Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı 40, S. 75-81, Eylül 2022

Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
Special Issue 40, pp. 75-81, September 2022

Research Article

Automatic Skull Stripping and Brain Segmentation with U-Net in MRI Database

Alperen Derin[1], Ahmet Furkan Bayram[2], Caglar Gurkan[3,*], Abdulkadir Budak[4] and Hakan Karatas[5]

[1] Department of Computer Engineering, Faculty of Computer and Information Science, Sakarya University, Sakarya (ORCID: 0000-0002-2276-0591), alperenderinn@gmail.com
[2] Department of Computer Engineering, Faculty of Engineering, Karadeniz Technical University, Trabzon (ORCID: 0000-0002-1304-9941), bayramahmet48@gmail.com
[3*] Department of Electrical and Electronics Engineering, Graduate School of Science, Eskisehir Technical University, Eskisehir, (ORCID: 0000-0002-4652-3363), caglargurkan@eskisehir.edu.tr
[3*] Department of Artificial Intelligence and Image Processing, Akgun Computer Inc., Ankara (ORCID: 0000-0002-4652-3363), caglar.gurkan@akgun.com.tr
[4] Department of Artificial Intelligence and Image Processing, Akgun Computer Inc., Ankara (ORCID: 0000-0002-0328-6783), kadir.budak@akgun.com.tr
[5] Department of Artificial Intelligence and Image Processing, Akgun Computer Inc., Ankara (ORCID: 0000-0002-9497-5444), hakan.karatas@akgun.com.tr

(1st International Conference on Innovative Academic Studies ICIAS 2022, September 10-13, 2022)

(**DOI:** 10.31590/ejosat.1173065)

ATIF/REFERENCE: Derin, A., Bayram, A.F., Gurkan, C., Budak, A. & Karatas, H. (2022). Automatic Skull Stripping and Brain Segmentation with U-Net in MRI Database. *European Journal of Science and Technology*, (40), 75-81.

Abstract

Skull stripping has an important in neuroimaging workflow. Skull stripping is a time-consuming process in the Magnetic resonance imaging (MRI). For this reason, skull stripping and brain segmentation are aimed in this study. For the this purpose, the U-NET architecture design, which is one of the frequently used models in the field of medical image segmentation, was used. Also, different loss functions such as Cross Entropy (CE), Dice, IoU, Tversky, Focal Tversky and their compound forms were tested on U-Net architecture design. The compound loss function of CE and Dice loss functions achieved the best performace with the average dice score of 0.976, average IoU score of 0.964, sensitivity of 0.972, specificity of 0.985, precision of 0.960 and accuracy of 0.981. As a result, skull stripping was performed to facilitate the detection of brain diseases.

Keywords: Brain, Skull stripping, MRI, Segmentation, U-Net.

MRG Veri Tabanında U-Net ile Otomatik Kafatası Çıkartma ve Beyin Segmentasyonu

Öz

Kafatasının çıkartılması beyin görüntüleme iş akışında önemli bir yere sahiptir. Kafatasının çıkartılması, Manyetik Rezonans Görüntülemede (MRG) zaman alan bir işlemdir. Bu nedenle bu çalışmada kafatası çıkartma ve beyin segmentasyonu amaçlanmaktadır. Bu amaçla tıbbi görüntü segmentasyonu alanında sıklıkla kullanılan modellerden biri olan U-Net mimari tasarımı kullanılmıştır. Ayrıca Cross Entropy (CE), Dice, IoU, Tversky, Focal Tversky gibi farklı kayıp fonksiyonları ve bunların bileşik formları U-Net mimari tasarımı üzerinde test edilmiştir. CE ve Dice kayıp fonksiyonlarının bileşik kayıp fonksiyonu, 0.976 ortalama dice skoru, 0.964 ortalama IoU skoru, 0.972 sensivity, 0.985 specificity, 0.960 presicion ve 0.981 accuracy ile en iyi performansı elde etmiştir. Sonuç olarak, beyin hastalıklarının tespitini kolaylaştırmak için kafatasının çıkartılması işlemi yapılmıştır.

