IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14440

Brain Age Estimation

Sahil Bendugade[1], Harsh Hate[2], Sahil Jadhav[3], Aryan Nangre[4], Shilpali Bansu[5]

Computer Department, University of Mumbai, Navi Mumbai, India[1-4]

AI-DS Department, University of Mumbai, Navi Mumbai, India[5]

Abstract: Brain age estimation is an emerging field in medical imaging, particularly useful for detecting neurological diseases and age-related cognitive decline. This project aims to develop a robust model for predicting brain age using T1-weighted MRI scans. By analyzing the structural patterns within these scans, the model will estimate the biological age of a patient's brain. The deviation between the predicted brain age and the chrono- logical age may indicate the presence of neurological diseases such as Alzheimer's, Parkinson's, or other neurodegenerative conditions. The project will leverage deep learning algorithms to process MRI data and predict brain age accurately. Various preprocessing steps will be applied to ensure high-quality input for the model, and advanced neural network architectures will be utilized for prediction. The ultimate goal is to provide a tool that aids in early diagnosis of neurological conditions by identifying patients whose brains show signs of accelerated aging. This system, if effective, can enhance early detection and interven- tion strategies, improving patient outcomes and contributing to personalized healthcare solutions.

I. INTRODUCTION

The "Brain Age Estimation" project aims to predict an indi- vidual's brain age based on neuroimaging data using advanced machine learning techniques. By analyzing brain structure and function, this project seeks to estimate the biological age of the brain, which can differ from chronological age due to various factors like health conditions or cognitive decline. Accurately estimating brain age can provide valuable insights into neurological health, helping in early detection of brain disorders and offering potential markers for aging-related research.

Brain age estimation is a technique used to predict the biological age of an individual's brain. This may differ from chronological age. This estimation is often done using ma- chine learning models. They are typically trained on brain imaging data such as MRI scans, features such as the thickness of the cerebral cortex. Brain volume and the in- tegrity of the white matter It has been used to predict the age of the brain. A brain age that is significantly longer than a person's actual age can indicate a potential risk for neurodegenerative disease. decreased awareness or other health problems.

This method relies on machine learning models trained on large data sets of brain scanning with marked ages. Functions extracted from scans - such as cortical thickness, brain volume and white matter integrity - when creating predictions. If the Identify applicable funding agency here. If none, delete this. estimated age of the brain is significantly higher than the chronological age, this may indicate early signs of neurodegen- erative diseases (eg Alzheimer's disease, Parkinson) or other health conditions such as stress and cardiovascular disease.

Brain age estimate has an application in medical diagnosis, aging research and personalized medicine, allowing timely intervention and monitoring of brain health over time.

II. MOTIVATION

The human brain undergoes extensive changes through- out life, from development in youth to degeneration with age. Understanding and estimating brain age is important because it provides insight into neurological health. While chronological age is an important determinant, brain age can often reflect underlying cognitive impairment or abnormal brain function, leading to the diagnosis of neurological diseases such as Early Alzheimer's disease, dementia, and other age- related conditions.

Advances in neuroimaging and machine learning have opened up new ways to accurately age the brain by analyzing structural and functional data A key motivation for this work is to harness this cutting-edge technology though assess brain age, and can identify individuals at risk for cognitive decline or accelerated brain aging Co,Expenditure- Provides an effective, efficient method

DOI: 10.17148/IJARCCE.2025.14440

III. LITERATURE SURVEY

The estimate of the age of the brain has been largely studied in neuroscience and medical imaging, with various techniques developing over time. Key progress can be divided into various approaches, including traditional statistical methods, machine learning and deep learning.

- **In The Paper titled ”Optimal Transport-Based Fea- ture Pyramid Fusion Network for Brain Age Estimation with 3D Overlapped ConvNext” [1]** , the authors propose the OTFPF network as an end-to-end neural network architecture designed for accurate and effective brain age estimation using T1-weighted MRI scans. The core algorithm employed in this study is Sinkhorn’s algo- rithm, which plays a vital role in the optimal transport- based fusion process. The model demonstrates promis- ing performance in estimating brain age with improved accuracy. This work was published in *Medical Image Analysis* in the year 2021.

