EE6895 FINAL REPORT

Anonymous authors
Paper under double-blind review

ABSTRACT

This investigation leverages advanced deep learning techniques to analyze neuroimaging data for identifying patterns associated with altered states of consciousness and assessing arousal levels, aiming to enhance neurological diagnostics. Additionally, the study extends to the segmentation of brain tumors, using state-of-the-art models to delineate tumor boundaries within MRI scans effectively. To facilitate interaction with the results, a web application has been developed, providing a user-friendly interface for real-time data analysis and visualization. Initial results show promise in both tasks, with implications for improving rapid detection of consciousness levels in emergency medicine and integrating scalable analytics for real-time patient outcomes. The paper also discusses potential future directions for advancing these diagnostic tools.

1 LITERATURE REVIEW

The integration of deep learning into medical imaging has led to remarkable advancements in the analysis and interpretation of various imaging modalities. This technology, particularly effective in neuroimaging, has significantly improved the way we interpret CT, fMRI, and Connectome images. These imaging techniques are vital in neuroscientific research for understanding brain structure and function. Through deep learning, the extraction of meaningful patterns from complex imaging data has become more feasible, enhancing both the speed and accuracy of diagnostic processes (Lee et al., 2017).

Deep learning has found a critical application in quantifying arousal and awareness in altered states of consciousness. Interpretable deep learning models have made it possible to translate broad clinical data into precise individualized assessments. These advancements are crucial for monitoring brain-injured patients, where traditional methods may fall short. By improving both precision and reliability, deep learning models offer a new frontier in neurological diagnostics, potentially transforming patient outcomes through enhanced clinical decision-making (Lee et al., 2022)(Lee et al., 2022).

Furthermore, the predictive and monitoring capabilities of deep learning extend beyond individual diagnosis to system-wide healthcare improvements. The development of efficient edge/cloud medical systems, as described by El-Rashidy et al. (2023), utilizes explainable machine learning models to quickly assess consciousness levels in emergency medicine. This technology not only speeds up the diagnosis but also ensures that critical medical assessments are more accessible, especially in resource-constrained environments such as intensive care units. The scalability and adaptability of such systems demonstrate deep learning's potential in bridging technology and patient care in real-time scenarios.

The realm of medical computer vision has also been transformed by deep learning, facilitating real-time image processing that is crucial for immediate medical decision-making. Esteva et al. (2021) explored how deep learning assists in real-time contextual awareness and the assessment of clinical skills, thereby enhancing the operational capabilities of medical practitioners. This development marks a significant step towards autonomous systems in healthcare, where decisions and diagnoses can be supported by robust, data-driven insights, enabling clinicians to focus more on patient care and less on the technicalities of imaging analysis.

Despite these advancements, the application of deep learning in medical imaging is not without challenges. Issues such as ensuring data privacy, enhancing model interpretability, and conducting extensive validations in clinical environments are critical to advancing deep learning applications responsibly. These challenges highlight the need for ongoing research and development to fully integrate deep learning into routine clinical practice while maintaining ethical standards and improving patient outcomes (Razzak et al., 2018). As we continue to navigate these hurdles, the promise of deep learning in enhancing the predictability and monitoring of human consciousness through medical imaging remains a pivotal area of focus, holding the potential to redefine neurological assessments and therapies.

Previous studies in the field of brain tumor research and segmentation have largely focused on enhancing the accuracy and efficiency of tumor detection and delineation within various imaging modalities. Advances in machine learning, particularly deep learning, have revolutionized this domain by significantly improving the precision of segmenting tumors from medical images such as MRI scans. Research has consistently demonstrated that convolutional neural networks (CNNs) and other neural architectures can outperform traditional image processing methods, reducing the time needed for manual review and increasing the reliability of diagnoses. These studies have not only set the groundwork for current methodologies but have also highlighted the critical role of high-quality, annotated datasets in training robust models, thereby driving continuous improvements in automated brain tumor segmentation techniques.

