MEDICAL IMAGE GENERATION

Generating the X-ray, MRI, CT-scan, ULTRA sound Images
using by Diffusion Model

TEAM - 29

DEEP NEURAL NETWORKS

TEAM MEMBERS

CH. VAMSHI KRISHNA BABU - SE22URAI038

CH. VAMSI KRISHNA - SE22UARI039

G.RITHVIK - SE22UARI144

V. RAJA - SE22UARI178

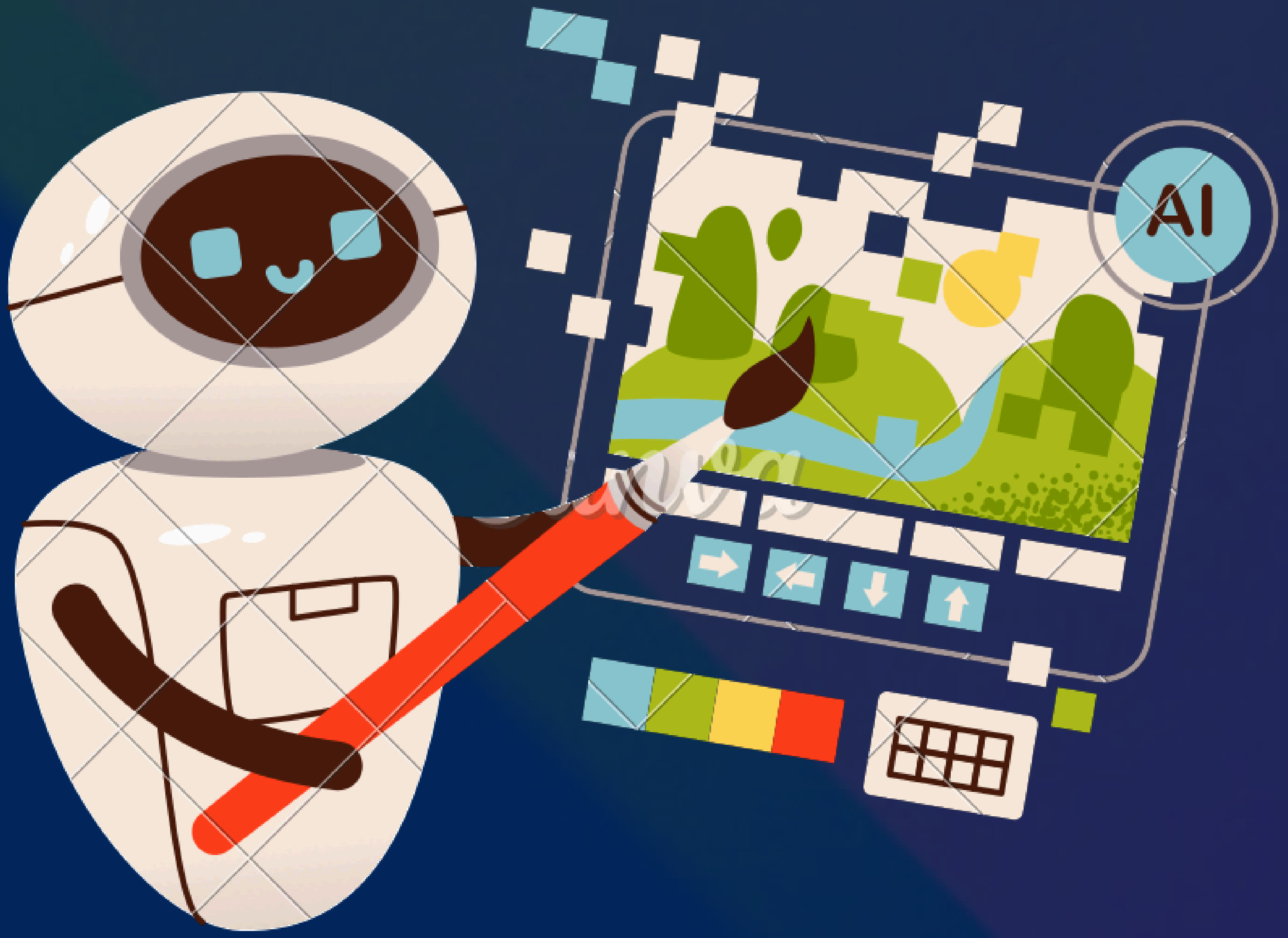
K. BHUVAN - SE22UARI212

00

CONTENT

PROBLEM STATEMENT

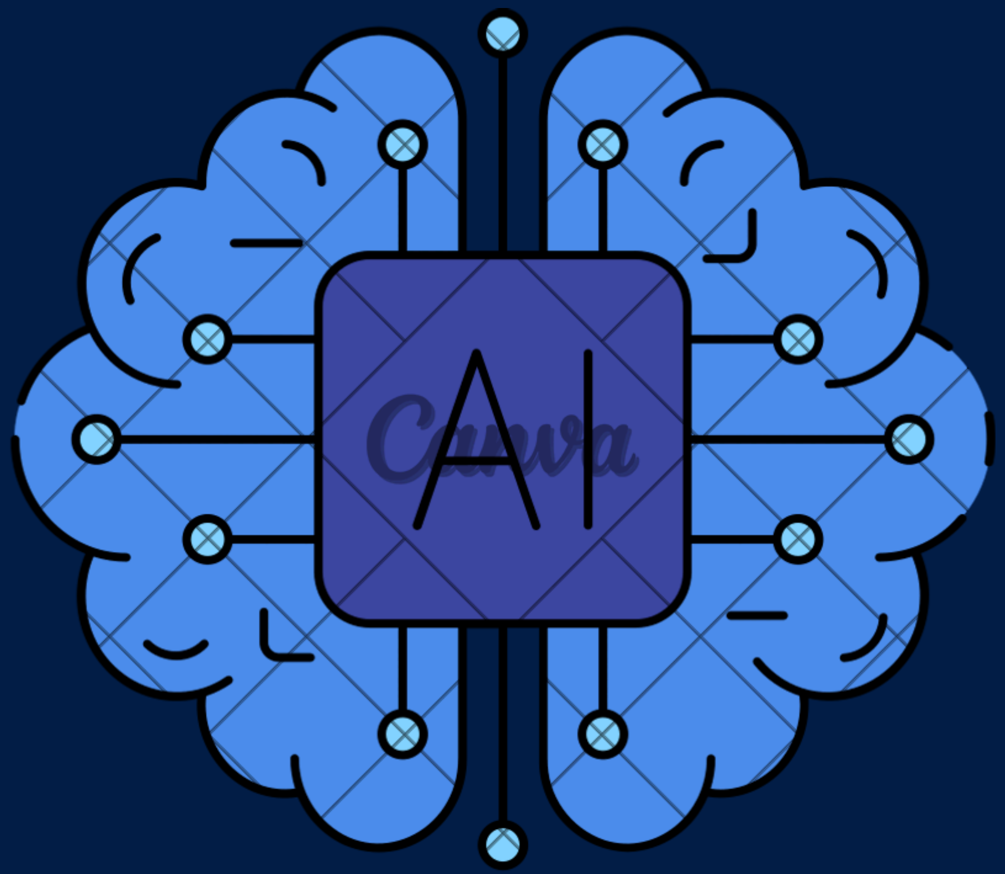
Developing the medical images generation model to synthesizing the X-ray, CT-scan, MRI, UltraSound Images to support the medical image research and application.



02

PROPOSED SOLUTION

Implement a Denoising Diffusion Probabilistic Model (DDPM) for training and a Denoising Diffusion Implicit Model (DDIM) for faster inference, using a UNet-based neural network to predict noise in the denoising process. The model will be trained on a dataset of grayscale CT scan images, resized to 128x128 pixels, and evaluated using SSIM to measure similarity to real images. The solution aims to generate realistic images that can augment medical imaging datasets.



DATASETS

X-RAY Image Dataset

MRI Images Dataset

CT-Scan Image Dataset

UltraSound Image Dataset

Taken Images are normal images without consists of diseases so that we cannot get the different image of diseased one.

MODEL ARCHITECTURE

04

```
[4]: from diffusers import UNet2DModel

model = UNet2DModel(
    sample_size=128,
    in_channels=1,   # 1 for grayscale MRI, 3 for RGB MRI
    out_channels=1,  # 1 for grayscale MRI, 3 for RGB MRI
    layers_per_block=3,
    block_out_channels=(128,128,256,256,512,512),
    down_block_types=(
        "DownBlock2D", "DownBlock2D", "DownBlock2D", "DownBlock2D", "AttnDownBlock2D", "DownBlock2D",
    ),
    up_block_types=(
        "UpBlock2D", "AttnUpBlock2D", "UpBlock2D", "UpBlock2D", "UpBlock2D", "UpBlock2D",
    ),
)
```

- The diffusion model defines the overall process of noise addition (forward) and removal (backward), including the noise schedule.
- The UNet2DModel is the neural network architecture used within the diffusion model to perform the denoising task in the reverse process.

WHY UNET2DMODEL IS PART OF THE DIFFUSION MODEL ?

- Functional Role: The diffusion model requires a neural network to learn the denoising function. The UNet2DModel serves this role by predicting the noise at each step of the reverse process.

