scientific reports

OPEN

Classifying and diagnosing Alzheimer's disease with deep learning using 6735 brain MRI images

Seyed Mohammad Mousavi[1,2], Khadijeh Moulaei[3,4]✉ & Leila Ahmadian[2]✉

Traditional diagnostic methods for Alzheimer's disease often suffer from low accuracy and lengthy processing times, delaying crucial interventions and patient care. Deep convolutional neural networks trained on MRI data can enhance diagnostic precision. This study aims to utilize deep convolutional neural networks (CNNs) trained on MRI data for Alzheimer's disease diagnosis and classification. In this study, the Alzheimer MRI Preprocessed Dataset was used, which includes 6735 brain structural MRI scan images. After data preprocessing and normalization, four models Xception, VGG19, VGG16 and InceptionResNetV2 were utilized. Generalization and hyperparameter tuning were applied to improve training. Early stopping and dynamic learning rate were used to prevent overfitting. Model performance was evaluated based on accuracy, F-score, recall, and precision. The InceptionResnetV2 model showed superior performance in predicting Alzheimer's patients with an accuracy, F-score, recall, and precision of 0.99. Then, the Xception model excelled in precision, recall, and F-score, with values of 0.97 and an accuracy of 96.89. Notably, InceptionResnetV2 and VGG19 demonstrated faster learning, reaching convergence sooner and requiring fewer training iterations than other models. The InceptionResNetV2 model achieved the highest performance, with precision, recall, and F-score of 100% for both mild and moderate dementia classes. The Xception model also performed well, attaining 100% for the moderate dementia class and 99–100% for the mild dementia class. Additionally, the VGG16 and VGG19 models showed strong results, with VGG16 reaching 100% precision, recall, and F-score for the moderate dementia class. Deep convolutional neural networks enhance Alzheimer's diagnosis, surpassing traditional methods with improved precision and efficiency. Models like InceptionResnetV2 show outstanding performance, potentially speeding up patient interventions.

Keywords Alzheimer, Artificial intelligence, Deep convolutional neural networks, Deep learning

Alzheimer's disease (AD) is a brain disorder that causes the progressive breakdown of nerve cells. It is characterized by the buildup of abnormal protein clumps called amyloid plaques and tangled fibers known as neurofibrillary tangles in specific areas of the brain, particularly the temporal lobe and neocortical regions[1]. These changes disrupt brain function and lead to the symptoms of Alzheimer's. Currently, around 50 million people worldwide are living with Alzheimer's, and this number is expected to rise to approximately 152 million by 2050[2]. The disease causes both structural and functional changes in the brain, leading to the death of nerve cells and the loss of brain tissue over time. Alzheimer's typically starts slowly and gradually worsens, affecting memory, thinking, and behavior. AD symptoms usually become noticeable after the age of 60. However, in some cases particularly for individuals with specific genetic mutations the disease can develop much earlier, between the ages of 30 and 50[3]. The onset and progression of Alzheimer's can be identified using clinical measures, but this is often a time-consuming process that requires specialized professionals to recognize the warning signs of the condition[4,5]. Additionally, early and accurate diagnosis of Alzheimer's necessitates a thorough evaluation by medical experts, including a detailed review of the patient's medical history, as well as physical and neurological examinations[4].

[1]Medical Informatics Research Center, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran. [2]Student Research Committee, Kerman University of Medical Sciences, Kerman, Iran. [3]Health Management and Economics Research Center, Health Management Research Institute, Iran University of Medical Sciences, Tehran, Iran. [4]Artificial Intelligence in Medical Sciences Research Center, Smart University of Medical Sciences, Tehran, Iran. ✉email: moulaei.kh91@gmail.com; ahmadianle@yahoo.com

Early diagnosis of Alzheimer's is crucial because it allows healthcare providers to offer preventive measures that can help slow the progression of the disease. It also helps patients become more aware of their condition, which can encourage them to cooperate with healthcare providers in adopting strategies to reduce the risk of disease progression[6]. Over the past few decades, neuroimaging data has been increasingly used to study AD. Researchers have applied machine learning and deep learning techniques to analyze this data, yielding promising results for improving the diagnosis of Alzheimer's[3]. One of the most widely used deep learning models for detecting and classifying Alzheimer's is called Convolutional Neural Networks (CNNs)[5]. CNNs are inspired by the human visual system and are similar to traditional neural networks, but they are specifically designed to work with two-dimensional image data, such as brain scans. This design allows CNNs to efficiently identify important patterns in images while reducing the number of parameters or variables that need to be processed, making them more efficient and effective. The structure of CNNs takes advantage of the spatial relationships within images, which helps reduce the complexity of the data and improves the network's ability to learn. This makes CNNs particularly good at tasks like medical imaging and disease diagnosis. For example, CNNs can extract complex, detailed features from brain scans, which can help identify signs of Alzheimer's disease. Additionally, CNNs are designed to reduce unnecessary information while preserving critical details in the images, which improves their performance and reduces the risk of overfitting (a problem where a model performs well on training data but poorly on new, unseen data)[5,6].

K. de Silva and H. Kunz[3] used a CNN to predict Alzheimer's disease from MRI brain images and achieved high accuracy, including an AUC score of 0.92. They optimized the CNN model by fine-tuning its hyperparameters and configurations to achieve the best performance. Their findings showed that using a CNN alongside MRI data can effectively diagnose Alzheimer's and offers a promising approach for early detection. Similarly, Liu et al.[5], developed an early detection system for Alzheimer's using two types of deep learning algorithms: CNN and Deep Metric Learning (DML), both applied to MRI images. They found that the DML model performed better at classifying early-stage Alzheimer's patients. Additionally, the DML algorithm significantly improved the clarity and quality of MRI images, enhanced the accuracy and stability of classifying early-stage Alzheimer's, sped up the model's learning process, and provided a new method for early prediction of Alzheimer's.

