YOLOv12 based Deep Learning Model for Real-Time Brain Tumor Detection in Computed Tomography Images

1st Babyrani Waikhom
Computer Science and Engineering Department
Amrita School of Computing, Amrita Vishwa Vidyapeetham
Chennai, India
babyraniw@gmail.com

2nd Nivethitha V
Computer Science and Engineering Department
Amrita School of Computing, Amrita Vishwa Vidyapeetham
Chennai, India
v_nivethitha@ch.amrita.edu

3rd W. Sharatchandra Singh
Nondestructive Evaluation Division
Indira Gandhi Centre for Atomic Research
Kalpakkam, India
ORCID: https://orcid.org/0000-0003-0031-378X

4th Suthir Sriram
Computer Science and Engineering Department
Amrita School of Computing, Amrita Vishwa Vidyapeetham
Chennai, India
s_suthir@ch.amrita.edu

***Abstract*—This paper proposes a YOLOv12 based deep learning model for real-time detection of brain tumors at an early stage and accurately finding their location in medical computed tomography (CT) images. The model was optimized to improve the detection accuracy and robustness. The results show that the YOLOv12 model is a powerful tool for the automated detection and classification of brain CT images. The model detects brain tumor and non-tumor with mean average precision, mAP50 of 0.905, precision of 0.761 and recall scores of 0.814. The detection results were also compared with other YOLO models namely, YOLOv5, YOLOv8 and YLOv11. The model outperforms existing YOLO models in terms of higher mAP50 and mAP50-90, showcasing the superiority of YOLOv12 for this purpose. The study demonstrates the potential of YOLOv12 for automated detection and classification of brain tumor and non-tumor, thereby offering a promising solution for earlier intervention and improved patient survival rate.**

***Index Terms*—YOLO, medical CT imaging, real-time tumor detection, deep learning**

I. Introduction

Brain tumor ranks among the top causes of tumor-related deaths and contributes to a large percentage of total tumor reports. As per the report by World Health Organization (WHO), the tenth position in the list of various cancer-related diseases is held by brain tumors [1]. The symptoms exhibited by this condition include severe headaches, seizures, visual and hearing impairments, and loss of intellectual functions. These factors, in turn, impact the quality of life and physiological conditions of patients. The general cause of this tumor is the accumulation of pathological cells in the brain [2]. It also has the capacity to invade surrounding tissues rapidly and cause the development of metastases, resulting in a potentially life-threatening situation. Brain tumors can be diagnosed according to types, origin, rate of growth, and stages of development. Benign (non-cancerous) and malignant (cancerous) brain tumors are primarily diagnosed depending upon the rates of growth, invasiveness, and ability to metastasize in brain tissue around the cancers. Malignant brain tumors grow rapidly along with metastasis; such conditions may increase complexities in treatment processes as well. Proper diagnosis of brain tumors is very significant in order to effectively deal with treatment processes to maximize outcomes of the treatment itself [3-4].

Computed tomography (CT) and magnetic resonance imaging (MRI) are widely used to determine the presence and extent of the brain tumor by visualizing abnormalities in the structural and functional of the brain. These medical imaging technologies have the potential to afford detailed internal structures non-invasively, which increases the diagnostic accuracy [5]. However, this type of imaging data is often unable to clearly distinguish characteristics of a tumor all by itself. Traditional brain tumor evaluation from the images of CT and MRI has mainly relied on visual inspection by doctors, which highly relies on the expertise of the doctor. This manual interpretation of images takes too much time and has prone chances of error. Therefore, there is an urge for the development of more reliable and automated systems with enhanced diagnostic accuracy in lesser time.

