

Article

# Spatial Feature Integration in Multidimensional Electromyography Analysis for Hand Gesture Recognition

Wensheng Chen , Yinxi Niu, Zhenhua Gan, Baoping Xiong \* and Shan Huang \*

School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350116, China; ichtenws@163.com (W.C.); 2211301016@smail.fjut.edu.cn (Y.N.); ganzh@fjut.edu.cn (Z.G.)

\* Correspondence: xiongbp@fjut.edu.cn (B.X.); 19872162@fjut.edu.cn (S.H.)

**Abstract:** Enhancing information representation in electromyography (EMG) signals is pivotal for interpreting human movement intentions. Traditional methods often concentrate on specific aspects of EMG signals, such as the time or frequency domains, while overlooking spatial features and hidden human motion information that exist across EMG channels. In response, we introduce an innovative approach that integrates multiple feature domains, including time, frequency, and spatial characteristics. By considering the spatial distribution of surface electromyographic electrodes, our method deciphers human movement intentions from a multidimensional perspective, resulting in significantly enhanced gesture recognition accuracy. Our approach employs a divide-and-conquer strategy to reveal connections between different muscle regions and specific gestures. Initially, we establish a microscopic viewpoint by extracting time-domain and frequency-domain features from individual EMG signal channels. We subsequently introduce a macroscopic perspective and incorporate spatial feature information by constructing an inter-channel electromyographic signal covariance matrix to uncover potential spatial features and human motion information. This dynamic fusion of features from multiple dimensions enables our approach to provide comprehensive insights into movement intentions. Furthermore, we introduce the space-to-space (SPS) framework to extend the myoelectric signal channel space, unleashing potential spatial information within and between channels. To validate our method, we conduct extensive experiments using the Ninapro DB4, Ninapro DB5, BioPatRec DB1, BioPatRec DB2, BioPatRec DB3, and Mendeley Data datasets. We systematically explore different combinations of feature extraction techniques. After combining multi-feature fusion with spatial features, the recognition performance of the ANN classifier on the six datasets improved by 2.53%, 2.15%, 1.15%, 1.77%, 1.24%, and 4.73%, respectively, compared to a single fusion approach in the time and frequency domains. Our results confirm the substantial benefits of our fusion approach, emphasizing the pivotal role of spatial feature information in the feature extraction process. This study provides a new way for surface electromyography-based gesture recognition through the fusion of multi-view features.



**Citation:** Chen, W.; Niu, Y.; Gan, Z.; Xiong, B.; Huang, S. Spatial Feature Integration in Multidimensional Electromyography Analysis for Hand Gesture Recognition. *Appl. Sci.* **2023**, *13*, 13332. <https://doi.org/10.3390/app132413332>

Academic Editor: Flavia Ravelli

Received: 5 November 2023

Revised: 27 November 2023

Accepted: 14 December 2023

Published: 18 December 2023

**Keywords:** EMG; multi-feature fusion; spatial feature information

## 1. Introduction

Electromyographic (EMG) signals are the physiological signals which are generated through muscular contractions during human movements. The signals consist of motor action potentials (MUAPs) generated by muscle fibers. Actually, EMG signals can be detected via surface or invasive method, rather than the two categories of EMG signals. The former is collected through electrodes above the skin, whilst the latter is collected through needle electrodes inserted in muscles [1,2]. Although the EMG signal obtained from the latter can accurately represent the current state of the muscles, it needs to invade and harm the human body, so surface EMG (sEMG) is generally used to further determine the current motion of muscles. Other sensors for human movement analysis, like inertial measurement units (IMUs) [3,4], cameras [5,6], near-infrared spectroscopy (NIRS) [7,8], force



**Copyright:** © 2023 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/>).

myography (FMG) [9,10], etc., are also capable of responding to human motion intent at the physical level. However, the above sensors only reflect the muscular state of human motion, but not the real movement intention of the subjects with limb abnormalities or neurological disorders.

The rich physiological information contained in sEMG signals enables the extraction of human motion intention data, thereby holding considerable practical value across various domains such as clinical medicine, human engineering, and rehabilitation medicine. Surface electromyography-based human–machine interaction plays a pivotal role in applications such as prosthetics [11], sign language recognition systems [12], and intelligent driving [1].

Precisely decoding extensive feature information from EMG signals has emerged as a pivotal aspect of human–machine interaction [13]. Two primary categories of sEMG-based human–machine interaction methods exist: one relies on constructing muscle models [14] and the other utilizes non-muscle model techniques [11]. The use of muscle–skeletal models for recognizing movement intent possesses inherent limitations and primarily applies to tasks with fewer degrees of freedom, often struggling to distinguish finer tasks like gesture recognition. On the other hand, non-muscle model methods leverage traditional machine learning and deep learning approaches for pattern recognition. The recognition process encompasses three fundamental stages: (1) data pre-processing, which entails noise removal from raw signals and segmentation of extended signals; (2) feature extraction, involving the derivation of temporal, spectral, time–frequency, spatial, or high-dimensional semantic features utilizing deep learning networks; (3) pattern recognition, encompassing the identification and classification of signals through deep learning networks or machine learning methodologies. It is imperative to emphasize the pivotal role played by feature extraction in this overall process.

Time-domain feature analysis [15], frequency-domain feature analysis [16], and time–frequency-domain feature analysis [17] stand out as the most commonly employed methods for feature extraction. Time-domain features are characteristics derived directly from the signal's amplitude with time as the independent variable. Due to their convenience and intuitive computation, time-domain features have gained widespread popularity in sEMG feature extraction [18]. In the time domain, sEMG signals are typically regarded as zero-mean stochastic signals with variance altering according to signal intensity [19]. Prominent time-domain features for surface electromyography signals include root mean square (RMS), zero crossing (ZC), waveform length (WL), variance (VAR), and mean absolute value (MAV). Frequency-domain features, on the other hand, involve spectral or power spectral (PS) features obtained through the application of fast Fourier transform (FFT) to the original sEMG signals, allowing for direct observation of the signal's frequency band distribution. Frequently employed surface electromyography frequency-domain features include PS, mean frequency (MNF), median frequency (MF), frequency ratio (FR), and autoregressive coefficients (AR). The fusion of time-domain and frequency-domain features enables the simultaneous assessment of the temporal and spectral aspects of sEMG signals, providing comprehensive insights into muscle physiological variations. Presently, primary methods for conducting time–frequency analysis of surface electromyography signals encompass short-time Fourier transform (STFT), Wigner–Ville transform (WVT), and wavelet transform. Traditional time-domain and frequency-domain surface electromyography analysis methods only describe time-domain features or only frequency-domain features, which may not fully reflect the characteristic information of sEMG signals. Nevertheless, time–frequency-domain feature analysis methods offer a means to overcome these constraints [20].

Electromyographic (EMG) signals, often characterized by temporal instability, present a unique challenge and opportunity in the domain of human movement analysis, while these signals exhibit temporal variations, they also retain essential spatial information across different sensor channels. Recent advancements in the field, such as the pioneering work by Xiong et al. [21], have harnessed this spatial consistency by introducing geometric methodologies that leverage the topological characteristics inherent to EMG signals. Their

approach emphasizes the fusion of spatial topological structural attributes with traditional time–frequency features. This fusion not only counters temporal feature instability but also significantly enriches the feature information crucial for precise and efficient EMG signal decoding.

In the quest for enhancing the decoding of EMG signals, the concept of feature-level fusion has gained prominence, especially in models reliant on multimodal data sources [22]. Traditional methods for analyzing surface EMG (sEMG) signals, often constrained to either temporal or spectral features alone, face limitations. To address this, researchers, such as Li et al., in their comprehensive review in 2020 [20], proposed an innovative approach incorporating time–frequency-domain feature analysis. This method has proven effective in mitigating temporal feature instability, thereby enhancing the overall feature information available for EMG signal decoding [20,21]. EMG signals' capability to classify human actions hinges on information gathered from various anatomical locations and the intrinsic features embedded within the signals. However, when information acquisition is restricted to a single channel, as noted in Parker's seminal work on motor control in 2004 [23], limitations arise. Each sEMG channel corresponds to the collective activity of numerous muscles, making it imperative to augment the number of channels. This augmentation allows for a finer dissection of the information contained within an individual channel, resulting in a more abundant resource of spatial and temporal information for subsequent feature extraction. This, in turn, contributes to the effectiveness and precision of EMG signal decoding.

While current methods have yielded promising results, they often overlook the spatial characteristics inherent in EMG channels. In this paper, we introduce a novel Space-to-space (SPS) framework for sEMG feature extraction. By adopting a spatial perspective, we integrate inter-channel correlations and employ a multi-view feature fusion approach to decode electromyographic signals for pattern recognition. Our model exhibits a significant performance improvement when compared to conventional methods.

