

# MAG-Net: Multi-task attention guided network for brain tumor segmentation and classification \*

Sachin Gupta, Narinder Singh Punn, Sanjay Kumar Sonbhadra, and Sonali Agarwal

Indian Institute of Information Technology Allahabad, India  
 {mit2019075, pse2017002, rsi2017502, sonali}@iiita.ac.in

**Abstract.** Brain tumor is the most common and deadliest disease that can be found in all age groups. Generally, MRI modality is adopted for identifying and diagnosing tumors by the radiologists. The correct identification of tumor regions and its type can aid to diagnose tumors with the followup treatment plans. However, for any radiologist analysing such scans is a complex and time-consuming task. Motivated by the deep learning based computer-aided-diagnosis systems, this paper proposes multi-task attention guided encoder-decoder network (MAG-Net) to classify and segment the brain tumor regions using MRI images. The MAG-Net is trained and evaluated on the Figshare dataset that includes coronal, axial, and sagittal views with 3 types of tumors meningioma, glioma, and pituitary tumor. With exhaustive experimental trials the model achieved promising results as compared to existing state-of-the-art models, while having least number of training parameters among other state-of-the-art models.

**Keywords:** Attention · Brain tumor · Deep learning · Segmentation.

## 1 Introduction

Brain tumor is considered as the deadliest and most common form of cancer in both children and adults. Determining the correct type of brain tumor in its early stage is the key aspect for further diagnosis and treatment process. However, for any radiologist, identification and segmentation of brain tumor via multi-sequence MRI scans for diagnosis, monitoring, and treatment, are complex and time-consuming tasks.

Brain tumor segmentation is a challenging task because of its varied behavior both in terms of structure and function. Furthermore, the tumor intensity of a person differs significantly from each other. MRI is preferred over other imaging modalities [4] for the diagnosis of brain tumor because of its non-invasive property that follows from without the exposure to ionizing radiations and superior image contrast in soft tissues.

---

\* All authors have contributed equally.

Deep learning has shown advancement in various fields with promising performance especially in the area of biomedical image analysis [24]. The convolutional neural networks (CNN) [2] are the most widely used models in image processing. The CNNs involve combination of convolution, pooling and activation layers accompanied with the normalization and regularization operations to extract and learn the target specific features for desired task (classification, localization, segmentation, etc.). In recent years various techniques have been proposed for identification (classification and segmentation) of the brain tumor using MRI images that achieved promising results [13,30]. However, most of the approaches use millions of trainable parameters that result in slower training and analysis time, while also having high variance in results in case of limited data samples.

In order to overcome the aforementioned drawbacks, Ronneberger et al. [26] proposed U shaped network (U-Net) for biomedical image segmentation. The model follows encoder-decoder design with feature extraction (contraction path) and reconstruction phases (expansion path) respectively. In addition, skip connections are introduced to propagate the extracted feature maps to the corresponding reconstruction phase to aid upsample the feature maps. Finally, model produces segmentation mask in same dimensions as the input highlighting the target structure (tumor in our case). Following the state-of-the-art potential of the U-Net model, many U-Net variants are proposed to further improve the segmentation performance. Attention based U-Net model [19] is one such variant that tend to draw the focus of the model towards target features to achieve better segmentation results. The attention filters are introduced in the skip connections where each feature is assigned weight coefficient to highlight its importance towards the target features. Despite achieving the promising results, these models have millions of trainable parameter which can be reduced by optimizing the convolution operation. This can be achieved by incorporating the depthwise convolution operations [8] that is performed in two stages: depthwise and pointwise convolutions. The reduction in the number of the parameters and multiplications as compared to standard convolution operation can be represented as  $1/r + 1/f^2$ , where  $r$  is the depth of the output feature map and  $f$  is the kernel height or width [21]. The achieved reduction in number of parameters and multiplications is  $\sim 80\%$ . Following this context, attention guided network is proposed that uses depthwise separable convolution for real time segmentation and classification of the brain tumor using MRI imaging. The major contribution of the present research work is as follows:

- A novel model, Multi-task (segmentation and classification) attention guided network (MAG-Net) is proposed for brain tumor diagnosis.
- Optimization of training parameters using depthwise separable convolution. The training parameters of the MAG-Net reduced from 26.0M to 5.4M.
- MAG-Net achieved significant improvement in classification and segmentation as compared to the state-of-the-art models while having limited data samples.

