Design Studies and Intelligence Engineering
L.C. Jain et al. (Eds.)

doi:10.3233/FAIA231428

Design of a Supernumerary Robotic Limb Based on Hybrid Control of Motion Imagination and Object Detection

Zhichuan TANG[a], Hang WANG[a], Zhixuan CUI[a], Jun DING[a,1], Lingtao ZHANG[a]
[a] *School of Design and Architecture, Zhejiang University of Technology, China.*

Abstract. Upper limb moto disorders are the main symptoms of stroke patients. Based on deep learning algorithms and object detection technology, we developed a brain-controlled supernumerary robotic limb system for upper-limb motion assistance. The system makes use of the motor imagery electroencephalogram (MI EEG) recognition model with graph convolutional network (GCN) and gated recurrent unit network (GRU) to obtain the patient's motion intentions and control the supernumerary robotic limb to move. The object detection technology can compensate for the disadvantages when using MI EEG alone like fewer control instructions and lower control efficiency. We also validated the feasibility and effectiveness of the system by designing model training experiment and target object grasping experiment. The results showed that the highest EEG classification accuracy using GCN+GRU algorithm achieved 92.32%, and the average success rate of grasping tasks achieved 88.67±3.77%.

Keywords. Stroke patients, Supernumerary robotic limb, Motor imagery, Object detection

1. Introduction

While the aging population continues to increase, the incidence rate of stroke and related diseases is also rising [1]. Patients with upper limb motor disorder caused by diseases such as amyotrophic lateral sclerosis (ALS), Parkinson's disease, progressive muscular atrophy (PMA), stroke, and spinal cord injury generally have limited mobility and daily activities [2]. These issues have already affected their daily activities, reduced the happiness of patients' lives, and brought enormous mental pressure to patients [3].

Supernumerary robotic limb, as wearable robot devices, can assist patients with motor disorder in various daily tasks [4]. Currently, the design and research of Supernumerary robotic limb remains an important challenge [5]. Most existing control methods use healthy limb to control the movement of the supernumerary robotic limb, which has problems such as high training time cost, high operational difficulty, and low overall coordination [6]. Therefore, many scholars have conducted in-depth application research on supernumerary robotic limb based on different physiological structures and application functions. Ciullo [7] propose a robotic supernumerary limb, the Soft-Hand X

[1] Corresponding author, School of Design and Architecture, Zhejiang University of Technology, 310023, China; E-mail: dingunun@163.com

(SHX) to re-enable hand use, and providing a degree of functionality and motivating against learned non-use. Charles[6] designed a robotic leg for assisting patients in walking with supernumerary limb assistance. Its more comfortable human-computer interaction mode makes it easier for patients to adapt to the inertial impact caused by the movement of supernumerary limb mechanical legs. Federico[8] designed a wearable supernumerary robotic limb that can provide support for the wearer with fixed objects around it, enabling it to complete corresponding tasks safely and stably.

Brain Computer Interface (BCI) as a system for human brain intention to interact with external devices [9]， it can be used to identify the limb motion intentions of patients with upper-limb motor disorder, quickly transmit control commands, and manipulate external devices to complete daily life tasks [10]. Patients with upper limb motor disorder can use the motor imagination paradigm to imagine their own limb movements and output EEG signals to control external devices [11]. This approach allows patients to interact more naturally with external devices, enhancing or replacing the damaged physical functions of disabled individuals[12]. However, due to the low signal-to-noise ratio of EEG signals during motor imagery （MI）, the decoding accuracy of EEG is low [13]. Some recent studies have used some feature selection related algorithms [14] , such as the Common Space Pattern (CSP) based on L1 Norm and Dempster Shafer theory [15], to find reliable features and improve the motor imagery classification performance [16]. Other studies have obtained deep features that can describe different MI classification through deep learning methods such as Deep Belief Networks (DBN)[17], and Long Short Term Memory (LSTM) networks[18]. Convolutional Neural Networks (CNN) [19] can be directly used for feature automatic extraction of raw input signals, and can obtain deeper and more differentiated feature information for EEG signal recognition [20,21]. However, traditional CNN methods do not consider the topological relationship and structural information of EEG electrodes, so they cannot grasp the topology relationship and structural information of EEG electrodes. Graph Convolutional Network (GCN) provides an effective way to describe the internal relationships between different nodes in a graph [22]. It is suitable for topological feature extraction of discrete spatial EEG signals, and combined with the temporal feature grasp by Gated Recurrent Unit network, it can fully extract the temporal and spatial feature information of EEG. However, due to the limitations of EEG decoding performance, it is still difficult to control external devices with multiple degrees of freedom to accurately reach and grasp the desired target in complex three-dimensional (3D) space through methods such as MI classification. Target detection as an auxiliary control method can solve the problem of insufficient control dimensions caused by brain computer interface control of supernumerary robotic limb.

