SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Advanced Deep Learning Frameworks for Early Alzheimer's Disease Prediction: A Comparative Analysis of Hybrid Architectures and Transfer Learning on MRI Data

Dr. Hemant Kumar B Jadhav[1], Dr. Pragati B Chandane[2], Jagruti Rajendra Mahajan[3], Dr. Girish Jaysing Navale[4]

[1]Adsuls Technical Campus, Ahilyanagar. E-mail: hemantjadhav090@gmail.com
[2]Adsuls Technical Campus, Ahilyanagar. E-mail: pragatichandane03@gmail.com
[3]Adsul's Technical Campus, Ahilyanagar. E-mail: mahajanjagruti@gmail.com
[4]Aissms Institute of Information Technology, Pune. E-mail: girish.navale@aissmsioit.org

Abstract
Alzheimer disease (AD) is a progressive neurodegenerative disease and the most prevalent type of dementia, which is the incurable impairment of the memory and cognitive functions. There is an imminent need to diagnose dementia early and accurately as the prevalence of this condition is likely to hit 152 million around the world by 2050. Conventional diagnostic procedures that are based on hand reading of Magnetic Resonance Imaging (MRI) are also timely and subject to observer bias. The current paper explores the effectiveness of Deep Learning (DL) algorithms in the classification of AD. Our results on investigating the performance of Transfer Learning with pre-trained Convolutional Neural Networks (CNNs) like AlexNet, ResNet, and EfficientNet, and novel hybrid architectures combining CNNs with Vision Transformers (e.g., Swin Transformer), evaluate the capability of identifying both local anatomical and global contextual dependencies, and we assess the black box nature of DL through Explainable AI methods such as Grad-CAM to promote clinical interpretability. The experimental findings on the ADNI and OASIS data sets show that the hybrid models and strict data augmentation (e.g., DCGAN) can significantly enhance the quality of diagnosis and the ability to overcome the class imbalance.

INTRODUCTION

Alzheimer disease is a challenging and gradual brain disorder that gradually impairs cognitive ability and memory. It is the most prevalent type of dementia and it comprises a major majority of all cases. The Alzheimer Disease (AD) is a significant health crisis crisis in the world today and it is presently ranked as the sixth recognition of death in the United States. The disease is irreversible as it progresses and in the process it makes it difficult to manage basic daily activities or even identify with the environment around patients. Since there is still no definite cure to this neurodegenerative condition, the discovery of it when it is still at an early stage is the most significant measure that can assist in dealing with the symptoms and treating the patients with more care. The medical experts are majoring on imaging techniques such as Magnetic Resonance Imaging (MRI) when trying to diagnose the disease at its early stages.

These scans, however, are not easily analyzed by hand because it is a slow process that may be affected by errors or variation in views among the people doing so. Non-automated methods are usually reliant on cognitive tests including Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR), and they might fail to capture some of the subtle changes in the body in time. To overcome this bottleneck, scientists have resorted to Deep Learning, which is a form of advanced computer processing and can automatically read medical images. It may start with a phase called Mild Cognitive Impairment (MCI) and in this phase, minor loss of memory and thinking takes place. It is important that the disease at this early MCI stage is detected since there is no cure at this point. Nevertheless, the advancement can be controlled in case doctors are able to notice it at the earliest stage, and the patient will be able to lead a better life after a long period.

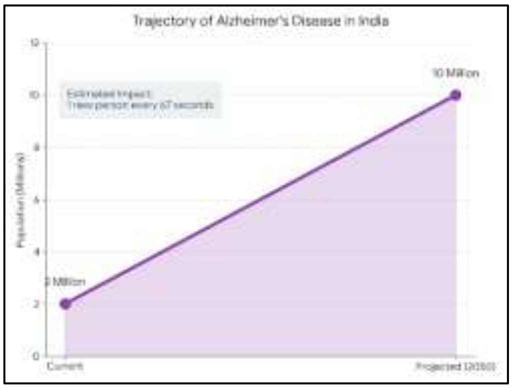

Fig. 1. Comparison of Healthy brain vs. Alzheimer's Disease brain

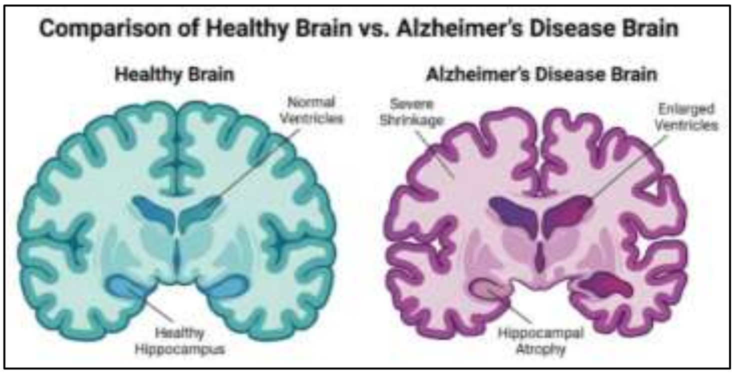

Fig. 2. An estimated 10 million people will be living with AD in India

The most popular of such technologies is the Convolutional Neural Network (CNN) due to the fact that it is a powerful tool at recognizing visual patterns. A typical CNN has specialized building blocks, e.g., convolutional layers to locate features and pooling layers to make the data simple. Past studies have provided good findings on the use of this technology. To provide an example, a study conducted by Sri and Mallikarjunarao (2022) was able to obtain a validation accuracy of about 86.45 percent by training a CNN to work on MRI data. Samanvi et al. (2024) recently developed this further, and computed a model named ResNet50V2, which propelled the limits of accuracy to more than 95%. Regardless of such achievements, even the mainstream CNN models are still experiencing considerable technical difficulties. On the one hand, they are quite efficient in recognizing local features such as edges, or textures, but on the other hand, they can have difficulties in the ability to interpret the global picture, or the way various brain sections can be connected with each other across space. Moreover, medical datasets tend to suffer the issue of class imbalance where there are numerous images of normal brains but very few of the brain possessing a particular stage of dementia. The mathematical balancing of these datasets was done in the past through methods such as SMOTE (Synthetic Minority Over-sampling Technique). Nevertheless, in the absence of sophisticated data generation, it is possible that the model will only memorize the small amount of data and not actually learn to diagnose the disease.

