Impact of Data-Augmentation on Brain Tumor Detection Using Different YOLO Versions Models

Abdelraouf Ishtaiwi
Faculty of Information Technology, University of Petra, Jordan
aishtaiwi@uop.edu.jo

Ali Ali
Faculty of Engineering, Al-Ahliyya Amman University, Jordan
a.mahmoud@ammanu.edu.jo

Ahmad Al-Qerem
Faculty of Information Technology, Zarqa University, Jordan
ahmad_qerm@zu.edu.jo

Yazan Alsmadi
Faculty of Information Technology, Zarqa University, Jordan
Y. Alsmadi@zu.edu.jo

Amjad Aldweesh
College of Computing and Information Technology, Shaqra University, Saudi Arabia
a.aldweesh@su.edu.sa

Mohammad Alauthman
Faculty of Information Technology, University of Petra, Jordan
mohammad.alauthman@uop.edu.jo

Omar Alzubi
Computer Engineering Department, Umm Al-Qura University, Saudi Arabia
orzubi@uqu.edu.sa

Shadi Nashwan
Faculty of Information Technology, Middle East University, Jordan
snashwan@meu.edu.jo

Awad Ramadan
College of Computing in Al-Qunfudah, Umm Al-Qura University, Saudi Arabia
amabker@uqu.edu.sa

Musab Al-Zghoul
Faculty of Information Technology
Isra University, Jordan
musab.alzgool@iu.edu.jo

Someah Alangari
College of Computing and Information Technology, Shaqra University, Saudi Arabia
salangari@su.edu.sa

Abstract: *Brain tumors are widely recognized as one of the world's worst and most disabling diseases. Every year, thousands of people die as a result of brain tumors caused by the rapid growth of tumor cells. As a result, saving the lives of tens of thousands of people worldwide needs speedy investigation and automatic identification of brain tumors. In this paper, we propose a new methodology for detecting brain tumors. The designed framework assesses the application of cutting-edge YOLO models such as YOLOv3, YOLO v5n, YOLO v5s, YOLO v5m, YOLOv5l, YOLOv5x, and YOLOv7 with varying weights and data augmentation on a dataset of 7382 samples from three distinct MRI orientations, namely, axial, coronal, and sagittal. Several data augmentation techniques were also employed to minimize detector sensitivity while increasing detection accuracy. In addition, the Adam and Stochastic Gradient Descent (SGD) optimizers were compared. We aim to find the ideal network weight and MRI orientation for detecting brain cancers. The results show that with an IoU of 0.5, axial orientation had the highest detection accuracy with an average mAP of 97.33%. Furthermore, SGD surpasses Adam optimizer by more than 20% mAP. Also, it was found that YOLO 5n, YOLOv5s, YOLOv5x, and YOLOv3 surpass others by more than 95% mAP. Besides that, it was observed that the YOLOv5 and YOLOv3 models are more sensitive to data augmentation than the YOLOv7 model.*

Keywords: *Data-augmentation, objects detection, brain tumor, yolov7, computer vision.*

Received February 22, 2024; accepted May 5, 2024
https://doi.org/10.34028/iajit/21/3/10

1. Introduction

The human brain is the principal organ of the human nervous system and the command-and-control center for all activities necessary to maintain a healthy, normal life. The brain receives impulses or stimuli from the many sensory organs, analyzes them, and responds accordingly. As a result of unchecked cell division or mutations, abnormal clusters of brain cells are generated, eventually leading to the development of a brain tumor. Not only can these cells be harmful to healthy tissue, but they may also disrupt normal brain function [17, 22].

Vomiting, Headaches, cognitive difficulties, personality changes, vision and speech impairments, and nausea and vomiting are all common symptoms of brain tumors. When a brain tumor grows, it can affect a person's personality, their way of thinking, and their ability to do just about anything else.

Brain tumors are classified into two types: those that are not cancerous and are known as benign and those that are cancerous (called malignant). Brain tumors considered benign grow slowly and are not malignant. This type of tumor is less dangerous since it cannot spread to other body parts.

Malignant tumors, on the other hand, are cancerous growths that are distributed rapidly. Malignant tumors are more prevalent. Furthermore, carcinogenic tumors are classified into two types: primary malignant tumors that start in the brain and spread to other body regions, and secondary malignant tumors that develop in other

parts of the body and spread to the brain [4]. Primary malignant tumors are more likely to kill the patient. Gliomas, pituitary tumors, and meningiomas are the three most common forms of brain tumors diagnosed. A meningioma is a tumor that develops in the meninges, the thin membranes (or tissues) surrounding and protecting the brain and spinal cord. Gliomas begin in the brain's glial cells, where they also begin. Tumors of the pituitary gland can develop when cells in the pituitary gland, located close to the brain, grow uncontrollably. A brain tumor is one of the diseases that can take someone's life the most quickly. Despite the presence of brain tumors, prompt identification and treatment are necessary to save lives. Machine learning (ML) algorithms could automatically diagnose individuals with brain tumors and classify them into specific groupings to address this issue. However, because of the wide range of sizes, shapes, and intensities these tumors can take, classifying brain cancers into meningioma, pituitary, and glioma tumors is more challenging [11]. This is because these tumors can take any of these shapes. Furthermore, meningiomas, pituitary tumors, and gliomas account for the great majority of occurrences of brain cancer [37].

Furthermore, the great resolution of brain MRI allows for an in-depth analysis of the brain's anatomy. As a result, magnetic resonance imaging (MRI) images substantially impact the automatic interpretation of medical images [18, 21, 36, 49]. Researchers heavily rely on MRI technology when detecting and evaluating brain cancers. Researchers have recently created many new automated algorithms for detecting and classifying brain cancers in MRI data. Traditional machine learning algorithms, such as Multi-Layer Perceptron (MLP) and SVM classifiers, are extensively used for brain tumor identification [30]. Deep learning (DL) [1] is a subfield of machine learning that builds a feature hierarchy by using low-level features to build high-level features.

Technological progress has enabled digital image processing to spread to fields including photogrammetry, remote sensing, and computer vision [8]. When an image is processed digitally, it undergoes a series of transformations that allow us to convert it to a digital format and extract valuable data. Computer vision applications of deep learning for digital image processing have expanded to include a wide range of tasks, from face recognition [2] to object detection and classification [3]. Remote sensing and photogrammetric images are ideal for using deep learning-based object detection techniques. Deep learning approaches to object detection can perform better with larger datasets and more robust models. Significant advances in object detection have been made thanks to R-CNNs and other region-based approaches [10]. Two main types of two-stage convolutional neural networks are used for object identification, and they are the single-stage networks. There are a few different types of two-stage CNNs, including R-CNN [14], Faster R-CNN [34], and R-FCN [12]. Two-stage methods are less efficient than one-stage methods, although they yield results eventually. One example of a single-stage strategy employed in the research is the "You Only Look Once" (YOLO) methodology. To locate objects with distinct bounding boxes in space, YOLO frames the task as a regression problem. How YOLO approaches the problem allows it to yield results faster than competing two-stage item identification algorithms [38].

DWI is an essential technique for diagnosing these malignant growths because it may show how brain tumors interfere with the normal free diffusion of water inside tissues [45]. DWI can reveal how brain tumors obstruct the usual free diffusion of water within tissues.

Because of the numerous features that can be extracted from MRI images, these images are a true goldmine of information that can be utilized to classify tumors. Learning how to describe data opens the stage for DL to eventually apply that knowledge to form inferences and carry out actions.

DL methods are used to classify diagnostic imaging studies. However, DL-based approaches are useful in various domains and specialties [5]. A large amount of training data is required for DL algorithms to perform successfully. In recent years, there has been an increase in the acceptability of DL techniques in general and the prominence of the CNN model in particular.

Some potentially fatal diseases have been difficult to diagnose, but recent advances in Computer-Aided Diagnosis (CAD) have eliminated that problem for many people CAD. This technology makes rapid and accurate identification by medical equipment possible, allowing doctors to extend a patient's life or improve their quality of life [9].

Further, potent new tools called Deep Convolutional Neural Networks (DCNN) were produced by fusing ML and CV. State-of-the-art models like Deep Convolutional Neural Networks (DCNNs) have successfully challenged CAD issues like recognition, classification, segmentation, and detection [29]. Figure 1 shows the process of detecting brain tumors using deep learning methods.

However, DCNNs form the basis of most existing CAD systems for identifying and detecting brain cancers. Unfortunately, these systems are not very powerful and perform poorly across most platforms [16]. Lighter classification models are not as effective for most DCNNs because they cannot pinpoint the tumor's location. The processing costs of a segmentation model, which uses a mask to detect the damaged area and identify the tumor, are greater. However, most existing CAD systems for identifying and locating brain and breast cancers rely on deep convolutional neural networks. However, these systems typically perform poorly across platforms and require a lot of computing power. Most DCNNs are restricted by the inability of lightweight classification models to pinpoint the precise site of the tumor [23]. Comparatively, a segmentation

model can utilize a mask to detect the damaged area and identify the tumor, but this model has greater processing costs [28].

The key contributions are as follows:

1. Several modern and ancient detection models exist.

In this paper, we propose a new methodology for comparing and evaluating the effectiveness of the most recent models of early brain cancer detection, such as YOLOv3, YOLOv5, and YOLOv7. The objective is to highlight the most effective models that can be used to identify various illnesses in the future.

2. Data quality is critical when utilizing supervised learning because of the scarcity of labeled datasets.

We apply a mixture of data augmentation techniques to expand the number of dataset components, lower the sensitivity of detection models, and improve their performance in the early detection of brain cancer utilizing other datasets.

3. Brain cancer may be diagnosed utilizing images with three distinct axes such as sagittal, coronal, and axial. As a result, in this study, we use three datasets that describe all distinct axes. Also, to emphasize the image dimensions with the best discriminatory strength of contemporary detection algorithms.

4. To our knowledge, the YOLOv7 model is the most recent detection model. As a result, we explain the utilization of the most well-known optimizers, such as Adam and Stochastic Gradient Descent (SGD), in terms of performance and detection speed rate based on a set of loss functions and evaluation metrics.

