

A Deep Learning Model to Classify and Detect Brain Abnormalities in Portable Microwave Based Imaging System

Amran Hossain (✉ amran.hossain@duet.ac.bd)

Dhaka University of Engineering & Technology

Mohammad Tariqul Islam

Universiti Kebangsaan Malaysia

Ali F. Almutairi

Kuwait University

Research Article

Keywords: YOLOv5 model, microwave imaging, tumor classification, tumor detection, brain abnormalities, augmentation technique

Posted Date: August 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-848377/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on April 15th, 2022. See the published version at <https://doi.org/10.1038/s41598-022-10309-6>.

A Deep Learning Model to Classify and Detect Brain Abnormalities in Portable Microwave Based Imaging System

Amran Hossain^{1,*}, Mohammad Tariqul Islam^{2,*}, Ali F. Almutairi³

¹ Department of Computer Science and Engineering, Dhaka University of Engineering & Technology, Gazipur, Gazipur 1707, Bangladesh

² Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia

³ Electrical Engineering Department, Kuwait University, Kuwait City 13060, Kuwait

*Correspondence: amran.hossain@duet.ac.bd; tariqul@ukm.edu.my; ali.almut@ku.edu.kw

ABSTRACT

Automated classification and detection of brain abnormalities like tumors from microwave head images is essential for investigating and monitoring disease progression. This paper presents the automatic classification and detection of human brain abnormalities through the deep learning-based YOLOv5 model in microwave head images. The YOLOv5 is a faster object detection model, which has a less computational architecture with high accuracy. At the beginning, the backscattered signals are collected from the implemented 3D wideband nine antennas array-based microwave head imaging (MWHI) system, where one antenna operates as a transmitter and the remaining eight antennas operate as receivers. In this research, fabricated tissue-mimicking head phantom with a benign and malignant tumor as brain abnormalities, which is utilized in MWHI system. Afterwards, the M-DMAS (modified-delay-multiply-and-sum) imaging algorithm is applied on the post-processed scattering parameters to reconstruct head regions image with 640×640 pixels. Three hundred sample images are collected, including benign and malignant tumors from various locations in head regions by the MWHI system. Later, the images are preprocessed and augmented to create a final image dataset containing 3600 images, and then used for training, validation, and testing the YOLOv5 model. Subsequently, 80% of images are utilized for training, and 20% are used for testing the model. Then from the 80% training dataset, 20% is utilized for validation to avoid overfitting. The brain abnormalities classification and detection performances with various datasets are investigated by the YOLOv5s, YOLOv5m, and YOLOv5l models of YOLOv5. It is investigated that the YOLOv5l model showed the best result for abnormalities classification and detection compared to other models. However, the achieved training accuracy, validation loss, precision, recall, F1-score, training and validation classification loss, and mean average precision (mAP) are 99.84%, 9.38%, 93.20%, 94.80%, 94.01%, 0.004, 0.0133, and 96.20% respectively for the YOLOv5l model, which ensures the better classification and detection accuracy of the model. Finally, a testing dataset with different scenarios is evaluated through the three versions of the YOLOv5 model, and conclude that brain abnormalities classification and detection with location are successfully classified and detected. Thus, the deep model is applicable in the portable MWHI system.

Keywords: YOLOv5 model, microwave imaging, tumor classification, tumor detection, brain abnormalities, augmentation technique

1. INTRODUCTION

Worldwide, brain abnormalities such as tumors are the prime reasons of fatality and incapacity since it damages the primary cell of the human brain. A brain tumor is the enhancement of anomalous tissues that creates inside the human head. Due to the brain abnormality effectiveness, the possibility of brain cancer in the head is increasing day by day, and it is the 9th dominant reason of death for women, men, and children [1]. Brain abnormalities might be fatal, crucially affecting the quality of longevity and changing the whole thing for patients and their families. Brain tumor as an abnormality is mainly classified into two categories: benign and malignant tumor [2]. The benign tumors are non-cancerous cells and have a homogeneous structure with a regular shape. In contrast, malignant tumors are cancerous cells and have a heterogeneous structure with an irregular shape [1]. In addition, benign tumors have a slower growth rate, and malignant tumors can grow uncontrollably. The percentage of death rate increases due to the invasive features of the tumors. But, early detection of brain tumors might increase the survival rate of the human. At present, different modern imaging modalities: PET (positron emission tomography), X-ray screening, MRI (magnetic resonance imaging), CT (computed tomography) scans, and ultrasound screening are used to identify brain tumors in medical applications [2-4]. The key disadvantages of the conventional medical imaging modalities are increasing cancerous risk due to high dose radiation, lower susceptibility, high ionizing with brain cells, costly, and danger for the old patient and pregnant women [2, 5-10]. In the last decades, microwave imaging (EMI) research has been growing in medical applications due to its abundant properties: non-ionizing radiation, inexpensive, non-invasive, and very safe for the human body compared to conventional medical imaging modalities [11-17]. An EMI system consists of an antenna array, where one antenna produces the microwave waves and send them to the objects, and the remaining antennas receive the scattering waves. Thereafter, applying image reconstruction algorithms on collected scattering signals and then produce reconstructed microwave images of the object of interest. In an microwave head imaging (MWHI) system, a wideband antenna design is indispensable, which will work between 1 to 4 GHz operating frequency band with unidirectional radiation characteristics, high gain, and high efficiency [11, 18-24]. Various MWHI systems have been implemented over time by using various types of two-dimensional(2D) [15, 18, 23-28] and three-dimensional(3D) antenna array to identify the brain tumor [14, 19, 22, 29-32].

Recently, the deep neural networks (DNNs) technique has been applied in medical imaging to identify brain abnormalities due to the benefit of making decisions with significantly less involvement from human interaction. The DNN filters the imaging data through the cascade of multiple layers and provides a more accurate result as an output, and it is faster than the traditional imaging system [33-36]. At present, researchers are applying the DNN technique in microwave imaging applications to classify the target object and imaging purposes [36-42]. An autoencoder based deep learning method was investigated to reconstruct the microwave image of the different shaped objects [40]. The technique used a 24×24 antenna array system, which generated EM

signals and propagated through the network to reconstruct the image. But no object detection mechanism was presented in the article. A convolutional neural network(CNN) based U-Net segmentation framework was used to masking the tumor object in microwave (MW) imaging [41]. The framework was investigated on only the simulated data but did not implement an experimental framework and detection algorithm. An automatic DNN based brain abnormalities classification and detection framework were presented in [43]. The approach classified only the tumor and non-tumor images and then masking the tumor regions in the image. But the process is failed to classify the benign or malignant tumor in the images. A handcrafted and deep learning features based brain tumor detection approach was presented in [44]. The approach used grab cut technique to segment the tumor regions, and then hand crafted and deep features were used to diminish the noise level in the image, but the abnormalities classification and specific tumor detection scenario were not presented in the paper. A fully CNN (F-CNN) method was applied to the brain tumor image dataset to detect the brain abnormalities [45]. The technique was used U-Net segmentation scheme with an adam optimizer algorithm to segment the tumor region in the images. From this approach, the tumor could be identified, but the appropriate location of the tumor is difficult to locate, and classification of abnormality is not possible because of the shortage of impartial predictors. The residual DNN technique was applied to identify and classify the tumor object in the brain [34]. The method used data augmentation and distillation procedures to mitigate the unwanted signal in the image, but tumor identification setup with exploratory analysis was not accomplished in the article. For MW imaging purposes, a DNN network based super-resolution algorithm (SRA) was proposed in [46]. The network used a 33×33 image dataset to train the model, and it produced two-dimensional images. But it was seen that the produced images were noisy, blurry, and the accuracy of abnormality classification was very low due to the small training dataset. Thus, detecting the precise position of the target tumor in the image was complicated, and the approach failed to classify the benign and malignant tumor in the images. However, especially stated deep learning-based approaches could not classify the benign and malignant tumor as brain abnormalities with their appropriate location in the images. Therefore, a deep learning-based framework with an object detection algorithm might be pioneering research findings in MW brain imaging application to classify the brain abnormalities as benign and malignant tumor and detection with appropriate localization.

It is observed that the stated MWHI systems used numerous traditional microwave imaging algorithms to identify the target tumor from the reconstructed images. These algorithms are not able to classify the tumors as brain abnormalities, and it is problematic to detect the precise position of the tumor because of the noisy, irregular pattern, and low-quality resolution. Also, classification as benign and malignant tumor and detection with their location of the tumor from the reconstructed images is more challenging to the non-technician or physician. Consequently, it is necessary to classify the tumor as benign or malignant and manually indicate the locations of the tumor in the images by the professional physician or technician, which is vigorous problem of the traditional MWHI system. Therefore, the inspiration of the research to mitigate the issue in the stated MWHI scheme relies on a DNN architecture that is able to automatically classify the benign and malignant tumor as brain abnormalities, then identify and localize the target tumor with a bounding box from the reconstructed EM head region images.

This research presents a deep learning-based YOLOv5 object detection model that can classify the brain abnormalities (i.e., tumor as benign or malignant) and detect with their position in the reconstructed brain images. The key contributions of the research work might be specified below:

- i) Until now, to the best knowledge of the author's, it is the first research work where a deep learning-based object detection method is employed on MW brain images to classify the brain tumor as benign and malignant, and detection with their location automatically.
- ii) We fabricated tissue-mimicking phantom with a benign and malignant tumor as brain abnormalities for experimental purposes.
- iii) Implemented the deep learning-based YOLOv5 object detection model with its three versions: YOLOv5s, YOLOv5m, and YOLOv5l, which can automatically classify the benign and malignant tumor and detect their position with a bounding box in the EM brain images.
- iv) The desired results of the research work are investigated and verified by the three versions of the YOLOv5 model and assessed tumor classification and detection performances by applying different MW brain image testing datasets.

The YOLOv5 object detection model is based on Bottleneck Cross Stage Partial Darknet (BCSPDarknet) backbone framework that is applied to classify the brain abnormalities (i.e., tumors as benign and malignant) and detect the target tumor with a bounding box in the MW based brain images. The YOLOv5 model with its three versions is implemented, executed by the Python high-level language with a powerful PyTorch machine learning library. A three dimensional (3D) wideband nine antenna array-based portable MWHI system with a fabricated head phantom model is developed for imaging purposes. The M-DMAS (modified-delay-multiply -and-sum) algorithm [19] is applied for MW imaging purposes. The generated image dimension is 640×640 pixels. Three hundred image samples from different cases are collected from the implemented MWHI system. After that, image pre-processing and augmentation techniques are applied to generate the final image dataset. Furthermore, the final image dataset is applied for training, validation, and testing the models. The final dataset consists of 3600 images together with original samples with (i) single benign and malignant tumor, (ii) double benign tumors, (iii) double malignant tumors, (iv) single benign and single malignant tumor. Afterwards, 80% of images are used to train the model, and 20% are used for testing. In further, from the 80% training dataset, 20% are utilized to validate the models. The brain tumor classification and detection performances with various datasets are investigated by the YOLOv5s, YOLOv5m and YOLOv5l. In addition, training vs validation loss, training classification loss vs validation classification loss, and training vs validation accuracy is illustrated. Moreover, the precision, recall, F1-score, and mean average precision (mAP) is also demonstrated. It is shown that the YOLOv5l model demonstrated better outcomes for each parametric than the other two models. The training accuracy, validation loss, precision, recall, F1 score, train classification loss, validation classification loss and mAP for YOLOv5l model are 99.84%, 9.38%, 93.20%, 94.80%, 94.00%, 0.004, 0.0133 and 96.20% respectively. Lower validation loss, higher accuracy, precision, recall, F1_score, and lower train and validation classification losses confirm better detection and classification performances. Also, in several cases, tumor detection and classification outcomes

are evaluated by YOLOv5s, YOLOv5m, and YOLOv5l. Therefore, it is seen from experimental outcomes, the tumor classification and detection with location are effectively classified and detected.

The remaining part of the article is structured as follows: Experimental investigation of the MWHI system is illustrated in section 2. A comprehensive study of the deep learning-based YOLOv5 object detection model with its background is discussed in Section 3. Overall research methodology is presented in Section 4. A detailed explanation of the training experiments is discussed in Section 5. Afterwards, the results and discussions are explained in Section 6, and last of all, conclude the paper in Section 7.

2. EXPERIMENTAL INVESTIGATION OF MWHI SYSTEM

In MWHI system, a wideband antenna is mandatory that must have the capability to operate a frequency band between the range of 1 GHz to 4 GHz with unidirectional radiation features [18, 19, 25, 26]. A 3D antenna is designed on cost-effective Rogers RT5880 substrate material having 1.575 mm thickness (T_h), 2.20 dielectric constant, and 0.0009 loss tangent (δ). The improved dimension of the antenna is $60 \times 25 \times 16.575 \text{ mm}^3$. The power network analyzer (PNA) is used to measure the $|S_{11}|$ (S-parameters) of the antenna. Figure 1(a) illustrates the measured and simulated $|S_{11}|$ parameters of the 3D antenna. It can be observed from Figure 1(a), the measured operating frequency band of the 3D antenna is 1.31 GHz to 3.65 GHz, while the simulated operating frequency band is 1.21 GHz to 3.67 GHz. It is noticeable that good agreement has been found in both measured and simulated outcomes except slightly shifted the resonances to the lower frequencies due to the little bit of fabrication error. Thereafter, the antenna is tested with a Hugo

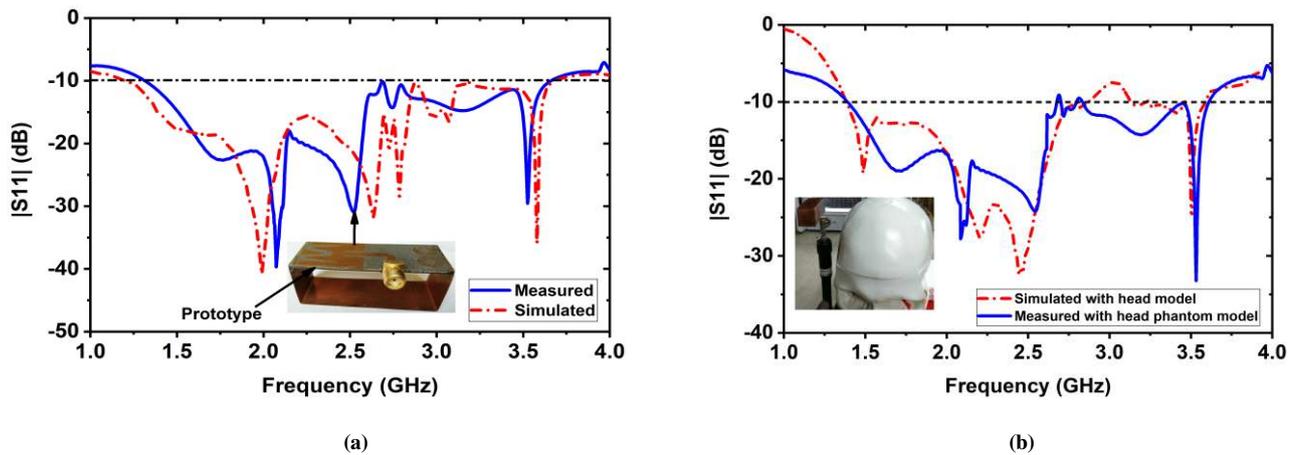


Figure 1: Measured and simulated outcomes of the 3D antenna: (a) $|S_{11}|$ parameters in open space, (b) $|S_{11}|$ parameters with the head phantom model.

