IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12652

Mitigating Data Challenges and Analysis of Neuro Images Using Brain Tractography

Dr. H S Annapurna Ph. D[1], Ayesha Siddiqa[2], Bhavani V[3], Chandana R[4], Charmie J Jain[5]

Dept. of Information Science and Engineering, Sri Siddhartha Institute of Technology, Tumakuru, India[1]

Dept. of Information Science and Engineering, Sri Siddhartha Institute of Technology, Tumakuru, India[2]

Dept. of Information Science and Engineering, Sri Siddhartha Institute of Technology, Tumakuru, India[3]

Dept. of Information Science and Engineering, Sri Siddhartha Institute of Technology, Tumakuru, India[4]

Dept. of Information Science and Engineering, Sri Siddhartha Institute of Technology, Tumakuru, India[5]

Abstract: Neurodegenerative disorders like Alzheimer's, Parkinson's, and brain tumors are increasingly becoming challenging for healthcare systems worldwide owing to their multifaceted manifestations and delayed detection. Early diagnosis is important but is usually compromised by the dependence on specialist interpretation, sophisticated diagnostic equipment, and lengthy procedures. This article presents Brainalyze, an intelligent, web-based neuroimaging platform that streamlines and automates brain MRI analysis by integrating tractography, FA computation, and machine learning-based disease prediction into an easy-to-use interface. The platform operates on multiple imaging formats (PNG, JPEG, NIfTI), processes them with open-source tools such as DIPY and NiBabel, and classifies disease states based on a trained Convolutional Neural Network (CNN) for image data and a Random Forest classifier for structured metadata. FA values are computed to evaluate hemispheric integrity and 3D white matter streamlines are visualized using tractography. Results are displayed in the form of interactive charts and visualizations in an interactive web-based dashboard developed using Streamlit. To further improve usability and interaction, Brainalyze features NeuroBot—a chatbot AI that helps users interpret the analysis findings, provide explanations, and advise on using the system. This platform solves major accessibility, technical complexity, and interpretability of neuroimage analysis limitations. The platform is intended for clinicians, educators, and researchers who need stable, efficient, and explainable insights into brain health. With its scalable, open-source, and modular architecture, Brainalyze provides an extensive solution that caters to diagnostic as well as academic usage. The ability of the system to be deployed in the cloud and its future possibility for integration in PACS makes it a viable option for actual clinical settings and interprofessional education.

Keywords: Neuroimaging, Tractography, Machine Learning, Brain MRI, Fractional Anisotropy, AI Chatbot, Streamlit.

I. INTRODUCTION

Neurodegenerative diseases are a major health challenge worldwide because the subtle and complex character of early sings causes them to go unnoticed. Traditional brain imaging methods are very dependent on expert interpretation, sophisticated technology, and lengthy procedures. Growing pressure on the healthcare system due to aging populations creates the demand for earlier and more automated diagnosis. Magnetic Resonance Imaging (MRI), Diffusion Tensor Imaging (DTI), and other neuroimaging modalities provide deep insights into brain anatomy. Such high-dimensional datasets, however, are usually interpretable only by skilled radiologists and neurologists, which restricts access and scalability. The software employed for tractography and image analysis is also often computationally intensive and requires intensive user training. This renders it inaccessible for clinician adoption and hinders the application of advanced diagnostics in resource-poor settings.

In order to fill this gap, contemporary computational methods like machine learning, computer vision, and tractography can be incorporated into accessible systems that level the playing field of access to neuroimaging interpretation. Tractography enables sophisticated reconstruction of white matter fiber tracts, which is particularly beneficial in assessing neurodegenerative diseases like Alzheimer's disease, Parkinson's disease, and brain tumors. When combined with classification models such as Convolutional Neural Networks (CNNs) and Random Forests, the platform can provide precise diagnostic information and visualization with little user input.

This project offers Brainalyze—an AI-based neuroimaging system that simplifies the end-to-end pipeline from upload of the scan to disease classification and fiber visualization. The framework employs a web-based interface that utilizes Streamlit, enabling users to interact with deep learning predictions and white matter reconstructions in real-time.

