Improved Decision Support System for Alzheimer's Diagnosis Using a Hybrid Machine Learning Approach with Structural MRI Brain Scans

Niranjan Kumar Parvatham, Lakshmana Phaneendra Maguluri*
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, India

***Abstract*—Alzheimer's disease (AD) causes damage to brain cells and their activities. This disease is typically caused by ageing, making people over the age of 65 more susceptible. As the disease progresses, it slowly destroys brain cells, making it harder to think clearly, recall things, and do everyday tasks. The end result of this is dementia. Metabolic disorders, such as diabetes and Alzheimer's disease, affect a substantial proportion of the world's population. While there is no permanent cure for AD, early diagnosis can help reduce damage to brain cells and support a faster recovery. Recent research has explored various machine learning approaches for early disease detection. However, traditional ML (Machine Learning) methods and deep learning techniques such as CNNs have not been individually effective in accurately detecting Alzheimer's disease (AD). In this proposed work, we developed a hybrid model that processes sMRI brain images to detect them as demented or non-demented. The model consists of two parts: the first part involves extracting significant features through a sequence of convolution and pooling operations; the second part uses these features to train SVM for binary classification. Data augmentation techniques such as horizontal flipping are used to balance dataset. We calculated key performance metrics essential for the healthcare domain, including sensitivity, specificity, accuracy, and F1-score. Notably, our model achieved an impressive accuracy of 99.60% in detecting AD, with a sensitivity of 99.83%, a specificity of 99.40%, and an F1-score of 99.58%. These results were validated using 15-fold cross-validation, enhancing the model's robustness for new data. This approach yields a more robust model, offering greater accuracy and precision compared to existing methods. This model can effectively support manual systems for detecting AD with greater accuracy.**

Keywords*—*Alzheimer's disease; binary classification; Convolutional Neural Network (CNN); horizontal flipping; healthcare decision support system; MRI images; Support Vector Machine(SVM)

I. INTRODUCTION

Alzheimer's disease is an illness of the brain that damages brain-active neurons, which is most likely due to aging. The aging population is increasing, leading to an increasing number of AD patients. It exerts detrimental effects on an individual's cognitive and interpersonal skills, progressively impeding their capacity to do routine activities [1], [2]. Alzheimer's disease (AD) is marked by the shrinkage of critical brain regions, such as the cerebral cortex and hippocampus, and this shrinkage results in damage to spatial and episodic memory, and disrupts the connection between the brain and body [3]. Due to progressive loss of brain cells, this leads to the accumulation of neurofibrillary tangles and amyloid plaques and a reduction in brain volume. The disease's rapid progression can weaken short-term memory, planning, and judgment [4].

The World Health Organization reports a significant increase in dementia cases, with the likelihood increasing with age, particularly among those over 65. Around 55 million people worldwide suffered from dementia in 2021, and experts predict that this number will increase to 139 million by 2050. On the other hand, only about 4% of young people experience early-onset dementia, which is generally caused by other health problems [5]. Alzheimer's disease (AD) affects roughly 5.7 million individuals in the United States. It has a substantial increase, multiplying by three in the mid-21st century. Alzheimer's disease was the sixth most common cause of death in the United States in 2015. An estimated 7.4% of India's population aged 60 and older suffers from dementia. This statistic indicates that approximately 8.8 million elderly people in the country live with this condition. The high prevalence of dementia among India's aging population highlights the growing need for increased awareness, support, and resources to address the challenges faced by those affected and their families [6].

The patient may exhibit symptoms of AD in relation to the loss of neurons in various parts of the brain. These symptoms can be detected in the clinical setting. In the mild stage, symptoms begin to disrupt some routine activities. In the moderate stage, symptoms interfere significantly with many daily activities. In the severe stage, symptoms profoundly impact almost all routine tasks. Currently, there is no permanent cure for Alzheimer's disease. However, if diagnosed at an early stage, effective medication can slow the progression of the disease. Early diagnosis allows for interventions that can significantly reduce brain damage, lower mortality rates, and enhance the quality of life for those affected.

AD Early diagnosis requires a comprehensive evaluation by a medical specialist, including both neurological and physical assessments [7], [8]. Treatment is more effective when it is given during the early stages of Alzheimer's disease [9]. Structural MRI, functional MRI, PET scans and other imaging modalities reveal significant changes in the brain associated with memory loss [10], [11]. Magnetic resonance imaging (MRI) [12] can be used as a biomarker to observe the size of damaged brain tissues [13]. To achieve higher levels of accuracy in predicting AD without false alarms, medical decision-support systems must incorporate an automated machine learning method that can

*Corresponding Author.

analyse MRI images at a deeper level in addition to human evaluations.

II. Related Work

In [14] a novel convolutional neural network (CNN) architecture that uses MRI data to accurately detect and classify Alzheimer's disease. The model achieves high accuracies of 99.43%, 99.57%, and 99.13% for categorizing AD across three, four, and five categories. The CNN architecture employs a hierarchical structure of convolutional, pooling, and fully connected layers to extract local and global patterns, enhancing clinical accuracy for early detection and disease monitoring. However, the study fails to address the applicability of the proposed CNN architecture to bigger and more varied datasets, and the focus on AD classification across three, four, and five categories may not cover the full spectrum of disease severity and subtypes. Further validation studies on external datasets are needed to evaluate the model's robustness and effectiveness.

The research [15] presented the use of DenseNet architecture for Alzheimer's disease classification, highlighting its effectiveness on MRI datasets. The model uses transfer learning techniques to improve accuracy and efficiency. Data augmentation is used to handle sparse data and generalize results. The model achieves an accuracy of 96.5% and an AUC of 99% in AD diagnosis. However, the study's limitations include prolonged computational time, manual hyperparameter tuning, and reliance on a single modality dataset.

