Alexandria Engineering Journal 119 (2025) 211–221

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Original article

YOLO-TumorNet: An innovative model for enhancing brain tumor detection performance

Jian Huang [a,b], Wen Ding [a], Tiancheng Zhong [c], Gang Yu [a,b,*]

[a] National Clinical Research Center for Child Health, National Children's Regional Medical Center, Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310052, China
[b] Sino-Finland Joint AI Laboratory for Child Health of Zhejiang Province, Hangzhou 310052, China
[c] Shenzhen Jancsitech Co., Limited, Shenzhen, 518100, China

ARTICLE INFO

Keywords:
IoT in healthcare
YOLO-TumorNet
Brain tumor detection
Medical imaging
Multi-scale feature extraction

ABSTRACT

Brain tumors are high-risk conditions where early detection and precise localization are crucial for improving patient prognosis. However, existing automated detection methods still exhibit limitations in robustness within complex backgrounds, boundary recognition, and the detection of small tumors, making it challenging to meet the high precision requirements of clinical applications. To address these issues, this paper proposes an improved YOLOv10-based model, YOLO-TumorNet. Specifically, YOLO-TumorNet integrates the InceptionNeXt architecture, Multi-Scale Spatial Pyramid Attention (MSPA), and Bidirectional Feature Pyramid Network (BiFPN) modules to enhance multi-scale feature extraction and channel attention mechanisms, thereby improving the model's accuracy and robustness in brain tumor detection. Additionally, extensive experiments conducted on the Br35H and Roboflow datasets demonstrate the superior performance of YOLO-TumorNet in terms of boundary clarity, detail capture, and small tumor detection.

1. Introduction

Brain tumors, especially malignant brain tumors, are one of the leading causes of death and disability worldwide. According to the World Health Organization (WHO), brain tumors rank tenth among all cancers, with a relatively high incidence among pediatric cancers. Brain tumors not only affect the physiological health of patients but also cause a series of symptoms such as severe headaches, seizures, vision and hearing impairments, and cognitive decline, significantly impacting the quality of life. Early diagnosis of brain tumors is crucial, as the growth rate and malignancy of the tumor directly affect treatment outcomes [1]. As the tumor expands within brain tissue, it may compress or damage critical neural structures, leading to irreversible loss of neurological function. Therefore, early detection and appropriate treatment methods, such as surgical resection, radiation therapy, or chemotherapy, can greatly improve the patient's survival rate. However, due to the diversity and complexity of brain tumors, traditional diagnostic methods, such as imaging examinations and biomarker detection, often face significant challenges. Therefore, more advanced technologies, particularly the application of artificial intelligence (AI) in medicine, are required [2].

In recent years, deep learning technologies have made remarkable progress in the field of medical image analysis [3–5], particularly in the automatic detection and diagnosis of brain tumors. Artificial intelligence (AI) can automatically learn latent features from images by training on large annotated datasets, enabling efficient tasks such as image classification, detection, and segmentation. In brain tumor detection, convolutional neural network (CNN)-based deep learning algorithms have been widely applied in the analysis of brain MRI (magnetic resonance imaging) images [6,7]. CNNs [8,9], by mimicking the processing methods of the human visual system, can automatically extract multi-level, multi-scale features from complex image data, effectively identifying the shape, location, and type of brain tumors. Specific applications include automatic tumor segmentation, boundary detection, pathological classification, and the prediction of tumor growth [10,11].

YOLO (You Only Look Once) is a deep learning model widely used in real-time object detection, with its greatest advantage being the ability to perform both object detection and classification in a single forward pass, achieving high detection speed [12,13]. This makes YOLO highly valuable in brain tumor detection, especially in

* Corresponding author at: National Clinical Research Center for Child Health, National Children's Regional Medical Center, Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310052, China.
E-mail addresses: huangjian427@zju.edu.cn (J. Huang), chdw@zju.edu.cn (W. Ding), zhongtiancheng@jancsitech.net (T. Zhong), yugbme@zju.edu.cn (G. Yu).

https://doi.org/10.1016/j.aej.2025.01.062
Received 28 November 2024; Received in revised form 1 January 2025; Accepted 13 January 2025
Available online 5 February 2025

real-time monitoring and diagnostic systems. However, YOLO's application also faces some technical bottlenecks, particularly for complex and fine-grained medical imaging tasks like brain tumor detection. A notable issue is YOLO's dependence on Non-Maximum Suppression (NMS) [14]. In the YOLO model, NMS is used to select the optimal bounding box from multiple predictions, which effectively reduces duplicate detections but also introduces additional computational overhead and, to some extent, slows down inference speed. This problem becomes more pronounced in high-resolution medical images, where NMS can lead to redundant computations, thus hindering the end-to-end deployment and real-time application of YOLO. Furthermore, while the components of the YOLO model are designed to be flexible, they lack sufficient optimization and integration [15, 16]. For example, there is considerable computational redundancy between YOLO's detection head and its regression part, limiting the potential for performance improvement. In YOLOv10, researchers have proposed a YOLO training algorithm without NMS based on consistency dual assignment, combined with an efficiency-accuracy-driven overall model design strategy, which enhances both the model's computational efficiency and detection accuracy [17].

Although YOLOv10 introduces innovations and addresses some of the issues found in traditional YOLO models, it still has certain limitations [18]. First, while YOLOv10 avoids the traditional NMS step and improves detection speed, it still faces issues such as false negatives or false positives when processing more complex medical images. This is especially true when handling multimodal image data, as YOLOv10 lacks strong adaptability to different image scales and types, which may result in the loss of subtle features. Furthermore, YOLOv10's network structure continues to focus on learning local features, with insufficient integration of global contextual information. Brain tumors exhibit diverse morphologies and their growth often involves various types of tissues and organ structures [19]. Therefore, the model needs to be capable of capturing a broader range of spatial information to more accurately locate and classify tumors.

To address the inherent limitations of YOLOv10, we present a refined approach that significantly enhances brain tumor detection by reengineering the network architecture and embedding advanced attention mechanisms. Central to our method is the integration of the InceptionNeXt convolutional structure, which partitions traditional convolution operations into four parallel branches along the channel dimension. These branches incorporate small square kernels, orthogonal band kernels, and unit mappings, thereby reducing computational redundancy and simultaneously augmenting the model's representational capacity. This design achieves an optimal balance between computational efficiency and predictive performance. To further advance feature extraction and fusion, we implement a Bidirectional Feature Pyramid Network (BiFPN). This architecture facilitates efficient multi-scale feature integration through bidirectional information exchange, enabling the model to capture intricate details across varying spatial resolutions. In pursuit of higher detection accuracy, we employ the Multi-Scale Spatial Pyramid Attention (MSPA) mechanism. This approach extracts spatial features at multiple scales via hierarchical residual-like connections, allowing the model to resolve tumor structures with exceptional granularity. This enhancement markedly improves the delineation of tumor morphology and boundaries. Complementing this mechanism, the Spatial Pyramid Recalibration (SPR) module fortifies the model's robustness, particularly when confronted with the complexities of real-world brain tumor imagery. Collectively, these innovations deliver a significant leap in detection performance, underscoring the efficacy and adaptability of our proposed framework.

The contributions of this paper are as follows:

- The YOLO-TumorNet model was designed and implemented by integrating the InceptionNeXt, MSPA, and BIFPN modules into the YOLOv10 framework, significantly enhancing the accuracy and robustness of brain tumor detection in MRI images.
- YOLO-TumorNet excels in boundary clarity, detail capture, and small object detection. Through the combination of multi-scale feature extraction and attention mechanisms, the model can more accurately identify brain tumor boundaries, especially achieving high precision in detecting small tumors, which meets the requirements of medical imaging analysis for detecting subtle features.
- Extensive experiments on the Br35H and Roboflow datasets demonstrate that YOLO-TumorNet outperforms mainstream detection models across multiple metrics.

The structure of this paper is organized as follows: Section 2 reviews related work, focusing on the application of machine learning in brain tumors and research on the YOLO algorithm. Section 3 describes the proposed method, YOLO-TumorNet, and explains its implementation. Section 4 presents the experimental results, including comparative experiments and ablation studies. Finally, Section 5 summarizes the findings of this study and discusses potential future research directions.

