SIIM CMIMI24
Oct 21-22 | Boston, MA
Conference On Machine Intelligence In Medical Imaging

Attention Variant Mechanism for Airways Segmentation

Chetana Krishnan, University of Alabama at Birmingham; Shah Hussain, PhD; Denise Stanford, PhD; Venkata Sthanam, PhD; Sandeep Bodduluri, PhD; Steve Rowe, PhD; Harrison Kim, PhD

Introduction/Background

Attention mechanisms enhance neural networks by focusing on relevant input features but often have limitations, such as addressing only single or dual aspects and struggling with diverse inputs. Our architecture overcomes these by integrating multiple attention strategies and adaptive embedding, ensuring dynamic, robust feature extraction and improved performance in small region segmentations (SRS). Integrating positional (POS), semantic (SEM), image (IM), cross-spatial (CS), and self-channel attentions (SC) with adaptive embedding will significantly enhance feature extraction and representation, improving accuracy and efficiency in SRS, such as airways in comparison to single/dual attentions.

Methods/Intervention

Non-contrast enhanced in vivo CT scans of the lungs were conducted on 25 ferrets, achieving a spatial resolution of 80 µm. The ferrets were anesthetized with inhaled isoflurane and gated to capture a single inspiratory phase using a µCT scanner (MiLabs, Utrecht, Netherlands). The ground truth airway was determined using a region-growing method. The proposed attention variant network (AVN) incorporates information from other pixels within the image by performing SEM, POS, SC, CS, and IM. AVN inputs feature maps from all locations at different scales and outputs refined feature maps. This approach captures and utilizes correlations between neighboring pixels, leading to more accurate segmentation. Multi-scale feature maps refine the attention mechanism, enabling precise adaptation to image data variations. SEM focuses on important semantic information, POS identifies where vital information is located, IM determines task-relevant regions, CS examines spatial relationships, and SC emphasizes relevant channel features. We trained AVN with 18 scans (over 500 slices per scan), and the model was tested with seven unseen scans. Performance was evaluated using the Dice similarity coefficient (DSC) and Intersection over Union score (IoU). AVN was compared with other popular deep-learning networks.

Results/Outcome

AVN achieved a higher DSC and exhibited the highest minimum DSC, indicating superior performance, as summarized in Table 1. Figure 1 illustrates the AVN structure. Figure 2 demonstrates the improved segmentation accuracy via AVN.

Conclusion

AVN dynamically captures spatial and channel information to address the challenge of SRS and the limitations of 2D networks.

Statement of Impact

AVN can improve the diagnosis, treatment planning, and airway branch and volume monitoring for clinical lung diseases.

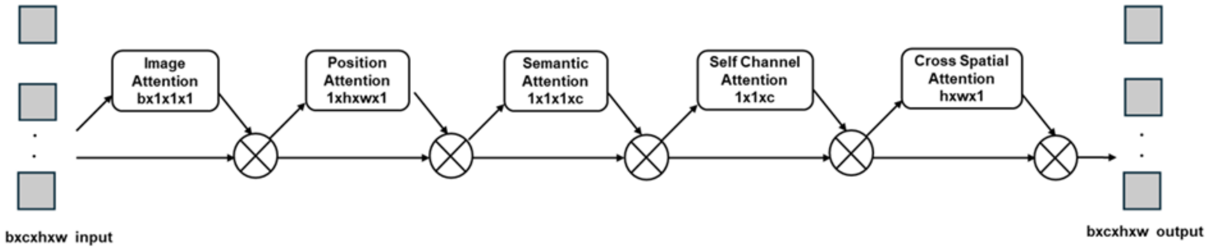

Figure 1. AVN Architecture. The input feature maps are fed into the module, where they go through five sequential sub-modules: the image, position, semantic, self-channel, and cross-spatial attention modules.

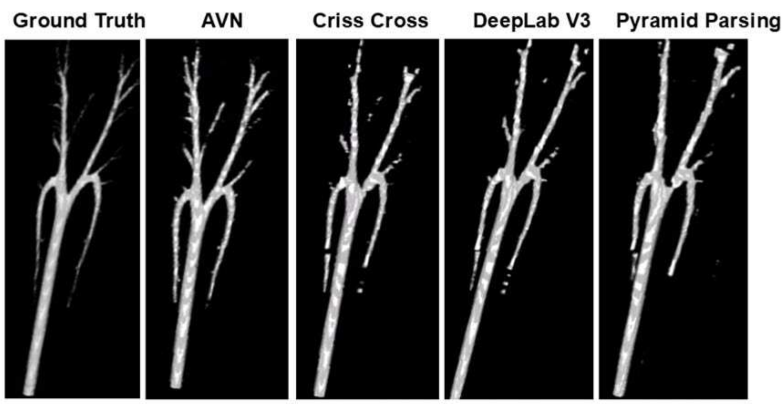

Figure 2. Segmentation Results. AVN demonstrates improved boundary, edge localization, and enhanced accuracy in capturing minute small branches. AVN captures spatial volumetric information accurately despite using 2D convolutions whereas breakage points are noticed in other networks.

Model	Test Average Dice Coefficient	Test Average IoU Score	Worst Case Dice Coefficient
AVN	0.91±0.017	0.83±0.028	0.88
Criss Cross Attention	0.85±0.008	0.73±0.011	0.84
DeepLab V3	0.85±0.008	0.73±0.012	0.83
Pyramid Parsing Network	0.85±0.010	0.74±0.015	0.83

Table 1. Performance Metrics. Average Dice Similarity Coefficient, Intersection over union (IoU) score, and worst case dice similarity coefficient of the proposed model (AVN) in comparison with three other popular networks on the test set.

Keywords

Attention; Segmentation; Spatial; Channel; Airways; Feature maps