Fractional Hybrid Election Based Optimization with DRN for Brain Thoughts to Text Conversion Using EEG Signal

Jyoti Prakash Botkar[1], Virendra V. Shete[2] and Ramesh Y. Mali[3]

[1]*Department of Electronics and Communication, School of Engineering, MIT ADT University, Rajbaugh Loni Kalbhor, Solapur Highway, Pune - 412201, Maharashtra, India*

[2]*School of Engineering, MIT ADT University, Rajbaugh Loni Kalbhor, Solapur Highway, Loni Kalbhor Railway Station, Pune - 412201, Maharashtra, India*

[3]*Department of Electric and Electronics Engineering, MIT School of Computing, MIT ADT University, Rajbaugh Loni Kalbhor, Solapur Highway Railway Station, Pune - 412201, Maharashtra, India*

Keywords: Electroencephalography, Brain Computer Interface, Maximum A Posteriori Probability, Deep Recedual Netork, Gaussion Filter.

Abstract: A Brain-Computer Interface (BCI) system based on electroencephalography (EEG) permits users to interconnect to contact the outer world using devices, like intelligent robots and wheelchairs by interpreting their brain EEG signals. Translating brain dynamics into natural language is critical to BCI, which has seen considerable development in recent years. This work proposes a novel method for brain thoughts-to-text conversion utilizing EEG signals. At first, the Brain EEG signal is taken from a database, which contains EEG signals related to tried imaginary thoughts of questions and statements. Afterward, signal preprocessing is done by exploiting a Gaussian Filter to reduce noise in EEG signals. Consequently, signal segmentation is done by the Maximum A posteriori probability (MAP) estimator. Later, feature extraction is done. After, word recognition is accomplished using Deep Residual Network (DRN) tuned by the proposed optimization technique called Fractional Hybrid Election-Based Optimization (FHEBO). Here, the FHEBO is developed by the amalgamation of Hybrid Election-Based Optimization (HEBO) and Fractional Calculus (FC). Further, language modelling is accomplished by utilizing the Gaussian Mixture Model (GMM). Moreover, the proposed approach is observed to record maximal text conversion accuracy at 91.765%, precision at 92.765%, F-measure at 94.241%, recall at 95.765%, and minimal error rate at 8.235%.

1 INTRODUCTION

Brain-Computer Interface (BCI) systems are progressed to decode an individual's intention, state of mind, and emotions, by observing person's brainwaves through sensors placed externally or internally on the human brain (Ullah & Halim, 2021). As a significant pathway across the human brain and the outside world, BCI systems allow users to interact or communicate with external devices like service robots or wheelchairs by their brain signals (Zhang et al., 2017) . In recent times, it is feasible to the availability of specific states of the electromagnetic field and neurons produced inside the brain. This is feasible due to the availability of various modalities, such as functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), and EEG. Among these technologies, EEG has benefits over others due to wireless connectivity, low cost, portability, easy handling, and portability (Kumar et al., 2018) . EEGs are a vital section of all BCI systems and are used for recording brain signals. The voltage fluctuations created through the movement of ions inside neurons in the brain in response to certain incitements are determined using EEG. There are two techniques of measuring voltage fluctuations utilizing EEG sensors, such as non-invasive and invasive. Sensors are surgically kept under the topmost part of the skull. Therefore, these techniques are challenging to implement, unsafe, and expensive, sensors are kept outside of the scalp in the non-invasive BCI. This makes non-invasive techniques straightforward, economical, handy, and safe (Ullah & Halim, 2021).

BCI technology engages the communication of data from one person's brain to external devices utilizing a wireless medium. A significant advantage of this technology is that it is non-invasive with no complications in utilizing bulky devices such as exoskeletons and comes at a low cost (Rajesh et al., 2020). BCI receives electrical signals and changes them into control commands such as a biological communication channel without compromise in a natural way (Junwei et al., 2019) . An EEG-based BCI was used as a user identification method. They also have many applications in the medical field. Emotion

Botkar, J. P., Shete, V. V. and Mali, R. Y.
Fractional Hybrid Election Based Optimization with DRN for Brain Thoughts to Text Conversion Using EEG Signal.
DOI: 10.5220/0013606100004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 2*, pages 852-861
ISBN: 978-989-758-763-4

recognition has been applied in several domains such as commerce, security, education, and others. Words associated with the detected emotion are first modified utilizing a correlation finder among emotion and words. Next, sentence correctness is checked using a language model depend on Long Short Term Memory (LSTM). The LSTM networks are widely used in speech recognition, time series prediction, grammar learning, handwriting recognition, etc (Gupta et al., n.d.). Visual/mental imagery (VI/MI) is the processing of visual data from the memory (rather than perceptual). Deep neural networks are utilized for EEG related tasks which includes emotion recognition, mental pressure detection and sleep analysis in addition to the MI-EEG task classification (Ullah & Halim, 2021). Deep Learning (DL) algorithms were introduced for classifying EEG signals. A convolutional neural network (CNN) has been effectively used in EEG-based BCIs for classification and end-to-end feature extraction as well as speech recognition and computer vision (Ahn & Lee, 2021).

