IE643 Course Project

Generative AI for Data augmentation in medical images

Vishnu

MVS Sri Harsha [1] (24M2151),Ritesh Sur Chowdary [2] (24M2154)

24m2151@iitb.ac.in[1],24m2154@iitb.ac.in[2]

IE 643
Deep Learning: Theory and Practice

Abstract

Recent deep learning advances, especially in diffusion models, have greatly improved medical image generation, addressing data scarcity in liver cancer diagnosis. In this project, we explored Variational Autoencoder (VAE) and Denoising Diffusion Probabilistic Models (DDPM) as a generator model on the LiTS dataset, finding DDPM superior in terms of Frechet Inception Distance (FID), Structural Similarity Index Measure (SSIM) and Sliced Wasserstein Distance (SWD) metrics. The U-Net segmentation model, which segments the generated images, demonstrated effective results.

Introduction

This project aims to solve liver cancer diagnosis by generating synthetic liver tumor images for data augmentation using the LiTS dataset. Liver imaging is challenging due to low contrast and complex structures, and limited datasets make data augmentation essential. Diffusion Models (DDPM) are explored for generating diverse, high-quality images, outperforming traditional methods like VAE and GAN. The project aims to enhance liver tumor detection task by creating quality data for model training, using performance metrics and interactive tools for visualization and testing.

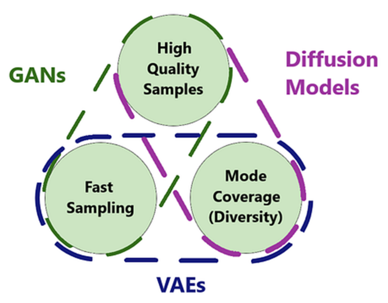

Figure 1: This figure shows the trade-off between VAE, GAN and Diffusion models.

We outline our major **contributions** below:

- We have pre-processed the dataset by reducing the number of images for efficient computation and utilizing only those images with ample amount of liver segments.
- Trained DDPM and generated fruitful results. Trained U-Net model to segment out the liver segments from the images. The segmentation model along with the generative model provides us the image, labels pairs.
- Evaluated performance of the proposed model for generation and segmentation tasks.
- Prepared an interactive tool.

Workflow

Figure 2 illustrates the workflow. The figure describes the work done in each phase.

