Stable Diffusion for Medical Imaging

Alaqian Zafar
New York University
aa7z118@nyu.edu

Chinmay Tompe
New York University
cst9314@nyu.edu

Akshay Gowda
New York University
ah6147@nyu.edu

Abstract

This paper explores the potential of stable diffusion processes for generating medical imaging data. Stable diffusion processes have shown promise in generating synthetic images that exhibit realistic features and statistical properties similar to real-world medical images. In this study, we propose a method that combines stable diffusion with a convolutional neural network to generate realistic medical images. We evaluate our method on a range of imaging tasks, including image classification and segmentation, and compare it with existing methods. Our strategy is to fine-tune a pre-trained Stable Diffusion model using a specific dataset to generate medical images that are similar in nature. We hope to demonstrate that stable diffusion can effectively generate high-quality medical images that can potentially be used for training machine learning models, data augmentation, and clinical applications. This paper provides insights into the application of stable diffusion in medical imaging and opens up new possibilities for synthetic data generation in the field.

I. Introduction

One of the biggest challenges in medical imaging is the limited availability of data for training deep learning models. Acquiring images through special techniques such as Diffusion MRI can be time-consuming and costly, leading to a shortage of data. For instance, the lack of dermatology data for darker skin tones is a problem that results in reduced model performance for underrepresented groups [1]. The confidentiality of this data, and the costs associated with labeling data stifles scientific development [2]. Stable Diffusion may help to overcome this scarcity by generating synthetic copies of medical images. This can be especially beneficial for startups that seek to use medical images but face limitations in accessing sensitive and private information.

II. Background

Recently, researchers at Stanford used Stable Diffusion to generate chest X-rays and found that training on a combination of real and synthetic images can increase classification performance [3][4]. Pinaya et al. (2023) used latent diffusion models to generate 3D MRI brain images, conditioned on several covariates [5]. Eschweiler & Stegmaier (2023) showed that these models can be used to generate realistic microscopy image data in 2D and 3D based on simulated sketches of cellular structures [6]. These synthetic images were used to train a segmentation model that accurately segmented cells in a real image dataset. These examples demonstrate the potential of Stable Diffusion for augmenting limited medical imaging datasets and improving deep learning model performance.

III. Project Schedule

We have a total of 12 weeks to complete the project, starting from the week of Monday, February 27$^{\text{th}}$. The project will be completed in four phases, with the midpoint of the project and the end of phase II being the week ending on April 6$^{\text{th}}$. Table I shows a gantt chart of the project schedule.

