DESIGN OF LOW-COST OBJECT IDENTIFICATION MODULE FOR CULINARY APPLICATIONS

Pranav Vijay Chakilam, Revanth B, Vyshnavi Muppirala, A. Anilet Bala, Vivek Maik

I. ABSTRACT

Using up the contents of your fridge often requires touch creativity, which is typically hard to come back by in a very busy world. This aim is to style a model that may scan the things in your fridge and recommend recipes that supported what you have got, even taking your preferences and dietary restrictions under consideration. Computers are commencing to process globe objects without the requirement for codes or human intervention in their own language. This project aims to create a module that uses object recognition to detect vegetables and fruits and display recipes that include those foods. This concept uses a camera that is installed in a refrigerator to show it into a sensible one. Using the camera, the system identifies objects that are placed inside using complex object recognition algorithms. If a vegetable or fruit is placed inside, the system will identify it and displays its name on the screen. It creates as an indication of the chances of object recognition. The platform enables users to put any number of various vegetables into the refrigerator. The system then checks a list of recipes that contain the ingredients inside the refrigerator and filters those that feature the vegetables available. Together with this, we also aim to watch the contents of the refrigerator and find the freshness and age using various sensors and also provide inventory and warnings when they are on the verge of completion.

II. INTRODUCTION

When we are shown a picture, the objects that are present inside the image are immediately recognized. At the same time, it requires heaps of time as well as preparatory training data for enabling machines with the same recognition capabilities [1], [2]. With recent advances in the field of Deep Learning (DL), hardware in terms of processing power and the advent of Big Data, the Computer Vision domain has become a significantly simpler and increasingly instinctive. Object Detection and its applications have become pervasive across a gamut of fields.From assisting self-driving cars to drive in the presence of traffic to spotting unruly incidents and rough conduct in packed places, Analyzing and constructing scouting reports to help sports teams to making sure that optimum internal control of manufacturing parts is guaranteed, its applications have become ubiquitous. Object Detection can be realized in the most efficacious manner with the help of DL.

A pair of methodologies by which deep learning can surpass existing object detection techniques are listed next. Rather than acquiring patches from the underlying image, the image is skilled in a neural network to scale back the size. The neural network is accustomed to suggest selective patches of the image[3], [4], [5], [6]. Now rather than preparing multiple neural networks to tackle the different aspects of object detection, A single, deep neural network will be attempted that can understand and provide a solution to the entire problem. The upside of taking this approach is that each of the discrete modules of a neural network will aid in the optimization of the contrary parts of the identical neural network. This can facilitate the combined training of the entire deep model[7], [8]. Training a model from scratch is often cumbersome and consumes a lot of time and data, to ameliorate a situation like this Transfer Learning can be employed where we can complete the training process with fewer data points and lesser time. In transfer learning the understanding, patterns and learning attained from one task is utilized and applied in another task. The idea is based on the logic the two similar tasks that appear completely unrelated, might have identical underlying patterns based on which the model detects the target so, we make use of this training in our task directly, bypassing the repetition of the extensive training which was initially performed. The improvement in accuracy while decreasing training time and data by employing Transfer learning has been repeatedly, widely and reliably demonstrated now.

I. STATEMENT OF PROBLEM

There has been massive research in the field of Deep Learning in recent years along with efforts to use it in culinary applications. But there has been no practical implementation in refrigerators. A Low Cost module is required that can convert existing normal refrigerators into smart ones. An Autonomous Sensor array is required that can constantly monitor and notify users about the state of the refrigerators contents. A Light, Accurate and Low-Latency model is required that can perform object-identification and recommends recipes.

II. SYSTEM DESIGN

A. Software Specifications:

I. Object Identification Module:

a) We use Convolutional Neural Network for object identification by implementing YOLO architecture. Python code is written on Google Collab Notebook.

b) Libraries required: pandas, numpy, os, pickle, matplotlib and darkflow.

II. Deploying the model on Cloud:

a) Google AI platform - To train the object identification model.

b) Google Compute Engine - To deploy the trained model on the cloud.

c) Libraries and packages required: Google Cloud SDK, inotify tools. Google Cloud VM(Virtual Machine) Instance details- Debian Operating System, 3.75GB RAM, 10GB Storage Memory.

The fundamental representation of the system is shown in the block diagram in figure 1 and the flow chart is shown in the figure 2.

