

A Computer-Aided Diagnosis of Brain Tumors Using a Fine-Tuned YOLO-based Model with Transfer Learning

Francis Jesmar P. Montalbo

College of Informatics and Computing Sciences, Batangas State University
Batangas City, Batangas 4200, Philippines
[e-mail: francismontalbo@ieee.org]

*Corresponding author: Francis Jesmar P. Montalbo

*Received August 5, 2020; revised November 4, 2020; accepted November 23, 2020;
published December 31, 2020*

Abstract

This paper proposes transfer learning and fine-tuning techniques for a deep learning model to detect three distinct brain tumors from Magnetic Resonance Imaging (MRI) scans. In this work, the recent YOLOv4 model trained using a collection of 3064 T1-weighted Contrast-Enhanced (CE)-MRI scans that were pre-processed and labeled for the task. This work trained with the partial 29-layer YOLOv4-Tiny and fine-tuned to work optimally and run efficiently in most platforms with reliable performance. With the help of transfer learning, the model had initial leverage to train faster with pre-trained weights from the COCO dataset, generating a robust set of features required for brain tumor detection. The results yielded the highest mean average precision of 93.14%, a 90.34% precision, 88.58% recall, and 89.45% F1-Score outperforming other previous versions of the YOLO detection models and other studies that used bounding box detections for the same task like Faster R-CNN. As concluded, the YOLOv4-Tiny can work efficiently to detect brain tumors automatically at a rapid phase with the help of proper fine-tuning and transfer learning. This work contributes mainly to assist medical experts in the diagnostic process of brain tumors.

Keywords: Brain Tumors, Computer-Aided Diagnosis, Deep Learning, Medical Image Analysis, Object Detection

This work would have not been possible without the support of Batangas State University and its research environment.

1. Introduction

The Central Nervous System (CNS) is home to some of the most critical human body components, including the brain and the spinal cord. These two mentioned parts perform sophisticated functions that, until today, are not fully understood. One of which is the development of brain tumors, where a massive uncontrolled growth of abnormal cells arises. Benign or malignant, these tumors tend to cause damage and discomfort to a person that can potentially lead to an early demise or life-long complications. In recent studies, the commonly diagnosed tumors consisted of Glioma, Meningioma, and Pituitary tumors. These primary brain tumors contribute to the entire list of cancer cases in the United States for about 1.4% having 20,500 new cases and 12,500 deaths per year [1, 2]. Cerebral glioma is an alarming type of tumor that originates in the glial cells of the brain. This tumor can become overly aggressive and spread quickly to other parts of the body [2]. Currently, glioma affects approximately 14,000 people in the United States a year and is continuously increasing over time [3]. If not diagnosed early, severe glioma can expand greatly and become physically visible [4, 5]. According to the World Health Organization (WHO), this type of tumor frequently falls under a grade of IV. In which the fourth level is considered malignant and fatal compared to its lower counterparts [6]. In contrast, a pituitary tumor is a less progressive brain tumor that develops in the pituitary gland. Unlike glioma, pituitary tumors do not cause violent spread to other parts of the body, and its likelihood to become cancerous is rare [7]. Another common non-glial intracranial tumor, meningioma, affects about six to eight out of 100,000 people a year. This tumor can originate from the brain's meningeal coverings or even on the meninges of the spine. Like the pituitary tumor, meningioma tumors are usually benign and do not require immediate treatment, and are asymptomatic. However, there are instances where it can grow more rapidly than usual, leading to a higher II or III grade where symptoms including frequent morning headaches, vision impairment, or even seizures start to arise [8-11]. Regardless of the degrees and severity, brain tumors must have a proper diagnosis and immediate treatment to reduce the chances of an untimely demise [12]. For that reason, patients experiencing symptoms related to these brain tumors must undergo a non-invasive method such as Magnetic Resonance Imaging (MRI) for an immediate diagnosis [13]. However, the task of analyzing the results produced by MRI scans requires meticulous observation and a high-level of proficiency and is impossible for an ordinary person to perform. Some hospitals and healthcare centers still lack the expertise to serve these needs and adding to the already lengthy diagnosis. [14].

Researchers and the likes had recently developed specialized solutions to overcome the difficulty of diagnosing several life-threatening diseases using automated Computer-Aided Diagnosis (CAD). Through this method, medical instruments can provide rapid and accurate detection to help medical experts lengthen the span or even improve the quality of a patient's life [15].

Moreover, with the help of Machine Learning (ML) and Computer Vision (CV), it produced formidable solutions called Deep Convolutional Neural Networks (DCNN). These DCNNs are state-of-the-art models that had solved complex CAD problems in terms of recognition, classification, segmentation, and even detection [16].

However, most existing CAD solutions for brain tumor detection and identification based on DCNNs does not run efficiently on most platforms and requires intensive computational resources. For most DCNNs, the lighter classification models have limitations as it cannot point out the tumor's exact location. On the other hand, a segmentation model can locate the affected area using a mask and identify the tumor but requires higher computing costs. Having

ordinary devices for such a solution can cause inefficient performance and unsatisfactory results [17-19]. Therefore, this work aims to exhibit and employ a recent DL model to produce a lighter solution with better performance to detect brain tumors from MRIs. Such a solution can emanate better flexibility for medical experts and small healthcare facilities to produce a faster diagnosis of brain tumors in most case scenarios.

2. Related Works

Recently, numerous DL models and various imaging techniques have enhanced CAD systems to distinguish glioma, meningioma, and pituitary brain tumors.

Cheng et al. [20] initiated the use of the Content-based Image Retrieval (CBIR) technique from a robust dataset of 3064 T1-weighted contrast-enhanced (CE) MRI images to segment brain tumors. Their study made use of a novel framework that augmented tumor regions to work as the region of interest (ROI) divided into subregions according to intensities with the use of adaptive spatial division. Successively, they used the Fisher kernel to accumulate all regions and produced an image-level signature, achieving a Mean Average Precision of (mAP) 94.68%. From the dataset contributed by Cheng et al.'s work, many began to explore DL's capability to classify the three brain tumors from MRIs. Through further research, DCNN became well-known due to its performance to classify multiple classes accurately compared to other methods. In the work of Swati et al. [21], they managed to use a pre-trained DCNN like the VGG19 with the help of transfer learning. Transfer learning provided the initial training needs for their custom classifier to use essential image recognition features immediately.

With fine-tuning, the previous VGG19 managed to classify images according to their requirement. Their work concluded that their fine-tuned block-wise VGG19 required a less demanding development cost than other models involving handcrafted feature extractors. The classification accuracy of their work achieved 94.82%. Deepak et al. [22] also applied the transfer learning method to perform a multi-class classification for CAD. Their work follows the patient-level five-fold cross-validation to classify three distinct brain tumors from a similar data source. Using the pre-trained DCNN GoogleNet and their proposed training approach, their work achieved an outstanding 98% accuracy. Their work further established that transfer learning and fine-tuning methods could favorably contribute to brain tumors classification when appropriately performed.

