

# Autonomous Restaurant Serving System using Image Detection and Voiceprint Detection in Android Application

Soham Dasgupta, Madhurupa Samaddar, Yogalakshmi CN, K Vijayan and Subhiksha

**Abstract**—In recent times hotel management has been pretty important when it comes to business meets and food chains. Currently in India multiple food chains are running out of which many food chains are running internationally too. One of the most important aspects of these chains is the management of customers and customer satisfaction. The first part of this project aims at intelligently predicting the choices and the personal details of the customer. This prediction is done through voice recognition where the voiceprint of a person is detected to access the database of the person which consists of his/her previous orders and special changes in the menu. This helps in knowing about the allergic reactions of a person to some specific food items and a customizable menu can be provided to the person. This would greatly help in management of customer databases and customer satisfaction as well as maintenance of the privacy of the person. The second part of the project deals with the food item detection algorithm using image processing so that veg and non veg food items do not get mixed up and the correct table receives the correct order.

**Index Terms**—Image Detection, CNN, Firebase, Voiceprint Analysis, Data Handling, Database Management and Android Development.

## I. INTRODUCTION

THE primary objective of this project is voiceprint recognition of the person so that the customer is not asked for additional personal details and also the database of the customer can be easily recognized using the voiceprint. This database can be used to customize the order for the customer making it easier for the restaurant or hotel to manage the customers [1-6]. The Secondary objective of the project is the detection of food items using image processing so that the correct order is always sent to the correct table. Moreover Vegan meals thus will never be mistaken for non veg meals [7]. The rest of the paper is organized as follows. The system model and android development are described in Section II and Section III respectively. The results and inferences is explained in Section IV. At last, Section V concludes the paper with conclusion of the work.

Soham Dasgupta, Madhurupa, Samddar, Yogalakshmi CN, Dr.K Vijayan and Subhiksha are with the Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Chennai. (email: [soham171097@gmail.com](mailto:soham171097@gmail.com), [madhurupa97@gmail.com](mailto:madhurupa97@gmail.com), [kalayoga2@gmail.com](mailto:kalayoga2@gmail.com), [vijayank@srmist.edu.in](mailto:vijayank@srmist.edu.in), [n.subhiksha@hotmail.com](mailto:n.subhiksha@hotmail.com))

## II. SYSTEM MODEL

### A. MFCC Methodology

The Mel Frequency Cepstral Coefficient is extracted from the voice signal is shown in Fig. 1. The output is in the form of a vector. This vector is compared to the test voice signal for detection. The feature extractions are done for the voice signal [8-10]. The features taken into account are pitch, frequency, loudness, bass, etc.

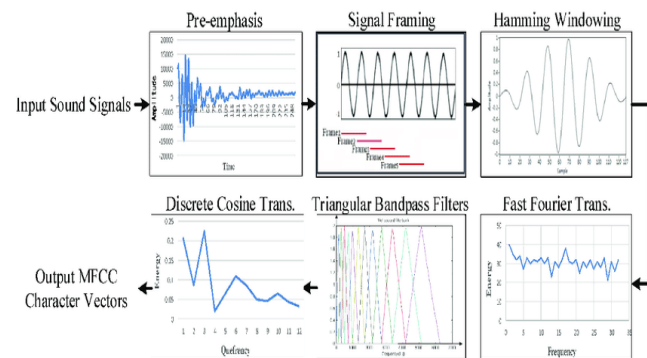


Fig. 1. MFCC Methodology for speech detection

### B. Image and Speech Detection Methodology

The image processing includes Adaptive Histogram Equalization for Contrast features is shown in Fig. 2. This filter enhances the outlines of the food images making it easier for the shapes to be detected. The voice detection is done through the voice database that is generated by multiple times storing multiple voice datasets [11]. The MFCC and the GMM are used for the detection of the voice data through feature extraction. The features vector is the stored. The vector is compared with the test voice vector.

### C. Voice Database Management

- MATLAB Audio Database Toolbox

This toolbox gives better access. Also we can filter auditory databases for eg:- TIMIT and YOHO by using their metadata. For accessing databases using customized coding and manual filtering, this database toolbox can easily replace it. This toolbox helps us utilize time of the database structure and makes it easier to focus on the code's algorithmic aspects.

