Proceedings of the 2024 IEEE
International Conference on Robotics and Biomimetics
December 10-14, 2024, Bangkok, Thailand

2024 IEEE International Conference on Robotics and Biomimetics (ROBIO) | 979-8-3315-0964-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ROBIO64047.2024.10907308

A Hybrid EEG-EMG Framework for Humanoid Control using Deep Learning Transformers

Parthan Olikkal, *Student Member, IEEE,* Branesh M. Pillai, Jackrit Suthakorn, Habib Ali, Ramana Vinjamuri, *Senior Member, IEEE*

***Abstract*— Brain-computer interfaces (BCIs) offer promising solutions for upper limb rehabilitation. Despite advancements in deep learning, traditional models for motor rehabilitation using electroencephalography (EEG) or electromyography (EMG) to control assistive devices require enhancement. This study aims to enhance motor control capabilities by integrating EEG and EMG signals using a Transformer-based deep learning model. Ten able-bodies subjects performed center-out tasks on a low-cost upper limb rehabilitation table, capturing 2D kinematic data, EEG, and EMG signals simultaneously. The tasks varied in complexity across four levels. Preprocessed EEG and EMG signals were fused and given as input to the proposed model, which was evaluated using three performance metrics. Results showed that the EEG-EMG combined model achieved 87.27% accuracy across all the four levels. Furthermore, the model's output successfully controlled a humanoid robot to replicate similar movements. These findings highlight the efficacy of combined EEG-EMG data in improving accuracy and performance in BCI applications, advancing assistive technologies and neurorehabilitation interventions.**

***Keywords*—Brain-computer interface, deep learning, EEG-EMG, transformer, humanoid robot, upper limb motor rehabilitation**

I. Introduction

Stroke is a leading cause of disability and mortality among the elderly, second only to heart disease. Approximately 80% of stroke survivors experience motor impairment, typically affecting one side of the body [1], [2]. Despite these impairments, many retain the ability to generate motor-related neural activities, similar to healthy individuals, but only a small fraction regains useful upper limb functions after prolonged physiotherapy. Enhanced upper limb function is crucial for post-stroke rehabilitation as many activities of daily living (ADLs) rely heavily on arm functions. As highlighted in [3], the projected increase in the stroke population will impose a significant economic burden on the society, underscoring the need for effective rehabilitation strategies.

Recent advancements in machine learning and signal processing have enabled researchers to decode brain signals into actionable outputs, allowing control of devices like wheelchairs, assistive robots, and autonomous vehicles [4], [5]. This possibility for the brain to act upon an environment through an alternate pathway has drawn attention to the field of brain-computer interaction (BCI) [6], [7], as a promising tool in rehabilitation for patients with limited movement functionalities. One of the main challenges involved in BCIs is decoding upper limb kinematics from brain signals. Studies have shown that invasive methods using arrays of microelectrodes directly in the motor cortex can successfully perform reach and grasp activities in primates [8] and individuals with upper limb motor disability [9]. Despite their potential, invasive approaches require surgery limiting its usage. Non-invasive methods, such as electroencephalography (EEG), involve applying conductive gel on electrodes to enhance the conductivity between the scalp and electrodes. However, the low signal-to-noise ratio of EEG signals makes it difficult to decode hand movements. Another interactive bio-signal widely used is electromyography (EMG), where surface electrodes detect muscle activities during attempted movements, serving as a control signal for device interaction. EMG has been proven [10], [11] as a viable alternative to BCI for detection movements in individuals with motor impairments.

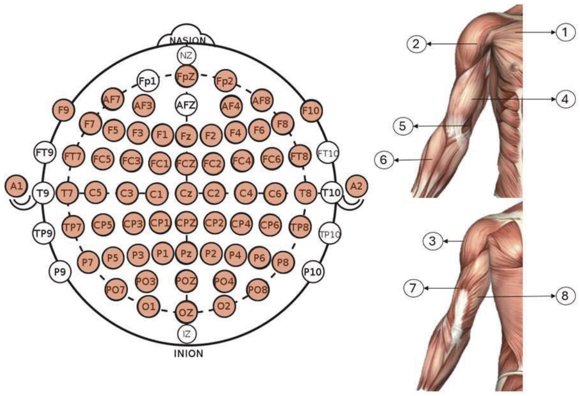

Fig. 1. EEG electrode locations on the scalp is shown on the left side and EMG electrode locations on the dominant arm is shown on the right side. All 64 EEG channels were located at the frontal, central, parietal and occipital areas to record the brain activities during center-out target experiment. 8 EMG sensors were placed strategically on – (1) PM, (2) Delt.A, (3) Delt.M, (4) Delt.L (5) B, (6) Brchl, (7) Tri. L, (8) Tri. Lat.

