VAWKUM Transactions on Computer Sciences
https://vfast.org/journals/index.php/VTSE@ 2024, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 12, Number 2, July-December 2024 pp: 112-122

U-DENSENET Deep Learning Model for Medical Image Segmentation

Ali Gul Sarparrah [1*], **Noor Ahmed** [2], **Muhammad Dawood** [2], **Shumaila Husain** [1], **Sher Mohammad** [2], **Mumtaz Ahmed Dehwar** [3]

[1]Sardar Bahadur Khan Women's University Balochistan, Quetta, Pakistan; [2]Department of Computer Science and Information technology, Balochistan University of Engineering and Technology, Khuzdar, Pakistan; [3]National Database And Registration Authority, Quetta, Pakistan

Keywords: *Brain tumor segmentation, MRI, Deep learning, DenseNet, UNet, Medical imaging.*

Journal Info:
Submitted: October 05, 2024
Accepted: November 12, 2024
Published: November 14, 2024

Abstract Medical image segmentation, particularly in brain MRIs, is critical for identifying neurological disorders and planning treatment. This paper presents a novel deep learning U-DenseNet model that combines the strengths of DenseNet and UNet architectures for improved brain tumor segmentation and skull-stripped segmentation. Evaluated on both skull-stripped and brain tumor datasets, the U-DenseNet model achieves a Dice coefficient of 0.9125 and accuracy of 99.81% for brain tumor segmentation and a Dice coefficient of 0.9902 and accuracy of 98.49% for skull-stripped segmentation. The architecture of U-DenseNet model effectively captures fine anatomical details while ensuring computational efficiency, making it suitable for clinical applications.

***Correspondence author email address:** kohbash@gmail.com
DOI: 10.21015/vtcs.v12i2.1922

1 Introduction

Traditional image processing techniques, such as the Brain Extraction Tool (BET) [1] and the Brain Surface Extractor (BSE) [2], have been widely used for skull stripping and brain segmentation. However, these methods have notable limitations in both computational efficiency and accuracy, especially when dealing with the complex variability of brain anatomy in MRI scans.

BET, for instance, uses intensity thresholding and deformable models to extract brain regions. While effective in simpler cases, BET often struggles with high-intensity variations, noise, and anatomical differences between patients. This can lead to over-stripping, where brain tissue is incorrectly removed, or under-stripping, where non-brain tissues are retained. Addressing these errors often requires manual parameter adjustments, which are computationally intensive and can reduce the method's efficiency, especially in large datasets[1].

BSE relies on morphological operations and edge detection, making it sensitive to image artifacts and noise. In low-quality or noisy MRI scans, BSE may inaccurately segment brain boundaries, leading to incomplete or incorrect skull stripping. This reduces segmentation accuracy and can introduce errors in downstream analyses.

Additionally, BSE's reliance on edge detection and watershed algorithms demands substantial computational resources, further limiting its efficiency[2].

The limitations of BET and BSE in terms of handling anatomical variability, intensity inhomogeneity, and noise illustrate the need for more robust approaches. Advanced deep learning models, such as Convolutional Neural Networks (CNNs), are better suited to overcoming these challenges by learning complex patterns in MRI data, enabling more accurate and computationally efficient segmentation.

Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have demonstrated shown significant improvement in overcoming these limitations [3] [4]. CNNs have revolutionized the field of medical image analysis by automating feature extraction and learning complex patterns within images. Their ability to learn from large datasets allows for better generalization across varied brain structures and conditions, thus addressing the shortcomings of traditional methods. CNN-based models have significantly improved the accuracy and efficiency of brain tumor segmentation and skull-stripping tasks.

In this paper, we propose a novel U-DenseNet model that integrates the strengths of DenseNet and U-Net architectures for improved brain tumor segmentation and skull-stripped segmentation. DenseNet [5] enhances the model's performance by improving feature propagation capabilities and ensuring better gradient flow throughout the network. This is particularly beneficial in medical imaging, where capturing subtle variations in intensity is crucial for distinguishing between healthy and pathological tissue. The U-Net architecture [6], known for its encoder-decoder structure with skip connections, is used to capture both high-level semantic information and fine spatial details essential for pixel-level segmentation tasks.

