Skelly: A Mind Booster Game

Manal Mohsen Tantawy
Scientific Computing Depatment
(Associate professor)
Ain Shams University
Cairo,Egypt
manalmt@cis.asu.edu.eg

Aya Nasser
Scientific Computing Depatment
(Teacher Assistant)
Ain Shams University
Cairo,Egypt
aya.naser@cis.asu.edu.eg

Ezz Khalil
Scientific Computing Department
(Student)
Ain Shams University
Cairo,Egypt
azz20191700386@cis.asu.edu.eg

Ahmed Awad
Scientific Computing Depatment
(Student)
Ain Shams University
Cairo,Egypt
ahmed20191700053@cis.asu.edu.eg

Mohammed Khalefa
Scientific Computing Depatment
(Student)
Ain Shams University
Cairo,Egypt
20201701161@cis.asu.edu.eg

Abdelrahman Emam
Scientific Computing Department
(Student)
Ain Shams University
Cairo,Egypt
20201700432@cis.asu.edu.eg

Mohammed Abd El-raouf
Scientific Computing Depatment
(Student)
Ain Shams University
Cairo,Egypt
mohammed20191700557@cis.asu.edu.eg

Caroline Atta
Scientific Computing Depatment
(Student)
Ain Shams University
Cairo,Egypt
caroline20191700445@cis.asu.edu.eg

***Abstract*—this study presents "Skelly," an innovative brain-controlled game designed to help individuals with disabilities engage in entertaining and rehabilitative activities. By using Brain-Computer Interface (BCI) technology, players can navigate through various maze levels using their brain signals (EEG), providing a unique gaming experience and therapeutic benefits. The BCI Competition IV dataset 2a, comprising motor imagery tasks from nine subjects, was utilized to train our deep learning model, Attention temporal convolution network (ATCNet), which demonstrated a high average accuracy of 87.16% in decoding EEG signals into control commands. "Skelly" illustrates the potential of integrating BCI technology into gaming, offering an inclusive platform for users to interact with digital environments through their thoughts, enhancing both their gaming experience and rehabilitation process.**

I. Introduction

The advent of Brain-Computer Interface (BCI) technology has opened new possibilities for enhancing the quality of life for individuals with physical disabilities and brain injuries. Traditional video games and interactive digital content often rely on controllers and keyboards, which are not always accessible for everyone. This limitation can prevent people with motor impairments from enjoying and participating in engaging digital experiences. "Skelly," a brain-controlled game, addresses these challenges by enabling users to navigate maze levels using their brain signals (EEG). This approach offers an inclusive entertainment platform and serves as a novel rehabilitation tool. By utilizing BCI technology, "Skelly" transforms the way individuals with disabilities interact with digital environments, allowing them to control the game character directly with their thoughts. This not only enhances their gaming experience but also provides therapeutic benefits. The motivation behind this study is rooted in the global prevalence of disabilities. Approximately 1.3 billion people, or 16% of the world's population, live with disabilities, including brain diseases such as Amyotrophic Lateral Sclerosis (ALS). ALS, affecting both middle-aged and younger individuals, results in severe motor function impairment, underscoring the need for accessible and engaging rehabilitation tools

"Skelly" leverages motor imagery (MI) EEG signals, which detect a user's intention from their brain activity when imagining movements. This capability allows users to control devices, communicate, and restore mobility through thought alone. By enhancing the accuracy, variability, and usability of existing MI EEG systems, "Skelly" aims to significantly improve the lives of those with disabilities or brain injuries. To develop "Skelly," we utilized the BCI Competition IV dataset 2a, containing data from nine subjects performing motor imagery tasks (left hand, right hand, feet, and tongue movements). This structured dataset was crucial for training our deep learning model, ATCNet, which achieved an impressive average accuracy of 87.16% in interpreting EEG signals for game navigation. The field of brain science, an interdisciplinary domain integrating neuroscience, psychology, biology, and computer science, has significantly advanced since the mid-20th century. Innovations in imaging techniques and computational modeling have deepened our understanding of neural connectivity, plasticity, and cognitive functions. BCI technology, a product of this interdisciplinary research, directly measures brain activity to create control signals, bypassing traditional neuromuscular pathways. "Skelly" demonstrates the potential of BCI technology in gaming and rehabilitation, offering a stimulating and accessible way for users to interact with digital environments using only their minds. This study showcases the profound impact that integrating brain science and technology can have on enhancing human experiences and therapeutic outcomes.