- **In the paper titled ”Deep Learning for Brain Age Estimation: A Systematic Review” [2]** , the authors explore various approaches utilizing Deep Neural Networks (DNNs) for brain age estimation. The primary objective of this study is to estimate brain age, which serves as an indicator of brain health by comparing the predicted age with the individual’s actual chronological age. This com- prehensive review highlights the effectiveness of DNN- based models in capturing neurobiological aging patterns. The paper was published by *Springer Nature* in the year 2022.

- **In the paper titled ”Multimodal Brain Age Es- timation Using Interpretable Adaptive Population-Graph Learning” [3]** , the authors present an end-to-end pipeline for adaptive population-graph learning tailored specifically for brain age estimation. The methodology integrates Multilayer Perceptrons (MLPs) and Graph Convolutional Networks (GCNs) to effectively model multimodal brain data while maintaining interpretability. This approach enhances the estimation accuracy by lever- aging the structural relationships within population data. The study was published by *Imperial College London, UK*.

- **In the paper titled ”Deep Learning for Brain Age Esti- mation: A Systematic Review” [4]**, the authors propose a deep learning model that combines a Convolutional Neural Network (CNN) with a Multi-Layer Perceptron (MLP) to enhance brain age prediction. The model esti- mates brain age and compares it with chronological age to determine the brain age gap, which serves as a potential biomarker for identifying neurodegenerative conditions. This research was published in *Scientific Reports* by
Nature Publishing Group, Volume 13, Article number 22388, in the year 2023.

- **In the paper titled ”Brain Age Gap Estimation Using Attention-Based ResNet Method for Alzheimer’s Dis- ease Detection” [5]** , the authors employ a combination of Relevance Vector Regression and Twin Support Vector Regression to predict brain age gaps. These gaps are used as critical indicators for detecting Alzheimer’s disease. The proposed method, based on an attention-enhanced ResNet architecture, achieves an impressive accuracy rate of 92%. This study was published in the journal *Brain Informatics*, Volume 11, Article Number 16, in the year 2024.

- **In the paper titled ”A Deep Learning Model for Brain Age Prediction Using Minimally Preprocessed T1w Images as Input” [6]** , the authors present a CNN- based approach for predicting biological brain age. The model utilizes raw T1-weighted MRI images registered to the MNI space, aiming to ensure accessibility and ease of implementation. Techniques such as FreeSurfer 6.0.0 and SmoothGrad were employed in the pipeline to enhance interpretability and performance. This study was published in the journal *Aging Neuroscience* in the year 2024.

IV. PROBLEM STATEMENT

Develop a machine learning model to estimate a person’s brain age from neuroimaging data. The goal is to predict the biological age of the brain, which may differ from chrono- logical age, to identify potential early signs of neurological diseases or cognitive decline.

V. PROPOSED SYSTEM

- Data Collection and Preprocessing: Gather a diverse and comprehensive dataset of neuroimaging data, including T1-weighted MRI scans, from multiple cohorts with de- tailed clinical information. Preprocess the data to ensure uniformity and quality, including registration to a stan- dard template (e.g., MNI space) and minimal processing steps to maintain the original image integrity.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14440

- Model Development and Training: Utilize deep learning frameworks such as PyTorch or TensorFlow to develop convolutional neural network (CNN) models for brain age estimation. Train the models using the collected and pre- processed neuroimaging data, incorporating techniques like hold-out validation and cross-validation to assess model performance.
- Model Evaluation and Validation: Evaluate the trained models using external test sets and diverse cohorts to as- sess generalizability and robustness across different pop- ulations. Validate the model's performance in predicting biological brain age and compare it with chronological age to ensure accuracy and reliability.
- Implementation and Accessibility: Make the developed CNN-based brain age estimation model publicly available for use in research and potentially clinical settings. En- sure the model's accessibility and ease of implementation by minimizing preprocessing steps and providing clear documentation for users.

