

Review

Multi-Task Deep Learning for Lung Nodule Detection and Segmentation in CT Scans: A Review

Runhan Li ^{1,2}  and Barmak Honarvar Shakibaei Asli ^{2,*} 

¹ College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; runhan.li@cranfield.ac.uk

² Centre for Life-Cycle Engineering and Management, Faculty of Engineering and Applied Sciences, Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

* Correspondence: barmak@cranfield.ac.uk

Abstract

Lung nodule detection and segmentation are essential tasks in computer-aided diagnosis (CAD) systems for early lung cancer screening. With the growing availability of CT data and deep learning models, researchers have explored various strategies to improve the performance of these tasks. This review focuses on Multi-Task Learning (MTL) approaches, which unify or cooperatively integrate detection and segmentation by leveraging shared representations. We first provide an overview of traditional and deep learning methods for each task individually, then examine how MTL has been adapted for medical image analysis, with a particular focus on lung CT studies. Key aspects such as network architectures and evaluation metrics are also discussed. The review highlights recent trends, identifies current challenges, and outlines promising directions toward more accurate, efficient, and clinically applicable CAD solutions. The review demonstrates that MTL frameworks significantly enhance efficiency and accuracy in lung nodule analysis by leveraging shared representations, while also identifying critical challenges such as task imbalance and computational demands that warrant further research for clinical adoption.

Keywords: multi-task; deep learning; image analysis; lung nodule detection; segmentation



Academic Editors: Morgado Dias,
Fabio Mendonca and Sheikh
Shanawaz Mostafa

Received: 16 June 2025

Revised: 17 July 2025

Accepted: 24 July 2025

Published: 28 July 2025

Citation: Li, R.; Honarvar Shakibaei Asli, B. Multi-Task Deep Learning for Lung Nodule Detection and Segmentation in CT Scans: A Review. *Electronics* **2025**, *14*, 3009. <https://doi.org/10.3390/electronics14153009>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Lung cancer is one of the most common and fatal cancers worldwide, with survival rates heavily dependent on early and accurate diagnosis [1,2]. According to Cancer Research UK, there were around 49,200 new lung cancer cases in the UK every year between 2017 and 2019, with around 34,800 lung cancer deaths every year [3]. Computed tomography (CT) is the primary imaging method for lung cancer screening due to its high spatial resolution and volumetric information [4]. However, the manual interpretation of CT scans is time-consuming and subject to variability across radiologists [2].

Pulmonary nodules are often early signs of lung cancer. The accurate detection and segmentation of these nodules are critical for diagnosis and treatment planning [5]. Traditionally, these tasks were approached separately using classical image processing or machine learning techniques [6]. Machine learning typically relies on handcrafted features and statistical classifiers, such as support vector machines or decision trees [7,8]. However, the rise of deep learning—particularly convolutional neural networks (CNNs)—has enabled end-to-end learning of feature representations from raw data, leading to significant improvements in performance [1,9].

Computer-aided diagnosis (CAD) systems have been developed to assist radiologists in analyzing pulmonary nodules, typically combining several image-processing tasks such as detection, segmentation, and classification [10]. These systems aid in the early detection and segmentation of various tumors across different organs, and also integrate feature extraction from imaging data for automated classification [11,12]. As these tasks are critical for assisting diagnosis and treatment decisions, various CAD systems have been proposed to support these goals [13]. Typically, such systems follow a structured pipeline that includes CT image acquisition, lung segmentation, nodule detection and segmentation, followed by feature extraction (e.g., volume, shape, and appearance). These features are then used to support malignancy assessment as illustrated in Figure 1 [10].

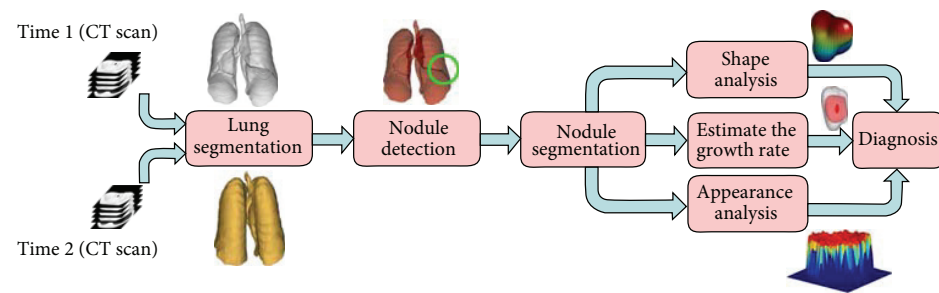


Figure 1. A typical CAD pipeline for lung cancer. The system takes CT scans as input, performs lung segmentation to define the region of interest, and applies nodule detection and segmentation. Key features such as volume, shape, and texture are extracted and used to support diagnostic classification [10].

Instead of treating these steps separately, recent studies have integrated related tasks—such as detection and segmentation, into unified Multi-Task Learning (MTL) frameworks [14–16]. MTL refers to a learning paradigm in which multiple related tasks are learned jointly, allowing knowledge to be shared across tasks to improve overall performance [17]. This approach enables shared representations, reduces redundant computations, and may improve performance by allowing tasks to benefit from each other [18].

Despite these benefits, designing efficient and accurate multi-task frameworks for lung nodule analysis remains challenging. Key issues include balancing task-specific learning, selecting efficient network architectures for 3D data, and maintaining performance without excessive model complexity [19]. These challenges highlight the need for lightweight, generalizable, and dynamically balanced MTL frameworks, supported by adaptive learning strategies and collaborative validation across diverse clinical settings.

This review summarizes recent progress in both traditional and deep learning methods for nodule detection and segmentation, with a special focus on MTL approaches. It provides a comprehensive overview of methods across the full CAD pipeline. Unlike previous reviews that primarily focus on either general CAD systems or deep learning architectures, this work systematically covers evaluation metrics, lung nodule analysis datasets, image preprocessing, lung segmentation, and both traditional and deep learning-based approaches to nodule detection and segmentation. It further highlights MTL as a promising strategy for integrating detection and segmentation tasks, classifies common architectural paradigms (e.g., parallel, cascaded, interacted, and hybrid), and critically summarizes unresolved methodological bottlenecks that limit clinical deployment. Figure 2 illustrates this conceptual flow, showing how clinical motivations guide methodological choices and ultimately enhance diagnostic capabilities. This review aims to serve as both a practical reference for understanding the CAD pipeline and a forward-looking synthesis that informs the development of efficient and clinically deployable MTL-based solutions.

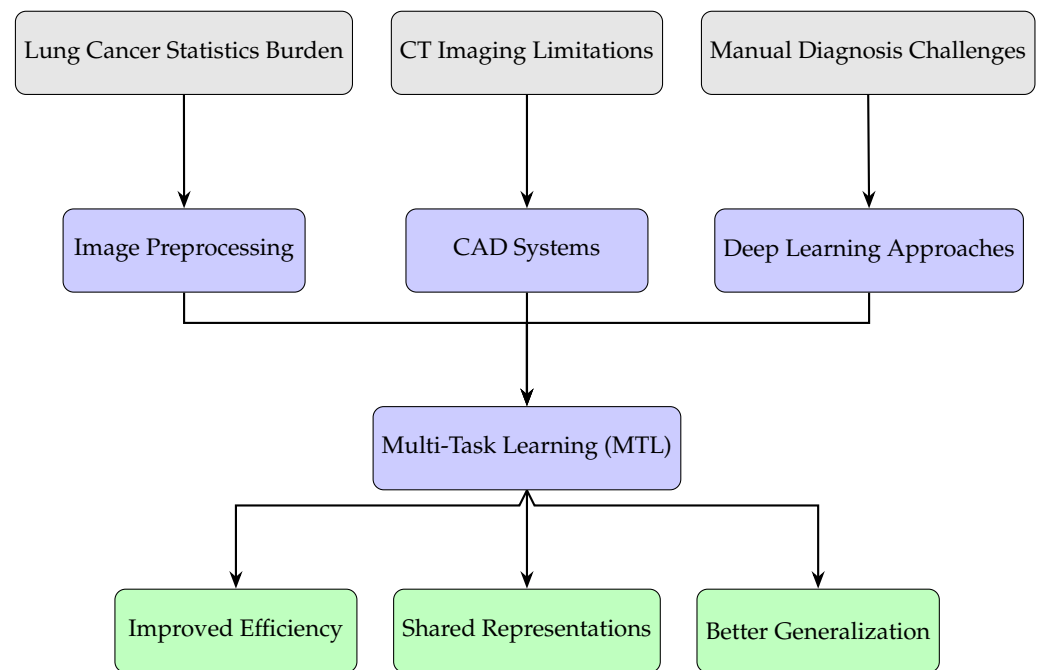


Figure 2. Conceptual flow of this review: clinical needs motivate the use of CAD systems and Multi-Task Learning approaches, leading to improved efficiency, representation sharing, and generalization.

2. Evaluation Metrics

Evaluation of multi-task models in lung CT analysis typically involves two types of metrics: image quality assessment (IQA) metrics for preprocessing or reconstruction, and task-specific performance metrics for nodule detection and segmentation. These metrics help to objectively assess the effectiveness of each component in the MTL pipeline.

2.1. IQA Metrics

IQA metrics are useful when the model includes image preprocessing steps such as denoising, deblurring, or enhancement. These metrics assess the quality of processed images either by comparing them to reference images as a Full-Reference (FR) approach or without references as a No-Reference (NR) technique [20]. Table 1 summarizes some common IQA metrics. The table includes Mean Squared Error (MSE), Structural Similarity Index Measure (SSIM), Peak Signal-to-Noise Ratio (PSNR), Perception-based Image Quality Evaluator (PIQE), Naturalness Image Quality Evaluator (NIQE) and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), which are commonly used in image analysis.

2.2. Task Performance Metrics for Detection and Segmentation

Detection and segmentation performance are assessed using classification-based and spatial overlap metrics. Here, detection refers to identifying the presence and location of pulmonary nodules, while segmentation involves delineating their precise boundaries within the lung region. Classification metrics assess whether a nodule is correctly identified, while overlap metrics evaluate how accurately the predicted segmentation matches the ground truth, which refers to manually annotated reference labels for both detection and segmentation tasks. These metrics are essential in medical imaging applications, where accurate detection and segmentation directly impact clinical decision-making [9]. Table 2 summarizes commonly used evaluation metrics, including their definitions and mathematical formulas.

Table 1. Common image quality assessment (IQA) metrics used in medical image analysis. Abbreviations: MSE (Mean Squared Error), SSIM (Structural Similarity Index), PSNR (Peak Signal-to-Noise Ratio), PIQE (Perception-based Image Quality Evaluator), NIQE (Naturalness Image Quality Evaluator), BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) [20–23].

Type	Metric	Description	Formula
FR	MSE	Calculates the Mean Squared Error between images. A smaller MSE indicates higher image similarity. A lower value indicates better quality.	$\text{MSE}(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2,$ where x_i and y_i are corresponding pixel values in the two images, and N is the total number of pixels.
	SSIM	Assesses perceptual similarity based on luminance, contrast, and structure.	$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$ where μ_x, μ_y are the mean intensities of images x and y , respectively; σ_x, σ_y are their standard deviations; σ_{xy} is the cross-covariance. C_1 and C_2 are constants to stabilize the division.
	PSNR	Measures signal fidelity between two images. PSNR is inversely related to MSE; as MSE decreases toward zero, PSNR increases, indicating better image quality.	$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right),$ where MAX is the maximum possible pixel value of the image.
NR	PIQE	Evaluates perceptual quality based on block-level distortion. It analyzes local blocks for artifacts and noise. Lower PIQE indicates better quality.	$\text{PIQE} = f(\text{BM}, \text{NM}),$ where BM is blockiness measure and NM is noise measure.
	NIQE	Estimates deviation from natural image statistics. A lower value indicates better quality.	$d = \sqrt{(\mu_1 - \mu_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\mu_1 - \mu_2)},$ where μ_1, μ_2 and Σ_1, Σ_2 are the mean vectors and covariance matrices of the test and natural images.
	BRISQUE	Predicts image quality using natural scene statistics. A lower score means better quality.	$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + C},$ where $\mu(i, j)$ and $\sigma(i, j)$ are local mean and standard deviation from MSCN coefficients. Features are modeled using AGGD, and quality is predicted via SVR.

These metrics allow for a reliable and standardized comparison of model performance across studies. In particular, metrics like Free-response Receiver Operating Characteristic (FROC) and Competition Performance Metric (CPM) are especially relevant when using datasets like LUNA16, where competition benchmarks demand robust sensitivity with minimal false positives [24].

Table 2. Common evaluation metrics for detection and segmentation tasks [1,9].

Metric	Description	Formula
Accuracy (ACC)	Correct predictions among all cases.	$\frac{TP + TN}{TP + TN + FP + FN}$
Sensitivity (Recall)	True positive rate.	$\frac{TP}{TP + FN}$
Specificity	True negative rate.	$\frac{TN}{TN + FP}$

Table 2. *Cont.*

Metric	Description	Formula
Precision (PPV)	Positive predictions that are correct.	$\frac{TP}{TP+FP}$
F1-Score	Harmonic mean of precision and recall.	$2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$
Dice Similarity Coefficient (DSC)	Similarity between prediction and ground truth.	$\frac{2 A \cap B }{ A + B }$
Intersection over Union (IoU)	Overlap ratio of prediction and ground truth.	$\frac{ A \cap B }{ A \cup B }$
AUC	Area under ROC curve.	–
ROC Curve	True positive rate vs. false positive rate.	–
FROC Curve	Sensitivity vs. false positives per scan.	–
CPM	Average sensitivity at 7 FPs/scan.	–
Mean Average Precision (mAP)	Mean of AP over all classes.	$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$

Notes: TP = True Positives; TN = True Negatives; FP = False Positives; FN = False Negatives. A and B are the predicted and true binary masks. AUC = Area Under the Curve, ROC = Receiver Operating Characteristic, FROC = Free-response ROC, CPM = Competition Performance Metric (used in LUNA16).

3. Datasets for Lung Nodule Analysis

Table 3 summarizes the most commonly used datasets for lung nodule detection and segmentation. These datasets differ in terms of size, annotation type, and clinical task suitability. While LIDC-IDRI and LUNA16 are widely used for detection benchmarking, Tianchi provides lesion-level clinical labels that also support diagnostic modeling.

Table 3. Summary of commonly used lung nodule datasets. Abbreviations: CT (computed tomography), DX (digital radiography), CR (computed radiography).

Dataset	Modality	Annotation Type	Size	Notes
LIDC-IDRI [25]	CT, DX, CR	Bounding boxes, malignancy ratings (4 readers)	1018 scans	Multi-reader annotated; basis for many derived datasets
LUNA16 [24]	CT	Filtered nodules ≥ 3 mm from LIDC-IDRI	888 scans	High-quality subset of LIDC-IDRI; used for 10-fold CV
NLST [26]	CT, X-ray	Patient outcome, lesion info	54,000+ participants	Large-scale clinical trial; mortality-focused
ANODE09 [27]	CT	True nodules and irrelevant findings	55 scans	Partially annotated; used for CAD evaluation
ELCAP [28]	CT	Nodule locations	50 scans	Early low-dose CT dataset; used for CAD benchmarking
Tianchi [29]	CT	3D nodule boxes, clinical labels	1000+ scans	5–30 mm nodules annotated by 3 doctors

4. Image Preprocessing

Preprocessing is an important step in lung CT images analysis. It helps reduce noise, improve image quality, and highlight lung structures that may look like nodules. Clean, enhanced inputs are essential for reliable detection and segmentation [13].

Many filters have been applied to reduce CT image noise and preserve structural detail [30]. Among them, the mean filter is a basic linear smoothing method that averages nearby pixels to reduce random noise, though it may blur important edges [31]. The median filter is widely used for suppressing “salt and pepper” noise while preserving edges [32]. It has been adopted in several studies to improve slice-wise consistency in CT scans [33–36].

Figure 3 compares the effects of mean and median filters on both normal and tumor CT slices, highlighting differences in noise suppression and edge retention.

