ISSN : 2455-9679
SJIF Impact Factor : 6.008

IJRTSM
INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"ALZHEIMER'S RECOGNITION IN HEALTHCARE BASED ON MEDICAL IMAGES VIA AI-DRIVEN DEEP LEARNING MODELS"

Sandeep Gupta [1]
SATI Vidisha
sandeepguptabashu@gmail.com

ABSTRACT

The degenerative neurological disorder known as Alzheimer's disease (AD) is typified by memory loss, diminished daily functioning, and cognitive decline. Eventually, AD leads to dementia. Early detection is challenging as symptoms appear years after the illness first manifests. Even while AD cannot be cured, its development can be significantly slowed with early identification and therapy. This study offered a platform for brain Magnetic Resonance Imaging (MRI)-based automated AD identification. In this work, a deep learning-based method for automatically classifying AD stages using T1-weighted structural MRI scans is presented. A publicly available dataset containing 5,121 grayscale brain scans was utilized, categorized into four clinical stages: Moderately Demented, Mild Demented, Very Mild Demented, and Non-Demented. Pre-processing techniques such as picture scaling, normalization, flipping, cropping, and enhancement were used to guarantee consistency and model resilience. The development of a bespoke Convolutional Neural Network (CNN) yielded 95.00% classification accuracy, 95.00% precision, 95.00% recall, and 95.00% F1-score for all classes. These findings show how well the suggested CNN model works for scalable and accurate AD detection in clinical settings.

Key Words: *AD, Structural MRI, Medical image analysis, Early diagnosis, Automated detection.*

I. INTRODUCTION

Human health inevitably deteriorates with age, making people more susceptible to illness. The brain is one area of the body that is most affected by the effects of ageing. Changes in intellectual function will occur in the brain, including memory problems and a delayed decision-making or action. [1]. Alzheimer's disease (AD) is one of the illnesses brought on by this brain condition. AD is a neurological condition that worsens over time, progressively impairing cognitive function and increasing the risk of mortality [2]. In older people, this illness is typically the cause of dementia. Most Alzheimer's patients will suffer from a number of conditions, including mood swings, personality or character changes, memory problems, and trouble connecting with others. Over the course of three to nine years, a person with AD will progressively develop a number of diseases [3][4]. Modern technology has made it feasible to use Magnetic Resonance Imaging (MRI) scans to identify AD. MRI is often the method most frequently used to diagnose and track the progression of AD [5][6][7]. This method allows for autonomous machine learning (ML)-based early AD detection. For computer-based diagnostic research, ML is crucial because it may sometimes forecast AD more accurately than medical professionals. ML is not appropriate for picture identification, despite its numerous benefits [8][9][10]. Deep learning (DL), which has better learning capabilities and is better suited to image recognition challenges, is now used in the field. In image recognition, several DL techniques like Convolutional Neural Networks (CNNs) perform better than ML techniques [11].

However, despite notable advancements in the application of ML and DL for Alzheimer's diagnosis, several limitations

ISSN : 2455-9679
SJIF Impact Factor : 6.008

still persist. Many existing models are limited to binary classification distinguishing only between Alzheimer's and non-Alzheimer's while not correctly distinguishing between the disease's progressing phases [12]. In addition, some approaches depend heavily on handcrafted features or require supplementary clinical data, which may not always be readily accessible [13][14][15]. The pathology of AD often spreads across the brain, with certain regions being impacted initially and others only becoming damaged in the most severe stages of the disease. The medial temporal lobe components, entorhinal and perirhinal cortex, and hippocampus exhibit the first alterations in this so-called topographic pattern that defines AD (Figure 1) [16]. This same pattern of disease progression made it possible to use imaging tools to examine specific brain regions and diagnose AD early.

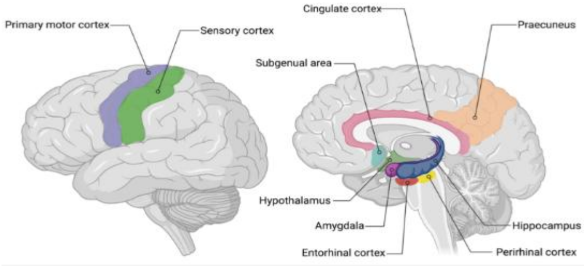

Fig.1. Regions Affected by Alzheimer's Disease

A DL framework based on a customized CNN for multi-class categorization of AD stages is proposed in this study to overcome these limitations Very mildly demented, mildly demented, moderately demented, and non-demented. The approach uses only T1-weighted structural MRI images and applies thorough pre-processing to improve image quality and learning. By automating feature extraction and training on class-specific data, the model aims to achieve high accuracy and robust classification. Its performance is also compared with traditional and hybrid methods to validate its effectiveness for early Alzheimer's diagnosis.

A. Motivation and Contribution

This research is driven by the pressing need for an accurate and timely diagnosis of AD, a neurodegenerative condition that worsens memory and cognitive function. Manual interpretation of structural MRI images is both time-intensive and susceptible to inter-observer variability, which can delay diagnosis and treatment. In order to help physicians make quicker and more accurate diagnoses, this project intends to automate the classification of Alzheimer's stages using DL models like CNN. This study's goal is to categorize AD. There are some key contributions as follows:

- Develop effective models for Alzheimer's disease detection by Alzheimer's MRI dataset.
- The proposed methodology delivers an automated and efficient DL framework for early Alzheimer's detection using T1-weighted structural MRI images, leveraging CNN to extract meaningful spatial features directly from brain scans.
- Robust pre-processing techniques including image resizing, normalization, flipping, cropping, and enhancement are applied to ensure consistency in input data and to improve model training quality.
- The study mitigates class imbalance through strategic data splitting and augmentation, enabling balanced learning across all four stages of AD, particularly addressing the limited samples in the Moderate Demented category.
- A unique CNN model is created and trained for multi-class classification, and it consistently performs well across important parameters, including good recall, accuracy, precision, and F1-score for every class.

