This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

CASA-Net: Confidence-Aware Spatial Attention Network for Efficient Detection-Guided Brain Tumor Segmentation

Somayeh Davar and Thomas Fevens

Department of Computer Science and Software Engineering,
Concordia University, Montreal, QC, Canada
{somayeh.davar,thomas.fevens}@concordia.ca

Abstract. In this paper we propose *CASA-Net*, a two-stage detection-guided framework for brain tumor segmentation. The proposed framework employs YOLOv11 for rapid tumor localization with an Attention U-Net for fine-grained segmentation. Unlike conventional approaches that rely on binary bounding-box attention masks, *CASA-Net* introduces a novel Confidence-Aware Spatial Attention (CASA) module, which transforms the detector confidence into a continuous spatial prior that guides the segmentation network. The detector first predicts a tumor bounding box and its associated confidence score, which are used to generate a confidence-aware attention map within an expanded region of interest (ROI). The resulting attention-guided ROI is then processed by the Attention U-Net, and the predicted segmentation mask is projected back to the original image space. Experimental results demonstrate that *CASA-Net* improves segmentation performance over existing detection-guided baselines while retaining the ROI-based design of the framework.

Keywords: Brain Tumor Segmentation · Attention U-Net · YOLOv11· Deep Learning · Medical Imaging · Detection-Guided Segmentation

1 Introduction

Brain tumors are among the most challenging neurological disorders, requiring accurate and timely diagnosis to support treatment planning, surgical intervention, and disease monitoring. Magnetic resonance imaging (MRI) plays a central role in brain tumor assessment because of its excellent soft-tissue contrast and high spatial resolution. However, accurate tumor segmentation remains difficult because of substantial variations in tumor size, morphology, location, intensity distribution, and boundary definition across patients [19,15,9]. Deep learning (DL) has advanced automated brain tumor analysis, enabling increasingly accurate localization and segmentation from MRI. However, precise tumor delineation remains challenging due to variations in tumor size, shape, location, and appearance. When applied to full MRI slices, segmentation networks must simultaneously localize the tumor and delineate its boundaries, even though tumors often occupy only a small portion of the image. Consequently, considerable computational capacity is devoted to anatomically irrelevant background regions.

Detection-guided segmentation addresses this limitation by first localizing the lesion and then performing fine-grained segmentation within a tumor-centered ROI, allowing the segmentation network to focus on the relevant anatomical region. Among object detectors, YOLO-based architectures provide an attractive balance between localization accuracy and inference efficiency [1,7]. A previous detection-guided approach employed YOLO-based localization followed by Attention U-Net segmentation using a binary attention mask derived from the detected bounding box [4]. Although such strategies effectively constrain segmentation to localized ROIs, binary spatial guidance treats all pixels within the detected region equally and does not explicitly incorporate detector confidence into the segmentation process. Consequently, the confidence information associated with the detector prediction was not explicitly exploited. To address this limitation, we propose *CASA-Net*, a confidence-aware detection-guided framework for brain tumor segmentation. The proposed framework introduces a novel *Confidence-Aware Spatial Attention (CASA)* module that transforms detector confidence into a continuous spatial prior, providing richer guidance than conventional binary masks. By combining YOLOv11 localization, confidence-aware spatial attention, and Attention U-Net segmentation within an ROI-based pipeline, CASA-Net establishes a more informative interaction between localization and segmentation while preserving the efficiency of the overall framework. The main contributions of this work are as follows:

(1) We propose the *CASA-Net* framework, which integrates YOLOv11 localization, the proposed Confidence-Aware Spatial Attention (CASA) module, and Attention U-Net segmentation within an efficient ROI-based pipeline.
(2) We introduce the proposed CASA module, which converts the detector confidence score into a continuous spatial prior, enabling richer information transfer from localization to segmentation than conventional ROI guidance.
(3) Experimental results on Figshare dataset demonstrate that CASA-Net improves segmentation performance over the previous detection-guided method.

