Decoding of Imagined Speech Neural EEG Signals Using Deep Reinforcement Learning Technique

Nrushingh Charan Mahapatra
Intel Technology India Pvt Ltd
Bengaluru 560103, India
School of Computer Engineering
Kalinga Institute of Industrial Technology
Bhubaneswar 751024,India
https://orcid.org/0000-0001-9166-5451

Prachet Bhuyan
School of Computer Engineering
Kalinga Institute of Industrial Technology
Bhubaneswar 751024,India.

2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC) | 978-1-6654-6109-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ASSIC55218.2022.10088387

***Abstract*—The basic objective of the study is to establish the reinforcement learning technique in the decoding of imagined speech neural signals. The purpose of imagined speech neural computational studies is to give people who are unable to communicate due to physical or neurological limitations of speech generation alternative natural communication pathways. The advanced human-computer interface based on imagined speech decoding based on measurable neural activity could enable natural interactions and significantly improve quality of life, especially for people with few communication alternatives. Recent advances in signal processing and reinforcement learning based on deep learning algorithms have enabled high-quality imagined speech decoding from noninvasively recorded neural activity. Most of the prior research focused on the supervised classification of collected signals, with no naturalistic feedback-based training of imagined speech models for brain-computer interfaces. We employ deep reinforcement learning in this study to create an imagined speech decoder artificial agent based on the deep Q-network (DQN), so that the artificial agent could indeed learn effective policies directly from multidimensional neural electroencephalography (EEG) signal inputs adopting end-to-end reinforcement learning. We show that the artificial agent, supplied only with neural signals and rewards as inputs, was able to decode the imagined speech neural signals efficiently with 81.6947% overall accuracy.**

Keywords*—brain-computer interface (BCI), imagined speech, neural signal processing, electroencephalography (EEG), deep reinforcement learning (DRL)

I. Introduction

Researchers have been interested in decoding neural electrical signals for decades in order to allow people with speech disorders to communicate properly with their surroundings. Advanced human-computer interaction, also known as brain-computer interface (BCI) or brain-machine interface (BMI), is a system that converts neural signals into machine-interpretable outputs [1]–[3]. As a result, practical BCI systems are the most viable way to respond to the challenges. The noninvasive EEG signals are frequently employed in the development of the BCI system involving information extraction from the brain [4]. Imagined speech is the internal articulation of a syllable or word prompt without any phonological motion. A reasonable vision is the BCI framework, which can decode imagined speech such as syllables and words directly from neural signals.

In the past, studies have looked at the decoding of imagined speech using EEG signals, with participants imagining certain speech stimulus syllables, words, or combinations of words, as well as syllables at specified time intervals within the imagined speech. Matsumoto and Hori [5] studied the classification of the imagined speech signals of vowels using common spatial patterns (CSP) signal preprocessing and nonlinear classification using support vector machine (SVM). Wang et al. [6] used the EEG signal feature vector extraction method CSP and a SVM classifier to demonstrate imagined speech classification of Chinese words. Idrees et al. [7] applied time domain features such as standard deviation and waveform length, and a linear classifier was used to classify pairwise imagined speech vowels. Coretto et al. [8] classified imagined speech combinations of five vowels and six words using the signal processing techniques of discrete wavelet transform (DWT) and a machine learning classifier, random forest (RF). Qureshi et al. [9] demonstrated imagined speech recognition of five words utilizing the feature vector extraction of EEG signals with the approach of coherence and covariance, and the data was classified using the extreme learning machine (ELM). Cooney et al. [10] exploited transfer learning models based on convolutional neural networks (CNN) to classify the vowels of imagined speech EEG signals. Panachakel et al. [11] used the DWT approach to extract the features from the EEG imagined speech signals and multiclass classification with a deep neural network to classify syllables and words as speech prompts. Tamm et al. [12] used the CNN model to classify imagined speech signals of vowels and words without the removal of artifacts. Zhang et al. [13] studied the application of CSP to extract feature vectors from imagined speech EEG signals of four Mandarin tones and SVM to classify these neural signals. Sarmiento et al. [14] experimented with the imagined speech signals of vowel decoding, which were preprocessed with filters, transformed with spectral analysis, and classified using the CNN model. Lee et al. [15] classified imagined speech EEG signals from two public datasets of six words each using deep metric learning feature extraction and a siamese neural network classifier. Bakhshali et al. [16] used inter-regional connection features extraction and SVM classification to show imagined speech decoding of EEG signals of four vowels. Asghari Bejestani et al. [17] demonstrated imagined speech decoding of six Persian words using EEG signal processing, fast fourier transform, and binary SVM classification.