The rest paper is organized as follows: Section 2 describes the crux of related work on brain tumor segmentation and classification. Section 3, talks about the proposed architecture, whereas Section 4 discusses the training and testing environment with experimental and comparative analysis. Finally, concluding remarks are presented in Section 5.

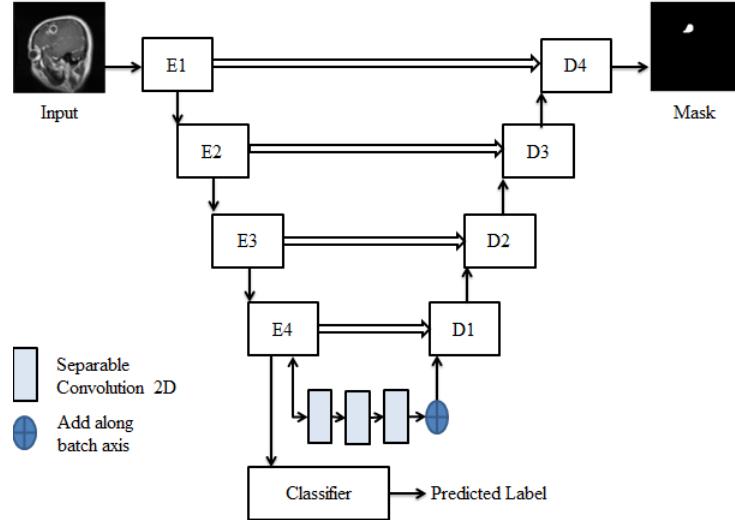
## 2 Literature review

Identifying the brain tumor is a challenging task for the radiologists. Recently, several deep learning based approaches are proposed to aid in faster diagnosis of the diseases. Segmentation of the infected region is most common and critical practice involved in the diagnosis. In addition, the segmented region can be provided with label (classification) to indicate what type of anomaly or infection is present in the image.

In contrast to the traditional approaches, Cheng et al. [7] proposed a brain tumor classification approach using augmented tumor region instead of original tumor region as RoI (region of interest). Authors utilized the bag of word (BOW) technique to segment and extract local features from RoI. Dictionary is used for encoding the extracted local features maps that are passed through SVM (support vector machine) classifier. The approach outperformed the traditional classification techniques with the accuracy of 91.28% but the performance is limited by the data availability. In similar work, Ismael et al. [14] proposed an approach of combining statistical features along with neural networks by using filter combination: discrete wavelet transform (DWT)(represented by wavelet coefficient) and Gabor filter (for texture representation). For classification of the tumor, three layered neural network classifier is developed using multilayer perceptron network that is trained with statistical features. In contrast to Cheng et al. [7], authors also achieved promising results on the limited data samples with an overall accuracy of. 91.9%.

Recently, capsule network [12] has shown great performance in many fields especially in biomedical image processing. Afshar et al. [1] proposed basic cap-snet with three capsules in last layer representing three tumor classes. However, due to varied behavior (background, intensity, structure, etc.) of MRI image, the proposed model failed to extract optimal features representing the tumor structure. The author achieved the tumor classification accuracy of 78% and 86.5% using raw MRI images and tumor segmented MRI images respectively. In another approach, Pashaei et al. [20] utilized CNN and kernel extreme learning machine that comprises one hidden layer with 100 neurons to increase the robustness of the model. With several experimental trials, the authors achieved an accuracy of 93.68% but detects only 1% of the positive pituitary tumor cases out of the total pituitary tumor case. Deepak et al. [9] proposed a transfer learning approach that uses pre-trained GoogleNet model to extract features (referred as deep CNN features) with softmax classifier in the output layer to classify three tumor classes. Furthermore, the authors combine the deep CNN features and SVM model to analyse the classification performance. The authors achieved 97.1% accuracy but

