Journal of Information Systems Engineering and Management
2025, 10(12s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Impact of Skull Segmentation in MRI images for Alzheimer's Diagnosis based on Transfer Learning Techniques

ZS. Khaleel[1,2]*, Amir Lakizadeh[1]
[1]*Computer Engineering and Information Technology Department, University of Qom, Qom, Iran*
[2]*Department of mathematic, Open Educational College, Basra Study Center, Iraq*
** Corresponding author's Email: zahsad@basrahaoe.iq*

ARTICLE INFO	ABSTRACT
Received: 25 Nov 2024 Revised: 26 Dec 2024 Accepted: 22 Jan 2025	This study aims to investigate how the process of skull stripping improves the performance of artificial intelligence models to predict early Alzheimer's disease through magnetic resonance imaging of the patient. The skull stripping process presented in this research was developed using thresholding, morphological, and U-net pre-trained transfer model on magnetic resonance imaging. The experimental results confirmed that the use of the skull stripping process can significantly improve the efficiency of Alzheimer's detection via VGG16, Inception, DenseNet121, and ensemble model ResNet+Bilstm. Using skull stripping using deep learning techniques with thresholding and morphological operations produced a very clear improvement over the AI models used in this research as follows: accuracy, f1 score, and Area under curve (AUC) were improved respectively to 99.5%, 95%, and 99% for ResNet+BILSM, up to 99.6%, 95.7%, and 99.14% for VGG16, enhance as 91%, 92.4%, and 95.47% for Inception, and improve DensNet121resultes as 79.5%, 82.2%, and 95.51%. **Keywords:** skull stripping, classification, MRI (Magnetic Resonance Imaging), Deep learning, Alzheimer's disease

1. Introduction

The degenerative neurological disorder known as Alzheimer's disease (AD) becomes worse with time and is typified by changes in behavior, memory loss, and cognitive ability degradation. As a person grows older, the possibility of developing Alzheimer's disease increases, and this is the main motivation for highlighting the necessity and importance of early and accurate diagnosis and medical intervention for the affected person to receive successful treatment The diagnosis and monitoring of AD depend critically on neuroimaging—more especially, magnetic resonance imaging (MRI). Eliminating non-brain tissues from neuroimaging studies is a crucial preprocessing step in MRI data. Skull stripping plays a crucial role in neuroimaging analyses by isolating brain tissues, which greatly improves the accuracy of subsequent image processing steps like tissue segmentation, cortical thickness measurement, and volumetric analysis. Although the significance of skull stripping in enhancing diagnostic accuracy for AD is widely recognized, it continues to be a subject of ongoing research. Various algorithms and techniques have been devised to carry out skull stripping, each with its own advantages and drawbacks[1,2]. There is currently no cure for AD, but there are treatments that can help alleviate and deal with several of its symptoms. There have been significant developments in the research and development of medications that attempt to slow down the progression of the disease, especially with the initial identification of disease biological markers[3]. In recent years, scientists have used machine learning and deep learning mechanisms to detect Alzheimer's at an early stage. During the initial stages of AD diagnosis, conventional machine learning methods commonly utilize two categories of features [4]. The identification of structural or functional problems in the brain heavily relies on fundamental assumptions, such as regional cortical thickness, hippocampal volume, and gray matter volume. [5, 6].

Conventional approaches rely on the extraction of features through manual procedures, which largely rely on technical expertise and repetitive trials. This process is perceived as time-consuming and subjective. Deep learning, namely convolutional neural networks (CNNs), is an efficient solution to address these issues [7]. CNN has the potential to enhance efficiency even further and has demonstrated significant effectiveness in diagnosing Alzheimer's disease. Additionally, it eliminates the need for manual extraction of characteristics as it automatically extracts them [8, 9].

In this study, it was proposed to combine the work of CNN in extracting features with the RNN algorithm for classification, but in their advanced stages, that means the use of the Resnet50 [10]CNN layers for accurate extraction of MRI features with the Bidirectional long-short-memory (BILSTM)algorithm derived to classification[11,12]. And utilized pre-trained deep learning models to study the influence of brain extraction such as VGG16, Inception, and DensNet121.

The rest of the paper is arranged in the following way. Section 2 provides a summary of the relevant studies for the suggested model Next, in Section 3, we provide a thorough explanation of our methodology. The results of the experiment are described in Section 4, Finally, Section 5 contains a summary of our findings

2. Related Work

Nikita Goenka et al [13] conducted a study on the detection of Alzheimer's disease using a 3D patch-based feature extraction technique. They used convolutional neural networks, specifically 3D networks, to classify MRI images obtained from the MIRIAD database into Normal Control and Alzheimer Disease categories. The images were pre-processed and augmented to increase the dataset size. The classification achieved an accuracy of 99.79 percent. The authors highlight the potential of this classification method in assisting with the detection of Alzheimer's disease, especially in situations where clinicians are not readily available, such as during a pandemic or in remote areas. Das, D., & Kalita, S. (2014). The focus of this paper is on the technical process of skull stripping, which involves removing non-brain tissues from brain MRI data to facilitate further analysis. The goal is to excise non-brain tissues in order to isolate the main region of the brain. An inherent challenge in skull stripping is distinguishing brain tissues from non-brain tissues owing to variations in intensity. To address this problem, the authors analyze five entropy-based thresholding approaches to determine an accurate threshold requirement for brain tissues. Maximal entropy sum, entropic correlation, Renyi's entropy, Tsallis entropy, and modified Tsallis entropy are among the approaches included. This work employed 50 T1p weighted coronal MR images for experimentation and achieved an accuracy rate of 80.4% via the usage of modified Ts. In their study, Snehashis Roy et al.[15] put forward a method called MONSTR, which utilizes sparse patch information to strip the brain of multiple contrasts. This method combines non-local patch information from atlases containing various MR sequences and reference brain masks to create a target brain mask In their study, Kleesiek, Jens, et al[16] introduced a 3D convolutional deep learning framework designed specifically for brain extraction from MR images. The method provided achieves state-of-the-art performance in terms of Dice score and specificity, while also addressing several issues present in prior methods. Their approach is not restricted solely to non-enhanced T1w pictures, but it can also handle contrast-enhanced scans. Furthermore, when properly trained, the method can effectively handle any number of modalities. In their study, Druzhinina and Kondrateva (17) investigate the application of artificial intelligence, particularly deep learning algorithms, in neuroimaging to investigate cognitive impairment in disorders such as Alzheimer's disease. Their main objective is to assess the stability of 3D computer vision models in the classification of Alzheimer's disease using ADNI data. They evaluate the performance of the model and analyze evidence of overfitting using traditional interpretive techniques and swap tests. The authors propose that the implementation of skull-stripping and information transfer techniques can have an effect on the resilience and replicability of acquired patterns. They also advocate the use of swap tests as a means to guarantee the stability of the model. The significance of early diagnosis and treatment for Alzheimer's disease (AD) and Mild Cognitive Impairment (MCI) is emphasized by Manu Raju, et al [18]. They suggest a technique that utilizes Meta-Heuristic Markov Random Field Segmentation to differentiate brain tissue and extract characteristics from the segmented Cerebral Spinal Fluid (CSF) and Grey Matter (GM). Additionally, Particle Swarm Optimization (PSO) and a Markov Random Field model are utilized to enhance segmentation. Brain tissue is segmented and features are extracted based on shape, intensity, and texture. A deep neural network with four layers is used for classification. The proposed method demonstrates an impressive accuracy rate of 97.5% when tested with a standard dataset from Alzheimer's disease-neuroimaging (ADNI), surpassing previous methods in AD and MCI classification. In this study, Pei, Linmin, and colleagues [19] The authors introduce an ensemble neural network (EnNet) approach for automating skull stripping in multiparametric MRI data. The performance of the system is examined on 15 distinct image modality combinations and compared to the most advanced algorithms currently available. A dataset of 815 patients, included both with and without glioblastoma multiforme (GBM), was gathered and the accuracy of skull stripping was confirmed by competent radiologists. The proposed approach provides superior performance and has several advantages, such as complete automation, applicability to various MRI modalities, accuracy in both healthy and GBM cases, ability to handle multicenter data, and being the pioneer in quantitatively comparing skull stripping performance across different modalities. Xing, Xin, et al.[20]. propose using a 2D CNN architecture for Alzheimer's disease classification using 3D MRI images. They address the time-consuming and computationally expensive nature of training a 3D CNN by using approximate rank pooling to transform the 3D MRI image volume into a 2D image. They demonstrate that their proposed CNN model achieves higher accuracy in Alzheimer's disease classification compared to baseline 3D models. Additionally, they highlight the efficiency of their method, requiring only a fraction of the training time.