2 Related Work

2.1 Deep Learning for Brain Tumor Segmentation

Deep learning has substantially improved brain tumor segmentation, with encoder–decoder architectures such as U-Net becoming the dominant paradigm because they effectively combine contextual information with precise spatial localization [16]. Numerous extensions, including residual, attention-based, and self-configuring architectures such as nnU-Net, have further improved segmentation accuracy and robustness [6,12,24,10]. Among these approaches, Attention U-Net is particularly well suited for brain MRI because its attention gates emphasize tumor regions while suppressing irrelevant background features, making it an effective backbone for ROI-based segmentation [14].

2.2 Detection-Guided Segmentation

Detection-guided segmentation improves computational efficiency by separating lesion localization from boundary delineation. Rather than processing the entire MRI slice, an object detector first localizes the tumor, after which segmentation is performed within a localized ROI. YOLO-based detectors are particularly attractive because they provide an effective balance between localization accuracy and inference efficiency, making them well suited for medical image analysis [7].

2.3 Attention Mechanisms

Attention mechanisms improve segmentation by emphasizing informative features while suppressing irrelevant responses. Architectures such as Attention U-Net and the Convolutional Block Attention Module (CBAM) learn channel and spatial attention directly from feature representations, enabling more discriminative feature learning [14,22]. However, these approaches do not explicitly incorporate prior spatial information obtained from an upstream lesion detector, limiting the interaction between localization and segmentation.

2.4 Research Gap

Although detection-guided segmentation enables the segmentation network to focus on a localized tumor region, the information transferred from the detection stage is typically limited to the predicted ROI or a binary spatial mask [3,4]. As a result, the detector confidence associated with the predicted bounding box is not explicitly utilized during segmentation. This limitation motivates the proposed *Confidence-Aware Spatial Attention (CASA)* module, which incorporates the detector confidence into a continuous spatial prior, enabling richer communication between the detection and segmentation stages.

3 Methodology

3.1 Dataset and Data Pre-processing

Dataset The proposed CASA-Net framework is evaluated on the publicly available Figshare brain MRI dataset [8], which is widely used as a benchmark for brain tumor analysis and has been extensively adopted for both classification and segmentation studies [8,21,13,2,20]. The dataset contains 3,064 contrast-enhanced T1-weighted (CE-MRI) slices acquired from 233 patients, including 708 meningioma, 1,426 glioma, and 930 pituitary tumor images. MRI scans are provided in axial, coronal, and sagittal views with a resolution of 512×512 pixels and predefined five-fold cross-validation splits. Each MRI slice is associated with a tumor label and a corresponding pixel-level tumor mask used as the segmentation ground truth. Ground-truth bounding boxes are generated from the tumor masks to train the YOLOv11 detector. The input MRI image, ground-truth segmentation mask, and corresponding bounding box are denoted as:

$$I \in \mathbb{R}^{H \times W}, \qquad Y \in \{0,1\}^{H \times W}, \qquad b = (x_{\min}, y_{\min}, x_{\max}, y_{\max}),$$

Data Pre-processing Prior to training, MRI images undergo a series of preprocessing steps to improve data consistency and facilitate robust feature learning. Each MRI image is resized from 512×512 to 256×256 pixels, reducing computational cost while preserving sufficient anatomical detail. Image intensities are subsequently standardized using z-score normalization to mitigate inter-image intensity variations, followed by Contrast Limited Adaptive Histogram Equalization (CLAHE) [11] to enhance local contrast and improve tumor visibility. During training, data augmentation is performed using random horizontal and vertical flipping, rotation, and translation to increase sample variability, reduce overfitting, and improve model generalization.