Anahtar Kelimeler: Beyin, Kafatasının çıkartılması, MRG, Segmentasyon, U-Net.

[*] Corresponding Author: caglar.gurkan@akgun.com.tr

1. Introduction

Clinical imaging systems are used by healthcare professionals for the detection of diseases or injuries. These systems are essential to ensure the correct response to disease or injury situations, to better detect the disease or injury, and to provide the proper treatment for the patient. Many medical imaging systems, i.e. Magnetic Resonance Imaging (MRI), Computed Tomography (CT), X-Ray, Ultrasound, Bone-Scan and Nuclear Imagination are used for clinical diagnosis. Among these techniques, MRI, CT and X-Ray are frequently used for disease detection. The MRI, unlike the CT and X-Ray imaging techniques, does not use radiation waves. MRI is an imaging method that has a considerable in medical imaging, which creating a layered output by scanning cross-sectional areas of body tissues with the help of magnetic waves (X-Rays, CT Scans and MRIs - OrthoInfo - AAOS, 2017).

Brain diseases such as Dementia, Brain Cancer, Trauma-Induced Epilepsy, Parkinson's, and Stroke, as well as many other brain damage caused by external effects, can be detected even in the early stages thanks to MRI technique. Early detection of these diseases is essential for the treatment of the disease. Furthermore, early diagnosis is critical issue day by day because of the fact that the brain is a vital organ in the human body. The use of artificial intelligence (AI) is vital in the health sector. AI facilitates the work of healthcare professionals. Also, AI provides the diagnosis of diseases in the early stages. Since diseases are not very obvious in the early stages, AI based models that provide more accurate should be developed. When studies are conducted on brain diseases, the skull is also included in the brain MR images. In this regard, it is necessary to segment the skull and brain regions in order to develop a more accurate AI-based model with the purpose of the analysis of brain diseases. However, there is still no solid solution to this problem (Kalavathi & Prasath, 2016).

In this paper, it is aimed to perform the skull stripping. This process is necessary before the detect brain diseases or damages. In this task, the Brain Tumor Progression data set was used for skull stripping. Considering the bones in the images, the brain region was labelled within the scope of this study. The data set was used in the training of the U-Net segmentation network. At this stage, different loss functions such as Cross Entropy (CE), Dice, IoU, Tversky, Focal Tversky and their compound forms were tested. As a result, the best predictive performance was obtained by the compound loss function of CE and Dice loss functions.

The rest of this paper is organised as: Section 2 presents literature survey of used several segmentation networks for brain skull stripping. Section 3 presents the utilized methodologies. Section 4 presents the results obtained by the U-Net architecture design using the several loss functions. Section 5 presents concluding remarks.

2. Related Works

It has been suggested by many scientific researchers that several segmentation techniques can be used for skull stripping.

Hwang et al. (Hwang et al., 2019) aimed to perform skull stripping to analyze brain images using MRI images. 3D U-NET segmentation network was used in the study. The used data set consists of MRI images of 125 patients. Segmentation performance was analyzed by Dice coefficient (DSC). Mean DSC score of 0.9903, sensitivity of 0.9853 and specificity of 0.9953 were achieved in the skull stripping.

Qamar et al. (Qamar et al., 2020) aimed to segment the brain of infants into tissues, such as cerebrospinal fluid (CSF), white matter (WM), and gray matter (GM) by using MRI images. In this study, a novel segmentation network, which is basic variant of 3D U-NET was proposed. The used data set consists of 144 sagittal images. Segmentation performance was analyzed by DSC. In this proposed approach, the dice scores of 0.95, 0.905 and 0.92 were achieved fot CSF, WM and GM tissues, respectively.