B. Novelty with Justification

The integration of sophisticated image pre-processing methods with a specially designed CNN model for the multi-class categorization of AD using structural MRI scans is what makes this study innovative. To guarantee consistent and high-quality input data, the process consists of scaling, normalization, flipping, cropping, and enhancing. Class imbalance is addressed through strategic data splitting and augmentation, particularly improving prediction for the underrepresented Moderate Demented class. The proposed CNN architecture effectively learns spatial patterns from MRI images, achieving high accuracy, precision, recall, and F1-score. This performance surpasses existing models like SVM, RF, and CNN+RNN, confirming its strength in handling complex medical image classification and its potential

ISSN : 2455-9679
SJIF Impact Factor : 6.008

for real-world clinical deployment.

C. Structure of the Paper

The study is structured as follows: Section II reviews existing literature on AD detection using ML and DL. Section III details the proposed methodology, including data collection, pre-processing, CNN model design, and performance metrics. Section IV presents experimental results, and Section V concludes the study with future research directions.

II. LITERATURE REVIEW

This section reviews the research on automated MRI images for AD early detection. It highlights existing approaches, datasets, and performance metrics used in related studies. Additionally, Table I provides a summary of the reviewed literature discussed below:

Goel et al. (2025) suggest combining MRI and PET images using Multimodality fusion using wavelet transform to add structural and metabolic data for the early identification of this fatal neurodegenerative illness. Additionally, the fused pictures' characteristics are extracted by the ResNet-50 DL model. To categorize the collected features, a single hidden layer random vector functional link (RVFL) is employed. An evolutionary method is used to optimize the initial RVFL network's weights and biases in order to achieve the highest accuracy. It is proven that the suggested approach is successful by all tests and comparisons conducted using the publicly accessible AD Neuroimaging Initiative (ADNI) dataset [17].

P et al. (2025) explore an automated approach using DL to detect Alzheimer's from MRI images. It proposes an ensemble of three models - EfficientNet, Vision Transformer (ViT), and SqueezeNet which were each chosen for their unique strengths in image analysis. EfficientNet excels in feature extraction, ViT captures long-range dependencies within images, and SqueezeNet offers computational efficiency. These models were then trained to identify four stages of cognitive decline: There are four types of dementia: mild, moderate, very mild, and non-demented. By utilizing the complementary strengths of these architectures, their ensemble approach achieved an impressive 90.50% accuracy on a publicly available Kaggle dataset, outperforming individual models [18].

Kumar et al. (2024) evaluate ML and DL methods to find AD using a dataset of 12,000 images, divided into four groups of 3,000 images each, and processed with representations of the Histogram of Oriented Gradients (HOG) features. Principal Component Analysis (PCA) was used in the ML analysis, Standard Scaling, and Stratified K-Fold cross-validation across four algorithms. The SVM achieved an accuracy of 79%, RF reached 67% DT obtained 51%, and LR recorded 65%. For the DL approach, utilized data augmentation and trained a CNN, achieving 87% accuracy. Transfer learning with the VGG16 model resulted in the highest accuracy of 91% [19].

Hossen Abir and Salam (2024) various ML techniques to assess AD detection based on eleven significant criteria. Developed predictive models using supervised learning algorithms, including RF, XGBoost, GB, CatBoost, SVM, KNN, LR, DT, and Voting Classifier. The model's performance was improved by applying grid search cross-validation and hyperparameter adjustment. With an accuracy of 96.30%, the RF, XGBoost, GB, and CatBoost classifiers were the most accurate, closely followed by the Voting Classifier. KNN and DT models performed well, with respective accuracy rates of 95.06% and 93.83%. At 92.59% and 90.12%, respectively, SVM and LR were less accurate [20].

Nan et al. (2024) A unique and repeatable multi-classification framework for AD early automated diagnosis is provided in this research to assess and validate the aforementioned concerns experimentally. Multimodal learning is then used to get some potential reasons for the early detection of AD using a suite of tests based on the ADNI-1 data. According to experimental findings, SNP has the greatest accuracy rate (57.09%) for AD early diagnosis. Multi-modal ML performance is enhanced by 3% to 7% when the Single Nucleotide Polymorphism modality is added to the modality combination. [21].

Lu and Gurram (2023) use ML to analyse handwriting and voice patterns in order to identify AD. Using Mel-Frequency cepstral coefficient testing, picture normalisation, binarization, and feature extraction, over 15,000 samples including handwriting, audio, and cognitive data from AD patients and controls were preprocessed. Six machine learning models were trained using an 80/20 split and a total F1-Score of 96.2% to detect AD based on handwriting and speech signals, including slurred speech, abrupt sentence endings, marked forgetfulness, readability, stroke information, and zone-based traits. React JavaScript and Python OpenCV were used to create the "revoAD" mobile app, which addressed healthcare inequities by providing low-cost screening, particularly in underprivileged regions,

ISSN : 2455-9679
SJIF Impact Factor : 6.008

and obtained 10 times faster diagnosis, 97.6% training accuracy, and 97.3% data validation accuracy [22].

Pallawi and Singh (2023) provide a framework for utilizing brain MRI to categorize different phases of this illness. The recent surge in interest in DL approaches can be attributed to their promising results in object recognition and classification problems. The most frequent problem with these systems, however, is the need for large datasets. Many academics are using the Transfer Learning (TL) approach to get around this problem and utilize pre-trained models. The EfficientNetB0 model has been refined in this work using TL on a Kaggle dataset that divides Alzheimer's disease into four stages. With a multi-class classification accuracy of 95.78%, the result shows that the model outperforms the most advanced techniques [23].

Sharma, Goel and Murugan, (2022) Patients suspected of having AD can be clinically examined non-invasively using magnetic resonance imaging (MRI). The research proposes an optimized deep learning network (DLN) model for automated diagnosis of AD, MCI, and CN. The primary phase in model design is choosing the hyperparameters, which are crucial to DLN training and success. The DLN layers are therefore trained by optimizing the hyperparameters using a whale optimization algorithm (WOA). The proposed optimal DLN is tested using the publicly accessible ADNI dataset, which contains structural MRI (sMRI) scans of AD, MCI, and CN. The suggested model's effectiveness is demonstrated by comparing the optimized DLN's performance to that of the state-of-the-art networks [24].

This study reviews recent research on AD detection using ML and DL, summarizing methodologies, model performance, key advantages, limitations, and future directions. A detailed comparison of techniques and outcomes across studies is presented in Table I.