Research has demonstrated the classification of upper limb center-out reaching tasks using EEG [12] and reaching-to-grasp tasks using EMG [13] signals for prosthetic control. Most studies extract key features from EMG and EEG signals to train machine learning models, typically using

Research supported by National Science Foundation CAREER Award HCC-2053498 and NSF IUCRC BRAIN CNS-2333292.

P. Olikkal, H. Ali, R. Vinjamuri are with the Vinjamuri lab in the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD, 21220, USA (email: rvinjam1@umbc.edu). B. M. Pillai and J. Suthakorn are with Center for Biomedical and Robotics Technology, Mahidol University, Salaya Nakorm, Pathom, Thailand.

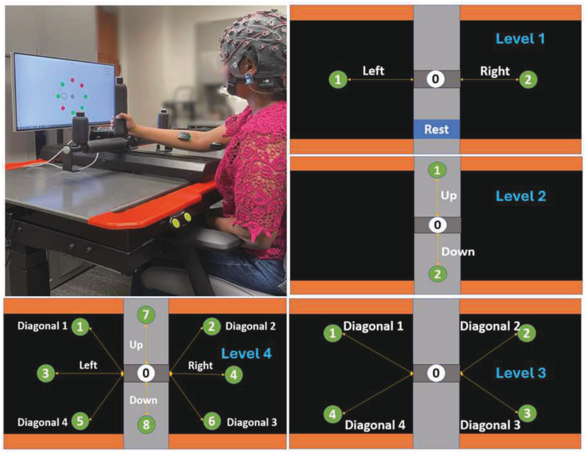

Fig. 2. Experiment setup. The subjects were asked to sit in front of the ArmAble^{TM} table. Their dominant hand rested on the unilateral handle with a grip hold to maneuver the device based on the target locations. Four different levels with 2, 2, 4 and 8 targets spread around the center were displayed on the screen. Target position was set randomly to increase engagement of the subjects.

either EEG or EMG alone. While EEG signals are complex to decode and do not directly measure movement, EMG signals provide a direct measure but are limited by factors like muscle fatigue. Although non-invasive BCI offers a promising alternative, few studies have investigated fusing scalp EEG [14] and surface EMG [15] for movement classification using signal decoding. Hybrid multimodal fusion frameworks [16]–[19] have been proposed, but often involve small sample size, limited EEG electrodes, and manual feature extraction followed by classical machine learning methods. Fusion techniques usually develop separate classifiers for EEG and EMG, combining their results through balanced weights or Bayesian approaches [16].

Deep learning algorithms have eliminated manual feature extraction, allowing preprocessed signals to directly classify movement tasks. By combining brain activity and muscle activity measurements, researchers can develop new technologies and therapies that assess changes in patients' brain and muscles during physiotherapy [20]. A practical framework combining EEG and EMG signals can help individuals with disabilities to perform ADL tasks with high precision through device or robot control. Although initial signal fusion show promise, the potential of deep learning for EEG and EMG fusion remains largely unexplored due to high computational complexity and data requirements for real-time movement recognition. Only few studies have thoroughly evaluated EEG and EMG efficacy for upper limb movement classification. While hybrid EEG-EMG-powered exoskeletons have shown considerable promise for gait movements [21], [22], the application of deep learning methods to hybrid EEG-EMG systems for upper limb movements still requires extensive exploration.

Stroke survivors often need to commute for physiotherapy sessions, and purchasing such devices for home use can place a significant financial burden on their families. Additionally, the lack of observable progress over time can diminish interest and motivation in continuing physiotherapy. Recent technology advancements have facilitated home-based rehabilitation, offering reduced travel burdens and increased flexibility. Such setup also allows patients to receive remote feedback from therapists. Chen et al. in [23] elaborates various home-based technologies for stroke rehabilitation. In account of this aspect, we employ a cost-effective, non-motor-driven interactive arm training device, designed for post-stroke recovery, namely ArmAble^{TM} (BeAble Health Pvt. Ltd., Telangana, India). This device utilizes interactive games to make upper limb rehabilitation more engaging and rewarding.

To address these concerns, this study evaluates the performance of EEG-EMG during center-out tasks using deep learning methods. Utilizing ArmAble^{TM}, EEG and EMG signals were collected from subjects performing center-out task scenarios of increasing complexity. In this preliminary study, a novel hybrid framework using both muscle and neural signals to classify movements was developed. It is hypothesized that supplementing neural signals with muscle signals will outperform methods using only neural signals. This novel framework aims to significantly improve classification accuracy. The results from this multimodal framework model were used to control a humanoid robot that moves the handle of the ArmAble^{TM} to the target locations based on the predictions.

