

Article

Optimizing sEMG Gesture Recognition: Leveraging Channel Selection and Feature Compression for Improved Accuracy and Computational Efficiency

Yinxi Niu ¹, Wensheng Chen ¹, Hui Zeng ¹, Zhenhua Gan ² and Baoping Xiong ^{1,*}

¹ School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350116, China; 2211301016@smail.fjut.edu.cn (Y.N.); ichenws@163.com (W.C.); x_data1@163.com (H.Z.)

² Fujian Key Laboratory of Automotive Electronics and Electric Drive, Fujian University of Technology, Fuzhou 350116, China; ganzh@fjut.edu.cn

* Correspondence: xiongbp@fjut.edu.cn

Abstract: In the task of upper-limb pattern recognition, effective feature extraction, channel selection, and classification methods are crucial for the construction of an efficient surface electromyography (sEMG) signal classification framework. However, existing deep learning models often face limitations due to improper channel selection methods and overly specific designs, leading to high computational complexity and limited scalability. To address this challenge, this study introduces a deep learning network based on channel feature compression—partial channel selection sEMG net (PCS-EMGNet). This network combines channel feature compression (channel selection) and feature extraction (partial block), aiming to reduce the model's parameter count while maintaining recognition accuracy. PCS-EMGNet extracts high-dimensional feature vectors from sEMG signals through the partial block, decoding spatial and temporal feature information. Subsequently, channel selection compresses and filters these high-dimensional feature vectors, accurately selecting channel features to reduce the model's parameter count, thereby decreasing computational complexity and enhancing the model's processing speed. Moreover, the proposed method ensures the stability of classification, further improving the model's capability of recognizing features in sEMG signal data. Experimental validation was conducted on five benchmark databases, namely the NinaPro DB4, NinaPro DB5, BioPatRec DB1, BioPatRec DB2, and BioPatRec DB3 datasets. Compared to traditional gesture recognition methods, PCS-EMGNet significantly enhanced recognition accuracy and computational efficiency, broadening its application prospects in real-world settings. The experimental results showed that our model achieved the highest average accuracy of 88.34% across these databases, marking a 9.96% increase in average accuracy compared to models with similar parameter counts. Simultaneously, our model's parameter size was reduced by an average of 80% compared to previous gesture recognition models, demonstrating the effectiveness of channel feature compression in maintaining recognition accuracy while significantly reducing the parameter count.



Citation: Niu, Y.; Chen, W.; Zeng, H.; Gan, Z.; Xiong, B. Optimizing sEMG Gesture Recognition: Leveraging Channel Selection and Feature Compression for Improved Accuracy and Computational Efficiency. *Appl. Sci.* **2024**, *14*, 3389. <https://doi.org/10.3390/app14083389>

Academic Editor: Andrea Prati

Received: 2 March 2024

Revised: 13 April 2024

Accepted: 15 April 2024

Published: 17 April 2024

Keywords: sEMG; hand gestures classification; channel feature compression; channel feature selection



Copyright: © 2024 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

Surface Electromyography (sEMG) allows for non-invasive detection of electrical activity generated by muscle fibers on the surface of the skin. These signals reflect muscle activity and provide information about limb movement [1]. Gesture recognition is one of the most crucial perceptual channels in human–computer interaction. It finds extensive applications in virtual reality, intelligent sign language translation for the deaf and mute [2], rehabilitation therapy and assessment [3,4], and bionic prosthetics [5], among other scenarios, showing vast potential across various applications. In the field of gesture recognition, surface electromyography signals serve as a common signal source, capturing muscle activity from electrical signals on the skin's surface [6]. sEMG signals offer numerous

advantages, including non-invasiveness, ease of acquisition, and suitability for dynamic gesture recognition. Consequently, sEMG has garnered significant attention from scholars and is widely applied in contexts where high-precision gesture recognition is demanded. Gesture recognition based on surface electromyography features high accuracy, ease of wear, and non-invasiveness, making it a focal point of research in the field.

Traditional gesture recognition frameworks based on sEMG consist of data preprocessing, feature extraction, feature selection, and gesture classification. Among these stages, feature extraction and gesture classification are two critical phases within the framework of sEMG-based gesture recognition. Common feature extraction methods encompass time-domain features [7], frequency-domain features [8], and time–frequency-domain features [9]. For instance, time-domain features often include metrics such as Mean Absolute Value (MAV), Root Mean Square (RMS), Mean Absolute Value Slope (MAV Slope), Waveform Length (WL), Slope Sign Changes (SSCs), Zero Crossing (ZC), and EMG Histogram (HIST), with EMG histogram being an extension of zero crossing. Following the feature extraction methods, traditional classification techniques are employed for gesture classification. Different classifiers have been introduced, such as k-Nearest Neighbors (KNN) [10], random forest [11], Linear Discriminant Analysis (LDA) [12], Support Vector Machine (SVM) [13], and Hidden Markov Models (HMMs) [14]. Khomami and Shamekhi [15] employed a KNN classifier using 25 features of sEMG and accelerometer signals to classify Persian sign language symbols, achieving an average accuracy of 96.13%. Altimemy et al. [16] classified 15 hand movements for individuals with intact limbs and 12 hand movements for amputees using linear discriminant analysis (LDA) and SVM with AutoRegressive (AR) features. Given that EMG signals represent sequential data, hidden Markov models are suitable for modeling the latent information in EMG signals. Yun et al. [17] used an HMM classifier to create an sEMG-based sign language recognition system. Despite the significant potential of sEMG technology, traditional gesture recognition methods still face several challenges. sEMG signals are usually of high dimensionality and complexity, involving multi-channel and high-sampling-rate data, which poses challenges for feature extraction and dimensionality management. Moreover, sEMG signals are susceptible to interference from muscle fatigue, electrode displacement, and environmental noise, which affects the stability and accuracy of the model. Finally, existing gesture recognition methods usually rely on complex feature engineering, which may limit the performance of the model.

In recent years, deep learning has gained substantial popularity and made groundbreaking advances in various domains, such as image processing [18] and speech recognition [19]. More recently, the use of deep learning for sEMG-based gesture recognition has started to capture researchers' attention. They have begun exploring the application of deep neural networks (DNNs) for gesture recognition. DNNs offer powerful feature learning and classification capabilities, eliminating the need for manual feature engineering and enabling the automatic learning of gesture feature representations, consequently significantly enhancing gesture recognition accuracy. Among various deep learning techniques used for sEMG-based gesture recognition, the Convolutional Neural Network (CNN) architecture stands out as one of the most widely employed. Researchers have categorized these into two primary types based on different evaluation methods. For example, Atzori et al. [5] conducted sEMG classification tasks on four publicly available datasets using a deep convolutional neural network (CNN) architecture comprising two convolutional layers. Their work demonstrated a performance improvement of 2–5% compared to existing machine learning classifiers such as KNN, SVM, random forests, and LDA [20]. Jia et al. [21] proposed a deep learning model that combines Convolutional Autoencoders (CAEs) and CNNs for classification of sEMG datasets consisting of ten hand gestures. The results indicated high levels of performance and robustness. Zhai et al. [22] fed the spectral representations of sEMG into a CNN for gesture recognition, but they achieved only 78.7% accuracy on the second sub-database of the NinaPro dataset. Furthermore, due to the temporal characteristics of sEMG, Recurrent Neural Networks (RNNs) and Long

Short-Term Memory (LSTM) architectures have been applied to hand-related problems based on surface electromyography. In [23], Koch et al. employed a ConvLSTM cascaded with an LSTM architecture for gesture sequence classification. In [24], a two-level network composed of a fully connected network and stacked RNNs was implemented to classify high-density (HD) and sparse sEMG signals. Amor et al. [25] collected sEMG signals using the Myo armband for sign language recognition and applied an RNN architecture to extract features from sequential data for analysis of sign language gestures.

