A Comprehensive Study of Machine Learning, Deep Learning and Attention-Based Models for Alzheimer's Disease Detection

Pranab Hazra[1], Mainak Dey[2], Anurupa Paul[2], Koushiki Mitra[2], Sandhya Pattanayak[1]

[1]Assistant Professor, Electronics & Communication Engineering, Narula Institute of Technology, Agarpara, Kolkata, West Bengal, India.

[2]Student of B-TECH in the Department of Electronics & Communication Engineering (ECE), Narula Institute of Technology, Agarpara, Kolkata, West Bengal, India.

Abstract: Alzheimer's Disease (AD) is a major public health concern globally. Timely and accurate diagnosis is vital for actual treatment and patient care. Recent advancements in deep learning have shown promise in improving disease detection through the analysis of neuroimaging data, such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) scans. The deep learning models examined in this study include convolutional neural networks (CNNs) and (SPECT). These tools aid in identifying structural and functional variations in the brain associated with the disease. By analyzing the substantial amount of data generated by these imaging and diagnostic tools, along with other clinical data such as physiological parameters and medical histories, modern statistical and computational tools like Machine Learning (ML) and Deep Learning (DL) can reveal patterns also relationships that may not be directly apparent to human forecasters. We rigorously analyze the effects of different optimizer algorithms, including the attention model along with deep learning. Additionally, we explore the influence of batch sizes during the model training process. One of the key contributions of this research is the in-depth examination of batch size selection in the context of AD detection. We demonstrate how batch size affects training dynamics, convergence speed, and memory requirements, shedding light on the trade-offs associated with different batch sizes. We employ rigorous evaluation metrics, including accuracy, sensitivity, specificity, and area Under the Receiver Operating Characteristic Curve (AUC-ROC), to quantify the performance of our models. In this project work, results showcase that no one-size-fits-all method exists, as the optimal combination of model, optimizer, and batch size depends on the specific problem domain and dataset characteristics. In this project comparative studies using machine learning, deep learning, and attention model. We have used MRI images of the ADNI dataset. By systematically exploring the interplay between different factors in model optimization, we offer a nuanced perspective on achieving superior model performance across diverse tasks and datasets.

Introduction

Alzheimer's Disease (AD) is a communal form of dementia caused by the degeneration of brain cells and tissues. It results in memory loss, impaired daily functioning, and behavioral changes. Individuals with Alzheimer's disease face significant challenges and require constant care. The disease is categorized by the presence of neurofibrillary tangles and absent-minded plaques in the hippocampus, caused by abnormal protein accumulation. Currently, approximately 44.4 million people worldwide suffer from dementia (Beheshti et al., 2013) with projections indicating a significant increase in the future.

AD is an irreversible genetic condition that significantly affects daily life. Patients struggle with basic tasks and lose memory and reasoning abilities due to brain tissue degeneration. The disease worsens over time and primarily affects individuals over 60. Symptoms include memory loss, communication difficulties, sleep pattern changes, anxiety, depression, and poor decision-making. The severity of this disease worsens over time. Unfortunately, no cure exists Low-income countries like Bangladesh face a high prevalence of dementia and limited resources for research and treatment. Additionally, being overweight may contribute to AD, emphasizing the importance of healthy lifestyles (Frisoni et al., 2010). Early detection is crucial, and machine learning algorithms show promise in identifying AD at a primary stage.

The significance of early detection of Alzheimer's disease using Machine Learning (ML) and Deep Learning (DL) techniques cannot be overstated. These advanced computational approaches offer promising avenues for identifying subtle patterns and biomarkers indicative of the disease at its earliest stages (Alzheimer's Association, 2016). By leveraging ML and DL algorithms on diverse datasets, researchers can potentially enhance diagnostic accuracy, enable timely intervention, and ultimately improve patient outcomes. This underscores the critical role those innovative technologies play in addressing the challenges of early Alzheimer's detection, ultimately contributing to more effective therapeutic strategies and patient care. This study goals to assess and compare the effectiveness of multiple machine learning (ML) and deep learning (DL) models for the detection of Alzheimer's disease. By evaluating different ML and DL techniques on relevant datasets, this research aims to identify which models exhibit superior act in terms of accuracy, sensitivity, and specificity, thereby providing insights into optimal methodologies for Alzheimer's Disease detection using computational approaches.

Here in this project work, we have done comparative studies using machine learning, deep learning, and attention models. We have used MRI images of the ADNI dataset. By using those methods, we have got to know we are getting the best accuracy for the attention model and the better accuracy for machine learning and deep learning techniques we have got average accuracy.

It is a neurogenerative disease. Alzheimer's is classified into three types Very Mild demented, Moderate demented and Non Demented. By checking the MRI images and after training and testing this dataset images through the models we can differentiate the severity and the level of this disease (Yang et al., 2017). MRI techniques from the Human Connective Project are going to be accustomed to mapping the consequences of AD on brain property. We can observe there is an optimal difference between CNN and SVM. These methods we have used to classify Alzheimer's disease comparing them with the attention model. Hence the attention model gives better results than both models. In the evaluation of Alzheimer’s classification in the final decade, Support Vector Machine, Naive Bayes, Artificial Neural Networks for machine learning, and deep learning are the most widely used highly observed. The optimization technique is the primary difference between convolution Neural Networks. Support Vector Machine, It offers a solution of the highest value, while ANN provides the most suitable solution. Feature extraction is an important step in every SVM and ANN.

