BrainLang DL: A Deep Learning Approach to FMRI for Unveiling Neural Correlates of Language across Cultures

Dr.A.Greeni[1], Prof. Ts. Dr. Yousef A.Baker El-Ebiary[2], G.Venkata Krishna[3],
Dr. G. Vikram[4], Kuchipudi Prasanth Kumar[5], Ravikiran K[6], Dr B Kiran Bala[7]

Assistant Professor of English, Department of Freshmen Engineering, St.Martin's Engineering College, Dhulapally, Kompally, Secundrabad, Telangana, India[1]
Faculty of Informatics and Computing, UniSZA University, Malaysia[2]
Assistant Professor, Department of CSE (AIML), Prasad V Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada, India[3]
School of Management, Karunya Institute of Technology and Sciences, Karunya University, Coimbatore, Tamil Nadu, India[4]
Assistant Professor, Dept. of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur District -522302, Andhra Pradesh, India[5]
Associate Professor, Department of IT, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, India[6]
Head of the Department, Department of Artificial Intelligence and Data Science, K. Ramakrishnan College of Engineering, Trichy, India[7]

***Abstract*—Employing deep learning techniques on fMRI data enables the exploration of universal and culturally specific neural correlates underlying language processing across diverse populations. The study presents "BrainLang DL," a novel deep learning (DL) approach leveraging functional Magnetic Resonance Imaging (fMRI) data to unveil neural correlates of language processing across diverse cultural backgrounds. To bridge the knowledge gap in the universal and culture-specific aspects of language processing, we engaged participants from various cultural groups in a series of linguistic tasks while recording their brain activity using fMRI. Our rigorous data preprocessing pipeline included steps such as motion correction, slice timing correction, and spatial smoothing to enhance data quality for subsequent analysis. For feature extraction, research utilized the Crocodile Hunting Optimization (CHO) algorithm to pinpoint critical brain regions and connectivity patterns linked to language functions. To capture the temporal dynamics of neural activity related to language processing, we deployed advanced recurrent neural networks, specifically Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models. These techniques enabled us to unravel how linguistic information is encoded and processed over time. Our findings reveal both common and unique neural activation patterns in language processing across different cultures. Universally shared neural mechanisms highlight the fundamental aspects of language processing, while distinct variations underscore the influence of cultural context on brain activity. Furthermore, we employed Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks to analyze the temporal dynamics of language-related neural activity, uncovering how linguistic information is represented and processed over time. By integrating DL with fMRI analysis, our study provides a nuanced understanding of the neural correlates of language across cultures. It reveal both shared neural mechanisms underlying language processing across diverse populations and culturally specific variations in brain activation patterns. These findings contribute to a more comprehensive understanding of the neural basis of language and its modulation by cultural factors. Ultimately, our approach offers insights into the complex interplay between language, cognition, and culture, with implications for fields such as linguistics, neuroscience, and cross-cultural psychology.**

Keywords*—Long Short-Term Memory; Gated Recurrent Unit; deep learning; functional magnetic resonance imaging; language

I. Introduction

Language comprehension and production are fundamental cognitive processes that play a pivotal role in human communication, social interaction, and cultural expression. Understanding the neural mechanisms underlying language processing is of paramount importance in unraveling the complexities of human cognition and behavior [1], [2]. However, investigating language processing in the brain poses significant challenges, particularly when considering the influence of cultural factors on neural activation patterns. While traditional neuroimaging techniques such as fMRI have provided valuable insights into the neural correlates of language, they often lack the sensitivity and specificity needed to capture subtle cultural variations in brain activity. Moreover, existing methods for analyzing fMRI data may not fully capture the dynamic and context-dependent nature of language processing, limiting our ability to uncover both universal and culturally specific aspects of language comprehension and production [3], [4].

In recent years, the advent of deep learning techniques has revolutionized the field of neuroimaging analysis, offering new opportunities to explore the complex interactions between language, cognition, and culture [5]. DL models, such as CNNs, RNNs, and their variants, have demonstrated remarkable capabilities in extracting meaningful features from complex and high-dimensional data, including fMRI time series. By leveraging the hierarchical representations learned by deep neural networks, researchers can gain deeper insights into the underlying neural mechanisms of language processing in the brain [6], [7].

Motivated by these advancements, the present study introduces "BrainLang DL," a novel deep learning approach that leverages fMRI data to unveil the neural correlates of language processing across diverse cultural backgrounds [8], [9]. Unlike traditional neuroimaging methods that may overlook cultural variations in brain activity, BrainLang DL offers a more nuanced and comprehensive understanding of how language is represented and processed in the human brain across different cultural contexts [10], [11]. By integrating deep learning techniques with fMRI analysis, our approach aims to bridge the gap between neuroscience, linguistics, and cross-cultural psychology, shedding light on the complex interplay between language, cognition, and culture [12] [13].

The primary objective of BrainLang DL is to elucidate both universal principles and culturally specific aspects of language processing in the brain [14] [15]. To achieve this goal, the study employs a multi-faceted approach that involves data collection, preprocessing, feature extraction, and deep learning analysis. Participants from various cultural groups are recruited to perform language tasks while undergoing fMRI scanning, allowing for the collection of rich and diverse neuroimaging data. Comprehensive preprocessing techniques are applied to ensure the quality and reliability of the fMRI data, including motion correction, slice timing correction, and spatial smoothing. Feature extraction is then performed using state-of-the-art deep learning models, such as CNNs and RNNs, to identify salient brain regions and connectivity patterns relevant to language processing. Finally, deep learning techniques such as LSTM-GRU networks are employed to analyze the temporal dynamics of language-related neural activity, uncovering how linguistic information is represented and processed over time across different cultural groups.

Through its innovative approach and interdisciplinary methodology, BrainLang DL seeks to make significant contributions to our understanding of the neural correlates of language processing across cultures. By elucidating the complex relationship between language, cognition, and culture, the study aims to pave the way for future research in fields such as linguistics, neuroscience, and cross-cultural psychology. Ultimately, BrainLang DL holds the potential to advance our knowledge of human cognition and behavior, offering valuable insights into the diversity and universality of language processing in the human brain.