Wang et al. (Wang et al., 2021) aimed to perform brain segmentation using MRI images. In the study, U-Net architecture design was used with/without transfer learning method. The U-Net architecture design supported by transfer learning method achieved better performance. The model achieved the DSC score of 0.9916 in skull stripping.

Kleesiek et al. (Kleesiek et al., 2016) aimed to perform brain segmentation using MR images, other than the T1-weighted MR images. Researchers used the 3D convolutional neureal network based deep learning architecture design. In this approach, architecture design considers any modality, including contrast scans. In this study, model achieved the DSC score of 95.19%.

Li et al. (Li et al., 2020) aimed to extraction the fetal brain using MRI images. In the study, researchers proposed a 2-step method to solve the automatic fetal brain extraction. In this method, two fully convolutional networks (FCN) were used. The first FCN removed the fetal's associated area of the brain (ROI). The second FCN produced the mask of the brain. In this proposed method, the location of the brain was found with 100% accuracy. In brain segmentation, 0.958 average DSC score, 0.95 sensitivity and 0.968 precision were achieved.

3. Methodology

The methodologies of the study are presented under the headings of experimental setup and performance evaluation metrics.

3.1. Experimental Setup

The Brain Tumor Progression data set, which is publicly available and includes data from a total of 20 patients MR images was used. The data set consists of T1, T2 and proton density (PD) weighted images. T1 and T2 weighted images are generally used in segmentation and detection problems. In this context, 329 T1-weighted MR images were selected, and the T1-weighted images were labelled for the purpose of skull stripping. Subsequently, images in data set were normalized by using min-max normalization technique. %90 of the images were used in the training of model, while %10 was used in testing of the trained model.

In this study, U-Net (Weng & Zhu, 2021) which is a state-of-the-art segmentation network was tested with several well-known loss functions, such as cross entropy (CE), Dice, IoU, Tversky, Focal Tversky and their compound form. Alpha, beta and gamma are the coefficient of the false negative, the coefficient of the false positive and the exponential coefficient

of the Tversky loss function, respectively. The alpha and beta values were set as 0.7 and 0.3 for Tversky loss function, respectively. The alpha, beta and gamma values were set as 0.7, 0.3 and 0.75 for Focal Tversky loss function, respectively. In the training phase for U-Net; optimizer, batch size, and epoch value were chosen as, the RMSprop, 1, and 50, respectively. The initial learning rate was chosen as 1e-4. If performance of segmentation models is not improvement for along the 15 epochs, the learning rate was multiplied by 0.1. PyTorch framework in Python programming language on Spyder Integrated Development Environment (IDE) was used in the experimental analysis procedure. NVIDIA GeForce RTX 3060 graphics card was used in training and testing procedures.

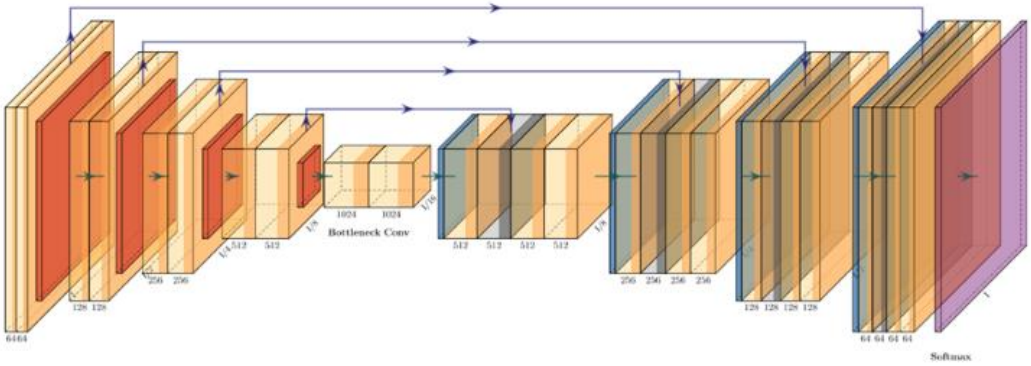

Figure 1. U-Net architecture design (HarisIqbal88/PlotNeuralNet: Latex Code for Making Neural Networks Diagrams, n.d.)