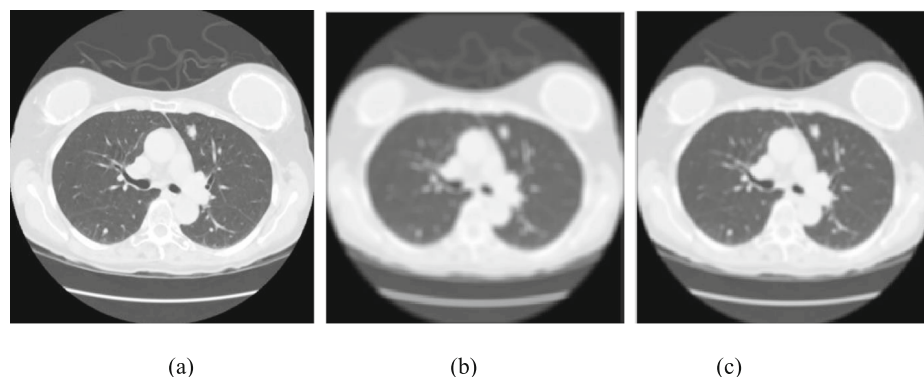


Figure 3. Comparison of preprocessing filters: (a) original CT slice, (b) mean-filtered image, and (c) median-filtered image. Note: Only a subset of the original figure is shown here to highlight differences in filter behavior [31].

The Gaussian filter is also widely used [37,38]. It smooths the image by reducing high-frequency noise and softens sharp edges [39]. Gaussian filters are effective in reducing speckle noise and making nodule-like regions easier to identify [40]. Gaussian blurring for CT images in both 1D and 2D is mathematically represented in Equation (1). σ denotes the standard deviation of the distribution, which controls the extent of blurring or noise spread. In the 1D case, $g_{\sigma}(x)$ models the probability density along a single spatial axis, while the 2D form $G_{\sigma}(x, y)$ represents the joint distribution over both spatial axes (x, y) . The Gaussian function is centered at the origin and decays smoothly as x and y move away from the center, making it suitable for smoothing operations in image processing [41]:

$$\begin{aligned} 1D : g_{\sigma}(x) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right), \\ 2D : G_{\sigma}(x, y) &= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \end{aligned} \quad (1)$$

The Wiener filter is helpful when both noise and blurring are present [42]. It works by minimizing the Mean Squared Error between the original and processed image. It can remove additive noise and also sharpen the image at the same time [36].

Histogram equalization and its improved form, contrast-limited adaptive histogram equalization (CLAHE), are used to adjust the contrast of CT scans [31,43]. These methods improve the visibility of lung regions by spreading out the intensity values across the image. CLAHE is especially useful because it avoids over-enhancing noise [44]. Equation (2) shows the rule of histogram equalization:

$$I_N = \frac{\text{number of pixels with available intensity } n}{\text{total number of pixels}}, \quad (2)$$

where $n = 0, 1, \dots, 255$.

Other techniques used in preprocessing include Gabor filters [45], which highlight texture information, log filters, and dot enhancement filters, which are good at enhancing small structures like nodules. Interpolation methods are used to adjust image resolution or fill in missing data between slices [46].

Deblurring is another important preprocessing step. It focuses on improving blurry CT images by recovering high-resolution (HR) images from low-resolution (LR) ones [47]. Equations (3)–(5) show the deblurring problem [23]. Deep learning models can be trained to

perform this task by learning the reverse process of image degradation [23]. In supervised settings, they learn from pairs of clear and blurry images and try to reduce the difference between the output and the target image [47]. Some common techniques include 3D-SRCNN [48], DCSRNN [49], and mDCSRNN [50]. The deblurring results of these methods are shown in Figure 4. Each section showcases the outputs from Tricubic interpolation, 3D SRCNN, DCSRNN, mDCSRNN, ParallelNet, and VolumeNet, alongside the ground truth (GT) images.

$$I_{LR} = \mathcal{A}I_{HR}, \quad (3)$$

where I_{LR} represents the blurred (LR) image, I_{HR} represents the original (HR) image, and $\mathcal{A}$ symbolizes the blurring process.

$$\arg \min_{\theta} \mathcal{L}(\hat{I}_{HR}, I_{HR}), \quad (4)$$

where $\hat{I}_{HR}$ denotes the estimated HR image generated by a deep neural network f with parameters θ , and $\mathcal{L}$ signifies the loss function that quantifies the difference between the estimated and original HR images. The objective is to determine the optimal θ that minimizes this reconstruction loss:

$$\mathcal{L}_{Per} = \left\| \phi(\hat{I}_{HR}) - \phi(I_{HR}) \right\|_2^2, \quad (5)$$

where $\mathcal{L}_{Per}$ stands for the perceptual loss, calculated as the squared L2 norm of the difference between feature representations ϕ extracted from a pre-trained network (e.g., VGG-16 [51]) for the estimated $\hat{I}_{HR}$ and the original I_{HR} . This loss measures the discrepancy in high-level image features rather than pixel-wise differences.

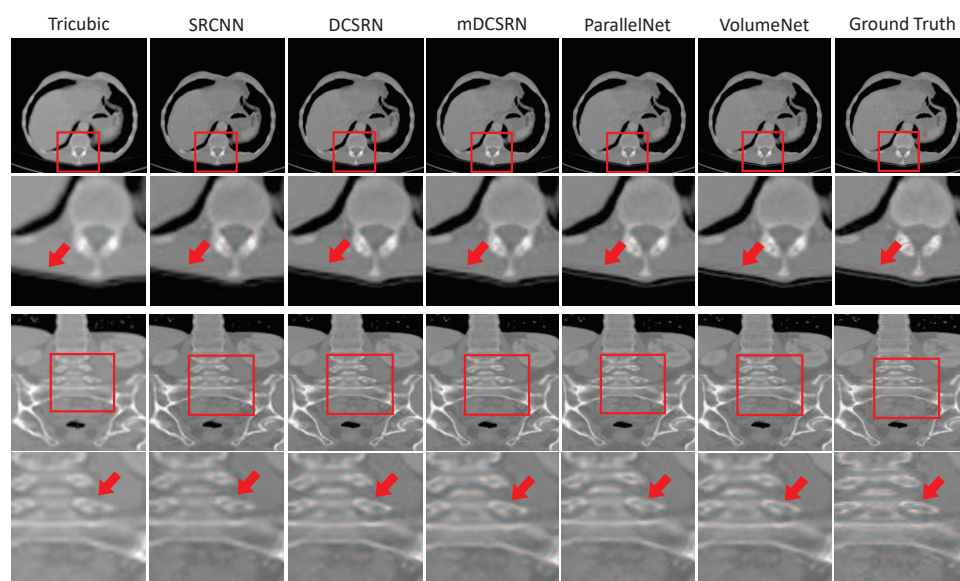


Figure 4. Deblurring results of different methods on the LiTS dataset. Compared methods include Tricubic interpolation, 3D-SRCNN, DCSRNN, mDCSRNN, ParallelNet, and VolumeNet. The ground truth (GT) represents the original high-quality CT image used as the reference for evaluation. The scale factor is $\times 2$ in each direction, with channel compression ratio $R = T/2$ in VolumeNet [52].

Recent methods also explore self-supervised deblurring, which does not need labeled data. Instead, they create supervision from the input data itself, such as by predicting missing slices between existing CT slices [53]. Some models also use meta-learning to adapt to different levels of blur or resolution. For example, Meta-SR can adjust to any scale factor to produce clearer images [54]. Furthermore, Meta-USR extends this idea by

using a conditional hyper-network to adapt to different degradation types, achieving better results on benchmark datasets [55]. The Meta-USR framework includes three sections: (a) the Residual Dense Block (RDB) from the RDN backbone, which is used for hierarchical feature extraction, (b) the feature learning module (FLM), which produces shared feature representations adaptable to various upscaling factors, and (c) the Meta-Representation Module (MRM), where each pixel in the super-resolved image is mapped to the low-resolution space. The MRM uses both position and scale information to dynamically generate convolution weights, enabling the precise reconstruction of HR outputs.

In summary, preprocessing helps prepare lung CT images by removing noise, improving contrast, and reducing blur [32]. These steps help improve the performance of the detection and segmentation models used later in the analysis.

5. Lung Segmentation

After preprocessing, the next important step is lung segmentation. It separates the lung region from other structures in CT images. This narrows the region of interest for subsequent analysis [10,56]. Many methods have been used for lung segmentation, and classic methods can be classified into four categories: methods based on thresholding, deformable boundaries, and shape or edge-based models [57].

Thresholding methods set pixel intensity limits to extract the lung region. Sometimes they are used together with other techniques, like region growing [58], morphological filters [59], and connected component analysis [60]. Figure 5 shows an example of lung segmentation using thresholding and region growing, where the two largest regions are selected while excluding the trachea. Hu et al. [61] proposed a fully automatic method for lung segmentation in 3D CT scans, combining gray-level thresholding, left–right lung separation, and morphological refinement. Dynamic programming is used to identify anatomical junctions, and post-processing smooths the mediastinal boundary. The process includes three main stages: first, the lungs are extracted from the CT scan; next, the left and right lungs are separated; finally, a smoothing step is applied to refine the lung boundaries. The method achieves sub-millimeter accuracy compared to manual annotations, demonstrating that thresholding-based approaches can be effective when guided by anatomical structure.

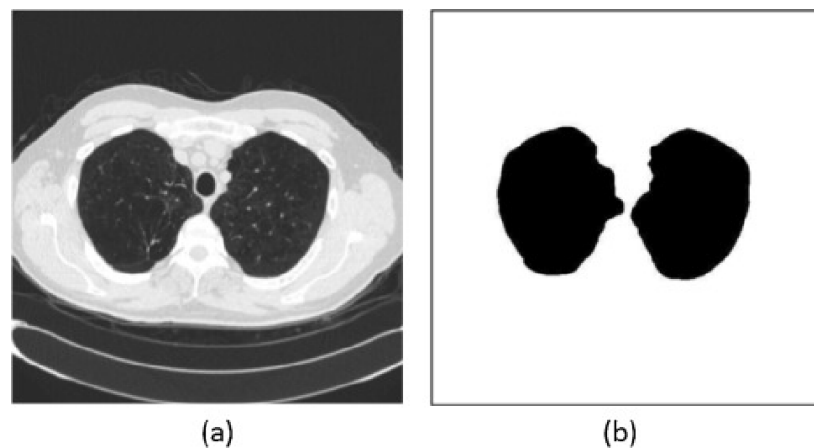


Figure 5. Example of lung segmentation: (a) original CT image and (b) segmentation mask showing the extracted lung regions (sometimes referred to as the segmented image in prior literature) [57].

The second group of methods uses deformable models. These include active contour models, which move shape outlines based on local image features. They are more flexible than thresholding but often slower [62]. To improve segmentation accuracy, some deformable models incorporate prior knowledge about lung shape. These evolve into shape-based methods, which use learned statistical models to guide anatomically consis-

tent segmentation. These methods perform well even when lung boundaries are deformed or poorly defined [63,64]. Edge-based methods use filters or transforms to detect strong edges around the lung. Techniques like Gaussian derivatives, Laplacian of Gaussian, or wavelets are often used. These methods are simple and effective when the lung boundary is clear and the contrast is high [65].

While many methods can achieve good results in lung segmentation, there is still no perfect solution. In recent years, deep learning methods have shown strong performance in lung segmentation, especially in challenging cases with pathologies. For instance, Hofmanninger et al. [66] demonstrated that a standard U-Net, when trained on diverse routine datasets, can outperform specialized methods, achieving a mean Dice score of 0.98 on pathological CT scans. These models benefit from being fast, reproducible, and scalable, though their performance still depends heavily on the quality and diversity of training data. Figure 6 presents segmentation outputs on routine CT cases. Each column corresponds to a different patient. The first row shows raw CT slices, the second row shows ground truth masks, and rows 3 to 5 display the predicted masks from different methods: U-net(R-231), CIP, and P-HNN. Tan et al. [67] proposed a lung segmentation method based on Generative Adversarial Networks (GANs), referred to as LGAN. This approach simplifies the segmentation process while achieving high accuracy and shape similarity on datasets like LIDC-IDRI and QIN, showing improved performance over existing methods. Figure 7 shows the LGAN framework, where a CNN-based generator predicts lung masks from CT slices, and a CNN-based discriminator compares them with ground truth masks to guide more accurate segmentation. Hu et al. [68] proposed a lung segmentation method based on Mask R-CNN combined with classical machine learning classifiers. They evaluated both supervised (SVM and Bayes) and unsupervised (K-means and GMM) models to refine the segmentation outputs. Among the tested combinations, Mask R-CNN with K-means clustering achieved the highest accuracy of 97.68% and the fastest runtime, demonstrating its effectiveness for automatic lung region extraction from CT images. As shown in Figure 8, segmentation results from different classifier combinations are compared.

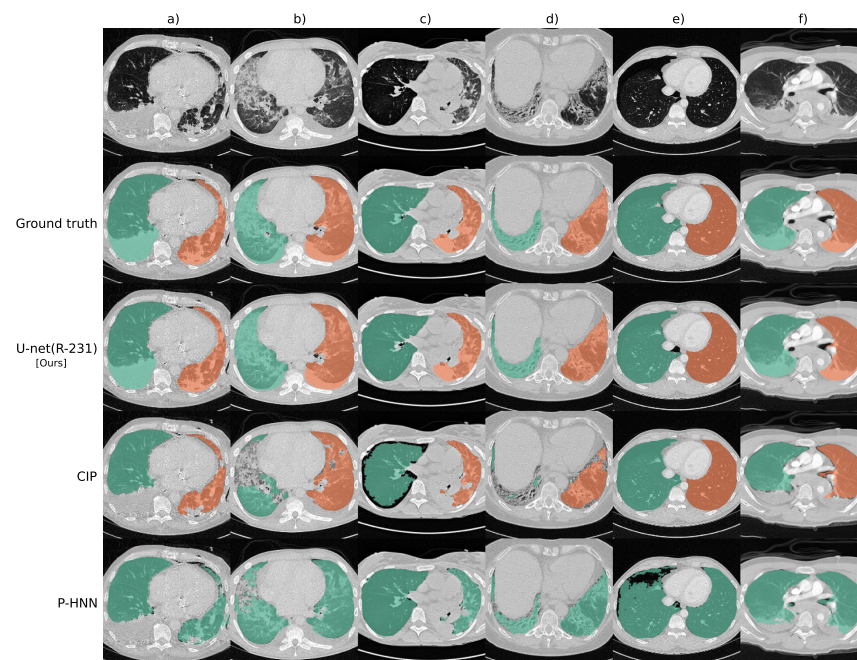


Figure 6. Sample lung segmentation results on routine CT cases: (a) pleural effusion, chest tubes, and consolidation, (b) minor effusions and ground-glass opacity, (c) asymmetrical ventilation and atelectasis, (d) fibrotic changes with traction bronchiectasis, (e) pneumothorax, and (f) trauma-related compression atelectasis [66].

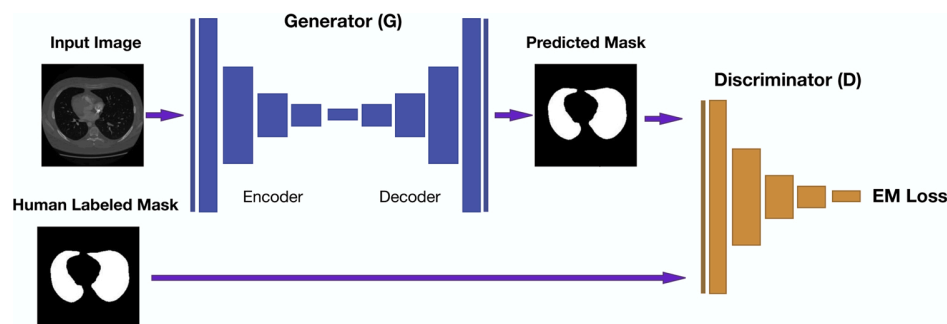


Figure 7. LGAN schema pipeline. A CNN-based generator produces lung masks from CT slices, while a discriminator guides learning by comparing predictions with ground truth. The adversarial setup enhances segmentation accuracy and structural realism [67].