3.2 Proposed CASA-Net Architecture

The proposed *CASA-Net* framework integrates object detection and semantic segmentation within a two-stage detection-guided architecture for brain tumor segmentation. As illustrated in Fig. 1, *CASA-Net* combines YOLOv11 for tumor localization with an Attention U-Net for fine-grained tumor segmentation. The overall framework consists of five sequential stages. First, YOLOv11 predicts a tumor bounding box together with its associated confidence score from the input MRI. The detected bounding box is then expanded and cropped to generate a tumor-centered ROI. Next, the proposed CASA module uses the predicted bounding box and its associated confidence score to generate a confidence-aware spatial prior, which is fused with the cropped ROI before being provided to the Attention U-Net. The Attention U-Net subsequently performs fine-grained segmentation. Finally, the predicted ROI mask is mapped back to its corresponding location in the original image space to obtain the final segmentation mask.

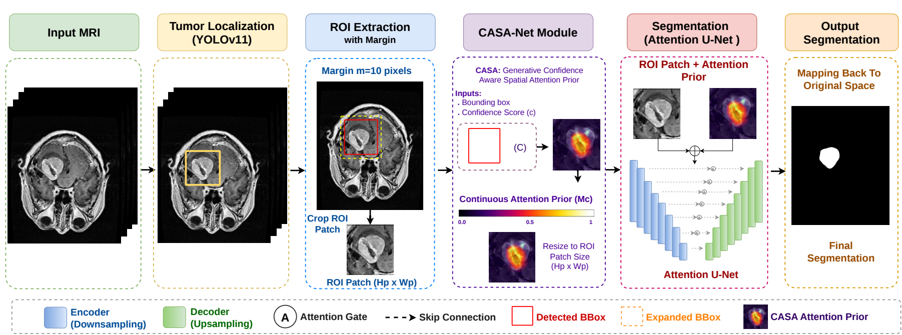

Fig. 1. Overview of CASA-Net. YOLOv11 localizes the tumor, CASA module generates a confidence-aware spatial prior, and Attention U-Net performs segmentation.

Stage 1: Tumor Localization The first stage of *CASA-Net* performs coarse tumor localization using YOLOv11, which provides an efficient one-stage detection framework for identifying the approximate tumor location within the input MRI image. The detector is trained using bounding-box annotations generated from the ground-truth tumor masks. During inference, YOLOv11 predicts both the tumor bounding box and an associated confidence score, providing the spatial information required for the subsequent ROI extraction and confidence-aware segmentation stages. Given an input MRI image I, the detector predicts a set of candidate bounding boxes together with their corresponding confidence scores,

$$\mathcal{B} = \{(b_i, s_i)\}_{i=1}^{N}, \tag{1}$$

where b_i denotes the i-th predicted bounding box and s_i represents its corresponding confidence score. After applying confidence thresholding and non-maximum suppression (NMS), the highest-confidence detection is selected as the tumor localization. The current framework therefore processes a single tumor ROI per slice;

$$\hat{b} = \arg\max_{b_i \in \mathcal{B}} s_i. \tag{2}$$

To preserve the surrounding anatomical context, the selected bounding box is expanded by a fixed margin $m = 10$ pixels, resulting in an enlarged ROI defined

$$\hat{b}^{+} = (x_{\min} - m,\ y_{\min} - m,\ x_{\max} + m,\ y_{\max} + m), \tag{3}$$

where the expanded coordinates are clipped to remain within the image boundaries. The resulting ROI is subsequently forwarded to the next stage.

Stage 2: ROI Patch Extraction The expanded predicted bounding box $\hat{b}^{+}$ obtained from the localization stage is used to extract a tumor-centered ROI from the input MRI. The same ROI coordinates are applied to the ground-truth segmentation mask Y to obtain the target ROI mask Y_r for training and evaluation. The cropped ROI and target mask are resized to 256×256, corresponding to the segmentation network input resolution;

$$I_r = \mathrm{Resize}\Big(\mathrm{Crop}(I, \hat{b}^{+})\Big), \qquad Y_r = \mathrm{Resize}\Big(\mathrm{Crop}(Y, \hat{b}^{+})\Big). \tag{4}$$

Here, I_r serves as the input to the CASA-guided segmentation stage, whereas Y_r is used only as the ground-truth target for training and quantitative evaluation. During inference, ROI extraction and CASA prior generation rely on the bounding box $\hat{b}$ and confidence score $\hat{s}$ predicted by YOLOv11; no ground-truth localization or segmentation information is used.