In the last decade, deep learning (DL) techniques are finding applications in medical imaging to detect and classify the images, because of its ability for automation with improved detection speed, consistency, and accuracy. Among various deep learning techniques, object detection models have shown considerable promise, particularly for detection and identification of specific targets within the medical images, such as brain tumors. The YOLO (You Only Look Once) family of models is particularly suited for this task as it offers real-time object detection, thereby enabling faster and more efficient

processing of large-scale datasets of medical images [6-7]. The YOLO model is capable of accurately identifying different objects faster and possesses a vast application potential within the domain of medical diagnosis. The YOLO approach divides the input image into a grid of cells while simultaneously predicting the bounded-boxes and classes corresponding to each cell. The bounded-boxes identify the position of different objects within the images, whereas the classes identify the type of object. The approach allows the combined detection of different objects and their corresponding positions within the images. All the detections happen simultaneously during the regression of the neural network. The quality metrics used for evaluating the efficiency of the YOLO approach includes accuracy, precision, recall rate, and F1-score. As the YOLO approach depends upon grid cell detection of objects through the entire images in a single pass of the neural network system, the approach enables the rapid examination of extensive datasets. The YOLO approach easily assists in the utilization of the approach for the rapid examination of extensive datasets of the medical domain. The rapid examination of extensive datasets becomes a significant aspect for improving the clinical applicability of the approach [8-9].

The objective of this study is to propose a YOLOv12 based deep learning model for real-time detection of brain tumors at an early stage and accurately finding their location in medical CT images. The study also aims to compare its performance with other existing YOLO models to detect the brain tumor in CT images.

II. Related Work

In the recent years, YOLO model has received considerable attention in the detection and analysis of medical images, because of its efficiency and accuracy in object detection [10-11]. Early versions of the YOLO model such as YOLOv3 could be found to use for real-time video analysis and object detection. All those models were of simple architecture, fast in their object detection and analysis speed. But these are not able to fulfil the precision and recall demands associated with medical imaging [12]. Consequently, the new versions have widened network depth and complexity for increasing model expressiveness, thus enabling the improved performance of more complex images [13-15]. Abdusalomov et al. [13] used YOLOv7 for the detection of the brain tumor of the MRI images. They reported that the model not only detects the presence of the brain tumor but is also able to identify the exact positions of the tumor on the MRI images. Almufareh et al. [14] used YOLOv5 and YOLOv7 models for the segmentation and classification phases of the brain tumors of the MRI images. Both models of YOLOv5 and YOLOv7 worked magnificently on the segmentation and exact classification of the type of tumor. Muksimova et al. [15] proposed a lightweight attention-driven YOLOv5m model for significant improvement of brain tumors detection in MRI images by integrating the Enhanced Spatial Attention (ESA) layer into the YOLOv5m framework. The above mentioned studies demonstrate the effectiveness of various YOLO versions in improving tumor detection accuracy in medical imaging by leveraging various properties of YOLO. Therefore, the YOLO models are effective tools for the improved analysis of large scale medical images, and further advancements in this regard will have a significant impact on patient diagnosis and effective treatment.

Although numerous studies were reported for detection and classification of brain tumor using deep learning, the use of YOLOv12 for this task has been, to our best knowledge, scarcely reported in the literature. Further, the use of deep learning for distinguishing between brain tumor and non-tumor (healthy) has not been thoroughly examined. To address these gaps, this paper proposes a YOLOv12 based deep learning model for real-time detection of brain tumors at an early stage and accurately finding their location in medical CT images. The study also aims to compare its performance with other existing YOLO models to the medical CT images. The rest of this paper is organized as follows: Section 3 presents the detailed of the YOLOv12 based deep learning model. Experimental results and related discussions are given in Section 4 and Section 5 respectively. The concluding remarks are given in Section 6.

III. YOLOv12 Based Deep Learning Model

A. Model Architecture

The YOLOv12 is first introduced by Ultralytics in 2025 as an object detection model. The architecture of the YOLOv12 model is given in Figure 1. It consists of three-stage components namely, backbone, neck and head. The function of Backbone is to extract critical features from the input image. The Neck processes and refines the extracted features before sending them to the detection head. It incorporates efficient attention mechanisms into the backbone, and preserves its real-time performance. Compared to its earlier versions of the YOLO model which are heavily dependent on CNN-based architectures, YOLOv12 introduces an "area attention" module, which does partitions of the feature map to reduce the quadratic complexity of full self-attention. It also employs residual efficient layer aggregation networks (R-ELAN) to enhance feature aggregation and training stability. Further refinements such as scaled residual connections and a reduced Multi-Layer Perceptron (MLP) ratio, enable YOLOv12 to harness the benefits of attention without sacrificing its speed.

