Intraoperative Glioma Segmentation with YOLO + SAM for Improved Accuracy in Tumor Resection

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

Gliomas, a common type of malignant brain tumor, present significant surgical challenges due to their similarity to healthy tissue. Preoperative Magnetic Resonance Imaging (MRI) images are often ineffective during surgery due to factors such as brain shift, which alters the position of brain structures and tumors. This makes real-time intraoperative MRI (ioMRI) crucial, as it provides updated imaging that accounts for these shifts, ensuring more accurate tumor localization and safer resections. This paper presents a deep learning pipeline combining You Only Look Once Version 8 (YOLOv8) and Segment Anything Model Vision Transformer-base (SAM ViT-b) to enhance glioma detection and segmentation during ioMRI. Our model was trained using the Brain Tumor Segmentation 2021 (BraTS 2021) dataset, which includes standard magnetic resonance imaging (MRI) images, and noise-augmented MRI images that simulate ioMRI images. Noised MRI images are harder for a deep learning pipeline to segment, but they are more representative of surgical conditions. Achieving a Dice Similarity Coefficient (DICE) score of 0.79, our model performs comparably to state-of-the-art segmentation models tested on noiseless data. This performance demonstrates the model's potential to assist surgeons in maximizing tumor resection and improving surgical outcomes.

1 Introduction

Gliomas are a common type of cancerous brain tumors that account for about 30% of all brain tumors and 80% of all malignant brain tumors [30]. Standard treatment modalities for gliomas include surgery, chemotherapy, and radiation therapy, with surgery often being the preferred option for most neurosurgeons [13]. The primary goal of surgery is to physically remove as much of the tumor as possible in a process known as resection [29], in which imaging technologies play a crucial role. Preoperative imaging, particularly MRI, is essential for diagnosing and planning the surgical approach.

However, the intraoperative success of glioma resection is frequently challenged by several factors. Brain shift, a phenomenon that occurs when the brain changes position during surgery, significantly hinders a surgeon's ability to accurately locate and resect the tumor [12, 17]. Another complication arises when gliomas infiltrate surrounding brain tissue, making it difficult to clearly delineate tumor margins[28, 27]. As a result, there runs a risk of leaving behind residual tumor cells, which can lead to the recurrence of the glioma or removing too much healthy tissue.

To address these challenges, neurosurgeons have adopted real-time imaging techniques using intra-operative magnetic resonance imaging (ioMRI), which has emerged as the preferred imaging tool for brain tumor operations [21, 11, 14, 25, 26]. ioMRI allows surgeons to update their view of the brain and tumor as the surgery progresses, compensating for brain shift and improving the accuracy of tumor resection. The interpretation of ioMRI images can be time-consuming due to the potential

for human error, which can prolong surgery and increase the risk of complications[3]. Moreover, the process of identifying tumor margins on ioMRI images is manually conducted and subject to human error and variability [23].

In this paper, we propos a pipeline that utilizes a YOLOv8 model for the detection of gliomas from ioMRI images, followed by the SAM model to refine the segmentation results, thereby ensuring higher accuracy and robustness. To evaluate the robustness of our model, we tested it on augmented MRI images that were simulated through the addition of Gaussian noise to MRI images. These augmented MRI images are similar to ioMRI images, which are generally noisier. Our model achieved a similar dice score to state-of-the-art tumor segmentation models and merits further exploration for use in improving glioma resection outcomes.

