge base update, the segment changes for the frames are used to identify the facial expressions and eye blinks of novice women drivers in different time frames. If the segment errors and textural variations are not identified, then the discrete and continuous frames will be grouped using the CNN, resulting in a high EBD precision.

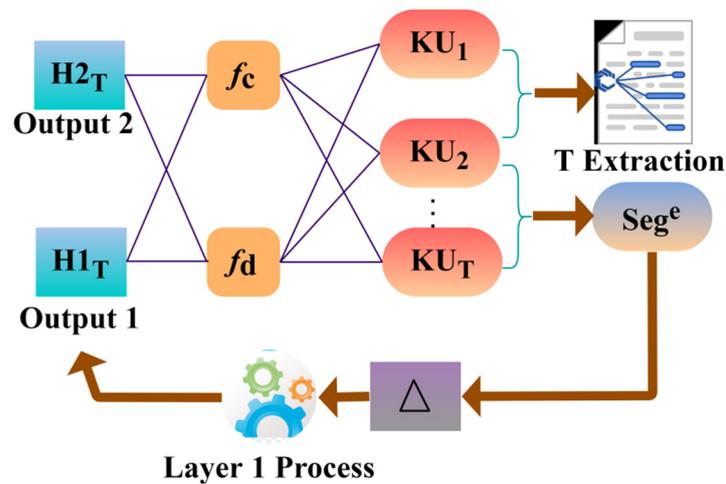


Figure 5. Layer 2 process for training.

5. Experimental Analysis

The experimental analysis is performed using the data provided in [33]. This dataset provides nearly 12,000 niqab-covered faces, from which 6000+ images are used for testing and 1200 images are used for training the $H1$ and $H2$. Using these data as a reference, a sample representation of continuous and discrete instances of f is illustrated in Figure 6.



Figure 6. Continuous and discrete sequence representations.

This illustration is presented here for understanding the sequences from which the variations are observed. The frames and segmentation processes performed for the above sequences are independently described in the following tables. Tables 2 and 3 present the segmentation processes for the continuous- and discrete-sequence-observed inputs.

Table 2. Segmentation for continuous sequence.

Input Image	Segmented Image	Variation Plot

Based on the variation plot, the segment changes are observed; the $H1$ and $H2$ processes are performed. The variations are observed for preventing errors in detecting the events and forwarded segments. Table 4 presents the actual extracted and the variation feature outputs.

Table 3. Segmentation for discrete sequence.

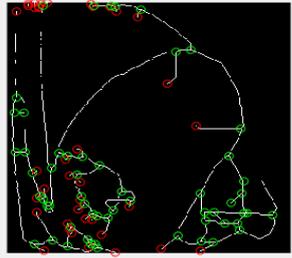
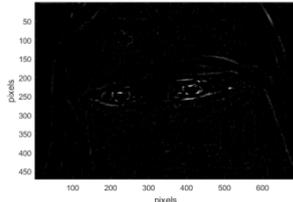
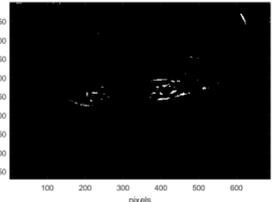
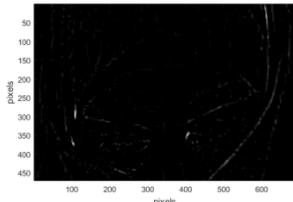
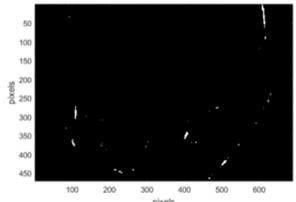
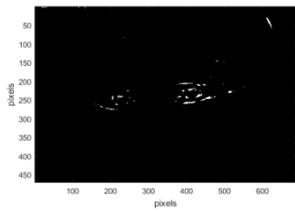
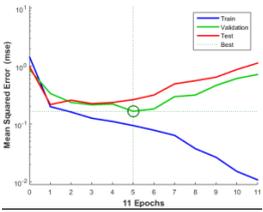
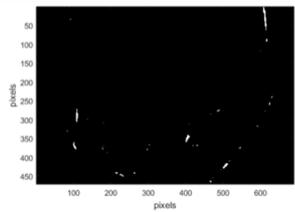
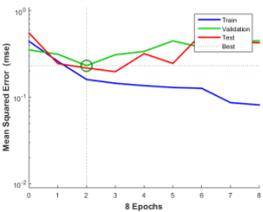
Input Image	Segmented Image	Variation Plot
		