TABLE I. COMPARATIVE SUMMARY OF RECENT STUDIES ON AD DETECTION USING ML AND DL TECHNIQUES

References	Methodology	Results Analysis	Advantages	Limitations & Future Work
Goel et al. (2025)	Wavelet-based fusion of MRI & PET + ResNet-50 for feature extraction + RVFL classifier optimized with evolutionary algorithm	High classification accuracy on ADNI dataset (exact value not specified)	Combines structural & metabolic data for better accuracy; evolutionary optimization enhances performance	Exact accuracy not mentioned; future work may include real-time clinical integration
P et al. (2025)	Ensemble of EfficientNet, ViT, SqueezeNet for multi-stage classification	Achieved 90.50% accuracy on Kaggle dataset	Combines models for robustness; effective for 4-stage dementia detection	May need larger datasets for generalization; future work can include temporal progression modeling
Kumar et al. (2024)	HOG feature extraction + PCA + SVM, RF, LR, DT + CNN + VGG16 Transfer Learning	VGG16 achieved highest accuracy: 91%, CNN: 87%, SVM: 79%	Compared classical ML and DL approaches; VGG16 shows strong transfer learning potential	Classical ML lags behind; DL approaches require more data augmentation
Hossen Abir & Salam (2024)	Used RF, XGBoost, CatBoost, etc. with GridSearchCV and hyperparameter tuning	Highest accuracy: 96.30% (RF, XGBoost, CatBoost), Voting Classifier: 96.29%	Comprehensive evaluation of models; high-performance ensemble classifiers	Focused only on structured data; integrating image data could enhance results
Nan et al. (2024)	Multimodal learning using ADNI-1 dataset + SNP integration	SNP modality improved accuracy by 3–7%, max accuracy: 57.09%	Introduces genetic modality; reproducible framework	Accuracy relatively low; needs integration with imaging modalities
Lu & Gurram, (2023)	Speech + handwriting analysis via MFCC, image preprocessing + 6 ML models + revoAD app	F1-score: 96.2%, Validation accuracy: 97.3%	Multimodal (audio/image) data improves detection; real-time low-cost app	Clinical validation needed; scalability for global use under research
Pallawi & Singh (2023)	Transfer learning using fine-tuned EfficientNetB0 on MRI dataset	Multi-class accuracy: 95.78%	High accuracy with fewer resources; tackles data scarcity	Models may overfit small datasets; expanding to multimodal data can improve generalization
Sharma, Goel & Murugan, (2022)	Deep Learning Network (DLN) + Whale Optimization Algorithm (WOA) for hyperparameter tuning	Not specified, but reported better than existing models	Automated hyperparameter tuning improves performance; tested on ADNI	Accuracy not quantified; requires benchmarking with new datasets

A. Research Gap

Although various studies have demonstrated high accuracy using individual modalities like MRI, SNP, and handwriting data, the integration of multimodal data for better Alzheimer's detection is lacking in unified systems. Most models are evaluated on limited or imbalanced datasets, with insufficient external validation. Furthermore, existing methods often

ISSN : 2455-9679
SJIF Impact Factor : 6.008

lack interpretability and are not optimized for real-time or resource-constrained environments. This highlights the need for robust, explainable, and generalizable AI models capable of handling diverse clinical data across real-world settings.

III. METHODOLOGY

The suggested approach for identifying AD by structural MRI image classification includes crucial phases in model training, data augmentation, and data preparation, as shown in Figure 2. The structural MRI dataset, sourced from Kaggle, undergoes pre-processing involving resizing of all images to a standardized resolution of 176×208 pixels, followed by normalization to ensure consistent pixel intensity distribution. Data augmentation techniques including cropping, picture improvement, and horizontal and vertical flipping are used to add diversity and enrich the training set while maintaining disease-relevant properties. Subgroups of the dataset are separated for training (80%), validation (10%), and testing (10%). To automatically extract discriminative characteristics and categories MRI scans into the appropriate phases of AD, the creation and training of a CNN. A balanced and thorough evaluation of classification efficacy is provided by the rigorous evaluation of model performance utilizing accuracy, precision, recall, and F1-score, especially when class imbalance is present, as is frequently seen in medical imaging datasets.

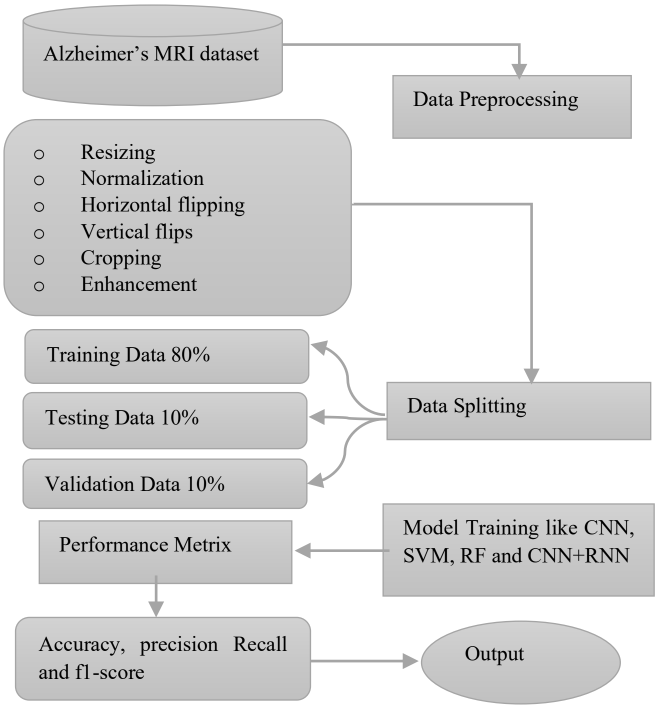

Fig. 2. Flowchart for AD on structural MRI Images

A. Data Collection

The study's Alzheimer's MRI dataset was acquired via Kaggle. The database known as the Open Access Series of Imaging Studies (OASIS) is one of its publicly available sources. There are four clinical classifications for the T1-weighted structural MRI brain scans in the dataset: There are four types of dementia: moderate, mild, non-demented, and very mild. The 5,121 MRI pictures show a class imbalance, with 2,560 listed as Non-Demented, 1,792 as Very Mild Demented, 717 as Mild Demented, and 52 as Moderate Demented. All images are grayscale sagittal slices, preprocessed and resized to a consistent resolution of 176×208 pixels, and organized into class-specific folders for ease of model training. No additional clinical metadata such as age or gender is provided, making the dataset specifically suitable for image-based classification tasks using DL techniques. Figure 3 below displays the dataset's sample images:

