International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Automatic Skull Stripping from MRI of Human Brain using Deep Learning Framework for the Diagnosis of Brain Related Diseases

C. Balaji[1*], Dr. S. Veni[2]

Submitted: 08/05/2023 **Revised**: 15/07/2023 **Accepted**: 06/08/2023

Abstract: In the technological world, health care systems play essential role in extending the human age through early detection. The grooming of Artificial Intelligence technology and deep learning enlighten many smart surgical devices and aid knowledge in disease diagnosis and planning. In particular, the brain disease diagnosis process embraces several devices, such as Magnetic Resonance (MR) machines etc. and software. Most of the brain disease diagnosis includes brain extraction (skull stripping) as the pre-processing step. Brain portion extraction from its non-brain tissues from MR image is a tedious process and takes 40 to 60 minutes per patient in manual process and the earlier model lacked accuracy. In the age of deep learning, accurate brain tissue detection using deep learning is essential and accounts for good results. This paper proposes a modified Unit for skull stripping from 2D MR images and comparing it with other deformable and deep learning models such as BET2, RoBEX and UNet3D. The proposed model provides good competency results, such as a 98 % Dice score and 1% higher than the UNet3D model**.**

Keywords**:** *skull striping, deep learning, CNN, brain, skull.*

1. Introduction

A patient is placed into a tunnel that has a magnetic field inside of it in order to obtain an MR image of them. As a result, the quantum spins of all the protons in the body 'align' to the same value. Then, to prevent this alignment, an oscillating magnetic field pulse is applied. An electromagnetic wave is emitted by the protons as they reach equilibrium. Different images will be produced depending on the amount of fat there the chemical makeup, and most critically the sort of stimulation (i.e., sequences) utilised to disrupt the protons. T1 weighted, the contrast enhanced T1 (T1C), T2, and Fluid Attenuation Inversion Recovery is the four commonly acquired sequences (FLAIR).

Skull removal known as skull stripping is a must for further procedures including volume estimation and disease pathology diagnosis. Experts must complete the time-consuming operation of creating gold standard data. To create a deep model, the majority of skull stripping data sets are quite modest in size. The image must go through a number of pre-processing stages that call for domain-specific information before being sent to the model. It's nice to build a model and have decent precision. A very good model, however, typically performs very poorly with data from the real world. When the model observes entirely different data than what it was trained on, this results from data drift. It might occur because of distinctions. Hence, several researches have been conducted on the skull stripping and segmentation work.

Several earlier and current researches have described various skull removal models. [1] Classified the methods into four groups: morphology, intensity, deformable surface- and at last based methods. The job of skull removal from brain MR images utilised morphological erosion and dilation operations that were suggested by [2]. A new edge detection operator called the Marr-Hildreh operator was employed along with the morphological operators to extract brain images from the skull [3]. The morphological methods are very sensitive to small intensity variations. Hence, the methods have challenges in finding proper thresholding and edge detection.

A deformable model was involved in the process of brain and non-brain tissue separation [4]. On those days, BET and BET 2 were utilised for T1 and T2 weighted MR brain images, which generated better results [5-7] used atlas-based methods. Recently, a robust skull stripping method has proposed by [8] that used a multi-contrast brain stripping method which uses sparse patch. The methods depend on image modalities and image qualities. Atlas-based methods suffer from multi parametric MR images. In pathological images with diseases or brain tumours, the techniques are not appropriate.

[1*]*Research Scholar, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, India*
[1*]*Email: balajicj80@gmail.com*
[2]*Professor& Head, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, India*
[2]*Email: venics@kahedu.edu.in*
[1*] *Corresponding Author Email: balajicj80@gmail.com*

Deep learning has become common in several fields over the past year as a result of advancements in computer hardware and the accessibility of big data., like image analysis, computer vision [9], and natural language processing. Subsequently, deep learning methods can be applied in human organ related image analysis, which includes skull stripping [10-11], brain segmentation [12], brain tumour classification [13]. Thus, the proposed method deals with the brain extraction process using a deep learning model. It utilises the U-Net model and has produced some changes in the algorithm and layer utilisation. The remaining part of the paper deals with the data set description, proposed method, result, and discussion

