Quantitative Assessment of Hand Motor Function for Post-Stroke Rehabilitation Based on HAGCN and Multimodality Fusion

Chenguang Li, *Student Member, IEEE*, Hongjun Yang, Long Cheng, *Senior Member, IEEE*, Fubiao Huang, Shuang Zhao, Dongyue Li, and Ruxiu Yan

Abstract— **Quantitative assessment of hand function can assist therapists in providing appropriate rehabilitation strategies, which plays an essential role in post-stroke rehabilitation. Conventionally, the assessment process relies heavily on clinical experience and lacks quantitative analysis. To quantitatively assess the hand motor function of patients with post-stroke hemiplegia, this study proposes a novel multi-modality fusion assessment framework. This framework includes three components: the kinematic feature extraction based on a graph convolutional network (HAGCN), the surface electromyography (sEMG) signal processing based on a multi-layer long short term memory (LSTM) network, and the quantitative assessment based on the multi-modality fusion. To the best of the authors' knowledge, this is the first study of applying a graph convolution network to assess the hand motor function. We also collect the kinematic data and sEMG data from 70 subjects who completed 28 types of hand movements. Therapists first graded patients using traditional motor assessment scales (Brunnstrom Scale and Fugl-Meyer Assessment Scale) and further refined the patient's motor assessment result by their experience. Then, we trained the HAGCN and LSTM networks and quantitatively assessed each patient based on the proposed assessment framework. Finally, the Spearman correlation coefficient (SC) between the assessment result of this study and the traditional scale are 0.908 and 0.967, demonstrating a significant correlation between the proposed assessment and the traditional scale scores. In addition, the SC value between the score of this study and the refined hand motor function is 0.997, indicating the "ceiling effect" of some traditional scales can be avoided.**

Index Terms— **Hand motor function, quantitative assessment, multi-modality fusion, graph convolutional network.**

I. INTRODUCTION

APPROXIMATELY 80% of stroke survivors suffer from motor dysfunction that affects one or both upper limbs, especially the coordination and flexibility of the hands [1]. Hand impairment is one of the major causes of functional limitations in individuals with post-stroke hemiparesis. Given that the hand provides 90% of the motor function of the upper limb, the patient's autonomy is reduced, thus affecting the performance of daily activities and reducing the quality of life [2]. An adequate treatment of hand motor function impairment is necessary for establishing a realistic prognosis, planning customized rehabilitation interventions, and evaluating the effectiveness of those interventions [3].

In the clinical setting, assessments of hand motor function are typically performed by the 'standardized' clinical scales and tests [4]. The Brunnstrom assessment (BA) scale [5] and Fugl-Meyer assessment (FMA) scale [6] are two most commonly used hand function assessment scales for stroke patients. Swedish physiotherapist Brunnstrom developed BA in the 1970s to assess movement disorders following central nervous system injuries. In his theory, stroke patients are divided into six stages of recovery. The FMA scale of the hand contains seven types of hand assessment movements. Each movement is assigned a qualitative rating with a score of 0, 1 or 2 depending on how well it is performed. However, since both scales are graded according to the therapist's own experience, both scales are subjective and have the disadvantage of the "ceiling effect".

The "ceiling effect" widely exists in assessing post-stroke upper extremity dyskinesia [7], which means that the scale is insensitive to the change of those patients at the top end of the recovery (*i.e.*, the scale resolution is low for patients with good motor recovery) [8]. In this study, the "ceiling effect" is suggested to occur when the patient's assessment score reaches over 80% of the maximum score. In this setting, approximately 80% of patients in the experiments conducted in this study face the "ceiling effect". It is obvious that the scale "ceiling effect"

affects the assessment accuracy, which should be addressed in practice.

Due to the above shortcomings, some papers in the literature used sensors to measure the hand kinematic information of stroke patients accurately and used artificial intelligence methods to evaluate their hand motor functions. Fang *et al.* allowed patients to perform seven FMA-specific movements above the leap motion (LM) sensor and evaluated the patients' FMA scores based on the range of each finger's angular changes detected by the LM [9]. Hamaguchi *et al.* used LM to record the finger angle changes of 24 stroke patients' hands during group flexion and extension within seven seconds, calculated the peak angle and normalized peak velocity, and then used a support vector machine method to classify the patients into six categories [10]. Li *et al.* designed a set of combined movements of the wrist and fingers and then utilized LM to measure the angle ranges of the patient's fingers when performing the movements to evaluate the patients by ensemble learning [11]. Song *et al.* designed a mobile phone-based automated Fugl-Meyer assessment system for stroke patients. Patients were asked to complete specific tasks when holding a mobile phone. Afterward, the motion information collected by the mobile phone was used in combination with the decision tree method to evaluate the patient's upper limb function [12]. Adams *et al.* built a virtual environment assessment system. The system assessed the patient based on the task completion time, the average speed of hand movement, and task scores. Spearman rank correlations showed a high and significant correlation between virtual world-derived measures and gold-standard assessments [13], [14]. Although the above studies used artificial intelligence methods to assess the hand motor function of patients directly, they all used the methods adopting features by the manual extraction manner. Nevertheless, the manually extracted features may not be optimal.

Deep learning belongs to the method which is capable of automatically extracting features. A typical deep learning network is the convolutional neural networks (CNNs). With the development of graph neural networks, graph convolution networks (GCNs), which extend CNNs to graphs of arbitrary structures, have received increasing attention and have been successfully used in various applications, such as image classification, document classification, and skeleton-based movement recognition [15]–[17]. In particular, Yan *et al.* proposed spatial-temporal graph convolutional networks for skeleton-based movement recognition and achieved a good classification result [18]. Whereas, hand joints can also be regarded as a graph network structure and the GCN-based hand motion assessment has yet to be explored.

Furthermore, besides kinematic signals, sEMG signals also reflect the hand motor function of stroke patients to a certain extent. Zhang *et al.* estimated the muscle strength by collecting sEMG signals and using a third-order polynomial fitting technique [19]. The muscle strength can reflect the motor function. Some studies have realized that studying both hand kinematics and sEMG data when conducting clinical trials can better understand the muscle coordination in functional recovery. However, they did not study how to implement the multi-modality fusion evaluation and only studied healthy people rather than stroke patients [20]–[23].

Given the above observations, the main contributions of this paper can be summarized as follows:

1) To automatically extract practical features and make full use of the spatial position information of the human hand joints, we propose a hand assessment graph convolution network (HAGCN). The network includes the graph convolution in the spatial domain and the temporal convolution in the time domain. To the best of the authors' knowledge, this is the first study of applying the GCN to assess the hand motor function.