B. Data Collection and Pre-Processing

1) Data Collection: In this research, a total of 4000 images of brain CT scans were used from Kaggle's Brain Tumor Dataset [17]. Of these 4000 images, 3200 were normal scans while 800 were tumor scans. In training the model, a total of 3200 images were used, where 1680 were tumor images while 1520 were normal images. The remaining 800 images (420 tumor images and 380 normal images) were used for testing how well this model attained effectiveness. These images were of the entire brain from the bottom of the skull to top, acquired through different directions by making use of different scanning processes like axiography and spirography scans. Each image was annotated using a bounding box notation used to

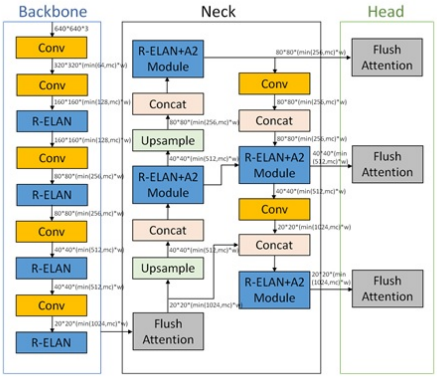

Fig. 1. Architecture of YOLOv12 model [16]

point out the presence or absence of tumors. These labels were formatted according to the demands for YOLOv12 model annotation and correlated image by image.

2) Data Pre-Processing: The images were resized to 640x640 pixels to ensure that the model processes data with a consistent input size and reduced computational cost. The Min-Max normalization technique is used for each pixel intensity between 0 and 1. Each pixel intensity value, x, is modified using the following equation:

$$x_{\text{normalized}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where x_{min} and x_{max} are the minimum and maximum values of pixels in the image respectively. This function increases the intensity as well as the contrast of the image and hence the clarity of recognition of features by the model.

C. Training and Validation

1) Training: The model was trained using the Stochastic Gradient Descent (SGD) optimization algorithm that iteratively updates a model's parameters by using a randomly selected small batch of data point at a time, instead of the entire dataset. The model was optimized to improve the detection accuracy and robustness. Table 1 gives the hyperparameters used in the training.

TABLE I
HYPERPARAMETERS USED IN THE TRAINING

Hyperparameters	Values
Epoch	100
Batch size	16
Learning rate	0.01
Momentum	0.937
Weight decay	0.0005

During the training of the YOLOv12 model for detection of brain tumor in the CT images, the loss values were determined from the weighted sum of bounding box loss (box_{loss}), classification loss (cls_{loss}) and distribution focal loss (dfl_{loss}) components to address different aspects of object detection. The training of the YOLOv12 model is performed until it minimizes this weighted total loss. The computation was carried out using an Intel Core (TM) i5- 8400 2.80GHz with YOLOv12 version, Python version 3.10.0, and Torch version 2.7.1+cpu.

2) Validation: The performance of the YOLOv12 model is evaluated by using a validation dataset which was not includedin the training set. Its performance was also compared against other YOLO models. The equations used for calculating the evaluation metrics are as follows [18]:

$$\text{AP}_i = \int_0^1 P(R)\, dR \tag{2}$$

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^{n} \text{AP}_i \tag{3}$$

where P refers to the precision rate and R denotes the recall rate, which is given by the formula:

$$P = \frac{TP}{TP + FP} \tag{4}$$

$$R = \frac{TP}{TP + FN} \tag{5}$$

where TP denotes the number of correctly detected targets, FP denotes the number of wrongly detected targets, and FN denotes the number of missed targets.

The F1-value is calculated based on the harmonic mean of the precision and recall values, which show the capability of the model, as described by the following formula:

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

IV. RESULTS

Figures 2 (a) -2(c) show the loss values for the box loss, classification loss, and distribution focal loss during training as a function of epoch. It can be observed that the values for box loss, classification loss, and distribution focal loss gradually decrease as the epoch increases, which shows that the values for all loss functions gradually improve during training. A similar trend is also observed for the loss values for validation, as can be seen from Figures 2 (d) -2(f).

