

Brain Extraction from Magnetic Resonance Images Using UNet modified with Residual and Dense Layers

Kali GÜRKAHRAMAN^{1*} , Çağrı DAŞGIN² 

¹ Sivas Cumhuriyet University, Faculty of Engineering, Department of Computer Engineering, Sivas, Türkiye

² Sivas Cumhuriyet University, Institute of Science, Department of Computer Engineering, Sivas, Türkiye

Kali GÜRKAHRAMAN ORCID No: 0000-0002-0697-125X

Çağrı DAŞGIN ORCID No: 0009-0006-9932-5789

*Corresponding author: abc@bingol.edu.tr

(Received: 08.08.2023, Accepted: 27.09.2023, Online Publication: 27.09.2023)

Keywords

Brain extraction,
Skull stripping,
Deep learning,
Dense connection,
Residual connection,
UNet

Abstract: The main goal of brain extraction is to separate the brain from non-brain parts, which enables accurate detection or classification of abnormalities within the brain region. The precise brain extraction process significantly influences the quality of successive neuroimaging analyses. Brain extraction is a challenging task mainly due to the similarity of intensity values between brain and non-brain structure. In this study, a UNet model improved with ResNet50 or DenseNet121 feature extraction layers was proposed for brain extraction from Magnetic Resonance Imaging (MRI) images. Three publicly available datasets (IBSR, NFBS and CC-359) were used for training the deep learning models. The findings of a comparison between different feature extraction layer types added to UNet shows that residual connections taken from ResNet50 is more successful across all datasets. The ResNet50 connections proved effective in enhancing the distinction of weak but significant gradient values in brain boundary regions. In addition, the best results were obtained for CC-359. The improvement achieved with CC-359 can be attributed to its larger number of samples with more slices, indicating that the model learned better. The performance of our proposed model, evaluated using test data, is found to be comparable to the results obtained in the literature.

Artık ve Yoğun Katmanlarla Değiştirilmiş UNet Kullanılarak Manyetik Rezonans Görüntülerinden Beyin Çıkarımı

Anahtar Kelimeler

Beyin çıkarımı,
Kafatası soyma,
Derin öğrenme,
Yığın bağlantı,
Artık bağlantı,
UNet

Öz: Beyin çıkarımının temel amacı, beyni beyin dışı kısımlardan ayırarak beyin bölgesi içindeki anormalliklerin doğru tespitini veya sınıflandırılmasını mümkün kılmaktır. Hassas beyin çıkarma işlemi, ardışık nörogörüntüleme analizlerinin kalitesini önemli ölçüde etkiler. Beyin çıkarımı, beyin ve beyin dışı yapılar arasındaki yoğunluk değerlerinin benzerliği nedeniyle zorlu bir görevdir. Bu çalışmada, Manyetik Rezonans Görüntüleme (MRG) görüntülerinden beyin çıkarımı için ResNet50 veya DenseNet121 özellik çıkarma katmanları ile geliştirilmiş bir UNet modeli önerilmiştir. Derin öğrenme modellerini eğitmek için IBSR, NFBS ve CC-359 adlı üç halka açık veri kümesi kullanılmıştır. UNet'e eklenen öznetelik çıkarma katman türleri arasındaki karşılaştırma sonuçları, ResNet50'den alınan artık bağlantıların tüm veri kümelerinde daha başarılı olduğunu göstermektedir. ResNet50 bağlantılarının, beyin sınır bölgelerindeki zayıf ancak önemli gradyan değerlerinin ayrımını artırmada etkili olduğu anlaşılmaktadır. Ayrıca, en iyi sonuçlar CC-359 için elde edilmiştir. CC-359 ile elde edilen gelişme, verisetinin daha fazla kesit ve örnek içermesinden dolayı modelin daha iyi öğrenmesinden kaynaklanmıştır. Önerilen modelin performansı, test verileri kullanılarak değerlendirildiğinde, literatürde elde edilen sonuçlarla karşılaştırılabilir bulunmuştur.

1. INTRODUCTION

The non-invasive Magnetic Resonance Imaging (MRI) technique provides rich information about the examined anatomical structure due to its high spatial resolution. It is widely used in the early diagnosis and assessment of many diseases, since it provides the detection of changes in brain structures that may develop even in micro dimensions over time. However, due to its high resolution, it also causes the detailed presence of non-brain structures such as skull, scalp and eyeballs in the MRI scans [1]. The removal of non-brain structures may have a significant impact on the subsequent analyses related to the brain.

