Identifying Alzheimer Disease Dementia Levels Using Machine Learning Methods

Md Gulzar Hussain, Ye Shiren*

School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, Jiangsu, China

*yes@cczu.edu.cn

Abstract

Dementia, a prevalent neurodegenerative condition, is a significant aspect of Alzheimer's disease (AD), causing progressive impairment in daily functioning. Accurate and timely classification of AD stages is crucial for effective management. Machine learning and deep learning models have shown promise in this domain. In this study, we proposed an approach that utilizes support vector machine (SVM), random forest (RF), and convolutional neural network (CNN) algorithms to classify the four stages of dementia. We augmented these algorithms with watershed segmentation to extract meaningful features from MRI images. Notably, our findings demonstrate that SVM with watershed features achieves an impressive accuracy of 96.25%, outperforming other classification methods. The evaluation of our method was performed using the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, and the inclusion of watershed segmentation was found to significantly enhance the performance of the models. This study highlights the potential of incorporating watershed segmentation into machine learning-based AD classification systems, offering a promising avenue for accurate diagnosis and treatment planning.

Keywords: Alzheimer's Disease, Dementia, CNN, Random Forest, SVM, MRI image, Computer-Aided Diagnostic.

Introduction:

The brain, being one of the highly vital and complicated organs in a human body, is responsible for a wide range of essential functions such as intellectual invention, problem-solving, thought, judgment, creativity, and memories. However, when an individual develops dementia, their cognitive abilities become impaired. Most of all dementia sufferers have Alzheimer's disease (AD), making it the most prevalent reason of dementia. AD gradually destroys brain cells, resulting in disconnection from surroundings, loss of recognition of loved ones, inability to recall childhood memories, familiar faces, and even basic procedures. Alarmingly, the number of individuals affected by AD is projected to increase significantly in the coming decades. It is estimated that by 2050, there will be over 150 million AD sufferers worldwide, up from 50 million in 2020. This rapid increase in AD cases poses a significant burden on patients, families, and healthcare systems, unless there are advancements in prevention measures utilizing modern medical technologies. The need for research and innovation in AD prevention and treatment is crucial to address this growing global health challenge. Efforts to develop effective medications, leveraging contemporary medical technologies, are essential to combat the rising prevalence of AD and alleviate the impact it has on individuals, families, and healthcare systems worldwide.

According to the estimates, there are 32, 69, and 315 million people worldwide who have prodromal AD, AD dementia, and preclinical AD, respectively. Combined, they made up 416 million people on the AD spectrum or 22% of all those 50 and older[1]. The burden of dementia grew globally between 1990 and 2019, increasing by 147.95% and 160.84%, correspondingly, in frequency and prevalence. In order to deal with the rising incidence of dementia, we should pay focus to the elderly society, prioritize programs that focus on dementia health conditions, and create plans of action[2]. Considering millions of people struggling with dementia globally, the prevalence of dementia has a significant necessity. The indications of AD and those of vascular dementia (VD) or ordinary aging intersect, making the identification of AD problematic[3]. Premature and precise identification of Alzheimer's disease (AD) is crucial for patient care, prevention, and treatment. Several research initiatives aim to utilize imaging of brain techniques, including the magnetic resonance imaging (MRI), for identifying AD. Neuroimaging methods like structural MRI (sMRI), amyloid PET, functional MRI (fMRI), fluorodeoxyglucose positron emission tomography (FDG-PET) imaging, and diffusion tensor imaging (DTI) have shown promise in evaluating abnormal brain changes associated with AD. By leveraging these advanced neuroimaging methods, researchers and clinicians can monitor the progression of AD, enable earlier and more accurate diagnosis, and facilitate appropriate patient care, therapy, and prevention strategies[4]. fMRI has aided AD investigators in assessing functionally active areas when performing a task to identify AD early in comparison to other neuro-imaging methods.

The hippocampus, a crucial brain region responsible for creating and retaining new memories, is one of the earliest areas to be influenced by Alzheimer's disease. As Alzheimer's progresses, the hippocampus undergoes structural changes, resulting in a decline in the brain's ability to create and retrieve memories. Advanced imaging techniques, such as 3D imaging through Alzheimer's MRI, can provide visual representations of the hippocampus, depicting the number of cells present and its size. By utilizing 3D imaging, researchers and clinicians can gain a clearer understanding of the extent of structural changes in the hippocampus caused by Alzheimer's[5]. Mild cognitive impairment (MCI), which is often considered as an intermediate stage in the middle of normal cognitive aging and Alzheimer's disease, can be detected using MRI imaging.

Modern developments in image processing have helped to improve quickly in developing modern methods. The use of image analysis has grown, particularly in the realm of healthcare. In comparison to the current approaches, deep learning offers more in the way of time and effectiveness. Great advancements have been achieved in the employment of machine learning (ML) and deep learning (DL) methodologies to the determination of AD, especially utilizing magnetic resonance imaging (MRI), which has been facilitated by the availability of open AD-related databases. In contrast to conventional ML techniques, DL has emerged as a paradigm-shifting tool in recent years. Since

DL can train immediately from images using neural networks, this has simplified the procedure without involving specialists in extracting features. This eliminates the need for feature extraction to be carried out physically and separately from the classifiers[6]. While there are many other ways to extract features, deep learning (DL) using multilayered ANNs, also known as deep neural networks (DNNs), provides a potential substitute by making feature extrication automated. Earlier feature extraction has been heavily used in earlier studies to lower the dimensions of actual-world inputs like MRI and expressed gene data[7]. Although this method can provide models with great accuracy, it can also result in models that do not take the fresh correlations found in the raw data into account. As a step of the learning process of the model, innovative methods to the structure of DNNs, including convolutional neural networks (CNNs) and, spiral neural networks (SNNs) enable the dimensions of the training dataset to be decreased, enabling the system to engage cooperatively with the raw input[8]. Convolutional neural networks (CNNs) has recently attained extremely high quality and reliability in image classification[6].

Depending on high-dimensional data obtained from various types of neuroimaging biomarkers, such as MRI, numerous machine learning techniques have been exhibited to assist in the diagnosis of Alzheimer's. Throughout this section, recent works that employed traditional DL and ML methods in AD diagnosis and prediction methods are

addressed. The authors of the work[9] suggested employing a previously trained CNN deep learning system called ResNet-50 as an automated features extraction technique for MRI image-based Alzheimer's disease diagnosis. The effectiveness of a CNN containing traditional SVM, Softmax, and RF was then assessed using several metrics, such as accuracy, and it was found to perform better, with a range of accuracy for classifiers using the MRI ADNI data of 85.7% to 99%. The correctness of CNN is 97.6%, that is 10.96% greater than the Spinl-Net technique, when using the XAI approach utilized in the study[10], local interpretative model agnostic explanation (LIME), for a straightforward human explanation. Support vector classifier offers greater accuracy when assessing gene expression data than other methods. The most crucial genes for a certain AD patient were identified from the gene expression data, and LIME demonstrates how genes were chosen for that patient. The adoption of the weighted map for the 2 young radiologists in the matter of EOAD vs. FTD dramatically enhanced diagnostic performance in article[11]. The increase in diagnosis accuracy was over 10 percent. The study [12]'s goal was to use an enhanced chicken swarm optimization (ICSO) approach to classify MRI pictures as demented (DM) and non-demented (ND). The traditional chickens swarm optimization method has an efficiency of 52.13% using statistical features and 52.99% without, meanwhile, the ICSO without statistical features has accurate results of 86.32%. The approach used in paper[13] can evaluate many classes in a singular installation with fewer

tagged training specimens and little previous domain knowledge than earlier methods by resolving the bottleneck problem. Authors of[14] suggested a depth-wise convolution operation in place of the existing DenseNet-121 architecture's convolution operation to reduce completion time. The model's efficiency was also enhanced by the new framework, on average by 90.22%.

