



Research article

Fatigue detection method for UAV remote pilot based on multi feature fusion

Lei Pan, Chongyao Yan, Yuan Zheng, Qiang Fu*, Yangjie Zhang, Zhiwei Lu, Zhiqing Zhao and Jun Tian

School of Computer Science, Civil Aviation Flight University of China, Guanghan 618307, China

* **Correspondence:** Email: cafczdm@163.com.

Abstract: In recent years, UAV industry is developing rapidly and vigorously. However, so far, there is no relevant research on the fatigue detection method for UAV remote pilot, which is the core technology to ensure the flight safety of UAV. Aiming at this problem, a fatigue detection method for UAV remote pilot is proposed in this paper. Specifically, we first build a UAV operator fatigue detection database (OFDD). By analyzing the fatigue features in the database, we find that multiple facial features are highly correlated to the fatigue state, especially the head posture, and the temporal information is essential for distinguish between yawn and speaking in the study of UAV remote pilot fatigue detection. Based on these findings, a fatigue detection method for UAV remote pilots was proposed by efficiently locating the related facial regions, a multiple features extraction module to extract the eye, mouth and head posture features, and an efficient temporal fatigue decision module based on SVM. The experimental results show that this method not only performs well on the traditional driver dataset, but also achieves an accuracy rate of 97.05%; and it achieves the highest detection accuracy rate of 97.32% on the UAV remote pilots fatigue detection dataset OFDD.

Keywords: fatigue detection; UAV remote pilot; database; feature fusion

1. Introduction

Civil unmanned aerial vehicles (UAVs) are growing rapidly in the last few years, thanks to their agility, flexibility and superior mobility. These benefits enable civil UAVs to be applied in a wide range of fields, such as film and television, logistics and distribution, agriculture and plant protection, land

surveying and mapping, energy and power line inspection.

Up to now, the Federal Aviation Administration (FAA) of the United States has issued 280,418 remote pilot licenses for UAVs [1] and 120,844 remote pilot licenses have been issued by the Civil Aviation Administration of China (CAAC) [2].

Fatigue has been proved to have a negative impact on people and reduce their work efficiency. N. Aung and P. Tewogbola [3] discussed the impact of emotional labor on the health of employees in the workplace, pointing out that such negative factors as burnout and fatigue will affect employees' emotions, not only reducing their work efficiency but also seriously affecting their physical and mental health. M. Hoang et al. [4] investigated the influence of call frequency, worry, fatigue and other negative factors on work efficiency. Studies [5–7] have consistently demonstrate a vigilance decrement of remote pilots about 20–35 minutes after task initiation and the dramatic increasement of fatigue after 12 hours working [8]. However, so far, there is no relevant work have applied in fatigue detection field for UAV remote pilot.

At present, the researches on fatigue detection are mainly concentrated in the field of traditional drivers. It is widely recognized that the fatigue can be detected by four different features, including physiological features [9], psychological features [10], speech features [11] and visual features [12]. Fatigue detection methods based on physiological features collect physiological signals and extract t electromyography signal and electrocardiograph signal through sensors [13]. Another natural idea is using psychological features. X. Li et al. [14] proposed a fatigue driving detection method based on psychoacoustic perceptual masking processing steps so as to the high fatigue sensitive frequency components in the speaking process can be highlighted. However, the above two kinds of methods are inapplicable to UAV remote pilot due to sensor intrusion [15,16]. The fatigue detection methods based on speech feature are only applicable to some specific scenarios with standard callouts and responses, and the conversation databases are limited or insufficient in quantity.

The methods based on visual features are the last kind of fatigue detection method. Some facial movement features always reflect the fatigue state, such as the number of blinks increases, the blinking speed becomes slower, the number of nods increases, and frequent yawning. Researchers have been working on techniques to detect the fatigue of driver based on these visual signs and those visual features based methods can be divided into two parts: based on expert rules or based on deep learning.

The first part of visual features based methods is fully rely on the rules design by experts. Those hand designed features and detection standards result in a slow detection speed and a low detection accuracy [17] for detecting driver fatigue. The second part of visual features based methods applies machine learning to detect driver fatigue. With the successful application of neural network architecture such as convolutional neural network in the field of computer vision, such as image classification, semantic segmentation and object detection [18]. In the field of face recognition, F. Liu et al. [19] comprehensively reviewed the existing deep learning based single sample face recognition methods, classifying them into virtual sample methods and generic learning methods, and discussed the advantages and the common defects of these methods. Y. Ed-doughmi et al. [20] proposed a fatigue detection method for drivers based on RNN; Z. Chang et al. [21] proposed a fatigue detection method based on Haar feature and extreme learning machine; R. Huang et al. [22] applied the functional calibration depth convolutional model RF-DCM to solve the problem of speaking and yawning distinguishing problems based on mouth feature fatigue detection; W. Gu et al. [23] employed multi-scale pooling for eye and mouth state detection, which was then applied in hierarchical CNN based model for driver fatigue detection. S. Dey et al. [24] used the positions of eyes, nose and mouth to

extract the deep information, such as the eye aspect ratio (EAR), the mouth opening ratio and the nose length ratio, by a histogram of oriented gradient (HOG) with support vector machine (SVM) pretrained model. Z. Xiao et al. [25] proposed a method for driver fatigue detection by deep cascaded multi-task long short-term memory (LSTM) model. W. Liu et al. [26] proposed a driver fatigue detection model based on multiple facial features and two stream convolution neural network, which can combine static and dynamic information. L. Geng et al. [27] proposed a real time fatigue detection method based on morphology infrared features and deep learning, which solves the problem of detecting driver fatigue in an underlit scene. F. Liu et al. [28] introduced the integration of RGB-D camera and deep learning where Generative Adversarial Networks and multi-channel scheme are used to improve the performance, proving that the features extracted by depth learning method are effective for fatigue detection.

To sum up, the current fatigue detection methods are aimed at the traditional driver, while the fatigue detection for UAV remote pilot, to our best knowledge, has not yet been carried out and they are inapplicable to UAV remote pilot fatigue detection due to the following three reasons: 1) the visibility of driver's working space is invariable and is not affected seriously by the outside environment. However, UAV remote pilots often work outdoors, greatly affecting by the visibility. 2) remote pilots always need to communicate with the air traffic controller frequently, which means a higher speaking time as well as the speaking frequency, resulting in a confusion between "yawn" and "speaking" that is critical and must be determined in fatigue detection practice. 3) the robustness of multi-facial features. Depending on the change of head posture, common fatigue state features are easily lost.

In view of the above problems, this paper proposed a fatigue detection method for UAV remote pilot. In brief, the main contributions of this paper are summarized as follows.

1) We establish a UAV remote pilots fatigue detection dataset OFDD that consists of 48 videos with diverse condition, along with the thorough analysis and finding on our dataset.

2) We propose a fatigue detection method for UAV remote pilots, which integrates multiple facial fatigue features, especially the head posture, and temporal information to solve above challenge.

3) The experimental results show that this method not only performs well on the traditional driver fatigue detection dataset with accuracy rate of 97.05%, but also achieves the highest detection accuracy rate of 97.32% on our OFDD dataset.

The rest of this paper is organized as follows. In Section 2, we first briefly review the related dataset for fatigue detection. Then, the UAV remote pilot fatigue detection dataset OFDD is established and analyzed. In Section 3, based on our findings from our OFDD, we propose a fatigue detection method for UAV remote pilot. In Section 4, experimental results are given to verify the performance of our proposed method. In Section 5, several conclusions are drawn based on the proposed method.

2. Database

At present, there are several public fatigue detection datasets available for traditional driver, such as YawDD yawning dataset released by the University of Ottawa [29], DriveFace driver driving scene facial feature dataset released by the University of California Irvine [30], CEW blinking dataset released by Nanjing University of Aeronautics and Astronautics [31] and NTHU-DDD driver sleepiness detection dataset of National Tsinghua University [32]. To the best of our knowledge, YawDD, DriveFace, CEW and NTHU-DDD are the most popular datasets for traditional driver fatigue detection, and they have been widely used in most recent driver fatigue detection work.

The YawDD yawning dataset is one of the earliest driver fatigue datasets. YawDD contains 322 videos captured by on board camera and consists of both male and female drivers, with and without glasses/sunglasses, from different ethnicities, and in 3 different situations: 1) normal driving (no talking), 2) talking or singing while driving, and 3) yawning while driving under natural and changing lighting conditions. For each video, the length ranges from 10 to 120 s. It is mainly used to develop and test fatigue detection algorithms and models, and can also be used for face and mouth recognition and tracking.

DriveFace dataset is a public image dataset that contains a sequence of object images when driving in a real scene. It consists of 606 samples, and the resolution of each sample is 640×480 . It was obtained from 4 drivers (2 women and 2 men) on different days. These drivers have various face accessories, such as wearing glasses and beard.

CEW blink dataset is provided by Nanjing University of Aeronautics and Astronautics. It contains photos of the 2423 testers with the open and closed eyes, including individual differences and various environments, such as lighting, blur, and glasses.