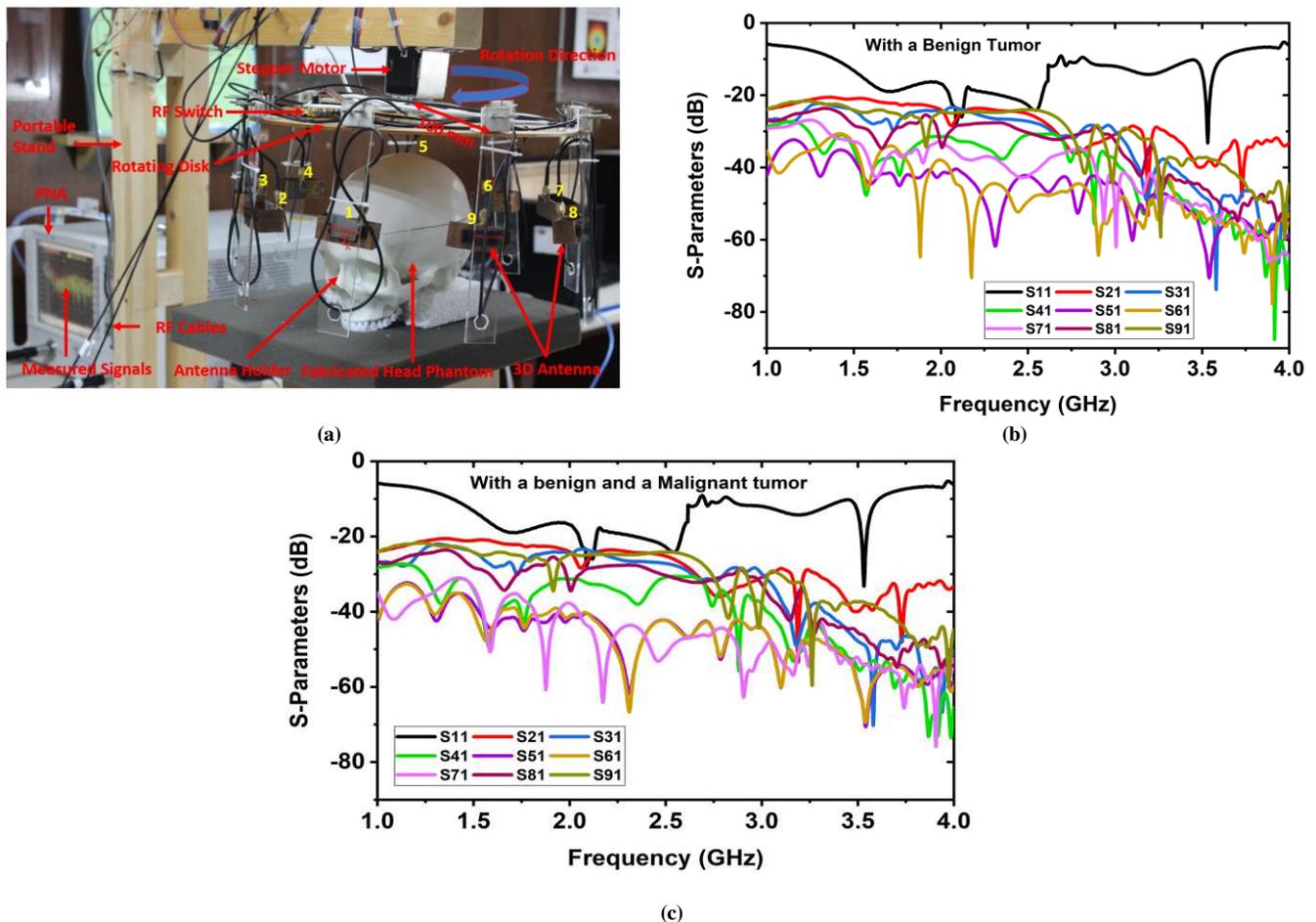


Figure 2: (a) Experimental MWHI system[19] (b) S-parameters with a benign tumor (c) S-parameters with a benign and a malignant tumor.

head model by the simulation and then measured with a fabricated tissue-mimicking head phantom to assess the performance steadiness of the fabricated antenna. Figure 1(b) depicts the measured and simulated $|S_{11}|$ parameters curve through the head phantom. It can be observed from Figure 1(b), the $|S_{11}|$ of the antenna in head proximity keeps a good contract between the measured and simulated outcomes. In this research, our fabricated nine-3D wideband antenna array is used in the developed MWHI framework [19] for imaging purposes, which is shown in Figure 2(a). After that, the system's performance is validated by our fabricated tissue-mimicking head phantom model including benign and malignant tumor as brain abnormalities. The system has a portable stand with a circular-shaped rotating disk (0° to 360°) and a stepper motor. All nine antennas are attached with the separate nine antenna holders, and each antenna is set 40° apart from each other. The distance of each antenna from the center of the disk is 100 mm. The mounted array of the antenna is rotated through the stepper motor from 0° to 360° around the fabricated tissue-mimicking head phantom. The brain tissues of the head phantom, the benign tumor, and the malignant tumor are fabricated according to the presented formula in [47]. The dielectric properties of the benign and the malignant tumor are considered according to the presented value in [48, 49]. The benign tumor is fabricated as almost circular with a regular shape, whereas the malignant tumor is fabricated as irregular shaped[49]. Then, the tumors are inserted into the different locations of the fabricated head model to verify the system performance. For instance, the S-parameters (scattering parameters) with a benign tumor is depicted in Figure 2(b), and with a benign and a malignant tumor is illustrated in Figure 2(c), when antenna 1 is excited, and others are received the scattered signals. The backscattered signals: $S_{11}, S_{21}, S_{31}, \dots, S_{91}$ are collected in each 7.2° degrees rotation. Thus, in-total $8 \times 9 \times 50$ locations are scanned to evaluate and produce desired high-resolution images. Afterwards, the M-DMAS (modified-delay-multiply-and-sum) imaging algorithm [19] is applied on the collected backscattered signals to reconstruct the images of the head regions.

Figure 3 illustrates the fabricated phantom, different tumor with location, and reconstructed head region images. The tissue-mimicking phantom is used inside the skull model to investigate the overall performances of the developed system. The left side phantom image of Figure 3(a) illustrates the five locations (Top left side, top right side, bottom left side, bottom right side, and center point) indication of the head model. Moreover, for imaging purposes, a benign tumor, a malignant tumor, double benign tumors, double malignant tumors, and a benign and a malignant tumor as brain abnormalities in various locations are used to collect three-hundred reconstructed samples as an original image dataset. After that, the image preprocessing and augmentation methods are applied on original dataset to create a training dataset as a final image dataset to train the deep neural network model. Furthermore, the YOLOv5

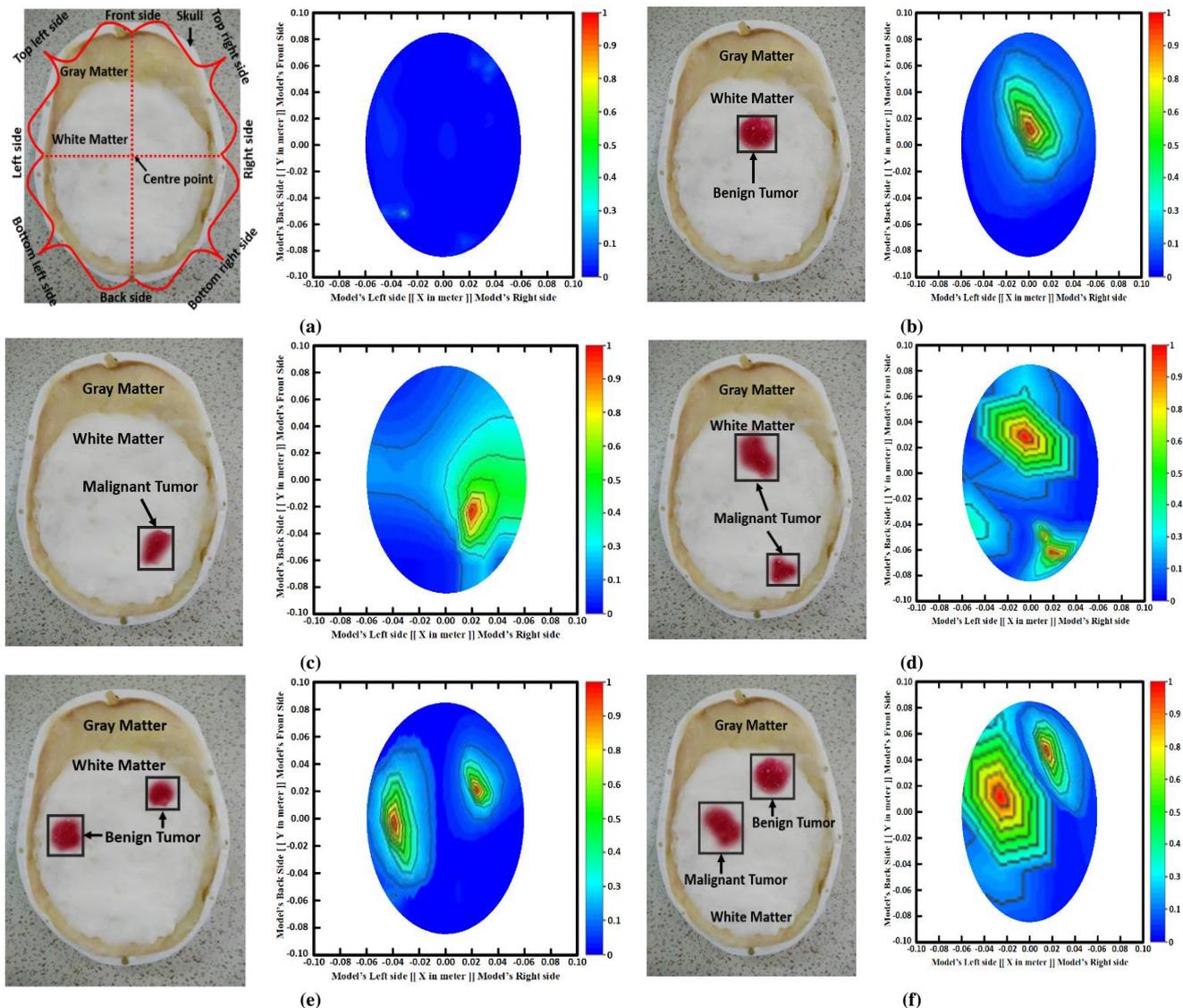


Figure 3: Fabricated head phantom model with brain abnormalities and corresponding reconstructed EM brain images (a) Without tumor image, (b) A single benign tumor image, (c) A single malignant tumor image, (d) Two malignant tumors image, (e) Two benign tumors image, (f) One benign and one malignant tumor image.

object detection model is used to automatically classify the tumor as the benign and the malignant, and then identify the tumor position from the produced reconstructed head images. The YOLOv5 model is implemented and executed by the Python high-level language with a PyTorch machine learning library.

3. BACKGROUND ANALYSIS OF YOLOv5 MODEL

3.1 Background of Yolov5 Algorithm

YOLO is the short name of the term "You look only once", which is the deep learning-based object detection algorithm invented by Joseph Redmon et al. [50]. YOLOv5 [51] is also the deep learning-based object detection algorithm, which improves the previous versions of YOLOv1 to YOLOv4, and it was recently released by the Ultralytics company [51] on 27 May 2020. It is much faster and has less computational architecture than the previous versions. The YOLOv5 model has the ability to balance detection accuracy and real-time performances. The YOLOv5 incorporates one stage deep convolutional neural network (DCNN) architecture that divided the images into a grid of cells. Straightly, each grid cell predicts a bounding box (BB) as well as to object categorization. Every cell is liable to detect the target object as an item and expects BB with the confidence scores (CS) [50]. In YOLOv5, the CS replicate whether a target object exists in a cell or not. Also, it predicts the object accurately. The CS is calculated by using the following formula (1):

$$CS = P_r(T_{object}) \times IoU(GT, PB) \quad (1)$$

Where $P_r(T_{object})$ is the prediction of the target object (T_{object}) and $P_r(T_{object}) \in \{0, 1\}$, intersection over union (IoU) calculates by between the GT and PB . GT means Ground Truth, and PB means predicted box. It is noticeable that if no target object presents in that grid cell, the CS should be one (1); otherwise, the CS should be the IoU between the GT and PB . In addition, each cell predicts C_L class probabilities for identifying the objects. However, the predicted output tensor (POT) is calculated by using equation (2) as follows:

$$POT = BB \times (5 + C_L) \quad (2)$$

Here, BB is the bounding boxes for each grid prediction, and it is considered as 3, "5" denotes every prediction box's four coordinates $[x, y, w, h]$ and one object confidence score (CS). In this instance, $[x, y]$ denotes the middle position coordinates of the box, and $[w, h]$ indicates the width and the height of the box, respectively. C_L represent the class probabilities. Hence an object can exist in numerous classes; thus, a single 1×1 convolutional layer, sigmoid, and Leaky ReLU activation functions are used in YOLOv5. In addition, the activation function Leaky ReLU is applied in intermediate and hidden layers, and the sigmoid function is involved in the ending prediction layer. The CS of each class is projected by the leaky ReLU and sigmoid activation function, and a threshold value identifies the object. In YOLOv5, the Binary Cross-Entropy with Logistic Loss (BCELL) function from the PyTorch library is used for calculating the loss of class probability and target object scores [52]. The estimated loss integrates a sigmoid layer and BCELoss in one class. The unreduced and multilevel classification loss functions in BCELL are as follows [52]:

$$L(x_i, y_i) = L = \{l_1, l_2, \dots, l_N\}^T \quad (3)$$

$$l_{ns} = -w_{ns} [y_{ns} \cdot \log \sigma(x_{ns}) + (1 - y_{ns}) \cdot \log(1 - \sigma(x_{ns}))] \quad (4)$$

$$L_c(x_i, y_i) = L_c = \{l_{1,c}, \dots, l_{N,c}\}^T \quad (5)$$

$$l_{ns,c} = -w_{ns,c} [p_c y_{ns,c} \cdot \log \sigma(x_{ns,c}) + (1 - y_{ns,c}) \cdot \log(1 - \sigma(x_{ns,c}))] \quad (6)$$

Here, N denotes the batch size, c represents the class number, (x_i, y_i) is the coordinate value of predicted boxes, the offsets of the center locations are presented by $\sigma(x_{ns})$ and $\sigma(x_{ns,c})$, ns is the number sample in the batch, p_c denotes the weight of the class c . If $c=1$, it would be a single-label binary classification and if $c>1$, it would be a multi-label binary classification.

