International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Transformative Insights: Swin Transformer's Impact on Alzheimer's Detection

Athira R. S. [1*], J. Charles [2]

Submitted: 29/01/2024 **Revised**: 07/03/2024 **Accepted**: 15/03/2024

Abstract: An irreversible neurological illness, Alzheimer's disease (AD) causes a rapid loss in mental function. It is primarily caused by a widespread degeneration of the brain cell and can strike at any age. Due to the progression of Alzheimer's disease, there is a great clinical, societal, and financial demand for early detection of AD. With neural network topologies, deep learning (DL) has become a potent tool for disease categorization, automatically deriving complex patterns from medical data. Deep learning models are useful for finding complex diseases because of their capacity to recognize minute patterns and connections. The major purpose of this work is to overcome the AD diagnosis challenges via MRI images of the brain by constructing a unique deep learning-based model employing a Swin Transformer. The Kaggle repository offers a well-curated and extensive dataset that is an invaluable tool for doing in-depth analysis and obtaining significant insights. The performance of the Swin Transformer study is demonstrated by its outstanding overall accuracy of 95.12%. Furthermore, a thorough investigation of the performance of current methods at the class level has been carried out by analyzing the parameters of the confusion matrix.

Keywords: *Alzheimer's disease, Neurodegenerative diseases, Patch Embedding, Swin Transformer.*

1. Introduction

A class of clinically and neurotically broad disorders known as neurodegenerative diseases(NDDs) [1] impact the neurological system and result in ongoing degradation of the nervous system. Three NDDs are typified by a long-term, steady deterioration in cognitive function that is likely to affect daily functioning. As illustrated in Figure 1, the most prevalent neurological diseases are Multiple System Atrophy (MSA) [2], Parkinson's disease (PD) [3], Alzheimer's disease, Dementia with Lewy bodies (DLB) [4], Huntington's disease (HD) [5], and Progressive Supranuclear Palsy (PSP) [6].

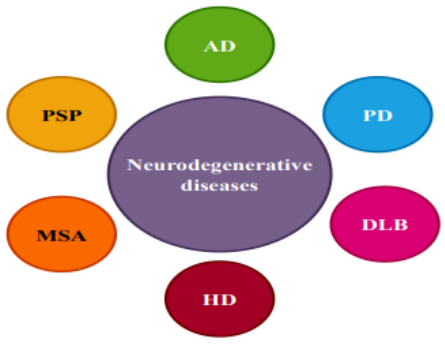

Fig. 1. Most severe neurodegenerative diseases

Growing interest has been shown in creating automatic Computer Aided Diagnosis (CAD) systems that can accurately identify patients with a variety of ailments in the current technological era. Artificial Intelligence is used by CAD systems at different stages of disease diagnosis. Artificial Intelligence is the capacity exhibited by computers and devices built for specialized tasks. These intelligent systems are typically designed to supplement and enhance human intelligence's capacity for making decisions.

Approximately 55 million cases of Alzheimer's disease are known to exist worldwide, with low- and middle-income countries accounting for more than 60% of cases [7]. Every year, over 10 million new cases are reported. It is a degenerative neurological condition that results in memory loss and intellectual disability due to brain cell destruction. AD bears Dr. Alois Alzheimer saw some amazing changes in the brain tissues of a patient in 1906 who had a rare psychological disorder characterized by memory loss, communication challenges, and strange mannerisms.

The part of brain responsible for memory and language is primary affected by AD. AD patients thus suffer from memory loss, disorientation, and issues with speaking, reading, and writing [8]. A notable reduction in memory, language, motor skill [9], reasoning, and other cognitive abilities together with a sharp drop in cortical measurement, are among the early indicators of AD. According to researchers, AD affects 30% of adults over 85 and 5% of those over 65. Figure 2 illustrates the expectation that over 150 million people will receive an AD diagnosis by 2050. As of right now, there is no drug that can stop AD from getting worse.

[1] *Department of Computer Application, Noorul Islam Centre for Higher Education, Tamil Nadu, India.*
.[2] *Department of Computer Application, Noorul Islam Centre for Higher Education, Tamil Nadu, India.*
**Athira.r.s@outlook.com*
* Corresponding Author Email: author@email.com

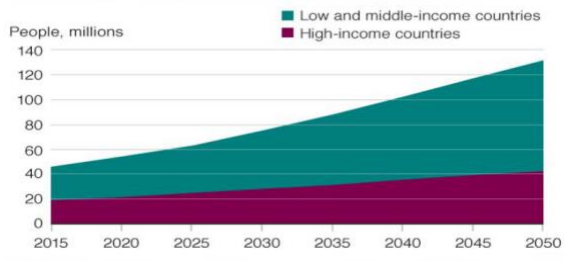

Fig. 2. Growth of AD over the years

It is quickly rising the list of causes of death, especially in the elderly, right behind heart disease and cancer. In AD brain cells are irreversibly destroyed and cannot heal. As neuro degeneration causes the brain to shrink, as seen in Figure 3. After observing the first major symptoms, Alzheimer's patients typically live for just eight years on average.