A. System Design and Architecture

Brain age estimation is a program designed to estimate brain age, typically using neu- roimaging data (such as an MRI scan) and machine learning algorithms. Estimated brain age is compared chronologically to assess neural health. Here is a brief layout process:

- Data collection : Neuroimaging data (such as MRI) are collected from the sub- jects, along with their chronological age.
- Preprocessing : Neuroimaging data undergo preprocessing steps such as noise reduction, normalization, and brain split- ting to extract relevant features.
- Feature emission: Important features (e.g., gray matter vol- ume, cortical thickness) are removed from images to represent brain structure and function.
- Model selection : Machine learning algorithms such as regression, convolutional neural networks (CNNs), or deep learning models are trained on these features to predict brain age.
- Training : The model is trained with labeled data with known target age sequences. Methods such as cross-validation are used to avoid overfitting. • Evaluation : Model performance is evaluated using metrics such as the mean absolute error (MAE) between predicted brain age and actual chronological age.
- Postprocessing : Corrections or adjustments (e.g., bias correction) are applied to estimate model biases or systematic errors in forecasts.
- Deployment : The final model can be incorporated into clinical tools to assess brain health in patients or to conduct research studies.

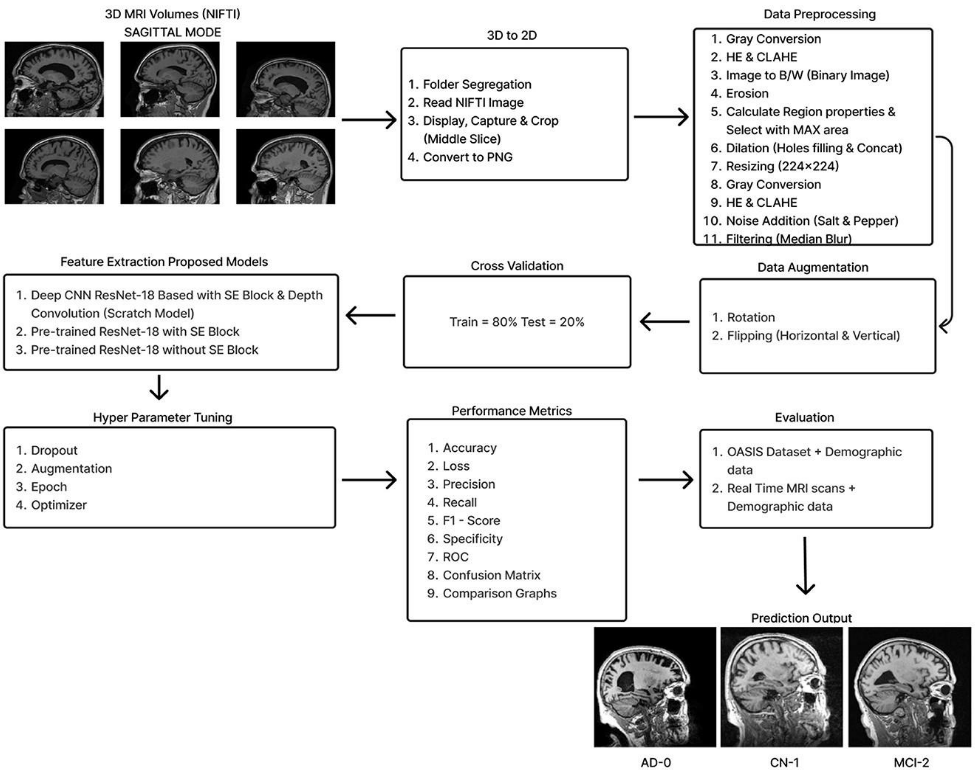

Fig. 1. Architecture design

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14440

B. System Architecture Overview

The proposed system for brain age estimation utilizes deep learning on sagittal-view 3D MRI scans, focusing on prepro- cessing, model training, and performance evaluation. Fig. **??** illustrates the complete workflow.

1) Data Acquisition and Conversion: Raw 3D MRI scans (in NIfTI format) are used as the primary input. These volumes are processed in the sagittal plane to extract 2D slices. Each image undergoes folder-wise segregation, mid-slice extraction, cropping, and conversion to PNG format for subsequent pro- cessing.