Table 4. Actual and variation feature outputs.

Sequence	Segmented Input	Actual ($H1$)	Extracted ($H2$)
f_c			
f_d			

Based on $H2$, the final error estimations from the experimental analysis are presented in Table 5.

Table 5. Error estimation.

Sequence	$H2$ Output	Error
f_c		
f_d		

6. Performance Assessment Using Comparative Analysis

The performance assessment is carried out using MATLAB experiments by varying the frames and segments. The metrics of the event detection, detection precision, sensitivity, variation error, and detection time are compared with existing methods using the results

obtained. The methods of HCNN + LSTM [27], 2s-STGCN [21], and CDFDM [25] are used for the combined analysis.

6.1. Event Detection

The identified event detection was high in this proposed method based on eye-region segmentation and textural feature extraction in the face image processing of novice women drivers in Saudi Arabia using in-car cam monitoring systems. This can be observed from the different instances depicted in Figure 7. This proposed ED-SF method satisfies women drivers that are less experienced and lacking in road safety knowledge by checking the different texture luminosities, using different inputs for identifying the eye blinks of women drivers. This process performed abnormal eye blink recognition through frame differentiation and textural feature variation computation at different time intervals for pixel extraction in image processing. The condition $\forall Eye_B(\Delta) \in F_{img}(\Delta)$ was defined for detecting and classifying the identified event occurrences in the discrete and continuous frames and analyzed using face image processing. The facial expressions were recognized until an identified abnormal event occurred, which was detected via physical observations and technical assistance. Therefore, the discrete and continuous frame differentiation was identified using the first hidden-layer processing for improving EBD. Hence, the segment errors detected using the CNN in different frames achieved a high event detection.

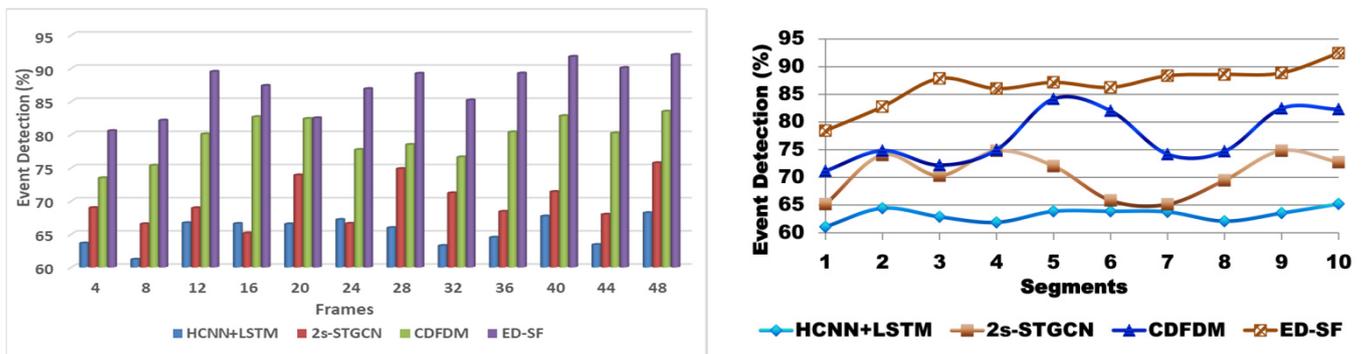


Figure 7. Event detection.

