**COMSATS University Islamabad,**

**Abbottabad Campus**

**SOFTWARE REQUIREMENTS SPECIFICATION   
(SRS DOCUMENT)**

**Text-To-Image Generator Using Machine Learning**

***By***

**Malik Muhammad Usman CIIT/SP20-BCS-053/ATD**

**Muhammad Haider Ashfaq CIIT/SP20-BCS-040/ATD**

**Supervisor****Ms. Kalsoom Ayaz**

*Bac*helor of Science in Computer Science (2020-2024)

**Table of Contents**

**Revision History 3**

**1.** **Introduction 5**

1.1 Purpose 5

1.2 Scope 5

**2.** **Overall description 5**

2.1 Product perspective 5

2.2 Operating environment 6

2.3 Design and implementation constraints 6

**3.** **Requirement identifying technique 7**

3.1 Use Case Diagram 7

3.2 Use Case Description 8

**4.** **Functional Requirements 11**

**5. Non-Functional Requirements 15**

5.1 Usability 15

5.2 Performance 15

5.3 Accuracy and Quality 15

5.4 Security 15

5.5 Scalability 15

**6.** **Project Gantt chart 16**

**Revision History**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Date** | **Reason for changes** | **Version** |
|  |  |  |  |
|  |  |  |  |

**Application Evaluation History**

|  |  |
| --- | --- |
| **Comments (by committee)**  **\*Include the ones given at scope time both in doc and presentation** | **Action Taken** |
|  |  |
|  |  |

**Supervised by**

**Ms. Kalsoom Ayaz**

Signature\_\_\_\_\_\_\_\_\_\_\_\_

**Introduction****:**

Text-to-image generation is a method used for generating images based on textual descriptions, with significant influence in various research areas and applications. Emphasize the benefits it offers, such as increased efficiency, improved creativity, reduced costs, enhanced realism, and expanded opportunities for visual content creation.

**Purpose:**

The main purpose of our application includes:

* Developing a software application that generates images based on textual descriptions.
* Utilizing the stable diffusion model to ensure semantically consistent image generation.
* Enhancing the accuracy of text-to-image generation by overcoming the limitations of existing algorithms.

The intended audience and users of our software application, such as content creators, designers, marketers, or individuals in need of generating images from textual descriptions can now definitely have an increased efficiency by automating the image creation process, improved creativity by generating innovative visuals, reduced costs by eliminating the need for professional photographers or illustrators, enhanced realism for use in virtual environments, product prototyping, and visual storytelling, easier, more accessible, and cost-effective visual content creation, expanded creative opportunities for users.

**Scope**

The image generation module uses deep learning techniques to create a realistic image that accurately represents the textual description. The scope of this project is to develop a system that can generate high-quality images from textual input with precision and accuracy.

The project's main functionalities include natural language processing (text and voice recognition), image generation, and image editing (Adjust contrast, sharpness, brightness and cropping). The natural language processing component is responsible for understanding and interpreting both textual and voice recognition, identifying the essential elements of the description, and extracting the relevant features for image generation.

**Overall description**

**Product perspective:**

The context of this system is basically to eliminate the reliance on professional photographers or illustrators, the text-to-image generator enhances efficiency and reduces costs. It empowers users to unleash their creativity by generating innovative visuals that match the given textual descriptions. The system provides a user-friendly interface, enabling users to input text or voice descriptions and obtain high-quality images efficiently.

Originating from the need for efficient and precise image generation, this product aims to revolutionize the way visual content is created. It addresses the limitations of existing algorithms for text-to-image generation by proposing a stable diffusion model, which ensures semantically consistent image generation.

Our current system is the next version of a mature system as we are using Stable diffusion model for our main module plus including some other modules as well. There are systems that implement stable diffusion model, but we are basically improving the efficiency this entire model by molding the dataset being used.