Supervised and unsupervised algorithms of machine learning have been used and it is observed that from SVM we have got the highest accuracy in Machine Learning. In the case of DL we have used the Resnet model, where we have got the highest accuracy in the Resnet 101 module. These methods we have used to classify Alzheimer's disease comparing them with the attention model where this model gives better results than both of these models. The evaluation of Alzheimer’s classification in the final decade, Support Vector Machine, for ML and DL and Resnet for deep learning method and attention model are the most widely used which is highly observed.(Kloppel et al., 2017) The optimization technique is the main difference between the deep learning Neural Networks (Resnet 101), Support Vector Machine and attention model. It provides a solution of the highest quality, while the attention model offers the most suitable solution. Feature abstraction is an essential step in every SVM and Resnet model and attention model where we get the comparatively highest accuracy in the attention model and Resnet 101 of the deep learning method.

This relative study on Machine Learning (ML), Deep Learning (DL), and attention-based DL models for Alzheimer's disease detection aims to address several key research questions. These include evaluating the performance of traditional ML algorithms like SVM and Random Forest against DL architectures such as CNN and RNN in terms of correctness for disease detection (Lao et al., 2018 & Simon et al., 2016). Additionally, the study seeks to assess the effectiveness of DL models, particularly CNN and RNN, in capturing complex patterns from Alzheimer 's-related brain imaging data compared to ML methods.

The investigation also focuses on understanding how incorporating attention mechanisms in DL models impacts accuracy and interpretability in disease detection. Other questions address generalization across diverse datasets, interpretability of models, computational efficiency (Weiner et al., 2013) and integration of multimodal data sources. By investigative these aspects, the study aims to provide insights into the optimal use of ML, DL, and attention-based DL models for Alzheimer's disease detection.

Methodology

In this sector, we deliberate our used dataset the evaluation of the model and the analysis of our result.

Datasets

We have used Python language libraries and firstly we have installed all the required libraries like pytorch, tensor flow etc. They imported the MRI scan brain images shown in Fig 1 from open source passed all data into our model and then preprocessed all datasets of all models extracted the features of all the models then did the EDA after that we split our datasets into two parts i.e. trained data and test data then transfer it in all the models like the machine learning, deep learning, deep learning with the attention network and then the models give the predicted output.

Machine Learning Model

A machine learning model regulates the output obtained after running a machine learning process on the collected data. Thus, choosing the correct model to give our desired output is important. We have used several machine learning models such as Logistic Regression, Decision Tree, Random Forest, SVM, gradient boosting, K- Neighbors, Principal Component Analysis, andIndependent Component Analysis. After selecting the model, we passed the processed data into the following models to find shapes and make forecasts. This allows the model to study from the data and perform a set of tasks. Evaluating the Model: After training the model (Schroeder et al., 2018), we need to check how the models are performing. This is done by testing the replica's performance against previously unseen data. The hidden data used is the test set that we earlier split the data into. Further, when used on testing data, we can compare the accuracy and speed of the different models. Such as Support Vector Machine (SVM) whereas a non-parametric classifier has no preliminary information available regarding its distribution. Training sets consist of paired input and output decision functions are obtained which are used to identify the input variable quantity within the new test and data set through these pairs (Reitz et al., 2016). The new transformed dimension is being investigated within the separator plane of the maximum margin. Decision Tree is a classifier algorithm construction of trees that are very simple. They work based on discrete value parameters. Random Forest is a well-known machine learning algorithm belonging to the SVM technique. Both classification and logistic regression algorithms is solved by Random Forest is also used to improve the predictive accuracy, which helps to prevent the problems of over fitting. One of the famous algorithms is logistic regression. It comes under the superintend classification technique Clustering is a technology system to catch classes of observations where a specific dataset is provided, that contributes identical features and where a data analysis technique is used to get a clue about the structure ofthe data. It is resolute and interprets the data which is one of the important features of this unsupervised clustering it also reforms compact clustering and can work on numerical data but it has limitations that highly depend on the original data and it is difficult to forecast the number of clusters. PCA is a famous non-superintended algorithm that has been used in different ways like analysis, compression, de-noising (Yang et al, 2017) reducing the dimension of data, etc. It helps to eliminate similar data in the line ofcomparison that does not need a bit to make data decisions. It is an unsupervised machine learning process that is used to reduce the dimension of a dataset although it collects the information as much as possible. Many techniques have been discovered for this process, but principal component analysis is the most widely used (Falahati et al., 2014). Its main motive is to reduce the size of a dataset. That's how machine learning is being implemented.

Deep Learning Methods

This process of Deep Learning is used to perform maximum computations on a huge some of data. The term Deep Learning has been taken from machine learning modules that use neural networks with more than three layers for more accurate output. Complex picture patterns, text, sound, and other kinds of audio-visual data can be recognized with the help of deep learning. Deep Learning methods are also used for the automation of different tasks which require human intelligence such as describing any images or transcription of audio files into a text document.

ResNet

Residual Network is a DL architecture that has meaningfully advanced in the field of computer vision and deep neural networks. It was designed to address the disappearing gradient problem that often occurs in very deep neural networks (Adrien et al., 2014).The key invention behind ResNet is the introduction of residual blocks, which enable the training of extremely deep neural networks more effectively. In a residual block, the input to a layer is collected with the output of the layer, creating a "shortcut connection" or "skip connection". This allows the network to study residual information, i.e., the difference between the output and input of a layer. As a result, even as the network depth increases, gradients can flow more easily during training, mitigating the vanishing gradient issue.