6.2. Detection Precision

Detection precision as a measure is computed using the correctly identified positive instances from the range of entire instances that the system recognized as positive. The detection precision concentrates on the accuracy of the positive prediction, which is measured as follows.

$$\text{Detection precision} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})} \tag{13}$$

Figure 8 represents the eye-region segmentation from both the discrete and continuous frames with abnormal event detections for identifying the eye blinks of women drivers. In the proposed method, the identified event detection was maximized through the proposed method for driving experience, and novice women drivers were highly exposed to vehicle-handling features. The first event detection occurrence from the discrete frame instances was identified in the face image processing. If any eye blink was identified in the image processing, training was provided to the women drivers for improving their practices. From the instances, the discrete frames were observed from the novice women drivers, such that the segment variation was identified and then the pixel extraction and textural feature variation were computed for $Eye_B(\Delta)$. The first-layer output and second-layer output were combined to improve the EBD precision and classify such event-occurring frames using the image processing. Detecting such segment errors in the face image processing maximized the pixel extraction for the training, such that a high EBD precision was achieved using the pixel extraction process.

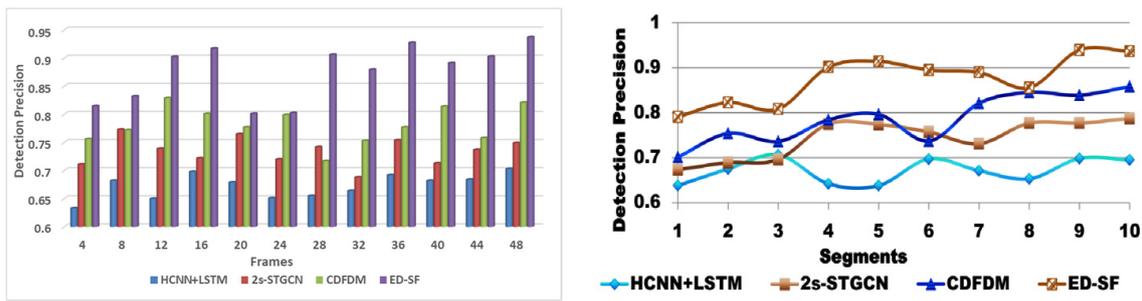


Figure 8. Detection precision.

6.3. Sensitivity

Sensitivity is computed from the correctly identified positive instances from the entire actual positive instances database. It concentrates on the how effectively the system predicted the positive features.

$$\text{Sensitivity} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})} \tag{14}$$

Figure 9 illustrates the sensitivity value of analyzing and detecting the EBD of novice women drivers in Saudi Arabia. Based on the eye-region segmentation process, the variation in the pixel extraction was identified for the precise detection of an eye blink. Both the discrete and continuous frames were separated for segmenting the eye part from the face image through a convolution neural network at different time intervals. This eye-region segmentation was processed for identifying Seg^e based on the textural feature variation identification using a convolutional neural network and the ED-SF method. The variation was identified in the pixel extraction and segment changes were identified from the different frames throughout the face image processing at different time intervals. There were two hidden-layer outputs in the sequential event detection from the continuous frames with the pixel extraction process for the precise eye-region segmentation without errors. Based on the continuous and discrete frame differentiation identification using one layer, a high sensitivity in the eye-region segmentation was satisfied by using the proposed method.