2. Data Set

Data from 125 people, aged 21 to 45, exhibiting a range of clinical and subclinical brain disorders are available in the NFBS collection. Each volume contains the raw images, images of the skull sripped, and images of the brain mask shown in Fig. 1. The repository has structural T1-weighted, single-channel MR images for every volume. The ground truth, or brain mask, is the brain's equivalent of an image mask. It is obtained by applying manual edits by subject-matter experts to remove non-brain tissue using the Beast (brain extraction based on nonlocal segmentation) approach. Fig. 1 shows three images: the original raw image, the skull stripped image, and a mask respectively. The first image in Fig. 1 is the original raw image, the second image is a stripped-down skull scan that only reveals the brain, and the third image is a mask with computer and human annotations. Fig 1 shows that Sample images from NFBS dataset. a) brain image with skull, b) brain extracted from its surrounding and c) brain mask.

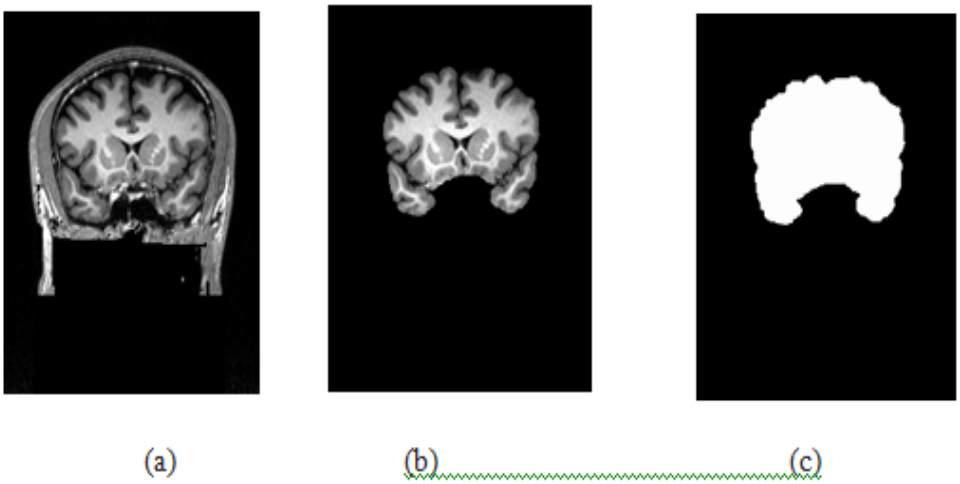


Fig 1: Sample images from NFBS dataset. a) brain image with skull, b) brain extracted from its surrounding and c) brain mask

3. Proposed Method

The presented method includes two stages, such as pre-processing and skull removal (brain extraction). The detailed structural design of this work is described in Fig. 2. The pre-processing step involves strengthening the edges around the brain. Further, the brain extraction process is done with the U-Net model to finely label the brain tissues. Fig 2 shows that Architecture of the proposed method

