



# TMCA-Net: A Compact Convolution Network for Monitoring Upper Limb Rehabilitation

Qi Liu<sup>1</sup>, Zihao Wu<sup>1,\*</sup> and Xiaodong Liu<sup>2</sup>

<sup>1</sup>The School of Computer Science, Nanjing University of Information Science and Technology, Nanjing, 210044, China

<sup>2</sup>The School of Computing, Edinburgh Napier University, Edinburgh, EH10 5DT, UK

\*Corresponding Author: Zihao Wu. Email: 20201221048@nuist.edu.cn

Received: 15 February 2023; Accepted: 24 March 2023; Published: 14 June 2023

**Abstract:** This study proposed a lightweight but high-performance convolution network for accurately classifying five upper limb movements of arm, involving forearm flexion and rotation, arm extension, lumbar touch and no reaction state, aiming to monitor patient's rehabilitation process and assist the therapist in elevating patient compliance with treatment. To achieve this goal, a lightweight convolution neural network TMCA-Net (Time Multiscale Channel Attention Convolutional Neural Network) is designed, which combines attention mechanism, uses multi-branched convolution structure to automatically extract feature information at different scales from sensor data, and filters feature information based on attention mechanism. In particular, channel separation convolution is used to replace traditional convolution. This reduces the computational complexity of the model, decouples the convolution operation of the time dimension and the cross-channel feature interaction, which is helpful to the target optimization of feature extraction. TMCA-Net shows excellent performance in the upper limb rehabilitation gesture data, achieves 99.11% accuracy and 99.16% F1-score for the classification and recognition of five gestures. Compared with CNN and LSTM network, it achieves 65.62% and 89.98% accuracy in the same task. In addition, on the UCI smartphone public dataset, with the network parameters of one tenth of the current advanced model, the recognition accuracy rate of 95.21% has been achieved, which further proves the light weight and performance characteristics of the model. The clinical significance of this study is to accurately monitor patients' upper limb rehabilitation gesture by an affordable intelligent model as an auxiliary support for therapists' decision-making.

**Keywords:** Rehabilitation monitoring; stroke; MEMS; gesture recognition

## 1 Introduction

Upper extremity hemiparesis is commonly seen in the stroke survivor population, where patients have to repetitively undergo extensive upper extremity rehabilitation under the guidance of a physician or professional rehabilitation therapist, in order to continuously stimulate damaged brain nerves,



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

remodel brain neural circuits, and achieve motor function recovery [1]. Traditional inpatient rehabilitation requires substantial human and medical resources, resulting in high costs that impose a significant burden on both society and individuals. Home-based remote rehabilitation systems, on the other hand, offer low costs and no time or space restrictions, enabling patients to engage in rehabilitation training anytime and anywhere [2]. Remote stroke rehabilitation systems monitor and recognize patients' posture, range of motion, and movement quality during rehabilitation training, record and provide feedback to rehabilitation therapists to timely assess patients' rehabilitation progress and formulate personalized treatment plans. Among them, camera-based rehabilitation monitoring systems are more mature, but suffer from limitations such as limited field of view, vulnerability to occlusion and lighting factors, and privacy concerns related to the personal data collected [3]. Inertial sensor-based rehabilitation monitoring systems, on the other hand, have better stability, and with the widespread commercial use of MEMS inertial sensors and the reduction of costs, more research is focusing on how to achieve stable and accurate rehabilitation monitoring with fewer inertial devices. Monitoring system based on wearable inertial sensors and intelligent algorithms are more stable, economical and secure. How to model medical data with the goal of high accuracy upper limb gesture recognition is the problem needs to be solved. Previous studies have achieved some substantive progress, but the cost of calculation and memory consumption of the model limit its rehabilitation monitoring in the home environment. Furthermore, Feature extraction is the key to rehabilitation gesture recognition algorithm research [4]. Classifiers based on hand-crafted feature in traditional machine learning methods rely on feature extraction design and proper selection of features, which is inefficient and complex. Existing stacked deep convolution neural networks can extract abstract features based on original sensor sequences, but they have problems such as high computational complexity and slow convergence. Not suitable for home equipment with limited computing resources. In response to the above problems and challenges, this paper presents an accurate, lightweight and fast convergence customized convolution neural network, TMCA-Net (Time Multiscale Channel Attention Convolutional Neural Network), which can accurately recognize patients' upper limb rehabilitation gesture based on the data recorded by a single inertial sensor fixed to the forearm.