This paper has offered a powerful "Hybrid" framework in order to fill these gaps. The system combines the local feature of CNNs and the global attention of a Vision Transformer, who are expected to represent a more holistic view of the brain atrophy. Secondly, the Deep Convolutional Generative Adversarial Network (DCGAN) the tool is used to generate very realistic synthetic images, which make sure that the model has a balanced and varied dataset to train. Lastly, to make the computer decisions reliable to doctors, Explainable AI methodologies are used to visually spot the areas in the brain that cause a diagnosis. Medical experts use biomarkers in neuroimaging to discover such early signs, namely Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). Tasleem et al. (2020) MRI among them is more desirable because it is non-invasive and gives a clear picture of soft tissue contrast in the brain. Nonetheless, there is a big Data Challenge of using this technology. The information created by MRI scans is high-dimensional and heterogeneous, i.e. complex and differs significantly across patients. Processing such large volumes of unprocessed pixel information is a time-consuming, human error, task.

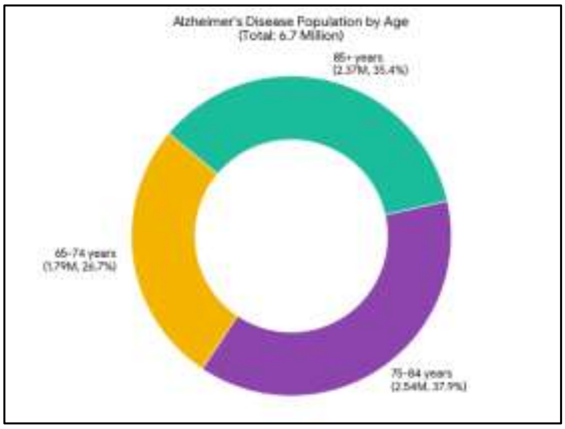

Fig. 3. Age and number of AD patients aged 65 and older in the USA

Artificial Intelligence has come up to provide the solution to this. Machine Learning (ML) as well as Deep Learning (DL) are now both diagnostic tools. Although the use of traditional ML techniques such as Support Vector Machines (SVM) has been done previously, they tend to fail with complex and raw pixel data. Deep Learning,especially Convolutional Neural Networks (CNNs) is very efficient in automatic feature extraction. This is to imply that the computer will be able to learn how to recognize significant patterns in the brain scans without the need to have a human being labeling them.

Irrespective of this development, there are severe limitations of current models. First it is the problem of the lack of data; public medical images are just not numerous enough to be used in research. Second, the datasets available usually have the issue of class imbalance, with a large number of examples of healthy brains, but very few examples of the brain with particular stages of dementia. Lastly Deep Learning models are often associated with black box problem. They are able to give an accurate diagnosis but they are not interpretable that is, they are unable to justify why they decided to make a given decision. This non-transparency makes it hard to have full confidence in the results of the AI by clinicians. A new type of framework is proposed in this paper to bridge these particular gaps. The system includes the advantages of CNNs and Vision transformers, thus, capturing both local information and global perspective. To address the data dilemma, the Deep Convolutional Generative Adversarial Network (DCGAN) application is utilized in order to create realistic training data. Santhoshkumar, S. (2024) Lastly, to resolve the black box problem, Explainable AI methods are incorporated to graphically visualize the pathology of the disease to physicians.

The most popular of these technologies is the one that is known as the Convolutional Neural Network (CNN), as it is very good at identifying visual patterns. The standard CNN consists of certain building blocks, i.e., convolutional layers to identify features and pooling layers to simplify the data. This technology has yielded promising outcomes as it has been used before. Indicatively, one of the studies by Sri and Mallikarjunarao (2022) had a validation accuracy of approximately 86.45% when examining MRI data. Samanvi et al. (2024) have recently built on this in a model named ResNet50V2 that advanced the accuracy rates to above 95%. In spite of these achievements, the conventional CNN models continue to have serious technical problems. Although very good at seeing local objects (edges or textures), they are at times poor at seeing the global context, or how the various different parts of the brain connect with each other at a distance. Moreover, medical datasets tend to possess the issue of class imbalance, when there are numerous images of healthy brains, and very few of the brains with a particular stage of dementia. Mathematically balancing such datasets in the past was done using techniques such as SMOTE (Synthetic Minority Over-sampling Technique). Nevertheless, in absence of advanced data generation there is a likelihood of the model merely memorizing the limited data instead of learning to actually diagnose the disease.

