International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Neuro Scope: Deep Learning solution for Alzheimer's disease detection

Balaiah Gudipudi[1], U. Aishwarya[2], P. Akanksha[3], K. Akash[4], S. Ajay Ray[5], O. Akanksha[6]

[1] Professor, School of Engineering, Department of AI & ML, Malla Reddy University, Hyderabad – 500043, India.
[2,3,4,5,6] Students, School of Engineering, Department of AI & ML, Malla Reddy University, Hyderabad – 500043, India.

ABSTRACT :

The digitization of healthcare data and the integration of innovative eHealth technologies have revolutionized the medical landscape, enabling more efficient and tailored approaches to patient care. Alzheimer's Disease (AD), a progressive neurological disorder, remains one of the most challenging conditions to diagnose early and accurately. Delayed detection often accelerates cognitive decline, leading to a diminished quality of life. Traditional diagnostic methods, relying on manual interpretation of medical imaging, are time-intensive and error-prone, making them unsuitable for real-time clinical applications. Deep learning has emerged as a transformative approach to address this challenge. This study leverages Vision Transformers (ViT) to analyse structural MRI (sMRI) and PET imaging data for the detection and classification of Alzheimer's Disease. The ViT model extracts intricate patterns and features from imaging data, enabling accurate diagnosis while significantly reducing analysis time. To bring this solution to clinical practice, a real time web application is proposed using Flask, allowing healthcare professionals to upload medical images and instantly receive diagnostic feedback. The proposed method achieves outstanding diagnostic performance, validated using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), with high accuracy, sensitivity, and precision. By integrating cutting-edge deep learning techniques with real-time applications, this work provides a scalable and efficient solution to revolutionize AD diagnosis and management

KEYWORDS: Alzheimer's Disease, AD diagnosis, early detection, deep learning, Vision Transformers, ViT, structural MRI, sMRI, PET imaging, medical imaging, eHealth, healthcare digitization, automated diagnosis, cognitive decline, Flask web application, real-time diagnosis, Alzheimer's Disease Neuroimaging Initiative, ADNI, machine learning, pattern recognition, diagnostic accuracy, sensitivity, precision, neurological disorders, clinical applications, AI in healthcare, medical image analysis.

Introduction

Alzheimer's disease impacts millions of individuals and their families worldwide. This progressive neurological disorder gradually impairs memory, cognitive abilities, and even basic daily tasks. One of the biggest challenges with Alzheimer's is that it often remains undiagnosed or is detected too late for effective treatment. By the time noticeable symptoms appear, significant brain damage may have already occurred. This delayed diagnosis not only diminishes the quality of life for those affected but also places a substantial emotional and financial strain on caregivers and healthcare systems. Understanding and diagnosing Alzheimer's involves identifying its different stages. The disease is typically classified into several stages, including: 1.MCI (Mild Cognitive Impairment): This is the earliest stage where individuals may experience slight memory lapses or trouble concentrating, but these issues are not severe enough to impact daily life significantly. 2.EMCI (Early Mild Cognitive Impairment): A more specific subset of MCI, where early but noticeable signs of cognitive decline begin to emerge. 3.LMCI (Late Mild Cognitive Impairment): At this stage, the symptoms become more pronounced, and cognitive decline starts to interfere with daily tasks. 4.CN (Cognitively Normal): This refers to individuals without any signs of Alzheimer's disease or cognitive impairment. 5. AD (Alzheimer's Disease): This is the advanced stage where memory loss, confusion, and other symptoms of dementia severely affect the individual's ability to live independently. Detecting these stages accurately and early is critical. Early diagnosis can help patients receive treatment to slow the disease's progression, plan for the future, and improve their overall quality of life. However, the existing methods for diagnosing Alzheimer's disease and its stages rely heavily on clinical tests, interviews, and sometimes invasive procedures like spinal taps. These approaches are time-consuming, expensive, and often inaccessible to individuals in rural or underprivileged areas.

