Image Segmentation Using MobileNetV2

Sachin Maurya
Roll Number: CS24MS002
cs24ms002@iitdh.ac.in

November 28, 2024

Abstract

This project presents a deep learning-based approach for binary image segmentation using MobileNetV2 as the encoder and a custom decoder. Two models were developed and evaluated:

- Model 1 with a frozen encoder.
- Model 2 with a fine-tuned encoder.

The dataset included 900 training images and 379 test images with their segmentation masks. Key performance metrics include:

- Intersection over Union (IoU).
- Dice Score.

Results indicate that Model 2 outperforms Model 1 in accuracy and generalization.

Introduction

Image segmentation is a vital task in computer vision, with applications in:

- Medical imaging.
- Autonomous driving.
- Object detection.

This study explores:

- MobileNetV2 as a lightweight encoder for feature extraction.
- A custom decoder for pixel-wise segmentation.

Dataset (ISIC Dataset)

- **Training Set:** 900 images resized to 128×128.
- **Testing Set:** 379 images resized to 128×128.

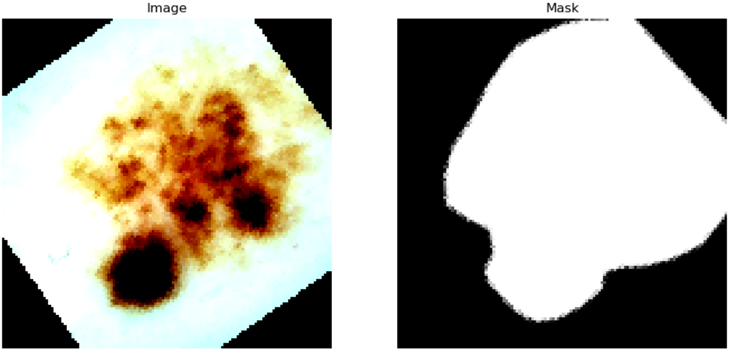

Figure: Visualization of input image and corresponding mask.

Model Architecture

- **Encoder:** MobileNetV2 pre-trained on ImageNet outputs feature maps of size $4 \times 4 \times 1280$.
- **Decoder:** Reconstructs feature maps to 128×128 resolution with:
  - Five convolutional layers.
  - Bilinear upsampling layers.
  - Batch normalization, ReLU activation, and dropout layers.

Model Parameters:

- Total params: 10,755,913.
- Trainable params: 7,251,041.
- Non-trainable params: 3,504,872.

Loss Functions

Model 1: Combined Dice Loss and Binary Cross-Entropy (BCE) Loss

$$\text{Loss} = 0.1 \cdot \text{Dice Loss} + \text{BCE Loss} \tag{1}$$

Model 2: Binary Cross-Entropy (BCE) Loss

$$\text{BCE Loss} = -\frac{1}{N} \sum_{i=1}^{N} [G_i \cdot \log(P_i) + (1 - G_i) \cdot \log(1 - P_i)] \tag{2}$$

Evaluation Metrics:

- Intersection over Union (IoU).
- Dice Score.

Results and Analysis

Model 1 (Frozen Encoder):

- IoU: 0.6354.
- Dice Score: 0.7675.
- Train Loss: 0.2911.
- Validation Loss: 0.2899.

Model 2 (Fine-Tuned Encoder):

- IoU: 0.6819.
- Dice Score: 0.7972.
- Train Loss: 0.2117.
- Validation Loss: 0.2169.

Comparative Analysis

Table: Comparison of Metrics for Model 1 and Model 2

Metric	**Model 1**	**Model 2**
IoU	0.6354	0.6819
Dice Score	0.7675	0.7972
Train Loss	0.2911	0.2117
Validation Loss	0.2899	0.2169

Radar Chart

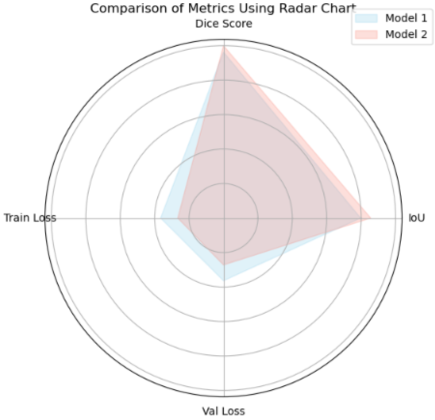

Figure: Radar chart comparing Model 1 and Model 2 across metrics.

Conclusion

- Model 2 consistently outperforms Model 1 across all metrics.
- Future work includes:
  - Exploring advanced loss functions to enhance segmentation performance.
  - Incorporating multi-scale feature aggregation techniques.

References

- A. Kanadath, J. A. Arul Jothi, and S. Urolagin, "Histopathology Image Segmentation Using MobileNetV2 based U-net Model," *2021 International Conference on Intelligent Technologies (CONIT)*, Karnataka, India, June 25–27, 2021.
- G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan, "Medical Image Segmentation based on U-Net: A Review," *Journal of Imaging Science & Technology*, vol. 64, no. 2, pp. 1–12, 2020.