A Project Report on

Alzheimer's Disease Detection Using Deep Learning

Submitted in partial fulfillment of the requirements for the Degree of

Bachelor of Engineering in Computer Engineering

BY

Walke Omkar Yashwant (B190354284)

Kottur Ashish Anil (B190354242)

Nisad Ravi Raju (B190354263)

Pawar Siddhiraj Sham (B190354268)

Under the guidance of

Dr. S. K. Shinde

Department of Computer Engineering

Vidya Pratishthans

Kamalnayan Bajaj Institute of Engineering and Technology

Baramati-413133, Dist-Pune (M.S.)

India

May 2023

VPKBIET, Baramati

Department of Computer Engineering

Certificate

This is to certify that the project entitled

Alzheimer's Disease Detection Using Deep Learning

submitted by

Walke Omkar Yashwant (B190354284)

Kottur Ashish Anil (B190354242)

Nisad Ravi Raju (B190354263)

Pawar Siddhiraj Sham (B190354268)

is a record of bona-fide work carried out by them, in the partial fulfillment of the requirement for the award of Degree of **Bachelor of Engineering** in Computer Engineering at Vidya Pratishthan's Kamalnayan Bajaj Institute Of Engineering and Technology, Baramati under the Savitribai Phule Pune University, Pune. This work is done during year 2022-23, under our guidance.

Dr. S. K. Shinde	Dr. C. S. Kulkarni	Dr. R. S. Bichkar
Project Guide	Head of Dept.	Principal

Examiner 1: - - - - - - -

Examiner 2: - - - - - - -

Acknowledgements

It gives immense pleasure in submitting this project report on “Alzheimer’s Disease Detection Using Deep Learning ”. I would like to express my deepest gratitude to my guide Dr.S.K.Shinde for his excellent guidance, caring, patience and providing an excellent atmosphere doing project. I am grateful to our college VPKBIET, Baramati for giving me the opportunity to provide necessary resources for this project.

I would like to take this time to thank our esteemed Principal, Dr.R.S. Bichkar, for his unwavering support and encouragement and Dr. C.S.Kulkarni Head of the Department (Computer Engineering), for their continuous encouragement and suggestions during my Project work.

Also, I would like to thank my family and friends for supporting me. Without this support, we would not have been able to do this project. Last but not least, I would like to thank everyone who helped and encouraged me to work on this Project...

Abstract

Alzheimer's disease remains a significant public health issue. Older persons are typically affected by this neurological disorder. Memory loss is the key symptom of Alzheimer's disease, which is followed over time by harder communication skills and other impairments. Other symptoms of Alzheimer's disease includes inability to learn new things, difficulty thinking and understanding, mental confusion, mood swing, depression, difficulty concentrating, forgetfulness. As a result, research on early AD detection has picked up in recent years. The project employed a two-step approach, starting with region of interest extraction to isolate the hippocampus region from brain images, which is closely associated with Alzheimer's disease. Subsequently, a CNN model, specifically VGG-16 architecture, was utilized for image classification. Transfer learning was applied by fine-tuning the pre-trained VGG-16 model on a dataset of brain images related to Alzheimer's disease. The model uses CNN to extract significant features from the brain scans, making it easier to divide them into groups: brains that are healthy(NonDemented) and those that have Alzheimer's disease(MildDemented & ModerateDemented). The system also featured a user interface that allowed users to register, login, and upload their brain images for analysis. The system provided accurate predictions and displayed the results to the users, aiding in the early detection of Alzheimer's disease. Extensive testing and evaluation were conducted to assess the system's performance.The results demonstrated high accuracy and efficiency in classifying brain images as normal or indicative of Alzheimer's disease. The utilization of deep learning and transfer learning techniques significantly enhanced the system's accuracy and reduced the training time.

Contents

Acknowledgements **1**

Abstract **2**

Notation and Abbreviations **7**

1 Introduction **8**

1.1 Motivation 10

1.2 Problem Definition and Objectives 11

1.2.1 Problem Definition 11

1.2.2 Objectives 11

1.3 Project Scope 12

2 Project Plan **13**

2.1 Project Estimate 14

2.2 Risk Management 15

2.3 Project Schedule 16

3 Literature Survey **17**

3.1 Early Detection of Alzheimer's Disease with Blood Plasma Proteins Using Support Vector Machines [1] 17

3.2 Application of KPCA and AdaBoost algorithm in classification of functional MRI of Alzheimer's disease [2] 18

3.3 Alzheimer's Disease Detection Using m-Random Forest Algorithm with Optimum Features Extraction [3] 19
3.4 On the detection of Alzheimer's disease using fuzzy logic-based majority voter classifier [4] 20
3.5 An Enhanced Fuzzy Based KNN Classification Method for Alzheimer's Disease Identification from SMRI Images [5] 21

4 Software Requirements Specification 23
4.1 User Classes and Characteristics 23
4.2 Functional Requirements
24
4.3 Nonfunctional Requirements 24
4.3.1 Safety and Security Requirements 24
4.3.2 Software Quality Attributes 25
4.4 System Requirements 25
4.4.1 Database Requirements 25
4.4.2 Software Requirement 26
4.4.3 Hardware Requirements: 26
4.5 SDLC Model Appplied 26

5 System Design 28
5.1 System Architecture 28
5.1.1 Preprocessing 29
5.1.2 Splitting Dataset 29
5.1.3 Training Dataset 30
5.1.4 Training CNN Classifier 30
5.1.5 Predicting Using CNN Model 31
5.2 Mathematical Model 32
5.3 Data Flow Diagrams 33
5.3.1 Data Flow Diagrams:Level-0 34
5.3.2 Data Flow Diagram:Level-1 35

5.4 Class Diagram . . . 36
5.5 UML Activity Diagram . . . 37
5.6 Sequance Diagram . . . 38
5.7 Use case Diagram . . . 39

6 Project Implementation 41
6.1 Overview of Project Modules . . . 41
6.2 Tools and Technologies Used . . . 44
6.3 Algorithm Details . . . 45
6.3.1 Convolutional neural network(CNN) . . . 45
6.3.2 Transfer Learning . . . 47
6.3.3 VGG-16 . . . 49

7 Software Testing 51
7.1 Type of Testing . . . 51
7.2 Test cases and Test Results . . . 53

8 Results & Discussion 55
8.1 Outcomes . . . 55
8.2 Screen Shots . . . 57

9 Advantages & Limitations 60
9.1 Advantages . . . 60
9.2 Limitations . . . 60

10 Future Scope 61

11 Conclusions 62

References 63

List of Figures

2.1 Timeline Chart 16

5.1 System Architecture 28

5.2 Data flow diagram level 0 34

5.3 Data flow diagram level 1 35

5.4 Class Diagram 36

5.5 Activity Diagram 37

5.6 Sequence Diagram. 38

5.7 Use case Diagram 39

6.1 CNN Architecture 45

6.2 Transfer Learning Architecture 47

6.3 VGG-16 Architecture 49

8.1 Validation Accuracy Graph 56

8.2 Validation Loss Graph 56

8.3 Home Page 57

8.4 User Registration & Login Page. 58

8.6 Upload Image & Predict Page 59

8.7 Result Page. 59

Notation and Abbreviations

AD	Alzheimer's Disease
CNN	Convolutional Neural Network
ROI	Region of Interest
VGG-16	Visual Geometry Group 16
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
ML	Machine Learning
DL	Deep Learning
CSV	Comma-Separated Values
API	Application Programming Interface
HTML	Hypertext Markup Language
CSS	Cascading Style Sheets
UI	User Interface
UX	User Experience
CPU	Central Processing Unit
GPU	Graphics Processing Unit
RAM	Random Access Memory
ReLU	Rectified Linear Unit
FC	Fully Connected

Chapter 1

Introduction

Alzheimer's disease is a neurodegenerative condition that affects the brain and causes a steady decline in memory and cognitive function. As the condition worsens, people may experience changes in behaviour and personality that make performing regular tasks challenging. Around the world, 50 million people suffer from dementia, with Alzheimer's disease accounting for two thirds of cases. With a new case of Alzheimer's occurring every three seconds worldwide, the illness has surpassed cancer cases to become the most dreaded in the United States. For Alzheimer's disease to be effectively treated and managed, early detection is essential. Current diagnostic techniques, however, have several limitations and can be pricy and invasive. Using non-invasive methods like brain imaging, recent developments in machine learning and artificial intelligence approaches have enabled the development of algorithms that can aid in the early detection of Alzheimer's disease. Machine learning and completely automatic segmentation techniques have shown impressive results in a variety of computer vision and image processing jobs during the past ten years. Early attempts to diagnose Alzheimer's disease using MRIs depended on pre-selected discriminative features. These characteristics include segmented cortical thickness and regional volumes from areas of the brain associated with or implicated in Alzheimer's disease-related memory loss and accelerated neurodegeneration. Deep convolutional neural networks (CNNs)-based modern machine learning techniques enable data-driven feature extraction from picture data.

Alzheimer's disease is a brain ailment that gradually impairs thinking and memory abilities as well as the capacity to do even the most basic tasks. The majority of Alzheimer's patients have their initial symptoms later in life.Experts estimate that more than 6 million Americans, the majority of whom are 65 or older, may have dementia brought on by Alzheimer's disease. These estimates differ.The suggested model consists of two key steps. The first stage is area of interest extraction, which relies on dividing the image into distinct blocks in order to extract only the portion of the brain that contains the hippocampus. The classification of images using the Convolutional Neural Network (CNN) and Transfer Learning deep learning algorithms constitutes the second phase. In the majority of image processing and computer vision tasks, these methods have been demonstrated to perform better than conventional strategies based on predetermined features.CNN-based techniques have shown promise in identifying new imaging biomarkers in the biomedical field. Deep learning methods have been utilized in several studies to diagnose mild Alzheimer's disease dementia using MRI data. Some notable examples include the use of 3D convolutional neural networks such as 3D AlexNet and 3D ResNet, patch-based models, Siamese networks, and auto-encoder-based models. However, previous approaches have been criticized in systematic reviews and surveys for their conceptual or validation flaws. Many of these studies primarily focus on distinguishing Alzheimer's disease dementia patients from healthy individuals.

Convolutional neural networks (CNNs) in particular are frequently employed in image analysis applications, such as Alzheimer's disease diagnosis. These networks can be trained to recognise visual cues that are important for categorising the images, such as anatomical alterations in the brain that are indicative of Alzheimer's disease. A CNN is often trained using a sizable dataset of MRI scans from both healthy people and people with Alzheimer's disease in deep learning models for Alzheimer's disease detection. The model is then put to the test on a different dataset to see how well it classifies brand-new photos. Another method for enhancing the efficiency of deep learning models for Alzheimer's disease diagnosis is transfer learning. This entails tweaking a pre-trained CNN on the Alzheimer's disease classification problem from a starting point, such as the VGG16 model. This strategy can decrease the amount of training data needed while increasing the model's accuracy.

In general, the suggested model is a promising method for automatically detecting Alzheimer's disease from MRI scans. This model has the potential to enhance early disease identification and therapy by concentrating on the hippocampal region and making use of CNNs and Transfer Learning. By including additional characteristics like age and gender and doing tests on larger datasets from various institutions, the model can be further enhanced.

1.1 Motivation

Alzheimer's disease is the extremely popular cause of dementia that causes memory loss. Alzheimer's disease is a neurodegenerative ailment that affects numerous brain functions and affects those who have it. Alzheimer's disease (AD) can be detected using a variety of conventional methods, including SVM, Random Forest, Fuzzy Logic, Adaboost, and others. Deep learning-based methods have recently been taken into consideration for the classification of neuroimaging data connected to AD. Because it is more effective and accurate than other methods, we investigate the use of convolutional neural networks (CNN) in this research for the identification of AD. The goal of the proposed research is to create an automated system for detecting Alzheimer's disease utilising deep learning techniques, such as CNNs, that is accurate and effective at detecting the disease in its early stages. The goal of the research is to properly identify and analyse changes in the area of interest (ROI) that is frequently impacted in Alzheimer's disease, such as the hippocampus, using cutting-edge image processing and deep learning methods. The proposed project has the potential to greatly improve the diagnosis and treatment of Alzheimer's disease as well as the quality of life for millions of individuals worldwide by creating an accurate and effective approach for Alzheimer's disease detection.

1.2 Problem Definition and Objectives

1.2.1 Problem Definition

To develop Convolutional Neural Network (CNN) based deep learning model for detection of Alzheimer's disease.

This study aims to address the need for a reliable and efficient method of diagnosing Alzheimer's disease (AD) using magnetic resonance imaging (MRI) scans. AD is a common and fatal neurodegenerative disorder that causes memory loss and other neurological symptoms. The management of the condition depends heavily on early diagnosis and therapy, but existing diagnostic techniques are frequently time-consuming, expensive, and lack the sensitivity and specificity needed for early detection. The suggested method tries to automatically identify and categorise AD from MRI scans using deep learning techniques, specifically Convolutional Neural Networks (CNNs). Creating a CNN model that can precisely recognise important areas of interest, such the hippocampus, and discover the underlying patterns that distinguish between MRI scans with and without AD is a difficult task.

1.2.2 Objectives

1. To create a CNN-based model with excellent sensitivity and specificity for AD detection and classification from MRI scans using the VGG16 architecture.
2. To enhance the performance, accuracy, and resilience of the VGG16-based CNN model architecture and hyperparameters.
3. To investigate using transfer learning with pre-trained VGG16 models to detect AD in order to get around the lack of large, diverse datasets.
4. To find out how the performance of the VGG16-based CNN model is affected by various preprocessing methods, such as image normalisation, registration, and segmentation.

5. To assess the effectiveness of the proposed VGG16-based CNN technique using common performance indicators including accuracy, sensitivity, and specificity.
6. to assess how well the suggested VGG16-based CNN methodology for detecting AD from MRI images performs in comparison to other established machine learning techniques like SVM, Random Forest, and Adaboost.
7. To investigate the possibility of the suggested VGG16-based CNN technique for AD early detection, which can assist patients with AD receive better care and live better lives.
8. To look at the VGG16-based CNN model's interpretability and find the pertinent characteristics and areas of interest that help with AD detection.
9. To provide a user-friendly interface that will enable medical practitioners to use the VGG16-based CNN AD detection system.
10. To explore the effect of the VGG16 architecture's layer count on the performance of the CNN model for AD detection and to look at several CNN architectures that might increase accuracy while lowering computational complexity.

1.3 Project Scope

This project's scope is to create and assess a CNN-based method for accurately detecting and classifying Alzheimer's disease (AD) from MRI scans utilising the VGG16 architecture and transfer learning. The pre-processing of MRI scans to extract pertinent regions of interest, such as the hippocampus, will be part of the research. After that, a CNN model built using the VGG16 architecture will be developed and trained. The performance of the CNN model will be enhanced by leveraging previously trained models through transfer learning. The study will also involve measuring the performance of the suggested approach using accepted measures including accuracy, sensitivity, and specificity. Investigating the CNN model's interpretability and creating a user-friendly interface for the AD detection system are also included in the project's scope.

