Brain Tumor Segmentation using U-Net and SegNet

Pankaj Kasar[1], Shivajirao Jadhav[2], Vineet Kansal[3]
[1]Department of Information Technology, ‡ Department of Computer Application,
[2]Dr. Babashaeb Ambedkar Technological University, Lonere, Raigad, India
[3]Institute of Engineering & Technology, Lucknow, UP, India
Email: *erpankajkasar@rediffmail.com, †smjadhav@dbatu.ac.in, ‡vineetkansal@ietlucknow.ac.in,

***Abstract*—The detection of tumors is the most difficult aspect of quantitative brain tumor evaluation. Magnetic Resonance Imaging (MRI) has gained popularity in recent years due to its non-invasive and powerful soft tissue contrast. MRI is a frequent imaging method used to detect brain cancers. The MRI produces a tremendous amount of data. Heterogeneity, isointensity, and hypointensity are characteristics of tumors that impede manual segmentation in a reasonable amount of time, hence limiting the use of valid quantitative measurements in clinical practice. In clinical practice, manual segmentation tasks are time-intensive and their performance is heavily dependent on the operator's level of expertise. Also required are accurate and automated tumor segmentation approaches; however, the high spatial and structural variability of brain tumors makes automatic segmentation a challenging task. This paper proposes fully automatic segmentation of brain tumors using convolutional neural networks with encoder-decoders. This work focuses on well-known deep neural networks for semantic segmentation, namely U-Net, and SegNet, for segmenting tumors from Brain MRI data. The networks are trained and evaluated using a pub- licly available standard dataset, with Dice Similarity Coefficient (DSC) as a metric for the entire predicted image (tumor and background). The average DSC for U-Net on the test dataset is 0.76, while the average DSC for SegNet is 0.67. The examination of results demonstrates that U-Net performs better than SegNet.**

***Index Terms*—U-Net, SegNet, Convolutional Neural Network (CNN), Brain tumor, segmentation, Magnetic Resonance Imaging (MRI).**

I. Introduction

In numerous academic domains, including Natural Lan- guage Processing, Image Analysis, and a variety of Expert Systems, today's deep learning approaches promise promising solutions. Due to developments in CNNs, the medical imaging field has achieved significant strides in recent years. It is also considered to be an important approach for various applications in the near future. Segmentation of abnormal- ities in medical imaging is one of the greatest challenges, including segmentation of brain tumors [1], cardiac ventricles [2], abdominal organs [3], cells in biological imaging [4], and detecting diabetic retinopathy [22][23][24].

In recent years, human understanding and advancements in the field of health care have demonstrated that diseases have declined, but even cancer, due to its erratic existence, remains a human plague. Brain tumor malignancy is regarded as one of the most fatal illnesses. The brain is the administrative and control center of the human body. It is responsible for all bodily functions, including respiration, muscle action, and our senses, through a complex network of neurons that

R. Murugan et al. (eds.), *Proceedings of the International Conference on Signal Processing and Computer Vision (SIPCOV 2023)*, Advances in Engineering Research 239,
https://doi.org/10.2991/978-94-6463-529-4_18

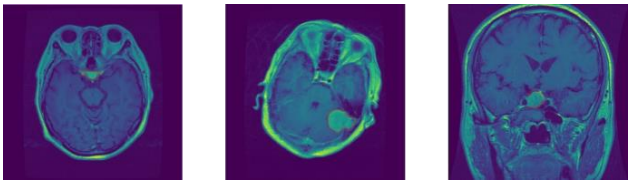
are interconnected internally. Brain tumors are formed by abnormal cell formation in the brain [5], which disrupts the function of the neurological system and, in the most severe cases, results in a very short lifespan. The diagnosis of early- stage brain tumors depends on the experience and level of competence of the physicians, making it simpler for the patient to recover and anticipate a continued life. The automated segmentation of brain tumors is an effective way for aiding clinicians in determining the best course of treatment. This automatic segmentation utilizing deep neural networks enables both inexperienced and skilled physicians to make correct decisions and resolve challenging situations. This method utilizes images captured by magnetic resonance (MR) imaging tools that are commonly utilized by radiologists for brain diagnosis.

