

Article

Fusion Learning for sEMG Recognition of Multiple Upper-Limb Rehabilitation Movements

Tianyang Zhong, Donglin Li, Jianhui Wang *, Jiacan Xu, Zida An and Yue Zhu

College of Information Science and Engineering, Northeastern University, Shenyang 110819, China; 1970815@stu.neu.edu.cn (T.Z.); 1910309@stu.neu.edu.cn (D.L.); 1410330@stu.neu.edu.cn (J.X.); 1900753@stu.neu.edu.cn (Z.A.); 20174129@stu.neu.edu.cn (Y.Z.)

* Correspondence: wangjianhui@mail.neu.edu.cn

Abstract: Surface electromyogram (sEMG) signals have been used in human motion intention recognition, which has significant application prospects in the fields of rehabilitation medicine and cognitive science. However, some valuable dynamic information on upper-limb motions is lost in the process of feature extraction for sEMG signals, and there exists the fact that only a small variety of rehabilitation movements can be distinguished, and the classification accuracy is easily affected. To solve these dilemmas, first, a multiscale time–frequency information fusion representation method (MTFIFR) is proposed to obtain the time–frequency features of multichannel sEMG signals. Then, this paper designs the multiple feature fusion network (MFFN), which aims at strengthening the ability of feature extraction. Finally, a deep belief network (DBN) was introduced as the classification model of the MFFN to boost the generalization performance for more types of upper-limb movements. In the experiments, 12 kinds of upper-limb rehabilitation actions were recognized utilizing four sEMG sensors. The maximum identification accuracy was 86.10% and the average classification accuracy of the proposed MFFN was 73.49%, indicating that the time–frequency representation approach combined with the MFFN is superior to the traditional machine learning and convolutional neural network.

Keywords: surface electromyogram; motion intention recognition; multiscale time–frequency information fusion representation; multiple feature fusion network; deep belief network



Citation: Zhong, T.; Li, D.; Wang, J.; Xu, J.; An, Z.; Zhu, Y. Fusion Learning for sEMG Recognition of Multiple Upper-Limb Rehabilitation Movements. *Sensors* **2021**, *21*, 5385. <https://doi.org/10.3390/s21165385>

Academic Editors: Ana Filipa Sequeira and Ana Maria Rebelo

Received: 30 June 2021
Accepted: 3 August 2021
Published: 9 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As an advanced technique, myoelectric control is mainly applied in aspects of human-support robots, industrial electronic equipment, or rehabilitation devices, such as surface electromyogram (sEMG)-based wheelchair controller [1,2], exoskeletons [3,4], industrial robots [5], diagnoses, and clinical applications [6,7]. These applications mainly focus on the identification of certain limited pattern types; however, stroke patients have strong demands to improve quality of life by rehabilitation training, which is completed with the assistance of the upper-limb rehabilitation system [8]. Moreover, the upper-limb rehabilitation system that operates with the active participation of the human brain brings great benefits [9,10]. Therefore, it is vital to recognize more types of upper-limb movements because the more upper-limb motion commands can be classified, the more the dexterity of the rehabilitation system that can be gained to reinforce the rehabilitation effect of stroke patients. Thus, this paper intends to investigate the sEMG-based identification of upper-limb locomotion for the rehabilitation system.

As a kind of random bioelectrical signal, sEMG signals are weak, low-frequency, and vulnerable to interference, often accompanied by artifacts, high dimension, and non-stationary problems [11]. Due to the significant differences among different individuals in age, gender, and physique, the characteristics of the sEMG become more prominent [12]. This phenomenon leads to the case that the existing algorithms can discriminate fewer kinds of upper-limb rehabilitation motions, and the classification accuracy is easily altered, which

dramatically influences the rehabilitation effect of stroke patients. Accurately identifying more types of upper-limb rehabilitation actions is extremely challenging in this stage [13], which has crucial theoretical significance and application value.

To solve the above problems, time–frequency domain techniques are widely adopted in feature extraction [14–29]. Englehart et al. utilized the short-time Fourier transform and wavelet transform methods and distinguished six types of upper-limb movements. Shi et al. [15] extracted the time-domain and frequency-domain features for four gestures. Ding et al. [16] gathered the wavelet packet transform coefficient, stationary wavelet transform coefficient, and time–frequency features for the recognition of five kinds of actions. Although the manually designed methods can extract the time–frequency features of sEMG signals, only part of the useful features were considered. However, some beneficial time–frequency features were not taken into account, which resulted in the loss of valid time–frequency features for the sEMG signals and cannot satisfy the discrimination of more rehabilitation actions.

Traditional machine learning algorithms combined with the above methods have been widely applied in intention identification [17–23], but performance and accuracy are unsatisfactory. Therefore, researchers began to combine various deep learning models for intention recognition [24]. Duan et al. [25] adopted the wavelet neural network combined with the wavelet transform, and the accuracy rate for six gestures was as high as 94.67%. Wei et al. [26] proposed a multistream CNN to recognize three public sEMG datasets, indicating that the neural network approach is superior to the traditional neural network method. Discrete wavelet transform [27,28] can decompose sEMG signals into regular time series with different frequency bands. Thus, Wu et al. [29] collected eight channels of sEMG signals and identified ten gestures using LSTM and CNN. However, the deep learning frameworks commonly applied for intention detection face the following problems: (1) some valuable time–frequency features would be lost during the process of feature extraction; (2) more kinds of upper-limb rehabilitation movements would easily influence the performance of classification algorithms.

To sum up, this paper mainly considers increasing the types of rehabilitation movements, extracting more complete time–frequency features of sEMG and heightening the stability of the identification accuracy for the intention recognition system. The contributions of this paper are as follows:

- Data representation stage: considering the validity and integrity of time–frequency characteristics, a multiscale time–frequency information fusion method is proposed to obtain the dynamic, related information of multichannel sEMG signals.
- Feature extraction stage: in this paper, a multiple feature fusion network is designed for the purpose of considering the relevant features of the current layer and cross-layer of CNN and reducing the loss of the time–frequency features for sEMG signals in the process of convolution operation.
- Classification stage: the deep belief network is introduced as a classification model of the multiple feature fusion network to realize the abstract expression and self-reconstruction for time–frequency features of sEMG signals corresponding to more types of upper-limb movements.

Figure 1 shows the overall framework flow chart of this work, and this paper is divided into the following parts. Section 2 describes the rehabilitation actions' design and sEMG acquisition, meanwhile, the data representation method and the architecture of the proposed MFFN are mainly introduced. Section 3 verifies the performance of the MFFN and then visualizes the learning effect. Finally, the conclusions are discussed in Section 4.

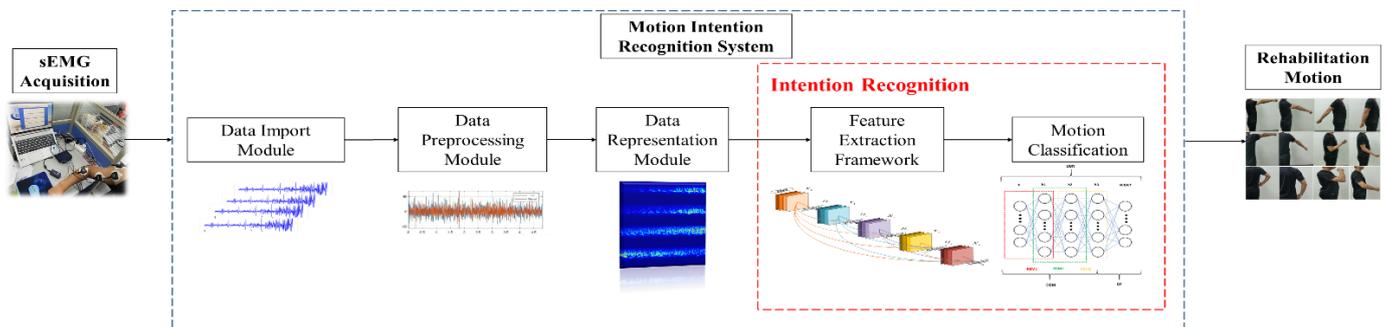


Figure 1. This is the schematic diagram of the proposed intention recognition method in this paper.

