Pertanika J. Sci. & Technol. 32 (4): 1629 - 1650 (2024)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Borderline-DEMNET: A Workflow for Detecting Alzheimer's and Dementia Stage by Solving Class Imbalance Problem

Neetha Papanna Umalakshmi[1]*, Simran Sathyanarayana[1], Pushpa Chicktotlikere Nagappa[1], Thriveni Javarappa[1] and Venugopal Kuppanna Rajuk[2]

[1]Department of CSE, University of Visvesvaraya College of Engineering, Bangalore, 560001 Karnataka, India
[2]Bangalore University, Bangalore, 560001 Karnataka, India

ABSTRACT

Alzheimer's Disease (AD) is the leading cause of dementia, a broad term encompassing memory loss and other cognitive impairments. Although there is no known cure for dementia, managing specific symptoms associated with it can be effective. Mild dementia stages, including AD, can be treated, and computer-based techniques have been developed to aid in early diagnosis. This paper presents a new workflow called Borderline-DEMNET, designed to classify various stages of Alzheimer's/dementia with more than three classes. Borderline-SMOTE is employed to address the issue of imbalanced datasets. A comparison is made between the proposed Borderline-DEMNET workflow and the existing DEMNET model, which focuses on classifying different dementia and AD stages. The evaluation metrics specified in the paper are used to assess the results. The framework is trained, tested, and validated using the Kaggle dataset, while the robustness of the work is checked using the ADNI dataset. The proposed workflow achieves an accuracy of 99.17% for the Kaggle dataset and 99.14% for the ADNI dataset. In conclusion, the proposed workflow outperforms previously identified models, particularly in terms of accuracy. It also proves that selecting a proper class balancing technique will increase accuracy.

Keywords: Alzheimer's disease, Borderline-SMOTE, dementia, DEMNET model, evaluation matrices

ARTICLE INFO

Article history:
Received: 27 July 2023
Accepted: 01 February 2024
Published: 16 July 2024

DOI: https://doi.org/10.47836/pjst.32.4.10

E-mail addresses:
neethapud@gmail.com (Neetha Papanna Umalakshmi)
simrankrithika@gmail.com (Simran Sathyanarayana)
pushpacn@gmail.com (Pushpa Chicktotlikere Nagappa)
drthrivenij@gmail.com (Thriveni Javarappa)
neethapuds@gmail.com (Venugopal Kuppanna Rajuk)
* Corresponding author

INTRODUCTION

Alzheimer's Disease (AD) is a progressive neurological condition that gradually impairs memory, cognitive abilities, and eventually the capacity to perform basic tasks. In the United States, AD is presently classified as

ISSN: 0128-7680
e-ISSN: 2231-8526

the seventh most common cause of mortality (National Institute on Aging, 2021). The prevalence of AD in the United States is rapidly increasing. The number of Americans affected by this condition is on the rise, with over 6 million individuals of various age groups being affected. As of 2023, it is estimated that approximately 6.7 million Americans aged 65 and above are living with Alzheimer's (Alzheimer's Association, 2023).

Dementia refers to the decline in both behavioral and cognitive abilities, reaching a point where it hampers an individual's ability to carry out daily activities, such as thinking, remembering, and rationalizing. The severity of dementia varies, starting from mild stages, where it minimally impacts one's functioning, to the most advanced stages, where individuals require complete assistance from others to perform basic tasks in their daily lives (National Institute on Aging, 2021).

Presently, there is a global prevalence of over 55 million individuals diagnosed with dementia, with approximately 60% residing in low- and middle-income nations. Annually, nearly 10 million new cases are reported. Dementia arises due to a range of ailments and brain injuries. AD stands as the predominant type of dementia, accounting for approximately 60% to 70% of cases (World Health Organization, 2023). Tangles and plaques remain recognized as fundamental features of AD. The progress in brain imaging technology has enabled researchers to directly observe alterations in brain structure and function, as well as the development and dissemination of abnormal tau and amyloid proteins in the living brain (National Institute on Aging, 2021).

Traditional imaging methods have had limited involvement in detecting AD, but contemporary imaging technologies have gained prominence in AD diagnosis. These advanced techniques aid in diagnosing AD and serve as crucial tools for assessing treatment effectiveness, making prognosis judgments, and facilitating drug development (Zeng et al., 2021). AI has demonstrated great potential in medicine, especially in neuroimaging. It encompasses computer systems that can carry out tasks typically requiring human intelligence. Deep Learning (DL) algorithms enable computers to learn the most effective data representations for a given problem. Machine Learning (ML) and DL aim to replicate the neural networks found in the human brain, resulting in Artificial Neural Networks (ANNs) consisting of nodes arranged in input, hidden, and output layers (Litjens et al., 2017) (Monsour et al., 2022).

The progress of technology and the ready availability of digital medical data have led to the widespread utilization of computer techniques in the medical domain (Liu et al., 2022) (Neetha et al., 2022). Imbalanced datasets pose a common challenge in health-related applications. The classification of medical data often faces the issue of uneven data distribution, where at least one class is disproportionately represented, comprising a considerably smaller portion of the dataset (Kotsiantis et al., 2005).

When rare events occur infrequently, they are frequently misinterpreted as unexplored or overlooked cases or disregarded as noise or anomalies. Consequently, this results in a greater occurrence of misclassifications for the positive class, especially when it represents the minority class, in contrast to the more prevalent class (Ali et al., 2013).

Many researchers have focused on addressing the classification challenges associated with multi-class imbalanced data in recent years. This particular task poses more significant obstacles when compared to binary imbalanced learning (Bi et al., 2019).

MOTIVATION

Our experiment is driven by the investigation into the factors contributing to AD and its impact on human life, particularly in relation to dementia. AD is the leading cause of dementia, a condition that can ultimately result in mortality. The imbalance of classes is a prevalent issue often encountered in real-time medical datasets. However, achieving a balanced dataset and ensuring high accuracy can be challenging. ML and DL techniques have proven effective in resolving various challenges in real-time scenarios. Hence, we attempted to implement specific techniques to address the class imbalance problem while improving classification accuracy (Pushpa et al., 2013).

CONTRIBUTIONS

Problem Statement

Designing effective prediction models for Alzheimer's or dementia data presents a major challenge, particularly due to the issue of class imbalance. Imbalanced datasets occur when one class has considerably larger data points than the other. In the realm of ML, various techniques have been developed to address imbalanced data, but not all techniques are considered the best for all datasets. We utilized an advanced technique that performs better in image processing. This approach is combined with a modified DL algorithm called DEMNET, resulting in significant improvements compared to other sampling methods. Also, getting good accuracy in Multi-Class situations is challenging.

Here are several notable contributions made by the work:

1. We have introduced a new workflow structure known as Borderline-DEMNET (DEMentia NETwork with Borderline SMOTE) to address the class imbalance challenge and enhance classification accuracy in four and five-class datasets.
2. The Borderline-DEMNET workflow has the potential to aid in the timely detection and diagnosis of diseases by focusing on class imbalance and multi-class classification issues sequentially.
3. Compared to previous studies on AD/dementia, our workflow demonstrates superior accuracy regarding four- and five-class problems.
4. It also demonstrates that employing an appropriate method for balancing class distribution can improve accuracy.

Related Works

We have organized the relevant papers into two categories: Binary Classification and Multi-class problem. Many papers and outcomes are available in the field of Binary Classification. In contrast, only a few papers address the specific disease in the context of multi-class problems involving three or more classes.

