# Computer Vision Applications to Improve Retail Shopping Experience

**Arun Poochelvan**
Dept of Business Management
Ligs University

***Abstract-*** *Computer vision is a technology which uses Artificial Intelligence that determines and analyses different objects for a wide range of automation. This technology could be used in various fields and one such application is in the retail stores.Retailers need to keep a constant check on their merchandise and accounting for the items. Most of them depend upon the barcode scanners and QR codes. These are not only time consuming but also involve human intervention and errors.In order to make this process quicker, reliable and convenient computer vision could be used. An AI powered mobile app created by using Skill Well software enables to keep track of merchandise and eliminates the possibility of human error. Microsoft has also been working with the advancements in the computer vision for the same purpose. Automatic retail checkout (ARC) relies on the convolution neural network which scans the object placed underneath it using a webcam fitted in it.*

***Keywords***- Retail, Merchandise, Supermarkets, Shopping, Products, Computer vision.

## I. INTRODUCTION

Retail stores and supermarkets have diverse range and different varieties of the products in huge catalogues. The retailer has to manually screen each product using the barcode. They also have to keep track of the merchandise, and this is weary and tiresome as it involves stacking up in the warehouse then into the storeroom and then finally into the shelves [1].

This invites not only for advancements in technology but also human errors. The latter depends upon the efficiency and knowledge of the staff members in the retail shops. Though the performance is based on the standard manual performance, the business completely depends upon the sales of the products. It is common to find duplicate entries of the same product in catalogues. This could be because of variety of reasons such as inappropriate marking of products on their arrival, mishandling on inventory while dispatching etc[2].

To address these problems a branch of artificial intelligence which is making breakthroughs by using innovative techniques and developments in deep learning could be used.[1] This is the computer vision which can tackle the issues in the retail industry and make the shopping between retailer and customer satisfactory. In today's world by using computer vision, it enables personalization and convenience [3].

This uses analysis of images and algorithms that are trained to identify objects and logos of companies in both picture and video format. The algorithms need less pre-processing and perform efficiently. Computer vision disables the poor performance of staff, errors and enables the high performance, efficiency by increasing sales and thus revenues. Thereby giving best customer reviews and happy shopping [1].

## II. REVIEW OF LITERATURE

In everyday life, there are repetitive tasks that are time consuming. In the supermarkets managing goods is hard task as there many shelves to be monitored and arranged. This paves way for devise an automated system ie computer vision. It helps in goods management and detection of products using a movable camera.

There are two ways for detection of object. One being the traditional way of detecting is by using deep learning networks. The other one is the one which sees the object as a regression problem. It uses a single network to return.

The objective of the computer vision is:

1. to generate the products available at that given point of time.
2. Validate the planogram.
3. shopping assistance to customers[6]

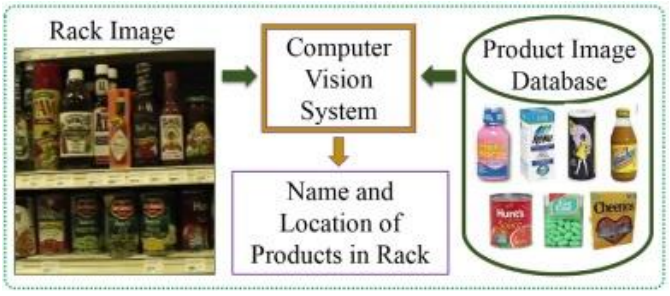



Fig1: Computer vision system for detection of products [6]

Amazon Go1system has a checkout free shopping experience. This involves the use of sensor function which classifies the products and subsequent bill in the mobile application. Computer vision could help visually impaired people from fetching products they want to buy Panasonic has a walk-throughself-checkout system. This is based on radio frequency identification tags. It is a cost-effective method. These are ubiquitous based systems [6][4][1].