Object detection technology can extract 2D or 3D information from images or videos, and use this information to recognize and locate target objects, thereby achieving auxiliary control for robots to grasp specific objects [23]. In this study, when the robotic arm enters an unfamiliar environment, the camera detects the measurement data of the target, obtains the corresponding environmental information and target features, guiding the robotic arm to carry out targeted movements, and helping the robotic arm determine and move its own position.

In this paper, we designed and developed a brain controlled supernumerary robotic limb system that meets the needs of patients with upper limb motor disorder. The system obtains the patient's motion intention through MI EEG signal recognition model based on graph convolutional neural network and gate loop unit network, achieving left and right motion control of the supernumerary robotic limb, and combines object detection technology to quickly grasp the target object, compensate for the limited control dimensions of MI control strategy while improving the control accuracy of the robotic arm.

2. External robot arm system

2.1. System framework

The supernumerary robotic limb system we have designed and developed is divided into software and hardware parts. Among them, the software part has two modules, including an MI recognition module and an object detection module. The MI recognition module is mainly used for obtaining EEG data and identifying motion intentions during patients engage in MI, then converting it into control signal and output. The object detection module is used to find, mark target objects, and calculate the grasping path. Combined with the control program, achieve the grasping and releasing function of the robotic limb. The hardware part includes a functional clothing, a robotic limb module, and a drive control module. The functional clothing is used to carry the robotic arm and related hardware, and the robotic arm module serves as a moving component to achieve grasping action. The workflow of the entire supernumerary robotic limb system is shown in Figure 1.

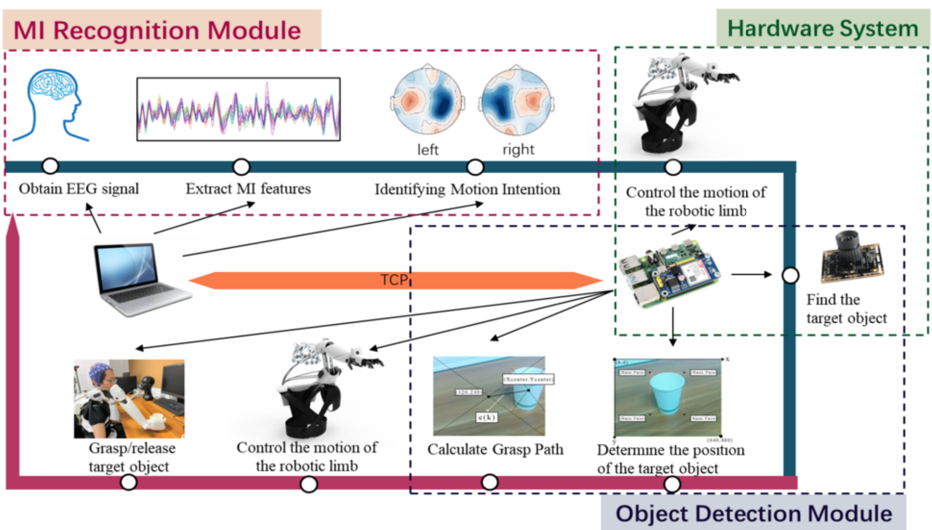

Figure 1. System framework diagram.

2.2. Hardware system

The hardware part of the supernumerary robotic limb we designed and developed is shown in Figure 2. The entire hardware system consists of a biomimetic robotic limb, a bionic hand, a camera, a functional clothing, a fixed base, a control backpack, and an EEG cap. The control backpack is embedded with hardware modules such as an electric motor drive board, a lithium battery, and a Raspberry Pi microcomputer. The entire robotic limb is fixed to the right shoulder of the functional clothing through a base, the control backpack is fixed to the back of the functional clothing, and the hardware part of the entire supernumerary robotic limb is fixed to the patient's upper torso through two elastic bands of the functional clothing. Both the robotic limb and the robotic finger control the servo motor drive through the drive board, and the camera is located at the end of the robotic limb to detect the target object in real-time, and cooperate with the robotic finger to achieve grasping action.