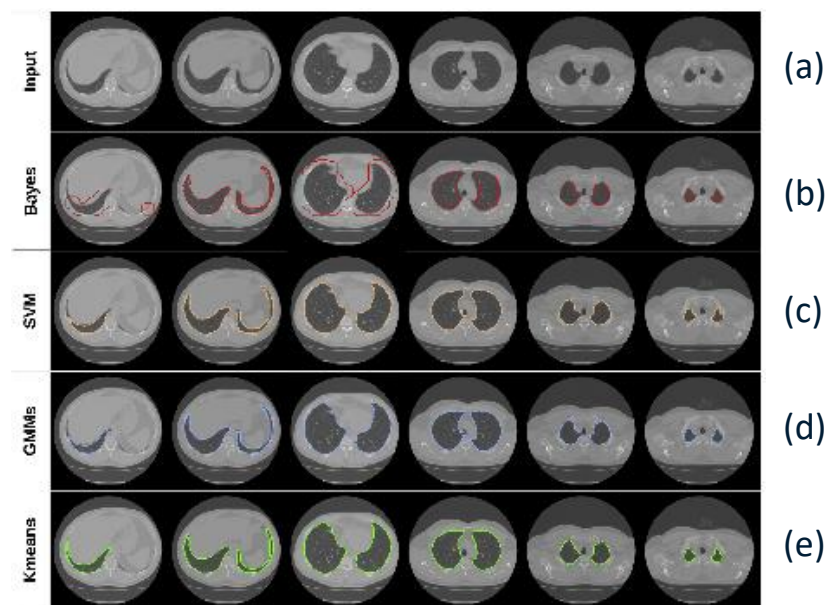


Figure 8. Segmentation results using Mask R-CNN with different classifiers: (a) Input CT image, (b) Mask R-CNN + Bayes, (c) Mask R-CNN + SVM, (d) Mask R-CNN + GMM, and (e) Mask R-CNN + K-means [68].

6. Lung Nodule Detection

After segmenting the lung region, the next step is to identify possible nodules. This is a critical task, as the early detection of lung nodules can improve patient survival rates [69,70]. However, the task is complex. Nodules often appear as round, low-contrast areas and can be easily confused with blood vessels, ribs, or other overlapping structures [57]. In addition, nodules can vary in size, shape, and density, making them harder to detect consistently [13] as shown in Figure 9.

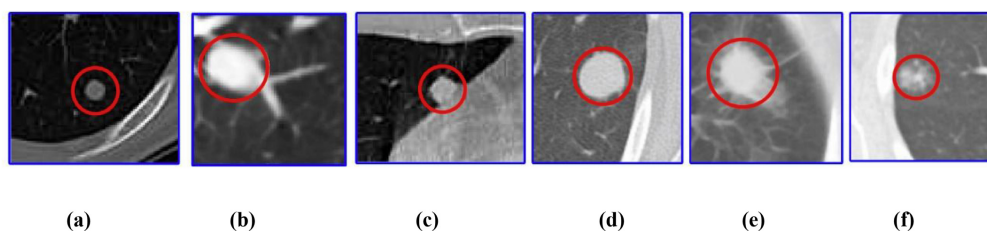


Figure 9. Different types of lung nodules: (a) well circumscribed, (b) juxta-vascular, (c) juxta-pleural, (d) solid, (e) part-solid, and (f) non-solid [13]. Note: Scale bars are not provided in the original figure, as the purpose is to illustrate morphological diversity rather than exact physical dimensions.

6.1. Traditional Methods

Traditional lung nodule detection methods often use handcrafted features and classical machine learning algorithms. Candidate nodules are selected using techniques like thresholding, region growing, or clustering-based methods [1].

Thresholding methods are among the simplest approaches for image segmentation. They use pixel intensity values to separate nodules from the background. If a pixel's intensity is greater than the threshold t , it is assigned to one class; otherwise, it is assigned to another. The mean grayscale $\mu(t)$ and inter-class variance $\sigma_b^2(t)$ can be represented by Equations (6) and (7). $\omega_0(t)$ and $\omega_1(t)$ represent the probabilities (or pixel ratios) of the two pixel classes separated by the threshold t , and $\mu_0(t)$, $\mu_1(t)$ are their respective mean intensity values. The goal is to find the threshold that maximizes $\sigma_b^2(t)$, which indicates the best class separation based on grayscale distribution [6]:

$$\mu(t) = \omega_0(t)\mu_0(t) + \omega_1(t)\mu_1(t), \quad (6)$$

$$\sigma_b^2(t) = \omega_0(t)\omega_1(t)(\mu_0(t) - \mu_1(t))^2. \quad (7)$$

Armato et al. [71] proposed a computerized scheme combining 2D and 3D analysis with gray-level thresholding, feature extraction, and linear discriminant analysis to automatically detect pulmonary nodules in CT scans, achieving an AUC of 0.93 in ROC analysis. Sousa et al. [58] proposed a fully automated and efficient lung nodule detection methodology based on thresholding and skeletonization techniques, achieving high accuracy and reproducibility while enabling large-scale, low-cost CT scan analysis without human intervention. Aresta et al. [72] developed a thresholding-based method using sliding windows and contrast enhancement to detect small juxta-pleural nodules in CT scans, demonstrating competitive results with a lightweight, integrable design.

Region-based methods detect nodules by grouping pixels with similar intensity. Region growing starts from a seed point and expands by including neighboring pixels with similar intensities. To begin, a pixel within the suspected nodule area is manually selected as the initial center $P(0)$. The algorithm then evaluates adjacent pixels $P(x)$ to determine whether they should be added to the region [6]. Growing requirements are shown in Equation (8), and the grayscale difference is shown in Equation (9). This method is simple but needs manual input and is not always accurate for complex structures like juxta-pleural or juxta-vascular nodules [73]:

$$N(n+1) = N(n) \cup P(x), \text{ for } \delta(x) < T, \quad (8)$$

where T is the set threshold, $N(n)$ is the current set of region pixels, $P(x)$ is the candidate neighboring pixel, and $\delta(x)$ is the grayscale difference used for comparison:

$$\delta(x) = g(x) - \frac{1}{n} \sum_{k=0}^n g(k), \quad (9)$$

where $g(x)$ is the grayscale value at position x , and n is the number of pixels currently in the region.

Suárez-Cuenca et al. [74] proposed a pulmonary nodule detection approach based on region growing and classifier combination, demonstrating that majority-vote fusion improves sensitivity over individual models at low false positive rates. Wu et al. [75] developed a region-based pulmonary nodule detection system that utilizes 3D geometric information and segmentation to identify solitary nodules, achieving good sensitivity while highlighting the need for improved ROI extraction to reduce false positives. Parveen et al. [76] proposed a region-based lung nodule detection method that combines threat pixel

identification with region growing to automatically segment suspicious areas, supporting improved diagnostic accuracy without human intervention.

Clustering-based methods group pixels into clusters using algorithms like K-means or fuzzy C-means. These methods help detect nodules by finding similar pixel patterns. They usually improve sensitivity but still face issues with noise and overlapping tissues [77]. Javaid et al. [78] proposed a clustering-based CAD system using K-means for detecting challenging nodules like juxta-vascular and juxta-pleural nodules, achieving high sensitivity and accuracy while reducing false positives through feature-based grouping and SVM classification.

Additional filtering and geometric feature analysis are also needed to improve candidate generation and reduce computation time [79]. Naqi et al. [80] proposed a hybrid-feature-based nodule candidate detection method and demonstrated that combining geometric features with AdaBoost classification achieved high accuracy and sensitivity, particularly improving the detection of challenging juxta-vascular and juxta-pleural nodules. Figure 10 shows the steps in the nodule candidate detection. Subfigure (a) presents the segmented regions of interest (ROIs) extracted from the original CT images. In subfigure (b), the contours of individual structures within each ROI are identified using active contour methods, which adaptively follow object boundaries. Subfigure (c) shows the final candidate regions selected for further analysis. Fotin et al. [81] proposed a multiscale Laplacian of Gaussian (LoG) filtering method for candidate nodule generation that achieved near-perfect sensitivity for both solid and nonsolid nodules, highlighting the importance of additional filtering techniques to optimize early-stage detection accuracy.

Overall, traditional methods are simple and work well in some cases, but their performance is limited by hand-designed features and their sensitivity to noise or image artifacts.

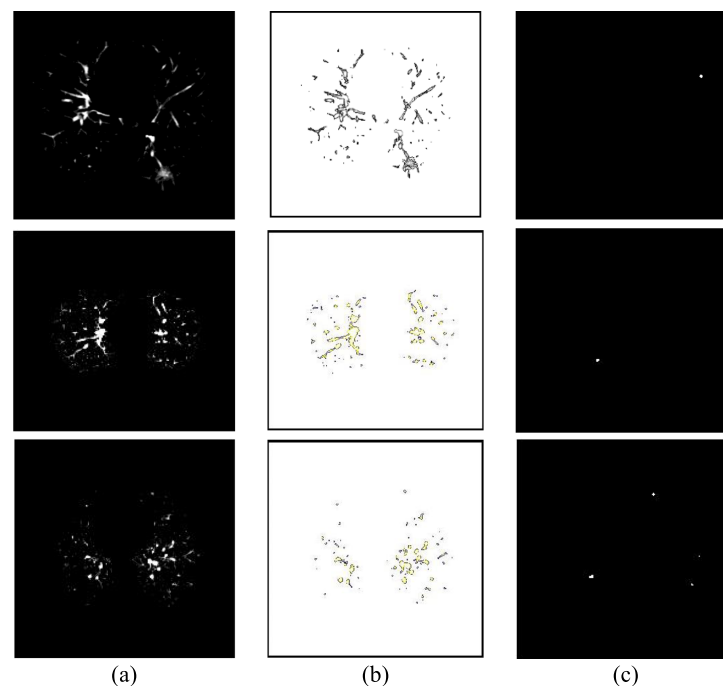


Figure 10. Candidates detection stages: (a) segmented ROIs, (b) corresponding contours of objects within the ROIs, and (c) final selected candidate regions [80]. Note: This image is reused from a published study and has been retained in its original form to preserve the integrity of the method being illustrated.

6.2. Deep Learning-Based Detection

With the rise of deep learning—a subset of machine learning that leverages multi-layer neural networks to automatically learn hierarchical feature representations—many researchers have applied deep neural networks (DNNs) to lung nodule detection [57]. These models can learn both simple and complex features from CT images, which helps improve detection performance. Most deep learning methods are based on CNNs, and they are widely used because of their strong feature learning ability and good results [79]. Figure 11 shows the architecture of the CNN. The model is trained using backpropagation to adjust parameters for feature extraction and classification. As shown in Figure 11, layers close to the input have fewer filters due to the larger spatial resolution, while deeper layers use more filters to capture complex patterns in smaller feature maps. The architecture includes convolutional layers for local feature detection, followed by pooling and fully connected layers for decision-making. MLP stands for multilayer perceptron, which connects the high-level features to the final output [82]. CNN-based methods can be grouped by the type of input they use. Some use 2D slices, while others use 3D volumes [1].

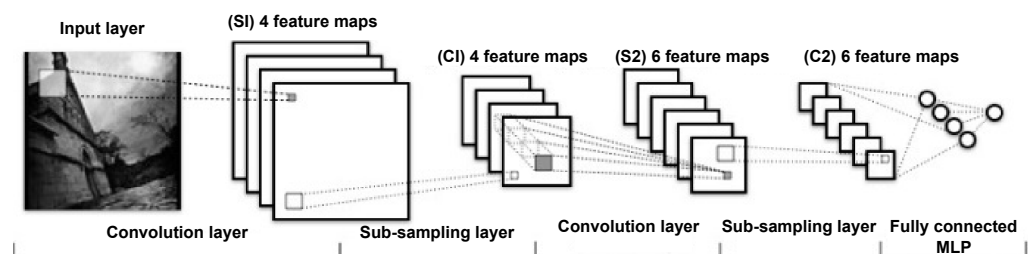


Figure 11. Architecture of a convolutional neural network. This model was constructed from training data using the gradient backpropagation method [82].

Two-dimensional CNNs work on individual CT slices and are easier to train. They work well with fewer resources and can still learn important features. Some 2D networks use slices from different directions (axial, coronal, and sagittal) to keep some 3D information [1]. Kumar et al. [83] proposed a CAD system that uses deep features extracted from a 2D autoencoder and a binary decision tree classifier to distinguish malignant from benign lung nodules, achieving higher accuracy than previous methods on the LIDC dataset. Ciompi et al. [84] proposed a method using an ensemble of classifiers based on 2D views of 3D nodules, leveraging features extracted from the pre-trained ‘OverFeat’ 2D CNNs to classify peri-fissural nodules, achieving near human-level performance (AUC = 0.868). Hua et al. [82] proposed using 2D CNNs and deep belief networks (DBNs) to classify lung nodules from CT images without handcrafted feature extraction, demonstrating that deep learning methods outperform traditional CAD systems in both efficiency and accuracy. Figure 12 shows the DBNs framework. The DBN is formed by stacking multiple restricted Boltzmann machines (RBMs), where each RBM is a two-layer network consisting of a visible layer x for input data and a hidden layer h_1 for feature extraction. There are no intra-layer connections, only inter-layer ones. The hidden layer of one RBM becomes the visible layer of the next. Training is carried out in a layer-wise manner using contrastive divergence and stochastic gradient descent. After the unsupervised pre-training, supervised fine-tuning is used to optimize the entire network for classification tasks. George et al. [85] proposed a real-time lung nodule detection system based on a 2D CNN using the DetectNet architecture derived from YOLO (You Only Look Once), achieving 93% accuracy with low false positives and high sensitivity by treating detection as a single-stage regression task. The network architecture of the nodule detector is shown in Figure 13. The architecture removes the original input, output, and final pooling layers from GoogLeNet.

It omits fully connected layers to allow flexibility for input images of various sizes and enables sliding-window style predictions with a stride of 16 pixels. The network's receptive field covers 555×555 pixels, and the number of feature maps at each stage is indicated in the diagram. Trajanovski et al. [86] proposed a two-stage deep learning framework that combines a 2D CNN-based nodule detector with a ResNet-inspired neural network to assess cancer risk from low-dose CT scans, achieving robust and comparable performance to radiologists across multiple datasets. The architecture of the network is shown in Table 4. The model begins with image input consisting of 10 nodules, each projected in three planes with a resolution of 28×28 . It stacks several convolutional layers with 3×3 kernels and 8 channels, each followed by batch normalization. Residual connections are introduced by merging earlier and deeper convolutional outputs. Dropout is applied for regularization, especially to prevent overfitting. The feature representations are passed through dense (fully connected) layers, and finally merged with nodule-level metadata (radius, coordinates, confidence scores). A final dense layer with a sigmoid activation estimates the malignancy probability for each nodule, and a global max pooling layer aggregates the outputs across all nodules to predict the final cancer risk. Xie et al. [87] proposed a two-stage automated pulmonary nodule detection framework based on 2D CNN and an improved Faster R-CNN with deconvolution and dual region proposal networks, followed by a boosting-based classifier for false positive reduction, achieving high sensitivity on the LUNA16 dataset.

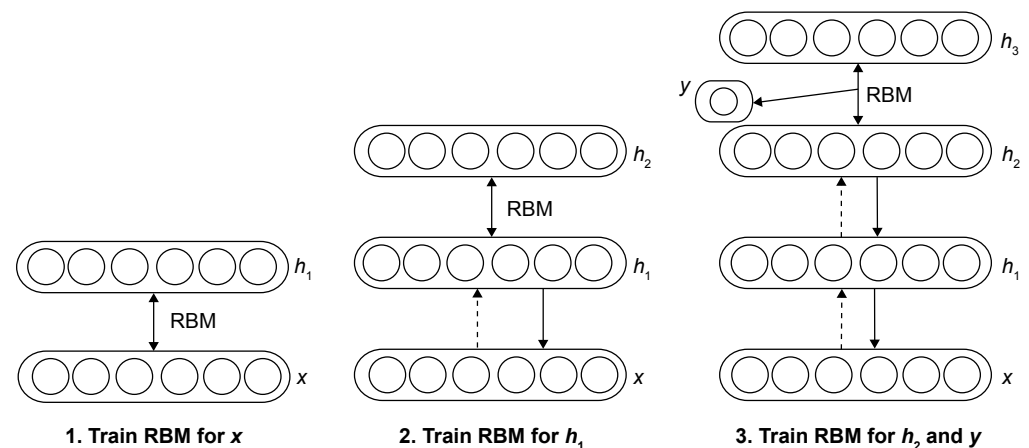


Figure 12. Deep belief network learning framework [82].

Table 4. Architecture of the deep and wide neural network [86].

Layer	Properties	Previous Layer (s)
1. Image input	$(10 \times 3 \times 28 \times 28)$	- Img: 10 nod, 3proj 28×28
2. Conv Layer + BN	$(3 \times 3, 8 \times)$, stride 1	1.
3. Conv Layer + BN	$(3 \times 3, 8 \times)$, stride 1	2.
4. Conv Layer + BN	$(3 \times 3, 8 \times)$, stride 1	3.
5. Conv Layer + BN	$(3 \times 3, 8 \times)$, stride 1	1.
6. Addition/Merge + BN	—	4., 5.
7. Dropout + BN	—	6.
8. Dense + BN	(64)	7.
9. Dropout + BN	—	8.
10. Dense + BN	(64)	9.
11. Numeric input	(10×1)	- Radius
12. Numeric input	(30×1)	- x, y, z nodule coordinates
13. Numeric input	(10×1)	- confidence score
14. Addition/Merge	—	10., 11., 12., 13.
15. Dense + sigmoid	(1)	14.
16. GlobalMaxPool	(10)	15.

