

Presentación de propuesta:

Conversión de imagen a textos y audio

Redes de computadores

Docente: Alexander Toscano Ricardo

Presentado por:

Lemuel David Marzola Baron

Sebastián López Sotelo

José Isaac ChantaK Zabaleta

Xiomara Hoyos Solano

Universidad de Córdoba

Montería Córdoba

2023-1

Nuestra propuesta consiste en desarrollar un complemento innovador para un api existente, que permita a los usuarios convertir imágenes en t este complemento permitirá a los usuarios convertir imágenes en texto y audio de manera eficiente y precisa. A través de esta api tendremos reconocimiento de imágenes avanzado y se utilizará algoritmos de aprendizaje automático y tecnologías de reconocimiento de imágenes para identificar y extraer texto de forma precisa de las imágenes cargadas por los usuarios, esto permitirá convertir documentos, letreros, pizarras y otros contenidos visuales en texto editable, y a su vez, se podrá en audio, la aplicación contará con tecnología de síntesis de voz de alta calidad que transformará el texto extraído en archivos de audio y los usuarios podrán ajustar parámetros como el tono y la velocidad para personalizar la experiencia auditiva. Esta aplicación será una herramienta útil para personas con discapacidades visuales, estudiantes, profesionales y cualquier persona interesada en extraer información de imágenes y compartirla de forma accesible.

Esta api contará con una interfaz intuitiva y accesible, ya que enfocaremos en desarrollar una interfaz de usuario sencilla y amigable, con una navegación intuitiva y compatibilidad con lectores de pantalla. De esta manera, la aplicación será accesible para personas con discapacidades visuales o dificultades de aprendizaje. Tendremos opción de iniciar sección o ingresar como invitado, y posterior a esto, aparecerá la opción de seleccionar la imagen que vamos a convertir. Una vez ingresada la imagen, automáticamente se convertirá a texto y aparecerá la opción para reproducir audio.

Con esta api se brindará una solución eficiente y accesible para extraer información de imágenes y convertirla en contenido legible y audible. Nuestro enfoque en la precisión, la facilidad de uso y la seguridad garantizará una experiencia positiva para los usuarios.

¿Cuál es el objetivo principal de la aplicación? ¿Cuál es el problema o necesidad que busca resolver?

El objetivo principal de la api es poder insertar cualquier tipo imágenes, y que a través de un escáner detallado, la api sea capaz de dar una descripción de la imagen en forma de texto, y que posteriormente se pueda reproducir en forma de audio, entonces lo que se busca es poder convertir las imágenes en una descripción a través de audio, esto con el fin de que personas con algún tipo de discapacidad visual, puedan tener noción del contenido de cualquier imagen, y en general cualquier persona que busque convertir una descripción de imagen a través de audio, pueda realizarlo por medio de la api.

¿Quiénes son los usuarios de la aplicación? ¿cuáles son sus características y necesidades específicas?

Los usuarios del api son personas que tienen alguna discapacidad visual y no pueden ver imágenes o leer texto y en general cualquier tipo de persona que busque tener una descripción auditiva de cualquier imagen. Entonces esta api busca facilitarles a estos usuarios que puedan comprender y saber de qué tratan las imágenes, a través de audios, en el cual les reproduzca una descripción detallada de la imagen. Esta api la puede usar cualquier Usuario ya que le facilita la lectura de la descripción de cualquier imagen en un Audio de manera más rápida y eficaz, entonces, las necesidades o características específicas de este tipo de usuarios, es que necesiten una descripción en forma de audio de cualquier imagen.

¿Qué funcionalidades se requieren para cumplir con el objetivo de la aplicación? ¿Cuáles son las más importantes?

En primer lugar, se requiere que el usuario tenga una imagen de la cual busca una descripción a través de audio, luego de esto se requiere que el usuario inserte la imagen a la api a través del canal dispuesto para esto, en este caso es un espacio para introducir archivos en formato de imagen, luego de esto aparecerá la opción de describir la imagen

y el usuario podrá escuchar de qué trata. La funcionalidad más importante en este caso es insertar la imagen de la cual se requiere una descripción.

¿Cómo debe ser la interfaz de usuario? ¿Qué elementos gráficos y de diseño son importantes para los usuarios?

La Interfaz al Principio es una pestaña con mucha claridad y concisión, donde se pide un registro por medio de correo electrónico o la opción de ingresar sin usuario a través de la opción de invitado, esto puede quedar a gusto de la persona que ingresa, si es a través de correo, al usuario luego de presionar en esta opción, se le desplegará una sección donde podrá elegir con qué tipo de cuenta de correo quiere ingresar, Hotmail, Gmail, o con alguna otra cuenta como Facebook, Instagram o Twitter, al ingresar con cualquiera de estos medios de correo electrónico, aparecerá una opción para ingresar los datos, correo, y contraseña, nombres y numero de teléfono, posterior a esto una pestaña de recaptcha para verificar que es un humano el que está ingresando, luego de ingresar su correo y su contraseña aparece un espacio con atractivo visual y la opción dispuesta para seleccionar archivos bien ya sea de la computadora o de cualquier dispositivo inteligente con el que se esté ingresando, al usuario seleccionar la imagen, se despliega un texto con la opción de reproducir el audio, un elemento importante para el usuario es que los botones para hacer selección, será grandes para que se pueda hacer una selección con bastante precisión.

¿Qué plataformas y dispositivos deben ser compatibles con la aplicación? ¿Deben ser móviles, de escritorio, o ambos?

Las plataformas compatibles son cualquier dispositivo inteligente con acceso a internet esto con el fin de no limitar a los usuarios a usar una plataforma en específico, debido a que no todos pueden contar con una computadora, o un teléfono o tableta, entonces se busca tener una diversidad de plataformas de uso y dar la opción al usuario de usar la plataforma que tenga a la mano o desee para ingresar.

¿Qué requisitos de seguridad y privacidad son necesarios para la aplicación?

Un requisito de seguridad es el inicio de sección con un correo electrónico, pero esta opción queda a gusto del usuario, como ya se mencionó en caso de que elija ingresar con correo u otra cuenta, se le pedirán una serie de datos personales los cuales estarán protegidos por parte de la plataforma para que no caigan en manos indebidas. En caso de no querer dar los datos personales, aparecerá también una opción para entrar como invitado a la aplicación, para que el usuario que lo desee, esto con el fin de no exigir datos personales.

¿Cuáles son los plazos y presupuestos para el desarrollo de la aplicación?

Se pretende crear la aplicación en el lapso de tiempo que sea necesario para efectuar su funcionamiento, podría decirse que un aproximado de 20 días, partiendo de que ya se tiene la api principal que se encarga de reconocer las imágenes y dar una descripción detallada de estas, ahora bien, el presupuesto con el que se cuenta para el desarrollo es el capital humano y los dispositivos inteligentes que hay lo cual es un capital fundamental para el desarrollo de la aplicación.

¿Hay alguna tecnología o herramienta específica que deba ser utilizada en el desarrollo de la aplicación?

En el caso de nuestra aplicación, al tratarse de un programa que examina imágenes para dar una descripción detallada de ella, se necesitará de una tecnología de reconocimiento de objetos y de lectura de textos, para que el programa sea capaz de describir en forma auditiva, el texto de la descripción de la imagen, esto se puede lograr mediante el uso de una api de reconocimiento de objetos y o imágenes, que nos permita tomar los detalles de la imagen y reproducirlos al tener la descripción ya lista.