

# **IMAGINIFY**

**A PROJECT REPORT  
for  
Project (KCA451)  
Session (2024-25)**

**Submitted by**

**DARREN JAMES  
2300290140050  
DIVYANSH SHARMA  
2300290140057  
ANSHOO YADAV  
2300290140028  
BHARGAVI  
2300290140048**

**Submitted in partial fulfilment of the  
Requirements for the Degree of**

## **MASTER OF COMPUTER APPLICATION**

**Under the Supervision of  
Ms. Komal Salgotra  
Assistant Professor**



**Submitted to  
DEPARTMENT OF COMPUTER APPLICATIONS  
KIET Group of Institutions, Ghaziabad  
Uttar Pradesh-201206  
(APRIL 2025)**

## **DECLARATION**

I hereby declare that the work presented in this report entitled “Imaginify” was carried out by me. I have not submitted the matter embodied in this report for the award of any other degree or diploma of any other University or Institute.

I have given due credit to the original authors/sources for all the words, ideas, diagrams, graphics, computer programs, experiments, results, that are not my original contribution.

I have used quotation marks to identify verbatim sentences and given credit to the original authors/sources.

I affirm that no portion of my work is plagiarized, and the experiments and results reported in the report are not manipulated. In the event of a complaint of plagiarism and the manipulation of the experiments and results, I shall be fully responsible and answerable.

**Name:**

**Roll No.:**

**Darren James**

**2300290140050**

**Divyansh Sharma**

**2300290140057**

**Anshoo Yadav**

**2300290140028**

**Bhargavi**

**2300290140048**

**Branch: MCA**

**Candidate's Signature:**

## **CERTIFICATE**

Certified that **Darren James 2300290140050, Divyansh Sharma 2300290140057, Bhargavi 2300290140048, Anshoo Yadav 2300290140028** have carried out the project work having “**Imaginify**” (**Major Project-KCA451**) for **Master of Computer Application** from Dr. A.P.J. Abdul Kalam Technical University (AKTU) (formerly UPTU), Lucknow under my supervision. The project report embodies original work, and studies are carried out by the student himself/herself and the contents of the project report do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

**Date:**

**Darren James 2300290140050**

**Divyansh Sharma 2300290140057**

**Bhargavi 2300290140048**

**Anshoo Yadav 2300290140028**

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date:

**Ms. Komal Salgotra**  
**Assistant Professor**  
**Department of Computer Applications**  
**KIET Group of Institutions, Ghaziabad**

**Dr. Akash Rajak**  
**Dean-MCA**  
**Department of Computer Applications**  
**KIET Group of Institutions, Ghaziabad**

# Imaginify

## ABSTRACT

This report outlines the design and development of an AI-based Software-as-a-Service (SaaS) platform focused on simplifying and enhancing image editing through the integration of intelligent tools. The platform provides users with a comprehensive suite of advanced image manipulation features powered by artificial intelligence, including image restoration, generative fill for aspect ratio adjustment, object removal, object recolouring, and background removal. These tools address common and time-consuming challenges in digital image editing by offering efficient, one-click solutions that were traditionally manual and skill-intensive.

The core objective of this platform is to democratize access to professional-grade image editing by making it accessible through a web-based interface, eliminating the need for complex software or high-end hardware. In many traditional image editing workflows, users are required to have in-depth knowledge of software like Adobe Photoshop or GIMP, which often involves a steep learning curve. Moreover, professional image editing requires considerable computing power, which can be a barrier for users on low-end devices. By offloading processing tasks to the cloud and streamlining interactions through a user-friendly interface, the platform enables users of all skill levels to achieve high-quality results without technical know-how or expensive infrastructure.

Leveraging the capabilities of Cloudinary's AI-based APIs, the platform handles the computationally intensive tasks in the cloud. Cloudinary is known for its media optimization and transformation capabilities, and by integrating its AI-driven APIs, the platform ensures high-speed processing, accurate outputs, and scalability. Each AI-powered feature is seamlessly embedded into the platform, allowing for a unified user experience. For example, users can remove unwanted objects from images with a single click, and the AI fills in the missing areas with context-aware data that preserves the natural look and feel of the image. Similarly, generative fill allows for aspect ratio adjustments without distortion, making it especially useful for e-commerce and marketing professionals who need consistent image sizes across various platforms.

The image restoration tool uses advanced neural networks trained on large datasets to remove noise, fix blurriness, and restore old or damaged photographs. This feature is particularly beneficial for archival purposes and personal use cases like restoring family photos. Background removal is another highly valuable feature that utilizes segmentation models to isolate the subject from the background, allowing users to replace or remove the background entirely. Object recolouring leverages semantic understanding to enable users to change the color of specific elements in an image while maintaining shading, lighting, and texture.

The technology stack is built on Next.js, a modern React framework that enables server-side rendering (SSR) and static site generation (SSG), both of which contribute to faster page load times and improved SEO. Next.js is well-suited for SaaS platforms due to its performance optimizations and support for incremental static regeneration, which ensures that users always access the most up-to-date content. The use of TypeScript adds an additional layer of reliability and maintainability to the codebase. By enforcing type constraints during development, TypeScript reduces the likelihood of bugs and runtime errors, resulting in a more stable application.

User authentication and management are handled through Clerk, a modern identity management solution designed for frontend frameworks like React. Clerk abstracts away the complexities of authentication, providing prebuilt components for login, registration, password reset, and user profiles. This not only accelerates development time but also ensures that user data is managed in a secure and compliant manner. Clerk supports multi-factor authentication (MFA), session management, and social login integrations, all of which enhance the platform's security and user experience.

To facilitate monetization, the platform integrates with Stripe, a globally trusted payment gateway. Stripe supports various pricing strategies, including monthly subscriptions, one-time payments, and usage-based billing. This flexibility allows the platform to cater to different user segments, from casual users who need occasional access to professionals and businesses requiring frequent or large-scale usage. Stripe's APIs handle payment processing, invoicing, and transaction history while ensuring compliance with global financial regulations like PCI-DSS and GDPR. The integration with Stripe also enables the platform to support promotional offers, discount codes, and tiered access based on user roles or subscription plans.

From a design perspective, the platform prioritizes a clean, intuitive user interface (UI) that minimizes friction and maximizes usability. Users can upload images directly from their devices or via URLs, apply edits using AI tools, preview changes in real-time, and download the result—all without leaving the platform. The UI is designed with accessibility in mind, ensuring compatibility with assistive technologies and compliance with standards such as WCAG. Responsive design principles are followed to ensure optimal performance across devices, including desktops, tablets, and smartphones.

The development approach embraces agile methodologies, allowing for iterative improvements and rapid feature deployment. Continuous integration and continuous deployment (CI/CD) pipelines are set up to automate testing and deployment, reducing the risk of bugs and downtime. Error monitoring and analytics tools are integrated to track user behaviour, identify bottlenecks, and optimize the platform continuously. The backend architecture is modular and scalable, enabling easy integration of new features and third-party services in the future.

Security is a key consideration in every aspect of the platform. All data transmissions are secured using HTTPS and encrypted using SSL/TLS protocols. User data, including

authentication credentials and payment information, is stored securely using industry best practices. Role-based access control (RBAC) is implemented to restrict access to sensitive administrative functions. Regular security audits and penetration testing are conducted to identify and mitigate vulnerabilities proactively.

The platform's architecture is designed for scalability and resilience. It utilizes serverless functions for processing user requests, ensuring that resources are allocated dynamically based on demand. This approach minimizes latency and operational costs while maximizing performance. Load balancing and content delivery networks (CDNs) are employed to distribute traffic efficiently and ensure high availability across different geographic regions. Looking ahead, the platform is poised for continuous growth and innovation. Planned enhancements include batch processing capabilities, allowing users to apply edits to multiple images simultaneously. This feature is particularly useful for e-commerce businesses managing large catalogs. Mobile compatibility is another priority, with native applications in development for iOS and Android platforms. These apps will offer offline capabilities and deeper integration with device features such as cameras and photo libraries.

Additional AI-powered features are also on the roadmap. These include facial recognition for automatic tagging and sorting, AI-generated captions for accessibility and SEO, and smart enhancement tools that adjust lighting, contrast, and color balance based on the content of the image. Integrations with popular content management systems (CMS) and e-commerce platforms like Shopify and WordPress are also planned, enabling seamless workflows for content creators and online retailers.

In summary, this AI-based SaaS platform represents a convergence of modern web technologies and artificial intelligence to deliver a powerful, scalable, and user-friendly image editing solution. It addresses a real market need by offering an accessible alternative to traditional image editing tools, and it lays the groundwork for future enhancements that will further empower users and expand the platform's capabilities. By focusing on usability, performance, and innovation, the platform positions itself as a competitive and forward-looking solution in the rapidly evolving digital media landscape.

## **ACKNOWLEDGEMENTS**

Success in life is never attained single-handedly. My deepest gratitude goes to my project supervisor, **Ms. Komal Salgotra** for her guidance, help, and encouragement throughout my project work. Their enlightening ideas, comments, and suggestions.

Words are not enough to express my gratitude to Dr. Akash Rajak, Professor and Head, Department of Computer Applications, for his insightful comments and administrative help on various occasions.

Fortunately, I have many understanding friends, who have helped me a lot on many critical conditions.

Finally, my sincere thanks go to my family members and all those who have directly and indirectly provided me with moral support and other kind of help. Without their support, completion of this work would not have been possible in time. They keep my life filled with enjoyment and happiness.

**Darren James**

**Divyansh Sharma**

**Bhargavi**

**Anshoo Yadav**

## TABLE OF CONTENTS

	Declaration	ii
	Certificate	iii
	Abstract	iv
	Acknowledgements	v
	Table of Contents	vi
1	Introduction	9-15
	1.1 Overview	9-10
	1.2 Motivation	10
	1.3 Problem Statement	11-12
	1.4 Expected Outcome	12-15
2	Literature Survey	16-22
3	Design	23-28
	3.1 Data Flow Diagram	23-26
	3.1.1 Level 0 DFD	23
	3.1.2 Level 1 DFD	24-26
	3.2 ER Diagram	27
	3.3 Use Case Diagram	28
4	Proposed Work	29-42
	4.1 Technology Description	29
	4.2 Approach Used	29-33
	4.3 Implementation Details	34-35
	4.4 Challenges Faced	36-37
	4.5 Future Enhancements	38-39
	4.6 Code	40-42
5	Results	43-47
	5.1 Screens and Explanation	43-47
6	Discussions	48-51
	6.1 Performance	48-49
	6.2 Future Research Directions	49-51
7	Conclusion	52-56



8	References	57-59
9	Bibliography	60-62
10	Curriculum Vitae	63-66
11	Mentor Remarks	67

# CHAPTER 1

## INTRODUCTION

### 1.1 Overview

This project details the development of an innovative Software-as-a-Service (SaaS) platform leveraging cutting-edge Artificial Intelligence (AI) to provide users with a suite of powerful image editing capabilities. Addressing the growing need for intuitive and efficient image manipulation tools, this platform offers features such as intelligent image restoration, generative fill for seamless aspect ratio adjustments, precise object removal, realistic object recoloring, and automated background removal.

Built with a modern technology stack, the platform utilizes Next.js for a performant and scalable frontend, enhanced by the type safety and developer experience of TypeScript. Secure user authentication and management are handled through Clerk, while payment processing is seamlessly integrated via Stripe. The core AI-powered image manipulation functionalities are delivered through Cloudinary, a robust cloud-based media management platform offering advanced AI capabilities.

#### **Key Features:**

**Intelligent Image Restoration:** Utilizing AI algorithms, this feature automatically detects and repairs imperfections in images, such as scratches, blemishes, and noise, bringing old or damaged photos back to life.

**Generative Fill for Aspect Ratio Adjustment:** This feature employs generative AI to intelligently extend the boundaries of an image, seamlessly filling in new content to accommodate different aspect ratios without cropping or distortion.

**Precise Object Removal:** Users can easily select and remove unwanted objects from their images. The AI intelligently fills the resulting gaps with plausible and contextually relevant content.

**Realistic Object Recoloring:** This feature allows users to change the color of specific objects within an image while maintaining realistic textures, shadows, and highlights.

**Automated Background Removal:** Leveraging AI-powered segmentation, this feature accurately isolates the foreground subject from the background, enabling users to create transparent backgrounds or replace them with new ones effortlessly.

This AI-powered SaaS platform offers a user-friendly and efficient solution for a wide range of image editing needs. By leveraging the power of AI, it simplifies complex tasks, reduces editing time, and empowers users with professional-quality results without requiring extensive technical expertise. The platform's subscription-based model provides accessible and scalable access to these advanced image manipulation capabilities.

## 1.2 Motivation

The genesis of this AI-powered image manipulation SaaS platform is deeply rooted in the aspiration to democratize sophisticated image editing. We recognized a significant gap between the advanced capabilities offered by professional-grade software and the accessibility afforded to everyday users, content creators, and small businesses. Traditional image editing tools often demand considerable technical proficiency, substantial financial investment, and significant time commitment, effectively creating a barrier for those who could greatly benefit from their power. Our core motivation is to dismantle these barriers by harnessing the transformative potential of Artificial Intelligence.

We envision a future where intricate image manipulation tasks, such as seamlessly restoring damaged photographs, intelligently adapting image aspect ratios for various platforms, precisely removing distracting elements, realistically recoloring objects to match branding or artistic vision, and efficiently isolating subjects from their backgrounds, are no longer the exclusive domain of skilled professionals. By integrating the cutting-edge AI functionalities provided by Cloudinary, we aim to abstract away the complexities of these processes, offering intuitive and user-friendly solutions accessible to anyone with an internet connection.

Furthermore, our motivation is fueled by the belief that AI can act as a powerful catalyst for creativity. By automating tedious and time-consuming editing tasks, we aim to free up users to focus on their artistic vision and narrative. Imagine a content creator effortlessly adapting their visuals for different social media formats without losing crucial details, or a small business owner quickly generating professional-looking marketing materials without needing to hire a graphic designer. This platform is designed to empower individuals and organizations to enhance their visual communication and achieve their creative goals with unprecedented ease and efficiency.

The selection of our technology stack—Next.js for a dynamic and performant user interface, TypeScript for enhanced code reliability and developer experience, Clerk for seamless and secure authentication, Stripe for frictionless payment processing, and Cloudinary for its robust AI-powered image manipulation APIs – directly reflects our commitment to delivering a high-quality, reliable, and user-centric platform. Each technology was chosen strategically to contribute to a seamless and empowering user experience. Ultimately, our driving force is to create a tool that not only simplifies image editing but also unlocks new creative possibilities, making advanced image manipulation accessible to all and fostering a more visually.

## 1.3 Problem Statement

In traditional image editing workflows, a range of challenges hinder accessibility, efficiency, and quality for both individuals and businesses. Professional software often demands a steep learning curve, specialized skills, and costly licenses, making high-quality image manipulation inaccessible to many users. Common editing tasks such as object removal, image restoration, background removal, or resizing to specific aspect ratios can be tedious, time-consuming, and prone to human error. Furthermore, existing tools are rarely optimized for real-time web-based experiences, and AI-based solutions available on the market often suffer from inconsistent performance or limited functionality. These limitations highlight the pressing need for an intelligent, automated, and user-friendly platform that can deliver high-quality image editing powered by AI—directly from the browser, with minimal manual effort:

- **Accessibility Barrier to Professional Editing Tools**

Traditional image editing software like Adobe Photoshop or CorelDRAW requires both technical proficiency and expensive licenses, making it difficult for non-designers, freelancers, and small businesses to access high-quality editing. This creates a gap in the market for users who need powerful yet simple tools without the associated complexity and cost.