2) Both the motion signal and the sEMG signal are analyzed, and the weighted decision fusion method is used to assess the hand motion function of patients with post-stroke hemiplegia, which improves the accuracy of the assessment. The SCs between the assessment result of this study and the traditional scales are 0.908 and 0.967, respectively, proving that there is a significant correlation between the proposed assessment and the traditional scale scores.

3) The therapists have refined the assessment results of 25 stroke patients facing the "ceiling effect". These stroke patients have been assessed by the proposed algorithm as well. The SC between the score of this study and the refined assessment is 0.997, indicating that the "ceiling effect" in some traditional scales can be avoided.

The remaining parts of this study are organized as follows: Section II introduces the proposed experimental setup and the acquisition of multi-modality data. Section III presents details on the assessment framework based on HAGCN, LSTM and multi-modality data fusion. Then, the assessment results and the related discussions are provided in Section IV and Section V, respectively. Finally, Section VI concludes the paper with final remarks.

II. EXPERIMENTAL METHODS

A. Participants

The experiments were performed in collaboration with the China Rehabilitation Research Center (Beijing Bo'ai Hospital) and we recruited 35 post-stroke hemiparetic patients (27 males, 8 females, mean age of 52.7 ± 12.7 years) from the hospital. The study imposes no subject requirements in terms of the minimum level of required motor function, as long as the subject has no cognitive deficits. Before the experiment, each post-stroke participant was examined by three experienced therapists for the Brunnstrom stage classification and the hand section of the Fugl-Meyer assessment. Then, according to the majority rule, the BS and FS of the patients are determined. The detailed clinical assessment results of the enrolled subjects are shown in Table I.

To eliminate the "ceiling effect", three therapists selected 25 patients with the Brunnstrom stage greater than III, touched these patients, felt the patient's gripping strength by shaking hands with them, and observed the patient's completion of some props tasks (such as drawing strokes and inserting nails)

TABLE I
CLINICAL ASSESSMENT RESULTS OF 35 POST-STROKE PATIENTS

Subject	Age/Sex	Etiology	Side	BS	FS
1	55/M	Ischemic	Left	5	12
2	58/M	Ischemic	Right	6	14
3	69/F	Ischemic	Left	5	11
4	52/M	Ischemic	Right	6	14
5	54/M	Ischemic	Left	5	11
6	55/F	Hemorrhagic	Left	4	8
7	57/F	Ischemic	Right	5	12
8	54/M	Ischemic	Right	6	14
9	51/M	Ischemic	Right	6	14
10	46/M	Hemorrhagic	Left	6	14
11	55/M	Ischemic	Left	5	12
12	63/M	Ischemic	Right	4	7
13	46/M	Hemorrhagic	Right	3	3
14	61/M	Ischemic	Right	5	10
15	48/F	Ischemic	Left	3	5
16	65/M	Ischemic	Left	1	1
17	45/M	Hemorrhagic	Right	5	11
18	65/M	Ischemic	Left	3	3
19	49/M	Ischemic	Right	5	10
20	69/M	Ischemic	Left	6	13
21	65/F	Ischemic	Left	2	4
22	44/F	Ischemic	Right	6	13
23	60/M	Hemorrhagic	Left	6	13
24	51/M	Ischemic	Left	2	3
25	49/M	Hemorrhagic	Left	6	14
26	52/M	Hemorrhagic	Left	3	3
27	51/F	Hemorrhagic	Right	6	14
28	68/M	Ischemic	Right	3	3
29	13/M	Hemorrhagic	Right	2	3
30	77/M	Hemorrhagic	Right	4	9
31	33/M	Hemorrhagic	Right	4	10
32	37/M	Hemorrhagic	Right	6	14
33	62/M	Ischemic	Left	6	13
34	34/F	Hemorrhagic	Right	5	10
35	32/M	Ischemic	Left	6	14

* BS refers to Brunnstom score.
* FS refers to FMA score.

in the occupational therapy room to further rank the patients. This step still follows the majority rule. The sorting result is shown in Order 1 of Table V.

This research was reviewed and approved by the Ethics Committee of the China Rehabilitation Research Center (approval number: 2021-108-1). Each subject signed a written informed consent form before enrollment.

B. Acquisition Setup

To explore the kinematic and muscular characteristics in normal and pathological movement patterns, we collect the kinematic and sEMG data of the subjects simultaneously.

1) Kinematics: Kinematic data are acquired by WISEGLOVE19 (Xintian Vision, Beijing, China) at a sampling rate of 200 Hz; the device collects data from 19 joint angles of the fingers. The data glove adopts an optical fiber to measure the angle, with a dynamic accuracy of 0.2 degrees.

Due to the use of optical fiber sensors, the maximum and the minimum values of the wearer's finger angle should be calibrated before using the data glove to collect data. Considering the patient's hand dysfunction, the volunteer needs to assist the patient in calibration by performing some calibration movements.

Fig. 1. The snapshot of the experiment.

Fig. 2. Proposed 28 hand movements in the assessment experiment.

2) Surface Electromyography: The muscular activity is gathered using a Thalmic Myo armband (Thalmic Labs, Ontario, Canada), a low-cost wireless armband containing eight single differential sEMG sensors. The sampling rate is also set to be 200 Hz, which is the same as the kinematic sampling rate. Eight electrodes are evenly wound around the forearm, keeping a constant distance from the radiohumeral joint just below the elbow. A snapshot of the experiment is shown in Fig. 1.

C. Experimental Paradigm

The hand movements in the experiment are proposed by therapists based on their clinical experience, aiming at evaluating the hand motor function, which are shown in Fig. 2. These movements are divided into two parts: one part contains five fundamental movements of the wrist (no. 1 – no. 5 in Fig. 2), four isometric and isotonic hand configurations (no. 6 – no. 9 in Fig. 2), and five combination movements (no. 10 – no. 14 in Fig. 2); the other part contains the left fourteen grasping movements.

The repetitions of one action are performed in one block, and the order of the blocks is the same as the sequence of
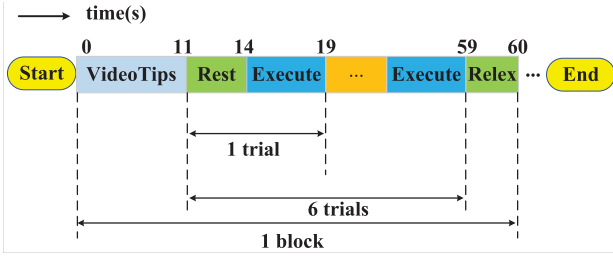
Fig. 3. The structure of one hand movement experiment block.

actions shown in Fig. 2. Figure 3 shows one block. There are video tips at the beginning of each block. The video of the hand movements are played twice, instructing the subject on what to do next. After resting for 3 seconds, subjects try their best to execute the hand movement within 5 seconds. This whole process is called a trial, and 6 trials form a block.

