TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 13/15 pp749–758
DOI: 10.26599/TST.2021.9010025
Volume 26, Number 5, October 2021

A Computer-Aided System for Ocular Myasthenia Gravis Diagnosis

Guanjie Liu, Yan Wei, Yunshen Xie, Jianqiang Li, Liyan Qiao, and Ji-jiang Yang*

Abstract: The current mode of clinical aided diagnosis of Ocular Myasthenia Gravis (OMG) is time-consuming and laborious, and it lacks quantitative standards. An aided diagnostic system for OMG is proposed to solve this problem. The values calculated by the system include three clinical indicators: eyelid distance, sclera distance, and palpebra superior fatigability test time. For the first two indicators, the semantic segmentation method was used to extract the pathological features of the patient's eye image and a semantic segmentation model was constructed. The patient eye image was divided into three regions: iris, sclera, and background. The indicators were calculated based on the position of the pixels in the segmentation mask. For the last indicator, a calculation method based on the Eyelid Aspect Ratio (EAR) is proposed; this method can better reflect the change of eyelid distance over time. The system was evaluated based on the collected patient data. The results show that the segmentation model achieves a mean Intersection-Over-Union (mIoU) value of 86.05%. The paired-sample T-test was used to compare the results obtained by the system and doctors, and the p values were all greater than 0.05. Thus, the system can reduce the cost of clinical diagnosis and has high application value.

Key words: ocular myasthenia gravis; computer-aided system; semantic segmentation; eyelid aspect ratio

1 Introduction

Myasthenia Gravis (MG) is an autoimmune disorder caused by autoantibodies acting against the nicotinic acetylcholine receptor on the postsynaptic membrane at the neuromuscular junction. It is characterized by weakness and fatigability of the voluntary muscles[1]. It is estimated that more than 700 000 people are affected worldwide, and the incidence ranges from 0.3 to 2.8 per 100 000[2]. According to Osserman et al.[3], adult patients with myasthenia can be divided into five different classifications: Ocular Myasthenia Gravis (OMG) is the mildest of them (i.e., only eye muscle weakness but no muscle weakness elsewhere). Its clinical manifestations include dropping of the eyelids or double vision, or both. In detail, double vision can occur in one eye or both eyes and worsen after activity. The symptoms can be relieved after resting or using anticholinergic drugs. Without immunotherapy, 30%−80% of OMG patients will progress to generalized MG (GMG) within two years[4].

Some methods are used to aid diagnosis of MG, such as the ice test, repetitive nerve stimulation, single fiber electromyography, and the acetylcholine receptor antibody test. The acetylcholine receptor antibody test is the most specific diagnostic test for MG[5, 6]. The neostigmine test is a kind of acetylcholine receptor antibody test, demonstrating a clinical improvement in patients affected by MG[7]. Eyelid distance, scleral distance, and palpebra superior fatigability test time are three important metrics in the neostigmine test, as

• Guanjie Liu, Yunshen Xie, and Jianqiang Li are with the Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China. E-mail: {ellzio0817, xieyunshen_2018}@163.com; lijianqiang@bjut.edu.cn.
• Yan Wei and Liyan Qiao are with the Neurology Department, The Second Affiliated Hospital of Tsinghua University, Beijing 100040, China. E-mail: weiyan9203@163.com; qiaoliyan@tsinghua.edu.cn.
• Ji-jiang Yang is with the Department of Automation, Tsinghua University, Beijing 100084, China. E-mail: yangjijiang@tsinghua.edu.cn.
* To whom correspondence should be addressed.
Manuscript received: 2021-02-27; accepted: 2021-03-15

illustrated in Figs. 1–3. Eyelid distance is the distance between upper and lower eyelids when the patient is in front view and maximum eyelid view; scleral distance is the length, in millimeters, of the displayed red line when the patient is gazing to the left or right; in the palpebra superior fatigability test, the cornea is regarded as a clock face, and the patient continues to open his eyes and look upward for 1 min. The time from drooping eyelids to the 9–3 point in the cornea is recorded. The left and right eyes are scored separately. The value of these indicators are helpful in judging the condition change of the patients before and after using neostigmine.

Doctors often employ methods of direct visual inspection of ptosis recovery or manual measurement in previous neostigmine tests. The results obtained may have certain subjective errors and lack quantitative standards. For doctors, the neostigmine test has many steps; some are easy to forget in the actual measurement.

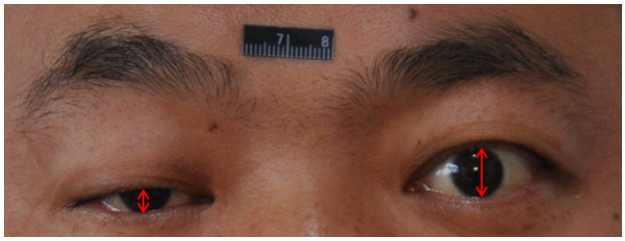
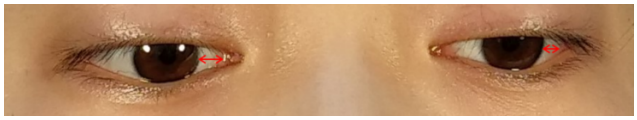


Fig. 1 Eyelid distance. Distance between upper and lower eyelids.



(a) Left view

(b) Right view

Fig. 2 Scleral distance. Distance between the edge of the iris and the corner of the eye when the eyes are moving horizontally.

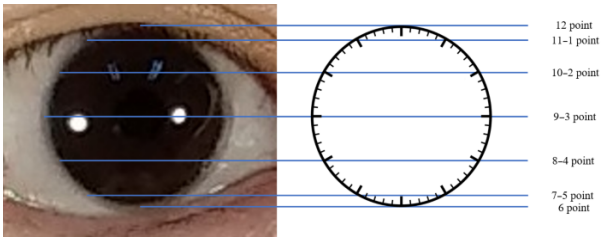


Fig. 3 Clock point.