Despite the promising outcomes in sEMG signal analysis, the broader application of existing methods is impeded by their complexity, computational demands, and extended training times. Current models predominantly focus on augmenting feature information from temporal and frequency domains [26], neglecting inter-channel correlations [27]. This oversight leads to increased model parameters and complexity due to excessive feature extraction, necessitating higher levels of computational resources.

Addressing this challenge requires methodologies capable of both extracting valuable inter-channel feature information and efficiently filtering and compressing redundant features. This paper introduces a novel deep learning-based approach for gesture recognition. The methodology presented in this paper was inspired by the work of Jia et al. [28], particularly the multimodal squeeze-and-excite feature fusion module in the SEN-DAL model. Given the unequal contributions of electroencephalogram (EEG) and electro-oculogram (EOG) data in different sleep stages, it is necessary to assign varying levels of importance to them. Hence, by modeling the interdependencies among channels through squeeze-and-excite (SE), SEN-DAL recalibrates the response contributions of channel features. Building upon the aforementioned process, this paper considers how to design modules to extract inter-channel feature information, thereby establishing the correlation between channel features and actions. Importance is assigned to different channel features based on their contributions during different action processes, thereby filtering and compressing the model's features. The model incorporates a channel feature selection unit to comprehensively address the interplay between distinct feature channels. It excels at amplifying the importance of valuable features, suppressing redundant attributes, and employs partial convolution to efficiently capture essential feature channel information while economizing on parameter count. This not only enhances algorithm robustness but also fosters generalization. Our innovative approach not only offers a fresh vantage point in sEMG signal processing and analysis but also ushers in new horizons for the application of deep learning in the realm of biomedical signal processing. Experimental results unequivocally underscore the effectiveness of our method and provide profound insights for the advancement of sEMG gesture recognition.

The structure of this paper is outlined as follows. First, we briefly introduce the importance of gesture recognition and its application areas to outline the research theme. Next, we provide an overview of existing gesture recognition methods and point out the existing issues. Then, we present the details of the method proposed in this paper. Subsequently, we describe the experimental results, including accuracy, recall, and other metrics, followed by a comparative analysis of the experimental results. Finally, we summarize the research achievements of this paper and provide insights into future research directions. The main contributions of this paper are summarized as follows:

- We introduce a channel selection mechanism empowered by gating unit mechanisms to selectively filter channel feature vectors. This innovation compresses and selects key features within high-dimensional feature vectors of the model, reducing its complexity and laying a foundation for further enhancements.
- We implement a combination of partial convolution and channel selection units. Through the collaborative operation of these modules, our approach not only extracts inter-channel features but also efficiently filters intrinsic redundant features in feature maps while maintaining recognition accuracy. Consequently, it significantly reduces the model's parameter count, marking progress in model efficiency and resource optimization.

- Our model is evaluated across various publicly available databases to demonstrate its effectiveness and robustness. Through comprehensive analysis and comparison with state-of-the-art deep learning methods, we show that our model achieves higher accuracy with fewer trainable parameters across different databases while maintaining exceptional recognition stability. These results provide strong evidence of the efficacy of our proposed model.

2. Methods

2.1. Patch Embedding

We utilize patch embedding to segment sEMG signals and employ one-dimensional convolutional layers to extract local features based on the temporal aspects of the sEMG signals, as depicted in Figure 1. Patch embedding plays a crucial role in transforming these segments into a representation suitable for neural networks. Furthermore, patch embedding subdivides sEMG signal segments into smaller patches, thereby reducing the parameter count at the input layer and making the network more lightweight. This enhancement enables the network to capture local contextual sEMG information more effectively. Additionally, patch embedding adds scalability to the network, enabling it to process sEMG segments of varying sizes without requiring modifications to the network structure. This significantly enhances its versatility.

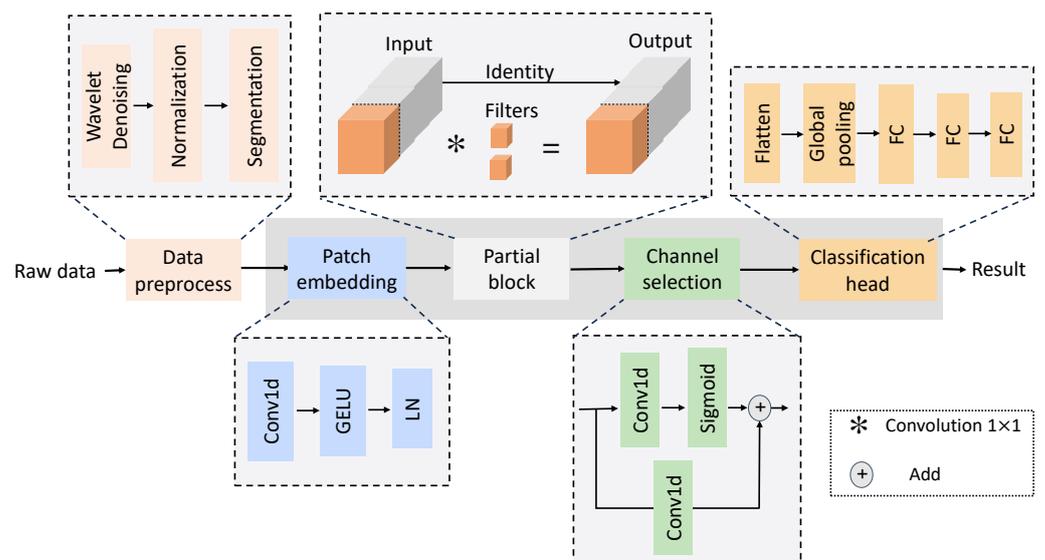


Figure 1. The general procedure of the proposed method. The gray area is the structure of PCS-EMGNet.

The employed one-dimensional convolutional layer comprises 128 filters with dimensions of 10×1 and a stride of 6. Convolutional neural networks (CNNs) are neural networks equipped with a “receptive field” to extract local features. In one-dimensional CNNs, the convolutional kernel convolves along the temporal dimension to extract local features from the time-based sEMG signal. Gaussian Error Linear Units (GELUs) are then employed as the activation function for the nonlinear transformation of the extracted features. Additionally, due to the strong correlation between adjacent time steps in the sEMG signal, Layer Normalization (LN) is applied to normalize the output.

2.2. Partial Block

In the realm of signal processing, precisely recognizing and utilizing essential features is a paramount challenge, especially in applications like sEMG-based gesture recognition. Our innovative approach tackles this challenge by introducing feature importance learning, a pivotal component that dynamically identifies and prioritizes the most critical features

within the sEMG signal. The feature importance learning process revolves around our feature importance network, a tailored neural network meticulously designed for this task. This network boasts a 1D convolutional layer that diligently evaluates the significance of each feature. By harnessing this layer, it quantifies the importance of features and generates importance scores. These scores, constituting a multidimensional importance distribution, undergo normalization via the softmax function. This ensures that they accurately reflect the relative importance of each feature in the input channels.

$$importance_scores = Softmax(Conv(x_{input})) \quad (1)$$

where x_{input} represents the input signal data within the partial block in Figure 1, which has is processed through patch embedding. The term $importance_scores$ denotes the importance scores of feature vectors, referred to as filters in Figure 1. Complementing our feature importance extraction is the partial conv module, a lightweight local convolutional layer that optimizes the signal processing pipeline. The partial conv module plays an indispensable role in feature processing and the reduction of computational overhead. The architecture of the partial conv module consists of a standard convolutional layer augmented by a GELU activation function and a partial convolution layer.

$$x = x_{input} \odot importance_scores \quad (2)$$

where x represents the processed output signal in Figure 1. The symbol \odot denotes the dot product operation. This combination enables the efficient processing of input channels while preserving the vital components of the signal. This is achieved by selectively applying convolution to a fraction of the input channels, a crucial element for efficient feature extraction.