3.2. Performance Evaluation Metrics

The performance of segmentation models was examined by using accuracy, precision, sensitivity, specificity, dice, and IoU performance evaluation metrics. Equations 1, 2, 3, 4, 5, and 6 indicates formulations of performance evaluation metrics.

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

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

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

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (4)$$

$$\text{Dice} = \frac{2*TP}{2*TP+FP+FN} \quad (5)$$

$$\text{IoU} = \frac{TP}{TP+FP+FN} \quad (6)$$

True Positive (TP) indicates to number of correctly classified positive class based on pixels. True Negative (TN) indicates to number of correctly classified negative class based on pixels. False Positive (FP) indicates to incorrectly classified positive class based on pixels. False Negative (FN) indicates to number of the incorrectly classified negative class based on pixels.

4. Results and Discussion

Table 1 shows the results obtained by testing several loss functions on the U-Net model. In addition to the results of the brain region, background results are also provided. Because the brain area is not small compared to the background. Therefore, it is important to take into account the brain and background results.

CE loss function achieved the dice score of 0.985 for background, dice score of 0.960 for brain, and average dice score of 0.972 for background and brain. CE loss function achieved the IoU score of 0.970 for background, IoU score of 0.922 for brain, and average IoU score of 0.946 for background and brain. CE loss function achieved the sensitivity of 0.960, specificity of 0.984, precision of 0.959, and accuracy of 0.978.

Dice loss function achieved the dice score of 0.986 for background, dice score of 0.964 for brain, and average dice score of 0.975 for background and brain. Dice loss function achieved the IoU score of 0.973 for background, IoU score of 0.930 for brain, and average IoU score of 0.951 for background and brain. Dice loss function achieved the sensitivity of 0.968, specificity of 0.984, precision of 0.959, and accuracy of 0.980.

IoU loss function achieved the dice score of 0.985 for background, dice score of 0.962 for brain, and average dice score of 0.974 for background and brain. IoU loss function achieved the IoU score of 0.971 for background, IoU score of 0.926 for brain, and average IoU score of 0.949 for background and brain. IoU loss function achieved the sensitivity of 0.975, specificity of 0.980, precision of 0.949, and accuracy of 0.979.

Tversky loss function achieved the dice score of 0.987 for background, dice score of 0.965 for brain, and average dice score of 0.976 for background and brain. Tversky loss function achieved the IoU score of 0.973 for background, IoU score of 0.933 for brain, and average IoU score of 0.953 for background and brain. Tversky loss function achieved the sensitivity of 0.974, specificity of 0.983, precision of 0.956, and accuracy of 0.981.

Focal Tversky loss function achieved the dice score of 0.986 for background, dice score of 0.962 for brain, and average dice score of 0.974 for background and brain. Focal Tversky loss function achieved the IoU score of 0.972 for background, IoU score of 0.928 for brain, and average IoU score of 0.950 for background and brain. Focal Tversky loss function achieved the sensitivity of 0.961, specificity of 0.986, precision of 0.964, and accuracy of 0.979. The compound loss function of CE and Dice loss functions achieved the dice score of 0.987 for background, dice score of 0.966 for brain, and average dice score of 0.976 for background and brain. The

compound loss function of CE and Dice loss functions achieved the IoU score of 0.974 for background, IoU score of 0.934 for brain, and average IoU score of 0.954 for background and brain. The compound loss function of CE and Dice loss functions achieved the sensitivity of 0.972, specificity of 0.985, precision of 0.960, and accuracy of 0.981.