ISSN : 2455-9679
SJIF Impact Factor : 6.008

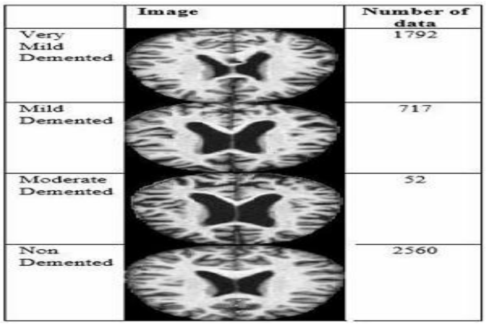

Fig.3. Sample Multiclass Images of the AD Dataset

Figure 3 displays the MRI brain scans from the multiclass dataset on AD, they are separated into four clinical phases: Mildly Demented (717 photos), Very Mild Demented (1,792 images), Moderately Demented (52 photographs), and Non-Demented (2,560 images). This class distribution highlights a significant imbalance, particularly with the limited number of Moderate Demented cases, which may influence model training and classification performance.

B. Data Preprocessing

Data pre-processing aids in keeping irrelevant information out of ML and DL models. It is essential to allow the models to learn patterns or features from several perspectives if the dataset is biased. This improves the model's performance [25]. The subsequent round of adjustments made to every training picture during augmentation:

1. Resizing

Resizing is a method for altering the size of pictures. If this is not addressed throughout the training process, compatibility problems arise. Therefore, the scaling approach was used after the enhancement procedure [26]. The resizing mathematical formula is stated as follows in Equation (1):

$$w_{new}, h_{new} = \frac{M}{\max(w,h)}(w,h) \quad (1)$$

2. Normalization

The pixel range is continuously distributed via normalization from a range of the lowest and maximum pixel values. It alters an image's perspective rather than its form [27]. The normalization technique's mathematical formula is stated as follows in Equation (2):

$$X_{normalized} = \frac{x - mean\,(x)}{x_{max} - x_{min}} \quad (2)$$

3. Horizontal and Vertical Flipping

There are two ways to flip an image: vertically and horizontally. This work made use of both. The active layer may be rotated from top to bottom or vertically. The horizontal flip makes the image appear as if it were reflected from a mirror. Each layer of the image is horizontally flipped from right to left or left to right. Only the pixel's position on the x-axis changes in this way, and no other information is lost. Equations (3) and (4) display the formula for both vertical and horizontal flipping as follows:

$$Horizontal\,(f(x)) = x^2 \quad (3)$$

$$Vertical\,(f(x)) = \sqrt{x} \quad (4)$$

4. Cropping

In addition, cropping was used while resizing the photos. Scaling is necessary after cropping, and augmentation is necessary to prevent the loss of important information. This is the second reason why this is necessary. It was essential

ISSN : 2455-9679
SJIF Impact Factor : 6.008

to resize to the original proportions since the training procedure was quick and trouble-free [28].

5. Enhancement

If the photographs in Figure 4 are resized, enhancement is a necessary step. The quality of images that need to be cropped or scaled may be compromised if they must be scaled back to their original size. Image augmentation is employed to overcome this obstacle. This is done to keep an image free of noise. Equation (5) for enhancement displays the following mathematical formulas:

$$g(x,y) = \begin{cases} a_1 f(x,y) f(x,y) < r_1 \\ a_2(f(x,y) - r_1) + s_1, \ r_1 \geq f(x,y) < r_2 \\ a_3(f(x,y) - r_2) + r_2, \ f(x,y) < r_3 \end{cases} \quad (5)$$

Where a_1, a_2, and a_3 are scaling factors for numerous greyscale regions and ss_1, s_2, r_1, and r_2He adjustable parameters, $g(x, y)$ are the image's output and $f(x, y)$ is the input pixel data.

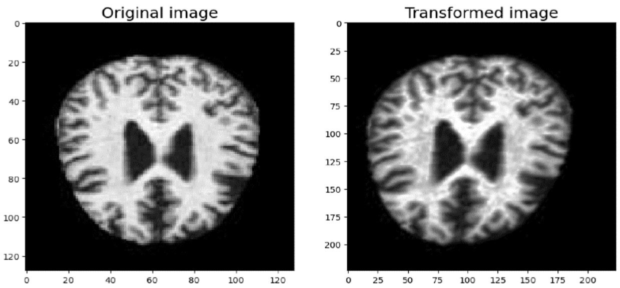

Fig.4. Original Vs. Transformed Image After Full Augmentation Sequence.

An example of a changed original image from the dataset is shown in Figure 4. After the entire augmentation process, the image on the right was produced. Despite drastically altering the way photos seem, the transformation preserves the traits that are used to categories AD.

C. Data Splitting

A rigorous performance evaluation was ensured by splitting the dataset using an 80:10:10 division into subsets for testing, validation, and training. As the model was being trained, the validation set tracked its generalization, while the test set evaluated the model's performance at the end.

D. Convolutional Neural Network (CNN) model for image classification

A DL architecture called a CNN was created especially for processing data having a grid-like layout, like images [29]. It leverages layers of convolutional operations followed by non-linear activations and pooling to extract hierarchical spatial features [30]. Unlike traditional neural networks, CNNs automatically learn relevant patterns (e.g., textures, edges, and shapes) through multiple layers, enabling robust feature extraction and classification. This architecture is particularly effective in medical image analysis, where subtle variations in structural patterns such as those in MRI scans are critical for disease detection and diagnosis. The mathematical Equation of the CNN is shown in Equation (6):

$$(f * g)[m,n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f[m,n] g[m-i, n-j] \quad (6)$$

Where, f[i,j] represents the input image, while $g[m-i, n-j]$ Denotes the convolution kernel (or filter) applied over the image. The indices i and j iterate over the input image's width and height, accordingly. The variables mmm and n correspond to the row and column positions in the output feature map where the convolution result is computed. The output $(f * g)[m,n]$Is obtained by summing the element-wise products between the image patch and the kernel at each position.