A. Phase I: Literature Survey

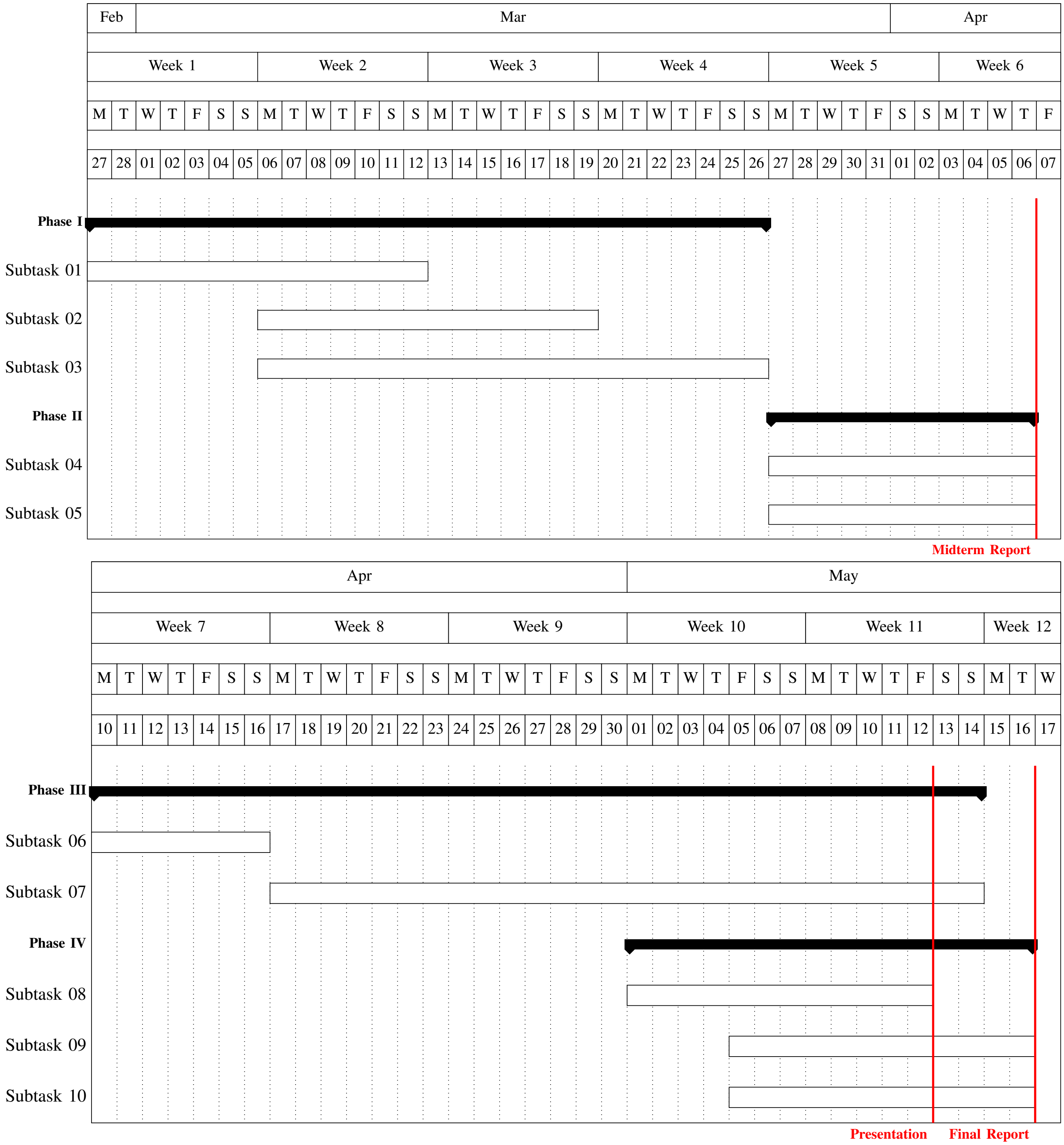
Time Period: Week 1 to Week 4

Phase I involves conducting research and testing the available code for its viability for the project. By the end of this phase, we will have all the information required to start formulating a concrete project scope.

1) Subtask 1: Stable Diffusion

Learn the theory of stable diffusion.
Time Period: Week 1 and Week 2
Assignee: Alaqian, Akshay, Chinmay
Things to consider:
- Learn the theory of stable diffusion
- Understand the mathematics for implementing stable diffusion
- Learn the architecture of the neural network
- Explore the degrees of freedom and modifications that can be made
- Check for available code implementations or tutorials

TABLE I
PROJECT SCHEDULE GANTT CHART



2) Subtask 2: Medical Imaging Papers
Study how stable diffusion has been applied in the medical imaging domain.
Time Period: Week 2 and Week 3
Assignee: Akshay, Chinmay
Things to consider:
- The objective of the study

- The problem they were trying to solve
- Their solution
- Modifications made for the application
- Dataset used
- Evaluation of the results
- Code implementations available

3) Subtask 3: Testing Code

Learn how to fine-tune an image-to-image stable diffusion model and generate synthetic images.

Time Period: Week 2 to Week 4

Assignee: Alaqian, Akshay

Things to consider:

- Resources required such as GPUs, HPC, etc.
- Time required to train the model
- Degrees of freedom
- Quality of images produced

B. Phase II: Project Scope Definition

Time Period: Week 5 and Week 6

Phase II involves finalizing a concrete scope and plan of completing the project. We will use what we learnt through the literature survey in our project plan.

1) Subtask 4: Project Scope

Consider all the research from the prior week to come up with an approach.