## 2 Related Work

Human gesture recognition has potential applications in human-computer interaction, medical rehabilitation, virtual reality and other fields. Currently, there are three main types of gesture recognition methods: based on vision, based on electromyography (EMG), and based on the MEMS sensor. With the rapid development of computer vision technology, visual-based human gesture recognition is becoming mature and widely used in various fields, Agarwal et al. recovering 3D human gesture from image sequences [5]. Fahn et al. proposed a real-time upper limb gesture recognition method based on particle filter and AdaBoost algorithm [6]. Brattoli et al. proposed a self-supervised LSTM for detailed behavioral analysis, which learns accurate representations of gesture and behavior through self-supervision to analyze motor function [7]. Htike et al. analyzed human activity [8] by converting video data into a dataset of static color images and using recognized gesture sequences. Chen et al. proposed Kinect data combined with deep convolutional neural network to estimate human gesture [9]. Neili et al. used CNN to predict the 2-dimensional spatial position of joint points as gesture characteristics to monitor whether accidents occurred in the elderly in order to help monitor their activities in the home environment [10].

Compared with the vision-based monitoring system, sensor-based method is not limited by environmental impact and measurement range. More importantly, sensor does not record sensitive data of patients, EMG is a technology that uses sensors to record human skeletal muscle electrical activity. It is often used in EMG data features for gesture recognition [11]. It has also been developing in recent years. It has the characteristics of low cost and high efficiency in clinical rehabilitation training [12], such as Lu et al. extracted features from the EMG data of the forearm muscles of stroke patients for human-computer interaction to achieve the purpose of auxiliary training [13], Bi et al. proposed a multi-signature reconstruction system based on the forearm EMG signal to strengthen the patient's autonomous control of FES and improve the therapeutic effect [14]. EMG is a powerful biological information for analyzing human motion, but the problem of EMG is that the measured signal intensity is small, In particular, the signal amplitude around the wrist is low [15], and different click positions will lead to inconsistent measured signals, which leads to strict requirements on the wearing position. Because EMG is a biological signal, the signal difference between different individuals is large [16], compared with the physical signal recorded by IMU sensor, the individual difference is relatively small, and the requirement for wearing is lower, and the actual performance is more stable, but there are also measurement errors and poor anti-jamming ability [17]. The commonly used method is usually based on acceleration signal data, using hand-crafted features [18], combined with machine learning method for classification, and the design of feature set is very key. Different task feature sets may be inconsistent, which requires professional a priori knowledge. For example, Liu et al. selected the mean and standard deviation of acceleration data to construct the feature set, combined a fully connected neural network for rehabilitation of upper limb gesture recognition [19]. Ferreira et al. proposed to use the artificial neural network to analyze the dynamic data recorded by IMU, to identify the dynamic gesture of Alzheimer's disease (AD), to find the optimal number of hidden layers and neurons, and to develop a multi-layer sensor ANN for the diagnosis of Alzheimer's disease [20]. Sanna et al. proposed an IMU sensor based on the Internet of Things (IOT) to monitor user activity and gesture transitions, transfer possible events to a home server or gateway, process the final process, and store the data in a telemedicine server [21]. With the rise of deep learning and the availability of large-scale sports data, the combination of deep neural network and gesture recognition or activity recognition begins to appear. Based on large-scale data from wearable inertial sensors, um et al. used convolutional neural network (CNN) to automatically extract and identify the generated images and classified 50 gym sports items with an accuracy of 92.1% [22]. Recently, the e-Health system based on the framework of the IOT has begun to emerge. Bisio et al. proposed a Smart-Pants medical system with multiple sensors and matching Android programs for remote real-time monitoring of lower extremity training of stroke patients [23]. Jones et al. proposed a data collection method based on mobile phones. The original training data will be uploaded to the cloud server for analysis and evaluation, and medical staff can use the web system for remote access [24]. Yuan and their team have designed a data glove device that integrates two arm rings and a 3D flexible sensor. They introduced a combination of residual convolutional neural network and LSTM module to capture the fine-grained movements of the arm and all finger joints [25]. Bianco et al. implemented arm posture recognition, user identification, and identity verification based on a custom wireless wristband inertial sensor combined with a recursive neural network. The U-WeAr study found that numerical normalization preprocessing can enhance the anti-interference ability of gesture data with different amplitudes and speeds [26]. Kang et al. achieved accurate gesture recognition during walking dynamics using a wristband-type inertial sensor, by combining empirical mode decomposition with a distribution-adaptive transfer learning method [27].