The research [16] explored about the early-stage diagnosis of Alzheimer's disease (AD) in patients who are Cognitively Normal (CN) and introduces a dense neural network designed for binary classification. The findings demonstrate that this model surpasses traditional machine learning algorithms, achieving higher accuracy in distinguishing between AD and CN, AD and Mild Cognitive Impairment (MCI), as well as MCI and CN. Additionally, the study emphasizes the crucial role of computer-aided diagnosis using MRI for precise AD classification. By utilizing different activation functions, the model's classification validation accuracy is further enhanced. However, the focus on binary classification tasks may restrict the study's applicability to more complex diagnostic scenarios.

This study [17] focused on both white and gray matter from MRI scan images using 3D MRI technology. The procedure entails acquiring 2D slices from the coronal, sagittal, and axial planes of the 3D MRI scans. The process involves finding the most pertinent segments and performing feature extraction using Multi-Layer Perceptron (MLP) and Support Vector Machine (SVM) techniques to forecast and categorize Alzheimer's disease. The researchers assess the system's performance by utilizing metrics like precision, recall, accuracy, and F1-score. However, it is important to note that this study examines only MRI scan images.

This research [18] presented a novel method for categorizing MR images in the detection of Alzheimer's disease (AD) by utilizing graph kernels obtained from textural features of structural MR (sMR) images. The method entails dividing MR brain pictures into several regions and obtaining 22 unique texture characteristics. These characteristics are subsequently utilized to define the qualities of graph nodes. Graph-kernel matrices are created in sequence and then classified using Support Vector Machines (SVMs).

The findings demonstrated that this approach has outstanding performance in differentiating between individuals who are cognitively normal (CN) and those with Alzheimer's disease (AD), as well as between CN and Mild Cognitive Impairment (MCI) instances. More precisely, the technique obtained an Area under the Curve (AUC) of 0.92 for distinguishing between CN and AD, and 0.81 for distinguishing between CN and MCI. Nevertheless, its efficacy in distinguishing between MCI and AD was limited, as indicated by an AUC of 0.78.

The research [19] presented a comparative analysis of transfer learning versus conventional machine learning techniques for the early diagnosis and prognosis of Alzheimer's disease using structural brain MRI. It underscores the potential benefits of transfer learning in enhancing accuracy and efficiency, possibly reducing the reliance on large annotated datasets. The study emphasizes the significance of applying transfer learning in medical imaging for neurodegenerative diseases like Alzheimer's. The results indicate that the fusion of conventional machine learning systems outperforms ensemble transfer learning approaches, achieving an AUC of 93.1% for distinguishing Alzheimer's disease from cognitively normal (CN) individuals, 89.6% for identifying MCI converters (MCIc) from CN individuals, and an AUC range of 69.1–73.3% for differentiating MCI converters from non-converters (MCInc). However, further optimization is necessary when using networks pre-trained on generic images.

In [20], the author developed a comprehensive algorithm to predict disease using brain volume, cognitive and biological features, clustering algorithms, and Fuzzy inference systems. Deterioration refers to a significant global decline in mental function, not due to careful adjustment. This study employs machine learning algorithms to analyze data acquired from neuroimaging technology to identify Alzheimer's disease in its early stages. Support Vector Machine and Gradient Boosting are highly effective algorithms for classification problems and exhibit excellent performance in this endeavour.

In [21], the author focused on making use of machine learning techniques, specifically Support Vector Machine (SVM), to identify Alzheimer's disease. The study utilized Grey Level Co-occurrence Matrix and Haralick features to extract features from MRI axial brain slices. These methods analyze the texture of brain images to identify patterns associated with Alzheimer's. The resulting model achieved an accuracy rate of 84% in AD detection. This demonstrates the potential of combining SVM with advanced feature extraction techniques for accurate detection of AD in its early stage. The approach highlights the importance of texture analysis in neuroimaging.

In [22] author(s) have proposed a decision-support model utilizing deep learning and machine learning techniques to predict the one-year conversion probability from Mild Cognitive Impairment (MCI) to Alzheimer's disease (AD). This addresses a gap in the existing literature. The methodology involved extracting features with the help of CNN and classifying them with a support vector machine (SVM) classifier employing different kernels (linear, polynomial, and RBF). The model

achieved high classification accuracies of 91.0%, 90.0%, and 92.3% respectively, demonstrating its effectiveness. While the results were promising, some limitations need to be addressed. The study's data was limited, even after augmenting it by randomly choosing and flipping images. Increasing the dataset size is expected to improve classification accuracy. Additionally, the image data used was from the ADNI database, which consists of Western patients.

The study [23] introduced a method for detecting Alzheimer's disease by employing Support Vector Machine (SVM) on brain MRI data. The Support Vector Machine (SVM) model categorizes the disease into three stages: mild cognitive impairment (MCI), moderate Alzheimer's disease (AD), and severe Alzheimer's disease (AD).The model was trained and tested with MRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), involving about 70 AD patients and 30 normal controls. This diverse dataset allowed for effective training and testing of the model, despite its relatively small size, which may impact generalizability. While the SVM algorithm showed promising classification accuracy, the study did not address the interpretability of the model's decisions, a key factor for understanding the features driving the classifications.

In this study [24] the authors investigated the application of Support Vector Machine (SVM) in the diagnosis of Alzheimer's disease through the analysis of brain MRI scans and the classification of its stages. The algorithm underwent training and testing using MRI data obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI). The dataset included 70 patients diagnosed with Alzheimer's disease and 30 individuals without any cognitive impairments. The method involves extracting feature points from MRI images using Speeded up Robust Features (SURF) and analyzing these features with the Gray Level Co-occurrence Matrix (GLCM).The study utilized brain images along with neuropsychological assessments, physical and neurological examinations, cognitive assessments, patient medical history, and baseline diagnosis and symptoms.

In [25] the authors suggested a technique to enhance feature detection models for Alzheimer's disease (AD) classification. They utilized AdaBoost and a weighted support vector machine (WSVM) that was tuned with particle swarm optimization (PSO) for feature selection. The proposed method effectively handles large, sparse datasets for brain image classification. The dataset included 198 AD cases and 229 normal controls (NC) for training and testing. The results showed a promising classification accuracy of 93%.