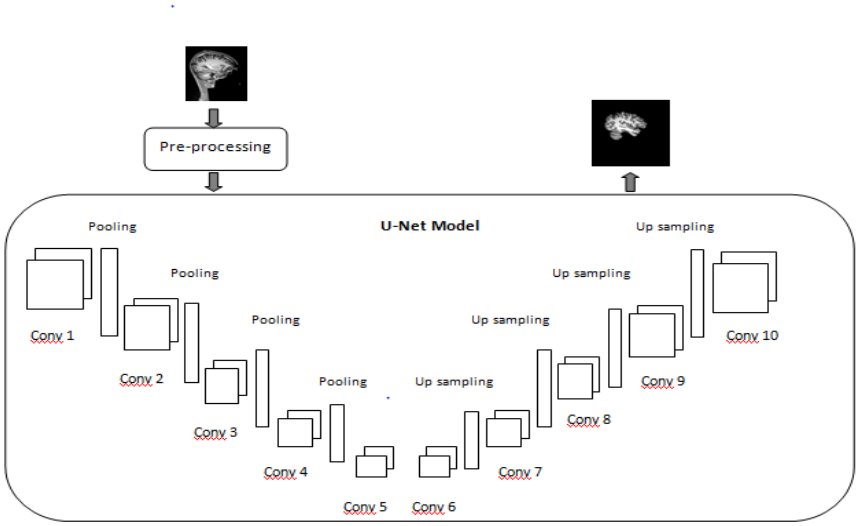


Fig 2: Architecture of the proposed method

Stage 1: Pre-processing

In MR images, the skull appears dark in T1-w images [14]. However, it accounts for high differences with its neighbouring pixels, as demonstrated in Fig. 3. In Fig. 3, column 1 explains the original MR images, column 2 be evidence for the differences among its neighbours, and column 3 shows the brain portion. The image provided in row 1 is only the skull and meningeal tissues and is taken at the end of the image volume. The image provided in row 2 contains very few brain tissues. The intensity differences among its neighbours are high and given in column 2, which is a binary image. Hence, the presented method takes derivatives of a pixel from each of its eight neighbours. Compares the differences and increases the pixel intensity values that have the largest differences. Thus, strengthening the boundary of the brain portion. Further, this method uses an enhanced non-local means filter to smooth the image. This filter uses a dynamic window to search for similar pattern pixels and ensure window size by using mean absolute deviation error. The window size is enlarged based on the error measure. The method provides contrast enhanced images and retains brain structure [15].

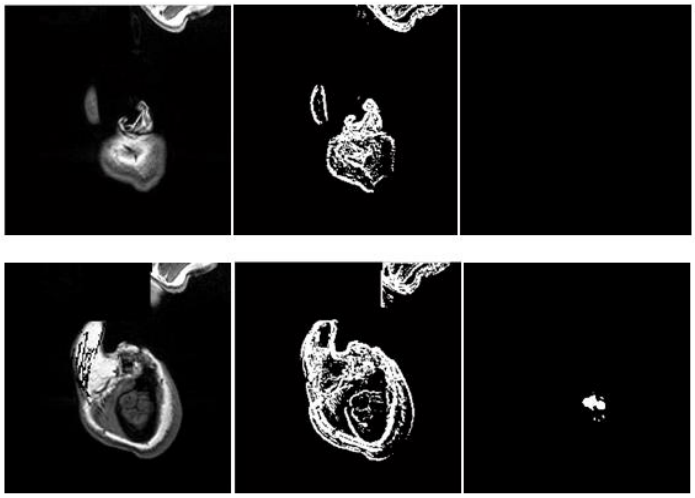


Fig 3: Pixel intensity in MR images. Column 1 shows the scanned images; column 2 shows the intensity difference of skull with neighbouring pixels and column 3 shows the brain mask

Stage 2: Brain Extraction

3.1 Convolution Neural Network

In this network, the fully connected layer employing the soft max function is often coupled with the convolution and subsampling layers. A set of many convolutional layers extract characteristics with increasing fineness at each layer as they move from the architecture's input to output layers. The completely linked layers that do classification are placed after the convolution layers. A subsampling or pooling layer often comes after every

convolution layer. The kernels or filters, which are collections of 2D neurons, make up each layer. When compared to other neural networks, a 2D, n-by-n-pixel picture only receives by CNN as input. The neurons gathered in the feature extraction layer of the network are not interconnected with the neurons in the neighboring levels. Instead, they are solely related to the fixed-size spatially mapped neurons; the layer above has partially overlapping neurons. The Local Receptive Field is the name given to the input region. (LRF). Due to the fewer connections, over-fitting is less likely and training time is reduced. Every neuron in a filter must be attached to the same quantity of neurons in the preceding input layer and have the same weights and biases (or feature map). These components facilitate learning and trim down the memory usages of the network. As a result, the neuron in a certain filter looks for the same pattern in several places throughout the provided image. The sub-sampling layers reduce the size of the network. Using the maximum or mean pooling, the sub-sampling layer lowers the size of the neurons. Fig 3 shows that Pixel intensity in MR images. Column 1 shows the scanned images; column 2 shows the intensity difference of skull with neighboring pixels and column 3 shows the brain mask

U-Net architecture is familiar with therapeutic image processing. The architecture of this model is very simple and has skip connections. It works with the encoder and feature map to help redirect the flow and improve training. The net model has a narrowing path and an immediate getting bigger path. Skip connection preserves the loss from the previous layers and strengthens the net. The final convolution layer has a single filter to confine the pixel class. The pooling process reduces the size of the image, which employs Max pooling to reduce the image size since the skull area is represented by a high intensity value.