resulted in poor performance by standalone GoogleNet model due to overfitting with limited training image dataset, and misclassifications in meningioma tumor. In another approach, Pernas et al. [11] proposed to process images in three different spatial scales along with multi pathways feature scales for classification and segmentation of brain tumor. The images are pre-processed with elastic transform for preventing overfitting. The model analyses entire image and classifies pixel by pixel in one of four possible output labels (i.e. 0-healthy, 1-menigoia, 2-glioma, and 3-pituitary tumor). The proposed approach outperformed existing approaches with 97.3% classification accuracy, but with poor segmentation performance. Following this context, in this article multi-task attention guided network (MAG-Net) is proposed based on the U-Net architectural design [26] that uses parallel depthwise separable convolution layers for multi-level feature extraction along with an attention mechanism to better extract tumor features for brain tumor classification and generate the corresponding tumor mask.



**Fig. 1.** Schematic representation of the architecture of MAG-Net model.

### 3 Proposed work

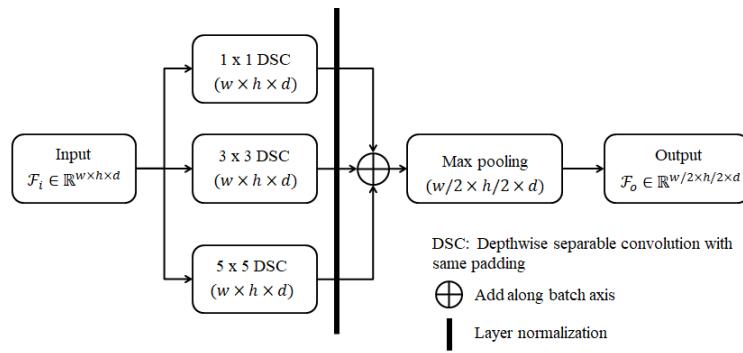
The proposed multi-task attention guided network (MAG-Net) model, as shown in Fig. 1, focuses on reducing overall computation, better feature extraction and optimizing the training parameters by reduction. The overall architectural design consists of an encoder, decoder, and classification module with 5.4M trainable parameters. The overall architectural design of the model is inspired by the U-Net encoder-decoder style [23]. Due to its state-of-the-art potential, this model is

the most prominent choice among the researchers to perform biomedical image segmentation [17].

In MAG-Net to reduce the number of training parameters without the cost of performance, standard convolution operations are replaced with depthwise separable convolution. In addition, the skip connections are equipped with attention filters [19] to better extract the feature maps concerning the tumor regions. The attention approach filters the irrelevant feature maps in the skip connection by assigning weights to highlight its importance towards the tumor regions. Besides, the encoder block is equipped with parallel separable convolution filters of different sizes, where the extracted feature maps are concatenated for better feature learning. These features are then passed to the corresponding decoder blocks via attention enabled skip connections to aid in feature reconstruction with the help of upsampling operation. The bottleneck layer connects the feature extraction path to the feature reconstruction path. In this layer filters of different sizes are used along with the layer normalization. Furthermore, the classification is performed using the extracted feature maps obtained from the final encoder block.

### 3.1 Encoder

To detect the shape and size of varying image like brain tumor it is required to use separable convolution of different sizes. Inspired from the concept of inception neural network [22] the encoder segment is consist of separable convolutions of  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$  kernels. Each of separable convolutions are followed by layer normalization. The extracted feature maps are fused with add operation that are downsampled by max pooling operation. Fig. 2, shows the proposed encoder architecture of MAG-Net model for some input feature map,  $\mathcal{F}_i \in \mathbb{R}^{w \times h \times d}$ , where  $w$ ,  $h$  and  $d$  are the width, height and depth of the feature map.



**Fig. 2.** Proposed MAG-Net 2D encoder module.

### 3.2 Decoder

The decoder component follows from the encoder block and that tend to reconstruct the spatial dimension to generate the output mask in same dimension as input. It consists of upsampling of the feature maps along with the concatenation with the attention maps followed by a separable convolution operation. Long skip connections [10] are used to propagate the attention feature maps from encoder to decoder to recover spatial information that was lost during downsampling in encoder. By using attention in the skip connection it helps the model to suppress the irrelevant features.