After a thorough review of the existing methodologies, we identified several research gaps. Some approaches rely on complex multiple deep learning techniques to extract features from MRI scan images, while others employ traditional machine learning techniques. However, neither approach has consistently yielded superior results. Additionally, some methods achieve better outcomes with small, imbalanced datasets, while hybrid methods using fusion traditional ML techniques show potential for improvement. Therefore, there is a need for a simpler yet more robust model to assist DSS (decision support system) in the accurate early-stage diagnosis of Alzheimer's disease (AD).

The primary objective of this research is to identify AD at an early stage using MRI images with a straightforward and robust model. We separated the feature extraction and classification tasks to design a hybrid model that leverages both deep learning and traditional machine learning techniques. Specifically, we used a Convolutional Neural Network (CNN) for feature extraction and a Support Vector Machine (SVM) for classification to improve the accuracy of Alzheimer's diagnosis. The CNN model extracted 512 features from high-quality MRI images, which the SVM then used to construct a hyperplane capable of classifying the images as AD or NC. This approach distinguishes itself from previous research by utilizing a balanced dataset with a sufficient number of images, all scaled to a resolution of 128 x 128. Furthermore, we ensure a balanced dataset using data augmentation techniques specifically tailored for binary classification.

Table I displays the results of our comparison study, which we conducted on the same sMRI dataset using the most recent techniques. The proposed model demonstrated superior performance in terms of accuracy and F1-score. The main contributions of this research are summarized below:

- Applying image preprocessing techniques to reduce noise and ensure label-wise balance in the dataset is crucial, particularly for binary classification.
- Extracting crucial features from MRI scan images using a fine-tuned CNN model.
- Predicting Alzheimer's disease using an SVM binary classifier with a linear kernel.
- Attaining a notable 99.60% accuracy in AD prediction.

TABLE I. A COMPARISON BETWEEN THE PROPOSED METHODOLOGY AND STATE-OF-THE-ART TECHNIQUES

Author Name	Model	Dataset	Accuracy (%)
(El-Assy et al., 2024)[14]	CNN	ADNI Dataset	99.57
(Saleh et al., 2023)[15]	DenseNet(feature extraction & Classification)	Kaggle Dataset	96.5
(De Mendonça and Ferrari, 2023) [18]	Support Vector Machine(feature extraction & Classification)	ADNI Dataset	92
(Kongala et al., 2023)[17]	Support Vector Machine(feature extraction & Classification)	ADNI Dataset	---
(Prajapati and Kwon, 2022)[16]	CNN(feature extraction & Classification)	ADNI Dataset	87.50
(Elahifasaee, 2022)[25]	AdaBoost and PSO(Feature extraction), SVM Algorithm for Classification	ADNI Dataset	93
(K et al., 2021)[21]	Support Vector Machine(feature extraction & Classification)	Kaggle Dataset	84

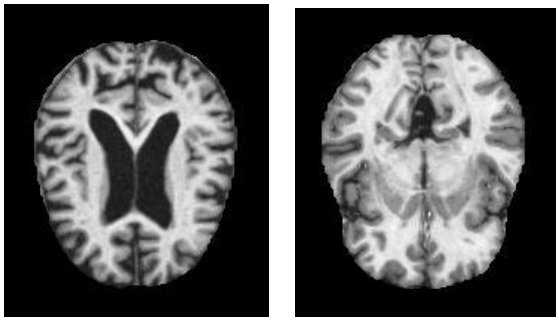
(Dwivedi et al., 2021)[23]	CNN for feature extraction , SVM Algorithm for Classification	ADNI Dataset	91.85
(Nanni et al.,2020)[19]	Fusion of Machine Learning algorithms and SVM Algorithm for Binary Classification	ADNI Dataset	93.30
(Lodha et al., 2018)[20]	Support Vector Machine(feature extraction & Classification)	ADNI Dataset	97.56
(Shen et al., 2018)[22]	Support Vector Machine(feature extraction & Classification)	ADNI Dataset	92.3
(NP and Varghese, 2018)[24]	Support Vector Machine(feature extraction & Classification)	ADNI Dataset	---
Proposed Model	Hybrid(CNN and SVM)	Kaggle Dataset	99.60

The paper divides the rest into sections. Section III provides a comprehensive overview of the MRI dataset and the architecture of the proposed methodology. Section IV describes results of proposed system, Section V discusses the results, while the final Section VI provides the concluding remarks.

III. Dataset Description and Proposed Methodology

A. Dataset Description

The dataset utilized in this research is obtained from Kaggle [26]. The data is gathered from diverse websites, with each label verified. The dataset consists of four types of images, 2,560 images of non-demented cases, 1,792 images of very mild demented cases, 717 images of mild demented cases, and 52 images of moderately demented cases. The dimensions of each MRI image are 176 × 208 and they are in the .jpg format. The objective of this research is to find people with or without dementia. The dataset is more suitable for binary classification rather than classification of multistage Alzheimer's disease due to the significant imbalance among the four labels, however, merging the labels into two better balance between the two classes. Fig. 1 depicts changes in sMRI images of the brain structure of subjects with and without dementia.

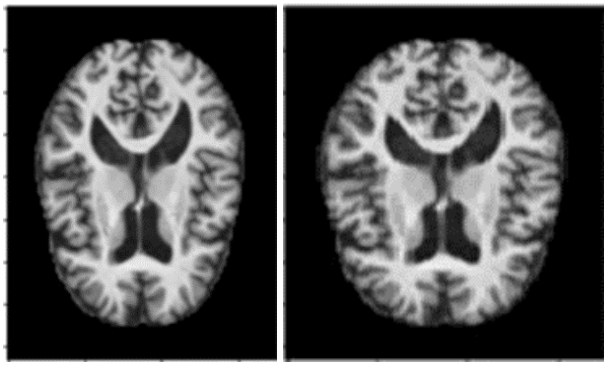


(a) (b)

Fig. 1. Shows (a) AD MRI brain scan image (b) Normal MRI brain Scan image.