Binary Classification

In the study by (Prajapati et al., 2021), Deep Neural Network (DNN) was employed for AD classification across different stages. The dataset, sourced from the ADNI website, encompassed AD, Normal Control (CN), and Mild Cognitive Impairment (MCI) stages. However, the model was trained only on two classes at a time. Consequently, the model achieved an accuracy of 85.19% for AD vs. CN, 76.93% for MCI vs. CN, and 72.73% for AD vs. MCI.

Basaia et al. (2019) introduced the use of a Modified Convolutional Neural Network (CNN) for AD classification, utilizing a dataset from the ADNI source with Healthy Control (HC) and AD classes. Given the Binary classification, the model achieved an accuracy of 99%.

Wang et al. (2019) presented the application of the Ensemble Method for AD classification at various stages, utilizing a dataset from the ADNI website covering AD, Normal, and MCI stages. Similar to previous studies, the model was trained with two classes at a time, resulting in accuracies of 98.83% for AD vs. CN, 98.42% for MCI vs. CN, and 93.61% for AD vs. MCI.

CNN was employed for AD classification across stages using the ADNI dataset (Basheera et al., 2019). Although the dataset covered three stages, the model was trained with two classes simultaneously. Subsequently, it was trained with all three classes simultaneously, achieving a 100% accuracy for AD vs. Normal only.

Richhariya et al. (2020) utilized Universum Support Vector Machine-based Recursive Feature Elimination (USVM-RFE) for AD classification, employing a dataset from the ADNI website covering AD and Normal stages. Like other studies, the model was trained with two classes at a time, resulting in a 100% accuracy for AD vs. Normal.

Alinsaif et al. (2021) introduced Shearlet-based descriptors and deep features for AD classification using a dataset from the Open Access Series of Imaging Studies (OASIS) and ADNI, covering AD and Normal stages. The model, trained with two classes at a time, achieved an accuracy of 80% for AD vs. Normal.

Table 1 presents a summary of research papers focusing on Binary Classification. The Table 1 demonstrates that Binary Classification problems generally yield favorable performance results. However, the same model tends to underperform when it comes to higher classes.

Table 1
A compilation of studies focusing on the binary classification problem is presented in the following summary

Authors	Method	Dataset	Binary Class	Accuracy
Prajapati et al., 2021	DNN	ADNI	AD/CN MCI/CN AD/MCI	85.19% 76.93% 72.73%
Basaia et al., 2019	Modified CNN	ADNI	AD/HC	99%
Wang et al., 2019	Ensemble Method	ADNI	MCI/AD MCI/Normal AD/Normal	93.61% 98.42% 98.83%
Basheera et al., 2019	CNN	ADNI	AD/CN	100%
Richhariya et al., 2020	Universum Support Vector Machine-based Recursive Feature Elimination (USVM-RFE)	ADNI	CN/AD	100%
Alinsaif et al., 2021	Shearlet-based descriptors and deep features	OASIS	CN/AD	80%

Multi-Class Classification

Murugan et al. (2021) introduced the DEMNET model for classifying four and five-class datasets from Kaggle and ADNI. The Kaggle dataset, with classes like Mild Dementia (MID), Non-Dementia (ND), Moderate Dementia (MOD), and Very Mild Dementia (VMD), achieved 95.23% accuracy, suggesting room for improvement.

Basheera et al. (2019) and (Wang et al., 2019) utilized CNN and ensemble methods, respectively, for AD classification in ADNI datasets. Accuracy for AD vs. Normal vs. MCI was 86.7% and 97.52%, indicating potential for further enhancement. (Neetha et al., 2022) employed D-DEMNET with DenseNet-121 for five class classifications from ADNI, obtaining 95.16% accuracy. However, it was comparatively less effective in five class classifications.

Raju et al. (2021) applied Transfer Learning with VGG16 using Fastai, achieving 99% accuracy in a four-class dataset. Suganthe et al. (2021) used a combination of Inception and ResNet V2, achieving 79.12% accuracy in a similar dataset, with plans to improve accuracy.

Raju et al. (2021) also presented Cascaded 3D CNN features and Multilayer Perceptron for ternary classification in a dataset covering AD, MCI, and NC, obtaining 96.66% accuracy.

Table 2 provides an overview of the studies focusing on Multi-Class Classification tasks. Several models mentioned in Table 1 are also mentioned here. Additionally, Table 2 presents the specific limitations addressed in this paper. Furthermore, it is observed that only a limited number of models in the Multi-Class classification domain accommodate datasets with more than three classes. As a result, there is a need to enhance the accuracy of classification problems involving four or more classes in the dataset.

Table 2
Synopsis of the studies focused on Multi-Class Classification

Authors	Method	Dataset	Multi-Class	Accuracy	Drawbacks
Murugan et al., 2021	DEMNET	Kaggle	MID/MOD/ND/VMD	95.23%	Accuracy is yet to Increase.
Basheera et al., 2019	CNN	ADNI	AD/CN/MCI	86.7%	Accuracy is yet to increase.
Neetha et al., 2022	D-DEMNET	ADNI	AD/lMCI/MCI/eMCI/NC	95.16%	Comparatively, it is less effective in five class classifications.
Raju et al., 2021 March	Transfer Learning with VGG16 using Fastai	Kaggle	MID/MOD/ND/VMD	99%	Yet to try on higher class
Wang et al., 2019	Ensemble Method	ADNI	AD/Normal/MCI	97.52%	For three-Class.
Suganthe et al., 2021	Combination of Inception and ResNet V2	Kaggle	MID/MOD/ND/VMD	79.12%	Accuracy is yet to increase.
Raju et al., 2021	Cascaded 3D CNN features and Multilayer Perceptron classifier	ADNI	AD/MCI/NC	96.66%	For three-Class.

MATERIALS AND METHODS

Dataset Description

The standard datasets used in our study included the four-class dementia dataset sourced from Kaggle and the five-class AD dataset obtained from ADNI, also found in Kaggle(AD). These datasets were preprocessed and oversampled to extract their unique features. Additionally, we categorized the datasets into training, testing, and validation sets.

Kaggle

Dementia is a progressive condition that tends to worsen over time. However, the progression of dementia varies from person to person. Nevertheless, most individuals experience symptoms that align with the different stages of dementia. The datasets used for AD analysis were obtained from Kaggle, a freely accessible platform. These datasets consist of a total of 6400 MR images, categorized into four groups: Mild Dementia (MID), Non-Dementia (ND), Moderate Dementia (MOD), and Very Mild Dementia (VMD). Each image in the dataset is 64 × 64 pixels in size. As dementia advances, individuals may require assistance from a loved one or a professional caregiver, as the condition can hinder daily tasks and activities (Sarvesh, 2019). There are varying quantities of images for each class in the dataset. Specifically, there are 2240 images for the Non-Demented (ND) class, 64 images for the Very Mild Demented (VMD) class, 896 images for the Mild Demented (MID) class, and 3200 images for the Moderate Demented (MOD) class.