The study of Rehman et al. [23] employed image processing methods to augment their dataset to improve their model performance. Various affine transformations of image samples helped their selected DCNNs like AlexNet, GoogleNet, and VGG16 to produce more features. Their classifiers attained accuracies of 97.39%, 98.04%, and 98.69%, respectively. Instead of using pre-trained models, Sultan et al. DCNNs [24] proposed a custom-built CNN model to perform a multi-class classification from the same set of brain tumors mentioned. The CNN structure used several activation functions, normalization, pooling, and dropout layers to control overfitting. Their work attained a significant performance of 98.7% accuracy rate that outperformed other state-of-the-art, giving importance to the impact of model tuning on image classification. CNN's limitless potential for image classification continued to progress with newer models with a more sophisticated approach. Noreen et al. [25] proposed using a more recent DCNN model, namely, DenseNet201 and InceptionV3, to diagnose brain tumors. Their approach developed a concatenated multi-stage feature extraction of tumors and generated accurate predictions. As a result of their work, they produced 99.34% for InceptionV3 and 99.51% for DenseNet201. Apart from the mentioned classification and segmentation studies

for brain tumors, Bhanothu et al. [26] used an object detection model called Faster R-CNN. Their selected work detected locations of brain tumors in MRIs with bound boxes and identified them specifically. However, due to the infancy stage of DCNN object detection methods, Faster R-CNN struggled to solve the problem and only attained an mAP of 77.60%.

The following works mentioned summarized in Table 1 displayed a remarkable performance in brain tumor diagnosis. Most of the works attained significant results mainly with classifications and handcrafted segmentation methods, unlike object detection. However, according to recently conducted medical imaging studies, a recent object detection model had improved to compete with existing solutions with better outcomes.

Table 1. Summary of several algorithms applied in brain tumor diagnosis

Author	Model	Approach	Performance	Year
Cheng et al.	Novel-segmentation model	Segmentation	mAP of 94.68%	2016
Bhanothu et al.	Faster R-CNN	Detection	mAP 77.60%	2020
Swati et al.	VGG19	Classification	Accuracy of 94.82%	2019
Deepak et al.	GoogleNet	Classification	Accuracy of 98%	2019
Rehman et al.	AlexNet, GoogleNet, VGG16	Classification	Accuracies of 97.39%, 98.04%, and 98.69%, respectively	2019
Sultan et. al	Custom-CNN	Classification	Accuracy of 98.7%	2019
Noreen et al.	DenseNet201 and InceptionV3	Classification	Accuracies of 99.34% and 99.51%	2020

In the study of Al-masni et al. [27], they applied the “You Only Look Once” (YOLO) detection model for the simultaneous detection of breast masses from multiple digital mammograms. The name YOLO pertains to its capability to assess the entire input image rather than just patches of regions during training and testing phases. This approach gave a substantial speed advantage and less overhead use of YOLO over the other region-based algorithms like Fast R-CNN [28] and the traditional window-sliding technique [29]. They concluded that YOLO has a significant potential to improve CAD as their work detected masses accurately with a score of 99.7%. Another study by Ünver et al. [30] also proved YOLO’s proficiency in medical imaging. Their YOLOv3 model trained using 2000 labeled images to detect skin lesions without any augmentation. Added with a segmentation algorithm, GrabCut, their work achieved 93.39% accuracy from 500 validation images and outperformed other works that used ResNet and U-Net models, concluding that YOLO can significantly help solve other medical imaging problems.

In the recent past, YOLO models required significantly high computational requirements with mediocre performance, making it difficult to trust for future adaptations and deployment [31]. With that said, only a few considered employing a state-of-the-art model like YOLO for medical imaging. However, with the recent advancements of YOLOv3 and the recently released YOLOv4 [32], object detection achieved a notable increase in performance compared to other similar solutions with less expense. Therefore, this work proposes to analyze the performance of the recent YOLOv4 model in terms of training and yielding an automated CAD detection model that focuses on glioma, meningioma, and pituitary brain tumors to assist medical experts in the diagnostic process. This work also aims to employ a lighter YOLOv4 model with lesser disk and computational consumption to ease deployment on most platforms. Given these statements and performance drawn from other medical imaging studies that used YOLO, automated brain tumor detection from MRIs can still improve. No other works have

yet employed the YOLOv4 for brain tumor detection in MRIs during this time and justified its overall precision.

The rest of the article consists of the given order: Section 3 includes the materials and methods to train the YOLO-based model with the sourced dataset. Section 4 evaluates the results of the following experiments conducted based on the proposed method. Lastly, Section 5 provides a conclusion based on the produced results.

3. Materials and Methods

3.1 The Proposed MRI Brain Tumor Detection Model

This section provides a brief discussion of the proposed MRI brain tumor detection model. As presented in Fig. 1, this work selected to use the YOLOv4-Tiny, a smaller version of the YOLOv4 model, to consume less computational resources [32]. The model initially pre-trained with the COCO dataset [33] to attain primitive image recognition features and alleviate data scarcity. However, the pre-trained YOLOv4-Tiny model did not immediately recognize MRI brain tumors due to the current hyper-parameter settings and pre-learned features acquired from the COCO dataset. Therefore, this work also fine-tuned and re-purposed the model to detect the object of interest only. Once pre-trained and fine-tuned, the model re-trained to initialize new weights from the sourced dataset of labeled MRI brain tumors. This process generated a YOLO-based model that detects meningioma, glioma, and pituitary tumors from MRIs.

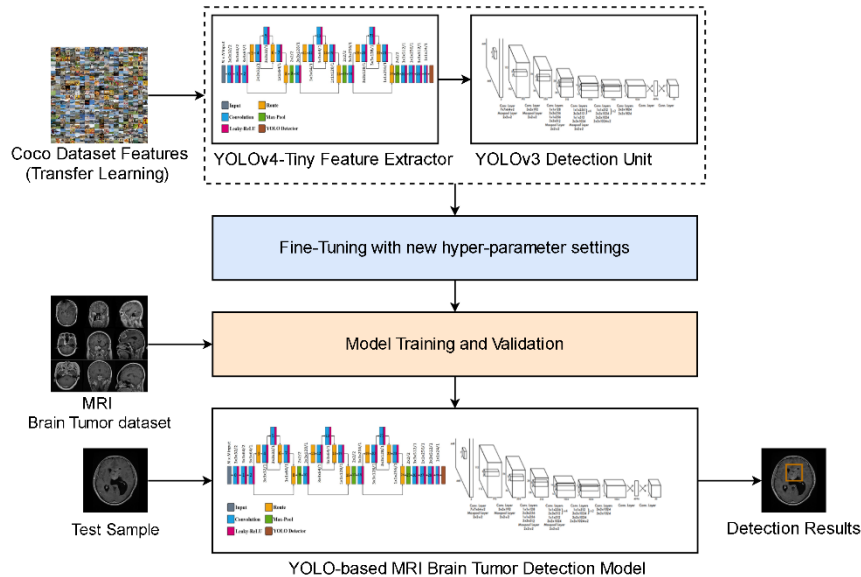


Fig. 1. The blueprint of the proposed YOLO-based MRI brain tumor detection model

3.2 Dataset Preparation


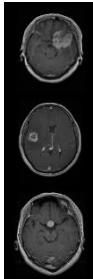
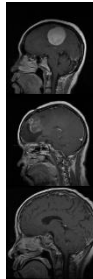






The dataset used in this work came from Cheng et al. [20]. In their work, the MRI scans within the dataset were collected initially from the Nanfang Hospital located in Guangzhou, China dating from 2005 to 2010. Table 2 defines the dataset specification, including 2D slices of T1-weighted CE-MRI images of 708 meningioma, 1426 glioma, and 930 pituitary brain tumor samples.