Brain extraction, also known as skull stripping, is the initial step in analyzing MRI and other neuroimaging data, involving the separation of the brain from non-brain parts [2-3]. Its primary goal is to remove the skull and provide leverage in obtaining high accuracy for the detection or classification of any abnormalities within the relevant brain region [3, 4]. Accurate brain extraction significantly impacts the quality of neuroimaging studies, including image registration, brain tumor or lesion segmentation, measurement of brain regions of interest for global and neurodegenerative diseases, detection of cortical thickness, and planning neurosurgical interventions [2].

Brain extraction is a challenging task mainly due to the similarity of intensity values between brain and non-brain structure [3]. The partial volume effect blurs the boundaries between two tissues, making it difficult to distinguish structures especially in low contrast brain images. Interpreting brain structures that lack sharp edges in brain images, and the presence of unwanted signals (artifacts) at air/tissue boundaries in brain images, further adds to the complexity. Moreover, motion artifacts from the patient and noise from the imaging environment can reduce image quality and increase the difficulty of brain extraction [1].

In the literature, manual segmentation is considered the “gold standard” for brain extraction [5-6]. However, this method is not only labor intensive and time consuming, but also exhibits significant inter-individual variability, potentially introducing analysis bias and thus impeding the reproducibility of clinical studies [2]. Therefore, in recent years, many semi-automatic or fully automatic brain extraction techniques have been proposed with the aim of overcoming the drawbacks encountered in manual segmentation [1].

In the morphological and intensity-based initial brain extraction methods, selecting the most appropriate threshold value(s) to separate foreground and background is often challenging [1, 3]. Brain Extraction Tool (BET) and BET2 techniques developed by Smith [5] and Jenkinson et al. [7], define an initial sphere by determining the center of gravity of the head and deforming it until it reaches the brain edge. Brain Surface Extraction (BSE) technique developed by Shattuck et al. [8] uses an edge-based approach with anisotropic

diffusion filtering. It has been reported that BET works poorly on neck-intense images and BSE has lower performance on low-resolution images [1]. Moreover, BET and BSE techniques require parameters to be optimized for each image, making them challenging to use in large-scale studies [1, 6].

Atlas/template-based brain extraction methods involve adapting an atlas/template MRI brain image to reveal relationships between brain regions, thereby separating structures with no relationship into brain and non-brain. The widely used atlas-based Brain Extraction using Nonlocal Segmentation Technique (BEaST) is fast and achieved successful performance on T1-weighted MRI images of both healthy individuals and Alzheimer's patients [9]. However, this method requires parameter optimization depending on the dataset used.

Hybrid techniques for brain extraction involve combining the results of multiple methods. Souza et al. [6] prepared a dataset called Calgary-Campinas-359 (CC-359), which contains 359 T1-weighted MRI images of healthy individuals. In their study, the images were segmented with eight different methods, and the segmentation error was reduced by evaluating the results of multiple methods using the expectation-maximization technique. However, the hybrid technique benefits from results of techniques like BEaST, BET, and BSE, which require parameter optimization [6].

Recently, deep learning (DL) techniques, which have achieved successful results in medical image analysis and imaging, have demonstrated the potential to outperform medical experts in solving specific problems [4, 10-12]. DL models proposed for brain extraction are also available in the literature [13]. Kleesiek et al. [14] developed a deep convolutional neural network (CNN) model for brain extraction from MRI images in three different open datasets and achieved high-performance results. Isensee et al. [2] developed a 3D-UNet-based method called HD-BET, which performed brain extraction by examining brain images of healthy individuals as well as those with various pathologies, apart from MRI images of healthy individuals. They trained the DL model with images obtained from different sequences (T1, T2, and FLAIR) and MRI devices, and tested it on open brain image datasets, achieving successful results. Similarly, Hwang et al. [15] used a modified 3D-UNet model from a 2D-UNet model for brain extraction from T1-weighted MRI images. The developed DL model achieved high performance compared to traditional models, but its performance on MRI images obtained from different devices is uncertain. Zhang et al. [16] modified the 3D-UNet model and developed the FRNET model, in which residual layers were added between encoder and decoder blocks, and a new boundary loss function was used during model training. FRNET was tested only on infant MRI dataset and achieved high dice score values for brain extraction. Similarly, Dasgin and Gurkahraman [17] showed the effectiveness of the 3D-UNet model modified with residual connections in brain extraction. Hoopes et al. [18] used generative DL model to synthesize medical images

and then generalized real brain images from these synthesized images using a 3D-UNet model, which they called SynthStrip. SynthStrip significantly improved performance values obtained by traditional methods for different datasets.

