International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-1, May 2020

# Deep Learning based Image Processing for Cashier-less Self-Checkout Methodology

**Sudeshna Thakur, Neha Patil, Soumya Patil, Nidhi Hegde, Amol Dumbare**

***Abstract***: ***In recent years, shopping experiences are becoming more advanced. These include the attempts of market shelves as well as the currently booming online shopping. Online shopping has a better convenience but not yet accepted on a large scale by many people. Retail shops still retain greater response by the users and thus the retailers are moving towards an attempt of cashier-less shopping. A major problem of retail shops is that the people have crunch-time for shopping and cannot afford the waiting time at the checkout counters. Addressing this problem, we have developed a shopping style which saves time of checkout and also the time of maintaining social distancing queues. This research paper presents a stereo vision-based AI system which is useful to monitor the customers while shopping and also the items which are added or replaced in the virtual cart. The customers can directly walk out of the store after shopping and the final order cost of the shopping will be evaluated. This amount will be charged to the customer's account. The system makes sure that there are no errors made during the evaluation and there are no charges for products which are not brought home. To achieve all this, the system uses image processing, object detection and face recognition algorithms that are widely practiced at present. The system also uses sensors like RFID tags and pressure sensors for weight measurement and detection of products on the shelves.***

***Keywords: Digital Image Processing, Face Recognition, Object Detection, stereo-vision cameras, RFID, Sensor.***

## I. INTRODUCTION

With the growing development in the industry of artificial intelligence, the technologies of object detection and face recognition are gradually been accepted on a large scale. The research in this paper involves the use of these technologies for the ease of daily local shopping. Moving online completely in particular categories of shopping is not feasible. Local shopping markets face many challenges for billing and statement generation during the checkout process. Each item has to be scanned individually which takes up much time of the customer and make the task tedious for the billing staff where hundreds of objects have to be scanned in a day. In the modern time, the people have more income to spend and lesser time to spare for grocery shopping. The technology advancements are moving towards an artificial intelligence era. Thus, the main goal of the research is to provide a better solution to this problem. It provides a real-world retail store shopping without having a checkout process. The customer grabs the items from the store and just walks out without worrying about the checkout lines, cashiers or the card payment problems. This research is not only of scientific significance, but also has specific economic and social benefits. Following are the specified concepts useful for this research.

### a. *STEREO VISION SYSTEM*

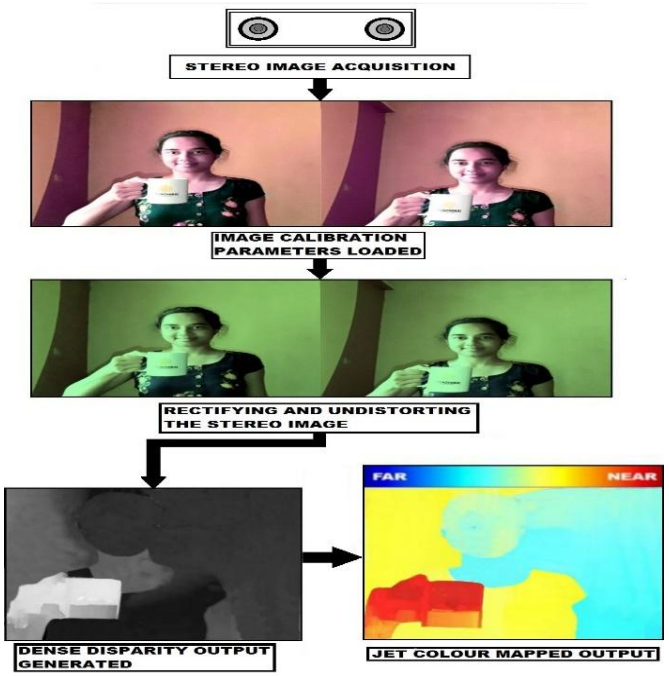


**Fig. 1.Depth mapping process.**

Stereo-vision is a field of study in Machine Vision for sensing depth information from the environment. The human vision system is recreated by using 2D representations of the particular view and derive 3D depth information about the view. This information is useful in tracking the motion of objects and humans in 3D space and obtain the distance information and features.