ADNI

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is a long-term, multicenter project aiming to develop biomarkers to detect and monitor AD early on. These biomarkers are based on various factors such as clinical information, imaging data, genetics, and metabolism. Over a decade ago, this collaboration between public and private entities was established and has significantly contributed to AD research by facilitating global data exchange among researchers. The ADNI dataset consists of 1296 images and categorizes AD into five groups: EMCI, MCI, LMCI, AD, and NC. To accommodate the DEMNET model, the images in the ADNI dataset are resized to 64 × 64 dimensions (Alzheimer's Disease Neuroimaging Initiative (Charan, 2022). The dataset consists of different classes, each containing a different number of images. These classes include Normal Control (NC) with 49 images, early Mild Cognitive Impairment (eMCI) with 204 images, late Mild Cognitive Impairment (lMCI) with 61 images, Mild Cognitive Impairment (MCI) with 198 images, and AD with 145 images.

Proposed Workflow

Our proposed methodology significantly enhances the accuracy of AD classification by employing the modified DEMNET approach to extract discriminative features. The suggested workflow is illustrated in Figure 1, encompassing several key phases: Data Pre-processing, Balancing the dataset using Borderline-SMOTE, the modified DEMNET phase, and Classification.

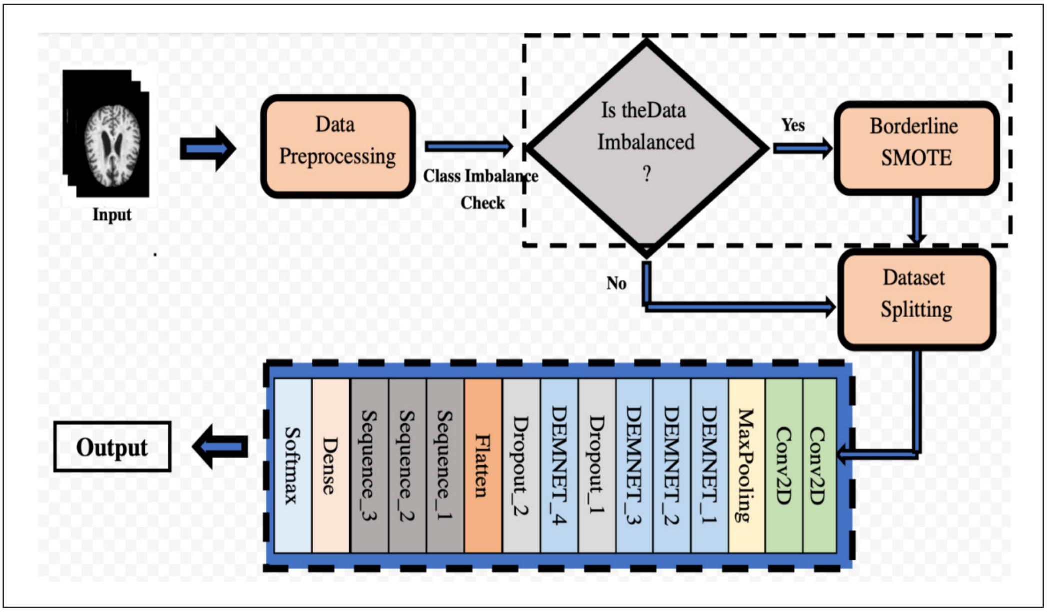


Figure 1. Proposed Borderline-DEMNET Workflow

Borderline-SMOTE for Class Imbalance

Certain DL techniques may struggle to effectively address the situation in cases of imbalance, leading to model malfunctions. Various balancing techniques can be employed to overcome this issue. Murugan et al. (2021) introduced the Synthetic Minority Over-sampling TEchnique (SMOTE) to tackle this problem, specifically in the context of images. However, utilizing SMOTE for image balancing was not widespread and was found best in text context. Consequently, we opted to implement Borderline SMOTE as specified in (Han et al., 2005), an extension of SMOTE that has proven effective in addressing image-related challenges.

Unlike alternative oversampling methods such as Random Oversampling or SMOTE (Synthetic Minority Over-sampling Technique), Borderline SMOTE often demonstrates advantages when addressing imbalanced datasets. This technique helps alleviate the risk of overfitting, commonly associated with generating synthetic examples for all instances in the minority class. Instead, it selectively enhances examples that pose greater challenges for the classifier to accurately classify. This targeted approach typically results in improved generalization performance and enhanced model robustness. Consequently, Borderline SMOTE proves to be a valuable tool in striving for a balanced and more precise classification in situations with imbalanced datasets. The Borderline SMOTE algorithm is presented in the accompanying Table 3.

Table 3
Borderline SMOTE algorithm

Algorithm BorderlineSMOTE()
Input: • *Training set T* • *Number of nearest neighbors ‘m’ for each example in the minority class P* • *Number of synthetic examples to generate ‘s’* • *Number of neighbors to consider when generating synthetic examples ‘k’* *Output:* • *Synthetic examples for the minority class P* Begin: Step 1: For each example p in the minority class P (i = 1, 2, ..., pnum): 1.1 Calculate the m nearest neighbors of p from the whole training set T. 1.2 Count the number of majority examples among the m nearest neighbors (m’). Step 2: For each p in the minority class P: 2.1 If m = m’, discard p as noise (not considered in the following steps). 2.2 If m / 2 ≤ m’ < m, add p to the set DANGER (easily misclassified). 2.3 If 0 ≤ m’ < m / 2, mark p as safe (not participating in the following steps). Step 3: Set DANGER = {p’1, p’2, ..., p’dnum} containing examples marked as DANGER. 3.1 For each example p’ in DANGER, calculate its k nearest neighbors from P. Step 4: For each example, p’ in DANGER: 4.1 Generate s synthetic examples: For j = 1 to s:

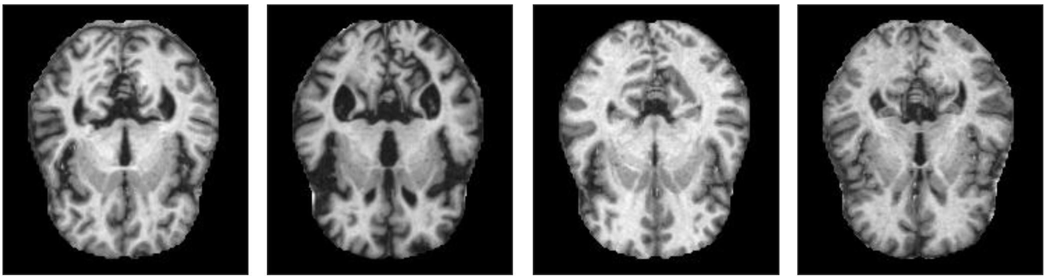
- Randomly select s nearest neighbors from p''s k nearest neighbors in P.
- Calculate the differences difj between p' and its s nearest neighbors.
- Multiply difj by a random number rj between 0 and 1.
- Generate a new synthetic minority example as p'' = p' + rj * difj.

End:
Output: Set of synthetic minority examples for the minority class P.

Borderline SMOTE represents an enhanced oversampling technique rooted in SMOTE. It strategically incorporates a limited number of class samples located on the border to generate new samples, thereby enhancing the distribution of sample categories. The Borderline SMOTE samples are categorized into Safe, Danger, and Noise. Ultimately, the oversampling is selectively applied to a small subset of Danger samples (Sun et al., 2022).

Our project's workflow illustrates the core functioning of our system. To begin, we acquired standard datasets: the four-class dementia dataset from Kaggle and the five-class AD dataset from ADNI. We then processed and oversampled these datasets to extract their unique characteristics. Additionally, we divided the dataset into three categories: training, testing, and validation. We employed the Borderline SMOTE approach for each dataset to create MRI images for minority groups. These images were utilized to train our modified DEMENT model, which underwent training with randomly selected images over 50 epochs.

