



ARTICLE

Multilevel Attention Unet Segmentation Algorithm for Lung Cancer Based on CT Images

Huan Wang¹, Shi Qiu^{1,2,*}, Benyue Zhang¹ and Lixuan Xiao³

¹Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China

²School of Biomedical Engineering, Fourth Military Medical University, Xi'an, China

³University of Illinois at Urbana Champaign, Champaign, USA

*Corresponding Author: Shi Qiu. Email: qiushi@opt.ac.cn

Received: 16 October 2023 Accepted: 07 December 2023 Published: 27 February 2024

ABSTRACT

Lung cancer is a malady of the lungs that gravely jeopardizes human health. Therefore, early detection and treatment are paramount for the preservation of human life. Lung computed tomography (CT) image sequences can explicitly delineate the pathological condition of the lungs. To meet the imperative for accurate diagnosis by physicians, expeditious segmentation of the region harboring lung cancer is of utmost significance. We utilize computer-aided methods to emulate the diagnostic process in which physicians concentrate on lung cancer in a sequential manner, erect an interpretable model, and attain segmentation of lung cancer. The specific advancements can be encapsulated as follows: 1) Concentration on the lung parenchyma region: Based on 16-bit CT image capturing and the luminance characteristics of lung cancer, we proffer an intercept histogram algorithm. 2) Focus on the specific locus of lung malignancy: Utilizing the spatial interrelation of lung cancer, we propose a memory-based Unet architecture and incorporate skip connections. 3) Data Imbalance: In accordance with the prevalent situation of an overabundance of negative samples and a paucity of positive samples, we scrutinize the existing loss function and suggest a mixed loss function. Experimental results with pre-existing publicly available datasets and assembled datasets demonstrate that the segmentation efficacy, measured as Area Overlap Measure (AOM) is superior to 0.81, which markedly ameliorates in comparison with conventional algorithms, thereby facilitating physicians in diagnosis.

KEYWORDS

Lung cancer; computed tomography; computer-aided diagnosis; Unet; segmentation

1 Introduction

Lung cancer is a neoplastic tumor that originates in the pulmonary system, with a precipitous escalation in morbidity and mortality, and it stands as one of the most pernicious malignancies affecting both the health and longevity of the general populace [1]. Numerous nations have reported a marked augmentation in the morbidity and mortality rates associated with lung cancer over the preceding five decades, with males predominating in both morbidity and mortality among all cancerous neoplasms, and females occupying the second position [2]. The etiology of lung cancer remains



enigmatic. The clinical manifestations of lung cancer are intricate, and the manifestation or absence, acuteness and temporal onset of symptoms are contingent upon variables such as the anatomical locus of the tumor, histological subtype, concomitant presence of metastasis and complications, as well as interindividual variability in physiological responsiveness and tolerance. The initial symptoms are frequently insidious, verging on unobtrusive. Symptoms associated with central lung tumors are pronounced and severe, whereas those stemming from peripheral lung cancer manifest belatedly and are often inconspicuous, or remain asymptomatic and are commonly discovered upon routine medical examination. The symptoms can be broadly categorized into: local symptoms, systemic symptoms, extrapulmonary symptoms, infiltrative, and metastatic symptoms [3].

1.1 Clinical Manifestations of Lung Cancer

The principal CT imaging phenotypes of lung cancer include: lobulation, spiculation, vacuole sign, vascular convergence sign, pleural indentation, and cavitory sign, among others [4]. 1) Lobulation: The margins of lesion sites are irregular and lobulated, attributable to heterogeneous differentiation of neoplastic cells and disparate growth kinetics across loci. Approximately 70%–80% of pulmonary nodules exhibiting lobulation are malignant lesions. 2) Spiculation: Lesions exhibit variances in the degree of fine, short protrusions at the margins, with spine-like or serrated morphological alterations, typically encountered at the juncture between pathologic and normal pulmonary tissue, with a prevalence rate of approximately 80%–85%, indicative of a common sign of lung malignancy. 3) Vacuole sign: Lesions with vacuolation may represent uninvolved pulmonary parenchyma or fine bronchial structures, or could be indicative of necrotic transformation post-excision. 4) Vascular convergence sign: Refers to the confluence of surrounding diminutive vascular structures toward the lesion, where the vessels either terminate at the lesion's periphery or penetrate the lesion itself; this manifestation is related to fibrous reactions within the neoplasm and hypertrophy of the supplying vessels. 5) Pleural indentation: Lesions situated peripherally within the pulmonary architecture result in pleural invagination oriented toward the lesion, commonly observed in pulmonary adenocarcinomas, and attributed to tumoral fibrotic contraction. 6) Cavitory sign: A spherical or ovoid air-like hypodense shadow exceeding 5 mm, having a likelihood of less than 10% in lung cancer radiographic assessments. The cavity wall is variably thickened, the internal wall is irregular, and the cavity is either central or eccentric in location. Lung cancer represents a heterogeneous pulmonary lesion depicted on imaging as a conglomeration of diverse signs, and the magnitude of irregularities in these signs directly correlates with the lesion's degree of malignancy.

In summary, lung cancer exerts a deleterious impact on human health and longevity; therefore, the precocious diagnosis and therapeutic intervention are of paramount import. Computed Tomography (CT) imaging serves as a visual conduit for pulmonary assessment. Utilizing such CT imagery, clinicians can glean vital information pertaining to pulmonary structure and function, thereby facilitating diagnosis. However, the physician's diagnostic acumen, coupled with unavoidable fatigue, can exert a direct influence on the resultant diagnostic outcome. Coinciding with the ascendancy of computational technologies, the implementation of artificial intelligence presents a viable avenue for augmenting diagnostic precision.

1.2 Artificial Intelligence-Based Algorithms for Lung Diagnosis

Computer-aided diagnosis (CAD) encompasses the amalgamation of imaging modalities, medical image processing techniques, and other putative physiological and biochemical markers, orchestrated with computational analytics, to facilitate lesion identification and enhance diagnostic veracity. CAD primarily operates within the realm of medical imaging and has acquired the epithet of the physician's

“third eye”. The ubiquitous deployment of CAD systems has been instrumental in amplifying both the sensitivity and specificity of clinicians’ diagnostic assessments [5,6]. Current research in the field of computer-assisted segmentation of lung neoplasms can be broadly bifurcated into traditional feature analysis and deep feature analysis.

In the domain of traditional feature analysis, Tripathi et al. [7] scrutinized the imaging characteristics inherent to pulmonary images and advanced a threshold-based segmentation algorithm. Mazzone et al. [8] employed mathematical statistics to conduct a comprehensive analysis of lung cancer pathology; Mathews et al. [9] conducted a comparative evaluation of the efficacy of disparate segmentation algorithms. Cai et al. [10] posited features derived from a clinical vantage point to achieve accurate pulmonary neoplasm segmentation. Alam et al. [11] devised a Support Vector Machine (SVM) algorithm to facilitate lung carcinoma extraction. Wanet et al. [12] used threshold algorithm to segmentation. Zhang et al. [13] generated a three-dimensional model to scrutinize lesion-specific attributes. Moon et al. [14] explored the low-dose CT characteristics of lung cancer and formulated a texture analysis framework for categorical classification. Avinash et al. [15] engineered watershed segmentation algorithm to achieve segmentation. Nazir et al. [16] developed a rapid segmentation model predicated on image feature analysis. Garau et al. [17] synthesized a model derived from both image features and clinical indicators. Lastly, Zhang et al. [18] employed topological structures and globally-enhanced graph reasoning paradigms to segment neoplastic regions.

In the realm of depth-based feature analysis, Yu et al. [19] architected a comprehensive deep learning framework for computer-aided detection to actualize pulmonary neoplasm identification. Shakeel et al. [20] employed cellular neural networks for neoplastic region delineation. Liu et al. [21] used a boundary attention model based on deep learning to extract lung parenchyma. Shamas et al. [22] conducted an exhaustive review of the advancements in deep learning techniques as applied to lung cancer segmentation and affirmed the efficacy of such algorithms. Yang et al. [23] achieved lung nodule segmentation by orchestrating feature-aware attention to modulate uncertainty and reported promising outcomes. Alakwaa et al. [24] formulated a 3D Convolutional Neural Network (CNN) to scrutinize the holistic features of lung malignancies. Skourt et al. [25] implemented a Deep Neural Network (DNN) to extract salient image attributes. Shaziya et al. [26] designed a Unet convolutional architecture to automate the segmentation of pulmonary neoplasms. Shen et al. [27] constructed a deep learning paradigm to prognosticate the development trajectory of lung cancer. Zhou et al. [28] used global-local attention YOLOV5 for lung tumor detection. Zhang et al. [29] harnessed the Resnet architecture to extract diagnostic features. Zhang et al. [30] combined atlas-based approaches with CNN to realize sophisticated image segmentation. Hu et al. [31] developed mask region-based CNNs for targeted neoplasm segmentation. Asuntha et al. [32] engineered a comprehensive deep network that unified both the detection and classification analyses. Ni et al. [33] devised a segmentation algorithm contingent upon the sequential correlations. Hou et al. [34] constructed an enhanced model to achieve target segmentation. Sousa et al. [35] formulated a Residual Unet to extract diagnostic markers for effective lung cancer segmentation. Shao et al. [36] explored the interrelation between semantic meanings and image data to achieve semantic segmentation. Alameen et al. [37] established a convolutional neural network tailored for pulmonary neoplasm segmentation.

