A Parallel Network for LRCT Segmentation and Uncertainty Mitigation with Fuzzy Sets

Shiyi Wang[1] **Yang Nan**[1] **Xiaodan Xing**[1] **Yingying Fang**[1] **Simon LF Walsh**[2,3] **Guang Yang**[1,2,3,4]

[1]Bioengineering Department and Imperial- X., Imperial College London., UK
[2]National Heart and Lung In- stitute., Imperial College London., UK
[3]Cardiovascular Research Centre., Royal Brompton Hospital., UK
[4] Imaging Sciences Dept., King's College London., UK

Abstract

Accurate segmentation of airways in Low-Resolution CT (LRCT) scans is vital for diagnostics in scenarios such as reduced radiation exposure, emergency response, or limited resources. Yet manual annotation is labor-intensive and prone to variability, while existing automated methods often fail to capture small airway branches in lower-resolution 3D data. To address this, we introduce **FuzzySR**, a parallel framework that merges super-resolution (SR) and segmentation. By concurrently producing high-resolution reconstructions and precise airway masks, it enhances anatomic fidelity and captures delicate bronchi. FuzzySR employs a deep fuzzy set mechanism, leveraging learnable t-distribution and triangular membership functions via cross-attention. Through parameters μ, σ, and d_f, it preserves uncertain features and mitigates boundary noise. Extensive evaluations on lung cancer, COVID-19, and pulmonary fibrosis datasets confirm FuzzySR's superior segmentation accuracy on LRCT, surpassing even high-resolution baselines. By uniting fuzzy-logic-driven uncertainty handling with SR-based resolution enhancement, FuzzySR effectively bridges the gap for robust airway delineation from LRCT data.

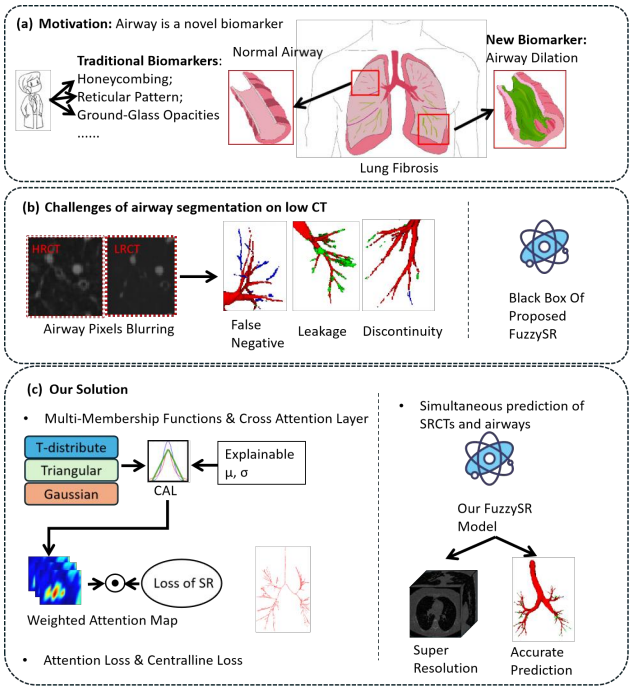

Figure 1: Figure 1: Introduction. (a) Motivation: need a new biomarker. (b) LRCT can result in false negatives, leakage, and discontinuities. (c) Our solution proposes a parallel network for super-resolution and segmentation. μ, σ is learnable for uncertainty.

1 INTRODUCTION

Ensuring CT images are sufficiently accurate is crucial for reliable diagnoses, and medical segmentation inevitably suffers from low-resolution scans Li et al. [2019]. In settings such as coronary CT angiography (CCTA) or vascular and airway imaging, segmentation quality depends heavily on image resolution Nagayama et al. [2023a], Jurek et al. [2020a]. Below a certain threshold, small but clinically significant structures become difficult to delineate, highlighting the need for methods that enhance resolution *and* improve segmentation performance. Airway dilation serves as a biomarker for diagnosing and prognosticating pulmonary fibrosis Nan et al. [2024a], Das et al. [2020b] (Figure 1(a)), calling for an automated approach under noisy, low-resolution conditions. Fuzzy set theory helps address intraclass heterogeneity and interclass homogeneity in LRCT, aiding radiologists' decision-making. Yet many LRCT scans are complicated by noise, making sequential super-resolution (SR) and segmentation both time-consuming and suboptimal Georgescu et al. [2020],

de Farias [2023], Nagayama et al. [2023b], Jurek et al. [2020b], Lu et al. [2017b]. Airway segmentation, especially near terminal bronchi, also remains challenging Singh et al. [2020], and while fuzzy sets are adept at managing uncertain boundaries Nan et al. [2024a], Das et al. [2020a], Huang et al. [2021], Ding et al. [2021], combining fuzzy logic with SR for airway segmentation remains underexplored (See Figure 1(b)). For example, Sale's and Hülleremeier's Hüllermeier et al. [2022], Sale et al. [2023] statistical-uncertainty models provide confidence intervals for the final mask and can estimate the mis-segmentation probability of individual branches. However, airway segmentation usually suffers from blurred boundaries and very small, partly invisible branches—limitations rooted in morphology and subjective ambiguity. In addition, the "black-box" nature of deep learning architectures remains an ongoing concern.

To address these challenges, we propose a parallel network called **FuzzySR**, which merges fuzzy logic with super-resolution and segmentation to enhance interpretability and reduce uncertainty. Rather than proceeding sequentially, FuzzySR performs super-resolution and airway segmentation concurrently, allowing outputs from the super-resolution decoder to directly benefit the segmentation process. We incorporate fuzzy attention to tackle borderline or noisy regions, where binary logic often omits fine details Jain and Sharma [2020]. Specifically, *triangular* and *t-distribution* membership functions enable channel-specific fuzzy attention, mitigating feature heterogeneity. By learning parameters (μ, σ), the network adapts to sub-branches that significantly vary yet still belong to the airway, while heavier tails in the t-distribution or localized windows in the triangular function refine airway versus vascular distinctions. Cross-attention further compares membership values adaptively, avoiding pointwise minima and thus more effectively capturing subtle intensity variations in LRCT (See Figure 1(c)).

Our comparative experiments yielded four critical findings:

a) Performing super-resolution and segmentation synchronously is more efficient and yields better outcomes than separate steps.

b) Including fuzzy attention led to significantly improved results compared to networks without fuzzy logic.

c) Our proposed FuzzySR achieved performance on LRCT comparable to HRCT-based UNet and FuzzyAttentionUnet, with no significant differences (p=0.89 and p=0.33).

d) Explainability: Visualizing μ, σ over training epochs and attention maps shows some channels becoming more sensitive to airway features.

2 METHODOLOGY

Joint multi-task learning in medical imaging typically outperforms sequential pipelines Amyar et al. [2020a], Lu et al. [2017a], Amyar et al. [2020b]. Accordingly, we adopt a Y-shaped network with a shared encoder and two parallel decoders—one for super-resolution (SR), one for segmentation—linked by a fuzzy-gating layer that forwards only SR features useful for segmentation. This design (i) provides mutual regularisation, with SR safeguarding pixel-level detail while segmentation enforces structural semantics; (ii) reduced error propagation - Unlike a sequential pipeline, the segmentation branch in our parallel framework is not solely conditioned on the SR output, thereby preventing the amplification of reconstruction errors. (iii) enables true end-to-end optimisation, letting the segmentation loss directly shape encoder filters; (iv) permits real-time cross-reinforcement between fine-detail and semantic cues that a serial SR→Seg cascade cannot exploit; and (v) handles boundary ambiguity through fuzzy weighting. These considerations furnish both the theoretical and empirical basis for expecting the parallel scheme to surpass sequential SR + Seg approaches.

2.1 SUPER RESOLUTION

Inspired by Luo's work Luo et al. [2023], our FuzzySR employs a super-resolution network resembling a plain CNN, which outperforms variants like Auto Encoder and UNet that utilize up/down sampling. Although UNet's skip connections slightly improve detail reuse, they still underperform compared to the plain CNN. In our design, eight convolutional layers are followed by LeakyReLU activation, and two fuzzy attention layers compute feature-weight maps for the 3rd and 4th encoder blocks, as well as the 1st and 2nd decoder blocks. Specifically, *att4* is used within the super-resolution model, while *att5* feeds information to the segmentation decoder. Figure 2 (left) outlines this architecture, and Table 3 illustrates how fuzzy approaches further enhance super-resolution performance.