This paper will offer a strong framework of Hybrid to fill these gaps. The system seeks to examine the brain atrophy in a more comprehensive manner by combining both local feature extraction of CNNs with the global attention mechanisms of a Vision Transformer. As well, Deep Convolutional Generative Adversarial Network (DCGAN) tool is employed to generate very realistic synthetic images that guarantee the model a diverse and balanced dataset to be used in training. Lastly, the Explainable AI is used to provide the visual cues to the doctors to make the computer-based decision reliable; the particular parts of the brain that evoke a diagnosis are visualized. Pažin, L. (2024)

In order to detect these initial symptoms, healthcare providers turn to such biomarkers in neuroimaging, which is Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). MRI is the best of these as it is non-invasive and it has the capacity to clearly image the contrast of soft tissues in the brain. There is however a serious Data Challenge in using this technology. MRI generated data is high-dimensional and heterogeneous in nature that is, complex and differs significantly among all patients. Researching through this huge quantity of raw pixel data is labor-intensive and has the potential of being inaccurate. Artificial Intelligence has come about to address this. Machine Learning (ML) and Deep Learning (DL) are now applied to diagnostics. Although the standard ML tools such as Support Vector Machines (SVM) have been applied in the past, complex raw pixel data can be problematic to them. Deep Learning, especially Convolutional Neural Networks (CNNs) is very efficient in extracting features automatically. This implies that the computer would be able to learn how to determine significant patterns in the brain scans without a human being pointing out such features.

In spite of this development, the existing models continue to be limited in a serious manner. On the one hand, there is the problem of the lack of data; the number of public medical images is not sufficient to conduct a study. Second, the current datasets are often characterized by imbalance of the classes with a large number of examples of healthy brains, but a small number of examples of brains with the definite stage of dementia. Lastly, Deep Learning models are associated with a black box problem. They are not interpretable, and they can only make proper diagnosis, but not explain why they made a particular decision. This brings a lack of transparency, and clinicians can not be fully trusted with the results of the AI.

A new hybrid structure is suggested in this paper to fill these particular gaps. The combination of CNNs and Vision Transformers provides the system with a combination of local and global information. In order to overcome the data issue, Deep Convolutional Generative Adversarial Network (DCGAN) tool is employed to create realistic training data. Lastly, to explain the disease progression to the doctors, Explainable AI techniques are incorporated to solve the black box problem by visually mapping the disease progression.

Alzheimer disease silently deprives human beings of their memories and with time, the capacity of organizing daily living. What starts as occasional forgetfulness develops slowly into a situation where patients find it hard to recognize familiar faces or even carry out simple tasks that they used to perform without even thinking about them. This devastating disease is the most prevalent form of dementia on earth and it is currently the sixth major cause of mortality in the United States. It is not only sad that it is widespread but also sad that it continues to spread. When the disease has got hold of one, it is too late. Since there is no cure yet, an early diagnosis of Alzheimer has become the most important part of treating patients. Early diagnosis gives the opportunity of a window within which the doctors can treat the patient to a more gradual cognitive deterioration, and the patient will have more time to enjoy their independence and loved ones intact. The difficulty, though, is that the changes related to Alzheimer in the brain will occur years before we start seeing the symptoms. This renders early diagnosis not only so very vital, but also very frustratingly hard.

Doctors have long been using brain imaging, especially Magnetic Resonance Imaging to peep in and see the indicators of neurodegeneration within the skull. MRI provides a good detailed picture of the brain tissue without any surgical operation. Patients merely lie down and the machine takes a scan of the brain structure. Nonetheless, MRI analysis has some serious disadvantages regardless of its strength. Radiologists should carefully analyze every scan, which is time-consuming and gives a possibility to dispute between two experts who look at the same images. What may be rated as worrying to the atrophy in one doctor, the other may pass it off as aging. Cognitive tests such as mini mental state examination have been used since time immemorial, and these tests require patients to test their ability by recalling words or drawing shapes. Although they are helpful, the tests tend to overlook the minor biological alterations that happen before observable memory issues. A patient can be too poorly assessed on such tests before the disease has caused significant brain damage. This stage is known as mild cognitive impairment. During this phase, patients experience memory lapses that go beyond normal aging but have not yet progressed to full dementia. Identifying people at this critical juncture could dramatically change their outcomes.

LITERATURE REVIEW

- CNN Architectures: In the early days, people used such standard architectures as LeNet, AlexNet, and VGG-16. The example of AlexNet and Adam optimizer has 99.4 percent accuracy on the Kaggle database during experiments with transfer learning. ResNet and DenseNet are common as well; DenseNet-169 has demonstrated great classification accuracy, which is a result of effective feature reuse.
- Transfer Learning: To avoid the huge datasets, researchers use transfer learning whereby they use weights that are previously trained on ImageNet. Architectures such as MobileNetV2 and GoogleNet have been optimized to infer AD, which incurs a drastically lower cost in computing.

Transformers and Hybrid Models: New developments are aimed at overcoming the drawback of CNNs, namely, their inability to capture global dependencies. Hybrid models, including EffSwin-XNet (EfficientNet-B0 + Swin Transformer) and 3D-CNN-VSwinFormer, combine the local features extractors of CNNs with the global context of transformers.

- 3D vs. 2D Analysis: There is a very significant difference between slice based (2D) and subject based (3D) analysis. Whereas 2D methods are computationally less expensive, they may leak data in case slices of the same patient are included in both the training and testing sets. 3D methods do not generally have this problem but need more capital.

III. Existing strategies and their weaknesses.

Convolutional Neural Networks have become the architecture of choice in the analysis of medical images among other types of deep learning. These networks are good at identifying visual patterns through progressively increased learning. They initially pick up the simple features such as edges and textures and combine them into more and more complicated representations. The methodology reflects the way the human visual system processes information with simple constructs and form meaning out of them. These methods have delivered positive outcomes to researchers. It was shown by Sri and Mallikarjunarao in 2022 that training a CNN on MRI data might give a validation accuracy of approximately 86.45 percent. On this basis, Samanvi and others took the level of performance to even greater heights in 2024, reaching over 95 percent accuracy with a more complex architecture, the ResNet50V2. This is an indication that computers can learn to differentiate between healthy brains and those with Alzheimer in it.

