Dean&Francis

ISSN 2959-409X

Deep Learning Approaches for Early Alzheimer's Detection Using Multimodal Imaging

Kailai Bao*

School of Electronics and Information Engineering, Hangzhou Dianzi University, Hangzhou, China

*Corresponding author: 18042007@hdu.edu.cn

Abstract:

This work explores the utility of multimodal imaging techniques such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) in the early detection of Alzheimer's disease (AD), as well as the importance of machine learning, notably convolutional neural networks (CNNs). In order to enhance the diagnostic performance of brain scans, the study employed a range of models, such as Visual Geometry Group 16 (VGG16), EfficientNetB7, and a hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) architecture, to collect both spatial and temporal data. Using datasets from Kaggle and the Alzheimer Disease Neuroimaging Initiative (ADNI), the models achieved very high accuracy. A 98.68% overall accuracy and an F-Score of 98.95% were shown by VGG16 and EfficientNetB7. The receiver operating characteristic (ROC) curve study demonstrated strong discriminating abilities, with an area under the curve (AUC) ranging from 0.80 to 0.83. The CNN-LSTM hybrid model effectively handled long-term dependencies in picture data, leading to better performance. These results show that deep learning has a great deal of promise for aiding in the early diagnosis of AD and providing a useful tool for enhancing diagnostic precision in clinical settings. Subsequent investigations may broaden the dataset and improve algorithms to amplify prediction resilience.

Keywords: Alzheimer's disease; spatial and temporal data; CNN-LSTM hybrid model.

1. Introduction

Sixty to eighty percent of dementia cases are thought to be caused by Alzheimer's Disease (AD), a neurodegenerative dementia [1, 2]. The complex etiology of AD involves the interaction of several factors that increase the risk of developing the disease: aging, genetic factors (particularly family history and specific gene mutations), lifestyle choices, and health conditions (such as obesity, diabetes, hypertension, cardiovascular disease, and physical inactivity). The disease's unique pathophysiology includes nerve cell death and a progressive loss of connections between brain neurons, which results in atrophy of the brain's tissue. AD symptoms increase from modest memory impairment to severe cognitive decline, including memory loss [3, 4], disorganized thinking, impaired judgment, verbal communication deficits, impaired visuospatial skills, mood swings (e.g., depression, anxiety, irritability), and personality changes, which ultimately affect the patient's ability to perform daily living activities and self-care, making it difficult for the patient to perform independently daily activities such as personal hygiene, eating, and safe mobility, and these symptoms gradually worsen, severely affecting patients' quality of daily life [5, 6].

The ability to differentiate between AD, normal cognitive aging, the Multi-Country Inventory (MCI), and other forms of dementia accurately calls for the clinical expertise of memory disorder specialists. However, patients' and families' timely access to memory clinics is frequently restricted, particularly in industrialized nations' remote rural areas and those whose economies are still in the early stages of growth, where a key issue is the absence of trained practitioners [7, 8]. In addition, the United States is anticipated to have a shortage of credentialed clinicians (such neurologists) in the coming decades as the demand for capable medical professionals grows. Due to a lack of medical professionals and increasing clinical demand, machine learning techniques are beginning to be used in neurological diagnostics [9]. An earlier study [10] demonstrated an interpretable deep learning approach that was able to distinguish between age-appropriate Normal Cognition (NC) participants and AD patients using age, gender, and Mini-Mental State Examination (MMSE) data, adding to the high diagnostic accuracy reported by other research groups [11].

The purpose of this study is to look at the possibility of using machine learning techniques for early AD diagnosis. It emphasizes the application of Convolutional Neural Network (CNN) models in particular, as well as the role played by sophisticated designs like Visual Geometry Group 16 (VGG16) and EfficientNetB7, which are both CNN versions that have been optimized, in the identification process. These models are designed to extract crucial features from medical images, such as brain scans, allowing for precise detection of early-stage Alzheimer's. VGG16 [12] is recognized for its depth and simplicity, while EfficientNetB7 [13] is notable for its scaling techniques that improve accuracy without significantly increasing computational costs. Furthermore, the paper examines the combination of CNN with Long Short-Term Memory (LSTM) networks, demonstrating how this hybrid model enhances image recognition tasks by capturing spatial and temporal patterns from sequential data, such as a series of brain scan images over time [14]. This study contributes by analyzing the relative performance of these models, comparing their strengths and weaknesses in the context of Alzheimer's detection. The findings provide insights into how machine learning, particularly deep learning models, can be optimized for medical applications, emphasizing both accuracy and computational efficiency.