2. Methodology

This paper's main data representation is based on continuous wavelet transform (CWT) [30,31]. In Section 2.1, data acquisition is introduced in detail. In Section 2.2, the approach of multiscale time–frequency information fusion is proposed. In Section 2.3, the basic frameworks and principles of DenseNet and DBN are introduced. In Section 2.4, a multiple feature fusion network based on DenseNet and DBN is presented, and the selection of network parameters is explained.

2.1. Data Acquisition

The electromyogram (EMG) signal is a complex biomedical signal produced during the muscle contraction process, which is influenced by anatomical or physiological properties of the muscles and environmental noises [32]. Therefore, the acquisition of EMG signals is a basic task influencing the next classification of upper-limb motion command recognition. In general, EMG signals can be acquired through either invasive or noninvasive electrodes. The EMG signals collected by invasive needles or wire electrodes are suitable for the study of deep muscle structure. However, due to noninvasiveness and convenient operation, sEMG techniques are mainly used to detect sEMG signals in biofeedback, rehabilitation medicine, movement analysis, and other fields. Thus, sEMG is employed as the intermediary carrier to control the upper-limb rehabilitation system.

First, considering the principle of human biological structure and the range of motion of each joint, twelve upper-limb rehabilitation exercises, commonly used in rehabilitation training for stroke patients, were designed in this paper, involving single-joint actions (posterior extension of shoulder, shoulder backstretch return, shoulder adduction, shoulder anterior flexion, shoulder forward flexion return, shoulder abduction, elbow flexion, and elbow extension) and compound joint actions (feeding action, feeding return action, pant move, and pant return move); and the action1–action12 are shown in Figure 2.

Then, the corresponding muscle groups of the upper limb needed to be determined. Although there are deep-layer and superficial-layer muscles that contribute to the above motions, the sEMG signals are mainly affected by the superficial-layer muscles. Therefore, four superficial-layer muscles were selected to fix the sEMG sensors, CH1 to CH4, which were the deltoid lateral muscle, anterior deltoid muscle, biceps brachii, and radial muscle, respectively, as shown in Figure 3.

To reduce the influence of skin, such as impedance, superficial oil content, and dead cell layer, the surface of the expected muscle groups on the upper limb was wiped with alcohol. Moreover, differential surface electrodes with large distances were utilized to collect sEMG signals after the skin dried. The sEMG signals were recorded with a sampled rate of 2048 Hz using FlexComp Infiniti (Montreal, Canada) acquisition equipment.

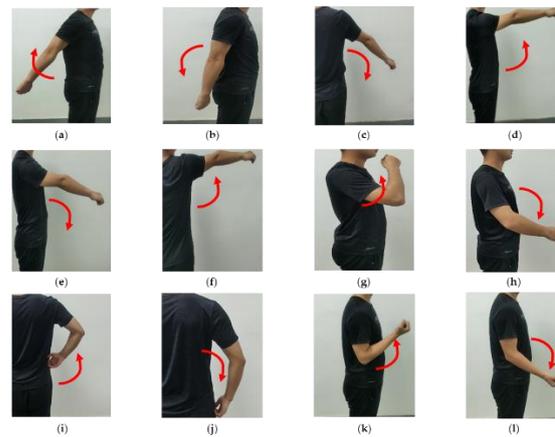


Figure 2. Twelve kinds of upper-limb rehabilitation movements designed for stroke patients are listed as follows: (a) posterior extension of shoulder (action1); (b) shoulder backstretch return (action2); (c) shoulder adduction (action3); (d) shoulder anterior flexion (action4); (e) shoulder forward flexion return (action5); (f) shoulder abduction (action6); (g) feeding action (action7); (h) feeding return action (action8); (i) pant move (action9); (j) pant return move (action10); (k) elbow flexion (action11); (l) elbow extension (action12).



Figure 3. The position of patch electrodes and the sEMG acquisition equipment.

Some rehabilitation experts have suggested that one should utilize healthy subjects for initial evaluation objectives, and the sEMG of healthy subjects is an appropriate emulation of the stroke patients' sEMG. Thereby, in this study, sEMG signals were collected from 14 healthy subjects (male and female) with ages ranging from 20 to 45 years old, with physiques including fat, overweight, normal, thin, and underweight. Each rehabilitation movement lasted for 5 s and was repeated five times. Each action was composed of 47 samples. Therefore, the total dataset consisted of $12 \text{ movements} \times 47 \text{ samples} \times 4 \text{ channels} \times 10,240 = 23,101,360$ sampling points.

2.2. Data Representation

Before the operation of data representation, the characteristics and attributes of the sEMG signal need to be considered. First, preprocessing of the sEMG signal from its own point of view was executed, and then the data representation operation was performed on the multichannel sEMG signal from the two aspects of the effectiveness and integrity of time–frequency conversion. Therefore, this section consists of two parts: preprocessing and data representation.

2.2.1. Signal Preprocessing

The signal preprocessing is the premise of subsequent data representation. The essence of preprocessing is to weaken the mixed noise in the signal and retain the useful components of the raw signal as much as possible. Thus, it is necessary to carry out a series of preprocessing steps on the raw sEMG signals. In the process of signal acquisition, the sEMG signal is greatly affected by 50 Hz power frequency interference. Therefore, a digital filter was adopted to remove 50 Hz power frequency noise. Moreover, the effective frequency range of sEMG signals is generally distributed in the range of 6~500 Hz, among which researchers have found that most part of the spectrum energy mainly concentrates on 50~150 Hz. Furthermore, the literature [33,34] pointed out that a high-pass filter can remove motion artifact noise while retaining most of the sEMG signal. To sum up, 6 Hz was selected as the cutoff frequency of the high-pass filter in this paper to filter the sEMG signals, which will obtain a more comprehensive sEMG signal without loss of useful information.

2.2.2. Multiscale Time–Frequency Information Fusion Representation

The functional frequency of the sEMG signal itself is 6~500 Hz, so this paper extracts time–frequency features within the corresponding frequency band, and each channel is done by the continuous wavelet transform (CWT), which can greatly reduce the amount of irrelevant data. From the perspective of the time–frequency feature fusion, there is no interaction among multichannel sEMG signals, and the order of input matrices does not affect the classification performance of the network [35]. Therefore, the multiscale time–frequency features could be represented as an RGB image about time, frequency, and amplitude. Finally, the target time–frequency features are extracted as the input of the network. In this paper, the multiscale decomposition module on time–frequency representation (MDTFR) is designed for a single-channel sEMG signal; and the MDTFR module is performed for the multichannel sEMG signals in the valid frequency band. Finally, the time–frequency features of the multichannel sEMG signals are seized comprehensively by fusion operation.

For the MDTFR module, as shown in Figure 4, the analytic Morlet wavelet, which shows good time–frequency characteristics and has no negative frequency component [36,37], is applied to analyze the time–frequency localization of the received sEMG signals within the effective frequency range. The time–frequency windows that vary with frequency can be provided by adjusting the scale factor, s , and the delay factor, a . Then, the CWT method is used to achieve time subdivision at high frequency and frequency subdivision at low frequency, which can automatically adapt the requirements of sEMG signal analysis so as to focus on the multiscale time–frequency details. This process can be defined as

$$\text{mm}[f(t)] = \{ S \mid S_i = \text{CWT}_{s_i, a_i}(f(t)), 1 \leq i \leq F \} \quad (1)$$

where s_i, a_i represent scale factor and delay factor; $f(t), F$ correspond to the time signal and its highest effective frequency; $\text{CWT}_{s_i, a_i}(\cdot)$ is the continuous wavelet transform under a combination of different s_i and a_i . The CWT is defined as the inner product of functions $f(t)$ and wavelet $\psi_{s,a}(t)$:

$$\text{CWT}_{s,a}(f(t)) = \langle f(t), \psi_{s,a}(t) \rangle = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t) \bar{\psi}_{s,a}(t) dt \quad (2)$$

where $\bar{\psi}_{s,a}(\frac{t-a}{s})$ is the conjugate function of $\psi_{s,a}(\frac{t-a}{s})$.