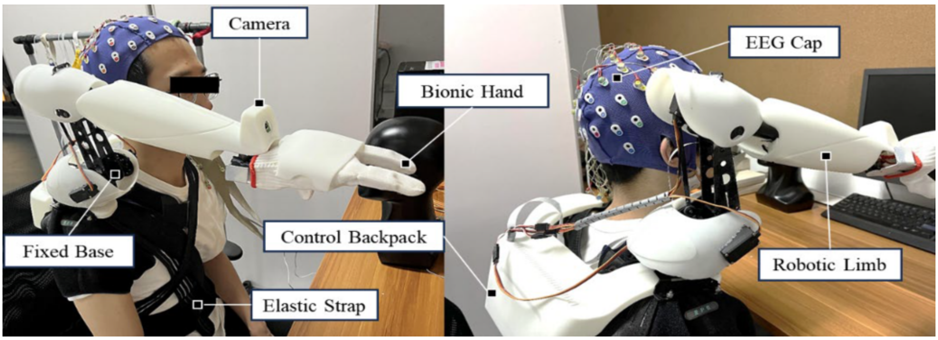

Figure 2. Hardware system.

2.3. Software System

The supernumerary robotic limb software we have developed and designed includes two parts: an MI recognition module and an object detection module, which are used to obtain and recognize the patient's motion intention and convert it into a robotic limb control signal. Combined with the object detection module, the supernumerary robotic limb grasping action is achieved.

2.3.1. MI recognition module

The MI recognition model we have developed and designed includes a feature extraction section and a classification section, specific framework as shown in Figure 3. Firstly, the collected EEG data is filtered across 11 frequency bands, and the EEG data from each frequency band is separately input into the MI recognition model. The data is divided into several time periods using the overlapping window method, and then reconstructed into graph data to extract spatial topology features into the GCN model. In the GCN model, each EEG channel corresponds to a node in the graph data, and the connection between two different nodes corresponds to the edges of the graph. Among them, the operation of graph convolution can be represented as

$$H^{(l+1)} = f(L^{sym} H^l W^l) \tag{2-1}$$

where H^l and $H^{(l+1)}$ are the l th graph convolutional layer and the *l*+1 th graph convolutional layer, f (·) is the ReLU activation function, W^l is the weight matrix of the lth graph convolutional layer, and L^{sym} is the symmetric normalized Laplace matrix.

The time series data learned from the GCN model is then input into the GRU model for temporal feature extraction. There are update gates in the GRU model $z_t$and reset door r_t. There are two types of gates, whose operations can be represented as

$$z_t = \sigma(W_{xz} x_t + W_{hz} h_{t-1} + b_z) \tag{2-2}$$

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_z) \tag{2-3}$$

Where σ represents the sigmoid function, W is the weight matrix and b is the deviation.

The Y value output by the GRU model will be sequentially transmitted to the fully connected layer, SoftMax layer, and classification output layer to generate category labels (left or right) for predicting MI motion intention. Finally, the best MI recognition model for the affected patient will be selected by comparing the best classification results of different EEG frequency bands.

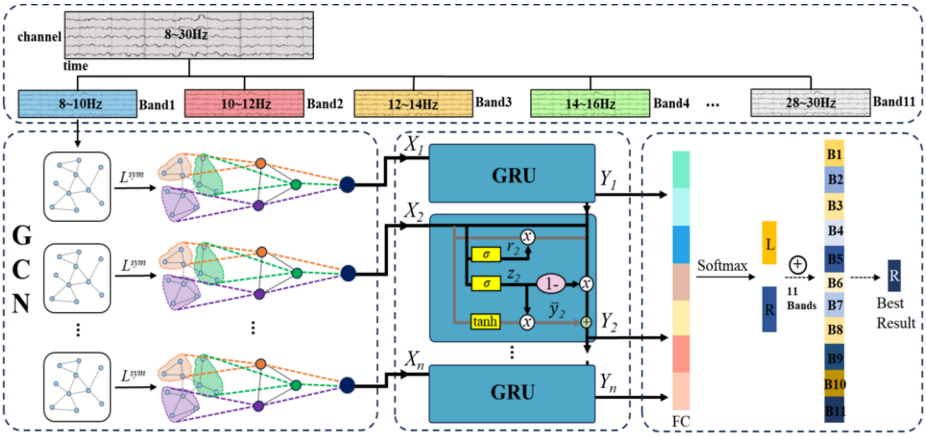

Figure 3. Framework diagram of MI recognition model.