From Figure 3, it can also be noted that precision, recall, and mAP values are increasing with the number of epochs. This reveals that the model is getting better at each epoch and accurately identifies the position of the brain tumors from the image. The training of the model on the dataset reveals effective detection of the brain tumor based on overall mAP and performance for the classes. The overall mAP50 (mAP at a crossing over an IoU threshold of 0.5) value for the model and the overall mAP50-95 (mAP computed across different loU thresholds ranging from 0.5 to 0.95) for the validation set were observed to be 0.905 and 0.671, respectively. This reveals that the model has the ability to accurately detect the brain tumors with high confidence.

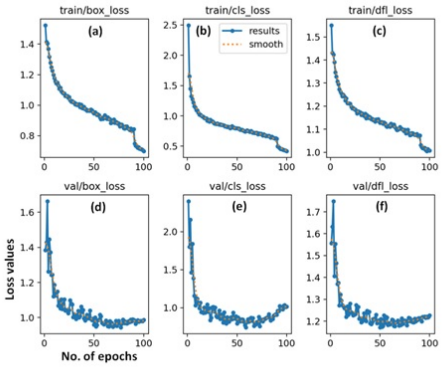

Fig. 2. Loss values as a function of epoch

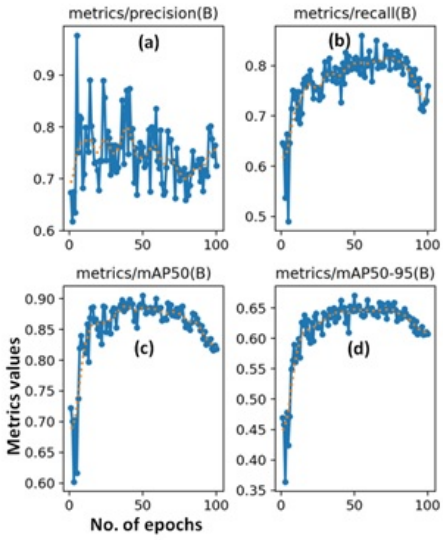

Fig. 3. Metric values of precision, recall and mAP as a function of epoch

Table 2 shows the performance evaluation of the YOLOv12 model for brain tumour detection in CT images. It can be obtained that the model shows a mAP50 of 0.905, precision of 0.761 and recall scores of 0.814 for all. Healthy class shows a mAP50 of 0.993, precision of 0.627 and recall scores of 0.998 while the tumor class shows a mAP50 of 0.818, precision of 0.895 and recall scores of 0.631. This observation reveals that the model is capable of accurate detection and classification of these classes. But the recall score for tumor class is relatively low at 0.631, which hints that it may miss some instances of this category possibly due to less number of tumor class.

The proposed model has proved to be quite effective while detecting and predicting brain tumors/healthiness. The proposed model acquired mAP50 scores of 0.993 and 0.818, respectively, for healthy and tumor situations, which verifies that it is the right choice for the detection and positioning of brain tumors in the image. The proposed model has been able to detect brain tumors with a precision of 0.895 and

TABLE II
PERFORMANCE EVALUATION OF THE YOLOV12 MODEL FOR BRAIN TUMOR DETECTION

Class	mAP@50	mAP@50–90	P	R
All	0.905	0.671	0.761	0.814
Healthy	0.993	0.818	0.627	0.998
Tumor	0.818	0.523	0.895	0.631

a value of 0.631 for recall, which verifies that it has been able to correctly identify and classify a large number of braintumors with a great deal of accuracy. The scores acquired by employing the YOLOv12 model signify that it has been able to efficiently classify and identify brain tumors in the image that we chose. Therefore, these results demonstrate the potential of the YOLOv12 model for detecting and classifying brain tumors in medical images, which could have significant implications for improving diagnosis and treatment outcomes.

Further, the performance of YOLOv12 model is analyzed with additional evaluation metrics. Figures 4-7 show the precision confidence curve, recall confidence curve, precision–recall (PR) curves and F1 score curves respectively. The precision is found to be higher with the increase in confidence level (Fig. 4). The higher precision with the increase in confidence level is because only the most confident predictions are kept, leading to fewer false positives. On the other hand, recall is found to be lower with the increase in confidence level (Fig. 5). The reason for a lower recall with an increase in the confidence level is due to the fact that only the most confident predictions are taken into account. The recall is 0.96, indicating the proportion of actual positive tumors correctly identified by the model. The PR curve demonstrates how a particular model balances between precision and recall at different levels of confidence (Fig. 6). Each point on the PR curve represents a specific precision and recall value achieved at a particular confidence threshold. Increasing the threshold generally leads to higher precision (fewer false positives) but lower recall (more missed true positives), and vice-versa. The F1 score of 0.75 indicates a well balance between precision and recall for the model's predictions (Fig. 7).