Figure 2 represents some sample images generated by the method. Tables 4 and 5 give the list of images in count after passing to the method.



(a) (b) (c) (d)

Figure 2. Sample images generated from borderline SMOTE: (a) MID; (b) MOD; (c) ND; and (d) VMD

Table 4
Dataset description after generation for four class dataset

	Original image count	After generation
ND	2240	4717
VMD	64	4052
MID	896	4892
MOD	3200	4960

Table 5
Dataset description after generation for four class dataset

	Original image count	After generation
MCI	198	2802
NC	493	2507
eMCI	204	2796
IMCI	61	2939
AD	145	2855

DEMNET for Classification

The task of the DNN is to automatically classify random MR images into different categories, leveraging the labeled images in the dataset for training. By analyzing the unique characteristics of each MR image and matching them with the appropriate dataset, the network progresses through its layers, transforming and transferring data from one layer to the next. This progressive learning makes the network more sophisticated and detailed with each layer. The most significant aspect of the DL model is its ability to learn independently, meaning it can quickly absorb information from the available data. Therefore, human knowledge or the quantity of minute features does not impact the network's learning process. We considered the DEMNET concept from Murugan et al. (2021). Figure 3 represents the configuration of the modified DEMNET architecture.

1. Input Layer: The Input Layer of the augmented MRI images can be compared to the base model, allowing for their integration.
2. Convolution Layer: The initial layers of the convolutional network gather information from the input image by applying a filter to it.
3. Pooling Layer: The primary objective of this layer is to minimize computational expenses by decreasing the image's spatial size and collecting trainable characteristics.
4. DEMNET Block: The DEMNET block is comprised of a series of two Convolutional Layers with ReLU activation, followed by a Batch Normalization Layer and a Maxpooling Layer.

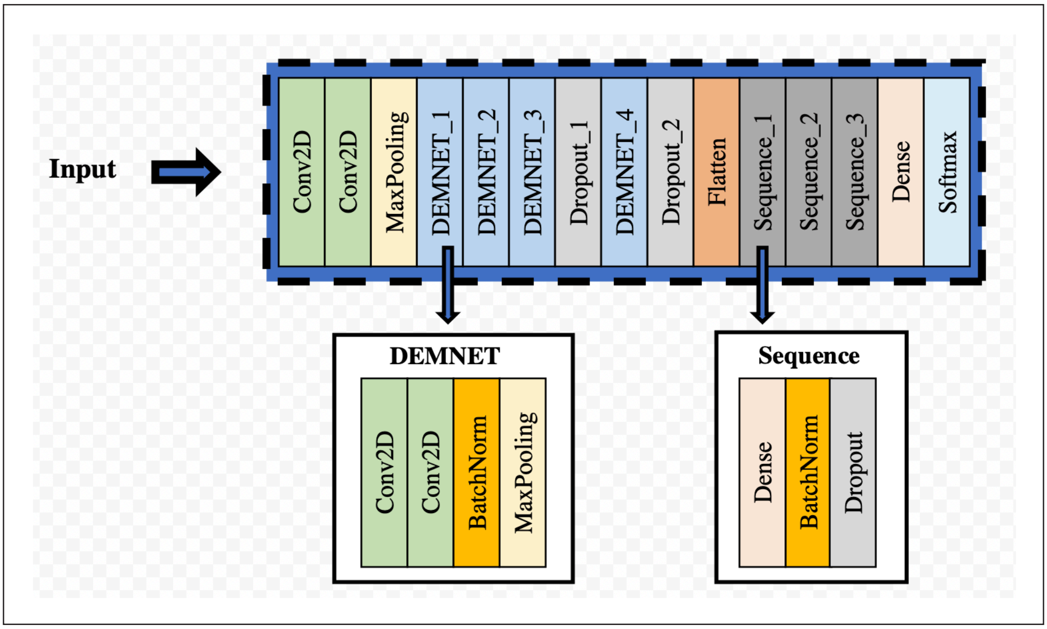


Figure 3. DEMNET Model with DEMNET block and sequence block

5. Dropout Layer: The dropout technique is a regularization method, miming the training process of multiple neural networks simultaneously, each with distinct architectures.
6. Sequence Block (Dense Block): The Sequence block is comprised of a Dense Layer with ReLU activation, followed by a Batch Normalization Layer and a Dropout Layer.
7. Dense Layer: The Dense Layer is a fully connected neural network layer where each neuron within the layer is intricately connected. It implies that every neuron in the Dense Layer receives input from all the neurons in the preceding layer.
8. Batch Normalization Layer: A CNN undergoes training using a gathered batch of input data rather than individual inputs. Similarly, Batch Normalization operates on batches, not singular inputs, to enhance the speed and stability of neural networks. This technique involves the addition of supplementary layers within a DNN.

To evaluate the performance of the modified DEMENT model, we employed a separate test dataset consisting of MRI data that was not used during the training phase. Samples from the training dataset were also utilized for validation purposes. In all scenarios, we generated results such as classification accuracy, AUC, loss, confusion matrix, and classification reports for both the Kaggle and ADNI datasets.

An overall workflow of the proposed framework is illustrated in Figure 1. The input undergoes the Data Preprocessing stage to resize the images for compatibility with the model. During this stage, we address the issue of Class-Imbalance by employing the Borderline technique to balance the dataset if any imbalance is detected. Once the dataset achieves class balance, it is divided into three sets: Training set, Validation set, and Test set, comprising 60%, 20%, and 20% of the dataset, respectively. After the dataset has been divided, the Training and Validation sets are forwarded to subsequent layers to train the model.

Borderline-DEMNET Flow

Table 6 provides an overview of the steps involved in the process in the form of an Algorithm. The process begins by loading the input data. Once the data is loaded, the MRI images undergo pre-processing. During this pre-processing stage, techniques such as Data Augmentation are applied, which involve initializing necessary parameters, adjusting zoom, brightness, rescaling, and other settings. A critical step at this stage was checking the balance of classes. If an imbalance is detected, the dataset is subjected to the Borderline SMOTE technique to generate additional images. After achieving balance, the images were resized to 64 * 64 dimensions. Table 7 gives the layers, kernel size, and parameters that workflow considers.

