Attention U-Net: Learning Where to Look for the Pancreas

CS300 : Midsem Project Evaluation

Sreejita Das

Department of Computer Science and Engineering
Indian Institute of Information Technology, Guwahati

3 March, 2025

Problem Statement

- **Challenge**: Accurate segmentation of anatomical structures, such as the pancreas, in medical imaging remains an obstacle due to significant inter-patient variability in shape and size, low tissue contrast in CT scans, and the dependency on computationally intensive multi-stage convolutional neural networks (CNNs) that inefficiently localize regions of interest.
- **Goal**: Develop an innovative, single-model architecture—Attention U-Net—that integrates attention gates to autonomously emphasize salient features and suppress irrelevant regions, achieving enhanced segmentation precision with reduced computational overhead.

Literature Survey

1. **Ronneberger et al. (2015) — U-Net: Convolutional Networks for Biomedical Image Segmentation**
    - Created U-Net to segment 2D medical images with a shrink-and-expand design for clear details.
    - Set the foundation for 3D models, aiding pancreas segmentation in CT scans.

2. **Çiçek et al. (2016) — 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation**
    - Upgraded U-Net to 3D for CT scans, linking slices for a full view.
    - Key for pancreas segmentation by capturing organ connections in 3D.

Literature Survey (Contd.)

3. **Milletari et al. (2016) — V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation**
    - Built V-Net for 3D CT/MRI scans, focusing on small organs like the pancreas.
    - Uses a smart scoring trick to handle uneven data, boosting accuracy.

4. **Gibson et al. (2017) — Automatic Multi-Organ Segmentation with Dense Dilated Networks**
    - Designed a method to segment multiple organs in CT scans at once.
    - Helps separate pancreas from nearby organs with a wider view.

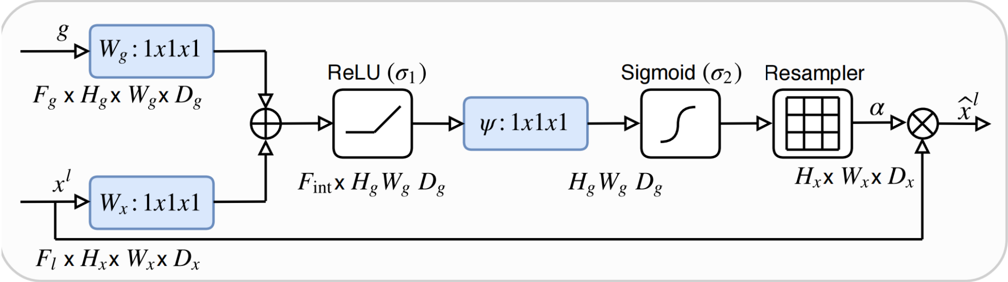
Research Focus: Attention U-Net - Learning Where to Look for the Pancreas

- **Medical Image Segmentation:** Divides medical images like CT scans into meaningful regions, identifying organs and tissues for disease diagnosis, treatment planning, and quantitative analysis. Deep Learning, particularly CNNs, automates feature extraction, boosting accuracy.
- **Attention U-Net:** Enhances pancreas segmentation by using **Attention Gates (AGs)** — mechanisms that focus on relevant regions while suppressing background noise, refining feature maps dynamically.
- **Loss Function:** Combines Dice Loss and Binary Cross-Entropy Loss (BCE) to optimize segmentation accuracy:

$$\mathcal{L} = \mathcal{L}_{Dice} + \mathcal{L}_{BCE}$$

Model Architecture

Main Paper Link: click here



Methodology: Attention Mechanism

1. **Input Processing:** Two parallel inputs:
   - **Local features** x_i^l from lower layers
   - **Gating signal** g_i from coarser scales
2. **Feature Transformation:** Both inputs undergo linear transformations via 1x1x1 convolutions:
$$W_x^T x_i^l \quad \text{and} \quad W_g^T g_i$$
3. **Feature Combination and Non-Linearity:** Additively combined and passed through ReLU:
$$\sigma_1(W_x^T x_i^l + W_g^T g_i + b_g)$$
4. **Attention Coefficients:** Transformed via linear mapping, normalized by sigmoid:
$$\alpha_i^l = \sigma_2\left(\psi^T \sigma_1(W_x^T x_i^l + W_g^T g_i + b_g) + b_\psi\right)$$
5. **Gated Output:** Final gated output via element-wise multiplication:
$$\hat{x}_{i,c}^l = x_{i,c}^l \cdot \alpha_i^l$$