3.2 Non-Maximum Suppression (NMS) Technique

An image contains multiple target objects, and the objects might be of different shapes and sizes. So, the target objects might be capture perfectly with a single bounding box. The YOLOv5 object detection model creates more than one overlapping bounding boxes (BB) in a single image to detect target objects, but needs to show only a single bounding box for each object in an image. Thus, we need a method that eliminates overlapping bounding boxes. The Non-Maximum Suppression technique is applied to eliminate the overlapping problem, which selects a single BB out of more than one overlapping BB to identify the objects in an image. The NMS method removes the redundant identifications and determines the best match for ending identification. The NMS technique is presented in Algorithm 1. Besides, YOLOv5 decide on GIoU loss[53] as the BB regression loss function to solve the inaccurate calculation of non-overlapping BB and defined as following equation [53]:

$$L_{GIoU} = 1 - IoU + \frac{|C_B - PB \cup B^{GT}|}{|C_B|} \quad (7)$$

Where B^{GT} represents the GroundTruth box, PB represents the predicted box, C_B represents the smallest box covering PB and

$$B^{GT}, \text{ and } IoU = \frac{PB \cap B^{GT}}{PB \cup B^{GT}}.$$

ALGORITHM 1: The NMS Technique Pseudocode

INPUT: $BB = \{bb_1, bb_2, bb_3, \dots, bb_{N-1}, bb_N\}$ $CS = \{s_1, s_2, s_3, \dots, s_{N-1}, s_N\}; T_V$ BB denotes the preliminary detection boxes. CS denotes the confidence score array that comprises the equivalent detection scores. T_V denotes the threshold value of NMS.OUTPUT: Total list of the last identification boxes D_B .

1. $D_B \leftarrow \{\}$
 2. While $BB \neq \emptyset$ do
 3. $M = \text{ArgMax } CS$;
 4. $D_B \leftarrow D_B \cup bb_M; BB \leftarrow BB - bb_M; CS \leftarrow CS - s_M$;
 5. For $bb_i \in BB$ do
 6. IF $\text{IoU}(bb_M, bb_i) \geq T_V$ Then
 7. $BB \leftarrow BB - bb_i$;
 8. $CS \leftarrow CS - s_i$;
 9. End IF
 10. End For
 11. End While
-

4. CLASSIFICATION AND DETECTION METHODOLOGY AND MATERIALS

This section discusses the study's overall classification and detection methodology, including dataset explanation, image preprocessing and data augmentation techniques, YOLOv5 architecture model with classification analysis, post-processing, and detection analysis. The complete methodology of the classification and detection is portrayed in Figure 4. This study used reconstructed microwave (MW) head images collected from the experimental MWHI system. The head images, including tumors and corresponding tumor object labelling in YOLO format, are provided to train the YOLOv5 model. As mentioned earlier, the study has mainly different types of images such as (i) single benign tumor images, (ii) single malignant tumor images, (iii) double benign tumors images, (iv) double malignant tumors images, and (v) single benign and single malignant. The work also explored different categories of reconstructed EM head images to investigate the performance of the YOLOv5 model to classify and detect tumor with localization in EM images.

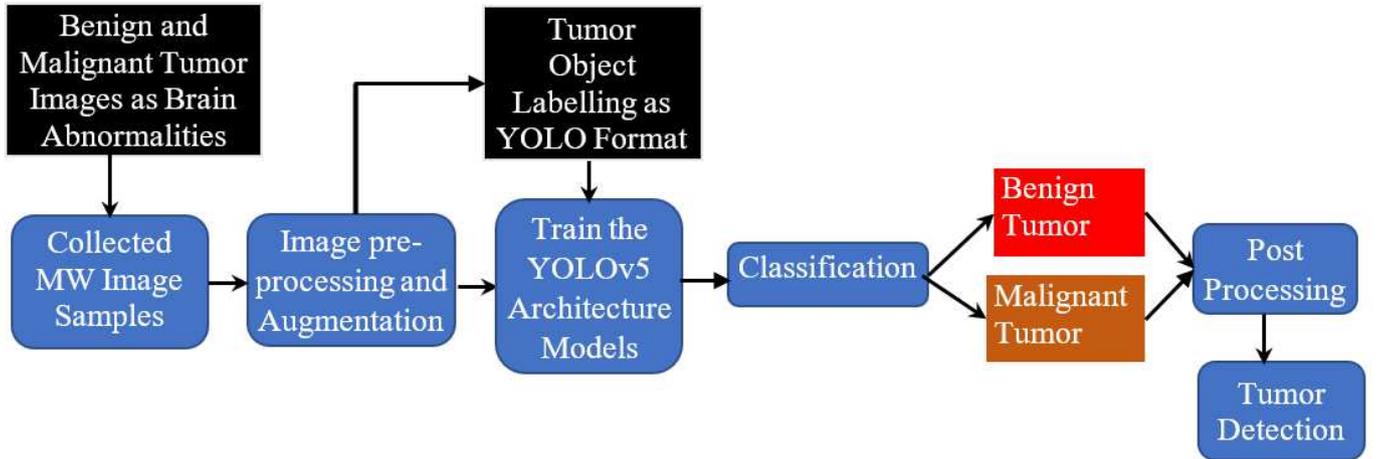


Figure 4: The proposed methodology diagram for tumor classification and detection.

4.1 Dataset Explanation

In this study, the original three hundred (300) reconstructed MW head image samples are initially collected from the implemented MWHI imaging system and then an original dataset is then created. The original dataset contains seventy-five (75) single benign tumor-based images, seventy-five (75) single malignant tumor-based images, fifty (50) double benign tumor-based images, fifty (50) double malignant tumor-based images, and fifty (50) single benign and single malignant tumor-based images. Image size is 640×640 pixels. After that, the image preprocessing and augmentation technique is applied to the original sample dataset to generate a final image dataset as a training dataset for training the YOLOv5 architecture model. After preprocessing and augmentation, the training dataset consists of a total of 3600 images. Some reconstructed MW head image samples with brain abnormalities from the original image dataset are illustrated in Figure 5.

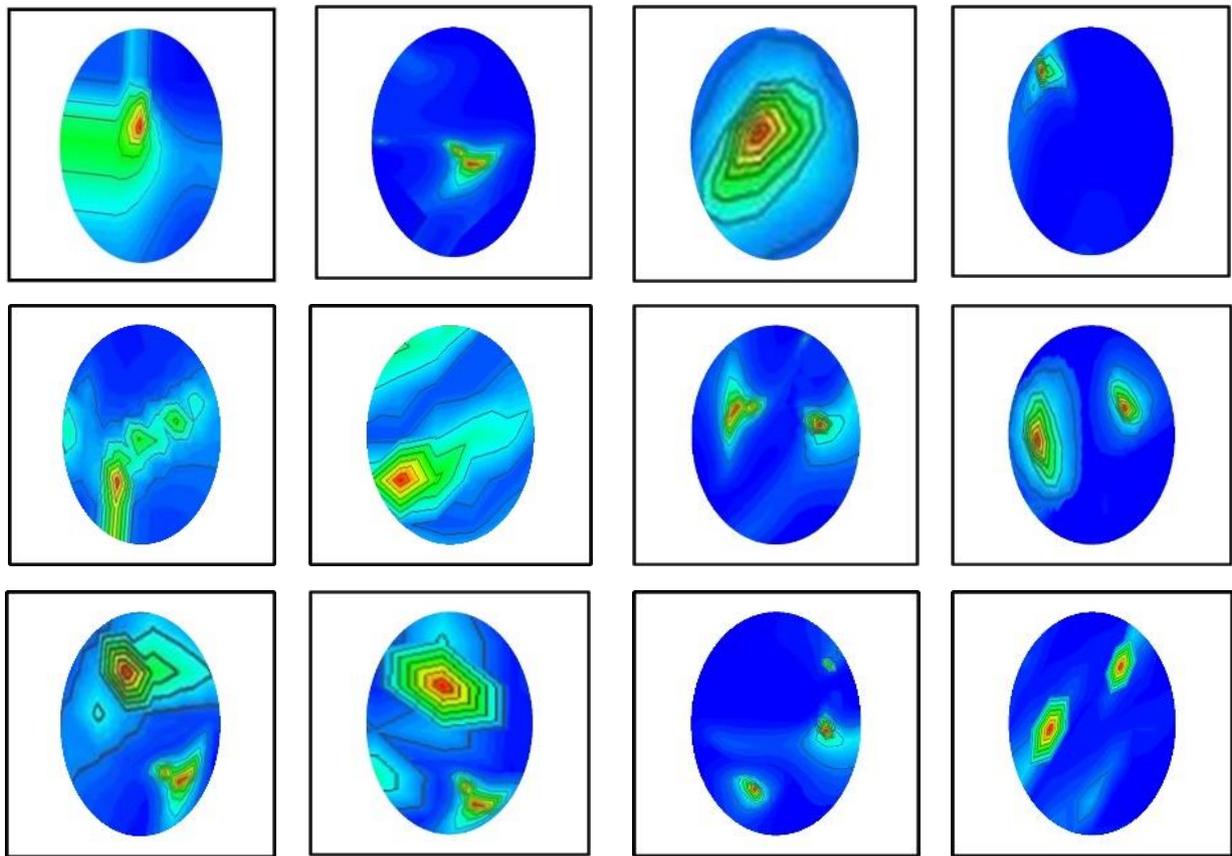


Figure 5: Randomly selected some reconstructed image samples with a benign and malignant tumor as brain abnormalities

4.2 Image Pre-processing and Augmentation Techniques

The image preprocessing technique is the first phase of a deep learning model due to its input image size requirements. For example, the YOLOv5 deep learning object detection model has an input size requirement. Thus, the images are preprocessed before training the network model. In this research, initially, from the original dataset, only the head region part of the images is cropped and then applied to resize and normalize preprocessing techniques to generate preprocessed images. Further, the images are resized to 640×640 pixels to fulfil the input requirements of the YOLOv5 model. In addition, by applying the Mean(M) and Standard Deviation (SD) of the images, the z-score normalization was completed to normalize the images. It is noticeable that the deep learning-based YOLOv5 model requires many images to be trained efficiently to classify and detect the target object locations in the EM image. Since we have an original three hundred image dataset, which is not suitable for training the model, thus, image augmentation techniques are applied to the original dataset in this study to generate an appropriate image dataset as a training dataset for effectively train the model. Image augmentation can enhance the classification and detection accurateness of the model by augmenting the existing image dataset rather than the collection of new samples or data. In addition, image augmentation can remarkably enhance the variety of data available for the training model and create a rich dataset from the insignificant image sample dataset for image categorization.

This study utilises eight different image augmentation strategies such as rotation, scaling, translating, horizontal flipping, vertical flipping, width shifting, height shifting, and zooming to generate the final image dataset. The rotation operation is done in the rotation scenario by rotating the images in the anticlockwise and clockwise direction with an angle between 3 to 30 degrees. As a result, shift the tumor objects at different positions in the images. Scaling is the reduction or magnification of the image frame size. So, 2% to 15% image magnifications are used for augmentation purposes in the scaling case. In addition, the image translation technique is done by translating the images vertically and horizontally by 3% to 10% to shift the tumor objects at different places in the images. Moreover, 5% to 15% width and height shifting are applied to the images for shifting the tumor objects. Also, the probability factor of 0.2 to 0.5 is applied to the images for vertical and horizontal flipping purposes. Thus, the tumor objects shift the position in the images. However, taking a single sample image from the original three-hundred (300) sample of images, and then applied all augmentation strategies separately to produce eleven (11) augmented images at a time by programmatically. As a result, obtained images are $11 \times 300 = 3300$ for the 300 samples. Finally, the final training image dataset comprises 3600 images, including the original 300 samples, which is further used to train the model. However, after image preprocessing and augmentation, some random image samples from the final training dataset are illustrated in Figure 6. For instance, only rotation and translation scenarios are illustrated in Figure 6.

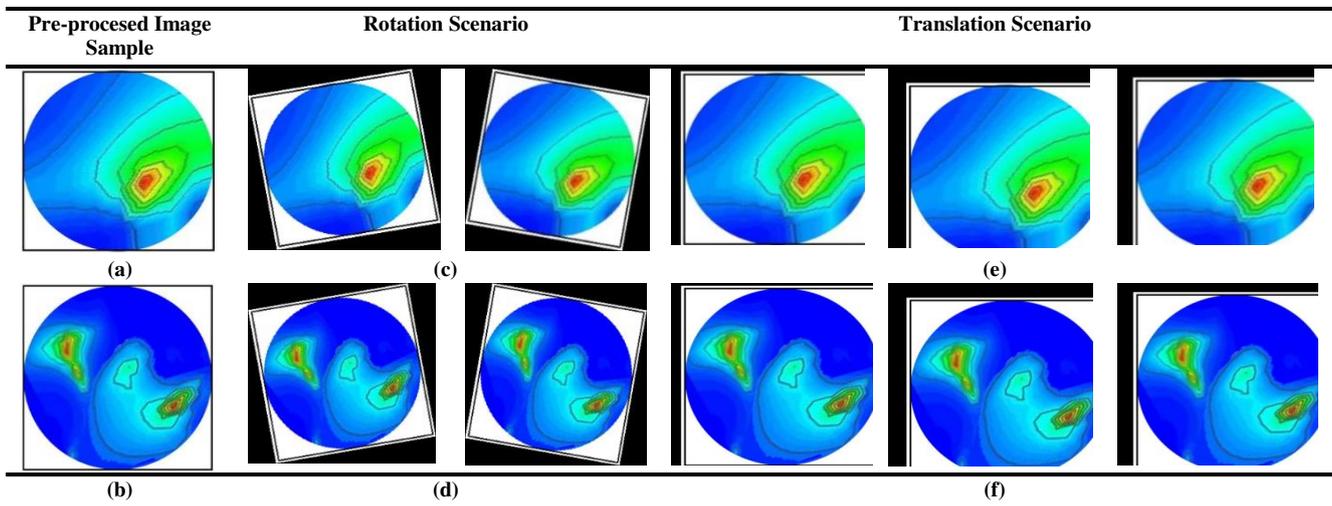


Figure 6. Randomly selected augmented image samples with abnormalities: (a) Single tumor image, (b) Double tumor image, (c-d) Image after rotation by 30 degrees anticlockwise and clockwise for single tumor and double tumor image, (e-f) Image after 3% horizontal, 5% vertical and horizontal, 5% horizontal and 3% vertical translation for single tumor and double tumor.