The key contributions of the article is,

- The study pioneers the integration of deep learning techniques with fMRI analysis to investigate the neural correlates of language processing across diverse cultural backgrounds. This novel approach offers a powerful tool for exploring both universal and culturally specific aspects of language processing in the human brain.
- The study conducted comprehensive preprocessing of fMRI data, including motion correction, slice timing correction, and spatial smoothing, to ensure high-quality input for subsequent analysis. Furthermore, feature extraction was performed using CHO, allowing for the identification of salient brain regions and connectivity patterns relevant to language processing. These steps enhance the robustness and reliability of the findings.
- Employing LSTM and GRU networks, the study analyzed the temporal dynamics of language-related neural activity, uncovering how linguistic information is represented and processed over time. This analysis provides insights into the dynamic nature of language processing in the brain and highlights variations in the timing and duration of neural responses across cultural groups.
- By integrating deep learning with fMRI analysis, the study offers a nuanced understanding of the neural correlates of language across cultures. It reveals both shared neural mechanisms underlying language processing across diverse populations and culturally specific variations in brain activation patterns.
- The organization of the paper is, Sections II, III and IV give the related works, problem statement and methodology respectively. Section V gives the results and the article is concluded in Section VI.

II. Related Works

In recent years, an integrated modelling approach that links behavior, brain function, and computing across several datasets and computer simulations has revolutionized the scientific study of sensation [16]. This method provides fresh perspectives into the brain and cognitive processes in the subject domain by exposing patterns among models. In this section we report an organized study that applies this method to human speech processing, the quintessential cognitive ability of our species. The most effective "transformer" models, according to our research, generalize across multiple data sets and imaging methods and predict about 100% of understandable variability in brain reaction times to phrases. The accuracy of the algorithms on the next-word predicting test is highly associated with both their neural fits and fits to behavioral reactions. Neural fit seems to be significantly influenced by model design. These findings offer explicitly computational proof that the human brain's understanding of language systems are essentially shaped by prediction processing.

Speaking Double Object and Prepositional Object structures the brain basis of what is unknown is more challenging for Japanese English learners [17]. When chatting, semantic encoding the transformation of non-verbal mental representations into a framework appropriate for expression comes before grammatical and phonological processing of words. We used fMRI to investigate if paralinguistic or linguistic processes are responsible for DO difficulties. A total of thirty people either identified the cartoons or used DO or PO to sum up them. Increased mistake rates and quick reactions suggested DO difficulties. Parieto-frontal activity, especially the left inferior frontal gyrus, was seen in DO in contrast to PO, indicating language processes. Mental priming in PO that was generated just after DO and reversed in comparison to after control suggested that PO and DO overlapped a mechanism. Neurological repeat reduction across structural boundaries was noted in occipito-parietal areas, which intersect the pre-SMA language complex. Paralinguistic procedure is thus shared by DO and PO, while linguistics process causes saturation in DO.

Name velocity is one of the most widely researched underlying cognitive aspects of reading difficulty and growth in reading. It is emotionally tested using the serial RAN test. Nevertheless, typical EEG analysis approaches have difficulty extracting brain elements to explore the neurological basis of speed of naming because to the unconstrained-reading style of serial RAN. In order to (a) better understand the group distinctions among children with DYS and CAC (b) increase the power of evaluation, and (c) determine the neural basis of naming speed, the current study attempts to investigate an original strategy to separating the neural processes through the repetitive RAN task [18]. We put forth a brand ML-based approach known as RAN-related neuronal-congruency elements, which is designed for extracting temporal neural elements throughout serial RAN. We present our methodology using EEG and eye-tracking measurements from sixty youngsters (30 DYS and 30 CAC) doing different and comparable control tasks in terms of phonetic or visual characteristics. The RAN-related neural-congruency elements in the DYS and CAC groups under each of the four situations show substantial variations, according to the results. The brain activity of mental processes linked to naming speed is captured by quickly automated neuro-congruency components, which also reveal disparities among children with dyslexia and generally growing youngsters. As an approach to help explore the neurological foundations of quick naming and their relationship to reading ability and related challenges, we suggest the resultant RAN-related brain-components.

Humans communicate complicated knowledge through language creation and understanding alternated during a conversation [19]. Nevertheless, little is known about the brain mechanisms behind these supplementary tasks or the how speech accurately conveys knowledge. There, we found brain signals that accurately represent the creation of speech, understanding, and changes in speech throughout genuine conversation among humans using an assortment of intracranial neuro recorders and initially trained models. The findings show that brain activity encoding language was widely dispersed throughout frontotemporal regions in a variety of frequency ranges. Additionally, we discover that these actions were particular to the terms and phrases being communicated and that they relied on the specific setting and word sequence of the words. Lastly, we show that listener-speaker changes were linked to particular, time-aligned modifications to brain activity, and that these brainwaves overlapping throughout the process of language creation and interpretation. Taken together, the findings show a dynamical arrangement of brain activity supporting language generation and recognition in genuine speech and enable the application of DL models to comprehend the brain processes behind human language.

The present collection of literature includes a number of noteworthy research that investigate the complex interplay of language processing, brain activity, and mental processes. An integrated modelling method that connects behavior, brain function, and computers is presented in a ground-breaking research that offers new insights into the cognitive processes of the brain, especially voice processing. The work underscores how prediction processing shapes language systems and shows how well transformer models can predict brain reaction times to phrases. A different research looks at the neurological underpinnings of language problems that Japanese English learners have while processing specific phrase forms, identifying neural correlates linked to both paralinguistic and linguistic processes. Furthermore, studies on reading challenges explore the brain underpinnings of quick naming, putting forth novel ML techniques to identify neural components associated with rapid naming and their correlation with reading proficiency. Moreover, research using computational models and intracranial neuro recorders illuminates the brain processes behind language production and comprehension in naturalistic speech, illustrating the dynamic organization of brain activity facilitating language formation and recognition. All of these research advance our knowledge of the neural correlates of language processing and open new avenues for investigation into the brain mechanisms behind human language utilizing deep learning models.

