Detection method of eyes opening and closing ratio for driver's fatigue monitoring

Qijie Zhao1,2 | Junye Jiang1 | Zhigao Lei1 | Jingang Yi3

1 School of Mechatronic Engineering and Automation, Shanghai Key Laboratory of Intelligent Manufacturing and Robotics, Shanghai University, Shanghai, China
2 Shanghai Key Laboratory of Intelligent Manufacturing and Robotics, Shanghai, China
3 Department of Mechanical and Aerospace Engineering, Rutgers University, Piscataway, New Jersey, United States of America

Abstract
Eyes opening and closing status is one of the most important components to monitor the driver's fatigue. The current research mainly considers eyes blink frequency and the closing duration to judge the driver's fatigue. To identify driver's fatigue level, eyes opening and closing ratio (EOCR) is a critical factor and therefore, it is desirable to detect EOCR for driver's fatigue monitoring. The proposed method aims to simultaneously segment images and measure the parameters of the EOCR. A BiSeNet-based iris and pupil segmentation network is first proposed and the Visual Geometry Group (VGG) ConvNet-based model to detect the EOCR value is provided by considering the main features rounding eyes area and the iris-pupil size for building test dataset. The comparison experiments are conducted with the proposed method and the other existing work in different datasets, such as CASIA-Iris-Thousand, CASIA-Iris-Interval, and UBIRIS.v2. The results demonstrate that the proposed method has superior detection effects on both infrared images and colour images than other existing approaches. Furthermore, the experiments of detecting the EOCR and iris and pupil segmentation are carried out with the test dataset and the results show that the proposed method can reliably identify driver's eye opening and closing degree.

1 INTRODUCTION

With the continuously increasing number of the cars on the road, traffic safety has become a world widely concerned problem. Driver fatigue account for a large proportion of road traffic accidents. Therefore, it is of great significance to warn drivers to pay attention to safety through effective fatigue detection methods. Among many driver fatigue detection methods, machine vision-based detection methods collect the driver's face image and estimate the driver's fatigue level by the camera installed in front of drivers. These cameras do not interfere with the driver's driving behaviour and computer vision is a desirable fatigue detection method.

To identify driver's fatigue state, eyes blink frequency is often considered, and a metric called PERCLOS (defined as the percentage of eyelid closure over the pupil over time) is widely adopted [1]. A driver is generally considered fatigue if PERCLOS value is at least 80%. Obviously, in this situation, the driver is already in severe fatigue state and it is desirable and helpful to detect the driver's fatigue level and give early fatigue warning before it reaches to large PERCLOS values. In order to get the driver's fatigue levels, additional eye features detection methods and metrics are needed. Eyes opening and closing ratio reflects a person's mind status and attention [2] and therefore, it can be potentially used to indicate driver fatigue levels. The ratio of iris and pupil area occluded precisely captures eye opening and closing ratio and can thus be used to predict fatigue levels. The work in [3] locate eyes region through facial landmarks and detect the eyes state by the ratio of the height to the width of the eyes. The accuracy of the method is limited by the facial landmark detection and biomechanic parameter values, such as eyes shapes. Because of these reasons, this method can be used to only determine whether the eye is open or not. Reference [4] uses principal components analysis (PCA) and linear discriminant analysis (LDA) algorithms to reduce image dimensions and support vector machine to classify the eye is whether open. However, the method cannot measure human eyes opening and closing ratio accurately. Although the above-mentioned methods...
are effective to determine eyes open or close states, fine segmentation of iris and pupil is needed for fatigue level detection.

Two classes of iris segmentation methods exist in literature, traditional image segmentation method and deep-learning method. Most of the traditional methods are based on Hough transform, image normalization, image gradient, active contour and threshold segmentation. For instance, references [5] and [6] use integro-differential operator (IDO) and circle Hough transformation (CHT) and these are commonly used methods of iris segmentation. Relative total variation with $L^1$ norm (RTV-$L^1$) can robustly suppress noisy texture pixels to obtain clear iris images, and then an improved circular Hough transform is used to detect iris and pupil circles on noise-free iris images [7]. The method in [7] uses a series of post-processing operations to locate iris boundaries accurately. Reference [8] evaluates the dataset with five different algorithms: Contrast-adjusted Hough transform (CAHT) [9], generalized structure tensor (GST) [10], iterative Fourier-based pulling and pushing (IFPP) [11], open source iris (OSIRIS) [12] and weighted adaptive Hough and ellipsopolar transforms (WAHET) [13]. All of these methods can obtain satisfactory results in high definition infrared images. However, when facing to the colour images, low-quality, or low-resolution iris images such as these images captured in driving condition, the segmentation accuracy is not high enough for features extraction.

