International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

Visual E-commerce Application using Deep Learning

Bhawana S. Dakhare[1], Amish Khatu[2], Hrushikesh Baliar Singh[3], Aniket Gupta[4]

Professor, Dept. of Information Technology, Bharati Vidyapeeth College of Engineering (Navi Mumbai),
[2,3,4]Student of Bharati Vidyapeeth College of Engineering (Navi Mumbai) , Maharashtra, India

---

Abstract - *The search functionality of ecommerce app is one of the most important parameter inorder to make it more user-friendly and earn maximum benefits. Many people search online before buying a product. But often times, a simple spelling mistake by the customers can make the E-commerce application lose potential customers. This is because many are still using Text-based search engines to handle on-site product queries. Textual search is good at finding keywords from a set of organized data, but it performs Searching for long tail keywords and handling misspelled words, understanding natural language expressions, Distinguishing product descriptors from product names. So, there is a need to create such a system which will solve this problem and make a user-friendly environment for the customer, which will also be performance efficient. Hence, we propose a system which will scans real time world objects and identifies it based on the deep learning algorithm. There are three phases for the searching technique: scanning, detection and recommendation. With the help of YOLO(deep learning object recognition technique) the image of the object will be scanned, once the object is detected it will recommend the desired product to the customer/user using web mining. Thus the technique will produce optimized and efficient results*.

***Key Words*: E-commerce application, *YOLO*, *deep learning*, Object recognition, Natural language expressions, Search functionality.**

1. INTRODUCTION

The digital world is getting crowded. The number of web sites have been hosted by millions of companies, supplement to their traditional merchandising avenues with e-commerce. Retail success is no longer all about physical stores. E-commerce has become a huge market throughout the world where there is direct contact between the producing company and customer. So benefited in both ways for customer as well as seller. There are many e-commerce web sites, but most of them only uses text-based input. Tradition search is a tedious process and should be now upgraded to smart search.

This proposed work is an attempt to create a system that scans real time objects of their choice using android application and get immediate results. Identification of the objects will be done using YOLO(You look only once) algorithm. There are three phases: scanning, detection and recommendation. YOLO is one of the most efficient algorithm for object scanning as compared to R-CNN and fast R-CNN. Once the object is recognized, the product results will be fetched from the online hosted database and shown to the user. User can then add to product to cart and proceed to place order or continue shopping. The user can view history of his ordered products. Object Detection is the process of finding the instances of the real world object. E.g.- mobile, ball, fruits etc. and E-commerce is the process of buying or selling any products on online services over an internet. Implementing Deep Learning technology for object- search will give rise to new era of ecommerce application, where user has to puts his minimum efforts.

1.1 Problem Statement

There are many e-commerce web sites hosted by different companies, but most of them only uses text-based input. When shopping on ecommerce websites, product search is a key component. Many times there arises a problem for the old age people and the illiterate people that they are unable to specify the thing or the product when they want to buy something. Thus, in spite of having money and mobile device(the platform to purchase) they are not able to make them useful. This also affects the e-commerce and business greatly. Sometimes, even the literate people are not able specify the name of the product. There can arise a problem, when the user is not sure about the product information or he/she has just seen the product and desires to buy it, searching for that product following the traditional way could be tedious. So, there is a need to create such a system which will solve this problem and make a user-friendly environment for the customer, which will be performance efficient. Hence, the system designed in this project will allow user for live image detection of the item or products and provide it as input on ecommerce sites/applications. The system will then recommend the user with similar products on the basis of scanned objects.

1.2 Scope of Project

Visual Ecommerce Application can be used in all kind of ecommerce applications and even in the applications where object search is required. This system can be used by common people as well as by business people for search operation on ecommerce website. The object recognition process can be proved helpful in medical sector, i.e. the doctors can scan for similar medicines and according to contents recommendations will be given. Application will be effective in various scientific fields i.e. in automated classification system for galaxies and stars in astronomy.

1.3 Objective

- To create an ecommerce application that can scans real time objects and identifies it based on the deep learning algorithm YOLO.
- To train model using YOLO.
- To provide fast and efficient searching process for users of ecommerce application.
- To enhance the ecommerce market for buying products.