Stage 3: Confidence-Aware Spatial Attention (CASA) To incorporate detector confidence into the segmentation process, we introduce the *Confidence-Aware Spatial Attention (CASA)* module. Conventional detection-guided approaches typically use the detected bounding box to define a binary spatial prior, providing localization information without explicitly accounting for the confidence associated with the detection. In contrast, CASA incorporates the detector confidence into a continuous spatial prior, providing confidence-aware spatial guidance to the segmentation network. Following ROI extraction, YOLOv11 provides the predicted bounding box $\hat{b}$ and its associated confidence score $c \in [0, 1]$. The predicted bounding-box coordinates are mapped from the original image space to the corresponding ROI coordinate system. Let (x_c, y_c) denote the center of the detected bounding box within the ROI. CASA constructs a confidence-aware spatial prior centered at (x_c, y_c), assigning higher attention weights to locations near the detected region and progressively decreasing the weights with increasing distance from the bounding-box center.

$$A_{\mathrm{CASA}}(x,y) = c \exp\left(-\frac{(x-x_c)^2 + (y-y_c)^2}{2\sigma^2}\right), \tag{5}$$

where $c \in [0,1]$ denotes the YOLOv11 detection confidence and $\sigma > 0$ controls the spatial spread of the Gaussian prior. The confidence score c is used as a detector-derived confidence cue rather than as a calibrated measure of uncertainty. The resulting spatial prior A_{CASA} is resized to match the spatial dimensions of the ROI patch I_r and concatenated with the corresponding image representation,

$$X = [I_r,\ A_{\mathrm{CASA}}], \tag{6}$$

Stage 4: Attention U-Net Segmentation The confidence-aware ROI representation generated by the CASA module is processed by an Attention U-Net to obtain the ROI segmentation, $\hat{Y}_r = f_\theta(X)$, where $X = [I_r, A_{\mathrm{CASA}}]$, f_θ denotes the Attention U-Net parameterized by θ, and $\hat{Y}_r$ is the predicted ROI segmentation. Attention U-Net is selected because its attention-gated skip connections suppress irrelevant activations while emphasizing task-relevant features [14]. These learned attention mechanisms operate in conjunction with the detector-derived spatial prior, allowing the segmentation to focus on the localized region. As a two-stage framework, CASA-Net remains dependent on the accuracy of the localization stage; inaccurate or missed detections may therefore propagate to the subsequent segmentation stage.

Stage 5: Mapping Back to Original Image Space The predicted ROI segmentation is resized to match the spatial dimensions of the expanded bounding box and mapped back to its corresponding location in the preprocessed MRI image space, producing the final segmentation mask $\hat{Y} \in \mathbb{R}^{H \times W}$.

3.3 Training Details and Implementation

The proposed CASA-Net framework is implemented in Python 3.11 using TensorFlow and Keras. Performance is evaluated using five-fold cross-validation. To prevent data leakage, the data is partitioned at the patient level, ensuring that MRI slices do not appear in both the training and testing sets. In each fold, approximately 80% of the data is used for training and 20% for testing, with a portion of the training set reserved for validation. The best-performing model is selected according to the validation loss. The localization and segmentation networks are trained independently in a two-stage manner. During segmentation, the network operates on the extracted ROI with its corresponding CASA spatial priors. The network is trained for 100 epochs with a batch size of 16 using stochastic gradient descent (SGD) with an initial learning rate of 10^{-3} and a momentum of 0.9. The Gaussian spread parameter was fixed to $\sigma = 10$ pixels. Training employs a hybrid Dice and Binary Cross-Entropy (BCE) loss, $\mathcal{L}_{seg} = \lambda_1 \mathcal{L}_{Dice} + \lambda_2 \mathcal{L}_{BCE}$, where $\lambda_1 = \lambda_2 = 0.5$.