In summary, the combination of deep artificial neural networks and wearable devices has been proven to be an effective rehabilitation gesture recognition solution in multiple studies. However, accurately extracting key features from sensor data is the core issue for achieving reliable gesture recognition. Convolutional neural network models have been shown to have efficient feature extraction capabilities, but previous research has mostly used monotonous same-size convolution kernels to construct feature extractors. Such network structures may fail to capture multi-scale information in sensor data. Additionally, traditional stacked convolutional networks increase model non-linear expression ability by stacking convolutional layers to deepen the network, which improves performance in gesture recognition tasks but ignores the heavy computational load that increased model complexity places on computing devices. This paper proposes a more practical rehabilitation gesture recognition model to address these issues and assist in the monitoring of stroke patients' rehabilitation.

### 3 Methodology

Stroke patients need to perform a large number of functional arm movements over a long period of time to ensure the reconstruction and recovery of upper limb function. Effective and accurate rehabilitation monitoring methods can feedback the current recovery progress of patients to develop corresponding treatment programs. Accurate recognition of upper limb gesture is the biggest challenge facing current monitoring scenes, in order to assist the efficient recovery of stroke patients, this paper will describe an arm data recording and intelligent recognition method proposed in the following two parts.

#### 3.1 Data Collection Device

Rehabilitation monitoring based on a single inertial sensor is a convenient and safe solution that avoids the need for multiple sensors on the upper extremity when performing rehabilitation actions, which can be confusing for patients or older people with inconvenient movements and may result in redundant recording data. A large number of studies have shown that upper extremity motion information can be captured based on a single inertial sensor. Based on this, this paper designs a data acquisition platform which is composed of a single inertial sensor (BWT901, Wit-motion), a PC with upper computer software, and a desktop device responsible for playing motion guidance. The inertial sensor is worn on the volunteer's forearm and wrist with a magic band. Data acquisition device is shown in [Fig. 1](#).



**Figure 1:** BWT901

BWT901 connects to PC through wireless Bluetooth to avoid limited transmission activities. The device captures the three-axis acceleration, three-axis angular velocity, motion direction and other

information when the upper limb moves at a frequency of 100 Hz. The transferred data will be retained in the format of text file and stored on the disk of PC for subsequent processing and analysis, the saved data frame is shown in Fig. 2.

14:34:32.329	0.01	0.015	1.042	0	0	0	0.067	30.058	-38.225	30.05
14:34:32.408	0.01	0.014	1.043	0	0	0	0.067	30.058	-38.225	30.05
14:34:32.528	0.01	0.012	1.044	0	0	0	0.067	30.058	-38.225	30.05
14:34:32.628	0.01	0.013	1.043	0	0	0	0.067	30.058	-38.225	30.05
14:34:32.748	0.01	0.014	1.043	0	0	0	0.1	30.075	-38.292	30.05
14:34:32.806	0.01	0.014	1.043	0	0	0	0.1	30.075	-38.292	30.05
14:34:32.928	0.009	0.013	1.043	0	0	0	0.1	30.075	-38.292	30.05
14:34:33.007	0.01	0.014	1.043	0	0	0	0.1	30.075	-38.292	30.05
14:34:33.129	0.009	0.013	1.043	0	0	0	0.092	30.158	-38.158	30.05
14:34:33.217	0.009	0.013	1.042	0	0	0	0.092	30.158	-38.158	30.05
14:34:33.346	0.01	0.015	1.043	0	0	0	0.092	30.158	-38.158	30.05
14:34:33.397	0.01	0.014	1.043	0	0	0	0.092	30.158	-38.158	30.05
14:34:33.518	0.01	0.013	1.043	0	0	0	0.075	30.075	-38.4	30.05
14:34:33.618	0.009	0.013	1.043	0	0	0	0.075	30.075	-38.4	30.05

