Sequential sEMG Recognition with Knowledge Transfer and Dynamic Graph Network based on Spatio-Temporal Feature Extraction Network

Zhilin Li, Xianghe Chen, Jie Li, Zhongfei Bai, Hongfei Ji, Lingyu Liu, and Lingjing Jin

***Abstract*— Surface electromyography (sEMG) signals are electrical signals released by muscles during movement, which can directly reflect the muscle conditions during various actions. When a series of continuous static actions are connected along the temporal axis, a sequential action is formed, which is more aligned with people's intuitive understanding of real-life movements. The signals acquired during sequential actions are known as sequential sEMG signals, including an additional dimension of sequence, embodying richer features compared to static sEMG signals. However, existing methods show inadequate utilization of the signals' sequential characteristics. Addressing these gaps, this paper introduces the Spatio-Temporal Feature Extraction Network (STFEN), which includes a Sequential Feature Analysis Module based on static-sequential knowledge transfer, and a Spatial Feature Analysis Module based on dynamic graph networks to analyze the internal relationships between the leads. The effectiveness of STFEN is tested on both modified publicly available datasets and on our acquired Arabic Digit Sequential Electromyography (ADSE) dataset. The results show that STFEN outperforms existing models in recognizing sequential sEMG signals. Experiments have confirmed the reliability and wide applicability of STFEN in analyzing complex muscle activities. Furthermore, this work also suggests STFEN's potential benefits in rehabilitation medicine, particularly for stroke recovery, and shows promising future applications.**

***Index Terms*— Surface electromyograph, Sequential feature analysis, Spatio-temporal feature extraction, Dynamic graph network, Knowledge transfer**

This work was supported by Shanghai Shenkang Hospital Development Center's Second Round of "Three-Year Action Plan to Promote Clinical Skills and Clinical Innovation in Municipal Hospitals" Research Physician Innovation Transformation Capability Training Program(SHDC2023CR001).

Zhilin Li, Xianghe Chen, Jie Li and Hongfei Ji are with the Translational Research Center, Shanghai Yangzhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Electronic and Information Engineering, Tongji University, Shanghai, China (e-mail: lizhilin@tongji.edu.cn; 2332032@tongji.edu.cn; nijanice@163.com; jhf@tongji.edu.cn).

Zhongfei Bai, Lingyu Liu, and Lingjing Jin are with the Department of Neurology and Neurological Rehabilitation, Shanghai Disabled Person's Federation Key Laboratory of Intelligent Rehabilitation Assistive Devices and Technologies, Yangzhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Medicine, Tongji University, Shanghai, China (e-mail: bzf567@163.com; happyneurologist@163.com; lingjingjin@163.com).

Lingjing Jin is also with the Neurotoxin Research Center of Key Laboratory of Spine and Spinal Cord Injury Repair and Regeneration of Ministry of Education, Neurological Department of Tongji Hospital, School of Medicine, Tongji University, Shanghai, China

This work was conducted under the supervision of the Ethics Review Committee of Shanghai Yangzhi Rehabilitation Hospital, with the ethical review approval letter number: Yangzhi Lun Shen Zi [2023]036.

I. Introduction

SURFACE Electromyographic signals (sEMG) are widely used for their ability to detect and analyze the electrical activity of muscles in various fields. These signals are acquired from skin-placed surface leads and analyzed through complex procedures, providing valuable information on muscle activity, strength, and coordination.

With technological progress, the use of sEMG has expanded, especially in rehabilitation medicine. It helps create personalized treatment plans by accurately monitoring muscle activity, aiding stroke recovery [1]. Additionally, sEMG is increasingly applied in various fields such as personal identification and entertainment, due to its accuracy in capturing the intent behind movements. Unlike other types of signals, sEMG directly mirrors muscle activity, making it an essential tool for understanding movements and identifying specific motion-related features.

Sequential actions, which are a series of ordered and continuous action patterns, offer a deep dive into human behavior's dynamic nature. They reveal how actions evolve over time and detail the transitions between different states. Studying these actions is key to grasping how movements are performed and controlled, particularly useful in rehabilitation. This understanding aligns with human cognitive processes, helping interpret the intentions behind movements more effectively.

Even though integrating sEMG technology with sequential actions shows a lot of potential, the research on sequential sEMG signals is still in the early stages and faces several challenges:

1) **Focus on Static Action Features in Algorithms:** Most current sEMG signal recognition algorithms concentrate on extracting and classifying features related to static actions. This emphasis results in a lack of representation for features tied to sequential actions within these models, and a gap in comprehensive sequential analysis.
2) **Confusion in Public Datasets Over Action Types:** Many sEMG datasets do not distinguish clearly between static and sequential signals, complicating the extraction of sequential features. This confusion contributes to a lack of datasets specifically designed for sequential actions, resulting in obstacles in the design of algorithms for sequential features right from the first step.

This paper introduces an algorithm designed to extract features from sequential signals, filling in the shortcomings of traditional static sEMG signal analysis methods in sequential sEMG feature exploration. The contributions of this paper are summarized below.

1) We set up an experiment for drawing Arabic numerals in the air to create the ADSE dataset, tackling the lack of sequential sEMG datasets.
2) We develop the STFEN model specifically for sequential sEMG signals, which includes a knowledge transfer module for time feature extraction and an adaptive dynamic graph neural network for spatial feature extraction.
3) We use the ADSE dataset and modified publicly available datasets to compare the performance of STFEN with other models and determine the excellent performance of STFEN in sequential sEMG signal processing.
4) We analyze the effectiveness of STFEN by visualizing intermediate features and explore the potential applications of sequential sEMG signal analysis in rehabilitation strategies.

II. Related Works

A. Analysis of Sequential sEMG Signals

Since Marey [2] first recorded surface muscle electrical signals in 1890, research into sEMG has steadily advanced. Sequential sEMG signals, which are recorded during human sequential actions and represent long time patterns, give a deeper time-based perspective than traditional static sEMG signals. They capture the changing sequential features over time, shedding light on the dynamic patterns of muscle activity. There is already a body of research concentrating on these sequential sEMG signals.

Chen *et al.* [3] developed a model that uses sequential sEMG to accurately predict the three-dimensional movements of the upper limb, aimed at rehabilitation robotics. Pan *et al.* [4] compared musculoskeletal models with data-driven approaches such as linear regression and classical artificial neural networks to predict sequential wrist and hand actions in sEMG interfaces. Liu *et al.* [5] worked on continuously decoding actions of the shoulder, elbow, and wrist from sequential sEMG, showing great promise for myoelectric-controlled exoskeletons used in upper limb rehabilitation.

Even with these advancements, there's a significant difference between these algorithms and those specifically designed for sequential sEMG, affecting the ability to identify sequential features effectively. To tackle this, we propose two considerations:

- Sequential signals essentially consist of static signals linked over time, indicating a strong similarity between two types of sEMG signals. We aim to use this similarity by applying knowledge transfer from the analysis of static actions to improve the recognition of sequential actions.
- Most current algorithms for sequential sEMG use Convolutional Neural Networks (CNNs), which do not specifically target the relationship between different leads. Considering the important spatial relationships in sequential actions, we suggest using graph networks to more accurately capture these relationships in sequential sEMG data.

B. Transfer Learning in sEMG Signal Analysis

Transfer learning is a key method in machine learning that improves a model's performance and its ability to be applied to new tasks by using knowledge gained from one task on another. This method has been widely used in fields such as natural language processing and image recognition. In the study of sEMG signals, transfer learning is becoming important for addressing issues like limited data and the need for models that perform well across different tasks. For example, Chen *et al.* [6] showed how transfer learning could enhance the accuracy of hand gesture recognition from sEMG data, leading to better results with less training on varied datasets. Ameri *et al.* [7] developed a CNN-based transfer learning approach that effectively handles the challenge of electrode displacement in sEMG recognition tasks. Rezaee *et al.* [8] successfully used deep transfer learning to detect Parkinson's disease using sEMG, with an accuracy rate of over 99%.

