ISSN 2616-5775 Vol. 6, Issue 9: 81-86, DOI: 10.25236/AJCIS.2023.060912

An Intelligent Facial Palsy Diagnostic System Based on Acupoint Identification

Keding Huang[1,a], Feiyu Zhu[1,b,*], Jicheng Zhang[2,c], Shun'e Tan[2,d], Ziyu Li[2,e]

[1]College of Computer Science, Sichuan University, Chengdu, China
[2]Sichuan Integrative Medicine Hospital, Chengdu, China
[a]keding75@126.com, [b]feiyuz@scu.edu.cn, [c]1808067047@qq.com, [d]514694194@qq.com, [e]449562892@qq.com
[*]Corresponding author

Abstract: *The methods for clinically diagnosing facial paralysis require doctors to possess a high degree of experience and specialized knowledge, often involving subjectivity. However, due to the uneven distribution of medical resources, many facial paralysis patients are unable to receive timely and accurate diagnosis and treatment. Traditional computer-assisted methods place high demands on hardware equipment and lack sufficient intelligence. With the continuous advancement of artificial intelligence, researchers have actively explored intelligent methods for facial paralysis detection. These methods mainly focus on extracting facial features and making judgments based on facial asymmetry, but they struggle to provide a scientific quantitative analysis of the severity of facial paralysis. This study is based on the lightweight network—MobileNetV2. By performing facial detection and processing on input images, it successfully identifies three groups of acupoints related to facial paralysis and conducts quantitative analysis based on this identification. Simultaneously, we have improved the network by constructing a two-stage network similar to object detection and regression, and optimizing the loss function. In the end, we compared the improved model with other mainstream frameworks through experiments. The results demonstrate that our proposed model achieves significant effectiveness in acupoint recognition and maintains low error in quantitative analysis.*

Keywords: *quantitative analysis, acupoint identification, facial paralysis, mobilenet*

1. Introduction

Facial paralysis, also known as Bell's palsy, is a common and prevalent facial nerve disorder. Patients often exhibit symptoms such as skewed mouth corners, difficulty raising eyebrows, facial twitching, and even speech impairment. It reported that incidence rate is approximately up to 0.053% [1]. This condition significantly affects the facial appearance, communication, and normal facial movements of patients, leading to substantial impacts on their physiological and psychological well-being.

Current methods for detecting facial paralysis mostly involve traditional facial observation or medical examination. Doctors guide patients to perform facial movements and then assess the degree of facial paralysis by observing facial limitations and using the House-Brackmann [2] grading scale.

In today's society, there exists a severe imbalance in different regions in terms of healthcare development. Many doctors in remote areas lack relevant experience and guidance, and patients often struggle with limited funds or transportation hindrances for treatment. Traditional medical methods heavily rely on doctors' experience and expertise, leading to subjective judgments and varying treatment approaches.

Therefore, the development of tools for objective facial paralysis diagnosis and assessment holds research significance and clinical value. Such tools could assist doctors in underdeveloped regions and aid patients in self-diagnosis.

In recent years, many scholars have turned their attention to using deep learning for intelligent facial paralysis assessment. For instance, Ngo [3] proposed a method based on Gabor and LBP features, which improved recognition rates but with longer response times. RIDHA [4] suggested a convenient application based on detecting key points and performing difference calculations.

Unlike mainstream facial feature research, this study introduces acupoints for assessment. These

ISSN 2616-5775 Vol. 6, Issue 9: 81-86, DOI: 10.25236/AJCIS.2023.060912

acupoints are derived from traditional Chinese medicine and are closely related to nerve endings and blood vessels. Extensive research indicates the effectiveness of acupoints in diagnosing and treating facial paralysis [5], with three sets of acupoints—ST4, EX-HN4, and EX-HN8—being particularly relevant to the facial nerve in clinical practice.

Considering the integration with mobile apps, this study builds upon the improved framework by using MobileNetV2[16]. It detects acupoints closely associated with facial paralysis and evaluates the asymmetry of acupoint groups. This innovative and lightweight approach offers valuable insights into facial paralysis diagnosis.