The compound loss function of CE and IoU loss functions achieved the dice score of 0.983 for background, dice score of 0.957 for brain, and average dice score of 0.970 for background and brain. The compound loss function of CE and IoU loss functions achieved the IoU score of 0.967 for background, IoU score of 0.918 for brain, and average IoU score of 0.943 for background and brain. The compound loss function of CE and IoU loss functions achieved the sensitivity of 0.976, specificity of 0.976, precision of 0.939, and accuracy of 0.976.

The compound loss function of CE and Tversky loss functions achieved the dice score of 0.985 for background, dice score of 0.960 for brain, and average dice score of 0.972 for background and brain. The compound loss function of CE and Tversky loss functions achieved the IoU score of 0.970 for background, IoU score of 0.923 for brain, and average IoU score of 0.946 for background and brain. The compound loss function of CE and Tversky loss functions achieved the sensitivity of 0.966, specificity of 0.982, precision of 0.954, and accuracy of 0.978.

The compound loss function of CE and Focal Tversky loss functions achieved the dice score of 0.985 for background, dice score of 0.960 for brain, and average dice score of 0.972 for background and brain. The compound loss function of CE and Focal Tversky loss functions achieved the IoU score of 0.970 for background, IoU score of 0.922 for brain, and average IoU score of 0.946 for background and brain. The compound loss function of CE and Focal Tversky loss functions achieved the sensitivity of 0.966, specificity of 0.982, precision of 0.953, and accuracy of 0.978.

The results obtained by the compound loss function of CE and Dice loss functions are given in Figure 2 and Figure 3.

Table 1. Results obtained by testing several loss functions on the U-Net model

Loss Functions	Dice Background	Dice Brain	Dice Average	IoU Background	IoU Brain	IoU Average	Sensitivity	Specificity	Precision	Accuracy
CE	0.985	0.960	0.972	0.970	0.922	0.946	0.960	0.984	0.959	0.978
Dice	0.986	0.964	0.975	0.973	0.930	0.951	0.968	0.984	0.959	0.980
IoU	0.985	0.962	0.974	0.971	0.926	0.949	0.975	0.980	0.949	0.979
Tversky	0.987	0.965	0.976	0.973	0.933	0.953	0.974	0.983	0.956	0.981
Focal Tversky	0.986	0.962	0.974	0.972	0.928	0.950	0.961	0.986	0.964	0.979
CE+Dice	0.987	0.966	0.976	0.974	0.934	0.954	0.972	0.985	0.960	0.981
CE+IoU	0.983	0.957	0.970	0.967	0.918	0.943	0.976	0.976	0.939	0.976
CE+Tversky	0.985	0.960	0.972	0.970	0.923	0.946	0.966	0.982	0.954	0.978
CE+Focal Tversky	0.985	0.960	0.972	0.970	0.922	0.946	0.966	0.982	0.953	0.978

In general analysis, the best dice score for background was obtained by the Tversky loss function and the compound loss function of CE and Dice loss functions. The best dice score for brain value was obtained by the compound loss function of CE and Dice loss functions. The best IoU score for background value was obtained by the compound loss function of CE and Dice loss functions. The best dice IoU for brain value was obtained by the Tversky loss function and the compound loss function of CE and Dice loss functions. The best sensitivity value was obtained by the compound loss function of CE and IoU loss functions. The best specificity value was obtained by the Focal Tversky loss function. The best precision value was obtained by the compound loss function of CE and Dice loss functions. The best accuracy value was obtained by the Tversky loss function and the compound loss function of CE and Dice loss functions. The Tversky loss function outperformed the Dice loss function. On the contrary, Dice loss function outperformed the Focal Tversky loss function. For this reason, different alpha, beta and gamma values should be tested for the Tversky and Focal Tversky loss functions.

In comparative analysis, while Tversky loss function achieved the best performance among the lean loss functions, CE loss function achieved the worst performance. While the compound loss function of CE and Dice loss functions achieved the best performance among the compound loss functions, the compound loss function of CE and IoU loss functions has the worst performance. When the lean and compound loss functions are considered together, while the compound loss function of CE and Dice loss functions has the best performance, the compound loss function of CE and IoU loss functions has the worst performance. While the compound form of the Dice loss function with the CE loss function achieved the better predictive performance, compound forms of other loss functions did not better predictive performance.