Meanwhile, for patients, OMG is a rare disease and requires lifelong monitoring. Many hospitals lack doctors who can diagnose and treat OMG.

With the development of Computer-Aided Diagnosis (CAD), many doctors use CAD to help improve their diagnostic decisions. CAD has a wide range of applications in breast cancer diagnosis, lung nodule detection in chest images, and Magnetic Resonance Imaging (MRI) images[8, 9]. However, to the best of our knowledge, there is little work incorporating deep learning methods in OMG diagnosis.

The remainder of the paper is structured as follows. Related work is discussed in Section 2. Section 3 introduces the details of our system design for OMG-aided diagnosis. Section 4 analyzes experimental results. Finally, conclusions are drawn in Section 5.

2 Related Work

Certain diseases can cause changes or deformities in the patient's face. Screening and diagnosis can be aided by the patient's facial image. Traditional facial-aided diagnosis mainly comprises five parts: image acquisition, face detection, image preprocessing, feature selection, and classifier. There will be some differences if using neural networks. Image acquisition is usually carried out in a unified shooting environment set by medical institutions. After this part is completed, the specific area in the face image is cropped out by face detection. Based on the needs of different methods used, it is necessary to preprocess the obtained images, e.g., gray-scale normalization and data augmentation. To convert facial features caused by diseases into analyzable data, feature extraction and selection is essential. Various facial feature of face images have been extracted, and good feature representation helps improve the effectiveness of the algorithm. In the last part, the classifier analyzes the extracted features and obtains the diagnostic result.

The method of aided diagnosis of face morphology based on the face image of the patient has been applied to many diseases. In 2013, Kosilek et al.[10] used a digital camera to collect the front and side photos of 20 patients with Cushing syndrome and 40 healthy people. After preprocessing, they imported the software facial image diagnostic aid. The software correctly classified 85.0% of the patients and 95.0% of the control group, and the overall classification accuracy was 91.7%. In 2015, Chen et al.[11] divided the face into several regions based on the clinical observation of the facial appearance of patients with chronic fatigue

syndrome, and constructed mixed appearance features based on the image features extracted from each region. The feature extraction methods used included Gabor wavelet filtering, threshold segmentation, and curvature fitting. Three methods of Principal Component Analysis (PCA), locality-preservation projection, and Manifold Preservation Projection (MPP) were used to reduce Gabor features. Finally, multi-modal facial feature fusion was performed based on the two-stage Adaptive Boosting (Adaboost) method. Among them, the MPP method showed the best result, and the accuracy of the test set was 88.32%. In 2018, Song et al.[12] extracted five features of subjects with Turner syndrome based on a trained 68 facial feature point model: forehead, melanocytic nevus, epicanthus, nasal bridge, and ocular distance, divided into three categories: global geometric features, global texture features, and local features. For global features, PCA was used to reduce dimensionality to fuse multiple features, and Support Vector Machine (SVM) was used for classification. For local features, Adaboost was used for classification after fusion, and ten-fold cross-validation was used. The best effect was the use of the Adaboost classification method after local feature fusion. The sensitivity and specificity were 75.6% and 91.2%, respectively. Children suffering from adenoid hypertrophy display adenoid facial features. In 2019, Hu et al.[13] used the Dlib face key point detection model to obtain the facial feature points of the child, using decision trees, SVM, K-Nearest Neighbor (KNN), and Extreme Gradient Boosting (XGBoost) to establish adenoid face and non-adenoid face prediction models. The reliability of prediction was evaluated by five-fold cross-validation. The best specificity was 89.19%, and the best sensitivity was 88.24%.

At present, many researchers focus on the use of manual extraction of predefined features to construct the corresponding facial disease automatic classifier, but this is a laborious heuristic method, and the predefined features may non-representative, incomplete, or redundant, or may even increase interference features. The deep learning method integrates feature learning into the process of establishing the model, thereby reducing the incompleteness caused by human design features.

As a powerful big data technology, deep learning has a wide range of applications in clinical research due to its huge advantages in feature and pattern recognition[14]. In 2017, Shukla et al.[15] proposed a new framework for detecting developmental disorders using facial images. This framework compares the Deep Convolutional Neural Network (DCNN) method with a similar method that uses predefined features with classifiers. The DCNN method achieved a maximum accuracy of 98.80%.

Therefore, in this paper, we propose an eye segmentation network to extract pathological features related to OMG. Based on the segmented mask, we calculate the values of the two evaluation indicators of the neostigmine test, and convert them to real values through distance conversion. For the time measurement of the palpebra superior fatigability test, this paper proposes a calculation method based on the key points of the human eye. Based on the above method, the OMG-aided diagnosis system is constructed to quantitatively analyze the condition of OMG patients, providing a novel idea for solving the aided diagnosis problem of OMG.

3 Model Structure and System Design

3.1 Dataset and preprocessing method

3.1.1 Collection standard

According to the items of the neostigmine test, we collect data from patients before and after the test (i.e., before injection, 30 min after injection, and 60 min after injection). The background is uniformly blue, and the illumination is controlled to ensure that the image is not too bright or too dark. To convert the measured pixel distance into the real distance, we uniformly pasted a 2 cm black bar on the forehead of patients for distance conversion. NIKON D300S and Xiaomi MI MAX2 were used for image acquisition. When shooting, the equipment was kept parallel to the patient's face, the position was fixed, and the distance was 1 m. For each of the above three time periods, four images were collected, namely, front view, maximum eyelid view, right view, and left view. The subjects in this experiment were all OMG patients. Based on the above collection criteria, a total of 550 facial images and 60 videos of OMG patients were collected. All photos were taken with the consent of patients.