B. Preprocessing Mri Images

High-resolution images have a larger number of pixels, which requires more computation and more main memory to save and process pixel values. The computational cost of the model can be greatly reduced by reducing the size of the images, leading to enhanced efficiency and quicker training and evaluation. Following the normalization process, the grayscale images undergo resizing to dimensions of 128x128 pixels, as depicted in Fig. 2.



(i) (ii)

Fig. 2. (i) Original image (176x208) & (ii) Resized image (128x128) Illustrates original and resized image.

C. Horizontal Flipping

In the realm of medical research, specifically in the domain of neuroimaging, the biggest problem is gathering an adequate quantity of images because of privacy concerns. Moreover, an insufficient and unbalanced dataset can cause overfitting problems, affecting the model's performance. In order to address these problems, data augmentation techniques are implemented on the original dataset. We experimented with all data augmentation techniques, such as adjusting brightness, rotation, zooming, etc., but we found that they did not improve the performance of our model. Therefore, we concluded to apply only the horizontal flipping technique [27] to enhance the size of the dataset as shown in Fig. 3. Post-processing, we ended with a balanced dataset consisting of two folders, "demented" and "non-demented," with 3200 images in each shown in Table II. We executed all preprocessing procedures using Python and built ML models by utilizing Keras libraries and Scikit-learn.

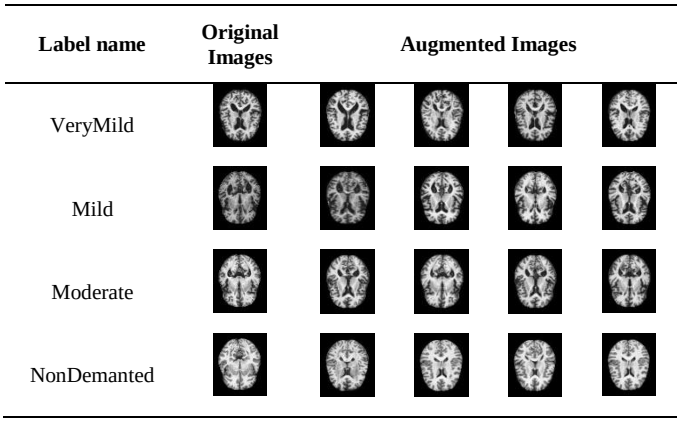


Fig. 3. Shows mirror images generated with horizontal flipping.

TABLE II. MRI DATASET AFTER DATA AUGMENTATION

Label Name	Number of Images
Non Demanted	3200
Demanted	3200
Total	6400

D. Feature Extraction

In the healthcare domain, where diagnosis involves image processing, CNN is highly effective in identifying patterns for classification [28] due to its ability to learn and extract relevant features directly from raw input data [29]. The proposed CNN model is structured with four convolutional blocks, each comprising two consecutive convolution operations followed by a pooling layer as shown in Fig. 4. The input layer accepts images of size 128x128 pixels with three color channels (RGB). The first convolutional block utilizes 64 filters and employs ReLU activation, Batch Normalization, MaxPooling, and Dropout to reduce spatial dimensions and prevent overfitting. This block outputs data in the shape of (64, 64, 64). The second block doubles the number of filters to 128 and follows a similar structure, outputting (32, 32, 128).

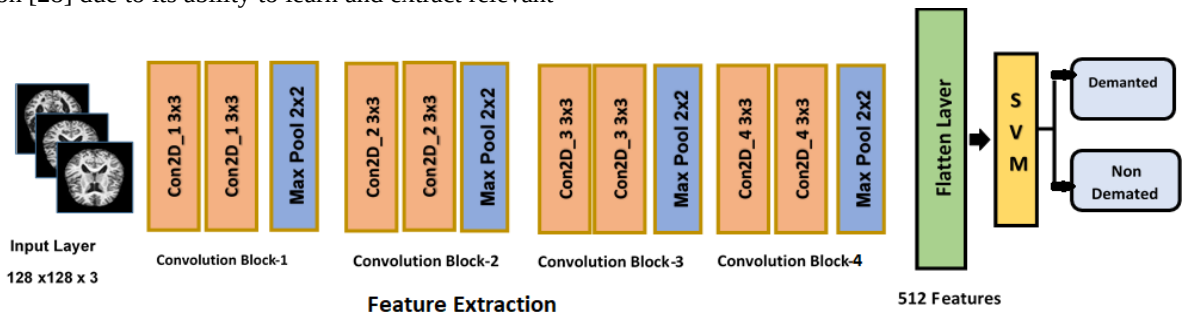


Fig. 4. Proposed model architecture.

The third convolutional block increases the filters to 256, maintaining the structure of convolutional layers, Batch Normalization, Max Pooling, and Dropout, resulting in an output shape of (16, 16, 256). The fourth block further increases the filters to 512 and follows the same pattern, producing an output shape of (8, 8, 512). Each block's design, with its combination of convolutional, normalization, pooling, and dropout layers, facilitates hierarchical feature extraction. This structure progressively enhances the model's ability to classify medical images by capturing increasingly complex patterns and features from the input data. The configuration of proposed CNN shown in Table III. The extracted features are given as input to classification algorithm for prediction.

TABLE III. PROPOSED CNN CONFIGURATION

Layer	Activation Function	Output Shape	Pool Size
2D Convolution Layer	ReLu	(128,128, 64)	
2D Max Pooling Layer	-	(64,64,64)	(2,2)
2D Convolution layer	ReLu	(64,64,128)	
2D Max Pooling Layer	-	(32, 32, 128)	(2,2)
2D Convolution Layer	ReLu	(32, 32, 256)	
2D Max Pooling Layer	-	(16, 16, 256)	(2,2)
2D Convolution Layer	ReLu	(16, 16, 512)	
2D Max Pooling Layer	-	(8, 8, 512)	(2,2)
Flatten Layer		(6400, 32768)	

Here, “Precn” – Precision, “Recl”- Recall,”Spcty”- Specificity, “Accry”- Accuracy,“F1S”- F1 score,”RF”- Random Forest,”LR”- Logistic Regression,DT- Decision Tree,NB- Naive Bayes Classifier,XGB- XGBoost,PM- Proposed Model.