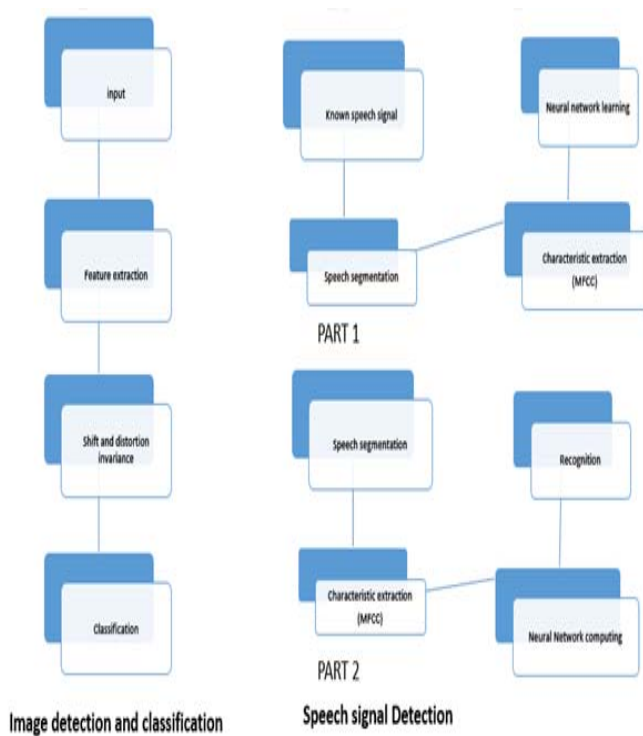


Fig. 2. Image and speech detection methodology

- TIMIT/NTIMIT supports for searching guidelines: phoneme ,word, sentence etc.

This part deals with data retrieving and data entering .As there are many food chains spread across many places around the world .One of the most important aspects of these chains is the management of customers and customer satisfaction. This is the first part of this project which aims at intelligently predicting the choices and the personal details of the customer [12,13].

### III. ANDROID DEVELOPMENT AND IMAGE DETECTION

#### D. Database

This part deals with data retrieving and data entering is shown in Fig. 3 and Fig. 4 respectively. As there are many food chains spread across many places around the world .One of the most important aspects of these chains is the management of customers and customer satisfaction. This is the first part of this project which aims at intelligently predicting the choices and the personal details of the customer. This helps in knowing about the allergic reactions of a person to some specific food items and a customizable menu can be provided to the person. This would greatly help in management of customer databases and customer satisfaction. This has been done by making an app so that it is easier to keep the records of all the customers. This is done by the following two ways:

#### i) Data Retrieval

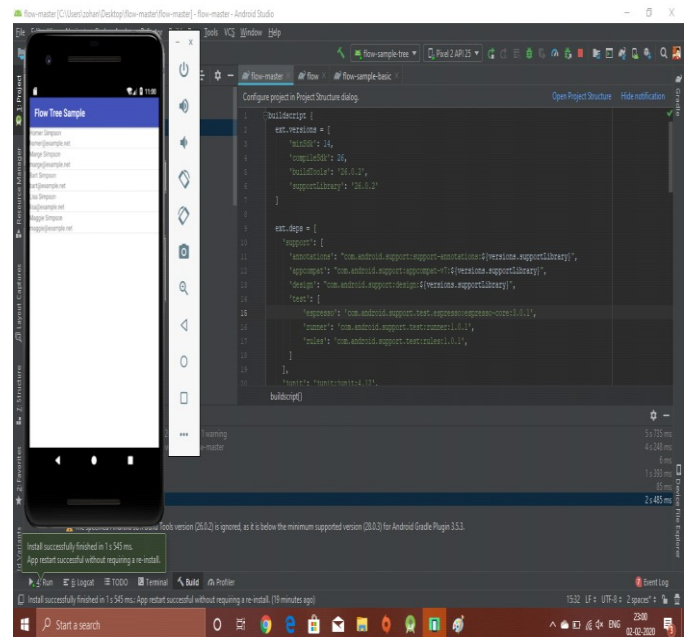


Fig. 3. Database Retrieval in flow template

Above we have shown some retrieved data using the emulator .The following entries have been retrieved and are shown as a part of the complete database. The data is procured through voice commands and each data is segregated using voice recognition.

#### ii) Data Entry

This module of the application contains the database entry portion for a customer. Here all his details related to his choices and allergies are entered. If the database of a person already exists, it can be updated in this part of the application.

