Name : Aditi Baheti
Roll No : M23CSA001
DL Assignment 3

Colab File : **M23SA001.ipynb**

Introduction :

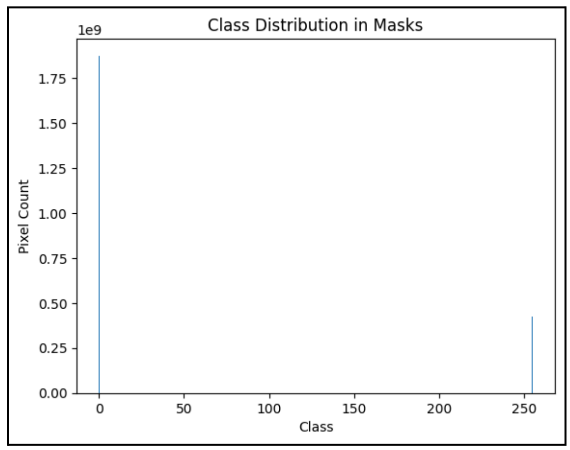
- Train a segmentation model using MobileNet pre-trained on ImageNet as encoder with a custom decoder for predicting segmented masks.

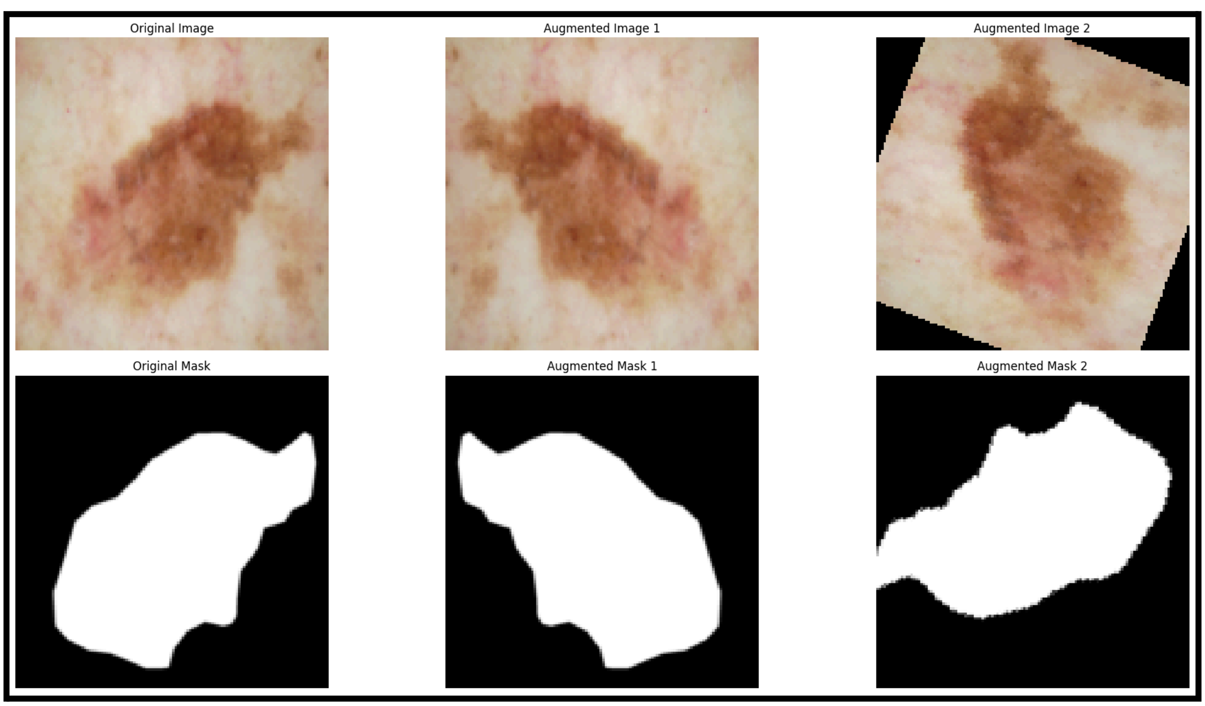
Dataset Description:

- The ISIC 2016 dataset consists of 900 training and 379 test dermoscopic lesion images along with their corresponding segmented masks.
- Each image is pre-processed and resized to 128x128 pixels, ensuring uniformity across the dataset for effective model training.