In summary, concomitant with the maturation of artificial intelligence paradigms, computational systems are increasingly proficient in augmenting clinicians in the diagnosis of lung carcinoma. Traditional feature-based algorithms hinge upon visual and quantifiable parameters such as textural attributes and gray scale values, yet these metrics possess constrained scalability and fail to efficaciously augment performance metrics. Deep learning approaches, however, construct an epistemological conduit between image data and semantic interpretations, predicated on neural transduction

mechanisms, and have shown noteworthy outcomes. Nonetheless, the dearth of theoretical substantiation for these algorithms raises questions about their generalizability, and they often yield false positives outside the lung regions. Moreover, these algorithms demonstrate attenuated interpretability, thus impeding their seamless integration into clinical workflows.

By scrutinizing the diagnostic methodologies employed by clinicians, we initially formulated an algorithm designed to isolate regions of diagnostic interest within the pulmonary landscape. Subsequently, we refined the Unet architecture to capture the unique feature sets pertinent to lung neoplasms, harnessed the three-dimensional rendering capabilities intrinsic to CT imaging, integrated attention mechanisms within the computational model, and optimized the loss function in the context of an imbalanced dataset comprising limited positive samples and an abundance of negative samples. This culminated in the precise extraction of pulmonary malignancies.

1.3 Unet

Unet exhibits superior training efficacy and demonstrates enhanced outcomes in the domain of medical image segmentation [38]. This architecture metamorphoses the rudimentary linear network topology into a symmetrical U-shaped configuration predicated on the Fully Convolutional Network (FCN) encoding-decoding framework, necessitating only a modest dataset for the realization of precise image segmentation, as illustrated in Fig. 1. The incorporation of skip-connection topologies within the network facilitates the amalgamation of both low-level and high-level feature vectors. Unet encompasses encoding modules, decoding modules, and skip-connection mechanisms.

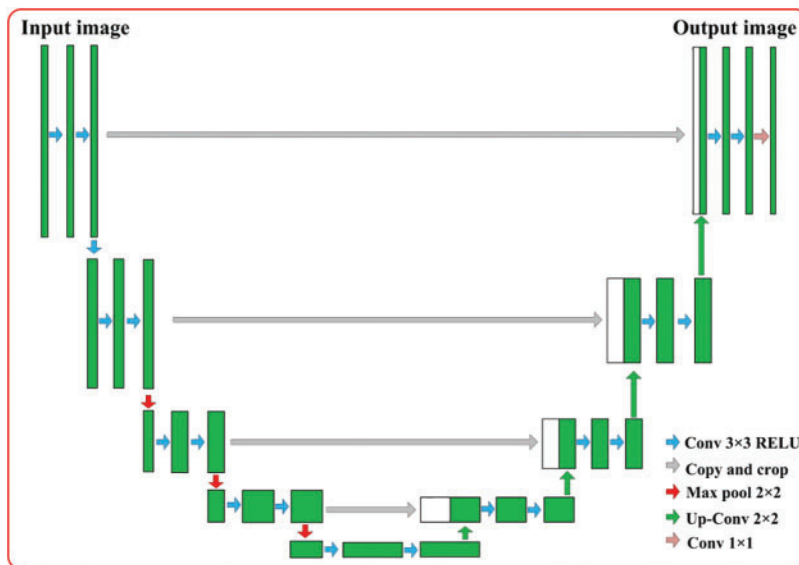


Figure 1: Traditional Unet network

The Unet architecture employs an encoding module to abstract image characteristics, composed of convolutional layers, activation functions, and maximal pooling operations. The decoding module of Unet is isomorphic to its encoding counterpart and includes convolution and upsampling procedures. Upsampling is executed via convolutional transpose operations, progressively restoring the image resolution through 2×2 “deconvolution”. The terminal yield is a probability heat map delineating the segmentation outcomes.

Unet augments the FCN model by appending skip-connection dyads between the encoding and decoding modules, thereby integrating the feature maps generated by each stratification of the encoding module into the upsampling phase of the corresponding decoding layer. This facilitates the confluence of low-level and high-level feature vectors. During this integration, both feature maps intended for fusion maintain dimensional congruence due to convolutional operations in the encoding phase leading to the diminution of boundary pixels. Unet opts for preliminary cropping of the generated feature maps, followed by their subsequent concatenation. Given that cropping is executed prior to each fusion event, the dimensions of the terminal output image are markedly reduced relative to the original, albeit the segmentation precision is substantially elevated due to the incorporation of skip-connection strategies.

1.4 Structure of the Paper

Predicated on an examination of the physician's diagnostic methodology, this manuscript introduces the Lung cancer segmentation algorithm based on explainable attention Unet, employing advanced machine learning techniques from an explainable analytic viewpoint. The structure of this paper is delineated as follows: [Section 2](#) delves into the nuances of the algorithmic architecture. [Subsection 2.1](#) introduces the database used in the experiment. [Subsection 2.2](#) focuses on the construction of the explainable model. [Subsection 2.3](#) discusses the target region extraction algorithm premised on intercept histogram methodology. [Subsection 2.4](#) elaborates on the Unet framework predicated on attention mechanisms, which serves to achieve precise lung cancer segmentation. [Subsection 2.5](#) introduces a multi-dimensional fusion loss function to amplify segmentation efficacy. [Section 3](#) presents the experimental results and subsequent analysis, serving to authenticate the algorithm's efficacy. Finally, [Section 4](#) encapsulates the novel contributions presented in this manuscript and proffers prospective directions for ensuing scholarly inquiry.

2 Materials and Methods

2.1 Materials

Our datasets originate from an aggregate of 31 collections of thoracic CT scans sourced from Therapy Response (RIDER), the Lung Image Database Consortium (LIDC), patient archives from Stanford University Medical Center and the Moffitt Cancer Centre, and the Columbia University/FDA Phantom, as well as 120 collections of lung CT data procured by the First Affiliated Hospital of Xi'an Jiaotong University via utilization of four Philips imaging apparatuses. The regions corresponding to pulmonary carcinomas are annotated in the axial orientation employing a double-blind methodology to establish a benchmark standard. The spatial resolution of the labeled dataset is 512×512 , and the slice thickness ranges between 0.5 and 1.5 mm, cumulatively yielding 76,512 frames of Digital Imaging and Communications in Medicine (DICOM) data. The particulars are delineated as follows: The prevalent images are encoded as 8-bit images with an intensity value domain spanning $[0,255]$, amenable for direct visual rendering on a display apparatus. In contrast, the 16-bit image encompasses a more expansive information spectrum, with an intensity value domain ranging between $[0,65535]$. Such images are incompatible with conventional visual display modalities and necessitate the calibration of window for appropriate mapping, as illustrated in [Fig. 2](#).

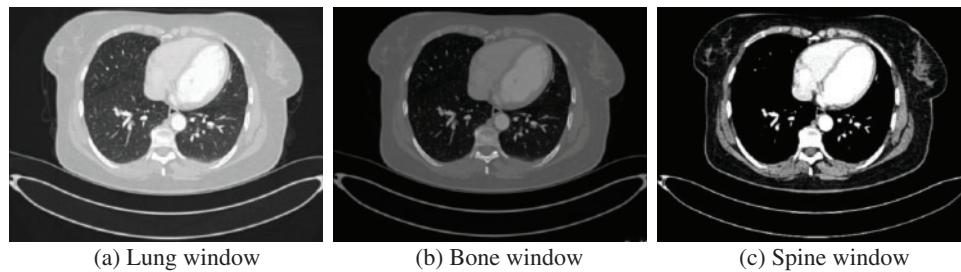


Figure 2: The mapping images by different window

2.2 Explainable Model Construction

CT images of the pulmonary system furnish a graphical elucidation of the cross-sectional morphology fundamental to a clinician's diagnostic reasoning. We have scrutinized the diagnostic paradigms employed by medical professionals, which are markedly process-oriented and interpretable: initially, they concentrate empirically on the loci where pathological alterations are plausible, subsequently executing a meticulous frame-by-frame examination of the suspect lesion. This cognitive approach is explicable via Gestalt psychology. Gestalt posits that human beings glean in excess of 80% of their sensory information through optical channels, and the capability for worldly perception emanates from the synergistic activity between the ocular system and the cerebral cortex [39]. Gestalt is governed by four cardinal principles of organization: figure-ground, proximity, similarity and continuity [40].

