ISSN 2616-5775 Vol. 7, Issue 6: 95-99, DOI: 10.25236/AJCIS.2024.070615

Classification of Facial Paralysis Based on Machine Vision Techniques

Lisha Bai[1], Yujian Li[1], Haojie Chang[1]

[1]School of Mechanical and Electronic Engineering, Wuhan University of Tecnology, Wuhan, China

Abstract: *With the development of computer-aided diagnosis, numerous studies have employed specialized algorithms to extract diagnostically valuable features from images, optimize the subsequent processing of these features, and finally perform classification evaluation on the processed features. This paper combines technologies such as facial recognition, image analysis and processing, and convolutional neural networks to design and implement a comprehensive facial nerve evaluation system. To comprehensively assess the severity of facial nerve disease in patients, this paper designs a Convolutional Neural Network (CNN) model for the extraction and grading of facial paralysis features. From data acquisition, preprocessing, data augmentation, model training to prediction and evaluation, a complete model for assessing and predicting the severity of facial paralysis in patients is established.*

Keywords: *Classification Evaluation, Facial Paralysis Features, Neural Network Model*

1. Introduction

Facial nerve paralysis, commonly known as facial palsy, is a prevalent health issue. Its primary characteristic is the impairment of facial muscle movement, which negatively affects various aspects such as facial expression, fluent speech, and interpersonal communication. Currently, there are significant differences in the grading standards for facial palsy both domestically and internationally, and these evaluations often rely on the subjective judgment of physicians. Traditional evaluation methods are not only inefficient but also pose a risk of potential misjudgment.

In recent years, many scholars at home and abroad have conducted research on automatic grading and evaluation of facial palsy based on computer vision. These studies have largely mitigated the subjectivity of manual diagnosis and reduced the burden on healthcare workers. Some of these research efforts have been successfully applied in clinical medical diagnosis, providing valuable references and suggestions for physicians. Barbosa et al. used a hybrid classification model to detect key points and analyzed the movement of these key points to study facial palsy [1]. He et al. utilized the Local Binary Pattern (LBP) model to extract horizontal and vertical feature information from facial regions to assess facial symmetry [2]. Wang et al. proposed a method for evaluating the severity of facial palsy that comprehensively considers both the static and dynamic features of the patient's face [3]. Guo et al. applied the Google-Net model in their research on facial palsy grading, achieving a nonlinear mapping from facial images to H-B grading levels [4]. This method enables more accurate grading of given facial palsy data.

Based on the current state of research both domestically and internationally, deep learning-based methods for evaluating the severity of facial palsy are considered an inevitable trend for future development. This paper will construct a neural network model for the extraction and grading of facial palsy features, enabling the rapid and accurate processing of large amounts of facial movement data. This approach avoids the complexity and time consumption of manual evaluation, providing a more convenient assessment tool for clinical practice.

2. Design of Facial Feature Extraction and Grading Model

This paper addresses the complex issue of evaluating the severity of facial palsy by developing a feature extraction method and establishing a scoring sample database, thereby resolving the input-output problems of model training. Consequently, an automated system for assessing facial symmetry has been constructed.

ISSN 2616-5775 Vol. 7, Issue 6: 95-99, DOI: 10.25236/AJCIS.2024.070615

2.1. Overall Process

The main content of this paper focuses on designing the assessment pathway for the severity grading of facial palsy patients, comprising 5 steps: data acquisition, preprocessing, data augmentation, model training, and prediction and evaluation. Figure 1 illustrates the overall process flowchart.

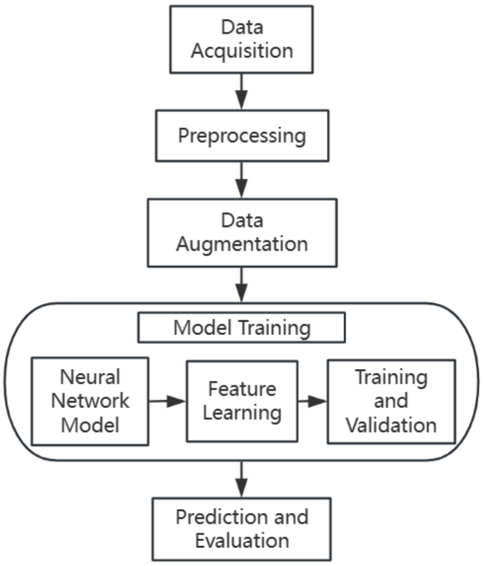

Figure 1: Overall Process Diagram.

