2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS 2019)

Image and Face Recognition using CV Lens Machine Learning Android Application

Dr.C.N.Vanitha
Department of Computer Technology
Kongu Engineering College
Perundurai,TamilNadu
vanithacn.ctug@kongu.ac.in

V.N.Jeevaa
Department of Computer Technology
Kongu Engineering College
Perundurai,TamilNadu
jeevaanallasamy@gmail.com

S.Prajith Shriman
Department of Computer Technology
Kongu Engineering College
Perundurai,TamilNadu
prajithshrimans@gmail.com

***Abstract* - This project deals with the camera application that can process the image to various processes. The impact of this project is made by the existing system where Google Lens process the image into text, translate, navigate and search about the text. The image is processed based on two major modules TEXT and IMAGE. TEXT module contains OCR, translation, maps, related Images. IMAGE module contains face emotions and image recognition. Where face emotion detects a face and reveals one's reactions such as happiness or sadness, etc. Whereas the image recognition module processes the entire object within an image and identifies each object in text form. This project is being developed using Android Studio with the Cognitive Services from Microsoft and Google API services.**

***Keywords* — text, image, maps, translation, face, detect, emotions, services.**

I. INTRODUCTION

Nowadays People travel across different places within and outside their country for their official purpose or it may be an excursion, most of them are not familiar with the language used within that place, recognizing the real-life object which mentioned in the other languages, routes they need to go. In order to solve this problem, the app is created based on a Camera application. At first, the user captures an image containing text or objects. If it contains the text is extracted using Machine Learning techniques and is translated into user chosen language, recognition of objects with relevant images, routes between source and destination. In addition to it, if the captured image contains objects like pen, hat, and mobile that are identified in a text format. If the object contains the face of a human then the emotions produced from the face like happiness or sadness are also identified. By integrating all these modules, it consumes less memory space than using separate apps for each purpose in the meanwhile it reduces the size of the app by satisfying the need of the people.

II. LITERARTURE SURVEY

Google Lens is based on recognition of image developed by Google in 2017. Using visual analysis bring up relevant information. With the help of phone's camera, Google Lens detects the object, text and show appropriate search results. This feature is merged with the Google Photos and Google Assistant. The application is similar to the previous app Google Goggles. Lens uses advanced deep learning routines techniques, used in other apps such as Samsung's Bixby Vision and *Image Analysis Toolset*; is used to identify the location, Optical Character Recognition (OCR) and artificial neural networks are used. Google launched its product Google Lens on October 4, 2017 with the in-built camera application into the Google Pixel Phones. A sneak peak of Google Lens had merged into the Google Photos app for Pixel phones. On March 5, 2018, Google released Google Lens embed into Google Photos for non-Pixel phones. On March 15, 2018 the Google Photos was launched to iOS. On May 2018, the Google Lens was made merged within Google Assistant to One Plus devices, and it is being integrated into in-built camera applications of Other Android phones. In the Month of June 2018, the First stable version of Google Lens app was available on Google Play store. Only limited phones has support, but it is not clear which Phones are not supported. It requires a minimum Android version Marshmallow (6.0) or newer.

III. TEXT MODULE

A. OCR

The technique to recognize the printed or written documents using a computer device is known as OCR. The process of OCR is to scan the photo for the text present in an image. It can be used for scanning for each character. In deep each characters are decoded into character codes, like as ASCII. ASCII represent the each character values in the form of binary codes. In OCR processing, the image are analyzed for bright and dark shades within the image. The extracted characters are decoded into an ASCII code. OCR uses predefined libraries to digitize and preserve their holdings. In general OCR is used for processing the Bank checks, credit card slips and magazines and letters are stored. In this project, the image containing text is processed and the characters are extracted. The Characters are listed out in a text format.

978-1-5386-9533-3/19/$31.00 ©2019 IEEE

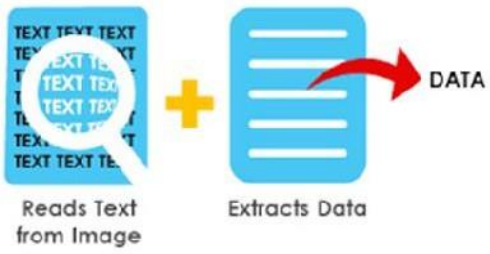


Figure 1: Extracting text from image

In Figure 1 if Image contains Text the OCR will extract the text from the image.

B. Translation

Translation is the process in which converts text from one language to other languages with the equivalent meaning. Translation can be done only when the text syncs with any language. It is necessary to introduce the language needed to be translated in the form of words, sentences, or any other form into the target-language. On the other side, the error words have relevant imported to translate from source-language words that have the enriched target languages. The Translators and early translators helped shaping the languages. Moreover, the development of the Internet has increased a market worldwide for translation services. In this project, the user selects the set of information to translate from the extracted information and choose the

Figure 2:Translating from source language to targeted language

targeted language thus produce the expected results.