- **Time-Consuming Manual Editing**

Tasks such as removing unwanted objects, restoring damaged images, or adjusting backgrounds typically require layer-based editing, masking, and retouching—all of which are labor-intensive and time-consuming. This slows down content creation pipelines, especially for users working under tight deadlines or managing bulk edits.

- **Inconsistent Results with Existing AI Tools**

While some AI tools offer basic automation, many fail to provide consistent, high-quality results across different image types, resolutions, or contexts. Poor object detection, unnatural fills, or hard edges during background removal result in subpar outputs that still require manual correction.

- **Limited Web-Based Editing Platforms**

Most high-end image editors are desktop-based and not accessible directly through a browser. This limits mobility and cross-device usage, which is increasingly important in today's remote-first, mobile-centric world. Users need a responsive and browser-native solution that doesn't compromise on quality.

- **Fragmented Editing Workflows**

Users often need to use multiple tools—one for background removal, another for recoloring, and yet another for cropping or resizing. This fragmented workflow increases friction, causes compatibility issues, and reduces overall productivity.

- **Lack of Integration with Modern SaaS Infrastructure**

Many existing solutions do not integrate well with modern SaaS stacks, such as user authentication, subscription billing, or cloud-based media storage. This makes it harder to

scale these tools commercially or provide a seamless experience to end users.

- **Growing Demand for Intelligent Automation**

As content consumption explodes across platforms like Instagram, Shopify, and YouTube, creators are increasingly seeking tools that can help them generate polished visuals faster and with minimal effort. Current solutions do not fully meet this demand for real-time, intelligent automation tailored to visual content creation

## **1.4 Expected Outcome**

The successful culmination of this project, marked by the robust development and effective launch of our AI-powered image manipulation SaaS platform, is projected to generate a cascade of positive outcomes that will resonate across multiple dimensions. These anticipated impacts span from significantly enhancing the productivity and creative capabilities of our users to establishing a sustainable and thriving platform ecosystem, and even contributing to our collective understanding of integrating advanced artificial intelligence within practical software solutions. We envision these outcomes unfolding both in the immediate aftermath of the launch and as the platform matures and evolves over time.

### **7 Enhanced User Productivity and Efficiency: Reclaiming Time and Simplifying Workflows**

At the forefront of our expected outcomes is a significant elevation in user productivity and efficiency. The current landscape of image editing often demands considerable time and effort, even for relatively straightforward tasks. Manual processes, intricate software interfaces, and the need for specialized skills can create bottlenecks and impede creative flow. Our platform is designed to dismantle these impediments by leveraging the power of AI to automate and streamline key image manipulation workflows.

**Drastic Reduction in Editing Time:** We anticipate a substantial reduction in the time users spend on core image editing tasks. Features such as automated background removal, which traditionally requires meticulous manual selection and refinement, will be executed almost instantaneously and with remarkable accuracy by our AI algorithms powered by Cloudinary. Similarly, intelligent object removal, often a painstaking process of cloning and patching, will be streamlined to a simple selection, with the AI intelligently filling the void with contextually appropriate and visually seamless content. Even tasks like image restoration, which can involve hours of careful digital repair, will be accelerated by AI algorithms capable of identifying and automatically correcting imperfections like scratches, blemishes, and noise. This liberation from time-consuming manual processes will allow users to focus on the more creative and strategic aspects of their work, ultimately boosting their overall output and efficiency.

**Simplification of Complex Editing Procedures:** Many powerful image editing tools come with a steep learning curve, requiring users to invest significant time in mastering complex interfaces and techniques. Our platform is designed with user-centricity at its core, aiming to

make advanced image manipulation accessible to individuals with varying levels of technical expertise. Features like generative fill for aspect ratio adjustment, which traditionally involves either cropping crucial image content or awkwardly stretching it, will be simplified to an intuitive selection of the desired new dimensions. The AI will then intelligently generate and seamlessly integrate new content to fill the expanded canvas, maintaining visual coherence and aesthetic appeal. Similarly, realistic object recoloring, which often requires careful masking and color adjustments to preserve textures and lighting, will be simplified to a selection and color choice, with the AI intelligently handling the nuances of shading and highlights to produce a natural-looking result. This simplification of complex tasks will empower a broader audience to achieve professional-quality results without the need for extensive training or specialized knowledge.

**Empowering Increased Output and Faster Turnaround Times:** By providing a suite of efficient and intuitive AI-powered tools, our platform aims to empower users to process a significantly larger volume of images in a shorter timeframe. This will be particularly beneficial for content creators who need to produce a high volume of visuals for various platforms, social media managers who require rapid adaptation of images for different campaigns, and small businesses that need to quickly generate marketing materials. The ability to achieve high-quality edits quickly will translate directly into faster turnaround times for projects, increased content output, and the ability to respond more agilely to dynamic visual demands.

## **2. Improved Image Quality and Enhanced Creative Possibilities: Unleashing Visual Potential**

Beyond mere efficiency, our platform is expected to contribute significantly to the quality of user-generated visuals and to unlock new avenues for creative expression. The integration of sophisticated AI algorithms will enable users to achieve results that often surpass what is feasible with traditional manual techniques, while also opening new possibilities for artistic exploration.

**Achieving Professional-Grade Results with Ease:** The AI algorithms underpinning our platform, leveraging the advanced capabilities of Cloudinary, are trained on vast datasets of images, enabling them to perform complex editing tasks with a high degree of accuracy and realism. This means that users, regardless of their technical skill level, will be able to achieve professional-looking results in areas such as image restoration, noise reduction, and detail enhancement. The AI will intelligently analyze image content and apply sophisticated processing techniques to optimize visual quality, ensuring that the final output is polished and impactful.

**Expanding the Boundaries of Creative Expression:** Features like generative fill and object recoloring are not just about efficiency; they are also powerful tools for creative exploration. Generative fill allows users to reimagine the composition of their images, seamlessly extending backgrounds or creating entirely new visual contexts. Object recoloring provides the ability to experiment with different color palettes and aesthetics, enabling users to align

visuals with branding guidelines or artistic visions in novel ways. These AI-powered capabilities will empower users to push the boundaries of their creativity and explore visual possibilities that might have been previously time-prohibitive or technically challenging.

**Preserving and Enhancing Visual Heritage through Intelligent Image Restoration:** The intelligent image restoration feature holds the potential to provide significant value by enabling users to revive and enhance old or damaged photographs. AI algorithms can intelligently identify and repair various forms of degradation, such as scratches, tears, fading, and discoloration, bringing cherished memories back to life. This capability not only improves the visual quality of these images but also contributes to their preservation for future generations.

### **3. Platform Growth and Sustainability: Building a Thriving Ecosystem**

The successful adoption and utilization of our platform are intrinsically linked to its long-term growth and sustainability. We anticipate that the value proposition offered by our AI-powered features and user-friendly interface will drive user acquisition and foster a thriving platform ecosystem.

**Attracting a Broad and Diverse User Base:** The platform's accessibility, coupled with its powerful and versatile features, is expected to attract a wide range of users. This includes individuals looking to enhance their personal photos, content creators seeking efficient tools for visual production, small businesses aiming to create professional marketing materials, and even larger organizations looking for scalable and cost-effective image editing solutions. This diverse user base will contribute to the platform's vibrancy and growth.

**Establishing a Robust and Recurring Revenue Stream through Subscription Model:** Our chosen subscription-based model, seamlessly integrated with Stripe's secure payment processing infrastructure, is projected to generate a stable and predictable recurring revenue stream. This predictable revenue will be crucial for ensuring the platform's long-term financial health, enabling us to invest in ongoing development, infrastructure maintenance, customer support, and the integration of new and innovative AI-powered features.

**Cultivating Positive User Feedback and Fostering Community:** We anticipate that the ease of use and the high-quality results delivered by our platform will lead to positive user feedback. This positive sentiment will be instrumental in driving organic growth through word-of-mouth referrals and positive reviews. Furthermore, we envision fostering a community around our platform, where users can share their creations, exchange tips and tricks, and provide valuable feedback that will contribute to the platform's continuous improvement.

**Laying the Foundation for Future Feature Expansion and Innovation:** The modular architecture of our platform, combined with our strategic partnership with Cloudinary and their continuously evolving AI capabilities, provides a solid foundation for future feature expansion and innovation. As AI technology advances, we will be well-positioned to integrate

new and cutting-edge image manipulation tools, ensuring that our platform remains at the forefront of the industry and continues to provide increasing value to our users.

#### **4. Technological Advancement and Knowledge Acquisition: Contributing to the AI Landscape**

Beyond the direct benefits to our users and the platform's sustainability, we also anticipate that this project will contribute to the broader landscape of technological advancement and knowledge acquisition.

**Demonstrating the Efficacy of AI Integration in SaaS Solutions:** This project will serve as a practical and tangible demonstration of how advanced AI algorithms can be effectively integrated into a SaaS platform to solve real-world problems and significantly enhance user experiences. By showcasing the power and versatility of AI in image manipulation, we hope to inspire further innovation and adoption of AI across various software domains.

**Gaining Valuable Insights into User Behavior and Preferences:** The development and subsequent usage of our platform will provide us with invaluable data and insights into how users interact with AI-powered image editing tools. By monitoring feature utilization, user feedback, and overall platform engagement, we will gain a deeper understanding of user needs and preferences, which will inform future development decisions and platform optimizations, ensuring that we continue to meet and exceed user expectations.

**Enhancing the Team's Expertise in Modern Web Technologies and AI Integration:** The process of developing this platform, utilizing technologies such as Next.js, TypeScript, Clerk, Stripe, and Cloudinary's AI APIs, will significantly enhance the skills and expertise of our development team. We will gain valuable hands-on experience in building scalable and performant web applications, implementing secure authentication and payment processing, and effectively integrating cloud-based AI services. This acquired knowledge and experience will be invaluable for future projects and will contribute to our collective understanding of building cutting-edge software solutions.

In conclusion, the expected outcomes of this AI-powered image manipulation SaaS platform extend far beyond simply providing a set of editing tools. We anticipate empowering our users with unprecedented efficiency and creative freedom, establishing a sustainable and thriving platform ecosystem, and contributing meaningfully to the advancement and understanding of artificial intelligence within the realm of practical software applications. The realization of these expected outcomes will mark a significant step towards our vision of democratizing advanced image editing and unlocking the full visual potential of individuals.



## CHAPTER 2

### LITERATURE SURVEY

This project details the development of an innovative Software-as-a-Service (SaaS) platform leveraging cutting-edge Artificial Intelligence (AI) to provide users with a suite of powerful image editing capabilities. Addressing the growing need for intuitive and efficient image manipulation tools, this platform offers features such as intelligent image restoration, generative fill for seamless aspect ratio adjustments, precise object removal, realistic object recoloring, and automated background removal.

Built with a modern technology stack, the platform utilizes Next.js for a performant and scalable frontend, enhanced by the type safety and developer experience of TypeScript. Secure user authentication and management are handled through Clerk, while payment processing is seamlessly integrated via Stripe. The core AI-powered image manipulation functionalities are delivered through Cloudinary, a robust cloud-based media management platform offering advanced AI capabilities.

#### Key Features:

- **Intelligent Image Restoration:** Utilizing AI algorithms, this feature automatically detects and repairs imperfections in images, such as scratches, blemishes, and noise, bringing old or damaged photos back to life.
- **Generative Fill for Aspect Ratio Adjustment:** This feature employs generative AI to intelligently extend the boundaries of an image, seamlessly filling in new content to accommodate different aspect ratios without cropping or distortion.
- **Precise Object Removal:** Users can easily select and remove unwanted objects from their images. The AI intelligently fills the resulting gaps with plausible and contextually relevant content.
- **Realistic Object Recoloring:** This feature allows users to change the color of specific objects within an image while maintaining realistic textures, shadows, and highlights.
- **Automated Background Removal:** Leveraging AI-powered segmentation, this feature accurately isolates the foreground subject from the background, enabling users to create transparent backgrounds or replace them with new ones effortlessly.

This AI-powered SaaS platform offers a user-friendly and efficient solution for a wide range of image editing needs. By leveraging the power of AI, it simplifies complex tasks, reduces editing time, and empowers users with professional-quality results without requiring extensive technical expertise. The platform's subscription-based model provides accessible and scalable access to these advanced image manipulation capabilities.

### **Literature Survey: AI-Powered Image Editing SaaS Platform**

The convergence of Artificial Intelligence (AI) and cloud computing has revolutionized numerous fields, and digital image editing is no exception. This literature survey provides a comprehensive overview of the recent advancements in AI-powered image editing, the underlying algorithms and models, and the technologies that enable the development and deployment of such solutions through Software-as-a-Service (SaaS) platforms. The survey will delve into the specific AI techniques employed for various image manipulation tasks, the architectural considerations for building scalable SaaS platforms, and the key technologies involved in frontend development, backend services, user management, and payment processing.

### **AI in Image Processing: A Paradigm Shift**

Traditional image editing often involved manual, time-consuming processes requiring significant expertise. AI, particularly deep learning, has automated and enhanced these processes, enabling complex transformations with unprecedented speed and accuracy. Deep learning models, especially Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), have become the cornerstone of AI-powered image editing.

- **Convolutional Neural Networks (CNNs):** CNNs are a class of deep neural networks designed to process structured grid data, such as images. Their ability to automatically learn hierarchical representations of visual features, from simple edges and textures to complex objects and scenes, has made them highly effective for various image processing tasks.
- **Generative Adversarial Networks (GANs):** GANs consist of two neural networks, a generator and a discriminator, competing in a zero-sum game. The generator learns to create synthetic data that resembles the real data, while the discriminator learns to distinguish between real and synthetic data. This adversarial process drives both networks to improve, resulting in the generation of highly realistic synthetic images.

### **Core AI-Powered Image Editing Techniques**

This section explores the specific AI techniques used to implement the core features of the

#### **Intelligent Image Restoration**

Image restoration aims to recover a degraded image to its original, pristine form. Degradation can arise from various sources, including noise, blur, and compression artifacts.

- **Denoising:** Removing noise from images is a fundamental restoration task. Traditional

denoising methods often rely on filtering techniques, which can smooth out noise but also blur important image details. Deep learning-based methods, such as DnCNN (Zhang et al., 2017), have demonstrated superior performance by learning to directly estimate and remove the noise component. DnCNN utilizes residual learning to focus on the noise itself, leading to more effective noise removal and better preservation of image details.

- **Deblurring:** Blur can occur due to camera shake, object motion, or out-of-focus optics. AI-powered deblurring techniques aim to reverse this process and recover a sharp image. Deep learning models can learn complex blur kernels and effectively restore sharp details, even in cases
- **Artifact Removal:** Compression artifacts, such as blocking and ringing, can degrade the visual quality of images, especially at low bitrates. AI-based methods can learn to identify and suppress these artifacts, improving the perceptual quality of compressed images. Autoencoders and other deep learning architectures can be trained to map degraded images to their clean counterparts, effectively removing compression artifacts.

### **Generative Fill for Aspect Ratio Adjustment**

Generative fill, also known as image inpainting, involves filling in missing or unknown regions of an image. This capability is crucial for tasks such as aspect ratio adjustment, object removal,

- **Context Encoders:** Pathak et al. (2016) introduced Context Encoders, a pioneering GAN-based approach for inpainting. Context Encoders consist of an encoder-decoder network trained to predict the missing content from the surrounding pixels. The generator network learns to synthesize plausible content, while the discriminator network ensures that the generated content is realistic and consistent with the image context.
- **Advanced Inpainting Models:** More recent models, such as LaMa (Suvorov et al., 2021) and those based on latent diffusion models, have further advanced the state-of-the-art in inpainting. LaMa utilizes a transformer-based architecture with a large receptive field, enabling it to handle larger missing regions and capture long-range dependencies in the image. Latent diffusion models, which have shown remarkable success in image generation, can also be adapted for inpainting, producing highly realistic and coherent fill-ins.