Attention Model

Attention is a mechanism that causes kind of mechanism to keep focus on the most important data which is learned by the algorithm. An attention model is actually in a deep learning neural network that lets the network subsequentially attention on a mini set of the input after processing it, and then revolution its focus to some other part of the input. This makes it easier to reason chronologically about the data, even if the data isn't sequential in nature (Moradi. Et al., 2015). An attention model can be learned by reinforcement learning, or by back propagation if the policy is differentiable.

Working With Model

Attention-based deep learning models have many advantages over traditional deep learning models: It has selective focus which allows the model to selectively focus on related parts of the input data, rather than processing the entire input uniformly. This selective attention enables the model to capture long-range needs and intricate patterns more effectively, leading to improved performance (cheng et al., 2017). It handles variable-length inputs such as Recurrent Neural Networks (RNNs) or convolutional neural networks (CNNs), often struggling with variable-length inputs. Attention-based models can handle variable-length inputs more. It provides a degree of interpretability to the model by highlighting the parts of the input data that are most relevant for making predictions (Ghazi et al., 2019).

This interpretability can be valuable in understanding why the model makes certain decisions, which is especially important in applications. It can sometimes reduce the computational complexity compared to

traditional deep learning architectures (Jude et al., 2018). By focusing computational resources on the most related parts of the input data. It can effectively handle memory-related tasks by allowing the model to selectively attend to relevant information stored in memory. Overall, attention-based deep learning models offer a more flexible and interpretable approach to learning from sequential data, with improved performance on tasks involving variable-length inputs and long- range dependencies.

The Performance Evaluation

Here we focus on the study. The most overall way for comparing algorithms is classification performance without convergingon a class. The empirical measures are used mostly, and the accuracy cannot be

distinguished between the number of accurate tables and the different types of class levels.

- **Accuracy**: Total number of records which are correctly classified by the classifier is mentioned to as accuracy. Accuracy can also be defined as the ratio of test set tuples that are classified accurately bythe model (Korolev et al., 2015).
- **Sensitivity**: The true positive rate that includes the amount of the positive tuples that are correctly classified is known as the sensitivity.
- **Specificity**: The rate by virtue of which a Test or a analytical method sets a standard diagnosis for a person who is not affected is known as Specificity.
- **Roc Curve**: ROC stands on the Receiver Operating Characteristics curve which displays both the specificity and sensitivity of the test. The comparison between TPR (True Positive Rate) and FPR (False Positive Rate) is known as the ROC curve (Lecun et al., 1998).

i.e., TPR = TP (TP+FN) and FPR = FP (FP+TN). ----------- (1)

$$AUC = \int_0^1 t_{pr}(f_{pr})df_{pr} = P(X1 > X0)$$ ------------- (2)

Where tpr is the True Positive Rate, fpr is the false positive rate, and X0 and X1 are the confidence scores for a negative and positive instance, respectively.

Result Analysis

The experiment in this project work was conducted using ADNI Dataset which is present on the Kaggle website (Kaggle kernels output Jeon woo park /alzheimer- detection-and- classification-98-7-acc –

p/path/to/test), it holds more than 40000 images marked as Non demented, Mild Demented, Moderate Demented and Very Mild Demented.

Required Experimental Data

We use our work with only 3200 images, from these 2000 images will be used as a training and the rest of them will be test sets. CNN model is trained by 2000 images by the size 224x224 using different types of Machine Learning and Deep Learning modules. The image slices were gained from the brain using Matplotlib which gave the right-angle, center, and left-angle views. These images were then augmented to increase the size. The transformations involved a zoom range of 0.07, a width shift range of 0.07, a hide shift range of 0.07, and a horizontal flip [2]. In this manner, the image dataset was formed for each of the four classes [6]. The dataset contained a total of 1700 images where each image had a dimension of 224x224 and a format of jpg. The image was processed in grayscale mode. Using this dataset, the model was trained with a technique of transfer learning (Fargo et al., 1998 & Rahman et al., 2013). In this project work we can use the many different types of Machine Learning and Deep Learning techniques based on CNN and the attention mechanism with deep learning and their accuracy is shown below. This experiment was conducted by Google Collaboratory with Epoch 60 and different types of backend functions.

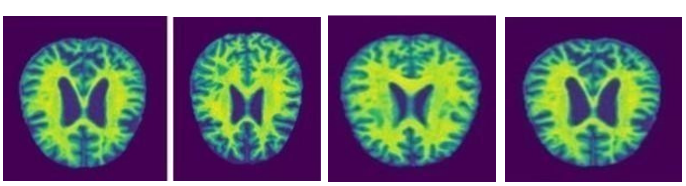


Fig 1. Different types of Datasets

In this study, we utilized the ADNI dataset and the universal challenge for the automated prediction of MCI and MRI datasets. Once the data is imported, we restructured and transformed it into a more organized pattern. We started by dropping unnecessary columns from both datasets and concatenated the ADNI datasets before performing further operations on the merged dataset (Shahbaz et al., 2019). To ensure the accuracy of our model, we removed null values from the CDR column, as we cannot train our model on null values. We used the simple imputer library to fill the null values in the SES and MMSE columns with the most frequent occurrences and median values, respectively. Additionally, to prevent redundancy, we removed duplicate values from all columns. Next, we used Label Encoder to convert the non-numerical values of the

group column (Demented and Non-Demented). This transformation makes it easier for the model to train. We generated a correlation matrix and heat map to visualize the relationship between features in the dataset.