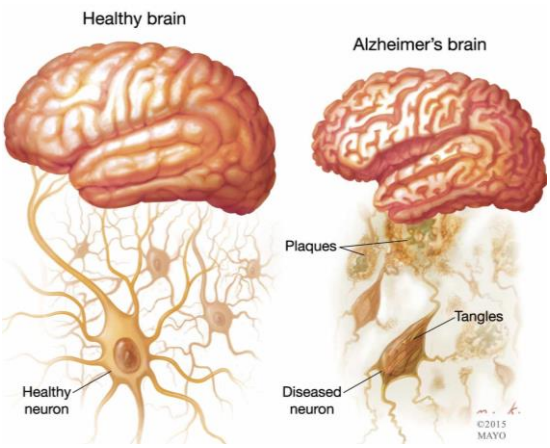

Fig. 3. Healthy brain vs Alzheimer's brain

1.1. Types of Alzheimer's

Three basic categories can be used to roughly categorize the course of AD progression:

- *Mild Alzheimer's*: People are able to function independently and require less help during this gentle stage. They could be able to work, drive, and engage in their favorite pastimes [10].
- *Moderate Alzheimer's*: Person will find it more difficult to perform daily tasks during this phase. Furthermore, individuals typically begin to move around without realizing where they are since they struggle to pinpoint their current location. The longest stage could last for a few of years [11]. The course of the affected person's disease will determine how much care they require.
- *Severe Alzheimer's*: This is the final stage of the disease, and the patients have significant symptoms. They are unable to respond, communicate, or control their own motions [12]. Their personality has completely altered, their memory and cognitive function have deteriorated, and they need a lot of help to perform daily duties.

1.2. Motivation

AD is so complex and requires early detection, it presents a significant challenge to clinicians. The advancement of the disease cannot always be precisely predicted or visualized using conventional methods. Unconventional and computationally intensive methods, such as DL, are required to solve this problem. In order to provide individualized and proactive treatment plans, medical professionals are depending more and more on DL approaches. Patients' lives are improved by this change in how health economists evaluate prices and physicians choose treatments. Our objective is to overcome the limitations in limited training data and achieve better prediction accuracy in the context of AD classification. We seek to improve the classification performance using MRI scans by investigating the use of DL approaches. Our objective is to overcome the drawbacks of traditional methods and create a reliable and effective AD classification model using the swin transformer. Our goal is to increase the precision of AD diagnosis, facilitate early detection, and improve the decision-making process by utilizing DL method.

2. Background of the Study

2.1. Related Works

Hazarika et al [13] discussed DL models and the outcomes of their application to AD classification. Alzheimer's disease Neuro imaging Initiative, dataset is the source of all brain MRI data is used in this study. The DenseNet-121 model attained an average performance rate of 88.78%. However, the DenseNet model has a drawback in that it is computationally slower than many of the other models presented because it makes extensive use of convolutional procedures. Al-Adhaileh [14] presented two deep neural methods for the identification and categorization of AD in their research using AlexNet and Restnet50. This study employed brain MRI data that were gathered from the Kaggle website to assess and test the model. To effectively classify AD, a convolutional neural network (CNN) technique along with Restnet50 and AlexNet transfer learning models were used.

A deep ensemble learning framework was presented by An et al. [15] with the goal of utilizing DL algorithms to integrate data from several sources and access the "wisdom of experts." In this work at each layer of the three-layer framework, a DL approach is used for AD classification is called Deep Learning. Authors tested DE Learning's performance with six representative ensemble learning techniques using the clinical data.

An automated DL-based approach was proposed by Raees et al. [16] to predict AD using a sizable MRI dataset of both healthy and diseased individuals. The 111 participants in the database were categorized into three groups: Normal, AD, and mild cognitive impairment (MCI). Deep Neural Network (DNN) models and other classification techniques such as Support Vector Machines (SVM) were examined. About 80–90% of AD predictions were made with good accuracy using DL techniques. For the diagnosis of AD and the stage-based categorization of medical imagery was put forth by Lokesh et al. [17]. The EfficientNet architecture serves as the foundation for this framework. There are several categories for each of the four AD periods. This classification utilizing a number of method, at first by utilizing previously trained models and transfer learning.

Tajammal et al. [18] introduced a method for Alzheimer's diagnose using functional MRI pictures from the publically accessible dataset. The scientific community has employed a variety of deep-learning models to automatically identify Alzheimer's patients. A limited amount of study is being done on the multiclass categorization of AD up to six different stages. These includes the binary classification of patient scans into MCI and AD stages. There are two steps to this investigation. Custom CNN is used in the initial phase to classify and in the second step, a subject's scan is multiclass classified into one of the six stages of using a combination of Custom CNN and several DL models.

Priyatama et al. [19] use transfer learning (VGG16 and VGG19) and convolutional neural networks (CNN) to classify images of Alzheimer's. This study classified images of AD into four groups that are acknowledged by medical professionals with accuracy this work. The study's output includes a number of assessment criteria. The dataset demonstrate that the algorithm is capable of categorizing ADMRIs into four groups that are well-known to medical professionals. For multi-label classification of six Alzheimer's stages Alorf et al. [20] presented techniques using rs-fMRI and DL. The suggested model uses two DL techniques—Brain Connectivity Graph Convolutional Network and Stacked Sparse Auto encoder—to handle the problem by obtaining the brain's functional connectivity networks using rs-fMRI data. The k-fold cross-validation method was used to evaluate the performance of the model. For multi-label classification, the average accuracy achieved was 77.13% and 84.03%, respectively.