4 Experimental Results

4.1 Performance Evaluation

The effectiveness of the proposed CASA-Net framework is evaluated on the publicly available Figshare dataset [8] using the predefined five-fold cross-validation. Performance is assessed using the Dice Similarity Coefficient (DSC) and Jaccard Index (IoU). CASA-Net obtained an average Dice score of 92.51% and IoU of 86.08%. Table 1 summarizes previously reported results for reference. The proposed method achieves Dice scores of 92.89%, 90.98%, and 93.66% for glioma, meningioma, and pituitary tumors, respectively, demonstrating consistent performance across different tumor types. The lower performance for meningioma may reflect its greater morphological variability, whereas pituitary tumors generally exhibit more compact and well-defined boundaries.

Table 1. Comparison of the proposed CASA-Net with representative methods.

Method	**Dice Score (%)**				**Jaccard Index (%)**			
	Glioma	Meningioma	Pituitary	Avg.	Glioma	Meningioma	Pituitary	Avg.
ResU-Net [23]	86.27	88.16	87.15	87.19	74.19	75.32	74.22	74.58
VGG19 U-Net [12]	80.33	–	86.00	83.16	–	–	–	–
Sheela et al. [18]	59.00	78.00	49.00	63.10	–	–	–	–
El Shafai et al. [17]	–	–	–	35.18	–	–	–	21.34
Edge U-Net [10]	91.76	88.80	87.28	89.28	84.78	79.85	77.43	80.69
Multiscale-CNN [5]	77.90	89.40	81.30	82.87	–	–	–	–
Proposed *CASA-Net*	92.89	90.98	93.66	92.51	86.72	83.45	88.07	86.08

4.2 Ablation Study

To isolate the CASA contribution, we compare spatial guidance strategies using identical five-fold splits and training settings; only the spatial prior is changed. Results are reported in Table 2 as mean $\pm$ standard deviation across the five cross-validation folds. CASA-Net achieved the highest mean performance, with a Dice score of $92.51 \pm 0.21\%$ and an IoU of $86.08 \pm 0.18\%$, supporting the benefit of incorporating detector confidence into the spatial prior.

Table 2. Ablation of spatial guidance strategies under identical experimental settings.

Spatial Guidance	**Dice Score (%)**	**Jaccard Index (%)**
Binary BBox Prior	92.27 ± 0.14	85.66 ± 0.08
Gaussian Prior ($c = 1$)	92.30 ± 0.07	85.90 ± 0.12
CASA-Net Prior	92.51 ± 0.21	86.08 ± 0.18

4.3 Comparison with Existing Methods

The performance of CASA-Net is compared with representative methods reported on the Figshare dataset, including ResU-Net [23], VGG19 U-Net [12], Edge U-Net [10], and Multiscale-CNN [5]. It should be noted that direct comparison should be interpreted with caution, as differences in preprocessing, data partitioning, training protocols, and evaluation settings may influence the reported performance. While these approaches have demonstrated strong performance, they primarily perform segmentation directly on full MRI images or rely solely on image-derived feature representations without explicitly incorporating detector-guided spatial priors into the segmentation process. In contrast to these image-based segmentation approaches, CASA-Net uses detector confidence as a continuous spatial prior to guide ROI-based segmentation and provides complementary localization information, enabling the network to more effectively delineate tumor boundaries across diverse anatomical structures. Table 1 is provided for contextual reference. The controlled contribution of this spatial guidance is evaluated separately through the matched ablation study in Table 2. Figure 2 illustrates qualitative results obtained using the proposed CASA-Net framework. The predicted masks closely correspond to the ground-truth annotations, demonstrating the effectiveness of the proposed method.