2. Methodology

2.1 Dataset Description

When studying brain diseases, various brain imaging techniques can be utilized. Magnetic Resonance Imaging (MRI), for instance, provides detailed images of brain structures and is well-suited for assessing anatomical changes. Positron Emission Tomography (PET), on the other hand, is used to observe brain metabolism and functional activity, making it particularly valuable in research on diseases like Alzheimer's. The datasets used in this review are described in this section. Multimodal imaging and the CNN in conjunction with the LSTM algorithm are discussed in the review, which makes use of 112 PET pictures from Munich and 512 MRI images from Kaggle. The remaining data was used for training, and a smaller portion was reserved for testing. The Kaggle Alzheimer's Classification Dataset provides 5,121 testing photographs and 1,279 training shots for the VGG16 model [15]. The Alzheimer's Disease Neuroimaging Initiative (ADNI) provided 48,000 MRI images for the EfficientNetB7 model. In specific, 300 subjects' 2D MRI scans were used: 75 had NC, 75 with AD, 75 with early mild cognitive impairment, and 75 with late mild cognitive impairment.

2.2 Proposed Approach

The three methods studied in this paper are all based on deep learning using CNN. To ascertain whether a person is experiencing the early stages of AD, they employ multimodal data, MRI, and PET scans. There are differences between the three approaches' accuracy and Receiver Op-

ISSN 2959-409X

erating Characteristic (ROC) curves. The five sections that make up the paper's structure are shown in Fig. 1. The fundamentals of AD and its detection techniques are covered in the first section. The basic ideas of CNN models are explained in the second part. The third section covers the principles of LSTM. The fourth section presents the detection results of the three different methods. Finally, the fifth section summarizes the comparative results of using various CNN models and methods to detect early Alzheimer's, discussing areas for improvement.

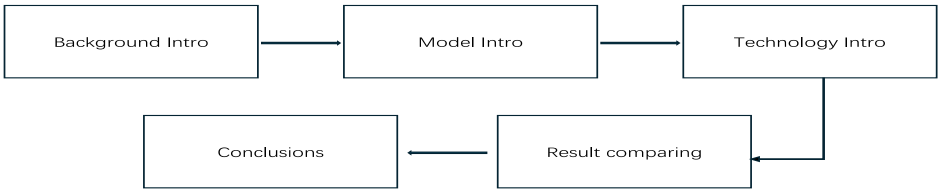

Fig. 1 The pipeline of this paper

2.2.1 Convolutional Neural Network (CNN)

Grid-structured data, such photos and videos, is processed by CNN, a deep learning model. It is still a fundamental component of contemporary computer vision, having demonstrated remarkable success in image recognition, video analysis, and natural language processing. CNN is specifically made to handle data that has a known grid-like topology, like 2D pictures or 1D time series data. Convolutional layers, activation functions, pooling layers, fully linked layers, and output layers make up the fundamental building blocks of CNN. The primary computations are performed by the convolutional layers, which also identify local features in images. ReLU and other activation functions provide non-linearity, which aids in the network's ability to acquire intricate characteristics. By reducing the spatial dimensions of the feature maps, pooling layers improve feature resilience by reducing the number of parameters and calculations. At the end of the network, the fully connected layers link the collected features to the output, which includes categorization labels. Finally, the task-specific outcomes are provided by the output layer.

CNN effectively mimics the layered processing of information in biological vision systems through mechanisms like local connectivity, weight sharing, and pooling. Local connections reduce the number of parameters, enhancing computational efficiency, while weight sharing allows the network to learn translation-invariant features within images. Pooling operations further improve the model's robustness by handling small distortions in the images. By employing multiple layers of convolution and pooling operations, CNN can extract increasingly abstract features from raw images, leading to a deeper understanding of image content and more accurate classifications.

2.2.2 CNN+LSTM Algorithm

This method proposes a hybrid deep learning approach that combines multimodal imaging, CNN, and LSTM algorithms for the early detection. The procedure is shown in Fig. 2. The technique uses CNN in conjunction with LSTM for model training, the Ant Colony Optimization (ACO) algorithm for noise reduction, and the Modified Fuzzy C-Means (MFCM) algorithm for picture segmentation. Adam algorithm has been updated and is the optimizer utilized. A specific type of Recurrent Neural Network (RNN) called LSTM was created to solve the vanishing and exploding gradient issues that arise frequently when processing lengthy sequences. Three gating mechanisms—the input gate, forget gate, and output gate—are incorporated into the long-term memory "cell state" that is introduced by LSTM. When deciding how much new information is added to the cell state, the input gate uses a sigmoid activation function to select which data to store and a Tanh function to generate candidate values. The forget gate employs a sigmoid function to generate a number between 0 and 1 that indicates the retention levels and decides which data should be erased from the cell state. To make sure that only pertinent information is passed along, the output gate regulates how the data in the cell state is output and affects the concealed state for the subsequent time step.