2. LITERATURE SURVEY

[1]. In the previous system, the method used for object recognition i.e. the algorithms were R-CNN, Selective search, fast R-CNN or region proposal. These methods were not efficient in detecting the objects as compared to YOLO. In R-CNN the process involved creating bounding boxes around the image. It looks at the image in randomly sized windows and tries to group the json pixels, we run the boxes through pre-trained Alexnet.[2] Keke Tang, Xiaoping Chen, Peng Song proposed their work in project named 3D Object Recognition in Cluttered Scenes with Robust Shape Description and Correspondence was published by IEEE. In this paper models have been implemented for identifying the patterns of the object by removing noise, occlusion and featureless objects. By this approach a 3D object can be recognized properly. They implemented methods in C++ and execute it on a desktop PC with an Intel Xeon E3-1230 v3 CPU (3.4GHz,4 cores) and 8GB memory without using any parallel computing techniques. They build and search HNSW graph for correspondence selection by employing the Non-Metric Space Library. To speed up descriptor comparison between models and a scene, they build another HNSW graph off-line to index sampled descriptors from all candidate models. First, they construct a Signature of Geometric Centroids descriptor that is descriptive and robust, and apply it to find high-quality potential feature correspondences. Second, they measure geometric compatibility between a pair of potential correspondences based on isometric and three angle-preserving components. Third, we perform effective correspondence selection by using both descriptor similarity and compatibility with an auxiliary set of "less" potential correspondences. [3]. Paper presented a brief review on Visual Search Engine for Product Images proposed by Xiaofan Lin, Baris Sumengen. This paper introduces a novel visual search engine for product images, which provides a brand-new way of visually locating products through Content-based Image Retrieval (CBIR) technology. This discusses the unique technical challenges, solutions, and experimental results in the design and implementation of this system. Query By Image Content (QBIC), presents the technologies allowing to organize digital pictures by their visual features. They are based on the application of computer vision techniques to the image Retrieval problem in large databases. Content-Based Image Retrieval (CBIR) consists of retrieving the most visually similar images to a given query image from a database of images. They used Content Based Image Retrieval (CBIR) based on evolutionary algorithm. Initially, the shape, color and texture feature is extracted for the given query image and also for the of the database images in a similar manner. Subsequently, similar images are retrieved utilizing an evolutionary algorithm based similarity. Thus, by means of the evolutionary algorithm, the required relevant images are retrieved from a large database based on the given query. The proposed CBIR system is evaluated by querying different images and the efficiency of proposed system is evaluated by means of the precision-recall value of the retrieved results.[4] This paper was named as Visual Commerce Application (Android) and the authors were Jay Shah, Tapan Desai , Pooja Shah. Their work presented SmartGlass Application for Visual Commerce based on Android Platform. Visual Commerce is a method of acquiring products online via a search based on image whereas E-Commerce works primarily on text based search. The primary purpose of this paper is to suggest an alternative solution for text based searching in the form of SmartGlass. Visual Commerce as mentioned above is an image based search i.e. searching is done on the basis of the image recognition and application displays the output in the form of a products from various online stores to the end user. Visual search module is based on Database Technology, Crawlers and Optical Character Recognition (OCR) engines. The user has to click an image of the text he/she wants to search. This text is then read by the OCR engine. The text is then passed as a query to the server. The server consists of MySQL database which daily populated by two crawlers. The database is then searched for the query. Once the results are found they are sent back to the mobile device. The response is quick and the results are displayed within a span of 3 seconds. The main feature of Visual Search is not using any readymade search engines but aggregating news in their own database and displaying results for the user. They used the algorithm with the following steps 1. Scan the word content 2. Extract the text content from the image 3. Spell check the retrieved content 4. Text sent to the server 5. Server Side Programming a. Split the sentence into words b. Remove stop words c. Search for relevant news articles based on scanned words d. JavaScript Object Notation (JSON) encode the retrieved articles and pass it to the device. 6. Decode the JSON11 response on the device. 7. Display results.[5] Project was created by Authors: Damir Demirovic, Emir Skejic, Amira Serifovic and project was named as "Performance of some image processing algorithms in Tensor Flow". Signal, image processing algorithms in recent time are used in a daily routine. Due to huge data and complexity, their processing is almost impossible in real time. Image processing algorithms are inherently parallel in nature. In this paper image processing algorithms were evaluated, which are capable to execute in parallel. All algorithms were tested in Tensor Flow, which is a novel framework for deep learning, but also for image processing. Tensor Flow Obtained speedups range from 3.6 to 15 times. In this paper Tensor Flow is widely used. The basic unit in Tensor Flow is a computational graph. Paper conducted the

