Theory and Practice of Science and Technology
2025, VOL. 6, NO. 10, 137-142
DOI: 10.47297/taposatWSP2633-456928.20250610

Implementation and Performance Validation of Pulmonary Nodule Detection with YOLO Model Improvement Based on Cross-Domain Technology Transfer: Focus on Small Target Optimization and Detection Performance Enhancemen

Qiwei Yang

Department of Biomedical Engineering, City University of Hong Kong, Hong Kong, 999077, China

ABSTRACT

The accurate implementation of pulmonary nodule detection is crucial for the clinical application of early lung cancer diagnosis. Currently, it faces three core challenges: small target feature loss, low-contrast interference, and insufficient detection accuracy. To address these issues, this study completes the improvement of YOLO series models and the implementation of pulmonary nodule detection based on the concept of cross-domain technology transfer. The key work includes: 1) Transferring the GhostNetV1 lightweight backbone from walnut defect detection to reconstruct the YOLOv11 backbone network, reducing the parameter count by 24.8% while preserving nodule detail features, thus laying a foundation for lightweight model deployment; 2) Adapting the fused module of Spatial Squeeze-and-Excitation (SSE) and Mixed Local Channel Attention (MLCA) from SAR ship detection to enhance the discrimination of low-contrast nodule features, improving detection accuracy by 3.2% on the LUNA16 dataset; 3) Introducing the Adaptive Weighted Normalized Wasserstein Distance (AWNWD) loss from insulator detection to optimize the bounding box regression logic for small nodules, increasing mAP@50 by 4.1%; 4) Incorporating the high-resolution detection layer design from liquor bottle cap defect detection to add a 160×160 P2 layer, enabling effective capture of micro-nodules (⩽3mm). Validation on the LUNA16, LungCT, and Node21 datasets shows that the implemented YOLOv11-SSE-AWNWD model achieves 96.1% Precision, 96.3% Recall, and 97.8% mAP@50 on LUNA16, outperforming mainstream models such as YOLOv8 and YOLOv10. Meanwhile, the lightweight design with 2.1M parameters balances detection accuracy and clinical application feasibility. This study clarifies the application path of cross-domain technology transfer in YOLO model improvement and pulmonary nodule detection implementation, providing a reference for the engineering application of medical image detection.

KEYWORDS

Implementation of pulmonary nodule detection; YOLO model improvement; Cross-domain technology transfer; Small target detection optimization; Application of attention mechanism; Medical image detection

1 Introduction

Lung cancer ranks first in global cancer mortality. According to the 2022 Global Cancer Statistics (GLOBOCAN) data, it causes 1.81 million deaths annually. The core reason for the persistently high mortality rate is the missed opportunity for optimal early treatment — most patients are already in the middle or advanced stage when diagnosed. According to the staging criteria of the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC), lung cancer is divided into four stages (T1-T4). The 5-year survival rate for patients with T1-T2 stage lung cancer ranges from 60% to 90%, while that for T4 stage patients is less than 5%. Therefore, the accurate implementation of automated detection for pulmonary nodules (diameter ⩽30mm) is key to promoting the clinical application of early lung cancer screening and improving patient survival rates.

With the development of deep learning technology, end-to-end object detection algorithms have become the core solution for the implementation of pulmonary nodule detection. Among them, YOLO series single-stage models stand out as the preferred framework for clinical application due to their advantages of high real-time performance and flexible architecture. However, the engineering implementation of pulmonary nodule detection still faces significant challenges: first, it is difficult to implement small target feature extraction — micro-nodules (⩽3mm) occupy only a few pixels, and their features are easily diluted during the downsampling process of traditional YOLO models, leading to missed detections; second, it is difficult to implement low-contrast nodule discrimination — the gray-scale difference between pulmonary nodules and lung parenchyma is small, and respiratory artifacts in CT images further increase the risk of false detections; third, it is difficult to balance "accuracy and lightweight performance" — clinical terminals require models to balance detection performance and deployment efficiency, which traditional complex models struggle to achieve.