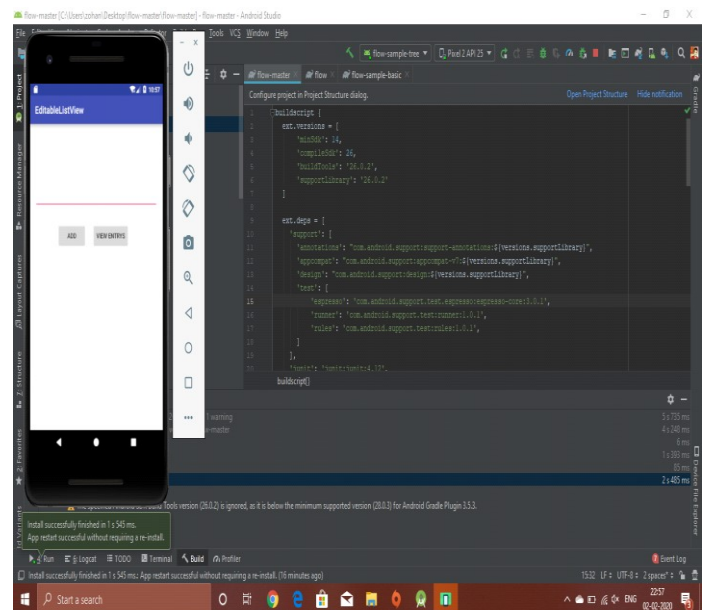


Fig. 4. New Entries in the database

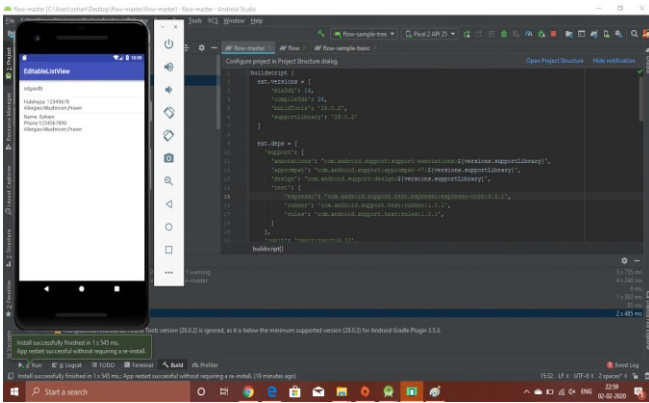


Fig. 5. Updating existing entries

In this step the user is given an interface to add the required information about a customer like their name, personal data like phone no, address and specific instructions like allergies and special likings etc which is shown in Fig. 5.

### E. Image Detection

Object recognition is a technique used for identifying objects in images or videos using computer vision so that we can classify events/objects in order to comprehend a real-time situation. In the domain of computer vision, we use programming and design algorithms to apprehend what is there in the pictures. One of the main objectives of deep learning and machine learning algorithms is object recognition. Normally when we view a photograph or see a video, we can easily discern humans, visual features and items. Our objective here is to train the computer to do what we can naturally do: to understand how to interpret the contents of an image. The connection and link are the most important elements for comprehending the scene.

- To use object recognition to differentiate between, unlike objects.

Two of the most useful and important jobs done by computer vision are detecting and recognizing objects. Object detection discovers the appearance of an object. It also can determine the area and the prominent locations in the image. The object represents the training database which recognizes the object class using object recognition. Object detection is generally given more preference than object recognition. It can also be seen as a two-step object recognition, where one step belongs to the object class and the other step belongs to the non-object class. Object detection can be additionally put into two types such as soft detection and hard detection out of which the former only determines the appearance of an entity/object and the latter determines both the appearance and site of the entity/object.

- Comparison between recognizing and detecting an object.

Both the techniques recognizing and detecting of objects are very alike in recognizing objects, but they are very different in their own ways of executing actions. Object detection is the method of recognizing objects/entity in

images. Object detection is actually a subgroup of object recognition according to an instance of deep learning which comes under data science. Under deep learning the object is just not spotted but also discovered in an image. This helps many objects/entities to be spotted and discovered inside the same image.

This project primarily focuses on detection of colored dots on plates as shown in Fig. 6 to Fig. 10 since the color coded dots can be detected to refer accordingly to veg or non veg platters. The veg and the non veg plates can be segregated using green and red dots similarly to the ones that are used officially as food standards. The following algorithm is used for detection of the colored dots from multicolored images.