2. Related work

In this section, we primarily discuss the application of machine learning and YOLO algorithms in brain tumor detection, drawing on a wide range of recent studies to highlight advancements and challenges in this field.

2.1. The application of machine learning in brain tumors

In the analysis of brain tumor MRI images, machine learning algorithms have been widely applied, especially in tumor detection and segmentation tasks, with different algorithms having their own strengths and weaknesses. Support Vector Machines (SVM) is a classic supervised learning method that performs classification by constructing an optimal hyperplane. It is suitable for high-dimensional data and can handle nonlinear problems. In brain tumor detection, SVM can effectively classify tumors, but due to its high computational complexity, it has longer training times on large-scale datasets and is sensitive to data labeling and feature selection, which may affect the model's generalization ability [20]. In contrast, Random Forest, an ensemble learning method, improves classification accuracy by constructing multiple decision trees. Its advantage lies in its robustness to noisy data and its ability to handle multi-dimensional features. However, despite its good classification performance, Random Forest has a relatively slow training process and poor model interpretability, which presents challenges for tumor diagnosis in medical imaging [21]. With the development of deep learning, CNNs have made significant progress in brain tumor detection. CNNs automatically learn image features and are particularly adept at handling large-scale image data, capable of recognizing the complex shapes and subtle features of tumors. CNNs are currently the mainstream method in brain tumor image analysis. Although CNNs perform excellently in terms of accuracy, their training process requires large amounts of labeled data and high computational resources. Additionally, in cases of small-sample learning or data imbalance, CNNs may not perform as well as traditional algorithms [22]. The K-Nearest Neighbors (KNN) algorithm has also been applied in some brain tumor classification tasks. KNN is a simple, instance-based learning method with an intuitive decision-making process, which can effectively handle small-scale datasets. However, it incurs high computational overhead, especially when working with large-scale image data, and its prediction speed can significantly decrease. Moreover, it is highly sensitive to noisy data, which can affect its detection accuracy. Finally, Decision Trees, as an interpretable classification model, are also widely used in brain tumor classification. The advantage of Decision Trees lies in their simplicity, transparency, and ease of understanding, making them suitable for clinical decision support. However, Decision Trees are prone to overfitting, especially when dealing with complex data, which can lead to unstable performance [23]. Therefore, pruning or ensemble methods are often needed to improve accuracy.

2.2. The application of YOLO in brain tumors

Since its introduction, the YOLO algorithm has been widely applied in various fields due to its efficient real-time object detection capability. YOLO transforms the object detection problem into a regression problem, enabling end-to-end training. This allows it to perform both classification and localization of objects in a single forward pass, a feature that shows great potential in medical image analysis, especially in brain tumor detection.

YOLOv3 is one of the earlier versions applied to brain tumor detection. By introducing multi-scale feature maps, YOLOv3 can detect tumors of different sizes, making it particularly adaptable to the diverse shapes of brain tumors. The advantage of YOLOv3 is its fast processing speed, enabling real-time medical image analysis, and compared to traditional CNNs, it requires fewer hardware resources, making it practically valuable in clinical applications. However, YOLOv3 still has some shortcomings in detecting small targets, particularly in the localization and segmentation of tiny brain tumors, and struggles with tumors in complex backgrounds, often leading to false positives [24]. With the release of YOLOv4, the performance of YOLO models was further enhanced. YOLOv4 introduced CSPDarknet53 as the backbone network and improved detection accuracy through optimizations such as the refined Anchor mechanism and the Mish activation function. In brain tumor detection, YOLOv4 showed better performance, particularly in handling noisy images, where it was more effective in suppressing background noise and improving tumor detection accuracy. Despite these improvements, YOLOv4 still exhibits slower inference speeds when processing high-resolution medical images and struggles with localizing tumor boundaries or overlapping regions accurately [25]. YOLOv5 quickly became the model of choice for many researchers due to its open-source nature and efficiency. YOLOv5 further optimized the model structure, adopting a more lightweight design, which significantly reduced computational costs while maintaining high accuracy. It is particularly suitable for real-time image analysis in brain tumor detection. However, due to its simpler model structure, YOLOv5 may show some limitations when dealing with complex brain tumor shapes or multiple tumors coexisting in an image [26]. Optimized versions such as YOLOv4-CSP and YOLOv4-tiny reduce the network size and computational load, offering better performance while maintaining high speed, making them ideal for resource-constrained devices. However, the accuracy and robustness of these versions are somewhat reduced compared to the full YOLOv4 model, particularly in detecting high-resolution images and complex cases, where they are more susceptible to occlusion and image noise [27]. Finally, YOLOv7, which further improves and optimizes YOLOv4 and YOLOv5, introduces the E-ELAN (Enhanced Efficient Layer Aggregation Network) architecture. This combines efficient feature aggregation and multi-resolution fusion, significantly improving the model's performance in small object detection and complex backgrounds. The main advantage of YOLOv7 is its balance between high accuracy and fast inference, maintaining high detection accuracy, especially in real-time analysis of complex brain tumor images [28]. YOLOv8 takes the advancements even further by improving detection accuracy, computational efficiency, and model flexibility. This version of YOLO places a strong emphasis on multi-scale feature fusion and efficient feature representation, making it particularly effective in brain tumor detection. With its refined feature fusion, YOLOv8 can more accurately identify different tumor shapes and boundaries, effectively reducing false positives and missed detections [29].

To overcome the limitations of the YOLO series algorithms in brain tumor detection, we propose an improved model based on YOLOv10, which combines the efficient inference capabilities of YOLOv10 with a refined design tailored for brain tumor imaging. To the best of our knowledge, this is the first model that integrates YOLOv10 with brain tumor detection, aiming to fully leverage YOLOv10's strengths in balancing inference speed and accuracy while addressing several key challenges in current YOLO models for brain tumor detection. Our improved model has been optimized in several aspects, including network architecture, feature fusion, attention mechanisms, and post-processing steps, all of which are customized to the characteristics of brain tumor images.

3. Method

In this section, we provide a detailed overview of the key components and architectural improvements of YOLO-TumorNet, focusing on the implementation of the Bidirectional Feature Pyramid Network, InceptionNeXt module, and Multi-Scale Spatial Pyramid Attention module.

3.1. Overview of our network

In this paper, we propose YOLO-TumorNet, with the network architecture shown in Fig. 1. The model is mainly divided into three parts: Backbone, Neck, and Head. The Backbone is responsible for feature extraction, the Neck performs multi-scale feature fusion, and the Head handles the final detection task. In this model, we combine the efficiency of YOLOv10 with a refined design tailored for brain tumor imaging. By incorporating the InceptionNeXt module, BIFPN, and MSPA modules, we enhance both detection accuracy and speed.

The Backbone section serves as the feature extraction layer of the entire model, responsible for extracting deep semantic information from the input brain tumor images. We introduce the InceptionNeXt module in the Backbone, which utilizes a multi-scale convolutional kernel design (including small square kernels, orthogonal band kernels, etc.) to effectively improve the model's expressive capability, enabling it to better capture the multi-scale features of brain tumors. The primary function of the Neck section is to effectively fuse the multi-scale features extracted by the Backbone, allowing for better detection in the Head section. In the Neck, we use the Bidirectional Feature Pyramid Network (BIFPN) structure, which employs a bidirectional information flow mechanism to fully fuse the feature layers at different scales. This enhances the model's robustness in handling multi-scale features. Additionally, the combination of the C2f module and Upsample operation further optimizes the precision of feature fusion, especially when detecting small tumors or tumors with blurred boundaries, significantly improving detection accuracy. In the Head section, we design a multi-layer detection output structure and combine it with the Multi-Scale Spatial Pyramid Attention (MSPA) module to further enhance the model's ability to focus on brain tumor regions. The v10Detect module generates detection results at multiple scales through multi-layer convolutions and concatenation operations, enabling the model to handle tumors of different sizes simultaneously. The MSPA module adapts to different scales of spatial information using a spatial pyramid structure and hierarchical residual connections, adjusting the model's focus on varying levels of spatial information. This improves the model's ability to detect brain tumors with low contrast and in complex backgrounds. Together with the Spatial Pyramid Recalibration (SPR) module, the model can more accurately distinguish between tumors and background noise regions, further optimizing detection accuracy and robustness.

3.2. Bidirectional feature pyramid network

Bidirectional Feature Pyramid Network (BIFPN) is a multi-scale feature fusion method designed to effectively combine feature information from different levels to enhance the model's ability to detect small objects and objects at different scales [30]. Traditional Feature Pyramid Networks (FPN) perform feature fusion in a top-down manner, where features are passed in a unidirectional information flow. However, BIFPN introduces bidirectional information flow, allowing bottom-up

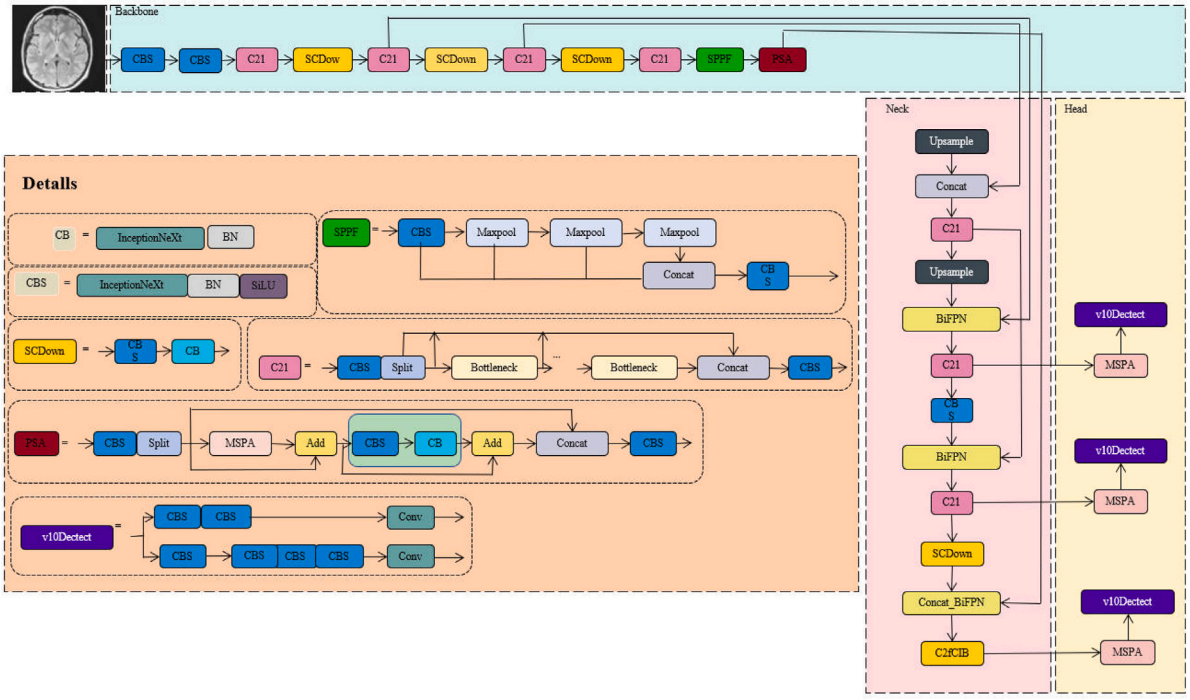

Fig. 1. YOLO-TumorNet overall network framework, including the backbone, neck, and head components.

information to also flow across layers, thereby achieving more comprehensive multi-scale feature fusion. This bidirectional feature propagation mechanism significantly improves the network's performance in detecting small objects and reduces information loss between different scales.

In the brain tumor detection task, there are significant differences in the shape and size of tumors, especially for small tumors or cases where the tumor boundaries are unclear. Single-scale features often fail to capture the complete information. By introducing BIFPN, we can fully integrate features of different resolutions, allowing the model to better capture the multi-scale information of tumors and improve detection accuracy for tumors of varying sizes and shapes. Additionally, the bidirectional information flow of BIFPN can effectively reduce information loss during the feature propagation process, enabling the model to focus more accurately on the subtle features of the tumor, which is especially crucial in complex medical imaging scenarios.

In this paper, we further optimize the bidirectional feature fusion process in BIFPN. The following equations describe the specific process of multi-scale feature fusion in BIFPN: First, the feature fusion in BIFPN begins with the bottom-up information flow, where feature aggregation is achieved by a weighted sum of features from different scales:

$$P_i^{up} = \alpha_i \cdot P_i + \beta_i \cdot P_{i+1}^{up} \tag{1}$$

where P_i^{up} represents the bottom-up feature fusion, and α_i and β_i are learnable weight coefficients that control the proportion of feature fusion.

Next, the top-down information flow is also performed by weighted summation to further strengthen the interaction between features at different scales:

$$P_i^{down} = \gamma_i \cdot P_i^{up} + \delta_i \cdot P_{i-1}^{down} \tag{2}$$

Here, γ_i and δ_i are learnable parameters in the top-down fusion process, which adjust the fusion weight between higher and lower level features.

To ensure the balance of information flow, we add a normalization operation after fusion to eliminate the imbalance of features across different scales:

$$P_i^{norm} = \frac{P_i^{up} + P_i^{down}}{\| P_i^{up} + P_i^{down} \|} \tag{3}$$

where $\| \cdot \|$ represents the normalization operation, ensuring that the fused features have a similar range of values.

Afterwards, the fused features are further processed through convolution operations to adapt to the input requirements of subsequent detection layers:

$$P_i^{final} = \sigma(W * P_i^{norm}) \tag{4}$$

Here, σ represents the activation function, typically ReLU or SiLU, and W is the convolution weight, with $*$ denoting the convolution operation.

To increase the model's sensitivity to small objects, we also introduce a scale-adaptive mechanism that dynamically adjusts the weighted coefficients of features at different scales:

$$\lambda_i = \frac{\exp(s_i)}{\sum_j \exp(s_j)} \tag{5}$$

where s_i is the weight parameter for the current scale feature, normalized using the Softmax function to ensure that each scale feature is optimally distributed in the multi-scale fusion process.

3.3. InceptionNeXt

InceptionNeXt is an efficient convolutional neural network structure based on large convolution kernels, designed to improve computational efficiency while maintaining model performance [31]. Unlike traditional deep convolution operations, InceptionNeXt decomposes large deep convolution kernels into multiple parallel convolution branches, including small square kernel convolutions, two orthogonal band kernel convolutions (in the horizontal and vertical directions), and an identity mapping branch. Through this decomposition, InceptionNeXt significantly reduces computational overhead while still maintaining a large receptive field, thus achieving a balance between performance and efficiency.

In this paper, the decomposition strategy of InceptionNeXt is particularly suitable, as brain tumor images often contain multi-scale lesion regions, and boundary and detail information require more refined feature representations. Through the multi-branch structure of InceptionNeXt, the model can reduce computational redundancy when

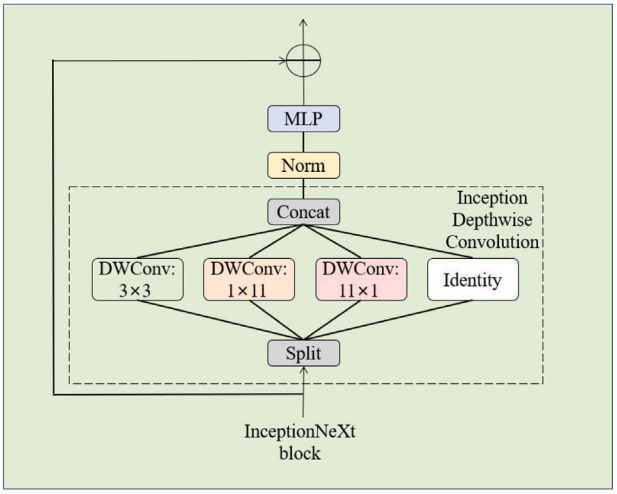

Fig. 2. InceptionNeXt network architecture diagram.

extracting multi-scale features, making the detection of small tumors and blurred boundaries more accurate. The following equations describe the operations of each convolution branch in InceptionNeXt: (see Fig. 2).