4.3 Target Object Labelling

The labelling of tumor objects in the image dataset is necessary for training the YOLOv5 model. The YOLOv5 object detection model is pre-trained on Microsoft Common Objects in Context (MSCOCO) dataset [51]. The dataset consists of 80 predefined classes, and the class names are stored in "yaml" file. But, our tumor object class names (i.e., target object) and dataset are not included in the predefined dataset of YOLOv5. Thus, we created a separate "yaml" class file and a YOLO tumor object labelling format file in this study. The "yaml" file consists of two classes: one is benign tumor class, and the other is malignant tumor class. On the other hand, the annotation process is applied to every image of the final image dataset to generate a tumor object labelling file as a YOLO format file. The YOLO format file is a text file with a similar name of each image that contains the tumor class id, tumor object's center coordinates, object's height and object's width (i.e., tumor labelling area). For tumor object labelling purposes, a graphical image annotation tool LabelImg is used to create an annotation file for every image. The tool uniformly labels each tumor object from the training and validation dataset. Finally, the training image dataset and corresponding YOLO format files as text files are used to train the network model. For instance, after annotation, the YOLO format file context is presented in the following Figure 7. The contents of the YOLO format file with a single benign tumor, single malignant tumor, and malignant and benign tumor objects are depicted in Figure 7. A file has three components for every target tumor object: tumor class id, center coordinates of the tumor bounding box (x_0 , y_0), tumor area (width, height). The id "0" denotes the benign tumor class, and "1" represents the malignant tumor class. The YOLO format file of benign tumor, malignant tumor, malignant and benign tumor labelling information are portrayed in Figures 7(a), 7(b), and 7(c), respectively.

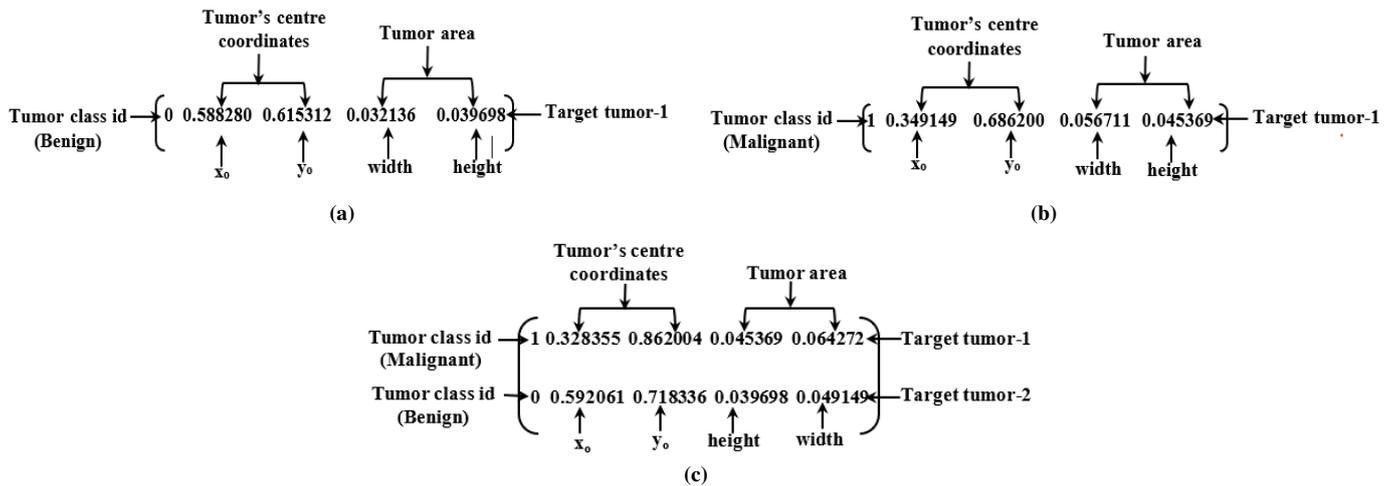


Figure 7: YOLO format file context: (a) For benign tumor as a single object, (b) For malignant tumor as a single object, (c) For malignant and benign tumor as two tumor objects.

4.4 Tumor-Yolov5 Architecture Model

In this research, we choose the YOLOv5 model due to some benefits: (i) The model has the advantages of accurate object classification with tumor localization, high speed, and high detection accuracy, (ii) The model can attain the identification of small tumor objects in noisy, blurry, foggy environment-based images. The YOLOv5 architectural model is portrayed in Figure 8. The architecture is divided into three parts: (i) Backbone, (ii) Neck, and (iii) Prediction. **Firstly**, the Bottleneck Cross Stage Partial Darknet (BCSPD) is the main backbone of the YOLOv5 framework. The input image dimension of YOLOv5 is 640×640×3 pixels. After feeding the images into the model, the output could be any combination of desired classes. The input image goes through the FOCUS module in the backbone. The FOCUS module slices the input image into four small ones and then concatenates them

together for convolutional operation. The $640 \times 640 \times 3$ pixels image is divided into four small images with the dimension of $320 \times 320 \times 3$, then that are concatenated into a $320 \times 320 \times 12$ pixels feature map. Thereafter, 32 convolutional kernels operation, it becomes a $320 \times 320 \times 32$ feature map. The CoBL module in the model is a fundamental convolutional module that epitomizes Conv2D + Batch Normalization (BN) + Leaky ReLU activation function. The BCSP mainly performs the feature extraction on the feature map, extracting major information from the feeding images. It solves the repeated gradient information duplication in CNNs optimization process and integrates the gradient changes into the feature map, thus reducing the input parameters and size of the model. The BCSP consists of two CoBL modules: one residual unit and two 1×1 Conv2D kernels. A residual unit contains the two CoBL modules and an adder, and the adder adds the features of the previous CoBL module output and two CoBL modules features output, then send the local features to one 1×1 Conv2D layer.

The four models with different input parameters can be attained, namely YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x by adjusting the width(w) and depth(d) of the BCSP module. Besides, the Spatial pyramid pooling (SPP) module integrates with the BCSP module in the backbone. The SPP module increases the receptive field of the network and acquires features of various scales. **Secondly**, in the neck part, YOLOv5 integrates Path Aggregation Network (PANet) to enhancement information flow [54]. The PANet is based on the feature pyramid network (FPN) structure that conveys robust semantic features from top to bottom. Also, FPN layers convey strong positioning features from bottom to top. In addition, PANet improves the propagation of low-level features and the utilization of accurate localization signals in bottom layers. Thus, enhance the location accuracy of the target object. **Thirdly**, the prediction layer is also called the detection or YOLO layer, generating (3) different feature maps to attain multi-scale prediction [55]. However, the model can classify and detect small, medium and large objects in the prediction layer. The YOLOv5 prediction process summary is discussed below:

Phase 1: Initially, the 640×640 pixels size of images are applied as input to the backbone. After that, the images are sliced by the FOCUS module. After executing many convolutions set of operation and two times BCSP1 operations, the feature map goes to the second concatenation layer. On the other hand, after one-time execution BCSP1, two times BCSP2, convolutions set of operation, and two times of upsampling, the feature map goes to the second concatenation layer. The second concatenation layer concatenates both of them. After executing the BCSP2 layers and 1×1 convolution operation, the 80×80 sized feature map as scale 1 is obtained.

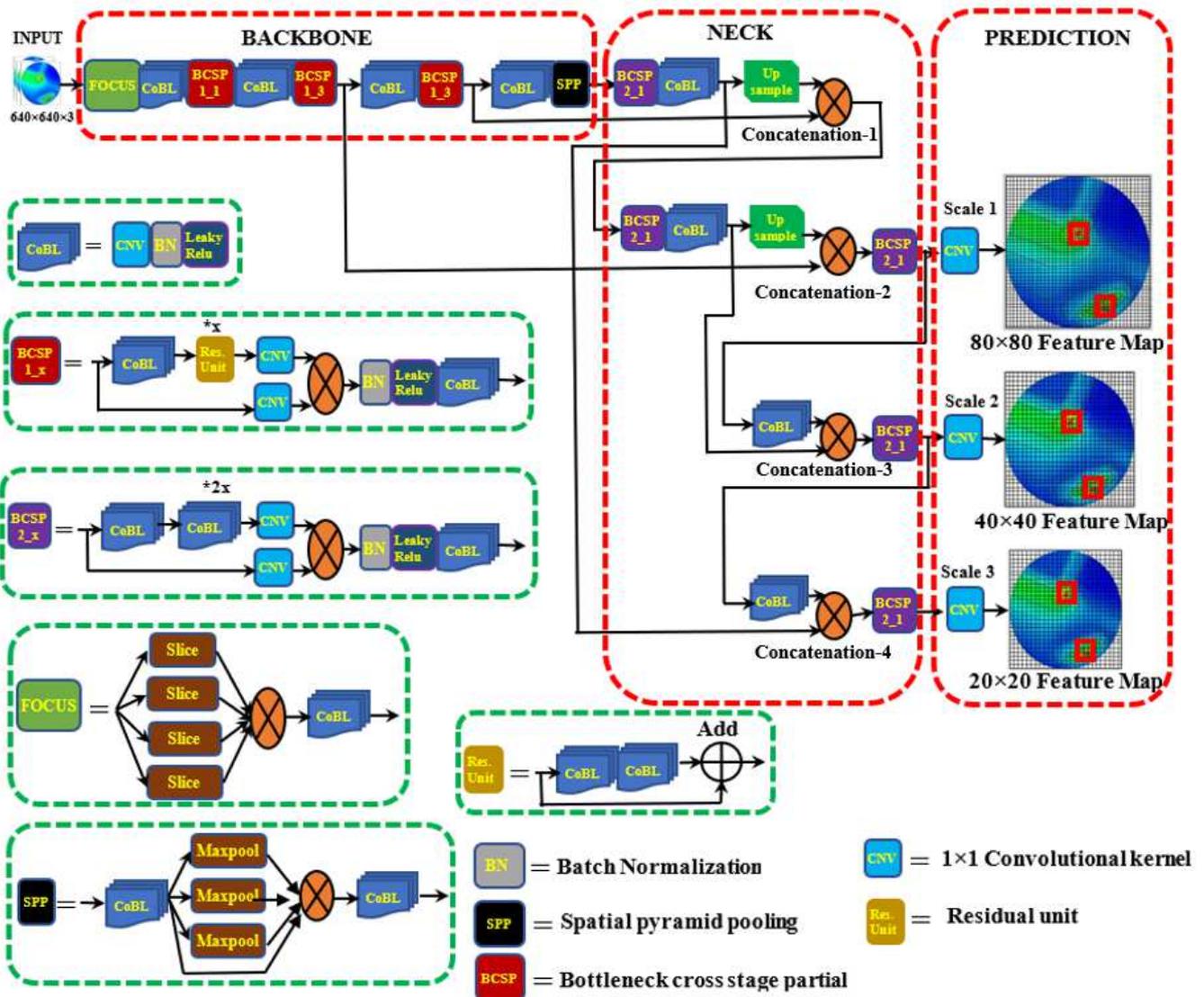


Figure 8: YOLOv5 architecture model

Phase 2: In the second step, the 80×80 sized feature map from phase 1 is processed by one 3×3 convolutional kernel and goes to the third concatenation layer. Besides, the feature map before the second upsampling is executed by one 1×1 convolutional kernel

and goes to the third concatenation layer. Then, the third concatenation layer concatenates both of them. After completing the BCSP2 layer and one 1×1 convolution operation, the 40×40 sized feature map as scale 2 is attained.

Phase 3: In the third phase, the 40×40 sized feature map from phase 2 is processed by one 3×3 convolutional kernel and goes to the fourth concatenation layer. Besides, the feature map before the first upsampling is executed by one 1×1 convolutional kernel and goes to the fourth concatenation layer. The fourth concatenation layer concatenates both of them. Then, executing by the BCSP2 layer and 1×1 convolution operation, the 20×20 sized feature map as scale 3 is obtained.

Phase 4: Finally, the feature maps of different sizes in scale 1 to 3 (i.e., 80×80, 40×40, 20×20) are improved for identifying the different size tumor objects with regression bounding boxes (BB). Henceforth, at every location, predicts three (03) regression bounding boxes for each feature map, thus creates 3 × (80×80, 40×40, 20×20) = 25200 regression bounding boxes. Finally, the predicted output of the model with BB is shown as a final tumor detection outcome.

4.5 Classifier, Post Processing and Detection

The YOLOv5 model is trained on MSCOCO dataset with 80 predefined object classes [51]. According to equation (2), the predicted output tensor (POT) dimension is 3 × (5+80) =255. Here, "3" denotes the three bounding boxes (BB) for each grid cell prediction, "5" denotes each prediction box's coordinates (x_o, y_o, w, h) and confidence score (CS), and "80" represents the predefined object class (C₁). However, in our study, we have two categories of objects in the EM head images: a benign tumor and a malignant tumor, respectively. Thus, we need to modify the classifier of the YOLOv5 model. So, according to equation (2), the predicted output tensor (POT) dimension is 3 × (5+2) =21 in our case. Due to the modification reduces the number of network parameters, computational overhead, and enhances detection accuracy. After training the network, the post-processing and detection phase is started. The testing image dataset evaluates the network model, and it shows the classification with tumor detection, either benign or malignant, with a bounding box and a detection score.

5. DETAILS EXPLANATION OF THE TRAINING EXPERIMENTS

The YOLOv5 deep learning model is implemented and executed on the anaconda distribution platform by using PyTorch machine learning library with Python 3.7 version. The experiments are carried out on 64-bit Windows 10 operating system with 128GB RAM and 64-bit Intel® Xeon®W-2016 @ of 3.30GHz CPU. In addition, a 32GB NVIDIA GeForce GTX 1082Ti GPU is used to enhance the network training performances. In this study, three versions: YOLOv5s, YOLOv5m, and YOLOv5l of the YOLOv5 model are carried out to evaluate the abnormality (i.e., tumor) classification and detection performances. In the experiment, the final training dataset and corresponding annotation class files are used in the model for training purposes. The final training dataset consists of 3600 images and related labelling files, where 80% of the total images are used for training and 20% for testing purposes. From the 80% training dataset, 20% is utilized for validation to avoid overfitting. Later, modified "ymal" class files and pre-trained learning weights for all models are used during the training period. The models are trained for 200 epochs with an initial learning rate of 0.01, and the batch size is 16. The stochastic gradient descent (SGD) optimizer is used to optimize the model outputs. The detailed training hyperparameters for the models are presented in Table 1.