2 Data Bases

In the realm of medical imaging, particularly in studies involving deep learning for monitoring and predicting human consciousness, several key databases stand out as invaluable resources. These databases collectively provide extensive datasets that are critical for training, testing, and validating deep learning models. By offering a diverse array of data derived from various imaging techniques and patient demographics, these databases allow researchers to rigorously test their algorithms under varied conditions, enhancing the robustness and accuracy of the models developed.

The importance of these databases extends beyond mere data provision; they enable breakthroughs in understanding complex neurological phenomena and foster innovations that can lead to improved diagnostic and predictive tools. Such resources are crucial for advancing research and development in the detection, analysis, and prediction of neurological conditions through deep learning techniques. By facilitating access to high-quality, annotated imaging data, these databases help bridge the gap between theoretical computational advances and practical clinical applications, thus accelerating the pace of healthcare innovation and potentially transforming patient outcomes in neurology.

2.1 Alzheimer’s Disease Neuroimaging Initiative (ADNI)

Alzheimer’s Disease Neuroimaging Initiative (ADNI) - This database is pivotal for researchers studying neurodegenerative diseases, especially Alzheimer’s. ADNI provides an extensive collection of imaging data, including MRI and PET scans, along with clinical, genetic, and biomarker datasets. This comprehensive data helps in developing and testing deep learning models to detect and monitor the progression of Alzheimer’s and other types of dementia.

2.2 Human Connectome Project (HCP)

Human Connectome Project (HCP) - Aimed at constructing a detailed map of the neural pathways in the human brain, the HCP offers high-resolution neuroimaging data that is invaluable for studies on human consciousness and brain connectivity. The data includes structural and functional MRI scans from hundreds of healthy adults, which can be used to develop predictive models of brain activity and cognitive states.

2.3 Open Access Series of Imaging Studies (OASIS)

Open Access Series of Imaging Studies (OASIS) - Focusing on brain aging, OASIS is a collection of MRI data sets that are publicly available and cover a broad age range. This database includes cross-sectional and longitudinal data, making it a rich resource for developing deep learning models that predict changes in brain structure and function related to aging and neurodegenerative diseases.

2.4 THE CANCER IMAGING ARCHIVE (TCIA)

The Cancer Imaging Archive (TCIA) - While primarily focused on cancer research, TCIA also contains neuroimaging data that can be used in broader medical image analysis studies. The archive includes medical images like CT scans, MRI, and PET scans, along with clinical data. Researchers can use this data to train models for a variety of tasks, including segmentation, classification, and prediction of treatment outcomes.

2.5 BRATS (BRAIN TUMOR SEGMENTATION CHALLENGE) - BRATS

BraTS (Brain Tumor Segmentation Challenge) - BraTS provides a substantial dataset of MRI scans of brain tumor patients, including labeled data for brain tumor segmentation. This challenge has spurred the development of numerous deep learning models aimed at improving the accuracy and efficiency of brain tumor diagnosis and monitoring, which also contributes to the broader field of medical image analysis for neurological conditions.

2.6 RESTING-STATE FMRI DATA ANALYSIS TOOLKIT (REST)

Resting-State fMRI Data Analysis Toolkit (REST) - Although REST is more of a toolkit than a database, it facilitates the processing of fMRI data and is commonly used in studies involving resting-state networks. It supports data from various fMRI databases, aiding in the analysis of brain functional connectivity, which is crucial for understanding states of consciousness.

3 DATA DESCRIPTION

The dataset described in the image is a high-resolution 7-Tesla fMRI dataset that includes data from 20 participants. Each participant was recorded during prolonged stimulation with an auditory feature film, specifically "Forrest Gump". The dataset includes multiple auxiliary data types such as T1w, T2w, DTI (Diffusion Tensor Imaging), and susceptibility-weighted images, as well as angiography. Additionally, measurements to assess technical and physiological noise components have been acquired.