Figure 2: The saved data frame

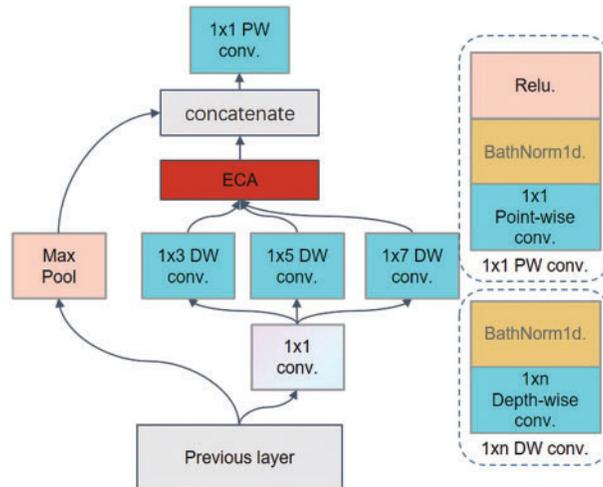
### 3.2 Proposed TMCA Net

To help therapists monitor and evaluate subsequent rehabilitation progress, portable monitoring devices and intelligent recognition algorithms are necessary in the rehabilitation monitoring system. Compared with visual sensor schemes, which encounter problems such as occlusion, illumination, privacy leaks, etc., wearable inertial sensors and intelligent algorithms-based monitoring schemes are more stable, economical and secure. The biggest challenge is the accurate extraction and recognition of posture features. Classifiers based on heuristic feature training based on traditional machine learning rely on the proper selection of feature extraction design and features, and are prone to recognition confusion caused by inadequate extraction of discriminable features. Existing stacked deep convolution neural networks can extract abstract features based on original sensor sequences, but they have problems such as high computational complexity and slow convergence. It is not suitable for home equipment with limited computing resources. This paper presents an accurate, lightweight and fast convergence customized convolution neural network TMCA-Net (Time-Multiscale Channel Attention Convolutional Neural Network). Inspired by the design idea of Network in Network Work, network can accelerate the stable convergence of network by adapting feature extraction tasks of different sizes with multilayer blocks and introducing residual connections. The proposed TMCA-Net structure is shown in Fig. 3.



Figure 3: The structure of TMCA-Net

Sensor data generated by patients during rehabilitation training have different characteristics at different time scales. For example, long-term features reflect the overall law of data during training execution, while short-term features pay more attention to the transient changes of upper limb movement data. In the feature extraction of rehabilitation inertia data, most previous studies used a fixed-size convolution core for automatic feature extraction, often ignoring the contribution of multiscale features in the activity sequence to the differentiation of different rehabilitation activities. In this paper, a multi-scale convolution module TMCA-Module based on channel constraints is designed to decouple the channel dimension convolution from the time dimension convolution. In the aspect of time feature extraction of motion data, a multi-branch structure is used to extract the multi-scale features of motion data based on different size convolution cores. In the channel dimension, pointwise  $1 \times 1$  Convolution is the feature between channels combined with the channel attention module, which fuses and selects features in the channel dimension, introduces the attention mechanism to inhibit the channel with low contribution, enhances the focus on the channel with beneficial features, and finally maps features to the correct upper limb training through the full connection layer. The detailed structure of the feature extraction block is shown in Fig. 4.



**Figure 4: TMCA-module**

The multi-branched structure in TMCA Module brings a variety of dense motion features. Among the rich and diverse sequence features, some key features are helpful to distinguish posture, but some of them can mislead posture distinction, such as outliers or noise in data, interference from gravitational acceleration characteristics in accelerometers, etc. to fit the data more accurately for the boost model. An adapted ECA [28] module is introduced to make the model more focused on the effective features while eliminating or suppressing the interference features.

#### 4 Experiment

The purpose of the experiment was to verify the feasibility and validity of the proposed scheme in the upper extremity rehabilitation monitoring. The experiment is divided into two parts. The first part is based on 3.1 data acquisition device to collect data and build data set for upper limb rehabilitation training. The second part experimentally validates the recognition performance of the model TMCA and the optimal hyper-parameter settings.

#### 4.1 Dataset Construction

Before data collection, the upper limb movements were selected and defined based on the FMA rehabilitation function evaluation scale and Brunnstrom rehabilitation stage theory. This paper focuses on the research of upper limb and arm rehabilitation monitoring. After communicating with the rehabilitation therapist, in order to ensure that the upper limb rehabilitation monitoring is accurate and not overly complex, five upper limb rehabilitation excises are designed to reflect the patients' upper limb functional status and rehabilitation progress, and have been confirmed by the rehabilitation therapist. Their specific definitions are shown in [Table 1](#).

**Table 1:** Pre-defined arm excises

Excises	Type
Initial state	No obvious arm reaction
Forearm flection	Partial synergy movement
Forearm pronation	Synergy movement
Lumbar touch	Partial isolated movement
Arm extension	Isolated movement

Initial state without noticeable reflex in the table corresponds to the arm function status of brunnstrom stage I and II patients. The forearm flexor posture can be used to determine whether the upper limbs show partial CO movement (pathological abnormal movement pattern, relying on adjacent joint muscles to compensate for completion), forearm pronation accompanied by obvious joint movement, corresponding to the three-stage state of sufficient co movement, hand touching lumbar movement accompanied by partial separation movement appears, corresponding to the fourth stage of rehabilitation, shoulder abduction is required for patients to be able to control joints and muscles to complete free separation movement, corresponding to the fifth stage of rehabilitation.