Typically, brain tumors are classified as benign or malignant. The fact that benign tumors grow consistently and originate in the brain suggests that this type of tumor will not spread across the body. It is therefore assumed to be non-cancerous (non-progressive) and less aggressive. Malignant tumors, on the other hand, are malignant, spreading rapidly with undeter- mined bounds and invading healthy body cells. If a primary malignant tumor is detected in the brain, it is identified as such. If it originates elsewhere in the body and spreads to the brain, it is a second malignant tumor [6].

Nevertheless, glioma, meningioma, and pituitary tumors are additionally common types of brain tumors. Meningiomas are the most common benign tumors seen in the thin membranes surrounding the spinal cord and brain. Gliomas are a category of brain tumors [7] that form within the brain's tissue. With a minimum survival rate of several years, high-grade gliomas are one of the most severe forms of brain cancer. In the pituitary gland of the brain, pituitary tumors are developed. Figure 1 depicts the uniform shape and inherent nature of all of these tumors.

T1-weighted contrast-enhanced imaging techniques are of- ten used to identify primary cancers such as meningioma (MEN), astrocytoma (AS), medulloblastoma (MED), glioblastoma multiforme (GBM), and metastatic tumors (MET). Through the administration of 0.150.20 mMol/kg contrast ma- terial (Gadolinium) to patients, T1-weighted contrast-enhanced pictures of these tumors are marginally increased [8].

Conventional algorithms rely significantly on the manual extraction of complex, expert-level features. In contrast, fea- ture extraction in deep learning is performed automatically.



(a) (b) (c)

Fig. 1: Illustrations of three typical brain tumors: (a) menin- gioma; (b) glioma; and (c) pituitary tumor. Red lines indicate the tumor border

Each layer of the neural network (i.e., a deep network with multiple hidden layers) learns higher-level features by utilizing the features of the preceding layer. This paper consequently proposes two cutting-edge segmentation networks, U-Net and SegNet, for separating abnormalities in MRI images. Due to the fact that these networks provide an end-to-end solution to a problem, they have rigorous computational capacity requirements for huge datasets. Typically, graphics processing units (GPUs) are utilized to meet this computational criterion and permit network training in a timely manner.

The remaining sections of the paper are organized as follows. Section 2 literature review of the existing models available for segmentation. Section 3. describes the detailed architecture of semantic segmentation deep neural networks 3.a.1 U-Net and 3.a.2 SegNet. The standard dataset used for training and testing is publicly available and is enumerated in Section 3.b. Different data augmentation operations and preprocessing steps applied before training the networks are demonstrated in section 3.c. The hardware and software resources used for training the network are mentioned in Section 4. The metrics used to measure the performance of the network with loss function is enumerated in Section 4.a. The dice similarity coefficient, precision, and accuracy obtained during training with loss convergence are specified in Section 4.b. The segmentation results of U-Net and SegNet are presented in section 4. Finally, the discussion along with the conclusion is drawn in section 5 with future improvements.

II. RELATED WORK

Segmentation models are deep learning models used for image recognition and computer vision tasks, with the aim of identifying and classifying different objects within an image. The following literature review explores the development and progress of segmentation models to date, with references to relevant research.

Fully Convolutional Networks (FCN): Fully Convolutional Networks (FCN) were first introduced in 2014 by Long et al. [9]. FCN is a deep learning model that replaces the fully connected layers with convolutional layers. FCN has become a foundational model for semantic segmentation and has been used in various applications, including autonomous driving, medical imaging, and robotics.