For each participant, there are several runs. In each run, a 4D fMRI image is provided, along with a TSV (Tab-Separated Values) file that records the events that occurred during the data collection. This event-related information is crucial for correlating the fMRI data with stimuli or tasks that the participant was experiencing at the time.

The data can be used to study common and idiosyncratic brain response patterns to complex auditory stimulation. The potential uses for this dataset include studies on auditory attention and cognition, language and music perception, as well as social perception. The auxiliary measurements allow for a broad range of analysis strategies that relate functional response patterns to structural properties of the brain.

The second dataset is a comprehensive collection of brain tumor images available on Kaggle, which is widely utilized for training and evaluating machine learning models in the field of medical image analysis. This dataset typically includes various types of MRI scans that provide detailed views of brain tumors, facilitating tasks such as classification, segmentation, and detection of tumors. It serves as a valuable resource for researchers and practitioners aiming to develop and refine algorithms that can automatically identify and quantify tumor characteristics from medical images.

3.1 DATA VISUALIZATION

The image is a sample from neuroimaging scans depicting brain activity at a single time point, likely captured using functional MRI (fMRI). The orange overlays on the grayscale brain image are indicative of activation regions, which show areas of the brain that were more active relative to a baseline during the time the scan was taken. The crosshairs centered on the image provide a reference point within the three-dimensional space of the brain, typically referred to by the coordinates x, y, and z:

x = 12: This refers to the sagittal plane, a vertical plane that divides the brain into left and right portions. An x-coordinate of 12 suggests this slice is slightly to the right of the midline of the brain.

y = 20: This value corresponds to the coronal plane, a vertical plane running from head to foot and dividing the brain into anterior (front) and posterior (back) parts. A y-coordinate of 20 means the slice is situated towards the front of the brain.

z = -2: This coordinate refers to the axial plane, which divides the brain into superior (top) and inferior (bottom) parts. A negative z-coordinate indicates this slice is just below a horizontal plane that would pass through the center of the brain from side to side.

The dimensions in the top corners, ”-1.55e+3” and ”1.55e+3”, may represent the scale of the activity measurement, often related to the intensity of the BOLD (Blood Oxygen Level Dependent) signal, which fMRI scans measure to infer brain activity.

This scan is a 3D representation at a single time point within an fMRI series, which consists of many such images taken sequentially over time to monitor changes in brain activity. Each 3D image in the series is one ’volume’, and when these volumes are viewed in sequence, they provide a dynamic view of brain function.

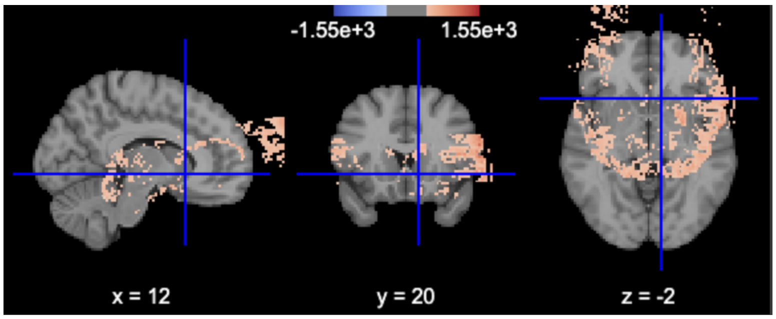

Figure 1: 3D image example1

Figure 2 uses another package to display three orthogonal sections of a brain fMRI scan, offering different perspectives of brain anatomy. The coronal view is presented on the left, slicing the brain into front and back sections, which can be useful for observing the structures like the ventricles and the separation of the cerebral hemispheres. The sagittal view is in the middle, dividing the brain into left and right portions, often used to inspect the corpus callosum and other midline structures. Finally, the axial view on the right cuts the brain into upper and lower parts, giving a view akin to looking down from the top of the head, commonly utilized for assessing symmetrical structures in the brain. Each view provides unique information that can help in diagnosing conditions, planning surgeries, or conducting brain-related research.