Brain thought to text conversion is effectuated utilizing EEG signal in this work., Initially, Input Brain EEG signal is taken from the dataset, which contains EEG signals associated to attempted imaginary thoughts or questions/statements. Afterward, signal preprocessing is effectuated utilizing a Gaussian Filter to lessen noise in the EEG signals. The signal segmentation is done with the help of MAP estimator. Feature extraction is done based on frequency-based features. Later, word recognition is carried out using DRN trained using the proposed optimization technique called FHEBO. The devised FHEBO is developed by the amalgamation of FC and HEBO. The HEBO is established by the integration of the Election Optimization Algorithm (EBOA) and Hybrid Leader Based Optimization (HLBO) .

2 LITERATURE REVIEW

(Ullah & Halim, 2021) designed Deep convolutional neural network (DCNN) for brain thoughts to text conversion. The technique recorded a superior recognition for all the alphabets in English language even with fewer parameters. Although, the performance was affected by the interference from other parameters. (Willett et al., 2021) introduced Recurrent Neural Network (RNN) for brain-to-text communication via handwriting. The method worked successfully with data-limited regimes and unlabelled neural sequences. However, the method failed to improve longevity and performance. (Zhang et al., 2017) proposed a Hybrid DL model based on a convolutional recurrent neural network for brain thoughts-to-text conversion. This approach exhibited high adaptability and worked well in multi-class scenarios, although the training time needed was high.(Kumar et al., 2018) designed a Random Forest (RF) classifier for brain thoughts-to-text conversion. The technique was robust and had a high accuracy. The method failed to be applied in several BCI applications like rehabilitation systems and artificial telepathy and rehabilitation systems.(Rajesh et al., 2020) designed a Novel tiny symmetric algorithm for brain thoughts-to-text conversion. The technique was lightweight and enabled the transfer of information from the patient's brain to the caretaker securely. The method failed to decrease the size of the system to be easily handled by patients.

2.1 Challenges

The major difficulties met by the various techniques in brain thoughts to text conversion are listed as,

- The DCNN-based method presented in (Ullah & Halim, 2021) was not extended to recognizing individual words or generating complete sentences.
- In (Zhang et al., 2017) , the hybrid DL model failed to focus on enhancing accuracy in a person-independent situation, where few subjects participate in training and the rest of the subjects participate in testing.
- The RF method (Kumar et al., 2018) failed to excerpt the robust features from EEG signals to enhance the system's recognition achievement.

The key challenges in developing BCI systems based on EEG signals are that the techniques suffer from high computational complexity and fluctuations in EEG signals due to the environmental noise limiting their performance.

3 PROPOSED FHEBO-DRN FOR BRAIN THOUGHTS TO TEXT CONVERSION

This work proposes a new method for brain thoughts-to-text conversion utilizing an EEG signal. The proposed technique is implemented as follows. Initially, input brain EEG signals is taken from the dataset, which contains EEG signals corresponding to attempted imaginary thoughts of questions/ statements. Next, signal preprocessing is carried out utilizing a Gaussian Filter (Kopparapu & Satish, 2011) to reduce noise in the EEG signals. Consequently, signal segmentation is utilizing MAP estimator (Popescu, 2021). After that, feature extraction is done based on frequency-based features, like spectral spread, power spectral density, total power ratio, spectral flux, and spectral centroid. Next, word recognition is carried out utilizing DRN (Chen et al., 2019) trained utilizing the proposed optimization technique called FHEBO. Here, the proposed FHEBO is developed by the combination of FC (Bhaladhare & Jinwala, 2014) and HEBO. The HEBO is developed by the combination of HLBO (Dehghani & Trojovský, 2022), and EBOA (Trojovský & Dehghani, 2022). Further, the Language modelling for speech recognition is carried out using the Gaussian Mixture Model (GMM) (Afy et al., n.d.) . Figure 1 displays the structural diagram of the FHEBO-DRN for brain thoughts-to-text conversion.

3.1 Data acquisition

The input EEG signal exploited in this work is taken from the dataset, and the dataset is described by

$$G = \{K_1, K_2, K_3 \cdots K_i, K_h\} \tag{1}$$

where G is the dataset, h specifies the number of EEG signals, K_i denotes the i^{th} EEG considered for processing.

3.2 Signal Preprocessing

The EEG signal K_i is forwarded to the signal pre-processing phase for removing the inherent noise in the signal. The Gaussian filter [9] effectively smooths the signal by reducing high-frequency noise. Gaussian filter convolves the signal with a Gaussian function, which has a bell-shaped curve. The probability distribution function of the filter is expressed as,

$$Z_k\left(\sigma_k, \mu_k^2, p\right) = \frac{1}{2\pi p} \exp^{-\left\{\frac{(p-\sigma_k)^2}{2\mu_k^2}\right\}} \tag{2}$$

where, σ_k indicates the mean, μ_k^2 symbolizes variance, p is the time interval. The output of this phase is the denoised signal depicted as M_i .

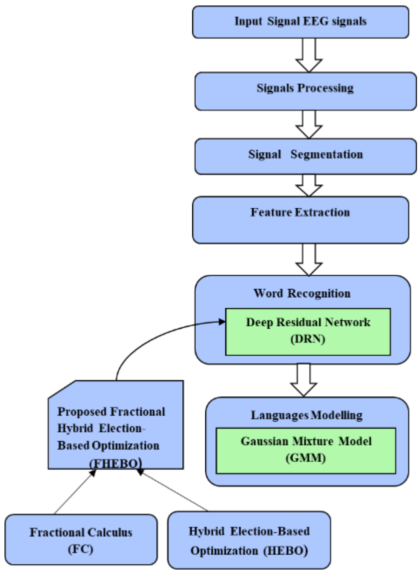

Figure 1: A systematic view of the FHEBO-DRN for brain thoughts to text conversion

3.3 Signal segmentation

The pre-processed signal M_i is subjected to signal segmentation (Popescu, 2021), which is accomplished using MAP estimation (Popescu, 2021) as it is easy to implement and has a low computational complexity. The conceptual description of the segmentation algorithm utilizing MAP estimation is provided below.