2 Methodology

2.1 Data Preprocessing

Our model was trained on the open access BraTS 2021 dataset, which is a collection of clinically acquired MR images of annotated glioma tumors from consenting patiends[4, 18, 7, 5, 6]. As the YOLO model can only process colorized images, an image processing function was developed to colorize the grayscale images. This was conducted by assigning an RGB value to each pixel based off its intensity. Another function was developed to create bounding boxes from the ground truth segmentation of the images. Following this function, all images and masks were resized to 256x256 pixels. Finally, the model was trained on both standard MR images and ioMR images that were synthesized through the addition of Gaussian noise. In order to accomplish a dataset of usable ioMR images, the signal-to-noise ratio (SNR) of the BraTS dataset was decreased to mimic ioMR images. ioMR images typically have an SNR of 25 under standard clinical conditions [10]. By decreasing the SNR and resolution, these modified images simulate the qualities of an ioMR image, demonstrated in Figure 1.
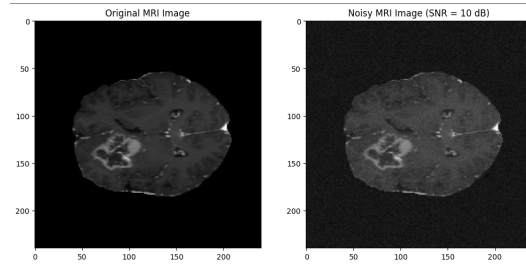

Figure 1: Left: regular MRI image. Right: augmented MRI image with SNR of 10

2.2 Architecture

The architecture of the model integrates two state-of-the-art algorithms, YOLO and SAM, to effectively detect and segment glioma tumors. The processed images are first fed into the YOLO model. The YOLO model then identifies the tumor and places a bounding box around it, additionally returning the middle coordinate of the bounding box. Following this general tumor detection, SAM is then used to precisely outline and segment the tumor based on the coordinates provided by YOLO. Figure 2 details this process.
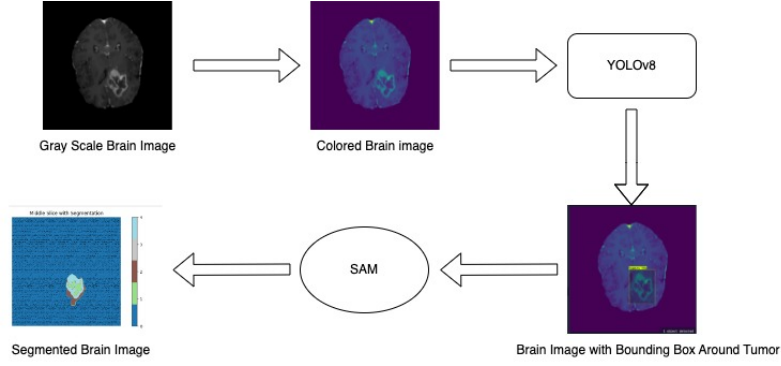
Figure 2: YOLO + SAM architecture; grayscale images are processed through an RGB assignment function, then passed through YOLO in which a bounding box is located around the tumor, the middle coordinate of the bounding box is passed into SAM, finally SAM produces a segmented brain image.

A pre-trained YOLOv8 model was chosen for our application to quickly and accurately detect the approximate location of tumors. YOLOv8 outperforms contemporary models and previous versions of the YOLO algorithm in speed and accuracy [16], making it highly suitable for real-time object detection tasks. Once the MRI images of the brain are passed into the YOLO model, it processes these images through a convolutional neural network (CNN), which extracts essential features and predicts bounding boxes around potential tumors. The YOLO model outputs the center coordinates of the predicted bounding box, which is then passed into SAM as a prompt.

The purpose of the SAM model in the pipeline is to refine the detection results provided by YOLOv8, ensuring that the tumors are accurately and precisely segmented for further analysis. The SAM ViT-b model was selected due to its lightweight nature, allowing for our model to be cost efficient while still maintaining high accuracy. Once the center coordinates of the YOLO bounding box are passed as a prompt into the SAM model, these inputs are used to perform precise segmentation, delineating the exact boundaries of the tumors. The SAM model then produces a detailed probability mask that delineates the tumor regions within the MRI images.

2.3 Training

The model was trained using the BraTS 2021 dataset using both standard MR images and the simulated ioMR images shown in Figure 1. In order to ensure consistency during training, middle slices from the axial plane (slices taken parallel to the X-axis) were extracted from the dataset by selecting the 78th slice of 155 from each image, thus converting the images from 3D to 2D. The middle slice is where the tumor is largest, making it the best choice for training. The specific MRI scan used was T1CE due to its tumor clarity within YOLO.

