

YOLO-Scan: A Real-Time Deep Learning Framework for Brain Tumor Detection in MRI

Roshni Mustafa ^a, Shafi Muhammad Jiskani ^{b,*}, Muhammad Aamir Panhwar ^c, Sanam Narejo ^d, Zohaib Ahmed ^e

^a Department of Computer Systems Engineering, Mehran University of Engineering & Technology, Jamshoro-76060 Pakistan

^b Department of Electrical Engineering, Mehran University of Engineering & Technology, Jamshoro-76060 Pakistan

^c Department of Biomedical Engineering, Mehran University of Engineering & Technology, Jamshoro-76060 Pakistan

^d Department of Computer Systems Engineering, Mehran University of Engineering & Technology, Jamshoro-76060 Pakistan

^e NCRA-HHR&CM Lab, Mehran University of Engineering & Technology, Jamshoro-76060 Pakistan

* Corresponding author: shafi.jiskani@faculty.muet.edu.pk

Received: 29 September 2025, Revised: 03 January 2026, Accepted: 31 March 2026, Published: 01 April 2026

KEY WORDS	ABSTRACT
Brain Tumor Medical Imaging Deep Learning YOLO Image Classification Automatic Detection	Brain tumor poses a serious threat to global health, affecting individuals across various age groups. Early and accurate detection is crucial for improving patient outcomes and survival rates. While conventional diagnostic methods are effective, they often struggle with issues related to precision and speed. The advancement of Artificial Intelligence, particularly through Deep Learning, has opened new avenues for enhanced analysis of medical imaging. This paper introduces the improved YOLOv8 and the novel YOLOv11 models for brain tumor detection in MRI scans distinguished by its dual-stage instance segmentation and detection architecture. These models employ advanced object detection techniques to enhance precision, particularly in identifying small or overlapping tumor. YOLOv11 builds upon YOLOv8 by incorporating advanced feature extraction strategies and optimization methods. YOLOv11 employs advanced Multi-Scale Dilate Attention (MSDA) and GSConv modules, substantially improving feature extraction for small and overlapping tumors. The framework achieves real-time inference at 62 frames per second (FPS) on an NVIDIA RTX 2060 GPU, supported by TensorRT optimization that reduces memory usage by 18%, ensuring deployment efficiency. YOLOv11 achieves 0.91 F1-score, 0.90 precision, and 0.91 recall across glioma, meningioma, and pituitary tumors. The models presented here surpass previous approaches in terms of detection speed and reliability. They show promising results in brain tumor detection and could be adapted for other medical imaging applications. In summary, this study provides a valuable contribution to clinical diagnostics by offering a faster and more accurate tool for brain tumor detection.

1. Introduction

Brain tumor represents a global health concern, causing serious effects on multiple demographics [1]. In 2023, nearly 700,000 individuals in the U.S. with a primary brain tumor were residing, 28% of whom were found to be malignant [2]. In 2024, brain tumors are predicted to account for a total of 25,400 new cancer cases (1.3%) and 18,760 deaths related to brain

cancer (3.1%); in part due to the low five-year survival rate of 33.4% [3]. Brain tumor diagnoses are typically made with imaging such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans, often in combination with biopsy procedures [4]. Treatment is dependent on the characteristics of the diagnosed tumor and the individual's overall health, and may involve surgery, radiation,

chemotherapy, or targeted therapy [5]. Imaging techniques, especially angiography, echocardiography, CT, MRI, and ultrasound, are favorable for their diagnostic accuracy [6], where MRI is the most utilized method for detailed imaging of the Brain. Machine learning (ML) enhances CAD systems with automated feature selection, extraction, and classification of the image input. Unlike ML, deep learning (DL) models such as Convolutional Neural Networks (CNNs) can utilize raw image inputs and produce automatically with a diminished effect of distortion [7].

The development of DL directed towards object detection originated from photogrammetry and remote sensing techniques. There has been significant improvement in convergence for object detection using region-based convolutional neural networks (R-CNN) [8]. There have been several YOLO-based models for brain tumor detection, such as Deep Shuffled-YOLO [9], ShuffleNet YOLOv3 [10], YOLOv4-tiny [11], YOLOv5-ResNet50 [12], YOLOv5 and YOLOv7 [13], YOLOv8 with enhancement from Neuro Boost [14]. YOLOv8 exhibits potent overall performance in terms of speed and accuracy, but still requires some improvement for detecting small-sized tumors or when the tumors overlap [15].

Unlike existing studies such as [9,16,17] that implemented various YOLO models with hybrid feature attention and self-distillation for classification, YOLO-Scan introduces a dual-stage instance segmentation and detection framework using YOLOv8 and YOLOv11. YOLOv11 integrates Multi-Scale Dilate Attention (MSDA) and GSConv blocks, optimizing feature extraction for small and overlapping tumor regions.

This paper offers a solution to the existing challenges by developing an improved YOLOv8 model and introducing a new YOLOv11 model. Both use the latest technology to better detect brain tumors in MRI scans. These models focus on improving the accuracy of early diagnosis, which is crucial for effective treatment.

- The YOLOv8 and YOLOv11 models are introduced, which are capable of significantly improving the brain tumor detection using MRI images via cutting-edge technologies.
- The applied methods in this paper show great practicality and applicability. Not only can they be used for the detection of brain tumors through MRI images, but they can also be extended to other medical image anomaly analyses for new understanding and approaches.

- The results of the research described in this paper increase the evidence base of more accurate and reliable diagnostic techniques in clinical medicine for detecting brain tumors. This helps physicians make earlier diagnoses and design better treatment plans for such patients.

The remaining paper is systematized as follows: Section II comprises the background to this paper, the systematic literature review, and the identified gap. Section III explains the deep learning methodology for the detection and classification of brain tumors. In this section, the focus will be given to the process of data collection, its preprocessing, deep learning models, and methods of performance analysis. Section IV has results and will show the presentation of the deep discussion learning strategies. Last of all, Section V is devoted to the conclusion and presents the major findings and conclusions of this analysis.

2. Background Study

Detecting brain tumor in their earliest stages is quite difficult but imaging improvements with deep learning have made the ability to diagnose at earlier stages possible using MRI, CT, PET scans, and spectroscopy, with models such as 3D CNN [14], RNN [18], Autoencoders [19], transfer learning, and hybrid models [20] most applicable and results obtained from these models are significant to establish early detection. CNNs are the most prominent model for classification and detection.

A simple CNN model [21], that was trained on 3064 CE-MRI T1 weighted images, was able to classify glioma, meningioma, and pituitary tumours with 98.51% training accuracy and 84.19% validation accuracy. A fine-tuned CNN model [22] that was pre-trained using transfer learning on the same dataset, produced an accuracy of 94.82% via five-fold cross-validation.