More recently, with the development of deep-learning method, significant progress has been made in intelligent transport systems and image classification [14–16]. Reference [17] proposes multi-scale fully convolutional networks (MFCNs) and hierarchical convolutional neural networks (HCNNs) which combines deep-learning method with iris segmentation. Reference [18] applies attention mechanism to atrous spatial pyramid pooling (ASPP) and pyramid pooling module, which adjusts the weights of different channels in feature maps and also re-estimates the spatial distribution of feature map. These methods improve the performance of iris segmentation and localization. In [19], fully residual encoder–decoder net (FRED-Net) with residual connectivity in both encoder and decoder is proposed to perform semantic segmentation of iris. Compared with the traditional method, deep-learning method always gets better result in colour images or low-quality iris images. Additionally, most iris segmentation methods mainly aim to obtain texture information of iris region for iris recognition. The images collected in this application are generally with high resolution and captured by a specific iris image acquisition instrument at close range. However, due to the change of illumination in driving environment and the camera installed usually has a large distance to driver's face, so the images obtained under this condition are often with low resolution and not clear, which makes accurate iris segmentation difficult. Furthermore, the segmentation used for iris recognition mainly focuses on the texture of the iris region and rarely distinguishes the boundaries between the iris and the pupil, which is useful for accurate EOCR calculations.

Considering the advantages of deep learning methods in different image segmentation and the difficulties in driver’s eye image feature extraction, the purpose of this paper is toward the application of driver fatigue level monitoring to study the segmentation of iris and pupil area using deep learning methods. The proposed method estimates the EOCR value according to the extracted features to provide reliable information for driver fatigue level detection. Based on the above motivation, we propose an integrated framework to segment iris and pupil images, and simultaneously measure the value of eyes opening and closing ratio. Furthermore, we design a novel BiSeNet-based network that can divide eyes image into iris part and pupil part. We also provide experimental dataset, validation and demonstration for the EOCR detection. We use the infrared camera to capture images to establish test dataset that contains both segmentation images and the ground truth of EOCR parameters. The images in our test dataset are all captured in driving environments. The proposed method can segment iris images and get eyes opening and closing ratio simultaneously and it would facilitate the monitoring and early warning of driver’s fatigue level.

### 2 | METHODOLOGY

Taking the advantages of both the BiSeNet in semantic segmentation and VGG convolutional network in accurate image feature extraction with deep layers [20], the EOCR detection framework is proposed as shown in Figure 1. The original image is first segmented by iris and pupil segmentation network to obtain the feature score map and segmented image, and then the feature score map and the original image are concatenated as the input of VGG model for accurate feature extraction. Finally, the detected value of the EOCR is obtained according to the extracted features. The iris and pupil segmentation network and the VGG based EOCR detection model are respectively described in Sections 2.1 and 2.3. The EOCR value calculation.
method and the establishment of the test dataset are illustrated in Section 2.2.

2.1 Iris and pupil segmentation network

Iris and pupil segmentation in this paper is a kind of image semantic segmentation. For deep learning methods, the semantic segmentation models commonly design and consider spatial information in shallow layer and context information in deep layer. A BiSeNet approach uses the strategy of detecting spatial information by spatial path and detecting context information by context path respectively, and then merging them with feature fusion module. A U-Net network method concatenates shallow features with deep features in depth direction in the decoding stage to fuse the context information and location information. These methods have achieved good results in image semantic segmentation such as street scene. Inspired by the above-mentioned BiSeNet and U-Net strategies, this paper proposes the iris and pupil image segmentation network as shown in Figure 2. The original images are respectively encoded to extract the spatial features and context features through spatial path and context path. The U-Net network is added into the spatial path for feature fusion so that the segmented image and context feature map are obtained in the decoding stage.