The rest of the paper is structured as follows: section 2 shows the state-of-the-art related works. Section 3 illustrates the proposed methodology for brain tumor detection, dataset description and analysis, and employed detection models. Section 4 shows the evaluation metrics. Section 5 discusses the experimental results. Finally, the conclusion and future directions are shown in section 6.

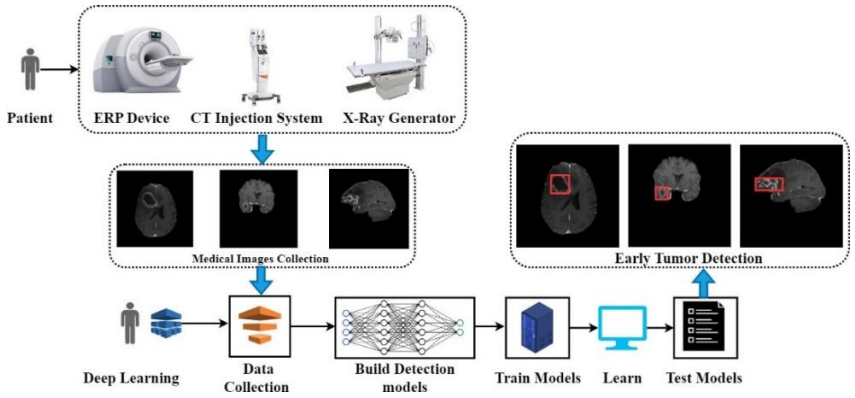


Figure 1. Brain tumor detection process using deep learning approaches.

2. Related Works

Detecting different objects in indoor, outdoor, and medical images is difficult since objects can be visible in low light. As a result, the pixel colors of the photographs are more closely associated with dark hues, notably black. For many years, prior contributions have used various ways to improve the detection process of medical diseases. However, a new version of the You Only Look Once (YOLO) algorithm that builds deep learning algorithms for object recognition and detection was presented in 2020. YOLOv5 shows better detection accuracy and performance compared to previous versions. In [41], a novel framework for detecting interior occupancy objects has been suggested. Using the anchor-free approach for parameter decrease and VariFocal loss for data balancing, the designed framework optimizes the utilization of YOLOv5.

Furthermore, a newly constructed dataset with 11,367 samples partitioned into training, testing, and validation sets was presented. In addition, a well-known dataset, Pascal-VOC2012, was used in the experiment. Also, the YOLOv5 improvement includes decoupling the head's layer to improve detection accuracy and performance. However, utilizing a resolution of 640 by 640 pixels. The proposed framework's findings were compared to eleven earlier models that used YOLO in various versions. The results of the tests reveal that the model can obtain an average accuracy of 93.9 at an Intersection Over Union (IOU) of 0.9. In the study in [28], a real-time experiment was conducted to identify indoor and outdoor objects by building an engineering system that utilizes camera sensors such as OS1-64 and OS0-128 that apply to the Lidar device. However, the primary contribution is ended by employing full 360-degree images with a resolution of 2048x128. The developed system compares the performance of FasterR-CNN, MaskRCNN, YOLOx, and YOLOv5. Sensor images define four interior and outdoor target classes: a person, a bike, a chair, and a car. The findings demonstrate that YOLOx outperforms others, successfully detecting over 80% of the interior and

outdoor objects with a precision of 100% and a recall of 95.3%. They also stated that YOLOx outperforms YOLOv5 regarding detection performance and speed.

Other research evaluated the evolution and improvement of various YOLO algorithm variants throughout time. In [19] a new framework was proposed to compare the efficiency of YOLOv3 with YOLOv5. The framework is intended to detect Apple items. However, a collection of 878 images with varying image resolutions describing a near and distant view at apples was employed. The testing results show that YOLOv5 surpasses YOLOv3 in recall, with 97.8%, compared to other models such as Faster-RCNN (81.4% recall) and DaSNet-v2 (86.8%). YOLOv5 generates better outcomes in terms of precision, recall, and F-measure. In [27] a methodology for detecting landing locations for autonomous flying systems was proposed. The developed frame analyzes the usage of several YOLO versions, such as YOLOv3, YOLOv4, and YOLOv5, to identify landing sweet spots to decrease flight system failure and boost safety. However, a collection of 11268 satellite images, called the DOTA, with up to 20,000x20,000 image resolution and 15 labels, was employed. The findings demonstrate that YOLOv5 with large network weights outperforms others, with a precision of 70%, recall of 61%, and mean average precision of 63%. Furthermore, YOLOv4 exceeds YOLOv3 with a recall of 57% and a mean average accuracy of 60%.

The detection of indoor and outdoor objects is helping to track down pandemic infections such as coronavirus. However, numerous research employed YOLO to recognize face masks. In [40] a real-time monitoring system that identifies face masks for the COVID-19 pandemic was proposed using YOLOv5. Closed-circuit television footage is supplied into the established framework. However, a dataset of 3,846 images with masked and unmasked labels was created. Several approaches were used to supplement the data, including Gaussian and motion blur. In addition, stacked ResNet-50, which incorporates transfer learning, was deployed. With a testing accuracy of 87% and a precision of 71%, stacked ResNet-50 surpasses other comparable models such as ResNet-50 and Convolutional Neural Network. In [47] a new face detection method based on YOLOv5 has been developed. ShuffleCANet is used as a new backbone YOLO layer in the designed system.

Furthermore, the AIZOO dataset was employed, which contains 7959 images of a face and a masked face. For image splicing and arrangement, the mosaic approach was used. Additionally, images were processed and scaled to 640 × 640. With a mean average accuracy of 95.2%, the proposed framework with YOLOv5 surpasses existing models such as YOLOv3. Furthermore, the proposed methodology with modified ShuffleCANet exceeds the original YOLOv5 findings by 0.58% in precision.

Other studies used YOLOv5 for indoor and outdoor fire and smoke detection. In [50], swin-YOLOv5, a new framework that improves on the original YOLOv5 architecture by improving feature extraction, was proposed. The developed framework can acceptably detect fire and smoke. Swin-fundamental YOLOv5's concept is to use a transformer across three headers. However, a dataset of 16,503 images of two target classes was employed for comparison. In addition, seven hyperparameters were fine-tuned. According to the data, swin-YOLOv5 outperforms the original by 0.7% mean average precision enhancement at 0.5 and 4.5% mean average precision enhancement at 0.5 to 0.95. In [43] an enhanced version of YOLOv5 was developed, incorporating dynamic anchor learning through the K-means++ algorithm. The approach developed seeks to limit fire damage by improving detection speed and performance. Furthermore, several loss functions such as CIOU and GIOU were applied to three distinct YOLOv5 models: YOLOv5 small, YOLOv5 medium, and YOLOv5 large. However, a self-created dataset of 4815 images was exposed to a synthetic system to increase the data size to 20,000. According to the findings, the modified model outperforms the original YOLOv5 by 4.4% mean average precision. It was also discovered that YOLOv5 performs better using the CIoU loss function, with a recall of 78% and a mean average precision of 87%.

Since the launch of YOLOv5 with many models, several studies have employed the YOLO algorithm to improve Internet of Things (IoT) technology. In [6], YOLOv5 was used to create a new framework to improve IoT devices with limited memory and computing power. The YOLOv4 model was also utilized to compare experiment findings. However, two separate datasets, including automobile license plates, were integrated into a single dataset to boost data size. The data set includes 5991 distinct 640x640-pixel images. Transfer learning was used to build the model using YOLOv5 with a tiny weight network and the Microsoft COCO dataset. Furthermore, Long Short-Term Memory (LSTM) based on the OCR engine was applied for automobile plate identification. The findings demonstrate that YOLOv5s outperform YOLOv4 with an mAP of 87% across 100 epochs.

In the context of medical data, YOLOv5 has shown an improvement in diagnosing cancer status. The study in [25] proposed a new methodology to improve the use of YOLOv5 for breast cancer detection. The intended work is assessed using four YOLOv5 weight models: YOLOv5 small, YOLOv5 medium, YOLOv5 large, and YOLOv5 x-large. However, the CBIS-DDSM dataset, which contains 10239 distinct 1000 x2000 pixel images, was employed. It specifies if the breast cancer is benign or malignant. The testing findings demonstrate that modified YOLOv5x outperforms the small, medium, and large weights, with a Matthews Correlation Coefficient (MCC) value of 93.6%. In addition, the

proposed improved version of YOLOv5m was compared to other models such as YOLOv3 and faster RCNN. It was discovered that modified YOLOv5m outperforms YOLOv3 and faster RCNN with an accuracy of 96.5% and mAP of 96%. In [26] a framework for detecting brain tumors via transfer learning is being developed. The proposed methodology employs the tiny YOLOv4 model for training and the YOLOv3 detection unit. However, the model was trained on the Microsoft Common Objects in Context (COCO) dataset with another gathered dataset that depicts 3064 magnetic resonance images with different regions for cancer tumors of 512x512 resolution, such as coronal, axial, and sagittal, using transfer learning. With a mean average precision of 0.9314, the findings show that the fine-tuned small YOLOv4 model with transfer learning surpasses others. The study in [35] used five distinct YOLOv5 models with transfer learning for identifying malignant brain tumors, including nano, small, medium, large, and x-large models. The proposed framework uses the Brats 2021 dataset, which contains 2,000 instances with 8000 scans at 240x240 resolution and three distinct kinds and locations, namely, T1, T2, and Flair. In addition, the Microsoft COCO dataset was utilized to train the model using transfer learning. With a mean average precision of 0.912, the findings demonstrate that the YOLOv5 x-large model outperforms the others. In this paper, we propose a new framework for detecting brain tumors using the three distinct YOLO models such as YOLOv3, YOLOv5, and YOLOv7 with different data augmentation techniques. Table 1 summarizes the state-of-the-art related works.

Table 1. Summary of the state-of-art research.