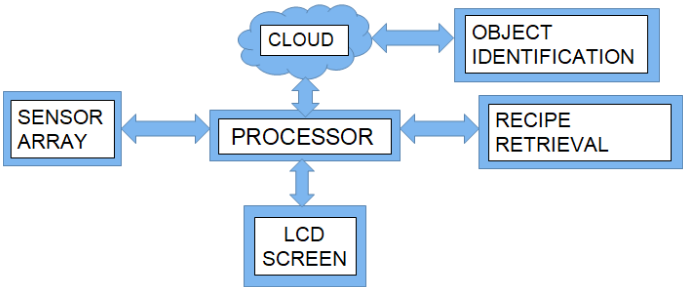

Fig. 1: Block Diagram

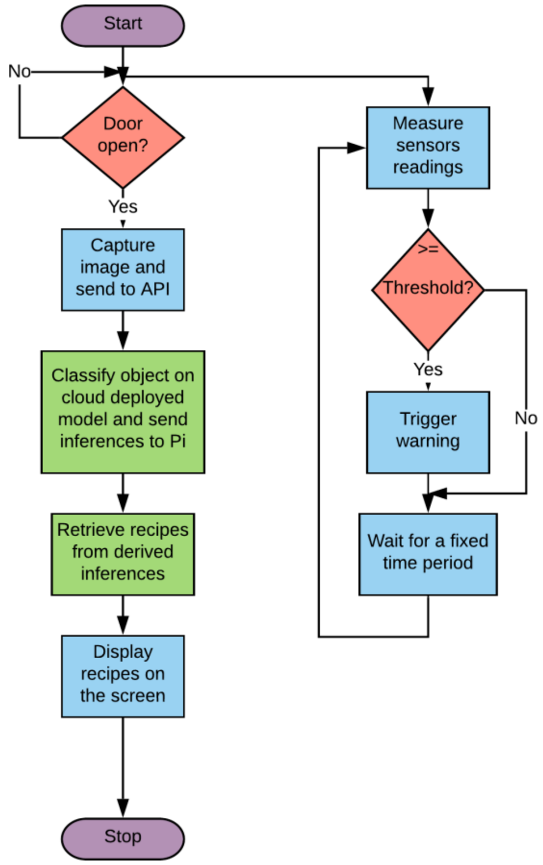

Fig. 2: Flowchart

B. YOLO Architecture

This variant of our model is an object detector because the name suggests it detects objects within images. It's different from the mobile net model in this it's ready to identify multiple objects within the identical image, as hostile a picture classifier, which only classifies opposed to

. Our network uses features from the whole image to predict each bounding box. The bounding box is an imaginary, 2-dimensional, rectangular box, drawn around an object as in figure 4, such it encompasses it. Here the prediction of every bounding box in the picture, covering all the classes is performed at the same time. This suggests that the reasoning done by the network is done in a global manner regarding the complete image as well as each object that the image consists. The Yolo algorithm allows real-time speeds and end-to-end training and does this without affecting the average precision and ensuring that it is kept high as in figure 3. The system dissects the input image in the form of an s × s grid as in figure 5 [9]. In case the middle of an object lies when the model thinks that there is no object inside the bounding region. Else, The Figure 3: YOLO architecture output score should ideally be calculated by the intersection over union (IOU) of the ground truth with the predicted box. Cell, irrespective of the amount of boxes O. Every bounding region outputs 5 predictions. They are, (x,y) which are the coordinates that represent the position of the bounding region's center with respect to the borders of the grid cell. (w,h) which stand for the width and height are anticipated with respect to the whole image. Also, every cell of the grid also estimates p probabilities which represent the conditional probability, i.e. given that an object is contained in the bounding box. This reflects the certainty with which the model thinks that the bounding box contains an object and that the accuracy which it attributes to the prediction that this object corresponds to a particular class[9].