II. Experimental Methods

A. Data Collection

Ten healthy subjects (4 female and 6 male subjects, age range: 20-33 yrs., mean age 27.4) with no history of upper extremity deformity or other musculoskeletal disorders were selected for the experiment. Subjects were informed about the experiment, and they provided written consent. The procedures were in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (IRB) ethical committee at the University of Maryland Baltimore County.

To capture muscle activities, eight wireless Avanti EMG sensors (Delsys, Natick, USA) were placed on the eight most representative muscles of the dominant hand: pectoralis major (PM), anterior deltoid (Delt.A), middle deltoid (Delt.M), lateral deltoid (Delt.L), biceps brachii (B), brachialis (Brchl), triceps brachii lateral head (Tri. Lat), triceps brachii long head (Tri. L) as shown in Fig. 1. These muscle locations were strategically selected due to their active engagement while performing the experiment. The skin was wiped and cleaned with alcohol swabs prior to placing the EMG electrodes and muscle signals were recorded at a sampling rate of 2000 Hz.

For neural activity, each subject wore an electrode cap to capture EEG signals recorded through an amplifier (g.HIamp, g.tec medical engineering GmbH, Graz, Austria) a sampling rate of 600 Hz. The cap contained 64 electrodes (shown in Fig. 1), with the left ear lobe as the reference and the forehead as ground.

Simultaneously, kinematic data were recorded while each subject controlled a unilateral handle of the ArmAble^{TM}, capturing the x-y coordinates of the cursor at a sampling rate of 50Hz, alongside the EMG and EEG (Fig.2)

TABLE I. EVALUATION METRICS OF EEG-EMG FUSED DATASET												
Subs	**Level 1**			**Level 2**			**Level 3**			**Level 4**		
	A	**F1**	**κ**	**A**	**F1**	**κ**	**A**	**F1**	**κ**	**A**	**F1**	**κ**
S1	0.833	0.832	0.667	0.883	0.883	0.767	0.789	0.787	0.72	0.833	0.829	0.805
S2	0.983	0.983	0.967	0.917	0.917	0.833	0.758	0.763	0.683	0.842	0.825	0.803
S3	0.967	0.967	0.933	0.917	0.916	0.833	0.841	0.84	0.788	0.813	0.796	0.767
S4	0.917	0.916	0.833	0.917	0.916	0.833	0.842	0.842	0.789	0.839	0.817	0.795
S5	0.817	0.814	0.633	0.9	0.899	0.8	0.867	0.868	0.823	0.84	0.838	0.815
S6	0.883	0.883	0.767	0.867	0.867	0.733	0.917	0.916	0.889	0.804	0.8	0.771
S7	0.933	0.933	0.867	0.933	0.933	0.867	0.833	0.831	0.776	0.852	0.846	0.823
S8	0.9	0.9	0.9	0.833	0.833	0.667	0.875	0.874	0.832	0.844	0.838	0.813
S9	0.967	0.967	0.933	0.933	0.933	0.867	0.892	0.866	0.822	0.804	0.796	0.768
S10	0.9	0.899	0.8	0.933	0.933	0.867	0.892	0.892	0.856	0.839	0.833	0.809
Mean	**0.91**	**0.909**	**0.83**	**0.903**	**0.903**	**0.807**	**0.851**	**0.848**	**0.798**	**0.827**	**0.822**	**0.797**

systems. Both EMG and EEG signals were recorded using MATLAB and SIMULINK, whereas the center-out-task experiment was designed using Unity and the kinematic data was recorded through in-built tracking methods.

B. Experimental Protocol

Subjects were seated comfortably in front of the ArmAble™ table, gripping the unilateral handle with their dominant hand in the rest position (Fig. 2). EEG and EMG

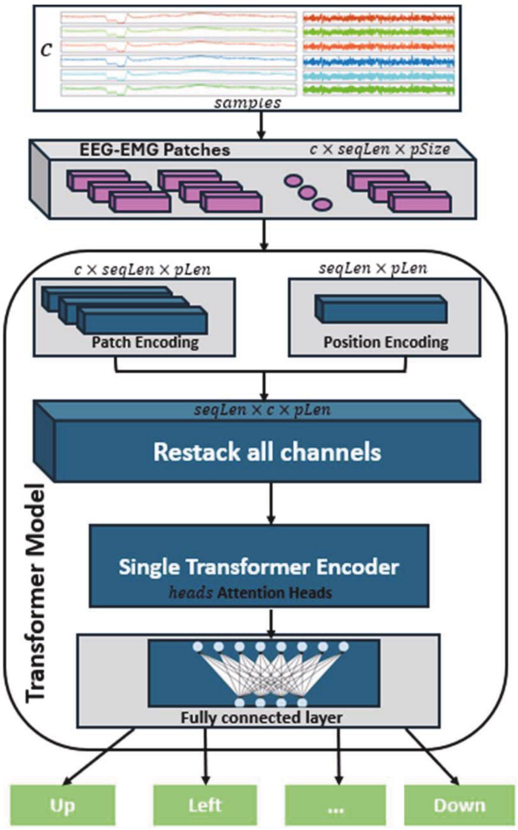

Fig. 3. Transformer Model Architecture. Input is EEG-EMG of $c \times samples$. Here c is the number of channels, $samples$ are the number of samples, $seqLen$ is the sequence length, $pSize$ is the length of each patch, $pLen$ is the patch vector length of encoded patch, $heads$ is the number of attention heads.

