Department of Electrical and Electronic Engineering
電 機 電 子 工 程 系

Explainable Machine Learning Algorithm and Hardware for Surgical Robotics

Student: Tian Qijia Supervisor: Dr. Can Li

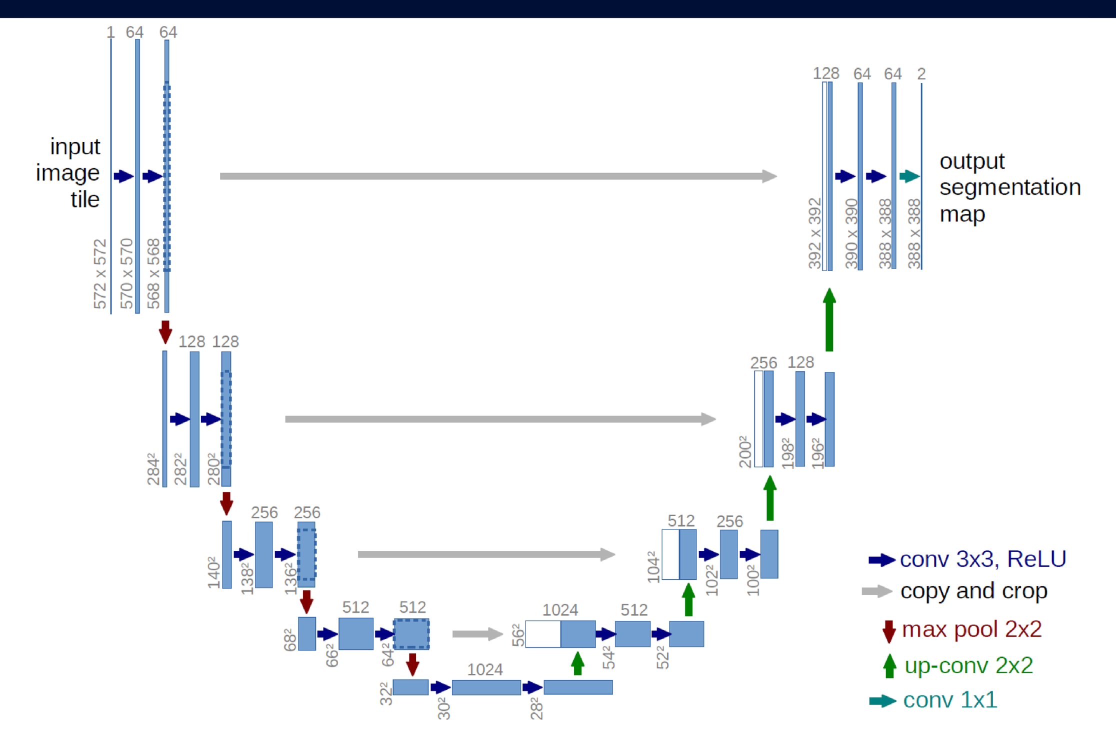
Abstract

Deep learning techniques have shown promise in accurate segmentation tasks in medical imaging, particularly in breast cancer diagnosis. This project endeavors to develop an AI-driven segmentation approach tailored for breast tumor delineation, focusing on efficiency and interpretability. The primary objective is to design an Attention UNet model capable of achieving high segmentation accuracy while addressing computational constraints commonly encountered in medical imaging tasks.
The methodology involves extensive experimentation, including data preprocessing, model training, and evaluation, culminating in robust segmentation performance metrics. The results reveal the model's high and stable accuracy in segmenting images, though the Intersection over Union in the evaluation metrics indicates some underperformance. This highlights the imperative for continued research aimed at not only improving the model's interpretability but also expanding its utility across a variety of medical imaging modalities.
Recommendations encompass acquiring more extensive annotated datasets and delving into sophisticated architectural enhancements to bolster the model's robustness and viability for deployment in real-world scenarios. Additionally, it is advised to swiftly design and implement high-performance, low-power hardware solutions to complement these advancements.

Objective

- Create an **efficient**, **accurate**, **explainable** image segmentation model.
- Build corresponding **hardware** for surgical robot guidance.

Theorectical Background



UNet Structure Proposed by Ronneberger et al. [1]

UNet Architecture Overview for Biomedical Segmentation

- **Purpose**: Precision in segmenting biomedical images with minimal data.
- **Structure**:
  - Contracting Path: Downscales images via convolution, reducing size but enhancing features.
  - Shortcut: A critical juncture that retains essential image features for detailed analysis.
  - Expansive Path: Upscales images, blending detailed features for accurate outlines.
- **Output**: Combines local details with overall context for precise segmentation.
- **Advantage**: Balances broad insights and specific details for superior accuracy.