Recognition of fruits was done by hybrid classification method based on FSCABC (fitness scaled chaotic artificial bee colony algorithm and feed forward neural network (FNN). The images of the fruits are captured by a camera and the background of each was removed by split and merge and algorithm. Next the colour histogram was extracted to compose a feature space. Finally, the reduced features were sent to FNN which is trained further by FSCABC algorithm [7].

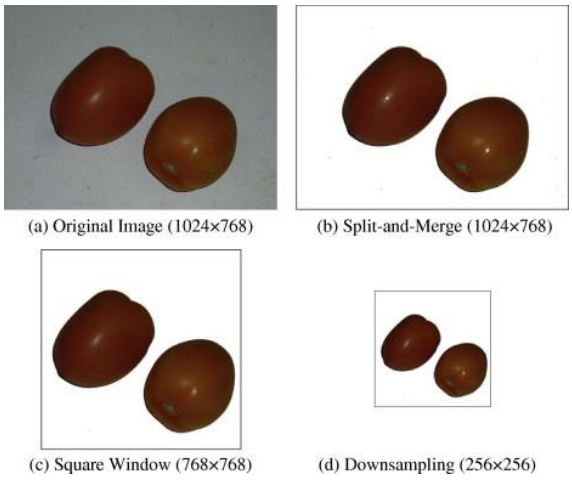



Fig 2-The Steps in preprocessing [7]

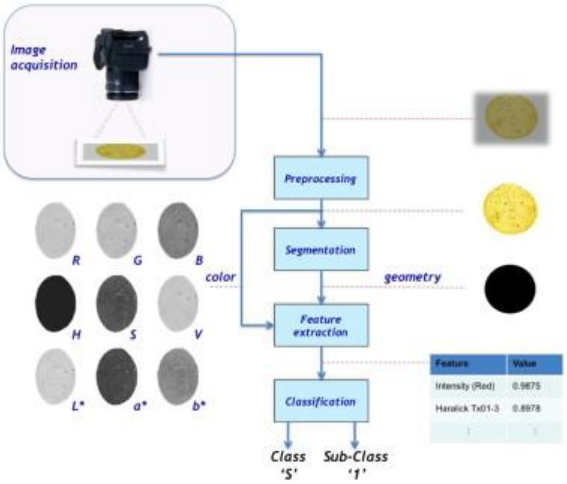



Fig3- : Computer vision schema used to determine quality [8]

Now A similar detecting is done for corn tortillas as well [8].

Intelligent self-checkout system (ISCOS) uses a single camera to detect many products real time. Using web scrapping, the data is products are extracted from various sources and the images are gathered for training. With the help of technologies, the computer vision could be trained to locate, identify the products on the shelves and quantify them. The information is sent to the server and hence the customer. Overall, the computer vision provides a cost effective alternate to the sensor-based approaches [9][6][2][1].

## III. ADVANTAGES

Retail stores and supermarkets have diverse range and different varieties of the products in huge catalogues. The retailer has to manually screen each product using the barcode. They also have to keep track of the merchandise, and this is weary and tiresome as it involves stacking up in the warehouse then into the storeroom and then finally into the shelves [1]. Computer vision eliminates the long waiting in queues and the time-consuming billing process. It helps increase in the business efficiency and its revenue. It allows companies to develop tools for the back-end processes which could be used amongst customers. The performance of the staff members is based on the standard manual performance.