electrodes were carefully placed on the scalp and muscles. The experiment followed a center-out protocol using the ArmAble™ system with a computer screen. Subjects manipulated the handle to control a cursor moving from the start position to the center and then to multiple targets around the center. The experiment had four levels: the first and second levels had two targets each, the third level had four targets, and the fourth level had eight targets. Each level included 35 repetitions, with each repetition lasting 5-20 seconds, followed by a 5-second break. After completing each level, a 5-minute break was provided to prevent fatigue and maintain data quality. Targets were randomized to enhance engagement and attention.

C. Data Preprocessing

For each subject, four sets of EEG, EMG, and cursor coordinate data corresponding to the four levels were collected. EEG and EMG signals were recorded using MATLAB and SIMULINK, while cursor coordinates were tracked using built-in methods. A custom timestamp method synchronized the EEG and EMG data with cursor coordinates. ArmAble™ started recording kinematic data when the subject moved the handle from the rest position to the center target (refer Fig.2) and stopped when the subject completed the level. This study focused on the data from the center to the target. Therefore, the raw EEG and EMG data were meticulously trimmed based on the timestamps of each subject's movements.

For EEG preprocessing, raw EEG data were passed through a 5th order high-pass filter with a cutoff frequency of 0.5 Hz, and the resulting 64-channel signals were cleaned using the ATAR algorithm [24] with soft-thresholding (β = 0.5) as it has shown its promise to subdue artifacts using wavelet decomposition. For EMG preprocessing, raw data were passed through a 4th order butter worth bandpass filter (25-500Hz) and a notch filter at 60Hz to eliminate power line interference. The root mean square (RMS) of the filtered signals was calculated and further subjected to Gaussian smoothing.

During the experiment, errors were observed when subjects reached incorrect targets, resulting in timeouts or quick target switches. To keep the consistency of the dataset for all the subjects, 30 out of 35 repetitions per level were selected, excluding those instances of incorrect target reach. To create the EEG-EMG fusion, preprocessed EEG signals were vertically concatenated with preprocessed EMG signals. EMG signals were down-sampled to match the length of EEG signals, resulting in a fused dataset with 64-channel EEG and 8-channel EMG data contributing to 72 channels for 30 repetitions per level. To prepare the data for analysis, z-score standardization was applied as expressed in Eq. (1).

$$\hat{X}_{t,c} = \frac{X_{t,c} - \bar{X}_c}{\sigma_c} \tag{1}$$

Here t, c represents the number of samples and channels, $\bar{X}_c$ represents the sample's mean on channel c, and σ_c represents the sample's standard deviation on channel c.

D. Proposed Model Architecture

To address the complexity of raw EEG and EMG signals, we employed advanced deep learning techniques. Our model architecture is inspired by the original Transformer [25], adapted to handle multi-channel EEG-EMG data. This adaptation leverages attention mechanisms to transform the data into highly distinguishing representations, capturing spatiotemporal pattens across EEG and EMG channels.

The core architecture comprises three key components: Patch Encoding, Positing Encoding, and a Single Transformer Encoder (Fig. 3). The Transformer Encoder uses multiple $heads$ corresponding to the number of EEG-EMG channels. The output from the encoder is then fed into a fully connected layer for final predictions.

P**atch Encoding**: Similar to word embedding in natural language processing (NLP), this component encodes each path of the fusion data to capture underlying patterns. EEG-EMG signals, initially structured as $channels \times samples$, were transformed into patches of size $patchSize$, resulting in a matrix of $channels \times seqLen \times patchSize$. This methodical division captures local patterns. The patches are then encoded using a linear layer, producing patch vectors of length $patchLen$, preserving intricate temporal dynamics across channels.

Position Encoding: To incorporate temporal information, position encoding is employed, inspired by conventional Transformer techniques. This process reshapes the data to $channels \times seqLen \times patchLen$ and adds temporal context, reorganizing the data into a 2D representation ($seqLen \times channels \times patchLen$). This matrix is provided as input to the Transformer encoder, enhancing the model's comprehension of temporal dynamics in the EEG-EMG data.