performance of Tensor Flow implementation of basic algorithms used in image processing. Parallel processing has become most dominant for high-performance computing. In this paper 2 data set is taken for all algorithm. Flow of process: 1) Convolution 2) Edge detection a) Image gradient b) Canny filter 3) Gaussian filtering 4) Matrix multiplication 5) Image resize 6) Image segmentation 7) Image rotation 8) Image rotation 9) Partial differential equation. [6] The Paper named You Only Look Once (YOLO) : Unified,Real-Time ObjectDetection whose authors where Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi . They presented YOLO, a new approach to object detection. Prior work on object detection repurposes classifiers to per-form detection. Instead, they frame object detection as a regression problem to spatially separated bounding boxes and associated class probabilities. A single neural network predicts bounding boxes and class probabilities directly from full images in one evaluation. Since the whole detection pipeline is a single network, it can be optimized end-to-end directly on detection performance. Their unified architecture is extremely fast. Their base YOLO model processes images in real-time at 45 frames per second. A smaller version of the network, Fast YOLO, processes an astounding 155 frames per second while still achieving double the map of other real-time detectors. Compared to state-of-the-art detection systems, YOLO makes more localization errors but is less likely to predict false positives on background. Finally, YOLO learns very general representations of objects. It outperforms other detection methods, including DPM and R-CNN, when generalizing from natural images to other domains like artwork.[7] Paper was published by the authors Shankar M. Patil and A.Z. Malik named as "Correlation based real-time data analysis of graduate students" in year 2019, this paper helped in the project implementation methodology and to follow an organized approach for project implementation as well as technical paper creation.

3. SYSTEM METHODLOGY

A. Block Diagram

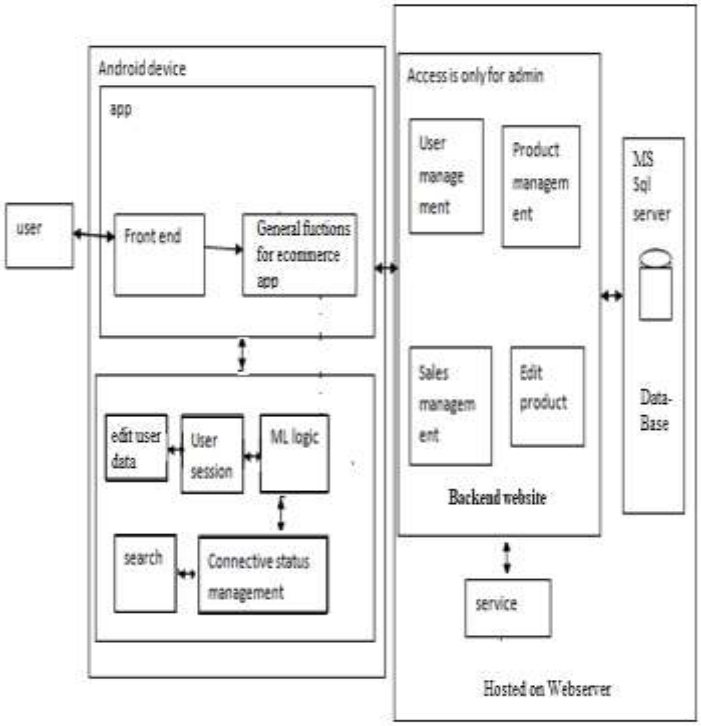


Fig-3.1 System Architecture

B. Backend Website Development

1) Post-Analysis Phase

As per the analysis, we determined the requirements of the project , what all parameters the user has to deal, while interacting with the application. Considering all the parameters, we designed the database schema and did the detail study of ER diagram of the project. We implemented the database on the local- server of PC and created database in visual studio. Database consist of 7 tables such as customer_master, category_master, Admin_login, Product-master, temp_cart, Transaction_details

2) Webpages development

Considering the database design, we constructed webpages in Asp.net using Visual studio 13. Webpages included: Login Page- provided Login functionality (only admin is permitted), Category_manager - To manage categories of the products, Product_manager - Adding, Editing products, Sales_history - Shows the details of the order placed by user and has functionality to send invoice, Customer_manager - to manage customer_data. Thus the entire website was created on a single machine and tested.