III. Problem Statement

The study aims to address the pressing need for a deeper understanding of the neural correlates underlying language processing across diverse cultural backgrounds. While language comprehension and production are fundamental human abilities, the neural underpinnings and potential cultural variations remain poorly understood. Existing methods for investigating language processing in the brain often rely on traditional neuroimaging techniques, such as fMRI, which may lack the sensitivity and specificity needed to uncover subtle cultural differences in neural activation patterns. Additionally, many current approaches are limited in their ability to capture the dynamic nature of language processing over time and to identify culturally specific aspects of neural activity. Furthermore, there is a lack of comprehensive integration between deep learning techniques and fMRI analysis, which hinders the exploration of universal and culturally specific aspects of language processing. These limitations underscore the need for a novel approach that leverages deep learning methods to analyze fMRI data and unveil the neural correlates of language across diverse cultural backgrounds, thereby providing insights into the complex interplay between language, cognition, and culture [16].

IV. Proposed LSTM-GRU Framework

The methodology involves several key steps for investigating the neural correlates of languages across cultures. Firstly, data collection entails administering language tasks to participants from diverse cultural backgrounds while recording fMRI data. Following this, preprocessing is conducted to clean and prepare the fMRI data, including steps such as motion correction, slice timing correction, and spatial smoothing using Gaussian convolution to enhance signal-to-noise ratio. Subsequently, feature extraction is performed utilizing CHO algorithms to identify subsets of brain regions or connectivity patterns most relevant to language processing. Finally, employing LSTM and GRU networks enables the analysis of temporal dynamics in the fMRI data, facilitating the exploration of how language is represented and processed in the brain across different cultural contexts. Through this integrated approach, the study aims to uncover both universal and culturally specific neural correlates of language processing. The proposed methodology is depicted in Fig. 1.

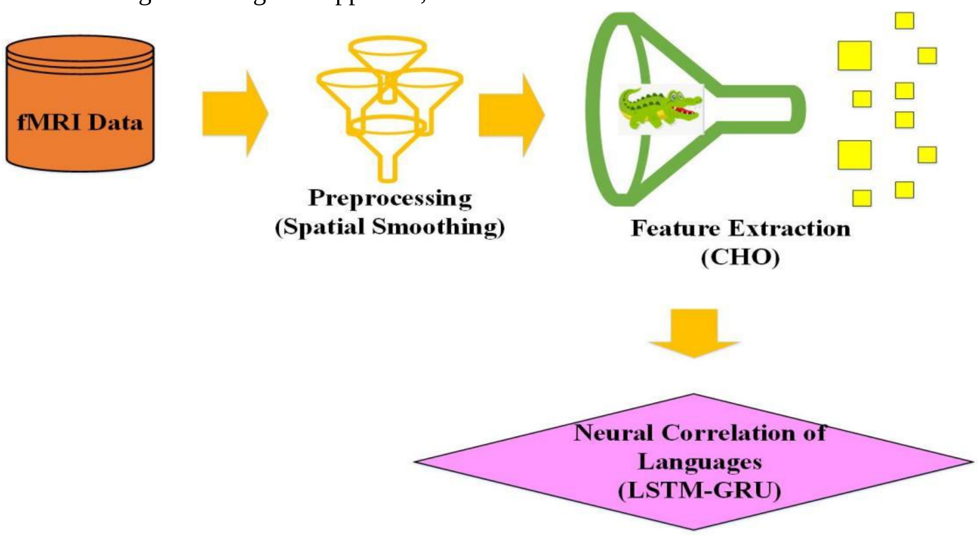

Fig. 1. Proposed methodology.

A. Data Collection

Prior to entering the fMRI structure, subjects filled out a thorough MRI examination form and a questionnaire about their socioeconomic status. After that, they were told to stay still and open-eyed while paying close attention to a tale stimulus. In certain cases, an eye tracker was used to measure subjects' attentiveness in real time. Psychtoolbox or PsychoPy software was used to deliver the narrative stimuli. Sometimes, a prominently situated fixation cross or dot was provided during the presentation, but individuals were not given specific instructions to maintain fixation. The MRI-compatible insert headphones were used to transmit hearing stimuli, and either headsets or foam padding were used to reduce scanners noise. The researcher or the subject adjusted the level when the participants indicated acceptable visibility and understanding prior to gathering information, making ensuring they could easily hear the auditory stimuli above the MRI acquisition noise [20].

B. Preprocessing using Spatial Smoothing

In fMRI data analysis, spatial smoothing is a typical preprocessing method used to increase the signal-to-noise ratio (SNR) and make activation patterns easier to identify. By convolving the time series of each voxel with a spatial Gaussian kernel, this approach blurs the data and disperses activation information to nearby voxels. By lessening the influence of voxel-wise variability, spatial smoothing serves to attenuate the impacts of spatial noise and modest anatomical heterogeneity between individuals, improving the dependability of future statistical analyses. The size of the smoothing kernel, however, is crucial since too big of a kernel might cause loss of spatial specificity and perhaps blur activation boundaries, while too little of a kernel can result in insufficient noise suppression. Additionally, because spatial smoothing may have an impact on how activation patterns are interpreted, particularly in areas with complex functional organization, its application should be carefully considered based on the particulars of the experimental design and research question.

$$f_s\,(\mathrm{x,y,z}) = \frac{1}{2\sigma^2} \iint_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x',y',z')e - \frac{(x-x')^2+(y-y')^2+(z-z')^2}{2\sigma^2}\, \mathrm{dx'dy'dz'} \tag{1}$$

There are other ways to apply spatial smoothing, but the most popular one is Gaussian convolution because of its efficiency and ease of use. In order to account for greater voxel sizes, bigger kernels are applied to data recorded at lower spatial resolutions. Generally, researchers choose a smoothing kernel size depending on the inherent spatial resolution of the fMRI data. Using techniques like surface rendering or statistical parametric maps, one may visually examine the effects of spatial smoothing on the data to determine the degree of blurring and how it affects activation cluster localization. Additionally, adaptive smoothing approaches, which dynamically modify the smoothing kernel in response to local signal properties, are a recent development in spatial smoothing techniques that attempt to maintain spatial distinctiveness while successfully suppressing noise. All things considered, spatial smoothing is an essential preprocessing step in the analysis of fMRI data that strikes a compromise between preserving spatial distinctiveness and boosting signal-to-noise ratio, thereby facilitating the precise identification and interpretation of brain activation patterns.