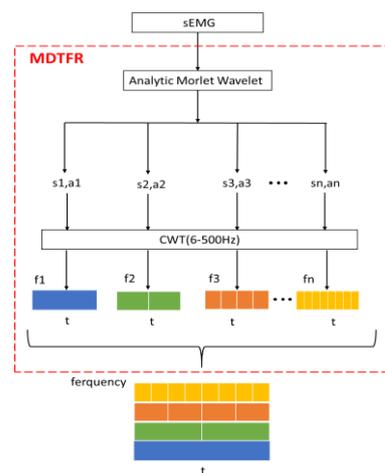


Figure 4. The multiscale decomposition on time-frequency representation for single-channel sEMG.

The process of MTFIFR is shown in Figure 5. Firstly, the denoised multichannel sEMG signals are processed by the MDTFR module, which aims at collecting the time–frequency features of the surface electromyography data about each channel. Then, these features concerning the valid frequency range of the sEMG signal are fused and represented by concatenating vertically. Finally, the multiscale time–frequency features corresponding to the dynamic information of upper-limb rehabilitation movements are acquired into a three-dimensional RGB image, which can be defined as

$$r = \text{concat}([m[ch1(t)], m[ch2(t)], m[ch3(t)], m[ch4(t)]]^T) \quad (3)$$

where $ch1 \sim ch4$ correspond to sEMG signals of four channels, and $\text{concat}(\cdot)$ represents the longitudinal connection to the feature map.

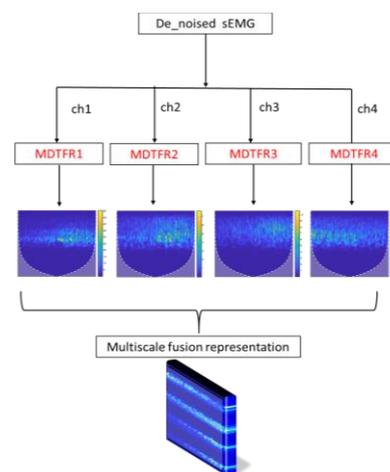


Figure 5. The whole data representation flowchart of multiscale time-frequency information fusion representation (MTFIFR).

2.3. Components of MFFN

The main data representation method in this paper is based on the above MTFIFR. For the purpose of acquiring complete time-frequency features of sEMG signals for upper-limb movements' identification, the framework of DenseNet was adjusted to make it more suitable for feature extraction and recognition of sEMG signals in this paper. In the classification stage, in order to improve the adaptability of the network and the stability of identification accuracy for various rehabilitation exercises, this paper realize the bidirectional evaluation

of the time-frequency features for sEMG and the corresponding rehabilitation movement by the deep belief network.

2.3.1. DenseNet

The architecture of a convolutional neural network consists of basic components such as input layer, convolution layer, pooling layer, and classification layer, and most architectures follow this process. In order to ensure maximum information flow among layers of the network, the dense block directly connects the output feature maps of all layers in a feedforward fashion, which reduces the loss of the features in the process of convolution operation [38]. Each layer receives additional input from all previous layers and passes its own feature maps to all subsequent layers, so that the network maintains feedforward characteristics. Figure 6 illustrates this architecture schematically, which is the core component of DenseNet. There is a direct connection between any two layers, and the output of the number L layer is defined as

$$x_L = H([x_0, x_1, \dots, x_{L-1}]) \quad (4)$$

where $H(\cdot)$ represents BN-ReLU-Conv (1×1)-BN-ReLU-Conv (3×3); x_0 is the input of the dense block; $[x_1, x_2, \dots, x_{L-1}]$ refers to the concatenation of the feature maps produced in layers $1, 2, \dots, L-1$; and dense blocks are connected by a transition layer, which is generally composed of BN-ReLU-Conv (1×1)-Ave pooling (2×2). For an L -layer network, the traditional convolutional neural network has L connections, while the DenseNet contains $L(L+1)/2$ connections, which significantly raises the transmission and reuse of features. Meanwhile, it is also observed that dense block modules have a regularization effect, which reduces the over-fitting problem of small sample data. Therefore, the DenseNet-based structure is capable of feature extraction for sEMG signals.

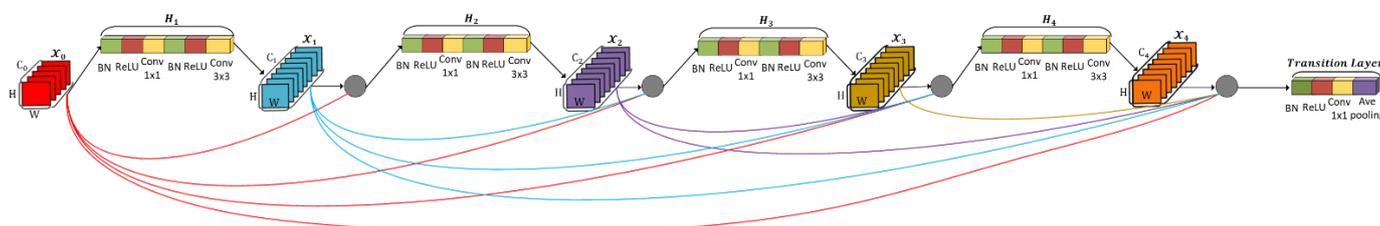


Figure 6. This is the basic framework of dense block and transition layer in DenseNet.

2.3.2. The Deep Belief Network

In the classification stage, the traditional fully connected layer has poor performance due to the complexity and high dimension of sEMG signals caused by individual differences. As a kind of probabilistic generation model, DBN establishes a joint probability distribution between data and labels. The deep belief network (DBN) is composed of a backpropagation (BP) neural network and a depth Boltzmann machine (DBM), and the DBM consists of a series of restricted Boltzmann machines (RBM) stacked in series. Figure 7 illustrates the DBN architecture schematically. The hidden-layer unit of DBN is trained to capture the correlation of high-order data expressed in the visual layer [39], and then features are extracted and reconstructed from the data layer by layer [40,41], and the basic principle of DBN can mainly be defined as

$$P(v, h^1, h^2, \dots, h^L) = P(v|h^1) \cdot P(h^1|h^2) \dots P(h^{L-2}|h^{L-1}) \cdot P(h^{L-1}|h^L) \quad (5)$$

where v is visible layer; h^i is the hidden layer of the number i ; $P(\cdot)$ is a probability distribution. In this way, the deep belief network extracts and reconstructs data layer by layer and simplifies a complex input pattern, which can promote the network's ability to represent and generalize data. Meanwhile, the DBN can prevent the network from falling

into local optimum to some extent [42]. Thus, the deep belief network is introduced as a classification model to realize the bidirectional evaluation of the time–frequency features of sEMG and the corresponding rehabilitation movement type. Meanwhile, this model can also be applied to capture the intrinsic characteristics of sEMG signals.

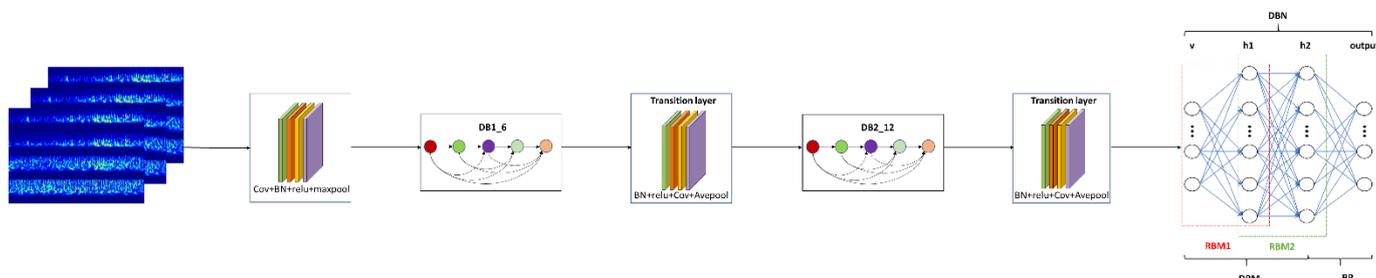


Figure 7. This is the main work framework of this paper, which consists of data representation (MTFIFR) and a learning model (MFFN). The structure of the MFFN, in this paper, is mainly composed of two dense blocks, two transition layers, and a deep belief network.