Paper	Dataset	Task	YOLO version	Findings
[41]	Newly constructed dataset with 11,367 samples and Pascal-VOC2012 with 640 x 640.	Indoor object detection.	YOLOv5	The best average accuracy of 93.9 at an intersection over union of 0.9.
[46]	Full 360-degree images with a resolution of 2048 x 12 were 8 collected by Lidar sensors.	Indoor and outdoor objects detection.	YOLOx and YOLOv5	YOLOx outperforms others with a precision of 100% and a recall of 95.3%.
[19]	Apple dataset with 878 images of different resolutions.	Detecting apple fruit.	YOLOv3 and YOLOv5	YOLOv5 outperforms YOLOv3 with a recall of 97.8%.
[27]	DOTA dataset with 11268 satellite images of 20,000 x 20,000 resolution and 15 target classes.	Detecting landing sweet spots.	YOLOv3, YOLOv4, and YOLOv5	YOLOv5 shows better results improvement with a precision of 70% and recall of 61%.
[40]	A dataset of 3,846 images of face masks was collected by CCTVs.	Detecting face masks.	YOLOv5	Stacked ResNet-50 outperforms others with a testing accuracy of 87% and a precision of 71%.
[47]	AIZOO dataset for face mask detection with 7959 images of 640 x 640.	Detecting face masks.	YOLOv3, and YOLOv5	ShuffleCANet as the backbone layer outperforms others with a mean average accuracy of 95.2%.
[50]	Dataset of 16,503 images of two target classes.	Fire and smoke detection.	YOLOv5	Swin-YOLOv5 outperforms others with an mAP of 0.7 improvements at an IOU of 0.5.
[43]	Self-build dataset of 4815 images of fires and smoke.	Fire detection.	YOLOv5	The improved model of YOLOv5 using K-means++ outperforms others by 4.4% mAP.
[6]	Google images, Microsoft COCO, and Indian number plates dataset	Automobile plate detection.	YOLOv4 and YOLOv5	YOLOv5s outperforms others with an mAP of 87%.
[25]	The CBIS-DDSM dataset with 10239 distinct 1000 x 2000 pixel images of breast cancer.	Breast cancer detection.	YOLOv5 and YOLOv3	YOLOv5x outperforms other models with an MCC of 93.6%. Also, it outperforms YOLOv3 with an accuracy of 96.5% and an mAP of 96%.
[26]	Microsoft COCO dataset and a collected dataset of 3064 brain cancer MRI images of 512 x 512.	Brain tumor detection.	Tiny YOLOv4 and YOLOv3	Tiny YOLOv4 with transfer learning shows the best results with an mAP of 93.14%.
[35]	Microsoft COCO and the Brats 2021 datasets with 240 x 240 brain cancer MRI images.	Brain tumor detection.	YOLOv5	YOLOv5 x-large model shows the best results with an mAP of 91.2%.

3. Proposed Methodology for Brain Tumor Detection

This section demonstrates the proposed framework for detecting brain tumors using various YOLO models such as YOLOv3, YOLOv7, and YOLOv5 with different weights and data augmentation. The detection process of various brain tumors of variable sizes and dimensions may be evaluated using various metrics regarding accuracy and loss functions. However, the weight and size of the neural networks have a significant influence on detection accuracy and speed, particularly in the situation of low-light magnetic resonance images. In this paper, we propose a novel framework for evaluating the usage of several YOLO models in order to find the optimal model for brain tumor detection, as shown in Figure 2. The proposed methodology contrasts traditional YOLO models such as YOLOv3 with cutting-edge models such as YOLOv7 and YOLOv5. We also test the YOLOv5 model with various network sizes, including nano, small, medium, large, and x-large networks. Data augmentation, on the other hand, improves the model's effectiveness by increasing the number of training samples. However, we apply several data augmentation techniques for improvement, such as image and bounding box flipping horizontally and vertically, which minimizes model sensitivity to varied orientations.

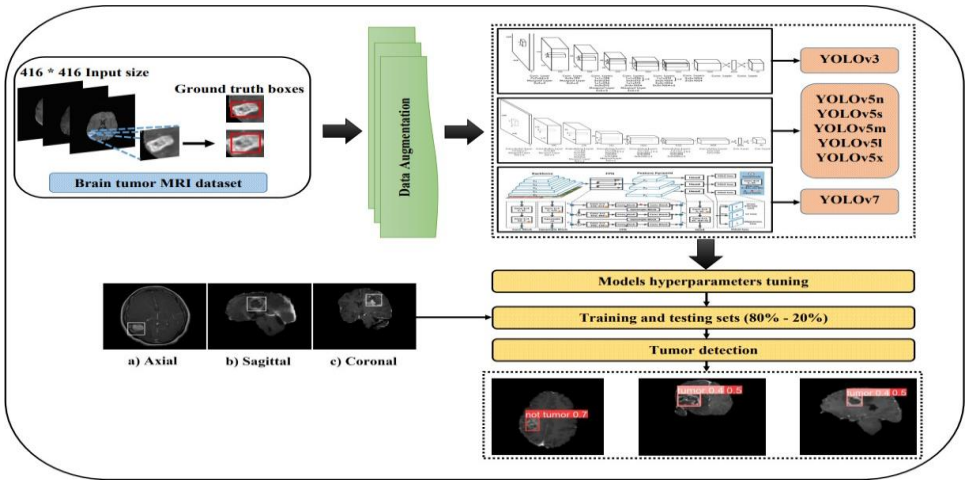


Figure 2. An Overview of the proposed framework for brain tumors detection.

Furthermore, it employs three distinct MRI imaging positions for brain tumors: axial, coronal, and sagittal. Our objective is to find the ideal posture for detecting cancer tumors. We maintain 20% of each dataset for model testing while evaluating multiple models. In addition, we compute several evaluation metrics for model comparisons, such as precision, recall, mean average precision, intersection over union, and three loss functions. Also, we compare different YOLO models with prior models, such as the faster region-based convolutional neural networks for object detection.

3.1. Dataset Collection and Image Processing

Magnetic Resonance Imaging (MRI) is a powerful diagnostic tool widely employed in medical imaging, capable of capturing detailed images of the body's internal structures from various orientations. The three primary orientations utilized in MRI are axial, coronal, and sagittal, each providing a unique perspective that enables healthcare professionals to comprehensively evaluate and diagnose a range of conditions, including brain tumors.

The axial orientation offers cross-sectional views perpendicular to the body's long axis, spanning from the top of the head to the bottom. These axial MRI images are invaluable for visualizing intricate details of the brain, spinal cord, and abdominal organs. Conversely, the coronal orientation presents a frontal view of the body, with the imaging plane perpendicular to the axial plane. Coronal MRI scans facilitate thorough assessments of the brain, eyes, facial structures, spinal cord, and abdominal organs.

Furthermore, the sagittal orientation provides a side view of the body, with the imaging plane parallel to its long axis. Sagittal MRI images are instrumental in examining the brain, spinal cord, pelvic region, and evaluating the integrity of various muscles and tendons. By incorporating these diverse MRI orientations into their study, researchers can comprehensively evaluate the performance of the YOLO models in detecting brain tumors from multiple vantage points, potentially enhancing the overall accuracy and robustness of the detection system.

In this work, we use an amassed dataset initially developed to identify malignancies in the brain using a variety of MRI orientations. The dataset is available online via the Kaggle repository at https://www.kaggle.com/datasets/davidbroberts/brain-tumor-object-detection-datasets. However, as shown in Figure 3, the dataset consists of three unique datasets representing three possible brain tumor orientations, namely axial, coronal, and sagittal, with two labels, tumor and non-tumor. It also includes 1218 images of varying resolutions. All Exchangeable Image File Format (EXIF) rotations were ignored during data preparation, and pixels were normalized. All images were also resized to 416×416. However, data analysis reveals that the axial dataset has 18 missing labels, and the coronal dataset contains one missing label. RoboFlow, an online platform, was used to manage missing classes and ground truth bounding boxes.

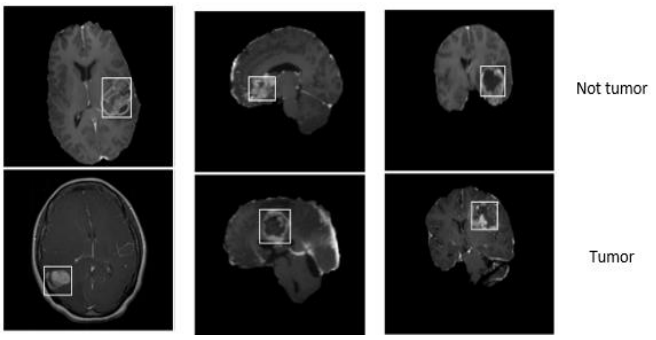


Figure 3. Illustrates a sample of different brain cancer MRI orientations and datasets labels.

On the other hand, data augmentation approaches were used to reduce the models' sensitivity to different orientations. We flip images and employ bounding boxes in horizontal and vertical directions to increase the number of data samples. The description and labels for the dataset are shown in Table 2. However, we maintain 20% of each dataset for testing for alternative

model comparisons by applying data augmentation techniques. The images were boosted three times more with 7382 training and testing samples. Table 3 shows the distribution of datasets following data augmentation.

Table 2. Datasets distribution following data augmentation.

Datasets labels	Axial	Sagittal	Coronal	Total
Tumor	1754	1023	1275	4052
Not tumor	1210	983	1137	3330
Total images	2964	2006	2412	7382
Total annotations	3141	2154	2566	7861
Average image size	0.17 megapixel	0.17 megapixel	0.17 megapixel	----
Median image ratio	416 x 416	416 x 416	416 x 416	----

Table 3. Dataset description and analysis following data-augmentation.

Evaluation sets	Axial	Sagittal	Coronal	Total
Training set (80%)	2721	1844	2215	6780
Testing set (20%)	243	162	197	602
Total	2964	2006	2412	7382

3.2. YOLOv5 Model

YOLOv5 [39, 48] was originally launched in May 2020. An object detection algorithm finds objects by looking at images all at once. The backbone, neck, and head layers are represented in Figure 4. All the layers are completely conventional networks. To begin, the backbone layer is utilized to extract significant and discriminative information from incoming images. In this YOLO version, the cross-stage partial network CSPNet is employed as the basic learner in the backbone layer to extract features. Second, the neck layer is utilized for feature pyramid creation, which aids in detecting the same objects with varying sizes and placements. YOLOv5, on the other hand, generates a features pyramid using the path aggregation network (PANet). Finally, the head layer, the YOLO layer, is employed for object recognition and prediction. It creates a vector containing the target class probability and bounding boxes. However, bounding boxes define object coordination regarding x and y cross points, height, and width. This layer generates many bounding boxes to improve detection accuracy and performance by computing the area of overlapping boxes. Intersection over Union (IoU) may, on the other hand, be computed to identify the best bounding overlapping boxes [31]. In this paper, we choose YOLOv5 because of its speed, performance, and accuracy. However, the critical differences between YOLOv5 and the prior versions are as follows:

1. It employs CSPDarknet53 at the backbone layer.
2. It employs the PANet in the neck layer.
3. It employs logistic and binary cross-loss functions.
4. It recognizes near and remote objects in the same input image.