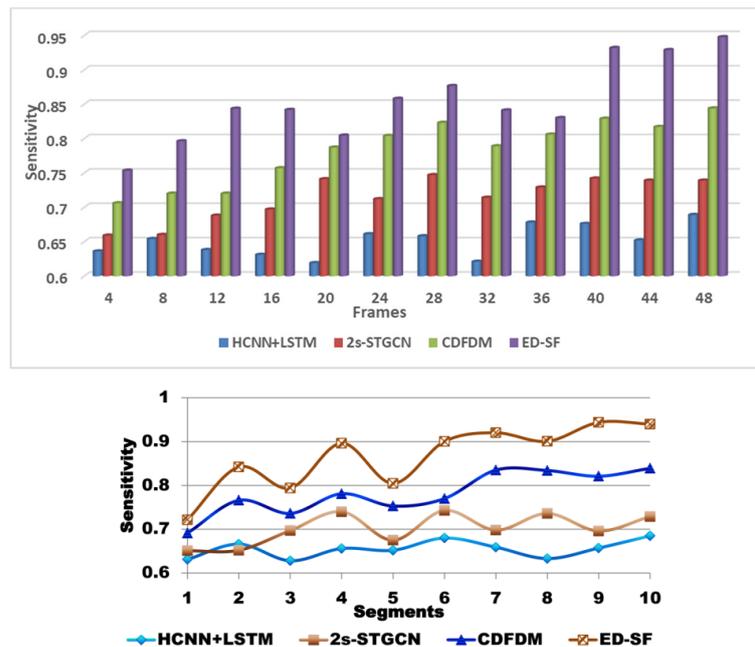


Figure 9. Sensitivity.

6.4. Variation Error

The variation error was computed by taking the difference between the actual and predicted output values. This proposed method using eye blink recognition reduced the occurrence of identified events and segment errors in the face image processing of novice women drivers in Saudi Arabia through a CNN, and did not contain the continuous frame sequence at any time interval. The textural feature variations were identified from the time frames observed in the discrete and continuous instances for improving the eye-region segmentation. In the first novice women drivers face image processing, the evaluation of tex_{f_c} , tex_{f_d} , $f(f_c)$, and $f(f_d)$ was processed for EBD using the CNN. The segment changes were identified for preventing errors in the pixel extraction process in the face image processing of women drivers for the $f_c = \{f_c \cup f(f_c)\}$ and $f_d = \{H1_T \cup H2_T \cap f(f_d)\}$ conditions. The time frames were differentiated for improving the textural feature variation identification using a convolutional neural network. If the occurrence of an identified event was observed from the continuous frames in the first image process, then layer two was processed for the eye-region segmentation. Based on the pixel extraction process, EBD was improved through the face image processing, preventing segment errors. The proposed method requires vehicle usage information, in which the proposed method achieves a minimum variation error, as represented in Figure 10.

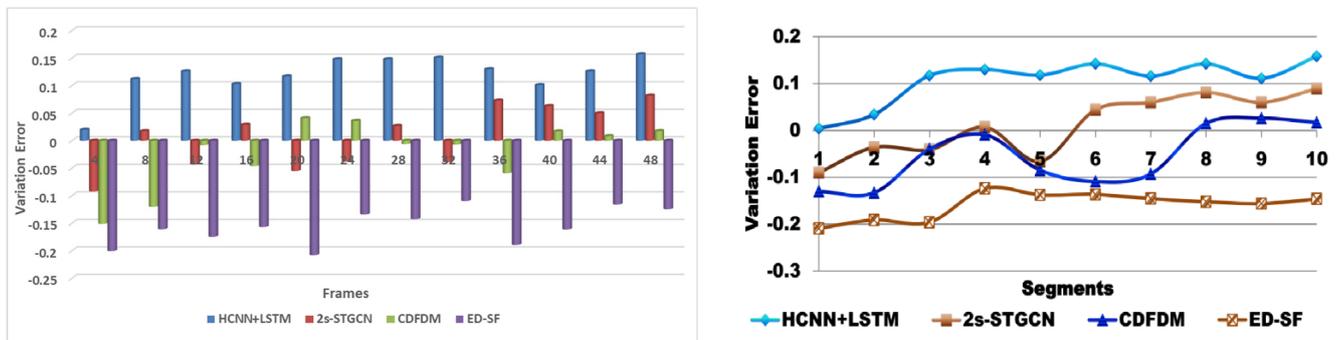


Figure 10. Variation error.