D. Data Acquisition and Data Preprocessing

After the experiment, we obtain two types of time-series signals. One is the kinematic signal, which is composed of 19 channels, and the other is the sEMG signal, which is composed of 8 channels. The kinematic data are filtered by a second-order two-way low-pass Butterworth filter with the cut-off frequency of 5 Hz [21]. The Thalmic Myo already presents a notch filter at 50 Hz, so the sEMG signal requires no extra filtering [24].

To expand the number of samples, we perform the sliding window method, and the window size and the sliding distance are consistent with the ones given in [25]. Therefore, each movement repetition has a window of 200 milliseconds (20 sampling points), with an overlap of 100 milliseconds (10 sampling points). All data preprocessing works are performed on MATLAB R2019a.

III. MULTI-MODALITY FUSION FRAMEWORK FOR FUNCTIONAL ASSESSMENT

After data collection, this section introduces the hand function assessment method based on the fusion of two data modalities. The overall framework is shown in Fig. 4. The method mainly includes the movement analysis based on HAGCN, the sEMG analysis based on LSTM, and the multi-modality fusion scheme. All methods are implemented by the Python language (version 3.6) based on the TensorFlow and PyTorch framework. We first introduce the classification task of this study.

A. Classification Task

It should be noted that the classification task is not designed to recognize the 28 hand movements but to distinguish patients of different levels through each hand movement. We expect to use the assessment framework to obtain more accurate assessment results and eliminate the "ceiling effect" of the scale, because the "ceiling effect" reduces the data interpretation accuracy and affects the effectiveness of the rehabilitation progress [26].

The therapists select and label 25 patients among 35 recruited patients to represent 25 categories at the top end of the recovery. The therapists suggest that 25 categories are sufficient to avoid the "ceiling effect". To distinguish these 25 categories, we have also added one extra category (the healthy category). We first collected data from 35 healthy subjects. Then, we randomly selected six groups of data generated by the same 28 hand movements of the healthy subjects to represent the healthy category. Since the selection is random, the selected data can represent the movements of most healthy subjects. In this way, there are a total of 26 categories in the classification task.

B. Motion Analysis Based on HAGCN

HAGCN includes a spatial graph convolution network (SGCN) and a temporal graph convolution network (TGCN).

1) Skeleton Graph Construction: In this work, we utilize a spatial graph to form a hierarchical representation of the hand skeleton sequence. The structure of SGCN is shown in Fig. 5. The yellow dots on the left in Fig. 5 are 19 key joints measured by the data glove. Leveraging the natural connections between these joints, we propose a graph structure. The structure can explicitly exploit the spatial relationship between the joints, which is crucial for understanding human actions. In the spatial graph, the internal edges between human joints are defined according to the natural connections of the human body.

2) Spatial Graph Convolutional Neural Network: In this study, we use $G(V, E)$ to graphically represent the hand skeletal structure, where V and E are the node set and the edge set of graph G, respectively. In the graph, the node set $V = \{v_{ti}|t = 1, \ldots, T, i = 1, \ldots, N\}$ contains all joints in the skeleton sequence, where T is the length of the sequence and $N = 19$ is the number of nodes. The framework of SGCN can be written as

$$Y = \sigma((D^{-1/2}AD^{-1/2})XW), \qquad (1)$$

where Y is the output matrix, $X \in \mathbb{R}^{19 \times L}$ is the input matrix, L is the input sequence length, $A \in \mathbb{R}^{19 \times 19}$ is the adjacency matrix, $D \in \mathbb{R}^{19 \times 19}$ is the degree matrix, $W \in \mathbb{R}^{19 \times 19}$ is the parameter matrix, and $\sigma(\cdot)$ is the nonlinear activation function.

3) Spatial Configuration Partitioning: When a person is performing hand movements, his/her finger joints play different roles in the hand motor assessment. Therefore, the adjacency matrix A is dismantled into several different matrices $A_n \in \mathbb{R}^{19 \times 19}$. Thus, (1) can be transformed into the following equation:

$$Y = \sigma(\Sigma((D_n^{-1/2}(K_nA_n)D_n^{-1/2})XW_n)), \qquad (2)$$

where Y is the output matrix, $X \in \mathbb{R}^{19 \times L}$ is the input matrix, $D_n \in \mathbb{R}^{19 \times 19}$ is the degree matrix corresponding to A_n, K_n is a learnable parameter, K_nA_n is the attentional mechanism, and $W_n \in \mathbb{R}^{19 \times 19}$ is the parameter matrix. To determine the value of n, we use the DeepWalk method to analyze the structure of the graph. DeepWalk proposed in [27] is one classical method of graph embedding. The graph embedding
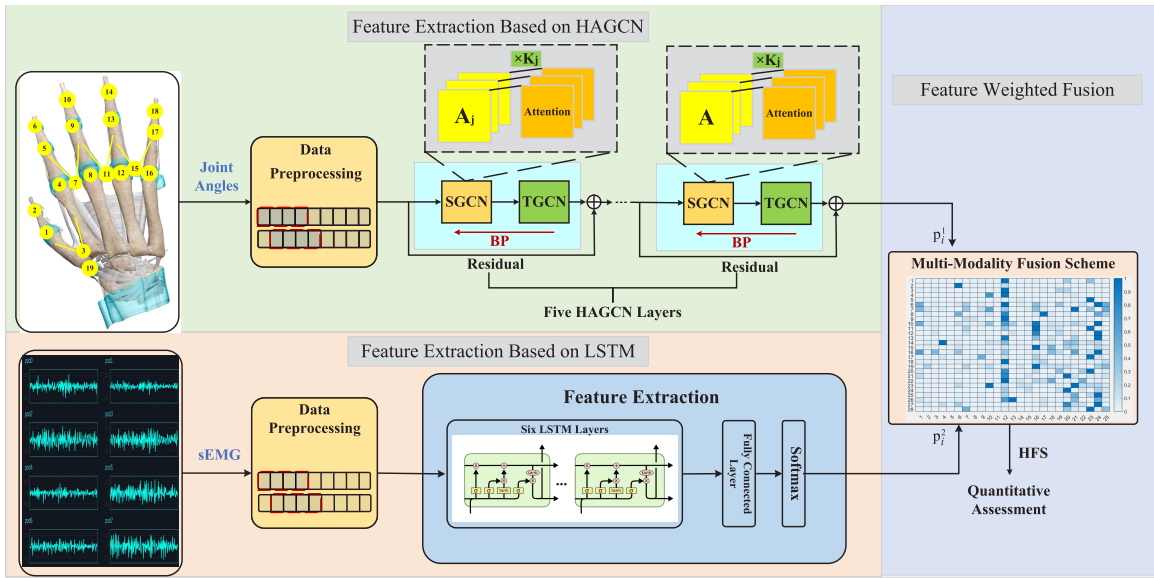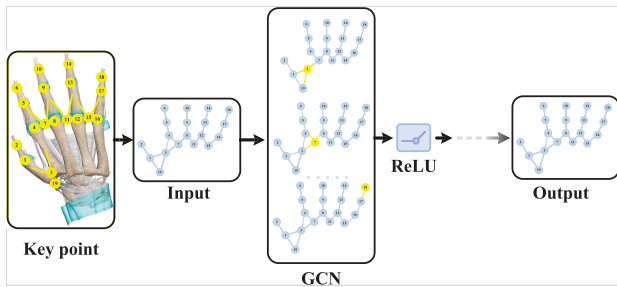
Fig. 4.   The overall framework of the study.