Table 2. Specification of the MRI brain tumor dataset

Class	Coronal	Axial	Sagittal	Total
Meningioma	232	208	268	708
Glioma	493	494	439	1426
Pituitary	321	291	318	930
Total	1046	993	1025	3064

Table 3 presents samples of each class in the dataset in various views, including axial, coronal, and sagittal from 233 anonymous patients pre-sorted and validated by an expert radiologist, according to the original author of the data source. The images presented have a standard dimension scale of 512x512 with a 49x49mm pixel size. Initially, the following images had a MAT format that required image processing methods. All the images were converted into JPG to provide an accessible image format presented in a 2D array for YOLO. The pixel intensities were also normalized using the min-max approach to eliminate any unwanted inconsistencies with future test samples [21].

Table 3. Brain tumor MRIs from various viewpoints

Class	Coronal	Axial	Sagittal	Patients
Meningioma				82
Glioma				89
Pituitary				62
Total				233

In preparation for training and testing, the entire dataset was divided into a training set and a testing set according to a specific MRI view and class, as shown in **Table 4**. This approach can help determine how well the models can detect from the test data from the 20% of the entire dataset using an adequate number of learnable patterns from the other 80%. Furthermore, with a blind stochastic selection of samples, this work prevented the chances of adding bias choices that may result in a poor or a pre-determined outcome [34].

Table 4. Dataset distribution for training and testing

Class	Train (Ground Truth)	Test	Total
Meningioma	566	142	708
Glioma	1141	285	1426
Pituitary	744	186	930
Total	2451	613	3064

Upon collection, the existing data did not have an appropriate label format for YOLO. Therefore, this work also performed a precise re-labeling approach to generate compatible ground truth images using an appropriate labeling tool [35]. **Fig. 2** presents the input image (a), pre-defined segment (b) from the work of Cheng et al. [20], and the bounding box (c) used in this work.

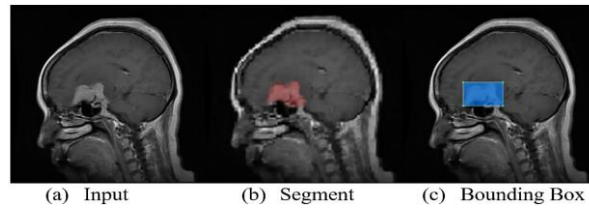


Fig. 2. Comparison of ground truth labels from a segment to a bounding box

3.3 YOLOv4-Tiny Model

This work mainly proposed using a lightweight YOLOv4 rather than the original structure. The tiny version had fewer layers of 29 (starting from 0) compared to the original with more than a hundred. Therefore, it made the training process faster with fewer computations without too much trade in detection performance. This work considered the approach and produced an easily deployable and replicable MRI brain tumor detection system. Furthermore, the compact version had better compatibility with most machines and smaller handheld devices like smartphones.

Fig. 3 presents a visual representation of the YOLOv4-Tiny from the yolov4-tiny.conv.29 [32]. The backbone includes several directly and indirectly connected components, an input layer, eighteen Convolution (Conv) layers with activations, nine routes, three Max-Pooling (MP) layers, and a YOLOv3 detector. The extractor begins with an input layer that takes an $N \times N$ image. The tiny version mainly uses a 416×416 by default. However, slight modifications of various sizes added further investigation on how the input layer impacts the overall model performance. One of the model's primary component is the Conv layers that extract a robust set of learnable features from the input image. The backbone comes with an interchanging 3×3 and 1×1 receptive filters that strides over the $N \times N$ input to produce a new set of feature maps or filters that pass through the network.

The following sequence of Conv layers dynamically changes throughout the network in terms of size, strides, and filters that affect each output. During convolutions, the activation function initializes the mapping of each successive input and outputs that help in the learning process [36]. For this work, the Leaky-ReLU performed the activation in every Conv layer in a non-linear fashion to increase feature size. This activation also prevented the dying ReLU problem [37]. Another component of the backbone are routes that act as shortcuts to improve gradient flow across the entire network [38]. The output of specific Conv layers connects directly and indirectly to other layers that concatenate various layers that generated fine-grained features [39]. Through this approach, the image maintained a rich resolution quality even after several downsamples. As part of the design to perform efficiently, the MP layer performed the downsizing of features into half, preventing a higher computing cost. However, even with the reduction process, the MP layer still preserved the highest output values generated using a consistent 2×2 filter with a stride of $/2$. This approach maintained the feature depth without the depletion of the entire resource during training.

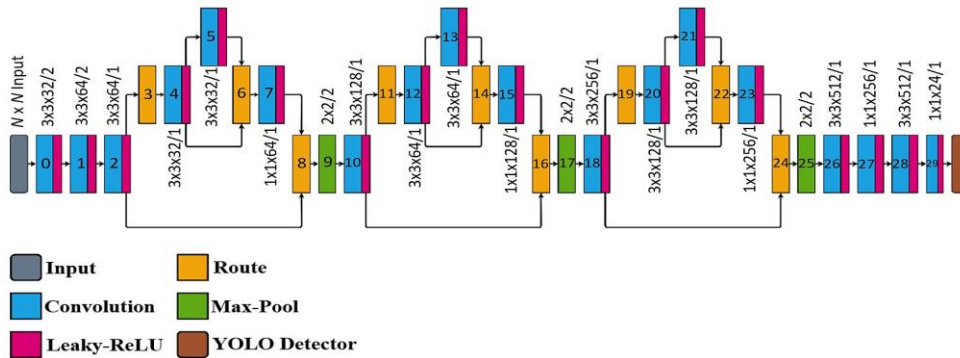


Fig. 3. The YOLOv4-Tiny feature extractor and backbone for the MRI brain tumor detection

3.4 The Detection Unit

In Fig. 4, the YOLOv3 detector served as the detection unit that takes its detection in the form of a regression problem instead of the conventional logistic approach to run at faster speeds. Based on the given structure, the network consists of 24 Conv layers with a filter size of 3x3 for the automated extraction of robust image features. It is also worth mentioning that the architecture’s main inspiration was a DCNN called GoogleNet [40]. However, instead of using Inception blocks like GoogleNet, YOLOv3 uses a 1x1 downsizing filter to reduce the feature size to improve its computational efficiency. With this design, the model propagates an entire image through the network rapidly while calculating highly accurate probabilities [41].

