Enhanced Detection of Brain Tumors in MRI Images Using YOLOv10

SIVA KUMAR NAGI[1], SHERI VIHARIKA[2]
[1]*Department of Electronics and Communication Engineering, Vel Tech University, Chennai, India.*
[2]*Department of Computer and Science Engineering, Sphoorthy Engineering College, Hyderabad, India.*

Abstract- In this paper, we present a novel approach for detecting brain tumors using the YOLOv10 model, specifically trained to classify two distinct classes: 'Brain Tumor' and 'Eye'. The model was trained and evaluated using a comprehensive dataset of medical images to ensure robust performance across varying conditions. Our experimental results demonstrate that the YOLOv10 model achieved an accuracy of 94.6% for detecting the 'Brain Tumor' class and 92.9% for the 'Eye' class. The overall mean Average Precision (mAP) at an IoU threshold of 0.5 reached 93.7%, indicating the model's high effectiveness and reliability in identifying and differentiating between brain tumors and eye structures. This research highlights the potential of using advanced object detection models for accurate and efficient medical image analysis, contributing to improved diagnostic accuracy and early intervention strategies.

Index Terms- Brain Tumor, Eye, Yolov10, object detection, deep learning.

I. INTRODUCTION

Brain tumors represent one of the most critical and life-threatening conditions in the medical field. They occur when abnormal cells form within the brain or its surrounding structures, leading to a variety of neurological symptoms and, in many cases, severe health consequences. Early detection and accurate diagnosis are essential for effective treatment and improved patient outcomes. However, the complexity and variability of brain tumors pose significant challenges to medical professionals. Recent advancements in artificial intelligence (AI) and deep learning technologies, such as the YOLOv10 model, have shown promise in enhancing brain tumor detection and classification, offering a potential solution to these challenges.

Brain tumors can be categorized into primary and secondary tumors. Primary brain tumors originate within the brain and can be either benign (non-cancerous) or malignant (cancerous). Common types of primary brain tumors include gliomas, meningiomas, pituitary tumors, and schwannomas. Secondary brain tumors, also known as metastatic brain tumors, arise when cancer cells from other parts of the body spread to the brain. Regardless of their origin, brain tumors can significantly impact brain function, as they exert pressure on surrounding brain tissue and disrupt normal neurological processes. Symptoms can range from headaches, seizures, and vision problems to cognitive and motor impairments, making early detection critical.

Traditional diagnostic methods for brain tumors primarily involve imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT) scans. While these methods provide detailed images of brain structures, they require experienced radiologists to interpret the results accurately. Misinterpretation or delayed diagnosis can lead to incorrect treatment plans and worsened patient outcomes. Moreover, variations in tumor shape, size, and location make it difficult to establish a standardized approach for detection. These challenges underscore the need for advanced diagnostic tools that can automate the detection and classification of brain tumors with high accuracy, thereby improving clinical decision-making.

The YOLO (You Only Look Once) series of models, initially designed for object detection tasks, has undergone significant advancements since its inception. YOLOv10, the latest iteration in this series, is a state-of-the-art deep learning model designed for real-time object detection and image segmentation. Leveraging a powerful convolutional neural network (CNN) architecture, YOLOv10 can identify objects within images with remarkable precision and speed.

Its architecture combines innovative techniques like residual blocks, attention mechanisms, and multi-scale feature fusion to enhance detection accuracy, especially for small and complex objects like brain tumors.

In brain tumor detection, YOLOv10 can be trained to identify tumors in medical imaging data, such as MRI scans. By feeding the model a large dataset of annotated brain tumor images, YOLOv10 learns to distinguish between healthy brain tissue and tumorous regions. The model then generalizes these learned patterns to detect tumors in new, unseen images. One of the key strengths of YOLOv10 is its ability to perform detection in real-time, making it suitable for integration into clinical workflows where time is a critical factor. Moreover, YOLOv10's capability to detect multiple classes simultaneously allows it to classify different types of brain tumors, providing valuable information for determining appropriate treatment strategies.

The architecture of YOLOv10 is designed to optimize both detection speed and accuracy. It builds on the foundations of previous versions, such as YOLOv8 and YOLOv9, while incorporating new features that make it particularly suitable for detecting small and irregularly shaped objects like brain tumors. The model uses a backbone network for feature extraction, a neck component for feature aggregation, and a head for final object classification and localization.