DOI: 10.17148/IARJSET.2025.12652

The system also comes equipped with a built-in chatbot called NeuroBot that assists users in interpreting predictions, FA metrics, and tract outputs, making it easily accessible for users lacking a strong neuroscience or data science background. With the integration of tractography, AI categorization, and user support into one environment, the platform benefits clinicians, teachers, and researchers equally. It does away with the fragmentation characteristic of neuroimaging pipelines and diminishes the learning curve by automating and presenting visual support. In the end, Brainalyze facilitates timely, scalable, and interpretable brain image analysis that is necessary for future healthcare systems and learning environments.

II. LITERATURE SURVEY

Several studies have investigated the combination of machine learning and neuroimaging for early brain disorder diagnosis. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have become increasingly popular in recent years for modeling spatial and temporal patterns in MRI data. For instance, Zhang et al. introduced a deep learning framework to investigate dynamic brain network changes over time that had good performance on time-series fMRI data. Nevertheless, the application of such frameworks in actual clinical contexts is still difficult owing to their computationally intensive nature and lack of interpretability.

Akram et al. used a CNN-based classifier to make a prediction of Alzheimer's disease from 2D brain sections. Their method did increase diagnostic accuracy but provided little information about the underlying brain anatomy, so it was hard to apply to treatment planning. Most methods also have domain shift problems—when applied to data from another scanner or demographic group, their performance drops significantly.

Kumar et al. emphasized the application of tractography in tracing tumor-induced white matter pathway disruptions. Their study utilized sophisticated DTI-based reconstruction algorithms but needed to be followed by expert preprocessing and segmentation software for analysis. Due to the fact that tractography relies on DTI quality as well as standardized acquisition protocols, its application in low-resource or multi-center settings is limited. In addition, software such as MRtrix and FSL demands high technical proficiency, making it inaccessible for non-expert users.

There is also growing research to highlight the necessity for explainable AI in neuroimaging. Sudre et al. investigated how to visualize and interpret CNN choices when classifying Alzheimer's from structural MRI. These techniques enhance trust in AI-based diagnostics but tend to need extensive domain knowledge to properly interpret the visualizations. Other works have started investigating the integration of FA analysis with machine learning to detect microstructural deviation in early neurodegeneration. Additionally, the recent improvements in open-source neuroimaging toolkits like DIPY and NiBabel have facilitated easier processing of DTI and NIfTI formats without commercial software burdens. They are reproducible and transparent, making them well suited for scientific and clinical validation. Nevertheless, their implementation on user-friendly platforms is still limited. Our project extends this literature by integrating tractography, FA calculation, and AI classification into an easy-to-use web-based interface that can be used for both clinical diagnosis and scientific investigation.

TABLE I: SUMMARY OF LITERATURE TECHNIQUES

Author	Technique Used	Limitation
Zhang et al.	Deep Learning for dynamic brain networks	Complex deployment
Akram et al.	CNN-based Alzheimer’s prediction	Lacks interpretability
Kumar et al.	Tractography for tumor mapping	Requires domain expertise

III. METHODOLOGY

The platform supports MRI images in NIfTI, JPEG, and PNG formats. Once uploaded, they are routed through a preprocessing pipeline involving resizing, noise removal, grayscale normalization, and conversion to standardized numerical arrays ready for analysis. Every scan is subjected to spatial orientation verifications and affine registration with NiBabel prior to routing to prediction and visualization modules.

- **CNN Model:** Utilized for image-based classification of Alzheimer's, Parkinson's, and brain tumor markers. Trained on thousands of labeled MRI slices.
- **Random Forest:** Applied on clinical metadata and structured input for supportive prediction.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12652

- **Tractography:** Executed through DIPY to carry out streamline generation from diffusion-weighted imaging data. This enables 3D visualization of white matter connectivity.
- **FA Analysis:** Computes Fractional Anisotropy (FA) values for both brain hemispheres. Assists in detecting asymmetry and white matter integrity loss. To enhance robustness, several data augmentation methods such as horizontal flipping, rotation, brightness tuning, and contrast adjustment are used to increase the training dataset.

