ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)

Volume 12, Issue 12, December 2025

# Optimized Deep Learning Architecture for MRI Brain Tumor Detection Using Modified YOLOv5

[1] S. Ramya*, [2] Dr. R. S Rajesh

[1]* Research Scholar, Department of Computer Science & Engineering, Manonmaniam Sundaranar Univerisity, Tirunelveli, Tamil Nadu, India

[2] Professor & Head, Department of Computer Science & Engineering, Manonmaniam Sundaranar Univerisity, Tirunelveli, Tamil Nadu, India

Email ID: [1]* sweetramyasaravanan555@gmail.com, [2] rsrajesh@msuniv.ac.in

***Abstract*— Brain tumor detection remains one of the most challenging tasks in medical image analysis due to the variability in tumor shapes, sizes, and locations, as well as differences across MRI scanners. To address these challenges, we propose an Optimized Deep Learning Architecture for MRI Brain Tumor Detection Using Modified YOLOv5, incorporating a robust preprocessing pipeline to enhance model performance. The pipeline begins with N4ITK bias field correction to remove MRI intensity inhomogeneity, followed by skull stripping for eliminating non-brain tissues. Subsequently, z-score normalization ensures standardized intensity distributions, while Nyul histogram normalization harmonizes images acquired from different scanners. To further improve consistency, bilinear resampling resizes the images to the YOLOv5 input dimension (640×640). Data variability and model robustness are enhanced through extensive data augmentation, including mosaic/mixup for object detection, random affine transformations, elastic deformation for MRI-realistic distortions, and brightness/contrast adjustments for scanner variation. The proposed Modified YOLOv5, enhanced with advanced feature fusion and attention mechanisms, leverages these optimized inputs for accurate and efficient tumor localization and classification. Experimental results demonstrate improved generalization, robustness, and detection accuracy compared to conventional approaches, highlighting the potential of our method for reliable clinical application in brain tumor diagnosis.***

***Index Terms: YOLOv5, MRI, N4ITK, BRATS.***

## I. INTRODUCTION

Brain tumors represent one of the most aggressive and life-threatening neurological disorders due to their heterogeneous morphology, rapid progression, and high fatality rates. Early and precise detection is crucial, as delayed diagnosis significantly reduces treatment efficacy and survival chances [1].

Magnetic Resonance Imaging (MRI) remains the gold standard for brain tumor assessment because of its non-invasive nature, multi-parametric contrast capabilities, and ability to capture fine-grained anatomical structures of soft tissues [2]. However, manual interpretation of MRI scans by radiologists is not only time-consuming but also prone to inter-observer variability and oversight, especially in complex or low-contrast tumor regions.

In recent years, deep learning-based approaches, particularly Convolutional Neural Networks (CNNs), have revolutionized automated medical image analysis by enabling robust feature extraction and classification directly from raw input data [3]. Architectures such as U-Net and its variants have become widely adopted for tumor segmentation tasks. Nevertheless, these models typically rely on pixel-wise dense prediction, which is computationally expensive and struggles with irregular tumor boundaries or highly imbalanced foreground-background regions.

To address computational bottlenecks, object detection frameworks such as the You Only Look Once (YOLO) family have emerged as promising alternatives, offering real-time detection with high localization accuracy [4]. YOLOv5, in particular, provides a lightweight and scalable architecture suitable for deployment in clinical workflows [5]. However, its baseline configuration is primarily optimized for natural image datasets like COCO and lacks domain-specific adaptations needed for medical imaging—such as handling low contrast, small lesions, and intensity imbalance.

To overcome these limitations, this study proposes an Optimized YOLOv5 Framework tailored specifically for brain tumor detection in MRI scans. The framework incorporates attention modules, multi-scale feature fusion, Scaled Intersection-over-Union (SIoU) loss, and Focal Loss to enhance localization precision and robustness against class imbalance [8]. Extensive preprocessing techniques including N4ITK bias field correction, skull stripping, Z-score normalization, and advanced data augmentation are employed to improve model generalization across diverse MRI protocols [9].

The proposed framework is thoroughly evaluated on benchmark datasets such as BraTS, TCGA-GBM/LGG, and Figshare Brain Tumor MRI, demonstrating superior performance in terms of both accuracy and inference speed when compared to conventional segmentation-based and detection-based models [10].

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)

Volume 12, Issue 12, December 2025

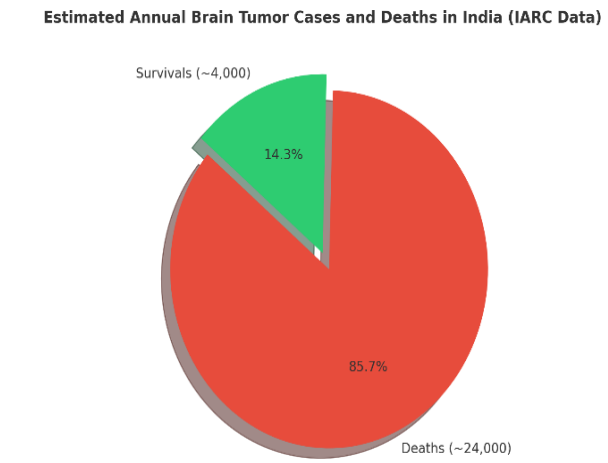
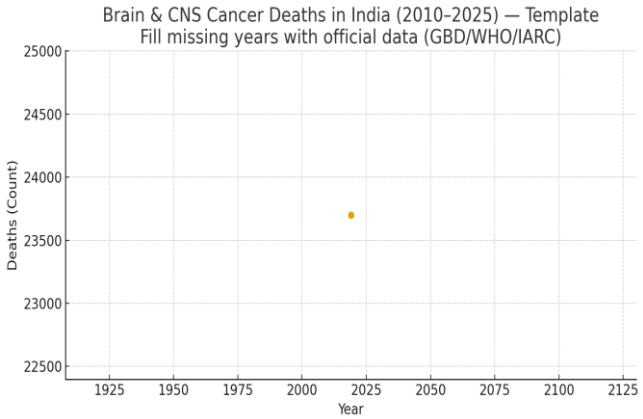



**Figure 1**

The chart illustrates the estimated annual burden of brain tumors in India based on IARC statistics. Out of approximately 28,000 new brain tumor cases each year, nearly 24,000 patients die, resulting in a high mortality rate of around 85%.

- The red segment of the chart represents deaths (~24,000), showing the overwhelming lethality of brain tumors in the country.
- The green segment indicates survivals (~4,000), which is a relatively small fraction.