PROJECT PIPELINE

06

1. **Data Preparation:** Load and preprocess images (resize, convert to tensors, normalize) using a custom dataset and DataLoader (batch_size=10) .
2. **Model Definition:** Define a UNet architecture (UNet2DModel) with parameters suited for grayscale images.
3. **Training:** Train the UNet within a DDPM framework to predict noise, using 5000 epochs and MSE loss.
4. **Model Saving:** Save the trained model for reuse.
5. **Inference:** Reload the model, use a DDIM pipeline with 20 steps to generate 4 new images from random noise.
6. **Visualization:** Display the generated images in a 1x4 grid using matplotlib.
7. **Evaluation:** Compute SSIM scores to compare generated images with ground truth, assessing perceptual similarity.

EVALUATION METRICS

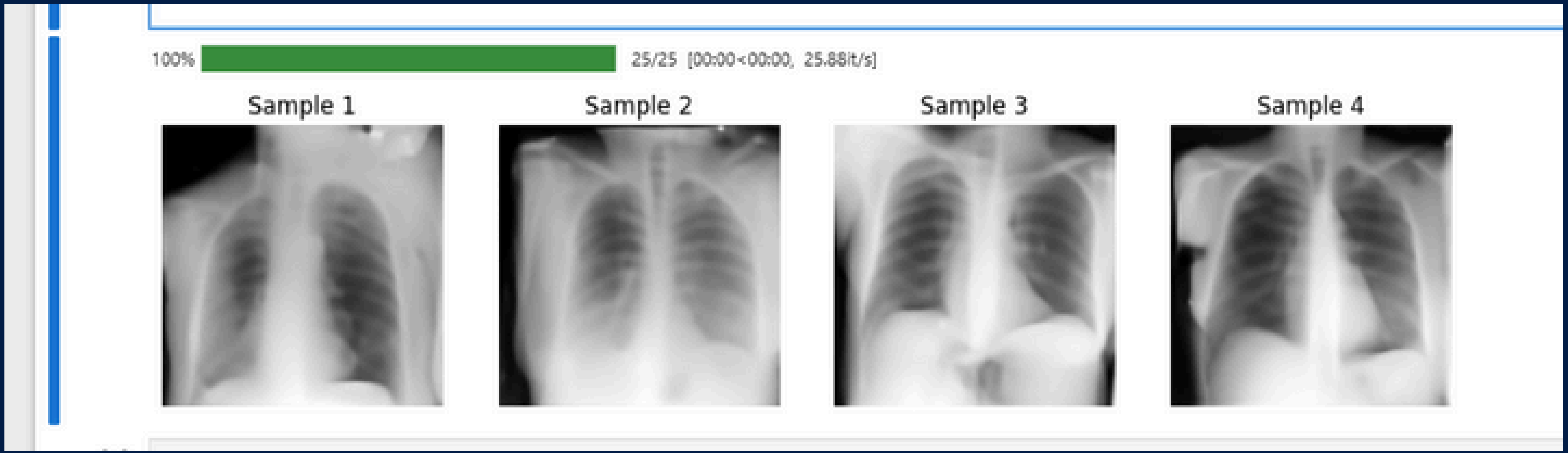
07

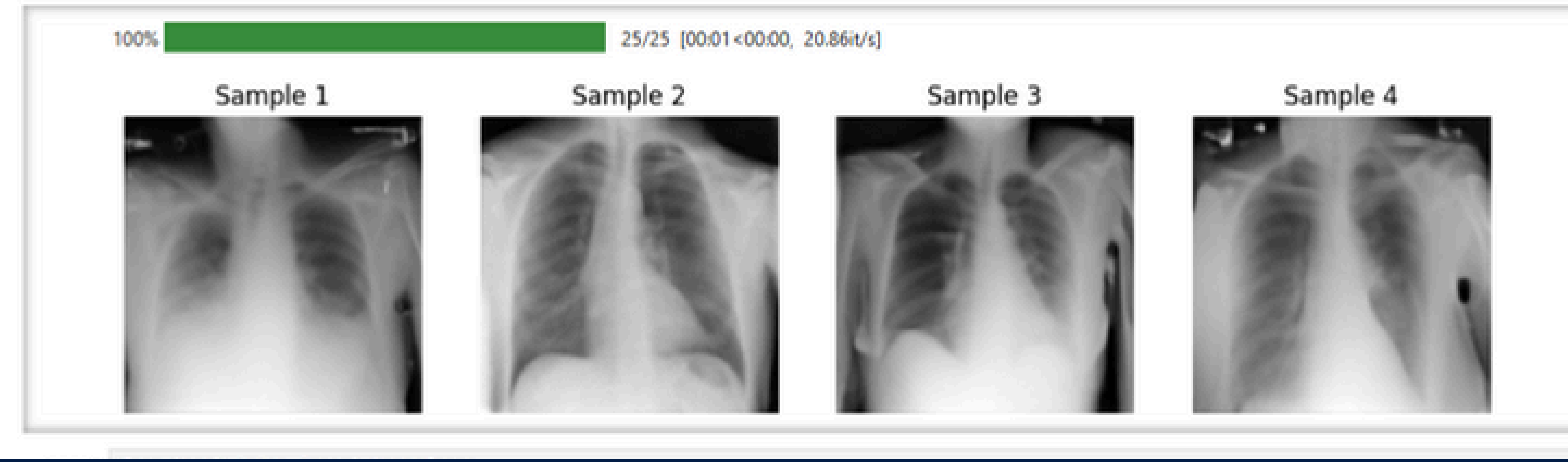
SSIM (Structural Similarity Index) is a metric that quantifies the similarity between generated images and ground truth by considering three key components: luminance (brightness), contrast, and structure.