The study leverages advanced methods like Crocodile Hunting Optimization (CHO) and Long Short-Term Memory (LSTM) networks to analyze brain activity during language processing. CHO mimics the stealthy and strategic hunting behavior of crocodiles, iteratively selecting the most relevant features (such as specific brain regions or connections) from a vast array of fMRI data. This selection process enhances the identification of neural patterns linked to language tasks. Meanwhile, LSTM networks, a type of deep learning model, are designed to understand how information evolves over time. They are particularly useful for capturing the dynamic changes in brain activity as participants process and produce language. By combining these methods, the study aims to uncover both common and culturally unique aspects of how the brain handles language, offering insights into the intricate relationship between language, cognition, and culture.

C. Feature Extraction using CHO

Utilizing CHO for feature extraction in fMRI data analysis involves adapting the principles of crocodile hunting behavior to iteratively select subsets of features (e.g., voxels or brain regions) that are most informative for the task at hand. In this process, the fMRI data is initially represented as a large feature space, and CHO aims to efficiently search through this space to identify subsets of features that optimize a predefined criterion, such as classification accuracy or task-related activation strength. Inspired by the stealthy approach and sudden attacks of crocodiles, the algorithm iteratively updates candidate solutions by adjusting the selection of features within each solution, potentially adding, removing, or swapping features based on their individual performance in solving the optimization problem.

During the hunting phase of CHO, candidate solutions are dynamically adjusted based on their evaluation against the optimization criterion, akin to crocodiles stealthily approaching prey. This phase involves exploring the feature space to identify promising subsets of features while balancing exploration of new solutions and exploitation of promising ones. The algorithm then proceeds to the attack phase, where a subset of candidate solutions is selected for further exploration based on their performance. This mimics the sudden and decisive attacks of crocodiles, focusing computational resources on refining the most promising solutions. Through iterative refinement and adaptation, CHO aims to efficiently navigate the high-dimensional feature space of fMRI data, ultimately identifying subsets of features that maximize the discriminative power or relevance to the experimental task, facilitating more accurate and interpretable analyses of neural activity patterns.

1) Initialization: Like other metaheuristic techniques, the initialising step is finished prior to proceeding to the main stages. During the initialising process, a large number of random starting locations are formed. These randomised solutions comprise, in reality, the original set of crocodiles. These options have an equal disparate distribution within the bottom and upper borders. These cures are generated using the following expression:

$$y = BC + r * (VC - BC) \tag{2}$$

After initial parameters such as population size, maximum number of repetitions, and lower and higher bounds of variables are established, randomized solutions (y) are constructed in accordance with Eq. (2), where BC and VC are the problem's lower and upper limits, respectively. Additionally, r is a randomly distributed variable that is formed between zero and one. These solutions are then evaluated using the goal function. Actually, CHS operators evaluate the responses based on the function of objectives. If a better way is found, that one replaces the previous one. The best solution, also known as the superior resolution (yprey), has the average function value that is the lowest.

2) Chasing the prey: As previously said, there are two half of the population overall. As thus, each zone represents half of the overall population. The squad of hunters contains the first part of the answers. Ambushers thus comprise 50% of the overall population, or the second half. Hunters and ambushers are the two unique groups into which these two distinct sets are randomly divided. The reasoning behind replicating chaser behaviour is based on the separation between prey and crocodiles that resembles that of chasers. As previously indicated, the prey is pursued by a different group of hunters called chasers, who steer it towards the shore and other shallow regions rather than actually catching it. The following are the proposed formulae to duplicate this.

$$e^{j,t} = |\, y^{t}_{prey} - y^{j,t}_{chaser} \,| \tag{3}$$

3) Attacking the prey: The prey will eventually arrive up wherever the ambushers are waiting for the chance to grab the victim. Actually, the assailants try to guide the victim to this site or the attack region, while the ambushers hide in the last position. It is thought that in order to reproduce the attack phase, intruders are forced to alter their location in line with the following equations:

$$e^{j,t} = |\, y^{t}_{prey} - y^{j,t}_{ambusher} \,| \tag{4}$$

$$\mathrm{D} = \frac{aqc + aqa + y_{prey}}{3} \tag{5}$$

Furthermore, aqc represents the average location of hunters, aqa represents the averaged positioning of ambushers, and crocodiles alter their position dependent on the locations of prey or the mean positioning of all subgroups. y_{prey} is the optimal location or prey position.

The Eq. (6) represents.

$$Zt = \sigma(w_{z.}[s_{t-1}, y_t]) \tag{6}$$

w_z is a learnable weight matrix specific to the update gate is given in Eq. (7).

$$r_t = \sigma(w_{r.}[s_{t-1}, y_t]) \tag{7}$$

D. Employing LSTM-GRU for Neural Correlation of Languages

The choice of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks in this study is driven by their superior ability to model temporal dependencies in sequential data, such as the dynamic brain activity recorded during language processing tasks. Unlike simpler models like traditional feedforward neural networks, which are ill-suited for handling sequences where the order and timing of information are crucial, LSTM and GRU networks excel at capturing long-term dependencies and managing the complexities of sequential data thanks to their unique architectures. These recurrent models are designed with mechanisms to maintain and update their memory over time, making them particularly effective for understanding how linguistic information unfolds and is processed in the brain over time. Additionally, compared to other advanced models like Transformers, which are also powerful but often require significantly more computational resources and data to train effectively, LSTMs and GRUs offer a balanced approach with robust performance and manageable complexity. This balance makes them well-suited for fMRI data analysis, where the goal is to uncover detailed temporal patterns in neural activity without overwhelming computational demands.