The segmentation issue is resolved by finding sequence $m^l = m_1, m_2, .., m_l$ that minimize the optimal criteria of the form,

$$\hat{m}_l = arg \min_{l \geq 1, 0 << .. << m_l = M} J(m^l) \tag{3}$$

where, m^l represents the groups of parameter vectors, noise scaling, and jump times, J indicates the sum of squared residuals.

To determine all segments. a linear regression model is used. For the measurements associated with the j^{th} segment, $e_{mj-1} + 1, \dots e_{mj} = e_{mj-1}^{mj} + 1$, the least squares assessment of its sample parameters and covariance matrix is determined:

$$\hat{\theta}(j) = H(j) \sum_{h=m_{j-1}+1}^{m_j} \phi_h F_h^{-1} e_h \tag{4}$$

$$(j) = \left(\sum_{h=m_j+1}^{m_j} \phi_h F_h^{-1} \phi_h^T\right)^{-1} \tag{5}$$

wherein, ϕ_h symbolizes the regressor, H, T is the transpose, F_h indicates nominal covariance matrix of the noise, and e_h is considered to be Gaussian and m_l specifies the time index.

For optimal segmentation algorithm, the following parameters are utilized:

$$J(j) = \sum_{h=m_{j-1}+1}^{m_j} \left(e_h - \phi_h^T \hat{\theta}(j)\right)^T F_h^{-1} \left(e_h - \phi_h^T \hat{\theta}(j)\right) \quad (6)$$

$$G(j) = -\log \det H(j) \quad (7)$$

$$M(j) = m_j - m_{j-1} \quad (8)$$

Here, M signifies the number of data in each segment and G denotes the logarithmic value of the determinant of the covariance matrix. The values in the m^l segmentation has a degree of freedom $l-1$ and needs 2^M segmentations, which makes the process extremely complex.

The MAP estimator is used to address this high dimensional complexity by considering the assumptions on noise scaling $\lambda(j)$, as follows:

$$Data : Signal\, e_{h,}\ h = 1 \ldots M \quad (9)$$

1: Analyse every segmentation, with the jump times m^l and the number of jumps m, for all cases.

2: For all segmentation results, the ideal model for all segments is calculated in the form of the covariance matrices $H(j)$ and least square estimates $\hat{\theta}(j)$.

3: For all segment calculate:

$$J(j) = \sum_{h=m_{j-1}+1}^{m_j} \left(e_h - \theta_h^T \hat{\theta}(j)\right)^T F_h^{-1} \left(e_h - \phi_h^T \hat{\theta}(j)\right) \quad (10)$$

$$G(j) = -\log \det H(j) \quad (11)$$

$$M(j) = m_j - m_{j-1} \quad (12)$$

4: To determine $\widehat{m}_l$, the MAP estimator is used based on three constraints on noise scaling, $\lambda(j)$, with $t(0<t<1)$ the transformation probability at every time instant.

(i)Known $\lambda_0 = \lambda(j)$,

$$\hat{m}^l = \arg\min_{m^l,l} \sum_{j=1}^{l} (G(j) + J(j)) + 2l \log \frac{1-t}{t} \quad (13)$$

where, $t(0<t<1)$ denotes the change probability at all instants.

ii) constant and unknown $\lambda(j) = \lambda$

$$\hat{m}^l = argmin_{m^l,l} \sum_{j=1}^{l} G(j) + (Mf - lb - 2) \times log \sum_{j=1}^{l} \frac{J(j)}{Mf-lb-4} + 2l\, log \frac{1-t}{t} \quad (14)$$

iii)unknown and changing $\lambda(j)$

$$\hat{m}^l = argmin_{m^l,l} \sum_{j=1}^{l} G(j) + (M(j) - lb - 2) \times log \sum_{j=1}^{l} \frac{J(j)}{M(j)-lb-4} + 2l\, log \frac{1-t}{t} \quad (15)$$

Results: Number l and locations m_j, $m^l = m_1, m_2, .., m_l$

Only one of the equations in step 4 is utilized to compute $\widehat{m}_l$, based on the hypothesis of noise scaling. The preprocessed signal can be split into various segments as given below,

$$L_L = \{L_1, L_2, \ldots L_y, L_e\} \quad (16)$$

where, λ is commonly selected as an independent function, $\lambda(j)$ is noise scaling based on segmentation. The segmented output produced by the MAP is signified as r.

