



## OPEN A non-invasive diagnostic approach for neuroblastoma utilizing preoperative enhanced computed tomography and deep learning techniques

Yuanyuan Wang<sup>1</sup>, Fangfang Wang<sup>2</sup>, Zixin Qin<sup>3</sup>, Yongcheng Fu<sup>1</sup>, Jingyue Wang<sup>1</sup>, Shangkun Li<sup>1</sup> & Da Zhang<sup>1</sup>✉

Neuroblastoma presents a wide variety of clinical phenotypes, demonstrating different levels of benignity and malignancy among its subtypes. Early diagnosis is essential for effective patient management. Computed tomography (CT) serves as a significant diagnostic tool for neuroblastoma, utilizing machine vision imaging, which offers advantages over traditional X-ray and ultrasound imaging modalities. However, the high degree of similarity among neuroblastoma subtypes complicates the diagnostic process. In response to these challenges, this study presents a modified version of the You Only Look Once (YOLO) algorithm, called YOLOv8-IE. This revised approach integrates feature fusion and inverse residual attention mechanisms. The aim of YOLO-IE is to improve the detection and classification of neuroblastoma tumors. In light of the image features, we have implemented the inverse residual-based attention structure (iRMB) within the detection network of YOLOv8, thereby enhancing the model's ability to focus on significant features present in the images. Additionally, we have incorporated the centered feature pyramid EVC module. Experimental results show that the proposed detection network, named YOLO-IE, attains a mean Average Precision (mAP) 7.9% higher than the baseline model, YOLO. The individual contributions of iRMB and EVC to the performance improvement are 0.8% and 3.6% above the baseline model, respectively. This study represents a significant advancement in the field, as it not only facilitates the detection and classification of neuroblastoma but also demonstrates the considerable potential of machine learning and artificial intelligence in the realm of medical diagnosis.

**Keywords** Deep learning, Artificial intelligence, Neuroblastoma, Machine learning, Diagnosis

Neuroblastoma, a neoplasm that arises from the sympathetic nervous system, represents the most prevalent and lethal extracranial solid malignant tumor in pediatric populations. The incidence of neuroblastoma is approximately 10.2 cases per million children under the age of 15, making it the most frequently diagnosed cancer during the first year of life<sup>1</sup>. The clinical manifestations of neuroblastoma can vary significantly, ranging from spontaneous regression without intervention to benign mature ganglioneuromas, as well as malignant cases with multiple systemic metastases or a combination of both benign and malignant components<sup>1</sup>. Different clinical phenotypes of neuroblastoma necessitate different treatment strategies and are associated with varying prognoses. Consequently, early diagnosis and accurate classification of neuroblastoma are critical for the advancement of diagnostic approaches. The diagnosis of neuroblastoma typically involves a combination of laboratory tests, radiological imaging, and pathological evaluation<sup>2</sup>. While pathology remains the gold standard for neuroblastoma diagnosis, tissue biopsy presents significant challenges, including the risk of incomplete sampling or rupture of tumors, as well as considerable safety concerns. In some cases, patients may require surgical intervention under general anesthesia. Therefore, the development of novel diagnostic modalities to

<sup>1</sup>Department of Pediatric Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450000, China. <sup>2</sup>Department of Electronic information, Zhengzhou University Cyberspace Security College, Zhengzhou 450000, Henan, China. <sup>3</sup>Department of Neurosurgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, Henan, China. ✉email: zhd1000@126.com

effectively differentiate among the various clinical phenotypes of neuroblastoma is essential for optimizing therapeutic management.

Deep learning models have profoundly impacted various areas of medical informatics, especially in applications based on computer vision. Numerous researchers have developed a range of deep learning models to address diverse medical tasks, including assisted diagnosis, disease screening, and lesion detection<sup>3</sup>. Artificial intelligence (AI)-driven deep learning algorithms are widely used in medical informatics for their outstanding data processing capabilities. Notably, these algorithms have demonstrated considerable advancements in tumor detection, diagnosis, and characterization. For instance, Mathivanan et al. employed transfer learning techniques to achieve an impressive accuracy of 99.75% in the detection of various brain tumors<sup>4</sup>. Furthermore, AI plays a pivotal role in surgical planning by accurately delineating tumor boundaries and normal tissues, thereby facilitating a balance between intervention and the preservation of quality of life. Furthermore, AI can predict complications, recurrence rates, and treatment responses. It can guide optimal follow-up strategies and provide personalized recommendations to patients through customized screening programs<sup>5</sup>. However, it is noteworthy that there is a scarcity of deep learning models specifically designed for the detection and diagnosis of extracranial tumors. This study aims to address this gap by focusing on the detection and diagnosis of extracranial neuroblastoma across different subtypes, utilizing an enhanced deep learning-based YOLO (You Only Look Once) model. The YOLO model is a sophisticated deep learning framework characterized by its robust image analysis and target detection capabilities, enabling the localization of multiple targets within a single image<sup>6</sup>. The YOLO algorithm operates as an end-to-end, single-stage prediction model based on a unified neural network, allowing for direct predictions of coordinates and object positions from input images. This model exhibits superior universality and transferability compared to traditional convolutional neural networks (CNNs)<sup>7</sup>. Consequently, the YOLO model holds significant promise for applications in medical classification and diagnostic tasks.

To date, there has been no available CT dataset for neuroblastoma that is suitable for classification tasks. The primary contributions of this paper are fourfold.

- (1) An NB-CT dataset comprising 233 cases was developed for the classification of neuroblastoma. This dataset includes 141 patients diagnosed with neuroblastoma, 37 patients with ganglioneuroma, and 55 patients simultaneously possessing ganglioneuroblastoma. In comparison to other datasets, our dataset offers several advantages: it encompasses a larger number of cases, integrates multiple classification tasks, and is exclusively focused on the NB-CT domain.
- (2) The YOLOv8-IE model is developed based on the YOLOv8 architecture to integrate both global and local information for the classification of neuroblastoma. The classification results obtained from the NB-CT dataset indicate that the YOLOv8-IE model outperforms the baseline model.
- (3) The results highlight the effectiveness of the proposed model in classifying extracranial neuroblastoma, suggesting its potential to improve diagnostic accuracy in medical image analysis.
- (4) The proposed model represents a significant advancement in the field of neuroblastoma diagnosis, indicating that artificial intelligence has the potential to transform the existing paradigm of medical diagnosis.

This paper is organized in the following manner. In Section “Related work”, we provide a review of the relevant literature. Section “Materials” outlines the specific process involved in constructing our dataset, along with the associated treatments. Section “Methods” details the methodology employed for conducting benchmark experiments on this dataset, as well as our proposed YOLO-IE method. Section “Results” presents an analysis of the experimental results. Section “Discussion” encompasses the discussion segment of this paper, while Section “Conclusion and prospect” concludes with a summary of findings and suggestions for future research.

## Related work

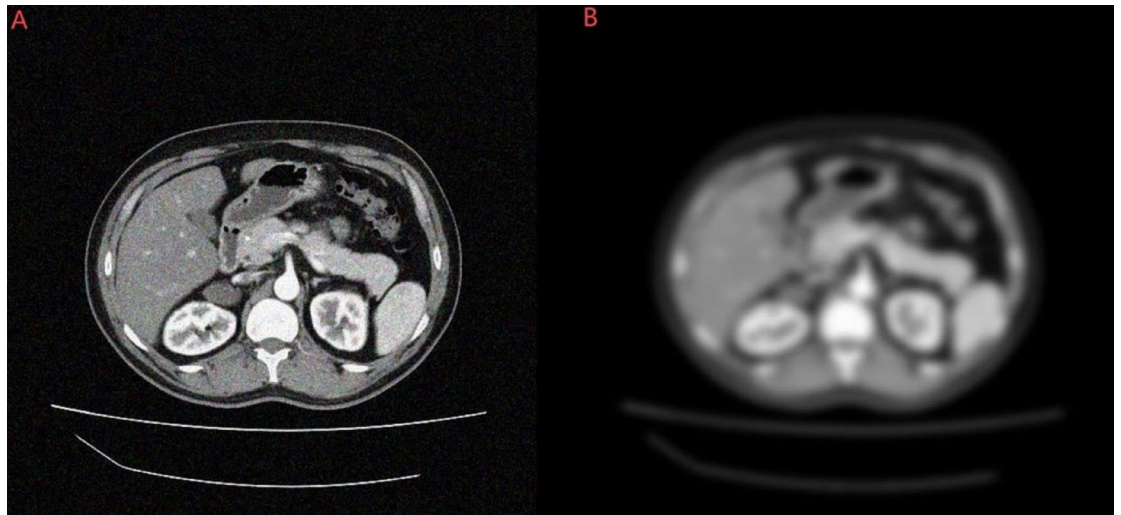
### Image enhancement

In this study, we employ the Image Data Generator, a pivotal technology, to augment the dataset for training a deep learning model aimed at the diagnosis of neuroblastoma. By generating modified copies of images that incorporate noise, blurring, scaling, gamut transformations, equalization, and color dithering, the model is exposed to a broader spectrum of variations, thereby enhancing its capacity to process novel data. This is critical for variability in medical imaging, especially in data-limited medical fields. An illustrative example is presented in Fig. 1, which displays an image prior to and following the application of blurring.