Time Period: Week 5 and Week 6

Assignee: Alaqian, Akshay, Chinmay

Things to determine:

- Define the objective of the study and the problem we want to solve
- Propose our solution or approach
- Select the dataset to use
- Determine the resources we need to request, such as GPUs or HPC
- Decide on the evaluation metrics for the model

2) Subtask 5: Midterm Report

Work on the midterm report and put together our research and refined project plan.

Time Period: Week 5 and Week 6

Assignee: Alaqian, Akshay, Chinmay

Requirements:

- Updated abstract
- Overview of the project
- Project accomplishment
- Remaining work and completion schedule
- Updated bibliography

C. Phase III: Design and Implementation of Experiments

Time Period: Week 7 to 11

Phase III will involve designing experiments, testing code, documenting the results.

1) Subtask 6: Preliminary Experiment Design

In week 7, we will design the experiments that we will carry out to implement our solution and make sure we have the resources to carry out the training/fine-tuning in the remaining time.

Time Period: Week 7 and Week 8

Asignee: Alaqian, Chinmay

Things to do:

- Write code to carry out training and evaluation
- Request any additional resources required such as GPUs, HPC, etc.
- Scale the experiments based on the time available
- Implement modifications for the application
- Load the Dataset
- Evaluation of the results

2) Subtask 7: Conducting Experiments
In weeks 8 to 11 we will carry out our experiments. We will also go back to our initial experiment design and adjust it.
Time Period: Week 8 to week 11
Asignee: Akshay, Chinmay
Things to consider:
- Quality of images produced
- Evaluating the images quantitatively
- Revisions need to be made to the preliminary design

D. Phase IV: Final Deliverables

Time Period: Week 10 to Week 12
During Week 10, we will start working on the project presentation and the final report. In the final week, we will finish the final report and document the GitHub repository.
1) Subtask 8: Presentation
Time Period: Week 10 and Week 11
Asignee: Alaqian, Akshay, Chinmay
Requirements:
- Project title and team members
- Overview of your project
- Describe each subtask: what you did and what you learnt/designed
- Demos if appropriate
- Summary

2) Subtask 9: GitHub Documentation
In the final 2 weeks, we will document the project code and the results on the GitHub repository.
Time Period: Week 11 and Week 12
Asignee: Alaqian, Chinmay
Things to do:
- Create a README file
- Document the code and results
- Provide instructions on how to use the code
- References

3) Subtask 10: Final Report
In the final two weeks, we will complete the final report documenting all the work done throughout the project.
Time Period: Week 11 and Week 12
Asignee: Alaqian, Akshay, Chinmay
Requirements:
- Updated abstract
- Overview of the project
- Project accomplishment
- Summary
- References
- Repository

REFERENCES

[1] L. W. Sagers, J. A. Diao, M. Groh, P. Rajpurkar, A. S. Adamson, and A. K. Manrai, "Improving dermatology classifiers across populations using images generated by large diffusion models," 2022.
[2] V. Fernandez, W. H. L. Pinaya, P. Borges, P.-D. Tudosiu, M. S. Graham, T. Vercauteren, and M. J. Cardoso, "Can segmentation models be trained with fully synthetically generated data?," 2022.
[3] P. Chambon, C. Bluethgen, C. P. Langlotz, and A. Chaudhari, "Adapting pretrained vision-language foundational models to medical imaging domains," 2022.
[4] P. Chambon, C. Bluethgen, J.-B. Delbrouck, R. Van der Sluijs, M. Połacin, J. M. Z. Chaves, T. M. Abraham, S. Purohit, C. P. Langlotz, and A. Chaudhari, "Roentgen: Vision-language foundation model for chest x-ray generation," 2022.
[5] W. H. L. Pinaya, P.-D. Tudosiu, J. Dafflon, P. F. da Costa, V. Fernandez, P. Nachev, S. Ourselin, and M. J. Cardoso, "Brain imaging generation with latent diffusion models," 2022.
[6] D. Eschweiler and J. Stegmaier, "Denoising diffusion probabilistic models for generation of realistic fully-annotated microscopy image data sets," 2023.