A deep ensemble model with transfer learning was suggested by Mujahid et al. [21] to identify AD patients from a multiclass dataset. The authors employed adaptive synthetic oversampling to correct for the significant imbalance in the Alzheimer disease dataset between the classes. When we compared the performance of another ensemble model to that of an individual EfficientNet-B2, the results were 1.46% better. Three two-dimensional (2D) CNNs are used by Angkoso et al[22] to gather discriminating data on all possible planes using multiplane technique. The approach is a fusion strategy designed to address the shortcomings of multi-view approaches based on shape. The three largest 3-planes are chosen by the framework to represent entire 3D-MRI objects. Numerous tests have demonstrated that the suggested approach can perform better than the 2DCNN method, which only makes use of one plane. The authors provide a new, simple, and effective method for identifying 3D objects by fully utilizing 2DCNN. According to the conducted trials, the approach obtains good precision for AD 93%, MCI 91%, and NC 95%, and multiclass AD-MCI-NC accuracy of 93%.

2.2. Research Gap

Current DL models are not very good at capturing minor differences in brain structure, it is challenging for them to categorize AD accurately. The creation of sophisticated DL models that are capable of capturing minute details and fine-tuning anatomical variations in the brain is crucial. Most of the current techniques for classifying AD do not sufficiently account for the spatial and dimensional connections that occur across different brain regions. While image biomarkers have shown promise in discovering potential neuro imaging-based biomarkers for differentiating people with AD, new research indicates that neuro image alone may not be adequate for classifying patients with the disease. Therefore, it is essential to create more effective DL techniques for classification of Alzheimer's by integrating the pertinent biomarkers

3. Materials and methods

Alzheimer's disease (AD) involves the progressive deterioration of brain cells due to underlying brain abnormalities, leading to the gradual loss of cognitive function. The most impacted parts of the brain are usually those that deal with memory, reasoning, and personality. The dataset for this investigation is MRI, a potent and adaptable medical imaging technique. DL models are useful for finding complex illness patterns because of their capacity to recognize minute patterns and connections.

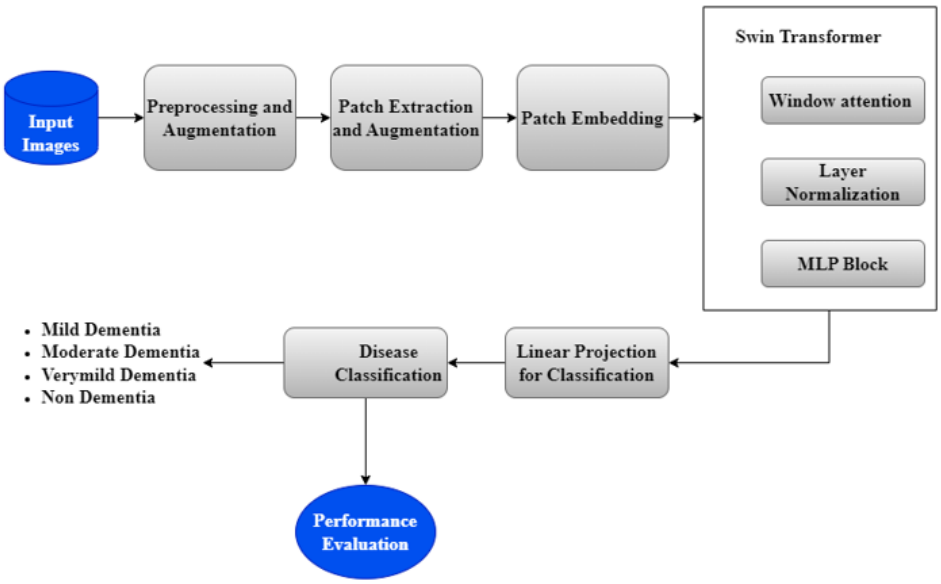

Fig. 4.Block diagram of the proposed model

This system starts by loading images from a given directory, preprocessing them, and then using TensorFlow and Keras to build a Swin Transformer model. Block diagram for the proposed model is illustrates in figure 4. Layer normalization, the MLP block, and the window attention mechanism define the architecture of the Swin Transformer. To capture complicated patterns in the data, the model incorporates multi-layer perceptron (MLPs), patch extraction and patch embedding with attention processes.