2. Related Work

2.1. Facial Feature Point Recognition and Acupoint Detection

In 1995, Cootes proposed the Active Shape Model (ASM) algorithm based on point distribution for facial keypoint detection [6]. While its advantage lies in clarity and simplicity of structure, its exhaustive search principle greatly hampers efficiency. With the development of deep learning, Sun et al. [7] first applied Convolutional Neural Networks (CNN) to keypoint detection, introducing the Deformable Convolutional Neural Network (DCNN) method. This method not only excels in robust feature extraction but also enhances the accuracy of keypoint detection. Subsequently, Face++ made improvements on this method. The Multi-task Cascaded Convolutional Networks (MTCNN) introduced in 2016 [8] achieved significant advancements in both speed and accuracy by simultaneously handling face detection and keypoint detection. In 2017, Kowalski et al. [9] proposed the Deep Alignment Network (DAN), which effectively addresses issues related to head pose and initialization.

Similarly, the recognition of traditional Chinese medicine acupoints can be seen as a specialized form of keypoint detection. However, the current methods for detecting acupoints often rely on external devices and suffer from low detection accuracy. Most detection methods are based on traditional Chinese acupuncture techniques, such as bone inch and cun measurements. Zhao et al. [10] proposed an infrared imaging-based method for acupoint detection, but this approach requires hardware support and lacks practicality, which contradicts the scope of our study. Chang [11] suggested a method that uses ASM for facial feature extraction and combines it with acupoint determination, but this approach is greatly influenced by facial feature extraction. Until now, limitations in recognition accuracy still exist.

In summary, facial feature point recognition has been a hot topic in recent years, while research into traditional Chinese medicine acupoint recognition remains relatively underexplored, offering substantial room for growth. Research in this direction heavily depends on acupoint-related datasets and requires the assistance of medical professionals, leading to challenges such as limited availability of relevant open-source datasets and how to enhance accuracy. Therefore, the direction of acupoint recognition applied to facial paralysis diagnosis and quantitative assessment holds both research and practical significance. We anticipate that our study will contribute positively to this field.

2.2. Facial Paralysis Detection

In the early stages, facial paralysis detection primarily relied on doctors observing whether patients had difficulty moving their facial muscles and diagnosing them based on the House-Brackmann grading scale. This assessment method was entirely subjective, heavily reliant on the diagnostic experience and professional expertise of the doctors. With the advancement of computers, research gradually introduced their assistance in diagnosis. NELLY et al. proposed a computer-assisted analysis method for grading facial movement functionality [14]. In recent years, Wang [15] proposed a novel approach by combining static and dynamic facial features to assess the severity of facial paralysis. Moreover, approaches like that of Wang Q and others [12] utilized Generative Adversarial Networks (GANs) for abnormal image detection to distinguish between facial paralysis and normal conditions. Xu and his team [13] introduced a method for facial paralysis grading assessment based on deep temporal features.

The method presented in this paper is grounded in the relationship between acupoints and facial nerves. It involves analyzing the relative positions and angles of the Di Cang acupoint, Shang Ying Xiang acupoint, and Yu Yao acupoint. This approach enables both facial paralysis detection and quantitative analysis. Among them, the Yu Yao acupoint (International Code: EX-HN4) is located just above the left pupil at the intersection of the eyebrow, where it is associated with the lateral branch of the supraorbital nerve, branches of the facial nerve, and lateral branches of the superior ophthalmic artery and vein. The

ISSN 2616-5775 Vol. 6, Issue 9: 81-86, DOI: 10.25236/AJCIS.2023.060912

Shang Ying Xiang acupoint (EX-HN8) is positioned at the top of the left nasolabial groove, housing the anterior ethmoid nerve, infraorbital nerve, branches of the inferior ophthalmic nerve, and facial veins. The Di Cang(ST4) acupoint is located at the left corner of the mouth, containing branches of the facial nerve and inferior ophthalmic nerve.

We will identify these acupoints, analyze their connections within the same group, study their relationship with the facial midline, and establish evaluation criteria in collaboration with authoritative medical professionals in the field. We set a 1-degree angle as the standard. If the angle is less than 1 degree, no damage is present, corresponding to level 0 damage. For angles greater than 1 degree, damage is increased by one level for every 5 degrees, with the highest level being 7 for angles beyond 31 degrees.

3. Method

Taking into consideration the impact of model efficiency and recognition speed on user experience during the deployment of the APP, we utilized the MobileNetV2 network for coarse keypoint extraction. MobileNet is a widely used lightweight network in current times, particularly suitable for deployment on mobile devices.