NTHU-DDD is another driver sleepiness detection dataset established by National Tsinghua University. The whole dataset contains 36 subjects of different ethnicities recorded with and without wearing glasses under a variety of simulated driving scenarios, including normal driving, yawning, slow blinking, falling asleep, laughing, etc., under different illumination conditions. Further, the 90 videos in NTHU-DDD consist of two scenarios, including drowsiness related symptoms (yawning, nodding, slow blink rate) and non-drowsiness related actions (talking, laughing, looking at both sides), are each recorded about 1.5 minutes long.

Although these datasets have been widely used in driver fatigue detection, they are not designed for UAV remote pilot fatigue detection. Specifically, YawDD yawning dataset and DriveFace dataset are video dataset recorded by onboard cameras. The scenes captured are limited to the main driver's seat, while UAV remote pilots often work outdoors and have a wide range of activities. Besides, some datasets are mainly designed for eye state, such as CEW. However, UAV remote pilots sometimes need to wear sunglasses so as to the eyes will be blocked and we cannot access the fatigue feature. Additionally, remote pilots always need to communicate with the air traffic controller frequently, resulting in a continuous mouth pose variation, which cannot be found easily in above datasets.

As analyzed above, the existing fatigue detection datasets are designed for traditional drivers, and the working scenes as well as the contents of UAV remote pilots are quite different from those of traditional drivers. Up to now, to our best knowledge, there is no dataset for UAV remote pilot fatigue detection. In this work, we have established a new UAV remote pilot fatigue detection dataset (OFDD) to facilitate future research. The details about our dataset are discussed in the Section 2.1.

2.1. Database establishment

We present our OFDD dataset from aspects of category, procedure and peculiarity.

Category. As pointed out in the literature [33], the face is a non-rigid complex structure. Fatigue is determined by many factors, including posture, light, facial expression, etc. Therefore, in order to diversify the content of our OFDD and represent the different working states for UAV remote pilot, we make our OFDD UAV fatigue detection dataset from the perspectives of gender, lighting, occlusion and action, and the keyword hierarchy of video categories is shown in Figure 1.

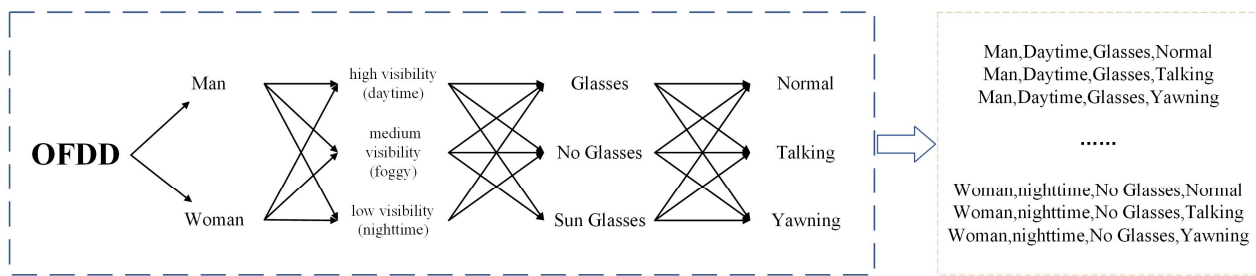


Figure 1. OFDD dataset classification by content. Classify the videos in OFDD according to the content. Categories are distinguished from gender, visibility, occlusion and action. Each category/subclass contains one video.

In practice, we asked UAV remote pilots of different genders to wear glasses/sunglasses or not; then they were asked to work in three different mouth conditions: 1) keep their mouth closed (without speaking), 2) speaking, 3) yawning. Because the UAV remote pilots have to work in different visibility environment, we also recorded our video from three different visibility environment: low visibility (nighttime), medium visibility (foggy) and high visibility (daytime). The right part of in Figure 1 shows the scenes after all conditions combined, and there are 48 combined scenes with different conditions.

Procedure. In practice, the video is collected by Hikvision E12 camera, and the camera is set in front of the UAV remote pilot. The detailed setting includes 640 x 480 resolution, 24-bit true color (RGB) and 30 frames per second. In Figure 2, it is a partial picture of our OFDD dataset, which can be viewed from left to right under different visibility conditions. During the experiment, each participant was asked to free control the UAV during nine different situations with randomly sequence: 1) high visibility (daytime), no talking, 2) high visibility (daytime), talking, 3) high visibility (daytime), yawning, 4) medium visibility (fog), no talking, 5) medium visibility (fog), talking, 6) medium visibility (fog), yawning, 7) low visibility (nighttime), no talking, 8) low visibility (nighttime), talking and 9) low visibility (nighttime), yawning. Also, each participant was required to do the same things with normal glasses and sun glasses orderly. The combination of these completed stages helps to create a more realistic scene, distinguish these situations at the same time, and accurately detect the yawning posture. The length of each video ranges from 30 to 100 seconds, and after 48 videos were collected, with a total duration of 63 minutes and 14 seconds, the audio was deleted in order to reduce the volume.

Peculiarity. Our OFDD UAV remote pilot fatigue detection dataset has all the following important features.

- 1) Our OFDD dataset is the first UAV remote pilot fatigue detection dataset.
- 2) The subjects were UAV remote pilots working in different visibility.
- 3) The position of the camera and the angle of capturing the remote pilot are the same as those in the actual control monitoring system.
- 4) The remote pilot shall be photographed under various mouth opening conditions, such as normal (no talking), talking or yawning conditions and etc.
- 5) Face covers, such as wearing glasses and sunglasses, are considered to test the robustness of the algorithm.
- 6) The dataset is enough to be statistically significant.



Figure 2. Six examples of our OFDD dataset. The situations, from top to bottom and left to right, are 1) male, medium visibility (fog), with glasses, without talking, 2) female, medium visibility (fog), without glasses, without talking, 3) male, high visibility (daytime), without glasses, yawning, 4) female, high visibility (daytime), with sun glasses, yawning, 5) male, low visibility (nighttime), with glasses, without talking, 6) female, low visibility (nighttime), with glasses, yawning, respectively.

2.2. Database analysis

In this section, we deeply mined our OFDD UAV remote pilot fatigue detection dataset, and further made statistics and analysis on each video. After that, we summarized the correlation and difference of fatigue features between traditional car driver and UAV remote pilot.

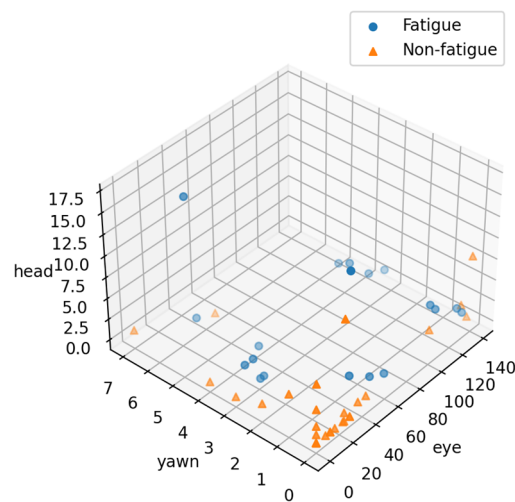


Figure 3. The number of fatigue symptoms including blinks times, yawning times and nodding times of our OFDD dataset (Each yellow and blue scatter point represents a non-fatigue case and a fatigue case, respectively.).

2.2.1. Correlation between multi-facial characteristic features and fatigue of UAV remote pilot

First, we investigated the relationship between multiple facial features and fatigue state in our OFDD dataset. We have counted the number of fatigue symptoms of our OFDD dataset, as shown in Figure 3, including blinking times, yawning times and nodding times. We can observe from Figure 3 that the number of fatigue symptoms of a fatigued UAV remote pilot is much greater than those of non-fatigued, which is in agreement with traditional car driver.

This result implies a high correlation between multiple facial features and fatigue of UAV remote pilot and we can further infer that the fatigue of UAV pilot can be identified by the frequency of three fatigue symptoms: blinking, nodding and yawning.

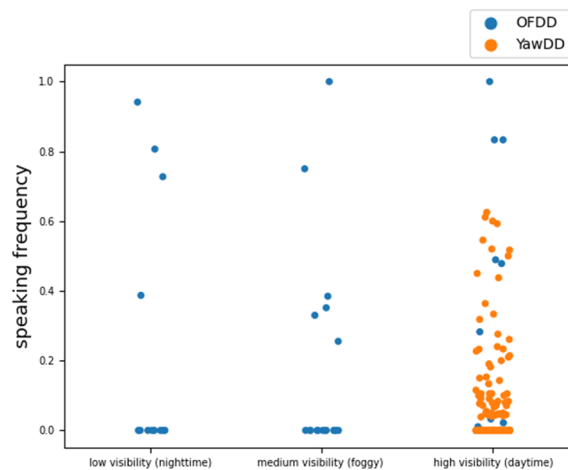


Figure 4. The speaking frequency distribution of samples under three different visibility conditions over OFDD and traditional car driver fatigue dataset YawDD (Each yellow and blue scatter point represents one sample in YawDD and OFDD, respectively.).