The output obtained after running respective machine learning and deep learning algorithms with the attention network on collected data is determined by the machine learning, deep learning model used. It is critical to select the correct model to obtain our desired output. In this study, we employed various Machine Learning Models, such as Logistic Regression, Decision Tree, Random Forest, SVM, Gradient Boosting, K-Neighbours, Principal Component Analysis, Independent Component Analysis, and various types of Deep learning models such as ResNet 18, ResNet 50, ResNet 101, ResNet 152 and combine the deep learning model with an attention mechanism to find patterns and make predictions on our processed data. The model learns from the data and performs a set of tasks to produce the desired output (Liu et al., 2018). To estimate the model's performance, we tested its accuracy and speed against previously unseen data, using the test set that we had formerly split from the data. This evaluation allows us to compare the different models' performance on the test data and determine which model works best for our specific task.

Discussion

In this paper, we have compared the Machine Learning models, the deep learning models and attention based deep learning models. As shown in Table 1, We can see that among the popular machine learning models, the SVM model gives the highest accuracy of 96.25%, the deep learning model (ResNet 101) gives the accuracy of 97.14% and the attention network with the deep learning techniques, we got our desired output with the highest accuracy and AUC with this combination that's is 98.93%, and we got this accuracy is the highest accuracy for the last 10 years of research project works. So, our comparative study will show the best result with the attention network.

For our comparative study, we can see that in the machine learning model (Jain et al., 2018), SVM gives the highest accuracy when we used the deep learning model and the accuracy increased and when additionally we combined the attention network with the deep learning we got the highest accuracy so, we can say that deep learning is better than the machine learning algorithm and if we added an attention network with the deep learning then the accuracy will further be increased and this stage we have got our best accuracy.

The present study is designed to conduct a comparative review of supervised and unsupervised machine learning algorithms and deep learning algorithms and the combination of attention networks for the early detection of Alzheimer's disease (AD) with a focus on classification as shown in Table 1. Alzheimer's disease is a complex neurodegenerative disorder, and early detection plays a critical role in timely intervention and action.

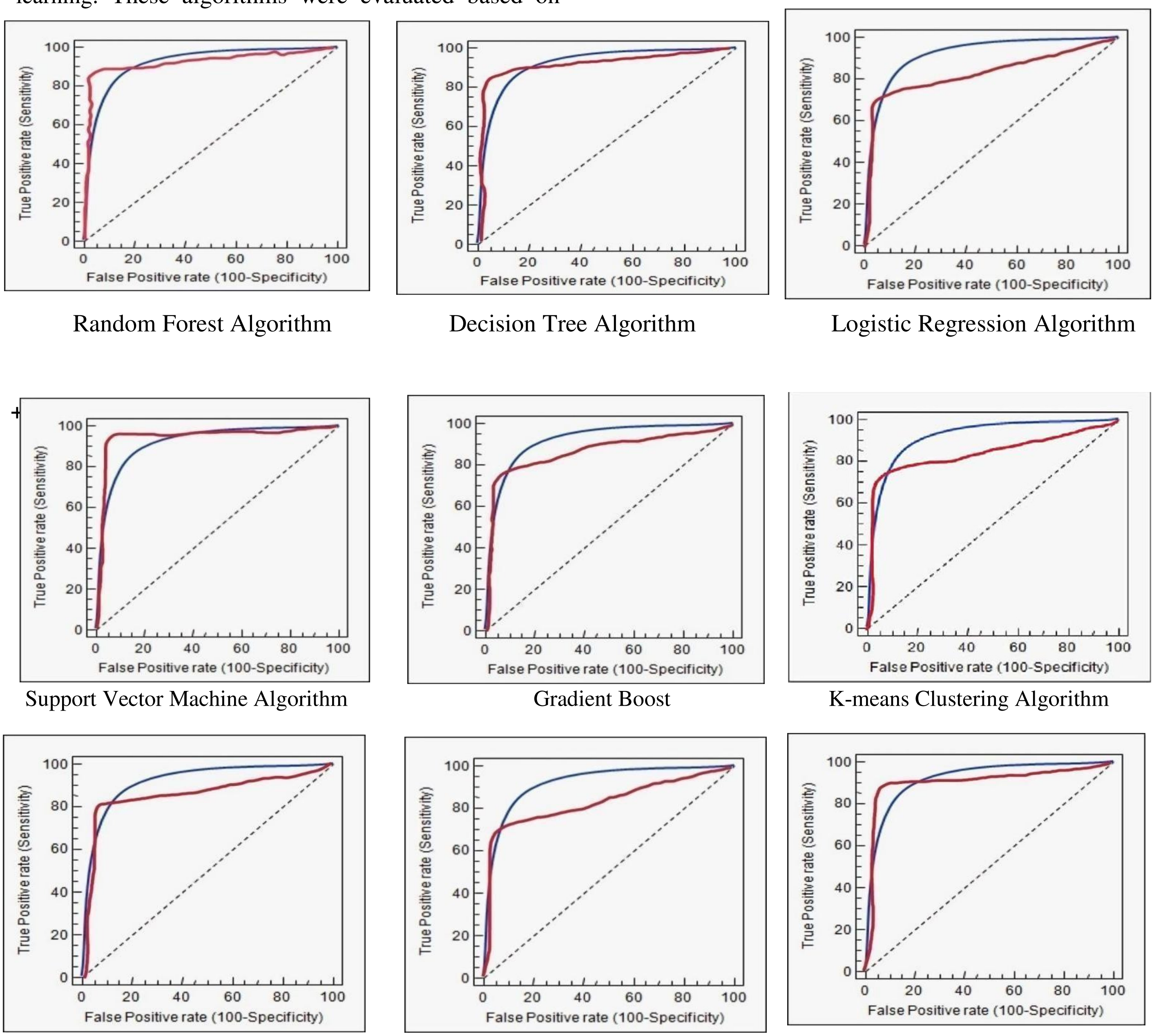
Table 1. Comparative Study of Various Models