U-Net: U-Net was introduced in 2015 by Ronneberger et al. [4]. It is a convolutional neural network that uses an encoder- decoder architecture for segmentation. U-Net has been widely

used for biomedical image segmentation and has been shown to achieve state-of-the-art performance on several datasets.

SegNet: SegNet was introduced in 2015 by Badrinarayanan et al. [10]. It is a deep convolutional encoder-decoder architecture that uses pooling indices to perform upsampling. SegNet has been used for various segmentation tasks, including road segmentation and medical image segmentation.

Mask R-CNN: Mask R-CNN was introduced in 2017 by He et al. [11]. It is an extension of the Faster R-CNN model and adds a branch that produces segmentation masks for each detected object. Mask R-CNN has achieved state-of-the-art performance on several segmentation benchmarks and has been used in various applications, including object detection, instance segmentation, and image segmentation.

DeepLab: DeepLab is a family of models for semantic seg- mentation introduced by Chen et al. in 2018 [12]. DeepLab uses dilated convolutions and atrous spatial pyramid pooling to extract features at multiple scales, and a decoder that combines the features to produce the final segmentation map. DeepLab has achieved state-of-the-art performance on various segmentation benchmarks.

HRNet: High-Resolution Network (HRNet) was introduced in 2019 by Sun et al. [13]. HRNet is designed to preserve high- resolution features throughout the network by using a multi- resolution fusion module that combines features from different resolutions. HRNet has achieved state-of-the-art performance on various segmentation benchmarks and has been used in var- ious applications, including semantic segmentation, instance segmentation, and pose estimation.

ViT-Seg: Vision Transformer Segmentation (ViT-Seg) is a recent model introduced in 2021 by Wang et al. [14]. ViT- Seg uses a transformer architecture for image segmentation, which has been shown to achieve state-of-the-art performance on several benchmarks.

In conclusion, segmentation models have come a long way since the introduction of FCN in 2014. With the devel- opment of new architectures and techniques, segmentation models have achieved state-of-the-art performance on various segmentation benchmarks and have been applied to various applications, including medical imaging, autonomous driving, and robotics.

III. METHODOLOGY

A. Semantic Segmentation

CNN-based classifier networks, like AlexNet [15], VGGNet [16], and GoogLeNet [17], etc., outperform the ImageNet classification task. These networks are currently configured by categorizing pixels in order to complete the segmentation of images. Unfortunately, the performance of the aforementioned networks is not identical, but the notion of down-sampling and up-sampling utilizing pooling and unpooling layers makes segmentation very efficient. In semantic segmentation, the image is segmented on a pixel-label basis, which is the act of assigning a class label to each pixel in an image. These labels can be a person, a car, a flower, a piece of furniture, or a tumor, among other things. Autonomous vehicles, human-

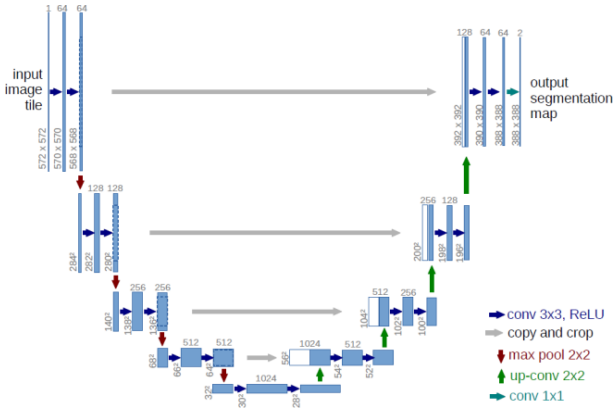


Fig. 2: The architecture of the U-net. The boxes correspond to a map of a multi-channel function. The number of channels is displayed at the top of the box. The x-y-size is seen at the bottom left edge of the box. White boxes reflect copied feature maps. The arrows signify the various operations [4]

computer interaction, robotics, photo editing/creativity tools, etc., are some of its principal applications. On each layer, the deep neural networks are trained in an end-to-end, pixel-by-pixel fashion. The following section will illustrate two neural networks for semantic segmentation: U-Net and SegNet.

