Image Analyzer using CNN

Rohit Kumar B[1], Sharanayya[2], Shashank M[3], Shankar D Navali[4], Maya B S[5]

Student, Computer Science and Engineering, Bangalore Institute of Technology, Bengaluru, India[1,2,3,4]

Associate Professor, Computer Science and Engineering, Bangalore Institute of Technology, Bengaluru, India[5]

Abstract: An Image Analyzer is an advanced system designed to process, analyze, and interpret visual data from digital images. By employing techniques from image processing and artificial intelligence, it automates the extraction of meaningful information, making it a powerful tool across diverse domains. The system typically consists of key modules are input handling, preprocessing, feature extraction, analysis, and output generation. The analysis module integrates traditional computer vision algorithms with modern AI techniques, including convolutional neural networks (CNNs), for tasks like object detection, classification, and segmentation. The results are presented in user-readable formats, enabling actionable insights.

Keywords: Image Analyzer, Artificial Intelligence, Computer Vision, Convolutional Neural Network.

I. INTRODUCTION

The Image Analyzer Project is an initiative aimed at developing a system capable of processing and interpreting digital images to extract meaningful insights automatically. In today's digital era, vast amounts of visual data are generated daily from sources like smartphones, surveillance systems, medical imaging devices, and industrial cameras. Analysing this data manually is time-consuming, prone to errors, and often infeasible at scale [1]. The Image Analyzer addresses this challenge by leveraging advanced image processing techniques and artificial intelligence (AI) to automate the analysis process.

The purpose of the Image Analyzer is to automate the process of interpreting and extracting meaningful information from digital images. It aims to replace manual analysis with efficient, accurate, and scalable solutions that can handle large volumes of image data. The system leverages advanced image processing, machine learning[4], and computer vision techniques to detect patterns, identify objects, and classify content within images.

II. BACKGROUND

There are several existing systems and technologies that enable the analysis of digital images, each leveraging different methods and algorithms to extract meaningful information. Existing image analysis systems predominantly rely on traditional image processing. Techniques or partially automated solutions. These Existing System lack the flexibility and scalability required to handle the growing volume and complexity of visual data[2-3]. These systems often depend on predefined rules and manual intervention, making them limited in their ability to adapt to diverse and evolving datasets. TensorFlow, an open-source machine learning framework developed by Google, and Keras (a high-level API for neural networks) are commonly used for image classification, object detection, and segmentation tasks[6-7].

III. METHODOLOGY

The proposed system consists of designing the mobile application for recognition of image. Generating the data of the captured image and converting the text to speech. The objective of proposed system is to automate image analysis tasks such as object detection, classification, and anomaly identification, system need to integrate machine learning and computer vision techniques into mobile application or system[5]. The design of Algorithms capable of extracting meaningful features and insights from images with minimal preprocessing. The block diagram shown in below figure1 outlines the flow of a system or application. It depicts the sequence of steps from the start to the end of the process, emphasizing user interaction and data flow. The process begins at the Start node, followed by navigating to the Welcome Page, which serves as an introductory page. From the Welcome Page, the user can navigate to the Home Page (branding page) or proceed directly to the Image Analyzer for analysis.
 The Image Analyzer processes input data, which can be obtained from a camera or storage. The input undergoes image processing to generate data, which is then converted into speech output through a Text-to-Speech (TTS) module. The process concludes at the End node, with the option to navigate back to the Home Page for further actions. The