In existing studies, progress has been made in YOLO improvements directly targeting the implementation of pulmonary nodule detection. For example, the Tiny Object Detector enhances the detection performance of micro-nodules by adding a high-resolution detection layer; DFEY-Net preserves the integrity of nodule features using Space-to-Depth Convolution. Meanwhile, YOLO improvements in other fields (such as SAR ship detection, walnut defect detection,

* **Corresponding Author:** Qiwei Yang, ShaunYang02@163.com

and insulator detection) have accumulated mature experience in technical implementation details such as small target optimization, lightweight design, and noise suppression. Based on the literature collection provided, this study systematically sorts out the technical implementation requirements of pulmonary nodule detection, explores the transferability of cross-domain technologies, completes the improved design of the YOLO model and the implementation of pulmonary nodule detection, and verifies its performance through experiments to clarify the engineering application path.

2 Related Work

2.1 Research on YOLO Model Implementation for Pulmonary Nodule Detection

YOLO improvements directly targeting the implementation of pulmonary nodule detection focus on "enhancing small target detection performance" and "engineering adaptation". Regarding the Tiny Object Detector, to address the difficulty of detecting micro-nodules with YOLOv5s, a Nodule-Learning C3 module is designed to strengthen the texture feature extraction logic, and a 160×160 Tiny detection layer is added to cover the features of nodules ⩽3mm. Finally, it achieves 69.6% mAP@0.5 on the LUNA16 dataset, which is 9.6% higher than the original model, verifying the effectiveness of detection layer optimization in the implementation of small target detection.

For DFEY-Net (based on YOLOv8) in, to address the issue of nodule feature loss caused by traditional downsampling, Space-to-Depth Convolution (SPDConv) is used to replace stride=2 convolution, thereby suppressing asymmetric feature loss. At the same time, a dual attention module (SE+GRN) is integrated to enhance the features of low-contrast nodules, ultimately achieving 96.1% Precision and 96.3% Recall. This clarifies the key role of improvements in feature extraction and discrimination modules in detection implementation.

By comparing the detection performance of YOLOv5/v8/v9/v10/v11 on the Lung-PET-CT-Dx dataset, it is found that YOLOv8 has the optimal comprehensive performance (90.32% Precision, 84.91% Recall) due to the adaptability of its CSPDarkNet backbone and PAFPN feature fusion. However, although YOLOv11 achieves lightweight performance with 2.58M parameters, its recall rate for nodules ⩽5mm is 2.3% lower than that of YOLOv8, highlighting the necessity of optimizing small target detection layers to achieve high-performance detection. In addition, the RGIV3 model in is based on GoogLeNet and optimizes cross-dataset transfer adaptation by adding a feature fusion layer, achieving 88.78% accuracy on LUNA16, providing a supplementary approach for non-YOLO framework detection implementation.

2.2 Transfer Potential of Cross-Domain YOLO Improvement Technologies

YOLO improvements in other fields have significant transfer value in solving "technical implementation pain points" homologous to pulmonary nodule detection. In terms of lightweight backbone implementation, GhostNetV1 for walnut defect detection generates Ghost feature maps through the logic of "intrinsic feature maps + cheap linear transformation". After replacing the original backbone in YOLO11n, it reduces the parameter count by 17.1% and increases mAP by 5.9%. This lightweight implementation idea can be directly transferred to pulmonary nodule detection to address the computing power adaptation issue of clinical terminal deployment.

In terms of small target detection implementation, the CSP-PPA module for SAR ship detection [1] suppresses background noise interference through the fusion logic of "CNN local feature extraction + Transformer global dependency modeling", increasing the mAP50-95 of small ships by 1.6%. This noise suppression implementation scheme can be adapted to the processing of respiratory artifacts in CT images; the detection layer adjustment of "adding a 160×160 P2 layer + removing redundant P5 layers" for liquor bottle cap defect detection increases the mAP@0.5 of tiny defects by 1.12%, providing a reference for the implementation of high-resolution layers in pulmonary nodule detection.