C. Development of Android Application

A user-friendly Ecommerce Android application for customers :

After creating the website, we started developing the front end GUI i.e. the Android application , which will be used by the user. We analyzed,, what activity pages will be required by a general user while dealing with the application and thus we created following activities:- i. Login ii. Registration iii. Forgot password iv. Home (providing all features of application to user). v. Camera scanning activity. vi. View cart vii. See cart viii. Contact Us. ix. Update_profile. After developing the layouts the corresponding Java files, we tested the application.

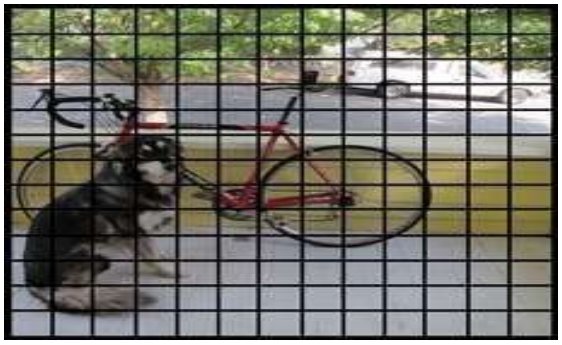
D. Implementing Object Recognition Search

i. YOLO (You Look Only Once) algorithm:

We then started the work related to object recognition module using deep learning algorithm(YOLO).

In object recognition algorithm, live object detection through camera is used for detection of objects, followed by the processing phase in which the original frame is passed on to the object detection algorithm **YOLO** which is based on convolutional neural networks (CNN). This algorithm divides the image into 13 *13 grid of cells. Each cell tries to predict the bounding boxes with confidence scores for those boxes and class probabilities. Each of these cells are responsible for predicting five bounding boxes. A bounding box describes the rectangle that encloses an object and the prediction is based on the following parameters.

- x, y, width, height for the bounding box's rectangle.
- The confidence score .
- The probability distribution over the 20 classes.



Fi: 3.2 Formation of Grid Cell

For each bounding box the cell also predicts the class and it gives probability distribution over all the classes also the confidence scores for the bounding box and the class prediction are combined into one final score that tells us the probability that the bounding box contains the object.

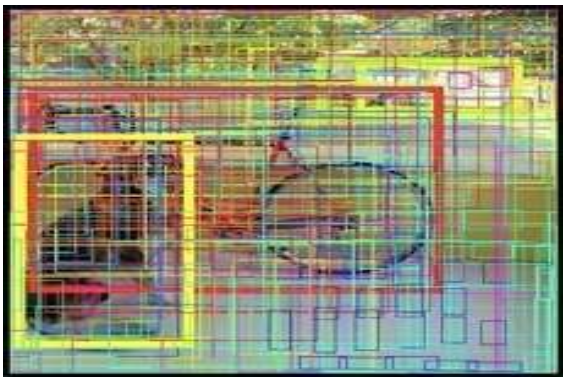


Fig: 3.3 Bounding box formation

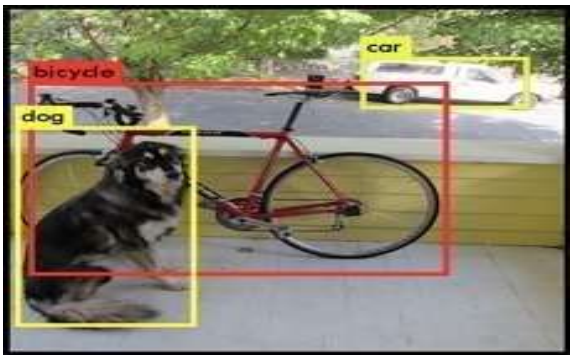


Fig: 3.4 Object Labeling

Confidence score is calculated as Pr (Object)*IOU (pred, truth). If no object exists in that cell, the confidence score is zero and for accurate prediction the confidence score should be equal to the intersection over union between the predicted box and the ground truth. Intersection over Union is an evaluation metric used to measure the accuracy of an object detector.

In order to apply IOU, we need:

- The ground-truth bounding boxes.
- The predicted bounding boxes from our model.

