A Novel Approach to Detect Dementia using ML

1st Mamtha Mohan
Dept. of ECE
Ramaiah Institute of Technology (VTU)
Bengaluru, India
mamtha.m@msrit.edu

2nd Pavitha U S
Dept. of ECE
Ramaiah Institute of Technology (VTU)
Bengaluru, India
pavitha@msrit.edu

3rd Saliq Arshid
Dept. of ECE
Ramaiah Institute of Technology (VTU)
Bengaluru, India
saliqkitab10@gmail.com

4th Shafkat Bano
Dept. of ECE
Ramaiah Institute of Technology (VTU)
Bengaluru, India
fauziapapa2020@gmail.com

5th Srigouri Anegundi
Dept. of ECE
Ramaiah Institute of Technology (VTU)
Bengaluru, India
srigourianegundi@gmail.com

6th Yuvraj Singh
Dept. of ECE
Ramaiah Institute of Technology (VTU)
Bengaluru, India
helloiamyuvraj@gmail.com

2023 International Conference on Network, Multimedia and Information Technology (NMITCON) | 979-8-3503-0082-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/NMITCON58196.2023.10276307

***Abstract*—Dementia presents a global health challenge, involving cognitive decline and impairment that seriously impacts daily functioning and quality of life. The most prevalent form of dementia is Alzheimer's disease (AD), which is a progressive neurodegenerative disorder posing a major global health concern. Timely detection of AD is crucial for effective intervention, improved patient outcomes, and the development of therapeutic strategies. The development of a proficient medical decision support system plays a critical role in early disease prediction and assists specialists in prescribing appropriate treatments. The advancement of computer technology has fostered collaboration between healthcare professionals and engineers, resulting in the creation of intelligent systems for medical diagnosis. Artificial Intelligence (AI) has appeared as a promising tool in healthcare, including the detection of dementia, by leveraging machine learning algorithms and data analysis techniques. AI algorithms, particularly those based on machine learning, have been useful to diverse data types, such as MRI scans, to identify patterns and markers indicative of AD. The utilization of AI technologies can significantly improve the accuracy, efficiency, and objectivity of dementia screening and diagnosis processes.**

Keywords— dementia, Alzheimer's disease, magnetic resonance imaging, machine learning algorithms

I. INTRODUCTION

Dementia is a syndrome marked by a progressive decline in brain function, leading to issues such as memory loss, decreased cognitive abilities, and slower thinking. It is triggered by various brain diseases and can manifest in different forms, such as Alzheimer's disease, vascular dementia, dementia with Lewy bodies, and frontotemporal dementia. Dementia affects people across social, economic, and geographical boundaries, with age being a significant risk factor, typically occurring after the age of sixty. The global impact of dementia is substantial, with a new case emerging every seven seconds worldwide. Memory loss is the most common cognitive weakening associated with dementia, affecting around 10% of individuals over the age of 70 and 20-40% of those over 85. In India, it is predictable that 5.3 million people aged 60 and above have dementia. In accordance with a report dated 2020 by the Alzheimer's and Related Disorders Society of India (ARDSI), the number of patients is expected to rise significantly. By 2050, 11.44 million Indians will be living with dementia, in comparison to 3.84 million in 2019, as per Global Burden of Disease study published in the Lancet Public Health. Considering the growing elderly population and the declining joint family system in the country, dementia presents a significant challenge. Therefore, our initiative aims to develop an AI-based assistance system for dementia detection.

The contents of the paper is as follows. After the introduction discussed, related words have been discussed which describe two different research papers that have been referred to. It is then followed by methodology. The results from the ML model built are mentioned in the next section. The final section concludes our project and also talks about the future scope.

II. RELATED WORK

Dementia is a chronic syndrome that impacts a person's daily functioning due to the deterioration of thinking, behavior, and memory. Alzheimer's disease (AD) is the leading cause of dementia, with approximately 1 million new cases reported in India each year.[1] AD is a serious, neurological problem. Magnetic Resonance Imaging (MRI) of the brain is a valuable tool for detecting cerebral and neural abnormalities associated with AD.