Pre-Processing and Data Augmentation

- Extensive exploratory data analysis revealed class imbalances and varying lesion sizes within the dataset.
- Based on these insights, a series of augmentations were applied to increase dataset diversity and robustness.
- Augmentations included random horizontal flips and rotations, which expanded the dataset to 3600 images.
- A custom data loader was implemented to manage image transformations and batch processing efficiently.



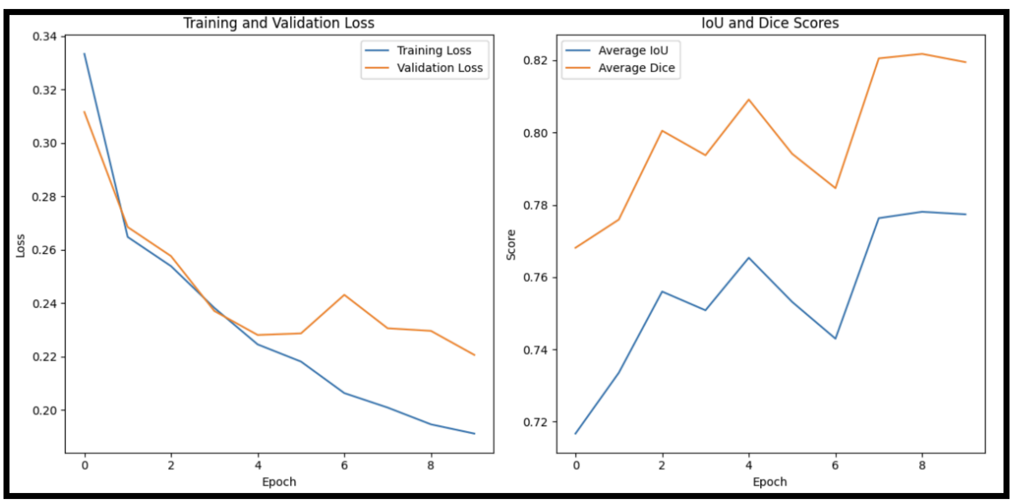


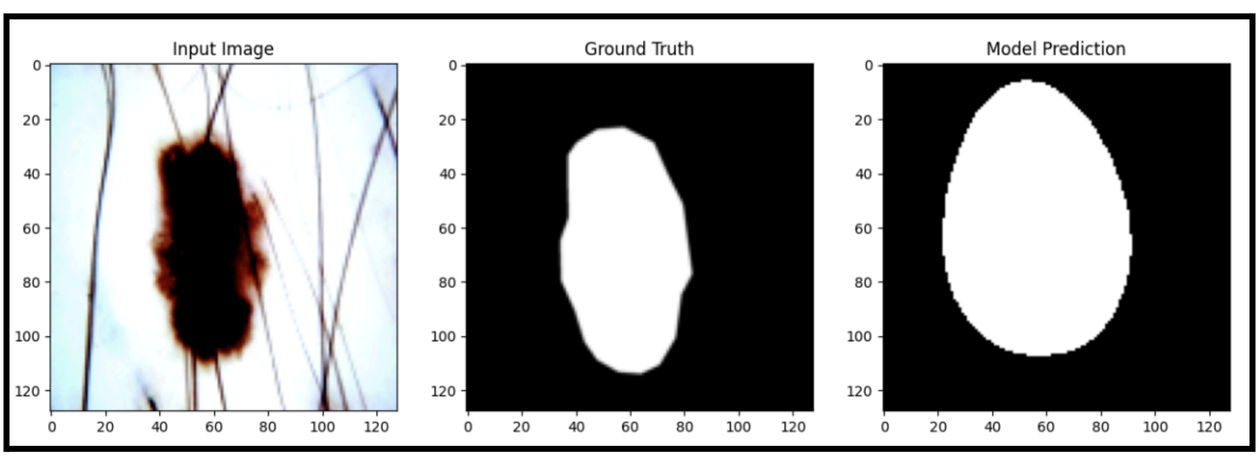
Experiment 1: Feature Extraction with Frozen Encoder