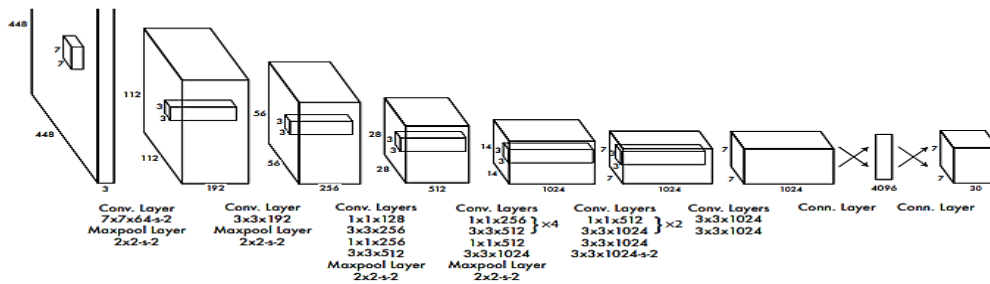


Fig. 4. The YOLOv3 detection unit [41]

It is worth to mention that YOLOv4 still relied on YOLOv3 as a head detection unit. However, YOLOv4 had an improved backbone that employed the Cross-Stage Partial Network (CSPNet), in the form of CSPDarknet53. The CSPNet managed the model’s input by separating it into two sections and passing it singularly to a dense block. Therefore, it reduced the previously large computational requirements to train custom models, that made YOLOv4 work on low-performing devices competently [42]. The neck section now consists of the Spatial Pyramid Pooling (SPP) and Path Aggregation Network (PAN). With SPP, a CNN model like YOLOv4 disregarded fixed image dimensions to improve its flexibility to learn and detect the various scaled images. This profound SPP method consists of arbitrary pooling of regions that produced fixed-length representations to train its detector with a single computation of features on an entire image [43]. Simultaneously, instead of using the previous Feature Pyramid Network (FPN) of YOLOv3, PAN took its place and improved the neural network propagation of YOLOv4 that boosted the flow of information over the entire network. Such an approach improved the propagation of features from the lower to upper network levels, which added further efficiency to the model performance [44].

3.5 The Detection Approach

This section briefly explains the detection process of the YOLO-based model. The process begins with the model interpreting an image using logical $S \times S$ grids and the weighted feature sets to create a probability on an area of cells. If the center of a probable object falls to one of the cells, a preliminary bounding box is produced based on the prediction probability given by the trained model in (1).

$$Pr(Object) = \begin{cases} 0, & \text{has potential objects} \\ 1, & \text{has no potential objects} \end{cases} \quad (1)$$

The model then predicts with the use of K various scaled boxes and extracts a 3D tensor based on (2), where C represents the defined number of classes, four as the t_x, t_y, t_w, t_h bounding box prediction coordinates, and one as the confidence of prediction for each bounding box [31].

$$S * S * (K * (4 + 1 + C)) \quad (2)$$

In Fig. 5, the bounding box prediction based on the width p_w and height p_h had offsets c_x and c_y from the cluster centroid. When the cell offset from the upper left by (c_x, c_y) and the bounding box has values of p_w and p_h , then the prediction corresponds to (3) [41].

$$\begin{aligned} b_x &= \sigma(t_x) + c_x \\ b_y &= \sigma(t_y) + c_y \\ b_w &= p_w e^{t_h} \\ b_h &= p_h e^{t_h} \end{aligned} \quad (3)$$

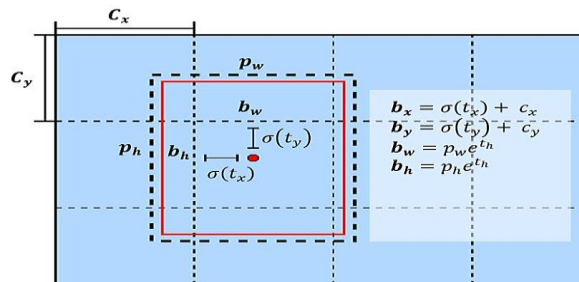


Fig. 5. Bounding box prediction with specifications

Simultaneously, in (4), the Intersection Over Union (IoU) calculates whether the prediction closely matches a ground truth image from the dataset during the creation of bounding boxes. However, (5) defines that when the initially predicted object does not closely resemble the ground truth, the confidence score decreases, resulting in an unsatisfied prediction [28].

Fig. 6 shows a visual sample of how the IoU works in evaluating an object where the bounding box surrounds the ground truth and a possible detection result from the algorithm.

$$IoU_{pred}^{truth} = \frac{area(B^{(truth)} \cap B^{(pred)})}{area(B^{(truth)} \cup B^{(pred)})} \quad (4)$$

$$Confidence(Object) = Pr(Object) * IOU_{pred}^{truth} \quad (5)$$

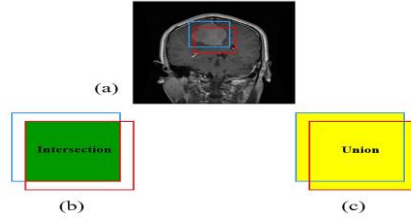


Fig. 6. Ground truth sample for detecting a brain tumor (a), with the use of intersection (b) over union (c).

For each grid cell, C is determined based on $Pr(Class_i|Object)$. Even with the multiple predictions over a specific object, only objects that meet the given threshold will achieve an initial bounding box. Eq. (6) formally presents the following [28].

$$Pr(Class_i|Object) * Pr(Object) * IOU_{pred}^{truth} = Pr(Class_i) * IOU_{pred}^{truth} \quad (6)$$

In most cases, during the initial detection, several potential object predictions can occur. However, with the application of the Non-Maximum Suppression algorithm [32], the detection model managed to retain only the prediction with the highest confidence score and eliminated any redundant boxes efficiently. Fig. 7 illustrates the detection process of the proposed YOLO-based model.

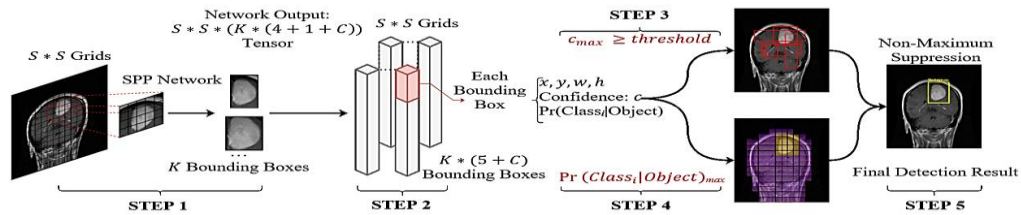


Fig. 7. The YOLOv4 brain tumor detection process

3.6 Transfer Learning, Fine-Tuning, and Model Training

In most DL tasks, using inadequate learning data can produce a weak and inaccurate performance. However, transfer learning paved the way to train models and acquire substantial results without the need for massive data [45]. Hence, this work adopted this technique and used the pre-trained weights from the COCO dataset to improve the model performance to detect several brain tumors. The previously learned COCO features supplied the model with additional image recognition essentials needed for the detection process. Also, to further optimize the pre-trained model, the application of fine-tuning adjusted the resource allocation and prevented the depletion of memory during training and testing [46].