A related work titled "Binary Classification of Alzheimer's Disease using MRI Images and Support Vector Machine" [2] proposed a prediction model that focused on observing textural feature differences in the affected brain region by analyzing axial brain slices. This study combines 2D axial brain slices from 2D MRI to construct a prediction model to distinguish between people with Alzheimer's disease (AD) and healthy controls. Dementia, in particular Alzheimer's disease (AD), is a chronic illness that impairs daily functioning because of declines in thinking, acting, and memory. The suggested method entails preprocessing the MRI images to improve image quality, followed by feature extraction utilizing Haralick features in the spatial domain and Gray Level Co- occurrence Matrix (GLCM) features.

979-8-3503-0082-6/23/$31.00 ©2023 IEEE

The spatial link between adjacent pixels is calculated using GLCM, which also produces characteristics like contrast, energy, correlation, and homogeneity. 13 pertinent features are provided by the Haralick features, which are obtained from the GLCM. Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Random Forest, and Linear Regression are some of the supervised machine learning algorithms used in the classification process.

In this work, the accuracy of these classifiers is compared. Without cross-validation, SVM performs better than the others, obtaining an accuracy of 82%.

The accuracy is then further increased by cross-validation, yielding 84% accuracy for a smaller sample of 200 photos and 73% accuracy for a bigger dataset of 6400 images. According to the study, early AD detection from MRI scans can be accomplished using Haralick features and SVM. It also highlights the potential for future research to incorporate further factors including age, gender, and Mini-Mental State Examination (MMSE) score for improved prognosis monitoring. Overall, the study shows how well MRI imaging and machine learning algorithms work together to help identify and classify Alzheimer's disease, which can have a big impact on early intervention and treatment planning.

Another literature related work titled “Detection of Alzheimer’s disease at Early Stage using Machine Learning” [3] uses different ML algorithms for detection of Dementia. In the paper, longitudinal MRI images was used as dataset which is taken from OASIS. Mini-Mental State Examination (MMSE), Estimated Total Intracranial Volume (eTIV), Clinical Dementia Rating (CDR), Normalized Whole Brain Volume, Atlas Scaling Factor are the five features used. Then preprocessing was performed. Handling missing data, Feature selection and Train-test split was done in this step. The best parameters are chosen using a K-fold cross validation, and the models are built using that. The following models are generated for detecting the AD. For Logistic Regression, Inverse of Regularization was used. For Support Vector Machine, Kernel Type, Kernel coefficient and Penalty parameters were used. For Decision Tree, max_depth of tree was used. And for Random Forest, the maximum number of trees in the forest (n_estimator), The maximum number of features to be considered for best split (max_features) and Tree’s maximum depth(max_depth) is used.

Finally, Precision, Recall and F1-score were calculated. Finally, the accuracy of the machine learning models on the test dataset was as follows: Logistic Regression achieved 76% accuracy, Support Vector Machine (SVM) achieved 81% accuracy,Random Forest achieved 83% accuracy, and Decision Tree also achieved 81% accuracy. In summary, this study demonstrates the potential of machine learning algorithms in early-stage Alzheimer's disease detection using longitudinal MRI data. The results suggest that Random Forest performed the best among the evaluated models, achieving the highest accuracy.

Hence, using the above-mentioned research papers, we built a ML model which can detect dementia using the MRI images of the brain. As mentioned in the first research paper, we used GLCM feature extraction to build the model using SVM algorithm. Later, as mentioned in the second research paper we also used Random Forest algorithm to fit the curve and also performed cross validation. However, the accuracy of the model wasn’t satisfactory. Hence, we used the Gabor Filter feature extraction to build the model. The final output after using the Gabor filter algorithm was quite better. The literature survey provided a strong foundation for our study and guided the selection of appropriate methodologies todevelop an effective and accurate machine learning model for early-stage dementia detection.