- In this phase, a custom decoder was integrated with a pre-trained MobileNet encoder, whose weights were kept frozen to leverage the existing learned features.
- The decoder architecture was designed with a series of upsample and refinement blocks to progressively reconstruct the segmentation masks from the encoded features.
- The decoder's upsample blocks consisted of transposed convolutions, which increased the spatial dimensions of the feature maps, while the refinement blocks, composed of standard convolutions and batch normalization, fine-tuned the details within the upsampled features.

Hyperparameter Set 1 :

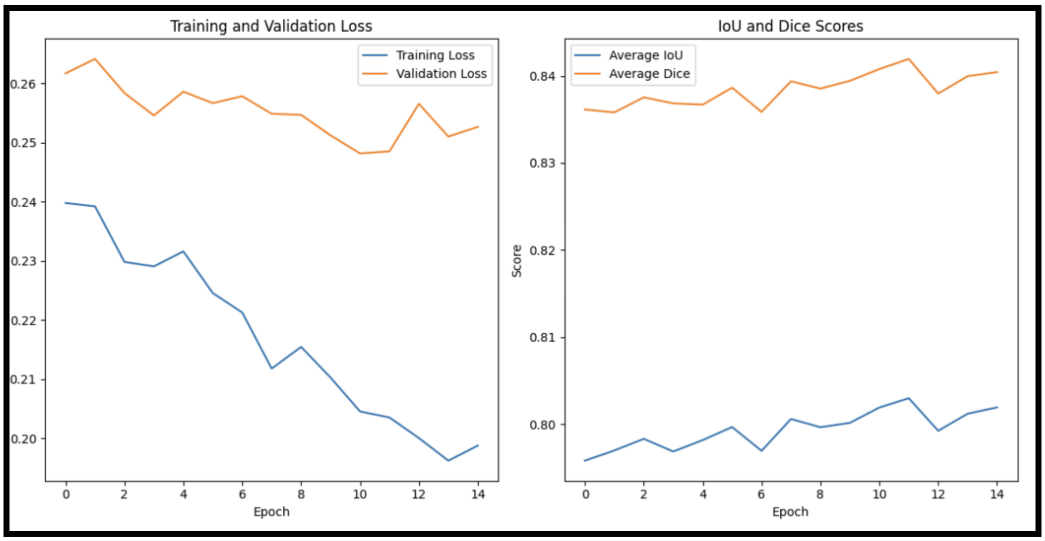
- Using Binary Cross-Entropy with Logits Loss (**BCEWithLogitsLoss**) and **RMSProp** optimizer. This configuration is standard for binary classification tasks where the output is a probability, and RMSProp is known for its adaptive learning rates.
- The training loss exhibited a consistent decline, indicating that the model was learning effectively from the data.
- The validation loss decreased in tandem with the training loss but showed some variability, suggesting the model was responsive to the validation data's nuances.
- The test loss was recorded at 0.2638, demonstrating the model's ability to generalize to unseen data.
- The average IoU of 0.7585 and Dice score of 0.8025 reflected the model's competency in accurately segmenting the images.

