IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 November 2023, accepted 19 January 2024, date of publication 29 January 2024, date of current version 2 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3359418

RESEARCH ARTICLE

Automated Brain Tumor Segmentation and Classification in MRI Using YOLO-Based Deep Learning

MARAM FAHAAD ALMUFAREH[1], MUHAMMAD IMRAN[2], ABDULLAH KHAN[2], MAMOONA HUMAYUN[1], AND MUHAMMAD ASIM[2]

[1]Department of Information Systems, College of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia
[2]Institute of Computer Sciences and Information Technology, Faculty of Management and Computer Sciences, The University of Agriculture, Peshawar, Peshawar 25130, Pakistan

Corresponding authors: Mamoona Humayun (mahumayun@ju.edu.sa) and Muhammad Imran (mimran@aup.edu.pk)

This work was supported by the Deputyship for Research and Innovation, Ministry of Education, Saudi Arabia, under Project 223202.

ABSTRACT Recent advancements in image processing and computer vision have brought significant transformations in healthcare technology, leading to significant improvements in diagnosis accuracy, cost-effectiveness, and time efficiency. Magnetic Resonance Imaging (MRI) is employed by the radiologist for its remarkable ability to detect even the most subtle brain abnormalities. This study considers a comprehensive analysis of the two prominent object identification frameworks, YOLOv5 and YOLOv7, leveraging state-of-the-art deep learning architectures to classify and detect brain cancers within MRI. The brain tumor dataset encompasses three distinct classes, including meningiomas, gliomas and pituitary tumors. To ensure precise segmentation of the tumor regions, the preprocessing phase incorporates advanced mask alignment techniques. This preprocessed dataset has been used to evaluate the performance of the deep learning models for brain tumor detection and classification. From the numerical results of YOLOv5, it was noticed that a recall score of 0.905 for box detection and 0.906 for mask segmentation, with a precision score of 0.94 and 0.936, respectively. At an IoU threshold of 0.5, both box detection and mask segmentation achieve a mAP of 0.947, whereas, at an IoU threshold of 0.5 to 0.95, they achieve mAPs of 0.666 and 0.657, respectively. In comparison, YOLOv7 exhibits strong performance with box detection accuracy of 0.936 and a mask segmentation accuracy of 0.935. The recall score are 0.904 for box detection and mask segmentation is 0.903. Notably, the mAP result at the IoU threshold of 0.5 are 0.94 for box detection and mask segmentation is 0.941. Over the broader IoU spectrum of 0.5 to 0.95, the mAP was 0.677 for box detection and 0.659 for mask segmentation. To underscore the novelty of the approach, the performance of the proposed framework is systematically compared with established methods such as RCNN, Faster RCNN, and Mask RCNN.

INDEX TERMS Brain tumor, deep learning, image processing, MRI, YOLO.

I. INTRODUCTION

Brain cancer is a widespread devastating illness that results in numerous fatalities, even in developed nations. In the United States, the death toll is particularly alarming, with approximately 20,000 lives lost to this formidable disease [1]. It is characterized by the uncontrolled growth and spread of cells within the body. The human body normally regulates cell growth and multiplication, generating new cells through cell division. When an old or damaged cell dies, they are promptly replaced by new ones. However, in cases where this well-organized process fails, abnormal or damaged cells can proliferate inappropriately. These cells can manifest as tissue masses called tumors. Primary brain tumors, for instance, originate in the brain and typically remain localized there [2], [3]. On the other hand, secondary brain tumors

The associate editor coordinating the review of this manuscript and approving it for publication was Wenming Cao.

begin elsewhere in the body and subsequently spread to the brain [4].

Tumors can be classified into malignant and benign [5]. Malignant tumors can infiltrate neighboring tissues and metastasize, forming new tumors in distant body parts. In contrast, benign tumors do not invade nearby tissues and lack the capacity to metastasize. Once removed, benign tumors usually do not regenerate, unlike malignant tumors. Although benign tumors do not spread, they can occasionally attain significant size and pose severe symptoms or life-threatening risks. Human brain is a crucial for controlling actions and decision-making, acting as the central hub of the nervous system. Protecting it from harm is essential and among potential threats, tumors are a significant concern. Specific types of brain tumors, such as meningioma, glioma, and pituitary tumors, result from abnormal cell growth. Meningiomas typically develop in the thin membranes surrounding the brain and are primarily benign [6]. Despite their benign nature, they still present a life-threatening disease that directly impacts human life.

Meningiomas, which make up 36.1% of all primary tumors, tend to cluster around the brain's outer and upper contours. The meninges are the three layers of tissue that cover and protect the brain and spinal cord, and these growths emerge as bumps there. Meningioma may be diagnosed based on its anatomical location, form, and cellular appearance. These tumors develop slowly and might cause serious side effects, including convulsions and vision loss. However, gliomas are tumors that develop from the brain's glial cells rather than its neurons. They comprise a large subset of tumors with a wide range of symptoms and danger levels. In addition, the pituitary glands at the base of the skull are surrounded by cells, and tumors may form as aberrant masses there. Due to their closeness to the pituitary glands, which play a vital role in regulating hormone synthesis and release, these tumors can affect many biological systems.

The ability to successfully prevent and treat brain tumors depends on thoroughly understanding the disease's many phases. Radiologists have used various imaging modalities, such as X-ray, Magnetoencephalography (MEG), Computed Tomography (CT), Ultrasound, Electroencephalography (EEG), and Magnetic Resonance Imaging (MRI), to examine brain tumors and aid in accurate diagnosis and selection of appropriate treatment in recent years. Primary brain tumors vary significantly in size, location, and other features, making early diagnosis difficult. Accurate visualization of tumors by non-invasive imaging methods relies on the absorption characteristics of tissues, making it necessary to precisely delineate absorption rates for efficient tumor imaging [7]. When comparing various imaging options, magnetic resonance imaging (MRI) stands out as the gold standard because it can provide information about the brain in both healthy and diseased states. In the event of anomalies, it aids in pinpointing the precise tumor type [8]. However, analyzing MRI scan results requires meticulous observation and a high level of proficiency, making it impossible for an ordinary person to perform. Unfortunately, some hospitals and healthcare centers still lack the expertise to meet these requirements, further contributing to the lengthy diagnostic process [9].

The rapid advancements in machine learning and computer vision have given rise to power solutions known as Convolutional Neural Networks (CNNs). These cutting-edge models have successfully addressed complex Computer-Aided Diagnosis (CAD) problems, encompassing recognition, classification, segmentation, and even detection [10], [11], [12], [13], [14]. However, many existing CAD solutions for brain tumor detection and identification relying on CNNs suffer from inefficiency across various platforms and demand substantial computational resources. The lighter variants of CNN classification models have inherent limitations when pinpointing the tumor's exact location precisely [15]. On the other hand, segmentation models can accurately identify the affected area using a mask, thus enabling tumor localization, albeit at the expense of higher computing costs. This poses challenges when employing ordinary devices, leading to inefficient performance and unsatisfactory results.

Several object detection algorithms have been developed to enhance the detection of various objects in images [16], [17]. Notable algorithms include the Single Shot Multibox Detector (SSD) [18], R-CNN [19], and Fast R-CNN [20]. Some recent literature also explores the use of self-supervised deep learning models to enhance performance by leveraging aggregated information related to semantics and position [21], [22]. However, the You-Only-Look-Once (YOLO) algorithm [23] has gained considerable attention due to its exceptional object detection system and employs single unified neural network. This algorithm revolutionized object detection by treating it as a regression problem, estimating bounding box coordinates and class probabilities directly from pixel-level information. It enables the simultaneous prediction of multiple bounding boxes and class probabilities, having higher performance in terms of both speed and accuracy. As a result, it has found widespread application in the medical field, such as its ability to accurately categorize and locate abnormalities in brain images, as demonstrated in this study.

In this work, we employed two variant of YOLO deep learning models, namely YOLOv5 and YOLOv7 that provides a more efficient solution for segmenting and classifying brain tumors from MRI scans. The main contributions of this manuscript are as follows:

1) Dataset alignment: Figshare brain tumor dataset [24] is used to train and validate the models. Henceforth, the dataset will be denoted as the BT dataset throughout the remainder of the article. Mask alignment scheme has been employed on the images in the dataset and are aligned using segmentation coordinates to ensure a standardized format, enhancing analysis outcomes and training for tumor detection.
2) Tumor segmentation and classification: (a). To the best of our knowledge, this study stands as a pioneering effort, being the first to utilize the BT dataset [24]

for training both YOLO variants. Additionally, it is noteworthy to highlight the absence of any documented instances of YOLOv7 implementation in the existing literature for brain tumor segmentation and classification. (b). Furthermore, YOLO V5 has not previously been utilized with the selected dataset, and its implementation has been included to validate the results by comparing them with the state of the art.

3) Evaluation metrics: To assess the performance of deep learning models in handling multiple-class MRI scans of brain tumors, various evaluation metrics are employed, including precision, recall, f1-score, and mean Average Precision (mAP) at different thresholds (mAP@0.5 and mAP@ 0.5 to 0.95). These metrics allow for a comprehensive comparison of the performance between the two YOLO variants.
4) Finally, the proposed framework is systematically compared with established methods such as RCNN, Faster RCNN, and Mask RCNN, to highlight the contribution of this work.

The remaining paper is organized as follows: Section II presents a systematic literature review, the scope of this paper, and highlights the research gap. Section III presents the research methodology for deep learning based on brain tumor detection and classification. This section covers data collection, preprocessing, deep learning models, and performance analysis methods. Section IV presents the numerical results and discusses the performance achieved by the considered deep learning strategies. In Section V, the performance of the proposed deep learning framework is compared with state-of-the art results from the literature and presents future work. Finally, Section VI serves as the concluding section, summarizing the key finding and insights.

II. LITERATURE REVIEW

Timely detection of brain tumors is challenging, but the evolution of deep learning algorithms has shown promising results in accurately detecting brain tumors using digital images. MRI images, CT scans, and EEG are standard methods used to detect brain tumors. MRI and CT scans combined with deep learning models have shown improved accuracy in correctly detecting brain tumors. Deep learning models based on CNNs [6], [7], [25], Recurrent Neural Networks [26], Auto-Encoders (AEs) [27], [28], [29], [30], transfer learning and hybrid algorithms [31], [32] are explored heavily by researchers in recent years because of their ability for early detection and high accuracy. CNNs, in particular, have become a choice of researchers for detecting and classifying brain tumors.

BrainMRNet [33] used an end-to-end model to detect tumors using MR images and CNN, in which dataset contains 253 images, with 155 tumors and 98 normal images. The model is evaluated in terms of standard parameters and achieved 96% accuracy, 92% precision, and 96% sensitivity. Similarly, a research work in [25] has proposed a 3D CNN that validates extracted features through feedforward NN as a classification model. The feature extraction process was accomplished using a pre-trained VGG19 network. Their model is trained and evaluated using BraTS 2015, 2016, and 2018 dataset with 98%, 96%, and 92% respective accuracy. Their work achieved higher accuracy in some cases, but the model result varies depending on the dataset.

A patch-based DNN is proposed in [34] to classify brain tumors using MR images. The proposed model is evaluated extensively using eight datasets, including BraTS 2012 to 2015, Ischemic Stroke Lesion Segmentation (ISLES) 2015 and 2017, and MICCAI. The model achieved a high Dice Similarity Coefficient of 99.8% on BraTS 2013 dataset and showed a consistent performance across different modalities and datasets. The authors in [35] have proposed an automatic system for detecting brain tumors using the segmentation method and DNN. The proposed model used MR images for training consuming the BraTS 2013 dataset. The model explored a unique method where local and global features are consumed for diagnosis. Their model achieved an 85% dice score, 93% specificity, and 80% sensitivity

The research work in [36] proposed a deep learning model to detect Gliomas type of brain tumor using a hybrid CNN consisting of two and three-path networks. Their model is evaluated on BraTS 2013 dataset and achieved 0.86 (dice score), 0.86 (sensitivity), and 0.91 (specificity). The dataset contains scans of 20 patients with four modalities resulting in a total of 80 images. A recent study has proposed a decision support system (DSS) using a pre-trained Densenet201 deep learning model for multi-class brain tumor classification [37]. Two different metaheuristic algorithms are employed for feature selection in the proposed DSS. The model is trained and evaluated using BraTS 2018 and BraTS 2019 datasets and achieves a high accuracy of 95% using the SVM classifier.