Single Transformer Encoder: At the core of our model, the Transformer encoder processes multi-channel EEG-EMG data. Each channel's information is processed independently, allowing a distinct understanding of diverse neural activities. Our encoder utilizes a single block, unlike the original Transformer's six blocks. The Transformer encoder employs $heads$ corresponding to the number of EEG-EMG channels, enabling simultaneous attention to various aspects of the input data. This sophisticated attention mechanism significantly enhances the model's ability to anticipate complex spatiotemporal patterns within the EEG-EMG data.

E. Scoring Performance

From the confusion matrix obtained, three indicators were employed to conduct a comprehensive evaluation for each level: accuracy, f1-score, and cohen-kappa score. These indicators provide a detailed analysis of the model's performance. The evaluation metrics are defined as follows:

Accuracy (A): Calculated as the mean accuracy of the sum of 10-fold cross-validation for each individual subject expressed in Eq. (2) as

$$Accuracy_{subject} = \sum \frac{(TP_i + TN_i)}{(TP_i + TN_i + FP_i + FN_i)} \tag{2}$$

Where TP_i is True Positive, $TN_i$is True Negative, FP_i is False Positive and FN_i is False Negative of the i^{th}fold.

F1-Score (F1): The harmonic mean of precision and recall, providing a balanced measure of the model's performance. It is expressed in Eq. (3) as a mean of the sum across 10-fold cross-validation as

$$F1_{subject} = \frac{2 \times Precision \times Recall}{(Precision + Recall)} \tag{3}$$

Cohen-Kappa Score (κ): Measures the inter-rater agreement for the true labels and predicted labels. Based on the measured score, the level of agreement can be a credible indicator to the model's performance.

The results obtained were then used to control a humanoid robot, Mitra (Invento Research Inc. Plano, TX, USA) [26], [27] with 21 degrees of freedom (Fig. 4).

III. Results

In this study, 64-channel EEG and 8-channel EMG data were collected from ten subjects performing center-our tasks on an upper limb rehabilitation device across four levels of increasing complexity, with each subject completing 30 trials per level. After preprocessing, a hybrid method, involving the vertical concatenation of down-sampled EMG signals with EEG data, was used to evaluate the performance of the fusion model. Three performance metrics were employed, and 10-fold cross-validation was conducted on the normalized dataset to ensure unbiased results. Table I represents the performance results for the classification of 2, 2, 4, 8 targets using EEG-EMG. The results indicates that the proposed model classifies the input fused EEG-EMG data with high accuracy.

As observed in Table I, the model demonstrates high binary classification accuracy, f1-score and kappa scores in Levels 1 and 2 across most subjects. The mean accuracy for Level 1 is 91% and for Level 2, it is approximately 90%. The f1-scores and kappa scores follow similar high trends. This indicates that the model performs exceptionally well in simpler tasks where the targets are fewer and easier to reach. As the task complexity increases in Levels 3 and 4, a slight decline in performance metrics is observed. For Level 3, the mean accuracy drops to 85%, and for Level 4, it further reduces to around 83%. f1-scores and cohen-kappa scores also follows a similar trend, indicating increased difficulty in accurately predicting the target positions as the number of targets and task complexity increases.

IV. Discussion

This study presents a novel deep learning Transformer model that uses EEG-EMG signals to classify center-out tasks performed on a low-cost upper limb rehabilitation table. Initially, the study was conducted on healthy subjects to ensure device safety, functionality, and model predictive capabilities before exploring its potential application in rehabilitation.

Although various fusion techniques have been proposed in other studies [14]–[18], [28], this research uniquely integrates EMG and EEG vertically and evaluates the classification performance using a light weight Transformer model with only one encoder compared to the traditional six encoder. Comprehensive reviews on various deep learning models for BCI using EEG data are detailed in [29]–[31]. Compared to models applied on similar datasets in [32], such as convolutional neural networks (CNN), long-short tern memory networks (LSTM), and EEG-Net, our proposed Transformer model shows superior performance. While CNN and EEG-Net achieved 93% and 70% accuracy for classifying 2 and 6 targets respectively, our model attained 98.3% accuracy for binary classification, 91.7% for 4-class classification, and 85.2% for 8-class classification – the highest among subjects. For 6-class classification, EEG-Net obtained 36% accuracy, whereas for 4-class, it achieved 61%. In comparison, our model achieved 85% and 83% for 4-class and 8-class, respectively, across subjects. In another study on decoding hand motor imagery tasks using EEG, the proposed CNN EEG-Net model [33] achieved 78.46% accuracy for 2-class classification and 76.72% for 3-class classification. Our Transformer-based model outperformed these results, achieving 91% accuracy for 2-class and 85.1% for 4-class classification. Aggregating performance metrics using EEG-EMG across four levels, our model achieved 87.27% accuracy, 87.05% f1-score, and 80.58% kappa score.