The YOLO model did not require any training on the BraTS dataset as it was already pre-trained on it. To fine-tune SAM, the middle coordinates of every bounding box in the training set, produced by YOLO, are fed as an initial prompt. SAM was trained on this data over 10 epochs. After the SAM model was finished being trained on the regular BraTS images, YOLO and SAM were then trained on the augmented version, or simulated ioMRI version, of the BraTS images.

3 Results

The proposed YOLO + SAM model was evaluated on an augmented version of the BraTS 2021 dataset. The model was evaluated using a Dice Similarity Coefficient (DICE) score, which is the similarity between two sets of data, in this case, predicted segmentation and ground truth, on a 0 -1 range with 1 indicating perfect overlap. The numerical value is calculated by 2 times the overlap area divided by the total area. The model achieved a DICE score of 0.79 on the augmented BraTS testing set for enhancing tumor (ET), which indicates a strong agreement between the predicted and ground truth segmentation. When compared to other state-of-the-art baseline models, YOLO + SAM has a comparable performance despite running on intentionally noised data. These models include

$E_1 D_3$ U-Net, Extended VAT method, and NVAUTO; created by Bukhari et al., Peiris et al., and Siddiquee et al. respectively, which were chosen as baselines as they are the state of the art trained on the BraTS 2021 dataset[9][19][20]. Their models achieved DICE scores of 0.826, 0.814, and 0.86 for ET. The inference times for these models are significantly higher, with estimates of 4 to 8 minutes, 3 to 6 minutes, and 45 to 90 seconds respectively, compared to 15 to 25 seconds for YOLO + SAM. This comparison shows the strong capability of the YOLO + SAM model as it achieved comparable performance to models that tested on images that were noiseless, while YOLO + SAM was tested on images that had extreme amounts of noise. The significantly lower inference time makes YOLO + SAM more suitable for real-world iMRI applications, providing faster and reliable results during surgery.

	Dice Score
$E_1 D_3$ U-Net	0.826
Extended VAT	0.814
NVAUTO	0.860
YOLO + SAM	**0.790**

(a) Comparison of model performances on the BraTS 2021 dataset. Note that, unlike the other models, YOLO+SAM achieved this score on *noised* data, demonstrating striking robustness.
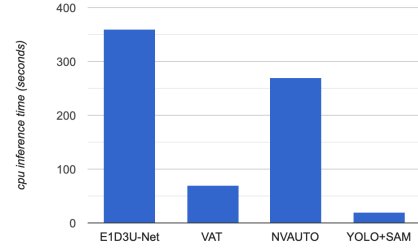
(b) Tumor Segmentation Model Inference Times. Note that, again, unlike the other models, YOLO+SAM achieved this score on *noised* data

Figure 3: DICE performance comparison (left) and inference times of various models (right).

4 Discussion

Physicians have used computed tomography (CT) scans, positron emission tomography (PET) scans, and MRI to detect and diagnose gliomas in patients [1]. Historically, machine learning applications for glioma imaging have focused on classification, diagnosis, and preoperative planning. For instance, Hua et al. implemented a cascaded V-Net model ensembling on segmented gliomas, which achieved high accuracy in delineating the whole tumor, tumor core, and enhanced tumor regions on the BraTS 2018 online validation set [15]. Another study by Shen et al. explores the use of a convolutional neural network combined with near-infrared II (NIR-II) fluorescence imaging, which achieves high sensitivity and specificity in the classification of tumor versus non-tumor intraoperatively [24]. The YOLO algorithm for object detection was then implemented by Abdulsalomov et al, who developed a YOLOv7 model for the detection of glioma tumors using MRI images, achieving 99.5% accuracy [2].

While the aforementioned models report high performances, they cannot be used intraoperatively and do not provide real-time imaging critical for glioma resection. The FL-CNN model proposed does have intraoperative capabilities, but it can only be used on fluorescent images, rendering it infeasible for ioMRI applications. This further clarifies the need for an ioMRI-specific model.