Figure 8 shows the confusion matrix for the number of true positives, true negatives, false positives, and false negatives for each class. As we can see, there is less number of false positives and false negatives which provides further insights into their classification performance. Therefore, the YOLOv12 demonstrates its effectiveness in reducing false detections, making it more reliable for clinical use.

Figure 9 shows the detection results of tumor (Column 1) and healthy (Coumns 2-4) using YOLOv12 model for different orientations of CT scans. As seen in Column 1, both the small and large sized tumors were detected with good accuracy. In Columns 2-4, the CT images are of healthy images for different orientations.

V. DISCUSSION

In the field of object detection models, accuracy alone is not a sufficient metric, due to the fact that it has not taken

TABLE III
COMPARISON OF YOLOV12 WITH OTHER YOLO MODELS

Model	mAP@50	mAP@50–90	P	R	GFLOPs	Computation Time (h)
YOLOv5	0.877	0.651	0.793	0.804	7.2	**31.310**
YOLOv8	0.873	0.648	0.747	**0.823**	**8.2**	33.642
YOLOv11	0.898	0.667	**0.795**	0.789	6.3	38.950
YOLOv12	**0.905**	**0.671**	0.761	0.814	6.5	51.103

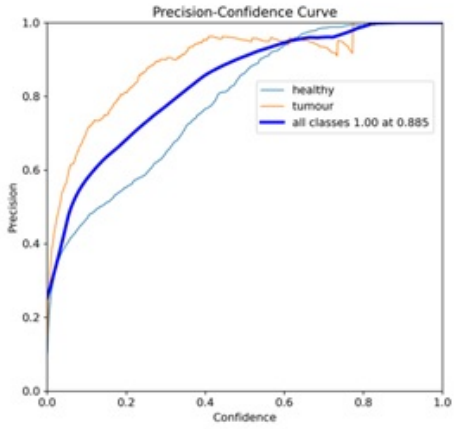

Fig. 4. Precision confidence curve

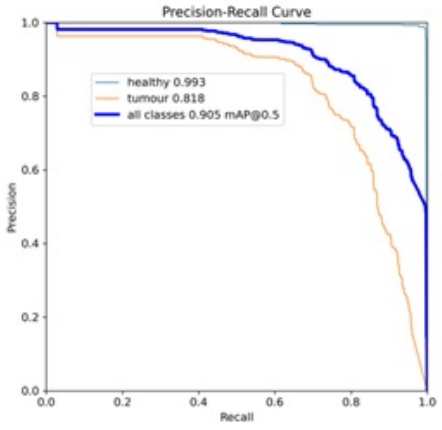

Fig. 6. Precision–recall curve

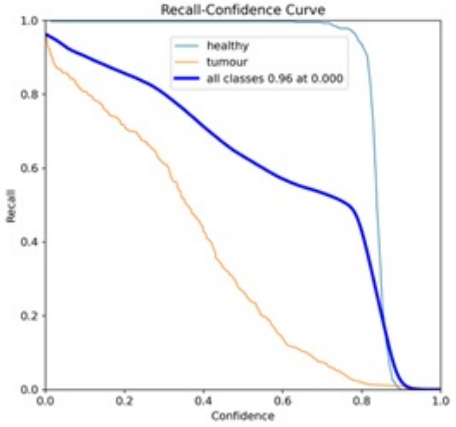

Fig. 5. Recall confidence curve

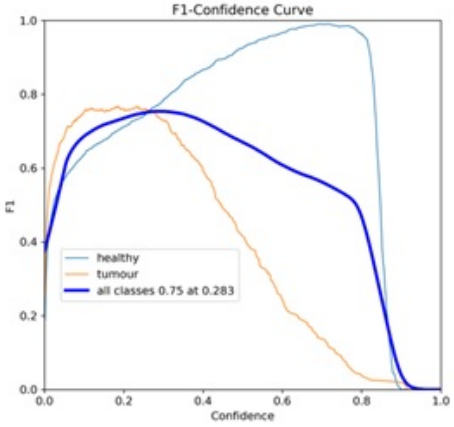

Fig. 7. F1 score curve

into account for false positives and false negatives. In this context, mean average precision (mAP) is generally used as it takes into account for both the precision and recall for all the objects at different levels of IoU thresholds. In our study, a mAP50 of 0.905 is found, indicating good performance of the model in detecting and classifying brain CT images.