Our unique methodology seamlessly intertwines feature importance learning and the partial conv module, leveraging the strengths of both components for optimized signal processing. The process begins with the feature importance network, which calculates importance scores for individual features. These scores serve as guidance for the selective feature weighting, emphasizing those features deemed most crucial. Following the feature importance-based weighting, the data are handed over to the partial conv module. This component, working in tandem with the importance scores, performs the processing. It partitions the input data into two segments, with one segment being selectively weighted based on the feature importance scores. By focusing computation on the most critical features, we substantially reduce redundancy and computational overhead.

Our approach marries feature importance learning with the lightweight partial conv module, presenting a powerful solution for efficient sEMG signal processing. This approach not only adapts to the dynamic importance of features in various gestures but also significantly streamlines computational complexity. It stands as an ideal choice for real-time applications by reducing computational load and enhancing the effectiveness of gesture recognition systems while preserving high accuracy.

2.3. Channel Selection Block

The channel selection block is at the core of this mechanism, playing a pivotal role in the process. During the forward propagation phase, it dynamically evaluates the significance of each input channel by computing gating signals denoted as g . These gating signals are generated based on the input data features, as seen in the following equation:

$$g = \sigma(Wg \times x + bg) \quad (3)$$

where g represents the gating signal, σ is the sigmoid activation function, and Wg and bg stand for the weight and bias of the gate convolution, respectively.

The gating signals are instrumental in determining which channel features should progress to the subsequent processing layer, selectively allowing certain feature channels to

pass through while deliberately suppressing others. This selective activation is described by the following equation:

$$x' = x \times g \quad (4)$$

Equation (4) demonstrates that the input features (x) are selectively modified by the gating signals (g) to produce x' , which represents the channel features after activation or suppression.

The convolution module complements this process by handling feature processing. It employs a convolutional layer specialized in processing input data features. Notably, it differs from standard convolutional layers in that the output of the convolution module is selectively influenced by the gating signals. This means that only those channel features deemed essential by the model are permitted to pass through the convolution module, while others are intentionally suppressed. This selective processing can be expressed as follows:

$$x'' = Conv(x') \quad (5)$$

where x'' represents the channel features after selective convolution processing.

Moreover, the unique strength of the channel selection mechanism lies in its adaptability. By autonomously learning and assigning importance to individual input channels, the model becomes adept at dynamically enabling or inhibiting specific channels as needed. This adaptability proves invaluable in effectively managing a wide range of tasks, as the relevance of different channels can fluctuate depending on the specific requirements of each task. In models that incorporate the channel selection mechanism, input data undergo processing, resulting in an output tensor. The output tensor exclusively comprises features from channels the model deems crucial, with features from less important channels consciously suppressed. This reduction in computational redundancy significantly simplifies the model, making it a robust and efficient solution for various gesture recognition tasks. The channel selection mechanism's adaptability and efficiency make it a valuable addition to the field of sEMG signal processing, where streamlined models capable of accommodating diverse gestures and tasks are in high demand.

2.4. Classification Head

In the final classification stage, we begin by flattening the preceding input features and pass them through a global average pooling layer, followed by three fully connected (FC) layers. The first FC layer consists of 128 neurons, corresponding to the features extracted from the two previously mentioned blocks. Subsequently, the feature vector is fed into a classification network comprising two fully connected layers, each consisting of 512 neurons. In both of these blocks, nonlinear GELU activations and layer normalization are applied after each FC layer. Within the classifier, we also apply a 20% dropout after the first and second fully connected layers to mitigate overfitting. This randomized dropout of neurons reduces the chances of common adaptation among parameters, subsequently decreasing interdependencies among neurons, thus mitigating the risk of overfitting.

3. Experimental Setup

3.1. Dataset Description

In this section, we briefly discuss the five publicly available benchmark datasets used in our experiments. These datasets are the Ninapro DB4 [29], Ninapro DB5 [29], BioPatRec DB1 [30], BioPatRec DB2 [30], and BioPatRec DB3 [30] datasets. sEMG sensors were placed at various muscle locations in the upper limbs during their respective measurement processes. The datasets encompass hand activities, broadly categorized into gestures, wrist movements, object grasping, and hand motions. Detailed descriptions of the datasets are provided in the references listed below.

The Ninapro DB5 dataset comprises muscle activity signals collected through two Thalmic Myo armbands, which contain 16 active single-differential wireless electrodes. The sampling frequency was 200 Hz. The dataset includes recordings of 10 participants

performing 53 gesture movements, as well as a rest condition, with each movement repeated six times. The 53 gestures are categorized into the following three groups: A, B, and C. Group A primarily documents fundamental finger movements, while Group B encompasses finger flexion and extension, along with wrist gestures. Group C focuses on grasping actions involving everyday objects. Furthermore, analysis was conducted using the Ninapro DB4 dataset, which utilized the same experimental design but was collected using 12 sensors with a sampling frequency of 2000 Hz. For training, we used trials numbered 1, 3, 4, and 6 of all 10 subjects, and for testing, we used trials numbered 2 and 5.

The BioPatRec DB1, DB2, and DB3 databases are subsets of the BioPatRec toolbox. BioPatRec DB1 and BioPatRec DB3 capture 10 hand movements using 4 EMG electrodes, while BioPatRec DB2 records 26 hand movements employing 8 EMG electrodes. These datasets comprise EMG signals collected from 20, 17, and 8 subjects, respectively. Each gesture repetition was performed thrice, with a 3 s relaxation period between repetitions. The sampling rate was set to 2000 Hz. The duration of muscle contraction is determined by a contraction time percentage, with a default value of 0.7. Electrodes were spaced at an inter-electrode distance of 2 cm and evenly distributed around the most proximal third of the forearm. For the training phase, data from trials 1 and 3 of all subjects were utilized, while for testing, data from trials 2 were employed. More details are described in Table 1.

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

Database	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

3.2. Dataset Preprocessing

The data preprocessing stage is mainly divided into the following three stages: wavelet denoising, normalization, and sliding-window segmentation. The sEMG signal 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 artefacts. 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 five 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 formula for min-max normalization is as follows:

$$\bar{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (6)$$

where x_{min} represents the minimum value of the signal vector (x), and x_{max} represents the maximum value of the signal vector (x).

To achieve real-time classification of muscle activity, patterns need to be captured rapidly within as short time windows as possible [31]. This allows for a quick response without physiological awareness of latency. In order to achieve this, we decomposed sEMG signals into small segments using the sliding-window strategy with an overlapped

windowing scheme [32]. The sliding window is used to simulate the real-time processing and classification of continuous-stream EMG signals to realize real-time gesture recognition.

The length of the window represents a compromise between time latency and classification accuracy. As described in [31], to satisfy the requirements of real-time control, the time latency should be less than 300 ms. Longer extended-window length leads to more controller delays, although it also increases classification accuracy [12,33,34]. To test the performance of the proposed algorithm in different scenes, in this study, we selected three different window sizes of 50 ms, 150 ms, and 250 ms and a stride of 25 ms for comparison with similar works. Figure 1 presents the segmentation and combination of sEMG signals. Figure 2 shows the process of segmentation.