The BT dataset used in this study has been explored by various researchers for brain tumor classification [38], [39], [40]. For instance, the research work in [38] uses a customized dual suppression encoding and factorized bilinear encoding plugged in to standard CNN model. The model is heavily tested on different dataset and achieved an accuracy of 95%. Similarly, the research in [39] proposed a residual network based on transfer learning for the classification of brain tumor. The model is checked in terms of accuracy with standard models including ConvNet, AlexNet, and VGG 16. The proposed model achieved accuracy of 95% using the BT dataset.

Many researchers have applied the YOLO model for object detection in medical images because of its speed and performance [41], [42], [43], [44]. The YOLOv5 object detection model, integrated into a portable Microwave Head Imaging system (MWHI), autonomously classifies and detects human brain abnormalities. Using 400 RMW image samples, encompassing non-tumor and tumor instances at various locations, the YOLOv5l model achieves notable metrics: 96.32% accuracy, 95.17% precision, 94.98% sensitivity, 95.28%

TABLE 1. Summary of tools, technologies, and models used for Brain Tumor segmentation and classification.

Ref.	Model	Dataset	Mechanism	Performance	Limitation/Remarks
[25]	3D CNN	BraTS 2015(192 Train, 82 Test), BraTS 2017(285 Train, 146 Test), BraTS 2018(285 Train, 191 Test)	Classification	98% Accuracy (2015) 96% Accuracy (2017) 92% Accuracy (2018)	Variance in result depending on train-test ratio and dataset
[33]	BrainMRNet (CNN-based network)	253 images, 2 classes, dataset containing MR images from field expert	Classification	96% Accuracy 92% Precision 96% Recall	Outperforms standard models. Small dataset
[34]	Patch-based DNN	BraTS2015 (274 Train, 110 Test)	Segmentation	95.1% Accuracy 97.2 % Precision 95.2% Recall	Consistent results across various datasets and modalities
[36]	Hybrid CNN	BraTS2013	Segmentation	85% Dice score 86% Recall 91% Specificity	High Complexity due to number of phases required for training
[38]	Bilinear network with customized FBE	BT dataset with two other dataset	Classification	Simulation results lack information for accuracy, precision and recall, specifically when considering all the classes collectively.	Interesting work but inconsistent performance across different datasets
[39]	Deep residual network with transfer learning	BT dataset	Classification	95% Accuracy 94% mAP 95% Recall	Lower accuracy compared to ResNet18 and ResNet101, High Complexity
[45]	Deep Learning Model YOLOv5	Reconstructed microwave brain images , 400 samples , 2 classes (Benign and Malignant)	Classification	95.3% Accuracy 95.17 % Precision 94.98% Recall 95.23% Specificity	Suitable for real-life application
[47]	ResNet-50 v2	BT dataset	Classification	99.68% Accuracy	Limited application due to high complexity
[48]	Automated Tumor Detector	BT dataset	Segmentation only	95% Accuracy	Interesting work but inconsistent performance across different datasets
[49]	CNN and Enhanced Sparrow Search Algorithm	BT dataset	Segmentation only	94.77% Accuracy 97.15 % Recall 67.16 Specificity	Limited application due to high complexity

specificity, 95.53% F1-score and 96.12% mean average precision (mAP). Similarly, the research work in [46] has applied YOLOv5 using a transfer learning approach to detect the brain using using BraTS 2020 dataset. The proposed model achieved 82-92% accuracy depending on the YOLO version, with YOLOv5 achieving the highest accuracy. Their work also suggests the tradeoff between the achieved accuracy and training time. Table 1 summarizes the significant research, existing tools, technologies, and models utilized to classify and detect brain tumors.

The literature review demonstrates the utilization of various object detection models in conjunction with classification models for detecting and classifying brain tumors. The accuracy of these models depends on factors such as the complexity of the underlying algorithm, the dataset utilized, and the extracted features. Among various object detection models, YOLO has exhibited remarkable performance in object detection across diverse domains, including medical imaging. Lightweight CNN classification models, in particular, have limitations in accurately identifying small objects. Segmentation models like YOLO address this limitation, albeit at a higher cost. This research aims to provide a balanced solution by combining the latest object detection models with classification models to achieve optimal performance.

In [47], an innovative deep learning approach incorporating transfer learning techniques was introduced, resulting in an impressive accuracy rate of 99.68% using ResNet50-v2. Furthermore, [48] developed an automated tumor identification and segmentation system, achieving a segmentation accuracy of 95% by utilizing mask images as labels. The work in [49] proposed a computer-aided system incorporating preprocessing, segmentation, and CNN-based classification, demonstrating improved efficiency compared to alternative techniques. However, it is worth noting that the provided illustrations portray some inaccurate in the segmentations.

The research work in [50] employed the Faster R-CNN algorithm for tumor detection, achieving an average precision of 77.60%. In [51], a CNN architecture was proposed, achieving a high accuracy rate of 96.56% for tumor classification. Lastly, [52] combined a pre-trained CNN with gray-level co-occurrence matrix features, resulting in a % accuracy rate of 96.5%. Previous work's lack of specific task details hinders a comprehensive analysis of overall performance. It becomes challenging to fully evaluate and compare the effectiveness of the related work and models by omitting essential metrics, such as recall or specific detection thresholds.

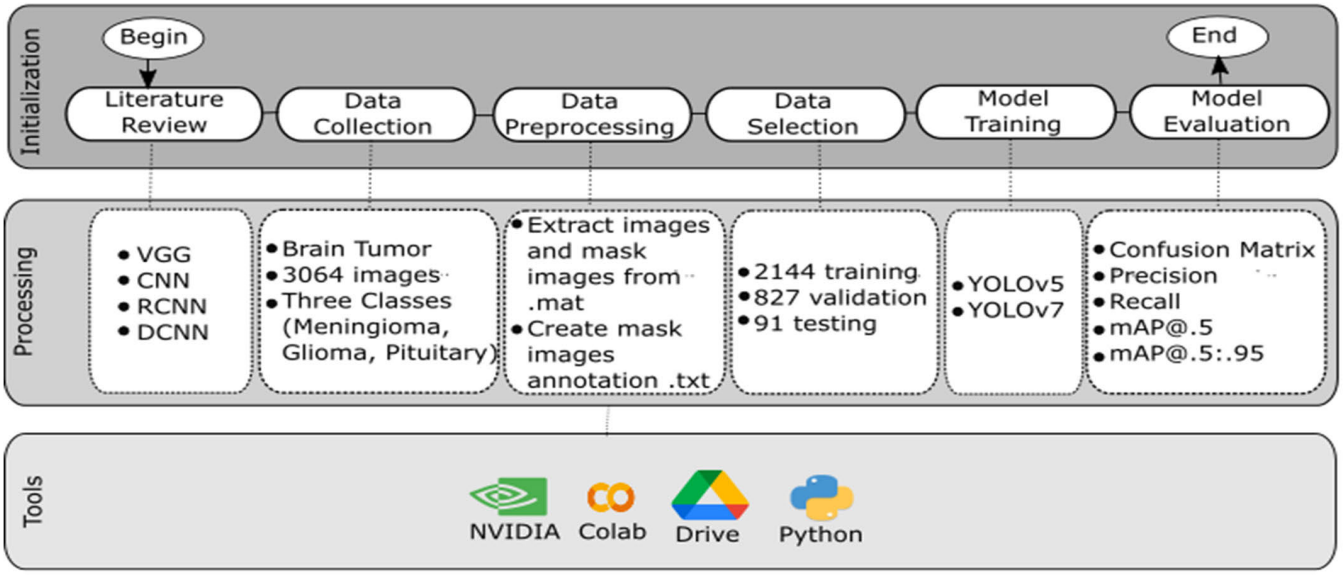


FIGURE 1. YOLO-based Brain Tumor Segmentation and Classification Research Flow Block Diagram.

III. PROPOSED METHODOLOGY

Figure 1 outlines the proposed research's key steps, starting with a thorough review of relevant literature, including the most recent papers, leading to dataset compilation. Subsequently, the data undergoes preprocessing and selection. The proposed model is trained and evaluated using standard parameters following these steps.

A. DATA COLLECTION

The Brain Tumor dataset of Southern Medical University, Guangzhou, is utilized for this research acquired from [24] and [53]. Figure 2 summarizes the dataset, revealing that it comprises 3064 images from 233 patients. All the images in the dataset belong to any of the three tumor classes: meningioma, glioma, and pituitary, with proportions of 23.11%, 30.35%, and 46.54%, respectively. Each slice is composed of 512 pixels.

B. DATA PREPROCESSING

The dataset is converted from .mat format to extract .png images and masks. Additionally, this study observes that the tumor border vector does not accurately align with the tumor in the images. As a result, the research extends to annotating the tumor using mask images. Figure 3 depicts the images and corresponding masks, which have been transformed from .mat to .png format. Subsequently, the masks are utilized to annotate the images using CV2 accurately.

Algorithm 1 presents the essential steps for converting .mat files to .png files. It begins by initializing the necessary paths for the dataset, output directory, and list.MAT files. Next, an outer for loop iterates over each file in the list (step 3). Within the loop, the transformation from MATLAB ".mat" images and masks to .png format is performed (step 4). Following the transformation, the brain tumor polygonal segmentation coordinates are extracted. This step involves utilizing image processing and analysis techniques to identify and extract the coordinates representing the tumor regions (step 5). Subsequently, an inner for loop is entered, iterating over the segmentation coordinates (step 6). Each coordinate in the list is converted to the YOLO format. This process entails transforming the coordinates into normalized values relative to the image dimensions and writing them into a .txt file (step 7). Once all the coordinates have been processed, the inner loop is exited (step 8). Finally, the outer loop is exited, completing the conversion and coordinate extraction for all the .mat files in the dataset (step 9).

Algorithm 2 aims to align the segmentation coordinates to a standardized format to enable further analysis and training for tumor detection tasks. It consists of preprocessing and correction steps by performing morphological closing on a

Algorithm 1 Pseudo Code for Converting .mat to .png Image

Step* 1** → ***Begin
Step* 2** → ***Initialize *path to the dataset, path to the output directory and a list of .mat files*
Step* 3** → ***Outer for Loop: *for file in files:*
Step* 4** → ***Transform : *MATLAB.*mat*images and masks to .png*
Step* 5** → ***Extract : *Brain tumor polygonal segmentation coordinates*
Step* 6** → ***Inner for loop : *for index, element in enumerate* (*seg_coords*):
Step* 7** → ***Convert : *The extracted coordinates to.txt YOLO format*
Step* 8** → ***End : *Inner for loop*
Step* 9** → ***End : *Outer for loop*
Step* 10** → ***End

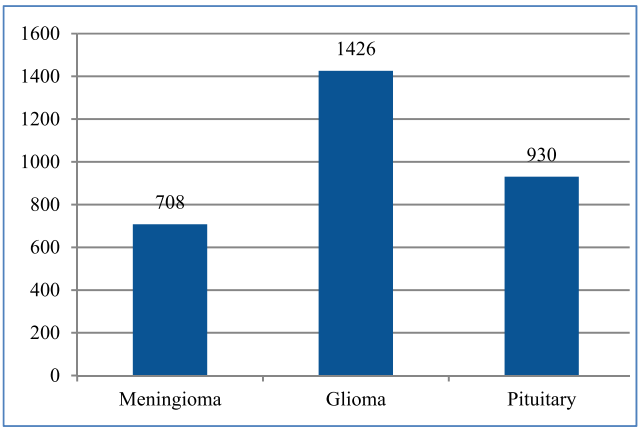


FIGURE 2. Distribution of Brain Tumor Slices in the Dataset.