2.2.2. Difference between traditional driver fatigue dataset and UAV remote pilot fatigue dataset

In addition, it is interesting to explore the difference between traditional driver fatigue dataset and UAV remote pilot fatigue dataset. We find that UAV remote pilots have two obvious differences compared with traditional car drivers: 1 Speaking frequency 2 Visibility condition. Specifically, we specially counted the speaking frequency under three different visibility conditions over OFDD and YawDD. As shown in Figure 4, the traditional driver is sitting in the driver seat (high visibility) and there is no data related to medium visibility (foggy) and low visibility (nighttime) scenes over YawDD dataset while UAV remote pilot need to work in all the above scenarios. Secondly, one also finds from Figure 4 that the speaking frequency of UAV remote pilots in OFDD dataset is significantly higher than that of traditional drivers in YawDD dataset. The reason may be that the UAV remote pilot always need to communicate with the air traffic controller frequently during working.

However, in the traditional driver fatigue detection, the opening and closing aspect ratio of driver's mouth is often used as the key element to judge fatigue, and the time sequence information do not be considered. This result implies the traditional fatigue detection method is not suitable for UAV remote pilots and temporal information should be introduced to distinguish yawning from speaking.

3. Proposed methods

3.1. Framework

For detecting fatigue state of UAV remote pilot, we develop a new DNN based detection network YPS network. In Figure 5, our whole detection network can be divided into face detection module, multi-feature extraction module, and temporal fatigue decision module. According to the findings in Section 2, multiple facial fatigue features is highly correlated to the fatigue state, and the temporal information is essential for fatigue detection of UAV remote pilots. As such, we first regard two contiguous frames X_t and X_{t-1} as the input of face detection module and the face detection module detects the human face through a state-of-the-art convolutional neural network. Then, the detected human faces serve as the input of multi-feature extraction module. After feature extraction, the temporal fatigue decision module integrates both the variety of different fatigue features and temporal information to make a final decision.

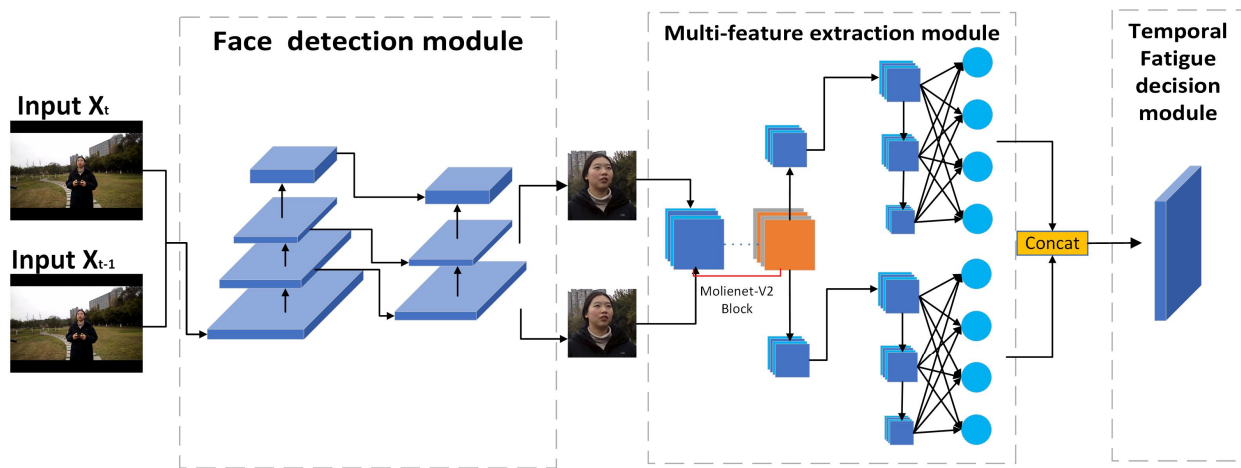


Figure 5. Our YPS network architecture is used for face detection, multi-feature extraction and fatigue determination, following Yolov5 and PFLD (Note that the training process is not annotated in the figure.).

3.2. Face contour positioning

In face detection module, we employ the state-of-the-art detection network, YOLOv5 [34]. YOLOv5 has four different models including YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. We choose YOLOv5s as our face detection model. YOLOv5s has four parts: input, backbone, neck and prediction and the inference time is about 0.007 s, that is 140 frame per second.

In practice, we respectively use the architecture of CSPDarknet53 with an SPP layer as our backbone, PANet as Neck. To further optimize the detection performance of human face, we follow the origin data enhancement framework of mosaic data enhancement and reset the anchor size to match the size of the network receptive field and the image scale. The details about our backbone choice and anchor setting are discussed in experiment.

3.3. Multi-feature extraction module

We present our multi-feature extraction module from four subcomponents: face key point detection, eye fatigue feature, mouth fatigue feature and head posture fatigue feature.

3.3.1. Face key point detection

First, following most of the existing fatigue detection technologies [35–37], we find the key point of the face. The number of face key points of traditional face key point detection methods includes 5 points, 68 points and 106 points, as shown in Figure 6. In practice, we choose 68 key points because 68 key points can accurately display the local features such as human eyes and mouth of a human head, reduce the operation time and improve the real time performance of face recognition [38].

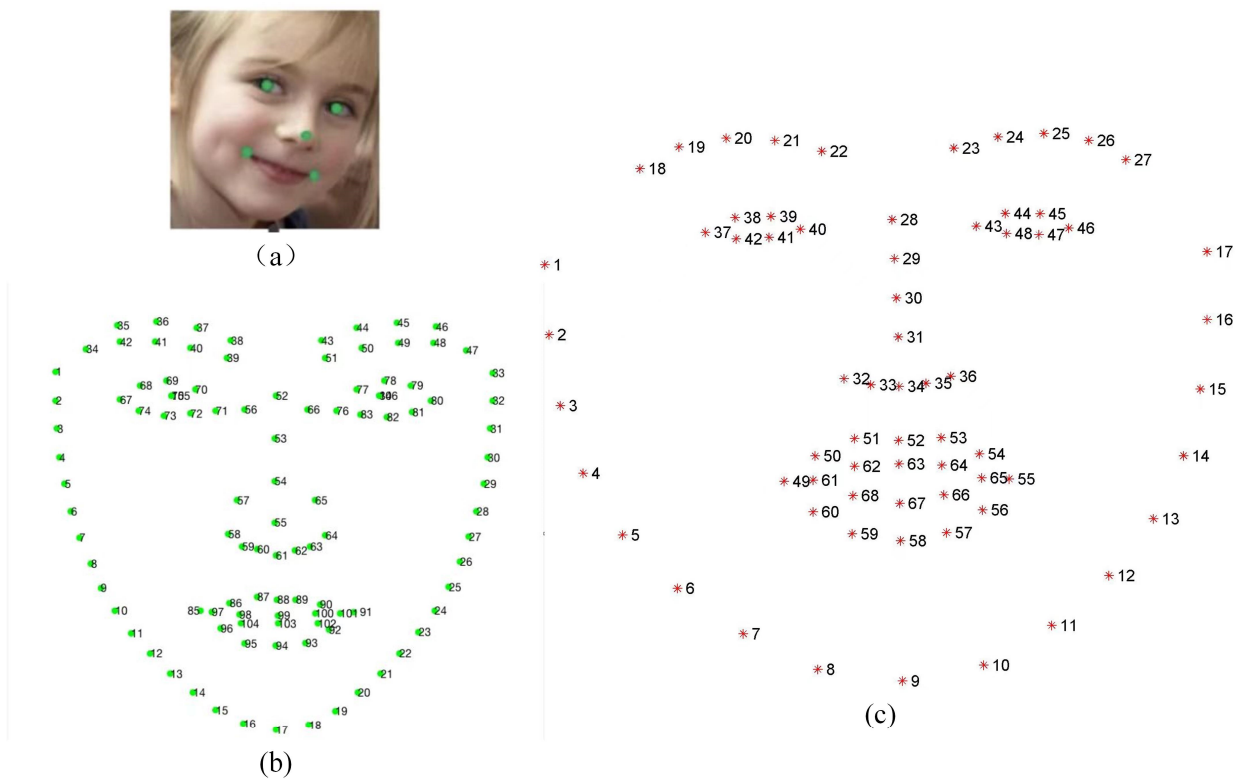


Figure 6. At present, there are three mainstream facial contour localization methods: (a) is the feature map of five key points of the face; (b) is the feature map of 106 key points of the face; (c) is a feature map of 68 key points of the face.

As for the face key point detection model, we choose PFLD face key point detection model [39]. Besides, based on our finding in OFDD, we introduce temporal feature and propose a face key point detection model for UAV remote pilots. The architecture of face key point detection model is shown in Figure 7.