3.4 Feature Extraction

The segmented signal r is subjected to feature extraction. Feature extraction is done based on frequency-based features. The frequency-based features like spectral flux, spectral centroid, spectral spread, power spectral density, and total power ratio are explained below,

i) Spectral flux

The spectral components (Mannepalli et al., 2017) of the signal are extracted with the help of the spectral flux. Spectral components are a significant feature due to the spectral contents of the signal change over time, and the efficiency of recognition decreases. The equation of spectral flux is defined by,

$$b_1 = \sum_{C=1}^{\frac{D}{2}} (|Z[(c)]| - |Z_{i-1}[c]|)^2 \quad (17)$$

where, D specifies the vector length, $Z(c)$ specifies the spectral magnitude of c^{th} instant and b_1 is the spectral flux.

ii) Spectral centroid

The Spectral Centroid (SC) (Hassan et al., 2016) indicates the center of mass of the spectrum. The SC is determined using the following expression,

$$b_2 = \frac{\sum_{x=1}^{D} xS(x)}{\sum_{x=1}^{D} S(x)} \quad (18)$$

where the amplitude value of x^{th} bin is specified as $S(x)$ and $b_2$represents the spectral centroid.

iii) Spectral spread

Spectral Spread (SS) (Hassan et al., 2016) is the spread of the spectrum around its centroid (ie) it measures the standard deviation of the spectral distribution. The spectral spread is defined by,

$$b_3 = \frac{\sum_{x=1}^{D}(x-SC)^2 S(x)}{\sum_{x=1}^{D} S(x)} \tag{19}$$

where b_3 indicates the spectral spread.

iv) Power Spectral Density (PSD)

PSD (Hong et al., 2018) specifies the strength of a signal as a function of frequency. The autocorrelation sequence for a specified data set is initially assessed for nonparametric techniques. The PSD is expressed as,

$$b_4 = Y_k^n = \frac{1}{D}\sum_{g=1}^{D}\left|Z_{gn}e^{-j2\pi gn}\right|^2 \tag{20}$$

where, Z_{gn} is the n^{th} sequence magnitude, Y_k^n designates the power of the n^{th} data sequence's k^{th} frequency band, D is the sample count and b_4 denotes the PSD.

v)Tonal power ratio

This is utilized to determine the tonalness of a speech signal (Mannepalli et al., 2017) . This is the percentage of the tonal power of the spectrum components to the whole power, and is defined by,

$$b_5 = \frac{O(a)}{\sum_{c=0}^{\frac{g}{2}}|Z(z,a)|^2} \tag{21}$$

where, $O(a)$ represents the tonal power that is calculated through totaling every bin z that lies above a threshold and is local maximal, $Z(z,a)$represents the preprocessed signal spectrum and b_5 indicates tonal power ration. The feature vector is denoted by $v = \{b_1, b_2, b_3, b_4, b_5\}$.

3.5 Word recognition

Word recognition utilizing Deep learning is generally the simplest approach to speech recognition. Here, the input feature $v$is applied to the DRN(Chen et al., 2019) for establishing word recognition. Further, the training process of the DRN is effectuated using the FHEBO algorithmic approach. Here, FHEBO is the combination of FC and HEBO.

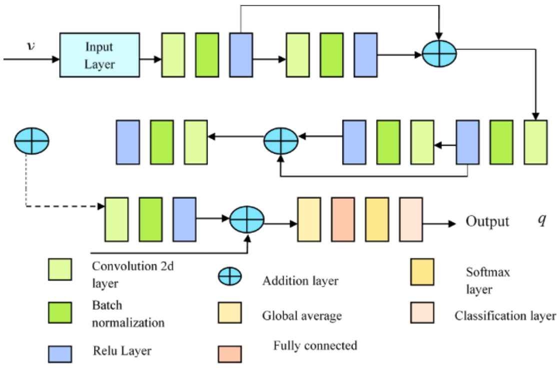

Figure 2: Architecture of DRN

3.5.1 Architecture of DRN

The DRN is a Deeper Neural Network (DNN) with low gradient vanishing or explosion and higher training speed. The DRN (Chen et al., 2019) was originally designed for complicated image classification processes, and it consists of fully connected layers and pooling layers, 2-dimensional convolution layers. A DRN structure contains various layers, such as (i) Convolutional layers (ii) pooling layers (iii) Convolutional layers (iv) Residual blocks (v) Linear Classifier. These layers in DRN are described below.

1)Convolutional Layer: A typical two-dimensional convolutional layer can significantly reduce the free parameters in the training procedure and improve performance to the benefits of the weight sharing and local receptive filed. The following equation is used to model the process established in the convolutional layer.

$$conv1d(v) = \sum_{w=0}^{B_{in}-1} W_w * v \tag{22}$$

where, v is the input applied to the DRN, W specifies the learnable kernel matrix, B^{in} is the input feature dimension and $*$ is the cross-correlation operator.

2)Pooling layer: The function of this layer is generally applied to subsequent convolution layers, and is mostly exploited for handling overfitting and decreasing the spatial extent of the feature maps.

3) Activation function: A activation function that is non-linear is utilized to higher the linearity of the extracted features. ReLU alleviate the vanishing gradient problem and significantly accelerates convergence.

4)Batch Normalization: Batch normalization minimizes internal covariate variation by scaling input layers and smoothing implementations, thus enhancing training speed and reliability, when mitigating gradient bursting/fading and overfitting issues with better learning rates.

5)Residual Blocks: The residual block had shortcut connection from input to output. The input is attached directly to the outputs, when output and input are of equal dimension, otherwise a dimension-matching factor is exploited.

