## **Project Title: Provision AI**

**Team Name: SVAG**

**Team Members:**

* P. Vamshi Krishna
* P. Swetha sri
* K. Alekhya
* G. Gayathri

**Phase-1: Brainstorming & Ideation**

**Objective:**

ProVisionAI aims to leverage Gemini Vision for high-precision automated image annotation, ensuring seamless integration with external APIs and cloud storage. The platform provides an intuitive user interface for manual refinements while maintaining security, compliance, and scalability for diverse industries.

**Key Points:**

1. **Problem Statement:**

ProVision AI is an innovative project aimed at revolutionizing image annotation through advanced AI capabilities, specifically leveraging Google's powerful Gemini Vision Pro model. This project provides users with a sophisticated platform for annotating images with rich descriptive captions and insightful information.

1. **Proposed Solution:**

ProVision AI enhances the image annotation process by leveraging the cutting-edge capabilities of the Gemini Vision Pro model. Users can upload images and receive detailed annotations, including descriptive captions and interesting facts about the content within the images.

1. **Target Users:**

**Content Creators:**

**Need**: Auto-generate engaging captions.

**Enhancement**: Include performance-based caption suggestions

**Marketing & Advertising Professionals:**

**Need**: Captions for ads and campaigns.

**Enhancement**: Multilingual support for global campaigns.

**Educators & Researchers:**

**Need**: Annotations for educational content.

**Enhancement**: Collaborative annotation tools.

**Healthcare & Medical Professionals:**

**Need**: Annotate medical images.

**Enhancement**: Integration with medical databases.

**Non-profits & NGOs:**

**Need**: Impactful captions for campaigns.

**Enhancement**: Emotionally resonant messaging for advocacy.

**Creative Agencies & Designers:**

**Need**: Descriptions for portfolios and projects.

**Enhancement**: Customizable style and tone.

**4. Expected Outcome:**

**1. Highly Accurate Image Annotations:**

Users will receive automated, highly accurate descriptions of the content within images, such as object identification, scene analysis, and activity recognition.

**2**. **Contextual Insights and Interesting Facts:**

Beyond basic annotations, **ProVisionAI** will provide insightful information, such as historical, geographical, or cultural context about the image’s content. For example, an image of a landmark could be annotated with interesting facts about its history.

This feature will offer added value by not only describing the image but also enriching the user experience with knowledge and engaging information.

**3. Integration with Various Industries:**

ProVisionAI can benefit a wide range of sectors, such as healthcare, e-commerce, education, and media. For example, in healthcare, medical images can be annotated with relevant diagnostic information or treatment suggestions

**4. Seamless User Experience:**

The platform will be intuitive, allowing users to simply upload images and receive automatic, accurate annotations without requiring manual input.

**Phase-2: Requirement Analysis**

**Objective:**

Leverage Gemini Vision to perform automated image annotation with high precision

**Key Points:**

**2.1 Image Upload and Processing**

* Users can upload images from local storage, cloud, or via API.
* Images should be pre-processed to enhance clarity and segmentation.
* Gemini Vision will analyze and annotate objects, labels, and regions of interest.