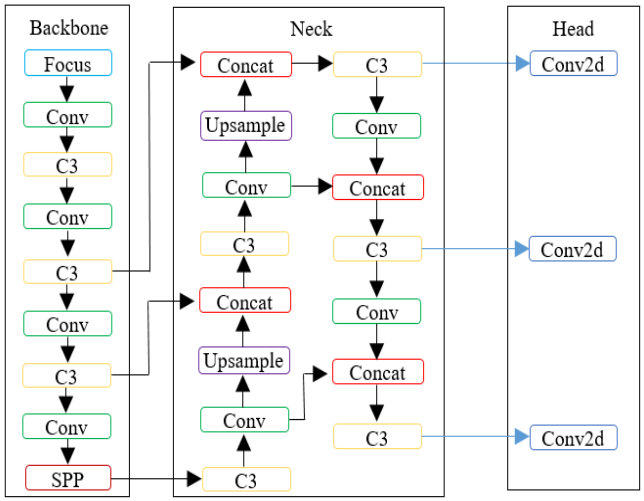


Figure 4. Illustration of YOLOv5 architecture.

3.3. YOLOv3 Model

In 2018, Redmon and Farhadi [32] came up with the idea for a new YOLO version, which they called YOLOv3 The revised version has an inference time of 22 milliseconds (ms) and a mean average precision of 28.2%. It does this by applying dimension clusters to the problem of predicting ground truth bounding boxes for anchor boxes. However, due to softmax's (the network classification layer's) poor performance, YOLOv3 uses the logistic regression function to minimize the confidence score. The confidence score is the probability of a certain item in a particular grid cell. Compared to YOLOv5, it employs the darknet-53 as the backbone layer, adding a more convolutional layer. In the neck layer, YOLOv5 extracts features using the path aggregation network. In light of the results, Redmon stated that the YOLOv3 detection model is faster and more accurate than other detection models, such as the YOLOv2 and Single-Shot Detector models (SSD). The architecture of the YOLOv3 network is shown in Figure 5. In modern times, YOLOv3 continues to be an effective detection model in various contexts. YOLOv3 was used by Magnuska *et al*. [24] to detect breast cancer tumors. According to the findings, the performance of YOLOv3 is superior to that of Viola-Jones in terms of the intersection over union measure. In addition, various variants, such as tiny-YOLOv3, were derived from YOLOv3 and used as the basis for their creation. In their experimental investigation, Zhang *et al*. [51] advocated using a K-means cluster as a technique for enhancing tiny YOLOv3 in order to raise the accuracy of pedestrian identification.

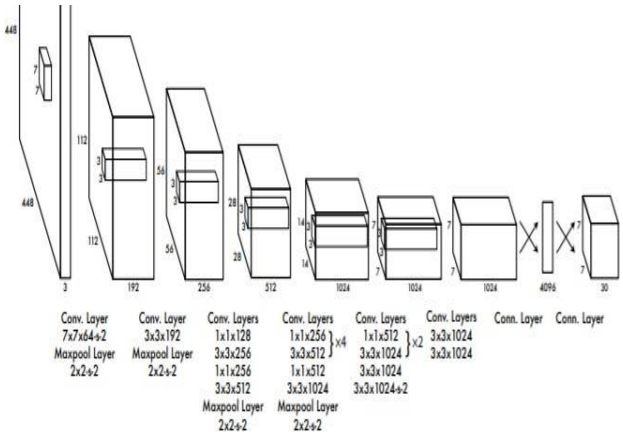


Figure 5. Illustration of YOLOv3 Architecture.

3.4. YOLOv7 Model

A new version of YOLO, known as YOLOv7, was developed by Wang *et al*. [42] in July 2022. Compared to its predecessors, the YOLOv7 model is quicker and more precise. In terms of real-time object identification with a mean average precision that ranges from 51.4% to 56.8%. However, the architecture of YOLOv7 is inspired by the original YOLOv4 and scaled versions of that design. The YOLOv7 architect is shown in Figure 6. In addition, the new Extended Efficient Layer Aggregation Network (E-ELAN) is used as the backbone layer of the system. The E-ELAN was developed to provide a detection system that is both more accurate and faster. However, in contrast to earlier networks, such as the first iteration of the ELAN and CSPVoVNet, the E-ELAN adds three extra components to the training layer. These components are known as shuffle, merge, and expand, respectively. One of the fundamental tenets of YOLOv7 is to enhance detection accuracy and performance while simultaneously decreasing the number of parameters and the amount of processing required.

On the other hand, YOLOv7 employs not one but two head layers, namely the lead head and the auxiliary head, for the detection layer. These two layers interact with one another to provide a more accurate representation of the correlation and distribution of the data. On the other hand, the authors' experimental findings demonstrate that YOLOv7 performs better than other models such as tiny YOLOv4, YOLOv4, and YOLOR. In modern times, YOLOv7 is beneficial in diagnosing several medical conditions. In the experimental investigation that Bayram *et al*. [7] conducted on diagnosing kidney disorders, the authors found that YOLOv7 has the highest mean average accuracy of 85% at IoU of 50%.

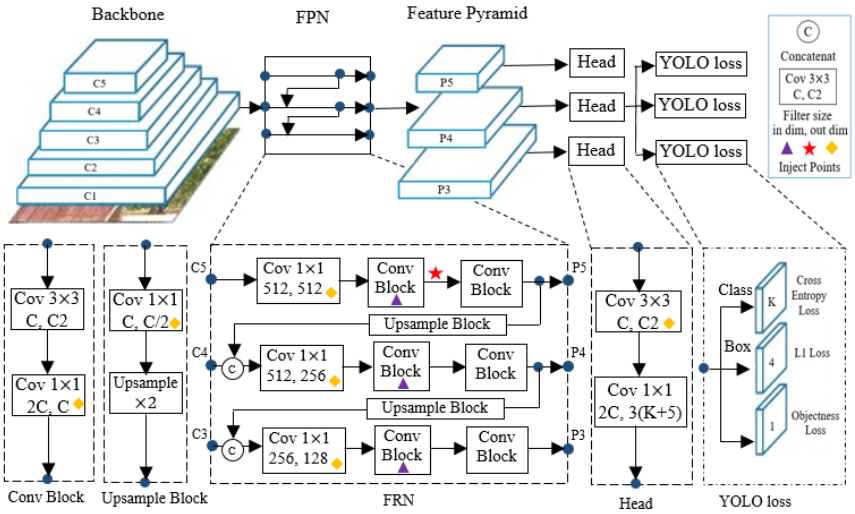


Figure 6. Illustration of YOLOv7 architecture.

4. Evaluation Metrics

We employ three evaluation metrics for the YOLO models: precision, recall, and mean Average Precision (mAP). We aim to determine the ideal YOLO network size and weight with the highest accuracy. The mAP metric, on the other hand, analyzes the effectiveness of object detection by calculating the mean of the average precision of the entire data class concerning the Intersection over Union (IoU) value [15]. The mAP value is derived from the confusion matrix, IoU, precision, and recall. Based on the ground truth bounding box, the confusion matrix summarizes the classification and detection outcomes regarding correctly and erroneously categorized objects. Precision is the proportion of correct predictions (True positive) relative to the total number of correct samples (True positive +false positive). The recall metric quantifies the proportion of accurate predictions (True positives) relative to the total number of relevant samples across all labels. A confusion matrix has four main qualities that may be used to construct alternative assessment metrics:

- True Positive (TP): indicates the number of brain tumors appropriately diagnosed.
- True Negative (TN): number of accurately identified non-tumors.

- False Positive (FP): the number of misclassified non-tumors that are tumors.
- False Negative (FN): the number of tumors misclassified as non-tumors.

The following metrics were computed and derived based on the confusion matrix:

$$Mean\ average\ precision\ (mAP) = \frac{1}{n}\sum_{k=1}^{k=n} Avereage\ precision\ of\ class\ k \quad (1)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (3)$$

$$Intersection\ over\ union\ (IoU) = \frac{Overlaped\ area\ between\ the\ predicted\ and\ ground\ truth\ boxes}{Area\ of\ union} \quad (4)$$

We also employ three loss functions for minimization and evaluation, including the bounding box regression score (loss), which may be used to assess non-overlapping bounding boxes [20]. The class probability score may be used to determine how well a bounding box matches the class of an item [33]. The objectness score (confidence score/GIoU) may be used to assess the likelihood of a certain object being in a grid cell [44].

$$Objectness\ score\ (confidence) = Propability(object) * IoU_{Predicted}^{Truth}) \quad (5)$$

$$Bounding\ box\ regression\ loss = Mean\ squared\ error\ (x^{predicted}, x^{truth}) \quad (6)$$

$$Class\ probability\ score = Propability(class_i|object) \quad (7)$$

Where:

$$Propability(object) = \begin{cases} 1, & There\ is\ no\ object \\ 0, & There\ is\ object \end{cases} \quad (8)$$

$$Probability(class_i) = The\ probability\ of\ object\ that\ belong\ to\ class_i \quad (9)$$

5. Experimental Results and Discussion

In cancer detection, supervised learning and detection models are used. The type and quality of data are thought to have the most significant influence on detecting operations. However, detection models need labeled datasets, which are costly and necessitate expert identification of illnesses to minimize confusion with symptoms or other conditions. As a result, we employ a variety of data augmentation approaches to boost the number of training components to get reliable detection results. On the other hand, brain tumor detection procedures have a specific and unique instance in which brain cancers may be discovered using a collection of magnetic resonance images of varying dimensions. Therefore, detecting tumors alone is not considered sufficient; instead, it is necessary to concentrate and conduct research on the side of the images with the highest discriminatory power of the automated detection compilations, highlighting their advantages and limitations while considering medical considerations.