2) Preprocessing Pipeline: The preprocessing involves sev- eral sequential image processing steps:

- Grayscale conversion, histogram equalization (HE), and CLAHE.
- Binarization, erosion, and region selection using maxi- mum area properties.
- Morphological operations like dilation for hole filling, and resizing to a uniform resolution (224×224).

Additional steps such as salt-and-pepper noise injection and median filtering are applied to improve robustness.

3) Data Augmentation and Cross-Validation: To mitigate overfitting, training data is augmented using horizontal/vertical flipping and image rotation. An 80/20 split is used for training and testing during cross-validation.

4) Modeling and Feature Extraction: Three model variants based on ResNet-18 are evaluated:

1) A custom deep CNN ResNet-18 with SE block and depthwise convolution (scratch training).
2) Pre-trained ResNet-18 with Squeeze-and-Excitation (SE) block.
3) Pre-trained ResNet-18 without SE block. Hyperparameter tuning includes experimenting with dropout rates, epochs, augmentation strategies, and optimizers.

5) Evaluation and Metrics: Model performance is eval-
uated using standard classification metrics: accuracy, loss, precision, recall, F1-score, specificity, ROC curves, and confu- sion matrices. Comparative graphs are generated for analysis. Evaluation uses the OASIS dataset and real-time MRI scans with associated demographic metadata.

6) Prediction and Output: The system outputs a classi- fication for brain age categories such as AD (Alzheimer's Disease), CN (Cognitively Normal), and MCI (Mild Cognitive Impairment), based on the processed MRI inputs.

7) Continuous Learning: To maintain performance over time, the system includes mechanisms for incorporating new data and retraining the models periodically, supporting scala- bility and adaptability for clinical environments.

This architecture combines rigorous image preprocessing, deep learning-based feature extraction, and evaluation on real- world datasets to deliver accurate and explainable brain age predictions.

C. Data Flow Diagram

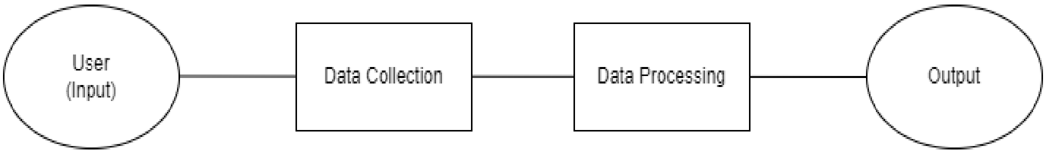

Fig. 2. Data Flow Diagram Level 0

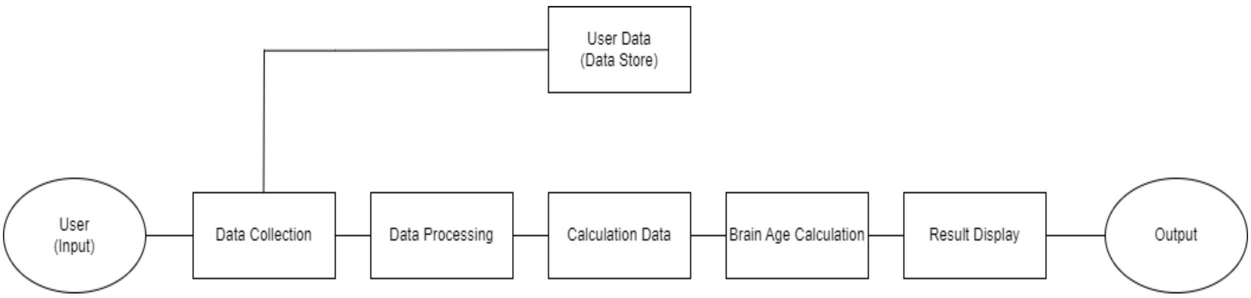

Fig. 3. Data Flow Diagram Level 1

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025
DOI: 10.17148/IJARCCE.2025.14440