To address a critical gap in the existing literature, this study investigates the underexplored area of large-scale, multi-class classification of Alzheimer's disease stages using advanced deep learning models. Specifically, it emphasizes the lack of comparative analyses evaluating multiple state-of-the-art CNN architectures on a sizable dataset, while incorporating transfer learning and ensemble techniques to enhance diagnostic accuracy. To the best of our knowledge, no prior study has focused on classifying Alzheimer's disease stages using 6,735 brain MRI images by comparing VGG16, VGG19, InceptionResNetV2, and Xception models. While previous works have employed models such as DenseNet, ResNet, or stacked autoencoders, many were constrained by binary classification tasks, extensive preprocessing, or limited dataset sizes[7–10]. In contrast, our study offers a robust comparative framework across four powerful CNN architectures, applied to a four-class classification task—Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented on a large dataset derived from the ADNI repository. We leverage transfer learning with customized fully connected layers to enhance model performance and reduce training complexity. Additionally, a majority voting–based ensemble method is employed to further improve diagnostic accuracy. The proposed end-to-end deep learning framework minimizes manual feature engineering and demonstrates practical applicability for clinical decision support. The key contributions of this work are summarized as follows:

- We conduct a large-scale, multi-class classification of Alzheimer's disease using 6,735 MRI images from the ADNI dataset.
- We present a comparative evaluation of four state-of-the-art CNN architectures: VGG16, VGG19, InceptionResNetV2, and Xception.
- We leverage transfer learning and fine-tune model-specific architectures to improve classification performance and efficiency.
- We implement an ensemble learning approach using majority voting to boost diagnostic accuracy.
- We develop an end-to-end framework requiring minimal feature engineering, with direct implications for clinical decision support in Alzheimer's diagnosis.

Method

Study design and setting

In this study, the Multi Disease dataset[11] was used, which includes 6735 structural MRI brain scan images. This dataset was compiled from several publicly available sources[12–18] and has been specifically introduced for the diagnosis of AD. The distribution of the data within the dataset can be seen in Fig. 1.

This study was conducted through several key stages: pre-processing and data augmentation, the application of data analysis models, hyperparameter optimization, and models evaluation. These stages are described in detail below.

Pre-processing and data augmentation

Pre-processing is a crucial step in preparing a dataset for training a classification model. In this study, the preprocessing process involved several steps:

1. Normalization: The pixel values of the images were scaled to a range between 0 and 1 to standardize the data.
2. Resizing: The images were resized to match the input size required by the neural network.

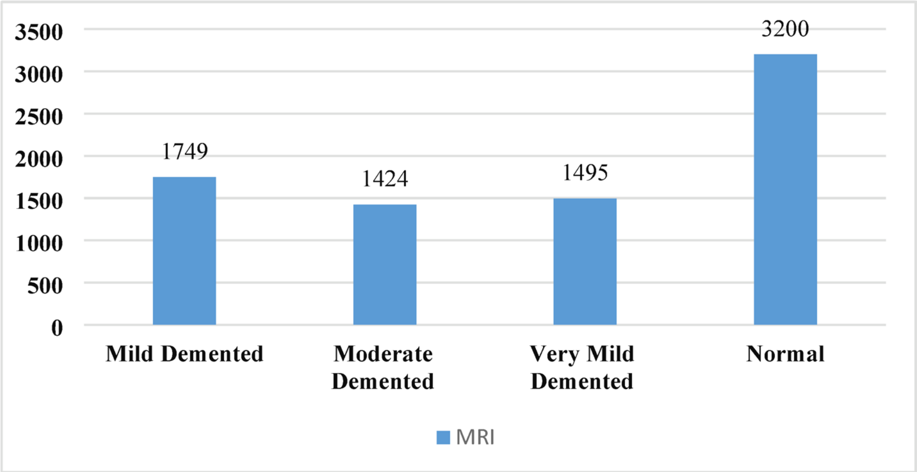

Fig. 1. Proposed dataset structure.

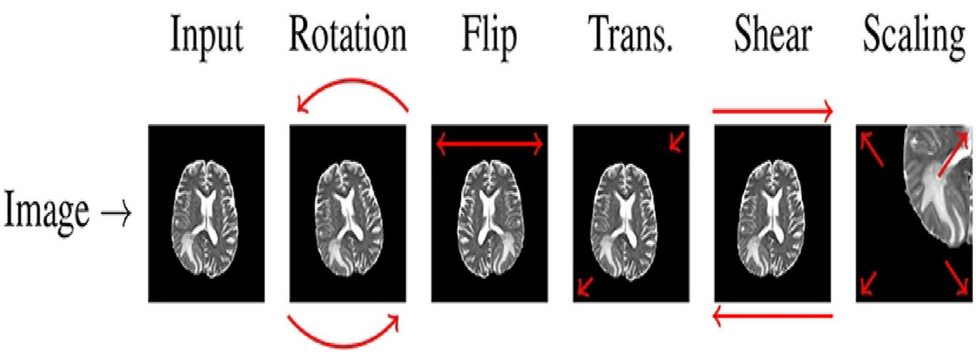

Fig. 2. Data augmentation process on an image.

3. Grayscale Conversion: To improve the learning process, the dataset images were first resized to the network's input dimensions. Then, all images were converted from color (RGB—Red, Green, Blue format, with three channels) to grayscale (single-channel).

After preprocessing, the dataset was randomly split into three subsets:

- Training set ($n = 4712$ images).
- Validation set ($n = 671$ images).
- Test set ($n = 1352$ images).

To further enhance the training process and improve classification performance, data augmentation was applied. Data augmentation involves applying various geometric transformations (e.g., rotations, flips, or shifts) to artificially increase the size and diversity of the training dataset. This helps the model generalize better to new data. Importantly, the test dataset remained unchanged during this process. Figures 2 and 3 show the results of applying data augmentation.

Data analysis models

Given the importance of accurately classifying MRI images of individuals with Alzheimer's disease, a convolutional neural network was used as a powerful method for this purpose. CNNs are one of the most reliable deep learning algorithms and are widely used for processing large volumes of data[19]. One of the key advantages of CNNs is that they do not require manual feature extraction, meaning the network automatically learns the most important features from the data. The architecture of a CNN is divided into two main parts: feature learning and classification. As shown in Fig. 4, CNNs are typically composed of three types of layers arranged in a hierarchical structure:

1. Convolutional Layer: Extracts features from the input images.
2. Pooling Layer: Reduces the size of the extracted features while retaining important information.
3. Fully Connected Layer: Uses the extracted features to classify the images into different categories.