3.1. Architecture Overview

Patch splitting module divides an RGB input image into non-overlapping patches in the first stage. Every patch is regarded as a "token," with its feature being defined as the RGB values of its raw pixels concatenated. Use a 4 × 4 patch size in the proposed work, each patch's feature dimension is 4 × 4 × 3 = 48. Each block of Swin Transformer generated intermediate feature contain semantic information of various scale, will be handled as the output feature at various encoding booster level. The first patch-merging input feature size $\frac{h}{2} * \frac{w}{2} * c$, since the block of transformer does not alter the feature scale. Patch merging is employed to down-sample and retrieve the multi-scale feature at various levels and the feature is first divided into two-by-two patch square. Then, the patches at the same location in each square are stacked in the channel dimension and rearranged in order to create $\frac{h}{4} * \frac{w}{4} * 2c$ feature."After the second merging the feature size will also be down to $\frac{h}{8} * \frac{w}{8} * 4c$." The swin Transformer pyramid's spatial dimensions continuously increase and the image size gradually decreases with patch merging.

3.2. Patch Embedding

In the embedding layer, initial low-level features are encoded using patch embedding and position embedding. Figure 5 showcasing the patch embedding operation. The patch size is 2 × 2 and the RGB input picture is $I \in rh * w * c$. Initially, create $\frac{h}{2} * \frac{w}{2}$ patch by image patching; each patch has a 4 × c dimension.

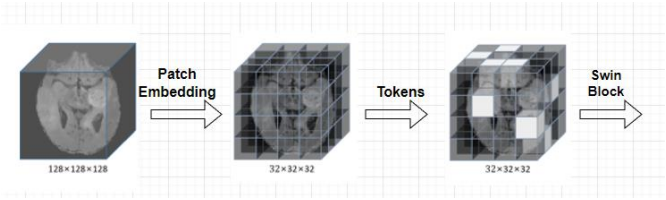

Fig. 5. Block diagram for patch embedding

In order to satisfy the feature fusion process that follows, project the patches to the original dimension c using a fully-connected layer. Consequently, following the embedding layer, the feature size is $\frac{h}{2} * \frac{w}{2} * c$. The initial layer of patch merging concatenates the features of every set of two-by-two contiguous patches. The resulting 4C-dimensional concatenated features were applied to the linear layer. As a result, the number of tokens is decreased and the output dimension is adjusted to 2C.After feature transformation, Swin Transformer blocks are applied, maintaining a resolution of $\frac{h}{8} \times \frac{w}{8}$. The process is performed twice, with output of $\frac{h}{16} \times \frac{w}{16}$ and $\frac{h}{32} \times \frac{w}{32}$. Together, these phases result in a hierarchical representation that has feature map resolutions comparable to that of conventional convolutional networks. We express the correlation between features using the self-attention technique. The following describes the self-attention mechanism:

$$Self_{attention}(q,k,v) = Softmax\left(\frac{qk^t}{\sqrt{d}}\right) \quad (1)$$

where the query, key, and value are represented by the matrices q, k, and v respectively. A scaling factor called d is employed to prevent the softmax (·) function's gradient from disappearing. Because it operates on an entire feature rather than just a small portion of receptive field, this type of attention mechanism is known as global attention. Self-attention is evident in the fact that, the input image is divided into several patches in the majority of implementations in order to shorten the sequences before they pass through the Transformer block." The linear projection of the same embedded word vector is q, v and k."

3.3. Swin Transformer

After the embedding the feature will go through the Transformer. The swin transformer block is illustrated in figure 6. Each encoder includes self-attention (SW-MSA), windowing _MSA or shifted-MSA, layer normalization (LN), and multi _layer perception (MLP). It is designed to use the shifted window design in order to implement the self-attention.