2.2 PROPOSED FUZZYSR

The fuzzy attention layers form the core of FuzzySR, promoting interaction between super-resolution (SR) and segmentation. The SR decoder's high-resolution output flows to the segmentation decoder, refining segmentation accuracy, while the segmentation model's attention map influences the MSE loss in SR. This bidirectional exchange fosters improved predictions of both SR images and airways. An attention gate suppresses irrelevant activations, and because different channels capture different features Oktay et al. [1804], the map must be channel-specific. Unlike conventional attention gates reliant on sigmoid thresholds Nan et al. [2024a], the fuzzy layer leverages T-distribution, Triangular, and Gaussian membership functions, endowing mean and

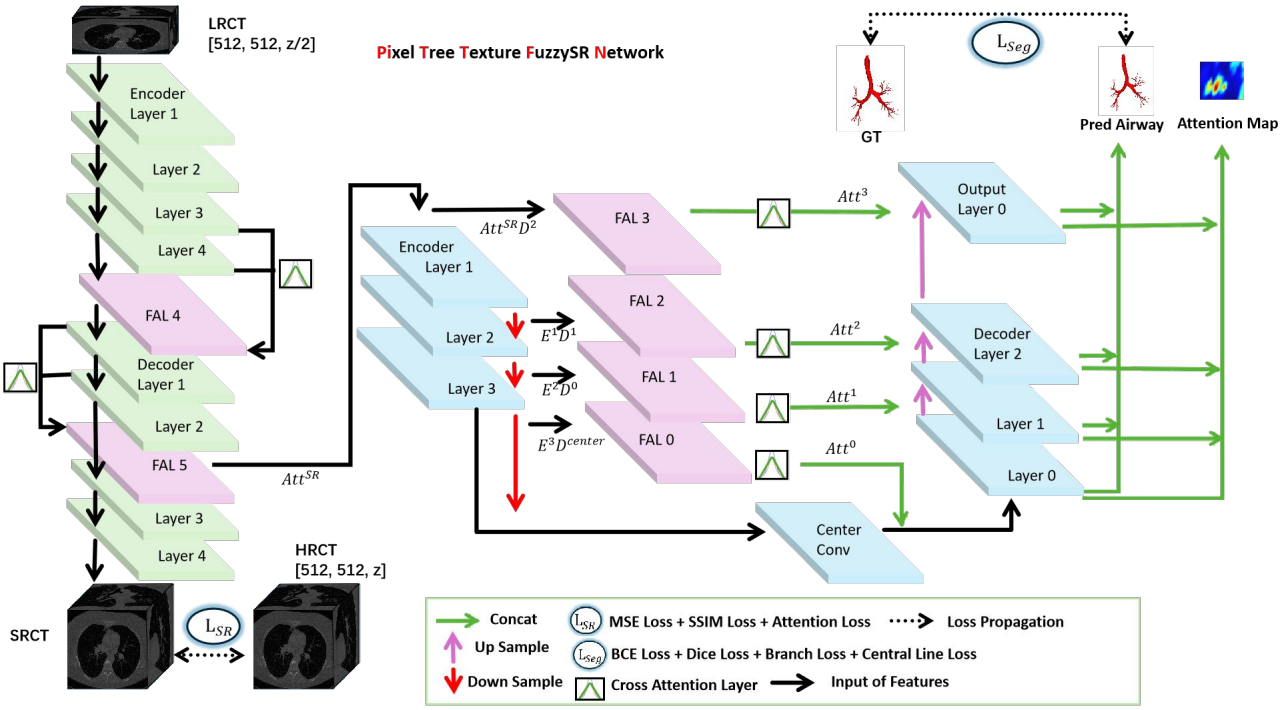

Figure 2: Network Structure of FuzzySR. *LR is low-resolution CT. HR is high-resolution CT. IN is InstanceNorm3d. $*E^i$ means the output of encoder. $*D^i$means the output of decoder. "512" denotes the resolution in the x and y axes, whereas "z" specifies the resolution along the z-axis in CT. FAL is Fuzzy Attention Layer.

variance with learnable parameters to address LRCT noise, heterogeneous features, and uncertainty.

To achieve parallel SR and airway segmentation within a single large network (FuzzySR), the third and fourth encoder layers for SR are processed by fuzzy attention layers and feed into the SR decoder. In turn, the first and second SR decoder layers pass through fuzzy attention for use as initial features and decoder inputs in the airway segmentation network. These inputs undergo typical UNet-like operations (convolution, up/downsampling). LRCT is provided to FuzzySR, which outputs a super-resolved CT (SRCT) and a predicted airway (Figure 2). In FuzzySR, the segmentation network fuses outputs from four decoding layers, iteratively computing each layer's loss against the ground truth to extract richer details from diverse convolutional kernels.

Fuzzy Attention Layer. In the segmentation network, the fuzzy attention layer is applied along the pathway of the skip connections, taking as input the spatial output features from both the encoder and decoder. These two input feature vectors are first passed through a 3-D convolution layer of size 1×1×1, instance normalization, and LeakyReLU for feature reconstruction. This is followed by voxel-level addition operations for information fusion, then another LeakyReLU. Subsequently, the feature representations are fed into the cross attention layer(CAL) combined with fuzzy logic for comparison, generating voxel-level attention maps (Figure 3).

Cross Attention Layer. After the final LeakyReLU in the Fuzzy Attention Layer, a 3D $1 \times 1 \times 1$ convolution and instance normalization produce a multi-channel attention map, which serves as input to the Cross Attention Layer (CAL) Das et al. [2020b] (Figure 4). This map is fed through membership functions with trainable parameters (df, σ, μ) to determine channel-specific feature importance. Query and Key feature representations are flattened (height, width, depth) and segmented into chunks (defined by chunk size) for the Cross Attention Gate (CAG) (Figure 4). After cosine similarity and softmax, the weighted Query and Key chunks are reconstructed into a weighted feature representation. Fuzzy logic then selects critical features by combining these weighted representations with encoder feature maps, producing an uncertainty map that reflects overall ambiguity.

Fuzzy Membership Functions. In fuzzy systems, Membership Functions are utilized to delineate the fuzziness of pixels, assigning a degree of membership (ranging between 0 and 1) to each element, thereby facilitating fuzzy inference and managing the complexity and uncertainty inherent in real-world scenariosZheng et al. [2021]. Membership Functions can be implemented through various methodologies. For instance, the triangular membership functionAzam et al. [2021] is defined by three parameters (a, b, c), forming a triangle where b represents the peak's position, and a and c are the base points.

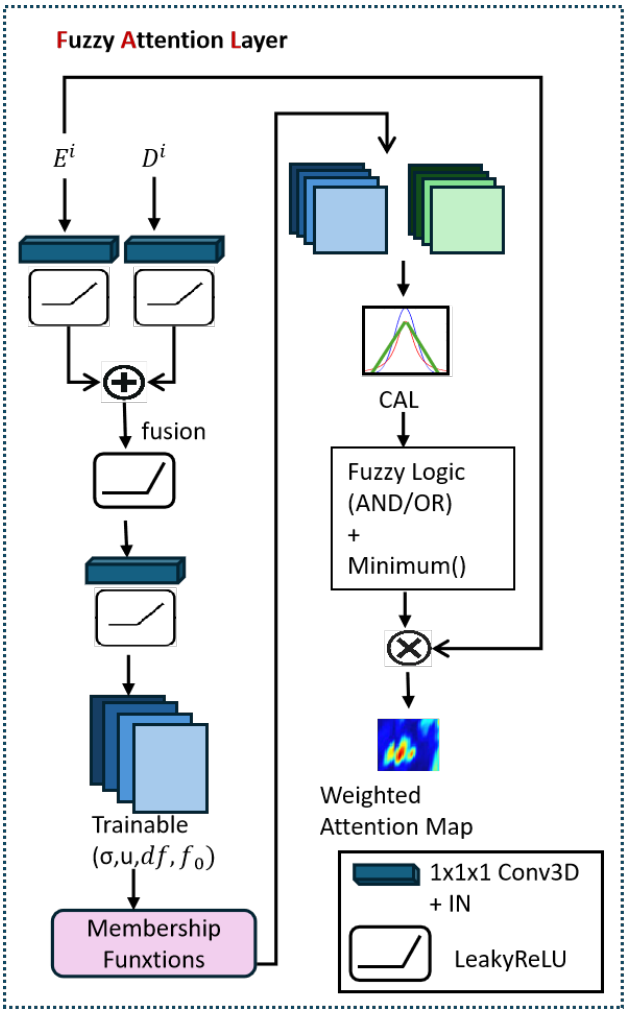

Figure 3: Fuzzy Attention Layer. *E^i means the output of encoder. *D^imeans the output of decoder. The parameters used for the membership functions, such as μ, σ and df are trainable. T-distribute,Triangular and Gaussian are fuzzy membership functions. CAL is Cross Attention Layer.

This function is straightforward and computationally efficient, making it suitable for applications demanding high computational performance:

$$\text{Triangular}(x) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right). \quad (1)$$

where, x represents the fusion generated by the outputs of encoder and decoder. a is the left endpoint, b is the peak point, c is the right endpoint. Where a = μ - σ, b = μ, c = μ + σ. μ is the mean, σ is the standard deviation. When σ is small and μ remains unchanged, the triangular shape becomes more "pointed," assigning higher membership values only near μ. Conversely, if σ is larger, the function grants higher membership across a broader range.

The t-distribution membership functionAli et al. [2023], a type of probability distribution function, resembles the standard normal distribution but with "heavier" tails. It is typically defined by the degrees of freedom parameter. The t-distribution is suitable for situations with smaller sample sizes, especially when estimating standard deviations. In fuzzy logic, it can be employed to better handle data uncertainty and skewed distributions. In Equation (2), ν is the degrees of freedom:

$$\text{T-distribution}(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}. \quad (2)$$

Concurrently, fuzzy logic operations (such as AND, OR) define the methodology for amalgamating the membership degrees of two or more fuzzy sets to derive a new membership degree. In the process of fuzzy logic reasoning, it is imperative to initially delineate the fuzzy sets and the membership degrees of elements via fuzzy membership functions. Subsequently, these membership degrees are combined through the application of fuzzy logic operations (e.g., AND, OR), thereby facilitating fuzzy inference and decision-makingKlir and Yuan [1996]. The outcomes of fuzzy logic operations (i.e., the new membership degrees) are then reflected upon the fuzzy membership functions. The higher the membership value, the greater the probability of occurrence. Ultimately, the precise output is derived through the process of defuzzificationGao et al. [2021].

