Averting Frustration in Geriatric Patients through Identification, 2D Mapping and Navigation Using Machine Learning

Alekha Malhotra

Class of 2023,
Cathedral and John Connon School, 6, Purushottamdas Thakurdas Rd, Marg, Azad Maidan, Fort, Mumbai,
Maharashtra 400001 India.
Email-id: alekha.malhotra1@gmail.com

Stuti Bhatia

Class of 2023,
Cathedral and John Connon School, 6, Purushottamdas Thakurdas Rd, Marg, Azad Maidan, Fort, Mumbai,
Maharashtra 400001 India.
Email-id: bhatia.stuti@gmail.com

Reetu Jain

Supervisor, IEEE member,
On My Own Technology Pvt. Ltd., 2A, Pace House, 7 Swastik Society, Gulmohar Road, Vile Parle West,
Mumbai - 400049, India.
Email-id: reetu.jain@onmyowntechnology.com
ORCID: 0000-0002-7199-2823

Abstract: **Alzheimer's patients experience agitation and frequently display aggressive behaviors due to their inability to retain simple information, such as people's names, and often misplace items, among other causes. The incurable nature of this disease creates the need for assistive technology to help patients. This proposed methodology, MemoryMap, consists of a pair of glasses, a camera, hardware for audio output and input and a microprocessor. The algorithm uses image-processing and machine learning to perform object and facial recognition, map out the locations of both the object and patient and provides a navigation system to guide patients to any object they have misplaced. This prototype continually records the patient's view through the camera and performs real-time analysis on the video feed using OpenCV, allowing the identification of the object or face desired by the user through a NumPy model. Each object is recognised and entered into a database in the form of an excel sheet using XLWT and XLRT to maintain the location and time frame. Visual feedback of the navigation route along with audio feedback for recognition enables the patient to become more self-sufficient. This prototype is targeted toward patients in the Mild Stage of Alzheimer's and works towards relieving their irritation caused due to memory loss.**

Keywords: Alzheimer, Object recognition, Facial recognition, 2D Mapping, Navigation

I. Introduction

Alzheimer's disease (AD) is the most prevalent neurodegenerative disorder in the world with patients experiencing progressive cognitive and functional deficits, including deficits in short-term memory, executive and visuospatial dysfunction, and praxis, as well as behavioural changes. Following symptoms, patients live for an average of 8 years, which can vary based on the patient's symptoms and the progression rate of the disease.

These symptoms are prevalent to different extents during each of the five stages of Alzheimer's disease: preclinical Alzheimer's, mild cognitive impairment, mild dementia, moderate dementia and severe dementia due to Alzheimer's. Mild symptoms start becoming prevalent during the mild cognitive impairment stage, worsening with each progressive stage. At the mild dementia stage, symptoms include forgetting recent events, misplacing items, and an impaired ability to navigate and solve problems. Inability to recognize their belongings or even remember the faces of relatives can lead patients to irritability and thus depression. Patients can also act out aggressively and get lost, putting both them and their caretakers at risk.

Care and management of patients, as well as research regarding treatments, have resulted in high costs. In 2015, an estimated $818 billion was spent on various dementia-related activities, a large component of which is due to costs of nursing homes, other healthcare facilities and caregivers. These expenses, coupled with risks of depression amongst patients and caregivers, have given rise to a need for alternative aids and caregiving methods. Consequently, much research has been carried out to develop assistive technology for Alzheimer's and other Dementia related conditions.