Chapter 2

Project Plan

The project plan outlines the key activities, resources, and timelines required for the successful completion of the Alzheimer's disease detection project. It provides a structured approach to ensure efficient project execution and delivery of desired outcomes. The following elements are included in the project plan:

1. **Project Scope:** The project scope establishes the parameters and goals of the project. It lists the precise attributes, capabilities, and outputs that will be created as a component of the Alzheimer's disease detection system.

2. **Project Objectives:** The project's goals and desired outcomes are described in the project objectives. These goals can include producing a reliable system, a user-friendly interface, an accurate model for detecting Alzheimer's disease, and a specific level of accuracy.

3. **Project Resources:** The human, technological, and financial resources needed for the project are all included in the project resources. The project team members, hardware and software tools, datasets, money, and any other resources required to complete the project's operations are included in this.

4. **Project Phases and Tasks:** In order to assist an organised and methodical execution, the project is separated into phases and tasks. The precise activities that must be finished inside each phase are referred to as tasks. Each phase signifies

a significant milestone. This division assists in tracking development, managing dependencies, and guaranteeing the timely delivery of project deliverables.

5. **Project Dependencies:** Project dependencies are the connections between several activities or phases that influence the project's execution and timetable. Managing interdependencies, planning tasks, and resolving any possible bottlenecks that may develop during the project are all made easier by identifying dependencies.

6. **Project Timeline:** The start and end dates for each phase and job are specified in the project timeline. It offers a visual depiction of the project timeline, complete with deadlines, key pathways, and milestones. The timeline aids in tracking development, spotting delays or deviations, and making the required corrections to keep the project on schedule.

7. **Project Budget:** The expected project costs are listed in the project budget. It covers costs for materials, tools, software licences, data collecting, and any other pertinent costs. Monitoring the project budget guarantees effective resource allocation and adherence to budgetary restrictions.

2.1 Project Estimate

The Alzheimer's disease detection project's resources, timeframe, and cost must all be estimated as part of the project estimate. It consists of a number of parts, including data collection, preprocessing, model training, evaluation, and visualisation of the results.

1. **Resource Estimate:** This entails calculating the amount of technology, software, and manpower that will be needed for the project. It takes into account elements like processing power, storage capacity, software licences, and team member knowledge. The best resources required to complete the project successfully are determined after thorough investigation.

2. **Time Estimate:** Time estimation is dividing the project's tasks into smaller chunks and calculating how long it will take to complete each one. Data gathering, preprocessing, model training, evaluation, result visualisation, and user interface

creation are some of the considerations that are made. The estimation is based on prior experience, professional opinion, and the project's complexity.

3. **Cost Estimate:** The main goal of cost estimation is to determine how much money will be needed for the project. It covers costs for staff, data collecting, hardware, software, and any other project-specific needs. The cost estimate guarantees that the project stays within the specified financial resources and aids in budgeting.

2.2 Risk Management

In order to successfully complete the Alzheimer's disease detection project, risk management requires recognising potential risks and creating plans to reduce or manage those risks. The objective is to reduce the influence of risks on project results. he risk management process includes the following steps:

1. **Risk Identification:** Determine and record any possible risks that might occur during the project. These dangers may be caused by technical problems, poor data, a lack of resources, or outside influences like shifting laws or funding.

2. **Risk Analysis:** Assess each indicated risk's likelihood and effect. Put the dangers in order of importance depending on their gravity and likelihood of happening. Understanding which dangers demand immediate attention and resources is made easier by this analysis.

3. **Risk Mitigation:** Create plans and methods to reduce or eliminate the identified risks. This may entail setting up backup plans, alternate data sources, extra funding, additional resources, or quality control procedures. The goal is to lessen the chance of or effect that risks will have on the project.

4. **Risk Monitoring:** Throughout the course of the project, keep an ongoing eye on and track of the identified risks. Assess the success of the applied risk reduction methods on a regular basis. Be prepared to address new hazards and adapt the solutions as necessary.

2.3 Project Schedule

The project schedule outlines the timeline and key milestones of the Alzheimer's disease detection project. The Gantt chart provides a visual representation of the project's tasks, dependencies, and durations. It helps in planning and tracking the progress of the project.

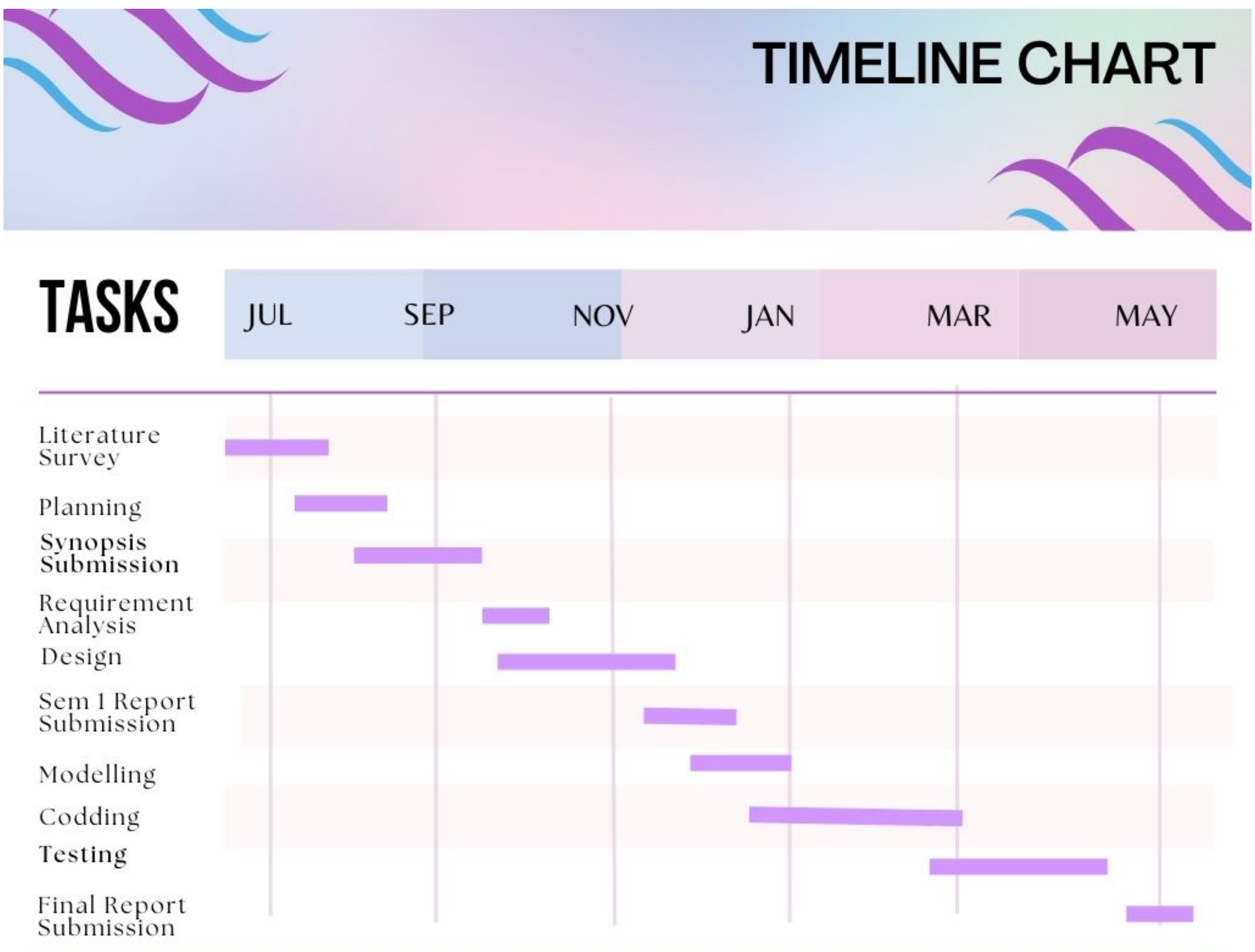

Fig.2.1 Timeline Chart

Chapter 3

Literature Survey

3.1 Early Detection of Alzheimer's Disease with Blood Plasma Proteins Using Support Vector Machines [1]

The aim of the research topic "Early Detection of Alzheimer's Disease with Blood Plasma Proteins Using Support Vector Machines" is to create a machine learning strategy that can aid in the identification of blood plasma proteins that may be used as potential biomarkers for the early diagnosis of Alzheimer's disease (AD). Beta-amyloid plaques and tau tangles build up in the brain as a result of the progressive neurodegenerative illness AD, which impairs cognition and causes memory loss. There is currently no known cure for AD, and there are few available treatments. Early disease detection is therefore essential for successful management and therapy. 82 people total were enrolled in the trial, comprising 41 AD sufferers and 41 healthy controls. All participants had blood plasma samples taken, which were then subjected to mass spectrometric analysis in order to recognise and count the proteins present. Based on the samples' protein patterns, the researchers classified the samples into AD and control groups using a support vector machine (SVM) technique. A popular supervised machine learning technique for classification tasks is SVM. A portion of the data was used to train the SVM algorithm, while the remaining samples were utilised to evaluate the model's effectiveness. The findings demonstrated that the SVM method had an accuracy of 93.5% in classifying the data into the AD and control groups. A group of 10 proteins were also found by the researchers to be significantly different between the AD and control groups. These proteins were linked to a number of biological functions, such as immune response, metabolism, and inflammation. The results of the study imply that blood plasma proteins may be useful biomarkers for the early identification of AD. The method may enhance the precision and effectiveness of AD diagnosis and open the door to the creation of efficient treatment plans for the condition. The study does, however, have certain shortcomings. The sample size was somewhat limited, and more extensive validation of the findings is required. Additionally, the study failed to take into account variables that can have an impact on protein profiles, such as age, sex, and drug use. Overall, the work emphasises the potential of employing blood plasma proteins and machine learning-based methods as biomarkers for early identification of AD. The results need to be confirmed by additional study in order to maximise the sensitivity and specificity of the technique.

3.2 Application of KPCA and AdaBoost algorithm in classification of functional MRI of Alzheimer's disease [2]

The paper "Application of KPCA and AdaBoost algorithm in classification of functional MRI of Alzheimer's disease" aims to develop a machine learning approach that can accurately classify functional magnetic resonance imaging (fMRI) data to distinguish between Alzheimer's disease (AD) patients and healthy controls. A non-invasive imaging method called functional magnetic resonance imaging (fMRI) can be used to analyse how the brain functions by detecting changes in blood flow and oxygenation. AD is a neurodegenerative disease that affects brain function and cognitive abilities, making fMRI an important tool for understanding the disease. This research developed a feature classification model for Alzheimer's disease based on AdaBoost algorithm and KPCA algorithm, and selected 21 patients with Alzheimer's disease in order to obtain successful classification of magnetic resonance images of Alzheimer's disease (AD). The experiment included 7 individuals with early Alzheimer's disease, 6 individuals with advanced Alzheimer's disease (LAD),and 8 healthy people (HC) who underwent various degrees of examination. The findings demonstrate that the KPCA algorithm is used in the article to get the maximum classification accuracy possible between the two groups: The node degree is the only characteristic that can be distinguished with a precision in the imaging diagnosis of AD. The classification of Alzheimer's disease-related magnetic resonance imaging can be greatly improved by the article.

By applying Adaboost, weak classifiers' accuracy can be increased. Adaboost is now used to categorise text and graphics instead of binary classification issues. Adaboost's key drawback is that a high-quality dataset is required. Before implementing an Adaboost algorithm, noisy data and outliers must be avoided. This outcome has significant clinical implications for the diagnosis and categorization of AD using magnetic resonance imaging and serves as a good test of the performance of the chosen algorithm.

3.3 Alzheimer's Disease Detection Using m-Random Forest Algorithm with Optimum Features Extraction [3]

The paper "Alzheimer's Disease Detection Using m-Random Forest Algorithm with Optimum Features Extraction" presents a machine learning approach to detect Alzheimer's disease (AD) using magnetic resonance imaging (MRI) data. The study aimed to develop a more accurate and efficient approach to AD detection, which is essential for early diagnosis and intervention. The proposed approach uses a combination of feature extraction techniques and machine learning algorithms.

The study used a dataset consisting of MRI pictures of 84 participants, comprising 44 patients with AD and 40 healthy controls. The data were preprocessed, and a set of features were extracted using wavelet transform and principal component analysis (PCA) techniques. The study then used an m-random forest algorithm, which is an extension of the traditional random forest algorithm, to classify the MRI data into AD and control groups based on their features. The m-random forest algorithm was optimized by selecting the optimal number of trees and features.

In this paper, an ensemble learning-based m-Random Forest classifier is used for Alzheimer's disease detection. Moreover, the proposed algorithm can be able to detect

Alzheimer's disease at an early stage. This paper propose an upgraded machine learning algorithm named Modified Random Forest (m-RF) to individualize between normal people and people with the risk of having Alzheimer's disease. Compared to other methods like Support Vector Machine, Adaptive Boosting, K-Nearest Neighbors, etc., accuracy is significantly higher. This model's weakness is its lack of accuracy. However, accuracy can be improved by extracting more features from a huge dataset. Because CNN are particularly good at making predictions from images, using a CNN-based model is therefore preferable to using a random forest model, which is normally used for tabular data with a mixture of numerical and categorical variables.

3.4 On the detection of Alzheimer's disease using fuzzy logic-based majority voter classifier [4]

The paper "On the detection of Alzheimer's disease using fuzzy logic-based majority voter classifier" proposes a machine learning approach to detect Alzheimer's disease (AD) using magnetic resonance imaging (MRI) data and fuzzy logic-based majority voter classifier. The study aims to establish a more reliable and efficient strategy to AD detection, which is critical for early diagnosis and treatments. The proposed approach uses a combination of feature extraction techniques and machine learning algorithms. The study used a dataset consisting of MRI images of 60 subjects, including 30 patients with AD and 30 healthy controls. The data were preprocessed, and a set of features were extracted using wavelet transform and gray-level co-occurrence matrix (GLCM) techniques.

Early identification of AD has attracted a lot of attention from researchers throughout the world in an effort to increase life expectancy. Proposed research suggests a fuzzy logic-based majority voter classifier approach for the classification of AD, MCI and HC subject from the volumetric information of MRI images. Main objective of our proposition is too fuzzy the training data (MRI images) on the basis of percentage of volumetric information of WM, GM and CSF of human brain with the aid of three different membership functions. During the testing phase, appropriate fuzzy class (either AD or MCI or HC) has been selected on the basis of max-membership principle and a majority voter classifier, that too with a superior classification accuracy. The study's conclusions point to the potential for the suggested method to be a highly effective and accurate tool for AD detection. The approach could also assist select the most significant traits for AD detection and improve our understanding of the disease's underlying causes.