In signal processing, convolution is a helpful procedure. Convolution may apply filters to images, producing various

ISSN : 2455-9679
SJIF Impact Factor : 6.008

effects like smoothing or sharpening. Three parameters govern convolutional layers in a neural network: The number of filters, the size of each filter determined by the kernel, and the number of pixels to "skip" before the filter is applied to the next block of pixels are all determined by the stride size. A convolutional network learns the optimal filters by minimizing their loss during filter parameter training. These learnt filters can identify helpful filters that can help later layers categories their data. This applies to a wide range of issues, including face recognition.

E. Performance Metrics

It is critical to evaluate a model's performance once it has been constructed. Four evaluation criteria were used in this study: F1-score, recall, accuracy, and precision. These metrics have the following definitions, and Table II displays the confusion metrics [31].

True Positive (TP): The percentage of times an AD detection system properly detects an image containing AD is known as the TP rate.

True Negative (TN): An AD detection system's TN rate is the proportion of times it correctly identifies a picture without AD.

False Positive (FP): The FP counts how many times an image is mistakenly classified as AD by the model.

False Negative (FN): The FN counts the instances in which the model misclassifies a picture as normal.

TABLE II. CONFUSION METRICS

Value	Actual positive	Actual Negative
Predictive positive	TP	FP
Predictive negative	FN	TN

1. Accuracy

It is a metric used to determine the percentage of correctly categorized results among all instances [32]. The statistical equation of the accuracy is shown in Equation (7):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{7}$$

2. Precision

The ratio of correctly predicted positive rates to all projected positive rates is computed in this way. The classifier is considered good when the precision value is 1. The precision is expressed in Equation (8):

$$Precision = \frac{TP}{TP+FP} \times 100 \tag{8}$$

3. Recall

A real positive rate is considered a good classifier if the recall is 1. The mathematical expression of the recall is shown below in Equation (9):

$$Recall = \frac{TP}{TP+FN} \times 100 \tag{9}$$

4. F1 Score

This metric takes into account the dimensions of precision and recall. The accuracy and recall measurements must both be 1 for the F1 score to be 1. It is expressed as Equation (10):

$$F1 - \text{score} = \frac{2 \times recall \times precision}{recall + precision} \tag{10}$$

These matrices are utilized to determine the ML models.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

IV. RESULT ANALYSIS AND DISCUSSION

This section details how a CNN model based on DL was used to analyze structural MRI data in order to identify AD from the experiment findings. The model was evaluated using standard classification metrics, such as F1-score, recall, accuracy, and precision. Python in Jupiter Notebook using Google Colab was used for model construction and testing, with the aid of libraries like TensorFlow, Keras, NumPy, Pandas, Seaborn, and Matplotlib. The Lenovo Legion 7 laptop utilized for all testing has a 10th Gen Intel® Core™ i7-10750H CPU (2.60 GHz, up to 5.00 GHz Turbo Boost, 6 cores, 12 threads, 12 MB cache), 32 GB DDR4 RAM (2933 MHz), and an NVIDIA® GeForce RTXTM 2070 Max-Q GPU with 8 GB VRAM. It also included a 1 TB SSD and a 15.6" Full HD (1920 x 1080) anti-glare display with a refresh rate of 144 Hz, which supplied the amount of processing power required to effectively train and assess the suggested model. Table III shows the performance of the CNN model for AD detection as follows:

TABEL III. CNN MODEL PERFORMANCE FOR THE DETECTION OF AD

Measure	CNN
Accuracy	95.00
Precision	95.00
Recall	95.00
F1-score	95.00

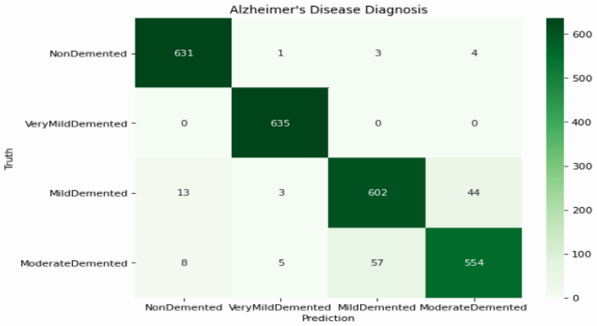

Fig.5. Confusion Matrix of Proposed CNN Model

The CNN model's confusion matrix, which summarizes its effectiveness in categorizing AD into four groups, is displayed in Figure 5: Non Demented, Very Mild Demented, Mild Demented, and Moderate Demented. The matrix displays TP and misclassified instances, with the diagonal values (631, 635, 602, and 554) representing correctly predicted cases for each class. This distribution shows the model performs effectively, with some confusion between adjacent dementia stages.

	precision	recall	f1-score	support
NonDemented	0.97	0.99	0.98	639
VeryMildDemented	0.99	1.00	0.99	635
MildDemented	0.91	0.91	0.91	662
ModerateDemented	0.92	0.89	0.90	624
micro avg	0.95	0.95	0.95	2560
macro avg	0.95	0.95	0.95	2560
weighted avg	0.95	0.95	0.95	2560
samples avg	0.95	0.95	0.95	2560

Fig. 6. Classification Report of CNN on Test Data

The classification report of the suggested CNN model on the test data for diagnosing AD across four classes is displayed in Figure 6: Non Demented, Very Mild Demented, Mild Demented, and Moderate Demented. The recall, F1-score, accuracy values of the model demonstrated outstanding performance measures. Mild Demented and Moderate

ISSN : 2455-9679
SJIF Impact Factor : 6.008

Demented classes scored slightly lower. The overall micro, macro, weighted, and sample averages for all metrics are consistently at 0.95, indicating a strong and balanced classification performance across all classes.