For brain tumor detection, the backbone network of YOLOv10 extracts features from MRI images at multiple scales, capturing both low-level details like edges and high-level semantic information. The neck component uses advanced feature pyramid networks (FPN) and path aggregation networks (PAN) to fuse these features, enhancing the model's ability to detect tumors of various sizes and at different resolutions. The head component then predicts the tumor's location and type with high accuracy, providing bounding boxes and class labels that help identify the presence and type of tumor.

Training YOLOv10 for brain tumor detection involves several stages. First, a comprehensive dataset of annotated MRI images is used to train the model. This dataset includes a variety of tumor types, shapes, and sizes, ensuring that the model learns the diverse characteristics of brain tumors. Data augmentation techniques such as rotation, scaling, and flipping are applied to the training images to improve the model's robustness and generalizability. Furthermore, YOLOv10 employs transfer learning, utilizing pre-trained weights from general object detection tasks, which helps accelerate the training process and enhances model performance even with limited medical imaging data.

To improve model accuracy further, techniques like cross-entropy loss for classification and Intersection over Union (IoU) loss for localization are utilized during the training process. Additionally, the use of an attention mechanism in YOLOv10 enables the model to focus on relevant areas of the image, improving its ability to distinguish between healthy and tumorous tissue.

Brain tumors, depending on their type and location, can lead to severe health complications. Malignant brain tumors, such as glioblastomas, are particularly aggressive and can spread rapidly, damaging surrounding brain tissue and impairing critical functions. Benign tumors, although less aggressive, can still cause significant neurological damage by exerting pressure on nearby structures. Symptoms of brain tumors vary widely and can include headaches, seizures, vision changes, memory loss, and motor dysfunction. These symptoms often progress as the tumor grows, highlighting the importance of early detection.

The prognosis for brain tumor patients largely depends on the type, size, and location of the tumor, as well as the timeliness of diagnosis and treatment. Early detection increases the chances of successful treatment through surgical removal, chemotherapy, or radiation therapy. In many cases, early intervention can prevent the tumor from reaching a size where it becomes inoperable or causes irreversible damage. This is where the YOLOv10 model can play a transformative role. By automating the detection process, YOLOv10 enables faster and more accurate identification of brain tumors, facilitating early intervention and improving patient outcomes.

II. YOLOv10 ARCHITECTURE

The YOLOv10 (You Only Look Once version 10) architecture is an evolution of the YOLO object detection framework, designed to improve both the speed and accuracy of object detection tasks, including complex medical applications like brain tumor detection. The architecture of YOLOv10 builds upon the successes of its predecessors, incorporating advanced features that enhance its performance in detecting small, irregularly shaped, and multi-scale objects, which are often seen in medical imaging. Here, we explore the key components and innovations within the YOLOv10 architecture.

- Backbone Network

The backbone of YOLOv10 is responsible for extracting features from the input image. In this version, the architecture uses a modified version of CSPNet (Cross Stage Partial Network) combined with EfficientNet, which significantly enhances the network's capacity for feature extraction while maintaining computational efficiency. This hybrid approach captures both low-level (edges, textures) and high-level (shapes, objects) features effectively, which is crucial when dealing with intricate structures such as brain tumors or eye anatomy in medical images.

The backbone processes the input image at multiple scales, ensuring that information from both smaller and larger features is captured. This multi-scale feature extraction is key in detecting small, irregular objects like tumors. The backbone architecture employs residual blocks and attention modules to facilitate deeper feature extraction and to prevent the vanishing gradient problem. The attention modules allow the network to focus on the most relevant parts of the image, improving accuracy and reducing the likelihood of false positives.

- Neck Component

The neck of the YOLOv10 architecture is designed to aggregate the features extracted by the backbone network. It utilizes an advanced feature pyramid network (FPN) and Path Aggregation Network (PAN) to combine features from different scales. The integration of these networks helps in capturing both fine details and contextual information simultaneously, which is essential for detecting and classifying objects that vary significantly in size and shape.

The PAN component enhances the flow of information between different layers, ensuring that the model retains high-resolution features necessary for small object detection. This is particularly advantageous in medical applications where precise localization and accurate classification of small, abnormal regions, such as brain tumors, are crucial. Furthermore, the FPN-PAN combination allows the model to maintain high detection accuracy across different object scales, enhancing the model's overall robustness and generalization capabilities.