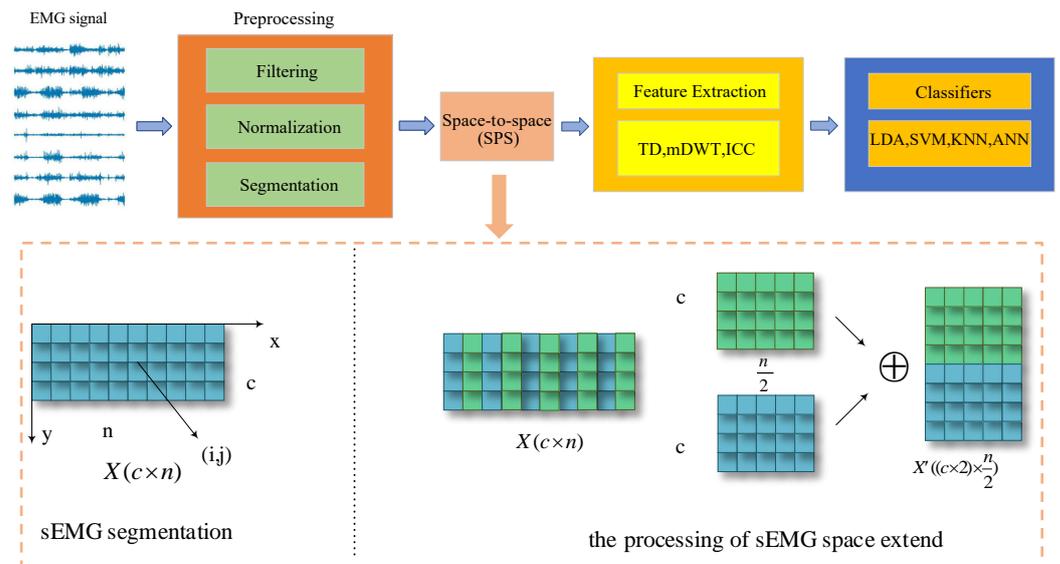
Our contribution to this work can be summarized as follows:

- This paper designs a new method for processing sEMG signals, which is based on expanding the original EMG signal segments in the space domain, thus releasing more feature information within a single channel to improve the accuracy of gesture action classification.
- A novel framework for sEMG-based gesture classification is established using inter-channel correlation (ICC) matrix representation extracted by the spatial sEMG information.
- Multi-information fusion involving time-domain and frequency-domain feature information with ICC matrix is implemented to build a pattern-recognition framework composed of multiple information.
- The effectiveness of the proposed approach is verified on hand gestures of six public datasets with various sEMG sensors and data formats, and a comprehensive analysis of the performance of multiple feature extraction combinations is completed.

## 2. SEMG Space Extend

This study proposes a space-to-space (SPS) framework for extracting features from sEMG signals, with the aim of describing specific signals by extracting structural characteristics from each sEMG channel. This framework is capable of extracting information from channel combinations, providing cross-channel patterns and background information on potential muscle synergy. The general procedure of the proposed method and the structure of SPS is shown in Figure 1.

In traditional methods, after signal down-sampling, a lot of information is lost along with the reduction in signal data. Among the lost information, there is a lot of potentially valuable information.



**Figure 1.** The general procedure of the proposed method. The bottom is the structure of space-to-space (SPS).

The extending operation mentioned in this paper is derived from the common down-sampling operation in images. For details, refer to SPD-Conv [22]. In ref. [22], the feature map of the original image is down-sampled to reduce the image resolution and avoid the loss of fine-grained information in the image. This paper refers to the idea of spatial conversion to depth and down-samples the original EMG segment according to time. Then, the down-sampled data are concatenated, releasing the time and space information in the EMG signal by transforming the EMG signal segment in space, and enriching the spatial domain information. The specific execution Equation (1) is as follows:

$$\begin{aligned}
 f_1 &= X[1 : c, 1 : scale : n], \\
 f_2 &= X[1 : c, 2 : scale : n], \\
 &\dots, \\
 f_{scale} &= X[1 : c, scale : scale : n]
 \end{aligned}
 \tag{1}$$

In Equation (1), the function  $f(\dots)$  signifies the specific down-sampling operation, extending the original signal segment into scale parts and ultimately merging them into the newly expanded sEMG signal segment. By transforming the original signal feature segment  $X(c \times n)$ , where  $n$  represents the length of the signal feature segments and  $c$  represents the number of signal channels, through signal folding, the original signal feature segment is transformed into  $X'((c \times scale) \times \frac{n}{scale})$ , where  $scale$  represents the extension factor (the extension factor can only be selected as a number that can be divided by  $n$ , and it is necessary to ensure that the folded signal feature segment  $\frac{n}{scale} > c \times scale$ ).

Taking  $scale = 2$  as an example shown in Figure 1, for a signal segment  $X(c \times n)$  with current time length  $n$  and channel count  $c$ , within  $f_1(\dots)$ , starting at time point 0 for each of the  $c$  channels, a sampling operation occurs at intervals of  $scale$  units until reaching the boundary of the time segment. Subsequent operations  $f_2(\dots), \dots, f_{scale}(\dots)$  mirror  $f_1(\dots)$ . By iteratively applying these steps and concatenating the down-scaled  $f(\dots)$ , the final result is the spatially expanded sEMG signal segment  $X'((c \times scale) \times \frac{n}{scale})$ , forming the basis for subsequent feature extraction.

For the extended EMG signal segment, extract the temporal and frequency domain and the spatial information between channels. It can more effectively extract the hidden feature information within a single channel, and then provide multi-perspective information for subsequent feature extraction.

### 3. Feature Extraction

One of the crucial factors determining classification and regression accuracy is the selection of appropriate feature representations. To maximize the representation of feature information within electromyographic signals, we propose a spatial domain feature-extraction method based on the covariance matrix. This method is combined with commonly used time-domain and frequency-domain approaches to extract more comprehensive feature information from electromyographic signals.

#### 3.1. Spatial Domain Feature Extraction

In mathematical statistics, covariance is commonly used as a parameter to investigate the relationship between two random variables, reflecting the second-order statistical properties between variables. Currently, there are limited studies utilizing covariance features to study EMG signal recognition issues [24,25]. The main reason is that EMG signals are one-dimensional time series, requiring the identification of an appropriate covariance matrix to explore the signal dynamics in high-dimensional space. Describing EMG information characteristics through the covariance matrix offers several advantages: it can integrate nonlinear feature information represented in the spatial domain; furthermore, the covariance matrix provides a method to filter out noise. By using spatial reconstruction to construct the covariance matrix from EMG signal sequences, potential signal dynamics information can be unearthed. An attempt is made to measure the similarity between covariance matrices using Euclidean distance, primarily achieved by vectorizing the matrices. Subsequently, through the use of classifiers, an effort is made to decode the underlying feature information and human kinematic information from higher dimensions, while restoring the spatial feature information between different channels of the EMG signals.

In this study, by constructing the inter-channel covariance matrix of the EMG signals, referred to as the inter-channel correlation (ICC) matrix, we aim to delve deeper into the hidden spatial feature information and human movement information between channels of the EMG signals. As shown in Figure 2, the spatial correlation of EMG signals is embedded within adjacent or correlated channels.

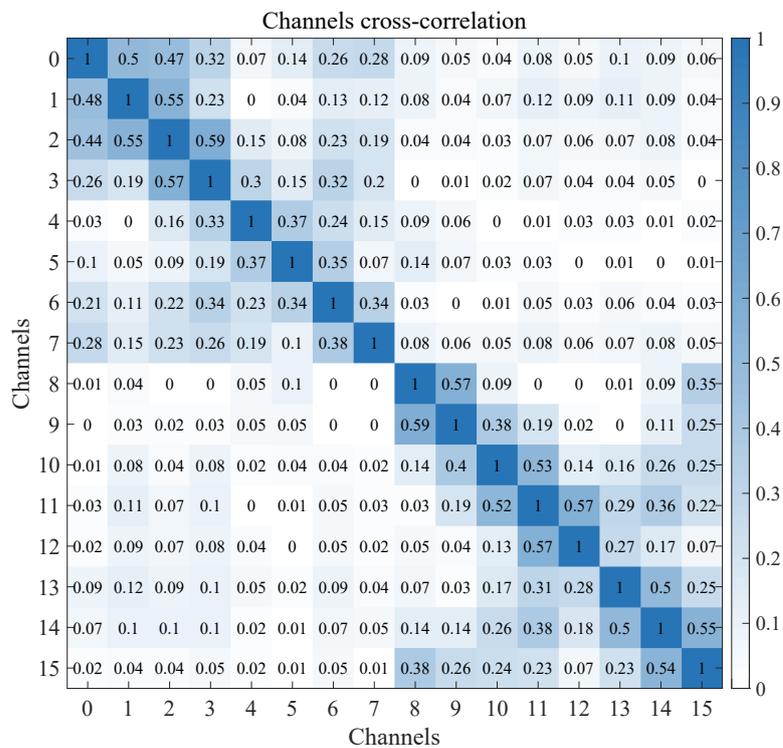


Figure 2. Example characteristics of EMG data. Cross correlation between channels of EMG data—note how many groups of channels have correlation.