The U-Net model consists of several layers, which are organised in the order of input, encoder, decoder and an output layer. In Fig. 2, each convolution block holds a pair of convolutional layers confined by the ReLu activation function. The left-side layers reduce the size of the image and are called the encoder block. It has five pairs of convolutional layers. The first pair of layers has 32 filters, which are 3 by 3. The following pooling process reduces the three by three pixels by one pixel and returns the maximum value among the nine pixels. The convolution layer 2 holds 64 filters in 3 by 3 sizes. Following that, layers 3, 4, and 5 contain 128, 256, and 512 filters, respectively. The right-side convolution blocks act as decoders and increase the image size; they hold the same order of filters in each convolution layer. The output has three labels that represent three classes, such as background, brain, and skull. The model classifies each pixel into the provided classes.

4. Results and Discussion

This section first shares the details about the data set and parameters. Second, reveal the quantitative measures, and then investigate the performance of the proposed model and computers with other methods. The data set was taken from the NFBS website, which contains the images as raw images jinni format. T1-weighted images with the skull stripped and the manually adjusted gold standard constitute each subject's data. Each voxel is 1 × 1× 1 mm^3 in size, and each scan is 256 × 256 × 192 in size. The first two dimensions 256 × 256 represent the size of a slice, and the third dimension 192 accounts for total number of slices in the3D dataset taken for the experiment. Micro software is employed to convert the 3D images into 2D slices of 1 × 1 mm pixel thickness in .tif compression format.

In the coding experiments, the training data frame contains 1500 images that are randomly selected from the 125 volumes of images. As a result, the training data frame was 1500 images long, the test data frame was 600 images long, and validation required 500 images from randomly selected volumes. During the training process, the average height and width of the image are selected to fix the image size. Because large images necessitate more layers and channels, In the 1500 training images, 500 images are mid-slice ranges from 92 to 98; 500 images are end slides, which only contain the skull without the brain portion; 500 images are slice ranges from 50 to 60; and 170 to 160 slices, which contain a very small number of brain tissues compared to non-brain tissues. The Adam optimizer is widely used for image processing and was used here with 0.001 as initial learning rate. The proposed use is implemented using Google Colab.

4.1 Evaluation Parameters

The quantitative validation only ensures the quality of work in machine vision. This work evaluated with the measures dice, precision, recall and accuracy.

$$Dice = \frac{TP}{2TP+FN+FP} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$Accuracy = \frac{TP+FP}{TP+TN+FP+FN} \tag{4}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{precision + Recall} \quad (5)$$

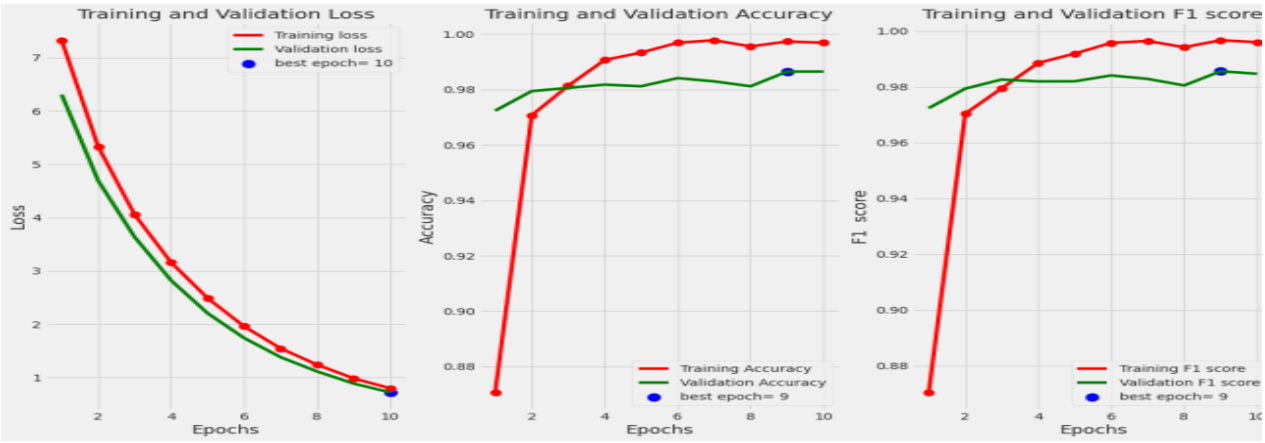
Dice measures the similarity ranges from to 1, 1 repress higher similarity. The precision indicates the relevant positive predictions. Recall is called as sensitivity; it represents how good a test is predicting correctly. The TP and FP stands for true positive and false positive respectively. TN and FN represents true negative and false negative. The measures are described in Fig 4.