Machine Learning Models					Deep Learning Models					Deep Learning Models with ECA-Net Attention Mechanism				
Models	**Accuracy**	**Sensitivity (SEN)**	**Specificity (SPE)**	**AUC**	**Models**	**Accuracy**	**Sensitivity (SEN)**	**Specificity (SPE)**	**AUC**	**Models**	**Accuracy**	**Sensitivity (SEN)**	**Specificity (SPE)**	**AUC**
Random Forest Algorithm	84.62	86.52	83.74	0.845	ResNet 18	78.23	76.51	79.85	0.784	ResNet 18	71.02	69.58%	72.98%	0.772
Decision Tree Algorithm	86.23	85.32	88.03	0.862	ResNet 50	88.23	89.02	84.21	0.893	ResNet 50	89.96	88.68%	89.20%	0.921
Logistic Regression Algorithm	73.42	73.21	76.98	0.732	**ResNet 101**	**97.14**	**96.61**	**97.92**	**0.973**	**ResNet 101**	**98.93**	**99.45%**	**98.62%**	**0.991**
Support Vector Machine Algorithm	**96.25**	**95.14**	**97.89**	**0.962**	ResNet 152	88.25	86.87	89.20	0.880	ResNet 152	96.52	95.04%	94.02%	0.973
Gradient boost	69.32	67.81	70.85	0.693	VGG19	75.09	79.53	78.52	0.801	VGG19	78.51	79.69	77.01	0.816
K-Means Clustering algorithm.	72.23	74.60	71.49	0.720	Efficient NetB0	73.51	74.02	75.11	0.786	Efficient NetB0	75.28	76.21	77.14	0.798
Principal Component Analysis.	80.28	81.27	79.10	0.801	Sequential	82.34	81.52	82.20	0.832	Sequential	84.21	85.34	86.76	0.853
Independent Component Analysis.	65.23	68.09	63.70	0.652	Xception	69.20	68.12	70.20	0.742	Xception	66.30	67.14	64.18	0.731
Adaboost	79.63	77.51	75.98	0.857	GAN	85.55	83.21	82.56	0.896	GAN	88.94	84.51	83.27	0.913

The present study aimed to conduct a relative review of supervised and unsupervised machine learning procedures and deep learning algorithms and the combination of attention networks for the early detection of Alzheimer's disease (AD) with a focus on classification. Alzheimer's disease is a complex neurodegenerative disorder, and early detection plays a crucial role in appropriate intervention and dealing (Helaly et al., 2022).

By this comparison, the study sought to identify the best comparative model and algorithms for the finding and progression of AD. The review encompassed a wide range of supervised and unsupervised ML Algorithms, Including Logistic Regression, Decision Tree, Random Forest, SVM, Gradient Boosting, K- Neighbours,

Principal Component Analysis., Independent Component Analysis and different types of Deep learning ResNet modules and additionally we can attach the attention network with the deep learning modelbecause we got the better accuracy in deep learning concerning machine learning. These algorithms were evaluated based on their performance in classifying individuals with AD from healthy controls, with a specific emphasis on assessing their effectiveness within the best model. Overall, these algorithms exhibited high accuracy rates, demonstrating their potential as valuable tools in clinical practice. Out of this model, several accuracies have been obtained from different models. Out of all these we have got highest accuracy in the SVM model, which is almost 96% as shown in Fig 2.



Random Forest Algorithm | Decision Tree Algorithm | Logistic Regression Algorithm

Support Vector Machine Algorithm | Gradient Boost | K-means Clustering Algorithm

Principal Component Analysis | Independent Component Analysis | Adaboost

Fig 2. AUC Graph for Different Machine Learning Algorithms

Additionally, the attention network of prevalence allowed for a deeper analysis of the disease's characteristics within deep learning. Firstly, we have used several machine learning models for the detection of AD.