In earlier studies, we discovered that researchers employed various signal processing techniques, statistical learning, machine learning, and deep learning frameworks for imagined speech decoding or classification. However, these frameworks lack the use of deep reinforcement learning (DRL). Deep learning and reinforcement learning are integrated into DRL frameworks. Reinforcement learning (RL) is a systematic strategy for an artificial agent to learn through interacting with its environment in the same way as biological agents learn. The artificial agent should be able to maximize the overall

978-1-6654-8695-8/22/$31.00 ©2022 IEEE

objectives in the form of cumulative rewards using information obtained from the environment.

In this study, we introduce the DRL framework for the deep Q-learning network based on a convolutional neural network approach for imagined speech classification. The proposed model is equipped with an ongoing trial-and-error capability such that artificial agents construct efficient representations from multi-dimensional neural signal inputs with rewards specific to the agent action interacting with the brain-computer interface environment and exploit them to generalize prior experience to the new observed state in the environment. The model's learning capabilities allow the artificial agent to continuously improve its objectives by maximizing the neural network's imagined speech signal classification accuracy.

II. MATRIALS AND METHODS

A. Neural EEG Signals

The noninvasive imagined speech neural signal collection for brain speech activity is obtained directly from the brain by the placement of EEG electrodes on the scalp in accordance with the ten-twenty international standard. The studies employed the imagined speech signals of five vowels from the open-access dataset recorded by Coretto et al. [8] and Table I outlines the features of the imagined speech neural signal database.

TABLE I. A BRIEF DESCRIPTION OF THE IMAGINED SPEECH NEURAL SIGNALS DATABASE

Attributes	Information
Signal Type	EEG
Channels	C3, F3, P3, C4, F4, P4
Sampling Rate	1 KHz
No. of Subjects	15
Speech Prompts	/a/, /e/, /i/, /o/, and /u/

B. Neural Signal Processing

The raw EEG signals were filtered with a band-pass filter with a minimum and maximum frequency range of 0.001–100 Hz to eliminate low-frequency patterns and high-frequency noises while preserving all frequency bands for feature extraction and analysis. The EEG signals were down-sampled at 512 Hz.

In imagined speech neural activity, EEG signals have the high temporal resolution and a low signal-to-noise ratio (SNR). In scalp EEG, independent component analysis (ICA) is often used to isolate neural signals from artifacts. ICA, a sophisticated approach for blind source separation, is use to eliminate various distortions or artifacts such as muscle activity or eye blinking [18]. The FastICA approach [19] was used to remove the artifacts by identifying the source EEG signals from noise and other background information in the signals in an efficient way.

C. Mathematical Representation

The fundamental form of imagined speech classification is described as the learning process of constructing the nonlinear function f: S → A parameterized with weights *W,* the equation is defined as in (1) that receives the input imagined speech EEG signal s ∈ S and output a decoded speech class a ∈ A, where signal s ∈ $R^{max\ channels \times sampling\ points}$, S is the imagined speech signal space and $A \in \{"/a/", "/e/", "/i/", "/o/", "/u/"\}$ action space is the speech class labels.

$$a = f(s; W) \tag{1}$$

D. Deep Reinforcement Learning-DQN Framework

In the model-free value-based DQN technique, the neural network model is employed as a Q-value function approximator, with the state as an input and the state-action value as an output. The DQN's Q-learning function is Q (s, a; W), where s represents the model input state, a represents the model output action, and W represents the model weights. The DQN architecture is comprised of three fundamental components, such as a Q-network deep learning model, a non-trainable target network model, and an Experience Replay component. The Q-network is a trainable agent trained over time steps that optimizes state-action value. The learning process in the DQN is converged by two components, such as the experience reply and the target network model [20]. Fig. 1 shows the DQN-based DRL frameworks for imagined speech-based BCI system.