**2.2 Annotation Capabilities**

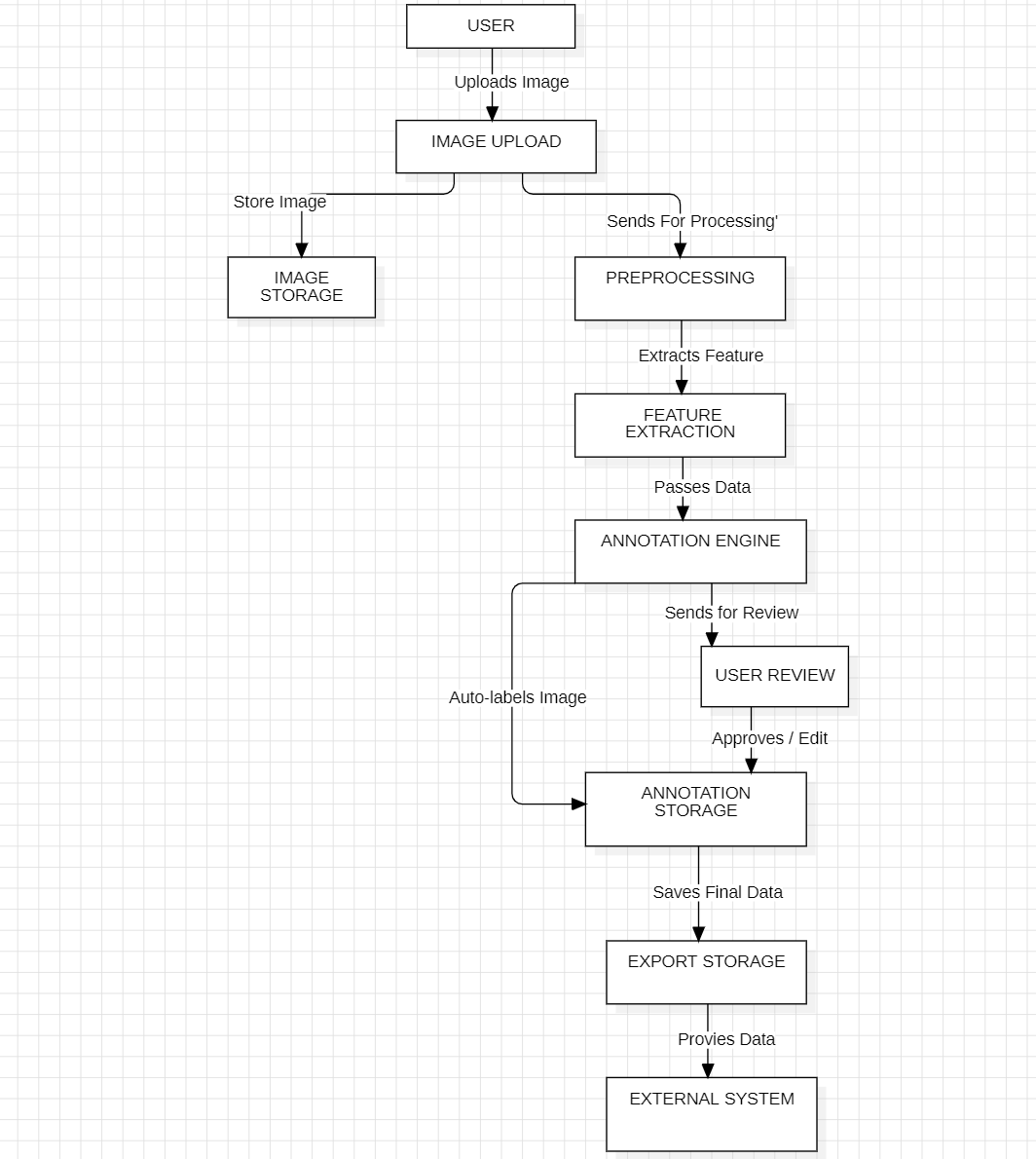
* Automatic annotation using Gemini Vision’s AI-based detection.
* Manual annotation tools (bounding boxes, polygons, semantic segmentation, tagging).
* Multi-layer annotation for hierarchical object classification.

**2.3 User Interface**

* Web-based and mobile-friendly UI with drag-and-drop functionality.
* Annotation visualization with zoom and pan capabilities.
* Collaboration tools for teams, including comments and revision tracking.

**Phase – 3: Project Design**

**Objective :**



**KEYPOINTS:**

**1.System Architecture:**

* **AI Core Engine:** Leverages **Gemini Vision** for image analysis and annotation.
* **Modular Microservices:** Ensures scalability, flexibility, and efficient data processing.
* **Cloud & On-Prem Support:** Deployable on cloud platforms or local infrastructure.
* **Secure Data Management:** Implements encryption and access control for sensitive data.

**2. User Flow:**

* **Image Upload & Preprocessing:** Users upload images/videos for annotation.
* **AI-Powered Annotation:** **Gemini Vision** auto-detects obje’vcts and applies labels.
* **Manual Review & Editing:** Users can refine annotations with intuitive tools.
* **Export & Integration:** Annotated data is exported in various formats for downstream applications.

**3. UI/UX Considerations:**

* **Minimalist & Intuitive Design:** Ensures ease of use with a clean and responsive interface.
* **Drag-and-Drop Functionality:** Simplifies image uploads and annotation edits.
* **Real-Time Feedback & Suggestions:** Enhances efficiency with AI-driven recommendations.
* **Customizable Annotation Tools:** Offers flexible settings for different use cases.