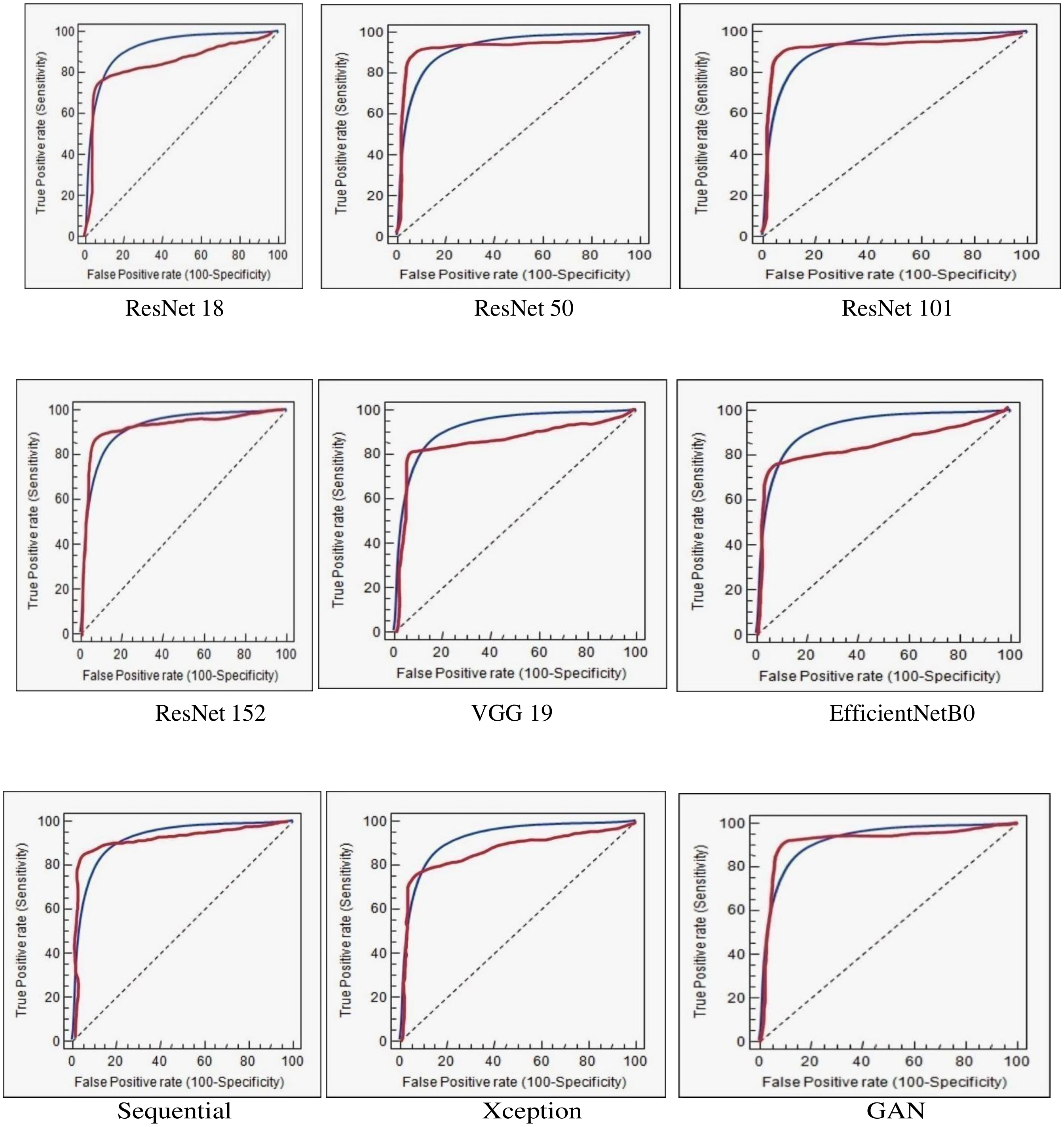
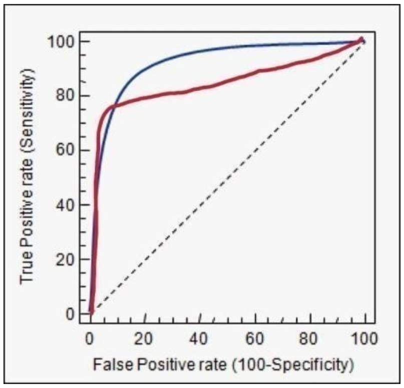
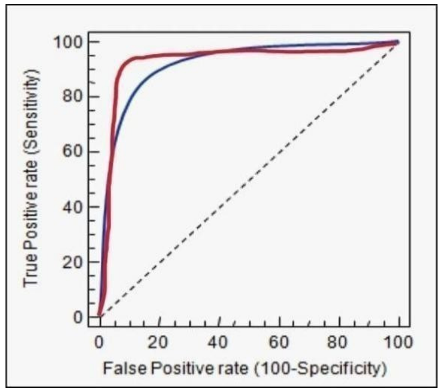


Fig 3. AUC Graph for Different Deep Learning Algorithms

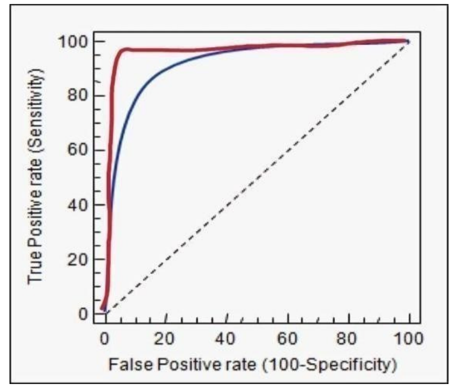
Then we transferred our dataset in several deep learning models after that we got the highest accuracy on the ResNet 101 model (98%) shown in Fig 3. For better accuracy, we used the attention mechanism along with the deep learning model and we got our highest accuracy (almost 99%) as shown in Fig 4, in this combination of the ECANet attention module with the ResNet 101 module. In Fig 4 the Deep learning with attention model we get the highest accuracy respect of other models.