The transition from traditional to machine- and deep-learning-based methods for skull stripping is explored by Anam Fatima et al [21], highlighting the importance of accurate and efficient techniques for brain segmentation tasks. Advanced methods have proven their superiority in diagnostic procedures, particularly in skull stripping, highlighting the importance of deep learning approaches over conventional and machine learning techniques. In their study, Zelin, Xu et al [22] propose a technique that utilizes deep learning and MRI images to diagnose AD. The method demonstrates impressive accuracy in differentiating between AD, MCI, and NC by extracting Gray Matter (GM) MRI images automatically. Skull stripping is done using FMRIB Software Library (FSL) software to eliminate the skull and cerebellum from MRI images. After that, the images are registered to improve the accuracy of the diagnosis process. Monica, Rajendiran et al [23] emphasize the importance of early detection in preventing memory loss in Alzheimer's disease (AD), and discuss the use of automated technologies for accurate identification of this neurodegenerative illness. Skull stripping is a valuable technique in MRI image processing that allows for the isolation of brain structures from non-brain tissues. This enables a more precise analysis of brain features related to Alzheimer's disease. Exploring deep learning models, such as transfer learning approaches, to enhance classification accuracy in AD detection. This research focuses on addressing challenges like multi-class classification and improving the performance of medical imaging

systems. AD is a widespread neurological disorder, highlighting the crucial importance of early diagnosis to enhance patient outcomes. Afiya Begum and Prabha Selvaraj [24] explore the application of image processing techniques, machine learning, and deep learning in the detection of AD. This text explores different techniques used for diagnosing AD, emphasizing the importance of image processing, feature extraction, optimization, and classification methods in improving AD recognition. Sajid Yousuf Bhat, Afnan Naqshbandi, and Muhammad Abulaish [25] Present a new method for skull stripping on different MRI modalities using thresholding and morphology techniques, the proposed method is user-friendly, requires minimal parameter adjustments, and produces satisfactory results. evaluation on three benchmark datasets and comparison with nine existing methods showed that the proposed method performs similarly to some of the best methods available, additionally the method not only extracts the brain mask but also generates the skull mask, which can be valuable for studying different skull pathologies. In their study, Stoleru and Iftene (26) evaluate the efficacy of transfer learning along with deep learning methods in precisely detecting Alzheimer's Disease (AD) using MRI scans. A comparative analysis is conducted on two models with prior training, ResNet-152 and AlexNet, using original supplier data taken from the ADNI dataset and data that has been subjected to skull stripping. The findings clearly demonstrate that the ResNet-152 were model has attained a remarkable accuracy rate of 99.96% when used on sagittal plane slices obtained from the original dataset. Furthermore, the study highlights the higher efficacy of models trained on authentic MRI images in comparison to those trained on pictures that have undergone skull stripping. E. Khaled and colleagues [27] have developed a highly efficient technique for identifying Alzheimer's disease (AD) by utilizing information from MRI scans. The researchers import a dataset obtained from the Alzheimer's disease neurological imaging initiative (ADNI) and employ a deep three-dimensional convolution network (3D CNN) together with a Transformer encoder to interpret the input of MRI images and genetic sequence. By experimenting with other classifiers, such as Support Vector Machine (SVM) and XGBoost, they determine that retraining the skull-stripped dataset with features derived from the CNN model and employing the XGBoost classifier yields a rather high accuracy rate of 70%. Hazarika, Ruhul Amin, and colleagues [28] want to enhance the accuracy of Alzheimer's disease (AD) diagnosis by utilizing brain magnetic resonance images and deep neural network (DNN) models. The researchers conducted a performance comparison of several deep neural network (DNN) models and determined that the DenseNet-121 model attained the highest level of performance, achieving an accuracy of 86.55%. Nevertheless, due to the high computational cost of

DenseNet-121, the authors suggest a hybrid approach that integrates the LeNet and AlexNet models. The hybrid model demonstrates superior performance compared to DenseNet, with an overall performance rate of 93.58%. The authors further observe that the put forward model possesses a reduced number of convolutional parameters, therefore rendering it a lightweight and computationally efficient alternative.

3. Materials and Methods

This section provides a description of the Alzheimer's disease dataset, pre-processing, and brain extraction technique from the skull using thresholding and morphology operation and then combined with deep learning to improve its accuracy, we adopted models of transfer learning and ensemble models (ResNet with BILSTM) to compare their outputs when we applied the proposed skull stripping and when it was not used, model evaluation metrics, and show the results of classification Fig.1. explain the ensemble proposed models (ResNet with BILSTM)

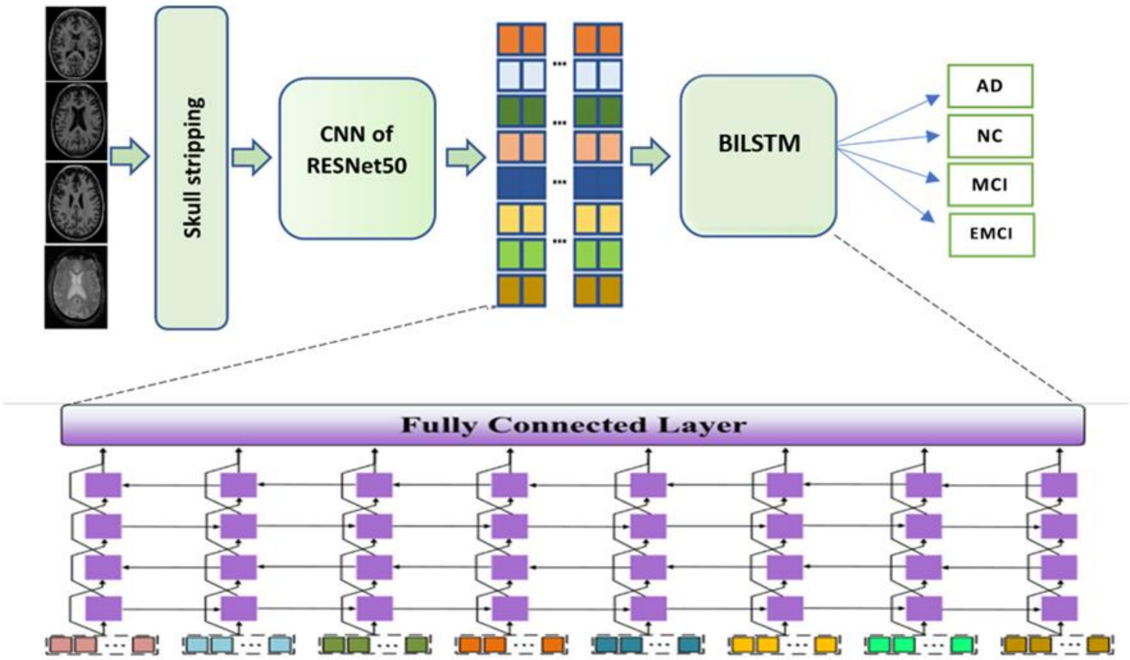

Figure. 1 Ensemble models (ResNet with BILSTM)