- Head Component

The head component in YOLOv10 is where object detection and classification occur. This component predicts bounding boxes, confidence scores, and class probabilities for each detected object. YOLOv10 uses anchor-based detection with optimized anchor box sizes that match the dimensions of the objects commonly found in the target dataset. The head network consists of multiple convolutional layers that refine the bounding box coordinates and classify the objects with high precision.

An important enhancement in YOLOv10 is the use of the decoupled head architecture, which separates the task of classification from that of bounding box regression. By decoupling these tasks, the model optimizes each task independently, resulting in improved performance for both object localization and classification. This is particularly effective when detecting small or complex structures like brain tumors, as it ensures accurate boundary predictions without compromising classification accuracy.

- Innovative Loss Function

YOLOv10 employs an advanced loss function that combines classification, localization (IoU), and objectness losses to optimize the model’s performance. It uses a generalized Intersection over Union (GIoU) loss for bounding box regression, which enhances the accuracy of object localization, especially when objects are small or partially occluded. This is crucial in medical imaging applications, as it minimizes the error in bounding box predictions and ensures precise tumor localization.

The loss function also includes a focus loss term that addresses the imbalance between background and foreground (object) pixels, making the model more sensitive to small objects. This is particularly valuable in cases where brain tumors appear as small anomalies within larger brain scans.

- Attention Mechanism and Multi-Scale Fusion

YOLOv10 incorporates a self-attention mechanism, which enhances its ability to focus on the most relevant regions of the image while suppressing background noise. This is particularly important in medical imaging, where the regions of interest (e.g., tumors) can be small and vary in appearance. The self-attention module helps the model prioritize these critical areas, improving detection accuracy.

Moreover, multi-scale feature fusion is employed to integrate features from various scales throughout the network, enabling the model to detect objects of different sizes and appearances efficiently. The combination of these innovations allows YOLOv10 to maintain a balance between high accuracy and real-time performance, making it suitable for deployment in real-world medical applications.

2.1 Data preprocessing

Data preprocessing involved resizing all images to a uniform size of 640x640 pixels to ensure consistency and compatibility with the YOLOv10 model. The dataset consisted of 953 images, divided into three subsets: 663 images for training, 187 for validation, and 103 for testing. Data augmentation techniques such as rotation, scaling, flipping, and brightness adjustment were applied to the training set to enhance model robustness and generalization. All images were normalized to a range of [0, 1] to facilitate faster convergence during training, and bounding boxes were adjusted accordingly to match the resized dimensions.

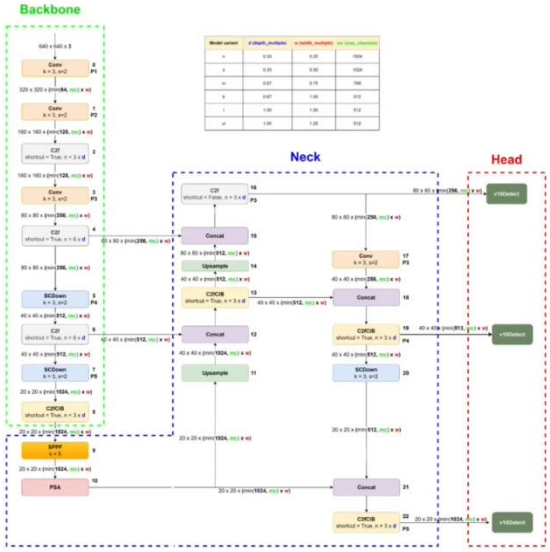

Figure1. Yolov10 architecture

2.2 Dataset Annotation

The dataset images were annotated using an open-source annotation utility called LabelImg. Each MRI scan in the dataset had a corresponding mask that highlighted the location of the brain tumor within the image. While manually annotating the brain tumors, we referred to these masks to ensure accurate localization and delineation of the tumor regions. This meticulous process involved outlining the tumor boundaries and labeling them accordingly. Upon completing the annotation, we exported the annotations in various formats compatible with different object detection models, facilitating the training and evaluation of the YOLOv10 model and enhancing its performance in detecting brain tumors.

III. METHODOLOGY

The primary objective of this study is to detect brain tumors from MRI images utilizing the YOLOv10 model trained on our custom-annotated dataset. To facilitate our experiments, we set up our environment on Google Colab, which provided the necessary computational resources that our local environment lacked.