Cross Attention for membership functions Cross Attention forms a similarity matrix (score_chunk) between query_chunk and key_chunk, leveraging global contextual information for feature fusion instead of purely local operations like element-wise minima Huang et al. [2019]. Chunk_size controls feature partitioning granularity in 3D or high-dimensional data, progressively accumulating local features into global representations. This multi-scale approach captures object boundaries and fine details more effectively than single-scale element-wise minima Lin et al. [2022], and dynamically balances uneven or noisy inputs via adaptive weight allocation Pan et al. [2022]. In the membership function processing, Q and K are query/key feature tensors, while Q_{flat} and K_{flat} denote flattened forms. The parameter (chunk_size) specifies how many channels C are processed in each chunk. Spatial size is $D \times H \times W$. Query/key chunks undergo cosine similarity S_i and softmax normalization. Weighted results W_q and W_k merge into the global tensor, and their element-wise minimum yields the final attention map, where smaller values indicate greater distance from the fuzzy center.

Attention Loss Functions

$$\mathcal{L}_{\text{att}} = \lambda \cdot \frac{1}{C}\sum_{i=1}^{C}\frac{1}{N}\sum_{n=1}^{N}\mathbf{A}_i^{(n)} \odot \left(\mathbf{P}_{\text{hr}}^{(n)} - \mathbf{H}^{(n)}\right)^2. \quad (3)$$

where The values $\mathbf{P}_j$ and $\mathbf{T}_j$ represent the predicted and target values at pixel j, where the intersection is calculated as $\sum_j \mathbf{P}_j^{(n)} \cdot \mathbf{T}_j^{(n)}$, the denominator is obtained by summing

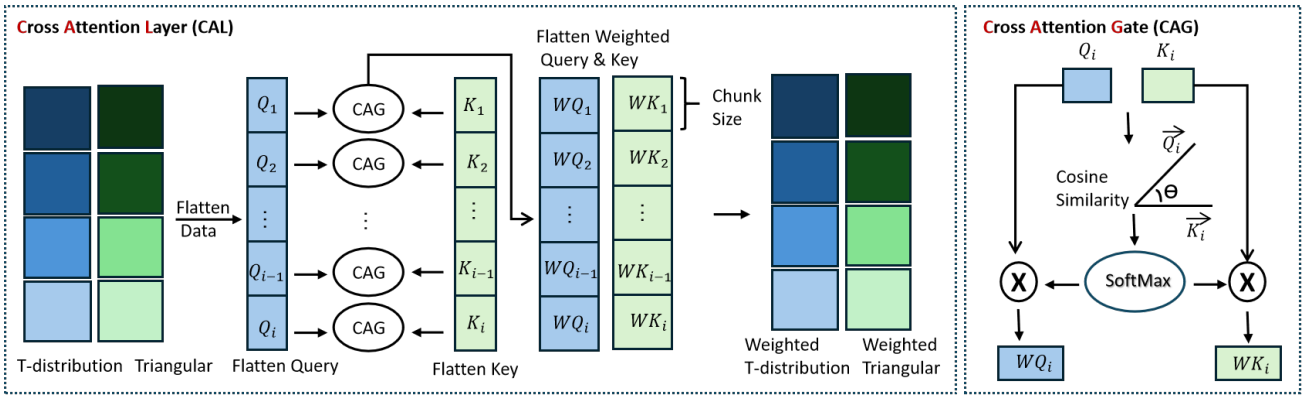

Figure 4: Cross Attention Layer & Cross Attention Gate (CAG). Q_i and K_i are multiple chunk size of flatten query and key arries. WQ_i and WK_i are weighted Q_i and K_i.

the target values $\sum_j \mathbf{T}_j^{(n)}$, and a smoothing term ϵ is added to avoid division by zero and improve numerical stability. The weighted sum of the attention map channels i (total C channels) is calculated, where each channel's weight map $\mathbf{A}_i$ is element-wise multiplied by the mean squared error $(\mathbf{P}_{\text{hr}} - \mathbf{H})^2$ using $\odot$, followed by batch averaging over N samples, and scaled by a fixed weight coefficient λ.

3 RESULTS AND DISCUSSION

3.1 DATASETS

We used 140 instances by combining EXACT'09 Lo et al. [2012], LIDC-IDRI Armato III et al. [2011], and an in-house 3D HRCT dataset. We split the entire dataset into training (90), validation (25), and testing (25), each containing original images and validated ground truths. As a preprocessing step, data were cropped into $[128, 96, 64]$ patches centered on pulmonary airways. The final 50 in-house cases were reserved for validation and testing to prevent overlap, ensuring distinct training/validation and test samples.

(1) Public binary dataset. The training set included three-dimensional CT scan images depicting typical lung structures, each accompanied by accurate ground truth.

(2) Our in-house dataset. The validation and test sets contained three-dimensional CT scan images of patients diagnosed with COVID-19 and pulmonary fibrosis, specifically for the lung airway segmentation task, effectively reducing the risk of model overfitting. Fibrosis data is now avaliable at MICCAI AIIB23 Nan et al. [2024b]. Meanwhile, this study has ethical approval from the Sydney Local Health District (protocol no. X14–0264) and the local Ethics Committee (code 934/2021/OSS/AOUPR -11.01.2022) at the University Hospital of Parma (UHP).

All models operate on input patches of size [1,1,128,96,144]. The proposed Parallel FuzzySR yields lower MACs (333.46 GMac) than AttentionUNet (364.68 GMac), and its parameter count (15.52 M) remains comparable to other high-performance networks. The inference time is approximately 1.08 seconds, which is still feasible for 3D medical datasets. Although this results in a slight increase in MACs and inference time, it achieves higher accuracy overall.

3.2 QUANTITATIVE AND QUALITATIVE FINDINGS

Comparison between methods with and without the fuzzy attention mechanism. In Table 1, FuzzyAttentionUnet using LRCT surpasses UNet in IoU, DSC, and Precision (0.7972, 0.8787, 0.8942). On HRCT, FuzzyAttentionUnet similarly outperforms UNet and AttentionUNet in IoU, DSC, and Precision (0.8243, 0.9032, 0.9027), while AttentionUNet attains higher TD/BD (0.9071/0.8761). For sequential setups, FuzzyAttentionUnet + SR yields higher IoU/DSC/Precision (0.7934, 0.8834, 0.8745) than AttentionUNet + SR but scores lower on TD/BD (0.8701/0.8963 vs. 0.8457/0.8587). In the parallel approach, FuzzySR with Gaussian MF achieves better IoU, DSC, Precision, and BD (0.8057, 0.8918, 0.8966, 0.8817) compared with AttentionUNet + SR, albeit slightly lower TD (0.9074/0.9136). Figure 5 shows that among the six baseline models (LRCT and HRCT), FuzzyUnet achieves the best IoU/DSC/Precision, with minimal false negatives and the capacity to predict small unannotated branches, while AttentionUnet achieves higher TD/BD but lower IoU/DSC/Precision. Although this suggests strong terminal-branch prediction, some outputs may reflect tiny branches missed by experts rather than leakage. Comparing both serial and parallel AttentionUnet + SR with FuzzyAttentionUnet + SR confirms that fuzzy-enabled models exhibit fewer false negatives, reduced leakage, and improved detection of small terminal branches.

The comparison between parallel and sequential multi-model networks. In comparing sequential vs. parallel AttentionUNet + SR, the parallel setup shows only minor dif-

Table 1: Comparative Analysis of Model Performance.

Our Models	Evaluation Metrics				
	IoU	DSC	TD	BD	Precision
UNet (Use LRCT)	$0.7815 \pm 0.067^*$	$0.8607 \pm 0.042^*$	$0.8436 \pm 0.076^*$	$0.8036 \pm 0.100^*$	$0.8587 \pm 0.048^*$
UNet (Use HRCT)	0.8162 ± 0.056	0.8978 ± 0.034	$0.8709 \pm 0.067^*$	$0.8269 \pm 0.095^*$	$0.8751 \pm 0.050^*$
AttentionUnet (Use LRCT)	$0.7914 \pm 0.050^*$	$0.8827 \pm 0.031^*$	$0.8744 \pm 0.068^*$	$0.8259 \pm 0.090^*$	$0.8697 \pm 0.042^*$
AttentionUnet (Use HRCT)Islam et al. [2020]	$0.7746 \pm 0.051^*$	$0.8720 \pm 0.033^*$	$0.9071 \pm 0.054^*$	$0.8761 \pm 0.074^*$	$0.8461 \pm 0.057^*$
FuzzyAttentionUnet (Use LRCT)	$0.7972 \pm 0.069^*$	$0.8787 \pm 0.069*$	$0.8329 \pm 0.095^*$	$0.7828 \pm 0.123^*$	$0.8942 \pm 0.063^*$
FuzzyAttentionUnet (Use HRCT)Nan et al. [2024a]	0.8243 ± 0.062	0.9032 ± 0.039	$0.8566 \pm 0.086^*$	$0.8392 \pm 0.111^*$	0.9027 ± 0.049
AttentionUNet + SR (sequence) (Use LRCT)	$0.7876 \pm 0.052*$	$0.8802 \pm 0.033*$	$0.8963 \pm 0.055*$	$0.8587 \pm 0.080*$	$0.8643 \pm 0.049*$
FuzzyAttentionUnet + SR (sequence) (Use LRCT)	$0.7934 \pm 0.064*$	$0.8834 \pm 0.040*$	$0.8701 \pm 0.066*$	$0.8457 \pm 0.090*$	$0.8745 \pm 0.062*$
AttentionUNet + SR (parallel) (Use LRCT)	$0.7866 \pm 0.045*$	$0.8799 \pm 0.028*$	$0.9136 \pm 0.046*$	$0.8808 \pm 0.067*$	$0.8582 \pm 0.036*$
FuzzySR (parallel) t-dis + triangular (Use LRCT)	0.8174 ± 0.043	0.8989 ± 0.026	0.9014 ± 0.056	0.8663 ± 0.080	0.9155 ± 0.032

All values are denoted by the mean value ± standard deviation. The red font indicates the best results and the bold font indicates our best proposed FuzzySR with best combination of membership functions. TD/BD: Tree/Branch detected ratio. "" represents statistical significance (with Wilcoxon Signed-Rank Test $p - value < 0.05$) compared with our proposed best FuzzySR t-dis + triangular.