(a)
(b)
(C)

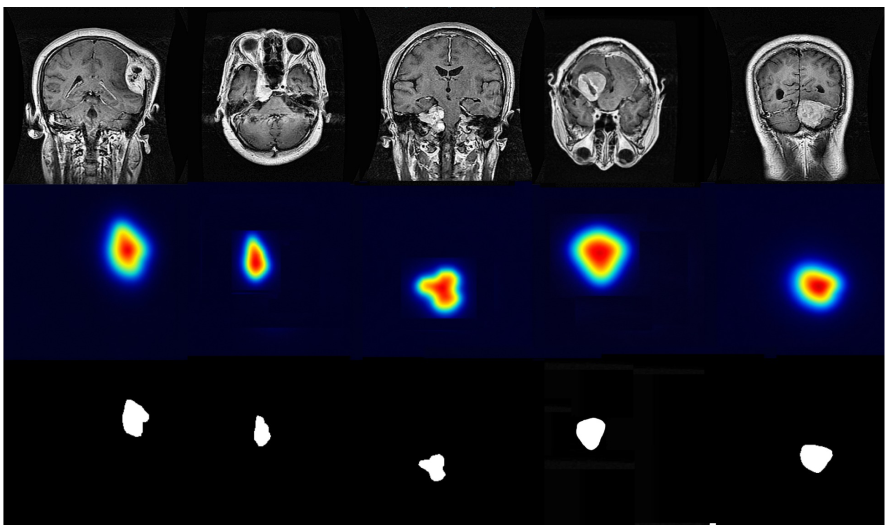

Fig. 2. Qualitative results obtained by the proposed CASA-Net framework. From left to right: (a) Input CE-MRI, (b) CASA Attention, and (c) Prediction.

5 Conclusions

This study introduces *CASA-Net*, a detection-guided framework that integrates YOLOv11 with Attention U-Net through the proposed *Confidence-Aware Spatial Attention (CASA)* module. Unlike conventional binary spatial guidance, CASA incorporates detector confidence into a continuous Gaussian spatial prior to guide tumor segmentation. Ablation experiments demonstrate that CASA improves segmentation over both binary bounding-box and fixed-confidence Gaussian guidance. The framework achieves consistent performance across glioma, meningioma, and pituitary tumors, supporting detector confidence as a useful source of spatial guidance. Future work will investigate adaptive priors in which Gaussian spread and anisotropy are conditioned on detection confidence and bounding-box geometry, as well as multimodal MRI and external validation.