The model weights with the highest accuracy from the cross-validation were saved during the training phase. Three random trials from each of the four levels were selected for each subject. These saved weights were loaded into the Transformer model for each level, and the model classified these trials for each subject. The results from the classification were then used to control the left hand of Mitra to the corresponding position as shown in Fig 4.

These results indicates that our proposed model not only surpasses the accuracy of existing models like CNN, LSTM, and EEG-Net for binary and multi-class classifications but also demonstrates a marked improvement in handling more complex classification tasks. The high performance across different classification level suggest that the integration of EEG and EMG signals and the use of a Transformer model enhance the robustness and reliability of the classification process, making it a promising approach for BCI technology for robot assistive rehabilitation settings.

V. Conclusion

This study proposes a novel transformer-based deep learning architecture for hand movements classification using EEG-EMG data. Ten subjects participated in data collection phase where they performed center-out tasks on a low-cost rehabilitation table. EEG and EMG signals were collected, preprocessed, and fused to create a combined dataset. The performance of the proposed model was evaluated using fused EEG-EMG data for binary (accuracy: 91% and 90.3% for level 1 and 2) 4-class (accuracy: 85.1% for level 3) and 8-class (accuracy: 82.7% for level 4) classification tasks. These results achieved superior accuracy compared to existing research employing similar datasets. Building on these promising findings, future work

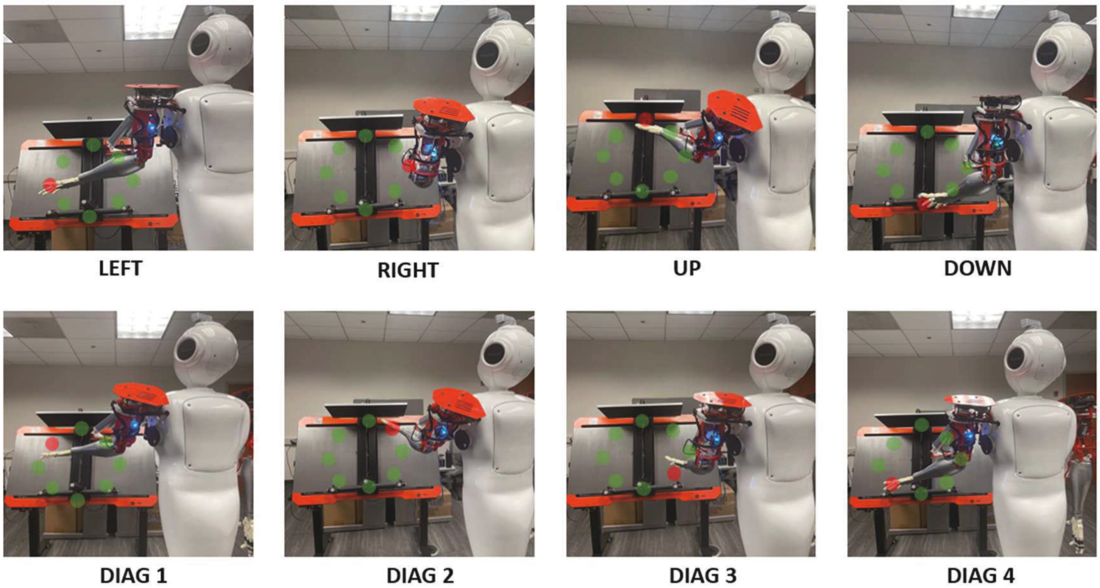

Fig.4. Three random trials were selected from EEG-EMG dataset and evaluated using the model at different levels. The outcomes from these evaluations were employed to guide the humanoid robot Mitra in reaching designated target locations. This illustration focuses on the scenario of Level 4, which involves eight target locations. The red circles denote the positions of the targets and utilizing EEG-EMG data as inputs, Mitra successfully navigated to each of these locations.

will include data collected from individuals undergoing upper limb rehabilitation. This framework has a potential to help individuals control robotic or assistive devices using EEG-EMG signals. The study also demonstrates the capability of a lightweight deep learning model to efficiently utilize both EMG and EEG data for device control. Further stages of this research will focus on refining the proposed model for real-time application and employing other fusion methods.

REFERENCES

[1] P. Langhorne, F. Coupar, and A. Pollock, "Motor recovery after stroke: a systematic review," *Lancet Neurol.*, vol. 8, no. 8, pp. 741–754, Aug. 2009, doi: 10.1016/S1474-4422(09)70150-4.