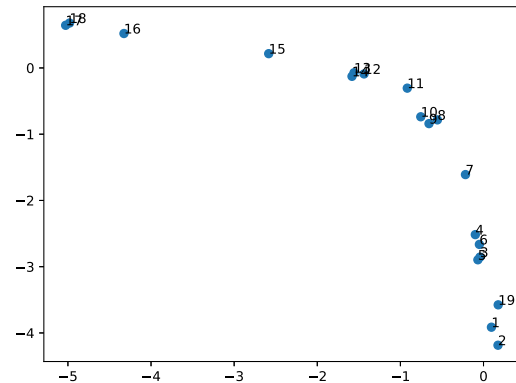Fig. 5.   The structure of the spatial graph convolution network.

Fig. 6.   The result of the DeepWalk.

Fig. 7.   The structure of the matrices A_0, A_1 and A_2.

algorithm represents the nodes in the graph as low-dimensional and dense vectors so that the similar nodes in the original graph are also similar in the low-dimensional representation space. The DeepWalk algorithm mainly includes two steps: the first step is to sample the node sequence by a random walk, and the second step is to use the skip-gram model to learn the representation vector. The 2-dimensional figure after the process of DeepWalk is shown in Fig. 6. The figure shows that the finger joints are symmetrical, and node 11 is located at the center symmetrical position. Therefore, the matrix A can be decomposed according to the symmetry of the graph. Considering the complexity of the calculation, we set n to be 2 or 3. The ablation test demonstrates that the classification accuracy when $n = 3$ is better than the one when $n = 2$ (*i.e.*, the centripetal group and centrifugal group mentioned below are combined into the same group), so n is set to be 3. We design the strategy to divide the neighbor set into three subsets, corresponding to A_0, A_1, and A_2 in Fig. 7. Three subsets are:

- Adjacency matrix group: the adjacency matrix A removes the centripetal group and the centrifugal group.
- Centripetal group: the neighboring nodes far from the gravity center (node 11), such as (8, 7), (11, 8), (8, 9), (11, 12), (12, 13), (12, 15).

- Centrifugal group: the neighboring nodes close to the gravity center (node 11), Such as (7, 8), (8, 11), (9, 8), (12, 11), (13, 12), (15, 12).

Matrix A_n is obtained according to this strategy. Figure 7 graphically shows the matrices A_0, A_1 and A_2. Note that the coordinates in Fig. 7 start from (0,0) and the key point in Fig. 5 starts at 1, then the point 0 in Fig. 7 is equivalent to point 1 in Fig. 5.

4) Temporal Graph Convolution Network: After constructing the SGCN, the task of modeling the TGCN within the skeleton sequence is performed. The process allows us to define a

Fig. 8. The structure of the temporal graph convolution network.

simple strategy for extending the SGCN to the spatial-temporal domain. The temporal graph is constructed by connecting the same joints in a continuous frame, as shown in Fig. 8. The three green points in Fig. 8 represent the three consecutive data points of the node in time, and the yellow and blue points are the adjacent nodes of the green joints. The nine points in the red grid are convolved in the spatial-temporal domain to obtain the red point. In this study, the temporal kernel size is set to be 3.

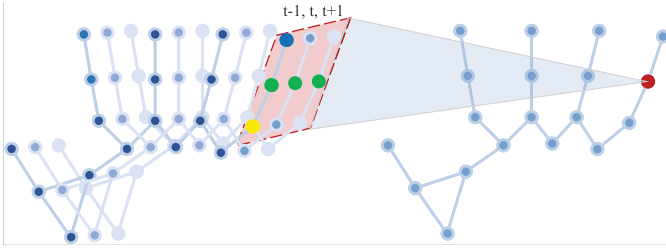*5) Network Architecture and Training:* One SGCN and one TGCN form the HAGCN. The whole network consists of five HAGCN layers. The output sizes of five HAGCN layers are 8, 16, 32, 64, and 32, respectively, and the input size is 19. The network is optimized by the residual connection. The whole model is trained in an end-to-end manner with backpropagation.

The division of the dataset adopts the method of leave-one-out (LOO) cross-validation, which is also known as the 6-fold cross-validation. Each movement has six trials, as shown in Fig. 3. Five trials are taken as the training set, and the data of the remaining trial are taken as the test set to test the classification result. In this way, six tests are performed with different validation sets. The average value of six tests represents the classification accuracy of the action. A total of 26 classification tasks for each of the 28 movements need to be completed.

C. sEMG Analysis Based on LSTM

In this study, a multi-layer LSTM network is introduced to extract the deep features of sEMG signals, which improves the generalization and robustness of the model compared with the manual feature extraction [25]. The network includes six LSTM layers and one fully connected layer. The input size of the first LSTM layer is eight. The data within each time window of sEMG are the input of the LSTM. The output size of each LSTM layer is 32, 64, 128, 64, 32, and 32. The output size of the fully connected layer is 26, which is activated by the softmax algorithm. Sparse categorical cross-entropy is used as the loss function. The Adams algorithm is used as the optimization method for the network training. The LSTM network training still employs the LOO method.

D. Multi-Modality Fusion Scheme

The proposed multi-modality fusion algorithm is given in Algorithm 1. To better understand this algorithm, we briefly explain its working principle and setting.