In Context Encoder Path, the down sampling layers (4× down to 32× down) are used to detect deep context features of images. Furthermore, inspired by the work in [21], we propose the global convolutional network (GCN)-based module to mine the context information for more boundary details of iris and pupil. Figure 3(a) illustrates the filters and kernel size of the GCN-based module. In this module, large kernel is used to gain larger receptive field, and considering that the resolution of the feature map after 16X and 32X down sampling is not too large, when the input is 128×96, the resolution of each feature map is 8×6 and 4×3, which is helpful to global convolution. To fuse with spatial information, the 4× up sampling and 2× up sampling are used to improve the resolution of the context information. In Spatial Encoder Path, we add the U-Net encoder part into the down sampling components (3×3 Conv, ReLU and Max pool) to get location information in images. In our model, different from the original U-Net, we only use 8× down sampling (three times of 3×3 Conv, ReLU and Max pool) in the encoder part, which could keep enough spatial information. The extracted spatial information and context information are fused in feature fusion module (FFM), which is the input of decoder part. Figure 3(b) shows the details of FFM module. In the decoding stage, the convolution (3×3 Conv, ReLU) and deconvolution (3×3 Deconv) operations are used to retain enough detail information. In addition, the feature map after deconvolution is concatenated with the spatial features obtained in the encoding stage, and then inputs into the residual module (Figure 3(c)). The design guarantees the algorithm to converge and the output of the network is the segmented image and feature score map.
2.2 | EOCR calculation and dataset

The images captured in driver fatigue monitoring environment usually include face and background. It is necessary to locate and extract the eye area for further iris and pupil segmentation. An entropy of regression trees algorithm is used to detect facial landmarks in infrared face images. The landmarks include the inner and outer eye corners, the upper and lower eyelid etc. The red points around eyes are the detected feature points as shown in Figure 4, from which the eye area is located.

The extracted eye area image is used as a sample of test dataset, different from other datasets containing only eye images for iris segmentation. The dataset in this paper includes both images and corresponding EOCR values, so the ground truth of iris and pupil region and the corresponding EOCR value of the sample image are also marked when the dataset is established. As shown in Figure 4, the red and green parts represent the pupil and the iris, and the value of EOCR is 0.961 in this sample image. In order to prevent the images from being deformed during the shrinking process and possibly affect the experiment results, we crop the region of the eye image by the following equation.

\[
\begin{align*}
\text{width} &= 1.4l \\
\text{height} &= 0.7l 
\end{align*}
\]

(1)

where \( l \) represents the distance between the inner and outer eye corners. To set the width and height of the sample images in test dataset, rather than ratio 4:3, we find that 2:1 is a better ratio of eye region because main classification error comes from the interference like eyebrow or some place which is away from the iris. At the same time, the ratio of 2:1 is closer to the real shape of human eyes. We resize the sample images into \( 128 \times 64 \) as the input of our network. We calculate the opening and closing ratio of eyes as

\[
\text{ratio} = \frac{O_i}{\bar{O}}
\]

(2)

where \( O_i \) is the detected pixel number of the iris and pupil area, \( \bar{O} \) is the ground truth of the iris and pupil area, the ground truth value is measured when a person is in good mental condition.
and eyes fully open. We manually labelled the ground truth of iris and pupil area with the Labelme software and calculated the corresponding value of eye opening and closing ratio according to Equation (2).

### 2.3 VGG-based EOOCR detection

VGG network model can accurately obtain detailed information by using small kernels in different convolution layers (3 × 3 Conv1a~Conv4b), maximum pooling layers (2 × 2 Pool1~Pool4), and full connection layers (FC) with increasing network depth (16 layers). We use these characteristics of VGG network in EOOCR detection, and the related parameters are shown in Table 1. We add the batch normalization and the ReLU unit to each convolution layer, and replace the first fully connected layer with global average pooling (GAP) layer. This arrangement reduces many parameters and improves the detection performance. Finally, the EOOCR value is obtained after the FC1 and FC2 layers.

### 3 EXPERIMENT

In this section, we present the evaluation of our proposed iris and pupil segmentation network with three public datasets, the CASIA-Iris-Thousand, the CASIA-Iris-Interval [22], and the UBIRIS.v2 [23]. We also conduct the experiment with our test dataset to evaluate the performance of the proposed method in iris and pupil segmentation and EOOCR detection.