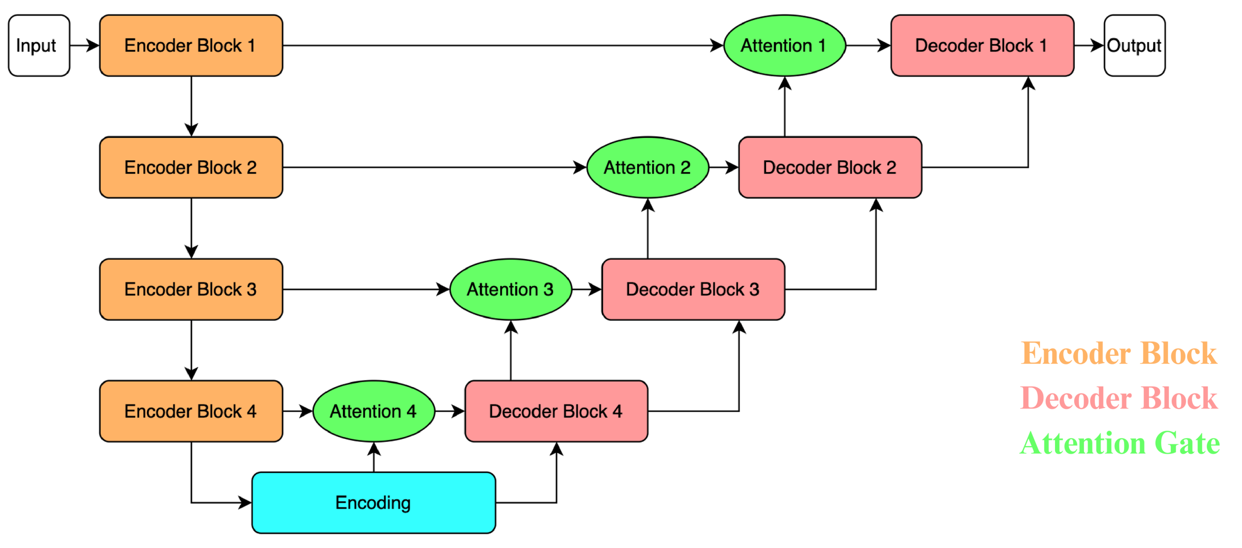
Evaluation Metrics [2]:

- **Accuracy**: Determines how often the model gets it right in spotting the target area (True Positive) and avoiding false alarms (True Negative).
- **Dice Coefficient** (Dice): Imagine two circles - one drawn by experts (ground truth) and one by our model. Dice tells us how much they overlap; the closer to 1 (or 100%), the better.
- **Intersection over Union** (IoU): This is like Dice but stricter. It divides the overlapping area by the total area they cover together. Closer to 1 means our model's predictions are spot-on.

Interpretability Methods [3]:

- **Saliency Maps**: These highlight what the model sees as important, using gradients to pinpoint image focus areas.
- **GradCAM**: Offers a heatmap visualization, showing where the model pays the most attention, thanks to the Attention layer.
- **Occlusion Sensitivity**: By hiding parts of the image and watching how the model reacts, we learn which parts are vital for its decisions.
- **Attention UNet**: This model's attention mechanism shines a light on the key regions, improving our grasp on its decision-making.