SSIM Score range from -1 to 1 .

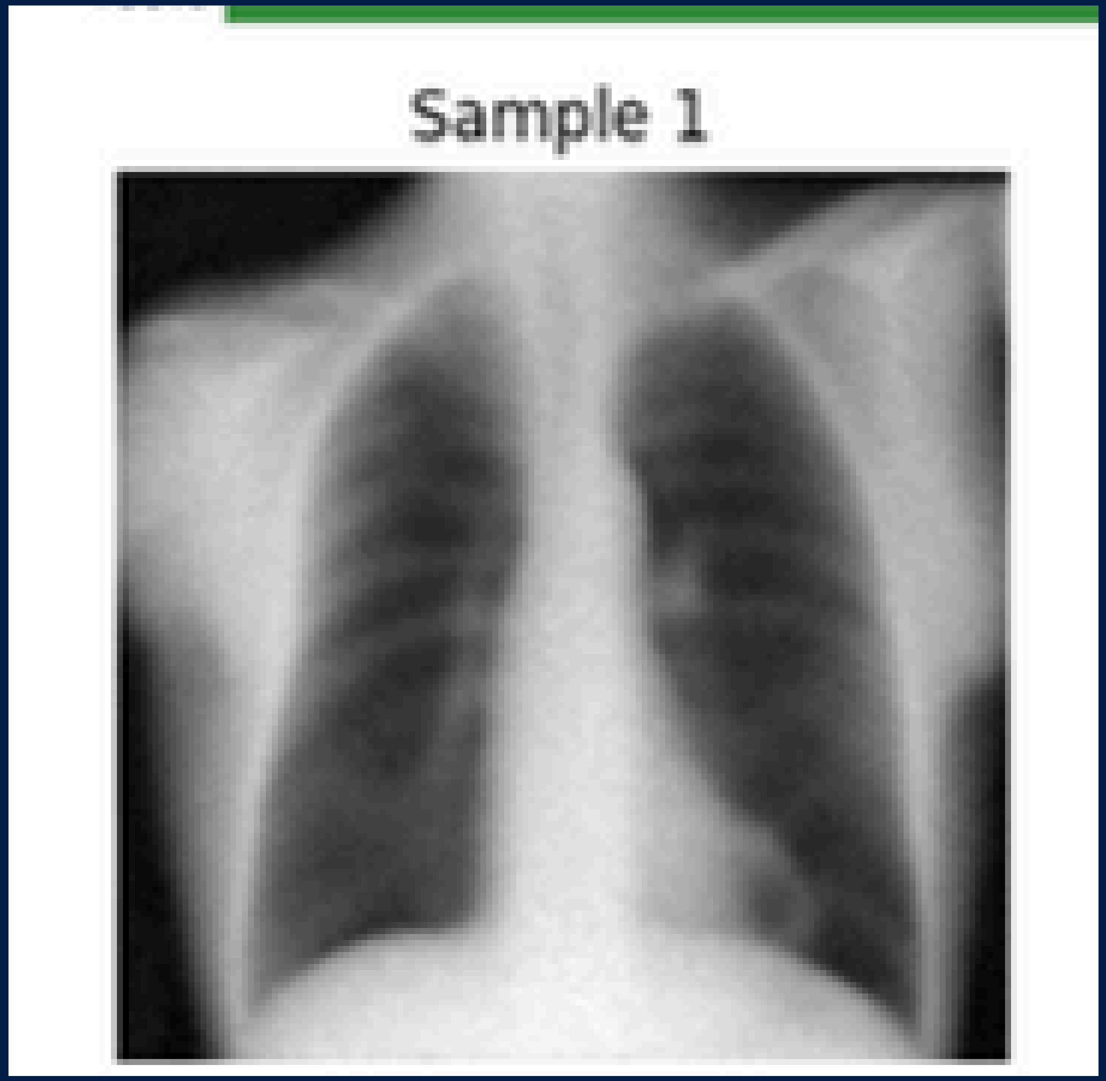
SSIM Range	Interpretation	Action Needed?
0.7 – 1.0	Excellent similarity	No (well-trained)
0.4 – 0.7	Moderate similarity	Maybe
0.1 – 0.4	Low similarity	Yes
< 0.1	Very poor similarity	Yes (retrain/tune)