Table 2: Comparative Analysis of Model Performance With Different Fuzzy Logic (AND/OR).

Our Models (Use HRCT)	Evaluation Metrics				
	IoU	DSC	TD	BD	Precision
Fuzzy_Unet (AND)	0.798 ± 0.072	0.886 ± 0.047	0.794 ± 0.141	0.735 ± 0.179	0.880 ± 0.041
Fuzzy_Unet (OR)	0.824 ± 0.047	0.903 ± 0.035	0.920 ± 0.067	0.879 ± 0.095	0.903 ± 0.049

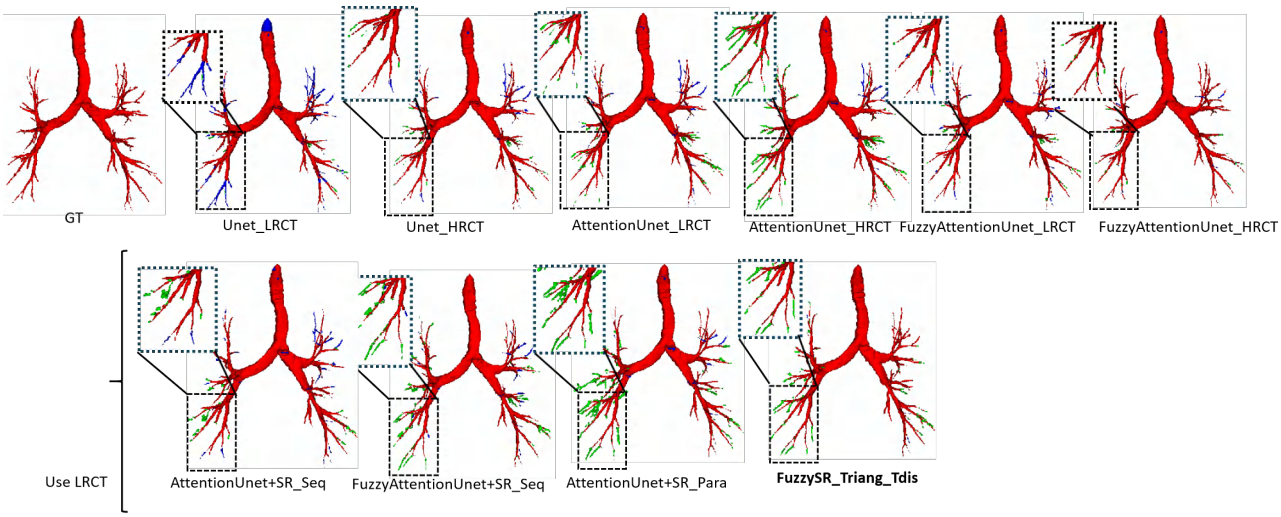

Figure 5: Visualization Results of all baseline models and proposed models. *Unet, AteentionUnet and FuzzyAttentionUnet are comparable baseline models. Green areas indicate false positive (FP), blue areas indicate false negative(FN). The blod font model is the proposed best FuzzySR model.

Table 3: Super-Resolution Network Comparison.

Our Models (Use LRCT)	Evaluation Metrics		
	SSIM	PSNR	RMSE
UNet	0.967 ± 0.026	74.97 ± 2.52	0.05 ± 0.02
AutoEncoder	0.956 ± 0.030	75.39 ± 3.03	0.05 ± 0.02
Plain CNN (no fuzzy)	0.953 ± 0.030	73.51 ± 2.25	0.06 ± 0.02
Plain CNN (with fuzzy)	0.971 ± 0.024	78.55 ± 3.84	0.002 ± 0.02
SRCNN Dong et al. [2014]	0.876 ± 0.021	32.26 ± 2.99	6.22 ± 0.19
ESPCN Shi et al. [2016]	0.872 ± 0.046	32.06 ± 2.93	6.36 ± 0.55
VDSR Kim et al. [2016]	0.886 ± 0.054	33.05 ± 3.22	5.68 ± 0.36
IMDN Hui et al. [2019]	0.884 ± 0.030	34.04 ± 1.98	5.06 ± 1.03
PAN Zhao et al. [2020]	0.885 ± 0.048	33.93 ± 2.76	5.13 ± 0.60
DRIDSR Chen et al. [2022]	0.887 ± 0.037	34.47 ± 3.11	4.82 ± 0.45
Proposed FuzzySR(with cl, MF)	0.924 ± 0.045	78.15 ± 4.83	0.04 ± 0.03

*MF is membership function. *cl is central line.

ferences in IoU, DSC, and Precision (0.001, 0.0003, 0.0061) but achieves higher TD and BD (+0.0173, +0.0221). With comparable leakage, the parallel AttentionUnet + SR yields fewer false negatives and better terminal airway detection (Figure 5). Meanwhile, the sequential FuzzyAttentionUnet + SR underperforms the proposed FuzzySR, which applies fuzzy logic (OR) and the T-Distribution + Triangular membership functions via cross-attention. This T-Distribution + Triangular setup attains near-zero leakage and minimal false negatives, scoring 0.8174 (IoU), 0.8989 (DSC), 0.9014 (TD), 0.8663 (BD), and 0.9155 (Precision) (Table 1), while preserving robust predictive capabilities for terminal airways.

Comparison between the use of LRCT and HRCT. The UNet using HRCT outperforms its LRCT counterpart on all five metrics (IoU, DSC, TD, BD, Precision) Li et al. [2019], Nagayama et al. [2023a], Jurek et al. [2020a], while AttentionUNet using LRCT surprisingly surpasses AttentionUNet with HRCT, showing minimal leakage and better terminal airway predictions on HRCT. FuzzyAttentionUnet further improves upon LRCT results by employing HRCT, achieving higher IoU, DSC, TD, BD, and Precision Nan et al. [2024a], Das et al. [2020b] (Figure 5). Notably, our FuzzySR with LRCT and T-Distribution + Triangular surpasses the classic UNet with HRCT, scoring 0.8174 (IoU), 0.8989 (DSC), 0.9014 (TD), 0.8663 (BD), and 0.9155 (Precision). Compared with FuzzyAttentionUnet using HRCT, FuzzySR achieves similar IoU, DSC, TD, BD, and even higher Precision, despite slightly lower IoU, DSC, and Precision stemming from more terminal branch predictions (some false positives). Table 1 reveals no significant difference (p=0.33 or p=0.89) between FuzzySR and other HRCT-based models, indicating their performance is statistically comparable. Although FuzzyAttentionUnet (HRCT) reports slightly better metrics, it can underperform on certain individual cases, suggesting weaker generalization.

Comparison of Fuzzy Logic. In determining the trachea's importance across channels and performing fuzzy logic operations, the AND (intersection) operator marks a voxel as airway only if multiple features concur, whereas the

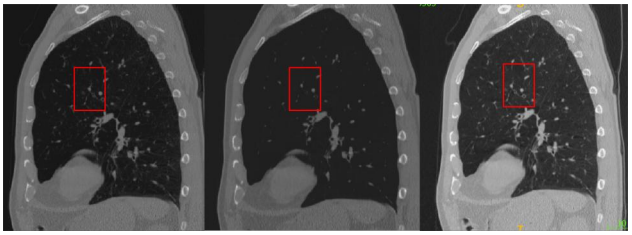

Figure 6: Comparison between original HRCT (Ground Truth) images, LRCT (Low-Resolution CT input images), and SRCT (Super-Resolution CT output images).

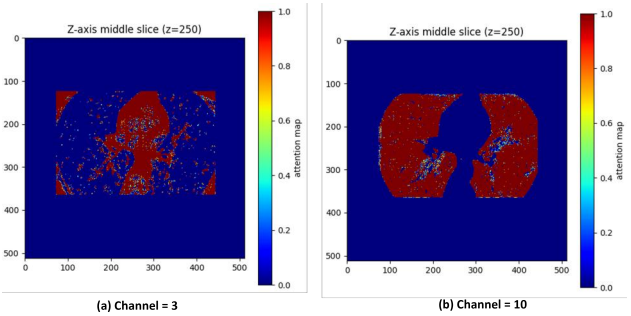

Figure 7: Visualization of attention maps for different feature channel indices. (a) Feature channel index = 3, (b) Feature channel index = 10.