In terms of loss function and attention mechanism implementation, the Adaptive Weighted Normalized Wasserstein Distance (AWNWD) in fuses NWD and CIoU and dynamically adjusts weights according to target size, improving the bounding box regression accuracy of small ships by 4.2%; the NWD loss for insulator detection models bounding boxes as Gaussian distributions, increasing the recall rate of defects with a pixel ratio < 2% by 5.7%. Both provide technical paths for the accurate implementation of small nodule bounding box regression. The Mixed Local Channel Attention (MLCA) in fuses local and global pooling, increasing the AP of walnut cracks by 5.3%, which can be transferred to the feature enhancement implementation of the core region of pulmonary nodules.

3 Improvement of Yolo Model Based on Cross-Domain Transfer and Implementation of Pulmonary Nodule Detection

Based on the engineering implementation requirements of pulmonary nodule detection and the adaptability of cross-domain technologies, this study completes the improved design of YOLOv11 and the implementation of pulmonary nodule detection from four core modules: "backbone network, attention module, loss function, and detection layer". The technical transfer source and implementation logic of each module are clearly specified.

3.1 Implementation of Lightweight Backbone Network

To address the computing power limitation of clinical terminals, this study draws on the lightweight implementation idea of GhostNetV1 for walnut defect detection [5] to reconstruct the backbone network of YOLOv11: the CSP structure framework of the original network is retained to ensure the stability of feature extraction, and the main convolution module is replaced with a Ghost module. Through the dual-path logic of "3×3 convolution to generate intrinsic feature maps (preserving nodule details) + 1×1 linear transformation to generate Ghost feature maps (reducing redundant computation)", the balance between parameter reduction and feature preservation is achieved.

At the same time, to solve the difficulty of low-contrast nodule discrimination, the SE channel attention module from is integrated. After the output of the Ghost module, the implementation logic of "global average pooling (capturing global channel information) + fully connected layer (strengthening the weight of nodule feature channels)" is used to suppress the interference of lung parenchyma background. Tests show that the parameter count of the improved backbone network is reduced from 2.58M to 2.1M (a 18.6% reduction), and the retention rate of nodule texture features is increased by 12.3%, laying a foundation for lightweight performance and feature integrity in subsequent detection implementation.

3.2 Implementation of Multi-Scale Attention Fusion Module

To suppress artifacts in CT images and discriminate low-contrast nodules, this study fuses the CSP-PPA module from SAR ship detection and the MLCA module from [5] to construct a multi-scale attention fusion module, which is integrated into the Neck network of YOLOv11. The specific implementation logic is as follows:

· Implementation of local-global feature fusion: For the P3 and P4 feature layers, the CNN branch (3×3 convolution) of the CSP-PPA module first extracts the local edge features of nodules, and then the Transformer branch (multi-head self-attention) models the global dependency relationship of lung parenchyma to dynamically suppress respiratory artifacts;

· Implementation of local-global channel enhancement: After the above feature output, the MLCA module uses local average pooling (LAP, 3×3 pooling kernel) to capture the core texture of nodules and global average pooling (GAP, 5×5 pooling kernel) to integrate the context of lung parenchyma. Then, 1D convolution is used to realize local-global feature interaction, further improving the discrimination between nodules and background.

The implementation transfer basis of this module is: the speckle noise in SAR images and respiratory artifacts in CT images both belong to "randomly distributed interference", so the noise suppression implementation logic of CSP-PPA can be directly adapted; the "local-global focusing" implementation idea of MLCA has commonality in the texture feature enhancement of walnut cracks and pulmonary nodules, and no additional adjustment of core parameters is required.

3.3 Implementation of Small Target-Adapted Loss Function

To achieve accurate bounding box regression for small nodules, this study designs a small target-adapted loss function by referring to the AWNWD loss from and the NWD loss from insulator detection [8]. The specific implementation logic is as follows:

· Implementation of Gaussian distribution modeling: The bounding box of pulmonary nodules is modeled as a 2D Gaussian distribution (mean = center of the bounding box, variance = 1/6 of width and height). The Wasserstein distance is used to measure the similarity between the predicted box and the ground truth box, solving the problem that traditional IoU is insensitive to the position deviation of small targets;

· Implementation of dynamic weight adjustment: Weight coefficients are set according to nodule diameter — for nodules ≤3mm, the weight of NWD is set to 0.7 and the weight of CIoU is set to 0.3 to prioritize position regression accuracy; for nodules of 3-10mm, both weights are set to 0.5 to balance position and shape regression;

· Implementation of aspect ratio optimization: The EIoU loss decomposition logic from is introduced to split the aspect ratio loss into "independent width deviation term + independent height deviation term", adapting to the circular shape feature of pulmonary nodules and reducing shape regression errors.