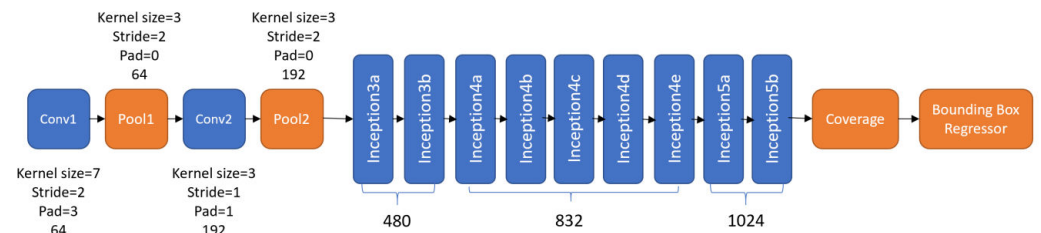


Figure 13. Network architecture of the nodule detector of DetectNet [1].

Three-dimensional CNNs can learn spatial patterns across multiple slices, which helps in detecting irregular or small nodules. However, 3D models require more memory and training time [1]. Dou et al. [88] proposed a 3D CNN framework that integrates multilevel contextual information to effectively reduce false positives in pulmonary nodule detection, achieving top performance in the LUNA16 challenge. The framework includes three separate 3D CNNs, each trained on a different level of spatial context. These networks process volumetric patches of increasing size to capture local to broader anatomical information. The final prediction is produced by combining the outputs (posterior probabilities) from the three subnetworks, enabling the model to make more accurate decisions in false positive reduction for pulmonary nodule detection. Anirudh et al. [89] proposed a 3D CNN-based lung nodule detection system that uses weak supervision with only point labels and estimated size, reducing the need for full 3D annotations while maintaining high sensitivity and low false positives through unsupervised label expansion and multi-scale context learning. Figure 14 shows the network design. The network consists of five convolutional layers with ReLU activation, two max-pooling layers, and a final softmax layer for binary classification. Dropout is applied to prevent overfitting. A multiscale strategy is adopted, where networks are trained at different scales with adjusted kernel sizes. Two of the five convolutional layers are fully connected layers implemented as 1×1 convolutions. Huang et al. [90] proposed a 3D CNN-based fusion model (A-CNN) that integrates multi-scale 3D convolutional networks and AdaBoost to improve pulmonary nodule detection, achieving high sensitivity on the LUNA16 dataset. Ali et al. [91] developed a reinforcement learning-based (RL) method using a deep 3D CNN architecture to detect lung nodules directly from raw CT images without preprocessing, achieving high training accuracy but moderate generalization, highlighting both the promise and data dependency of RL approaches in medical imaging. A flowchart of the convolutional neural network architecture is shown in Figure 15. The network begins with an input layer (blue), followed by multiple convolutional layers (red) with ReLU activation functions to extract features. Pooling layers (purple) reduce spatial dimensions, helping to retain key information. A dropout layer (cyan) helps prevent overfitting. The fully connected layer (green) integrates features for classification, and the final softmax output (yellow) provides the class prediction.

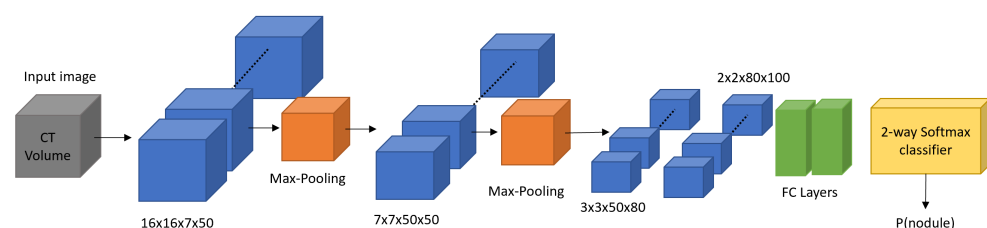


Figure 14. Overall design of the 3D CNNs trained for lung nodule detection [1].

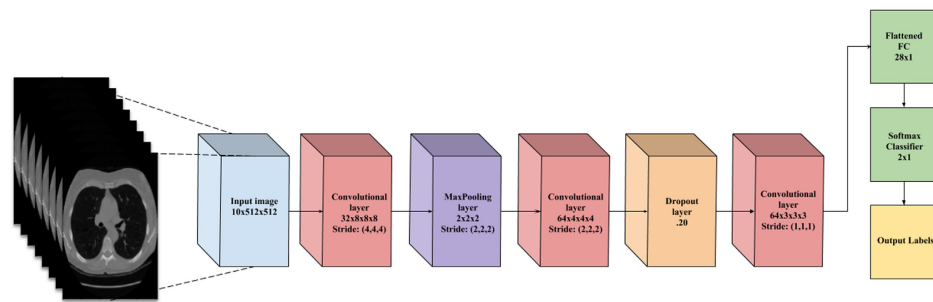


Figure 15. A flowchart of the CNN architecture [91].

Transfer learning is also popular. Pre-trained networks like VGG or ResNet, trained on large datasets, can be used to extract features from CT scans [30]. These features are then used to detect or classify nodules. Paul et al. [92] proposed a hybrid approach for pulmonary nodule detection and classification by combining features extracted from a 3D CNN trained on the NLST dataset with deep features from a pre-trained VGG network and traditional radiomics features, achieving improved performance through multi-channel input and feature fusion. The flowchart of the hybrid approach is shown in Figure 16. This design fuses features from two branches: one processes input images through a max-pooling path, and the other applies a convolution followed by pooling. Both branches produce feature maps of the same size, which are then merged. A final convolution and pooling layer refines the combined features before classification.

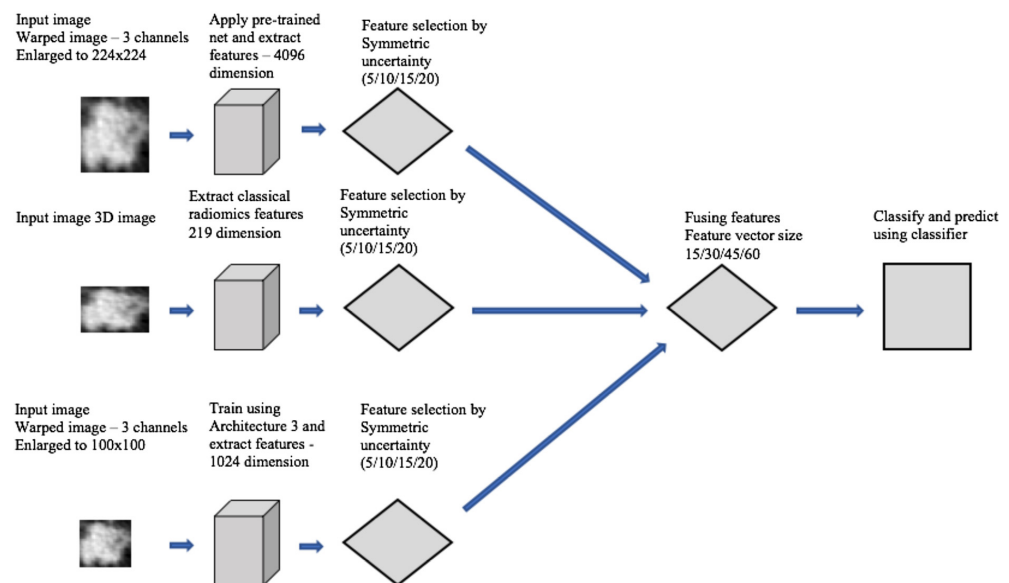


Figure 16. Flowchart of feature fusion approach [92].

Compared to 3D CNNs, 2D models are computationally efficient and require less memory, making them suitable for training with limited resources. However, they often fail to capture inter-slice contextual information, which is critical for detecting small or subtle nodules. On the other hand, 3D models provide richer spatial context and generally achieve better performance in volumetric analysis, but they are computationally expensive and sensitive to variation in slice thickness. Hybrid 2.5D strategies, such as multi-slice inputs, attempt to balance efficiency and contextual learning.

In short, deep learning is flexible and has shown strong performance in many studies, making it a popular and effective approach for lung nodule detection today. Table 5 summarizes the representative detection models previously introduced, detailing their

architectural types, datasets, and key metrics such as accuracy, sensitivity, and AUC. Notable architectural features are also noted to highlight innovations. This table complements the preceding discussion and enables a quick comparison of different detection strategies. Among the listed models, the Faster R-CNN variant [87] shows notably high AUC, likely due to its use of a dual RPN structure and deconvolution layers, which enhance multi-scale feature extraction and improve localization of small nodules. In contrast, the point-supervised method [89] sacrifices performance for annotation efficiency, which makes it more suitable for settings with limited labeling resources but less optimal for high-sensitivity tasks.

Table 5. Comparison of representative deep learning-based detection models. Only reported metrics are shown. Abbreviations: ACC (Accuracy), Sens. (Sensitivity), AUC (Area Under Curve).

Type	Model/Method	Dataset	ACC (%)	Sens. (%)	AUC	Notes
2D CNN-based	DetectNet [85]	LIDC-IDRI	93.00	89.00	–	F1-score = 90.96, single-stage
	Faster R-CNN Variant [87]	LUNA16	–	86.42	95.4	Deconv + dual RPN
3D CNN-based	Multilevel CNN [88]	LUNA16	–	94.40	–	Multi-context fusion
	Point-supervised [89]	LUNA16	–	80.00	–	Weak supervision
Hybrid	3D CNN + VGG + Radiomics [92]	NLST	76.79	–	78	Feature-level fusion

7. Lung Nodule Segmentation

After detecting the possible lung nodules, the next step is to segment them. This means accurately outlining the shape and size of each nodule [93]. Like detection, segmentation is also very important. It helps measure the nodules and supports further diagnosis or classification. However, nodule segmentation is also challenging because nodules come in different sizes, shapes, and textures. Some are attached to the lung wall or blood vessels, which makes their boundaries unclear [94].

7.1. Traditional Methods

Traditional segmentation methods use simple image processing rules, like thresholding, morphological operations and region growing [10].

Thresholding is one of the simplest ways. It checks the brightness of each pixel and separates nodules from other parts by choosing a value (threshold) [95]. It works fast but may fail when nodules are close to other tissues or have unclear edges [10]. Zhao et al. [96–99] developed an early multi-criterion segmentation method based on thresholding, which uses density, gradient, and shape constraints to separate nodules from surrounding structures in helical CT images, helping to improve edge accuracy and support early malignancy assessment. This approach was adapted by several early methods as a foundation for threshold-based segmentation. Wiemker et al. [100] proposed a fast and fully 3D thresholding algorithm to segment lung nodules by finding the optimal Hounsfield threshold based on iso-surface gradients and sphericity, achieving real-time performance and high accuracy.

Mathematical morphology uses shape operations like dilation and erosion to separate nodules from vessels or the chest wall. It is useful for cleaning the segmented result but needs careful tuning. Kubota et al. [101] proposed a general-purpose segmentation algorithm using mathematical morphology and figure–ground separation techniques,

which effectively handles various nodule types—including juxta-pleural and vascularized nodules—without requiring prior lung wall separation, and shows comparable or better performance than existing methods on LIDC datasets. Halder et al. [102] proposed an adaptive morphology-based segmentation method using adaptive structuring elements and morphological filters to improve nodule detection and segmentation in CT images, achieving high accuracy on both public and private datasets by combining shape, texture, and intensity features with an SVM classifier.

Region growing starts from one or more seed points and adds nearby pixels with similar intensity. It can be more flexible than thresholding but results depend on the choice of seeds and similarity rules. Dehmeshki et al. [103] proposed a region growing segmentation method based on fuzzy connectivity and adaptive contrast, which showed high reproducibility and good performance in segmenting various types of pulmonary nodules in CT images. Soltani-Nabipour et al. [104] proposed an improved region growing algorithm with multipoint initialization and edge correction, achieving 98% segmentation accuracy on lung tumors and showing its potential for other cancers like brain and breast tumors. Equation (10) shows the rules when these points will be added to the points array (seedval is the color intensity of this point stored in the base value; threshval is the maximum radius expected for a tumor, which is entered by the user):

$$\begin{aligned} \text{seedval} - \text{threshval} &\leq \text{pointval} \leq \text{seedval} + \text{threshval}, \\ \text{point} &\in \text{lungregion}, \\ \text{distance}(\text{seed}, \text{point}) &\leq 2 \times \text{radiusval}. \end{aligned} \quad (10)$$

Other traditional methods include graph cuts, watershed, statistical classification, and model fitting. While these approaches can work well, they often need manual steps or careful parameter tuning. Also, their performance drops when dealing with small or irregularly shaped nodules [10].

7.2. Deep Learning-Based Segmentation

Recently, deep learning has become the most popular approach for lung nodule segmentation. These methods use neural networks to learn patterns directly from data. The most common network is the CNN [4].

Many works use well-known CNNs for nodule segmentation [9]. U-net [105] is widely used because of its simple encoder–decoder structure with skip connections, which helps keep important features. Tong et al. [106] proposed an improved U-net-based segmentation method for pulmonary nodules by introducing residual connections and batch normalization to enhance network depth, training stability, and segmentation accuracy using the LUNA16 dataset. Figure 17 shows a 45-layer network architecture, including 19 convolutional layers. The downsampling path uses repeated 3×3 convolutions and 2×2 max pooling with ReLU activation to extract high-dimensional features. In the upsampling path, the spatial resolution is doubled while feature channels are halved. Each upsampled feature map is concatenated with its corresponding feature map from the downsampling path to recover boundary details. A final 1×1 convolution maps the output to the target classes. V-Net [107] is a 3D version that works well for volumetric CT data. Ma et al. [108] proposed an improved V-Net model that combines pixel threshold separation and attention mechanisms to enhance lung nodule segmentation, achieving high Dice scores and outperforming most existing U-net-based models. The improved Dig-CS-VNet model shown in Figure 18 enhances V-Net by adding a feature separation module and 3D attention blocks to better capture edge and semantic details. Skip connections between the encoder and decoder help improve segmentation accuracy. ResNet [109] has also been used to capture deeper

and more complex patterns. Cao et al. [110] proposed a Dual-branch Residual Network (DB-ResNet) that combines intensity-based and convolutional features using a novel central intensity-pooling layer, achieving higher segmentation accuracy than experienced radiologists on the LIDC-IDRI dataset, demonstrating the strength of ResNet-based methods for lung nodule segmentation. Figure 19 shows the proposed architecture of DB-ResNet which includes three key components: (a) a main structure where average pooling (AP) and concatenation (Concate) operations are used, and the central intensity-pooling module can be inserted at marked positions; (b) a convolution block (ConvBlock) design; and (c) a residual block (ResBlock) layout. The variables k , m , and n represent the channel numbers in each layer. This modular design enhances feature extraction and network depth while keeping the model efficient.

To improve accuracy, some methods adopt multi-view or multi-scale strategies. Multi-view CNNs process CT slices from different angles (axial, sagittal, and coronal) to learn more spatial features [9]. Wang et al. [111] proposed a multi-view CNN (MV-CNN) architecture that combines three orthogonal image views for lung nodule segmentation, achieving strong performance on the LIDC-IDRI dataset without requiring predefined shape assumptions or parameter tuning. The MV-CNN architecture consists of three branches, each designed to extract features from axial, coronal, and sagittal views. Each branch processes two-scale patches and includes six convolutional layers, two max-pooling layers, and a fully connected layer. The outputs of all branches are combined in another fully connected layer. Liu et al. [112] proposed MSMV-Net, a multi-scale and multi-view CNN framework that improves lung tumor segmentation by fusing 2D views and applying uncertainty-based deep supervision, achieving strong performance on public datasets. Figure 20 shows the comprehensive architecture of MSMV-Net. The model has two parts: (a) multi-scale multi-view fusion combines sagittal, coronal, and axial views, each weighted by attention parameters α , β , and γ ; (b) multi-scale uncertainty-based deep supervision uses softmax-derived uncertainty to guide training and improve accuracy.