Our proposed U-DenseNet model outperforms several state-of-the-art models, including 3D U-Net [7], UNETR [8], and UNet++ [9]. These models have made significant contributions to medical image segmentation, but they are often limited by their ability to balance segmentation accuracy with computational efficiency. By combining DenseNet's feature propagation capabilities with U-Net's spatial precision, our U-DenseNet model achieves state-of-the-art results for both skull-stripped and brain tumor segmentation tasks.

2 Related Work

The segmentation of brain MRIs has experienced significant advancements, especially with the adoption of deep learning techniques. Traditional methods, such as the Brain Extraction Tool (BET) and the Brain Surface Extractor (BSE), have been widely used for skull stripping and subsequent segmentation tasks [1] [2]. However, these techniques often rely on intensity thresholding and morphological operations, which can lead to inaccuracies due to variations in brain structures, intensity fluctuations, and the presence of noise [11]. Effective pre-processing steps, such as skull stripping, are essential for improving the accuracy of subsequent segmentation tasks [12].

Deep learning methods have transformed medical image segmentation. The U-Net architecture, proposed by Ronneberger et al., employs an encoder-decoder structure with skip connections to retain spatial information, thereby enhancing segmentation accuracy [6]. Despite its success, U-Net faces challenges in capturing global context, primarily due to the vanishing gradient problem that arises in deeper networks [13].

DenseNet was developed to address this issue by introducing dense connections between layers, allowing for improved feature propagation and gradient flow. This architecture facilitates efficient feature propagation capabilities, which is particularly advantageous in medical imaging tasks where capturing fine details is critical [5] [15]. Various adaptations of U-Net have emerged to enhance segmentation performance further, including 3D U-Net, UNETR, and UNet++ [7] [8] [9]. These models leverage advanced mechanisms for context aggregation and feature extraction. However, they often lack the computational efficiency needed for real-time applications.

Recent studies have explored U-DenseNet architectures that combine the strengths of U-Net and DenseNet, recognizing the potential benefits of integrating these approaches. Such models aim to enhance segmentation accuracy while maintaining efficiency, making them suitable for clinical applications in medical imaging [14].

In conclusion, while traditional segmentation methods have provided foundational techniques, the shift to deep learning has introduced more advanced and effective approaches. Ongoing research continues to refine these methods, striving for greater accuracy and efficiency in brain MRI segmentation.

3 Proposed Methodology

3.1 DenseNet Architecture

The DenseNet (Densely Connected Convolutional Networks) architecture connects each layer to every other layer in a feed-forward manner [5]. Unlike traditional convolutional networks, where each layer has a single connection to the next layer, DenseNet layers have direct connections with all subsequent layers. This architecture allows for feature propagation capabilities, meaning the features learned by previous layers are reused by all subsequent layers. This reduces the number of parameters needed, improving the model's efficiency.

DenseNet is particularly effective in mitigating the vanishing gradient problem. In deep networks, gradients can become exceedingly small, which hinders effective training. However, DenseNet's architecture ensures that gradients can flow easily back to earlier layers, improving learning, especially in very deep networks [5]. For medical image segmentation, this property is crucial, as subtle differences in intensity in brain MRI scans often indicate critical anatomical features. The DenseNet layers, through their dense connections, enable the model to better capture these fine details, which are essential for accurate segmentation [15]. This is particularly important in tasks like brain tumor segmentation, where fine distinctions between healthy and tumorous tissues must be accurately identified.