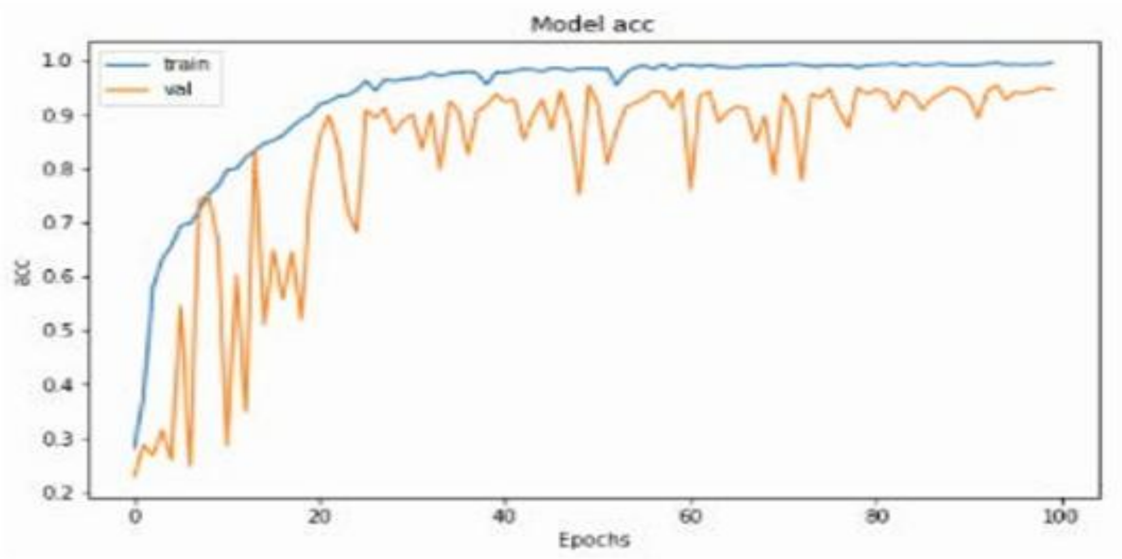

Fig.7. Training and Validation Accuracy Curve of CNN Model

The accuracy curve for the training and validation of the suggested CNN model, demonstrating its performance across 100 epochs, is displayed in Figure 7. As the training accuracy gradually rises to about 99%, it shows that the training data is being learnt effectively. The validation accuracy rises sharply during the initial epochs and fluctuates, with noticeable variance. The model has some overfitting, as seen by the discrepancy between the training and validation curves, given that it consistently performs well on training data but behaves strangely on unknown validation data. This pattern highlights the need for regularization or data augmentation to improve generalization.

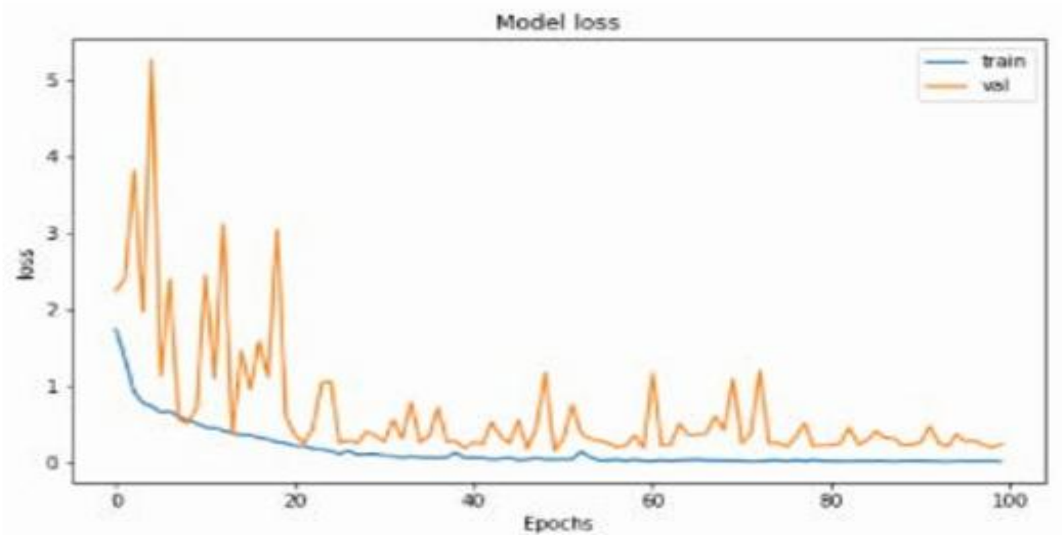

Fig.8. Training and Validation Loss Curve of CNN Model

The training and validation loss curves for the proposed CNN model, which display its performance across 100 epochs, are displayed in Figure 8. The training loss steadily decreases and converges to nearly zero, indicating that the training set is helping the model learn efficiently. In contrast, the validation loss exhibits significant fluctuations, especially in the first 30 epochs, with peaks exceeding 5, before gradually stabilizing around 2 to 5. This apparent discrepancy between the model does well on training data but struggles to generalize consistently on unknown validation data, which makes overfitting a possibility according to the training and validation loss curves.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

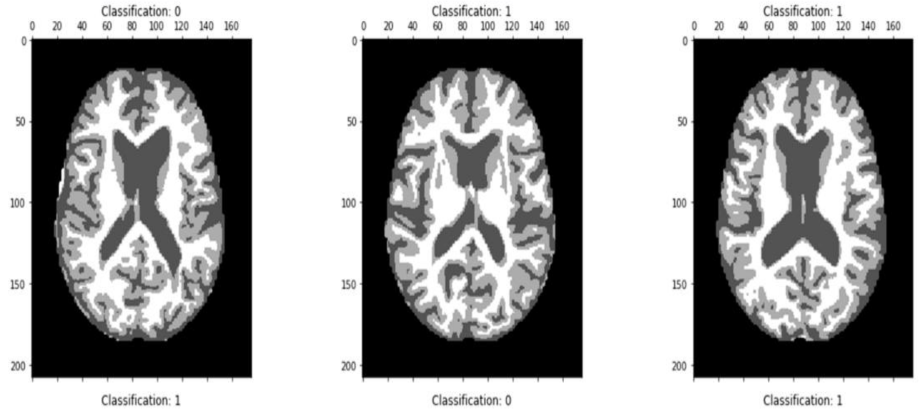

Fig.9. MRI Image Classification

The MRI image classification findings in Figure 9 indicate that the illness is not present, with a classification value of 0. Classification 1 displays the brain imaging of an individual with AD.

A. Discussion

This section provides a comparative analysis of structural MRI scans for the detection of AD. The suggested Convolutional Neural Network (CNN) model is compared to the Support Vector Machine (SVM) [33], Random Forest (RF) [34], and a CNN+RNN [35] hybrid model for AD diagnosis in Table IV. Across all assessment measures, the suggested CNN performs better. It achieves the highest accuracy (95%), outperforming CNN+RNN (93%), SVM (92%), and RF (88%). Similarly, it records the highest precision, recall, and F1-score, all at 95%, indicating balanced and robust classification capabilities. In contrast, the CNN+RNN model, though better than SVM and RF, trails the proposed CNN, while RF shows the lowest performance across all metrics.