### **Precise Object Removal**

Object removal typically involves a combination of object segmentation and image inpainting.

- **Object Segmentation:** Accurate object segmentation is essential for isolating the object to be removed from its background. Deep learning models, such as U<sup>2</sup>-Net (Qin et al., 2020), have achieved high accuracy in generating saliency maps, which highlight the objects of interest in an image. These models learn to identify objects based on their visual features and spatial context,

- **Seamless Inpainting:** Once the object is segmented, inpainting techniques are used to fill the resulting gap with plausible background content. Models like DeepFill v2 (Yu et al., 2019) employ advanced generative networks and contextual attention mechanisms to synthesize background regions that seamlessly blend with the surrounding areas.

### **Realistic Object Recoloring**

Object recoloring involves changing the color of specific objects within an image while preserving their original texture, shading, and lighting.

- **Colorization Networks:** Early approaches to image colorization focused on adding color to grayscale images. Zhang et al. (2016) proposed a colorization network that leverages class-rebalancing and feature embeddings to produce realistic colorizations.
- **Conditional GANs (cGANs):** cGANs can be used for more sophisticated recoloring tasks, where the desired color is specified as a condition. By conditioning the generator network on the target color, cGANs can generate recolored images that adhere to the specified color.

### **Automated Background Removal**

Background removal aims to isolate the foreground object from its background, enabling various applications such as image compositing and virtual background replacement.

- **Semantic Segmentation:** Semantic segmentation plays a crucial role in background removal by classifying each pixel in an image as either foreground or background. Deep learning models, such as DeepLabV3+ and MODNet (Ke et al., 2020), have achieved state-of-the-art performance in semantic segmentation, providing accurate and detailed segmentation maps.
- **Matting:** While semantic segmentation provides a coarse foreground/background separation, matting techniques estimate the opacity of each pixel, allowing for more precise isolation of the foreground object, particularly around the edges.

### **SaaS Architecture for AI Applications: Key Considerations**

Deploying AI-powered image editing capabilities through a SaaS model requires careful consideration of the underlying architecture. The architecture must be designed to handle the computational demands of AI processing, ensure scalability to accommodate a growing user base, and provide a seamless and responsive user experience.

- **Scalability:** AI-powered image editing can be computationally intensive, especially for high-resolution images and complex editing tasks. The SaaS architecture must be able to scale its resources dynamically to handle fluctuating user loads and ensure that the platform remains responsive even during peak usage. Cloud computing platforms, such as AWS, Azure, and Google Cloud, provide the necessary infrastructure and services to achieve scalability.
- **Performance:** Users expect image editing operations to be performed quickly and efficiently.

The SaaS architecture must be optimized for performance to minimize processing time and deliver a smooth user experience. This can involve techniques such as caching, load balancing, and asynchronous processing.

- **Cost-Effectiveness:** Building and maintaining a SaaS platform can be expensive. The architecture should be designed to optimize resource utilization and minimize costs. Cloud computing platforms offer various pricing models, such as pay-as-you-go, which can help SaaS providers manage their expenses effectively.
- **Microservices Architecture:** A microservices architecture, where the application is decomposed into a collection of small, independent services, can offer several advantages for AI-powered SaaS platforms. Microservices can be scaled independently, allowing for more efficient resource allocation. They can also be developed and deployed independently, enabling faster development cycles and easier maintenance.

### **Frontend Technologies: Building the User Interface**

The frontend of the SaaS platform is responsible for providing users with an intuitive and user-friendly interface for accessing the AI-powered image editing features.

- **Next.js:** Next.js is a popular React framework that offers several features that are particularly beneficial for building SaaS applications, including server-side rendering (SSR), static site generation, and API routing. SSR improves SEO and initial page load performance, while static site generation enables the creation of highly performant and scalable websites.
- **TypeScript:** TypeScript is a superset of JavaScript that adds static typing to the language. Using TypeScript in the frontend development process can improve code quality, enhance developer productivity, and reduce the risk of runtime errors.
- **React:** React is a JavaScript library for building user interfaces. Its component-based architecture allows developers to create reusable UI elements, making it easier to build and maintain complex applications.

### **Backend Services: Handling AI Processing and Data Management**

The backend of the SaaS platform is responsible for handling the AI processing tasks, managing user data, and providing APIs for the frontend to interact with.

- **Cloudinary:** Cloudinary is a cloud-based media management platform that provides a comprehensive suite of services for storing, transforming, and delivering images and videos. Its AI-powered features, such as background removal, object detection, and image enhancement, can be easily integrated into a SaaS platform through its APIs. By leveraging Cloudinary's capabilities, developers can offload the computationally intensive AI processing

tasks to a dedicated platform, simplifying the backend development process and improving scalability.

- **Serverless Computing:** Serverless computing platforms, such as AWS Lambda, Azure Functions, and Google Cloud Functions, allow developers to run code without provisioning or managing servers. This can be a cost-effective solution for handling AI processing tasks, as the platform automatically scales resources based on demand.

#### **User Management and Authentication: Securing User Data**

Secure user management and authentication are essential for any SaaS platform, especially those that handle sensitive user data.

- **Clerk:** Clerk is a user authentication and management platform that provides pre-built UI components and APIs for handling user signup, login, password management, and multi-factor authentication. Its seamless integration with frontend frameworks like Next.js simplifies the process of adding secure user authentication to SaaS applications.

#### **Payment Processing: Enabling Subscriptions and Transactions**

For SaaS platforms that offer paid subscriptions or usage-based pricing, a reliable and secure payment processing system is crucial.

- **Stripe:** Stripe is a popular payment processing platform that provides a comprehensive set of APIs and tools for handling online payments. It supports various payment methods, including credit cards, debit cards, and digital wallets, and offers features such as recurring billing, subscription management, and fraud prevention. Stripe's robust infrastructure and PCI DSS compliance ensure the secure handling of sensitive payment data.

- 

#### **The Competitive Landscape: Existing Solutions and Opportunities**

Several commercial platforms offer AI-based image editing features, each with its strengths and limitations.

- **Remove.bg:** Remove.bg specializes in AI-powered background removal, providing a simple and efficient solution for this specific task.
- **Cleanup.pictures:** Cleanup.pictures focuses on object removal, utilizing inpainting techniques to seamlessly remove unwanted objects from images.
- **Adobe Firefly and Canva Magic Edit:** Adobe Firefly and Canva Magic Edit integrate generative AI into their respective design platforms, enabling users to perform tasks such as image expansion and generative fill within a broader design workflow.

While these platforms offer valuable AI-powered image editing capabilities, there remains an opportunity for a more comprehensive and versatile solution. A SaaS platform that combines multiple AI-powered editing features, is built on a modern technology stack, and offers a

developer-friendly API could cater to a wider audience of users and developers.

### **Challenges and Future Directions**

The field of AI-powered image editing is constantly evolving, with ongoing research and development leading to new and improved techniques. However, several challenges remain:

- **Ethical Considerations:** The ability of AI to manipulate images raises ethical concerns regarding the potential for misuse, such as creating deepfakes or spreading misinformation. It is crucial to develop guidelines and safeguards to ensure the responsible use of this technology.
- **Bias and Fairness:** AI models can perpetuate and amplify biases present in the training data, leading to unfair or discriminatory outcomes. It is essential to address these biases and develop AI algorithms that are fair and equitable.
- **Computational Cost:** Some AI-powered image editing tasks can be computationally expensive, requiring significant processing power and memory. Developing more efficient algorithms and hardware is crucial for enabling real-time or near real-time performance.

Despite these challenges, the future of AI-powered image editing is bright. Advancements in deep learning, computer vision, and cloud computing will continue to drive innovation in this field, leading to even more powerful and accessible image editing tools.

### **Conclusion**

AI has revolutionized digital image editing, empowering users with unprecedented capabilities to manipulate and enhance images. The development of SaaS platforms that leverage AI has made these advanced tools more accessible and affordable, democratizing access to professional-quality image editing. By combining cutting-edge AI algorithms with modern web technologies, it is possible to create innovative SaaS solutions that cater to a wide range of image editing needs. This literature survey has provided an overview of the key AI techniques, architectural considerations, and technologies involved in building such platforms, highlighting the potential of AI to transform the future of digital image editing.

## CHAPTER 3

### DESIGN

#### 3.1 Data Flow Diagram

##### 3.1.1 Level 0 Data Flow Diagram

Level 0 Data Flow Diagram will explain the basic flow of data in a system which shows how the new or old user will interact with the system.

Fig. 3.1 Level 0 DFD of Imaginify

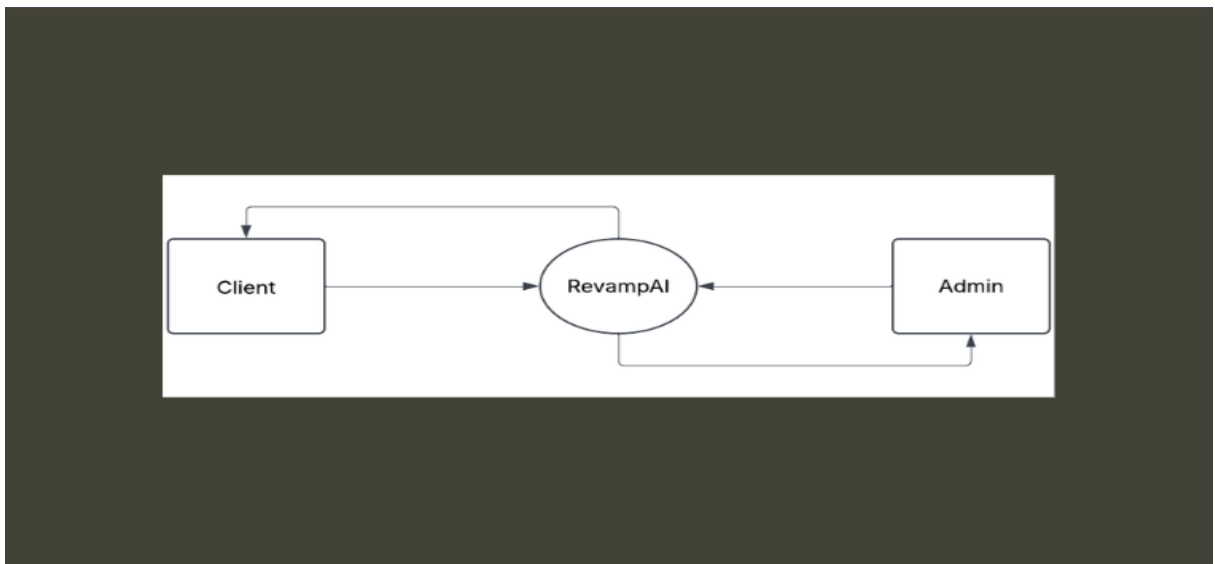


Fig. 3.1 The Level 0 Data Flow Diagram (DFD) provides a high-level overview of the AI-based SaaS image editing platform. It represents the system as a single process and illustrates its interaction with external entities such as users, authentication services, payment gateways, and the AI media processing engine. At this level, the platform receives input from users in the form of image uploads, editing requests (such as image restoration, object removal, generative fill, background removal, and recoloring), and account management activities. Users also interact with the system to retrieve the processed images and manage their profiles or subscriptions.



### 3.1.2 Level 1 Data Flow Diagram

Level 1 Data Flow Diagram will explain the basic flow of data in a system which shows how the new or old user will interact with the system with different processes.

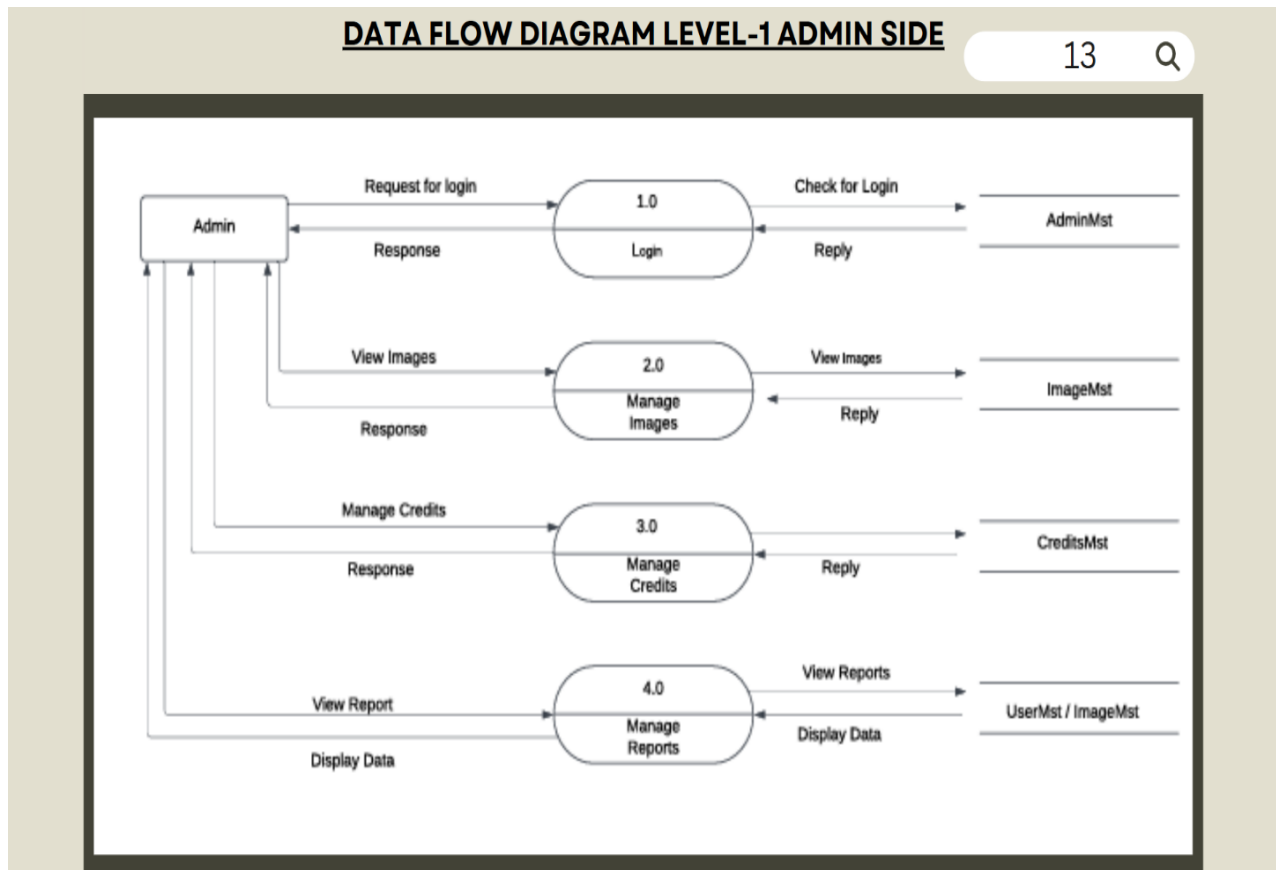


Fig. 3.2 Level 1 DFD of Imaginify

The admin accesses the system through a secure authentication process managed by Clerk, which verifies login credentials and grants access to the admin dashboard. Once authenticated, the admin can perform several functions including user management, subscription monitoring, and system analytics review.

In the user management module, the admin can view registered users, manage user roles, deactivate accounts, and monitor platform activity. Through the subscription module, the admin receives data from Stripe, allowing them to track billing information, subscription status, and payment histories. The analytics and monitoring module provides insights into

platform usage, AI feature usage statistics, and error logs, typically stored in a database or analytics dashboard integrated with the platform.