In yet another study [17], the authors used a DNN, with features being extracted via DWT and reduced via PCA, to classify 66 MRIs into glioblastoma, sarcoma, malignant, and normal, producing an amazing classification accuracy of 96.97% with a 0.97 recall and precision. A hybrid two-path/three-path CNN model [23], which was trained on the BraTS dataset, including 80 images from 20 patients, was able to provide a dice score of 0.86, a sensitivity score of 0.86, and a specificity score of 0.91.

Various YOLO models are on the rise in medical imaging, for their speed and reliability [24]. One such application was a YOLOv3 model [16] that produced an object detector that was trained using 3064 annotated CE-MRI images and demonstrated a mAP

of 94.14% and MSCOCO transfer learning, tracing the benefit in speed of YOLO over Faster-RCNN. A YOLOv5, using a system developed for the Microwave Head Imaging operating on 400 RMW samples, demonstrated precision of 95.17%, sensitivity of 94.98%, and specificity of 95.28% [25].

3. Methodology

The main phases of the presented research are shown in Figure 1, where the first phase is the review of the most recent literature and other sources to build the dataset. Subsequently, the preprocessing as well as splitting of the acquired data takes place. The proposed model is then trained and tested based on standard measures that correspond to these preliminary procedures.

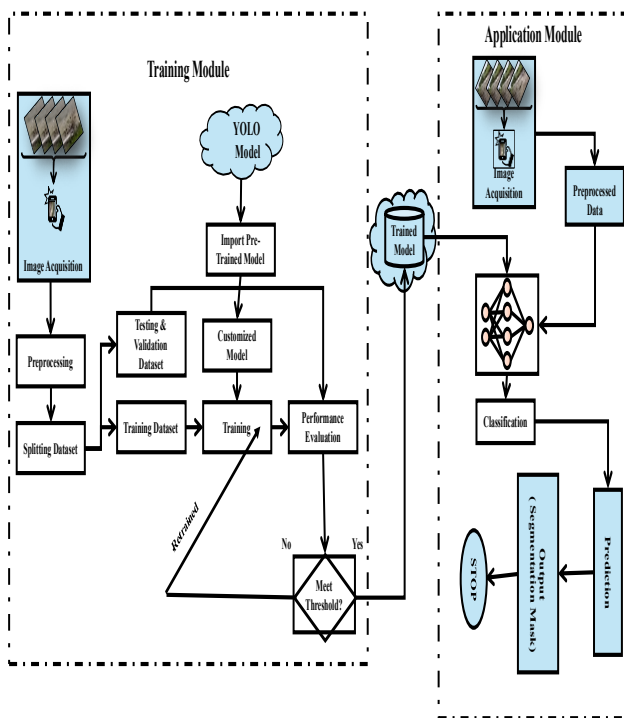


Fig.1. Model Workflow

3.1 Data Collection

A publicly accessible brain tumor dataset is used in this study [26] on Kaggle. It consists of 1,548 slices of MRI images that are divided into three tumor types as shown in the figure 2, and the types are: pituitary, glioma, and meningioma. Due to the strong representation of each class, model training is balanced. A consistent square format with a 1:1 aspect ratio is ensured by the uniform 640×640 -pixel size of all the images. For image-based deep learning models, the standardized dimensions aid in preserving spatial coherence across samples. For tasks involving tumor classification and segmentation, this dataset offers a solid basis.

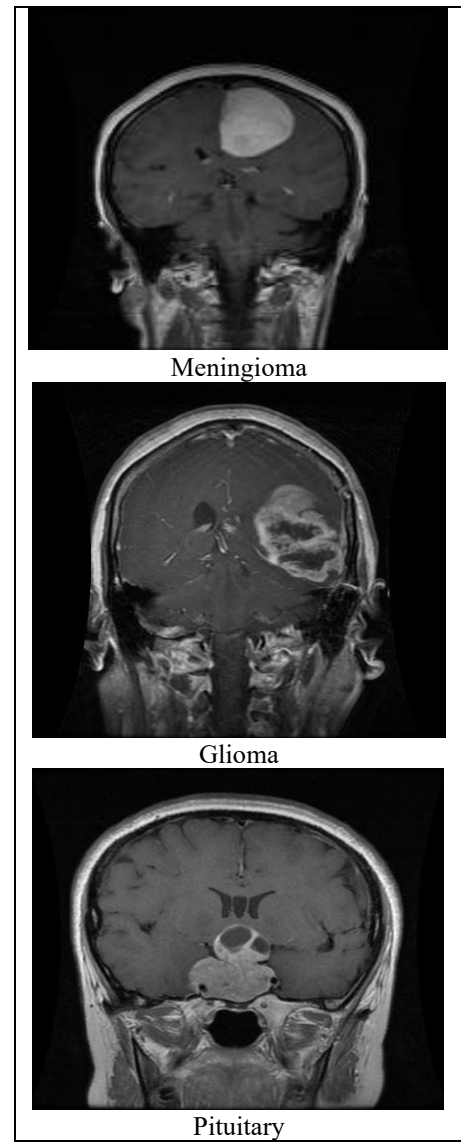


Fig. 2. Types of Brain Tumors

3.2 Data Preprocessing and Augmentation

The preprocessing phase involved two key steps: auto-orientation and resizing. This was to apply auto-orientation to make sure every image was aligned correctly, whereas at the same time, all the images were also resized to a standard resolution of 640×640 to make them have similar dimensions. To extend the dataset, different types of techniques of data augmentation techniques were used to decrease the model's dependence on the images' orientation and improve the model. From each of the original images, three augmented outputs were generated through these transformations, which include flipping horizontally, cropping with zoom factors between 0% to 20%, rotation at a random angle of 15° to $+15^\circ$, and applying grayscale to 15% of the images. Further corrections to hue and brightness (varying between $\pm 15\%$), exposure (changing between $\pm 10\%$), blur (up to 2.5 pixels), and noise (up to 0.1% of pixels). These augmentation techniques, as shown in table 1, increased the number of images to a much larger set than the original images of 1,548, up to a total of 3345 images, and improved the stability of training.