Using the volumetric data of the white matter (WM), grey matter (GM), and cerebro spinal fluid, this research offers a novel attempt to categorise the brain MRI pictures into three classes, namely Alzheimer's disease (AD), mild cognitive impairment (MCI), and healthy controls (HC) (CSF). This classification was completed using a majority voter classifier and a fuzzy logic-based technique. Finally, multiple brain MRI images compiled from the ADNI dataset are used to test our hypothesis However, the proposition suffers from the indecision of majority voter classifier in case three intermediate classifiers decide in favor of three different classes and consequently no absolute majority is obtained against any of the three classes.

3.5 An Enhanced Fuzzy Based KNN Classification Method for Alzheimer's Disease Identification from SMRI Images [5]

The paper "An Enhanced Fuzzy Based KNN Classification Method for Alzheimer's Disease Identification from SMRI Images" proposes a machine learning approach to identify Alzheimer's disease (AD) using structural magnetic resonance imaging (SMRI) data and a fuzzy-based k-nearest neighbor (KNN) classifier. The study aimed to develop a more accurate and efficient approach to AD identification, which is essential for early diagnosis and intervention. The proposed approach uses a combination of feature extraction techniques, fuzzy logic, and KNN classification.

The study used a dataset consisting of SMRI images of 96 subjects, including 48 patients with AD and 48 healthy controls. The data were preprocessed, and a set of features were extracted using wavelet transform and gray-level co-occurrence matrix (GLCM) techniques. The study then used a fuzzy-based KNN classifier to classify the SMRI data into AD and control groups based on their features. The classifier combines the outputs of multiple classifiers to improve the classification accuracy.

In this research work, the fuzzy model combined with the enhanced KNN method for the detection of AD, CN, and MCI based on the hippocampus region from the MRI images. The study's findings suggest that the proposed approach could be a promising tool for AD identification, with high accuracy and efficiency. The approach could also help identify the most relevant features for AD identification and improve our understanding of the disease's underlying mechanisms. However, the study has some limitations, including the relatively small sample size and the lack of validation on an independent dataset. Further research is needed to validate the approach's performance on larger datasets and optimize the algorithm's sensitivity and specificity.

Summary: The literature survey of the five selected papers provides valuable insights into different approaches for Alzheimer's disease detection and diagnosis. These studies demonstrate how people with Alzheimer's disease may be successfully identified and categorised using machine learning methods such support vector machines, AdaBoost, m-random forest, fuzzy logic-based classifiers, and enhanced K-nearest neighbours. These algorithms, combined with different data preprocessing strategies, feature extraction approaches, and classification models, have demonstrated promising outcomes in terms of accuracy and disease early detection. According to these results, our method makes use of convolutional neural networks and transfer learning, utilising the strength of deep learning and pre-trained models to extract significant characteristics from brain images and achieve high accuracy in Alzheimer's disease classification. Our approach intends to increase the effectiveness and accuracy of Alzheimer's disease detection, supporting early intervention, by integrating CNN and transfer learning.

Chapter 4

Software Requirements Specification

4.1 User Classes and Characteristics

Classes

- User
- System
- User Interface
- Model
- Alzheimer's Dataset

Attributes-

- User Attributes- Username , Email , Password
- Dataset Attributes – Non-Demented, Mild-Demented, Moderate-Demented
- User Interface Attributes- Login and registration

Operations-

- User- Register(), Login(), Tutorial(), Logout(), Home(), Predict()
- User Interface- getalzhemer'sdata(), displayprediction()
- Model- trainmodel(), testmodel(), predict()
- Dataset- pre-processing()

4.2 Functional Requirements

- The user should register on Alzheimer's disease detection system.
- The user should login after registration.
- The system should validate user.
- The web application should let customer upload MRI images.
- The web application should collect responses from the user.
- The dataset should be provided by Kaggle to the system
- The model should be trained on given dataset using CNN Algorithm.
- The model should be tested by the system
- The model should predict result

4.3 Nonfunctional Requirements

- The user should register with appropriate details.
- The application should be connected with Internet.

4.3.1 Safety and Security Requirements

Safety requirement establishes the safety functional requirements for the Safety Instrumented Systems. CNN deep learning models are built using a fundamental set of safety requirements. These include:

- Distinct layers that are not physically related to one another.
- Data preprocessing, activation functions, and gradient descent parameters all lack bias in the model.
- The model has not used a library or framework to construct its architecture (e.g., random forests).

4.3.2 Software Quality Attributes

- Adaptability: This system works on minimum computing power.
- Availability: It is nothing but the probability that the system is operating properly when it is requested to use.
- Reliability: It is the probability of failure-free software operation for a specified period of time.
- Correctness: It determines that how users can interact with the software and how the software should behave when it is used correctly.

4.4 System Requirements

4.4.1 Database Requirements

In the Alzheimer's disease detection project using CNN and VGG16, a database is an essential component.The initiative calls for a sizable database collecting MRI scans of the brains of both people with Alzheimer's disease and healthy people. The patient's age, sex, and other pertinent medical information must be noted in the applicable metadata on the high-quality MRI scans. The database must also be trustworthy and secure to guarantee that patient information is safeguarded and only accessible by authorised employees.Large amounts of data should be efficiently stored, retrieved, and processed by the database. Both organised and unstructured data formats should be supported. The project team should be able to access and understand the data thanks to proper organisation. Additionally, as the project advances, the database should be scalable and capable of handling a growing volume of data. Additionally, it must be able to smoothly connect with the project's software tools and technology. The database should allow SQL queries and offer tools for validating, transforming, and manipulating data.The Alzheimer's disease detection project employing CNN and VGG16 requires a high-quality, dependable, secure, and scalable database. The database should support the project's software tools and technologies, be capable of handling huge amounts of organised and unstructured data efficiently, and be simple to use and comprehend.

4.4.2 Software Requirement

- Tool: Anaconda Navigator

  Anaconda Navigator is a graphical user interface that allows users to easily manage and access the various components of the Anaconda distribution.It contains a variety of programmes and tools that are frequently used in data science and machine learning applications, including RStudio, Spyder, and Jupyter Notebook.

- Language: Python

  Python is a programming language known for its emphasis on readability and simplicity, making it easy to learn and understand. Python code is typically simpler for developers to read and understand than code written in other languages.

4.4.3 Hardware Requirements:

- RAM: Greater than 4GB
- PROCESSOR: Core 2 and above
- DISK: Greater than 10GB
- Operating system:Windows 7 or higher

4.5 Analysis Models: SDLC Model to be applied

Using a software engineering approach basically means creating a software development methodology that organises, plans, and controls the system development process. It contains definitions for all project team items that are created and completed. For this modelling, there are a few options:

1 Waterfall a linear framework.

2 Prototyping an iterative framework.

3 Incremental: a combined linear an iterative framework.

4 Spiral a combined linear iterative framework.

5 Rapid application development (RAD) an iterative framework.

6 Extreme Programming.

The Waterfall model is a conventional technical method to software engineering in which the entire software development process is broken down into steps. The following are the phases of the waterfall model:

* Requirement analysis: For designing software projects, perform a requirement analysis that includes both software and hardware perspectives. Requirements engineering stresses the use of systematic and repeatable approaches to guarantee that system requirements are full, consistent, and relevant. Virtual environments such as platform and running environment are included on the software side. The physical environment, such as hard dises and RAM, is included in the hardware side The goal of this project is to come up with a cost-effective and efficient solution. The impact of changing parameters on the outcome is included in the requirements

* Software Design: This phase reviews and prepares the system based on the requirements specified in the preceding phase. System design aids in the development of overall system architecture as well as the specification of hardware and system requirements. The development of

  – Algorithm Design

  – Planning and Scheduling

  – System Architecture

  – Unified Modelling Design

Chapter 5

System Design

5.1 System Architecture

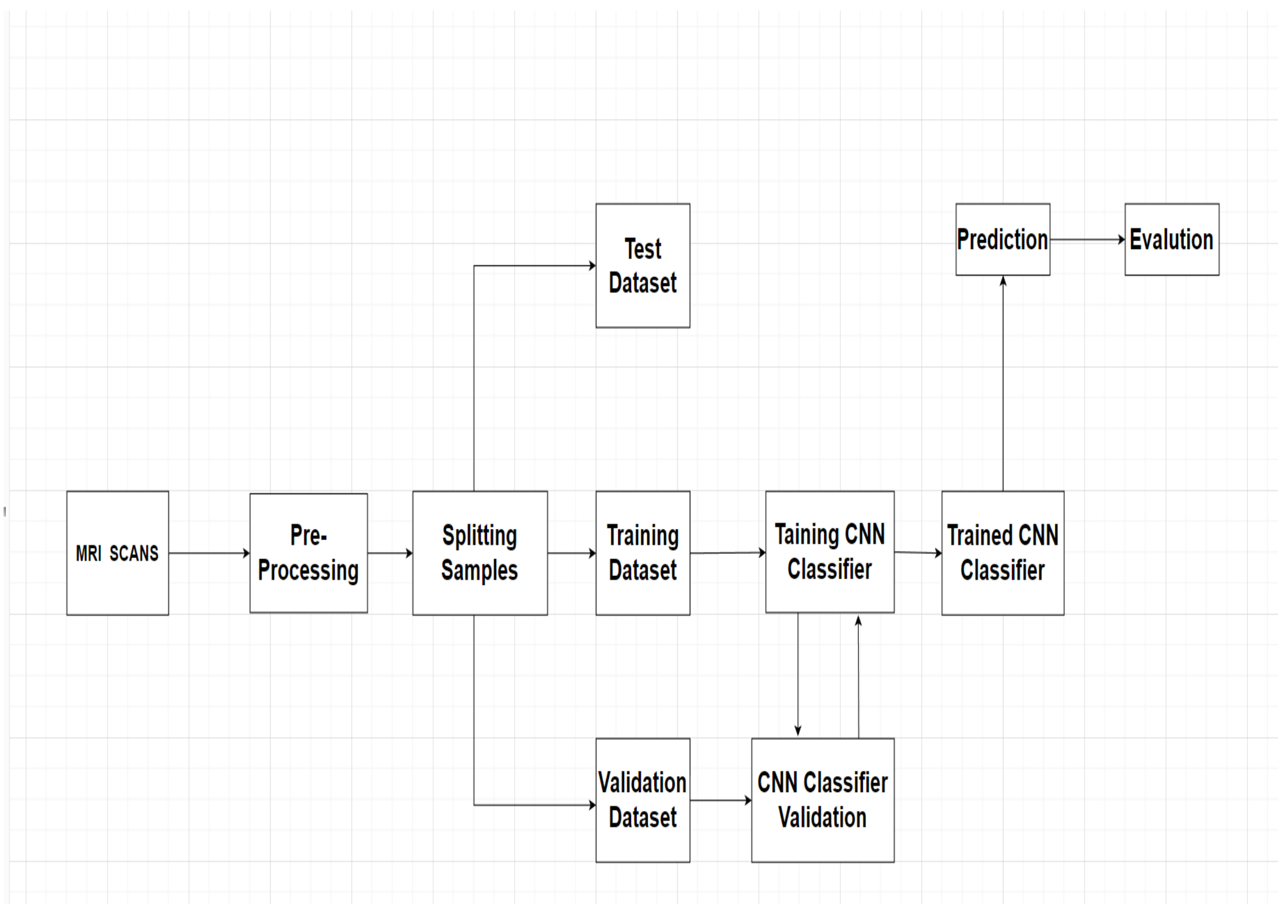

Fig.5.1 System Architecture

5.1.1 Preprocessing

In this section, we will describe preprocessing of MRI images in the data processing pipeline. The resizing and normalisation processes make up the preprocessing. We need to combine many photos from a single recording session into a single image, thus resizing is required. Each image's size is determined by the length of the recording session, which can vary depending on the patient. Therefore, before computing statistics or carrying out other operations on any photographs, we must resize them altogether. Many engineering (such as signal processing) and biological (such as microscopy) areas routinely resize objects. Depending on the intended size of the output image, there are various ways for resizing images, such as linear scaling or non-linear scaling. The most basic technique is linear scaling, which involves multiplying each pixel value by its matching factor to create new pixel values. There is no information loss as a result of resizing because the pixel values in this instance are linearly connected to one another.

5.1.2 Splitting Dataset

In Alzheimer's disease detection using VGG16 and transfer learning, splitting the dataset refers to dividing the available data into training, validation, and test sets. This step is essential for creating and assessing a CNN model since it enables the model to be trained and tested on various data subsets. The training set, the validation set, and the test set are typically the three components of the dataset. The validation set is used to fine-tune the model's hyperparameters and assess the model's performance during training, whereas the training set is used to train the CNN model. The test set is then used to assess how well the CNN model performs on new data and to predict how well it generalises. To make sure that the subsets reflect the distribution of the entire dataset, the dataset must be carefully divided. This aids in avoiding overfitting, a situation in which the CNN model performs admirably on training data but poorly on untrained data. To enable comparison and replication by other people, the splitting should also be random and repeatable.

5.1.3 Training Dataset

The training dataset typically consists of a large number of labeled images, where each image is associated with a label indicating whether it belongs to a normal group or an Alzheimer's group. In the objective of this project, the training dataset would most likely include brain MRI scans from both healthy and Alzheimer's disease-infected patients. Following that, the CNN model is trained by minimising a loss function that gauges the discrepancy between the CNN model's anticipated output and the actual label of the training data. In order to reduce the loss function, the weights of the CNN model are then modified iteratively using an optimisation approach like stochastic gradient descent (SGD). The CNN model gains knowledge of the patterns and characteristics that are most important for differentiating between normal and Alzheimer's brain scans during training. Starting with random weights, the model learns to modify these weights as it is trained to reduce the loss function. An essential first step in creating a reliable Alzheimer's detection system is training the CNN model. Metrics like accuracy, precision, recall, and F1-score can be used to assess the model's performance on the training dataset. The model can be used to generate predictions on fresh, unforeseen data once it has been trained and proven to work satisfactorily.

5.1.4 Training CNN Classifier

Training the CNN classifier refers to the process of training the classifier on top of the pre-trained VGG16 model to classify the brain MRI scans as either normal or Alzheimer's disease. First, the pre-trained VGG16 model is loaded, and the learned weights are frozen to stop future updates. A new fully connected layer is used in place of the pre-trained VGG16 model's final fully connected layer, which was originally intended to categorise photos from the ImageNet dataset into 1,000 classes. Two outputs are available from this newly connected layer, one for regular brain MRI scans and the other for scans for Alzheimer's disease. The new completely connected layer's weights are initially randomised. The complete model is then trained using the backpropagation algorithm on the training dataset after the CNN classifier's architecture has been established. The VGG16 convolutional layers are run on the input MRI images throughout this phase, and their feature maps are then retrieved. The new completely linked layer receives these feature maps and categorises the input scans as normal or Alzheimer's disease-related. During training, the weights of the next fully connected layer are changed to minimise the loss function, which calculates the discrepancy between the model's anticipated output and the actual label of the training data. By minimising the categorical cross-entropy loss function, the CNN classifier is trained. Adam optimizer, a stochastic gradient descent optimizer that employs an adaptive learning rate, is the optimizer utilised in this project. To make sure that the model converges to the ideal solution, the learning rate of the optimizer must be tweaked during the training phase. After the CNN classifier has been trained, a validation dataset is used to assess its performance. A portion of the training dataset called the validation dataset is not used for training. Metrics like accuracy, precision, recall, and F1-score are used to assess how well the CNN classifier performs. The trained model can then be used to generate predictions on fresh, upcoming MRI scans if the performance is adequate.

