TEXT, FACE AND OBJECT DETECTION BY USING MACHINE LERNING

Shweta Sahu*1, Mansi Sahu*2, Anjali Choudhary*3,

Sanjay Kashyap*4, Soniya Wadhwa*5

*1,2,3,4,5Computer Science & Engineering, Government Engineering College Bilaspur C.G., India.

ABSTRACT

Text, Face & Object Detection using Machine Learning is an application which offers users to apply machine learning-based functionalities in the real-life event. This application will give the users to recognize text and translate the text as well as facial recognition, detecting objects and barcode scanning by using their smart phone camera. purpose of this project is to develop a user-friendly android mobile application that will help to understand and interact this purpose we used toolkit and languages like Android Studio, Machine learning kit, java, etc. Google Lens is a set of vision-based computing capabilities that can understand what you are looking at and use that information to copy or translate text, identify plants and animals, explore locales or menus, discover products, find visually similar images, and take other useful actions. Google lens is an image recognition technology, designed to bring up relevant information related to objects it identifies using visual analysis based on a neural network.

Keywords: Google Lens, Android Studio, Firebase, ML Kit, Java.

I.   INTRODUCTION

Machine Learning has started to reshape how we live, everything around us now a days has a touch of ML in it, so has mobile applications. Google has re-presented Firebase with the help of some prepared to utilize machine learning models like-Text recognition, Face detection, Image labeling, and more and backing to actualize custom models utilizing Tensor Flow Lite for iOS, Android and Web applications. Smart phones are more advanced than ever. They're filled with plenty of features, like the front- and rear-facing cameras, pulse rate monitors, fingerprint screen lock, and access to apps galore. There's no denying it, from the minute we get up to the minute we head to sleep, we're altogether stuck to our cell phones. Another work messages. Another Facebook notification.

The weather application discloses to us a tempest is headed. There's constantly another motivation uplift our smart phones since they do as such much. The smart phone is the most widely used mobile technology right now. We can't think of life without a smart phone. It became a part of our day to day life. We do a lot of day to day activities with our smart phone. We all have one in our pocket and that has cameras. Google lens is best described as a search engine for the real world. It uses artificial intelligence to identify text and objects both within images and in a live view from your phones camera, and it then lets you learn about and interact with those elements in all sorts of interesting ways. Cloud vision allows to easily vision detection features within applications, including image labeling, face and landmark detection, optical character recognition(OCR), and tagging of explicit content. It can detect text in images, text in files, handwriting in images and detect faces

Text Recognition

•Identifying the content from a picture. It very well may be utilized for photographs of signs, names, labels, etc. It can be useful for

many cases like detecting text from an image and make a copy of that text.

Translate Text

•Translate text from English to Bengali. It can translate a sentence or also word by word like a dictionary.

Image Labeling aking a picture, and identifying substances it contains, for example, objects, creatures, natural products, exercises and that's just the beginning.

Face Detection

•Not to be mistaken for face acknowledgment, which can perceive who is the individual in the picture, or realizing that we see a similar individual on numerous photographs. This is about detecting face features, whether the person is smiling or not or eyes isopen or not.

Barcode Scanning

Peruse information encoded utilizing most standard standardized tag groups.React astutely when a client examines a standardized tag.

Language Identification

Identify a language based on the user input.

II.   LITERATURE SERVEY

Dr. S. M. Mukane et.al[1], proposed to have a method for classification of flowers using Artificial Neural Network (ANN)classifier. They proposed a method which was based on textural features such as Gray level co-occurrence matrix (GLCM) and discrete wavelet transform (DWT). Using a threshold based method a flower image was segmented. They used a database of flower images which is a combination of images captured byre searchers or downloaded from World Wide Web. The ANN classifier has been trained using 50 samples in order to classify 5classes of flowers. They achieved more than 85% classification accuracy using GLCM features only.

Fadzilah Sirajet.al.[2], conducted a study with a dataset consists of 180 patterns with 7 attributes for each type of flower has been gathered. The finding from the study reveals that the number of images generated to represent each type of flower influences the classification accuracy. One interesting observation is that duplication of very hard to learn images assist Neural Network toimprove its classification accuracy. This is also another area thatcould lead to better understanding towards the behaviour of images when applied to Neural Network classification.