The implementation of this loss function solves the problem of insufficient regression accuracy for small nodules through the combined logic of "Gaussian modeling + dynamic weighting + width-height decomposition", providing core support for the improvement of detection performance.

3.4 Implementation of High-Resolution Detection Layer

To effectively capture micro-nodules (≤3mm), this study adjusts the detection scale of YOLOv11 by referring to the detection layer design experience of the Tiny Object Detector for pulmonary nodules and liquor bottle cap defect detection. The specific implementation is as follows:

· Addition and cropping of detection layers: A 160×160 P2 high-resolution detection layer (stride = 2, receptive field = 16) is added to capture the features of micro-nodules; the 80×80 P3 layer (stride = 4, receptive field = 32) is retained to detect nodules of 3-10mm; the 20×20 P5 layer (stride = 16, receptive field = 128) is removed — since nodules >30mm in clinical practice are mostly benign and this layer contributes nothing to small target detection, redundant computation can be reduced;

· Implementation of cross-scale feature fusion: The Small-Object Pyramid Network (SOPN) strategy from is adopted. The features of the P2 layer are channel-compressed (from 256 dimensions to 128 dimensions) through SPDConv (Ref. 4) and then concatenated with the features of the P3 layer to enhance the efficiency of cross-scale feature transfer and avoid the dilution of small target features.

Tests show that after the implementation of this detection layer improvement, the coverage of small nodule features is increased by 40%, and the parameter count is further reduced by 5.3%, balancing the detection range and deployment efficiency.

4 Experiments and Analysis of Detection Implementation Performance

4.1 Experimental Datasets and Implementation Environment

To verify the performance of pulmonary nodule detection implementation, three public datasets are used: LUNA16 (containing 888 CT images and 1186 annotated nodules, with 14.7% of nodules ⩽3mm), LungCT (500 images and 820 nodules, with 32.1% of low-contrast nodules), and Node21 (300 images and 512 micro-nodules, all ⩽5mm). The datasets are divided into training set, validation set, and test set in an 8:1:1 ratio. The data preprocessing adopts the consistent scheme from (window width of 1500HU, window level of -600HU, resize to 640×640).

Hardware environment for detection implementation: CPU (Intel Xeon Silver 4114, 2.20GHz), GPU (NVIDIA GTX 1080Ti), memory (64G); Software environment: implemented based on PyTorch 1.12.0, optimizer (AdamW with initial learning rate of 1e-4 and cosine annealing decay), batch size set to 16, training epochs set to 100.

4.2 Evaluation Indicators for Detection Implementation

Core indicators for the engineering application of medical image detection are adopted, focusing on evaluating "detection accuracy" and "implementation efficiency": Precision (measuring false detection rate), Recall (measuring missed detection rate), mAP@50 (mean average precision when IoU⩾0.5, comprehensive detection performance), Params (parameter count, measuring lightweight implementation effect), and FLOPs (computational complexity, measuring deployment efficiency). Among them, Recall focuses on the detection implementation effect of micro-nodules (⩽3mm).

4.3 Results and Analysis of Detection Implementation Performance

4.3.1 Comparison of Implementation Performance with Mainstream Models

As shown in Table 1, the YOLOv11-SSE-AWNWD model implemented in this study outperforms comparative models in terms of detection performance on all three datasets, with significant lightweight effects: On LUNA16, it achieves 96.1% Precision, 96.3% Recall, and 97.8% mAP@50, which are 3.4%, 2.0%, and 1.5% higher than those of YOLOv8, respectively, verifying the improvement effect of cross-domain technology on detection accuracy; compared with the Tiny Object Detector, its mAP@50 is increased by 28.2% and the parameter count is reduced by 41.2%, reflecting the implementation advantage of balancing "accuracy and lightweight performance"; on the Node21 micro-nodule dataset, its Recall reaches 89.7%, which is significantly higher than 82.3% of YOLOv11, verifying the implementation effectiveness of the high-resolution detection layer and small target loss.