In practical applications, combining LSTM with CNN is particularly effective for tasks like video analysis or image sequence processing. This process typically involves three steps: first, CNN extracts spatial features from individual frames through multiple layers of convolution and pooling operations, capturing local features and structural information within the images. Next, the extracted feature vectors are input into LSTM, which leverages its strong temporal modeling capabilities to capture trends and relationships over time. Finally, based on LSTM's output, classification, generation, or prediction tasks. By combining CNN ability to understand single-image content with LSTM temporal dependency modeling, this approach significantly improves the handling of dynamic scenes.

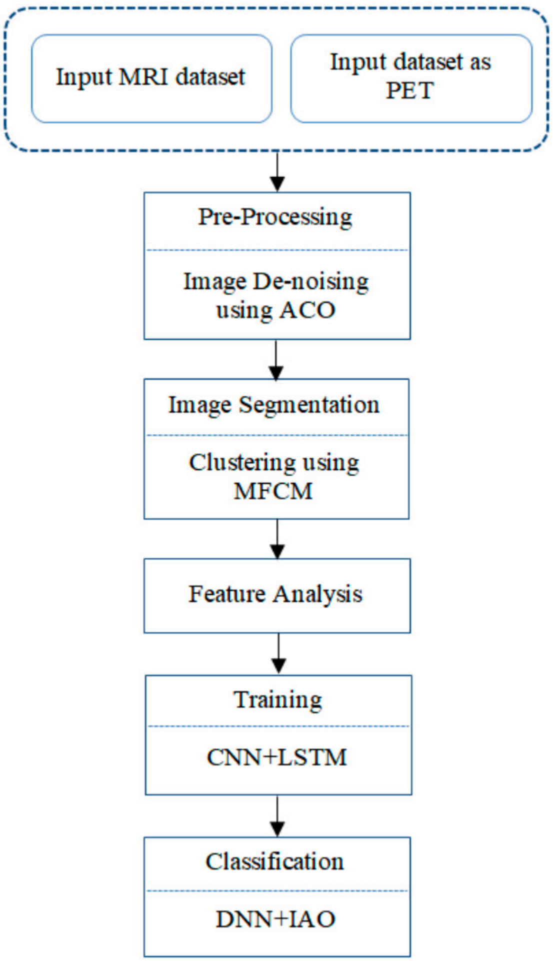

Fig. 2 The pipeline of CNN+LSTM algorithm [14]

2.2.3 Pretrained VGG16 Architecture

The pretrained VGG16 model was fine-tuned and transferred learning-based for Alzheimer's detection. Convolutional and pooling layers made up the initial sixteen frozen layers. Three fully linked layers were then added, resulting in a Sigmoid output layer. The procedure is shown in Fig. 3. A machine learning technique called transfer learning makes use of information from a related activity or domain to enhance or increase performance in a new task. Transfer learning in deep learning is based on applying a pretrained model—one that has been trained on extensive datasets—to novel tasks. This method expedites training and frequently enhances model performance by lowering the requirement for large amounts of labeled data in new jobs.

2.2.4 EfficientNetB7 Pretrained Model

The EfficientNetB7 architecture adopts a unique compound scaling strategy, which increases the network depth and width while also improving input image resolution. This approach optimizes the balance between model performance and computational cost, achieving higher accuracy and faster inference. During data preprocessing, Gaussian filtering was applied to remove noise from MRI images, and all images were resized to 256x256 pixels to ensure input consistency and processing efficiency. The CNN architecture of EfficientNetB7 employs compound coefficients to uniformly scale depth, width, and resolution. Each model weight is determined through training to minimize the loss function. During feature extraction, convolutional layers capture critical information from the MRI images. ReLU activation functions then process these features, while pooling layers reduce the dimensionality of the feature maps. Finally, the fully connected layers map the feature maps to target categories, completing the transformation from features to classifications.