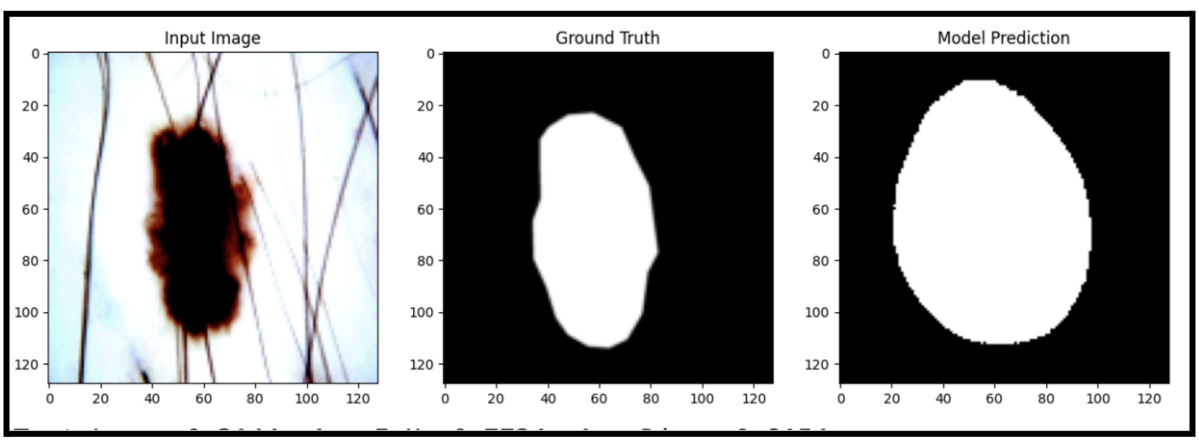




Hyperparameter Set 2 :

- Using a custom **IoU loss function** and the **Adam** optimizer. The IoU loss provides direct optimization of the segmentation overlap, and the Adam optimizer is recognized for its efficiency in various conditions due to its momentum and adaptive learning rate features.
- The training and validation loss trends showed improvement over epochs, though the validation loss displayed some fluctuations.
- This variation could indicate a better model fit to the validation set due to the direct optimization of the IoU metric.
- A test loss of 0.3144 was observed, **slightly higher than in hyperparameter set 1**, which could suggest that IoU optimization is more challenging to generalize.
- However, the model achieved a higher average IoU of 0.7724 and Dice score of 0.8154, indicating a marginal improvement in the segmentation's precision and recall.





For both configurations, the performance was quantitatively evaluated using the average Intersection over Union (IoU) and Dice scores.

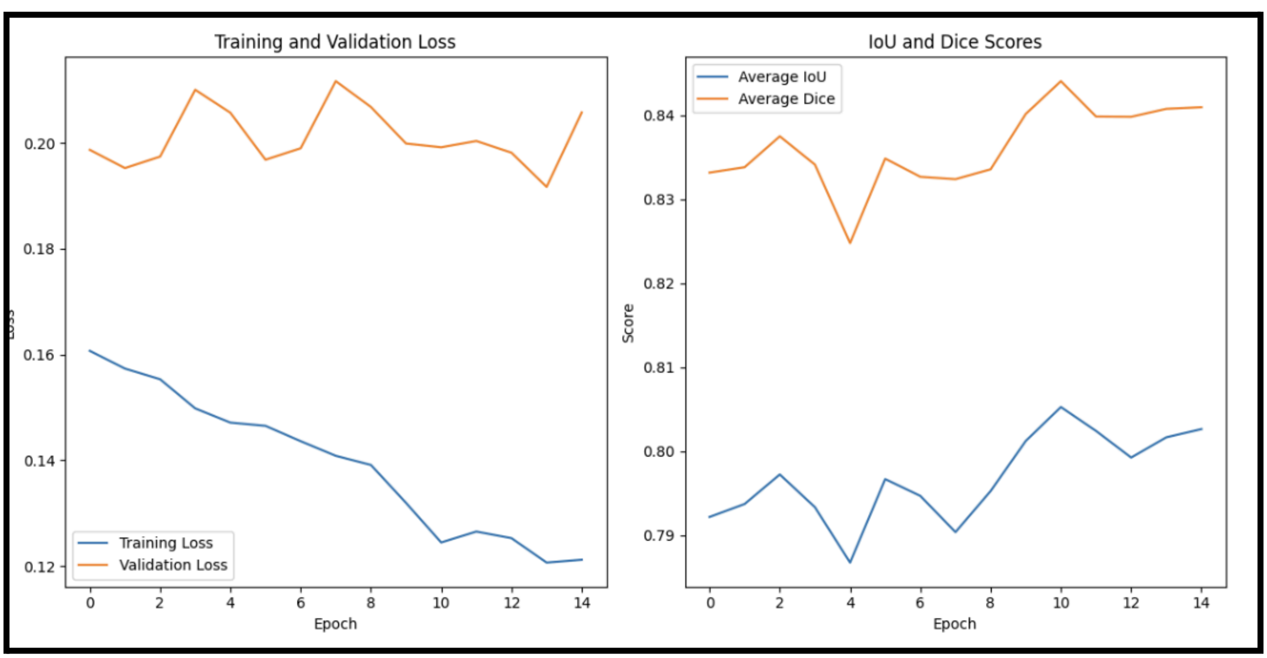
The **comparative analysis suggests that while BCEWithLogitsLoss paired with RMSProp provides a robust model with lower test loss, optimizing directly for IoU with Adam gives slightly better segmentation quality, as evidenced by the higher IoU and Dice scores.**