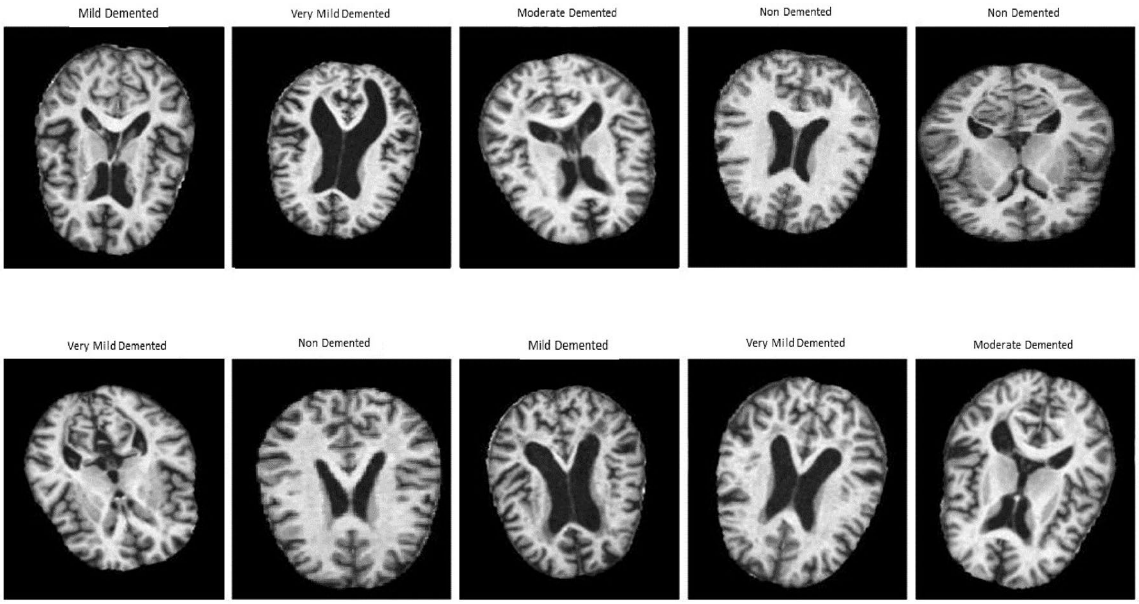

Fig. 3. Some images of the dataset after applying data augmentation.

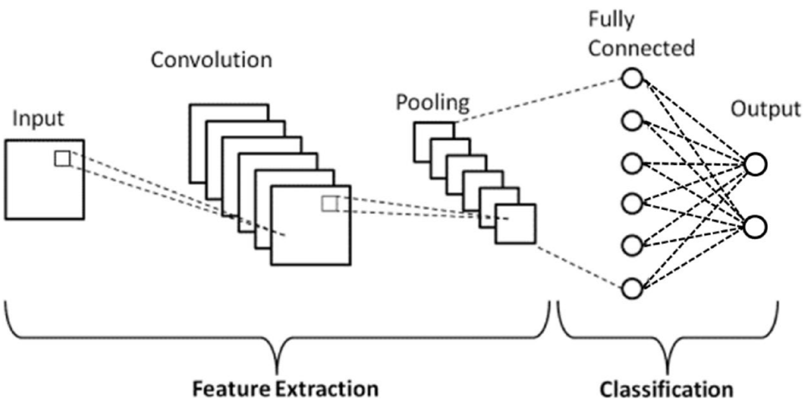

Fig. 4. Convolutional neural network architecture.

This layered structure allows CNNs to efficiently learn and classify complex patterns in data, making them highly effective for tasks like medical image analysis (Fig. 4).

The pre-processed images were used to train the proposed models based on CNNs. In this step, four pre-trained models Xception, VGG19, VGG16, and InceptionResNetV2 were utilized with transfer learning and fine-tuning of key parameters to extract the most relevant features from the images. Tsransfer learning allows these models to leverage knowledge from previously trained tasks, improving their performance on the new task. Moreover, to enhance the accuracy of the models for classification, a Global Average Pooling 2D layer was added at the end of each architecture, followed by three fully connected layers. After feature extraction, a Flatten layer was used to connect the feature extraction part of the network to the classification part. The classification part consisted of three fully connected layers with sizes 2048 and 1024, along with several Dropout layers placed at different positions to prevent overfitting. Also, the classification part of each network was responsible for categorizing the images into four classes based on the extracted features: Non-Demented, Mild Demented, Moderate Demented, and Very Mild Demented.

Therefore, to distinguish our work from previous studies[20–24], we implemented a standardized and comparative fine-tuning approach using these pre-trained architectures, tailored to the classification of Alzheimer's disease stages (Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented). In contrast to many earlier works that primarily performed binary classification (e.g., AD vs. healthy), we designed and evaluated a clinically relevant multi-class classification task. Moreover, for each model, we customized the architecture by adding a Global Average Pooling layer, multiple fully connected layers with specific neuron configurations (2048 and 1024), and Dropout layers in strategic positions to prevent overfitting. These modifications, combined with a consistent preprocessing pipeline and hyperparameter tuning (e.g., learning rate decay, early stopping), enabled a fair and optimized performance comparison. This methodological setup represents a novel experimental contribution, offering practical insights into the suitability and adaptability of these models for multi-class medical image classification in Alzheimer's diagnosis.

Hyperparameter optimization

To prevent the problem of overfitting, techniques such as generalization and hyperparameter tuning were used. Hyperparameters are settings that control the learning process of a neural network, and they are typically adjusted experimentally to optimize performance. For hyperparameter optimization, a wide range of common values for each parameter was evaluated, as shown in Table 1, to find the best combination. The ReLU and sigmoid activation functions were used, and the weights of the network were generated using the Adam optimizer with an initial learning rate of 0.003. The networks were trained for 30 epochs with a dynamic learning rate to improve accuracy and prevent overfitting. Additionally, a decaying learning rate was employed, which decreases over time to tune the model as training progresses.

During training, there comes a point where the model's performance no longer improves. To address this, an early stopping technique was used, which halts training when the validation error reaches its minimum. This ensures the model does not overfit by continuing to learn from noise in the training data. Training was conducted over 30 epochs using a decaying learning rate, with early stopping based on the lowest validation error. The structure of the proposed models can be seen in Table 2.

Models evaluation

In the last step, the proposed models were evaluated using standard criteria: accuracy, precision, recall, and F1-score. Additionally, the cost function was calculated. The following are these criteria and the cost function:

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

$$Precision = \frac{(TP)}{(TP + FP)} \tag{2}$$

$$Recall = \frac{(TP)}{(TP + FN)} \tag{3}$$

$$F - score = \frac{2 * (precision * recall)}{(precision + recall)} \tag{4}$$

$$CCE = -\sum_{k=0}^{m-1} y_k \log\left(\hat{y}_k\right) \tag{5}$$

True positive (TP): Correct Prediction for positive class.

