
Executive Summary: Deep Learning-based Comparative Segmentation of Thyroid Ultrasound Images

---

1. Introduction and Problem Statement

The diagnosis of thyroid diseases heavily relies on **ultrasound imaging**, where the accurate boundary definition, or segmentation, of the thyroid gland or nodules is crucial. Manual segmentation is inefficient, time-consuming, and prone to variability. This capstone project addresses this limitation by developing and comparatively analysing advanced **Deep Learning (DL)** models for automated, high-accuracy thyroid ultrasound image segmentation.

The primary objective was to rigorously compare three state-of-the-art Convolutional Neural Network (CNN) architectures : **U-Net, SegNet**, and the enhanced **Residual U-Net**, to select a model that not only achieves high accuracy but also demonstrates the consistent and stable performance required for clinical integration.

2. Methodology and Technical Approach

2.1. Data Preparation and Preprocessing

The project utilized a comprehensive **Thyroid Dataset** consisting of 3,585 high-quality image-mask pairs.

- **Data Split:** The dataset was partitioned into a **training set of 2,509 samples** and a **testing set of 1,076 samples**.
- **Preprocessing:** All input images underwent preprocessing steps including conversion to grayscale, normalization, and resizing to **256x256 pixels** to standardize inputs for all models.

2.2. Deep Learning Architectures

Three distinct semantic segmentation architectures were implemented using the **PyTorch** framework:

1. **U-Net:** A classic encoder-decoder structure utilizing **skip connections** to preserve spatial detail.
2. **SegNet:** Characterized by its use of Max-Pooling indices for efficient upsampling.
3. **Residual U-Net:** An advanced variant that incorporates **Residual Blocks** into the U-Net design to stabilize training, mitigate the vanishing gradient problem, and ensure more consistent feature learning.

2.3. Training Environment and Evaluation

All models were trained across a combination of the **Kaggle GPU environment** and the **Colab GPU environment**, utilizing an **Adam Optimizer** (learning rate of $1e-4$) and **Binary Cross Entropy (BCELoss)** for **25 epochs**.

The core performance metric was the **Dice Coefficient**, which measures the spatial overlap between the predicted and ground truth masks.

3. Key Findings and Model Selection

3.1. Final Model Performance Comparison

The comparative evaluation on the independent test set yielded the following results:

Model Architecture	Final Test Dice Coefficient	Comparative Performance/Stability
U-Net	**0.9765**	Highest peak accuracy, but required careful monitoring for stability.
Residual U-Net	**0.9732**	Extremely high accuracy with the most **consistent** and stable performance throughout training and testing.
SegNet	0.7783	Significantly lower performance, indicating poor feature reconstruction for this task.

3.2. Selection Rationale: Residual U-Net

While the standard U-Net achieved a marginally higher peak Dice score, the **Residual U-Net** was ultimately selected as the model for deployment. The model's integration of **Residual Blocks** yielded the **most stable training curves and reproducible high accuracy** (Dice: **0.9732**). In a capstone project aiming for a deployable clinical solution, the inherent robustness and reliability of the Residual U-Net were deemed superior to the marginal accuracy gain of the U-Net, ensuring consistent and trustworthy results in a production environment.

---

4. Conclusion and Future Scope

This project successfully implemented and evaluated deep learning architectures for automated thyroid segmentation, validating the potential of DL to improve diagnostic efficiency. The **Residual U-Net** architecture was chosen for its excellent performance (Dice Coefficient of **0.9732**) combined with its superior stability and resilience, making it an ideal candidate for clinical integration.

Future Work:

- **Advanced Architecture Search:** Integrate pre-trained backbones (e.g., ResNet) within the Residual U-Net to explore potential further improvements in feature extraction.
- **Clinical Integration:** Focus efforts on deploying the trained Residual U-Net model as a service that can interface with a hospital's medical imaging system (PACS) for real-time validation.
- **Extension to Multi-Nodule:** Expand the model’s capability to perform multi-class segmentation, enabling it to distinguish and segment different types of thyroid nodules.