1) U-Net: The U-Net network architecture is illustrated in Fig 2. This is a very popular semantic segmentation network for biomedical imaging proposed by Olaf Ronneberger in May 2015 [4]. It consists of three parts i) The contract- ing/downsampling path ii) the Bottleneck iii) The expand- ing/upsampling path. This architecture is considered as an extension of a Fully Convolutional Network [9] in a way that,
i) U-Net is symmetric, ii) the skip connections between the contracting path and the expanding path apply a concatenation operator instead of a sum.

During upsampling, the skip connections are responsible for providing local information to global information. Owing to its symmetric structure, the upsampling path has a high number of feature maps that facilitate the transfer of data.
Contract/downsampling path: consists of 4 blocks, each of which consists of two 3x3 Convolution Layer + Activation functions, followed by a 2x2 Max Pooling layer. Notice that the number of feature maps doubles at each pooling layer, starting with 64 feature maps for the first block, 128 for the second block, and so on. The aim of this contracting path is to collect the context of the input image and via skip connections, this rough contextual information will then be passed to the upsampling path.
Bottleneck: This section of the network is the juncture between contracting and expanding routes. It's constructed from only two convolution layers (with batch normalization).
Expanding/upsampling path: It is also a collection of 4 blocks with each block consisting of - Deconvolution layer with stride 2, Concatenation with the corresponding

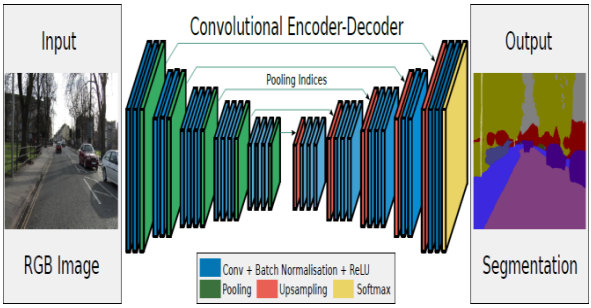


Fig. 3: SegNet architecture [10]

cropped feature map from the contracting path, and two 3x3 Convolution layer + activation function. The purpose of this expanding route is to allow for a specific position combined with the qualitative details of the contracting path.

2) SegNet: In 2015, Badrinarayanan et al. [10] suggested a CNN-based deep neural network for semantic segmentation. As depicted in Figure 3, it is composed of downsampling- upsampling blocks. The downsampling path consists of five encoders with thirteen convolutional layers borrowed from the initial thirteen layers of the VGG16 network [17]. Each encoder layer corresponds to a decoder layer, therefore the decoder route also consists of 13 levels. Each encoder permits a convolution with a filter set to generate a sequence of feature maps for batch normalization using rectified linear unit (ReLU) operations, followed by Max pooling with a 2 x 2 window and stride 2. The use of max-pooling indices in decoders to upsample low-resolution feature maps is a basic principle of SegNet. Pooling indices provide the benefit of maintaining high-frequency information in processed images with lower decoder training settings.

B. Dataset Description

We used a publicly accessible brain tumor dataset-Figshare [18]-to train and test the U-Net and SegNet segmentation networks. The dataset contains 3064 brain MRIs obtained from 233 patients. It comprises three types of brain tumors: meningioma (708), pituitary (930), and glioma (1426). The data set is already available ".mat" format in MATLAB. The configuration of the ".mat" file comprising a patient ID, a unique mark indicating the type of brain tumor, 512 × 512 image data in uint16 format, a vector containing a tumor boundary with distinct point coordinates, and ground truth in binary mask image.

