Indian Institute of Information Technology, Surat

Mid-Semester Report on

MINI PROJECT (CS604)

Submitted by

Aditya Kumar (UI22CS03)

Shruti Paulastye (UI22CS76)

Course Faculty

Ms. Nancy Sukhadia

Department of Computer Science and Engineering

Indian Institute of Information Technology Surat

Gujarat-394190, India

2025-2026

Title:

Zen AI Personalized AI-Driven Image Generation Platform for Professional Photo Transformation

Team Members:

- ***Aditya Kumar***
- ***Shruti Paulastye***

Subject: Mini Project

Course: EC604/CS604

Faculty Mentor: Ms. Nancy Sukhadia

Submission Date: March 7, 2025

Acknowledgment

We extend our deepest gratitude to everyone who contributed to the successful completion of our project, ZEN-AI.

First and foremost, we express our sincere appreciation to Prof. Rajeev Shorey, Director of Indian Institute of Information Technology, Surat, for his visionary initiative in introducing this project-based subject. His foresight and commitment to fostering research and practical application have been instrumental in our development and the realization of this project.

We are profoundly thankful to Ms. Nancy Sukhadia ma'am, our esteemed faculty advisor, for her unwavering encouragement, insightful guidance, and invaluable feedback throughout the project journey. Her mentorship was pivotal in shaping our work, ensuring its successful execution, and providing us with a strong foundation in Generative AI and Machine Learning.

Our sincere thanks also go to our project examiners for their constructive suggestions and continuous academic support. Their expertise and critical analysis significantly enhanced the technical and research aspects of our project.

Furthermore, we extend our appreciation to our industry mentors and external collaborators, who generously shared their real-world insights, technical advice, and valuable perspectives on system integration and scalability. Their contributions were crucial in bridging the gap between academic theory and practical application, ensuring the project's relevance and potential for real-world impact.

This project, ZEN-AI, stands as a testament to collective effort and collaboration. We are truly fortunate to have had such an exceptional team and supportive network throughout this enriching process.

Abstract

This mini-project focuses on solving the challenge of **consistent and personalized image generation**, a growing demand in the era of AI-driven creativity. Often, generating high-quality, identity-preserving images of an individual across various styles and contexts is difficult due to limitations in generic models. To address this, we developed a full-stack **AI SaaS platform** that enables users to generate hyper-personalized and realistic images of themselves anytime, using advanced model fine-tuning techniques.

Our system allows users to upload a dataset of at least **20 diverse images** from different angles, which are then used to train and fine-tune a custom AI model on their likeness. For example, we successfully created a personalized model for **"Elon Musk"**, ensuring that his facial features and identity are consistently maintained across multiple generations. We leveraged techniques like **LoRA (Low-Rank Adaptation)** and **Flux-based model optimization** to fine-tune large language models (LLMs) and image generation models, allowing seamless customization with minimal compute overhead.

Technologically, the application integrates **Next.js**, **Shadcn UI**, and **Supabase** for secure and scalable frontend and backend architecture, while using Replicate API for managing **model training** and inference workflows. Additionally, Stripe is integrated to handle subscription-based billing, ensuring a sustainable SaaS business model.

This report outlines the problem statement, technical architecture, model training methodologies, experimental results, and the potential for future improvements. With its real-world applicability and innovative use of personalization techniques, the project showcases strong potential for academic publication and industry adoption in the personalized AI generation domain.

Table of Contents

List of Figures

List of Tables

Abbreviations and Nomenclature

- **AI:** Artificial Intelligence
- **SaaS:** Software as a Service
- **UI:** User Interface
- **SSR:** Server-Side Rendering
- **API:** Application Programming Interface
- **JWT:** JSON Web Token
- **LoRA:** Low-Rank Adaptation
- **SSR:** Server-Side Rendering

Chapter 1: Introduction

1.1 Background and Motivation

The rapid advancement of Artificial Intelligence (AI) has revolutionized content creation, particularly in image generation. However, the domain of personalized image generation, where AI accurately replicates an individual's likeness, faces significant challenges. Current platforms, predominantly foreign-based, necessitate the upload of sensitive personal data, including facial imagery, raising critical privacy concerns. This data is often stored on servers outside India, leaving users with limited control and transparency over its usage.

Motivated by these privacy and data sovereignty issues, we are developing an Indian-based Generative AI (GenAI) platform dedicated to personalized image generation. Our system prioritizes data security by storing all user data—photos, models, and outputs—exclusively within India, adhering to stringent privacy guidelines. Utilizing advanced AI techniques such as LoRA and Flux-based model fine-tuning, we enable users to train custom models from approximately 20 diverse personal photos, consistently generating high-quality, realistic images that authentically capture their appearance and style.

Furthermore, we are driven to contribute to India's burgeoning AI ecosystem. Aligned with the Indian government's focus on developing indigenous Large Language Models (LLMs) and AI tools, we aspire to collaborate with national bodies to bolster India's AI capabilities, fostering the creation of secure, sovereign, and 'Made-in-India' AI technologies.

1.2 Objectives

The primary objectives of this project are:

- To build a personalized AI-driven image generation platform.
- To explore existing personalized AI image generation techniques and their limitations.
- To research and apply advanced fine-tuning methods like LoRA for improved personalized Image Generation.
- To ensure privacy preserving model training by studying secure data handling practices.
- To contribute research insights towards developing ethical and efficient Generative AI systems
- To support India's AI mission through collaboration with government initiatives, contributing to the development of indigenous GenAI models.

1.3 Problem Statement

With the rise of AI-generated content, there is a growing demand for **personalized image generation**—where people can create high-quality images that reflect their own unique appearance. However, most existing services:

- **Foreign Data Storage:** Reliance on foreign-based services raises critical concerns regarding the storage and usage of sensitive personal data, particularly facial images.
- **Data Sovereignty:** Data stored on servers outside India lacks the protection of Indian privacy laws, leaving users vulnerable.
- **Lack of Transparency and Control:** Users lack adequate control and transparency over the handling of their photos and trained models.