Training Results from the Project Source Code

- **Dataset**: Pancreas CT-82 (TCIA), currently trained on its reduced version (60 patients out of total of 82)
- **Number of Epochs**: 70

	A	B	C
1	Metric	Predicted 3D U-Net	Attention U-Net
2	Dice Score	65%–70%	81.5% ± 6.8%
3	Precision	67%–71%	81.5% ± 9.3%
4	Recall	68%–72%	83.5% ± 5.7%
5	S2S Distance (mm)	3.0–3.5	2.333 ± 0.856 mm

Future Improvement: Enhancing Attention U-Net

- **Limitation:** The current model uses one attention focus, which can miss complex organ shapes and their broader connections—like the pancreas near the spleen.
- **Proposed Upgrade:** Introduce multi-head attention to split focus into multiple views, capturing both small details and the big picture simultaneously.

  Example: One head highlights the pancreas boundary near the stomach, another maps its overlap with the spleen blood vessels.
- **Benefit:** This could boost accuracy for tricky organs like the pancreas with little extra effort—an idea I'm eager to test.

Related Work: Hierarchical 3D Fully Convolutional Networks for Multi-Organ Segmentation

- **What They Did:** Used a 3D U-Net with a hierarchical two-step method, a **coarse-to-fine** approach to detect **seven abdominal organs** (artery, vein, liver, spleen, stomach, gallbladder, pancreas) in CT scans.
- **First Step:** Roughly outlines organs by scanning **40% of the image** with a **simple body mask** to focus on relevant regions, separating the background from the organs.
- **Second Step:** Zooms into **10% of the image** based on the first guess, sharpening edges for **precise boundaries**.
- **Key Wins:** Big accuracy boost for **small organs** and adapts to **new hospital testing data**.

Dice	liver	spleen	pancreas	liver	spleen	pancreas	liver	spleen	pancreas
Mean	93.6	89.7	68.5	94.9	91.4	81.2	95.4	92.8	82.2
Std	2.5	8.2	8.2	2.1	8.9	10.2	2.0	8.0	10.2
Median	94.2	91.8	70.3	95.4	94.2	83.1	96.0	95.4	84.5
Min	78.2	20.6	32.0	80.4	22.3	1.9	80.9	21.7	1.8
Max	96.8	95.7	82.3	97.3	97.4	91.3	97.7	98.1	92.2

Conclusion

In this project, a robust approach for pancreas segmentation from CT images was successfully developed leveraging advanced deep learning techniques.The methodology effectively addressed key challenges such as boundary ambiguity and class imbalance, achieving notable improvements in segmentation accuracy.

The results underscore the potential of AI-driven solutions in medical image analysis, paving the way for more precise and automated diagnostic tools. Future work may focus on enhancing model generalization across diverse datasets and integrating clinical insights to further refine segmentation performance.

References

Gibson, E., Giganti, F., Hu, Y., et al. (2017).
Automatic multi-organ segmentation with dense dilated networks.
In *MICCAI*.

Milletari, F., Navab, N., and Ahmadi, S.-A. (2016).
V-net: Fully convolutional neural networks for volumetric medical image segmentation.
In *3DV*.

Ronneberger, O., Fischer, P., and Brox, T. (2015).
U-net: Convolutional networks for biomedical image segmentation.
In *MICCAI*.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016).
3d u-net: Learning dense volumetric segmentation from sparse annotation.
MICCAI.

Thank You