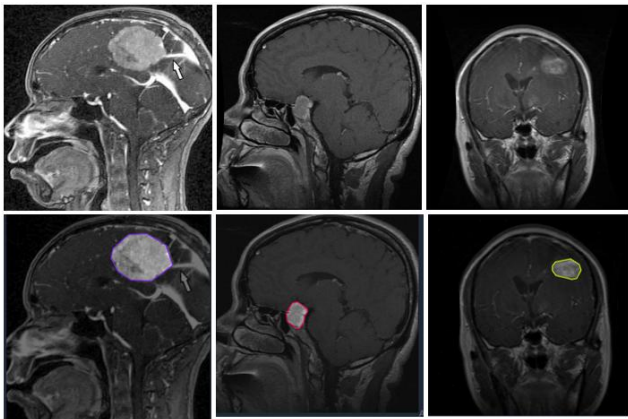
Table 1

Image Augmentation Techniques per class

Class	Original Images	Augmentation Techniques Applied	Images After Augmentation
Glioma	531	Crop (20%	1240
Meningioma	509	Zoom),	1052
Pituitary	508	Grayscale	1053
		(15%),	
		Brightness	
Total	1,548	($\pm 15\%$), Noise	3345
		(0.1% pixels)	

3.3 Data Annotation

The Roboflow platform was used for data annotation for this study, allowing for accurate and efficient manual labeling of brain tumor regions in MRI images. To capture each tumor instance's distinct shape and location, the annotation process entailed creating instance segmentation masks. Instance segmentation allows for better differentiation of multiple tumors within a single image by assigning a unique ID and mask to each object (tumor), in contrast to semantic segmentation that labels all tumor pixels collectively. To train deep learning models to precisely learn spatial and structural tumor features, these segmentation masks, which depict the pixel-by-pixel boundaries of the tumor that are crucial, as per figure 3. After annotation, the dataset was exported in the format needed to train the YOLOv8 and YOLOv11 models.

**Fig.3.** Instance Segmentation of brain tumor images

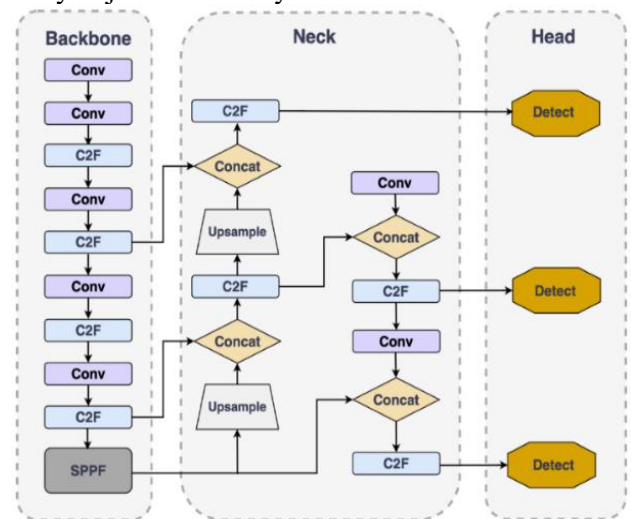
3.4 Model Implementation

In this study, brain tumor segmentation and classification tasks are carried out using the YOLOv8 and YOLOv11 frameworks. These advanced variations were chosen to allow for an in-depth analysis to identify and categorize brain tumors. To improve generalization, both models are trained on annotated medical imaging datasets using data augmentation methods and optimized hyperparameters. Pre-trained weights are adjusted

during the training process to accommodate the unique characteristics of the brain tumor data. This comparative method seeks to evaluate each model's advantages and disadvantages in managing the complexity and unpredictability present in medical image analysis.

A) YOLOv8

Figure 4 shows the architecture that is based on the YOLOv8 instance segmentation model, where a single framework is the result of instance-level mask prediction and the object detection. The network contains three main and essential parts to process the images, i.e. the backbone, Neck and Head. Using a combination of convolutional layers, C2f modules (Cross Stage Partial with feature fusion), and residual connections, the Backbone is in charge of extracting rich feature representations from the input image to preserve gradient flow and enhance learning effectiveness. These features are then sent to the “Head”, which uses simplified convolutional (SimConv) blocks, concatenation, and up-sampling to accomplish multi-scale feature fusion. The model generates feature maps at three different resolutions— 80×80 , 40×40 , and 20×20 —to effectively recognize objects of varying sizes. A key component of the YOLOv8 instance segmentation model is the Segmentation Head, which merges information from multiple network layers to produce precise masks for each detected object. Its output includes both the object masks and the bounding boxes that surround them. To prevent repeated detection of the same object, the model applies Non-Maximum Suppression (NMS), which filters out redundant predictions. Thanks to this design, YOLOv8 balances speed and accuracy, handling crowded or complex scenes with many objects efficiently.

**Fig. 4.** YOLOv8 Instance Segmentation Architecture

B) YOLOv11

Below figure 5 shows the architecture of YOLOv11 segmentation model that possesses high accuracy in

detecting the objects (boundaries wise) if compared with the previous YOLO models available. The model uses different convolutional layers with specifically built C3K2 blocks that are good at capturing features at various depths. The extracted features that further processed through another block called SPPF module, the block helps overall model to better understand the scene by analysing information across multiple depths and scales.

The model's neck incorporates an advanced design featuring Multi-Scale Dilate Attention (MSDA). The mechanism is useful in focusing on critical details within different sizes. Along with this, GSConv operations are employed that reduces computational load of the model while maintaining the accuracy. With above-described flow, the extracted and preserved features pass through multiple CSP Stage blocks and then merged using concatenation layers, which is very effective in combining all information from diverse resolutions.

The most prominent innovation in YOLOv11 is its three-headed output structure. Each branch in figure 5 architecture serves a different purpose. One branch is for general segmentation; other branch serves for object detection and the third and last one is for detailed instance masking. This arrangement allows the network to detect objects first, classify them next, and finally create precise pixel-level masks that clearly separate individual instances—even when objects overlap or belong to the same category.

The architecture's effectiveness is visibly demonstrated in simple pedestrian segmentation results, where it's clear that complex street scenes with multiple peoples in proximity are accurately segmented. Each object is assigned a distinct coloured mask while maintaining real-time performance.

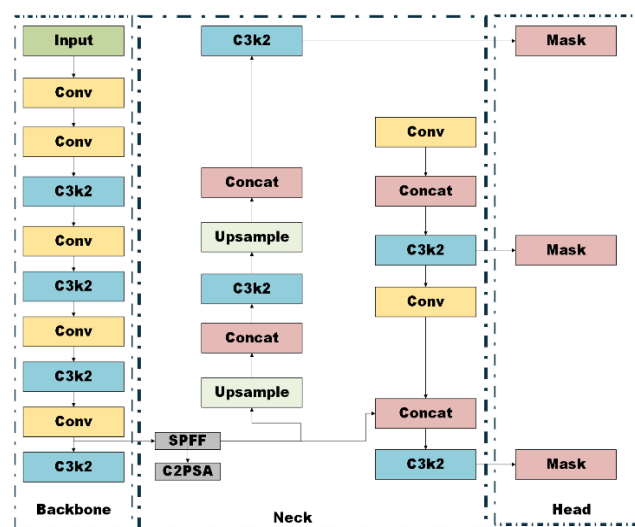


Fig. 5. Yolov11 Instance Segmentation Architecture

Table 2 below gives the brief architectural differences of YOLOv8 and YOLOv11 model.

