

---

# Brain Tumor Detection

---

**Damilola Olaiya\***  
Carleton University  
Ottawa, Canada

damilolaolaiya@cmail.carleton.ca

**Hamza Cecen\***  
Carleton University  
Ottawa, Canada

hamzacecen@cmail.carleton.ca

**WooSeok Kim\***  
Carleton University  
Ottawa, Canada  
wooseokkim@cmail.carleton.ca

\*These authors contributed equally to this work.

## Abstract

Magnetic Resonance Imaging (MRI) has a central role in the detection and clinical assessment of brain tumors. Recent advancements in machine learning and object detection, especially YOLO-related frameworks, have significantly improved automated tumor localization in 2D multiplanar MRI slices. One of these, Pretrained Knowledge Guided YOLO (PK-YOLO) has demonstrated state-of-the-art performance by integrating SparK-pretrained RepViT backbone and specialized optimization techniques. However, the original PK-YOLO design exhibits notable limitations, primarily stemming from a plane-shift domain gap, with its backbone is pretrained exclusively on axial images, resulting in substantial performance disparities across planes.

To solve the limitations, we propose a 3D Multiplanar Fusion framework. In this method, three different backbones, each of which is pretrained on an anatomical plane (*i.e.*, axial, coronal, or sagittal), are used. Then the outputs of the backbones are combined using an adaptive router module that learns how much weight to give each plane.

## 1 Introduction

Spotting brain tumors matters a lot when looking at medical scans, since finding them early and correctly helps doctors figure out what’s wrong, decide on treatments, and keep track of how patients are doing. MRIs are common for this job because they show fine details and distinguish soft tissues well. It is especially useful when multi-angle MRI views are used, such as top, front, and side cuts, which give different body viewpoints, making it easier for physicians to pin down where the tumor sits and understand its shape.

Over the past few years, advances in machine learning have opened up the possibilities to integrate deep learning methods into the medical domain[1]. Throughout active research, several papers employ multiple approaches (*e.g.*, image processing, convolutional networks, attention-based networks) to improve brain tumor detection performance [3]. Of these tools, YOLO-based [4] models shine, thanks to its speed, solid results, and live detection power. However, applying YOLO-based models in medical domain is not a simple task, because MRIs vary due to its unique layouts, noise types, and hidden features.

---

\*Equal contribution.

To tackle these issues, M. Kang *et al.* [2] proposed a new You Only Look Once (YOLO)-based [4] brain tumor detection model by leveraging 2D multiplanar Magnetic Resonance Imaging (MRI) slices for model training and fine-tuning. The model employs a pre-trained RepViT backbone, fine-tuned with SparK masking, to effectively capture and transfer brain tumor-specific knowledge. What helps it stand out is how it blends features across layers while swapping standard loss with Focaler-IoU, boosting precision, especially on tiny growths. The experiments show that it achieves the state-of-the-art performance, scoring the highest metrics on all three scan types. Despite its performance, we have identified gaps in the model performance demonstrated by the plane-shift domain gap.

## 2 Limitations of the Previous Work

## 3 Related Work

Research on brain tumor detection using deep learning largely intersects with three key areas: DETR framework, brain tumor detection, and multiplanar MRI analysis. The PK-YOLO paper introduces contributions in all three areas, and the prior work they reference can be grouped into the following themes.

### 3.1 DETR Framework

Object detection approaches based on the DETection TRansformer (DETR) architecture have evolved significantly in recent years, especially through improved training strategies and backbone pretraining. Early variants such as UP-DETR introduced unsupervised pretraining for Transformers, showing that self-supervised representation learning could improve downstream detection accuracy. Subsequent works such as Group DETR v2 and Co-DETR further advanced DETR performance by exploring encoder-decoder pretraining strategies and hybrid assignments, achieving strong results on benchmarks such as MS COCO.

Lightweight or real-time DETR models have also emerged, including:

- RT-DETR, which allows flexible inference speed by varying decoder layers without retraining,
- LW-DETR, which improves efficiency using pretraining strategies,
- Saliency DETR, which introduces saliency-guided supervision to accelerate and stabilize training.

Despite these advances, the PK-YOLO paper points out that DETR-based models have not been sufficiently explored for brain tumor detection in MRI images, and prior DETR variants struggle to match YOLO-based detectors in medical settings due to computation cost and limited domain adaptation.

### 3.2 Brain Tumor Detection

While YOLO models dominate natural image detection tasks, their application to medical imaging, especially brain tumors, remains relatively limited.

Prior work includes:

- RCS-YOLO, which integrates region concentration modules to improve localization of tumors,
- BGF-YOLO, which enhances YOLOv8 using multiscale attentional feature fusion,
- Traditional YOLO variants such as YOLOv5, YOLOv8, and YOLOv9 have been used but still face challenges detecting small tumors and generalizing across multiplanar MRI slices.

These YOLO-based approaches generally perform well in speed and accuracy, but lack mechanisms to integrate domain-specific pretrained knowledge, especially for complex medical imagery.

### 3.3 Multiplanar MRI Analysis

Multiplanar analysis (axial, coronal, sagittal slices) is critical in medical imaging because tumors may appear differently across orientations.

Previous works include:

- Piantadosi et al. used ensembles of 2D CNNs for MRI tissue segmentation,
- MPS-FFA model introduced multiplane and multiscale feature fusion for Alzheimer’s disease classification,
- Studies evaluating object detection on multiplanar MRI slices (e.g., Barbato & Menga with YOLOv5m) reported low mean Average Precision, highlighting the difficulty of multiplane data.

Challenges identified in previous studies:

- Significant variation in lesion size and position across planes,
- Increased number of small lesions due to slice orientation,
- Difficulty training a single model to perform well on all three anatomical planes.

## 4 Method

here

## 5 Experiments

here

## 6 Conclusion

## References

- [1] S. Asif, Y. Wenhui, S. ur Rehman, Q. ul ain, K. Amjad, Y. Yueyang, S. Jinhai, and M. Awais. Advancements and prospects of machine learning in medical diagnostics: unveiling the future of diagnostic precision. *Archives of Computational Methods in Engineering*, 32(2):853–883, 2025.
- [2] M. Kang, F. F. Ting, R. C.-W. Phan, and C.-M. Ting. Pk-yolo: Pretrained knowledge guided yolo for brain tumor detection in multiplanar mri slices. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 3732–3741. IEEE, 2025.
- [3] N. Rasool and J. I. Bhat. Brain tumour detection using machine and deep learning: a systematic review. *Multimedia tools and applications*, 84(13):11551–11604, 2025.
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection, 2016.