First, the input feature map X is split into four branches along the channel dimension: X_{hw}, X_w, X_h, and X_{id}, where X_{hw} represents the small square convolution branch, X_w and X_h represent the band convolution branches in the horizontal and vertical directions, and X_{id} represents the identity mapping branch:

$$X_{hw}, X_w, X_h, X_{id} = \text{Split}(X) \tag{6}$$

The small square convolution branch X_{hw} is processed by a $k \times k$ depthwise convolution operation (e.g., $k = 3$):

$$X'_{hw} = \text{DWConv}_{k\times k}(X_{hw}) \tag{7}$$

The horizontal band convolution branch X_w uses a $1\times k_b$ convolution kernel (e.g., $k_b = 11$) to expand the receptive field in the horizontal direction:

$$X'_w = \text{DWConv}_{1\times k_b}(X_w) \tag{8}$$

The vertical band convolution branch X_h uses a $k_b \times 1$ convolution kernel to expand the receptive field in the vertical direction:

$$X'_h = \text{DWConv}_{k_b\times 1}(X_h) \tag{9}$$

Finally, the feature maps from each branch are concatenated to obtain the final output feature map:

$$X' = \text{Concat}(X'_{hw}, X'_w, X'_h, X_{id}) \tag{10}$$

3.4. Multi-scale spatial pyramid attention

The Multiscale Spatial Pyramid Attention (MSPA) mechanism is an advanced attention module designed to enhance the performance of CNNs in multiscale spatial feature extraction, structural regularization, and long-range channel dependencies (Fig. 3). MSPA addresses the shortcomings of traditional attention mechanisms in capturing multiscale feature representations and establishing channel dependencies [32]. By combining the Hierarchical Phantom Convolution and Spatial Pyramid Recalibration (SPR) modules (Fig. 4), and using a Softmax operation to recalibrate the attention weights between channels, MSPA improves the model's feature representation capability. MSPA first extracts spatial information at multiple scales through the HPC module, utilizing hierarchical residual connections so that feature information from different scales can be extracted and fused with finer granularity. Given an input feature map $F \in \mathbb{R}^{C\times H\times W}$, the HPC module processes it to generate multiscale enhanced feature maps $\hat{F}$, as follows:

$$\hat{F} = \text{HPC}(F) \tag{11}$$

On this basis, the Spatial Pyramid Recalibration (SPR) module adaptively combines global and local features to generate channel attention weights. For each scale feature map $\hat{F}_i$, the SPR module generates the corresponding channel attention weight V_i:

$$V_i = \text{SPR}(\hat{F}_i) \tag{12}$$

The channel attention weights V_i from different scales are concatenated to allow for information interaction, resulting in the total attention matrix V:

$$V = \text{Concat}(V_1, V_2, \ldots, V_s) \tag{13}$$

To establish long-range channel dependencies, the attention weights for each channel are Softmax-normalized, generating the final channel attention weight S_i:

$$S_i = \text{Softmax}(V_i) = \frac{\exp(V_i)}{\sum_{j=1}^{s}\exp(V_j)} \tag{14}$$

Next, the enhanced feature maps $\hat{F}_i$ from each scale are reweighted according to the channel attention weights, producing the final output feature maps $\tilde{F}_i$:

$$\tilde{F}_i = S_i \otimes \hat{F}_i \tag{15}$$

Finally, the reweighted feature maps from all scales are concatenated to form the final output:

$$\tilde{F} = \text{Concat}(\tilde{F}_1, \tilde{F}_2, \ldots, \tilde{F}_s) \tag{16}$$

The SPR module models the inter-channel relationships on the multiscale features output by the HPC module, enhancing the model's sensitivity to information channels and improving decision quality. The SPR module uses SPA to obtain global and local context information, achieving structural regularization and rich feature representations. Meanwhile, the Channel Interaction module captures the relationships between channels, further enhancing the expressive power of the feature maps. MSPA significantly improves the feature representation and multiscale representation capabilities of the model by integrating multiscale spatial information and cross-channel attention mechanisms. This makes the model more effective in complex tasks such as brain tumor detection.

4. Experiments

4.1. Datasets

In this study, we selected two publicly available brain tumor detection datasets—the Br35H dataset [33] and the Roboflow dataset [34]. These datasets cover various types of brain tumor images, with diverse features and rich annotations, making them suitable for model training and evaluation. Fig. 5 shows samples from these datasets.

The Br35H dataset is specifically designed for brain tumor detection tasks and contains a large number of annotated MRI (Magnetic

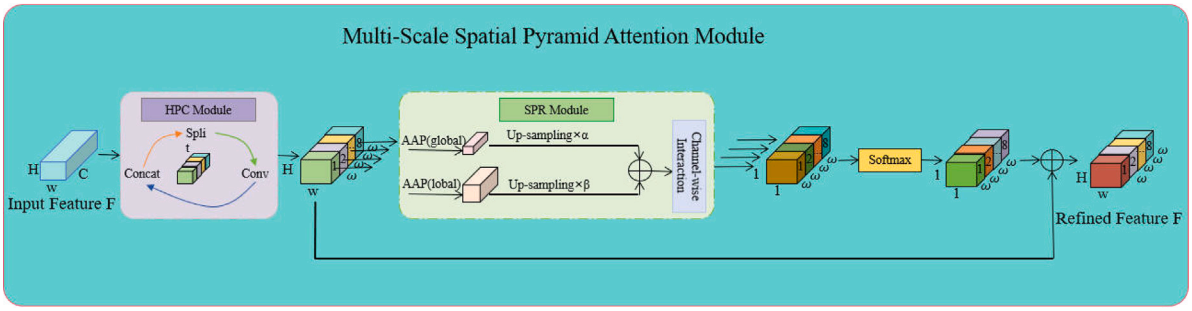

Fig. 3. A detailed illustration of the proposed HPC module, where Split means equally splitting in the channel dimension, Conv represents a 3 × 3 standard convolutional layer followed by batch normalization, and Concat refers to concatenating features in the channel dimension [32].

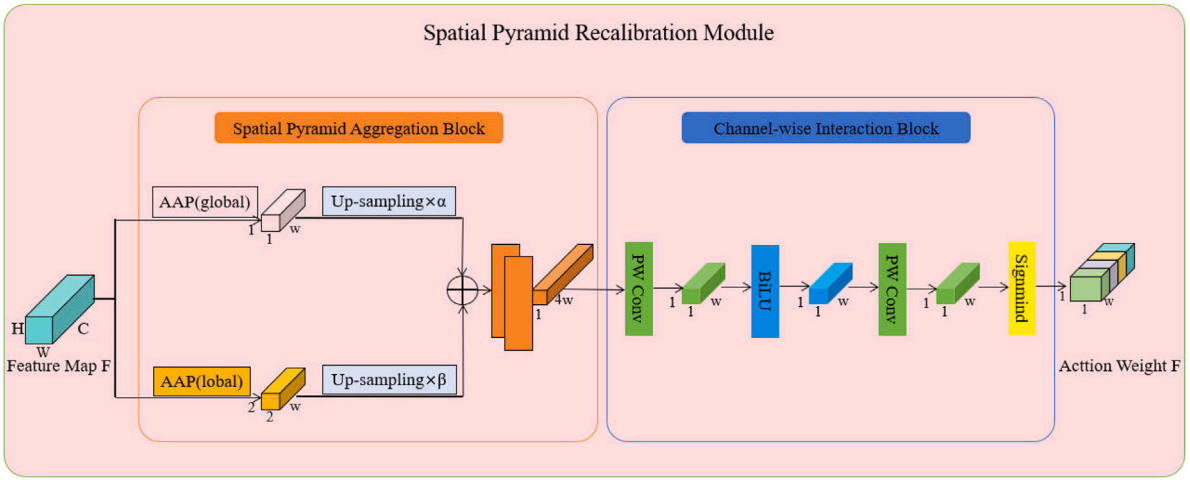

Fig. 4. Diagram of the proposed SPR module. It comprises two essential components, i.e., spatial pyramid aggregation block and channel-wise interaction block [32].

Table 1
The data set partitioning is in the Br35H dataset and the Roboflow dataset.

Dataset	Training set	Testing set	Validation set
Br35H	561	160	80
Roboflow	60	9	7

Resonance Imaging) brain tumor images. This dataset provides annotations for both benign and malignant tumors, including gliomas, meningiomas, and other types of brain tumors. The images are manually annotated by experts, ensuring label accuracy and providing high-quality training data for the model. The Roboflow dataset is a general-purpose image dataset generation platform that provides sample sets for various image processing tasks. In this study, we selected specific brain tumor images from the Roboflow dataset, further enriching the data source for model training. The images in the Roboflow dataset typically cover various imaging angles and conditions, allowing the model to adapt to different clinical imaging environments, thus enhancing robustness in real-world applications. Table 1 presents the specific dataset split. Through experiments on these two datasets, we are able to comprehensively verify the effectiveness and robustness of our method under different data conditions.

