**Requirements Specification Document for Voice-to-Image Recognition System**

**Main Objective**

The system will take voice input and provide an image output based on the input description.

**Scope**

The voice-to-image recognition system will:

* Take voice input in the form of an audio file.
* Convert the audio file into a text description using speech recognition.
* Use the text description to retrieve the most relevant image from a database of images.
* Provide the image output to the user.

**Functional Requirements**

* The system shall convert the voice input into text using a speech recognition module.
* The system shall retrieve images from a database of labeled images based on the text description.
* The system shall use a pre-trained image retrieval model to retrieve the most relevant image.

**Non-Functional Requirements**

* The system shall be able to process a voice input and provide the corresponding image.
* The system shall be developed using Python programming language.
* The system shall use TensorFlow and Keras deep learning libraries.
* The system shall be deployed on a cloud server.

**Conclusion**

This requirements specification document defines the functional and non-functional requirements for a voice-to-image recognition system. The document will serve as a guide for the development and testing of the system.