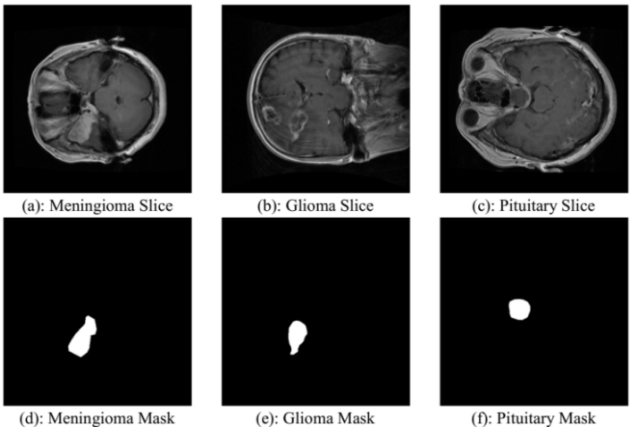


FIGURE 3. Brain Tumor Slices and Masks.

binary mask image (step 3) which closes small holes in masked brain tumor, finding tumor contours (step 4), approximating contours with polygons (step 6), and converting the coordinates to the YOLO format (step 8). The aligned segmented annotations are saved in a .txt file, completing the preprocessing and correction step.

Algorithm 2 Pseudo code for correction of annotations

Input: *Binary Mask Image*
Output: *Aligned Segmentation Coordinates in .TXT*
Step 1 → ***Begin***
Step 2 → ***Initialize*** *Provide binary mask image of brain tumor, list of coordinates*
Step 3 → ***Perform***: *Morphological closing* $C(x, y) = E(D(f(x, y)), B)$
Step 4 → ***Find***: *Contours of tumor in image*
Step 5 → ***for loop***: *for contour in contours*:
Step 6 → ***Approximate***: *Contour with polygon*
Step 7 → ***Append***: *Polygonal coordinates to list of coordinates*
Step 8 → ***Convert***: *List of coordinates to YOLO annotation format*
Step 9 → ***End***: *for loop*
Step 10 → ***End***

The results of Algorithm 2 are illustrated in Figure 4, where 4(a) shows sample images before correction, and 4(b) displays the images after modification. It is evident that Algorithm 2 effectively aligns and corrects the labels and coordinates accordingly.

C. DATA SELECTION

After the necessary preprocessing, the entire dataset is split into three portions for model training, validation, and testing. The training set comprises 2144 images, which account for 69.97% of the data. The validation set contains 827 images, representing 26.99% of the data. The testing set consists of 91 images comprising 2.97% of the data. Note that all three portions of the dataset have additional mask annotations.

D. MODEL TRAINING

Two variants of the YOLO framework, i.e., YOLOv5 and YOLOv7, are employed in this research work for brain tumor segmentation and classification. The selection of these variants is aimed at facilitating a comprehensive assessment of their capabilities and gauging their effectiveness in addressing the challenge of tumor detection and classification.

1) YOLO V5 MODEL

The YOLOv5 is a robust architecture for object detection. It consists of 225 layers and 7,413,608 parameters. It is a deep convolutional neural network type, where various layers extract meaningful features, and object localization is achieved within images, as shown in Figure 5.

The model begins with a series of Convolutional (Conv) layers, which learn low-level features from the input images. These layers are followed by C3 layers, which help in feature aggregation and information fusion. The C3 layers, responsible for collecting more general and high-level information, use larger filters. Down-sampling is achieved using Conv layers with stride and kernel size settings to improve the model's capacity to collect multi-scale features by lowering the spatial dimensions of the feature maps. The input picture is down-sampled so that the model can better identify items across various sizes. A Spatial Pyramid Pooling Fusion (SPPF) component is also a part of the YOLOv5 design. The SPPF component is built to execute pooling operations of varying sizes on the feature maps, thereby capturing contextual information at various scales.

This broadens the model understanding of the images contents, which improves the accuracy with which it can recognize things. The YOLOv5 model last levels are the Conversion (Conv) and Concatenation (Concat) layers. The Conv layers aid in the fine-tuning of the features, while the Concat layers combine multi-scale feature maps to create detailed object representations. The YOLOv5 model final layer, called YOLOv5 Segment, performs the actual object detection task.

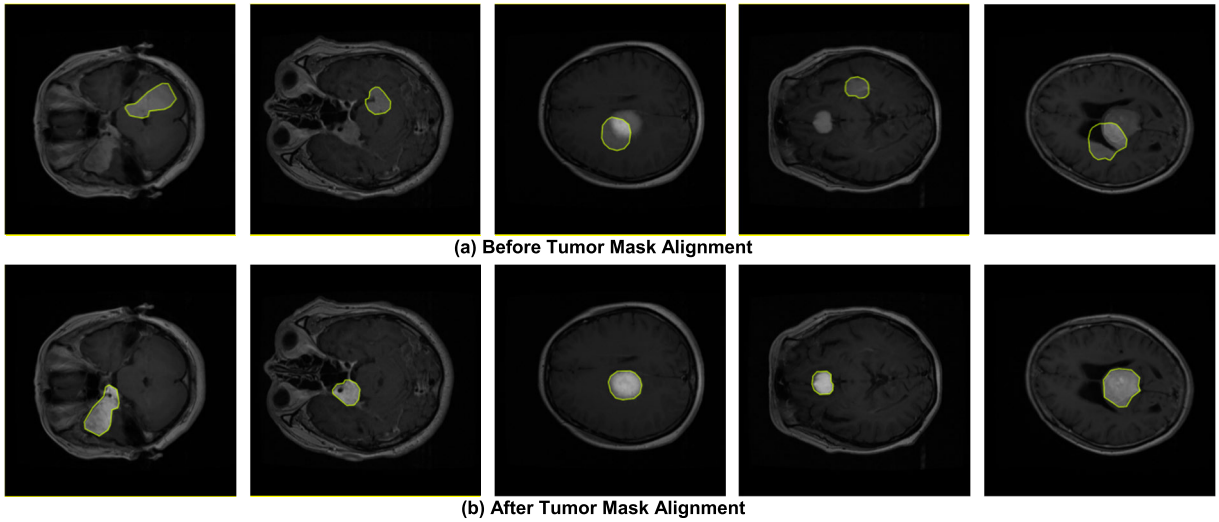


FIGURE 4. Tumor Mask Alignment for Correction of Annotations or Label.

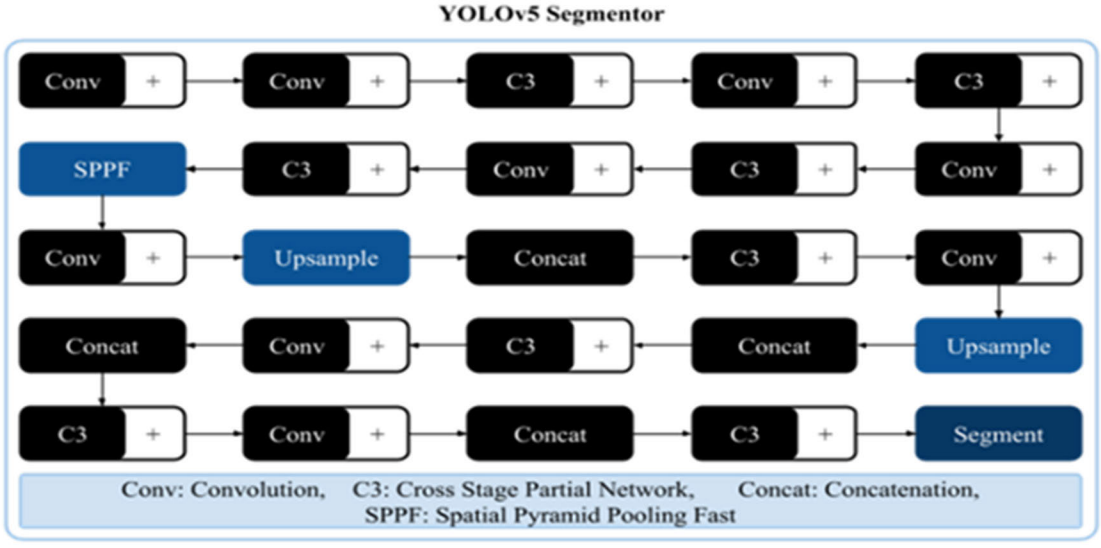


FIGURE 5. YOLOv5 Architecture for Brain Tumor Segmentation and Classification.

It generates bounding box predictions for three scales, allowing the model to detect objects of various sizes and aspect ratios. These bounding boxes are predicted based on anchor boxes as illustrated in Figure 6.

The mathematical formulas used to predict the bounding boxes in YOLO which involves determining the width and height as offsets from cluster centroids and the center coordinates relative to the filter application location using a sigmoid function [54]. In particular, the network predicts five coordinates for each bounding box, t_x, t_y, t_w, t_h and t_o.If the cell is positioned with an offset from the top left corner of the image by (C_x, C_y), and the bounding priors have widths and heights represented as p_w, p_h, then the predictions correspond to b_x, b_y, b_w, b_h, as depicted in the Figure 6.

The pseudo code of the employed model to segments meningioma, glioma, and pituitary for the given input images. It starts by initializing the segmenters optimized weights, image dimensions, and confidence score (step 2). Subsequently, each layer is executed, which includes the convolutional layer (step 6), C3 layer (step 8), SPPF (step 10), upsample layer (step 12), concatenation layer (step 14), and the segment layer (step 16). After processing all the layers, the output is obtained as the segmented image of the any particular class of the brain tumor.

2) YOLO V7 MODEL

YOLOv7 is a refined variant of the YOLO framework, introducing a range of advancements to achieve highly efficient and accurate results. It incorporates several enhancements

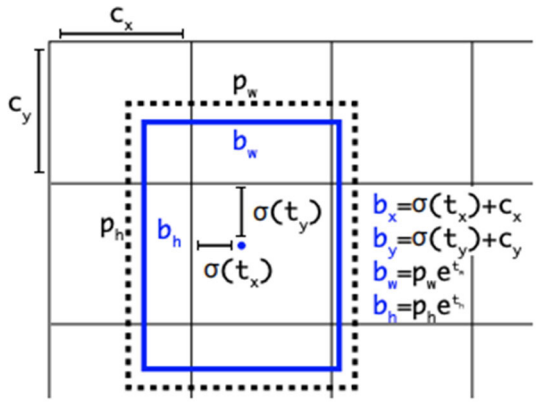


FIGURE 6. Bounding Box with dimension prior and location prediction [54].

Algorithm 3 Pseudo code for YOLOv5 Model for Brain Tumor Detection

Input: *Brain Tumor Dataset*
Output: *Segment Tumor (Meningioma, Glioma, Pituitary)*
Step 1 → ***Begin***
Step 2 → ***Initialize***: *Segmentor with best weights*, *image size*, *and confidence value*
Step 3 → ***Input***: ***Camera or image***
Step 4 → ***for loop***: *for layer in layers*:
Step 5 → ***if condition***: *if layer in Conv_layers*:
Step 6 → *Conv(layer)*
Step 7 → ***elseif condition***: *layer in C3*:
Step 8 → C3 (layer)
Step 9 → ***elseif condition***: *layer in SPPF*:
Step 10 → *SPPF*(*layer*)
Step 11 → ***elseif condition***: *layer in Upsample_layers*:
Step 12 → *Upsample (layer)*
Step 13 → ***elseif condition***:*layer in Concat_layers*:
Step 14 → *Concat*(*layer*)
Step 15 → ***elseif condition***: *layer in segment_layers*
Step 16 → *Segment* (*layer*)
Step 17 → ***Endif***
Step 18 → ***EndFor***
Step 19 → ***End***

to improve performance and handle objects across different scales effectively. It excels at capturing contextual information from diverse scales by integrating multi-scale feature fusion, which enhances object detection accuracy. Additionally, spatial pyramid pooling enables the model to extract information at multiple scales and effectively handle objects of varying sizes. Moreover, it integrates upsampling layers to facilitate the detection of smaller objects by enlarging the feature maps. It employs concatenation operations and multiple convolutional layers to iteratively enhance the representation of objects and elevate the precision of bounding box predictions.