OR (union) operator requires at least one feature to indicate airway membership. Although the AND operator enforces stricter segmentation, it may under-segment terminal bronchi, while the OR operator often predicts more terminal branches but admits some false positives. Table 2 confirms that AND logic causes a larger decline in TD/BD metrics (by 0.162 and 0.144) compared to OR logic, whereas IoU, DSC, and Precision drop by 0.026, 0.017, and 0.004, respectively. This finding aligns with the theoretical trade-off and justifies using OR logic for all airway prediction models Lu et al. [2017b].

Comparison of the super resolution network. In line with Section I.A, simpler architectures without up/downsampling (e.g., Plain CNN) can deliver better results without extra computational load. Table 3 compares PlainCNN with and without the fuzzy algorithm, revealing that PlainCNN with fuzziness yields superior SSIM, PSNR, and RMSE (0.971, 78.55, and 0.002) than UNet and AutoEncoder (0.967/0.956, 74.97/75.39, and 0.05/0.05). Also, PlainCNN (with fuzzy) also superior the metrics than PlainCNN (without Fuzzy) in SSIM, PSNR and RMSE (0.953, 73.51, 0.06). In FuzzySR, the super-resolution CT output alongside airway predictions outperforms most SR methods Dong et al. [2014], Shi et al. [2016], Kim et al. [2016], Hui et al. [2019], Zhao et al. [2020], Chen et al. [2022], achieving SSIM, PSNR, and RMSE of 0.924 ± 0.045, 78.15 ± 4.83, and 0.038 ± 0.028

Table 4: State-Of-The-Art Tansformer Based Models Performed On HRCT Compares To Proposed FuzzySR Performed On LRCT.

Models	Evaluation Metrics				
	IoU	DSC	TD	BD	Precision
FuzzyAttentionUnet (Use HRCT)Nan et al. [2024a]	0.8243 ± 0.062	0.9032 ± 0.039	0.8566 ± 0.086*	0.8392 ± 0.111*	0.9027 ± 0.049
Transformer-based 3D U-Net (HRCT)Wu et al. [2023]	0.7106 ± 0.124*	0.8232 ± 0.098*	0.8454 ± 0.067*	0.7975 ± 0.100*	0.8227 ± 0.156*
Segformer (HRCT)Perera et al. [2024]	0.7350 ± 0.066*	0.8455 ± 0.045*	0.7220 ± 0.101*	0.6244 ± 0.124*	0.8337 ± 0.043*
SwinUnet (HRCT)Cai et al. [2023]	0.8134 ± 0.048	0.8963 ± 0.029	0.7955 ± 0.097*	0.7284 ± 0.119*	0.9166 ± 0.030
Proposed FuzzySR (Use LRCT)	0.8174 ± 0.043	0.8989 ± 0.026	0.9014 ± 0.056	0.8663 ± 0.080	0.9155 ± 0.032

All values are denoted by the mean value ± standard deviation. The red font indicates the best results and the bold font indicates our proposed FuzzySR. "" represents statistical significance (with Wilcoxon Signed-Rank Test $p - value < 0.05$) compared with our proposed best FuzzySR t-dis + triangular.

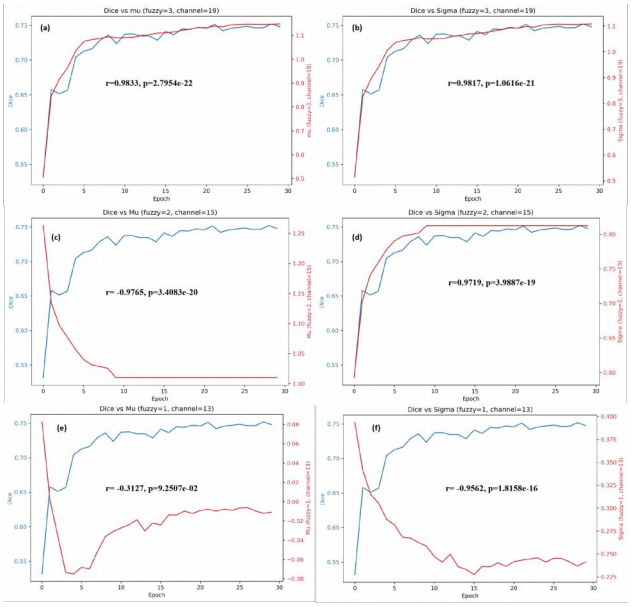

Figure 8: Evolution of the learnable μ (mu) and σ (sigma) across different feature channels, as training epochs increase and Dice scores vary. fuzzy in this figure means different fuzzy subsets. r,p were calculated by Pearson correlation coefficient.

(Table 3). As shown in Figure 6, the SRCT predicted by FuzzySR exhibits finer textures and is visually closer to HRCT than LRCT.

SOTA Transformer-based Model Compares With Our Proposed FuzzySR. Due to the proposed FuzzySR employs a t-distribution and triangular MF for cross-attention fusion, we not only compared it against the baseline AttentionUnet but also against several state-of-the-art methods, including our benchmark model FuzzyAttentionUnet (using HRCT)Nan et al. [2024a], Transformer-based 3D U-Net (HRCT)Wu et al. [2023], Segformer (HRCT)Perera et al. [2024], and SwinUnet (HRCT)Cai et al. [2023]. Except for the benchmark model, the other three transformer-based methods—despite utilizing HRCT for training—performed worse than the proposed FuzzySR framework, which was trained solely on LRCT, across all five evaluation metrics (See Table 4). Although SwinUnet's performance is close, it remains slightly below that of the proposed FuzzySR in IoU (0.8134 vs. 0.8174) and DSC (0.8963 vs. 0.8989). Two metrics are notably lower than those of the proposed FuzzySR—TD (0.7955 vs. 0.9014) and BD (0.7284 vs. 0.8663). SwinUnet achieves a marginally higher Precision value, surpassing the proposed FuzzySR by only 0.0011.

Explainability of Our Proposed FuzzySR. In medical image segmentation, certain structures (e.g., pulmonary airways) often occupy a limited grayscale range but may face uncertainty from noise or pathologies. The membership function provides higher membership near μ and decreases it for more distant intensities. To validate learnable μ, σ in our FuzzySR for mitigating airway-segmentation uncertainty, we examined two feature channels—one (a) Channel $= 3$ targeting brighter intensities and another (b) Channel $= 10$ focusing on darker intensities (Figure 7), revealing distinct types of airway boundaries or noise regions. As shown in Figure 8, the network gradually shifts each channel's μ toward relevant intensity ranges (a, c, e), indicating stronger resilience to fuzziness and noise. Once μ, σ stabilize, the channel gains "confidence" in that intensity range: a high μ highlights brighter zones (a), while a low μ focuses on lower-intensity areas (c, e). A larger σ tolerates outliers (b, d), whereas a smaller σ narrows the focus, improving detail yet risking noise sensitivity (f). Pearson's correlation analysis reveals significant positive or negative links between Dice scores and μ, σ in certain channels ($p < 0.05$ or $p < 0.01$), confirming strong interpretability for both high- and lower-intensity regions.

In Figure 8, certain high–sensitivity channels appear to map to specific airway sub-structures. Because the pulmonary tree forms a hierarchical network whose branches vary in calibre and morphology, some channels specialise in detecting peripheral bronchioles, segmental bronchi, or the main bronchus. Abrupt shifts in a channel's learnable μ or σ thus mark strong responsiveness and exert a major influence on the final segmentation. Channels that focus on ultra–thin branches or high–contrast edges rapidly adapt their fuzzy functions during training and can markedly raise the Dice score. When the channel-wise (μ, σ) values cor-

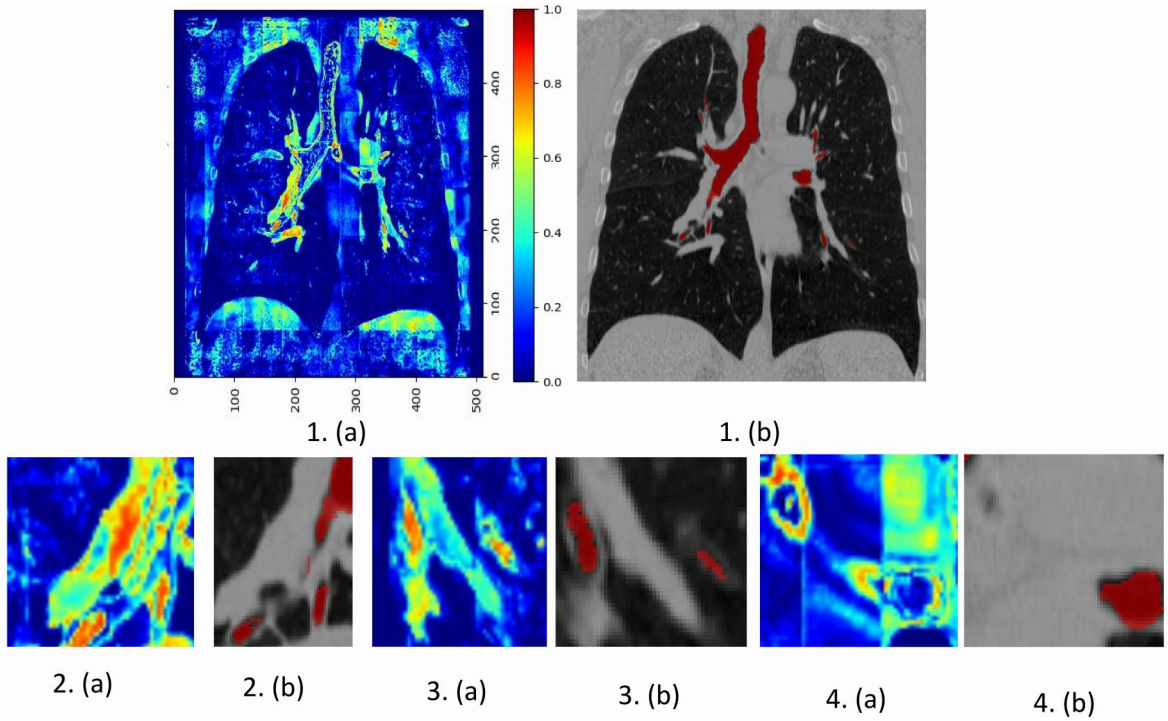

Figure 9: Comparison between the ground-truth airway annotation and the corresponding fuzzy membership values. 1(a) and 1(b) visualize the global airway fuzzy membership, where the color gradient from blue to red indicates membership values ranging from high to low (0–1). Meanwhile, 2(a)/2(b), 3(a)/3(b), and 4(a)/4(b) each display the zoomed-in view of local membership regions.