We also carried out a comparative analysis of YOLOv12 with other YOLO models and are given in Table 3. It is seen from the table that YOLOv12 model outperforms compared to the other YOLO models in terms of the mAP50 and mAP50-90. These results demonstrate the potential of YOLOv12 model for accurate brain tumor detection and classification. However, among the models studied, YOLOv8 shows higher recall and Giga Floating-Point Operations per Second (GFLOPS) while YOLOv11 shows higher precision. As expected, YOLOv5 shows less computation time compared to other models, due to its simple architecture. Therefore, further improvements can be achieved by fine-tuning the hyperparameters of the model and repeated training with stable convergence.

In a future work, we plan to use diverse medical imaging datasets to strengthen the model's generalization capabilities

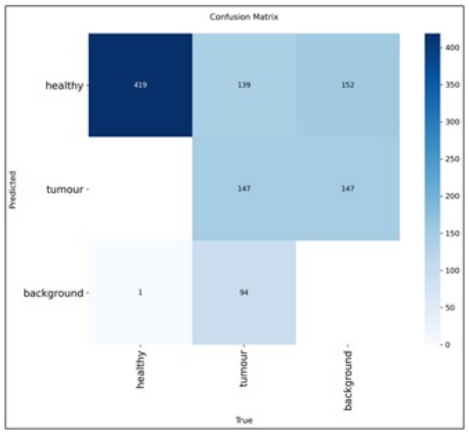

Fig. 8. Confusion matrix of the YOLOv12

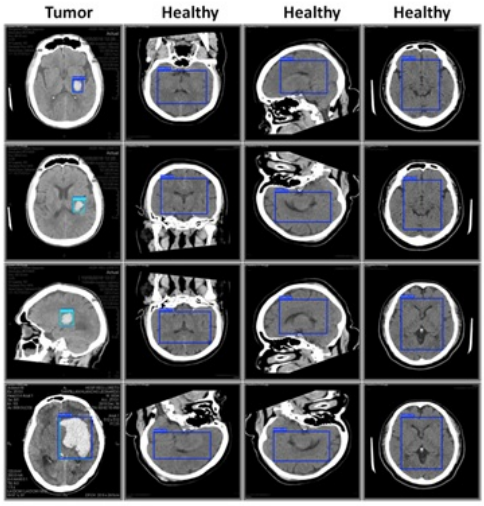

Fig. 9. Detection results of tumor (Column 1) and healthy (Columns 2-4) using YOLOv12 model

and to develop lightweight models for enhancing the reliability and computational efficiency of YOLO. Further, it will be beneficial to develop hybrid models incorporating other algorithms to the YOLO in the field of medical diagnostics.

VI. Conclusions

A deep learning model based on YOLOv12 has been proposed for detecting brain tumors in real time within CT images. The model has been optimized to enhance detection accuracy and robustness. The results revealed that the YOLOv12 model could be a great tool for automatically detecting and classifying CT images of the human brain. This has shown high accuracy in identifying images pertaining to brain tumor and healthy ones with the mAP50 of 0.905, precision of 0.761 and recall scores of 0.814. We further performed the comparison analysis of YOLOv12 with other YOLO models. Our proposed technique performed better than other YOLO models in terms of mAP50 and mAP50-90, showing the effectiveness of YOLOv12 in this aspect. Therefore, this study demonstrates the potential of YOLOv12 model for automated detection and classification of tumor and healthy in brain CT images, offering a real time diagnosis and improved patient outcomes with reduced workload of healthcare professionals and possibility of human-error.

References

[1] J. Huang, W. Ding, T. Zhong, and G. Yu, "YOLO-TumorNet: An innovative model for enhancing brain tumor detection performance," *Alexandria Engineering Journal*, vol. 119, pp. 211–221, 2025.