3.1.2 Image preprocessing

Inspired by Ref. [13], Dlib was also used in our paper to get eye and black bar patches. Based on the key points around the eyes, we extended a part of the pixels in each of the four directions to ensure the integrity of the eye in the cropped image. The black bar area was also cropped in terms of the key points of the eyebrows. After cropping the facial images of OMG patients, a total of 1100 cross-resolution eye images and 550 black

bar images were obtained. Also, to increase the number of eye images, we used data augmentation to expand the number of images to 4400 through image flipping, random cropping, and the addition of Gaussian noise. The Image Labeler of MATLAB was used to label eye images. Then, a Myasthenia Eye Patches (MEP) database was established.

3.2 Eye segmentation model

3.2.1 Network structure

Encoder-decoder networks such as Segnet[16] and U-Net[17] have remarkable performance in image segmentation. Owing to this, we designed an encoder-decoder network for OMG patient eye image segmentation, which we called OMG-Net. The backbone was MobileNet[18], and the use of depth separable convolution considerably decreased the number of parameters and calculations, which was more conducive to running on low-configuration machines. The decoder part consisted of upsampling layers and a Softmax layer, which were used to restore the image size and make predictions of the category of each pixel. Our network is depicted in Fig. 4.

OMG-Net is divided into the following parts:

(1) Input layer. The read image is an $m \times n \times k$ matrix, where m and n are the length and width of the input image, and k is the number of channels of the matrix. To process input images of different sizes, the image size is uniformly scaled to 416×416 here.

(2) Depthwise separable convolution. This is made up of two parts: depthwise convolutions and pointwise convolutions. One convolution kernel of depthwise convolutions is responsible for one channel, and one channel is convolved by only one convolution kernel. If the depthwise convolutional kernel of size is $D_K \times D_K$ and M is the number of channels of input feature map. The computational cost of depthwise convolution is $D_K \times D_K \times M$. The operation of pointwise convolutions is similar to that of standard convolution. The difference is that the size of the convolution kernel is $1 \times 1 \times M$. We assume that the channels of output feature map is N, then the depthwise separable convolutions cost is $D_K \times D_K \times M + M \times N$. For standard convolutions, the input feature map size is $D_K \times D_K \times M$, and the output feature map size is $D_F \times D_F \times N$, where D_K and D_F represent the spatial width and height of the square input and output feature maps, respectively. The computational cost of standard convolution is $D_K \times D_K \times M \times N$, then the computational cost ratio of the depthwise separable convolution and the standard convolution is

$$\frac{D_K \times D_K \times M + M \times N}{D_K \times D_K \times M \times N} = \frac{1}{N} + \frac{1}{D_K^2} \quad (1)$$

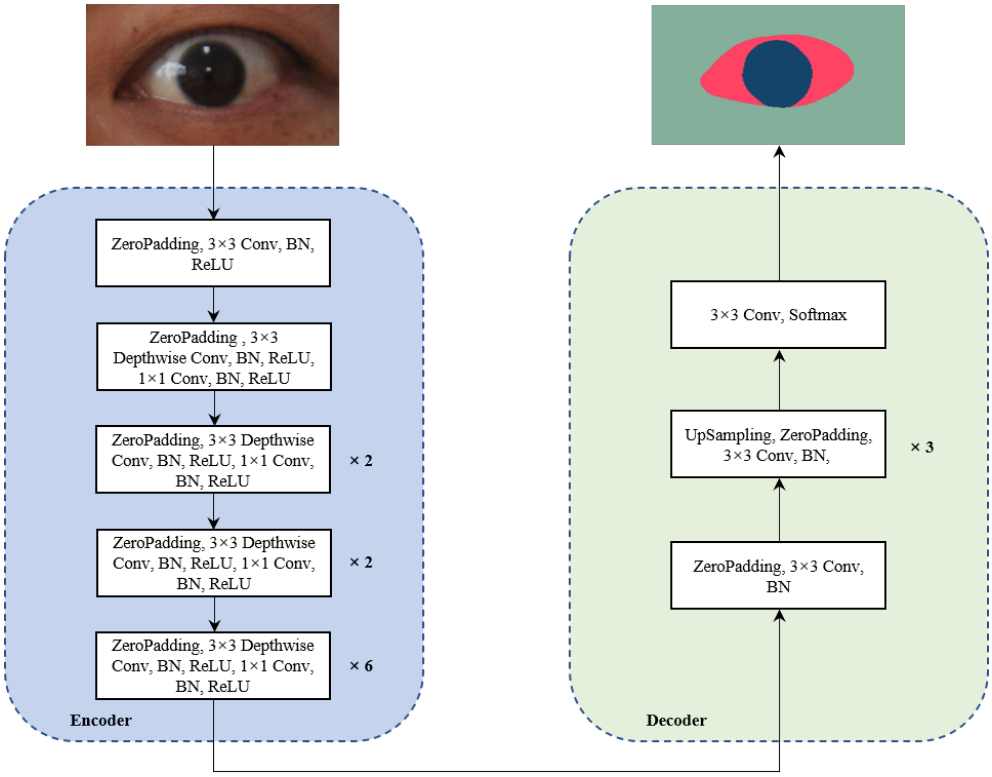


Fig. 4 Overview of proposed OMG-Net.

MobileNet uses 3×3 depthwise separable convolutions, which uses 8 to 9 times less computation than standard convolutions.

(3) Upsampling. Common methods of upsampling include bilinear interpolation, transposed convolution[19], and unpooling[20]. In this paper, the bilinear interpolation method is used. The core idea of the bilinear interpolation method is to perform linear interpolation in two directions separately, without learning any parameters. The calculation speed is fast and the operation is simple.