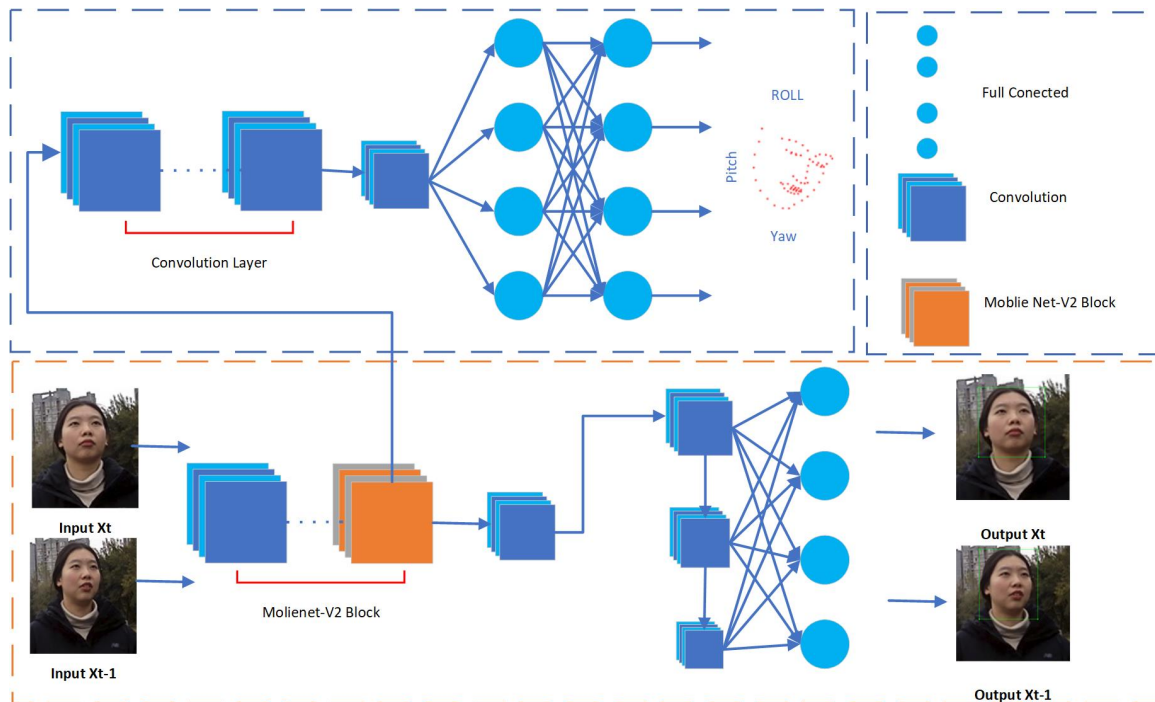


Figure 7. Architecture of face key point detection model for predicting 68 face key points with extra temporal features, following PFLD.

The L2 loss function of key point detection module can be expressed as

$$L_2 = \frac{1}{W} \sum_{w=1}^W \sum_{n=1}^N [\sum_{c=1}^C \omega_n^c \sum_{k=1}^K (1 - \cos \theta_n^k)] \|d_n^w\|_2^2 \quad (1)$$

where W is the number of samples; N is the number of feature points; $\|d_n^m\|$ is the distance measurement of the n -th feature point in w samples; C is the number of different human faces (side face, front face, head up, head down, expression and occlusion); ω_n^c is the given weight corresponding to category C ; $K = 3$ is the three dimensions of face pose, which is namely yaw, pitch, roll angle (the higher the angle, the greater the weight); θ is the deviation value of predicted value in yaw angle, pitch angle and roll angle.

3.3.2. Eye fatigue feature

In order to better describe the eye state, the midpoint of key points 38 and 39, as shown in Figure 6(c), is used to represent the upper eyelid position of the left eye and the midpoint of key points 41 and 42 is used to represent the lower eyelid position of the left eye. As for the right eye, we use the midpoint of key points 44 and 45 to represent the upper eyelid position, and use the midpoint of key points 47 and 48 to represent the lower eyelid position. As a result, the eye state can be calculated by following formula.

$$E = \frac{\text{dist}(E_{lu}, E_{ld})}{\text{dist}(E_{olu}, E_{old})} + \frac{\text{dist}(E_{ru}, E_{rd})}{\text{dist}(E_{oru}, E_{ord})} \quad (2)$$

$$E_{lu} = \left(\frac{x_{38} + x_{39}}{2}, \frac{y_{38} + y_{39}}{2} \right) \quad (3)$$

$$E_{ld} = \left(\frac{x_{41}+x_{42}}{2}, \frac{y_{41}+y_{42}}{2} \right) \quad (4)$$

$$E_{ru} = \left(\frac{x_{44}+x_{45}}{2}, \frac{y_{44}+y_{45}}{2} \right) \quad (5)$$

$$E_{rd} = \left(\frac{x_{47}+x_{48}}{2}, \frac{y_{47}+y_{48}}{2} \right) \quad (6)$$

Among them, E_{lu} , E_{ld} , E_{ru} and E_{rd} respectively represent the upper and lower eyelid positions of the left and right eyes, which are calculated according to the coordinates of eye key points in the detection process of each frame; $dist_{olu}$, $dist_{old}$, $dist_{oru}$, $dist_{old}$ represent the distance between the upper and lower eyelids in the initial state, which only needs to be calculated once in the initial state of the system. The value of distance is calculated by the Euclidean distance.

3.3.3. Mouth fatigue feature

According to the findings in Section 2, the temporal information is essential for fatigue detection of UAV remote pilots, especially in distinguishing yawning and speaking with mouth fatigue feature.

There are usually three states of the mouth: normal, speaking and yawning. People yawn more frequently when they are tired. Therefore, the features of the mouth are also an important feature for judging fatigue. From the Section 2, it can be seen from the statistical data of OFDD dataset that the frequency of yawning can effectively reflect the fatigue state of UAV remote pilots. When the UAV remote pilot yawns due to fatigue, the mouth will open to the greatest extent, and the mouth will increase in height and decrease in width. In order to measure the fatigue state represented by these two indicators, the mouth aspect ratio M is introduced, and its calculation formula is

$$M = \frac{(|y_{51}-y_{59}|)+(|y_{53}-y_{57}|)}{2(x_{55}-x_{49})} \quad (7)$$

where x_{49} and x_{55} are the abscissa of the two key points on the left and right of the mouth; y_{51} , y_{53} , y_{57} and y_{59} are the ordinates of the four key points above and below the mouth.

When determining the fatigue state according to the mouth state, we calculate the ratio between the origin distance between left and right corners of the mouth and the current distance between left and right corners of the mouth. If the ratio is greater than the given threshold value of 1.2, it is recognized as yawning.

In order to eliminate the interference of speaking and other actions, it is not reliable to judge the fatigue state only by the mouth state of current frame, so we introduce extra temporal information of mouth feature from last frame to improve the accuracy of judging the fatigue state. As a result, if the value of M is greater than the given threshold value of 1.2 for a continuous period of time (2S), it can be determined that the current UAV remote pilot is in a fatigue state.

3.3.4. Head posture fatigue feature

People tend to nod their heads frequently or lower head their heads for a long time when they are in fatigue. Therefore, nodding (lower head) situation can be used as a fatigue characterization parameter. When people lower their heads, the distance from the midpoint of eyes to mouth in the

photos taken by the camera will be significantly shorter than that when people raise their heads. Therefore, the change of this distance is used in this work as the change of nod (lower head) posture.

In Figure 6(c), the center position of left eye $[x_{left}, y_{left}]$ and right eye $[x_{right}, y_{right}]$ can be expressed as:

$$\begin{aligned}x_{left} &= \frac{1}{4}(x_{38} + x_{39} + x_{41} + x_{42}) \\y_{left} &= \frac{1}{4}(y_{38} + y_{39} + y_{41} + y_{42}) \\x_{right} &= \frac{1}{4}(x_{44} + x_{45} + x_{47} + x_{48}) \\y_{right} &= \frac{1}{4}(y_{44} + y_{45} + y_{47} + y_{48})\end{aligned}\quad (8)$$

Then the linear fitting function of the center of both eyes is

$$\frac{x - x_{left}}{x_{right} - x_{left}} = \frac{y - y_{left}}{y_{right} - y_{left}} \quad (9)$$

where x_{left} , x_{right} , y_{left} and y_{right} are the horizontal and vertical coordinates of the midpoint of the left and right eyes respectively; x , y is the horizontal and vertical coordinates of the linear function of the two eyes. When Eq (7) is transformed into standard $Ax + By + C = 0$, the values of A , B , C are

$$\begin{cases} A = y_{right} - y_{left} \\ B = x_{left} - y_{right} \\ C = x_{right}y_{left} - x_{left}y_{right} \end{cases} \quad (10)$$

Then, the distance from the midpoint of the two eyes of the face to the line is calculated by using point 52, that is, the distance d from the midpoint of the two eyes on the image to the mouth can be calculated by

$$d = \frac{\left| \frac{Ax_{52} + By_{52} + C}{\sqrt{A^2 + B^2}} \right|}{\cos(\text{pitch}_1)} \quad (11)$$

With reference to p80 standard [40], the threshold value of distance d is set to 0.8. In other words, when $d \leq 0.8d_{max}$, it can be determined that the driver is in a low head state. If the remote pilot hold bow state more than 5 s, it can be directly judged that the pilot is in the fatigue state. If the remote pilot hold less than 5 s, the percentage of bow state can be calculated by

$$H = \frac{t_{down}}{T} \times 100\% \quad (12)$$

where t_{down} is the head lowering time; T is the total time; H is the head lowering frequency. The greater the value, the deeper the fatigue degree. H can also be used to determine the fatigue state. The Euler angle of the remote pilot's head posture is divided into pitch angle, yaw angle and roll angle, that is, the Euler angle of the remote pilot's head rotating around the x , y and z axes respectively. According to the research [41], the maximum angle range of the human head rotating around the X axis of the central axis is $[-60.4^\circ, 69.6^\circ]$, the maximum range of rotating around the Y axis is $[-90^\circ, 75^\circ]$,

and the maximum range of rotating around the Z axis is $[-40.9^\circ, 36.3^\circ]$. Thus, the abnormal angle can be calculated according to the remote pilot's real time pitch, yaw and roll. In this paper, the judgment of abnormal angle refers to p80 standard [41] in PERCLOS, that is, when the remote pilot's head angle exceeds 80% of the normal range, it is determined as abnormal angle.