In figure 2 the text contains other language contents are converted into User chosen languages.

C. Maps

The process of collecting the data and designed in to virtual image is known as Digital mapping. The main purpose of this technology is that it show maps, with an higher accuracy with representations of a particular area, detailing major road arteries and other points of interest. The Map also gives the calculation of distances from one place to another. The main use of the Digital maps is with the Global Positioning System, or GPS satellite network, used in standard automotive navigation systems.The GPS collects the data from any four satellites that orbiting the earth and it calculates the position in the form of three dimension representation. The GPS receives and it utilizes the satellite data and receives position to provide GPS coordinates such as latitudes and longitudes with appropriate direction from GPS satellites. The output accuracy ranges between 10-20 meters approximately of the actual location. The beginning point is entered via GPS coordinates, and the ending point will be given input by the user, and then the data will be entered into the digital mapping software. The mapping software outputs will be in the visual representation of the route.

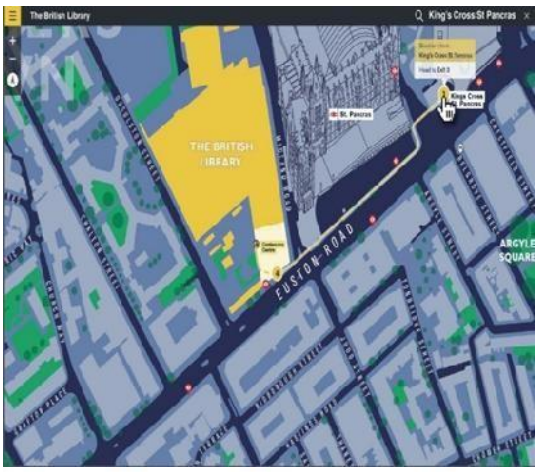


Figure 3: Source and Destination in Map

In figure 3 if the image contains the Location based contents can be viewed in maps

D. Image Search

Image search is a technique to search or find the relevant pictures. For search an image, a user will be providing the query or search terms like keyword, image or click on some image, and the system will return images similar to the query. The criteria used for the search could be meta tags, color distribution in images, region/shape attributes, etc. Image meta search – searching metadata of a keyword or text associated with the image. Content-based image retrieval (CBIR) – retrieves the image by the computer vision. Instead of concentrating in textual description, CBIR concentrates at retrieves images based on similarities in their contents such as textures, colors, shapes, etc., to a user entered query user-specified image features. List of CBIR Engines which search for images based on image visual content such as color, texture, shape or object, etc., Image collection exploration is the process of searching the images based on the use of novel exploration paradigms. An image retrieval system is for browsing, searching and retrieval of images from a large database of digital images. More traditional and common methods of image retrieval utilize some methods of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. The increase in social web applications have inspired the development of several web-based image annotation tools.

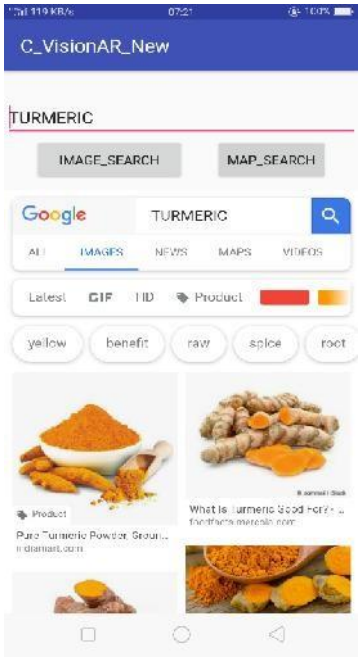


Figure 4: Retrieving Images from the Text

IV. IMAGE MODULE

A. Face Emotions

Emotion recognition is the process of identifying human emotion, most usually from facial expressions. Emotional expressions are in the theory, based on psychology. These expressions are in people while talking noticeably these verbal and nonverbal behaviors are seen common in people, which communicate an emotional or affective. The expression may be facial movements such as smiling or bliss, or actions like laughing or crying or angry or sad or happy or thankful. Emotional expression can occur with or without acknowledgment. Probably, individual peoples have the conscious control of their emotional expressions, in order to express emotion they need not have conscious awareness of their emotional state. From the past years, researchers have found different models for explaining emotion and emotional expression. All philosophers in emotion are strongly supportive that all normal, operational and humans express emotions with their voices, faces, and bodies. While the expressions of loving feelings are shaped by traditional and public factors.

Figure 5: Results of face emotions with numeric values

In figure 5 if the user snaps the human faces, the face emotion can be calculated using facial expressions.