		Actual	
		Positive	Negative
Predicted	Positive	TP	FP
	Negative	FN	TN

Fig 4: Relationship with actual and predicted values

4.2 Discussion

During training, training and validation loss, training and validation accuracy, and training and F1 score are depicted in Fig. 5. The training and validation loss are given in Fig. 5 (a). The loss is higher in training but lower when working with the validation dataset. The loss value is reduced for subsequent epochs. This chart confirmed that 10 epochs are enough. Fig. 5 (b) represents the accuracy during the training and validation processes. From the first epoch, the training accuracy ranges from 0.6 to 0.99 at the ninth epoch. The F1 score is high during training at the 7th epoch, which is illustrated in Fig. 5 (c) though the F1 score in the validation data set is high at the 9th epoch. At the 9th epoch, all parameter values were saved and used for the testing process.



(a) (b) (c)

Fig 5: Relation between loss and epochs. (a) Training, validation loss and epochs, (b) Accuracy measure in training data and validation data and (c) F1 score

The comparative results of presented method with other methods which done the similar task is essential to ensure the strength and weaknesses of the proposed method. In the same way, for the comparison three previously stated methods are taken. They are called Brain Extraction Tool (BET) [4], Robust Learning-Based Brain Extraction (RoBEX) [10], and UNet3D [11]. The results of the existing methods are extracted from [13]. Dice value of the presented method with other early stated methods are depicted in Fig 6(a), which

reveals that the BET achieves less than 90%, RoBEX obtains a 95% dice score, the UNet3D and proposed methods get almost equal values, and the proposed method obtains a higher value than the UNet3D; the standard deviation of RoBEX is 0.0179, and the standard deviation of the proposed method is 0.0163. The proposed method attains a higher precision value, as given in Fig. 6 (b). The true value prediction of the proposed methods is higher than that of other methods. The sensitivity of the proposed method is 98%, 1% more than the UNet3D model, and 5% more than the RoBEX method. The overall result of the proposed method is higher than the other methods, ensuring outperformance of the proposed method. The proposed and UNet3D models follow the same architecture with less variation, though the pre-processing methods are different, and the latter provides the highest result of the proposed method. As per the discussion in [16] the RoBEX and UNet3D models have been using 100s of epochs to deal with the abnormal and normal images, but the proposed method is using the NFBS dataset alone and results in 9 epochs. [17-18].