This enhancement strategy contributes to the acquisition of a more extensive and diverse training dataset, which facilitates improved generalization of the model across different scenarios. The utilization of the Image Data Generator during the training of models offers two primary advantages. Firstly, it ensures that deep learning models are exposed to a more enriched training set, which aids in the learning of intricate patterns and features. Secondly, the automatic generation of augmented images diminishes the risk of model overfitting and enhances the model's resilience to variations in input, thereby augmenting its robustness. This enhancement-oriented approach has been acknowledged for its effectiveness in improving the overall performance of deep learning models, consequently enhancing their accuracy and adaptability in real-world applications<sup>8</sup>.

### Attention mechanisms

The attention mechanism has been proposed to mitigate the computational complexity associated with image processing while simultaneously enhancing performance by employing a model that concentrates on specific regions of an image rather than the entirety of the image<sup>9,10</sup>. Attention mechanisms have been applied across various visual tasks, including image classification<sup>11</sup>, target detection<sup>12</sup>, semantic segmentation<sup>13</sup>, 3D vision<sup>14</sup>, and so on. Existing attention methods can be categorized into three primary types: channel attention<sup>15</sup>, spatial



**Fig. 1.** Picture enhancement (A) Picture before blurring. (B) Picture after blurring.

attention<sup>16</sup>, and combined channel and spatial attention<sup>17</sup>. Channel attention addresses the issue of determining which information warrants focus, whereas spatial attention pertains to identifying the specific locations that require attention. Spatial attention emphasizes global location information over local location information to resolve the issue of spatial positioning. Medical images, known for their high resolution and intricate details, offer opportunities to improve the attention mechanism's ability to focus on specific regions. This is particularly important for detecting tumor targets that may appear randomly within the context of this study.

### Characteristic pyramid network architecture

Feature Pyramid Network (hereinafter referred to as FPN) is a network architecture designed to construct multi-scale feature representations, primarily addressing the challenges posed by variations in target scale in tasks such as object detection and image segmentation. FPN generates multi-scale feature representations enriched with semantic information by integrating feature maps across different scales, so as to improve the performance of the model in detecting or segmenting targets at different scales. Li et al. introduced the Cross-Layer Feature Pyramid Network (CFPN), which integrates multi-scale features from different levels into feature maps, which can access both high-level and low-level information for a more comprehensive contextual understanding. Direct cross-layer communication can improve the progressive fusion of salient target detection and reduce the loss of important information during the feature fusion process<sup>18</sup>. Furthermore, the Quasi-Balanced Pyramid Network introduced by Song et al. utilizes an implicit function to model a balanced state of the feature pyramid at infinite depth, creating a more realistic convergence state model that significantly improves network performance<sup>19</sup>. Furthermore, Dang et al. proposed a new Hierarchical Attention Feature Pyramid Network (HA-FPN), which employs multi-scale convolutional features and self-attention mechanisms to capture contextual information among markers. The Channel Attention Mechanism (CAM) is utilized to select channels with rich information, thereby reducing the loss of channel data. HA-FPN markedly improves the accuracy of bounding box detection, leading to enhanced recognition and localization of target objects<sup>20</sup>. Medical images often present at various scales (e.g., cells, tissues, organs) exhibit distinct features and sizes, necessitating reliance on global and local information. This scenario highlights the potential for further advancements in the feature pyramid structure.

### YOLO

YOLO models are widely used in target detection, computer vision, and other fields. For example, Gope et al. compared different versions of the YOLO model and found that YOLOv8 showed excellent performance in detecting green coffee beans<sup>21</sup>. YOLO models are also widely used in the medical field. Rong et al. used a histology-based YOLO detection model (HD-YOLO), a new method for significantly accelerating the segmentation of cell nuclei and TME quantification. The study demonstrated that HD-YOLO outperforms existing WSI analysis methods in terms of nuclei detection, classification accuracy, and computation time. The advantages of the system were validated on three different tissue types: lung, liver, and breast cancer. For breast cancer, the nuclei feature of HD-YOLO was more significant in terms of prognostic significance than both estrogen receptor status and progesterone receptor status by immunohistochemistry<sup>22</sup>. Xie et al. proposed SMLS-YOLO, an instance segmentation method based on YOLOv8n-seg. The study demonstrated that SMLS-YOLO has a promising application in image segmentation of pathologic myopia<sup>23</sup>. Zhang et al. proposed an automatic model (MA-YOLO) for MA detection in fluorescein angiography (FFA) images. The results showed that the MA-YOLO model had the best performance in MA detection with the best metrics including recall, precision, F1 score, and AP of 88.23%, 97.98%, 92.85%, and 94.62%, respectively. The proposed MA-YOLO model is generally suitable for the automatic detection of MA in FFA images, which helps ophthalmologists diagnose the progression of diabetic retinopathy<sup>24</sup>.

In addition to this, the YOLO model has also been applied in the field of tumor detection and diagnosis. Dinesh et al. aimed to analyze the medical imaging data (mainly CT scans) by analyzing the medical data, and using a convolutional neural network (CNN) and a CNN based on the YOLO model. YOLO-based CNN (YCNN) model to identify important features and cancerous growths in the pancreas, to create a deep learning based system for early prediction of pancreatic cancer. The YCNN method was shown to perform well by a percent accuracy compared to other modern techniques in a thorough review of comparative findings<sup>25</sup>.

## Materials

### Data collection

The study was conducted in accordance with the guiding principles outlined in the Declaration of Helsinki. Due to the retrospective nature of the enrollment process, the Ethics Committee of the First Affiliated Hospital of Zhengzhou University agreed that we waived informed consent. The study protocol received approval from the Ethics Committee of the First Affiliated Hospital of Zhengzhou University (2024-KY-1520).

We collected electronic medical record data from patients diagnosed with neuroblastoma at the First Affiliated Hospital of Zhengzhou University between January 2019 and January 2024. The inclusion criteria for cases were as follows: (1) Pathologically confirmed diagnoses of neuroblastoma, ganglion cell neuroblastoma, and both neuroblast and neuroectodermal components. (2) Completion of enhanced computed tomography (CT) examinations at our hospital prior to surgery, which included the non-contrast-enhanced phase (NC phase), arterial phase (ART phase), and portal venous phase (PV phase). The exclusion criteria were as follows: (1) An interval exceeding 90 days between pathological findings and preoperative CT; (2) Poor image quality resulting in a lack of clarity; (3) Absence of specific staging in preoperative enhanced CT. We meticulously screened the cases in accordance with the established inclusion and exclusion criteria and incorporated them into the diagnostic dataset.

### CT acquisition protocol

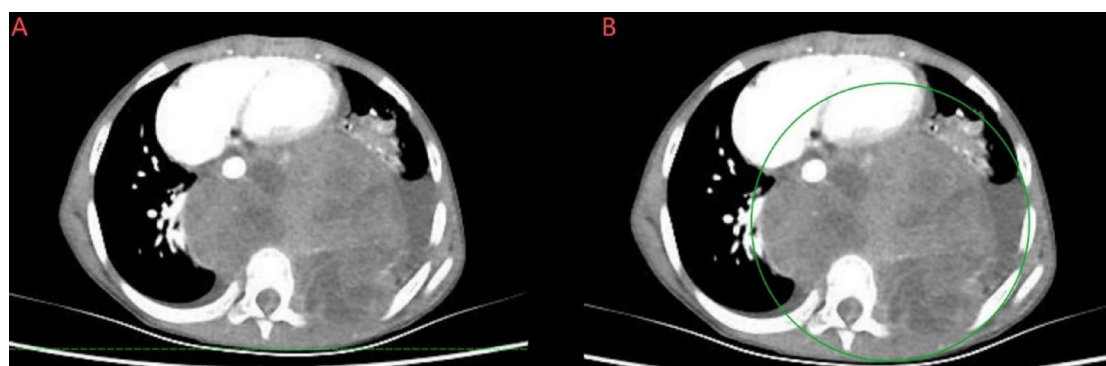
All patients underwent either 16-slice or 64-slice spiral computed tomography (CT). Each patient fasted for more than four hours prior to the CT scan. Due to significant individual differences among children, the parameters for contrast-enhanced CT (CECT) may vary. The key parameters include:

- Detector collimation: 1 mm.
- Pitch: 0.9.
- Gantry rotation: 0.5 s.
- Tube voltage: 80–120 kV.
- Tube current: 110–240 mA.
- Matrix size: 512 × 512.
- Slice thickness: 0.625–5 mm.
- Reconstruction interval: 1 mm.

Scans were performed at fixed equilibrium times of 25 s and 65 s during the arterial and portal venous phases, respectively, following the intravenous injection of 100 mL of iodinated contrast at a rate of 3 mL/s using an autoinjector. The polyphase-enhanced CT images were exported from the Image Archiving and Communication System (PACS) and saved as BMP files. We manually extracted the three-phase CT image slice that displayed the largest cross-sectional area of the tumor, which was then used as the CT feature image for the corresponding patient. Finally, we applied a transformation function to normalize all CT images and resized them to 224 × 224 pixels to meet the input requirements for the model.