(4) Output layer. We calculate the probability that each pixel belongs to each semantic category through the Softmax classifier, and select the category corresponding to the maximum probability as the final result output. The form of the Softmax function is usually as follows:

$$p_i = \frac{\exp(s_i)}{\sum_{j=1}^{k} \exp(s_j)} \tag{2}$$

where s_i and p_i indicate the output of the last hidden layer and the discriminative probability of category i, respectively. j is the category, and $j = 1, 2, \ldots, k$.

3.2.2 Distance conversion

After obtaining the segmentation result of the eye image, we could calculate the pixel distance of the two items in the neostigmine test in light of the pixel position of the segmented mask. In order to know whether our calculation result was accurate, it needed to be contrasted with the result manually measured by the doctor, so the calculated pixel distance needs to be converted into the actual distance.

We propose an algorithm to solve this problem (Algorithm 1), i.e., the pixel side length of the black bar is obtained by finding the smallest circumscribed rectangle of the black bar on the patient's forehead in the image. The actual side length is two centimeters, so the conversion ratio of pixel distance and actual distance can be obtained.

Algorithm 1 Calculating the ratio of pixel distance and real distance

Input: The image with black bars, I;
Output: Conversion ratio, R;
/* C is the set of the sum of $c1$ and $c2$ */
/* TH is the threshold of contour */
/* RCM is the set of the ratio of ca and mbra */

$Gray \Leftarrow I$
$HSV \Leftarrow I$
$H \Leftarrow HSV$
$c1 \Leftarrow$ contours from $Gray$
$c2 \Leftarrow$ contours from H
$C \Leftarrow c1 + c2$
for *each* $c \in C$ **do**
  if $c < TH$ **then**
    remove c from C
for *each* $c \in C$ **do**
  $ca \Leftarrow$ contour area
  $mbra \Leftarrow$ minimum bounding rectangle area
  $RCM \Leftarrow ca/mbra$
for *each* $a \in RCM$ **do**
  if a *is maximum in* RCM **then**
    long side of rectangle $\Leftarrow$ coordinate of mbra
$R \Leftarrow$ long side of rectangle / 2 cm
return R

3.3 OMG-aided diagnostic system

With the massive amount of data generated daily and the development of deep learning, the framework and system will reduce the cost of manpower[21, 22]. To assist doctors in disease diagnosis, we developed an OMG-aided diagnosis system based on the above methods. First, the system processes the images of the patient front view, maximum eyelid view, right view, left view, and the video of the palpebra superior fatigability test in the neostigmine test. Then, the size of the eyelid distance, the scleral distance, and the palpebra superior fatigability test time can be obtained based on image processing and deep learning methods. The overall framework is shown in Fig. 5.

There are too many redundant features in the face image. In order to eliminate these redundant features and further refine the critical areas of the human eye, we used the face key point detection model in the Dlib library to obtain the key points of the human eye. According to the indicators that the system needs to calculate, the OMG-aided diagnosis system can be divided into two parts: static image analysis and dynamic video analysis.

In the static image analysis, the system first receives the facial image uploaded by the user and extracts the left and right eye regions of the patient based on the Dlib face key point detection model. The left and right eye images are sent to the pre-trained OMG-Net model to obtain the segmentation results of the eye region, including the iris, sclera, and background regions. According to the results of segmentation, the following two parts can be completed: measurement of the size of the eyelid distance and scleral distance. The iris and sclera constitute the main area of the palpebral fissure, and the maximum vertical distance of this area is the eyelid distance. Similarly, for the scleral area, the maximum distance from the horizontal corner of the eye to the iris