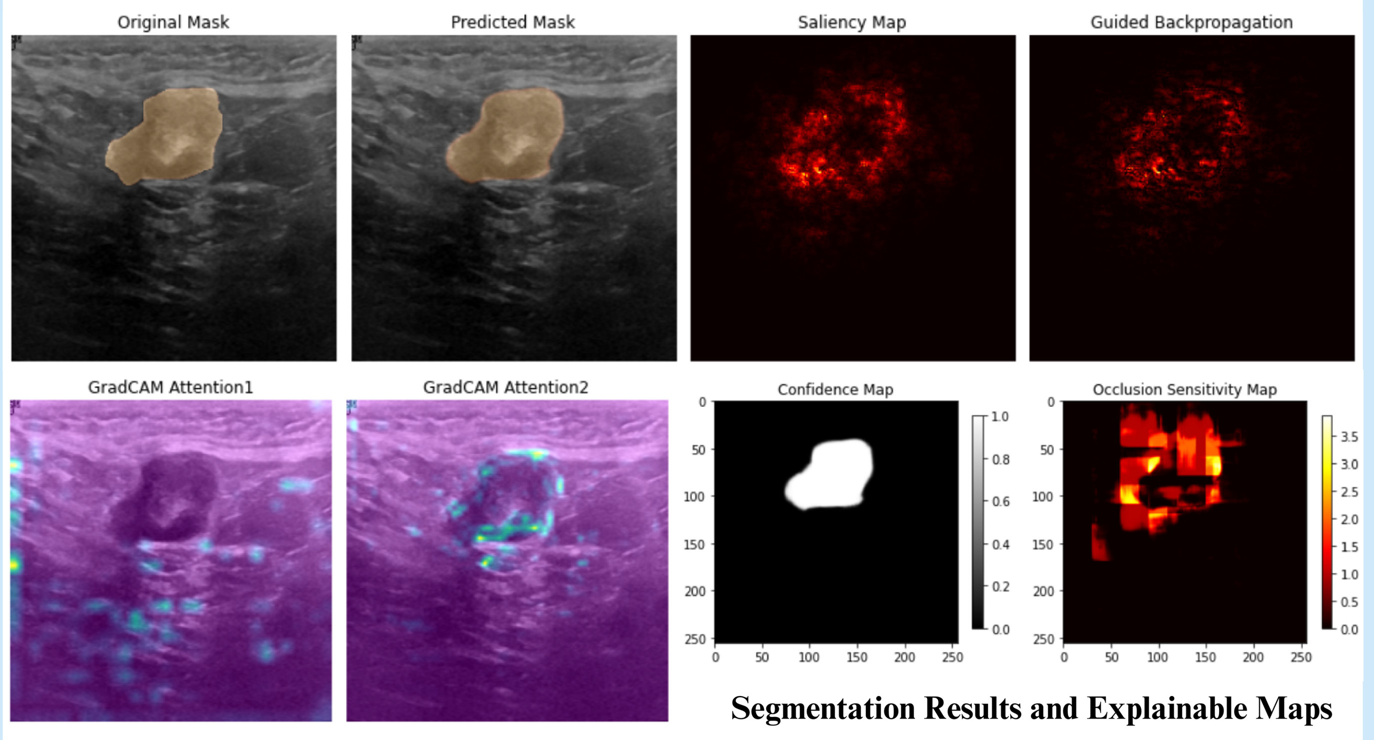
Procedure



Attention UNet Structure (Modular Design)

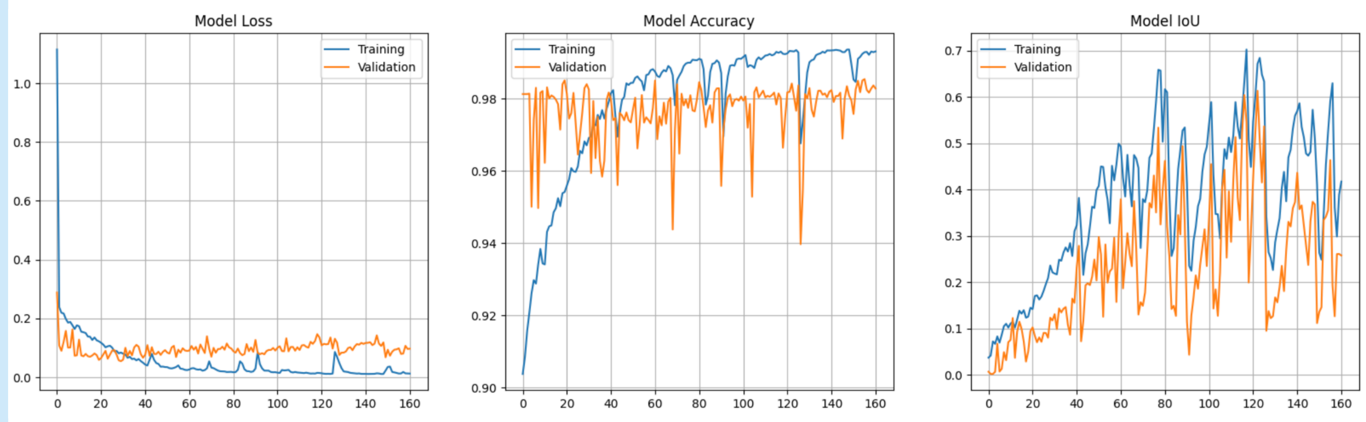
The model introduces **Attention Gates** in UNet Structure.
Process: Data preprocessing - Building Components - Training - Inference & Evaluation

Results



Segmentation Results and Explainable Maps

- **Precision in mask**: Exact tumor boundaries identified.
- **Saliency Maps**: Tumor edges are key focus areas.
- **Attention Gate**: Prioritizes critical regions, like folds.
- **Occlusion Sensitivity**: Covering edges significantly affects accuracy.