TThis necessitates an Indian-based solution that provides:

- **Secure Data Storage:** A safe platform ensuring all data is stored on Indian servers.
- **Data Privacy:** Complete adherence to Indian privacy laws.
- **Advanced Personalization:** High-quality, personalized image generation using cutting-edge AI techniques.
- **National AI Contribution:** Support for India's AI industry through the development of indigenous AI technologies.

By solving these problems, our project aims to create a trusted, scalable, and privacy-first AI image generation platform made specifically for **Indian users** and beyond.

1.4 Need of the Study

In today's digital world, people love to create unique and professional images of themselves using AI. But most of the platforms available for personalized AI image generation are developed by **foreign companies**. These platforms require users to upload sensitive personal data, especially **facial photos**, which raises serious concerns like:

- **Privacy Risks** – Users don’t have control over where their data is stored or how it’s used.
- **Data Security** – Information is stored on global servers, outside India’s legal protection.
- **Lack of Local Solutions** – There are very few Indian platforms that offer high-quality, personalized AI image generation with a focus on privacy.

At the same time, the **Government of India** is focusing on developing **India-first AI technologies** to make the country self-reliant in artificial intelligence. This creates a perfect opportunity to build a **safe, Indian-based GenAI platform** that:

- **Ensures Data Sovereignty:** Protects user privacy by storing data within India.
- **Delivers High-Quality Personalization:** Provides advanced AI image generation capabilities.
- **Contributes to National AI Growth:** Supports India's AI innovation and development.

Thus, there is a **strong need** for this project to fill the gap of **secure, local, and advanced personalized AI image generation** services.

1.5 Scope of the Study

This study focuses on developing an India-based AI image generation platform that ensures user data privacy by storing all personal data within the country. The goal is to create a secure, scalable, and personalized image generation service using techniques like LoRA and Flux.

We aim to reduce reliance on foreign platforms and work towards training our own custom models in India. For this, high GPU power and advanced computing resources are required. To make the system more efficient and affordable, we plan to optimize models to reduce computation costs and enable local deployment.

Additionally, the project aims to support national AI initiatives by contributing towards the development of indigenous generative AI technologies while maintaining ethical data usage and privacy.

Chapter 2: Literature Review

Literature Review: Generative AI Development

2.1. Introduction

Generative AI (Gen AI) represents a transformative class of artificial intelligence models that create new content—ranging from text and images to music and code—by learning patterns from existing data. Over the past decade, deep learning, particularly through transformer architectures, has revolutionized how these models work, enabling them to generate content that is increasingly complex and realistic. This review examines the evolution, current state, applications, challenges, and future directions of generative AI, highlighting its relevance to our project on personalized image generation.

2.2. Evolution of Generative AI

The journey of generative AI began with simple probabilistic models such as Markov chains and Gaussian Mixture Models (GMMs), which, despite their limitations, set the groundwork for understanding data patterns. The field took a significant leap forward with the advent of deep learning. The introduction of Generative Adversarial Networks (GANs) by **Goodfellow** et al. in 2014 and Variational Autoencoders (VAEs) by **Kingma and Welling** in 2013 marked a turning point. These models could generate more complex data and opened the door for future innovations.

The real breakthrough came with the transformer revolution. The transformer architecture, introduced by Vaswani et al. in 2017, enabled the development of large-scale language models like **GPT and BERT**, which not only excelled in language understanding and generation but also paved the way for multimodal applications, where models generate images and other content types from text prompts.

2.3. Key Developments in Generative AI

Recent years have witnessed rapid advancements driven by improvements in model architectures, training techniques, and the availability of powerful computational resources. Notably, large language models (LLMs) such as **OpenAI's GPT-3** and **GPT-4** have set new performance benchmarks in natural language generation, while models like Google's PaLM and Meta's LLaMA have contributed significantly through open-source initiatives.

In parallel, multimodal models like DALL·E, Stable Diffusion, and Imagen have pushed the boundaries of text-to-image generation, achieving remarkable levels of detail and fidelity. Beyond creative fields, generative AI has also made strides in code generation, with tools like GitHub Copilot and Amazon CodeWhisperer assisting developers in writing and debugging code.

2.4. Applications of Generative AI

Generative AI is being applied across diverse domains. In healthcare, models like AlphaFold are revolutionizing drug discovery and protein structure prediction, while synthetic data generation aids research and training. In the business realm, productivity tools powered by AI, such as Microsoft 365 Copilot, are automating repetitive tasks, and in education, personalized AI tutors and content generators are transforming learning experiences. Moreover, in entertainment, AI-generated art, music, and video content are increasingly integral to creative production.

2.5. Challenges and Ethical Considerations

Despite its tremendous potential, generative AI faces several challenges. One major concern is bias—models often inherit and amplify biases present in their training data, leading to unfair outputs. The realistic nature of AI-generated content also raises concerns about misinformation and deepfakes. Additionally, the use of copyrighted data for training these models has sparked debates over intellectual property rights. Furthermore, the high computational resources required for training these models contribute significantly to environmental impacts, an increasingly important consideration in today's eco-conscious world.

Chapter 3: Problem Statement

3.1 Problem Background

Current photo generation systems, while advanced, often produce generic outputs that fail to capture the unique traits and visual styles of individual users. Models like DALL·E, Stable Diffusion, and MidJourney, though impressive, lack personalization, leading to outputs misaligned with user-specific characteristics. Key challenges include:

- **Generic Outputs:** Inability to replicate personal style or features.
- **Limited Customization:** Lack of user-driven personalization.
- **High Computational Costs:** Resource-intensive fine-tuning.
- **Data Privacy Concerns:** Overseas data storage in existing solutions.

3.2 Research Gap

Despite advancements in personalization techniques like DreamBooth and LoRA, there's a gap in accessible, scalable solutions that:

- Enable non-expert fine-tuning.
- Operate with reduced computational costs.
- Provide a user-friendly web interface.
- Ensure local data privacy in India.