The proposed methodology in article[15] correctly categorized instances of constitutionally normal (CN) and Alzheimer's disease (AD) employing the Alzheimer's disease Neuroimaging Initiative (ADNI) dataset having accuracy scores of 99.6 percent, 99.8 percent, and 97.8 percent, correspondingly. In multi-classification investigations, the proposed architecture achieved a categorization accuracy of 97.5 percent on the ADNI dataset. To categorize the several levels of Alzheimer's Disease using the magnetic resonance imaging (MRI) pictures, VGG16 influenced CNN (VCNN) and Deep Convolution Neural Network (DCNN) architectures had been established in article[16]. The findings of experimentation on an ADNI dataset demonstrate that the proposed approaches achieved outstanding accuracy. To recognize the phases of dementia using MRI, a DEMentia NETwork (DEMNET) is proposed in article[17]. Using the Kaggle dataset, this DEMNET acquires an accuracy score of 95.23 percent, an area under curve (AUC) of 97%, and a Cohen's Kappa score value of 0.93. A deep learning-based algorithm is suggested in the research[4] which can forecast MCI, late MCI (LMCI), early MCI

(EMCI), and AD. In their study for testing, the 138-subject ADNI fMRI dataset was employed. For the AD vs. EMCI, AD vs. LMCI, and EMCI vs. MCI classifying settings, the fine-tuned ResNet-18 network acquired an accuracy score of classification of 99.99%, 99.95%, and 99.95%, correspondingly. In the work[18], the categorization of undertaken EEG signals from FTD and AD patients was examined across six supervised machine-learning approaches. The accuracy ratings for the suggested approach were 86.3% for detection of FTD using random forests and 78.5% for detection of AD using decision tree. The determination proof from a three-year clinical follow up was compared with ML predictions in the paper[3]. The most effective method for differentiating AD from VD was ANFIS, which used a minimal feature pattern to obtain an accuracy of classification greater than 84%. Using the most frequent T1-weighted sequencing magnetization produced fast accumulation with gradient echo (MPRAGE), the presented structure in[19] produced an accuracy of 91.83%. To extract certain useful features from PSD images and carry out the associated two and three-dimensional classification problems, a modified CNN including one processed element of convolution, pooling layer, and Rectified Linear Units (ReLu) is constructed in[20]. Its accuracy obtained in binary classification is 89.8%, while in three classifications, it is 83.3%. In the study[21], research on disease detection over the past five years utilizing transfer learning based on deep learning for only brain MRI scans is conducted. This research suggests that a

composite approach focused on transfer learning provided over an accuracy of 90% in most scenarios with a little training period.

The watershed algorithm using the conception of topography, which is fundamentally a region-growing technique, differs in that it starts development from the local minimum throughout the images. There will be a bunch of pseudo-minima because of the impact of the darker disturbance and darker roughness in the photo, and these pseudo-minimums result in the associated pseudo-water basins in the picture. Hence, the watershed method separates out every pseudo-real minimum's minimum amount into a distinct region. According to the literature, there are many methods for classifying AD employing machine and deep learning. Nevertheless, the multi-label AD classification's significant model parameters and class imbalance remain a problem. Feature extraction by the watershed algorithm is employed and for classification, RF, SVM and CNN model with fewer parameters is developed & compared to address this issue and forecast four levels of dementia that can guide to AD in this research work. The target of this study is to evaluate watershed algorithm-based extraction of features from MRI for the automatic detection and categorization of four dementia levels of Alzheimer's disease depending on the superior execution of support vector machine (SVM), random forest (RF), and convolutional neural network (CNN) approach in a variety of images classification challenges. Alzheimer's disease utilizing MRI images is being

diagnosed using RF, SVM, and CNN classifier models, and the effectiveness of the models was evaluated between them.

Methods

The primary objective of this research is to evaluate and improve the performance of the classifier of MRI images for the classification of dementia levels of Alzheimer's disease using RF, SVM, and CNN. Data gathering, pre-processing, Deep Learning based performance tuning and classifying, and impact assessment are all part of the research technique. The MRI data were obtained from a famous Alzheimer's disease database[22]. Figure 1 depicts the suggested model's process flow chart.
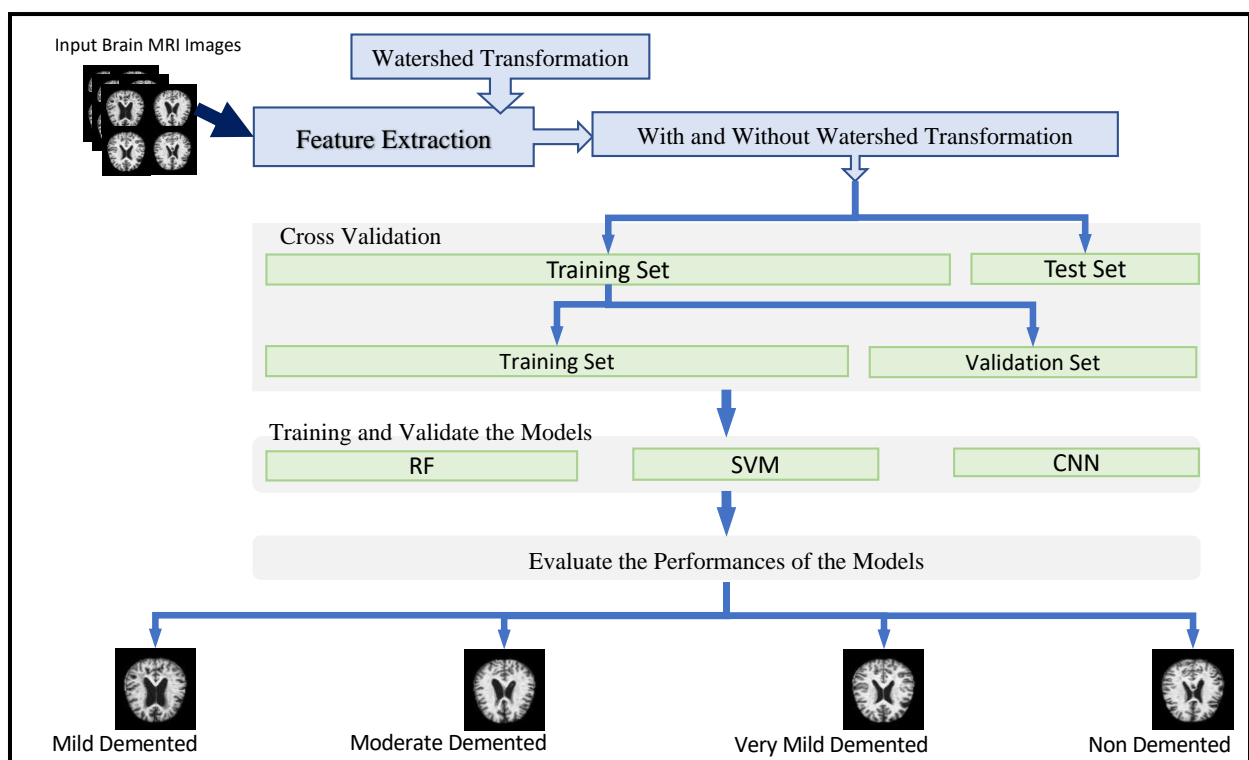

Figure 1 Flow Diagram of the Suggested Method in this Research Work.

Data Collection:

For this study, wez used MRI scans of the brain acquired from the ADNI (Alzheimer's Disease Neuroimaging Initiative) dataset[23] and downloaded from[22]. The dataset is comprised of 6400 pre-processed MRI (Magnetic Resonance Imaging) images of four types. All the pictures have been scaled to 128 x 128 pixels. Sample images of each category are given in Figure 2.
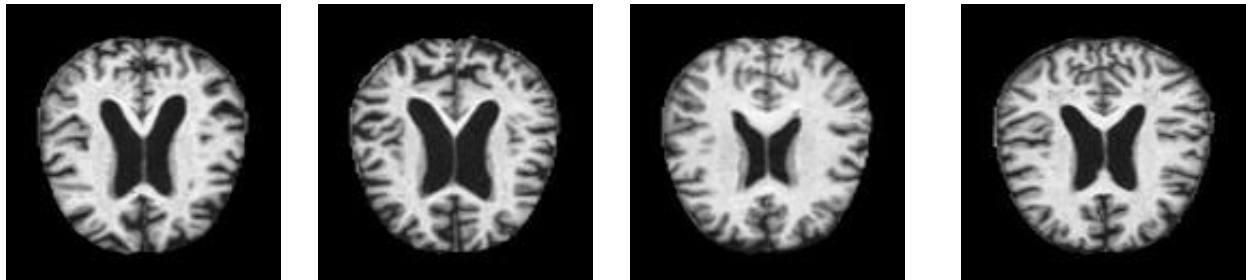
Figure 2 Sample Images of Four Classes in the Dataset.