The initial step to fine-tune the model was to replace the default class numbers from 80 to three, in which the three corresponds to the brain tumors, namely, glioma, meningioma, and pituitary, as the default number of 80 corresponds to the previous classes from COCO. With the newly defined class size, every Conv filters must also shift from the default 255 to 24 as formally defined in (7) where C corresponds to the number of classes, five as the YOLO coordinates, and three as the various scaled bounding boxes K .

$$filters = 3 * (5 + C) \quad (7)$$

Subsequently, other hyper-parameters like the batch size, subdivisions, learning rate, momentum, decay, and iterations are tailor-fitted for this work, as shown in [Table 5](#). The fine-tuned model trained with a batch size of 64, a subdivision of 8, and iterations of 6000. The learning rate, momentum, and decay for the training process were based on the current resource and adjusted optimally with values of 0.00261, 0.9, and 0.0005. Furthermore, to provide initial results of the performance during the training phase, the weights serialized automatically every 1000 iterations.

Table 5. Hyper-Parameter configuration to fine-tune the model

Hyper-Parameter	Value
Batch Size	64
Subdivisions	8
Learning Rate	0.000261
Momentum	0.9
Decay	0.0005
Iterations	6000

3.7 Evaluation Metrics

Once the model completed both the training and testing phases, the next step is to perform an appropriate performance measurement with standardized evaluation metrics for object detection to measure overall performance.

This work selected a threshold k of 0.5 to evaluate the Intersection over Union (IoU), Precision (PR), Recall (RE), and the mAP. The following are calculated based on the number of True Positives (TP), False Positives (FP), and False Negative (FN) detections by the model using a test set of 613 MRIs. The TP signifies a correctly detected tumor class with a correct label, while FPs are the detected non-tumors, and FNs are tumors that the model did not detect properly. This work did not include True Negatives as the dataset did not have any negative samples (MRI without lesions or tumor) included. Thus, an F1-score plays a more relevant metric to measure the harmonic mean between FNs and FPs of the unbalanced dataset rather than accuracy [\[47\]](#).

In a global standard, Average Precision (AP) is the metric used to determine the overall detection prowess of object detection models rather than accuracy [\[48\]](#). This metric pertains to the number of correctly and incorrectly classified samples of a specific class instance. Where the $P(k)$ refers to the precision at a specifically given threshold k , and $\Delta r(k)$ as the shift in the Recall (RE). Eq. (8) formally presents the AP [\[49\]](#).

$$AP = \frac{1}{N} \sum_{k=1}^N P(k) \Delta r(k) \quad (8)$$

The mAP calculates the mean of all AP for each category. Using the mAP as the primary key indicator can justify a model that worked best overall to detect brain tumors specifically. Eq. (9) formally presents the mathematical equation for the mAP [\[31\]](#).

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (9)$$

The Intersection over Union (IoU) determines the overlap between two bounding boxes.

Eq. (10) calculates the IoU by having the intersection area divided by the area of union [28].

$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}} \quad (10)$$

In Eq. (11), the Miss Rate (MR) measures the FNs ratio compared to all the detected objects within the given image. In contrast, (12) measures the FPs instead. Both measures indicate better performance with lower values [50].

$$MR = \frac{FN}{FN + TP} \quad (11)$$

$$FPPI = \frac{FP}{FP + TP} \quad (12)$$

In the medical field, the PR, RC, and the F1-Score metrics play a vital role to justify the ratio of correctly predicted positives towards all detections, potential detections, and the harmony between PR and RE, respectively. Similarly, in DL, the given metrics can determine a model's effectiveness and establish its trustworthiness towards a specific task. The following are formally calculated based on (13), (14), and (15) [30].

$$PR = \frac{TP}{TP + FP} \quad (13)$$

$$RC = \frac{TP}{TP + FN} \quad (14)$$

$$F1 - Score = \frac{2 * (PR * RC)}{PR + RC} \quad (15)$$

4. Experimental Results and Discussion

This section presents the overall performance results from the conducted experiments. It is worth to mention that this work trained several models with various input sizes for the proposed work to identify the most probable variant to solve the given problem.

4.1 Average Loss and Convergence

In this work, to evaluate if the model trained adequately with the training data, the loss graphs visualized the convergence of loss towards the local minima. As presented in Fig. 8, each graph's blue lines state that the models did not have much loss or errors committed while validating from the image dataset. This characteristic of the loss line is ideal and indicates that the models validated the images correctly in a precise manner with minimal error and that the model learned gradually without fitting problems. Simultaneously, the red line indicated that the model progressively increased its precision from the training dataset over time. Fig. 8 (b), compared to others, attained the lowest average loss of 0.1953, followed by (a) with 0.2084 and (c) with 0.2284. The following results justified that all models trained well in a stable manner through the span of 6000 iterations. The CIoU loss function [32] worked effectively without showing signs of over or underfitting and delivered a rapid convergence during the

phase of training.

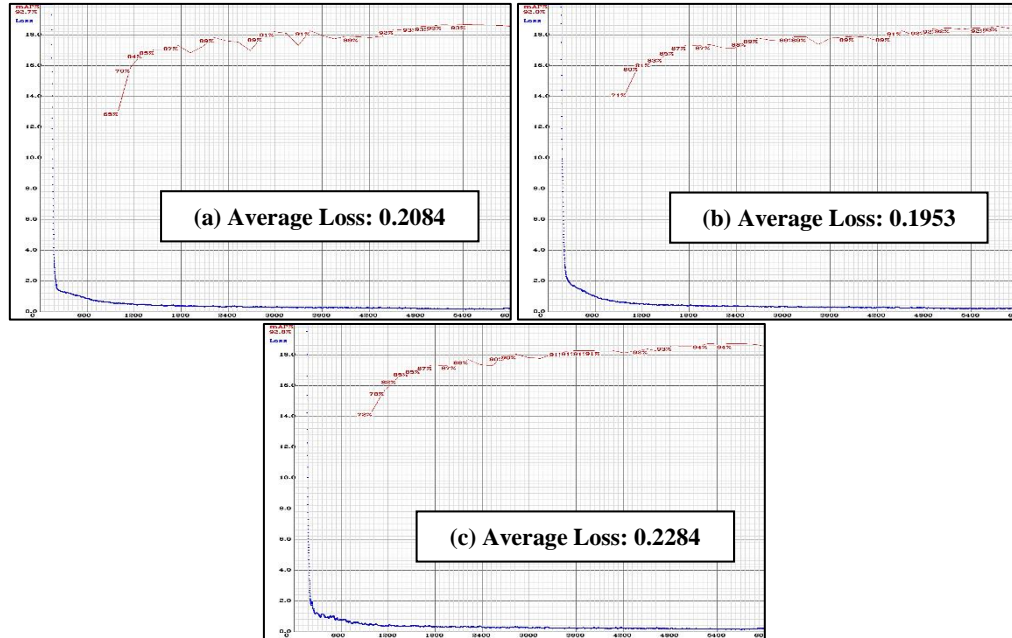


Fig. 8. Left-Right: YOLOv4-Tiny 608x608 (a), 416x416 (b), and 320x320 (c)

4.2 Detection Count and Average Precision

Table 6 presents the AP based on the number of TP and FP detections per class of each model. According to the results, YOLOv4-Tiny(416x416) had the highest AP for detecting meningioma tumors with 98.73%. The 608x608 follows this result with 98.53% and 320x320, with 97.85%. In terms of detecting glioma tumors, the 320x320 outmatched the rest, with 86.13% having the 608x608 at 83.31% and the 416x416 at 84.68%. For the pituitary tumor detection, the 608x608 prevailed as the most precise, 96.23%, followed by 416x416 with 96.01%, and 320x320 with 94.47%.