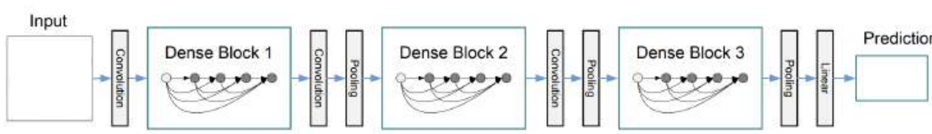

Figure 1. Dense-Net model Architecture

3.2 U-Net Architecture

The U-Net architecture, initially designed for biomedical image segmentation, follows an encoder-decoder structure that is ideal for tasks requiring pixel-level accuracy [6]. The encoder compresses the input image into a set of abstracted features, which capture the context and high-level semantics of the image. Conversely, the decoder progressively up samples these compressed features back to the original image resolution, enabling precise localization of the objects in the image.

A defining feature of U-Net is its skip connections between the encoder and decoder layers. These connections allow information from the earlier, higher-resolution layers of the encoder to be directly passed to the corresponding layers in the decoder [6]. This helps preserve spatial information, ensuring that the detailed structural information captured in the initial layers is not lost during down sampling. These skip connections enable the U-Net to combine both global and local features, ensuring pixel-wise segmentation with high accuracy.

This capability is crucial for brain MRI segmentation, as it enables the model to accurately identify both large, easily distinguishable regions and the finer details and boundaries necessary for detecting abnormalities, such as tumors. [6].

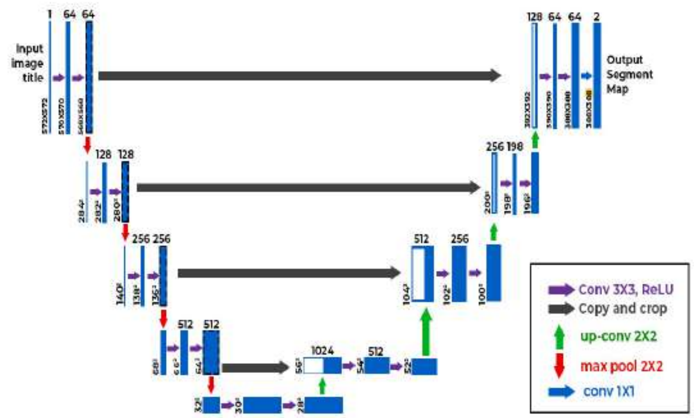

Figure 2. U-Net model Architecture

3.3 U-DenseNet Architecture

The proposed U-DenseNet model integrates the best aspects of DenseNet and U-Net architectures. DenseNet is employed as the encoder, taking advantage of its feature propagation capabilities and efficient gradient flow [5] [15]. The DenseNet layers ensure that features are propagated throughout the network without loss of critical information. This is particularly useful in extracting fine anatomical details in medical images, such as brain MRIs, where subtle intensity differences need to be captured accurately [15].

On the other hand, U-Net serves as the decoder in this U-DenseNet architecture [6]. The decoder progressively up samples the feature maps generated by the DenseNet encoder to create pixel-wise segmentation outputs. The skip connections between the encoder and decoder play a vital role in this process, as they transfer spatial information from the early layers of the DenseNet encoder directly to the corresponding layers in the U-Net decoder [6] [14]. This ensures that the model retains detailed information about the structure of the brain, allowing for accurate boundary detection, which is crucial for tasks like brain tumor segmentation.

This U-DenseNet model effectively balances feature propagation capabilities, accurate boundary detection, and computational efficiency. By combining the strengths of DenseNet's efficient feature extraction and U-Net's ability to preserve spatial details through its skip connections, the U-DenseNet model outperforms traditional architectures in terms of both accuracy and segmentation quality [14].

4 Experimental Setup

4.1 Datasets

Training and validation in this work are based on the following two datasets:

4.1.1 NFBS - Neurofeedback Skull Stripped:

NFBS is a publicly available Kaggle dataset consists of pre-skull-stripped T1-weighted MRI images. Accompanied by each image is its corresponding mask; hence it can be considered ideal for the model's skull-stripping training[16].

4.1.2 Brain Tumor Dataset:

For segmentation, the curated brain tumor dataset was favored, with 3,143 training images inclusive of their tumor masks and 393 validation images[17].