Volunteers learned the normative execution of the trained movements through upper extremity action demonstration while watching the video before data acquisition, after which they wore the calibrated inertia sensor in the dominant hand forearm position in a sitting position, with the arms naturally placed in the initial posture after wearing it. Volunteers performed the predefined postures sequentially as indicated by the video, each action was repeated to follow the rehabilitation training truth, the action execution was slightly different according to individual habits, the execution speed was slightly slower than normal, each action was performed 20 times in one group, each action was performed with an interval of 1–2 min rest between completion, and the initial posture was acquired by default for  $60 \pm 5$  s. The data acquisition process is illustrated in [Fig. 5](#).

The motion data is segmented by fixed window size 600, and the overlap rate is 50%. Unify the length and size, and construct a data sample dimension of  $600 \times 6$ . Take 100 Hz as the sampling frequency, and the actual interception time of a single sample is 6 s. In order to adapt to the data modeling of deep convolution network, the data set is divided according to the object, of which 80% is used for the source data of training set, and 20% is used for the source data of model test.



**Figure 5:** Volunteer in data collection process

#### 4.2 TMCA-Net Performance

This section validates the performance of TMCA-Net on upper extremity rehabilitation monitoring by experimentation. The stacked CNN, ResNet network and LSTM network modeled by time series are selected as comparison objects. The basic settings of the experiment are shown in [Table 2](#).

**Table 2:** Experiment setting

Object	Value
Operation system	Ubuntu 20.04
GPU device	GeFore RTX 3090
Compared model	LSTM, Bi-LSTM, CNN, ResNet, Rehab-Net [26], iSPLInceptiont [27]
Input size	(600, 6)
Epoch	128
Optimizer	Adam
Loss function	Cross entropy loss

The task of upper extremity rehabilitation monitoring and recognition using a deep convolution neural network is a multi-classification problem. This section describes the evaluation indexes used to evaluate the performance of the proposed methods and the meaning of their respective representatives. In order to verify the validity and robustness of the method, four evaluation indexes, Accuracy, precision, recall and F1-score, are used to evaluate the identification method comprehensively. The calculation formulas for each index are shown in [Table 3](#).

**Table 3:** Model evaluation indices

Object	Value
Accuracy	$\frac{(TP + TN)}{(TP + FN + FP + TN)}$
Precision	$\frac{TP}{(TP + FP)}$
Recall	$\frac{TP}{(TP + FN)}$
F1-score	$\frac{(2 * Precision * Recall)}{(Precision + Recall)}$

In the aspect of model performance test, the TMCA-Net network based on multi-scale features proposed in this paper has significant recognition accuracy advantages on the upper extremity rehabilitation posture recognition dataset by comparing the collected dataset with the public dataset. The comparative evaluation results of each model are shown in [Table 4](#). The F1-score indicates that the model performs more balanced in recognition performance, which proves the validity of the method proposed in this paper. In networks with similar performance, iSPLInception network and adapted Resnet show better performance in identifying tasks, but the large number of model parameters is the main factor that restricts the scenario applicable to traditional stacked deep convolution network. Baseline CNN network also shows good performance. Compared with LSTM, the strong feature extraction capability of CNN convolution network plays a role. Rehab-Net, a deep learning framework that focuses on upper extremity rehabilitation monitoring, performed relatively well in this experiment, inferring that recognition performance depends to some extent on pre-treatment methods. The performance of the TMCA-Net model on public datasets is shown in [Table 5](#).

**Table 4:** Model evaluation results

Model	Parameters (million)	Epoch	Accuracy	Precision	Recall	F1-score
LSTM	0.120379	128	65.62%	65.93%	67.20%	63.77%
Bi-LSTM	0.238869	128	69.64%	74.43%	71.99%	71.90%
CNN	0.135460	128	87.95%	91.68%	87.73%	85.98%
ResNet18	4.374725	128	92.86%	92.92%	92.84%	92.79%
Rehab-Net [29]	0.157519	128	91.52%	91.47%	91.64%	91.47%
IspInception [30]	1.327089	128	95.98%	96.47%	95.71%	95.66%
TMCA-Net <sup>[Our]</sup>	0.061041	128	<b>99.11%</b>	<b>99.23%</b>	<b>99.09%</b>	<b>99.16%</b>

**Table 5:** Model performance in public dataset

Dataset	Epoch	Accuracy	Precision	Recall	F1-score
UCI-HAR	128	95.21%	95.23%	95.23%	95.20%
OPPO	128	85.23%	80.35%	79.60%	79.48%
PMAP	128	95.91%	95.97%	95.40%	95.65%

The performance of the model is affected by many factors, including the size of data split window, loss function and optimizer algorithm. In this paper, the hyper-parameters of the proposed network are optimized to achieve the optimal recognition performance for the upper limb rehabilitation posture recognition task. The execution time of different upper extremity postures is inconsistent, resulting in different time span sizes of complete data. In order to explore the optimal time window for data segmentation, the sensor multivariate sequence data streams are segmented with window sizes of 400 and 600, overlapping 50%, and the effect of sample dimensions on the performance of recognition is explored, and the optimal segmented window is found. At the same time, different combinations of hyper-parameters were set up. Considering the balance of computing resources, the fixed model structure was a two-layer TMCA stacking structure. Based on the rehabilitation posture dataset, training and testing were carried out. The results of three experiments were averaged under each setting, and the accuracy of model recognition was used as the evaluation index to explore the best combination of hyper-parameters.