It is common to find duplicate entries of the same product in catalogues. This could be because of variety of reasons such as inappropriate marking of products on their arrival, mishandling on inventory while dispatching etc. [2]

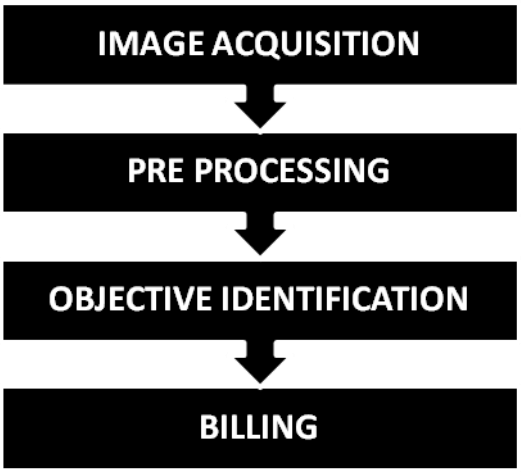
By using innovative ideas in deep learning and segmentation it is possible to overcome the issue. The long-term objective is to make the process more automated and

self-learning. It allows performing higher value tasks to humans than wasting time in data entry.[2]

## IV. METHODOLOGIES

The operating procedure involves the steps as shown in the flowchart below:

Image acquisition involves keeping the product on the conveyor belt which enters the hood. Inside the hood the product interferes with the LASER LDR module. This indicates the arrival of the object within the field view of the camera that is fitted in it. The micro controller switches on and off the conveyor belt. For this case it turns off. In the meantime, signals are sent are sent to extract a frame from the continuous stream of camera feed. This is input for the next steps. Once the image is acquired, the microcontroller switches on and the conveyor belt starts again. Now the product exits the hood. This is done to avoid motion blur in the frame.




InIn preprocessing unit, the image captures pass through 6 preprocessing steps. The extra and unnecessary parts are cropped out from the picture. A binary convex hull of the image is obtained, and information is extracted through the spatial distribution of foreground pixels. By standardizing the orientation of each product, it lowers the difficulty of the object identification model thereby allowing it to focus on the characteristic features. Upon image acquisition a detector called canny edge performs non maximal suppression and hysteresis thresholding to give binary image containing edges. Next it is subjected to morphological closing using 7x7 rectangular structuring elements in order to fill out voids and gaps. Along the object lights and the empty background also gets captured. To overcome this, a minimum area is proximate around the object and then it is cropped into a rectangle. The edges of the images are sharpened by using a tool called sobel kernel. Histogram equalization is applied onto the light component of the image. This enhances the contrast and makes it clearer and more prominent.

Object identification is done by using convolutional neural networks (CNNs). These belong to deep neural networks they have wafers of convolutional layers. The operation involves applying a kernel element to the entire input orientationeach unit of the score map is connected to the input pixels via kernel weights. CNNs are highly efficient and are less prone to errors. More than one kernel is applied to each layer of the waffle as a result multiple features are extracted in parallel. The outputs in each case are checked by subjecting it to a nonlinear function. They are joined simultaneously from a single layer to the subsequent layers. In fully connected layers, features are combined, and it performs the classification task without any discrimination.

Another method for identification is by using activation functions. Non linearities applied to the multiple layers enable the network for complex functions. The last level of layer is subjected to SoftMax activation whose results depend upon certain factors. The factors like units of layer, activations.Pooling operations combines the nearby features in a map which eliminates redundant features and retaining significant ones. The feature maps are more compact and robust to local translational invariance. This is achieved by convolving the input with the kernel. Initialization of parameters controls the signals. If neglected it retains the forward and back propagated signals which vanishes or explodes. Gaussian distribution with a standard deviation of p 2/N (N is the number of incoming nodes to one neuron) is used and initialized to 0.1.The last step is billing. Once the product has been recognized it is added to the shopping cart along with its price. The customer checks out and receives a receipt from the printer.[1]

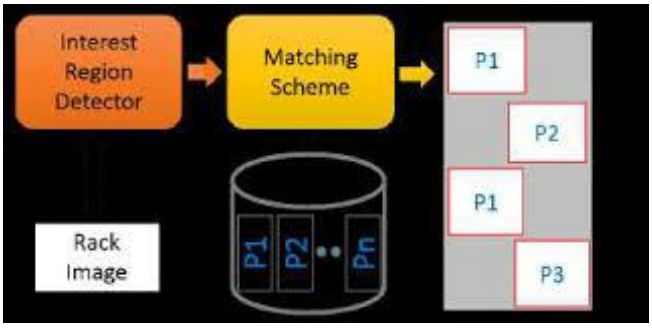



Fig4: Block diagram of a detector based method: P1, P2, · · · ,Pn are the products.[6]