relate strongly with Dice (e.g., $r \approx 0.97$–0.98), the channel is likely performing critical *discrimination* or *selection*, thereby driving the overall improvement. Clinically, if a channel "locks onto" a distinctive morphology—such as tiny peripheral airways—and exhibits such a high performance correlation, it indicates refined differentiation of those structures or pathologies. Hence, these high–sensitivity channels capture regions that are otherwise difficult to delineate (e.g., very fine or distorted bronchi), implying enhanced discriminative power precisely where segmentation is most challenging and providing potentially valuable insights for diagnosis and intervention.

Figure 9 visualises global and local uncertainty maps: high-uncertainty regions (green) cluster along airway boundaries, indicating that the super-resolution branch—by reconstructing fine details—also contributes to boundary awareness instead of relying solely on the segmentation branch. In clinical airway assessment, one must evaluate not only airway location but also subtle deformities, stenoses, and other fine-scale changes. By retaining boundary uncertainty, fuzzy logic supports more precise downstream quantification—for example, detecting incipient luminal narrowing. A hard-threshold mask risks discarding the borderline voxels (e.g., a value of 0.49 forced to 0), thereby masking early pathology. Fuzzy membership maps, in contrast, allow graded updates, prevent abrupt structural jumps, and give clinicians a smoother, more reliable basis for local refinement and intervention planning.

In two separate models (super-resolution and segmentation), feature maps from the super-resolution decoder do not undergo fuzzy attention nor interact with the segmentation decoder. By contrast, our proposed FuzzySR embeds the fuzzy algorithm within the SR network and fuses intermediate four-layer features through a fuzzy attention gate, subsequently feeding them into the segmentation network (Fig. 2). We also integrate central line assistance, fuzzy logic (OR), central line loss, attention loss, cross-attention fusion, and triangular/t-distribution membership functions. As shown in Table 1 and Table 4, this combination achieves the best airway prediction accuracy. FuzzySR significantly outperforms sequential SR+segmentation models across all metrics ($p < 0.01$), while the single SR model (PlainCNN with fuzzy) surpasses baseline UNet and AutoEncoder in super-resolution precision. FuzzySR likewise outperforms many state-of-the-art SR approaches. Furthermore, fuzzy membership functions combined with logical operations (OR/AND/min) enable "multiple parallel judgments" across scales and channels, preserving high-likelihood branches and avoiding overly strict thresholds that may exclude relevant structures.

Acknowledgements

Guang Yang was supported in part by the ERC IMI (101005122), the H2020 (952172), the MRC (MC/PC/21013), the Royal Society (IEC\NSFC\211235), the NVIDIA Academic Hardware Grant Program, the *SABER* project supported by Boehringer Ingelheim Ltd, the NIHR Imperial Biomedical Research Centre (RDA01), the Wellcome Leap *Dynamic Resilience* programme (co-funded by Temasek Trust), UKRI guarantee funding for Horizon Europe MSCA Postdoctoral Fellowships (EP/Z002206/1), the UKRI MRC Research Grant, TFS Research Grants (MR/U506710/1), and the UKRI Future Leaders Fellowship (MR/V023799/1).

References

Wajid Ali, Miin-Shen Yang, Mehboob Ali, and Saif Ud-Din. Fuzzy model-based sparse clustering with multivariate t-mixtures. *Applied Artificial Intelligence*, 37(1):2169299, 2023.

Amine Amyar, Romain Modzelewski, Hua Li, and Su Ruan. Multi-task deep learning based ct imaging analysis for covid-19 pneumonia: Classification and segmentation. *Computers in biology and medicine*, 126:104037, 2020a.

Amine Amyar, Romain Modzelewski, Hua Li, and Su Ruan. Multi-task deep learning based ct imaging analysis for covid-19 pneumonia: Classification and segmentation. *Computers in biology and medicine*, 126:104037, 2020b.

Samuel G Armato III, Geoffrey McLennan, Luc Bidaut, Michael F McNitt-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I Henschke, Eric A Hoffman, et al. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Medical physics*, 38(2):915–931, 2011.

Muhammad Hamza Azam, Mohd Hilmi Hasan, Saima Hassan, and Said Jadid Abdulkadir. A novel approach to generate type-1 fuzzy triangular and trapezoidal membership functions to improve the classification accuracy. *Symmetry*, 13(10):1932, 2021.

Yimin Cai, Yuqing Long, Zhenggong Han, Mingkun Liu, Yuchen Zheng, Wei Yang, and Liming Chen. Swin unet3d: a three-dimensional medical image segmentation network combining vision transformer and convolution. *BMC medical informatics and decision making*, 23(1):33, 2023.

Yihan Chen, Qianying Zheng, and Jiansen Chen. Double paths network with residual information distillation for improving lung ct image super resolution. *Biomedical Signal Processing and Control*, 73:103412, 2022.

Rangan Das, Sagnik Sen, and Ujjwal Maulik. A survey on fuzzy deep neural networks. *ACM Computing Surveys (CSUR)*, 53(3):1–25, 2020a.

Rangan Das, Sagnik Sen, and Ujjwal Maulik. A survey on fuzzy deep neural networks. *ACM Comput. Surv.*, 53(3), may 2020b. ISSN 0360-0300. doi: 10.1145/3369798. URL `https://doi.org/10.1145/3369798`.

Erick Costa de Farias. Impact of gan-based lesion-focused medical image super-resolution on radiomic feature robustness. Master's thesis, Universidade NOVA de Lisboa (Portugal), 2023.

Weiping Ding, Mohamed Abdel-Basset, Hossam Hawash, and Witold Pedrycz. Multimodal infant brain segmentation by fuzzy-informed deep learning. *IEEE Transactions on Fuzzy Systems*, 30(4):1088–1101, 2021.

Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 184–199, Cham, 2014. Springer International Publishing.

Yuan Gao, Shiwei Ma, Jingjing Liu, Yanyan Liu, and Xianxia Zhang. Fusion of medical images based on salient features extraction by pso optimized fuzzy logic in nsst domain. *Biomedical Signal Processing and Control*, 69: 102852, 2021.

Mariana-Iuliana Georgescu, Radu Tudor Ionescu, and Nicolae Verga. Convolutional neural networks with intermediate loss for 3d super-resolution of ct and mri scans. *IEEE Access*, 8:49112–49124, 2020. doi: 10.1109/ACCESS.2020.2980266.

Kuan Huang, Yingtao Zhang, Heng-Da Cheng, Ping Xing, and Boyu Zhang. Semantic segmentation of breast ultrasound image with fuzzy deep learning network and breast anatomy constraints. *Neurocomputing*, 450:319–335, 2021.

Zilong Huang, Xinggang Wang, Lichao Huang, Chang Huang, Yunchao Wei, and Wenyu Liu. Ccnet: Criss-cross attention for semantic segmentation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 603–612, 2019.

Zheng Hui, Xinbo Gao, Yunchu Yang, and Xiumei Wang. Lightweight image super-resolution with information multi-distillation network. In *Proceedings of the 27th acm international conference on multimedia*, pages 2024–2032, 2019.

Eyke Hüllermeier, Sébastien Destercke, and Mohammad Hossein Shaker. Quantification of credal uncertainty in machine learning: A critical analysis and empirical

comparison. In *Uncertainty in Artificial Intelligence*, pages 548–557. PMLR, 2022.

Mobarakol Islam, VS Vibashan, V Jeya Maria Jose, Navodini Wijethilake, Uppal Utkarsh, and Hongliang Ren. Brain tumor segmentation and survival prediction using 3d attention unet. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I 5*, pages 262–272. Springer, 2020.

Arpit Jain and Abhinav Sharma. Membership function formulation methods for fuzzy logic systems: A comprehensive review. *Journal of Critical Reviews*, 7(19): 8717–8733, 2020.

Jakub Jurek, Marek Kociński, Andrzej Materka, Marcin Elgalal, and Agata Majos. Cnn-based superresolution reconstruction of 3d mr images using thick-slice scans. *Biocybernetics and Biomedical Engineering*, 40(1):111–125, 2020a. ISSN 0208-5216. doi: https://doi.org/10.1016/j.bbe.2019.10.003. URL https://www.sciencedirect.com/science/article/pii/S0208521619304711.