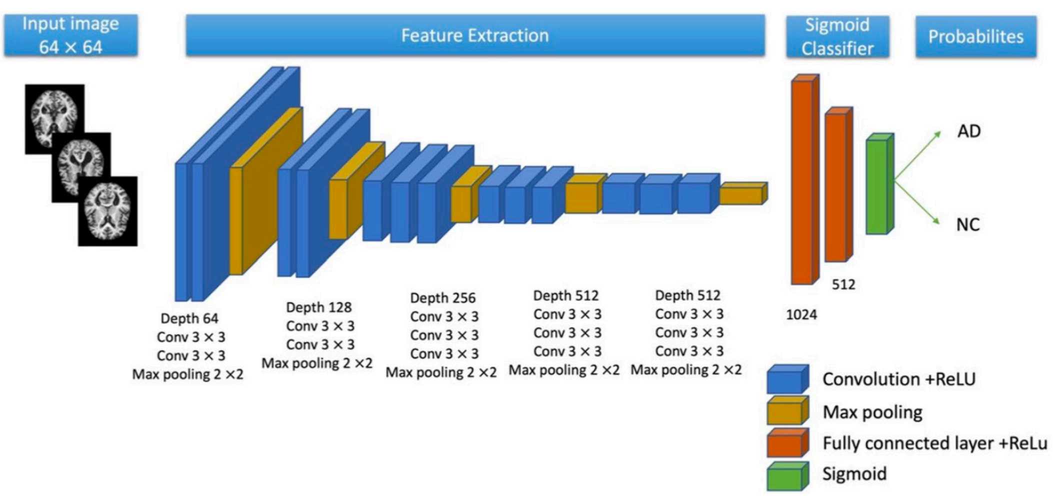

Fig. 3 VGG16 model in disease detection [12]

ISSN 2959-409X

3. Result and Discussion

The proposed CNN with LSTM models demonstrated exceptional accuracy in identifying early-stage AD, with a test set accuracy reaching 98.5%. This outcome clearly beats the random classification baseline, proving the model's dependability and persistence in identifying AD patients from healthy individuals. This high degree of accuracy was made possible by the use of CNNs in conjunction with enhanced preprocessing methods, demonstrating the efficacy of deep learning in identifying patterns in complicated medical imaging data. The model's Area Under the Curve (AUC), which ranges from 0.80 to 0.83, was examined in addition to accuracy. The model's ability to distinguish between genuine positives (AD cases) and true negatives (healthy controls) is demonstrated by this AUC score. This range, as seen in Fig. 4, demonstrates a dependable trade-off between specificity and sensitivity, demonstrating that the model can function effectively under clinical conditions when precise and timely AD detection is essential.

To achieve optimal performance, the Adam optimizer was employed with an epoch count of 512 and a batch size of 64. Under these conditions, the VGG16 model reached its highest accuracy of 97.44%, as depicted in Fig. 5. The choice of the optimizer and fine-tuned hyperparameters were pivotal in enhancing the model's learning process, allowing it to converge efficiently during training. These results demonstrate the importance of parameter optimization in boosting model performance, particularly in medical imaging applications. In addition, the Deep Learning for Early Detection of Alzheimer's Disease (DL-EDAD) system performed well on a number of criteria, such as F-Score, sensitivity, specificity, and accuracy. The system achieved a sensitivity of 98.08%, meaning it accurately detected almost all AD cases, while maintaining a specificity of 98.20%, indicating it could reliably identify healthy individuals without misclassifying them as AD patients. The overall accuracy reached 98.68%, further emphasizing the model's robustness, and the F-Score was an impressive 98.95%, reflecting its balance between precision and recall. These findings, which are shown in Fig. 6, show how well the system handles intricate medical imagery, including MRI scans, and can identify minute structural alterations linked to the advancement of AD.

The findings highlight the efficiency of deep learning in AD detection, especially with regard to CNN-based EfficientNetB7 models. The combination of multimodal imaging and advanced machine learning techniques significantly enhances the ability to detect early AD with high accuracy. By fine-tuning hyperparameters and optimizing the training process, the proposed approach shows great potential for real-world clinical applications, paving the way for improved diagnostic tools in Alzheimer's research.