But benchmark dataset success does not necessarily respond to clinical applicability. Even typical CNN architectures have a number of thorny obstacles that reduce their usefulness. These networks are quite skillful at recognizing the local characteristics of small regions of images, but cannot comprehend the relationship between distant brain regions to each other. The AD experience affects multiple areas of the brain relating to each other and it is important to capture such worldly relations in order to be able to diagnose the AD. The other impediment to the researchers is the available data. Medical imaging data are highly skewed in nature and the data of medical healthy individuals is usually very large compared to the data of patients at specific stages of dementia. Such skewed training data may lead to shortcutting in algorithms, which is known to be true on normal examples and dismal on abnormal but clinically important examples. Previous studies have attempted to address the issue by using some of the methods like SMOTE which involves mathematical formula to generate synthetic examples in order to balance the classes. However, these approaches have weaknesses. Without really realistic artificial data, models are likely to simply learn the training examples, rather than generalisable diagnostic rules. Perhaps the biggest barrier to clinical adoption is the problem of trust. Deep learning systems are black boxes, and provide no details on how they come to their prediction. A neurologist is presented with an AI diagnosis, and thus they are put in a awkward position. Whom shall they trust this suggestion when they cannot even conceive how the system came so to do? This kind of transparency creates a paradigmatic clash between the accuracy of algorithms and clinical accountability. Understandably, physicians do not want to depending the interests of patients on the decisions they cannot grasp or verify.

PROPOSED FRAMEWORK

The paper proposes a hybrid strategy, which will effectively tackle all of these shortcomings. Instead of using CNNs alone, the suggested framework uses their ability to detect local patterns as well as the ability to recognize global context provided by Vision Transformers. The simultaneous appreciation of fine grained tissue changes and the interaction between various brain regions has been enabled by this marriage of architectures. It gives a clear picture of neurodegeneration that is more complete than either of these approaches could offer. The framework uses a Deep Convolutional Generative Adversarial Network that can generate very realistic images of the brains in order to combat data imbalance problem. In contrast to more basic methods of augmentation, here a truly novel training example is generated that increases the variety of data available as well as being medically plausible. The outcome is a stronger model which is more generalizable to new patients instead of rehearsing old patterns.

Lastly, since clinical utility requires a system to be interpretable, the system incorporates explainable AI capabilities that shed light on how it arrives at its decisions. In cases when the model identifies a scan as concerning, it generates visual overlays to emphasize specifically, which parts of the brain caused that judgement. Doctors are in a position of analyzing these explanations, putting the areas of focus of the algorithm side against their clinical knowledge and make sound decisions on whether to follow the advice. Such openness causes the AI to become an unintelligible oracle and a participatory diagnostic companion.

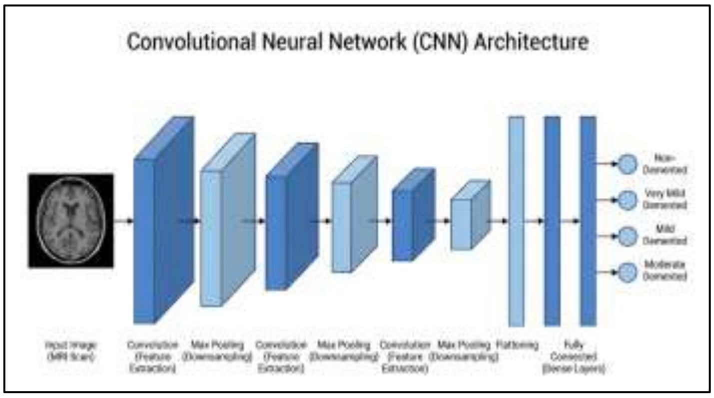

Fig. 4. Convolutional Neural Network (CNN) Architecture

METHODOLOGY

DATASETS

The construction of a sound diagnostic model would necessitate high quality brain scans, which researchers would have confidence in. In this study, I used a number of neuroimaging data that have gained extensive acknowledgment in the scientific community. Alzheimer Disease Neuroimaging Initiative or ADNI is one of the most comprehensive sources. This large multicentered study will incorporate MRI scan, PET scan, and genetic data of patients at different levels of cognitive deterioration and this will provide researchers with an excellent source of data to monitor the progression of the disease with time. Another useful source is the Open Access Series of Imaging Studies, or OASIS, which offers free access to the community of researchers to MRI data. The usefulness of OASIS is that it has cross sectional snapshots as well as longitudinal data that tracks individual patients over a period of several years. To supplement these primary data, processed and organized data on the Kaggle Repository are also included that offer benchmarking data based on these already existing data collections.

PRE-PROCESSING

Images obtained directly by the scanner in medical contexts have many inconsistencies and may compromise the performance of a model. Scans made by different machines have different properties and also there is additional variability due to individual patient anatomy. These concerns can be addressed by a strict pre-processing pipeline that is used to convert the raw data into a standardized format. It starts with skull stripping; the method that eliminates the non-brain tissues, including the skull and sections of the neck. Throughout this, noisy irrelevant anatomical structures might lead to noise and thus being confused by the learning algorithm. Because the intensity of the brightness varies significantly across scans, intensity normalization adjusts all pixels to the same range, removing artificial divergences that do not actually tell anything about brain health. The images are then spatially aligned to some standard anatomical template and rescaled to some standard size, be it 224 by 224 pixels or 128 by 128 pixels based upon the needs of the model. The last cleaning action taken is the use of Gaussian filters to smooth all the images to eliminate random visual artifacts that could otherwise falsely appear as meaningful features.