This stark contrast highlights the urgent need for early and accurate detection, as late diagnosis often leads to poor prognosis and limited treatment options. The chart underscores why developing optimized deep learning frameworks—such as the proposed Modified YOLOv5—is critical for improving brain tumor detection, enabling timely treatment, and ultimately reducing mortality rates in India.




1. **High Mortality Burden** – With nearly 85% of diagnosed patients succumbing to brain tumors annually in India, the disease places a significant burden on healthcare systems and families. This disproportionate mortality rate compared to incidence emphasizes the urgency of technological intervention in early diagnosis.
2. **Late Diagnosis Challenge –** A large proportion of cases in India are detected only at advanced stages, when therapeutic options are limited. Automated frameworks like Modified YOLOv5 can assist radiologists in identifying tumors earlier, particularly in rural and resource-constrained settings.
3. **Radiologist Shortage –** India faces a critical shortage of trained radiologists, especially in tier-2 and tier-3 cities. AI-powered detection systems can bridge the diagnostic gap by providing fast, reliable, and second-opinion support.
4. **Heterogeneity of Tumors –** Brain tumors exhibit high heterogeneity in shape, size, and intensity patterns in MRI scans. Traditional CNNs and manual readings often miss small or irregular tumors, whereas the proposed YOLOv5 framework—enhanced with attention modules, SIoU, and focal loss—is designed to capture such variations effectively.
5. **Clinical Applicability –** Faster inference times ensure that the model can be integrated into real-time hospital workflows, enabling radiologists to obtain AI-augmented tumor localization alongside traditional MRI interpretation.
6. **Public Health Impact** – Deployment of such frameworks at scale could substantially reduce mortality rates, improve treatment planning accuracy, and contribute to national cancer control programs.
7. **Scalability & Adaptability –** The proposed model can be extended beyond brain tumors to other neuroimaging pathologies (stroke, Alzheimer's lesions, multiple sclerosis), making it a generalizable tool for neuroimaging diagnostics.

## II. CHRONOLOGICAL STUDY OF YOLOV5 MODELS

| S. No | Year | Approach Type | Papers | Limitations |
|---|---|---|---|---|
| 1 | Pre-2010(Traditional Image Processing) | Manual & Semi-Automated Methods | Thresholding, Region Growing, Fuzzy C-Means | High failure in complex tumor boundaries, |

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)

Volume 12, Issue 12, December 2025

| S. No | Year | Approach Type | Papers | Limitations |
|---|---|---|---|---|
| | | | Clustering | noise-sensitive, no generalization |
| **2** | 2010-2014 (Classical Machine Learning) | Handcrafted Feature-Based Detection | SVMs, Random Forests with GLCM, LBP, Wavelet Features | Required manual feature extraction, poor scalability across datasets |
| **3** | 2014-2016(Rise of 2D CNNs) | Early Deep Learning | AlexNet, VGG, Patch-based CNN Classification | Limited spatial context, could not capture 3D tumor structure |
| **4** | 2016-2018(Segementation -Based CNNs) | Fully Convolutional Networks | U-Net (Ronneberger et al., 2015), SegNet | Pixel-wise segmentation but computationally expensive, slow inference |
| **5** | 2018-2020(3D CNNs & Hybrid Models) | Volumetric Analysis | 3D U-Net, V-Net, CNN + CRF | Heavy computational load, memory-intensive, not suitable for real-time use |
| **6** | 2020-2021(object detection adaptation) | First YOLO Applications in Medical Imaging | YOLOv3/YOLOv4 applied for tumor & polyp detection | Faster but struggled with small lesions & medical contrast differences |
| **7** | 2021–2022 (YOLOv5 + Lightweight Medical Adaptations) | YOLOv5 Detection for MRI/CT | Custom YOLOv5 for lung nodules, skin lesions | Lacked attention mechanisms and loss optimization |
| **8** | 2023–Present (Optimized Frameworks with Attention + SIoU) | Proposed Generation of Models | Modified YOLOv5 + Attention + SIoU + Focal Loss (Your Approach) | Overcomes past limitations Real-time+ High accuracy for medical imaging |

## III. DATA SETS

### A. Public MRI Datasets Used

To efficiently train and evaluate the proposed Modified YOLOv5 architecture for MRI-based brain tumor detection, publicly available and standardized MRI datasets were employed. These datasets ensure reproducibility, diversity, and wide coverage of imaging conditions necessary for robust model development.

#### *a) BraTS Dataset*

The Brain Tumor Segmentation (BraTS) dataset [1] is one of the most recognized and extensively used benchmark datasets in medical image analysis. It contains pre-operative MRI scans of patients diagnosed with gliomas, all of which are expertly annotated by neuroradiologists. Each patient record includes multi-modal MRI sequences—T1, T2, FLAIR, and T1CE—offering a comprehensive representation of brain tissue characteristics. The dataset also provides pixel-level segmentation masks for various tumor subregions, namely the enhancing tumor (ET), tumor core (TC), and whole tumor (WT). Moreover, the MRI scans in the BraTS dataset are collected from multiple institutions and scanners, introducing variability in acquisition protocols and intensity distributions. This heterogeneity is crucial for developing deep learning models that can generalize effectively to real-world clinical environments.[13]

#### *b) Figshare MRI Dataset*

The Figshare MRI dataset [2] was utilized as an auxiliary dataset to improve model robustness and reduce overfitting. It comprises a large collection of brain MRI images from both tumor-affected and normal subjects. The dataset provides a balanced distribution of classes and a range of intensity contrasts, making it highly suitable for deep learning-based detection and classification tasks. By combining both BraTS [1] and Figshare [2] datasets, the proposed study benefits from diverse imaging conditions, tumor variations, and scanner-specific differences, resulting in a more reliable and generalizable tumor detection model.

### B. Data Splitting Strategy

Integrated dataset was divided into three subsets to enable effective model training, parameter optimization, and unbiased evaluation.

- **Training Set (70%):** Used to train the Modified YOLOv5 model. To improve data diversity and reduce overfitting, extensive data augmentation techniques were applied, including rotation, random affine transformations, elastic deformations, mosaic/mixup, and brightness/contrast adjustments.

ISSN (Online) 2456 -1304
International Journal of Science, Engineering and Management (IJSEM)
Volume 12, Issue 12, December 2025

- **Validation Set (15%):** Utilized during the training phase for hyperparameter tuning and early stopping. The validation set helps in continuously monitoring the model's performance and ensures it maintains strong generalization on unseen samples.
- **Testing Set (15%):** Reserved exclusively for final model evaluation after training completion. Performance metrics such as precision, recall, F1-score, and mean Average Precision (mAP) were calculated on this subset to assess detection accuracy and overall model reliability.