3.1 DESCRIPTION OF DATASETS AND PREPROCESSING

There are a plethora of internet data sets accessible for AD categorization. Several AD data sets are unsuitable for this research due to their CSV format. Organizations such as ADNI provide their data sets for educational and scientific purposes. However, the data sets are in the form of three-dimensional images. The ADNI organization is a research endeavor that seeks to enhance comprehension of Alzheimer's disease through the examination of its advancement and the creation of novel diagnostic and therapeutic approaches. The MRI data obtained from ADNI are highly helpful for investigating the advancement of Alzheimer's disease, finding biomarkers, and creating novel imaging techniques for the early detection and monitoring of the disease. We selected the ADNI dataset obtained from Kaggle due to its identical key features. Offers a comprehensive collection of neuroimaging data, allowing for the examination of disease progression over a prolonged period through the availability of longitudinal data. The dataset comprises of de-identified MRI scan images together with their related classification labels. The data collection comprises a multi-class dataset with multiple views and four separate classes. AD refers to Alzheimer's Disease, CN stands for Cognitive Normal, EMCI represents Early Mild Cognitive Impairment, and MCI denotes Mild Cognitive Impairment. As shown in Fig 2

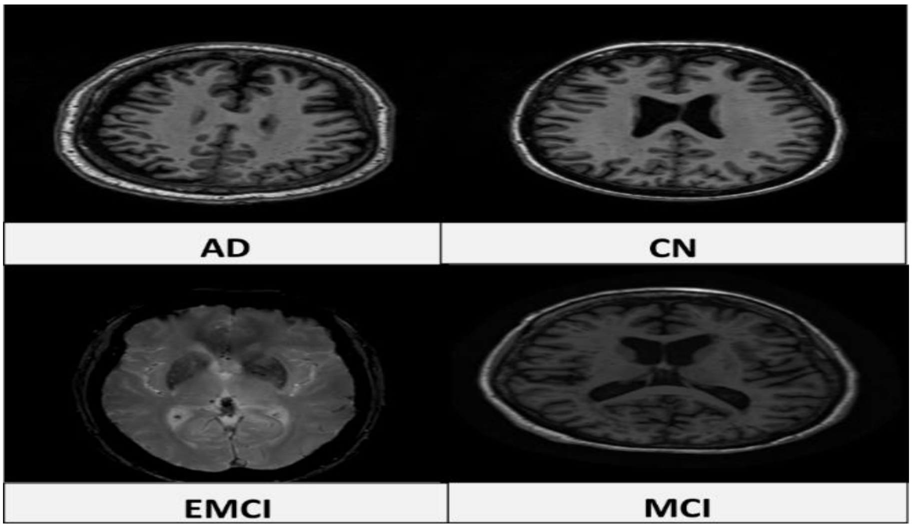

Figure 2. shows the four classes of ADNI

Table 1. Clears the distribution of ADNI datasets

Classes	No. of images+gender
Alzheimer's Disease	1124 M/F
Cognitive Normal	1440 M/F
Early Mild Cognitive Impairment	904 M/F
Mild Cognitive Impairment	1590 M/F

To mitigate the issue of overfitting, we utilized augmented processing approaches. Overfitting is a prevalent issue in machine learning and statistical modelling This problem often occurs when developers train the model too much or have too little training data, leading to inefficient performance when applied to unseen data. Augmented processing, this process helps improve the performance of the training model by increasing the number of training data by introducing variables such as applying transformations, shifting, flipping, incorporating noise, or producing synthetic samples. Utilizing augmented processing can help address the issue of overfitting by increasing the training set to include a broader range of patterns and variations. By implementing this approach, the model achieves greater resilience and the capacity to generate adaptable representations, leading to enhanced performance when faced with new data. For this task, we utilized various augmentation techniques adapted to our specific dataset. These techniques included manipulating the original images through rotation in the range [10, -10], both vertical and horizontal, shifting in the range (-0.1, 0.1), flipping vertically and horizontally randomly, and shear of the original images in the range (-0.1, 0.1). By incorporating augmented data, significant enhancements were made to the training process, leading to a more streamlined and reliable model that exhibits superior generalization abilities. With this enhancement, the model becomes more adept at handling unseen data, ultimately mitigating the detrimental impact of overfitting. The table 2 displays the number of images after the Augmentation process for the datasets.

Tabel2. Number of ADNI images after augmentation

Classes	No. of images
Alzheimer's Disease	10924
Cognitive Normal	11300
Early Mild Cognitive Impairment	10986
Mild Cognitive Impairment	11194

3.2 Skull Stripping

3.2.1 Skull Stripping with Threshold and Morphological Operations

A key stage in neuroimaging analysis, skull stripping aims to eliminate non-brain structures (skull, scalp, and other soft tissues) from the brain picture. Correct brain volume measurements, segmentation, and other analysis depend on this procedure. Two widely used methods for skull stripping are thresholding and morphological procedures. Thresholding is a simple and effective technique that divides the image into segments based on the intensity levels of its pixels. By selecting a threshold value, we can easily identify which pixels represent brain tissue and which ones do not. These algorithms in image processing utilize the shape and organization of objects within the image. Applying these techniques can improve the results of thresholding and address small isolated objects or minor gaps in the brain mask. Several frequently employed morphological techniques for skull stripping include Erosion shrinks the foreground objects by removing pixels from their boundaries, Expanding the foreground objects by adding pixels to their boundaries, opening combines erosion and dilation to remove small protrusions and smooth the object boundaries and Closing combines dilation and erosion to fill small holes and smooth the object boundaries. Thresholding and morphological operations can be combined to achieve better skull-stripping results. Typically, thresholding is used to obtain an initial segmentation, and then morphological operations are applied to refine the mask and remove any remaining non-brain tissues. The benefits of this combination are straightforward and effective techniques, can be readily executed, and is appropriate for various kinds of images but sometimes sensitive to noise and artifacts in the image, and may not apply to all categories of photos and manual correction.

3.2.2 Skull stripping with U-Net

In recent years, the U-Net architecture has become a potent tool for skull stripping. It is a convolutional neural network uniquely developed for segmenting biomedical images. It has various advantages compared to conventional techniques such as thresholding and morphological procedures. The U-Net design consists of two main paths, known as encoders: The proposed approach employs a series of convolutional and pooling layers to extract information from the input image. Each layer in the network decreases the spatial resolution of the feature maps while simultaneously increasing the number of feature channels, (decoder): The purpose of this process is to increase the resolution of the feature maps obtained from the contracting path, it achieves this by combining them with the matching feature maps from the contracting path utilizing skip connections, this enables the network to regain spatial information that was lost during the contracting path. The last layer of the U-Net produces a probability map for every pixel, indicating the possibility of it being part of the brain or non-brain tissue. Advantages of U-Net for skull stripping it demonstrates superior accuracy in skull stripping compared to conventional techniques, particularly when trained on extensive datasets, robustness means U-Net exhibits more resilience to noise and artifacts in the image as compared to the techniques of thresholding and morphological operations, Automated learning U-Net has the capability to autonomously acquire the required characteristics for skull stripping from the training data, hence obviating the necessity for manual parameter selection and U-Net can undergo end-to-end training, where the entire network is tailored specifically for the skull stripping task., therefore it has become a highly effective and popular instrument for skull stripping, providing automated learning capabilities, robustness, and high accuracy. U-Net can accomplish state-of-the-art performance in skull stripping through the careful selection of training data, optimization of parameters, and network design, thereby significantly enhancing the accuracy and efficiency of neuroimaging analysis.

3.2.3 A Novel Approach to Skull Stripping by Integrating Thresholding, Morphological Operations, and U-Net

The integration of thresholding, morphological procedures, and U-Net can as shown in Fig .3 yield a potent and resilient pipeline for skull stripping. This hybrid strategy utilizes the advantages of each method to attain superior accuracy and efficiency.