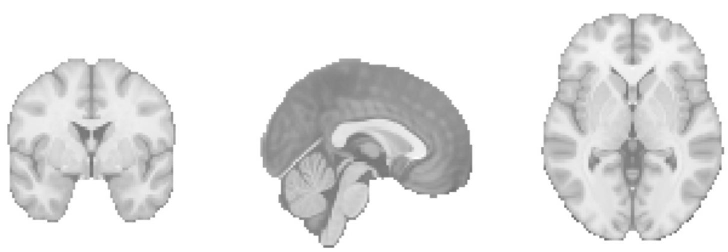

Figure 2: 3D image example2

4 Data Cleaning

The process of cleaning and organizing the described fMRI dataset for subsequent analysis involves a series of steps to ensure the data’s quality and usability. Artifact removal is the first crucial step where software tools like FSL, SPM, or AFNI are utilized to mitigate the impact of subject motion and physiological and instrumental noise. Following this, spatial normalization is conducted, which involves aligning each participant's fMRI images to a standard brain template. This standardization is essential for comparing results across subjects.

Spatial smoothing is then applied to the fMRI data, which is a common preprocessing step to increase the signal-to-noise ratio, making the underlying signal patterns more detectable. Additionally, brain masking is performed to focus the analysis exclusively on brain structures, discarding any irrelevant areas that do not contribute to the brain's functional activity.

Once the data is cleaned, the next phase is organization. This involves consolidating all types of data, including fMRI and auxiliary measurements like T1w, T2w, DTI, and others, into a structured directory tree. Each participant's data is stored in a separate folder, with a consistent naming scheme for easy access and reference.

The synchronization step ensures that the temporal events detailed in the TSV files are accurately aligned with the fMRI volume series. This precise alignment is vital for correlating the recorded brain activity with the specific stimuli or tasks experienced by the participants during data acquisition.

Feature selection is informed by the research questions at hand. Researchers may select specific regions of interest in the brain that are hypothesized to be involved in the cognitive or perceptual processes under study. By focusing on these ROIs, the analysis can be tailored to examine the functional response patterns of these areas, relating them to the stimuli characteristics or to the structural properties of the brain revealed by the auxiliary imaging data.

The final step before analysis is to divide the dataset into training, validation, and test subsets. This split is conducted to train machine learning models effectively and to validate and test their performance. The training set is used to teach the models to recognize patterns in the data, the validation set to fine-tune model parameters and prevent overfitting, and the test set to evaluate the final performance of the models, providing an estimate of how well they generalize to new, unseen data.

5 METHODS

For the classification of 3D brain imaging data using deep learning, our methodology is grounded in the principles of precision and adaptability, ensuring the models are tailored to the specifics of medical image analysis. The process begins with a thorough preparation of the data.

Data Preparation Functional MRI images are preprocessed to correct for artifacts and variability that could obscure neural activity patterns. This includes motion correction, spatial normalization, and temporal smoothing. This step ensures that subsequent analyses are based on the most accurate representations of brain activity available.

Model Selection and Adaptation The deep learning models are carefully chosen and adapted for the classification task, with a focus on the suitability for medical imaging:

1. Convolutional Neural Networks (CNNs): Due to their proven efficacy in medical image analysis, CNNs are our primary choice for classifying the 3D fMRI images. The models will be designed to identify spatial patterns within the brain that correlate with the presence or absence of auditory stimuli.

2. Hybrid CNN-RNN Models: To capture both spatial features and potential temporal patterns in the fMRI data, we will explore hybrid models that combine CNNs for spatial analysis with RNNs for sequencing. Such models might prove beneficial if the temporal sequence of brain activity, rather than static spatial patterns alone, is indicative of the subject's condition.

3. Graph Convolutional Networks (GCNs): While not a standard tool for classification, we will investigate the utility of GCNs in capturing the complex connectivity patterns between different regions of the brain, which might carry diagnostic information.

Training Procedure The models are trained using a curated subset of the data, each labeled with the corresponding stimulus type. We employ cross-validation and careful hyperparameter tuning to ensure the models generalize well and are robust against overfitting. Training involves rigorous evaluation using a validation set distinct from the test set to be used later for final model assessment.