C. Graph Networks in sEMG Signal Analysis

Graph Neural Networks (GNNs) have become a powerful type of deep learning used to work with data that has a graph structure, like social media, traffic systems, and biology research. Recently, researchers have started using GNNs to analyze sEMG signals. For instance, Massa *et al.* [9] applied GNNs to high-density sEMG data to recognize action intentions, reaching a low classification error rate of 8.75% over 65 gestures. Vijayvargiya *et al.* [10] used a GNN based on Pearson correlation for analyzing lower limb sEMG signals and achieved a high accuracy of up to 99.36% in recognizing different activities. Li *et al.* [11] enhanced the performance of action prediction by integrating a Graph Isomorphic Network with models that learn over sequences, for both Electroencephalography and sEMG, attaining an accuracy of 88.89%.

These studies highlight how well GNNs model the connections between sEMG sensors, performing better than traditional methods. However, GNNs sometimes miss the changing interactions during actions because they usually work with fixed graph structures. We suggest using a dynamic graph structure, which can adjust the graph as needed. Dynamic graph networks are already helping improve recommendations, predict traffic, model 3D shapes, and other fields. For sEMG analysis, Lee *et al.* [12] developed a network that updates its adjacency based on what it learns from stretchable sEMG leads, managing to recognize actions with 95% accuracy across 18 different gestures.

III. Signal Acquisition

A. Acquisition Device

To acquire the sequential sEMG signal, we use acquisition leads shown as Fig. 1(a). Signal acquisition and annotation are conducted via a PC-based program, following the

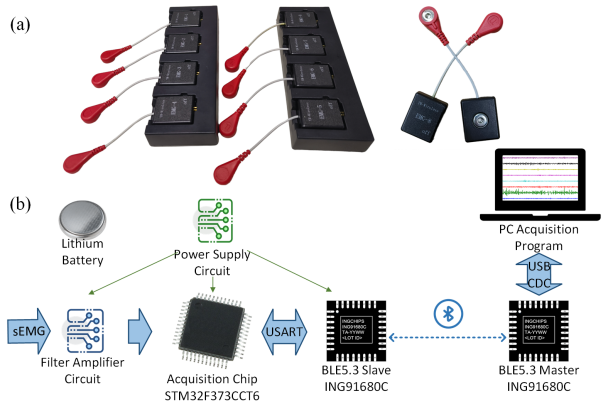

Fig. 1. (a) Experimental Acquisition Equipment: Comprising 8 circuits, it simultaneously acquires sEMG signals from 8 distinct muscles. (b) Signal Acquisition Process Module.

workflow in Fig. 1(b). Our system includes an acquisition module, transmission module, and power supply. The acquisition module initially enhances and filters the sEMG signal, followed by analog-to-digital (AD) conversion using the STM32F373CCT6 chip. The transmission module, centered on the BLE 5.3 ING91680C Bluetooth chip, connects to the PC and acquisition module, enabling signal transmission from eight muscle locations. The power supply module powers the whole device.

B. Paradigm Design

Our paradigm focuses on sequential sEMG signals by evaluating action sequence representation, subject familiarity, and action discrimination. We use an experimental setup involving drawing Arabic numerals 0-9, each for 5 seconds as shown in Fig. 2(a). As Arabic numerals are ubiquitous symbols in everyday life, it's unnecessary to provide special training to the participants. Additionally, the process of drawing these numerals naturally creates a specific sequence of actions, which fits the goals of our study well. We also acquire static sEMG signals with subjects' hands in five positions (center, up, down, left, right) for 20 seconds as shown in Fig. 2(b). These static signals help train the initial model in transfer learning approach, improving the model's understanding of

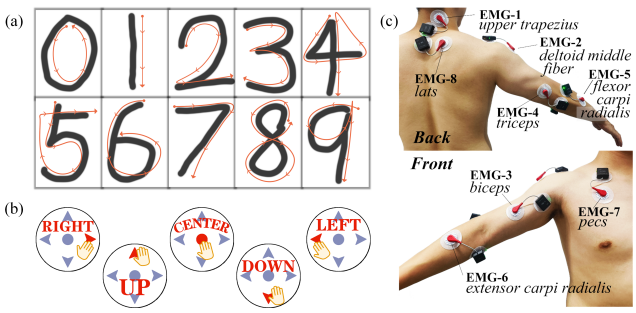

Fig. 2. (a) Sequential Action Paradigm: The upper limbs draw 10 Arabic digits from 0 to 9 in the air. (b) Static Action Paradigm: The upper limbs are located in five positions in front of the body: the center, up, down, left, and right. (c) Schematic Diagram of 8 Muscle Groups Pasting Positions of the Corresponding Acquisition Leads.

feature extraction. For muscle selection, focusing on upper limb movements, we identify eight key muscle groups for sequential sEMG signal collection: upper trapezius, deltoid middle fiber, biceps, triceps, flexor carpi radialis, extensor carpi radialis, pecs, and lats, with corresponding lead placements shown in Fig. 2(c).

C. Acquisition of ADSE Dataset

Following this paradigm, we acquired data from 10 male adults aged between 22 and 28, each contributing 100 sequential sEMG sequences and 10 static signal sets. We also measured the distances between the leads to analyze the relationships between different lead locations. This effort resulted in the Arabic Digit Sequential Electromyography (ADSE) Dataset, available at http://ieee-dataport.org/12775.

IV. Spatio-Temporal Feature Extraction Network

A. Basic Framework of STFEN

In our study, we develop the Spatio-Temporal Feature Extraction Network (STFEN) shown as Fig. 3 to effectively process sequential sEMG signals. STFEN combines two primary modules: a temporal feature extraction module that utilizes knowledge transfer, and a spatial feature extraction module that employs both dynamic and static graph networks. These modules work in parallel to extract temporal and spatial features from sequential sEMG signals. The two features are jointly used for the classification of the sequential signal after concatenating.

The temporal feature extraction starts by training a static action recognition network based on static sEMG datasets. STFEN acts as a knowledge base for transferring insights to sequential sEMG signal recognition. The sequential sEMG signals are first partitioned along the temporal axis into multiple 'static sEMG atoms'. Then, recognizing each atom's category probability for specific static actions allows for the creation of a feature sequence that captures the dynamic changes in sEMG characteristics through the sequential action.

The spatial feature extraction module prepares the sEMG sequences by partitioning them into time windows and converting these into frequency band energy vectors. A graph neural network, informed by inter-lead correlations, processes these vectors. STFEN utilizes two kinds of adjacency graphs: one static, based on inter-lead distances, and another dynamic, formed based on signal similarity. This approach enables a detailed spatial feature analysis, considering both static spatial relationships and dynamic signal characteristics.

After extracting features from both modules, the sequential sEMG signal provides temporal and spatial feature vectors. These vectors are concatenated to form a composite feature for the input sequential sEMG signal. By applying this composite feature in a fully connected neural layer, digit drawing actions based on the sequential sEMG data can be classified.