This strategy provides numerous benefits such as enhanced accuracy the integration of U-Net with conventional techniques can yield superior accuracy compared to the utilization of each method independently, The U-Net model has the ability to acquire intricate characteristics from the data, while the utilization of thresholding and morphological processes ensures a strong and reliable initial segmentation, U-Net necessitates a substantial quantity of training data, particularly for intricate tasks such as skull peeling but the use of a pre-trained U-Net or commencing with a proficient first mask can diminish the necessity for extensive training data, the amalgamation of techniques can optimize the overall efficiency of the skull stripping procedure. U-Net is capable of effectively managing intricate scenarios, whereas smaller scenarios can be easily addressed via thresholding and morphological processes that lead to Enhanced Efficiency

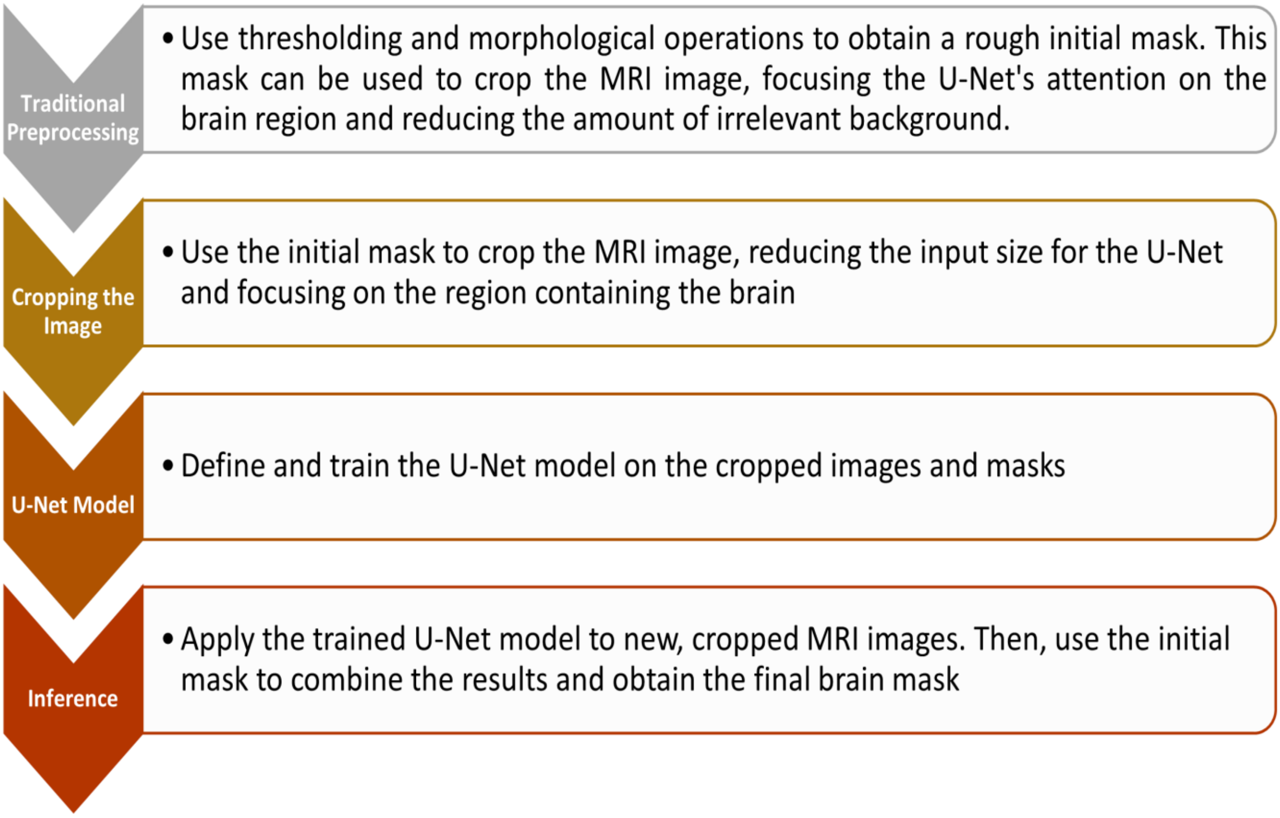

Figure 3. Workflow for Combining Thresholding, Morphological Operations, and U-Net for Skull Stripping

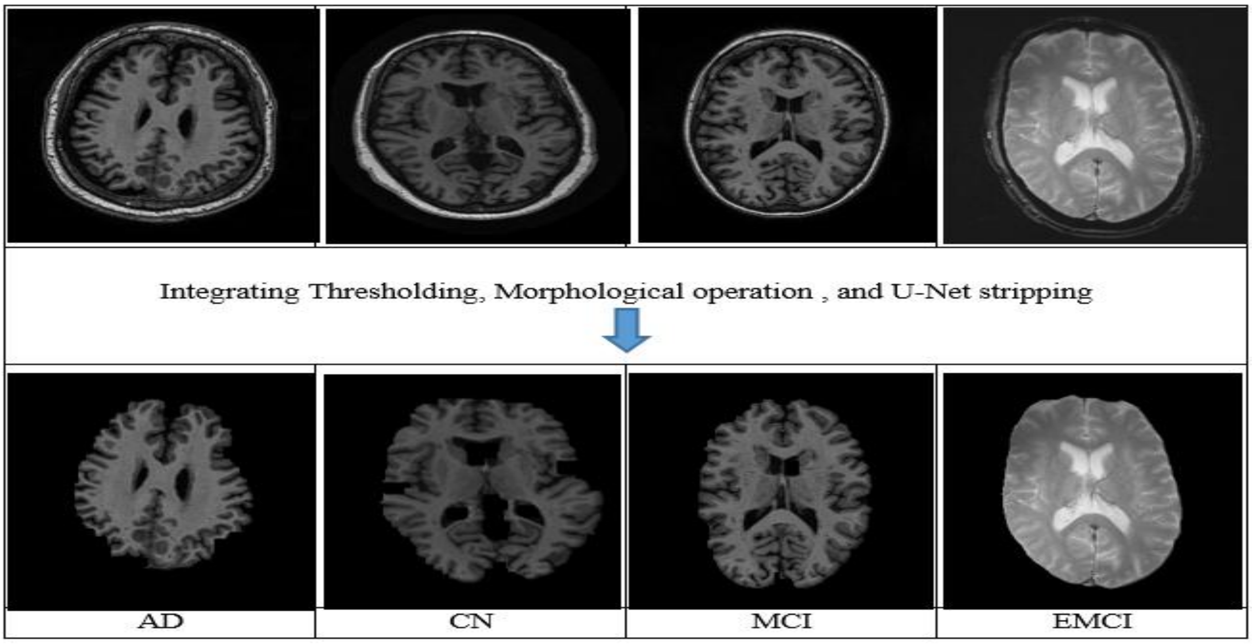

Figure 4. Hybrid skull stripping

Figure 4 shows pictures of all four types of dementia after applying the brain extraction process. It is very clear that the proposed method works accurately in removing the non-brain area from the bones and brain tissues. A pre-processing operation applied also for enhancement the performance of the proposed model to improve the quality of the data and extract relevant features such as image rescaling, noise removal, image normalization

3.3 Fine-Tuned Individual Deep Learning Models

3.3.1 VGG16

VGG-16 is a deep convolutional neural network (CNN) architecture pioneered by the visual geometry group (VGG) at the University of Oxford in 2014. An extensively employed model for image identification tasks, it has attained cutting-edge performance in several computer vision (CV) benchmarks. VGG16's architecture was designed with 16 layers, comprising 13 convolutional layers and 3 fully-connected layers. Convolutional layers are constructed with compact 3 × 3 filters and are stacked atop one another, hence augmenting the network's depth. Utilizing tiny filters with a small stride size enhances the preservation of spatial information and facilitates the network's acquisition of more intricate characteristics [29].

3.3.2 DenseNet121

DenseNet121 [30] is a typically used convolutional neural network (CNN) architecture for image classification problems. Introduced in 2017, it was developed as an enhancement to the earlier widely used architectures like VGG and ResNet. The DenseNet121 model utilizes a dense connection network, in which each layer gets feature maps from all preceding layers and transmits its own feature maps to all subsequent layers. The high level of connectedness enables improved management of gradient flow and parameter efficiency, thereby minimizing issues related to disappearing gradients. The architecture employs 121 layers, encompassing convolutional, pooling, and dense blocks, and has attained exceptional performance on various benchmark datasets, including ImageNet

3.3.3 Inception

Google researchers created the Inception model, a deep convolutional neural network that uses Inception modules in a novel way to capture characteristics at several scales through the parallel application of different convolutional filters and pooling operations. Improved computational efficiency and simplified models are the results of combining this method with dimensionality reduction using 1x1 convolutions and global average pooling rather than fully connected layers. Auxiliary classifiers used in intermediary layers also help with training stability and gradient flow [31,32].