3.4. Temporal fatigue decision module

In temporal fatigue decision module, the cross-time temporal information integration procedure and multi-fatigue feature fusion SVM are proposed. In particular, the cross-time temporal information integration procedure can be used to encode temporal information from different frames and the multi-fatigue feature fusion SVM is developed to cumulate multi-fatigue feature and make a final decision.

3.4.1. Cross-time temporal information integration procedure

The fatigued state of the human face can be comprehensively expressed according to the eyes (open, closed), mouth (closed, talking, yawning) and head yaw angle, pitch angle and roll angle. After the PFLD obtains 68 key points of the face, for the video image frames collected on the continuous time node sequence (T_t, T_{t-1}) , each frame image can be used as a state according to the key points of the eye, mouth and head posture. Through 68 key points and attitude angles, the fatigue features E , M and H can be obtained according to Eqs (2), (7) and (12). As for the Euler angle, a standard face is defined firstly (the average value is taken from a stack of frontal faces), and 11 key points are fixed on the main plane of the face as the reference for all training faces. Then, using the corresponding 11 key points and the reference matrix of the estimated rotation matrix, the Euler angle is calculated by the rotation matrix.

Finally, combining the above features, we can get the following feature matrix (13).

$$\begin{bmatrix} T_t & E_t & M_t & H & yaw_t & pitch_t & roll_t \\ T_{t-1} & E_{t-1} & M_{t-1} & H & yaw_{t-1} & pitch_{t-1} & roll_{t-1} \end{bmatrix} \quad (13)$$

where yaw, pitch and roll are the output vector of the PFLD auxiliary subnet (upper left blue box in Figure 7) to predict the human face posture during training.

3.4.2. Multi-fatigue feature fusion SVM

Support vector machine (SVM) is a highly accurate and efficient supervised machine learning algorithm, which is mainly used to solve the data classification problem in the field of pattern recognition. After cumulating the deep fatigue features, this work uses SVM algorithm to establish a fatigue judgment model for UAV remote pilots.

Because the collected data of eye, mouth and head features are not regular, this kind of problem belongs to nonlinear classification problem. According to the time point when the characteristics change, the awake state and the fatigue state are distinguished. The identified fatigue characteristics of eyes, mouth and head are classified as fatigue state, and the classification label is 1; classify others as awake, and the classification label is 0.

4. Experimental results

In this section, our fatigue detection method of UAV remote pilot is tested, and a comparative experiment is carried out on OFDD dataset and other dataset to verify the effectiveness of the algorithm.

4.1. Settings

In our experiment, videos in 48 subcategories of OFDD UAV remote pilot fatigue detection dataset were randomly divided into training set, verification set and test set. Specifically, we intercepted continuous frame images of the video in the 48 subcategories of OFDD dataset. In order to balance the data in the experiment, we only clipped the first 30 s of videos in each subcategory for training. The resolution of each continuous frame image is 640×480 , the horizontal resolution is 96 dpi, the vertical resolution is 96 dpi. In order to evaluate the performance of our proposed fatigue detection method for UAV remote pilots, we compared 10 other more effective fatigue detection methods. All experiments were conducted on a computer using Intel (R) core (TM) i7-4770. These include CPU@3.4 GHz, 16 GB ram and two NVIDIA GeForce GTX 3080 GPU.

The accuracy of face region and key point marking samples in the self-collected dataset is crucial to the accuracy of in depth learning. A picture is intercepted every 2 frames from the videos in the 48 subcategories in the OFDD dataset. A total of 21,860 frames of pictures are intercepted and the added face frame and key point positions are marked, like in Figure 8.

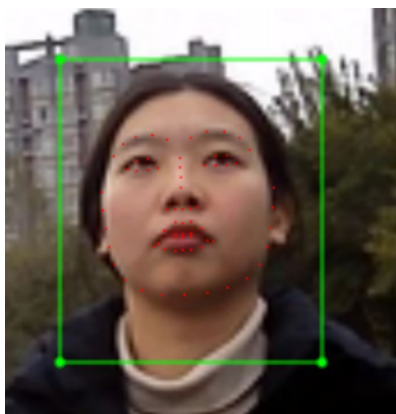


Figure 8. The video in OFDD dataset is intercepted and annotated, and the annotation example of the JPG picture of 06611.

4.2. Performance analysis of our method

4.2.1. Improved anchor settings

The setting of anchor is very important for face position. At present, most popular face detection algorithms use Anchor mechanism in the detection process, such as Face R-CNN [42], FDRNet [43], DSFD [44], RetinaFace [45], DFS [46], S³fd [47], FAN [48], etc. The anchor should cover the target to be detected, and the scale of the Anchor should match the size of the network receptive field and the image scale.

In order to ensure that human face is in the picture frame of the camera, the camera's angle of view should be as wide as possible, so there will be a lot of interference information in the background. In order to reduce the impact of other backgrounds, this paper sets an appropriate Anchor according to the receptive field of the human head and the scale of the feature map.

According to the Anchor setting method in DFS, FAN, the aspect ratio of the anchor is set to 1:1.5 in the three higher-level face detection modules FDM1, FDM2 and FDM3, while the aspect ratio of the anchor is set to 1:1 in the lower-level detection modules FDM4 and FDM5. This is because the smaller size face is closer to a square than the large-scale front face. In addition, each point in the feature map of each face detection module is set with three Anchors of different sizes to improve the detection accuracy. The setting of Anchor size and proportion in each level of feature map in the five detection modules is shown in Table 1.

Table 1. In our face detection module, we modify the anchor box size according to the particularity of the face. Lists the size and scale of each anchor.

Detection module	Stride	Anchor (width)	Aspect ratio
FDM5	4	121,620	1:1
FDM4	8	243,240	1:1
FDM3	16	486,480	1:1.5
FDM2	32	96,128,160	1:1.5
FDM1	64	192,256,320	1:1.5

Where the stride indicates the relationship between the size of the current feature map and the input image. For example, if Stride is 4, the width and height of the current feature map are respectively 1/4 of the width and height of the input image. The corresponding Stride of the detection module determines its Anchor interval in the input image. For example, the Stride in DM5 is 4 and the Anchor size is 12×12 , which Indicates that there is a 12×12 Anchor for every 4 pixels on the input image, and the ratio of anchors is 3–5 times that of Stride, which ensures that anchors of different scales have the same density on the image, and faces of various scales can roughly match the same number of anchors. In the training process, if the IOU of face annotation box and Anchor is greater than 0.5, it is considered that the box contains positive face samples. If the IOU is less than 0.3, it is considered that the box contains negative background samples without faces. If the IOU is between 0.3–0.5, it is ignored in the training process.

The labeled ODFD dataset is divided into training set, verification set and test set according to the ratio of 8:1:1. After introducing the training set data into the modified Anchor, yolov5 is trained. The trained model is fitted and verified with the validation set data, and then tested with the test dataset. The test identification results are shown in the table.

It can be seen from Table 2 that the recognition rate of face area reaches 99.5%, which is 1.2% higher than the original yolov5. This experimental result shows that our improved anchor setting for UAV remote pilot face detection is effective.

Table 2. Compare the performance of face detection module with modified anchor frame and object detection module with yolov5s.

Face detection network	Number of face photos (piece)	Number of recognized faces (piece)	Recognition rate (%)
Yolov5s	21,860	21,488	98.3%
Ours	21,860	21,751	99.5%

4.2.2. Backbone effect comparison

In this subsection, we compare the performance of different backbone in our UAV remote pilot fatigue detection model, including ResNest50 [49], MobileNetV2 [50], HRNetV2 [51] and EfficientNet [52]. Table 3 demonstrates the performance of different backbone.

Table 3. Performance of our method with different backbone structures.