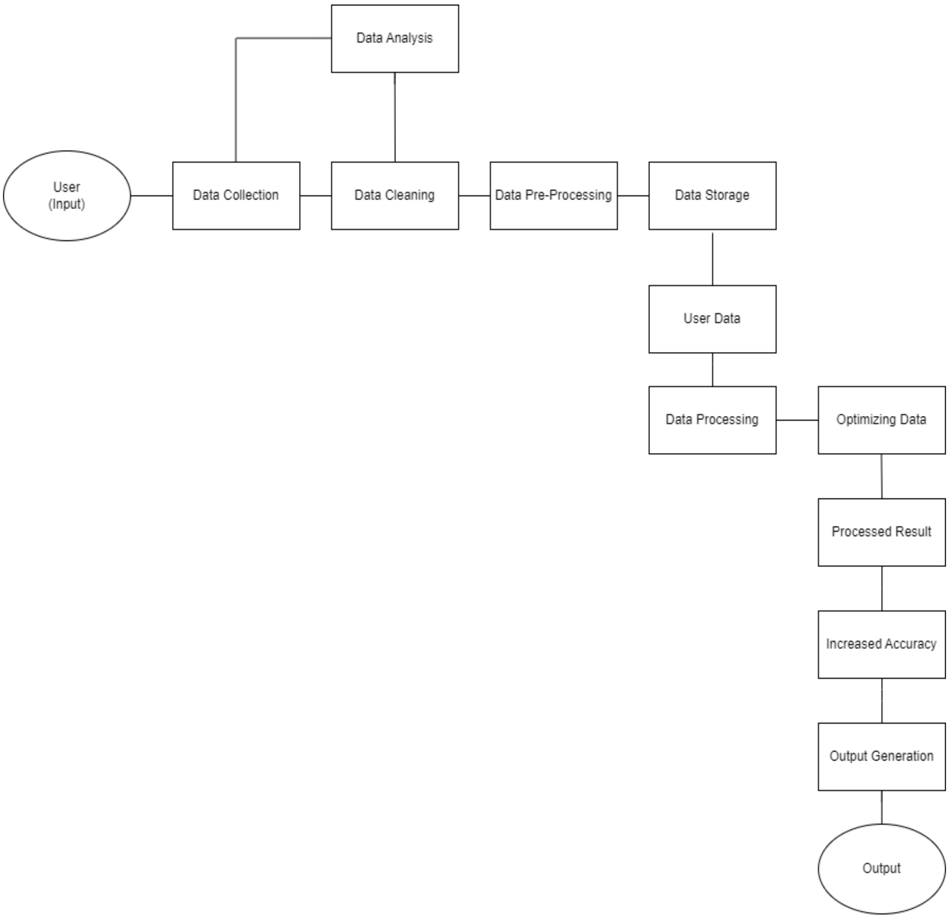

Fig. 4. Data Flow Diagram Level 2

VI. HARDWARE AND SOFTWARE REQUIREMENTS

A. *Software Requirements*

- Programming Language : Python (for data analysis and model development).
- Libraries and Frameworks:

NumPy, Pandas (for data manipulation) Scikit-learn (for machine learning models)
TensorFlow or PyTorch (for deep learning models) OpenCV (for image processing, if using neuroimaging data) Matplotlib, Seaborn (for data visualization) Nibabel or other neuroimaging libraries (for handling MRI or fMRI data).

- Development Environment : Jupyter Notebook or any Python IDE (e.g., PyCharm, VS Code)
- Data Storage and Management : Local storage for smaller datasets Cloud services (e.g., AWS S3, Google Cloud Storage) for larger datasets.
- Version Control : Git and GitHub or GitLab for version control and collaboration.

B. *Hardware Requirements*

- Computing Resources : A high-performance laptop or desktop with at least 8GB of RAM and a multi-core CPU Preferably, a system with a dedicated GPU (e.g., NVIDIA) for deep learning model training.
- Storage : At least 100GB of available storage space for dataset storage and processing.

Cloud Computing Resources (optional but recommended for large datasets or complex models) : Access to cloud-based GPU/TPU instances from platforms like AWS,Google Cloud, or Azure.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⸎ Peer-reviewed & Refereed journal ⸎ Vol. 14, Issue 4, April 2025
DOI: 10.17148/IJARCCE.2025.14440

VII. SYSTEM DEVELOPMENT AND OUTPUT

A. Algorithms

- **VGG-16 CNN Model :**

A Convolutional Neural Network (CNN) [7] architecture is a deep learning model designed for processing struc- tured grid-like data, such as images. It consists of mul- tiple layers, including convolutional, pooling, and fully connected layers. CNNs are highly effective for tasks like image classification, object detection, and image segmentation due to their hierarchical feature extraction capabilities.