[2] M. A. Rahman, M. I. Masum, K. M. Hasib, M. F. Mridha, S. Alfarhood, M. Safran, and D. Che, "Glioma CNN: An effective lightweight CNN model in assessment of classifying brain tumor from magnetic resonance images using explainable AI," *CMES–Computer Modeling in Engineering & Sciences*, vol. 140, no. 3, 2024.

[3] R. Krishnan, P. G. Gokul, G. Sujith, T. Anjali, and S. Abhishek, "Enhancing brain tumor diagnosis: A CNN-based multi-class classification approach," in *Proc. IEEE Int. Conf. on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, vol. 2, 2024, pp. 1–6.

[4] N. Thota, M. Vallapuri, and B. V., "Genetic algorithm based feature selection and optimized edge detection for brain tumor detection," in *Proc. 7th Int. Conf. on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, Kolkata, India, 2023, pp. 1–5.

[5] T. Ruba, R. Tamilselvi, M. P. Beham, and N. Aparna, "Accurate classification and detection of brain cancer cells in MRI and CT images using nano contrast agents," *Biomedical and Pharmacology Journal*, vol. 13, no. 3, pp. 1227–1237, 2020.

[6] M. H. O. Alnageeb and M. H. Supriya, "Real-time brain tumour diagnoses using a novel lightweight deep learning model," *Computers in Biology and Medicine*, vol. 192, Art. no. 110242, 2025.

[7] B. Aldughayfiq, F. Ashfaq, N. Z. Jhanjhi, and M. Humayun, "YOLO-based deep learning model for pressure ulcer detection and classification," *Healthcare*, vol. 11, Art. no. 1222, 2023.

[8] M. Santos, M. Aguiar, D. Welfer, and B. Belloni, "A new approach for detecting fundus lesions using image processing and deep neural network architecture based on YOLO model," *Sensors*, vol. 22, no. 17, Art. no. 6441, 2022.

[9] K. Liu, "STBi-YOLO: A real-time object detection method for lung nodule recognition," *IEEE Access*, vol. 10, pp. 75385–75394, 2022.

[10] N. V., S. Sriram, G. K. Praadeep, and S. Prarthana, "Hierarchical bone fracture detection and classification using YOLOv9 with Grad-CAM++ visualization," in *Proc. Int. Conf. on Data Science and Business Systems (ICDSBS)*, Chennai, India, 2025, pp. 1–6.

[11] S. Sriram, A. P., A. K. T. P., N. Vijayaraj, and T. Murugan, "Enhanced YOLOv10 framework featuring DPAM and DALSM for real-time underwater object detection," *IEEE Access*, vol. 13, pp. 8691–8708, 2025.

[12] M. A. R. Alif and M. Hussain, "YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain," arXiv:2406.10139, 2024.

[13] B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Brain tumor detection based on deep learning approaches and magnetic resonance imaging," *Cancers*, vol. 15, no. 16, Art. no. 4172, 2023.

[14] M. Almufareh, M. Imran, A. Khan, M. Humayun, and M. Asim, "Automated brain tumor segmentation and classification in MRI using YOLO-based deep learning," *IEEE Access*, vol. 12, pp. 16189–16207, 2024.

[15] S. Muksimova, S. Umirzakova, S. Mardieva, N. Iskhakova, M. Sultanov, and Y. I. Cho, "A lightweight attention-driven YOLOv5m model for improved brain tumor detection," *Computers in Biology and Medicine*, vol. 188, Art. no. 109893, 2025.

[16] A. N. Chitari, S. Inamadar, and P. Salve, "Deep learning-based NIR face detection under adverse illumination with explainable AI," *Journal of Information Systems Engineering and Management*, vol. 10, no. 36s, pp. 613–625, 2025.

[17] Ultralytics, "YOLO object detection," Kaggle. [Online]. Available: https://www.kaggle.com/code/givkashi/yolov10-object-detection

[18] R. Guo, P. Ji, J. Hu, Y. Zhang, X. Li, W. Liu, M. Li, T. Xu, and Y. Jiang, "ASFM: A multi-scale feature-focused steel defect inspection model," *Engineering Research Express*, vol. 7, Art. no. 035227, 2025.