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

Global Local Attention for Prostate Zonal Segmentation

Chetana Krishnan, University of Alabama at Birmingham; Ezinwanne Onuoha, PhD; Alex Hung, PhD; Kyung H. Sung, PhD; Harrison Kim, PhD

Introduction/Background

We focus on representation learning for large-scale image segmentation. Besides backbones, training pipelines, and loss functions, key methods have explored various spatial pooling and attention mechanisms essential for creating robust global image representations. Attention mechanisms differ based on feature tensor interactions (local vs. global) and the dimensions they target (spatial vs. channel). However, most studies examine only one or two forms of attention. Focusing on global and local descriptors, we can provide empirical evidence of the interaction of all forms of attention and improve the state of the art on standard benchmarks.

Methods/Intervention

The proposed GLCSA network uses multi-stream processing to capture comprehensive contextual information from images. The local attention stream (LAS) focuses on detailed information at individual spatial locations and specific feature channels, highlighting fine-grained patterns and textures. The global attention stream (GAS) models interactions across the entire spatial dimension and among feature channels, ensuring broader relationships are captured. The LAS uses fine-scale convolutions to discern intricate details, while the GAS leverages self-attention to integrate long-range dependencies. The information from these streams is embedded into feature maps, which are then fused into a unified feature map. Finally, a pooling operation distills the combined information into a compact representation for robust image analysis. We trained GLCSA with 34 prostate scans (over 20 slices per scan), and the model was tested with ten unseen scans. Performance was evaluated using the Dice similarity coefficient. Several networks were compared with and without GLCSA.

Results/Outcome

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

Conclusion

GLCSA dynamically captures global–local spatial and channel information to address the challenge of prostate segmentations and the limitations of 2D networks.

Statement of Impact

GLCSA can improve the diagnosis, treatment planning, and monitoring of prostate pathology and volume for clinical diseases.

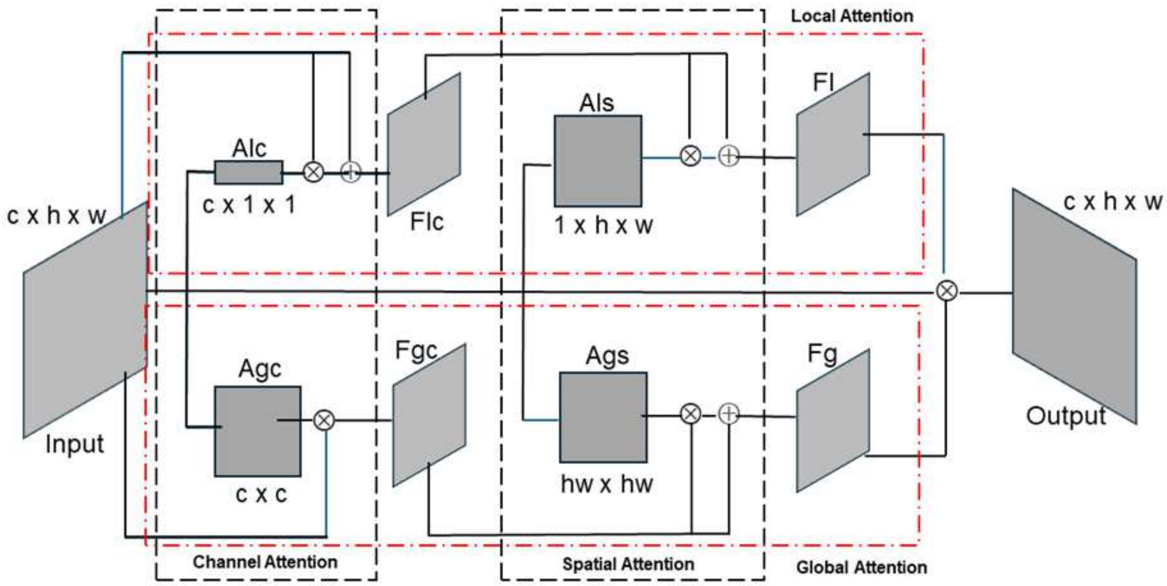

Figure 1. GLCSA Architecture integrates both channel and spatial attention mechanisms, as well as local and global attention strategies. Local attention applies independent weighting to channels or locations based on contextual information obtained through pooling, while global attention considers pairwise interactions between channels or locations. Consequently, four distinct attention maps are utilized: local channel (Alc), local spatial (Als), global channel (Agc), and global spatial (Ags). The input is modulated into local (Fl) and global (Fg) attention feature maps. These maps are then combined with input to produce the final global-local attention feature map (output).