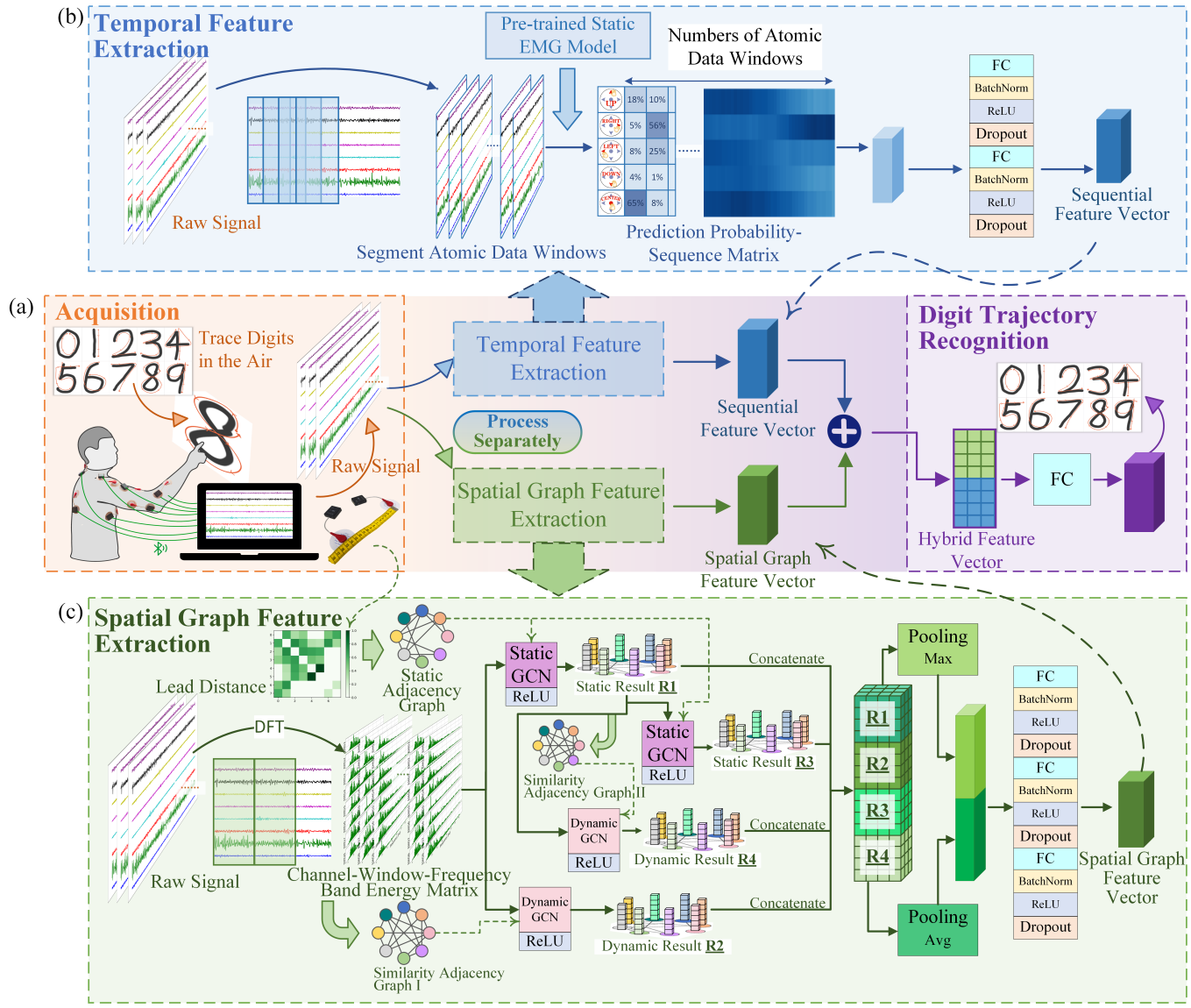

Fig. 3. (a) The Main Process of STFEN: The acquired signals go through the temporal feature extraction module based on "static-sequential" knowledge transfer and the spatial feature extraction module based on adaptive dynamic graph network, and then the temporal features together with spatial features are merged for sequential action recognition. (b) Temporal Feature Extraction Process: It divides the sequential signal into multiple overlapping continuous atomic windows, uses the pre-trained static sEMG recognition model to identify the state features of each window, and then flattens the temporal sequential feature matrix and feature embedding. (c) Spatial Feature Extraction Process: It converts the sequential signal into multiple frequency domain windows, performs spatial feature analysis through a two-layer static graph network and a two-layer dynamic graph network, and then performs feature embedding after merging and pooling.

B. Temporal Feature Extraction Module Based on "Static-Sequential" Knowledge Transfer

Knowledge transfer uses knowledge and features obtained from a source task to improve the performance of target task, enhancing the generalization of the target model, particularly in tasks with intricate features.

In STFEN, the high cost and small size of sequential sEMG data limit the model's adaptability, making it hard to identify common features that apply across different muscle activities. To tackle this, STFEN involves constructing a source model using static sEMG signals. This source model learns universal sEMG feature extraction through extensive static data. Then, through knowledge transfer, STFEN applies the insights from the source model to sequences of static-like sEMG segments in sequential signals, helping to identify their sequence-specific features. Notably, during the construction of the source model, we find that transforming signals into the frequency domain often leads to higher recognition accuracy than directly recognizing them in the time domain.

1) Transformation from Time Domain to Frequency Domain: Frequency band energy, important in signal processing, illustrates energy distributions within specific frequency ranges. Fourier Transform (FT) is as follows:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \tag{1}$$

Through FT, a temporal signal $x(t)$ is converted into its frequency representation $X(f)$, performing the energy at var-

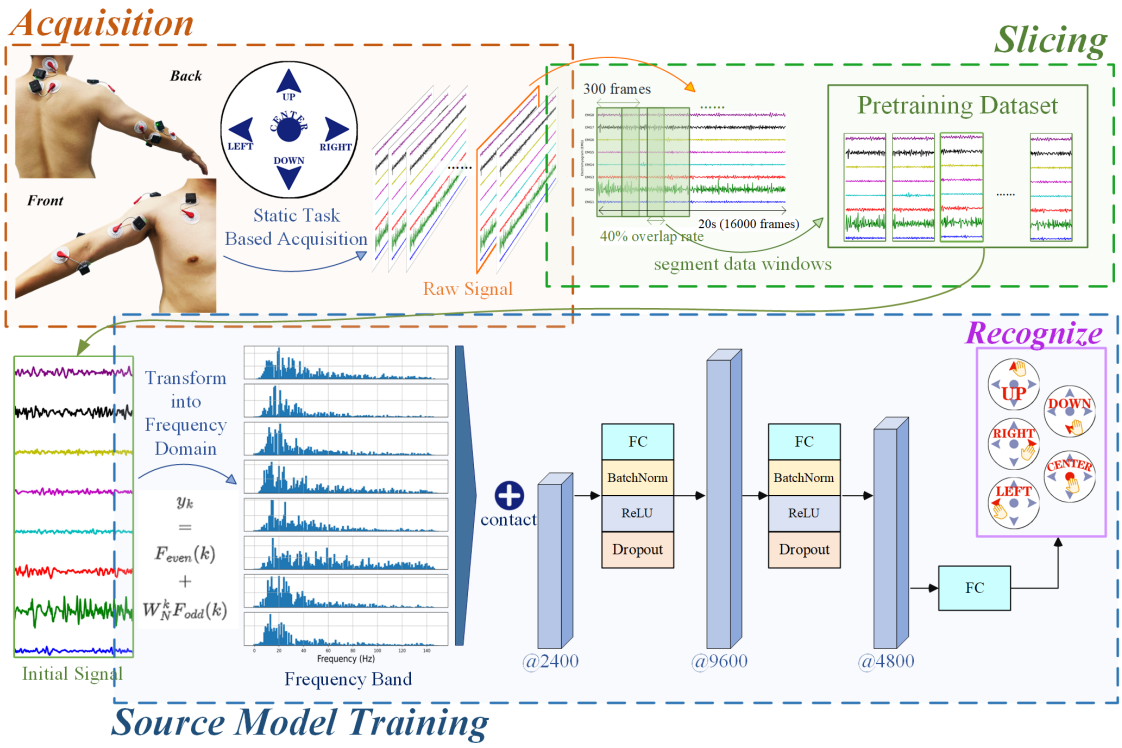

Fig. 4. Training Process of the Static sEMG Source Model: After acquiring the sEMG signals corresponding to the static paradigm actions, the data are augmented through overlapping slices, and then sent to the paradigm recognition network composed of a DFT module and a three-layer fully connected network.

ious frequencies. For sEMG signal analysis, Discrete Fourier Transform (DFT) as 2 is preferred because it deals with discrete voltage values from leads. It calculates frequency band energy y_k for discrete frequencies k from signals $x(n)$ in the time domain.

$$y_k = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi kn/N} \tag{2}$$

Compared to time-domain signals, frequency band energy offers insights into signal periodicity, frequency distribution, and energy levels, enhancing feature distinctiveness and facilitating subsequent models' capabilities in discerning different types of sEMG statuses and signal patterns.

2) Construction of sEMG Source Model: For analyzing static sEMG signals, we partition static signals into 300-frame lengths with a 40% overlap between each window, forming the training set for our source model. As shown in Fig. 4, the source model, based on a multilayer perceptron (MLP), concentrates the DFT-converted frequency band energies from each channel into a single input vector. Through fully connected layers, the MLP classifies static actions by using features from the frequency domain for better signal understanding, calculating probabilities for classifying static actions.

3) Knowledge Transfer from Static Features to Sequential Features: Addressing the temporal complexity of sequential sEMG signals, STFEN adopts a sliding data window method, facilitating the knowledge transfer from static to sequential features while capturing their temporal features. With windows covering 300 data frames and a 60% overlap, each sEMG atom presents a segment of static muscle activity.