X-RAY IMAGES

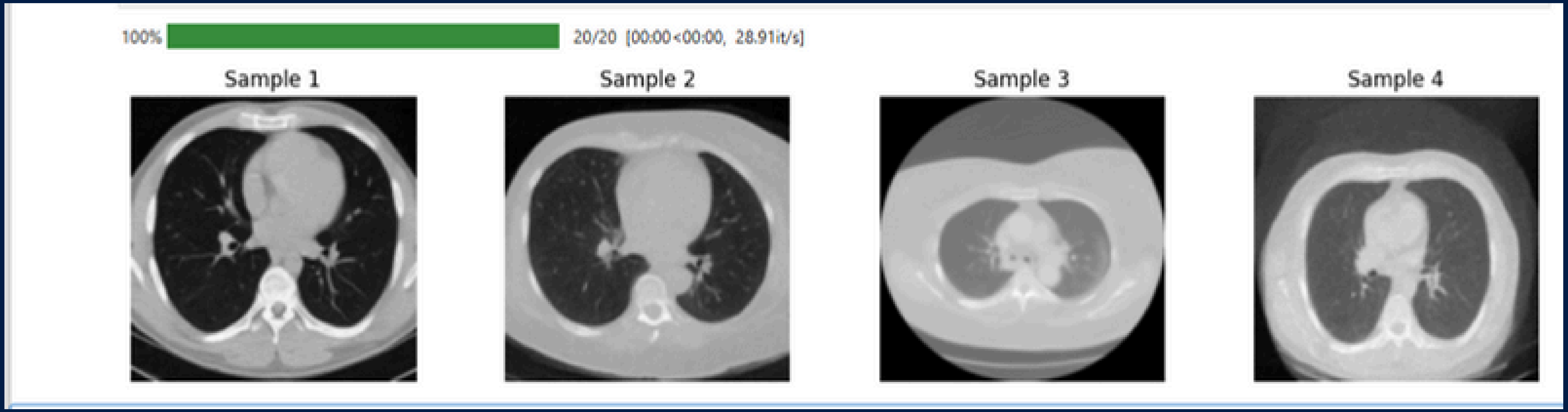




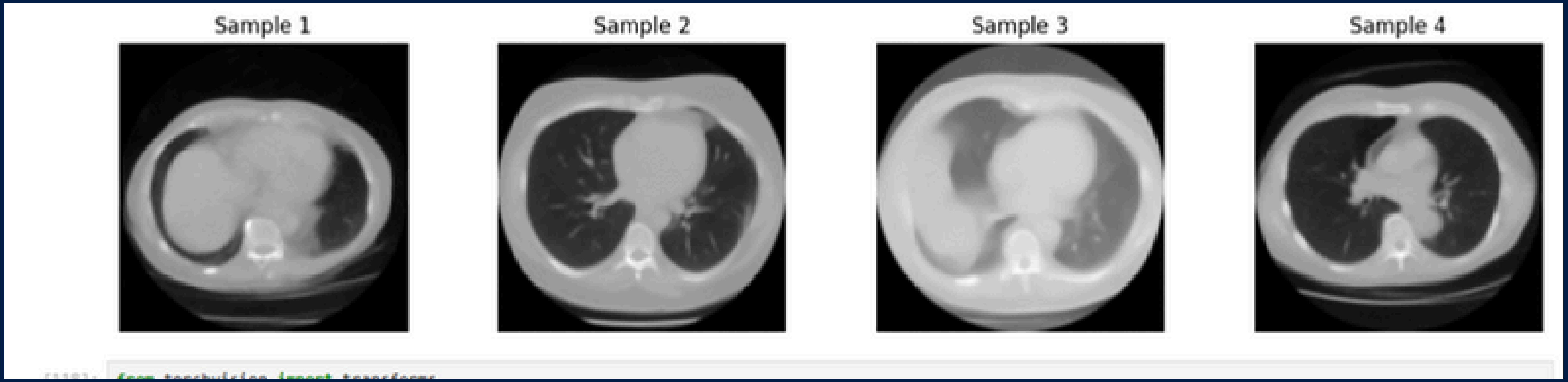
Dataset Size - 4999
Epochs - 50
SSIM value - 0.4600