TABLE IV COMPARISON BETWEEN PROPOSED CNN AND EXISTING MODELS FOR ALZHEIMER'S DISEASE DETECTION

Measure	CNN	SVM	RF	CNN+RNN
Accuracy	95	92	88	93
Precision	95	91.9	88	92
Recall	95	91.9	88	91
F1-score	95	91.9	87	91.5

This end-to-end DL architecture effectively extracts and trains discriminative features from brain imaging data, making it the ideal CNN model for AD diagnosis. It uses pooling layers and stacked convolutional layers to capture both low-level and high-level spatial patterns associated with the progression of illness. Without the need for human participation or manually created features, the model can produce reliable predictions due to its hierarchical feature extraction. The model's main benefit is its strong generalization across a range of input data, which consistently yields excellent results across all assessment measures. Its robustness, scalability, and automation make it an effective method for accurately and early AD identification.

V. CONCLUSION AND FUTURE WORK

AD is a degenerative neurological disorder that primarily impacts cognitive, memory, and behavior. It is one of the primary causes of dementia in the elderly. For prompt treatment and better patient outcomes, early and precise AD identification is essential. A CNN model based on DL was proposed to automatically categories AD phases using structural

MRI data. The CNN model achieved 95.00% accuracy, precision, recall, and F1-score, demonstrating consistent good performance in identifying significant spatial features from MRI data. This illustrates the model's dependability and suitability for application in clinical situations. A crucial component of practical implementation in healthcare settings, the model's consistency across performance indicators also implies that it reduces FP and FN. The model can assist neurologists and radiologists in early diagnosis by providing precise stage-by-stage categorization of AD, progression tracking, and treatment planning, especially in resource-limited or underserved areas where access to specialized diagnostics may be limited.

The model can be further enhanced in further research by addressing class imbalance and overfitting through advanced augmentation and regularization techniques. Furthermore, using multimodal data sources including genetic biomarkers, PET imaging, and cognitive tests might improve the accuracy of diagnosis. Finally, applying explainable AI (XAI) techniques will improve model interpretability, fostering greater trust and integration into real-world clinical workflows.

REFERENCES

[1] [1] M. W. Oktavian, N. Yudistira, and A. Ridok, "Classification of Alzheimer's Disease Using the Convolutional Neural Network (CNN) with Transfer Learning and Weighted Loss," IAENG Int. J. Comput. Sci., vol. 50, no. 3, pp. 1–10, 2023.

[2] V. H. Jasodanand et al., "AI-driven fusion of neurological work-up for assessment of biological Alzheimer's disease," Mar. 17, 2025. doi: 10.1101/2025.03.12.25323862.

[3] M. A. Mostafiz, "Machine Learning for Early Cancer Detection and Classification : AI-Based Medical Imaging Analysis in Healthcare," Int. J. Curr. Eng. Technol., vol. 15, no. 3, pp. 251–260, 2025, doi: https://doi.org/10.14741/ijcet/v.15.3.7.

[4] R. P. Mahajan, "Transfer Learning for MRI image reconstruction: Enhancing model performance with pretrained networks," Int. J. Sci. Res. Arch., vol. 15, no. 1, pp. 298–309, Apr. 2025, doi: 10.30574/ijsra.2025.15.1.0939.

[5] R. Dattangire, D. Biradar, R. Vaidya, and A. Joon, "A Comprehensive Analysis of Cholera Disease Prediction Using Machine Learning," in Congress on Intelligent Systems, 2025, pp. 555–568. doi: 10.1007/978-981-96-2697-7_39.

[6] I. Malik, A. Iqbal, Y. H. Gu, and M. A. Al-antari, "Deep Learning for Alzheimer's Disease Prediction: A Comprehensive Review," Diagnostics, vol. 14, no. 12, 2024, doi: 10.3390/diagnostics14121281.

[7] S. Singamsetty, "Neurofusion: Advancing Alzheimer's Diagnosis with Deep Learning and Multimodal Feature Integration," Int. J. Educ. Appl. Sci. Res., vol. 8, no. 01, pp. 23–32, 2021.

[8] V. Singh, D. K. Pathak, and P. Gupta, "Integrating Artificial Intelligence and Machine Learning into Healthcare ERP Systems: A Framework for Oracle Cloud and Beyond," ESP J. Eng. Technol. Adv., vol. 3, no. 2, pp. 171–178, 2023, doi: 10.56472/25832646/JETA-V3I6P114.

[9] N. Patel, "Quantum Cryptography In Healthcare Information Systems: Enhancing Security In Medical Data Storage And Communication," J. Emerg. Technol. Innov. Res., vol. 9, no. 8, 2022.

[10] S. Pahune and N. Rewatkar, "Large Language Models and Generative AI's' s Expanding Role in Healthcare," no. January, 2024, doi: 10.13140/RG.2.2.20109.72168.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

[11] L. Cardinali et al., "Early Detection of Alzheimer's Disease via Machine Learning-Based Microwave Sensing: An Experimental Validation," Sensors, vol. 25, no. 9, 2025, doi: 10.3390/s25092718.

[12] S.-K. Kim, J. Bin Kim, H. Kim, L. Kim, and S. H. Kim, "Early Diagnosis of Alzheimer's Disease in Human Participants Using EEGConformer and Attention-Based LSTM During the Short Question Task," Diagnostics, vol. 15, no. 4, p. 448, Feb. 2025, doi: 10.3390/diagnostics15040448.

[13] W. M. van Oostveen and E. C. M. de Lange, "Imaging techniques in alzheimer's disease: A review of applications in early diagnosis and longitudinal monitoring," 2021. doi: 10.3390/ijms22042110.

[14] P. Prasad Jalaja, D. Kommineni, A. Mishra, R. Tumati, C. Anna Joseph, and R. Veer Samara Sihman Bharattej Rupavath, "Predictors of Mortality in Acute Myocardial Infarction: Insights From the Healthcare Cost and Utilization Project (HCUP) Nationwide Readmission Database," Cureus, vol. 17, no. 5, p. e83675, May 2025, doi: 10.7759/cureus 83675.