This section shows the experimental findings of comparing several YOLO models for identifying brain cancers. Our objective is to find the optimum YOLO model and MRI orientation for tumor detection in terms of accuracy and performance. However, image processing and classification are known to have high equipment requirements, such as large amounts of RAM and a powerful GPU. The experiment setup and device qualification are shown in Table 4.

Table 4. Experiment setup and simulation device qualifications.

Device Specification	Description
Processor	Intel(R) Core i7 10^{th} generation.
RAM	8 Gigabits.
Operating system	Windows x64
CPU	1.50 GHz
GPU	NVIDIA GeForce MX230

For hyperparameters tuning, YOLO models comprise around 29 distinct parameters in total. However, as shown in Table 5, we set up twelve parameters. The parameters are loss gain functions, learning rates, optimizers, and IoU threshold. All images were resized to 416×416 for all models as input image size. However, due to low-weight networks such as Nano and small models, we increase the number of epochs to 100 iterations in YOLOv5 for detection results improvement. To make the comparison more accurate, we set all other YOLOv5 models, including medium, large, and x-large, to 100 epochs.

Table 5. Hyperparameter tuning and data-augmentation processing.

Parameters	Detection Models		
	YOLOv3	YOLOv5	YOLOv7
Initial learning rate (lr0)	0.01	0.01	0.01
Final learning rate (lrf)	0.1	0.01	0.1
Momentum	0.937	0.937	0.937
Box loss gain	0.05	0.05	0.05
Classification loss gain	0.5	0.5	0.3
Objectness loss gain	1.0	1.0	0.7
IoU training threshold	0.2	0.2	0.2
Optimizer	SGD	SGD	SGD/Adam
Anchors per output layer	6.14	6.14	6.02
Image input size	416 x 416	416 x 416	416 x 416
Batches	16	16	16
Epochs	50	100	60
Data-Augmentation	For images (Flip horizontally and vertically). For bounding boxes (Flip horizontally and vertically)		

In contrast to YOLOv5, we chose 50 epochs to train the YOLOv3 model and 60 epochs to train the YOLOv7 model. However, all models were set up to 16 batches compatible with the small chosen learning rates and device qualifications in terms of RAM and GPU, as discussed in Table 4. For clarity, only four images will be fed at once into the model in each iteration. We use the SGD optimizer for optimizers in YOLOv3 and YOLOv5 models. However, to our knowledge, the Adam optimizer performs worse than the SGD optimizer even though the Adam optimizer converges

faster [13]. Therefore, we compare the state-of-the-art YOLOv7 model using two different optimizers, SGD and Adam optimizer.

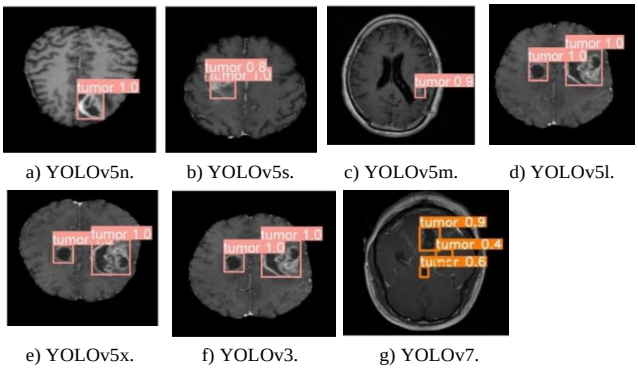
The efficiency of several YOLO models for tumor detection across each dataset is discussed separately in the following sections. Tables 6 and 7 show the findings of brain tumor detection over the axial orientation using different YOLO models before and after data augmentation. Also, to emphasize the improvement in outcomes, the average was calculated. However, when comparing the outcomes before and after data augmentation. It was discovered that there was a result enhancement with 0.432 average precision, 0.242 average recall, 0.431 average mAP at IoU of 0.5, and 0.503 average mAP at IoU of 0.5 to 0.95. Furthermore, YOLOv3 surpasses others with an mAP of 62.4% at IoU of 0.5 and a precision of 63.1% before data augmentation. YOLOv7 also has the lowest mAP of 23.5% at IoU of 0.5 and precision of 32.4%. However, YOLOv7 has the highest recall rate of 82.4% compared to other models. Data augmentation revealed that the YOLOv5 x-large surpasses others, with an mAP of 99.5% at IoU of 0.5, a precision of 99%, and an mAP of 93% for IoU of 0.5 to 0.95. YOLOv7, on the other hand, has inferior detection outcomes with 0.027 classification loss gain and 0.66 box loss gain. As a result, the YOLOv3 and all YOLOv5 models outperform the cutting-edge YOLOv7 model on the axial dataset. Also, in contrast to YOLOv3 and YOLOv5, it was found that YOLOv7 is more affected by data augmentation. However, as shown in Figure 7, the YOLOv3 and YOLOv5 models have a high tumor detection accuracy ranging from 90% to 100%. YOlOv7, on the other hand, achieves detection rates ranging from 30% to 90% across all samples, despite occasional misclassifications.

Table 6. Shows tumor detection results over the axial dataset before augmentation with different YOL weights.

Model	Precision	Recall	mAP 50	mAP 50-95	Obj loss	Cls loss	Box loss
YOLO v5n	0.61746	0.70329	0.61684	0.43355	0.0052418	0.034826	0.025097
YOLO v5s	0.60834	0.68293	0.59927	0.43204	0.0054891	0.036905	0.024193
YOLO v5m	0.51929	0.73614	0.56837	0.40287	0.0054119	0.045909	0.023632
YOLO v5l	0.51097	0.69296	0.49397	0.35077	0.005594	0.058388	0.022582
YOLO v5x	0.56775	0.68737	0.58189	0.41925	0.0054035	0.044001	0.022247
YOLO v3	0.63109	0.69901	0.62469	0.4512	0.0052822	0.037102	0.02407
YOLO v7	0.3242	0.8242	0.3869	0.2358	0.005279	0.01698	0.08777
Average	0.52694	0.72044	0.54252	0.38199	0.00541	0.039881	0.034082

Table 7. Shows tumor detection results over the axial dataset following data-augmentation with different YOLO weights.

Model	Precision	Recall	mAP 50	mAP 50-95	Obj loss	Cls loss	Box loss
YOLO v5n	0.99076	0.98204	0.99009	0.8639	0.0025082	0.0000361	0.012857
YOLO v5s	0.99898	0.99401	0.99497	0.90149	0.0021793	0.0000289	0.010736
YOLO v5m	0.99938	0.99401	0.99485	0.91338	0.0019784	0.0000256	0.0092766
YOLO v5l	0.99932	0.99401	0.99494	0.92649	0.0020109	0.0000254	0.0086069
YOLO v5x	0.999	0.997	0.995	0.930	0.0081	0.0018	0.000021
YOLO v3	0.986	0.994	0.994	0.903	0.0021	0.00008	0.0105
YOLO v7	0.7657	0.8007	0.8661	0.731	0.003841	0.02762	0.06685
Average	0.95806	0.96229	0.97331	0.88423	0.0033683	0.00493	0.0176651

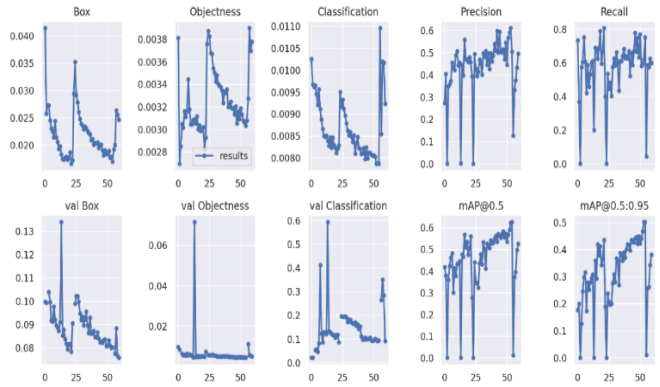


a) YOLOv5n. b) YOLOv5s. c) YOLOv5m. d) YOLOv5l.

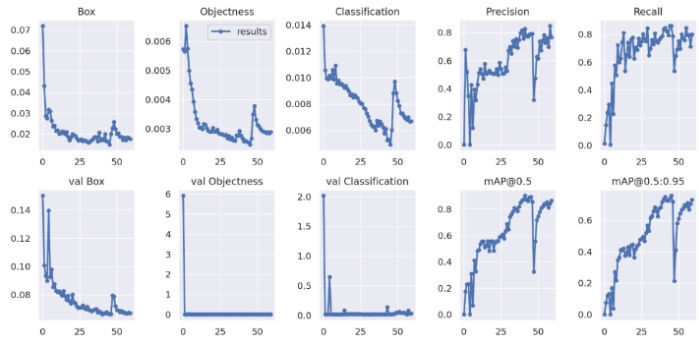
e) YOLOv5x. f) YOLOv3. g) YOLOv7.

Figure 7. Shows a sample of improved brain tumor detection results and accuracy over the axial dataset following data-augmentation.

For the YOLOv7 optimizers' comparison, it was discovered that YOLOv7 does not perform well when utilizing the Adam optimizer, most likely owing to numerous placed labels (number of detected labels). However, the Adam optimizer only recognized labels ranging from 0 to 7 in each epoch. As a result, YOLOv7 yields 40 to 56% mAP. In SGD, identified labels rose from 7 to 29 in each epoch, indicating improved outcomes. Therefore, we choose the SGD optimizer for the YOLOv7 detection model. Figure 8 shows the results of comparing the Adam and SGD optimizers using YOLOv7.



a) Adam optimizer performance.



b) SGD optimizer performance.

Figure 8. Compares adam and SGD optimizers performance using YOLOv7 Model over axial dataset.

Tables 8 and 9 show the tumor detection accuracy and performance for the coronal orientation dataset using YOLO models before and after data augmentation. However, compared to others before data

augmentation, YOLOv3 has the highest detection outcomes with an mAP of 78.2% at IoU of 0.5. Furthermore, YOLOv5 x-large and medium weights rank second in terms of mean average precision. YOLOv7, like the axial orientation dataset, performs the worst, with a mAP of 54% and an IoU of 50%. Also, compared to the axial orientation dataset and other detection models, YOLOv7 is the least affected by data augmentation, with results enhancement ranging from 1 to 3% of total assessment metrics.