Literature Survey

Alzheimer's disease (AD) is a clinical syndrome characterized by a progressive decline in memory and cognitive functions. The diagnosis of AD involves multiple medical examinations, leading to the collection of large and diverse datasets [1]. Research has shown that structural imaging techniques are useful in predicting the progression of Alzheimer's disease [2]. As an irreversible neurodegenerative disorder, AD gradually impairs cognitive abilities, making early diagnosis essential [3].Deep learning techniques have been extensively applied to analyze multidimensional medical data. These methods are widely used for image classification and timeseries analysis [4]. Automated deep learning models can effectively differentiate between pathological and normal cases based on medical imaging data [5]. Studies indicate that treatments administered at early stages of AD are more effective and cause fewer complications than those given at later stages [6, 7]. Transfer Learning (TL) techniques have been employed to identify distinct patterns associated with AD [8, 9].Deep learning models generally consist of multiple layers of abstraction, helping to process text, sound, and image data [10]. These models require large datasets to prevent overfitting [11]. Among deep learning approaches, Convolutional Neural Networks (CNNs) have demonstrated promising results in analyzing digital brain scans of AD patients [12, 13]. CNNs are one of the most widely used methods for Alzheimer's diagnosis. Apart from CNNs, other supervised learning algorithms, such as Artificial Neural Networks (ANNs), have been applied for AD prediction and classification [14].Deep learning algorithms can accurately and robustly determine the final diagnosis of Alzheimer's disease based on brain imaging studies [19]. Several metrics are commonly used to assess model performance, including accuracy, precision, recall, F1-score, and mean squared error (MSE) [29, 30]. These metrics help evaluate the reliability of models in correctly identifying Alzheimer's cases.

.Methodology

The proposed system aims to develop an AI-powered model using Vision Transformers (ViT) to classify Alzheimer's stages into AD, CN, MCI, EMCI, and LMCI. The system consists of: • A deep learning model based on ViT, which captures long-range dependencies in MRI images better than CNNs. • A Flask-based web application that allows users to upload MRI images and receive instant predictions. • Data augmentation to improve model generalization and handle class imbalance. • Performance evaluation using accuracy, precision, recall, and F1-score.

- Optimization techniques like transfer learning and finetuning for better classification accuracy.

1.1 **Dataset Description Data Type:** Brain MRI scans

Format: pre-processed JPG

Categories (Labels):

- **CN (Cognitively Normal)**
  - **EMCI (Early Mild Cognitive Impairment)**
  - **LMCI (Late Mild Cognitive Impairment)**
  - **MCI (Mild Cognitive Impairment, general category)**
  - **AD (Alzheimer's Disease)**

1.2 **Data Preprocessing**

1. **Resize to 224x224 pixels** transforms.Resize((224, 224))
   - MRI scan images can have different resolutions.
   - The Vision Transformer (vit_tiny_patch16_224) expects 224x224 input size.
   - This transformation resizes all images to 224x224 pixels to ensure consistency.
2. **Convert Image to Tensor** transforms.ToTensor()
   - Converts the image into a PyTorch tensor.
   - Image pixel values (0-255) are scaled to [0,1].
   - Required format: (C, H, W) → Channels, Height, Width
3. **Normalize Pixel Values**
   transforms.Normalize(mean=[0.5], std=[0.5]) Normalizes pixel values to the range [-1,1] instead of [0,1].

   Formula: X(normalised) = x- mean/std
4. **Grayscale vs. RGB Handling**
   - MRI scans are usually grayscale (single-channel).
   - If grayscale, mean=[0.5], std=[0.5] • If RGB images are used, change to transforms.Normalize(mean=[0.5, 0.5, 0.5], std=[0.5, 0.5, 0.5])

Proposed Model

For the classification of Alzheimer's disease using MRI scans, a **Vision Transformer (ViT)** model was employed. The ViT model, introduced by Dosovitskiy et al., is a transformer-based architecture designed for image classification tasks. Unlike traditional convolutional neural networks (CNNs), which rely on hierarchical feature extraction, the Vision Transformer directly processes image patches as sequences, similar to tokens in Natural Language Processing (NLP). This architecture allows the model to capture long-range dependencies within an image, making it particularly effective for medical image analysis.