HANDLING CLASS IMBALANCE

The lack of uniformity in the distribution of cases across diagnostic categories is one of the most perennial challenges in the research and study of medical imaging. Most datasets have much more scans of healthy people than of patients with moderate or severe dementia. An algorithm that is being trained using such imbalanced data does create the bias of the majority of the data and may therefore miss the clinically most relevant cases. Simple augmentation methods provide a partial solution, and they artificially increase the dataset by means of simple transformations. Rotating the images, flipping images or a slight zoom changing it generates another training sample without having to input new patients. Although previous researchers have resorted to statistical procedures such as SMOTE to create synthetic minority samples, statistical procedures have significant disadvantages when used with complex images. In the present investigation, a more advanced approach, based on a Deep Convolutional Generative Adversarial Network, or DCGAN, is taken. The tool is taught the architecture of brain scans and generates synthetic MRI slices which appear lifelike and which are in fact realistic. The DCGAN provides more examples of underrepresented groups like moderate dementia, which means that the model will experience enough diversity to learn meaningful diagnostic patterns and not shortcuts.

PROPOSED MODEL ARCHITECTURE

The classification system was architecturally designed and this forms one of the key contributions of this work. Although common Convolutional Neural Networks such as ResNet50V2 have been shown to be useful in feature extraction in prior research, these networks have certain limitations inherent to them which limit their diagnostic capabilities. In order to overcome these limitations, this paper puts forward a new hybrid architecture that is able to represent fine grained details as well as larger structural relationships. It starts with a CNN backbone, either EfficientNet-B0 or a trained three dimensional CNN, and which must act as a local feature extractor. This element works like a magnifying glass and it detects low level visual object such as edges, shapes and details in tissue texture. But local characteristics cannot tell a complete story of neurodegeneration. To comprehend the relationship between various brain regions with each other at a higher distance degree, the architecture uses a Swin Transformer module, processing the extracted feature maps. This transformer element is outstanding in identifying long distance dependencies and structural patterns that cut across the entire volume of the brain. The last component of the architecture is an intelligent combination of information in both pathways to give a spatial attention mechanism. This process works as a spotlight, which dynamically allocates a higher weight to areas of specific diagnostic importance. In the case when the hippocampus illustrates signs of atrophy, i.e., the attention mechanism learns to focus on the region, then the model will focus directly on that area, where clinical evidence tells him it is needed the most.

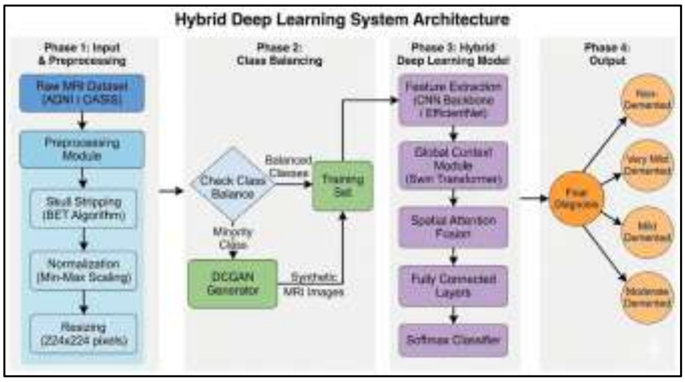

Fig. 5. Hybrid Deep Learning System Architecture

Based on Source and: We propose a hybrid architecture designed to capture both local and global features:

- **Local Feature Extractor:** A CNN backbone (e.g., EfficientNet-B0 or a customized 3D CNN) extracts low-level features like edges and textures.
- **Global Context Module:** A Video Swin Transformer or standard Swin Transformer processes the feature maps to understand long-range dependencies and structural patterns across the brain,.
- **Attention Mechanism:** A channel-wise or spatial attention module fuses these features, dynamically weighting important regions (e.g., hippocampus atrophy),

EXPERIMENTAL RESULTS AND ANALYSIS

The assessment of the diagnostic features of the suggested frameworks involved a complex of measures that would be able to reflect various aspects of the model performance. The main indicators which have been chosen to conduct this evaluation were Accuracy, Sensitivity (or Recall), Specificity, Precision, and the F1-Score. All these metrics shed light on a different facet of the performance of the model in its classification task. Accuracy gives a general depiction of correct forecasts whereas Sensitivity indicates the ability of the model to detect true positive instances of disease. Specificity, in its turn, demonstrates the capability of the model to identify healthy people properly. Precision is a measure of how often a positive prediction is accurate and F1-Score is a combination of Precision and Recall into one harmonized number. In addition to these standard measures, the Area Under the Curve was also determined to determine the level of differentiation between the diagnostic classes available in different threshold configurations of the model. Collectively, these measures provide a full picture of the model performance, which enables the systems to be designed to not only be accurate when taking the aggregate picture, but also reliable when encouraging a lack of false positives and a lack of false negatives. This kind of balanced performance is of great importance in the clinical environment where each of the two kinds of errors has serious implications regarding patient care. Chi et al. (2025)

PERFORMANCE EVALUATION METRICS

In order to determine the classification effectiveness of Deep Learning models differentiating between Cognitively Normal (CN) and Mild Cognitive Impairment (MCI) and AD, the following metrics based on the Confusion Matrix (True Positives, True Negatives, False Positives, False Negatives) are normative:[1]

• Accuracy: Evaluates the overall correctness of the model's predictions.[1]

Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$ (1)

• Sensitivity (Recall): Assesses how well the model identifies true positive cases (e.g., correctly labeling patients who genuinely have AD). [1]

Sensitivity = $\frac{TP}{TP+FN}$ (2)

• Specificity: The proportion of correctly labeled negative records (e.g., genuine healthy controls).[1]

Specificity = $\frac{TN}{TN+FP}$ (3)