The mechanism underpinning the physicians' identification of focal points can be elucidated by employing the aforementioned principles. 1) Figure-ground: when confronted with CT image datasets, the human cognition accentuates regions of clinical pertinence, thereby rendering them into the foreground, whilst relegating irrelevant zones to the background. 2) Proximity: spatially adjacent regions are classified into a cohesive group. 3) Similarity: elements with the same characteristics are more closely related, such as shape, size, orientation, grey value, etc. 4) Continuity: CT imaging constitutes a scanned modality in which the brain intrinsically synthesizes these discrete pixel points to yield an authentic depiction of the pulmonary architecture.

Through the preceding analytical discourse, we have formulated a computer-aided diagnostic schema as depicted in Fig. 3. Predicated upon the principle of Continuity, the fragmented pixel constituents of CT scans are treated as a contiguous visual landscape, mirroring the pulmonary topography. Following the Figure-ground principle, a threshold segmentation technique is devised to focalize on regions manifesting suspect pulmonary neoplasia. Local attention algorithms are architected to excavate spatial tri-dimensional relationships pursuant to the principle of Proximity. Due to the similarity between the training dataset and the testing dataset, as well as the morphological similarities in lung cancer, we have improved the Unet network based on the principle of similarity, thus enabling the precise demarcation of focal points.

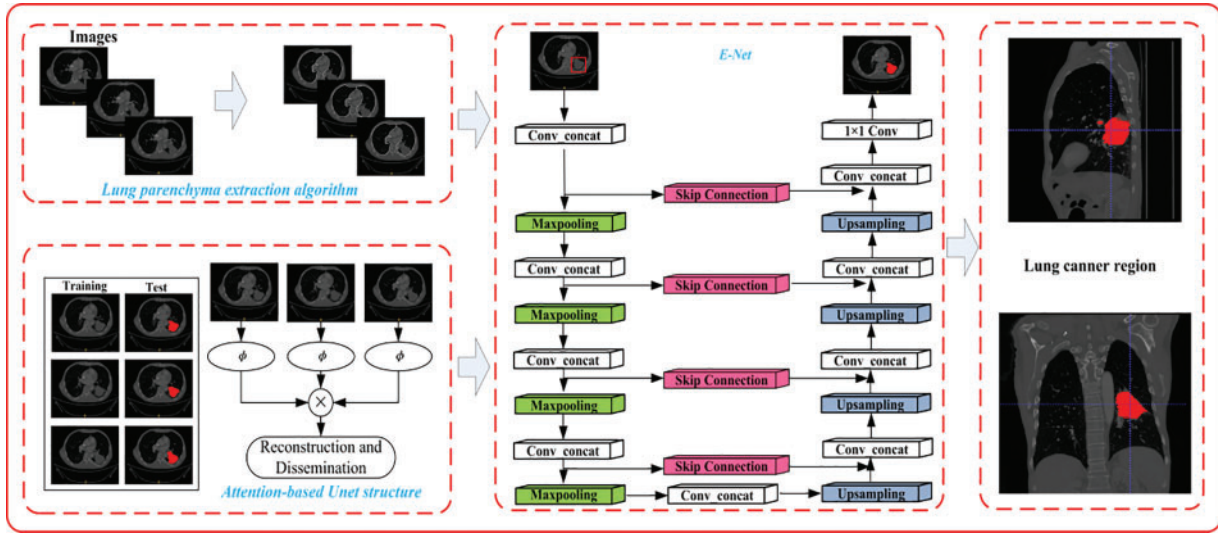


Figure 3: Algorithm structure

2.3 Lung Parenchyma Extraction Algorithm Based on Intercept Histogram

Lung cancer is a lesion embodied in the region of the lung, which presents a specific shape in the image. It would be computationally inefficient to carry out a study of fine segmentation of lung cancer by the entire image sequence. Therefore, we emulate the clinician's diagnostic methodology by initially undertaking a rudimentary segmentation of the image sequence to delineate regions with a high likelihood of neoplastic manifestation. The extant OTSU algorithm [41] examines the grayscale distribution architecture of the image and employs a bivariate histogram to articulate both the grayscale values at individual pixel locations and their spatial proximities. It partitions the histogram into four orthogonal sections demarcated by the threshold coordinates (s, t) , and the optimal threshold for image segmentation is a bivariate vector procured when the bivariate metric criterion reaches its zenith. Let f, g represent the coordinate axes of the bivariate histogram, and $T = f + g$ is a line orthogonal to the principal diagonal across the threshold point. However, traditional 2D OTSU assumes that the likelihood of two orthogonal sections deviating from the diagonal is nil, thus compromising the precision of image segmentation.

To enhance both the algorithm's efficacy and its resilience against noise perturbations, the linear intercept histogram algorithm employs the intercept T orthogonal to the principal diagonal as the threshold parameter. Assuming (x, y) symbolizes a binary tuple comprising the gray intensity of a pixel and the mean gray intensity of its surrounding pixels, the probability that $x + y = k$ pixels exist within the image is defined as follows:

$$p_k = \frac{n_k}{M \times N}, k = 0, 1, \dots, 2(L - 1) \quad (1)$$

where $M \times N$ is the image resolution and L is the image gray level.

Let \mathbf{A} be the set of gray intensities and mean gray intensities for all pixels in the image: $\mathbf{A} = \{(x, y) | x, y \in (0, 1, \dots, L-1)\}$. Let \mathbf{B} represent the set of all intercept values. Let g be the mapping $g = x + y$ from $\mathbf{A} \rightarrow \mathbf{B}$. The linear intercept histogram OTSU algorithm calculates the threshold by considering the entire histogrammatic landscape, thereby somewhat fortifying its noise robustness.

However, due to the low signal-to-noise ratio inherent in pulmonary imaging of lung cancer, further optimization of the algorithm is requisite. The algorithm tends to obliterate detailed nuances in the lower-intensity spectrum of the image during the segmentation process, thus necessitating improvements for more refined image partitioning. For this purpose, we propose the directional fuzzy derivative intercept histogram. The derivative of the central pixel (i, j) in a specified direction D is defined as the disparity among adjacent pixels along the same directional vector D .

$$V_N(i, j) = I(i, j - 1) - I(i, j) \quad (2)$$

where $V_N(i, j)$ is the blurred cepstrum value of pixel (i, j) in the N direction.

Assume an edge traverses the pixel (i, j) along the southwest-northeast (SW-NE) vectorial axis, its corresponding derivative magnitude $V_{NW}(i, j)$ would be augmented, and the positional derivatives of the neighboring pixels orthogonal to the edge trajectory would similarly be amplified. Consequently, if an edge intersects (i, j) along the SW \rightarrow NE directional axis, the derivative magnitudes of the pixels located at $(i + 1, j - 1)$ and $(i - 1, j + 1)$ in the northwest (NW) vectorial direction will be more pronounced. The arithmetic mean of these three derivative magnitudes serves as the fuzzy derivative metric in the NW vectorial direction at (i, j) and is denoted as $g(i, j)$. A diminished $g(i, j)$ suggests the absence of an edge within this localized region, implying that the image characteristics in the SW \rightarrow NE directional axis are homogenous and belong to a singular region. To attenuate the noise and accentuate the granular detail within the image, fuzzy derivative metrics for the 8 proximate neighborhoods of pixel points are computed individually. The maximal fuzzy derivative metric is selected to supplant the domain-specific mean value at the corresponding pixel point.