True negative (TN): Correct Prediction for negative class.

False positive (FP): False prediction for negative class.

False negative (FN): False prediction for pos clas.

Result

The loss and accuracy for both the training and validation set are shown in Fig. 5 for all three models during the training process. The InceptionResNetV2 and VGG19 models learned faster than the other models, reached the early stopping point sooner, and required fewer training epochs to achieve optimal performance.

Table 3 shows the overall performance evaluation results for all four models. The InceptionResNetV2 model achieved the best performance based on the evaluation metrics, with an accuracy, F-score, recall, and precision of 0.99. The Xception model performed the best in terms of precision, recall, and F-score, with values of 0.97 and an accuracy of 96.89.

Table 4 shows the performance of each model for the four groups: non-demented, mild demented, moderate demented, and very mild demented. The deep learning models, particularly the InceptionResNetV2 and

Hyperparameters	
Activation function	ReLU, Sigmoid
Epochs	30
Batch size	64
Optimizer	Adam
Loss function	Categorical Cross Entropy
Drop out	0.5
Learning rate decay	5 Epochs
Early stopping	12 Epochs

Table 1. Hyperparameters of the proposed model.

#Parameter	Output	Layer
VGG16		
14,714,688	(None, 9, 9,512)	VGG16
0	(None, 2048)	global_ average_ pooling2d
0	(None, 100352)	flatten
0	(None, 100352)	dropout_1
525,312	(None, 1024)	dense_1
1,049,600	(None, 1024)	dense_2
0	(None, 1024)	dropout_2
4100	(None, 4)	dense_3
Total number of parameters: 16,293,730 Trainable parameters: 1,579,042 Non-trainable parameters: 14,714,688		
VGG19		
20,024,384	(None, 9, 9,512)	VGG19
0	(None, 512)	global_ average_pooling2d
0	(None, 512)	Flatten
0	(None, 512)	dropout_1
525,312	(None, 1024)	dense_1
1,049,600	(None, 1024)	dense_2
0	(None, 1024)	dropout_2
4100	(None, 4)	dense_3
Total number of parameters: 21,603,426 Trainable parameters: 1,579,042 Non-trainable parameters: 20,024,384		
inception_resnet_v2		
54,336,736	(None, 8, 8, 1536)	inception_resnet_v2
0	(None, 1536)	global_ average_pooling2d
0	(None, 1536)	flatten
0	(None, 1536)	dropout_1
1,573,888	(None, 1536)	dense_1
1,049,600	(None, 1024)	dense_2
0	(None, 1024)	dropout_2
4100	(None, 4)	dense_3
Total number of parameters: 56,964,354 Trainable parameters: 2,627,618 Non-trainable parameters: 54,336,736		
Xception		
12,642,880	(None, 9, 9, 1664)	Xception
0	(None, 1664)	global_ average_pooling2d
0	(None, 1664)	flatten
0	(None, 1664)	dropout_1
1,704,960	(None, 1024)	dense_1
1,049,600	(None, 1024)	dense_2
0	(None, 1024)	dropout_2
4100	(None, 4)	dense_3
Total number of parameters: 15,401,570 Trainable parameters: 2,758,690 Non-trainable parameters: 12,642,880		

Table 2. Types of layers and number of parameters used in the network.

Xception, demonstrated impressive performance in classifying Alzheimer's disease severity levels using MRI data. The InceptionResNetV2 model achieved the highest overall performance, with a precision, recall, and F-score of 100% for both the mild demented and moderate demented classes. Then Xception model also performed exceptionally well, attaining a precision, recall, and F-score of 100% for the moderate demented class and 99–100% for the mild demented class. Additionally, the VGG16 and VGG19 models showed strong results, with the VGG16 model reaching 100% precision, recall, and F-score for the moderate demented class.

Figure 6 also shows the results of classification of MRI images in the form of a confusion matrix.