The ICC matrix is utilized to learn spatial representation from EMG signals. We constructed the ICC matrix in the following, let  $x_t \in \mathbb{R}^c$  denote the sEMG signal vector at a specific time point  $t$ , with  $c$  denoting the number of recording channels. The ICC matrix is formally defined by  $\Sigma = E(x_t - E\{x_t\})(x_t - E\{x_t\})^T$ , where  $E\{\cdot\}$  denotes the expected value. In ICC matrix we consider short-time sEMG signal, in the form of a matrix  $X_i = [x_{(t+T_i)} \cdots x_{(t+T_i+T_n-1)}] \in \mathbb{R}^{c \times T_n}$ , which corresponds to the  $i$ -th segments of movement started at time  $t = T_i$ . Here,  $T_i$  denotes the number of sampled time points in each segment.

$$C_i = \frac{1}{T_n - 1} X_i * X_i^T \quad (2)$$

where  $X$  is EMG signal pre-processed using the procedure in Section 4.2. It is with shape  $(c \times n)$ , where  $n, c$  are the signal's fragment length and the number of channels. The ICC matrix is known to be an unbiased estimator of the covariance matrix  $\Sigma$  provided that the number of observations  $T_n$  is much than the number of variables  $c$  [23].

The ICC matrix is based on the covariance matrix using Equation (2), revealing spatial correlations among different channels and partially presenting the distribution characteristics of EMG signal data within each channel. Given that covariance reflects the distribution and variability of data, the ICC matrix not only captures spatial correlations among different channels but also characterizes the distribution of EMG signal vectors within the same channel. The variable  $X_i$  represents the  $i$ -th signal segment containing a signal vector of length  $n$  and  $c$  channels, the covariance matrix between different channels of this signal segment is computed using Equation (2). Here,  $T_n$  denotes the time length  $n$  of the signal vector for the respective signal segment. Ultimately, the covariance matrix for signal segment  $i$  is a symmetric matrix with dimensions  $(c \times c)$ . This matrix reflects the degree of correlation among different channels and also illustrates the data distribution within the same channel.

As shown in Figure 2, this study has chosen to employ the Pearson correlation coefficient matrix for analyzing the correlation among various channels of EMG signals. The Pearson correlation coefficient gauges the strength and direction of the linear relationship between two variables, calculated through the product of covariance and the standard deviations of the two variables. By leveraging the elements of the covariance matrix, we can determine the covariance between any two variables, subsequently deriving their Pearson correlation coefficient. This correlation assists in quantifying the linear relationships between variables, providing a comprehensive understanding of the associations between data structures and features. In sEMG analysis, it can be employed to investigate the synergy among different muscles, contributing to the understanding of physiological features related to movement and posture control.

### 3.2. Time Domain Feature Extraction

Common procedures are used to detect muscle activation. These are described by the observable lobes appearing in the sEMG time series. We choose the time-domain (TD) features set involving root mean square (RMS), mean absolute value (MAV), slope–sign change (SSC), waveform length (WL), and zero crossing (ZC) [26–28].

#### 3.2.1. Root Mean Square (RMS)

Root mean square is a commonly used time-domain feature of sEMG signals, which reflects the effective value of a sEMG signal segment. Furthermore, it indicates the contribution of various muscle tissues during the completion of gesture movement. It is defined by Equation (3):

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (3)$$

where  $x$  is amplitude in  $i$ th sample and  $N$  is total number of samples.

### 3.2.2. Mean Absolute Value (MAV)

The mean absolute value (MAV) feature represents one of the most direct measures of signal amplitude and is a commonly used time-domain feature. When the signal manifests in the form of a Laplacian random process, MAV serves as the maximum likelihood estimate of its signal amplitude. Its computational formula is as follows:

$$MAV = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (4)$$

Here,  $N$  represents the number of sample points in a given channel of the electromyographic signal, and  $x_i$  denotes the energy value (amplitude) of a signal sample point, where  $i = 1, 2, \dots, N$ .

### 3.2.3. slope–Sign Change (SSC)

Another feature that can offer insight into frequency content is the count of instances where the slope of the waveform changes its sign. Once more, an appropriate threshold must be selected to mitigate noise-induced fluctuations in slope.

$$SSC = \frac{1}{N} \sum_{i=2}^{N-1} f((x_{i+1} - x_i)(x_i - x_{i-1})) \quad (5)$$

$$f(x) = \begin{cases} 1, & \text{if } x > Th \\ 0, & \text{otherwise} \end{cases}$$

Considering three consecutive samples,  $x_{i-1}$ ,  $x_i$ , and  $x_{i+1}$ , the slope sign change count, denoted as SC, is incremented if  $x > Th$ ,  $f(x) = 1$ . This process is applied to three sequential segments using the threshold function Th to mitigate background noise in the EMG signal. This method captures frequency information from the time domain.

### 3.2.4. Waveform Length (WL)

Waveform length refers the signal feature obtained by accumulating the length of the signal energy (amplitude) within the number of data points and normalizing it. It reflects multiple physical quantities of signal such as action time, signal frequency, and amplitude.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (6)$$

### 3.2.5. Zero Crossing (ZC)

Zero crossing (ZC) is a metric of frequency information in the time domain of EMG signals. This feature can to some extent reflect the jitter and oscillation of the signal during a specific action segment. It also indicates the number of times the amplitude of the EMG signal crosses the zero level. To avoid capturing low-voltage fluctuations or background noise, a threshold condition is added. The specific calculation process is as follows:

$$ZC = \frac{1}{N} \sum_{i=1}^{N-1} f_1(x_{i+1} * x_i) \cap f_2(x_{i+1} - x_i) \quad (7)$$

$$f_1(x) = \begin{cases} 1, & \text{if } x < Th \\ 0, & \text{otherwise} \end{cases}$$

$$f_2(x) = \begin{cases} 1, & \text{if } |x| > Th \\ 0, & \text{otherwise} \end{cases}$$

where  $f_1(\dots)$  indicates whether  $x_{i+1}$  and  $x_i$  have opposite signs. Only when they are opposite and below the threshold can it be counted as 1.  $f_2(\dots)$  indicates that  $x_{i+1}$  and  $x_i$  are not the same point, and there is a distance between them that surpasses the threshold, only then can it be counted as 1. In the end, the two functions are intersected to count the number of zero crossings.

### 3.3. Time–Frequency-Domain Feature Extraction

Commonly employed techniques for extracting time–frequency features include the Fourier transform (FT) and wavelet transform (WT). FT, owing to its inherent characteristics, exhibits inefficiency when processing non-stationary EMG signal data and offers only a single resolution, alongside other limitations [27]. Consequently, we opt for an analytical approach with increased resolution and adaptability to data length. In the spectral analysis of biological signals, WT proves more suited than FT. WT's proficiency in feature extraction can be attributed to the scale and translation properties of its mother wavelet, which enhances its ability to effectively process and analyze the amplitude, frequency, and time characteristics of the signal [29]. As an exemplar of the WT family, we select the marginal discrete wavelet transform (mDWT), which does not necessitate wavelet time instants. Instead, it relies on the energy accumulated within a signal segment, rendering it a superior discriminant criterion [29].

#### Marginal Discrete Wavelet Transform (mDWT)

Discrete Wavelet Transform (DWT) discretizes the scale and translation of the basic wavelet. The basis functions of any wavelet transform fundamentally involve scaling and translating the mother wavelet and father wavelet (also known as the scaling function). Here,  $x_t$  represents the muscle EMG signal segment at time  $t$ , which undergoes transformation and decomposition using the wavelet basis  $\psi_{l,\tau}(t)$ . This process separates the high-frequency and low-frequency components of the signal. After undergoing multilevel decomposition, both the frequency-domain and time-domain features of the signal are adequately represented.

$$\hat{x}_l = \sum_{\tau=0}^{T/2^l-1} \sum_{t=1}^T x_t \psi_{l,\tau}(t) \quad (8)$$

$$\psi_{l,\tau}(t) = 2^{-\frac{1}{2}} \psi(2^{-l}t - \tau) \quad (9)$$

where  $\psi_{l,\tau}(t)$  represents the wavelet basis. For this study, the Daubechies Wavelet basis is chosen for decomposing the EMG signal. Here,  $x$  signifies the variable, where  $\tau$  takes values from 0 to  $2^l - 1$  inclusive. The parameter  $l$  denotes the scale factor, altering which scales the function values.  $\tau$  is the translation parameter, which when modified, shifts the function along the  $t$ -direction.