Image Pre-processing:

Pre-processing of data is one crucial process that prepares the dataset for corresponding classification model training. The resolution of MRI images of the dataset is 128 x 128, with a significant number of characteristics. To process high-resolution images high configured machine is necessary, which is expensive too. One of our targets is to observe the performance of low-resolution images. As a result, to reduce resource utilization for classifying these images, the resolution must be decreased. So, the MRI images are compressed to 32 x 32 to fit in the proposed approach.

Feature Extraction:

Feature extraction is a leap in the dimension retrenchment procedure that splits and alleviates an initial amount of fresh data to more accessible classes. Consequently, processing procedure will be easier. The minimization of data assists in the construction of the model with little machine activity, as well as the efficiency of the training and classification phases in the machine learning model. Due to this reason, researchers and the deep learning communities proposed several methods and techniques to decrease the data dimensionality for performance optimization. The Watershed Algorithm was used in this work as a feature-extraction approach to retrieve the features from the dataset.

Watershed Algorithm:

Watershed algorithm is a well-received segmentation technique[24], which provides numerous benefits for picture segmentation, such as ensuring closed section boundaries and producing reliable findings. The morphological watershed transform is a well-known image segmentation approach that uses mathematical morphology to divide a picture due to irregularities. Watersheds are defined by considering a picture in three dimensions: 2-dimensional co-ordinates against the intensity[25]. The watershed algorithm developed on gathering certain foreground and background information, and then utilizing markers to run the watershed and discover the exact borders. This technique is useful for detecting contacting and overlapped objects in images.

The marker-based watershed transformation technique is an enhancement of the watershed method, which is employed to establish a marker in a picture. The marking can be a line, a spot, or a region; the placement is more significant than the indicated shape. Every tag presents a picture of a finalized separation, and choosing a mark is an important aspect in deciding segmentation. By adjusting the threshold, one can obtain all the low values. With all these lowlands as a marker, the watershed transformation generates the final segmentation image[26]. With the relation between its aims of breaking each small area, and in the middle of every target, the position of neighboring pixels with the same gray level of the formulation adds up to the local highest platform region strength, and noise is removed by filtration and image renovation. The seeding section is the largest target

location. This seeding section is the extracted marker region. Image in grayscale of the destination within the marker contain a significant number of non-seed areas with the noise or areas. The gray threshold is strongly tied to the tag area amount. The image with the highest gray level is from the decision's height. By deleting the marked portion of the regional encounters de-noising impact, the gray value is reduced. A targeted area of the markers is merged by obtaining the maximum gray level. The target particle size was chosen as a threshold image's gray level. If the gray threshold is too significant, all markers under the threshold mark of tiny target regions have been deleted. If the gray threshold is too minimum, extremely tiny particles have also been determined as a target. The watershed segmented picture is created after identifying the reproduced binary image.
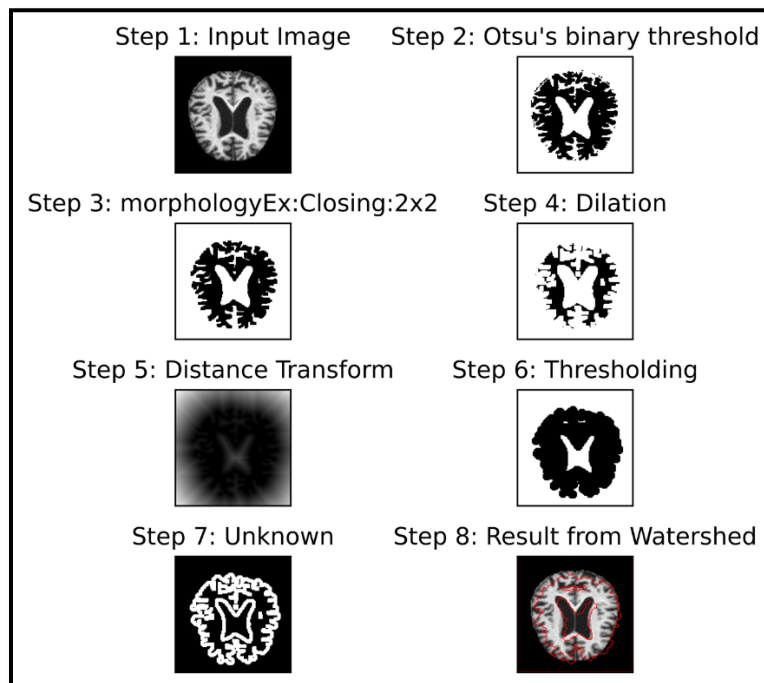

Figure 3 Steps of Marker Based Watershed Algorithm.

Figure 3 is showing the steps of the marker based watershed segmentation technique employed in this experimental work. These steps are expressed in the previous paragraph. For this research work, the generated image showed in step 7 of Figure 3 is used to generate the features of the MRI images.

Classification Models:

This section of the paper describes the machine learning models employed to identify the level of dementia in patients. For training, from ensemble algorithm, Random Forest (RF), from support vector networks, Support Vector Machine Classifier (SVM) and from neural network, convolutional neural network (CNN) approach is employed in this research work. The subsections provide a summary of the RF, SVM, and CNN models used to meet the clinical goal of dementia classification.

Random Forest (RF):

A group or ensemble of classifier and regression Trees is known as a Random Forest. In comparison to other machine learning methods, RF is regarded to be steadier in the existence of abnormalities and higher dimensional parameters spaces because it adheres to specified directives for tree building, self-testing, tree combination, and post-processing. It is also vigorous to overfitting[27]. In this research, we set the tree's maximum depth as 16 with zero random states.

Support Vector Machine Classifier (SVM):

One of the highest popular kernel-based learning methods is the SVM technique, which was first developed by Vapnik and his team in the late 1970s and is notably popular in image classification tasks. SVM is a member of the class of supervised nonparametric algorithms, which are unaffected by the dispersion of the underlying data. That is one of the benefits of using SVMs over other statistical methods. SVM is a linear binary classifier that distinguishes just one border between two classes in its most basic configuration[28]. SVMs utilize a subset of the training specimen that is located nearest to the ideal decision boundary in the feature space, acting as support vectors, to increase the dissociation or border. In this research a polynomial kernel function is used with a degree value of 3.

Convolution Neural Networks (CNN):

In terms of image processing, CNN happens to be one of the finest and most widely used Deep Learning techniques. In this experimental work, the target is to use minimum layers and only three 2D Convolution layers are used. Figure 4 demonstrates the CNN classification structure in this research.
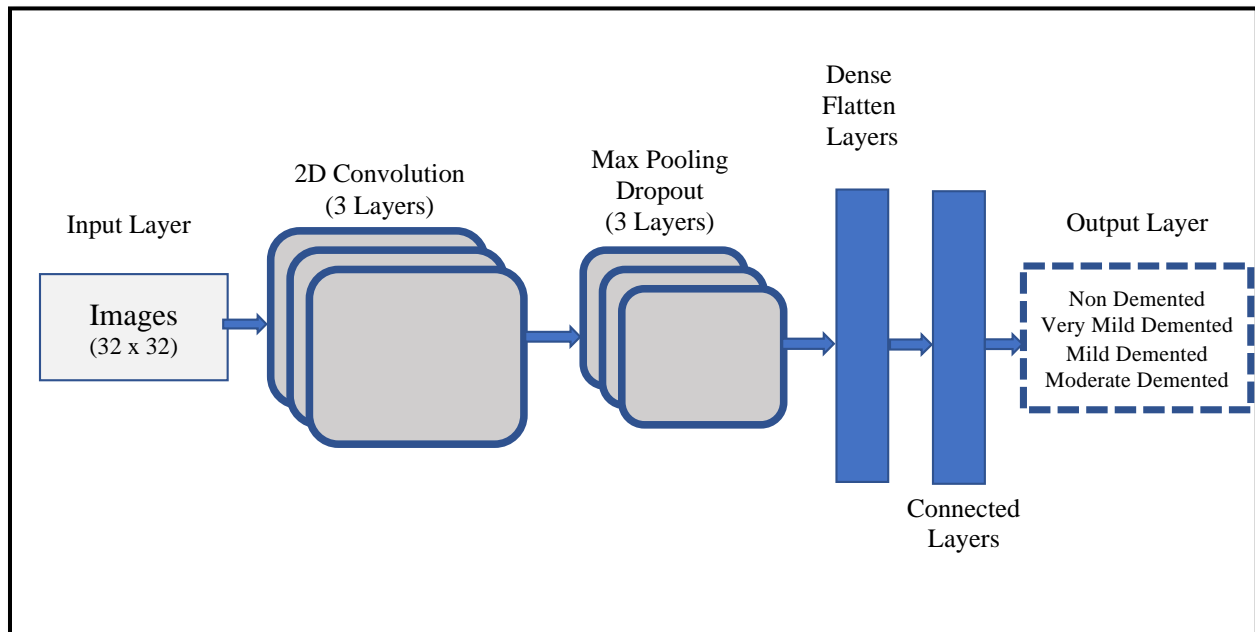
Figure 4 Architecture of CNN Used in this Research Work.