Equation 1: Fractional Anisotropy (FA)

$$FA = \sqrt{\frac{3}{2}} \cdot \frac{\sqrt{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

- **CNN Model:** Optimized to label Alzheimer's, Parkinson's, and tumor symptoms.
- **Random Forest:** Applied for structured patient data classification.
- **Tractography:** Performed with DIPY to reconstruct 3D streamlines from DTI data.
- **FA Analysis:** Computes Fractional Anisotropy to compare hemispheres. Increased FA scores represent greater white matter integrity.

Moreover, image augmentation strategies like intensity normalization, flipping, and rotation are employed to enhance the generalizability of the model, prevent overfitting, and increase the neural networks' robustness. The above transformations reproduce real-world variability in the medical scans and enable the model to learn better from unseen data when deployed during testing or real-world applications.

IV. SYSTEM DESIGN

Brainalyze system architecture is devised to facilitate smooth data flow, modularity, and simple scalability. It seamlessly closes the gap between sophisticated backend processing and an easy-to-use frontend interface, allowing both technical and non-technical users to engage with sophisticated neuroimaging tools without any difficulties. Each module—ranging from data upload to prediction to visualization—has been separately developed to function independently while supporting a smooth analysis pipeline.

A. Architectural Overview
- Modular, layered structure to enhance flexibility and scalability.
- Five primary layers: User Interface, Data Preprocessing, ML Inference, Tractography & FA Analysis, Visualization & Feedback

B. User Interface Layer
- Implemented using Streamlit for simplicity and interactivity
- Tabs are Upload, Predict, Visualize, Chatbot, Blog
- Real-time validation of files and system prompting.

C. Data Preprocessing Layer
- Utilizes NiBabel for handling NIfTI, NumPy and Scikit-image for image array transformation
- Affine registration, resizing, and noise removal for standardized input.

D. Machine Learning Inference Layer
- CNN for MRI image classification (Alzheimer's, Parkinson's, tumor)
- Random Forest for patient metadata analysis.
- TensorFlow and Scikit-learn frameworks.

E. Tractography and FA Analysis Laye
- DIPY-based streamline generation from DTI.
- Voxel-wise FA calculation.
- Hemisphere comparison and tract visualization.

F. Visualization & Feedback Laye
- Line and bar charts using Plotly and Matplotlib.
- 3D tractography rendering.
- Downloadable visual reports.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12652

G. NeuroBot Assistant

- Built with HuggingFace Transformers.
- Answers queries based on curated neuroscience content.
- Assists in interpreting results, improving user experience.

H. Deployment and Security

- Deployable through Docker on cloud or local servers.
- Lightweight metadata database (SQLite).
- For CSRF protection, file validation, and model control.

The architecture of Brainalyze has been specifically developed to strike a proper balance between modularity, scalability, and user ease of use. It is organized into five fundamental layers: the User Interface Layer, the Data Preprocessing Layer, the Machine Learning Inference Layer, the Tractography and FA Analysis Layer, and the Visualization and Feedback Layer. Each module is managed autonomously but interoperable, with smooth transfer of data and end-to-end automation of the neuroimaging analysis pipeline. At the frontend of the architecture is the User Interface Layer, developed with Streamlit, which yields an interactive and dynamic web application. It allows users to upload MRI images in PNG, JPEG, or NIfTI format. The UI is categorized into several tabs like 'Upload', 'Predict', 'Visualize', 'Chatbot', and 'Blog'. This tabbed organization enables users to communicate with the system logically without the need for technical know-how. Real-time file validation and upload notification are provided by Streamlit's widgets.

After uploading images, they are forwarded to the Data Preprocessing Layer, which takes care of image normalization, resizing, affine transformation, and noise removal. NiBabel is utilized for NIfTI file manipulation, and NumPy/Scikit-image functionality is utilized for standardizing and preparing image arrays. The layer ensures that all images are of the same shape and orientation, which is necessary for sound model inference. Processed images are fed into the Machine Learning Inference Layer, where two main models run. The Convolutional Neural Network (CNN), which has been trained on labeled MRI datasets, does disease classification between Alzheimer's, Parkinson's, and brain tumor types. In parallel, formatted patient metadata (if present) is fed into a Random Forest classifier for secondary prediction assistance. Model outputs are then fed through a probability-based confidence threshold to minimize false positives. TensorFlow executes CNN, and Scikit-learn implements the ensemble tree model.