## 4. Experimental Setup

### 4.1. Dataset Description

This section provides an introduction to the six public datasets used in this paper, Ninapro DB4 [30], Ninapro DB5 [30], BioPatRec DB1 [31], BioPatRec DB2 [31], BioPatRec DB3 [31], and Mendeley Data [32]. The above dataset captures the hand and wrist movement gestures. As indicated in [33], the hand gestures are distinguished between static and dynamic. The primary distinction lies in the fact that static movements do not alter the finger angles, whereas dynamic movements involve changes in finger angles as the hand moves. Based on this description, the motions within the selected datasets exclusively fall under the category of static hand gestures. Here, is a brief overview of each analyzed dataset:

1. Ninapro DB4: In the dataset, the sEMG signals are collected from 10 subjects. Each subject is recorded 53 motions and divided into three groups: exercises A, exercises B and exercises C. The exercises A consists of basic fingers movements. The exercises B

includes the multiple finger flexion and extensor with wrist gestures. The exercises C set consists of grasping the common daily objects. The duration of each movement is 5 s, followed by a 3 s rest, with six repetitions of each motion. NinaPro DB4 contains 12 channels. Columns 1–8 are the electrodes equally spaced around the forearm at the height of the radio humeral joint. Columns 9 and 10 contain signals from the main activity spot of the muscle flexor and extensor digitorum superficialis, while columns 11 and 12 contain signals from the main activity spot of the muscle biceps brachii and triceps brachii.

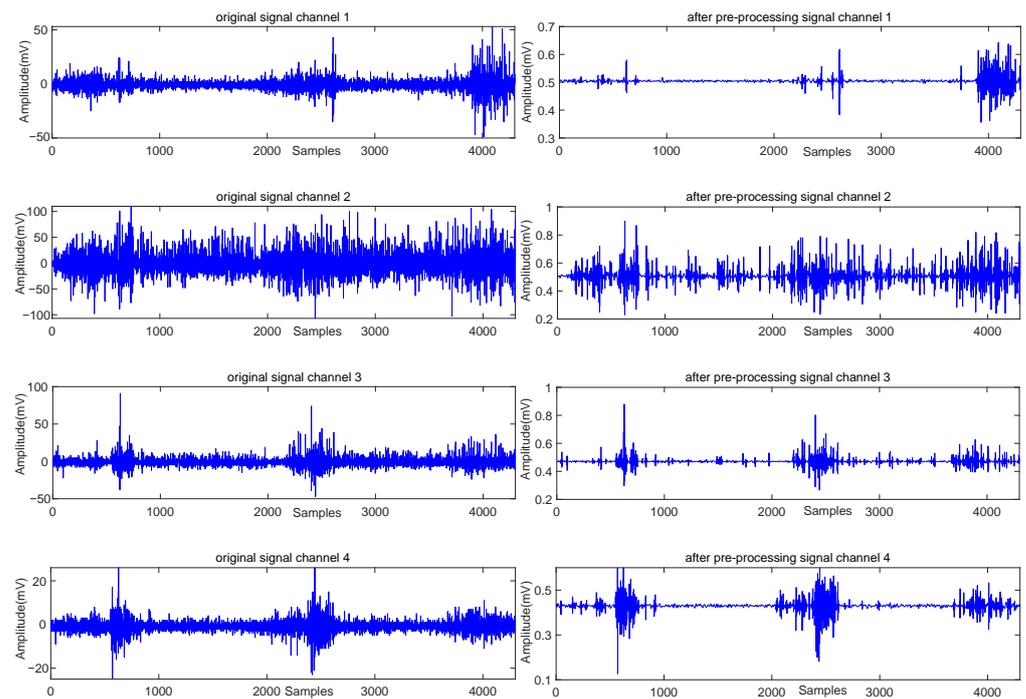
2. NinaPro DB5: In this dataset, sEMG signals are acquired from 10 subjects. The hand gestures are identical to those in NinaPro DB4, with the primary distinctions being the number of channels and sampling frequency. NinaPro DB5 contains 16 channels, columns 1–8 are the electrodes equally spaced around the forearm at the height of the radio humeral joint. Columns 9–16 represent the second Myo, tilted by 22.5 degrees clockwise.
3. BioPatRec DB1: This dataset contains 10 hand and wrist movements collected from 20 subjects. It is including of hand open/close, wrist flex/extend, pro/supination, fine/side grip, pointer (index extension), and agree or thumb up. Each movement has three repetitions and rests for 3 s after a duration of 3 s.
4. BioPatRec DB2: This dataset is collected from 17 subjects including 26 hand activities. It is similar to BioPatRec DB1, except for fine/side grip, pointer (index extension), and agree or thumb up. The duration of each motion is 3 s, followed by 3 s rest, with three repetitions.
5. BioPatRec DB3: This dataset is collected from eight subjects including ten movements with three different devices. The hand gestures are same as BioPatRec DB1. The duration of each hand gesture is 3 s long with 3 s resting in between two motions and three repetitions. BioPatRec DB1, DB2, and DB3 have four, eight, and four electrodes, respectively. The diameter of the electrodes is 1 cm, and the distance between the electrodes of the dipole is 2 cm. The electrodes are equidistantly distributed around the most proximal third of the forearm.
6. Mendeley Data: The dataset records 10 different hand gestures from 40 subjects. The hand gestures performed by the subjects are rest, extension/flexion of the wrist, radial deviation of the wrist, grip, abduction/adduction of all fingers, supination, and pronation. The duration of each gesture is 4 s long with 4 s resting in between two hand gestures and five repetitions. The four electrode positions of Mendelay are Extensor Carpi Ulnaris, Flexor Carpi Ulnaris, Extensor Carpi Radialis, and Flexor Carpi Radialis.

#### 4.2. Dataset Pre-Processing

The sEMG can be disturbed by various types of noise, such as noise from electronic devices (from 0 Hz to thousands of Hz) and noise from motion artifacts [34]. Therefore, filtering operations are needed to remove the noise information in order to preserve the original signal's characteristic information as much as possible. The same filtering method is used for the six public datasets to filter the original EMG signals. Wavelet denoising is used to filter the original EMG signals, and the third-level mother wavelet "db7" is selected for wavelet filtering. The filtered data are then normalized using min–max normalization to ensure that all data are distributed between [0,1]. The filtered and normalized EMG signal segments are shown in Figure 3. The formula for min–max normalization is as follows:

$$\bar{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (10)$$

where  $x_{min}$  represents the minimum value of the signal vector  $x$ , and  $x_{max}$  represents the maximum value of the signal vector  $x$ . Finally, the sEMG signal is split into a collection of signal segments with a window size of 150 ms and a sliding step of 25 ms [35].



**Figure 3.** A visual juxtaposition of selected signal segments from NinaPro DB5 before and after pre-processing is presented. On the left are the raw signal segments without filtering or normalization, while on the right are the corresponding EMG signal segments post-processing, involving wavelet denoising and min–max normalization.

#### 4.3. Pattern Classification and Setting

Following feature extraction from the EMG data, various machine learning algorithms and artificial neural networks (ANNs) are deployed for pattern recognition. The model is trained using the training set's data, and its performance is assessed using the testing dataset. We employ four supervised pattern-recognition algorithms typically used in related hand gesture recognition tasks involving EMG signals [36,37]. These methods encompass linear discriminant analysis (LDA), K-nearest neighbors (KNNs), support vector machine (SVM), and ANN for hand gesture classification.

Linear discriminant analysis (LDA), k-nearest neighbors (KNNs), and support vector machine (SVM) are widely used machine learning algorithms. LDA, a supervised learning method, aims to identify the optimal linear projection that separates different classes in the data. KNN is an instance-based algorithm that classifies new data points by measuring distances to the nearest K neighbors and making decisions based on their labels. SVM, a powerful binary classifier, seeks to find the best hyperplane that maximizes the margin between classes while minimizing classification errors. In those classifications prediction of the test data are evaluated by measuring distance from the training samples. In this work, to determine the class of the testing data Euclidean distance is calculate and the number of neighbors  $K = 10$  has been considered. SVM classification consists of a hyper surface in the space that is used for classification of the data set. In our work, RBF (radial basis function) has been used as the SVM kernel. LDA involves projecting multidimensional data features onto one or lower-dimensional space to maximize the distance between different categories while minimizing the distance within the same category.