5.1.5 Predicting Using CNN Model

To make predictions using the trained CNN model, we first preprocess the input MRI scans in the same way as we did for the training and validation datasets. This involves resizing the photos to a defined size, pixel value normalisation, and grayscale conversion. After preprocessing the input scans, we run them through the trained CNN model. A probability score for each class, such as normal and Alzheimer's disease, is the model's output. The class for the input scan is projected to be the one with the highest likelihood score. It is significant to emphasise that care should be used when interpreting the CNN model's predictions. Although the model may correctly categorise MRI images as either normal or Alzheimer's disease, it does not offer a conclusive diagnosis of the condition. Only a medical practitioner can make an accurate diagnosis of Alzheimer's disease after conducting a thorough evaluation that includes a clinical history, physical exam, and diagnostic tests. In general, predicting with the CNN model entails picking the class with the highest probability as the predicted class for the input scan by running preprocessed input MRI scans through the trained model to get probability scores for each class.

5.2 Mathematical Model

Let X be the input image matrix of size $(W \times H \times C)$, where W represents the width, H represents the height, and C represents the number of channels (typically 3 for RGB images).

Convolutional Layer: The convolutional layer applies a set of filters to the input image to extract features. Let F be the number of filters, and each filter is represented by a weight matrix W^l of size $(K \times K \times C)$, where K is the size of the filter kernel. The convolution operation is performed using a stride of S, and the output feature maps are obtained as follows:

$$Z^l = \text{Convolve}(X, W^l, S)$$
$$A^l = \text{ReLU}(Z^l)$$

Max Pooling Layer: The max pooling layer downsamples the feature maps to reduce spatial dimensions. Let P be the size of the pooling filter. The max pooling operation is performed as follows:

$$M^l = \text{MaxPool}(A^l, P)$$

Fully Connected Layer: The fully connected layer connects all neurons from the previous layer to the current layer. Let N^l be the number of neurons in the fully connected layer. The output of the fully connected layer is computed as:

$$O^l = \text{Activation}(W^l \cdot M^l + b^l)$$

Here, W^l represents the weight matrix of size $(N^l \times M)$, b^l represents the bias vector, and Activation represents the activation function (e.g., ReLU or softmax).

Training Objective: The training objective of the model is typically to minimize a loss function. For binary classification, commonly used loss functions include binary cross-entropy or sigmoid cross-entropy. For multi-class classification, categorical cross-entropy is often employed.

Optimization Algorithm: To optimize the model parameters, an optimization algorithm such as stochastic gradient descent (SGD) or Adam is used. The algorithm adjusts the weights and biases of the model during the training process based on the gradients of the loss function with respect to the parameters.

Transfer Learning: Transfer learning is applied by initializing the CNN model with pre-trained weights obtained from training on a large dataset, such as ImageNet. The pre-trained model is then fine-tuned on the Alzheimer's disease dataset specific to this project.

5.3 Data Flow Diagrams

A data flow diagram (DFD) shows how data moves through a system or process graphically. It demonstrates the inputs, outputs, and changes that occur as data passes through various elements and processes. A data flow diagram can offer a visual depiction of how data is processed and analysed in relation to the Alzheimer's disease detection project. The project's various stages from data collection and preprocessing to model training, evaluation, and prediction are shown visually in the data flow diagram. It facilitates a better comprehension and analysis of the project's data processing pipeline by illuminating the overall data flow and interactions between various components.

5.3.1 Data Flow Diagrams:Level-0

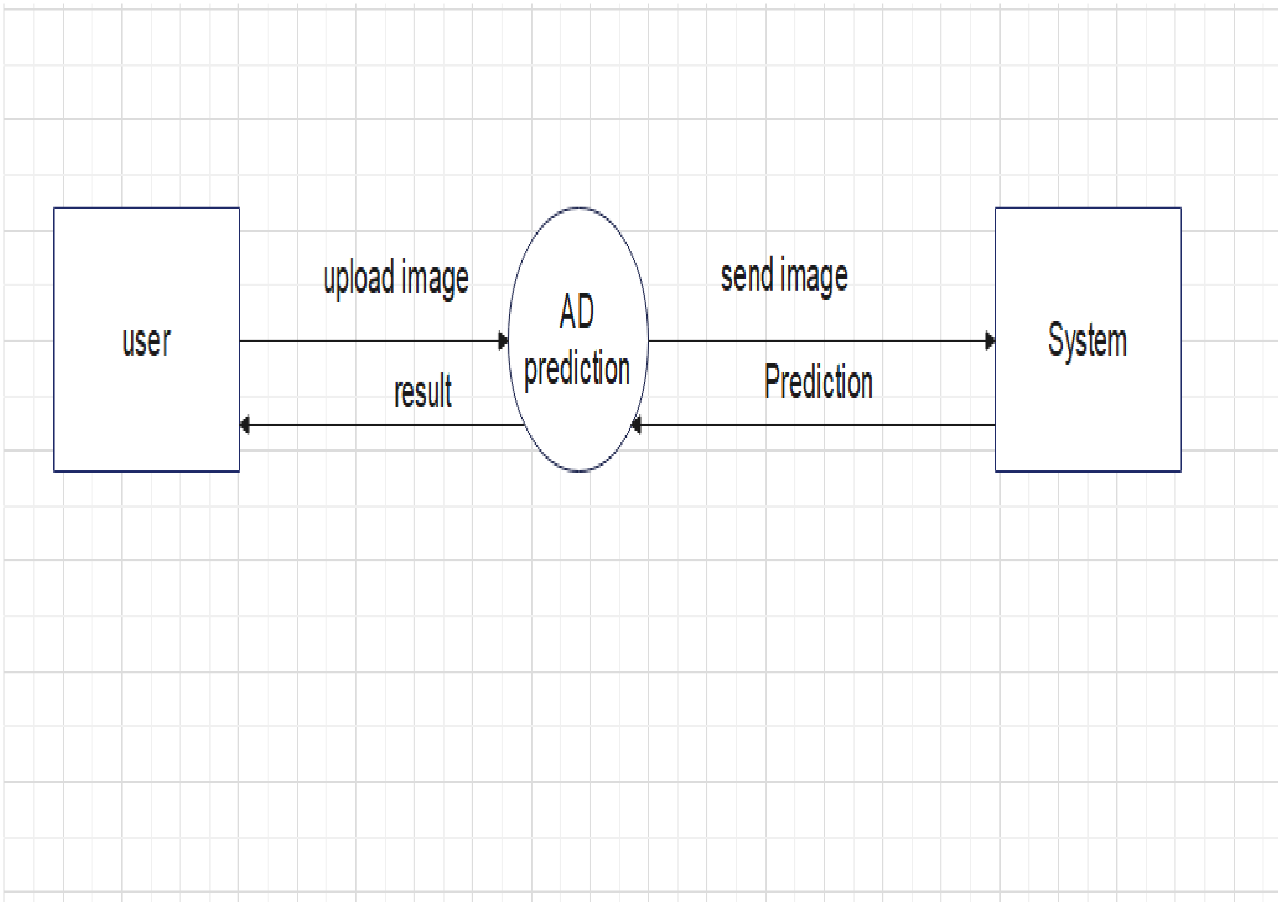

Fig.5.2 Data flow diagram level 0

The greatest degree of abstraction for an overview of the system is a degree-0 Data Flow Diagram (DFD).An overview of the system's components and interactions is given at a high level in a Level-0 DFD, with the system as a whole as the main focus. It includes all relevant external parties as well as the main operations and the data flow between them. The core workings of processes and the precise data items involved are not covered in detail by the Level-0 DFD.Understanding a system's general data structure and flow is helped by using Level-0 DFDs. At lower levels, where processes from the Level-0 DFD are broken down into more specialised subprocesses, they serve as a starting point for developing more intricate DFDs. This makes it possible to comprehend the system's operation and data flow better.

5.3.2 Data Flow Diagram:Level-1

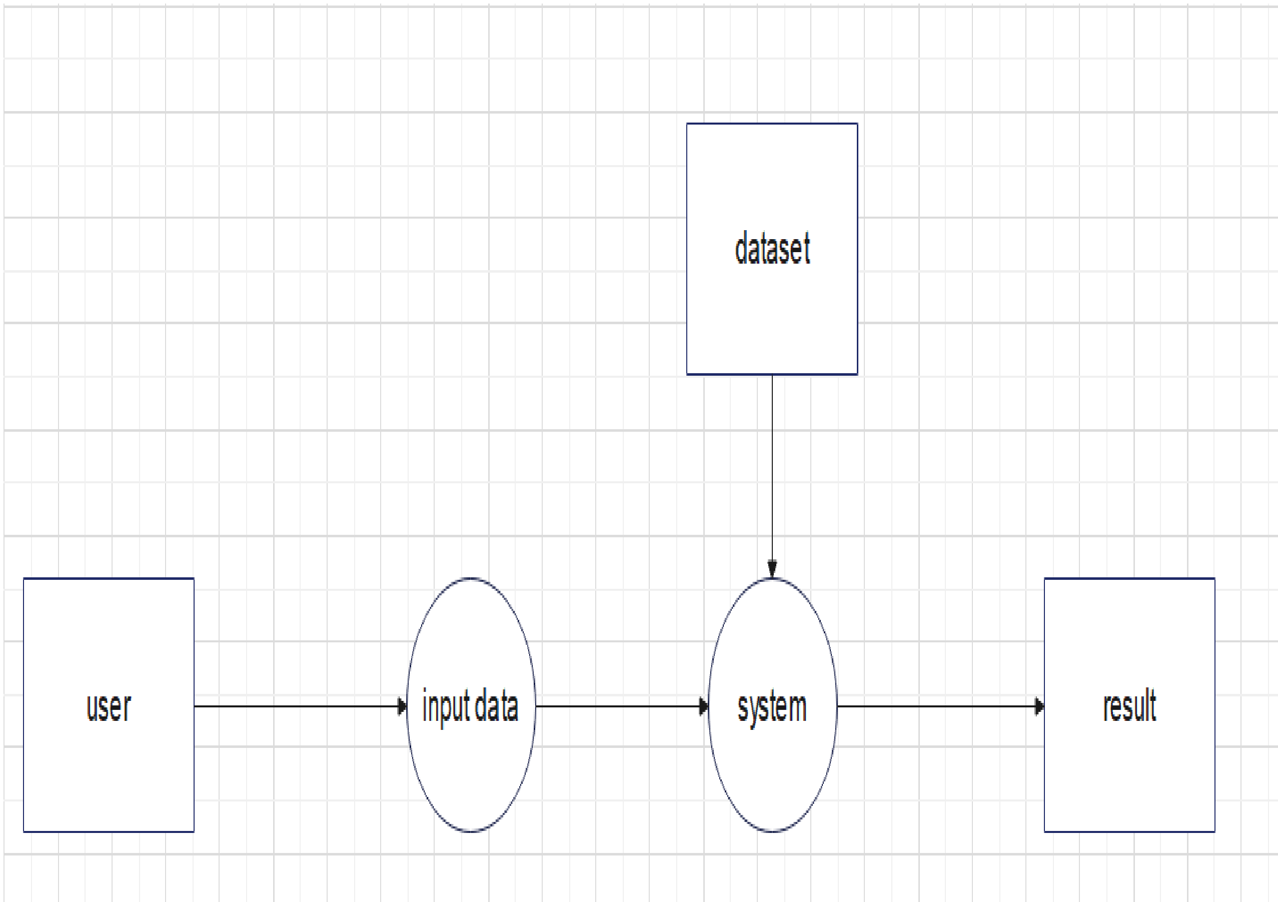

Fig.5.3 Data flow diagram level 1

A Level-1 Data Flow Diagram (DFD) depicts a system in greater detail than a Level-0 DFD.In a Level-1 DFD, the emphasis switches from a broad perspective of the system to a more in-depth analysis of subprocesses and their interactions. It offers a greater comprehension of how data moves throughout the system and how different subprocesses alter or process the data. Each subprocess may have a unique set of data flows for input and output, which helps the system as a whole function.

5.4 Class Diagram

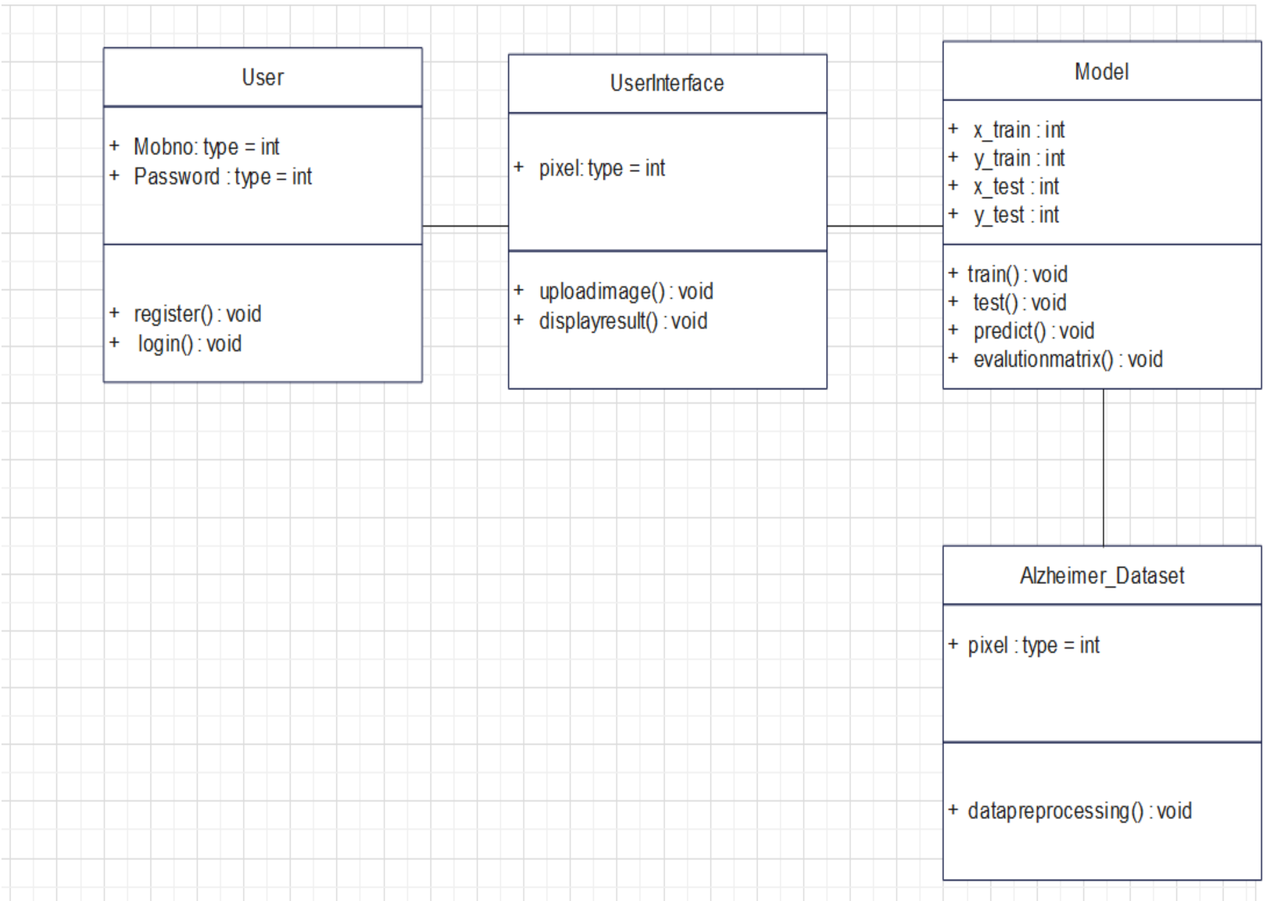

Fig.5.4 Class Diagram

The project consists of four main classes: User, UserInterface, Model, and Alzheimer's Dataset. The User class handles user registration, login, and profile management. The UserInterface class represents the web application's interface, allowing users to interact with the system through features like uploading images and obtaining predictions. The Model class contains the deep learning model used for Alzheimer's disease forecasting, including methods for preprocessing input data and making predictions. The Alzheimer's Dataset class provides access to the dataset used for training the model, including methods for retrieving and processing the data.