Name	Params	Mean error	Failure rate
ResNest-50	122.27 M	0.046	0.038
MobileNetV2_0.25	1.09 M	0.075	0.174
MobileNetV2_1.00	7.28 M	0.065	0.127
HRNetV2	545.07 M	0.066	0.125
EfficientNet-B0	16.67 M	0.064	0.119
EfficientNet-B1	26.37 M	0.075	0.149
EfficientNet-B2	30.85 M	0.071	0.145
EfficientNet-B3	42.29 M	0.099	0.136
EfficientNet-B4	68.34 M	0.077	0.164
EfficientNet-B5	109.34 M	0.094	0.173
EfficientNet-B6	156.34 M	0.081	0.175
EfficientNet-B7	244.03 M	0.081	0.196

The performance analysis of the backbone structure in the network structure is shown above. In the experiment, we tested 13 different versions of the backbone structure. Here, we mainly use the backbone parameter size, average error rate and failure rate to evaluate the performance of the adopted backbone structure. It can be seen from the results in Table 3 that the average error rate and failure rate of resnet50 are 0.046 and 0.038 respectively, reaching the lowest point in the experiment, but the parameters of the backbone structure are too large, about 122.27 M; the average error rate and failure rate of Blaze landmark are 0.069 and 0.131 respectively. The backbone parameter size is 7.52 M, and the first iteration time (seconds) is 0.171. The backbone parameter size is in the third place. The average error rate and failure rate of HRNetV2 are 0.066 and 0.125 respectively, but the parameter is 545.07 M. The big backbone structure is not suitable for actual deployment of UAV fatigue detection, so we choose efficient as the backbone of our model.

We also tested eight EfficientNet backbone structures from efficientnet-b0 to efficient-net-b7. The experimental results show that with the increase of backbone size, their average error rate and failure rate do not decrease significantly, while the backbone structure parameters increase gradually. The network backbone structure with the best performance in the experiment are MobileNetV2_0.25 and MobileNetV2_1.00, not only are the size relatively small, but also the average error rate and failure

rate relatively low. For MobileNetV2_0.25 and MobileNetV2_1.00, although both are lightweight network backbone, but the average error rate and failure rate of MobileNetV2_1.00 are only 0.065 and 0.124, both better than MobileNetV2_0.25. These experimental results show that the error rate of MobileNetV2_1.00 selected in the paper is low, the detection accuracy is higher, and the comprehensive detection effect is better. The model used in this paper is used to detect the fatigue of UAV remote pilots, which is easy to be applied in low computing power devices such as mobile devices, and is suitable for the fatigue state detection of UAV remote pilots in outdoor scenes.

4.2.3. Ablation experiment of characteristic parameters

According to the statistical analysis of OFDD UAV remote pilot fatigue detection dataset, the fatigue state of UAV remote pilot is highly correlated to multi-facial fatigue characteristics, including eye fatigue feature, mouth fatigue feature and head posture fatigue feature, and also the temporal information is essential for fatigue detection of UAV remote pilots. In order to further verify the accuracy of our proposed method, we used eye feature parameter E, mouth feature parameter M, head feature parameter H, pitch angle pitch, yaw angle yaw and roll angle to perform ablation experiments. The results are shown in Table 4.

Table 4. Influence of different fatigue characteristic parameters on the accuracy of fatigue testing.

Fatigue characteristic parameters	Accuracy rate (%)
E + pitch + yaw + roll	93.26
M + pitch + yaw + roll	90.13
H + pitch + yaw + roll	86.76
E + M + H + pitch + yaw + roll	97.32

It can be seen from Table 4 that when the eye feature parameter E is used to train the network with the angle of yaw, pitch and roll, the detection accuracy is 93.26%; when the mouth feature parameter M and yaw, pitch and roll angles are used to train the network, the detection accuracy is 90.13%; when the head feature parameters H and yaw, pitch and roll angles are used to train the network, the detection accuracy is 86.76%. It can be seen that when a single feature parameter is used for training, the eye feature parameter E has the highest accuracy rate, followed by the mouth feature parameter M, and finally the head feature parameter H. Ablation experiments show that eye feature parameters, mouth feature parameters and head feature parameters are all important for UAV remote pilot fatigue detection.

4.2.4. Ablation experiment of face key point detection models

We conducted ablation experiments on face key point detection modules and compared five face key point detection modules: DSFD [44], RetinaFace [45], PFLD [39] (no auxiliary) and PFLD [39], where the first three model do not include head posture features (pitch + yaw + roll).

It can be seen from Table 5 that although the parameters and FLOP of DSFD and RetinaFace are very low, the accuracy of the face detection module is about 1 and 2% lower than that of PFLD (no auxiliary) respectively. For a PFLD module with the auxiliary, it can not only output 68 key points of face coordinates through the main network but also the yaw, pitch and roll attitude Euler angles of the

head through its auxiliary network, which is helpful to improve the detection accuracy. As a result, PFLD module not only achieves the highest accuracy, but also has the smallest computational burden.

Table 5. Accuracy and computational complexity of different face key point detection mode architectures.

Module	Head Posture Feature	Backbone	Accuracy	Params (M)	Flops (G)
DSFD	×	MobileNetV2_0.25	86.25	10.26	22.55
RetinaFace	×	MobileNetV2_0.25	87.78	8.43	18.34
PFLD (no auxiliary)	×	MobileNetV2_0.25	88.64	7.68	16.92
PFLD	√	MobileNetV2_0.25	97.32	8.21	17.53

4.3. Evaluation on our OFDD database

In this section, in order to verify the application particularity of UAV remote pilot fatigue detection and the superiority of the method proposed in this paper, we will compare the accuracy of our proposed method for UAV remote pilot fatigue detection with that of other traditional driver's fatigue detection methods. Including the LSTMs of literature [20], the Haar + ELM of literature [21], the RF-DCM of literature [22], the MSP-Net of literature [23], the HOG-SVM of literature [24], the FDRNet of literature [25] and the GFDN of literature [26]. The results are shown in Table 5.

Table 6. Comparison of our fatigue detection method with other fatigue detection methods on OFDD dataset.

Literature	Algorithm	Accuracy rate (%)
[20]	LSTMs	89.23
[21]	Haar + ELM	92.71
[22]	RF-DCM	90.12
[23]	MSP-Net	93.69
[24]	HOG-SVM	94.04
[25]	FDRNet	94.61
[26]	GFDN	94.88
Ours	Our method	97.32

In our experiments, we measure the performance of the algorithm with accuracy. As can be seen from Table 6, when did the experiment on our OFDD dataset, among all the comparison methods, the accuracy of LSTMs [20] is the lowest with 89.23% accuracy. The accuracy rates of Haar + ELM [21], RF-DCM [22], MSP-Net [23], HOG-SVM [24], FDRNet [25] and GFDN [26] are 92.71, 90.12, 93.69, 94.04, 94.61 and 94.88% respectively. Among all the experimental methods, the accuracy of the fatigue detection method for UAV remote pilots proposed by us is the highest, reaching 97.32%, 8.09% higher than LSTMs, and 2.44% higher than GFDN. The experimental results show that our method can detect the fatigue of UAV remote pilots with state-of-the-art performance. We hold the view that the main reasons for the superior performance of our method in OFDD UAV remote pilot fatigue detection dataset are as follows: 1) Our method adjusts the size of network input and anchor box in the face detection part, which improves the accuracy of face detection. 2) The PFLD module is fused to extract

the features of eye, mouth and head features, which improves the extraction accuracy of face key points. 3) The temporal information is introduced to reduce the misjudgment of mouth movements such as speaking and improve the accuracy of fatigue detection.

We selected some samples in OFDD dataset to analyze the detection results of different algorithms, shown in Figure 9. The first picture shows the normal working scene of a male remote pilot without glasses under daytime conditions. The second picture shows the normal working scene of a female remote pilot wearing sunglasses during the daytime. The third picture shows a male remote pilot without glasses yawning in fog. The fourth picture shows the working scene of a female remote pilot without glasses talking in fog. The fifth picture shows the normal working scene of a male remote pilot without glasses at night. The sixth picture shows the normal working scene of a female remote pilot wearing glasses at night.






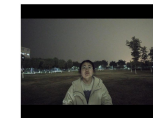
Input						
LSTM_s	Successfully	Successfully	Failed	Failed	Successfully	Successfully
Haar+ELM	Successfully	Successfully	Failed	Successfully	Successfully	Successfully
RF-DCM	Successfully	Successfully	Failed	Successfully	Successfully	Successfully
MSP-Net	Successfully	Successfully	Successfully	Successfully	Failed	Failed
HOG-SVM	Successfully	Failed	Successfully	Successfully	Successfully	Failed
FDRNet	Successfully	Failed	Successfully	Successfully	Successfully	Failed
GFDN	Successfully	Failed	Successfully	Successfully	Successfully	Successfully
Ours	Successfully	Successfully	Successfully	Successfully	Successfully	Successfully

Figure 9. Screenshots of videos of 6 categories randomly selected from the test set of OFDD dataset (Through the six videos describing different visibility, gender, occlusion and status, the proposed method is compared with other fatigue detection methods.).

It can be seen from Figure 9 that the LSTMs algorithm fails to detect in medium visibility (fog). We think the reason for the failure may be that LSTMs is a fatigue detection method based on human posture. However, in the fog scene, the remote pilots are far away from the camera and their body postures are not clear. Haar + ELM and RF-DCM do not perform well in the graph of yawning in foggy days. This can be caused by the above two algorithms are the fatigue detection methods for traditional drivers based on mouth features, which may mislead the decision in frequent speak scene. MSP-Net method successfully distinguishes yawning and speaking, but the detection effect is poor in the night scene. Although GFDN has successfully extracted the eye feature parameters in the scene of wearing glasses at night, this method fails in the scene of wearing sunglasses with serious masking. The fatigue detection method proposed in this paper is suitable for UAV remote pilots. In view of the frequent

speech of UAV remote pilots and the working scene under high visibility, it combines the eye fatigue feature, mouth fatigue feature and head posture fatigue feature, and achieves the best detection accuracy in OFDD dataset.