D.S. Guruet.al.[3], investigated the effect of texture features forthe classification of flower images. They segmented a flower image through eliminating the background and thus making useof threshold-based method. They extracted texture features, namely the color texture moments, gray-level co-occurrence matrix, and Gabor responses then made combinations of thesethree features in classification of flowers. They used aprobabilistic neural network as aclassifier. They used 35 classes in their own dataset where each set included 50 samples in order]to verify efficiency of the proposed method.

Viswanadhuni et al.[4]proposed a domain knowledge embedding model as a content provider for AI camera solution (e.g., Google Lens, Bixby Vision). The output of this model aids in obtaining appropriate information from the content provider. Their research Intent Classification model predicted three types of intents with91 percent accuracy including beauty product purchase interest, generic information seeking, and movie information seeking intents.

Sergeeva et al.[5]examined the possibility of utilizing Augmented Reality in teaching foreign language. Forty students took part in VR/AR course introducing AR tools (e.g., Google Lens, AR browsers, and WallaMe service). Their studies reported that AR can be applied in foreign language classes as it contributed to the optimization of the educational process by filling it with information, involving students, and successfully influencing the process of developing the student's foreign language competency.

Rebekah et al.[6]utilized Google Lens as a means to count the number of employees based on the facial detection feature of Google Lens. The corresponding image frames of the detected person were kept extracting desirable features such as ID card and Shoe.

Kumar et al.[7]utilized Google Lens to split the text from the image, which was then processed as independent entities and delivered to the text and image analytics modules.

Lucia et al.[7]investigated the consequences of altered set of spatial, gestural and cognitive relations in order to better understand how locative media may impact everyday users' interactions and experiences with their surroundings. They argued that the application advances reductive representations of complicated sets of relationships formed by locative media and augmented reality. Locative media, in this sense, is defined as the media linked to a given geographical location.

III.   METHODOLOGY

Our developed application is based on Google's own firebase machine learning kit which was first introduced in Google I/O 2018. After researching about related work, we found on "**Google Lens**" which matches with our developed application.

Google Lens is one of the most up to date increments to the universe of ML and AI reasoning administrations. Presented in 2017 at Google I/O, Google Lens is a picture acknowledgment versatile application that can educate you regarding an article, milestone, or item by examining it through a perspective.
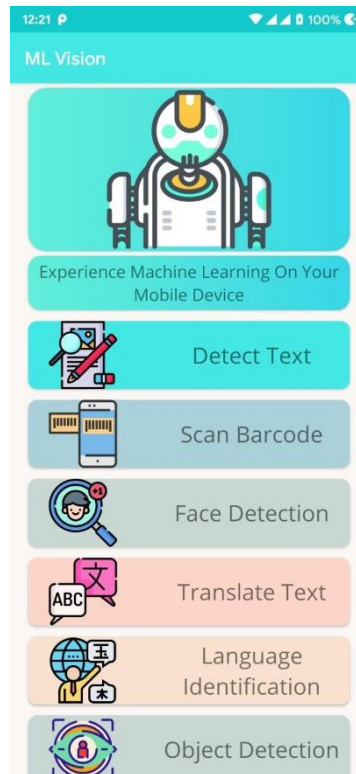

IDENTIFIED DATA

An ML algorithm would be prepared on preparing information to 'learn' to perceive facial milestone. Where a straightforward calculation would not be fit for playing out this undertaking, an ML additionally would not exclusively have the option to recognize the photographs as prepared, it would consistently gain from testing information and add to its "learning" to turn out to be progressively exact in its forecasts.