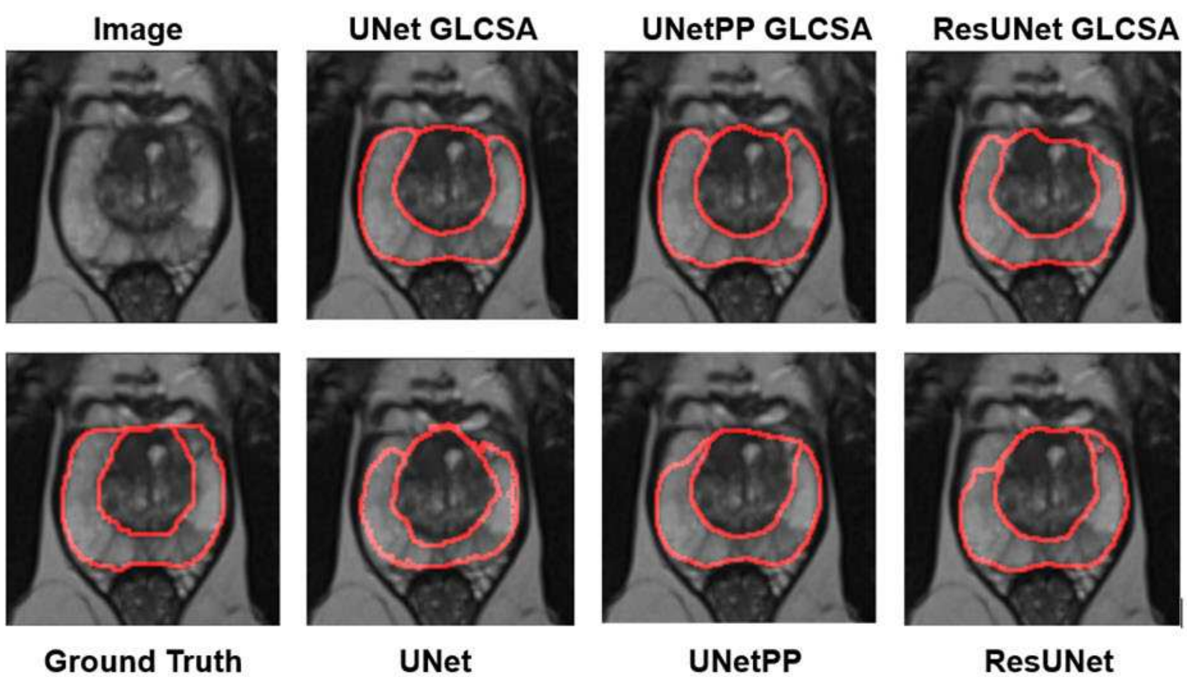

Figure 2. Segmentation Results. GLSCA-based models demonstrate improved boundary, edge localization, and enhanced accuracy in capturing minute small branches. GLCSA captures spatial volumetric information accurately despite using 2D convolutions whereas breakage points are noticed in other networks.

Model	Zone	Test Average Dice Coefficient		Worst Case Dice Coefficient	
		With GLCSA	Without GLCSA	With GLCSA	Without GLCSA
UNet	PZ	0.75±0.086	0.60±0.104	0.58	0.45
	TZ	0.93±0.010	0.84±0.034	0.91	0.76
UNet Plus Plus	PZ	0.75±0.081	0.73±0.081	0.60	0.63
	TZ	0.92±0.013	0.91±0.016	0.90	0.88
Residual UNet	PZ	0.71±0.078	0.66±0.111	0.90	0.59
	TZ	0.93±0.012	0.89±0.025	0.84	0.52

Table 1. Performance Metrics. Average Dice Similarity Coefficient, and worst case dice similarity coefficient of the proposed model (GLCSA) in comparison with networks without the use of GLSCA of the peripherial (PZ) and transitional (TZ) prostate zones on the test set.

Keywords
Attention; Segmentation; Spatial; Channel; Prostate; Global-Location