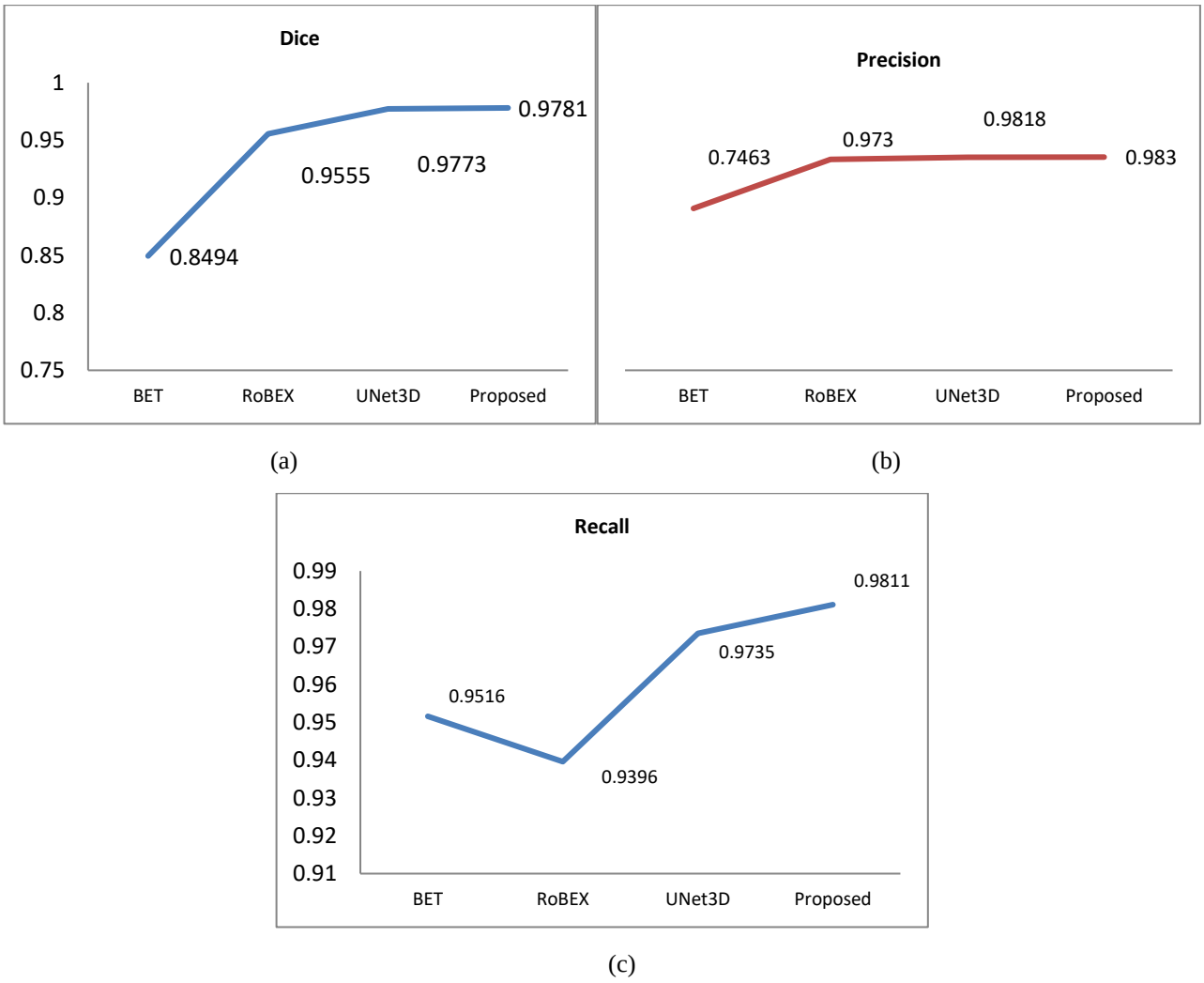


Fig 6: Quantitative charts. (a) Dice Score, (b) Precision and (c) Recall

5. Conclusion

In this work, we suggest an updated U Net-based deep learning method for brain extraction. It uses a completely automated computer-assisted process. Additionally, the proposed strategy produces a robust trained model. It can be used for tough multi center data as well as internal private data. The proposed method provides the good results when evaluate to the outcomes of the previously proposed methods. In addition, we first compare the effect of different techniques work for skull stripping using quantitative measures. We intend to expand the multi-cantered dataset used as training data in the future and create a sophisticated deep learning model for brain extraction.

References

[1] L. Pei, M. Ak, N. H. Tahon, S. Zenkin, S. Alkarawi, A. Kamal, M. Yilmaz, L. Chen, M. Er, N. Ak, and R. Colen, “A general skull stripping of multiparametric brain MRIs using 3D convolutional neural network,”

Scientific Reports. vol. 12, no. 1, pp. 10826, Jun 2022.

[2] M. E. Brummer, R. M. Mersereau, R. L. Eisner, and R. R. Lewine, "Automatic detection of brain contours in MRI data sets," IEEE Transactions on medical imaging. vol.12, no.2, pp.153-66, Jun 1993

[3] D.W. Shattuck, S.R. Sandor-Leahy, K.A. Schaper, D.A. Rottenberg, and R.M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," NeuroImage. vol.13, no.5, pp.856-76, May 2001

[4] S.M. Smith, "Fast robust automated brain extraction," Human brain mapping. vol.17, no.3, pp.143-55, Nov 2002

[5] M. Jenkinson, M. Pechaud, and S. Smith, "BET2: MR-based estimation of brain, skull and scalp surfaces," InEleventh annual meeting of the organization for human brain mapping (vol. 17, no. 3, p. 167), Jun 2005