In the proposed CNN model, the count of convolutional layers and completely linked layers are three. The edges and color of the image are mostly captured by the first convolutional layer. As 2D images are used all the convolutional layers are two-dimensional. Explanation of the inputs, outputs, and number of parameters that need to be adjusted for all layers of the model along three convolution layers are given below:

First Layer: Every gray scale image in the database is in 32x32 dimension. The first convolution layer receives this image as input, which is then processed using a single convolution filter. There are 1664 parameters in this layer that must be trained. Since each filter generates an image with a size of 28x28, the outcome of this layer is 28 x 28 x 1.

Second Layer: The second convolutional layer utilizes an input with 64 filters and a size of 28x28x1. There are 73856 parameters throughout this layer that must be taught. As every filter produces an image of size 28x28 and every one of the 64 filters produces 64 pictures of size 28x28, the outcome of this layer is 28 x 28 x 1.

Third Layer: Images with 128 filters are used in the third layer. There are 65664 parameters in this layer that must be taught. As every filter produces an image with size 6x6 and the 128th filter produces 128 pictures of size 6x6, the product of this layer is 6 x 6 x 128.

Output Layer: The last layer utilizes the soft-max activation function, which has four possible outcomes, and is linked. This layer approximates the class values and assigns the four classes after receiving input from a fully linked layer. In this layer, 516 parameters have been modified. This classification indicates the class the subject belongs within.

Results:

This segment of the article discusses the outcomes of the experiment and explains how they were conducted. We trained random forest, support vector machine classifier, and convolutional neural networks to perform four class classifications, such as Moderate Demented (MoD), Mild Demented (MD), Very Mild Demented (VMD), and Non-Demented (ND). The dataset comprises 6400 MRI images in total and is imbalanced. This study examines the efficacy of RF, SVM, and CNN with watershed and without watershed features for classifying dementia levels from MRI images. The dataset was split into 80% and 20% ratios for training and testing for RF and SVM evaluation. On the other hand, it is split into a 70% training, 10% validation, and test 20% split ratio for the CNN evaluation.

Result Based on RF:

Throughout this research, the first step in classifying the four classes with the RF model and to determine the results using training, and test data. We have done it with and without the watershed feature of the image data. This model without watershed achieved an accuracy result of 91.25% and with watershed 89.53%. The confusion matrix of the random forest is stated in the Figure 5 and Figure 6.

Figure 5 is showing the confusion matrix for random forests without watershed and Figure 6 is for those with the watershed. From these matrices, we can observe that though the overall accuracy is higher without watershed than with watershed, for the non-demented class the number of true positives is greater with the watershed. For other classes, the number of true positives is higher without watershed. For the non-demented class number of images was maximum. We can say that, if the dataset were balanced and the other three classes had more images than with watershed the RF could have given higher accuracy.
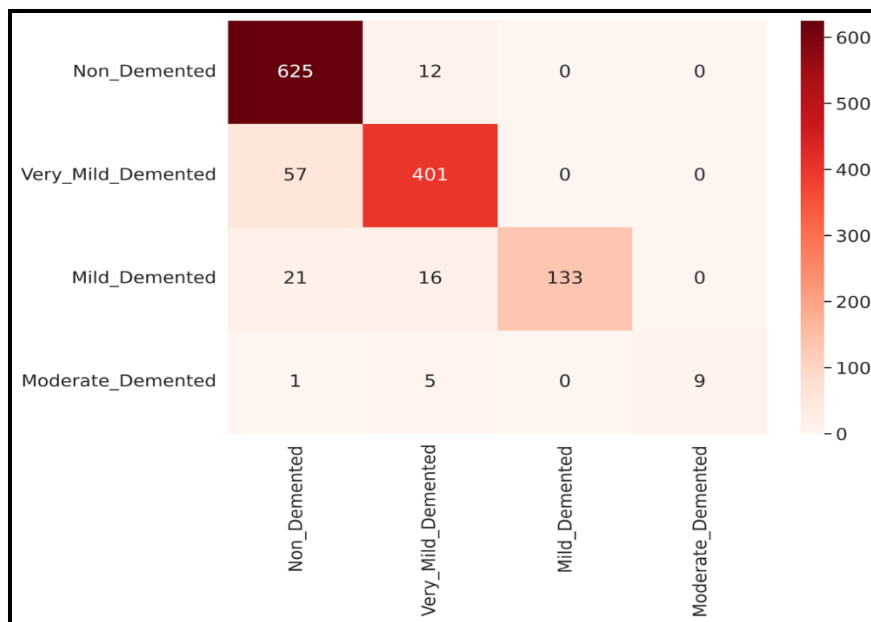

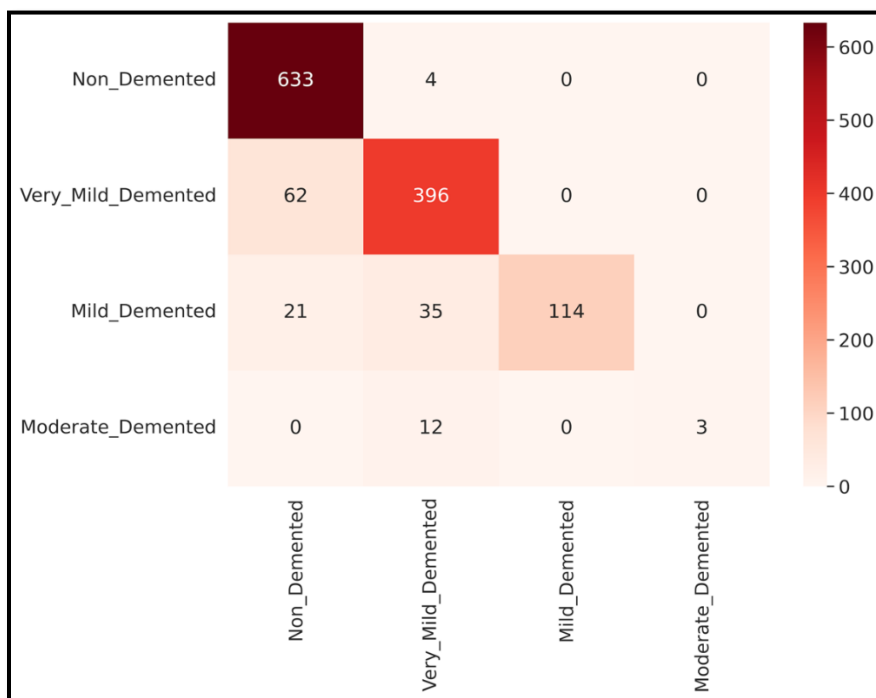*Figure 5 Confusion Matrix for Random Forest Without watershed.*

Figure 6 Confusion Matrix for Random Forest With watershed.