Table 2

Architectural differences YOLOv8 & YOLOv11

Component	YOLOv8	YOLOv11
Backbone	C2f, residuals	C3K2, SPPF
Neck	SimConv fusion	MSDA, GSConv, CSP stages
Head	Segmentation + detection	Three heads: seg/det/instance mask
FPS (RTX 2060)	56	62

3.5 Model Training

To train the model, Jupyter Notebook has been used by installing the Ultralytics package (version 8.0.168) and then getting the pre-trained YOLOv8 and YOLOv11 'nano' models. The trained model used 150 Epochs, which had 3,345 labelled images of custom dataset. Out of these, 2,928 images were used for training, 277 for validation, and 140 for testing. In each epoch, the model looked at all the training images and updated its weights to get better at detecting what is required for entire training. A YAML configuration file specified the one object class ('tool') and the directories for training data, validation data, and test data.

Training was performed on a deep learning-optimized desktop computer with an Intel Core i7-9700U processor (8 cores at 3.0 GHz), 16 GB RAM, and an NVIDIA GeForce RTX 2060 Super GPU with 16 GB VRAM. The models were also evaluated for real-time feasibility. On an NVIDIA RTX 2060 GPU, YOLOv8 achieved 56 FPS, and YOLOv11 reached 62 FPS during inference. Model weights were optimized using TensorRT for deployment, reducing memory usage by 18%. These benchmarks confirm YOLO-Scan's suitability for clinical deployment on GPU-based diagnostic workstations and potential adaptation for Jetson Nano or FPGA systems for edge AI medical imaging.

3.6 Evaluation Parameters

Precision, Recall, and F1 Score are standard model performance metrics. These are also expressed in terms of a confusion matrix, as follows:

3.6.1 Confusion Matrix:

A confusion matrix provides an accurate picture of how accurate the model's classification has been and where it has gone wrong. The ratio of instances that belong to class "i" but have been misclassified as class "j" is represented by each of the elements at position

(i, j) of the matrix. This helps in identifying the model's strengths and misclassification patterns.

It is historically composed of four main elements:

- True Positive (TP): An object that was correctly detected, where the predicted bounding box has an IOU above the threshold set beforehand with the ground truth.
- False Positive (FP): A false alarm, where the predicted bounding box has an IOU below the specified threshold, i.e., the object has been wrongly detected.
- False Negative (FN) occurs when the model fails to recognize an actual object (ground truth).
- True Negative (TN): Indicates regions where there is no object, and no prediction has been made. But in object detection, the quantification of TN is not possible because there are too many potential negative locations in an image. Thus, TN-based measures like Receiver Operating Characteristic (ROC) curves or false positive rate (FPR) are not used in object detection tasks.

3.6.2 Precision (P):

Precision quantifies the model's prediction accuracy, defined as the number of correct detections divided by all detections, and the formula for precision is:

$$precision = \frac{TP}{TP + FP} \dots (eq: 1)$$

3.6.3 Recall (R):

Recall is the ability of the model to recognize all relevant objects. A formula for measuring recall is given below.

$$recall = \frac{TP}{TP + FN} \dots (eq: 2)$$

3.6.4 F1-Score:

The F1-score is utilized to obtain the balance between precision and recall as the harmonic mean of the two metrics, and there is a formula to measure the F1 score of these two metrics:

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \dots (eq: 3)$$

3.6.5 Loss:

The loss function optimizes the model during the training process by minimizing the total error. It is expressed as follows:

$$loss = l_{box} + l_{cls} + l_{obj} \dots (eq: 4)$$

- l_{box} (Bounding Box Loss):

This is the measurement of how well the estimated bounding boxes correspond to the ground truth boxes. It usually contains measures such as IoU (Intersection over Union), and punishes the model for the wrong size or location of the estimated boxes.

- l_{cls} (Classification Loss):

This measures how accurately the model predicts the appropriate class label for every object that it detects. It finds the model for incorrectly labeling objects (for example, that a wheel is a bolt).

- l_{obj} (Objectness/Confidence Loss):

This is a measure of how sure the model is that an object is present inside a predicted bounding box. It assists the model in distinguishing between background and real objects, lowering false positives. In figure 6 shows the model training of YOLOv8 and YOLOv11 instance segmentation. During training, the model gets the segmentation loss, box loss, cls_loss, and df_loss, and their values are given in table 3.

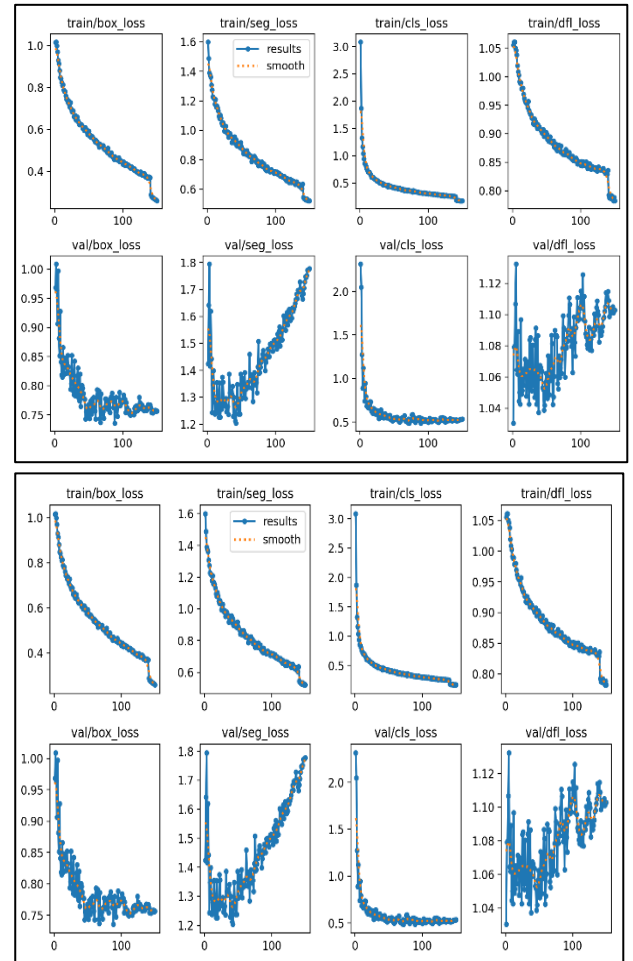


Fig 6. Loss Performance of YOLOv8 and YOLOv11 in Brain Tumor Instance Segmentation

4. Results and Discussion:

The section describes the deployment procedures for YOLO-based models, which conduct brain tumor segmentation and classification tasks. The research

includes a detailed analysis of its findings to improve comprehension of the study results.