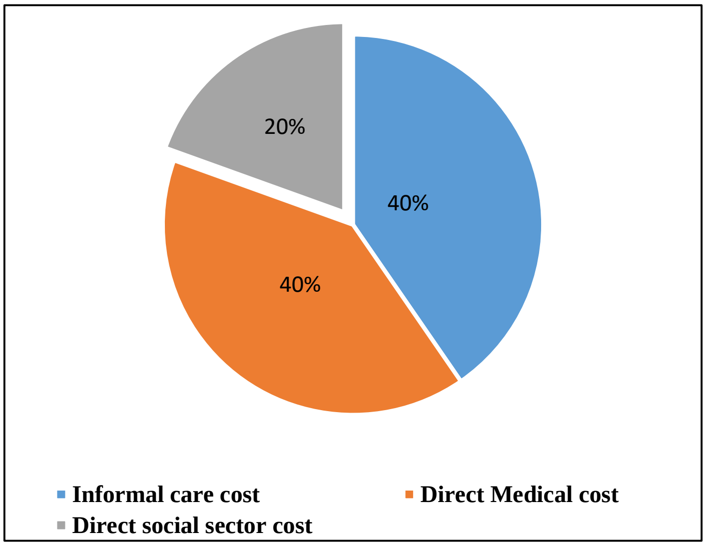

Figure 1: Graphical View of Alzheimer's total cost (%)

A. Literature Review

The last few years have seen the development of several specialized aids for Alzheimer's patients designed to tackle symptoms and behaviors such as memory loss, wandering, and language impairment. Location tracking is a very commonly used solution in several assistive devices. A majority of these types of aids employ Global Positioning Systems (GPS) to enable location tracking and navigation beyond patients' homes. Some aids, used for indoor tracking, commonly use RFID Tags, whose working relies on the presence of routers scattered within patients' homes and other commonly visited indoor areas. These are amongst the most common solutions in the existing literature. However, most of the devices proposed require additional equipment such as routers and networks provided by third party GPS providers. These add to costs and are inaccessible in certain regions. Additionally, some of these devices are highly specialized, offering limited services that still require much active involvement from caregivers. More holistic solutions have, however, been explored wherein Artificial Intelligence and Machine Learning have been repurposed to aid patients in tasks such as maintaining a daily routine and consuming medication at the prescribed time and in prescribed quantities. Although these systems do not affect the progression of the disease, they have been found to significantly reduce pressure and anxiety in patients and caregivers by compensating for everyday memory concerns.

B. Motivation and Novelties

This paper aims to evaluate and apply Machine Learning methods via a piece of assistive technology to aid Alzheimer's and other memory-loss patients in improving their quality of life and reducing the burden on caregivers. Wandering is common among Alzheimer's patients across the spectrum of stages. Although a vast literature exists on methods to track the location of patients outdoors, little R&D has been conducted to track locations indoors using expensive methods. Additionally, patients tend to lose their ability to recognize familiar places and faces, which serves as a source of frustration. In order to mitigate this frustration, the proposed device comprises a pair of spectacles with audio feedback that integrates facial and object recognition, indoor location tracking, and navigation for object-finding. Its multifunctional, compact, and cost-effective nature makes it a more ideal aid for AD patients.

II. Methodology

MemoryMap's user interface adopts an audio feedback approach, using audio from the users as input and outputting the solutions for the user through audio feedback and images. The algorithm relies on keywords, further explained in section II – E, to decipher whether the user wants the name of a person, assistance finding an object, or assistance navigating to a room, calling the necessary algorithm accordingly.

The device's hardware comprises a pair of glasses, a 620x420p camera to continuously record the patient's view, a speaker and microphone for output and input respectively, and a RaspberryPi to process the feed. The feed from the camera is constantly fed into the algorithm which identifies faces and objects in each frame. This data is recorded and used to assist the patients, as explained in Figure 2.