The upcoming sections of this paper will cover the following topics: the related works (Section 2), providing a review of existing research and technologies relevant to our project; the system architecture (Section 3), detailing the methodology, including data preprocessing, model training, and the overall framework of "Skelly"; and the results (Section 4), presenting the performance metrics and findings from our experiments.

II. Related Works

Feng Li et al. [1] utilized BCI Competition IV Dataset 2b to validate their CWT-SCNN algorithm for classifying motor imagery (MI) EEG signals. The dataset included nine subjects with recordings from electrodes C3, Cz, and C4 at 250 Hz across five sessions per subject. They focused on sessions 04E and 05E, each containing 160 trials. Their algorithm applied continuous wavelet transform (CWT) to transform raw MI-EEG signals (filtered from 4–35 Hz) into time-frequency images, emphasizing mu and beta rhythms. The SCNN architecture comprised six layers: input, two convolutional layers with batch normalization and ReLU activation, a flatten layer, a fully connected layer, and an output layer. The C2 and C3 layers used eight and 16 filters, respectively, for frequency and time-domain features. Achieving 83.2% average accuracy and a kappa value of 0.651 on Dataset 2b, their approach outperformed competitors by 11.9%, demonstrating improved classification with reduced training time suitable for real-time BCI applications.

Tariq Sadiq et al [2]. introduced the Multivariate Empirical Wavelet Transform (MEWT) algorithm for robust motor imagery (MI) task decoding. Their approach enhances noise robustness with multiscale principal component analysis and automates channel selection for MI task decoding. They innovate by introducing a sub-band alignment method using MEWT for joint amplitude and frequency component extraction in MI applications. They also employ correlation-based feature selection to reduce computational complexity effectively. Experiments on BCI Competition III datasets show significant improvements: achieving up to 98% sensitivity, specificity, and classification accuracy in subject-specific cases, and 93%, 92.1%, and 91.4% in subject-independent cases, with enhancements of 23.50% and 18.14%, respectively, over previous methods. The algorithm demonstrates perfect classification accuracy (100%) for subjects with limited training data and in subject-independent cases using a single source subject, highlighting its practical viability for MI EEG signal classification.

Hassanpour et al [3]. proposed Gradient Deep Learning (GDL) for motor imagery (MI) EEG signals, emphasizing end-to-end learning to minimize preprocessing steps. Their multilevel GDL-based scheme employs Deep Belief Network (DBN) and Stacked Sparse Autoencoder (SSAE) models, outperforming FBCSP with 78.5074% accuracy and 0.6278 kappa. The Multilevel Binary Classifier (MBC) integrates preprocessing, feature extraction, and classification into a single block, omitting CSP for data augmentation, band-pass filtering, and frequency conversion. MBC's binary tree structure and similarity metrics optimize classification by partitioning multi-class problems effectively, validated with an M × M similarity matrix for enhanced MI-EEG signal classification.

Xu et al [4]. focused on enhancing motor imagery task recognition for BCI applications by proposing a deep multi-view feature learning method for EEG signals. They addressed the challenge of overlooking critical motor imagery features in deep learning models. Their approach involved multi-view feature representation, integrating distinct EEG signal characteristics and different feature types (time domain, frequency domain, time-frequency domain, spatial features). They utilized a deep restricted Boltzmann machine (RBM) network with t-distributed stochastic neighbor embedding (t-SNE) to learn these features, reducing redundancy and enhancing feature recognizability. Classification was performed using a support vector machine (SVM). Tested on the BCI Competition IV 2a dataset, their method significantly improved motor imagery task accuracy, underscoring its potential to advance BCI applications.

From the review of these four studies in motor imagery EEG signal classification for BCI applications introduced innovative methodologies to enhance accuracy. Feng Li et al. and TARIQ SADIQ et al. focused on CWT-SCNN and MEWT algorithms, respectively, emphasizing robust feature extraction and noise resilience. Hassanpour et al. explored GDL with DBN and SSAE models, while Xu et al. developed a multi-view feature learning approach. These studies demonstrated high accuracies, up to 98%, leveraging deep learning for complex signal patterns. They prioritized robustness against noise and computational efficiency, validated on benchmark datasets like BCI Competition IV, highlighting advancements in practical BCI implementations.

III. System Architecture

The proposed ATCNet model consists of three main blocks: convolutional (CV) block, attention (AT) block, and temporal convolutional (TC) block, as shown in Figure 1. CV block encodes low-level spatio-temporal information within the MI-EEG signal through three convolutional layers: temporal, channel depth-wise, and spatial convolutions. The output of the CV block is a temporal sequence with a higher-level representation. The AT block then highlights the most important information in the temporal sequence using a multihead self-attention (MSA). Finally, the TC block extracts highlevel temporal features within the temporal sequence using TCN and feeds them into a fully connected (FC) layer with a SoftMax classifier.