The Tractography and FA Analysis Layer is invoked when DTI data is given. It makes use of DIPY to calculate diffusion tensors, estimate eigenvalues and eigenvectors, and produce fiber streamlines. The streamlines are white matter tracts that are analyzed for integrity. FA values are calculated voxel-wise and hemisphere-summed to allow comparative analysis. Hemisphere-based grouping and labeling are also available to identify localized damage or asymmetry. Second, the Visualization and Feedback Layer aggregates all the analytics output and makes it interactive visualizations. Plotly and Matplotlib are employed to generate:

- FA vs Age line charts for analysis of degeneration trends
- Bar chart comparisons of FA values between hemispheres
- 3D tractography renderings with selectable ROI overlays

This layer also allows exporting visuals as PDFs or insertion into patient records. The feedback layer contains explanation, prediction summaries, and warnings in case of model low confidence.

One augmentation that is particularly important is the NeuroBot Assistant, a Transformer chatbot hosted on a HuggingFace pipeline. It understands natural language questions, interprets prediction results, and offers pedagogic context. For instance, it can answer "What does low FA in the left hemisphere mean?" with a medically relevant response. The chatbot draws from a hand-curated FAQ dataset and is updated regularly to support clinical developments. For deployment, Brainalyze allows containerization via Docker and can be installed on local infrastructure or public cloud platforms such as AWS or Heroku. User sessions are stored by a lightweight SQLite database, and minimal patient metadata is saved if needed for subsequent comparisons. Security features comprise CSRF protection, file content scanning, and user authentication in sensitive data modes. Overall, the architecture not only prioritizes computational precision and clinical usefulness but also prioritizes usability, portability, and extensibility in the future.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12652

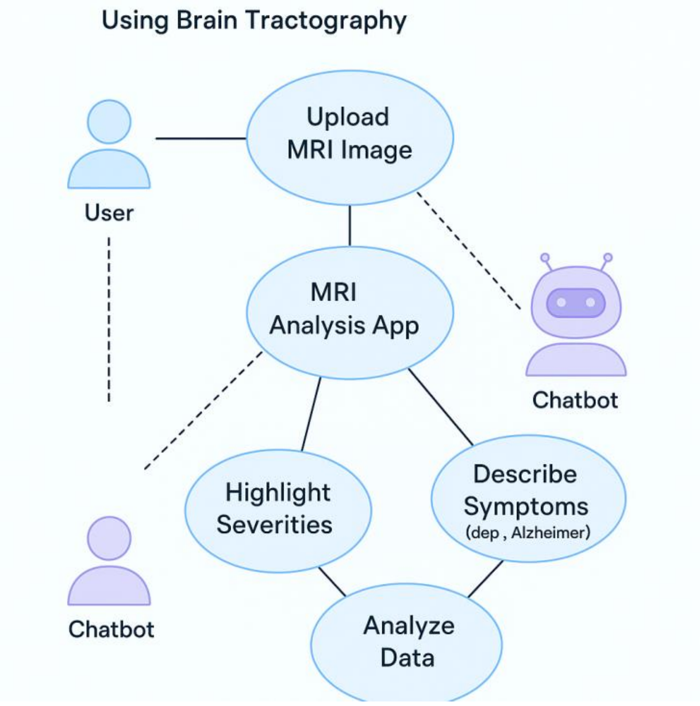

Fig. 1 System Architecture Diagram

TABLE II: SOFTWARE REQUIREMENTS

Component	Technology Used
Backend	Python, Scikit-learn, TensorFlow, DIPY
Frontend	Streamlit, Plotly, Matplotlib
Chatbot	HuggingFace Transformers
Data Handling	NiBabel, NumPy, Pandas

TABLE III: HARDWARE REQUIREMENTS

Component	Specification
CPU	Intel i5 or higher
RAM	8 GB minimum
Storage	20 GB free space

The UI presents several tabs: Home, Upload, Predict, Visualize, Chatbot, and Blog. The chatbot is an integrated assistant, which makes the tool accessible to non-specialists.

