`

Enhancing Medical Image Segmentation: A Study of U-Net Architecture and Performance Optimization on the BCSS Dataset

Ren-Di Wu

whats2000mc@gmail.com

1. Introduction:

This study aims to construct and apply the U-Net architecture from scratch for medical image segmentation, focusing on the BCSS dataset. U-Net, renowned for its effectiveness in image boundary delineation, is critical in medical diagnostics. Building U-Net ourselves allows for a deeper understanding and potential enhancements of its structure.

The experiment's core is to explore U-Net's capabilities in the precise segmentation of medical images, an essential aspect in the diagnosis and treatment of diseases. Through this process, we seek to contribute to advancements in medical image analysis and potentially improve disease detection accuracy.

2. Methodology:

Dataset Configuration

- **Data Overview:** Utilized the Breast Cancer Semantic Segmentation (BCSS) dataset, composed of medical images for segmentation.
- **Image Details:**
    - Total Training Images: 30,760
    - Total Validation Images: 5,429
    - Total Testing Images: 4,021
    - Image Resolution: Each image is of size 224 x 224 pixels.
    - Segmentation Classes: The ground truth is segmented into 3 classes.

Experiment Approach

- **U-Net Implementation:** Developed the U-Net architecture from scratch, adhering closely to its original design to understand its mechanics and behavior in medical image segmentation.
- **Data Augmentation Trials:** Experimented with various data augmentation techniques to enhance the model's ability to generalize and perform effectively on unseen data.
- **EMA Adoption:** Incorporated the Exponential Moving Average (EMA) strategy, known for its success in object detection, to assess its impact on segmentation performance.
- **Training Strategy:**
    - Utilized label smoothing with a value of 0.1 to stabilize training and potentially improve generalization.
    - Employed metrics such as mean Intersection over Union (mIOU) and accuracy for model evaluation and performance assessment.

`

U-Net Architecture

- **Building Blocks:**
  - **`DoubleConv`:** A convolutional block consisting of two convolutional layers, each followed by batch normalization and ReLU activation.
  - **`Down`:** Encapsulates pooling and double convolution, forming the contracting path of the network.
  - **`Up`:** Manages the expansive path, including up-sampling and concatenation with features from the contracting path.
  - **`OutConv`:** The final convolution layer to produce segmentation maps.
- **Model Flow:**
  - The U-Net model starts with the `DoubleConv` block, followed by successive `Down` blocks that reduce spatial dimensions while increasing feature maps.
  - The lowest resolution is then expanded through `Up` blocks, which combine up-sampled output with corresponding features from the contracting path.
  - The `OutConv` block finalizes the segmentation map prediction.
- **Implementation Details:** The U-Net model is constructed with careful consideration of each layer's role in achieving precise segmentation. The architecture's ability to capture both context and localization details is leveraged to address the challenges in medical image segmentation.

3. Results:

The performance outcomes of the U-Net model are meticulously detailed in Table 1. This table highlights the accuracy and mean Intersection over Union (mIOU) metrics for both the training and validation datasets. The mIOU metric, a crucial indicator of segmentation precision, is also presented for the test dataset. (Table 1, Figure 1~3)

4. Analysis & Interpretation

Analysis of Results

The performance data, as shown in *Table 1*, reveals insightful trends and outcomes:

- **Training Set Performance:** The base U-Net model achieved an accuracy of 75.69% and an mIOU of 58.77%. The absence of EMA model metrics in this phase suggests a focused evaluation of the base model during training.
- **Validation Set Performance:** Here, both models were closely matched in terms of accuracy and mIOU. The base U-Net recorded slightly higher accuracy (77.05% vs.

`

76.67%) and mIOU (60.58% vs. 60.21%) compared to the EMA model.

- **Test Set Performance:** The mIOU for the base U-Net was marginally higher at 60.95% compared to 60.76% for the EMA model.

Interpretation

- **Model Efficacy:** Both models demonstrated good performance, with the base U-Net slightly outperforming the EMA variant in validation and test metrics. This indicates the robustness of the base U-Net architecture in handling medical image segmentation tasks.
- **EMA Impact:** The expectation that EMA would significantly boost performance was not strongly evidenced in this case. However, its close alignment with the base model's results suggests stability and consistency in model predictions.
- **Generalization Capability:** The mIOU scores on the test set for both models suggest a competent level of generalization, crucial for practical medical applications.

5. Conclusion

The results affirm the U-Net architecture's strong capability in medical image segmentation, as demonstrated on the BCSS dataset. Both the base U-Net and the EMA variant exhibited impressive segmentation abilities, with the base model slightly surpassing the EMA model in terms of accuracy and mIOU on the validation and test datasets.

Interestingly, the EMA model did not significantly outperform the base model as hypothesized. This could indicate that for this specific application and dataset, the benefits of EMA in stabilizing and improving performance are somewhat limited. However, its close performance to the base model underscores its potential as a reliable alternative, especially in scenarios where prediction stability is paramount.

This study reaffirms the strength of U-Net in medical image analysis and opens avenues for further research, particularly in exploring other enhancement techniques or architectural modifications that could yield more pronounced improvements in segmentation tasks.