A. Preprocessing and Input Representation

We feed raw MI-EEG signals into the proposed model without preprocessing, that is, the full frequency band, all channels, and without artifact removal. ATCNet model takes as input a motor imagery trial consisting of C channels (EEG electrodes) and T time points. The objective of the ATCNet model is to map the input MI trial X to its corresponding class

y. For the BCI-2a dataset, T = 1125 time points, C = 22 EEG channels, n = 4 MI classes, and m = 5184 MI trials.

B. Convolutional (CV) block

The CV block is similar to the EEGNet architecture proposed in [5]. CV block differs from EEGNet by using 2D convolution instead of separable convolution, which showed better performance. CV block also uses different parameter values than those used in [5].

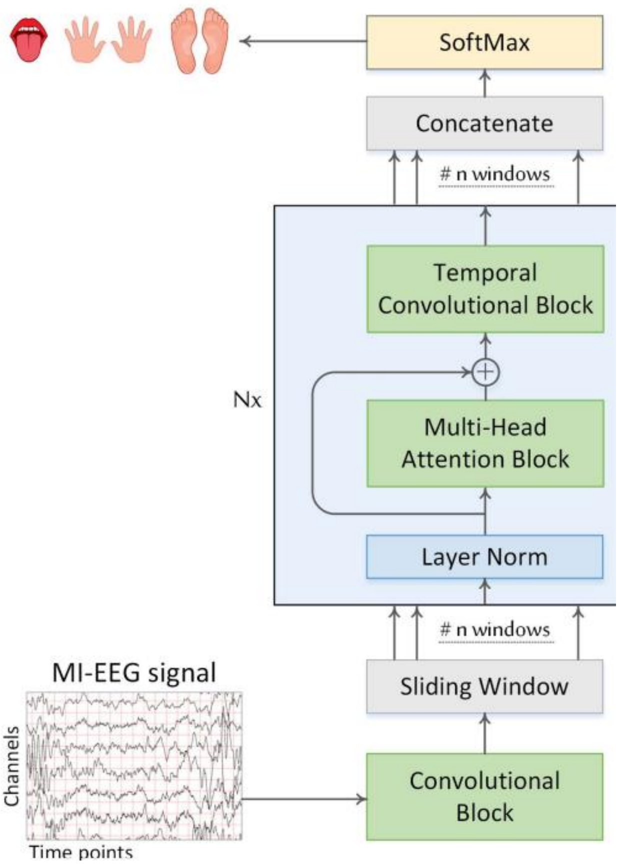

Figure 1. The components of the proposed ATCNet model.

CV block consists of three convolutional (conv) layers, as shown in Figure 2. The first layer performs a temporal convolution using F filters of size (1, K), where K is the filter length in the time axis. K was set to be one-fourth of the sampling rate (64 for BCI-2a). This allows the filters to extract temporal information associated with frequencies above 4 Hz. The output of this layer is F temporal feature maps. The second layer is a depth-wise convolution with F filters of size (C, 1), where C is the number of EEG channels. Using depth-wise convolution, each filter extracts spatial features (i.e., related to EEG channels) from a single temporal feature map. Therefore, the output of this layer is $F \times D$ feature maps, where D is the number of filters linked to each temporal feature map in the previous layer. D is set empirically to 2. $F \times D$ determine the output dimension of the CV block. The depth-wise convolution is followed by an average pooling layer of size (1, 8) to abstract the temporal data by a factor of 8. This reduces the sampling rate of the signal to ~32Hz. The third convolutional layer consists of F filters of size (1,) . K was set to 16 to decode MI activities within 500 ms (for 32 Hz sampled data). Finally, a second average pooling layer with a size of (1, P) is used to reduce the sampling rate to ~32/P Hz. P is used to control the length of the temporal sequence produced by CV block. The second and third conv layers are followed by batch normalization to speed up network training and then by exponential linear unit (ELU) activation for nonlinearity

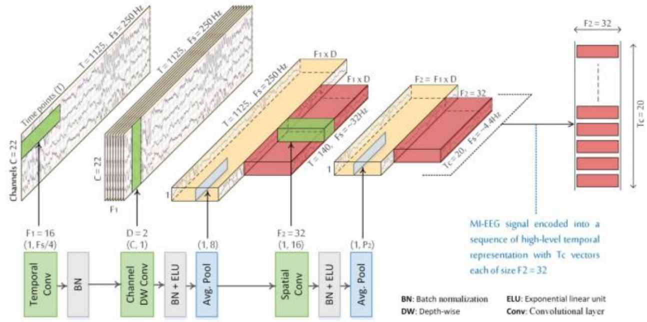

Figure 2. CV block performs spatio-temporal encoding through three convolutional layers.