**Operating environment**

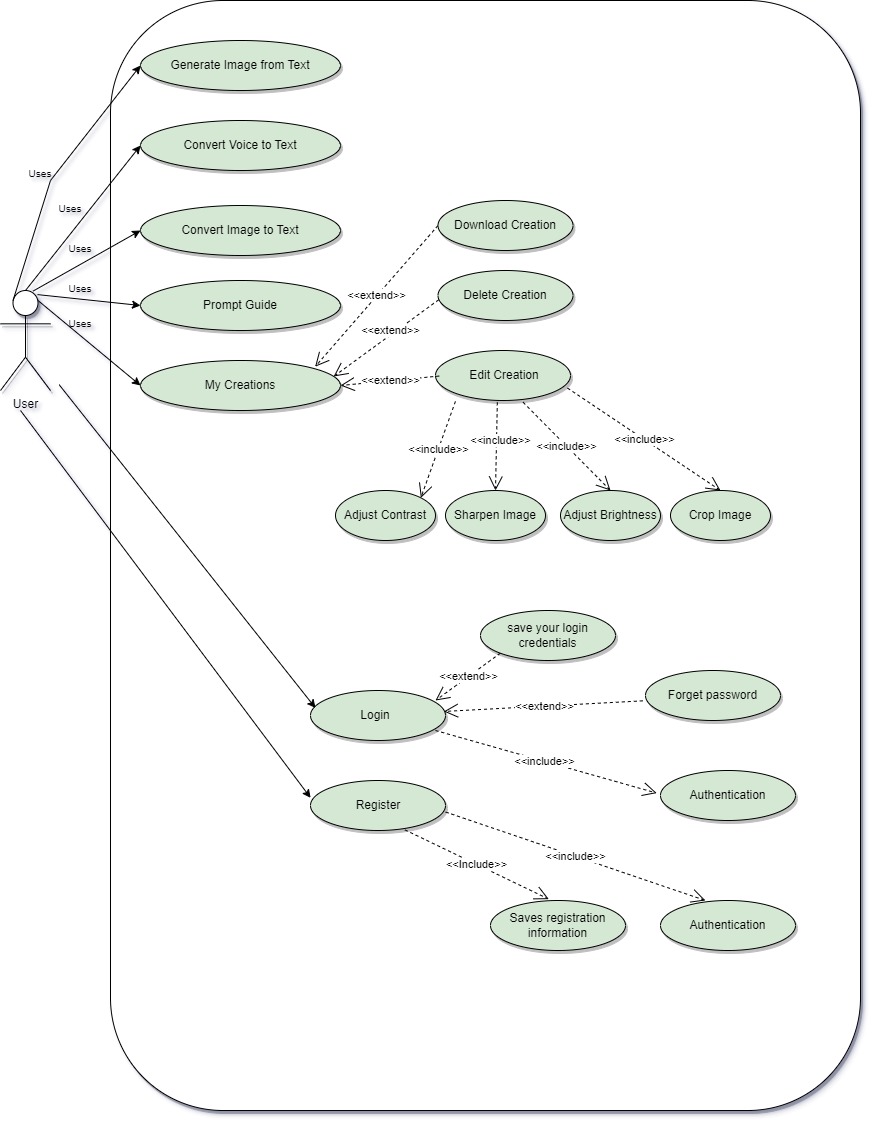
The environment in which our system will be operatable will be a mobile app containing an android operating system and the latest of its version. Geographical location is not a constraint in this regard as we are operating it on a mobile app and that is globally usable for all and the database, we are using is the Cloud Fire-store Database.

**Design and implementation constraints**

* The system shall use Python as the programming language for developing the text-to-image generator.
* The system shall utilize an established machine learning library, such as TensorFlow for model development and training.
* The system shall utilize an image generation and incorporate natural language processing (NLP) model.
* The system shall require a GPU (Graphics Processing Unit) for efficient training and generation of high-quality images.

**Requirement identifying technique:**

**Use case diagram:**



**Use case description:**

**Module 1: Text-To-Image generator**

|  |  |
| --- | --- |
| **Use Case ID:** | UC1 |
| **Use Case Name:** | Generate Image from Text |
| **Actors:** | User |
| **Description:** | Allows users to create images by writing textual description. |
| **Trigger:** | User provides input text. |
| **Preconditions:** | Internet working |
| **Postconditions:** | An image is generated based on the input text. |
| **Normal Flow:** | 1. User provides the input text. 2. The system processes the text using machine learning algorithms. 3. The system generates an image based on the processed text. 4. The system displays the generated image to the user. |
| **Alternative Flows:** | None |
| **Exceptions:** | None |
| **Business Rules:** | The generated image should accurately reflect the input text. |
| **Assumptions:** | The system has the necessary resources and models to generate images. |