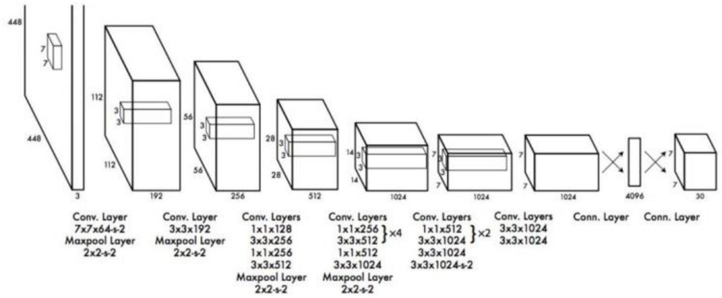

Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

Fig. 3: Architecture

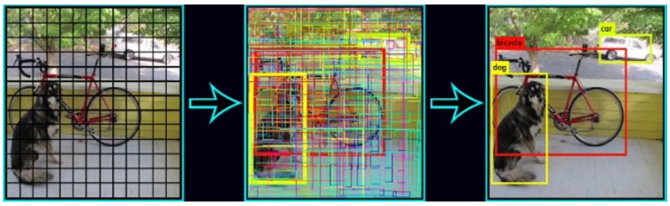

Fig. 4: YOLO

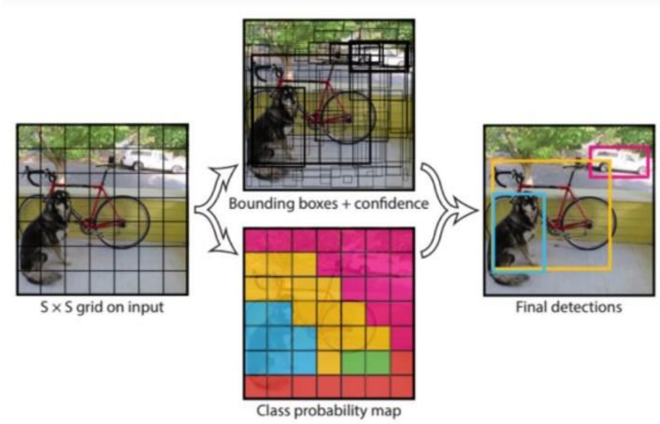

Figure 2: The Model. Our system models detection as a regression problem. It divides the image into an $S \times S$ grid and for each grid cell predicts B bounding boxes, confidence for those boxes, and C class probabilities. These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor.

Fig. 5: YOLO mapping

C. Hardware Description:

The major concern pertaining to food wastage is Food hygiene and safety. The food standard must be checked and kept safe from rotting and spoiling by the climatic components like light, temperature, moistness. In this manner, it is valuable to introduce quality checking gadgets at food stores and homes. This gadget keeps a watch on the natural factor that causes rot of the nourishment. Refrigeration and vacuum storage controls many other natural factors. The hardware kit (to be added to the refrigerator) consists of a camera, screen, a Raspberry Pi and a sensors array consisting of DHT 11 sensor, which is a humidity and temperature sensor , an MQ3 sensor, which is used to identify and detect the presence of ethanol and alcohol and a Nodemcu, which is a firmware that is used for open-source prototyping board design . In this project, The quality checking device keeps a watch on natural factors like alcohol, spoilage, temperature, light, and humidity[10]. The main processor board used is Node MCU- 8266. It is interfaced with sensors like MQ3 to detect food quality and spoilage. It is an IoT device that sends the sensor values to a database that is connected to a dashboard. The IoT dashboard is used to log and monitor the sensor's data. The dashboard used here is Cayenne. Because of the internet of things, the data can be logged from anywhere, anytime, and from any gadget. The position of these sensors are shown in figure 6.