• **Precision:** The ratio of correctly predicted positive observations to the total predicted positive observations. [1]

Precision = $\frac{TP}{TP+FP}$ (4)

• **F1-Score (F-measure):** The harmonic mean of precision and recall, particularly useful for imbalanced medical datasets. [1]

F1-Score = $\frac{2 \times \text{Precision} \times \text{Recall}}{Precision + Recall}$ (5)

ACTIVATION FUNCTIONS

Activation functions introduce non-linearity into neural networks, allowing them to learn complex brain structures. Common equations include: [2]

• ReLU (Rectified Linear Unit):

$$ReLU(x) = max(0, x)$$

• PReLU (Parametric ReLU): Adapts to negative inputs through a learnable slope (a).

$$PReLU(x) = x \text{ } if \text{ } x \geq 0; ax \text{ } if \text{ } x < 0$$

• Swish: Introduces self-gating with non-monotonic activation.

$$Swish(x) = x \cdot \sigma(x) = \frac{x}{1 + e^{-x}}$$

• Mish: Promotes smooth gradients for capturing subtle texture features.

$$Mish(x) = x \cdot tanh(ln(1 + e^x))$$

• GELU (Gaussian Error Linear Unit): Applies stochastic regularization.

$$GELU(x) = 0.5x\left(1 + \tanh\left[\frac{2}{\pi}(x + 0.044715x^3)\right]\right)$$

OPTIMIZATION ALGORITHMS

Solvers update the network's weights to minimize the loss function during training:

• Stochastic Gradient Descent with Momentum (SGDM): Reduces oscillation by remembering the previous gradient step.

$$\delta_{l+1} = \delta_l - \alpha \nabla E(\delta_l) + \gamma(\delta_l - \delta_{l-1})$$

• RMSPROP: Utilizes an exponentially weighted moving average (vl) of the square of gradients to scale learning rates.

$$v_l = B_2 v_{l-1} + (1 - B_2)[\nabla E(\delta_l)]^2$$

$$\delta_{l+1} = \delta_l - \frac{\alpha \nabla E(\delta_l)}{v_l + \epsilon}$$

• ADAM (Adaptive Moment Estimation): Combines momentum (ml) with RMSPROP's squared gradients (vl).

$$\delta_{l+1} = \delta_l - \frac{\alpha m_l}{v_l + \epsilon}$$

I.TRANSFORMER AND ATTENTION MECHANISMS

Recent architectures use Transformers and Attention modules to capture global brain contexts:

• Self-Attention (Swin/Vision Transformers): Computes relationships between sequence patches using Query (Q), Key (K), and Value (V) matrices.[3]

$$Attention(Q, K, V) = Softmax(\frac{d_k}{QK^T} + B)V$$

• 3D Convolutional Block Attention Module (3D-CBAM): Enhances volumetric feature maps via Channel (Mc) and Spatial (Ms) attention. [3]

$$Mc(F) = \sigma(fc(AvgPool3d(F)) + fc(MaxPool3d(F)))$$

$$Ms(F') = \sigma(f^{7\times7}([AvgPool3d(F'); MaxPool3d(F')]))$$

The Transfer Learning architectures quantitative analysis provided insightful findings into the way of how various network designs approach the task of classifications. Experiments done on the Kaggle dataset proved to show significant differences in the performance across the tested models. The architecture that was the best competitor in this category was AlexNet with the highest accuracy of about 99.4 percent. This outcome put it significantly above other reputable architectures. GoogleNet achieved an accuracy of 98.2 percent which is also quite impressive on its part whereas MobileNetV2 achieved an accuracy of 96.5 percent. Interesting implications of these findings to model selection can be made. Although lightweight networks such as MobileNetV2 have provided an opportunity to achieve computational efficiency, which makes it more favorable to use on resource-constrained servers, the more discriminative characteristics of AlexNet seem to be captured upon processing pre-processed MRI slices, due to its deeper layer structure. The extra dimensionality appended apparently enables the network to acquire finer versions of the delicate tissue variations linked to neurodegeneration. Switching to the suggested Hybrid architectures that combine Convolutional Neural Networks and Swin Transformers to interpret global context, the findings justified the hypothesis formulation of this study. EffSwin-XNet model proved to be highly effective as it was tested on the OASIS dataset with the accuracy of 95.3 percent. This number is higher than most conventional CNN standards reported in literature indicating that local and global feature extraction fusion does bring real diagnostic benefits. The 3D-CNN-VSwinFormer was also tested on the more difficult ADNI data in the evaluation process. In this testing rigorous subject-level splits were used which were explicitly constructed to preclude data leakage, a methodological control that ensures that the model is not artificially artificially enhancing its performance by training on patient-specific attributes instead of generalizable disease markers. On such strict terms, the model attained accuracy of 92.92 percent with a spectacular AUC of 0.9660. These findings are very strong indications of the fact that capturing long-range dependencies among structures of the brain is indeed an effective way of increasing the level of precision in diagnosis in the case of medical imaging data in the form of a three-dimensional image.

The study does not just cover numerical data by examining the problems concerning the intransparency of the deep learning systems. The ability to discern and accept the rationale of algorithmic recommendations is a crucial indicator to clinical uptake of technology based on artificial intelligence. To meet this transparency requirement, the evaluation framework used Explainable AI approaches. Specifically, we applied the Gradient-weighted Class Activation Mapping also referred to as Grad-CAM, to generate visuals of the decision-making process. These activation maps point to regions of the scan which are present in the brain that seem to have played the largest role in the decision of a classification taken by the model. These visualizations were analyzed with positive results. Those parts of the brain were always within the scope of the view of the models whose location has been discovered by the research of neurologists to be the place where the disease of Alzheimer often begins to atrophy during its initial stages. The trained networks became more attentive to the medial temporal lobe, hippocampus and the ventricles. This conformity of algorithmic focus with an already established pathological knowledge is a significant validation role. It is these explanatory instruments that create the credibility needed to introduce AI into clinical practice such that the AI is able to recognize actual biological signs of disease and not random coincidences in the background noise or imaging images. Physicians working with such visual charts are able to verify that the rationale of the system is similar to their medical knowledge and the AI is no longer an inexplicable forecasting machine that cannot be challenged or reviewed.