ML Kit's base APIs cover:

- Barcode scanning, to scan and process barcodes.
- Text recognition.
- Face detection.
- Landmark detection which identify popular landmarks.
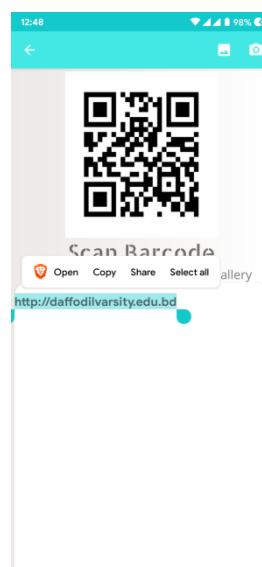- Image labeling, which identify objects, activities, products etC
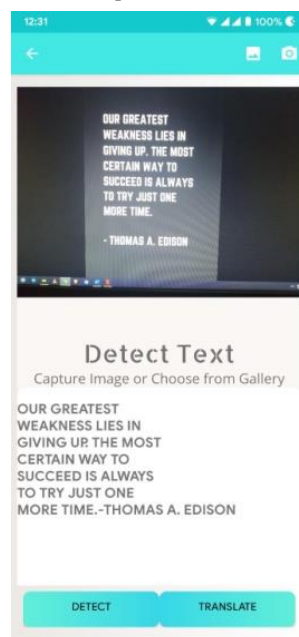
TEXT DETECTING USER INTERFACE

SCAN BARCODE: This part of the user interface is very simple. The user just needs to choose the barcode image from the gallery or capture it using a camera.
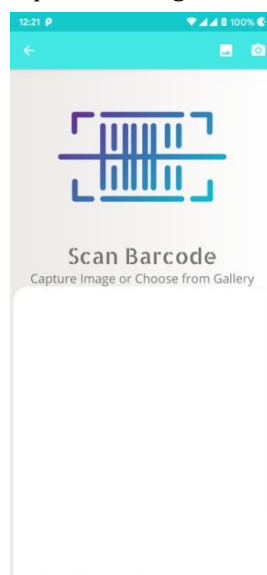
BARCODE SCANNING USER INTERFACE   OUTPUT OF THE BARCODE SCANNING

Text Translation

Translate text from English to Bengali. It can translate a sentence or also word by word like a dictionary

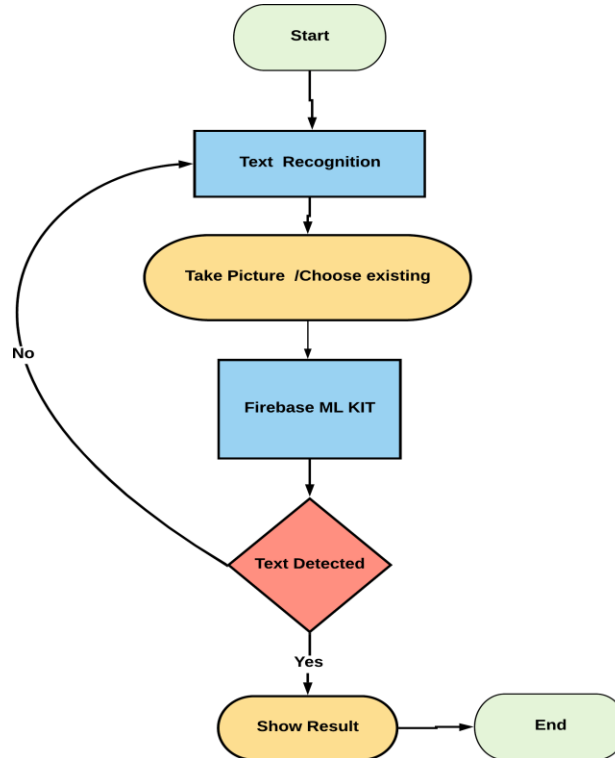TEXT TRANSLATOR USER INTERFACE

Object Detection

Here user can take an image, and detect entities it contains, such as objects, animals, fruits, activities and more