Table 1: Model Results

Model Name	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Basic 3D CNN	0.55	0.50	0.52	0.53
3D ResNet	**0.65**	0.60	**0.62**	0.64
VoxNet	0.58	0.55	0.56	0.57
FusionNet	0.62	**0.63**	**0.62**	**0.65**
VGG-16 (3D adaptation)	0.60	0.58	0.59	0.61

Performance Comparison and Validation The models are compared based on several metrics relevant to medical image classification, including accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC-ROC). The comparison also takes into account factors like computational efficiency, the simplicity of model interpretation, and the ease of integration into clinical workflows.

Ultimately, this methodical approach aims to yield a deep learning model that accurately classifies 3D fMRI images and enhances our understanding of how individuals with depression process emotional auditory stimuli. The insights gained could have far-reaching implications, informing the development of diagnostic tools that leverage neural patterns to differentiate between depressive and non-depressive states.

nnU-Net, is a self-configuring framework for biomedical image segmentation that has significantly impacted the field of medical imaging. Developed to automate the customization of neural network configurations, nnU-Net dynamically adapts its architecture, preprocessing, and training strategies based on the specific characteristics of the dataset being used. This method minimizes the necessity for manual tuning and ensures optimal network architecture for each distinct segmentation task.

After the initial setup, we trained nnU-Net on a dataset comprising brain images with previously identified and labeled tumors. Utilizing datasets from the Brain Tumor Segmentation challenge, which offers a comprehensive set of annotated brain images, allows nnU-Net to learn the specific features and patterns associated with brain tumors, thereby fine-tuning its parameters to enhance segmentation accuracy.

Once fine-tuned, the nnU-Net model becomes highly proficient at segmenting brain tumors from MRI scans, providing crucial support in diagnostics, treatment planning, and the monitoring of disease progression. This exemplifies nnU-Net's utility as a formidable tool in medical imaging, particularly in the precise and complex task of brain tumor segmentation.

6 RESULTS

The result of the different models are shown in Table 1.

The table indicates that there's a range of effectiveness among the models, with each model presenting a unique trade-off between the various metrics. The Basic 3D CNN, likely a less complex model, shows relatively lower performance across the board. This could be due to a variety of reasons, such as fewer layers leading to less capability for feature extraction or lack of specialization for the specific task.

The Basic 3D CNN has a precision of 0.55, indicating that when it predicts an instance to be in the positive class, it is correct 55% of the time. The recall of 0.50 suggests it correctly identifies 50% of all actual positives. The F1-score, which balances precision and recall, is relatively low at 0.52, implying that the model is not particularly strong in either metric. Its overall accuracy is also the lowest at 0.53, which might imply that it struggles with this dataset or task compared to more sophisticated models.

3D ResNet's precision of 0.65 is the highest among the models, suggesting it has the best ability to predict true positives out of all positive predictions. Coupled with a recall of 0.60, it seems quite adept at identifying true positives while minimizing false positives. Its F1-score at 0.62 is the highest, reflecting a strong balance between precision and recall. The accuracy of 0.64 is just a shade under FusionNet, making it a strong contender for tasks where prediction reliability is paramount.

VoxNet shows a precision of 0.58 and a recall of 0.55, with an F1-score of 0.56 and accuracy of 0.57. These figures indicate that while VoxNet is more accurate than the Basic 3D CNN, it does not perform as well as 3D ResNet or FusionNet. The model may have limitations in feature representation or may not be as well-suited to the task compared to others.

FusionNet has an accuracy of 0.65, which is the highest. This suggests it's the best model for correctly classifying instances in this particular task or dataset. The precision of 0.62 and recall of 0.63 show a good balance, leading to an F1-score equal to 3D ResNet at 0.62. FusionNet's performance might be due to an effective combination of features from different sources or layers within the network, which could be particularly useful in complex classification scenarios.