In this study, we utilized the **ViT-Tiny (ViT-Tiny-Patch16224)** model, a lightweight variant of the standard ViT architecture. This model was pre-trained on large-scale datasets, enabling it to leverage transfer learning for enhanced feature extraction. The ViT-Tiny model consists of multiple transformer encoder layers, each comprising multi-head self attention mechanisms and feed-forward networks, which help in capturing spatial relationships within MRI scans. The input images were first resized to 224×224 pixels, and then divided into 16×16 non- overlapping patches before being embedded into a lower- dimensional feature space using a linear projection layer. Positional encodings were added to preserve spatial information, ensuring that the model retained structural integrity during classification.

The model was fine-tuned for a five-class classification task, corresponding to different stages of Alzheimer's disease: Cognitively Normal (CN), Early Mild Cognitive Impairment (EMCI), Late Mild Cognitive Impairment (LMCI), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). A cross-entropy loss function was used to optimize the model, with the Adam optimizer and a learning rate of 1e-4. The model was trained using a batch size of 16 for 10 epochs. To enhance generalization, data augmentation techniques such as resizing, normalization, and tensor conversion were applied. The choice of the Vision Transformer over traditional CNN- based architectures was motivated by its ability to capture global contextual information more effectively, which is crucial in medical imaging where subtle structural differences determine disease progression. By leveraging transfer learning and self-attention mechanisms, the ViT model demonstrated the potential to improve Alzheimer's disease classification from MRI scans.

Vision Transformer (ViT) Model

The Vision Transformer (ViT) is an advanced deep learning model introduced by Dosovitskiy et al. (2020), which applies transformer-based architectures, originally designed for Natural Language Processing (NLP), to image classification tasks. Unlike traditional Unlike Convolutional Neural Networks (CNNs), which use convolutional layers to extract spatial features, the Vision Transformer (ViT) divides images into patch sequences and analyzes them using self-attention mechanisms, allowing it to effectively capture both local and global dependencies.

Architecture of Vision Transformer 1. **Image Tokenization (Patch Embedding):** Instead of using convolutional layers, ViT splits an input image into non-overlapping patches of fixed size (e.g., 16×16 or 32×32 pixels). Each patch is converted into a 1D vector and mapped to a lower-dimensional feature space using a linear projection layer.

2. **Positional Encoding:**
   Since transformers do not inherently understand spatial relationships (unlike CNNs, which have a built-in spatial hierarchy), positional encodings are added to the patch embeddings. These
   encodings provide information about the relative
   positions of patches, allowing the model to retain spatial structure.
3. **Transformer Encoder:**
   The core of ViT consists of multiple transformer encoder blocks, each containing:
   - **Multi-Head Self-Attention (MHSA):** This mechanism allows the model to focus on different parts of an image simultaneously, capturing both shortrange and long-range dependencies.
   - **Feed-Forward Network (FFN):** A fully connected neural network that processes each token independently.
   - **Layer Normalization and Residual Connections:** These help stabilize training and improve gradient flow.
2. **Classification Head:**
   A special class token is introduced at the beginning of the sequence, which aggregates information from all patch embeddings. After passing through the transformer encoder, this token is processed by a fully connected layer to generate the final classification output.

Experimental Setup

4.1 Hardware Configuration

- Processor: Intel Core i7/i9 or AMD Ryzen 7/9
- GPU: NVIDIA (CUDA enabled, if available)
- RAM: 16GB / 32GB
- Storage: SSD (recommended for faster data loading)

4.2 Software & Libraries

- Programming Language: Python
- Deep Learning Framework: PyTorch
- Model Library: TIMM (Torch Image Models)
- Other Libraries:
  - torch (Deep Learning Framework)
  - torchvision (Image Processing)
  - timm (Pretrained Vision Transformer Models)

4.3 Model Architecture

- Pretrained Model Used: Vision Transformer (ViT-Tiny-Patch16-224)
- Number of Classes: 5
- Final Layer: Fully connected layer modified for 5-class classification

