Autonomous assistance for Alzheimer’s patients

*Nithin Kumar*

Department of Information Technology Engineering

*Hindustan Institute of Technology and science*Chennai, India

julianmr@hindustanunic.ac.in

*Sathyalakshmi S*

Department of Computer Science Engineering  
*Hindustan Institute of Technology and science*Chennai, India

slakshmi@hindustanuniv.ac.in

Rishwanth Praveen   
Department of Computer Science Engineering  
*Hindustan Institute of Technology and science*Chennai, India  
rishwanthpraveen@gmail.com

Julian Menezes  
Department of Information Technology Engineering  
*Hindustan Institute of Technology and science*Chennai, India  
julianmr@hindustanunic.ac.in

Autonomous assistance for Alzheimer’s patients

Abstract

Alzheimer's disease (Ad) is a disease that becomes more and more inhibiting of memory with a progressive course over time it also affects thinking and causes visual impairment. Visual impairment is a customary symptom of AD, and recent studies have shown that visual interventions may improve the functioning of AD patients. This paper aims to use image processing techniques to detect AD at an early stage using Magnetic Resonance Imaging (MRI) of the brain. The proposed technique also includes camera-based detection for visually impaired or AD individuals to recognize text, objects and known people in real-time using face recognition and sign language. Optical Character Recognition (OCR) method is used to transform pictures to text, and ML is used to detect known faces, objects, and sign language. The output will be expressed in the form of an audio stream with the help of text to speech.

Keywords — Face detection, Ocr , Sign detection , Tesseract , MRI, API, AD,

# 1. Introduction

Visually impaired individuals face many challenges in their daily lives, including difficulties in reading printed text, recognizing faces, and identifying objects. Assistive technology has the prospective ability to alleviate challenges and enhance the standard of living for individuals with visual impairments.. The number of people with near or distance vision impairments worldwide is estimated to be at least 2.2 billion. One billion - or almost half - of these cases are yet to be treated for vision impairments. Recent advancements in OCR and face recognition technologies have made it possible to develop camera-based systems that can accurately recognize text, faces, and objects. However, these systems have not yet been fully explored for their potential to assist visually impaired individuals. There are several studies that have developed OCR and face recognition systems for the visually impaired, but these systems have limitations such as not being able to recognize text in real-time, not being able to recognize faces and objects. We additionally add sign to speech as an additive functionality in the integrated system that will help other people understand the sign language being expressed by the affected

Jennifer R. Evans et al [1] conducted a study in which 13900 individuals resulted showing that the presence of depression was rife within individuals who were visually impaired to those who are not visually impaired. This suggests that there may be a link between visual impairment and depression. Likely it is possible that individuals with visual impairment are to a greater extent prone to experiencing difficulties in functioning their daily lives. These difficulties can have a negative impact on an individual's mental well-being, leading to depression which could lead to social isolation, which is another risk factor for depression. In conclusion, this research highlights the importance of recognizing the potential impact of visual impairment on mental health, and the need for interventions that can help alleviate the burden of depression in this population.

Maisha Mashiata et al. [2] conducted an all-inclusive sweeping review on the status quo and anticipated prospects of assistive technology catered to the visually impaired individuals. The review aimed to provide an overview of the various assistive devices that are currently available, as well as their features and limitations. Various types of assistive devices are being used for visually impaired individuals , including capturing devices , working hours, response times , coverage areas and feedback . The devices reviewed include those that use image processing, computer vision, and machine learning techniques to assist users with visual impairments. The review also discussed the features of the assistive devices, such as the type of sensors used, their working hours, response time, coverage area, and feedback mechanisms. The devices were also evaluated on their ability to provide assistance to users in different scenarios, such as indoor and outdoor environments. However, the review concluded that at the time of the review, there was no single device that met all the requirements for designing assistive technology for visually impaired individuals. In conclusion, the review highlights the need for continued research and development in assistive technology for visually impaired individuals to create devices that can meet the needs of all users. The review suggests that it is important to consider the trade-offs between different features when designing assistive technology for visually impaired individuals, in order to provide the best possible assistance to users in different scenarios. For this reason , we aim to provide a fully provisioned system for persons with visually impaired that will contain 5 modules which are MRI Alzheimer’s detection , OCR , Sign to Speech , Object detection and face detection that will help the visually impaired to perform basic functions