### 3.3 Classification

This module classifies the brain tumor MRI images into respective classes i.e meningioma, glioma, and pituitary tumor by utilizing the features extracted from the encoder block. This encoder block act as backbone model for both classification and segmentation, thereby reducing the overall complexity of the model. In this classification block the feature maps of the last encoder block act as input that are later transformed into 1D tensor by using global average pooling. The pooled feature maps are then processed with multiple fully connected layers. The classification output is generated from the softmax activated layer that generates the probability distribution of the tumor classes for an image.

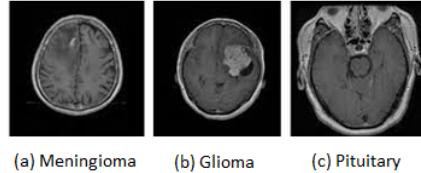
## 4 Experiment and Results

### 4.1 Dataset Setup

The present research work utilizes Figshare [6] dataset that comprises of 2D MRI scan with T1-weighted contrast-enhanced modality acquired from 233 patients to form a total of 3064 MRI scans. The T1 modality highlight distinct features of the brain tumor with three classes representing the type of brain tumor i.e. meningioma (708 slices), glioma (1426 slices), and pituitary (930 slices) forming 23%, 46.5%, and 30% class distribution in the dataset respectively. The sample MRI slices of different tumor classes are presented in Fig. 3. Dataset is randomly split into 80% training and 20% of the validation set. The training and testing composition kept the same throughout the experiment trails for comparative analysis.

### 4.2 Training and Testing

The MAG-Net model is trained and evaluated on the Figshare dataset. The training phase is accompanied with early-stopping [3] to tackle the overfitting problem, and Adam as a learning rate optimiser [27]. Cross entropy based loss functions are most popularly used for model training and validating segmentation and classification tasks. Following this, binary cross entropy and categorical cross entropy functions are employed for training the model for binary tumor mask generation and classification respectively. Binary cross entropy (BCE,



**Fig. 3.** A slice of MRI scan with T1 modality showing different tumor classes: meningioma, glioma, and pituitary

shown in Eq. 1) is a sigmoid activation [16] followed by cross entropy loss [29] that compares each of the predicted probabilities to actual output. Categorical cross entropy (CE, shown in Eq. 2) is a softmax activation function followed by cross-entropy-loss that compares the output probability over each tumor class for each MRI image.

$$\mathcal{L}_{BCE} = -\frac{1}{N} \sum_{i=1}^N (y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - P(y_i))) \quad (1)$$

where  $y$  represents actual tumor mask,  $p(y)$  represents predicted tumor mask and  $N$  is the total number of images.

$$\mathcal{L}_{CE} = \sum_i^C t_i \log(f(s_i)) \quad (2)$$

where  $C$  is the no. of class,  $f(s_i)$  is the probability of occurrence of each class  $t_i$  represents 1 for true label and 0 for others.

For segmentation the most popular evaluation metrics are dice coefficient (shown in Eq. 3) and intersection-over-union (IoU / Jaccard index) (shown in Eq. 4), and hence are utilized to evaluate the trained MAG-Net model. TP defines correctly classified predictions FP defines wrongly classified, and FN defines missed objects of each voxel.

$$DiceCoefficient = \frac{2 * TP}{2 * TP + FP + FN} \quad (3)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (4)$$

To evaluate classification module of the MAG-Net, accuracy, precision, recall, f1-score and micro average metrics are considered for better quantification and visualization of the performance of the model. Precision of the class, as shown in Eq. 5, quantifies about the positive prediction accuracy of the model. Recall is the fraction of true positive which are classified correctly (shown in Eq. 6). F1-score quantifies the amount of correct predictions out of all the positive predictions (shown in Eq. 7). Support quantifies the true occurrence in the specified dataset of the respective class. Micro average ( $\mu_{avg}$ ) (shown in Eq. 8, Eq. 9 and Eq. 10) is

calculated for precision, recall, and F1-score. To compute micro average ( $\mu_{avg}$ ), the test dataset is divided into two sub dataset, on each of which the true positive, false positive and false negative predictions are identified.