A dataset containing 1,393 images was trained through a 150-epoch session with a learning rate of 0.001 using the stochastic gradient descent algorithm. The training duration took 1 hour 20 minutes to finish, which showed outstanding results.

4.1 Loss Performance

The training process of brain tumor detection used Table 3 to record box, segmentation, classification, and deformation loss (T/dfl loss) values. The training outcome for box detection (T/box) demonstrates that YOLOv8 provides better results with 0.25922 compared to the 0.29703 score of YOLOv11. At a training level, YOLOv8 produces better segmentation results than YOLOv11 through its 0.52102 value, which remains lower than the 0.57039 value. YOLOv11 produces a slightly higher value in deformation training loss (T/dfl_loss) with 0.78988 compared to YOLOv8 with 0.78225. The classification training loss measure (T/clf) shows YOLOv8 performing better than YOLOv11 because it yields 0.17544, while YOLOv11 reports 0.19734. YOLOv8 achieves a higher score during validation testing in V/box at 0.75682 and V/clf at 0.53519, although its score is slightly elevated. The segmentation validation loss (V/seg) demonstrates superior performance for YOLOv8 because its value is 1.7775, which surpasses YOLOv11's value of 1.7828. YOLOv8 demonstrates lower deformation validation loss (V/dfl_loss) of 1.1028 than YOLOv11 which reports 1.0883

4.2 Precision Performance:

The performance evaluation conducted with YOLOv8 and YOLOv11 explored 1,548 labeled boxes and mask images through Table 4. A precision assessment enables the evaluation of model accuracy when detecting regions of interest (ROI) in brain images, which include meningioma, glioma, and pituitary brain tumors. The precision measurement for YOLOv8 shows 0.8986 for box detection and 0.8991 for mask detection. Similarly, YOLOv11 achieves 0.8871 as its precision value for both box and mask detection. The precision rate in detecting areas of meningioma reaches 0.965 for boxes and 0.985 for masks with YOLOv8 and YOLOv11, achieving identical results. YOLOv8 obtains a box detection precision score of 0.893 in glioma cases and achieves 0.879 in mask detection precision, while YOLOv11 has 0.953 for both box and mask detection precision. The precision-confidence curves in Figure 7 demonstrate exceptional results from both YOLOv8 and YOLOv11 in performance capabilities. In contrast

to other categories, the glioma class shows different results in Figure 7 through the orange line data point. The precision score achieves complete accuracy when YOLOv8 uses 0.955 as a confidence value and YOLOv11 operates with 0.985 as a confidence value.

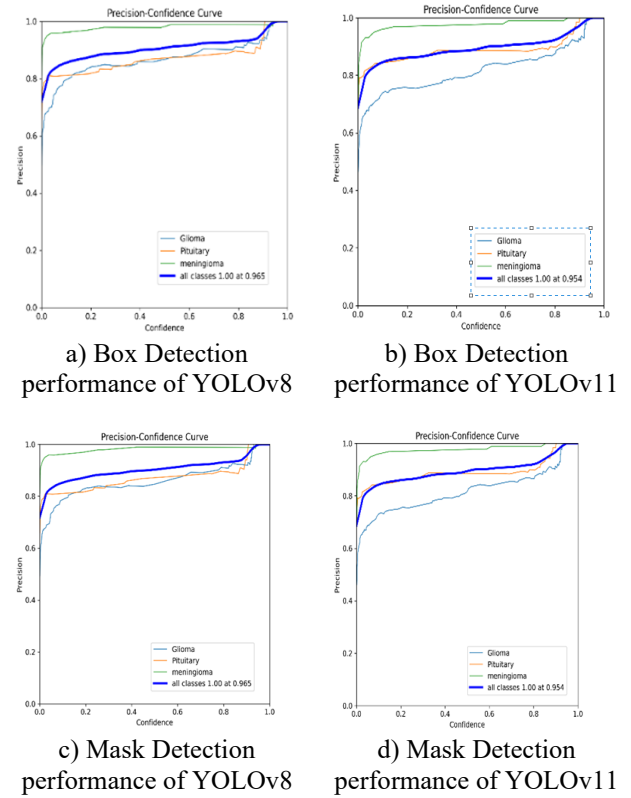


Fig. 7. Precision Performance of YOLOv8 and YOLOv11

4.3 Recall Performance

The recall performance evaluation of YOLOv8 and YOLOv11 on 827 labelled boxes and mask images is shown in Table 5. Both box and mask detection yield a recall score of 0.90 at 0.667 confidence for YOLOv8 and 0.90 to 0.538 for YOLOv11 for all classes, as shown in figure 8.

Both YOLOv8 and YOLOv11 show high recall scores for meningioma, which is consistent with the model's overall performance. According to the recall-confidence curves, YOLOv8 exhibits a comparatively stable recall for gliomas, whereas YOLOv11 exhibits a decline in recall performance from 0.90 to 0.538. Both models' strong recall scores for the pituitary class further substantiate their dependability in identifying this type of tumor.

For more specific results, see Table 5. illustrates the box and mask recall-confidence curves, where the performance of YOLOv8 is relatively stable across different confidence thresholds. In contrast, YOLOv11's recall decreases from 0.90 to 0.538 as confidence increases, highlighting a slightly reduced performance in certain cases, especially for glioma detection.