6)Linear Classifier: This layer encompasses a Softmax function and fully connected (FC) layer. The following illustrates the output of this layer,

$$q = T_{P\times U} v_{U\times V} + u_{P\times V} \tag{23}$$

where, $T_{P\times U}$ specifies the weight matrix of $P \times U$ dimension, $v_{U\times V}$ represents input feature map of $U \times V$ dimension and u is the bias of dimension $P \times V$. The output is designated as q. Figure 2 illustrations the architecture of DRN.

3.5.2 Training of DRN with FHEBO

The DRN is structurally optimized by using the FHEBO developed by the combination of FC (Bhaladhare & Jinwala, 2014) and HEBO. The HEBO is developed by the combination of HLBO (Dehghani & Trojovský, 2022) and EBOA (Trojovský & Dehghani, 2022). A novel optimization algorithm called EBOA (Trojovský & Dehghani, 2022) is created that imitates the voting procedure to elect a leader. The basic impetus of EBOA is the procedure of voting, electing a leader and the cause of degree of awareness on electing a leader. The EBOA people are led by a search space directed by an elected leader. The HEBO is developed based on the process of guiding a solution to optimal one under the supervision of a hybrid leader. Instead of considering a specific member for updating the population, three members are taken into consideration for updating the solutions, thus avoiding convergence to local minima. FC (Bhaladhare & Jinwala, 2014) plays a significant role in increasing the efficiency of many methods like curve fitting, filtering, modeling, edge detection and pattern matching. The integration of FC in HEBO enhances the convergence speed and assists in attaining a global optimal solution.

i)Initialization

All members of the population specify a solution to the issue in FHEBO. In the mathematical perspective, the population is specified through a matrix utilizing the following equation,

$$I = \begin{bmatrix} I_1 \\ \vdots \\ I_a \\ \vdots \\ I_E \end{bmatrix}_{E\times t} = \begin{bmatrix} t_{1,1} \cdots t_{1,g} \cdots t_{1,t} \\ \vdots \ddots \vdots \iddots \vdots \\ t_{a,1} \cdots t_{a,g} \cdots t_{a,t} \\ \vdots \iddots \vdots \ddots \vdots \\ t_{E,1} \cdots t_{E,g} \cdots t_{E,t} \end{bmatrix}_{E\times t} \tag{24}$$

The primary location of member is estimated randomly as follows.

$$t_{a,g} = bp_g + n.\left(jp_g - bp_g\right), a = 1,2 \ldots, E, g = 1,2, \ldots, t, \tag{25}$$

where, n is a random number ranging in [0,1], and bp_g and jp_g represents the lower and upper limit of the g^{th}variable. I_a represents the a^{th} member, $I$is the population matrix, t indicates the count of decision variables, E specifies the population size, $t_{a,g}$ indicates the g^{th} problem variable value represented by a^{th} population member.

ii)Fitness function

The solution is considered to be the optimal solution with the minimal Mean Square Error (MSE), and the MSE is proved as follows.

$$MSE = \frac{1}{y}\sum_{j-1}^{y}\left(q_j^* - q_j\right)^2 \tag{26}$$

where, q_j^* specifies the anticipated value, y denotes the sample count, and $q_j$characterizes the recognized output by the DRN.

iii) Phase 1: Exploration

The members take part in the election depending on their awareness and vote for a candidate. A person's awareness is considered in terms of the goodness and quality of the objective function value. The updating process in this phase is modelled utilizing the equation given below,

$$t_{a,g}^{new,F1} = \begin{cases} t_{a,g} + n.\left(H_g - A.t_{a,g}\right), & BT_H < BT_g \\ t_{a,g} + n.\left(t_{a,g} - H_g\right), & else \end{cases} \tag{27}$$

where, H specifies the elected leader, A represents an integer with value as 1 or 2, and BT_H is its objective function value, H_g is EBOA g^{th} dimension, $t_{a,g}^{new,F1}$ refers the new created position for the a^{th} member, $t_{a,g}^{new,F1}$ is its g^{th} dimension, and $BT_{a}^{new,F1}$ indicates the objective function value.

This update process is modified by integrating the HLBO in the EBOA, and the updated equation of HEBO is expressed as,

$$t_{a,g}(x+1) = \frac{1}{2.n.A}\left[(H_g.n)(1+n.A) - (n.TH_{a,g})(1-n.A)\right] \tag{28}$$

In order to apply FC, subtracting $t_{a,g}(x)$on both sides,

By applying Fractional calculus [12],

$$t_{a,g}(x+1) - t_{a,g}(x) = \frac{1}{2.n.A}\left[(H_g - n)(1+n.A) - (n.TH_{a,g})(1-n.A)\right] - t_{a,g}(t) \tag{29}$$

$$T^{\alpha}\left[t_{a,g}(x+1)\right] = \frac{1}{2.n.A}\left[(H_g.n)(1+n.A) - (n.TH_{a,g})(1-n.A)\right] - t_{a,g}(x) \tag{30}$$

$$t_{a,g}(x+1) - \alpha t_{a,g}(x) - \frac{1}{2}\alpha t_{a,g}(x-1) - \frac{1}{6}(1-\alpha)t_{a,g}(x-2) + \frac{1}{24}\alpha(1-\alpha)(2-\alpha)t_{a,g}(x-3) = \frac{1}{2.n.A}\left[(H_g.n)(1+n.A) - (n.TH_{a,g})(1-n.A)\right] - t_{a,g}(x) \tag{31}$$