Additionally, the admin can also manage AI model performance by accessing logs or usage reports from Cloudinary to assess the health and efficiency of image processing tasks. These subprocesses work together to ensure smooth platform operation, service quality, and secure administration.

This Level 1 DFD reveals the back-end responsibilities of the admin role, emphasizing how various internal modules are connected to external services and system databases to maintain control, security, and scalability of the platform.

This Level 1 DFD reveals the back-end responsibilities of the admin role, emphasizing how various internal modules are connected to external services and system databases to maintain control, security, and scalability of the platform.

## DATA FLOW DIAGRAM LEVEL-1 USER SIDE

12

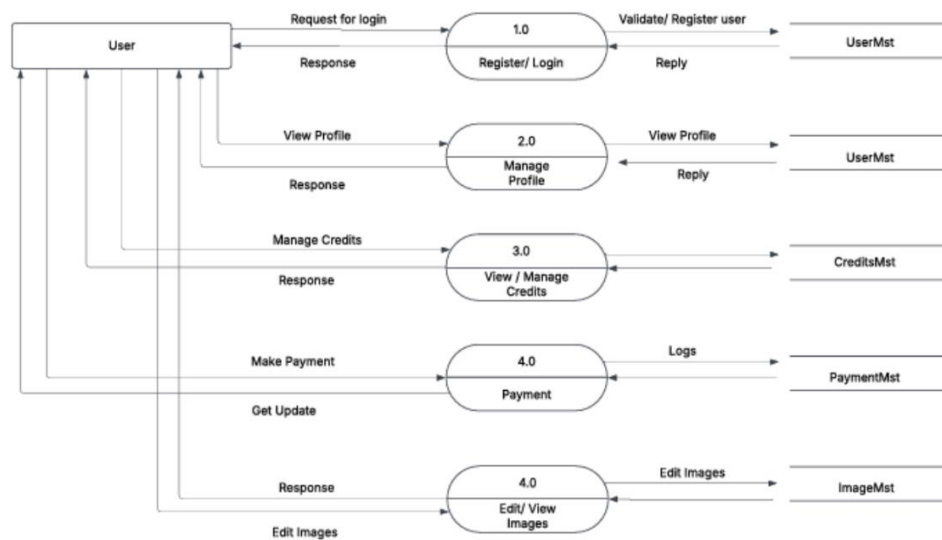


Fig. 3.2 Level 1 DFD of Imaginify

Fig. 3.2 the user side breaks down the main functionalities available to end users, offering a detailed view of the internal operations involved in the user interaction with the platform. After accessing the platform, the user first interacts with the authentication module, which communicates with Clerk to handle secure login or registration. Upon successful authentication, users gain access to the main dashboard where they can upload images and select various AI-powered editing features.

The image upload module handles the user's input, passing uploaded media to Cloudinary for secure storage and retrieval. Users can then choose specific tools like image restoration, generative fill, background removal, object removal, or object recoloring—each of which routes through the image processing module. This module acts as a middleware that sends editing requests to Cloudinary's AI engine or other integrated APIs, then receives and delivers the processed images back to the user interface.

## 3.2 ER Diagram

An Entity Relationship Diagram is a diagram that represents relationships among entities in a database.

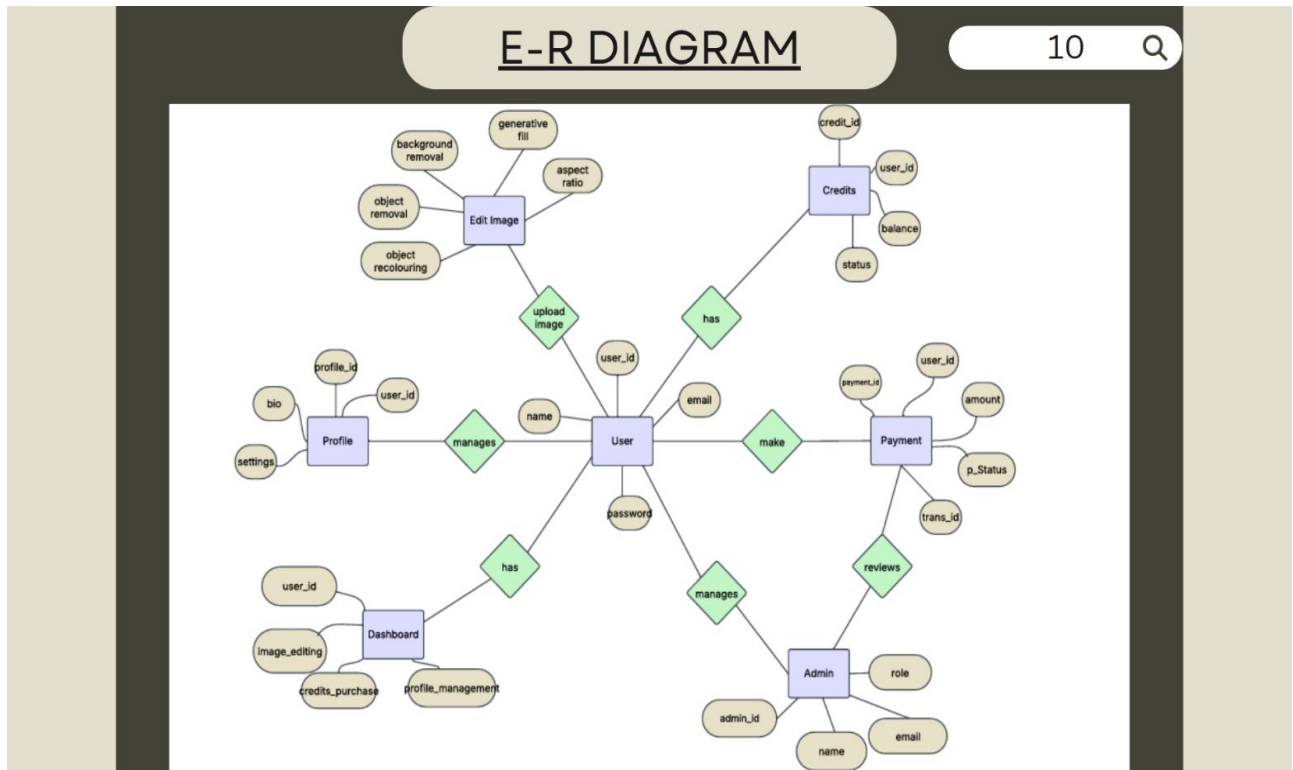


Fig. 3.3 E-R Diagram of Imaginify

### 3.3 Use Case Diagram

In Use Case Diagram we elaborate about the purpose, actor, pre-condition, post-condition, basic flow, and alternate flow of all the use cases. In our system there are two actors, one is a client and other is the admin who interacts with the use cases of the course and quizzes. It explains the details and conditions of the system to be fulfilled to successfully complete each use case.

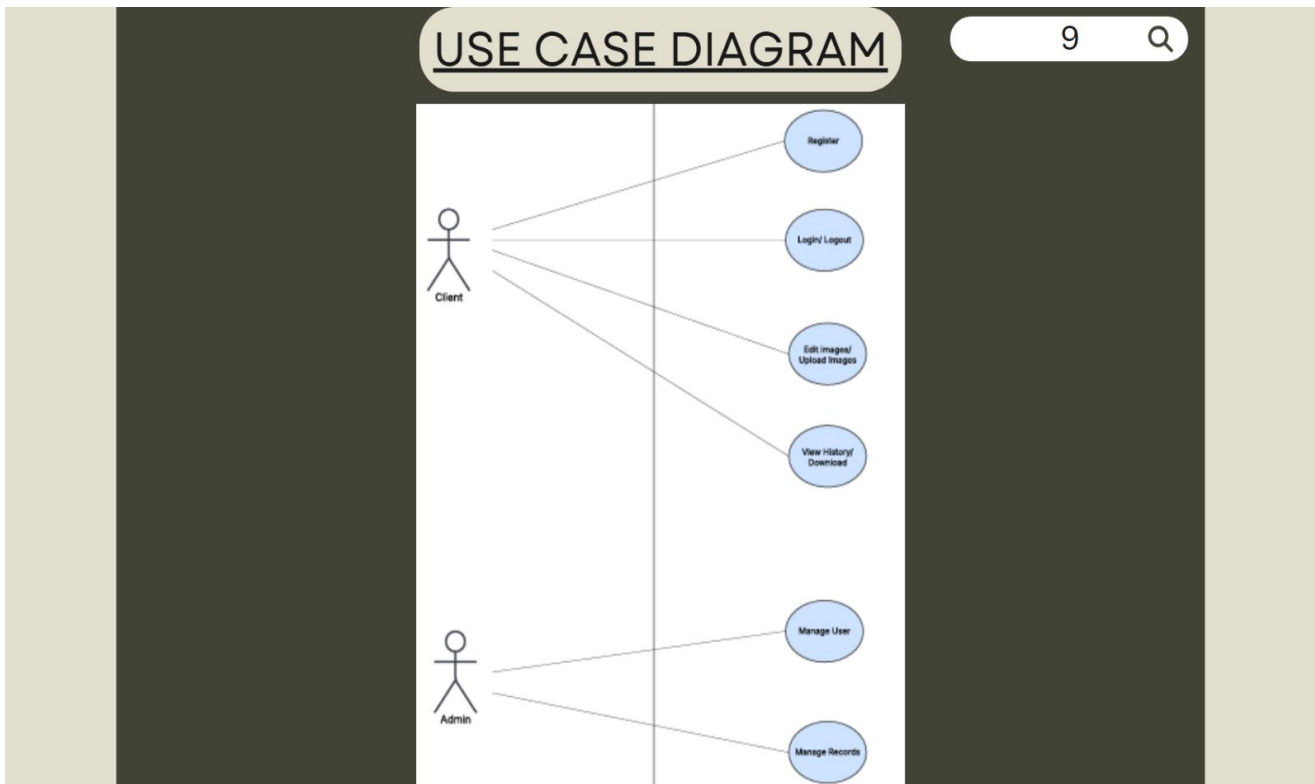


Fig. 3.4 Use Case Diagram of Imaginify

## CHAPTER 4

### PROPOSED WORK

#### 4.1 Technology Description

- **Selection of Operating System:** Our website is platform independent, so it does not depend on the operating system.
- **Selection of Software:** Next.js, TypeScript, Cloudinary, Stripe, Clerk.
- **Languages Used:** TypeScript, Next.js.

#### 4.2 Approach Used

To overcome the limitations of traditional, locally-installed image editing tools and cater to a modern, cloud-centric audience, this project adopts a **modular, cloud-native, and AI-first approach**. The platform's architecture is designed to support extensibility, high availability, and seamless user interaction. Below is a breakdown of its key architectural components and design choices:

- **Frontend Development**

The frontend is developed using **Next.js** and **TypeScript**, forming the backbone of the user interface. The advantages of this setup include:

1. **Server-Side Rendering (SSR) and Static Site Generation (SSG):** Improves performance and SEO.
2. **Client-Side Navigation:** Enables fast, smooth transitions between views.
3. **Component-Based Architecture:** Promotes reusable and maintainable UI components.
4. **Tailwind CSS:** Used for consistent, responsive design and rapid styling with utility classes.

The result is a clean, modern interface that is both fast and intuitive, allowing users to engage with complex image editing tools without a steep learning curve.

- **Authentication & User Management**

User authentication and access control are handled by Clerk, which offers:

- Secure sign-in/sign-up with support for email/password, social logins, and OTP.
- Session handling across tabs and devices.
- Role-based access for free and premium users.
- Built-in support for secure password resets, two-factor authentication (2FA), and user profile management.

This ensures that only authenticated users can access their saved work and premium features, while also simplifying user management for the development team.

- **Payment Integration**

**Stripe** is employed to manage subscriptions, billing logic, and secure payment processing. Key benefits include:

- **Checkout Flows:** Pre-built, responsive checkout pages that streamline payments.
- **Subscription Plans:** Support for multiple billing tiers (e.g., free, monthly, yearly).
- **Webhook Integration:** Real-time event handling to update user access status based on payments, cancellations, or trial periods.
- **Usage Monitoring:** For future metered billing implementations.

The Stripe-Clerk integration ensures smooth state synchronization between payment and user access status, enhancing reliability and preventing unauthorized access to paid features.

- **Image Upload and Storage**

The platform uses **Cloudinary** for handling image storage and transformation. Cloudinary offers:

- Optimized storage with automatic image compression and format conversion.
- CDN-based delivery for lightning-fast global access.
- AI-based editing tools like background removal and auto-tagging.
- On-the-fly image transformation through URL-based APIs.

All user-uploaded images are securely stored in Cloudinary buckets and are accessible only to authorized users. Transformations are applied either in real-time or through queued serverless processes depending on the task complexity

- **AI-Based Image Processing**

AI-driven image transformation is the platform's core functionality. These features are built using a combination of **Cloudinary APIs** and third-party or proprietary **deep learning models** deployed on serverless infrastructure. Key AI features include:

- **1. Image Restoration**

**Super-Resolution GANs (SRGAN)** are used to upscale and restore old or blurry images by reconstructing missing details and removing artifacts.

- **2. Object Removal & Generative Fill**

Uses models based on **Mask R-CNN**, **inpainting GANs**, and **partial convolution networks** to intelligently erase unwanted objects and reconstruct missing background pixels.

- **3. Background Removal**

Employs segmentation models such as **U-2-Net** and **custom-trained Unets** to accurately detect and remove the background. Users can then replace it with solid colors, gradients, or uploaded images.

- **4. Object Recoloring**

Utilizes **CNN-based colorization** and **style transfer learning** techniques to change object colors without affecting lighting or textures.

These AI models are triggered via API calls or serverless function execution depending on performance demands and transformation type.

- **Backend & API Layer**

- The backend layer is implemented using **Next.js API routes**, allowing the project to use serverless functions. These functions handle:
- Secure calls to **Stripe APIs** for payment verification.
- Requests to **Cloudinary** for applying AI transformations.
- User access validation via **Clerk** tokens.
- Transformation logging, usage tracking, and temporary caching.
- (Optional) Database integrations for storing project history or presets.

This lightweight backend model ensures scalability, reduced maintenance overhead, and cost efficiency



- **Deployment & Hosting**

The platform is deployed on **Vercel**, which offers seamless integration with Next.js, CI/CD pipelines, and automatic scaling. Media assets are CDN-delivered via Cloudinary for fast, global access

#### 4.2.1 Technologies Used

- ✓ **Frontend:** Next.js, TypeScript, React.js, Tailwind CSS.
- ✓ **Backend:** Stripe, Clerk, Cloudinary, Next.js.

#### 4.2.2 Features

The platform offers a comprehensive set of AI-powered image editing features designed to simplify complex editing tasks, deliver professional-quality results, and enhance user productivity. Below are the core features:

##### **Image Restoration**

Utilizes AI models to automatically restore old, damaged, or low-quality images. This includes fixing scratches, blurriness, noise, and missing parts, bringing images back to life with improved clarity and detail.

##### **Generative Fill (Aspect Ratio Adjustment)**

Powered by generative AI and inpainting techniques, this feature intelligently fills in extended parts of an image to match a desired aspect ratio—ideal for creating social media posts, banners, or thumbnails without cropping important elements.

##### **Object Removal**

Allows users to remove unwanted objects or elements from an image seamlessly. The AI model predicts and fills the removed area with appropriate background texture, making the edit appear natural and undetectable.

##### **Object Recoloring**

This feature enables users to change the color of specific objects (e.g., clothing, furniture, products) while maintaining lighting, shading, and texture integrity. Ideal for product design previews or creative visualizations.