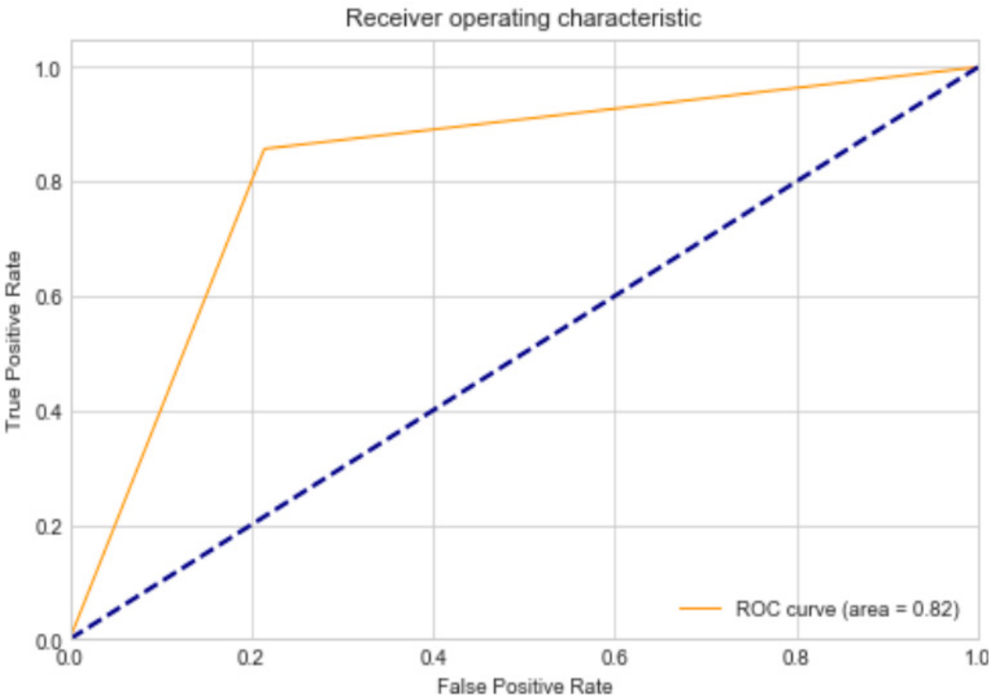

Fig. 4 ROC curve analysis chart [14]

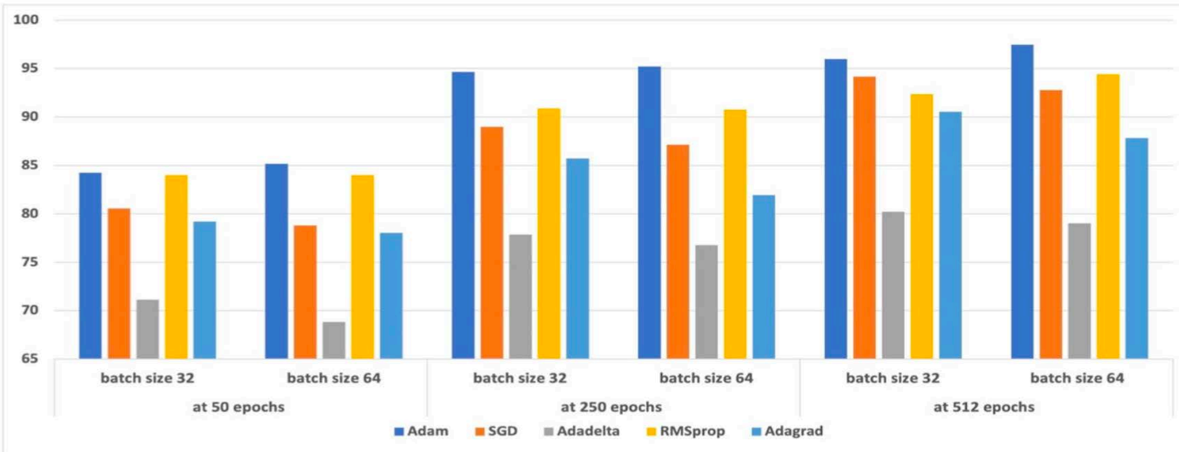

Fig. 7 Comparison for VGG16 model with different optimizers (Adam, SGD, Adadelta, RMSprop, Adagrad) and at different epochs and batch size

Fig. 5 Epoch analysis of different optimizers [12]