6.5. Detection Time

The differentiation of the discrete and continuous frames was identified for improving the texture luminosity in the precise eye-region segmentation of women drivers' face image processing and is illustrated in Figure 11. In this event detection, using segmented frames for precise EBD, less detection time and segment errors were satisfied through a computer-based analysis for the differentiation of the discrete and continuous frames, initiating the two hidden-layer processes, as per the above equation. In this CNN process, the segment errors and textural feature variation were identified in using the conditions f_d and $Eye_B(\Delta)$ at different time frame sequences. In this process, the discrete and continuous frame differentiation was detected for identifying the segment changes with f_d and $Eye_B(\Delta)$. The precise event detection was processed using the variation pixel extraction process from the different frames using the CNN. In this proposed identified event detection, the EBD improved through using a convolutional neural network with two hidden-layer processes. In this article for segmenting the eye region devouring the textural variation, the detection time was less at a different time frames.

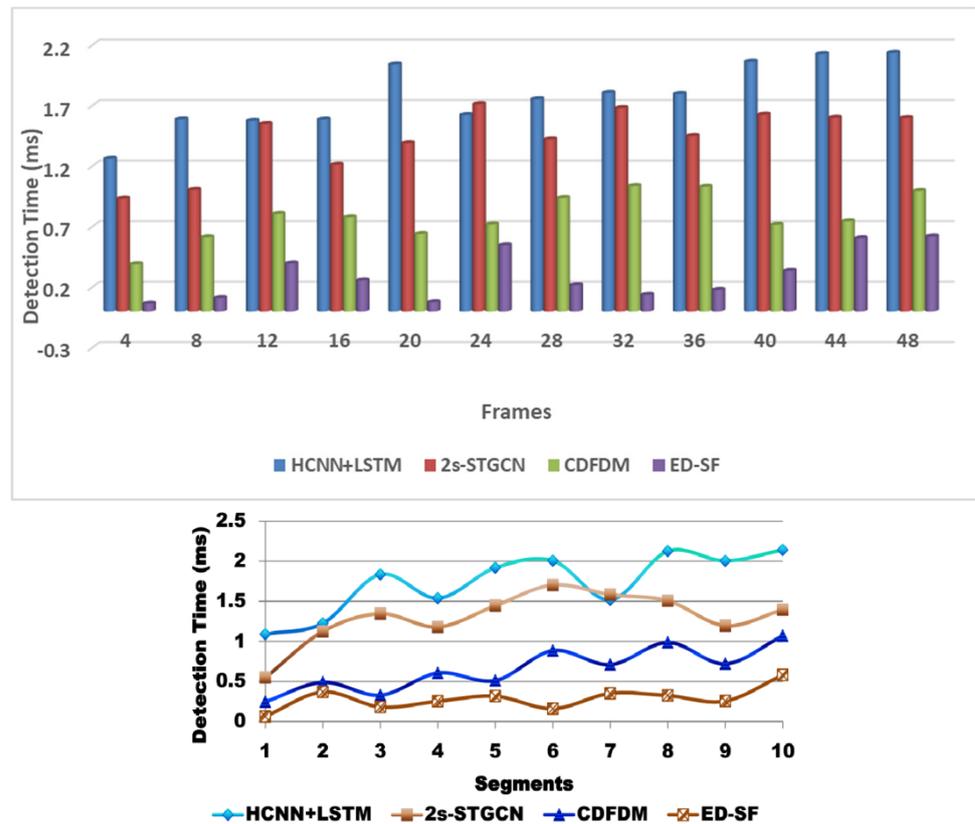


Figure 11. Detection time.

