

Real Time Communication System Powered By AI For Specially Abled

Abstract:

The sixth sense is a multi-platform app for aiding the people in need that is people who are handicapped in the form of lack of speech (dumb), lack of hearing (deaf), lack of sight (blind), lack of judicial power to differentiate between objects (visual agnosia) and people suffering from autism (characterized by great difficulty in communicating and forming relationships with other people and in using language and abstract concepts). Our current implementation of the product is on two platforms, namely, mobile and a web app. The mobile app even works for object detection cases in offline mode. What we want to achieve using this is to make a better world for the people suffering from disabilities as well as an educational end for people with cognitive disabilities using our app. The current implementation deals with object recognition and text to speech and a speech to text converter. The speech to text converter and text to speech converter utilized the Web Speech API (Application Program Interface) for the website and text to speech and speech to text library for the mobile platform. The object recognition wouldn't fetch enough use out of a website. Hence, it has been implemented on the mobile app utilizing the Firebase ML toolkit and different pre-trained models, which are both available offline as well as online. Keywords: Sixth sense, disabilities, Web Speech API, Firebase ML toolkit, cognitive disabilities.

Feature Extraction:

CNN:

A convolutional neural network (CNN) is a type of artificial neural network used in image recognition and processing that is specifically designed to process pixel data.

CNNs are powerful image processing, artificial intelligence (AI) that use deep learning to perform both generative and descriptive tasks, often using machine vision that includes image and video recognition, along with recommender systems and natural language processing (NLP).

Artificial Intelligence:

Artificial intelligence is a field that combines computer science and robust datasets to enable problem-solving. Expert systems, an early successful application of AI, aimed to copy a human's decision-making process. In the early days, it was time-consuming to extract and codify the human's knowledge.

AI today includes the sub-fields of machine learning and deep learning, which are frequently mentioned in conjunction with artificial intelligence. These disciplines are comprised of AI algorithms that typically make predictions or classifications based on input data. Machine learning has improved the quality of some expert systems, and made it easier to create them.

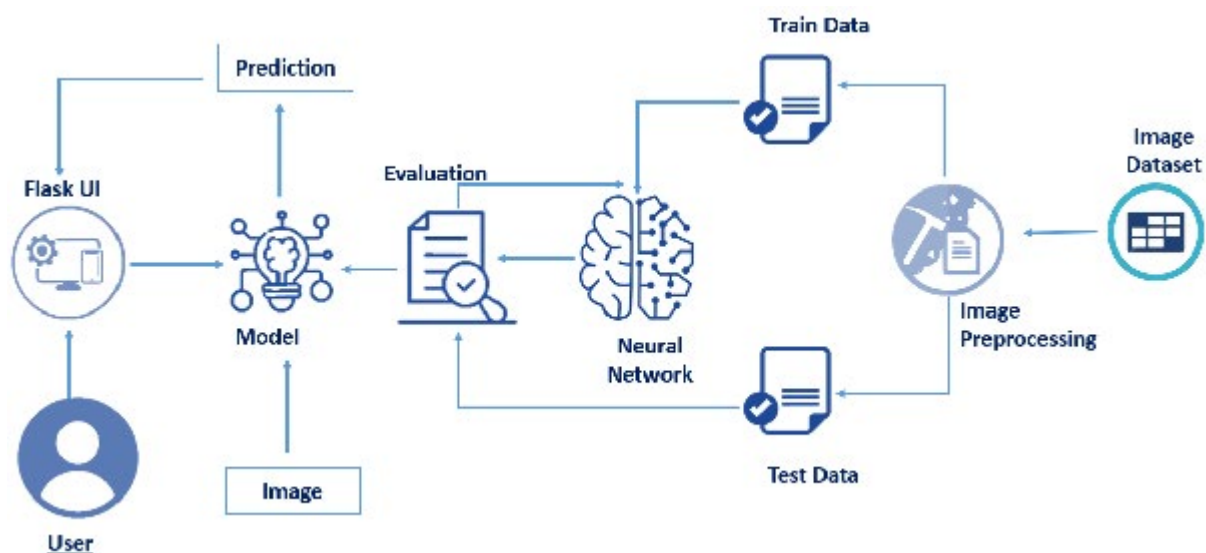
Today, AI plays an often invisible role in everyday life, powering search engines, product recommendations, and speech recognition systems.

Deep learning:

Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to “learn” from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimize and refine for accuracy.

Deep learning drives many artificial intelligence (AI) applications and services that improve automation, performing analytical and physical tasks without human intervention. Deep learning technology lies behind everyday products and services (such as digital assistants, voice-enabled TV remotes, and credit card fraud detection) as well as emerging technologies (such as self-driving cars).

Technical Architecture:



Conclusion and Future Scope :

Implemented the object tracking, recognition & classification, and character recognition in offline mode and guarded the app to shrink the size of the app. The main highlight of the project hence came forward, which was being an application that provided a one-stop-shop solution to all the sections of differently-abled people. Integration has provided a seamless User interface/experience for the initial setup. Another point achieved here was no extra hardware; hence, no additional cost to utilize the service. The application still does depend on the camera picture quality for object detection and OCR but is still high enough in confidence level (70%) for most of the cases which were covered. Higher accuracy could be achieved in the future scope of the implementation through the use of custom models for object detection and text recognition as it could take into account the cases of objects from differently-abled people and work on those only yielding faster and accurate results.