Let $[x, g(i, j)]$ represent the binary set comprising the fuzzy derivative metrics and the gray intensity values of the pixels, along with their angular orientations. Let p_k denote the likelihood that pixels within the image satisfy $x + g(i, j) = k$. Within the framework of the intercept histogram, the image undergoes segmentation via a threshold parameter T^* .

$$w_0(T) = \sum_{k=0}^T p_k, \quad w_1(T) = 1 - w_0(T) \quad (3)$$

$$\mu(T) = \sum_{k=0}^T k p_k \quad (4)$$

$$\sigma^2(T) = w_0(T) w_1(T) [\mu_0(T) - \mu_1(T)]^2 \quad (5)$$

Then the corresponding optimal threshold T^* satisfies.

$$\sigma^2(T^*) = \max_{0 \leq T \leq 2L-2} \sigma^2(T) \quad (6)$$

After obtaining the optimal threshold T^* , pixel classification is achieved. Since lung cancer presents connected regions, we use 8-connected region labeling for all pixel points.

2.4 Attention-Based Unet Structure

The Unet network topology has manifested superlative efficacy in medical image segmentation, comprising primarily of encoding, decoding, and skip-connection frameworks. To enhance this architecture, we propose the augmented Unet (referred to as E-Unet). A Conv_E module is constructed in the encoding phase to achieve optimized feature amalgamation, whilst the maximal pooling layer accomplishes downsampling by retaining only the most salient features and discarding insignificant

attributes. The decoding segment incorporates the Conv_E module and upsampling to recompense for the information attrition incurred during pooling. A 1×1 convolution operation subsequently categorizes the pixels within the image as per classification requisites, culminating in the final segmentation outcome.

The E-Unet module we have engineered is delineated in Fig. 4, and the deviations from the traditional Unet architecture are primarily articulated as follows: 1) The feature vector generated via convolution is amalgamated with the antecedent input module attributes and subsequently utilized as the input for the ensuing convolution. This augmentation facilitates feature transference, achieves feature recycling, and mitigates the attrition of image features during the training regimen, yet still forfeits image peripheries and the image dimensions remain incongruent pre- and post-convolution. Therefore, to synchronize the feature maps before and after convolution within the module, it is imperative to pad the feature maps prior to convolution to ensure homogeneous feature map resolution antecedent to and subsequent to convolution. 2) The incorporation of a Batch Normalization (BN) layer subsequent to each convolution in E-Unet standardizes the data, thereby attenuating the network’s reliance on the initialization parameters.

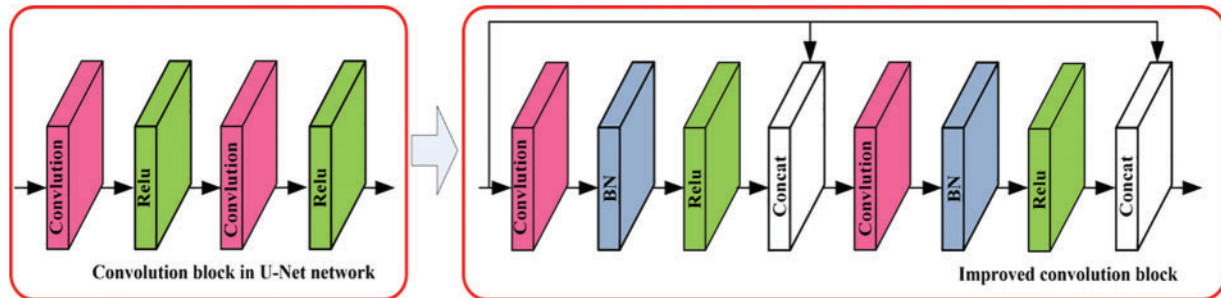


Figure 4: Convolutional block and modified convolutional block of traditional Unet network

To circumvent the adverse impact on network optimization engendered by feature distribution, the normalized data x is inversely transformed, which can be interpreted as an adjustment in offset and scaling coefficients.

$$y_i = \gamma_i x_i + \beta_i \tag{7}$$

where γ_i is the variance and β_i is the mean.

We examine lung image sequences characterized by robust inter-frame correlation. To capitalize on this correlation to aid in segmentation, we have architected a profound neural network. This methodology eschews the necessity for labeled data; the network is formulated directly from the image data. We architect the Bottleneck layer within a self-supervised learning paradigm, thereby enabling the network to discern the set of pixels in the reference frame that optimally align with the extant pixel feature, and reconstitute said feature.

Given the image where the lung cancer is located as I_i , the neural network feature encoder with training is $\phi(\bullet; \theta)$; the features extracted from the input frames after the feature encoder are represented as $f_i = \phi(g(I_i); \theta)$, and $g(I_i)$ denotes Bottleneck in self-supervision. The similitude between a specific frame is calculated via a soft attention mechanism, culminating in the acquisition of the similarity matrix A . Contemplating the spatial contiguity between frames, the features are synthesized.

$$\hat{I}_i = \sum_{j \in N} A_i^j I_{t-1}^j \quad (8)$$

$$A_i^j = \frac{\exp \langle f_i^i, f_{t-1}^i \rangle}{\sum_{n \in N} \exp \langle f_i^i, f_{t-1}^n \rangle}, \quad N = \{\forall n \in I_{t-1}, |n - i| \leq c\} \quad (9)$$

where c denotes the neighborhood radius of pixel i ; $\langle . \rangle$ denotes the number product operation. Through training, the feature encoder endowed with Huber loss is attained, i.e., the features elucidated by the encoder strive to minimize the Euclidean distance between the pixel features in I_i and those in I_{t-1} , while maximizing the distance from other irrelevant pixel attributes, and the optimization criterion for the network is articulated as

$$\hat{\theta} = \arg \min_{\theta} \vartheta (I_i, \hat{I}_i) \quad (10)$$

For the test, the similarity matrix is calculated from the feature map of the image frames to obtain the corresponding segmentation results.

$$\hat{y}_i^j = \sum_{j \in N} A_i^j y_{t-1}^j \quad (11)$$

To exhaustively investigate the interrelations among images, we incorporated an auxiliary attention module to glean information from the images for the purpose of target reconstruction. Given I_q and K reference frames retained in the attention module, these are designated as I_r . We orchestrated two feature encoders predicated upon the neural network architecture to procure image features, denoted as f_q and f_r respectively. During the training process, θ was used in dynamic update mode: $\theta_r \leftarrow m\theta_r + (1 - m)\theta_q$.

2.5 Multi-Dimensional Fusion Loss Function

In the realm of medical image segmentation, the prevalently employed loss functions encompass Binary Cross Entropy Loss (BCE Loss), BCEWith Logits Loss, and Dice Loss, among others. Taking into account the specific spatial and pixel value distributions pertinent to lung cancer, the merits of BCEWith Logits Loss and Dice Loss are harnessed to formulate the loss function.

BCEWith Logits Loss: Computes the cross entropy to depict the loss. It is imperative that the samples reside within the interval $[0,1]$. The ubiquitously deployed Sigmoid function has a propensity to engender unstable computational outcomes and induce the vanishing of gradients. For this rationale, the Sigmoid and BCE Loss are aggregated into a singular class, thereby rendering computational outcomes more homogeneous, which is propitious for stable gradient back-propagation.

$$L_{BCEW} = - \sum_{i=1}^n y_i \log (\sigma (x_i)) + (1 - y_i) \log (1 - \sigma (x_i)) \quad (12)$$

where x_i is the training sample, y_i is the sample label, n is the number of samples, and $\sigma(x_i)$ is the Sigmoid function. This loss function appraises each category equitably, which is inimical to the refinement of segmentation networks with disproportionate categories. Conversely, in the lung tumor images requiring processing in this manuscript, the foreground pixels (tumor region) are significantly outnumbered by the background pixels (non-tumor region); thus, the background pixels will predominate the loss function and orient this network towards optimizing the background.

The Dice coefficient characterizes the overlapping part of the two samples and takes values in the range $[0, 1]$.

$$D(G, R) = 2 \frac{|G \cap P|}{|G \cup P|} \quad (13)$$

where G denotes Ground Truth and P denotes Prediction. Then the corresponding DissLoss is

$$L_{Dice} = 1 - \frac{2 \sum G \bullet P}{\sum (G^2 + P^2)} \quad (14)$$

Dice Loss avoids the network from falling into a local minimum of the loss function by increasing the weight of the foreground region, i.e., avoiding the occurrence of overfitting. Therefore, Dice Loss is more suitable for the case of extreme sample imbalance.

$$\frac{\partial L_{Dice}}{\partial P} = \frac{2G^2}{(G + P)^2} \quad (15)$$

when the sum of G and P is minimal, the gradient values fluctuate acutely, leading to exceedingly volatile training that is detrimental to network optimization. Therefore, although Dice Loss ameliorates the perturbations instigated by category imbalance on network training, it occasionally becomes unreliable and adverse for back-propagation. To ameliorate this, we construct a novel objective function to leverage the advantages of both.

$$Loss = \lambda L_{BCEW} + (1 - \lambda) L_{Dice} \quad (16)$$

3 Results and Discussion

Utilizing publicly accessible datasets as well as datasets amassed autonomously by medical institutions, the proffered algorithm was juxtaposed with conventional algorithms to ascertain its efficacy in lung extraction and lung cancer segmentation.