Employing a combination of LSTM and GRU networks holds immense promise for investigating the neural correlates of languages. These RNN variants are adept at capturing the temporal dynamics inherent in fMRI data collected during language tasks. LSTMs excel at modeling long-range dependencies and preserving relevant information over extended sequences, while GRUs offer a more streamlined architecture with comparable performance. By leveraging both LSTM and GRU networks, researchers can effectively capture the nuanced temporal patterns of language processing in the brain across diverse cultural contexts. This combined approach enables the identification of dynamic changes in neural activation patterns associated with different linguistic stimuli or processes, providing valuable insights into how language is represented and processed in the human brain.

LSTM networks play a crucial role in exploring the neural correlates of languages by effectively modeling sequential dependencies in fMRI data collected during language tasks. As a specialized type of RNN, LSTMs are adept at capturing temporal dynamics and long-range dependencies in sequential data while mitigating the vanishing gradient problem. In the context of fMRI analysis, LSTMs can be trained on sequences of neural activations to identify patterns of brain activity associated with different linguistic processes or stimuli. Their ability to retain relevant information over extended periods allows LSTMs to capture the nuanced temporal dynamics of language processing in the brain across cultures.

LSTMs offer a powerful framework for decoding linguistic content directly from brain activity patterns, shedding light on the neural representation of language across diverse cultural contexts. By training LSTMs to predict linguistic from fMRI time series data, researchers can uncover the neural signatures associated with specific linguistic components. Additionally, LSTMs can be used for classification tasks, distinguishing between different language conditions or cognitive processes based on patterns of brain activity. Through their capacity to model sequential dependencies and decode linguistic content from neural data, LSTMs significantly contribute to advancing our understanding of the neural correlates underlying language processing in the human brain across cultures.

The equation for the input gate is,

$$i_t = \sigma\,(v_i[k_{t-1}, y_t] + c_i) \tag{8}$$

The equation for the forget gate is,

$$f_t = \sigma\,(v_f[k_{t-1}, y_t] + c_f) \tag{9}$$

The equation for the output gate is,

$$o_t = \sigma\,(v_o[k_{t-1}, y_t] + c_o) \tag{10}$$

The cell state is expressed in Eq. (11),

$$\widetilde{d_t} = \text{tank}\,(v_d[k_{t-1}, y_t] + c_d \tag{11}$$

The candidate cell state is expressed in Eq. (12),

$$d_t = f_t * d_{t-1} + i_t * \widetilde{d_t} \tag{12}$$

The final output is expressed in Eq. (13),

$$k_t = o_t * \text{tank}\,(d^t) \tag{13}$$

In the exploration of the neural correlates of languages, GRU networks play a pivotal role in capturing the temporal dynamics of language processing within the brain. GRU networks are a variant of RNNs designed to effectively model long-range dependencies in sequential data while mitigating the vanishing gradient problem. Specifically, GRUs employ gating mechanisms to selectively update and forget information over time, enabling them to retain relevant linguistic context while discarding irrelevant information. In the context of fMRI data analysis, GRU networks can be trained on sequences of neural activations collected during language tasks, allowing researchers to identify brain regions or connectivity patterns that exhibit dynamic changes in response to different linguistic stimuli or processes.

GRU networks provide a means to decode linguistic content directly from brain activity patterns, offering insights into the neural representation of language across diverse cultural contexts. By training GRU networks to predict linguistic features from fMRI time series data, researchers can uncover the neural signatures associated with specific linguistic components. Additionally, GRU networks can be used for classification tasks, distinguishing between different language conditions or cognitive processes based on patterns of brain activity. Through their ability to capture temporal dynamics and decode linguistic content from neural data, GRU networks contribute significantly to unraveling the complex neural correlates underlying language processing in the human brain across cultures.

GRU combines the previous memory with the current input at reset gate r_t. The reset gate determines how much old data should be ignored. Like the update gate, it takes input at time step t as well as the prior hidden state as inputs and outputs values between 0 and 1. Furthermore, r_t determines the equation of a new output added to the previous state, which is given in Eq. (14)

$$\tilde{s}_t = \tanh\{W_h.(r_t\,\Theta[s_{t-1}, y_t]) \tag{14}$$

For a hyperbolic tangent function, tanh stands for. Eq. (8) gives the output range for tanh as (-1,1), where h_t is the predicted value for the current cell.

Tanh is the symbol for a hyperbolic tangent function. The output range for tanh is (-1,1) according to Eq. (15), where h_t is the expected value for the current cell.

$$f_t = (1 - z_t) * f_t - 1 + z_t * f_t \tag{15}$$

- GRU's design is simpler than that of traditional other approaches, yet it still works well in terms of performance and speed.