In our study, the architecture of our ANN network is structured as follows: it consists of three fully connected layers, where the number of neurons in the hidden layer is dynamically adjusted based on the varying feature quantities. The number of neurons in the output layer is also adapted to the number of categories in different datasets. We have opted for the LeakyReLU activation function [38]. The specific parameter configurations

are as follows: the learning rate is set at 0.01, with a learning rate adjustment strategy of CosineAnnealingLR. We employ the AdamW optimization function, a batch size of 256, and 200 iterations. To address potential distribution differences between the training and test data, a total trail is used for the training and testing. Specifically, Ninapro DB4, the trials numbered 1, 3, 4, and 6 of all 10 subjects were used for training, and the trials numbered 2 and 5 were used for testing. More detail is provided in Table 1.

**Table 1.** Specifications of the sEMG databases used in this paper.

Name	Gestures	Subjects	Channels	Trials	Training	Testing	Sampling Rate
Ninpro DB4	53	10	12	6	1, 3, 4, 6	2, 5	2000 Hz
Ninapro DB5	53	10	16	6	1, 3, 4, 6	2, 5	200 Hz
BioPatRec DB1	10	20	4	3	1, 3	2	2000 Hz
BioPatRec DB2	26	17	8	3	1, 3	2	2000 Hz
BioPatRec DB3	10	8	4	3	1, 3	2	2000 Hz
Mendeley Data	10	40	4	5	1, 2, 4	3, 5	2000 Hz

Regarding the hardware setup, we utilized an Intel Core i5-12400F with a 3.9 GHz CPU, 32 GB of RAM, and an Nvidia RTX 3060 GPU. The project was developed using the PyTorch deep learning library, and all comparative experiments were conducted in this aforementioned environment.

#### 4.4. Performance Evaluation Standards

In this work, the performance of methods is evaluation by classification accuracy and F1-score. Accuracy is the most commonly used classification evaluation index. The formula of it is defined by Equation (11):

$$Accuracy = \frac{\text{Correctedly predicted samples}}{\text{All samples}} \times 100 \quad (11)$$

The F1-score is selected as another evaluation index due to the presence of a large number of similar actions in the data, it and accuracy forms the evaluation index system of the method. The formula of it is defined by Equation (12):

$$F1 \text{ score} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 2 \quad (12)$$

in which precision and recall are determined with the following Equation:

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN} \quad (13)$$

where the operators  $TP$ ,  $FP$ , and  $FN$  are true positives, false positives, and false negatives, respectively. The results for each label are weighted by the number of samples in each class to calculate the F1-score.

## 5. Experiment and Results

### 5.1. Comparison of the Original Data and Data after Extending Function

In this section, we compared the performance of different feature-extraction method combinations and analyzed the classification accuracy when employing the space-to-space (SPS) technique. Our test involved classifiers such as LDA, SVM, KNN, and ANN, conducted on the BioPatRec DB1 dataset, and the results are presented in Table 2 and the accompanying graph.

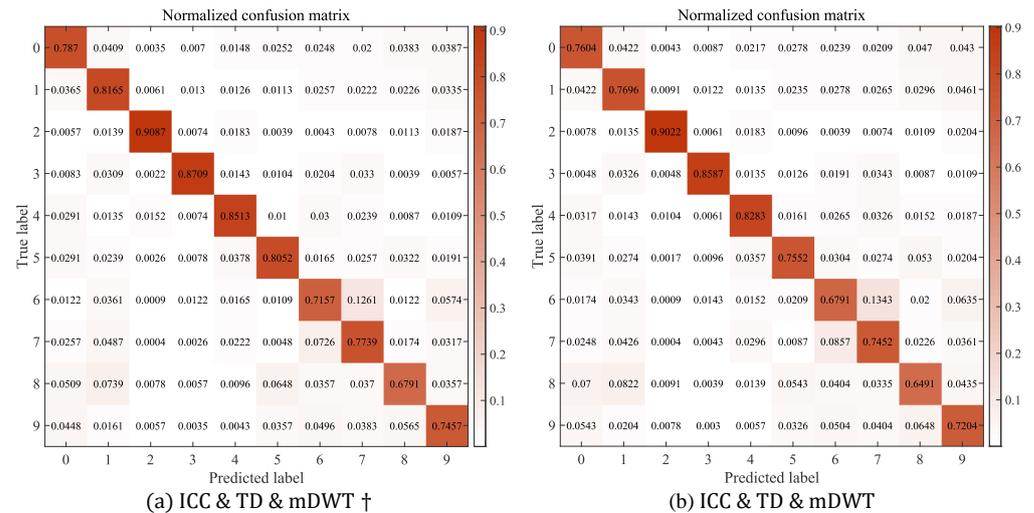
**Table 2.** Ablation experiments of SPS technique, where † represents the operation of expanding sEMG signals, testing on BioPatRec DB1 database. To clearly describe the final data input format for each feature domain, let us consider signal segments  $X(c \times n)$  and  $X'((c \times scale) \times \frac{n}{scale})$  as examples. Here,  $n$  represents the time length of the signal segment,  $c$  denotes the number of channels, and  $scale$  signifies the scaling factor for down-sampling.  $X$  and  $X'$  represent different electromyographic signal segments before and after SPS processing. In the table, † indicates whether a feature participates. After SPS processing, the time length and number of channels of the signal segment are reduced to  $\frac{n}{scale}$  and  $c \times scale$ , respectively. The shape of  $X$ , untreated by SPS, after feature extraction through TD, mDWT, and ICC is  $(c \times 5)$ ,  $(c \times 3)$ , and  $(c \times c)$ , respectively. On the other hand, the shapes of  $X$ , post SPS processing, after feature extraction through TD, mDWT, and ICC, are  $((scale \times c) \times 5)$ ,  $((scale \times c) \times 3)$ , and  $((scale \times c) \times (scale \times c))$ , respectively. Bold indicates the best result under the current method.

Method	LDA	SVM	KNN	ANN
TD	50.88	30.17	64.43	85.09
TD †	<b>52.15</b>	<b>36.59</b>	<b>65.65</b>	<b>85.95</b>
mDWT	47.29	63.90	74.57	79.53
mDWT †	<b>48.55</b>	<b>65.19</b>	<b>77.05</b>	<b>81.52</b>
ICC	39.04	<b>39.39</b>	69.59	76.56
ICC †	<b>41.76</b>	38.94	<b>70.51</b>	<b>78.08</b>
TD and mDWT	53.27	61.85	78.71	85.19
TD and mDWT †	<b>55.60</b>	<b>63.90</b>	<b>81.41</b>	<b>87.25</b>
TD and ICC	52.39	43.60	79.07	85.91
TD and ICC †	<b>55.59</b>	<b>44.86</b>	<b>79.27</b>	<b>86.84</b>
ICC and mDWT	50.48	<b>60.27</b>	74.66	84.01
ICC and mDWT †	<b>53.43</b>	59.28	<b>77.24</b>	<b>85.09</b>
ICC and TD and mDWT	54.60	59.69	78.71	86.36
ICC and TD and mDWT †	<b>57.81</b>	<b>61.00</b>	<b>81.57</b>	<b>88.40</b>

From the table, it is evident that the utilization of SPS for expanding the original sEMG signal segments, in conjunction with various feature-extraction methods, has led to a significant improvement in overall accuracy. The average classification accuracy of the four classifiers increased by 2.39%, 1.5%, 1.7%, and 2.49% after using the SPS technique. This underscores the effectiveness of SPS, as it allows for the extraction of more feature information within individual channels, thereby enhancing the accuracy of gesture classification. Furthermore, It is worth noting that when utilizing a single-feature-extraction method, the use of SPS results in a more pronounced enhancement in classification performance. This phenomenon may be attributed to the inability of a single-feature-extraction method to fully exploit the latent features within sEMG signals. The incorporation of SPS allows for the extraction of additional features, enriching the feature set and ultimately improving recognition accuracy.

Additionally, Figure 4 illustrates the normalized confusion matrix obtained using our proposed approach with an KNN classifier for the BioPatRec DB1 database, with gestures represented by numbers 0 through 9. A majority of the original labels were correctly predicted, and the proposed approach achieved an average classification accuracy of 88.4% across all 10 gestures, with the best class achieving 95.52% accuracy. This further confirms the robust performance of our proposed approach.

Figure 3 illustrates the different classification results with and without the involvement of SPS. It is evident that the inclusion of SPS leads to an enhancement in the classification performance of the classifier. This improvement is particularly notable in categories where the previous recognition accuracy was lower, consequently resulting in an overall elevation of the accuracy in EMG pattern recognition.