### Image labeling

YOLO (You Only Look Once) is a deep learning algorithm utilized for target detection and classification. Consequently, it is essential to manually delineate the target regions, as illustrated in Fig. 2. We employed a labeling tool developed in Python to annotate the tumor lesion regions in all computed tomography (CT)



**Fig. 2.** Picture annotation (A) Image prior to labeling. (B) Image following the application of labels.



images. To ensure the accuracy of the annotations, two radiologists, each possessing over ten years of clinical experience, were engaged to label each case independently.

## Methods

### Inverse residual attention module (iRMB)

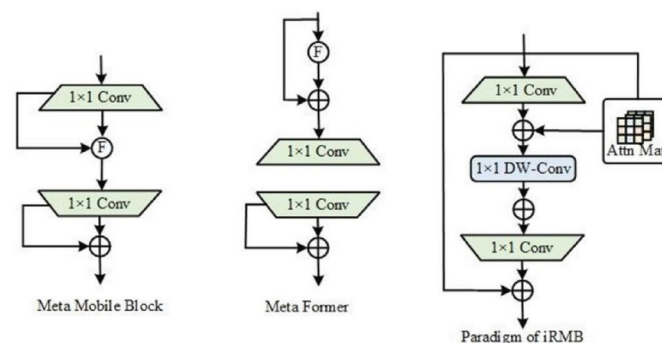
Inverse Residual Attention Module (iRMB) is a deep learning architecture designed for image classification and object detection, which integrates the benefits of both inverse residual and attention mechanisms. In iRMB, the term refers to the inversion of the conventional residual block, where the convolution operation and batch normalization are placed at the end of the block. This configuration enhances the model's capacity for nonlinear representation while simultaneously reducing the number of parameters. Furthermore, iRMB also incorporates an attention mechanism that employs two parallel convolutional branches: one dedicated to global information extraction and the other to local information extraction. The outputs from these two branches are subsequently fused through adaptive pooling and a channel attention mechanism, enabling the model to prioritize significant features. The architecture of the model is illustrated in Fig. 3.

By integrating the lightweight characteristics of Convolutional Neural Networks (CNN) with the dynamic modeling capabilities of Transformers, the Inverted Residual Mobile Block (iRMB) is well-suited for intensive prediction tasks on mobile devices. The iRMB employs an inverted residual block design alongside the Meta-Mobile Block. This approach not only extends the traditional Inverted Residual Block (iRMB) of CNNs to an attention-based model but also enhances the model's flexibility and efficiency through the application of various scaling ratios and efficient operators. Consequently, this study incorporates the iRMB into the C2F board of YOLOv8, thereby enabling the model to focus more effectively on critical features<sup>26</sup>.

Neuroblastoma has a variety of imaging manifestations, and its tumor morphology may be irregular, sometimes with unclear borders, and may occur in various parts of the sympathetic nervous system, such as the adrenal glands, the neck, the chest, the abdomen, etc. The relationship between different parts and the surrounding tissues varies, which makes it difficult to accurately identify and classify the tumor. The relationship between the tumor and the surrounding tissues varies in different parts of the body, which increases the difficulty of accurate identification and classification. Tumors may have necrosis, hemorrhage, calcification, and other components that are intertwined in the image, making it difficult to accurately extract and analyze the features for classification and identification<sup>27</sup>. iRMB can capture multi-dimensional features, and can learn and extract neuroblastoma image features from both spatial and channel dimensions at the same time. For example, in CT images, attention can be paid to the spatial features such as shape, size, and location of the tumor, as well as the characteristics of different tissues and cell types associated with the tumor in the channel, so as to more comprehensively portray the characteristics of the tumor, and provide rich information for accurate identification. iRMB can highlight the key features, and through the attentional mechanism, automatically learn and amplify the feature information in the images of neuroblastoma, which is significant to the classification and identification, such as the information of the tumor in the channel, the information of the neuroblastoma in the image of the neuroblastoma. iRMB can highlight key features, through the attention mechanism, it can automatically learn and amplify the feature information in the neuroblastoma image that is important for classification and recognition, such as the edge of the tumor, the internal texture, etc., and at the same time, it suppresses the interference information in the background and irrelevant areas, so that the model focuses on the tumor features and improves the recognizability of the features<sup>28</sup>.

### Centralized feature pyramid enhanced visual computing

The centered feature pyramid EVC module primarily facilitates the fusion of local and global features. It comprises two parallel interconnected components: a lightweight multi-layer perceptron (MLP) and a learnable visual centroid mechanism. The lightweight MLP is designed to capture global long-range dependencies, thereby extracting global information, while the learnable visual centroid mechanism focuses on preserving critical local region information from the input image. By integrating both global and local feature information, the EVC module effectively generates rich visual centroid data for subsequent global centralized conditioning (GCR). This data is instrumental in enhancing shallow feature conditioning, which can effectively obtain a comprehensive



**Fig. 3.** A schematic diagram illustrating the structure of the inverse residual block (iRMB).

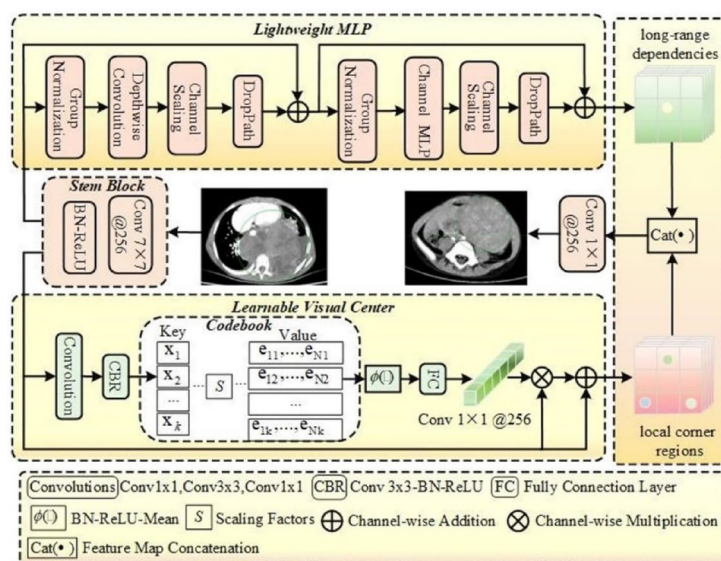
and discriminative feature representation. The introduction of the EVC module into the YOLOv8 architecture not only achieves efficient feature fusion, but also enhances the model's accuracy.

Traditional feature fusion is achieved by multi-layer feature pyramids and multi-scale fusion, but this usually increases the loss of feature resolution and computational cost. The introduction of the CFP and EVC modules within the centralized feature pyramid of YOLOv8 enhances the quality of the feature pyramid while preserving low computational complexity. The CFP and EVC modules are designed to effectively manage targets of varying sizes while simultaneously emphasizing both globally and locally significant details. Consequently, YOLOv8 demonstrates superior performance in detecting small-sized objects or targets within complex scenes, making it particularly suitable for intricate medical imaging classification tasks. The structure of YOLOv8 is illustrated in Fig. 4<sup>29</sup>.

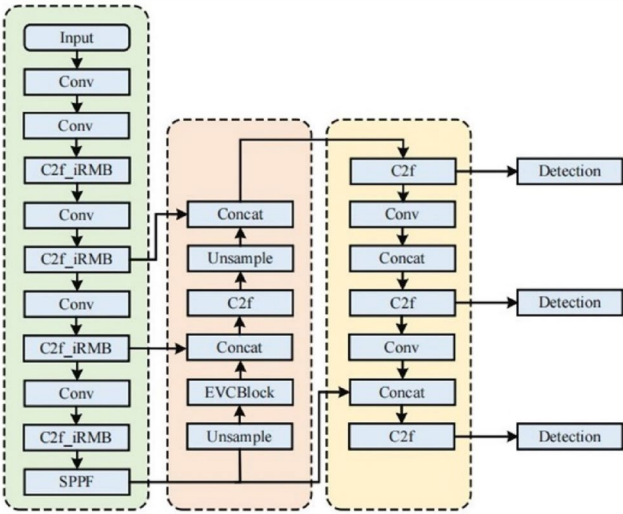
### YOLOv8-IE

YOLOv8 represents a significant advancement in the domain of object detection, building upon the strengths of its predecessors within the YOLO series. As a result, YOLOv8 exhibits enhanced performance in identifying small objects or targets within complex scenes, making it especially well-suited for intricate medical imaging classification tasks. The architecture of YOLOv8 is depicted in Fig. 4. This innovative design not only enhances the model's detection accuracy but also increases its adaptability. Furthermore, YOLOv8 employs a decoupled head structure to distinctly separate the tasks of classification and detection. The network architecture of YOLOv8 effectively extracts and fuses features by integrating design principles from CSPNet and PANet. Additionally, the decoupling of the head structure, along with the implementation of the DFL loss function, contributes to improved prediction accuracy and classification performance in relation to object localization. With these improvements, YOLOv8 can now perform exceptionally well in target detection tasks, obtaining excellent detection accuracy and speed<sup>30–33</sup>.