diagram highlights navigation buttons such as "Get Started," "Eye," and "Home" that facilitate the flow between components.
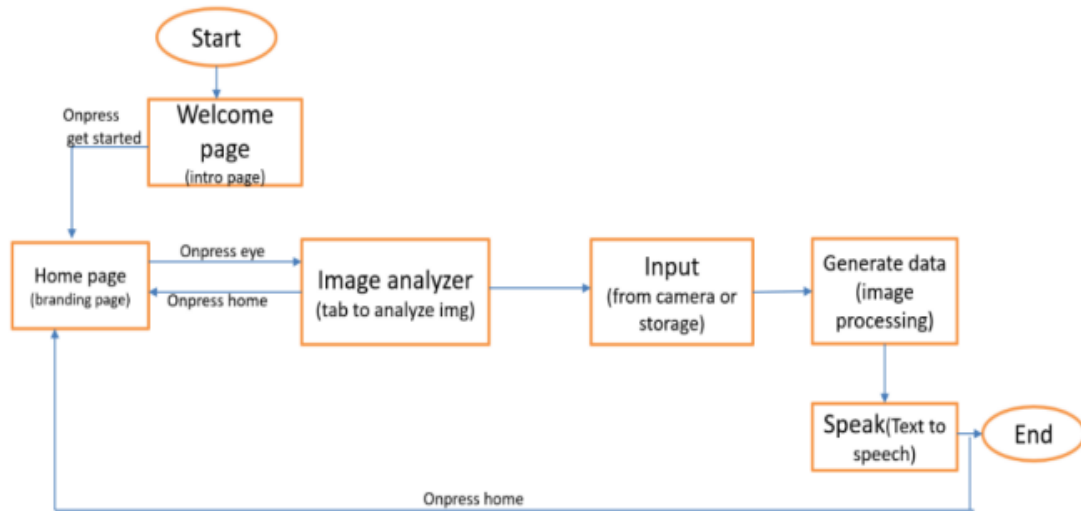

Fig .1 Proposed system of the Image Analyzer system

Modules Description

A. **React Native:** React Native is an open-source framework developed by Facebook that enables developers to build cross-platform mobile applications using JavaScript and React. Leveraging a single codebase, React Native allows for the creation of apps that run natively on both iOS and Android, significantly reducing development time and effort.

B. **Expo-image picker**

The Expo Image Picker is a popular library within the Expo ecosystem, designed to simplify the process of selecting images or videos in React Native applications. It provides an intuitive and customizable interface that allows users to either pick media files from their device's gallery or capture new ones using the camera. Expo Image Picker is a library in the Expo ecosystem that allows React Native developers to easily access device media, such as photos and videos [9]. It provides a simple interface for selecting images or videos from the camera roll or capturing new ones using the device camera, streamlining media integration in mobile applications. It also supports customizable options like media type selection, aspect ratio, and quality settings, giving developers flexibility to tailor the image or video picking experience to their app's needs.

C. **Expo Speech**

The block diagram represents an application designed for image analysis and text-to-speech conversion, offering a seamless and user-friendly experience. The process begins at the Start node, where users are directed to a Welcome Page that introduces the application. From there, users can navigate to the Home Page, which serves as a branding hub showcasing key features[8]. The core functionality lies in the Image Analyzer, where users can input images either by capturing them through a camera or uploading them from storage.

D. **Expo-router**

Expo Router is a routing library for React Native applications that simplifies navigation by bringing web-like routing capabilities to mobile apps. Built on top of React Navigation [10], it allows developers to define routes in a filebased structure, similar to frameworks like Next.js. This approach makes it intuitive to organize and manage navigation, supporting nested and dynamic routes effortlessly. Expo Router integrates seamlessly with Expo projects, enabling features like deep linking, dynamic parameters, and stack-based navigation.

E. **Gemini**

Gemini AI, AI model designed to seamlessly integrate and process diverse data types, including text, images, audio, and video. Its sophisticated reasoning abilities enable it to perform advanced tasks such as natural language understanding, image recognition, and coding with state-of-the-art precision. Gemini's architecture is built for scalability and efficiency, optimized using Google's advanced Tensor Processing Units (TPUs). It comes in various configurations, including Ultra, Pro, and Nano, catering to applications from high-performance computing to mobile devices.

IV. RESULTS

Performance analysis involves evaluating the efficiency and scalability of the image analysis system, such as how quickly the system processes images, how much memory it uses, and how it scales with larger datasets or more complex images. The image Analyzer generates the caption for the given image, which contains the entrance of Bangalore Institute of