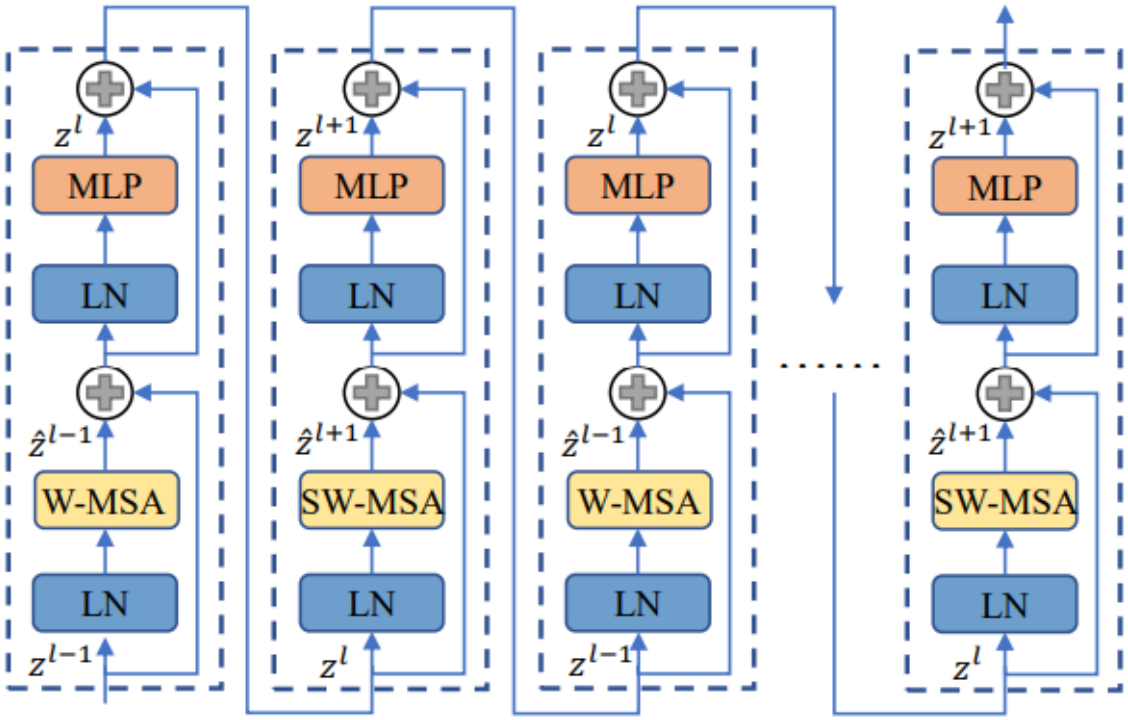

Fig. 6.TheSwin transformer block

Because it is a sequence-to-sequence model, the 2D features are flattened to sequences prior to input. The output sequences are scaled to two dimensional features to function as the merging layer's input. The following formula can be used to characterize the feature processing in the block:

$$\hat{Y}^i = Windowing\_MSA\left(LN\left(Y^{i-1}\right)\right) + Y^{i-1}{}_{,i=1,\ldots,I} \quad (2)$$

$$Y^i = Perception\left(LN\left(\hat{Y}^i\right)\right) + \hat{Y}^i{}_{,i=1,\ldots,I} \quad (3)$$

$$\hat{Y}^{i+1} = Shifting\_MSA\left(LN\left(Y^i\right)\right) + Y^i{}_{,i=1,\ldots,I} \quad (4)$$

$$Y^{i+1} = Perception\left(LN\left(\hat{Y}^{i+1}\right)\right) + \hat{Y}^{i+1}{}_{,i=1,\ldots,I} \quad (5)$$

where $^{i,\hat{Y}}$and Y stand for the number of encoders, Windowing_MSA and MLP output sequences, respectively. The i^{th} self-attention calculation is realized by encoder containing Windowing_MSA computation, and the $i+1^{th}$ self-attention calculation is realized Shifting_MSA computation. Empirical evidence indicates that the training platform still needs a lot of memory and processing power because to the self-attention performed on every patch, even after the patching embedding. To further improve the computational complexity, we employ the shifted-window architecture to overcome this challenge. Figure 7 delves into the intricacies of the shift operation within the Swin block. The full patched image is further divided into shifted windows in a swin design based on patching, and self-attention calculation will be limited within window. Because each window contains fewer patches than the entire image, the complexity can be greatly decreased.

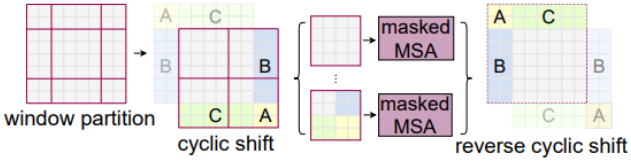

Fig. 7. Shift operation in Swin block

Window shifting increases the number of areas, which contribute to rebound complexity even if it can generate cross-window feature links. The freshly created zones, which come in different sizes, present another difficulty for the uniform scale self-attention computation. As a result, the areas that have been moved and those that have not can be joined to create a new feature that can still be divided into four 4 x 4 windows. As a result, this additional feature can be used for the standard self-attention computation, just like in the i^{th}computation. These two types of encoder pairs each realize a complete cycle of the self-attention computation with shifted windows. Table 1 details the hyper parameters utilized in the Swin Transformer model, providing a comprehensive overview of the key settings and configurations that govern the model, contributing to its performance in diverse tasks. Algorithm 1 outlines a novel approach for Alzheimer's disease classification.

Table1.Hyper parameters

Parameters	Values
Patch Size	(2,2)
Dropout rate	0.1
num_heads	8
Embed_dim	64
num_mlp	2
Window_size	7
Shift_size	1
Image_dimension	128
Learning rate	1e-3
Batch size	128
Weight decay	0.0001
Label_smoothing	0.1

Algorithm 1: Alzheimer's disease classification

Input: OASIS MRI dataset (X) with labels Mild dementia, Moderate dementia, Very mild dementia, Non-demented.

Output: Predictions of Alzheimer's disease probability for each image

Begin

1.Data Preparation

- Preprocess images: Rescale, normalize, and augment the dataset for enhanced training
- Split dataset: Divide the preprocessed dataset into X_train for training and X_test for testing.

2. Model Building and Training

- Defines the input layer of the model with a shape of (128, 128, 1).
- Defines the dimensions of the input image, patch size for patch extraction, and the embedding dimension for each patch.
- Calculate the number of patches in the x and y directions based on the given image dimension and patch size.
- Set the multi-head self-attention, shift operation, layers in the multi-layer perceptron,value in linear projections and a dropout rate for regularization.
- Defines a dense layer with 256 units with softmax activation for classification.
- Compile the model
- Using X_train, train the model.

3. Model Evaluation

- To assess the performance, evaluate the model on X_test
- Calculate metrics such as accuracy, precision, recall, and F1-score.