#### 3.1 Dataset

The CASIA-Iris-Thousand dataset is a subset of the CASIA-IrisV4 dataset, which contains 20,000 iris images from 1000 subjects with resolution of 640 × 480. The dataset is suitable for iris features extracting and iris segmentation algorithm testing. With this dataset, we mainly test the segmentation effect of our proposed method under low quality image set. We take the testing approach [24] and the testing contains three steps. Firstly, we reduce the image contrast by the following relationship:

\[
y = \text{norm} \left( \text{tanh} \left( 3 \times \left( \frac{x}{255} - 0.5 \right) + u \left( -0.3, 0.3 \right) \right) \times 255 \right)
\]

where \(x\) is the input intensity in the range [0, 255], \(y\) is the output intensity in the same range, \(u\) \((a, b)\) is the uniform distribution between \(a\) and \(b\), and the norm function normalize the output between 0 and 1, tanh is a hyperbolic tangent function. Second, we add some shadows to the image in order to make it closer to the low quality image as:

\[
y = \text{norm} \left( \text{tanh} \left( \text{randSign} \times (x - 0.5 + u (-0.1, 0.3)) \right) \right) + u (0, 0.1)
\]

where randSign generates a random coefficient in the set \((-1, 1)\). Finally, we pass the shadowed image through a motion blur filter applying the linear camera motion by \(u(5, 10)\) pixels in the direction \(u(-\pi, \pi)\). Figure 5 shows some processed image examples.

The CASIA-Iris-Interval dataset is another subset of the CASIA-IrisV4, which contains 2636 images in 249 groups. Due to the unique infrared LED array, the images in this dataset contain rich texture information, which is suitable for detailed texture research and iris recognition. Some samples are shown in Figure 6.

The UBIRIS.v2 dataset contains 2250 colour images from 20 subjects with resolution of 400 × 300. Images in this dataset were captured on moving, at-a-distance, and involve noises, such as illumination variance, motion/defocus blur and occlusion of glasses and eyelids. Figure 7 shows some samples.

The images in our test dataset are shown in Figure 8. The images were captured in driving environment with infrared camera and the active infrared auxiliary lighting source. The environment illumination is variable with lighting from day to night, including eyes state from open to closure. From the original images, we cropped and resized the images with the resolution of 128 × 64, and manually labelled the ground truth of iris and pupil, the dataset labelled by one operator and checked by another person. Our test dataset contains 1015 images from six subjects. There are 907 images in training set and 108 images in test set.

#### 3.2 Training

We trained our proposed iris and pupil segmentation network with CASIA-IrisV4 and UBIRIS.v2 Datasets respectively, and resized the input images to the resolution of 128 × 96. We also used the mean subtraction, random horizontal flip and random rotation on the input images to augment the dataset in training process. The loss function in iris and pupil segmentation is

| Layer name | Shape-in | Shape-out | Kernel |
|------------|----------|-----------|--------|
| Conv1a     | 64 × 128 × 3 | 64 × 128 × 64 | 3 × 3 × 64.1 |
| Conv1b     | 64 × 128 × 64 | 64 × 128 × 64 | 3 × 3 × 64.1 |
| Pool1      | 64 × 128 × 64 | 32 × 64 × 64 | 2 × 2 × 1.2 |
| Conv2a     | 32 × 64 × 64 | 32 × 64 × 128 | 3 × 3 × 128.1 |
| Conv2b     | 32 × 64 × 128 | 32 × 64 × 128 | 3 × 3 × 128.1 |
| Pool2      | 32 × 64 × 128 | 16 × 32 × 128 | 2 × 2 × 1.2 |
| Conv3a     | 16 × 32 × 128 | 16 × 32 × 256 | 3 × 3 × 256.1 |
| Conv3b     | 16 × 32 × 256 | 16 × 32 × 256 | 3 × 3 × 256.1 |
| Pool3      | 16 × 32 × 256 | 8 × 16 × 256 | 2 × 2 × 1.2 |
| Conv4a     | 8 × 16 × 256 | 8 × 16 × 512 | 3 × 3 × 512.1 |
| Conv4b     | 8 × 16 × 512 | 8 × 16 × 512 | 3 × 3 × 512.1 |
| Pool4      | 8 × 16 × 512 | 4 × 8 × 512 | 2 × 2 × 1.2 |
| GAP        | 4 × 8 × 512 | 1 × 1 × 512 | — |
| FC1        | 1 × 1 × 512 | 1 × 1 × 256 | — |
| FC2        | 1 × 1 × 256 | 1 × 1 × 1 | — |
the mean binary cross-entropy between the output and ground truth given by:

\[
\text{loss}_1 = \frac{1}{M \times N \times B} \sum_{k=1}^{B} \sum_{j=1}^{N} \sum_{i=1}^{M} \log (p_{ij}) - (1 - t_{ij}) \log (1 - p_{ij})
\]  