In summary, DL-based automatic brain extraction techniques are faster than traditional methods and do not require parameter optimization. However, the selection of the brain extraction method depends on the problem and is influenced by the characteristics of MRI images. Therefore, when choosing a brain extraction technique, factors such as the model of the MRI scanner and magnetic field intensity (Tesla) should be taken into account. In image segmentation studies, previous studies have shown that the 3D- UNet model outperformed CNN models and that enriching the latent layer of the 3D- UNet model with residual and dense connections between encoder and decoder blocks improved accuracy performance. Therefore, this study aims to perform brain extraction on MRI images obtained from different MRI scanners with different magnetic field intensity values by adding feature extraction layers of ResNet and DenseNet architectures to the UNet model.

2. MATERIAL AND METHOD

In this study, a UNet model improved with residual and dense connections was proposed for brain extraction from MRI images. In the following subsections, the details of the datasets used and the developed method are presented.

2.1. Datasets

In this study, three different publicly available datasets, The Internet Brain Segmentation Repository (IBSR) [19], The Neurofeedback Skull-stripped (NFBS) [20], and CC-359 [6], were utilized for training and testing the 3D-DL model.

The IBSR dataset [19] includes 3D-T1 weighted gray images acquired from 18 healthy subjects using a 1.5 Tesla Siemens Magnetom MRI scanner, along with manually segmented brain mask images by experts.

The NFBS dataset [20] consists of images from 125 participants with various clinical and psychiatric histories, acquired using a 3T Siemens Magnetom TIM Trio scanner with a resolution of $1 \times 1 \times 1 \text{ mm}^3$. The brain mask images were also manually segmented by experts.

The CC-359 dataset [6] includes brain images with a slice thickness of $1 \times 1 \times 1 \text{ mm}^3$ from 359 healthy individuals aged between 29 and 80 years, acquired using 1.5/3.0

Tesla Philips, Siemens, and GE MRI scanners. In this study, the images were saved in NIFTI (.nii) format without data loss, and both gray and mask images were created. Note that not all images in the dataset have manual segmentations available. Thus, the brain mask images recommended as the silver standard by Souza et al. [6], using the Simultaneous Truth and Performance Level Estimation (STAPLE) technique, were used as ground truth in this study.

2.2. UNet-based 3D Deep Learning Model

Autoencoder is a data compression algorithm that automatically perform compression and decompression functions specific to the data, and the autoencoder in the DL model implements these functions using neural networks. To create an autoencoder, a coding function, a decoding function, and an error function showing the information loss between the compressed and uncompressed representations of the data are required [21]. The UNet model, fundamentally an autoencoder, consists of an encoder that encodes the input image into low-level features at multiple levels and a decoder that reflects these features into pixel space. With its symmetrical structure, UNet uses skip connections between the encoder and decoder to preserve image information in greater detail [22, 23].

Figure 1 shows the 3D UNet model modified with the feature extraction layers of the ResNet50 [24] and DenseNet121 [25] models proposed for ImageNet. The model consists of two different blocks, encoder, and decoder, each containing convolution and deconvolution layers. In the middle part of this general structure, the residual/dense blocks of ResNet50 and DenseNet121 models were adapted to achieve better compression of information, and the brain extraction results were compared.

In the encoder blocks, four convolution blocks with filter sizes of 8, 32, 64, and 64 were used, respectively. Each block consists of a $3 \times 3 \times 3$ convolution operation, followed by batch normalization (BN) and ReLU activation. This process is repeated twice, and finally, the feature maps' dimensions are reduced using max pooling. In the decoder blocks, there are four upsampling blocks with filter sizes of 64, 64, 32, and 8, respectively. The operations in the decoder blocks consist of upsampling (transposed convolution), followed by two repetitions of Convolution, BN, and ReLU. The activation function used in the final block is sigmoid.