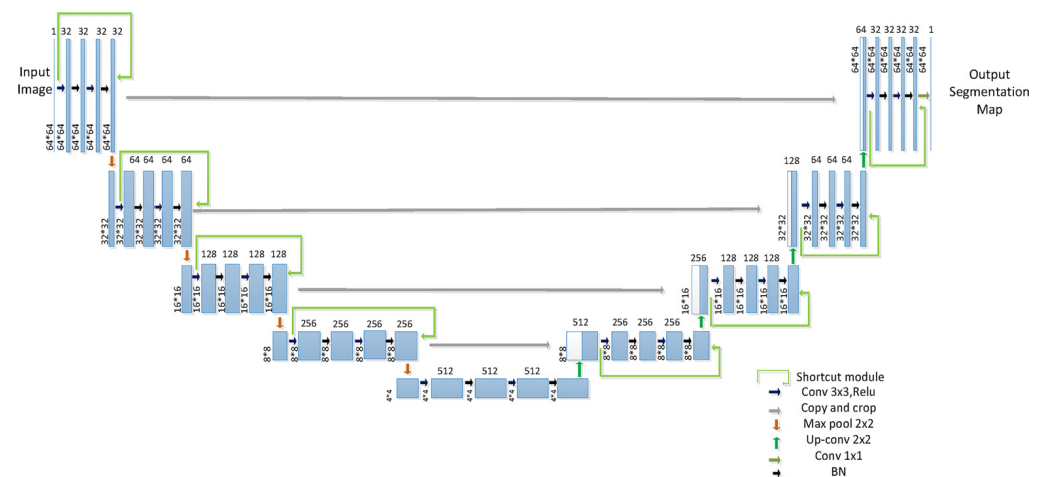


Figure 17. Improved U-net network structure. This 45-layer model enhances pulmonary nodule segmentation by integrating residual connections and batch normalization. The encoder path (left) extracts features using repeated convolutions and max pooling, while the decoder path (right) restores spatial resolution via upsampling and skip connections. A final 1×1 convolution produces the segmentation map [106].

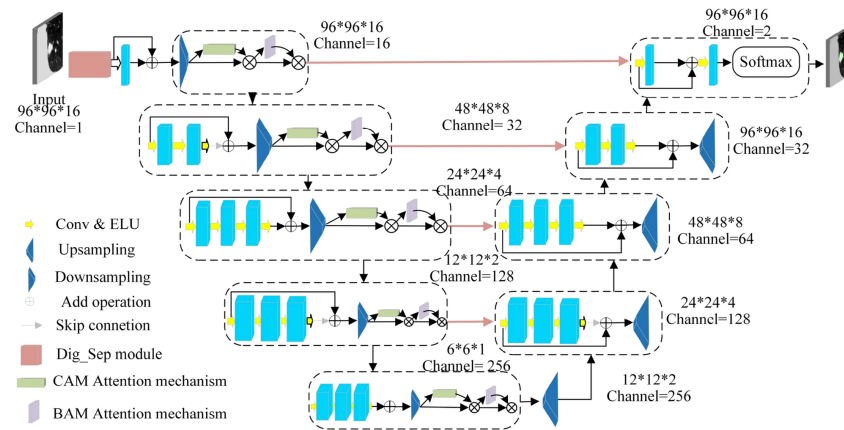


Figure 18. Improved Dig-CS-VNet network architecture [108].

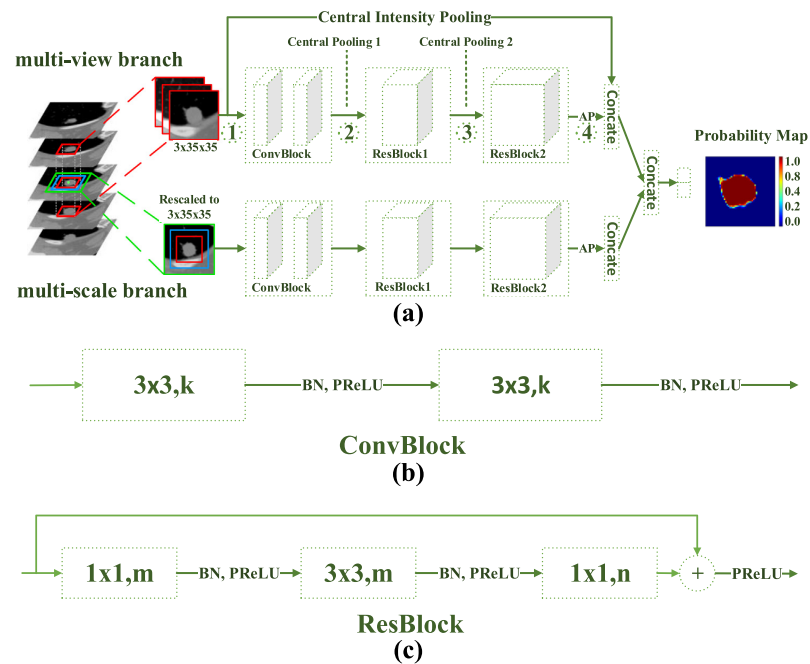


Figure 19. DB-ResNet network architecture: (a) Main architecture, with Average Pooling (AP), Concatenate (Concat), and Central Intensity-Pooling placement, (b) Diagram of the convolution block. k , m , and n indicate channel numbers, and (c) Diagram of the residual block [110].

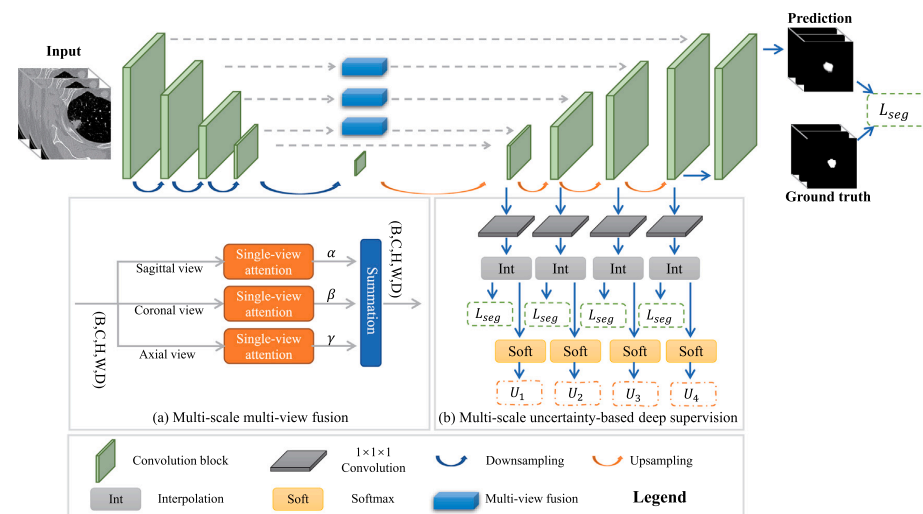


Figure 20. MSMV-Net network architecture [112].

Some models combine CNNs with other advanced techniques. Attention-based networks help the model focus on key areas, such as nodules, while ignoring irrelevant regions [113]. Bruntha et al. [114] proposed Lung_PAYNet, an attention-based pyramidal network that combines inverted residual blocks and swish activation to enhance lung nodule segmentation in low-dose CT scans, achieving a Dice score of 95.7% on the LIDC-IDRI dataset. GANs [115] have been applied to enhance segmentation quality by generating more realistic and structurally consistent masks. Tyagi et al. [116] proposed a 3D conditional GAN (CSE-GAN) for lung nodule segmentation. The generator adopts a U-Net backbone with squeeze-and-excitation (CSE) modules, while the discriminator employs spatial squeeze and channel excitation (sScE) blocks to better distinguish real and generated masks. As shown in Figure 21, the model architecture is designed to enhance both feature representation and mask discrimination. Evaluated on LUNA16 and a local dataset, the method achieved Dice scores of 80.74% and 76.36%, demonstrating clear improvements over standard CNN-based methods. As illustrated in Figure 22, the proposed CSE-GAN performs well across different nodule sizes—small (Case 1), medium (Case 2), and large (Case 3). Each case includes the original CT slice with enlarged patches (a), the ground truth mask (b), and the predicted segmentation (c), highlighting the model's ability to produce accurate and consistent outputs. Other methods use Transformers [117] to capture long-range dependencies in the image, which is useful for precise boundary segmentation. Usman et al. [118] proposed a two-stage segmentation framework using a dual-encoder hard attention network (DEHA-Net) and transformer-inspired attention modules to achieve accurate and robust 3D lung nodule segmentation across multiple CT views.

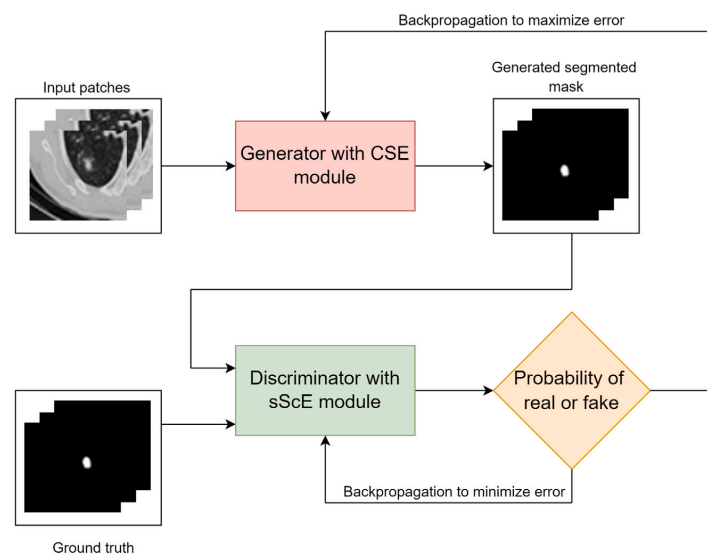


Figure 21. Overview of the CSE-GAN architecture [116].

Deep learning methods have shown strong performance in lung nodule segmentation. Standard CNNs are simple and effective. Multi-view and multi-scale models improve feature learning. Hybrid networks add flexibility and precision, while two-stage systems help target difficult cases. Each method has strengths, and the choice often depends on the available data and task needs. Table 6 presents a comparative overview of the key segmentation models applied to lung nodules previously introduced. It reports performance metrics such as DSC and Sensitivity, along with dataset details and unique architectural components. This summary supports the prior discussion by clarifying how segmentation performance varies across design paradigms. Models such as Dig-CS-VNet [108] demonstrate superior Dice performance. This is largely due to their integration of 3D attention mechanisms and feature separation modules, which help better preserve spatial

continuity, especially for small or low-contrast nodules. Transformer-based models like DEHA-Net [118] further leverage global attention to enhance boundary delineation, but their complexity may limit scalability in real-time clinical applications. While GAN-based models such as CSE-GAN [116] introduce adversarial refinement to improve mask realism, their performance can be unstable and highly dependent on discriminator design and training balance.

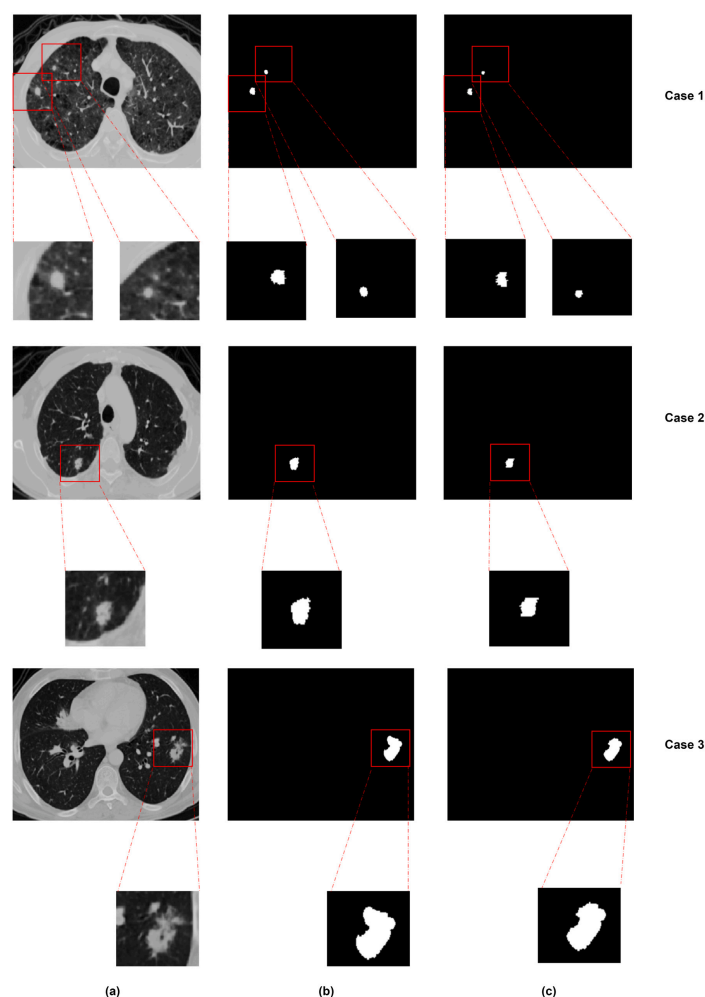


Figure 22. Segmentation results for nodules of different sizes: (a) original image, (b) ground truth, and (c) predicted segmentation [116].

Table 6. Comparison of representative deep learning-based segmentation models. Only the reported metrics are shown. Abbreviation: DSC (Dice Similarity Coefficient) and Sens. (Sensitivity).

Type	Model/Method	Dataset	DSC (%)	Sens. (%)	Notes
2D CNN-based	Improved U-Net [106]	LUNA16	73.6	–	Residual connections, batch normalization, skip connections
3D CNN-based	Improved Dig-CS-VNet [108]	LUNA16, LNDb	94.9, 81.1	92.7, 76.9	Feature separation, 3D attention blocks
ResNet-based	DB-ResNet [110]	LIDC-IDRI	89.40	–	Dual-branch structure, intensity pooling, modular ResBlocks

Table 6. Cont.

Type	Model/Method	Dataset	DSC (%)	Sens. (%)	Notes
Multi-view CNN	MV-CNN [111]	LIDC-IDRI	82.74	–	Axial/coronal/sagittal branches, late fusion
Multi-scale + view	MSMV-Net [112]	LUNA16, MSD	55.60, 59.94	–	2D view fusion, attention-weighted deep supervision
GAN-based	CSE-GAN [116]	LUNA16, ILND	80.74, 76.36	85.46, 82.56	3D U-Net + CSE attention, sScE discriminator
Transformer-based	DEHA-Net [118]	LIDC-IDRI	87.91	90.84	Dual-encoder, transformer modules, hard attention

8. MTL in Medical Image Analysis

MTL has gained increasing attention in medical image analysis due to its ability to improve model performance by learning shared representations across related tasks. In the context of lung CT imaging, where detection and segmentation tasks are often interconnected, MTL offers a promising framework for optimizing learning efficiency and reducing annotation cost. This section reviews the principles of MTL and explores its specific applications in medical imaging, with a focus on lung nodule analysis.

8.1. Overview of MTL

MTL means training a model to perform more than one task at the same time. Instead of training separate models for each task, MTL combines them and shares useful information between tasks [119]. This helps the model learn better, especially when one task has less labeled data but is related to another task with more data [19]. By sharing the learning process, MTL can improve accuracy, reduce overfitting, and use less memory. It is also helpful when tasks are related or have overlapping features [18]. As illustrated in Figure 23, the model usually includes two main parts: shared layers and task-specific layers. The shared layers are used to extract common features across tasks, while the task-specific layers process these shared features further to produce outputs for each individual task.

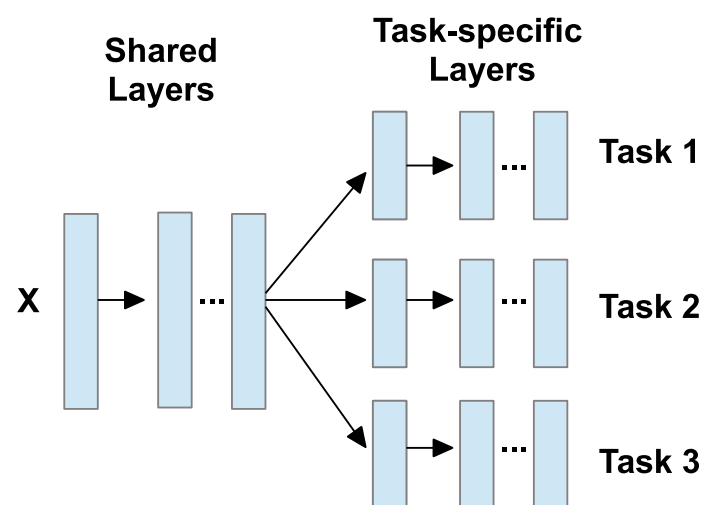


Figure 23. MTL model framework [18].