Manuscript received on April 30, 2020.
Revised Manuscript received on May 06, 2020.
Manuscript published on May 30, 2020.
* Correspondence Author
**Sudeshna Thakur**, Student, Pimpri Chinchwad College of Engineering & Research Ravet, Pune, PIN: - 412101, India. Email: t.sudeshna44@gmail.com
**Neha Patil**, Student, Pimpri Chinchwad College of Engineering & Research Ravet, Pune, PIN: - 412101, India. Email: neha.pat1998@gmail.com
**Soumya Patil**, Student, Pimpri Chinchwad College of Engineering & Research Ravet, Pune, PIN: - 412101, India. Email: soumyapatil031@gmail.com
**Nidhi Hegde**, Student, Pimpri Chinchwad College of Engineering & Research Ravet, Pune, PIN: - 412101, India. Email: nidhih000@gmail.com
**Prof. Amol Dumbare,** Assistance Prof. in Dept. of Computer, Pimpri Chinchwad College of Engineering & Research, Pune, Pin: -412101, India. Email: amol.dumbare@pccoer.in

These systems achieve good efficiency which is closest to the human perception of vision.
Our developed system uses stereo cameras for achieving stereo vision depth information in order to track the customer's motion in the retail store. It also tracks the actions of adding or removing any item to or from their cart. Further image processing on the depth information is performed using the OpenCV library.

### b. *QR CODE GENERATION*

The advancement in technology has led to the development of QR (Quick Response) code which is an essential medium concerning the information exchange carried within the limited time. It is a storage of various encoded records. QR code has the advantage of having higher information storage capacity. It can encode or encrypt various types of data, like textual data, URL, phone number and Email. QR code also benefits the process of verification of user information from the data centre.
The system used this advanced technology in for fetching the customer's information.
We have developed an application in which the customer provides his/her identification information like name, phone number and residential address. After entering these details, the customer receives a unique QR code generated for use. This QR code comprises of complete information about the customer in an encrypted format. A smartphone with Android or any other OS is useful for extraction of the same encrypted data from the QR code. In this way, our system retrieved the customer information at the entry of the retail store. This data is stored at the store's database and is provided to the AI system for further customer mapping with the shopping items.
Use of QR code helps fetch the details of all the customers effortlessly furthermore easing the customers to save time.

### c. *OBJECT DETECTION*

Object detection is a concept which is useful to identify or discover objects which belong to a particular class. It is a methodology achieving the classification of objects from their digital images by implementing image processing. The concept of creating a bounding box around the item or object in the image is called segmentation.
For the detection of objects, we have provided the system with a pre-developed dataset.
The dataset used for this purpose is COCO (Common Objects in Context) dataset.
COCO dataset is useful for large-scale object detection. It is also applicable for the segmentation as well as captioning. This dataset comprises various images, bounding boxes and labels and includes images of 91 objects types which are easily recognizable. Hence, we have provided this dataset for our model training.
For training of our object detection model, a pre-trained model YOLOv3 was beneficial.
YOLO stands for You Only Look Once, and for the detection of objects, this model uses the concepts of deep learning as well as convolutional neural networks. The reason for selecting this model for our system is that this algorithm is one of the fastest object detection algorithms since it needs to see an image only once to get trained.
YOLOv3 is the third gradually incremented version of YOLO and is responsible for demonstrating real-time object detection. Furthermore, this model has lower identification exactness, but since its recognition speed is very high, it is the best model. Moreover, this method provides ideal outcomes in the case of consistent execution. The excellent accuracy of this model for our system and high speed of execution makes it a better model for our application.