The experimental results in [Table 6](#) show that the performance of the split window with 400 window size is slightly worse than that with 600 window size, the performance under different windows sizes is shown in [Fig. 6](#). The main misclassification is the forearm pronation. The analysis shows that the 400 window cannot include the complete forearm pronation posture due to the long time-consuming part of the forearm pronation. The ablation results show that the optimal optimization algorithm is Adam, and Relu shows stronger task applicability in terms of activation function.

**Table 6:** Hyper-parameter setting

Model	Hyper-parameter	Value	Accuracy
TMCA-Net	Window size	400	95.74%
		600	98.11%
		Relu	97.91%
	Loss Function	ERelu	96.87%
		LeakyRelu	97.27%
		SGD	95.32%
	Optimizer	Adam	99.11%
		Adamax	94.67%

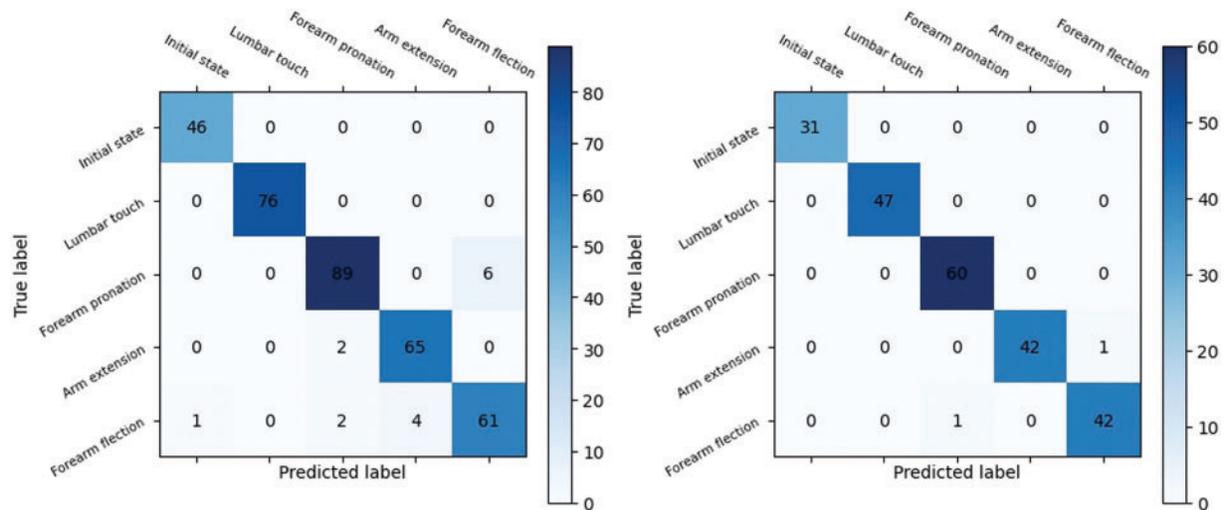


Figure 6: Performance under different windows size (left-400, right-600)

## 5 Conclusion

In view of the current problems faced by the remote upper limb rehabilitation monitoring system, this paper puts forward a set of reliable upper limb rehabilitation monitoring scheme, which uses a single inertial sensor fixed on the forearm to achieve effective motion information collection. The adaptive algorithm is also the core work of this paper. For the current feature extraction methods are complex, inaccurate and incomplete, based on the use of multi-branch convolution TMCA module. Automatically extracts multi-scale features from inertial data, combines attention mechanism to learn distinguishable features efficiently, and achieves accurate upper extremity rehabilitation movement recognition. Experiments verify the performance and task adaptability of the proposed model. Considering the limited memory of the device in the home environment, network parameters are kept at a low level by optimizing the structure design and replacing the convolution method. This is of great significance to the actual deployment of the upper extremity rehabilitation monitoring program. The future work will further focus on researching upper limb rehabilitation posture recognition and monitoring, and attempt to introduce reinforcement learning algorithms to automatically adjust applicable models in different scenarios, in order to improve the stability of monitoring solutions. Alternatively, edge computing technology may be combined to deploy deep models on edge devices, achieving low-latency remote real-time upper limb rehabilitation posture recognition.