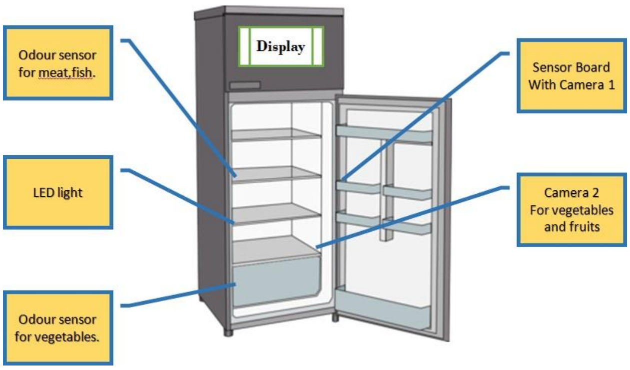

Fig. 6: Position of hardware components

D. Interpretation

Our system has Raspberry Pi (local system), camera(s), and sensors array inside the refrigerator. The screen is outside the fridge to have access to the data. Once it's powered through, it connects to the strongest wi-fi based on user input. The camera is placed to achieve a bird's-eye view of the vegetables cabinet in our model, but in real time refrigerators, we will have cameras in multiple positions to cover all possible locations inside the refrigerator. When the refrigerator door is opened, the lights turn on and LDR detects the light that triggers the camera (connected to Raspberry Pi). The camera captures the picture of the cabinet and sends it to the Pi as shown in figure 9. Our object recognition model (YOLO v2) is deployed on the cloud (Google Cloud Platform) for fast and efficient processing. Hence, the Pi uploads the captured image to the Cloud Compute virtual machine which contains our object detection model. Here, the compute machine performs complex computations on the uploaded picture by passing it through the model for object identification and it takes approximately 8 seconds to recognize the items. Once the recognition is done, the output is received by the local system (Raspberry Pi) in the form of a list as shown in the figure 10. This output is then used in the recipe retrieval process (on the Pi using an online browser) and the recipes for the objects detected are displayed on the screen as shown in the figure 10. This uploading and downloading of the image, and the recipe retrieval display consumes approximately 20 seconds bringing the total time of the complete process to nearly 30 seconds. The IoT device is installed inside the fridge. The sensors read the data continuously and transfer them via wi-fi. The DHT 11 is a temperature and humidity sensor as mention before. It is a digital sensor that detects temperature and humidity every 2 seconds. The sensor with a voltage supply of 3.5 to 5.0 v and the temperature ranges from 00 c to 500 c. The sensor operates on a 1 – wire protocol which is implemented on the firmware because of which the sensor cannot interface with digital pins. The 40 bytes data read by the sensor consists of temperature and humidity. The mq3 sensor detects gases like ethanol and alcohol. It has to be placed where conservative foods (fruits and vegetables) are kept[11][12]. The sensor keeps on detecting the concentration of ethanol. Once the concentration reaches the threshold level, the sensor signals to the analog pin. The prototype board has an inbuilt ADC. All the sensor values keep sending to the cayenne server. The cayenne server stores the data and displays it on the dashboard. The work flow is shown in the figure 7 below:

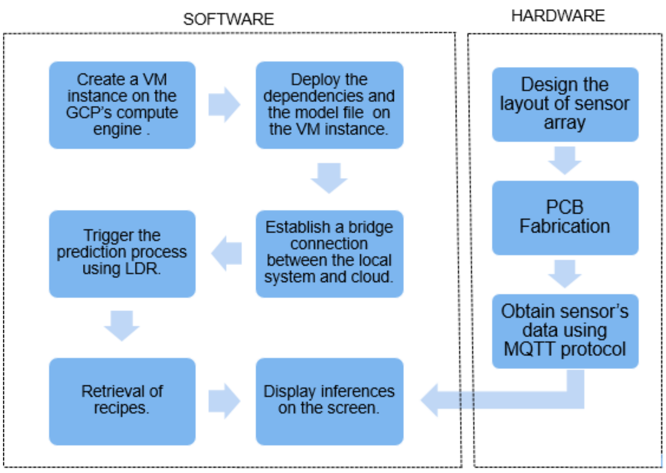

Fig. 7: Work flow

IMAGE	No. of Objects	Classes of Objects	confidence score
1	1	Banana	0.45
2	1	Apple	0.63
3	1	Orange	0.7
		Apple	0.5
4	2	Orange	0.6
		Banana	0.67
5	2	Apple	0.43
		Apple	0.49
		Banana	0.69
6	3	Orange	0.57