C. Convolutional-based sliding window (SW)

Instead of entering the whole T samples of z to the later layers, a sliding window has been used to divide the temporal sequences into multiple windows. This helps to augment the data and enhance the decoding accuracy. However, the sliding window raises the computations, because it requires the input data to be passed through the DL model n times (instead of once), where n stands for the number of windows. As a result, the computations are incremented n times. But in our approach, we used a sliding window as integration with convolutional layers (in the convolutional block). In this approach, convolution computations are performed once for all windows, which reduces training and inference time by parallelizing the process. This technique was originally used in sliding-window based object detection.

D. Attention (AT) Block

In psychology, the cognitive process of selectively focusing on one or a few things while disregarding others is known as attention. In deep neural networks, the attention mechanism is an effort to emulate the human brain behavior of selectively focusing on a few significant elements while ignoring others. In the visual world, subjects use both volitional and nonvolitional cues to selectively focus attention. The former is taskdependent, and the latter is based on the conspicuity and saliency of things in the surroundings. Inspired by the voluntary and involuntary attention cues, the attention mechanism can be emulated using three components: values (sensory inputs), keys (nonvolitional cues), and queries (volitional cues). The interaction of queries and keys creates attention pooling that biases the selection of values, as demonstrated in Figure 3. The attention mechanism can be implemented based on attention scores or by different machine learning algorithms such as reinforcement learning. This research adopts an attention scores-based approach, i.e., MSA, due to its large success in various fields such as NLP and computer vision. The attention block consists of an MSA layer,MSA consists of several self-attention layers (i.e., scaled dot-product attention) called heads, as shown in Figure 3.

Each self-attention layer consists of three main components: query Q, keys K, and values V. Interactions between query and keys produce attention scores that guide selection bias over values.

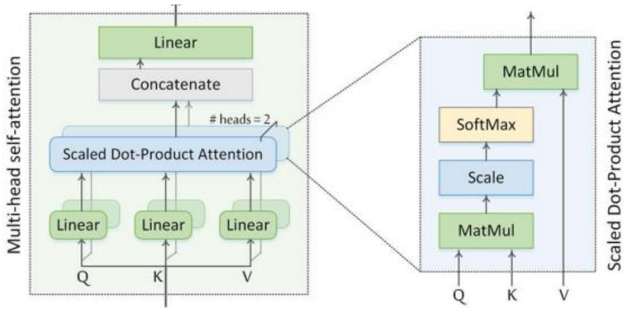

Figure 3. Multi-head self-attention.

E. *Temporal Convolutional (TC) block*

The TC block has the same architecture as the TCN. TCN consists of a stack of residual blocks. The residual block composes of two dilated causal convolutional layers, each one followed by batch normalization and ELU activation, as shown in Figure 4

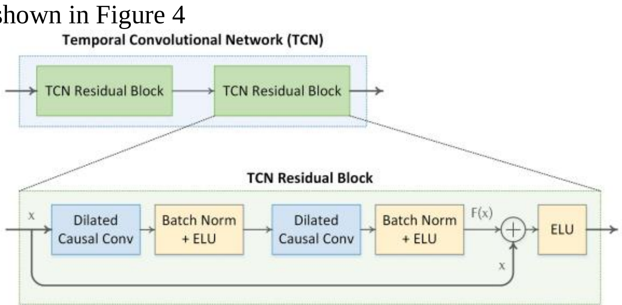

Figure 4. The architecture of the temporal convolutional network (TCN) consisting of two residual blocks

IV. Results

BCI Competition IV-2a (BCI-2a) dataset is used to train and evaluate the proposed model. BCI-2a is a well-known public MI-EEG dataset created by Graz University of Technology in 2008. BCI-2a has been widely used in the research community and is thus considered a benchmark dataset in MI-EEG decoding. It contains a limited number of samples captured under uncontrolled conditions with a considerable amount of artifacts, which makes decoding MI tasks using this dataset a challenging process.