4.2. Experiments environment

The experimental environment in this study utilizes high-performance hardware and software configurations to support the efficient training and inference of deep learning models. In terms of hardware, it includes a high-performance GPU with large memory capacity, a powerful multi-core CPU, ample RAM, and fast solid-state storage, ensuring speed and stability when processing high-resolution medical imaging data. The software environment includes the Ubuntu operating system, Python programming language, and mainstream deep learning frameworks such as PyTorch, along with CUDA and cuDNN acceleration, as well as commonly used data processing and visualization libraries such as NumPy, OpenCV, and Matplotlib, ensuring the efficiency and reproducibility of model development and experimental analysis. The specific configuration is shown in Table 2:

Table 2
Hardware and Software configuration.

Hardware configuration	
Component	Specification
CPU	Intel Core i9-11900K @ 3.50 GHz
GPU	NVIDIA RTX 3090 (24 GB VRAM)
RAM	64 GB DDR4
Storage	1 TB NVMe SSD
Software configuration	
Software	Version/Details
Operating System	Ubuntu 20.04 LTS
Programming Language	Python 3.8
Deep Learning Frameworks	PyTorch 1.9.0, TensorFlow 2.5
CUDA Toolkit	CUDA 11.2
cuDNN	cuDNN 8.2
Libraries	NumPy, OpenCV, Matplotlib, scikit-learn
Version Control	Git

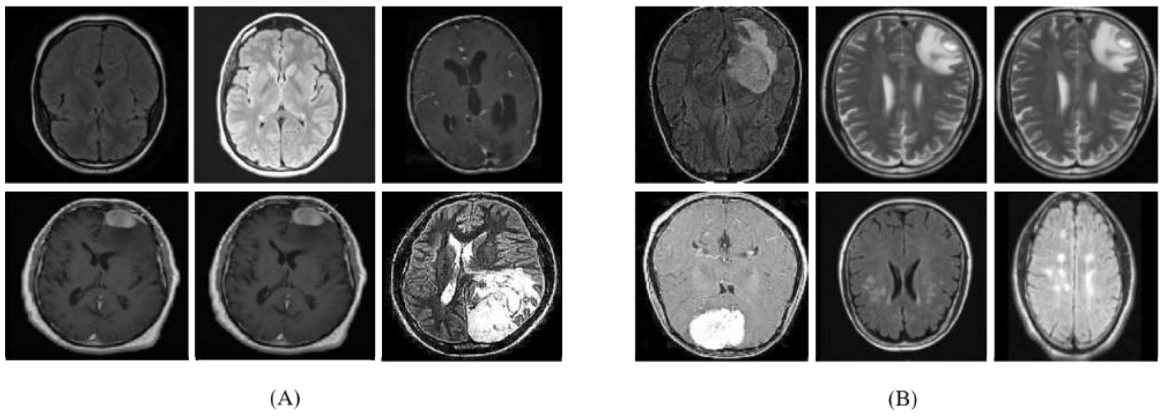

Fig. 5. Sample displays of the Br35H dataset and the Roboflow dataset. (a) shows a sample from the Br35H dataset. (b) shows a sample from the Roboflow dataset.

4.3. Evaluation matrix

This paper uses several evaluation metrics to measure the performance of the model in brain tumor detection tasks, including mean Average Precision (mAP@50), Precision (PR), Accuracy (AC), and F1-Score. The definitions and calculation methods for these evaluation metrics are as follows.

mAP@50 is a key evaluation metric used in detection tasks to measure the matching degree between predicted boxes and ground truth boxes. It is calculated by determining the Average Precision (AP) for each class at an Intersection over Union (IoU) threshold of 0.5, and then averaging across all classes, yielding mAP@50. The formula is as follows:

$$\text{mAP@50} = \frac{1}{N}\sum_{i=1}^{N}\text{AP}_i \tag{17}$$

where N is the number of classes, and AP_i is the Average Precision for the ith class.

Precision represents the proportion of correct positive predictions among all the predicted positive samples, used to evaluate the accuracy of the model's predictions. It is defined as:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{18}$$

where TP is the number of true positives, and FP is the number of false positives.

Accuracy represents the proportion of correctly classified samples out of the total samples, reflecting the overall performance of the model. It is calculated as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{19}$$

where TN is the number of true negatives, and FN is the number of false negatives.

The F1-Score is the harmonic mean of Precision and Recall, used to evaluate the model's precision and coverage comprehensively. It is defined as:

$$\text{F1-Score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{20}$$

where Recall is defined as:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{21}$$

4.4. Implementation details

During the training process, parameters such as learning rate, batch size, weight decay, and number of training epochs were tuned to achieve optimal model performance. Additionally, to balance the model complexity and computational efficiency, this paper selects the appropriate number of layers and parameter scale for the task. The specific parameter settings are shown in Table 3:

Table 3
Training parameters.

Parameter	Value
Learning Rate	0.001
Batch Size	16
Weight Decay	0.0005
Epochs	300
Layers	245
Parameters	10,236,528

4.5. Results

4.5.1. Comparison to prior work

In this study, we compared the performance of different models on the Roboflow and Br35H datasets. The experimental results are shown in Table 4. Compared to existing mainstream models, our proposed YOLO-TumorNet achieved significant improvements across all metrics, particularly in precision, accuracy, and F1 score, demonstrating superior performance.

On the Roboflow dataset, YOLO-TumorNet achieved an mAP@50 of 97.70%, which is an improvement of approximately 1.79 percentage points over the best-performing YOLOv8n . This improvement is primarily attributed to the introduction of the MSPA and SPR modules in our model, which enable more effective extraction of multi-scale spatial information and enhance detection performance for brain tumors of different sizes. Additionally, YOLO-TumorNet's PR increased by approximately 2.91 percentage points, reaching 98.66%, indicating a significant advantage in reducing false positive rates.

On the Br35H dataset, YOLO-TumorNet further improved its mAP@50 to 99.51%, outperforming YOLOv8n by about 1.80 percentage points. The Br35H dataset contains more complex brain tumor shapes, and our model's excellent performance on this dataset further confirms its robustness in complex contexts. YOLO-TumorNet also achieved an AC of 99.50%, which is 1.88 percentage points higher than YOLOv5n's 97.62%. This can be attributed to the model's finely-tuned architecture and sensitivity to small tumors, enabling it to maintain high accuracy even in complex lesion environments. Furthermore, in terms of F1 score, YOLO-TumorNet reached 97.62% and 99.44% on the Roboflow and Br35H datasets, respectively, surpassing YOLOv8n by approximately 1.56 and 1.63 percentage points. The improvement in F1 score reflects the model's enhanced ability to balance precision and recall, enabling more effective detection of true tumor locations while minimizing false negatives. We speculate that this improvement is due to the multi-scale feature fusion and channel-adaptive weight allocation, which enhance the model's adaptability to the diversity of brain tumor shapes.

Table 4
Comparison of model performance on Roboflow and Br35H datasets.