In DQN, the CNN deep learning model was utilized to approximate the Q-value function. The observed state is fed into the CNN model, and the model output is the Q-value of all feasible actions. The input layer, number of one-dimensional convolutional layers, fully connected, and output layer encompass the Q-network CNN feedforward model architecture. In the convolutional layer and fully connected layer, the activation function used is ReLU. Softmax is used in the output layer's activation function. Fig. 2 shows the Q-network CNN model. The DRL system environment consists of EEG signal acquisitions of imagined speech with labels, as well as components for EEG signal processing. Instead of performing the action, the environment calculates the reward using the label signal sample and the agent's action imagined speech classification. The DQN Agent (imagined speech classifier), a value-based DRL agent, interacts with the BCI environment to train the CNN model to approximate the Q-value pair based on observation state (imagined speech signal) and likely action spaces (speech prompts), resulting in future reward values. During the DQN model training process, an epsilon-greedy exploration strategy would be used by the agent to explore the action space. The main notions of DRL systems, such as observation state, action, and delayed reward, are as follows.

- *State (s):* The observation state is represented by a preprocessed neural imagined speech EEG signal sample with regard to a certain speech prompt.
- *Action (a)*: The DQN agent operates in distinct action spaces. The action is the decoded imagined speech class, and the action set is the five speech vowel classes. The observed state is fed into the Q-network model, and the model output is the Q-value of all feasible actions.
- *Reward (r):* The purpose is to appropriately classify the imagined speech signals. The environment rewards the action with a reward of 1 when the artificial agent DQN performs the activity that corresponds to the relevant imagined speech class. The activity related to the different imagined speech classes is performed by the artificial agent