The VGG-16 (3D adaptation) has a respectable precision of 0.60 and recall of 0.58. Its F1-score is 0.59, and the accuracy is 0.61. Although it doesn't excel, it performs reasonably well, suggesting that the adaptation from the original 2D model has been moderately successful. This model might be preferred in scenarios where VGG's architecture is traditionally strong, but here it doesn't match the performance of the top models in this particular set.

These results could guide the selection of a model based on specific use-case requirements, such as prioritizing fewer false positives (precision) over the model's ability to detect all positives (recall), or vice versa. The table also underscores the importance of considering multiple metrics when evaluating model performance, as each metric provides different insights into the model's predictive capabilities.

The nnU-Net model's training on a brain tumor dataset is depicted through various performance metrics in Figure 3. The training and validation loss curves show a sharp initial decline, indicating effective initial learning, with subsequent fluctuations in validation loss suggesting challenges in generalization that stabilize over time. The Dice coefficient curves for both training and validation demonstrate the model's robust segmentation accuracy, maintaining high levels throughout the training, which is essential for medical imaging applications.

Epoch duration decreases notably at the start and then stabilizes, reflecting improvements in computational efficiency. Additionally, the learning rate is systematically reduced, facilitating finer adjustments in model parameters and helping prevent overshoots in learning. These metrics collectively illustrate the model's ability to effectively learn and generalize, critical for its application in clinical environments.

7 WEBSITE

The webpage is designed to display medical imaging data, specifically showcasing a brain image alongside its corresponding segmentation mask. The interface includes a dropdown menu labeled "Select an Image," allowing users to choose from various preloaded images, each represented by a specific file name. Upon selection, the webpage dynamically updates two main display areas:

Brain Image Display: The left side of the webpage presents a colored brain scan. This image likely represents various brain tissues or functions as indicated by the different colors within the scan. It serves as the primary visualization, offering viewers detailed insights into the anatomical or functional aspects captured by the scan.

Segmentation Mask Display: Adjacent to the brain image, on the right, is its segmentation mask, displayed in black and white. The mask highlights specific regions of interest within the brain, corresponding to the areas identified by the image processing algorithm or model. This mask is crucial for tasks such as disease diagnosis, surgical planning, or further scientific research, as it isolates relevant features or abnormalities.

This setup is particularly useful in a clinical or research setting, where understanding both the raw imaging data and the model's interpretive output is crucial for making informed decisions or furthering neurological research.