3.3 Objectives

Our objectives are:

- **Personalization:** Fine-tune models with user-uploaded images.
- **Accessibility:** Develop an intuitive user interface.
- **Data Privacy:** Store data on Indian servers.
- **Cost Optimization:** Reduce fine-tuning computational overhead.
- **Scalability:** Build a robust, scalable SaaS architecture.
- **Indigenous AI:** Contribute to India's AI ecosystem.

3.4 Proposed Solution

Our platform leverages LoRA to fine-tune a base Flux Dev model using user images. Key components include:

- **Custom Token System:** Unique tokens for personalized style.
- **Efficient Fine-Tuning:** Resource-efficient model adaptation.
- **Web-Based Interface:** User-friendly image uploads and outputs.
- **Scalable Architecture:** SaaS-based, multi-user support.
- **Local Data Storage:** Data privacy within India."

3.5 Significance of the Study

"This study addresses personalization and data privacy shortcomings in AI image generation. The platform will:

- **Enhance User Experience:** Deliver personalized, professional images.
- **Broaden Applications:** Expand use in photography, marketing, and content creation.
- **Support Indigenous AI:** Develop privacy-conscious, local AI solutions.
- **Reduce Costs:** Optimize fine-tuning for broader accessibility.

This research establishes a foundation for a scalable, private, and cost-effective GenAI platform tailored for professional use in India."

Key Refinements:

- **Conciseness:** Removed redundant phrases and streamlined sentences.
- **Bullet Points:** Used bullet points to clearly highlight key points.
- **Emphasis on Key Features:** Highlighted the unique aspects of your solution (custom token system, local data storage).
- **Stronger Language:** Used more impactful verbs and phrases to emphasize the significance of your work.
- **Clarity and Flow:** Improved the overall flow and readability of the chapter.

Chapter 4: Project Plan & Timeline

4.1 Deliverables and Outcomes

- **Research Output:**
  - A publication-ready paper detailing our approach and findings.
  - A comprehensive technical report covering system architecture, methodologies, experiments, and results.
- **Open-Source Contribution:**
  - Release of the complete source code for academic and personal use, fostering community development and collaboration.
- **Demo and Training:**
  - Live demo sessions and internal workshops to showcase the platform's features and functionalities.
- **Intellectual Property:**
  - Evaluation of novel techniques for potential patents or copyrights.

4.2 Timeline and Work Items

Work Item	**January**	**February**	**March**	**April**	**Outcome**
Major Work Item 1: Project Setup & Environment Configuration	✅				Initial project repository setup and CI/CD pipeline established.

Sub Work Item 1.1: UI/UX Development (Next.js, Shadcn UI)	✅	✅			Development of a responsive and user-friendly interface.
Sub Work Item 1.2: Supabase Integration for Auth & Storage			✅		Secure login and robust data management system in place.
Major Work Item 2: AI Model Training & Integration			✅	✅	Fine-tuned model on user-specific data using DreamBooth/LoRA.
Sub Work Item 2.1: Implement Replicate API Integration			✅	✅	Custom training pipeline integrated with the Replicate API.
Major Work Item 3: Payment and Subscription Module (Stripe)			✅	✅	Billing, credits, and subscription management are fully functional.
Major Work Item 4: Testing, Deployment & Feedback Integration			✅	✅	Deployed application with user feedback loop established.

Chapter 5: Methodology

5.1 System Architecture

Our system is designed as a multi-layered architecture to ensure a smooth, secure, and scalable user experience for personalized image generation. The key layers include:

- **Frontend:**
  Built on Next.js to enable fast, SEO-friendly pages with server-side rendering (SSR). The user interface is styled using Tailwind CSS and Shadcn UI, providing a modern and responsive design.
- **Backend:**
  A Node.js-based API manages core functionalities such as authentication, image uploads, and job scheduling. This layer ensures that requests are handled efficiently and securely.
- **Database & Storage:**
  Supabase is employed for user data management, secure image storage, and session management. By using Supabase, all sensitive data is stored locally within India to address data privacy concerns.
- **AI Model Training:**
  The platform integrates the Replicate API for fine-tuning a pre-trained diffusion model (e.g., Flux Model) using advanced techniques of LoRA. This allows for custom model training based on user-uploaded images and a unique personal token.
- **Payment Gateway:**
  Stripe is integrated to handle secure billing, subscription management, and credit handling, ensuring smooth and safe financial transactions.
- **Asynchronous Processing:**
  Celery/Redis-based job queues are utilized to manage long-running tasks, such as model training and fine-tuning, ensuring that the system remains responsive.**5.2 Technology Stack**

The following technologies and frameworks form the backbone of our solution:

- **Frontend:**
  - Next.js, React
  - Tailwind CSS, Shadcn UI
- **Backend:**
  - FastAPI, Node.js, Express (if applicable)

- **Database & Storage:**
  - Supabase (PostgreSQL)
- **AI & Machine Learning:**
  - PyTorch, HuggingFace Diffusers
  - Replicate API (for model training and inference)
  - Techniques: LoRA
- **Payment Processing:**
  - Stripe
- **State Management:**
  - Zustand
- **Notifications:**
  - Resend for email notifications