Algorithm 1 Multi-Modality Fusion Algorithm

Input: Hand kinematic data and sEMG data from 25 patients and 35 healthy subjects

Output: Hand function score

1 $H_i^k \leftarrow$ train HAGCN;
2 $L_i^k \leftarrow$ train LSTM;
3 **for** $m = 1,2$ **do**
4   **for** $i = 1,...,28$ **do**
5   **for** $k = 1,...,6$ **do**
6   **if** m == 1:
7             $p_{m,i,j}^k \leftarrow$ softmax$(O(H_{i,j}^k))$;
8   **else** :
9             $p_{m,i,j}^k \leftarrow$ softmax$(O(L_{i,j}^k))$;
10   **end**
11   **end**
12     $H_m^k \leftarrow$ concatenate$(p_{m,i,j}^k)^*$;
13     $\hat{H}_m \leftarrow \sum_k (\text{norm}(H_m^k))$;
14     $s_m \leftarrow \sum \text{column}(\hat{H}_m)$
15 **end**
16 $s \leftarrow c_1 \times s_1 + (c - c_1) \times s_2$

The softmax function is often used as the last activation function of a neural network to normalize the output of a network to a probability distribution [28], which can be used as a feature to achieve the satisfactory classification tasks [29]. In this study, the number of softmax outputs is 26, and the final output represents the probability value of the input sample being recognized as a healthy person by the neural network. The patient's hand motor function is assessed by calculating and comparing the average probability value of each type of the patient's hand movement to be recognized as a healthy subject's movement. The assessment consists of the kinematic-based assessment and the sEMG-based assessment. We use the decision fusion approach for the multi-modality assessment. The total score is the weighted sum of the kinematic modality score and the sEMG modality score.

The details of Algorithm 1 are as follows. After training HAGCN and LSTM networks, 336 ($2 \times 28 \times 6 = 336$) trained neural networks are obtained, which are written as H_i^k and L_i^k, $i = 1, \ldots, 28, k = 1, \ldots, 6$, respectively. Here, i represents the i-th movement, and k represents the k-th iteration of the 6-fold cross-validation. Then, all stroke patients' kinematic data (including the training set and the test set) are input to the HAGCN, and the output of the HAGCN is $O(H_{i,j}^k) \in \mathbb{R}^{1 \times 26}$, where $j = 1, \ldots, 25$ indicates the order of the stroke patient. Similarly, all stroke patients' sEMG data are input into the LSTM network, which outputs $O(L_{i,j}^k) \in \mathbb{R}^{1 \times 26}$. Next, $O(H_{i,j}^k)$ and $O(L_{i,j}^k)$ are input to the layer of softmax, and then generates $P_{m,i,j}^k \in \mathbb{R}^{1 \times 26}$, where $m = 1$ refers to the output of $O(H_{i,j}^k)$, and $m = 2$ refers to the output of $O(L_{i,j}^k)$. The last element of $P_{m,i,j}^k$ represents the similarity between the i-th movement of the j-th patient and the movement of a healthy subject. We use $(p_{m,i,j}^k)^*$ to represent the last element of $p_{m,i,j}^k$ and use $(p_{m,i,j}^k)^*$ to construct matrix $H_m^k \in \mathbb{R}^{28 \times 25}$.
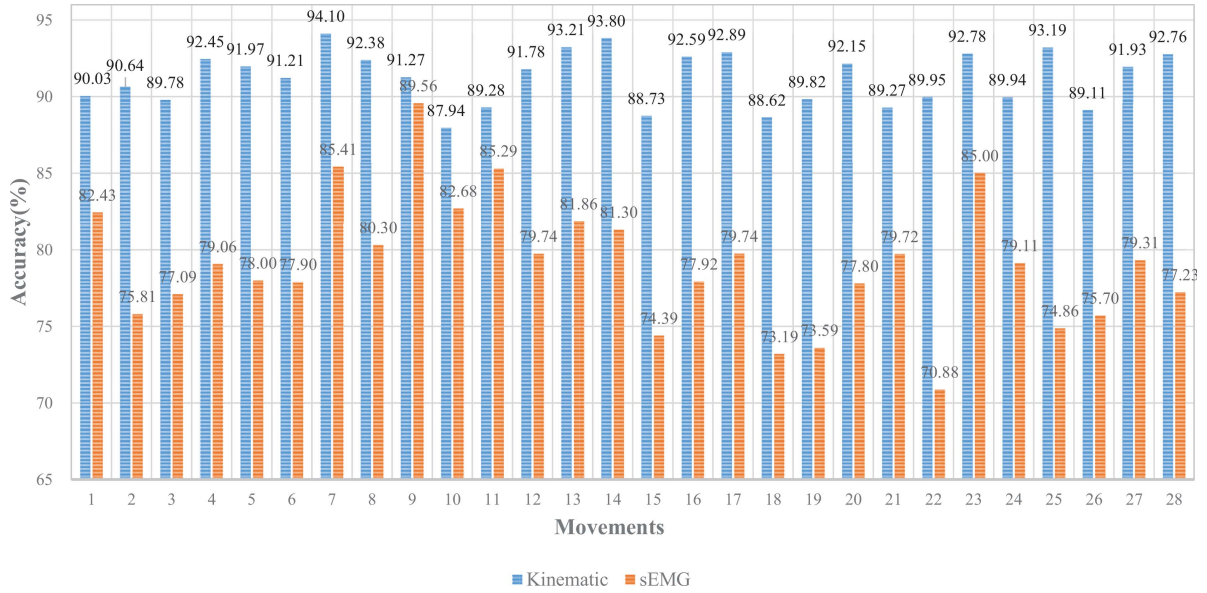
Fig. 9. The classification accuracy by the kinematic information based on HAGCN and the classification accuracy by the sEMG based on LSTM.

Here, $(p_{m,i,j}^k)^*$ is the i-th row and j-th column element of H_m^k. Then, each matrix is normalized and they are added to obtain $\hat{H}_1$ and $\hat{H}_2$, respectively. $\hat{H}_1$ and $\hat{H}_2$ represent the scoring matrices of kinematic signals and sEMG signals versus healthy people. Each column in the matrix represents a patient, and each row represents a movement. Therefore, by calculating the sum of each matrix's column separately, the patient's movement score is obtained. There are two types of scores for each patient: the kinematic score s_1 and the sEMG score s_2. Finally, s_1 and s_2 are fused according to (3) to obtain the final score s.

$$s = c_1 \times s_1 + (c - c_1) \times s_2, \qquad (3)$$

where c_1 ranges from 0 to $c = 10$. Equation (3) belongs to the weighted decision fusion strategy.

IV. RESULTS

A. Data Collection

Thirty-five stroke patients and thirty-five healthy subjects participated in collecting trial data. We design a visual interface that displays the movement of the hand in real-time to ensure the patients to follow the guide. We also evaluate the effect of experimental factors on the range of the joint angles. Factors affecting the joint angles include the individual joint, the movement, the subject, and the movement repetition. The evaluation results show that the data collection quality is reliable.

