Training Requirements & Optimization Report

Training Requirements

Model Architecture: Modified 2D U-Net
Framework: PyTorch with MONAI
Training Mode: Slice-based (axial, sagittal, or coronal)
Loss Function: Dice Loss (binary or categorical depending on segmentation setup)

Resource	Requirement
GPU	NVIDIA Tesla T4 or equivalent (6–8GB VRAM)
CPU	2–4 cores sufficient for pipeline and I/O
RAM	~12 GB minimum
Storage	~4 GB for dataset and checkpoints
Training Time	~1 hour for 1000 steps on T4 GPU

Training Time ~1 hour for 1000 steps on T4 GPU

Suggestions for Making Training More Efficient

1. Model Architecture

- Consider switching to a **3D U-Net** to leverage volumetric spatial context, especially for anatomical continuity.
- Use **depthwise separable convolutions** or **efficient blocks** (e.g., Residual or Attention gates) to reduce computation without losing accuracy.
- Consider **group normalization** instead of batch norm for small batch sizes.

2. Data Pipeline

- Use **region-of-interest (ROI) cropping** to eliminate irrelevant background and focus on hippocampal regions.
- Optimize data loading with **pre-caching** and **parallel workers** using `num_workers > 0` in `DataLoader`.

3. Loss Functions

- Combine **Dice Loss + Cross-Entropy** for more stable gradients.
- Try **Focal Loss** if encountering class imbalance between hippocampus and background.

4. Data Augmentation

- Effective:
  - **Random flips** (horizontal/vertical)
  - **Elastic deformations**
  - **Small rotations** (+/- 10 degrees)
- Not Useful:
  - **Color jitter / brightness/contrast** — not relevant in grayscale MRI
  - **Heavy affine transforms** — can distort anatomical shapes
  - **Gaussian noise** — may degrade anatomical boundary visibility

---

Performance Analysis: Best vs Worst Volumes

Best Performing Volumes

- Typically mid-sagittal slices with clear, high-contrast boundaries between the hippocampus and surrounding tissue (e.g., CSF).
- Slices where both anterior and posterior sections are fully visible within frame.

Worst Performing Volumes

- Edge slices where the hippocampus is partially or not fully present, leading to false negatives.

- Volumes with poor contrast or motion artifacts — model struggles with subtle textures and ambiguous outlines.

Possible Causes

- The model’s receptive field may be too small to capture larger anatomical context in edge slices.
- Label noise or inter-rater variability in ground truth could also contribute to discrepancies.

Summary

This version of the model achieves acceptable performance under constrained compute settings. Efficiency gains can be made by transitioning to 3D architectures, improving ROI focus, and refining loss functions. Augmentations should remain anatomically plausible to preserve clinical relevance.