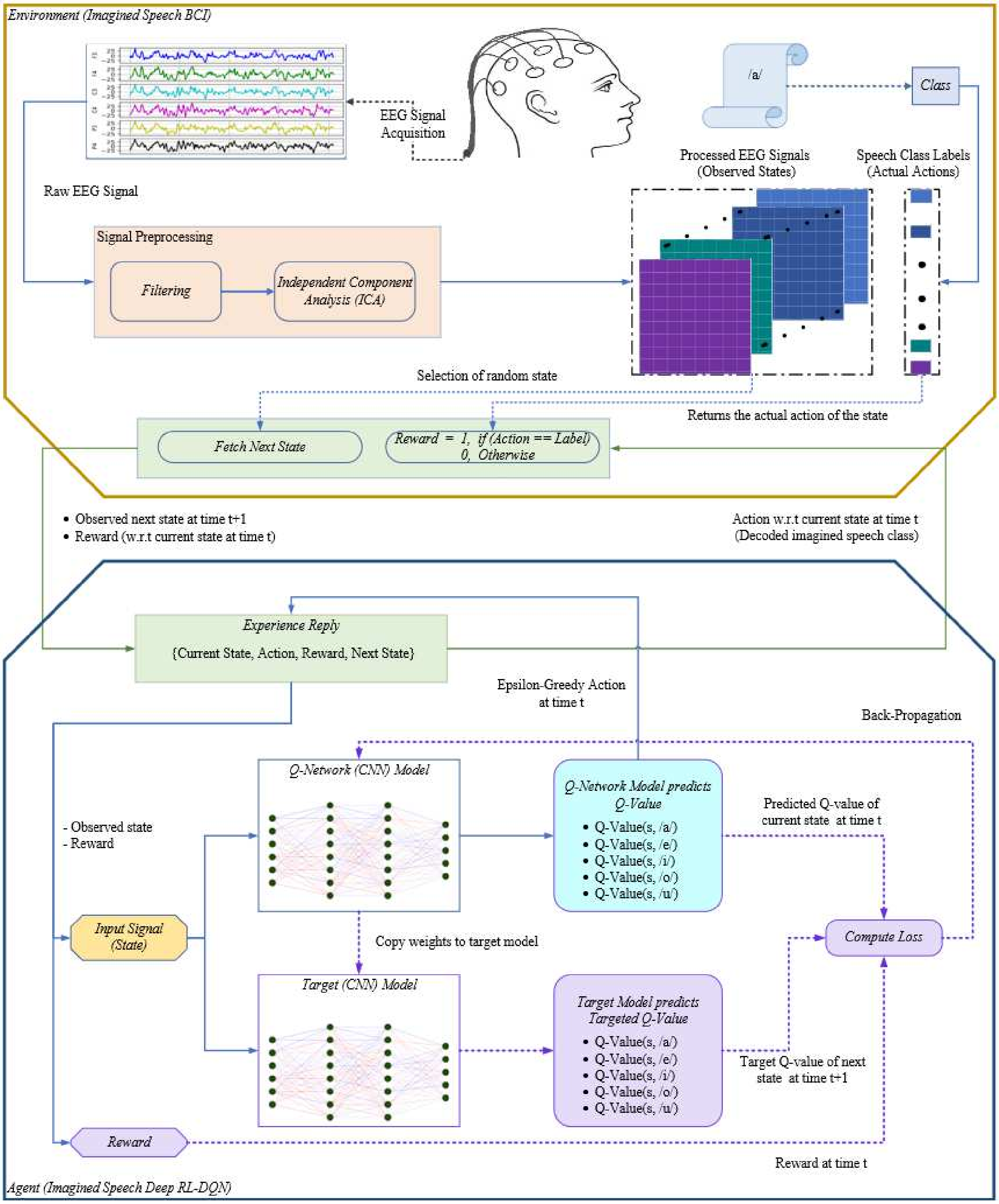

Fig. 1. The imagined speech decoder's architecture is based on DQN deep reinforcement learning. The observation states and rewards are collected from the environment, and the observation states are fed into the DQN model to produce the desired action.

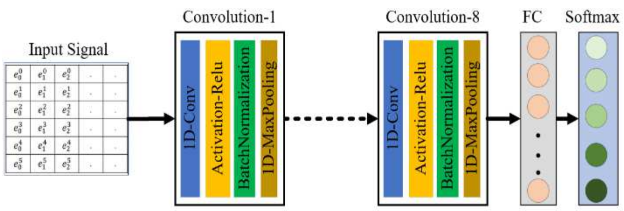

Fig. 2 The diagram illustrates the Q-network CNN model framework..

DQN, and the environment rewards the action with a reward of 0.

The DRL system based on the DQN approach allows the imagined speech signal decoder agent to leverage multiple state-action value combinations in order to get the largest feasible reward for improving system performance over time-step t = 0, 1, 2, 3, …, max time (T). The DQN agent observes

s_t the current state at time instant t and uses the Q-network to identify an action a_t, after which it receives a delayed reward r_t for the state-action value (s_t, r_t) and the next state s_{t+1} interacting with the environment at time instant t+1. The Q-network model learns to enhance performance by using the loss calculated by Q-values derived from the current state s_t, reward r_t and the next state s_{t+1}.

E. Learning flow of DQN

Initialize the Q-network model's weights (W) with random values and replicate them to the target model weights (W'). The DQN agent learning process is structured into several episodes, one of which is represented below.

- The observed state or processed EEG signal (s_t) was fed into our Q-network model, which outputs the Q-value of all possible actions or decoded imagined speech classes (a_t) for the observed state.
- The agent selects a random action with a probability of epsilon and leverages the most favorable action with a probability of 1-epsilon based on the epsilon-greedy criteria. The epsilon-greedy action is performed by the Experience Replay, which then obtains the next state and reward from the environment.
- The experiment replay stored the observed sample data e_t = {current state (s_t), action (a_t), reward (r_t), next state (s_{t+1})} in the replay memory. The replay memory M = $\{e_0, e_1, e_2, \ldots\}$ contains the replay samples used for Q-network training with limited memory capacity. Select a training batch (size = n) of random samples from the replay memory as input for the Q-network and the target network.
- The target network computes the target Q-value using the next state as input and predicts the optimal Q-value from all feasible actions. Subtract the targeted Q-value from the predicted Q-value to calculate the mean squared error (MSE) loss as specified in (2). Using gradient descent, back-propagate the loss and upgrade the weights of the Q-network model. No loss is calculated and no back-propagation is conducted since the target model is not trained and remains unaltered. This completes a time-step in the learning and training process of the Q-network.

$$loss = \frac{1}{n}\sum_n (predict - target)^2, \quad (2)$$

where predicted Q-value = $Q(s_t, a_t; W)$,
target Q-value = $(r_t + \gamma\, max_{a_{t+1}}\, Q(s_{t+1}, a_{t+1}; W'))$,
and γ is the discount factor.

- After specified time steps (T), copy the Q-network weights to the target network. As a result, the target network is able to enhance overall model weights and accurately predict Q-values.

In this experiment, we used the Adam optimizer [21] to train the proposed DQN model. Table II summarizes the experimental parameters of the DRL agent CNN model framework. Table III summarizes the training and learning parameters of the DRL model.

TABLE II. AN OUTLINE OF THE PARAMETERS IN THE DQN AGENT CNN MODEL FRAMEWORK

Parameter	*Value*
Number of layer	8
Convolution	Conv1D
Kernel size	3
No. of filters	256
Dropout	0.2
Activation	Relu
Pooling	maxpooling1D (2)
Fully connected layer	5
Activation	softmax

TABLE III. THE BRIEF OUTLINE OF THE DRL MODEL TRAINING AND LEARNING PARAMETERS

Parameter	*Value*
optimiser	adam
model initial weight	normal
epislon-gready	0.01
num-episode	100000
discount factor	0.1
minibatch size	32
learning rate	0.001
replay memory size	100000

III. RESULTS

We evaluated the efficacy of the proposed framework for imagined speech classification based on EEG data using the possible generalization of information in five-fold cross-validation. To evaluate classification performance, accuracy, precision, and recall were employed as assessment criteria. To avoid providing incorrect information about a classifier's performance on high-dimensional data, we used the harmonic mean f1-score.

The proposed DQN model's multiclass imagined speech signal classification five-fold cross-validation accuracy for all participants is 81.6947% ± 01.1334% (mean ± standard deviation). Table IV shows the classification performance measures (f1-score, precision, and recall) of the DQN model classifier results for each imagined speech class.

TABLE IV. THE RESULTS OF THE IMAGINED SPEECH SIGNALS CLASSIFICATION MODEL PERFORMANCE METRICS (MEAN ± STANDARD DEVIATION %): PRECISION, RECALL, AND F1-SCORE

Speech Prompt	*Precision (mean ± std)*	*Recall (mean ± std)*	*F1-Score (mean ± std)*
/a/	0.8157±0.0218	0.8079±0.0180	0.8115±0.0138
/e/	0.8121±0.0416	0.8143±0.0347	0.8122±0.0269
/i/	0.7895±0.0640	0.8289±0.0509	0.8049±0.0221
/o/	0. 8629±0.0298	0.8217±0.0268	0.8411±0.0143
/u/	0.8202±0.0297	0.8173±0.0222	0.8185±0.0227

Fig. 3 illustrates the confusion matrix of the proposed model's evaluations. The predicted imagined speech class is represented by the rows in the confusion matrix, whereas the actual imagined speech class is represented by the columns. The primary diagonal cells in the confusion matrix indicate the proportions of correctly classified imagined speech signals. The remaining cells, on the other hand, represent the erroneously classified imagined speech signal proportion.

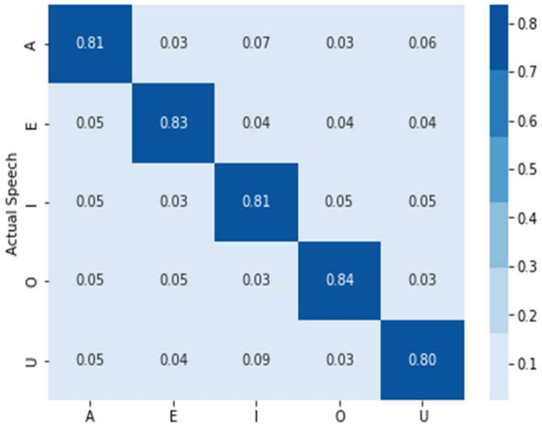

Fig. 3. The model confusion matrix's classification performance is indicated by the diagonal, which shows the reliability of all imagined speech classes.

IV. DISCUSSIONS

In this study, we attempted to examine whether the RL based on DQN classifiers was able to learn and train with feedback from the imagined speech BCI environment and classify the imagined speech signals.

The DQN algorithm can exploit the difference between separate imagined speech vowel signals, as indicated by its performance mean accuracy of 81.6947% with a standard deviation of 01.1334. According to the confusion matrix, the true-positive rates for the imagined speech prompts /a/, /e/, /i/, /o/, and /u/ were 0.81, 0.83, 0.81, 0.84, and 0.80, showing that the DQN accomplished balanced multiclass classification across all speech classes. The proposed DQN model f1-score for the imagined speech class was in the range of 80-84%, which is a statistic that shows the balance between recall and precision. Furthermore, the proposed DRL model was compared against state-of-the-art imagined speech vowel decoding performance, and it performed with substantial reliability. Table V displays the state-of-the-art comparison of the overall accuracy classification performance of imagined speech vowels. Moreover, the DRL agent could learn to identify distinct temporal properties of imagined speech EEG signals throughout the training phase by employing the reward obtained for past signal decoding results, steadily increasing model performance across all the imagined speech vowels.

TABLE V. THE COMPARISON OF THE OVERALL ACCURACY PERFORMANCE OF THE PROPOSED MODEL WITH THE RELEVANT WORK IN IMAGINED SPEECH CLASSIFICATION

Authors	*Model*	*Accuracy (%)*
Coretto et al., 2017 [8]	SVM	22.32
Cooney et al., 2019 [10]	CNN	35.68
Tamm et al., 2020 [12]	CNN	23.98
Sarmiento et al., 2021 [14]	CNN	65.62
Bakhshali et al., 2022 [16]	SVM	81.1
Proposed Method	DRL	81.69

The results of the experiments demonstrated that the proposed CNN-based DQN agent architecture performed effectively in classifying neural EEG imagined speech in general. The results in precision and recall suggested that utilizing reinforcement learning as classifiers in decoding imagined speech signals was feasible, which might pave the way for future research into real-time or online synthesis of imagined speech. The experiment's shortcoming was that it only used one database, despite the fact that signals were collected from several individuals. Therefore, future research might explore incorporating distinct imagined speech stimuli from different signal databases to explore DRL model learning capacity with the intent of providing generalization ability, taking into consideration individual variance in EEG signals measured for the same imagined speech stimuli.

V. CONCLUSION

We focused on the imagined speech decoding strategy in this work to exemplify the potential of reinforcement learning in the training process of a deep learning model by employing feedback or reward for the model classification of neural imagined speech class.

In the proposed method, the integrated Q-Network with a convolutional neural network based reinforcement learning agent to communicate with the imagined speech BCI environment, where noninvasive EEG signals are measured and preprocessed, and imagined speech signals are classified by DQN agents using a trial-and-error strategy with environment feedback. We describe the structure of our proposed technique, which includes fundamental DQN notions such as the agent and the environment, as well as Q-values, action selection based on epsilon-greedy criteria, and feedback rewards. We investigated several DQN agent parameters to determine the optimal fine-tuning techniques for classifying imagined speech tasks in order to improve the learning capabilities of the proposed method.

REFERENCES

[1] X. Gu *et al.*, "EEG-Based Brain-Computer Interfaces (BCIs): A Survey of Recent Studies on Signal Sensing Technologies and Computational Intelligence Approaches and Their Applications," IEEE/ACM Trans. Comput. Biol. Bioinform., vol. 18, no. 5, pp. 1645–1666, Sep. 2021, doi: 10.1109/TCBB.2021.3052811.

[2] Tripathy, H. K., Mishra, S., Suman, S., Nayyar, A., & Sahoo, K. S. (2022). Smart COVID-shield: an IoT driven reliable and automated prototype model for COVID-19 symptoms tracking. Computing, 1-22.

[3] Suman, S., Mishra, S., Sahoo, K. S., & Nayyar, A. (2022). Vision Navigator: A Smart and Intelligent Obstacle Recognition Model for Visually Impaired Users. Mobile Information Systems, 2022.

[4] J. Xu, S. Mitra, C. Van Hoof, R. F. Yazicioglu, and K. A. A. Makinwa, "Active Electrodes for Wearable EEG Acquisition: Review and Electronics Design Methodology," *IEEE Rev. Biomed. Eng.*, vol. 10, pp. 187–198, 2017, doi: 10.1109/RBME.2017.2656388.

[5] M. Matsumoto and J. Hori, "Classification of silent speech using support vector machine and relevance vector machine," *Appl. Soft Comput.*, vol. 20, pp. 95–102, Jul. 2014, doi: 10.1016/j.asoc.2013.10.023.