To corroborate the algorithm's performance, we incorporate the Area Overlap Measure (AOM), a widely-recognized evaluative metric in the domain of medical imaging. Additionally, Area Over-segmentation measure (AVM), Area Under-segmentation measure (AUM), and Combination Measure (CM) are delineated as [42]

$$AOM = \frac{R_s \cap R_g}{R_s \cup R_g} \quad (17)$$

$$AVM = \frac{R_s - R_g}{R_s} \quad (18)$$

$$AUM = \frac{R_g - R_s}{R_g} \quad (19)$$

$$CM = \frac{1}{3} \{AOM + (1 - AVM) + (1 - AUM)\} \quad (20)$$

where R_g is the result of physician labeling and R_s is the result derived by the algorithm. AOM and CM are proportional to the segmentation result. AUM and AVM are inversely proportional to the segmentation result.

3.1 Lung Parenchyma Extraction Effect

The precision of lung extraction serves as a sine qua non for accurate lung cancer detection, aiming to maximize the retention of the neoplastic lung region and obviate false negatives. The advanced algorithm is juxtaposed against canonical algorithms for evaluation, as shown in [Table 1](#).

Table 1: Lung parenchyma extraction effect

Algorithm	<i>AOM</i>	<i>AVM</i>	<i>AUM</i>	<i>CM</i>
T [12]	0.81	0.26	0.31	0.75
Watershed [15]	0.87	0.24	0.25	0.79
OTSU [41]	0.90	0.21	0.21	0.83
B_Attention [21]	0.91	0.19	0.18	0.85
D_Guass [39]	0.93	0.18	0.14	0.87
Ours	0.96	0.15	0.12	0.90

The manual Threshold method (T method) [12] employs an interactive modality for threshold delineation, capable of discerning neoplastic zones with varied indicia. Nevertheless, this technique establishes the threshold under the auspices of a known neoplastic condition, rendering the algorithm's scalability constrained, and its segmentation effect is relatively poor. The Watershed method [15] draws inspiration from topographical distributions to fabricate a segmentation model. In the context of images manifesting attenuated lung cancer signals, the demarcation between foreground and background domains is nebulous, thus engendering detection lapses. The OTSU method [41] ascertains the threshold based on statistical histogram analysis, solely utilizing the distributional attributes of pixel grayscale values, while disregarding spatial attributes, leading to inadequate sensitivity in detecting incipient lung carcinomas. Liu et al. [21] proposed a boundary attention model based on deep learning to extract lung parenchyma, which achieves good results for lung images with relatively intact parenchyma. However, its performance is less satisfactory for lung images with significant pathological changes. The Double Gaussian (D_Guass) fitting algorithm [39] scrutinizes the pixel distribution within pulmonary images, treating the pixel distributions of background and foreground as bimodal Gaussian models. It opts for the local nadir as the segmentation threshold, thereby accomplishing neoplastic region extraction. All aforementioned algorithms develop models predicated on pixel-level considerations, yet neglect localized spatial data pertaining to lesions. In antithesis, the proposed algorithm incorporates spatial parameters, exploits features within less illuminated regions, mitigates noise, amplifies the detection sensitivity for subdued targets, and attains an AOM of 0.96. An intuitive visual representation of the advanced algorithm is elucidated in [Fig. 5](#), which demonstrates efficacious lung extraction.

3.2 Segmentation Performance Effect

To ascertain the efficacy of the neural network architecture, we depict the training curves for Unet, Residual Unet [35], VGG Unet [36], E-Unet, and E-L-Unet, as elucidated in [Fig. 6](#). It is manifest that the encoder-decoder topology, constructed on the foundation of the Unet network, can facilitate lung carcinoma segmentation. Residual Unet incorporates a residual architecture to delve more profoundly into image feature space, thereby augmenting performance metrics. The VGG Unet combines the VGG network to extract prominent features from the images, and then inputs them into Unet, thereby enhancing its representational capacity. E-Unet standardizes the image predicated on the delineated

regions of suspicion, implements an attention mechanism, and concentrates computational resources on the domain where the lung cancer is situated. This approach diminishes the computational overhead for lung cancer segmentation and elevates the accuracy of segmentation. E-L-Unet refines the loss function predicated on E-Unet; although Dice loss mitigates the perturbations introduced by category imbalance in network training, it exhibits computational instability, which is inimical to gradient back-propagation. In response to this, we amalgamate BCEWith Logits Loss and Dice loss functions, thereby proposing a novel loss function that enhances both the stability and rate of convergence for the network.

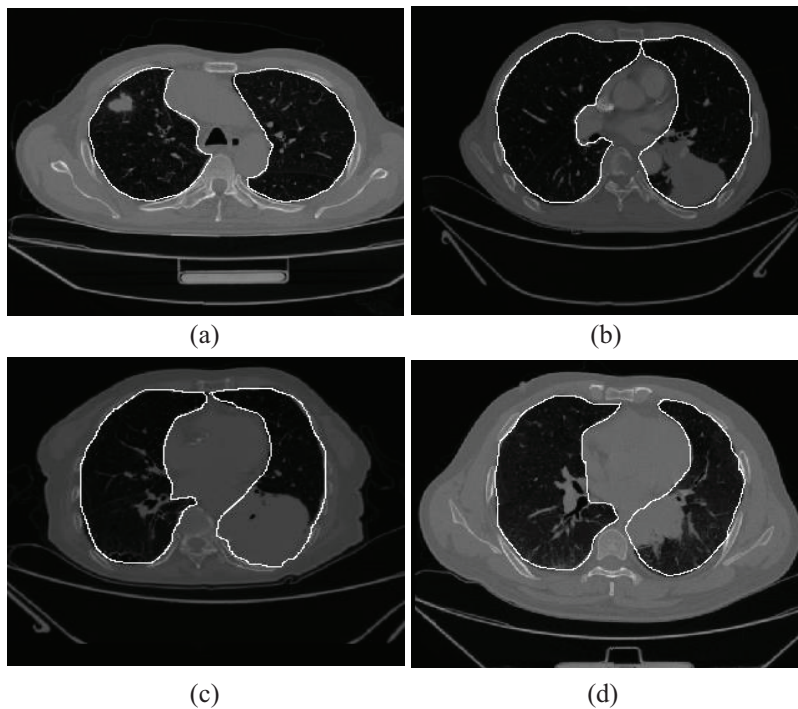


Figure 5: Lung parenchyma extraction regions. (a) Lung carcinoma lesion situated in the left lung, characterized by minimal area and a homogenous texture. (b) Lung carcinoma lesion adhered to the thoracic wall. (c) Lung carcinoma located in the lower quadrant of the right lung, which has induced a blurring effect on the thoracic wall. (d) Lung carcinoma with indistinct boundaries

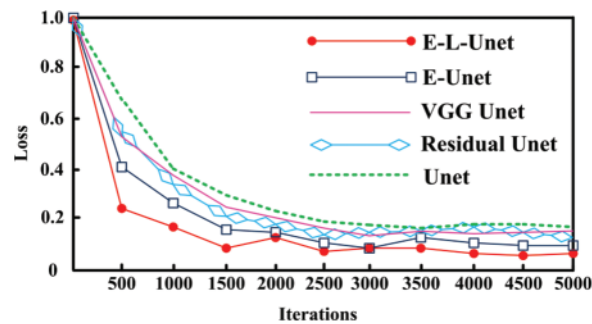


Figure 6: Training curve

The proffered algorithm emulates the diagnostic methodology of clinicians, directing focus sequentially onto the regions delineated as suspicious. Initially, a thresholding technique is utilized to ascertain the region of focal interest. This modality employs classical features and boasts rapid computational performance. Subsequently, segmentation is executed via a deep learning architecture. Although this phase incurs a more protracted time investment during training, ranging from 1–2 weeks, upon completion of the network’s training, the test outcomes are swiftly generated. The predominant operations entailed in this step comprise addition and multiplication. Hence, the algorithmic complexity remains relatively subdued, and the segmentation efficacy is contingent upon the dimensions of the focal region. On average, the algorithm necessitates approximately 13 s for detection, in stark contrast to a minimum of 10 min required for a clinician’s examination.