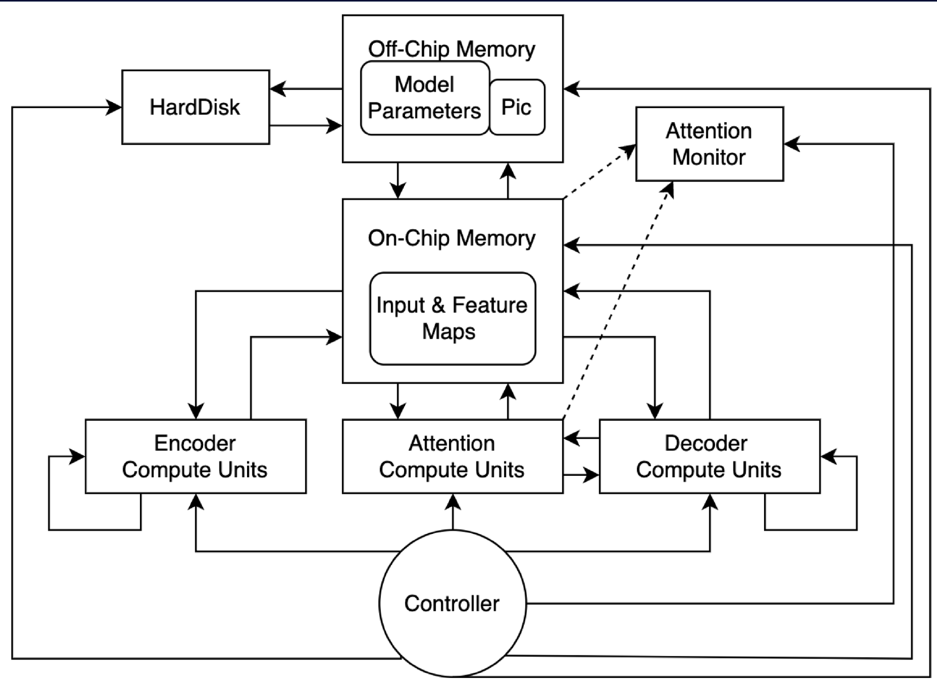


Evaluation Metrics - Loss, Accuracy, IoU

The evaluation, depicted with blue for **training** and orange for **validation** performance, shows promising trends and areas for improvement:

- **Loss**: Rapidly decreases with training, stabilizing around 0.1, indicating effective learning.
- **Accuracy**: Remains high at 0.98, showcasing the model's reliable segmentation capability.
- **IoU** Fluctuations: Displays noticeable volatility in its upward trend, suggesting some instability during training. The converging gap between training and validation IoU suggests a reduced risk of overfitting, potentially tied to optimizer adjustments.
- **Gap** with State of the Art: Despite improvements, there's at least a 30% performance gap compared to leading methods [2], highlighting room for further optimization.

Hardware Design



Hardware Structure

The hardware, designed for Attention UNet integration in medical devices, balances energy efficiency with power. It features **dual-memory**: off-chip for large storage and on-chip for fast data processing, enhancing responsiveness. The structure supports swift adjustments and efficient data flow, minimizing off-chip memory access. The system comprises three units—Encoder, Attention, and Decoder—each tailored for specific segmentation phases. This design streamlines workflow, improves efficiency, and ensures seamless operation with medical imaging tasks, all coordinated by a central controller for operational harmony.

Conclusion

This project developed the Attention UNet model for detecting breast cancer in ultrasound images, demonstrating AI's potential to enhance diagnostic accuracy and reduce healthcare burdens. The model's design allows for efficient operation with limited resources, and its interpretability helps gain medical professionals' trust.
Despite challenges such as irregular tumor shapes, the model showed promising results with high accuracy and an improving trend in IoU. It aims to automate image preprocessing, lightening the workload for physicians.
Future work will address dataset limitations, refine metrics, and explore new architectures like Transformer UNet to enhance performance. Developing specialized hardware is also a priority to ensure this technology can be effectively used in clinical settings, making significant strides in medical imaging analysis.

Reference

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," arXiv.org, https://arxiv.org/abs/1505.04597 (accessed Mar. 7, 2024).
[2] V. K. Singh et al., "Breast tumor segmentation in ultrasound images using contextual-information-aware deep adversarial learning framework," Expert Systems with Applications, vol. 162, p. 113870, Dec. 2020. doi:10.1016/j.eswa.2020.113870
[3] Q. Teng, Z. Liu, Y. Song, K. Han, and Y. Lu, "A survey on the interpretability of deep learning in medical diagnosis," Multimedia Systems, vol. 28, no. 6, pp. 2335–2355, Jun. 2022. doi:10.1007/s00530-022-00960-4