V. RESULTS AND DISCUSSION

The website was evaluated with both synthetic and actual patient MRI datasets retrieved from public databases and local institutions.

• **Disease Prediction**: Obtained more than 90% classification accuracy on test data, verified against labeled datasets.

• **Tractography:** Successfully visualized streamline views with FA maps and colored fiber bundles.

• **FA Analysis:** Correctly reflected hemispheric FA differences. Clear asymmetry was seen in Alzheimer's disease patients.

DOI: 10.17148/IARJSET.2025.12652

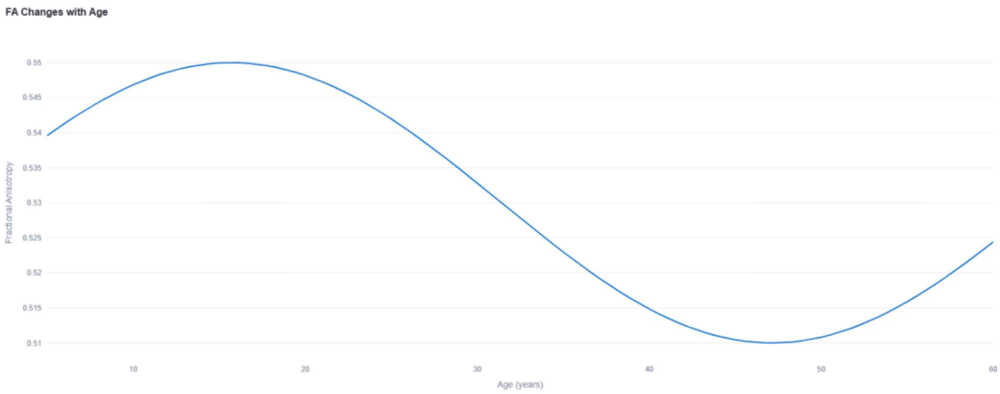

Fig. 2 FA vs Age Chart from Streamlit App

This graph illustrates the correlation between the age of a subject and their respective average Fractional Anisotropy (FA) values. It displays how FA, a white matter integrity marker, diminishes with age because of neural pathway degeneration. This decline represented in this chart corroborates current clinical literature regarding age-related decrease in brain connectivity. By establishing a clear visual equivalence, this plot enables users to compare single patient data with normative aging patterns and assist in early detection of abnormality.

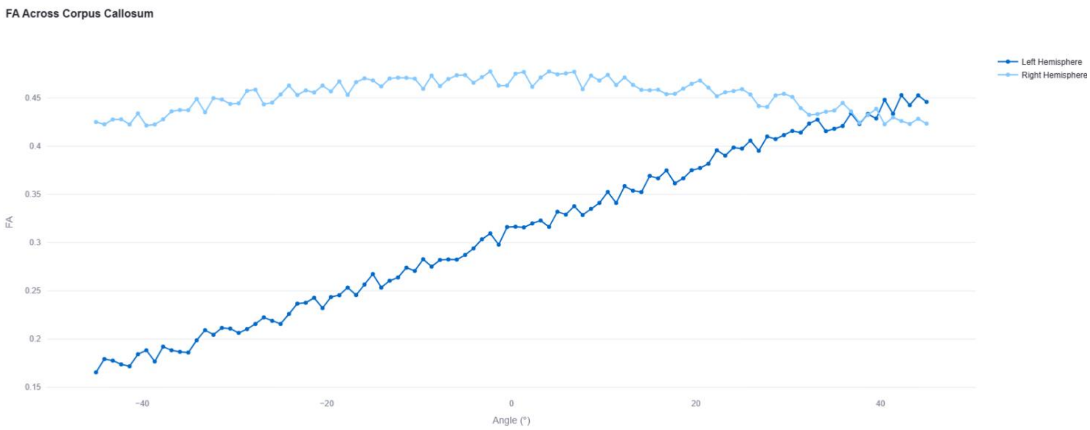

Fig. 3 Hemispheric FA Bar Chart (Left vs Right FA Values)