4. End

4. Result and Discussion

4.1. Experimental Setup

One of the obstacles encountered during implementation was the necessity of modifying the development environment to conform to the requirements. The experiments and simulations were carried out using the Python programming language. To make the datasets accessible while creating the model, they were transferred to Google Drive. Additionally, an elite configuration workstation with an Intel i7 processor, 32GB of RAM, and an 8GB Nvidia graphics processing unit was used for image processing.

4.2. Performance Evaluation

A sample of preprocessed and enhanced brain MRI images is shown in the figure 8, along with the critical steps that were done to improve the model's robustness. The images have been adjusted to a standardized pixel value range and scaled to guarantee uniformity during the preparation stage. To lessen data volatility, mean subtraction has also been used. Numerous transformations, including random rotations, flips, zooming, and shearing, have been used to enhance the dataset.

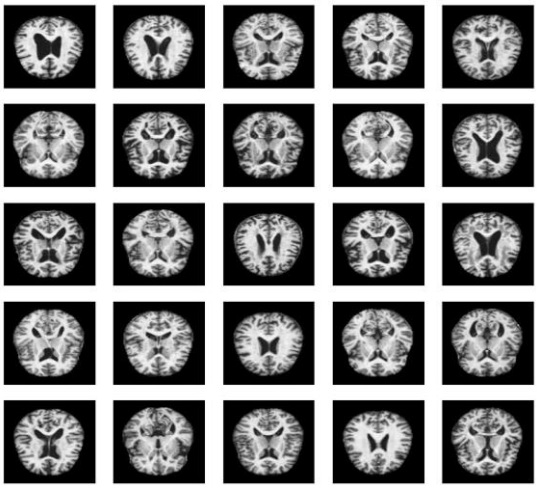

Fig. 8.Preprocessed and Augmented Images

Performance parameters are necessary to assess the model's efficiency. Table 2 presents an overview of the performance parameters. Performance metrics that emphasize a model's capacity to accurately categorize instances and reduce errors, such as precision, recall, accuracy, and F1 score, offer a quantitative evaluation of a model's classification abilities.

Model's classification report in table 3gives a thorough overview of how well it performs in various classes. It contains important metrics for every class, including F1 score, recall, accuracy, and precision.

Table 2. Performance parameters

Performance Metrics	Equation
Accuracy	$\frac{(TP+FP)}{(TP+FP+TN+FN)}$
Precision	$\frac{(TP)}{(TP+FP)}$
Recall	$\frac{(TP)}{(TP+FN)}$
F1-Score	$2\times\frac{(Precision\ \times Recall)}{(Precision+Recall)}$
where TP= True Positive, TN= True Negative, FP= False Positive, FN= False Negative.	

Table 3. Classification report

Class	Precision	Recall	F1-Score
0	92.12	93.23	92.67
1	93.34	94.43	93.88
2	94.23	92.12	93.87
3	92.01	92.83	92.47

The accuracy and loss plot provide a visual depiction of the model's training progress throughout epochs. The accuracy plot in figure 9 provides information on the convergence and learning capabilities of the model by showing the percentage of correctly identified examples. In addition, a decreasing loss indicates better model performance. The loss plot in in figure 10 illustrates the optimization process of the model and shows how well it reduces the difference between anticipated and actual values during training. By examining these plots, one can ascertain whether the model is learning correctly, identify any possible over fitting or under fitting problems, and make decisions about additional training modifications.

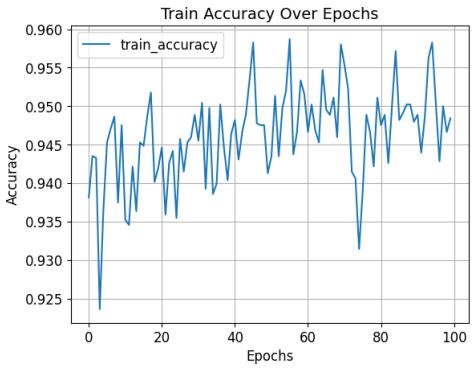

Fig. 9. The accuracy plot of the proposed system

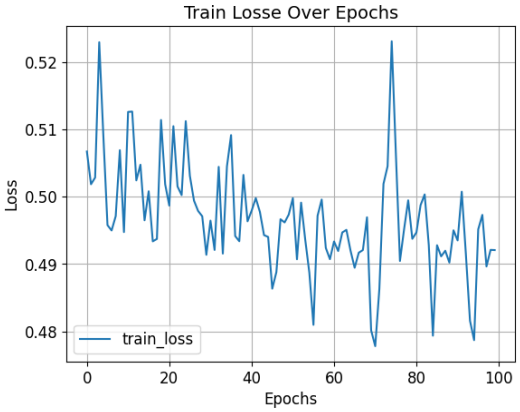

Fig. 10. The loss plot of the proposed system

The confusion matrix is an essential tool for assessing a classification model's performance. The matrix in figure 11 provides a clear synopsis that facilitates a more profound comprehension of the model's advantages and disadvantages about the accuracy and mistakes of class-wise categorization. Finding misclassification trends and directing changes can be facilitated by analyzing the confusion matrix.