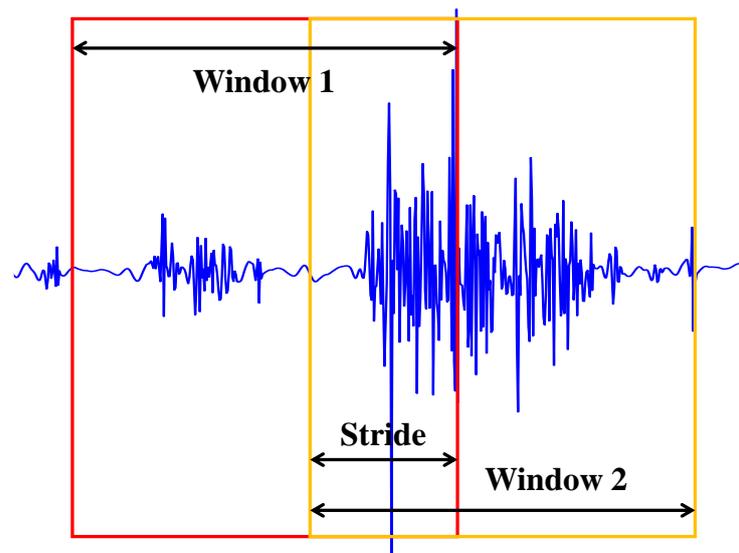


Figure 2. Process of segmentation. The red box indicates the preceding sliding window, the yellow box indicates the subsequent sliding window. The stride represents the overlapping region between the two sliding windows.

3.3. Network Settings and Evaluation Indicators

The proposed architecture was implemented within the PyTorch 1.11 deep learning framework and trained using the AdamW optimizer, which demonstrated faster empirical convergence and achieved higher accuracy in this work. Model training parameters were carefully selected through a series of experiments. The dropout rate was set to 0.2, and the model was trained for 200 epochs with a batch size of 1024. To ensure reproducibility, the parameters were initialized using a fixed random seed, and the initial learning rate was set to 5×10^{-3} . All experiments are conducted on a computer with an Intel CPU i5-13600KF 3.50 GHz and an Nvidia RTX 4090 GPU, which are sourced from Intel Corporation and Nvidia Corporation, respectively, in Santa Clara, USA. The detailed structure and parameters of the proposed model are listed in the Table 2.

Table 2. Detailed parameters of each block used in the proposed model.

Block	Type	Output Shape	Kernel Size	Activation Unit	Parameters
Patch embedding	1D Convolution	[1, 128, 16]	5	GELU	10,368
	Layernorm	[1, 128, 16]	-	-	-
	1D Convolution	[1, 128, 16]	3	GELU	49,280
	Dropout(0.2)	[1, 128, 16]	-	-	-
Partial block	1D Convolution	[1, 32, 16]	3	-	4128
	Softmax	[1, 32, 16]	-	-	-
	1D Convolution	[1, 32, 16]	1	-	3104

Table 2. Cont.

Block	Type	Output Shape	Kernel Size	Activation Unit	Parameters
Channel selection block	1D Convolution	[1, 64, 16]	3	Sigmoid	24,640
	1D Convolution	[1, 64, 16]	3	-	24,640
Classification head	AdaptiveAvgPool	[1, 64, 1]	-	-	-
	FC layer	[1, 256]	-	GELU	16,640
	LayerNorm	[1, 256]	-	-	512
	Dropout(0.2)	[1, 256]	-	-	-
	FC layer	[1, 256]	-	GELU	65,792
	LayerNorm	[1, 256]	-	-	512
	Dropout(0.2)	[1, 256]	-	-	-
FC layer	[1, 53]	-	-	13,621	

During training, all the data are randomly shuffled. To address potential distribution differences between the training and test data, a total trial is used for the training and testing [26]. Specifically, NinaPro DB4 trials numbered 1, 3, 4, and 6 of all 10 subjects were used for training, while trials numbered 2 and 5 were used for testing.

In this work, the performance of methods is evaluated by classification accuracy and F1-score. Accuracy is the most commonly used classification evaluation metric. The formula is defined by Equation (7).

$$Accuracy = \frac{\text{Correctly predicted samples}}{\text{All samples}} \times 100 \quad (7)$$

The F1-score is selected as another evaluation index due to the presence of a large number of similar actions in the data, which, together with accuracy, forms the evaluation index system of the method. Its formula is defined by Equation (8).

$$F1 \text{ score} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 2 \quad (8)$$

in which precision and recall are determined by the following equation:

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN} \quad (9)$$

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.

4. Experiment and Results

4.1. Evaluation of the Proposed Network

Our proposed method's classification accuracy and F1-score on the Ninapro DB4 and DB5 and the BioPatRec DB1, DB2, and DB3 datasets are presented in Table 3. To comprehensively assess the model's performance, we conducted tests with window sizes of 50 ms, 150 ms, and 250 ms. For Ninapro DB4, the classification accuracy and F1-score were 83.0% and 83.1%, respectively, with a window size of 250 ms. In the case of Ninapro DB5, a window size of 250 ms resulted in a classification accuracy of 87.6% and an F1-score of 87.4%. For BioPatRec DB1, DB2, and DB3, the best results were obtained with a window size of 250 ms, with identical classification accuracies and F1-scores of 91.4%, 91.3%, and 91.3%, respectively. This consistency demonstrates the stability of the model.

Our experimental results suggest that larger window sizes generally lead to higher accuracy but introduce increased latency. Striking a balance between accuracy and latency is crucial. The results indicate that a window size of 150 ms represents a favorable choice, offering performance close to the 250 ms window with minimal performance loss while reducing real-time recognition latency, and it outperforms the 50 ms window in terms of accuracy. Figure 3 illustrates the confusion matrices for BioPatRec DB1 at 50 ms, 150 ms,

and 250 ms. It is apparent from the figures that the model can correctly predict most classes, achieving single-class recognition accuracies of 94%, 95%, and 96% for the three different window sizes, respectively.

Table 3. Performance analysis using accuracy and F1-score with different window sizes for different databases.

Database	Evaluation Metric	50 ms	150 ms	250 ms
NinaPro DB4	Accuracy	77.0	81.2	83.0
	F1-score	76.5	81.1	82.7
NinaPro DB5	Accuracy	82.0	87.5	88.3
	F1-score	81.7	87.4	88.2
BioPatRec DB1	Accuracy	84.1	89.7	91.5
	F1-score	84.1	89.0	91.2
BioPatRec DB2	Accuracy	84.3	88.4	91.6
	F1-score	84.3	88.2	91.6
BioPatRec DB3	Accuracy	84.7	89.0	90.9
	F1-score	84.7	88.7	90.6

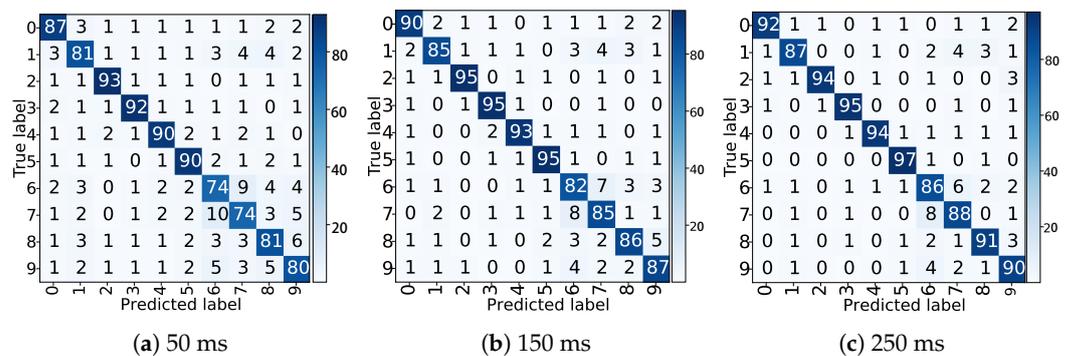


Figure 3. The normalized confusion matrix of BioPatRec DB1 with different window sizes.