We perceive facial paralysis recognition as a six-point screening of the human face. For detecting eyes, nose, and mouth, patients need to be photographed in different facial poses. Consequently, each patient requires three images for screening. However, predicting results based on two points in these three images would make the network learn which facial pose corresponds to which specific part of the face, which overburdens the network and hampers its training effectiveness. Furthermore, our analysis of using MobileNetV2 for coarse keypoint extraction led to two conclusions:

1) The Euclidean distance between predicted and actual points remains within a very small range.

2) Network results often exhibit substantial discrepancies, and the model struggles to fit our desired values, instead prioritizing the minimization of global loss.

Based on the observations and reflections on these conclusions, we refined the network. We designed a two-stage network resembling object detection followed by regression: In our data, the two key points in a single image are sparse, with only a tiny fraction of information being relevant, while the rest is redundant. To address this, we adopted a two-stage regression approach. In the first stage, we used the MobileNet network for coarse regression of key points. We took the regression point as the center and cropped a patch from the original image with an appropriate width and height, ensuring that the patch contains the actual point. We scaled and matched the coordinates of the target point. Finally, we concatenated the required six patches and fed them into a simple convolutional network with ReLU and max pooling layers. A linear layer at the end provided the regression output for predicting target points. The specific process is shown in Figure 1.

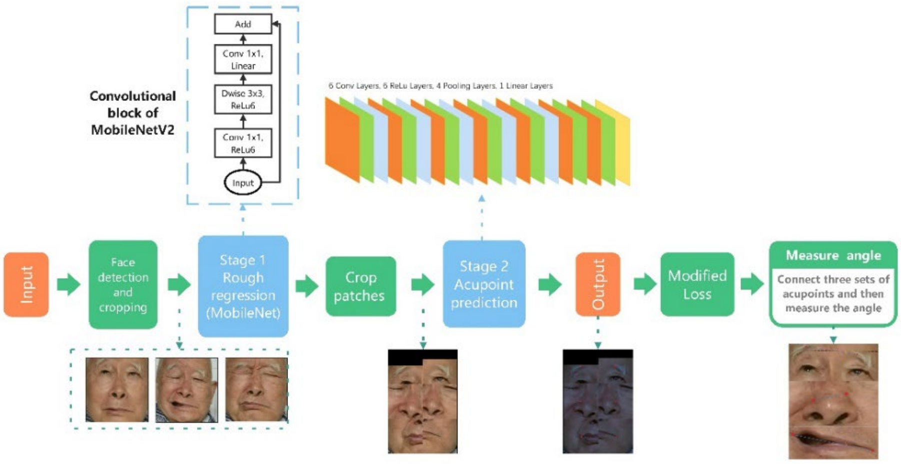

Figure 1: The process of intelligent facial palsy diagnostic system

Furthermore, we enhanced the Loss function. While we are regressing six points from a single image, only two points hold reference value during individual actions. Hence, we only calculate the Loss for the two points corresponding to the relevant action. Moreover, our evaluation target requires an error in

ISSN 2616-5775 Vol. 6, Issue 9: 81-86, DOI: 10.25236/AJCIS.2023.060912

relation to the horizontal angle, specifically vertical error / horizontal error. It's evident that the eyes, nose, and mouth have certain horizontal spacing, and vertical error will play a dominant role.

$$Dist = |\frac{1}{n}\sum_{i=1}^{n} Lx_i - Rx_i|$$

$$f(n) = |\frac{1}{n}\sum_{i=1}^{n} \frac{(\widehat{Lx_{n_i}} - Lx_{n_i}) - (\widehat{Rx_{n_i}} - Rx_{n_i})}{Dist_n} + (\widehat{Ly_{n_i}} - Ly_{n_i}) - (\widehat{Ry_{n_i}} - Ry_{n_i})|$$

In the above, "*Dist*" represents the distance between the acupoints on the left and right sides within the same group. "*Lx/Rx*" signifies the x-coordinate of the left/right acupoint within each acupoint set, while "*Ly/Ry*" corresponds to the y-coordinate. Lx_{n_i} / $\widehat{Lx_{n_i}}$represent the real value and the predicted value.Using this information, we can derive the following new loss function:

$$Loss_{new} = f(ST4) + f(EX - HN4) + f(EX - HN8)$$

4. Experiment

4.1. Dataset

The data for this experiment consists of facial photographs, with labels corresponding to the six acupoints on the face (two on the brow, two on the nose, and two on the mouth). The primary objective of the experiment is to calculate the angular deviation between the line connecting the two points on different parts of the face and the horizontal direction. The maximum value among these deviations is considered as the deviation angle for that particular subject.