Computing Intersection over Union is calculated as:

IOU = Area of Overlap / Area of Union

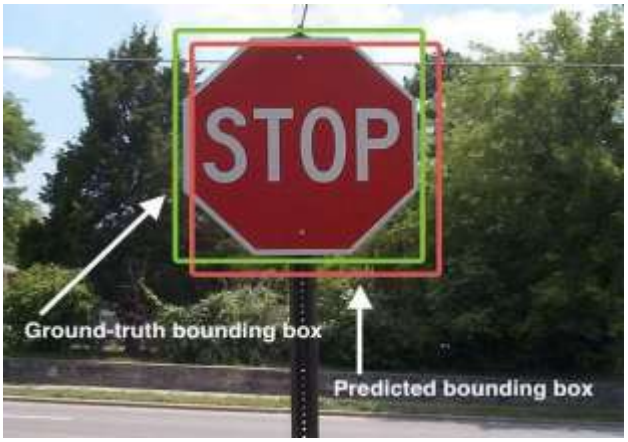


Fig: 3.5 IOU Example

YOLO model works efficiently on the predefined dataset in which 20 data items are present, they are: bottle, TV/monitor, sofa, chair, bicycle, motorbike, bus, car, cat, dog, horse, sheep, train, aeroplane, person, dining table, birds, boat, cow, and potted plant.

We used the YOLO model for implementing the object recognition process, we designed the search by scanning objects based on TensorFlow libraries which also helped in classification of objects. Object recognition search process needs only the object in front of camera and then generate a rectangular bounding-box around the object accurately and instantly. Search also gives the probability of the scanned objects while classifying the objects. This search is included in the camera_scanning activity in the Android application. Thus, the implemented object recognition in Android Application was successful.

ii. Introduction to CNN

Convolutional Neural Network(CNN) is one of the most efficient technique used in mathematics that converts the image into multiplication matrix. Mostly, TensorFlow and Keras are the libraries which are used in training CNN model. These techniques help in classification. **TensorFlow** is an open source library created for Python by the Google Brain team. TensorFlow compiles many different algorithms and models together, enabling the user to implement deep neural networks for use in tasks like image recognition/classification.

The CNN Network consists of three layers:

(a) **Convolutional layer**: This is the first layer of CNN which takes input image and maps features from it. The first layer of a neural network takes in all the pixels within an image Once the data has been fed into model, the neural network will apply different filters on it. A common filter used in CNN is 3. The filter examines the height and width into 3*3 pixels.

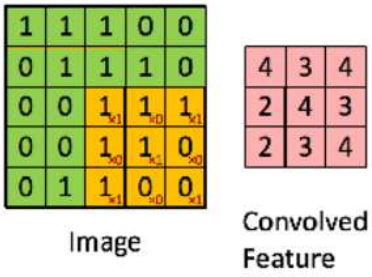


Fig 3.6- Convolved Feature

After the feature map of the image has been created, the values that represent the image are passed through an activation function or activation layer. The typical activation function used to accomplish this is a **Rectified Linear Unit** (ReLU)

(b) **Pooling Layer:** After applying the activation function to the image, it is then sent to the pooling layer. The pooling layer makes the network more flexible and adaptable at recognizing the images based on the relevant features. This layer removes the unnecessary part present on the image with some specified size. Pooling layer controls the further process as it makes the decision about the relevant features from which the model will learn.

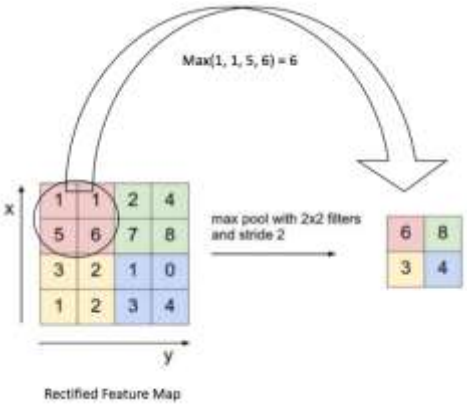


Fig 3.7- Consideration of Maxpooling

(c) **Flattening:** The final layers of the CNN Network , the densely connected layer, which processes the data into vector form.

(d) **Fully Connected Layer:** The final layer which is also called as Artificial Neural Network(ANN), that train the model completely. The function of the ANN is to analyze the input images and combine it different attributes. This layers are collection of neurons, which gets activated on providing input image and classifies the image.

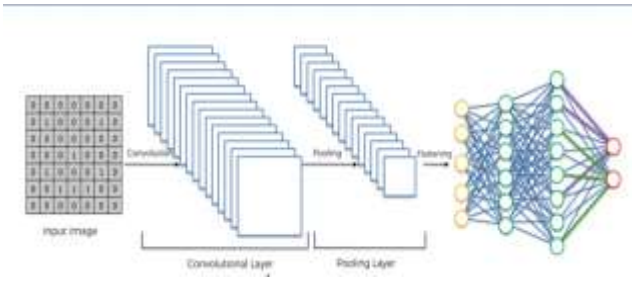


Fig 3.8- CNN Network

Hence, the above contents were the hardware and software details that identifies objects digitally.