2.3.2. Object detection and control module

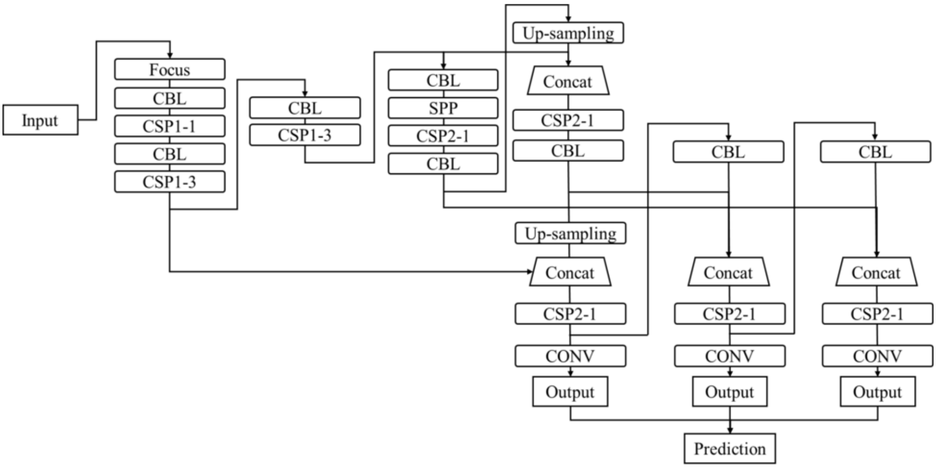

Figure 4. Framework diagram of the YOLO model.

The object detection module we have developed and designed is used to identify different target objects within the target area and provide position information of the object to further accurately control the motion of the robotic limb. Among them, we chose the more concise and faster YOLO [24] framework as the algorithm for the object detection module. The basic process includes two processes: pre-setting anchor boxes to locate the target and identifying the located target. The framework of the YOLO model is shown in Figure 4.

3. Experiment

3.1. Experimental Procedure

For our designed supernumerary robotic limb system, we conducted validation experiments on the MI recognition model and overall functionality, to evaluate the accuracy of the model and the grasping and collaboration capabilities of the supernumerary robotic limb. Firstly, we selected 10 subjects and trained them with MI models, obtaining MI models for each subject in different EEG frequency bands. Based on the model training results, we selected the model with the highest recognition accuracy in different frequency bands as the final model for this subject. Afterwards, the subject wears a 64 channel EEG cap and fixed the robotic limb to the right shoulder using two elastic bands. Each subject is required to conduct 25 grasping experiments on each target object, totaling 75 grasping experiments. They grasped different objects (paper cups, cloth bags, or plates), and the experimental scene is shown in Figure 5. The success rate of grasping is defined as the percentage of times the subject successfully grasp the target object.

Figure 5. Experimental Scenarios for Grasping Three Types of Objects.

3.2. Results of Experiment

In the MI model training experiment, we obtained MI recognition models of different subjects in different EEG frequency bands, and the model recognition accuracy is shown in Table 1. We found that the accuracy of the models trained by the same subject in different frequency bands was different. For example, subject 1 had the lowest MI classification accuracy (78.19%) in the 26-28Hz frequency band, and the highest MI classification accuracy (91.28%) in the 18-20Hz frequency band. The highest MI classification accuracy corresponding to each subject is shown in Figure 6. We selected MI recognition models from different subjects within the frequency band with the highest MI classification accuracy as the optimal model for each subject in the grasping experiment. To ensure the optimal results can be obtained in the grasping experiment. In the grasping experiment, the subject controlled the supernumerary robotic limb to perform grasping tests on three types of objects, achieving high target object recognition accuracy. The success rate of identifying and grasping three types of items by each subject is shown in Table 2. The results showed that the average success rate of grasping the cloth bag task was the highest, at 87.2 ± 3.60%. The average success rates of grasping the paper cup task and the plate task were 80.4 ± 3.67% and 81.6 ± 2.15%, respectively. Figure 7 shows the average success rate of each subject in three types of grasping tasks. The results showed that subject 10 had the highest average grasp success rate, which was 88.67 ± 3.77%, while subject 5 had the lowest average grasp success rate, which was 78.00 ± 3.27%.