(5)

where \( t_{ij} \) is the label of pixel \((i, j)\) in the ground truth, \( p_{ij} \) is the label of pixel \((i, j)\) in the output image for the image of the size \( M \times N \) and \( B \) is the batch size which equals to 8 in this work. In EOCR detection, the loss function is given by:

\[
\text{loss} = \text{loss}_1 + a |x - x^*|
\]  

(6)

FIGURE 5  Processed sample images in CASIA-Iris-Thousand

FIGURE 6  Sample images with corresponding ground truth in CASIA-Iris-Interval

FIGURE 7  Sample images with corresponding ground truth in UBIRIS.v2

FIGURE 8  Sample images with corresponding EOCR and image ground truth in our test dataset
where $x$ is the output value of the EOCR, $x^*$ is ground truth of the EOCR, $\alpha$ in this work set as 1. We use these loss functions in our test dataset to evaluate the segmented image and the detected eye opening and closing ratio.

### 3.3 Test and result

#### 3.3.1 Test in CASIA-Iris-Thousand

In CASIA-Iris-Thousand dataset, we mainly want to test the effect of our method in low-quality image set. So we trained the network by the images after augmentation and adding random noises, part of the segmented results as shown in Figure 9. The black region in the figure means the correct segmentation, the green one for the false positive error, and the red one for the false negative error. We can see that most of the iris and pupil areas are detected correctly, and only a few detection errors (red and green dots) appear on the iris and pupil edges. These results imply that the proposed method can get correct segmentation results even in noisy and low-quality images.

**FIGURE 9** Part segmentation results of the degraded images in CSAIA-Iris-Thousand dataset

#### 3.3.2 Test in CASIA-Iris-Interval and UBIRIS.v2

In this section, we compare our method with other methods in iris segmentation with CASIA-Iris-Interval and UBIRIS.v2 datasets. The images are respectively inferred images and colour images and some test examples are shown in Figure 10. We use the following calculation to measure the average segmentation error rate that is used in noisy iris challenge evaluation (NICE) I competition [19].

$$\text{error} = \frac{1}{N \times m \times n} \sum_{i,j \in (m,n)} G(i,j) \oplus M(i,j) \quad (7)$$

The performance is shown in Table 2. In Equation (7), $N$ is the total number of test images, $m$ is the rows and $n$ is the columns of the sample image. $G$ and $M$ are respectively the ground truth and the output. $i$ and $j$ are the column and row coordinates of pixels in $G$ and $M$. The operator $\oplus$ represents the XOR operation to evaluate the inconsistent pixels between $G$ and $M$. 
In order to evaluate the method more fairly, three parameters: Recall ($R$), precision ($P$) and $F$-measure protocol ($RPF$-Measure) [8] are used. The $RPF$-Measure can evaluate the segmentation performance on the basis of corresponding ground truths with a smart measure of strength and weakness of an algorithm [19]. This measurement has the prerequisite calculations of true positive, false negative, and false positive pixels from the ground truth image and output image, and it is calculated as,

\[
\begin{align*}
R &= \frac{tp}{tp + fn} \\
\mu &= \frac{\sum \text{tp}}{\sum \text{output}} \\
P &= \frac{tp}{tp + fp} \\
\sigma &= \sqrt{\frac{\sum (tp - \mu)^2}{\text{output}}} \\
F &= \frac{2RP}{R + P}
\end{align*}
\]  

(8)

where $tp$ is the number of true positive pixels, $fn$ is the number of false negative pixels, and $fp$ is the numbers of false positive pixels. The segmentation performance is compared by $RPF$-Measure as shown in Table 3, $\mu$ and $\sigma$ are respectively the mean and variance of the corresponding indicators. It can be seen
TABLE 3  Performance comparison with different methods