Experiment 2: Fine-Tuning the Entire Model

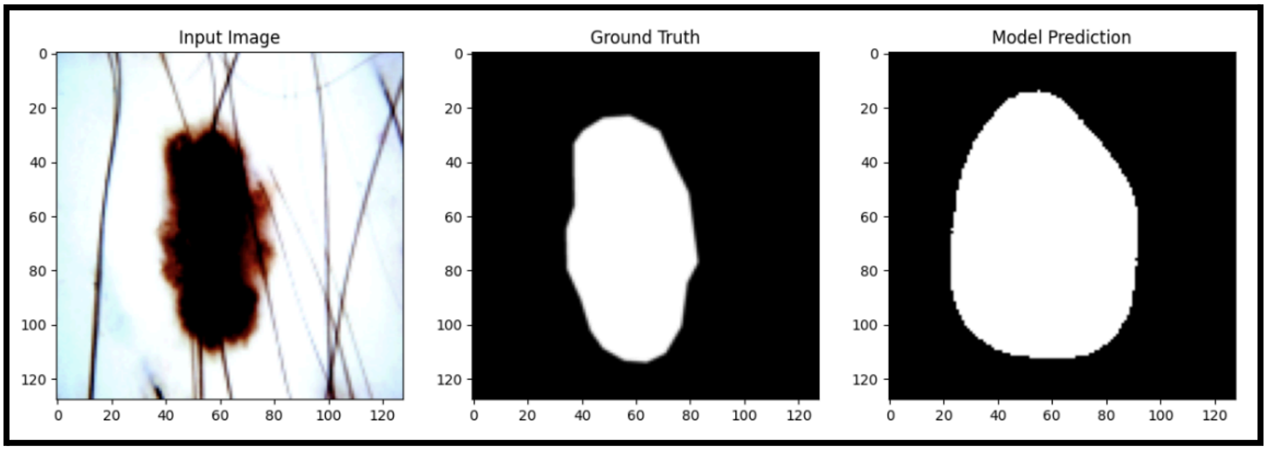
- Maintaining the same decoder architecture, the previously frozen encoder was unfrozen, allowing all weights in the network to be updated during training.
- This approach aimed to refine the pre-trained features specifically for the segmentation task at hand.
- The MobileNet V2 encoder was unfrozen, making its weights trainable to allow for end-to-end learning.
- This is based on the hypothesis that while pre-trained weights offer a good starting point, allowing them to adjust to The model was trained for 15 epochs, using Binary Cross-Entropy with Logits Loss as the criterion, which is a standard practice for binary segmentation tasks.
- The Adam optimizer was selected for its adaptive learning rate capabilities, known to be effective in converging deep learning models.the new data could potentially improve performance.

Performance Metrics :

- Training Loss and Validation Loss: The model exhibited a decreasing trend in training loss, indicating effective learning, while the validation loss fluctuated, suggesting that the model was learning new representations and adapting to the dataset's variability.
- Average IoU and Dice Scores: These scores increased over epochs, a positive indicator of the model's improving capability to accurately segment lesions .
- Test Loss: 0.2660, which suggests that the model has generalized well to unseen data.
- Average IoU: 0.7659, and Average Dice: 0.8094, both metrics indicating a high-quality segmentation performance.



- A visual inspection of the input image, ground truth, and model prediction shows that the model is capable of producing segmentation masks that are close to the ground truth, with boundaries that are well delineated



- Fine-tuning the entire model, including the encoder, yielded a noticeable improvement in the model's ability to segment skin lesions, outperforming the previous experiment where only the decoder was trainable.
- The use of Adam as an optimizer and the end-to-end training strategy were likely contributors to this enhanced performance.

Comparative Analysis of Segmentation Model Training Strategies

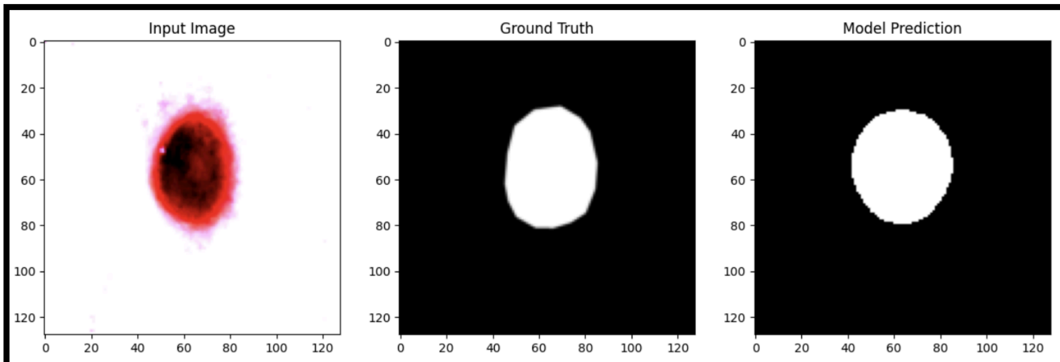
- **Training and Validation Loss: Method 2 exhibited a more significant reduction** in training loss over epochs compared to Method 1, suggesting a more effective learning process. Although both methods had fluctuations in validation loss, **Method 2's validation loss was generally lower**, indicating better generalization.

- **Average IoU and Dice Scores:** Both methods improved these metrics over time. However, **Method 2 achieved a slightly higher IoU and Dice score**, suggesting a more precise match between the predicted segmentations and the ground truth.

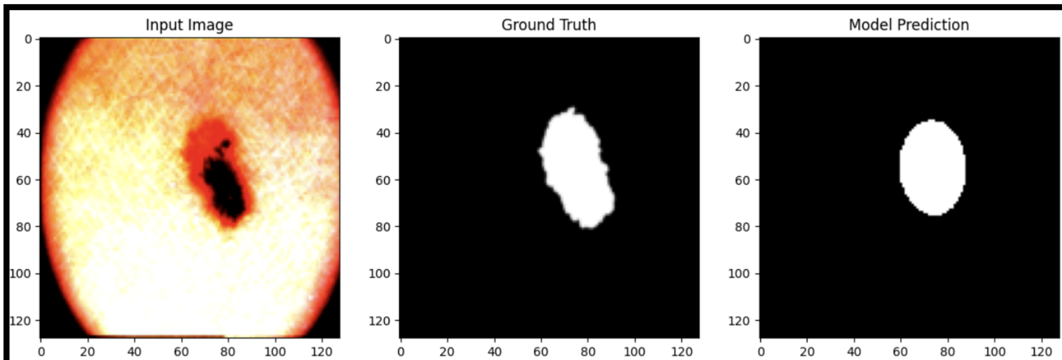
- **Test Loss: Method 2 again showed a marginally lower test loss**, further indicating its superior generalization capabilities.