Recent research has shown an abundance of high-resolution, preoperative MRI data, prompting efforts to leverage this data as a proxy for ioMRI. Fei et al. addressed this by simulating low-field interventional MRI images to align real-time interventional MRI images with high-resolution MRI images [10]. By adding noise and creating thicker slices, they successfully simulated 3D images that matched the signal-to-noise ratio of interventional MRI images [10]. Given that interventional MRI and ioMRI have the same fundamental qualities, their method can be used to simulate the dataset necessary to train an effective model [8].

In this context, we introduce a novel method using the YOLO algorithm combined with SAM to identify and detect glioma tumors in real time during ioMRI. In this study, we introduced a novel YOLO + SAM model capable of detecting and segmenting glioma tumors using ioMRI images. The model achieved a DICE score of 0.79 for ET and inference time of 15 to 25 seconds, which displays a robust ability for efficient and effective tumor segmentation. This can have a profound impact in the field of glioma surgery as integrating this model with an ioMRI machine could result in improved patient outcomes and more successful surgeries.

To address the limitations of this model, several areas for improvement have been identified. The first is that the YOLO + SAM model we produced was trained solely on simulated ioMRI images, and for

future research a model trained on proper clinical ioMRI images could have better performance and accuracy. The second is that the SAM model used for this currently only supports 2D inputs, which according to Zhang et al. could "result in a loss of context information", so an application of SAM to 3D data could be a promising venture [31]. A possible method for incorporating 3D data with SAM is by using TomoSAM which is a 3D slicer extension that uses SAM to help with the segmentation of 3D data from tomography or other imaging methods [22].

References

[1] Glioma diagnosis. `https://www.mskcc.org/cancer-care/types/glioma/glioma-diagnosis`. Accessed: 2024-06-17.

[2] A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo. Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers*, 15(16):4172, 2023.

[3] L. Abernethy, S. Avula, G. Hughes, E. Wright, and C. Mallucci. Intra-operative 3-t mri for paediatric brain tumours: challenges and perspectives. *Pediatric radiology*, 42:147–157, 2012.

[4] U. Baid, S. Ghodasara, S. Mohan, M. Bilello, E. Calabrese, E. Colak, K. Farahani, J. Kalpathy-Cramer, F. C. Kitamura, S. Pati, et al. The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint arXiv:2107.02314*, 2021.

[5] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, and et al. Segmentation labels and radiomic features for the pre-operative scans of the TCGA-GBM collection. The Cancer Imaging Archive, 2017. Data set.

[6] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, and et al. Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection. The Cancer Imaging Archive, 2017. Data set.

[7] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data*, 4(1):1–13, 2017.

[8] R. T. Blanco, R. Ojala, J. Kariniemi, J. Perälä, J. Niinimäki, and O. Tervonen. Interventional and intraoperative mri at low field scanner–a review. *European journal of radiology*, 56(2):130–142, 2005.

[9] S. T. Bukhari and H. Mohy-ud Din. E1d3 u-net for brain tumor segmentation: submission to the rsna-asnr-miccai brats 2021 challenge. In *International MICCAI Brainlesion Workshop*, pages 276–288. Springer, 2021.

[10] B. Fei, J. L. Duerk, and D. L. Wilson. Automatic 3d registration for interventional mri-guided treatment of prostate cancer. *Computer Aided Surgery*, 7(5):257–267, 2002.

[11] N. Foroglou, A. Zamani, and P. Black. Intra-operative mri (iop-mr) for brain tumour surgery. *British journal of neurosurgery*, 23(1):14–22, 2009.

[12] I. J. Gerard, M. Kersten-Oertel, K. Petrecca, D. Sirhan, J. A. Hall, and D. L. Collins. Brain shift in neuronavigation of brain tumors: A review. *Medical image analysis*, 35:403–420, 2017.

[13] R. Goldbrunner, M. Ruge, M. Kocher, C. W. Lucas, N. Galldiks, and S. Grau. The treatment of gliomas in adulthood. *Deutsches Ärzteblatt International*, 115(20-21):356, 2018.

[14] D. H. Haydon, M. R. Chicoine, and R. G. Dacey Jr. The impact of high-field-strength intraoperative magnetic resonance imaging on brain tumor management. *Neurosurgery*, 60:92–97, 2013.