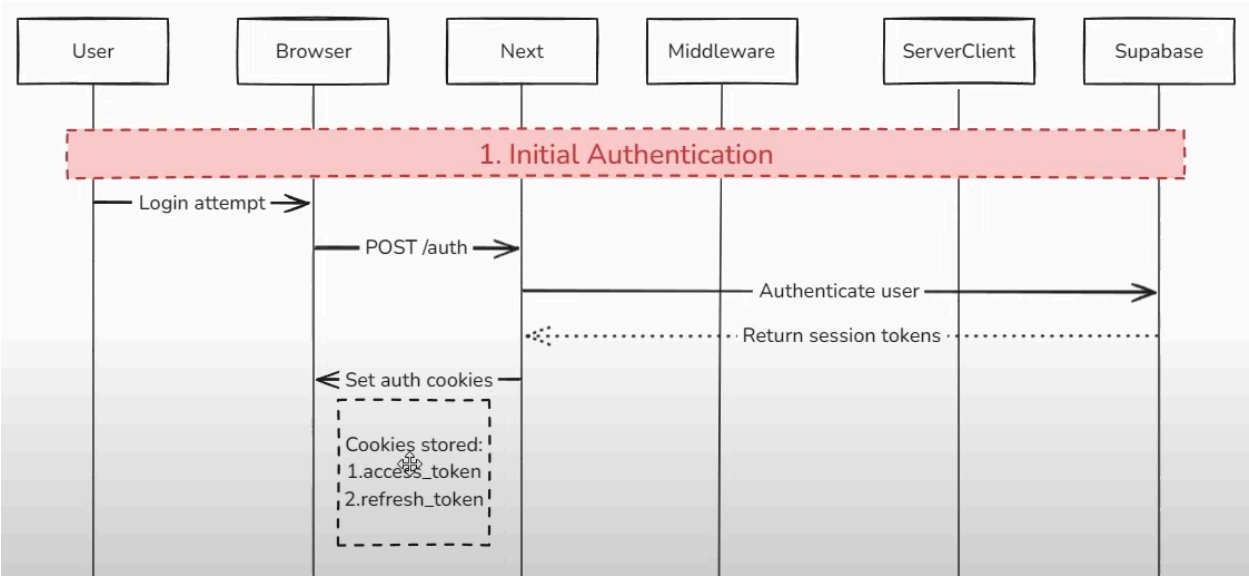

Figure 1: System architecture

5.2 Model Training and Personalization

The core of our platform is the AI model training and personalization process, which includes:

- **Image Preprocessing:**
  User-uploaded images are validated, resized, and normalized to meet the input requirements of the base model. This step ensures consistent quality and prepares the data for fine-tuning.
- **Fine-Tuning Process:**
  The system employs advanced techniques such as DreamBooth and LoRA to adapt a pre-trained diffusion model (e.g., Stable Diffusion) to a user's specific features.
    - A unique token is assigned to each user, representing their personal visual style.
    - Fine-tuning is performed using a curated dataset (typically at least 20 images from various angles) to capture the user's distinctive traits.
    - Training iterations are monitored closely to prevent overfitting and to maintain high image quality while optimizing GPU usage.
- **Inference Pipeline:**
  Once the model is fine-tuned, users can generate personalized images by submitting text prompts that include their custom token.
    - The inference module processes these prompts to generate images that reflect the user's unique characteristics.
    - Generated images are post-processed and stored in the user's gallery for easy access and further use.