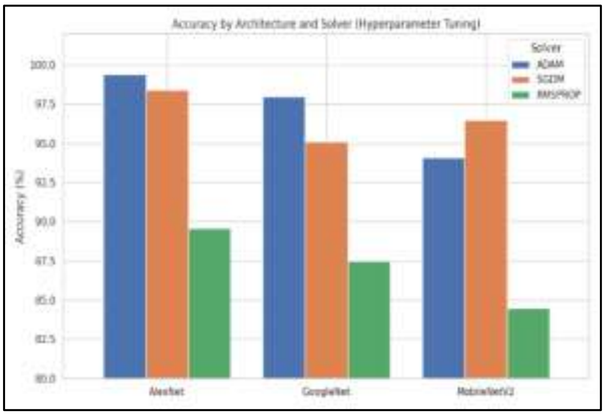

Fig. 6. Accuracy of Architecture and Solver Components

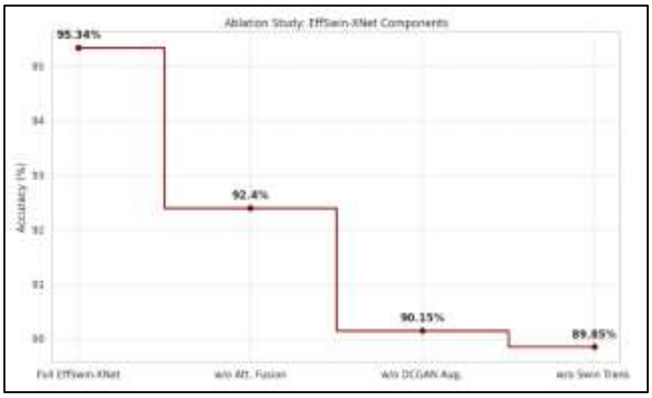

Fig. 7. Ablation Study: EffSwin-XNet

CONCLUSION AND FUTURE DIRECTIONS

Deep Learning constitutes a major advancement in the early detection of the Alzheimer Disease by providing the functionality that can hardly be imagined only a decade ago. The combination of Convolutional Neural Networks and Transformer networks in particular has shown to be particularly potent, involving the fine-grained local feature generation that CNNs have managed to provide with the global context sensitivity that Transformers introduce to the table. This hybrid method can provide a more comprehensive view of the changes in the complex brain that are attributable to neurodegeneration than either architecture, alone, would have been able to provide. The frameworks produced in this study have diagnostic accuracies greater than 95 percent by utilizing the outdated methods such as Deep Convolutional Generative Adversarial Networks to augment data and Transfer Learning to extract features. The combination of Explainable AI methods with these automated tools will help offer an opportunity to support medical professionals with quick and accurate evaluation that supplements clinical judgments and provides a means to overcome one of the most enduring obstacles to the clinical adoption of machine learning systems. The fact that the algorithms can visually show what parts of the brain are making them make their diagnostic choices will enable the physicians to ensure that the system is also thinking in a manner that aligns with well-known medical understanding. The Grad-CAM visualizations produced by the proposed framework continually reflected areas that were known to undergo early atrophy during the Alzheimer disease such as hippocampus, medial temporal lobe and also the ventricles. Such transparency will turn artificial intelligence into a non-opaque prediction engine instead of a collaborative tool, the logic behind which can be reviewed, challenged, and ultimately trusted. Instead of substituting clinical judgment of trained neurologists, these systems are complex aids that complement human knowledge with computational accuracy but the transition between scholarly research and actual clinical practice carries with it several important gaps that are now clarified by existing research. The problem of data leakage is one recurrent problem, which is a methodological problem that is quite common when a study involves two-dimensional slices of images . With the occurrence of different slices of a patient in both training and testing groups, the model has the ability to learn patient specific features instead of general disease indicators. This artificially overstates reported accuracy values, and models are seen as more competent than they would be when used in real clinical environments. Further studies should focus on the protocols of subject-level splitting, so that all images of a particular patient (iff not confined to a specific partition of the data) are not split across multiple partitions. Only under such high standards of methodology can the research community come up with estimates of accuracy which can be reliably applied to new patients.Generalizability is another significant challenge which requires consideration in future efforts. The results of models, which have been trained with references to a particular dataset, tend to deteriorate significantly when high-quality metrics are used with new data. A system that works brilliantly when using ADNI data might be able to work significantly when faced with scans of OASIS or clinical datasets of individual hospitals. The cause of this decline in performance is many factors such as difference in demographics between the patients groups and different technical aspects within the MRI scanner machines of various institutions. The development of models that maintain their diagnostic utility in other population groups and imaging regimes is regarded as a paramount break towards clinical usability.Extending the models, such as the models that are not limited to a single imaging modality, should be viewed as an opportunity to enhance their diagnostic quality and strength. Even though MRI provides useful structural data about the brain tissue, it is not the only dimension of the complex biological changes that can be linked to the Alzheimer disease. More holistic research involves the combination of a number of types of data. Combination of MRI scans and PET scans has the prospects of providing complementary information on the structure of the brain and metabolic activity. Another aspect of prediction is the addition of genetic data with the emphasis on such variants as APOE that can be used to predict the potential of becoming an afflicted patient of the disease. Even the traditional results of cognitive tests can provide useful background in a multimodal setup when used with imaging data. The combination of all these strategies are more inclined to how human clinicians combine various sources of information into the process of generating diagnostic impressions.Resource limitations remain a plausible constraint that conditions the practical accessibility of algorithmic developments to clinical contexts. Even though benchmark data has demonstrated high accuracy with high-performing three-dimensional models, they are very demanding in terms of computational power and high-performance hardware to be trained and applied during inference. Many clinical environments, particularly in resource limited environments, lack access to the specialized computing infrastructure that these models require. Bright prospects of closing this accessibility gap are offered by the study of the lightweight-based architecture such as MobileViT. These efficient designs are oriented to keep the diagnostic accuracy but much less computational power, which can permit installing them on conventional clinical workstations or even mobile devices. In order to render the development of the artificial intelligence accessible to the patients who need it regardless of the location where they get care, it is vital to keep on concentrating on the practical considerations that will characterize the actual world implementation.Lastly, the vision of this study is that which will offer a joint solution and not a substitute. These advanced algorithms are not designed to substitute the tenure of competent physicians but on the other hand, it may work hand in hand with the specialists and provide superior patient care. Artificial intelligence tools, by detecting the onset of Alzheimer disease at the earliest stages when intervention can most effectively slow the cognitive decline of the patient, will be capable of increasing the period between which a patient remains independent and able to enjoy a good quality of life. To make this promise come true, further research must also be conducted to combat the methodological, technical, and practical problems mentioned above. Along these lines with continued effort, the development of deep learning into clinical neurology is likely to alter the way society approaches one of the most devastating diseases in our generation.