2.4. Designing Architecture of MFFN

In this paper, the framework of DenseNet is improved to realize the extraction and representation of relevant features for sEMG signals through the convolution operation in the time–frequency domain. First of all, the properties of network input need to be taken into account before designing a feature extraction network. In terms of time–frequency domain features, the denoised sEMG signal, including $10,240 \times 4$ sample points, is processed by the MTFIFR method to obtain the RGB image, which contains the main time–frequency domain information of multichannel sEMG signals; and the RGB image fed into the convolutional layer is in the form of three-dimensional data. The characteristic of 3D convolution is that the number of channels to the input feature map is equal to the number of convolution kernel channels, and there is no need to slide in the channel direction. Therefore, this paper chose a 3D filter whose kernel size is the number of input channels to fuse the information of each channel, instead of a small size convolution of channels. Secondly, convoluting across time–frequency domain information about sEMG signals will lead to deep learning of the dynamic properties and frequency components hidden in sEMG signals. In this paper, time–frequency domain convolution is realized by the optimized DenseNet. Third, in the classification stage, in order to heighten the recognition algorithm’s performance for various rehabilitation motions, the time–frequency information of sEMG signals is reconstructed through neurons to capture the essential attributes of sEMG and establish a probabilistic approximation model between sEMG and the corresponding rehabilitation movement by the deep belief network.

As shown in Figure 7, the inputs to the network are first convolved through the channels and then placed into the modified DenseNet. Within each dense block, there are a certain number of dense block layers, which are responsible for fusing the features of sEMG signals at each learning stage and splicing together the learned time–frequency features of sEMG signals. Meanwhile, in order to prevent too large or too small a range of features from affecting the convergence speed of the network, the normalization layers and activation layers are added after the network input layer [43–45]. Subsequently, the time–frequency features of the sEMG signals are concatenated and sent to the transition layer, which controls the feature map’s size and considers the compatibility of the feature map with the computation burden of the network [46]. For the purpose of acquiring more comprehensive features about rehabilitation actions [47], the above operations are performed repeatedly to gain the deeper features of sEMG signals. After repeating this phase, the feature map is sent to the deep belief network for the identification of upper-limb movements. Finally, various parts in the MFFN such as the convolution layer, dense block module, transition

layer, and deep belief network structure are selected by the cross-validation method, which will be demonstrated and discussed in the experimental results.

3. Experimental Results and Discussion

For experiments, the signal preprocessing and network construction that were carried out in MATLAB 2020b environment used 16 GB RAM and CPU of Intel® Core™ i7-10750H CPU @2.60 GHz 2.59 GHz. As for deep learning, the NVIDIA GeForce GTX 1660 Ti GPU with 7.9 GB of RAM was employed. The whole process is supported by the Windows 10 operating system.

3.1. Parameter Selection

Each sample has a total of $10,240 \times 4$ sample points. Therefore, the appropriate input size was firstly selected for the network framework. Then, the key parameters of MFFN such as the convolution kernel size, the number of dense blocks, and DBN structure were optimized by cross-validation. Finally, the training parameters such as learning rate and iteration number were adjusted by the grid searching method, and the cross-entropy and Adam were chosen as the loss function and optimization algorithm of the MFFN. Based on the above analysis, parameter selection and tuning of the MFFN were executed, and detailed information is shown in Table 1.

Table 1. This is the network configuration of the proposed architecture (MFFN).

Layers	Output Size	Multiple Feature Fusion Network
Input Layer	$224 \times 224 \times 3$	–
Convolution	$112 \times 112 \times 64$	filtersize: [7,7], numfilters: 64, stride 2, padding: [3,3,3,3]
BN	$112 \times 112 \times 64$	channels: 64
ReLu	$112 \times 112 \times 64$	–
Pooling	$56 \times 56 \times 64$	[3,3] max pool, stride: [2,2]
Dense Block (1)	$56 \times 56 \times 256$	[BN, ReLu, 1×1 conv, BN, ReLu, 3×3 conv] $\times 6$
Transition Layer (1)	$56 \times 56 \times 256$	BN
	$56 \times 56 \times 256$	ReLu
	$56 \times 56 \times 128$	[1,1] conv
	$28 \times 28 \times 128$	[2,2] average pool, stride 2
Dense Block (2)	$28 \times 28 \times 512$	[BN, ReLu, 1×1 conv, BN, ReLu, 3×3 conv] $\times 12$
Transition Layer (2)	$28 \times 28 \times 512$	BN
	$28 \times 28 \times 512$	ReLu
	$28 \times 28 \times 256$	[1,1] conv
	$14 \times 14 \times 256$	[2,2] average pool, stride 2
Classification Layer	1024×1	$1024 \times 50,176$ RBM
	1500×1	1500×1024 RBM
	12×1	12D Fully Connected, softmax

3.2. Experimental Comparison and Analysis

In terms of dataset splitting, this paper took the data of 12 rehabilitation movements for each subject as a whole, and then randomly selected all rehabilitation movement data for some of the subjects among all subjects as the test set, and the rest as the training set. Several experiments were carried out to evaluate the performance of the network framework. The experiment was mainly executed from two dimensions. In the vertical aspect, the performance of DBN and traditional fully connected layer as MFFN classification model was compared. In the horizontal aspect, the recognition performance of different network frameworks for 12 rehabilitation movements were contrasted, and the specific schemes and results are given subsequently. Finally, this paper analyzed and discussed the performance of the proposed method.

3.2.1. Longitudinal Contrast Experiment

In order to verify the ability of the DBN to capture time–frequency features of sEMG signals, the performance of traditional fully connected layer and DBN structure as a classification model of the MFFN were compared respectively; and the architectures were selected by the cross-validation method.

In Table 2, the experimental results show that the proposed method can effectively improve the accuracy of the intention recognition system in each test, which proves that the DBN has a better effect on the abstract representation and reconstruction of time–frequency features for sEMG signals. Furthermore, the accuracy of the DBN model with the MFFN is 1.64% higher than that of the fully connected layer in an average sense. Thus, the results indicate that the proposed DBN as the classification layer of MFFN is more effective in upper-limb movement identification.

Table 2. Comparison results of DBN and FC (fully connected) as a classification module of the MFFN.

Methods	MFFN_FC	MFFN_DBN
subject 1, subject 2	83.94%	86.10%
subject 3, subject 4	63.23%	65.11%
subject 5, subject 9	79.17%	79.37%
subject 6, subject 10	75.00%	77.78%
subject 3, subject 9	62.33%	65.89%
subject 1, subject 4	76.39%	77.39%
subject 7, subject 10	62.69%	62.78%
subject 2, subject 5	83.33%	85.10%
subject 6, subject 7	64.17%	66.56%
subject 7, subject 9	68.28%	68.83%
AVE	71.85%	73.49%

MFFN_FC: multiple feature fusion network proposed in this paper combined with fully connected layer.
MFFN_DBN: multiple feature fusion network proposed in this paper combined with deep belief network.