C. Preprocessing and Data Augmentation

The images and ground truths extracted are reduced to size 256 X 256 to have a low computational cost. The image data is considered in colored (RGB channels) whereas ground truths are in binary as shown in figure 4. We have divided 3064 images of a dataset into 80 % training data and 20 % testing

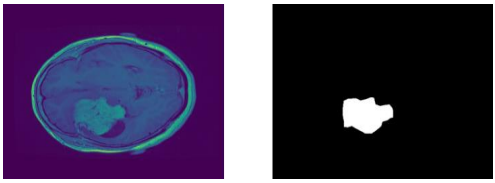


Fig. 4: Colored image and binary ground truth

data. Thus 2451 images in the training set and 613 images in images in the test set.
To avoid overfitting [19] and to address the class imbalance problems, the following techniques are used:

- Data augmentation: Data augmentation can be used to generate more training data from the existing data by ap- plying transformations such as rotation, scaling, flipping, shearing, and cropping. This can help balance the classes and prevent overfitting to the majority class.
- Class weighting: Assigning weights to each class based on its frequency can help balance the classes. The weight for each class can be inversely proportional to its frequency so that the model pays more attention to the minority classes.

Hence, we have increased the size of the training data set 10 times, i.e., the total number of training images is now 24510, and class weights are calculated, which are subsequently utilized in loss functions.

IV. EXPERIMENTATION AND RESULTS

The implementation of U-Net and SegNet has been carried out with KERAS using TensorFlow as the backend in python. The networks have been trained for a number of epochs using an NVIDIA TESLA V100 graphics card with 16 GB of dedicated memory with 192 GB of DDR4, 2666 MHz, and an Intel Xeon SKL G-6148 processor. Hyperparameters of the network, such as the rate of learning, size of the batch, and the number of epochs, are tuned on a trial-and-error basis. The training optimizer used is Adam [20] with an initial 0.0001 learning rate. Adam, an extension of the stochastic gradient descent (SGD) algorithm, preserves the learning rate for each network weight and adapts it independently as learning continues. We used the ”ReduceLROnPlateau” Keras callback to adaptively minimize the learning rate when the parameter has stopped improving. Models also benefit from the reduction of the learning rate by a factor of 10 after learning has stagnated. This callback tracks the quantity and if there is no change in the ”patience” -a number of epochs, the learning rate is decreased.

A. Evaluation metrics

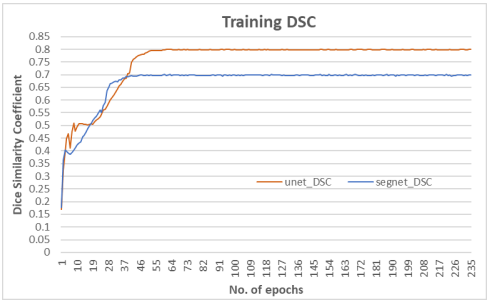
The following are the metrics to analyze the performance of segmentation networks: I

1) Dice Similarity Coefficient: The statistic used to calcu- late the efficiency of the segmentation is the Dice sim- ilarity coefficient. It’s the proportion of the area to the union area between the ground truth image (X) and the segmented image (Y) and can be found using the following equation

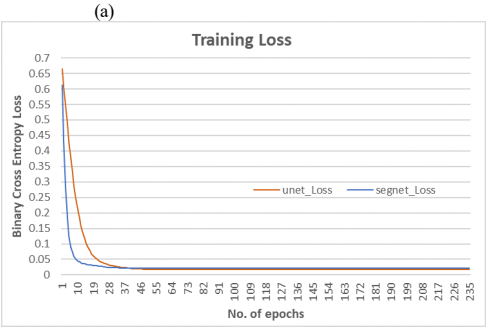
$$DSC = \frac{2\,|X \cap Y|}{|X| + |Y|} \qquad (1)$$

2) Accuracy: Often known as classification accuracy, which is the ratio of accurate predictions to the total number of predictions made by the classifier. It can be shown by using the equation