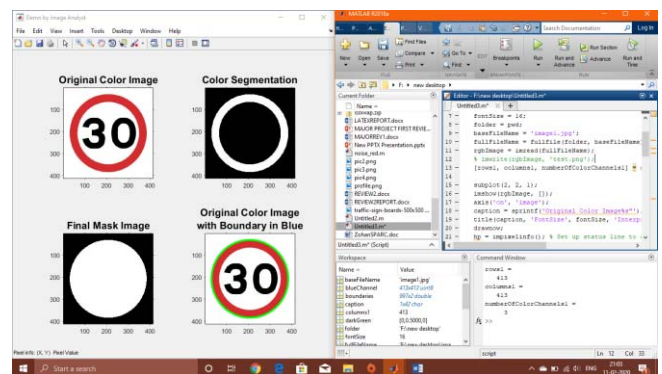


Fig. 6. The largest red circle is recognized

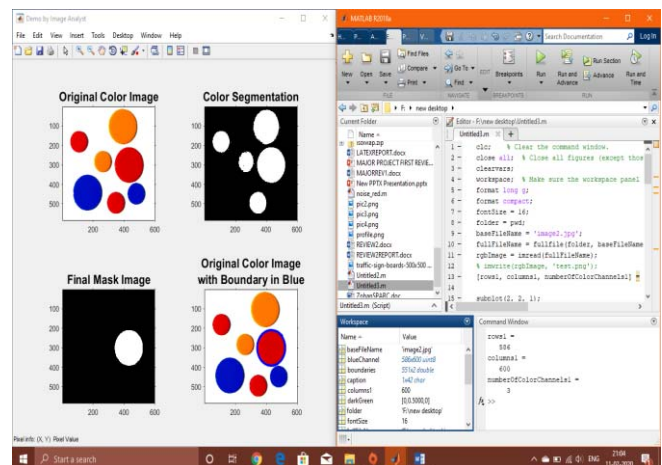


Fig. 7. The red dot with maximum red intensity and size is recognized

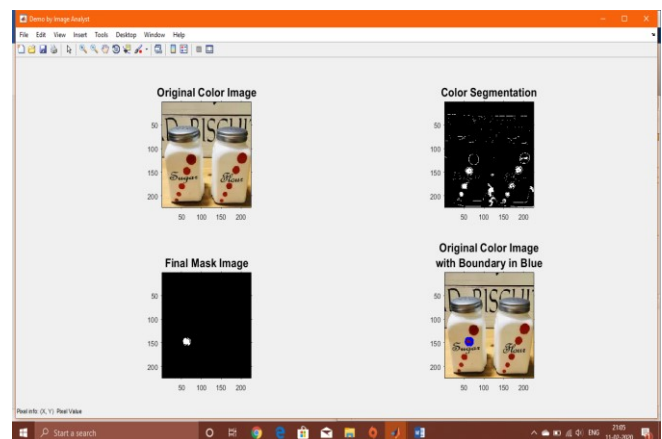


Fig. 8. Red dot with maximum intensity is recognized



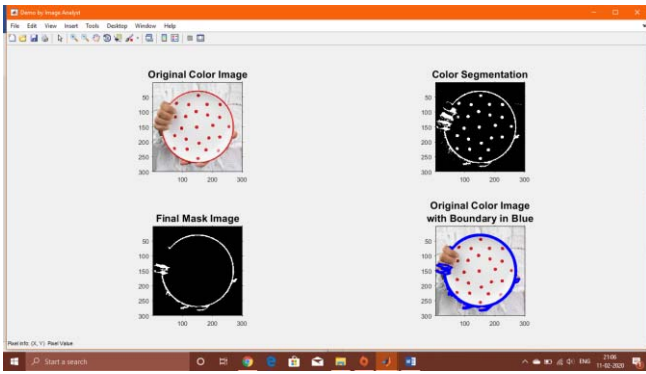


Fig. 9. The largest red circle is recognized

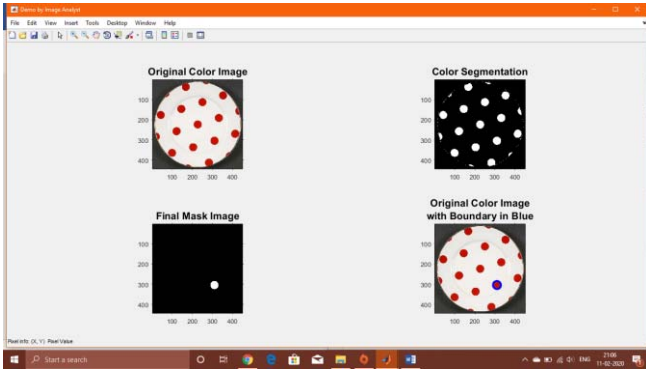


Fig. 10. The inner red dot is recognized.

#### IV. RESULTS AND INFERENCES

##### A. Android app development

The app is connected with online google firebase console which acts as the common server for all the devices using the application currently for database management is shown in Fig. 11. The android app contains the voice detection so that the customers can simply go for voice recognition to unlock their database of details and preferences in the respective food outlets.

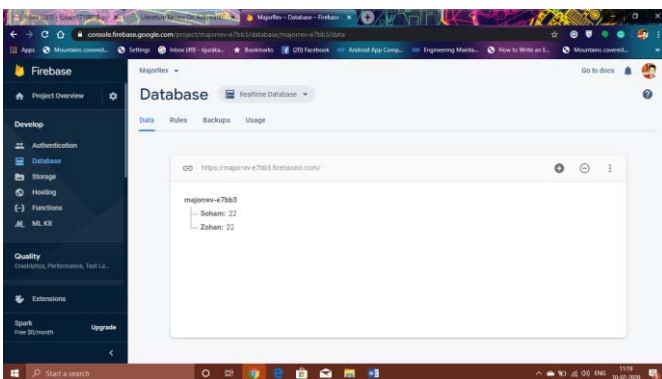


Fig. 11. Google firebase console for the research work

##### B. Image detection for food classification

This algorithm shown in the paper effectively segregates the food into veg and non veg depending upon the dots. This method can be very helpful in automated restaurants, where minimum human interaction is essential. This automatic detection technique would help in minimizing the cost of automation and maximizing the output as well.