- **SVR :**

Support Vector Regression (SVR) [8] is an extension of the Support Vector Machine (SVM) algorithm used for regression tasks. Unlike traditional linear regression, which minimizes the error between predicted and actual values, SVR aims to find a hyperplane (or line) that fits the data with a margin of tolerance, where most data points fall within this margin. The SVM algorithm creates a hyperplane with the largest gap between positive and negative instances in the feature space. For data that can be linearly separated, linear SVM is frequently employed. The SVR is a regression analysis model based on the SVM. [9]

- **CNN :**

Convolutional Neural Network [10] and Multi-layer Per- ceptron Algorithms. MLP is a feed- forward artificial neural network (ANN) that is trained using a back- propagation algo- rithm. An MLP is composed of input nodes at each layer that form a directed graph between the output and input layers. An MLP is a neural network that connects many layers in a directed graph, which means that the data signal is routed through the nodes of the graph in only one direction. [11] The search range was (0.00001, 0.00002, 0.00004, 0.00008, 0.00016, 0.00032, 0.00064, 0.00128, 0.00256, 0.00512) for learning-rate.

- **XgBoost and CatBoost :**

XgBoost: XgBoost is an implementation of gradient boosted decision trees. which has frequently appeared in winning solutions in Kaggle competitions. In order to get the best XgBoost model, grid-searching was performed on parameter combinations of learning-rate, max- imum depth of the tree, and number of estimators. The search range was: (100, 300, 500, 1000), (2, 10, 1), and (0.01, 0.03, 0.05, 0.1, 0.3, 0.5) for the number of estimators, maximum depth of the tree and learning rate respectively.

CatBoost: CatBoost is also an algorithm for gradient boosting on decision trees. It offers a new strategy for handling categorical features that can address the gradient bias and predic- tion shift issues. Grid-searching was used in learning-rate, maximum depth of the tree, and number of estimators were estimated. The search range was (100, 300, 500, 1000), (2, 10, 1), and (0.01, 0.03, 0.05, 0.1, 0.3, 0.5) for the number of estimators, maximum depth of the tree, and learning-rate, respectively. [12]

B. Output/Result

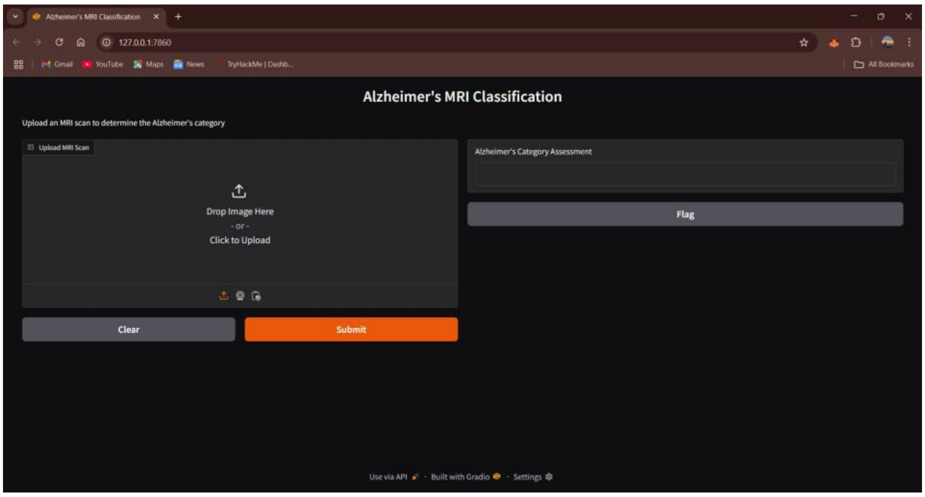

Fig. 5. Interface

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⪫ Peer-reviewed & Refereed journal ⪫ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14440