References

1. Adarsh, P., Rathi, P., Kumar, M.: Yolov3-tiny:object detection recognition using one stage improved model. In: ICACCS. pp. 687–694 (2020). https://doi.org/10.1109/ICACCS48705.2020.9074315
2. Badza, M.M., Barjaktarovic, M.C.: Classification of brain tumors from mri images using a convolutional neural network. Applied Sciences **10**(6) (2020). https://doi.org/10.3390/app10061999, https://www.mdpi.com/2076-3417/10/6/1999
3. Davar, S., Fevens, T.: Enhanced U-Net for brain tumor localization & segmentation in T1-weighted MRI. IEEE Transactions on Circuits & Systems II: Express Briefs **72**(8), 993–997 (2025). https://doi.org/10.1109/TCSII.2025.3556846
4. Davar, S., Fevens, T.: Hybrid deep-learning approach for brain tumor analysis in MRI: from detection to delineation. In: Gimi, B.S., Krol, A. (eds.) Medical Imaging 2026: Clinical and Biomedical Imaging. vol. 13929, p. 139292P. SPIE (2026). https://doi.org/10.1117/12.3087813, https://doi.org/10.1117/12.3087813
5. 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) (2021). https://doi.org/10.3390/healthcare9020153
6. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. In: Journal Nature Methods. vol. 18, pp. 203–211. Nature Publishing Group (2021). https://doi.org/10.1038/s41592-020-01008-z
7. Jocher, G., Qiu, J.: Ultralytics yolo11 (2024), https://github.com/ultralytics
8. Jun, C.: Brain tumor dataset,3064 T1-weighted contrast-enhanced images with three kinds of tumor (2017), https://doi.org/10.6084/m9.figshare.1512427.v5
9. Liu, Z., Tong, L., Chen, L., Jiang, Z., Zhou, F., Zhang, Q., Zhang, X., Jin, Y., Zhou, H.: Deep learning based brain tumor segmentation: a survey. Complex and Intelligent Systems **9**, 1001–1026 (2023). https://doi.org/10.1007/s40747-022-00815-5
10. M. Gab Allah, A., M. Sarhan, A., et al.: Edge U-Net: Brain tumor segmentation using MRI based on deep U-Net model with boundary information. Expert Systems Applications **213**, 118833 (2023). https://doi.org/https://doi.org/10.1016/j.eswa.2022.118833
11. Mgbejime, G.T., Hossin, M.A., Nneji, G.U., Monday, H.N., Ekong, F.: Parallelistic convolution neural network approach for brain tumor diagnosis. Diagnostics **12**(10) (2022). https://doi.org/10.3390/diagnostics12102484
12. Nawaz, A., Akram, U., Salam, A.A., Ali, A.R., Ur Rehman, A., Zeb, J.: VGG-UNET for brain tumor segmentation and ensemble model for survival prediction. 2021 International Conference on Robotics and Automation in Industry (ICRAI) pp. 1–6 (2021). https://doi.org/10.1109/ICRAI54018.2021.9651367
13. Noreen, N., Palaniappan, S., Qayyum, A., Ahmad, I., Imran, M., Shoaib, M.: A deep learning model based on concatenation approach for the diagnosis of brain tumor. IEEE Access pp. 55135–55144 (2020). https://doi.org/10.1109/ACCESS.2020.2978629
14. 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. arXiv **abs/1804.03999** (2018), https://api.semanticscholar.org/CorpusID:4861068

15. Ostrom, Q.T., et al.: Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the united states. Neuro-Oncology **25**(Supplement_5), v1–v99 (2023)
16. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015. pp. 234–241. Springer International Publishing (2015)
17. Shafai, W.E., Mahmoud, A.A., El-Rabaie, E.S.M., et al.: Hybrid segmentation approach for different medical image modalities. Computers, Materials and Continua **73**(2), 3455–3472 (2022). https://doi.org/10.32604/cmc.2022.028722
18. Sheela, C.J.J., Suganthi, G.: Automatic brain tumor segmentation from MRI using greedy snake model and fuzzy c-means optimization. Journal of King Saud University - Computer and Information Sciences **34**(3), 557–566 (2022). https://doi.org/https://doi.org/10.1016/j.jksuci.2019.04.006
19. Siegel, R.L., Miller, K.D., Fuchs, H.E., Jemal, A.: Cancer statistics. CA: A Cancer Journal for Clinicians **74**(1), 17–48 (2024)
20. Sumit Tripathi, A.V., Sharma, N.: Automatic segmentation of brain tumour in MR images using an enhanced deep learning approach. Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization **9**(2), 121–130 (2021). https://doi.org/10.1080/21681163.2020.1818628
21. Swati, Z.N.K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., Lu, J.: Content-based brain tumor retrieval for MR images using transfer learning. IEEE Access **7**, 17809–17822 (2019)
22. Woo, S., Park, J., Lee, J.Y., Kweon, I.S.: Cbam: Convolutional block attention module. In: ECCV (2018)
23. Zhang, Z., et al: Road extraction by deep residual u-net. IEEE Geoscience& Remote Sensing Letters **15**(5), 749–753 (2018). https://doi.org/10.1109/LGRS.2018.2802944
24. Zhao, L., Ma, J., Shao, Y., Jia, C., Zhao, J., Yuan, H.: MM-UNet: A multimodality brain tumor segmentation network in MRI images. Frontiers in Oncology **12** (2022). https://doi.org/10.3389/fonc.2022.950706