3.3.5 ResNet with BILSTM

A Residual Network (ResNet) is a neural network architecture designed to tackle the issue of disappearing gradients in extremely deep networks by incorporating residual learning. The network comprises "residual blocks," in which skip connections facilitate the learning of residual functions, therefore simplifying optimization and permitting the implementation of somewhat more complex structures. The fundamental component consists of convolutional layers with batch normalization and ReLU activation, preceded by an identity shortcut that appends the input of the block to its respective output. ResNet designs, such as ResNet-50 and ResNet-101, are extensively employed in image recognition applications because of their exceptional precision and effectiveness [33]. Bidirectional Long Short-Term Memory (BiLSTM) is a recurrent neural network (RNN) that integrates two LSTM networks: one for forward processing of the input sequence (from beginning to end) and another for backward processing (from end to beginning). This architectural design effectively captures information from both past (backward) and future (ahead) contexts, rendering it very efficient, in order to address the vanishing gradient problem, BiLSTM employs LSTM cells that exhibit superior long-term dependency compared to conventional RNNs. By concatenating the forward and backward hidden states, a BiLSTM produces an output that offers a more comprehensive description of the input sequence. Integration of ResNet with BiLSTM as shown in Figure 1 above entails the utilization of ResNet for extracting spatial characteristics from input images, which are then inputted into a BiLSTM model to acquire temporal dependencies. Typically, the ResNet model is pre-trained and used without its final classification layers, where it generates high-level feature vectors. The input vectors are subsequently sent into a BiLSTM, which progressively processes them in both forward and backward directions to capture temporal context. The outputs of the BiLSTM are routed into a last dense layer for the purpose of classification.

4. Experiments

4.1 Evaluation of the models

The experiments were on a Google Colab hosted Jupiter Notebook service in subscription mode with runtime type Python 3 and hardware accelerator T4GPU Tesla T4 is a GPU card based on the Turing architecture and targeted at deep learning model inference acceleration with system RAM 52 GB, the evolution of model was conducted utilizing the cross-validation which is part of the dataset, A categorical cross-entropy was used to train and test individual models for Alzheimer's Disease (AD), Cognitive Normal (CN), Early Mild Cognitive Impairment (EMCI), and Mild Cognitive Impairment (MCI). The performance of these models was optimized using the Adam optimizer. Using several measures based on test data ensures a model is resilient from all angles Successful model training depends on an extensive understanding of these results, for example, high accuracy (over 90%) does not necessarily indicate an excellent model Other factors include f1-score and Area under the curve used to evaluate the performance of our models.

4.1.1. Accuracy

Accuracy is the measure of the total of correct predictions out of all accurate ones, and it is calculated using the following formulas:

$Accuracy = (TP + TN) / (TP + FN + FP + TN)$ (1)

TP, TN, FN, and FP represent True Positive, True Negative, and False Positive values, respectively.

4.1.2. F1—Score

An ideal classification model has precision and recall values of 1.0. The F1 score represents the harmonic average of precision and recall. The F1 score graph is distinctive in that it displays an individual line for each class designation. The F1 score is computed using the following formula:

$F1 = 2 (Precision * Recall / Precision + Recall)$ (2)

4.1.3 . Area Under the curve

AUC, also known as Area Under the ROC Curve, is a quantitative measure utilized to assess the effectiveness of classification models. A single numerical value quantifies the model's capacity to differentiate between positive and negative classes.