8.2. Applications in Medical Imaging

In medical imaging, MTL is useful because many tasks are linked. For example, segmentation and classification often go together. MTL has been used in many medical areas, such as brain [120,121], chest [122,123], cardiac [124,125], and abdominal imaging [126,127]. Common tasks include identifying organs, segmenting tumors, classifying disease types, or estimating clinical scores [19,128]. Depending on the task, researchers design different MTL structures. These include cascaded (task B depends on task A), parallel (tasks are done side-by-side), interacted (tasks share features), and hybrid (a mix of the above) [19]. Table 7 summarizes each type, and Figure 24 shows the embodiments of these four types. Y_A and Y_B are the outputs of tasks T_A and T_B . In the figure, these are marked with gray and green blocks; F_A and F_B are the networks designed for each task separately. They are shown as yellow and orange blocks; F_C stands for the layers that both tasks use together, shown in blue in the figure; and $F_{A,B}$ and $F_{B,A}$ describe how the networks for tasks T_A and T_B pass information to each other. These connections are shown as dashed black lines between yellow and orange blocks.

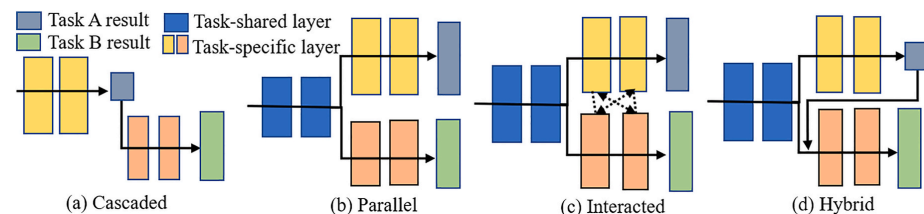


Figure 24. Four embodiments of MTL network architectures: (a) cascaded, (b) parallel, (c) interacted, and (d) hybrid [19].

Table 7. Common MTL architectures [19].

Architecture	Description and Usage	Example	Formula
Cascaded	Tasks run one by one. The following task builds upon the outcome of the preceding task. Best for strong task dependency.	TS-MDL framework first segments teeth with iMeshSegNet, then detects landmarks with PointNet-Reg based on segmentation output [129].	$\{Y_A, Y_B\} = F_B(F_A(\mathbf{X}))$, where $Y_A = F_A(\mathbf{X}); Y_B = F_B(Y_A)$.
Parallel	A shared encoder feeds multiple separate decoders. Tasks are related but not dependent.	COMiT-Net uses one encoder for CXR images and separate decoders for COVID classification and segmentation [130].	$\{Y_A, Y_B\} = \{F_A(f), F_B(f)\}$, $f = F_C(\mathbf{X}); Y_A = F_A(F_C(\mathbf{X})); Y_B = F_B(F_C(\mathbf{X}))$.
Interacted	Task-specific layers exchange features. Useful when tasks benefit from each other.	AMTA-Net segments bladder and rectum first, then uses attention to guide prostate bed segmentation [131].	$\{Y_A, Y_B\} = \{F_{A,B}(f), F_{B,A}(f)\}$, where $f = F_C(\mathbf{X}); Y_A = F_{A,B}(F_C(\mathbf{X})); Y_B = F_{B,A}(F_C(\mathbf{X}))$.
Hybrid	Mix of cascaded, parallel, and/or interaction. Suited for complex task combinations.	DeepSC-COVID jointly performs 3D lesion segmentation and COVID diagnosis using multiple interacting subnets [132].	$\{Y_A, Y_B\} = \{F_A(f), F_B(f, F_A(f))\}$, where $f = F_C(\mathbf{X}); Y_A = F_A(f); Y_B = F_B(f, Y_A)$.

Each architecture has its own strengths and is suited to different application scenarios. Cascaded models are typically preferred when strong task dependencies exist, such as segmentation guiding classification. Parallel architectures are widely adopted for their implementation simplicity and are effective when tasks are loosely coupled. Interacted designs allow for richer feature exchange, which can benefit highly correlated tasks like lesion detection and boundary refinement. Hybrid models combine the advantages of multiple strategies and are well-suited for complex workflows, albeit with increased computational cost. Ultimately, the choice of architecture should be guided by the specific task requirements, data availability, and deployment constraints.

8.3. MTL in Lung Nodule Detection and Segmentation

In lung imaging, MTL is often used to detect and segment nodules. Some systems first use a detection model like Faster R-CNN to find possible nodules. Then, they apply a segmentation model such as U-net or Mask R-CNN to obtain the exact shape. This two-step process works well for small or unclear nodules and helps reduce false positives [19].

Tang et al. [14] introduced NoduleNet, an end-to-end 3D DCNN that jointly performs nodule detection, false positive reduction, and segmentation. To reduce interference between tasks, they decoupled detection and FP reduction features and added a segmentation refinement module. Although not originally labeled as MTL, their model follows a multi-task structure and demonstrates that joint training improves detection accuracy by over 10% compared to single-task baselines. Figure 25 shows the overview of NoduleNet, including three main stages arranged in sequence: first, an initial stage for detecting possible nodule regions; second, a stage to reduce false positives; and third, a stage for refining the segmentation results. The variable k refers to the number of anchor boxes used in detection, while FC denotes a fully connected layer used in the classification or regression parts of the network.

Song et al. [15] proposed a weakly supervised MTL segmentation framework combining 2D U-net and ConvLSTM to leverage sequential context in CT slices. They used annotated data from different tasks to extract shared features and addressed class imbalance with a dynamic multi-task loss weighting strategy. The method improved MIOU and Dice to 79.23% and 82.26% on LIDC-IDRI, showing the benefit of MTL under limited supervision. The framework combines a U-shaped encoder–decoder network with ConvLSTM modules to capture spatial and sequential features from 2D CT slice sequences. It also integrates a multi-branch design for multiple tasks and applies a prior-based loss to improve performance under limited annotation.

Luo et al. [16] proposed a 3D U-net-based MTL framework (ECANodule) for joint pulmonary nodule detection and segmentation. They introduced dense skip connections and a lightweight channel attention module to enhance semantic feature learning. Their approach integrates detection and segmentation via shared encoder parameters and a refined decoder, achieving 91.1% detection sensitivity and 83.4% Dice score on LIDC-IDRI. The use of online hard example mining (OHEM) further addresses class imbalance during training. This framework includes three parts: (1) an improved U-net with a region proposal module to detect nodule candidates, (2) a false positive reduction branch to refine candidate boxes and scores, and (3) a mask branch to enhance segmentation. Each cube shows spatial dimensions and channel number.

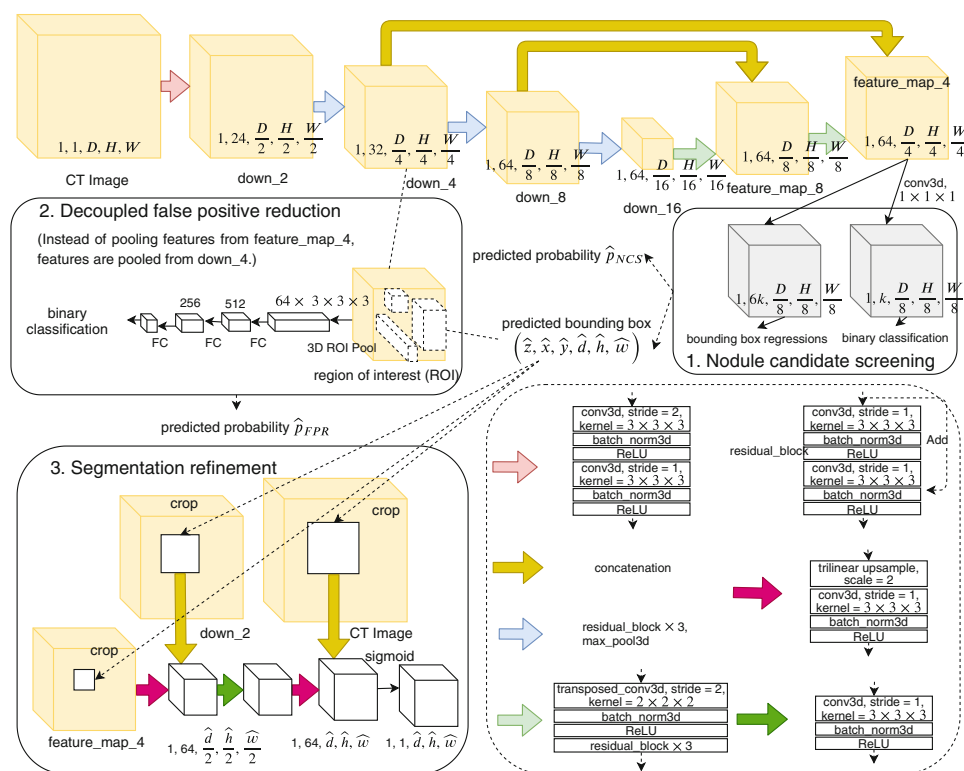


Figure 25. Overview of NoduleNet [14].

Nguyen et al. [133] developed MANet, an MTL model based on a full 3D U-net structure with deep supervision and multi-branch attention mechanisms. The model includes three new modules—projection, fast cascading context, and boundary enhancement—to capture both global and local features. Their refinement step improves detection and segmentation performance, achieving superior FROC and IoU/DSC scores compared to existing methods.

Tang et al. introduced modularity by decoupling detection and segmentation tasks, reducing task interference but lacking dynamic coordination between branches. Song et al. focused on weak supervision and sequential context modeling, making their approach suitable for low-annotation scenarios, though the reliance on 2D slices may limit spatial precision. Luo et al. and Nguyen et al. emphasized architectural refinements, such as dense skip connections, attention modules, and deep supervision, resulting in improved segmentation accuracy. However, these improvements often come with increased model complexity. For instance, Nguyen's MANet achieved superior FROC and Dice scores but required high memory, which can constrain deployment. Table 8 compares representative MTL frameworks previously introduced, presenting key performance metrics such as FROC and Dice, dataset sources, and core architectural strategies like feature decoupling, attention modules, or weak supervision. These comparisons highlight the trade-offs in design choices and their impact on model effectiveness. Multi-task frameworks like NoduleNet [14] benefit from task decoupling, which mitigates performance degradation and reduces false positives by refining spatial boundaries through segmentation. ECANodule [16], on the other hand, excels at detecting small or ambiguous nodules by integrating attention mechanisms and hard example mining, effectively concentrating the model's attention on diagnostically critical regions.

Table 8. Comparison of representative MTL-based models for lung nodule detection and segmentation. Only reported metrics are shown. Abbreviations: Weakly-sup. (Weak supervision), FROC (Free-response ROC), DSC (Dice Similarity Coefficient).

Type	Model/Method	Dataset	FROC (%)	DSC (%)	Key Features
Cascaded	NoduleNet [14]	LIDC-IDRI	+10.27% vs. single-task	83.1	Decoupled tasks, FP reduction, segmentation refinement module
Parallel (Weakly-sup.)	Multi-branch U-net + ConvLSTM [15]	LIDC-IDRI	+6.89% vs. single-task	82.26	Sequential context, dynamic loss weighting, weak labels
Hybrid	ECANodule [16]	LIDC-IDRI	91.1	83.4	Dense skip connections, attention, OHEM training
Hybrid (Deep-attention)	MANet [133]	LIDC-IDRI	88.11	82.74	Deep supervision, multi-branch attention, boundary enhancement

9. Research Gaps and Future Directions

Despite significant progress in lung nodule analysis using deep learning, there are still several open issues that this study aims to address.

First, task imbalance is a common issue in MTL frameworks. Many existing models treat detection and segmentation equally, often ignoring their intrinsic differences. While detection focuses on localizing nodules, segmentation requires precise boundary delineation. Shared features without proper task-specific adaptation may cause one task—typically detection—to dominate, especially in cases with small or indistinct nodules. Most studies adopt fixed loss weights, and adaptive strategies such as uncertainty-based weighting or gradient normalization are rarely investigated in the context of lung CT [19].

Second, while heavy models like full 3D U-nets or large transformers are common in MTL research, there is limited exploration of more efficient backbones such as 3D ResNet-18 in multi-task pipelines. It remains under-explored how these lighter backbones perform when combined with specialized modules like YOLO and U-net to achieve accurate results with reduced computational cost [14,16,133].

Third, although most current methods are developed using datasets like LIDC-IDRI and LUNA16, their real-world generalizability remains questionable. In clinical practice, variations in scanner types, imaging protocols, and patient populations can significantly affect performance. However, few studies have conducted cross-institutional validation or robustness testing, and domain adaptation or self-supervised learning remains under-explored in this context [15]. This lack of external validation presents a major barrier to clinical deployment and must be addressed in future research.

These gaps suggest the need for more adaptive, efficient, and generalizable MTL frameworks to support lung nodule analysis in real-world environments. Beyond improving model architectures, future research should also consider collaborative benchmark design and task-specific evaluation criteria, which would support more consistent comparison and accelerate clinical adoption.

10. Conclusions

This review has examined recent developments in deep learning-based methods for lung nodule detection and segmentation, with a particular focus on MTL frameworks. We have categorized MTL strategies into parallel, cascaded, interacted, and hybrid architectures, and reviewed how these frameworks share representations and improve efficiency and accuracy in pulmonary nodule analysis.

While MTL approaches have shown strong potential, several important challenges remain. Many models lack adaptive coordination between tasks, leading to imbalanced learning that may hinder performance—especially when nodules are small or difficult to segment. The reliance on large 3D architectures also limits real-world clinical deployment due to high computational demands. Additionally, current models often generalize poorly beyond curated datasets, and fully annotated training data can be difficult to obtain.

Addressing these challenges will require further research into dynamic task-balancing strategies, such as attention-based feature sharing and adaptive loss weighting. Lightweight architectures should be explored and optimized for clinical settings, with careful evaluation of efficiency metrics. Self-supervised or weakly supervised learning techniques offer promise for reducing annotation costs and improving robustness. Ultimately, validating these models in diverse, real-world settings is essential for achieving practical, trustworthy CAD tools. As the field advances, closer collaboration between researchers, radiologists, and healthcare providers will be essential toward building efficient, generalizable, and clinically impactful lung nodule analysis systems.