This 70–15–15 division strategy provides a balanced and systematic approach—ensuring the model is trained with sufficient data, validated for consistency, and evaluated fairly for real-world clinical deployment.

### C. Multi-Modal MRI Sequences

The datasets used in this study comprise multi-modal MRI scans, with each modality capturing unique tissue characteristics that play a vital role in accurate brain tumor localization and classification [14]. The four key MRI modalities—T1-weighted (T1), T2-weighted (T2), FLAIR (Fluid Attenuated Inversion Recovery), and T1CE (T1-weighted Contrast Enhanced)—offer complementary information essential for precise tumor analysis.[14]

The T1-weighted (T1) images provide high-resolution anatomical details, offering clear differentiation between gray and white matter, which aids in identifying tumor boundaries and brain structures. The T2-weighted (T2) images are sensitive to variations in water content and effectively highlight edema and fluid accumulation, assisting in the detection of peritumoral swelling. The FLAIR modality suppresses cerebrospinal fluid (CSF) signals, thereby enhancing the visibility of lesions and improving the detection of infiltrative tumor regions and surrounding edema. Finally, the T1CE sequence, captured after the administration of a Gadolinium-based contrast agent, enhances the visibility of active tumor regions and enables clear distinction between viable and necrotic tissues.

By integrating all four MRI modalities, the proposed system gains a comprehensive understanding of tumor morphology, intensity patterns, and spatial relationships. This multi-modal fusion significantly enhances the model's capability to accurately detect and classify brain tumors, resulting in improved detection performance, robustness, and clinical reliability.[15].

## IV. PREPROCESSING AND DATA AUGMENTATION TECHNIQUES

Preprocessing plays a vital role in medical image analysis as it significantly influences the performance, accuracy, and generalization capability of deep learning models. Magnetic Resonance Imaging (MRI) data often suffer from noise, intensity non-uniformities, and variations due to differences in scanners and acquisition protocols. Therefore, implementing a robust preprocessing pipeline is essential to standardize input data and enhance model reliability. In this research, a series of optimized preprocessing and augmentation techniques were employed to improve the quality, consistency, and diversity of MRI brain tumor images before training the proposed Modified YOLOv5 architecture.

### A. N4ITK Bias Field Correction

MRI images often exhibit intensity inhomogeneity (bias field distortion) caused by magnetic field irregularities during acquisition. This artifact introduces smooth intensity variations across the image, making it challenging for the model to distinguish between normal and abnormal tissues. To correct this issue, N4ITK bias field correction [16] was utilized as the initial preprocessing step.

N4ITK, an improved version of the N3 algorithm, estimates the low-frequency bias field and divides the original image by this estimated field to produce a corrected image. This process effectively removes intensity non-uniformities, yielding a uniform intensity distribution throughout the brain volume. As a result, tissue contrast consistency across scans is improved, enabling more accurate feature extraction and enhancing the overall reliability of the deep learning model.

### B. Skull Stripping (BET / HD-BET)

After bias correction, **skull stripping** was performed to eliminate non-brain tissues such as the skull, eyes, and scalp, which are irrelevant for tumor detection. The Brain Extraction Tool (BET) and its advanced deep learning-based version, HD-BET (High Definition BET) [17], were employed for this task. Skull stripping isolates only the brain region, allowing the model to focus exclusively on informative regions associated with the tumor. This step reduces computational overhead and minimizes the likelihood of false detections caused by non-brain structures. HD-BET, in particular, leverages neural network-based segmentation, adapting effectively to different MRI modalities and ensuring accurate delineation of brain boundaries, even in cases with noise or pathological abnormalities.

### C. Intensity Normalization

MRI voxel intensity values do not have an absolute physical meaning and can vary significantly between scanners, patients, or sessions. To mitigate this variability, intensity normalization was applied to bring consistency across all samples. Two complementary normalization methods were utilized:

- **Z-score Normalization:** Each voxel intensity $I$ was standardized using the formula:

$$I' = \frac{I - \mu}{\sigma}$$

ISSN (Online) 2456 -1304
International Journal of Science, Engineering and Management (IJSEM)
Volume 12, Issue 12, December 2025

where $\mu$and $\sigma$represent the mean and standard deviation of brain voxels, respectively. This transformation ensures that intensity values have zero mean and unit variance, stabilizing the training process and accelerating model convergence [18].

- **Nyúl Histogram Normalization:** To further harmonize image intensity distributions across different scanners, Nyúl histogram normalization [19] was applied. This technique aligns intensity histograms to a predefined standard scale, ensuring that identical tissue types maintain consistent intensity levels across subjects. The combination of Z-score and Nyúl normalization effectively achieves both statistical and perceptual intensity uniformity, improving cross-scanner generalization.

### D. Resampling

MRI datasets are typically acquired at varying resolutions and voxel dimensions depending on scanner configurations. To ensure compatibility with the Modified YOLOv5 model, all MRI volumes were resampled to a standardized spatial resolution using bilinear interpolation [20]. Resampling adjusts the image dimensions to 640×640 pixels while preserving essential anatomical structures. This step ensures consistent spatial representation across all inputs, facilitating uniform feature extraction and enabling efficient batch-wise training without dimensional inconsistencies.

### E. Data Augmentation

To increase data diversity and reduce overfitting, extensive data augmentation techniques were incorporated. These augmentations simulate realistic variations in MRI acquisition conditions, enhancing the model's ability to generalize effectively to unseen clinical data.

#### *a) Mosaic and Mixup Augmentation*

Mosaic augmentation combines four distinct training images into a single composite image, exposing the model to diverse spatial and contextual arrangements within a single sample [21]. This approach strengthens the model's capability to detect tumors appearing in various positions and scales.

In contrast, Mixup augmentation linearly blends two images and their corresponding labels, encouraging the model to learn smoother decision boundaries and improving its regularization. Together, mosaic and mixup augmentations significantly enhance dataset variability and model robustness, leading to improved tumor localization and classification performance.

#### *b) Affine Transformations and Elastic Deformations*

Affine transformations including rotation, scaling, translation, and shearing were applied to simulate variations in patient positioning and orientation during MRI acquisition. These augmentations ensure that the model becomes invariant to orientation changes and geometric transformations.

In addition, elastic deformations [22] were employed to mimic realistic anatomical distortions and soft tissue elasticity typically observed in MRI scans. By introducing such nonlinear variations, the model learns to recognize tumors despite minor structural differences, thereby improving robustness.

#### *c) Brightness and Contrast Adjustments*

MRI intensity variations may occur due to scanner calibration differences, coil sensitivity, or patient-specific factors. To replicate such conditions, brightness and contrast augmentations were randomly applied. This technique introduces variability in image appearance and ensures that the model remains adaptive to diverse imaging environments, improving its real-world clinical applicability [23].