Table 1. MI classification accuracy across EEG frequency bands for all subjects

Brands	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	Subject 9	Subject 10
8-10 Hz	0.8836	0.7847	0.9005	0.8108	0.8937	0.8747	0.8864	0.7765	0.7894	0.7861
10-12 Hz	0.7891	0.7409	0.9011	0.8306	0.8463	0.8629	0.8672	0.7895	0.8037	0.7971
12-14 Hz	0.8906	0.7761	0.9164	0.8557	0.871	0.8341	0.8903	0.8147	0.7686	0.9232
14-16 Hz	0.9122	0.7849	0.8896	0.8893	0.8977	0.8334	0.9011	0.8904	0.8064	0.9011
16-18 Hz	0.8902	0.8428	0.8548	0.8766	0.9026	0.8016	0.9047	0.8874	0.8791	0.8962
18-20 Hz	0.9128	0.8133	0.8247	0.8679	0.9032	0.7983	0.8945	0.8659	0.8447	0.8871
20-22 Hz	0.8439	0.8569	0.9017	0.7908	0.8907	0.7769	0.7839	0.8744	0.8806	0.7891
22-24 Hz	0.8017	0.8667	0.8948	0.7984	0.8528	0.6981	0.8896	0.8537	0.9148	0.8833
24-26 Hz	0.8022	0.8436	0.8961	0.8458	0.7837	0.7832	0.8738	0.8647	0.8904	0.7886
26-28 Hz	0.7819	0.8111	0.9018	0.8573	0.7739	0.7468	0.8457	0.8328	0.8746	0.775
28-30 Hz	0.7871	0.8005	0.8761	0.8366	0.7814	0.7233	0.8168	0.7906	0.8622	0.7351

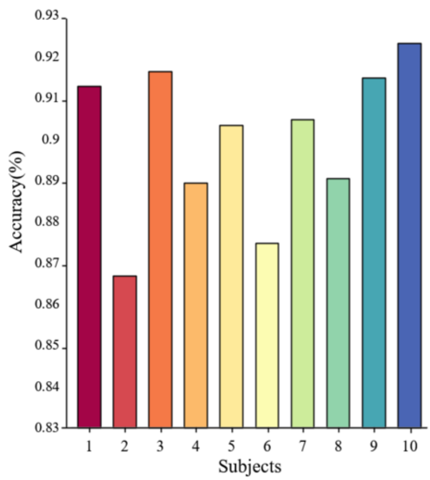

Figure 6. Maximum MI classification accuracy for each subject

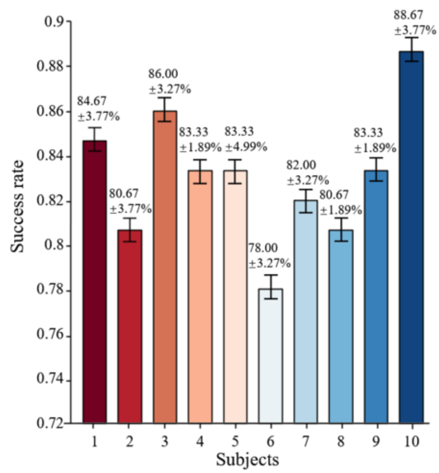

Figure 7. The average success rate of each subject in three types of crawling tasks.

Table 2 The success rate of each subject's three types of the grasping experiment

subject	paper cup	cloth bag	Ceramic plates
1	82.00%	90.00%	82.00%
2	78.00%	86.00%	78.00%
3	86.00%	90.00%	82.00%
4	82.00%	86.00%	82.00%
5	78.00%	90.00%	82.00%
6	74.00%	82.00%	78.00%
7	78.00%	86.00%	82.00%
8	78.00%	82.00%	82.00%
9	82.00%	86.00%	82.00%
10	86.00%	94.00%	86.00%
Average±std	80.4±3.67%	87.2±3.60%	81.6±2.15%

4. Conclusion and future work

This article designs a supernumerary robotic limb system based on BCI and object detection technology to assist patients with upper-limb motor disorder in completing daily grasping tasks. The MI recognition model and overall functional feasibility of the system are experimentally verified, providing a reference for solving the problems of high training cost and high operational difficulty of supernumerary robotic limb. Among them, the hybrid control strategy designed based on GCN+GRU network model and combined with object detection technology has made some preliminary theoretical contributions to improving the overall coordination and accurate recognition and transformation of motion intentions of Supernumerary robotic limb.