[6] K.K. Leung, J. Barnes, M. Modat, G.R. Ridgway, J.W. Bartlett, and N.C. Fox, S, Ourselin, "Alzheimer's Disease Neuroimaging Initiative. Brain MAPS: an automated, accurate and robust brain extraction technique using a template library," Neuroimage. vol.55, no.3, pp.1091-108, Apr 2011

[7] S.F. Eskildsen, P. Coupé, V. Fonov, J.V. Manjón, K.K. Leung, N. Guizard, S.N. Wassef, L.R. Østergaard, and D.L. Collins, "Alzheimer's Disease Neuroimaging Initiative," BEaST: brain extraction based on nonlocal segmentation technique. NeuroImage. vol.59, no.3, pp.2362-73, Feb 2012

[8] S. Roy, J.A. Butman, and D.L. Pham, "Alzheimers Disease Neuroimaging Initiative Robust skull stripping using multiple MR image contrasts insensitive to pathology," Neuroimage. vol.146, pp.132-47, Feb 2017

[9] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review. Computational intelligence and neuroscience," vol.2018, Feb 2018

[10] J.E. Iglesias, C.Y. Liu, P.M. Thompson, and Z. Tu, "Robust brain extraction across datasets and comparison with publicly available methods," IEEE transactions on medical imaging. vol.30, no.9, pp.1617-34, Apr 2011

[11] H. Hwang, H.Z. Rehman, and S. Lee, "3D U-Net for skull stripping in brain MRI," Applied Sciences. vol.9, no.3, pp.569, Feb 2019

[12] H. Chen, Q. Dou, L. Yu, J. Qin, and P.A. Heng, "Vox ResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," NeuroImage. vol.170, pp.446-55, Apr 2018

[13] P.H. Scherrer, R.S. Bogart, R.I. Bush, J.A. Hoeksema, A.G. Kosovichev, J. Schou, W. Rosenberg, L. Springer, T.D. Tarbell, A. Title, and C.J. Wolfson, "The solar oscillations investigation—Michelson Doppler imager," The SOHO Mission. pp.129-88, 1995

[14] J. Swiebocka-Wiek. "Skull stripping for MRI images using morphological operators," In Computer Information Systems and Industrial Management: 15th IFIP TC8 International Conference, CISIM 2016, Vilnius, Lithuania, September 14-16, Proceedings 15 2016 (pp. 172-182). Springer International Publishing. 2016

[15] G. Santulli, "Epidemiology of cardiovascular disease in the 21st century: Updated updated numbers and updated facts," Journal of Cardiovascular Disease Research. vol.1, no.1, Jul 2013

[16] L. Pei, M. Ak, N.H. Tahon, S. Zenkin, S. Alkarawi, A Kamal, M Yilmaz, L Chen, M Er, N Ak, and R Colen, "A general skull stripping of multi parametric brain MRIs using 3D convolutional neural network," Scientific Reports. vol.12, no.1, pp.10826, Jun 2022

[17] H.K. Hahn, and H.O. Peitgen. "The skull stripping problem in MRI solved by a single 3D watershed transform," In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2000: Third International Conference, Pittsburgh, PA, USA, October 11-14, 2000. Proceedings 3 (pp. 134-143). Springer Berlin Heidelberg. 2000

[18] I. Good fellow, Y. Bengio, and A. Courville. "Deep learning," MIT press. Nov 2016

[19] Sikri, A. ., Singh, N. P. ., & Dalal, S. . (2023). Analysis of Rank Aggregation Techniques for Rank Based on the Feature Selection Technique. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 95–108. https://doi.org/10.17762/ijritcc.v11i3s.6160

[20] Kevin Harris, Lee Green, Juan González, Juan Garciam, Carlos Rodríguez. Automated Content Generation for Personalized Learning using Machine Learning. Kuwait Journal of Machine Learning, 2(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/180

[21] Keerthi, R.S., Dhabliya, D., Elangovan, P., Borodin, K., Parmar, J., Patel, S.K. Tunable high-gain and multiband microstrip antenna based on liquid/copper split-ring resonator superstrates for C/X band communication (2021) Physica B: Condensed Matter, 618, art. no. 413203