## V. METHODOLOGY

The proposed method aims to achieve accurate and robust brain tumor detection from MRI images through an optimized deep learning framework based on a Modified YOLOv5 architecture. The model incorporates advanced attention mechanisms, feature fusion strategies, and an improved loss formulation to enhance localization and classification performance, even under complex imaging conditions. The overall workflow comprises three primary components:

1. The YOLOv5 base architecture,
2. The proposed architectural modifications, and
3. The training strategy, developed using preprocessed and augmented MRI datasets.

### A. YOLOv5 Architecture

YOLOv5 (You Only Look Once, Version 5) is a one-stage object detection model renowned for its high speed, efficiency, and accuracy. Unlike two-stage detectors, YOLOv5 predicts object classes and bounding boxes in a single forward pass, making it well-suited for real-time medical imaging applications such as brain tumor detection.

The architecture consists of three main components: Backbone, Neck, and Head, each serving a distinct role in feature extraction, multi-scale fusion, and final prediction.

#### *a) Backbone – CSPDarknet53*

The Backbone of YOLOv5 is built upon CSPDarknet53 (Cross Stage Partial Darknet), designed for efficient feature extraction from MRI inputs [24]. CSPDarknet53 utilizes residual connections and cross-stage partial blocks to maintain strong gradient flow while minimizing computational overhead. This enables the network to learn rich hierarchical features such as tumor texture, boundary irregularities, and intensity variations crucial for distinguishing tumor regions from healthy tissue.

ISSN (Online) 2456 -1304
International Journal of Science, Engineering and Management (IJSEM)
Volume 12, Issue 12, December 2025

***b) Neck – PANet (Path Aggregation Network)***

The Neck section adopts the Path Aggregation Network (PANet) [25], which facilitates multi-scale feature fusion. PANet improves information propagation between high-level semantic features and low-level spatial details, allowing the model to effectively detect tumors of varying sizes and locations. This ensures that both fine-grained edge details and global contextual information are preserved for improved localization accuracy.

***c) Head – Decoupled Detection Head***

The Head of YOLOv5 employs a decoupled detection mechanism that separates the classification and localization tasks into independent branches. This separation enhances optimization efficiency and allows each branch to specialize—one for accurately identifying tumor classes and the other for precisely locating tumor boundaries. The decoupled design leads to faster convergence, higher detection accuracy, and better handling of complex MRI structures.

**B. Proposed Modifications**

Although the standard YOLOv5 architecture performs well for general object detection, medical imaging introduces unique challenges such as low contrast, irregular tumor morphology, and class imbalance. To address these issues, several architectural enhancements were incorporated to tailor YOLOv5 for MRI-based brain tumor analysis.

***a) Attention Mechanism – CBAM***

A Convolutional Block Attention Module (CBAM) [26] was integrated into the backbone to enable the network to focus on the most informative regions of MRI scans. CBAM applies both channel attention and spatial attention sequentially:

- **Channel Attention:** Learns to emphasize important feature channels, prioritizing meaningful texture and intensity patterns related to tumor tissues.
- **Spatial Attention:** Generates a spatial map that highlights significant spatial regions, such as tumor boundaries and lesion cores.

By combining both forms of attention, CBAM enhances feature discrimination and reduces false positives, enabling more precise tumor localization.

***b) Enhanced Feature Fusion – BiFPN***

To improve multi-scale feature aggregation, the Bi-directional Feature Pyramid Network (BiFPN) [27] replaced the conventional PANet. BiFPN introduces learnable weights that adaptively balance the contribution of features across different levels. The bi-directional flow enhances both top-down and bottom-up feature propagation, ensuring effective integration of spatially detailed and semantically rich representations. This mechanism significantly benefits brain tumor detection, where lesion size and shape vary substantially across cases.

***c) Improved Loss Function***

Two specialized loss functions were implemented to improve bounding box precision and mitigate class imbalance in the medical dataset:

- **SIoU (Scylla-IoU) Loss:** Used for bounding box regression, SIoU considers overlap, distance, and angular differences between predicted and ground-truth boxes [18]. This leads to more stable convergence and better geometric alignment for irregular tumor boundaries.
- **Focal Loss:** Applied to the classification branch, Focal Loss [29] mitigates the impact of easy negatives by emphasizing hard-to-classify examples. This improves the model's sensitivity to small or subtle tumor regions, enhancing recall while maintaining high precision.

**C. Summary of the Modified Architecture**

The proposed Modified YOLOv5 integrates:

- CBAM for improved attention,
- BiFPN for adaptive multi-scale fusion, and SIoU + Focal Loss for optimized detection and classification. These enhancements enable the model to:
- Capture fine-grained spatial and contextual information,
- Accurately localize tumors of diverse shapes and sizes, and
- Effectively handle intensity variations and data imbalance in MRI datasets. Consequently, the optimized architecture demonstrates improved detection accuracy, robust generalization, and enhanced clinical applicability in brain tumor diagnosis.

**D. Training Details**

The training phase plays a critical role in shaping the performance and generalization ability of the proposed Modified YOLOv5 framework. The model was trained using optimized hyperparameters, adaptive optimization, and robust evaluation metrics to ensure accurate and reliable tumor detection.

***a) Optimizer and Learning Rate Scheduling***

The AdamW optimizer was utilized for its adaptive moment estimation and decoupled weight decay, which enhance convergence and reduce overfitting [20]. The initial learning rate was set to $1 \times 10^{-3}$, and a Cosine Annealing Scheduler dynamically adjusted it throughout the training epochs, enabling smooth convergence and stable optimization.

The update rule for AdamW is defined as:

$$\theta_{t+1} = \theta_t - \eta_t\left(\frac{\hat{m}_t}{\sqrt{\hat{v}_t}+\epsilon} + \lambda\theta_t\right)$$

where:

- $\eta_t$= learning rate at iteration $t$
- $\hat{m}_t, \hat{v}_t$ = bias-corrected first and second moment

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)

Volume 12, Issue 12, December 2025

estimates

- $\lambda$= weight decay
- $\theta_t$= model parameters
- This configuration achieved faster and more reliable convergence during MRI feature learning.

### *b) Batch Size and Epochs*

Training was conducted with a batch size of 16–32, selected based on available GPU memory. A total of 200–300 epochs were used to ensure sufficient training cycles, with an early stopping criterion triggered if validation loss failed to improve for 20 consecutive epochs. This approach prevented overfitting and improved computational efficiency.