| Dataset          | Method          | \( R \)  | \( P \)  | \( F \)  |
|------------------|-----------------|---------|---------|---------|
|                  |                 | \( \mu \) | \( \sigma \) | \( \mu \) | \( \sigma \) | \( \mu \) | \( \sigma \) |
| CASIA-Iris-Interval | GST [10]        | 85.19   | 18      | 89.91   | 7.37     | 86.16   | 11.53     |
|                  | OSIRIS [12]     | 97.32   | 7.93    | 93.03   | 4.95     | 89.85   | 5.47      |
|                  | WAHET [13]      | 94.72   | 9.01    | 85.44   | 9.67     | 89.13   | 8.39      |
|                  | IFPP [11]       | 91.74   | 14.74   | 83.5    | 14.26    | 86.86   | 13.27     |
|                  | CAHT [9]        | 97.68   | 4.56    | 82.89   | 9.95     | 89.27   | 6.67      |
|                  | IDO [5]         | 71.34   | 22.86   | 61.62   | 18.71    | 65.61   | 19.96     |
|                  | IrisSeg [27]    | 94.26   | 4.18    | 92.15   | 3.34     | 93.1    | 2.65      |
|                  | SegNet-Basic [28]| 99.60  | 0.66    | 91.86   | 2.65     | 95.55   | 1.40      |
|                  | IrisDenseNet [29]| 97.10  | 2.12    | 98.10   | 1.07     | 97.58   | 0.99      |
|                  | Proposed method | 97.50   | 1.64    | 96.41   | 1.37     | 97.00   | 2.00      |
| UBIRIS v2.0      | CAHT [9]        | 18.02   | 26.04   | 11.39   | 19.89    | 12.55   | 20.06     |
|                  | GST [10]        | 42.59   | 40.53   | 39.91   | 34.39    | 39.93   | 36.55     |
|                  | IFPP [11]       | 44.38   | 37.42   | 23.71   | 26.94    | 28.52   | 29.68     |
|                  | OSIRIS [12]     | 26.46   | 31.52   | 17.29   | 28.81    | 18.65   | 28.47     |
|                  | WAHET [13]      | 27.40   | 34.48   | 22.76   | 31.03    | 23.68   | 31.90     |
|                  | MFCNs [17]      | N/A     | N/A     | N/A     | N/A      | 90.78   | 4.70      |
|                  | IrisParseNet [18]| N/A    | N/A     | N/A     | N/A      | 91.82   | 4.26      |
|                  | FCDNN [24]      | 93.98   | N/A     | 94.88   | N/A      | 93.90   | N/A       |
|                  | Proposed method | 96.49   | 3.18    | 92.38   | 4.95     | 94.39   | 4.36      |

FIGURE 11  Examples of segmentation results in our test dataset

that the average value \( \mu \) in our proposed method generally is larger than the one in others. Meanwhile, the deviation \( \sigma \) in our proposed method generally is smaller than the one in others methods. This implies that the segmentation results are more reliable with the proposed method both in infrared images and colour images.

3.3.3  Test in our dataset

With our test dataset and the proposed method, we also conduct the test of both segmenting iris and pupil images and detecting the value of EOCR. Figure 11 shows some results. Since the value of EOCR mainly reflects eyes opening degree, while PERCLOS denotes the percentage of eyelid closure over the pupil [1], their relationship can be described by following equations.

\[
EOCR = 1 - PERCLOS 
\]

\[
PERCLOS = \frac{\sum_{i=1}^{n} \text{Close}[i]}{n} \times 100\%
\]

where \( n \) is the total number of frames analysed, \( \sum_{i=1}^{n} \text{Close}[i] \) is the number of frames where the eye is closed.
In light of different eyes open and closure perception, the parameter of PERCLOS defined as P70 or P80, respectively describing eye closure over 70% or 80%. In our proposed method, we not only segment iris and pupil information, but also directly detect the EOCR value. The EOCR value is between 0.0 and 1.0, which represents detailed information of eye’s open and closure. By the relationship of EOCR and PERCLOS, if the EOCR is detected, the PERCLOS can be obtained by Equation (9). Since PERCLOS is a popular parameter in fatigue detection, by using the EOCR value, the fatigue degree also can be identified. In order to evaluate the results of detected EOCR value, we calculate the standard deviation $sd$ of EOCR in test dataset and test videos as follows.

$$sd = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - x^*_i)^2}.$$  (11)

where $x_i$ is the detected EOCR value and $x^*_i$ is calculated by Equation (2). The value of sd represents the deviation between the calculated value and the detected value. We captured 25 testing videos of drivers in driving environment to detect the value of EOCR. Figure 12 illustrates comparison results between the detected EOCR and calculated EOCR values. The consistency
between the two values is small and these results show that the proposed method has good results in detecting EOOCR. Furthermore, we also use the test dataset to train the VGG model to conduct the same experiment with the testing videos. The performance comparison is listed in Table 4. It can be seen that the parameter sd in our proposed method is smaller than that in the VGG model, which implies that the proposed method is more robust in detecting eye status than the VGG model.