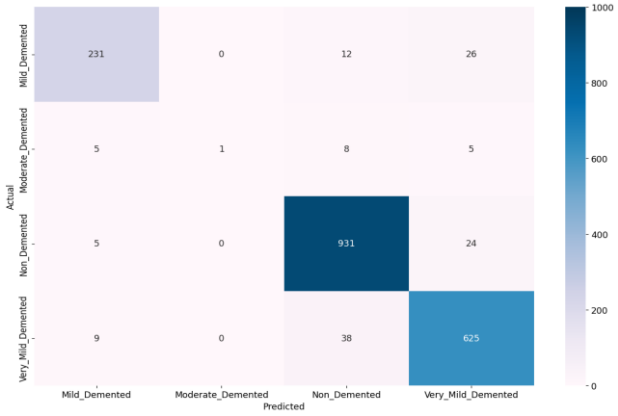

Fig. 11. Confusion matrix of the proposed system

To evaluate the performance, a comparative analysis was carried out utilizing the suggested model in conjunction with the most recent six techniques. Figure 12 displays a sample of classification outputs generated by the proposed model. It was discovered via thorough testing and analysis that the suggested model performed more accurately than the other models. The results demonstrated the superior predictive skills of the suggested model, which reached an excellent accuracy rate of 95.12%. This notable increase in accuracy serves as more evidence of the model's efficacy and promise.

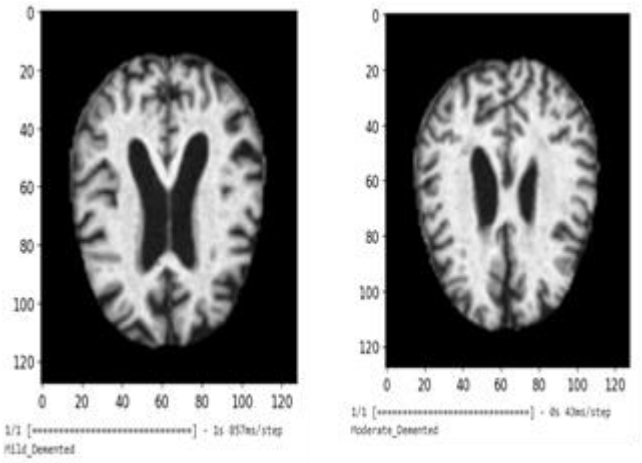

Fig. 12. Sample of Classification Outputs

Table 4. Comparison of the proposed work with existing system

Author &Reference	**Accuracy**
Raees et al [23]	87.29
Suganthe et al [24]	79.12
Liu et al [25]	83.45
Turkson et al [26]	90.15
Zhang et al [27]	0.95
Sinha et al [28]	84.79%
Proposed model	**95.12%**

5. Conclusion

With more than 20 million Alzheimer's sufferers globally, and with the number predicted to double by 2050, early diagnosis of the illness becomes crucial. Therefore, early precise diagnosis of AD is essential for prompt therapy and intervention. Image classification methods can be used to make sense of changes in the macroscopic structure of the brain, which neuro imaging reveals. The preprocessed images were fed into a Swin Transformer architecture, which successfully extracted important AD features and classified AD with an astounding accuracy of 95.12%. The architecture of Swin Transformer, with its window-based processing and self-attention mechanism, is excellent for identifying complex patterns and relationships found in medical imaging linked to Alzheimer's disease. By allowing the model to concentrate on pertinent areas of the input images, the attention mechanism of the Swin Transformer facilitates a more detailed comprehension of the neuro anatomical features linked to Alzheimer's disease. Its capacity to deliver precise, comprehensible, and flexible solutions corresponds with the increasing demand for advanced instruments in contemporary healthcare, ultimately leading to enhanced patient treatment and results. In comparison to many cutting-edge computer-aided diagnostic techniques, our suggested approach has shown superior or at least equivalent categorization performance.

References

[1] Hansson, O. (2021). Biomarkers for neurodegenerative diseases. Nature medicine, 27(6), 954-963.

[2] Fanciulli, A., Stankovic, I., Krismer, F., Seppi, K., Levin, J., &Wenning, G. K. (2019). Multiple system atrophy. International review of neurobiology, 149, 137-192.

[3] Ball, N., Teo, W. P., Chandra, S., & Chapman, J. (2019). Parkinson's disease and the environment. Frontiers in neurology, 218.

[4] Kane, J. P., Surendranathan, A., Bentley, A., Barker, S. A., Taylor, J. P., Thomas, A. J., ... & O'Brien, J. T. (2018). Clinical prevalence of Lewy body dementia. Alzheimer's research & therapy, 10, 1-8.

[5] Stoker, T. B., Mason, S. L., Greenland, J. C., Holden, S. T., Santini, H., & Barker, R. A. (2021). Huntington's disease: Diagnosis and management. Practical neurology.

[6] Giagkou, N., Höglinger, G. U., &Stamelou, M. (2019). Progressive supranuclear palsy. International review of neurobiology, 149, 49-86.

[7] Adewale, B. A., Coker, M. M., Ogunniyi, A., Kalaria, R. N., & Akinyemi, R. O. (2023). Biomarkers and Risk Assessment of Alzheimer's Disease in Low-and

Middle-Income Countries. Journal of Alzheimer's Disease, (Preprint), 1-11.