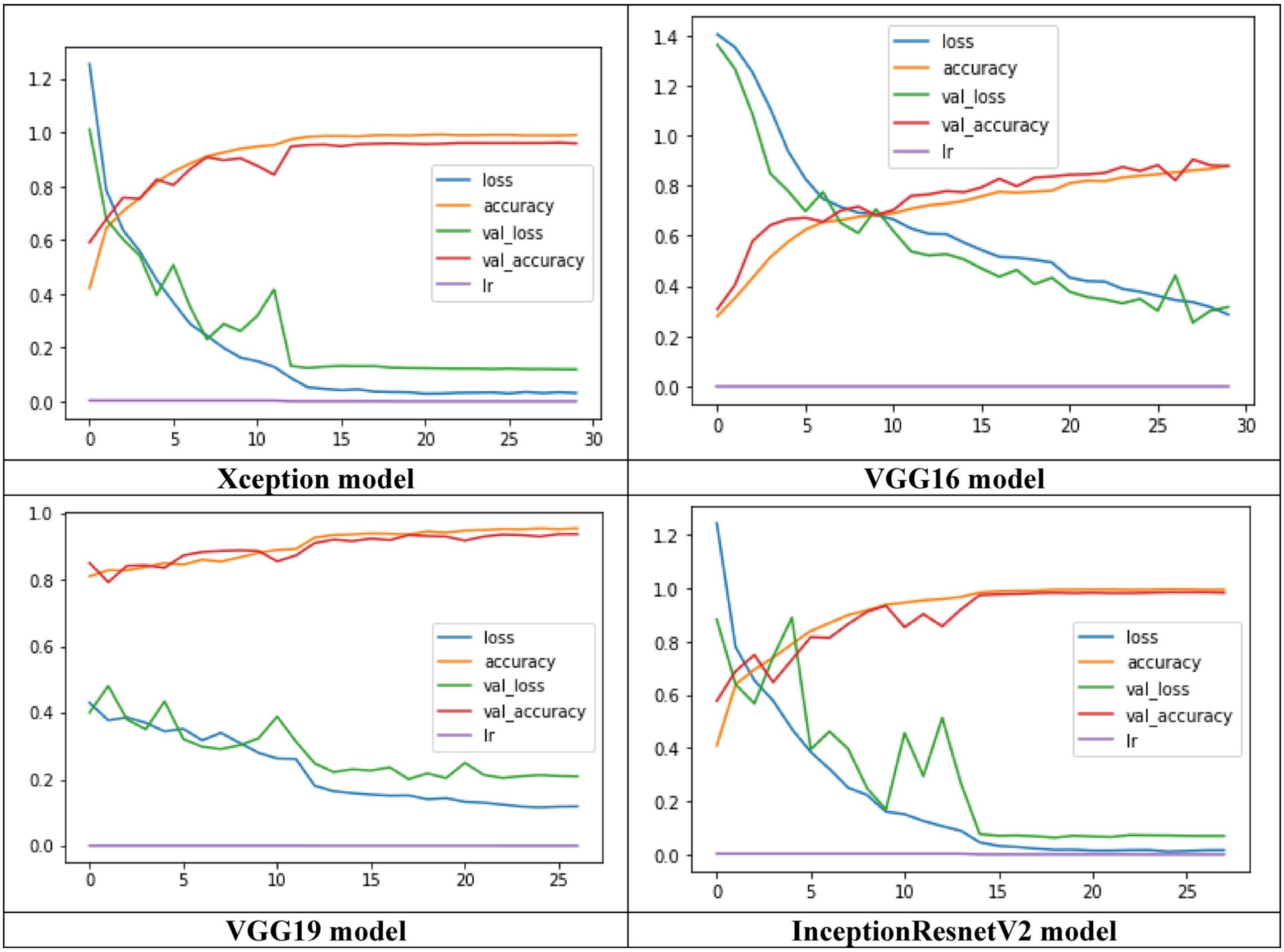

Fig. 5. Loss and accuracy values for training and validation sets during the model training process using MRI datasets.

Model	Precision(%)	Recall(%)	F-score(%)	Accuracy(%)
VGG16	89.0	89.0	89.0	89.20
VGG19	94.0	94.0	94.0	93.56
InceptionResNetV2	**99.0**	**99.0**	**99.0**	**98.52**
Xception	**97.0**	**0.97**	**0.97**	**96.89**

Table 3. Evaluation performance results obtained from VGG16, VGG19, inceptionResNetV2, Xception and inception V3 models. Note: Significant values are bolded to highlight top-performing models.

Discussion

In this study, we utilized deep convolutional neural networks to diagnose and classify Alzheimer's disease using MRI data. According to our findings, the InceptionResnetV2 model demonstrated superior performance in diagnosing Alzheimer's disease and distinguishing healthy individuals, as evidenced by its high scores in recall and F-score evaluation criteria. According to the study findings, the InceptionResNetV2 model demonstrated the best performance based on evaluation metrics. Then, the Xception model also performed exceptionally well in terms of precision, recall, and F-score. The InceptionResNetV2 model achieved the highest overall performance, reaching maximum precision, recall, and F-score for the mild and moderate demented classes. The Xception model also showed outstanding results, attaining top scores for the moderate demented class.

In our study, the InceptionResNetV2 model distinguished itself as the leading performer in detecting healthy individuals, outperforming other models in both precision and recall metrics. This effectiveness indicates its ability to accurately identify negative cases while minimizing false positives, which is vital for effective screening. These results align with the findings of Mondal et al., who also reported the superior capabilities of InceptionResNetV2 in differentiating between COVID-19 patients and healthy individuals in medical imaging tasks. Dash et al.[1], introduced a CNN model based on an enhanced InceptionResNetV2 architecture

Model	Class	Precision(%)	Recall(%)	F-score(%)
VGG16	Non demented	85	86	85
	Mild demented	93	98	96
	Moderate demented	**100**	**100**	**100**
	Very mild demented	80	73	76
VGG19	Non demented	90	92	91
	Mild demented	97	99	98
	Moderate demented	**100**	**100**	**100**
	Very mild demented	86	86	86
InceptionResnetV2	Non demented	98	98	98
	Mild demented	**100**	**100**	**100**
	Moderate demented	**100**	**100**	**100**
	Very mild demented	97	96	97
Xception	Non demented	95	95	95
	Mild demented	**99**	**100**	**100**
	Moderate demented	**100**	**100**	**100**
	Very mild demented	94	93	93

Table 4. Class-wise performance results for all the studied models using MRI datasets. Note: Significant values are bolded to highlight top-performing models.

for the segmentation and classification of the Cervical Transformation Zone (TZ). By integrating features from Reduction-A and Reduction-B, their model achieved notable results, including 81.24% accuracy, 81.24% sensitivity, 90.62% specificity, 87.52% precision, a 9.38% false positive rate (FPR), 81.68% F1 score, 75.27% Matthews Correlation Coefficient (MCC), and a 57.79% Kappa coefficient, surpassing other CNN approaches and existing methodologies. Baldassarre et al.[2], asserted that leveraging Inception-ResNet-v2 as a high-level feature extractor enhances the accuracy and precision of image content analysis, thereby facilitating the colorization process. Other study[25] investigated the use of deep learning algorithms for predicting Alzheimer's disease through MRI. They evaluated three transfer learning models: EfficientNetV2B1, InceptionResnetV2, and InceptionV3. The training accuracies were 87.12% for EfficientNetV2B1, 99.23% for InceptionResnetV2, and 98.40% for InceptionV3, with InceptionV3 achieving the highest accuracy on unseen test data. These findings collectively highlight the critical role of InceptionResNetV2 across various medical imaging applications, from disease detection to image segmentation and analysis. Its consistent superior performance across multiple studies underscores its potential as a transformative tool for enhancing healthcare diagnostics and image processing tasks.

Moreover, recent research has also explored the application of deep learning architectures in challenging medical imaging tasks beyond Alzheimer's disease. For instance, Ullah and Javed[26], conducted a comparative analysis of deep features for detecting COVID-19 using chest radiographs, showcasing the capability of convolutional neural networks in differentiating between complex pathological patterns. Similarly, another study developed ChestCovidNet, a deep learning-based model that effectively classified COVID-19, lung opacity, and pneumonia from chest X-ray images, further underscoring the strength of DL models in multi-class classification problems in radiographic data[27]. A separate study by Ullah et al.[28], introduced DeepCRINet, integrating multiple datasets to enhance the robustness of lung disease classification while employing occlusion sensitivity to improve interpretability. In the context of histopathological imaging, Ullah et al.[29], proposed a novel deep learning framework for classifying colon and lung cancers, demonstrating the versatility of CNN-based models across various medical image modalities. These studies reinforce the broader applicability and efficacy of deep learning models such as InceptionResNetV2 and Xception in addressing diverse and complex diagnostic tasks in medical imaging.