Each individual in the dataset has three images, and the example is shown in Figure 2. These images are taken under different photographic pose requirements, resulting in three images for each person. Each image corresponds to a set of acupoints for testing. The experimental data originates from real clinical data at the Sichuan Provincial Hospital of Integrated Traditional Chinese and Western Medicine. The facial images are clear and centered.

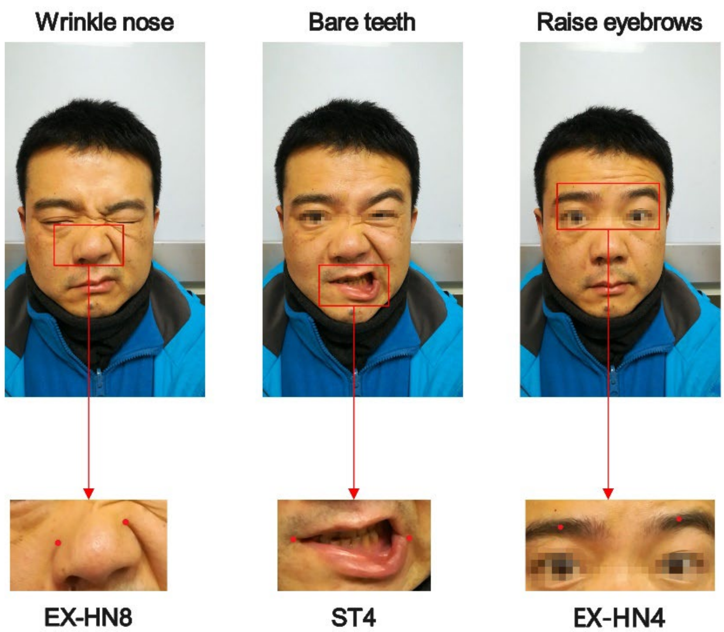

Figure 2: Locations of the three groups acupuncture points

4.2. Training

During the training process, we first need to detect the faces in the dataset images and crop out irrelevant information to reduce the model's workload. After cropping and resizing to a uniform size, the images are then fed into the model.

Then we conducted training on the two-stage network using images of patients at different levels of severity from the dataset. To enhance the model's training capability, we pre-trained the model by extracting features for each of the three sets of acupoints marked on every image, capturing distinct acupoint characteristics. After obtaining the pre-training results, we cropped and concatenated the three images corresponding to different acupoint sets for each patient and performed accurate coordinate regression on the concatenated image.

We train the model using hardware equipment with CPU 13th Gen Intel(R) Core(TM) i9-13900HX 2.20 GHz and GPU NVIDIA GeForce RTX 4060.

The training was conducted in batches of 32 images, with shuffling within each batch. We employed the Adam optimizer for gradient updates in both stages of training. The initial learning rate was set at 0.01, with a weight decay of 0.0005. The learning rate was decreased by 0.1 every 30 epochs.

4.3. Evaluate

In terms of performance evaluation, we conducted comparative experiments using models such as VGG16. The evaluation metric we employed was the Normalized Mean Error (NME), which is commonly used in tasks such as image alignment and keypoint detection. NME measures the average distance between predicted values and ground truth. A smaller NME indicates that the predicted results are closer to the ground truth, signifying better algorithm performance.

$$NME = \frac{1}{n}\sum_{i=1}^{n}\frac{||y_n - y_n'|_2|}{d}$$

Through comparative experiments, we have obtained the following results in table 1. It is evident from these results that our improved model has achieved optimal performance. This validates the effectiveness of our work.

Table 1: Comparison of the different models

Network	NME	Loss	average_angle_cost
ResNet50	0.0054	0.7636	6.36
VGG16	0.0031	0.4125	4.32
MobileNetV2	0.0032	0.5035	5.63
Ours	**0.0031**	**0.2744**	**3.78**

5. Conclusion

This paper presents a facial paralysis diagnostic method based on acupoint recognition. We construct a two-stage network combining object detection and regression, and refine the Loss function tailored to our task. This lightweight approach demonstrates excellent performance in quantitative assessment of facial paralysis. To validate the framework's performance, we compared it with detection results from ResNet50 and other models. Experimental results reveal significant improvements in acupoint detection and angle measurement accuracy achieved by our enhanced model. These outcomes affirm the contribution of our work.