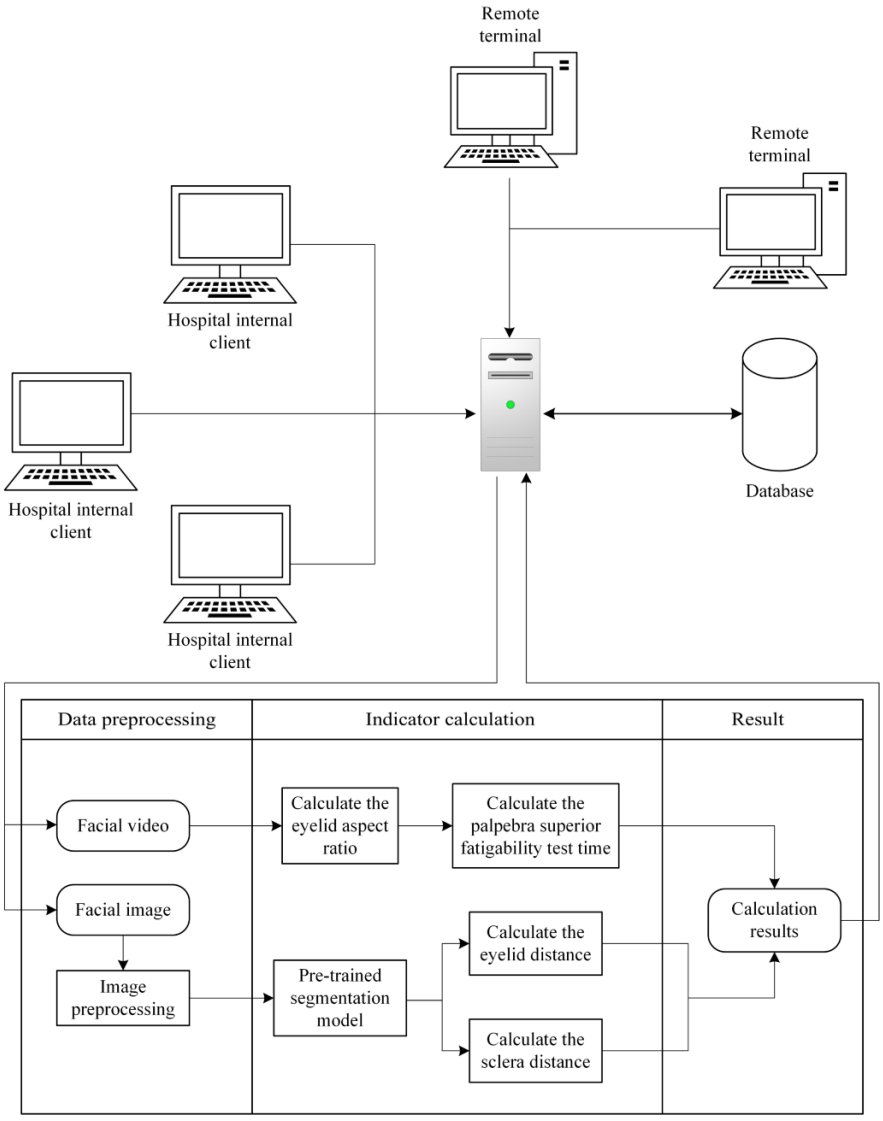


Fig. 5 OMG-aided diagnosis system framework.

in this area is the scleral distance.

For dynamic video analysis, inspired by Soukupová and Čech[23], we added two points on the basis of the original six key points of the eye, which are the position where the maximum height of the blepharoplasty intersects the upper and lower eyelids, as shown in Fig. 6, where Fig. 6a is the normal eye and Fig. 6b is the patient eye. The key points of the upper and lower eyelids are fitted by the least square method to acquire two parabolas, so that points P3 and P7 are obtained. Based on these eight key points around the eye, we propose the concept of Eyelid Aspect Ratio (EAR), which is the ratio of eye height to width. The specific calculation method is

$$\mathrm{EAR} = \frac{\| p_2 - p_8 \| + \| p_3 - p_7 \| + \| p_4 - p_6 \|}{3 \| p_1 - p_5 \|} \tag{3}$$

We calculated the EAR value for each frame of the video. If the EAR value was less than or equal to the set threshold, the time from the beginning of the video to this frame was returned, so that the palpebra superior fatigability test time could be obtained.

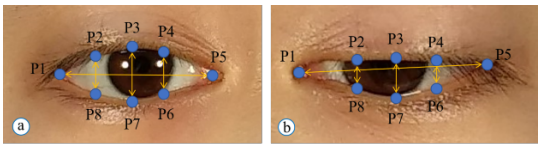


Fig. 6 Key points around the eyes.

4 Experiment and Discussion

4.1 Implementation details

The running environment of the experiment was

Windows10, memory 16 GB, Nvidia GeForce GTX 1060 graphics card. The ratio of train set to test set was 4:1. The proposed network structure was implemented using the Keras framework, the initial learning rate was 1×10^{-4}, ten-fold cross-validation was used, the epoch was set to 60, the optimization algorithm was Adam, and the loss function used was the cross-entropy function. During the training process, if the loss of the verification set did not decrease three consecutive times, the learning rate was decreased to continue training. If the loss did not decrease ten consecutive times, the model was considered to be trained and the training was stopped.

4.2 Performance verification of OMG-aided diagnostic system

4.2.1 Evaluation indicators

Common evaluation indicators for semantic segmentation included pixel accuracy, mean pixel accuracy, and mean Intersection-Over-Union (mIoU). To evaluate the results of the experiment more accurately, we used mIoU as the evaluation index. For the segmented mask, we calculated the mIoU values of the iris, sclera, and whole image, respectively named I-mIoU, S-mIoU, and Mean mIoU.

To verify the accuracy of the calculated results of the OMG-aided diagnosis system, this study used SPSS24.0 statistical software and a paired-sample T-test to contrast the differences between the system and the manual measurement results of doctors.

4.2.2 Comparison with other segmentation approaches

We compared OMG-Net with FCN[24], U-Net[17], and PSPNet[25]. All methods were re-trained on the same training set (i.e., MEP dataset) during this experiment. The segmentation methods are trained and tested in the same setting as OMG-Net. Visualization of eye segmentation results of OMG patients is shown in Fig. 7. Table 1 illustrates the mIoU scores and inference time. It can be clearly seen that OMG-Net performed better than FCN and U-Net. The mIoU values of OMG-Net and PSPNet were very close, but OMG-Net was far better than PSPNet in the inference time.

Through lateral contrast observation, it is not difficult to find that, for the segmentation results of the two regions of iris and sclera, the segmentation of iris is better than that of the sclera. Combining with Fig. 7, we can understand that segmentation errors are primarily localized in the corner of the eye. For OMG patients, when their eyes are in left or right view, the iris part cannot reach the corner of the eye, and a part of the sclera will be exposed. This area is sometimes small, which easily divides the area into two other categories.

4.2.3 Cross-resolution comparison

The iBUG Eye Segmentation Dataset[26] is used to verify the performance of OMG-Net on human eye images of different resolutions. It contains 8882 annotated human eye images, all of which are low-resolution images in the field, with an image size of 160 × 80 (Table 2).