7. Conclusions

Saudi Arabian women are novices to driving, as the ban was uplifted in 2018. Therefore, pursuing driving rules is novel for them. Considering this fact, eye-region-based fatigue or drowsiness detection was presented in this article. Events such as eye blinks or eye closing were identified for improving women drivers’ safety. Computer-vision-based event detection through segmented frames was exclusively designed for abnormality identification. This method relied on the textural variations observed from the discrete and continuous sequences through continuous monitoring. The input image was analyzed using two layers of a convolution neural network. In the first layer, the discrete and continuous frames for the observed sequences were verified and pursued by the precise frame extraction in the second layer. This provided a segmented image, regardless of the textures and features extracted in each frame. The error-causing variation sequences were identified and used for training the neural network in order to prevent additional errors in the upcoming sequences. Therefore, the proposed method was reliable in identifying the eye-blink events for novice women drivers to provide safety assistance.

Findings: From the comparative analysis, the proposed method improved the event detection by 9.5%, detection precision by 7.84%, and sensitivity by 9.46%, and reduced the variation error by 13.32% and detection time by 10.37% compared to other models. Thus, the proposed method improved the EBD accuracy by 9.5% compared to HCNN + LSTM, 2s-STGCN, and CDFDM

Future Work: However, the introduced method failed to recognize the eye blinks by detecting faces. Inspired by the recent trends in face detection, facial-expression-based driver assistance and safety regulations are planned to be incorporated into future work. A face covered with a Niqab or mask is to be considered for this purpose for differentiating facial expressions along with event detection.

Author Contributions: Methodology, M.S.A.-R.; Validation, Y.A.A.; Formal analysis, I.A.; Investigation, I.A.; Resources, M.S.A.-R.; Data curation, Y.A.A.; Writing—original draft, Y.A.A.; Writing—review & editing, I.A.; Supervision, M.S.A.-R.; Project administration, M.S.A.-R. and Y.A.A.; Funding acquisition, M.S.A.-R. All authors have read and agreed to the published version of the manuscript.

Funding: The authors present their appreciation to King Saud University for funding this research through Researchers Supporting Program number (RSP2023R206), King Saud University, Riyadh, Saudi Arabia.

Data Availability Statement: The data that support the findings of this study are openly available in [33].