Collectively, these studies underscore the growing trend of employing advanced deep learning strategies such as data fusion, interpretability mechanisms, and specialized architectures—to address diagnostic complexities in diverse medical imaging domains. Their findings complement the present study by reaffirming the value of deep learning approaches like InceptionResNetV2 and Xception in managing heterogeneous datasets and enhancing classification accuracy across a spectrum of clinical scenarios. This body of work not only contextualizes the current findings but also highlights key methodological strategies, such as multi-dataset integration and model explainability, that may guide future research in the field.

Our study provides valuable insights into the diagnosis and classification of Alzheimer's disease. The Xception model demonstrated remarkable precision and high F-score values, highlighting its ability to accurately identify Alzheimer's cases while minimizing false positives. Additionally, it achieved the highest recall rate, underscoring its effectiveness in capturing true positive cases. These findings underscore the potential of deep learning models in facilitating the early and accurate diagnosis of Alzheimer's disease, a critical factor in implementing timely interventions and improving patient outcomes. Balancing precision and recall is particularly important, as early detection of Alzheimer's disease can slow its progression and enhance patients' quality of life. Various studies[30–32] have also confirmed that models like Xception and InceptionResNetV2 can significantly improve the timely and accurate diagnosis of diseases. For instance, Aparna et al.[33], introduced a hybrid deep learning framework

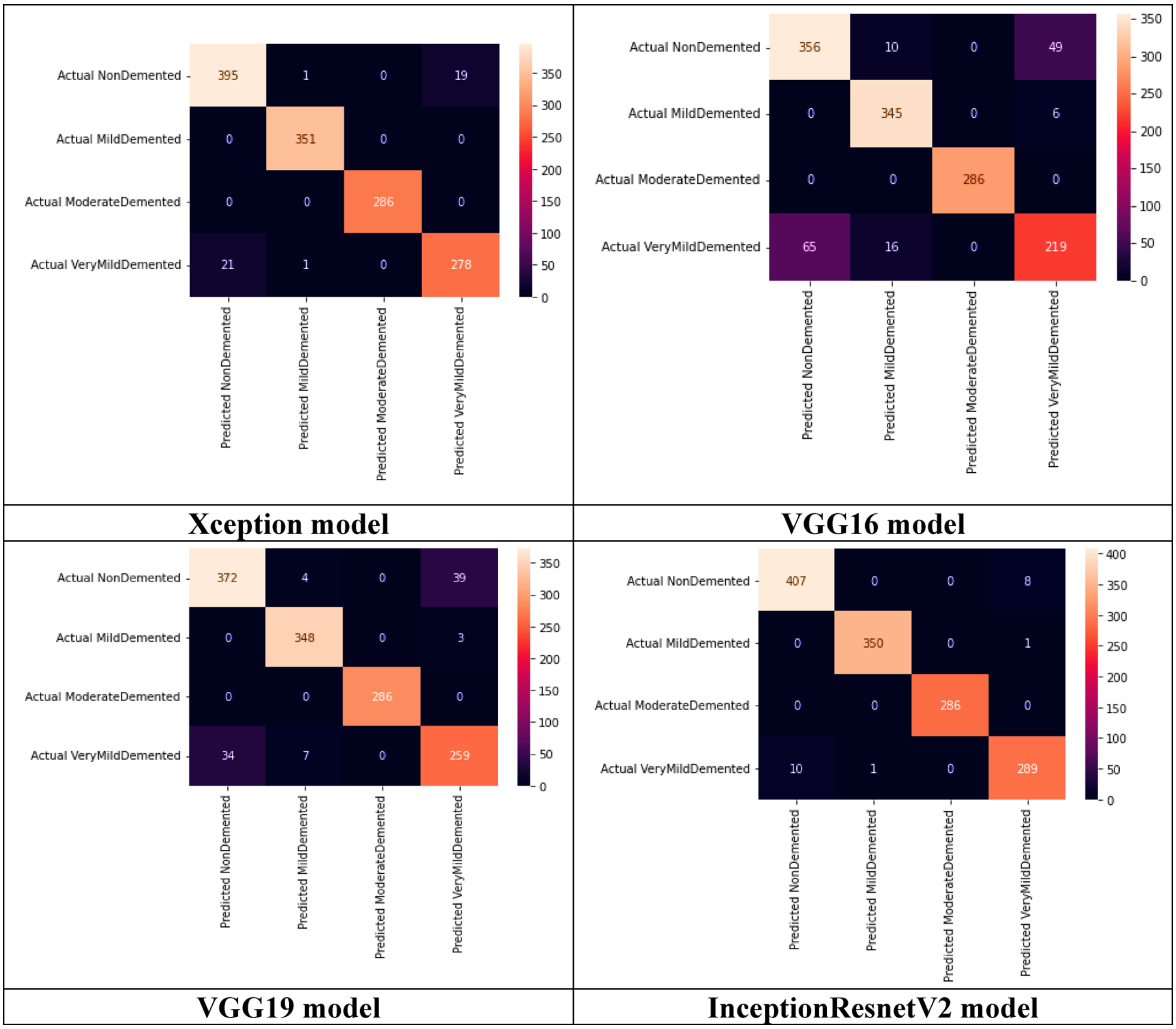

Fig. 6. Confusion matrix of MRI test set.

combining Xception and FractalNet to classify Alzheimer's disease into five stages. Using a Unet + + network for brain tissue segmentation, the Xception model extracted high-level features, achieving an impressive accuracy of 99.06% in multi-class AD classification, outperforming conventional and state-of-the-art methods.