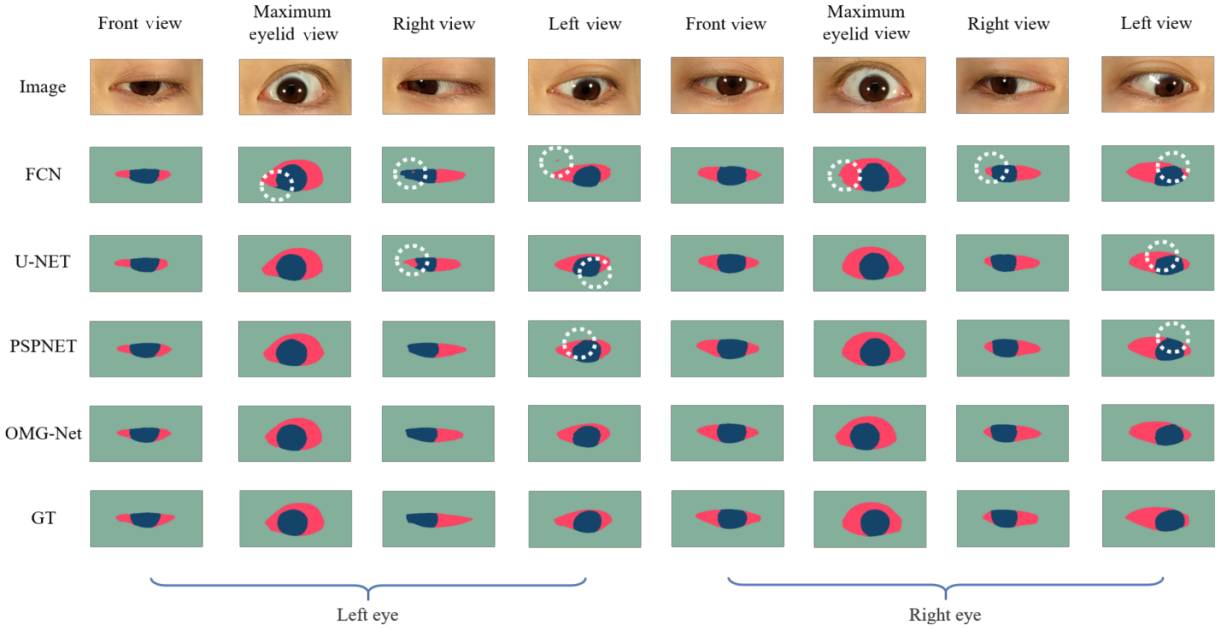


Fig. 7 Eye images segmentation results of patients with different methods.

Table 1 mIoU and average inference speed achieved by OMG-Net and other segmentation methods.

Method	S-mIoU	I-mIoU	Mean mIoU	Inference time (s)
FCN	76.28%	81.56%	78.92%	0.042
U-Net	77.04%	83.28%	80.16%	0.067
PSPNet	**82.49%**	89.23%	85.86%	0.113
OMG-Net	82.37%	**89.73%**	**86.05%**	**0.031**

Table 2 Statistical information of iBUG Eye Segmentation Dataset.

Name	Value
Total number of faces	4461
Total number of eye patches	8882
Non-frontal faces proportion	18.35%
Low-resolution eye patches proportion	57.58%
Proportion of images with any kind of occlusion	16.05%

In this part, we used OMG-Net to conduct experiments on three different datasets. The three datasets were the MEP dataset, the iBUG eye segmentation dataset, and the combined dataset (Table 3). First, among the segmentation results of OMG-Net on different datasets, the best performance was the segmentation result based on the MEP dataset, and the Mean mIoU was 86.05%. The second was MEP combined with iBUG, and the last one was iBUG. At the same time, the increase in the number of eye images did not increase mIoU. The results showed that OMG-Net has better segmentation results on high-resolution images than low-resolution images, and is more suitable for the task of segmenting OMG patient eye images.

Table 3 OMG-Net segmentation results on different datasets.

(%)

Dataset	S-mIoU	I-mIoU	Mean mIoU
MEP	**82.37**	**89.73**	**86.05**
iBUG	64.85	82.00	73.43
MEP+iBUG	70.21	84.81	77.51

4.2.4 Calculation results comparison

In this section, we explore the closeness between the calculated results based on the OMG-aided diagnosis system and the results manually measured by the doctor. We used the data of nine subjects before and after the neostigmine test to calculate the evaluation indicators. The difference between the two was tested by a paired-sample T-test, and a total of 10 paired samples were formed. The test results are shown in Table 4.

It can be seen from Table 4 that the p values of the ten paired samples are all greater than 0.05, and the difference is not statistically significant. The calculated results of the OMG-aided diagnosis system were consistent with the manual measurement results of the doctor. To show the calculated results of the palpebra superior fatigability test more intuitively, we calculated the EAR value for each frame of the 1 min video and extracted the smallest EAR value within 1 s to draw a line chart, as shown in Fig. 8.

In Fig. 8, the EAR value of normal people fluctuates up and down in a fixed range within 1 min and is always above the threshold. For OMG patients, with increasing time, their upper eyelids cannot be supported and gradually begin to droop; the EAR value of their eyes gradually becomes smaller and lower than the threshold. By calculating the EAR value, we can get not only the time of the palpebra superior fatigability test, but also the change in EAR value, which can reflect the change of patient palpebral fissure in 1 min, thus providing an effective reference for doctors to assist in the diagnosis of patient condition.

5 Conclusion and Future Work

The contributions of this paper are mainly reflected in

Table 4 Paired-sample T-test results.

No.	Paired difference					t-value	Degree of freedom	Significant 2-tailed
	Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
				Lower	Upper			
Pair 1	0.22852	0.61297	0.11797	–0.01396	0.47100	1.937	26	0.064
Pair 2	0.24519	0.70866	0.13638	–0.03515	0.52552	1.798	26	0.084
Pair 3	–0.09741	0.56044	0.10786	–0.31911	0.12429	–0.903	26	0.375
Pair 4	0.01667	0.70925	0.13649	–0.26390	0.29724	0.122	26	0.904
Pair 5	0.24889	0.79800	0.15358	–0.06679	0.56457	1.621	26	0.117
Pair 6	–0.41926	1.21189	0.23323	–0.89867	0.06015	–1.798	26	0.084
Pair 7	0.19630	0.56991	0.10968	–0.02915	0.42175	1.790	26	0.085
Pair 8	–0.28741	1.57595	0.30329	–0.91083	0.33602	–0.948	26	0.352
Pair 9	3.852	15.745	3.030	–2.377	10.080	1.271	26	0.215
Pair 10	3.926	22.812	4.390	–5.098	12.950	0.894	26	0.379