Acknowledgments: The authors present their appreciation to King Saud University for funding this research through Researchers Supporting Program number (RSP2023R206), King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jastania, Z.; Abbasi, R.A.; Aslam, M.A.; Khanzada, T.J.S.; Ghorri, K.M. Analyzing Public Discussions about #SaudiWomenCanDrive Using Network Science. *IEEE Access* **2022**, *10*, 4739–4749.
2. Al-Garawi, N.; Kamargianni, M. Exploring the factors affecting women’s intention to drive in Saudi Arabia. *Travel Behav. Soc.* **2022**, *26*, 121–133. [[CrossRef](#)]
3. Al-Razgan, M.; Alrowily, A.; Al-Matham, R.N.; Alghamdi, K.M.; Shaabi, M.; Alssum, L. Using diffusion of innovation theory and sentiment analysis to analyze attitudes toward driving adoption by Saudi women. *Technol. Soc.* **2021**, *65*, 101558. [[CrossRef](#)]
4. Sattari, N. Women driving women: Drivers of women-only taxis in the Islamic Republic of Iran. *Women’s Stud. Int. Forum* **2020**, *78*, 102324. [[CrossRef](#)]
5. Al-Garawi, N.; Kamargianni, M. Women’s modal switching behavior since driving is allowed in Saudi Arabia. *J. Transp. Geogr.* **2021**, *96*, 103192. [[CrossRef](#)]
6. Jannusch, T.; Shannon, D.; Völler, M.; Murphy, F.; Mullins, M. Cars and distraction: How to address the limits of Driver Monitoring Systems and improve safety benefits using evidence from German young drivers. *Technol. Soc.* **2021**, *66*, 101628. [[CrossRef](#)]
7. Corcoba, V.; Paneda, X.G.; Melendi, D.; García, R.; Pozueco, L.; Paiva, S. COVID-19 and its effects on the driving style of spanish drivers. *IEEE Access* **2021**, *9*, 146680–146690. [[CrossRef](#)]
8. Vasconez, J.P.; Viscaino, M.; Guevara, L.; Cheein, F.A. A fuzzy-based driver assistance system using human cognitive parameters and driving style information. *Cogn. Syst. Res.* **2020**, *64*, 174–190. [[CrossRef](#)]
9. Shi, G.; Gao, C.; Wang, D.; Su, Z. Automatic 3D virtual fitting system based on skeleton driving. *Vis. Comput.* **2021**, *37*, 1075–1088. [[CrossRef](#)]
10. Buzon, L.G.; Figueira, A.C.; Larocca, A.P.C.; Oliveira, P.T.M. Effect of speed on driver’s visual attention: A study using a driving simulator. *Transp. Dev. Econ.* **2021**, *8*, 1. [[CrossRef](#)]
11. Sohn, K.; Jang, G. Ground Vehicle Driving by Full Sized Humanoid. *J. Intell. Robot. Syst.* **2020**, *99*, 407–425. [[CrossRef](#)]
12. Giot, C.; Hay, M.; Chesneau, C.; Pigeon, E.; Bonargent, T.; Beauflis, M.; Chastan, N.; Perrier, J.; Pasquier, F.; Polvent, S.; et al. Towards a new approach to detect sleepiness: Validation of the objective sleepiness scale under simulated driving conditions. *Transp. Res. Part F Traffic Psychol. Behav.* **2022**, *90*, 109–119. [[CrossRef](#)]
13. Bitkina, O.V.; Park, J.; Kim, H.K. The ability of eye-tracking metrics to classify and predict the perceived driving workload. *Int. J. Ind. Ergon.* **2021**, *86*, 103193. [[CrossRef](#)]
14. Li, K.; Gong, Y.; Ren, Z. A fatigue driving detection algorithm based on facial multi-feature fusion. *IEEE Access* **2020**, *8*, 101244–101259. [[CrossRef](#)]
15. Chen, W.; Sawaragi, T.; Hiraoka, T. Comparing eye-tracking metrics of mental workload caused by NDRTs in semi-autonomous driving. *Transp. Res. Part F Traffic Psychol. Behav.* **2022**, *89*, 109–128. [[CrossRef](#)]
16. Xue, Z.; Chen, L.; Liu, Z.; Lin, F.; Mao, W. Rock segmentation visual system for assisting driving in TBM construction. *Mach. Vis. Appl.* **2021**, *32*, 77. [[CrossRef](#)]
17. Cori, J.M.; Turner, S.; Westlake, J.; Naqvi, A.; Ftouni, S.; Wilkinson, V.; Vakulin, A.; O’Donoghue, F.J.; Howard, M.E. Eye blink parameters to indicate drowsiness during naturalistic driving in participants with obstructive sleep apnea: A pilot study. *Sleep Health* **2021**, *7*, 644–651. [[CrossRef](#)]
18. Kır Savaş, B.; Becerikli, Y. Behavior-based driver fatigue detection system with deep belief network. *Neural Comput. Appl.* **2022**, *34*, 14053–14065. [[CrossRef](#)]
19. Jordan, A.A.; Pegatoquet, A.; Castagnetti, A.; Raybaut, J.; Le Coz, P. Deep learning for eye blink detection implemented at the edge. *IEEE Embed. Syst. Lett.* **2020**, *13*, 130–133. [[CrossRef](#)]
20. Mou, L.; Zhou, C.; Xie, P.; Zhao, P.; Jain, R.C.; Gao, W.; Yin, B. Isotropic Self-supervised Learning for Driver Drowsiness Detection With Attention-based Multi-modal Fusion. *IEEE Trans. Multimed.* **2021**, *25*, 529–542. [[CrossRef](#)]