OBJECT DETECCTING USER INTERFACE

BARCODE SCANNING: RECOGNIZE TEXT

Barcode detection USE case diagram Text recognizing use case diagram

Use Case Model and Description

This following Use Case Diagram shows the basic model of our application also the relation between a user and a functional model.
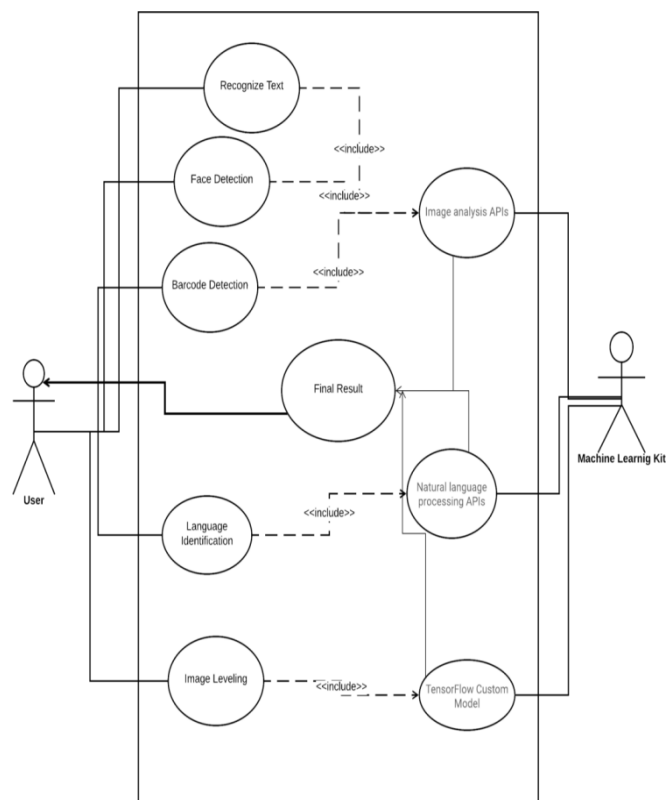

Fig: Use case diagram

IV. THE PROBLEM

Scope of Problems

As we wanted to make an application by which a user can interact with what's in our mobile device cameras viewfinder. Our developed application working principle is dependent on the device camera. Considering that these are the problems we faced while developing this application.

Camera hardware of the device may cause problems. As our application is fully dependent on the device camera so if the camera of the user's device is not capable to capture a good quality image then it may not deliver expected accurate outcome. Machine Learning has started to reshape how we live, everything around us nowadays has a touch of ML in it, so has mobile applications. Google has re-presented Firebase with the help of some prepared to utilize machine learning models like-Text recognition, Face detection, Image labeling, and more and backing to actualize custom models utilizing Tensor Flow Lite for iOS , Android and Web applications.

V. CONCLUSION

It was a magnificent and learning knowledge for us while taking a shot at this venture. This venture took us through the different periods of task improvement and gave us a genuine knowledge into the universe of computer science. The delight of working and the rush include while handling the different issues and difficulties gave us a vibe of being in developer industry. From the very begging of this project, we tried our best to develop this application the way we dreamed of. We are hopeful about our work. It would be satisfactory if our user gets benefited using our application.

Our project can be enhanced with more features and contents to help the user. We have some planned about adding some new features to be implemented and also redesigning all the current features with more functionality in future work to make the system more useful.

VI. REFERENCES

[1] Google Lens, available at << https://en.wikipedia.org/wiki/Google_Lens>>, last accessed on 05-07-19

[2] Android Platform, available at:<<https://www.hokuapps.com/blogs/5-reasons-to-choose-android-application-development-over-other-platforms/>>, last accessed on 04-07-19

[3] IDE (Android Studio), available at:
<<https://searchmobilecomputing.techtarget.com/definition/Android-Studio>>, last accessed on 06-07-19

[4] Firebase, available at <<https://firebase.google.com/>> last accessed on 04-07-19

[5] Machine Learning, available at:
<<https://www.quora.com/What-is-machine-learning-4>> last accessed on 05-07-19

[6] Machine Learning Kit, available at <<https://www.infoworld.com/article/3270731/ml-kit-google-brings-machine-learning-apis-to-mobile-developers.html>> last accessed on 05-07-19

[7] Rapid Application Development Method, available at <<https://blog.capterra.com/what- is-rapid-application-development/>> last accessed on 05-07-19