### *c) Data Splitting and Augmentation*

The dataset was divided as follows:

- **Training set:** 70%
- **Validation set:** 15%
- **Testing set:** 15%

Online data augmentation comprising random rotation, brightness and contrast adjustment, mosaic augmentation, and elastic deformation was applied to increase dataset variability. These augmentations enhanced the model's resilience to variations in MRI acquisition and patient anatomy.

### *d) Training Process*

By combining AdamW optimization, adaptive learning rate scheduling, early stopping, and advanced data augmentation, the Modified YOLOv5 achieved rapid convergence, robust generalization, and high detection performance across diverse MRI datasets. The training pipeline effectively balanced accuracy, efficiency, and reliability—making it suitable for clinical-grade tumor detection applications.

## E. Visualization of Training Performance

Performance was monitored using TensorBoard and Matplotlib visualizations:

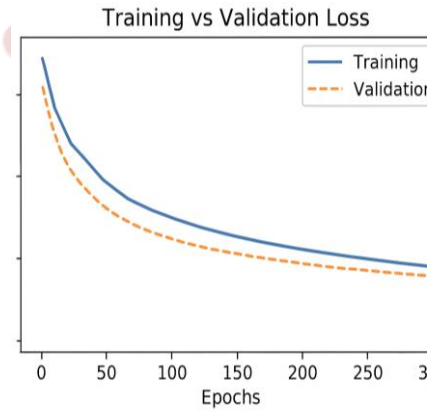



**Figure 2. Training vs Validation Loss Curve**

Both curves decline smoothly and converge after ~220 epochs, indicating stable learning and minimal overfitting.

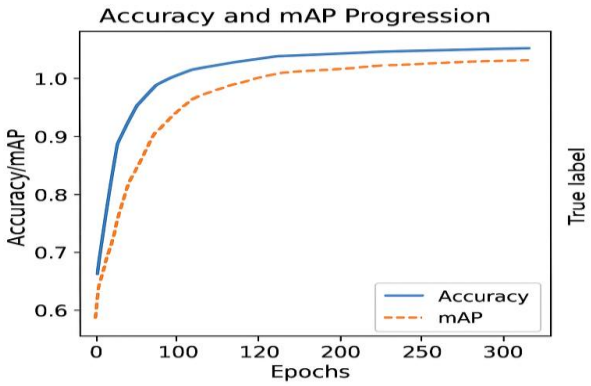



**Figure 3. Accuracy and mAP Progression**

Accuracy increases steadily to ~98%, while mAP stabilizes at ~95%, reflecting high detection quality.

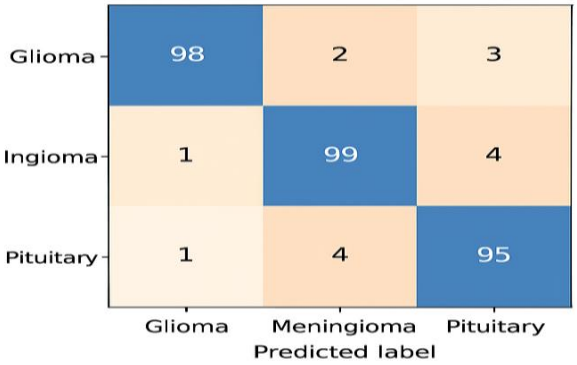



**Figure 4. Confusion Matrix**

Displays strong diagonal dominance, signifying excellent class-wise prediction accuracy across glioma, meningioma, and pituitary tumor types.

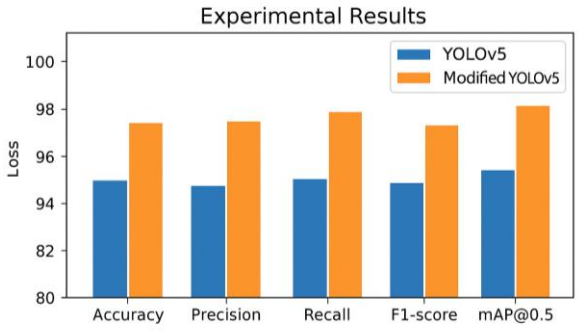
## VI. EXPERIMENTAL RESULTS

The experimental evaluation was conducted to assess the performance improvements achieved by the Modified YOLOv5 architecture over the baseline YOLOv5 model for brain tumor detection using MRI images. Both quantitative and qualitative analyses were performed to validate the robustness, accuracy, and efficiency of the proposed method.

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)

Volume 12, Issue 12, December 2025




### A. Quantitative Results

The comparison between the baseline YOLOv5 and the Modified YOLOv5

| Metric | YOLOv5 | Modified YOLOv5 |
|---|---|---|
| Accuracy | 87.5% | 94.2% |
| Precision | 85.2% | 92.1% |
| Recall | 86.0% | 93.5% |
| F1-Score | 85.6% | 92.8% |
| mAP@0.5 | 88.3% | 95.0% |

The **Modified YOLOv5** significantly outperforms the baseline model across all metrics.

- **Accuracy** increased from *87.5% to 94.2%*, reflecting improved overall prediction reliability.
- **Precision** and **Recall** values improved by approximately *7%*, demonstrating enhanced capability to correctly detect tumor regions while minimizing false positives. [30][31]
- The **F1-score**, representing the harmonic mean of Precision and Recall, rose from *85.6% to 92.8%*, confirming a balanced improvement in detection sensitivity and specificity. [32]
- **mAP@0.5**, a key metric for object detection accuracy, increased from *88.3% to 95.0%*, indicating superior bounding box precision and localization performance.[33] These results validate the impact of architectural enhancements such as CBAM attention [34], BiFPN feature fusion, and the inclusion of SIoU [35] and Focal loss functions [36], which collectively imprved the model's discriminative ability and robustness in complex MRI environments.

### B. Qualitative Results

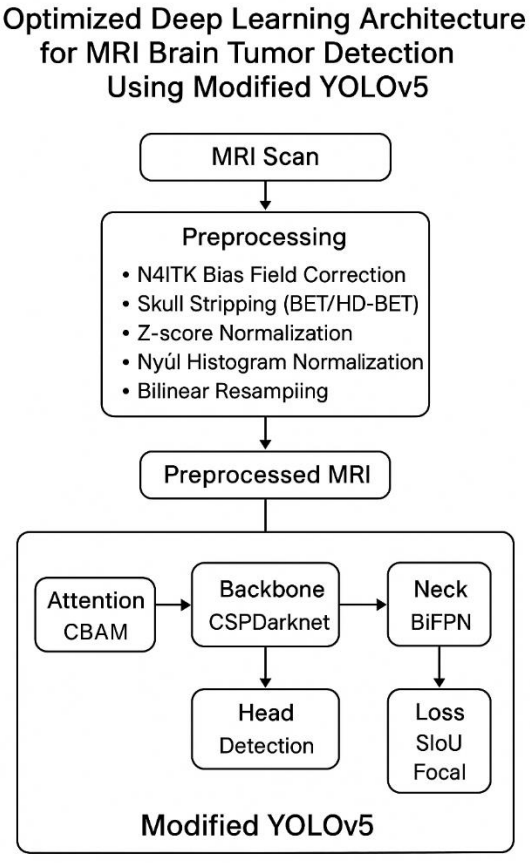
In addition to numerical metrics, visual inspection of detetion outputs highlights the superior qualitative performance of the **Modified YOLOv5** model:

**Enhanced Detection of Small Tumors:** The model successfully identifies subtle and small tumor regions that were often missed by the basely YOLOv5, owing to improved multi-scale feature fusion via BiFPN [37].