Table 6
Proposed Borderline-DEMNET flow

Input: MRI images different classes.
Ouput: Classification results including Accuracy, Precision, Recall and F1 scores.

begin
Step 1:
Load the MRI data
Step 2:
Data pre-processing of the MRI images

- Perform Data Augmentation
- Initialize the parameters required for data augmentation
- For each image, call function ImageDataGenerator to perform Zoom, Brightness Range, Horz Flip and Rescaling of the MRI images.
- Store the augmented image in the working directory
- Image Normalization
- Perform Over-Sampling of the images using Borderline SMOTE, as the classes are imbalanced
- Store the re-sampled data in synthetic samples and synthetic labels
- Concatenate the synthetic labels to augmented labels and synthetic samples to augmented samples
- Resize the pixels of the augmented images to size 64 * 64

Step 3:
Apply the Sequential() function to define the CNN model

- Building Model with ReLU as activation function
- Apply categorical-cross entropy
- RMSProp optimizer to train the model

Step 4:
Pass each MRI image to the convolution process

- Process each image of dimensions 64 * 64 * 3 processed depth-wise separate convolutions and convert the image into dimensions
- Drop Out the processed matrix/images by 0.5

Step 5:
The image is converted/fattened into a single-dimensional array
Step 6:
Apply the Dense layer with the softmax activation function and then apply dropout by 0.5 to the resultant array
Step 7:
Repeat Step 6 with a different set of neurons, apply dropout for repeated learning and activate the neurons
Step 8:
Plot the ACC and AUC curves for the trained model
Step 9:
Apply the confusion matrix, fetch the classification report results, and calculate the accuracy of test data
end

Table 7
The layers, kernel size and parameter details

Layer (type)	Output Shape	Param #
Conv2D+ReLU	(None, 64, 64, 16)	448
Conv2D+ReLU	(None, 64, 64, 16)	2320
MaxPooling2D	(None, 32, 32, 16)	0

Table 7 *(continue)*

Layer (type)	Output Shape	Param #
DEMNET Block_1	(None, 16, 16, 32)	14016
DEMNET Block_2	(None, 8, 8, 64)	55680
DEMNET Block_3	(None, 4, 4, 128)	221952
Dropout	(None, 4, 4, 128)	0
DEMNET Block_4	(None, 2, 2, 256)	886272
Dropout	(None, 2, 2, 256)	0
Flatten	(None, 1024)	0
Sequential_1 (Dense Block)	(None, 512)	526848
Sequential_2 (Dense Block)	(None, 128)	66176
Sequential_3 (Dense Block)	(None, 64)	8512
Dense	(None, 4)	260
Total params: 1,782,484 **Trainable params: 1,780,116** **Non-trainable params: 2,368**		

RESULTS AND DISCUSSION

Evaluation Metrics

Assessing the performance of a specific model is a crucial stage in creating a successful ML model. Various measures, known as Performance Metrics or Evaluation Metrics, are employed to gauge the model's effectiveness and quality. These performance indicators evaluate how effectively the model handles the provided data.

- Accuracy: Equation 1 is the primary metric to evaluate the model's performance in accurately predicting positive and negative events.

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

The value is classified as TP (True Positive) when the model accurately identifies the actual label and the image as normal. Similarly, the value is classified as TN (True Negative) when the algorithm correctly predicts an abnormal image and the actual label is also abnormal. In cases where the model predicts the image to be normal, but the actual label is incorrect, the value is classified as FP (False Positive). Conversely, when the algorithm predicts an abnormal image, but the actual label is normal, the value is classified as FN (False Negative).

- Precision: The Equation 2, denoted as PR, is determined by the proportion of accurately predicted positive observations compared to the total number of positive observations.

$$Precision = \frac{(TP)}{(TP+FP)} \quad [2]$$

- Recall: The parameter of Recall (REC) (Equation 3), alternatively referred to as sensitivity, evaluates the classifier's ability to identify all positive samples effectively.

$$Recall = \frac{(TP)}{(TP+FN)} \quad [3]$$

- F1 score: The performance of a classification model is assessed by calculating the F-score (Equation 4) or F1 Score, which considers the model's predictions specifically for the positive class.

$$F1score = \frac{(2*Precision*Recall)}{(Precision+Recall)} \quad [4]$$

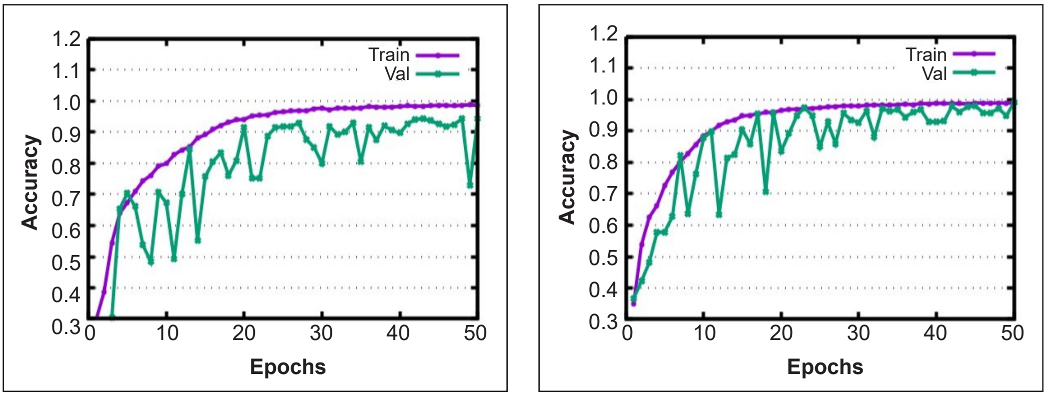
Experimental Setup

The proposed model was experimented with on Windows 10 edition with a device specification of Intel(R) Core(TM) i5-6200U CPU @ 2.30GHz–2.40 GHz with an 8.00 GB CPU. The suggested model underwent training using 50 epochs, a batch size 16, and an initial learning rate of 0.001. The required libraries for implementation were TensorFlow 2.7, Keras, Pandas, NumPy, and Matplotlib. The RMSProp optimizer was employed to train the algorithm. Additionally, the Area Under Curve (AUC) was computed for each epoch to evaluate the model's ability to accurately differentiate between positive and negative classes.

DEMNET with SMOTE and Borderline SMOTE for Four Class Dataset

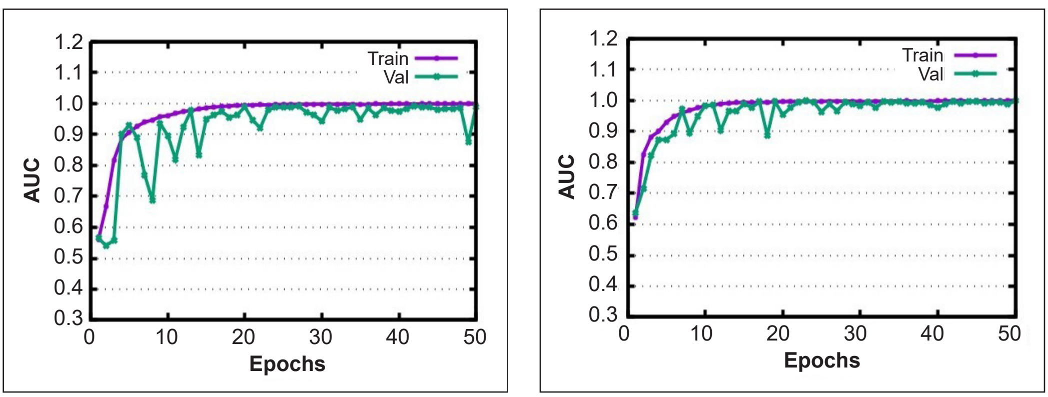
The related work presents the DEMNET model, which is compared to the performance analysis of the Borderline-DEMNET model. The evaluation of these models includes metrics such as accuracy, precision, recall, and F1-measure. During the training phase, the four-class dataset yielded an accuracy of 98.97%, while the accuracy during the validation process was 99.17%.