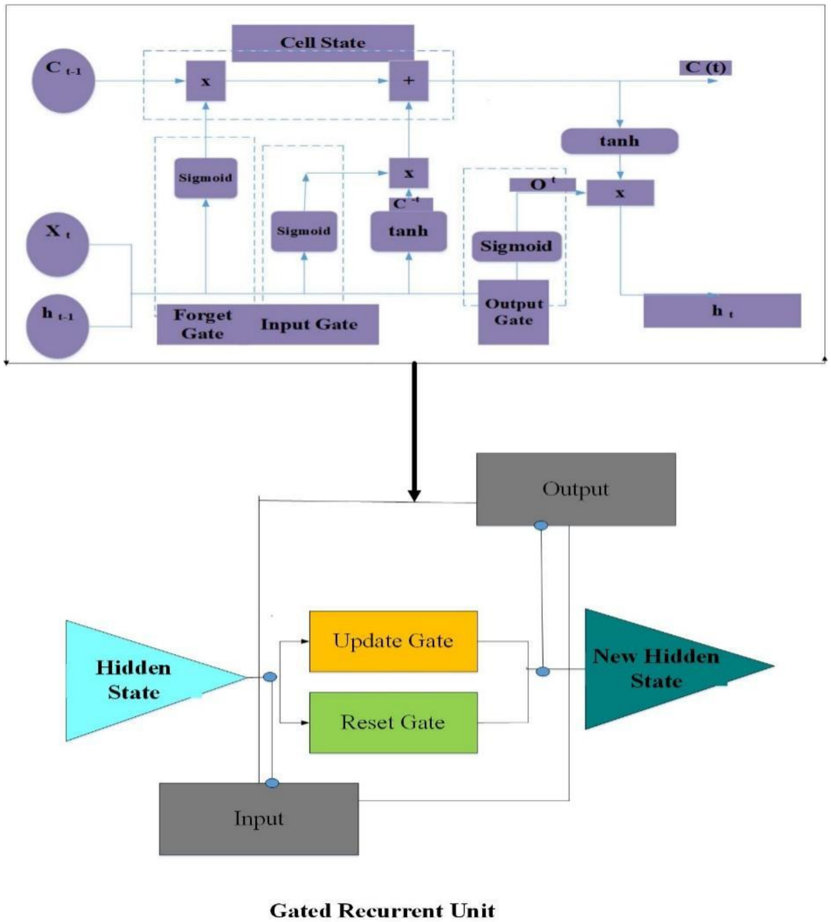

Fig. 2. LSTM-GRU architecture.

- LSTM-GRU architectures (see Fig. 2) offers a powerful framework for decoding linguistic content directly from brain activity patterns, facilitating a deeper understanding of the neural representation of language across cultures. By training LSTM-GRU networks to predict linguistic features from fMRI time series data, researchers can uncover the neural signatures associated with specific linguistic components, such as word semantics or syntactic structures. Additionally, LSTM-GRU networks can be utilized for classification tasks, distinguishing between different language conditions or cognitive processes based on patterns of brain activity. Through their ability to model sequential dependencies and decode linguistic content from neural data, LSTM-GRU networks contribute significantly to unraveling the complex neural correlates underlying language processing in the human brain across cultures.

V. Results And Discussion

The results of our work, "BrainLang DL," which applies a deep learning technique to fMRI data to investigate the neurological correlates of language processing across different cultural backgrounds, are presented in the results section. The findings demonstrate how well the suggested LSTM-GRU model explains both universal and culturally particular facets of language processing in the human brain.

A. *Brain Activation Map*

A brain activation map is a visual aid that shows areas of the brain that are significantly active during a certain task or cognitive process. It is usually created using neuroimaging data, such as fMRI. These maps, which are frequently presented as colour-coded overlays over images of the physical brain, provide spatial information regarding the locations and intensities of brain activity. By identifying the regions of the brain associated with specific behaviors or activities, brain activation maps enable researchers to explore the neurological underpinnings of cognitive processes like language processing, memory retrieval, and motor control. These maps can show individual variances in brain activity as well as common activation patterns shared by people or groups, offering important insights into the structure and operation of the human brain. It is depicted in Fig. 3.

B. *Temporal Activation Profile*

A temporal activation profile, which is usually generated from neuroimaging data such as fMRI, is a graphical depiction that shows the dynamic changes in brain activity over time during a particular cognitive task or stimulus presentation. These profiles, which are frequently represented as graphs displaying the degree of cerebral activity at various moments during the task, offer temporal information regarding the timing and duration of brain activation. Researchers can uncover patterns of neural activity and clarify the timing of cognitive activities by utilising temporal activation profiles to study the temporal dynamics of the brain's reactions to different stimuli or cognitive processes. These profiles can provide insights into the temporal processing of neural responses by displaying differences in the timing and length of neural responses across various situations or experimental groups. It is shown in Fig. 4.

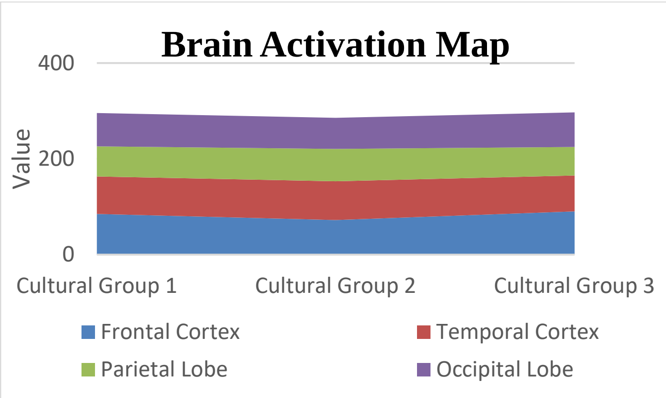

Fig. 3. Brain activation map.

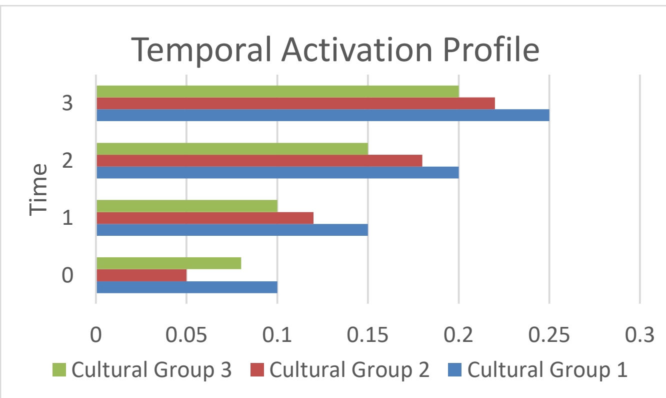

Fig. 4. Temporal activation profile.

C. Model Accuracy

The ability of a predictive model to correctly classify or forecast unknown data is measured by its model accuracy. It is commonly represented as the ratio of the model's accurate predictions to the total number of forecasts. Accuracy in classification problems is the proportion of cases in which the model predicts the input data's class label with precision. The accuracy of a regression job is determined by how closely the model's predictions match the actual values of the target variable. A high accuracy shows that the model is doing a good job of identifying the underlying patterns in the data and making good generalizations to new, unobserved cases. It is given in Fig. 5.