[6] L. Wang, X. Zhang, X. Zhong, and Z. Fan, "Improvement of mental tasks with relevant speech imagery for brain-computer interfaces," *Measurement*, vol. 91, pp. 201–209, Sep. 2016, doi: 10.1016/j.measurement.2016.05.054.

[7] B. M. Idrees and O. Farooq, "EEG based vowel classification during speech imagery," *2016 3rd Int. Conf. Comput. Sustain. Glob. Dev. INDIACom*, pp. 1130–1134, Mar. 2016.

[8] G. A. Pressel Coretto, I. E. Gareis, and H. L. Rufiner, “Open access database of EEG signals recorded during imagined speech,” in *SPIE 10160, 12th International Symposium on Medical Information Processing and Analysis*, Jan. 2017, p. 1016002. doi: 10.1117/12.2255697.

[9] M. N. I. Qureshi, B. Min, H.-J. Park, D. Cho, W. Choi, and B. Lee, “Multiclass Classification of Word Imagination Speech with Hybrid Connectivity Features,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 10, pp. 2168–2177, Oct. 2018, doi: 10.1109/TBME.2017.2786251.

[10] C. Cooney, R. Folli, and D. Coyle, “Optimizing Layers Improves CNN Generalization and Transfer Learning for Imagined Speech Decoding from EEG,” in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy, Oct. 2019, pp. 1311–1316. doi: 10.1109/SMC.2019.8914246.

[11] J. T. Panachakel, A. G. Ramakrishnan, and T. V. Ananthapadmanabha, “Decoding Imagined Speech using Wavelet Features and Deep Neural Networks,” in *2019 IEEE 16th India Council International Conference (INDICON)*, Rajkot, India, Dec. 2019, pp. 1–4. doi: 10.1109/INDICON47234.2019.9028925.

[12] M.-O. Tamm, Y. Muhammad, and N. Muhammad, “Classification of Vowels from Imagined Speech with Convolutional Neural Networks,” *Computers*, vol. 9, no. 2, p. 46, Jun. 2020, doi: 10.3390/computers9020046.

[13] X. Zhang, H. Li, and F. Chen, “EEG-based Classification of Imaginary Mandarin Tones,” *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2020, pp. 3889–3892, Jul. 2020, doi: 10.1109/EMBC44109.2020.9176608.

[14] L. C. Sarmiento, S. Villamizar, O. López, A. C. Collazos, J. Sarmiento, and J. B. Rodríguez, “Recognition of EEG Signals from Imagined Vowels Using Deep Learning Methods,” *Sensors*, vol. 21, no. 19, p. 6503, Sep. 2021, doi: 10.3390/s21196503.

[15] D.-Y. Lee, M. Lee, and S.-W. Lee, “Decoding Imagined Speech Based on Deep Metric Learning for Intuitive BCI Communication,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1363–1374, 2021, doi: 10.1109/TNSRE.2021.3096874.

[16] M. A. Bakhshali, M. Khademi, and A. Ebrahimi-Moghadam, “Investigating the neural correlates of imagined speech: An EEG-based connectivity analysis,” *Digit. Signal Process.*, vol. 123, p. 103435, Apr. 2022, doi: 10.1016/j.dsp.2022.103435.

[17] M. R. Asghari Bejestani, Gh. R. Mohammad Khani, V. R. Nafisi, and F. Darakeh, “EEG-Based Multiword Imagined Speech Classification for Persian Words,” BioMed Res. Int., vol. 2022, pp. 1–20, Jan. 2022, doi: 10.1155/2022/8333084.

[18] Mishra, S., Jena, L., Tripathy, H. K., & Gaber, T. (2022). Prioritized and predictive intelligence of things enabled waste management model in smart and sustainable environment. PloS one, 17(8), e0272383.

[19] Mohapatra, S. K., Mishra, S., Tripathy, H. K., & Alkhayyat, A. (2022). A sustainable data-driven energy consumption assessment model for building infrastructures in resource constraint environment. Sustainable Energy Technologies and Assessments, 53, 102697.

[20] V. Mnih *et al.*, “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015, doi: 10.1038/nature14236.

[21] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *ArXiv14126980 Cs*, Jan. 2017, [Online]. Available: http://arxiv.org/abs/1412.6980