TABLE I  
ACCURACY VALUES

Serial	element	With 100 images	With 200 images
1	accuracy	97%	95%

##### C. Voice detection

The Voice detection has been done in an open CV platform and has been integrated with android studios for immediate database access of the user. The speech recognizer module provided by the android studios helps in retrieval of speech data of the user and thus recognition of the speech data is done.

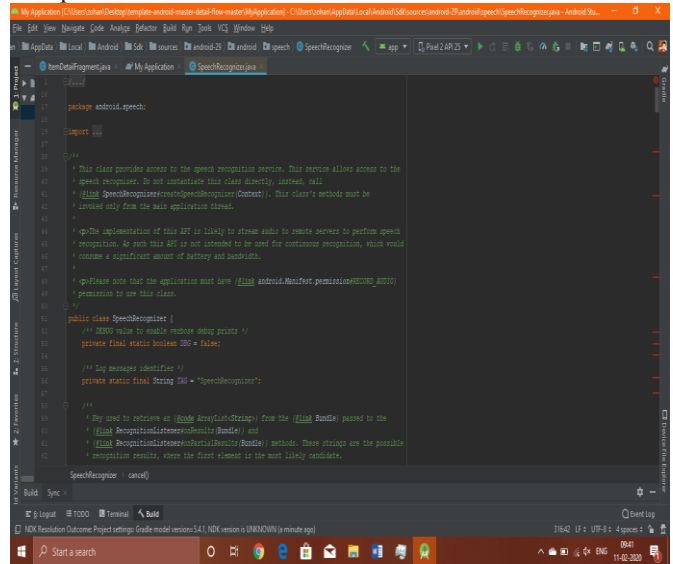


Fig. 12. SpeechRecognizer class

This speech recognizer class shown in Fig. 12 is used for the speech recognition in Android Studios. This helps the user in segregating the customers based on different voiceprints.

For countries like India, Spending huge amounts of money on automated restaurants is neither feasible nor profitable. This project aims and provides with a solution of a simpler and efficient customer management system, keeping in mind the privacy as well as the satisfaction of the customer. This project can essentially help restaurants in managing crowds with higher efficiency and lower error rates.