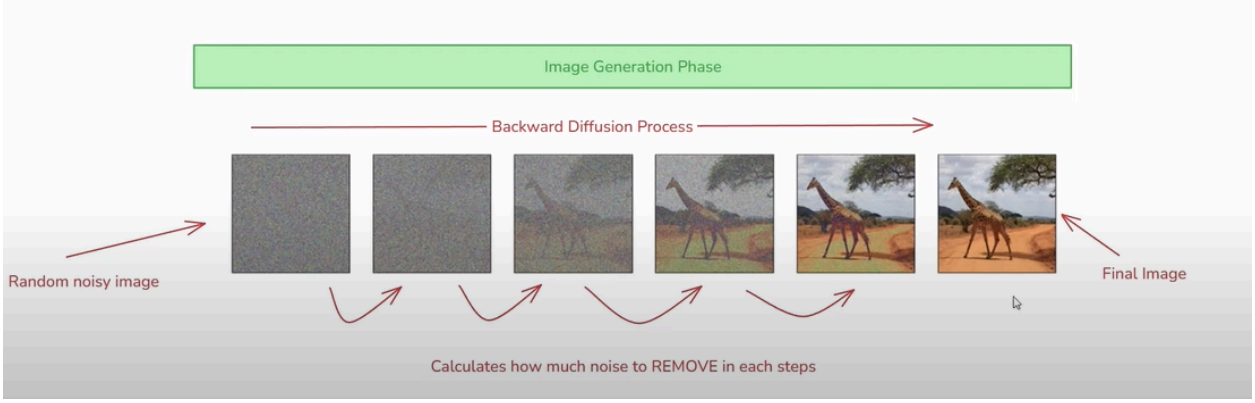

Figure 2: Image generation Model Working

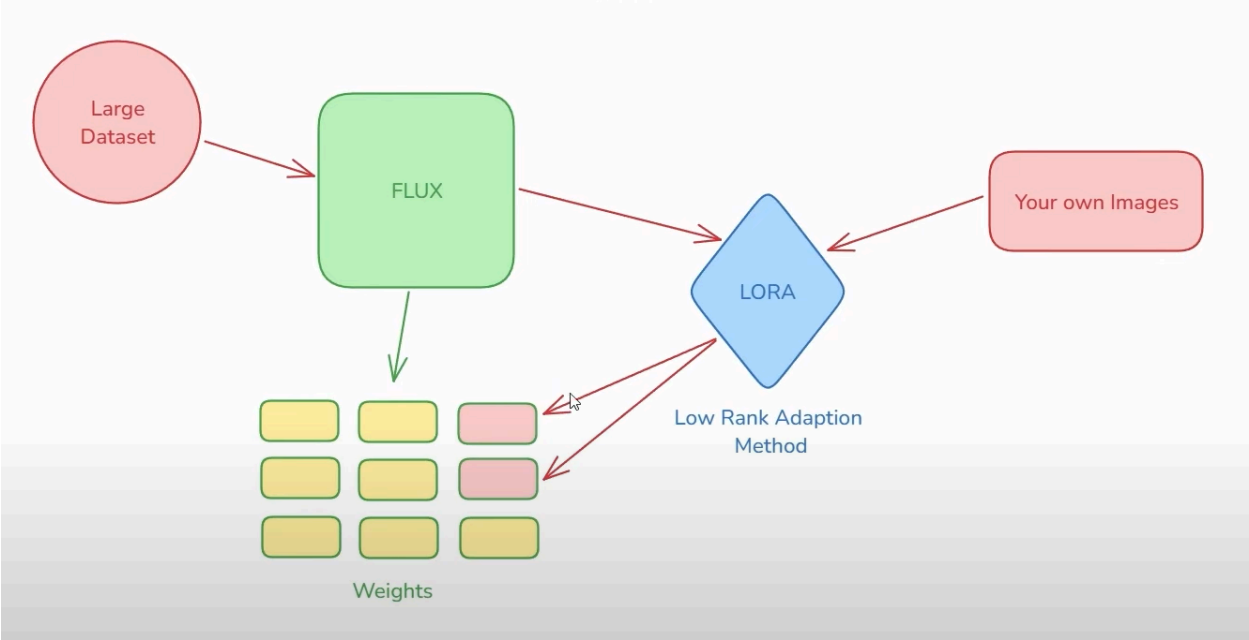

Figure 3: Lora Image Training

Chapter 6: Implementation

6.1 Frontend Development

- **User Authentication & Dashboard:**
  - Integration of Supabase for secure login and session management.
  - Development of a responsive dashboard for image uploads and prompt submissions.
- **Interactive Components:**
  - Image gallery component for viewing uploaded and generated images.
  - Real-time status updates on training jobs using webhooks

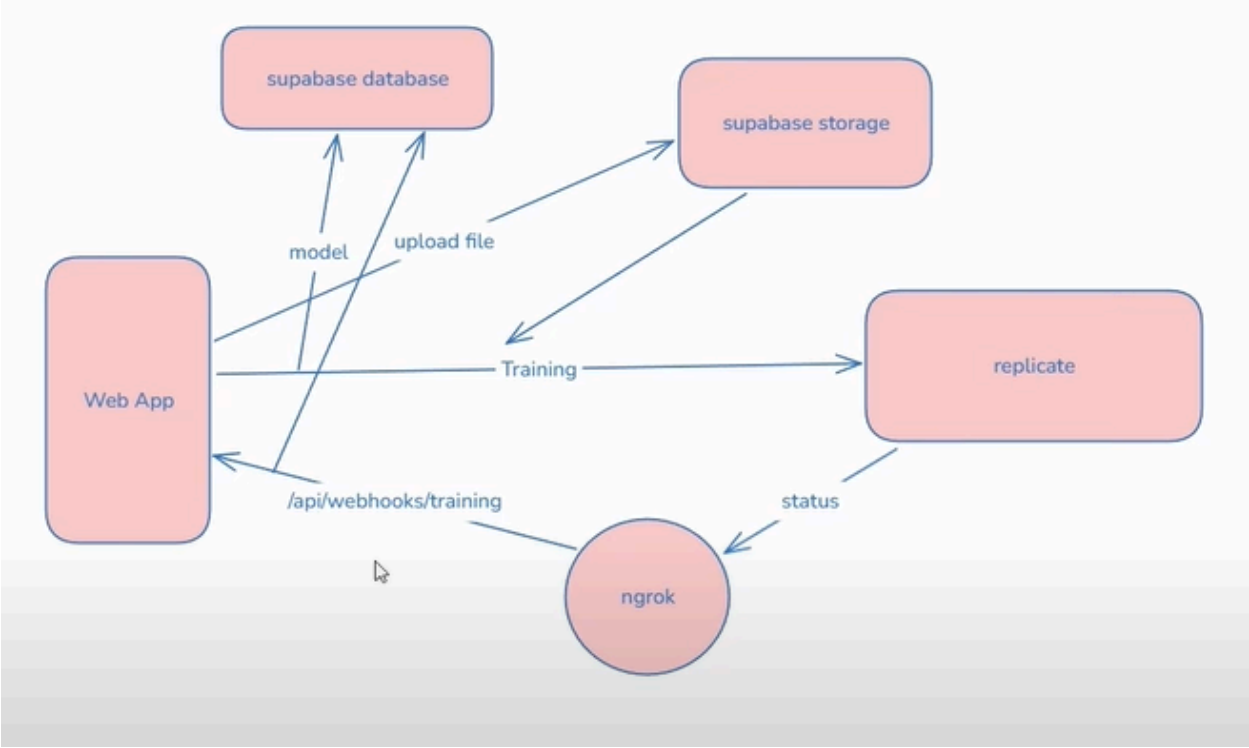

Figure 4: System Architecture

6.2 Backend & API Integration

- **Endpoints:**
  - /upload: For handling image uploads and storage.
  - /train: To initiate AI model fine-tuning.
  - /generate: For processing text prompts and returning AI-generated images.
- **Job Management:**
  - Implement asynchronous processing using Celery/Redis to handle long-running training tasks.
  - Webhook integration to update the system on training status.

6.3 AI Model Fine-Tuning Pipeline

- **Training Module:**
  - Fine-tune a pre-trained diffusion model using user images and a unique token.
  - Use Replicate API to trigger and monitor training jobs.
- **Inference Module:**
  - Dynamically load the fine-tuned model or token embedding during image generation.
  - Optimize latency through caching and efficient resource management.

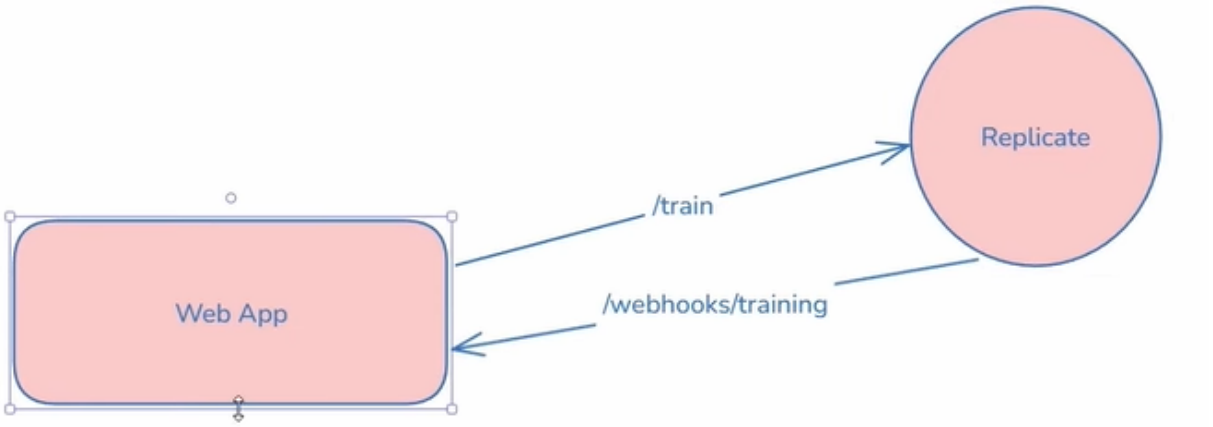

Figure 5: API Architecture

6.4 Payment Integration and Billing

- **Stripe Integration:**

    - Implement subscription models, credits, and billing workflows.
    - Secure endpoints to restrict premium features to paying users.
- **Usage Monitoring:**
    - Track image generations and training sessions to manage user credits effectively.