Furthermore, Dhillon et al.[34], employed the Xception deep learning algorithm within an ensemble approach for Alzheimer's disease detection. This methodology significantly enhanced model performance, achieving an accuracy of 98.5%, sensitivity of 98%, precision of 95%, and an F1-score of 96.47%. These results highlight the potential of ensemble-based deep learning models in enhancing predictive capabilities for early Alzheimer's diagnosis, addressing the critical need for effective diagnostic techniques in healthcare. Another study by Kumar Singh et al.[35], utilized the Xception model as one of four pre-trained CNNs to extract features from MRI brain scans for Alzheimer's diagnosis. Their approach, which integrated features from Xception with those from ResNet50, VGG16, and InceptionV3, achieved an accuracy of 91% on the DS-160 dataset, demonstrating the effectiveness of deep learning models in the early detection and diagnosis of Alzheimer's disease. Additionally, another study[36] incorporated the Xception pre-trained CNN model into its methodology for detecting Alzheimer's disease using MRI brain images. By leveraging transfer learning and fine-tuning techniques, the research achieved significant accuracy improvements in AD detection. Although the VGG19 model attained a perfect accuracy score of 100%, Xception was also explored, reinforcing the potential of advanced deep learning techniques to enhance early Alzheimer's diagnosis. Overall, these studies emphasize the crucial role of deep learning models, particularly Xception, in facilitating the rapid and accurate diagnosis of Alzheimer's disease. By providing a well-balanced combination of precision and recall, these models serve as valuable tools for healthcare practitioners in managing neurodegenerative diseases. Therefore, the further development and refinement of these models could significantly enhance healthcare quality and treatment outcomes for Alzheimer's patients.

Finally, it should be noted that, despite the high performance of deep learning models in diagnosing Alzheimer's disease in the present study, one of the main challenges in deploying these models in clinical settings is their inherently "black-box" nature. Deep models typically lack transparency in decision-making processes, which can hinder their acceptance by healthcare professionals. In critical applications such as diagnosing brain disorders using MRI, Explainable Artificial Intelligence (XAI) plays a pivotal role in building trust, supporting clinical decision-making, and meeting ethical requirements. XAI techniques can provide insights into which features or regions of the image contributed most to the model's predictions, enabling clinicians to better understand and evaluate the model's reasoning. Ullah et al.[37] demonstrated the successful application of XAI in medical imaging by introducing the DeepEBTDNet model and applying the LIME method for brain tumor detection, effectively improving clinicians' understanding of the model's behavior. Furthermore, in another study, Ullah et al.[38], emphasized the significance of XAI by exploring its importance, stages, challenges, and output types across various domains, particularly healthcare. Their work highlights the necessity of integrating interpretability into deep learning frameworks to facilitate broader clinical adoption.

In the context of Alzheimer's disease diagnosis, implementing techniques such as Grad-CAM, SHAP, and LIME can provide valuable visual explanations of the brain regions influencing the model's decisions. This transparency allows clinicians to verify or question model outputs, increasing confidence in the results. Therefore, it is recommended that future research incorporate XAI frameworks as part of the model development process. Doing so would not only enhance the clinical reliability and acceptance of deep learning models but also represent a significant step toward developing trustworthy tools for the early and accurate diagnosis of neurodegenerative diseases such as Alzheimer's.

Study limitations

The study's reliance on the Alzheimer MRI Preprocessed Dataset, consisting of 6735 brain structural MRI scan images, may limit the generalizability of findings to other datasets or populations due to potential dataset-specific characteristics or biases. To address this limitation, future research could consider validating findings using multiple datasets representing diverse populations. Additionally, expanding the dataset size or incorporating data from different sources could enhance the robustness and generalizability of the study's conclusions.

Conclusion

This study investigated the efficacy of deep convolutional neural networks in improving Alzheimer's disease diagnosis using MRI data. Deep convolutional neural networks trained on MRI data offer a promising avenue for enhancing the diagnosis and classification of Alzheimer's disease. This study highlights the superiority of models like InceptionResNetV2 and Xception, which demonstrate remarkable performance in predicting Alzheimer's patients and potentially expediting interventions. Leveraging deep learning techniques could revolutionize Alzheimer's diagnosis, paving the way for more efficient and accurate patient care pathways.

Integrating deep convolutional neural networks into Alzheimer's diagnosis represents a significant advancement in medical imaging technology. The exceptional performance of models such as InceptionResNetV2 and Xception underscores the potential of AI-driven approaches to transforming disease diagnosis and management. By surpassing traditional methods in precision and efficiency, these models hold great promise for accelerating patient interventions and improving overall healthcare outcomes for Alzheimer's disease.

Data availability

The data is available through the following link: https://www.kaggle.com/datasets/praneshkumarm/multidiseasedataset.

Received: 20 March 2025; Accepted: 18 June 2025
Published online: 02 July 2025

References

1. Dash, S., Sethy, P. K. & Behera, S. K. Cervical transformation zone segmentation and classification based on improved Inception-ResNet-V2 using colposcopy images. *Cancer Inform.* **22**, 11769351231161477 (2023).
2. Ferreira, C. A. et al. Classification of breast cancer histology images through transfer learning using a pre-trained inception resnet v2. in *International conference image analysis and recognition* 763–770 (Springer, 2018).
3. De Silva, K. & Kunz, H. Prediction of alzheimer's disease from magnetic resonance imaging using a convolutional neural network. *Intelligence-Based Med.* **7**, 100091 (2023).
4. Basaia, S. et al. Automated classification of alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clin.* **21**, 101645 (2019).
5. Bi, X., Liu, W., Liu, H. & Shang, Q. Artificial intelligence-based MRI images for brain in prediction of alzheimer′ s disease. *J. Healthc. Eng.* **2021**(1), 8198552 (2021).
6. Srivastava, S., Ahmad, R. & Khare, S. K. Alzheimer's disease and its treatment by different approaches: A review. *Eur. J. Med. Chem.* **216**, 113320 (2021).
7. Balne, S. & Elumalai, A. Machine learning and deep learning algorithms used to diagnosis of alzheimer's. *Mater.Today: Proc.* **47**, 5151–5156 (2021).
8. Liu, S. et al. Early diagnosis of Alzheimer's disease with deep learning. in *2014 IEEE 11th international symposium on biomedical imaging (ISBI): 2014* 1015–1018 (IEEE, 2014).
9. Al Shehri, W. Alzheimer's disease diagnosis and classification using deep learning techniques. *PeerJ Comput. Sci.* **8**, e1177 (2022).
10. Yamanakkanavar, N., Choi, J. Y. & Lee, B. MRI segmentation and classification of human brain using deep learning for diagnosis of alzheimer's disease: a survey. *Sensors* **20**(11), 3243 (2020).
11. Multidiseasedataset https://www.kaggle.com/datasets/praneshkumarm/multidiseasedataset.
12. Adni dataset https://adni.loni.usc.edu/.
13. Alzheimers https://www.alzheimers.net/.
14. MRI and alzheimers https://www.kaggle.com/datasets/jboysen/mri-and-alzheimers.
15. alzheimers disease. and healthy dataset https://catalog.data.gov/dataset/alzheimers-disease-and-healthy-aging-data.