4.4. Evaluation on YawDD databases

In order to evaluate the generalization ability of our proposed method and further verify the effectiveness of our method. We conducted experiments on the public dataset YawDD dataset. The algorithm in this paper is compared with other latest fatigue detection algorithms in YawDD dataset, and the results are shown in Table 7.

Table 7. Comparison of our fatigue detection method with other fatigue detection methods on YawDD dataset.

Literature	Algorithm	Accuracy rate (%)
[20]	LSTMs	92.71
[21]	Haar + ELM	93.90
[22]	RF-DCM	89.42
[23]	MSP-Net	95.36
[24]	HOG-SVM	96.40
[25]	FDRNet	98.96
[26]	GFDN	97.06
Ours	Our method	97.05

It can be seen from Table 7 that the detection accuracy of the method proposed in this paper is 97.15% in the public dataset of YawDD. The method proposed in this paper is higher than LSTMs [20], Haar + ELM [21], RF-DCM [22], MSP-Net [23] and HOG-SVM [24], but lower than FDRNet [25] and GFDN [26]. Considering that the traditional driver's working scene is relatively fixed and the working environment is relatively single, while the UAV remote pilot's working scene is complex and the environment is diverse, the main research and experimental object of the method proposed in this paper is the UAV remote pilot. The experimental results show that the algorithm can basically meet the fatigue detection of traditional drivers as well as the fatigue detection of UAV remote pilots.

The above experimental results show that, compared with the traditional fatigue driving detection algorithm, the fatigue detection algorithm based on deep learning can automatically extract the features of the input image, mine the deep information of the target image and do not rely on manual participation.

Summing up, compared with other in depth learning models, this paper effectively improves the recognition ability of fatigue characteristics of the model, improves the accuracy of fatigue detection, has strong robustness to the complex environment of the UAV remote pilot's working scene, and can basically realize the real time detection of the UAV remote pilot's fatigue state.

4.5. Analysis

In addition, we additionally tested and analyzed the accuracy of different methods in different working scenarios. Specifically, in order to verify the application particularity of UAV remote pilot fatigue detection and the superiority of the method proposed in this paper, we will compare the

accuracy of our UAV remote pilot fatigue detection method with other traditional pilot fatigue detection methods in the daytime, foggy day and nighttime.

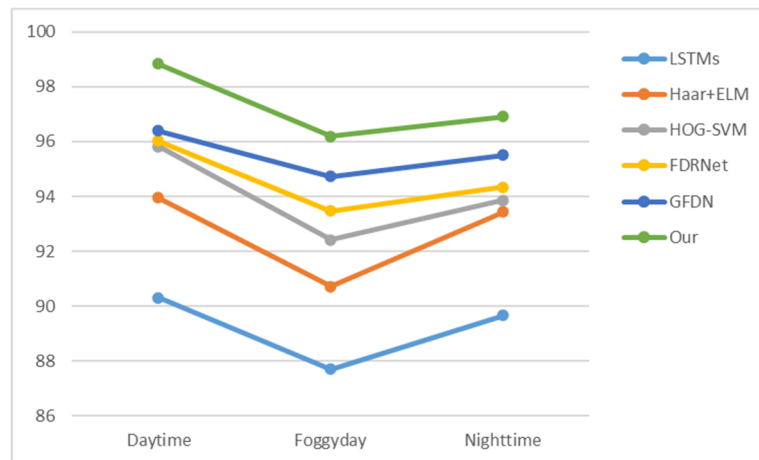


Figure 10. Accuracy comparison of LSTMs [20], Haar + ELM [21], RD-DCM [22], MSP-Net [23], HOG-SVM [24], FDRNet [25], GFDN [26] and our methods in daytime, foggy day and nighttime.

It can be seen from Figure 10 that all algorithms have the best performance in daytime scenes, poor performance in nighttime scenes, and the worst performance in foggy scenes. This can be explained as: firstly, the lighting conditions in foggy days and at night are poor, which is not conducive to obtaining high precision face key points. Secondly, comparing with the nighttime, the sunglasses may block the eye features of remote pilots in foggy days.

As above analysis, we can draw a conclusion that the detection accuracy can keep rising by image enhancement module in foggy and nighttime scenes.

5. Conclusions

Based on the Yolov5 network, this paper integrates the PFLD module to detect the fatigue of UAV operators, which is convenient for application in low computing power devices such as mobile devices. The whole network input and the size of the anchor frame, and temporal feature is introduced to propose a fatigue detection method for UAV operators. This method not only performs well in the traditional motor vehicle driver data and achieves a detection success rate of 97.05%. And it achieves the highest detection success rate of 97.32% on the UAV operator's fatigue detection dataset OFDD. This method has good efficiency and accuracy, and can meet the requirements of UAV operator fatigue control detection based on visual features.

Acknowledgments

This research was funded by science and technology plan project of Sichuan Provincial Department of Science and Technology, grant number 21RKX0103. This research was funded by Civil Aviation Flight University of China "Smart Civil Aviation" special project, grant number ZHMH2022-005. This research was funded by key science and technology project of Civil Aviation Flight University of

China, grant number ZJ2021-11. This research was funded by independent research project of civil aviation flight technology and flight safety key laboratory of Civil Aviation Flight University of China, grant number FZ2022ZZZ06.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. Drones by the Numbers, Federal Aviation Administration, UAS quarterly activity reports, 2022. Available from: https://www.faa.gov/uas/resources/by_the_numbers.
2. Civil Aviation Administration of China, Annual report of Chinese civil aviation pilot development 2021, Dalian maritime university press, Dalian, China, 2022. Available from: https://pilot.caac.gov.cn/jsp/phone/airPhoneNewsDetail.jsp?uuid=abbc4b8e-1d42-4aaa-a0b4-6ab4ef2eb1df&code=Statistical_info#down.
3. N. Aung, P. Tewogbola, The impact of emotional labor on the health in the workplace: a narrative review of literature from 2013–2018, *AIMS Public Health*, **6** (2019), 268–275. <https://doi.org/10.3934/publichealth.2019.3.268>
4. M. Hoang, E. Hillier, C. Conger, D. N. Gengler, C. W. Welty, C. Mayer, et al., Evaluation of call volume and negative emotions in emergency response system telecommunicators: a prospective, intensive longitudinal investigation, *AIMS Public Health*, **9** (2022), 403–414. <https://doi.org/10.3934/publichealth.2022027>
5. R. Parasuraman, D. R. Davies, Decision theory analysis of response latencies in vigilance, *J. Exp. Psychol. Hum. Percept. Perform.*, **2** (1976), 578–590. <https://doi.org/10.1037/0096-1523.2.4.578>
6. G. P. Krueger, Sustained military performance in continuous operations: combatant fatigue, rest and sleep needs, in *Handbook of Military Psychology*, John Wiley & Sons press, (1991), 255–277.
7. C. D. Wickens, W. S. Helton, J. G. Hollands, S. Banbury, *Engineering Psychology and Human Performance*, Routledge Press, New York, USA, 2021. <https://doi.org/10.4324/9781003177616>
8. A. P. Tvaryanas, *Human factors considerations in migration of unmanned aircraft system (UAS) operator control*, Defense Technical Information Center press, Brooks, USA, 2006.
9. X. Qiu, F. Tian, Q. Shi, Q. Zhao. B. Hu, Designing and application of wearable fatigue detection system based on multimodal physiological signals, in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Seoul, Korea (South), (2020), 716–722. <https://doi.org/10.1109/BIBM49941.2020.9313129>
10. T. Iida, Y. Ito, M. Kanazashi, S. Murayama, T. Miyake, Y. Yoshimaru, et al., Effects of psychological and physical stress on oxidative stress, serotonin, and fatigue in young females induced by objective structured clinical examination: pilot study of u-8-OHdG, u-5HT, and s-HHV-6, *Int. J. Tryptophan Res.*, **14** (2021). <https://doi.org/10.1177/11786469211048443>
11. X. Li, G. Li, L. Peng, L. Yan, C. Zhang, Driver fatigue detection based on speech feature transfer learning, *J. China Railw. Soc.*, **42** (2020), 74–81.
12. G. Zhao, Y. He, H. Yang, Y. Tao, Research on fatigue detection based on visual features, *IET Image Proc.*, **16** (2022), 1044–1053. <https://doi.org/10.1049/ipr2.12207>