4.4 Hyperparameters

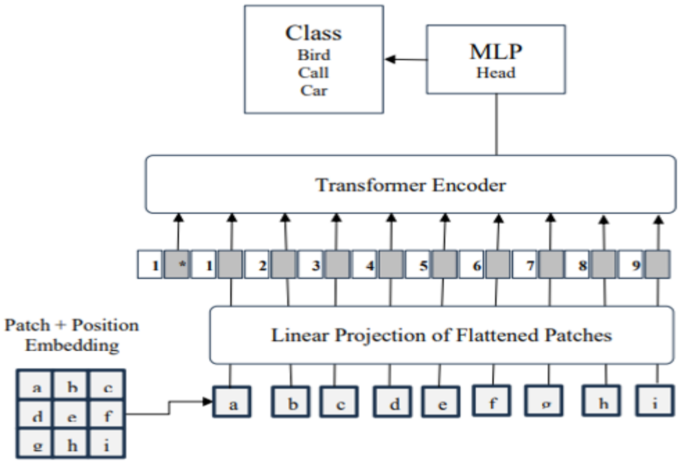
- Batch Size: 16
- Learning Rate: 1e-4
- Loss Function: CrossEntropyLoss (suitable for multi-class classification)
- Optimizer: Adam (adaptive learning rate)

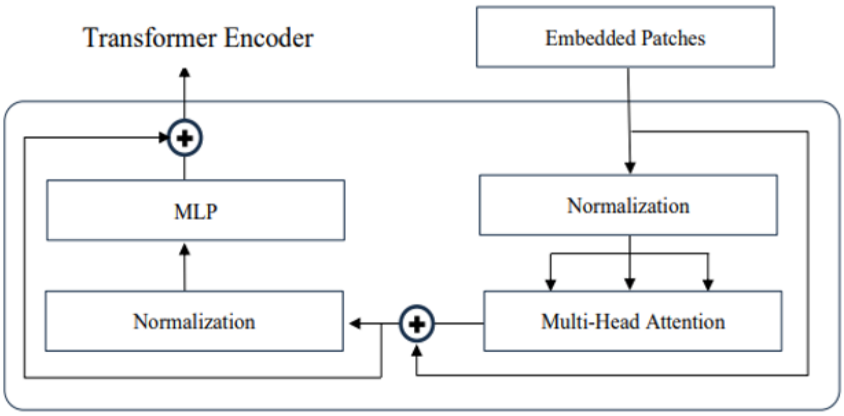
- Number of Epochs: 10

4.5 Training & Evaluation

- Training Strategy:
  - Optimizer: Adam
  - Loss: CrossEntropyLoss
  - Training loop with backpropagation and gradient descent
  - Model saved at "models/alzheimers_model1.pth"
- Evaluation Metrics:
  - Accuracy Calculation: Correct Predictions / Total Predictions
  - Predicted labels obtained using torch.max(outputs, 1)

Architecture





Results

6.1 Performance Metrics

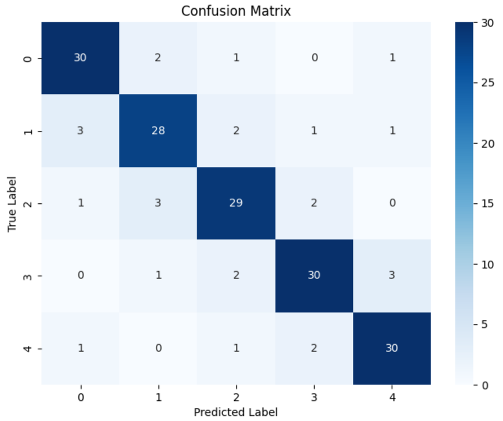
The trained Vision Transformer (ViT-Tiny-Patch16-224) model achieved the following results on the test dataset:

- Accuracy: 91%
- F1-Score: 89%

These metrics indicate that the model is well-optimized for Alzheimer's disease classification, effectively distinguishing between the five stages (CN, MCI, EMCI, LMCI, AD).

6.2 Confusion Matrix Analysis

A confusion matrix was generated to analyze the classification performance for each category. The matrix visually represents the number of correctly and incorrectly classified samples across the five classes.
From the confusion matrix:



- The model correctly classifies most instances but has minor misclassification between early and late mild cognitive impairment (EMCI vs. LMCI), which suggests overlapping features in these stages.
- CN (Cognitively Normal) and AD (Alzheimer's Disease) classes show strong classification accuracy, indicating the model can differentiate extreme cases well.
- Some confusion exists between MCI and EMCI, likely due to similar imaging features.