$$t_{a,g}(x+1) = \frac{1}{2.n.A}\left((H_g.n)(1+n.A) - (n.TH_{a,g})(1-n.A)\right) + (\alpha-1)t_{a,g}(x) + \frac{1}{2}\alpha t_{a,g}(x-1) + \frac{1}{6}(1-\alpha)t_{a,g}(x-2) + \frac{1}{24}\alpha(1-\alpha)(2-\alpha)t_{a,g}(x-3) \tag{32}$$

where, A specifies randomly selected number in (1,2), n is the random variable (0,1), TH represents hybrid leader, $t_{a,g}(x-1)$ is where the solution is located at iteration $(x-1)$, $t_{a,g}(x-3)$ at $x-3$, and $t_{a,g}(x-2)$ is the position at iteration $x-2$.

iv)Phase 2: Exploitation

The awareness of person in society in the voting and election procedure has a high effect on their decisions. Additionally, each person's activities and thoughts and the leader's authority on a person's awareness can enhance the people's awareness. From the mathematical view point, an optimal solution can be found depending on the local search near any designed solution and is formulated as,

$$t_{a,g}^{new,F2} = t_{a,g} + (1-2n).G.\left(1-\frac{x}{C}\right).t_{a,g}, \tag{33}$$

where is refers to the recently generated position for a^{th} member and its g^{th}dimension is specified as $t_{a,g}^{new,F2}$, C refers to the maximal count of iterations and x refers to iteration contour.

v) Re-evaluating the fitness

The objective of the modified solution is evaluated by equation (34) once updation is complete, and the solution that succeeds in attaining the lowest objective is deemed ideal.

vi)Termination

The above process is continued until the maximal iteration is grasped.

3.6 Language Model

Here, GMM is used to determine the text transmitted by EEG signals. Once the words are recognized by the DRN, it is subjected to the GMM. The GMM (Afy et al., n.d.) can map words from discrete space to continuous space. It consists of a linear layer that maps a vector in word space to a continuous parameter space. The vocabulary size and vector size are considered to be same. After that, these words are applied to a multi-layer perceptron (MLP) and combined based on the history. MLP selects an output word for all input words or assigns a probability value to every word in the vocabulary. Let us assume a vocabulary χ with dimension ε_1 and every word $y, 1 \leq y \leq \varepsilon_1$ is represented by a vector with value '1' at y^{th}position and remaining positions have a value '0'.The vector $\xi_r$is mapped to a vector e_r with less dimension ε_2 with the support of a matrix W as below,

$$e_r = W\xi_r \tag{34}$$

where, matrix W has a dimension $\varepsilon_1 \times \varepsilon_2$. Further, all the mapped words can be combined and represented as,

$$v_D = e_E(\beta_{g-1}) \dots e_E(\beta_1) \tag{35}$$

Here, β_k specifies the k^{th} history and E denotes to the argument's word identity. The GMM is built by performing linear mapping of every history into a small space given as follows,

$$t_D = Qv_D \tag{36}$$

where, t is the new feature space, which is textual format contained in the brain thoughts. Thus, the EEG signals are converted into text and Q specifies the linear mapping.

4 RESULT AND DISCUSSION

In this section, the assessment of the outcomes obtained by the DRN-FHEBO for conversion of brain thoughts to text is portrayed. Also, performance metrics and the dataset are discussed below.

4.1 Experimental setup

The designed FHEBO-DRN for conversion of brain thoughts to text is executed by the MATLAB tool.

4.2 Dataset description

Here, conversion from brain thoughts-to-text is carried out based on the data taken from the handwriting BCI dataset (GitHub - Fwillett/HandwritingBCI: Code from the Paper "High-Performance Brain-to-Text Communication via Handwriting," n.d.). This dataset includes neural activity in attempted handwriting, recorded with 43,501 characters in 1000 sentences over a period of 10.7 hours. The neuronal activity was recorded by placing two microelectrode arrays with 96 electrodes fixed to the hand region of the motor cortex. Also, it includes the signature output of BCI in real-time.

4.3 Evaluation measures

The effectiveness of FHEBO-DRN is estimated by utilizing assessment metrics such as Recall, F-Measure, Text conversion accuracy, and Precision.

i) Precision: Precision estimates the fraction of correctly classified samples or events among those that are positively classified and is given as follows,

$$Precision = \frac{TP}{TP+FP} \tag{37}$$

wherein, TP is the True Positive and FP specifies False Negative.

ii) F-Measure: This metric yields the Weighted Harmonic Mean of recall and precision, and is given by,

$$F-Measure = \frac{2TP}{(2TP+FP+FN)} \tag{38}$$

where, FN is the False Negative.

iii)Recall: Recall is a measure of how frequently a machine learning method exactly identifies positive instances from every true positive sample in the dataset. It is expressed as,

$$Recall = \frac{TP}{TP+FN} \tag{39}$$

iv)Text conversion accuracy: The accuracy of text conversion is determined by determining the proportion of the number of correctly converted words to the whole number of words.

v)Error rate: The error rate is the measure of the prediction error of the generated model concerning the true model. Error rate is often used in the context of classification models.

$$Errorrate = \frac{Numberoferrors}{Totalnumberoftaskattempts} \tag{40}$$

4.4 Comparative methods

The designed FHEBO-DRN is examined based on the techniques such as DCNN (Ullah & Halim, 2021), RNN (Junwei et al., 2019), Hybrid DL (Willett et al., 2021), and RF-classifier (Zhang et al., 2017) , HEBO-Distributed Long Short-Term Memory (DLSTM).