**Figure 4.** For the selected KNN classifier, confusion matrices were generated to evaluate the classification performance with and without the involvement of SPS in the methods ICC, TD, and mDWT. The dataset used for this evaluation was BioPatRec DB1. **(a)** The confusion matrix with the inclusion of the SPS method. **(b)** The confusion matrix without the SPS method. The symbol † indicates an SPS operation that extends the sEMG signal.

## 5.2. Comparison of Different Feature Extraction Combinations

To evaluate the performance of the SPS method on different datasets, we selected six publicly available datasets for validation. Additionally, to assess the improvement in EMG pattern recognition achieved through the combination of spatial-domain, frequency-domain, and time-domain features from multiple perspectives, we tested various combinations of feature domains. Furthermore, to mitigate any biases introduced by different classifiers, we conducted tests using four classifiers: LDA, SVM, KNN, and ANN, covering both machine learning and deep learning methods to ensure the generalizability and practicality of our approach.

The results presented in Table 3 demonstrate that employing a multi-perspective fusion approach consistently yields superior outcomes across most classifiers, with the deep learning classifier ANN showing the most significant performance improvement. Across all six datasets, the performance of the artificial neural network (ANN) surpasses those of other machine learning classifiers. The utilization of the ICC, TD, and mDWT-fused feature-extraction method exhibits superior performance compared to other feature-extraction approaches, showcasing average classification accuracy improvements of 2.04%, 3.17%, 4.27%, 2.27%, 4.93%, and 4.26% on Ninapro DB4, DB5, BioPatRec DB1-DB3, and Mendelay Data, respectively. Notably, the smallest enhancement is observed in Ninapro DB4, while the most significant improvement is witnessed in BioPatRec DB3.

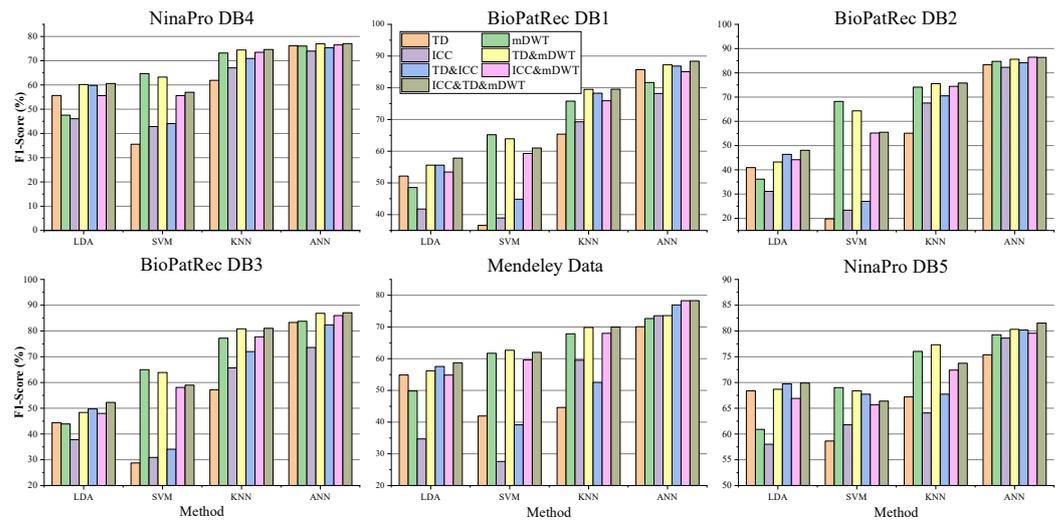
Additionally, Table 3 indicates that the inclusion of multiple feature perspectives leads to performance improvements in most classifiers. In cases where spatial domain features are combined, there is a performance increase of 2–3% compared to methods that lack such features. Moreover, multi-perspective feature classification, when compared to single-feature domain methods, results in enhancements of at least 3% or more.

However, it is important to note that SVM and LDA exhibit distinctive trends. The SVM classifier demonstrates greater sensitivity to frequency domain feature information, achieving higher accuracy in feature recognition when compared to multi-perspective fusion methods. It excels at handling feature information in the frequency domain but struggles with multi-dimensional data. In contrast, LDA exhibits a sensitivity to time-domain feature information, performing well in both single-feature domain and time–frequency-domain combination methods. Nevertheless, its capabilities in managing multi-dimensional feature information are comparatively weaker than other classifiers.

**Table 3.** Classification accuracy (%) on Ninapro DB4, DB5, BioPatRec DB1, DB2, DB3, and Mendeley Data using four classifiers. Furthermore, all methods are processed by SPS. Bold indicates the best results under the same method for the same dataset.

Datasets	Method	LDA	SVM	KNN	ANN
Ninapro DB4	TD	58.96	50.89	63.85	75.78
	mDWT	56.18	<b>66.47</b>	73.03	75.79
	ICC	55.92	54.76	69.74	76.61
	TD and mDWT	57.26	65.30	74.94	75.11
	TD and ICC	62.27	55.55	75.51	76.09
	ICC and mDWT	60.37	62.30	74.73	76.19
	ICC and TD and mDWT	<b>62.91</b>	63.15	<b>75.70</b>	<b>77.64</b>
Ninapro DB5	TD	69.43	66.54	69.88	75.37
	mDWT	67.08	70.97	73.78	78.83
	ICC	66.68	68.86	69.38	79.12
	TD and mDWT	69.55	70.27	74.52	79.73
	TD and ICC	70.55	70.91	70.91	79.19
	ICC and mDWT	69.21	70.89	74.47	79.97
	ICC and TD and mDWT	<b>70.59</b>	<b>71.25</b>	<b>75.41</b>	<b>81.88</b>
BioPatRec DB1	TD	52.15	36.59	65.65	85.95
	mDWT	48.55	<b>65.19</b>	77.05	81.52
	ICC	41.76	38.94	70.51	78.08
	TD and mDWT	55.60	63.90	80.41	87.25
	TD and ICC	55.59	44.86	79.27	86.84
	ICC and mDWT	53.43	59.28	77.24	85.09
	ICC and TD and mDWT	<b>57.81</b>	61.00	<b>81.57</b>	<b>88.40</b>
BioPatRec DB2	TD	40.93	19.78	55.11	83.33
	mDWT	36.17	<b>68.20</b>	76.59	83.74
	ICC	31.12	23.37	70.96	82.31
	TD and mDWT	43.24	64.32	77.29	84.60
	TD and ICC	46.39	27.03	72.15	85.16
	ICC and mDWT	44.18	55.19	77.15	85.45
	ICC and TD and mDWT	<b>48.06</b>	55.52	<b>78.67</b>	<b>86.37</b>
BioPatRec DB3	TD	44.39	28.76	57.18	82.31
	mDWT	43.90	64.97	77.20	82.77
	ICC	37.82	30.92	65.67	73.59
	TD and mDWT	48.36	<b>63.87</b>	79.83	85.84
	TD and ICC	49.74	34.10	71.98	82.36
	ICC and mDWT	47.99	58.08	77.73	85.99
	ICC and TD and mDWT	<b>52.27</b>	59.00	<b>81.04</b>	<b>87.08</b>
Mendeley Data	TD	54.87	41.93	44.62	70.14
	mDWT	49.83	61.72	66.90	72.64
	ICC	34.70	27.61	59.60	73.59
	TD and mDWT	56.19	61.67	69.92	73.50
	TD and ICC	57.55	39.15	52.54	76.77
	ICC and mDWT	54.84	59.59	68.11	77.13
	ICC and TD and mDWT	<b>58.71</b>	<b>62.00</b>	<b>70.06</b>	<b>78.23</b>

Furthermore, we assessed F1-score performance on these datasets. The F1-score is employed to characterize performance in scenarios with sample imbalances for each label, as illustrated in Figure 5. We observed that classifiers with stronger decoding capabilities could extract more information from electromyographic signals as the feature information increased. Even classifiers with relatively weaker decoding abilities exhibited good performance. Importantly, the ANN classifier demonstrated performance improvements after the fusion of spatial domain features, further enhancing the multidimensional perspective.



**Figure 5.** The F1-score of all datasets is tested by five classifiers on seven methods.

### 5.3. Comparison of Result with Recent Feature-Extraction Methods

In this section, we conducted a comparative analysis of our proposed method against existing sEMG-based gesture recognition techniques to underscore the superiority of our approach, as illustrated in Table 4. Recognizing the substantial variations in data collection conditions and environmental factors, we followed the settings outlined in [21] and selected a set of machine learning and traditional deep learning methods for comparison. To ensure fairness and minimize the influence of data quantity variations due to different acquisition devices, we adopted a uniform approach, using sEMG data from eight channels. Our comparison was conducted using the BioPatRec DB2 dataset, which encompasses data from 17 subjects and includes 26 distinct gestures, making it the dataset with the highest gesture diversity among those considered for comparison, thus presenting greater recognition complexity.