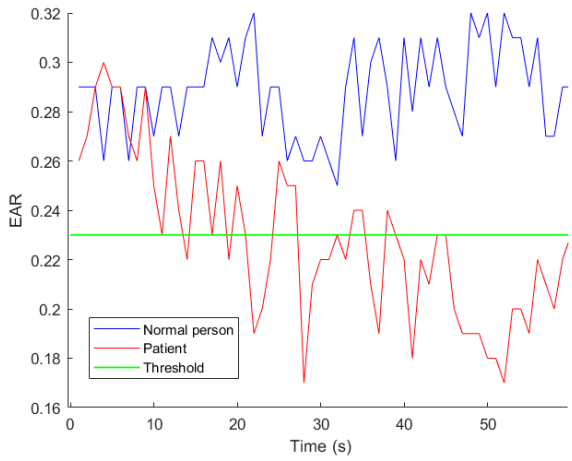


Fig. 8 EAR line chart.

the following three parts: (1) Based on the semantic segmentation method, OMG-Net for segmenting OMG patient eye images is proposed, and the effectiveness of OMG-Net on the collected patient eye image data set is verified. The mean mIoU result was 86.05%; (2) A method for calculating the palpebra superior fatigability test time based on the EAR value is proposed. This can directly reflect the changes of the patient eyelid distance within 1 min; (3) For that portion of OMG-aided diagnosis that has not yet been involved in relevant research, an OMG-aided diagnosis system for measuring three indicators of eyelid distance, scleral distance, and palpebra superior fatigability test time is proposed.

This paper only designs an aided diagnosis system for the three clinically-aided methods of OMG; more methods can be incorporated into the system in the future. The semantic segmentation model in this study needs to improve the segmentation results of small target areas such as the corners of the eyes, also providing ideas for continued improvement of the system in the next step.

References

[1] B. R. Thanvi and T. C. Lo, Update on myasthenia gravis, *Postgraduate Medical Journal*, vol. 80, no. 950, pp. 690–700, 2004.

[2] J. C. W. Deenen, C. G. C. Horlings, J. J. G. M. Verschuuren, A. L. M. Verbeek, and B. G. M. van Engelen, The epidemiology of neuromuscular disorders: A comprehensive overview of the literature, *Journal of Neuromuscular Diseases*, vol. 2, no. 1, pp. 73–85, 2015.

[3] K. E. Osserman, P. Kornfeld, E. Cohen, G. Genkins, H. Mendelow, H. Goldberg, H. Windsley, and L. I. Kaplan, Studies in myasthenia gravis: Review of two hundred eighty-two cases at the Mount Sinai Hospital, New York City, *A.M.A. Archives of Internal Medicine*, vol. 102, no. 1, pp. 72–81, 1958.

[4] S. H. Wong, A. Petrie, and G. T. Plant, Ocular myasthenia gravis: Toward a risk of generalization score and sample size calculation for a randomized controlled trial of disease modification, *Journal of Neuro-ophthalmology*, vol. 36, no. 3, pp. 252–258, 2016.

[5] M. Benatar, A systematic review of diagnostic studies in myasthenia gravis, *Neuromuscular Disorders*, vol. 16, no. 7, pp. 459–467, 2006.

[6] K. Scherer, R. S. Bedlack, and D. L. Simel, Does this patient have myasthenia gravis? *JAMA*, vol. 293, no. 15, pp. 1906–1914, 2005.

[7] G. Sciacca, E. Reggio, G. Mostile, A. Nicoletti, F. Drago, S. Salomone, and M. Zappia, Clinical and CN-SFEMG evaluation of neostigmine test in myasthenia gravis, *Neurological Sciences*, vol. 39, no. 2, pp. 341–345, 2018.

[8] K. Doi, Computer-aided diagnosis in medical imaging: Historical review, current status and future potential, *Computerized Medical Imaging and Graphics*, vol. 31, nos. 4&5, pp. 198–211, 2007.

[9] J. Liu, Y. Pan, M. Li, Z. Y. Chen, L. Tang, C. Q. Lu, and J. X. Wang, Applications of deep learning to MRI images: A survey, *Big Data Mining and Analytics*, vol. 1, no. 1, pp. 1–18, 2018.

[10] R. P. Kosilek, J. Schopohl, M. Grunke, M. Reincke, C. Dimopoulou, G. K. Stalla, R. P. Würtz, A. Lammert, M. Günther, and H. J. Schneider, Automatic face classification of Cushing's syndrome in women–A novel screening approach, *Experimental and Clinical Endocrinology & Diabetes*, vol. 121, no. 9, pp. 561–564, 2013.

[11] Y. H. Chen, W. J. Liu, L. Zhang, M. Y. Yan, and Y. J. Zeng, Hybrid facial image feature extraction and recognition for non-invasive chronic fatigue syndrome diagnosis, *Computers in Biology and Medicine*, vol. 64, pp. 30–39, 2015.

[12] W. A. Song, Y. Lei, S. Chen, Z. X. Pan, J. J. Yang, H. Pan, X. L. Du, W. B. Cai, and Q. Wang, Multiple facial image features-based recognition for the automatic diagnosis of turner syndrome, *Comput. Ind.*, vol. 100, pp. 85–95, 2018.

[13] X. Hu, Q. Y. Zhang, J. J. Yang, Q. Wang, Y. Lei, and J. L. Wu, Photographic analysis and machine learning for diagnostic prediction of adenoid hypertrophy, in *Proc. 2019 IEEE 16th Int. Conf. Networking, Sensing and Control*, Banff, Canada, 2019, pp. 7–11.