To further analyze the performance of the model and attempt to visualize it, the machine learning method t-SNE (t-distributed Stochastic Neighbor Embedding) [35] was employed to perform dimensionality reduction analysis on the model’s output. This was done to measure the Euclidean distance between different categories and samples of the same category, thereby facilitating an analysis of the model’s performance. t-SNE is a nonlinear dimensionality reduction technique designed to map high-dimensional data into two or three-dimensional space for visualization. By applying t-SNE for nonlinear dimensionality reduction and data visualization techniques, we conducted an in-depth visual analysis of the model’s classification performance. It effectively projected complex high-dimensional feature information into a two-dimensional space. As illustrated in Figure 4, as the window segment length increased from 50 ms to 250 ms, the model’s performance in recognizing different hand gesture actions progressively improved. This improvement was evident in the increased clustering of similar gestures and the more pronounced distinction between different gestures. Notably, when the window lengths were set at 150 ms and 250 ms, there was no significant difference in the model’s performance in classifying various hand gestures. This finding suggests that within a certain range of window lengths, further increasing the length has a limited effect in terms of enhancing the model’s classification capabilities.

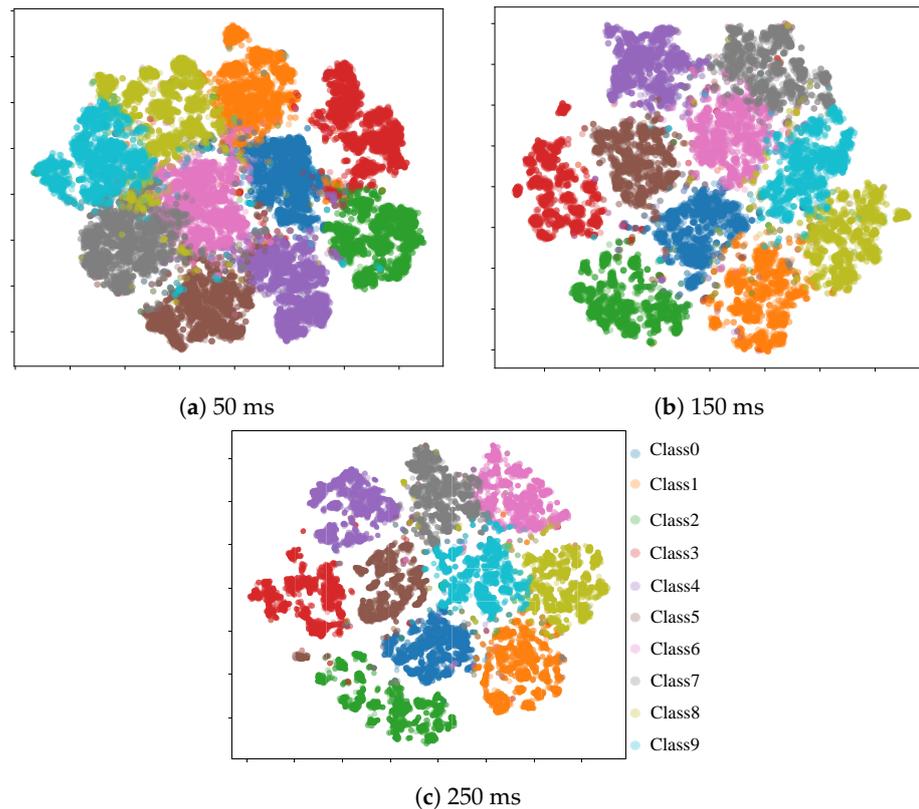


Figure 4. The t-SNE visualization of BioPatRec DB1 with different window sizes.

4.2. Ablation Studies of the Proposed Network

In this section, to demonstrate the effectiveness of our proposed feature extraction method with partial convolutional fusion, we conduct ablation experiments. Additionally, we perform ablation experiments on the channel selection unit to demonstrate its capability in reducing computational complexity and compressing the channel features. We compare the recognition accuracy of partial convolution with feature importance learning. These comparisons were performed continuously with three different window sizes (50 ms, 150 ms, and 250 ms) on the NinaPro DB4 and DB5 and BioPatRec DB1, DB2, and DB3 datasets while maintaining constant experimental settings. The results are shown in Table 4.

Table 4. Classification accuracy (%) ablation of the proposed method.

Database	Ablation	Params (M)	50 ms	150 ms	250 ms
NinaPro DB4	Full	0.22	77.0	81.2	83.0
	Without partial block	0.21	76.0	80.6	82.1
	Without channel selection	0.48	77.7	81.8	83.3
NinaPro DB5	Full	0.21	82.0	87.5	88.3
	Without partial block	0.20	80.5	85.3	86.3
	Without channel selection	0.47	82.1	87.4	87.3
BioPatRec DB1	Full	0.20	84.1	89.7	91.5
	Without partial block	0.19	81.8	86.8	87.8
	Without channel selection	0.45	84.4	90.3	92.3
BioPatRec DB2	Full	0.21	85.0	88.4	91.6
	Without partial block	0.20	83.9	86.2	89.7
	Without channel selection	0.46	85.5	89.2	91.8
BioPatRec DB3	Full	0.20	85.4	89.0	90.9
	Without partial block	0.19	80.9	84.8	86.8
	Without channel selection	0.45	85.6	89.1	90.2

As shown in Table 4, the use of partial convolution consistently enhances accuracy across all datasets. Across the five datasets and three different window sizes, the average improvement in classification accuracy due to partial convolution is 3%, 2%, 2.23%, 1.03%, and 3.93%, respectively. This highlights the important role of partial convolution, as it is effective in extracting valuable features from the sEMG signals, thus improving the training efficiency. In addition, the introduction of the channel selection unit can significantly reduce the number of parameters with minimal impact on accuracy.

To validate the effectiveness of the proposed model in channel selection, a comprehensive analysis was conducted using Shapley Additive exPlanations (SHAP), which is python package shap 0.44.1, as illustrated in Figure 5. Initially, SHAP was employed to delve into the correlation between each sEMG channel and the final gesture action. Specifically, we compared the model integrating both a partial block and a channel selection block with the models lacking either of these blocks in terms of the correlation scores between channel features and gesture action outcomes. Figure 5a,c reveal that the introduction of the channel selection block did not alter the model’s classification performance, nor did it significantly change the focus on respective channel features. However, the incorporation of this block led to a nearly 50% reduction in the model’s parameter count, substantially decreasing its size and thereby highlighting the significant role of the channel selection block in filtering and focusing on relevant channel features. Concurrently, Figure 5a,b demonstrate the impact of the presence or absence of the partial block on the model’s performance in channel feature extraction, as well as a more dispersed representation of related channel features.