Nonetheless, when comparing data augmentation outcomes, we find that increased data affects YOLO models, resulting in improved results, which is unsurprising given that data augmentation methods lower detector sensitivity. The results reveal that detection results improve by 0.317 average precision, 0.239 average mAP at IoU of 0.5, and 0.343 average mAP at IoU of 0.5 to 0.95. Similarly to the axial dataset, the YOLOv5 small weight has the highest detection results after data augmentation, with an mAP of 99.3%, demonstrating that the small weight is the most favorably affected by data increase. However, as shown in Figure 9, all models have significant detection accuracy over coronal orientation. Compared to axial orientation, YOLOv5n, YOLOv5s, YOLOv5x, and YOLOv3 have the highest detection stability.

Table 8. Shows tumor detection results over the coronal dataset before data-augmentation with different YOLO weights.

Model	Precision	Recall	mAP 50	mAP 50-95	Obj loss	Cls loss	Box loss
YOLO v5n	0.65431	0.67047	0.6923	0.46631	0.0041891	0.031565	0.024857
YOLO v5s	0.5559	0.79502	0.68246	0.4927	0.0040685	0.042924	0.023766
YOLO v5m	0.67332	0.74448	0.71691	0.53649	0.0041634	0.038446	0.021936
YOLO v5l	0.57298	0.73733	0.6436	0.48649	0.0045528	0.053788	0.02338
YOLO v5x	0.61021	0.82697	0.71946	0.5315	0.0045794	0.043106	0.023826
YOLO v3	0.68968	0.79627	0.78289	0.59829	0.0044065	0.032529	0.02301
YOLO v7	0.4808	0.8538	0.5405	0.4063	0.004962	0.02106	0.08383
Average	0.59715	0.79231	0.68097	0.50863	0.0044554	0.038642	0.033291

Table 9. Shows tumor detection results over the coronal dataset following data-augmentation with different YOLO weights.

Model	Precision	Recall	mAP 50	mAP 50-95	Obj loss	Cls loss	Box loss
YOLO v5n	0.97445	0.9856	0.99283	0.88498	0.0023926	0.0001683	0.010644
YOLO v5s	0.9994	0.98735	0.99359	0.92035	0.0019996	0.0002368	0.0089049
YOLO v5m	0.99415	0.9835	0.99185	0.92973	0.0017931	0.0003926	0.0079115
YOLO v5l	0.99303	0.99028	0.99109	0.93772	0.0017671	0.0005378	0.0074157
YOLO v5x	0.9934	0.98055	0.99053	0.94204	0.0017561	0.0001419	0.0068884
YOLO v3	0.994	0.994	0.991	0.926	0.001877	0.0009482	0.0085525
YOLO v7	0.5115	0.8996	0.5564	0.4532	0.003962	0.02854	0.06537
Average	0.91425	0.97255	0.91908	0.85151	0.0021925	0.0051329	0.0175072

a) YOLOv5n. b) YOLOv5s. c) YOLOv5m. d) YOLOv5l. e) YOLOv5x.

f) YOLOv3. g) YOLOv7.

Figure 9. Shows a sample of improved brain tumor detection results and accuracy over the coronal dataset following augmentation.

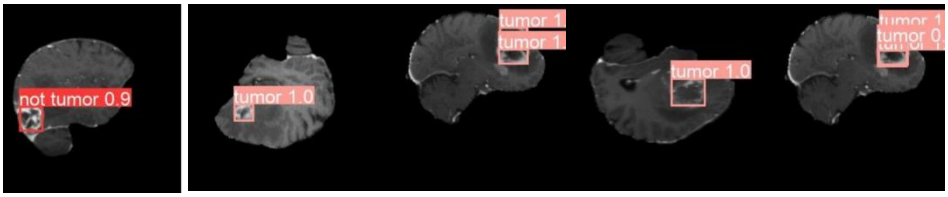
Tables 10 and 11 show the tumor detection accuracy and performance using YOLO models before and after data augmentation for the sagittal orientation dataset. However, following data augmentation, detection accuracy improves by 0.434 average precision, 0.388 average mAP at IoU=0.5, and 0.207 average recall. Furthermore, we found that the YOLOv5 nano weight surpasses others with an mAP of 96.7% at an IoU of 50%. It is worth noting that, like with the coronal orientation dataset, increasing network weight significantly impacts detection accuracy. However, building a large weight network does not necessarily improve detection outcomes. As a result, it is critical to identify the relationship between data augmentation and detector weights. YOLOv3, on the other hand, has a high positive sensitivity to data augmentation, with precision increased by 59.8%, mAP by 42.9% at IoU of 0.5, and 0.077 classification loss gain reduction. YOLOv7 is the least susceptible to data increase, yet it improves outcomes by 0.064 compared to coronal orientation. Furthermore, as compared to the axial and coronal orientations, the sagittal dataset has the lowest detection accuracy, with an average mAP of 89.9%. However, as shown in Figure 10, all models exhibit significant detection performance and accuracy.

Table 10. Shows tumor detection results over the sagittal dataset before data-augmentation with different yolo weights.

Model	Precision	Recall	mAP 50	mAP 50-95	Obj loss	Cls loss	Box loss
YOLO v5n	0.48354	0.83131	0.55339	0.35948	0.0051868	0.028463	0.0299
YOLO v5s	0.46457	0.85612	0.54414	0.3966	0.0042979	0.028271	0.027036
YOLO v5m	0.49976	0.78951	0.55119	0.3871	0.0044313	0.057906	0.027464
YOLO v5l	0.42547	0.57456	0.46628	0.34213	0.0053369	0.072441	0.028687
YOLO v5x	0.46981	0.57894	0.49138	0.34809	0.0049407	0.071224	0.030903
YOLO v3	0.36036	0.86259	0.52743	0.36598	0.0047536	0.083155	0.031451
YOLO v7	0.4622	0.7664	0.4912	0.3465	0.004968	0.01662	0.08737
Average	0.44703	0.73802	0.51194	0.3644	0.0047881	0.054936	0.038819

Table 11. Shows tumor detection results over the sagittal dataset following data-augmentation with different yolo weights.

Model	Precision	Recall	mAP 50	mAP 50-95	Obj loss	Cls loss	Box loss
YOLO v5n	0.94046	0.94527	0.96799	0.85926	0.0026652	0.0056824	0.013165
YOLO v5s	0.96164	0.94119	0.95512	0.8765	0.002439	0.0069707	0.01133
YOLO v5m	0.95574	0.9523	0.95932	0.8893	0.002121	0.00835	0.010057
YOLO v5l	0.95606	0.94729	0.9502	0.89003	0.0021369	0.0085233	0.0094977
YOLO v5x	0.95589	0.94631	0.95412	0.9	0.0020497	0.0083765	0.0083264
YOLO v3	0.95817	0.92698	0.9566	0.87865	0.002192	0.0061747	0.0106
YOLO v7	0.5014	0.9595	0.6201	0.552	0.002978	0.021	0.06583
Average	0.88148	0.9456	0.89924	0.83108	0.0023194	0.0098992	0.0192735



a) YOLOv5n. b) YOLOv5s. c) YOLOv5m. d) YOLOv5l. e) YOLOv5x.

f) YOLOv3. g) YOLOv7.

Figure 10. Shows a sample of improved brain tumor detection results and accuracy over the sagittal dataset following data-augmentation.

For medical considerations, it should be noted that the objective of employing MRI layers is to cover all of the critical components by which cancers may be diagnosed precisely and clearly, where the tumor's location, size, and kind may be determined. On the other hand, the study's findings revealed a considerable improvement in the accuracy of identifying brain cancers utilizing magnetic resonance imaging in all dimensions. However, utilizing axial images yielded the greatest detection results. This is owing to the nature of the axis' dimensions, centered on the X and Y points compared to others. Furthermore, it provides an upper and precise coverage of the right and left sides of the brain, providing more recognized data patterns.

To highlight the state-of-the-art of the study compared to other studies. First, while reviewing a variety of internet datasets for brain tumor detection, we discovered that many of the data lacked labeling. As a result, in this work, we apply a series of data augmentation approaches to enhance the training set and lessen the sensitivity of detection models from future detection operations using any other data set than the one used in the study. Also, this may be utilized to construct detection models that can learn from unlabeled data sets by discovering methods to enhance the data at the start and then transferring this knowledge to specialized categorization procedures. Second, in this work, we employ all axes of magnetic resonance imaging, including axial, sagittal, and coronal, to find the ideal image dimensions that may be used to achieve the maximum discriminatory accuracy in detection. Third, we compare the most recent detection models, such as YOLOv5, YOLOv7, and YOLOv3, to find the most accurate model. Fourth, with the introduction of novel detection models such as YOLOv7, evaluating the model's performance using well-known optimizers such as Adam and SGD is useful. Finally, we analyze the performance of several YOLO models using a comprehensive set of evaluation metrics to demonstrate detection speed, performance, and detection error.

Concerning the study's limitations, it should be highlighted that altering the parameters of detection models may change from one trial to the next, resulting in erroneous detection findings. Furthermore, the quality of the images used to diagnose brain tumors significantly enhances detection accuracy. Since we used images at a resolution of 416*416 for this study, enlarging the image could result in a lower-quality image. Furthermore, we discovered that all YOLO models are susceptible to data augmentation strategies, with the YOLOv7 model being the least affected. Finally, this study did not discover any micro-tumor. The study was restricted to the early detection of cancers with two primary categories (tumor and no tumor). This is owing to a paucity of online addressable data sets, particularly for detection operations, whose addressing procedure is quite costly and necessitates using radiologists for addressing operations.

In short, the results show that the YOLOv5 and YOLOv3 models are more sensitive to data augmentation than the YOLOv7 model. In addition, we show that the axial orientation has higher tumor

detection accuracy than the other orientations. However, based on the statistical results, large-weight models are more likely to recognize data samples than uncover data patterns. As a result, YOLOv7 performs the worst compared to others, in which the number of identified labels (Classes) in each epoch ranged from 0 to 7. Nonetheless, all models revealed significant detection accuracy.