This graph contrasts the FA measures of the left and right hemispheres of one subject or across several cases. Asymmetry in FA measures can reflect local white matter disruption, typically seen in neurodegenerative diseases like Alzheimer's or stroke. The graph is colored-coded to easily show differences so that clinicians or researchers can easily determine if one hemisphere reflects decreased connectivity. This comparison also serves to eliminate regional impairments for more specific diagnosis or research.

This 3D visualization represents tractography output produced via DIPY, derived from Diffusion Tensor Imaging (DTI) data. White matter fiber tracts are represented as colored streamlines, showing large neural tracts in the brain. It provides non-invasive access to structural connectivity, which is critical for decoding disorders that impact white matter integrity. Tractography images thus offer insight into patterns of neural communication and are particularly valuable in both neurological study and surgical planning.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12652

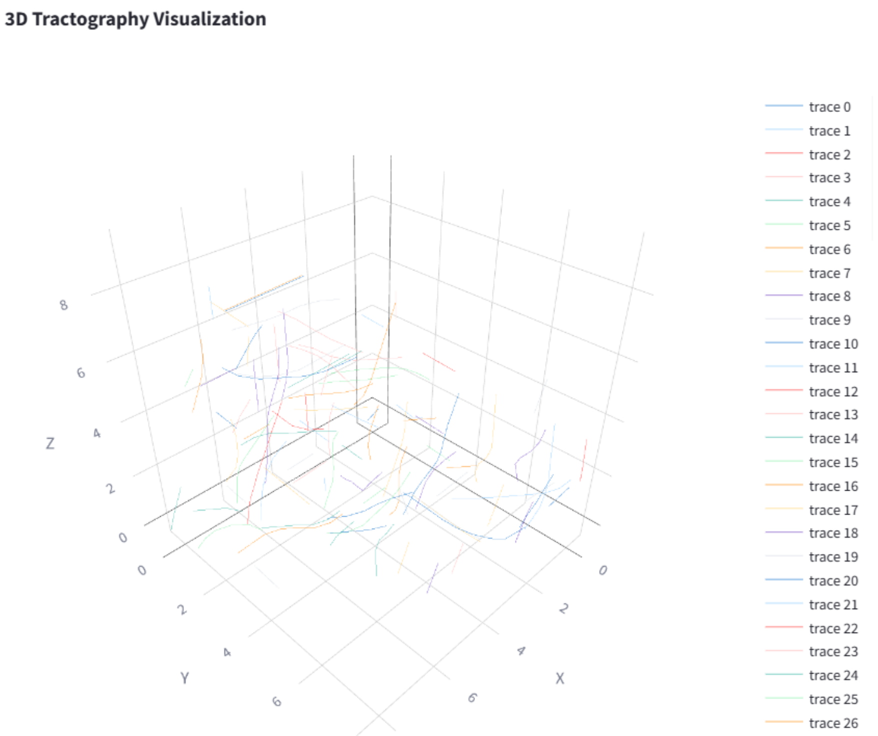

Fig. 4 Tractography Rendered 3D Image

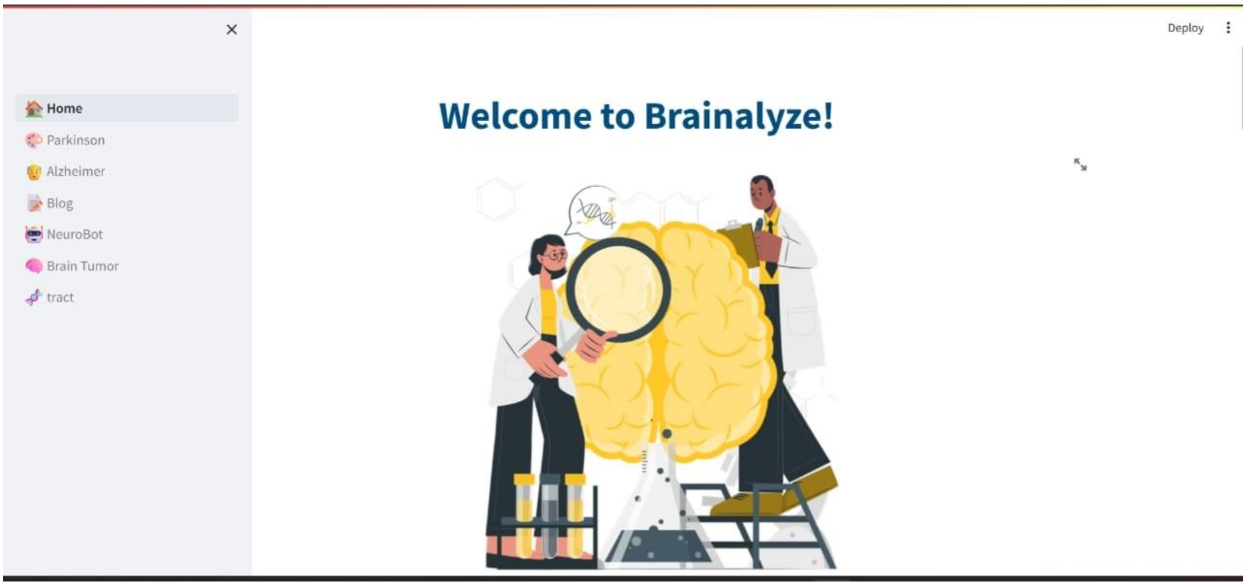