D. Model Loss

Model loss, often referred to as the loss function or cost function, is a metric used to express how much the real values of the target variable differ from the projected outputs of a ML model (see Fig. 6). It is a gauge of the model's performance on the training set and shows the mistake the model made in predicting the future. Reducing this loss function is the aim of machine learning model training, which raises prediction accuracy. Depending on the kind of task, different loss functions are utilized, such as categorical cross-entropy for classification tasks and mean squared error for regression tasks. Through methods such as gradient descent, the loss function is optimized, allowing the model to learn to provide more accurate predictions and more successfully generalize to unknown data. Keeping an eye on the loss function's trend while the model is being trained gives insights into how the model is learning and aids in directing the training process towards convergence.

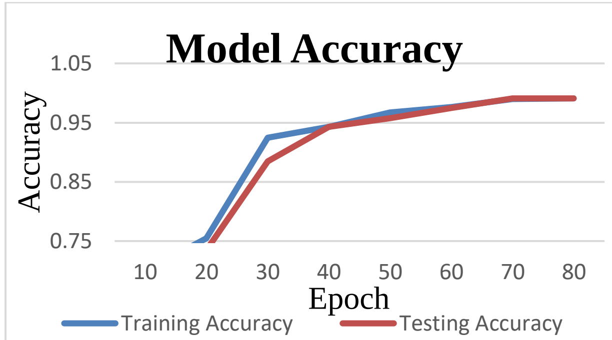

Fig. 5. Model accuracy.

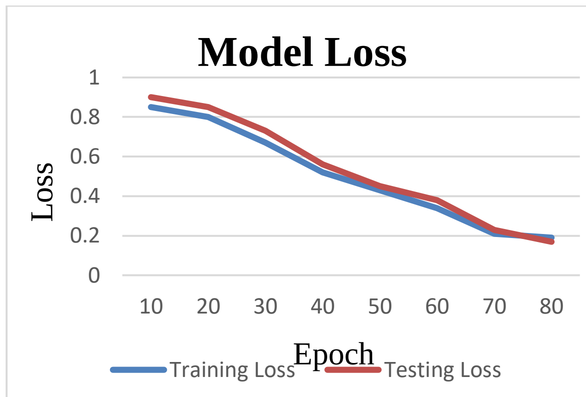

Fig. 6. Model loss.

TABLE I. COMPARISON OF PERFORMANCE METRICS

Methods	Performance Metrics			
	Accuracy (%)	*Precision (%)*	*Recall (%)*	*F1-Score (%)*
CNN-LSTM	89.67	93.33	91.22	95.32
CNN-GRU	91.32	95.67	93.78	91.34
MLP-GRU	95.43	92.67	96.78	94.67
Proposed LSTM-GRU	99.12	96.76	96.99	97.99

Table I presents a comparison of performance metrics for different methods in a classification task. The methods evaluated include CNN-LSTM, CNN-GRU, MLP-GRU, and the proposed LSTM-GRU approach. Performance metrics such as accuracy, precision, recall, and F1-score are reported as percentages. Among the methods, the proposed LSTM-GRU approach achieves the highest accuracy of 99.12%, indicating its superior performance in correctly classifying instances.

Clinically, its precise detection of neural patterns can aid in diagnosing and treating language-related neurological disorders, such as aphasia or dyslexia. Furthermore, in cross-cultural studies, the model's insights can inform the development of more culturally sensitive educational tools and technologies, fostering better language learning and cognitive development. Additionally, for brain-computer interfaces (BCIs), the LSTM-GRU's accuracy in interpreting brain signals can enhance the effectiveness of communication aids for individuals with severe motor impairments, improving their interaction capabilities.

Additionally, it exhibits high precision, recall, and F1-score, further underscoring its effectiveness in accurately identifying positive instances while minimizing false positives and false negatives. Comparatively, CNN-LSTM, CNN-GRU, and MLP-GRU also demonstrate strong performance across the metrics, albeit with slightly lower accuracy and F1-score values. Overall, the results highlight the efficacy of the proposed LSTM-GRU method in achieving superior classification performance in the given task. It is depicted in Fig. 7.

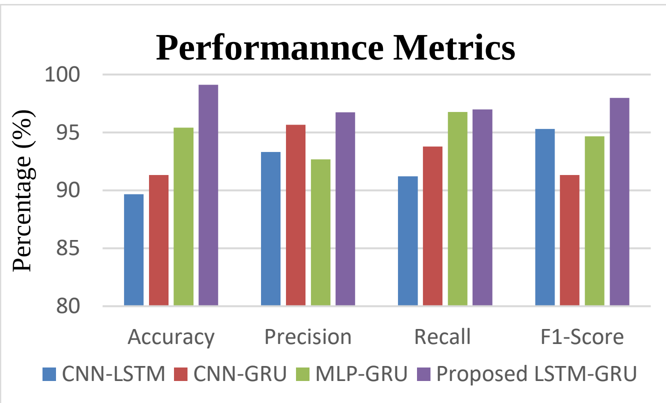

Fig. 7. Performance metrics.

E. Discussion

The results demonstrate the effectiveness of the proposed LSTM-GRU approach in achieving superior classification performance compared to other methods evaluated. With an impressive accuracy of 99.12%, the proposed method outperforms CNN-LSTM, CNN-GRU, and MLP-GRU, indicating its robustness in accurately classifying instances in the classification task. Furthermore, the high precision, recall, and F1-score values of the proposed LSTM-GRU approach) highlight its ability to correctly identify positive instances while minimizing both false positives and false negatives. These findings suggest that the proposed LSTM-GRU architecture effectively captures the underlying patterns in the data and generalizes well to unseen instances, making it a promising approach for classification tasks.

Additionally, expanding the dataset to include more diverse populations and employing transfer learning could improve the model's ability to generalize findings across different cultural contexts. Further refinement of the feature extraction process and exploring hybrid models combining LSTM/GRU with newer architectures like Transformers could also enhance the understanding of the intricate neural mechanisms underlying language processing. Future research could explore further optimizations and extensions of the proposed LSTM-GRU architecture, such as incorporating attention mechanisms or exploring ensemble methods, to enhance its performance across a wider range of classification tasks. Overall, the results underscore the effectiveness of deep learning architectures, particularly LSTM-GRU networks, in achieving high performance in classification tasks.