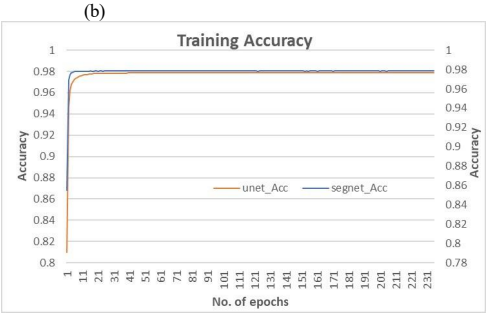
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (2)$$



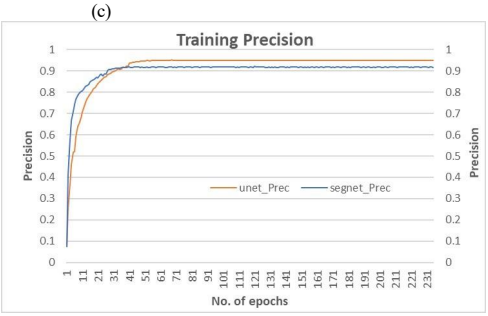
(a)



(b)



(c)



(d)

Fig. 5: Performance on Training Dataset a. DSC, b. Loss, c. Accuracy, d. Precision

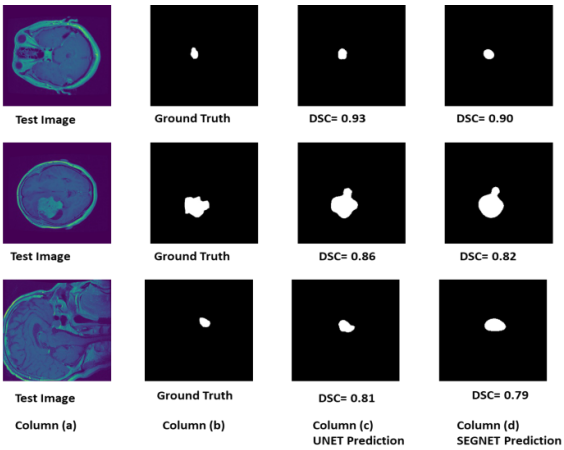


Fig. 6: Visualization of Segmented Results a. Test images, b. Ground Truths, c. Prediction by U-Net d. Prediction by SegNet

where, TP represents the number of True Positives, i.e. pixels that are properly classified as tumor area, True Negative (TN) shows the number of real negative ones, i.e. pixels belonging to the background, properly classified as background, False Positive (FP) indicates the number of false positives, i.e. the number of wrongly classified as tumor pixels, and False Negative (FN) denotes the number of false negative, i.e. pixels belonging to tumor region falsely classified as background [21].

3) Precision: Percentage of positive instances out of the total predicted positive instances. Here denominator is the model prediction done as positive from the whole given dataset. Take it as to find out 'how much the model is right when it says it is right?'. It is used to find the actual tumorous pixels.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

4) Loss Function: The function used to measure the loss is Binary cross entropy also called Sigmoid Cross-Entropy loss. The equation for Binary Cross-Entropy loss given in eqn 4.

$$BCE(y, y\hat{}) = -(y \log y\hat{} + (1 - y) \log(1 - y\hat{})) \quad (4)$$

where y is the true label (either 0 or 1), and y^ is the predicted probability of the positive class.

TABLE I: Training Performances with parameters

Network Architecture	U-Net	SegNet
Optimizer	Adam	
Initial Learning Rate	0.0001	
No. of epochs	235	
Images	24510	
Binary cross entropy loss	0.01795996	0.021663887
Dice Similarity Coefficient	0.800096512	0.701303184
Accuracy	0.978710473	0.978591979
Precision	0.951362908	0.921810567

B. Training performance

The training performance of U-Net and SegNet is shown in figure 5. Various training experiments with different hyperpa- rameters are performed on both networks. The best training performance with suitable final hyperparameters is shown in table 1. As depicted in figure 5b loss of U-Net is minimum than SegNet. The training Dice Similarity Coefficient is raised to 0.80 and 0.70 for U-Net and SegNet respectively as shown in 5a. Fig 5c and 5d denote the Accuracy and Precision of both networks respectively. U-Net performed well as compared to SegNet in all respects.