Figure 7 shows the employed YOLOv7 segmentor architecture for brain tumor segmentation and classification. The segmentor consists of 82 convolution layers, 15 concatenation layers, five max-pooling layers, an SPPCSP layer, two upsample layers, and a segment layer as the foundation of the algorithm for brain tumor segmentation. Further, the segmentor is trained with 2144 images with 512 image size, 64 batch size, and yolov7-seg.pt weights for 100 epochs. The employed framework accurately segment meningioma, glioma, and pituitary tumors using using MR images.

The pseudo code of the emoloyed model to segment meningioma, glioma, and pituitary by using the given MR images. It begins by initializing the segmentor optimized weights, image dimensions, and confidence score (Step 2). Subsequently, each layer is executed based on its respective nature: the convolutional layer conducts feature extraction by convolving input data with filters (step 6), the concatenation layer merges feature maps from diverse network stages (step 8), the max-pooling layer partitions the input into non-overlapping regions and retains the maximum value within each region (step 10), the SPPCSPC layer effectively captures multi-scale information using the SPP layer and guarantees the proper flow of gradients and information across different stages of the network using the CSP connection (step 12), the upsampling layer enhances the spatial resolution of the feature maps (step 14), and the segment layer generates predictions for segmentation boundaries and class probabilities for the segmented objects (step 16). After processing all the layers, the output is obtained as the segmented image of the brain tumor.

Algorithm 4 Pseudo Code for YOLOv7 Model for Brain Tumor Detection

Input: *Brain Tumor Dataset*
Output: *Segment Tumor (Meningioma, Glioma, Pituitary)*
Step 1 → ***Begin***
Step 2 → ***Initialize***: *Segmentor with best weights*, *image size*, *and confidence value*
Step 3 → ***Input***: *Camera or image*
Step 4 → ***for loop***: *for layer in layers*:
Step 5 → ***if condition***: *if layer in Conv_layers*:
Step 6 → *Conv*(*layer*)
Step 7 → ***elseif condition***: *layer in Concat_layers:*
Step 8 → *Concat (layer)*
Step 9 → ***elseif condition***: *layer in MP_layers*
Step 10 → *MP* (*layer*)
Step 11 → ***elseif condition***: *layer in SPPCSPC_layer*
Step 12 → *SPPCSPC (layer)*
Step 13 → ***elseif condition***:*layer in Upsample_layers*
Step 14 → *Upsample* (*layer*)
Step 15 → ***elseif condition***: *layer in segment_layers*
Step 16 → Segment (layer)
Step 17 → ***End if***
Step 18 → ***End For***
Step 19 → ***End***

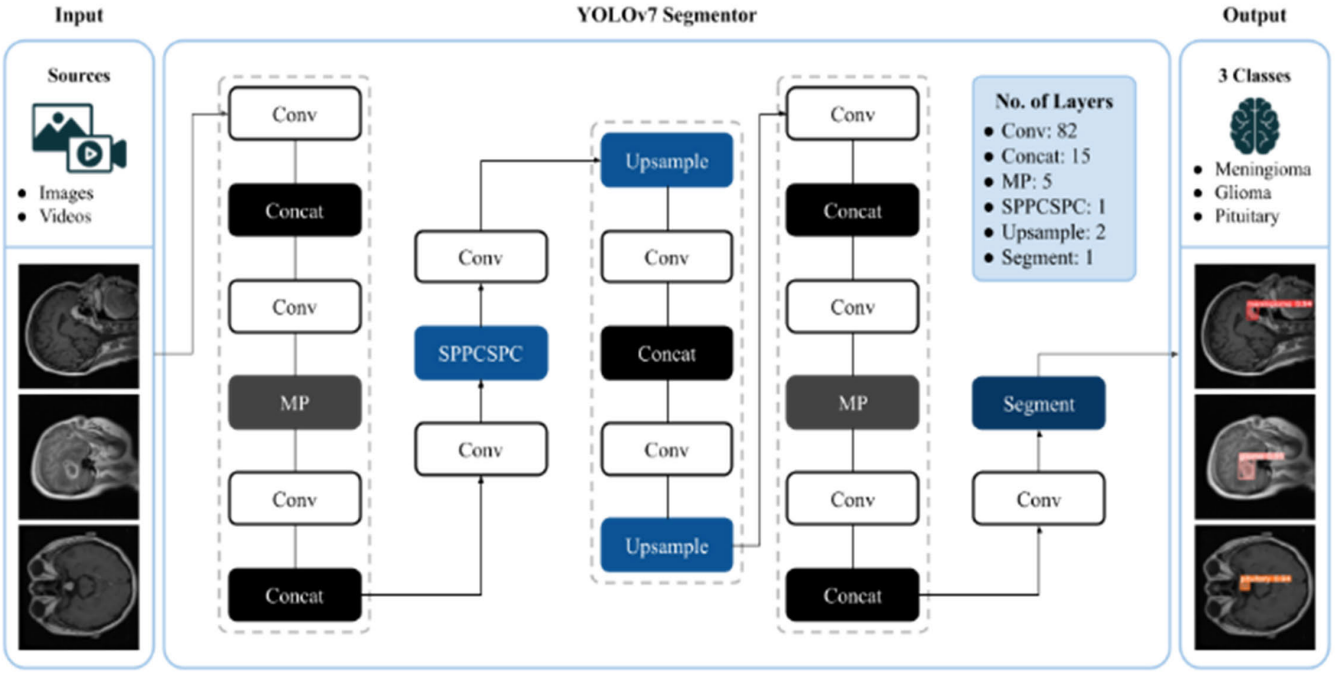


FIGURE 7. YOLOv7 Architecture for Brain Tumor segmentation and classification.

E. EVALUATION PARAMETERS

Precision, Recall, and F1 Score are widely used performance evaluation metric, which are discussed below.

1) CONFUSION MATRIX

The metric provide insights in to the precision and inaccuracies of the proposed models. Each element (i, j) within the matrix denotes the proportion of instances that are labeled as class "i" while predicted as a class "j".

2) LOSS

The loss function optimizes the training phase of the model to minimize the loss and is given as:

$$loss = l_{box} + l_{cls} + l_{obj} \tag{1}$$

where l_{box} represents the bounding box regression, l_{cls} represents the classification loss, and l_{obj} represents confidence loss. Equation (2) – (4) presents the loss for each category as follows:

$$l_{box} = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{obj} bj\,(2 - w_i \times h_i) \times \left[\left(x_i - \hat{x}_i^j\right)^2 + \left(y_i - \hat{y}_i^j\right)^2 + \left(w_i - \hat{w}_i^j\right)^2 + \left(h_i - \hat{h}_i^j\right)^2 \right], \tag{2}$$

$$l_{cls} = \lambda_{cls} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{obj} \sum_{c \in classes} p_i(c) \log\left(\hat{p}_l(c)\right), \tag{3}$$

$$l_{obj} = \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{noobj} \left(c_i - \hat{c}_l\right)^2 + \lambda_{obj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{i,j}^{obj} \left(c_i - \hat{c}_l\right)^2, \tag{4}$$

where coefficient $\lambda_{co??rd}$ is the weight given to positional loss, while coefficient λ_{cls} is the weight given to category loss. The variables $\hat{x}$ and $\hat{y}$ are the actual central coordinate of the target while $\hat{w}$ and $\hat{h}$ are the target width and height. If the anchor box located at position (i, j) contains targets, the corresponding value $I_{i,j}^{obj}$ is set to 1; otherwise, 0. The variable $p_i(C)$ denotes the probability of the target belonging to a specific category, and $p_l(C)$ is the actual category value. The lengths of both variables equal the total number of categories C.

3) PRECISION

This metric assess the correctness of the bounding box (bbox) predictions by calculating the proportion of true positive detections to the combined count of true positive and false positive using (5)

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

whereas TP and FP are the true positive and false positive values, respectively.

4) RECALL

It is determined by calculating the ratio of true positive detections to the total of true positive and false negative using (6)

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

whereas TP and FN are the true positive and false negative values, respectively.

TABLE 2. Loss performance of the YOLO models.

YOLO	T/box	T/seg	T/obj	T/cls	V/box	V/seg	V/obj	V/cls
v5	**0.017893**	0.017295	0.005704	0.000391	**0.026912**	0.03365	0.005083	0.0022
v7	0.012063	**0.014582**	**0.00412**	**0.000368**	0.025752	**0.036826**	**0.005529**	**0.001796**

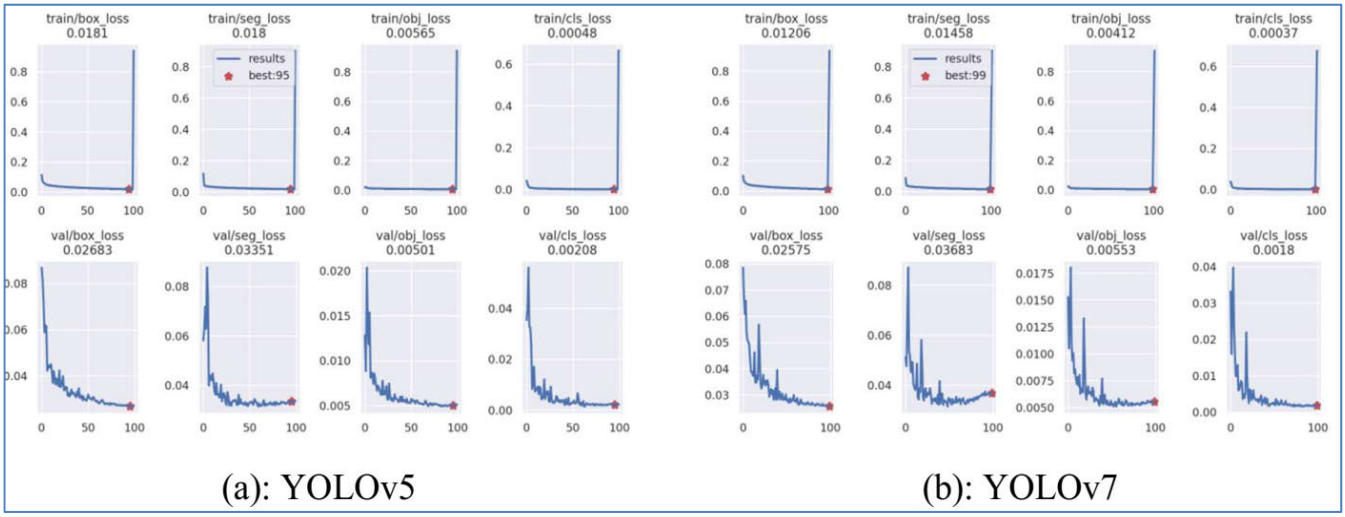


FIGURE 8. Loss Performance for the YOLOv5 and YOLOv7 for Brain Tumor segmentation and classification.

5) MEAN AVERAGE PRECISION(MAP@.5 & MAP@.5:.95)

The metric mAP@0.5 represents the mean Average Precision (mAP) at an Intersection over a Union (IoU) threshold of 0.5, while mAP@0.5:0.95 stands for the average mAP calculated across various IoU thresholds ranging from 0.5 to 0.95 and calculated using (7)

$$mAP = \frac{1}{n}\sum_{n=1}^{k=n} AP_k \tag{7}$$

where n are the number of classes and AP_k refers to the average precision of class k.

6) F1-CONFIDENCE CURVE

It is a graphical illustrates the correlation between F1 score and the confidence threshold for object detection. Understanding the tradeoff between precision and recall at different degrees of confidence may be facilitated by examining the F1-confidence curve.

IV. RESULTS AND DISCUSSION

This section provides all the necessary detail required to deploy the YOLO based models for brain tumor segmentation and classification. This section also outlines an in-depth assessment and findings for understanding the outcome of this study.

A. PRELIMINARY STUDY

Google Colab integrated development environment i.e., Google Colab Pro has been used for training and evaluating the performance of the deployed models for brain tumors with A100-SXM4-40GB GPU. The GitHub repositories of YOLOv5 and YOLOv7 are cloned to Google Drive. The models are evaluated using a dataset of 827 images of brain tumors as discussed in section III (c). Both models have been trained using hyper parameters with the stochastic gradient decent learning algorithm, employing a learning rate of 0.001 over the course of 100 epochs.