- **Visual Assessment of Model Predictions:** The **predictions from Method 2 more closely resemble the ground truth masks**. This visual agreement reinforces the quantitative findings and suggests that fine-tuning the entire model leads to more accurate segmentations.

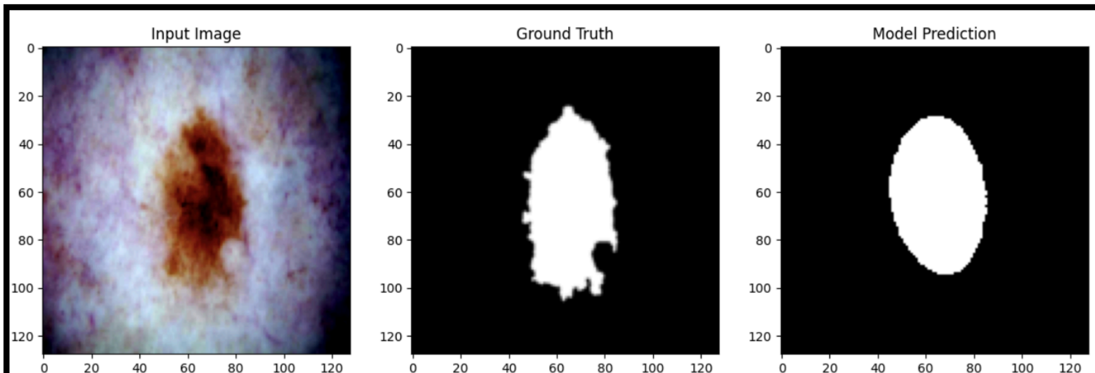
Following are some of the visualizations done between epochs to see how well model performs on validation data :

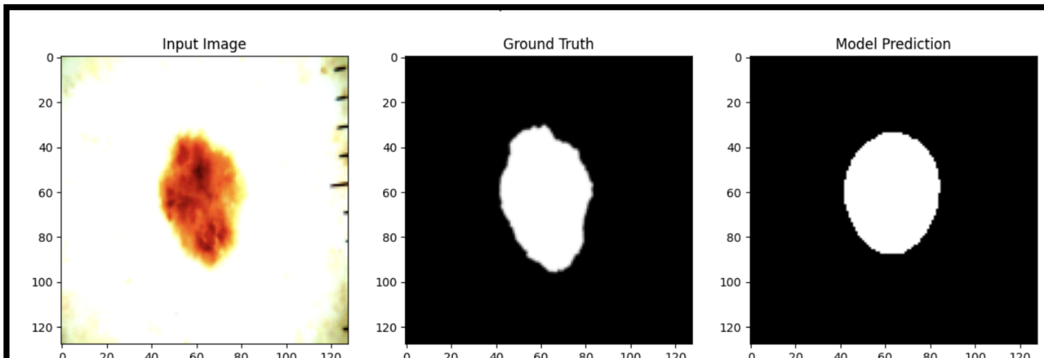


Epoch 2/15, Train Loss: 0.2392, Val Loss: 0.2642, Avg IoU: 0.7970, Avg Dice: 0.8358



Epoch 5/15, Train Loss: 0.2316, Val Loss: 0.2586, Avg IoU: 0.7982, Avg Dice: 0.8367





Epoch 13/15, Train Loss: 0.1253, Val Loss: 0.1982, Avg IoU: 0.7992, Avg Dice: 0.8398