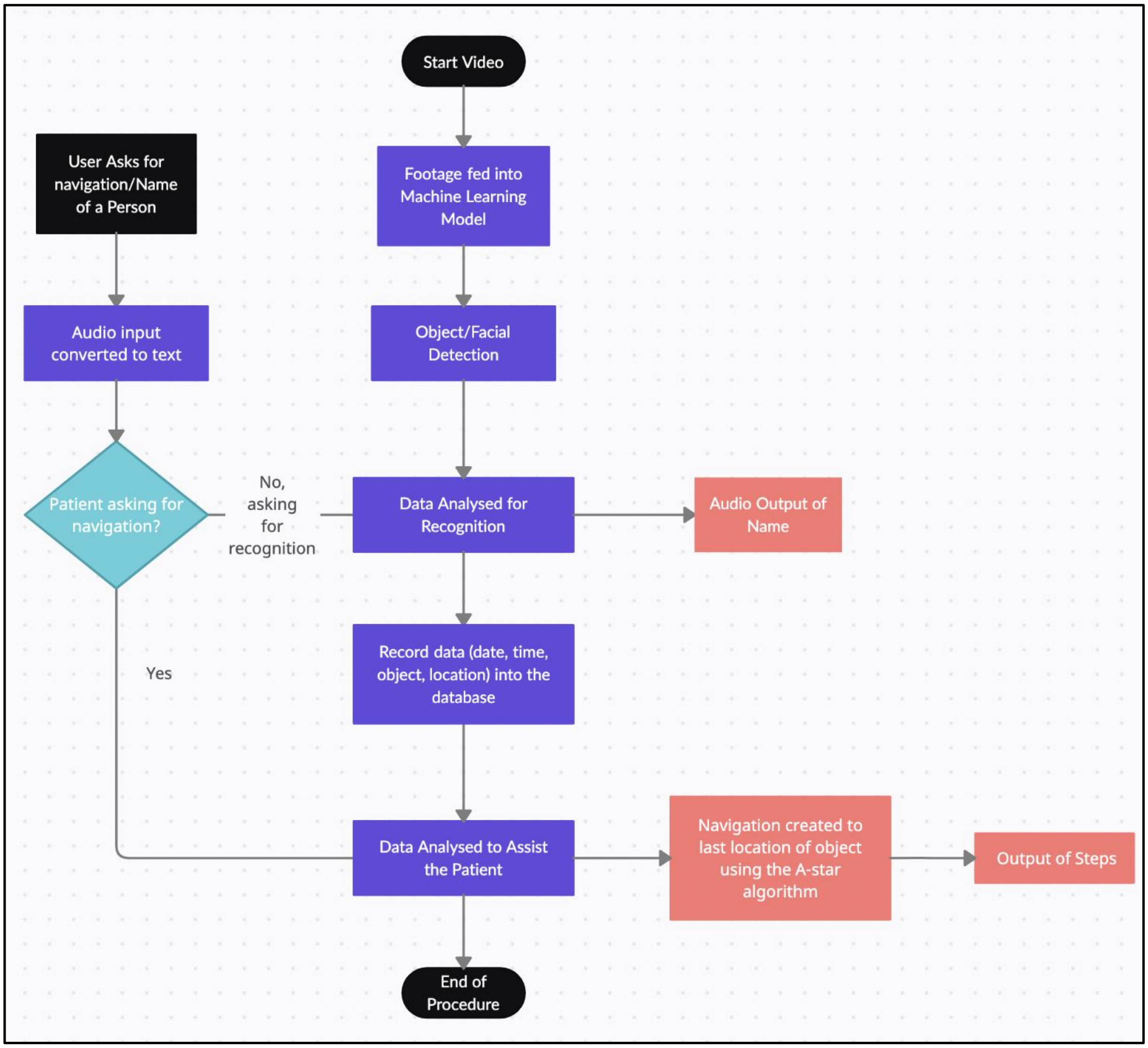

Figure 2: Flowchart of the working of MemoryMap

A. Facial Recognition

One of the most frustrating traits of AD involves the inability to recognize faces. This is particularly disheartening when patients are unable to identify family members and their caregivers. The device consists of a facial recognition feature to combat this. Patients will have the option of adding a new person that they wish to be able to recognize to the dataset. On encountering one of these people, the algorithm will recognize the individual and deliver their name to the patient in the form of audio feedback after giving the necessary speech command (outlined in section II – E).

a) Image Dataset: The dataset used for the process consists of 100 images for each person the patient wishes to be able to recognise. The images for preparation can be taken by the 720p camera embedded in the frame of the glasses or any conventional or web camera with the same or higher resolution.

b) Training the Dataset: A python script is written which resizes the images by cropping them to save computational time and power. The images are then converted into grayscale and translated into a numpy array. The *Haar Cascade Classifier* file detects the faces in each of the images, followed by a scan of each of the faces. The faces are then added to the datasets of the corresponding IDs or people. This concludes the training of the facial recognition model. It is stored in a trainer file that uses the extension '.yml'.

c) Facial Recognition: The trainer is utilized to correlate faces detected in the live feed and faces present in the device's database using the *Local Binary Pattern Histogram* (*LBPH*) algorithm. LBPH was found to be amongst the most preferred algorithms according to the OpenCV website due to its optimal usage of computation time and power, latency and accuracy.

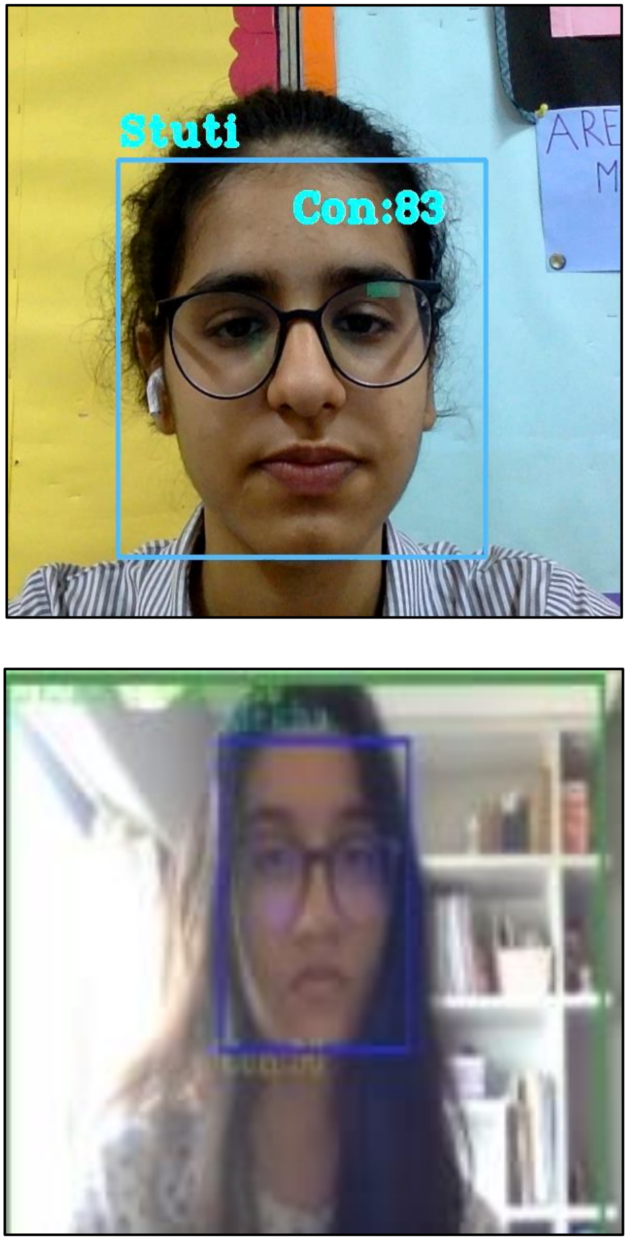

Figure 3: Facial Recognition Demonstration

B. Object Recognition

The device employs deep learning via SSD (Single-Shot Detector) as proposed by Liu et al., 2016. to extract the object's image from the overall capture. Literature suggests that there are two more commonly used object detection and recognition algorithms. Some, such as R-CNN and Fast(er) R-CNN use a double-step procedure during which regions where objects are expected to be found are identified and then to detect objects in those regions only. Contrastingly, single-shot algorithms such as You Only Look Once (YOLO) and SSD employ a procedure in which all objects in an image can be found in a single pass. These single stage detectors are not as accurate as the region proposal ones, but are faster and more efficient with adequate accuracy. Therefore, SSD was found to be most suited to the proposed device to be able to provide patients with immediate and accurate assistance.

Imutils is used to resize data from the video feed. The pre-trained neural network will then identify objects existing in the predefined and expandable classes, based on the patients' requirements. A NumPy array outlines the specific object. A confidence rate, time of recognition and current location will be outputted based on neighboring objects and using the Datetime library.

Figure 4: Object Recognition Demonstration

C. Database Management System

In order to keep track of the data for each patient, this algorithm uses Microsoft Excel as a database management system. As the algorithm recognises any object or face, it uses the recordInfo() method to append an Excel file with the object(s)/face(s), date, time of recognition and location using *XLRD* and *XLWT*. This file is iterated through whenever the user requests navigation to an object, feeding the location corresponding to the latest time that the object was recognised to the navigation algorithm.