Upon evaluation, all three models had shown different performances in detecting various types of tumors, making it challenging to determine which model is the most efficient. Therefore, with additional metrics, an additional evaluation set is performed to identify the most formidable brain tumor detection model.

Table 6. The number of detections and average precision of each model per class

Model	Meningioma-0			Glioma-1			Pituitary-2		
	TP	FP	AP	TP	FP	AP	TP	FP	AP
YOLOv4-Tiny (608x608)	139	8	98.53%	220	57	83.31%	174	9	96.23%
YOLOv4-Tiny (416x416)	138	6	98.73%	231	39	84.68%	174	13	96.01%
YOLOv4-Tiny (320x320)	137	9	97.85%	234	42	86.13%	168	12	94.47%

4.3 Overall Performance

Table 7 presents the effect of input dimensions on the overall performance. The calculated results from the test dataset of 613 images generated the highest mAP of 93.14% from the 416x416 variant. The 320x320 then followed with 92.82%. Surprisingly, the 608x608 had no improvements even with the increased dimension and had the lowest mAP of 92.69%. However, it had a slight increase in IoU having 71.77%, in which 320x320 only had 71.45%. In terms of the overall performance, 416x416 attained the highest mAP, IoU of 72.51%, RE of 88.58%, and an F1-Score of 89.45%. The least MR belonged to 608x608, with 1.11% leaving the rest with 2.16 for 416x416 and 1.82% for 320x320. The 608x608 had the highest FPPI of 12.19%, followed by 320x320 with 10.47%, and 416x416 with only 9.65%.

Table 7. Comparison of the overall detection performance with various input dimensions

Model	mAP	IoU	PR	RC	F1	MR	FPPI
YOLOv4-Tiny (608x608)	92.69	71.77	87.80	86.94	87.37	1.11	12.19
YOLOv4-Tiny (416x416)	93.14	72.51	90.34	88.58	89.45	2.16	9.65
YOLOv4-Tiny (320x320)	92.82	71.45	89.53	87.92	88.72	1.82	10.47

4.4 Detections

Fig. 9 presents the following detection made by each model from randomly selected test data. Surprisingly, the 416x416 with the highest mAP failed to detect the tumor in samples (a) and (b). Also, the 320x320 even outperformed it and detected the tumor in (a). Overall, only the 608x608 managed to attain complete detections from the given set of randomly selected samples in the given figure. However, this does not entirely summarize the overall performance of each model. This section simply demonstrates how YOLO can detect various small to large brain tumors from the test set.

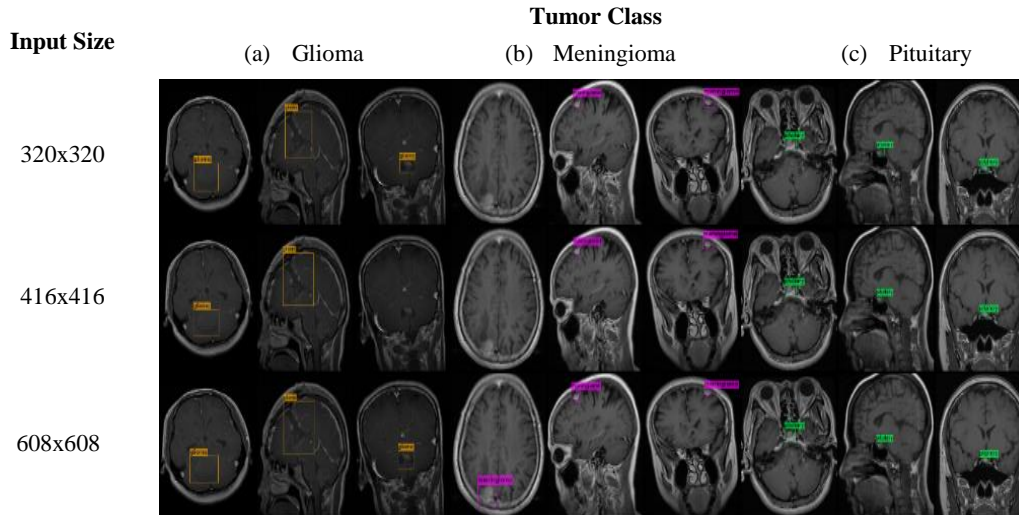


Fig. 9. Detections from a set of randomly selected test data

4.5 Size Allocation and Computational Cost of YOLOv4-Tiny

Currently, object detection seeks further improvements in precision and less storage consumption that implicates the concept of ubiquitous computing with DL. A lesser costing model allows users to use robust DL applications across multiple computing platforms [51].

Fig. 10 presents the computational cost (GFLOPS), detection speed (s), and size allocation (MB) of the following brain tumor detection models trained with various YOLO-tiny versions. The evaluated results show that YOLOv4-Tiny had the smallest allocation size requirement than the older versions with 22.989MB. Due to the slightly larger backbone size, YOLOv4-Tiny had a slightly higher GLOPS of 6.79 while the v2 and v3 had 5.346 and 5.451. However, compared to others with higher GLOPS, v2 still consumed the most with 43.124MB, followed by v3 with 33.91MB. Overall, v4 and v2 attained excellent performance in terms of detection speed of only 3 seconds, unlike v3, that required 5s.

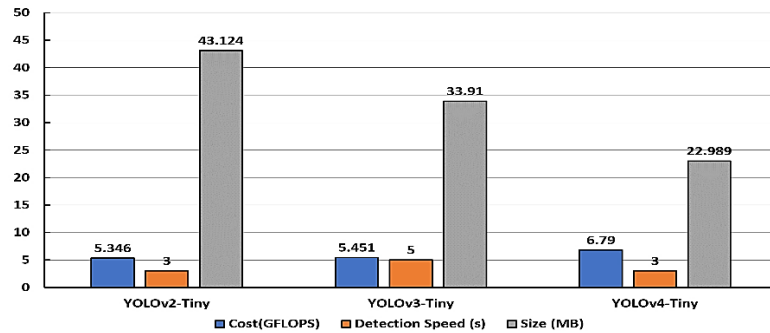


Fig. 10. Comparison of computing cost, detection speed, and disk consumption of the YOLOv4-Tiny with previous versions

4.6 Comparison of Performance with Other Object Detection Models

This section compares the results of this work from some studies that performed brain tumor detection from MRIs. However, this work does not entirely provide a direct comparison due to variations in methods. The primary purpose is to present how the proposed method improved brain tumor detection status in object detection. It is worth to mention that only a handful of studies exist that have used bounding boxes to detect brain tumors in MRIs. Classification and segmentation methods are the usual studies conducted. However, object detection already had plenty of results in other medical fields, indicating its effectiveness [27, 30]. Object detection can present a new perspective to other researchers into diving more profoundly with object detection for brain tumor detection in MRIs through this work.