3.2.2. Transverse Contrast Experiment

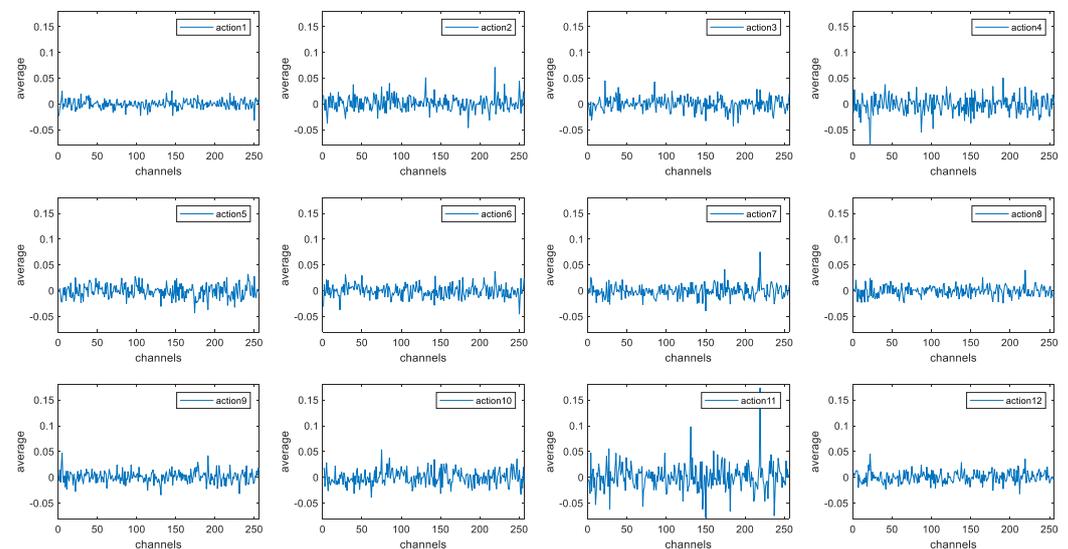
To validate the performance of the proposed time–frequency representation method and the MFFN framework in motion intention classification, the proposed method was compared with traditional machine learning algorithms and common deep learning network frameworks. The identification results of different frameworks and the proposed MFFN for 12 rehabilitation movements are shown in Table 3, and the average accuracy is better than in the previous approaches. The results of the first two columns are obtained by using traditional, manual-design time–frequency features and machine learning approaches; and the MFFN increases by 7.35% and 3.63% versus LDA (linear discriminant analysis) and SVM (support vector machine). It is obvious that the identification effect of the MFFN is more noticeable. Moreover, the first three methods in the table all use the time–frequency features of sEMG, while we choose the MTFIFR approach as the data representation. Therefore, compared with the previous three ways, the presented method retains more complete time–frequency features of sEMG signals, does not lose too much intermediate information in the convolution operation, and learns more relevant features. At the same time, the classification accuracy of the test for number 2, 5, 8 in the table are slightly less than in the LCNN (long short-term memory network and convolutional neural network) [26] method, but, overall, the accuracy of the other items was boosted, which illustrates the advantages of continuous wavelet features combined with MFFN in motion intention recognition. Compared with the end-to-end network frameworks including G_CNN (convolutional neural network with gate) and LCNN, the identification effect of the MFFN is superior, indicating that the proposed time–frequency representation method combined with MFFN is more effective in rehabilitation movement classification.

Table 3. Comparison results of different frameworks for recognizing 12 rehabilitation movements.

Methods	LDA [14]	SVM [48]	G_CNN [49]	LCNN [26]	Proposed Method
subject 1, subject 2	80.26%	85.03%	85.29%	80.11%	86.10%
subject 3, subject 4	52.43%	60.29%	62.59%	65.29%	65.11%
subject 5, subject 9	70.21%	72.98%	78.39%	75.22%	79.37%
subject 6, subject 10	65.16%	69.26%	75.00%	70.92%	77.78%
subject 3, subject 9	59.98%	65.53%	65.39%	65.98%	65.89%
subject 1, subject 4	70.56%	72.20%	75.65%	74.61%	77.39%
subject 7, subject 10	60.92%	60.89%	61.85%	61.79%	62.78%
subject 2, subject 5	78.23%	78.29%	80.59%	85.49%	85.10%
subject 6, subject 7	62.51%	64.93%	65.29%	63.46%	66.56%
subject 7, subject 9	61.09%	69.22%	67.92%	62.34%	68.83%
AVE	66.14%	69.86%	71.80%	70.52%	73.49%

3.2.3. Analysis and Discussion

For analyzing the performance of the MFFN, we first extracted the features of Transition Layer (2) for the proposed feature extraction framework and carried out an average operation on each feature map channel. The visualization results are shown in Figure 8, and the data distribution of each action from action1 to action12 has no significant mutation and remains within a certain range, which proves that the MFFN can extract stable time–frequency features of sEMG signals. At the same time, the data of action1, action2, and other actions have great differences in the amplitude change process, and the corresponding rehabilitation movements can be well distinguished through the dynamic change process of the data, illustrating that MFFN can better extract the useful time–frequency features of complex sEMG signals.

**Figure 8.** The extraction effect of multiple feature fusion network for the time–frequency features of sEMG signals corresponding to 12 kinds of upper-limb rehabilitation movements.

To evaluate the mutation of the extracted feature for each class of action and to measure the distinctions in Figure 8, first, the statistics of quartile difference (QD), mean difference (MD), and variance (STD) were introduced to evaluate the mutation of the extracted feature for each class of action on the basis of Figure 8, in which,

$$QD = Q_3 - Q_1 \quad (6)$$

$$MD = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} \quad (7)$$

$$STD = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (8)$$

where Q_1 is the next quartile; Q_3 denotes the upper quartile; x_i , \bar{x} , and n represent the i -th element, average, and length of the data, respectively. The greater the three above, the greater the fluctuation in the data and vice versa. The results are shown in Figure 9 as follows. On the whole, the trend of three statistics describing the mutation of 12 kinds of action features is consistent. At the same time, it can be seen that the corresponding statistical values of action2, action4, and action11 are larger, and those of some actions are smaller, e.g., action1 and action8. This shows that there are great differences in the process of feature mutation corresponding to the 12 types of actions, which indicates that the features extracted by the proposed method are beneficial to the classification of multiple actions from the aspect of mutation degree.

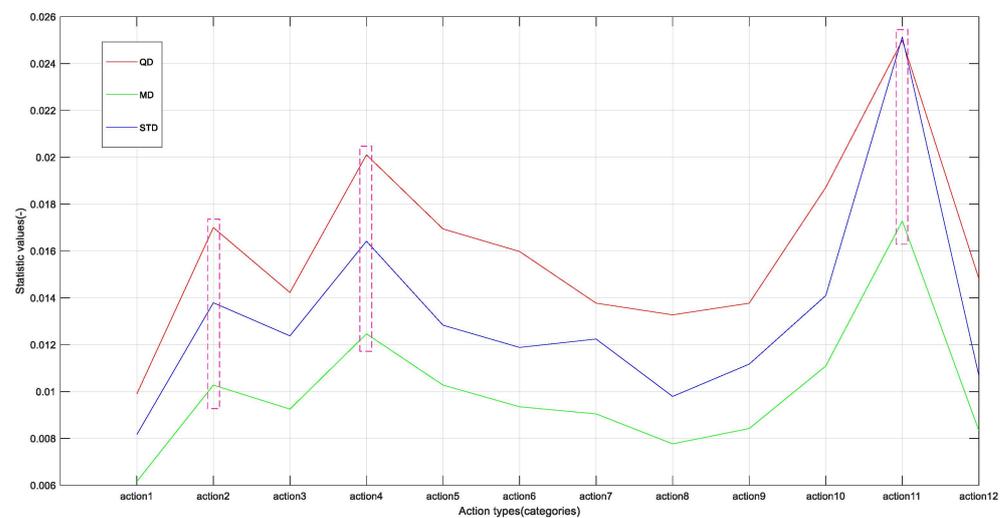


Figure 9. Evaluating the mutation of the extracted feature for each class action by statistics (QD, MD, and STD).