6. Conclusions

The developed methodology evaluates the performance of state-of-the-art YOLO models on a dataset of 7382 samples from three different MRI orientations (axial, coronal, and sagittal) using different weights and degrees of data augmentation. Many data augmentation approaches were used to reduce detector sensitivity and enhance detection accuracy. Furthermore, a comparison was made between the Adam and SGD optimizers. We need to determine the optimal network weight and MRI orientation to detect brain tumors with MRI. With an IoU of 0.5, the results show that the average mAP for axial orientation is 97.33 percent.

Additionally, SGD outperforms Adam optimizer by over 20% mAP. In addition, YOLOv5n, YOLOv5s, YOLOv5x, and YOLOv3 were discovered to have a mAP of greater than 95%. Furthermore, the YOLOv5 and YOLOv3 models were more sensitive to data augmentation than the YOLOv7 model. Using the proposed framework for brain tumor diagnosis has a moderate computational cost and a small space requirement; as a result, it is capable of running on most systems.

Acknowledgement

The authors would like to thank the Deanship of Scientific Research at Shaqra University for supporting this research.

References

[1] Abiwinanda N., Hanif M., Hesaputra S., Handayani A., and Mengko T., "Brain tumor Classification Using Convolutional Neural Network," *in Proceedings of the World congress on Medical Physics and Biomedical Engineering*, Prague, pp. 183-189, 2019. DOI: 10.1007/978-981-10-9035-6_33

[2] Atik M. and Duran Z., "Deep learning-based 3d Face Recognition Using Derived Features from Point Cloud," *The Proceedings of the 3rd International Conference on Smart City Applications*, Karabukp, pp. 797-808, 2020.

[3] Atik S. and Ipbuker C., "Integrating Convolutional Neural Network and Multiresolution Segmentation for Land Cover and Land Use Mapping Using Satellite Imagery," *Applied Science*, vol. 11, no. 12, p. 5551, 2021. https://doi.org/10.3390/app11125551

[4] Badran E., Mahmoud E., and Hamdy N., "An Algorithm for Detecting Brain Tumors in MRI Images," *in Proceedings of the International Conference on Computer Engineering and Systems*, Cairo, pp. 368-373, 2010.

[5] Bakator M., and Radosav D., "Deep Learning and Medical Diagnosis: A Review of Literature," *Multimodal Technology Interact*, vol. 2, no. 3, pp. 47, 2018. https://doi.org/10.3390/mti2030047

[6] Batra P., Hussain I., Abdul Ahad M., Casalino G., and Alam M., "A Novel Memory and Time-Efficient ALPR System Based on YOLOv5," *Sensors*, vol. 22, no. 14, pp. 5283, 2022. https://doi.org/10.3390/s22145283

[7] Bayram A., Gurkan C., Budak A., and Karataş H., "A Detection and Prediction Model Based on Deep Learning Assisted by Explainable Artificial Intelligence for Kidney Diseases," *European Journal of Science and Technology*, no. 40, pp. 67-74, 2022. DOI: 10.31590/ejosat.1171777

[8] Cepni S., Atik M., and Duran Z., "Vehicle Detection Using Different Deep Learning Algorithms from Image Sequence," *Baltic Journal Modern Computing*, vol. 8, no. 2, pp. 347-358, 2020. DOI:10.22364/bjmc.2020.8.2.10

[9] Chan H., Hadjiiski L., and Samala R., "Computer-aided Diagnosis in the Era of Deep Learning," *Medical Physics*, vol. 47, no. 5, pp. e218-e227, 2020. DOI: 10.1002/mp.13764

[10] Chen C., Liu M., Tuzel O., and Xiao J., "R-CNN for Small Object Detection," *in Proceedings of the Asian Conference on Computer Vision*, Taipei pp. 214-230, 2016.

[11] Cheng J., Huang W., Cao S., Yang R., Yang W., Yun Z., Wang Z., and Feng Q., "Enhanced Performance of Brain Tumor Classification Via Tumor Region Augmentation and Partition," *PLoS One*, vol. 10, no. 10, p. e0140381, 2015. 10.1371/journal.pone.0140381

[12] Dai J., Li Y., He K., and Sun J., "R-FCN: Object Detection Via Region-Based Fully Convolutional Networks," *in Proceedings of the 30th Conference on Neural Information Processing Systems*, Barcelona, vol. 29, 2016.

[13] Gupta A., Ramanath R., Shi J., and Keerthi S., "Adam vs. SGD: Closing the Generalization Gap on Image Classification," *in Proceedings of the 13th Annual Workshop on Optimization for Machine Learning*, 2021.

[14] He K., Gkioxari G., Dollár P., and Girshick R., "Mask R-CNN," *in Proceedings of the IEEE International Conference on Computer Vision*, Venice, pp. 2961-2969, 2017. https://doi.org/10.48550/arXiv.1703.06870

[15] Henderson P. and Ferrari V., "End-to-end Training of Object Class Detectors for Mean

Average Precision," *in Proceedings of the Asian Conference on Computer Vision*, Taipei, pp. 198-213, 2016. https://doi.org/10.48550/arXiv.1607.03476

[16] Huang X., Yue X., Xu Z., and Chen Y., "Integrating General and Specific Priors into Deep Convolutional Neural Networks for Bladder Tumor Segmentation," *in Proceedings of the International Joint Conference on Neural Networks*, Shenzhen, pp. 1-8, 2021. doi: 10.1109/IJCNN52387.2021.9533813

[17] Kavitha R., Chitra L., and Kanaga L., "Brain Tumor Segmentation Using Genetic Algorithm with SVM Classifier," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 5, no. 3, pp. 1468-1471, 2016. DOI:10.15662/IJAREEIE.2016.0503043

[18] Khambhata K. and Panchal S., "Multiclass Classification of Brain Tumor in MR Images," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 4, no. 5, pp. 8982-8992, 2016.

[19] Kuznetsova A., Maleva T., and Soloviev V., *Cyber-Physical Systems: Modelling and Intelligent Control*, Springer, 2021. https://doi.org/10.1007/978-3-030-66077-2_28

[20] Lee S., Kwak S., and Cho M., "Universal Bounding Box Regression and Its Applications," *in Proceedings of the Asian Conference on Computer Vision*, Perth, pp. 373-387, 2018. https://doi.org/10.1007/978-3-030-20876-9_24

[21] Litjens G., Kooi T., Bejnordi B., Setio A., Ciompi F., Ghafoorian M., Laak J., Ginneken B., and Sánchez C., "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017. https://doi.org/10.1016/j.media.2017.07.005

[22] Logeswari T. Karnan M., "An Improved Implementation of Brain Tumor Detection Using Segmentation Based on Soft Computing," *Journal of Cancer Research and Experimental Oncology*, vol. 2, no. 1, pp. 006-014, 2010.

[23] Lundervold A. and Lundervold A., "An Overview of Deep Learning in Medical Imaging Focusing on MRI," *Zeitschrift Für Medizinische Physik*, vol. 29, no. 2, pp. 102-127, 2019. https://doi.org/10.1016/j.zemedi.2018.11.002

[24] Magnuska Z., Theek B., Darguzyte M., Palmowski M., and Stickeler E., "Influence of the Computer-Aided Decision Support System Design on Ultrasound-Based Breast Cancer Classification," *Cancers*, vol. 14, no. 2, pp. 277, 2022. DOI: 10.3390/cancers14020277

[25] Mohiyuddin A., Basharat A., Ghani U., Peter V., and Abbas S., "Breast Tumor Detection and Classification in Mammogram Images Using Modified YOLOv5 network," *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022. doi: 10.1155/2022/1359019

[26] Montalbo F., "A Computer-Aided Diagnosis of Brain Tumors Using a Fine-Tuned YOLO-Based Model with Transfer Learning," *KSII Transactions on Internet and Information Systems*, vol. 14, no. 12, pp. 4816-4834, 2020. DOI: 10.3837/tiis.2020.12.011

[27] Nepal U. and Eslamiat H., "Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs," *Sensors*, vol. 22, no. 2, pp. 464, 2022. https://doi.org/10.3390/s22020464

[28] Nogales A., Garcia-Tejedor A., Monge D., Vara J., and Antón C., "A Survey of Deep Learning Models in Medical Therapeutic Areas," *Artificial Intelligence in Medicine*, vol. 112, pp. 102020, 2021. doi: 10.1016/j.artmed.2021.102020.

[29] Oza P., Sharma P., Patel S., Kumar P., "Deep Convolutional Neural Networks for Computer-Aided Breast Cancer Diagnostic: A Survey," *Neural Computing and Applications*, vol. 34, no. 6, pp. 1-22, 2022. DOI: 10.1007/s00521-021-06804-y

[30] Pan Y., Huang W., Lin Z., Zhu W., Zhou J., Wong J., Ding Z., "Brain Tumor Grading Based on Neural Networks and Convolutional Neural Networks," *in Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Milan, pp. 699-702, 2015. DOI: 10.1109/EMBC.2015.7318458

[31] Rahman M. and Wang Y., "Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation," *International Symposium on Visual Computing*, pp. 234-244, 2016.

[32] Redmon J. and Farhadi A., "Yolov3: An Incremental Improvement," 2018.

[33] Redmon J., Divvala S., Girshick R., and Farhadi A., "You Only Look Once: Unified, Real-Time Object Detection," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, pp. 779-788, 2016. doi: 10.1109/CVPR.2016.91

[34] Ren S., He K., Girshick R., and SunJ., "Faster r-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[35] Shelatkar T., Urvashi D., Shorfuzzaman M., Alsufyani A., and Lakshmanna K., "Diagnosis of Brain Tumor Using Light Weight Deep Learning Model with Fine-Tuning Approach," *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022. https://doi.org/10.1155/2022/2858845

[36] Singh L., Chetty G., and Sharma D., "A Novel Machine Learning Approach for Detecting the Brain Abnormalities from Mri Structural Images,"

in Proceedings of IAPR International Conference On Pattern Recognition in Bioinformatics, Tokyo, 2012, pp. 94-105.

[37] Swati Z., Zhao Q., Kabir M., Ali F., Ali Z., Ahmed S., and Lu J., "Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning," *IEEE Access*, vol. 7, pp. 17809-17822, 2019. 10.1109/ACCESS.2019.2892455

[38] Tian L., Thalmann N., Thalmann D., Fang Z., and Zheng J., "Object grasping of humanoid robot based on YOLO," *in Proceedings of the Computer Graphics International Conference*, Calgary 2019, pp. 476-482: Springer.

[39] Vengaloor R. and Muralidhar R., "Deep Learning Based Feature Discriminability Boosted Concurrent Metal Surface Defect Detection System Using YOLOv-5s-FRN," *The International Arab Journal of Information Technology*, vol. 21, no. 1, pp. 94-106, 2024. https://doi.org/10.34028/iajit/21/1/9

[40] Walia I., Kumar D., Sharma K., Hemanth J., and Popescu D., "An Integrated Approach for Monitoring Social Distancing and Face Mask Detection Using Stacked ResNet-50 and YOLOv5," *Electronics*, vol. 10, no. 23, pp. 2996, 2021.https://doi.org/10.3390/electronics10232996

[41] Wang C., Zhang Y., Zhou Y., Sun S., Zhang H., and Wang Y., "Automatic Detection of Indoor Occupancy Based on Improved YOLOv5 model," *Neural Computing and Applications*, vol. 23, pp. 2575-2599, 2023. https://doi.org/10.1007/s00521-022-07730-3

[42] Wang C., Bochkovskiy A., and Liao H., "YOLOv7: Trainable Bag-Of-Freebies Sets New State-of-The-Art for Real-Time Object Detectors," 2023 IEEE/CVF Conference on Computer *in Proceedings of the Vision and Pattern Recognition*, Vancouver, 2022.

[43] Wang Z., Wu L., Li T., and Shi P., "A Smoke Detection Model Based on Improved YOLOv5," *Mathematics*, vol. 10, no. 7, pp. 1190, 2022. https://doi.org/10.3390/math10071190

[44] Wenkel S., Alhazmi K., Liiv T., Alrshoud S., and Simon M., "Confidence Score: The Forgotten Dimension of Object Detection Performance Evaluation," *Sensors*, vol. 21, no. 13, pp. 4350, 2021. https://doi.org/10.3390/s21134350

[45] White N., McDonald C., Farid N., Kuperman J., Kesari S., and Dale A., "Improved Conspicuity and Delineation of High-Grade Primary and Metastatic Brain Tumors Using "Restriction Spectrum Imaging": Quantitative Comparison with High B-Value DWI and ADC," *American Journal of Neuroradiology*, vol. 34, no. 5, pp. 958-964, 2013.

[46] Xianjia Y., Salimpour S., Queralta J., and Westerlund T., "Analyzing General-Purpose Deep-Learning Detection and Segmentation Models with Images from a Lidar as a Camera Sensor," *arXiv Preprint*, vol. arXiv:2203.04064v1, pp. 1-6, 2022. https://arxiv.org/pdf/2203.04064

[47] Xu S., Guo Z., Liu Y., Fan J., and Liu X., "An Improved Lightweight YOLOV5 Model Based on Attention Mechanism for Face Mask Detection," *in Proceedings of the International Conference on Artificial Neural Networks*, Bristol, pp. 531-543, 2022. https://doi.org/10.1007/978-3-031-15934-3_44

[48] Yan B., Fan P., Lei X., Liu Z., and Yang F., "A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5," *Remote Sensing*, vol. 13, no. 9, pp. 1619, 2021. https://doi.org/10.3390/rs13091619

[49] Zacharaki E ., Wang S., Chawla S., Yoo D., Wolf R., Melhem E., and Davatzikos C., "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, 2009. doi: 10.1002/mrm.22147.

[50] Zhang S., Zhang F., Ding Y., Li Y., "Swin-YOLOv5: Research and Application of Fire and Smoke Detection Algorithm Based on YOLOv5," *Computer Intelligence Neuroscience*, vol. 2022, 2022. doi: 10.1155/2022/6081680

[51] Zhang Y., Shen Y., and Zhang J., "An Improved Tiny-YOLOv3 Pedestrian Detection Algorithm," *Optik*, vol. 183, pp. 17-23, 2019. DOI:10.1016/j.ijleo.2019.02.038

Abdelraouf Ishtaiwi is a highly accomplished academic with over 22 years of experience in teaching and research in the field of Artificial Intelligence (AI). He earned his Master's degree in AI from Griffith University, Brisbane in 2001, followed by a Ph.D. in the same field in 2007. Dr. Ishtaiwi's academic career has been dedicated to advancing the field of AI through exceptional research and teaching skills. His expertise in the field has led to numerous top scientific contributions, including groundbreaking research on machine learning, local search algorithms, and optimization methods. Throughout his career, Dr. Ishtaiwi has published more than 23 research papers in highly regarded academic journals, demonstrating his significant impact on the field of AI. His research has been widely cited and has received recognition from the academic community for its innovative approach to AI. In addition to his research accomplishments, Dr. Ishtaiwi is an experienced teacher and mentor. He has taught a wide range of courses in AI, including advanced topics in machine learning and optimization. His dedication to teaching has earned him accolades from his students and colleagues alike.

Ali Ali is an assistant professor at Communications and Computer Engineering Department, Faculty of Engineering, Al-Ahliyya Amman University. He received a PhD in computer and communications engineering from the University of Huddersfield, UK in 2021. His primary research interests are in the analysis of communication system reliability using complex modelling techniques, network security, machine learning optimization techniques as well as approaches to WLAN optimization.

Ahmad Al-Qerem graduated in applied mathematics and MSc in Computer Science at the Jordan University of Science and Technology and Jordan University in 1997 and 2002, respectively. After that, he was appointed as full-time lecturer at the Zarqa University. He was a visiting professor at Princess Sumaya University for Technology (PSUT). He obtained a PhD from Loughborough University, UK. His research interests are in performance and analytical modeling, mobile computing environments, protocol engineering, communication networks, transition to IPv6, machine learning and transaction processing. He has published several papers in various areas of computer science. Currently, he has a full academic post as a full professor at computer science department at Zarqa University-Jordan.

Yazan Al Smadi is a seasoned software engineering expert, holding a Master's degree in Software Engineering from Zarqa University and a Bachelor of Science degree in Computer Science from Al Balqaa University. With a passion for technology and innovation, Yazan has honed his skills and expertise in the field of software development.
Throughout his academic journey, Yazan demonstrated exceptional dedication and a keen understanding of computer science principles. His academic achievements served as a solid foundation for his professional endeavors. Currently serving as a software development consultant in the private sector, Yazan leverages his extensive knowledge and experience to provide invaluable insights and solutions to various technological challenges. His role involves collaborating with teams, analyzing complex problems, and implementing innovative software solutions tailored to meet the specific needs of clients

Amjad Aldweesh is a computer assistant professor interested in the Blockchain and Smart contracts technology as well as cyber security. Amjad has a Bachelor degree in computer science. He has a MSc degree in advanced computer science and security from the University of Manchester with distinction. Amjad is the second in the UK and the first in the middle east to have a PhD in the Blockchain and Smart contracts technology from Newcastle University.

Mohammad Alauthman Received PhD degree from Northumbria University Newcastle, UK in 2016. He received a B.Sc. degree in Computer Science from Hashemite University, Jordan, in 2002, and received M.Sc. degrees in Computer Science from Amman Arab University, Jordan, in 2004. Currently, he is Assistant Professor and senior lecturer at Department of Information Security, Petra University, Jordan. His main research areas cyber-security, Cyber Forensics, advanced machine learning and data science applications.

Omar Alzoubi is an Assistant Professor of Computer Engineering at the University of Umm Al-Qura in Saudi Arabia. With a focus on Computer Engineering, Dr. Alzoubi brings a wealth of expertise to his role. He is committed to advancing the field through both his research and teaching endeavors. As a respected academic, Dr. Alzoubi is dedicated to nurturing the next generation of computer engineers, guiding students to excel in their studies and research pursuits. His contributions to the university community are invaluable as he works towards furthering knowledge and innovation in the field of Computer Engineering.

Shadi Nashwan was born in Amman, Jordan, in 1978. He received the B.Sc. degree in computer science from Al-Azhar University, Palestine, in 2001, the M.Sc. degree in computer science from University of Jordan, Jordan, in 2003, and the Ph.D. degree in computer and network security from Anglia Ruskin University, U.K., in 2009. From 2009 to 2010, he was an Assistant Professor with the Department of Computer Science, Al-Zaytoonah University, Jordan. At the end of 2010, he became an Assistant Professor with the Computer Science Department, Jouf University, Saudi Arabia. In 2018, he was promoted to the position of associate professor in cybersecurity. In 2022, he was promoted to the position of full professor in cybersecurity. Currently, he has been appointed as a head of cybersecurity and Software Engineering departments at Middle East University, Jordan, in 2023. He has published several articles in the area of authentication protocol, recovery techniques, analytic model, and mobility management. His research interests include authentication protocol for mobile networks, security of healthcare system, security Agriculture system and security of wireless networks, such as NFC, RFID, WSNs, and WMSNs.

Awad Ramadan is a dedicated academician with a strong commitment to excellence in education and administration. He holds the position of Lecturer in the Computer Science Department at the College of Computing in Al-Qunfudah, Umm Al-Qura University, Saudi Arabia, a role he has served in since 2007. Additionally, he plays a vital role as a member of the Academic Oversight Committee, ensuring quality standards in education since 2021. With a wealth of experience and a proactive approach, Awad Mohamed Ramadan continues to make significant contributions to the academic community at Umm Al-Qura University.

Musab Al-Zghoul, PhD, received his doctorate in Computer Science from Don State Technical University. Currently, he serves as an Assistant Professor at Isra University's Information Technology College. His research interests encompass machine learning, data mining, and IoT, and he is enthusiastic about collaborating with both national and international research teams. With a rich academic background, Dr. Al-Zghoul spent three years at Zarqa University in Jordan and eleven years at Umm Al-Qura University in Saudi Arabia. He holds a patent in Russia since 2008 for a “Software Benchmark for Studying Caching Algorithms.” Over the past five years, his publications have centered on classification and NLP in Islamic Research.

Someah Alangari is an Assistant Professor, at College of Science and Humanities Dawadmi at Shaqra University, Saudi Arabia. Dr. Someah has got her Ph.D in Computer Science at University of Southampton in the UK. She received her master's degree in Software Engineering at University of Southampton in the UK, and BSc in Computer Science at King Saud University in Saudi Arabia. Her research interests include ML , software engineering, and information systems. Dr. Someah has over 10 years of working experience in the academic sector.