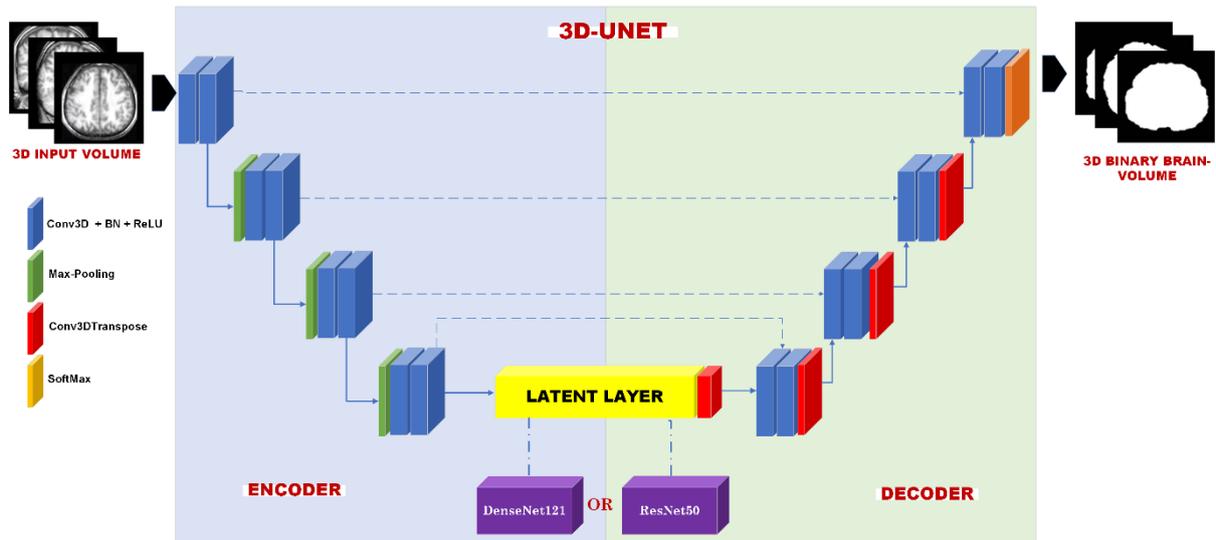


Figure 1. 3D-UNet model modified with feature layers of the DenseNet121 and ResNet50.

For the experiments, due to the limited number of samples in the IBSR dataset (only 18 samples), the datasets were split into 80% for training and 20% for testing, with the test set also serving as the validation set. A dataset's optimal train-test split ratio for DL applications is not clearly defined. In this study, we adopted the widely favored experimental train-test ratio of 80%-20%, as commonly seen in the DL medical image analysis literature [26-27]. The optimization technique used during training was ADAM, with a learning rate of 0.0002 and a momentum coefficient of 0.8. The number of epochs was set to 500 for IBSR and 100 for the other datasets, with early stopping as an option. The batch size was set to 4 due to hardware limitations for training the 3D model. The 3D-UNet model was implemented using the Keras library in Python 3.9 based on Tensorflow. The experiments were conducted on a PC with a NVIDIA RTX A6000 48 GB GPU, Intel i9 12900 KS @ 3.40 Hz CPU, and 64 GB RAM.

To compare the performances of the dense and residual blocks in the latent layer of the model under the same conditions, the rest of the architecture was designed to be the same, and the traditional loss function, binary-cross entropy, was used.

2.3. Performance Metrics

The segmentation performance of the UNet model was assessed using the Dice coefficient, sensitivity, and specificity metrics. The Dice coefficient, as given in Equation 1, is calculated by dividing twice the intersection of the real (R) and predicted (P) masks by the sum of the areas of both masks.

$$Dice = \frac{2|P \cap R|}{|P| + |R|} = \frac{2TP}{2TP + FP + FN} \quad (1)$$

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

Sensitivity, also known as recall, assesses the proportion of brain tissue that is successfully included in the segmentation. Specificity gauges the proportion of non-brain tissue that is accurately excluded from the segmentation (Equation 2).

$$Sensitivity = \frac{TN}{TN + FP} \quad (2)$$

$$Specificity = \frac{TP}{TP + FN} \quad (3)$$

3. EXPERIMENTAL RESULTS AND DISCUSSION

The experiments were performed by training and testing two separate UNet architectures modified with DenseNet121 and ResNet50 feature extraction layers. Both architectures were trained and tested on the IBSR, NFBS, and CC359 datasets. The training and testing datasets were randomly split to ensure the model was tested on data it had not seen during the training process. The training and testing procedures were repeated five times, and the average performance values were computed.