Three pretrained networks were shortlisted to detect Alzheimer’s in a MRI image VGG-16, Resnet50, SE Resnet50 of which Resnet seemed to be more accurate than the others [3].The Haar cascade classifier is used due to its quick result , low computational costs and reliability [5].The Haar cascade for face detection then use the LBPH ( local binary pattern histogram ) algorithm for face recognition due to its known performance and ability to recognize the face from the frontal , rear or side view [6].The sign to speech module will be carried out using a regular CNN with tuned pretrained models [7].Pytesseract the wrapper for google ocr and python will be used for the OCR module [8,9]. Object detection is carried out using the Caffe model [10]. The output is in form of an audio stream which is converted with the help of GTTS (Google text to speech )[11].

# 2. Related Works

In the field of assistive technology for the visually impaired, there have been several studies and projects related to camera-based text recognition and object identification that has ranged over a long period of time with significant developments. Jiri Martinek et al [12] presented an OCR system that can perform two crucial tasks: analysing the layout of a page, including identifying text blocks and lines, and recognizing text using OCR. Jongkil Hyun et al.[13] An architecture for face detection using a Haar classifier was proposed that aims to enhance the computational speed of the system. This is achieved by eliminating unnecessary iterations during the classification process hence increasing the processing speed by 4.46% , which can be beneficial for real-time applications such as security systems and video surveillance. Additionally, the use of Haar classifiers in this architecture is a conventional system for detecting faces in images, making it a reliable and robust solution

Chunming Wu et al.[14] an experiment with a 99.85 accuracy rate was achieved by the improved facenet in the experiment which demonstrates the ability to effectively recognize faces in real-time. A face detection and recognition system that uses this technology can detect and recognize faces with high accuracy

Sustainability is a key concern in face detection systems, as they often rely on small processing units.

Gabriel B Holanda Et al [15] proposed a reading system was proposed that utilizes a portable device and combines ML techniques , segmentation and feature extraction to achieve an accuracy of 99.86% and a specificity of 99.3%. The system integrates a central moments extractor with a multi-layer perceptron.

Sheirg wang et al. [16] 400 training images and 100 gaze attention maps were classified using a classification algorithm which has Resnet 50 as a backbone. The first initial epochs don’t show considerable improvement but after 10 epochs there was significant improvement Resnet 50 proved to be a conventional system to classify medical images.

# 3. Proposed Methodology

Our objective is to develop a Dual Aid system that comprises two main components: the Doctor's Aid and the Patient's Aid. The Doctor's Aid module is designed to offer doctors a preliminary understanding and a starting point for analyzing MRI scans of patients with Alzheimer's disease. This model can be deployed on a server to provide local access within the hospital's network. Considering that visual impairment is a common symptom in Alzheimer's disease, the Patient's Aid is divided into four sub-modules that can help patients with Alzheimer's perform basic activities. These sub-modules include face detection for identifying family members, object detection for recognizing common objects in the environment, OCR for reading large print signs, and sign-to-speech, which allows patients who communicate via sign language to converse directly with their doctors.

## Doctors aid

We propose an ensemble method that concatenates Resnet 50 and Xception that will perform a multiclass classification of four categories of MRI images, namely Mild Demented, Moderate Demented, Very Mild Demented, and Non-Demented. An Augmented dataset from Kaggle containing 39000 images for training and 6000 images for testing was used.