B. Classification Results

After 6-fold cross verification, the classification accuracy is shown in Fig. 9. The blue histogram represents the classification accuracy by the kinematics, with an average accuracy of 91.2%. The red histogram shows the classification accuracy by the sEMG signal, with an average accuracy of 79.1%. Classification accuracy shows that we can distinguish patients

with the same Brunnstrom grade or the same FMA score. Table II shows the order of the classification accuracy, and the following facts can be observed:

- In the traditional FMA scale, the diameter of the cylinder is not considered as an affecting factor. In this experiment, three types of cylinder grasping actions with different cylinder diameters are designed (marked by the blue background in Table II). Among them, the classification accuracy of grasping a large diameter cylinder is the highest. This finding shows that the more challenging the movement is, the easier it is to distinguish patients at different levels. This finding is also consistent with the physiological characteristics of stroke patients. They are easy to bend fingers but not easy to stretch fingers.
- The three-finger sphere grasp has the highest classification accuracy among three spherical grip movements designed in this experiment (marked by the red background in Table II). Therefore, this movement can be used as a representative action of the spherical grasp, which is consistent with our previous kinematic analysis results [30].
- Among the top 6 movements with the highest classification accuracy based on two kinds of signals, the same movements are: (1) the abduction of all fingers and (2) the lateral grasp. This finding suggests that the hand extension and thumb movement are more effective in identifying the patient's hand movement level.

C. Performance of Quantitative Assessment

To prove the validity of the proposed quantitative assessment, we need to select a performance metric. Generally, the Pearson correlation coefficient (PC) is the most commonly used metric to prove the correlation between two datasets. However, PC can only be applied if the sample is normally distributed. Therefore, the first step is to verify whether the sample follows a normal distribution. Considering that the

TABLE II
ORDER OF THE CLASSIFICATION ACCURACY

| Type	Kinematic	sEMG
Average accuracy	91.2%	79.1%
Sequence No.	Movements	Movements
1	Abduction of all fingers	Thumb adduction
2	Thumb point to little finger	Abduction of all fingers
3	Little finger extension with closed hand and wrist	Litter finger extension with closed hand
4	Prismatic four fingers grasp	Lateral grasp
5	Large diameter grasp	Index finger extension with closed hand
6	Lateral grasp	Wrist extension
7	Tip pinch	Little finger extension with closed hand and wrist
8	Medium diameter grasp	Thumb point to little finger
9	Wrist supination	Hook fingers
10	Hook fingers	Large diameter grasp
11	Three finger sphere grasp	Index finger extension with closed hand and wrist
12	Wrist pronation	Precision sphere grasp
13	Finger pinch	Finger pinch
14	Index finger extension with closed hand and wrist	Writing tripod grasp
15	Thumb adduction	Wrist pronation
16	Fingers flexed together in fist	Wrist supination
17	Ulnar deviation	Medium diameter grasp
18	Wrist extension	Fingers flexed together in fist
19	Power disk grasp	Three finger sphere grasp
20	Writing tripod grasp	Tip pinch
21	Power sphere grasp	Radial deviation
22	Radial deviation	Ulnar deviation
23	Litter finger extension with closed hand	Stick grasp
24	Precision sphere grasp	Prismatic four fingers grasp
25	Stick grasp	Small diameter grasp
26	Small diameter grasp	Power sphere grasp
27	Fixed hook grasp	Fixed hook grasp
28	Index finger extension with closed hand	Power disk grasp
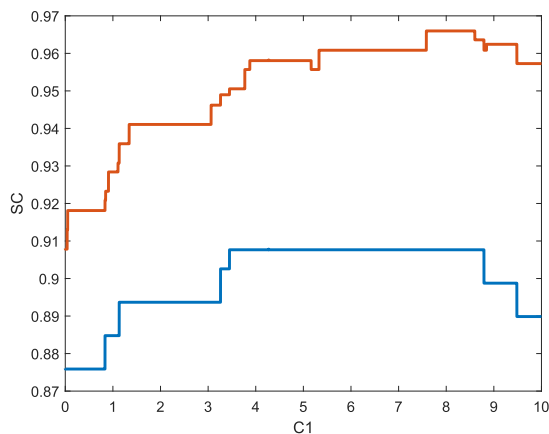
Fig. 10. SC under different values of c_1 (the red line is the SC based on FS and the blue line is the SC based on BS).

TABLE III
SC BETWEEN THE ASSESSMENT RESULT BASED ON THE SINGLE
MODALITY AND THE TRADITIONAL ASSESSMENT

Relation	BS&Kinematic	BS&sEMG	FS&Kinematic	FS&sEMG
SC	0.890	0.876	0.956	0.918

[*] BS refers to Brunnstom score.
[*] FS refers to FMA score.

results based on the kinematic signal are more consistent with two traditional methods.

Figure 10 shows the changing trend of SC under different c_1 values, and the resolution of c_1 is 0.001. By Fig. 10, when c_1 equals to 8, SC reaches the maximum value. Table IV shows the SC of the fusion scheme when c_1 is an integer. As noted in the table, when c_1 equals 8, SC reaches the maximum value, which is higher than that based on one modality. This finding suggests that the fused assessment results are closer to the traditional assessment results, as shown in Fig. 11. More detailed assessment results are provided in Table V. From Table V, the SC between FUS and Order 1 is 0.997, indicating that the multi-modality fusion assessment results are significantly correlated with the therapist's refined judgement on the patient's hand recovery level and can avoid the "ceiling effect" in some traditional scales.

V. DISCUSSION

The main objective of this study is to develop a hand motor function assessment system for the quantitative analysis

number of samples is less than 50, the Shapiro-Wilk test is used to assess whether the samples followed a normal distribution. As a result, the p-values of the Shapiro-Wilk test results of the Brunnstrom score and FMA score are both less than 0.05, indicating that the sample does not exhibit the normal distribution. Therefore, Spearman's rank correlation coefficient (SC) is selected as the metric. In general, an SC greater than 0.8 indicates a strong correlation.