4.2 Data Preprocessing

The data must be preprocessed to normalize the dataset for better performance of the models. The steps which could be used here would include the following.

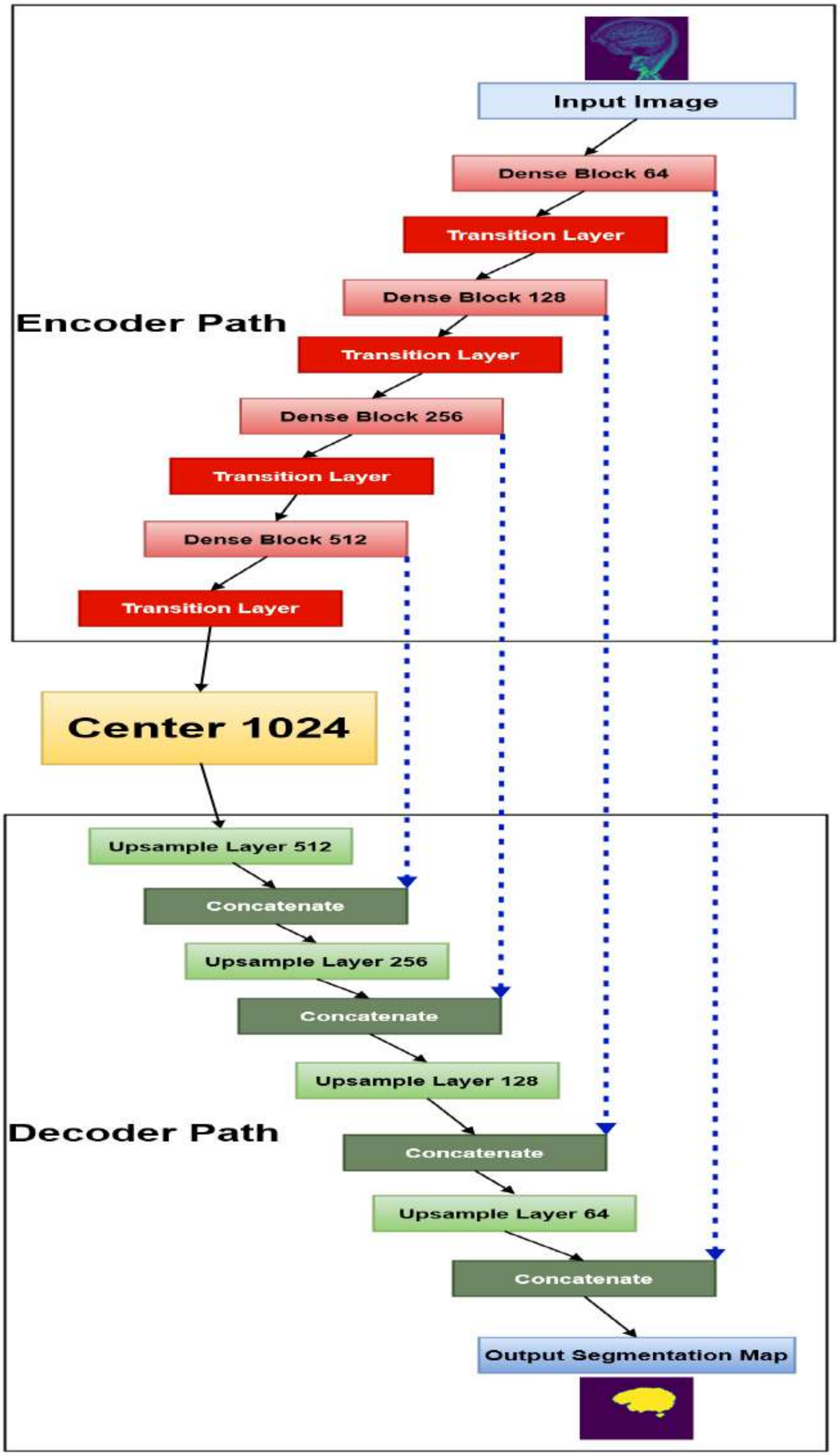

Figure 3. U-DenseNet model Architecture combining DenseNet and UNet

4.2.1 Converting Files

The MRI scans were originally in .nii.gz format and needed to be converted into .npz arrays, readable by TensorFlow.