[14] Y. Yu, M. Li, L. L. Liu, Y. H. Li, and J. X. Wang, Clinical big data and deep learning: Applications, challenges, and future outlooks, *Big Data Mining and Analytics*, vol. 2, no. 4, pp. 288–305, 2019.

[15] P. Shukla, T. Gupta, A. Saini, P. Singh, and R. Balasubramanian, A deep learning frame-work for recognizing developmental disorders, in *Proc. 2017 IEEE Winter Conf. Applications of Computer Vision*, Santa Rosa, CA, USA, 2017, pp. 705–714.

[16] V. Badrinarayanan, A. Kendall, and R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[17] O. Ronneberger, P. Fischer, and T. Brox, U-Net:

Convolutional networks for biomedical image segmentation, presented at the Int. Conf. Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015, pp. 234–241.

[18] A. G. Howard, M. L. Zhu, B. Chen, D. Kalenichenko, W. J. Wang, T. Weyand, M. Andreetto, and H. Adam, MobileNets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv: 1704.04861, 2017.

[19] V. Dumoulin and F. Visin, A guide to convolution arithmetic for deep learning, arXiv preprint arXiv: 1603.07285, 2016.

[20] H. Noh, S. Hong, and B. Han, Learning deconvolution network for semantic segmentation, in *Proc. 2015 IEEE Int. Conf. Computer Vision*, Santiago, Chile, 2015, pp. 1520–1528.

[21] W. Zhong, N. Yu, and C. Y. Ai, Applying big data based deep learning system to intrusion detection, *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 181–195, 2020.

[22] W. W. Jiang and L. Zhang, Geospatial data to images: A deep-learning framework for traffic forecasting, *Tsinghua Science and Technology*, vol. 24, no. 1, pp. 52–64, 2019.

[23] T. Soukupová and J. Čech, Real-time eye blink detection using facial landmarks, in *Proc. 21st Computer Vision Winter Workshop*, Rimske Toplice, Slovenia, 2016, pp. 1–8.

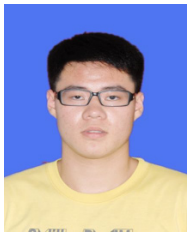
[24] J. Long, E. Shelhamer, and T. Darrell, Fully convolutional networks for semantic segmentation, in *Proc. 2015 IEEE Conf. Computer Vision and Pattern Recognition*, Boston, MA, USA, 2015, pp. 3431–3440.

[25] H. S. Zhao, J. P. Shi, X. J. Qi, X. G. Wang, and J. Y. Jia, Pyramid scene parsing network, in *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 2017, pp. 6230–6239.

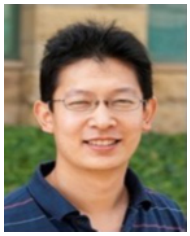
[26] B. Luo, J. Shen, Y. Wang, and M. Pantic, The ibug eye segmentation dataset, presented at the 2018 Imperial College Computing Student Workshop, London, UK, 2019.

Guanjie Liu is currently a postgraduate student in the Faculty of Information Technology, Beijing University of Technology, China. His current research interests include medical image processing and deep learning.

Yan Wei received the master degree in neurology from Peking University, Beijing, China in 2018. She has been working in the Second Affiliated Hospital of Tsinghua University since graduation. Her specialties include neuroimmunity (myasthenia gravis, neuromyelitis optica spectrum disease and multiple sclerosis), Alzheimer's disease, and cerebrovascular disease.



Yunshen Xie received the MS degree at the Faculty of Information Technology, Beijing University of Technology, China in 2020. His research interests include medical image processing and machine learning.



Jianqiang Li received the BS degree in mechatronics from Beijing Institute of Technology, Beijing, China in 1996, and the MS and PhD degrees in control science and engineering from Tsinghua University, Beijing, China in 2001 and 2004, respectively. He joined Beijing University of Technology, Beijing, China, in 2013 as a Beijing distinguished professor. His research interests are in Petri nets, enterprise information system, business process, data mining, information retrieval, semantic web, privacy protection, and big data. He has over 40 publications including 1 book, more than 10 journal papers, and 37 international patent applications (19 of them have been granted in China, US, or Japan). He served as PC members in multiple international conferences and organized the IEEE workshop on medical computing. He served as guest editor to organize a special issue on information technology for enhanced healthcare service in Computer in Industry.

Liyan Qiao received the PhD degree in neurology from Peking Union Medical College, Beijing, China in 2002, and training and studying in the Department of Neuropathology, the University of Alabama at Birmingham, Birmingham, USA in 2003. She has been working in the Second Affiliated Hospital of Tsinghua University since 2002. Her specialties include Alzheimer's disease, Parkinson's disease, neuroimmunity (multiple sclerosis and myasthenia gravis), blepharospasm, and dystonia. She has published more than 30 papers in domestic and foreign core journals, including 4 SCI papers. She presided over 2 projects of the Capital Medical Development Fund, the Ministry of Education's Research Startup Fund for Returned Overseas Chinese Scholars, and 1 sub-project of the 863 Project. She participated in a number of funds including the National Natural Science Foundation of China.

Ji-jiang Yang received the BS and MS degree from Tsinghua University (China) in 1990 and 1995, and the PhD in industrial engineering from National University of Ireland (Galway) in 2004. His interesting research areas involve in business process collaboration, information resource management, system integration and architecture, medical big data, intelligent assisted diagnosis, privacy protection, etc. Since 2009, he has focused mainly on digital medicine and health area, dedicating to using information and communication technologies to promote medical service capability and quality. He has undertaken/participated a few important projects from Ministry of Science & Technology on Digital Medical Service technologies. He is the member of IEEE, and the State Standard Committee of Medical Information (China).