Technology. The figure 2 shows the entrance to Bangalore Institute of Technology and Vindhana Soudha. A grand gateway with the institute's name prominently displayed in English and Kannada. Several people are seen near the gate. Suggesting a campus environment. The image Analyzer generates the caption for the given image, which contains legislature building Called Vindhana Soudha. It can also read aloud the generated Image caption.
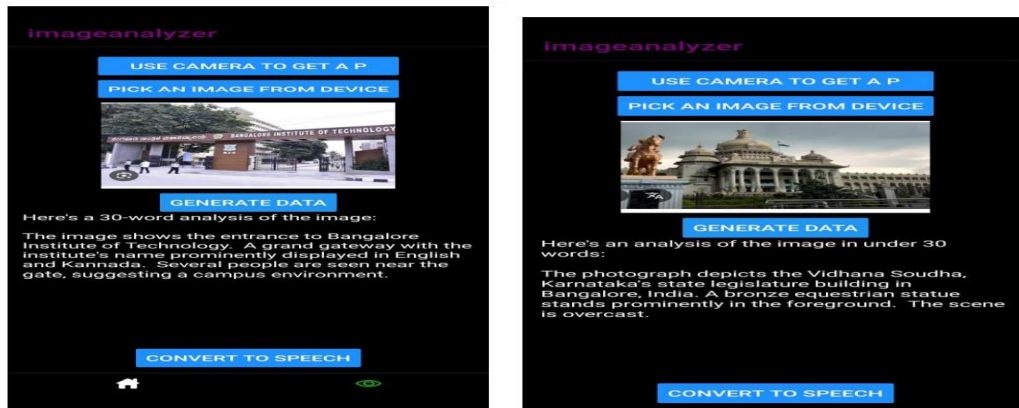

Fig 2: Performance of Image Anayzer

V.   CONCLUSION

The image analyzer is a versatile and powerful application that utilizes advanced technologies such as machine learning, deep learning, and computer vision to interpret and extract meaningful information from images. It finds applications across diverse domains, including healthcare, accessibility, security, agriculture, and more, making it an indispensable tool in modern technology. Image caption generators bridge the gap between visual content and textual interpretation, enhancing accessibility, automation, and efficiency across various domains. By leveraging AI, they streamline workflows, improve content management, and open new possibilities in accessibility, marketing, and user experience.

REFERENCES

[1]. Neha Sharma, Vibhor Jain, Anju Mishra, " An Analysis Of Convolutional Neural Networks For Image Classification, Procedia Computer Science, Volume 132,,Pages 377-384, 2023.

[2]. C. Titouna and F. Naït-Abdesselam, "A Convolutional Neural Network-based Approach For Image Analysis and Injection Detection," *2024 IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Niagara Falls, ON, Canada, pp. 1-6, 2024.

[3]. Tripathi, Milan, "Analysis of Convolutional Neural Network based Image Classification Techniques", Journal of Innovative Image Processing., Vol.3, pp.100-117, 2021.

[4]. Leelavathi Arepalli ,P Indu ,P Rohitha ,K.T.P Dharani ,G Devi ,D Gandhi ,P Sri Sowmya , " Medical Image Analysis by Using CNN Approach " , International Journal of Advance Computational Engineering and   Networking (IJACEN) , pp. 7-11, Volume-12,Issue-5, 2024.

[5]. Giriraj Gautam, Anita Khanna,"Content Based Image Retrieval System Using CNN based Deep Learning Models",Procedia Computer Science,Volume 235,2024,

[6]. Lee, Daeun & Yoon, Jaehong & Cho, Jaemin & Bansal, Mohit," VideoRepair: Improving Text-to-Video Generation via Misalignment Evaluation and Localized Refinement", 2024.

[7]. Rui Zhao and Hangjie Yuan and Yujie Wei and Shiwei Zhang and Yuchao Gu and Lingmin Ran and Xiang Wang,"EvolveDirector: Approaching Advanced Text-to-Image Generation with Large Vision-Language Models",The Thirty-eighth Annual Conference on Neural Information Processing Systems, 2024.

[8]. Gaurav Meena, "Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach", International Journal of Information Management Data Insights, Vol.3, Issue.1, 2023.

[9]. S, Prasantha,"Image Text To Speech Conversion In Desired Language. International Journal Of Creative Research Thoughts", 11. b361-b370,2023.

[10]. Shaik Karishma , Siddu Devi Naga Susmitha , Nanditha Katari , G. Sirisha," Speech to Image Conversion", SSRG International Journal of Computer Science and Engineering, Volume 10 Issue 10, 1-5, October 2023.