**Table 4.** Comparison with recent pattern-recognition research. Bold indicates the optimal result among all methods.

Method	Participants	Gesture	Channels	Accuracy
Machine learning with multi-window majority Voting [18]	40	6	8	80.70%
Deep Learning [38]	10	10	8	77.78%
Machine learning with the sequential floating feature selection method [39]	21	17	8	84.50%
Machine Learning [36]	4	7	8	89.20%
Machine Learning with SPD-manifold based features [21]	10	11	8	84.85%
This work	17	26	8	<b>86.37 %</b>

There are a total of five methods participating in the comparison, as shown in Table 4. The authors of [18] investigated the effects of different window sizes and overlaps on machine learning performance, and developed a novel multi-window majority voting strategy to improve hand gesture recognition accuracies using electromyography signals. The authors of [38] introduced an innovative hand gesture recognition system that utilizes long short-term memory (LSTM) deep learning algorithms to classify hand gestures by training and testing the collected inertial measurement unit (IMU), electromyographic

(EMG), and finger and palm pressure data. The authors of [39] studied the feasibility of using electromyography (EMG) signals for hand gesture recognition, with a focus on comparing the effects of recording EMG signals from the forearm versus the wrist. The results showed that wrist EMG signals have higher signal quality metrics for subtle finger motions, while maintaining comparable quality for wrist motions. The authors of ref. [36] presented an EMG acquisition and gesture recognition system based on an embedded platform, which employs customized analog front-end circuits and digital signal processing algorithms to achieve high-precision and real-time gesture recognition. The authors of ref. [21] proposed a myoelectric pattern-recognition method based on the SPD manifold, which utilizes the SPD manifold to extract spatial structural information from electromyography signals as features for hand gesture recognition.

Table 4 clearly demonstrates that our proposed method is highly competitive when juxtaposed with recent studies. This work achieved an accuracy rate of 86.37%; while this is slightly below the 89.2% reported in ref. [36], it is essential to note that [36] focused on classifying only 7 gestures, whereas our method addresses 26 gestures and still attains remarkably close results. This underscores the remarkable effectiveness of our approach.

Notably, in the context of the machine learning with multi-window majority voting technique, which utilized data from 40 subjects performing six different gestures, our method showcased a performance improvement of 5.67%. In the case of the deep learning method based on LSTM, incorporating data from 10 subjects and encompassing 10 different gestures, our approach exhibited substantial performance enhancements of 8.59%. It is important to highlight that deep learning methodologies have been increasingly adopted in EMG pattern recognition, as they aim to automatically extract more informative features and identify EMG patterns end-to-end, despite the associated increase in computational complexity.

Furthermore, with regard to the machine learning with sequential floating feature selection approach, which involved data from 21 subjects encompassing 17 gestures, our method demonstrated a performance improvement of 1.87%. The SPD-manifold achieved 84.85% with 10 subjects and 11 gestures. However, it is worth noting that the feature extraction process in this method is relatively complex and can be time-consuming, potentially presenting challenges for real-time recognition in practical scenarios.

## 6. Discussion

This paper proposes a surface electromyography (sEMG) feature-extraction method based on spatial expansion to improve gesture recognition accuracy. The innovation lies in introducing the “Space-to-Space (SPS)” framework, which folds the signals in the time dimension to expand the original sEMG data points in the spatial dimension, releasing more feature information within a single channel. On the SPS processed signals, time-domain features (including RMS, MAV, SSC, WL, ZC, etc.), frequency-domain features (using wavelet transform), and spatial features (ICC matrix constructed from inter-channel covariance matrix) are extracted to achieve multi-dimensional feature fusion. The experimental results demonstrate that compared to using only time-domain and frequency-domain features, this method achieves significant recognition performance improvements of 2.04% to 4.93% in average accuracy on the Ninapro DB4, DB5, BioPatRec DB1–DB3, and Mendeley Data datasets. This clearly shows that spatial features play a key role in sEMG decoding. In addition, comparative analysis with other state-of-the-art methods also proves the superiority of this method in terms of higher gesture recognition accuracy.

While adopting the multi-view feature fusion strategy enriches the feature expression of sEMG signals and improves gesture recognition accuracy, different classifiers demonstrate varying sensitivities towards features. For instance, the SVM classifier is more sensitive to frequency-domain features and can achieve higher accuracy using only frequency-domain features, but struggles in handling multi-dimensional features. In contrast, the LDA classifier is more sensitive to time-domain features and performs well in both individual time-domain features and combinations of time–frequency-domain fea-

tures, yet has weaker capabilities in dealing with multi-dimensional features. Overall, this article has made useful explorations in the expression and fusion of multi-view sEMG features, which provides a reference for the selection of classifiers and feature-extraction methods in practical applications, and provides new ideas for surface electromyographic signal analysis.

## 7. Conclusions

To better decode additional feature information from electromyographic signals, thereby achieving the recognition task of movement postures, this paper introduces a method for EMG pattern recognition using space-to-space (SPS), which integrates multiple feature perspectives from the spatial domain (ICC), time domain, and frequency domain. By combining spatially-extended SPS with subsequent feature extraction techniques, it extensively explores spatial, temporal, and frequency-domain features within and across different channels of electromyographic signals. To evaluate the performance of the proposed architecture, experiments were conducted on six sEMG benchmark databases: Ninapro DB4, Ninapro DB5, BioPatRec DB1, BioPatRec DB2, BioPatRec DB3, and Mendeley Data. The results demonstrate that the feature-extraction method based on multi-perspective fusion can better extract kinematic information from sEMG signals and decode them, leading to improved recognition accuracy compared to traditional feature-extraction methods. The experimental findings underscore the superiority of our approach and offer a novel method for electromyographic pattern recognition.

In the future, we plan to explore alternative feature-extraction methods in Euclidean space, projecting them into Riemannian space to seek more suitable representations of electromyographic signals. Additionally, we aim to enhance the algorithm's robustness and stability by combining these methods with existing feature extraction techniques. Furthermore, we will investigate the fusion of modal information from other domains with electromyographic signals and establish a unified decoding framework, providing a new option for human-computer interaction.

**Author Contributions:** Conceptualization, W.C. and Y.N.; Formal analysis, B.X.; Methodology, W.C. and Y.N.; Validation, Z.G. and S.H.; Writing—original draft, W.C. and Y.N.; Writing—review and editing, B.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the National Natural Science Foundation of China under Grant 62073271, the Natural Science Foundation of Fujian Province (2020J01890, 2020J01891), and the Scientific Fund Projects in Fujian University of Technology (GY-Z17144), and in part by the Scientific Research Projects of the Science and Technology Department in Fujian of China (2020J01890, 2023I0024), the Provincial Project of Education Department in Fujian of China (JT180344, FBJY20230078, FJJKBK23-045), and the Scientific Research Projects in Fujian University of Technology (KF-X19002, KF-19-22001, GY-H-21191, GY-Z21049, GY-Z220210).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are openly available in public domain at <https://ninapro.hevs.ch>, <https://github.com/biopatrec/biopatrec> and <https://data.mendeley.com/datasets/ckwc76xr2z/2> (accessed on 4 November 2023).