E. Creating Web Service

As the website and the application was ready, there was a need to connect the Front end GUI and Back end website. For this purpose, we created a web service project in .Net using Visual studio. All the users actions are directly connected to the website which will be hosted on the server using this service. E.g. - users getting registered to the application on Android device are reflected on the website using service project. For this purpose the service URL is included in the connectivity module of the Android application. The object recognized and searched will retrieve the results from the online database (MS SQL server) and provide it to user.

F. Hosting project on webserver

As all the modules are ready, we hosted the website on the server with IP address:- 103.21.58.192 , database name: Object_recognition_db and the hosted website link is : http://demoproject.in/Object_recognition_shopping_website/Login.aspx. So can be managed by the admin easily. Also the created service project was hosted for the same. MS SQL server 2014 is used for database management (online). After successful hosting, the testing and debugging of the Android application was conducted to run efficiently with the online hosted data providing accurate and instant results to the user or customers.

2. RESULTS

(a) Website webpages:

Fig 4.1: Login for admin

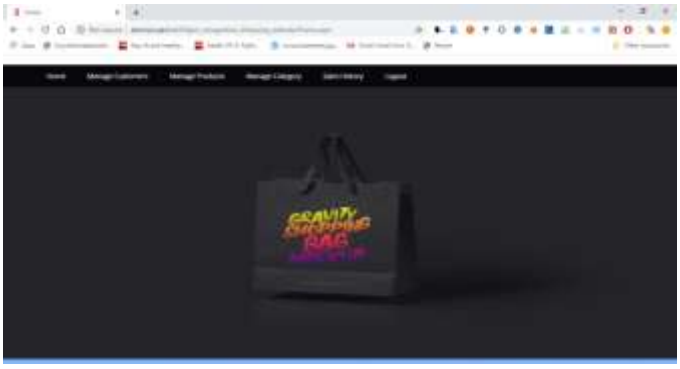


Fig 4.2 : Home page

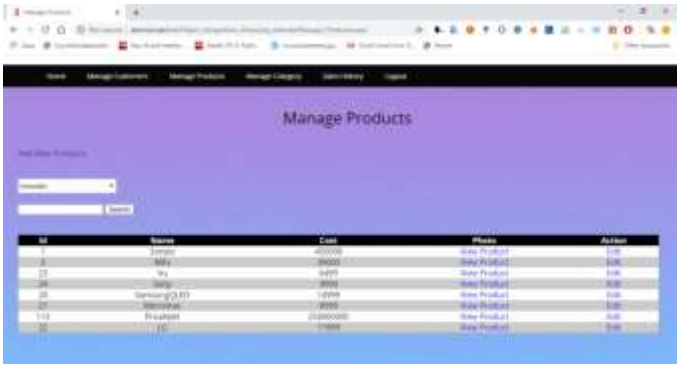


Fig 4.3: Webpage for managing product details

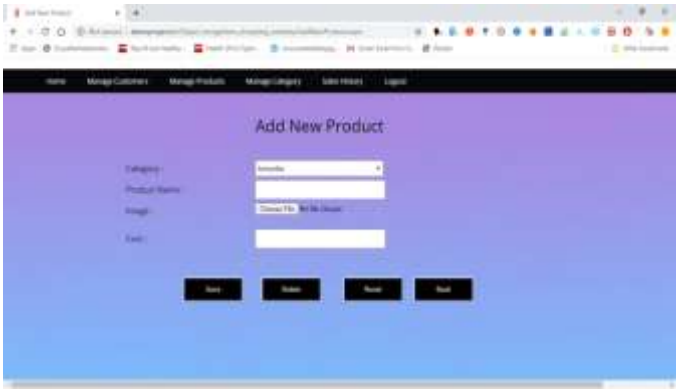


Fig 4.4: Webpage for add new product

Fig 4.5: Webpage for edit product details

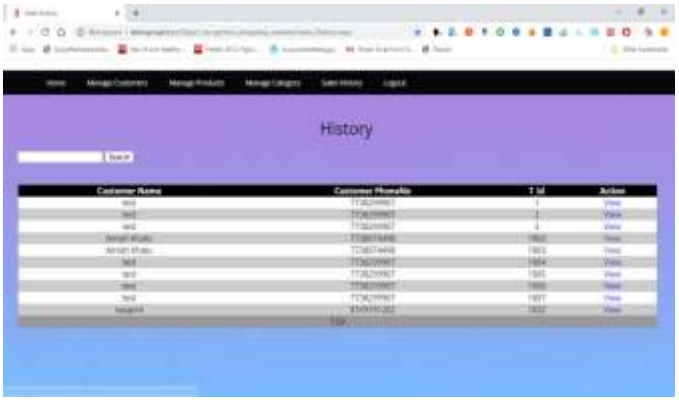


Fig 4.6: Webpage for sales history details

(b) Implementation of object search and recognition on android platform :