Models	Roboflow dataset				Br35H dataset			
	mAP50	PR	AC	F1-Score	mAP50	PR	AC	F1-Score
YOLOv4 [35]	93.90	93.66	93.86	93.02	95.32	95.46	95.64	95.84
SSD [36]	94.46	94.36	94.54	94.67	96.24	96.16	96.33	96.41
YOLOv7-Tiny [37]	94.84	94.44	94.76	94.93	96.93	96.23	96.53	96.77
VitDet [12]	94.96	94.53	94.66	94.16	96.70	96.34	96.47	96.91
YOLOv5n [38]	95.66	95.56	95.86	95.76	97.44	97.34	97.62	97.52
YOLOv8n [39]	95.91	95.75	96.06	96.06	97.71	97.55	97.81	97.81
YOLO-TumorNet	**97.70**	**98.66**	**97.71**	**97.62**	**99.51**	**99.47**	**99.50**	**99.44**

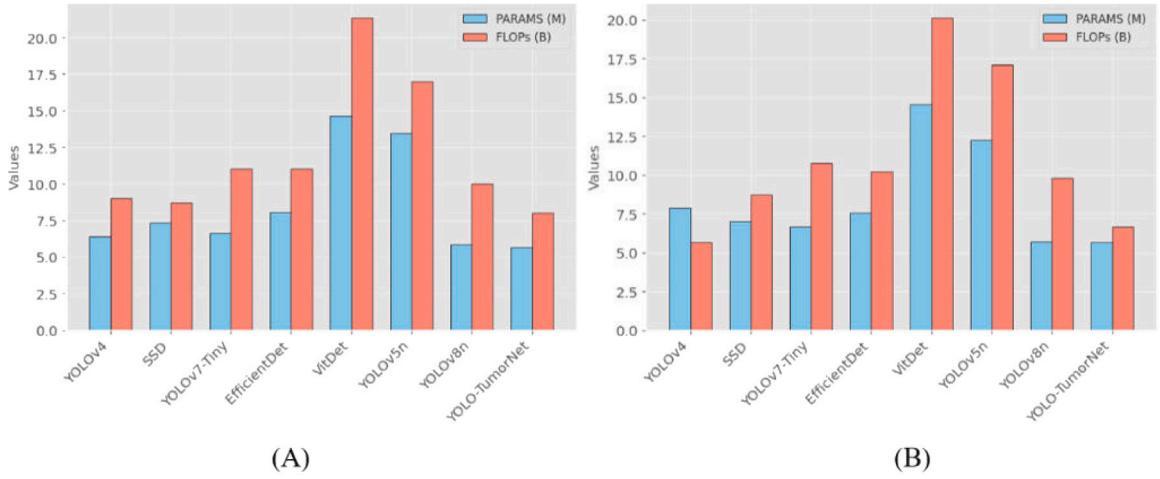

Fig. 6. Different model parameter comparisons.

Table 5
Comparison of Model Parameters (PARAMS) and Floating Point Operations (FLOPs) on Roboflow and Br35H datasets.

Model	Roboflow dataset		Br35H dataset	
	PARAMS	FLOPs	PARAMS	FLOPs
YOLOv4	6.39 M	9.00 B	7.85 M	5.68 B
SSD	7.34 M	8.71 B	6.99 M	8.74 B
YOLOv7-Tiny	6.64 M	11.02 B	6.64 M	10.77 B
EfficientDet	8.06 M	11.02 B	7.55 M	10.22 B
VitDet	14.67 M	21.33 B	14.52 M	20.11 B
YOLOv5n	13.47 M	19.01 B	12.22 M	17.09 B
YOLOv8n	5.87 M	10.00 B	5.69 M	9.80 B
YOLO-TumorNet	5.69 M	8.02 B	5.66 M	6.64 B

4.5.2. Comparison of model parameters

As shown in Table 5, the proposed YOLO-TumorNet demonstrates significant advantages in terms of model parameters (PARAMS) and floating point operations (FLOPs). Compared to other models, YOLO-TumorNet achieves superior computational efficiency while maintaining a relatively small number of parameters. Specifically, YOLO-TumorNet has fewer parameters and FLOPs than models such as YOLOv5n and VitDet, yet it outperforms them in detection performance. This indicates that YOLO-TumorNet has been optimized in terms of model architecture, reducing computational burden while ensuring high detection accuracy. This efficiency is particularly suitable for applications in medical image analysis, where heavy computation is required, and it can perform exceptionally well in resource-constrained environments. Therefore, YOLO-TumorNet not only maintains high performance but also offers good computational efficiency and a lightweight model, making it highly suitable for large-scale medical image detection tasks. The Fig. 6 visualizes the table, providing a more intuitive view of the advantages of our model.

4.5.3. Qualitative results

Fig. 7 shows the performance of YOLO-TumorNet in brain tumor detection on MRI images, highlighting the model's significant advantages in boundary clarity and detail capture. YOLO-TumorNet can more accurately identify tumor boundaries and maintain high detection precision even in complex backgrounds. Additionally, our experiments demonstrate YOLO-TumorNet's excellent performance in detecting small tumors (see Fig. 8). In tests with tumors of varying sizes, YOLO-TumorNet not only performs well in detecting large tumors but also shows strong accuracy and robustness in identifying small tumors. Small abnormalities are often more difficult to detect but can hold critical clinical significance. YOLO-TumorNet's outstanding capability in this area is especially valuable for medical imaging analysis, providing reliable support for early diagnosis and timely treatment.

Finally, as shown in Fig. 9, we conducted feature visualization for the MSPA module. From the figure, it can be observed that MSPA significantly enhances the model's focus on key regions, enabling the model to effectively capture subtle tumor features across different scales. This visualization further validates the importance of the MSPA mechanism in enhancing model performance. It also demonstrates YOLO-TumorNet's strong potential for application in complex brain tumor detection tasks, providing clinicians with more accurate and reliable diagnostic support.

4.6. Limitations and future work

Despite its strong performance in brain tumor detection tasks, YOLO-TumorNet faces several challenges in practical applications. First, the model's reliance on large-scale, high-quality annotated datasets poses a significant limitation. In the medical field, particularly for brain tumor MRI images, precise annotations require expertise from radiologists, which incurs substantial time and labor costs. This scarcity of annotated data restricts the model's scalability across diverse medical institutions and datasets. Additionally, the relatively complex architecture of YOLO-TumorNet, while improving detection accuracy, leads

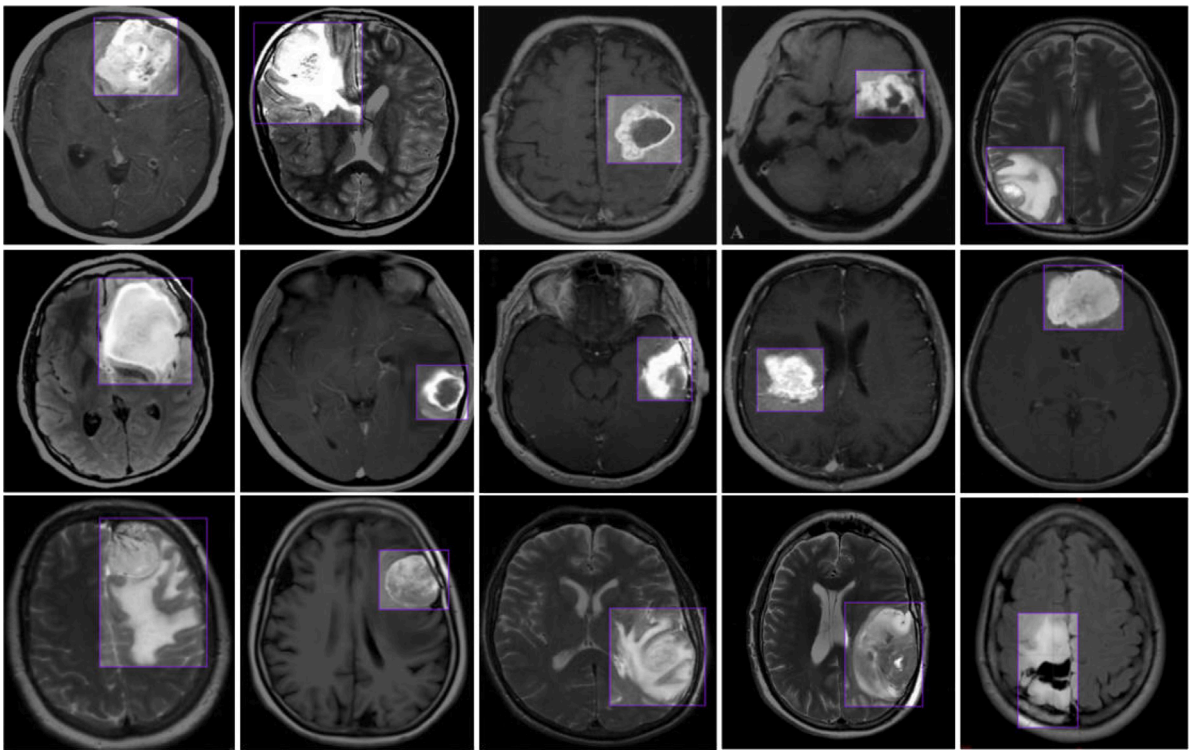

Fig. 7. YOLO-TumorNet brain tumor detection visualization.