6. Appendices

The Appendices contain comprehensive tables and plot that detail the results of the experiment.

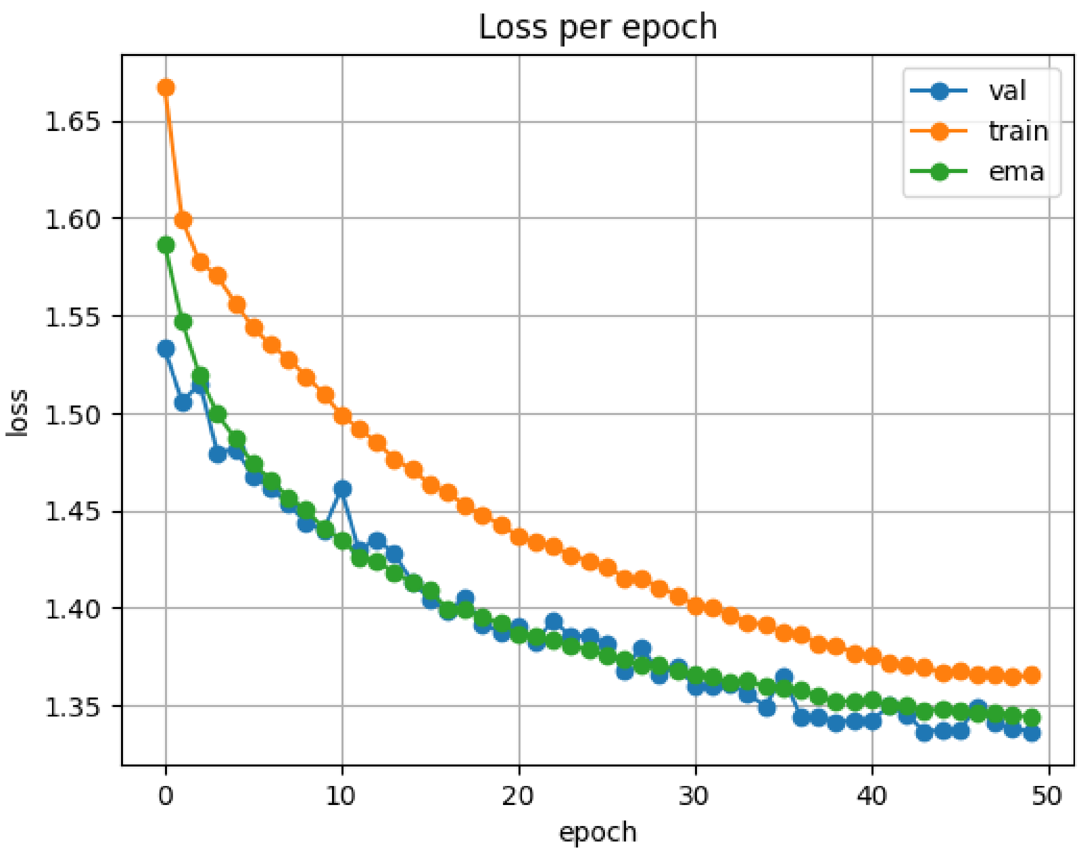
Table 1

Performance Metrics of U-Net Model

Dataset	Metric	Base U-Net Accuracy	Base U-Net mIOU	EMA Model Accuracy	EMA Model mIOU
Training Set	Accuracy, mIOU	0.756866	0.587662	-	-
Validation Set	Accuracy, mIOU	0.770524	0.60576	0.766657	0.602057
Test Set	mIOU	-	0.6095	-	0.60758

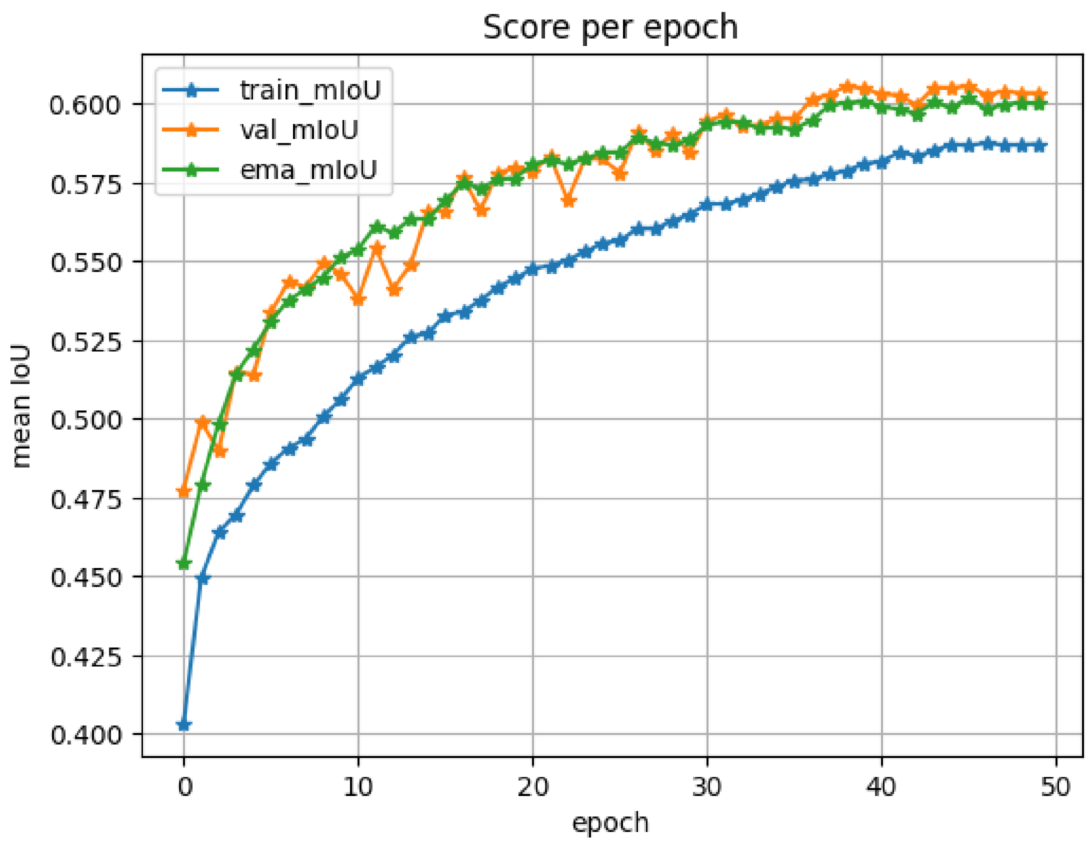
Note. The accuracy is not provided on test set