Chapter 7: Results and Analysis

7.1 System Performance

- **Model Accuracy:**
  - Evaluation metrics comparing generated images against user expectations.
- **User Engagement:**
  - Analytics on user interaction, training time, and image generation frequency.
- **Scalability:**
  - Testing the robustness of the application under concurrent training and inference requests.

7.2 Challenges and Mitigation

- **Training Latency:**
  - Mitigated by job queues and asynchronous processing.
- **Data Privacy:**
  - Secure storage protocols and encrypted transmission for user images.
- **Model Overfitting:**
  - Utilized data augmentation and early stopping techniques.

Chapter 8: Conclusion and Future Work

8.1 Conclusion

The personalized AI image generation platform successfully integrates state-of-the-art web technologies and advanced AI methodologies to offer a unique, user-centric service. The project demonstrates how personalized fine-tuning can transform generic models into tailored solutions, with real-world applicability in professional photo generation.

8.2 Future Work

- **Continuous Learning:** Implement dynamic model updates based on ongoing user feedback.
- **Multi-Modal Inputs:** Expand the platform to support additional inputs such as sketches or voice descriptions.
- **Enhanced Personalization:** Explore further fine-tuning techniques and customization options.
- **Scalability Improvements:** Optimize resource allocation and incorporate advanced container orchestration for global deployment.
- **Extended Monetization:** Experiment with tiered subscription models and incentive programs.

Chapter 9: References

1. Ruiz, N., et al. (2023). DreamBooth: Fine-Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
2. Hu, E. J., et al. (2021). LoRA: Low-Rank Adaptation of Large Language Models. arXiv preprint arXiv:2106.09685.
3. Rombach, R., et al. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
4. Brown, T., et al. (2020). Language Models are Few-Shot Learners. arXiv preprint arXiv:2005.14165.
5. Industry White Papers on SaaS Application Architecture. Microsoft Azure, AWS, and Google Cloud.
6. https://dreambooth.github.io/
7. https://huggingface.co/black-forest-labs/FLUX.1-dev
8. https://arxiv.org/abs/2412.18653
9. https://medium.com/@drmarcosv/how-does-flux-work-the-new-image-generation-ai-that-rivals-midjourney-7f81f6f354da

Chapter 10: Appendices

- **Appendix A:** Detailed code snippets.