$$Precision = \frac{TP}{(TP + FP)} \quad (5)$$

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

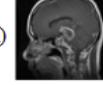
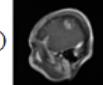
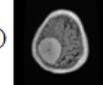
$$F1 - score = \frac{2 * Recall * Precision}{(Recall + Precision)} \quad (7)$$

$$\mu_{avg}(Precision) = \frac{TP_1 + TP_2}{(TP_1 + TP_2 + FP_1 + FP_2)} \quad (8)$$

$$\mu_{avg}(Recall) = \frac{TP_1 + TP_2}{(TP_1 + TP_2 + FN_1 + FN_2)} \quad (9)$$

$$\mu_{avg}(F1 - score) = HM(\mu_{avg}(Precision), \mu_{avg}(Recall)) \quad (10)$$

where  $TP_1$ ,  $FP_1$ , and  $FN_1$  belong to the first set and  $TP_2$ ,  $FP_2$ , and  $FN_2$  belongs to the different sets.  $HM$  is the harmonic mean.

		G.T	Predicted		
	Input	Mask	Label	Mask	Label
(a)			Meningioma		Meningioma
(b)			Glioma		Glioma
(c)			Pituitary		Pituitary

**Fig. 4.** Qualitative results of brain tumor segmentation and classification on MRI images (a, b, and c of different tumor classes) using MAG-Net model.

### 4.3 Results

The MAG-Net outputs the segmented mask of a given MRI image consisting of tumor region corresponding to meningioma, glioma, and pituitary as classified by the model. For randomly chosen MRI slices, Fig. 4 presents the segmentation and classification results of model. The visual representation confirms that the results are close to the ground truth of respective tumor classes.

**Table 1.** Comparative analysis of the MAG-Net with the existing segmentation models on test dataset.

Model	Accuracy	Loss	Dice coefficient	Jaccard index	Parameters
U-Net	99.5	0.024	0.70	0.55	31M
wU-Net	99.4	0.034	0.66	0.49	31M
Unet++	99.5	0.028	0.65	0.49	35M
MAG-Net	<b>99.52</b>	<b>0.021</b>	<b>0.74</b>	<b>0.60</b>	<b>5.4M</b>

\*bold quantities indicate the best results.

Table 1 represents the result of the proposed work for segmentation in the form of accuracy, loss, dice coefficient, Jaccard index, and trainable parameters along with comparative analysis with other popular approaches. The proposed framework outperforms the other approaches in segmenting tumor with the dice and IoU score of 0.74 and 0.60 respectively. In contrast to other models, MAG-Net achieved best results with minimal trainable parameters. The other popular approaches taken in comparative analysis for segmentation are U-Net [15], U-Net++ [18, 31], and wU-Net. [5].

**Table 2.** Comparative analysis of the MAG-Net with the existing classification models on test dataset using confusion matrix.

Model	Acc.	Loss		Meningioma	Glioma	Pituitary
VGG16	93.15	0.26	Meningioma	114	25	3
			Glioma	13	271	1
			Pituitary	0	1	185
VGG19	93.8	0.25	Meningioma	114	21	7
			Glioma	11	274	0
			Pituitary	0	1	185
ResNet50	94.2	0.31	Meningioma	123	12	7
			Glioma	16	266	3
			Pituitary	1	0	185
MAG-Net	<b>98.04</b>	<b>0.11</b>	Meningioma	<b>134</b>	<b>7</b>	<b>1</b>
			Glioma	<b>1</b>	<b>282</b>	<b>2</b>
			Pituitary	<b>1</b>	<b>0</b>	<b>185</b>

\*bold quantities indicate the best results.