E. Classification Algorithm

Support Vector Machines (SVM) [30] are a strong choice for binary classification tasks. SVM is well-suited for tasks where the objective is to distinguish between two classes, such as distinguishing between diseased and non-diseased patients in medical imaging. As shown in Fig. 5, the main goal is to create a hyperplane that effectively separates the input data points into two output classes. Eq. (1) illustrates the decision function is used to classify new data points.

$$f(x) = w \cdot x + b \quad (1)$$

Where,

The vector w corresponds to the weights, b to the bias, and x to the input features. We choose the hyperplane that gives the maximum distance between the two closest data points in each group. Eq. (2) represents the training dataset.

$$(x_1, y_1) \ldots\ldots, (x_{n,} y_n), x_i \in R_d \; and \; y_i \in (-1, +1) \quad (2)$$

Where

xi-feature vector and yi-output, Rd-set of feature vectors

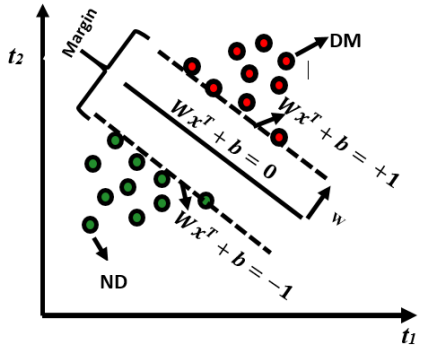


Fig. 5. Illustrates a linear SVM model with two classes.

For each element in the training dataset, the parameters 'b' and 'w' must satisfy the inequality specified in Eq. (3) and (4).

$$if\ \ f(x) > 0 \quad then \quad y_i = +1 \quad (Demanted) \tag{3}$$

$$if\ f(x) < 0 \ \ then \ \ y_i = -1 \quad (Non\ Demented) \tag{4}$$

Eq. (5) defines a function to classify a new data object (x).

$$f(x) = sign(w \cdot x + b) \tag{5}$$

Where sign (.) returns +1 or -1, indicating the class of the data point.

In this study, we evaluated various classification algorithms using extracted features. The Proposed model outperformed all other models across all metrics, showcasing its excellent ability in binary classification tasks. Proposed Model achieved the highest precision (99.35), recall (99.83), specificity (99.40), accuracy (99.60), and F1-score (99.58). Logistic Regression (LR) also performed well, though slightly lower than Proposed Model in specificity and recall. Random Forest (RF) and XGBoost (XGB) showed moderate results, while Naive Bayes (NB) and Decision Tree (DT) significantly lagged in most metrics, especially recall and F1-score. Overall, the Proposed Model demonstrates outstanding performance in binary classification shown in Table IV.

TABLE IV. COMPARING PROPOSED SVM CLASSIFIER WITH OTHER CLASSIFICATION ALGORITHMS

Model	Precn	Recl	Spcty	Accry	F1S
RF	87.64	89.34	86.29	87.89	88.48
LR	99.10	99.10	99.02	99.06	99.09
DT	76.47	76.01	74.55	75.31	76.23
NB	74.38	49.62	81.40	64.84	59.52
XGB	81.73	79.16	80.75	79.92	80.42
PM	99.35	99.83	99.40	99.60	99.58

F. SVM - Kernel Functions

A kernel function can boost the performance of a Support Vector Machine as shown in (6), it improves the performance of SVM by generating non-linear models in higher dimensions. It transforms complex problems into simpler ones by converting non-linear problems into linear ones. In a multi-dimensional space, this would necessitate sophisticated computations, but in this case, it speeds up the calculations.

$$K(t_1, t_2) = < f(t_1), f(t_2) > \tag{6}$$

K-kernel function, t1 and t2 are inputs that have M dimensions. Function f maps the input features from a space with M dimensions to a space with N dimensions. (t1, t2) is the dot product of the inputs.

We experimented with SVM with different kernels to achieve the best performance. The linear kernel performed exceptionally well, attaining the greatest accuracy. It surpassed other kernel functions. Conversely, The Polynomial and RBF kernels [31] are more appropriate for nonlinear data, although they exhibit somewhat poorer accuracy in comparison to the linear kernel. The sigmoid kernel performed poorly, indicating it is not suitable for this dataset. Fig. 6 shows accuracy comparison of various kernels.

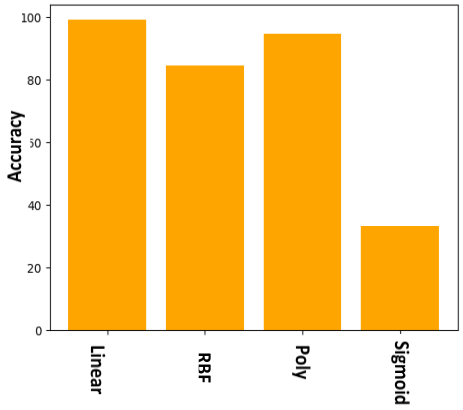


Fig. 6. Comparing the accuracy of different kernels using SVM.

G. Proposed Methodology Architecture

The key steps involved in the proposed methodology are depicted in Fig. 7. First, the data undergoes pre-processing to resize the images from 170x208 to 128x128 pixels. Data augmentation is then applied to increase the dataset size from 5,121 to 6,400 images, ensuring a balanced dataset. Key feature maps are extracted from each image using a CNN model, resulting in 512 features per image, yielding a resultant shape of (6400, 8, 8, 512). The dataset is then split into two sets: 80% for model training and 20% for testing. The SVM algorithm undergoes rigorous training to discover patterns in the data. Next, we perform model validation using a separate testing dataset. We conduct performance evaluation using widely recognized performance indicators like accuracy and F1 score.