References

[1] I. Miyai, A. D. Blau, M. Reding, and B. T. Volpe, Patients with Stroke Confined to Basal Ganglia Have Diminished Response to Rehabilitation Efforts, Neurology **48** (1997), 1, 95–101.
[2] P. Raghavan, "Upper Limb Motor Impairment After Stroke, Physical Medicine and Rehabilitation Clinics of North America **26** (2015), 4, 599–610.
[3] R.-R. Lu et al., "Motor imagery-based brain-computer interface control of continuous passive motion for wrist extension recovery in chronic stroke patients," Neuroscience Letters **718** (2020), 134727.
[4] H. Xie, Z. Ding, S. Yoshida, T. Chong, T. Torii, and T. Fukusato, Augmenting Human with Compact Supernumerary Robotic Limbs, in 13th Augmented Human International Conference, Winnipeg MB Canada: ACM, May 2022, pp. 1–4.
[5] P. Maciejasz, J. Eschweiler, K. Gerlach-Hahn, A. Jansen-Troy, and S. Leonhardt, A survey on robotic devices for upper limb rehabilitation, J NeuroEngineering Rehabil **11** (2014), 1, 3.
[6] C. Khazoom, P. Caillouette, A. Girard, and J.-S. Plante, A Supernumerary Robotic Leg Powered by Magnetorheological Actuators to Assist Human Locomotion, IEEE Robot. Autom. Lett. **5** (2020), 4, 5143–5150.
[7] F. Parietti and H. Asada, Supernumerary Robotic Limbs for Human Body Support, IEEE Trans. Robot., 32 (2016), 2, 301–311.
[8] L. F. Nicolas-Alonso and J. Gomez-Gil, Brain Computer Interfaces, a Review, Sensors **12** (2012), 2, 1211–1279.
[9] A. Remsik et al., A review of the progression and future implications of brain-computer interface therapies for restoration of distal upper extremity motor function after stroke, Expert Review of Medical Devices, **13** (2016), 5, 445–454.
[10] M. A. Khan, R. Das, H. K. Iversen, and S. Puthusserypady, "Review on motor imagery based BCI systems for upper limb post-stroke neurorehabilitation: From designing to application," Computers in Biology and Medicine, vol. 123, p. 103843, Aug. 2020.
[11] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," J. Neural Eng. **4** (2007), 2, R1–R13.
[12] S. R. Soekadar et al., "Hybrid EEG/EOG-based brain/neural hand exoskeleton restores fully independent daily living activities after quadriplegia," Sci. Robot., **1** (2016), 1, 3296.
[13] J. Jin, R. Xiao, I. Daly, Y. Miao, X. Wang, and A. Cichocki, "Internal Feature Selection Method of CSP Based on L1-Norm and Dempster–Shafer Theory," IEEE Trans. Neural Netw. Learning Syst. 32 (2021), 11, 4814–4825.
[14] J. Jin, Y. Miao, I. Daly, C. Zuo, D. Hu, and A. Cichocki, Correlation-based channel selection and regularized feature optimization for MI-based BCI, Neural Networks 118 (2019), 262–270.
[15] B. Sun, X. Zhao, H. Zhang, R. Bai, and T. Li, EEG Motor Imagery Classification With Sparse Spectrotemporal Decomposition and Deep Learning, IEEE Trans. Automat. Sci. Eng., 18 (2021), 2, 541–551.
[16] Z. Tang, C. Li, and S. Sun, Single-trial EEG classification of motor imagery using deep convolutional neural networks, Optik **130** (2017), 11–18.
[17] Y. Chu, X. Zhao, Y. Zou, W. Xu, J. Han, and Y. Zhao, A Decoding Scheme for Incomplete Motor Imagery EEG With Deep Belief Network," Front. Neurosci., **12** (2018), 680.
[18] G. Dai, J. Zhou, J. Huang, and N. Wang, HS-CNN: a CNN with hybrid convolution scale for EEG motor imagery classification, J. Neural Eng. **17** (2020), 1, 016025.

[19] B. Xu et al., Wavelet Transform Time-Frequency Image and Convolutional Network-Based Motor Imagery EEG Classification, IEEE Access, **7** (2019), 6084–6093.
[20] Y. Hou et al., Deep Feature Mining via Attention-based BiLSTM-GCN for Human Motor Imagery Recognition. arXiv, Dec. 02, 2021. Accessed: Aug. 08, 2023. [Online].
[21] R. Ying, R. He, K. Chen, P. Eksombatchai, W. L. Hamilton, and J. Leskovec, Graph Convolutional Neural Networks for Web-Scale Recommender Systems, in Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London United Kingdom: ACM, Jul. 2018, pp. 974–983.
[22] J. Qi et al., An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease, Computers and Electronics in Agriculture, vol. 194, p. 106780, Mar. 2022.
[23] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement. arXiv: April. 08, 2018. Accessed: August 8, 2023. [Online].