**Acknowledgement:** Z. Wu author would like to thank the Q. Liu author for his theoretical and practical guidance, thanks to X. Liu for his writing guidance. The authors would like to thank J. Gan, H. Ling and Z. Wang for providing us with data source and images. We would like to thank the anonymous reviewers for their helpful remarks.

**Funding Statement:** This work has received funding from the Key Laboratory Foundation of National Defence Technology under Grant 61424010208, National Natural Science Foundation of China (Nos. 41911530242 and 41975142), 5150 Spring Specialists (05492018012 and 05762018039), Major Program of the National Social Science Fund of China (Grant No. 17ZDA092), 333 High-Level Talent Cultivation Project of Jiangsu Province (BRA2018332), Royal Society of Edinburgh, UK and

China Natural Science Foundation Council (RSE Reference: 62967\_Liu\_2018\_2) under their Joint International Projects funding scheme and basic Research Programs (Natural Science Foundation) of Jiangsu Province (BK20191398 and BK20180794).

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] S. I. Lee, C. P. A. Dester, M. Grimaldi, A. V. Dowling, P. C. Horak *et al.*, “Enabling stroke rehabilitation in home and community settings: A wearable sensor-based approach for upper-limb motor training,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1–11, 2018.
- [2] J. Bai and A. Song, “Development of a novel home-based multi-scene upper limb rehabilitation training and evaluation system for post-stroke patients,” *IEEE Access*, vol. 7, pp. 9667–9677, 2019.
- [3] I. Chiuchisan, D. -G. Balan, O. Geman, I. Chiuchisan and I. Gordin, “A security approach for health care information systems,” in *2017 E-Health and Bioengineering Conf. (EHB)*, Sinaia, Romania, pp. 721–724, 2017. <https://doi.org/10.1109/EHB.2017.7995525>
- [4] M. X. Sun, Y. H. Jiang, Q. Liu and X. D. Liu, “An auto-calibration approach to robust and secure usage of accelerometers for human motion analysis in FES therapies,” *Computers, Materials & Continua*, vol. 60, no. 1, pp. 67–83, 2019.
- [5] A. Agarwal and B. Triggs, “Recovering 3D human pose from monocular images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 44–58, 2006.
- [6] C. -S. Fahn and S. -L. Chiang, “Real-time upper-limbs posture recognition based on particle filters and AdaBoost algorithms,” in *2010 20th Int. Conf. on Pattern Recognition*, Istanbul, Turkey, pp. 3854–3857, 2010. <https://doi.org/10.1109/ICPR.2010.939>
- [7] B. Brattoli, U. Büchler, A. -S. Wahl, M. E. Schwab and B. Ommer, “LSTM self-supervision for detailed behavior analysis,” in *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 3747–3756, 2017. <https://doi.org/10.1109/CVPR.2017.399>
- [8] K. K. Htike, O. O. Khalifa, H. A. Mohd Ramli and M. A. M. Abushariah, “Human activity recognition for video surveillance using sequences of postures,” in *The Third Int. Conf. on e-Technologies and Networks for Development (ICeND2014)*, Beirut, Lebanon, pp. 79–82, 2014.
- [9] K. Chen and Q. Wang, “Human posture recognition based on skeleton data,” in *2015 IEEE Int. Conf. on Progress in Informatics and Computing (PIC)*, Nanjing, China, pp. 618–622, 2015. <https://doi.org/10.1109/PIC.2015.7489922>
- [10] S. Neili, S. Gazzah, M. A. El Yacoubi and N. E. Ben Amara, “Human posture recognition approach based on ConvNets and SVM classifier,” in *2017 Int. Conf. on Advanced Technologies for Signal and Image Processing (ATSIP)*, Fez, Morocco, pp. 1–6, 2017. <https://doi.org/10.1109/ATSIP.2017.8075518>
- [11] Y. Zhang, Y. Chen, H. Yu, X. Yang and W. Lu, “Learning effective spatial–Temporal features for sEMG armband-based gesture recognition,” *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6979–6992, 2020.
- [12] T. H. Wagner, A. C. Lo, P. Peduzzi, D. M. Bravata, G. D. Huang *et al.*, “An economic analysis of robot-assisted therapy for long-term upper-limb impairment after stroke,” *Stroke*, vol. 42, no. 9, pp. 2630–2632, 2011.
- [13] Z. Bi, Y. Wang, H. Wang, Y. Zhou, C. Xie *et al.*, “Wearable EMG bridge—A multiple-gesture reconstruction system using electrical stimulation controlled by the volitional surface electromyogram of a healthy forearm,” *IEEE Access*, vol. 8, pp. 137330–137341, 2020.
- [14] T. Oyama, Y. Mitsukura, S. G. Karungaru, S. Tsuge and M. Fukumi, “Wrist EMG signals identification using neural network,” in *2009 35th Annual Conf. of IEEE Industrial Electronics*, Porto, Portugal, pp. 4286–4290, 2009. <https://doi.org/10.1109/IECON.2009.5415065>