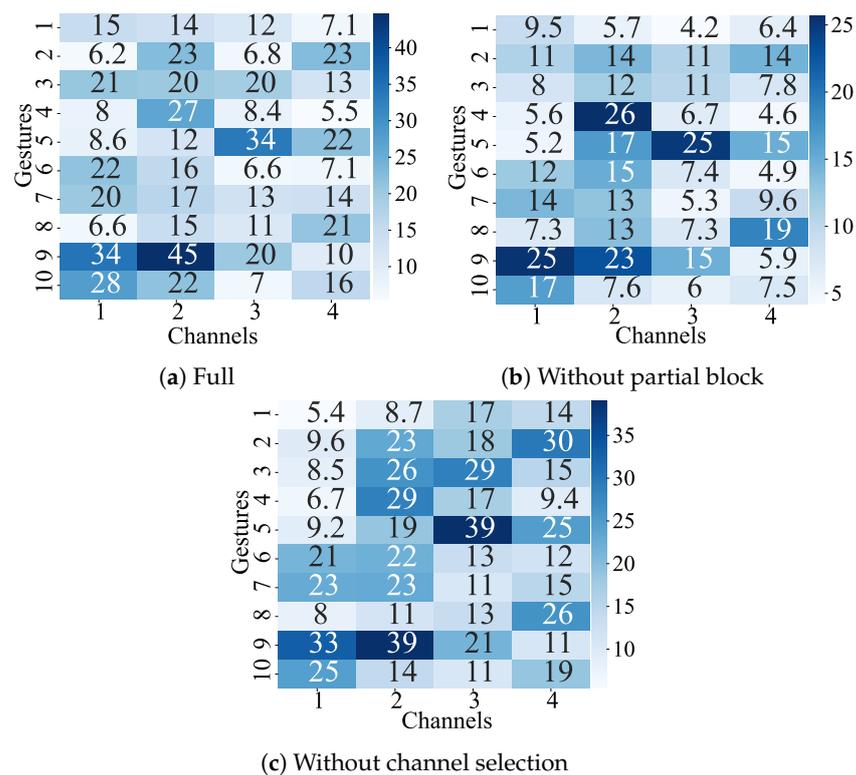


Figure 5. The SHAP visualization of our model relevance scores of channels (4 channels on x-axis) for different gestures in BioPatRec DB1 with a window size of 250 ms. The various colors represent the differential emphasis of model on signal channels corresponding to distinct actions. Intensified colors denote heightened model focus on particular signal channels, indicative of heightened correlation with the final classification outcome. Conversely, light colors suggest diminished focus, reflecting a weakened correlation with the final classification outcome.

To further showcase the model’s focus on channel features, a segment of the sEMG signal corresponding to the agree gesture and fine grip gesture in the BioPatRec DB1

dataset were selected as a case study, as shown in Figures 6 and 7. The red highlight areas in Figures 6 and 7 align with the model's focus on channel features for the corresponding action, further substantiating the precision and practicality of our model in terms of channel feature attention.

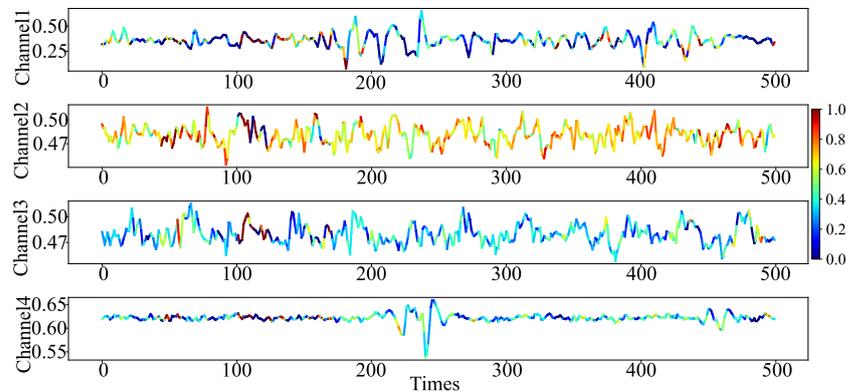


Figure 6. The SHAP visualization of our model for the agree gesture (class 9) in BioPatRec DB1 with a window size of 250 ms and 4 channels (red highlight areas indicate higher correlations).

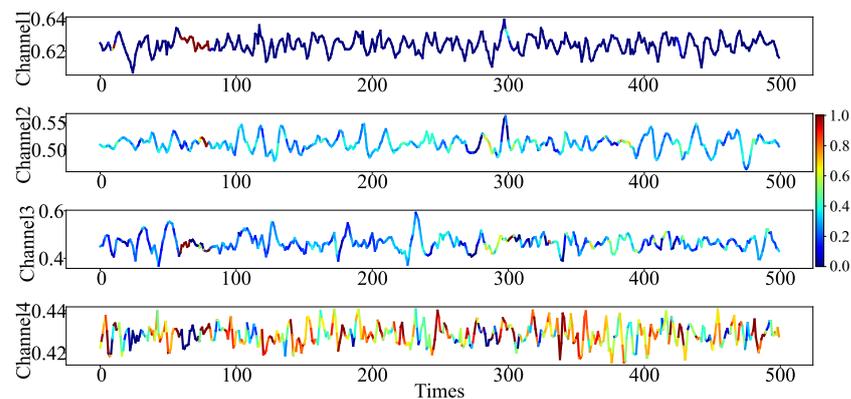


Figure 7. The SHAP visualization of our model for the fine grip gesture (class 8) in BioPatRec DB1 with a window size of 250 ms and 4 channels (red highlight areas indicate higher correlations).

4.3. Comparison With the Current Gesture Recognition Methods

In this section, we conduct a performance comparison with existing research methods on five publicly available datasets, namely Multi-view CNN [26], HVPN [36], Attention sEMG [37], EMGHandNet [38], TDCT [39], and SE-CNN [40]. The SE-CNN [40] method utilizes a strategy that integrates squeeze-and-excite (SE) modules within a CNN framework to suppress irrelevant features while enhancing important ones. This approach integrates SE, CNN, and attention mechanisms to recognize features. EMGHandNet [38] is built upon a hybrid architecture of CNN and Bi-LSTM for the learning of inter-channel and temporal features. Spatial and short-term temporal relationships are encoded by convolutional layers, while long-term temporal relationships are learned by the Bi-LSTM layer. The proposed framework is capable of extracting cross-channel and temporal features, where one-dimensional convolution encodes cross-channel and short-term temporal information, while Bi-LSTM encodes long-term temporal information in both forward and backward directions. In contrast, Attention sEMG [37] and TDCT [39] are variants based on attention mechanisms to extract global feature information, facilitating the learning of global long-term features for gesture classification. Attention sEMG utilizes a feed-forward simple attention mechanism to extract representation features in the time domain from multiple channels, while TDCT enhances the extraction of local temporal and channel correlation features in sEMG by replacing the linear transformation in the multi-head self-attention mechanism with temporal depth convolution. This modification boosts feature-learning

capabilities and decreases parameter size. Both Multi-view CNN [26] and HVPN [36] employ similar approaches by integrating multiple sets of feature information. They extract feature information from various perspectives and integrate it using different network architectures. In multi-view learning, each perspective has specific viewpoint features, and all perspectives can access common viewpoint features. These models learn feature information from multi-channel sEMG signals from both local and global perspectives. The key aspect of these models lies in their rich feature information, which may, however, introduce feature redundancy.

Due to variations in the window sizes used by different models, for the sake of fairness, we unify the comparison based on each model's best classification results, as shown in Table 5.

Table 5. Classification accuracy (%) with different models. Bold indicates the best results under different models for the same dataset.