Neuroblastoma in the early stage is small in size, and the imaging performance may be atypical, only showing localized nodules or slight density changes, which are easy to be missed or misdetected. YOLOv8 abandons the traditional anchor boxes, and directly predicts the center coordinates, width, and height of the target. This approach avoids the problem of missed or erroneous detection caused by the mismatch between the preset size and scale of anchor boxes in the early identification of neuroblastoma, and is more adaptable to the scale change and irregular shape of early neuroblastoma, and can locate the small target more accurately. Advanced neuroblastoma often invades the surrounding tissues and organs and develops distant metastasis, and the complexity of the tumor's imaging manifestations makes it difficult to accurately determine the tumor's origin and extent of invasion. Some metastatic tumors of other malignant tumors may also involve the common metastatic sites of neuroblastoma, such as the liver and other organs, and their imaging manifestations may be similar to those of neuroblastoma metastatic tumors, which increases the difficulty of classification. During the training process of YOLOv8, the mosaic data enhancement in the input module enables the model to better adapt to complex neuroblastoma images with different morphologies, sizes, and locations. The multi-scale prediction of YOLOv8 can predict the target at multiple scales, so as to extract more effective features and improve the detection accuracy of neuroblastoma. Due to the fact that ganglioneuroblastoma has both neuroblastoma and ganglioneuroma components, the imaging performances are complicated and diverse, and it is easy to misdetect it as simple neuroblastoma or ganglioneuroblastoma. Path Aggregation Network (PANet) of YOLOv8 is used as a feature fusion module to fuse the features of different levels of the backbone network of the model, transfer the semantic information of the high-level features to the low-level network, and at the same



**Fig. 4.** A schematic diagram illustrating the structure of the centralized feature pyramid within the EVC framework.



**Fig. 5.** A schematic diagram illustrating the structure of the YOLOv8-IE model.

Tumour subtypes	Quantity	Percentage
Neuroblastoma	141	60.5%
Ganglioneuroma	37	15.8%
Ganglioneuroblastoma	55	23.6%

**Table 1.** The statistical data for the NT-CT dataset.

time, transfer the semantic information of the low-level features to the low-level network. Network and transfer the semantic information of the low-level features to the low-level features to the high level, so that the model can learn richer and more representative target features, and thus improve the detection ability of ganglion cell neuroblastoma<sup>27,34–36</sup>.

Given the benefits of YOLOv8 for detection and classification tasks, we improved the YOLOv8 framework by adding the centered feature pyramid (EVC) structure and the inverse residual attention module (iRMB) in the c2f slab. The integration of these two components facilitates the efficient detection of tumor targets. The specific architecture is illustrated in Fig. 5.

**Results**  
**Parameter configuration**

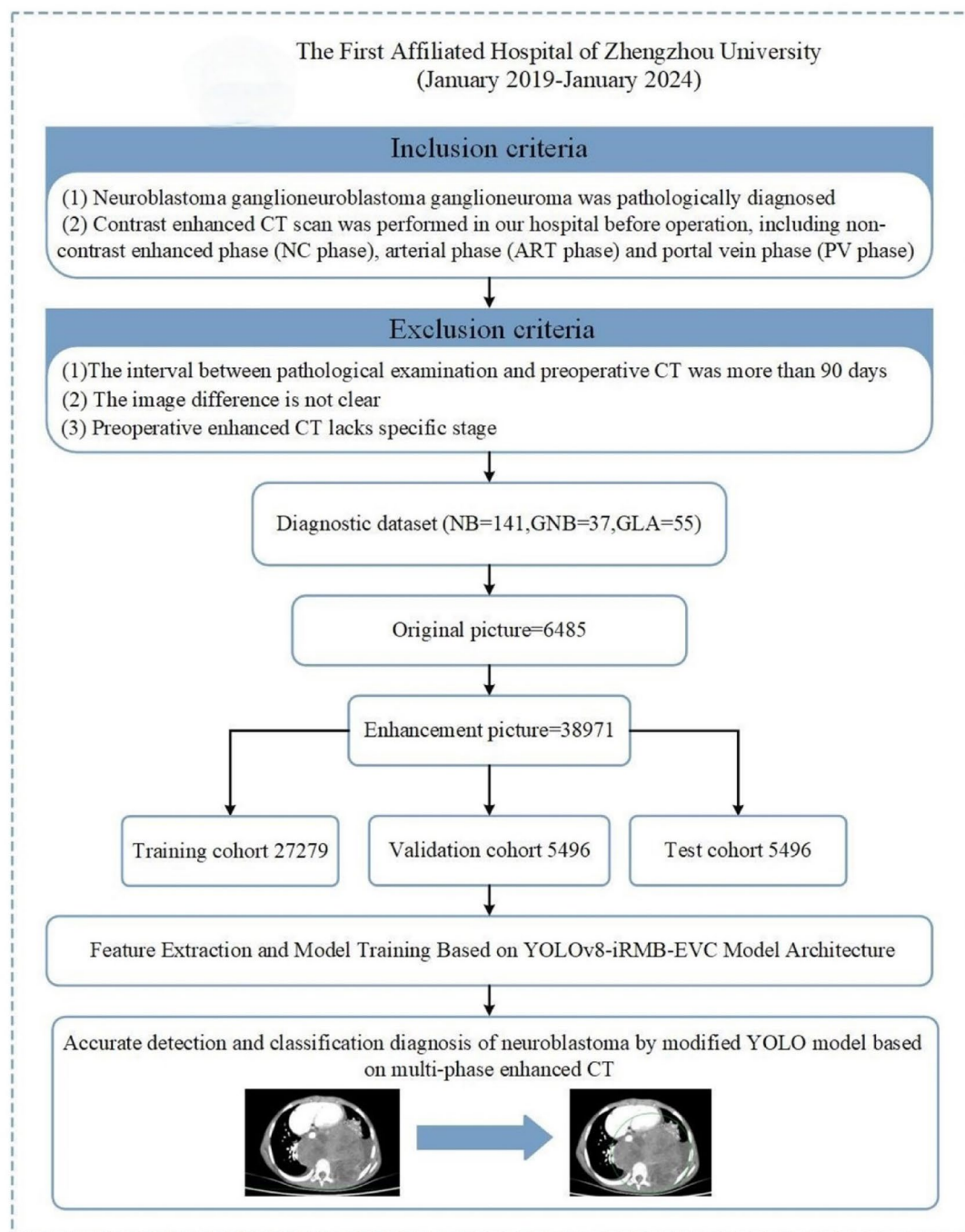
In this experiment, all models were trained on the Ubuntu 18.04 operating system. The GPU utilized was the NVIDIA GeForce RTX 4090, which is equipped with 24 GB of video memory. The CUDA version employed was 11.1, while Python version 3.8 and PyTorch version 1.9.0 were used as the deep learning framework.

Utilizing a computed tomography (CT) dataset of neuroblastoma patients admitted to Zhengzhou University’s First Affiliated Hospital between 2019 and 2024, we examined the identification and categorization of neuroblastoma in this study. The dataset included 141 cases of neuroblastoma, 37 cases of ganglioneuroblastoma, and 55 cases of both neuroblast and neuroectodermal components, as detailed in Table 1. Frame-by-frame labeling of the CT images was done with great care; neuroblastoma was labeled as “NT”, ganglioneuroma as “GLA”, and ganglioneuroblastomaas “GNT”. Out of the tagged photos, 38,971 images and their associated JSON files were produced. The dataset was partitioned into training, testing, and validation sets in a ratio of 8:1:1, resulting in 27,279 images for the training set, and 5496 images each for the test and validation sets. The experimental parameters for the training process are presented in Table 2. The experimental workflow of this study is illustrated in Fig. 6.

**Evaluation of indicators**  
*mAP0.5*

Target detection algorithms are typically assessed using metrics such as Precision, Recall, mean Average Precision at IoU threshold 0.5 (mAP0.5), and mean Average Precision across multiple IoU thresholds (mAP0.5:0.95). Precision quantifies the likelihood that an image is accurately detected among samples with positive predictions, while Recall indicates the probability that all targets within the image are correctly identified among samples with true positive predictions. The mean Average Precision (mAP) serves as a comprehensive measure of the effectiveness of the target detection model. The specific formulas for these indices are presented below. In Eqs. (1) and (2), true positives, which refer to the total number of target images accurately identified by the

Parameter name	Parameter values
Initial learning rate	0.001
Batch size	16
Momentum	0.937
Non-extremely suppressed	0.5
Learning rate decay factor	0.1
The number of iterations	300

**Table 2.** Description of network parameters.**Fig. 6.** A schematic diagram illustrating the process employed in this study.



model. False positives represent the total number of target images incorrectly identified as positive by the model. False negatives correspond to the total number of target images that were not identified as positive by the model. Furthermore, Eq. (3) represents Precision, the Recall rate, and the total number of sample categories.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$mAP = \frac{\sum_{n=1}^N \int_0^1 p(r) dr}{N} \quad (3)$$

### Confusion matrix

The confusion matrix offers a comprehensive overview of the results, detailing the counts of true positives, true negatives, false positives, and false negatives for each category. This matrix provides valuable insights into the model's performance across various categories. The P-R curve serves as another visual representation of the test machine's performance. By analyzing the shape and dynamics of the curves, one can assess the model's performance in terms of precision and recall at different thresholds, thereby facilitating the selection of the optimal threshold for prediction.