III. METHODOLOGY

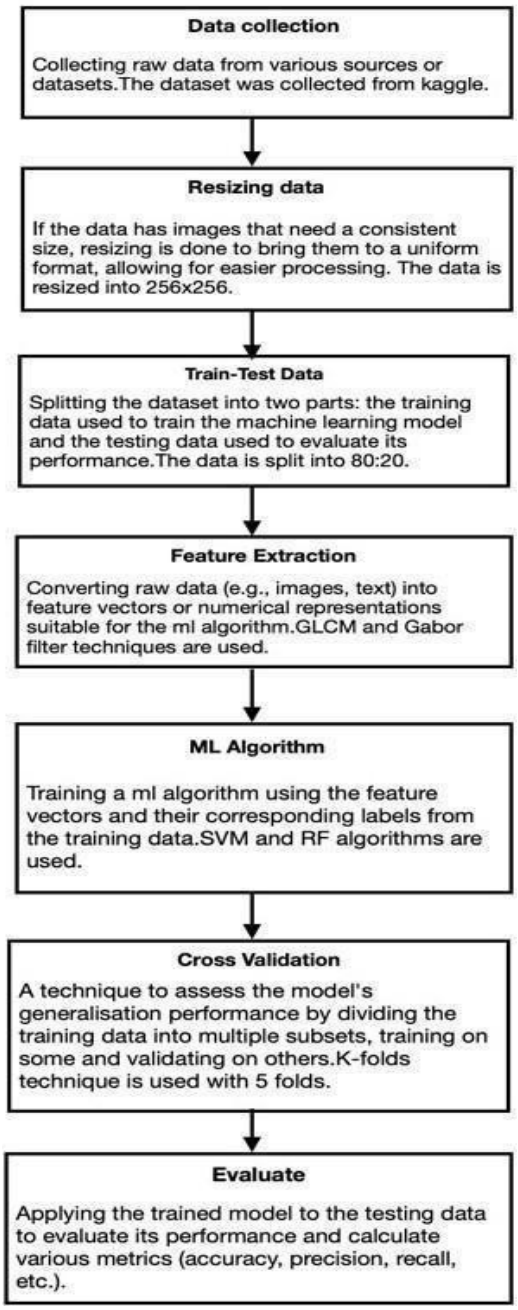


Fig. 1. Flowchart for methodology

A. *Collection of Dataset*

The dataset used is MRI images of the brain. The dataset is taken from Kaggle. The dataset consists of total 6400 MRI images.

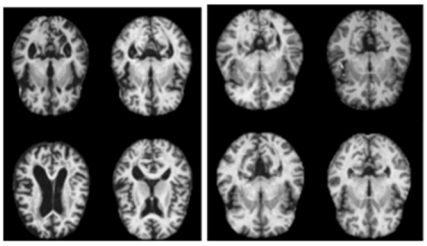
Class-1: Mild Demented (896 images)

Class-2: Moderate Demented (64 images)

Class-3: Non-Demented (3200 images)

Class-4: Very Mild Demented (2240 images)

The dataset consists of MRI images of patients suffering dementia extending from very mild to moderate. This makes sure that the detection can be done in early stages.[4] The dataset has various parameters but only the significant parameters that greatly help in predicting the disease



(A) (B)

Fig. 2. Dataset visualization for (A) Demented (B) Non-Demented

B.*Resizing*

Machine learning algorithms often need inputs of a consistent size. Resizing images to a standard size guarantee that they are all the same size. Large images contain a lot of data and processing them can be computationally expensive. Pre-processing is basically performed to ensure the removal of noise from input MRI images [5]. By resizing images to a smaller size, the amount of data that needs to be processed is reduced,which can improve the speed of training. Resizing images can help expand the generalization of a machine learning model. The images were resized to 256x256. In this step the images were divided in to 2 groups: NON for not demented and ALZ for demented.

C.*Train-Test data*

It's important to split your data into training and testing sets because if you evaluate the performance of the model on the same data you used to train it, you may end up with a model that is overfit to the training data. This means that the model may have memorized the training data and won't generalize well to new, unseen data. In this the ratio of training: testing data is 80:20.

D.*Feature Extraction*

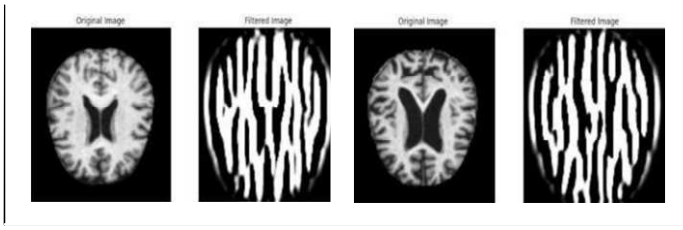
In machine learning, feature extraction is the method of selecting and extracting relevant attribute from raw data. In case of image analysis, feature extraction is to identify important characteristics, or features, of an image to be used to classify or identify the image. It is a developing practice for mathematical detection of changes in MRI that are not noticeable among image pixels, thus giving a quantitative method of extracting image features [6] We worked with 2 different feature extractions:

Gray-Level Co-occurrence Matrix (GLCM): It is a texture analysis technique used to extract features from images. The GLCM algorithm calculates the frequency of occurrence of pixel pairs with specific spatial relationships in an image. These pairs can be used to compute the statistical properties, such as contrast, homogeneity, entropy, and correlation. In the model we have used the standard features of the GLCM 4 matrix, i.e., contrast, energy, correlation and homogeneity along with an extra feature ASM (Algorithmic State Machines).

TABLE I FEATURES FOR DEMENTED AND NON-DEMENTED

Features	Demented	Non-Demented
ASM	0.182300	0.1877
Homogeneity	0.4984	0.4992
Energy	0.42697	0.43325
Correlation	0.9767	0.9732
Contrast	346.76	256.0

a) Gabor Filter: It is a texture analysis technique which uses a set of bandpass filters. It generates a matrix which gets convoluted with the image to give us a new image with highlighted features. It allows it to capture both the frequency and orientation of an image pattern. The parameters required are size of matrix, standard deviation, orientation, wavelength, gamma and offset.



(A) (B)

Fig 3. Dataset after Gabor Filter for (A) Demented (B)Non-Demented

E. *Machine learning model*

Machine learning model is an procedure or set of procedures that can be trained on a dataset to study patterns and relations in the data and used to make calculations or conclusions about new data.

The algorithms used are:

SVM (Support Vector Machines): It is a type of supervised learning model used for classification and regression analysis. SVM yields a good amount of generality as it is based on structural risk. In SVM, the goal is to generate

the optimal hyperplane which divides the data into classes with the largest margin possible minimization.[7]. SVM with linear kernel is used. To get neat and preprocessed data which is linearly separable, SVM creates a hyperplane separating the support vectors known as decision boundary [8].

Random Forest: It is an algorithm used for classifying and regression. It is a popular machine learning algorithm that uses combination of multiple decision trees to make more accurate predictions. In this, a set of decision trees are trained on parts of the input, with each tree making a prediction independently. The final prediction is then based on the majority vote of the individual tree predictions. The model implements 100 trees.

F. *Evaluation*

The evaluation is done based on accuracy. Accuracy is defined as the number of correct predictions made. Cross Validation is also performed. Cross-validation is a technique used in machine learning to evaluate the performance and generalization ability. The cross validation used is k fold with 5 folds. K-Fold cross validation used, using which the best parameters are chosen, and the best parameters are used in the model [9]

IV. RESULTS & DISCUSSION

In this study, two different feature extraction methods were employed to analyze the given dataset: Grey-Level Co-occurrence Matrix (GLCM) and Gabor Filter.

A. Accuracy after GLCM analysis

The GLCM method extracted five features from the images, namely: Angular Second Moment (ASM), Homogeneity, Energy, Correlation, and Contrast. SVM and Random Forest classifiers were applied to the feature arrays obtained from the GLCM feature extraction process, resulting in accuracy values of approximately 62.81% and 69.84%, respectively.

TABLE II ACCURACY AFTER GLCM TEXTURE ANANLYSIS

Algorithm Used	Accuracy
SVM	62.81%
RF	69.84%

B. Accuracy after Gabor filter analysis

Next, the Gabor Filter was utilized for feature extraction, and it significantly outperformed the GLCM method. The Gabor Filter achieved much higher accuracy of about 93.67% and 93.75% for SVM and Random Forest classifiers, respectively. This indicates that the Gabor Filter's ability to extract relevant features from the images was more effective in discriminating between classes compared to the GLCM method.