Result Based on SVM:

Support vector machine classifier without watershed achieved an accuracy result of 80.70% and with watershed 96.25%. The confusion matrix of support vector machine classifier is stated in Figure 7 and Figure 8.
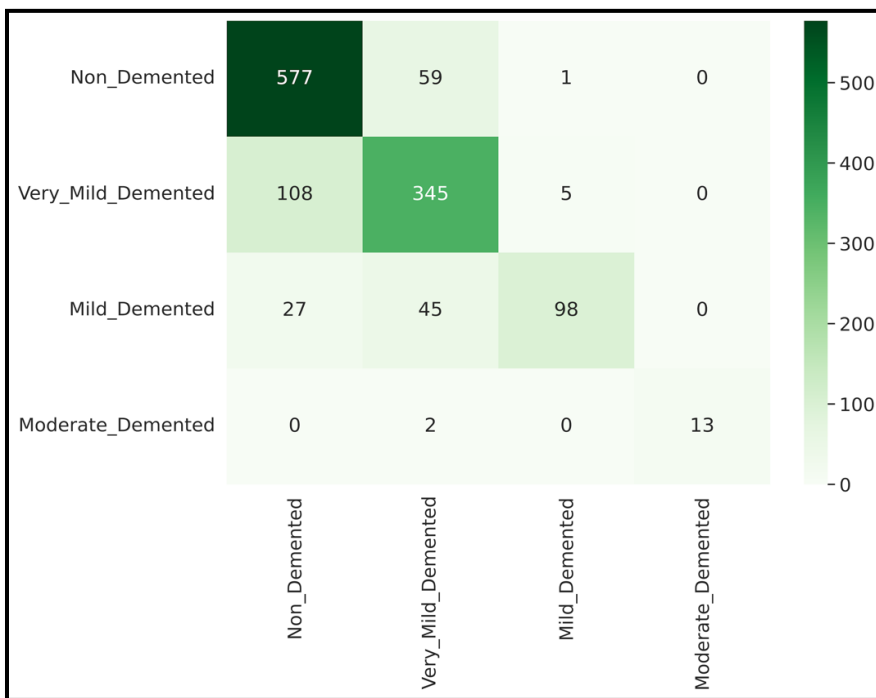

Figure 7 Confusion Matrix for Support Vector Machine Classifier Without watershed.
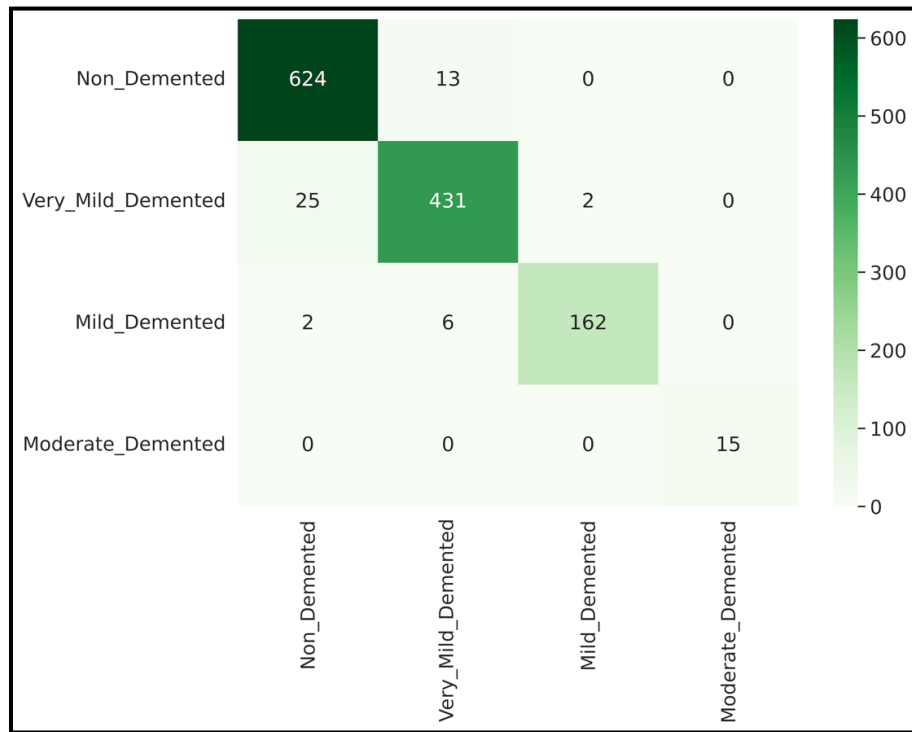
Figure 8 Confusion Matrix for Support Vector Machine Classifier With watershed.

Figure 7 is showing the confusion matrix of support vector machine without watershed and Figure 8 is for those with the watershed. From these, we can observe that for all the classes the number of true positives is higher with the watershed than without the watershed. The difference is notable and for this reason, the overall accuracy with watershed is much higher than without watershed. We can say that, if the dataset were balanced and the other three classes had more images than with watershed the SVM could have given a much better performance.

Result Based on CNN:

The CNN framework was trained and validated using 50 epochs to classify the four levels of dementia. The graph of training & validation accuracy without watershed is given in the Figure 9 and Figure 10. The training & validation loss graph with watershed is demonstrated in Figure 11 and Figure 12. In Figure 9, Figure 10, Figure 11, Figure 12, it can be noticed that without watershed the training accuracy is around 92% and loss is around 10% and validation accuracy is around 94% and loss is around 7%. But with watershed, the training accuracy is around 98% and loss around 3% and validation accuracy are around 94% and loss are around 8%. From Figure 9, Figure 10, Figure 11, Figure 12, for all the graphs we can observe that the opening among training accuracy and validation accuracy or between training loss and validation loss is not so large. So, 50 epochs are enough for training and validating our models with and without the watershed.

From Figure 9 we can see that after 40 epochs the accuracy for training and validation is in a stable position without watershed but in Figure 10, both the training and validation accuracy are stable after around 25 epochs. So, we can say the watershed algorithm helped the model to be trained and validate with fewer epochs.
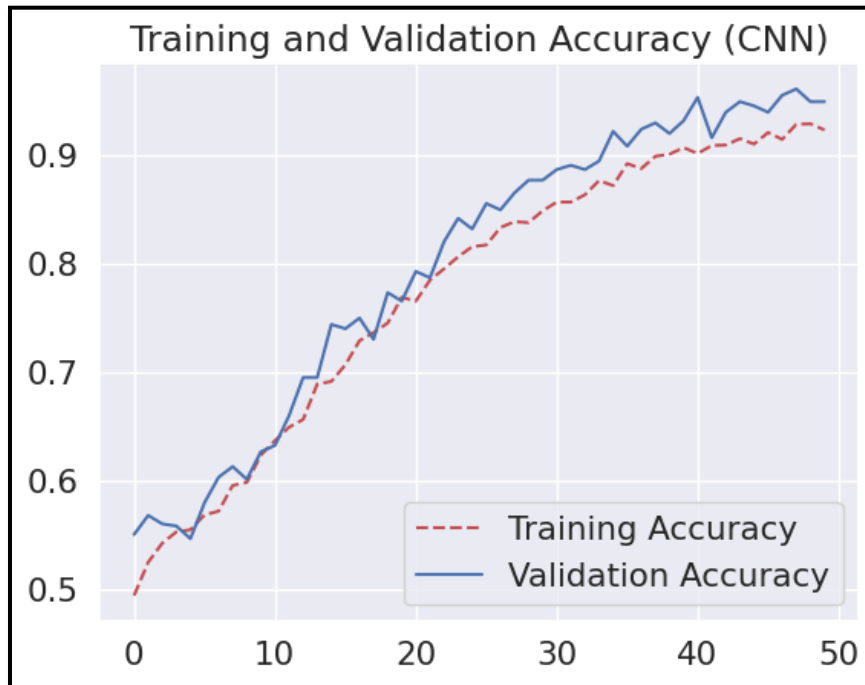