Fig. 8: Confidence score table

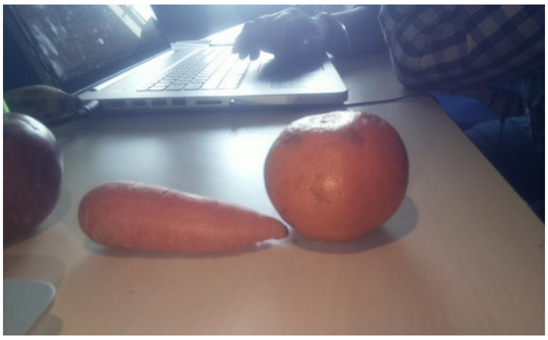

Fig. 9: Image captured

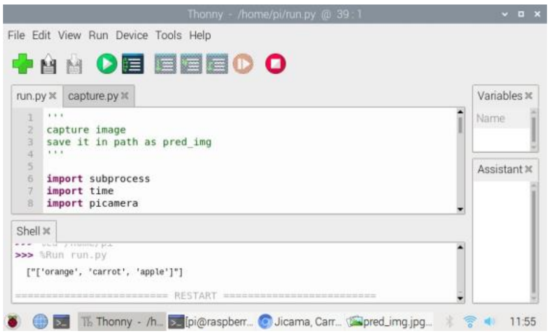

Fig. 10: Result obtained from the given input

V. RESULTS AND CONCLUSION

The model's efficiency was empirically observed to be 93.77 percent. To test the real-time performance of our model, multiple objects were placed in the same picture and tested on. The model was successfully able to identify all the images. The metric used is confidence score as output by the model. The objects were kept at a distance of 1m. As observed, the confidence score is independent of the number of classes present as shown in figure 8. This is because the model attempts to draw bounding boxes around objects that it identifies and predicts the confidence of that box alone. It was observed that even on occlusion (figure 9 - apple is occluded), the confidence score of the occluded object decreased considerable, yet the model was still able to make a fairly confident prediction (figure 10). The recipes that feature the objects detected are retrieved and displayed on the screen as shown in figure 11. The status of the contents of the refrigerator is also displayed on the screen as shown in figure 12. And when there is any rotten food, the odour sensor takes it's measurement to shoot up which sends a mail to notify the same as shown in figure 13. In a nutshell any regular refrigerator can be given this smart module that facilitates: one to receive recommendations on recipes from the available vegetables/ fruits, intimation on food expiry, and a display unit with advanced and user friendly UI.