The results of all experiments are presented in Table 1. Each cell in the table shows the average and standard deviation of five test results. When comparing the different feature extraction layer types added to the UNet architecture, the most successful results were achieved with ResNet50 across all datasets. However, when considering individual datasets, the most successful outcomes were obtained with CC-359. The residual connections in the ResNet50 architecture particularly improved the distinction of weak but important gradient values in the brain boundary regions. On the other hand, according to the datasets, the improvement can be attributed to the larger number of samples and more slices in the CC-359 dataset, indicating that the model learned better.

Table 1. Average performance values of modified UNet models.

Dataset	Models	Dice Coefficient	Sensitivity	Specificity
IBSR	ResNet50	0.9613 (± 0.0036)	0.9590 (± 0.0017)	0.9931 (± 0.0010)
	DenseNet121	0.9473 (± 0.0063)	0.9211 (± 0.0230)	0.9955 (± 0.0016)
NFBS	ResNet50	0.9875 (± 0.0018)	0.9840 (± 0.0048)	0.9990 (± 0.0005)
	DenseNet121	0.9838 (± 0.0067)	0.9812 (± 0.0146)	0.9985 (± 0.0010)
CC-359	ResNet50	0.9887 (± 0.0028)	0.9899 (± 0.0058)	0.9985 (± 0.0009)
	DenseNet121	0.9872 (± 0.0025)	0.9874 (± 0.0069)	0.9985 (± 0.0011)

The sample results obtained with the IBSR, NFBS, and CC-359 datasets are presented in Figure 2-4. Upon examining the generated binary segmentation masks, they demonstrate consistency with the Dice scores given in Table 1. It is evident that the key factor influencing the segmentation performance is how well the model has learned the boundary regions. Across all cases, the most successful segmentation masks are produced by the model with residual connections. Comparing the results from all figures, the most promising outcomes are achieved for the CC-359 dataset, as shown in Figure 4. Our observation that the model learned better with the CC-359 dataset due to more samples and more slices is also supported by the segmentation results.

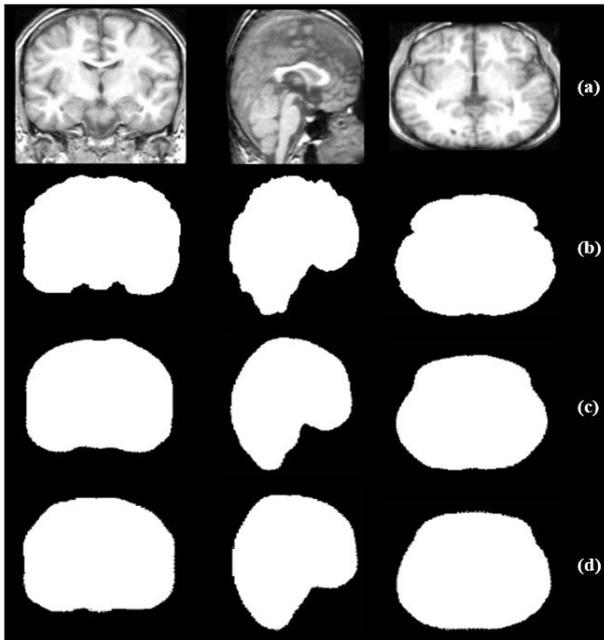


Figure 2. IBSR results, a) original gray, b) binary ground-truth, c) UNet with ResNet binary result, d) UNet with DenseNet binary result