## **Background Removal**

Automatically detects and removes backgrounds from images with high accuracy using deep learning-based segmentation models. Users can replace it with solid colors, gradients, or custom backgrounds.

## **Secure User Authentication**

Powered by **Clerk**, the platform offers secure and seamless authentication, including user registration, login, password recovery, and session management.

## **Subscription and Payment Integration**

Users can subscribe to premium plans through **Stripe**, enabling access to advanced features. The billing system includes support for monthly/yearly plans, usage limits, and real-time payment status updates.

## **Cloud-Based Storage and Processing**

All image uploads and transformations are handled via **Cloudinary**, providing fast, scalable, and secure cloud storage, along with AI-based media optimization.

## **User-Friendly Interface**

Built with **Next.js** and **TypeScript**, the platform delivers a fast, responsive, and intuitive user experience with real-time previews and easy-to-use editing tools.

## **Cross-Device Accessibility**

As a SaaS application, the platform is accessible via any modern browser across devices—desktops, laptops, and tablets—without the need for installation.

## 4.3 Implementation Details

### Frontend Development

The frontend is developed using **Next.js**, a React-based framework that enables server-side rendering and static site generation for better performance and SEO. **TypeScript** is used for type safety and maintainability, ensuring scalability as the application grows. The UI includes responsive design principles for cross-device compatibility and a smooth user experience.

### Authentication and Authorization

User authentication is managed through **Clerk**, which handles login, sign-up, multi-factor authentication (MFA), and session management. Role-based access control (RBAC) is implemented to differentiate between regular users and admin-level users for dashboard access and feature permissions.

### Payment Integration

The subscription and billing functionality is powered by **Stripe**. Users can choose between different pricing tiers and manage their subscriptions through a secure billing portal. Webhooks from Stripe are used to trigger access control changes in real time based on payment status.

### Image Processing Workflow

Users upload images via the web interface, which are then sent to **Cloudinary**. Cloudinary handles:

1. **Storage** of uploaded images.
2. **Transformation** using built-in AI features for background removal, object detection, image restoration, and more.
3. **Delivery** via optimized CDN URLs for fast image retrieval and download.
4. Each editing request is managed through an API endpoint that validates user authentication and subscription status before processing the image via Cloudinary's SDK and transformations.

### AI Features and Models

The AI editing features leverage a mix of **Cloudinary's AI APIs** and optionally external AI models for more advanced tasks:

1. **Image Restoration** using deep learning-based super-resolution models.

2. **Generative Fill** powered by inpainting techniques and GANs.
3. **Object Removal** based on segmentation and completion models like Mask R-CNN.
4. **Recoloring and Enhancement** using colorization and style-transfer models.
5. **Background Removal** using deep semantic segmentation.

### **Admin Panel Functionality**

Admins access a protected dashboard where they can:

1. View user activity and edit history.
2. Monitor feature usage analytics.
3. Manage user accounts and subscriptions.
4. Review logs and error reports for performance optimization.

### **Database and Storage**

User metadata, subscription status, and image history are stored in a backend database (e.g., PostgreSQL or Firebase, depending on chosen stack). Cloudinary handles all media file storage and transformation logic.

### **Security and Compliance**

Security measures include:

1. HTTPS for secure communication.
2. OAuth2.0 and JWTs for session validation.
3. Data validation and sanitization on all endpoints.
4. Compliance with data protection standards (e.g., GDPR) for handling user images and personal data.

## 4.4 Challenges Faced

Throughout the development and deployment of the AI-based image editing SaaS platform, a range of technical and operational challenges emerged. These challenges spanned across architecture design, third-party service integration, user experience, and resource management. Overcoming these obstacles required iterative improvements, deep technical understanding, and user-centric design thinking. Below is a detailed account of the most significant challenges encountered during the platform's lifecycle:

### 1. Integrating AI Features with Frontend in Real Time

1. One of the most technically demanding aspects of the platform was the real-time integration of advanced AI features—such as image inpainting, object removal, background substitution, and recoloring—into the frontend. These operations are computationally intensive and typically run on cloud GPUs, which naturally introduce latency.
2. To bridge this gap and provide a seamless user experience, asynchronous communication patterns were established between the frontend and backend. Background processes handled requests while users received visual feedback through loading indicators or progress bars. Lazy-loading and preview generation strategies were implemented to allow near-instantaneous previews for lower-resolution images, followed by higher-quality versions rendered in the background.
3. Balancing processing time, image fidelity, and responsive interaction was essential to prevent user drop-off and ensure smooth engagement.

### 2. Dependency on Third-Party APIs

The platform relies on services like **Cloudinary** for AI features, **Clerk** for authentication, and **Stripe** for billing. Ensuring these services worked harmoniously and handling their rate limits, API changes, and occasional downtimes added complexity to system stability and maintenance.

### 3. Maintaining Image Quality after Transformation

AI transformations, particularly those involving object removal or generative fill, often result in artifacts or loss of detail if not handled carefully. Selecting the right model configurations and applying post-processing filters were necessary to maintain high-quality outputs that meet user expectations.

### 4. User Subscription & Access Control Logic

Implementing fine-grained access control based on subscription tiers through Stripe was non-trivial. Ensuring that premium features were only accessible to subscribed users and handling edge cases (like subscription cancellation or payment failure) required careful webhook

management and state synchronization.

## **5. Cost Management of Cloud Services**

The frequent use of cloud-based AI transformations, especially during testing, increased Cloudinary API usage, impacting budget allocation. Rate-limiting and optimizing transformation logic were necessary to reduce unnecessary API calls and stay within service quotas.

## **6. Handling Large Image Uploads**

Users sometimes uploaded very high-resolution images, which introduced performance issues during upload, transformation, and preview. Implementing pre-upload compression, size checks, and background processing helped mitigate these bottlenecks.

## **7. Cross-Browser and Device Compatibility**

Ensuring that the application rendered and behaved consistently across different devices and browsers (e.g., Chrome, Safari, mobile browsers) presented UI/UX challenges. Responsive design and extensive testing were essential to deliver a seamless experience.

## **8. User Feedback and UX Improvements**

As a creative tool, the platform's usability is critical. Gathering real user feedback and improving UI/UX elements such as tool previews, loading indicators, undo/redo functionality, and real-time feedback was a continuous process

## 4.5 Future Enhancements

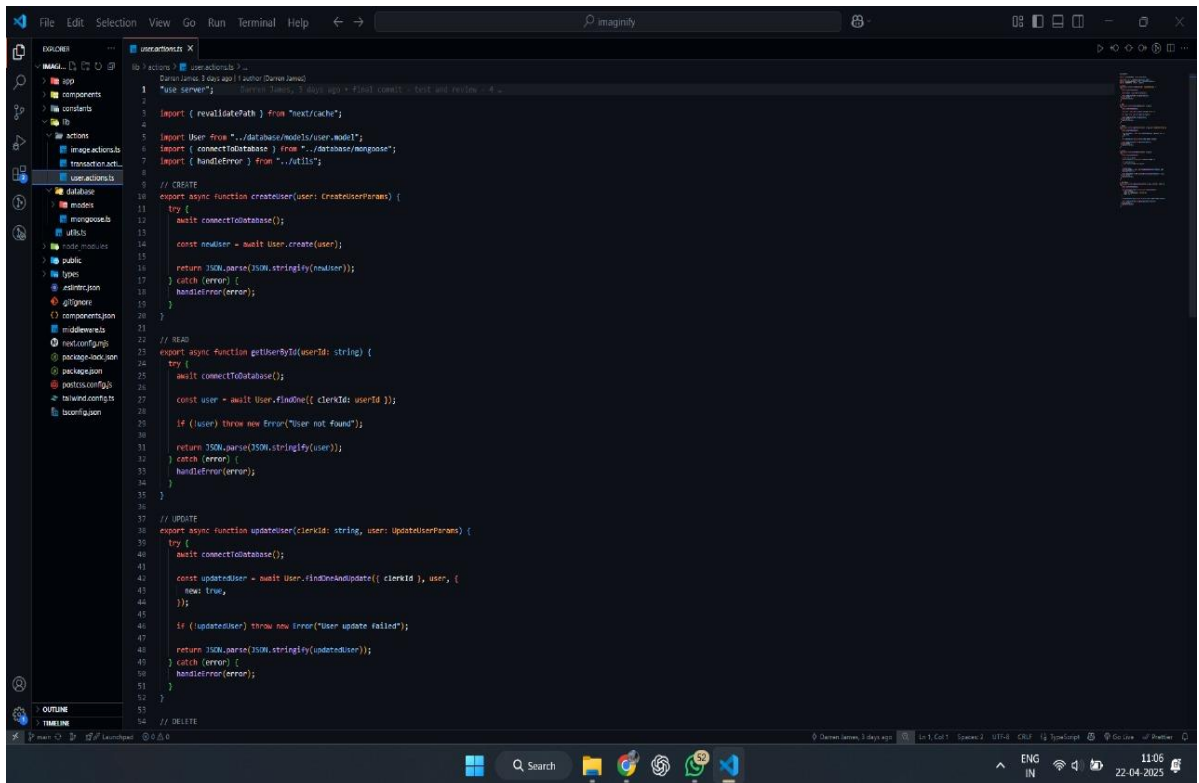
1. **Batch Image Processing**  
Allow users to upload and edit multiple images simultaneously, enabling faster workflows for photographers, marketers, and e-commerce businesses who deal with bulk image requirements.
2. **Video Editing Support**  
Extend AI functionalities such as background removal, object tracking, and color enhancement to videos—broadening the platform’s reach to content creators, influencers, and video editors.
3. **Mobile Application**  
Develop a mobile version of the platform for Android and iOS, allowing users to access image editing features on-the-go, with a focus on intuitive UI/UX tailored for smaller screens.
4. **Custom AI Training Interface**  
Allow premium users to upload their own datasets and fine-tune AI models (e.g., for consistent product backgrounds or brand-specific styles), providing personalized and adaptive outputs.
5. **User Collaboration Tools**  
Introduce real-time collaboration features where multiple users can co-edit images, comment, and share drafts—ideal for design teams and remote creative projects.
6. **Advanced AI Models**  
Integrate more sophisticated models such as Stable Diffusion for creative edits, StyleGAN for face enhancements, and OpenAI’s DALL·E for contextual content generation.
7. **AI-Based Image Suggestions**  
Provide smart suggestions like cropping, colour grading, or enhancement options based on image content and intended use (e.g., e-commerce, social media, portfolios).
8. **Marketplace for AI Templates**  
Launch a marketplace where users can share or purchase pre-configured AI editing presets, filters, or workflows, building a creator-driven ecosystem around the platform.
9. **Improved Accessibility & Localization**  
Add multi-language support, accessibility features for visually impaired users, and simplified UI modes for users with limited tech proficiency.

#### 10. Offline & Desktop Versions

Offer a downloadable desktop version with local AI model support for offline editing, especially beneficial for users with slow or intermittent internet connections.



## 4.6 CODE



The screenshot shows a code editor with a file explorer on the left and a code editor on the right. The file explorer shows a project structure with folders like 'app', 'components', 'constants', 'actions', 'interactions', 'transaction.actions', and 'user.actions'. The 'user.actions' folder is selected, and the file 'user.actions.js' is open in the editor. The code in 'user.actions.js' implements three functions: 'createUser', 'getUserById', and 'updateUser'. Each function uses 'mongoose' to interact with a database. The 'createUser' function takes a 'CreateUserParams' object and returns a 'JSON.parse(JSON.stringify(newUser))' object. The 'getUserById' function takes a 'userId' string and returns a 'JSON.parse(JSON.stringify(user))' object. The 'updateUser' function takes a 'userId' string and a 'UserUpdateParams' object and returns a 'JSON.parse(JSON.stringify(updatedUser))' object. The code is written in JavaScript and uses 'async/await' for database operations. The code is as follows:

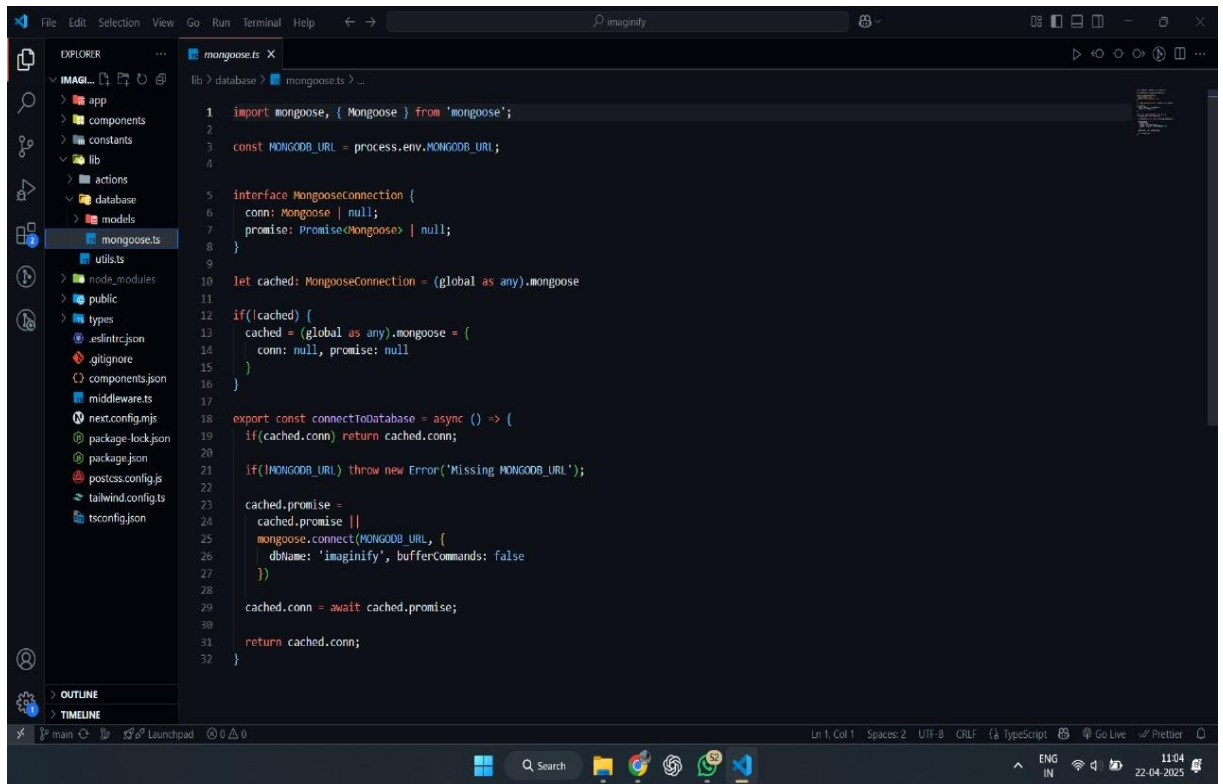
```
1 // CREATE
2 import { next } from 'next/cache';
3 import { connectToDatabase } from '../database/mongoose';
4 import { handleError } from '../utils';
5
6 // CREATE
7 export async function createUser({ createUserParams }) {
8   try {
9     await connectToDatabase();
10    const newUser = await User.create(newUserParams);
11    return JSON.parse(JSON.stringify(newUser));
12  } catch (error) {
13    handleError(error);
14  }
15 }
16
17 // READ
18 export async function getUserById(userId: string) {
19   try {
20     await connectToDatabase();
21     const user = await User.findOne({ clerkId: userId });
22     if (!user) throw new Error('User not found');
23     return JSON.parse(JSON.stringify(user));
24   } catch (error) {
25     handleError(error);
26   }
27 }
28
29 // UPDATE
30 export async function updateUser(userId: string, { updateUserParams }) {
31   try {
32     await connectToDatabase();
33     const updatedUser = await User.findOneAndUpdate({ clerkId: userId }, {
34       ...updateUserParams,
35     }, {
36       new: true,
37     });
38     if (!updatedUser) throw new Error('User update failed');
39     return JSON.parse(JSON.stringify(updatedUser));
40   } catch (error) {
41     handleError(error);
42   }
43 }
44
45 // DELETE
```

```
File Edit Selection View Go Run Terminal Help
imagind

transactionactions.js
1 // use server
2
3 import { redirect } from 'next/navigation'
4 import Stripe from 'stripe'
5 import { handleError } from '../utils'
6 import { connectToDatabase } from '../database/mongoose'
7 import Transaction from '../database/models/transaction.model'
8 import { updateCredits } from './user.actions'
9
10 export async function checkoutCredits(transaction: CheckoutTransactionParams) {
11   const stripe = new Stripe(process.env.STRIPE_SECRET_KEY!)
12
13   const amount = Number(transaction.amount) * 100;
14
15   const session = await stripe.checkout.sessions.create({
16     line_items: [
17       {
18         price_data: {
19           currency: 'usd',
20           unit_amount: amount,
21           product_data: {
22             name: transaction.plan,
23           },
24         },
25         quantity: 1
26       },
27     ],
28     metadata: {
29       plan: transaction.plan,
30       credits: transaction.credits,
31       buyerId: transaction.buyerId,
32     },
33     mode: 'payment',
34     success_url: `${process.env.NEXT_PUBLIC_SERVER_URL}/profile`,
35     cancel_url: `${process.env.NEXT_PUBLIC_SERVER_URL}/`,
36   })
37   redirect(session.url)
38 }
39
40 export async function createTransaction(transaction: CreateTransactionParams) {
41   try {
42     await connectToDatabase()
43
44     // Create a new transaction with a buyerId
45     const newTransaction = await Transaction.create({
46       ...transaction,
47       buyerId: transaction.buyerId
48     })
49
50     await updateCredits(transaction.buyerId, transaction.credits)
51
52     return JSON.parse(JSON.stringify(newTransaction))
53   } catch (error) {
54     handleError(error)
55   }
56 }
```

```
File Edit Selection View Go Run Terminal Help
imagind

imageactions.js
1 // use server
2
3 import { revalidatePath } from 'next/cache'
4 import { connectToDatabase } from '../database/mongoose'
5 import { handleError } from '../utils'
6 import User from '../database/models/user.model'
7 import Image from '../database/models/image.model'
8 import { redirect } from 'next/navigation'
9
10 import { v2 as cloudinary } from 'cloudinary'
11
12 const populateUser = (query: any) => query.populate({
13   path: 'author',
14   model: User,
15   select: '_id firstName lastName clerkId'
16 })
17
18 // use from
19 export async function addImage({ image, userId, path }: AddImageParams) {
20   try {
21     await connectToDatabase()
22
23     const author = await User.findById(userId)
24
25     if (!author) {
26       throw new Error('user not found')
27     }
28
29     const newImage = await Image.create({
30       ...image,
31       author: author._id,
32     })
33
34     revalidatePath(path)
35
36     return JSON.parse(JSON.stringify(newImage))
37   } catch (error) {
38     handleError(error)
39   }
40 }
41
42 // update image
43 export async function updateImage({ image, userId, path }: UpdateImageParams) {
44   try {
45     await connectToDatabase()
46
47     const imageToUpdate = await Image.findById(image_id)
48
49     if (!imageToUpdate || imageToUpdate.author.toString() !== userId) {
50       throw new Error('unauthorized or image not found')
51     }
52
53     const updatedImage = await Image.findByIdAndUpdate(
54       imageToUpdate._id,
55       image,
56       { new: true }
57     )
58
59     revalidatePath(path)
60
61     return JSON.parse(JSON.stringify(updatedImage))
62   } catch (error) {
63     handleError(error)
64   }
65 }
```



```

> > utils.ts > ...
9  export function cn(...inputs: ClassValue[]) {
10
11  }
12
13  // ERROR HANDLER
14  export const handleError = (error: unknown) => {
15    if (error instanceof Error) {
16      // This is a native JavaScript error (e.g., TypeError, RangeError)
17      console.error(error.message);
18      throw new Error(`Error: ${error.message}`);
19    } else if (typeof error === "string") {
20      // This is a string error message
21      console.error(error);
22      throw new Error(`Error: ${error}`);
23    } else {
24      // This is an unknown type of error
25      console.error(error);
26      throw new Error(`Unknown error: ${JSON.stringify(error)}`);
27    }
28  };
29
30  // PLACEHOLDER LOADER - while image is transforming
31  const shimmer = (w: number, h: number) => `
32  <svg width="${w}" height="${h}" version="1.1" xmlns="http://www.w3.org/2000/svg" xmlns:xlink="http://www.w3.org/1999/xlink">
33    <defs>
34      <linearGradient id="g">
35        <stop stop-color="#7986AC" offset="20%" />
36        <stop stop-color="#68769e" offset="50%" />
37        <stop stop-color="#7986AC" offset="70%" />
38      </linearGradient>
39    </defs>
40    <rect width="${w}" height="${h}" fill="#7986AC" />
41    <rect id="r" width="${w}" height="${h}" fill="url(#g)" />
42    <animate xlink:href="#r" attributeName="x" from="-${w}" to="${w}" dur="1s" repeatCount="indefinite" />
43  </svg>`;
44
45  const toBase64 = (str: string) =>
46    typeof window === "undefined"

```

## CHAPTER 5

### RESULTS

#### 4.7 Screens and Explanations

This chapter will include all the screens available in the project such as home page, sign up page, login page, along with saved prediction detailed explanation of each screen and its functionality. Screens available in the system are as follows

##### **Screen 1: Sign-Up Page**

The Sign-Up Page of the platform is the first step in onboarding new users to Imaginify, ensuring a secure and user-friendly registration experience. It is designed with a clean, minimalistic UI that aligns with modern SaaS aesthetics, built using Next.js, React, Tailwind CSS, and integrated with Clerk for authentication.

##### Key Features

- Form Fields:
  - Name – Full name input for personalization.
  - Email – User's email for account verification and login.
  - Password & Confirm Password – Ensures the user enters a secure password and confirms it to prevent typos.

The screenshot shows a web browser at the URL `imaginify-cloud.vercel.app/register`. The page features a central white card with the 'Imaginify' logo and the text 'Create an account'. The registration form includes fields for 'Name' (filled with 'john doe'), 'Email' (filled with 'john.doe@example.com'), 'Password' (masked with '\*\*\*\*\*'), and 'Confirm Password' (masked with '\*\*\*\*\*'). Each password field has an eye icon for toggling visibility. Below the form is a dark blue 'Register' button and a 'Continue with Google' button with the Google logo.

The screenshot shows a web browser at the URL `imaginify-cloud.vercel.app/login`. The page features a central white card with the 'Imaginify' logo and the text 'Welcome back'. The login form includes fields for 'Email' (filled with 'john.doe@example.com') and 'Password' (masked with '\*\*\*\*\*'), with an eye icon for the password field. Below the form is a 'Forgot password?' link, a dark blue 'Login' button, and a 'Continue with Google' button with the Google logo. At the bottom of the card is a link that says 'Don't have an account?'.

Fig. 5.1 Sign-Up/Login Page

## Screen 2: Features Screen

The **Object Removal** feature is a core AI-powered functionality of the Imaginify platform, allowing users to seamlessly erase unwanted elements from images with minimal effort. Whether it is a photobomber in the background, an irrelevant item, or text overlay, this tool uses advanced deep learning techniques to intelligently reconstruct the erased area, preserving the natural look of the surrounding environment.

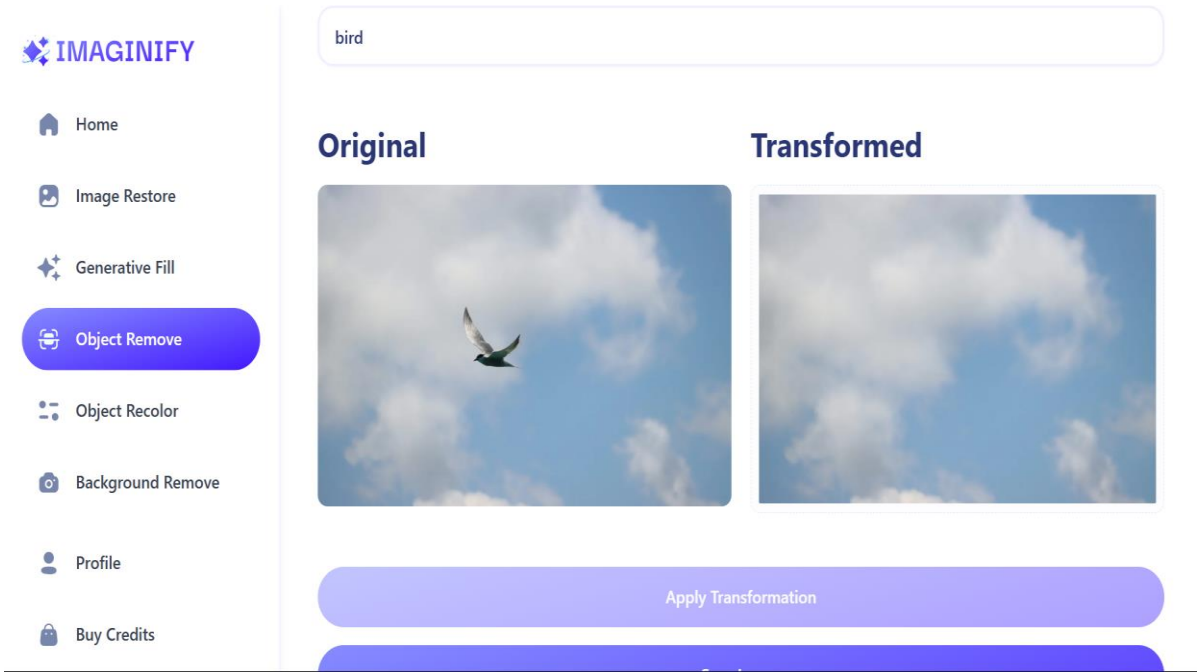


Fig. 5.2 Image Object Removal Screen

The **Object Recolor** feature empowers users to modify the color of specific elements within an image using AI-powered detection and transformation techniques. This feature is especially useful for designers, marketers, and e-commerce professionals looking to visualize products or elements in different colors without needing complex manual editing tools.

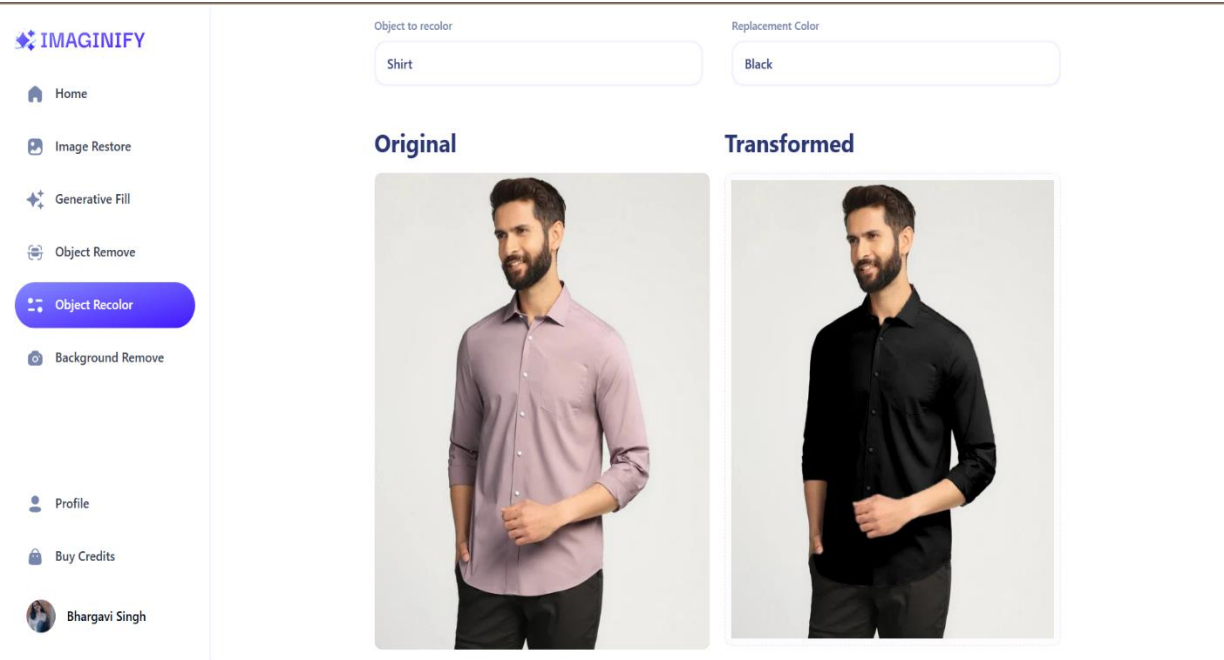


Fig. 5.2 Image Object Recolor Screen

The **Generative Fill** feature is an advanced AI-powered tool designed to intelligently extend or modify parts of an image while maintaining a natural and seamless appearance. This capability is particularly useful for tasks such as **aspect ratio adjustment**, **removing gaps after object removal**, or **filling in missing parts** of an image with realistic content.

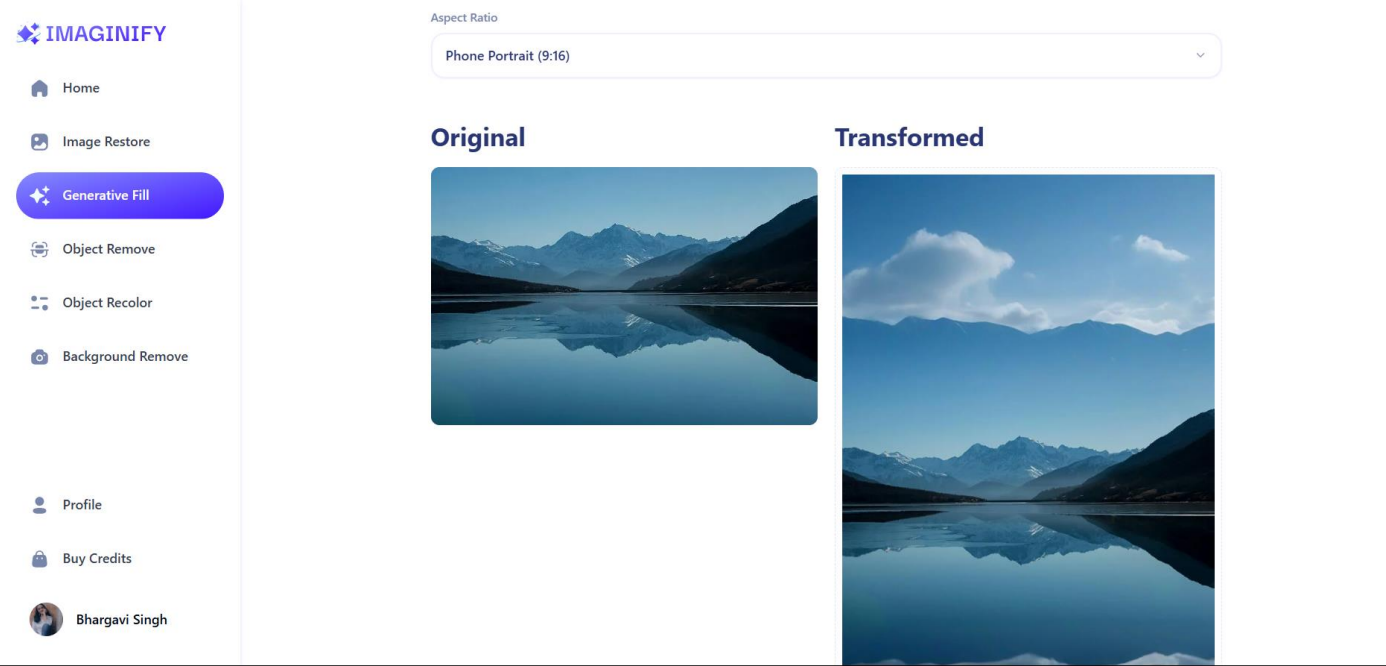


Fig. 5.2 Image Generative Fill Screen



## CHAPTER 6

### DISCUSSIONS

The AI-based SaaS image editing platform represents a significant step toward democratizing access to advanced image editing tools. Traditional image manipulation typically requires specialized knowledge, expensive software, and considerable manual effort. In contrast, this platform enables users to achieve professional-grade results effortlessly through intelligent automation.