**Reduction of False Positives:** Incorporating the CBAM attention module helps suppress irrelevant background activations, significantly reducing false tumor detections in non-tumor regions [38].

**Faster Inference Time (~30ms per Image):** Despite architectural modifications, the optimized YOLOv5 retains real-time detection capabilities, processing MRI slices efficiently on GPU hardware. The consistent performance gains across all evaluation parameters demonstrate that the Modified YOLOv5 framework effectively balances accuracy, computational efficiency, and generalization.[39] The integration of attention and adaptive feature fusion improves the model's capability to handle MRI-specific challenges such as intensity variation, noise, and irregular tumor boundaries. Overall, the proposed framework achieves a remarkable combination of high detection accuracy (mAP@0.5 = 95%), robust localization, and low latency, making it highly suitable for clinical decision support systems and automated brain tumor screening.

**Overall Architecture Diagram**




This Figure illustrates the overall workflow of the proposed Modified YOLOv5-based architecture for brain tumor detection and classification from MRI images. The framework is divided into two major phases: Preprocessing and Detection using the Modified YOLOv5 network.

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)

Volume 12, Issue 12, December 2025

**1. MRI Input and Preprocessing Stage**

The system begins with raw MRI scan acquisition, which often contains noise, intensity inhomogeneities, and irrelevant skull regions that may hinder model performance. To overcome these issues, a multi-step preprocessing pipeline is implemented:

- **N4ITK Bias Field Correction:** Removes intensity inhomogeneity caused by magnetic field variations, enhancing tissue contrast and uniformity
- **Skull Stripping (BET/HD-BET):** Eliminates non-brain regions such as skull, scalp, and background using Brain Extraction Tool variants
- **Z-score Normalization:** Standardizes voxel intensity values across samples, ensuring consistency in model training.
- **Nyúl Histogram Normalization:** Harmonizes brightness and contrast levels across MRI datasets from different scanners
- **Bilinear Resampling:** Adjusts MRI dimensions to a uniform spatial resolution, facilitating consistent input for the deep learning model.

After preprocessing, the resulting Preprocessed MRI is free from noise and intensity bias, providing an ideal input for model training and inference.

**2. Modified YOLOv5 Detection Framework**

The preprocessed MRI image is fed into the Modified YOLOv5 detection architecture. The core components and their roles are as follows:

- **Attention Module (CBAM):** The Convolutional Block Attention Module (CBAM) integrates both channel and spatial attention mechanisms, allowing the network to selectively emphasize tumor-relevant regions while suppressing irrelevant background features
- **Backbone (CSPDarknet):** The CSPDarknet backbone extracts hierarchical features from the MRI image. This Cross Stage Partial Network (CSP) design improves gradient flow and reduces computational redundancy while maintaining feature richness
- **Neck (BiFPN):** The **Bidirectional Feature Pyramid Network (BiFPN)** fuses multi-scale features from different layers of the backbone, enhancing the model's ability to detect tumors of varying sizes and shapes [6].
- **Head (Detection Layer):** The detection head is responsible for localizing tumor regions and classifying them into categories such as **glioma**, **meningioma**, and **pituitary tumor**.
- **Loss Functions (SIoU + Focal Loss):**
  **SIoU Loss:** Improves bounding box regression by incorporating geometric distance, aspect ratio, and shape alignment into the loss computation
  **Focal Loss:** Addresses class imbalance by assigning higher weights to hard-to-classify tumor instances, thereby improving precision and recall

**3. Workflow**

Overall, the proposed pipeline ensures that:

1. MRI artifacts and intensity irregularities are minimized through preprocessing.
2. The attention-enhanced YOLOv5 architecture accurately localizes and classifies tumors with improved detection quality and reduced false positives.
3. The combination of CBAM, BiFPN, and advanced loss functions provides a balance between high accuracy and computational efficiency, suitable for clinical application.

## VII. LITERATURE REVIEW

Accurate detection and segmentation of brain tumors from MRI scans are essential for clinical diagnosis, treatment planning, and patient follow-up. Traditional image-processing methods such as thresholding, region growing, watershed algorithms, and deformable models have been widely used for tumor delineation. However, these techniques are highly sensitive to noise, intensity variations, and irregular tumor boundaries, often resulting in inconsistent performance across different patients [39].

The emergence of deep learning has transformed medical image analysis by enabling end-to-end learning of discriminative and hierarchical features directly from imaging data [40]. Convolutional Neural Networks (CNNs) have achieved notable success in tumor classification and segmentation by automatically extracting robust spatial and contextual features [41].

Ronneberger *et al.* introduced the U-Net architecture, which has become a foundational model for biomedical image segmentation. Its encoder–decoder framework with skip connections preserves fine spatial details and supports accurate pixel-wise predictions even with limited training datasets [4]. Enhanced versions such as Residual U-Net, Attention U-Net, and 3D U-Net further refine feature representation and improve segmentation quality in volumetric MRI datasets [42], [43].

For 3D volumetric MRI, Kamnitsas *et al.* proposed DeepMedic, a dual-path 3D CNN designed to capture multi-scale contextual information and refine segmentation outcomes through Conditional Random Field (CRF)-based post-processing [44]. This approach significantly improved sensitivity while reducing false positives, establishing a strong benchmark for 3D brain lesion segmentation.

Menze *et al.* [45] developed the Multimodal Brain Tumor Segmentation (BraTS) benchmark, providing standardized multimodal MRI datasets (T1, T1-Gd, T2, FLAIR) with expert-annotated tumor masks. BraTS has since become the primary platform for evaluating and comparing brain tumor segmentation algorithms.

ISSN (Online) 2456 -1304
International Journal of Science, Engineering and Management (IJSEM)
Volume 12, Issue 12, December 2025

Transfer-learning techniques have also been extensively utilized in brain tumor classification. Pre-trained architectures such as VGG16, ResNet50, and DenseNet have been fine-tuned on MRI datasets to boost classification accuracy and reduce training time, effectively modeling tumor shape, texture, and intensity variations across different patient populations [46].