4 | DISCUSSION

This paper aims to study the detection method of eye opening and closing ratio for driver fatigue monitoring, and the three group of experiments were conducted to evaluate the proposed method. The first experiment used the CASIA-Iris-Thousand dataset to test the detection effect of the proposed iris and pupil segmentation network on the degraded image. From the results in Figure 9, we see that the iris and pupil region is well extracted in the random noisy and shadowed images, while only a small number of segmentation errors appear on the edges. This indicates that the proposed method has good processing performance for low-quality images.

The second experiment carried out iris segmentation test using the proposed method in CASIA-Iris-Interval and UBIRIS.v2 datasets. We comprehensively compared the performance of the proposed method and other algorithms using NICE I and RPF-Measure evaluation indicators. From Figure 10, it can be seen that there are a small number of segmentation errors distributed in iris and pupil edges in colour images with low contrast, while for most colour images and infrared images, good segmentation results were obtained. From the NICE I average segmentation error rate in Table 2, although the proposed method is not the best (IrisParseNet obtained the minimum error rate of 0.42%), the average error rate is also relatively small (0.75%). Compared with the RPF-Measure indicators in Table 3, although the proposed method is not the best in single statistical indicators (such as μ or σ), and the best results have been obtained by IrisDenseNet), the proposed method respectively got that the μ is 97.0 and 94.39, and the σ is 2.0 and 4.36, which means the proposed method is also stable and shows best accuracy in segmentation by comprehensively analysing the results of the F parameters in the two datasets.

In the third experiment, the iris and pupil segmentation and EOOCR detection were carried out simultaneously with the proposed method using test dataset and test videos. It can be seen from Figure 11 that the proposed method correctly segmented the iris and pupil regions and simultaneously obtained the EOOCR values. By analysing the data in Table 4 and Figure 12, the detected EOOCR values by the proposed method are consistent with the calculated values. The standard deviation is about 0.0544 and these results further illustrated that the proposed method accurately detected the EOOCR values while segmenting the images.

Combined with the above analysis and research motivation, it is of importance for the proposed method to obtain the image segmentation and the EOOCR detection simultaneously. The comprehensive performance of the detection results is satisfactory, which provides a basis for driver’s fatigue degree monitoring. However, the method proposed in this paper synthesized the characteristics of the complex network structure with high-cost computation, which is the main limitation of this method.

5 | CONCLUSION

In this paper, we proposed a novel method to detect the eye opening and closing ratio for identifying driver’s fatigue level. A network based on BiSeNet model was first proposed by combining a GCN-based module and U-Net to get accurate iris and pupil features. Moreover, we designed the network based on VGG model to measure the EOOCR. To test the detecting effect of the proposed method, we also established our test dataset in which all the images were captured in driving environment. Experiments have been conducted and compared with different methods in CASIA-Iris-Thousand, CASIA-Iris-Interval, UBIRIS.v2, and our test datasets. The results showed that the proposed method has good detection effect on both infrared images and colour images. The experimental results further showed that the proposed method was reliable in detecting the eyes opening degree and therefore, the detected EOOCR value can be used to identify driver’s fatigue level. In addition, fatigue monitoring is a real-time process, which requires high processing speed, but the method proposed in this paper needs high computational cost. Therefore, optimizing the network structure to improve efficiency will be the focus of the future research direction.

ORCID

Qi ji Zhao https://orcid.org/0000-0002-5695-0043

REFERENCES

1. Dinges, D.F., Grace, R.: PERCLOS: A Valid Psychophysiological Measure of Alertness as Assessed by Psychomotor Vigilance. Federal Highway Administration, Washington DC (1998). Publication Number FHWA-MCRT-98-006
2. Zhao, Q.-J., et al.: Eye moving behaviors identification for gaze tracking interaction. J. Multimodal User Interfaces 9(2), 89–104 (2015)
3. Bai, Z.-H., et al.: Driving fatigue detection based on active shape model and fuzzy inference. Chin. J. Sci. Instrum. 36(4), 768–775 (2015)
4. Jo, J., et al.: Detecting driver drowsiness using feature-level fusion and user-specific classification. Expert Syst. Appl. 41(4), 1139–1152 (2014)
5. Daugman, J.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. 14(1), 21–30 (2004)
6. Wildes, R.P.: Iris recognition: an emerging biometric technology. Proc. IEEE 85(9), 1348–1363 (1997)
7. Zhao, Z., Kumar, A.: An accurate iris segmentation framework under relaxed imaging constraints using total variation model. In: Proc. Int. Conf. IEEE International Conference on Computer Vision. Santiago, Chile, pp. 3828–3836 (2015)
8. Hofbauer, H., et al.: A ground truth for iris segmentation. In: Proc. Int. Conf. the 22nd International Conference On Pattern Recognition. Stockholm, Sweden, pp. 527–532 (2014)