Table1: Training hyperparameters of YOLv5 models

Parameter's name	Value	Parameter's name	Value
Initial learning rate	0.01	Weight decay	0.0005
Learning factor	0.2	Warmup epochs	3.0
Momentum	0.937	Warmup momentum	0.8
Optimizer	SGD	Warmup initial bias learning rate	0.1
IoU training threshold	0.20	Box loss gain	0.05
Anchors	3.0	Anchor multiple thresholds	4.0

5.1 Performance Evaluation Metrics

The tumor object classification with detection performance of the YOLOv5 network model is evaluated by the five evaluation matrices such as (i) Precision (P), (ii) Recall (R), (iii) F1-score (Fs), and (iv) Mean average precision (mAP). Precision represents the ability of a model to detect only the relative objects. On the other hand, recall represents the ability of a model to find out all the relevant cases. The evaluation matrices are calculated by using the following equations:

$$P = \frac{N_{TP}}{(N_{TP} + N_{FP})} \tag{8}$$

$$R = \frac{N_{TP}}{(N_{TP} + N_{FN})} \tag{9}$$

$$F_s = \frac{(2 \times N_{TP})}{(2 \times N_{TP} + N_{FN} + N_{FP})} \tag{10}$$

$$mAP = \sum_{i=1}^N P(i) \times \Delta R(i) \quad (11)$$

Where N_{TP} represents the number of tumor images were identified as tumors (either benign or malignant), N_{FP} represents the number of images was identified as tumor errors (either benign or malignant), N_{FN} represents the number of images that were missed identified as tumors (either benign or malignant), $P(i)$ is the precision, and $\Delta R(i)$ is the change in recall from the i^{th} detection.

6. RESULT AND DISCUSSION

This section discusses the experimental analysis of the training and testing performances, tumor classification, and evaluation of the tumor detection in EM brain images of the YOLOv5 models. The final training dataset consists of 3600 images, including benign and malignant tumor, where 80% of the total images are used for training the model and 20% for testing purposes. In further, from the 80% training dataset, 20% is utilized for validation to avoid overfitting. The experimental set-up of the model is discussed in section V.

6.1 Training and Validation Performance Analysis

We have implemented three versions of YOLOv5s, YOLOv5m, and YOLOv5l of the YOLOv5 model to investigate the training versus validation accuracy and loss performances by the different training datasets. It is seen that the YOLOv5l model showed the best performances. For simplicity, the training and validation accuracy graphs of the YOLOv5l model by the different training datasets are illustrated in Figure 9. It is observed from Figure 9, when 700 images were used as a training dataset then, initially, the training and validation accuracy is very low. However, after 78 epochs, both are reached 99.99% accuracy. Also, when training dataset size is increased as 1400, 2100, and 2880 to train the model, then the training and validation accuracy is increased to 99.99%, after 60, 35, and 22 epochs, respectively. In contrast, the training and validation loss graphs of the YOLOv5l model by the different training datasets are illustrated in Figure 10. It is observed from Figure 10, when 700 images were used as a training dataset, then initially, the training and validation loss is very high. However, after 72 epochs, both losses are reached approximately 0. Also, when training dataset size is increased as 1400, 2100, and 2880 to train the model, then the training and validation losses are decreased to approximately 0 after 38, 29, and 22 epochs, respectively. The overall average accuracy and loss investigation results are shown in Table 2.

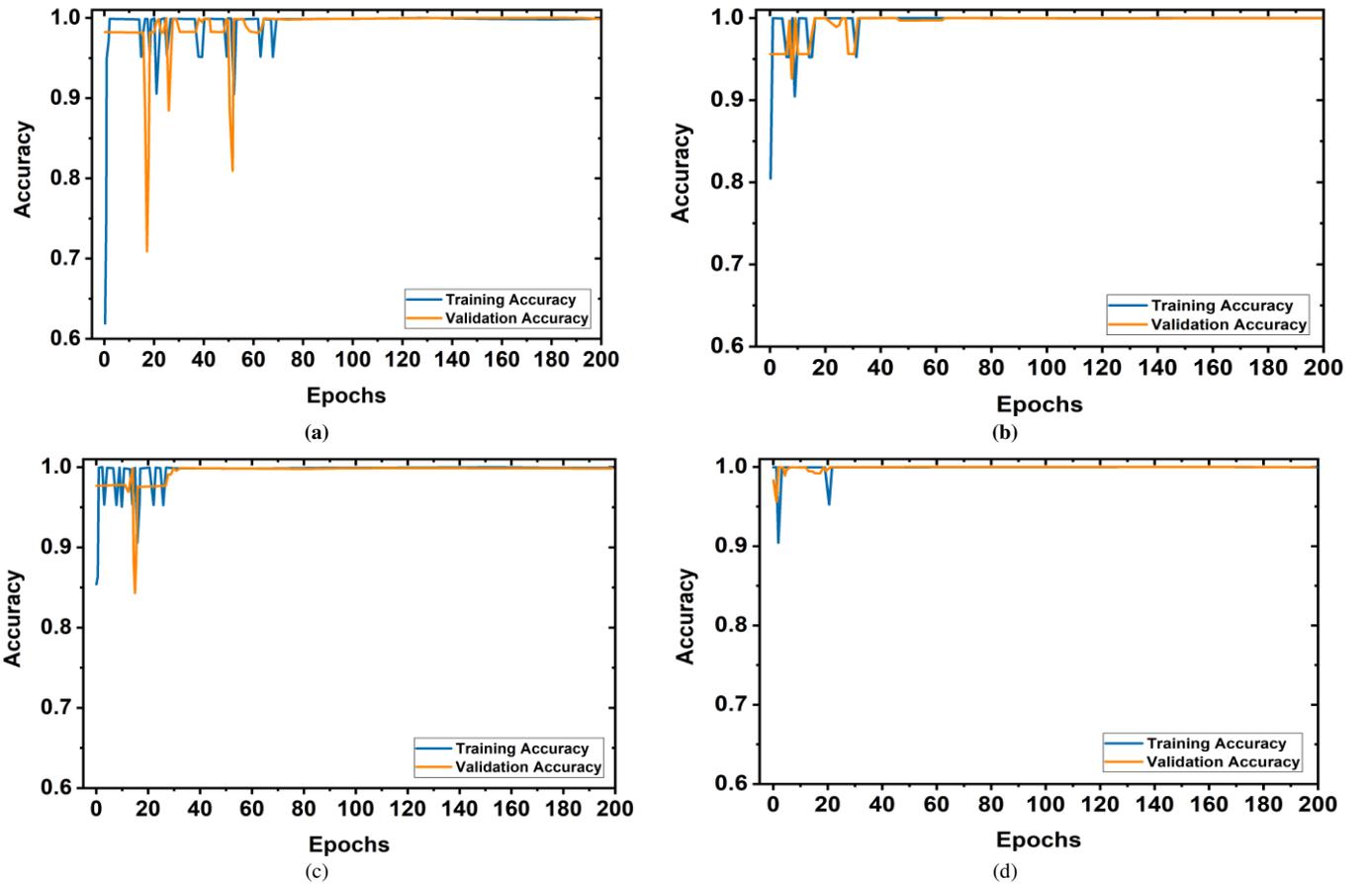


Figure 9: Training versus validation accuracy of YOLOv5l model for different training dataset: (a) For 700 images, (b) For 1400 images, (c) For 2100 images, (d) For 2880 images.

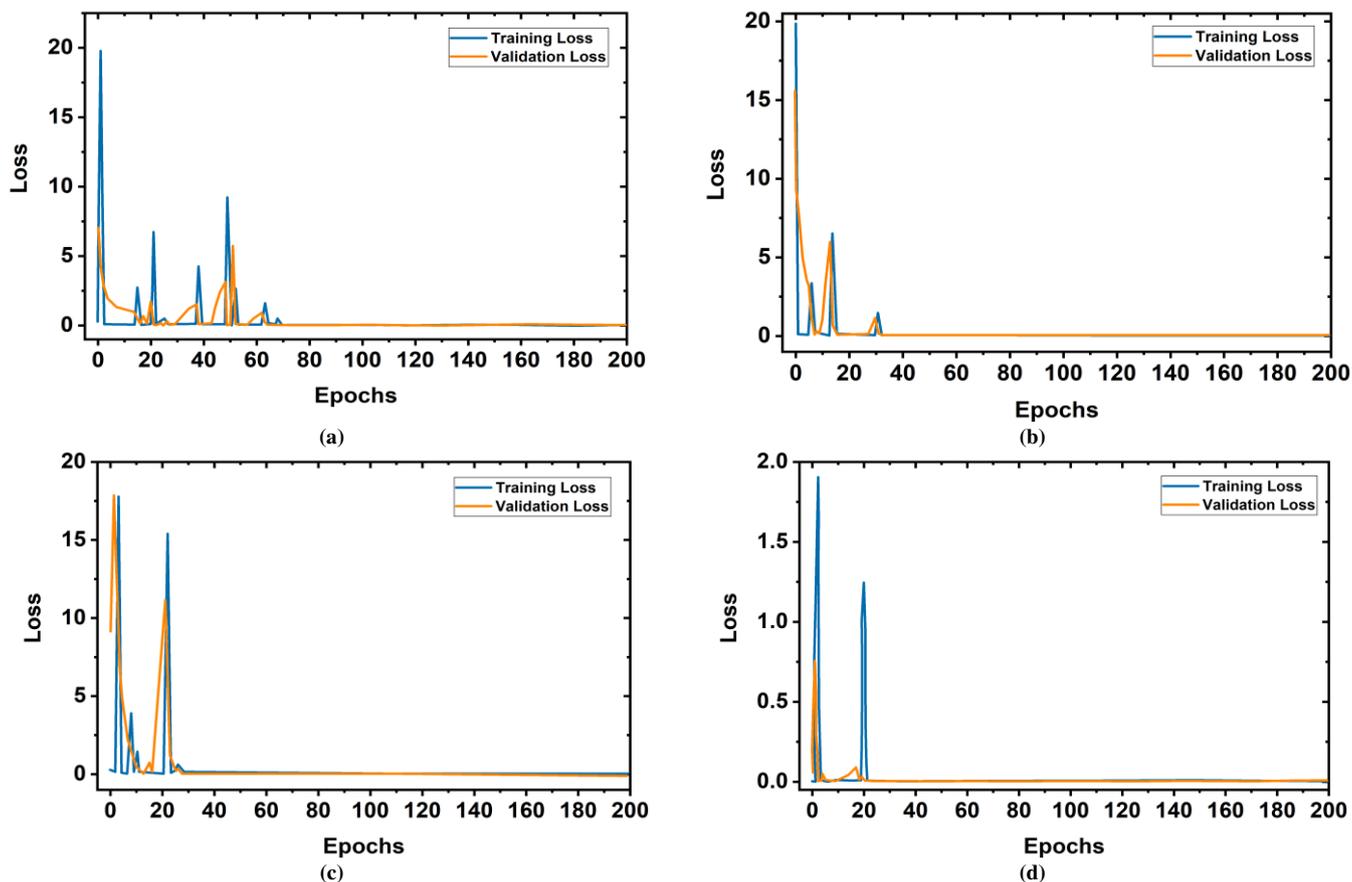


Figure 10: Training versus validation losses of YOLOv5l model for different training datasets: (a) For 700 images, (b) For 1400 images, (c) For 2100 images, (d) For 2880 images.

Table 2: Overall outcomes of training versus validation accuracy and loss

Training Dataset	Average Training Accuracy	Average Validation Accuracy	Average Training Loss	Average Validation Loss
700	95.98%	94.96%	2.38%	3.28%
1400	96.80%	95.45%	2.15%	2.79%
2100	97.67%	97.24%	1.04%	1.06%
2880 (80%)	99.84%	99.38%	0.0436%	0.0938%

6.2 Investigation of Tumor Object Classification and Prediction

6.2.1 Classification and Prediction Parametric Analysis

We have investigated two classes as brain abnormalities: benign tumor class and malignant tumor class performances, and tumor detection performances through the three versions such as YOLOv5s, YOLOv5m, and YOLOv5l of the YOLOv5 model by using the different EM brain image training dataset. *Firstly*, the precision, recall, training classification loss, and validation classification loss performance metrics with respect to epochs of three versions of the YOLOv5 model are investigated. Precision is the ability of a model to classify and predict only the target tumor objects. The precision curve of YOLOv5s, YOLOv5m, and YOLOv5l is illustrated in Figure 11(a), where the final image dataset is utilized. Figure 11(a) shows that the precision rate is initially very low in all models, but it increases with respect to increasing the epochs. In addition, the precision rate of the YOLOv5s model is varied in between 75% to 85%, whereas the rate of YOLOv5m is varied in between 78% to 89%, and the rate of YOLOv5l is varied in between 83% to 95%. As a result, the benign and malignant tumor classification and prediction rates of the YOLOv5l model are higher than the other two models. On the other hand, recall is the ability of a model to find out all the relevant tumor classes. The recall curve of YOLOv5s, YOLOv5m, and YOLOv5l is depicted in Figure 11(b), where the final image dataset is employed. From Figure 11(b), it is examined that the recall rate is initially low (50% to 80%) for the YOLOv5s and YOLOv5m models, but it is 60% to 95% for the YOLOv5l model. It is also observed that the recall rate increases with respect to increasing the epochs. Moreover, the recall rate of the YOLOv5s model is varied in between 80% to 90%, whereas the rate of the YOLOv5m is varied in between 85% to 95%, and the rate of the YOLOv5l is varied in between 88% to 98%. Thus, the target tumor classification and prediction rate of the YOLOv5l model is higher than the other two models. Besides, the mean average precision (mAP) for all models is observed, and it is shown in Figure 11(c). It is seen that the mAP of the YOLOv5l model is also higher than the YOLOv5m and YOLOv5s models. The overall investigation outcomes are presented in Table 3.