Figure 9 Training and Validation Accuracy Without watershed of CNN.
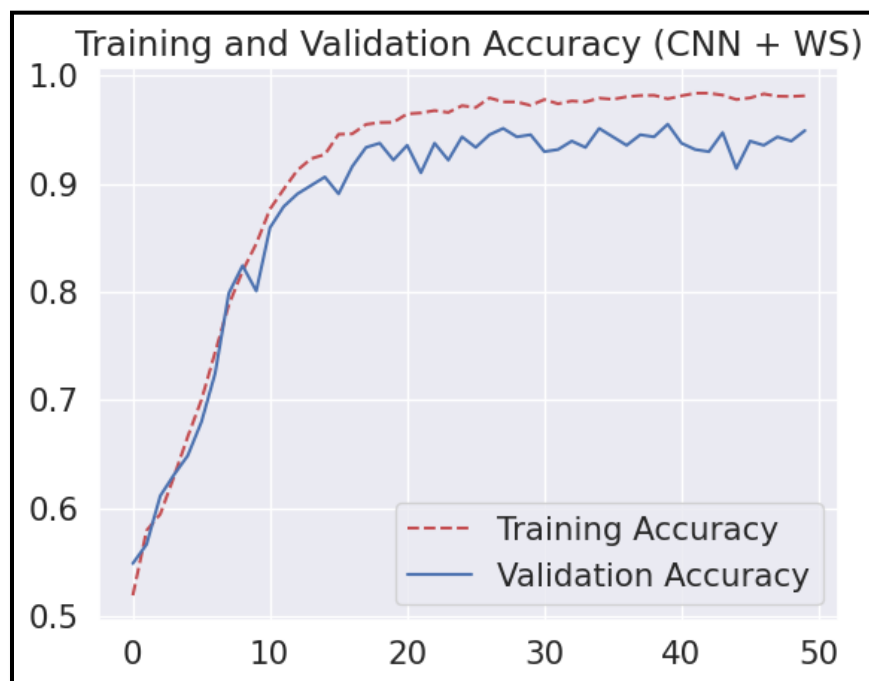

Figure 10 Training and Validation Accuracy with watershed of CNN.
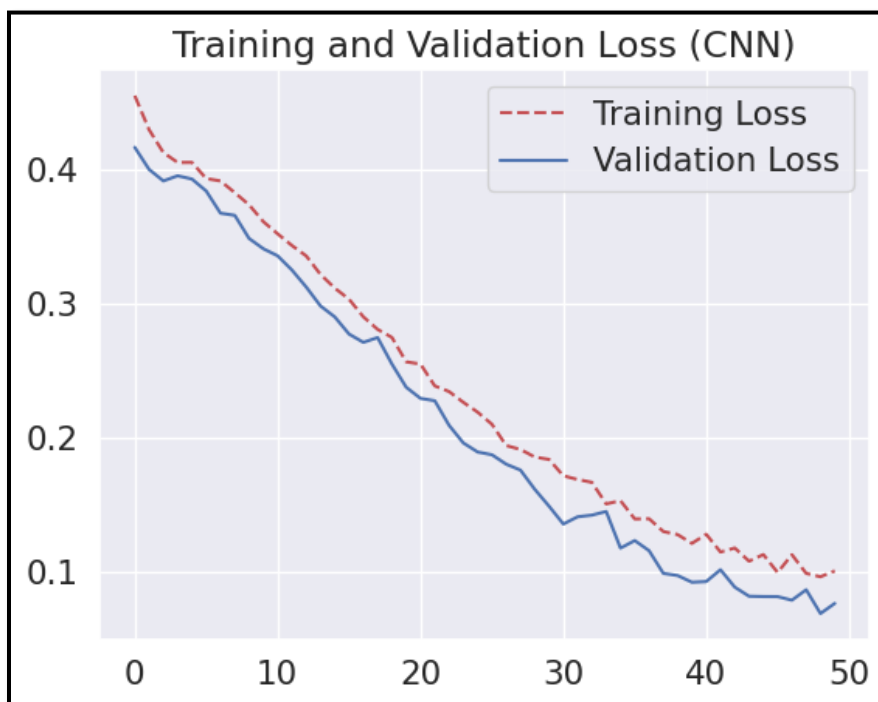
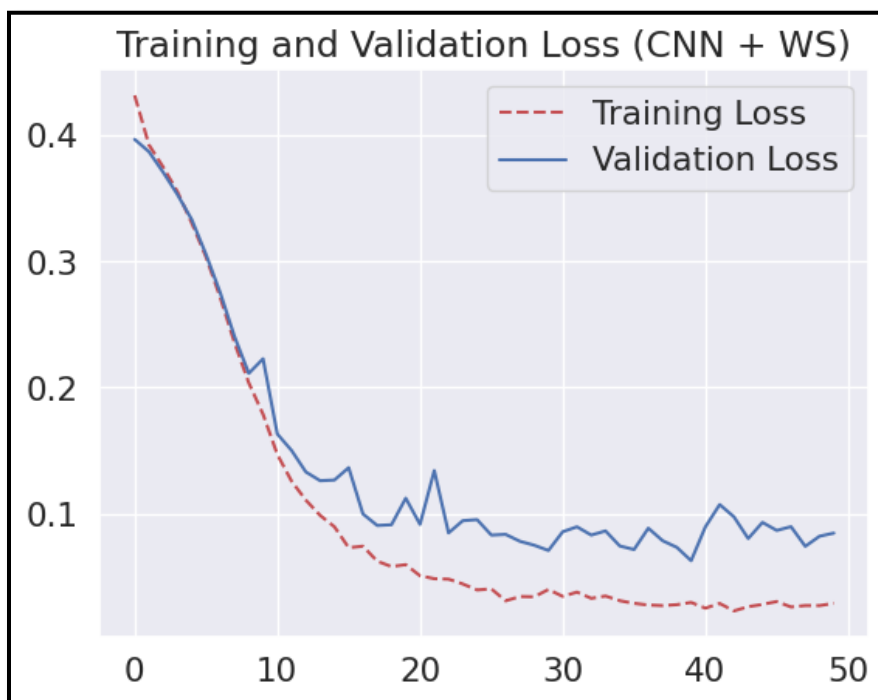*Figure 11 Training and Validation Loss Without watershed of CNN.*

Figure 12 Training and Validation Loss with watershed of CNN.

After the system has been trained, it is evaluated using the testing set in which the images that are present were not given to the system throughout training. The CNN model

achieves accuracy results of 93.98% without watershed feature, and 95.16% with watershed respectively for the test set. The projected class and labeled classes from the four separate categories are plotted on the confusion matrix. The model effectiveness over the training dataset is provided by the confusion matrix. Figure 13 and Figure 14 depicts the confusion matrix used by this structure to categorize the phases of dementia.
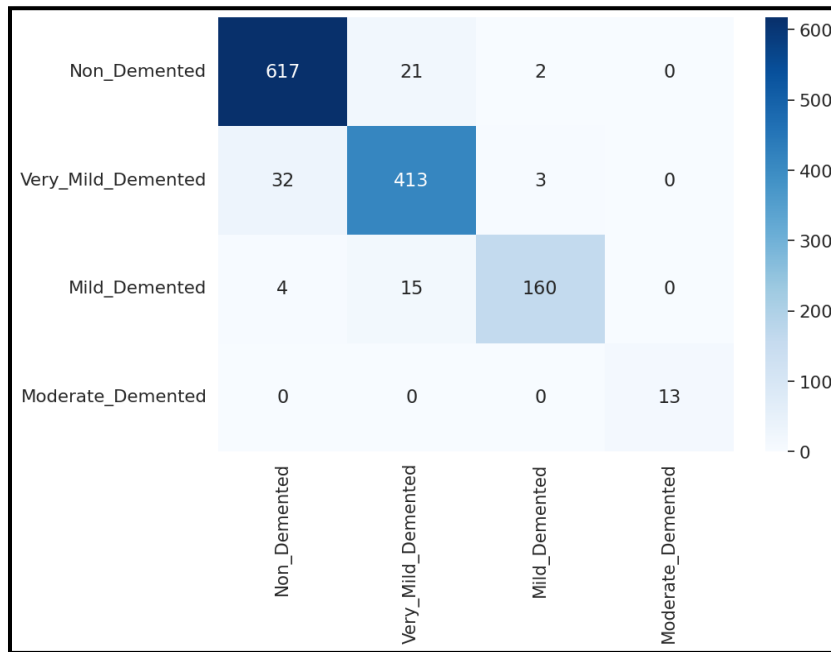