### d. *FACE RECOGNITION*

The innovation of technologies and today's growing technology requirements are the reason for the development of this technique called face recognition. It is a concept of identification and detection of an individual's face whose image data is present in the data set.
It is a method which is useful in security applications in a beneficial manner.
For facial recognition, our developed system uses the concept of transfer learning. This is a machine learning methodology in which after the development of a model for a particular task, the same model is useful as an initial point for a model on the next incremented task.
In other words, pre-trained models are used as a starting point on the NLP tasks as well as on computer vision. The knowledge gained in one training is exploited for improvisation of generalization in the training of another model.
Transfer learning is an optimization that permits an improvement in performance at the time of modelling the second task. In this method, we train a base network on a dataset initially and then repurpose the features that we learn to the next target network for its training. In some scenarios where we do not have much data for the model, the use of transfer learning helps to develop skilful models.
VGG16 is an architecture of the convolutional neural network and this model is pre-trained by providing a huge dataset for achieving better representation of the low-level features like spatial, edges, rotation, shapes. Due to this, the model is best for knowledge transfer and acts as a feature extractor for new additional images. We utilise this model in our system for facial recognition.
The steps we follow are:

- Importing all the libraries.
- Importing ImageDataGenerator from keras.preprocessing. ImageDataGenerator is useful to import data with labels easily into the model. This class is beneficial as it doesn't affect the data stored on the system and alters the data dynamically while passing it to the model.
- Creating and passing the objects of ImageDataGenerator for both training and testing data.

The ImageDataGenerator will automatically label all the data we provide. Therefore, the data is ready and we pass it to the neural network. Furthermore, our system uses the Haar Cascade algorithm for fetching the frontal face features of the customer. It is an algorithm useful to identify objects, humans in an image or video. This approach is based on machine learning and is implemented in four steps. Haar feature selection is the initial stage in which the Haar-like features are selected and their integral images are created in the next stage for making it speedy. In this algorithm, a cascade function is trained from huge number of positive-negative images. Cascading Classifiers with Adaboost training are the next stages for rejecting the negative samples as fast as possible.

## II. METHODOLOGY

The system is associated with a unique generated QR code for each customer. Every customer entering the retail store would be tagged by scanning the generated QR code at the entrance of the shop. This will enable the sensors of the system application in the customer's smartphone. The stereo cameras present in the store monitors when the customer picks an item from the shelf and add into the cart or remove any item from the cart and put back on the shelf. This in collaboration with the weight sensors analyses the vision, weight and stock location data to ensure if the shopping is done efficiently. This data is provided to the object detection algorithm YOLOv3 for detection of the item that is picked. The customer is identifiable with the help of Bluetooth beacons which fetches the nearest mobile device. The application in the customer's mobile device communicates with the store's Bluetooth network signals. Facial Recognition is achieved using transfer learning VGG16 and Haar-cascade classifier algorithm. Furthermore, the customer information is mapped with the face recognition data which may include images of the user and the other details about the customer. All this data is fed to an AI system. The product uses OpenCV for keeping track of items picked from the rack to the shopping cart. The customers upon leaving the shop are charged on their account that is linked with the store's system application. This system ensures that there are no errors in the evaluation of the bill, and the customer is not chargeable for any item which is not purchased.

Description of accompanying figures: The present invention has been summarized with the help of schematic artworks and overall internal details concerning inventing matter of subject.