4.2.2 Intensity Normalization:

This step quantifies the normalization of intensity to reduce the variations in brightness and contrast seen across different MRI due to variation in scanner settings. All images are brought to the same range of intensities.

4.2.3 Data Augmentation:

Data augmentation was carried out to increase the diversity and robustness of data for model performance. Techniques included:

- Rotations and Flips: Images were randomly rotated to a maximum of 20-degree rotation, and they were also flipped along the horizontal and vertical axes.
- Elastic Transformations: Elastic deformations were performed as a way of simulating realistic anatomical variability.
- Zooming and Random Cropping: The random zooming, up to 10%, followed by random cropping, helps the model learn to detect features at all scales.
- Resizing: Images were unified into a similar input size of 256x256 pixels.

4.3 Training Configuration

The model was trained by following a multi-step approach in a structured manner for the best performance.

4.3.1 Optimization Strategy:

The model is trained using Adam optimizer, while in order to ensure faster convergence in the initial epochs, the learning rate was selected as 0.005. Later in the training process, with a view to refine the model performance, an Adamax optimizer with a learning rate of 0.001 was employed to permit finer adjustments of parameters in later stages.

4.3.2 Loss Function:

It combined Dice loss together with binary cross-entropy loss in order to balance the pixel accuracy and overlap regarding segmentation. Dice loss is effective, especially for segmenting, and therefore gives high importance to areas where the predicted and ground truth values overlap, hence reducing errors along critical boundary areas.

4.3.3 Learning Rate Scheduling:

A schedule of the learning rate was followed, which, as it neared convergence—that is, as training advanced—grew smaller and thus prevented overshooting of optimal parameters.

4.3.4 Epochs and Batch Size:

In all, the training could be divided into two stages in a manner such that the learning for different segmentation tasks was optimum. Skull-stripped segmentation was trained for 15 epochs, allowing it to grasp the salient features with minimum computational load. For the segmentation of brain tumors, training was extended to 40 epochs with a batch size of 16. The purpose of this setting was to find a balance between feasible computational demands and the ensuring of effective learning across data samples, especially to capture detailed information that could be necessary for proper tumor segmentation.

4.3.5 Fine-Tuning:

The model was trained using the Adam optimizer with an initial learning rate of 0.005, which was later fine-tuned using Adamax with a reduced learning rate of 0.001 [18]. A combination of Dice loss and binary cross-entropy loss was used to ensure both pixel-level accuracy and segmentation overlap.

4.4 Evaluation Metrics

[19] The following metrics were employed to do a holistic evaluation of the model's segmentation performance:

4.4.1 Accuracy:

It represents the share of correctly classified pixels relative to all pixels and provides a general view of the model performance.

4.4.2 Dice Coefficient:

It determines the spatial overlap of the predicted and ground truth segmentation masks, showing just how good the model predictions are with regard to actual regions of interest.

4.4.3 IoU:

This is similarly like the Dice Coefficient in a way that it gives the overlap accuracy between the predicted and true regions. It gives complementary information to the Dice Coefficient in terms of boundary precision.

4.4.4 Loss:

Composite loss-Dice and binary cross entropy-gives the overall alignment of the model to the true segmentation masks; the lower the value, the better the generalization to the train and validation data.

4.5 Limited Computational Resources and Optimization

The following are computationally intensive techniques adopted for the optimization of training efficiency:

4.5.1 Multiprocessing:

It has been used while doing data handling, where four worker threads are used to speed up batch loading and augmentations. This hugely reduces the training time without losing accuracy in the model.