Once an object has been recognised, the first step is writing the recorded data into an Excel file. Using Python's *datetime()* method, the date and time are returned during the time of recognition in their respective methods.

Having collected the Strings of the location and object/person from the recognition methods, the algorithm will check whether the path for the Excel file exists, using the *path.exists()* method, creating a new file by the name of "Record" if it does not, or opening this file if it does. Following this, the algorithm fetches the row to use using *nrows*. Having collected all this information, the algorithm appends this to the file using the *write()* method, resulting in a file such as in figure .

01:11:2021	09:18:19:PI	Stuti	Bedroom	Desk
01:11:2021	09:18:19:PI	Alekha	Bedroom	chair
01:11:2021	09:18:20:PI	Alekha	Bedroom	Desk
01:11:2021	09:18:21:PI	Alekha	Bedroom	bottle
01:11:2021	09:18:22:PI	Stuti	Bedroom	bottle
01:11:2021	09:18:23:PI	Stuti	Bedroom	bottle
01:11:2021	09:18:26:PI	Stuti	Bedroom	Desk
01:11:2021	09:18:27:PI	Stuti	Bedroom	chair
01:11:2021	09:18:28:PI	Stuti	Bedroom	chair

Figure 5: Image of Dataset for one patient

In order to fetch data, the application uses Python's *Speech Recognition* library to convert the audio input of the object, recognized based on key phrases discussed in section II – E. The String of the object name that the patient desires to be navigated to is passed into a loop that stores the last known location of the object, corresponding with the last row that it is in, along with the date and time in a variable, which is then passed to the Navigation algorithm.

D. Mapping and Navigation

The navigation module of this prototype uses the A* algorithm to find the shortest possible distance between the patient and the object. The location of the patient is detected based on their surroundings, using the same method as that of detecting the location of objects. The last known location of the object is extracted from the database using the aforementioned system. These two parameters are passed into the A* algorithm, which works in the following way.

This algorithm calculates the value ***f*** for each potential step, choosing the lowest ***f*** value. This value is the sum of the values ***g*** and ***h***. The value ***g*** represents the cost of going from the current step to the potential next step, while ***h*** represents the heuristic value, i.e. the potential cost of going from this potential next step to the final destination.

The heuristic value can be calculated in multiple ways but for the purpose of this algorithm the method of calculating the Manhattan distance is easily chosen, i.e. the sum of distances as measured along perpendicular axes. This method was chosen for its more realistic view around household objects, as well as it being convertible to easy-to-follow audio navigation instructions in the future. The Manhattan distance between the green and red boxes can be represented by the yellow boxes below.

To calculate these values, the layout of the house is converted into a grid, where each square is considered to be 1x1 units. These squares are all of the same size. The starting square, i.e. the current position of the patient, is coloured yellow, while the square representing the final position of the object is colored green. To account for the obstacles, such as walls, those squares are colored gray and appended to the *obstacle_list* of the algorithm, which indicates that the algorithm cannot consider it as a potential choice for the next step.

Using the above information, the cost of each potential path to the object, choosing the shortest path to display to the user, as seen in figure 6.

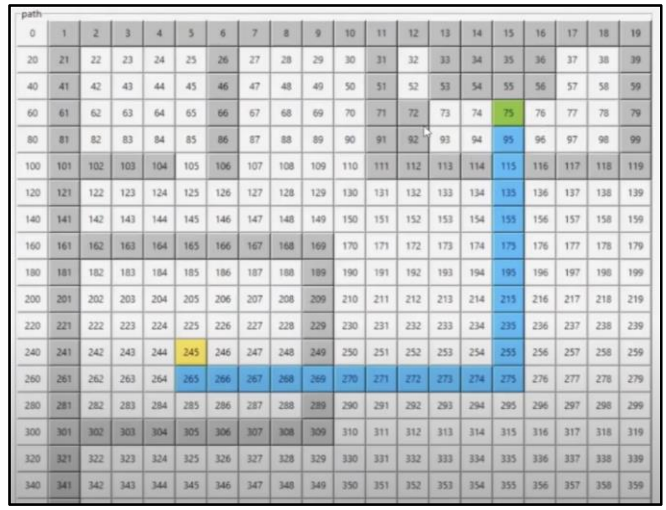

Figure 6: A* algorithm Demonstration

E. User Interface

The model is trained to recognise 3 key phrases which will prompt it to give the user output. The phrases are as follows:

1. Who/Face: This command prompts the algorithm to provide audio feedback of the faces that it recognises in the current frame.