## III. DRAWING AND FIGURES

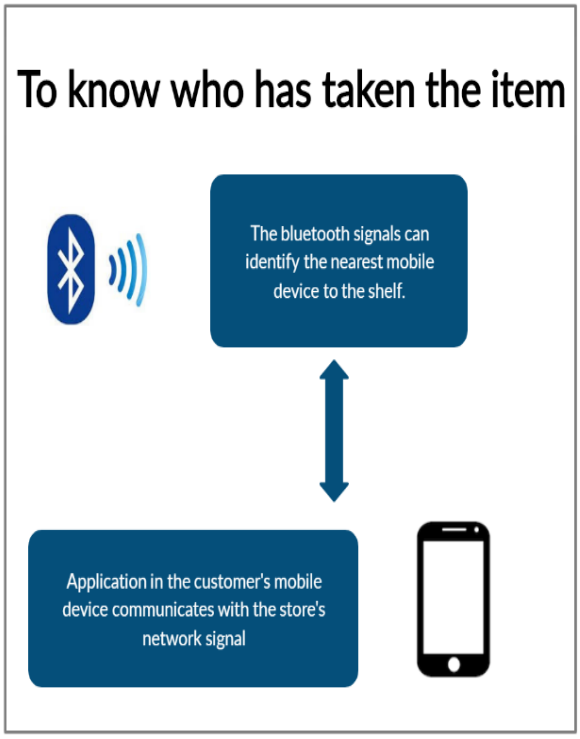


**Fig. 2.Explains who has taken the item**

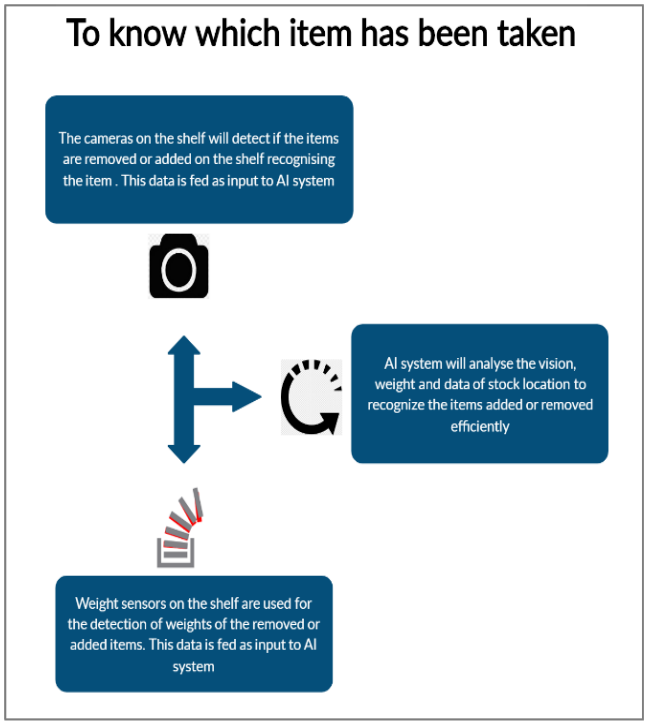


**Fig. 3.Explains what item was taken**

The drawing illustrates: Fig. 2. Explains who has taken the item in accordance with proposed of system. Fig. 3. Explains what item was taken in accordance with proposed system.

## IV. FLOWCHART

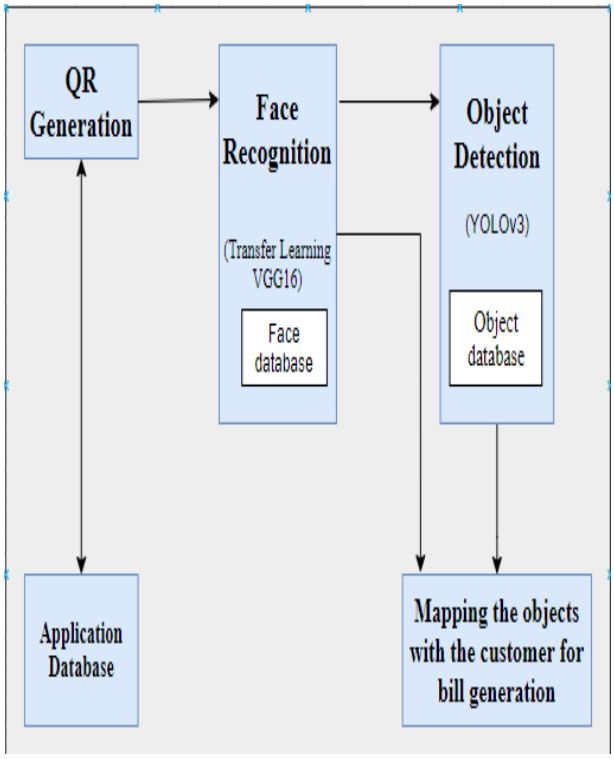


**Fig. 4.Work flow**

*Retrieval Number: A2958059120/2020©BEIESP*
*DOI:10.35940/ijrte.A2958.059120*
*Journal Website: www.ijrte.org*