Jakub Jurek, Marek Kociński, Andrzej Materka, Marcin Elgalal, and Agata Majos. Cnn-based superresolution reconstruction of 3d mr images using thick-slice scans. *Biocybernetics and Biomedical Engineering*, 40(1):111–125, 2020b. ISSN 0208-5216. doi: https://doi.org/10.1016/j.bbe.2019.10.003. URL https://www.sciencedirect.com/science/article/pii/S0208521619304711.

Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1646–1654, 2016.

George J Klir and Bo Yuan. Fuzzy sets and fuzzy logic: theory and applications. *Possibility Theory versus Probab. Theory*, 32(2):207–208, 1996.

Yawei Li, Vagia Tsiminaki, Radu Timofte, Marc Pollefeys, and Luc Van Gool. 3d appearance super-resolution with deep learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

Hezheng Lin, Xing Cheng, Xiangyu Wu, and Dong Shen. Cat: Cross attention in vision transformer. In *2022 IEEE international conference on multimedia and expo (ICME)*, pages 1–6. IEEE, 2022.

Pechin Lo, Bram Van Ginneken, Joseph M Reinhardt, Tarunashree Yavarna, Pim A De Jong, Benjamin Irving, Catalin Fetita, Margarete Ortner, Rômulo Pinho, Jan Sijbers, et al. Extraction of airways from ct (exact'09). *IEEE Transactions on Medical Imaging*, 31(11):2093–2107, 2012.

Fang Lu, Fa Wu, Peijun Hu, Zhiyi Peng, and Dexing Kong. Automatic 3d liver location and segmentation via convolutional neural network and graph cut. *International journal of computer assisted radiology and surgery*, 12: 171–182, 2017a.

Ping Lu, Livia Barazzetti, Vimal Chandran, Kate Gavaghan, Stefan Weber, Nicolas Gerber, and Mauricio Reyes. Highly accurate facial nerve segmentation refinement from cbct/ct imaging using a super-resolution classification approach. *IEEE transactions on biomedical engineering*, 65(1):178–188, 2017b.

Weixun Luo, Xiaodan Xing, and Guang Yang. Is autoencoder truly applicable for 3d ct super-resolution? In *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, pages 1–4. IEEE, 2023.

Andrea Maggiordomo, Paolo Cignoni, and Marco Tarini. Texture defragmentation for photo-reconstructed 3d models. In *Computer Graphics Forum*, volume 40, pages 65–78. Wiley Online Library, 2021.

Yasunori Nagayama, Takafumi Emoto, Yuki Kato, Masafumi Kidoh, Seitaro Oda, Daisuke Sakabe, Yoshinori Funama, Takeshi Nakaura, Hidetaka Hayashi, Sentaro Takada, et al. Improving image quality with super-resolution deep-learning-based reconstruction in coronary ct angiography. *European Radiology*, 33(12):8488–8500, 2023a.

Yasunori Nagayama, Takafumi Emoto, Yuki Kato, Masafumi Kidoh, Seitaro Oda, Daisuke Sakabe, Yoshinori Funama, Takeshi Nakaura, Hidetaka Hayashi, Sentaro Takada, et al. Improving image quality with super-resolution deep-learning-based reconstruction in coronary ct angiography. *European Radiology*, 33(12):8488–8500, 2023b.

Yang Nan, Javier Del Ser, Zeyu Tang, Peng Tang, Xiaodan Xing, Yingying Fang, Francisco Herrera, Witold Pedrycz, Simon Walsh, and Guang Yang. Fuzzy attention neural network to tackle discontinuity in airway segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, 35(6):7391–7404, 2024a. doi: 10.1109/TNNLS.2023.3269223.

Yang Nan, Xiaodan Xing, Shiyi Wang, Zeyu Tang, Federico N Felder, Sheng Zhang, Roberta Eufrasia Ledda, Xiaoliu Ding, Ruiqi Yu, Weiping Liu, Feng Shi, Tianyang Sun, Zehong Cao, Minghui Zhang, Yun Gu, Hanxiao Zhang, Jian Gao, Pingyu Wang, Wen Tang, Pengxin Yu, Han Kang, Junqiang Chen, Xing Lu, Boyu

Zhang, Michail Mamalakis, Francesco Prinzi, Gianluca Carlini, Lisa Cuneo, Abhirup Banerjee, Zhaohu Xing, Lei Zhu, Zacharia Mesbah, Dhruv Jain, Tsiry Mayet, Hongyu Yuan, Qing Lyu, Abdul Qayyum, Moona Mazher, Athol Wells, Simon LF Walsh, and Guang Yang. Hunting imaging biomarkers in pulmonary fibrosis: Benchmarks of the aiib23 challenge. *Medical Image Analysis*, 97:103253, 2024b. ISSN 1361-8415. doi: https://doi.org/10.1016/j.media.2024.103253. URL https://www.sciencedirect.com/science/article/pii/S1361841524001786.

Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention u-net: Learning where to look for the pancreas. arxiv 2018. *arXiv preprint arXiv:1804.03999*, 1804.

Liangrui Pan, Hetian Wang, Lian Wang, Boya Ji, Mingting Liu, Mitchai Chongcheawchamnan, Jin Yuan, and Shaoliang Peng. Noise-reducing attention cross fusion learning transformer for histological image classification of osteosarcoma. *Biomedical Signal Processing and Control*, 77:103824, 2022.

Shehan Perera, Pouyan Navard, and Alper Yilmaz. Segformer3d: an efficient transformer for 3d medical image segmentation, 2024. URL https://arxiv.org/abs/2404.10156.

Yusuf Sale, Michele Caprio, and Eyke Höllermeier. Is the volume of a credal set a good measure for epistemic uncertainty? In *Uncertainty in Artificial Intelligence*, pages 1795–1804. PMLR, 2023.

Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883, 2016.

Satya P Singh, Lipo Wang, Sukrit Gupta, Haveesh Goli, Parasuraman Padmanabhan, and Balázs Gulyás. 3d deep learning on medical images: a review. *Sensors*, 20(18): 5097, 2020.

Unknown. Mimics innovation suite | medical image analysis software, June 4, 2024a. URL https://www.materialise.com/en/healthcare/.

Unknown. 3d slicer image computing platform | 3d slicer, June 4, 2024b. URL https://www.slicer.org/.

Unknown. Itk-snap home, June 4, 2024c. URL http://www.itksnap.org/pmwiki/pmwiki.php.

Yanan Wu, Shouliang Qi, Meihuan Wang, Shuiqing Zhao, Haowen Pang, Jiaxuan Xu, Long Bai, and Hongliang Ren. Transformer-based 3d u-net for pulmonary vessel segmentation and artery-vein separation from ct images. *Medical & Biological Engineering & Computing*, 61(10): 2649–2663, 2023.

Hengyuan Zhao, Xiangtao Kong, Jingwen He, Yu Qiao, and Chao Dong. Efficient image super-resolution using pixel attention. In *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*, pages 56–72. Springer, 2020.

Yuanhang Zheng, Zeshui Xu, and Xinxin Wang. The fusion of deep learning and fuzzy systems: A state-of-the-art survey. *IEEE Transactions on Fuzzy Systems*, 30(8):2783–2799, 2021.

A Parallel Network for LRCT Segmentation and Uncertainty Mitigation with Fuzzy Sets
(Supplementary Material)

Shiyi Wang[1] **Yang Nan**[1] **Xiaodan Xing**[1] **Yingying Fang**[1] **Simon LF Walsh**[2,3] **Guang Yang**[1,2,3,4]

[1]Bioengineering Department and Imperial- X., Imperial College London., UK
[2]National Heart and Lung In- stitute., Imperial College London., UK
[3]Cardiovascular Research Centre., Royal Brompton Hospital., UK
[4] Imaging Sciences Dept., King's College London., UK

A INTRODUCTION

Integration of Central Lines and Attention Loss. (1) Central-Line-Based Loss. We incorporate manually extracted and expert-modified central lines of the airways into the loss function, thus improving the detection around branching areas. This step addresses the difficulty in capturing fine, distal airways and ensures that the parallel network remains sensitive to small structural details. (2) Attention Loss. We additionally propose an Attention Loss, which uses the fuzzy attention maps generated by the segmentation branch to guide the super-resolution branch, ensuring that terminal airway regions receive higher emphasis. This design further facilitates the capture of subtle airway structures in low-resolution scans and reinforces the synergy between super-resolution and segmentation.

B ADDITIONAL METHODOLOGY

B.1 CENTRAL LINE EXTRACTION

Certain Python packages offer functions frequently used for segmenting tubular structures like airways and vessels. For instance, one function from the Skimage package processes a three-dimensional Boolean array (a pre-processed binary image) and produces another Boolean array of identical dimensions, indicating the skeleton of the original shape. This skeletonization process iteratively strips away outer pixels until only the core structure remains. However, its output can be inaccurate, often retaining redundant looping centrelines that diminish segmentation precision. Consequently, we manually extract centrelines during data preprocessing and then rely on expert corrections to ensure more accurate centreline segmentation.