Author Contributions: Conceptualization, R.L. and B.H.S.A.; methodology, R.L. and B.H.S.A.; resources, B.H.S.A. and R.L.; writing—original draft preparation, B.H.S.A. and R.L.; writing—review and editing, B.H.S.A. and R.L.; supervision, B.H.S.A.; visualization, R.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Riquelme, D.; Akhloufi, M.A. Deep learning for lung cancer nodules detection and classification in CT scans. *Ai* **2020**, *1*, 28–67. [CrossRef]
2. Marinakis, I.; Karampidis, K.; Papadourakis, G. Pulmonary Nodule Detection, Segmentation and Classification Using Deep Learning: A Comprehensive Literature Review. *BioMedInformatics* **2024**, *4*, 2043–2106. [CrossRef]
3. Cancer Research UK. Lung Cancer Statistics. 2025. Available online: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer#lung_stats1 (accessed on 5 June 2025).
4. Wang, Y.; Mustaza, S.M.; Ab-Rahman, M.S. Pulmonary Nodule Segmentation using Deep Learning: A Review. *IEEE Access* **2024**, *12*, 119039–119055. [CrossRef]
5. MacMahon, H.; Naidich, D.P.; Goo, J.M.; Lee, K.S.; Leung, A.N.; Mayo, J.R.; Mehta, A.C.; Ohno, Y.; Powell, C.A.; Prokop, M.; et al. Guidelines for management of incidental pulmonary nodules detected on CT images: From the Fleischner Society 2017. *Radiology* **2017**, *284*, 228–243. [CrossRef]
6. Zheng, L.; Lei, Y. A review of image segmentation methods for lung nodule detection based on computed tomography images. *MATEC Web Conf.* **2018**, *232*, 02001. [CrossRef]
7. Pehrson, L.M.; Nielsen, M.B.; Ammitzbøl Lauridsen, C. Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: A systematic review. *Diagnostics* **2019**, *9*, 29. [CrossRef] [PubMed]
8. Alshayegi, M.H.; Abed, S. Lung cancer classification and identification framework with automatic nodule segmentation screening using machine learning. *Appl. Intell.* **2023**, *53*, 19724–19741. [CrossRef]
9. Li, R.; Xiao, C.; Huang, Y.; Hassan, H.; Huang, B. Deep learning applications in computed tomography images for pulmonary nodule detection and diagnosis: A review. *Diagnostics* **2022**, *12*, 298. [CrossRef]
10. El-Baz, A.; Beache, G.M.; Gimel farb, G.; Suzuki, K.; Okada, K.; Elnakib, A.; Soliman, A.; Abdollahi, B. Computer-aided diagnosis systems for lung cancer: Challenges and methodologies. *Int. J. Biomed. Imaging* **2013**, *2013*, 942353. [CrossRef]
11. Zhu, Z.; Sun, Y.; Honarvar Shakibaei Asli, B. Early Breast Cancer Detection Using Artificial Intelligence Techniques Based on Advanced Image Processing Tools. *Electronics* **2024**, *13*, 3575. [CrossRef]
12. Sun, Y.; Zhu, Z.; Asli, B.H.S. Automated Classification and Segmentation and Feature Extraction from Breast Imaging Data. *Electronics* **2024**, *13*, 3814. [CrossRef]

13. Zhang, G.; Jiang, S.; Yang, Z.; Gong, L.; Ma, X.; Zhou, Z.; Bao, C.; Liu, Q. Automatic nodule detection for lung cancer in CT images: A review. *Comput. Biol. Med.* **2018**, *103*, 287–300. [CrossRef]
14. Tang, H.; Zhang, C.; Xie, X. Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, 13–17 October 2019; Proceedings, Part VI 22; Springer: Berlin/Heidelberg, Germany, 2019; pp. 266–274.
15. Song, G.; Nie, Y.; Zhang, J.; Chen, G. Multi-task weakly-supervised learning model for pulmonary nodules segmentation and detection. In Proceedings of the 2020 International Conference on Innovation Design and Digital Technology (ICIDDT), Zhenjing, China, 5–6 December 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 343–347.
16. Luo, D.; He, Q.; Ma, M.; Yan, K.; Liu, D.; Wang, P. ECANodule: Accurate Pulmonary Nodule Detection and Segmentation with Efficient Channel Attention. In Proceedings of the 2023 International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 18–23 June 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–8.
17. Zhang, Y.; Yang, Q. A survey on multi-task learning. *IEEE Trans. Knowl. Data Eng.* **2021**, *34*, 5586–5609. [CrossRef]
18. Thung, K.H.; Wee, C.Y. A brief review on multi-task learning. *Multimed. Tools Appl.* **2018**, *77*, 29705–29725. [CrossRef]
19. Zhao, Y.; Wang, X.; Che, T.; Bao, G.; Li, S. Multi-task deep learning for medical image computing and analysis: A review. *Comput. Biol. Med.* **2023**, *153*, 106496. [CrossRef] [PubMed]
20. Chow, L.S.; Paramesran, R. Review of medical image quality assessment. *Biomed. Signal Process. Control* **2016**, *27*, 145–154. [CrossRef]
21. Kang, S.H.; Park, M.; Yoon, M.S.; Lee, Y. Quantitative evaluation of total variation noise reduction algorithm in CT images using 3D-printed customized phantom for femur diagnosis. *J. Korean Phys. Soc.* **2022**, *81*, 450–459. [CrossRef]
22. Wang, Z.; Bovik, A.C. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Process. Mag.* **2009**, *26*, 98–117. [CrossRef]
23. Lei, Y.; Niu, C.; Zhang, J.; Wang, G.; Shan, H. CT image denoising and deblurring with deep learning: Current status and perspectives. *IEEE Trans. Radiat. Plasma Med. Sci.* **2023**, *8*, 153–172. [CrossRef]
24. Setio, A.A.A.; Traverso, A.; De Bel, T.; Berens, M.S.; Van Den Bogaard, C.; Cerello, P.; Chen, H.; Dou, Q.; Fantacci, M.E.; Geurts, B.; et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Med. Image Anal.* **2017**, *42*, 1–13. [CrossRef]
25. Armato III, S.G.; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.R.; Henschke, C.I.; Hoffman, E.A.; et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med. Phys.* **2011**, *38*, 915–931. [CrossRef] [PubMed]
26. Trial Summary—Learn—NLST—The Cancer Data Access System. Available online: <https://cdas.cancer.gov/learn/nlst/trial-summary/> (accessed on 14 July 2025).
27. Van Ginneken, B.; Armato III, S.G.; de Hoop, B.; van Amelsvoort-van de Vorst, S.; Duindam, T.; Niemeijer, M.; Murphy, K.; Schilham, A.; Retico, A.; Fantacci, M.E.; et al. Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study. *Med. Image Anal.* **2010**, *14*, 707–722. [CrossRef] [PubMed]
28. ELCAP Public Lung Image Database. Available online: <http://www.via.cornell.edu/lungdb.html> (accessed on 14 July 2025).
29. Organizers., A.T.C. Tianchi Medical AI Competition Dataset. 2017. Available online: <https://tianchi.aliyun.com/competition/entrance/231601/information> (accessed on 14 July 2025).
30. Halder, A.; Dey, D.; Sadhu, A.K. Lung nodule detection from feature engineering to deep learning in thoracic CT images: A comprehensive review. *J. Digit. Imaging* **2020**, *33*, 655–677. [CrossRef] [PubMed]
31. Asuntha, A.; Srinivasan, A. Deep learning for lung Cancer detection and classification. *Multimed. Tools Appl.* **2020**, *79*, 7731–7762. [CrossRef]
32. Mridha, M.; Prodeep, A.R.; Hoque, A.M.; Islam, M.R.; Lima, A.A.; Kabir, M.M.; Hamid, M.A.; Watanobe, Y. A comprehensive survey on the progress, process, and challenges of lung cancer detection and classification. *J. Healthc. Eng.* **2022**, *2022*, 5905230. [CrossRef]
33. Chen, B.; Kitasaka, T.; Honma, H.; Takabatake, H.; Mori, M.; Natori, H.; Mori, K. Automatic segmentation of pulmonary blood vessels and nodules based on local intensity structure analysis and surface propagation in 3D chest CT images. *Int. J. Comput. Assist. Radiol. Surg.* **2012**, *7*, 465–482. [CrossRef]
34. Soltaninejad, S.; Keshani, M.; Tajeripour, F. Lung nodule detection by KNN classifier and active contour modelling and 3D visualization. In Proceedings of the 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012), Shiraz, Iran, 2–3 May 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 440–445.
35. Tun, K.M.M.; Soe, K.A. Feature extraction and classification of lung cancer nodule using image processing techniques. *Int. J. Eng. Res. Technol.* **2014**, *3*, 2204–2210.
36. Sangamithraa, P.; Govindaraju, S. Lung tumour detection and classification using EK-Mean clustering. In Proceedings of the 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 23–25 March 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 2201–2206.

37. Namin, S.T.; Moghaddam, H.A.; Jafari, R.; Esmaeil-Zadeh, M.; Gity, M. Automated detection and classification of pulmonary nodules in 3D thoracic CT images. In Proceedings of the 2010 IEEE International Conference on Systems, Man and Cybernetics, Istanbul, Turkey, 10–13 October 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 3774–3779.
38. Rossetto, A.M.; Zhou, W. Deep learning for categorization of lung cancer ct images. In Proceedings of the 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), Philadelphia, PA, USA, 17–19 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 272–273.
39. Young, I.T.; Van Vliet, L.J. Recursive implementation of the Gaussian filter. *Signal Process.* **1995**, *44*, 139–151. [\[CrossRef\]](#)
40. Neyenssac, F. Contrast enhancement using the Laplacian-of-a-Gaussian filter. *CVGIP: Graph. Model. Image Process.* **1993**, *55*, 447–463. [\[CrossRef\]](#)
41. Elaiyaraja, G.; Kumaratharan, N.; Chandra Sekhar Rao, T. Fast and efficient filter using wavelet threshold for removal of Gaussian noise from MRI/CT scanned medical images/color video sequence. *IETE J. Res.* **2022**, *68*, 10–22. [\[CrossRef\]](#)
42. Khireddine, A.; Benmahammed, K.; Puech, W. Digital image restoration by Wiener filter in 2D case. *Adv. Eng. Softw.* **2007**, *38*, 513–516. [\[CrossRef\]](#)
43. Ashwin, S.; Ramesh, J.; Kumar, S.A.; Gunavathi, K. Efficient and reliable lung nodule detection using a neural network based computer aided diagnosis system. In Proceedings of the 2012 International Conference on Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM), Chennai, India, 13–15 December 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 135–142.
44. Punithavathy, K.; Ramya, M.; Poobal, S. Analysis of statistical texture features for automatic lung cancer detection in PET/CT images. In Proceedings of the 2015 International Conference on Robotics, Automation, Control and Embedded Systems (RACE), Chennai, India, 18–20 February 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–5.
45. Mehrotra, R.; Namuduri, K.R.; Ranganathan, N. Gabor filter-based edge detection. *Pattern Recognit.* **1992**, *25*, 1479–1494. [\[CrossRef\]](#)
46. Liu, X.; Hou, F.; Qin, H.; Hao, A. Multi-view multi-scale CNNs for lung nodule type classification from CT images. *Pattern Recognit.* **2018**, *77*, 262–275. [\[CrossRef\]](#)
47. Li, Y.; Sixou, B.; Peyrin, F. A review of the deep learning methods for medical images super resolution problems. *IRBM* **2021**, *42*, 120–133. [\[CrossRef\]](#)
48. Pham, C.H.; Tor-Díez, C.; Meunier, H.; Bednarek, N.; Fablet, R.; Passat, N.; Rousseau, F. Multiscale brain MRI super-resolution using deep 3D convolutional networks. *Comput. Med. Imaging Graph.* **2019**, *77*, 101647. [\[CrossRef\]](#)
49. Song, Y.; Ermon, S. Generative modeling by estimating gradients of the data distribution. *Adv. Neural Inf. Process. Syst.* **2019**, *32*. Available online: <https://proceedings.neurips.cc/paper/2019/hash/3001ef257407d5a371a96dcd947c7d93-Abstract.html?ref=https://githubhelp.com> (accessed on 15 June 2025).
50. Chen, Y.; Shi, F.; Christodoulou, A.G.; Xie, Y.; Zhou, Z.; Li, D. Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 91–99.
51. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
52. Li, Y.; Iwamoto, Y.; Lin, L.; Xu, R.; Tong, R.; Chen, Y.-W. VolumeNet: A Lightweight Parallel Network for Super-Resolution of MR and CT Volumetric Data. *arXiv* **2020**, arXiv:2010.08357. [\[CrossRef\]](#)
53. Fang, C.; Wang, L.; Zhang, D.; Xu, J.; Yuan, Y.; Han, J. Incremental cross-view mutual distillation for self-supervised medical CT synthesis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 20677–20686.
54. Hu, X.; Mu, H.; Zhang, X.; Wang, Z.; Tan, T.; Sun, J. Meta-SR: A magnification-arbitrary network for super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 1575–1584.
55. Hu, X.; Zhang, Z.; Shan, C.; Wang, Z.; Wang, L.; Tan, T. Meta-USR: A unified super-resolution network for multiple degradation parameters. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4151–4165. [\[CrossRef\]](#)
56. Cascio, D.; Magro, R.; Fauci, F.; Iacomì, M.; Raso, G. Automatic detection of lung nodules in CT datasets based on stable 3D mass-spring models. *Comput. Biol. Med.* **2012**, *42*, 1098–1109. [\[CrossRef\]](#)
57. Valente, I.R.S.; Cortez, P.C.; Neto, E.C.; Soares, J.M.; de Albuquerque, V.H.C.; Tavares, J.M.R. Automatic 3D pulmonary nodule detection in CT images: A survey. *Comput. Methods Programs Biomed.* **2016**, *124*, 91–107. [\[CrossRef\]](#)
58. da Silva Sousa, J.R.F.; Silva, A.C.; de Paiva, A.C.; Nunes, R.A. Methodology for automatic detection of lung nodules in computerized tomography images. *Comput. Methods Programs Biomed.* **2010**, *98*, 1–14. [\[CrossRef\]](#)
59. Gupta, A.; Saar, T.; Martens, O.; Moullec, Y.L. Automatic detection of multisize pulmonary nodules in CT images: Large-scale validation of the false-positive reduction step. *Med. Phys.* **2018**, *45*, 1135–1149. [\[CrossRef\]](#)

60. Saien, S.; Pilevar, A.H.; Moghaddam, H.A. Refinement of lung nodule candidates based on local geometric shape analysis and Laplacian of Gaussian kernels. *Comput. Biol. Med.* **2014**, *54*, 188–198. [[CrossRef](#)]
61. Hu, S.; Hoffman, E.A.; Reinhardt, J.M. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. *IEEE Trans. Med. Imaging* **2001**, *20*, 490–498. [[CrossRef](#)]
62. Silveira, M.; Nascimento, J.; Marques, J. Automatic segmentation of the lungs using robust level sets. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 4414–4417.
63. Shi, Y.; Qi, F.; Xue, Z.; Chen, L.; Ito, K.; Matsuo, H.; Shen, D. Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics. *IEEE Trans. Med. Imaging* **2008**, *27*, 481–494. [[PubMed](#)]
64. Annangi, P.; Thiruvankadam, S.; Raja, A.; Xu, H.; Sun, X.; Mao, L. A region based active contour method for x-ray lung segmentation using prior shape and low level features. In Proceedings of the 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Rotterdam, The Netherlands, 14–17 April 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 892–895.
65. Campadelli, P.; Casiraghi, E.; Artioli, D. A fully automated method for lung nodule detection from postero-anterior chest radiographs. *IEEE Trans. Med. Imaging* **2006**, *25*, 1588–1603. [[CrossRef](#)]
66. Hofmanninger, J.; Prayer, F.; Pan, J.; Röhrich, S.; Prosch, H.; Langs, G. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. *Eur. Radiol. Exp.* **2020**, *4*, 1–13. [[CrossRef](#)] [[PubMed](#)]
67. Tan, J.; Jing, L.; Huo, Y.; Li, L.; Akin, O.; Tian, Y. LGAN: Lung segmentation in CT scans using generative adversarial network. *Comput. Med. Imaging Graph.* **2021**, *87*, 101817. [[CrossRef](#)] [[PubMed](#)]
68. Hu, Q.; Souza, L.F.d.F.; Holanda, G.B.; Alves, S.S.; Silva, F.H.d.S.; Han, T.; Reboucas Filho, P.P. An effective approach for CT lung segmentation using mask region-based convolutional neural networks. *Artif. Intell. Med.* **2020**, *103*, 101792. [[CrossRef](#)]
69. Henschke, C.I. Early lung cancer action project: Overall design and findings from baseline screening. *Cancer* **2000**, *89*, 2474–2482. [[CrossRef](#)]
70. Camarlinghi, N. Automatic detection of lung nodules in computed tomography images: Training and validation of algorithms using public research databases. *Eur. Phys. J. Plus* **2013**, *128*, 110. [[CrossRef](#)]
71. Armato, S.G.; Giger, M.L.; Moran, C.J.; Blackburn, J.T.; Doi, K.; MacMahon, H. Computerized detection of pulmonary nodules on CT scans. *Radiographics* **1999**, *19*, 1303–1311. [[CrossRef](#)]
72. Aresta, G.; Cunha, A.; Campilho, A. Detection of juxta-pleural lung nodules in computed tomography images. In Proceedings of the Medical Imaging 2017: Computer-Aided Diagnosis, Orlando, FL, USA, 11–16 February 2017; SPIE: Bellingham, WA, USA, 2017; Volume 10134, pp. 952–958.
73. Zawish, M.; Siyal, A.A.; Shahani, S.H.; Junejo, A.Z.; Khalil, A. Brain tumor segmentation through region-based, supervised and unsupervised learning methods: A literature survey. *J. Biomed. Eng. Med. Imaging* **2019**, *6*, 8–26.
74. Suárez-Cuenca, J.J.; Guo, W.; Li, Q. Automated detection of pulmonary nodules in CT: False positive reduction by combining multiple classifiers. In Proceedings of the Medical Imaging 2011: Computer-Aided Diagnosis, Lake Buena Vista, FL, USA, 15–17 February 2011; SPIE: Bellingham, WA, USA, 2011; Volume 7963, pp. 927–932.
75. Suiyuan, W.; Junfeng, W. Pulmonary nodules 3D detection on serial CT scans. In Proceedings of the 2012 Third Global Congress on Intelligent Systems, Kota Kinabalu, Malaysia, 6–8 November 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 257–260.
76. Parveen, S.S.; Kavitha, C. Detection of lung cancer nodules using automatic region growing method. In Proceedings of the 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 4–6 July 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 1–6.
77. Zhang, Y.F.; Mao, J.L.; Xiong, Z.Y. An efficient clustering algorithm. In Proceedings of the 2003 International Conference on Machine Learning and Cybernetics, Xi'an, China, 2–5 November 2003; (IEEE Cat. No. 03EX693); IEEE: Piscataway, NJ, USA, 2003; Volume 1, pp. 261–265.
78. Javaid, M.; Javid, M.; Rehman, M.Z.U.; Shah, S.I.A. A novel approach to CAD system for the detection of lung nodules in CT images. *Comput. Methods Programs Biomed.* **2016**, *135*, 125–139. [[CrossRef](#)]
79. Cao, W.; Wu, R.; Cao, G.; He, Z. A comprehensive review of computer-aided diagnosis of pulmonary nodules based on computed tomography scans. *IEEE Access* **2020**, *8*, 154007–154023. [[CrossRef](#)]
80. Naqi, S.M.; Sharif, M.; Lali, I.U. A 3D nodule candidate detection method supported by hybrid features to reduce false positives in lung nodule detection. *Multimed. Tools Appl.* **2019**, *78*, 26287–26311. [[CrossRef](#)]
81. Fotin, S.V.; Yankelevitz, D.F.; Henschke, C.I.; Reeves, A.P. A multiscale Laplacian of Gaussian (LoG) filtering approach to pulmonary nodule detection from whole-lung CT scans. *arXiv* **2019**, arXiv:1907.08328.
82. Hua, K.L.; Hsu, C.H.; Hidayati, S.C.; Cheng, W.H.; Chen, Y.J. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Oncotargets Ther.* **2015**, *8*, 2015–2022.
83. Kumar, D.; Wong, A.; Clausi, D.A. Lung nodule classification using deep features in CT images. In Proceedings of the 2015 12th Conference on Computer and Robot Vision, Halifax, NS, Canada, 3–5 June 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 133–138.