In Figure 4, a comparison is presented between the training and validation percentages of the DEMNET model and the Borderline-DEMNET model in terms of accuracy. In Figure 5, a comparison is presented between the training and validation percentages of the DEMNET model and the Borderline-DEMNET workflow in terms of AUC parameters. Similarly, Figure 6 presents the comparison of loss parameters. The results indicate that our model performs at least as well as or better than the base model due to Borderline SMOTE. Borderline SMOTE tackles the problem of imbalanced class distribution by specifically targeting minority class samples located near the decision boundary.



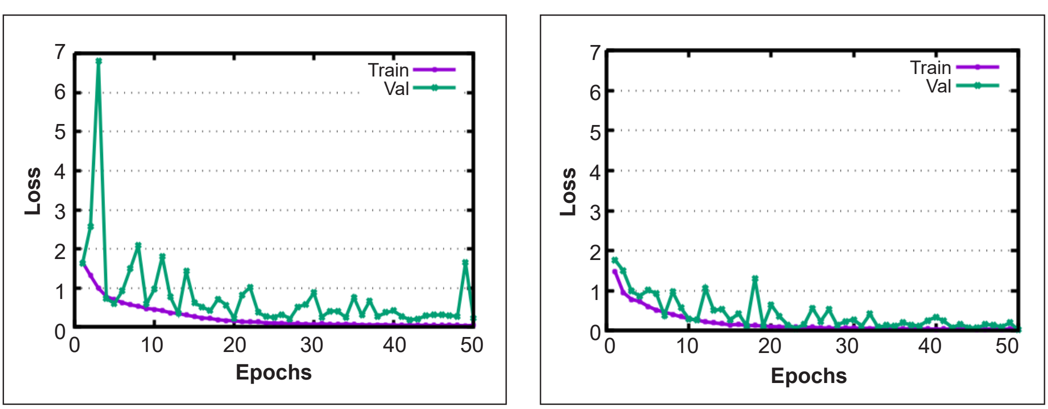
(a) (b)

Figure 4. Training and validation curves of accuracy obtained: (a) DEMNET; and (b) Borderline-DEMNET Workflow



(a) (b)

Figure 5. Training and validation curves of AUC obtained: (a) DEMNET; and (b) Borderline-DEMNET Workflow



(a) (b)

Figure 6. Training and validation curves of loss obtained: (a) DEMNET; and (b) Borderline-DEMNET Workflow

Figure 7 displays the average precision, recall, and F1-Score values, comparing the two workflows. Furthermore, the DEMNET model achieved a testing accuracy of 94% (Figure 8), while the Borderline-DEMNET workflow achieved a higher testing accuracy of 99.17% compared to the base model.

The performances of each class are presented in Tables 8 and 9, respectively, for both the workflows, which are used to calculate the average performance. The findings from classifying four separate categories indicate that the Borderline-DEMNET outperforms the DEMNET model in accuracy, precision, recall, and F1-measure.

Table 8
Performance indices for individual classes in DEMNET are evaluated

	Precision	Recall	F1-score
ND	0.98	0.96	0.97
VMD	1.00	1.00	1.00
MID	0.91	0.93	0.92
MOD	0.91	0.90	0.91

Table 9
Indices for the performance of individual classes in the new workflow

	Precision	**Recall**	**F1-score**
ND	1.00	1.00	1.00
VMD	1.00	1.00	1.00
MID	0.98	0.99	0.99
MOD	0.99	0.98	0.98

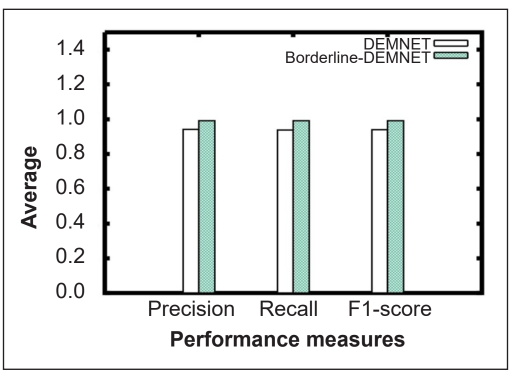


Figure 7. A comparison was conducted on the four-class dataset to evaluate the average precision, recall, and F1-score performance measures

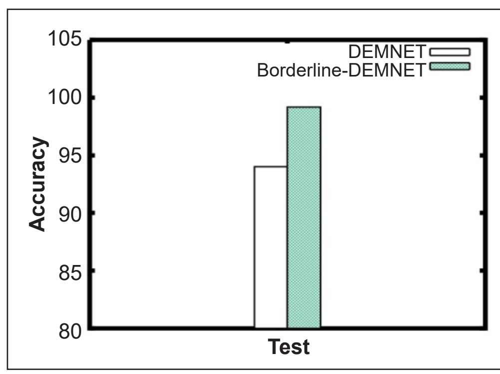


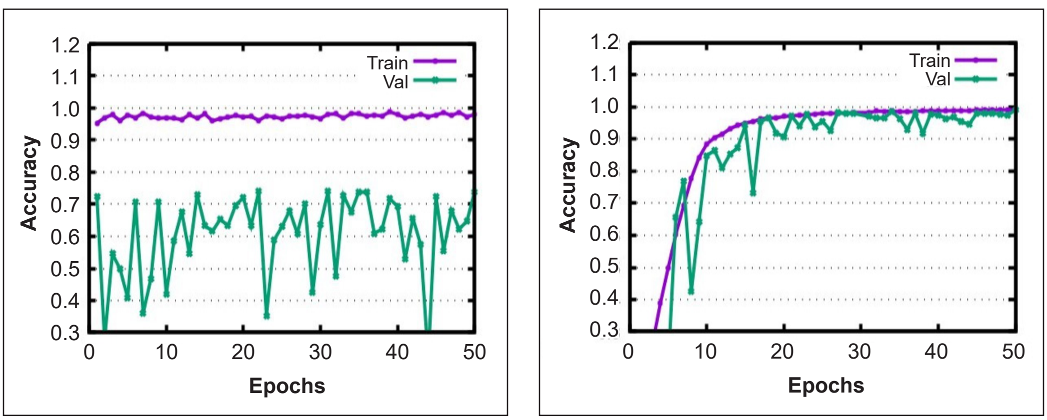
Figure 8. Comparison of testing accuracy among different workflows

Comparison of DEMNET with SMOTE and Borderline SMOTE for Five Class Dataset

In order to assess the effectiveness of the Borderline-DEMNET framework on additional MRI datasets related to AD, subsequent to achieving improved outcomes on the four-class dataset, an experiment was conducted. The experiment involved a five-class classification of AD based on data obtained from the ADNI database. The five AD classes included in the analysis were NC, eMCI, MCI, lMCI, and AD.