B. LOSS PERFORMANCE

Table 2 depicts the training loss of box, segmentation, object, and classification in brain tumor detection. In box training loss (T/box), YOLOv5 has slightly high than YOLOv7 with values of 0.017893 and 0.012063, respectively. In segmentation training loss (T/seg), YOLOv7 performs better than YOLOv5 with 0.017295 compared to 0.014582. YOLOv5 also exhibits classification training loss (T/cls) and object training loss (T/obj) compared to YOLOv7 with values of 0.000391 and 0.005704 for YOLOv5, and 0.000368 and 0.00412 for YOLOv7, respectively. Regarding validation loss, YOLOv5 shows slightly higher values for box validation loss (V/box) and classification validation loss (C/cls). The segment validation loss (V/seg) compared to YOLOv7 is low. However, YOLOv7 outperforms YOLOv5 in box validation loss (V/box) and classification validation loss (V/cls) with lower values. Figure 8 displays the data in Table 2 graphically to facilitate a clear understanding of the employed models. To maintain brevity, we refrain from presenting additional results obtained from the analysis.

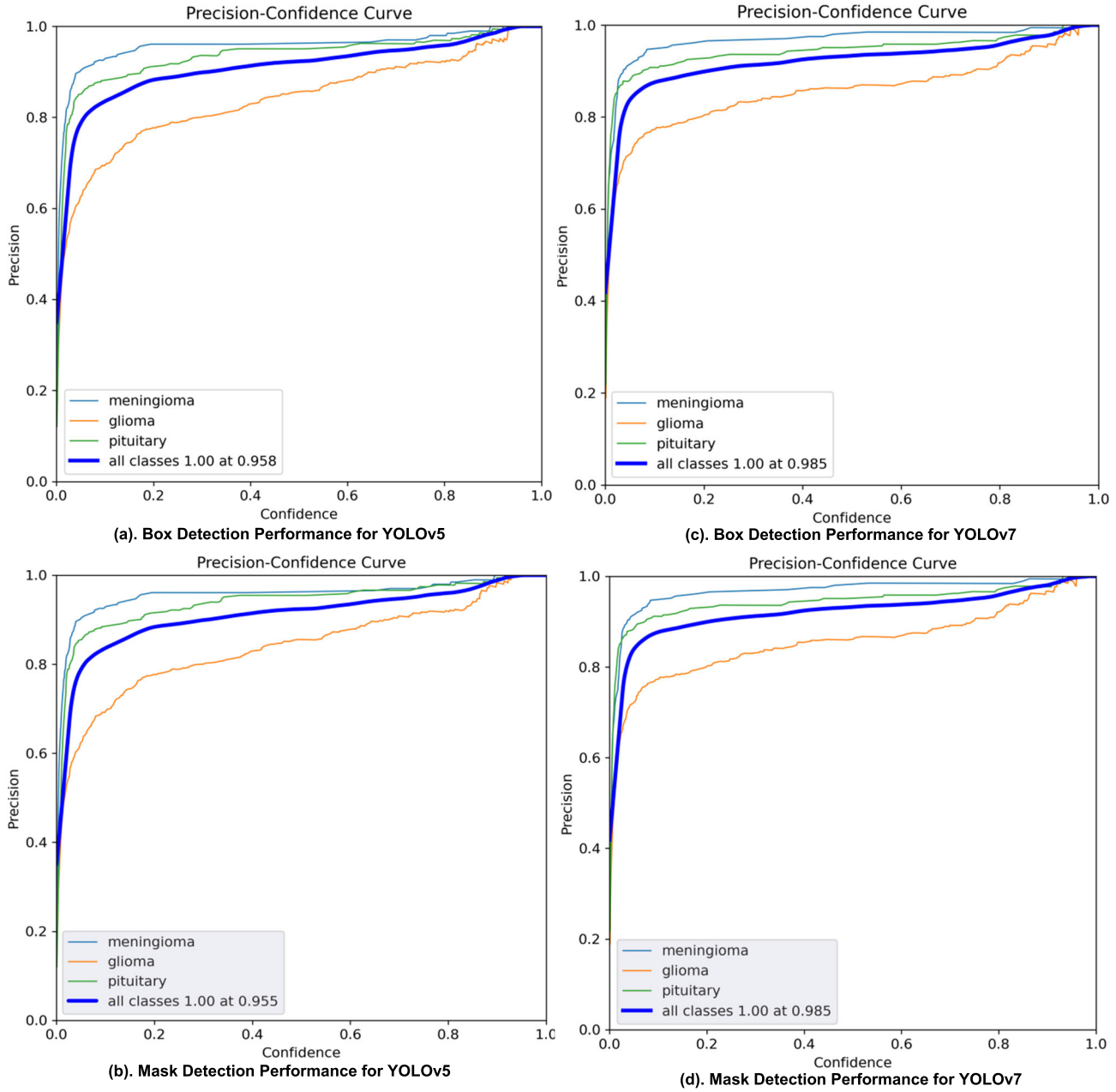


FIGURE 9. Precision confidence curve for YOLO models for Brain Tumor segmentation and classification.

C. PRECISION PERFORMANCE

The performance evaluation of YOLOv5 and YOLOv7 on 827 labeled boxes and mask images are presented in Table 3. Meningioma, glioma, and pituitary brain tumors are evaluated for the precision metric. The precision measures provide the model ability to accurately identify the regions of interest (ROI) in images. For all the classes, YOLOv5 has a precision score of 0.94 for box detection and 0.936 for mask detection, while YOLOv7 has a precision score of 0.936 for box detection and 0.935 for mask detection.

For meningioma, both YOLOv5 and YOLOv7 have the highest precision scores of 0.965 for box detection and 0.985 for mask detection. For glioma, YOLOv5 has a box detection precision score of 0.893 and a mask detection precision score of 0.879, while YOLOv7 has a box detection

TABLE 3. Precision performance of YOLO models.

Class	Images	Labels	YOLOv5		YOLOv7	
			Box	Mask	Box	Mask
All	827	827	**0.94**	**0.936**	0.936	0.935
Meningioma		202	0.965	0.965	**0.985**	**0.985**
Glioma		382	**0.893**	**0.879**	0.87	0.867
Pituitary		243	**0.962**	**0.964**	0.953	0.953

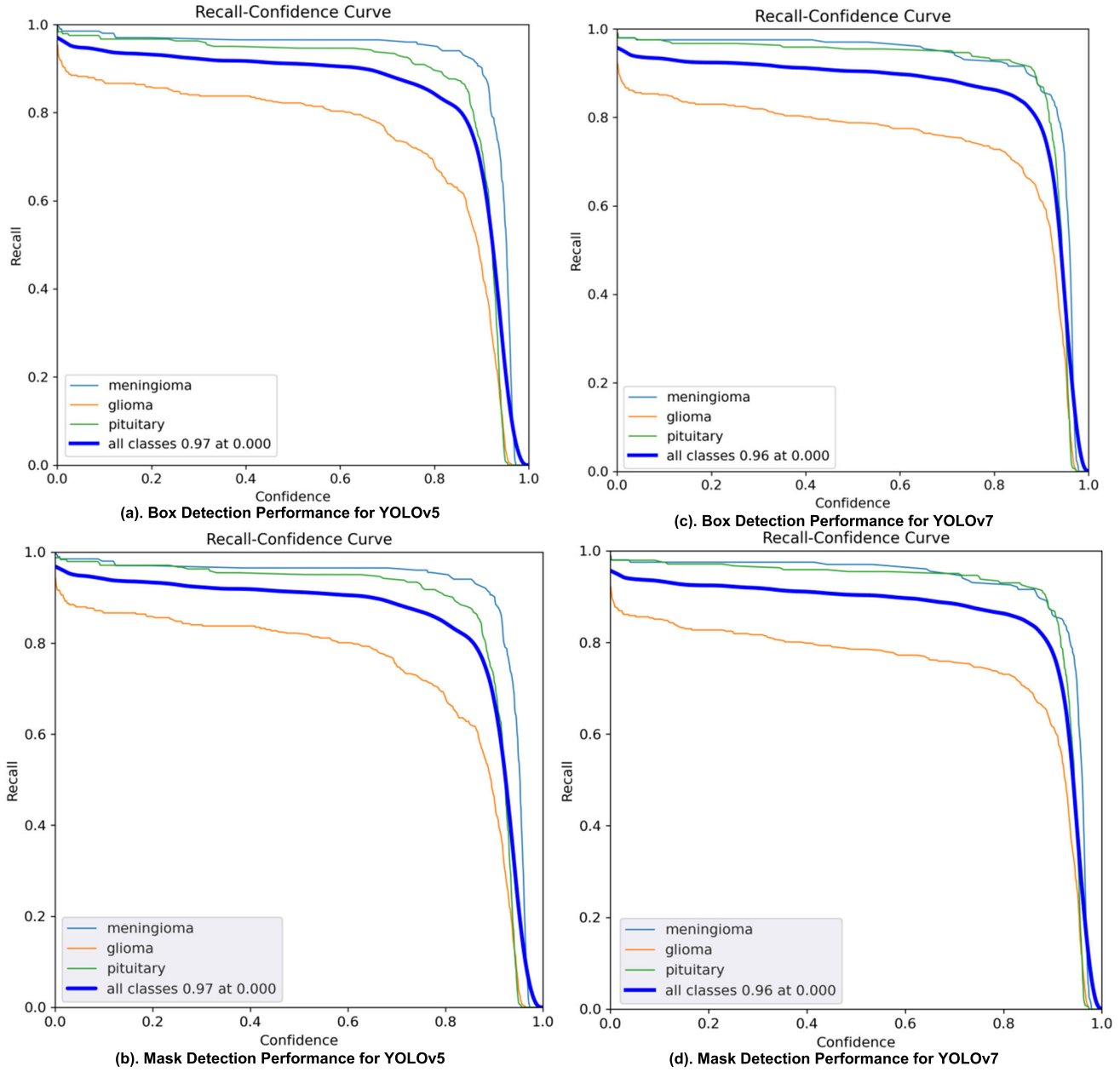


FIGURE 10. Recall-confidence curve for YOLO models for Brain Tumor segmentation and classification.

precision score of 0.962 and a mask detection precision score of 0.964, while YOLOv7 has box detection precision score of 0.953 and a mask detection precision score of 0.953.

Figure 9 depicts the precision-confidence curves for YOLOv5 and YOLOv7, showing the exceptional performance of both models; overall, with one exception in the glioma class, which is represented by the orange line. The score for 100% precision is at a confidence value of 0.955 and 0.985 for YOLOv5 and YOLOv7, respectively.

D. RECALL PERFORMANCE

Table 4 presents the recall performance evaluation of YOLOv5 and YOLOv7 on 827 labeled boxes and masks images. For all the classes, YOLOv5 has a recall score of 0.905 for box detection and 0.906 for mask detection, while YOLOv7 has a recall score of 0.904 for box detection and 0.903 for mask detection. For meningioma, both YOLOv5 and YOLOv7 have the highest recall scores of 0.965 for box detection and 0.97 for mask detection. For glioma, YOLOv5 has a box detection recall score of 0.799 and a mask detection recall score of 0.801, while YOLOv7 has a box detection recall score of 0.787 and a mask detection recall score of 0.785. For the pituitary, both YOLOv5 and YOLOv7 have high recall scores of 0.951 for box detection and 0.955 for mask detection. Refer to the Table 4, for more results clarity.