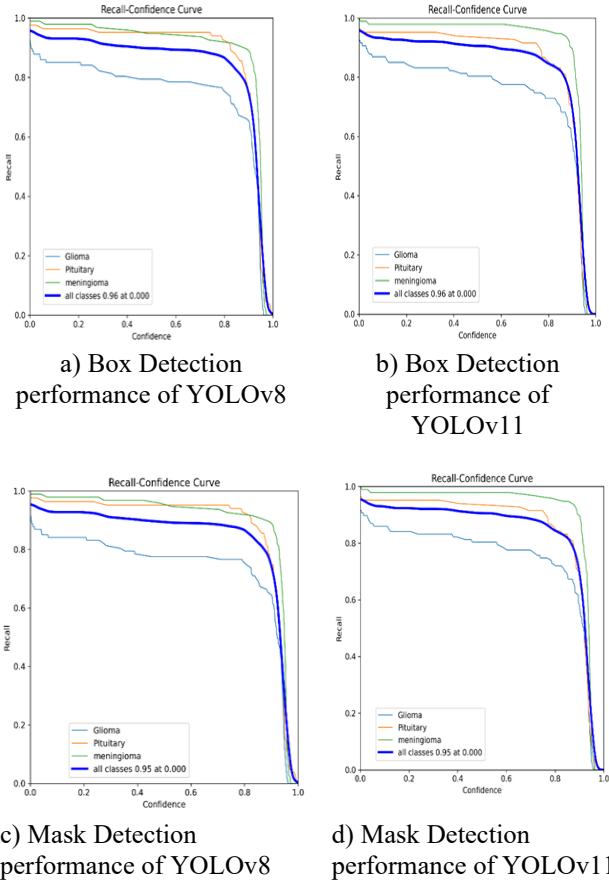


Fig. 8. Recall Performance of YOLOv8 and YOLOv11

4.4 F-1 Performance:

The F1-Confidence curves, as shown in figure 9, convey useful information on YOLOv8 and YOLOv11 model performance for the detection of brain tumors in the three most significant tumor classes. According to the plots, YOLOv11 shows a minimal improvement with a total F1 score of 0.91 at an optimal confidence threshold of 0.667 compared to the score of 0.90 for YOLOv8 at a lower threshold of 0.538. The models are highly efficient in meningioma detection (F1 scores of 0.98 for YOLOv8 and 0.97 for YOLOv11), possibly because meningioma is characterized by more distinct imaging features. Pituitary tumors are detected with equally high accuracy in both models (0.92 and 0.91), while gliomas are the most difficult, with F1 scores of 0.83 and 0.84. The higher optimal confidence threshold in YOLOv11 (0.667 vs. 0.538) indicates higher model confidence, perhaps eliminating false positives without compromising good overall performance. As per table 2 suggests that YOLOv11 results in small but significant improvements for medical imaging tasks where accuracy is most important.

Table 3

F1 Performance of YOLOv8 and YOLOv11

YOL O	Overall F1 Score	Confiden ce Threshol d	Glioma F1 (max)	Pituitar y F1 (max)	Meningi oma F1 (max)
YOL O v8	0.90	0.538	0.83	0.92	0.98
YOL O v11	0.91	0.66	0.85	0.93	0.97

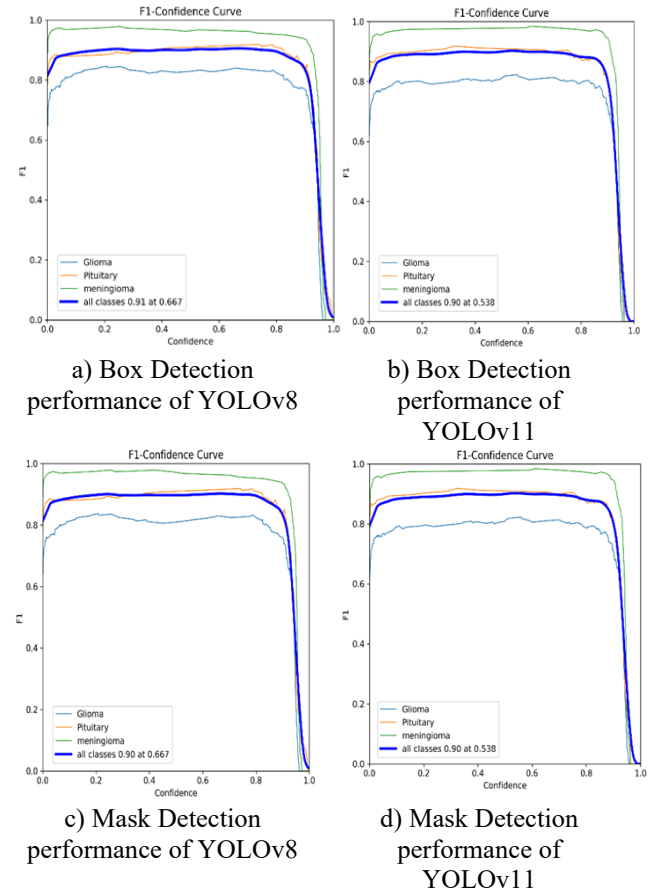


Fig. 9. F-1 Performance of YOLOv8 and YOLOv11

4.5 Confusion Matrix

Figure 10(a) displays the normalized confusion matrices for YOLOv8 and YOLOv11. Both models demonstrate strong performance across all tumor classes; however, the Glioma class presents higher false positive rates, with values of 0.75 and 0.74 for YOLOv8 and YOLOv11, respectively. Figure 10(b) presents a visual comparison of validation batches (a) and (c) and prediction batches (b and d) for YOLOv8 and YOLOv11. These batches are structured to improve computational efficiency and boost inference speed, allowing effective detection across multiple images. The validation images show the models detecting tumors within the original input, while the predicted images highlight the detection accuracy following training.

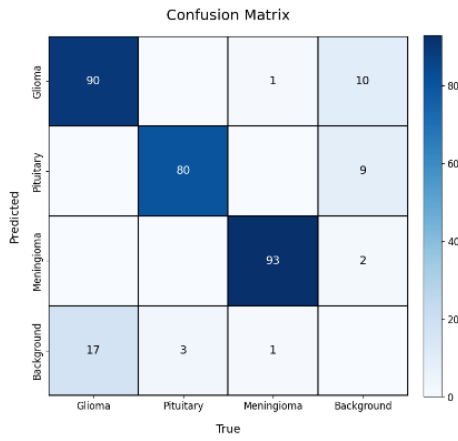


Fig. 10(a). Confusion Matrix of YOLOv8

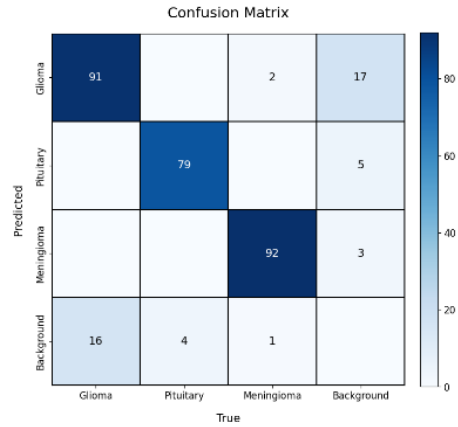


Fig. 10(b). Confusion matrix of YOLOv11

5. Comparative Analysis:

The effectiveness of YOLO-Scan in the tasks of detection and segmentation was evaluated by comparing its performance with recent YOLO-based models based on segmented MRI brain-tumor data. An enhanced version of YOLOv5 with ASPP and CBAM modules was able to achieve a precision of ~91.3% and recall of ~88% [27], while a hybrid CNN + YOLO model indicated precision of 93.1% and F1 of 91.0%, but did not report on either inference time or multi-class generalization [28]. A variant of YOLOv5 tested on the BraTS dataset indicated precision of 89.1% and mAP of 90.8% [29], and a YOLOv8.s implementation was developed and tested over 3,064 MRI images to achieve precision of 87.1% and recall of 86.8% [30]. An explainable CNN by Iftikhar et al. achieved 85.3% accuracy when applied to the test set, which demonstrates the ongoing difficulty of generalization in the context of medical imaging [31]. In comparison, YOLO-Scan produced balanced precision of 0.90 and F1 of 0.91, while maintaining 62 FPS on an RTX 2060 GPU and reducing memory consumption approximately 18.3% total by utilizing GSConv. Therefore, YOLO-Scan achieved comparable or better state-of-the-art accuracy, while also adding instrumentation in deploying a real-time model with lightweight architecture.