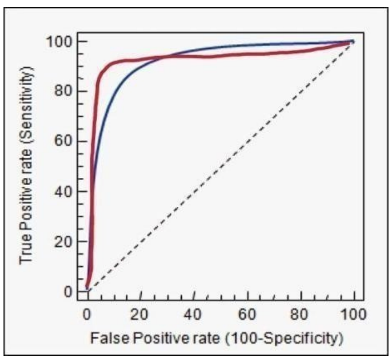
ResNet18



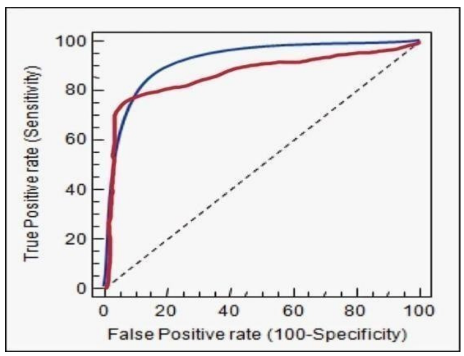
ResNet50



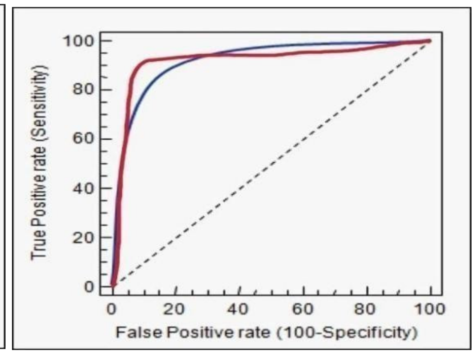
ResNet101



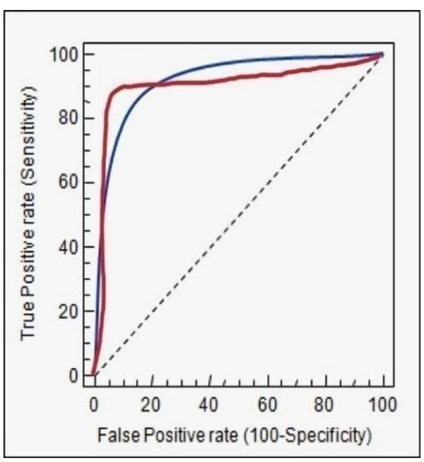
ResNet 152



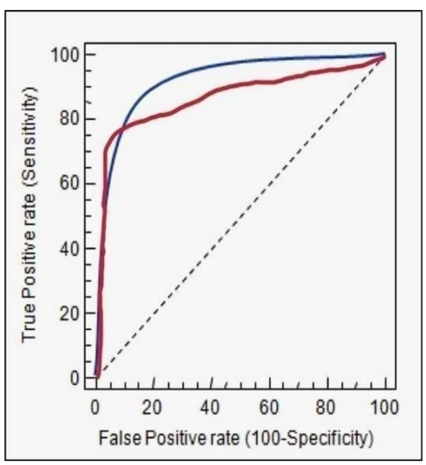
VGG19



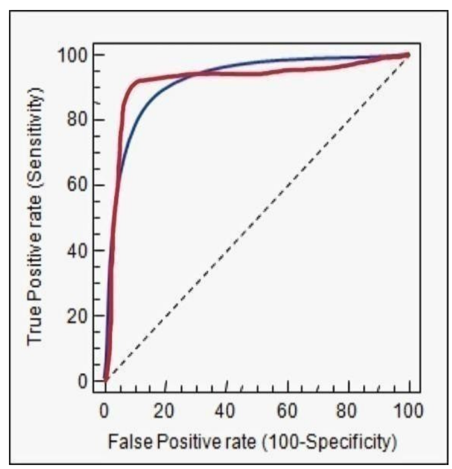
EfficientNetB0



Sequential



Xception



GAN

Fig 4. AUC Graph for Different Deep Learning With ECANet Attention Algorithms

Conclusions

This research study is completely based on the comparative study between the machine learning model, deep learning model and deep learning models with the attention network which is to be applied to the ADNI datasets of Alzheimer's disease taken from the Kaggle website. Much research has been done before using machine learning using different optimizers and batch sizes. Each of the models gives different ac curacies.

Alzheimer's is a neurodegenerative and non-curable disease. It is needed to detect early to reduce the severity of the disease in a patient. After our research observation, we can believe that using medical imaging which we have used MRI images of ADNI datasets can early diagnose the disease. The highest accuracy gives the efficiency of the model. Our work compares the models with each other and gives results with better and better accuracy. We have used the CNN model with the deep learning model,

where we have achieved our best accuracy at the Resnet 101 model which is almost 99% if we look forward to machine learning and deep learning models we have got high ac-curacies but not more than the attention network model. So this model can cover up the limitations of the early detection of medical technical issuesrather than the other two models.

References

Beheshti, H., Demirel, A & Initiativ, D. N. (2013). Feature- ranking-based Alzheimer’s disease classification from structural MRI, *Magnetic resonance imaging*, vol. 34, pp. 252-263 .