[8] Budson, A. E., & Solomon, P. R. (2021). Memory Loss, Alzheimer's Disease, and Dementia-E-Book: A Practical Guide for Clinicians. Elsevier Health Sciences.

[9] Du, Y., Krakauer, J. W., & Haith, A. M. (2022). The relationship between habits and motor skills in humans. Trends in cognitive sciences, 26(5), 371-387.

[10] Morris, J. C., Storandt, M., Miller, J. P., McKeel, D. W., Price, J. L., Rubin, E. H., & Berg, L. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. Archives of neurology, 58(3), 397-405.

[11] Doody, R. S., Thomas, R. G., Farlow, M., Iwatsubo, T., Vellas, B., Joffe, S., ... & Mohs, R. (2014). Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. New England Journal of Medicine, 370(4), 311-321.

[12] Voisin, T., &Vellas, B. (2009). Diagnosis and treatment of patients with severe Alzheimer's disease. Drugs & aging, 26, 135-144.

[13] Hazarika, R. A., Kandar, D., & Maji, A. K. (2022). An experimental analysis of different deep learning based models for Alzheimer's disease classification using brain magnetic resonance images. Journal of King Saud University-Computer and Information Sciences, 34(10), 8576-8598.

[14] Al-Adhaileh, M. H. (2022). Diagnosis and classification of Alzheimer's disease by using a convolution neural network algorithm. Soft Computing, 26(16), 7751-7762.

[15] An, N., Ding, H., Yang, J., Au, R., & Ang, T. F. (2020). Deep ensemble learning for Alzheimer's disease classification. Journal of biomedical informatics, 105, 103411.

[16] Raees, P. M., & Thomas, V. (2021, May). Automated detection of Alzheimer's Disease using Deep Learning in MRI. In Journal of Physics: Conference Series (Vol. 1921, No. 1, p. 012024). IOP Publishing.

[17] Lokesh, K., Challa, N. P., Satwik, A. S., Kiran, J. C., Rao, N. K., &Naseeba, B. (2023). Early Alzheimer's Disease Detection Using Deep Learning. EAI Endorsed Transactions on Pervasive Health and Technology, 9.

[18] Tajammal, T., Khurshid, S. K., Jaleel, A., Qayyum Wahla, S., &Ziar, R. A. (2023). Deep Learning-Based Ensembling Technique to Classify Alzheimer's Disease Stages Using Functional MRI. Journal of Healthcare Engineering, 2023.

[19] Priyatama, A., Sari, Z., &Azhar, Y. (2023). Deep Learning Implementation using Convolutional Neural Network for Alzheimer's Classification. Jurnal RESTI (RekayasaSistem dan TeknologiInformasi), 7(2), 310-317.

[20] Alorf, A., & Khan, M. U. G. (2022). Multi-label classification of Alzheimer's disease stages from resting-state fMRI-based correlation connectivity data and deep learning. Computers in Biology and Medicine, 151, 106240.

[21] Mujahid, M., Rehman, A., Alam, T., Alamri, F. S., Fati, S. M., & Saba, T. (2023). An efficient ensemble approach for Alzheimer's disease detection using an adaptive synthetic technique and deep learning. Diagnostics, 13(15), 2489.

[22] Angkoso, C. V., Agustin Tjahyaningtijas, H. P., Purnama, I., & Purnomo, M. H. (2022). Multiplane Convolutional Neural Network (Mp-CNN) for Alzheimer's Disease Classification. International Journal of Intelligent Engineering & Systems, 15(1).

[23] Raees, P. M., & Thomas, V. (2021, May). Automated detection of Alzheimer's Disease using Deep Learning in MRI. In Journal of Physics: Conference Series (Vol. 1921, No. 1, p. 012024). IOP Publishing.

[24] Suganthe, R. C., Geetha, M., Sreekanth, G. R., Gowtham, K., Deepakkumar, S., &Elango, R. (2021). Multiclass classification of Alzheimer's disease using hybrid deep convolutional neural network. NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal| NVEO, 145-153.

[25] Liu, J., Li, M., Luo, Y., Yang, S., Li, W., & Bi, Y. (2021). Alzheimer's disease detection using depthwise separable convolutional neural networks. Computer Methods and Programs in Biomedicine, 203, 106032.

[26] Turkson, R. E., Qu, H., Mawuli, C. B., &Eghan, M. J. (2021). Classification of Alzheimer's disease using deep convolutional spiking neural network. Neural Processing Letters, 53, 2649-2663.

[27] Zhang, P., Lin, S., Qiao, J., & Tu, Y. (2021). Diagnosis of Alzheimer's disease with ensemble learning classifier and 3D convolutional neural network. Sensors, 21(22), 7634.

[28] Sinha, S., Thomopoulos, S. I., Lam, P., Muir, A., & Thompson, P. M. (2021, December). Alzheimer's disease classification accuracy is improved by MRI harmonization based on attention-guided generative adversarial networks. In 17th international symposium on medical information processing and analysis (Vol. 12088, pp. 180-189). SPIE.