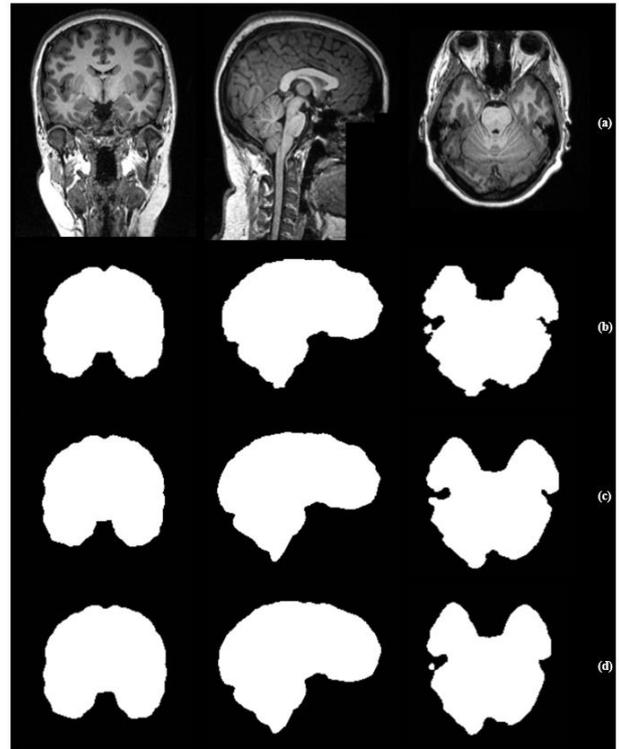


Figure 3. NFBS results, a) original gray, b) binary ground-truth, c) UNet with ResNet binary result, d) UNet with DenseNet binary result

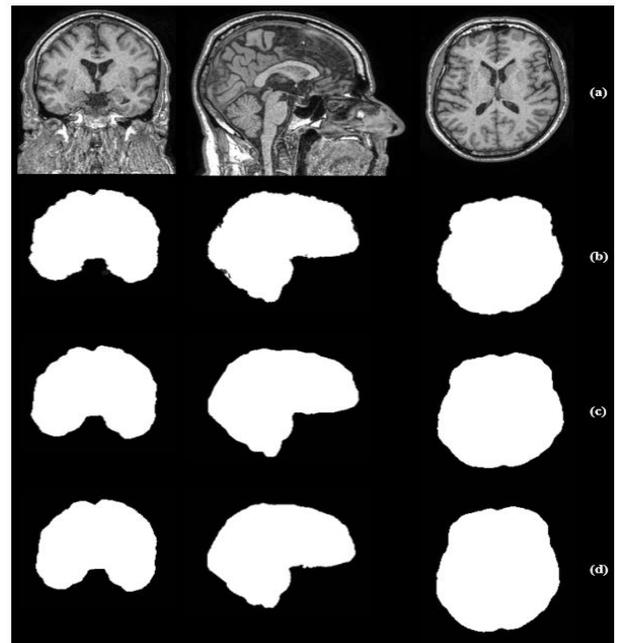


Figure 4. CC-359 results, a) original gray, b) binary ground-truth, c) UNet with ResNet binary result, d) UNet with DenseNet binary result

Table 2 provides a comparison of our model with the studies from the literature which used DL architecture and the same datasets. The comparisons were performed using the results of the 3D-UNet model based on ResNet50, which yielded better results than the DenseNet-based architecture. Kleesiek et al. [14] achieved Dice scores, sensitivity, and specificity values of 0.9632, 0.9501, and 0.9961, respectively, for the IBSR dataset using their proposed CNN model. Hwang et al. [15] compared their 3D-UNet architecture results, obtained for the NFBS dataset, with traditional techniques such as BSE [8], Robust Brain Extraction (ROBEX) [28], and CNN model's results [14], demonstrating superior performance

values. As evident from Table 2, the performance values obtained in this study for the IBSR dataset are competitive with Kleesiek et al.'s [14] results. For the NFBS dataset, the average Dice and sensitivity values are comparable to those of Hwang et al. [15], while our specificity value is higher. Moreover, the Dice values obtained for the NFBS and CC-359 datasets are higher than those achieved by Isensee et al.'s [2] HD-BET method, which reported Dice coefficient values of 0.9820 for NFBS and 0.9690 for CC-359 datasets. Although this study exclusively used NFBS and CC-359 datasets for testing, the model was trained on a large dataset consisting of 6586 samples collected from 25 different institutions.