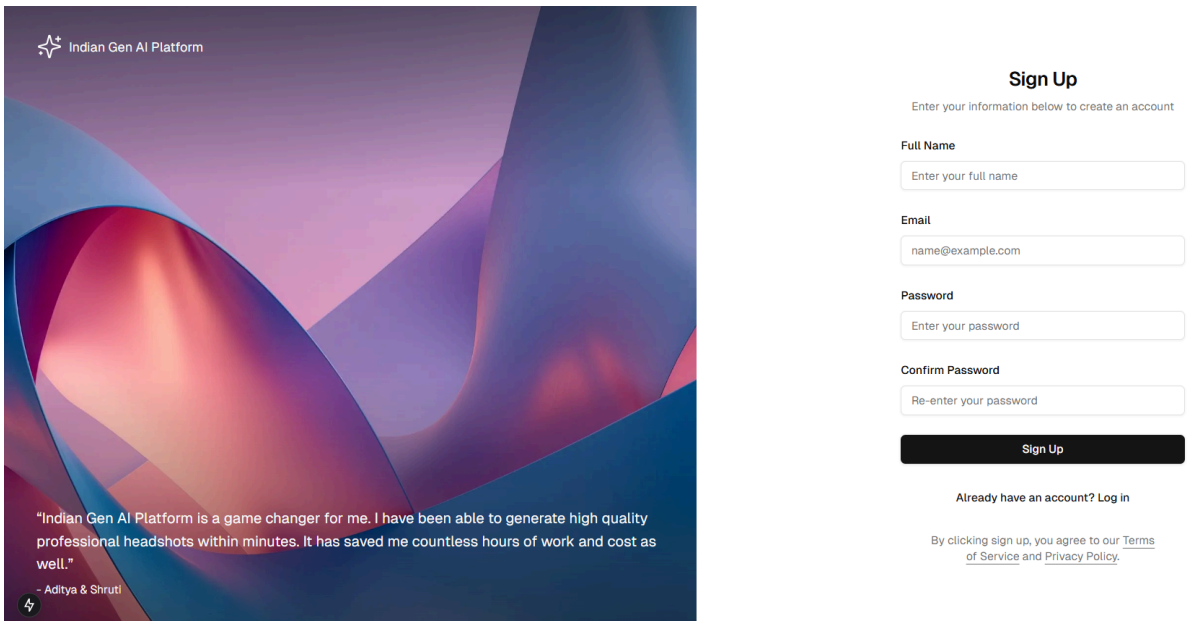

Figure 6: Signup Page

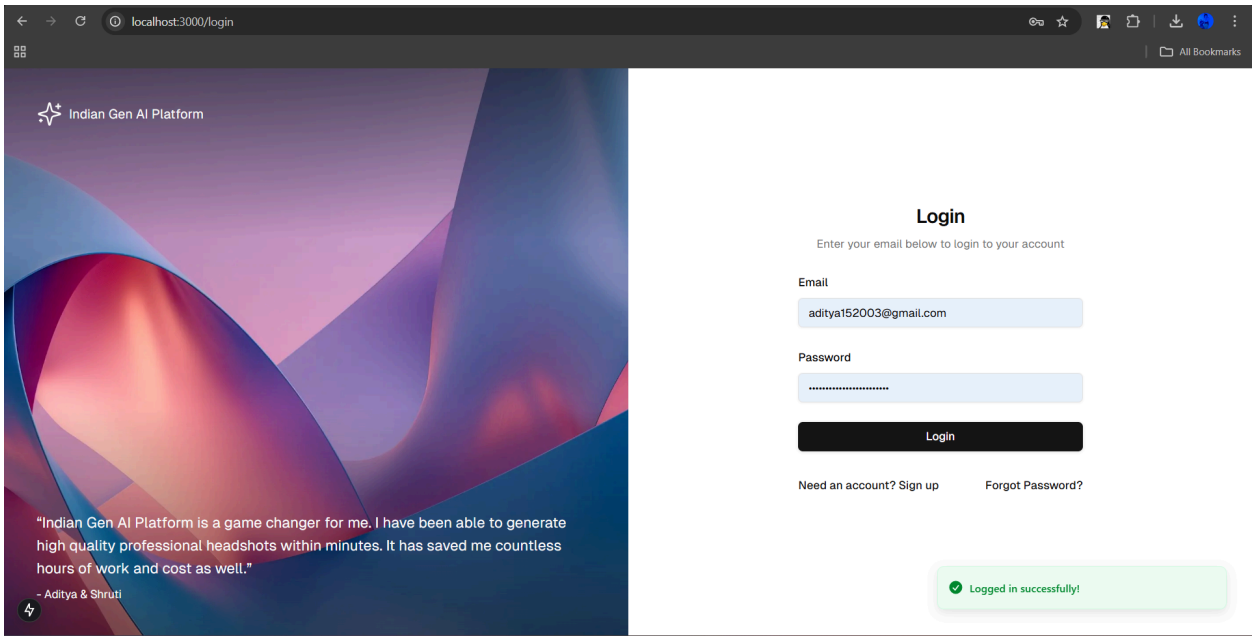

Figure 7: Login Page

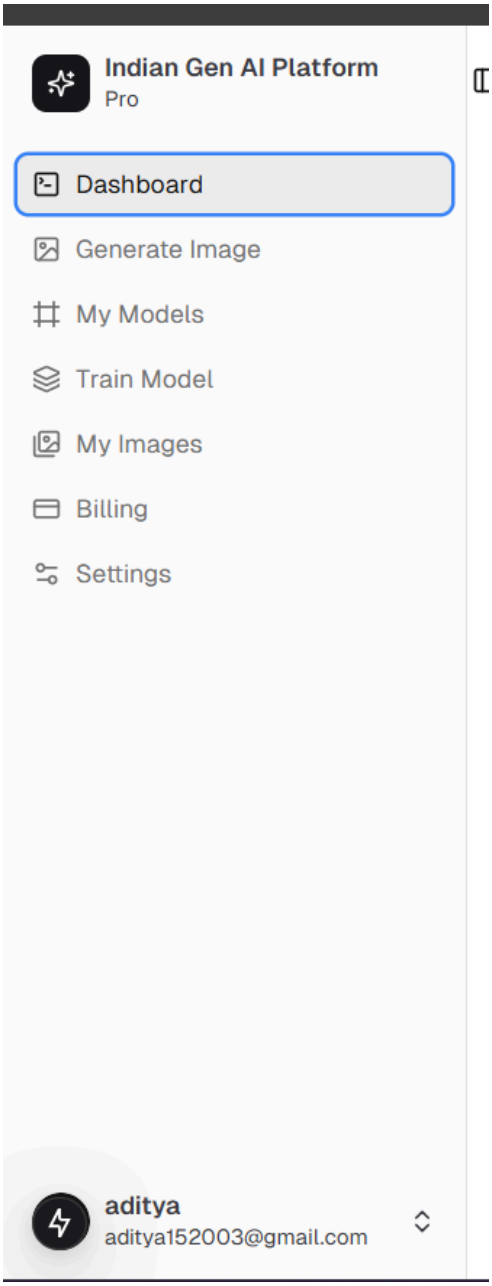

Figure 8: SideBar

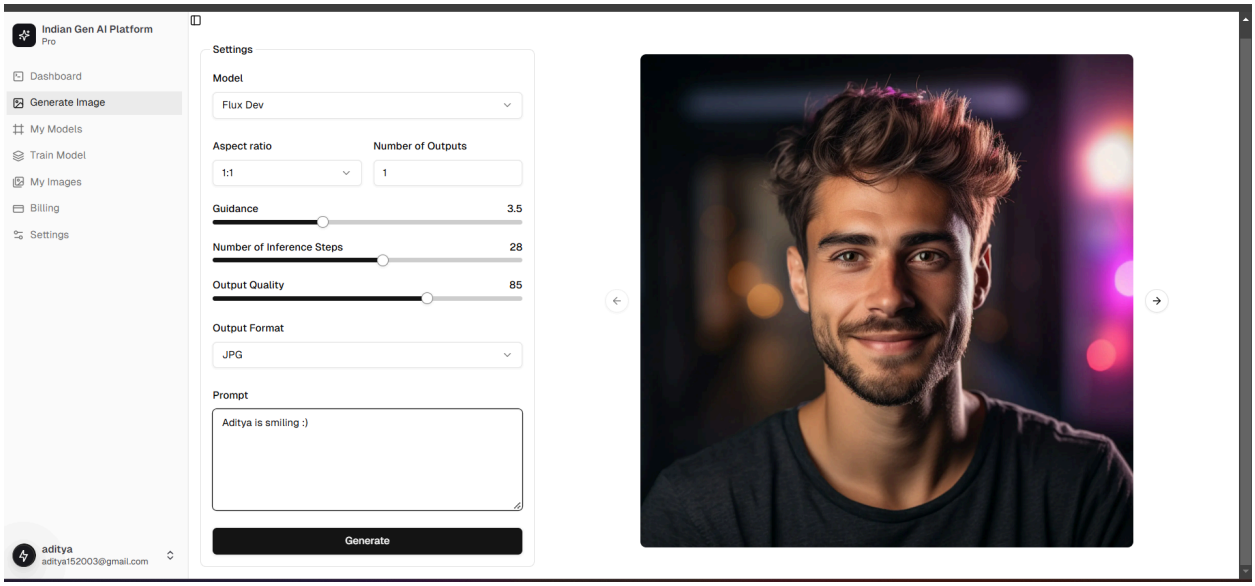

Figure 9 : Frontend Part of Generate Image