Database/ Methodology	NinaPro DB4	NinaPro DB5	BioPatRec DB1	BioPatRec DB2	BioPatRec DB3
Multi-View CNN [26]	54.3	90.0	78.2	94.0	50.5
HVPN [36]	67.0	87.1	-	-	-
Attention sEMG [37]	73.0	87.0	-	-	-
EMGHandNet [38]	89.5	-	-	83.9	-
TDCT [39]	-	72.8	-	-	-
SE-CNN [40]	77.6	87.4	-	-	-
Ours	83.0	88.3	91.5	91.6	90.9

From Table 5, it is evident that our proposed model delivers strong performance across multiple datasets, achieving an average classification accuracy of 88.98% across the five datasets. For the NinaPro DB4 dataset, our model achieved an accuracy of 83.3%—slightly lower than EMGHandNet's 89.5%. However, it is important to note that our model maintains a significantly lower parameter count, with only 0.20 million parameters compared to EMGHandNet's 6.4 million. This underscores the effectiveness of our approach.

For the Ninapro DB5 dataset, our model achieved a classification accuracy of 87.6%, coming close to the best results. In contrast, Multi-View CNN exhibits notable instability and significant performance variations across different datasets, with an average classification accuracy of 73.4% across the five datasets, trailing our model by 15.4%. Our model leverages compact convolution to learn temporal dependencies and dynamically model correlations between different channels, enabling strong generalization across diverse datasets.

Similar to the Ninapro DB5 results, for the BioPatRec DB2 dataset, our model maintained excellent classification performance, outperforming EMGHandNet by a substantial margin of 7.4%. For the BioPatRec DB1 and BioPatRec DB3 datasets, our model achieved remarkable classification accuracies of 91.4% and 91.3%, respectively, significantly surpassing Multi-View CNN. While our model may not achieve the absolute highest accuracy on certain databases, it consistently ranks second overall, comparable to the average accuracy of all other models, and this is accomplished with an impressively minimal parameter count.

4.4. Computational Complexity Analysis

In this section, we compare the computational complexity of the model proposed in this study with popular sEMG-based hand gesture recognition methods to highlight the advantages of our approach in managing complexity. Drawing on the previously mentioned dataset, we selected the widely used NinaPro dataset as a benchmark and compared models with window lengths of up to 300 ms. Given the lack of a unified standard for window lengths in prior research, we opted for a comparison using a 250 ms window length.

In Table 6, we present the recognition performance and parameter size of current popular gesture recognition models on the NinaPro DB4 and DB5 datasets. Compared to the existing optimal model, EMGHandNet, which boasts an accuracy of 89.5%, our proposed

model is slightly behind by 6.5% in accuracy. However, in terms of parameter size, our model accounts for only 3.8% of EMGHandNet. This significant difference demonstrates that our model substantially reduces parameter size while maintaining relatively high accuracy, thereby greatly saving on the resource consumption of the model. Apart from the optimal model, compared to other models, our model exhibits clear advantages in both accuracy and parameter size.

Table 6. Comparison of the accuracy and computational complexity on NinaPro databases with different models. Bold indicates the best results under different models for the same dataset.

Database	Methodology	Accuracy (%)	Number of Parameters (M)
NinaPro DB4	Multi-View CNN [26]	54.3	3.44
	HVPN [36]	67.0	4.76
	Attention sEMG [37]	73.0	1.44
	EMGHandNet [38]	89.5	5.76
	Ours	83.0	0.22
NinaPro DB5	Multi-View CNN [26]	90.0	4.03
	HVPN [36]	87.1	5.54
	Attention sEMG [37]	87.0	1.44
	TDCT [39]	72.8	1.20
	Ours	88.3	0.21

In comparison to the HVPN and Multi-View CNN models, which employ a multi-feature approach on the NinaPro DB5 dataset, our proposed model exhibits a slight decrease in accuracy—approximately 1.7% lower than that of the Multi-View CNN model. However, our model achieves this performance with only 5% of its parameter scale. Unlike methods focusing on the fusion of multiple features, our approach, centered on channel feature selection and compression, appears to be better-suited for relevant recognition tasks. Despite a minor performance loss compared to the optimal model, the reductions in parameter scale and computational complexity contribute to an overall enhancement of our model’s performance. Additionally, when contrasted with attention mechanism-based models such as Attention sEMG and TDCT, our model outperforms them both in terms of accuracy and parameter count. This suggests that in recognition tasks, improving model performance through local feature attention and the selection of key features, as opposed to a global focus on sEMG signal features, may be pivotal.

5. Conclusions

In this paper, we introduce a PCS-EMGNet based on a channel selection mechanism powered by gating units mechanisms to tackle the enduring challenges within the domain of sEMG gesture recognition. By recognizing the imperative for more precise and computationally efficient methodologies, our proposed model significantly enhances feature extraction capabilities, yielding a model that is both more generalizable and robust. The empirical validation of our model on diverse datasets, including NinaPro DB4, NinaPro DB5, BioPatRec DB1, BioPatRec DB2, and BioPatRec DB3, demonstrated an impressive average classification accuracy of 88.34%, with a notable increase in accuracy and a substantial reduction in computational requirements compared to existing methods. Our findings indicate that our model not only achieves superior performance in terms of accuracy but also in computational efficiency, marking a significant advancement in the practical application of sEMG gesture recognition technologies.

Furthermore, by incorporating a novel partial convolution module alongside a channel feature selection unit, our approach effectively capitalizes on the redundancy within feature maps. This methodical application of convolution to select portions of input channels drastically reduces the model’s parameter count, thereby streamlining the computational

process. This innovation is poised to set a new benchmark for resource efficiency in the field of sEMG gesture recognition.

Gesture recognition through sEMG signals is rapidly gaining traction for its non-invasive nature, ease of acquisition, and high signal stability, making it an exemplary choice for dynamic gesture recognition applications. However, the field continues to grapple with challenges stemming from the diversity and complexity of gestures, as well as vulnerability to signal interference. Current gesture recognition techniques are often limited by their low accuracy and high computational demands. In response to these challenges, future research will focus on optimizing model architectures, advancing feature extraction methods, and broadening the scope of this technology's applications to tackle a wider array of real-world issues, thereby enhancing the accessibility and efficacy of sEMG-based gesture recognition systems.

Author Contributions: Conceptualization, Y.N. and W.C.; formal analysis, B.X.; methodology, Y.N. and W.C.; validation, H.Z. and Z.G.; writing—original draft, Y.N. and W.C.; 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 original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

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

References

- Wei, J.; Meng, Q.; Badii, A. Classification of Human Hand Movements Using Surface EMG for Myoelectric Control. In *Advances in Computational Intelligence Systems: Contributions Presented at the 16th UK Workshop on Computational Intelligence, Lancaster, UK, 7–9 September 2016*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 331–339.
- 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](#)] [[PubMed](#)]
- Fan, Y.; Yin, Y. Active and Progressive Exoskeleton Rehabilitation Using Multisource Information Fusion From EMG and Force-Position EPP. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 3314–3321. [[CrossRef](#)] [[PubMed](#)]
- Thielbar, K.O.; Triandafilou, K.M.; Fischer, H.C.; O'Toole, J.M.; Corrigan, M.L.; Ochoa, J.M.; Stoykov, M.E.; Kamper, D.G. Benefits of Using a Voice and EMG-Driven Actuated Glove to Support Occupational Therapy for Stroke Survivors. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 297–305. [[CrossRef](#)] [[PubMed](#)]
- 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](#)] [[PubMed](#)]
- Merletti, R.; Botter, A.; Barone, U. *Surface Electromyography: Physiology, Engineering, and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2016; Chapter 3, pp. 1–37.
- Gu, Y.; Yang, D.; Huang, Q.; Yang, W.; Liu, H. Robust EMG pattern recognition in the presence of confounding factors: Features, classifiers and adaptive learning. *Expert Syst. Appl.* **2018**, *96*, 208–217. [[CrossRef](#)]
- 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](#)]
- Kakoty, N.M.; Hazarika, S.M.; Gan, J.Q. EMG Feature Set Selection Through Linear Relationship for Grasp Recognition. *J. Med. Biol. Eng.* **2016**, *36*, 883–890. [[CrossRef](#)]
- Kim, J.; Mastnik, S.; André, E. EMG-based hand gesture recognition for realtime biosignal interfacing. In *Proceedings of the 13th International Conference on Intelligent User Interfaces: Association for Computing Machinery, Gran Canaria, Spain, 13–16 January 2008*; pp. 30–39. [[CrossRef](#)]