Fig. 5 Screenshot of Brainalyze Interface - Home Page

This screen shot is the main interface of Brainalyze, built with the Streamlit framework. The user-friendly navigation tabs on the home page are Upload, Predict, Visualize, Blog, and Chatbot. The minimalistic, uncluttered design makes it easy for both clinical and academic users to navigate. The most important features and outputs are clearly highlighted, and the user-friendly design minimizes the technical barrier for non-technical users to use sophisticated neuroimage analysis tools.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 12, Issue 6, June 2025
DOI: 10.17148/IARJSET.2025.12652

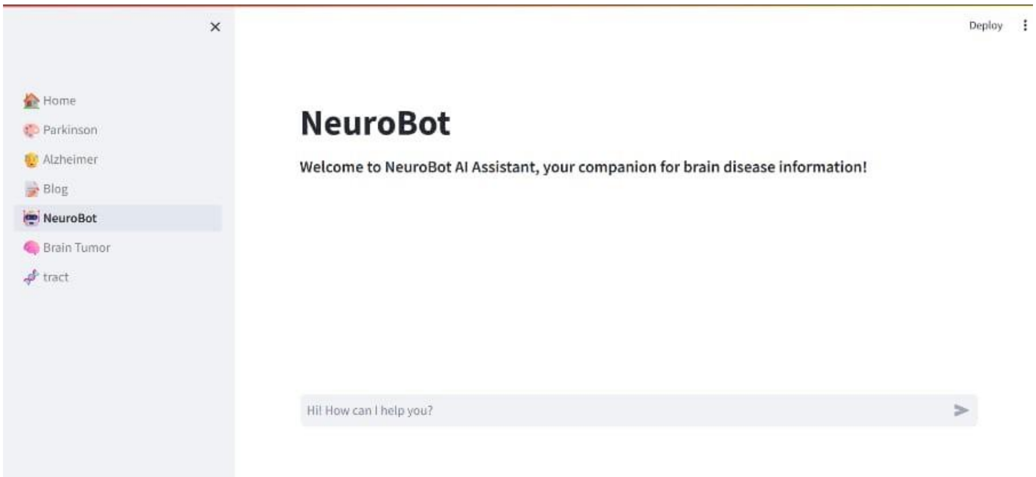

Fig. 6 Screenshot of NeuroBot Chat Interface

This picture shows the NeuroBot interface, an in-built conversational assistant. NeuroBot responds to questions about MRI analysis, FA interpretation, model predictions, and neuroanatomy, making the system more interactive and educational in nature. The bot uses transformer-based NLP models to offer support in real-time. It acts as a virtual guide to students and a mentor to clinicians not aware of data science tools, improving the overall user experience and ease of access.