Table 2. Literature comparison

Study	Model	Dataset	Dice	Sensitivity	Specificity
Kleesiek et al. [14]	CNN	IBSR	0.9632 (± 0.0100)	0.9501 (± 0.0200)	0.9961 (± 0.0030)
Hwang et al. [15]	3D-UNet	NFBS	0.9903 (± 0.0016)	0.9853 (± 0.0040)	0.9953 (± 0.0022)
Isensee et al. [2]	HD-BET	NFBS	0.9820 (± 0.0020)	-	-
	(3D-UNet)	CC-359	0.9690 (± 0.0020)	-	-
This study	ResNet-based	IBSR	0.9613 (± 0.0036)	0.9590 (± 0.0017)	0.9931 (± 0.0010)
	3D-UNet	NFBS	0.9875 (± 0.0018)	0.9840 (± 0.0048)	0.9990 (± 0.0005)
		CC-359	0.9887 (± 0.0028)	0.9899 (± 0.0058)	0.9985 (± 0.0009)

The process of refining brain images by eliminating non-brain regions is crucial for improving the performance of artificial intelligence (AI) algorithms, particularly in medical applications like classification and disease analysis. It offers several advantages: Firstly, it enhances accuracy by ensuring AI algorithms focus exclusively on pertinent brain structures, minimizing diagnostic errors. Secondly, it boosts efficiency by simplifying data, enabling faster processing, which is vital for real-time healthcare applications. Additionally, it promotes better generalization, helping AI models perform well on new cases without being affected by irrelevant data. Its clinical value lies in facilitating precise and timely diagnoses, aiding healthcare professionals in informed decision-making. Lastly, purified brain data advances neuroscience and medical research, allowing for a deeper understanding of neurological conditions. In essence, the purification of brain images by excluding non-brain regions is a critical preprocessing step that enhances AI algorithms' performance in medical fields, benefiting clinical practice and scientific research.

DL-based brain extraction has potential to improve the accuracy and efficiency of clinical applications significantly. It not only saves time but also reduces the risk of human error, making it a valuable tool in modern healthcare for diagnosing, monitoring, and researching neurological conditions and brain function. As DL algorithms continue to evolve, we can expect even greater advancements in the field of brain image analysis such as disease diagnosis and monitoring, functional brain imaging, and treatment planning. Disease diagnosis and monitoring includes such as tumor detection and analysis, stroke assessment, and neurodegenerative diseases. Accurate brain extraction helps in identifying and characterizing brain tumors. It can aid in tumor volume measurement and tracking changes over time, which is crucial for treatment planning and monitoring. In stroke diagnosis, it's important to accurately segment the brain to identify regions affected by ischemia or hemorrhage. This information assists in determining the extent of brain

damage and guiding treatment decisions. DL-based brain extraction can improve the accuracy of identifying brain regions affected by neurodegenerative diseases like Alzheimer's and Parkinson's. It can facilitate early diagnosis and disease progression monitoring. For the treatment planning, in both neurosurgery and radiation therapy for brain tumors, accurate brain segmentation plays a pivotal role. Precise brain segmentation provides surgeons with a clear view of the brain's structure, enabling them to plan the safest and most effective approach for tumor removal or other procedures in neurosurgical interventions. In radiation therapy, knowing the exact boundaries of the brain is essential to prevent damage to healthy tissue, and accurate brain extraction helps define the treatment target.

4. CONCLUSION

This study proposed an improved UNet architecture including residual and dense feature extraction layers to address the challenging problem of extracting brain from surrounding structures, primarily due to pixel intensity similarities. The obtained results demonstrate that the model's performance is particularly dependent on its ability to accurately segment the boundary regions of the brain and other structures. The superior performance of the architecture with residual connections can be attributed to its capability to preserve weak gradient values in deep layers. The findings also emphasize the importance of large-scale training data, as evidenced by the best results obtained for the CC-359 dataset. It highlights the significance of extensive training data for achieving improved performance. Comparing our proposed model to relevant literature studies utilizing the same dataset and UNet-based models, it yields comparable performance results.

In conclusion, for future research in this domain, focusing on architectures and loss function selection that target the distinguishing ability of boundary regions could be a

realistic approach, given its significant impact on the model's performance.

Acknowledgement

This paper is resulted from the Master of Science (MS) thesis study of the second author.

Conflict of Interest Statement

There is no conflict of interest between the authors.