4.2 . Results and Discussion

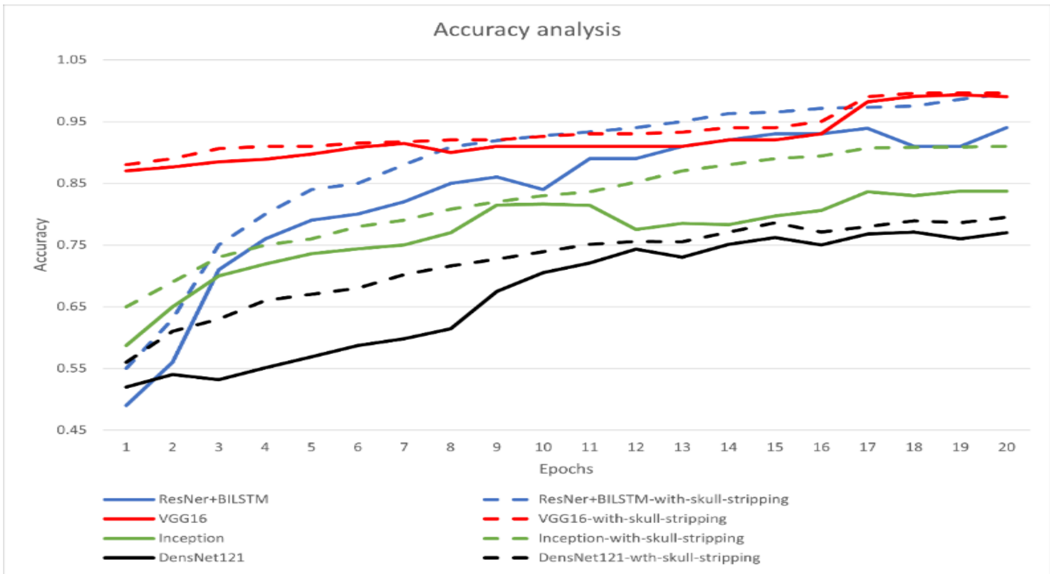

Figure 5. Accuracy results of models with and without skull stripping

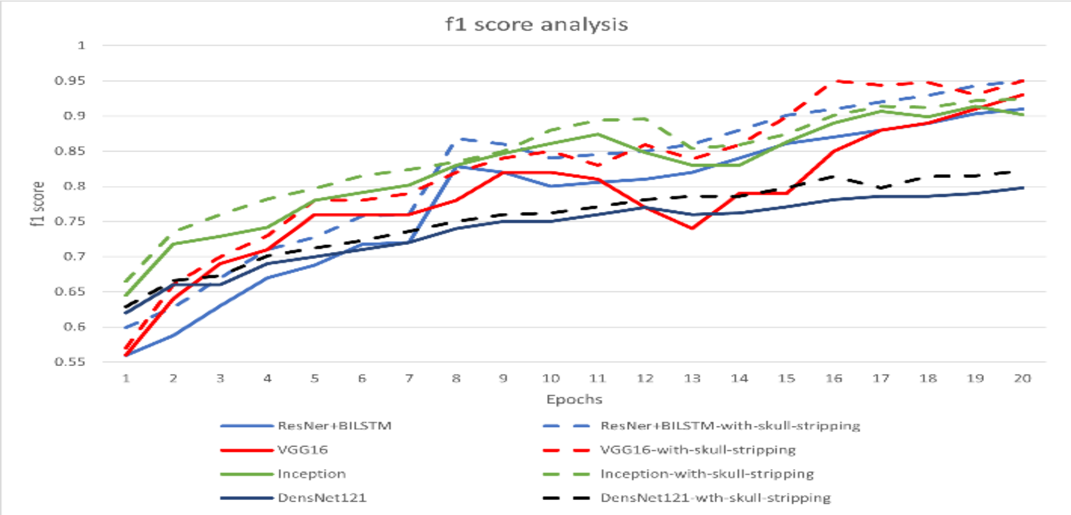

Figure 6. f1 score results of models with and without skull stripping

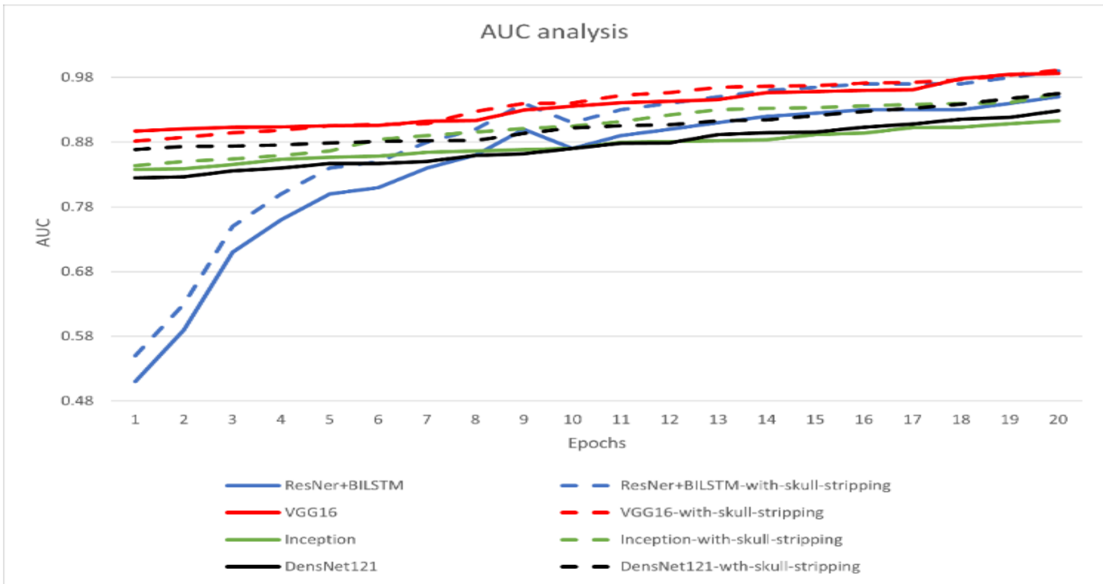

Figure 7. AUC results of models with and without skull stripping

Models	Accuracy	F1 score	Auc
ResNet+BILSM	94%	91%	95%
VGG16	99%	93%	98.62%
Inception	85.7%	90.2	91.27%
DensNet121	77%	78.8%	92.86%
With using Skull stripping			
ResNet+BILSM	99.5%	95%	99%
VGG16	99.6%	95.7%	99.14%
Inception	91%	92.4%	95.47%
DensNet121	79.5%	82.2%	95.51%

Table 3. Results of models after the use of skull stripping

Figs 5,6, and 7 with Table 3 depict the ability of models improved by 5.5 for ResNet+BILSM, 1.6 for VGG16, 5.3 for Inception, and for DenseNet121 is 2.5 which means the model's accuracy in classifying brain tissue is enhanced by avoiding misclassification caused by irrelevant features, resulting in more accurate predictions overall. The improvement in the F1 score to 95% for ResNet+BILSM, 99.6% for VGG16, 91% Inception, and 79.5% for DensNet121 after applying skull stripping focusing more on specific and detailed brain images, the model improves its ability to recognize and distinguishing between different classes accurately, this improvement likely results from reduced noise, enhanced feature quality, and improved data consistency, Furthermore, we observe a rise in AUC values with skull stripping enhances the model's capacity to differentiate between classes across all potential threshold values, by excluding non-brain tissues, the model can acquire a more advanced understanding of the unique characteristics of many brain disorders, resulting in a stronger performance that sustains a high level of sensitivity (true positive rate) and specificity (true negative rate).