### F1 curves

F1 curves are usually obtained by calculating F1 scores at different confidence thresholds. In object detection, a confidence threshold is used to determine whether a prediction is considered valid object detection. The abscissa usually represents the confidence threshold, which generally ranges from 0 to 1. As the confidence threshold increases, the number of positive examples (detected objects) predicted by the model will increase. The ordinate represents the F1 score, which is an indicator that combines precision and recall. Its value is also between 0 and 1. The higher it is, the better the balance between precision and recall of the model is.

### The loss function

In the YOLO model, the loss function is mainly composed of three parts: bounding box loss, class loss, and confidence loss. The abscissa usually represents the iteration of the training (Epochs) or the number of training steps. With the abscissa, the model is constantly training and updating parameters. The ordinal coordinate represents the loss value, which is a synthesis of position loss, category loss, and confidence loss. The lower the loss value, the closer the predicted results of the model are to the real results, and the better the performance of the model. Figure 7 below shows the loss curve of our proposed model YOLO-IE.

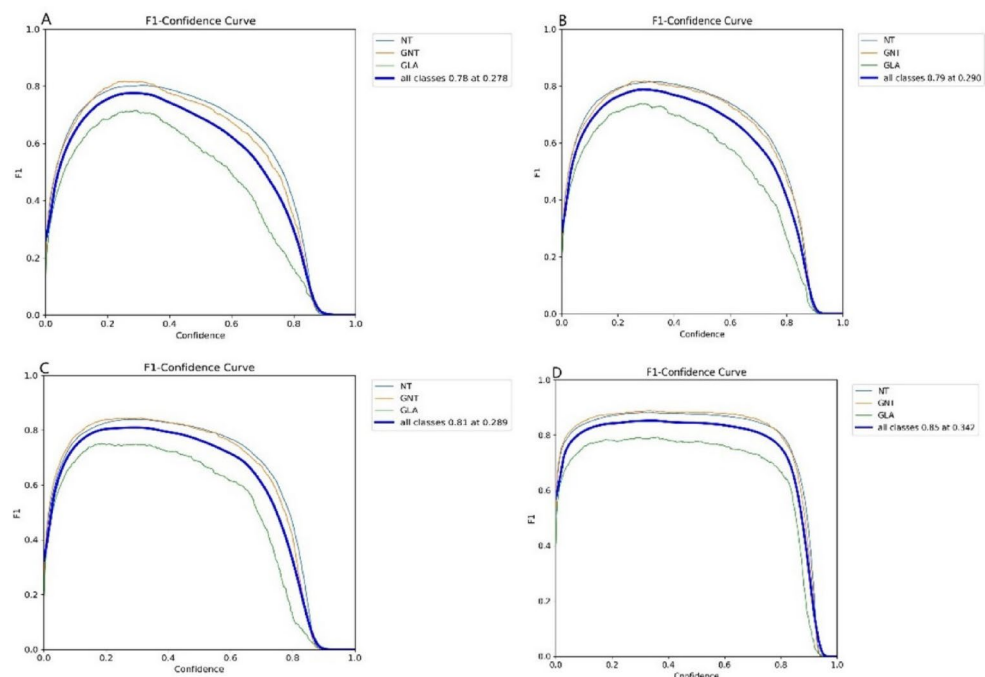


Fig. 7. YOLO-IE-loss graph.

Submodels	Pr	Rc	mAP-0.5	mAP-0.5:0.95
YOLOv8	0.82	0.74	0.81	0.49
YOLOv8-iRMB	0.82	0.76	0.82	0.54
YOLOv8-EVC	0.85	0.77	0.85	0.57
YOLOv8-IE	0.86	0.84	0.89	0.72

**Table 3.** The comprehensive training outcomes are derived from both the training and validation datasets.

	Positive projections	Predicting negative numbers
Actual positive numbers	TP	FN
Actual negative numbers	FP	TN

**Table 4.** Formulas for the confusion matrix.

The visual presentation of the prediction results

The visual presentation of the prediction results of the model is also essential. It shows how the model predicts the tumor target in the image in a batch of the validation set. In this image, we can visually see which tumors the model believes are in the image, as well as the location and category of these tumors in the image.

Analysis and comparison of results

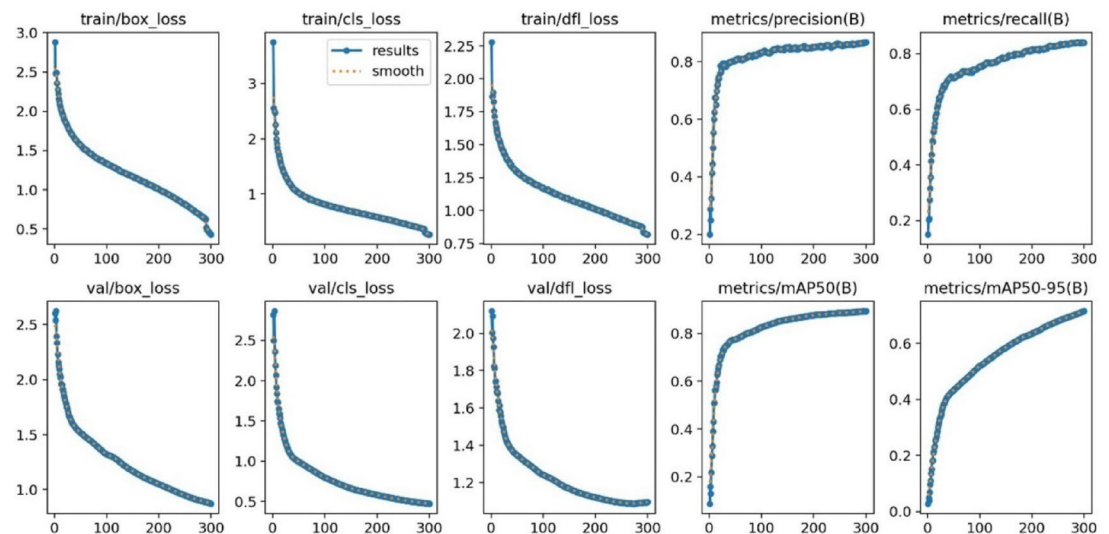
In this paper, we present a series of enhancements to the YOLOv8 model, culminating in the development of a high-precision variant, designated as YOLOv8-IE. The mean Average Precision (mAP) of the YOLOv8-IE model is 89.3%, representing a substantial improvement over the original YOLOv8 model, which achieved a mAP of 81.4%. This reflects an enhancement of 7.9% compared to the pre-improved version. Notably, we also evaluated the performance of YOLOv8-iRMB, which incorporates the iRMB module independently, as well as YOLOv8-EVC, which integrates the EVC module independently.

The results indicate that the mean Average Precision (mAP) of the YOLOv8-iRMB model is 82.2%, which represents an improvement of 0.8% over the baseline model. This enhancement may be attributed to the inverse residual structure, which reduces computational demands while simultaneously increasing the model's expressiveness. This is achieved by initially expanding the number of channels before performing the convolution operation, in contrast to traditional inverse residual architectures. Such a structure effectively preserves the feature information of the input image and facilitates the detection of subtle lesion characteristics, making it particularly suitable for medical classification tasks. Furthermore, the model incorporates an attention mechanism that adaptively focuses on critical regions, such as lesion areas or key features, within the input image. This not only bolsters the model's robustness and generalization capabilities when dealing with complex medical image data but also significantly enhances the representation of lesion features, thereby improving classification accuracy.