9. Rathgeb, C., et al.: Iris Biometrics: From Segmentation to Template Security. Springer-Verlag, New York (2013)

10. Alonso-Fernandez, F., Bigun, J.: Iris boundaries segmentation using the generalized structure tensor. A study on the effects of image degradation. In: Proc. Int. Conf. the 5th International Conference on Biometrics: Theory, Applications and Systems. Arlington, VA, USA, pp. 426–431 (2012)

11. Uhl, A., Wild, P.: Multi-stage visible wavelength and near infrared iris segmentation framework. In: Proc. International Conference Image Analysis and Recognition. Aveiro, Portugal, pp. 1–10 (2012)

12. Othman, N., et al.: OSIRIS: An open source iris recognition software. Pattern Recognit. Lett. 82(2), 124–131 (2016)

13. Uhl, A., Wild, P.: Weighted adaptive hough and ellipsopolar transforms for real-time iris segmentation. In: 5th IEEE International Conference On Biometrics. New Delhi, India, pp. 283–290 (2012)

14. Yang, Y., et al.: Chinese vehicle license plate recognition using kernel-based extreme learning machine with deep convolutional features. IEEE Trans. Intell. Transp. Syst. 12(3), 213–219 (2018)

15. Wang, Y., et al.: Deep learning-based vehicle detection with synthetic image data. IEEE Trans. Intell. Transp. Syst. 13(7), 1097–1105 (2019)

16. Yépez, J., et al.: Deep learning-based embedded license plate localisation system. IEEE Trans. Intell. Transp. Syst. 13(10), 1569–1578 (2019)

17. Liu, N.-F., et al.: Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In: IEEE International Conference on Biometrics. Halmstad, Sweden, pp. 1–8 (2016)

18. Wang, C.-Y., et al.: Joint iris segmentation and localization using deep multitask learning framework. Comput. Sci. (2019) https://arxiv.org/abs/1901.11195

19. Muhammad, A., et al.: FRED-Net: Fully residual encoder–decoder network for accurate iris segmentation. Expert Syst. Appl. 122, 217–241 (2019)

20. Yu, C.-Q., et al.: BiSeNet: bilateral segmentation network for real-time semantic segmentation. In: The European Conference On Computer Vision. Munich, Germany, pp. 325–341 (2018)

21. Peng C., et al.: Large kernel matters—improve semantic segmentation by global convolutional network. In: IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, pp. 1743–1751 (2017)

22. CASIA Iris Image Database [Online]. http://biometrics.idealtest.org/ (2020). Accessed 28 April 2020

23. Proença, H., et al.: The UBIRIS.v2: A database of visible wavelength iris images captured on-the-move and at-a-distance. IEEE Trans. Pattern Anal. Mach. Intell. 32(8), 1529–1535 (2010)

24. Shabab, B., et al.: An end to end deep neural network for iris segmentation in unconstrained scenarios. IEEE Trans. Neural Networks 106, 79–95 (2018)

25. Tan, T.-T., et al.: Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. IEEE Trans. Image and vision computing 28(2), 223–230 (2010)

26. Hindl, M., Krupička, M.: Unsupervised detection of non-iris occlusions. Pattern Recognit. Lett. 57, 60–65 (2015)

27. Gangwar, A., et al.: IrisSeg: A fast and robust iris segmentation framework for non-ideal iris images. In: International Conference on Biometrics. Halmstad, Sweden, pp. 1–8 (2016)

28. Badrinarayanan, V., et al.: SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39(12), 2481–2495 (2017)

29. Muhammad, A., et al.: IrisDenseNet: Robust iris segmentation using densely connected fully convolutional networks in the images by visible light and near-infrared light camera sensors', IEEE Trans. Sensors 18(5), 1–30 (2018)

---

How to cite this article: Zhao Q, Jiang J, Lei Z, Yi J. Detection method of eyes opening and closing ratio for driver’s fatigue monitoring. IET Intell Transp Syst. 2021;15:31–42. https://doi.org/10.1049/itr2.12002