^ *Table 1: Performance Metrics of U-Net Model*

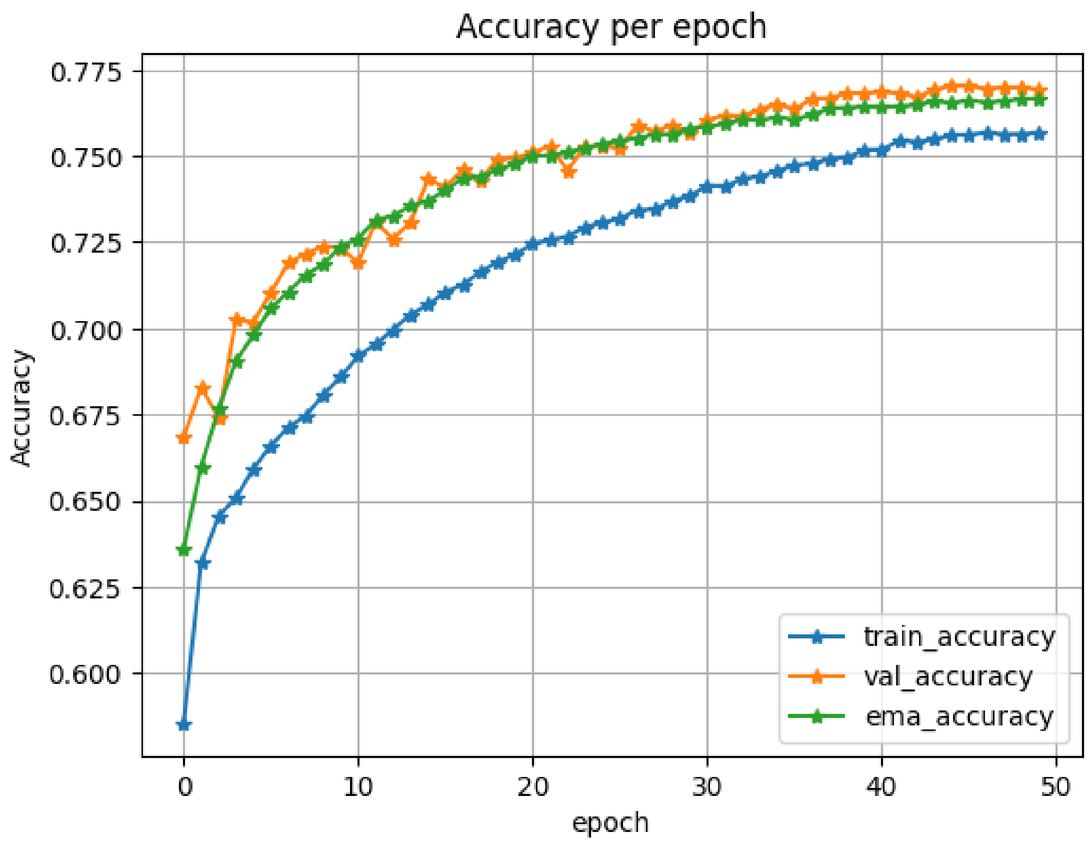


^ *Figure 1: Loss Per Epoch*

`



^ *Figure 2: Score Per Epoch*



^ *Figure 3: Accuracy Per Epoch*

References

[1] OpenAI. (2023). ChatGPT [Large language model]. https://chat.openai.com

[2] Amgad M, Elfandy H, ..., Gutman DA, Cooper LAD. Structured crowdsourcing enables convolutional segmentation of histology images. Bioinformatics. 2019. doi: 10.1093/bioinformatics/btz083