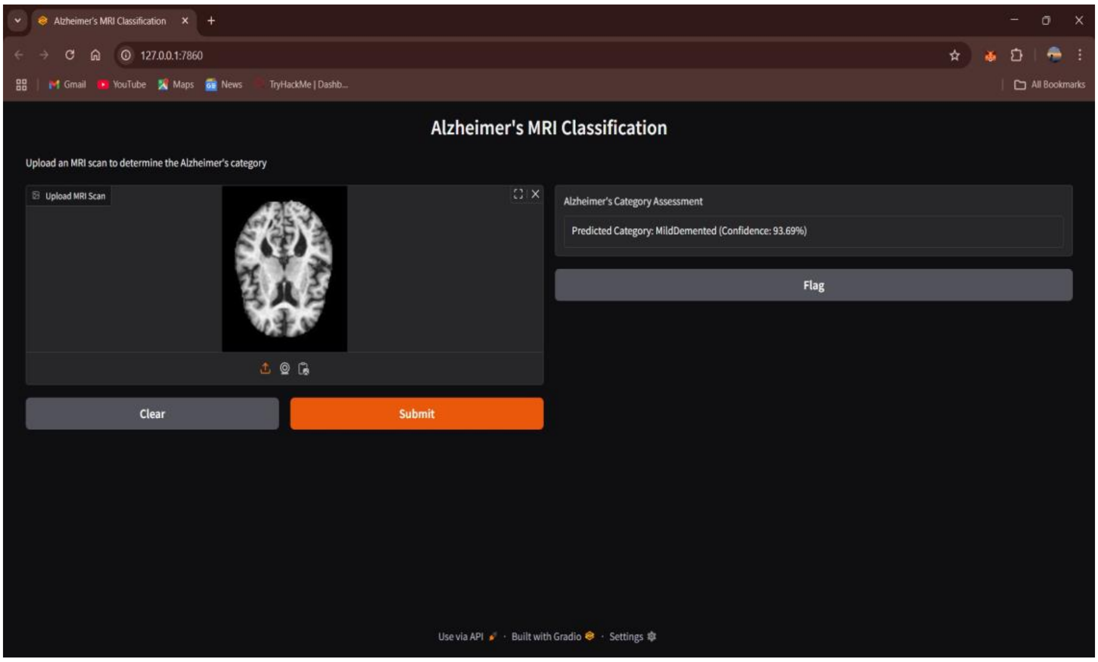

Fig. 6. Uploaded Image

VIII. FUTURE WORK

Multi-Modal Data Integration Explore the integration of multi-modal neuroimaging data, such as functional MRI (fMRI) and diffusion tensor imaging (DTI), to capture a more comprehensive understanding of brain aging and improve the accuracy of age estimation models.

Explainable AI (XAI) Integrate explainable AI techniques to provide insights into the features and regions of interest driving the age predictions, enhancing the interpretability and trustworthiness of the model outputs for clinical and research applications. Lifelong Learning and Adaptive Systems In- vestigate the development of adaptive brain age estimation systems that can continuously learn from new data, adapt to evolving imaging technologies, and dynamically update model parameters to reflect changes in brain aging patterns over time.

IX. CONCULSION

In this project, we have laid the groundwork for a compre- hensive brain age estimation system utilizing neuroimaging data, particularly T1-weighted MRI scans. Through metic- ulous data collection and preprocessing, we have ensured that our dataset is diverse and of high quality, which is essential for training robust machine learning models. Our systematic approach to model development and evaluation has prioritized accuracy and generalizability, which are crucial for the clinical applicability of our findings.

The proposed system integrates advanced deep learning techniques, leveraging convo- lutional neural networks (CNNs) to analyze neuroimaging data. We have outlined a clear work- flow that includes essential steps such as data preprocessing, feature extrac- tion, model training, and evaluation. This struc- tured approach enables the model to learn meaningful patterns in the data, ultimately providing reliable predictions of brain age that can be compared to chronological age. Furthermore, the planned accessibility of our developed model will facilitate its use in both research and clinical settings, allowing health- care professionals to gain insights into neurological health and potentially aiding in the early detection of cognitive decline or neurodegenerative diseases.