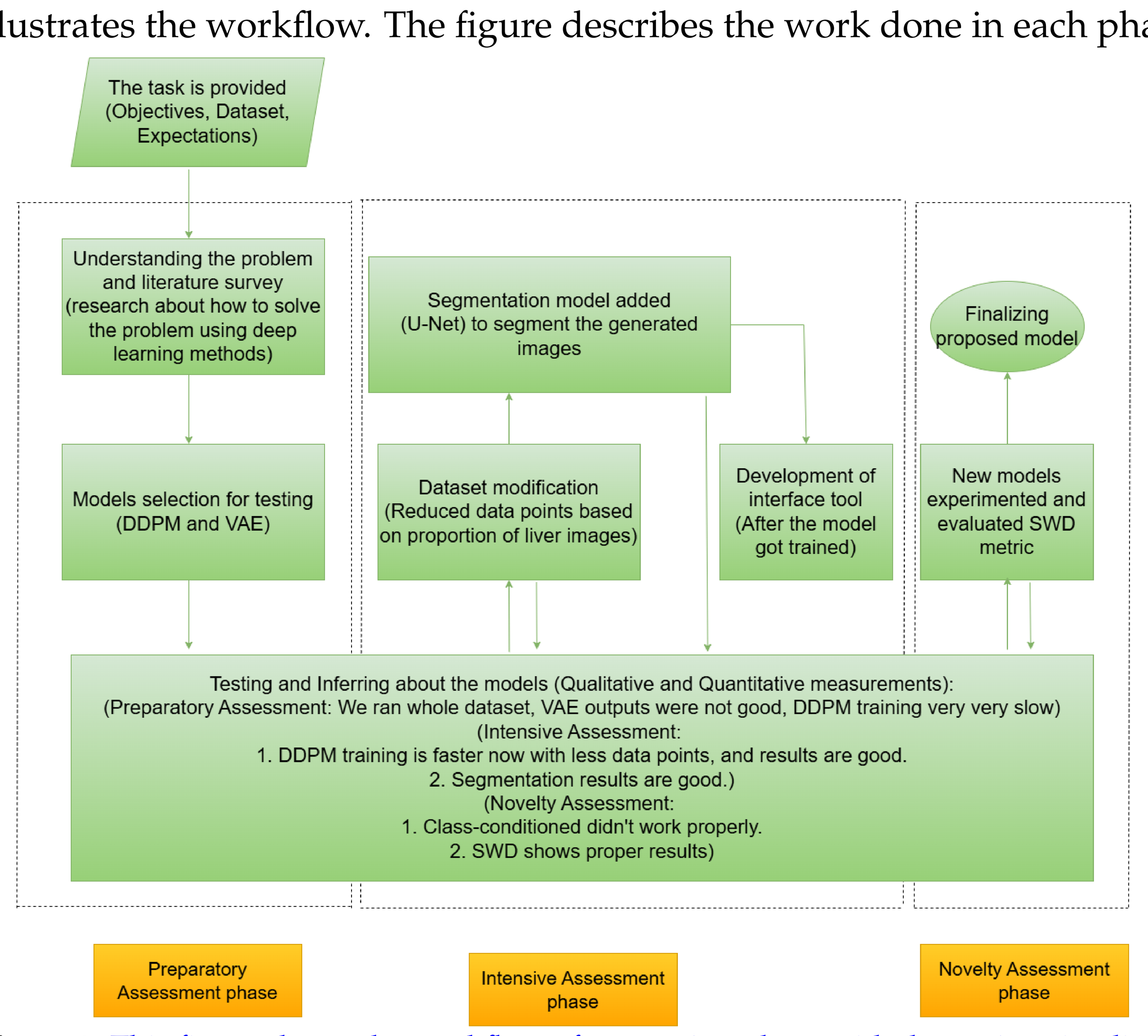

Figure 2: This figure shows the workflow of our project along with the project timeline.

Methodology

Figure 3 illustrates the architecture. The proposed methodology uses image generation and segmentation modules using DDPM and U-Net models respectively.