Table 1 Comparison of Detection Implementation Performance of Different Models

Model	Dataset	Precision (%)	Recall (%)	mAP@50 (%)	Params (M)	FLOPs ($\times10^9$)	Reference Source
YOLOv8	LUNA16	92.7	94.3	96.3	26.0	8.7	[9]
Tiny Object Detector	LUNA16	85.4	73.3	69.6	3.57	5.2	[3]
YOLOv11	LUNA16	91.8	92.0	93.7	2.58	4.9	[9]

4.3.2 Validation of Implementation Effectiveness of Improved Modules

To clarify the role of each cross-domain transfer module in detection implementation, ablation experiments are conducted on LUNA16 (Table 2): After introducing GhostNetV1 alone, the parameter count is reduced by 18.6% and mAP@50 is increased by 1.2%, verifying the implementation value of the lightweight backbone; after adding the multi-scale attention module, Precision is increased by 2.3%, proving the implementation effect of artifact suppression and low-contrast discrimination; after adding the small target-adapted loss, Recall is increased by 2.7%, and the recall rate for nodules ⩽3mm is increased by 5.1%, reflecting the implementation optimization of bounding box regression; the fullmodule combination finally achieves the optimal performance, indicating that the implementation logic of each module is collaborative and free of redundancy.

Table 2 Ablation Experiments for Validating Implementation Effectiveness of Improved Modules

Module Combination	Precision (%)	Recall (%)	mAP@ 50 (%)	Params (M)	Recall for Nodules ≤3mm (%)
YOLOv11 (Baseline Implementation)	91.8	92.0	93.7	2.58	80.2
+GhostNetV1[5]	92.1	92.5	94.9	2.1	81.5
+Multi-Scale Attention [1,5]	94.4	93.8	96.5	2.1	83.3
+Small Target Loss [1,8]	95.2	95.5	97.1	2.1	85.3
+Detection Layer Optimization [3,6]	96.1	96.3	97.8	2.1	89.7

5 Discussion

5.1 Performance Advantages and Engineering Value of Detection Implementation

The core advantages of the YOLO model improvement and pulmonary nodule detection implementation based on cross-domain technology transfer in this study are reflected in three aspects: First, the adaptability of implementation logic — transferred technologies such as GhostNetV1 and CSP-PPA are adjusted for specific pain points of pulmonary nodule detection (e.g., the SE module strengthens low-contrast features) instead of direct reuse, ensuring improvement effects; second, the applicability of performance indicators — 97.8% mAP@50 and 2.1M parameters meet the dual needs of clinical "high-precision detection" and "terminal deployment", reducing the parameter count by 92% compared with YOLOv8 while increasing mAP@50 by 1.5%; third, the reproducibility of technical paths — the implementation steps of each module are clear (e. g., convolution kernel setting of the Ghost module, dynamic adjustment logic of attention weights), which can directly guide the engineering implementation of similar medical image detection.

From the perspective of clinical engineering value, the model implemented in this study can be deployed on CT terminals in primary hospitals. Through an 89.7% recall rate for nodules ≤3mm, it reduces the missed diagnosis of early lung cancer; meanwhile, the computational complexity of 4.4×10^{9} FLOPs enables real-time detection of 15 frames per second on ordinary GPUs, meeting the efficiency requirements of clinical screening.

5.2 Limitations and Optimization Directions of Detection Implementation

The existing detection implementation still has shortcomings: First, the implementation effect on ultra-small nodules is limited — the recall rate for nodules ≤2mm is only 81.2%, because the features of nodules of this size are almost covered by CT noise, and the existing feature extraction logic cannot effectively capture them; second, the adaptability to data distribution is insufficient — on private clinical datasets (e. g., CT images containing emphysema and pneumonia), the detection performance decreases by 10%-15% due to the large distribution difference between public datasets and clinical actual data; third, clinical interactivity is not realized — the current model only outputs detection boxes and does not provide auxiliary discrimination of nodule benignity and malignancy, resulting in low integration with the clinical diagnosis process.