We employed both conventional algorithms [8,13] and deep learning algorithms [27,30,38] to scrutinize their segmentation efficacies, as delineated in Table 2. The application of mathematical statistical methods [8], involving the construction of an analytical model, enables lung carcinoma segmentation. However, the inherent variability in the sample set leads to rudimentary computations that fall short of fulfilling practical application requirements, culminating in suboptimal segmentation performance. The Hybrid Geometric Active Contour (HGAC) algorithm [13], predicated on the concept of energy minimization, actualizes image segmentation when internal and external forces are commensurate. This algorithm yields superior results in scenarios where there is a pronounced discrepancy between the lung carcinoma and background regions. Traditional algorithms are unencumbered by reliance on pre-existing large datasets, and their feature selection demonstrates robust interpretability. However, the inherent limitations of the chosen features preclude significant advancements in segmentation accuracy. Recent developments in deep learning, such as the Unet [38], utilize a U-shaped network topology to extract depth features pertinent to lung carcinoma segmentation. Nonetheless, this architecture is solely designed from the vantage point of individual image frames, neglecting to account for the spatial morphology constraints inherent to lung carcinomas. CNN networks amalgamated with Atlas [30] hone in on key regions to accomplish lung carcinoma segmentation, but they exhibit shortcomings when applied to smaller samples characterized by diminished area and attenuated signal intensity. WS-LungNet [27] employs a two-tiered, weakly-supervised framework for lung carcinoma detection and diagnosis, mitigating the issue of weak labels. The integration of an attention mechanism further augments the segmentation of lung carcinoma.

Table 2: Algorithm segmentation effect

Representation	Algorithm	<i>AOM</i>	<i>AVM</i>	<i>AUM</i>	<i>CM</i>
Lung carcinoma lesion situated in the left lung, characterized by minimal area and a homogenous texture.	Mathematical statistics [8]	0.74	0.23	0.31	0.73
	HGAC [13]	0.79	0.22	0.29	0.76
	Unet [38]	0.84	0.21	0.27	0.79
	Atlas + CNN [30]	0.86	0.18	0.21	0.82
	WS-LungNet [27]	0.92	0.15	0.19	0.86
	Ours	0.93	0.13	0.15	0.88

(Continued)

Table 2 (continued)

Representation	Algorithm	<i>AOM</i>	<i>AVM</i>	<i>AUM</i>	<i>CM</i>
Lung carcinoma lesion adhered to the thoracic wall.	Mathematical statistics [8]	0.71	0.33	0.29	0.70
	HGAC [13]	0.73	0.28	0.27	0.73
	Unet [38]	0.81	0.26	0.25	0.77
	Atlas + CNN [30]	0.84	0.22	0.19	0.81
	WS-LungNet [27]	0.87	0.18	0.18	0.84
	Ours	0.90	0.16	0.16	0.86
Lung carcinoma located in the lower quadrant of the right lung, which has induced a blurring effect on the thoracic wall.	Mathematical statistics [8]	0.69	0.29	0.31	0.70
	HGAC [13]	0.72	0.25	0.30	0.72
	Unet [38]	0.74	0.24	0.26	0.75
	Atlas + CNN [30]	0.79	0.24	0.23	0.77
	WS-LungNet [27]	0.85	0.21	0.21	0.81
	Ours	0.88	0.19	0.17	0.84
Lung carcinoma with indistinct boundaries.	Mathematical statistics [8]	0.58	0.35	0.35	0.63
	HGAC [13]	0.63	0.33	0.32	0.66
	Unet [38]	0.69	0.29	0.30	0.70
	Atlas + CNN [30]	0.75	0.28	0.28	0.73
	WS-LungNet [27]	0.80	0.26	0.25	0.76
	Ours	0.84	0.22	0.20	0.81

Four illustrative sequences were selected to validate the algorithm's performance. Fig. 7a depicts a circumscribed lung carcinoma lesion situated in the left lung, characterized by minimal area and a homogenous texture. Owing to its biopsy-confirmed status as lung carcinoma, the lesion proves amenable to straightforward segmentation, eliciting favorable outcomes across all algorithms. Fig. 7b exhibits a substantial lung carcinoma lesion adhered to the thoracic wall; however, the prominence of the thoracic wall complicates the segmentation task, resulting in compromised performance across all algorithms. Fig. 7c reveals a sizable lung carcinoma located in the lower quadrant of the right lung, which has induced a blurring effect on the thoracic wall, further exacerbating the complexity of the segmentation task and leading to diminished algorithmic performance. Fig. 7d portrays a lung carcinoma with indistinct boundaries, rendering accurate localization infeasible and thus undermining algorithmic performance. In particular, traditional, feature-based algorithms experience a precipitous decline in efficacy due to their incapacity to accurately represent features. Deep learning-based algorithms, which construct a semantic bridge via neural propagation mechanisms, exhibit a more gradual degradation in performance. The algorithm proposed in this paper, built upon an interpretable conceptual framework and incorporating an attention-based Unet network, synthesizes both deep and traditional features. While its performance has also seen some attenuation, it retains a degree of optimality.

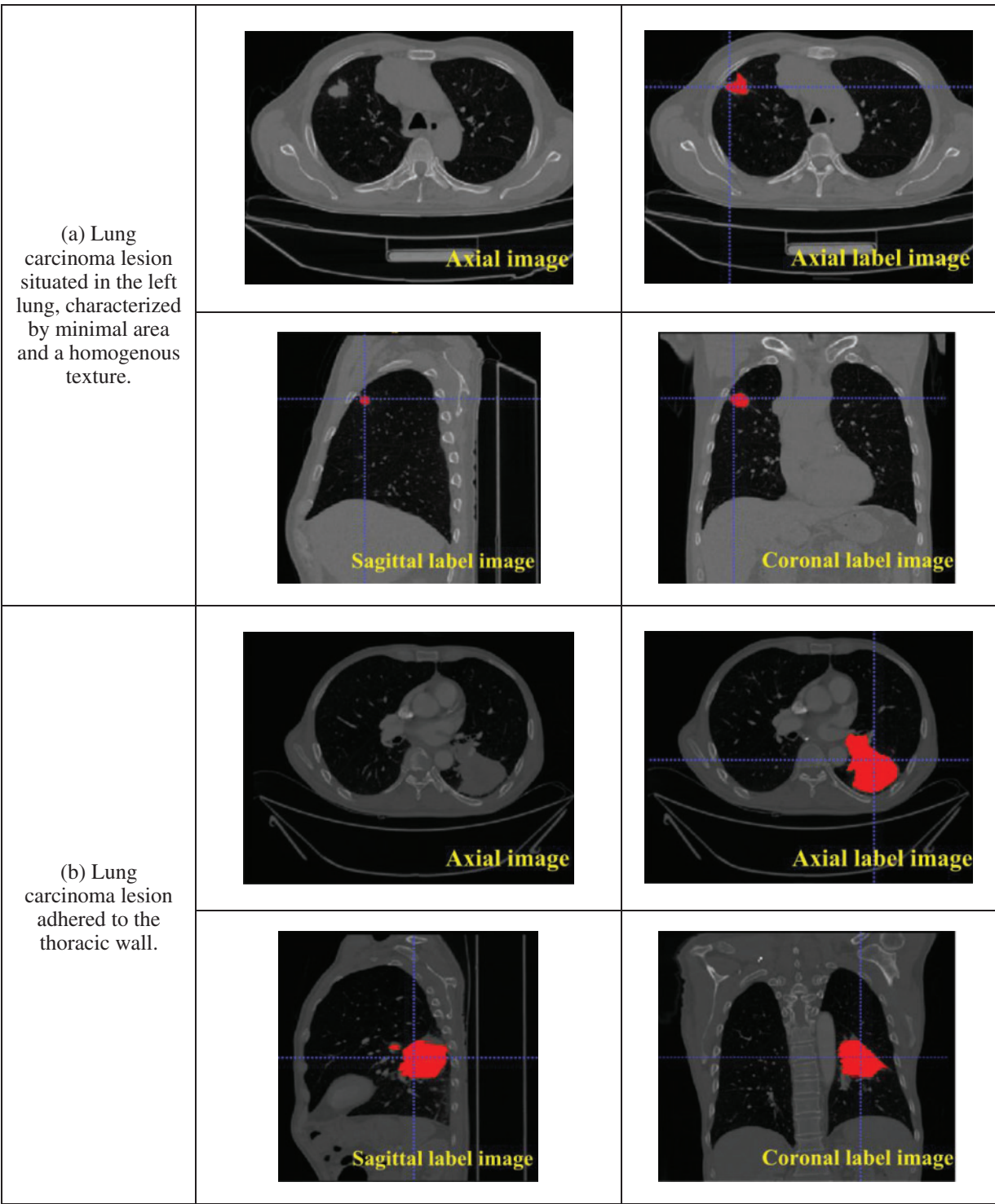


Figure 7: (Continued)