16. Bae, J. B. et al. Identification of alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. *Sci. Rep.* **10**(1), 22252 (2020).
17. Kamal, M. S. et al. Alzheimer's patient analysis using image and gene expression data and explainable-AI to present associated genes. *IEEE Trans. Instrum. Meas.* **70**, 1–7 (2021).
18. The European Prevention of Alzheimer's Dementia. https://cordis.europa.eu/article/id/429468-the-final-epad-dataset-is-now-available-on-the-alzheimer-s-disease-workbench.
19. Owais, M. et al. Light-weighted ensemble network with multilevel activation visualization for robust diagnosis of COVID19 pneumonia from large-scale chest radiographic database. *Appl. Soft Comput.* **108**, 107490 (2021).
20. Islam, J. & Zhang, Y. A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data. in *Brain Informatics: International Conference, BI 2017, Beijing, China, November 16-18, 2017, Proceedings: 2017* 213–222 (Springer, 2017).
21. Helaly, H. A., Badawy, M. & Haikal, A. Y. Deep learning approach for early detection of alzheimer's disease. *Cogn. Comput.* **14**(5), 1711–1727 (2022).
22. Mandal, P. K. & Mahto, R. V. Deep multi-branch CNN architecture for early alzheimer's detection from brain MRIs. *Sensors* **23**(19), 8192 (2023).
23. Oktavian, M. W., Yudistira, N. & Ridok, A. Classification of Alzheimer's disease using the Convolutional Neural Network (CNN) with transfer learning and weighted loss. ArXiv Preprint *arXiv:220701584* (2022).
24. Savaş, S. Detecting the stages of alzheimer's disease with pre-trained deep learning architectures. *Arab. J. Sci. Eng.* **47**(2), 2201–2218 (2022).
25. Jha, S., Mahajan, A. D., Thenraj, P. G., Karuppasamy, S. & Lakshminarayanan, S. Comparative Evaluation of Transfer Learning Models on Dementia Prediction. in *6th International Conference on Electronics, Communication and Aerospace Technology: 2022* 1520–1526 (IEEE, 2022).
26. Ullah, N. & Javed, A. Deep features comparative analysis for COVID-19 detection from the chest radiograph images. in *2021 International Conference on Frontiers of Information Technology (FIT): 2021* 258–263 (IEEE, 2021).
27. Ullah, N. et al. Chestcovidnet: An effective dl-based approach for covid-19, lung opacity, and pneumonia detection using chest radiographs images. *Biochem. Cell Biol.* (2024).
28. Ullah, N., Khan, J. A., De Falco, I. & Sannino, G. Bridging clinical gaps: multi-dataset integration for reliable multi-class lung disease classification with DeepCRINet and occlusion sensitivity. in *2024 IEEE Symposium on Computers and Communications (ISCC): 2024*1–6 (IEEE, 2024).
29. Ullah, N., De Falco, I. & Sannino, G. A novel deep learning approach for colon and lung cancer classification using histopathological images. in *2023 IEEE 19th International Conference on e-Science (e-Science): 2023* 1–10 (IEEE, 2023).
30. Suganthe, R. et al. Multiclass classification of alzheimer's disease using hybrid deep convolutional neural network. *NVEO-Nat. Volatiles Essent. Oils J.* 145–153 (2021).
31. Sharma, C. M., Goyal, L., Chariar, V. M. & Sharma, N. Lung disease classification in CXR images using hybrid inception-ResNet-v2 model and edge computing. *Journal of Healthcare Engineering* 2022. (2022).
32. Rashid, J., Qaisar, B. S., Faheem, M., Akram, A. & Amin Ru, Hamid, M. Mouth and oral disease classification using InceptionResNetV2 method. *Multimed. Tools Appl.* 1–19 (2023).
33. Aparna, M. & Rao, B. S. Xception-Fractalnet: hybrid deep learning based Multi-Class classification of alzheimer's disease. *Comput. Mater. Cont.* **74**(3) (2023).
34. Dhillon, A., Sharma, C., Kaur, K. & Kumar, N. An Ensembled Deep Learning Approach for the Prediction of Alzheimer Disease. in *2024 International Conference on Cybernation and Computation (CYBERCOM): 2024* 409–414 (IEEE, 2024).
35. Singh, S. K. et al. An Automated Deep Learning Model for Dignosis of Alzheimer's Disease using Deep Feature Fusion. in *14th International Conference on Computing Communication and Networking Technologies (ICCCNT): 2023* 1–7 (IEEE, 2023).
36. Prasetyo, S. Y. Alzheimer's Disease Identification through Brain MRI Images: Leveraging Pretrained CNN Models. in *IEEE 10th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA): 2024* 321–325 (IEEE, 2024).
37. Ullah, N., Hassan, M., Khan, J. A., Anwar, M. S. & Aurangzeb, K. Enhancing explainability in brain tumor detection: A novel DeepEBTDNet model with LIME on MRI images. *Int. J. Imaging Syst. Technol.* **34**(1), e23012 (2024).
38. Ullah, N., Khan, J. A., De Falco, I. & Sannino, G. Explainable artificial intelligence: Importance, use domains, stages, output shapes, and challenges. *ACM Comput. Surv.* **57**(4), 1–36 (2024).

Acknowledgements
The authors express their sincere gratitude to the Student Research Committee of Kerman University of Medical Sciences. This study was supported by Kerman University of Medical Sciences (Grant: 403000428). The funder had no role in study design, data gathering, or analysis.

Author contributions
Conceptualization: S.M.M., L.A., K.M. Data curation: S.M.M., K.M. Funding: S.M.M., L.A., Project administration: S.M.M., L.A. Design and Modeling: S.M.M., K.M. Resources: S.M.M., L.A., K.M. Supervision: S.M.M., L.A. Writing—original draft: S.M.M., L.A., K.M. Writing—review & editing: S.M.M., K.M.

Funding
We hereby declare that no financial support was received for the conduct of this research.

Declarations

Competing interests
The authors declare no competing interests.

Ethical approval
The ethics committee of Kerman University of Medical Sciences provided the code of ethics under number IR.KMU.REC.1404.089.

Additional information
Correspondence and requests for materials should be addressed to K.M. or L.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.