**Module 2: Voice-To-Text Converter**

|  |  |
| --- | --- |
| **Use Case ID:** | UC2 |
| **Use Case Name:** | Convert Voice to Text |
| **Actors:** | User |
| **Description:** | Enables users to input text descriptions by speaking, rather than typing. |
| **Trigger:** | User speaks the text. |
| **Preconditions:** | Microphone permission required. Internet working |
| **Postconditions:** | The spoken text is converted into written text. |
| **Normal Flow:** | 1. User activates the voice-to-text functionality. 2. The system listens to the user's speech until stopped. 3. The system converts the speech into text using speech recognition technology. 4. The system displays the converted text to the user in input box. |
| **Alternative Flows:** | None |
| **Exceptions:** | 1a. Mic access not granted. |
| **Business Rules:** | The voice-to-text conversion should accurately reflect the intended message. |
| **Assumptions:** | The system has reliable and accurate speech recognition technology. |

**Module 3: Image-To-Text Conversion**

|  |  |
| --- | --- |
| **Use Case ID:** | UC3 |
| **Use Case Name:** | Convert Image to Text |
| **Actors:** | User |
| **Description:** | Allows users to convert images, such as text photos, into text format. |
| **Trigger:** | User provides an image for conversion. |
| **Preconditions:** | Files & Camera permission required. Internet working |
| **Postconditions:** | Text is extracted from the image & is displayed in the text box. |
| **Normal Flow:** | 1. User provides the image for conversion. 2. The system processes the image using optical character recognition (OCR) technology. 3. The system extracts the text from the image. 4. The system displays the extracted text to the user. |
| **Alternative Flows:** | None |
| **Exceptions:** | 1a. Files & Camera access not granted. |
| **Business Rules:** | The image-to-text conversion should accurately extract the text from the image. |
| **Assumptions:** | The system has reliable OCR technology. |

**Module 4: Prompt Example**

|  |  |
| --- | --- |
| **Use Case ID:** | UC4 |
| **Use Case Name:** | Prompt Example |
| **Actors:** | User |
| **Description:** | Allows users to experiment with pre-existing text examples. |
| **Trigger:** | User selects a pre-existing text example. |
| **Preconditions:** | Pre-existing text examples can be modified. Internet working |
| **Postconditions:** | User can modify and experiment with the selected example. |
| **Normal Flow:** | 1. User selects a pre-existing text example. 2. The system displays the selected example to the user. 3. User modifies or manipulates the example to experiment with the functionality. |
| **Alternative Flows:** | None |
| **Exceptions:** | None |
| **Business Rules:** | The pre-existing text examples should accurately demonstrate how the application or feature works. |
| **Assumptions:** | Pre-existing text examples are provided by the system. |

**Module 5: My Creations**

|  |  |
| --- | --- |
| **Use Case ID:** | UC5 |
| **Use Case Name:** | Manage My Creations |
| **Actors:** | User |
| **Description:** | Allows users to view and manage their own content or creations. |
| **Trigger:** | User accesses the "My Creations" section. |
| **Preconditions:** | User must have at least a single image in creation tab. Internet working |
| **Postconditions:** | User can edit, delete, and download their images. |
| **Normal Flow:** | 1. User accesses the "My Creations" section. 2. The system displays the user's content or creations. 3. User selects an item to perform actions like editing or downloading. 4. The system allows the user to perform the selected action on the item. |
| **Alternative Flows:** | None |
| **Exceptions:** | 3a. Files access not granted.(Download) |
| **Business Rules:** | Users should have centralized control over their own content within the application. |
| **Assumptions:** | Users have created content within the application. |

**Module 6: Download Option**

|  |  |
| --- | --- |
| **Use Case ID:** | UC6 |
| **Use Case Name:** | Download Content |
| **Actors:** | User |
| **Description:** | Allows users to download and save content from within the application. |
| **Trigger:** | User selects the "Download" option for specific content. |
| **Preconditions:** | Content is available for download. Files permission required. Internet working |
| **Postconditions:** | User downloads the selected content. |
| **Normal Flow:** | 1. User selects the "Download" option for specific content. 2. The system verifies the user's request and prepares the content for download. 3. The system provides the user with a download link or saves the content directly to the user's device. |
| **Alternative Flows:** | None |
| **Exceptions:** | 1a. Files access not granted. |
| **Business Rules:** | Users should be able to download and save content securely and easily. |
| **Assumptions:** | Content is available for download within the application. |

**Module 7: Editing Option**

|  |  |
| --- | --- |
| **Use Case ID:** | UC7 |
| **Use Case Name:** | Edit Content |
| **Actors:** | User |
| **Description:** | Allows users to edit content within the application. |
| **Trigger:** | User selects the "Editing" option for specific content. |
| **Preconditions:** | Content is available for editing. Internet working |
| **Postconditions:** | User saves the modified image. |
| **Normal Flow:** | 1. User selects the "Editing" option for specific content. 2. The system provides users with various editing tools and options. 3. User applies desired modifications such as contrast enhancement, sharpness, brightness, or cropping. 4. The system applies the modifications to the content. |
| **Alternative Flows:** | None |
| **Exceptions:** | None |
| **Business Rules:** | Users should have access to a range of editing tools and features to modify content effectively. |
| **Assumptions:** | Content is available for editing within the application. |