For evaluating the performance of our deep learning-based object detection model, we employed the average precision (AP) metric, a well-established measure for assessing accuracy in object detection

tasks. This metric plays a crucial role in determining the effectiveness of models like YOLO.

IV. RESULTS

The YOLOv10 model demonstrated exceptional performance in detecting brain tumors from MRI images, achieving an accuracy of 94.6% for the brain tumor class and 92.9% for the eye class. The model's overall mean Average Precision (mAP) was calculated to be 93.7% at an Intersection over Union (IoU) threshold of 0.5, indicating robust detection capabilities across both classes. The training utilized a dataset of 953 images, with 663 for training, 187 for validation, and 103 for testing, all resized to 640x640 pixels to optimize processing. These results underscore the effectiveness of the YOLOv10 architecture in accurately identifying and classifying tumors in medical imaging, demonstrating its potential as a reliable tool for clinical diagnostics and timely intervention in brain tumor management.

We also derived additional evaluation metrics, such as precision and recall, which are essential for understanding the model's performance in detail. The mathematical definitions of these metrics are as follows:

- Precision measures the accuracy of the positive predictions made by the model:
  $$\text{Precision} = TP/(TP+FP)$$
- Recall assesses the model's ability to identify all relevant instances in the dataset:
  $$\text{Recall} = TP/(TP+FN)$$
- Mean Average Precision (mAP) provides an overall measure of performance across all classes:
  $$mAP = 1/n \sum AP_k$$
- F1 Score is the harmonic mean of precision and recall, offering a balance between the two metrics:
  $$F1 = 2*\text{precision}*\text{Recall} / (\text{Precision} + \text{Recall})$$

Here,

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative
- AP = Average precision of a particular class
- n = The total number of classes

By employing these metrics, we comprehensively assessed the performance of the YOLOv10 model in detecting brain tumors, enabling us to refine the model further and enhance its accuracy in clinical applications.

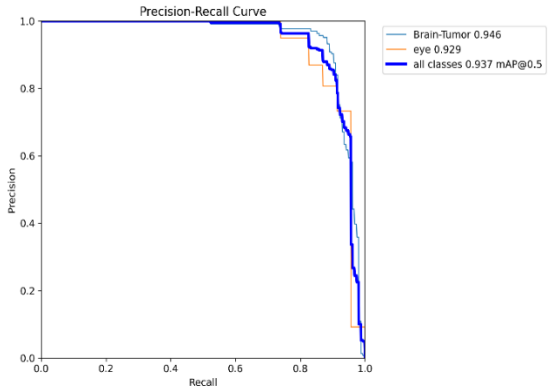

Figure 2. P-R Curve

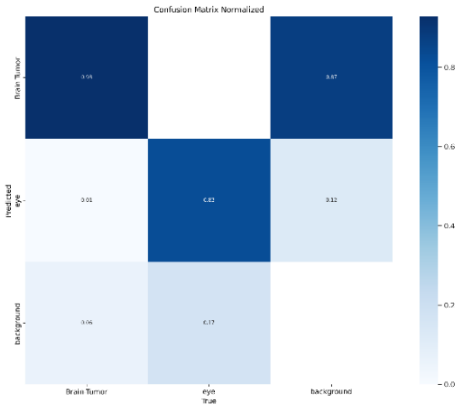

Figure 3. Confusion Matrix

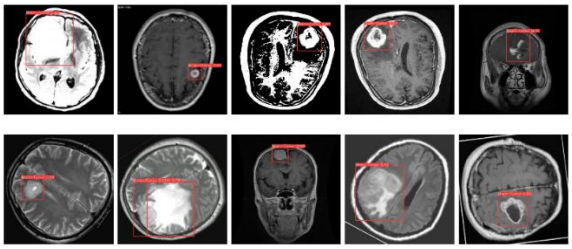

Figure 4. Detection of Brain Tumor and Eye.