5.5 UML Activity Diagram

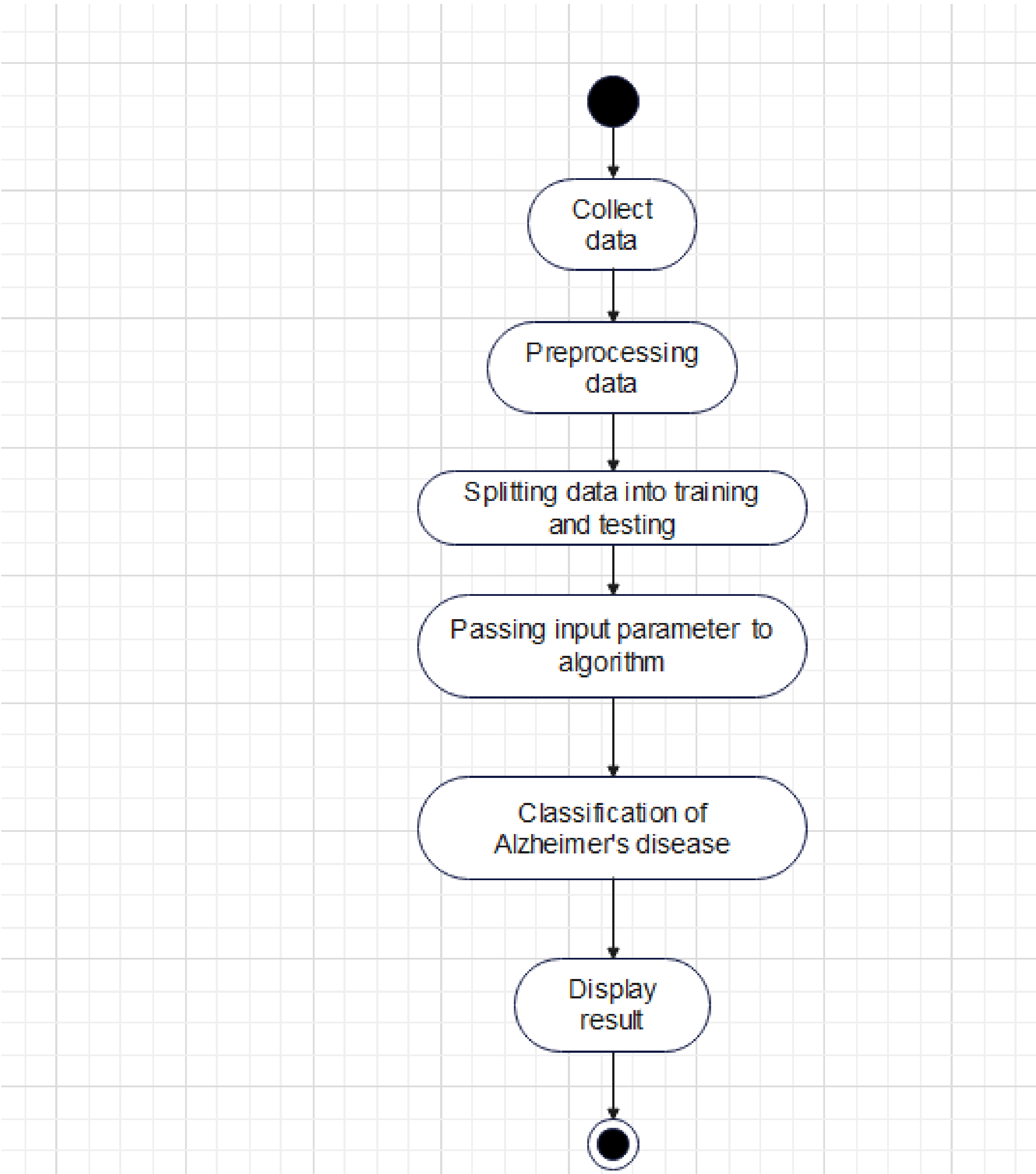

Fig.5.5 Activity Diagram

A visual representation of the flow of activities or actions within a system or process is an activity diagram. Using deep learning techniques, the activity diagram shows the successive stages involved in diagnosing Alzheimer's disease. It emphasises the crucial steps in the detection process, such as preprocessing, data splitting, classification, and result presentation, making it easier to comprehend how things work.

5.6 Sequance Diagram

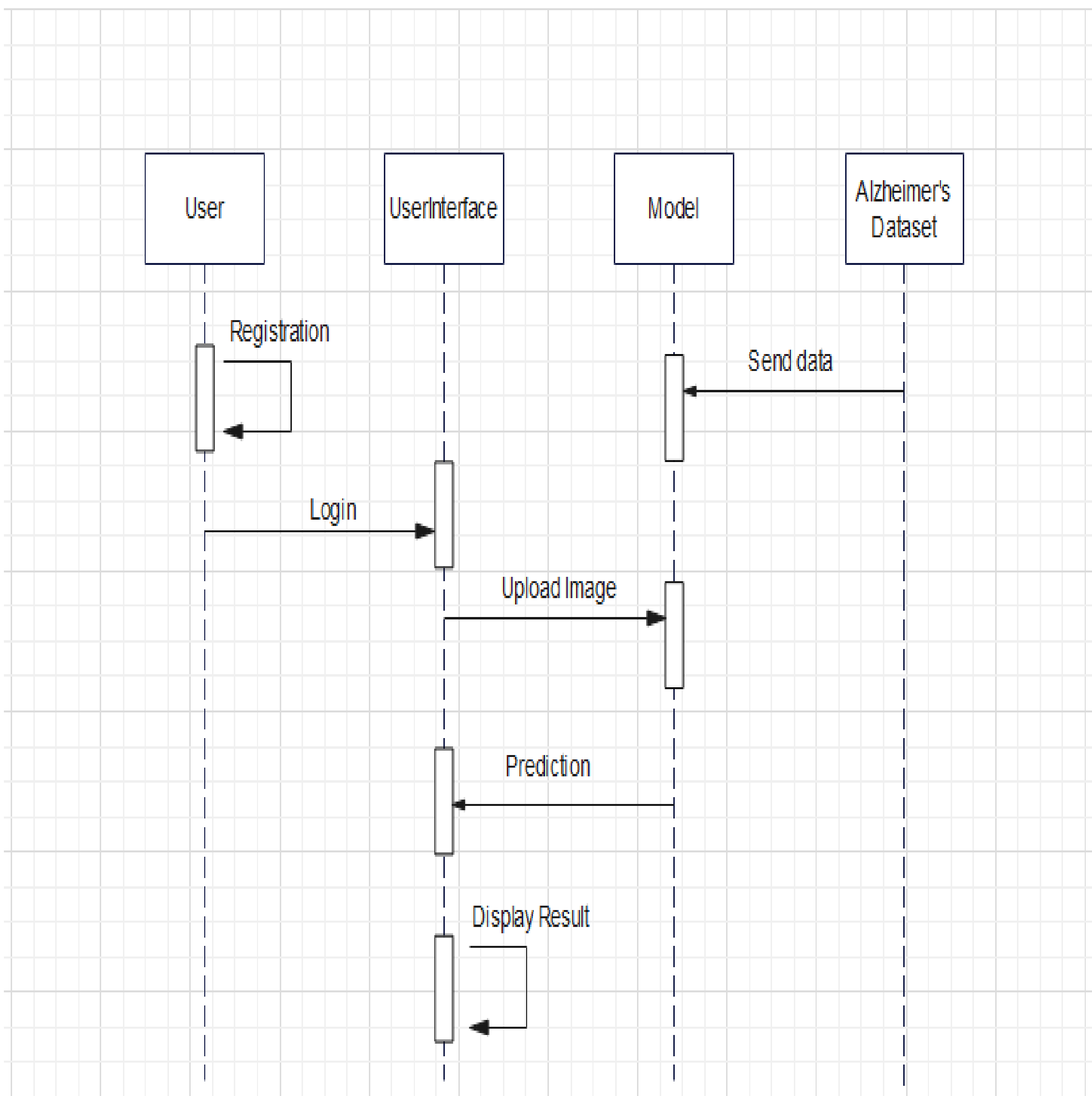

Fig.5.6 Sequence Diagram

A sequence diagram depicts the interactions between several modules or components of a system in a time-based manner.The interaction between the user, system components, and the deep learning-based Alzheimer's detection process is depicted in the sequence diagram. It shows how several processes work, such as user authentication, image uploads, preprocessing, classification, and result presentation.

5.7 Use case Diagram

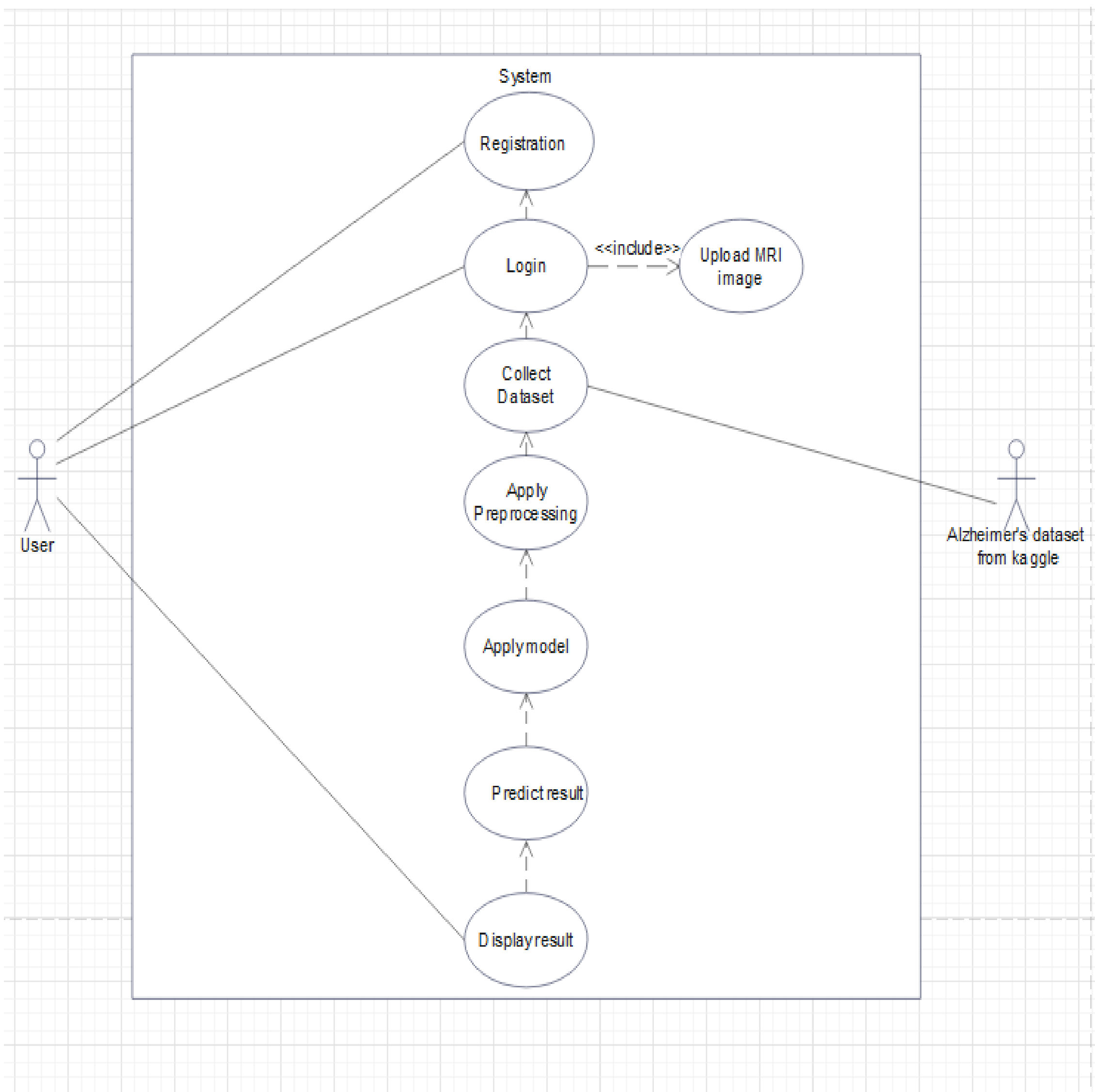

Fig.5.7 Use case Diagram

An Alzheimer's disease use case diagram represents the interactions between the system and its actors, illustrating the various use cases or functionalities provided by the system.