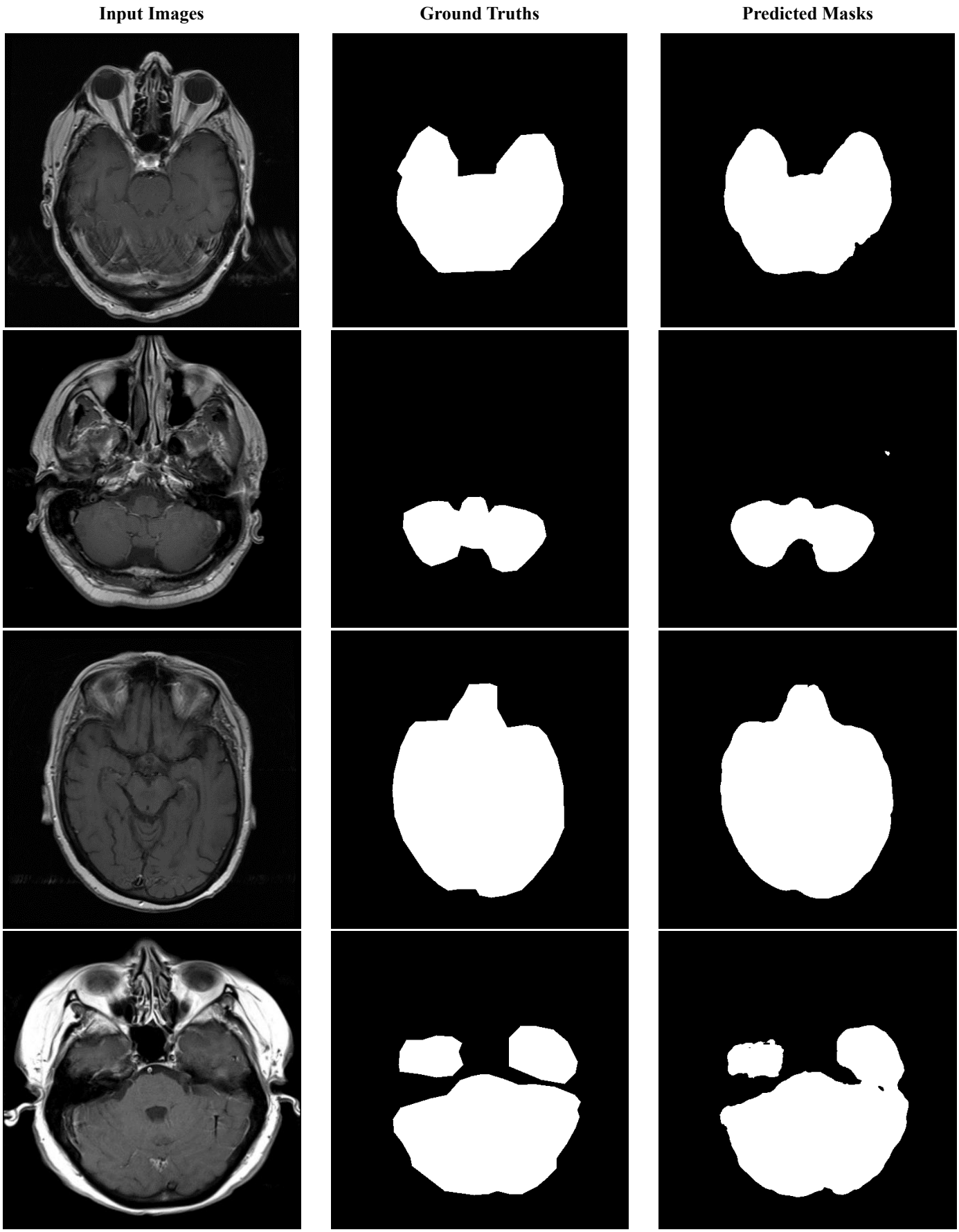

Figure 2. Results obtained by compound loss function of CE and Dice loss functions