11. Robinson, C.P.; Li, B.; Meng, Q.; Pain, M.T. Pattern Classification of Hand Movements using Time Domain Features of Electromyography. In Proceedings of the 4th International Conference on Movement Computing: Association for Computing Machinery, London, UK, 28–30 June 2017; pp. 27:1–27:6. [[CrossRef](#)]
12. Menon, R.; Di Caterina, G.; Lakany, H.; Petropoulakis, L.; Conway, B.A.; Soraghan, J.J. Study on Interaction Between Temporal and Spatial Information in Classification of EMG Signals for Myoelectric Prostheses. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 1832–1842. [[CrossRef](#)] [[PubMed](#)]
13. Chen, B.W.; Chen, C.Y.; Wang, J.F. Smart Homecare Surveillance System: Behavior Identification Based on State-Transition Support Vector Machines and Sound Directivity Pattern Analysis. *IEEE Trans. Syst. Man, Cybern. Syst.* **2013**, *43*, 1279–1289. [[CrossRef](#)]
14. Wu, Z.; Wang, X.; Jiang, Y.G.; Ye, H.; Xue, X. Modeling Spatial-Temporal Clues in a Hybrid Deep Learning Framework for Video Classification. In Proceedings of the 23rd ACM International Conference on Multimedia: Association for Computing Machinery, Brisbane, Australia, 26–30 October 2015; pp. 461–470. [[CrossRef](#)]
15. Khomami, S.A.; Shamekhi, S. Persian sign language recognition using IMU and surface EMG sensors. *Measurement* **2021**, *168*, 108471. [[CrossRef](#)]
16. Al-Timemy, A.H.; Khushaba, R.N.; Bugmann, G.; Escudero, J. Improving the Performance Against Force Variation of EMG Controlled Multifunctional Upper-Limb Prostheses for Transradial Amputees. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *24*, 650–661. [[CrossRef](#)] [[PubMed](#)]
17. Tan, T.S.; Lum, K.Y.; Anuar, R.; Yahya, Z.; Yahya, A.B.; Kadir, M.R.A. Sign language recognition system using SEMG and hidden markov model. In Proceedings of the Conference on Recent Advances in Mathematical Methods Intelligent Systems and Materials, Lemesos, Cyprus, 19–24 March 2013; pp. 50–53.
18. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
19. Graves, A.; Mohamed, A.R.; Hinton, G. Speech recognition with deep recurrent neural networks. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 6645–6649. [[CrossRef](#)]
20. Atzori, M.; Gijssberts, A.; Castellini, C.; Caputo, B.; Hager, A.G.M.; Elsig, S.; Giatsidis, G.; Bassetto, F.; Müller, H. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Sci. Data* **2014**, *1*, 140053. [[CrossRef](#)] [[PubMed](#)]
21. Jia, G.; Lam, H.K.; Liao, J.; Wang, R. Classification of electromyographic hand gesture signals using machine learning techniques. *Neurocomputing* **2020**, *401*, 236–248. [[CrossRef](#)]
22. Zhai, X.; Jelfs, B.; Chan, R.H.M.; Tin, C. Self-Recalibrating Surface EMG Pattern Recognition for Neuroprosthesis Control Based on Convolutional Neural Network. *Front. Neurosci.* **2017**, *11*, 379. [[CrossRef](#)] [[PubMed](#)]
23. Koch, P.; Dreier, M.; Maass, M.; Phan, H.; Mertins, A. RNN With Stacked Architecture for sEMG based Sequence-to-Sequence Hand Gesture Recognition. In Proceedings of the 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, The Netherlands, 18–21 January 2021; pp. 1600–1604. [[CrossRef](#)]
24. Ketykó, I.; Kovács, F.; Varga, K.Z. Domain Adaptation for sEMG-based Gesture Recognition with Recurrent Neural Networks. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–7. [[CrossRef](#)]
25. Amor, A.B.H.; Ghoul, O.; Jemni, M. Toward sign language handshapes recognition using Myo armband. In Proceedings of the 2017 6th International Conference on Information and Communication Technology and Accessibility (ICTA), Muscat, Oman, 19–21 December 2017; pp. 1–6. [[CrossRef](#)]
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. 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](#)] [[PubMed](#)]
28. Jia, Z.; Cai, X.; Jiao, Z. Multi-Modal Physiological Signals Based Squeeze-and-Excitation Network With Domain Adversarial Learning for Sleep Staging. *IEEE Sens. J.* **2022**, *22*, 3464–3471. [[CrossRef](#)]
29. 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](#)] [[PubMed](#)]
30. 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](#)]
31. Hudgins, B.; Parker, P.; Scott, R. A new strategy for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.* **1993**, *40*, 82–94. [[CrossRef](#)]
32. Ding, Z.; Yang, C.; Tian, Z.; Yi, C.; Fu, Y.; Jiang, F. sEMG-Based Gesture Recognition with Convolution Neural Networks. *Sustainability* **2018**, *10*, 1865. [[CrossRef](#)]
33. 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.* **2011**, *19*, 186–192. [[CrossRef](#)] [[PubMed](#)]

34. Kuzborskij, I.; Gijbsberts, A.; Caputo, B. On the challenge of classifying 52 hand movements from surface electromyography. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 4931–4937. [[CrossRef](#)]
35. van der Maaten, L.; Hinton, G. Visualizing High-Dimensional Data Using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
36. Wei, W.; Hong, H.; Wu, X. A Hierarchical View Pooling Network for Multichannel Surface Electromyography-Based Gesture Recognition. *Comput. Intell. Neurosci.* **2021**, *2021*, 6591035. [[CrossRef](#)] [[PubMed](#)]
37. Josephs, D.; Drake, C.; Heroy, A.; Santerre, J. sEMG Gesture Recognition with a Simple Model of Attention. In Proceedings of the Machine Learning for Health NeurIPS Workshop, Virtual, 11 December 2020; Voume 136, pp. 126–138.
38. Karnam, N.K.; Dubey, S.R.; Turlapaty, A.C.; Gokaraju, B. EMGHandNet: A hybrid CNN and Bi-LSTM architecture for hand activity classification using surface EMG signals. *Biocybern. Biomed. Eng.* **2022**, *42*, 325–340. [[CrossRef](#)]
39. Wang, Z.; Yao, J.; Xu, M.; Jiang, M.; Su, J. Transformer-based network with temporal depthwise convolutions for sEMG recognition. *Pattern Recognit.* **2024**, *145*, 109967. [[CrossRef](#)]
40. Xu, Z.; Yu, J.; Xiang, W.; Zhu, S.; Hussain, M.; Liu, B.; Li, J. A Novel SE-CNN Attention Architecture for sEMG-Based Hand Gesture Recognition. *Comput. Model. Eng. Sci.* **2023**, *134*, 157–177. [[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.