[2] V. L. Feigin *et al.*, "Global, regional, and national burden of stroke and its risk factors, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019," *Lancet Neurol.*, vol. 20, no. 10, pp. 1–26, 2021, doi: 10.1016/S1474-4422(21)00252-0.

[3] S. Anwer *et al.*, "Rehabilitation of Upper Limb Motor Impairment in Stroke: A Narrative Review on the Prevalence, Risk Factors, and Economic Statistics of Stroke and State of the Art Therapies," *Healthcare*, vol. 10, no. 2, Feb. 2022, doi: 10.3390/HEALTHCARE10020190.

[4] A. Schwarz, M. K. Höller, J. Pereira, P. Ofner, P. Ofner, and G. R. Müller-Putz, "Decoding hand movements from human EEG to control a robotic arm in a simulation environment," *J. Neural Eng.*, vol. 17, no. 3, 2020, doi: 10.1088/1741-2552/ab882e.

[5] A. Cruz, G. Pires, A. Lopes, C. Carona, and U. J. Nunes, "A Self-Paced BCI with a Collaborative Controller for Highly Reliable Wheelchair Driving: Experimental Tests with Physically Disabled Individuals," *IEEE Trans. Human-Machine Syst.*, vol. 51, no. 2, pp. 109–119, 2021, doi: 10.1109/THMS.2020.3047597.

[6] M. A. L. Nicolelis, "Actions from thoughts," *Nature*, vol. 409, no. 6818, pp. 403–407, 2001, [Online]. Available: http://www.nature.com/articles/35053191

[7] J. D. R. Millán *et al.*, "Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges," *Front. Neurosci.*, vol. 4, no. SEP, pp. 1–15, 2010, doi: 10.3389/fnins.2010.00161.

[8] J. M. Carmena *et al.*, "Learning to Control a Brain–Machine Interface for Reaching and Grasping by Primates," *PLoS Biol.*, vol. 1, no. 2, p. e42, Oct. 2003, doi: 10.1371/journal.pbio.0000042.

[9] L. R. Hochberg *et al.*, "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, 2012, doi: 10.1038/nature11076.

[10] S. Balasubramanian, E. Garcia-Cossio, N. Birbaumer, E. Burdet, and A. Ramos-Murguialday, "Is EMG a Viable Alternative to BCI for Detecting Movement Intention in Severe Stroke?," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 12, pp. 2790–2797, Dec. 2018, doi: 10.1109/TBME.2018.2817688.

[11] A. B. Barreto, S. D. Scargle, and M. Adjouadi, "A practical EMG-based human-computer interface for users with motor disabilities," *J. Rehabil. Res. Dev.*, vol. 37, no. 1, pp. 53–64, 2000.

[12] A. Úbeda, J. M. Azorín, R. Chavarriaga, and J. del R Millán, "Classification of upper limb center-out reaching tasks by means of EEG-based continuous decoding techniques," *J. Neuroeng. Rehabil.*, vol. 14, no. 1, pp. 1–14, Feb. 2017, doi: 10.1186/S12984-017-0219-0.

[13] K. T. Kim, S. Park, T. H. Lim, and S. J. Lee, "Upper-Limb Electromyogram Classification of Reaching-to-Grasping Tasks Based on Convolutional Neural Networks for Control of a Prosthetic Hand," *Front. Neurosci.*, vol. 15, p. 733359, Oct. 2021, doi: 10.3389/FNINS.2021.733359/BIBTEX.

[14] J. Tryon, E. Friedman, and A. L. Trejos, "Performance evaluation of EEG/EMG fusion methods for motion classification," *IEEE Int. Conf. Rehabil. Robot.*, vol. 2019-June, pp. 971–976, 2019, doi: 10.1109/ICORR.2019.8779465.

[15] V. De Seta *et al.*, "Towards a hybrid EEG-EMG feature for the classification of upper limb movements: Comparison of different processing pipelines," *Int. IEEE/EMBS Conf. Neural Eng. NER*, vol. 2021-May, pp. 355–358, 2021, doi: 10.1109/NER49283.2021.9441390.

[16] A. Manolova, G. Tsenov, V. Lazarova, and N. Neshov, "Combined EEG and EMG fatigue measurement framework with application to hybrid brain-computer interface," *2016 IEEE Int. Black Sea Conf. Commun. Networking, BlackSeaCom 2016*, Apr. 2017, doi: 10.1109/BLACKSEACOM.2016.7901569.

[17] A. Sarasola-Sanz *et al.*, "A hybrid brain-machine interface based on EEG and EMG activity for the motor rehabilitation of stroke patients," *IEEE Int. Conf. Rehabil. Robot.*, pp. 895–900, Aug. 2017, doi: 10.1109/ICORR.2017.8009362.