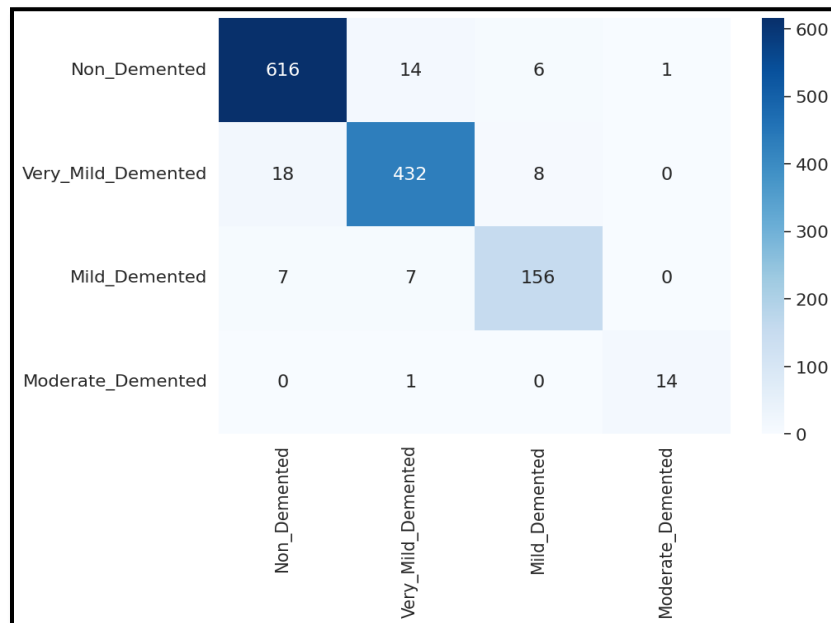*Figure 13 Confusion Matrix for Convolutional Neural Network Without watershed.*

Figure 14 Confusion Matrix for Convolutional Neural Network With watershed.

Figure 13 is showing the confusion matrix for CNN without watershed and Figure 14 is for those with the watershed. From these matrices, we can observe that though the overall accuracy is higher with watershed than without watershed, for the non-demented & mild demented class the number of true positives is fewer with the watershed, but the difference is very low. For other classes, the number of true positives is higher with watershed and the difference is high. So, we can say that if the dataset were balanced and larger, then with watershed the CNN could have given much higher accuracy.

Discussion:

In this research work, the performances metrics, such as recall, precision, and f1-score were also measured for all the models explored. The all-inclusive classification performance of the proposed models is demonstrated in Table 1. It compares the performance of respective RF, SVM, and CNN for the four classes classification on the dataset used in this research work.

Table 1 Accuracy, Precision, Recall and F1-Score Result of Models Used.

Model	Accuracy %	Precision %	Recall %	F1-Score %
RF	91.25	92.0	91.0	92.0
WS + RF	89.53	90.0	90.0	89.0
SVM	80.70	81.0	81.0	80.0
WS + SVM	**96.25**	**98.0**	**97.0**	**97.0**
CNN	93.98	94.0	94.0	94.0
WS + CNN	95.16	95.0	95.0	95.0

The performance comparison presented in Table 1 demonstrates the effectiveness of the SVM model with watershed feature in terms of accuracy, precision, recall, and f1-score. The SVM model consistently outperforms the other models utilized in this research work, indicating its superior ability to classify the different dementia classes accurately. Notably, all the models exhibit relatively high precision values compared to the accuracy scores, which is advantageous for multi-class classifications as shown in Table 1. A higher precision implies a higher likelihood of

correctly predicting an image to its respective class, which is particularly valuable for balanced and larger datasets. Another interesting observation is that despite the low resolution of the MRI images (32 x 32), all the models achieve accuracies exceeding 80 percent. This suggests that even with lower-resolution MRI images, which consume fewer computational resources, it is still possible to effectively identify and classify different dementia levels. It is important to acknowledge that while the performance of the models in this research work is promising, there are

certain limitations to consider. The dataset used in this study may not be representative of the entire population, potentially introducing biases and affecting generalizability. Additionally, the reliance on low-resolution MRI images may impact the ability to capture subtle details and nuances, potentially leading to reduced classification accuracy.

In the case of SVM and CNN models, the inclusion of the watershed algorithm as a feature has shown improvements in accuracy and overall performance. The SVM model, in particular, exhibits a significant enhancement when using the watershed feature, resulting in the best performance among the three models. On the other hand, incorporating the watershed algorithm as a feature in the RF model slightly decreased its performance. This discrepancy could be attributed to the imbalanced and relatively small dataset used in this research work. It is worth noting that while CNN models generally benefit from larger datasets[29], the performance achieved with the available dataset of 6400 images is still noteworthy. The CNN model achieves an accuracy of around 94 percent without the watershed feature and 95 percent with the watershed feature. However, it is important to highlight that the small dataset size may limit

the full potential of the CNN model. To further enhance the performance of the CNN model with the watershed feature, it would be advisable to expand and balance the dataset, allowing for more comprehensive training. Overall, the results suggest that the watershed segmentation algorithm holds promise as a valuable feature in Alzheimer's disease classification using MRI images. It has demonstrated its potential to improve the performance of SVM and CNN models in this research work. To fully exploit the benefits of the watershed algorithm and enhance the classification accuracy further, future directions should focus on obtaining larger and more balanced datasets for training the models effectively.

Performance Comparison with Some Other Works:

The mentioned method is affirmed by contrasting the results to earlier research that focused on the earlier diagnosis of Alzheimer's disease using a classification system of four classes, as shown in the Table 2. The suggested SVM model performs superior in terms of accuracy, precision, and F1-score, classifying the four types of dementia with 96.25% accuracy, 98.0% precision, and 97% F1 score.

Table 2 Results Comparison with Existing Works.

Reference	Methodology	Accuracy	Precision	Recall	F1-Score
This Study	WS + SVM	96.25	98.0	97.0	97.0
Article [30]	CNN	89.0	89.0	89.0	89.0
Article [31]	HT-MGWRO	93.16	90.74	94.24	92.45
Article [32]	Voting 1 (RF, GB, XGB	92.0	91.0	91.0	91.0

Reference	Methodology	Accuracy	Precision	Recall	F1-Score
Article [33]	SVM	92.5	92.0	93.0	92.49
Article [34]	Random Forest	94.0	93.0	98.0	95.43
Article [35]	DNN	85.19	76.93	72.73	74.77
Article [36]	BOF + SVM	93.0	NA	NA	NA
Article [37]	DCNN	81.94	NA	NA	NA

Conclusion:

In this study, a methodology is proposed for classifying Alzheimer's disease (AD) dementia levels using feature extraction with the watershed technique in combination with RF, SVM, and CNN architectures. The standard ADNI data is utilized to train and assess the models for classifying the four stages of dementia. Results indicate that SVM with watershed features outperforms other models on the ADNI dataset, achieving an impressive overall accuracy of 96.25% when tested with data from all four dementia levels. This approach effectively identifies brain areas connected with AD and serves as a valuable decision supporting system for doctors to estimate the seriousness of AD grounded on dementia level. However, it should be noted that the dataset used in this research has a class imbalance issue, which was not addressed in this study. Future work could focus on addressing this limitation to further enhance the model's performance. This proposed approach has the potential to be employed as a standalone framework for dementia level evaluation and AD diagnosis support in clinical settings and can be further validated on diverse datasets in future studies.

Corresponding Author:

Ye Shiren

School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, Jiangsu, China

Email: yes@cczu.edu.cn

Conflicts of Interest Statement

The authors have no conflicts of interest to declare.