Table III presents the SC of the quantitative assessment based on the single modality. The quantitative evaluation
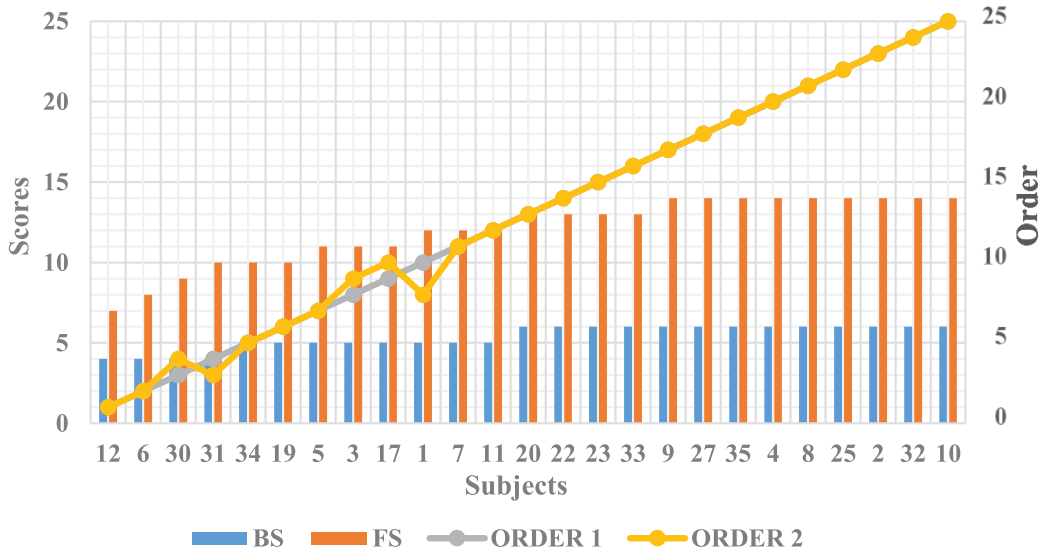
Fig. 11. Scores of traditional scales and the score obtained by the proposed algorithm (BS, FS and ORDER 1 are Brunnstrom scores, FMA scores, and the order of patients assessed by the therapists, respectively; ORDER 2 is the order of patients assessed in this study. The x-axis represents the patient's sequence number.).

TABLE IV

SC BETWEEN THE ASSESSMENT RESULT BASED ON THE MULTI-MODALITY FUSION AND THE TRADITIONAL ASSESSMENT

FUS & Brunnstrom Score		FUS & FMA Score	
c_1	SC	c_1	SC
$c_1 = 1$	0.885	$c_1 = 1$	0.928
$c_1 = 2$	0.894	$c_1 = 2$	0.941
$c_1 = 3$	0.894	$c_1 = 3$	0.941
$c_1 = 4$	0.908	$c_1 = 4$	0.958
$c_1 = 5$	0.908	$c_1 = 5$	0.958
$c_1 = 6$	0.908	$c_1 = 6$	0.961
$c_1 = 7$	0.908	$c_1 = 7$	0.961
$c_1 = 8$	0.908	$c_1 = 8$	0.967
$c_1 = 9$	0.889	$c_1 = 9$	0.962

* FUS refers to fusion score.

TABLE V

DETAILS ON THE ASSESSMENT RESULTS

Subjects	BS	FS	Kinematic Score	sEMG Score	Fusion Score	Order 1	Order 2
12	4	7	1.02	1.05	10.26	1	1
6	4	8	1.01	1.59	11.26	2	2
30	4	9	1.13	2.68	14.40	3	4
31	4	10	1.25	1.76	13.52	4	3
34	5	10	1.51	2.48	17.04	5	5
19	5	10	2.21	3.98	25.64	6	6
5	5	11	3.29	3.88	34.08	7	7
3	5	11	4.03	5.32	42.88	8	9
17	5	11	5.02	3.63	47.42	9	10
1	5	12	3.54	5.21	38.74	10	8
7	5	12	4.54	6.58	49.48	11	11
11	5	12	5.61	4.32	53.52	12	12
20	6	13	4.57	11.89	60.34	13	13
22	6	13	8.26	4.78	75.64	14	14
23	6	13	8.15	9.22	83.64	15	15
33	6	13	7.54	12.32	84.96	16	16
9	6	14	8.42	9.56	86.48	17	17
27	6	14	10.03	9.48	99.20	18	18
35	6	14	8.97	15.20	102.16	19	19
4	6	14	10.51	11.86	107.80	20	20
8	6	14	11.68	12.37	118.18	21	21
25	6	14	12.15	11.86	130.55	22	22
2	6	14	13.41	13.92	135.12	23	23
32	6	14	15.56	16.70	157.88	24	24
10	6	14	21.66	38.91	251.10	25	25

* BS refers to Brunnstom score.
* FS refers to FMA score.

of motor impairment in patients with post-stroke hemiplegia. The system is constructed based on the kinematic data and the sEMG signals which are collected synchronously during 28 well-designed hand movements. Under the framework of multi-modality fusion, the quantitative evaluation results of different modalities are well weighted and integrated, which results in a comprehensive assessment of the hand motor function.

The proposed HAGCN can achieve a classification accuracy of 91.2% based on the kinematic modality. Although, in the sEMG modality, the LSTM network could only achieve a classification accuracy of 79.1%. By exploiting the complementarity between motor characteristics of different modalities, the obtained fused results show that the clinical relevance can be enhanced by fusing the multi-modal information. It should be noted that the classification accuracy solely based on the sEMG is relatively low. Because the sEMG signal can only reflect the neuromuscular activity to certain extents, the classification accuracy using sEMG signal is naturally

inferior to the one using the motion information. However, it is reasonable to keep the sEMG modality in the assessment system. This is because if the motion information is collected by some non-contact devices like Leap Motion (Leap Motion's price is much lower than the data glove used in this paper), the hand movement measurement is susceptible to illumination and occlusion. This can cause the measurement missing problem, while the sEMG signal can be stably measured by the wearable bracelet. More importantly, the muscle strength is also crucial for the rehabilitation assessment. In Brunnstrom

and FMA, there are also some items reflecting the muscle strength. The muscle strength can be well estimated by the sEMG [31]. Therefore, from the perspective of the assessment scalability, it is necessary to keep the sEMG modality.

In this study, as shown in Table V, each patient can be given a specific score by the proposed assessment algorithm rather than a rough grade, and even patients with the same scale can be distinguished, avoiding the "ceiling effect" of the traditional scales.

VI. CONCLUSION

In this study, we propose a multi-modality (kinematics and sEMG) fusion assessment framework based on HAGCN and LSTM, and apply this framework to the self-collected dataset to quantitatively evaluate the rehabilitation levels of 25 stroke patients. The SCs between the assessment results of this study and the traditional scales (Brunnstrom scale and Fugl-Meyer assessment scale) are 0.908 and 0.967, respectively, providing a significant correlation between the proposed assessment and the traditional scale scores. In addition, the SC value between the score of this study and the refined rehabilitation level of patients with the same grade is 0.997, suggesting that the quantitative assessment of 25 stroke patients can avoid the "ceiling effect" of traditional scales to some extends.

REFERENCES

[1] Y. Tomita, M. R. M. Rodrigues, and M. F. Levin, "Upper limb coordination in individuals with stroke: Poorly defined and poorly quantified," *Neurorehabil. Neural Repair*, vol. 31, nos. 10–11, pp. 885–897, Nov. 2017.