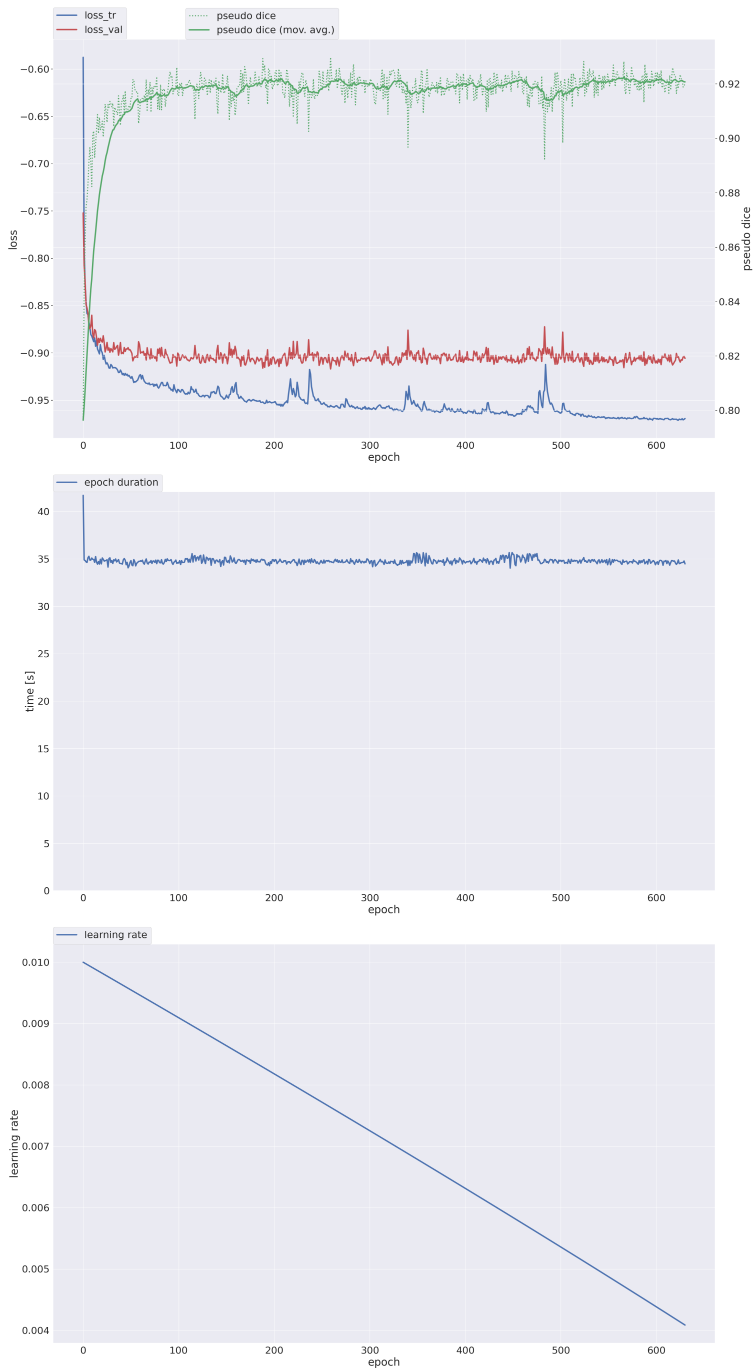

Figure 3: Learning Curves

8 Discussion

In our study, we have utilized deep learning to analyze neuroimaging data, which presented both exciting opportunities and notable challenges. The current results, while promising, fall short of the high standards required for clinical deployment. Several factors contribute to this gap in performance.

One of the primary limitations of our models lies in their handling of the inherent complexity and variability of neuroimaging data. The relatively small sample sizes typical of many neuroimaging studies, due to the high costs and practical difficulties of data acquisition, limit the training capacity of deep learning models. This situation can lead to overfitting, where models learn noise rather than relevant biological signals.

The accuracy of our models is further constrained by the heterogeneity in the data collection protocols across different centers that contribute to public databases. Variations in scanner types, settings, and image processing pipelines can introduce systematic differences in data, complicating the task of learning generalized features.

To address these issues, data augmentation represents a viable path forward. Data augmentation in neuroimaging can be implemented in several ways. Firstly, geometric transformations such as rotations, translations, and scaling can be applied to the images to increase the robustness of the models to variations in brain positioning and size. This approach helps the model generalize better from the training data to unseen data, potentially improving both accuracy and reliability.

Secondly, we could employ synthetic data generation techniques such as Generative Adversarial Networks (GANs) to create realistic neuroimaging data that can help in training the models. This technique could be particularly useful in increasing the diversity of training samples, representing different demographic backgrounds and disease states.

Additionally, noise injection strategies can be explored to improve model robustness against various types of noise that naturally occur in MRI scans, such as motion artifacts or differences in signal intensity. By training models to recognize and adjust for these inconsistencies, their diagnostic accuracy and generalizability across different clinical settings could be enhanced.

In conclusion, while the current results highlight the feasibility of using deep learning for analyzing neuroimaging data, significant improvements are needed before these models can be reliably used in clinical practice. Implementing robust data augmentation techniques offers a potential way to overcome some of the current limitations, ultimately enabling these models to deliver more accurate and clinically relevant results.