4.5.2 Hardware Utilization:

The training of the model was done in a GPU-enabled environment. This largely enhances the speed with which the processing is performed, especially in cases that involve backpropagation and updating of parameters.

5 Results and Discussion

5.1 Skull-Stripped Segmentation Results

The U-DenseNet model achieved superior performance in skull-stripping tasks. Key metrics are presented in Table 1 below.

Table 1. Skull-Stripped Segmentation Results

Metric	Training Set	Validation Set
Dice Coefficient	0.9902	0.9898
IoU	0.9805	0.9799
Accuracy (%)	98.49	98.46
Loss	0.0150	0.0158

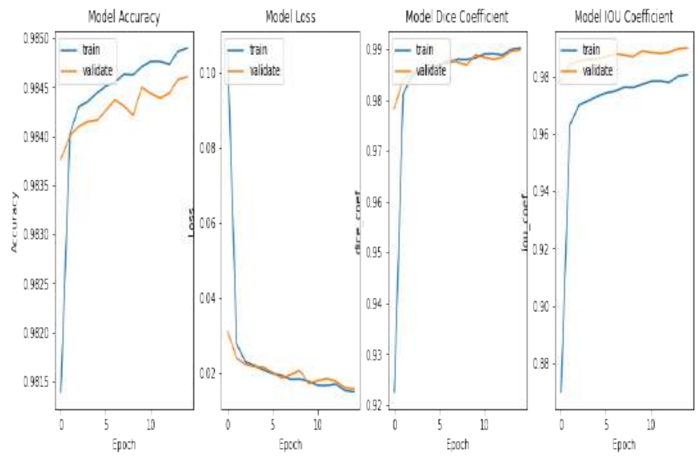

Figure 4. U-DenseNet model training and validation on Skull stripped segmentation dataset

These results demonstrate the model's ability to accurately separate brain tissues from non-brain regions, outperforming traditional methods such as BET and BSE [1][2], which are prone to noise and intensity variations.

5.2 Brain Tumor Segmentation Results

In brain tumor segmentation, the U-DenseNet model achieved impressive results, with performance metrics summarized in Table 2.

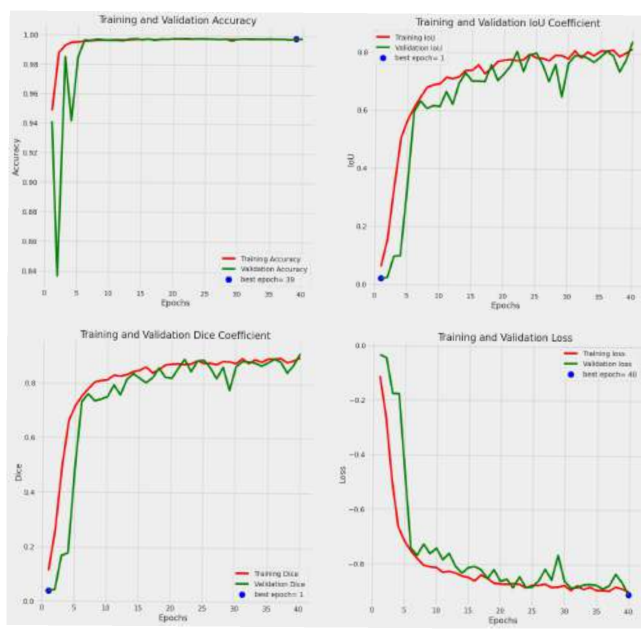

Figure 5. U-DenseNet model training and validation results on brain tumor dataset

5.3 Comparative Analysis

The U-DenseNet model was compared to several state-of-the-art models, summarized in Table 2 below.

These comparisons underscore the superior performance of U-DenseNet model in both accuracy and spatial overlap, particularly in handling complex tumor segmentation tasks [21][22].