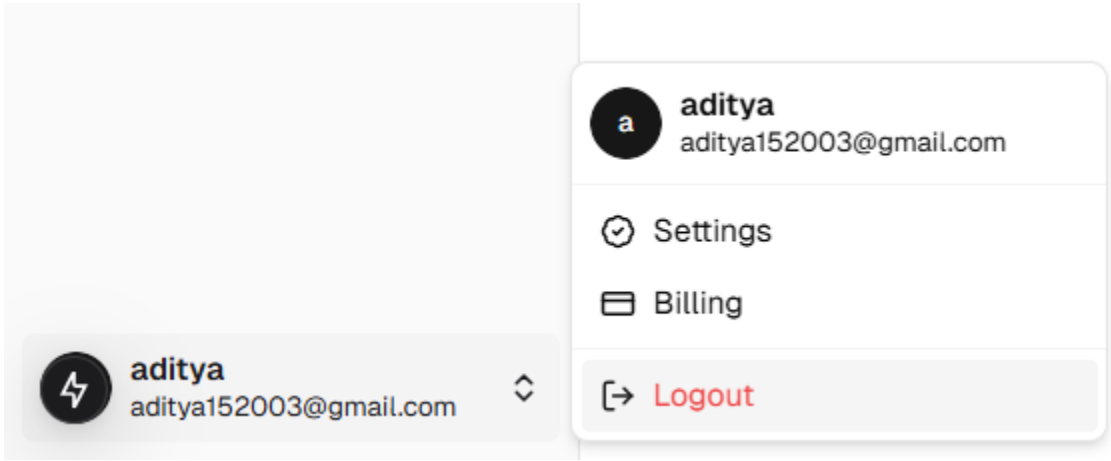

Figure 10: Logout and Profile Button

- **Appendix B:** Backend

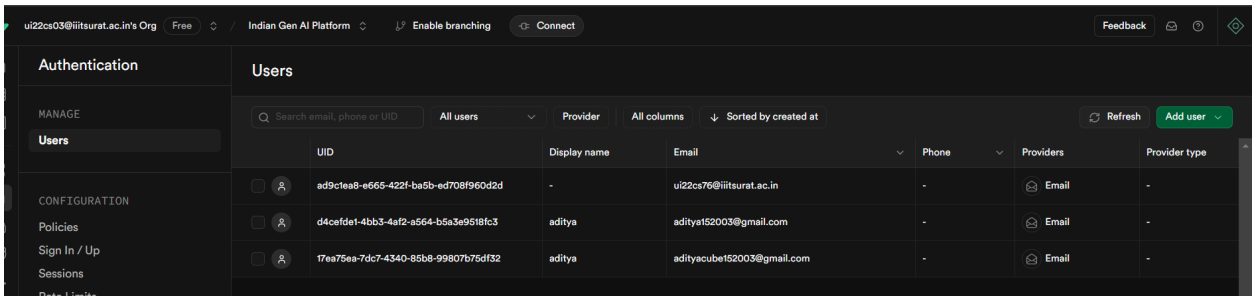

Figure 11: Backend SupaBase Authentication

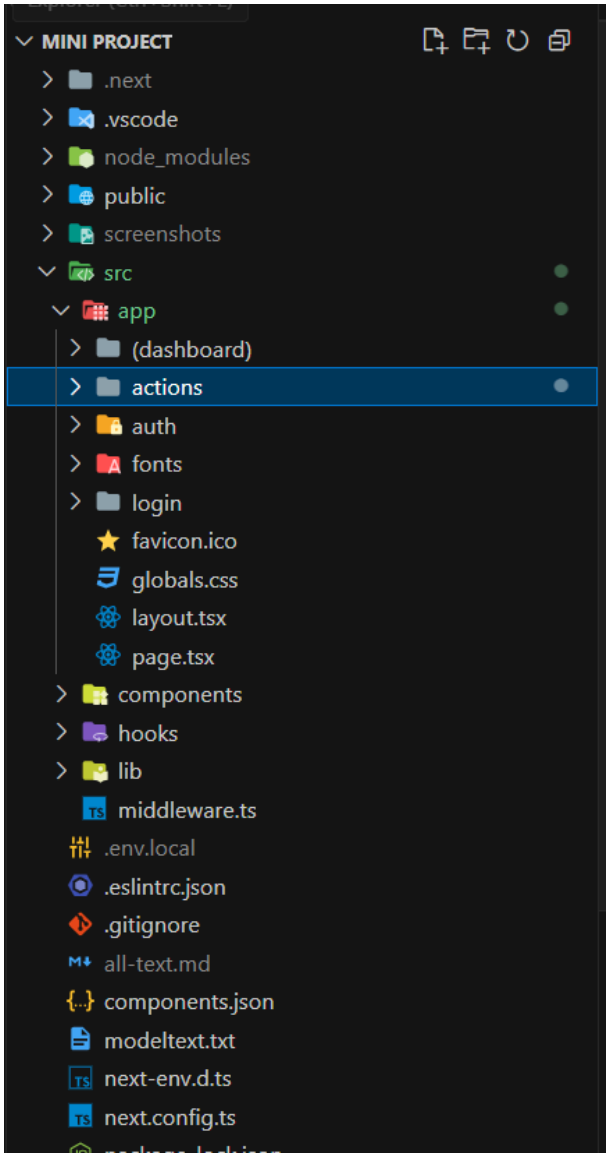

Figure 12: Codes