Then, the pre-trained static sEMG model analyzes each atom to identify class probabilities for static actions, which are then compiled into a feature matrix representing the sequential signals' temporal evolution. Since each static sEMG atom window matches an individual element within the overall sequential sEMG feature, stacking the class probability vectors for each sEMG atom yields a feature matrix. This matrix is processed through two fully connected layers after being flattened, thereby extracting the temporal feature vector corresponding to the input sequential sEMG signal.

C. Spatial Feature Extraction Module Based on Adaptive Dynamic Graph Network

Compared to traditional deep networks like CNNs, graph neural networks offer a wider perspective by considering the relationships between different leads in signal data, improving the analysis and extraction of features from sEMG signals. This inclusive view greatly enhances the algorithm's generalization capability.

In STFEN, spatial feature extraction begins with partitioning data into windows and transforming these into frequency band energy via DFT. Upon stacking the band energy matrices across each channel, the transferred data are set for further spatial analysis with graph networks.

STFEN uses two kinds of adjacency graphs to model internode relationships: a static graph based on inter-lead distances

and a dynamic graph based on signal similarities. Through layers of graph convolutional networks that process both static and dynamic graph data, STFEN captures comprehensive spatial features from the sEMG signals by concatenating and pooling.

1) Constructing Static Adjacency Matrix Based on Inter-Lead Distance: Constructing a static graph based on inter-lead distances stands as a key component within the spatial feature extraction module of STFEN, which determines the direction of diffusion and aggregation of signal features within the graph. This process involves normalizing inter-lead distances to form a weighted adjacency matrix, which reflects the spatial relationships. Leads positioned closer together are assigned higher weights, indicating stronger interactions. STFEN uses a threshold mechanism in the static adjacency graph to maintain flexibility, and ensure precision in representing spatial relationships, balancing the adjacency based on physical proximity.

The formula for the static adjacency matrix A^{static} is as follows:

$$A_{i,j}^{static} = \begin{cases} 0, & i = j \\ \omega_{max}, & i \neq j \ \& \ d_{i,j} \leq dis_{lower} \\ \omega_{min}, & i \neq \ j \ \& \ d_{i,j} \geq dis_{upper} \\ \omega_{min} + \dfrac{\left(\omega_{max} - \omega_{min}\right)\left(dis_{upper} - d_{i,j}\right)}{dis_{upper} - dis_{lower}}, & otherwise \end{cases} \tag{3}$$

where $A_{i,j}^{static}$ represents the static adjacency weight between leads i and j, $d_{i,j}$ signifies the actual measured inter-lead distance between leads i and j, dis_{upper} and dis_{lower} are respectively set as the maximum and minimum thresholds for skin distances (in STFEN, empirically dis_{upper} is set to $40cm$ and dis_{lower} is set to $5cm$), ω_{max} and ω_{min} are the maximum and minimum threshold values for adjacency weights (in STFEN, empirically ω_{max} is set to 1.0 and ω_{min} is set to 0.2). Notably, during measurements, for very far apart leads like the pecs and the flexor carpi radialis, their adjacency weight is set to 0.

2) Constructing Dynamic Adjacency Matrix Based on Cosine Similarity of Signals: Cosine similarity is a useful method for measuring how similar the directions of two vectors are. It calculates the angle between two vectors in a multi-dimensional space, independent of the vector magnitude. Its applicability spans across various domains such as natural language processing, recommendation systems, and signal processing. The formula to compute cosine similarity is as follows:

$$Similarity(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} \tag{4}$$

where A and B denote two vectors, and $\|\mathbf{A}\|$ and $\|\mathbf{B}\|$ represent the norms of A and B, respectively.

For sEMG signal analysis, cosine similarity is advantageous as it reduces the impact of different muscle intensities on signal amplitudes. It focuses solely on the direction of the sEMG signals from various leads, effectively showing the degree of association between various muscles. The more alike two signals are, the smaller the cosine angle between them, indicating a closer dynamic adjacency between the corresponding leads. Furthermore, the simplicity and efficiency of cosine similarity calculations enable its incorporation into deep learning models, making it easier to quickly build the dynamic adjacency graph.

Within STFEN, the formula for the dynamic adjacency matrix $A^{dynamic}$ is as follows:

$$A_{i,j}^{dynamic} = \begin{cases} 0, & i = j \\ Similarity(\mathbf{S_i}, \mathbf{S_j}), & i \neq j \end{cases} \tag{5}$$

where $A_{i,j}^{dynamic}$ signifies the dynamic adjacency weight between leads i and j, and S_i and S_j are the sEMG signals on leads i and j. By employing cosine similarity between S_i and S_j to construct the dynamic graph, it helps to capture complex associations among different sEMG leads beyond the physical distances represented by static adjacency matrix. This dynamic graph allows the depiction of cooperative muscle associations that might be non-dominant but possess underlying synergistic effects.

3) Extracting Spatial Features Using Graph Convolutional Networks: Graph Convolutional Network (GCN) is a fundamental algorithm that uses convolutional techniques to handle graph-structured data. Its main strategy is to use the information from a node's neighbors to update its own, allowing for the spread of information across the graph. GCNs are adept at learning the connections between nodes in such data, applicable in areas like node classification, graph analysis, and predicting connections.

GCNs consist of several convolutional layers where each one updates the information of its nodes by aggregating information from adjacent nodes, along with diffusing their own features to others, resulting in improved information for each node. The node update process in GCNs can be shown as:

$$h_i^{(l+1)} = \sigma\left(\sum_{j=1}^{N} \frac{1}{c_{ij}} \cdot h_j^{(l)} \cdot A^{static/dynamic}\right) \tag{6}$$

where $h_i^{(l)}$ is the information of node i's at layer l, σ is the activation function, c_{ij} are normalization factors, and A is the weighted adjacency matrix. In STFEN, A includes both the static A^{static} and the dynamic $A^{dynamic}$ adjacency matrix.

STFEN uses two graph convolutional layers, each with a static and a dynamic GCN, to extract spatial features from the input signal based on physical proximity and signal similarity. The initial inputs for the first layer's GCNs come from the frequency band energy obtained after signal partition and DFT. The second layer's GCNs take the output from the first layer's static GCN.

In the convolution process, the static graph network consistently employs the current subject-normalized static adjacency matrix A^{static}, while the dynamic graph network recalculates its dynamic adjacency matrix $A^{dynamic}$ based on the latest data before each operation. To thoroughly extract spatial features from the sequential sEMG signals across multiple dimensions and depths, STFEN retains the results of the four node updates from two GCN layers.

TABLE I
ACCURACY OF STFEN - SOURCE MODEL ON STATIC ADSE SEMG DATASET

Subject	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	Mean
Accuracy	0.8882	0.943	0.962	0.8249	0.9747	0.9219	0.9916	0.8608	0.9705	0.9684	0.9306

TABLE II
PERFORMANCE COMPARISON OF STFEN AND BASELINE MODELS BASED ON ACCURACY ON THE ADSE DATASET

Methods		s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	Mean
Conventional Algorithms	SVM	0.1238	0.1429	0.1619	0.1905	0.1714	0.1333	0.1524	0.1143	0.1809	0.1048	0.1476
	MLP	0.1429	0.1810	0.1714	0.2381	0.2095	0.1905	0.2286	0.1619	0.2000	0.1524	0.1876
	RF	0.4381	0.4762	0.4667	0.4476	0.4286	0.4952	0.4571	0.4857	0.4571	0.4667	0.4619
sEMG Feature Analysis Algorithms	Tsagkas [14]	0.5714	0.6190	0.5905	0.6095	0.5619	0.6476	0.6000	0.5810	0.5524	0.6286	0.5962
	Betthauser [15]	0.8095	0.8190	0.8381	0.7905	0.8476	0.8000	0.8286	0.7714	0.8571	0.7809	0.8143
	Wu [16]	0.7333	0.7143	0.7809	0.6952	0.7619	0.7238	0.7429	0.7048	0.7714	0.7524	0.7381
	Garg [17]	0.6667	0.7238	0.6857	0.6476	0.7048	0.6190	0.6571	0.5333	0.6762	0.6381	0.6552
	Karnam [18]	0.6667	0.6857	0.6762	0.6476	0.6571	0.6952	0.6381	0.7143	0.7619	0.6286	0.6771
	Zabihi [19]	0.8095	0.7429	0.7809	0.8667	0.7143	0.7524	0.8381	0.7619	0.7905	0.7333	0.7790
	Sun [20]	0.8476	0.8190	0.8286	0.8095	0.8667	0.8000	0.8571	0.7905	0.8952	0.8762	0.8390
sEMG Transfer Learning Algorithms	Yu [21]	0.8476	0.8762	0.8286	0.8571	0.8095	0.8857	0.8381	0.8190	0.8952	0.8667	0.8524
	Lu [22]	0.7429	0.7905	0.7238	0.7714	0.7524	0.8095	0.7333	0.7809	0.8476	0.8000	0.7752
	STFEN(Ours)	**0.8667**	**0.9143**	**0.9048**	**0.8762**	**0.9429**	**0.8762**	**0.9143**	**0.8381**	**0.9524**	**0.9333**	**0.9019**

4) Pooling for Spatial Feature Extraction: Pooling is a widely used technique in deep learning that aims to lessen the dimensions of feature maps and computational demands while preserving essential feature information. In GCNs, even though local information gets combined at each node, the broader information of the whole graph might not be fully captured. Therefore, pooling is crucial for further extracting this wide-ranging graph information, simplifying feature dimensions, reducing the model's complexity, and improving robustness.