Table 2 and Table 3 represent the results of the proposed work for classification in the form of accuracy, loss, confusion matrix, and classification report for meningioma, glioma, and pituitary tumor along with comparative analysis with other state-of-the-art approaches: VGG-16 [28], VGG-19 [28], and ResNet50 [25]. With exhaustive experimental trials it is observed that MAG-Net outperformed the existing approaches with significant margin in all the metrics.

**Table 3.** Comparative analysis of the MAG-Net with the existing classification models on test dataset considering classification report as evaluation parameter.

Model	Classes	Precision	Recall	f1-Score	Support
VGG16	Meningioma	0.90	0.80	0.85	142
	Glioma	0.91	0.85	0.93	285
	Pituitary	0.98	0.99	0.99	186
	Micro avg.	0.93	0.93	0.93	613
VGG19	Meningioma	0.91	0.80	0.85	142
	Glioma	0.93	0.96	0.94	285
	Pituitary	0.96	0.99	0.98	186
	Micro avg.	0.93	0.93	0.93	613
ResNet-50	Meningioma	0.88	0.87	0.87	142
	Glioma	0.93	0.99	0.94	285
	Pituitary	0.95	0.99	0.97	186
	Micro avg.	0.94	0.94	0.94	613
MAG-Net	Meningioma	<b>0.99</b>	<b>0.94</b>	<b>0.96</b>	142
	Glioma	<b>0.98</b>	<b>0.99</b>	<b>0.98</b>	285
	Pituitary	<b>0.98</b>	<b>0.99</b>	<b>0.99</b>	186
	Micro avg.	<b>0.98</b>	<b>0.98</b>	<b>0.98</b>	613

\*bold quantities indicate the best results.

It is observed that unlike other state-of-the-art models, the MAG-Net model achieved promising results due to the reduction in the overall computation, better feature extraction and training parameters optimization. As shown in Table 1 raw U-Net displayed similar performance but at the cost of large number of trainable parameters. In the MAG-Net model, the encoder block is developed by replacing convolution layers with parallel depthwise separable convolution of various sizes connected in parallel which resulted in better multi-scale feature learning for varying shapes and sizes of the tumor. For reducing spatial loss during feature reconstruction, attention mechanism is used in skip connections for better feature reconstruction. To reduce the overall complexity of the model the feature extracted by encoder blocks are reused to classify the type of brain tumor.

## 5 Conclusion

In this paper, the complex task of brain tumor segmentation and classification is addressed using multi-task attention guided network (MAG-Net). This a

U-Net based model that features reduction in the overall computation, better feature extraction and training parameters optimization. The proposed architecture achieved significant performance on the Figshare brain tumor dataset by exploiting the state-of-the-art advantages of U-Net, depthwise separable convolution and attention mechanism. The MAG-Net model recorded the best classification and segmentation results compared to the existing classification and segmentation approaches. It is believed that this work can also be extended to other domains involving classification and segmentation tasks.

## Acknowledgment

We thank our institute, Indian Institute of Information Technology Allahabad (IIITA), India and Big Data Analytics (BDA) lab for allocating the centralised computing facility and other necessary resources to perform this research. We extend our thanks to our colleagues for their valuable guidance and suggestions.

## References

1. Afshar, P., Mohammadi, A., Plataniotis, K.N.: Brain tumor type classification via capsule networks. In: 2018 25th IEEE International Conference on Image Processing (ICIP). pp. 3129–3133. IEEE (2018)
2. Albawi, S., Mohammed, T.A., Al-Zawi, S.: Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). pp. 1–6. Ieee (2017)
3. Brownlee, J.: Use early stopping to halt the training of neural networks at the right time. <https://machinelearningmastery.com/how-to-stop-training-deep-neural-networks-at-the-right-time-using-early-stopping/> (2018), [Online; accessed April 17, 2021]
4. Cancer.Net: Brain tumor: Diagnosis. <https://www.cancer.net/cancer-types/brain-tumor/diagnosis> (2020), [Online; accessed March 20, 2021]
5. CarryHJR: Nested unet. <https://github.com/CarryHJR/Nested-UNet/blob/master/model.py>. (2020), [Online; accessed March 11, 2021]
6. Cheng, J.: brain tumor dataset (4 2017), [https://figshare.com/articles/dataset/brain\\_tumor\\_dataset/1512427](https://figshare.com/articles/dataset/brain_tumor_dataset/1512427)
7. Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z., Wang, Z., Feng, Q.: Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PloS one* **10**(10), e0140381 (2015)
8. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1251–1258 (2017)
9. Deepak, S., Ameer, P.: Brain tumor classification using deep cnn features via transfer learning. *Computers in biology and medicine* **111**, 103345 (2019)
10. Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C.: The importance of skip connections in biomedical image segmentation. In: Deep learning and data labeling for medical applications, pp. 179–187. Springer (2016)