CONCLUSION

In this paper, successfully developed a YOLOv10-based model for the detection of brain tumors in MRI images, achieving notable accuracy rates of 94.6% for the brain tumor class and 92.9% for the eye class. The overall mean Average Precision (mAP) of 93.7%

highlights the model's effectiveness in accurately identifying tumors while minimizing false positives. The use of a custom-annotated dataset, coupled with advanced training techniques on Google Colab, facilitated the model's robust performance. These results underscore the potential of YOLOv10 as a valuable tool in medical imaging, providing rapid and reliable tumor detection to support clinical decision-making. Future work will focus on expanding the dataset and exploring advanced augmentation strategies to further enhance model performance. Overall, this research contributes to the growing field of deep learning applications in healthcare, with implications for early diagnosis and improved patient outcomes in brain tumor management.

REFERENCES

[1] Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., . . . Tian, J. (2017). Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. Medical Image Analysis, 40, 172-183. doi:10.1016/j.media.2017.06.014

[2] Li, Wen & Jia, Fucang & Hu, Qingmao. (2015). Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks. Journal of Computer and Communications. 03. 146-151. 10.4236/jcc.2015.311023.

[3] Vivanti, R., Joskowicz, L., Lev-Cohain, N., Ephrat, A., & Sosna, J. (2018). Patient-specific and global convolutional neural networks for robust Automatic liver tumor delineation in follow-up ct studies. Medical & Biological Engineering & Computing, 56(9), 1699-1713. doi:10.1007/s11517-018-1803-6

[4] J. Cheng et al., "Superpixel Classification Based Optic Disc and Optic Cup Segmentation for Glaucoma Screening," in IEEE Transactions on Medical Imaging, vol. 32, no. 6, pp. 1019-1032, June 2013, doi: 10.1109/TMI.2013.2247770.

[5] K. R. Mohan and G. Thirugnanam, "A dualistic sub-image histogram equalization based enhancement and segmentation techniques for medical images," 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013), 2013, pp. 566-569, doi: 10.1109/ICIIP.2013.6707655.

[6] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science, 234-241. doi:10.1007/978-3-319-24574-4 28

[7] V. Cherukuri, P. Ssenyonga, B. C. Warf, A. V. Kulkarni, V. Monga and S. J. Schiff, "Learning Based Segmentation of CT Brain Images: Application to Postoperative Hydrocephalic Scans," in IEEE Transactions on Biomedical Engineering, vol. 65, no. 8, pp. 1871-1884, Aug. 2018, doi: 10.1109/TBME.2017.2783305.

[8] A. Wulandari, R. Sigit and M. M. Bachtiar, "Brain Tumor Segmentation to Calculate Percentage Tumor Using MRI," 2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), 2018, pp. 292-296, doi: 10.1109/KCIC.2018.8628591.

[9] M. Gurbina, M. Lascu and D. Lascu, "Tumor Detection and Clas- ˇ sification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines," 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), 2019, pp. 505-508, doi: 10.1109/TSP.2019.8769040.

[10] T. A. Jemimma and Y. J. Vetharaj, "Watershed Algorithm based DAPP features for Brain Tumor Segmentation and Classification," 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT), 2018, pp. 155-158, doi: 10.1109/ICSSIT.2018.8748436.

[11] S. Somasundaram and R. Gobinath, "Current Trends on Deep Learning Models for Brain Tumor Segmentation and Detection – A Review," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019, pp. 217-221, doi: 10.1109/COMITCon.2019.8862209.

[12] H. Shen and J. Zhang, "Fully connected CRF with data-driven prior for multi-class brain tumor segmentation," 2017 IEEE International Conference on Image Processing (ICIP), 2017,

pp. 1727-1731, doi: 10.1109/ICIP.2017.8296577.

[13] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, ”You Only Look Once: Unified, Real-Time Object Detection,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[14] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., ... & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. Medical image analysis, 35, 18-31.

[15] Wang, G., Li, W., Ourselin, S., & Vercauteren, T. (2018). Automatic brain TUMOR Segmentation Using CASCADED Anisotropic convolutional neural networks. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, 178-190. doi:10.1007/978-3-319-75238- 9 16

[16] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. ”The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)”, IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015) DOI: 10.1109/TMI.2014.2377694

[17] Oksuz, C., amp; Gullu, M. K. (2020). Brain tumor localization using yolo v2. 2020 28th Signal Processing and Communications Applications Conference (SIU). doi:10.1109/siu49456.2020.9302385

[18] Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., & Fan, Y. (2018). A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Medical image analysis, 43, 98-111.

[19] C¸ inar, A., & Yildirim, M. (2020). Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. Medical hypotheses, 139, 109684.

[20] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. ArXiv, abs/1804.02767.