2. Where and Object Name: This command prompts the algorithm to run the navigation method and display the shortest route to the object.

3. Where and Room Name: This command prompts the algorithm to run the navigation method and display the shortest route to the specified room.

The algorithm uses the *Speech Recognizer Library* to convert the audio file to text, which is then searched through for the keyphrases. The *recognize_google()* method is used to transcribe the audio file. This text file is iterated through to search for the keywords as specified.

In the case of facial recognition, once the user's command has been executed, the audio output is created for the user using *pyttsx3()*. In order to return audio feedback for the name of a face, the *say()* command is used, having the person's name as a String parameter passed from the faceRecognition() method.

For the navigation algorithm, the same form of audio input is used, however a picture displaying the route to the object is outputted instead, using the OpenCV library.

III. Prototype

The hardware of the current prototype is shown in figure 7 consists of a 620*420 camera, a speaker, a pair of spectacles and a Raspberry Pi 4 that will be enclosed in a protective casing. The camera can be attached to the top corner of either side of the frame. Similarly, the speaker can be fastened to either temple tip. Patients also have the option of adapting Memory Map to their existing glasses if they do not wish to purchase the spectacles that complement the rest of the hardware.

There is a wired connection between the camera and the Raspberry Pi 4 at present. The casing can be placed in a patient's pocket or attached to their attire through a clip for convenience.

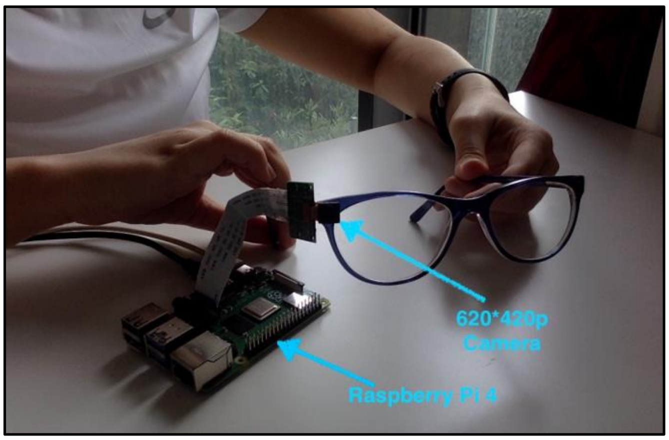

Figure 7: Prototype

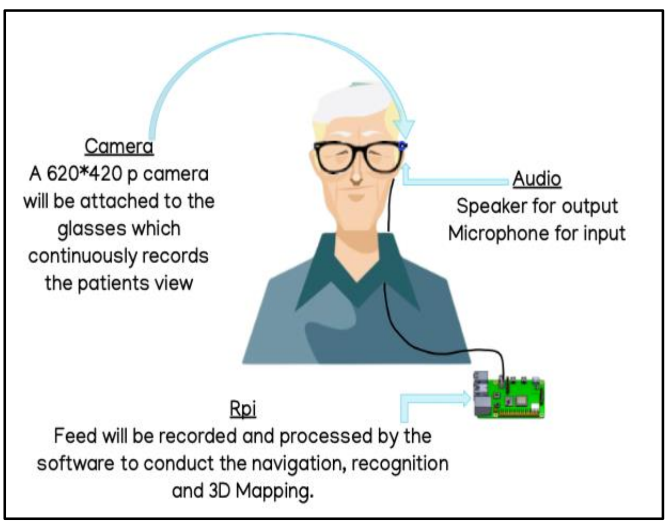

Figure 8: *Device Concept*

TABLE I: HARDWARE COSTS

Sl. No.	**Component**	**Cost (INR)**
1.	Raspberry Pi 4B	6500
2.	8MP Camera Module V2	1500
3.	Spectacle Frame (optional)	500
4.	Speaker	500
5.	9V rechargeable battery	250
	Total	9250