[2] N. Sun, G. Li, and L. Cheng, "Design and validation of a self-aligning index finger exoskeleton for post-stroke rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1513–1523, 2021.

[3] M. C. Sara, L. Roberto, B. Adrián, S. M. Bárbara, and S. A. Pilar, "Validity, reliability, and sensitivity to motor impairment severity of a multi-touch app designed to assess hand mobility, coordination, and function after stroke," *J. Neuroeng. Rehabil.*, vol. 18, pp. 1–15, Dec. 2021.

[4] C. E. Lang, M. D. Bland, R. R. Bailey, S. Y. Schaefer, and R. L. Birkenmeier, "Assessment of upper extremity impairment, function, and activity after stroke: Foundations for clinical decision making," *J. Hand Therapy*, vol. 26, no. 2, pp. 105–115, Sep. 2012.

[5] K. A. Sawner, J. M. LaVigne, and S. Brunnstrom, *Brunnstrom's Movement Therapy Hemiplegia: A Neurophysiological Approach*, 2nd ed. Philadelphia, PA, USA: Lippincott, 1992.

[6] A. R. Fugl-Meyer, L. Jääskö, I. Leyman, S. Olsson, and S. Steglind, "The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance," *Scand. J. Rehabil. Med.*, vol. 7, no. 1, pp. 13–31, 1975.

[7] H. H. Lee *et al.*, "Revisiting the proportional recovery model in view of the ceiling effect of Fugl–Meyer assessment," *Stroke*, vol. 52, no. 10, pp. 3167–3175, Oct. 2021.

[8] D. J. Gladstone, C. J. Danells, and S. E. Black, "The fugl-meyer assessment of motor recovery after stroke: A critical review of its measurement properties," *Neurorehabil. Neural Repair*, vol. 16, no. 3, pp. 232–240, Sep. 2002.

[9] Q. Fang, S. S. Mahmoud, X. Gu, and J. Fu, "A novel multistandard compliant hand function assessment method using an infrared imaging device," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 2, pp. 758–765, Mar. 2019.

[10] T. Hamaguchi *et al.*, "Support vector machine-based classifier for the assessment of finger movement of stroke patients undergoing rehabilitation," *J. Med. Biol. Eng.*, vol. 40, no. 1, pp. 91–100, Feb. 2020.

[11] C. Li, L. Cheng, H. Yang, Y. Zou, and F. Huang, "An automatic rehabilitation assessment system for hand function based on leap motion and ensemble learning," *Cybern. Syst.*, vol. 52, no. 1, pp. 3–25, Oct. 2020.

[12] X. Song, S. Chen, J. Jia, and P. B. Shull, "Cellphone-based automated Fugl–Meyer assessment to evaluate upper extremity motor function after stroke," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 2186–2195, Oct. 2019.

[13] R. J. Adams, M. D. Lichter, E. T. Krepkovich, A. Ellington, M. White, and P. T. Diamond, "Assessing upper extremity motor function in practice of virtual activities of daily living," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 2, pp. 287–296, Mar. 2015.

[14] R. J. Adams *et al.*, "Virtual activities of daily living for recovery of upper extremity motor function," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 1, pp. 252–260, Jan. 2018.

[15] I. Spinelli, S. Scardapane, and A. Uncini, "Adaptive propagation graph convolutional network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 10, pp. 4755–4760, Oct. 2021.

[16] X. Xu, T. Wang, Y. Yang, A. Hanjalic, and H. T. Shen, "Radial graph convolutional network for visual question generation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1654–1667, Apr. 2021.

[17] Y. Ren, Y. Shi, K. Zhang, Z. Chen, and Z. Yan, "Medical treatment migration prediction based on GCN via medical insurance data," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 9, pp. 2516–2522, Sep. 2020.

[18] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in *Proc. AAAI Conf. Artif. Intell.*, vol. 32, no. 1, Apr. 2018, pp. 7444–7452.

[19] C. Zhang, X. Chen, S. Cao, X. Zhang, and X. Chen, "A novel HD-sEMG preprocessing method integrating muscle activation heterogeneity analysis and kurtosis-guided filtering for high-accuracy joint force estimation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 9, pp. 1920–1930, Sep. 2019.

[20] N. J. Jarque-Bou, M. Atzori, and H. Müller, "A large calibrated database of hand movements and grasps kinematics," *Sci. Data*, vol. 7, no. 1, pp. 1–10, Jan. 2020.

[21] N. J. Jarque-Bou, M. Vergara, J. L. Sancho-Bru, V. Gracia-Ibáñez, and A. Roda-Sales, "A calibrated database of kinematics and EMG of the forearm and hand during activities of daily living," *Sci. Data*, vol. 6, no. 1, p. 270, Nov. 2019.

[22] M. Atzori *et al.*, "Characterization of a benchmark database for myoelectric movement classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 1, pp. 73–83, Jan. 2015.

[23] M. Atzori *et al.*, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Sci. Data*, vol. 1, no. 1, Dec. 2014, Art. no. 140053.

[24] S. Pizzolato, L. Tagliapietra, M. Cognolato, M. Reggiani, H. Müller, and M. Atzori, "Comparison of six electromyography acquisition setups on hand movement classification tasks," *PLoS ONE*, vol. 12, no. 10, Oct. 2017, Art. no. e0186132.

[25] F. Stival, S. Michieletto, M. Cognolato, E. Pagello, H. Müller, and M. Atzori, "A quantitative taxonomy of human hand grasps," *J. NeuroEng. Rehabil.*, vol. 16, no. 1, pp. 1–17, Feb. 2019.

[26] K. Scrivener, K. Schurr, and C. Sherrington, "Responsiveness of the ten-metre walk test, step test and motor assessment scale in inpatient care after stroke," *BMC Neurol.*, vol. 14, no. 1, p. 129, Dec. 2014.

[27] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2014, pp. 701–710.

[28] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[29] C. Wang, L. Peng, Z.-G. Hou, J. Li, T. Zhang, and J. Zhao, "Quantitative assessment of upper-limb motor function for post-stroke rehabilitation based on motor synergy analysis and multi-modality fusion," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 4, pp. 943–952, Apr. 2020.

[30] C. G. Li, H. J. Yang, S. Y. Sun, and L. Cheng, "Quantitative taxonomy of hand kinematics based on long short-term memory neural network," in *Proc. 28th Int. Conf. Neural Inf. Process.*, Dec. 2021, pp. 221–229.

[31] Y. Zou, L. Cheng, and Z. Li, "A multimodal fusion model for estimating human hand force: Comparing surface electromyography and ultrasound signals," *IEEE Robot. Autom. Mag.*, early access, Jun. 10, 2022, doi: 10.1109/MRA.2022.3177486.