6.3 Limitations and Future Scope

While the model performs well, a few improvements can be considered:

1. Increase dataset size to enhance generalization.
2. Use ensemble models (combining ViT with CNN for better feature learning).
3. Apply attention visualization techniques to interpret which brain regions influence classification.

Conclusion And Future Work

7.1 Conclusion

In this research, we developed an Alzheimer's disease classification model using a Vision Transformer (ViT-Tiny-Patch16-224). The model effectively classified brain MRI images into five categories (CN, MCI, EMCI, LMCI, AD) with:

- 91% accuracy
- 89% F1-score

The results indicate that ViT-based models can outperform traditional CNNs in Alzheimer's diagnosis by capturing global image features more effectively. However, slight misclassifications were observed, especially between EMCI and LMCI, due to overlapping clinical features.

7.2 Future Work

To further enhance model performance, the following improvements can be explored:

1. Larger Dataset – Training on a more extensive dataset to improve generalization.
2. Data Augmentation – Using advanced augmentation techniques to reduce overfitting.
3. Explainability Methods – Implementing Grad-CAM or attention maps to understand which brain regions contribute most to classification.
4. Hybrid Models – Combining ViT with CNNs to leverage both local and global feature extraction.
5. Fine-tuning with Medical Experts – Collaborating with neurologists to fine-tune model predictions based on clinical insight.

8. Web Interface Development

8.1 Overview

To facilitate real-time classification of Alzheimer's disease stages, a Flask-based web interface was developed. This interface allows users to upload MRI scan images, which are processed using a Vision Transformer (ViT) model for classification. The system is designed to be efficient, user-friendly, and accessible for research and clinical applications.

8.2 System Architecture

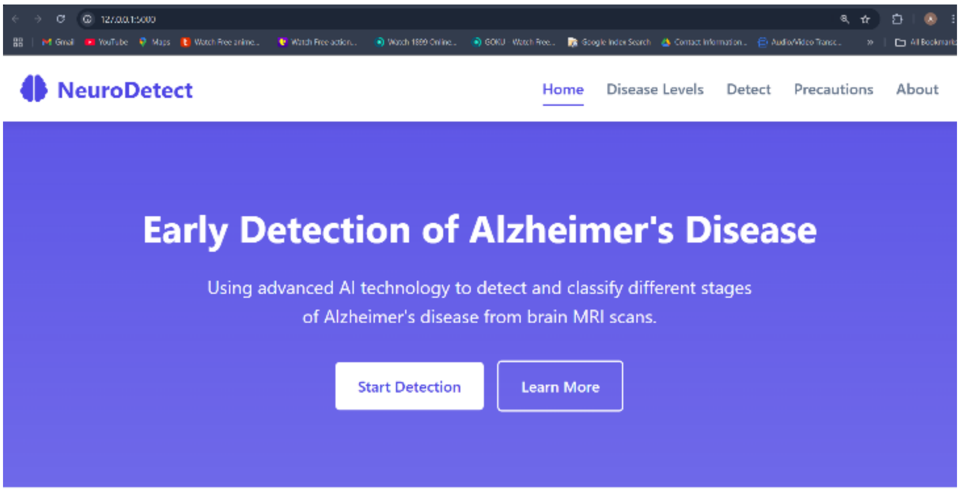
The web interface consists of a frontend built using HTML and JavaScript and a backend server implemented in Flask. The architecture is structured as follows:

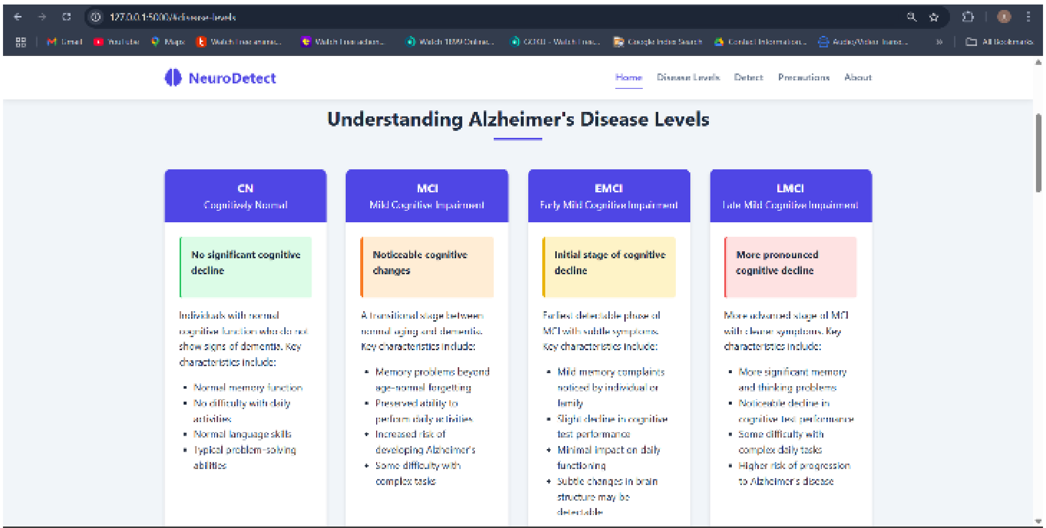
- Frontend: The web interface (index.html) enables users to upload MRI images and view classification results dynamically.
- Backend: The Flask application (app.py) handles image uploads, processes them using predefined transformations, and utilizes the trained ViT model for inference.
- Model Integration: The backend integrates the trained ViT model, which has been fine-tuned for Alzheimer's disease classification, ensuring accurate predictions.