IV. Results

This application was created using Python 3.8 and was run on multiple systems, including MacOS Monterey (v12.4, 10 cores) and MacOs Big Sur (v11.6, 5 cores) To test the functionality of the system, the algorithm was tested on its ability to recognize faces, recognize objects, estimate object location and map a navigation route to the object. This algorithm was trained on approximately 970 free size images in the database for each object, 50 binary images of size 224x224 for each face and 30 images for each location.

Based on testing, this algorithm operates with an accuracy rate of 88%, 83%, 98% and 72% for object recognition, facial recognition, location finding of the object and speech recognition respectively. The Navigation model is trained on Teachable Machine, which uses the Google AI Machine with an accuracy of 96%. The overall accuracy rate is 92%. The following data forms a subset of the model's object and facial recognition testing, resulting in an accuracy rate of 88%:

TABLE II: OBJECT RECOGNITION SUMMARY TABLE

Object to be detected	**Items in the Image**	**Hit/Miss**	**Time taken for identification (s)**	**Confidence**
Football	Football and Volleyball	Hit	1.1	76%
Football	Football and Stuti	Miss	1.0	43%
Football	Football and Stuti	Hit	1.1	76%
Volleyball	Football and Volleyball	Miss	1.1	43%
Volleyball	Volleyball	Hit	1.0	88%
Volleyball	Football and Volleyball	Miss	1.1	77%
Stuti	Football, Volleyball and Stuti	Hit	1.0	92%
Stuti	Football, Volleyball and Stuti	Hit	1.0	85%

Stuti	Football, Volleyball and Stuti	Hit	1.0	97%
Football	Football, Volleyball and Stuti	Hit	1.1	77%
Volleyball	Football, Volleyball and Stuti	Hit	1.0	92%
Volleyball	Football, Volleyball and Stuti	Hit	1.0	78%
Volleyball	Football and Volleyball	Hit	1.0	79%
Volleyball	Football and Volleyball	Hit	1.0	89%
Volleyball	Football and Volleyball	Hit	1.0	99%
Football	Football and Stuti	Miss	1.0	31%
Football	Football and Volleyball	Hit	1.0	94%
Football	Football and Volleyball	Hit	1.0	95%
Stuti	Football, Volleyball and Stuti	Hit	1.0	100%
Stuti	Football, Volleyball and Stuti	Hit	1.0	92%

The above data indicates a sufficiently high level of accuracy for object and facial recognition, suggesting a bench mark of 60% for the confidence rating as the cut-off for accurate detection. Using this cut-off allows the algorithm to be more certain of its response. Using this data, the following confidence matrix was created, displaying 17 accurate detections (true positives), 3 false positives, 3 false negatives and 27 true negatives.

TABLE III: CONFUSION MATRIX TABLE

	Actual values	
Predicted value	***1***	***0***
1	17	3
0	3	27

The computed test accuracy for the model is 88%. An error distribution graph, which graphed the absolute number of errors for the football, volleyball, and Stuti, was created to assess the working of the algorithm.

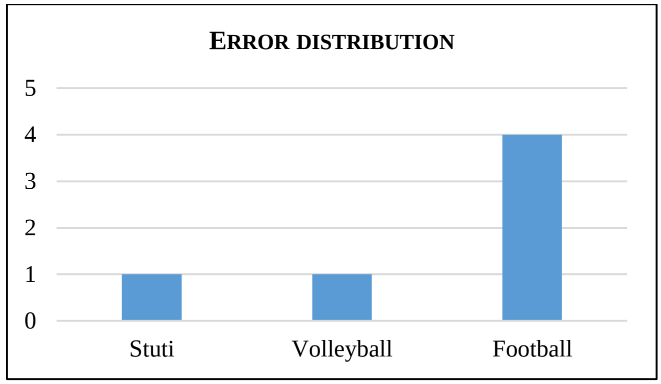

Figure 9: Error Distribution Graph

This graph indicates that the facial recognition accuracy is probably higher than the overall accuracy calculated. This data allows a better understanding for future development of the algorithm. While the accuracy rate is sufficiently high, it could be increased in the future by increasing the number of images that the object recognition algorithm is trained on.