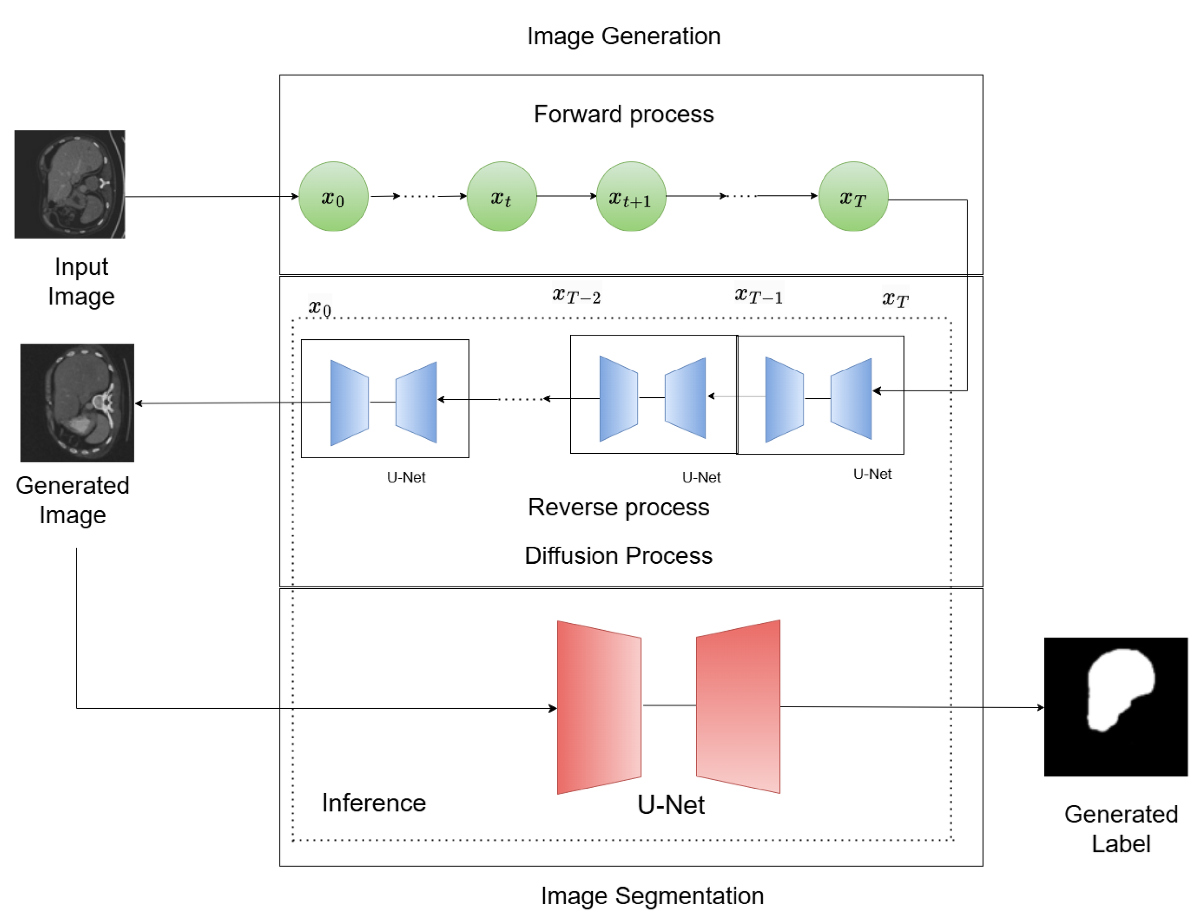

Figure 3: This figure shows the model used for this task. We used DDPM model for Image Generation task followed by U-Net model for Image Segmentation task

Dataset Details

The LiTS dataset [1] contains 3D CT scan volumes for liver tumor segmentation, with 131 train and 70 test volumes in NifTI format, each sized $512 \times 512 \times N_i$. Pre-processing involved clipping voxel values between -100 and 300, scaling intensities to 0-255, applying gamma correction, and merging liver and tumor labels into one class. The resulting 5791 2D grayscale images ($1 \times 128 \times 128$) focus on liver regions with intensity values from 0 to 255.

Novelty Assessment

During the Novelty Assessment, we have attempted the following tasks:

- We have tried class-conditioning DDPM (tumor-based), which didn't work, as the tumor detection is difficult in this dataset and scaling the model efficiently is difficult for higher dimension.
- We have added SWD metric as performance metric, which performs good for medical datasets and doesn't use pre-trained classifier, avoiding bias towards dataset.
- We have shown the SSIM and SWD metrics in the interactive tool.