The use of **Next.js** and **TypeScript** has allowed for the creation of a scalable, maintainable, and high-performance frontend, ensuring a fast and responsive user experience. **Cloudinary** plays a crucial role in handling AI-based image transformations, enabling powerful features such as image restoration, background removal, object recoloring, and generative fill without requiring heavy local processing. This cloud-centric approach improves scalability and makes the platform accessible from any device with an internet connection.

#### 6.1 Performance

The performance of the AI-based SaaS image editing platform is a critical factor in delivering a smooth, reliable, and responsive user experience. Various architectural and technological decisions were made to ensure optimal system efficiency and scalability. The following key performance aspects were observed:

##### Frontend Responsiveness-

Built using Next.js with TypeScript, the frontend loads quickly and performs well across modern browsers and devices. Server-side rendering (SSR) and static site generation (SSG) help in reducing load times and improving SEO. By using dynamic imports and lazy loading for non-critical components, the application ensures that initial page loads remain lightweight and fast. Image previews, transformation effects, and user dashboards are all optimized to ensure seamless interactions, even with high-resolution media.

##### Cloudinary-Powered Image Processing

Image transformation tasks such as restoration, object removal, and generative fill are offloaded to Cloudinary, which handles them using its optimized cloud infrastructure. This approach reduces client-side processing load and ensures that results are delivered back to users in real-time or within a few seconds, depending on image complexity. Additionally, Cloudinary's CDN capabilities ensure that transformed media is cached and delivered rapidly to global users.

## **Scalability**

The platform is designed to scale seamlessly using a serverless architecture and third-party cloud services. This means that traffic spikes or high numbers of concurrent image edits do not affect the core system performance. Load is distributed via CDN (Cloudinary) and edge functions (Vercel), allowing for efficient resource utilization. Auto-scaling mechanisms ensure that serverless functions can handle spikes in usage without manual intervention, and global edge deployments reduce latency for international users.

## **Authentication and Billing Efficiency**

User authentication via Clerk is fast and secure, with minimal latency during login and session validation. Clerk's use of JSON Web Tokens (JWT) ensures stateless session handling, which contributes to fast API responses. Subscription management using Stripe webhooks works in near real-time, updating user permissions automatically based on billing events such as upgrades, cancellations, or renewals.

## **Error Handling and Logging**

Robust error tracking and logging mechanisms are implemented to catch failures in API requests, image uploads, and payment processes. Services such as Sentry and LogRocket are integrated to provide real-time alerts and comprehensive debugging tools. These tools help developers trace issues quickly and implement fixes with minimal downtime, thereby improving platform reliability and user trust.

## **API Efficiency**

APIs are designed to handle requests asynchronously where possible, especially for operations like image uploads and transformations. This improves perceived performance and allows users to continue interacting with the app while edits are being processed. Rate limiting and retry mechanisms ensure system stability even under heavy load, while GraphQL-based querying may be considered in future iterations to optimize data fetching.

## **Testing and Optimization**

Extensive performance testing was done using real-world scenarios to identify bottlenecks. Frontend bundle sizes were reduced by removing redundant dependencies and using tree-shaking. Image files are compressed before upload to reduce bandwidth consumption. Browser caching and lazy hydration techniques are used to accelerate repeated interactions with the platform.

## **6.2. Future Research Directions**

As AI and web technologies continue to evolve rapidly, there are several promising directions for future research and development that can further enhance the platform's capabilities, performance, and impact:

### **Advanced Generative AI for Creative Editing**

Future research can explore the integration of more sophisticated generative models (e.g., GANs, diffusion models) to enable highly realistic and customizable edits such as scene expansion, 3D-to-2D rendering, and context-aware inpainting. These techniques can allow users to modify or generate entirely new content that blends seamlessly with original images, opening possibilities for product design, advertising, and digital storytelling.

### **Personalized AI Editing**

Building adaptive AI models that learn user preferences over time can lead to personalized editing suggestions, automated styling, and one-click enhancements based on editing history or image context. For instance, users who frequently brighten their images may be presented with pre-adjusted previews. Leveraging machine learning, the platform can provide a tailored experience that adapts to user habits.

### **Real-Time Edge Processing**

Research into edge AI computing may enable faster, real-time image processing directly in the browser or on edge networks. Technologies like Web Assembly and TensorFlow.js can be explored to run models locally within the user's browser. This could significantly reduce latency and dependence on cloud servers, improving performance on mobile and low-bandwidth environments and enabling privacy-sensitive applications.

### **Integration of Multimodal Inputs**

Future versions of the platform could allow users to provide voice commands, sketches, or text prompts to guide image edits, combining natural language processing (NLP) and computer vision for a more intuitive editing experience. For example, a user could describe changes like “make the sky more vibrant” or “replace the background with a sunset,” and the AI would interpret and execute the command accordingly.

### **Ethical and Responsible AI**

As AI-generated content becomes more realistic, it's essential to invest in research around watermarking, content authenticity verification, and bias mitigation to ensure ethical use of the platform. Tools like image provenance tracking, tamper detection, and usage disclaimers could be integrated to promote transparency and avoid misuse of AI-generated content.

### **AI-Powered Video Editing**

Expanding research into AI-based video editing—including object removal, background editing, and frame-by-frame enhancements—would broaden the platform's appeal and utility for content creators. Real-time video filters, AI-based trimming, scene transitions, and even generative frame interpolation can transform the platform into a comprehensive multimedia editing solution.

### **Collaboration & Feedback Loops**

Researching real-time collaboration features, version control, and feedback-based iterative improvements can make the platform more suitable for teams and design studios. Users could comment, suggest edits, and revert to previous versions of their work, much like tools used in code collaboration (e.g., GitHub). This would significantly enhance the platform's value for professional environments.

### **Low-Resource AI Models**

To optimize cost and accessibility, future efforts could explore lightweight AI models that perform well on-device or with limited computational resources. These models can be deployed to support offline functionality or lower-tier subscription plans. Research into model quantization, pruning, and distillation techniques can contribute to this goal, ensuring the platform remains inclusive for users with lower-spec devices or limited internet access.

### **Integration with the Creator Ecosystem**

Future enhancements may include seamless integration with tools used by creators and professionals, such as Adobe Creative Cloud, Figma, Canva, and productivity suites like Notion or Google Workspace. APIs could allow content from these platforms to be imported, edited, and exported without breaking the user's workflow, increasing adoption in enterprise and educational sectors.

### **Augmented Reality (AR) and Virtual Reality (VR) Compatibility**

With the rise of immersive technologies, future research may explore AR/VR-compatible image editing interfaces. Users could manipulate images in a 3D space using gestures or AR overlays, offering a highly interactive and visual editing experience suitable for creative professionals in fields such as architecture, game design, and education.

## CHAPTER 7

### CONCLUSION

The evolution of artificial intelligence has profoundly impacted digital imaging, transitioning from traditional manual editing techniques to sophisticated, automated AI-driven processes. This transformation has not only enhanced the quality and efficiency of image manipulation but has also made advanced tools accessible to non-experts across various industries. The AI-based SaaS platform discussed in this report represents a critical juncture in this transformation—merging the capabilities of powerful AI models with the scalability and accessibility of cloud-based web applications.

At the heart of this platform is the intelligent orchestration of cutting-edge AI models for various image editing tasks: image restoration, generative fill for aspect ratio adjustments, object removal, object recoloring, and background removal. Each of these features addresses common, high-demand use cases that span across creative industries, digital marketing, e-commerce, social media, and personal use. By automating these previously manual tasks, the platform enhances productivity and expands creative possibilities for users with minimal technical expertise.

These AI-driven features are built on the foundation of deep learning, leveraging convolutional neural networks (CNNs), GANs (Generative Adversarial Networks), and transformer-based vision models. For example, image restoration uses deep convolutional networks trained on degraded and high-quality image pairs to learn noise reduction, deblurring, and artifact removal. Generative fill applies contextual learning, completing missing image areas while maintaining consistent textures and structures. These models are continually improving with advances in computer vision research, enabling more refined and realistic results over time.

From a technological standpoint, the platform’s architecture is built on a foundation of modern and efficient tools. **Next.js**, paired with **TypeScript**, provides a performant and developer-friendly environment, enabling the creation of fast, scalable, and maintainable applications. Next.js supports server-side rendering (SSR), static site generation (SSG), and incremental static regeneration (ISR), all of which contribute to a dynamic and fast user experience. TypeScript enforces strong typing, minimizing runtime errors and improving code reliability—especially vital in large-scale applications.

**Stripe** offers a secure and flexible payment solution that supports various business models—from subscriptions to usage-based billing—thus accommodating a wide range of customers. Its support for global currencies, webhooks, and invoicing tools streamlines the monetization strategy and ensures compliance with international financial standards such as PCI-DSS and GDPR. Stripe’s modular billing infrastructure enables fine-grained control over pricing, metering, and customer segmentation.

**Clerk** simplifies authentication and user management while ensuring data privacy and compliance with modern security standards. It provides ready-to-use components for login, registration, session handling, and account management, including support for social login, two-factor authentication, and role-based access control (RBAC). These features minimize development time while maximizing security and user convenience.

Most importantly, **Cloudinary** provides robust and scalable media processing powered by AI, which allows the platform to offload heavy image transformation tasks to the cloud. Cloudinary’s image and video APIs include built-in AI models for object detection, background removal, image enhancement, and automatic cropping. Offloading computationally intensive operations to Cloudinary ensures that image processing is not dependent on the user’s device, allowing seamless access from low-end hardware.

These technologies work in tandem to form a seamless, integrated system that is both powerful and user-friendly. The decision to utilize pre-trained models and third-party APIs like Cloudinary’s AI engine ensures that the platform can be developed and scaled quickly without the overhead of training and hosting custom models from scratch. This strategic use of existing infrastructure lowers development costs and minimizes time-to-market—two crucial factors in the competitive SaaS space.

What differentiates this platform from single-purpose tools or heavyweight design suites is its **modularity and focus on usability**. While tools like Adobe Photoshop offer a comprehensive set of image editing features, they come with steep learning curves and are often overkill for users who only need a few specific capabilities. On the other hand, niche AI tools like Remove.bg or Cleanup.pictures focus on one task, limiting their overall value. By bundling multiple high-demand features into a single platform and presenting them through an intuitive user interface, this platform aims to strike a balance between simplicity and versatility—offering both convenience and creative power.

Another critical strength lies in its **SaaS delivery model**. By deploying the platform entirely online, users are not required to download or install software. This enhances accessibility, particularly for users on low-spec machines or mobile devices. The SaaS model also enables continuous updates, feature rollouts, and performance improvements without requiring any action from the user. It provides a foundation for agile product development, rapid feedback implementation, and responsive customer support.

To further improve performance, the platform implements **serverless architecture** using edge functions, which allocate compute resources dynamically based on traffic. It also integrates content delivery networks (CDNs) to cache and distribute images globally with low latency. This architectural setup not only reduces server costs but also ensures availability and reliability during traffic surges or geographic scaling.

Despite its strengths, the platform faces several **challenges and limitations**. Inference latency, particularly with large or high-resolution images, can affect the real-time usability of certain features. While Cloudinary’s infrastructure mitigates this to some extent, performance optimization remains a key area of future work. Potential improvements include asynchronous editing queues, progress feedback mechanisms, and batch processing support for commercial users.

Similarly, model reliability and bias are non-trivial issues in AI applications. AI-generated outputs can sometimes be unpredictable, especially with diverse image inputs or underrepresented contexts. A generative fill might produce artifacts, or a recoloring tool might misinterpret object boundaries. Building in fallback mechanisms, manual overrides, user previews, and feedback buttons can help manage user expectations and improve trust. Future development may also include fine-tuned models trained on diverse datasets to reduce algorithmic bias.

**Cost** is another consideration. High-quality AI processing, especially if hosted on GPU-backed infrastructure, can be expensive. Balancing user needs with operational costs will require a combination of smart pricing models and usage tracking. This is where Stripe’s flexible billing capabilities will play a vital role—enabling usage-based, tiered, or freemium plans to accommodate different customer segments without compromising on sustainability. Additionally, features like usage dashboards and quotas can help users self-manage their consumption and reduce waste.

**User experience (UX)** is central to the success of any SaaS platform. While AI enables automation, not all users will be satisfied with “one-click” edits. Providing fine-tuning options, adjustable parameters, and real-time feedback can greatly enhance the perceived control and satisfaction of users. Future updates may include interactive sliders for brush sizes

in object removal, adjustable color palettes for recoloring, or aspect ratio presets in generative fill. Offering before-and-after comparisons, undo history, and customization presets can further empower users.

Looking forward, the platform is well-positioned to evolve alongside rapidly advancing AI technologies. The image editing capabilities can be expanded to include **multi-modal input**, such as text-to-image editing (e.g., “replace the sky with sunset”), or **video frame editing**. Such advancements will require integration with large multimodal models or video-based AI models, which are becoming more accessible through APIs provided by companies like OpenAI, RunwayML, and Pika.

Integration with design tools like **Figma**, **Canva**, or **WordPress plugins** could broaden the user base by allowing creative professionals to incorporate AI edits directly into their existing workflows. These integrations would enable “edit-in-place” workflows, boosting convenience and reducing context switching.

Furthermore, the platform has the potential to introduce **collaborative features**. Real-time collaboration, version history, shared projects, and team workspaces are essential functionalities in many SaaS tools today. Introducing such features would elevate the platform from a utility to a productivity suite, increasing retention and enterprise appeal. Admin roles, team billing, and content review workflows could be added to attract business users and agencies.

Another strategic direction involves the incorporation of **AI feedback loops**. By collecting anonymized data (with user consent) on how users interact with the platform—such as which edits are most used, which results are accepted or rejected—the system can improve through reinforcement learning or adaptive UI recommendations. This data-driven approach can also inform marketing, customer support, and feature prioritization. Heatmaps, edit frequency analytics, and A/B testing interfaces can help shape future product decisions.

**Security, privacy, and ethical AI usage** will continue to be non-negotiable pillars as the platform scales. Ensuring **GDPR compliance**, implementing watermarking or usage policies to prevent misuse, and maintaining transparency in how AI models work is critical for maintaining trust. User-uploaded content must be stored and processed securely, with clear opt-in policies for data use. The platform must also educate users about the limitations of AI-generated content to avoid overreliance or misinformation.



Future roadmap items may also include accessibility enhancements, such as voice-guided editing, screen reader compatibility, and AI-generated alt text. These features would help the platform serve a broader audience and comply with standards like WCAG.