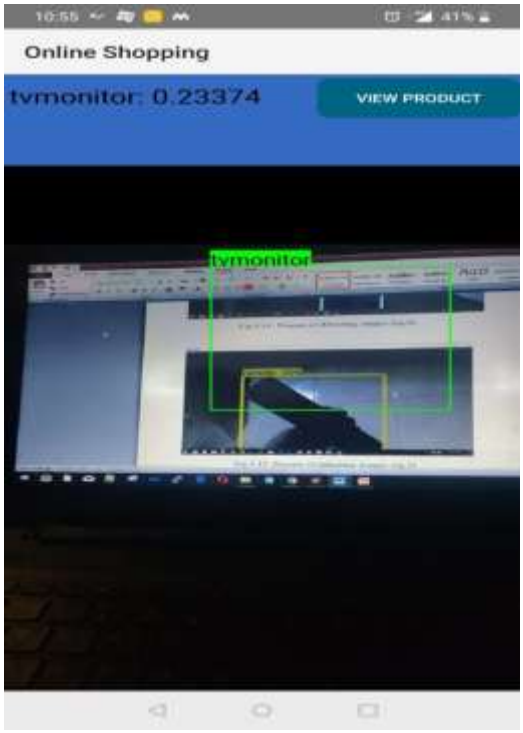


Fig 4.7: Process of detecting images (eg.1)

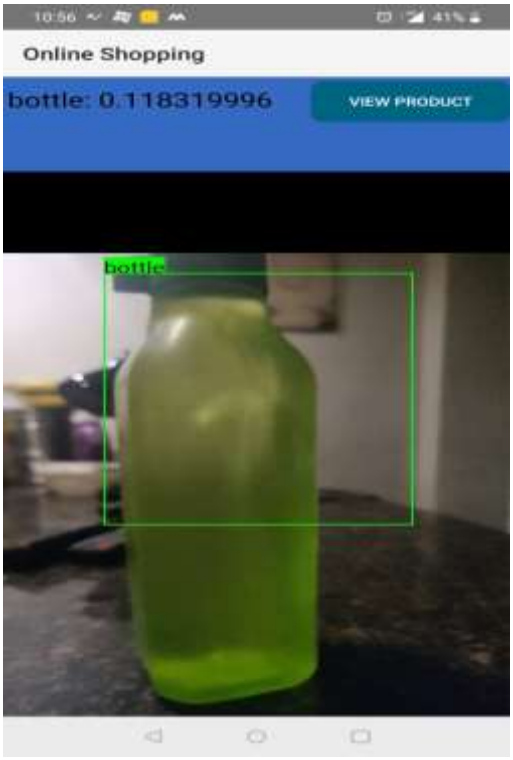


Fig 4.8: Process of detecting images (eg.2)

Fig 4.9: Process of detecting images (eg.3)

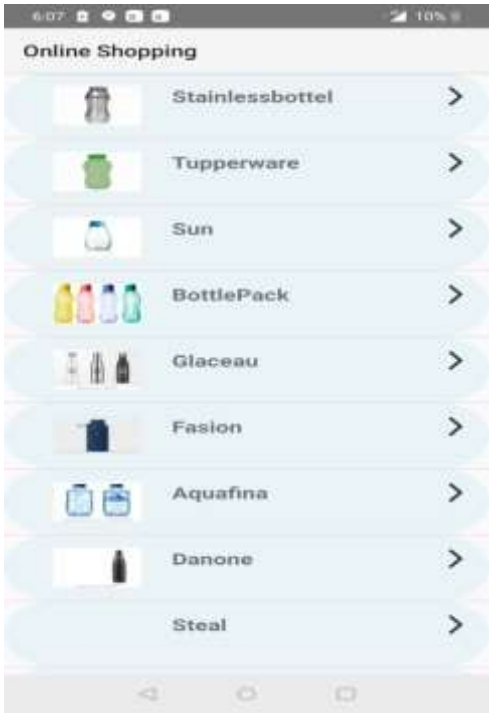


Fig 4.10: Results fetched from webserver, after scanning the object

Fig 4.11: Home Screen for app

(C) Result Analysis

Performance is evaluated using different parameters such as: Accuracy, Recall, Precision, & Specificity for Object Recognition in E-commerce Application[17].

i. Accuracy

It's the ratio of the correctly labeled subjects to the whole pool of subjects.[17]
Accuracy is the most intuitive one.