Table 8 presents a few of the only studies that employed object detection methods with bounding boxes to locate brain tumors in MRIs. In comparison, this work attained the highest mAP of 93.14% with the YOLOv4-Tiny model while the rest acquired a lesser rate. These results show that the YOLOv4-Tiny in this work surpassed the other existing brain tumor detection methods.

Table 8. Comparison with existing object detection approach for brain tumors from MRIs

Method	Dataset Size	Classes	Dataset	Results
YOLOv4-Tiny	3064	3	[20]	93.14%
YOLOv3-Tiny	3064	3	[20]	91.98%
YOLOv2-Tiny	3064	3	[20]	79.90%
Faster R-CNN [52]	3064	3	[20]	91.66%
Faster R-CNN [26]	2406	3	[20]	77.60%

5. Conclusions

This work presented the efficiency of employing transfer learning and fine-tuning on a recent YOLOv4-Tiny model to perform the detection of glioma, meningioma, and pituitary brain tumors in MRIs. The dataset used had 3064 T1-weighted CE-MRIs containing several axial, coronal, and sagittal views for each given class. Several data pre-processing methods included the min-max normalization of pixel contrast, file conversion, and generation of training labels for the tumor coordinates. The 29-layer YOLOv4-Tiny served as a backbone to extract a robust set of features to enable the detector to learn valuable patterns. The model also used the pre-learned weights from COCO through transfer learning and the newly initialized feature sets generated by the extractor from the MRI dataset. With the fine-tuned network, the model trained end-to-end optimally without issues. Both mentioned techniques provided additional leverage for a quick and robust training process. Upon evaluation, the compact YOLOv4-Tiny with a 416x416 input dimension precisely detected the distinct brain tumors with a 93.14% mAP set at a 0.5 threshold. With such performance yielded in this work, the YOLOv4-Tiny (416x416) only consumed a 22.989MB of disk space with 6.79GFLOPS and a 3s detection speed.

This work concludes that object detection models pre-trained and fine-tuned like the YOLOv4-Tiny can efficiently diagnose brain tumors from MRIs. Compared to classification methods, this work localized brain tumors from the MRIs and classified it specifically with less comprehension required. Unlike segmentation methods, the proposed work can run on most platforms due to the relatively small space requirement and low computational cost. Moreover, compared to existing works that employed bounding box detection methods for meningioma, glioma, and pituitary brain tumors, this work prevailed as the most precise.

However, this work still has certain caveats in terms of having bounding boxes to detect tumors. The use of bounding boxes still limits the precise selection of tumors compared to a segmentation approach. With that said, YOLO can capture excess areas due to the complex morphology of tumors compared to the limited shape of the bounding box. Training YOLO-based models and other similar models also require lengthy and tedious dataset labeling compared to a classification approach. YOLO can also become sensitive to the lack of data, requiring additional MRI images for future works. Nonetheless, the trade-offs are relatively minimal compared to the emitted solution as YOLO can still scale and evolve through continuous research to resolve the mentioned concerns.

References

- [1] J. McFaline-Figueroa and E. Lee, "Brain Tumors," *The American Journal of Medicine*, vol. 131, no. 8, pp. 874-882, 2018. [Article \(CrossRef Link\)](#)
- [2] C. Gladson, R. Prayson, and W. Liu, "The Pathobiology of Glioma Tumors," *Annual Review of Pathology: Mechanisms of Disease*, vol. 5, no. 1, pp. 33-50, 2010. [Article \(CrossRef Link\)](#)
- [3] P. Wen and S. Kesari, "Malignant Gliomas in Adults," *New England Journal of Medicine*, vol. 359, no. 5, pp. 492-507, 2008. [Article \(CrossRef Link\)](#)
- [4] L. DeAngelis, "Brain Tumors," *New England Journal of Medicine*, vol. 344, no. 2, pp. 114-123, 2001. [Article \(CrossRef Link\)](#)
- [5] J. Baehring, W. Bi, S. Bannykh, J. Piepmeyer, and R. Fulbright, "Diffusion MRI in the early diagnosis of malignant glioma," *Journal of Neuro-Oncology*, vol. 82, no. 2, pp. 221-225, 2006. [Article \(CrossRef Link\)](#)

- [6] A. Gupta and T. Dwivedi, "A Simplified Overview of World Health Organization Classification Update of Central Nervous System Tumors 2016," *Journal of Neurosciences in Rural Practice*, vol. 08, no. 04, pp. 629-641, 2017. [Article \(CrossRef Link\)](#)
- [7] I. Shimon and S. Melmed, "Pituitary Tumor Pathogenesis", *The Journal of Clinical Endocrinology & Metabolism*, vol. 82, no. 6, pp. 1675-1681, 1997. [Article \(CrossRef Link\)](#)
- [8] L. Rogers, I. Barani, M. Chamberlain, T. Kaley, M. McDermott, J. Raizer, D. Schiff, D. Weber, P. Wen, and M. Vogelbaum, "Meningiomas: knowledge base, treatment outcomes, and uncertainties. A RANO review", *Journal of Neurosurgery*, vol. 122, no. 1, pp. 4-23, 2015. [Article \(CrossRef Link\)](#)
- [9] R. Buerki, C. Horbinski, T. Kruser, P. Horowitz, C. James, and R. Lukas, "An overview of meningiomas," *Future Oncology*, vol. 14, no. 21, pp. 2161-2177, 2018. [Article \(CrossRef Link\)](#)
- [10] B. Holleczeck, D. Zampella, S. Urbschat, F. Sahn, A. von Deimling, J. Oertel, and R. Ketter, "Incidence, mortality, and outcome of meningiomas: A population-based study from Germany," *Cancer Epidemiology*, vol. 62, p. 101562, 2019. [Article \(CrossRef Link\)](#)
- [11] I. Whittle, C. Smith, P. Navoo, and D. Collie, "Meningiomas," *The Lancet*, vol. 363, no. 9420, pp. 1535-1543, 2004. [Article \(CrossRef Link\)](#)
- [12] L. Zhao and K. Jia, "Multiscale CNNs for brain tumor segmentation and diagnosis," *Comput. Math. Methods Med.*, vol. 2016, Feb. 2016, Art. no. 8356294. [Article \(CrossRef Link\)](#)
- [13] D. Saloner, A. Uzelac, S. Hetts, A. Martin, and W. Dillon, "Modern meningioma imaging techniques," *Journal of Neuro-Oncology*, vol. 99, no. 3, pp. 333-340, 2010. [Article \(CrossRef Link\)](#)
- [14] J. Watts, G. Box, A. Galvin, P. Brochie, N. Trost, and T. Sutherland, "Magnetic resonance imaging of meningiomas: a pictorial review," *Insights into Imaging*, vol. 5, no. 1, pp. 113-122, 2014. [Article \(CrossRef Link\)](#)
- [15] G. Litjens, T. Kooi, B. E. Bejnordi, A. Setio, F. Ciompi, M. Ghafoorian, J. Laak, B. Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, December 2017. [Article \(CrossRef Link\)](#)
- [16] 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. [Article \(CrossRef Link\)](#)
- [17] J. Y. Chiao, K. Y. Chen, K. Y. K. Liao, P. H. Hsieh, G. Zhang, and T. C. Huang, "Detection and classification the breast tumors using mask R-CNN on sonograms," *Medicine*, vol. 98, no. 19, May 2019, Art. no. e15200. [Article \(CrossRef Link\)](#)
- [18] Z. Zhao, P. Zheng, S. Xu, and X. Wu, "Object Detection With Deep Learning: A Review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212-3232, Nov. 2019. [Article \(CrossRef Link\)](#)
- [19] O. Alsing, "Mobile object detection using tensorflow lite and transfer learning," Degree Project, KTH Royal Institute of Technology School of Electrical Engineering and Computer Science, Stockholm, Sweden, 2018. [Article \(CrossRef Link\)](#)
- [20] J. Cheng, W. Yang, M. Huang, W. Huang, J. Jiang, Y. Zhou, R. Yang, J. Zhao, Y. Feng, Q. Feng, and W. Chen, "Retrieval of Brain Tumors by Adaptive Spatial Pooling and Fisher Vector Representation," *PloS one*, 2016. [Article \(CrossRef Link\)](#)
- [21] 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, July 2019. [Article \(CrossRef Link\)](#)
- [22] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput. Biol. Med.*, vol. 111, Aug. 2019, Art. no. 103345. [Article \(CrossRef Link\)](#)
- [23] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Syst., Signal Process.*, vol. 39, pp. 757-775, Sep. 2019. [Article \(CrossRef Link\)](#)
- [24] H. H. Sultan, N. M. Salem, and W. Al Atabany, "Multi-classification of brain tumor images using deep neural network," *IEEE Access*, vol. 7, pp. 69215-69225, 2019. [Article \(CrossRef Link\)](#)