BCI-2a dataset consists of 5184 trials (samples) of MI-EEG data recorded using 22 EEG electrodes from 9 subjects (576rials per subject). MI trials last 4 seconds and were sampled at 250 Hz and filtered between 0.5 and 100 Hz. Each trial belongs to one of four MI tasks: imagining of movement of the left hand (class 1), right hand (class 2), both feet (class 3), and tongue (class 4). For each subject, two sessions were recorded on different days. Each session consists of 288 trials per subject. One of these sessions is used to train the model and the other for evaluation.

The proposed model is evaluated using subject-dependent (subject-specific) and subject-independent approaches. For subject-dependent, we used the same training and testing data as the original competition, i.e., 288 x 9 trials in session 1 for training, and 288 x 9 trials in session 2 for testing and its results in Table 1.

Table 1. summarizes the accuracy of the proposed ATCNet model using the BCI-2a dataset

Subject	Accuracy
Subject 1	**84.48%**
Subject 2	**77.59%**
Subject 3	**89.66%**
Subject 4	**81.03%**
Subject 5	**87.93%**
Subject 6	**77.59%**
Subject 7	**96.55%**
Subject 8	**94.83%**
Subject 9	**94.83%**
Average Acc.	**87.16%**

V. Conclusion

"Skelly" represents a significant advancement in the application of Brain-Computer Interface (BCI) technology for gaming and rehabilitation. By leveraging EEG signals, this study provides an inclusive entertainment platform and a novel therapeutic tool for individuals with physical disabilities or brain injuries.

The high classification accuracy achieved by the ATCNet model, trained on the BCI Competition IV dataset 2a, underscores the effectiveness of this approach. With an average accuracy of 87.16%, the model reliably translates users' motor imagery into game commands, enabling them to navigate through complex mazes using their thoughts alone.

This study not only highlights the potential of BCI technology to enhance the gaming experience but also its practical applications in rehabilitation. By engaging users in mentally stimulating tasks, "Skelly" supports cognitive and motor skill development, offering a valuable resource for therapeutic interventions.

The success of "Skelly" demonstrates the feasibility of integrating BCI technology into interactive digital environments, paving the way for future innovations that further bridge the gap between human intention and machine action.

VI. References

[1] Li, F., He, F., Wang, F., Zhang, D., Xia, Y., & Li, X. (2020). A novel simplified convolutional neural network classification algorithm of motor imagery EEG signals based on deep learning. *Applied Sciences*, *10*(5), 1605.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Sadiq, M. T., Yu, X., Yuan, Z., Zeming, F., Rehman, A. U., Ullah, I., ... & Xiao, G. (2019). Motor imagery EEG signals decoding by multivariate empirical wavelet transform-based framework for robust brain–computer interfaces. *IEEE access*, *7*, 171431-171451.

[3] Xu, J., Zheng, H., Wang, J., Li, D., & Fang, X. (2020). Recognition of EEG signal motor imagery intention based on deep multi-view feature learning. *Sensors*, *20*(12), 3496..

[4] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[5] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces,” J. Neural Eng., vol. 15, no. 5, p. 56013, 2018

[6] D. Li, J. Xu, J. Wang, X. Fang, and J. Ying, “A Multi-Scale Fusion Convolutional Neural Network based on Attention Mechanism for the Visualization Analysis of EEG Signals Decoding,” IEEE Trans. Neural Syst. Rehabil. Eng., 2020.

[7] G. A. Altuwaijri, G. Muhammad, H. Altaheri, and M. Alsulaiman, “A Multi-Branch Convolutional Neural Network with Squeeze-and-Excitation Attention Blocks for EEG-Based Motor Imagery Signals Classification,” Diagnostics, vol. 12, no. 4, p. 995, 2022

[8] S. U. Amin, H. Altaheri, G. Muhammad, M. Alsulaiman, and A. Wadood, “Attention-Inception and Long Short-Term Memory-based Electroencephalography Classification for Motor Imagery Tasks in Rehabilitation,” IEEE Trans. Ind. Informatics, 2022.

[9] D. Zhang, K. Chen, D. Jian, and L. Yao, “Motor imagery classification via temporal attention cues of graph embedded EEG signals,” IEEE J. Biomed. Heal. informatics, vol. 24, no. 9, pp. 2570–2579, 2020.

[10] M. Qamhan, H. Altaheri, A. H. Meftah, G. Muhammad, and Y. A. Alotaibi, “Digital Audio Forensics: Microphone and Environment Classification Using Deep Learning,” IEEE Access, vol. 9, pp. 62719–62733, 2021.