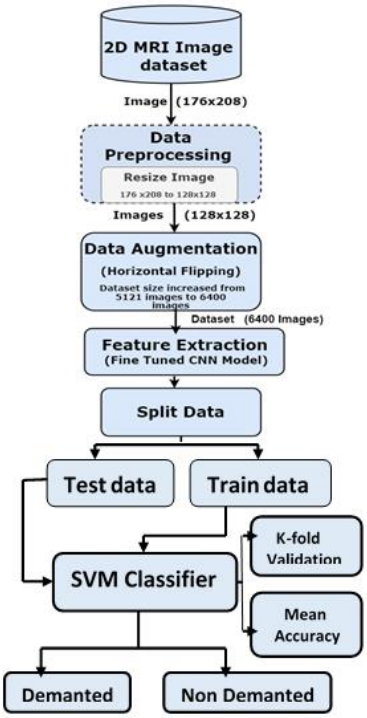


Fig. 7. The proposed methodology flowchart.

H. K-Fold Cross Validation

In order to mitigate the issue of overfitting, the model is tested with the cross-validation technique. This method is essential for evaluating the ability of a model by dividing the dataset into ‘t’ folds of equal size. In every iteration, the model trained on 't-1' folds, leaving one fold for the purpose of testing. This iterative process exposes the model to diverse data subsets, thereby mitigating overfitting. The StratifiedKFold cross-validation ensures consistent class distribution across folds. In this study, as illustrated in Table V, we used 15-fold cross-validation to demonstrate the model's consistent and robust performance across all folds, consistently observing high accuracy. This consistency indicates that the model can generalize to new data.

TABLE V. SHOWS 15-FOLD CROSS-VALIDATION RESULTS OF PROPOSED MODEL

Fold	AC(%)	AVG
1	99.63	
2	99.56	
3	99.73	
4	99.48	
5	99.63	
6	99.58	
7	99.58	
8	99.68	99.58
9	99.44	
10	99.68	
11	99.55	
12	99.30	
13	99.80	
14	99.66	
15	99.50	

Here, AC(%)-Accuracy, AVG-Average

IV. RESULTS

This section discusses model performance metrics, particularly those related to the health care domain, such as sensitivity and specificity. Sensitivity and specificity play crucial roles in assessing the model's accuracy in predicting positive and negative cases. Sensitivity (recall) measures the model's ability to correctly identify positive cases (Demanted) among all actual positive cases, while specificity gauges its accuracy in identifying negative cases (Non-Demanted).

The proposed model attains sensitivity of 99.83%, which implies its rare omission of AD cases and its accurate detection of positive cases, thereby potentially extending lifespan through early detection and appropriate treatment. Meanwhile, the model demonstrates a specificity of 99.40%, signifying high accuracy in identifying individuals without AD. Sensitivity helps us with early detection and treatment, while specificity provides a ratio of false alarms, reducing unnecessary interventions and costs.

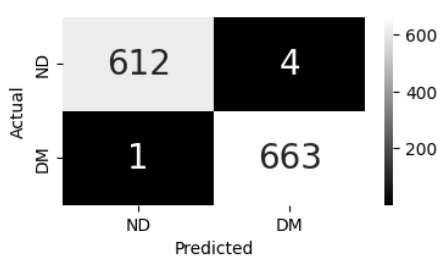
Precision is a quantitative measure that determines the ratio of correctly predicted positive cases to the total number of predicted positive cases. The model achieved a precision of 99.35%, indicating a minimal occurrence of incorrect positive predictions. The F1-score offers a comprehensive assessment of the model's performance by incorporating both precision and recall in its calculation. The model's F1-score of 99.58 indicates an excellent a balance between precision and recall, which is essential for accurate healthcare diagnosis.

V. DISCUSSION

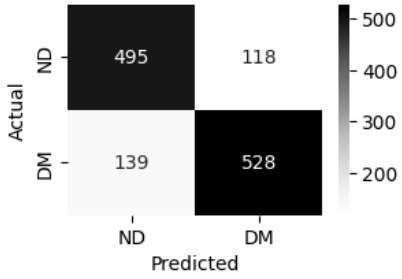
The proposed hybrid model, which uses fine-tuned CNN to pull out features from MRI images and SVM with a linear kernel for classification, achieved impressive metrics: 99.60% accuracy, 99.83% sensitivity, 99.40% specificity, and a 99.67 F1-score in detecting dementia patients. The above metrics collectively show how well the model is able to recognise AD cases while limiting false alarms, indicating its potential usefulness in clinical practice. Fig. 8 shows a comparison between our model's confusion matrix and the confusion matrix of other competitive classifiers. In the medical field, incorrect predictions can have serious consequences. The model we propose demonstrates a minimal number of such errors.

In this study, we applied only horizontal flipping for data augmentation to maintain the image's orientation. Other studies in the literature have employed various data augmentation techniques, which can cause their models to deviate in extracting key features. After several trial-and-error attempts, we determined that 128x128 was the optimal image size, whereas other studies used different sizes such as 64x64 or 256x256, which may result in the loss of important features or introduce noise into the image.

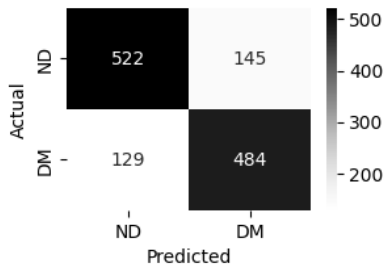
The proposed model successfully extracted crucial key features from well-preprocessed MRI images using a fine-tuned CNN model, outperforming state-of-the-art techniques mentioned in the literature. The proposed model produced superior accuracy compared to recent state-of-the-art techniques as displayed in Fig. 9. This model can serve as an automated computer-based decision support system, using structural MRI images to provide pattern-based decisions that assist manual systems in accurately identifying AD quickly, enabling timely medication to increase the lifespan of affected individuals.