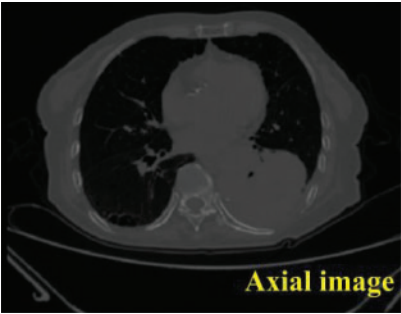
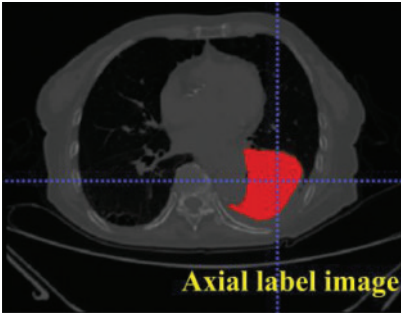
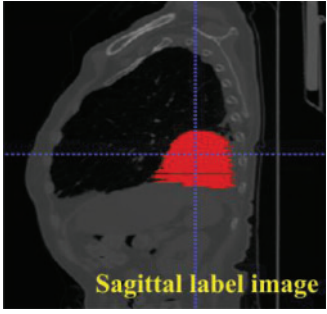
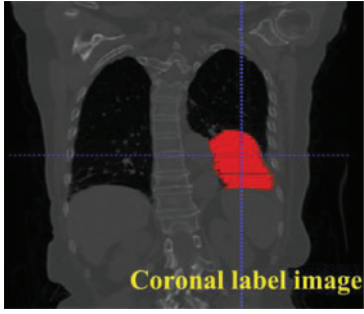
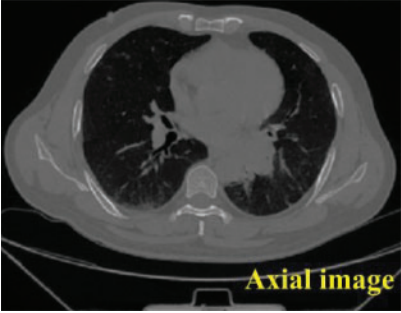
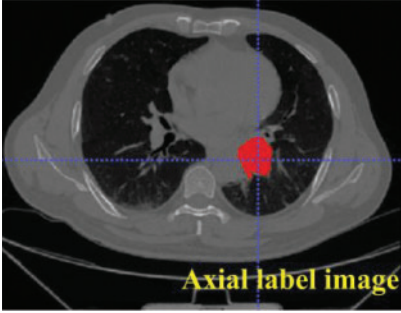
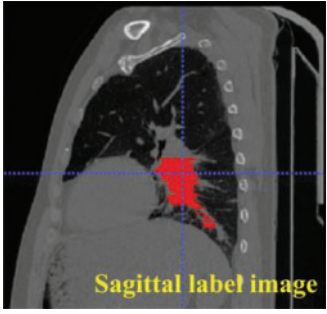
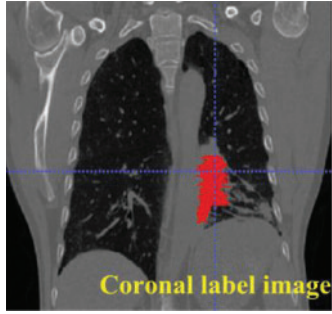
<p>(c) Lung carcinoma located in the lower quadrant of the right lung, which has induced a blurring effect on the thoracic wall.</p>	 <p>Axial image</p>	 <p>Axial label image</p>
	 <p>Sagittal label image</p>	 <p>Coronal label image</p>
<p>(d) Lung carcinoma with indistinct boundaries.</p>	 <p>Axial image</p>	 <p>Axial label image</p>
	 <p>Sagittal label image</p>	 <p>Coronal label image</p>

Figure 7: The algorithm labels images for lung cancer

The algorithm articulated in this paper examines the three-dimensional morphological attributes of pulmonary carcinoma imagery predicated on Unet and incorporates an attention mechanism to concentrate on neoplastic pulmonary regions, thereby facilitating precise segmentation. The algorithm's efficacy is substantiated through axial, coronal, and sagittal segmentations of lung cancer. The loss function has been refined to enable accurate segmentation under conditions characterized by a paucity of positive samples.

4 Conclusion and Future Directions

Pulmonary carcinoma manifests a prevalent incidence rate, thereby rendering early diagnostic and therapeutic interventions imperative. Given the voluminous corpus of computed tomography (CT) imagery pertinent to pulmonary structures and the inherent complexities in demarcating neoplastic lesions, this treatise advances a lung carcinoma segmentation algorithm, conceptualized from the vantage point of medical diagnostic protocols, and predicated on explainable attention Unet. The salient innovations are itemized herewith:

1. Leveraging traditional features, we proffer an intercept histogram algorithm with a focus on the pulmonary parenchyma.
2. Predicated on the deep learning architecture of the Unet network, we advance an attention-based Unet for the precise segmentation of pulmonary carcinoma.
3. In consideration of sample imbalance phenomena, we introduce a multi-dimensional fusion loss function to augment network convergence.

Experimental evaluations, undertaken utilizing both publicly accessible datasets and clinically amassed databases, corroborate that the proposed algorithm is capable of segmenting an array of morphologies and lesions associated with both diminutive and extensive pulmonary carcinomas. However, the algorithm in this paper was only studied on existing datasets, which have a good signal-to-noise ratio and exhibit good internal consistency within each dataset. The research on lesion segmentation in non-uniform data with lower signal-to-noise ratio is a key focus for future work.

Acknowledgement: The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

Funding Statement: This work is supported by Light of West China (No. XAB2022YN10).

Author Contributions: Huan Wang and Shi Qiu performed the experiments. Benyue Zhang and Lixuan Xiao analyzed the data. All authors conceived and designed research, and contributed to the interpretation of the data and drafting the work.

Availability of Data and Materials: The data used to support the findings of this study are available from the corresponding author upon request.

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

References

- [1] E. Barrows, M. Blackburn and S. Liu, "Evolving role of immunotherapy in small cell lung cancer," *Seminars in Cancer Biology*, vol. 86, pp. 868–874, 2022.
- [2] J. Chaft, Y. Shyr, B. Sepesi and P. Forde, "Preoperative and postoperative systemic therapy for operable non-small-cell lung cancer," *Journal of Clinical Oncology*, vol. 40, no. 6, pp. 546–555, 2022.