Figure 9 shows the box and mask recall-confidence curves of YOLOv5 and YOLOv7, which depict that both models have good performance for two classes i.e., meningioma and

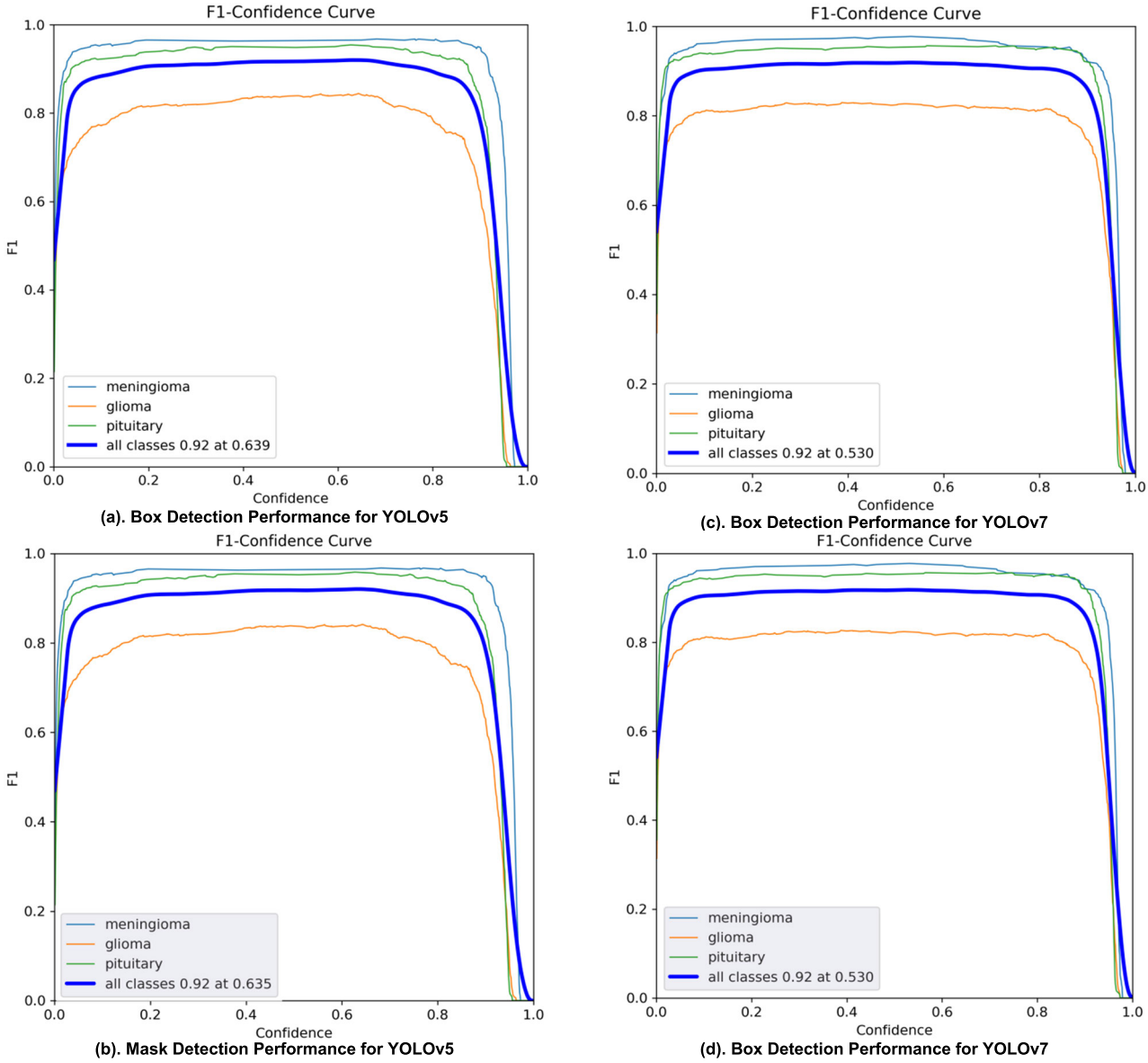


FIGURE 11. F1-confidence curve for YOLO models for brain tumor segmentation and classification.

TABLE 4. Recall performance of the YOLO models.

Class	Images	Labels	YOLOv5		YOLOv7	
			Box	Mask	Box	Mask
All	**827**	**827**	**0.905**	**0.906**	0.904	0.903
Meningioma		**202**	0.965	0.965	**0.97**	**0.97**
Glioma		**382**	**0.799**	**0.801**	0.787	0.785
Pituitary		**243**	0.951	0.951	**0.955**	**0.955**

TABLE 5. Mean average precision (MAP@.5).

Class	Images	Labels	YOLOv5		YOLOv7	
			Box	Mask	Box	Mask
All	827	827	**0.947**	**0.947**	0.94	0.941
Meningioma		202	**0.99**	**0.99**	0.985	0.985
Glioma		382	**0.872**	**0.872**	0.86	0.862
Pituitary		243	**0.978**	**0.978**	0.976	0.975

pituitary. However, the recall-confidence curve for glioma indicates slightly poor performance, highlighted as orange line in the Figure 9.

E. MEAN AVERAGE PRECISION (mAP@.5)

Table 5 shows the mean Average Precision (mAP) for the YOLOv5 and YOLOv7 models on 827 labeled boxes and masks images, evaluated at the IoU threshold of 0.5. For all the classes combined, YOLOv5 has a mAP score of 0.947 for box detection and 0.947 for mask detection, while YOLOv7 has a mAP score of 0.94 for box detection and 0.941 for mask detection. For meningioma, both YOLOv5 and YOLOv7 have the highest mAP scores of 0.99 for box detection and 0.985 for mask detection.

For glioma, YOLOv5 has a box detection mAP score of 0.872 and a mask detection mAP score of 0.872, while

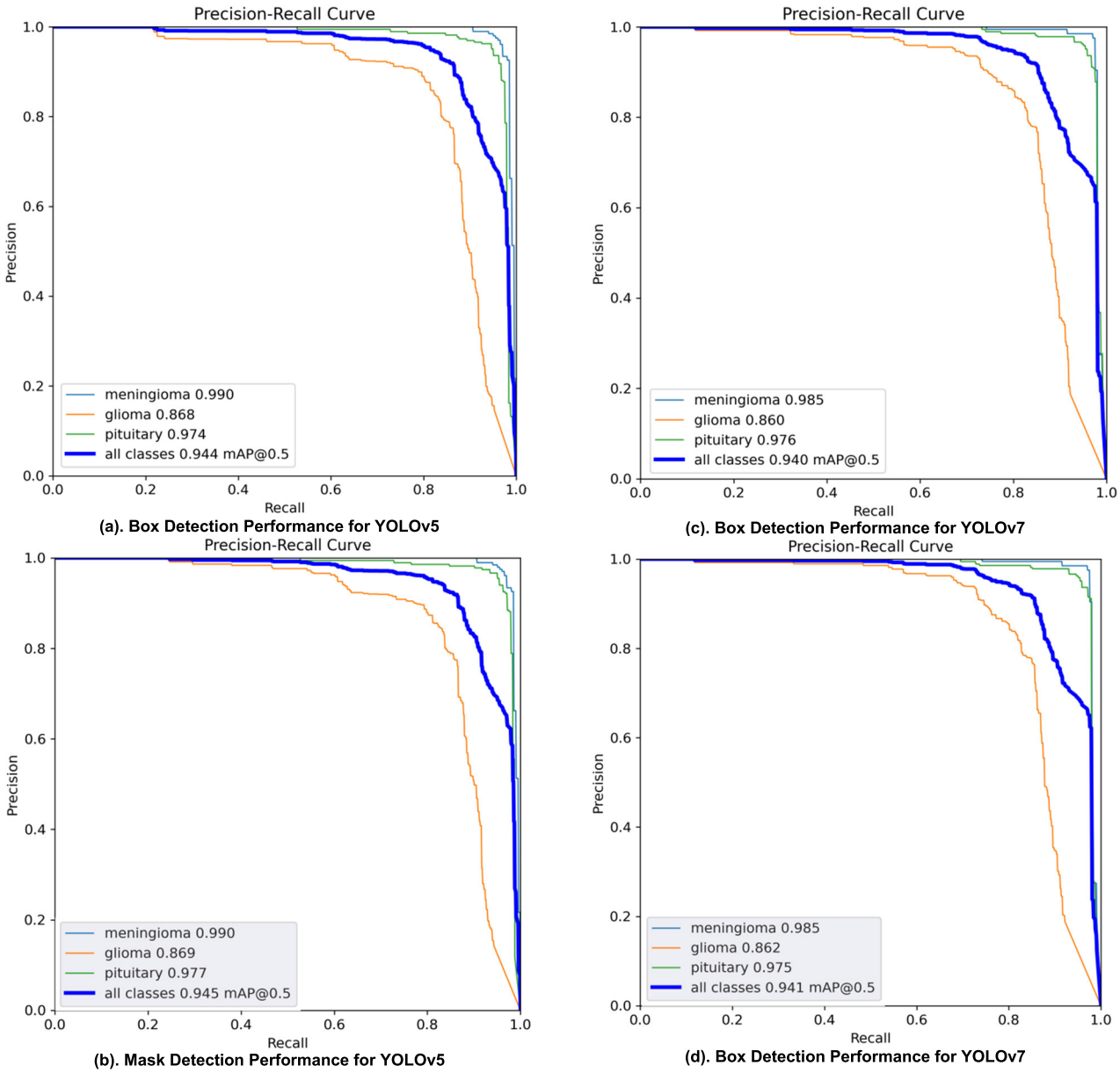


FIGURE 12. Precision-recall curve for YOLO models for brain tumor segmentation and classification.

YOLOv7 has a box detection mAP score of 0.86 and a mask detection mAP score of 0.862. For pituitary, YOLOv5 has a box detection mAP score of 0.978 and a mask detection mAP score of 0.978, while YOLOv7 has a box detection mAP score of 0.976 and a mask detection mAP score of 0.975.

Figure 11 presents the F1-confidence curves for YOLOv5 and YOLOv7. Both the models perform good for two classes i.e., meningioma and pituitary. However, the glioma has a low curve indicating poor performance relating to other classes. Further, it depicts that the ideal box confidence value is 0.639 and 0.53 while mask confidence values are 0.635 and 0.53 for YOLOv5 and YOLOv7, respectively, to achieve a 0.92 f1-score.

F. MEAN AVERAGE PRECISION (mAP@.5:.95)

Table 6 shows the mAP for the YOLOv5 and YOLOv7 models on 827 labeled boxes and masks images, evaluated at IoU threshold of 0.5-0.95. The mAP@.5:.95 score means that the model considers an object detected if it has an IOU threshold score between 50- 95%.

The mAP@.5 score for all the classes for YOLOv5 has 0.666 for box detection and 0.657 for mask detection, while YOLOv7 has mAP@.5 score of 0.677 for box detection and 0.659 for mask detection. For meningioma, both YOLOv5 and YOLOv7 have the highest scores of 0.79 for box detection and 0.799 for mask detection. For glioma, YOLOv5 has a score of 0.546 for box detection and 0.53 for mask detection,

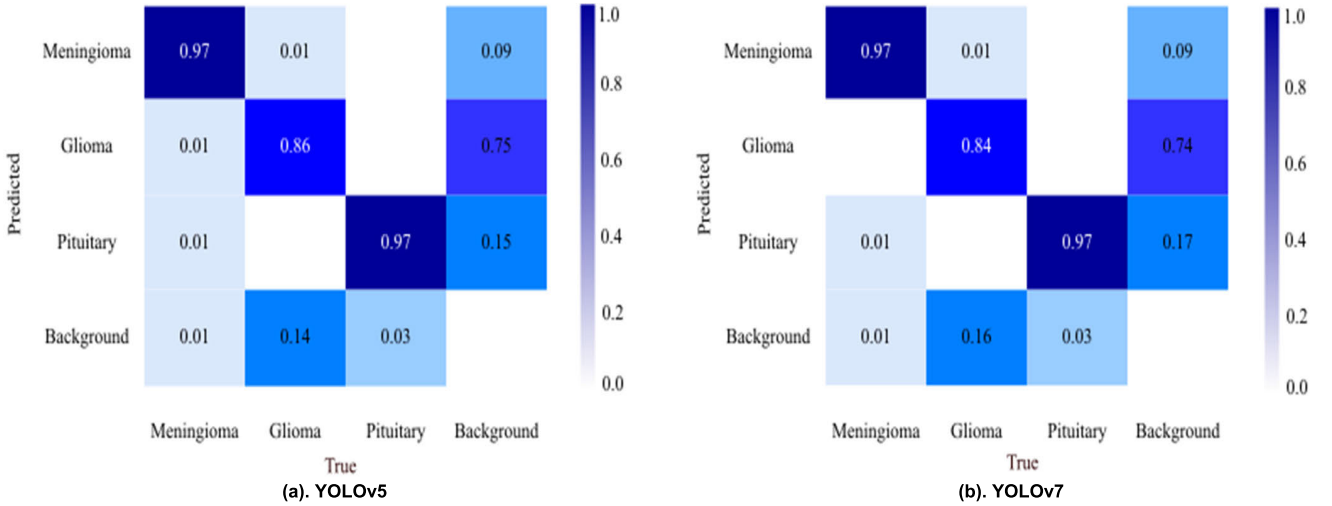


FIGURE 13. Confusion Matrix for YOLO models for Brain Tumor segmentation and classification.