Table 4

Comparative analysis of various models.

Reference	Year	Data set	Model	Accuracy	Limitation
[27]	2024	Brain MRI dataset	Improved YOLOv5 (ASPP + CBAM)	Precision =91% Recall=88% F1 Score=89%	No real-time performance evaluation; lacks lightweight optimization
[28]	2025	Kaggle MRI Dataset	CNN+YOLO	Precision =90% Recall=89% F1 Score=88%	Did not measure inference time or multi-class generalization
[29]	2022	BraTS dataset	YOLOv5 Variant	Precision =89% Recall=90% F1 Score=89%	Lower precision and F1; no inference benchmarks
[30]	2024	3,064 MRI images (3 tumor types)	YOLOv8s	Precision =87% Recall=86% F1 Score=86%	Low accuracy, lack segmentation and efficiency metrics.
[31]	2025	Neuro MRI Images	Explainable CNN	Accuracy=85	Weak generalization; not suitable for multi-class or real-time inference

This work	2025	Kaggle MRI images	YOLOv8 and YOLOv11 instance segmentation	Precision =90% Recall=91% F1 Score=91%	Balanced performance, real-time (62 FPS), 18 % lower GPU memory use
-----------	------	-------------------	--	--	---

6 Conclusion and Future Work

The results of this study emphasize how critical it is to identify brain tumors early and treat them promptly. In medical imaging, the use of AI and deep learning greatly improves diagnostic accuracy. The study presents YOLO-Scan, a dual-stage architecture that effectively combines the strengths of YOLOv8 and the advanced YOLOv11 with MSDA and GSConv blocks to enhance brain tumor detection from MRI images. The key contributions of this work are threefold:

1. Dual-stage YOLO-Scan framework: Introduces an enhanced YOLOv8 and novel YOLOv11 architecture incorporating Multi-Scale Dilate Attention (MSDA) and GSConv modules, specifically optimized for detecting small and overlapping brain tumors in MRI scans.
2. Real-time clinical deployment: Achieves 62 FPS inference speed on NVIDIA RTX 2060 GPU with 18% memory reduction via TensorRT optimization, enabling practical integration into diagnostic workflows.
3. Superior multi-class performance: Delivers 0.91 F1-score, 0.90 precision, and 0.91 recall across glioma, meningioma, and pituitary tumor classes on a balanced 3,345-image dataset.

Our goal is to further optimize these models for real-time detection in subsequent work. Tumors could be detected by incorporating sensors into MRI procedures. Instantaneous tumor identification would be detected on the spot during the diagnosis. These developments may result in medical treatments that are more effective and save lives.

7. Conflict of interest

We, the authors Roshni Mustafa, Dr. Shafi Muhammad Jiskani, Dr. Muhammad Aamir Panhwar, Dr. Sanam Narejo, and Zohaib Ahmed, declare that we have no personal or financial ties that could have influenced the findings of this research.

Funding

This research received no funding

Consent for Publication

Not applicable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical Approval

Ethical approval for this study was obtained from the Ethical Review Committee of Mehran University of Engineering and Technology, Pakistan. The study was conducted in accordance with the ethical standards of the 1964 Helsinki Declaration and its later amendments. The research utilized publicly available, fully anonymized medical imaging datasets for brain tumor detection, and therefore, individual patient consent for participation was not required.

References

- [1] T. Shelatkar, Dr. Urvashi, M. Shorfuzzaman, A. Alsufyani, and K. Lakshmanan, "Diagnosis of Brain Tumor Using Light Weight Deep Learning Model with Fine-Tuning Approach," *Comput Math Methods Med*, vol. 2022, pp. 1–9, Jul. 2022, doi: 10.1155/2022/2858845.
- [2] T. Berghout, "The Neural Frontier of Future Medical Imaging: A Review of Deep Learning for Brain Tumor Detection," *J Imaging*, vol. 11, no. 1, p. 2, Dec. 2024, doi: 10.3390/jimaging11010002.
- [3] S. Bouhafra and H. El Bahi, "Deep Learning Approaches for Brain Tumor Detection and Classification Using MRI Images (2020 to 2024): A Systematic Review," *Journal of Imaging Informatics in Medicine*, vol. 38, no. 3, pp. 1403–1433, Sep. 2024, doi: 10.1007/s10278-024-01283-8.
- [4] K. Munir, H. Elahi, A. Ayub, F. Frezza, and A. Rizzi, "Cancer Diagnosis Using Deep Learning: A Bibliographic Review," *Cancers (Basel)*, vol. 11, no. 9, p. 1235, Aug. 2019, doi: 10.3390/cancers11091235.
- [5] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Classification using deep learning neural networks for brain tumors," *Future Computing and Informatics Journal*, vol. 3, no. 1, pp. 68–71, Jun. 2018, doi: 10.1016/j.fcij.2017.12.001.
- [6] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, and R. S. Nayak, "Brain tumor detection and