Frisoni, G. B.,Fox, N. C., Jack Jr, C. R., Scheltens, P.& Thompson P. M. (2010). The clinical use of structural MRI in Alzheimer disease, *Nature Reviews Neurology*, vol. 6, p. 67.

Yang, J.,Zhao, Y. Q.& ChanJ. C.-W. (2017).: Learning and transferring deep joint spectral–spatial features for hyperspectral classification, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 4729-4742 .

Kloppel, S., Stonnington, C., Chu, C., Draganski, B., Scahill, R., Rohrer, J. N., Ashburner, G.& Frackowiak, R. (2008). Automatic classification of MR scans in Alzheimer’s disease, *Brain*, vol. 131, no.3, pp.681–689.

Lao, Z., Shen, D., Xue, Z., Karacali, B., Resnick, S. M.& Davatzikos (2018). C.: Morphological classificationof brains via high-dimensional shape transformations and machine learning methods, *Neuroimage*, vol.21, pp. 46-57.

Weiner, M. W. (2013). The Alzheimer’s disease neuroimaging initiative: A review of project works published since its inception, *Alzheimer’s Dementia*, vol. 9,no. 5, pp. e111–e194.

Schroeder, M. L., Stein, T., Maslowski, N.& Neumann J. (2016). Neural correlates of Alzheimer’s disease and mild cognitive impairment: A systematic and quantitative meta-analysis involving 1351 patients, *Neuroimage*, vol. 10, pp. 102-108.

Yang, C.H.& Yu C.J. (2017). Alzheimer's disease classification algorithm based on fusion principal component analysis and linear discriminant analysis, *J Xiamen Univ NatS ci*, vol. 56(2), pp. 226–230.

Falahati, F., Westman, E.& Simmons, A.(2014) Multivariate data analysis and machine l learning in Alzheimer's disease with a focus on structural magnetic resonance imaging, *J A Disease* vol. 41, no.3, pp. 685–708 .

Adrien, P., & Giovanni M. (2015). Predicting Alzheimer’s disease: a neuroimaging study with 3D convolutional neural networks, arXiv:1502.02506 [cs.CV] .

Moradi, E., Pepe, A., Gaser, C., Huttunen, H.& Tohka J. (2015). Machine learning framework for early MRI based Alzheimer’s conversion prediction in MCI subjects, *Neuroimage*, vol. 104, pp. 398– 412.

Cheng, B., Liu, M., Shen, D., Li, Z., Zhang, D.& Initiative A. D. N. (2017). Multi-Domain Transfer Learning for Early Diagnosis of Alzheimer’s Disease, *Neuroinformatics*, vol. 15, no. 2, pp. 115.

Jude Hemanth, D., Vijila, C. K. S., Selvakumar, A. I.& Anitha J. (2018). Performance Improved IterationFree Artificial Neural Networks for Abnormal Magnetic Resonance Brain Image Classification, *Neurocomputing*, vol. 130, pp. 98–107.

Korolev, S.,Safiullin, A.,Belyaev, M. & Dodonova Y. (2015). Residual and plain convolutional neuralnetworks for 3D brain MRI classification, *Proc. - Int. Symp. Biomed. Imaging*, pp. 835–838.

Lecun, Y., Bottou, L., Bengio, Y.& Haffner, P. (1998). Gradient-based learning applied to document recognition, *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324 .

Chollet& F.: Keras (2015). [Online]. Available: https://keras.io/.

Fargo, k. (1998). Alzheimer's Association Report: 2014 Alzheimer's disease facts and figures, *Alzheimer's Dement*., vol. 10, no. 2, pp. e47–e92 .

Shahbaz, M., Ali, S., Guergachi, A., Niazi, A.& Umer (2019). A.: Classification of Alzheimer's disease using machine learning techniques, *DATA*, pp. 296–303.

Liu, M., Cheng, D., Wang, K., Wang, Y.&Initiative (2018). A. S. D. N.: Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis, *Neuroin formatics*, vol. 16, pp. 295 308.

Jain, R., Jain, N., Aggarwal, A.& Hemanth (2018). D. J.: Convolutional neural network based disease classification from magnetic resonance brain images, *Cognitive Systems Research*,vol. 57, pp. 147-159 .

Alzheimer's Association (2019). Alzheimer's disease facts and figures: Includes a special report on the financial and personal benefits of early diagnosis.

Simon, F., Deo, M., Selvam, V.&Babu R. (2016). An overview of machine learning and its applications, *International Journal of Electrical Sciences & Engineering*, vol. 1, pp.22–24 (2016).

Reitz, C., Brayne, C.&Mayeux, R. (2016). Epidemiology of Alzheimer's disease, *Nature Reviews Neurology*, vol.7, No. 3, pp.137–152

Ghazi, M.M., Nielsen, M., Pai, A., Cardoso, M.J., Modat, M., Ourselin.&S., Sørensen L. .(2019). Training recurrent neural networks robust to incomplete data: Application to Alzheimer's's disease progression modeling, *Med. Image Anal.*,Vol. 53, pp.39–46 .

Rahman, F., Salam &M. Islam et al., (2013). Alzheimer's disease-an update, Bangladesh *Journal of Neuroscience*, vol. 28, No. 1, pp.52–58 .

Helaly, H. A., Badawy, M., Haikal A. Y. (2022). Deep Learning Approach for Early Detection of Alzheimer's Disease, vol. 14 pp.1711-1727.