2.2. Model Principles

Convolutional Neural Network (CNN) is a type of deep learning model particularly suited for processing data with grid-like structures such as images and videos. It consists of multiple layers of neurons, and its core idea is to extract features from input data through convolutional operations. The main components of CNN include Convolutional Layers, Pooling Layers, and Fully Connected Layers.

In the convolutional layer, a series of convolutional filters (kernels) slide over the input image with a specified stride. Each filter computes a weighted sum of pixels within the current window, generating elements in the feature map. This process extracts local features and progressively constructs more complex feature representations layer by layer, capturing edges, textures, and other information in the image for feature extraction and representation. The formula for the feature map is:

$$\omega' = (\omega + 2p - k) - s + 1 \tag{1}$$

Where ω represents the size of the input matrix, p denotes the number of zero-padding layers, k indicates the size of the convolutional kernel, and s represents the stride.

In the convolutional layer, each individual convolutional kernel can only extract a specific type of feature. The pooling layer is an indispensable component of convolutional neural networks, effectively reducing the dimensions of the feature maps. This not only simplifies the computation process but also significantly enhances the model's translational invariance.

The fully connected layer plays the role of a classifier in the neural network. Its primary function is to map the feature maps processed through the pooling layers to the final output class space. This process is achieved through the weights and biases within the fully connected layer, ultimately generating classification results from the output layer, thereby accurately classifying the input data.

2.3. Model Design

The convolutional layers are effective in extracting image features by using local receptive fields to capture spatial structural information within the images. Through a stacked design of convolutional layers, the model progressively captures local details and texture features of the images, thereby gradually extracting more abstract and complex image features. This hierarchical learning structure enables the model to gain a deeper understanding of the internal structure of images, thereby enhancing its performance across various visual tasks. The network parameters for the facial paralysis feature extraction and grading model are shown in Table 1.

Table 1: Network Parameters of the Model.

Network Layer	Input Size	Convolution Kernel Size	Stride	Activation Function
Input Layer	256×256×1			
Convolutional Layer 1	254×254×12	5×5	1	LeakyReLU
Pooling Layer 1	127×127×12	2×2	2	
Convolutional Layer 2	125×125×12	5×5	1	LeakyReLU
Pooling Layer 2	62×62×12	2×2	2	
…	…	…	…	…
Fully Connected Layer 1	13824			LeakyReLU
Fully Connected Layer 2	512			LeakyReLU
Fully Connected Layer 3	64			

The input images are grayscale with 1 channel. Each convolutional layer is followed by a batch normalization layer and LeakyReLU activation function. The fully connected layers flatten the output, with the last layer having an input size of 64 and output size of 4, corresponding to 4 categories.

LeakyReLU activation function introduces a small slope in its negative region, which helps improve gradient behavior during backpropagation. This small slope ensures that gradients do not drop to near-zero for negative inputs, thereby mitigating the vanishing gradient problem. This approach enhances the model's sensitivity to negative inputs, facilitating better learning and adaptation to different data features during training.

Batch Normalization layers expedite the model training convergence process and reduce sensitivity to initial parameter selection. By normalizing the input data to have zero mean and unit variance, Batch Normalization enhances the stability and robustness of the model, reduces internal covariate shift, and improves generalization ability.

The facial feature extraction model can extract the following facial features:

1) Eyelid Movement: Recognizing the open and closed states of eyelids to detect if the eyes can close properly.

2) Mouth Corner Movement: Detecting upward or downward movement of mouth corners to assess asymmetry in smiling or mouth movements.

3) Eyebrow Movement: Analyzing the degree of eyebrow elevation and descent to assess if eyebrows can move symmetrically.

4) Lip Shape Changes: Detecting changes in lip shape to assess if the patient can express emotions normally.

5) Facial Muscle Symmetry: Assessing the symmetry of facial muscles on both sides; patients with facial paralysis often exhibit symptoms of one side having less flexible or weaker facial muscle movements.

3. Experimental Results and Analysis

Due to the personal privacy of facial paralysis patients, there is currently no publicly available dataset for automatic diagnosis in the field of facial paralysis research. To validate the effectiveness of the model designed in this paper, images were collected from YouTube Facial Paralysis (YFP) [5] and from various sources on the internet. The data was annotated based on facial paralysis assessment features set by the House-Brackmann scale [6], which was annotated and cross-validated by two professional doctors. Finally, accuracy was used as the metric to evaluate the model performance. Figure 2 shows examples of the images.

ISSN 2616-5775 Vol. 7, Issue 6: 95-99, DOI: 10.25236/AJCIS.2024.070615

Figure 2: Image Examples.