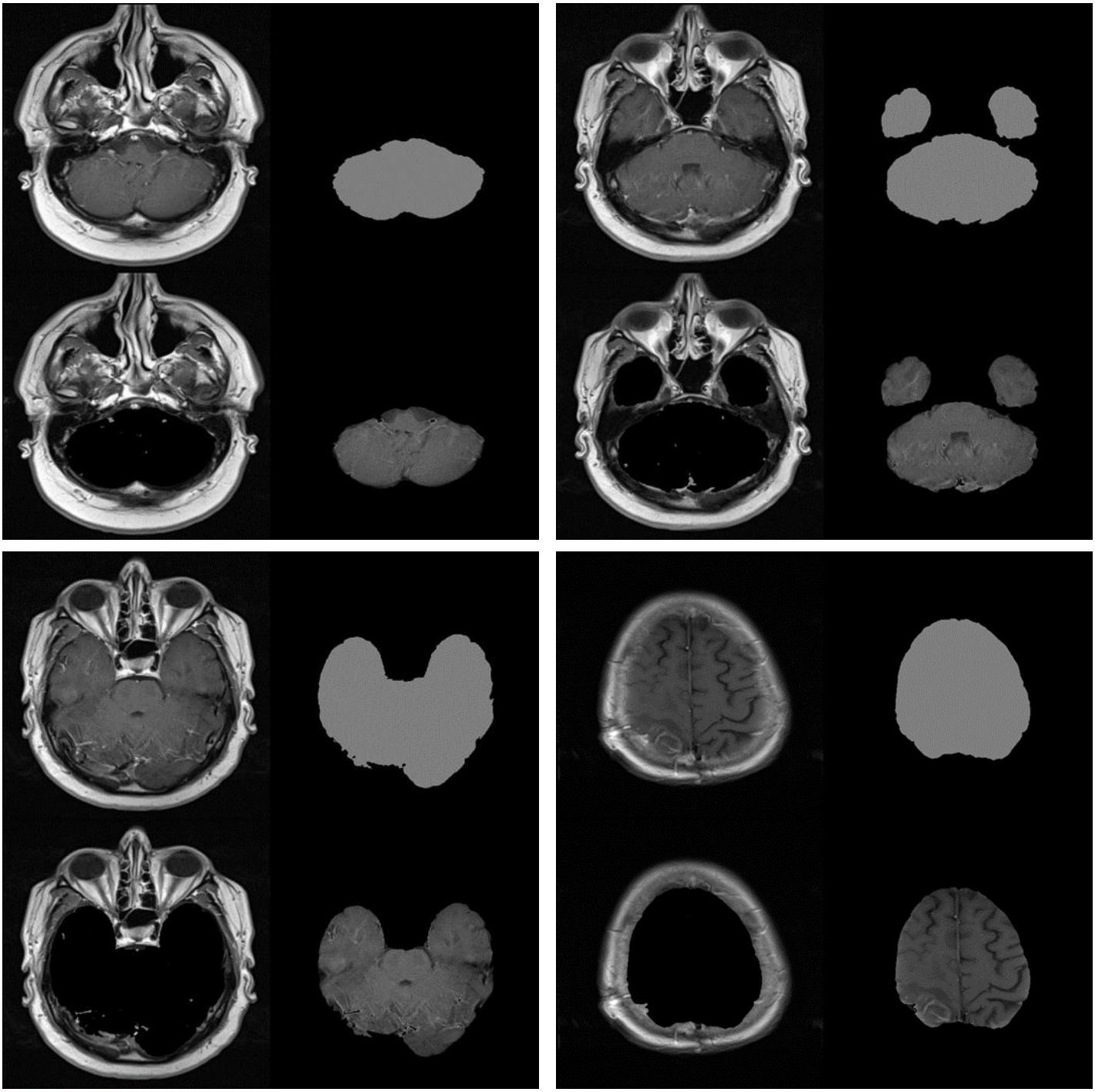

Figure 3. Results after skull stripping process

4. Conclusion

In this study, the U-Net architecture design was used for skull stripping. The U-Net was tested with several well-known loss functions, such as CE, Dice, IoU, Tversky, Focal Tversky and their compound form. The U-Net by using with together the different loss functions, a suggestion was presented to other researchers about segmentation performances in the skull stripping process. In comparative analysis, the average dice score of 0.976, average IoU score of 0.964, sensitivity of 0.972, specificity of 0.985, precision of 0.960 and accuracy of 0.981 were achieved by the compound loss function of CE and Dice loss functions. As a result, the workload of radiologists can be reduced by easily separating the relevant region (ROI) thanks to the automatic skull stripping and brain segmentation system. In future studies, different alpha, beta and gamma values will be tested for Tversky and Focal Tversky loss functions. The size of the data set will be enlarged. Other segmentation networks will be tested. Lastly, we will deal with brain tumor segmentation using brain MR images.

5. Acknowledge

This paper has been prepared by AKGUN Computer Incorporated Company. We would like to thank AKGUN Computer Inc. for providing all kinds of opportunities and funds for the execution of this project.

References

X-rays, CT Scans and MRIs - OrthoInfo - AAOS (pp. 1–4). (2017). https://orthoinfo.aaos.org/en/treatment/x-rays-ct-scans-and-mris/

Kalavathi, P., & Prasath, V. B. S. (2016). Methods on Skull Stripping of MRI Head Scan Images—a Review. In Journal of Digital Imaging (Vol. 29, Issue 3, pp. 365–379). Springer. https://doi.org/10.1007/s10278-015-9847-8