- [25] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, M. Imran, and M. Shoaib, "A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor," *IEEE Access*, vol. 8, pp. 55135-55144, 2020. [Article \(CrossRef Link\)](#)
- [26] Y. Bhanothu, A. Kamalakannan, and G. Rajamanickam, "Detection and Classification of Brain Tumor in MRI Images using Deep Convolutional Network," in *Proc. of 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2020, pp. 248-252. [Article \(CrossRef Link\)](#)
- [27] 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. [Article \(CrossRef Link\)](#)
- [28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, 2016. [Article \(CrossRef Link\)](#)
- [29] P. Viola and M. Jones, "Robust Real-Time Face Detection," *Int'l J. Computer Vision*, vol. 57, no. 2, pp. 137-154, May 2004. [Article \(CrossRef Link\)](#)
- [30] H. Ünver and E. Ayan, "Skin Lesion Segmentation in Dermoscopic Images with Combination of YOLO and GrabCut Algorithm," *Diagnostics*, vol. 9, no. 3, p. 72, 2019. [Article \(CrossRef Link\)](#)
- [31] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in *Proc. of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 6517-6525. [Article \(CrossRef Link\)](#)
- [32] A. Bochkovskiy, C.Y. Wang, and H.Y. Mark Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv:2004.10934*, 2004. [Article \(CrossRef Link\)](#)
- [33] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," *Proc. Eur. Conf. Comput. Vis.*, pp. 740-755, 2014. [Article \(CrossRef Link\)](#)
- [34] H. Jiang and O. Nachum, "Identifying and correcting label bias in machine learning," in *Proc. of the Twenty Third International Conference on Artificial Intelligence and Statistics (PMLR)*, vol. 108, pp. 702-712, 2020. [Article \(CrossRef Link\)](#)
- [35] Tzutalin, "LabelImg," GitHub, 2015. [Online]. Available: <https://github.com/tzutalin/labelImg>
- [36] R. Girshick, "Fast R-CNN," in *Proc. of 2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, pp. 1440-1448, 2015. [Article \(CrossRef Link\)](#)
- [37] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," in *Proc. of 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1026-1034, 2015. [Article \(CrossRef Link\)](#)
- [38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016. [Article \(CrossRef Link\)](#)
- [39] C. Dewi, R. Chen, and S. Tai, "Evaluation of Robust Spatial Pyramid Pooling Based on Convolutional Neural Network for Traffic Sign Recognition System," *Electronics*, vol. 9, no. 6, p. 889, 2020. [Article \(CrossRef Link\)](#)
- [40] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-9, 2015. [Article \(CrossRef Link\)](#)
- [41] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv:1804.02767*, 2018. [Article \(CrossRef Link\)](#)
- [42] C. Wang, H. Mark Liao, Y. Wu, P. Chen, J. Hsieh, and I. Yeh, "CSPNet: A New Backbone that can Enhance Learning Capability of CNN," in *Proc. of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1571-1580, 2020. [Article \(CrossRef Link\)](#)
- [43] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904-1916, Sep. 2015. [Article \(CrossRef Link\)](#)

- [44] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path Aggregation Network for Instance Segmentation," in *Proc. of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8759-8768, 2018. [Article \(CrossRef Link\)](#)
- [45] H. C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Uao, D. Mollura, and R. M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics, and Transfer Learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285-1298, May 2016. [Article \(CrossRef Link\)](#)
- [46] S. Sharma and R. Mehra, "Effect of layer-wise fine-tuning in magnification-dependent classification of breast cancer histopathological image," *The Visual Computer*, vol. 36, pp. 1755-1769, 2020. [Article \(CrossRef Link\)](#)
- [47] P. Henderson and V. Ferrari, "End-to-end training of object class detectors for mean average precision," in *Proc. of The Asian Conference on Computer Vision (ACCV)*, vol. 10115, 2017. [Article \(CrossRef Link\)](#)
- [48] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," in *Proc. of 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 237-242, 2020. [Article \(CrossRef Link\)](#)
- [49] R. Huang, J. Pedoem, and C. Chen, "YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers," in *Proc. of 2018 IEEE International Conference on Big Data (Big Data)*, pp. 2503-2510, 2018. [Article \(CrossRef Link\)](#)
- [50] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: A benchmark," in *Proc. of 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 304-311, 2009. [Article \(CrossRef Link\)](#)
- [51] C. Chen, P. Zhang, H. Zhang, J. Dai, Y. Yi, H. Zhang, and Y. Zhang, "Deep Learning on Computational-Resource-Limited Platforms: A Survey," *Mobile Information Systems*, vol. 2020, pp. 1-19, 2020. [Article \(CrossRef Link\)](#)
- [52] E. Avşar and K. Salçin, "Detection and classification of brain tumours from MRI images using faster R-CNN," *Tehnički glasnik*, vol. 13, no. 4, pp. 337-342, 2019. [Article \(CrossRef Link\)](#)



Francis Jesmar P. Montalbo received both of his BS and MS degrees in Information Technology from Batangas State University in 2012 and 2018, respectively. He is currently a candidate for the degree of Doctor in Information Technology at the Technological Institute of the Philippines and working as a full-time faculty researcher. Part of his career is being a current member of IEEE and the IEEE-Computational Intelligence Society (CIS) under the Philippine section. He also contributed several works presented and published in various Scopus indexed journals and conferences that correspond to his research interest in applied deep learning and computer vision.