[18] R. Leeb, H. Sagha, R. Chavarriaga, and J. D. R. Millán, "A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities," *J. Neural Eng.*, vol. 8, no. 2, Apr. 2011, doi: 10.1088/1741-2560/8/2/025011.

[19] P. Olikkal, D. Pei, T. Adali, N. Banerjee, and R. Vinjamuri, "Data Fusion-Based Musculoskeletal Synergies in the Grasping Hand," *Sensors (Basel).*, vol. 22, no. 19, Oct. 2022, doi: 10.3390/S22197417.

[20] T. D. Lalitharatne, K. Teramoto, Y. Hayashi, and K. Kiguchi, "Towards Hybrid EEG-EMG-Based Control Approaches to be Used in Bio-robotics Applications: Current Status, Challenges and Future Directions," *Paladyn*, vol. 4, no. 2, pp. 147–154, Dec. 2013, doi: 10.2478/PJBR-2013-0009/MACHINEREADABLECITATION/RIS.

[21] S. Tortora, L. Tonin, C. Chisari, S. Micera, E. Menegatti, and F. Artoni, "Hybrid Human-Machine Interface for Gait Decoding Through Bayesian Fusion of EEG and EMG Classifiers," *Front. Neurorobot.*, vol. 14, no. November, 2020, doi: 10.3389/fnbot.2020.582728.

[22] Z. Li *et al.*, "Hybrid Brain/Muscle Signals Powered Wearable Walking Exoskeleton Enhancing Motor Ability in Climbing Stairs Activity," *IEEE Trans. Med. Robot. Bionics*, vol. 1, no. 4, pp. 218–227, 2019, doi: 10.1109/TMRB.2019.2949865.

[23] Y. Chen, K. T. Abel, J. T. Janecek, Y. Chen, K. Zheng, and S. C. Cramer, "Home-based Technologies for Stroke Rehabilitation: A Systematic Review," *Int. J. Med. Inform.*, vol. 123, p. 11, Mar. 2019, doi: 10.1016/J.IJMEDINF.2018.12.001.

[24] N. Bajaj, J. Requena Carrión, F. Bellotti, R. Berta, and A. De Gloria, "Automatic and tunable algorithm for EEG artifact removal using wavelet decomposition with applications in predictive modeling during auditory tasks," *Biomed. Signal Process. Control*, vol. 55, p. 101624, 2020, doi: 10.1016/j.bspc.2019.101624.

[25] A. Vaswani, "Attention Is All You Need," no. Nips, 2017.

[26] P. Olikkal, D. Pei, B. K. Karri, A. Satyanarayana, N. M. Kakoty, and R. Vinjamuri, "Learning Hand Gestures using Synergies in a Humanoid Robot," *2023 IEEE Int. Conf. Robot. Biomimetics, ROBIO 2023*, pp. 1–6, 2023, doi: 10.1109/ROBIO58561.2023.10354698.

[27] P. Olikkal, D. Pei, B. K. Karri, A. Satyanarayana, N. M. Kakoty, and R. Vinjamuri, "Biomimetic learning of hand gestures in a humanoid robot," no. July, pp. 1–14, 2024, doi: 10.3389/fnhum.2024.1391531.

[28] P. Olikkal, D. Pei, T. Adali, N. Banerjee, and R. Vinjamuri, "Musculoskeletal Synergies in the Grasping Hand," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2022-July, pp. 3649–3652, 2022, doi: 10.1109/EMBC48229.2022.9871023.

[29] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, 2019, doi: 10.1088/1741-2552/ab0ab5.

[30] K. M. Hossain, M. A. Islam, S. Hossain, A. Nijholt, and M. A. R. Ahad, "Status of deep learning for EEG-based brain–computer interface applications," *Front. Comput. Neurosci.*, vol. 16, no. i, 2023, doi: 10.3389/fncom.2022.1006763.

[31] F. Safavi *et al.*, "Emerging Frontiers in Human–Robot Interaction," *J. Intell. Robot. Syst. Theory Appl.*, vol. 110, no. 2, 2024, doi: 10.1007/s10846-024-02074-7.

[32] M. M. Abdelhamid *et al.*, "Recognizing Hand Movements Using EEG-Signal Classification," *Proc. - 11th IEEE Int. Conf. Intell. Comput. Inf. Syst. ICICIS 2023*, pp. 463–469, 2023, doi: 10.1109/ICICIS58388.2023.10391145.

[33] D. Achanccaray and M. Hayashibe, "Decoding Hand Motor Imagery Tasks within the Same Limb from EEG Signals Using Deep Learning," *IEEE Trans. Med. Robot. Bionics*, vol. 2, no. 4, pp. 692–699, 2020, doi: 10.1109/TMRB.2020.3025364.