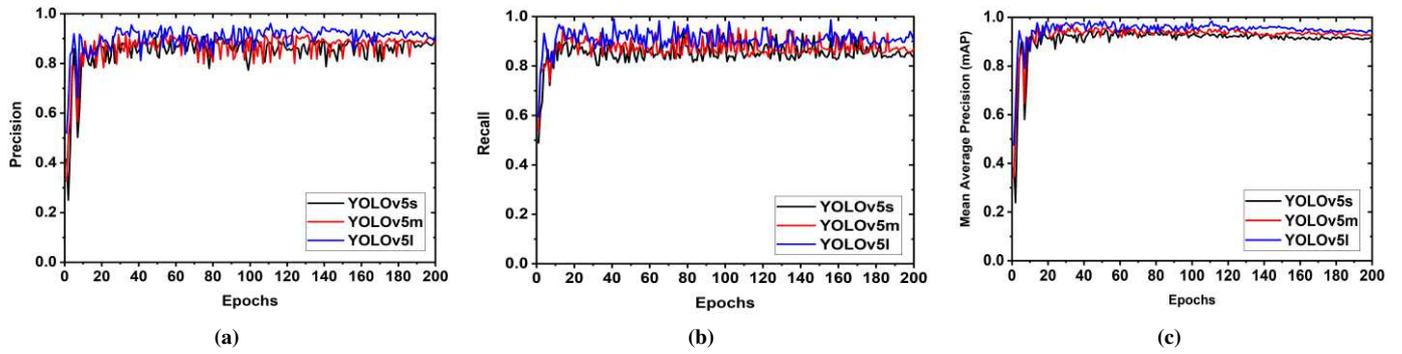


Figure 11: Tumor object classification and detection analysis curve with respect to epochs: (a) Precision (P), (b) Recall (R), (c) Mean average precision(mAP).

Secondly, precision, recall, F1 score, and PR curve (Precision \times Recall) are also investigated with respect to confidence. The precision curve is illustrated in Figure 12(a). From Figure 12(a), it is observed that at the beginning, the precision is very low for all models, but the precision of benign tumor and malignant tumor classes are gradually increased with increasing confidence. Moreover, all class precision is 100%, when confidence scores are 0.926, 0.923, and 0.896 for YOLOv5s, YOLOv5m, and YOLOv5l models. So, it should be noted that all classes precision of YOLOv5l is high at reasonable confidence (0.895) compared to other models. On the other hand, the recall curve is illustrated in Figure 12(b). In Figure 12(b), it is examined that at the beginning the

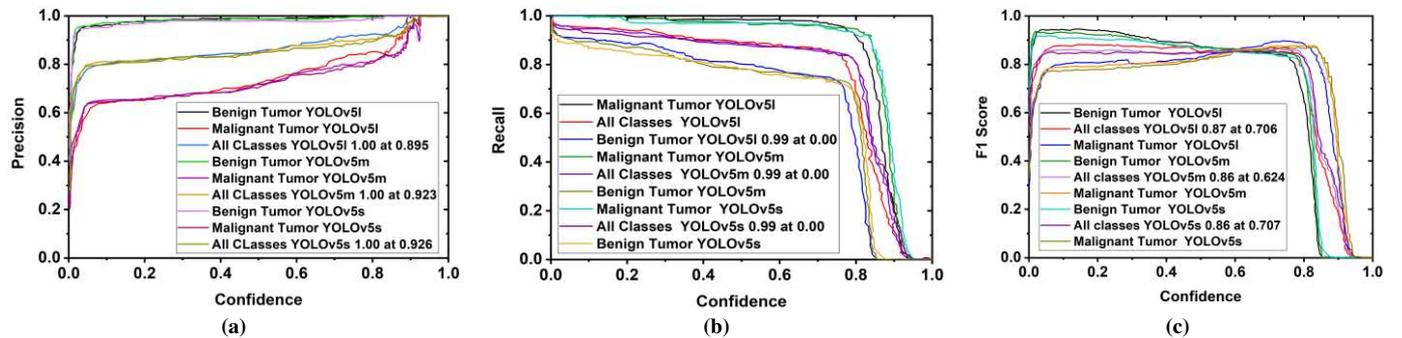


Figure 12: Benign and the malignant tumor classification performance curve with respect to confidence: (a) Precision (P), (b) Recall (R), (c) F1 score.

recall for benign and malignant tumor class is 99% for all models, but the recall of benign and malignant tumours class gradually decreases with increasing confidence up to 0.80. Although all classes recall is 0%, when confidence scores are between 0.85 to 1 for all models. However, the recall for benign and malignant tumor class of the YOLOv5l is higher than that of the other two models. In addition, another parametric F1 score is examined. The F1 curve is depicted in Figure 12(c). The F1 score is calculated based on precision and recall. Figure 12(c) shows that F1 scores are 86%, 86%, and 87% in all classes, when confidences are 0.707, 0.624, and 0.706 for YOLOv5s, YOLOv5m, and YOLOv5l, respectively. Thus, it should be noted that all classes F1 score of the YOLOv5l is high at moderate confidence (0.706) compared to the other two models. Furthermore, the Precision \times Recall curve is investigated to evaluate the performances of the tumor detection of the YOLOv5 model, as the confidence is changed for each class (i.e., benign and malignant tumor). It helps to evaluate tumor prediction ability when the precision maintains a significant value with the increase in the recall. The Precision \times Recall curve of YOLOv5s, YOLOv5m, and YOLOv5l is portrayed in Figure 13(a-c), respectively. It is observed from Figure 13 the better performance outcomes are displayed closer to the right corner in the Figures. The cross values for benign tumor class are 94.2%, 96.0%, and 96.0% for YOLOv5s, YOLOv5m, and YOLOv5l models, respectively, whereas the cross values for malignant tumor class are 89.7%, 88.7%, and 90.9% for YOLOv5s, YOLOv5m, and YOLOv5l models, respectively. Likewise, the mAP of all classes of YOLOv5s, YOLOv5m, and YOLOv5l models are 91.9%, 92.4%, and 93.5% at the rate of 0.5, respectively.

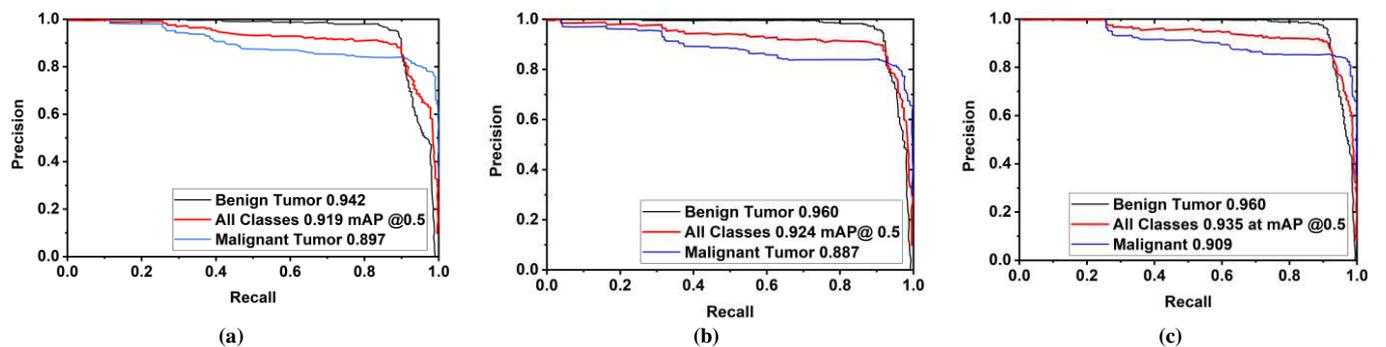


Figure 13: Precision \times Recall analysis curve for three models: (a) For YOLOv5s, (b) For YOLOv5m, (c) For YOLOv5l.

6.2.2 Tumor Classification Loss Analysis

The investigation of tumor class loss is significant to predict the detection score of tumor object in the EM images. So, the tumor classification loss is investigated during the training. There are two classification losses (training classification loss and validation classification loss) are observed when the EM head image training dataset is utilized. The training classification loss of the tumor objects with respect to the epochs is illustrated in Figure 14(a). Starting of the training, the training classification loss of the tumor object is relatively high (i.e., 0.25) in the three models. But the training classification loss is gradually decreased when increasing the iteration. We used 200 epochs for both training and validation classification in the investigation to get better classification outcomes. The training classification loss is comparatively higher in the YOLOv5s model and lowers in the YOLOv5l model. After completing 200 epochs, the train classification losses are 0.004, 0.003, 0.002 for YOLOv5s, YOLOv5m, and YOLOv5l, respectively. These incidences are found due to variation of pretrained weights of the YOLOv5 models. The best weight is used to train all models. The YOLOv5l model has large pretrained weights from the other two models and takes slightly more time than other models to train the model. But the YOLOv5l model shown a better result than other models.

In contrast, the validation classification loss of the tumor objects with respect to the epochs is illustrated in Figure 14(b). At the starting of the validation, the validation classification loss of the tumor object is relatively high such as 0.034 for YOLOv5s, 0.024 for YOLOv5m, and 0.017 for YOLOv5l models. Still, after 5 epochs, it abruptly decreased to 0.015, 0.013, and 0.007 for YOLOv5s, YOLOv5m, and YOLOv5l, respectively. Afterwards, it is also investigated from Figure 14(b), after completion of 200 epochs, the validation classification losses are 0.022, 0.015, and 0.012 for YOLOv5s, YOLOv5m, and YOLOv5l, respectively. But the losses for all models are slightly increased when increasing the epochs. It is noticeable that the YOLOv5l model shown better outcomes from the other two models. Finally, it is decided, the YOLOv5l model showed better tumor classification and detection results. The overall investigation outcomes are presented in Table 3.

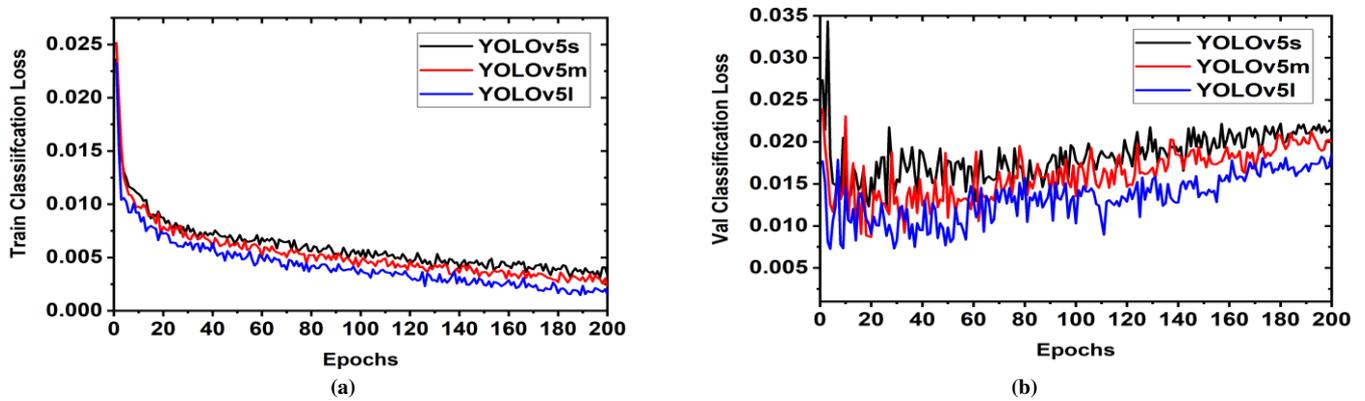


Figure 14: Tumor object classification loss: (a) Training classification loss, (b) Validation classification loss.

Table 3: Overall investigation outcomes of different classification parametric

Training Models	Weight Size	No. of Training image dataset	Precision (P) (%)	Recall (R) (%)	F1 score (%)	mAP (%)	Train Classification Loss	Validation Classification Loss
YOLOv5s	15 MB	700	81.9	87.5	84.6	87.8	0.00665	0.0225
		1400	82.2	88.0	85.0	88.3	0.00645	0.0198
		2100	83.1	88.2	85.5	88.7	0.00623	0.0192
		2880 (80%)	85.2	88.7	86.9	91.0	0.00620	0.0183
YOLOv5m	43 MB	700	82.7	87.7	85.2	87.9	0.00562	0.0195
		1400	83.8	88.3	85.9	88.5	0.00553	0.0189
		2100	86.8	88.8	87.7	89.0	0.00550	0.0175
		2880 (80%)	87.8	91.1	89.4	92.1	0.00545	0.0161
YOLOv5l	95 MB	700	83.2	87.9	85.4	88.0	0.00435	0.0172
		1400	84.8	88.4	86.5	89.7	0.00421	0.0166
		2100	89.8	91.0	90.39	93.1	0.00415	0.0150
		2880 (80%)	93.2	94.8	94.00	96.2	0.00400	0.0133

6.3 Tumor Classification and Detection Results in Different Case

The section discusses the tumor classification and detection outcomes in different cases by the trained YOLOv5s, YOLOv5m, and YOLOv5l models. Different cases are (i) images with one benign tumor in various locations, (ii) images with one malignant tumor in various locations, (iii) images with two benign tumor in different locations, (iv) images with two malignant tumor in different locations, and (v) images with one benign tumor and one malignant tumor in different locations. The confusion matrix of the YOLOv5l model is presented in Figure 15. It is realized from the confusion matrix; the model can classify the tumor objects into two classes and predict the benign and malignant tumours in EM head images.