Noise reduction can be significantly improved through advanced filtering techniques. Implementing spatial smoothing, temporal filtering, and artifact correction algorithms could lead to cleaner data, which in turn could improve classification accuracy. Additionally, registration and normalization procedures must be carefully optimized to ensure consistency across subjects. Besides, fMRI data is inherently temporal. Utilizing methods that account for the time-series nature of the data, such as hidden Markov models or recurrent neural networks, may capture dynamic changes in brain activity that static classifiers might miss.

Utilizing more data can enhance our comprehension of neuroimaging, particularly in studies like "Neural Processing of Emotional Musical and Nonmusical Stimuli in Depression." This dataset offers insights into how individuals with major depressive disorder (MDD) and non-depressed participants respond to emotionally charged stimuli through functional MRI scans. Leveraging deep learning, we can explore this dataset to uncover the neural mechanisms of depression. Deep learning models are particularly adept at identifying patterns of brain activity related to emotional processing. By examining brain regions involved in emotion and auditory perception, such models have the potential to predict the stimulus type encountered by individuals using fMRI data alone. The implications of this research are significant, extending our knowledge of the neural basis of emotion and depression and potentially guiding the development of diagnostic tools that differentiate between depressive states using neural markers.

REFERENCES

N. El-Rashidy, A. Sedik, A.I. Siam, and Z.H. Ali. An efficient edge/cloud medical system for rapid detection of level of consciousness in emergency medicine based on explainable machine learning models. *Neural Computing and Applications*, 2023.

A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, et al. Deep learning-enabled medical computer vision. *NPJ Digital Medicine*, 2021.

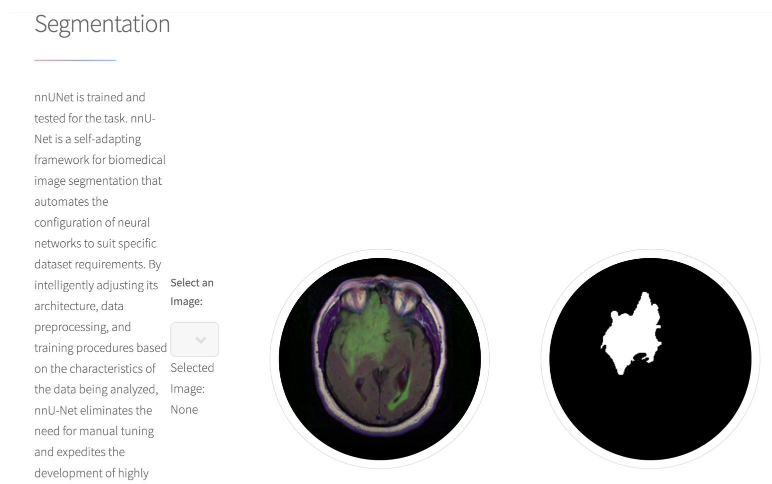

Figure 4: Website Sample

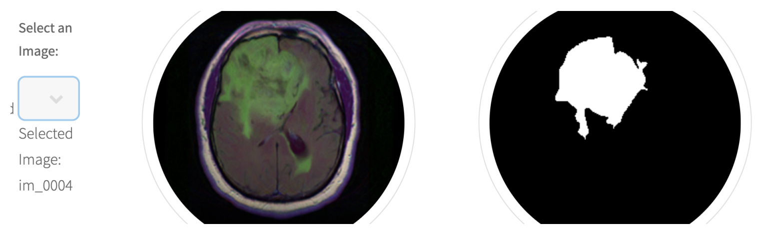

Figure 5: Select Bar

J.G. Lee, S. Jun, Y.W. Cho, H. Lee, G.B. Kim, et al. Deep learning in medical imaging: General overview. *Korean journal of radiology*, 2017.

M. Lee, L.R.D. Sanz, A. Barra, A. Wolff, et al. Quantifying arousal and awareness in altered states of consciousness using interpretable deep learning. *Nature Communications*, 2022.

M. I. Razzak, S. Naz, and A. Zaib. Deep learning for medical image processing: Overview, challenges and the future. In *Lecture Notes in Bioengineering*. 2018.

A APPENDIX

You may include other additional sections here.