Then, the correlation coefficient method was used to analyze the correlation of the features extracted from the 12 types of actions to observe the distinctions among actions. The expression of correlation coefficient is

$$\rho_{X,Y} = \frac{con(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - EX)(Y - EY)]}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \quad (9)$$

where $con(\cdot)$ is the covariance, σ is the standard deviation, and E denotes the mathematical expectation; the correlation coefficient takes values in the range $[-1,1]$, and in general, the closer the correlation coefficient is to 1, the more positive the correlation is; the closer it is to -1 , the more negative the correlation is; and the closer it is to 0, the less relevant it is. The results are shown in Figure 10. It is obvious that the correlation coefficients in red marked the maximum (the correlation coefficient between action4 and action6) and the closest to zero (the correlation coefficient between action5 and action11) are 0.6687 and -0.0062 , respectively, indicating that there is a certain correlation between action4 and action6, and there is a very weak correlation or no correlation between action5 and action11.

	action1	action2	action3	action4	action5	action6	action7	action8	action9	action10	action11	action12
action1	1	0.0568	0.0324	0.0381	-0.2610	0.3616	0.2810	-0.0273	0.3683	-0.0709	0.1097	0.0829
action2	0.0568	1	0.4790	-0.2176	-0.0390	-0.0797	0.1919	0.2288	0.0619	0.3726	0.3943	0.1947
action3	0.0324	0.4790	1	0.1084	0.3556	0.1697	0.0974	0.3097	-0.1305	0.2713	0.3707	0.0362
action4	0.0381	-0.2176	0.1084	1	0.2780	0.6687	0.3518	0.1632	0.1728	-0.0146	0.3881	-0.3922
action5	-0.2610	-0.0390	0.3556	0.2780	1	0.0568	-0.2105	0.2364	-0.3513	0.1541	-0.0062	-0.1973
action6	0.3616	-0.0797	0.1697	0.6687	0.0568	1	0.4145	0.0718	0.2918	-0.1193	0.3400	-0.3682
action7	0.2810	0.1919	0.0974	0.3518	-0.2105	0.4145	1	0.4254	0.3456	0.1266	0.6548	0.0501
action8	-0.0273	0.2288	0.3097	0.1632	0.2364	0.0718	0.4254	1	-0.0903	0.2136	0.3954	0.0180
action9	0.3683	0.0619	-0.1305	0.1728	-0.3513	0.2918	0.3456	-0.0903	1	0.3133	0.0265	0.0806
action10	-0.0709	0.3726	0.2713	-0.0146	0.1541	-0.1193	0.1266	0.2136	0.3133	1	0.0215	0.2894
action11	0.1097	0.3943	0.3707	0.3881	-0.0062	0.3400	0.6548	0.3954	0.0265	0.0215	1	0.1542
action12	0.0829	0.1947	0.0362	-0.3922	-0.1973	-0.3682	0.0501	0.0180	0.0806	0.2894	0.1542	1

Figure 10. The correlation coefficient matrix among the 12 types of actions.

To further compare the distinctions in features across action1–action12, the distribution of the correlation coefficients was visualized by means of a box plot, which is a method of describing the distribution of data using five statistical values (minimum, first quartile, median, third quartile, maximum, accompanied by outlier points). The horizontal axis indicates the action types, and the vertical axis represents the distribution of the correlation coefficients, where the green part denotes the correlation coefficient between each action class and itself, with a constant value equal to 1; each box means the correlation among the current action class and the rest. The corresponding results are shown in Figure 11, it can be seen clearly that the overall correlation coefficient is mainly distributed between -0.2 and 0.43 . At the same time, there are significant differences in the position of each box, which shows that there is a relatively weak correlation among 12 kinds of actions, that is, there are big distinctions among the features of different types of actions, which proves the effectiveness of the proposed method in this paper.

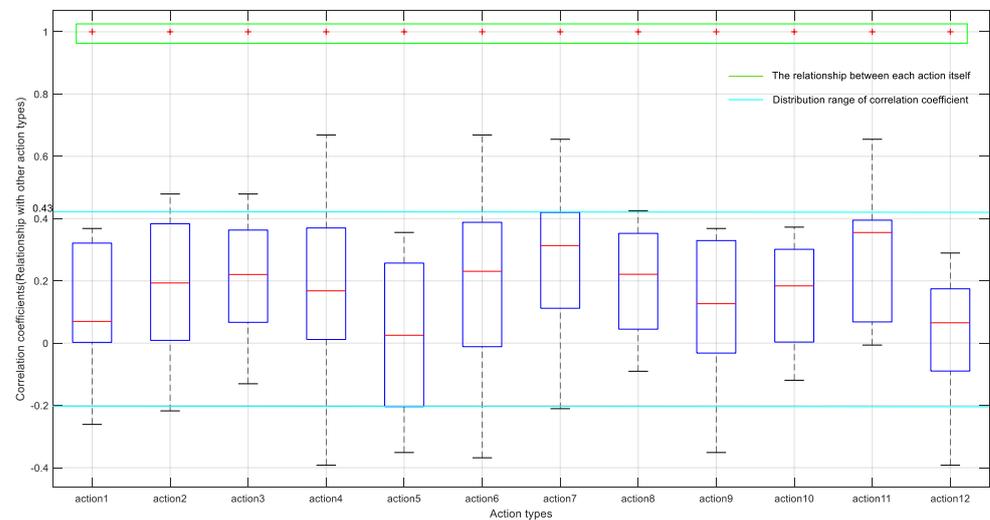


Figure 11. Box plot of correlation coefficient distribution among 12 action types.

Then, this paper further analyzed the learning effect of DBN as the classification model of the MFFN. As shown in Figure 12, the output features of the DBN model are reduced in dimension and visualization by the t-SNE (t-Distributed Stochastic Neighbor Embedding, which returns a matrix of two-dimensional embeddings of the high-dimensional rows of the input) approach. It can be observed in the distribution of features that although there is a small amount of data overlap, the 12 classes can be well distinguished. For example, there is a large distance among action1, action2, and action6, which indicates that DBN introduced as the classification model of the MFFN can effectively capture the time–frequency features of sEMG signals and perform well in recognition of more types of rehabilitation actions.

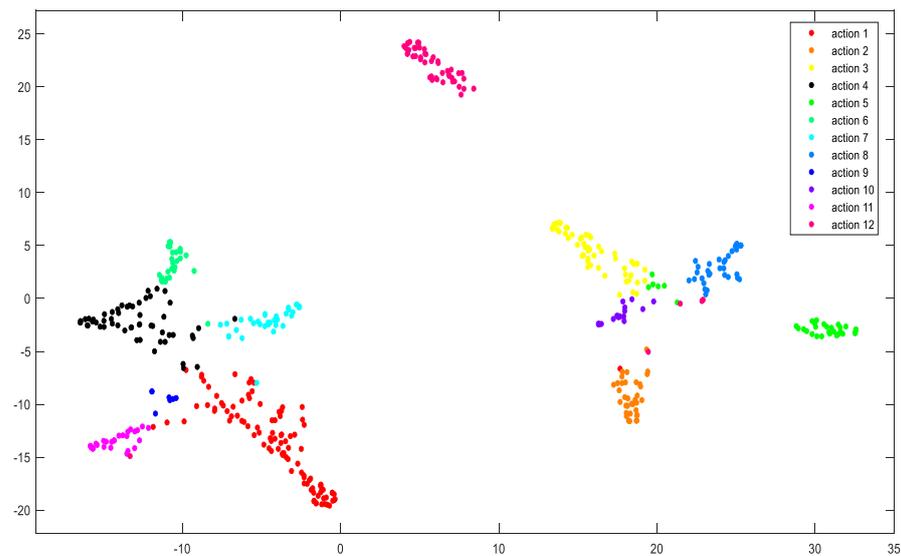


Figure 12. Feature distribution of 12 types of rehabilitation motions, obtained by visualizing the last layer features of DBN using the t-SNE method.

Finally, the identification results of 12 rehabilitation exercises were presented in the form of a confusion matrix, as shown in Figure 13. It is relatively serious that action4 was mistaken as action6, and the main reason may derive from the high similarity of movements that are mainly driven by the same muscles. As a result, the time–frequency features of sEMG signals are not obvious enough to distinguish these two actions, which also verifies the overlap of a small number of features in Figure 12. However, most actions, such as action1, action2, action6, action7, action10, and action11, can be classified well, which demonstrates that the proposed time–frequency representation method combined with MFFN is more effective in recognizing more types of rehabilitation movements.