## **Phase-4: Project Planning (Agile Methodologies)**

### **Objective:**

Break down development tasks for efficient completion.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Task** | **Priority** | **Duration** | **Deadline** | **Assigned To** | **Dependencies** | **Expected Outcome** |
| Sprint 1 | Environment Setup & API Integration | 🔴 High | 2.5 hours (Day 1) | 11:30 AM (Day 1) | Vamshi | Google API Key, Python, Flask setup | API connection established & working |
| Sprint 1 | Frontend UI Development | 🟡 Medium | 1 hour (Day 1) | 12:30 PM (Day 1) | Swetha | API response format finalized | Basic UI with input fields |
| Sprint 2 | Image Processing & Annotation | 🔴 High | 2 hours (Day 1) | 3:30 PM (Day 1) | Swetha & Vamshi | API response, UI elements ready | Image annotation functionality using Gemini Vision |
| Sprint 2 | Error Handling & Debugging | 🔴 High | 1 hour (Day 1) | 4:30 PM (Day 1) | Swetha & Vamshi | API logs, UI inputs | Improved API stability |
| Sprint 3 | Testing & UI Enhancements | 🟡 Medium | 1 hour (Day 1) | 5:30 PM (Day 1) | Gayathri & Alekhya | API response, UI layout completed | Responsive UI, better user experience |
| Sprint 3 | Final Testing & UI Fixes | 🟡 Medium | 2 hours (Day 2) | 11:00 AM (Day 2) | Gayathri & Alekhya | API response, UI layout completed | Bug-free, refined UI & improved stability |
| Sprint 3 | Final Presentation & Deployment | 🟢 Low | 2 hours (Day 2) | 1:00 PM (Day 2) | Entire Team | Working prototype | Demo-ready project |

**Sprint Planning with Adjusted Priorities & Timeline**

**Sprint 1 – Setup & Integration (Day 1, Morning)**

* 🔴 High Priority – Set up environment & install dependencies.
* 🔴 High Priority – Integrate Google Gemini Vision API.
* 🟡 Medium Priority – Develop basic UI with input fields.

**Sprint 2 – Core Features & Debugging (Day 1, Afternoon)**

* 🔴 High Priority– Implement image processing & annotation functionalities.
* 🔴 High Priority– Debug API issues & handle errors in responses.
* 🟡 Medium Priority– Initial UI Testing & Enhancements.

**Sprint 3 – Testing, Enhancements & Submission (Day 2, 4 Hours)**

* 🟡 Medium Priority -Test API responses, refine UI, & fix UI bugs.
* 🟢 Low Priority – Final demo preparation & deployment.
* 🟢 Low Priority – Buffer time for last-minute fixes & presentation finalization.

**Phase-5: Project Development**

**5.1 Development Roadmap**

* **Phase 1**: Research & Planning – Define project scope, feasibility analysis, and requirement gathering.
* **Phase 2**: Prototyping & Design – Create UI/UX wireframes, develop mockups, and finalize system architecture.
* **Phase 3**: Core Development – Implement core functionalities such as image processing, annotation tools, and AI integration.
* **Phase 4**: Testing & Optimization – Conduct unit testing, performance testing, and user feedback-based refinements.
* **Phase 5**: Deployment & Maintenance – Deploy the application, provide user training, and establish ongoing support and updates.

**5.2 Technology Stack**

* **Frontend**: Flask, Python, Streamlit
* **Backend:** Flask, Streamlit
* **AI/ML Integration:** Gemini Vision API, Hugging Face, OpenCV
* **Cloud & Hosting:** Google Cloud

**5.3** **Quality Assurance & Security**

* Comprehensive testing strategy covering functional, performance, and security aspects.
* Implementation of encryption, authentication, and access controls.
* Continuous monitoring and logging for issue tracking and improvements.

**Phase-6: Functional & Performance Testing**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Test Case ID** | **Category** | **Test Scenario** | **Expected Outcome** | **Status** | **Tester** |
| TC-001 | Functional Testing | Verify that users can upload images in supported formats (JPEG, PNG). | Relevant budget cars should be displayed. | ✅ Passed | Tester 1 |
| TC-002 | Functional Testing | Test the accuracy of AI-generated annotations on various image categories. " | Seasonal tips should be provided. | ✅ Passed | Tester 2 |
| TC-003 | Performance Testing | Ensure API response time for image processing is under 500ms. | API should return results quickly. | ⚠ Needs Optimization | Tester 3 |
| TC-004 | Bug Fixes & Improvements | Fixed incorrect AI annotation errors and improved accuracy. | Data accuracy should be improved. | ✅ Fixed | Developer |
| TC-005 | Final Validation | Ensure UI responsiveness across different devices (mobile, desktop). | UI should work on mobile & desktop. | ❌ Failed - UI broken on mobile | Tester 2 |
| TC-006 | Deployment Testing | Deploy the app on Streamlit Sharing and verify accessibility. | App should be accessible online. | 🚀 Deployed | DevOps |