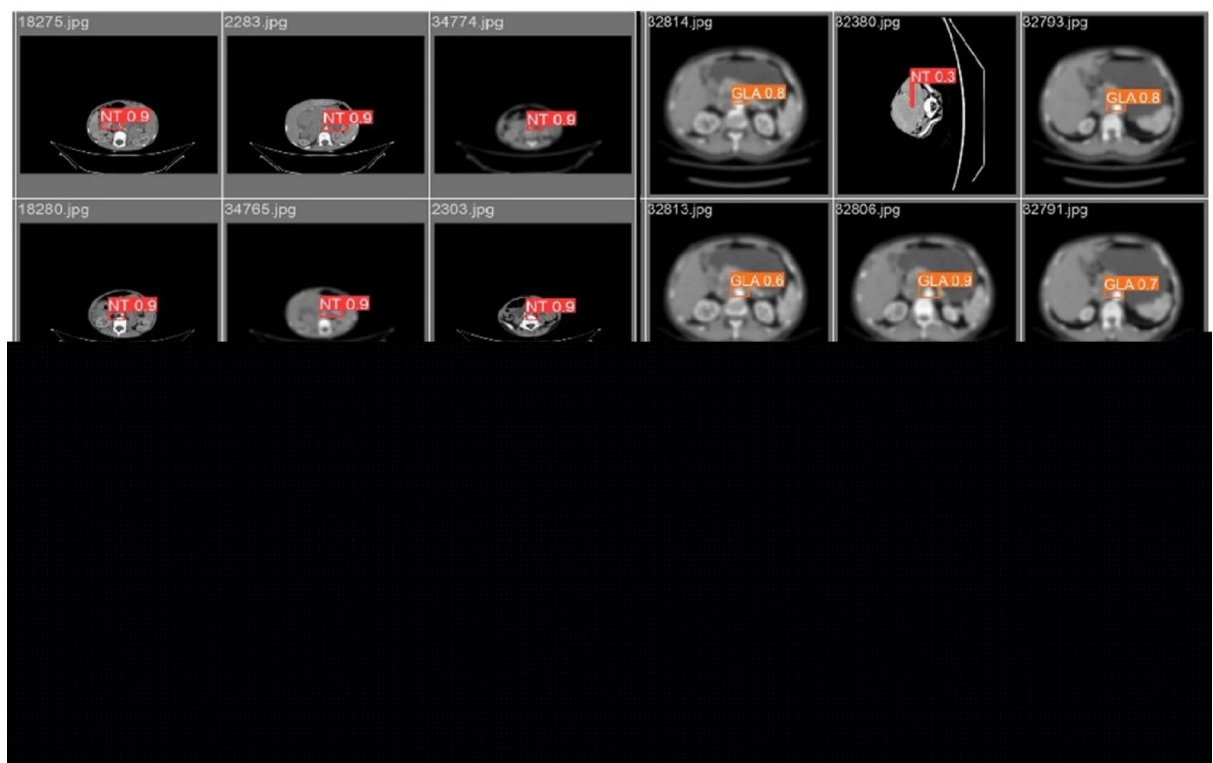
The mAP of the YOLOv8-EVC model is reported at 85%, which is 3.6% higher than that of the baseline model. This improvement is attributed to the centered feature pyramid structure, which effectively extracts features at various scales, thereby enhancing the representation of objects of differing sizes and shapes within the image. In the context of medical classification tasks, this capability allows the model to more effectively capture multi-scale features of the lesion region, thereby improving classification accuracy. Additionally, the centered feature pyramid structure facilitates the fusion of features across different scales, optimizing the utilization of multi-scale information present in the image. This comprehensive understanding of the lesion area contributes to increased reliability in classification outcomes. The YOLOv8-IE model combines the benefits of the centered feature pyramid with the inverse residual iRMB, showing notable gains in the diagnostic tasks related to neuroblastoma classification and detection. The overall performance results of the four sub-models, as evaluated on the training and validation sets, are presented in Table 3.

The confusion matrix is a specific square tabular representation utilized primarily for visualizing the performance of algorithms, particularly in classification tasks. It offers an intuitive framework for understanding the accuracy of model predictions, the nature of errors, and the degree of confusion among categories. This matrix provides a comprehensive overview of model performance. Table 4 contains the formulas related to the confusion matrix, where TP denotes true positives, FP represents false positives, FN indicates false negatives, and TN signifies true negatives. In this context, the tumor categories were systematically labeled, with each letter identifier corresponding to a specific tumor type. This systematic labeling approach provides a clear indication of the classification outcomes of the model. A detailed examination of the confusion matrix reveals that the performance of the YOLOv8-IE model is commendable. Specifically, the model accurately identified 5225 images categorized as "NT", classified 1127 images as "GNT", and recognized 460 images as "GLA", as illustrated in Fig. 8. The Precision-Recall (P-R) curve for the proposed model is shown in Fig. 9.

The F1 curve can be used to compare the four YOLO models we proposed (such as YOLOv8, YOLOv8-iRMB, YOLOv8-EVC, YOLOv8-IE, etc.). Under the same data set and task, we compared their F1 curves and found that in this task, the YOLO-IE model has better overall performance in precision and recall (Figs. 10, 11).



**Fig. 8.** Confusion matrix. **(A)** The YOLOv8 baseline model. **(B)** The YOLOv8-iRMB model after the introduction of the inverse residual attention iRMB in YOLOv8. **(C)** The YOLO-EVC with the introduction of the centred eigenpyramid EVC module. **(D)** The YOLOv8-IE.

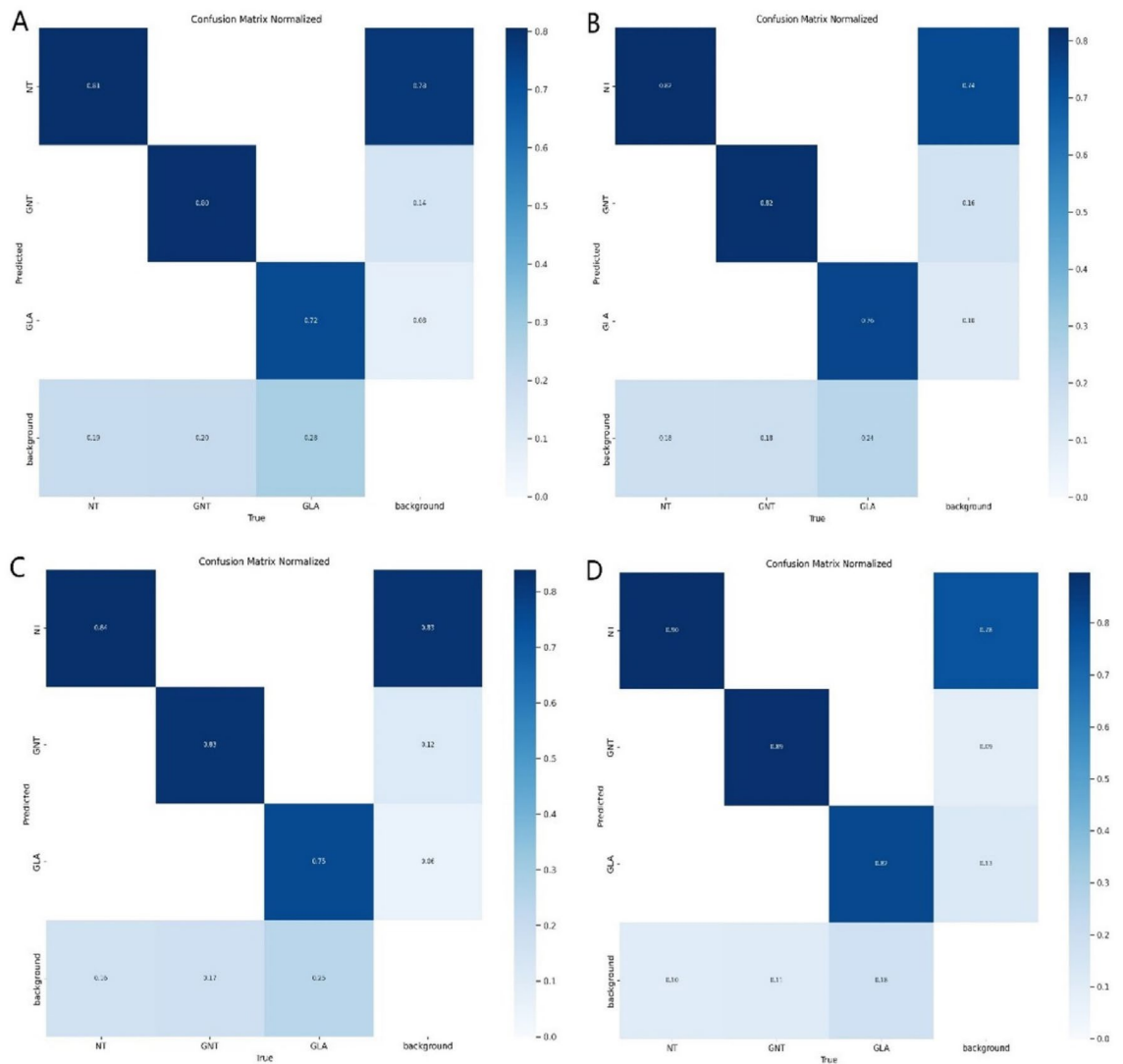


**Fig. 9.** The Precision-Recall (P-R) curves for the YOLOv8-IE model.

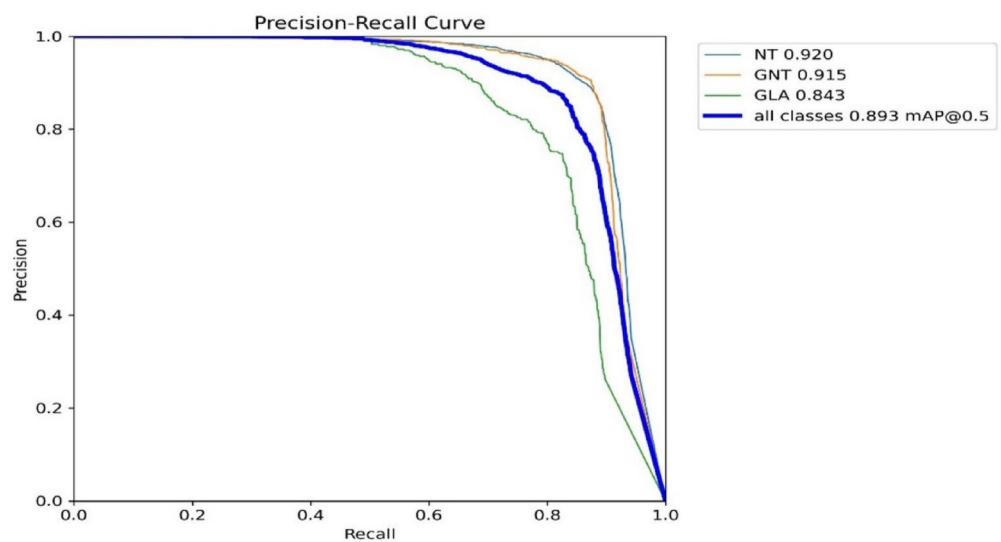
### Multi-model comparison experiments

In order to verify the detection performance of the proposed model for neuroblastoma subtypes, the model is compared with the current mainstream targeting algorithms SSD, Faster R-CNN, YOLOv3-tiny, YOLOv5-s, YOLOv7-tiny, YOLOv10, and the improved YOLOv8-IE based on the same dataset with the experimental environment. The experimental results are shown in Table 5.