Acknowledgements

This work is supported by Chengdu Science and Technology Bureau (2022-YF05-01924-SN), the postdoc research funding of Sichuan University (2022SCU12079), and the postdoc inter-disciplinary innovation funding of Sichuan University.

References

[1] Reginald F. Baugh et al. Clinical Practice Guideline: Bell's Palsy[J]. Otolaryngology–Head and Neck Surgery, 2013, 149(2_suppl): P 13-P14.
[2] House J W, Brackmann D E. Facial nerve grading system [J]. Otolaryngology-head and neck surgery:

ISSN 2616-5775 Vol. 6, Issue 9: 81-86, DOI: 10.25236/AJCIS.2023.060912

official journal of American Academy of Otolaryngology-Head and Neck Surgery, 1985, 93(2): 146-148.
[3] T. H. Ngo, M. Seo, N. Matsushiro and Y. -W. Chen, "Quantitative analysis of facial paralysis based on filters of concentric modulation, " 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, 2015, pp. 1758-1763, doi: 10. 1109/FSKD. 2015. 7382213.
[4] A. M. Ridha, W. Shehieb, P. Yacoub, K. Al-Balawneh and K. Arshad, "Smart Prediction System for Facial Paralysis, " 2020 7th International Conference on Electrical and Electronics Engineering (ICEEE), Antalya, Turkey, 2020, pp. 321-327, doi: 10. 1109/ICEEE49618. 2020. 9102600.
[5] Wang Bolu, Wu Xiaomei. Application of Artificial Intelligence in Facial Paralysis Recognition and Assessment [J]. Chinese Journal of Medical Devices, 2022, 46(1): 6.
[6] Cootes T F, Taylor C J, Cooper D H, et al. Active shape models-their training and application[J]. Computer vision and image understanding, 1995, 61(1): 38-59. 3
[7] Y. Sun, X. Wang and X. Tang, "Deep Convolutional Network Cascade for Facial Point Detection, " 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 2013, pp. 3476-3483, doi: 10. 1109/CVPR. 2013. 446.
[8] Zhang K, Zhang Z, Li Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks [J]. IEEE signal processing letters, 2016, 23(10): 1499-1503.
[9] Kowalski M, Naruniec J, Trzcinski T. Deep alignment network: A convolutional neural network for robust face alignment[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017: 88-97.
[10] Zhao Y , Zhang D , Wang Y .Automatic location of facial acupuncture-point based on content of infrared thermal image[C]//International Conference on Computer Science & Education. IEEE, 2010.DOI:10.1109/ICCSE.2010.5593668.
[11] Chang M, Zhu Q. Automatic location of facial acupuncture-point based on facial feature points positioning[C]//Research Institute of Management Science and Industrial Engineering. Proceedings of 2017 5th International Conference on Frontiers of Manufacturing Science and Measuring Technology (FMSMT 2017). Atlantis Press, 2017:5.
[12] Wang Qi, Chen Gong, Hu Wenxin, et al. Research on Facial Paralysis Recognition Application Using GANomaly Network [J]. Journal of Software Engineering, 2022, 25(03): 29-33. DOI: 10. 19644/j. cnki. issn2096-1472. 2022. 003. 007.
[13] Xu Pengfei, Yang Chunlei, Sun Haojie, et al. Facial Paralysis Grading Assessment Method Based on Deep Temporal Features [J/OL]. Acta Automatica Sinica. http://kns. cnki. net/ kcms/detail/ 11. 2019. TP. 20200414. 1951. 005. html.
[14] Neely J G, Joaquin A H, Kohn L A, et al. Quantitative assessment of the variation within grades of facial paralysis [J]. Laryngoscope, 1996, 106: 438-442
[15] Wang T, Zhang S, Liu L, et al. Automatic evaluation of the degree offacial nerve paralysis [J]. Multimedia Tools & Applications, 2016, 75(19): 11893-11908
[16] Sandler M, Howard A, Zhu M, et al. Mobilenetv 2: Inverted residuals and linear bottlenecks[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 4510-4520.