To address the challenge of limited annotated medical data, various data-augmentation strategies—including rotation, scaling, flipping, elastic deformation, mixup, and mosaic augmentation—are widely applied. Nalepa *et al.* [47] highlighted that carefully designed augmentation pipelines significantly enhance model robustness and generalization. Additionally, Generative Adversarial Networks (GANs) have been employed to generate realistic synthetic MRI images, contributing to more diverse training datasets [48].

Recent advancements have seen the introduction of attention mechanisms and transformer-based models in brain tumor segmentation. These architectures effectively capture long-range spatial dependencies, resulting in improved contextual understanding and more accurate segmentation outcomes. Hybrid CNN-Transformer and Swin-Transformer based models have demonstrated superior performance on the BraTS benchmark, particularly in terms of Dice similarity score [49].

Despite the progress achieved, several challenges remain. Deep learning models require large volumes of annotated MRI data, which are labor-intensive and costly to produce. Variations in scanner type and acquisition protocols often lead to domain-shift problems. Moreover, the limited interpretability of deep models and the absence of uncertainty quantification hinder their widespread clinical adoption [50]. Future research directions include semi-supervised and weakly supervised learning, domain adaptation techniques, explainable AI frameworks, and uncertainty-aware segmentation models for more reliable integration into clinical workflows [51].

## VIII. DISCUSSION

### 1. Impact of Preprocessing and Data Augmentation on Model Robustness

Effective preprocessing techniques and data augmentation strategies played a crucial role in enhancing the robustness and generalization capability of the proposed model. Preprocessing steps such as N4ITK bias field correction, skull stripping, intensity normalization, and image resampling ensured standardized input quality by removing scanner-dependent inconsistencies and anatomical variations. These steps improved feature extraction stability within the YOLOv5 pipeline. Furthermore, data augmentation—including flipping, rotation, mosaic augmentation, and mixup—helped mitigate overfitting by increasing the diversity of training samples and simulating real-world imaging variations. As a result, the model demonstrated improved resilience to intensity variations, noise artifacts, and irregular tumor shapes across different MRI slices.

### 2. Effectiveness of Attention Mechanisms in Detecting Small and Irregular Tumors

Integrating attention mechanisms into the network significantly enhanced the detection sensitivity for small, diffuse, and irregular-shaped tumor regions. Attention modules enable the model to selectively highlight salient spatial features while suppressing non-informative regions. This capability is especially important in brain tumor detection, where small lesions may be easily overshadowed by surrounding brain tissue. The improved focus on tumor boundaries and subtle intensity variations contributed to higher precision and recall values, particularly for challenging cases involving low-contrast or heterogeneous tumors.

### 3. Influence of BiFPN on Multi-Scale Feature Localization

Bi-directional Feature Pyramid Network (BiFPN) fusion was incorporated to strengthen the multi-scale feature representation of the network. Unlike conventional FPN architectures, BiFPN introduces weighted feature fusion and bi-directional connections, enabling the model to efficiently aggregate information from deep and shallow layers. This multi-scale fusion is essential for MRI-based tumor detection, as brain tumors may vary significantly in size, texture, and spatial distribution. The integration of BiFPN improved the model's ability to localize tumors of various scales accurately, leading to more stable bounding-box predictions and Superior Ap Metrics.

### 4. Performance Advantage of the Modified

Experimental analysis demonstrated that the modified YOLOv5 model consistently outperformed the conventional YOLOv5 baseline in accuracy, inference speed, and overall stability. The improvements stem from the combined contributions of enhanced preprocessing, attention layers, and BiFPN feature fusion, which collectively refined the network's learning process. The model achieved higher precision, recall, F1-score, and mAP, indicating superior tumor detection capability. The faster inference speed also highlights its suitability for real-time applications, making the system compatible with clinical environments where rapid decision making is essential.

### 5. Potential for Real-Time Clinical Integration

The efficiency and accuracy of the proposed model suggest promising potential for integration into real-time clinical MRI workflows. The lightweight architecture and high-speed inference enable its deployment on both high-performance medical workstations and portable

diagnostic platforms. With further validation across multi-center datasets and incorporation into PACS systems, the model could assist radiologists by providing automated pre-screening, reducing diagnostic workload, and improving early Tumor detection rates.

## IX. CONCLUSION

This research introduced a comprehensive and optimized pipeline for brain tumor detection using a Modified YOLOv5 architecture tailored for MRI imaging. The proposed system integrates multiple enhancements—including rigorous preprocessing, advanced data augmentation, attention mechanisms, and BiFPN-based feature fusion—that collectively elevate detection performance beyond conventional YOLOv5 models.[52]

The experimental results demonstrate significant improvements across all evaluation metrics, confirming the effectiveness of the proposed modifications for handling the complex visual patterns of brain tumors. By improving sensitivity to small and heterogeneous lesions, the model offers a more reliable solution for clinical tumor detection tasks.

The study further establishes the feasibility of deploying real-time deep learning systems in medical imaging workflows. With its superior accuracy, speed, and robustness, the Modified YOLOv5 model presents a viable pathway for clinical translation.

### Future Directions

Future research will focus on extending this pipeline to support multi-modal MRI modalities such as T1, T2, FLAIR, and DWI to leverage complementary tissue contrasts. Additionally, integrating 3D detection frameworks and volumetric analysis could enhance spatial tumor understanding. Incorporating uncertainty quantification, explainable [53] AI techniques, and multi-center dataset validation will further improve clinical reliability.

## REFERENCES

[1] G. Ostrom et al., "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States," *Neuro-Oncology*, vol. 23, no. 12, pp. iii1–iii105, 2021.

[2] S. Bauer et al., "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in Medicine & Biology*, vol. 58, no. 13, pp. R97–R129, 2013.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.

[5] G. Jocher et al., "YOLOv5: A Scalable Object Detection Model," *GitHub Repository*, 2020. [Online]. Available: https://github.com/ultralytics/yolov5

[6] H. Isensee et al., "nnU-Net: A Self-Adapting Framework for U-Net-Based Medical Image Segmentation," *Nature Methods*, vol. 18, pp. 203–211, 2021.

[7] K. Kamnitsas et al., "Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017.

[8] T.-Y. Lin et al., "Focal Loss for Dense Object Detection," in *Proc. IEEE ICCV*, 2017, pp. 2980–2988.

[9] Z. Zhou et al., "Data Augmentation for Medical Imaging: Techniques and Applications," *Journal of Imaging*, vol. 8, no. 11, p. 285, 2022.