84. Ciompi, F.; de Hoop, B.; van Riel, S.J.; Chung, K.; Scholten, E.T.; Oudkerk, M.; de Jong, P.A.; Prokop, M.; van Ginneken, B. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. *Med. Image Anal.* **2015**, *26*, 195–202. [[CrossRef](#)] [[PubMed](#)]
85. Ramachandran, S.; George, J.; Skaria, S.; Varun, V.V. Using YOLO based deep learning network for real time detection and localization of lung nodules from low dose CT scans. In Proceedings of the Medical Imaging 2018: Computer-Aided Diagnosis, Houston, TX, USA, 12–15 February 2018; SPIE: Bellingham, WA, USA, 2018; Volume 10575, pp. 347–355.
86. Trajanovski, S.; Mavroeidis, D.; Swisher, C.L.; Gebre, B.G.; Veeling, B.S.; Wiemker, R.; Klinder, T.; Tahmasebi, A.; Regis, S.M.; Wald, C.; et al. Towards radiologist-level cancer risk assessment in CT lung screening using deep learning. *Comput. Med. Imaging Graph.* **2021**, *90*, 101883. [[CrossRef](#)] [[PubMed](#)]
87. Xie, H.; Yang, D.; Sun, N.; Chen, Z.; Zhang, Y. Automated pulmonary nodule detection in CT images using deep convolutional neural networks. *Pattern Recognit.* **2019**, *85*, 109–119. [[CrossRef](#)]
88. Dou, Q.; Chen, H.; Yu, L.; Qin, J.; Heng, P.A. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Trans. Biomed. Eng.* **2016**, *64*, 1558–1567. [[CrossRef](#)]
89. Anirudh, R.; Thiagarajan, J.J.; Bremer, T.; Kim, H. Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data. In Proceedings of the Medical Imaging 2016: Computer-Aided Diagnosis, San Diego, CA, USA, 14–16 September 2016; SPIE: Bellingham, WA, USA, 2016; Volume 9785, pp. 791–796.
90. Huang, W.; Xue, Y.; Wu, Y. A CAD system for pulmonary nodule prediction based on deep three-dimensional convolutional neural networks and ensemble learning. *PLoS ONE* **2019**, *14*, e0219369. [[CrossRef](#)]
91. Ali, I.; Hart, G.R.; Gunabushanam, G.; Liang, Y.; Muhammad, W.; Nartowt, B.; Kane, M.; Ma, X.; Deng, J. Lung nodule detection via deep reinforcement learning. *Front. Oncol.* **2018**, *8*, 108. [[CrossRef](#)] [[PubMed](#)]
92. Paul, R.; Hawkins, S.H.; Schabath, M.B.; Gillies, R.J.; Hall, L.O.; Goldgof, D.B. Predicting malignant nodules by fusing deep features with classical radiomics features. *J. Med. Imaging* **2018**, *5*, 011021. [[CrossRef](#)] [[PubMed](#)]
93. Kamble, B.; Sahu, S.P.; Doriya, R. A review on lung and nodule segmentation techniques. In *Advances in Data and Information Sciences: Proceedings of ICDIS 2019*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 555–565.
94. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [[CrossRef](#)] [[PubMed](#)]
95. Jardim, S.; António, J.; Mora, C. Image thresholding approaches for medical image segmentation—short literature review. *Procedia Comput. Sci.* **2023**, *219*, 1485–1492. [[CrossRef](#)]
96. Zhao, B.; Yankelevitz, D.; Reeves, A.; Henschke, C. Two-dimensional multi-criterion segmentation of pulmonary nodules on helical CT images. *Med. Phys.* **1999**, *26*, 889–895. [[CrossRef](#)]
97. Zhao, B.; Reeves, A.P.; Yankelevitz, D.; Henschke, C.I. Three-dimensional multi-criterion automatic segmentation of pulmonary nodules of helical computed tomography images. *Opt. Eng.* **1999**, *38*, 1340–1347.
98. Yankelevitz, D.F.; Gupta, R.; Zhao, B.; Henschke, C.I. Small pulmonary nodules: Evaluation with repeat CT—Preliminary experience. *Radiology* **1999**, *212*, 561–566. [[CrossRef](#)]
99. Yankelevitz, D.F.; Reeves, A.P.; Kostis, W.J.; Zhao, B.; Henschke, C.I. Small pulmonary nodules: Volumetrically determined growth rates based on CT evaluation. *Radiology* **2000**, *217*, 251–256. [[CrossRef](#)]
100. Wiemker, R.; Zwartkuis, A. Optimal thresholding for 3D segmentation of pulmonary nodules in high resolution CT. In *Proceedings of the International Congress Series*; Elsevier: Amsterdam, The Netherlands, 2001; Volume 1230, pp. 653–658.
101. Kubota, T.; Jerebko, A.K.; Dewan, M.; Salganicoff, M.; Krishnan, A. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. *Med. Image Anal.* **2011**, *15*, 133–154. [[CrossRef](#)]
102. Halder, A.; Chatterjee, S.; Dey, D.; Kole, S.; Munshi, S. An adaptive morphology based segmentation technique for lung nodule detection in thoracic CT image. *Comput. Methods Programs Biomed.* **2020**, *197*, 105720. [[CrossRef](#)] [[PubMed](#)]
103. Dehmeshki, J.; Amin, H.; Valdivieso, M.; Ye, X. Segmentation of pulmonary nodules in thoracic CT scans: A region growing approach. *IEEE Trans. Med. Imaging* **2008**, *27*, 467–480. [[CrossRef](#)] [[PubMed](#)]
104. Soltani-Nabipour, J.; Khorshidi, A.; Noorian, B. Lung tumor segmentation using improved region growing algorithm. *Nucl. Eng. Technol.* **2020**, *52*, 2313–2319. [[CrossRef](#)]
105. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; proceedings, part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
106. Tong, G.; Li, Y.; Chen, H.; Zhang, Q.; Jiang, H. Improved U-NET network for pulmonary nodules segmentation. *Optik* **2018**, *174*, 460–469. [[CrossRef](#)]
107. Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 565–571.

108. Ma, X.; Song, H.; Jia, X.; Wang, Z. An improved V-Net lung nodule segmentation model based on pixel threshold separation and attention mechanism. *Sci. Rep.* **2024**, *14*, 4743. [\[CrossRef\]](#)
109. Xu, W.; Fu, Y.L.; Zhu, D. ResNet and its application to medical image processing: Research progress and challenges. *Comput. Methods Programs Biomed.* **2023**, *240*, 107660. [\[CrossRef\]](#)
110. Cao, H.; Liu, H.; Song, E.; Hung, C.C.; Ma, G.; Xu, X.; Jin, R.; Lu, J. Dual-branch residual network for lung nodule segmentation. *Appl. Soft Comput.* **2020**, *86*, 105934. [\[CrossRef\]](#)
111. Wang, S.; Zhou, M.; Gevaert, O.; Tang, Z.; Dong, D.; Liu, Z.; Jie, T. A multi-view deep convolutional neural networks for lung nodule segmentation. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju Island, Republic of Korea, 11–15 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1752–1755.
112. Liu, C.; Liu, H.; Zhang, X.; Guo, J.; Lv, P. Multi-scale and multi-view network for lung tumor segmentation. *Comput. Biol. Med.* **2024**, *172*, 108250. [\[CrossRef\]](#)
113. Alhajim, D.; Ansari-Asl, K.; Akbarizadeh, G.; Soorki, M.N. Improved lung nodule segmentation with a squeeze excitation dilated attention based residual UNet. *Sci. Rep.* **2025**, *15*, 3770. [\[CrossRef\]](#)
114. Bruntha, P.M.; Pandian, S.I.A.; Sagayam, K.M.; Bandyopadhyay, S.; Pomplun, M.; Dang, H. Lung_PAYNet: A pyramidal attention based deep learning network for lung nodule segmentation. *Sci. Rep.* **2022**, *12*, 20330. [\[CrossRef\]](#)
115. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [\[CrossRef\]](#)
116. Tyagi, S.; Talbar, S.N. CSE-GAN: A 3D conditional generative adversarial network with concurrent squeeze-and-excitation blocks for lung nodule segmentation. *Comput. Biol. Med.* **2022**, *147*, 105781. [\[CrossRef\]](#) [\[PubMed\]](#)
117. Wang, S.; Jiang, A.; Li, X.; Qiu, Y.; Li, M.; Li, F. DPBET: A dual-path lung nodules segmentation model based on boundary enhancement and hybrid transformer. *Comput. Biol. Med.* **2022**, *151*, 106330. [\[CrossRef\]](#) [\[PubMed\]](#)
118. Usman, M.; Shin, Y.G. Deha-net: A dual-encoder-based hard attention network with an adaptive roi mechanism for lung nodule segmentation. *Sensors* **2023**, *23*, 1989. [\[CrossRef\]](#) [\[PubMed\]](#)
119. Suganyadevi, S.; Seethalakshmi, V.; Balasamy, K. A review on deep learning in medical image analysis. *Int. J. Multimed. Inf. Retr.* **2022**, *11*, 19–38. [\[CrossRef\]](#)
120. Nandakumar, N.; D'Souza, N.S.; Manzoor, K.; Pillai, J.J.; Gujar, S.K.; Sair, H.I.; Venkataraman, A. A multi-task deep learning framework to localize the eloquent cortex in brain tumor patients using dynamic functional connectivity. In *Proceedings of the Machine Learning in Clinical Neuroimaging and Radiogenomics in Neuro-Oncology: Third International Workshop, MLCN 2020, and Second International Workshop, RNO-AI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, 4–8 October 2020*; Proceedings 3; Springer: Berlin/Heidelberg, Germany, 2020; pp. 34–44.
121. Zhou, T.; Fu, H.; Chen, G.; Shen, J.; Shao, L. Hi-net: Hybrid-fusion network for multi-modal MR image synthesis. *IEEE Trans. Med. Imaging* **2020**, *39*, 2772–2781. [\[CrossRef\]](#)
122. Park, S.; Kim, G.; Oh, Y.; Seo, J.B.; Lee, S.M.; Kim, J.H.; Moon, S.; Lim, J.K.; Ye, J.C. Multi-task vision transformer using low-level chest X-ray feature corpus for COVID-19 diagnosis and severity quantification. *Med. Image Anal.* **2022**, *75*, 102299. [\[CrossRef\]](#)
123. Zhang, L.; Zhong, L.; Li, C.; Zhang, W.; Hu, C.; Dong, D.; Liu, Z.; Zhou, J.; Tian, J. Knowledge-guided multi-task attention network for survival risk prediction using multi-center computed tomography images. *Neural Netw.* **2022**, *152*, 394–406. [\[CrossRef\]](#)
124. Jafari, M.H.; Girgis, H.; Van Woudenberg, N.; Liao, Z.; Rohling, R.; Gin, K.; Abolmaesumi, P.; Tsang, T. Automatic biplane left ventricular ejection fraction estimation with mobile point-of-care ultrasound using multi-task learning and adversarial training. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 1027–1037. [\[CrossRef\]](#) [\[PubMed\]](#)
125. Yu, C.; Gao, Z.; Zhang, W.; Yang, G.; Zhao, S.; Zhang, H.; Zhang, Y.; Li, S. Multitask learning for estimating multitype cardiac indices in MRI and CT based on adversarial reverse mapping. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 493–506. [\[CrossRef\]](#) [\[PubMed\]](#)
126. Keshwani, D.; Kitamura, Y.; Ihara, S.; Iizuka, S.; Simo-Serra, E. TopNet: Topology preserving metric learning for vessel tree reconstruction and labelling. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, 4–8 October 2020*; Proceedings, Part VI 23; Springer: Berlin/Heidelberg, Germany, 2020; pp. 14–23.
127. Du, B.; Liao, J.; Turkbey, B.; Yan, P. Multi-task learning for registering images with large deformation. *IEEE J. Biomed. Health Informatics* **2020**, *25*, 1624–1633. [\[CrossRef\]](#)
128. Zhang, Y.; Yang, Q. An overview of multi-task learning. *Natl. Sci. Rev.* **2018**, *5*, 30–43. [\[CrossRef\]](#)
129. Wu, T.H.; Lian, C.; Lee, S.; Pastewait, M.; Piers, C.; Liu, J.; Wang, F.; Wang, L.; Chiu, C.Y.; Wang, W.; et al. Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3D intraoral scans. *IEEE Trans. Med. Imaging* **2022**, *41*, 3158–3166. [\[CrossRef\]](#)
130. Malhotra, A.; Mittal, S.; Majumdar, P.; Chhabra, S.; Thakral, K.; Vatsa, M.; Singh, R.; Chaudhury, S.; Pudrod, A.; Agrawal, A. Multi-task driven explainable diagnosis of COVID-19 using chest X-ray images. *Pattern Recognit.* **2022**, *122*, 108243. [\[CrossRef\]](#)

131. Xu, X.; Lian, C.; Wang, S.; Zhu, T.; Chen, R.C.; Wang, A.Z.; Royce, T.J.; Yap, P.T.; Shen, D.; Lian, J. Asymmetric multi-task attention network for prostate bed segmentation in computed tomography images. *Med. Image Anal.* **2021**, *72*, 102116. [[CrossRef](#)]
132. Wang, X.; Jiang, L.; Li, L.; Xu, M.; Deng, X.; Dai, L.; Xu, X.; Li, T.; Guo, Y.; Wang, Z.; et al. Joint learning of 3D lesion segmentation and classification for explainable COVID-19 diagnosis. *IEEE Trans. Med. Imaging* **2021**, *40*, 2463–2476. [[CrossRef](#)]
133. Nguyen, T.C.; Nguyen, T.P.; Cao, T.; Dao, T.T.P.; Ho, T.N.; Nguyen, T.V.; Tran, M.T. MANet: Multi-branch attention auxiliary learning for lung nodule detection and segmentation. *Comput. Methods Programs Biomed.* **2023**, *241*, 107748. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



DSPACE

<https://dspace.org/>

Multi-task deep learning for lung nodule detection and segmentation in CT scans: a review

Li, Runhan; Honarvar Shakibaei Asli, Barmak

2025-07-28

MDPI

<https://doi.org/10.3390/electronics14153009>

Li R, Honarvar Shakibaei Asli B. (2025) Multi-task deep learning for lung nodule detection and segmentation in CT scans: a review. Electronics, Volume 14, Issue 15, July 2025, Article number 3009