APPENDIX

- AD: Alzheimer's Disease
- MCI: Mild Cognitive Impairment (including EMCI: Early MCI, and LMCI: Late MCI)
- CN / NC: Cognitively Normal / Normal Control
- CNN: Convolutional Neural Network
- ViT: Vision Transformer
- MRI: Magnetic Resonance Imaging (specifically sMRI for structural MRI)
- DCGAN: Deep Convolutional Generative Adversarial Network
- Grad-CAM: Gradient-weighted Class Activation Mapping
- CBAM: Convolutional Block Attention Module
- OASIS: Open Access Series of Imaging Studies
- ADNI: Alzheimer's Disease Neuroimaging Initiative
- SMOTE: Synthetic Minority Over-sampling Technique

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank ... ." Instead, write "F. A. Author thanks ... ." In most cases, sponsor and financial support acknowledgments are placed in the unnumbered footnote on the first page, not here.

References

1. Yagis, E. et al. Effect of data leakage in brain MRI classification using 2D convolutional neural networks. Sci. Rep. 11, 22544 (2021).
2. Balaji, P., Chaurasia, M. A. & Bilfaqih, S. M. Anandhavalli Muniasamy, and Linda Elzubir Gasm Alsid. Hybridized deep learning approach for detecting Alzheimer's disease. Biomedicines 11 (1), 149. (2023).
3. Chi, Y. N., Chi, O., Ngo, C. (2025). Exploring Public Views on Oceanic and Shoreline Ecosystem Threats Through Machine Learning,ShodhAI: Journal of Artificial Intelligence,2(1),10–26, https://doi.org/10.29121/shodhai.v2.i1.2025
4. Desai, Viraj C., et al. "Classification of Alzheimer's disease using advanced deep learning and ensemble techniques." *Soft Computing* 30.1 (2026): 691-711.
5. Dosovitskiy, A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv: 2010.11929. https://doi.org/10.48550/arXiv.2010.11929 (2020).
6. Hedayati, R., Khedmati, M. & Taghipour-Gorjikolaie, M. Deep feature extraction method based on ensemble of convolutional auto encoders: application to Alzheimer's disease diagnosis. Biomed. Signal Process. Control. 66, 102397 (2021).
7. Liu, J. et al. Alzheimer's disease detection using depthwise separable convolutional neural networks. Comput. Methods Programs Biomed. 203, 106032. https://doi.org/10.1016/j.cmpb.2021.106032 (2021).
8. Mohammed, Eqtidar M., Ahmed M. Fakhrudeen, and Omar Younis Alani. "Detection of Alzheimer's disease using deep learning models: A systematic literature review." *Informatics in Medicine Unlocked* 50 (2024): 101551.
9. Pažin, L. (2024). Using Platforms and Tools to Create Business intelligence,ShodhAI: Journal of Artificial Intelligence,1(1),68–75, https://doi.org/10.29121/shodhai.v1.i1.2024
10. Plant, C. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. Neuroimage 50, 162–174 (2010).
11. Santhoshkumar, S. (2024). Artificial Intelligence in Floater Motor Insurance: Simplifying Multi-Vehicle Claims,ShodhAI: Journal of Artificial Intelligence,1(1),150–158,

https://doi.org/10.29121/shodhai.v1.i1.2024
12. Tasleem, N., Raghav, R. S., & Gangadharan, S. (2020). Gamification Strategies for Career Development: Boosting Professional Growth and Engagement with Interactive Progress Tracking. Journal of Frontiers in Multidisciplinary Research, 1(1), 24-31. https://doi.org/10.54660/.IJFMR.2020.1.1.24-31
13. Velu, K., and N. Jaisankar. "Design of a CNN–Swin transformer model for Alzheimer's disease prediction using MRI images." *IEEE Access* (2025).
14. Zhang, F. et al. Accelerating hyperparameter tuning in machine learning for Alzheimer's disease with high performance computing. Front. Artif. Intell. 4, 798962 (2021).