In summary, the AI-based SaaS platform described in this report is a forward-looking solution that brings the power of artificial intelligence to a wide audience through the convenience of a web application. It successfully combines technical excellence with user-centric design and is built on a scalable architecture that allows for continuous evolution. By offering a diverse range of features through a single interface, it meets real user needs while providing a foundation for long-term innovation.

The convergence of AI, cloud computing, and SaaS delivery models represents one of the most transformative trends in software development today. Platforms like this not only reflect this convergence but also push the boundaries of what's possible with creative automation. As the platform continues to grow—incorporating user feedback, emerging AI models, and new business opportunities—it stands to become a central tool in the future of image editing, where creativity is no longer constrained by technical skill, and everyone has the power to create, modify, and share with ease

.

## CHAPTER 8

### REFERENCE

1. Next.js Documentation

Vercel. (n.d.). Next.js Documentation. Retrieved from <https://nextjs.org/docs>

2. TypeScript Documentation

Microsoft. (n.d.). TypeScript Handbook. Retrieved from <https://www.typescriptlang.org/docs/>

Stripe Documentation

- a. Stripe. (n.d.). Stripe Docs – Payments, Subscriptions & Billing. Retrieved from <https://stripe.com/docs>

3. Clerk Authentication Documentation

Clerk.dev. (n.d.). Clerk Docs – Authentication and User Management. Retrieved from <https://clerk.dev/docs>

4. Cloudinary Documentation

Cloudinary(n.d.). Cloudinary Docs – Media Optimization and AI Features. Retrieved from <https://cloudinary.com/documentation>

5. Generative Fill & Inpainting with AI

NVIDIA. (2018). Image Inpainting for Irregular Holes Using Partial Convolutions. Retrieved from <https://arxiv.org/abs/1804.07723>

6. AI-Based Object Detection and Removal

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. Retrieved from <https://arxiv.org/abs/1703.06870>

7. Background Removal using Deep Learning

- Xiong, B. et al. (2019). A Background Matting Approach Using Deep Learning. Retrieved from <https://arxiv.org/abs/1905.00737>
- Ledig, C. et al. (2017). Photo-Realistic Single Image
8. Super-Resolution Using a Generative Adversarial Network. Retrieved from <https://arxiv.org/abs/1609.04802>
  9. React Documentation  
Meta. (n.d.). React – A JavaScript library for building user interfaces. Retrieved from <https://reactjs.org/docs/getting-started.html>.
  10. Building a SaaS Product with Next.js and Stripe  
Lee, J. (2023). How to Build a SaaS App with Next.js, Stripe, and Clerk. Retrieved from <https://leejjon.medium.com> (or similar blog/tutorials depending on your source)
  11. Understanding SaaS Architecture  
Amazon Web Services. (n.d.). SaaS Architecture Center. Retrieved from <https://aws.amazon.com/saas/>
  12. Modern Web Application Security Practices  
OWASP Foundation. (n.d.). OWASP Top 10 Web Application Security Risks. Retrieved from <https://owasp.org/www-project-top-ten/>
  13. Image Recoloring with AI  
Zhang, R., Isola, P., & Efros, A. A. (2016). Colorful Image Colorization. Retrieved from <https://arxiv.org/abs/1603.08511>
  14. Responsible AI and Ethical Considerations Microsoft. (n.d.). Responsible AI Principles.  
Retrieved from <https://www.microsoft.com/en-us/ai/responsible-ai>
  15. Cloudinary AI Features Overview Cloudinary. (n.d.). AI-Based Media Enhancements with

Cloudinary. Retrieved from <https://cloudinary.com/blog>

16. Generative AI in Design Tools

Adobe. (2023). Introducing Generative Fill in Photoshop. Retrieved from <https://www.adobe.com/sensei/generative-ai.html>

17. OpenAI APIs for Image Processing

OpenAI. (n.d.). DALL·E and Inpainting APIs. Retrieved from <https://platform.openai.com/docs/guides/images>

18. Frontend Optimization with Next.js

Cassidy, C. (2022). Performance Optimization in Next.js Apps. Retrieved from <https://vercel.com/blog>

19. Using AI for Real-Time Editing

Huang, X. et al. (2020). Real-Time User-Guided Image Editing with GANs.

## CHAPTER 9

### BIBLIOGRAPHY

For the AI-based SaaS platform designed to provide intelligent image editing capabilities—such as image restoration, object removal, background replacement, aspect ratio enhancement through generative fill, and object recoloring—the references have been carefully selected to reflect cutting-edge advancements in relevant fields.:

1. AWS. (n.d.). *SaaS architecture best practices*. Amazon Web Services. Retrieved from <https://aws.amazon.com/partners/saas-factory/architecture/>
2. Clerk. (n.d.). *Secure authentication for modern web apps*. Retrieved from <https://clerk.com/docs/security>
3. Cloudinary. (n.d.). *AI-powered image & video APIs*. Retrieved from <https://cloudinary.com/documentation>
4. Cloudinary. (n.d.). *Use cases: AI-based media optimization*. Retrieved from <https://cloudinary.com/customers>
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative adversarial nets*. In Advances in Neural Information Processing Systems (NeurIPS).
6. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). *Mask R-CNN*. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2961–2969.
7. Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). *Image-to-image translation with conditional adversarial networks*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1125–1134.

8. Microsoft. (n.d.). *TypeScript documentation*. Retrieved from <https://www.typescriptlang.org/docs/>
9. Next.js. (n.d.). *Next.js documentation*. Vercel. Retrieved from <https://nextjs.org/docs>
10. Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional networks for biomedical image segmentation*. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234–241.
11. Sharma, G. (2019). *Building microservices with Node.js: A practical guide to implementing microservices in Node.js*. Packt Publishing.
12. Stripe. (n.d.). *SaaS billing with Stripe*. Retrieved from <https://stripe.com/solutions/saas>
13. Stripe. (n.d.). *Stripe billing documentation*. Retrieved from <https://stripe.com/docs/billing>
14. Xu, N., Price, B., Cohen, S., Yang, J., & Huang, T. S. (2017). *Deep image matting*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2970–2979.
15. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., & Huang, T. S. (2018). *Generative image inpainting with contextual attention*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5505–5514.
16. Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). *Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising*. IEEE Transactions on Image Processing, 26(7), 3142–3155.
17. Microsoft. (n.d.). *TypeScript Handbook*. Retrieved from <https://www.typescriptlang.org/docs/>
18. Stripe. (n.d.). *Stripe Docs – Payments, Subscriptions & Billing*. Retrieved from

<https://stripe.com/docs>

19. Clerk.dev. (n.d.). *Clerk Docs – Authentication and User Management*. Retrieved from <https://clerk.dev/docs>
20. Cloudinary. (n.d.). Cloudinary Docs – Media Optimization and AI Features. Retrieved from <https://cloudinary.com/documentation>
21. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. Retrieved from <https://arxiv.org/abs/1703.06870>
22. NVIDIA. (2018). Image Inpainting for Irregular Holes Using Partial Convolutions. Retrieved from <https://arxiv.org/abs/1804.07723>
23. Ledig, C. et al. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. Retrieved from <https://arxiv.org/abs/1609.04802>
24. Cloudinary. (n.d.). Cloudinary Docs – Media Optimization and AI Features. Retrieved from <https://cloudinary.com/documentation>
24. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. Retrieved from <https://arxiv.org/abs/1703.06870>

## Curriculum Vitae

### DARREN JAMES

#### DETAILS

##### ADDRESS

Nagpur  
India

##### PHONE

8668505561

##### EMAIL

darrenmanueljames4080@gmail.com

#### LINKS

[LinkedIn](#)

[Hackerrank](#)

[Leetcode](#)

#### SKILLS

C

C++

HTML & CSS

JavaScript

Node.js

Git

MongoDB

Java

#### PROFILE

- Innovative problem solver.
- Cyber security enthusiast.
- Adept at financial analysis and a keen observer of the stock market, leveraging insights to make informed investment decisions.

#### EDUCATION

##### MCA, KIET Group Of Institutions

Ghaziabad

2023 — 2025

##### Bsc in Information Technology, Rashtrasant Tukadoji Maharaj Nagpur University

Nagpur

2020 — 2023

- Secured an outstanding academic record, achieving 83.48%.
- Demonstrated a strong commitment to academic excellence throughout the course of my undergraduate studies.

##### 12th, Maharashtra State Board

Nagpur

2020

##### 10th, St. Francis De Sales High School

Nagpur

2018

#### ACHIEVEMENT

##### University Topper, Nagpur University, 2023

#### PROJECTS

##### Image Generation Web-App

**Description:** Generates random images based on the query given by the user, uses *Unsplash* API for image retrieval. Also stores the previously searched images in the "Community showcase section" of the WebApp

**Live Project Link:** <https://image-gen-frontend-blond.vercel.app/>

**Github Link:** <https://github.com/darrenjames8793/image-Gen>

##### News Web-App

**Description:** A News Web App, which simply fetches the news from the news-api and presents it to the user. The user can categorize the news based on category and location(country)

**Live Project Link:** <https://news-web-app-api.vercel.app/>

**Github Link:** <https://github.com/darrenjames8793/News-Web-App>



# Anshoo Yadav

✉ anshoo.2325mca1088@kiet.edu

📍 Ghaziabad, India

🌐 github.com/Anuydv07

📞 7060392514

🌐 linkedin.com/in/anshoo-yadav-47a856205

## EDUCATION

— **Masters of Computer Application,M.C.A**  
Dr.APJ Abdul Kalam Technical  
University,Lucknow

*Pursuing*

— **Bachelors of Computer  
Applications,B.C.A**  
Chaudhary Charan Singh University,Meerut

*07/2020 - 07/2023*

— **Intermediate,12th**  
T.R.M.Public School,Modinagar

*04/2019 - 04/2020*

— **High School,10th**  
T.R.M.Public School,Modinagar

*04/2017 - 04/2018*

## CERTIFICATES

Learning Java 11 At LinkedIn Learning (08/2023)

Advance Java Programming At LinkedIn Learning  
(08/2023)

## SKILLS

Java Data Structure And Algorithms C  
HTML CSS JavaScript Servlet SQL

## CO-CURRICULARS

Attended Virtual Global Conclave On Theme  
'Traversing the VUCA Landscape With Soft Skills'  
(12/2023)


Attended Guest Lecture on 'Discovering New-Age  
AI Tehnologies' (11/2023)

## ACHIEVEMENTS

Secured 1st Position At Ideathon 2024 At KIET  
Group Of Institutions

## BHARGAVI

bhargavi02.work@gmail.com@gmail.com | (+91) 6369808825

 /bhargavi-singh

### EDUCATION

Master of Computer Application   KIET Group of Institutions, Ghaziabad   SGPA: 7.87 Affiliated from AKTU, Lucknow	Batch (2023-25)
Bachelor's of Computer Application   ITS, Mohan Nagar   Percentage: 73% Affiliated from CCSU, Meerut	Batch (2019-22)
Senior Secondary   Kendriya Vidyalaya No.02, Chennai   Percentage: 73% CBSE Board	Session(2018-19)
Secondary   Kendriya Vidyalaya No.02, Chennai   CGPA: 9.0 CBSE Board	Session(2016-17)

### SKILLS

❖ C | C++ | Python | MySQL | Java | MS-Excel | HTML/CSS |

### PROJECTS

- ❖ House Price Prediction (Oct'24-Nov'24)
  - The House Price Prediction Model is a powerful machine learning solution.
  - It is designed to estimate the market value of residential properties based on different factors.
  - Tools used are: Python, pandas and scikit-learn
- ❖ Task Tracker (April'24-June'24)
  - This application is a To-Do List for the users to track their tasks.
  - This page will have different functionality for the users like: Add, Delete, and completed task.
  - Tools used are: HTML, CSS, JavaScript.

### CERTIFICATION

- ❖ Networking Basics: Cisco
- ❖ Java 8 Essential training: LinkedIn
- ❖ HTML, CSS and JavaScript, building the web: LinkedIn
- ❖ Learning SQL Programming: LinkedIn
- ❖ Python Quick Star: LinkedIn

### ACHIEVEMENTS / HOBBIES

- ❖ Secured 1st position in Emerging Technology in Ideathon '23 at KIET Group of Institution.
- ❖ Earned a skill badge for OOP on LinkedIn.
- ❖ Secured 1st position in Hackathon & Logo design competition.
- ❖ Participated in the District Badminton Championship.

### POSITION OF RESPONSIBILITY

- ❖ **Head Girl | BCA** (2021-22)  
Managed and led a team of 10 prefects to organize college events, ensuring smooth execution and engagement.  
Acted as a mentor to junior students, promoting a positive college culture and teamwork
- ❖ **Placement Coordinator | BCA** (2021-22)  
Maintained and analyzed placement-related data to track progress and report results to the administration.

# DIVYANSH SHARMA

## FULL-STACK WEB DEVELOPER

I am a highly motivated MCA graduate with a strong foundation in software development, problem-solving, and database management. Passionate about coding and technology, I specialize in full-stack development, API integrations, and data-driven applications.

### CONTACT

- +91 9773611053
- sdivyansh462@gmail.com
- Ghaziabad, India
- www.linkedin.com/in/divyansh-sharma-6456b81b1
- github.com/Divyansh053

### SKILLS

- JAVA
- DBMS
- C/C++
- Express JS
- HTML/CSS
- ReactJS
- MongoDB
- SDLC
- NodeJS
- JavaScript

### CERTIFICATES

- JavaScript, LinkedIn Learning
- HTML Essentials, LinkedIn Learning
- Tailwind CSS 3, LinkedIn Learning
- Java 17, LinkedIn Learning
- Project Management, Coursera

### CO-CURRICULAR

- Attended Webinar on Management and Leadership skills.
- Co-ordinated Frescos 2025 at KIET Group of Institutions.
- Volunteered as Management staff in VOLFEST 2024-2025.

### HOBBIES

- Watching tech review videos
- Travel
- Gaming
- Learning new infra projects

### EDUCATION

**MCA** 2023 - PRESENT  
KIET Group of Institutions, Ghaziabad Affiliated to AKTU, Lucknow

- 7.53 SGPA till 3<sup>rd</sup> Sem.

**BCA** 2019-2022  
Institute of Management & Research, Ghaziabad Affiliated to CCS University, Meerut

- 62.97%

**INTERMEDIATE** 2019  
Green Field Public School Ghaziabad, CBSE Board, New Delhi

- 64%

**HIGH SCHOOL** 2017  
Green Field Public School Ghaziabad, CBSE Board, New Delhi

- 68.4%

### PROJECTS

**News Aggregator Website** 2024

- A API-based News Web Application built on the MERN stack dynamically fetches and displays the latest news from external news APIs. It offers a seamless user experience with categorized news, search functionality, and real-time updates.
- MERN Stack
- Individual

**House Price Predictor** 2024

- The House Price Predictor is a Python-based machine learning model to predict house prices based on parameters like area, BHK, and location. It features user authentication, price prediction & the ability to save searches in user profiles for personalized insights.
- Front-End- HTML/CSS, Bootstrap
- Backend- Express JS, MongoDB, NodeJS
- Machine Learning- Lasso Model with 82% accuracy using Python
- College Group Project- My role was to work on datasets and test and train the algorithm with supporting documents.

### EXTRA-CURRICULAR

- Former Member of Robin Hood Army(NGO), Ghaziabad (2023).
- Supported the needy with food resources and community support through work with RHA.
- Active Member of HOV(NGO), Ghaziabad (Since 2024).
- Participated in Smart India Hackathon in KIET Group of Institutions.

## **MENTOR REMARKS**

Remark: Report is fine.  
Project is complete.  
Good team Coordination

Rating: 10

(Mentor's Signature):