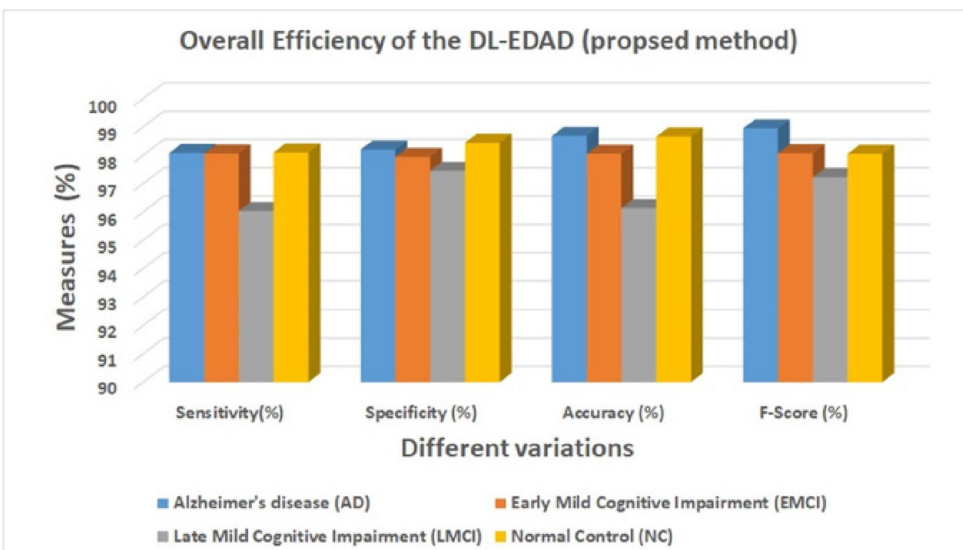

Fig. 6 Comprehensive indicator analysis under different variables [13]

4. Conclusion

This study examined the application of CNNs and hybrid deep learning models in the early detection of AD using multimodal imaging methods. The results demonstrated the remarkable sensitivity, specificity, and accuracy that the models—particularly VGG16 and EfficientNetB7—achieved in identifying AD in MRI and PET scans. Improved diagnostic performance was attained by combining LSTM networks with CNNs to further improve the models' capacity to capture temporal and spatial data. The DL-EDAD system showed significant potential, with an accuracy rate of 98.68% and an F-Score of 98.95%, outperforming traditional diagnostic methods. These findings emphasize the importance of using deep learning for early AD detection, where accurate and timely diagnosis can have a profound impact on patient care and treatment. Future work could focus on expanding the dataset, refining models, and incorporating other forms of neuroimaging data to further enhance the robustness of early detection systems.

References

[1] Singh S P, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyás B. 3D Deep learning on medical images: a review, 2020, 1–13.

[2] Jo T, Nho K, Saykin AJ. Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using

ISSN 2959-409X

neuroimaging data, Frontiers in Aging Neuroscience. 2019, 11.
[3] Physicians P C. 2020 Alzheimer's disease facts and figures. Alzheimer's Dement. 2020, 16(3): 391–460.
[4] Yang Y, Li X, Wang P, Xia Y, Ye Q. Multi-Source transfer learning via ensemble approach for initial diagnosis of Alzheimer's disease, IEEE J Transl Eng Heal Med. 2020, 1–10.
[5] Wen J, et al. Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation, Medical Image Analysis. 2020, 63: 101694.
[6] Altinkaya E, Polat K, Barakli B. Detection of Alzheimer' s disease and dementia states based on deep learning from MRI images: a comprehensive review. 2020, 39–53.
[7] Dall, T. M. et al. Supply and demand analysis of the current and future US neurology workforce. Neurology. 2013, 81:470–478.
[8] Dall, T. M. Physician workforce shortages: what do the data really say? Acad. Med 90. 2015, 1581–1582.
[9] Pedersen, M. et al. Artificial intelligence for clinical decision support in neurology. Brain Commun. 2020, 2: fcaa096.
[10] Lu D, Popuri K, Ding G W. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. 2018, 8: 1–13.
[11] Qiu, S. et al. Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. Brain. 2020, 143: 1920–1933.
[12] Balaji P, Chaurasia M A, Bilfaqih S M. Hybridized Deep Learning Approach for Detecting Alzheimer's Disease. Biomedicines, 2023, 11: 149.
[13] Suchitra S, Krishnasamy L, Poovaraghan R J. A deep learning-based early alzheimer's disease detection using magnetic resonance images. Multimedia Tools and Applications, 2024: 1-22.
[14] Balaji P, Chaurasia M, Bilfaqih S M, Muniasamy A. Hybridized Deep Learning Approach for Detecting Alzheimer's Disease. Biomedicines. 2023, 11: 149.
[15] tourist. Alzheimer's Dataset (4 class of Images). Kaggle. Retrieved on 2024, Retrieved from: https:// www. kaggle.com/ datasets/ tourist55/ alzheimers-dataset-4-class-of-images.