### Resnet50 and Xception

The reason for choosing this approach is that ResNet50 and Xception have complementary strengths - ResNet50 is proficient at capturing long-range dependencies, while Xception excels at capturing local features. Since the details in MRI images are often subtle, combining both these architectures allowed us to capture both local and global features accurately. We selected ResNet50 for its complexity, having 50 layers, which makes it more effective at capturing complex patterns present in MRI images, crucial for dementia classification. Furthermore, ResNet50 is a conventional backbone algorithm for classifying medical images, making it a preferred choice. We leveraged the pre-trained weights of ResNet50 from ImageNet dataset to bring prior knowledge to the task of dementia classification. Similarly, Xception, known to work well with ImageNet weights, was combined with ResNet50. The availability of a wide range of resources for both ResNet50 and Xception made it feasible to use them together to form an Ensemble method.Resnet 50 is a adaptation of the Resnet model having 50 layers which consists of 48 convolutional layer with 1 Average pool and 1 Max pool layer.Furthermore, Xception uses skip or residual connections to facilitate the flow of information from earlier layers to later layers, allowing the network to preserve significant information throughout the training process. This is accomplished by adding shortcut connections between the layers of the network, allowing for direct communication between them.

## Patients Aid

# The primary module of the patients aid is the face detection module where the Haar cascade classifier is used due to its quick result , low computational costs and reliability .The Haar cascade for face detection then use the LBPH ( local binary pattern histogram ) algorithm for face recognition due to its known performance and ability to recognize the face from the frontal , rear or side view .Optical character recognition also known as OCR will be carried out by pytesseract that will help the patients read big signs .For object detection we are using the caffe model that can detect multiple objects in a screen Sign to speech will be carried out with a Pretrained CNN Sequential model. These algorithms have been selected upon extensive literature survey so that the selected algorithms work properly in all conditions where speed or accuracy is not compromised for reliance. All these algorithms are pretrained network hence the inbuilt databases were used.

# 4. Simulation and Analysis

We performed simulations to investigate the behaviour of different algorithms with different optimizers. We experimented with several algorithms and optimizers, and after analysing the results, we shortlisted two algorithms: Xception and Resnet 50. Xception is a variant of the Inception network, while Resnet 50 is a variant of the Resnet network. Furthermore, we shortlisted three known optimizers for these algorithms: Stochastic Gradient Descent (SGD) optimizer, Adaptive Moment Estimation (Adam) optimizer, and Root Mean Square Propagation (Rmsprop) optimizer. These optimizers are widely used in deep learning for optimizing the parameters of neural networks. The shortlisting process was based on the performance of the algorithms with the different optimizers, computational efficiency, or other relevant factors. The shortlisted algorithms and optimizers was hence chosen to be further analysed and compared to determine the best combination for a specific task.

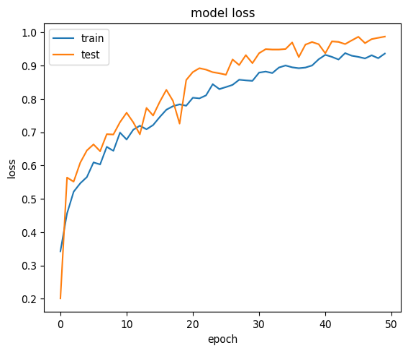
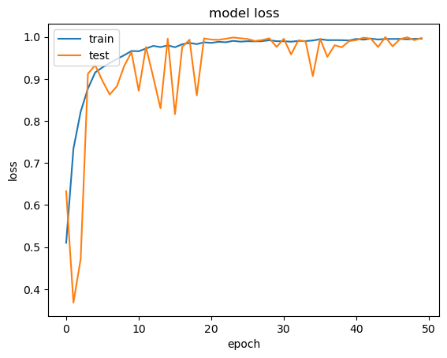
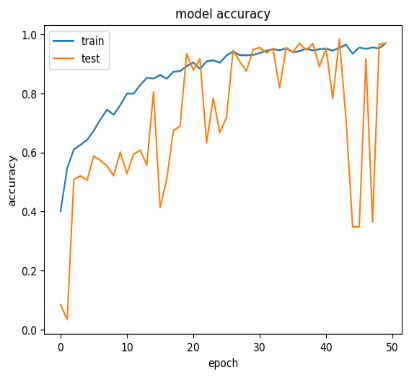