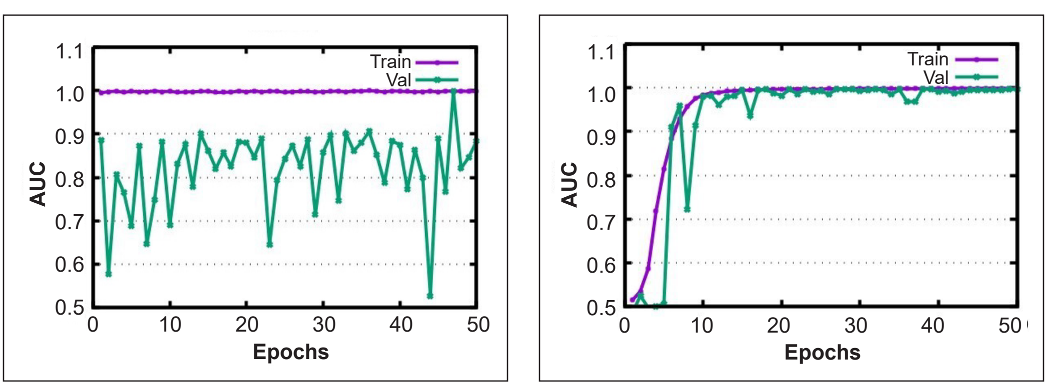
The evaluation of models involves using metrics such as accuracy, precision, recall, and F1-measure (Pushpa et al., 2013). The Borderline-DEMNET model was trained using a total of 1,780,181 parameters. During the training of the dataset, the model achieves a training accuracy of 99.25% and a validation accuracy of 99.14%. Figures 9, 10, and 11 illustrate both flows' training and validation curves. Upon comparing the two workflows, it is observed that our workflow performs on par with or better than the base model. Figure 12 comprehensively compares the average precision, recall, and F1-Score values.

The base architecture achieved training accuracy, validation accuracy, and testing accuracy rates of 97.89%, 73.74%, and 73.12%, respectively. In contrast, the proposed framework achieved higher accuracy rates, with 99.25% for training and 99.14% for validation and testing (Figure 13). In Figure 13, we added another model for comparison, which is known as D-DEMNET. It was our previous work on the same concept. Performance indices for individual classes can be found in Tables 10 and 11. The proposed workflow was also found to be effective in the five-class dataset.



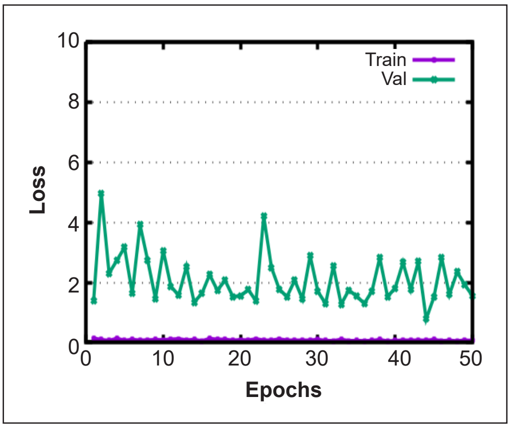
(a) (b)

Figure 9. The accuracy training and validation curves of the five-class dataset were analyzed: (a) DEMNET; and (b) Borderline-DEMNET Workflow

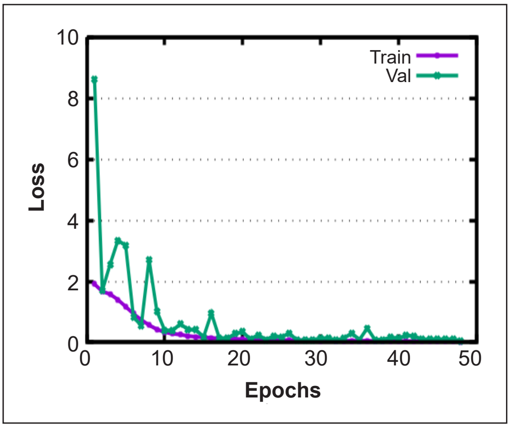


(a) (b)

Figure 10. The AUC training and validation curves were generated using a five-class dataset: (a) DEMNET Model; and (b) Borderline-DEMNET Workflow



(a)



(b)

Figure 11. The loss training and validation curves were generated using a five-class dataset: (a) DEMNET Model; (b) and Borderline-DEMNET Workflow

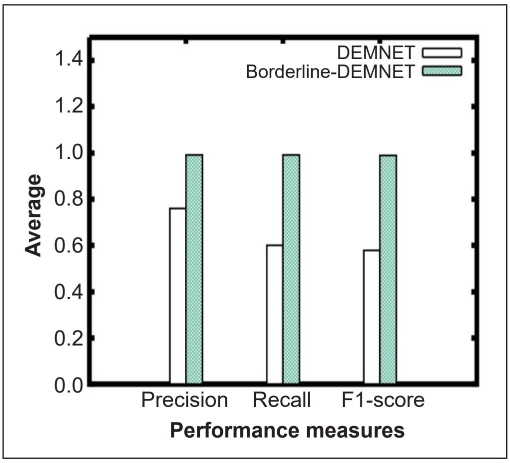


Figure 12. A study evaluated the performance of five-class datasets by comparing their average precision, recall, and F1-score

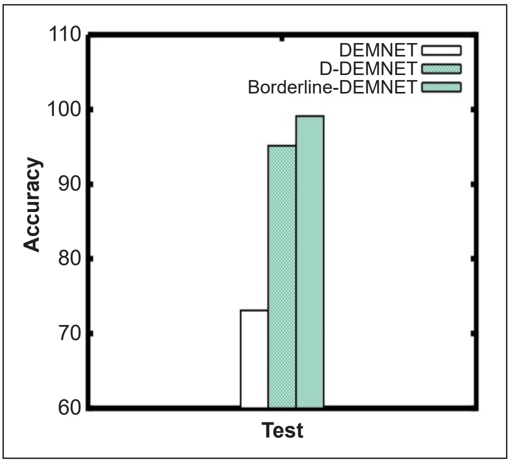


Figure 13. Comparison of testing accuracy among different workflows for a five-class dataset

Table 10
The performance indices for each class of the DEMNET Model were evaluated using a five-class dataset

	Precision	Recall	F1-score
AD	0.81	0.78	0.79
CN	0.38	0.89	0.53
eMCI	0.84	0.27	0.41
lMCI	0.87	0.91	0.89
MCI	0.94	0.15	0.27

Table 11
The performance metrics of the borderline-DEMNET architecture on a dataset consisting of five classes are evaluated

	Precision	Recall	F1-score
AD	1.00	0.99	1.00
CN	0.98	0.99	0.98
eMCI	0.99	1.00	0.99
lMCI	1.00	0.99	0.99
MCI	0.99	0.99	0.99

CONCLUSION

The paper presents a novel framework named DEMentia NETwork with Borderline technique (Borderline-DEMNET) for Multi-Class classification problems to early diagnose or detect AD/dementia. This framework proves that using better techniques in class balancing can also improve the model's accuracy. It is considered an extension of DEMNET, which is used for classifying different stages of dementia and AD as specified in the related work. This Borderline-DEMNET framework is evaluated and compared using evaluation matrices such as accuracy, precision, recall, and F1-score. The framework results in 99.17% accuracy for the four-class dataset, whereas the framework gives 99.14% accuracy for the five-class dataset. So, from the results, the paper can conclude that the Borderline-DEMNET framework performs better in terms of evaluation matrices considered in the paper.