TABLE 6. Mean average precision (MAP@.5:.95).

Class	Images	Labels	YOLOv5		YOLOv7	
			Box	Mask	Box	Mask
All	827	827	0.666	0.657	**0.677**	**0.659**
Meningioma		202	0.79	0.799	**0.803**	**0.793**
Glioma		382	0.546	0.53	**0.553**	**0.531**
Pituitary		243	0.661	0.642	**0.675**	**0.653**

while YOLOv7 has a score of 0.553 for box detection and 0.531 for mask detection. For the pituitary, YOLOv5 has a score of 0.661 for box detection and 0.642 for mask detection, while YOLOv7 has a score of 0.675 for box detection and 0.653 for mask detection.

Figure 12 demonstrate the precision-recall curves of YOLOv5 and YOLOv7 models. It depicts that both models perform better for all classes except glioma. The glioma curve shows a more false positive than meningioma and pituitary.

G. DETECTION TIME

The models are evaluated on a dataset of 91 frames using an image size of 512 pixels. For YOLOv5, the inference time is 7.0 ms, and the NMS time is 1.2 ms. The Frames Per Second (FPS) for YOLOv5 is 142.857, indicating that the model can process approximately 143 frames per second on average. For YOLOv7, the inference time is 10.0 ms, and the NMS time is 1.1 ms per image. The FPS for YOLOv7 is 100, indicating that the model can process approximately 100 frames per second on average.

H. CONFUSION MATRIX

Figure 13 shows the normalized confusion matrix of YOLOv5 and YOLOv7, illustrating that both models perform well for all classes except for glioma, which exhibits high false positives with values of 0.75 and 0.74, respectively.

Figure 14 illustrates the visual representation of the validation (a) and (c) batches, as well as prediction batches (b) and (d) for both YOLOv5 and YOLOv7 models. These batches serve to optimize computational efficiency and enhance inference speed, enabling the detection of instances across multiple images. The validation images showcase tumor detection on the original images; while the predicted images demonstrate the accuracy of the brain tumor detection after training on the original images.

I. DISCUSSION

The performance of the YOLOv5 and YOLOv7 is evaluated in segmenting three distinct classes of brain tumors, namely meningioma, glioma, and pituitary. The performance of these models relies on several metrics, including the confusion matrix, precision, recall, F1-curve, and inference criteria. The finding revealed that both models exhibited superior performance across all classes except for glioma, which displayed an elevated false positives rate of 0.75 and 0.74 for YOLOv5 and YOLOv7, respectively. Notably, YOLOv5 excelled in detecting meningioma, while it also outperformed in glioma and pituitary tumor detection. In contrast, YOLOv7 demonstrated comparable performance across all three classes, albeit with slightly lower precision scores than YOLOv5. Furthermore, precision scores for box detection are generally exceeded those for mask detection.

The study also presented the precision-confidence curves, which underscored models robust performance, except in the case of glioma. Remarkably, recall scores for all classes exceeded 0.78, with the highest recall observed for meningioma in both YOLOv5 and YOLOv7. Based on these results, it can be concluded that both models offer exceptional overall performance, with YOLOv7 demonstrates superior performance across numerous evaluation criteria, particularly with regard to mAP@0.5-0.95as evidence in Table 6.

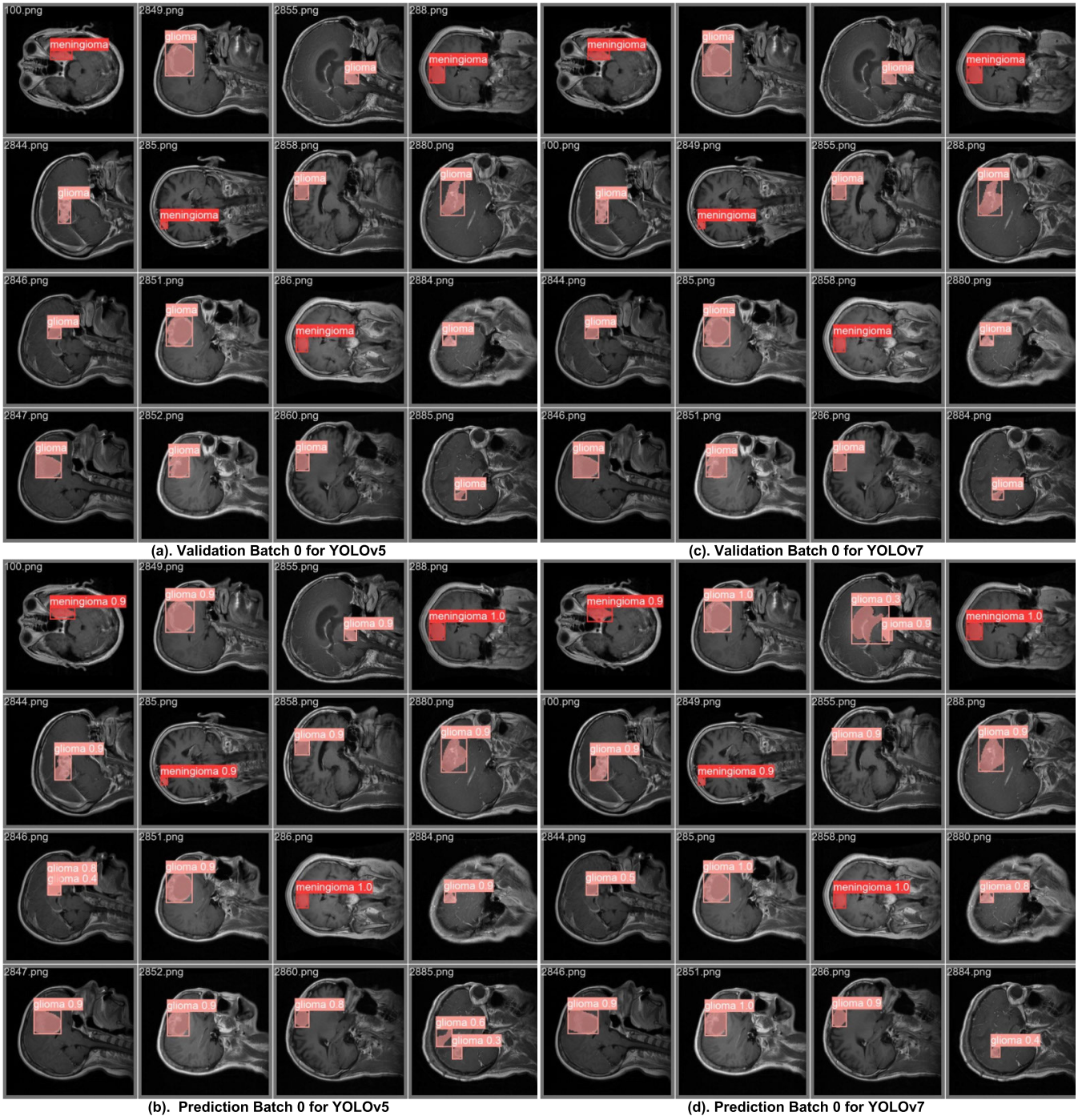


FIGURE 14. Validation and prediction batches for tumor segmentation and classification.

V. COMPARATIVE ANALYSIS OF SEGMENTATION MODELS

Table 7 showcase various segmentation models, including RCNN, Faster RCNN, and Mask RCNN [49], all of which employed the same dataset. The RCNN models achieved a commendable high recall rate with 95%, though precision and mAP remain unspecified. In contrast, YOLOv5 exhibited a precision of 93.6%, a recall rate of 90.6%, and an mAP of 94.7% at an IoU of 0.5, with a slightly reduced mAP of 65.7% in the broader IoU range of 0.5 to 0.95. Similarly, YOLOv7 achieved a precision of 93.5% precision, a recall rate of 90.3%, and an mAP of 94.7% at an IoU of 0.5, with a corresponding mAP of 65.7% within the IoU range of 0.5 to 0.95. This work presents a state-of-the-art implementation of YOLO based brain tumor detection and classification. However, it is essential to note that this study relies exclusively on

TABLE 7. Comparison of the proposed models.

Models	Precision	Recall	mAP@.5	mAP@.5:.95
RCNN [55]	×	95.0%	91.0%	×
Faster RCNN [55]	×	94.0%	94.0%	×
Mask RCNN [55]	×	95.0%	94.8%	×
YOLOv5 (Proposed)	93.6%	90.6%	94.7%	65.7%
YOLOv7 (Proposed)	93.5%	90.3%	94.1%	65.9%

a single dataset, namely the brain tumor dataset for evaluating the employed model. Additionally, a limitations of this work lies in the substantial computational resources required to analyze larger datasets with these models. This challenge may be mitigated by augmenting available computing resources and amalgamating multiple datasets, including BraTS, and to assess variations in accuracy, training time, and other pertinent performance metrics. Conversely, transitioning to more lightweight deep learning models [46] or investigating self-supervised learning [22], holds significant potential, especially in handling diverse data sources.

VI. CONCLUSION

This study presents a comprehensive evaluation of the YOLO-based deep learning segmentation and classifications of brain tumors, in particular meningioma, glioma, and pituitary. Both YOLOv5 and YOLOv7 demonstrated superior performance in segmenting and correctly identifying the specific tumor class. These models exhibited remarkable proficiency in accurately detecting meningioma, with YOLOv5 excelling in identifying glioma and pituitary tumors. Furthermore, YOLOv7 displayed comparable performance across all three tumor categories, with only slight difference in precision scores compare to YOLOv5. Recall scores for all categories consistently exceeded the threshold of 0.78, with the highest recall scores observed for meningioma in both YOLOv5 and YOLOv7. These findings validate the potential of YOLO models in precisely detecting brain tumors, particularly in the case of meningioma. They also offer valuable insights into the performance characteristics and limitations of these models, paving the way for further advancements in the computer vision and medical field.

ACKNOWLEDGMENT

The authors extend their appreciation to the Deputyship for Research Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number 223202.

REFERENCES

[1] (2023). *Key Statistics for Brain and Spinal Cord Tumors*. Accessed: May 18, 2023. [Online]. Available: https://www.cancer.org/cancer/types/brain-spinal-cord-tumors-adults/about/key-statistics.html

[2] S. Lapointe, A. Perry, and N. A. Butowski, "Primary brain tumours in adults," *Lancet*, vol. 392, no. 10145, pp. 432–446, Aug. 2018.

[3] E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem, and C. Davatzikos, "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme," *Magn. Reson. Med.*, vol. 62, no. 6, pp. 1609–1618, Dec. 2009.

[4] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, and Y. Pan, "A survey of MRI-based brain tumor segmentation methods," *Tsinghua Sci. Technol.*, vol. 19, no. 6, pp. 578–595, Dec. 2014.

[5] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in *Proc. World Congr. Med. Phys. Biomed. Eng.* Prague, Czech Republic: Springer, 2019, pp. 183–189.

[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 Intell. Syst.*, vol. 8, no. 4, pp. 3161–3183, Aug. 2022.

[7] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, and J. Lu, "Brain tumor classification for MR images using transfer learning and fine-tuning," *Computerized Med. Imag. Graph.*, vol. 75, pp. 34–46, Jul. 2019.

[8] J. Watts, G. Box, A. Galvin, P. Brotchie, N. Trost, and T. Sutherland, "Magnetic resonance imaging of meningiomas: A pictorial review," *Insights Imag.*, vol. 5, no. 1, pp. 113–122, Feb. 2014.

[9] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[10] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift Für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, May 2019.

[11] G. Alwakid, W. Gouda, M. Humayun, and N. Z. Jhanjhi, "Diagnosing melanomas in dermoscopy images using deep learning," *Diagnostics*, vol. 13, no. 10, p. 1815, May 2023.