Figure 1: Ensemble Model trained with Rms prop, Adam and Sgd optimizers respectively

As observed for the given dataset Resnet 50 performed better with the Sgd optimizer giving us a more stable test accuracy and also higher highs and less lower lows as compared to Adam optimizer. Similar observations were observed for Xception paired with Sgd and Adams optimizers It was additionally notices that xception wasn’t able to get trained effectively during initial epochs and there was an accuracy drop of around 10-12 % in the testing accuracy.Resnet 50 and Xception concatenated with SGD optimizer performs with increased stability through the epochs. Extra Augmented data was added to the ensemble model. All the simulations were carried out for 10 epochs with a learning rate of 0.001. Rms prop doesn’t seem to be a competent optimizer for this ensemble method as we can see there is a difference of about 15 % and fluctuation in the test accuracy.

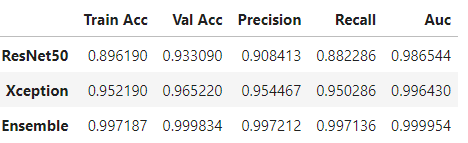


Figure 2. Accuracy Table

The ensemble method had 10 % better training accuracy than Resnet50 individually and 4 % better training accuracy than Xception individually. The validation accuracies were also better by 6 % and 3% respectively for Resnet50 and Xception models.

The Patient’s Aid module was able perform the respective functions during testing which was run on a simulative software that had 5 buttons that have different functions

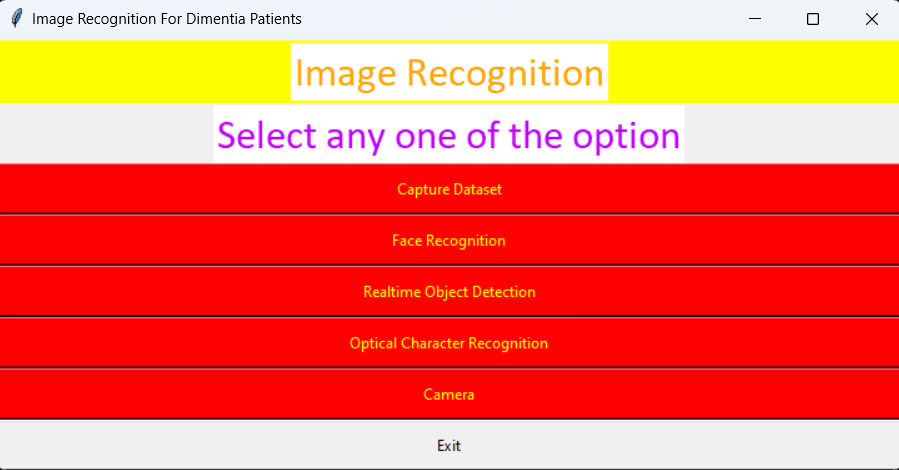


Figure 3. Patient’s Aid User Interface

# Conclusion

In conclusion, The Patient’s Aid module was able to run in a controlled environment with minimal errors upon observational analysis of confidence percentage and an effective framework has been built to for a multiclass classification where an ensemble method of concatenating the Xcpetion and Resnet 50 algorithm paired with the Sgd optimizer has proved to a viable method to classify the available MRI axial cut images for the given dataset .