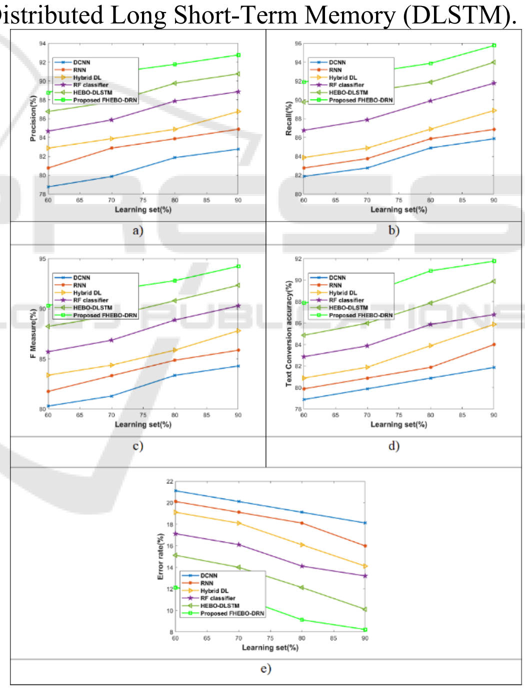

Figure 3: Comparative analysis of FHEBO-DRN based on learning set with a) precision, b) recall c) F-measure d) text conversion accuracy and e) error rate

4.5 Comparative Analysis

The assessment of FHEBO-DRN for brain thoughts to text conversion is executed with respect to learning set. The assessment of FHEBO-DRN for brain thoughts-to text conversion is done by learning set and is shown in figure 3 Figure 3a) illustrates the analysis of FHEBO-DRN based on precision. The precision recorded by FHEBO-DRN, with a learning set of 90% is 92.765% and the precision computed by DCNN is 82.765%, RNN is 84.878%, Hybrid DL is 86.766%, RF-classifier is 88.865% and HEBO-DLSTM is 90.756%. This illustrates FHEBO-DRN was successful in generating an enhanced performance of 2.17% than the HEBO-DLSTM. Figure 3b) describes the evaluation of FHEBO-DRN based on recall. The recall obtained by FHEBO-DRN is 95.765% with the learning set of 90% and the recall figured by DCNN is 85.865%, RNN is 86.867%, Hybrid DL is 88.867%, RF-classifier is 91.766% and HEBO-DLSTM is 93.987%. The improved performance of 9.29% is figured by FHEBO-DRN than existing RNN. Figure 3c) explains the analysis of FHEBO-DRN considering F-measure. The F-measure figured by FHEBO-DRN, with a learning set of 90% is 94.241% and the F-measure computed by DCNN is 84.287%, RNN is 85.861%, Hybrid DL is 87.804%, RF-classifier is 90.292% and HEBO-DLSTM is 92.343%. This demonstrates FHEBO-DRN successfully recorded an enhanced performance of 6.83% than Hybrid DL. In figure 3d), the valuation of FHEBO-DRN considering text conversion accuracy is illustrated. With learning set of 90%, the text conversion accuracy recorded by FHEBO-DRN is 91.765%. The text conversion accuracy valuated by DCNN is 81.867%, RNN is 83.998%, Hybrid DL is 85.877%, and RF-classifier is 86.786%, HEBO-DLSTM is 89.887%. The FHEBO-DRN achieved an improved performance by 5.42% than RF-classifier. Figure 3e) demonstrates the evaluation of FHEBO-DRN on the basis of error rate. The error rate obtained by FHEBO-DRN is 8.235%, with the learning set of 90% and the error rate figured by DCNN is 18.132%, RNN is 16.001%, Hybrid DL is 14.123%, RF-classifier is 13.213% and HEBO-DLSTM is 10.112%.

The comparative discussion of the FHEBO-DRN is demonstrated in table 1.

The application of FHEBO for tuning the DRN improved the convergence rate leading to accurate word recognition. Further, the application of MAP for signal segmentation and the high accuracy of DRN all enabled effective attainment of conversion of brain thoughts to text, resulting in excellent outputs.

Table 1: Comparative discussion

Metrics	HEBO-DLSTM	Proposed FHEBO-DRN	DCNN	RNN	Hybrid DL	RF classifier
Precision (%)	90.756	92.765	82.765	84.878	86.766	88.865
Recall (%)	93.987	95.765	85.865	86.867	88.867	91.766
F-measure (%)	92.343	94.241	84.287	85.861	87.804	90.292
Total conversion accuracy(%)	89.887	91.765	81.867	83.998	85.877	86.786
Error rate (%)	10.112	8.235	18.132	16.001	14.123	13.213

5 CONCLUSION

The conversion of brain thoughts to text depending on BCI is seen as an emerging field. In this paper, a technique for conversion of brain thoughts to text is designed. Here, Input Brain EEG signals are taken from the dataset, which contains EEG signals corresponding to attempted imaginary thoughts of questions/statements. Subsequently, the gaussian filter is used for removing the noise in EEG signal and then, signal segmentation and feature extraction are processed. Later, word recognition is carried out using DRN structurally optimized using the FHEBO. Here, the proposed FHEBO is developed by the combination of FC and HEBO. The HEBO is developed by the combination HLBO, and EBOA. Further, the Language modelling for speech recognition is carried out utilizing the GMM. Furthermore, the efficacy of the designed technique is analysed and the FHEBO-DRN obtained a maximal value of precision is 92.765%, F-measure of 94.241%, recall is 95.765% text conversion accuracy of 91.765% and minimal error rate is 8.235%. In future, advanced deep learning models will be developed for enhancing the performance further. In addition to this, EEG signals from other datasets can be considered to validate the generalizability of the approach.