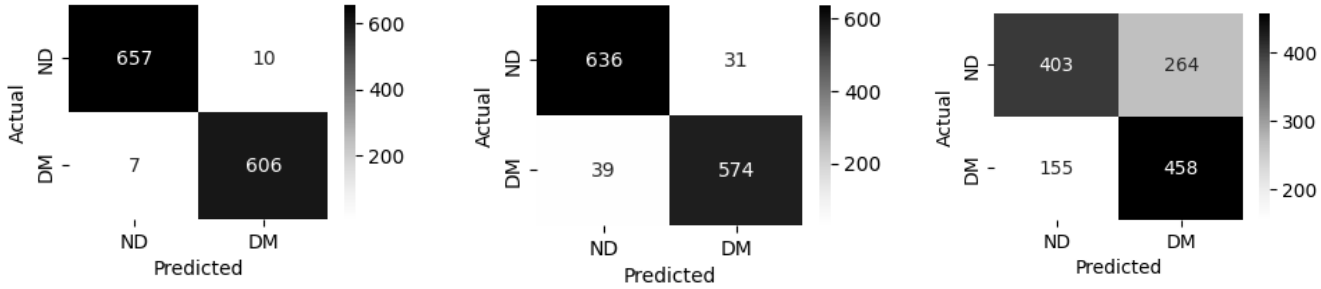
(i) Proposed



(ii) XG boost



(iii) Decision tree



(iv) Logistic regression (v) Random gorest (vi) Naïve bayes

Fig. 8. From (i) to (vi) demonstrates the CM for the proposed method and other classification algorithms. Here DM-"Demanted", ND-"Non_Demanted",CM-confusion matrix.

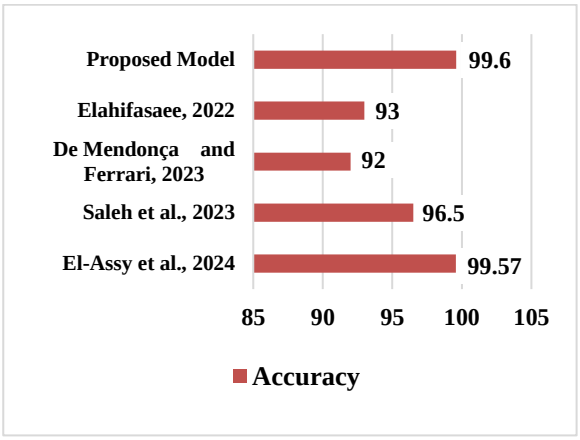


Fig. 9. Shows comparison of the proposed model's accuracy with state-of-the-art techniques.

VI. Conclusion

AD is a serious health concern, and early detection and treatment are crucial. The hybrid model is proposed to tackle this challenge. The model has demonstrated superior performance compared to contemporary techniques in various classification performance measures, indicating its potential as a robust tool for early diagnosis. Although the results are promising, there remains potential for additional enhancement and refinement. In this research, we used a dataset with sMRI images, integrating additional imaging modalities, such as functional MRI (fMRI), Fluid-Attenuated Inversion Recovery (FLAIR), Arterial Spin Labeling (ASL), Susceptibility Weighted Imaging (SWI), and Positron Emission Tomography (PET), could provide more comprehensive and precise insights into the disease's progression and characteristics. These modalities offer different perspectives on brain structure and function, which could enhance the accuracy and reliability of the diagnostic model. Furthermore, this study focused only on binary classification. Advanced deep learning techniques can extend this work to multistage classification using various modalities. This can lead to accurate, timely diagnosis, ultimately improving patient care and outcomes in the battle against Alzheimer's disease.

References

[1] Arafa, D.A., Moustafa, H.E.-D., Ali, H.A., Ali-Eldin, A.M.T., Saraya, S.F., 2023. A deep learning framework for early diagnosis of Alzheimer's disease on MRI images. Multimedia Tools and Applications 83, 3767–3799.https://doi.org/10.1007/s11042-023-15738-7.

[2] Kong, Z., Zhang, M., Zhu, W., Yi, Y., Wang, T., Zhang, B., 2022. Multi-modal data Alzheimer's disease detection based on 3D convolution. Biomedical Signal Processing and Control 75, 103565.https://doi.org/10.1016/j.bspc.2022.103565.

[3] Mehmood, A., Maqsood, M., Bashir, M., Shuyuan, Y., 2020. A Deep Siamese Convolution Neural Network for Multi-Class Classification of Alzheimer Disease. Brain Sciences 10, 84. https://doi.org/10.3390/brainsci10020084.

[4] 2023 Alzheimer's disease facts and figures, 2023. Alzheimer's & Dementia 19, 1598–1695. https://doi.org/10.1002/alz.13016.

[5] Hebert, L.E., Weuve, J., Scherr, P.A., Evans, D.A., 2013. Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. Neurology 80, 1778–1783. https://doi.org/10.1212/wnl.0b013e31828726f5.

[6] Arevalo-Rodriguez, I., Smailagic, N., Roqué-Figuls, M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., Pedraza, O.L., Bonfill Cosp, X., Cullum, S., 2021. Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). Cochrane Database of Systematic Reviews 2021. https://doi.org/10.1002/14651858.cd010783.pub3.

[7] Harrell, L.E., Marson, D., Chatterjee, A., Parrish, J.A., 2000. The Severe Mini-Mental State Examination: A New Neuropsychologic Instrument for the Bedside Assessment of Severely Impaired Patients With Alzheimer Disease. Alzheimer Disease and Associated Disorders 14, 168–175. https://doi.org/10.1097/00002093-200007000-00008.

[8] Pangman, V.C., Sloan, J., Guse, L., 2000. An examination of psychometric properties of the Mini-Mental State Examination and the Standardized Mini-Mental State Examination: Implications for clinical practice. Applied Nursing Research 13, 209–213. https://doi.org/10.1053/apnr.2000.9231.

[9] Leifer, B.P., 2003. Early Diagnosis of Alzheimer's Disease: Clinical and Economic Benefits. Journal of the American Geriatrics Society 51. https://doi.org/10.1046/j.1532-5415.5153.x.