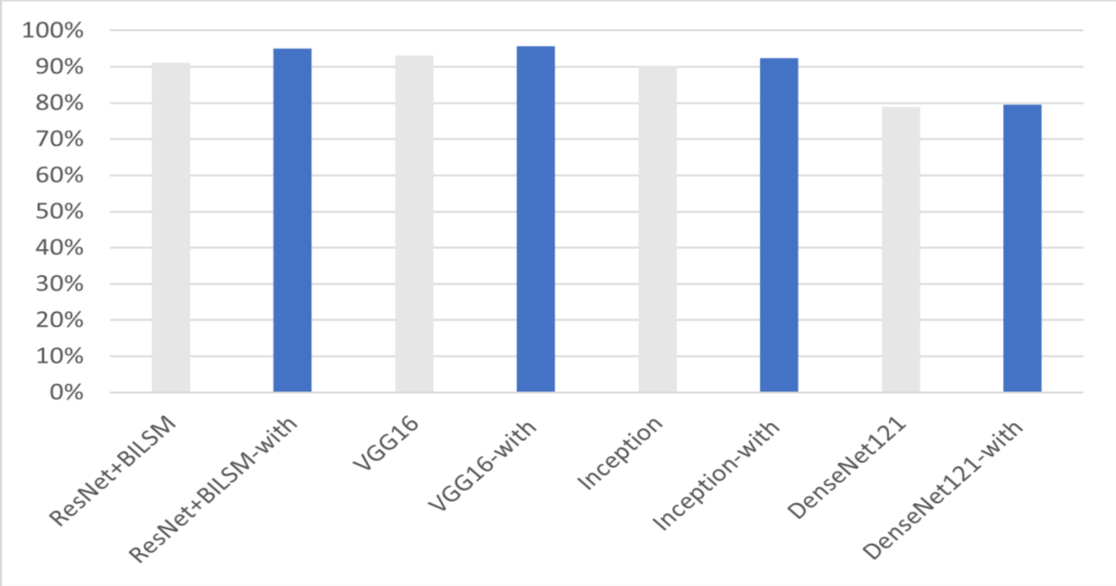

Figure 8. Accuracy results of models with /without skull stripping

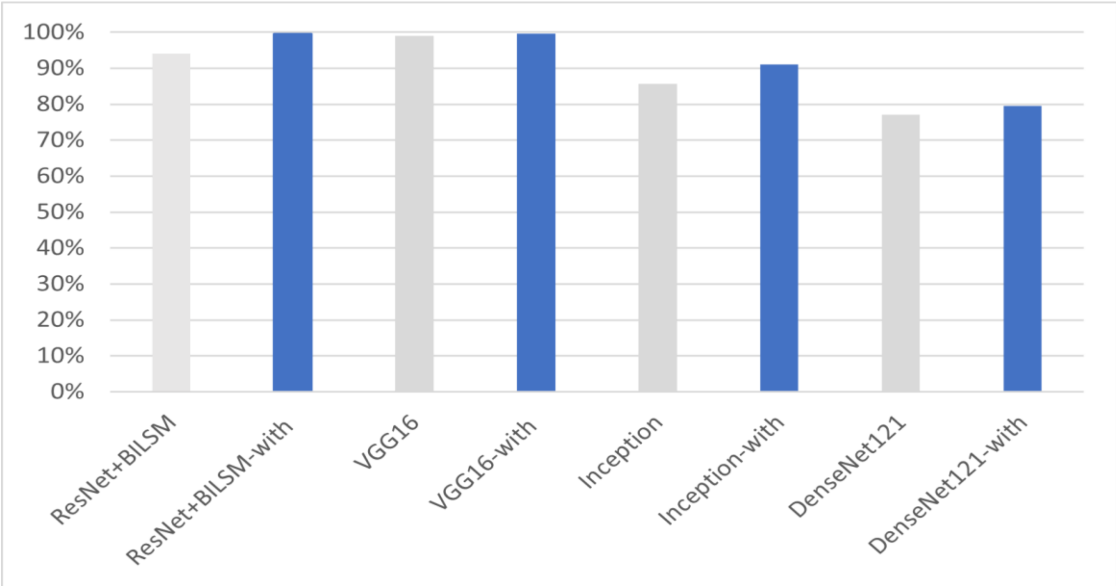

Figure 9. F1 score results of models with /without skull stripping

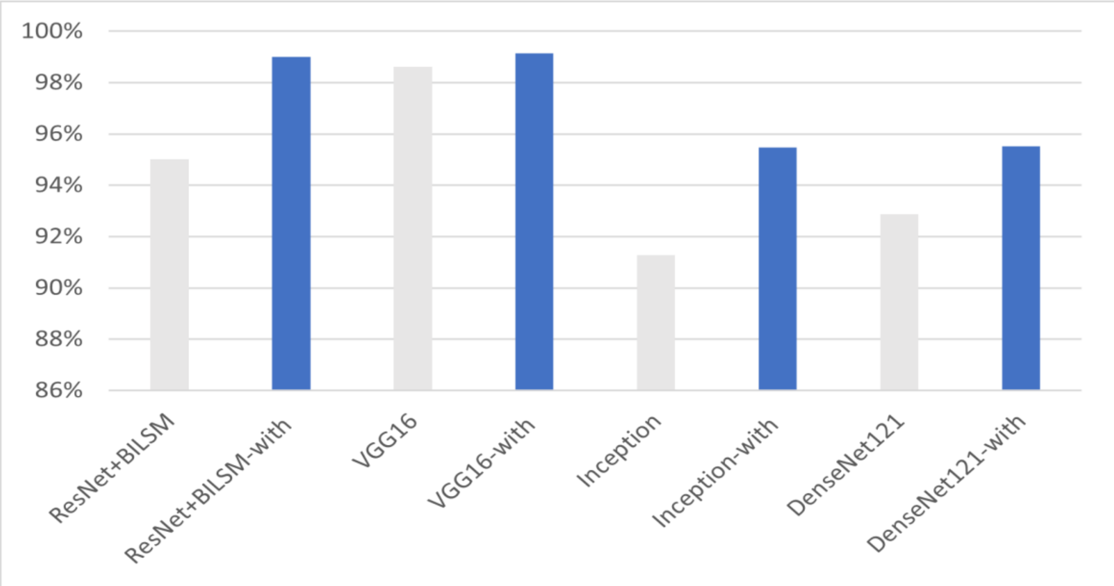

Figure 10. AUC results of models with/without skull stripping

5. CONCLUSION

In this research, the effect of the skull abstraction process on the models is studied, four models were utilized for this study VGG16, Inception, DensNet121, and the proposed model which integrates CNN features in extracting features from MRI images, and RNN for classification, researchers resorted to using CNN ResNet50 layers, due to its high ability to extract features and integrate it with the Bidirectional long short-term memory algorithm(BiLSTM). we use reliable datasets from ADNI with four classes AD, NC, EMCI, and MCI. We suggest improving the work of the models by increasing the accuracy of MRI images, so we proposed a novel hybrid process for skull stripping by integrating both thresholding and morphological Operations with U-net our method deals with various medical imaging modalities involving MRI, and it achieved highly accurate comparing with others method extracting the brain from a human skull and improve the results for VGG16 accuracy 99.6%, f1 score 95.7%, and AUC 99.14%, Inception accuracy 91%, f1 score 92.4%, and AUC95.4%, DenseNet121 accuracy79.5%, f1 score82.2, and AUC 95.51, and ResNet+Bilstm accuracy 99.5%, f1 score 95%, and AUC 99%.There are also some challenges recognized in this study such as training U-Net can be time-consuming, especially with large datasets, despite these challenges, combining thresholding, morphological operations, and U-Net represents a promising approach for achieving accurate and efficient skull stripping. This hybrid approach can be particularly beneficial for large-scale studies where both accuracy and efficiency are crucial.

References

[1] Arafah, A., Azher, M., et al., "The Future of Precision Medicine in the Cure of Alzheimer's Disease," *Biomedicines*, Vol. 11, No. 2, p. 335, Jan. 2023.

[2] van der Flier, W. M., de Vugt, M. E., Smets, E. M. A., et al., "Towards a future where Alzheimer's disease pathology is stopped before the onset of dementia," *Nature Aging*, Vol. 3, pp. 494–505, 2023.

[3] Kumar, A., Sidhu, J., Goyal, A., Tsao, J. W., and Doerr, C., "Alzheimer disease (nursing)," *Artificial Intelligence Review*, Vol. 53, pp. 5455–5516, 2021.

[4] Angkoso, C. V., Tjahyaningtijas, H. P. A., Purnama, I., and Purnomo, M. H., "Multiplane Convolutional Neural Network (Mp-CNN) for Alzheimer's disease classification," *International Journal of Intelligent Engineering & Systems*, Vol. 15, No. 1, pp. 1–9, 2022.

[5] Suresha, H. S., and Parthasarathy, S. S., "Diagnosis of Alzheimer Disease Using Fast Independent Component Analysis and Otsu Multi-level Thresholding," *International Journal of Intelligent Engineering & Systems*, Vol. 11, No. 5, pp. 32–39, 2018.

[6] Yamanakkanavar, N., Choi, J. Y., and Lee, B., "MRI segmentation and classification of the human brain using deep learning for diagnosis of Alzheimer's disease: a survey," *Sensors (Switzerland)*, Vol. 20, No. 11, pp. 1–31, 2020.

[7] Noor, M. B. T., Zenia, N. Z., Kaiser, M. S., Mamun, S. A., and Mahmud, M., "Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease, and schizophrenia," *Brain Informatics*, Vol. 7, pp. 1–21, 2020

[8] Li, F., Tran, L., Thung, K. H., Ji, S., Shen, D., and Li, J., "A robust deep model for improved classification of AD/MCI patients," *IEEE Journal of Biomedical and Health Informatics*, Vol. 19, No. 5, pp. 1610–1616, 2015.

[9] Voulodimos, A., Doulamis, N., Doulamis, A., and Protopapadakis, E., "Deep learning for computer vision: A brief review," *Computational Intelligence and Neuroscience*, Vol. 2018, No. 1, pp. 7068349, 2018.

[10] X. Li, L. Ding, L. Wang, and F. Cao, "FPGA accelerates deep residual learning for image recognition," In *Proc. of the 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 837–840, December 2017.

[11] Katrawi, A. H., Abdullah, R., Anbar, M., AlShourbaji, I & ,.Abasi, A. K. (2020b). Straggler handling approaches in mapreduce framework: a comparative study. International Journal of Power Electronics and Drive Systems/International Journal of Electrical and Computer Engineering, 11(1), 375 . https://doi.org/10.11591/ijece.v11i1.pp375-382

[12] Singh, A., & Ramachandran, R. (2014). Study on the effectiveness of smart board technology in improving the psychological processes of students with learning disability. *Sai Om Journal of Arts & Education*, *1*(4), 1-6.

[13] Kaur, C., Al Ansari, M. S., Rana, N., Haralayya, B., Rajkumari, Y., & Gayathri, K. C. (2024). A Study Analyzing the Major Determinants of Implementing Internet of Things (IoT) Tools in Delivering Better Healthcare Services Using Regression Analysis. *Advanced Technologies for Realizing Sustainable Development Goals 5G, AI, Big Data, Blockchain and Industry 4.0 Applications*, 270.

[14] Band, S. S., Ardabili, S., Danesh, A. S., Mansor, Z., AlShourbaji, I & ,.Mosavi, A. (2022). Colonial competitive evolutionary Rao algorithm for optimal engineering design. Alexandria Engineering Journal, 61(12), 11537-11563 .https://doi.org/10.1016/j.aej.2022.05.018

[15] Wadeea, S. J. (2024). The Cultural Impact Of Generation Z On Neologism And Its Translation: Internet Slangs As A Model. *Educational Administration: Theory and Practice*, *30*(6), 2751-2756.

[16] Al-Shourbaji, I., Alhameed, M., Katrawi, A., Jeribi, F., & Alim, S. (2022). A Comparative Study for Predicting Burned Areas of a Forest Fire Using Soft Computing Techniques. In ICDSMLA 2020: Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications (pp. 249-260). Springer Singapore. https://doi.org/10.1007/978-981-16-3690-5_22

[17] Wadeea, S. J. (2024). The Cultural Impact Of Generation Z On Neologism And Its Translation: Internet Slangs As A Model. *Educational Administration: Theory and Practice*, *30*(6), 2751-2756.