Another methodology was done using unity and SDK Vuforia to work on the AR powered glasses and mobile phones. To visualize the AR, basic SDK allows object

tracking and actualize the inventory management software. Software unity was also adopted as it could be used for all platforms. [2]. Microsoft is testing the computer vision technology based on deep learning and segmentation which can spot duplicates in the catalogue by the separating foreground i.e., image and background.[3]

## V. PROPOSED METHOD

The method implemented involves certain steps.

Data curation- curation of the dataset received from the retail items was done. Each product was made to pass the hood and the picture was captured at different orientation. By doing so it has pictures of the empty sides and the nutrient declaration as well. The underlying hypothesis is that, just like humans it should be able to identify the object in any way it is placed. In this case study 31000 images were acquired, and they performed 65-25-10 split for training, validation and testing.[1]

Training of neural network- for training and validation purpose 32 images from the mini batch was performed which consists of 16 images from the preprocessed dataset. The weights and biases are penalized with weight decay of 0.01. After each round the decay rate reduced to 0.96 rill 20th epochs and later reduces to 0.75. If there is validation loss, then the decay value also reduces to 1/10. In a case study it was proposed that they used Google Collaboratory service for training. It has an accuracy of 94.76% in training and 95.24% in validation.[1]

## VI. IMPLEMENTATION

In a case study in the fashion retail stores, the major challenge is stock inventory management. This costs retailers billable. It is common to find duplicate entries of the same object in huge catalogues. The reasons this happens because the stock was incorrectly marked on arrival or while it got dispatched. It is difficult to manually check for all these entries are there are so many products. Computer vision can be implemented to avoid this situation. The reason why there are supplicated is because of the amount of light, folds, sizes and the way objects photograph has been taken. Microsoft is testing a newly developed computer vision technology including thresholding and tiramisu semantic segmentation which detects the duplicates in the catalogues very accurately. The underlying principle is that it separates the foreground from the background and then compares it with the copies in the catalogues. This helps in saving a lot of time and the staff can-do high-level tasks instead of data entry. By using machine learning functions, the computer vision can check if the object is already present in the catalogue. Computer vision is also used for quality classification of corn tortillas, salmon for fruits.[2][3][7][8]

## VII. APPLICATIONS

Computer vision in retail finds in applications in many places. The first being self-checkout. Instead of waiting in long queues it would be more convenient to customers for self-checkout. It makes customer service automation possible. Self-checkout is nothing, but manual screening of the products and the computer vision cameras detected the products without relying on barcodes. This improves customer experience and a happy shopping.

Computer vision has also found its use in inventory management as the retailers are looking for data driven solutions to optimize inventory.

By installing vision cameras, customers patterns and hot maps can also be detected which improves the store layout management.

Virtual mirrors may enhance customer experience. They can show a range of contextual information which helps the customers to find the products according to their tastes ie personalization. With the scanning connected to the smart phones online ratings and reviews could be brought to the physical states and makes the job very much easier [10].

## VIII. RESULTS AND DISCUSSION

The system design in the methodology has certain assumptions that the product that is feed into the hood at any point has no objects within the field view of camera. The objects are neither small nor big. Lastly the conveyor is stopped by the micro controller whenever the frame is extracted to avoid motion blur.[1]

Australian start-up tilter developed a recognition system that deals with high theft rates. It also eliminated the need for wrapping fruits and vegetables.Abto software has a cashier less technology that enables the computer vision powered unit to process more than 100 items in less than a minute.