**Functional Requirements:**

**Module 1: Text-To-Image generator**

|  |  |
| --- | --- |
| **Identifier** | REQ001 |
| **Title** | Text-To-Image Generator Functionality |
| **Requirement** | The text-to-image generator module shall have the following functionalities:   1. Generate images that accurately reflect the input text. 2. Provide a user-friendly interface for inputting text and generating images. 3. Generate images quickly and efficiently. 4. Ensure high-quality results of the generated images. |
| **Source** | User |
| **Rationale** | The requirement aims to provide users with a simple, efficient, and flexible tool for creating images based on written descriptions or input text. |
| **Business Rule** | None |
| **Dependencies** | None |
| **Priority** | High |

**Module 2: Voice-To-Text Converter**

|  |  |
| --- | --- |
| **Identifier** | REQ002 |
| **Title** | Voice-to-Text Conversion |
| **Requirement** | The voice-to-text conversion module shall accurately and reliably convert spoken words into text, allowing users to input text descriptions by speaking instead of typing. The conversion process should have a high level of accuracy, ensuring that the generated text reflects the intended message. |
| **Source** | User |
| **Rationale** | The requirement aims to improve user experience and accessibility by providing an alternative input method for users who have difficulty typing or prefer speaking over typing. |
| **Business Rule** | The voice-to-text conversion should accurately reflect the intended message. |
| **Dependencies** | None |
| **Priority** | Medium |

**Module 3: Image-To-Text Conversion**

|  |  |
| --- | --- |
| **Identifier** | REQ003 |
| **Title** | Image-to-Text Conversion |
| **Requirement** | The system shall provide image-to-text conversion functionality, allowing users to convert images, particularly text photos, into text format. |
| **Source** | User |
| **Rationale** | The requirement aims to provide users with the capability to extract information from images using OCR and assist users with accessibility needs. |
| **Business Rule** | The image-to-text conversion should accurately extract the text from the image. |
| **Dependencies** | None |
| **Priority** | Medium |

**Module 4: Prompt Example**

|  |  |
| --- | --- |
| **Identifier** | REQ004 |
| **Title** | Prompt Example |
| **Requirement** | The system shall provide a " Prompt Example " option that allows users to experiment with or test out an application or feature using pre-existing Text examples. The " Prompt Example " option should offer users a variety of pre-existing Text examples that demonstrate how the application or feature works. Users should be able to modify or manipulate these Text examples to experiment with the functionality and learn how to use the application or feature effectively. |
| **Source** | User |
| **Rationale** | The requirement aims to enhance user experience by providing a hands-on learning and testing capability for the application or feature. |
| **Business Rule** | The system should store a repository of pre-existing Text examples for each application or feature and provide a user-friendly interface for selecting and modifying them. |
| **Dependencies** | None |
| **Priority** | Medium |

**Module 5: My Creations**

|  |  |
| --- | --- |
| **Identifier** | REQ005 |
| **Title** | My Creations |
| **Requirement** | The user shall be able to view and manage their own content or creations within the application through the "My Creations" option. This includes the ability to edit, delete, and download items. |
| **Source** | User |
| **Rationale** | The requirement aims to provide users with a centralized location to manage their content, enhancing organization and accessibility. |
| **Business Rule** | The system should only allow the user to manage their own content and not provide access to other users' creations. |
| **Dependencies** | None |
| **Priority** | Medium |

**Module 6: Download Option**

|  |  |
| --- | --- |
| **Identifier** | REQ006 |
| **Title** | Download Option |
| **Requirement** | The application shall provide a "Download" option for users to download and save content. |
| **Source** | User |
| **Rationale** | The "Download" option enhances user experience by allowing users to save and access content offline, improving engagement and providing greater control over their content. |
| **Business Rule** | The application must ensure that downloaded content is secure and virus-free. |
| **Dependencies** | None |
| **Priority** | Medium |