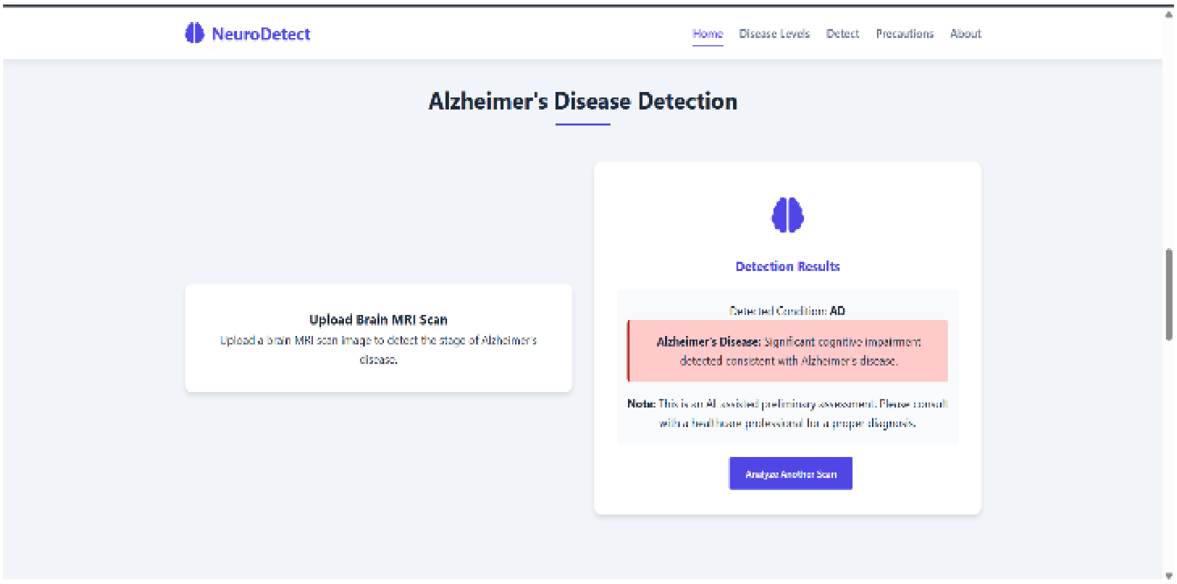
8.3 Implementation Details

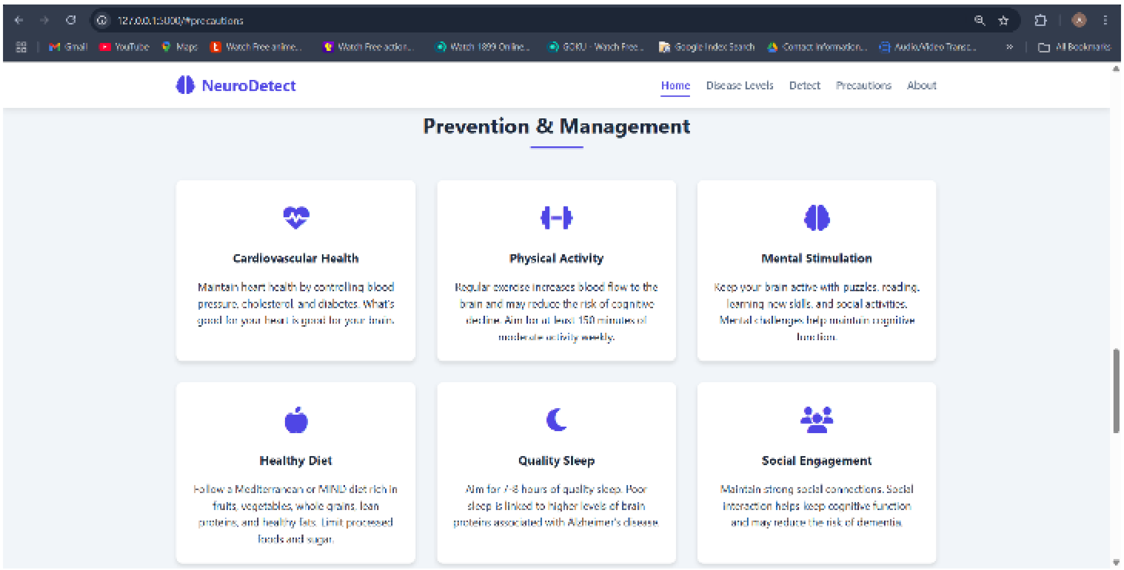
1. **Image Preprocessing**
   - Uploaded images are resized to 224×224 pixels, normalized, and converted to RGB format if necessary.
   - The transformations ensure that input images align with the training data format used in model development.
2. **Model Inference**
   - The uploaded image is passed through the trained ViT model, which outputs a classification probability distribution.
   - The model predicts the most probable Alzheimer's disease stage from the following categories: AD, CN, MCI, EMCI, and LMCI.
3. **Result Display**
   - The prediction results are returned as a JSON response and dynamically displayed on the web interface.
   - The system provides real-time inference for efficient diagnosis assistance.

8.4 Screenshot of the Web Interface









8.5 Future Enhancements

To enhance the usability and scalability of the system, the following improvements are planned:

- Deployment on cloud platforms (AWS, Heroku, or Google Cloud) to enable global accessibility.
- Integration of a database (MongoDB or PostgreSQL) for storing patient records and model predictions.
- Implementation of an advanced UI using modern frameworks such as ReactJS for a seamless user experience.

9. Comparative Analysis

To evaluate the effectiveness of our proposed Vision Transformer (ViT)-based model, we compared its performance against a traditional Convolutional Neural Network (CNN) model. The comparison was conducted on the same dataset using identical preprocessing techniques and hyperparameter settings to ensure a fair evaluation.

9.1 Model Performance Comparison

Model	Accuracy	F1 score
CNN	85	81
ViT	91	89

The ViT model outperformed the CNN by a significant margin, achieving a 6% higher accuracy and an improved F1-score by 6.5%.

9.2 Observations

- Feature Extraction: CNNs rely on local feature extraction, making them less effective for complex patterns. In contrast, ViTs use self-attention mechanisms, allowing them to capture global dependencies more efficiently.
- Generalization Ability: The CNN model exhibited overfitting, indicating it struggled to generalize well across diverse samples. ViT's transformer-based architecture helped improve generalization, leading to better performance.
- Computational Complexity: The CNN model was computationally lighter and faster during inference. However, the ViT model required more computational resources, making it more demanding in terms of hardware requirements.

Conclusion from Comparison

The results indicate that ViT is superior in terms of accuracy and F1-score for Alzheimer's disease classification. However, for resource-constrained environments, CNNs may still be a viable option due to their lower computational requirements. Future work can explore hybrid architectures combining CNNs and transformers to achieve a balance between efficiency and accuracy.