[12] M. F. Almufareh, S. Tehsin, M. Humayun, and S. Kausar, "A transfer learning approach for clinical detection support of monkeypox skin lesions," *Diagnostics*, vol. 13, no. 8, p. 1503, Apr. 2023.

[13] M. Humayun, M. I. Khalil, S. N. Almuayqil, and N. Z. Jhanjhi, "Framework for detecting breast cancer risk presence using deep learning," *Electronics*, vol. 12, no. 2, p. 403, Jan. 2023.

[14] A. Khan, J. Bukhari, J. I. Bangash, A. Khan, M. Imran, M. Asim, M. Ishaq, and A. Khan, "Optimizing connection weights of functional link neural network using APSO algorithm for medical data classification," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 6, pp. 2551–2561, Jun. 2022.

[15] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019.

[16] M. Imran, H. Anwar, M. Tufail, A. Khan, M. Khan, and D. A. Ramli, "Image-based automatic energy meter reading using deep learning," *Comput., Mater. Continua*, vol. 74, no. 1, pp. 203–216, 2023.

[17] G. Alwakid, W. Gouda, and M. Humayun, "Deep learning-based prediction of diabetic retinopathy using CLAHE and ESRGAN for enhancement," in *Proc. MDPI*, 2023, p. 863.

[18] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Vision—ECCV*. Amsterdam, The Netherlands: Springer, 2016, pp. 21–37.

[19] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.

[20] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

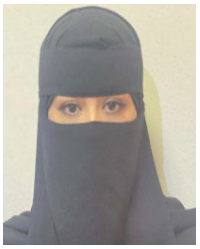
[21] W. Zhao, C. Li, W. Zhang, L. Yang, P. Zhuang, L. Li, K. Fan, and H. Yang, "Embedding global contrastive and local location in self-supervised learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 5, pp. 2275–2289, May 2023.

[22] W. Zhao, L. Yang, W. Zhang, Y. Tian, W. Jia, W. Li, M. Yang, X. Pan, and H. Yang, "Learning what and where to learn: A new perspective on self-supervised learning," *IEEE Trans. Circuits Syst. Video Technol.*, early access, Jul. 26, 2023, doi: 10.1109/TCSVT.2023.3298937.

[23] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.

[24] (2023). *Figshare Brain Tumor Dataset*. Accessed: May 23, 2023. [Online]. Available: https://www.kaggle.com/datasets/ashkhagan/figshare-brain-tumor-dataset

[25] A. Rehman, M. A. Khan, T. Saba, Z. Mehmood, U. Tariq, and N. Ayesha, "Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture," *Microsc. Res. Technique*, vol. 84, no. 1, pp. 133–149, Jan. 2021.
[26] 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 Comput. Informat. J.*, vol. 3, no. 1, pp. 68–71, Jun. 2018.
[27] C. Baur, S. Denner, B. Wiestler, N. Navab, and S. Albarqouni, "Autoencoders for unsupervised anomaly segmentation in brain MR images: A comparative study," *Med. Image Anal.*, vol. 69, Apr. 2021, Art. no. 101952.
[28] K. Munir, H. Elahi, A. Ayub, F. Frezza, and A. Rizzi, "Cancer diagnosis using deep learning: A bibliographic review," *Cancers*, vol. 11, no. 9, p. 1235, Aug. 2019.
[29] J. Amin, M. Sharif, N. Gul, M. Raza, M. A. Anjum, M. W. Nisar, and S. A. C. Bukhari, "Brain tumor detection by using stacked autoencoders in deep learning," *J. Med. Syst.*, vol. 44, no. 2, pp. 1–12, Feb. 2020.
[30] M. Togaçar, B. Ergen, and Z. Cömert, "Application of breast cancer diagnosis based on a combination of convolutional neural networks, ridge regression and linear discriminant analysis using invasive breast cancer images processed with autoencoders," *Med. Hypotheses*, vol. 135, Feb. 2020, Art. no. 109503.
[31] G. B. Praveen and A. Agrawal, "Hybrid approach for brain tumor detection and classification in magnetic resonance images," in *Proc. Commun., Control Intell. Syst. (CCIS)*, Nov. 2015, pp. 162–166.
[32] M. Sharma, G. N. Purohit, and S. Mukherjee, "Information retrieves from brain MRI images for tumor detection using hybrid technique K-means and artificial neural network (KMANN)," in *Networking Communication and Data Knowledge Engineering*, vol. 2. Berlin, Germany: Springer, 2018, pp. 145–157.
[33] M. Togaçar, B. Ergen, and Z. Cömert, "BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model," *Med. Hypotheses*, vol. 134, Jan. 2020, Art. no. 109531.
[34] J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "Big data analysis for brain tumor detection: Deep convolutional neural networks," *Future Gener. Comput. Syst.*, vol. 87, pp. 290–297, Oct. 2018.
[35] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.
[36] S. Sajid, S. Hussain, and A. Sarwar, "Brain tumor detection and segmentation in MR images using deep learning," *Arabian J. Sci. Eng.*, vol. 44, no. 11, pp. 9249–9261, Nov. 2019.
[37] M. I. Sharif, M. A. Khan, M. Alhussein, K. Aurangzeb, and M. Raza, "A decision support system for multimodal brain tumor classification using deep learning," *Complex Intell. Syst.*, vol. 8, no. 4, pp. 3007–3020, Aug. 2022.
[38] G. Xiao, H. Wang, J. Shen, Z. Chen, Z. Zhang, and X. Ge, "Synergy factorized bilinear network with a dual suppression strategy for brain tumor classification in MRI," *Micromachines*, vol. 13, no. 1, p. 15, Dec. 2021.
[39] H. S. Ali, A. I. Ismail, E. M. El-Rabaie, and F. E. A. El-Samie, "Deep residual architectures and ensemble learning for efficient brain tumour classification," *Exp. Syst.*, vol. 40, no. 6, Jul. 2023, Art. no. e13226.
[40] A. A. Waskita, J. M. Amda, D. S. K. Sihono, and H. Prasetio, "EfficientNetV2 based for MRI brain tumor image classification," in *Proc. Int. Conf. Comput., Control, Informat. Appl. (IC3INA)*, Oct. 2023, pp. 171–176.
[41] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A review of YOLO algorithm developments," *Proc. Comput. Sci.*, vol. 199, pp. 1066–1073, Jan. 2022.
[42] M. A. Al-masni, M. A. Al-antari, J.-M. Park, G. Gi, T.-Y. Kim, P. Rivera, E. Valarezo, M.-T. Choi, S.-M. Han, and T.-S. Kim, "Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system," *Comput. Methods Programs Biomed.*, vol. 157, pp. 85–94, Apr. 2018.
[43] H. M. Ünver and E. Ayan, "Skin lesion segmentation in dermoscopic images with combination of YOLO and GrabCut algorithm," *Diagnostics*, vol. 9, no. 3, p. 72, Jul. 2019.
[44] J. George, S. Skaria, and V. V. Varun, "Using YOLO based deep learning network for real time detection and localization of lung nodules from low dose CT scans," *Proc. SPIE*, vol. 10575, pp. 347–355, Feb. 2018.
[45] A. Hossain, M. T. Islam, and A. F. Almutairi, "A deep learning model to classify and detect brain abnormalities in portable microwave based imaging system," *Sci. Rep.*, vol. 12, p.6319, Apr. 2022, doi: 10.1038/s41598-022-10309-6.
[46] T. Shelatkar and U. Bansal, "Diagnosis of brain tumor using light weight deep learning model with fine tuning approach," in *Machine Learning and Computational Intelligence Techniques for Data Engineering*, vol. 2. Berlin, Germany: Springer, 2023, pp. 105–114.
[47] M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, M. A. J. Pramanik, S. Aryal, M. A. A. Almoyad, K. F. Hasan, and M. A. Moni, "An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning," *Exp. Syst. Appl.*, vol. 230, May 2023, Art. no. 120534.
[48] L. A. Salman, A. T. Hashim, and A. M. Hasan, "Automated brain tumor detection of MRI image based on hybrid image processing techniques," *Telkomnika*, vol. 20, no. 4, pp. 762–771, Aug. 2022.
[49] T. Liu, Z. Yuan, L. Wu, and B. Badami, "An optimal brain tumor detection by convolutional neural network and enhanced sparrow search algorithm," *Proc. Inst. Mech. Eng., H, J. Eng. Med.*, vol. 235, no. 4, pp. 459–469, Apr. 2021.
[50] Y. Bhanothu, A. Kamalakannan, and G. Rajamanickam, "Detection and classification of brain tumor in MRI images using deep convolutional network," in *Proc. 6th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Mar. 2020, pp. 248–252.
[51] M. M. Badža and M. C. Barjaktarovic, "Classification of brain tumors from MRI images using a convolutional neural network," *Appl. Sci.*, vol. 10, no. 6, p. 1999, Mar. 2020.
[52] O. N. Belaid and M. Loudini, "Classification of brain tumor by combination of pre-trained VGG16 CNN," *J. Inf. Technol. Manag.*, vol. 12, no. 2, pp. 13–25, 2020.
[53] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0140381, doi: 10.1371/journal.pone.0140381.
[54] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.
[55] M. Masood, T. Nazir, M. Nawaz, A. Javed, M. Iqbal, and A. Mehmood, "Brain tumor localization and segmentation using mask RCNN," *Frontiers Comput. Sci.*, vol. 15, no. 6, Dec. 2021, Art. no. 156338.



MARAM FAHAAD ALMUFAREH received the Ph.D. degree in computer sciences from Claremont Graduate University, USA. She has five-plus years of teaching and administrative experience nationally and internationally. She has an extensive background in teaching, research supervision, and administrative work. She has experience in teaching advanced Era technological courses, including mobile application development (Android), cyber security, net framework programming besides other undergraduate and postgraduate courses, graduation projects, and thesis supervisions.

MUHAMMAD IMRAN received the Ph.D. degree (Hons.) in computer science (specifically in the area of cloud computing and data preservation) from the University of Vienna, Austria, in 2014. He is currently an Assistant Professor and a Member Postgraduate Advisory Committee with the Institute of Computer Sciences and Information Technology, The University of Agriculture, Peshawar, Pakistan. He has published several research articles in international journals, presented his work at many conferences of international repute, and chaired many conferences. His research interests include cloud computing, data preservation, provenance, machine learning, and service-oriented architectures.

ABDULLAH KHAN received the B.Sc. degree from the University of Malakand, the M.Sc. degree in computer science from the University of Science and Technology, Bannu, and the Ph.D. degree from Universiti Tun Hussein Onn Malaysia, in 2010. He is currently an Assistant Professor with the Faculty of Management and Computer Sciences, The University of Agriculture, Peshawar, Pakistan. He has published a variety of research papers in reputed journals and conferences. His research interests include the domain of artificial intelligence, machine learning, and deep learning.

MAMOONA HUMAYUN received the Ph.D. degree in computer sciences from the Harbin Institute of Technology, China. She has 15 years of teaching and administrative experience internationally. She has an extensive teaching, research supervision, and administrative work background. She has authored several research papers, supervised many postgraduate students, and has an external thesis examiner to her credit. She has strong analytical, problem-solving, interpersonal, and communication skills. Her research interests include cyber security, wireless sensor networks (WSN), the Internet of Things (IoT), requirement engineering, global software development, and knowledge management. She is currently the guest editor and a reviewer of several reputable journals and conferences around the globe.

MUHAMMAD ASIM received the B.S. degree in electrical engineering from CECOS University, Pakistan, in 2009, the M.Eng. degree in information and communication engineering from Chosun University, South Korea, in 2012, under the NIPA Scholarship Program, and the Ph.D. degree in information technologies from the University of Parma, Italy, in 2016. Currently, he is an Assistant Professor with the Institute of Computer Sciences and Information Technology, ICS/IT, Faculty of Management and Computer Sciences, The University of Agriculture, Peshawar, Pakistan. His research interests include channel coding schemes, soft decision-directed iterative receivers in fading channels, reconfigurable architectures, and computer vision techniques.