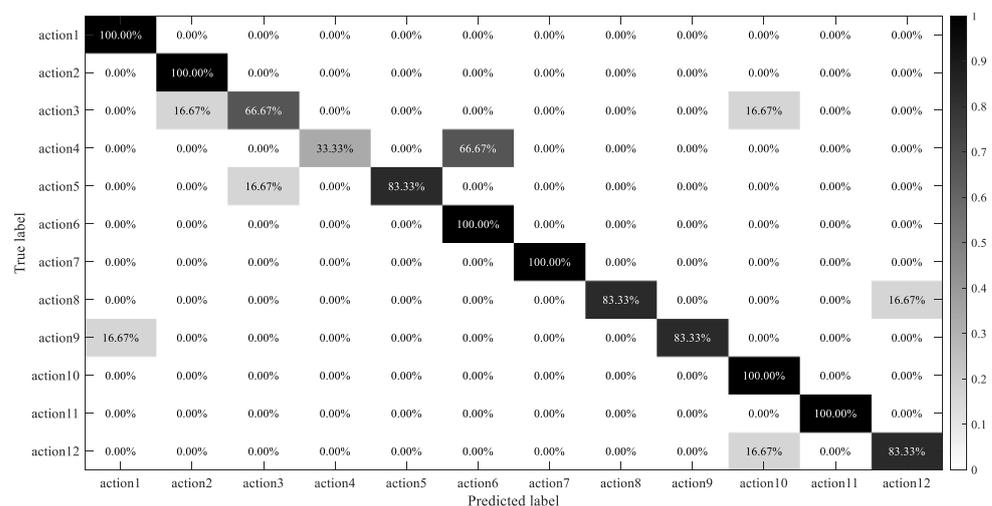


Figure 13. Confusion matrix of motion classification.

4. Conclusions and Future Outlook

In this research, first, the data representation method can fuse more complete time–frequency dynamic information about sEMG signals. At the same time, most of the invalid frequency domain information of the sEMG signal can be effectively suppressed, and the valid frequency domain information can be well preserved, which is beneficial to boost the learning efficiency. Then, this paper mainly proposes the application of a multiple feature fusion network in sEMG signals. The MFFN does not learn the time–

frequency features of sEMG signals from the depth and breadth of the architecture but strengthens the network through reusing or fusing features. The feature loss in each part of the network and the number of parameters are reduced. Finally, the deep belief network is introduced as a classification model of the MFFN in the classification stage, which can realize the bidirectional evaluation of the time–frequency features of sEMG and the corresponding rehabilitation movement. Meanwhile, the deep belief network can capture the intrinsic properties of surface EMG signals, which is conducive to improving the generalization and accuracy of the multiclassification system. The final experimental results and visualization analysis prove that the multiple feature fusion network can effectively distinguish more types of upper-limb rehabilitation movements. Compared to SVM and other CNN frameworks, the multiple feature fusion network can produce better results, demonstrating that the multiple feature fusion network structures make sense for sEMG signals.

However, in the current architecture of the network there mainly exists a problem that needs to be verified and solved, that is, there may be redundant features in the process of time–frequency feature extraction from sEMG signals. To sum up, the MFFN provides an effective network framework for extracting more stable and valuable features from sEMG signals. In the future, we will integrate the above issue further to enhance the performance of the feature fusion network framework and apply it to the field of rehabilitation medicine.

Author Contributions: Conceptualization, T.Z. and J.W.; data curation, T.Z.; methodology, T.Z.; software, T.Z.; formal analysis, T.Z.; validation, T.Z. and D.L.; investigation, T.Z.; resources, J.W. and Y.Z.; writing—original draft preparation, T.Z.; writing—review and editing, D.L., J.W., T.Z., J.X., and Z.A.; visualization, T.Z. and D.L.; supervision, J.W. and D.L.; project administration, J.W.; funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation of China grant number 61733003.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Oonishi, Y.; Oh, S.; Hori, Y. A New Control Method for Power-Assisted Wheelchair Based on the Surface Myoelectric Signal. *IEEE Trans. Ind. Electron.* **2010**, *57*, 3191–3196. [CrossRef]
2. Giho, J.; Choi, Y. EMG-based continuous control method for electric wheelchair. In Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, IL, USA, 14–18 September 2014; pp. 3549–3554.
3. Veneman, K.R.; Hekman, E.E.G.; Hori, Y.; Ekkelenkamp, R.; Van, E.H.F.; Van, H. Design and Evaluation of the LOPES Exoskeleton Robot for Interactive Gait Rehabilitation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 379–386. [CrossRef] [PubMed]
4. Kiguchi, K.; Hayashi, Y. An EMG-Based Control for an Upper-Limb Power-Assist Exoskeleton Robot. *IEEE Trans. Syst. Man Cyber.* **2012**, *42*, 1064–1071. [CrossRef] [PubMed]
5. Clearpath Robotics Drives Robot with Arm Motions. Available online: http://www.clearpathrobotics.com/press_release/drive-robot-with-arm-motion/ (accessed on 20 February 2014).
6. Kamali, T.; Boostani, R.; Parsaei, H. A Multi-Classifer Approach to MUAP Classification for Diagnosis of Neuromuscular Disorders. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 191–200. [CrossRef] [PubMed]
7. Nair, S.S.; French, R.M.; Laroche, D.; Thomas, E. The application of machine learning algorithms to the analysis of electromyographic patterns from arthritic patients. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2010**, *18*, 174–184. [CrossRef]
8. Xie, Q.; Meng, Q.; Zeng, Q.; Fan, Y.; Dai, Y.; Yu, H. Human-Exoskeleton Coupling Dynamics of a Multi-Mode Therapeutic Exoskeleton for Upper Limb Rehabilitation Training. *IEEE Access* **2021**, *9*, 61998–62007. [CrossRef]
9. Chen, S.H.; Lien, W.M.; Wang, W.W. Assistive Control System for Upper Limb Rehabilitation Robot. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *24*, 1199–1209. [CrossRef]
10. Wang, W.Q.; Hou, Z.G.; Cheng, L. Toward Patients' Motion Intention Recognition: Dynamics Modeling and Identification of iLeg-An LLRR Under Motion Constraints. *IEEE Trans. Syst. Man Cyber. Syst.* **2016**, *46*, 980–992. [CrossRef]
11. Chu, J.; Moon, I.; Lee, Y.; Kim, S.; Mun, M. A Supervised Feature-Projection-Based Real-Time EMG Pattern Recognition for Multifunction Myoelectric Hand Control. *IEEE/ASME Trans. Mecha.* **2007**, *12*, 282–290. [CrossRef]