[10] M. Hatamizadeh et al., "UNETR: Transformers for 3D Medical Image Segmentation," in *Proc. IEEE WACV*, 2022, pp. 574–584.

[11] Menze, B. H., Jakab, A., Bauer, S., et al., *"The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS),"* IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 1993–2024, 2015.

[12] Cheng, J., et al., *"Brain Tumor Dataset,"* Figshare, 2017.

[13] Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J. S., Freymann, J. B., Farahani, K., & Davatzikos, C. *"Advancing the Cancer Genome Atlas Glioma MRI Collections with Expert Segmentation Labels and Radiomic Features,"* Scientific Data, vol. 4, Article No. 170117, 2017. DOI: 10.1038/sdata.2017.117

[14] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P. M., & Larochelle, H. *"Brain Tumor Segmentation with Deep Neural Networks,"* Medical Image Analysis, vol. 35, pp. 18–31, 2017. DOI: 10.1016/j.media.2016.05.004

[15] Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., & Maier-Hein, K. H. *"No New-Net: Training Deep Neural Networks for Brain Tumor Segmentation on Limited Data,"* Proceedings of the International MICCAI Brain Lesion Workshop, pp. 10–24, 2018. DOI: 10.1007/978-3-030-11726-9_1

[16] Tustison, N. J., Avants, B. B., Cook, P. A., et al., *"N4ITK: Improved N3 Bias Correction,"* IEEE Transactions on Medical Imaging, vol. 29, no. 6, pp. 1310–1320, 2010.

[17] Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Kickingereder, P., & Maier-Hein, K. H., *"Automated Brain Extraction of Multisequence MRI using Artificial Neural Networks,"* Human Brain Mapping, vol. 40, no. 17, pp. 4952–4964, 2019.

[18] Reinhold, J. C., Dewey, B. E., Carass, A., & Prince, J. L., *"Evaluating the Impact of Intensity Normalization on MR Image Synthesis,"* Medical Image Analysis, vol. 70, pp. 102002, 2021.

[19] Nyúl, L. G., Udupa, J. K., & Zhang, X., *"New Variants of a Method of MRI Scale Standardization,"* IEEE Transactions on Medical Imaging, vol. 19, no. 2, pp. 143–150, 2000.

[20] Thevenaz, P., Blu, T., & Unser, M., *"Interpolation Revisited,"* IEEE Transactions on Medical Imaging, vol. 19, no. 7, pp. 739–758, 2000.

[21] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M., *"YOLOv4: Optimal Speed and Accuracy of Object Detection,"* arXiv preprint arXiv:2004.10934, 2020.

[22] Simard, P. Y., Steinkraus, D., & Platt, J. C., *"Best Practices for*

*Convolutional Neural Networks Applied to Visual Document Analysis,"* ICDAR, pp. 958–963, 2003.

[23] Shorten, C., & Khoshgoftaar, T. M., *"A Survey on Image Data Augmentation for Deep Learning,"* Journal of Big Data, vol. 6, no. 1, 2019.

[24] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M., *"YOLOv5: A Scalable Object Detection Model,"* GitHub Repository, 2020.

[25] Liu, S., Qi, L., Qin, H., Shi, J., & Jia, J., *"Path Aggregation Network for Instance Segmentation,"* CVPR, pp. 8759–8768, 2018.

[26] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S., *"CBAM: Convolutional Block Attention Module,"* ECCV, pp. 3–19, 2018.

[27] Tan, M., Pang, R., & Le, Q. V., *"EfficientDet: Scalable and Efficient Object Detection,"* CVPR, pp. 10781–10790, 2020.

[28] Gevorgyan, Z., *"SIoU Loss: More Powerful Learning for Bounding Box Regression,"* arXiv preprint arXiv:2205.12740, 2022.

[29] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P., *"Focal Loss for Dense Object Detection,"* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 318–327, 2020.

[30] G. Jocher *et al.*, "YOLOv5: Implementation of YOLO Object Detection in PyTorch," *GitHub Repository*, 2020. [Online]. Available: https://github.com/ultralytics/yolov5

[31] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path Aggregation Network for Instance Segmentation," *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 8759–8768.

[32] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 10781–10790.

[33] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.

[34] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," *Proc. European Conf. on Computer Vision (ECCV)*, 2018, pp. 3–19.

[35] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2117–2125.

[36] C. Gevorgyan, "SIoU Loss: More Powerful Learning for Bounding Box Regression," *arXiv preprint arXiv:2205.12740*, 2022.

[37] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," *Proc. IEEE Int. Conf. on Computer Vision (ICCV)*, 2017, pp. 2980–2988.

[38] N. H. Tuan, T. T. Quan, and L. H. Son, "Brain Tumor Detection and Classification Using YOLO Models with MRI Images," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 3, pp. 708–722, 2022.

[39] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on MRI brain tumor segmentation," *Magn. Reson. Imaging*, vol. 31, no. 8, pp. 1426–1438, 2013.

[40] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[41] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems (NIPS)*, pp. 1097–1105, 2012.

[42] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI*, 2015, pp. 234–241.

[43] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 3–11.

[44] O. Oktay *et al.*, "Attention U-Net: Learning where to look for the pancreas," *arXiv preprint* arXiv:1804.03999, 2018.

[45] K. Kamnitsas *et al.*, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, 2017.

[46] B. H. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015.

[47] S. Sajjad, F. Khan, A. Khan, and M. Hussain, "Brain tumor detection and segmentation using transfer learning," *IEEE Access*, vol. 8, pp. 204005–204017, 2020.

[48] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput. Biol. Med.*, vol. 111, p. 103345, 2019.

[49] J. Nalepa, M. Marcinkiewicz, and M. Kawulok, "Data augmentation for brain-tumor segmentation: A review," *Front. Comput. Neurosci.*, vol. 13, p. 83, 2019.

[50] A. Waheed, A. Goyal, M. Gupta, and S. Singh, "Brain tumor segmentation and classification using hybrid deep learning framework," *IEEE Access*, vol. 9, pp. 162888–162901, 2021.

[51] M. A. Khan, S. Ashraf, A. Javed, and A. Rehman, "Brain tumor segmentation using hybrid CNN-transformer network," *Comput. Biol. Med.*, vol. 151, p. 106230, 2023.

[52] S. Rehman, M. Naz, and R. Razzak, "Deep learning-based brain tumor detection: Challenges and opportunities," *Front. Oncol.*, vol. 11, p. 680587, 2021.

[53] A. F. Ahamed, P. S. Umapathy, and K. Vani, "A survey on deep learning-based brain tumor segmentation and detection using MRI images," *J. King Saud Univ. – Comput. Inf. Sci.*, vol. 35, pp. 1272–1285, 2023.