REFERENCES

Afy, M., Siohan, O., & Sarikaya, R. (n.d.). GAUSSIAN MIXTURE LANGUAGE MODELS FOR SPEECH RECOGNITION.

Ahn, H.-J., & Lee, D.-H. (2021). Decoding 3D Representation of Visual Imagery EEG using Attention-based Dual-Stream Convolutional NeuralNetwork. http://arxiv.org/abs/2112.07148

Bhaladhare, P. R., & Jinwala, D. C. (2014). A Clustering Approach for the l -Diversity Model in Privacy Preserving Data Mining Using Fractional Calculus-Bacterial Foraging Optimization Algorithm . Advances in Computer Engineering, 2014, 1–12. https://doi.org/10.1155/2014/396529

Chen, Z., Chen, Y., Wu, L., Cheng, S., & Lin, P. (2019). Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions. Energy Conversion and Management, 198. https://doi.org/10.1016/j.enconman.2019.111793

Dehghani, M., & Trojovský, P. (2022). Hybrid leader based optimization: a new stochastic optimization algorithm for solving optimization applications. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-09514-0

GitHub - fwillett/handwritingBCI: Code from the paper "High-Performance Brain-to-Text Communication via Handwriting." (n.d.). Retrieved December 17, 2024, from https://github.com/fwillett/handwritingBCI

Gupta, A., Sahu, H., Nanecha, N., Kumar, P., Pratim Roy, P., & Chang, V. (n.d.). Enhancing text using emotion detected from EEG signals.

Hassan, A. R., Siuly, S., & Zhang, Y. (2016). Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating. Computer Methods and Programs in Biomedicine, 137, 247–259. https://doi.org/10.1016/j.cmpb.2016.09.008

Hong, K. S., Khan, M. J., & Hong, M. J. (2018). Feature Extraction and Classification Methods for Hybrid fNIRS-EEG Brain-Computer Interfaces. In Frontiers in Human Neuroscience (Vol. 12). Frontiers Media S.A. https://doi.org/10.3389/fnhum.2018.00246

Junwei, L., Ramkumar, S., Emayavaramban, G., Vinod, D. F., Thilagaraj, M., Muneeswaran, V., Pallikonda Rajasekaran, M., Venkataraman, V., & Hussein, A. F. (2019). Brain Computer Interface for Neurodegenerative Person Using Electroencephalogram. IEEE Access, 7, 2439–2452. https://doi.org/10.1109/ACCESS.2018.2886708

Kopparapu, S. K., & Satish, M. (2011). Identifying optimal Gaussian filter for Gaussian noise removal. Proceedings - 2011 3rd National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics, NCVPRIPG 2011, 126–129. https://doi.org/10.1109/NCVPRIPG.2011.34

Kumar, P., Saini, R., Roy, P. P., Sahu, P. K., & Dogra, D. P. (2018). Envisioned speech recognition using EEG sensors. Personal and Ubiquitous Computing, 22(1), 185–199. https://doi.org/10.1007/s00779-017-1083-4

Mannepalli, K., Sastry, P. N., & Suman, M. (2017). A novel Adaptive Fractional Deep Belief Networks for speaker emotion recognition. Alexandria Engineering Journal, 56(4), 485–497. https://doi.org/10.1016/j.aej.2016.09.002

Popescu, T. D. (2021). Signal segmentation using maximum a posteriori probability estimator with application in eeg data analysis. International Journal of Circuits, Systems and Signal Processing, 15, 1336–1345. https://doi.org/10.46300/9106.2021.15.144

Rajesh, S., Paul, V., Menon, V. G., Jacob, S., & Vinod, P. (2020). Secure Brain-to-Brain Communication With Edge Computing for Assisting Post-Stroke Paralyzed Patients. IEEE Internet of Things Journal, 7(4), 2531–2538. https://doi.org/10.1109/JIOT.2019.2951405

Trojovský, P., & Dehghani, M. (2022). A new optimization algorithm based on mimicking the voting process for leader selection. PeerJ Computer Science, 8, 1–40. https://doi.org/10.7717/peerj-cs.976

Ullah, S., & Halim, Z. (2021). Imagined character recognition through EEG signals using deep convolutional neural network. Medical and Biological Engineering and Computing, 59(5), 1167–1183. https://doi.org/10.1007/s11517-021-02368-0

Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M., & Shenoy, K. V. (2021). High-performance brain-to-text communication via handwriting. Nature, 593(7858), 249–254. https://doi.org/10.1038/s41586-021-03506-2

Zhang, X., Yao, L., Sheng, Q. Z., Kanhere, S. S., Gu, T., & Zhang, D. (2017). Converting Your Thoughts to Texts: Enabling Brain Typing via Deep Feature Learning of EEG Signals. http://arxiv.org/abs/1709.08820