V. CONCLUSION

Our current results are based on the usage of 970 free-size images in the database for each object, 50 binary images of size 224 x 224 for each face, and 30 images for each location. The accuracy rates can be easily improved by further training the model and increasing the number of images in our database. While our accuracy rates are high, we hope to improve them to be able to aid patients as seamlessly as possible. In line with this, part of the future scope for this device involves incorporating an automated 2D mapping process for navigation for indoor spaces and integrating that with a GPS for outdoor areas. Moreover, Memory Map can be made wireless through Bluetooth. Assistive technology for Alzheimer's is an emerging field, but few products work to mitigate frustration in patients by making them feel more self-sufficient. The proposed device is aimed at being a more suitable alternative by making patients feel more independent in not one but several ways by blending facial and object recognition, location tracking, and navigation.

REFERENCES

[1] García-Betances, R. I., Jiménez-Mixco, V., & Arredondo, M. T. (2014, August 7). Using Virtual Reality for Cognitive Training of the Elderly. *American Journal of Alzheimer's Disease & Other Dementias*. https://journals.sagepub.com/doi/full/10.1177/1533317514545866

[2] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016, December 29). SSD: Single Shot MultiBox Detector. *arXiv*. Retrieved June 29, 2022, from https://arxiv.org/abs/1512.02325?e05802c1_page=1

[3] Maliha, R., & Habib, M. A. (2021, October 16th). *A smartphone-based wander management system for Bangla speaking patients with Alzheimer's disease*. Springer Link. Retrieved June 17, 2022, from

https://link.springer.com/article/10.1007/s41870-021-00761-4

[4] Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., Prina, A. M., Winblad, B., Jönsson, L., Liu, Z., & Prince, M. (2017). The worldwide costs of dementia 2015 and comparisons with 2010. Alzheimer's & dementia : the journal of the Alzheimer's Association, 13(1), 1–7. https://doi.org/10.1016/j.jalz.2016.07.150

[5] Stages of Alzheimer's Disease. (n.d.). Johns Hopkins Medicine. Retrieved July 1, 2022, from https://www.hopkinsmedicine.org/health/conditions-and-diseases/alzheimers-disease/stages-of-alzheimer-disease

[6] Alzheimer's stages: How the disease progresses. (n.d.). Mayo Clinic. Retrieved July 1, 2022, from https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/in-depth/alzheimers-stages/art-20048448

[7] Xie, F., Oremus, M., & Gaebel, K. (2012). Measuring health-related quality-of-life for Alzheimer's disease using the general public. Quality of Life Research, 21(4), 593–601. http://www.jstor.org/stable/41445045

[8] Viviana Perilli, Fabrizio Stasolla, Stefania Maselli, Isabel Morelli (2018) Assistive Technology and Cognitive-Behavioral Programs for Promoting Adaptive Skills of Persons with Alzheimer Disease: A Selective Review. Journal of Neurological Research And Therapy - 2(3):14-29

[9] Balasubramanian, G.V., Beaney, P. & Chambers, R. (2021, November 15). Digital personal assistants are smart ways for assistive technology to aid the health and wellbeing of patients and carers. BMC Geriatr 21, 643. https://doi.org/10.1186/s12877-021-02436-y

[10] Hélène Imbeault, Nathalie Bier, Hélène Pigot, Lise Gagnon, Nicolas Marcotte, Tamas Fulop & Sylvain Giroux (2014) Electronic organiser and Alzheimer's disease: Fact or fiction?, Neuropsychological Rehabilitation, 24:1, 71-100, DOI: 10.1080/09602011.2013.858641

[11] Schumann, C., Alexopoulos, P., & Perneczky, R. (2019, April 26). Determinants of self- and carer-rated quality of life and caregiver burden in Alzheimer disease. International Journal of Geriatric Psychiatry, 34(10), 1378-1385. https://doi.org/10.1002/gps.5126