References:

1. Gustavsson A, Norton N, Fast T, et al. Global estimates on the number of persons across the Alzheimer's disease continuum. *Alzheimers Dement.* 2023;19(2):658-670. doi:10.1002/alz.12694

2. Li X, Feng X, Sun X, Hou N, Han F, Liu Y. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2019. *Front Aging Neurosci.* 2022; 14:937486. doi:10.3389/fnagi.2022.937486

3. Castellazzi G, Cuzzoni MG, Cotta Ramusino M, et al. A Machine Learning Approach for the Differential Diagnosis of Alzheimer and Vascular Dementia Fed by MRI Selected Features. *Front Neuroinformatics.* 2020;14:25. doi:10.3389/fninf.2020.00025

4. Odusami M, Maskeliūnas R, Damaševičius R, Krilavičius T. Analysis of Features of Alzheimer's Disease: Detection of Early Stage from Functional Brain Changes in Magnetic Resonance Images Using a Finetuned ResNet18 Network. *Diagnostics.* 2021;11(6): 1071. doi:10.3390/diagnostics11061071

5. Developer. Role of MRIs in Detecting Alzheimer's. Envision Radiology. Published July 8, 2020. Accessed April 28, 2023. https://www.envrad.com/role-of-mris-in-detecting-alzheimers/

6. Liu S, Liu S, Cai W, Pujol S, Kikinis R, Feng D. Early diagnosis of Alzheimer's disease with deep learning. In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI).* IEEE; 2014:1015-1018. doi:10.1109/ISBI.2014.6868045

7. Tanveer M, Richhariya B, Khan RU, et al. Machine Learning Techniques for the Diagnosis of Alzheimer's Disease: A Review. *ACM Trans Multimed Comput Commun Appl.* 2020;16(1s):1-35. doi:10.1145/3344998

8. Wen J, Thibeau-Sutre E, Diaz-Melo M, et al. Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. *Med Image Anal.* 2020;63:101694. doi:10.1016/j.media.2020.101694

9. AlSaeed D, Omar SF. Brain MRI Analysis for Alzheimer's Disease Diagnosis Using CNN-Based Feature Extraction and Machine Learning. *Sensors.* 2022;22(8):2911. doi:10.3390/s22082911

10. Kamal MdS, Northcote A, Chowdhury L, Dey N, Crespo RG, Herrera-Viedma E. Alzheimer's Patient Analysis Using Image and Gene Expression Data and Explainable-AI to Present Associated Genes. *IEEE Trans Instrum Meas.* 2021;70:1-7. doi:10.1109/TIM.2021.3107056

11. Chagué P, Marro B, Fadili S, et al. Radiological classification of dementia from anatomical MRI assisted by machine learning-derived maps. *J Neuroradiol.* 2021;48(6):412-418. doi:10.1016/j.neurad.2020.04.004

12. Bharanidharan N, Rajaguru H. Improved chicken swarm optimization to classify dementia MRI images using a novel controlled randomness optimization algorithm. *Int J Imaging Syst Technol.* 2020; 30(3):605-620. doi:10.1002/ima.22402

13. Haque S, Thaneeghaivel R, Gangwar M, Singh S. A Deep Learning Model in the Detection of Alzheimer Disease. *Turk J Comput Math Educ.* 2021;12(10):4013-4022. doi: https://doi.org/10.17762/turcomat.v12i10.5113

14. Hazarika RA, Kandar D, Maji AK. An experimental analysis of different Deep Learning based Models for Alzheimer's Disease classification using Brain Magnetic Resonance Images. *J King Saud Univ - Comput Inf Sci.* 2022;34(10):8576-8598. doi:10.1016/j.jksuci.2021.09.003

15. AbdulAzeem Y, Bahgat WM, Badawy M. A CNN based framework for classification of Alzheimer's disease. *Neural Comput Appl.* 2021;33(16):10415-10428. doi:10.1007/s00521-021-05799-w

16. Suganthe RC, Latha RS, Geetha M, Sreekanth GR. Diagnosis of Alzheimer's Disease from Brain Magnetic Resonance Imaging Images using Deep Learning Algorithms. *Adv Electr Comput Eng.* 2020; 20(3):57-64. doi:10.4316/AECE.2020.03007

17. Murugan S, Venkatesan C, Sumithra MG, et al. DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images. *IEEE Access.* 2021;9:90319-90329. doi:10.1109/ACCESS.2021.3090474

18. Miltiadous A, Tzimourta KD, Giannakeas N, et al. Alzheimer's Disease and Frontotemporal Dementia: A Robust Classification Method of EEG Signals and a Comparison of Validation Methods. *Diagnostics.* 2021;11(8):1437. doi:10.3390/diagnostics11081437

19. Hu J, Qing Z, Liu R, et al. Deep Learning-Based Classification and Voxel-Based Visualization of Frontotemporal Dementia and Alzheimer's Disease. *Front Neurosci.* 2021; 14:626154. doi:10.3389/fnins.2020.626154

20. Ieracitano C, Mammone N, Bramanti A, Hussain A, Morabito FC. A Convolutional Neural Network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings. *Neurocomputing.* 2019;323:96-107. doi:10.1016/j.neucom.2018.09.071

21. Chandaran SR, Muthusamy G, Sevalaiappan LR, Senthilkumaran N. Deep Learning-based Transfer Learning Model in Diagnosis of Diseases with Brain Magnetic Resonance Imaging. *Acta Polytech Hung.* 2022;19(5):127-147. doi:10.12700/APH.19.5.2022.5.7

22. Alzheimer MRI Preprocessed Dataset. Accessed April 28, 2023. https://www.kaggle.com/datasets/sachinkumar413/alzheimer-mri-dataset

23. Jack CR, Bernstein MA, Fox NC, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging.* 2008;27(4):685-691. doi:10.1002/jmri.21049

24. Kornilov A, Safonov I, Yakimchuk I. A Review of Watershed Implementations for Segmentation of Volumetric Images. *J Imaging.* 2022;8(5):127. doi:10.3390/jimaging8050127

25. Shaik F, Prasad MNG. Feature Extraction from Micrograph Images Using Watershed Segmentation Approach. *Int J Appl Eng Res.* 2012;5(23-24):3665-3674.

26. Zhang W, Jiang D. The marker-based watershed segmentation algorithm of ore image. In: *2011 IEEE 3rd International Conference on Communication Software and Networks*. IEEE; 2011:472-474. doi:10.1109/ICCSN.2011.6014611

27. Sarica A, Cerasa A, Quattrone A. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review. *Front Aging Neurosci*. 2017;9:329. doi:10.3389/fnagi.2017.00329

28. Sheykhmousa M, Mahdianpari M, Ghanbari H, Mohammadimanesh F, Ghamisi P, Homayouni S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2020;13:6308-6325. doi:10.1109/JSTARS.2020.3026724

29. Chaganti SY, Nanda I, Pandi KR, Prudhvith TGNRSN, Kumar N. Image Classification using SVM and CNN. In: *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*. IEEE; 2020:1-5. doi:10.1109/ICCSEA49143.2020.9132851

30. De Silva K, Kunz H. Prediction of Alzheimer's disease from magnetic resonance imaging using a convolutional neural network. *Intell-Based Med*. 2023;7:100091. doi:10.1016/j.ibmed.2023.100091

31. Bharanidharan N, Harikumar R. Modified Grey Wolf Randomized Optimization in Dementia Classification Using MRI Images. *IETE J Res*. 2022;68(4):2531-2540. doi:10.1080/03772063.2020.1715852

32. Bharati S, Podder P, Thanh DNH, Prasath VBS. Dementia classification using MR imaging and clinical data with voting based machine learning models. *Multimed Tools Appl*. 2022;81(18):25971-25992. doi:10.1007/s11042-022-12754-x

33. Herzog NJ, Magoulas GD. Brain Asymmetry Detection and Machine Learning Classification for Diagnosis of Early Dementia. *Sensors*. 2021;21(3):778. doi:10.3390/s21030778

34. Mohammed BA, Senan EM, Rassem TH, et al. Multi-Method Analysis of Medical Records and MRI Images for Early Diagnosis of Dementia and Alzheimer's Disease Based on Deep Learning and Hybrid Methods. *Electronics*. 2021;10(22):2860. doi:10.3390/electronics10222860

35. Prajapati R, Khatri U, Kwon GR. An Efficient Deep Neural Network Binary Classifier for Alzheimer's Disease Classification. In: *2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. IEEE; 2021:231-234. doi:10.1109/ICAIIC51459.2021.9415212

36. Bansal D, Khanna K, Chhikara R, Dua RK, Malhotra R. Classification of Magnetic Resonance Images using Bag of Features for Detecting Dementia. *Procedia Comput Sci*. 2020;167:131-137. doi:10.1016/j.procs.2020.03.190

37. Bidani A, Gouider MS, Travieso-González CM. Dementia Detection and Classification from MRI Images Using Deep Neural Networks and Transfer Learning. In: Rojas I,

Joya G, Catala A, eds. *Advances in Computational Intelligence.* Vol 11506. Lecture Notes in Computer Science. Springer International Publishing; 2019:925-933. doi:10.1007/978-3-030-20521-8_75