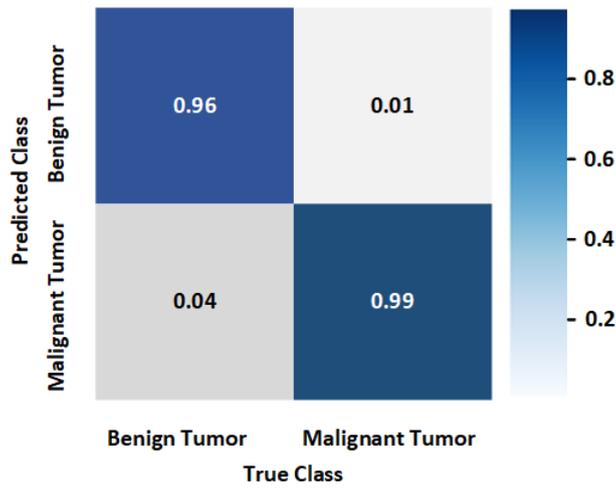
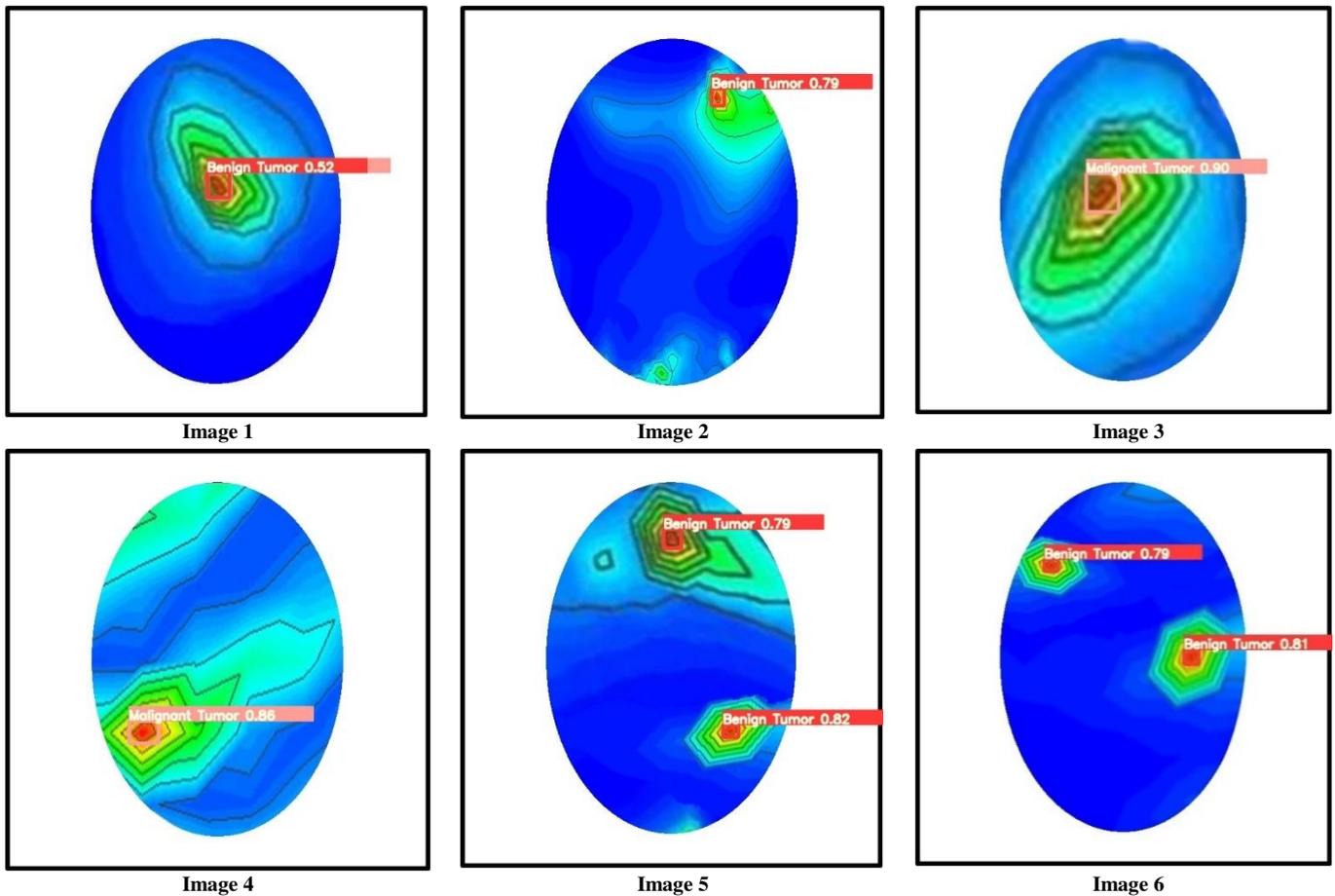


Figure 15: Confusion matrix of the YOLOv5l model

In this investigation, 2080 benign and 1780 malignant tumor instances exist in the final image dataset. It can be noticed that with the best performing model of YOLOv5l for two classes EM images, 4% out of 100% benign tumor is misclassified as a malignant tumor, and 1% out of 100% malignant tumor is misclassified as a benign tumor. From the final image dataset, 20% of images as a test set are used to detect the tumor objects in EM brain images. However, the experimental tumor detection outcomes by the YOLOv5s, YOLOv5m, and YOLOv5l are illustrated in Figures 16,17, and 18, respectively. For simplicity, only nine image detection results of three models are illustrated in this paper. All models automatically classified the tumor objects as benign and malignant tumours with their detection score accurately. But it is noticeable that the YOLOv5l model showed better results than the YOLOv5s and YOLOv5m models. In addition, the YOLOv5l model detected the target tumor object with their class, location, and detection accuracy more appropriately. The detection accuracy of the YOLOv5l model has increased 2% to 30% compared to other models for different cases, shown in the detection result Figures. However, an overall comparison of detection results with their classification and detection accuracy is presented in Table 4. In the end, it is decided that the YOLOv5l model shown appropriate tumor classification and detection outcomes in the portable MWHI system because of high precision, recall, F1 score, training accuracy, and low classification loss and validation loss.



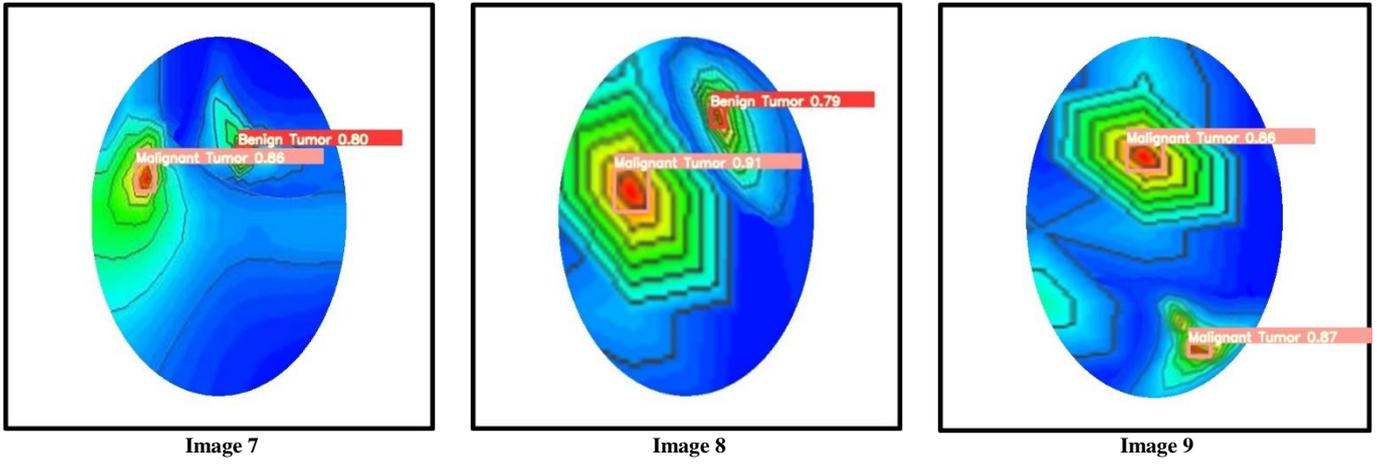


Figure 16: Tumor classification and detection results of the YOLOv5s model

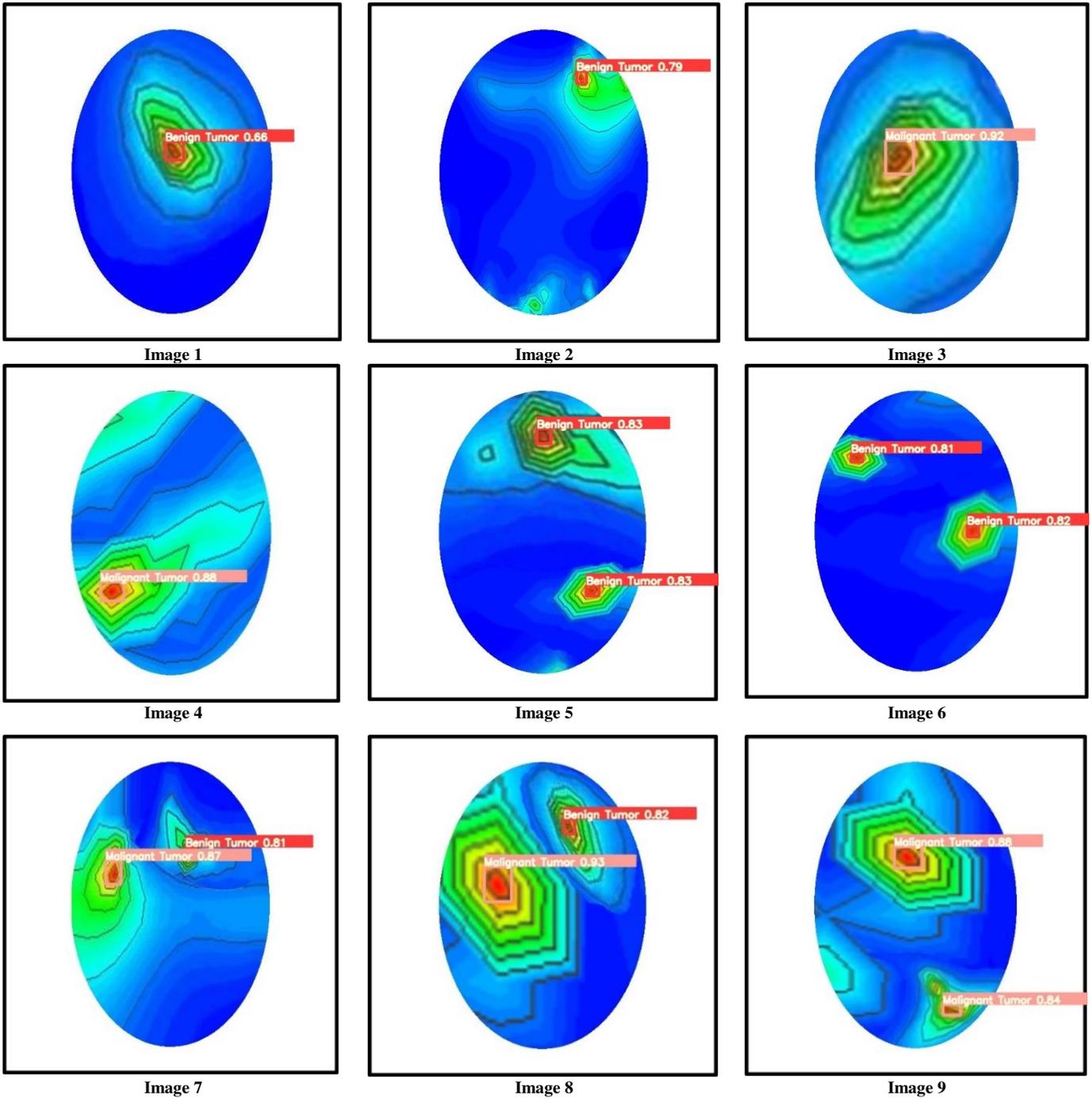


Figure 17: Tumor classification and detection results of the YOLOv5m model

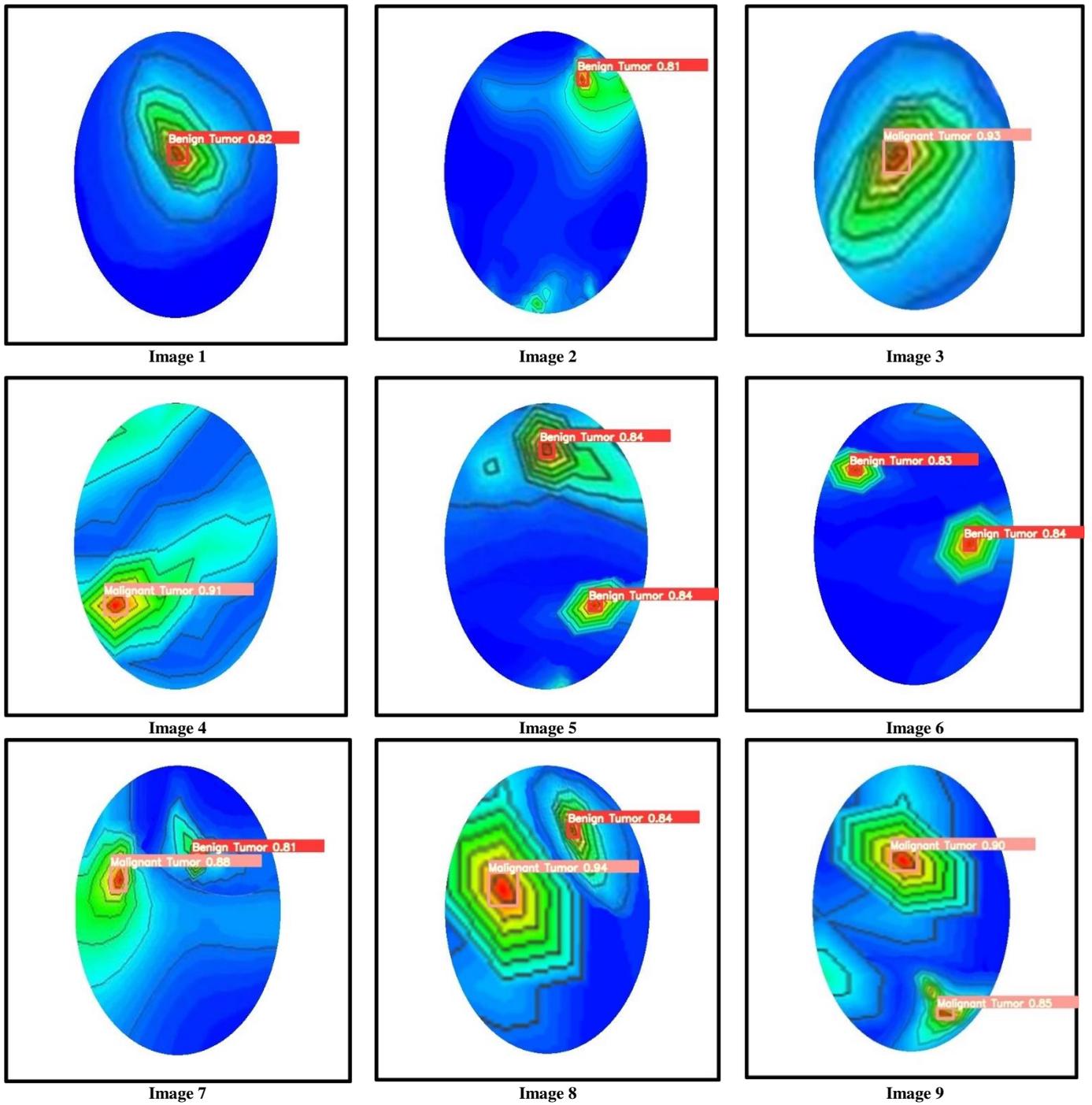


Figure 18: Tumor classification and detection results of the YOLOv5l model

Table 4: Overall comparison of tumor detection results with their classification and detection accuracy by the three models

Image No	Tumor classification				Detection Score of three models		
	Benign Tumor 1	Benign Tumor 2	Malignant Tumor 1	Malignant Tumor 2	YOLOv5s	YOLOv5m	YOLOv5l
1	Yes	No	No	No	0.52	0.66	0.82
2	Yes	No	No	No	0.79	0.79	0.81
3	No	No	Yes	No	0.90	0.92	0.93
4	No	No	Yes	No	0.86	0.88	0.91
5	Yes	Yes	No	No	0.79 and 0.82	0.83 and 0.83	0.84 and 0.84
6	Yes	Yes	No	No	0.79 and 0.81	0.81 and 0.82	0.83 and 0.84
7	Yes	No	Yes	No	0.80 and 0.86	0.81 and 0.87	0.81 and 0.88
8	No	Yes	No	Yes	0.79 and 0.91	0.82 and 0.93	0.84 and 0.94
9	No	No	Yes	Yes	0.86 and 0.87	0.88 and 0.84	0.90 and 0.85

7. CONCLUSION AND FUTURE WORK

This article uses a deep learning-based YOLOv5 model to automatically classify and detect brain abnormalities (i.e., tumors) in a portable microwave head imaging (MWHI) system. A 3D wideband nine antenna array-based MWHI system is utilized to analyze the backscattered signals and reconstruct the EM brain images. For the head region image reconstruction purposes, the system set up and tissue-mimicking head phantom including the benign and malignant tumor as brain abnormalities are employed to collect

backscattered signals, which are further post-processed by the M-DMAS image processing algorithm. The generated image dimension is 640×640 pixels. Three hundred sample images are collected from different locations in head regions by the developed MWHI system, which is later preprocessed and augmented to create a final image dataset. The final image dataset consists of 3600 images including original samples, then utilized for training, validation, and testing the YOLOv5 network model. Afterward, 80% of images are utilized to train the model, 20% are utilized to test the effectiveness of the model. The CoBL module, BCSP module, Leaky ReLU activation function, and many 3×3 and 1×1 convolutional layers are used in the YOLOv5 network model to optimize the network features, diminishing the size of the feature map on three scales. After investigation, it is noticeable that the YOLOv5l model demonstrated better outcomes than YOLOv5s and YOLOv5m models regarding tumor classification and detection performances. Moreover, the results of the YOLOv5l model shown high precision, recall, F1 score, training accuracy, mAP with low validation and training loss, validation and training classification loss. Furthermore, the testing stage of the best model YOLOv5l determines the overall MWHI system's ability to classify and detect the brain tumors automatically with a bounding box of its appropriate location, including a detection score in the MW brain images. Our future work aims to improve the tumor detection accuracy score by improving the YOLOv5 model.