- Actors:
  1 User: Encapsulates the system's users who communicate with the Alzheimer's disease detection system. They carry out tasks including signing up, logging in, uploading photographs, and getting the findings.
  2 Alzheimer's Dataset:Represents an external entity or source that provides the dataset of labeled images for training the model.
- Use Cases:
  1 Registration: In this use scenario, the user registers with the system by giving the required information, such as a username, email address, and password. To utilise the system's functionality, the user may set up an account.
  2 Login: To utilise the system's features, the user must log in using their login information (username and password).
  3 Collect Dataset: In order to gather the labelled dataset for deep learning model training, the system interacts with the Alzheimer's Dataset. In this use case, the dataset is obtained from the external source and integrated into the system.
  4 Apply Preprocessing: In this use case, the system preprocesses the user-provided photos or the dataset that has been gathered. Resizing, normalising, and converting the photos to an appropriate format for additional analysis may all be part of the preprocessing procedure.
  5 Apply Model: For the purpose of detecting Alzheimer's disease, the system uses the preprocessed images and a deep learning model, such as a CNN with VGG16 architecture. In this use case, important features are extracted for classification and fed into the trained model together with the preprocessed photos.
  6 Predict Result: The system predicts the result or diagnosis for each input image after applying the model. According to the model's predictions, this use case entails categorising the photos as being indicative of Alzheimer's disease or not.

Chapter 6

Project Implementation

6.1 Overview of Project Modules

The project consists of several modules that collectively contribute to the development and implementation of an Alzheimer's disease detection system using deep learning. Here is an overview of the project modules:

1. **Data Collection and Preprocessing:** The Alzheimer's disease detection project's Data Collection and Preprocessing module entails gathering and getting ready the dataset for deep learning model training and evaluation. Gathering particular brain images relating to Alzheimer's illness is the initial stage. These photos can be found in a variety of places, including medical databases, research centres, and hospitals with a focus on neuroimaging. It is important to take care to make sure the dataset includes a variety of instances, including both healthy people and those with Alzheimer's disease.

2. **Model Training:** An essential component of creating a deep learning-based system for Alzheimer's disease detection is model training. This lesson focuses on teaching convolutional neural networks (CNNs), a type of deep learning model, to recognise patterns and features in brain images that are suggestive of Alzheimer's disease. There are several crucial elements in the training process. The architecture of the deep learning model is established. In this situation, the base model might be a CNN architecture like VGG-16. To take advantage of the characteristics that were learned

from a big dataset, like ImageNet, the pre-trained weights of the CNN can be loaded. The process of initialising the model with useful and generalizable representations of visual information is referred to as transfer learning. The training procedure starts after the model architecture has been established. There are training and validation sets for the dataset. The training procedure is repeated across a number of epochs, with each epoch denoting a thorough scan of the full training dataset. The model continuously improves in accuracy and performance as it learns to recognise pertinent patterns and traits unique to Alzheimer's disease.

3. **Model Evaluation and Validation:** In this module, we evaluate the trained models to determine how well they function and how accurate they are at diagnosing Alzheimer's disease. To assess the model's efficacy, evaluation measures like accuracy, precision, recall, and F1 score are computed. To make sure the results are robust, validation techniques like cross-validation or train-test split may be utilised.

4. **Prediction and Classification:** The Alzheimer's disease detection system's Prediction and Classification module is a key element. The deep learning models are prepared to make predictions on fresh, unused photos once they have been trained on the preprocessed brain images. The module uses the input of the preprocessed photos and the trained models to categorise them as either being indicative of Alzheimer's disease or not. The trained models, which have learned to recognise patterns and features associated with Alzheimer's disease, are fed the preprocessed images during the prediction phase. Based on learnt representations, the models analyse the input images and produce predictions. The prediction process's output tells us how likely it is that the input image contains indications of Alzheimer's disease. The goal of the classification step is to divide the input photos into groups that commonly represent cases of Alzheimer's disease or other conditions. The learned features and patterns are used by the models to produce precise classifications depending on the properties of the input photos. The techniques used by the models, such softmax, assign probabilities to each class, allowing one to identify which class is most likely for a particular image.

5. **User Interface and Interaction:** The Alzheimer's disease detection project's User Interface and Interaction module focuses on developing a user-friendly interface that enables users to efficiently engage with the system. To ensure a flawless user experience, it has a variety of parts and features.

   1 Home Page: The web application's home page acts as its front page. It gives a general summary of the system's features. A navigation menu or buttons that link to various application parts, such as login, registration, instructional, and prediction, are frequently included on the home page.

   2 Login: By entering their login information, including their email and password, individuals can access their personal accounts. Users can access the prediction page and other customised features after successfully logging in.

   3 Registration: New users can register by completing a registration form on the registration page. Typically, users are asked for personal information such their name, email address, and password.

   4 Tutorial: Guide or instructions on how to utilise the website or the Alzheimer's disease detection system are provided in the tutorial section. Users can contribute photographs for prediction, evaluate the results, and navigate between pages with its assistance.

   5 Prediction Page: The key element of the user interface where people input brain photos for Alzheimer's disease detection is the prediction page. Users can choose and upload their brain photos using a file upload tool that is provided. The trained models then evaluate the submitted photos to determine whether Alzheimer's disease is likely to occur. Users can make sense of the outcomes by looking at them on the prediction page.

   6 Logout: Users can safely log out of their accounts and end sessions using the logout feature. After a successful logout, it often redirects back to the home page and is usually reachable from any page inside the application.

6.2 Tools and Technologies Used

1. **Anaconda Navigator**

   The main development environment for the study on Alzheimer's disease detection utilising CNN and VGG16 was Anaconda Navigator. It enabled for the establishment of a virtual environment that was especially suited to the needs of the project and offered a user-friendly interface for maintaining and upgrading packages.The Python code for data preprocessing, model training, and evaluation was written and executed using Jupyter Notebook, one of the tools available in Anaconda Navigator. Additionally, Anaconda Navigator made it easier to install deep learning packages like Keras and TensorFlow, which were necessary for putting the CNN and VGG16 models into practise.

2. **Python Flask**

   Flask is a popular Python web framework that allows developers to easily and quickly create web apps. It is classified as a micro web framework, which means that it is basic and lightweight, making it simple to learn and use. As a result, it is an excellent solution for developing small web applications or prototypes. The first step in creating a website for Alzheimer's disease detection would be to put up a Flask application. It is necessary to create a Python file that defines the application and its routes.

3. **Xamp**

   XAMPP is an open-source cross-platform web server solution stack package, consisting mainly of the Apache HTTP Server, MySQL database, and interpreters for scripts written in the PHP and Perl programming languages. With the aid of XAMPP, a web interface for user login and registration can be built for an Alzheimer's disease detection system. PHP can be used for server-side processing while HTML, CSS can be used to create the web interface. A MySQL database can be used to construct the login and registration system because it is part of the XAMPP package.

6.3 Algorithm Details

6.3.1 Convolutional neural network(CNN)

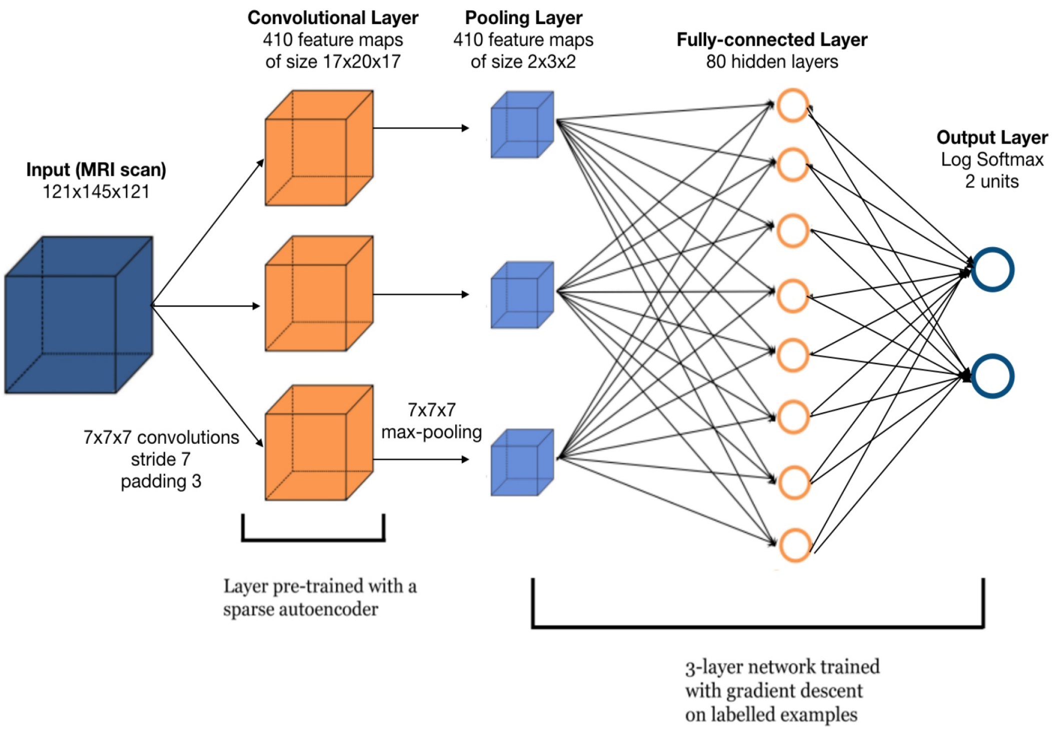

Fig.6.1 CNN Architcture

A deep learning architecture called CNN, or Convolutional Neural Network, is particularly made for processing and analysing visual input, such as photographs. By obtaining cutting-edge performance in a variety of image-related tasks like image classification, object recognition, and picture segmentation, it has completely changed the field of computer vision. The architecture of a CNN is made up of a number of interconnected layers that work together to identify significant elements in the input data and generate predictions.

1. **Convolutional Layer:** This is the foundational element of a CNN. It performs a convolution operation on the input image by applying a series of learnable filters (sometimes referred to as kernels). This layer's function is to record the input image's regional patterns and details, including edges, textures, and forms.

2. **Activation Function:** To add non-linearity to the network, an activation function is applied element-wise after each convolutional operation. The Rectified Linear Unit (ReLU), which sets negative values to zero and leaves positive values unaltered, is the most widely used activation function in CNNs.

3. **Pooling Layer:** The most crucial details are preserved while the spatial dimensions (width and height) of the feature maps are reduced by this layer. This is accomplished by using techniques like maximum pooling or average pooling to downsample the feature maps. Pooling aids in lowering computational complexity and increases the network's resistance to minute changes in input.

4. **Fully Connected Layer:** The features are flattened and run through fully connected layers following a number of convolutional and pooling layers. The final classification or regression is carried out by the fully connected layers using the learnt representations after they have mastered complicated combinations of features.

5. **Output Layer:** The CNN's last layer generates the required results. For instance, in image classification, a softmax activation function is frequently used to assign class probabilities to the input image. The class that has the greatest likelihood is referred to as the projected class.

6.3.2 Transfer Learning

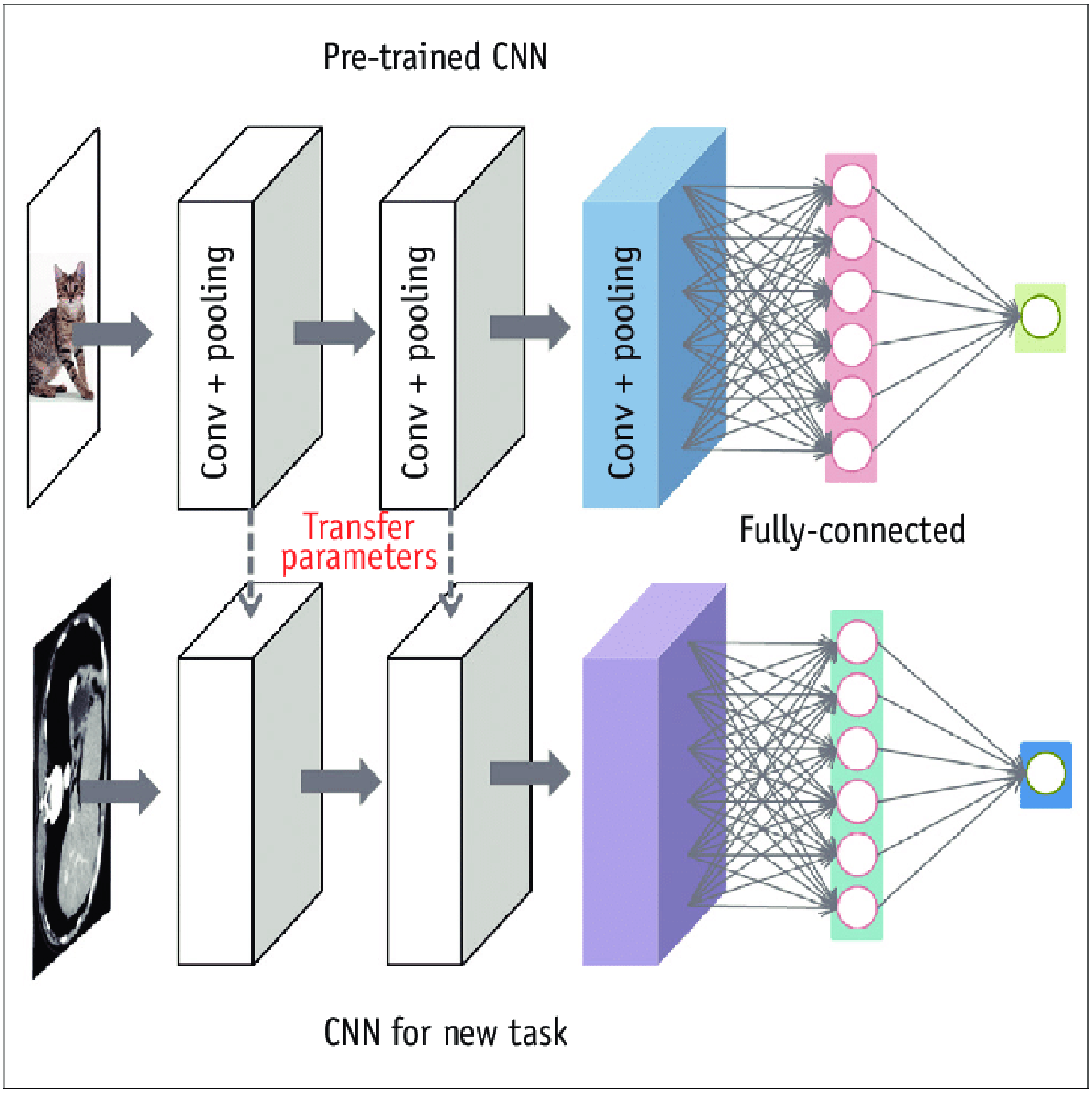

Fig.6.2 Transfer Learning Architecture

A deep learning technique called transfer learning uses a pre-trained model—typically a Convolutional Neural Network (CNN)—as a starting point for tackling a new, related job. Transfer learning uses the information and learnt characteristics of the pre-trained model to speed up training and enhance performance on the new task rather than training a CNN from start on a big dataset.

- **Pre-trained CNN:** A CNN model is pre-trained on a big dataset in the first stage. This dataset is often general-purpose, like ImageNet, and comprises a broad variety of images from many categories. In order to pre-train the CNN, the images from the dataset are fed into it, and the weights are then optimised using methods like backpropagation and gradient descent. Consequently, the CNN gains the ability to extract significant characteristics from the input photos and generalise well on a variety of visual patterns and objects. Convolutional layers for feature extraction and fully linked layers for classification are among the layers that make up the pre-trained CNN. While the deeper layers of the network tend to catch more high-level and abstract data, the initial layers of the network often collect low-level elements like edges and textures.