Typical pooling methods include max pooling and average pooling. Max pooling takes the highest value within the pooling area as the outcome, described as:

$$P_{\max}(i,j) = \max_{p,q \in window(i,j)} F(p,q) \tag{7}$$

where $P_{\max}$ is the max pooling matrix, window(i,j) is the central point of the ith row and jth column within the pooling area, and F is the matrix being pooled. In contrast, average pooling calculates the mean value of features within the pooling area, described as:

$$P_{\text{avg}}(i,j) = \frac{1}{m \times n} \sum_{p,q \in \ window(i,j)} F(p,q) \tag{8}$$

where P_{avg} is the average pooling matrix, with m and n indicating the number of rows and columns in the pooling window, respectively.

For sEMG signal processing, max pooling focuses on local features, highlighting critical aspects that are responsive to abrupt changes or significant occurrences. Average pooling, meanwhile, sums up broader information, providing a steady summary of signal traits and effectively spotting overall signal changes.

STFEN employs both max and average pooling after graph convolution and concatenation to utilize their benefits. The pooled results are then merged and processed through MLPs to derive a spatial feature vector. This approach of using both pooling types boosts the model's capacity to generalize across different sEMG signals.

V. RESULTS AND ANALYSIS

A. STFEN Performance on the ADSE Dataset

We first train the static sEMG source model within STFEN and assess its effectiveness through a subject-wise cross-validation approach on the ADSE static dataset.The results, shown in Table I, showcase an average accuracy of over 93%, proving the static model's effectiveness as a transfer source model.

After training the source model, the comprehensive STFEN framework is trained on the ADSE dataset, using a cross-validation approach to assess it as well. For comparison, ten benchmark models, including traditional machine learning methods like SVM and MLP, recent sEMG feature extraction methods, along with two advanced sEMG transfer learning methods, are evaluated on the ADSE dataset. The comparative results, presented in Table II, highlight STFEN's superior classification performance against both contemporary sEMG feature extraction methods and transfer learning methods on the ADSE task. This comparison not only confirms STFEN's strong classification ability but also shows the overall advantage of deep learning over traditional machine learning methods. Specifically, transfer learning-based methods show a significant advantage, benefiting greatly from the use of static information, compared to some end-to-end sEMG feature analysis methods.

B. Performance of STFEN on the Modified NinaPro DB2

To assess STFEN's effectiveness in recognizing sequential sEMG actions across different datasets, the NinaPro DB2

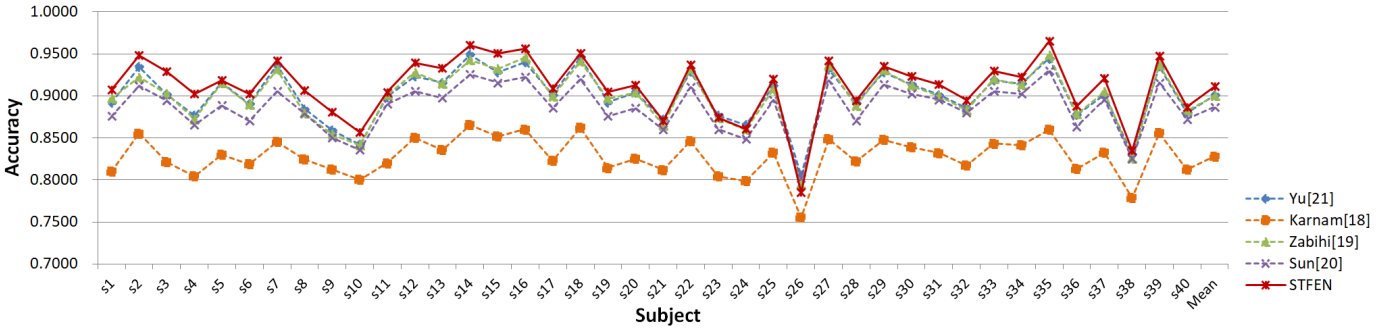

Fig. 5. Comparison of STFEN and Baseline Models based on Accuracy on the Modified NinaPro DB2

TABLE III
STFEN'S ABLATION EXPERIMENT BASED ON ACCURACY ON ADSE DATASET

Subject	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	Mean
Omission of Temporal Feature Extraction Module	0.7238	0.7619	0.7524	0.7333	0.7810	0.7333	0.7619	0.7524	0.7905	0.7714	0.7562
Omission of Spatial Graph Feature Extraction Module	0.8571	0.8571	0.8476	0.8190	0.8762	0.7714	0.8571	0.8286	0.8857	0.8286	0.8428
Substitution of Temporal Feature Extraction with LSTM Method	0.7524	0.8000	0.7905	0.7619	0.8190	0.7619	0.8000	0.6952	0.8381	0.8095	0.7829
Substitution of Temporal Feature Extraction with GRU Method	0.7048	0.7524	0.7143	0.6857	0.7905	0.7238	0.7429	0.6476	0.8381	0.8286	0.7429
GCN Utilizing Solely Static Graph Adjacency Matrix	0.8571	0.9048	0.8667	0.8476	0.8952	0.8762	0.8762	0.8000	0.9143	0.9048	0.8743
GCN Utilizing Solely Dynamic Graph Adjacency Matrix	0.8476	0.8667	0.8571	0.8286	0.8952	0.8286	0.8667	0.7905	0.9048	0.8857	0.8572
Sole Use of Max Pooling in GCN Pooling Layer	**0.8857**	**0.9143**	0.8762	0.8667	0.9143	**0.8762**	0.8952	**0.8381**	0.9333	0.9143	0.8914
Sole Use of Avg Pooling in GCN Pooling Layer	0.8476	0.8952	0.8857	0.8667	**0.9429**	0.8667	0.8952	0.8190	0.9238	0.9048	0.8848
Direct Processing of sEMG without DFT	0.7429	0.7810	0.7524	0.7238	0.8095	0.7143	0.7810	0.6857	0.8190	0.7714	0.7581
STFEN(Ours)	0.8667	**0.9143**	**0.9048**	**0.8762**	**0.9429**	**0.8762**	**0.9143**	**0.8381**	**0.9524**	**0.9333**	**0.9019**

dataset [13] is chosen for a detailed study. We re-categorize actions into sequential and static groups and modify data segmentation to better represent the sequential features in sequential sEMG signals.

The NinaPro DB2 dataset includes sEMG signals from 40 subjects performing 49 unique actions six times, with each action lasting 5 seconds, separated by 3 seconds of rest. It is recorded using the Delsys Trigno Wireless system with 12 leads at a $2kHz$ sampling rate.