Virtual mirrors may enhance customer experience. They can show a range of contextual information which helps the customers to find the products according to their tastes ie personalization. With the scanning connected to the smart phones online ratings and reviews could be brought to the physical states and makes the job very much easier [10].

Amazons just walk out system is also another cashier less technology. It makes use of sensors, cameras and deep learning so customers can walk out by picking up the items they want. The customer's card is automatically charged for items they picked.Selfie uses a computer vision camera mounted on shelves. It alerts staff about gaps, misplaced goods on shelves. Tally is a mobile robot designed by robotics captures visual data and detects damaged packaging and inaccurate pricing.

Samsung uses computer vision and data analytics .it identifies what zones have a huge impact on the conversion when the store was understaffed. This enabled the store to adjust store layout real time. Similarly in fashion retails find mine offers a virtual fitting storeroom. It provides real time fashion recommendations based on their current outfit. Here computer vision detected the outfit, accessories they are wearing and then provides outfits based on an algorithm that complement their looks.Last but not the least Scandits barcode application allows the customers to get information and reviews of a product by just scanning the barcode with the Smartphone camera. It uses the AR guitar centre for its feature.

## IX. CONCLUSION

Computer vision can be used in the retail stores. It makes use of image acquisition coupled with the convolutional neural network with the help of the camera fitted in. One of the long-term goals is to keep it cost effective, efficient, error less. Implementation costs were kept moderate. But working with the web camera has its own cons. One major concern is the blurring. If we use a camera with a higher shutter speed, there will not be stopping of the conveyor belt by the micro controllers. It could be combined with an object detection technique to separate multiple objects in real time. Another long-term goal is to make the whole process automated and self-learning. This allows humans to perform high value tasks rather than the time-consuming data entry. Computer vision can deliver measurable efficiencies and better outcomes. It has a huge scope and could be put to use with improving deep learning technologies and machine learning [1][2][3].

## REFERENCES

[1] Talha Bukhari, S., Wahab Amin, A., Abdullah Naveed, M., &Rzi Abbas, M. (2021). ARC: A Vision-based Automatic Retail Checkout System. arXiv e-prints, arXiv-2104.

[2] https://medium.com/@aryeh_62634/seeing-is-believing-ai-and-computer-vision-for-retail-business-41279f1f048c

[3] https://medium.com/ynap-tech/how-computer-vision-is-making-us-more-effective-at-our-jobs-22748fd70c11

[4] Qiao, S., Shen, W., Qiu, W., Liu, C., & Yuille, A. (2017). Scalenet: Guiding object proposal generation in supermarkets and beyond. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1791-1800).

[5] Do, H. T., & Pham, V. C. (2021). Deep Learning Based Goods Management in Supermarkets. Journal of Advances in Information Technology Vol, 12(2).

[6] Santra, B., & Mukherjee, D. P. (2019). A comprehensive survey on computer vision-based approaches for automatic identification of products in retail store. Image and Vision Computing, 86, 45-63.

[7] Zhang, Y., Wang, S., Ji, G., & Phillips, P. (2014). Fruit classification using computer vision and feedforward neural network. Journal of Food Engineering, 143, 167-177.

[8] Mery, D., Chanona-Pérez, J. J., Soto, A., Aguilera, J. M., Cipriano, A., Veléz-Rivera, N., ... & Gutiérrez-López, G. F. (2010). Quality classification of corn tortillas using computer vision. Journal of food engineering, 101(4), 357-364.

[9] B.-F. Wu, W.-J. Tseng, Y.-S. Chen, S.-J. Yao, and P.-J. Chang, "An intelligent self-checkout system for smart retail," in 2016 International Conference on System Science and Engineering (ICSSE). IEEE, 2016, pp. 1–4.

[10] https://www.itransition.com/blog/computer-vision-retail

11] Ribas Fernandez, J. (2020). Recognition of supermarket products using Computer Vision (Master's thesis, Universitat Politècnica de Catalunya).