3.1. Handling of Experimental Data

Splitting the collected dataset into training and validation sets, as well as a test set, with a ratio of approximately 4:1. The training and validation sets comprise about 80% of the dataset, with the remaining 20% reserved for testing. Table 2 shows the number of images for each grade.

Table 2: Number of Images for Each Grade.

H-B Grade	Quantity
I (Normal)	234
II,III (Mild to Moderate Dysfunction)	249
IV (Moderate to Severe Dysfunction)	261
V (Moderate to Severe Dysfunction)	256
Total	1000

During the data splitting process, random and stratified sampling was employed to ensure representation of all grades of facial paralysis patients in both training and testing datasets, maintaining an equal distribution across the dataset.

All facial images underwent the following preprocessing steps: face alignment; grayscale conversion to eliminate color bias; histogram equalization to remove shadow and lighting variations; cropping to obtain the largest facial image while removing unnecessary backgrounds. Subsequently, data augmentation was performed. One inherent limitation of CNNs is overfitting during testing phases. Overfitting occurs when a CNN has low training errors but high validation errors.

3.2. Experimental Results and Statistical Analysis

To validate the effectiveness and accuracy of the model proposed in this paper for facial paralysis grading, experiments were conducted to compare its performance parameters with classical neural network models including AlexNet, VGG16, and Inception. These models all utilize single convolutional neural network architectures for facial paralysis grading classification. The accuracy, precision, and F1 score differences between traditional methods and the approach proposed in this paper are shown in Table 3.

Table 3: Comparison of Experimental Results Using Different Methods.

	Accuracy	Precision	F1
AlexNet	0.625	0.612	0.546
VGG16	0.715	0.703	0.688
Inception	0.782	0.688	0.732
in this paper	0.828	0.845	0.830

From Table 3, it can be observed that compared to traditional methods, the method proposed in this paper has improved accuracy by 20.3%, 11.3%, and 4.6%; precision by 23.3%, 14.2%, and 15.7%; and F1 score by 28.4%, 14.2%, and 9.8%, respectively. Based on these metrics, the method proposed in this paper outperforms traditional convolutional neural network models.

4. Conclusion

In real-life clinical practice, the diagnosis of the severity of facial paralysis has traditionally relied on

ISSN 2616-5775 Vol. 7, Issue 6: 95-99, DOI: 10.25236/AJCIS.2024.070615

the subjective judgment and clinical experience of specialized doctors. Therefore, improving the convenience of facial paralysis grading assessment and helping reduce the workload of medical professionals is highly meaningful. Many scholars both domestically and internationally have contributed to the automatic grading of facial paralysis, achieving certain effectiveness. Building upon this foundation, this paper designs a Convolutional Neural Network (CNN) model for extracting facial features and grading facial paralysis patients, applying it to the study of facial paralysis grading assessment.

The combined structure of layers in this network effectively learns features specific to facial paralysis grading tasks, including five features: eyelid movement, mouth corner movement, eyebrow movement, lip shape changes, and facial muscle symmetry. These features demonstrate strong representational and generalization capabilities. From the results of experimental and statistical analyses, it is evident that compared to traditional methods, this model shows significant improvements in accuracy, precision, and F1 score, aligning closely with diagnoses made by professional doctors.

References

[1] Gaber A , Taher M F , Wahed M A ,et al.Classification of facial paralysis based on machine learning techniques[J].Biomedical engineering online, 2022, 21(1):65.DOI:10.1186/s12938-022-01036-0.
[2] Fei X , Mengmeng S , Xinming G ,et al.Comprehensive evaluation of facial paralysis classification by combining the facial texture features and optical flow[J].Journal of Northwest University(Natural Science Edition), 2019.
[3] B. , Tanwar D P .Evaluation of 3d facial paralysis using fuzzy logic [J].International Journal of Engineering & Technology, 2018. DOI:10.14419/ijet.v7i4.13619.
[4] Barbosa J, Seo W K, Kang J .ParaFaceTest: An ensemble of regression tree-based facial features extraction for efficient facial paralysis classification [J].BMC Medical Imaging, 2019, 19(1).DOI:10.1186/s12880-019-0330-8.
[5] Chuang C C , Chang N J , Lu C Y .Postparalysis Facial Synkinesis: Clinical Classification and Surgical Strategies[J].Plastic and Reconstructive Surgery – Global Open, 2015, 3.DOI:10.1097/ GOX.0000000000000283.
[6] Anguraj K, Padma S. A Precise Facial Paralysis Degree Evaluation with Severity Classification Using Image Processing and Neural Network[J].International Review on Computers and Software, 2013, 8(2):569-576.