ACKNOWLEDGEMENT

The collection and sharing of data for the Alzheimer's Disease Neuroimaging Initiative (ADNI) is funded by the National Institute on Aging through National Institutes of Health Grant U19 AG024904. The Northern California Institute for Research and Education serves as the grantee organization. ADNI has also historically received support from the National Institute of Biomedical Imaging and Bioengineering, the Canadian Institutes of Health Research, and private sector donations through the Foundation for the National Institutes of Health (FNIH). Generous contributions have been made by various organizations, including AbbVie, the Alzheimer's Association, the Alzheimer's Drug Discovery Foundation, Araclon Biotech, BioClinica, Inc., Biogen, Bristol-Myers Squibb Company, CereSpir, Inc., Cogstate, Eisai Inc., Elan Pharmaceuticals, Inc., Eli Lilly and Company, EuroImmun, F. Hoffmann-La Roche Ltd and its affiliate Genentech, Inc., Fujirebio, GE Healthcare, IXICO Ltd., Janssen Alzheimer Immunotherapy Research & Development, LLC., Johnson & Johnson Pharmaceutical Research & Development LLC., Lumosity, Lundbeck, Merck & Co., Inc., Meso Scale Diagnostics, LLC., NeuroRx Research, Neurotrack Technologies, Novartis Pharmaceuticals Corporation, Pfizer Inc., Piramal Imaging, Servier, Takeda Pharmaceutical Company, and Transition Therapeutics.

REFERENCES

ADNI. (2021). *Alzheimer's Disease Neuroimaging Initiative*. https://adni.loni.usc.edu/

Charan, M. (2022). *Alzheimer's Disease 5 Class Dataset ADNI.* Kaggle. https://www.kaggle.com/datasets/madhucharan/alzheimersdisease5classdatasetadni

Ali, A., Shamsuddin, S. M., & Ralescu, A. (2013). Classification with class imbalance problem: A review. *International Journal Advance Computing Application, 5*(3), 176-204.

Alinsaif, S., & Lang, J. (2021). 3D shearlet-based descriptors combined with deep features for the classification of Alzheimer's disease based on MRI data. *Computers in Biology and Medicine*, *138*, Article 104879. https://doi.org/10.1016/j.compbiomed.2021.104879

Alzheimer's Association. (2023). *Alzheimer's Disease Facts and Figures*. https://www.alz.org/alzheimers-dementia/facts-figures

Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., Filippi, M., & Alzheimer's disease neuroimaging initiative (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical, 21,* Article 101645. https://doi.org/10.1016/j.nicl.2018.101645

Basheera, S., & Ram, M. S. S. (2019). Convolution neural network–based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation. *Alzheimer's & Dementia: Translational Research & Clinical Interventions, 5*(1), 974-986. https://doi.org/10.1016/j.trci.2019.10.001

Bi, J., & Zhang, C. (2019). An empirical comparison on state-of-the-art multi-class imbalance learning algorithms and a new diversified ensemble learning scheme. *Knowledge-Based Systems, 158*, 81-93. https://doi.org/10.1016/j.knosys.2018.05.037

Han, Hui., Wang, Wen-Yuan & Mao, Bing-Huan. (2005). Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning. In D. S. Huang, X.P. Zhang, & G. B. Huang (Eds.), *Advances in Intelligent Computing. ICIC 2005. Lecture Notes in Computer Science* (Vol. 3644, p. 878-887). Springer. https://doi.org/10.1007/11538059_91

Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. E. (2005). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering, 30*(1), 25-36.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Laak, J. A. V. D., Ginneken, B. V., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42*, 60-88. https://doi.org/10.1016/j.media.2017.07.005

Liu, L., Wu, X., Li, S., Li, Y., Tan, S., & Bai, Y. (2022). Solving the class imbalance problem using ensemble algorithm: Application of screening for aortic dissection. *BMC Medical Informatics and Decision Making Volume, 22* Article 82. https://doi.org/10.1186/s12911-022-01821-w

Monsour, R., Dutta, M., Mohamed, A. Z., Borkowski, A., & Viswanadhan, N. A. (2022). Neuroimaging in the era of artificial intelligence: Current applications. *Federal Practitioner, 39*(1), 14-20. https://doi.org/10.12788/fp.0231

Murugan, S., Venkatesan, C., Sumithra, M. G., Gao, X. Z., Elakkiya, B., Akila, M., & Manoharan, S. (2021). DEMNET: A deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images. *IEEE Access*, *9*, 90319-90329. https://doi.org/10.1109/ACCESS.2021.3090474

National Institute on Aging. (2021). *Alzheimer's disease fact sheet*. https://www.nia.nih.gov/health/alzheimers-disease-fact-sheet

Neetha, P. U., Simran, S., Sunilkumar, G., Pushpa, C. N., Thriveni, J., & Venugopal, K. R. (2022). D-DEMNET framework for classifying the Alzheimer's disease for detection. In H. L. Gururaj, M. R. Pooja & F. Francesco (Eds.), *Recent trends in computational sciences. Proceedings of the Fourth Annual International Conference on Data Science, Machine Learning and Blockchain Technology* (pp. 1-5). CRC Press. https://doi.org/doi: 10.1201/9781003363781-3

Prajapati, R., Khatri, U., & Kwon, G. R. (2021). An efficient deep neural network binary classifier for Alzheimer's disease classification. In *2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)* (pp. 231-234). IEEE Publishing. https://doi.org/10.1109/ICAIIC51459.2021.9415212

Pushpa, C. N., Patil, A., Thriveni, J., Venugopal, K. R., & Patnaik, L. M. (2013, December). Web page recommendations using radial basis neural network technique. In *2013 IEEE 8th International Conference on Industrial and Information Systems* (pp. 501-506). IEEE Publishing. https://doi.org/10.1109/ICIInfS.2013.6732035

Pushpa, C. N., Thriveni, J., Venugopal, K. R., & LM, P. (2013). Web search engine based semantic similarity measure between words using pattern retrieval algorithm. *Computer Science & Information Technology*, 1-11. https://doi.org/10.5121/csit.2013.3101

Raju, M., Gopi, V. P., & S., A. V. (2021). Multi-class classification of Alzheimer's disease using 3DCNN features and multilayer perceptron. In *2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 368-373). IEEE Publishing. https://doi.org/10.1109/WiSPNET51692.2021.9419393

Raju, M., Thirupalani, M., Vidhyabharathi, S., & Thilagavathi, S. (2021, March). Deep learning based multilevel classification of Alzheimer's disease using MRI scans. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1084, No. 1, pp. 012017). IOP Publishing. https://doi.org/10.1088/1757-899X/1084/1/012017

Richhariya, B., Tanveer, M., Rashid, A. H., & Alzheimer's disease neuroimaging initiative. (2020). Diagnosis of Alzheimer's disease using universum support vector machine based recursive feature elimination (USVM-RFE). *Biomedical Signal Processing and Control*, *59*, Article 101903. https://doi.org/10.1016/j.bspc.2020.101903

Sarvesh, D. (2019). *Alzheimer's dataset (4 class of images).* Kaggle. https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images

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*, *8*(5), 145-153.

Sun, Y., Que, H., Cai, Q., Zhao, J., Li, J., Kong, Z., & Wang, S. (2022). Borderline SMOTE algorithm and feature selection-based network anomalies detection strategy. *Energies, 15*(13), Article 4751. https://doi.org/10.3390/en15134751

Wang, H., Shen, Y., Wang, S., Xiao, T., Deng, L., Wang, X., & Zhao, X. (2019). Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer's disease. *Neurocomputing, 333,* 145-156. https://doi.org/10.1016/j.neucom.2018.12.018

World Health Organization. (2023). *Dementia*. https://www.who.int/news-room/fact-sheets/detail/dementia

Zeng, H. M., Han, H. B., Zhang, Q. F., & Bai, H. (2021). Application of modern neuroimaging technology in the diagnosis and study of Alzheimer's disease. *Neural Regeneration Research, 16*(1), 73-79. https://doi.org/10.4103/1673-5374.286957