C. Segmentation Results

The U-Net and SegNet networks are evaluated on the test dataset of size 613 images. The average DSC of a testingdataset for U-Net and SegNet is 0.76 and 0.67 respectively with a precision score of 0.90 for both. The accuracy and precision of the segmented image are often higher than the DSC, as it also involves the correctly classified background area. In addition, it should not rationally present the predictive performance of the model since it suffers from the contradic- tion of accuracy (which suggests that a trained model with a given degree of accuracy can have more predictive power than the models do better accuracy). Thus, when analyzing the segmentation networks, we put more emphasis on DSC in addition to accuracy and precision. The visualization of pre- dictions of some of the testing image segmentations is shown in figure 6. It is clearly indicated that U-Net outperformed the SegNet.

V. DISCUSSION AND CONCLUSION

In this study, we demonstrate the effectiveness of semantic segmentation networks, namely U-Net and SegNet, in the au- tomatic segmentation of brain tumors. We have demonstrated that U-Net outperforms SegNet on the figshare brain MRI image dataset for segmentation because U-Net is designed to handle high-resolution images and maintain fine details in the segmentation results, making it particularly well-suited for

medical image analysis tasks where precise segmentation is important, such as in tumor or lesion segmentation. SegNet, on the other hand, is a deep encoder-decoder network that uses pooling indices from the encoder to perform upsampling in the decoder. It is designed to be more computationally efficient than U-Net, but it can result in some loss of spatial resolution in the segmentation results. In terms of accuracy, U- Net generally performs better than SegNet on medical image segmentation tasks, due to its ability to maintain fine details in the segmentation results. This is due to the use of skip connections, which help to preserve high-level features and improve the accuracy of the segmentation results.

We intend to increase the performance of both networks by partitioning the training dataset image into a number of small patches. Future dice similarity coefficients should be reported separately for tumor and background to improve segmented image analysis. In addition, the effect of enlarging encoder- decoder blocks in each network can be further examined. After effective segmentation with high DSC, we attempt to automatically classify tumors as gliomas, meningiomas, or pituitary tumors.

ACKNOWLEDGMENT

The support and the resources provided by 'The PARAM Shakti Facility' under the National Supercomputing Mission, Government of India at the Indian Institute of Technology, Kharagpur are gratefully acknowledged. The authors are also grateful to the anonymous referees for the insights and the priceless recommendations provided in order to progress the initial version of this manuscript.

FUNDING

Financial support under the scheme 'Collaborative Research Innovation Project (CRIP)', TEQIP-III, Government of India through Dr. Babasaheb Ambedkar Technological University, Lonere Raigad, M.S. India is gratefully acknowledged.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

CONSENT TO PARTICIPATE

Informed consent was obtained from all individual partici- pants included in the study.

AUTHORS' CONTRIBUTIONS

PEK, SMJ, and VK have contributed to the research in the order they appear. PEK and VK are the principal investigator and co-investigator of the research project respectively. SMJ is the research supervisor. Also, all authors have discussed the results as well as improved the final manuscript.

REFERENCES

[1] de Brebisson, A., Montana, G. (2015). Deep neural networks for anatomical brain segmentation. In CVPR Workshops (pp. 20-28).

[2] Avendi, M., Kheradvar, A., Jafarkhani, H. (2016). A combined deep- learning and deformable-model approach to fully automatic segmenta- tion of the left ventricle in cardiac MRI. Medical Image Analysis, 30, 108-119.

[3] Roth, H. R., Oda, H., Hayashi, Y., Oda, M., Shimizu, N., Fujiwara, M., Misawa, K., Mori, K. (2017). Hierarchical 3D fully convolutional networks for multi-organ segmentation. arXiv preprint:1704.06382.