- **CNN for New Task:**The pre-trained CNN is employed as a feature extractor for a new assignment in the second step. The pre-trained weights of the CNN are frozen, and only the final few levels (completely connected layers) are updated or altered to meet the needs of the new job rather than training the entire CNN from scratch on the new dataset. The final layer of the pre-trained CNN outputs high-level features, which are fed into the new or changed layers. Regression or classification that is task-specific is carried out by these levels. The labelled data from the new task is used to train the new layers, which have a random initialization process. The pre-trained layers that have been frozen serve as a feature extractor, giving the new layers useful representations of the input data. The pre-trained layers' weights do not change during training; instead, the weights of the new layers are updated via backpropagation and gradient descent. As a result, the pre-trained layers' information can be used by the new layers to learn task-specific features.

6.3.3 VGG-16

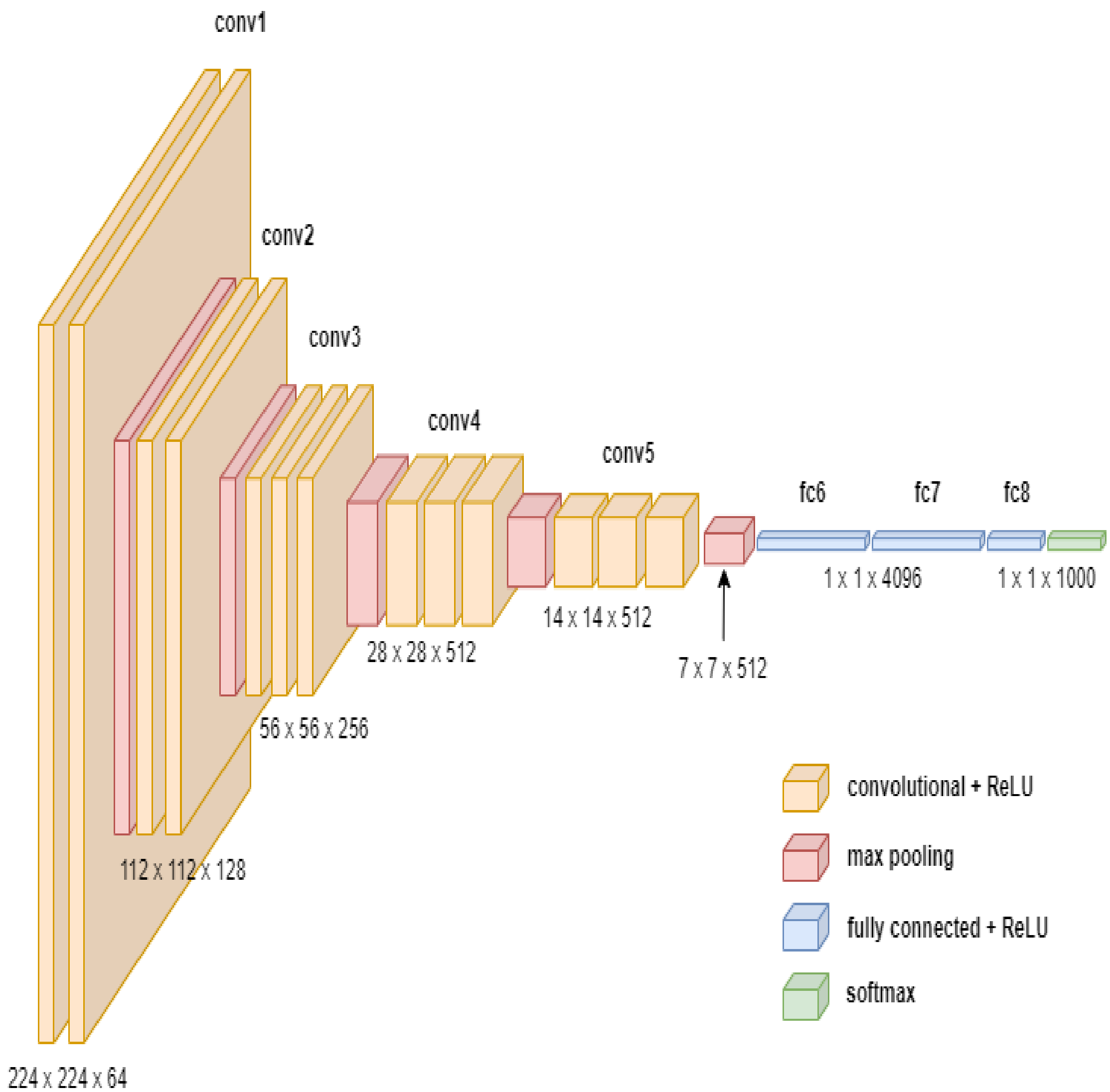

Fig.6.3 VGG-16 Architecture

A well-known convolutional neural network model is the VGG-16 architecture, which was created by the Visual Geometry Group (VGG) at the University of Oxford. It is frequently employed in many computer vision tasks, such as segmentation, object identification, and image classification. The VGG-16 model contains 16 layers total, including 3 fully connected layers and 13 convolutional layers.

The VGG-16 architecture consists of stacked convolutional layers, followed by fully connected layers. Here is a detailed explanation of each component:

1. **Input Layer:** The VGG-16 network receives an RGB image with a 224x224 pixel input.

2. **Convolutional Layers:** There are 13 convolutional layers in the VGG-16 architecture. Each convolutional layer employs a tiny (3x3 filter) receptive field with a 1-pixel stride. From 64 to 512 filters are used, progressively more complex patterns and features in the input image are captured.

3. **Max Pooling Layers:** A max pooling layer is introduced after every two convolutional layers. The max pooling operation shrinks the feature maps' spatial dimensions while retaining their most important data. The filter used by the pooling layers is a 7x7 filter with a 7-pixel stride.

4. **Fully Connected Layers:** The stack of convolutional and pooling layers are followed by three fully connected layers in the VGG-16 architecture. Based on the features that have been learned, these layers carry out the final categorization. Each fully connected layer starts out with 4096 neurons, and the final layer has 1000 neurons, which corresponds to the number of classes in the ImageNet dataset.

5. **Softmax Classifier:** Class probabilities are typically generated using the VGG-16 architecture in the final layer using softmax activation. It is appropriate for multi-class classification tasks since Softmax makes sure that the projected probability add up to 1.

The model can recognise pertinent patterns and features in brain images and help classify Alzheimer's disease by utilising the learned properties of the VGG-16 architecture. A powerful feature extractor, the pre-trained VGG-16 model enables the network to learn discriminative representations and attain excellent accuracy in the task of detecting Alzheimer's disease.

Chapter 7

Software Testing

7.1 Type of Testing

In our Alzheimer's disease detection project, we performed several types of testing to ensure the quality and reliability of the software. The following are the types of testing conducted:

1 **Unit Testing:** To test individual software modules or components in isolation, unit testing was done. Each component, including the user interface, CNN classifier was examined to ensure that it operated as intended and to look for any faults or errors.

2 **Integration Testing:** Integration testing focused on testing the integration and interaction between different modules of the system. We verified that the modules worked seamlessly together and that data and information were accurately exchanged between them.

3 **Functional Testing:** To determine whether the system complied with the stated functional requirements, functional testing was done. It required evaluating the system's capabilities, including user login, image upload, prediction, and result display. Functional testing made sure the system accurately carried out its intended functions and generated the desired results.

4 **System Testing:** The purpose of system testing was to examine the behaviour and functionality of the system as a whole in a comprehensive environment. End-to-end testing was done to make sure that all the modules and parts functioned properly as a whole and that the system offered the capabilities and functions that were anticipated.

5 **Performance Testing:** To evaluate the system's performance under various workloads and conditions, performance testing was carried out. We evaluated each functionality's reaction time, searched for performance bottlenecks, and verified that the system could support a manageable number of concurrent users without noticeably suffering performance loss.

6 **Usability Testing:**Usability testing was primarily concerned with assessing the system's user interface and interaction features. In order to evaluate the system's usability, intuitiveness, and user satisfaction. It assisted in detecting any usability flaws or potential areas for enhancement in the user experience, navigation, and interface.

7 **Security Testing:** Security testing aimed to identify vulnerabilities and ensure the system's security. We conducted tests to validate that user data was handled securely, sensitive information was encrypted, and appropriate access controls were in place to protect user privacy.

8 **Regression Testing:** Regression testing was performed whenever changes were made to the system or new features were added. It ensured that existing functionalities were not affected by the updates and that the system continued to work as expected.

7.2 Test cases and Test Results

Test Case ID	Description	Expected Result	Pass/Fail
Backend Test Cases			
BT-001	Verify database connectivity	Successful connection to the database	Pass
BT-002	Test data preprocessing function	Input data is correctly preprocessed	Pass
BT-003	Test model training function	Models are trained successfully	Pass
BT-004	Test prediction function	Correct prediction is generated for input data	Pass
BT-005	Test database update function	Database is updated with prediction results	Pass
BT-006	Test API response time	API responds within an acceptable time limit	Pass
BT-007	Test error handling	Appropriate error messages are displayed for invalid input	Pass
BT-008	Test data encryption	Sensitive user data is securely encrypted	Pass
BT-009	Test data backup	Data is backed up and can be restored successfully	Pass
BT-010	Test scalability	System handles a large number of requests without performance degradation	Pass

Table 7.1 Backend test cases & test results

Test Case ID	Description	Expected Result	Pass/Fail
Frontend Test Cases			
FT-001	Test user registration form validation	Invalid form inputs display appropriate error messages	Pass
FT-002	Test login functionality with incorrect credentials	Error message is displayed for invalid login attempt	Pass
FT-003	Test image upload feature with large file size	Successfully handle images with large files	Pass
FT-004	Test prediction display for Alzheimer's disease	System accurately predicts and displays Alzheimer's disease diagnosis	Pass
FT-005	Test prediction display for normal brain	System accurately predicts and displays normal brain diagnosis	Pass
FT-006	Display no image uploaded	System shows no image uploaded error	Fail
FT-007	Test logout functionality	User is successfully logged out and redirected to the login page	Pass
FT-008	Test tutorial section functionality	User can access and navigate through the tutorial section easily	Pass
FT-009	Go back after prediction to upload page	User can go back to upload page after prediction	Fail
FT-0010	Test system security	Validate that the system implements necessary security measures	Pass

Table 7.2 Frontend test cases & test results

Chapter 8

Results & Discussion

8.1 Outcomes

In order to facilitate understanding, graphs of the validation accuracy and validation loss are also supplied. In order to facilitate understanding, graphs of the validation accuracy and validation loss are also supplied.

1. **Model Loss and Accuracy:** The trained model produced outstanding accuracy and loss values. The effectiveness of the model was assessed using parameters like loss and accuracy. The results that were attained are as follows:

Metric	Value
Loss	0.0920
Accuracy	0.9728
Validation Loss	1.0104
Validation Accuracy	0.7390

Table 8.1 Model loss & Accuracy

The model's predictions are near to the actual values, as evidenced by the achieved loss of 0.0920, leading to a very accurate classification of Alzheimer's disease. The accuracy score of 0.9728 demonstrates the model's high degree of accuracy in classifying brain pictures as normal or Alzheimer's disease.

2. **Validation Accuracy Graph:**

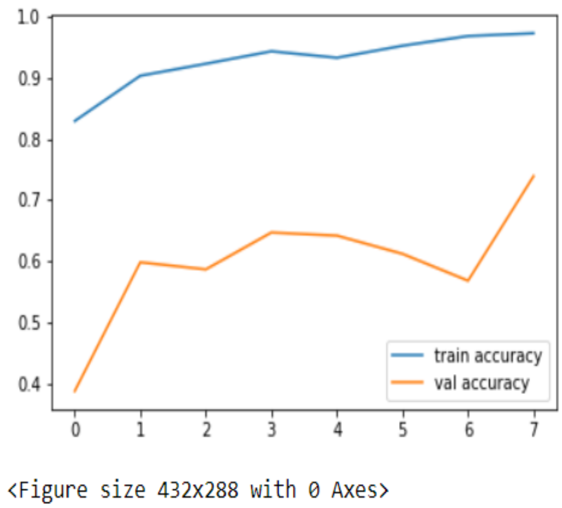

Fig.8.1 Validation Accuracy Graph:

The validation accuracy graph gives a graphic illustration of how the model's accuracy changes throughout training or remains stable. At each epoch or iteration, it displays the accuracy that was attained on the validation dataset. The graph shows how the accuracy of the model gradually rises over time, demonstrating successful learning and efficient classification.

3. **Validation Loss Graph:**

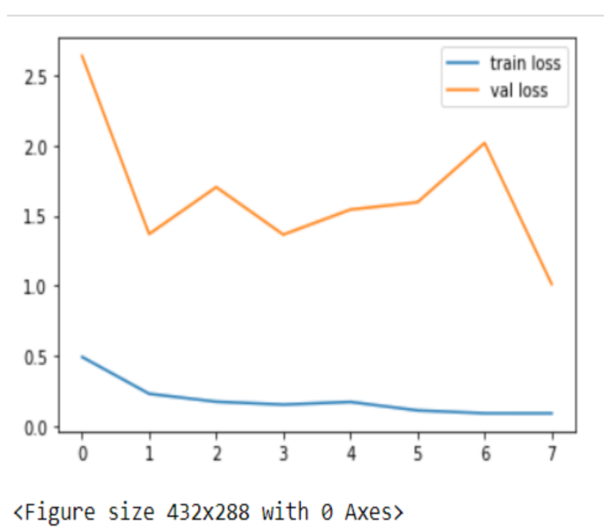

Fig.8.2 Validation loss Graph:

The model's loss on the validation dataset at each epoch or iteration is shown on the validation loss graph. It shows how the loss lowers with time, demonstrating the model's capacity to optimise its parameters and reduce training-related errors. A declining validation loss indicates that the model is improving its prediction accuracy by taking into account the data.