[15] S. S. S. Neeli, "Heart Disease Prediction For A Cloud-Based Smart Healthcare Monitoring System Using Gans And Ant Colony Optimization," Int. J. Med. Public Heal., vol. 14, no. 4, p. 11, 2024.

[16] R. P. Mahajan, "Development of Predictive Models for Early Detection of Alzheimer s Disease Using Machine Learning," Int. J. Curr. Eng. Technol., vol. 15, no. 2, pp. 115–123, 2025.

[17] T. Goel, R. Sharma, M. Tanveer, P. N. Suganthan, K. Maji, and R. Pilli, "Multimodal Neuroimaging Based Alzheimer's Disease Diagnosis Using Evolutionary RVFL Classifier," IEEE J. Biomed. Heal. Informatics, vol. 29, no. 6, pp. 3833–3841, 2025, doi: 10.1109/JBHI.2023.3242354.

[18] B. P, R. Panda, A. Boddu, and N. Joshy, "Alzheimer's Disease Detection Using Ensemble Techniques on MRI Images," in 2025 International Conference on Advanced Computing Technologies (ICoACT), IEEE, Mar. 2025, pp. 1–5. doi: 10.1109/ICoACT63339.2025.11005279.

[19] N. K. Kumar, V. D. Ambeth Kumar, D. Quamar, B. S. K. Reddy, R. Yogendra, and N. Poojitha, "Enhanced Alzheimer's Disease Prediction through Advanced Imaging: A Study of Machine Learning and Deep Learning Approaches," in 2024 5th International Conference on Smart Electronics and Communication (ICOSEC), 2024, pp. 1177–1182. doi: 10.1109/ICOSEC61587.2024.10722116.

[20] M. I. Hossen Abir and T. Salam, "Comparative Analysis and Prediction of Machine Learning Algorithms for MRI-Based Alzheimer's Detection Using Multi-modal Data," in 2024 IEEE International Conference on Computing, Applications and Systems (COMPAS), IEEE, Sep. 2024, pp. 1–5. doi: 10.1109/COMPAS60761.2024.10797119.

[21] F. Nan et al., "A Multi-Classification Assessment Framework for Reproducible Evaluation of Multimodal Learning in Alzheimer's Disease," IEEE/ACM Trans. Comput. Biol. Bioinforma., vol. 21, no. 4, pp. 559–572, 2024, doi: 10.1109/TCBB.2022.3204619.

[22] B. M. Lu and A. S. Gurram, "revoAD: Revolutionizing Alzheimer's Disease Diagnosis through Multimodal Machine Learning for Universal Screening via Speech and Handwriting Patterns," in 2023 8th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), 2023, pp. 493–494. doi: 10.1109/ICIIBMS60103.2023.10347810.

[23] S. Pallawi and D. K. Singh, "Detection of Alzheimer's Disease Stages Using Pre-Trained Deep Learning Approaches," in 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), 2023, pp. 252–256. doi: 10.1109/ICCCMLA58983.2023.10346730.

[24] R. Sharma, T. Goel, and R. Murugan, "An Optimized Deep Learning Network for Prognosis of Alzheimer's Disease using Structural Magnetic Resonance Imaging," in IEEE Region 10 Humanitarian Technology Conference, R10-HTC, 2022. doi: 10.1109/R10-HTC54060.2022.9929626.

[25] F. Mehmood, A. Mehmood, and T. K. Whangbo, "Alzheimer's Disease Detection in Various Brain Anatomies Based on Optimized Vision Transformer," Mathematics, vol. 13, no. 12, 2025, doi: 10.3390/math13121927.

[26] R. Ashtagi et al., "Deep learning-enhanced mri imaging for early alzheimer's detection," Int. J. Comput. Digit. Syst., vol. 17, no. 1, pp. 1–14, 2025.

[27] S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani, and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," in 2024 9th International Conference on

ISSN : 2455-9679
SJIF Impact Factor : 6.008

Communication and Electronics Systems (ICCES), IEEE, Dec. 2024, pp. 1149–1154. doi: 10.1109/ICCES63552.2024.10859381.

[28] Z. Yao et al., "Artificial intelligence-based diagnosis of Alzheimer's disease with brain MRI images," Eur. J. Radiol., vol. 165, Aug. 2023, doi: 10.1016/j.ejrad.2023.110934.

[29] P. K. Mandal and R. V. Mahto, "Deep Multi-Branch CNN Architecture for Early Alzheimer's Detection from Brain MRIs," Sensors, 2023, doi: 10.3390/s23198192.

[30] S. R. Sagili, S. Chidambaranathan, N. Nallametti, H. M. Bodele, L. Raja, and P. G. Gayathri, "NeuroPCA: Enhancing Alzheimer's disorder Disease Detection through Optimized Feature Reduction and Machine Learning," in 2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), IEEE, Jul. 2024, pp. 1–9. doi: 10.1109/ICEEICT61591.2024.10718628.

[31] T. Ghosh, M. I. A. Palash, M. A. Yousuf, M. A. Hamid, M. M. Monowar, and M. O. Alassafi, "A Robust Distributed Deep Learning Approach to Detect Alzheimer's Disease from MRI Images," Mathematics, vol. 11, no. 12, 2023, doi: 10.3390/math11122633.

[32] C. Kavitha, V. Mani, S. R. Srividhya, O. I. Khalaf, and C. A. Tavera Romero, "Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models," Front. Public Heal., 2022, doi: 10.3389/fpubh 2022.853294.

[33] M. Bari Antor et al., "A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease," J. Healthc. Eng., vol. 2021, no. 1, pp. 1–12, Jul. 2021, doi: 10.1155/2021/9917919.

[34] S. Bringas, S. Salomón, R. Duque, C. Lage, and J. L. Montaña, "Alzheimer's Disease stage identification using deep learning models," J. Biomed. Inform., vol. 109, no. February, p. 103514, 2020, doi: 10.1016/j.jbi.2020.103514.

[35] M. Sarkar, "Integrating machine learning and deep learning techniques for advanced Alzheimer's disease detection through gait analysis," J. Bus. Manag. Stud., vol. 7, no. 1, pp. 140–147, 2025.