##### D. Inference

The speech recognizer recognizes the voice print based on the mel frequency value. The error rate from the speech recognizer is pretty less which makes it effective even in commercial purposes. The image detection algorithm when trained with 500 images gives an error rate of 8-9% which can be effectively used in commercial purposes too. The example of 100 images and 200 images are reported in Table I. The firebase database services are used for free for the project by accessing only the realtime database systems. The paid version of the Firebase database management can be used in further stages for handling large amounts of data. The cost effectiveness of the complete project makes this project pretty feasible as most of the platforms used in the project are open source and the databases used for training purpose are open source too.

## V. CONCLUSION

In this paper the ways of automating restaurant serving system and database access of customers is proposed. However there will always be some sources of error if we consider the factor of human error too. However the color coding procedure matched with the food image detection procedure shall lower the chances of error a lot. The cost of the whole application will be much less only to consider the cost fixed by the developer for the project and the license of the firebase servers. This cost is much less compared to the fully automated restaurants that are operating currently. This low cost helps small restaurants to operate the whole system only by installing the application in their systems. The project aims at effectively lowering the cost of the entire system and as well as make the whole thing as much smart as possible.

## ACKNOWLEDGMENT

This research work would not have been possible without the guidance of Dr.K. Vijayan. With the help of his knowledge and skills, a very innovative insight is added to the research work. We would like to thank Dr. K Vijayan for his guidance and help with the research work.

## REFERENCES

- [1] S. Sasano, X. Han and Y. Chen, "Food recognition by combined bags of color features and texture features," 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Datong, 2016, pp. 815-819.
- [2] M. Wazumi, X. Han, D. Ai and Y. Chen, "Auto-recognition of food images using SPIN feature for Food-Log system," 2011 6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT), Seogwipo, 2011, pp. 874-877.
- [3] W. Zhang, D. Zhao, W. Gong, Z. Li, Q. Lu and S. Yang, "Food Image Recognition with Convolutional Neural Networks," 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), Beijing, pp. 690-693.
- [4] G. K. Berdibaeva, O. N. Bodin, V. V. Kozlov, D. I. Nefed'ev, K. A. Ozhikenov and Y. A. Pizhonkov, "Pre-processing voice signals for voice recognition systems," 2017 18th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM), Erlagol, 2017, pp. 242-245.
- [5] J. Meng, J. Zhang and H. Zhao, "Overview of the Speech Recognition Technology," 2012 Fourth International Conference on Computational and Information Sciences, Chongqing, 2012.
- [6] Y. Yu, "Research on Speech Recognition Technology and Its Application," 2012 International Conference on Computer Science and Electronics Engineering, Hangzhou, 2012, pp. 306-309.
- [7] G. Frewat, C. Baroud, R. Sammour, A. Kassem and M. Hamad, "Android voice recognition application with multi speaker feature," 2016 18th Mediterranean Electrotechnical Conference (MELECON), Lemesos, 2016, pp. 1-5.
- [8] S. Primorac and M. Russo, "Android application for sending SMS messages with speech recognition interface," 2012 Proceedings of the 35th Int Convention MIPRO, Opatija, 2012, pp. 1763-1767.
- [9] M.P.Nirmala, C.Muniyakala, V.Nagarajan (FEB 2014), "Secured Data Embedding in Encrypted Image Using RDH Algorithm for Multimedia Communication", International Journal of Electronic Communications Engineering Advanced Research, Vol.2, SP-1, Feb.2014, pp. 120-124
- [10] V. H. Goh and Y. Wen Hau, "Android-based Mobile Application for Home-based Electrocardiogram Monitoring Device with Google Technology and Bluetooth Wireless Communication," 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Sarawak, Malaysia, 2018, pp. 205-210.
- [11] A. Alsalemi *et al.*, "Real-Time Communication Network Using Firebase Cloud IoT Platform for ECMO Simulation," 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter, 2017, pp. 178-182.
- [12] M. T. Islam, B. M. N. Karim Siddique, S. Rahman and T. Jabid, "Image Recognition with Deep Learning," 2018 Int Conf on Intelligent Info and Biomedical Sci (ICIIBMS), 2018, pp. 106-110.
- [13] King-Sun Fu and Rosenfeld, "Pattern Recognition and Image Processing," in *IEEE Transactions on Computers*, vol. C-25 Dec. 1976