**Module 7: Editing Option**

|  |  |
| --- | --- |
| **Identifier** | REQ007 |
| **Title** | Editing Option |
| **Requirement** | The system shall provide an "Editing" option in the user interface that allows users to modify content within the application. The "Editing" option should include tools and options for contrast, sharpness, brightness, and cropping. |
| **Source** | User |
| **Rationale** | The "Editing" option is essential for enhancing user experience and enabling users to create and modify content effectively within the application. |
| **Business Rule** | The system must provide a user-friendly interface for the "Editing" option, ensuring ease of use and intuitive access to editing tools and features. |
| **Dependencies** | None |
| **Priority** | Medium |

**Non-Functional Requirements:**

**Usability:**

* The system should have an intuitive and user-friendly interface to facilitate easy interaction with the text-to-image generator.
* The user should be able to input text and generate an image with minimal effort and without the need for advanced technical knowledge.
* The system should provide clear and concise instructions or guidance on how to use the text-to-image generator effectively.

**Performance:**

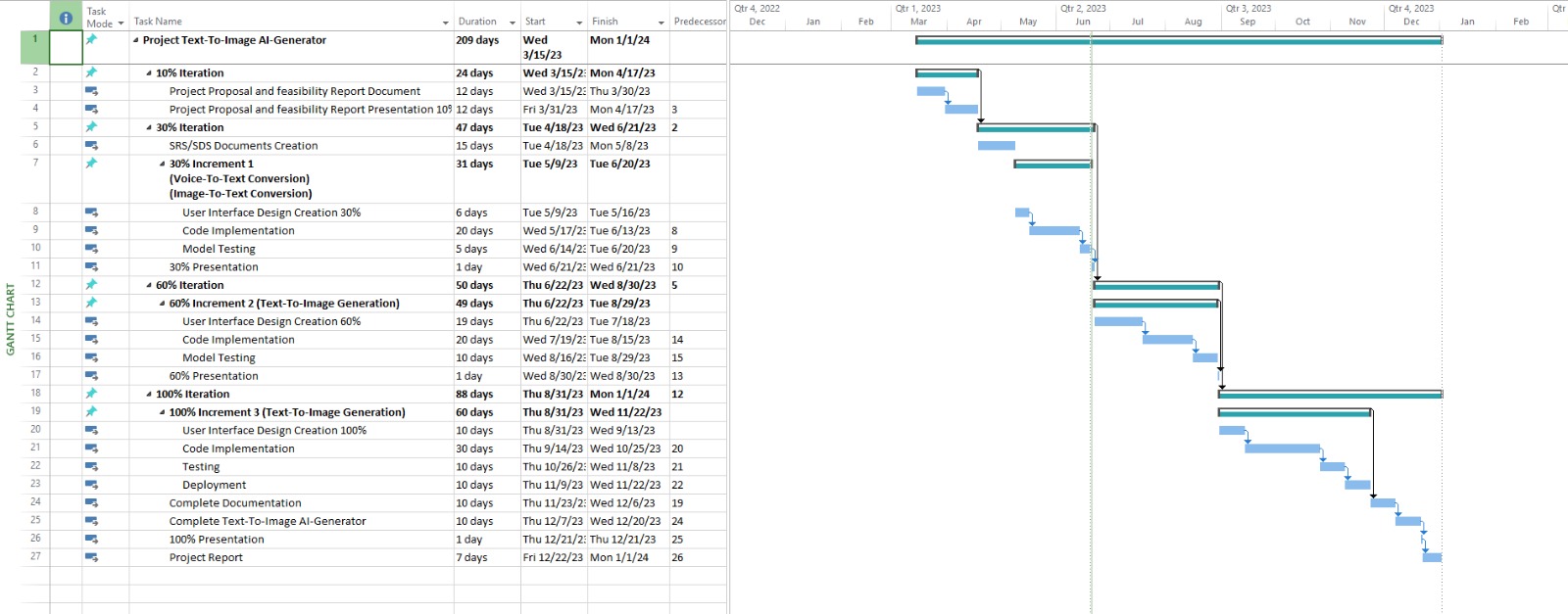
* The text-to-image generation process should have a fast response time, generating images within an acceptable timeframe.
* The system should be capable of handling multiple concurrent user requests without significant performance degradation.
* The system should be capable of handling high volumes of text data efficiently, processing and generating images within a reasonable timeframe.

**Accuracy and Quality:**

* The generated images should accurately represent the input text and reflect the intended meaning or concept.
* The system should strive for high-quality image generation, producing visually appealing and realistic images.

**Security:**

* The system should incorporate appropriate security measures to protect user data and prevent unauthorized access or misuse.
* Any sensitive or personal information provided by the user should be handled securely and in compliance with relevant privacy regulations.

**Project Gantt chart:**