# Future Work

However, we need to extend the functionality of this framework to include the coronal and sagittal brain slices and add other data modalities such as, FDG-Positron Emission Tomography(PET), FDG-Positron Emission Tomography(PET), , Structural Magnetic resonance Imaging (SMRI), Computer Tomography(CT) and Functional Magnetic Resonance(Fmri). The face detection module, Sign to speech, object detection and ocr module are good solutions that can help the visually impaired to do basic functionalities which was achieved in a simulative process

# References

[1] Jennifer R. Evans, Astrid E. Fletcher, Richard P.L. Wormald,Depression and Anxiety in Visually Impaired Older People,Ophthalmology,Volume 114, Issue 2,2007,Pages 283-288,ISSN 0161-6420,

[2] Maisha Mashiata et al. Towards assisting visually impaired individuals: A review on current status and future prospects, Biosensors and Bioelectronics: X, Volume 12, 2022, 100265, ISSN 2590-1370,

[3] Theckedath, D., Sedamkar, R.R. Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks. SN COMPUT. SCI. 1, 79 (2020).

[4] Besnassi, M., Neggaz, N. & Benyettou, A. Face detection based on evolutionary Haar filter. Pattern Anal Applic 23, 309–330 (2020).

[5] J. Hyun, J. Kim, C. -H. Choi and B. Moon, “Hardware Architecture of a Haar Classifier Based Face Detection System Using a Skip Scheme,” 2021 IEEE International Symposium on Circuits and Systems (ISCAS), Daegu, Korea, 2021, pp. 1-4,

[6] Hwang, J.J., Kim, Y.M., Rhee, K.H. (2019). Faces Recognition Using HAARCASCADE, LBPH, HOG and Linear SVM Object Detector. In: Hwang, S., Tan, S., Bien, F. (eds) Proceedings of the Sixth International Conference on Green and Human Information Technology. ICGHIT 2018. Lecture Notes in Electrical Engineering, vol 502. Springer, Singapore.

[7] Ahmed, H.A., Mustafa, S.Y., Braim, S.Z. et al. An Intelligent Kurdish Sign Language Recognition System Based on Tuned CNN. SN COMPUT. SCI. 3, 481 (2022).

[8] Tafti, A.P., Baghaie, A., Assefi, M., Arabnia, H.R., Yu, Z., Peissig, P. (2016). OCR as a Service: An Experimental Evaluation of Google Docs OCR, Tesseract, ABBYY FineReader, and Transym. In: , et al. Advances in Visual Computing. ISVC 2016. Lecture Notes in Computer Science(), vol 10072. Springer, Cham.

[9] Hegghammer, T. OCR with Tesseract, Amazon Textract, and Google Document AI: a benchmarking experiment. J Comput Soc Sc 5, 861–882 (2022).

[10] Ikeda, M., Oda, T. & Barolli, L. A vegetable category recognition system: a comparison study for caffe and Chainer DNN frameworks. Soft Comput 23, 3129–3136 (2019).

[11] Kaur, N., Singh, P. Conventional and contemporary approaches used in text to speech synthesis: a review. Artif Intell Rev (2022).

[12] J. Martínek, L. Lenc, P. Král, A. Nicolaou and V. Christlein, “Hybrid Training Data for Historical Text OCR,” 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 2019, pp. 565-570,

[13] J. Hyun, J. Kim, C. -H. Choi and B. Moon, “Hardware Architecture of a Haar Classifier Based Face Detection System Using a Skip Scheme,” 2021 IEEE International Symposium on Circuits and Systems (ISCAS), Daegu, Korea, 2021, pp. 1-4,

[14] Chunming Wu, Ying Zhang MTCNN and FACENET Based Access Control System for Face Detection and Recognition. Aut. Control Comp. Sci. 55, 102–112 (2021).

[15] Gabriel B. Holandaet al., Development of OCR system on android platforms to aid reading with a refreshable braille display in real time, Measurement,Volume 120,2018,Pages 150-168,ISSN 0263-2241,

[16] S. Wang, X. Ouyang,et.al, "Follow My Eye: Using Gaze to Supervise Computer-Aided Diagnosis," in IEEE Transactions on Medical Imaging, vol. 41, no. 7, pp. 1688-1698, July 2022,