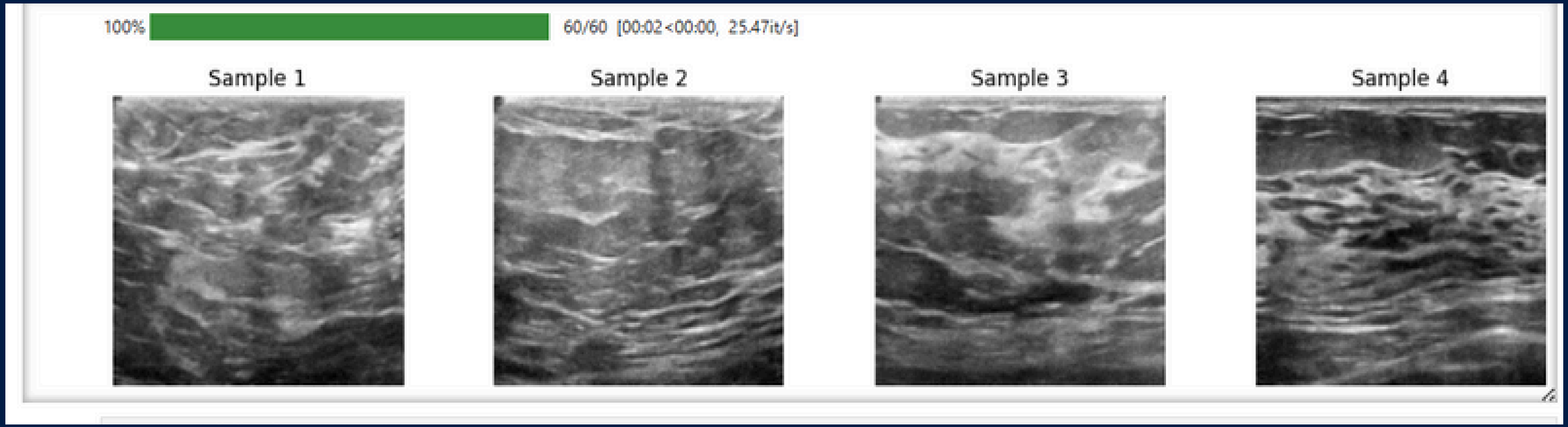
MRI IMAGES

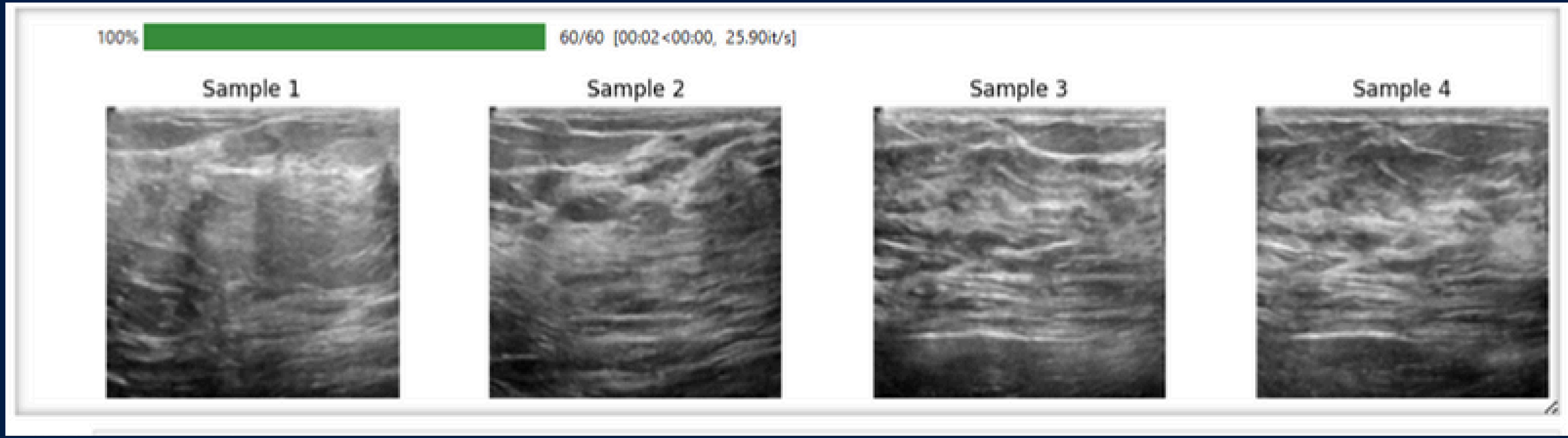


Dataset Size ~ 2700
Epochs - 1000
SSIM value - 0.4172

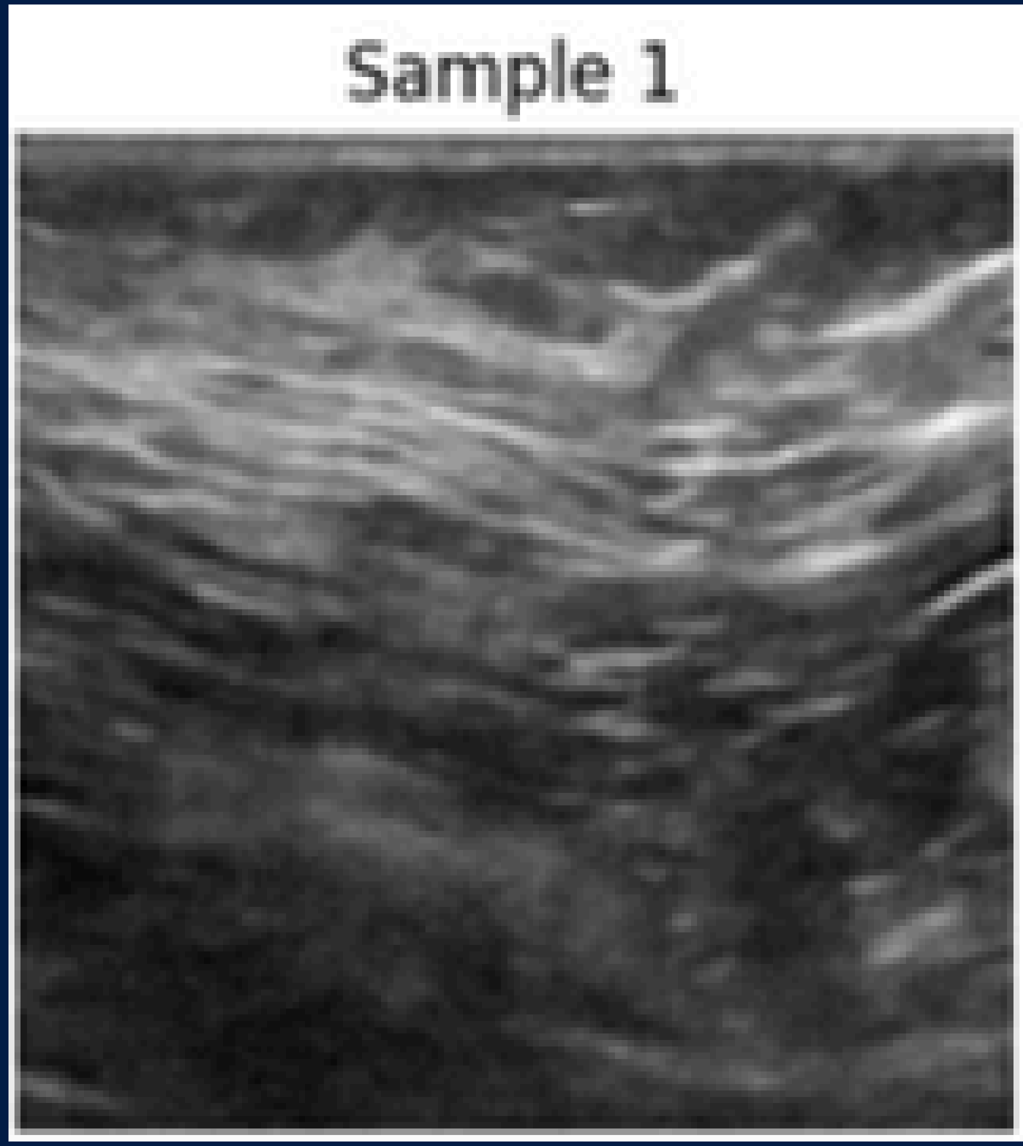


ULTRA SOUND IMAGES

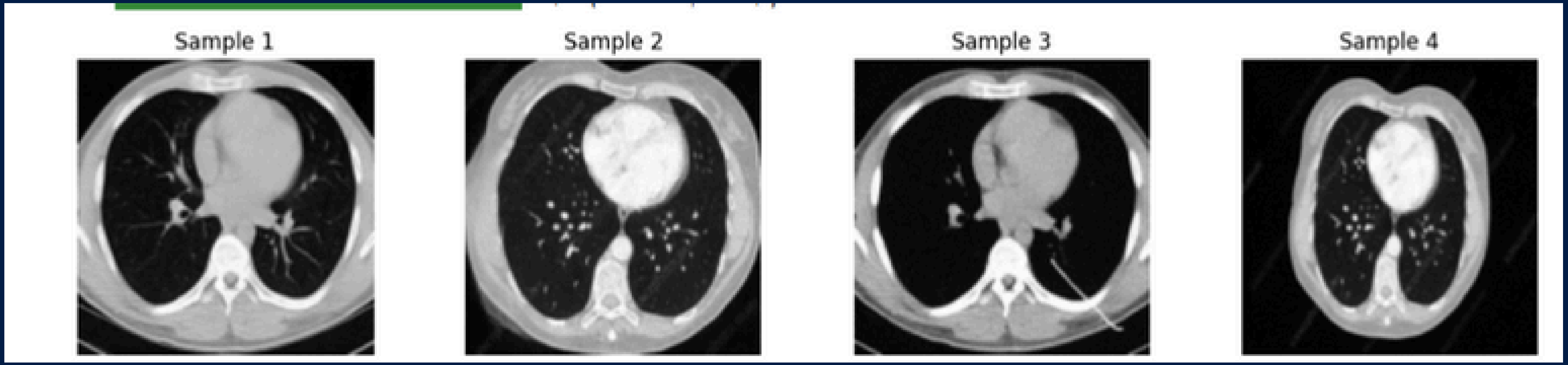


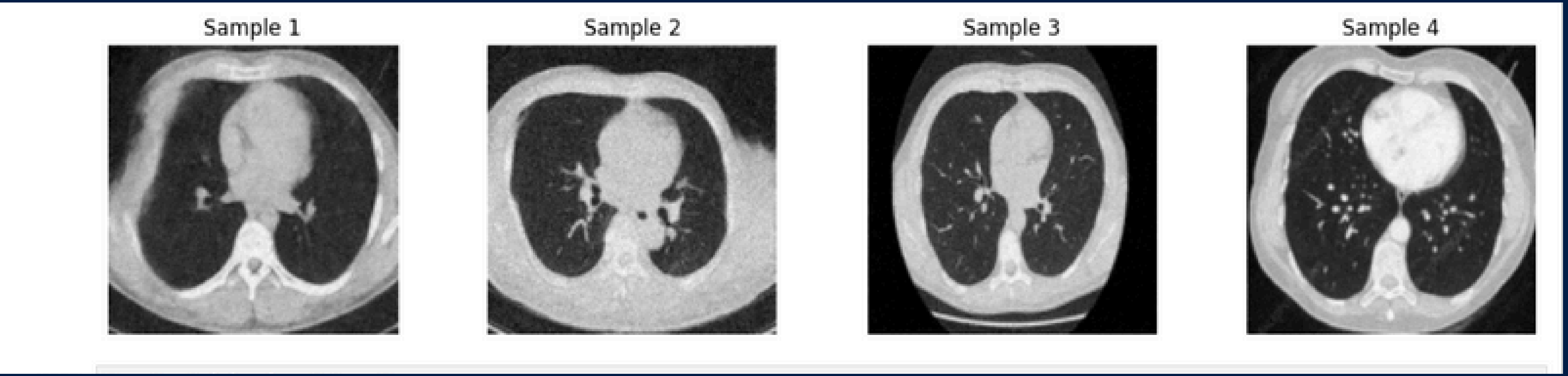


Dataset Size ~ 200
Epochs - 5000
SSIM value - 0.7664



CT-SCAN IMAGES





Dataset Size ~ 230
Epochs - 5000
SSIM value - 0.4056

CHALLENGES FACED

Memory Constraints: Out of Memory errors from the complex UNet and 2000-timestep schedule necessitated a small batch size (10), increasing training duration.

Training Efficiency: Training for 5000 epochs was computationally intensive; no learning rate scheduling or early stopping reduced efficiency.

Inference Speed and Efficiency: Slow inference due to the iterative nature of diffusion models (even with DDIM's 20 steps) challenged practical deployment; balancing speed and image quality was difficult on limited hardware.

CONCLUSION

The project, which implemented a diffusion model to generate high-quality X-ray, MRI, UltraSound, CT scan images using PyTorch and the diffusers library, successfully addressed the core objectives outlined in the problem statement but faced notable challenges that impacted its overall effectiveness. By leveraging a Denoising Diffusion Probabilistic Model (DDPM) for training and a Denoising Diffusion Implicit Model (DDIM) for inference, with a UNet-based architecture, the project demonstrated a robust technical approach to synthesizing grayscale Images (128x128 pixels) for potential medical imaging applications.

THANK YOU

TEAM - 29

DEEP NEURAL NETWORKS