11. Díaz-Pernas, F.J., Martínez-Zarzuela, M., Antón-Rodríguez, M., González-Ortega, D.: A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. *Healthcare* **9**(2), 153 (2021). <https://doi.org/10.3390/healthcare9020153>, <https://app.dimensions.ai/details/publication/pub.1135094000andhttps://www.mdpi.com/2227-9032/9/2/153/pdf>
12. Hinton, G.E., Sabour, S., Frosst, N.: Matrix capsules with em routing. In: International conference on learning representations (2018)
13. İşin, A., Direkoglu, C., Şah, M.: Review of mri-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science* **102**, 317–324 (2016)
14. Ismael, M.R., Abdel-Qader, I.: Brain tumor classification via statistical features and back-propagation neural network. In: 2018 IEEE international conference on electro/information technology (EIT). pp. 0252–0257. IEEE (2018)
15. Jain, A.: brain tumor segmentation u-net. <https://github.com/adityajn105/brain-tumor-segmentation-unet> (2020), [Online; accessed January 08, 2021]
16. Jamel, T.M., Khammas, B.M.: Implementation of a sigmoid activation function for neural network using fpga. In: 13th Scientific Conference of Al-Ma'moon University College. vol. 13 (2012)
17. Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., Terzopoulos, D.: Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021)
18. MrGiovanni: U-net++ keras. <https://github.com/MrGiovanni/UNetPlusPlus> (2020), [Online; accessed March 12, 2021]
19. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D.: Attention u-net: Learning where to look for the pancreas (2018)
20. Pashaei, A., Sajedi, H., Jazayeri, N.: Brain tumor classification via convolutional neural network and extreme learning machines. In: 2018 8th International conference on computer and knowledge engineering (ICCKE). pp. 314–319. IEEE (2018)
21. Punn, N.S., Agarwal, S.: Chs-net: A deep learning approach for hierarchical segmentation of covid-19 infected ct images. *arXiv preprint arXiv:2012.07079* (2020)
22. Punn, N.S., Agarwal, S.: Inception u-net architecture for semantic segmentation to identify nuclei in microscopy cell images. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **16**(1), 1–15 (2020)
23. Punn, N.S., Agarwal, S.: Multi-modality encoded fusion with 3d inception u-net and decoder model for brain tumor segmentation. *Multimedia Tools and Applications* pp. 1–16 (2020)
24. Punn, N.S., Agarwal, S.: Modality specific u-net variants for biomedical image segmentation: A survey. *arXiv preprint arXiv:2107.04537* (2021)
25. raghakot: keras-resnet. <https://github.com/raghakot/keras-resnet> (2017), [Online; accessed March 18, 2021]
26. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
27. Ruder, S.: An overview of gradient descent optimization algorithms (2017)
28. Thakur, R.: step by step vgg16 implementation in keras for beginners. <https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c> (2019), [Online; accessed March 20, 2021]
29. Zhang, Z., Sabuncu, M.R.: Generalized cross entropy loss for training deep neural networks with noisy labels. *arXiv preprint arXiv:1805.07836* (2018)

30. Zhou, T., Ruan, S., Canu, S.: A review: Deep learning for medical image segmentation using multi-modality fusion. *Array* **3**, 100004 (2019)
31. Zhou, Z., Siddiquee, M., Tajbakhsh, N., Liang, J.U.: A nested u-net architecture for medical image segmentation. arXiv preprint arXiv:1807.10165 (2018)