12. Duan, F.; Dai, L. Recognizing the Gradual Changes in sEMG Characteristics Based on Incremental Learning of Wavelet Neural Network Ensemble. *IEEE Trans. Ind. Electron.* **2017**, *64*, 4276–4286. [[CrossRef](#)]
13. Xiao, F.; Wang, Y.; He, L.; Wang, H.; Li, W.; Liu, Z. Motion Estimation from Surface Electromyogram Using Adaboost Regression and Average Feature Values. *IEEE Access* **2019**, *7*, 13121–13134. [[CrossRef](#)]
14. He, C.; Zhuo, T.; Ou, D.; Liu, M.; Liao, M. Nonlinear Compressed Sensing-Based LDA Topic Model for Polarimetric SAR Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 972–982. [[CrossRef](#)]
15. Shi, W.T.; Lyu, Z.J.; Tang, S.T.; Chia, T.L.; Yang, C.Y. A bionic hand controlled by hand gesture recognition based on surface EMG signals: A preliminary study. *Biocyber. Biomed. Eng.* **2017**, *38*, 126–135. [[CrossRef](#)]
16. Ding, H.J.; He, Q.; Zeng, L.; Zhou, Y.J.; Shen, M.M.; Dan, G. Motion intent recognition of individual fingers based on mechanomyogram. *Patt. Recogn. Lett.* **2017**, *88*, 41–48. [[CrossRef](#)]
17. Khezri, M.; Jahed, M. A Neuro-Fuzzy Inference System for sEMG-Based Identification of Hand Motion Commands. *IEEE Trans. Ind. Electron.* **2011**, *58*, 1952–1960. [[CrossRef](#)]
18. Chu, J.U.; Moon, I.; Mun, M.S. A Real-Time EMG Pattern Recognition System Based on Linear-Nonlinear Feature Projection for a Multifunction Myoelectric Hand. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 2232–2239. [[PubMed](#)]
19. Tsujimura, T.; Hashimoto, T.; Izumi, K. Genetic reasoning for finger sign identification based on forearm electromyogram. *Int. Conf. Appl. Electron.* **2014**, *7*, 297–302.
20. Ajiboye, A.B.; Weir, R.F. A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthesis control. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2005**, *13*, 280–291. [[CrossRef](#)]
21. Momen, K.; Krishnan, S.; Chau, T. Real-Time Classification of Forearm Electromyographic Signals Corresponding to User-Selected Intentional Movements for Multifunction Prosthesis Control. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 535–542. [[CrossRef](#)]
22. Geethanjali, P.; Ray, K.K. A Low-Cost Real-Time Research Platform for EMG Pattern Recognition-Based Prosthetic Hand. *IEEE/ASME Trans. Mech.* **2015**, *20*, 1948–1955. [[CrossRef](#)]
23. Kiguchi, K.; Tanaka, T.; Fukuda, T. Neuro-fuzzy control of a robotic exoskeleton with EMG signals. *IEEE Trans. Fuzzy Syst.* **2004**, *12*, 481–490. [[CrossRef](#)]
24. Yang, Y.K.; Duan, F.; Ren, J.; Xue, J.N.; Lv, Y.Z.; Zhu, C.; Yokoi, H.S. Performance Comparison of Gesture Recognition System Based on Different Classifiers. *IEEE Trans. Cognit. Dev. Syst.* **2021**, *13*, 141–150. [[CrossRef](#)]
25. Duan, F.; Dai, L.; Chang, W.; Chen, Z.; Zhu, C.; Li, W. sEMG-Based Identification of Hand Motion Commands Using Wavelet Neural Network Combined with Discrete Wavelet Transform. *IEEE Trans. Ind. Electron.* **2016**, *63*, 1923–1934. [[CrossRef](#)]
26. Wei, W.; Wong, Y.; Du, Y.; Hu, Y.; Kankanhalli, M.; Geng, W. A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface. *Patt. Recogn. Lett.* **2019**, *119*, 131–138. [[CrossRef](#)]
27. Azami, H.; Hassanpour, H.; Escudero, J.; Sanei, S. An intelligent approach for variable size segmentation of non-stationary signals. *J. Adv. Res.* **2015**, *6*, 687–698. [[CrossRef](#)]
28. Ercan, G.; Abdulhamit, S. Comparison of decision tree algorithms for EMG signal classification using DWT. *Biomed. Signal Process. Control.* **2015**, *18*, 138–144.
29. Wu, Y.; Zheng, B.; Zhao, Y. Dynamic Gesture Recognition Based on LSTM-CNN. In Proceedings of the 2018 Chinese Automation Congress (CAC), Xi'an, China, 30 November–2 December 2018; pp. 2446–2450.
30. Alkan, A.; Günay, M. Identification of EMG signals using discriminant analysis and SVM classifier. *Expert Syst. Appl.* **2012**, *39*, 44–47. [[CrossRef](#)]
31. Godfrey, A.; Conway, R.; Leonard, M.; Meagher, D.; O'laighin, G.M. A Continuous Wavelet Transform and Classification Method for Delirium Motoric Subtyping. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2009**, *17*, 298–307. [[CrossRef](#)]
32. Shi, X.; Qin, P.; Zhu, J.; Zhai, M.; Shi, W. Feature Extraction and Classification of Lower Limb Motion Based on sEMG Signals. *IEEE Access* **2020**, *8*, 132882–132892. [[CrossRef](#)]
33. De, C.J.; Donald, L.; Mikhail, K.; Roy, S.H. Filtering the surface EMG signal: Movement artifact and baseline noise contamination. *J. Biomech.* **2010**, *43*, 1573–1579.
34. Potvin, J.R.; Brown, S.H.M. Less is more: High pass filtering, to remove up to 99% of the surface EMG signal power, improves EMG-based biceps brachii muscle force estimates. *J. Electromyogr. Kinesi.* **2004**, *14*, 389–399. [[CrossRef](#)]
35. Li, D.L.; Wang, J.H.; Xu, J.C.; Fang, X.K. Densely Feature Fusion Based on Convolutional Neural Networks for Motor Imagery EEG Classification. *IEEE Access* **2019**, *7*, 132720–132730. [[CrossRef](#)]
36. Lilly, J.M.; Olhede, S.C. Generalized Morse Wavelets as a Superfamily of Analytic Wavelets. *IEEE Trans. Signal Process.* **2012**, *60*, 6036–6041. [[CrossRef](#)]
37. Lilly, J.M.; Olhede, S.C. Higher-Order Properties of Analytic Wavelets. *IEEE Trans. Signal Process.* **2009**, *57*, 146–160. [[CrossRef](#)]
38. Gao, H.; Liu, Z.; Laurens, V.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.
39. Sarikaya, R.; Hinton, G.E.; Deoras, A. Application of Deep Belief Networks for Natural Language Understanding. *IEEE/ACM Trans. Aud. Spe. Lan. Process.* **2014**, *22*, 778–784. [[CrossRef](#)]
40. Hinton, G.E.; Osindero, S.; Teh, Y.W. A Fast Learning Algorithm for Deep Belief Nets. *Neural Comput.* **2006**, *18*, 1527–1554. [[CrossRef](#)]
41. Chen, Y.; Zhao, X.; Jia, X. Spectral-Spatial Classification of Hyperspectral Data Based on Deep Belief Network. *IEEE J. Sel. Top. Appl. Ear. Obser. Rem. Sens.* **2015**, *8*, 2381–2392. [[CrossRef](#)]

42. Pan, T.Y.; Tsai, W.L.; Chang, C.Y.; Yeh, C.W.; Hu, M.C. A Hierarchical Hand Gesture Recognition Framework for Sports Referee Training-Based EMG and Accelerometer Sensors. *IEEE Trans. Cyber.* **2020**, *10*, 1–12.
43. Kalayeh, M.M.; Shah, M. Training Faster by Separating Modes of Variation in Batch-Normalized Models. *IEEE Trans. Patt. Analy. Mach. Intelli.* **2020**, *42*, 1483–1500. [[CrossRef](#)]
44. Chen, Z.D.; Deng, L.; Li, G.Q.; Sun, J.W.; Hu, X.; Liang, L.; Ding, Y.F.; Xie, Y. Effective and Efficient Batch Normalization Using a Few Uncorrelated Data for Statistics Estimation. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 348–362. [[CrossRef](#)]
45. Banerjee, C.; Mukherjee, T.; Pasilio, E. Feature representations using the reflected rectified linear unit (RReLU) activation. *Big Data Min. Anal.* **2020**, *3*, 102–120. [[CrossRef](#)]
46. Deng, R.; Liu, S. Relative Depth Order Estimation Using Multiscale Densely Connected Convolutional Networks. *IEEE Access* **2019**, *7*, 38630–38643. [[CrossRef](#)]
47. Wang, Z.; Liu, G.; Tian, G. A Parameter Efficient Human Pose Estimation Method Based on Densely Connected Convolutional Module. *IEEE Access* **2018**, *6*, 58056–58063. [[CrossRef](#)]
48. Matsubara, T.; Morimoto, J. Bilinear Modeling of EMG Signals to Extract User-Independent Features for Multiuser Myoelectric Interface. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 2205–2213. [[CrossRef](#)]
49. Ye, Q.; Chu, M.; Grethler, M. Upper Limb Motion Recognition Using Gated Convolution Neural Network via Multi-Channel sEMG. In Proceedings of the 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), Shenyang, China, 22–24 January 2021; pp. 397–402.