Given the lack of explicit lead distance measurements in the NinaPro dataset, we base on the spatial layout on the lead setup during the DB2 data acquisition. The first eight sensors are arranged in an arm band at equal distances, while leads 9 to 12 are placed on specific muscle areas. For STFEN's analysis, facilitating modeling of spatial relationships among lead locations, this setup informs the establishment of the static adjacency matrix A^{static} as:

$$A^{static}_{ninaPro} = \begin{bmatrix} 0.0 & 1.0 & 0.5 & 0.3 & 0.3 & 0.3 & 0.5 & 1.0 & 0.2 & 0.2 & 0.1 & 0.1 \\ 1.0 & 0.0 & 1.0 & 0.5 & 0.3 & 0.3 & 0.3 & 0.5 & 0.2 & 0.2 & 0.1 & 0.1 \\ 0.5 & 1.0 & 0.0 & 1.0 & 0.5 & 0.3 & 0.3 & 0.3 & 0.2 & 0.2 & 0.1 & 0.1 \\ 0.3 & 0.5 & 1.0 & 0.0 & 1.0 & 0.5 & 0.3 & 0.3 & 0.2 & 0.2 & 0.1 & 0.1 \\ 0.3 & 0.3 & 0.5 & 1.0 & 0.0 & 1.0 & 0.5 & 0.3 & 0.2 & 0.2 & 0.1 & 0.1 \\ 0.3 & 0.3 & 0.3 & 0.5 & 1.0 & 0.0 & 1.0 & 0.5 & 0.2 & 0.2 & 0.1 & 0.1 \\ 0.5 & 0.3 & 0.3 & 0.3 & 0.5 & 1.0 & 0.0 & 1.0 & 0.2 & 0.2 & 0.1 & 0.1 \\ 1.0 & 0.5 & 0.3 & 0.3 & 0.3 & 0.5 & 1.0 & 0.0 & 0.2 & 0.2 & 0.1 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.0 & 0.8 & 0.3 & 0.3 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.8 & 0.0 & 0.3 & 0.3 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.3 & 0.3 & 0.0 & 0.8 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.3 & 0.3 & 0.8 & 0.0 \end{bmatrix} \quad (9)$$

We conduct a comparative analysis between STFEN and four benchmark models previously performing well on the NinaPro DB2 dataset, as shown in Fig. 5. for accuracy comparison. STFEN achieves the highest accuracy in 37 out of 40 subjects, surpassing the other models. This comparison confirms STFEN's effectiveness and adaptability in recognizing sequential sEMG tasks, demonstrating its ability to extract relevant patterns from datasets like NinaPro DB2, which are originally designed for static action analysis.

C. Ablation Study on STFEN

An ablation study is carried out on the STFEN framework to determine the impact of each component on its overall

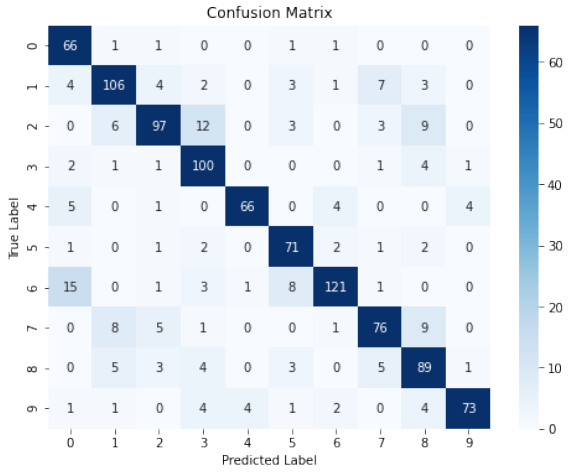

Fig. 6. Confusion Matrix of STFEN for Identification across All Subjects in the ADSE Dataset: The vertical axis represents the true labels, while the horizontal axis represents the labels predicted by STFEN.

effectiveness. By removing or replacing specific parts or structures, we make corresponding adjustments to the rest of the model. Nine variations of the model are evaluated: **a.** Without the temporal feature extraction module; **b.** Without the spatial feature extraction module; **c.** Substituting the temporal feature extraction's knowledge transfer with a 2-layer LSTM; **d.** Substituting with a 2-layer GRU; **e.** Using only static adjacency in GCN for spatial features; **f.** Using only dynamic adjacency; **g.** Applying only Max pooling; **h.** Applying only Average pooling; **i.** Bypassing DFT for direct sEMG feature extraction.

Results, shown in Table III, highlight several key observations:

a) Spatial and Temporal Modules: Including both spatial and temporal modules significantly improves movement detection, underscoring their collective importance.

b) Temporal Feature Extraction: The knowledge transfer method outperforms RNN-based algorithms like LSTM and GRU networks in temporal feature analysis, emphasizing the value of leveraging static sEMG knowledge to boost temporal feature extraction.

c) Graph Structures in GCN: Employing both dynamic and static adjacency graphs for spatial analysis offers a view of spatial relationships in sEMG signals from more perspectives, better than relying on just one type of adjacency graph.

d) Pooling Methods: While both Max and Average pooling contribute to higher accuracy, Max pooling shows particular efficacy with certain subjects, indicating it may better capture local signal features. However, STFEN's approach of using both pooling methods consistently deliver competitive and superior results, showcasing its more comprehensive ability in feature extraction.

e) Frequency Band Energy Transformation: Extracting features directly from sEMG signals without using DFT is less effective than transforming sEMG signals into frequency band energy, indicating DFT's critical role in precise recognition of sequential sEMG actions.

TABLE IV
TOP 5 DIGIT PAIRS MOST FREQUENTLY MISIDENTIFIED BY STFEN

Pairs	Occurrences of Misclassifications
0-6	16
1-7	15
7-8	14
2-3	13
2-8	12

D. Confusion Matrix Analysis

Evaluating STFEN's ability to classify actions on the ADSE validation set results in the confusion matrix shown as Fig. 6. Table IV lists the top five digit pairs that are most frequently confused. These misclassifications mainly fall into two categories based on their characteristics.

The first category includes digit pairs '0-6' and '1-7,' which share a high degree of similarity. This similarity mirrors typical difficulties in distinguishing handwritten digits, where alike shapes might cause mix-ups. The second category includes digit pairs '7-8', '2-3' and '2-8', characterized by the fact that the trajectory of one digit is part of another. If only the sEMG signals corresponding to the initial part of the trajectory drawn for the digit '8' are acquired, it becomes easy to recognize the drawing as the digit '7' or '2' (after all, based on the acquired data, the subject indeed only drew a trajectory resembling the digit '7' or '2'). The same logic applies to the digit pairs '2-3'. Given the variability in the timing when subjects start the action, there is a certain possibility of such incomplete data acquisition occurring.

The identified misclassifications, mainly between digit pairs with similarities or trajectory overlaps, are exactly extreme cases in this task. Therefore, these mistakes can be considered acceptable within the context of this study.

E. Inter-lead Correlation in the Final Layer GCN Dynamic Adjacency Matrix

The duplication of STFEN's dynamic adjacency matrix from the final layer GCN during the learning process with the ADSE dataset is shown in Fig. 7. Analysis of matrices from 10 subjects reveals patterns of self-learned dynamic adjacency, showing consistent yet unique inter-lead correlations, highlighting three main observations:

a) Upper Arm Muscle Group Correlation: A significant correlation is observed among five leads linked to the deltoid, biceps, triceps, flexor carpi radialis, and extensor carpi radialis muscles, highlighted within the red-boxed area **a**. This group, particularly the four muscles excluding the deltoid, shows a marked correlation across nine of ten subjects, suggesting their combined action during digit drawing tasks.

b) Core Muscle Group Correlation: Leads linked to the lats and pecs, highlighted within the orange-boxed area **b**, show strong correlations across 9 of 10 subjects, pointing to a coordinated activation pattern within core muscles during digit drawing tasks.

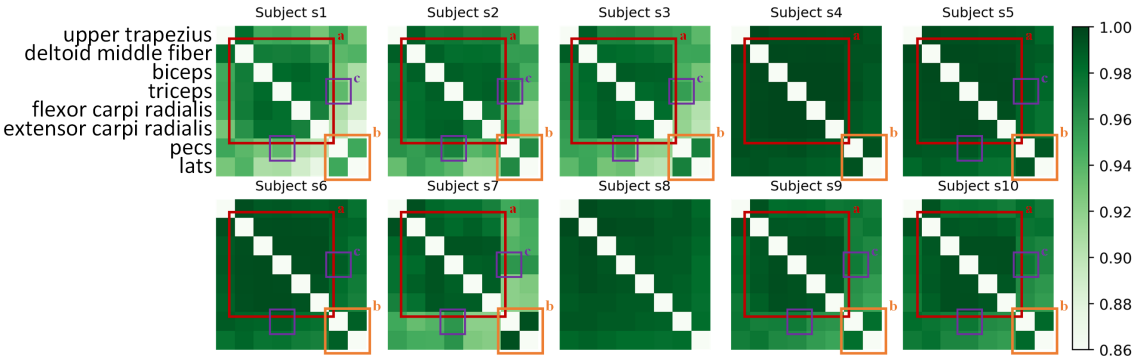

Fig. 7. Dynamic Adjacency Matrices of the Last Layer GCN for 10 Subjects: The red box denotes the upper arm muscles, encompassing the deltoid middle fiber, biceps, triceps, flexor carpi radialis, and extensor carpi radialis; the orange box denotes the torso muscles, including the pecs and lats; meanwhile, the purple box identifies a deeper interrelation between the pecs and triceps. The muscles corresponding to the leads of the three boxes all show strong correlation inside.