8.2 Screen Shots

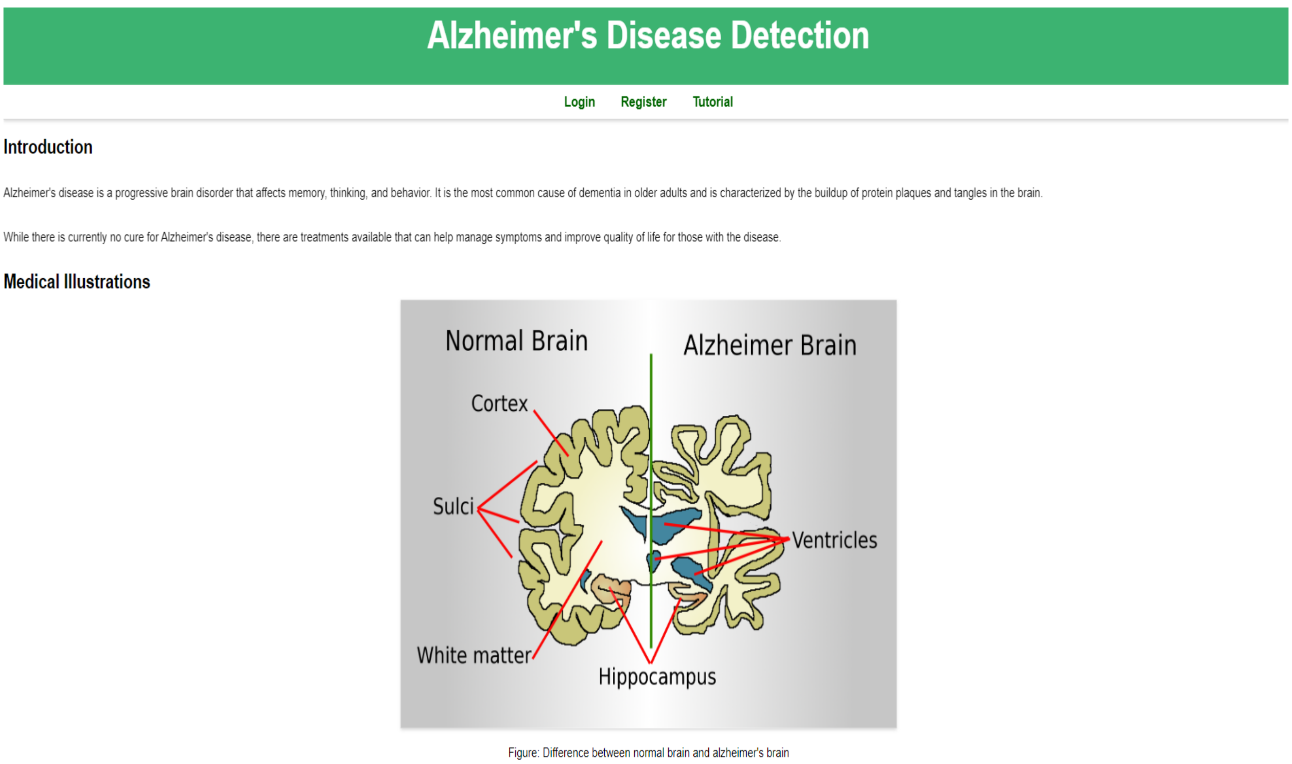

Fig 8.3 Home Page

The web application's home page acts as its front page. It gives a general summary of the system's features. A navigation menu or buttons that link to various application parts, such as login, registration, tutorial are frequently included on the home page.

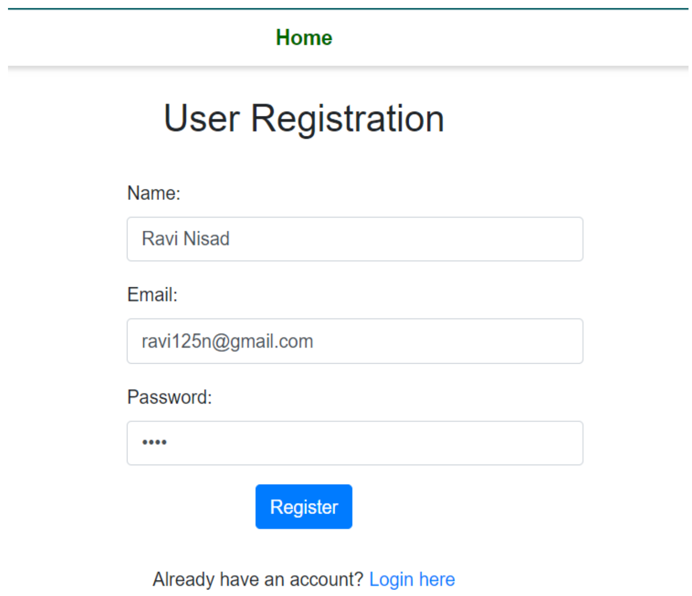

Fig.8.4 User Registeration

New users can register by completing a registration form on the registration page.

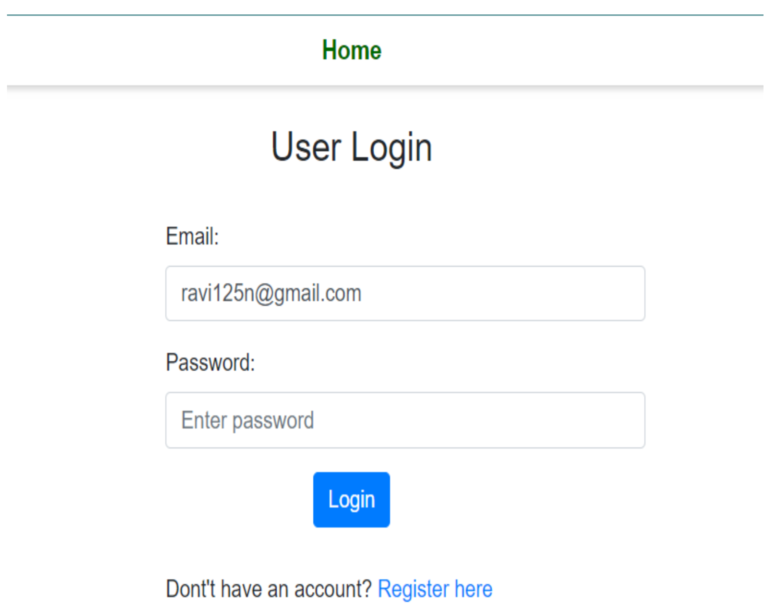

Fig.8.5 User Login

By entering their login information, including their email and password, individuals can access their personal accounts.

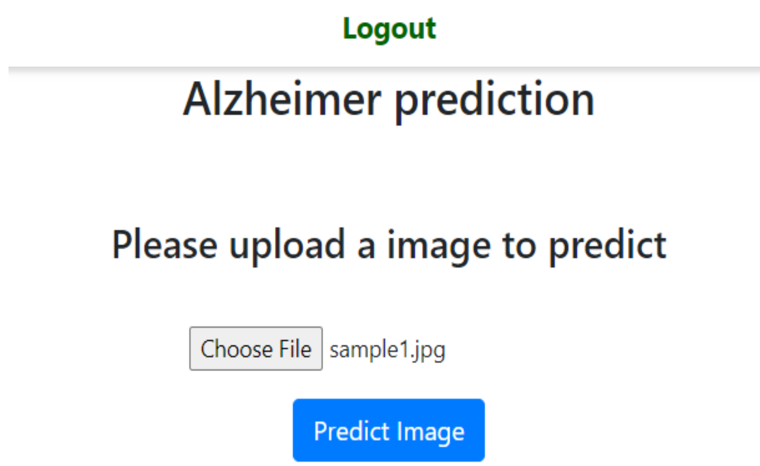

Fig.8.6 Upload Image & Predict

Users can choose and upload their MRI image using a file upload tool that is provided.

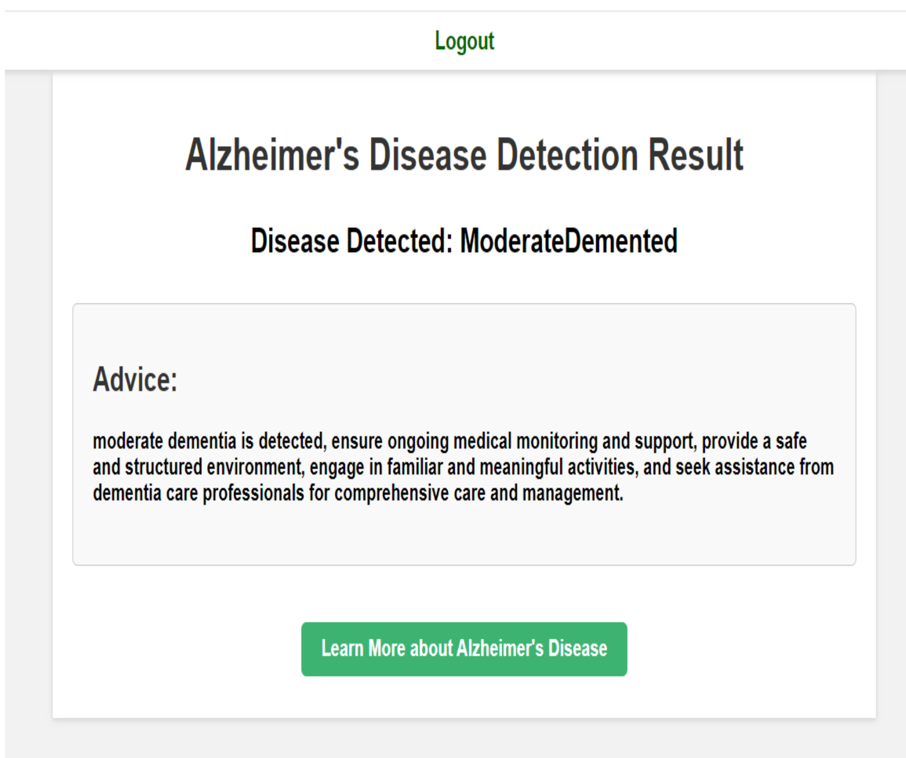

Fig.8.7 Result

User can get result of the disease detected and advice according to it.

Chapter 9

Advantages & Limitations

9.1 Advantages

1. **High Accuracy:** The system achieved accurate classification of brain images related to Alzheimer's disease.
2. **Efficient Diagnosis:** The system offers a fast and efficient diagnostic process, enabling early detection.
3. **User-Friendly Interface:** The user interface provides an intuitive experience for easy registration, image upload, and result viewing.
4. **Potential for Clinical Application:** The system has potential as a screening tool for Alzheimer's disease in clinical settings.

9.2 Limitations

1. **Dataset Limitations:** The system's performance relies on the quality and size of the training dataset.
2. **Limited Diagnostic Modalities:** The system focuses on Alzheimer's disease detection through brain image analysis, excluding other diagnostic approaches.
3. **Hardware Requirements:** The system may require powerful hardware resources for optimal performance.

Chapter 10

Future Scope

The project has many potential paths for growth and development in the future. First, in order to solve the shortcomings of the current system, additional research can concentrate on resolving issues such dataset size restrictions, expanding the CNN model's interpretability, and boosting the robustness of the classification method. The system's accessibility and usefulness can be substantially improved by making it a mobile application, which would enable users to carry out Alzheimer's disease detection anywhere. A complete and flexible platform for early detection and diagnosis would be provided by broadening the system's detection capabilities to include a variety of diseases, such as Parkinson's disease or other neurodegenerative disorders. These upcoming improvements would help identify a wider spectrum of neurological diseases and improve healthcare services.

Chapter 11

Conclusions

In conclusion, the project aimed to develop an effective Alzheimer's disease detection system using deep learning techniques.To improve the precision and effectiveness of the detection process, the project made use of approaches like Convolutional Neural Networks (CNN), transfer learning, region of interest extraction, and data preprocessig. The created technology showed encouraging results after comprehensive testing and review. The VGG-16 architecture and CNN-based model demonstrated greater performance in correctly categorising brain images as normal or suggestive of Alzheimer's disease. Using transfer learning reduced training time and increased overall accuracy by using pre-trained models and fine-tuning them for the specific goal of Alzheimer's disease identification. The developed system's high accuracy and user-friendly interface make it a valuable asset in the field of medical diagnostics. Further research and improvements can be made to expand the system's capabilities, incorporate additional diagnostic modalities, and enhance its overall performance in detecting Alzheimer's disease.

References

[1] Chima S. Eke, Emmanuel Jammeh, Xinzhong Li, Camille Carroll, Stephen Pearson and Emmanuel Ifeacho, “Early Detection of Alzheimer’s Disease with Blood Plasma Proteins Using Support Vector Machines”, IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 25, NO. 1, JANUARY 2021.

[2] Zhao Fan, Fanyu Xu, Cai Li and Lili Yao ,“Application of KPCA and AdaBoost algorithm in classification of functional magnetic resonance imaging of Alzheimer’s disease”, Neural Computing and Applications (2020).

[3] Md Shahin Ali, Md. Khairul Islam, Jahurul Haque, A Arjan Das, D S Duranta and Md Ariful Islam ’“Alzheimer’s Disease Detection Using m-Random Forest Algorithm with Optimum Features Extraction”, 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA).

[4] Subhabrata Roy · Abhijit Chandra, “On the detection of Alzheimer’s disease using fuzzy logic-based majority voter classifier”, Journal of algebraic statistics, volume13, No.3, 2022.

[5] Mathews Emmanuel, Dr. J. Jabez, “An Enhanced Fuzzy Based KNN Classification Method for Alzheimer’s Disease Identification from SMRI Images”, Multimedia tools and applications January 2022.

[6] X. Chen, D. Xie, Q. Zhao, and Z.-H. You,“MicroRNAs and complex diseases: from experimental results to computational models”, Briefings Bioinformatics, vol. 20, no. 2, pp. 515–539, 2019.

[7] Satoko Ito; Marcel Wira; Ruck Thawonmas,"User Friendly Minecraft Mod for Early Detection of Alzheimer's Disease in Young Adults", IEEE Games, Entertainment, Media Conference (GEM), 2022.

[8] S. Saravanakumar; T. Saravanan,"Early Alzheimer's Disease Detection Using Semi-Supervised GAN Based on Deep Learning", IEEE VLSI Device Circuit and System (VLSI DCS), 2022.

[9] Shichen Jin; Peini Zou; Ying Han; Jiehui Jiang,"Unsupervised detection of individual atrophy in Alzheimer's disease", 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021.

[10] G. Palacios-Navarro; J. Buele; S. Gimeno Jarque; A. Bronchal García,"Cognitive Decline Detection for Alzheimer's Disease Patients Through an Activity of Daily Living (ADL)", IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022.

[11] Nayan Anand Vats; Purva Barche; Ganesh S Mirishkar; Anil Kumar Vuppala,"Exploring High Spectro-Temporal Resolution for Alzheimer's Dementia Detection", IEEE International Conference on Signal Processing and Communications (SPCOM), 2022.

[12] Dalin Yang; Ruisen Huang; Kunqiang Qing; Keum-Shik Hong,"Exploring the possibility for early detection of Alzheimer's disease with spatial-domain neural images",2020 20th International Conference on Control, Automation and Systems (ICCAS), 2020.

[13] D. Deepa; M. Sundar Raj; S. Gowthami; K. Hemalatha; P. Thangavel ;C. Poongodi;,"Identification and Analysis of Alzheimer's Disease using DenseNet Architecture with Minimum Path Length Between Input and Output Layers", Smart Technologies, Communication and Robotics (STCR), 2022.

[14] M. Lavanya; Radha Raman Chandan,"Machine Learning-based Alzheimer's Disease Prediction using Personalized Methods", ICOSEc, 2022.