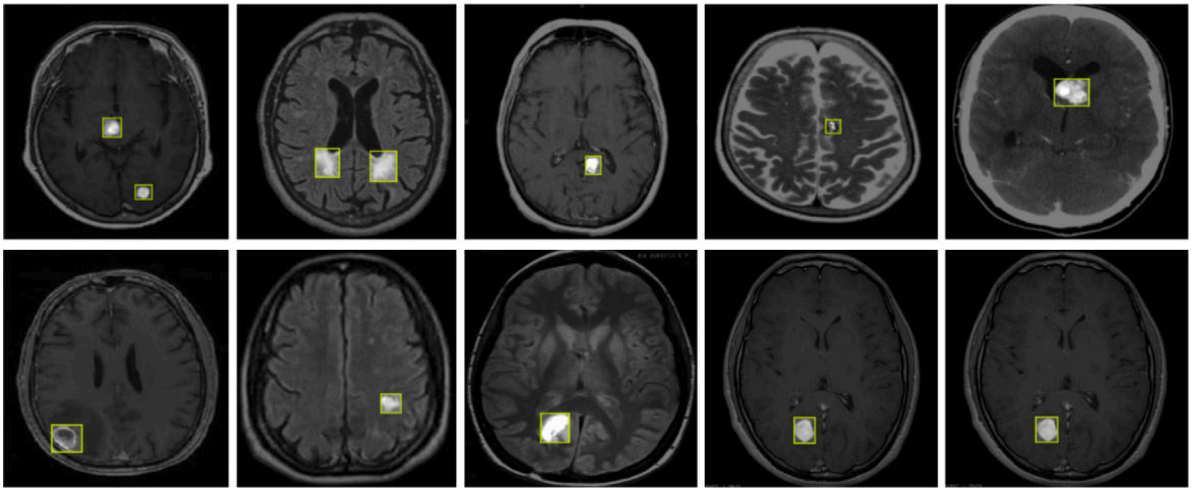

Fig. 8. YOLO-TumorNet for small brain tumor detection.

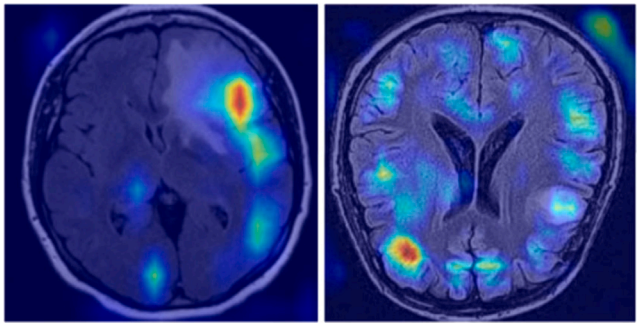

Fig. 9. Visualization of MSPA feature visualization.

to higher computational demands. This poses challenges for deployment on resource-constrained devices, such as portable medical imaging equipment or mobile platforms. Ensuring real-time performance and operational efficiency under such constraints remains difficult, impacting its suitability for on-site diagnostics and edge computing scenarios.

To address these challenges, future research could explore several promising directions. Firstly, unsupervised or weakly supervised learning approaches offer potential for reducing dependency on large, accurately labeled datasets. Techniques such as pre-training on unlabeled data and fine-tuning with minimal labeled data can significantly lower annotation costs while improving the model's generalization capabilities. Secondly, optimizing model structure by adopting lightweight

Table 6
Experiment results for each component.

Method	nceptionNeXt	MSPA	BIFPN	mAP50	mAP50-95
YOLOv8n	✗	✗	✗	83.26	85.37
YOLOv10-1	✔	✗	✗	85.48	84.46
YOLOv10-2	✗	✔	✗	83.58	85.15
YOLOv10-3	✗	✗	✔	85.13	82.27
YOLOv10-4	✔	✔	✗	90.25	87.66
YOLOv10-5	✗	✔	✔	95.37	92.46
YOLO-TumorNet	✔	✔	✔	**97.70**	**95.48**

architectures can enhance efficiency. Strategies like knowledge distillation, model pruning, and quantization can help reduce computational overhead, making the model more practical for deployment in resource-constrained environments. Lastly, incorporating adaptive multi-scale feature extraction mechanisms could enhance the model's versatility and detection accuracy, particularly for tumors with varying shapes and sizes. This advancement would further strengthen its applicability in complex clinical settings, enabling more robust and reliable diagnostic support.

4.7. Ablation study

To verify the contribution of each module in YOLO-TumorNet, we conducted ablation experiments as shown in Table 6, progressively adding the InceptionNeXt, MSPA, and BIFPN modules and observing their impact on model performance. The experimental results demonstrate that the inclusion of each module plays a crucial role in improving the overall performance of the model. After introducing the InceptionNeXt module into the baseline model, the detection accuracy was significantly improved. This is mainly due to the multi-branch convolutional structure of InceptionNeXt, which enhances the model's feature extraction capabilities while maintaining computational efficiency, allowing the model to better capture complex features. In medical imaging, this enhancement is especially important for recognizing brain tumors with diverse shapes. The addition of the MSPA module further improved the model's stability under different IoU thresholds, as it enhanced the model's sensitivity to spatial information through a multi-scale spatial pyramid structure, particularly excelling in detail detection and boundary recognition. MSPA provides crucial support for capturing subtle features in medical images. The BIFPN module plays a key role in multi-scale feature fusion. When combined with other modules, the model's multi-scale detection ability is significantly enhanced, as BIFPN facilitates the exchange of information across different scales and balances feature weights, making the model more accurate in detecting objects of varying sizes. When all modules are combined in YOLO-TumorNet, the model achieves its highest performance. The experimental results indicate that the combination of these modules enables the model to more efficiently handle the diversity and complexity of brain tumors, improving detection accuracy and robustness.

5. Conclusion

This paper introduces YOLO-TumorNet, an advanced model for automatic brain tumor detection. By integrating InceptionNeXt, MSPA, and BIFPN mechanisms into YOLOv10, the model achieves significant improvements in accuracy and robustness for detecting brain tumors in MRI images. Comprehensive experiments on the Br35H and Roboflow datasets highlight YOLO-TumorNet's superior performance in boundary clarity, detail capture, and small tumor detection, surpassing existing state-of-the-art models. In conclusion, YOLO-TumorNet offers a highly efficient and accurate solution for automatic brain tumor detection, with strong potential for clinical applications. Future efforts will focus on optimizing the model architecture, minimizing reliance on large-scale annotated datasets, and enhancing efficiency in resource-constrained environments, thereby extending its applicability to medical image analysis and clinical decision-making.

CRediT authorship contribution statement

Jian Huang: Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization. **Wen Ding:** Writing – original draft, Investigation, Formal analysis, Data curation. **Tiancheng Zhong:** Writing – review & editing, Supervision, Resources. **Gang Yu:** Writing – review & editing, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

This work was partially supported by the National Key R&D Program of China (grant number 2023YFC2706400) and National Natural Science Foundation of China (grant number 62076218).