Hwang, H., Ur Rehman, H. Z., & Lee, S. (2019). 3D U-Net for skull stripping in brain MRI. Applied Sciences (Switzerland), 9(3), 569. https://doi.org/10.3390/app9030569

Qamar, S., Jin, H., Zheng, R., Ahmad, P., & Usama, M. (2020). A variant form of 3D-UNet for infant brain segmentation. Future Generation Computer Systems, 108, 613–623. https://doi.org/10.1016/j.future.2019.11.021

Wang, X., Li, X. H., Cho, J. W., Russ, B. E., Rajamani, N., Omelchenko, A., Ai, L., Korchmaros, A., Sawiak, S., Benn, R. A., Garcia-Saldivar, P., Wang, Z., Kalin, N. H., Schroeder, C. E., Craddock, R. C., Fox, A. S., Evans, A. C., Messinger, A., Milham, M. P., & Xu, T. (2021). U-net model for brain extraction: Trained on humans for transfer to non-human primates. NeuroImage, 235, 118001. https://doi.org/10.1016/j.neuroimage.2021.118001

Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., & Biller, A. (2016). Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. NeuroImage, 129, 460–469. https://doi.org/10.1016/j.neuroimage.2016.01.024

Li, J., Luo, Y., Shi, L., Zhang, X., Li, M., Zhang, B., & Wang, D. (2020). Automatic fetal brain extraction from 2D in utero fetal MRI slices using deep neural network. Neurocomputing, 378, 335–349. https://doi.org/10.1016/j.neucom.2019.10.032

Weng, W., & Zhu, X. (2021). INet: Convolutional Networks for Biomedical Image Segmentation. IEEE Access, 9, 16591–16603. https://doi.org/10.1109/ACCESS.2021.3053408

HarisIqbal88/PlotNeuralNet: Latex code for making neural networks diagrams. (n.d.). Retrieved September 5, 2022, from https://github.com/HarisIqbal88/PlotNeuralNet