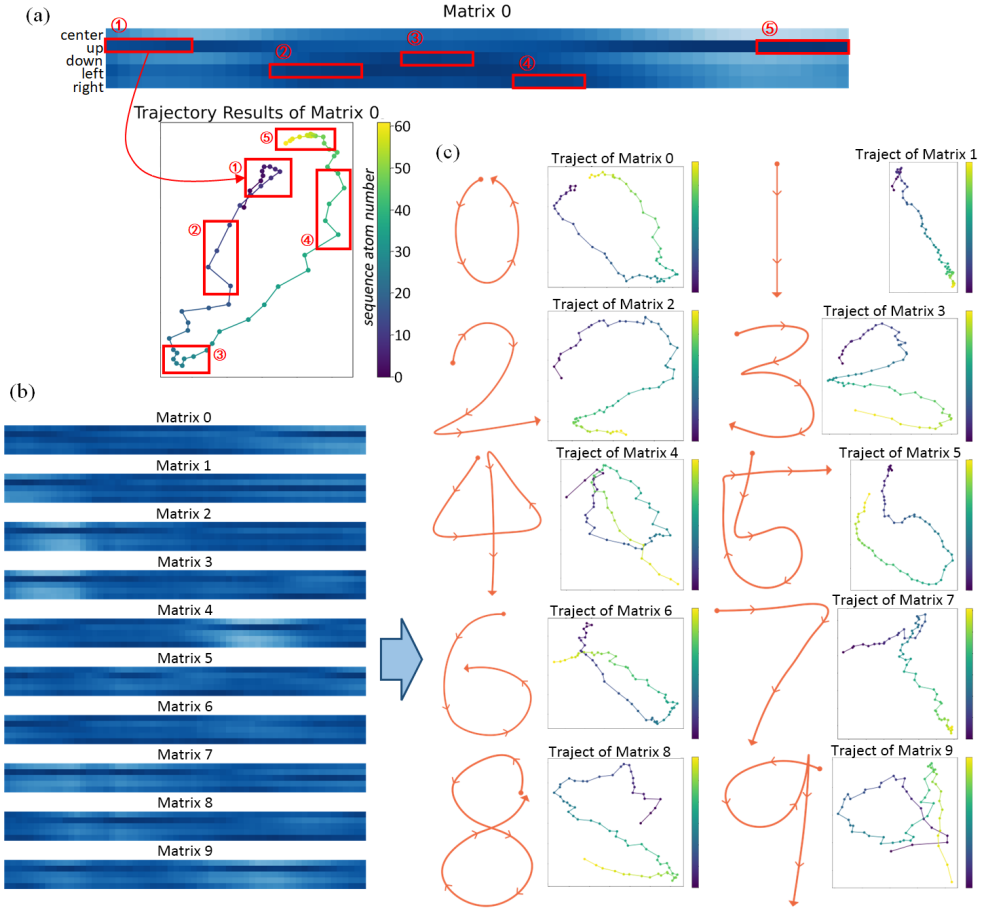

Fig. 8. (a) Relationship between Sequential Temporal Features after Knowledge Transfer and Their Corresponding Trajectories in Space: Based on the hand position probabilities identified from sEMG atom by the static sEMG model, the spatial position corresponding to each sEMG atom can be calculated. By connecting these spatial positions in sequential order, the trajectory in space can be obtained. (b) Hand Position Probability Matrix Obtained through Knowledge Transfer of Ten Digits. (c) The Visualization of Spatial Trajectories Corresponding to These Probability Matrices.

c) Deep Pattern Association between Triceps and Pecs: Despite the physical distance between their respective leads, the triceps and pecs, highlighted within the purple-boxed area **c**, exhibit a notable synergistic correlation across 8 of 10 subjects, suggesting a deep interaction pattern that goes beyond simple proximity of lead placement.

These insights confirm the dynamic graph structure of STFEN's ability to reveal spatial interactions between sEMG signal sources, with correlations in upper arm and core muscle groups reflecting anatomical and functional partnerships. Moreover, the unexpected correlation between triceps and pecs highlights STFEN's ability to uncover complex spatial relationships, enhancing its classification accuracy. This analysis highlights the model's success in using dynamic inter-lead connections to interpret sequential sEMG data accurately.

Unlike other subjects, Subject s8 exhibited no pronounced

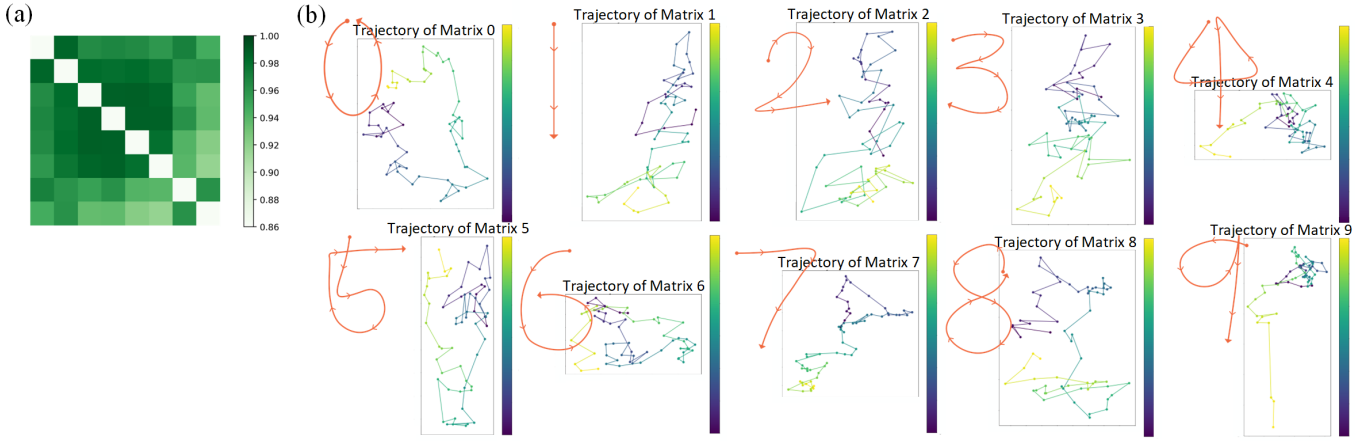

Fig. 9. (a) Dynamic Adjacency Matrices of Stroke Patients: The muscle synergy patterns in stroke patients are similar to those of healthy individuals, showing significant internal correlation between muscles. (b) Visual Trajectories of Stroke Patients Drawing 10 Digits: Compared to healthy individuals, stroke patients exhibit disorganized spatial trajectories, making it difficult to identify the digits they have drawn.

spatial adjacency characteristics between different electrodes. Instead, all leads showed a strong adjacency relationship. We believe this might be because this subject engages more muscles in coordination during the process of drawing numbers, as reflected by the dynamic adjacency matrix reflects. Therefore, we conclude that Subject s8 does not impact the overall effectiveness of STFEN in capturing spatial features.

F. Visualization of the Sequential Temporal Features

The process of extracting sequential features includes averaging the predictions for hand positions from the static sEMG model for each digit, as shown in Fig. 8(b). In this sequence matrix, each column represents the predicted positions of a sEMG atom, while the rows show the probability of the hand being in center, up, down, left, or right. To make these results clearer, we develop a new way to visualize them.

For each recognition result of an sEMG atom (p_{center}, p_{up}, p_{down}, p_{left}, p_{right}), we come up with a method to calculate as:

$$x = \frac{p_{right} - p_{left}}{p_{left} + p_{center} + p_{right}} \tag{10}$$

$$y = \frac{p_{up} - p_{down}}{p_{up} + p_{center} + p_{down}} \tag{11}$$

where x and y are the coordinates of the position in space corresponding to the sequential features.