13. L. Wang, C. Zhang, X. Yin, R. Fu, H. Wang, A non-contact driving fatigue detection technique based on driver's physiological signals, *Automot. Eng.*, **40** (2018), 333–341. <https://doi.org/10.19562/j.chinasae.qcgc.2018.03.014>
14. X. Li, G. Li, J. Shi, L. Peng, Fatigue driving detection based on speech psychoacoustic analysis, *Chin. J. Sci. Instrum.*, **39** (2018), 166–175. <https://doi.org/10.19650/j.cnki.cjsi.J1702539>
15. W. Zheng, K. Zheng, G. Li, W. Liu, C. Liu, J. Liu, et al., Vigilance estimation using a wearable EOG device in real driving environment, *IEEE Trans. Intell. Transp. Syst.*, **21** (2020), 170–184. <https://doi.org/10.1109/TITS.2018.2889962>
16. X. Wang, C. Xu, Driver drowsiness detection based on non-intrusive metrics considering individual difference, *Accid. Anal. Prev.*, **95** (2016), 350–357. <https://doi.org/10.1016/j.aap.2015.09.002>
17. W. Feng, Y. Cao, X. Li, W. Hu, Face fatigue detection based on improved deep convolutional neural network, *Comput. Intell. Neurosci.*, **20** (2020), 5680–5687.
18. Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou, A survey of convolutional neural networks: analysis, applications, and prospects, *IEEE Trans. Neural Networks Learn. Syst.*, **2021** (2021), 1–21. <https://doi.org/10.1109/TNNLS.2021.3084827>
19. D. Chen, F. Liu, Z. Li, Deep learning based single sample per person face recognition: A survey, preprint, arXiv:2006.11395.
20. Y. Ed-doughmi, N. Idrissi, Driver fatigue detection using recurrent neural networks, in *Proceedings of the 2nd International Conference on Networking, Information Systems & Security (NISS19)*, ACM, Rabat, Morocco, (2019), 1–6. <https://doi.org/10.1145/3320326.3320376>
21. C. Zheng, B. Xiaojuan, W. Yu, Fatigue driving detection based on Haar feature and extreme learning machine, *J. China Univ. Posts Telecommun.*, **23** (2016), 91–100. [https://doi.org/10.1016/S1005-8885\(16\)60050-X](https://doi.org/10.1016/S1005-8885(16)60050-X)
22. R. Huang, Y. Wang, Z. Li, Z. Lei, Y. Xu, RF-DCM: multi-granularity deep convolutional model based on feature recalibration and fusion for driver fatigue detection, *IEEE Trans. Intell. Transp. Syst.*, **23** (2022), 630–640. <https://doi.org/10.1109/TITS.2020.3017513>
23. W. Gu, Y. Zhu, X. Chen, L. He, B. Zheng, Hierarchical CNN-based real-time fatigue detection system by visual-based technologies using MSP model, *IET Image Proc.*, **12** (2018), 2319–2329. <https://doi.org/10.1049/iet-ipr.2018.5245>
24. S. Dey, S. A. Chowdhury, S. Sultana, M. A. Hossain, M. Dey, S. K. Das, Real time driver fatigue detection based on facial behaviour along with machine learning approaches, in *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, IEEE, Dhaka, Bangladesh, (2019), 135–140. <https://doi.org/10.1109/SPICSCON48833.2019.9065120>
25. Z. Xiao, Z. Hu, L. Geng, F. Zhang, J. Wu, Y. Li, Fatigue driving recognition network: fatigue driving recognition via convolutional neural network and long short-term memory units, *IET Intel. Transport Syst.*, **13** (2019), 1410–1416. <https://doi.org/10.1049/iet-its.2018.5392>
26. W. Liu, J. Qian, Z. Yao, X. Jiao, J. Pan, Convolutional two-stream network using multi-facial feature fusion for driver fatigue detection, *Future Internet*, **11** (2019), 115. <https://doi.org/10.3390/fi11050115>
27. L. Geng, X. Liang, Z. Xiao, Y. Li, Real-time driver fatigue detection based on morphology infrared features and deep learning, *Infrared Laser Eng.*, **47** (2018), 203009. <https://doi.org/10.3788/IRLA201847.0203009>

28. F. Liu, D. Chen, J. Zhou, F. Xu, A review of driver fatigue detection and its advances on the use of RGB-D camera and deep learning, *Eng. Appl. Artif. Intell.*, **116** (2022), 105399. <https://doi.org/10.1016/j.engappai.2022.105399>
29. S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, B. Hariri, YawDD: A yawning detection dataset, in *Proceedings of the 5th ACM Multimedia Systems Conference*, ACM, Singapore, (2014), 24–28. <https://doi.org/10.1145/2557642.2563678>
30. K. Diaz-Chito, A. Hernández-Sabaté, A. M. López, A reduced feature set for driver head pose estimation, *Appl. Soft Comput.*, **45** (2016), 98–107. <https://doi.org/10.1016/j.asoc.2016.04.027>
31. F. Song, X. Tan, X. Liu, S. Chen, Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients, *Pattern Recognit.*, **47** (2014), 2825–2838. <https://doi.org/10.1016/j.patcog.2014.03.024>
32. C. Weng, Y. Lai, S. Lai, Driver drowsiness detection via a hierarchical temporal deep belief network, in *Asian Conference on Computer Vision (ACCV)*, Springer Cham, (2016), 117–133. https://doi.org/10.1007/978-3-319-54526-4_9
33. R. Gross, Face databases, in *Handbook of Face Recognition*, Springer, (2005), 301–327. https://doi.org/10.1007/0-387-27257-7_14
34. G. Jocher, A. Stoken, J. Borovec, A. Chaurasia, T. Xie, C. Y. Liu, Ultralytics/yolov5: v5.0-yolov5-p6 1280 models, *Zenodo*, **2021** (2021). <https://doi.org/10.5281/zenodo.4679653>
35. J. F. Cohn, A. J. Zlochower, J. Lien, T. Kanade, Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding, *Psychophysiology*, **36** (1999), 35–43. <https://doi.org/10.1017/s0048577299971184>
36. S. Dubuisson, F. Davoine, M. Masson, A solution for facial expression representation and recognition, *Signal Process. Image Commun.*, **17** (2002), 657–673. [https://doi.org/10.1016/S0923-5965\(02\)00076-0](https://doi.org/10.1016/S0923-5965(02)00076-0)
37. N. A. Rahman, K. C. Wei, J. See, RGB-H-CbCr skin colour model for human face detection, *Fac. Inf. Technol.*, **4** (2007), 1–6.
38. M. Sun, L. Liang, H. Wang, W. He, L. Zhao, Facial landmark detection based on cascade convolutional neural network, *J. Univ. Chin. Acad. Sci.*, **37** (2020), 562–569. <https://doi.org/10.7523/j.issn.2095-6134.2020.04.017>
39. X. Guo, S. Li, J. Yu, J. Zhang, J. Ma, L. Ma, et al., PFLD: A practical facial landmark detector, preprint, arXiv:1902.10859.
40. P. Yan, D. Yan, C. Du, Design and implementation of a driver's eye state recognition algorithm based on PERCLOS, *Chin. J. Electron.*, **4** (2014), 669–672.
41. V. F. Ferrario, C. Sforza, G. Serrao, G. Grassi, E. Mossi, Active range of motion of the head and cervical spine: a three-dimensional investigation in healthy young adults, *J. Orthop. Res.*, **20** (2002), 122–129. [https://doi.org/10.1016/S0736-0266\(01\)00079-1](https://doi.org/10.1016/S0736-0266(01)00079-1)
42. H. Wang, Z. Li, X. Ji, Y. Wang, Face R-CNN, preprint, arXiv:1706.01061.
43. B. Yan, C. Yang, F. Chen, K. Takeda, C. Wang, FDNet: a deep learning approach with two parallel cross encoding pathways for precipitation nowcasting, *J. Comput. Sci. Technol.*, **1** (2021). <https://jcst.ict.ac.cn/EN/10.1007/s11390-021-1103-8>
44. J. Li, Y. Wang, C. Wang, Y. Tai, J. Qian, J. Yang, et al., DSFD: dual shot face detector, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Long Beach, USA, (2019), 5060–5069. <https://doi.org/10.1109/CVPR.2019.00520>

45. J. Deng, J. Guo, E. Ververas, I. Kotsia, S. Zafeiriou, RetinaFace: single-shot multi-level face localisation in the wild, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Seattle, USA, (2020), 5203–5212. <https://doi.org/10.1109/cvpr42600.2020.00525>
46. W. Tian, Z. Wang, H. Shen, W. Deng, Y. Meng, B. Chen, et al., Learning better features for face detection with feature fusion and segmentation supervision, preprint, arXiv:1811.08557.
47. S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, S. Z. Li, S³fd: single shot scale-invariant face detector, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, IEEE, Venice, Italy, (2017), 192–201. <https://doi.org/10.1109/iccv.2017.30>
48. J. Wang, Y. Yuan, G. Yu, Face attention network: an effective face detector for the occluded faces, preprint, arXiv:1711.07246.
49. H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Lin, Z. Zhang, et al., ResNeSt: Split-attention networks, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, New Orleans, USA, (2022), 2736–2746. <https://doi.org/10.1109/CVPRW56347.2022.00309>
50. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L. Chen, Mobilenetv2: Inverted residuals and linear bottlenecks, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Salt Lake, USA, (2018), 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
51. K. Sun, Y. Zhao, B. Jiang, T. Cheng, B. Xiao, D. Liu, et al., High-resolution representations for labeling pixels and regions, preprint, arXiv:1904.04514.
52. M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, in *International Conference on Machine Learning (ICML)*, PMLR, Long Beach, USA, (2019), 6105–6114.



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)