After medical experts annotated the tracheal central line, we incorporated it as a corrective measure in the training process. Comparative analyses demonstrated pronounced improvements in quantitative metrics (e.g., Dice Similarity Coefficient, Intersection over Union, and precision) once the central line was included. Meanwhile, removing falsely segmented regions (false positives) led to minor decreases in the branch score (BD) and tree detected ratio (TD). Nonetheless, these findings underscore the clinical importance of integrating the central line, which enhances the model's predictive accuracy and delivers better diagnostic support for airway-related pathologies—consistent with the primary objective of our study. **Refinement of Pulmonary Airway Central Lines:** The skeletonize_3d() function in the skimage.morphology module of the SciKit-Image Python library is frequently used to obtain object skeletons in 3D images. However, since it is designed for general 3D image processing rather than tracheal anatomy, it may yield tracheal centerlines lacking the smoothness and branching clarity needed for accurate analyses. In contrast, MIMICSUnknown [June 4, 2024a] provides superior centerline extraction compared to skeletonize_3d(), though further radiologist refinement remains necessary to ensure data reliability. Consequently, expert manual corrections are indispensable for obtaining accurate tracheal centerlines, even after the use of skeletonize_3d() and MIMICS (see Figure 10).

Figure 10 highlights the limitations of traditional centerline extraction functions, which, although convenient, are not tailored to the trachea's spatial configuration and can generate central lines with looping and nodules. Conversely, medical software

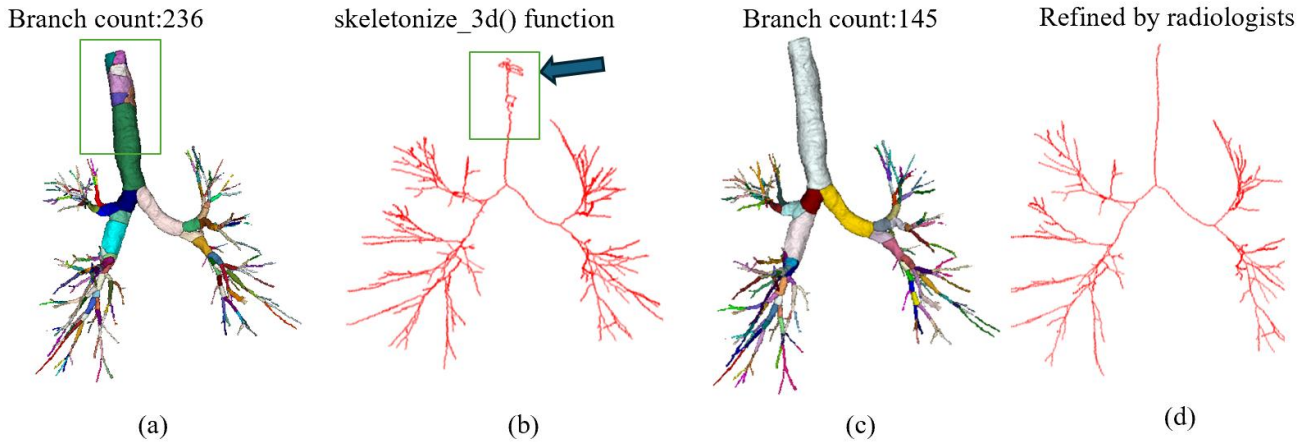

Figure 10: Visualization of Airway Parsing Results. (a) represents the tree parsing results based on the central line extracted by skeletonize_3d() function. (b) is the inaccurate central line extracted by skeletonize_3d() function. (c) shows the accurate tree parsing results based on the central line refined by radiologists. (d) represents an accurate, smooth, and nodule-free refined central line.

specifically designed for pulmonary airways removes most looping and nodules to produce smoother, more coherent tracheal centerlines. Nonetheless, small branches still require radiologist intervention. These refined centerlines afford more precise airway parsing, reducing inaccurate branch counts and diminishing erroneous segments from 236 to 145. In Section IV (RESULTS AND DISCUSSION), all airway prediction metrics reported are derived using the refined centerlines described in this section to support model training.

Proposed Loss Functions The proposed FuzzySR incorporates two novel loss functions: the Central Line Loss and the Attention Loss. Each branch is assigned a unique numerical identifier, which is determined through computational processes. The key to this calculation lies in the hierarchical structure defined by the parent-child relationships. This computation employs tree parsing methodologies to delineate the branches pertinent to each segment. The intersection metric is derived by multiplying the forecasted branch for every segment with its respective ground truth (GT) branch and then aggregating these multiplicative values. The aggregate of the branches delineated in the ground truth across all segments constitutes the calculation's denominator. Therefore, the formula for branch loss is given by equation (3) and (4). The formula for central line loss is given by following equations:

$$\mathcal{L}_{\text{CL}} = \frac{1}{N}\sum_{n=1}^{N}\left(1 - \frac{\sum_j \mathbf{P}_j^{(n)} \cdot \mathbf{T}_j^{(n)} + \epsilon}{\sum_j \mathbf{T}_j^{(n)} + \epsilon}\right). \tag{4}$$

$$\begin{aligned}\mathcal{L}_{\text{airway}} = \lambda \cdot \sum_{i=1}^{C} \left(\mathcal{L}_{\text{BCE}}(\mathbf{P}_i, \mathbf{T}) + \mathcal{L}_{\text{Dice}}(\mathbf{P}_i, \mathbf{T})\right) \\ + \lambda \cdot \sum_{i=1}^{C} \mathcal{L}_{\text{CL}}(\mathbf{P}_i, \mathbf{C}).\end{aligned} \tag{5}$$

Inaccuracies in airway central lines extracted via a Python library function, including imprecise branching and looping, led our experts to refine these lines. The improved lines were then used as a loss metric in our model's training. The proposed attention loss is designed by weighting the computed mean squared error (MSE) with the weighted attention map extracted from FuzzySR. This mechanism provides feedback to the super-resolution regions based on the varying importance of airway feature distributions. Theoretically, the proposed attention loss allows FuzzySR to allocate more focus to the pixels surrounding the airway features, thereby improving the accuracy of airway predictions.

C EXPERIMENT SETUP

Coding Environment and Radiologist Annotation tools. The study employed a quad-GPU setup featuring an NVIDIA

Algorithm 1 Cross Attention

class CROSS_ATTENTION(Q,K)
  initialize:
    $\mathrm{Q}_{\text{flat}}, K_{\text{flat}} \in \mathbb{R}^{B \times C \times (DHW)}$
    $W_Q, W_K \in \mathbb{R}^{B \times C \times D \times H \times W}$
    $\hat{S}_i \in \mathbb{R}^{B \times \text{chunk_size} \times \text{chunk_size}}$
function FORWARD(Q, K)
  for i = 0 to spatial_size by chunk_size:
    $Q_i = Q_{\text{flat}}[:, :, i : i + \text{chunk_size}]$
    $K_i = K_{\text{flat}}[:, :, i : i + \text{chunk_size}]$

    $S_i = \text{cosine_similarity}(Q_i, K_i) = \frac{Q_i^\top \cdot K_i}{\|Q_i\| \cdot \|K_i\|}$

    $\hat{S}_i = \text{softmax}\left(\frac{S_i}{\text{scale}}\right)$

    $W_{Q_i} = \hat{S}_i \cdot Q_i^\top, \quad W_{K_i} = \hat{S}_i \cdot K_i^\top$
    $W_Q[:, :, i : i + \text{chunk_size}] += W_{Q_i}$
    $W_K[:, :, i : i + \text{chunk_size}] += W_{K_i}$

  $attention_map = \min(W_Q, W_K)$
  return $attention_map$

TITAN X (12GB VRAM). For annotation, we utilized 3D SlicerUnknown [June 4, 2024b], ITK-SnapUnknown [June 4, 2024c], MeshLabMaggiordomo et al. [2021], and MIMICSUnknown [June 4, 2024a]. Specifically, experts used ITK-Snap for the annotation of the pulmonary airways and corrections of the central line. MeshLab converted airway ground truths (annotated in ITK-Snap in NIFTI format) into meshes. 3D Slicer transformed 3D CT data into DICOM, a medical imaging standard. These DICOM CT scans and meshes were then processed in MIMICS to generate central lines, followed by expert annotation. We plan to share our findings, including source code, with the scientific community in due course. **Performance Evaluation Metrics:** The efficacy of the model was assessed through predictions conducted on the test dataset, subsequently benchmarked against the ground truth annotations provided by radiologists. Four key metrics were employed for this evaluation: the Dice Similarity Coefficient (DSC), Precision, Tree Detected Ratio (TD), and Branch Detected Ratio (BD). These measures facilitated a comprehensive analysis of the model's performance in terms of accuracy, precision, and its ability to accurately detect tree and branch structures within the pulmonary airway. To validate the performance of the proposed FuzzySR in super-resolution, three evaluation metrics were introduced: SSIM(Structural Similarity Index Measure), PSNR(Peak Signal-to-Noise Ratio), and RMSE(Root Mean Square Error). SSIM is an index that measures the similarity in visual effects between two images. It considers changes in luminance, contrast, and structure of the images, with values ranging from -1 to 1, where 1 indicates identical images. This measure aligns more closely with human visual perception of image quality. PSNR is one of the most commonly used metrics for assessing the quality of image reconstruction. It evaluates image quality by calculating the relationship between the maximum possible pixel value of the original image and the mean squared error (MSE) with the reconstructed image. A higher PSNR value indicates better image quality. RMSE directly measures the pixel-level differences between the original and reconstructed images. A lower RMSE value indicates that the image is closer to the original, signifying higher image quality.