These calculations help map the position of each sEMG atom to a detailed spatial path by linking sequential points, as illustrated in Fig. 8(a), leading to the trajectory visualization of 10 digits in Fig. 8(c). This method calculates the x and y coordinates separately, allowing the spatial distribution to cover the full $(-1, -1)$ to $(1, 1)$ area, greatly improving the clarity of trajectory visualization.

The trajectory visualizations clearly display the path characteristics for each digit as interpreted by the static sEMG network. This visualization validates the sequential feature extraction capability of STFEN's designed sequential feature extraction module based on knowledge transfer.

VI. Discussion

A. Comprehensive Review of STFEN

STFEN stands out from other sEMG feature extraction models with its specialized focus on sequential sEMG signals. It introduces a method for extracting sequential features that benefits from transferring knowledge from static to sequential sEMG signals, proving its effectiveness in identifying sequential features through different tests.

Additionally, STFEN specifically targets the gathering of spatial data found in the connections between leads during data acquisition. It develops a spatial feature extraction approach using GCN, which utilizes both static and dynamic adjacency matrices to represent physical distances between leads and their signal similarity correlation. Experimental validations highlight the enriched perspective these matrices bring to sequential feature extraction.

Furthermore, STFEN's implementation of GCNs and knowledge transfer mechanisms enhances its interpretability. The capacity to show intermediate results, such as sequential features and changing spatial connections, sheds light on how STFEN processes data, providing a clearer understanding of its functionality. This attribute stands as one of STFEN's advantages over other sEMG models.

B. Prospects of STFEN in Rehabilitation Medicine

Using the STFEN framework, we carry out a study with sEMG data from a subset of stroke patients, focusing on sequential action tasks at a partner medical center. Taking the data from one patient as an example, when analyzing the dynamic adjacency matrix of the patient, as shown in Fig. 9(a), we find that there is a muscle synergy relationship in the adjacency matrix that is similar to that of healthy people, which shows that the muscles of stroke patients can still function normally. However, action trajectories derived from sequential temporal features, as shown in Fig. 9(b), reveal irregularities in the patient's actions, and highlight the difficulties the patient experienced in performing the sequential action tasks.

Rehabilitation medicine increasingly recognizes the value of sEMG signal analysis in aiding recovery for patients, particularly those recuperating from strokes. The STFEN model's capability to detail muscle action features opens up valuable possibilities for real-time monitoring and assessing rehabilitation progress. Furthermore, sequential sEMG signals could control rehabilitative devices such as electrical stimulators, enhancing patient engagement in their recovery process.

REFERENCES

[1] I. L. Kopyt *et al.*, "A Novel sEMG Triggered FES-Hybrid Robotic Lower Limb Rehabilitation System for Stroke Patients," *IEEE Trans. Med. Rob. Bionics.*, vol. 2, no. 4, pp. 631–638, Nov. 2020, doi: 10.1109/TMRB.2020.3019081.

[2] J. R. Cram, G. S. Kasman, and J. Holtz, "Chapter 11 The History of Muscle Dysfunction and Surface Electromyography," in Introduction to Surface Electromyography, 1st ed. Gaithersburg, MD, USA: Aspen Publishers, 1998, ch. 11, pp. 175.

[3] Y. Chen *et al.*, "A continuous estimation model of upper limb joint angles by using surface electromyography and deep learning method," *IEEE Access*, vol. 7, pp. 174940–174950, Dec. 2019, doi: 10.1109/ACCESS.2019.2956951.

[4] L. Pan, D. L. Crouch, and H. Huang, "Comparing EMG-based human-machine interfaces for estimating continuous, coordinated movements," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 2145–2154, Aug. 2019, doi: 10.1109/TNSRE.2019.2937929.

[5] J. Liu *et al.*, "EMG-based real-time linear-nonlinear cascade regression decoding of shoulder, elbow, and wrist movements in able-bodied persons and stroke survivors," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 5, pp. 1272–1281, Aug. 2019, doi: 10.1109/TBME.2019.2935182.

[6] X. Chen *et al.*, "Hand Gesture Recognition based on Surface Electromyography using Convolutional Neural Network with Transfer Learning Method," *IEEE J. Biomed. Health. Inf.*, vol. 25, no. 4, pp. 1292–1304, Jul. 2020, doi: 10.1109/JBHI.2020.3009383.

[7] A. Ameri *et al.*, "A deep transfer learning approach to reducing the effect of electrode shift in EMG pattern recognition-based control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 370–379, Dec. 2019, doi: 10.1109/TNSRE.2019.2962189.

[8] K. Rezaee *et al.*, "A hybrid deep transfer learning-based approach for Parkinson's disease classification in surface electromyography signals," *Biomed. Signal Process. Control*, vol. 71, pp. 103161, Dec. 2022, doi: 10.1016/j.bspc.2021.103161.

[9] S. M. Massa, D. Riboni, and K. Nazarpour. (2022, November). Graph neural networks for HD EMG-based movement intention recognition: An initial investigation. presented at the 2022 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE). [Online]. Available: http://ieeexplore.ieee.org/document/9989657

[10] A. Vijayvargiya, R. Kumar, and P. Sharma, "PC-GNN: Pearson Correlation-Based Graph Neural Network for Recognition of Human Lower Limb Activity Using sEMG Signal," *IEEE Trans. Hum.-Mach. Syst.*, Oct. 2023, doi: 10.1109/THMS.2023.3319356.

[11] H. Li *et al.*, "A sequential learning model with GNN for EEG-EMG-based stroke rehabilitation BCI," *Front. Neurosci.*, vol. 17, pp. 1125230, Apr. 2023, doi: 10.3389/fnins.2023.1125230.

[12] H. Lee *et al.*, "Stretchable array electromyography sensor with graph neural network for static and dynamic gestures recognition system," *npj Flexible Electron.*, vol. 7, no. 1, pp. 20, Apr. 2023, doi: 10.1038/s41528-023-00246-3.

[13] M. Atzori *et al.*, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Sci. Data*, vol. 1, no. 1, pp. 1–13, Dec. 2014, doi: 10.1038/sdata.2014.53.

[14] N. Tsagkas, P. Tsinganos, and A. Skodras. (2019, July). On the use of deeper CNNs in hand gesture recognition based on sEMG signals. presented at the 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA). [Online]. Available: http://ieeexplore.ieee.org/document/8900709

[15] J. L. Betthauser *et al.*, "Stable responsive EMG sequence prediction and adaptive reinforcement with temporal convolutional networks," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 6, pp. 1707–1717, Sept. 2019, doi: 10.1109/TBME.2019.2943309.

[16] L. Wu *et al.*, "Metric learning for novel motion rejection in high-density myoelectric pattern recognition," *Knowledge-Based Syst.*, vol. 227, pp. 107165, Sept. 2021, doi: 10.1016/j.knosys.2021.107165.

[17] N. Garg *et al.*. (2021, July). Signals to spikes for neuromorphic regulated reservoir computing and EMG hand gesture recognition. Presented at the International Conference on Neuromorphic Systems 2021 (ICONS 2021). [Online]. Available: https://doi.org/10.1145/3477145.3477267

[18] N. K. Karnam *et al.*, "EMGHandNet: A hybrid CNN and Bi-LSTM architecture for hand activity classification using surface EMG signals," *Biocybern. Biomed. Eng.*, vol. 42, no. 1, pp. 325–340, Jan. 2022, doi: 10.1016/j.bbe.2022.02.005.

[19] S. Zabihi *et al.*, "Trahgr: Transformer for hand gesture recognition via electromyography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, Oct. 2023, doi: 10.1109/TNSRE.2023.3324252.

[20] B. Sun *et al.*, "A multi-scale feature extraction network based on channel-spatial attention for electromyographic signal classification," *IEEE Trans. Cognit. Dev. Syst.*, Apr. 2022, doi: 10.1109/TCDS.2022.3167042.

[21] Z. Yu *et al.*, "Surface EMG-based instantaneous hand gesture recognition using convolutional neural network with the transfer learning method," *Sensors*, vol. 21, no. 7, pp. 2540, Jan. 2021, doi: 10.3390/s21072540.

[22] J. Lu *et al.*. (2023, April). EffiE: Efficient Convolutional Neural Network for Real-Time EMG Pattern Recognition System on Edge Devices. Presented at the 2023 11th International IEEE/EMBS Conference on Neural Engineering (NER). [Online]. Available: https://doi.org/10.1109/NER52421.2023.10123741