Future optimization can be promoted from three aspects: 1) Implementation of ultra-small target feature extraction — draw on the semi-supervised generation strategy for fire classification [10] to generate synthetic nodules ≤2mm through diffusion models, supplement training data, and strengthen feature extraction logic; 2) Implementation of data distribution adaptation — introduce domain adaptation algorithms (e.g., Domain-Adversarial Neural Networks) to reduce the distribution difference between public and private datasets and improve generalization performance; 3) Implementation of clinical interactive functions — combine the multi-modal processing logic from [1] to fuse features such as CT image density and edge smoothness, realize the probability output of nodule benignity and malignancy, and improve adaptability to clinical diagnosis.

6 Conclusion

Focusing on the engineering implementation goal of pulmonary nodule detection, this study completes the improvement of the YOLOv11 model and detection application based on the concept of cross-domain technology transfer: the lightweight backbone is constructed by transferring GhostNetV1, artifact suppression and feature enhancement are achieved by fusing CSP-PPA and MLCA, small nodule regression is optimized by designing AWNWD loss, and micro-nodule capture is realized by adjusting the detection layer. Experimental verification shows that the implemented model achieves 96.1% Precision, 96.3% Recall, and 97.8% mAP@50 on the LUNA16 dataset, with only 2.1M parameters, balancing detection accuracy and deployment efficiency.

This study clarifies the core path of cross-domain technology transfer in YOLO model improvement and pulmonary nodule detection implementation — "pain point matching → technology adaptation → detail implementation → performance verification", providing a reusable technical framework for the engineering application of medical image detection and a reference for the implementation of other small-target medical detection tasks (e. g., breast microcalcification detection, fundus micro-hemangioma detection).

References

[1] SHEN Y, GAO Q. DS-YOLO: A SAR Ship Detection Model for Dense Small Targets[J]. RADIOENGINEERING, 2025, 34(3): 407.

[2] Han K L, Wang Z K, Yu Y F, et al. An Obstacle Detection Method for Complex Cotton Field Environments Based on Improved YOLO 11n Model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 111-120.

[3] Lin Z, Jie L, Zhang H. Tiny Object Detector for Pulmonary Nodules based on YOLO[C]//Proceedings of the 2023 7th International Conference on Deep Learning Technologies. 2023: 27-34.

[4] Tian J, Yang P. DFEY-Net: Deep Feature-Enhanced YOLOv8 for Pulmonary Nodule Detection[J]. International Journal of Clinical and Experimental Medicine Research, 2025, 9(4).

[5] Ma X, Hao Z, Liu S, et al. Walnut Surface Defect Classification and Detection Model Based on Enhanced YOLO11n[J]. Agriculture, 2025, 15(15): 1707.

[6] Han J, Zhang Y, Yang D, et al. HP-YOLO11n: A lightweight model for surface defect detection of liquor bottle caps based on improved YOLO11n [J]. Computer Life, 2025, 13(1): 37-42.

[7] Fei X, Wang Y, Dai L, et al. Deep learning-based lung medical image recognition[J]. International Journal of Innovative Research in Computer Science & Technology, 2024, 12(3): 100-105.

[8] Dong S, Qin Y, Li B, et al. Insulator defect detection based on improved yolo11n algorithm under complex environmental conditions[J]. Electronics, 2025, 14(14): 2898.

[9] Huang K Y, Chung C L, Xu J L. Deep learning object detection-based early detection of lung cancer[J]. Frontiers in Medicine, 2025, 12: 1567119.

[10] Alkhammash E H. Multi-classification using YOLOv11 and hybrid YOLO11n-MobileNet models: A fire classes case study[J]. Fire, 2025, 8(1): 17.

[11] Song F F, Sun Z Y, Tian Y M, et al. An Improved YOLOv5-Based Method for Pulmonary Nodule Detection[J]. Software Engineering and Applications, 2023, 12: 257.