*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

## V. LIMITATIONS

The system developed experiences some limitations due to the technologies used. Using Computer Vision and OpenCV does not allow the object detection and processing for very small products. This affects the cost-effectiveness since there are many small products which most of the customers will be purchasing. Keeping such products on the shelf will result in a lot of ambiguities in case of object recognition. Furthermore, the packed items such as rice, dals are not efficient to keep on the shelves. The children entering the shop and adding any item to the cart will not be tracked since the children does not have a smartphone and thus the items picked up by them will not be charged. If the customer fails to put back the item on its correct shelf it will be difficult for tracking.

## VI. RESULT

The system is tested for various use cases in handling the stereo vision video feed. The system is feasible for checkout-free grocery shopping and provides the best experience for the customers while shopping. It is a new methodology of next-gen retail shopping. The test results were good with precise outcomes for object detection, face recognition as well as the depth mapping process of stereo cameras. However, the depth mapping is not very precise in the use case of finding actual distances of the customers. This can be acceptable because the accurate recognition of the shopping items in 3D space and tracking the motion of the customers is enough for further processing. Deep learning algorithms monitor the products shopped frequently and provide the stock renewal alerts as well as the pattern for product discounts to the shopkeeper. This ensures the customer satisfaction and customer relationship is maintained effectively.

## VII. CONCLUSION

The strategy of our system is an evolution in technology. The advanced technology with Artificial Intelligence system permits the customer to pick the products from shelves, add those in their carts, and then leave the retail shop without checkout of the products. Besides the convenience to the customer, the system profits the store by reducing the checkout clerks.

## REFERENCES


1. Nivedha N, Prof. Ganesh Pai, Dr. M Sharmila Kumari “Recent Trends in Face Detection Algorithm”.
2. Chaitra T. K, M.C.Chandrashekhar, Dr. M. Z. Kurian ”Face Detection and Feature Extraction” .
3. K. Wankhede. Dr. B Wukkadada, V Nadar “Just Walk Out Technology and its challenges”.
4. G.Bakkiyaraj, V.Dinesh. “The Process of Face Detection and Recognition in Multi-Videos”.
5. Sergey Kosov, Thorsten Thorm¨ahlen, Hans-Peter Seidel “Rapid Stereo-Vision Enhanced Face Recognition”.
6. Yufeng Zheng, Erik Blasch “The Advantages of Stereo Vision in a Face Recognition System”.
7. Mohammad Dawi, Mohamad Adnan Al-Alaoui, Mohammed Baydoun.“ 3D Face Recognition Using Stereo Images”.
8. Naohide Uchida, Takuma Shibahara, Takafumi Aoki “3D Face Recognition Using Passive Stereo Vision”.
9. Chaohui Lü, Xi Wang, Yinghua Shen. “A stereo vision measurement system based on OpenCV”.
10. K Goyal, K Agarwal, R Kumar “Face detection and tracking”.


## AUTHORS PROFILE

**Sudeshna Thakur** is currently pursuing bachelor’s degree from Pimpri Chinchwad College of engineering and Research in the final year of Computer department. And is currently working on the image processing of retail shopping system project and conducting research for same.

**Neha Patil** is currently pursuing bachelor’s degree from Pimpri Chinchwad College of engineering and Research in the final year of Computer department. And is currently working on the image processing of retail shopping system project and conducting research for same.

**Soumya Patil** is currently pursuing bachelor’s degree from Pimpri Chinchwad College of engineering and Research in the final year of Computer department. And is currently working on the image processing of retail shopping system project and conducting research for same.

**Nidhi Hegde** is currently pursuing bachelor’s degree from Pimpri Chinchwad College of engineering and Research in the final year of Computer department. And is currently working on the image processing of retail shopping system project and conducting research for same.

**Amol Dumbare** is currently working as Assistant Professor for Computer engineering department in Pimpri Chinchwad College of engineering and Research, Ravet, Pune, Maharashtra, India.