21. Bai, J.; Yu, W.; Xiao, Z.; Havyarimana, V.; Regan, A.C.; Jiang, H.; Jiao, L. Two-stream spatial–temporal graph convolutional networks for driver drowsiness detection. *IEEE Trans. Cybern.* **2021**, *52*, 13821–13833. [[CrossRef](#)]
22. Li, X.; Schroeter, R.; Rakotonirainy, A.; Kuo, J.; Lenné, M.G. Effects of different non-driving-related-task display modes on drivers' eye-movement patterns during take-over in an automated vehicle. *Transp. Res. Part F Traffic Psychol. Behav.* **2020**, *70*, 135–148. [[CrossRef](#)]
23. Liang, N.; Yang, J.; Yu, D.; Prakah-Asante, K.O.; Curry, R.; Blommer, M.; Swaminathan, R.; Pitts, B.J. Using eye-tracking to investigate the effects of pre-takeover visual engagement on situation awareness during automated driving. *Accid. Anal. Prev.* **2021**, *157*, 106143. [[CrossRef](#)]
24. Akrouf, B.; Mahdi, W. A novel approach for driver fatigue detection based on visual characteristics analysis. *J. Ambient Intell. Humaniz. Comput.* **2023**, *14*, 527–552. [[CrossRef](#)]
25. Zeng, L.; Zhou, K.; Han, Q.; Wang, Y.; Guo, G.; Ye, L. An fNIRS labeling image feature-based customized driving fatigue detection method. *J. Ambient. Intell. Humaniz. Comput.* **2022**, 1–17. [[CrossRef](#)]
26. Yamabe, S.; Kawaguchi, S.; Anakubo, M. Comfortable awakening method for sleeping driver during autonomous driving. *Int. J. Intell. Transp. Syst. Res.* **2022**, *20*, 266–278. [[CrossRef](#)]
27. Guo, J.M.; Markoni, H. Driver drowsiness detection using hybrid convolutional neural network and long short-term memory. *Multimed. Tools Appl.* **2019**, *78*, 29059–29087. [[CrossRef](#)]
28. Li, D.; Zhang, X.; Liu, X.; Ma, Z.; Zhang, B. Driver fatigue detection based on comprehensive facial features and gated recurrent unit. *J. Real-Time Image Process.* **2023**, *20*, 19. [[CrossRef](#)]
29. Wijnands, J.S.; Thompson, J.; Nice, K.A.; Aschwanden, G.D.; Stevenson, M. Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks. *Neural Comput. Appl.* **2020**, *32*, 9731–9743. [[CrossRef](#)]
30. Cui, Z.; Sun, H.M.; Yin, R.N.; Gao, L.; Sun, H.B.; Jia, R.S. Real-time detection method of driver fatigue state based on deep learning of face video. *Multimed. Tools Appl.* **2021**, *80*, 25495–25515. [[CrossRef](#)]
31. Ghosh, R.; Phadikar, S.; Deb, N.; Sinha, N.; Das, P.; Ghaderpour, E. Automatic Eye-blink and Muscular Artifact Detection and Removal from EEG Signals Using k-Nearest Neighbor Classifier and Long Short-Term Memory Networks. *IEEE Sens. J.* **2023**, *23*, 5422–5436. [[CrossRef](#)]
32. Egambaram, A.; Badruddin, N. An Investigation to Detect Driver Drowsiness from Eye blink Artifacts Using Deep Learning Models. In Proceedings of the 2022 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia, 7–9 December 2022; pp. 23–29. [[CrossRef](#)]
33. Alashbi, A.; Sunar, M.S.; Alqahtani, Z. Deep-Learning-CNN for Detecting Covered Faces with Niqab. *J. Inf. Technol. Manag.* **2022**, *14*, 114–123.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.