[15] R. Hua, Q. Huo, Y. Gao, H. Sui, B. Zhang, Y. Sun, Z. Mo, and F. Shi. Segmenting brain tumor using cascaded v-nets in multimodal mr images. *Frontiers in computational neuroscience*, 14:9, 2020.

[16] G. Jocher, A. Chaurasia, and J. Qiu. Ultralytics yolo (version 8.0.0) [computer software], 2023. URL `https://github.com/ultralytics/ultralytics`.

[17] G. Lab. Brain shift. `https://golbylab.bwh.harvard.edu/brain-shift/`, 2024. Accessed: 2024-06-28.

[18] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2014.

[19] H. Peiris, Z. Chen, G. Egan, and M. Harandi. Reciprocal adversarial learning for brain tumor segmentation: a solution to brats challenge 2021 segmentation task. In *International MICCAI Brainlesion Workshop*, pages 171–181. Springer, 2021.

[20] M. M. Rahman Siddiquee and A. Myronenko. Redundancy reduction in semantic segmentation of 3d brain tumor mris. In *International MICCAI Brainlesion Workshop*, pages 163–172. Springer, 2021.

[21] C. M. Rogers, P. S. Jones, and J. S. Weinberg. Intraoperative mri for brain tumors. *Journal of neuro-oncology*, 151:479–490, 2021.

[22] F. Semeraro, A. Quintart, S. F. Izquierdo, and J. C. Ferguson. Tomosam: a 3d slicer extension using sam for tomography segmentation. *arXiv preprint arXiv:2306.08609*, 2023.

[23] R. Shamsuddeen. *Assessment of the Role of Different MRI Sequences and Contrast in Determining Tumor Margins in Musculoskeletal Tumors*. PhD thesis, Rajiv Gandhi University of Health Sciences (India), 2018.

[24] B. Shen, Z. Zhang, X. Shi, C. Cao, Z. Zhang, Z. Hu, N. Ji, and J. Tian. Real-time intraoperative glioma diagnosis using fluorescence imaging and deep convolutional neural networks. *European Journal of Nuclear Medicine and Molecular Imaging*, 48(11):3482–3492, 2021.

[25] S. T. Solís, C. de Quintana Schmidt, J. G. Sánchez, I. F. Portales, M. d. Á. de Pedro, V. R. Berrocal, R. D. Valle, and G. de trabajo de la SENEC. Intraoperative imaging in the neurosurgery operating theatre: A review of the most commonly used techniques for brain tumour surgery. *Neurocirugía (English Edition)*, 31(4):184–194, 2020.

[26] M. C. Staff. Intraoperative magnetic resonance imaging (imri). https://www.mayoclinic.org/tests-procedures/intraoperative-magnetic-resonance-imaging/about/pac 20394451#: :text=In Accessed: 2024-06-24.

[27] L. Van Hese, S. De Vleeschouwer, T. Theys, et al. The diagnostic accuracy of intraoperative differentiation and delineation techniques in brain tumours. *Discover Oncology*, 13:123, 2022. doi: 10.1007/s12672-022-00585-z. URL https://doi.org/10.1007/s12672-022-00585-z.

[28] G. A. Whitfield, S. R. Kennedy, I. Djoukhadar, and A. Jackson. Imaging and target volume delineation in glioma. *Clinical Oncology*, 26(7):364–376, 2014.

[29] V. Wykes, A. Zisakis, M. Irimia, I. Ughratdar, V. Sawlani, and C. Watts. Importance and evidence of extent of resection in glioblastoma. *Journal of Neurological Surgery Part A: Central European Neurosurgery*, 82 (01):075–086, 2021.

[30] T. Zeng, D. Cui, and L. Gao. Glioma: an overview of current classifications, characteristics, molecular biology and target therapies. *Front Biosci (Landmark Ed)*, 20:1104–1115, 2015.

[31] P. Zhang and Y. Wang. Segment anything model for brain tumor segmentation. *arXiv preprint arXiv:2309.08434*, 2023.