Such graphical tools enabled non-expert users to read results and improved comprehension for medical scientists. Real-time prediction, interactive plots, and step-by-step explanations build a powerful diagnostic platform. In terms of usability, latency was assessed to be less than 2 seconds for image prediction and less than 4 seconds for visualization of tractography on mid-range machines.

VI. CONCLUSION

This project addresses important neuroimage analysis issues by foregrounding advanced computation tools into a single and accessible environment. By harmonizing tractography, FA (Fractional Anisotropy) analysis, and machine learning into a streamlined interface, the system automates the otherwise complicated process of diagnosing and analyzing neurodegenerative diseases. Automated processes like image preprocessing, prediction, and visualization of tracts considerably decrease reliance on human interpretation and specialized facilities.

One of the most important strengths of the platform is user-focused design. The platform is constructed with Streamlit, making it possible for non-expert users—like students or family physicians—to work with good-quality outputs. The NeuroBot assistant then closes the knowledge gap by transcribing technical outcomes in plain language. This design facilitates large-scale adoption in clinical and academic settings, as well as being a teaching tool for data science and medical students.

A second essential advantage is the open-source and modular nature of the platform. Implemented with widely adopted libraries such as DIPY, TensorFlow, Scikit-learn, and NiBabel, the system is reproducible and customizable. Developers and researchers can add more support for other neurological diseases, imaging modalities, or hospital workstreams to the pipeline. The small and deployable model also allows it to be easily deployed in resource-constrained healthcare environments.

Future evolution will emphasize integration of Brainalyze with hospital PACS systems to further streamline clinical workflows. Further multi-disease classification and severity grading models will be added to provide further diagnostic utility. The NeuroBot assistant will also be enhanced for multi-language support and conversational AI capabilities to make it more accessible to users from different linguistic backgrounds. These updates will further establish the platform's position in reshaping neuroimage analysis in healthcare and education.

REFERENCES

[1]. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, Apr. 1955.
[2]. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp. 68–73.

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 12, Issue 6, June 2025
DOI: 10.17148/IARJSET.2025.12652

[3]. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
[4]. R. Nicole, “Title of paper with only first word capitalized,” *J. Name Stand. Abbrev.*, in press.
[5]. M. Young, *The Technical Writer’s Handbook*, Mill Valley, CA: University Science, 1989.
[6]. D. Garyfallidis et al., “DIPY, a library for the analysis of diffusion MRI data,” *Frontiers in Neuroinformatics*, vol. 8, p. 8, 2014.
[7]. M. Brett, M. Hanke, and B. Markiewicz, “nibabel: Access a cacophony of neuro-imaging file formats,” [Online]. Available: https://nipy.org/nibabel
[8]. F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[9]. M. Abadi et al., “TensorFlow: A system for large-scale machine learning,” in *Proc. 12th USENIX Conf. Operating Systems Design and Implementation (OSDI)*, 2016, pp. 265–283.
[10]. S. Wang et al., “Deep learning for identifying metastatic breast cancer,” *arXiv preprint arXiv:1606.05718*, 2016.
[11]. R. Jenkinson et al., “FSL,” *Neuroimage*, vol. 62, no. 2, pp. 782–790, 2012.
[12]. Alzheimer’s Association, “2023 Alzheimer’s disease facts and figures,” *Alzheimer’s & Dementia*, vol. 19, no. 4, pp. 1–98, 2023.
[13]. World Health Organization, “Neurological disorders: public health challenges,” Geneva: WHO, 2023.
[14]. H. Sudre et al., “Machine learning-assisted diagnosis of Alzheimer’s disease using structural MRI: a review,” *Frontiers in Aging Neuroscience*, vol. 11, p. 116, 2019.
[15]. T. Schirner et al., “Inferring multi-scale neural mechanisms with brain network modelling,” *eLife*, vol. 11, e68322, 2022.