- [3] D. Ettinger, D. Wood, D. Aisner, W. Akerley, J. Bauman *et al.*, “Non-small cell lung cancer, version 3.2022, NCCN clinical practice guidelines in oncology,” *Network*, vol. 20, pp. 497–530, 2022.
- [4] S. Qiu, B. Li, T. Zhou, F. Li and T. Liang, “Multi-view auxiliary diagnosis algorithm for lung nodules,” *Computers, Materials & Continua*, vol. 72, no. 3, pp. 4897–4910, 2022.
- [5] T. Zhou, Q. Cheng, H. Lu, Q. Li, X. Zhang *et al.*, “Deep learning methods for medical image fusion: A review,” *Computers in Biology and Medicine*, vol. 160, pp. 106959, 2023. <https://doi.org/10.1016/j.compbimed.2023.106959>
- [6] T. Zhou, Q. Li, H. L. Lu, Q. R. Cheng and X. X. Zhang, “GAN review: Models and application of medical image fusion,” *Inf Fusion*, vol. 91, pp. 134–148, 2023.
- [7] P. Tripathi, S. Tyagi and M. Nath, “A comparative analysis of segmentation techniques for lung cancer detection,” *Pattern Recognition and Image Analysis*, vol. 29, pp. 167–173, 2019.
- [8] P. Mazzone, N. Obuchowski, M. Phillips, B. Risius, B. Bazerbashi *et al.*, “Lung cancer screening with computer aided detection chest radiography: Design and results of a randomized, controlled trial,” *PLoS One*, vol. 8, no. 3, pp. e59650, 2013.
- [9] A. Mathews and M. Jeyakumar, “Analysis of lung tumor detection using various segmentation techniques,” in *2020 Int. Conf. on Inventive Computation Technologies (ICICT)*, Coimbatore, India, IEEE, pp. 454–458, 2020.
- [10] W. L. Cai and G. B. Hong, “Quantitative image analysis for evaluation of tumor response in clinical oncology,” *Chronic Diseases and Translational Medicine*, vol. 4, no. 1, pp. 18–28, 2018.
- [11] J. Alam, S. Alam and A. Hossan, “Multi-stage lung cancer detection and prediction using multi-class svm classify,” in *2018 Int. Conf. on Computer, Communication, Chemical, Material and Electronic Engineering IEEE (IC4ME2)*, Rajshahi, Bangladesh, pp. 1–4, 2018.
- [12] M. Wanet, J. A. Lee, B. Weynand, M. de Bast, A. Poncelet *et al.*, “Geets gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: A comparison with threshold-based approaches, CT and surgical specimens,” *Radiotherapy and Oncology*, vol. 98, no. 1, pp. 117–125, 2011.
- [13] W. Zhang, X. Wang, X. Li and J. Chen, “3D skeletonization feature based computer-aided detection system for pulmonary nodules in CT datasets,” *Computers in Biology and Medicine*, vol. 92, pp. 64–72, 2018.
- [14] S. Moon, J. Kim, J. Joung, H. Cha, W. Park *et al.*, “Correlations between metabolic texture features, genetic heterogeneity, and mutation burden in patients with lung cancer,” *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 46, pp. 446–454, 2019.
- [15] S. Avinash, K. Manjunath and S. Kumar, “An improved image processing analysis for the detection of lung cancer using Gabor filters and watershed segmentation technique,” in *2016 Int. Conf. on Inventive Computation Technologies (ICICT)*, Coimbatore, India, IEEE, vol. 3, pp. 1–6, 2016.
- [16] I. Nazir, I. Haq, M. Khan, M. Qureshi, H. Ullah *et al.*, “Efficient pre-processing and segmentation for lung cancer detection using fused CT images,” *Electronics*, vol. 11, no. 1, pp. 34, 2021.
- [17] N. Garau, C. Paganelli, P. Summers, D. Bassis, C. Lanza *et al.*, “A segmentation tool for pulmonary nodules in lung cancer screening: Testing and clinical usage,” *Physica Medica*, vol. 90, pp. 23–29, 2021.
- [18] T. Zhang, K. Wang, H. Cui, Q. Jin, P. Cheng *et al.*, “Topological structure and global features enhanced graph reasoning model for non-small cell lung cancer segmentation from CT,” *Physics in Medicine & Biology*, vol. 68, no. 2, pp. 025007, 2023.
- [19] H. Yu, Z. Zhou and Q. Wang, “Deep learning assisted predict of lung cancer on computed tomography images using the adaptive hierarchical heuristic mathematical model,” *IEEE Access*, vol. 8, pp. 86400–86410, 2020.
- [20] P. M. Shakeel, M. A. Burhanuddin and M. I. Desa, “Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier,” *Neural Computing and Applications*, vol. 34, pp. 1–14, 2022.
- [21] X. Liu, H. Shen, L. Gao and R. Guo, “Lung parenchyma segmentation based on semantic data augmentation and boundary attention consistency,” *Biomedical Signal Processing and Control*, vol. 80, pp. 104205, 2023.

- [22] S. Shamas, S. N. Panda and I. Sharma, "Review on lung nodule segmentation-based lung cancer classification using machine learning approaches," in *Artificial Intelligence on Medical Data: Proc. of Int. Symp., ISCMM 2021*, Singapore, Springer Nature Singapore, pp. 277–286, 2022.
- [23] H. Yang, L. Shen, M. Zhang and Q. Wang, "Uncertainty-guided lung nodule segmentation with feature-aware attention," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, Cham, Springer Nature Switzerland, pp. 44–54, 2022.
- [24] W. Alakwaa, M. Nassef and A. Badr, "Lung cancer detection and classification with 3D convolutional neural network (3D-CNN)," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 8, 2017.
- [25] B. Skourt, A. El Hassani and A. Majda, "Lung CT image segmentation using deep neural networks," *Procedia Computer Science*, vol. 127, pp. 109–113, 2018.
- [26] H. Shaziya, K. Shyamala and R. Zaheer, "Automatic lung segmentation on thoracic CT scans using U-net convolutional network," in *2018 Int. Conf. on Communication and Signal Processing (ICCSPP)*, IEEE, Chennai, India, 2018.
- [27] Z. Shen, P. Cao, J. Yang and O. R. Zaiane, "WS-LungNet: A two-stage weakly-supervised lung cancer detection and diagnosis network," *Computers in Biology and Medicine*, vol. 154, pp. 106587, 2023. <https://doi.org/10.1016/j.compbiomed.2023.106587>
- [28] T. Zhou, F. Liu, X. Ye, H. Wang and H. Lu, "CCGL-YOLOV5: A cross-modal cross-scale global-local attention YOLOV5 lung tumor detection model," *Computers in Biology and Medicine*, vol. 165, pp. 107387, 2023. <https://doi.org/10.1016/j.compbiomed.2023.107387>
- [29] F. Zhang, Q. Wang and H. Li, "Automatic segmentation of the gross target volume in non-small cell lung cancer using a modified version of resNet," *Technology in Cancer Research & Treatment*, vol. 19, pp. 533033820947484, 2020.
- [30] T. Zhang, Y. Yang, J. Wang, K. Men, X. Wang *et al.*, "Comparison between atlas and convolutional neural network based automatic segmentation of multiple organs at risk in non-small cell lung cancer," *Medicine*, vol. 99, no. 34, pp. e21800, 2020.
- [31] Q. Hu, L. Souza, G. Holanda, S. Alves, F. Silva *et al.*, "An effective approach for CT lung segmentation using mask region-based convolutional neural networks," *Artificial Intelligence in Medicine*, vol. 103, pp. 101792, 2020.
- [32] A. Asuntha and A. Srinivasan, "Deep learning for lung cancer detection and classification," *Multimedia Tools and Applications*, vol. 79, pp. 7731–7762, 2020.
- [33] B. Ni, Z. Liu, X. Cai, M. Nappi and S. Wan, "Segmentation of ultrasound image sequences by combing a novel deep siamese network with a deformable contour model," *Neural Computing and Applications*, vol. 35, no. 20, pp. 14535–14549, 2023.
- [34] S. Hou, T. Zhou, Y. Liu, P. Dang, H. Lu *et al.*, "Teeth U-Net: A segmentation model of dental panoramic X-ray images for context semantics and contrast enhancement," *Computers in Biology and Medicine*, vol. 152, pp. 106296, 2023.
- [35] J. Sousa, T. Pereira, F. Silva, M. Silva, A. Vilares *et al.*, "Lung segmentation in CT images: A residual U-Net approach on a cross-cohort dataset," *Applied Sciences*, vol. 12, no. 4, pp. 1959, 2022.
- [36] W. Shao, Z. Wang, Z. Zhang, Q. Zhou, R. Wang *et al.*, "A segmentation method of airway from chest CT image based on VGG-Unet neural network," in *2022 IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM)*, IEEE, Las Vegas, NV, USA, pp. 1702–1705, 2022.
- [37] A. Alameen, "Smart lung tumor prediction using dual graph convolutional neural network," *Intelligent Automation & Soft Computing*, vol. 36, no. 1, pp. 369–383, 2023.
- [38] Z. Zhou, M. Rahman, N. Tajbakhsh and J. Liang, "UNet++: A nested U-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support 4th Int. Workshop, DLMIA 2018, and 8th Int. Workshop, ML-CDS 2018*, Granada, Spain, pp. 3–11, 2018.
- [39] S. Qiu, D. Wen, Y. Cui and J. Feng, "Lung nodules detection in CT images using Gestalt-based algorithm," *Chinese Journal of Electronics*, vol. 25, no. 4, pp. 711–718, 2016.

- [40] O. Bader and T. Fuchs, “Gestalt perception and the experience of the social space in autism: A case study,” *Psychopathology*, vol. 55, no. 3–4, pp. 1–8, 2022.
- [41] F. Essaf, Y. Li, S. Sakho, P. K. Gadosey and T. Zhang, “An improved lung parenchyma segmentation using the maximum inter-class variance method (OTSU),” in *Proc. of the 2020 6th Int. Conf. on Computing and Artificial Intelligence*, Tianjin, China, pp. 204–212, 2020.
- [42] S. Qiu, Y. Jin, S. Feng, T. Zhou and Y. Li, “Dwarfism computer-aided diagnosis algorithm based on multimodal pyradiomics,” *Information Fusion*, vol. 80, pp. 137–145, 2022.