[18] Kaur, C., Al Ansari, M. S., Dwivedi, V. K., & Suganthi, D. (2024). Implementation of a Neuro-Fuzzy-Based Classifier for the Detection of Types 1 and 2 Diabetes. *Advances in Fuzzy-Based Internet of Medical Things (IoMT)*, 163-178..

[19] Al-Janabi, S & ,.Al-Shourbaji, I. (2016). Cooperative Methodology to Generate a New Scheme for Cryptography. The 3rd International Congress on Technology, Communication and Knowledge (ICTCK), At: Islamic Azad University – Mashhad Branch, 1-9.

[20] Shiju, K. K., Breja, M., Mohanty, N., Ramachandran, R., & Patra, I. (2023). Importance of Special Education and Early Childhood General Education Teachers' Attitudes toward Culturally Linguistically Diverse People. *Journal for ReAttach Therapy and Developmental Diversities*, *6*, 1544-1549.

[21] Elkady, G., Sayed, A., Priya, S., Nagarjuna, B., Haralayya, B., & Aarif, M. (2024). An Empirical Investigation into the Role of Industry 4.0 Tools in Realizing Sustainable Development Goals with Reference to Fast Moving Consumer Foods Industry. *Advanced Technologies for Realizing Sustainable Development Goals 5G, AI, Big Data, Blockchain and Industry 4.0 Applications*, 193.

[22] Al-Janabi, S & ,.Al-Shourbaji, I. (2016, December). A smart and effective method for digital video compression. In 2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT) (pp. 532-538). IEEE. 10.1109/SETIT.2016.7939927

[23] Ramachandran, R., & Singh, A. (2014). The Effect of Hindustani Classical Instrumental Music Santoor in improving writing skills of students with Learning Disability. *International Journal of Humanities and Social Science Invention*, *3*(6), 55-60.

[24] Sravanthi, A. L., Al-Ashmawy, S., Kaur, C., Al Ansari, M. S., Saravanan, K. A., & Vuyyuru, V. A. (2023). Utilizing Multimodal Medical Data and a Hybrid Optimization Model to Improve Diabetes Prediction. *International Journal of Advanced Computer Science & Applications*, *14*(11).

[25] Subudhi, S., Aarif, M., Kumar, S., Younis, D., Verma, M. K., Ravi, K., & Shivakumari, G. (2024). Evaluating Blockchain's Potential for Secure and Effective Digital Identity Management. In *Recent Technological Advances in Engineering and Management* (pp. 100-104). CRC Press.

[26] Al-Khateeb, M. O., Hassan, M. A., Al-Shourbaji, I & ,.Aliero, M. S. (2021c). Intelligent Data Analysis approaches for Knowledge Discovery: Survey and challenges. İlköğretim Online, 20(5), 1782–1792. http://ilkogretim-online.org/?mno=59960

[27] Schuster, M., and Paliwal, K. K., "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, Vol. 45, No. 11, pp. 2673–2681, 1997.

[28] Graves, A., and Schmidhuber, J., "Framewise phoneme classification with bidirectional LSTM networks," *Proceedings of the 2005 IEEE International Joint Conference on Neural Networks*, Vol. 4, IEEE, pp. 2047–2052, 2005.

[29] Goenka, N., Goenka, A., and Tiwari, S., "Patch-based classification for Alzheimer disease using sMRI," *2022 International Conference on Emerging Smart Computing and Informatics (ESCI)*, IEEE, pp. 89–94, 2022.

[30] Das, D., and Kalita, S. K., "Skull stripping of brain MRI for analysis of Alzheimer's disease," *International Journal of Biomedical Engineering and Technology*, Vol. 36, No. 4, pp. 331–349, 2021.

[31] Roy, S., Butman, J. A., Pham, D. L., and Alzheimer's Disease Neuroimaging Initiative, "Robust skull stripping using multiple MR image contrasts insensitive to pathology," *NeuroImage*, Vol. 146, pp. 132–147, 2017.

[32] Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., and Biller, A., "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, Vol. 129, pp. 460–469, 2016.

[33] Druzhinina, P., and Kondrateva, E., "The effect of skull-stripping on transfer learning for 3D MRI models: ADNI data," *Medical Imaging with Deep Learning*, pp. 1–8, 2022.

[34] M. Raju, T. V. Sudila, V. P. Gopi, and V. S. Anitha, "Classification of mild cognitive impairment and Alzheimer's disease from magnetic resonance images using deep learning," In *Proc. of the 2020 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, pp. 52–57, November 2020.

[35] Pei, L., Ak, M., Tahon, N. H. M., Zenkin, S., Alkarawi, S., Kamal, A., and Colen, R., "A general skull stripping of multiparametric brain MRIs using 3D convolutional neural network," *Scientific Reports*, Vol. 12, No. 1, pp. 10826, 2022.

[36] Xing, X., Liang, G., Blanton, H., Rafique, M. U., Wang, C., Lin, A. L., and Jacobs, N., "Dynamic image for 3D MRI image Alzheimer's disease classification," *European Conference on Computer Vision*, Cham: Springer International Publishing, pp. 355–364, 2020.

[37] Anam, F., Ahmad, R., Shahid, B., Raza, M., Tahir, M., Madni, U., and Iqbal, J., "State-of-the-art traditional to the machine-and deep-learning-based skull stripping techniques, models, and algorithms," *Journal of Digital Imaging*, Vol. 33, pp. 1443–1464, 2020.

[38] Zelin, X., Deng, H., Liu, J., and Yang, Y., "Diagnosis of Alzheimer's Disease Based on the Modified TresNet," *Electronics*, Vol. 10, No. 16, Article ID: 10161908, 2021. doi: 10.3390/ELECTRONICS10161908.

[39] Monica, R., Sanal, K. P., Kumar, S., and Nair, A. H., "Detection of Alzheimer's disease in MRI images using different transfer learning models and improving the classification accuracy," *International Journal of Health Sciences (IJHS)*, Vol. 6, No. 3, pp. 11851–11869, 2022. doi: 10.53730/ijhs.v6ns3.8944.

[40] Afiya, B., and Prabha, S., "Multiclass diagnosis of Alzheimer's disease analysis using machine learning and deep learning techniques," *International Journal of Image and Graphics*, Vol. 23, No. 1, Article ID: 0219467824500311, 2023. doi: 10.1142/s0219467824500311.

[41] Bhat, S. Y., Naqshbandi, A., and Abulaish, M., "Skull stripping on multimodal brain MRI scans using thresholding and morphology," *The Imaging Science Journal*, Vol. 72, No. 4, pp. 466–478, 2024.

[42] Stoleru, G. I., and Iftene, A., "Transfer learning for Alzheimer's disease diagnosis from MRI slices: A comparative study of deep learning models," *Procedia Computer Science*, Vol. 225, pp. 2614–2623, 2023.

[43] Khaled, E., Amr, F., Essam, F., Nouh, M., Sherif, Y., Yakan, Z., and AbdelRaouf, A., "Alzheimer's Disease Classification Approach Using MRI Imaging," *2023 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, IEEE, pp. 465–472, September 2023.

[44] Hazarika, R. A., Maji, A. K., Kandar, D., Jasinska, E., Krejci, P., Leonowicz, Z., and Jasinski, M., "An approach for classification of Alzheimer's disease using deep neural network and brain magnetic resonance imaging (MRI)," *Electronics*, Vol. 12, No. 3, pp. 676, 2023.

[45] Simonyan, K., and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, Vol. 1409, No. 1556, pp. 1–15, 2014.

[46] Chen, L., Wang, R., Yang, J., Xue, L., and Hu, M., "Multi-label image classification with recurrently learning semantic dependencies," *Visual Computer*, Vol. 35, pp. 1361–1371, 2019.

[47] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, and A. Rabinovich, "Going deeper with convolutions," In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

[48] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, 2016.

[49] X. Li, L. Ding, L. Wang, and F. Cao, "FPGA accelerates deep residual learning for image recognition," In *Proc. of the 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 837–840, December 2017.