REFERENCES

- [1] Kalavathi P, Prasath VS. Methods on skull stripping of MRI head scan images-a review. *Journal of Digital Imaging*. 2016; 29: 365-379.
- [2] Isensee F, Schell M, Pflueger I, Brugnara G, Bonekamp D, Neuberger U, et al. Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*. 2019; 40(17): 4952-4964, 2019.
- [3] Bhat SY, Naqshbandi A, Abulaish M. Skull stripping on multimodal brain MRI scans using thresholding and morphology. *The Imaging Science Journal*, 2023; 1-13.
- [4] Karakis R, Gurkahraman K, Mitsis GD, Boudrias MH. Deep learning prediction of motor performance in stroke individuals using neuroimaging data. *Journal of Biomedical Informatics*. 2023; 141: article number 104357.
- [5] Smith SM. Fast robust automated brain extraction. *Human Brain Mapping*. 2002; 17: 143-155.
- [6] Souza R, Lucena O, Garrafa J, Gobbi D, Saluzzi M, Appenzeller S, et al. An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement. *NeuroImage*. 2018; 170: 482-494.
- [7] Jenkinson M, Pechaud M, Smith S. BET2 - MR-based estimation of brain, skull and scalp surfaces. *Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB)*, Oxford, 2005.
- [8] Shattuck DW, Sandor-Leahy SR, Schaper KA, Rottenberg DA, Leahy RM. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 2001; 13 (5): 856-876.
- [9] Eskildsen SF, Coupe P, Fonov V, Manjon JV, Leung KK, Guizard N, et al. BEaST: brain extraction based on Non-local segmentation technique. *NeuroImage*. 2012; 59 (3): 2362-2373.
- [10] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M., et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*. 2017; 42: 60-88.
- [11] Yapici M, Karakis R, Gurkahraman K. Improving Brain Tumor Classification with Deep Learning Using Synthetic Data. *Computers, Materials and Continua*. 2023; 74 (3): 5049-5067.
- [12] Gurkahraman K, Karakis R. Brain tumors classification with deep learning using data augmentation. *Journal of the Faculty of Engineering and Architecture of Gazi University*. 2021; 36 (2): 997-1011.
- [13] Rehman HZU, Hwang H, Lee S. Conventional and deep learning methods for skull stripping in brain MRI. *Applied Sciences*. vol. 10, no. 5, article number 1773, 2020.
- [14] Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bendszus M, et al. Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. *NeuroImage*. 2016; 129: 460-469.
- [15] Hwang H, Rehman HZU, Lee S. 3D U-Net for skull stripping in brain MRI. *Applied Sciences*. 2019; 9 (3): article number 569.
- [16] Zhang Q, Wang L, Zong X, Lin W, Li G, Shen D. FRNET: Flattened Residual Network for Infant MRI Skull Stripping. *arXiv 2019; arXiv:1904.05578*.
- [17] Daşgın Ç, Gürkahraman K. Artık Bağlantılar ile Düzenlenen U-Net Mimarisi Kullanarak Beyin Çıkarımı. *International Conference on Applied Engineering and Natural Sciences ICAENS 2023*. Konya, Turkey: 2023. p.348.
- [18] Hoopes A, Mora JS, Dalca AV, Fischl B, Hoffmann M. SynthStrip: Skull-stripping for any brain image. *NeuroImage*. 2022; 260: article number 119474.
- [19] IBSR [Internet]. The Internet Brain Segmentation Repository (IBSR) [cited 2023 July 27]. Available from: <https://www.nitrc.org/projects/ibsr>
- [20] Puccio B, Pooley JP, Pellman JS, Taverna EC, Craddock RC. The preprocessed connectomes project repository of manually corrected skull-stripped T1-weighted anatomical MRI data. *GigaScience*. 2016; 5: article number 45.
- [21] Kingma DP, Welling M. Auto-encoding variational bayes. 2013; *arXiv preprint, arXiv:1312.6114*.
- [22] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*. Munich, Germany: 2015, Proceedings, Part III 18, Springer International Publishing; 2015. p. 234-241.
- [23] Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference*. Athens, Greece: 2016, Proceedings, Part II 19, Springer International Publishing; 2016. p. 424-432.
- [24] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA; 2016. p. 770-778.
- [25] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA; 2017. p. 4700-4708.
- [26] Montagnon E, Cerny M, Cadrin-Chênevert A, Hamilton V, Derennes T, Ilinca A, et al. Deep learning workflow in radiology: a primer. *Insights into Imaging*. 2020; 11: 1-15.
- [27] Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis:

diagnosis to detection. *Multimedia Tools and Applications*. 2023; 82(17): 26731-26769.

- [28] Iglesias JE, Liu CY, Thompson PM, Tu Z. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Transactions on Medical Imaging*, 2011; 30 (9): 1617-1634.