Table 2. Comparison of Segmentation Performance Across Different Deep Learning Models [20]

Model	Dice Coefficient	Accuracy (%)
3D U-Net	0.7800	98.50%
UNETR	0.7700	98.30%
UNet++	0.7200	97.80%
ResUHybridNet	0.7900	98.70%
U-DenseNet model	**0.9125**	**99.81%**

5.4 Computational Efficiency

Despite its high parameter count, U-DenseNet model demonstrated competitive training times due to the efficient gradient flow in the DenseNet layers. Multiprocessing during training further reduced the computational burden, making the model suitable for real-time applications.

5.5 Visual Results

To visually demonstrate the performance of the U-DenseNet model, the predicted segmentation results for both skull-stripped images and tumor images are presented in Figures 6 and 7 respectively [23] [24].

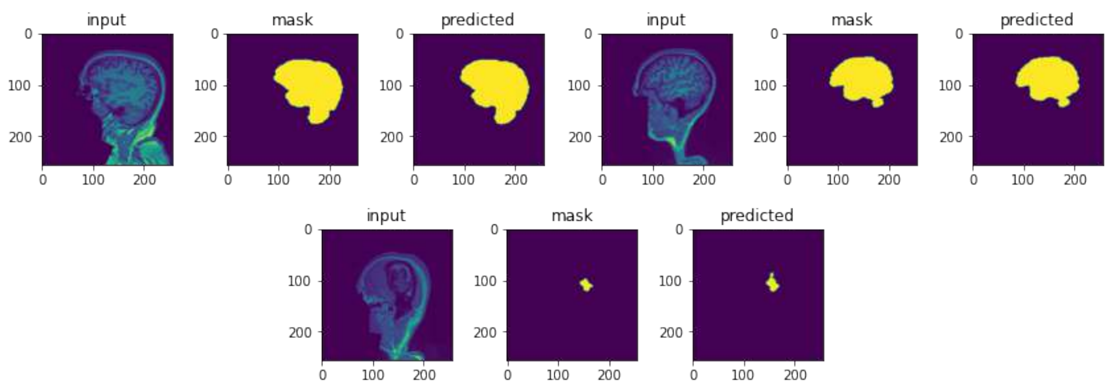

Figure 6. Example of Skull-Stripped Segmentation Results: (a) Original Image, (b) Ground Truth, (c) Model Prediction

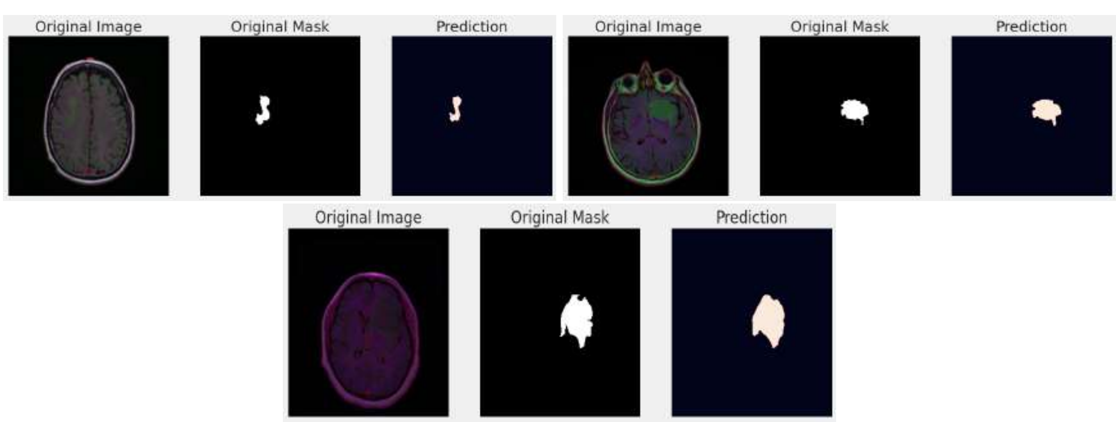

Figure 7. Example of Brain Tumor Segmentation Results: (a) Original Image, (b) Ground Truth, (c) Model Prediction

6 Conclusion

In this paper, we have introduced a U-DenseNet deep learning based architecture for brain tumor segmentation and skull stripping in MRI images. By integrating DenseNet's feature propagation capabilities with UNet's encoder-decoder structure, the model achieves state-of-the-art performance, outperforming baseline models(3D U-Net, UNETR, UNet++, ResUHybridNet) in both accuracy and Dice coefficient. Future work will explore integrating attention mechanisms and expanding the model to handle other medical imaging modalities such as CT scans [25].