Accuracy = (TP+TN)/(TP+FP+FN+TN) *100

= (20+23)/(20+5+2+23) *100

= (43/50)*100

= 86%

ii. Precision

Precision is the ratio of the **correctly** positive labelled by our program to all positive labelled subjects[18].

Precision = TP/(TP+FP) *100

= (20)/(20+5) *100

= (20/25)*100

= 80%

iii. Recall (Sensitivity)

Sensitivity is the proportion of cases classified positive in relation to all cases tested positive[18].
Recall = TP/(TP+FN) *100

= (20)/(20+2) *100

= (20/22)*100

= 90%

iv. Specificity

Specificity is the proportion of the true negatives correctly identified by a test. It suggests how good the test is at identifying normal (negative) condition.

Specificity = TN/(TN+FP) *100

= (23)/(23+5) *100

= (23/28)*100

= 82%

4. CONCLUSION

The use of object detection in ecommerce application makes it very user friendly. We have successfully implemented object recognition searching technique using YOLO. Determining the objects by just clicking the pictures and providing the related information is very beneficial for the customers. YOLO model is trained to recognize the objects and categorized into different classes based on various parameters. The scanned image is detected and similar products are recommended to the user.

REFERENCES

[1] Link https://www.youtube.com/watch?v=4eIBisqx9_g

[2] Authors: Keke Tang, Xiaoping Chen, Peng Song, published paper named "3D Object recognition in cluttered scenes with robust shape description and corresponding selection" (IEEE) in year 2017.

[3] Authors: Xiaofan Lin, Baris Sumengen published paper named "Visual Search Engine for Product Images" published by ISOE in year 2018.

[4] Authors: Jay Shah, Tapan Desai , Pooja Shah published paper named "SmartGlass: Visual Commerce Application (Android)" published by ICACTA in 2015.

[5] Authors: Damir Demirović, Emir Skejić, Amira Šerifović published paper named "Performance of some image processing algorithms in Tensor Flow" in year 2018 published by IEEE.

[6] "A YouOnlyLookOnce (YOLO) algorithm : Unified RealTimeObjectDetection" is the paper published by the authors named Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi in may-2016.

[7] Authors: Shankar M. Patil and A. K. Malik published paper named 'Correlation based real-time data analysis of graduate students behavior' in 2019.

[8] Authors: J. Dong, Q. Chen, S. Yan, and A. Yuille published paper named "Unified object detection and semantic segmentation in Computer Vision–ECCV" 2014, pages 299–314. Springer, 2014

[9] Authors: D.Erhan,C.Szegedy, A.Toshev, and D.Anguelov published paper named "Scalable object detection using deep neural networks",in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 2155–2162. IEEE, 2014.

[10] Authors: S. Ren, K. He, R. B. Girshick, X. Zhang, and J. Sun published paper named "Object detection networks on convolutional feature maps",in year 2015.

[11] Authors: Thadagoppula, P. K., & Upadhyaya, V published paper named "Speed detection using image processing", in 2016 International Conference on Computer, Control, (IC3INA). doi:10.1109/ic3ina.2016.7863015

[12] Link of Website referred for Image Processing basics: http://www.idi.ntnu.no/~blake/gbimpdet.htm

[13] Dataset link : https://pythonprogramming.net/videotensorflow-object-detection-api-tutorial/

[14] https://pjreddie.com/darknet/yolo

[15] https://hackernoon.com/understanding-yolof5a74bbc7967

[16] Link for OD https://medium.com/@jonathan_hui/real-timeobjectdetection-with-yolo- yolov2-28b1b93e2088

[17] Author named Cristina Oprea published a paper named "Performance evaluation of the data mining classification methods".

[18] Author: Shankar M. Patil published paper named "Content based image retrieval using color, texture and shape' in year 2012".