SSD is a classical single-stage target detection and classification model. It usually uses pre-trained convolutional neural networks such as VGG16, Res Net, Mobile Net, etc., as the feature extractor, and extracts the feature maps of the image in different layers of the base network to obtain the multi-scale features. Faster



**Fig. 10.** F1 curves of four models. (A) YOLOv8. (B) YOLOv8-iRMB. (C) YOLOv8-EVC, DYOLO-IE.



**Fig. 11.** Visualization of YOLO-IE model prediction results.



Model	Params/M	Flops/G	mAP0.5	Size/MB
SSD	24.5	59.4	0.75	104.4
Faster-RCNN	165	47.9	0.82	186.5
YOLOv3-tiny	8.7	12.9	0.82	17.4
YOLOv5s	7.1	16.5	0.86	14.4
YOLOv7-tiny	6.02	13.0	0.83	12.3
YOLOv10	8.04	24.5	0.84	16.6
Ours	8.5	11.5	0.89	14.1

**Table 5.** Multi-model comparison experiments.

R-CNN is a classical two-stage target detection algorithm, which usually uses pre-trained deep convolutional neural networks such as VGG16, Res Net, etc., as the backbone for feature extraction. It usually uses pre-trained deep convolutional neural networks such as VGG16, Res Net, etc., as the backbone for feature extraction. After a series of convolutional layers and pooling layers, different levels of feature maps are obtained.

Compared with Faster-RCNN and SSD, YOLO-IE has less parameter counts and computational power, while Precision, Recall, and mAP are greatly improved, and mAP is improved by 14% and 7%, respectively. At the same time, Precision, Recall, and mAP are greatly improved, and mAP is improved by 14% and 7% points, respectively. YOLO-IE is lower than YOLOV3-tiny in terms of parameter and computation, and the mAP value is 7% points higher than that of YOLOv3-tiny, which is an obvious improvement in precision. YOLO-IE is more accurate than YOLOv5-s. YOLOv7-tiny and YOLOv10, and YOLOv7-tiny and YOLOv10 are more precise than YOLOv7-tiny, YOLO-IE, compared with YOLOv5-s, YOLOv7-tiny and YOLOv10, the number of parameters and computation amount are slightly increased, but the mAP accuracy is increased by 3%, 6%, and 5% points, respectively, which improves the accuracy of the detection and meets the demand of real-time detection on the premise of slightly increasing the number of parameters and computation amount. In summary, YOLO-IE has excellent performance in terms of parameter count and accuracy, and also has certain advantages in terms of computational resources and storage space. YOLO-IE can be applied to the detection of neuroblastoma subtypes. In this paper, we summarize the results of YOLO-IE.

Discussion

It is widely recognized that a definitive diagnosis is a crucial prerequisite for the individualized treatment of patients with tumors. Complete surgical resection of primary tumors constitutes a significant component of the treatment for abdominal solid tumors in pediatric patients and serves as a vital method for establishing a definitive diagnosis<sup>37</sup>. Pathological examination of surgically resected specimens is currently regarded as the gold standard for diagnosing neuroblastoma; however, this approach is more invasive. Distinguishing between neuroblastoma and ganglioneuroma, as well as identifying cases with ganglioneuroblastoma that show favorable clinical regression, poses challenges due to their similar clinical manifestations and laboratory findings. To address this issue, we have developed a preclinical model utilizing multiphase enhanced computed tomography (CT) for the preoperative detection and classification of neuroblastoma. This advancement enhances the management of neuroblastoma and represents a groundbreaking development in the field, potentially transforming our clinical diagnostic paradigm through the application of artificial intelligence.

The YOLO model plays a crucial role in this study by facilitating image segmentation, specifically by delineating tumor regions from background areas. However, tumor regions within an image are often continuous and exhibit complex geometries. Consequently, the target frames identified by the model, whether through automated detection or manual labeling, frequently do not correspond precisely to the actual tumor lesion areas. To develop models that achieve higher precision and recall in tumor region segmentation, convolutional neural network (CNN) architectures, such as U-Net or Deep Lab, may be employed. Nevertheless, the object region labeling associated with these models is generally more intricate than that of the YOLO model.

The YOLO model is an advanced object detection algorithm that has demonstrated good performance in the classification and diagnosis of neuroblastoma subtypes, but it also has some limitations. The acquisition of medical image data is often subject to various restrictions, such as patient privacy protection and high data collection costs. Compared to other fields, the number of medical images available for training is relatively small, which may lead to the YOLO model being unable to fully learn the characteristics of various diseases during the training process, thereby affecting its diagnostic accuracy and generalization ability. To pursue real-time detection speed, the YOLO model has undergone certain simplifications in its network structure, which may result in insufficient learning capability for subtle lesion features in abnormal sample images. Lesions in medical images often exhibit diversity and complexity; for instance, the morphology, size, location, boundaries of tumors, and their relationship with surrounding tissues all require precise identification and judgment. The YOLO model may not be able to fully capture these abnormal feature information<sup>34</sup>.

Medical diagnosis relies not only on image information but also requires a comprehensive assessment that includes the patient's clinical symptoms, medical history, family history, and other information. The YOLO model focuses solely on object detection within images and cannot directly integrate other clinical context information, which somewhat limits its practical application value in clinical diagnosis. In the medical field, doctors need to have a clear understanding and explanation of the diagnostic results in order to communicate them to patients and other healthcare personnel. However, as a deep learning model, the internal decision-

making process of the YOLO model is complex and difficult to interpret, making it challenging for doctors to intuitively understand how the model arrives at its diagnostic conclusions. This may affect the doctors' trust and acceptance of the model's results.

In summary, our proposed YOLOv8-IE model has an impressive accuracy of 89.3%. This suggests that the model can effectively identify and predict the presence of neuroblastoma cells in CT images, and is an important tool to assist diagnosis.

## Conclusion and prospect

In this study, we developed the YOLOv8-IE model utilizing multiphase-enhanced computed tomography (CT) for the differential diagnosis of various subtypes of neuroblastoma. The YOLOv8-IE model represents a promising tool for clinicians, facilitating the clarification of tumor diagnoses and enabling a comprehensive assessment of the child's condition. This capability can significantly inform clinical decision-making regarding treatment options, ultimately enhancing the overall prognosis for children diagnosed with neuroblastoma.

Future research endeavors could investigate the applicability of the proposed model to other imaging modalities, such as magnetic resonance imaging (MRI), positron emission tomography (PET), and ultrasound. Such exploration would not only yield a more comprehensive understanding of the model's suitability and validity across a broader spectrum of medical imaging data but also enable a thorough evaluation of the model's performance. Each imaging modality possesses distinct strengths and characteristics; thus, the exploration of multiple imaging modalities not only enhances the model's capabilities but also has the potential to expand its impact within the realm of medical image classification. In conclusion, our YOLOv8-IE model demonstrates significant potential in advancing medical image classification and substantially improving clinical diagnostic processes. However, this study is not without limitations. It did not assess the model's performance on alternative datasets, and the dataset employed may not adequately represent the full spectrum of patient types, which could lead to biased predictions. Additionally, the study did not address the costs associated with training and implementing the model, which may pose challenges for its application in clinical settings. Despite these limitations, future work will focus on enhancing the model's precision and recall, exploring various model architectures to achieve higher performance metrics, and investigating the impact of U-Net on segmentation in CT images.

## Data availability

The datasets generated and analyzed during the current study are not publicly available due to privacy and ethical restrictions; however, they can be obtained from the corresponding author upon reasonable request.