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

## References

1. Merletti, R.; Parker, P.J. *Electromyography: Physiology, Engineering, and Non-Invasive Applications*; John Wiley and Sons: Hoboken, NJ, USA, 2004; Volume 11.
2. Merletti, R.; Farina, D. Detection and Conditioning of Surface EMG Signals. In *Surface Electromyography: Physiology, Engineering, and Applications*; IEEE: Piscataway, NJ, USA, 2016; pp. 1–37.
3. Siddiqui, N.; Chan, R.H.M. Multimodal hand gesture recognition using single IMU and acoustic measurements at wrist. *PLoS ONE* **2020**, *15*, e0227039. [[CrossRef](#)] [[PubMed](#)]

4. Ahmad, N.; Ghazilla, R.A.R.; Khairi, N.M.; Kasi, V. Reviews on Various Inertial Measurement Unit (IMU) Sensor Applications. *Int. J. Signal Process. Syst.* **2013**, *1*, 256–262. [[CrossRef](#)]
5. Došen, S.; Cipriani, C.; Kostić, M.; Controzzi, M.; Carrozza, M.C.; Popović, D.B. Cognitive vision system for control of dexterous prosthetic hands: Experimental evaluation. *J. Neuroeng. Rehabil.* **2010**, *7*, 42. [[CrossRef](#)] [[PubMed](#)]
6. Haroon, M.; Altaf, S.; Rehman, Z.u.; Soomro, M.W.; Iqbal, S. Human hand gesture identification framework using SIFT and knowledge-level technique. *ETRI J.* **2022**, *early view*. [[CrossRef](#)]
7. Guo, W.; Fang, Y.; Sheng, X.; Zhu, X. Measuring Motor Unit Discharge, Myofiber Vibration, and Haemodynamics for Enhanced Myoelectric Gesture Recognition. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–10. [[CrossRef](#)]
8. Guo, W.; Sheng, X.; Liu, H.; Zhu, X. Toward an Enhanced Human–Machine Interface for Upper-Limb Prosthesis Control with Combined EMG and NIRS Signals. *IEEE Trans. Hum.–Mach. Syst.* **2017**, *47*, 564–575. [[CrossRef](#)]
9. Ha, N.; Withanachchi, G.P.; Yihun, Y. Performance of Forearm FMG for Estimating Hand Gestures and Prosthetic Hand Control. *J. Bionic Eng.* **2019**, *16*, 88–98. [[CrossRef](#)]
10. Asfour, M.; Menon, C.; Jiang, X. A machine learning processing pipeline for reliable hand gesture classification of FMG signals with stochastic variance. *Sensors* **2021**, *21*, 1504. [[CrossRef](#)]
11. Atzori, M.; Cognolato, M.; Müller, H. Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands. *Front. Neurobot.* **2016**, *10*, 9. [[CrossRef](#)]
12. Geng, W.; Du, Y.; Jin, W.; Wei, W.; Hu, Y.; Li, J. Gesture recognition by instantaneous surface EMG images. *Sci. Rep.* **2016**, *6*, 36571. [[CrossRef](#)]
13. Peng, X.; Zhou, X.; Zhu, H.; Ke, Z.; Pan, C. MSFF-Net: Multi-Stream Feature Fusion Network for surface electromyography gesture recognition. *PLoS ONE* **2022**, *17*, e0276436. [[CrossRef](#)] [[PubMed](#)]
14. Crouch, D.L.; Huang, H. Lumped-parameter electromyogram-driven musculoskeletal hand model: A potential platform for real-time prosthesis control. *J. Biomech.* **2016**, *49*, 3901–3907. [[CrossRef](#)] [[PubMed](#)]
15. Tkach, D.; Huang, H.; Kuiken, T.A. Study of stability of time-domain features for electromyographic pattern recognition. *J. NeuroEng. Rehabil.* **2010**, *7*, 21. [[CrossRef](#)] [[PubMed](#)]
16. Merletti, R.; Lo Conte, L.R. Advances in processing of surface myoelectric signals: Part 1. *Med Biol. Eng. Comput.* **1995**, *33*, 362–372. [[CrossRef](#)] [[PubMed](#)]
17. Cifrek, M.; Medved, V.; Tonković, S.; Ostojić, S. Surface EMG based muscle fatigue evaluation in biomechanics. *Clin. Biomech.* **2009**, *24*, 327–340. [[CrossRef](#)] [[PubMed](#)]
18. Wahid, M.F.; Tafreshi, R.; Langari, R. A multi-window majority voting strategy to improve hand gesture recognition accuracies using electromyography signal. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *28*, 427–436. [[CrossRef](#)] [[PubMed](#)]
19. Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Feature reduction and selection for EMG signal classification. *Expert Syst. Appl.* **2012**, *39*, 7420–7431. [[CrossRef](#)]
20. Oladazimi, M.; Molaei-Vaneghi, F.; Safari, M.; Asadi, H.; Aghay Kaboli, S. A review for feature extraction of EMG signal processing. In Proceedings of the 4th International Conference on Computer and Automation Engineering (ICCAE 2012), Mumbai, India, 14–15 January 2012; pp. 85–94.
21. Xiong, D.; Zhang, D.; Zhao, X.; Chu, Y.; Zhao, Y. Learning Non-Euclidean Representations with SPD Manifold for Myoelectric Pattern Recognition. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 1514–1524. [[CrossRef](#)]
22. Sunkara, R.; Luo, T. No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Grenoble, France, 19–23 September 2022; pp. 443–459.
23. Smith, S.T. Covariance, subspace, and intrinsic crame/spl acute/r-rao bounds. *IEEE Trans. Signal Process.* **2005**, *53*, 1610–1630. [[CrossRef](#)]
24. Gopan K, G.; Prabhu, S.S.; Sinha, N. Sleep EEG analysis utilizing inter-channel covariance matrices. *Biocybern. Biomed. Eng.* **2020**, *40*, 527–545. [[CrossRef](#)]
25. Barachant, A.; Bonnet, S.; Congedo, M.; Jutten, C. Multiclass brain–computer interface classification by Riemannian geometry. *IEEE Trans. Biomed. Eng.* **2011**, *59*, 920–928. [[CrossRef](#)] [[PubMed](#)]
26. Wei, W.; Dai, Q.; Wong, Y.; Hu, Y.; Kankanhalli, M.; Geng, W. Surface-electromyography-based gesture recognition by multi-view deep learning. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 2964–2973. [[CrossRef](#)] [[PubMed](#)]
27. Kuzborskij, I.; Gijsberts, A.; Caputo, B. On the challenge of classifying 52 hand movements from surface electromyography. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 10 November 2012; pp. 4931–4937.
28. Li, W.; Shi, P.; Yu, H. Gesture recognition using surface electromyography and deep learning for prostheses hand: State-of-the-art, challenges, and future. *Front. Neurosci.* **2021**, *15*, 621885. [[CrossRef](#)] [[PubMed](#)]
29. Lucas, M.F.; Gaufriau, A.; Pascual, S.; Doncarli, C.; Farina, D. Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. *Biomed. Signal Process. Control* **2008**, *3*, 169–174. [[CrossRef](#)]
30. Pizzolato, S.; Tagliapietra, L.; Cognolato, M.; Reggiani, M.; Müller, H.; Atzori, M. Comparison of six electromyography acquisition setups on hand movement classification tasks. *PLoS ONE* **2017**, *12*, e0186132. [[CrossRef](#)]
31. Ortiz-Catalan, M.; Brånemark, R.; Håkansson, B. BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms. *Source Code Biol. Med.* **2013**, *8*, 11. [[CrossRef](#)]

32. Ozdemir, M.A.; Kisa, D.H.; Guren, O.; Akan, A. Dataset for multi-channel surface electromyography (sEMG) signals of hand gestures. *Data Brief* **2022**, *41*, 107921. [[CrossRef](#)]
33. Reifinger, S.; Wallhoff, F.; Ablassmeier, M.; Poitschke, T.; Rigoll, G. Static and dynamic hand-gesture recognition for augmented reality applications. In Proceedings of the Human–Computer Interaction, HCI Intelligent Multimodal Interaction Environments: 12th International Conference, HCI International 2007, Beijing, China, 22–27 July 2007; pp. 728–737.
34. Nazmi, N.; Abdul Rahman, M.A.; Yamamoto, S.I.; Ahmad, S.A.; Zamzuri, H.; Mazlan, S.A. A review of classification techniques of EMG signals during isotonic and isometric contractions. *Sensors* **2016**, *16*, 1304. [[CrossRef](#)]
35. Smith, L.H.; Hargrove, L.J.; Lock, B.A.; Kuiken, T.A. Determining the optimal window length for pattern recognition-based myoelectric control: Balancing the competing effects of classification error and controller delay. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2010**, *19*, 186–192. [[CrossRef](#)]
36. Benatti, S.; Casamassima, F.; Milosevic, B.; Farella, E.; Schönle, P.; Fateh, S.; Burger, T.; Huang, Q.; Benini, L. A versatile embedded platform for EMG acquisition and gesture recognition. *IEEE Trans. Biomed. Circuits Syst.* **2015**, *9*, 620–630. [[CrossRef](#)]
37. Gijsberts, A.; Atzori, M.; Castellini, C.; Müller, H.; Caputo, B. Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-Based Hand Movement Classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 735–744. [[CrossRef](#)] [[PubMed](#)]
38. Zhang, X.; Yang, Z.; Chen, T.; Chen, D.; Huang, M.C. Cooperative sensing and wearable computing for sequential hand gesture recognition. *IEEE Sens. J.* **2019**, *19*, 5775–5783. [[CrossRef](#)]
39. Botros, F.S.; Phinyomark, A.; Scheme, E.J. Electromyography-based gesture recognition: Is it time to change focus from the forearm to the wrist? *IEEE Trans. Ind. Inform.* **2020**, *18*, 174–184. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.