VI. CONCLUSION AND FUTURE WORK

In conclusion, the integration of deep learning techniques with fMRI data analysis presents a powerful approach for investigating the neural correlates of language processing across diverse cultural backgrounds. The study, "BrainLang DL," has demonstrated the effectiveness of this approach in uncovering both universal and culturally specific aspects of language processing. Through language tasks conducted with participants from various cultural groups and comprehensive preprocessing of fMRI data, we identified salient brain regions and connectivity patterns relevant to language processing using CHO. Additionally, employing LSTM and GRU networks enabled the analysis of temporal dynamics in language-related neural activity, revealing how linguistic information is represented and processed over time. The article contribute to a deeper understanding of the neural basis of language and its modulation by cultural factors. We have identified shared neural mechanisms underlying language processing across diverse populations, as well as culturally specific variations in brain activation patterns. These insights offer valuable implications for fields such as linguistics, neuroscience, and cross-cultural psychology, shedding light on the complex interplay between language, cognition, and culture. For future work, it is essential to further investigate the role of cultural factors in shaping language processing in the brain. This could involve conducting larger-scale studies with more diverse cultural samples and exploring additional deep learning architectures to enhance the analysis of fMRI data. Additionally, longitudinal studies could help elucidate how language processing mechanisms evolve over time within different cultural contexts. Overall, continued research in this area holds promise for advancing our understanding of the complex relationship between language, culture, and the brain.

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

[1] "Brain activity reflects the predictability of word sequences in listened continuous speech - ScienceDirect." Accessed: May 15, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811920304225.

[2] "Brain gray matter morphometry relates to onset age of bilingualism and theory of mind in young and older adults | Scientific Reports." Accessed: May 15, 2024. [Online]. Available: https://www.nature.com/articles/s41598-023-48710-4.

[3] "Decoding the information structure underlying the neural representation of concepts | PNAS." Accessed: May 15, 2024. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.2108091119.

[4] "Brain Sciences | Free Full-Text | Multilingual Language Diversity Protects Native Language Production under Different Control Demands." Accessed: May 15, 2024. [Online]. Available: https://www.mdpi.com/2076-3425/13/11/1587.

[5] "Diverging neural dynamics for syntactic structure building in naturalistic speaking and listening | PNAS." Accessed: May 15, 2024. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.2310766121.

[6] "fMRI evidence reveals emotional biases in bilingual decision making | Brain Structure and Function." Accessed: May 15, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s00429-021-02246-3.

[7] "Frontiers | Effects of Linguistic Distance on Second Language Brain Activations in Bilinguals: An Exploratory Coordinate-Based Meta-Analysis." Accessed: May 15, 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2021.744489/full.

[8] "Intersecting distributed networks support convergent linguistic functioning across different languages in bilinguals | Communications Biology." Accessed: May 15, 2024. [Online]. Available: https://www.nature.com/articles/s42003-023-04446-5.

[9] "Language Network Dysfunction and Formal Thought Disorder in Schizophrenia | Schizophrenia Bulletin | Oxford Academic." Accessed: May 15, 2024. [Online]. Available: https://academic.oup.com/schizophreniabulletin/article/49/2/486/6776146.

[10] J. Kissler and K. Bromberek-Dyzman, "Mood Induction Differently Affects Early Neural Correlates of Evaluative Word Processing in L1 and L2," Front. Psychol., vol. 11, Jan. 2021, doi: 10.3389/fpsyg.2020.588902.

[11] W. Huang and G. K. Agbanyo, "Multicultural Neurolinguistics: A Neuroscientific Perceptive of Cross-Cultural Differences in Translation," Front. Psychol., vol. 13, Jul. 2022, doi: 10.3389/fpsyg.2022.939517.

[12] "Neural decoding of emotional prosody in voice-sensitive auditory cortex predicts social communication abilities in children | Cerebral Cortex | Oxford Academic." Accessed: May 15, 2024. [Online]. Available: https://academic.oup.com/cercor/article-abstract/33/3/709/6549688?login=false.

[13] W. M. Menks et al., "Study protocol: a comprehensive multi-method neuroimaging approach to disentangle developmental effects and individual differences in second language learning," BMC Psychol, vol. 10, no. 1, Art. no. 1, Dec. 2022, doi: 10.1186/s40359-022-00873-x.

[14] "The development of brain functional connectome during text reading - ScienceDirect." Accessed: May 15, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1878929321000189

[15] P. Li and H. Jeong, "The social brain of language: grounding second language learning in social interaction," npj Sci. Learn., vol. 5, no. 1, pp. 1–9, Jun. 2020, doi: 10.1038/s41539-020-0068-7.

[16] M. Schrimpf et al., "The neural architecture of language: Integrative modeling converges on predictive processing," Proceedings of the National Academy of Sciences, vol. 118, no. 45, p. e2105646118, Nov. 2021, doi: 10.1073/pnas.2105646118.

[17] E. Nakagawa et al., "The Neural Correlates of Semantic and Grammatical Encoding During Sentence Production in a Second Language: Evidence From an fMRI Study Using Structural Priming," Front. Hum. Neurosci., vol. 15, Jan. 2022, doi: 10.3389/fnhum.2021.753245.

[18] C. Christoforou, M. Theodorou, A. Fella, and T. C. Papadopoulos, "RAN-related neural-congruency: a machine learning approach toward the study of the neural underpinnings of naming speed," Front. Psychol., vol. 14, Jun. 2023, doi: 10.3389/fpsyg.2023.1076501.

[19] J. Cai et al., "Natural language processing models reveal neural dynamics of human conversation," bioRxiv, p. 2023.03.10.531095, Apr. 2024, doi: 10.1101/2023.03.10.531095.

[20] S. A. Nastase et al., "The 'Narratives' fMRI dataset for evaluating models of naturalistic language comprehension," Sci Data, vol. 8, p. 250, Sep. 2021, doi: 10.1038/s41597-021-01033-3.