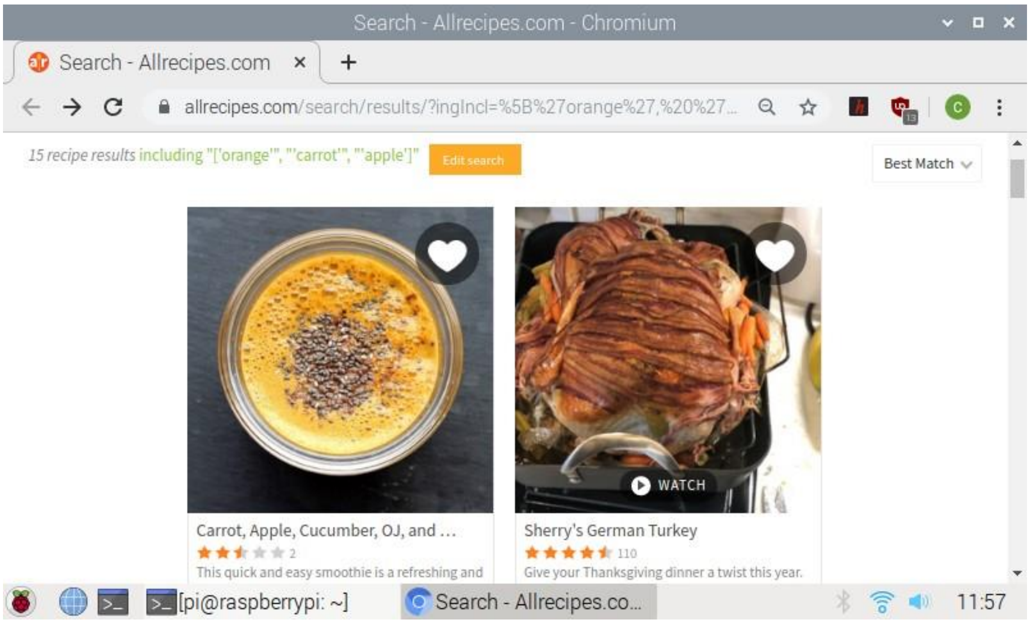

Fig. 11: Recipes displayed

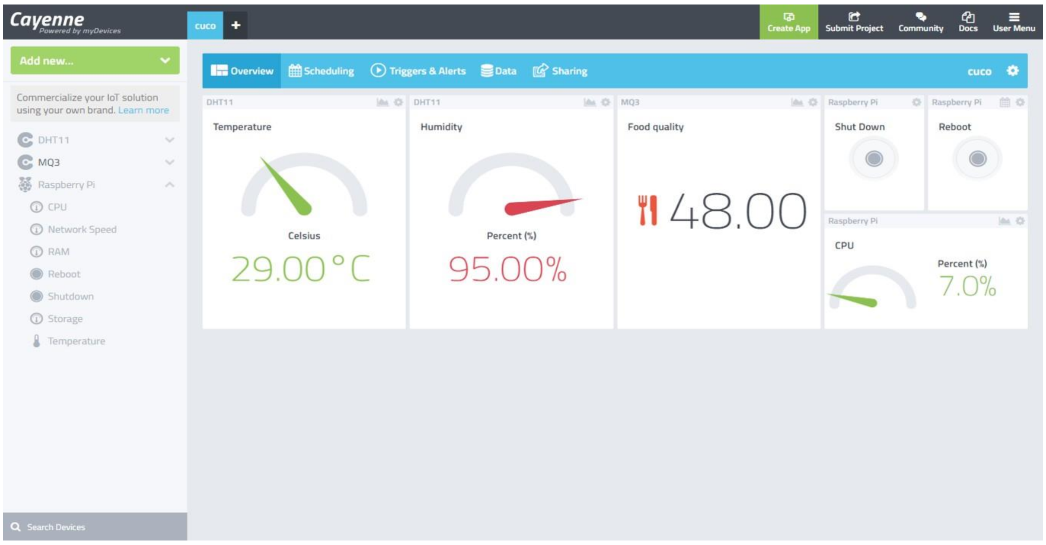

Fig. 12: Dashboard on the screen

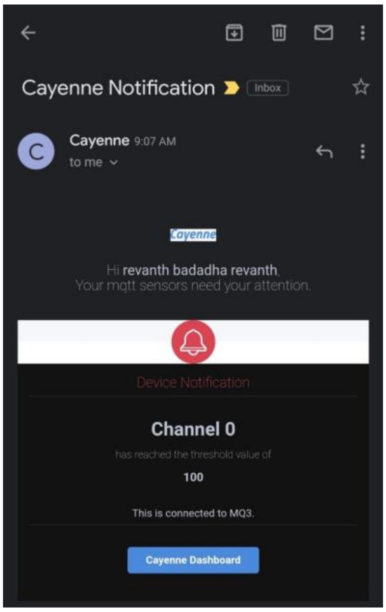

Fig. 13: Notification email

VI. REFERENCES

[1] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma,Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, And L. Fei-Fei, "Imagenet Large Scale Visual Recognition Challenge," In International Journal Of Computer Vision (Ijcv),2015.

[2] A. Krizhevsky, I. Sutskever, And G. Hinton, "Imagenet Classification With Deep Convolutional Neural Networks," In Neural Information Processing Systems (Nips), 2012.

[3] Shaoqing Ren, Kaiming He, Ross Girshick, And Jian Sun," Faster R-Cnn: Towards Real-Time Object Detection With Region Proposal Networks", Doi 10.1109/Tpami.2016.2577031, Ieee Transactions On Pattern Analysis And Machine Intelligence.

[4] Peng Bian1, Wanwan Li2,3, Yi Jin2,3* And Ruicong Zhi4,5 "Ensemble Feature Learning For Material Recognition With Convolutional Neural Networks" Bian Et Al. Eurasip Journal On Image And Video Processing (2018) 2018:64

[5] S. Ren, K. He, R. Girshick, And J. Sun, "Faster R-Cnn: Towards Real-Time Object Detection With Region Proposal Networks," In Neural Information Processing Systems (Nips), 2015.

[6] R. Girshick, "Fast R-Cnn," In IEEE International Conference On Computer Vision (Iccv), 2015.

[7] Ajeet Ram Pathaka,*, Manjusha Pandeya, Siddharth Rautaraya " Application Of Deep Learning For Object Detection " International Conference On Computational Intelligence And Data Science (Iccids 2018).

[8] Parisa Pouladzadeh, Shervin Shirmohammadi, "Mobile Multi-Food Recognition Using Deep Learning Acm Trans. Multimedia Comput. Commun. Appl., Vol. 13, No. 3s, Article 36, Publication Date: August 2017.

[9]You Only Look Once: Unified, Real-Time Object Detection Joseph Redmon , Santosh Divvala†, Ross Girshick¶ , Ali Farhadi† University Of Washington, Allen Institute For Ai† , Facebook Ai Research¶

[10] S. Qiao, H. Zhu, L. Zheng And J. Ding, "Intelligent Refrigerator Based On Internet Of Things," 2017 Ieee International Conference On Computational Science And Engineering (Cse) And Ieee International Conference On Embedded And Ubiquitous Computing (Euc), Guangzhou, 2017, Pp. 406-409, Doi: 10.1109/Cse-Euc.2017.262.

[11] Haidawati Nasir, Wan Basyar Wan Aziz, Fuead Ali, Kushsairy Kadir "The Implementation of IoT based Smart Refrigerator System", 2018 2nd International Conference on Smart Sensors and Application (ICSSA)

[12] Hsin-Han Wu, Yung-Ting Chuang, Low-Cost Smart Refrigerator, 2017 IEEE 1st International Conference on Edge Computing