- [15] C. Shen, Z. Pei, W. Chen, J. Wang, J. Zhang *et al.*, "Toward generalization of sEMG-based pattern recognition: A novel feature extraction for gesture recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022.
- [16] C. Yunfang, Z. Yitian, Z. Wei and L. Ping, "Survey of human posture recognition based on wearable device," in *2018 IEEE Int. Conf. on Electronics and Communication Engineering (ICECE)*, Xi'an, China, pp. 8–12, 2018. <https://doi.org/10.1109/ICECOME.2018.8644964>
- [17] E. Barkallah, J. Freulard, M. J. D. Otis, S. Ngomo, J. C. Ayena *et al.*, "Wearable devices for classification of inadequate posture at work using neural networks," *Sensors*, vol. 17, no. 9, pp. 24, 2017.
- [18] Q. Liu, X. Wu, Y. Jiang, X. Liu, Y. Zhang *et al.*, "A fully connected deep learning approach to upper limb gesture recognition in a secure FES rehabilitation environment," *Int. J. Intell. Syst.*, vol. 36, no. 5, pp. 2387–2411, 2021.
- [19] J. Ferreira, M. F. Gago, V. Fernandes, H. Silva, N. Sousa *et al.*, "Analysis of postural kinetics data using artificial neural networks in Alzheimer's disease," in *2014 IEEE Int. Symp. on Medical Measurements and Applications (MeMeA)*, Lisboa, Portugal, pp. 1–6, 2014.
- [20] A. Cristiano, A. Sanna and D. Trojaniello, "Daily physical activity classification using a head-mounted device," in *2019 IEEE Int. Conf. on Engineering, Technology and Innovation (ICE/ITMC)*, Valbonne Sophia-Antipolis, France, pp. 1–7, 2019. <https://doi.org/10.1109/ICE.2019.8792621>
- [21] T. T. Um, V. Babakeshizadeh and D. Kulić, "Exercise motion classification from large-scale wearable sensor data using convolutional neural networks," in *2017 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Vancouver, BC, Canada, pp. 2385–2390, 2017. <https://doi.org/10.1109/IROS.2017.8206051>
- [22] I. Bisio, C. Garibotto, F. Lavagetto and A. Sciarrone, "When eHealth meets IoT: A smart wireless system for post-stroke home rehabilitation," *IEEE Wireless Communications*, vol. 26, no. 6, pp. 24–29, 2019.
- [23] M. S. Karunarathne, S. A. Jones, S. W. Ekanayake and P. N. Pathirana, "Remote monitoring system enabling cloud technology upon smart phones and inertial sensors for human kinematics," in *2014 IEEE Fourth Int. Conf. on Big Data and Cloud Computing*, Sydney, NSW, Australia, pp. 137–142, 2014.
- [24] I. Chiuchisan, D. Balan, O. Geman, I. Chiuchisan and I. Gordin, "A security approach for health care information systems," in *2017 E-Health and Bioengineering Conf.*, Sinaia, Romania, pp. 721–724, 2017.
- [25] G. Yuan, X. Liu, Q. Yan, S. Qiao, Z. Wang *et al.*, "Hand gesture recognition using deep feature fusion network based on wearable sensors," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 539–547, 2021.
- [26] S. Bianco, P. Napoletano, A. Raimondi and M. Rima, "U-WeAr: User recognition on wearable devices through arm gesture," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 4, pp. 713–724, 2022.
- [27] P. Kang, J. Li, B. Fan, S. Jiang and P. B. Shull, "Wrist-worn hand gesture recognition while walking via transfer learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 3, pp. 952–961, 2022.
- [28] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo *et al.*, "ECA-net: Efficient channel attention for deep convolutional neural networks," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 11531–11539, 2020.
- [29] M. Panwar, D. Biswas, H. Bajaj, M. Jöbges, R. Turk *et al.*, "Rehab-net: Deep learning framework for arm movement classification using wearable sensors for stroke rehabilitation," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 11, pp. 3026–3037, 2019.
- [30] M. Ronald, A. Poulouse and D. S. Han, "iSPLInception: An inception-ResNet deep learning architecture for human activity recognition," *IEEE Access*, vol. 9, pp. 68985–69001, 2021.