References

[1] M. Woźniak, J. Siłka, M. Wieczorek, Deep neural network correlation learning mechanism for CT brain tumor detection, Neural Comput. Appl. 35 (20) (2023) 14611–14626.
[2] E. Belykh, K.V. Shaffer, C. Lin, V.A. Byvaltsev, M.C. Preul, L. Chen, Blood-brain barrier, blood-brain tumor barrier, and fluorescence-guided neurosurgical oncology: delivering optical labels to brain tumors, Front. Oncol. 10 (2020) 739.
[3] J. Selvaraj, A. Jayanthy, Design and development of artificial intelligence-based application programming interface for early detection and diagnosis of colorectal cancer from wireless capsule endoscopy images, Int. J. Imaging Syst. Technol. 34 (2) (2024) e23034.
[4] J. Selvaraj, A. Jayanthy, Automatic polyp semantic segmentation using wireless capsule endoscopy images with various convolutional neural network and optimization techniques: A comparison and performance evaluation, Biomed. Engineering: Appl. Basis Commun. 35 (06) (2023) 2350026.
[5] S. Jothiraj, J.A. Kandaswami, Localization and semantic segmentation of polyp in an effort of early diagnosis of colorectal cancer from wireless capsule endoscopy images, in: 2022 Seventh International Conference on Parallel, Distributed and Grid Computing, PDGC, IEEE, 2022, pp. 749–754.
[6] M.M. Badža, M.Č. Barjaktarović, Classification of brain tumors from MRI images using a convolutional neural network, Appl. Sci. 10 (6) (2020) 1999.
[7] F.J.P. Montalbo, A computer-aided diagnosis of brain tumors using a fine-tuned YOLO-based model with transfer learning, KSII Trans. Internet Inf. Syst. (TIIS) 14 (12) (2020) 4816–4834.
[8] T. Wang, Z. Yu, J. Fang, J. Xie, F. Yang, H. Zhang, L. Zhang, M. Du, L. Li, X. Ning, Multidimensional fusion of frequency and spatial domain information for enhanced camouflaged object detection, Inf. Fusion (2024) 102871.
[9] Q. Chen, F. He, G. Wang, X. Bai, L. Cheng, X. Ning, Dual guidance enabled fuzzy inference for enhanced fine-grained recognition, IEEE Trans. Fuzzy Syst. (2024) 1–14, http://dx.doi.org/10.1109/TFUZZ.2024.3427654.
[10] F.J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, D. González-Ortega, A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network, in: Healthcare, vol. 9, (2) MDPI, 2021, p. 153.
[11] N.A. Samee, N.F. Mahmoud, G. Atteia, H.A. Abdallah, M. Alabdulhafith, M.S. Al-Gaashani, S. Ahmad, M.S.A. Muthanna, Classification framework for medical diagnosis of brain tumor with an effective hybrid transfer learning model, Diagnostics 12 (10) (2022) 2541.
[12] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, M. Bendechache, Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images, Sci. Rep. 11 (1) (2021) 1–17.

[13] M.A. Al-Masni, W.R. Kim, E.Y. Kim, Y. Noh, D.H. Kim, Automated detection of cerebral microbleeds in MR images: A two-stage deep learning approach, NeuroImage: Clin. 28 (2020) 102464.
[14] M. Kang, C.M. Ting, F.F. Ting, R.C.W. Phan, RCS-YOLO: A fast and high-accuracy object detector for brain tumor detection, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2023, pp. 600–610.
[15] F.J. Díaz-Pernas, M. Martínez-Zarzuela, M. Antón-Rodríguez, D. González-Ortega, A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network, in: Healthcare, vol. 9, (2) MDPI, 2021, p. 153.
[16] H. Chegraoui, C. Philippe, V. Dangouloff-Ros, A. Grigis, R. Calmon, N. Boddaert, F. Frouin, J. Grill, V. Frouin, Object detection improves tumour segmentation in MR images of rare brain tumours, Cancers 13 (23) (2021) 6113.
[17] B.K. Ahir, H.H. Engelhard, S.S. Lakka, Tumor development and angiogenesis in adult brain tumor: glioblastoma, Mol. Neurobiol. 57 (2020) 2461–2478.
[18] U. Baid, S. Ghodasara, S. Mohan, M. Bilello, E. Calabrese, E. Colak, K. Farahani, J. Kalpathy-Cramer, F.C. Kitamura, S. Pati, et al., The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification, 2021, arXiv preprint arXiv:2107.02314.
[19] E. Irmak, Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework, Iran. J. Sci. Technol. Trans. Electr. Eng. 45 (3) (2021) 1015–1036.
[20] M. Nazir, S. Shakil, K. Khurshid, Role of deep learning in brain tumor detection and classification (2015 to 2020): A review, Comput. Med. Imaging Graph. 91 (2021) 101940.
[21] F. Isensee, P.F. Jäger, P.M. Full, P. Vollmuth, K.H. Maier-Hein, Nnu-net for brain tumor segmentation, in: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6, Springer, 2021, pp. 118–132.
[22] A. Raza, H. Ayub, J.A. Khan, I. Ahmad, A. S. Salama, Y.I. Daradkeh, D. Javeed, A. Ur Rehman, H. Hamam, A hybrid deep learning-based approach for brain tumor classification, Electronics 11 (7) (2022) 1146.
[23] Z. Liu, L. Tong, L. Chen, Z. Jiang, F. Zhou, Q. Zhang, X. Zhang, Y. Jin, H. Zhou, Deep learning based brain tumor segmentation: a survey, Complex Intell. Syst. 9 (1) (2023) 1001–1026.
[24] J. Amin, M. Sharif, A. Haldorai, M. Yasmin, R.S. Nayak, Brain tumor detection and classification using machine learning: a comprehensive survey, Complex Intell. Syst. 8 (4) (2022) 3161–3183.
[25] J. Jeong, Y. Lei, S. Kahn, T. Liu, W.J. Curran, H.K. Shu, H. Mao, X. Yang, Brain tumor segmentation using 3D mask R-CNN for dynamic susceptibility contrast enhanced perfusion imaging, Phys. Med. Biol. 65 (18) (2020) 185009.
[26] M. Toğaçar, B. Ergen, Z. Cömert, Tumor type detection in brain MR images of the deep model developed using hypercolumn technique, attention modules, and residual blocks, Med. Biol. Eng. Comput. 59 (1) (2021) 57–70.
[27] M.S.I. Khan, A. Rahman, T. Debnath, M.R. Karim, M.K. Nasir, S.S. Band, A. Mosavi, I. Dehzangi, Accurate brain tumor detection using deep convolutional neural network, Comput. Struct. Biotechnol. J. 20 (2022) 4733–4745.
[28] M.A. Talukder, M.M. Islam, M.A. Uddin, A. Akhter, M.A.J. Pramanik, S. Aryal, M.A.A. Almoyad, K.F. Hasan, M.A. Moni, An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning, Expert Syst. Appl. 230 (2023) 120534.
[29] S. Rani, B.K. Singh, D. Koundal, V.A. Athavale, Localization of stroke lesion in MRI images using object detection techniques: a comprehensive review, Neurosci. Informatics 2 (3) (2022) 100070.
[30] J. Chen, H. Mai, L. Luo, X. Chen, K. Wu, Effective feature fusion network in BIFPN for small object detection, in: 2021 IEEE International Conference on Image Processing, ICIP, IEEE, 2021, pp. 699–703.
[31] W. Yu, P. Zhou, S. Yan, X. Wang, Inceptionnext: When inception meets convnext, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 5672–5683.
[32] Y. Yu, Y. Zhang, Z. Cheng, Z. Song, C. Tang, Multi-scale spatial pyramid attention mechanism for image recognition: An effective approach, Eng. Appl. Artif. Intell. 133 (2024) 108261.
[33] G.A. Amran, M.S. Alsharam, A.O.A. Blajam, A.A. Hasan, M.Y. Alfaifi, M.H. Amran, A. Gumaei, S.M. Eldin, Brain tumor classification and detection using hybrid deep tumor network, Electronics 11 (21) (2022) 3457.
[34] Magesh, Brain tumor dataset, 2024, https://universe.roboflow.com/magesh-kctcd/brain-tumor-3rrwu. URL https://universe.roboflow.com/magesh-kctcd/brain-tumor-3rrwu visited on 2024-06-03.
[35] S. Mei, H. Jiang, L. Ma, YOLO-lung: A practical detector based on imporved YOLOv4 for pulmonary nodule detection, in: 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE, 2021, pp. 1–6.
[36] C.S. Rao, K. Karunakara, Efficient detection and classification of brain tumor using kernel based SVM for MRI, Multimedia Tools Appl. 81 (5) (2022) 7393–7417.
[37] A. Jlassi, K. ElBedoui, W. Barhoumi, Glioma tumor's detection and classification using joint YOLOv7 and active contour model, in: 2023 IEEE Symposium on Computers and Communications, ISCC, IEEE, 2023, pp. 1–5.
[38] M. Kumar, U. Pilania, S. Thakur, T. Bhayana, YOLOv5x-based brain tumor detection for healthcare applications, Procedia Comput. Sci. 233 (2024) 950–959.
[39] B. Selcuk, T. Serif, Brain tumor detection and localization with YOLOv8, in: 2023 8th International Conference on Computer Science and Engineering, UBMK, IEEE, 2023, pp. 477–481.