B. Object Recognition

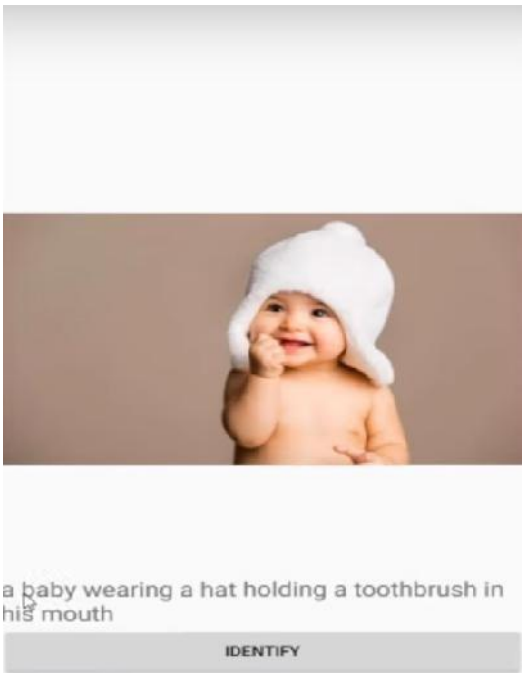
Image Recognition the taking out of important information from images, generally from digital images and by means of digital image processing practices. Image analysis can be done as identifying a person from their face. The human visual cortex is a best image analysis apparatus, especially for extracting higher-level information, and for many applications including medicine, security, and remote sensing human analysts still cannot be replaced by a computer machine. Many significant image analysis tools such as edge detectors and neural networks are inspired by human visual perception models.

Figure 6: Objects are identified using Image Recognition

In figure 6 if the Image contains any objects that are identified and formed as words.

CONCLUSION AND FUTURE ENHANCEMENT

The project "COMPUTER VISION LENS" has been designed as per the specification. The project is very simple



and easy understanding for users. Through this Android Application, the user feels more feasible and resolvable at any time in exploring new things. In order to use navigation by routes via with vehicle facility, which may leads to consume more time. It also provides easier solution. Provide flexible platform between user and the society. The new system eliminates the difficulties in the existing system. It is

developed user-friendly manner. The system is very fast according to the user wish that can be viewed or retaken at any level the project successfully verified which is specified in the abstract and system analysis concept.

In future enhancements, the payment module will be implemented to transfer the money without any external application. It includes with data transfer between two users by using two methods, Local Wi-Fi and Mobile data. Sharing the text by means of converting to document format or to a PDF (Portable Document File) format.AR stickers will be used to cover the objects and snap the pictures and AR measurements will be implemented for measuring the objects using certain conversion methods. Voice Access method will be invoked throughout the application and while reading the text using OCR the content will read by using text to speech method with pause and play content. The last method is, focusing the Objects and it will display the attributes of the object (for e.g. If a shirt is the object, then it will show the size of the shirt and it can show the nearest shopping store).

REFERENCES

[1] Marian Schluter, Carsten Niebuhr, Jan Lehr, Jorg Kruger, "Vision Based Identification Service For Remanufacturing Sorting", Elsevier, Procedia Manufacturing Volume 21, pp. 384 – 391, 2018.

[2] Barret Zoph, Vijay Vasudevan, Jonathan Shlens, Quoc V.Le, "Learning Transferable Architectures For Scalable Image Recognition", The IEEE Conference on Computer Vision and Pattern Recognition, pp. 8697-8710, 2018.

[3] C. V. Arulkumar, P. Vivekanandan, "Multi-feature based automatic face identification on kernel eigen spaces (KES) underunstable lighting conditions", Advanced Computing and Communication Systems 2015 International Conference on. IEEE, 2015.

[4] V. Arulkumar. "An Intelligent Technique for Uniquely Recognising Face and Finger Image Using Learning Vector Quantisation (LVQ)-based Template Key Generation," International Journal of Biomedical Engineering and Technology 26, no. 3/4 (February 2, 2018): 237-49.

[5] Yueqi Duan, Jiwen Lu, Jianjiang Feng, Jie Zhou, "Context-Aware Local Binary Feature Learning For Face Recognition", IEEE Transaction on Pattern Analysis and Machine Intelligence, Volume 40, Issue 5, pp. 1139-1153, 2017.

[6] https://developer.android.com/docs/

[7] https://firebase.google.com/docs/ml-kit/

[8] C.V. Arulkumar, G. Selvayinayagam and J. Vasuki, "Enhancement in face recognition using PFS using Matlab," International Journal of Computer Science & Management Research, vol. 1(1), pp. 282-288, 2012

[9] https://developers.google.com/maps/documentation/

[10] H. Anandakumar and K. Umamaheswari, "A bio-inspired swarm intelligence technique for social aware cognitive radio handovers," Computers & Electrical Engineering, vol. 71, pp. 925–937, Oct. 2018.

[11] https://cloud.google.com/translate/docs/