TABLE III ACCURACY AFTER GABOR FILTERTEXTURE ANANLYSIS

Algorithm Used	Accuracy
SVM	93.67%
RF	93.75%

C. Cross Validation:

The main idea behind cross-validation is to partition the available data into multiple subsets, or "folds," and then use each fold alternatively as a validation set while the remaining folds form the training set. The cross validation used is k fold with 5 folds. After the Gabor filter texture analysis and using the random forest algorithm, the cross validation was performed. Mean accuracy for 5 folds was 88.71%. Hence, detection of AD in early stages, is important as it may increase the odds of proactive actions in the future [10]

V. CONCLUSION

In conclusion, we have successfully developed an AI-based assistance system that utilizes MRI scan images to detect Alzheimer in individuals. This application holds great potential to be an accessible and valuable tool for patients suffering from dementia. While our current achievements are promising, there remains ample opportunity for further improvement. Currently, the model focuses on binary classification, determining whether a person has dementia or not. However, as a future direction, we aim to enhance the model's capabilities to provide a more nuanced and finer-grained classification, distinguishing between mild, moderate, and very mild stages of dementia. This expansion will enable us to better understand and diagnose the varying degrees of dementia severity in individuals.

The results of this study demonstrate that the Gabor Filter is a superior feature extraction method compared to GLCM for the given dataset. The Gabor Filter provided a substantial boost in classification accuracy, achieving impressive results with both SVM and Random Forest classifiers. The 5-fold cross-validation further confirmed the reliability of the Gabor Filter's performance, demonstrating its ability to maintain high accuracy across different subsets of the data. Therefore, we recommend using the Gabor Filter in combination with the Random Forest classifier for similar image classification tasks, as it yields superior results and offers better generalization capabilities compared to the GLCM method.

In the future, we plan to build a safety device for patients suffering from Dementia. The device could be used to track the location of the patient, detect falls and also display the name of the person in front of the patient using face recognition.

REFERENCES

[1] N. P. D, H. M. V, D. C, S. S and A. K. S, "Alzheimer's Disease Prediction Using Machine Learning Methodologies," 2022 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2022, pp. 1-6, doi: 10.1109/ICCCI54379.2022.9740942.

[2] U. R. K, S. S. S, U. M. G and V. B. C, "Binary Classification of Alzheimer's disease using MRIimages and Support Vector Machine," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), Hassan, India, 2021, pp. 423-426, doi: 10.1109/MysuruCon52639.2021.9641661.

[3] S. Pavalarajan, B. A. Kumar, S. S. Hammed, K. Haripriya, C. Preethi and T. Mohanraj, "Detection of Alzheimer's disease at Early Stage using Machine Learning," 2022 International Conference on Advanced Computing Technologies and Applications (ICACTA), Coimbatore, India, 2022, pp.1-5,doi:10.1109/ICACTA54488.2022.9752827.

[4] Sumit Salunkhe, Mrinal Bachute, Shilpa Gite, Nishad Vyas, Saanil Khanna, Keta Modi, Chinmay Katpatal and Ketan Kotecha, "Classification of Alzheimer's Disease Patients Using Texture Analysis and Machine Learning", Received: 6 June2021 / Revised: 27 July 2021 / Accepted: 28 July 2021 / Published: 4 August 2021.

[5] V. Patil and S. L. Nisha, "Detection of Alzheimer's Disease Using Machine Learning and Image Processing," 2021 International Conference on Smart Generation Computing, Communication and Networking, Pune, India, 2021, pp.1-5, doi:10.1109/SMARTGENCON51891.2021.96457.

[6] Cai, J.-H. et al. (2020) "Magnetic resonance texture analysis in Alzheimer's disease," Academic Radiology, 27(12), pp. 1774–1783.

[7] P. Lodha, A. Talele and K. Degaonkar, "Diagnosis of Alzheimer's Disease Using Machine Learning," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, pp. 1-4, 2018.

[8] J. Neelaveni and M. S. G. Devasana, "Alzheimer Disease Prediction using Machine Learning Algorithms," 2020 6th International Conference on Advanced Computing and Communication Systems, Coimbatore, India, 2020, pp. 101-104, doi:10.1109/ICACCS48705.2020.9074248.

[9] S. Dahiya, S. Vijayalakshmi and M. Sabharwal, "Alzheimer' s Disease Detection using Machine Learning: A Review," 2021 3rd International Conference on Advances in Computing,Communication Control and Networking (ICAC3N), Greater Noida, India, 2021, pp. 288-293, doi:10.1109/ICAC3N53548.2021.9725

[10] Shah, D. Lalakiya, S. Desai, Shreya and V. Patel, "Early Detection of Alzheimer's Disease UsingVarious Machine Learning Techniques: Comparative Study," 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), Tirunelveli, India, 2020, pp. 522-526, doi: 10.1109/ICOEI48184.2020.9142975.