Acknowledgements

The authors would like to acknowledge the Kuwait Foundation for the Advancement of Sciences (KFAS) for support of this project under the grant number CR20-13EE-01.

Author Contributions

Amran Hossain: Conception, Methodology, Design, Measurement, Programming, Experimental work, Result Investigation, and Writing - Original Draft. **Mohammad Tariqul Islam:** Experimental work, Project administration, Supervision, Updating the original draft, and Funding acquisition. **Ali F. Almutairi:** Conception, Program analysis, Measurement, and Writing- Reviewing and Editing the article, and Funding acquisition.

Conflicts of interest

The authors declare no conflict of interest.

REFERENCES

- [1] M. W. D. M. L. Schapira. Brain Tumor: Statistics [Online]. Available: <https://www.cancer.net/cancertypes/brain-tumor/statistics>
- [2] M. Tariq, A. A. Siddiqi, G. B. Narejo, and S. Andleeb, "A Cross Sectional Study of Tumors Using Bio-medical Imaging Modalities," *Current Medical Imaging*, vol. 15, pp. 66-73, 2019.
- [3] H. A. Ahmad, H. J. Yu, and C. G. Miller, "Medical imaging modalities," in *Medical Imaging in Clinical Trials*, ed: Springer, 2014, pp. 3-26.
- [4] A. F. Frangi, S. A. Tsafaris, and J. L. Prince, "Simulation and synthesis in medical imaging," *IEEE transactions on medical imaging*, vol. 37, pp. 673-679, 2018.
- [5] E. B. Adamson, K. D. Ludwig, D. G. Mummy, and S. B. Fain, "Magnetic resonance imaging with hyperpolarized agents: methods and applications," *Physics in Medicine & Biology*, vol. 62, p. R81, 2017.
- [6] R. L. Cazzato, J. Garnon, B. Shaygi, G. Koch, G. Tsoumakidou, J. Caudrelier, et al., "PET/CT-guided interventions: indications, advantages, disadvantages and the state of the art," *Minimally Invasive Therapy & Allied Technologies*, vol. 27, pp. 27-32, 2018.
- [7] S. Chakraborty, S. Chatterjee, A. S. Ashour, K. Mali, and N. Dey, "Intelligent computing in medical imaging: A study," in *Advancements in Applied Metaheuristic Computing*, ed: IGI global, 2018, pp. 143-163.
- [8] E. Dougeni, K. Faulkner, and G. Panayiotakis, "A review of patient dose and optimisation methods in adult and paediatric CT scanning," *European journal of radiology*, vol. 81, pp. e665-e683, 2012.
- [9] M. A. Jacobs, T. S. Ibrahim, and R. Ouwkerk, "MR imaging: brief overview and emerging applications," *Radiographics*, vol. 27, pp. 1213-1229, 2007.
- [10] K. M. Jones, K. A. Michel, J. A. Bankson, C. D. Fuller, A. H. Klopp, and A. M. Venkatesan, "Emerging magnetic resonance imaging technologies for radiation therapy planning and response assessment," *International Journal of Radiation Oncology* Biology* Physics*, vol. 101, pp. 1046-1056, 2018.
- [11] A. S. Alqadami, K. S. Bialkowski, A. T. Mobashsher, and A. M. Abbosh, "Wearable electromagnetic head imaging system using flexible wideband antenna array based on polymer technology for brain stroke diagnosis," *IEEE transactions on biomedical circuits and systems*, vol. 13, pp. 124-134, 2018.
- [12] A. Hossain, M. T. Islam, A. F. Almutairi, M. S. J. Singh, K. Mat, and M. Samsuzzaman, "An octagonal ring-shaped parasitic resonator based compact ultrawideband antenna for microwave imaging applications," *Sensors*, vol. 20, p. 1354, 2020.
- [13] A. Hossain, M. T. Islam, M. Islam, M. E. Chowdhury, H. Rmili, and M. Samsuzzaman, "A Planar Ultrawideband Patch Antenna Array for Microwave Breast Tumor Detection," *Materials*, vol. 13, p. 4918, 2020.
- [14] A. Hossain, M. T. Islam, M. S. Islam, M. E. Chowdhury, A. F. Almutairi, Q. A. Razouqi, et al., "A YOLOv3 Deep Neural Network Model to Detect Brain Tumor in Portable Electromagnetic Imaging System," *IEEE Access*, 2021.
- [15] A. Mobashsher, K. Bialkowski, A. Abbosh, and S. Crozier, "Design and experimental evaluation of a non-invasive microwave head imaging system for intracranial haemorrhage detection," *Plos one*, vol. 11, p. e0152351, 2016.
- [16] A. T. Mobashsher, A. M. Abbosh, and Y. Wang, "Microwave system to detect traumatic brain injuries using compact unidirectional antenna and wideband transceiver with verification on realistic head phantom," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, pp. 1826-1836, 2014.
- [17] J. A. Tobon Vasquez, R. Scapatucci, G. Turvani, G. Bellizzi, N. Joachimowicz, B. Duchêne, et al., "Design and experimental assessment of a 2D microwave imaging system for brain stroke monitoring," *International Journal of Antennas and Propagation*, vol. 2019, 2019.
- [18] A. Hossain, M. T. Islam, M. E. Chowdhury, and M. Samsuzzaman, "A grounded coplanar waveguide-based slotted inverted delta-shaped wideband antenna for microwave head imaging," *IEEE Access*, vol. 8, pp. 185698-185724, 2020.
- [19] M. S. Islam, M. T. Islam, A. Hoque, M. T. Islam, N. Amin, and M. E. Chowdhury, "A Portable Electromagnetic Head Imaging System Using Metamaterial Loaded Compact Directional 3D Antenna," *IEEE Access*, vol. 9, pp. 50893-50906, 2021.
- [20] B. J. Mohammed, A. M. Abbosh, S. Mustafa, and D. Ireland, "Microwave system for head imaging," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, pp. 117-123, 2013.
- [21] M. Rahman, M. Islam, M. Mahmud, and M. Samsuzzaman, "Compact microstrip patch antenna proclaiming super wideband characteristics," *Microwave and Optical Technology Letters*, vol. 59, pp. 2563-2570, 2017.
- [22] M. Rokunuzzaman, A. Ahmed, T. C. Baum, and W. S. Rowe, "Compact 3-D antenna for medical diagnosis of the human head," *IEEE Transactions on Antennas and Propagation*, vol. 67, pp. 5093-5103, 2019.
- [23] A. Salleh, C. Yang, T. Alam, M. Singh, M. Samsuzzaman, and M. Islam, "Development of microwave brain stroke imaging system using multiple antipodal vivaldi antennas based on raspberry Pi technology," *J. Kejuruteraan*, vol. 32, pp. 1-6, 2020.
- [24] A. E. Stancombe, K. S. Bialkowski, and A. M. Abbosh, "Portable microwave head imaging system using software-defined radio and switching network," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, vol. 3, pp. 284-291, 2019.

- [25] A. S. Alqadami, K. S. Bialkowski, A. T. Mobashsher, and A. M. Abbosh, "Wearable electromagnetic head imaging system using flexible wideband antenna array based on polymer technology for brain stroke diagnosis," *IEEE transactions on biomedical circuits and systems*, vol. 13, pp. 124-134, 2019.
- [26] A. S. Alqadami, N. Nguyen-Trong, B. Mohammed, A. E. Stancombe, M. T. Heitzmann, and A. Abbosh, "Compact unidirectional conformal antenna based on flexible high-permittivity custom-made substrate for wearable wideband electromagnetic head imaging system," *IEEE Transactions on Antennas and Propagation*, vol. 68, pp. 183-194, 2019.
- [27] I. Merunka, A. Massa, D. Vrba, O. Fiser, M. Salucci, and J. Vrba, "Microwave tomography system for methodical testing of human brain stroke detection approaches," *International Journal of Antennas and Propagation*, vol. 2019, 2019.
- [28] R. Scapatucci, J. Tobon, G. Bellizzi, F. Vipiana, and L. Crocco, "Design and numerical characterization of a low-complexity microwave device for brain stroke monitoring," *IEEE Transactions on Antennas and Propagation*, vol. 66, pp. 7328-7338, 2018.
- [29] A. Mobashsher and A. Abbosh, "Compact 3-D slot-loaded folded dipole antenna with unidirectional radiation and low impulse distortion for head imaging applications," *IEEE Transactions on Antennas and Propagation*, vol. 64, pp. 3245-3250, 2016.
- [30] M. Rokunuzzaman, M. Samsuzzaman, and M. T. Islam, "Unidirectional wideband 3-D antenna for human head-imaging application," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 169-172, 2016.
- [31] S. A. Rezaeieh, A. Zamani, and A. Abbosh, "3-D wideband antenna for head-imaging system with performance verification in brain tumor detection," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 910-914, 2015.
- [32] M. A. Ullah, T. Alam, M. S. Alam, S. Kibria, and M. T. Islam, "A unidirectional 3D antenna for biomedical microwave imaging based detection of abnormality in human body," *Microsystem Technologies*, vol. 24, pp. 4991-4996, 2018.
- [33] M. Böhle, F. Eitel, M. Weygandt, and K. Ritter, "Layer-wise relevance propagation for explaining deep neural network decisions in MRI-based Alzheimer's disease classification," *Frontiers in aging neuroscience*, vol. 11, p. 194, 2019.
- [34] D. Lu, N. Polomac, I. Gacheva, E. Hattingen, and J. Triesch, "Human-expert-level brain tumor detection using deep learning with data distillation and augmentation," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 3975-3979.
- [35] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, pp. 102-127, 2019.
- [36] E. A. Rashed, J. Gomez-Tames, and A. Hirata, "Deep learning-based development of personalized human head model with non-uniform conductivity for brain stimulation," *IEEE transactions on medical imaging*, vol. 39, pp. 2351-2362, 2020.
- [37] T. A. Anjit, R. Benny, P. Cherian, and P. Mythili, "Non-Iterative Microwave Imaging Solutions for Inverse Problems Using Deep Learning," *Progress In Electromagnetics Research M*, vol. 102, pp. 53-63, 2021.
- [38] V. Khoshdel, "Using deep learning approaches for microwave imaging," 2021.
- [39] V. Khoshdel, M. Asefi, A. Ashraf, and J. LoVetri, "Full 3D microwave breast imaging using a deep-learning technique," *Journal of Imaging*, vol. 6, p. 80, 2020.
- [40] W. Shao and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Transactions on Antennas and Propagation*, vol. 68, pp. 5626-5635, 2020.
- [41] A. Yago, M. Cavagnaro, and L. Crocco, "Deep Learning-Enhanced Qualitative Microwave Imaging: Rationale and Initial Assessment," in *2021 15th European Conference on Antennas and Propagation (EuCAP)*, 2021, pp. 1-5.
- [42] M. Ambrosanio, S. Franceschini, F. Baselice, and V. Pascazio, "Machine learning for microwave imaging," in *2020 14th European Conference on Antennas and Propagation (EuCAP)*, 2020, pp. 1-4.
- [43] A. Ari and D. Hanbay, "Deep learning based brain tumor classification and detection system," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 26, pp. 2275-2286, 2018.
- [44] T. Saba, A. S. Mohamed, M. El-Affendi, J. Amin, and M. Sharif, "Brain tumor detection using fusion of hand crafted and deep learning features," *Cognitive Systems Research*, vol. 59, pp. 221-230, 2020.
- [45] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks," in *annual conference on medical image understanding and analysis*, 2017, pp. 506-517.
- [46] P. Shah and M. Moghaddam, "Super resolution for microwave imaging: A deep learning approach," in *2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, 2017, pp. 849-850.
- [47] A. Mobashsher and A. Abbosh, "Three-dimensional human head phantom with realistic electrical properties and anatomy," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 1401-1404, 2014.
- [48] Y. Cheng and M. Fu, "Dielectric properties for non-invasive detection of normal, benign, and malignant breast tissues using microwave theories," *Thoracic cancer*, vol. 9, pp. 459-465, 2018.
- [49] M. Lazebnik, D. Popovic, L. McCartney, C. B. Watkins, M. J. Lindstrom, J. Harter, *et al.*, "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign and malignant breast tissues obtained from cancer surgeries," *Physics in medicine & biology*, vol. 52, p. 6093, 2007.
- [50] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779-788.
- [51] G. Jocher, ultralytics/yolov5 [Online]. Available: <https://github.com/ultralytics/yolov5>
- [52] G. Jocher, BCEWithLogitsLoss [Online]. Available: <https://pytorch.org/docs/master/generated/torch.nn.BCEWithLogitsLoss.html#bcewithlogitsloss>
- [53] C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 390-391.
- [54] K. Wang, J. H. Liew, Y. Zou, D. Zhou, and J. Feng, "Panet: Few-shot image semantic segmentation with prototype alignment," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 9197-9206.
- [55] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.