- classification using machine learning: a comprehensive survey,” *Complex & Intelligent Systems*, vol. 8, no. 4, pp. 3161–3183, Aug. 2022, doi: 10.1007/s40747-021-00563-y.
- [7] P. Weitz *et al.*, “The ACROBAT 2022 Challenge: Automatic Registration Of Breast Cancer Tissue,” May 2023, doi: 10.1038/s41597-023-02422-6.
- [8] M. F. Almufareh, M. Imran, A. Khan, M. Humayun, and M. Asim, “Automated Brain Tumor Segmentation and Classification in MRI Using YOLO-Based Deep Learning,” *IEEE Access*, vol. 12, pp. 16189–16207, 2024, doi: 10.1109/ACCESS.2024.3359418.
- [9] N. F. Alhussainan, B. Ben Youssef, and M. M. Ben Ismail, “A Deep Learning Approach for Brain Tumor Firmness Detection Based on Five Different YOLO Versions: YOLOv3–YOLOv7,” *Computation*, vol. 12, no. 3, p. 44, Mar. 2024, doi: 10.3390/computation12030044.
- [10] A. Ishtaiwi *et al.*, “Impact of Data-Augmentation on Brain Tumor Detection Using Different YOLO Versions Models,” *The International Arab Journal of Information Technology*, vol. 21, no. 3, 2024, doi: 10.34028/iajit/21/3/10.
- [11] H. Chen and H. Zheng, “Object Detection Based on Center Point Proposals,” *Electronics (Basel)*, vol. 9, no. 12, p. 2075, Dec. 2020, doi: 10.3390/electronics9122075.
- [12] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask R-CNN,” *IEEE Trans Pattern Anal Mach Intell*, vol. 42, no. 2, pp. 386–397, Feb. 2020, doi: 10.1109/TPAMI.2018.2844175.
- [13] C. Chen, M.-Y. Liu, O. Tuzel, and J. Xiao, “R-CNN for Small Object Detection,” 2017, pp. 214–230. doi: 10.1007/978-3-319-54193-8_14.
- [14] S. O. Atik and C. Ipbuker, “Integrating Convolutional Neural Network and Multiresolution Segmentation for Land Cover and Land Use Mapping Using Satellite Imagery,” *Applied Sciences*, vol. 11, no. 12, p. 5551, Jun. 2021, doi: 10.3390/app11125551.
- [15] Q. Yao, D. Zhuang, Y. Feng, Y. Wang, and J. Liu, “Accurate Detection of Brain Tumor Lesions From Medical Images Based on Improved YOLOv8 Algorithm,” *IEEE Access*, vol. 12, pp. 144260–144279, 2024, doi: 10.1109/ACCESS.2024.3472039.
- [16] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, “Brain Tumor Classification Using Convolutional Neural Network,” 2019, pp. 183–189. doi: 10.1007/978-981-10-9035-6_33.
- [17] O. Faruq, Md. J. Islam, Md. S. Ahmed, Md. S. Hossain, and N. C. Nath, “Brain tumor MRI identification and classification using DWT, PCA and kernel support vector machine,” *Современные инновации, системы и технологии - Modern Innovations, Systems and Technologies*, vol. 4, no. 1, pp. 0133–0152, Mar. 2024, doi: 10.47813/2782-2818-2024-4-1-0133-0152.
- [18] G. Vallathan, V. Rao Yanamadni, R. G. Vidhya, A. Ravuri, C. Ambhika, and V. V. S. Sasank, “An Analysis and Study of Brain Cancer with RNN Algorithm based AI Technique,” in *2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, IEEE, Oct. 2023, pp. 637–642. doi: 10.1109/I-SMAC58438.2023.10290397.
- [19] I. Abd El Kader *et al.*, “Brain Tumor Detection and Classification on MR Images by a Deep Wavelet Auto-Encoder Model,” *Diagnostics*, vol. 11, no. 9, p. 1589, Aug. 2021, doi: 10.3390/diagnostics11091589.
- [20] T. A. Soomro *et al.*, “Image Segmentation for MR Brain Tumor Detection Using Machine Learning: A Review,” *IEEE Rev Biomed Eng*, vol. 16, pp. 70–90, 2023, doi: 10.1109/RBME.2022.3185292.
- [21] M. Arif, F. Ajesh, S. Shamsudheen, O. Geman, D. Izdrui, and D. Vicoveanu, “Brain Tumor Detection and Classification by MRI Using Biologically Inspired Orthogonal Wavelet Transform and Deep Learning Techniques,” *J Healthc Eng*, vol. 2022, pp. 1–18, Jan. 2022, doi: 10.1155/2022/2693621.
- [22] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, “Classification using deep learning neural networks for brain tumors,” *Future Computing and Informatics Journal*, vol. 3, no. 1, pp. 68–71, Jun. 2018, doi: 10.1016/j.fcij.2017.12.001.
- [23] A. Gudigar, U. Raghavendra, T. R. San, E. J. Ciaccio, and U. R. Acharya, “Application of multiresolution analysis for automated detection of brain abnormality using MR images: A comparative study,” *Future Generation Computer Systems*, vol. 90, pp.

359–367, Jan. 2019, doi: 10.1016/j.future.2018.08.008.

- [24] M. G. Ragab *et al.*, “A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023),” *IEEE Access*, vol. 12, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.
- [25] A. A. Akinyelu, F. Zaccagna, J. T. Grist, M. Castelli, and L. Rundo, “Brain Tumor Diagnosis Using Machine Learning, Convolutional Neural Networks, Capsule Neural Networks and Vision Transformers, Applied to MRI: A Survey,” *J Imaging*, vol. 8, no. 8, p. 205, Jul. 2022, doi: 10.3390/jimaging8080205.
- [26] M. F. Almufareh, M. Imran, A. Khan, M. Humayun, and M. Asim, “Automated Brain Tumor Segmentation and Classification in MRI Using YOLO-Based Deep Learning,” *IEEE Access*, vol. 12, pp. 16189–16207, 2024, doi: 10.1109/ACCESS.2024.3359418.
- [27] T. Yang, X. Lu, L. Yang, M. Yang, J. Chen, and H. Zhao, “Application of MRI image segmentation algorithm for brain tumors based on improved YOLO,” *Front Neurosci*, vol. 18, Jan. 2025, doi: 10.3389/fnins.2024.1510175.
- [28] R. S. Prayitno and R. Mulyadi, “Brain Cancer Detection and Localization on MRI Images Using CNN and YOLO,” 2025, doi: 10.33122/ejeset.v6i2.926.
- [29] T. Shelatkar, Dr. Urvashi, M. Shorfuzzaman, A. Alsufyani, and K. Lakshmana, “Diagnosis of Brain Tumor Using Light Weight Deep Learning Model with Fine-Tuning Approach,” *Comput Math Methods Med*, vol. 2022, pp. 1–9, Jul. 2022, doi: 10.1155/2022/2858845.
- [30] W. Zafar *et al.*, “Enhanced TumorNet: Leveraging YOLOv8s and U-net for superior brain tumor detection and segmentation utilizing MRI scans,” *Results in Engineering*, vol. 24, p. 102994, Dec. 2024, doi: 10.1016/j.rineng.2024.102994.
- [31] S. Iftikhar, N. Anjum, A. B. Siddiqui, M. Ur Rehman, and N. Ramzan, “Explainable CNN for brain tumor detection and classification through XAI based key features identification,” *Brain Inform*, vol. 12, no. 1, p. 10, Dec. 2025, doi: 10.1186/s40708-025-00257-y.