[4] Ronneberger, O., Fischer, P., Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In MICCAI (pp. 234-241).

segmentation. In MICCAI (pp. 234-241).

[5] Razzak, M. I., Imran, M., Xu, G. (2018). Efficient brain tumor seg- mentation with multi-scale statistics two pathway- group conventional neural networks. IEEE Journal of Biomedical and Health Informatics. https://doi.org/10.1109/JBHI.2018.2874033

[6] Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A., Mengko, T. R. (2019). Brain tumor classification using convolutional neural network. In World Congress on Medical Physics and Biomedical Engineering (pp. 183-189). Springer.

[7] Abir, T. A., Siraji, J. A., Ahmed, E., Khulna, B. (2018). Analysis of a novel MRI-based brain tumor classification using probabilistic neural network (PNN). International Journal of Scientific Research in Science, Engineering, and Technology, 4(8), 65-79.

[8] Rehman, A., Naz, S., Razzak, M. I., Akram, F., Imran, M. (2019). A deep learning-based framework for automatic brain tumor classifica- tion using transfer learning. Circuits, Systems, and Signal Processing. https://doi.org/10.1007/s00034-019-01246-3

[9] Long, J., Shelhamer, E., Darrell, T. (2015). Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431-3440.

[10] Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmenta- tion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481-2495.

[11] He, K., Gkioxari, G., Dollar, P., Girshick, R. (2017). Mask R-CNN. Proceedings of the IEEE International Conference on Computer Vision.

[12] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A. L. (2018). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE transactions on pattern analysis and machine intelligence, 40(4), 834-848.

[13] Sun, K., Xiao, B., Liu, D., Wang, J. (2019). Deep high-resolution representation learning for visual recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

[14] Wang, Z., Liu, Q., Wu, S., Li, Z., Zou, J. (2021). Vision transformer segmentation. arXiv preprint arXiv:2103.10619.

[15] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). Imagenet classifi- cation with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097–1105).

[16] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. In Proceedings of the international conference on learning representations. arXiv:1409.1556.

[17] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. (2015). Going deeper with convolutions. In Computer vision and pattern recognition (CVPR). arXiv:1409.4842.

[18] Cheng, J. (n.d.). Brain tumor dataset [Data set]. figshare. https://doi.org/10.6084/m9.figshare.1512427.v5

[19] Wang, J., Perez, L. (2017). The effectiveness of data augmentation in image classification using deep learning (Tech. Rep.). Technical report.

[20] Kingma, D., Ba, J. (2015). Adam: A method for stochastic optimization. In Proceedings of the 3rd international conference on learning represen- tations (ICLR).

[21] Mittal, A., Hooda, R., Sofat, S. (2018). LF-SegNet: A Fully Convo- lutional Encoder-Decoder Network for Segmenting Lung Fields from Chest Radiographs. Wireless Personal Communications, 98(1), 149-165.

[22] N. J. Mohan, R. Murugan, T. Goel and P. Roy, "Exudate Detection with Improved U-Net Using Fundus Images," 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 560-564, doi: 10.1109/ComPE53109.2021.9752239.

[23] Nagula, J. M., Murugan, R., Goel, T. (2023). Role of Machine and Deep Learning Techniques in Diabetic Retinopathy Detection. In Multi- disciplinary Applications of Deep Learning-Based Artificial Emotional Intelligence (pp. 32-46). IGI Global.

[24] Jagan Mohan, N., Murugan, R., Goel, T. (2022). Deep Learning for Dia- betic Retinopathy Detection: Challenges and Opportunities. In: Tripathy, B.K., Lingras, P., Kar, A.K., Chowdhary, C.L. (eds) Next Generation Healthcare Informatics. Studies in Computational Intelligence, vol 1039. Springer, Singapore. https://doi.org/10.1007/978-981-19-2416-3$_1$2.