[10] Kavitha, C., Mani, V., Srividhya, S.R., Khalaf, O.I., Tavera Romero, C.A., 2022. Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models. Frontiers in Public Health 10. https://doi.org/10.3389/fpubh.2022.853294.

[11] Tufail, A.B., Ma, Y.-K., Zhang, Q.-N., 2020. Binary Classification of Alzheimer's Disease Using sMRI Imaging Modality and Deep Learning. Journal of Digital Imaging 33, 1073–1090. https://doi.org/10.1007/s10278-019-00265-5.

[12] EL-Geneedy, M., Moustafa, H.E.-D., Khalifa, F., Khater, H., AbdElhalim, E., 2023. An MRI-based deep learning approach for accurate detection of Alzheimer's disease. Alexandria Engineering Journal 63, 211–221. https://doi.org/10.1016/j.aej.2022.07.062.

[13] Herrera, L.J., Rojas, I., Pomares, H., Guillen, A., Valenzuela, O., Banos, O., 2013. Classification of MRI Images for Alzheimer's disease Detection. 2013 International Conference on Social Computing. https://doi.org/10.1109/socialcom.2013.127.

[14] El-Assy, A.M., Amer, H.M., Ibrahim, H.M., Mohamed, M.A., 2024. A novel CNN architecture for accurate early detection and classification of Alzheimer's disease using MRI data. Scientific Reports 14. https://doi.org/10.1038/s41598-024-53733-6.

[15] Saleh, A.W., Gupta, G., Khan, S.B., Alkhaldi, N.A., Verma, A., 2023. An Alzheimer's disease classification model using transfer learning Densenet with embedded healthcare decision support system. Decision Analytics Journal 9, 100348. https://doi.org/10.1016/j.dajour.2023.100348.

[16] Prajapati, R., Kwon, G.-R., 2022. A Binary Classifier Using Fully Connected Neural Network for Alzheimer's disease Classification. Journal of Multimedia Information System 9, 21–32. https://doi.org/10.33851/jmis.2022.9.1.21.

[17] Rao, K.N., Gandhi, B.R., Rao, M.V., Javvadi, S., Vellela, S.S., Khader Basha, S., 2023. Prediction and Classification of Alzheimer's disease using Machine Learning Techniques in 3D MR Images. 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS). https://doi.org/10.1109/icscss57650.2023.10169550.

[18] de Mendonça, L.J.C., Ferrari, R.J., 2023. Alzheimer's disease classification based on graph kernel SVMs constructed with 3D texture features extracted from MR images. Expert Systems with Applications 211, 118633. https://doi.org/10.1016/j.eswa.2022.118633.

[19] Nanni, L., Interlenghi, M., Brahnam, S., Salvatore, C., Papa, S., Nemni, R., Castiglioni, I., 2020. Comparison of Transfer Learning and Conventional Machine Learning Applied to Structural Brain MRI for the Early Diagnosis and Prognosis of Alzheimer's Disease. Frontiers in Neurology 11. https://doi.org/10.3389/fneur.2020.576194.

[20] Lodha, P., Talele, A., Degaonkar, K., 2018. Diagnosis of Alzheimer's Disease Using Machine Learning. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). https://doi.org/10.1109/iccubea.2018.8697386.

[21] K, U.R., S, Sharvari.S., G, Umesh.M., C, Vinay.B., 2021. Binary Classification of Alzheimer's disease using MRI images and Support Vector Machine. 2021 IEEE Mysore Sub Section International Conference (MysuruCon). https://doi.org/10.1109/mysurucon52639.2021.9641661.

[22] Shen, T., Jiang, J., Li, Y., Wu, P., Zuo, C., Yan, Z., 2018. Decision Supporting Model for One-year Conversion Probability from MCI to AD using CNN and SVM. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). https://doi.org/10.1109/embc.2018.8512398.

[23] Dwivedi, S., Goel, T., Sharma, R., Murugan, R., 2021. Structural MRI based Alzheimer's Disease prognosis using 3D Convolutional Neural Network and Support Vector Machine. 2021 Advanced Communication Technologies and Signal Processing (ACTS). https://doi.org/10.1109/acts53447.2021.9708107.

[24] Thulasi N.P., K., Varghese, D., 2018. A Novel Approach for Diagnosing Alzheimer's Disease Using SVM. 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI). https://doi.org/10.1109/icoei.2018.8553789.

[25] Elahifasaee, F., 2022. Optimized SVM using AdaBoost and PSO to Classify Brain Images of MR. 2022 International Conference on Machine Vision and Image Processing (MVIP). https://doi.org/10.1109/mvip53647.2022.9738549.

[26] Alzheimer's Dataset (4 class of Images) [WWW Document], 2019. . Kaggle. URL https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images, Accessed on

[27] Mehmood, A., Maqsood, M., Bashir, M., Shuyuan, Y., 2020. A Deep Siamese Convolution Neural Network for Multi-Class Classification of Alzheimer Disease. Brain Sciences 10, 84. https://doi.org/10.3390/brainsci10020084.

[28] Mehmood A, Abugabah A, AlZubi AA, Sanzogni L (2022) Early Diagnosis of Alzheimer's Disease Based on Convolutional Neural Networks. Comput Syst Sci Eng 43(1):305–315. https://doi.org/10.32604/csse.2022.018520.

[29] "What is Feature Extraction? Feature Extraction in Image Processing | Great Learning." (n.d.) https:// www.mygreatlearning.com/blog/feature-extraction-in-image-processing/. Accessed 27 Feb 2024.

[30] Applications of Support Vector Machine (SVM) Learning in Cancer Genomics, 2018. Cancer Genomics & Proteomics 15. https://doi.org/10.21873/cgp.20063.

[31] Neffati, S., Taouali, O., 2017. An MR brain images classification technique via the Gaussian radial basis kernel and SVM. 2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA).
https://doi.org/10.1109/sta.2017.8314948.