Received: 2 November 2024; Accepted: 21 April 2025

Published online: 26 April 2025

## References

1. Maris, J. M. Recent advances in neuroblastoma. *N Engl. J. Med.* **362**(23), 2202–2211. <https://doi.org/10.1056/NEJMra0804577> (2010).
2. Maris, J. M., Hogarty, M. D., Bagatell, R. & Cohn, S. L. Neuroblastoma. *Lancet* **369**(9579), 2106–2120. [https://doi.org/10.1016/S0140-6736\(07\)60983-0](https://doi.org/10.1016/S0140-6736(07)60983-0) (2007).
3. Hahn, U. & Oleynik, M. Medical information extraction in the age of deep learning. *Yearb Med. Inf.* **29**(1), 208–220. <https://doi.org/10.1055/s-0040-1702001> (2020).
4. Mathivanan, S. K. et al. Employing deep learning and transfer learning for accurate brain tumor detection. *Sci. Rep.* **14**(1), 7232. <https://doi.org/10.1038/s41598-024-57970-7> (2024).
5. Cè, M. et al. Artificial intelligence in brain tumor imaging: a step toward personalized medicine. *Curr. Oncol.* **30**(3), 2673–2701. <https://doi.org/10.3390/curroncol30030203> (2023).
6. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 779–788. <https://doi.org/10.1109/CVPR.2016.91> (2016).
7. Hniewa, M. & Radha, H. Integrated multiscale domain adaptive YOLO. *IEEE Trans. Image Process.* **32**, 1857–1867. <https://doi.org/10.1109/TIP.2023.3255106> (2023).
8. Islam, M. M., Uddin, M. R., Ferdous, M. J., Akter, S. & Akhtar, M. N. BdSLW-11: dataset of Bangladeshi sign Language words for recognizing 11 daily useful BdSL words. *Data Brief.* **45**, 108747. <https://doi.org/10.1016/j.dib.2022.108747> (2022).
9. Chen, X., Zheng, H., Tang, H. & Li, F. Multi-scale perceptual YOLO for automatic detection of clue cells and Trichomonas in fluorescence microscopic images. *Comput. Biol. Med.* **175**, 108500. <https://doi.org/10.1016/j.compbiomed.2024.108500> (2024).
10. Itti, L., Koch, C. & Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(11), 1254–1259. <https://doi.org/10.1109/34.730558> (1998).
11. Li, R., Zheng, S., Duan, C., Yang, Y. & Wang, X. Classification of hyperspectral image based on double-branch dual-attention mechanism network. *Remote Sens.* **12**(3), 582. <https://doi.org/10.3390/rs12030582> (2020).
12. Gong, H. et al. Swin-transformer-enabled YOLOv5 with attention mechanism for small object detection on satellite images. *Remote Sens.* **14**(12), 2861. <https://doi.org/10.3390/rs14122861> (2022).
13. Li, H. et al. SCAttNet: semantic segmentation network with Spatial and channel attention mechanism for high-resolution remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **18**(5), 905–909. <https://doi.org/10.1109/LGRS.2020.2988294> (2021).
14. Dong, M., Fang, Z., Li, Y., Bi, S. & Chen, J. AR3D: attention residual 3D network for human action recognition. *Sensors* **21**(5), 1656. <https://doi.org/10.3390/s21051656> (2021).
15. Hu, J., Shen, L. & Sun, G. Squeeze-and-Excitation networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 7132–7141. <https://doi.org/10.1109/CVPR.2018.00745> (2018).
16. Carion, N. et al. End-to-end object detection with transformers. In *Proceedings of the 16th European Conference on Computer Vision (ECCV)* 213–229. [https://doi.org/10.1007/978-3-030-58452-8\\_13](https://doi.org/10.1007/978-3-030-58452-8_13) (2020).
17. Woo, S., Park, J., Lee, J. Y. & Kweon, I. S. CBAM: convolutional block attention module. In *Proceedings of the 15th European Conference on Computer Vision (ECCV)* 3–19. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1) (2018).
18. Li, Z. et al. Cross-layer feature pyramid network for salient object detection. *IEEE Trans. Image Process.* **30**, 4587–4598. <https://doi.org/10.1109/TIP.2021.3072811> (2021).

19. Song, Y., Tang, H., Zhao, M., Sebe, N. & Wang, W. Quasi-equilibrium feature pyramid network for salient object detection. *IEEE Trans. Image Process.* **31**, 7144–7153. <https://doi.org/10.1109/TIP.2022.3220058> (2022).
20. Dang, J., Tang, X. & Li, S. HA-FPN: hierarchical attention feature pyramid network for object detection. *Sensors* **23**(9), 4508. <https://doi.org/10.3390/s23094508> (2023).
21. Gope, H. L., Fukai, H., Ruhad, F. M. & Barman, S. Comparative analysis of YOLO models for green coffee bean detection and defect classification. *Sci. Rep.* **14**(1), 28946. <https://doi.org/10.1038/s41598-024-78598-7> (2024).
22. Rong, R. et al. A deep learning approach for histology-based nucleus segmentation and tumor microenvironment characterization. *Mod. Pathol.* **36**(8), 100196. <https://doi.org/10.1016/j.modpat.2023.100196> (2023).
23. Xie, H. et al. SMLS-YOLO: an extremely lightweight pathological myopia instance segmentation method. *Front. Neurosci.* **18**, 1471089. <https://doi.org/10.3389/fnins.2024.1471089> (2024).
24. Zhang, B. et al. An improved microaneurysm detection model based on swinir and YOLOv8. *Bioengineering* **10**(12), 1405. <https://doi.org/10.3390/bioengineering10121405> (2023).
25. Dinesh, M. G., Bacanin, N., Askar, S. S. & Abouhawwash, M. Diagnostic ability of deep learning in detection of pancreatic tumour. *Sci. Rep.* **13**(1), 9725. <https://doi.org/10.1038/s41598-023-36886-8> (2023).
26. Zhang, J. et al. Rethinking mobile block for efficient attention-based models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* 1389–1400. <https://doi.org/10.1109/ICCV51070.2023.00134> (2023).
27. Sheng, J. et al. Evaluation of clinical and imaging features for differentiating rhabdomyosarcoma from neuroblastoma in pediatric soft tissue. *Front. Oncol.* **14**, 1289532. <https://doi.org/10.3389/fonc.2024.1289532> (2024).
28. Lv, H. et al. Seizure detection based on lightweight inverted residual attention network. *Int. J. Neural Syst.* **34**(8), 2450042. <https://doi.org/10.1142/S0129065724500424> (2024).
29. Quan, Y., Zhang, D., Zhang, L. & Tang, J. Centralized feature pyramid for object detection. *IEEE Trans. Image Process.* **32**, 4341–4354. <https://doi.org/10.1109/TIP.2023.3297408> (2023).
30. Cao, Q. et al. Pyramid-YOLOv8: a detection algorithm for precise detection of rice leaf blast. *Plant. Methods.* **20**(1), 149. <https://doi.org/10.1186/s13007-024-01275-3> (2024).
31. Chen, H., Zhou, G. & Jiang, H. Student behavior detection in the classroom based on improved YOLOv8. *Sensors* **23**(20), 8385. <https://doi.org/10.3390/s2320838> (2023).
32. Li, S. et al. A glove-wearing detection algorithm based on improved YOLOv8. *Sensors* **23**(24), 9906. <https://doi.org/10.3390/s23249906> (2023).
33. Ma, N., Su, Y., Yang, L., Li, Z. & Yan, H. Wheat seed detection and counting method based on improved YOLOv8 model. *Sensors* **24**(5), 1654. <https://doi.org/10.3390/s24051654> (2024).
34. Tulbure, A. A., Tulbure, A. A. & Dulf, E. H. A review on modern defect detection models using DCNNs—deep convolutional neural networks. *J. Adv. Res.* **35**, 33–48. <https://doi.org/10.1016/j.jare.2021.03.015> (2021).
35. Zheng, Y., Zheng, W. & Du, X. A lightweight rice pest detection algorithm based on improved YOLOv8. *Sci. Rep.* **14**(1), 29888. <https://doi.org/10.1038/s41598-024-81587-5> (2024).
36. Zhao, B. et al. Modular YOLOv8 optimization for real-time UAV maritime rescue object detection. *Sci. Rep.* **14**(1), 24492. <https://doi.org/10.1038/s41598-024-75807-1> (2024).
37. Matthysens, L. E. et al. A novel standard for systematic reporting of neuroblastoma surgery: the international neuroblastoma surgical report form (INSRF): a joint initiative by the pediatric oncological cooperative groups SIOPEN\*, COG\*\*, and GPOH\*\*\*. *Ann. Surg.* **275**(3), e575–e585. <https://doi.org/10.1097/SLA.0000000000003947> (2022).

## Author contributions

All authors contributed to the conception and design of the study. Y.Y.Wang was responsible for data collection and the visualization of experimental results. F.F.Wang authored the initial draft, while Z.X.Qin organized the data. The review of the first draft was conducted by D.Zhang, Y.C.Fu, J.Y.Wang, and S.K.Li undertook the revision of the initial draft. D.Zhang led and directed the entire research project and provided financial.

## Funding

This research was funded by the Key Scientific Research Project of Universities in Henan Province (Grant No. 24A320078). The Key Scientific Research Project of Colleges and Universities in Henan Province (24A320078). The Science and Technology Research Program of Henan Province (242102310256). The Science and Technology Research Program of Henan Province (252102311122).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-99451-5>.

**Correspondence** and requests for materials should be addressed to D.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025