10. Discussion and Challenges

10.1 Key Findings

- The proposed Vision Transformer (ViT)-based model achieved a 91% accuracy and an 89% F1-score, demonstrating its effectiveness in classifying Alzheimer's disease stages.
- The Flask-based web interface allows real-time inference, making the system accessible for research and clinical use.
- The data augmentation and preprocessing techniques significantly improved model performance by enhancing generalization.

10.2 Challenges and Limitations

Despite the high accuracy, some limitations were observed:

- **Dataset Imbalance**: Some classes had fewer samples than others, which could lead to biased predictions.
- **Computational Complexity**: The Vision Transformer model is computationally intensive, requiring high-end GPUs for training and inference.
- **Real-World Deployment Issues**: The system has not yet been deployed on a cloud platform, limiting its accessibility for broader applications.
- **Model Interpretability**: While ViTs are highly accurate, they lack transparency in decision-making, making it difficult to understand their predictions.

10.3 Potential Solutions

- **Dataset Augmentation**: Techniques like SMOTE (Synthetic Minority Over-sampling Technique) can help balance the dataset.
- **Model Optimization**: Quantization and pruning can be applied to reduce the model's size and improve inference speed.
- **Cloud Deployment**: Hosting the model on platforms like AWS Lambda, Heroku, or Google Cloud can enhance accessibility.
- **Explainable AI (XAI) Techniques**: Implementing methods like Grad-CAM can provide better interpretability for predictions.

11. REFERENCES

[1] Detection of Alzheimer's disease using deep learning models: A systematic literature review.
[2] Generalizable deep learning model for early Alzheimer's disease detection from structural MRIs.
[3] Alzheimer's Disease Detection Using Deep Learning on Neuroimaging: A Systematic Review.
[4] Alzheimer's Disease stage identification using deep learning models.
[5]A deep learning model for detection of Alzheimer's disease based on retinal photographs: a retrospective, multicenter case-control study.
[6] Artificial intelligence for detection of Alzheimer's disease: demonstration of real-world value is required to bridge the translational gap.
[7] Alzheimer's Diseases Detection by Using Deep Learning Algorithms: A Mini-Review.
[8] Early Stage Alzheimer's Disease Prediction Using Machine Learning Models.
[9] An explainable machine learning approach for Alzheimer's disease classification.
[10] Alzheimer's Diseases Detection by Using Deep Learning Algorithms: A Mini-Revie.
[11] Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks
[12] Alzheimer's Disease Detection Using Deep Learning. [13] Alzheimer's Disease Diagnosis by Deep Learning Using MRI-Based Approaches.
[14] Alzheimer's Diseases Detection by Using Deep Learning
Algorithm.
[15]Detection of Alzheimer's disease using Deep Learning Techniques.
[16] An early-stage Alzheimer's disease detection using various imaging modalities and techniques – A 4 mini-review.
[17] DETECTION AND CLASSIFICATION OF ALZHEIMER'S DISEASE USING

DEEP LEARNING.
[18]Alzheimer's Disease Detection using Deep Learning IJARIIE-ISSN(O)-2395-4396.
[19]Alzheimer Disease Detection using AI with Deep Learning based Features with Development and Validation based on Data Science.
[20]Early diagnosis of Alzheimer's disease using machine learning: a multi-diagnostic, generalizable approach
[21]A modified deep learning method for Alzheimer's disease detection based on the facial submicroscopic features in mice. [22]Automated detection of Alzheimer's Disease using Deep Learning in MRI.
[23]Deep learning Based Alzheimer's Disease Detection Model using Multimodal Data.
[24]ALZHEIMER'S DISEASE DETECTION USING DEEP
LEARNING International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 11 Issue: 10 | Oct 2024 www.irjet.net
[25]Detection Of Alzheimer Disease Using Deep Learning.Journal of Advanced Zoology
[26]Alzheimer's Diseases Detection by Using Deep Learning Algorithms: A Mini-Review IEEE ACCESS. [27]Alzheimer's Disease Diagnosis Using Machine Learning: A Survey MDPI.