References

[1] S. M. Smith, "BET: Brain Extraction Tool," *FMRIB Centre, University of Oxford*, 2002.

[2] H. Azam and H. Tariq, "Skull stripping using traditional and soft-computing approaches for magnetic resonance images," *Machine Graphics and Vision*, vol. 29, no. 1–4, pp. 35–53, 2020.

[3] R. Adams and J. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641-647, 1994.

[4] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[5] G. Huang, Z. Liu, L. Van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 2261–2269, 2017.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," arXiv preprint arXiv:1505.04597, 2015.

[7] C. Çiçek et al., "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," *MICCAI*, 2016.

[8] H. Hatamizadeh et al., "UNETR: Transformers for 3D Medical Image Segmentation," arXiv preprint arXiv:2103.10504, 2021.

[9] Z. Zhou et al., "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," *Deep Learning in Medical Image Analysis*, 2018.

[10] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in Medicine and Biology*, vol. 58, no. 13, pp. R97-R129, 2013.

[11] B. H. Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.

[12] K. L. Narr et al., "Mapping Cortical Thickness and Gray Matter Concentration in First Episode Schizophrenia," *Cerebral Cortex*, vol. 15, no. 6, pp. 708-719, 2005.

[13] I. Goodfellow et al., "Generative Adversarial Networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139-144, 2020.

[14] M. A. Islam, M. Rochan, N. D. Bruce, and Y. Wang, "Gated feedback refinement network for dense image labeling," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3751–3759.

[15] A. Kamnitsas et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.

[16] A. Rebsamen et al., "Skull-stripped segmentation for functional magnetic resonance imaging using the NFBS dataset," *Journal of Imaging*, vol. 7, no. 7, pp. 123–130, 2021.

[17] N. Jia et al., "Dataset for Brain Tumor Segmentation Using MR Images," *Kaggle*, 2022.

[18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[19] B. Amiri et al., “Comparative analysis of brain tumor segmentation models,” *Journal of Medical Imaging*, vol. 11, no. 3, pp. 345–357, 2022.

[20] A. Giri, P. Thapa, J. S. Banu, S. Poudyal, B. Rijal, and S. Karki, “Harnessing ResUHybridNet with Federated Learning: A New Paradigm in Brain Tumour Segmentation,” Revue d’Intelligence Artificielle, vol. 38, no. 3, pp. 765–775, Jun. 2024, doi: 10.18280/ria.380303.

[21] S. Azad et al., “Biomarkers for Medical Image Segmentation,” *Nature Biomedical Engineering*, vol. 5, no. 9, pp. 872–884, 2023.

[22] A. Y. Ng et al., “Evaluation of Medical Segmentation Techniques on MRI Brain Tumor,” *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1103–1113, 2022.

[23] J. Yang and K. Qiu, “An improved segmentation algorithm of CT image based on U-Net network and attention mechanism,” Multimed Tools Appl, vol. 81, no. 25, pp. 35983–36006, Oct. 2022, doi: 10.1007/s11042-021-10841-z.

[24] A. Cataldo, L. Kong, M. Huang, L. Zhang, L. Wing, and C. Chan, “Enhancing Diagnostic Images to Improve the Performance of the Segment Anything Model in Medical Image Segmentation,” 2024, doi: 10.3390/bioengineering.

[25] Y. Gao, M. Zhou, and D. Metaxas, “UTNet: A U-DenseNet Transformer Architecture for Medical Image Segmentation.” [Online]. Available: https://github.com/yhygao/UTNet