As we move into the next phase of our project in Semester VIII, we will focus on the execu- tion of our proposed system. This will involve the implementation of the trained model, conducting further evaluations, and refining the model based on real-world data. Con- tinuous improvement and feedback integration will be vital as we seek to enhance the model's accuracy and applicability in various populations. Overall, this project not only aims to advance the field of neuroimaging and machine learning but also to contribute to our understanding of brain health and aging, paving the way for future research and clinical interventions.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14440

REFERENCES

[1]. Yu Fu, Yanyan Huang, Yalin Wang, Shunjie Dong, Le Xue, Xunzhao Yin, Qianqian Yang, Yiyu Shi, and Cheng Zhuo. Otfpf: Optimal transport-based feature pyramid fusion network for brain age estimation with 3d overlapped convnext. *arXiv preprint arXiv:2205.04684*, 2022.

[2]. Kaizhu Huangg Laif, Yu-Dong Zhangh, Javier Del Seri, and Chin-Teng Link. Deep learning for brain age estimation: A systematic review. *arXiv preprint arXiv:2212.03868*, 2022.

[3]. Kyriaki-Margarita Bintsi, Vasileios Baltatzis, Rolandos Alexandros Potamias, Alexander Hammers, and Daniel Rueckert. Multimodal brain age estimation using interpretable adaptive population-graph learning. In *International Conference on Medical Image Computing and Computer- Assisted Intervention*, pages 195–204. Springer, 2023.

[4]. Muhammad Tanveer, MA Ganaie, Iman Beheshti, Tripti Goel, Nehal Ahmad, Kuan-Ting Lai, Kaizhu Huang, Yu-Dong Zhang, Javier Del Ser, and Chin-Teng Lin. Deep learning for brain age estimation: A systematic review. *Information Fusion*, 96:130–143, 2023.

[5]. Atefe Aghaei, Mohsen Ebrahimi Moghaddam, and Alzheimer's Dis- ease Neuroimaging Initiative. Brain age gap estimation using attention- based resnet method for alzheimer's disease detection. *Brain Informat- ics*, 11(1):16, 2024.

[6]. Caroline Dartora, Anna Marseglia, Gustav Ma˚rtensson, Gull Rukh, Junhua Dang, J-Sebastian Muehlboeck, Lars-Olof Wahlund, Rodrigo Moreno, Jose´ Barroso, Daniel Ferreira, et al. A deep learning model for brain age prediction using minimally preprocessed t1w images as input. *Frontiers in Aging Neuroscience*, 15:1303036, 2024.

[7]. N Deepa and SP Chokkalingam. Optimization of vgg16 utilizing the arithmetic optimization algorithm for early detection of alzheimer's disease. *Biomedical Signal Processing and Control*, 74:103455, 2022.

[8]. MA Ganaie, M Tanveer, and Iman Beheshti. Brain age prediction using improved twin svr. *Neural Computing and Applications*, 36(1):53–63, 2024.

[9]. Iman Beheshti, MA Ganaie, Vardhan Paliwal, Aryan Rastogi, Imran Razzak, and Muhammad Tanveer. Predicting brain age using machine learning algorithms: A comprehensive evaluation. *IEEE Journal of Biomedical and Health Informatics*, 26(4):1432–1440, 2021.

[10]. Yoonji Joo, Eun Namgung, Hyeonseok Jeong, Ilhyang Kang, Jinsol Kim, Sohyun Oh, In Kyoon Lyoo, Sujung Yoon, and Jaeuk Hwang. Brain age prediction using combined deep convolutional neural network and multi- layer perceptron algorithms. *Scientific Reports*, 13(1):22388, 2023.

[11]. Yoonji Joo, Eun Namgung, Hyeonseok Jeong, Ilhyang Kang, Jinsol Kim, Sohyun Oh, In Kyoon Lyoo, Sujung Yoon, and Jaeuk Hwang. Brain age prediction using combined deep convolutional neural network and multi- layer perceptron algorithms. *Scientific Reports*, 13(1):22388, 2023.

[12]. Min Xiong, Lan Lin, Yue Jin, Wenjie Kang, Shuicai Wu, and Shen Sun. Comparison of machine learning models for brain age prediction using six imaging modalities on middle-aged and older adults. *Sensors*, 23(7):3622, 2023.