Results

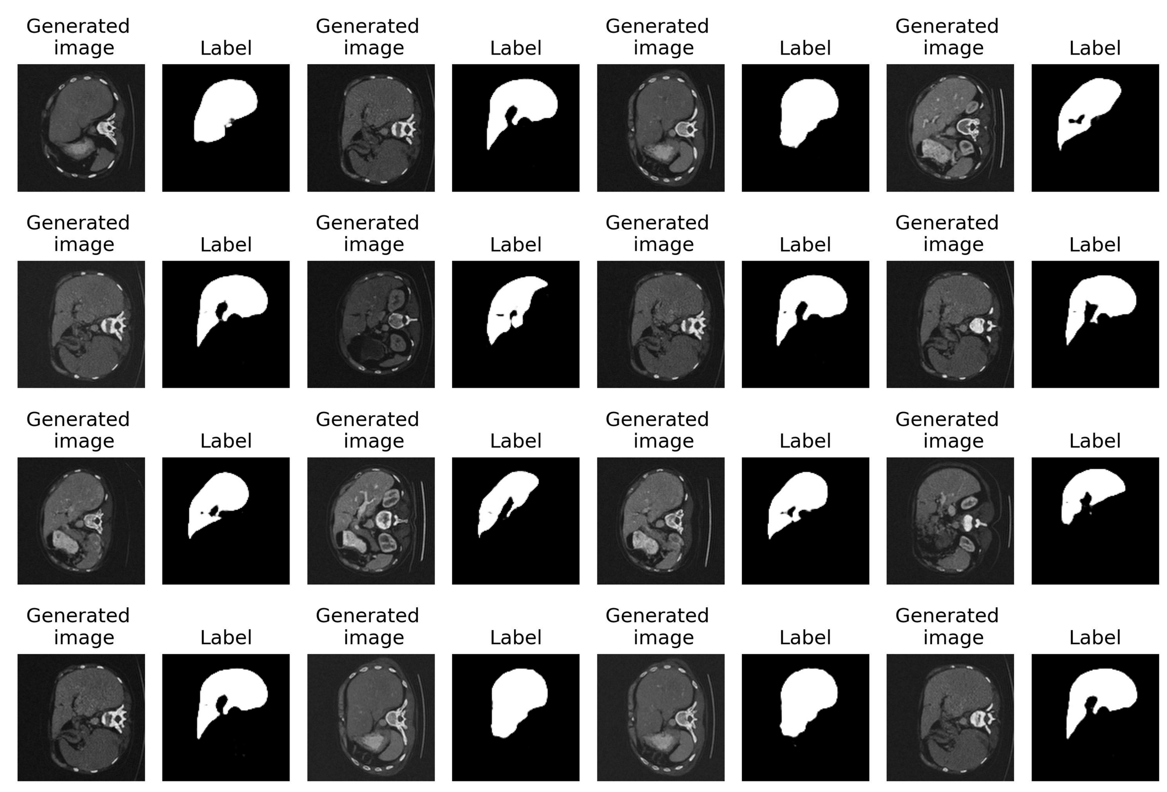

Figure 4: This figure shows the generated images using DDPM model and the corresponding labels generated by U-Net.

Model	FID Score	SSIM	SWD [3]
VAE	212.35	0.32	0.15
DDPM [2]	42.51	0.37	0.13

Table 1: Performances between the VAE and DDPM models in terms of FID score, SSIM and SWD.

Figure 4 and table 1 showcase the qualitative and quantitative outcomes of the generator model, which show the outstanding performance of DDPM model on the dataset. The liver segmentation on the input dataset by the segmentation model is almost perfect.

Conclusion

- **Objective:** The project aims to augment liver tumor datasets by generating synthetic liver images using generative models.
- **Methodology:** A generative module (DDPM and VAE) and a segmentation module (U-Net) are used, with DDPM yielding better image quality than VAE.
- **Results and Future Works:** DDPM produces high-quality images, demonstrated by better FID score, SSIM and SWD. The Tumor-based class-conditioning can be exercised in future.

References

[1] Patrick Bilic, Patrick Christ, Hongwei Bran Li, Eugene Vorontsov, Avi Ben-Cohen, Georgios Kaissis,Adi Szeskin, Colin Jacobs, Gabriel Efrain Humpire Mamani, Gabriel Chartrand, et al. The livertumor segmentation benchmark (lits). Medical Image Analysis, 84:102680, 2023. $https://www.sciencedirect.com/science/article/pii/S1361841522003085.$

[2] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. Advances inneural information processing systems, 33:6840–6851, 2020. $https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.$

[3] Soheil Kolouri, Gustavo K Rohde, and Heiko Hoffmann. Sliced wasserstein distance for learning gaussian mixture models. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3427–3436, 2018. $https://openaccess.thecvf.com/content\_cvpr\_2018/html/Kolouri\_Sliced\_Wasserstein\_Distance\_CVPR\_2018\_paper.html.$

Acknowledgments

We thank **Prof. P. Balamurugan** and the concerned TAs for assisting in completing the project.

Github Link and Demo Video Link

Github Link:

Demo Video Link: