Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

J Component Report

Programme : B.Tech CSE With AI & ML (SCOPE)

Course Title : VIDEO ANALYTICS

Course Code : CSE3043

Slot : B2

Title: Video Analytics based on Retail Consumer Behaviour

Team Members: Anshuman Mohanty | 19BAI1156

Reuben Varghese Joseph | 19BAI1133

Soumyae Tyagi | 19BAI1126

Akhil Rudrawar | 19BAI1066

Faculty: Dr. Malathi G

Sign:

Date: 29/4/2022

Scanned with CamScanner

DECLARATION

I hereby declare that the report titled "**Video Analytics based on Retail Consumer Behaviour"** submitted by me to VIT Chennai is a record of bona-fide work undertaken by me under the supervision of **Dr. Malathi G**, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai.

Signature of the Candidate

CERTIFICATE

Certified that this project report entitled "**Video Analytics based on Retail Consumer Behaviour"** is a bonafide work of **Reuben Varghese Joseph (19BAI1133), Anshuman Mohanty (19BAI1156), Soumyae Tyagi (19BAI1126) and Akhil Rudrawar (19BAI1066)** and they carried out the Project work under my supervision and guidance for CSE3043 – Video Analytics.

Dr. Malathi G

SCOPE, VIT Chennai

ACKNOWLEDGEMENT

We wish to express our sincere thanks and deep sense of gratitude to our project guide, **Dr. Malathi G,** School of Computer Science Engineering for his consistent encouragement and valuable guidance offered to us in a pleasant manner throughout the course of the project work.

We are extremely grateful to **Dr. Ganesan R,** Dean of the School of Computer Science Engineering (SCOPE), VIT University Chennai, for extending the facilities of the school towards our project and for his unstinting support.

We express our thanks to our **Head of The Department Dr. Priyadarshani J (for B.Tech-CSE with specialization in AI and ML)** for her support throughout the course of this project.

We also take this opportunity to thank all the faculty of the school for their support and their wisdom imparted to us throughout the courses till date.

We thank our parents, family, and friends for bearing with us throughout the course of our project and for the opportunity they provided us in undergoing this course in such a prestigious institution.

ABSTRACT

AI technologies have been on the rise and been applied to various commercial domains. With recent advancements in video processing technology, we can now record subtle human behavior in a retail shop setting, which is crucial for business management. Customer behavior, when combined with specific demographic data, provides considerably more relevant business intelligence insights to the retail industry. It is key area for any retail scenario as shopping experience plays a vital role in ensuring the success of the company in the context of retail. Besides, it also helps predict which time of the day, a particular store is crowded, composition of male/female population visiting the store, predicting the preference of the customers and subsequently placing advertisements. In our work, we would be working on keeping track of real-time footfall analysis, demographic analysis on the basis of customer age groups and gender detection, traffic detection in particular sections of the shop and motion heatmaps to highlight the areas with the most customer traffic activity (i.e., hot zones). At the end of the work, the results generated suggested that the proposed system could be effectively employed in order to effectively support the analysis of customer experience in the context of retail.

INTRODUCTION

In the digital world, there are a million things that shops must keep track of, and there just isn't enough time in the day. Retailers face a slew of new issues these days, including diminishing sales, strong competition from online-only companies, and shifting consumer tastes. Despite these obstacles, some conventional stores continue to flourish year over year, shattering sales records. The victors are using smart retail analytics to do something unusual, something that not only helps them survive but also thrive in this rapidly developing retail catastrophe.

Retail data analytics is the act of obtaining and analyzing retail data (such as sales, inventory, pricing, and so on) in order to discover trends, predict outcomes, and make better business decisions. When done effectively, data analytics may help retailers obtain a better understanding of their stores', items', customers', and vendors' performance — and then use that information to boost profits. Retailers must be able to successfully target and forecast consumer demands in order to supply the right things at the right price at the right time, which necessitates the use of analytics. By identifying areas for development and optimization, analytics can assist retailers in making the best marketing decisions, improving their company operations, and providing better overall consumer experiences. The retail industry may benefit from analytics in a variety of ways. To obtain a competitive edge, retailers may utilize retail analytics to forecast customer requirements and make company changes. Retailers may use sales data to spot developing trends and anticipate client demands.

The discovery, analysis, and transmission of relevant patterns in data derived from video material, as well as the application of those patterns to successful decision-making, are all part of retail video analytics. Today, retail video analytics extends beyond security and loss prevention to provide retailers with actionable business knowledge. Computer vision technology has led to the development of retail video analytics software. The programme can count individuals and generate heat maps to highlight which areas of the store are busiest (and which areas are being ignored). This helps store managers to properly organise their stores and determine the busiest times of day for scheduling purposes. Retail video analysis insights may also be utilised to develop successful promotions and display designs that drive greater interaction. Retailers can better link promotions, displays, and sales by monitoring trends in engagement with product displays. It may even be able to find fresh cross-merchandising concepts to entice clients. The layout of a retail business is a critical aspect in creating a positive shopping experience. Retailers may improve product placement and in-store traffic patterns to increase sales by studying how consumers normally traverse around the shop and how long they stay in various spots.

Retail video analysis software enables brick-and-mortar retailers to gain a deeper understanding of in-store customers in order to customise interaction and increase retention and sales. The software data may be utilised to learn how to better engagement with various demographics in order to expand the consumer base, support merchandising strategy, or locate new pop-up stores. Instead of guessing about client demographics, merchants can use video analytics data to figure out who they're attracting and better tailor customer experiences based on the information acquired. Video technology may assist in providing a positive customer experience during their visit to the shop. Peak hours, traffic patterns, and high-occupancy zones may all be identified using video data collecting and analytics. Given the above use cases and advantages, video analytics appears to be a must-have for every retail company looking to improve customer experience and bottom lines.

However, they come with their own set of challenges such as occlusion scenario, non-uniform distribution of illumination, variations in appearance for intrascene and interscene, scale and perspective etc. We have made an attempt in order to reduce the effect of these problems and provide detailed insights to the retailers.

MOTIVATION

The main objective of this work is to support the retailers in devising effective strategies by supporting their decisions with the required insights or data. Video analytics provides technical solutions to queries like "How many people visited my store today/this week/this month?" or "How many customers exhibited interest in the item on sale?". Information offers retailers with useful insights that allows them to enhance merchandising/marketing and improve the consumer experience, resulting in increased profitability. Footfall analysis is a crucial metric for retailers because it enables them to compute the store conversion rate (i.e., the percentage of visitors converted to buyers). The conversion rate is an important metric for retailers because it indicates how well a store is functioning. By having knowledge on hot zones, Store managers can optimize the layout of their stores for better product placement. This information can also be used to assess or improve the performance of sales or advertising displays.

LITERATURE SURVEY

In 2020, Qingyang Xu,Wanqiang Zheng,Xiaoxiao Liu, and Punan Jing proposed a deep learning framework that uses a skeleton detection algorithm to recognize the human skeleton, which outperforms the object detection algorithm in human detection. The skeleton data can be used to determine the outline of a human. Then, based on the store's surveillance video, this paper employs a compound deep learning technique to generate crowd counting and density maps. Following this, vendors can predict consumer preferences and modify the product arrangement based on customer information. The goal of human tracking is to follow the human motion trail in order to predict motion or to draw a density map. As for the drawback, the ability of the model to predict depth correctly in the proposed method has a significant impact on the accuracy and effectiveness of the proposed frameworks.

In 2016, Lokesh Boominathan, Srinivas S S Kruthiventi, and R. Venkatesh Babu proposed a novel deep learning approach for predicting crowd density from static images of densely packed crowds. To predict the density map for a given crowd image, they used a combination of deep and shallow fully convolutional networks. A combination like this is used to effectively capture both high-level semantic information and low-level features, which are required for crowd counting under large scale variations. They also demonstrated that the challenge of varying scales, as well as the inherent challenges in dense crowds, can be effectively addressed by augmenting the training images. The disadvantage is that the model's ability to predict correctly is heavily dependent on the training dataset.

In 2008, Antoni B. Chan, Zhang-Sheng John Liang, and Nuno Vasconcelos proposed a privacy-preserving system for estimating the size of inhomogeneous crowds made up of pedestrians traveling in different ways, without the use of explicit object segmentation or tracking. Using the mixture of dynamic textures motion model, the crowd is first segmented into homogeneous motion components. Second, from each segmented region, a set of simple holistic features is extracted, and the correspondence between features and the number of people per segment is learned using Gaussian Process regression. It was discovered that effective crowd segmentation is required for successful crowd counting. Concerning the disadvantage, the model's ability to predict correctly is highly dependent on the positioning of the camera taking input and the pre-processing applied to that input.

A Fully Convolutional Neural Network for Predicting Human Eye Fixations was proposed by Srinivas S S Kruthiventi, Kumar Ayush, and R. Venkatesh Babu in 2015. A fully convolutional neural network for accurate saliency prediction that is the first of its kind unlike previous works that categorise the saliency map using numerous hand-crafted features, our algorithm accurately learns features in a hierarchical fashion and predicts the saliency map from start to finish. Using network layers with very large receptive fields, this model is designed to collect semantics at different scales while taking global context into account. Introducing LBC - Location Biased Convolutional Filters, a novel technique that allows the deep network to learn location dependent patterns, though environmental conditions such as lightning in the dataset also affect accuracy.

In 2019, Osama T. Ibrahim, Walid Gomaa, and Moustafa Youssef proposed A Deep Learning System for Device-free Human Counting via WiFi. It is a precise and reliable deep-learning-based human count estimator that estimates the human count in an area of interest using a single WiFi link. The central idea is to use the temporal link-blockage pattern as a discriminant feature that is more resistant to wireless channel noise than signal strength, resulting in a ubiquitous and accurate human counting system. The model addresses a number of deep learning challenges, such as class imbalance and training data augmentation, as part of its design to improve model generalizability. As for drawback the accuracy of model highly depends on availability of Wi-Fi in area of interest.

Aibek Musaev, Jiangping Wang, Liang Zhu, Cheng Li, Yi Chen, Jialin Liu, Wanqi Zhang, Juan Mei, and De Wang presented a study in 2020 that looked at the topic of computer vision-based client monitoring in the retail business. To that purpose, they presented a dataset gathered through a webcam in an office setting, in which participants imitated various grocery consumer behaviours. They also provided an example of how this dataset may be used to follow participants using a head tracking model to reduce occlusion errors. They also presented a method for identifying consumers and employees based on their movements. A real-world dataset gathered over a 24-hour period in a supermarket is used to assess the model. The model is tested using a real-world dataset gathered over a 24-hour period in a supermarket and found to be 98 percent accurate during training and 93 percent accurate during assessment.

Harikrishna G.N. Rai, Kishore Jonna, and P. Radha Krishna offered a paper in 2011 in which they presented a computer vision-based system for monitoring client whereabouts inside shopping malls by identifying individual shopping carts in order to support location-based services. They proposed a two-stage technique to cart recognition, consisting of cart detection and then cart recognition. For identification, a binary pattern is inserted between two pre-defined colour markers and fastened to each cart. The system receives a live video feed from the cameras positioned along the shopping mall's aisles and analyses frames in real time. Color segmentation, feature extraction, and classification are utilised in the cart detection step to detect binary patterns and colour indicators. To decode the cart identifying number, a segmented binary strip is processed utilising spatial image processing algorithms at the recognition step.

In 2018, Andrea Generosi, Silvia Ceccacci et al. proposed a deep learning framework to determine the customers' age, sex, ethnicity, and emotions by using convolution neural network which takes different frames of a video stream as input and provide an estimation of the aforementioned details. The system is modular and does not implement any invasive solutions by allowing the monitoring of the emotive status of customers without being aware of it. Easy and remotely accessible web-based interface allows users to access data in a cloud-based environment securely. The accuracy and effectiveness of the proposed frameworks are strongly affected by environmental conditions (e.g., illumination conditions, distance between customer and camera).

In 2020, Earnest Paul Ijjina, Goutham Kanahasabai et al. proposed a deep learning framework for determining various demographics such as customer age, gender, and expressions. Wide Residual Networks and Xception deep learning models was used to predict age, gender, and facial expressions of consumers to devise effective marketing strategies in order to maximize sales. The effectiveness of the proposed approach is evaluated on real-life garment store surveillance video, which is captured by low resolution camera, under non-uniform illumination, with occlusions due to crowding, and environmental noise. Using the proposed framework, a high accuracy is achieved. The model achieves significantly higher accuracy when the subject faces towards the camera and the performance drops drastically when customers do not face the camera directly.

In this paper Valerio Nogueira Jr., Hugo Oliveira et al. proposed a low-cost deep learning approach to real-time count of people in retail stores in real-time and visualize hot spots. They use a supervised learning strategy based on a CNN model to address the people counting problem by using RGBP image to clearly observe humans in the images. Using merely a cheap security camera, this article proposed a low-cost deep learning strategy for estimating the number of people in retail outlets and detecting hot zones in real-time. According to the findings, it is observed that a trained model cannot be used in many locations or views inside the same location. In other words, training is tailored to each location and point of view. To avoid viewpoint shifts, the surveillance camera must remain stationary.

MODULE DESCRIPTION

Gender Demographic Analysis

In this module, we will examine the demographic information of people in a retail store. Because customer demographics primarily dictate their preferences, identifying and utilizing customer gender information in sales forecasting may maximize retail sales. We would draw a box around the people and indicate their gender. This would aid in the promotion of gender-based marketing. Gender marketing can be successful if individuality and diversity are considered, which this module will do so by collecting gender information about purchased products.

Traffic Counting system

In this module, we would count the number of people in a retail store in a specific section at a specific time. It would provide detailed analytics on the number of people visible at various time intervals in various sections. Using this information, retailers can determine the peak hours for customer visits to a specific section. It will not only assist owners in product management in various sections, but it will also assist them in better advertising different products. It would also provide information on the sales of various products at different times of day, which would aid in the better management of products with a short shelf life.

Traffic Color Density Heatmap

The heat maps depict the business, establishment, or public location where the most people have passed. It is used to examine consumer traffic patterns in a retail establishment. It enables us to identify the parts of the shop where our consumers pass, pause, and pay closer attention, as well as the so-called "dead zones of the store," or regions where customers do not normally pass. In the form of a colour density heatmap, it shows where the most traffic is focused in the business. The software creates a heat map using these data, with the red tones indicating the areas with the most traffic and high consumer attention, the orange areas a medium-high attention, the green areas a medium-low attention, the areas blue low attention, and the unpainted areas indicating the dead zones, where consumers practically do not pass. Retailers may use these figures to better understand how to optimise product positioning and the efficacy of promotional displays.

Footfall Counter

We'll be tracking people's footfall at a retail store in this module. It would offer detailed statistics on the number of persons seen at various time intervals. Many shops are missing critical data points such as precise counts of consumer foot traffic at any one time and knowing how customers move across a store. These kinds of data may be used to improve not only product placement in stores, but also personnel levels and even safety. Retailers may also acquire the peak hours of consumer visits in the shop using this data. These data may then be used to augment store inventory planning, product placement strategy, better evaluate personnel needs, and even guarantee that the number of consumers in the shop does not exceed the legal limit. These insights may also be used to examine regular travel patterns, fine-tune staffing to account for peaks and troughs in foot traffic, and learn how external variables like as holidays and weather affect foot traffic.

PROPOSED METHODOLOGY

Traffic Color Density Heatmap

First, we define a video capture method using cv2.VideoCapture(). The frames of the video are captured by the read() method. Each of the frames read are supplied for background subtraction using cv2.bgsegm.createBackgroundSubtractorMOG(). It is done so that a mask could be generated. Then thresholding is performed to the mask to remove the small number of movements and to set the accumulation value for every iteration. The result of the threshold is added to an accumulation image (one that starts out at all zero and gets added to each iteration without removing anything), which is what records the motion. Finally, a colour map is applied to the gathered image to make the motion more visible. Finally, a colour map is applied to the gathered image to make the trails of the human motion more visible. In order to achieve the overlay result, the current-coloured image and the previous coloured image are combined with the first frame in order to achieve the overlay.

Footfall Counter

To implement In-Out footfall counter, we have made use of MobileNet SSD as it was trained to detect people. The model is loaded from the disk and the video capture method is defined using cv2.VideoCapture() method. We have made use of centroid tracking mechanism in order to detect the people. For centroid tracking, we accept a set of bounding boxes and compute the centroids as shown in Fig. 1.

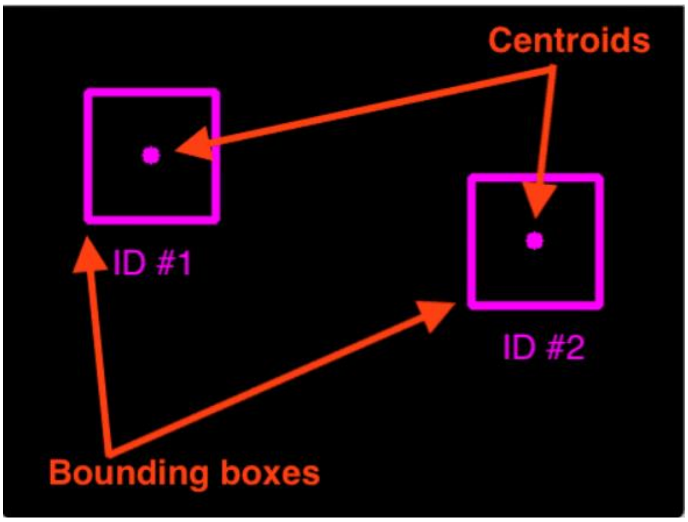


Fig. 1. Bounding boxes around the person with centroid points

The bounding boxes are created using SSDs. The next step is to compute Euclidean distance between the newer centroids (highlighted in yellow in Fig. 2.) and existing centroids (highlighted in purple in Fig. 2.).

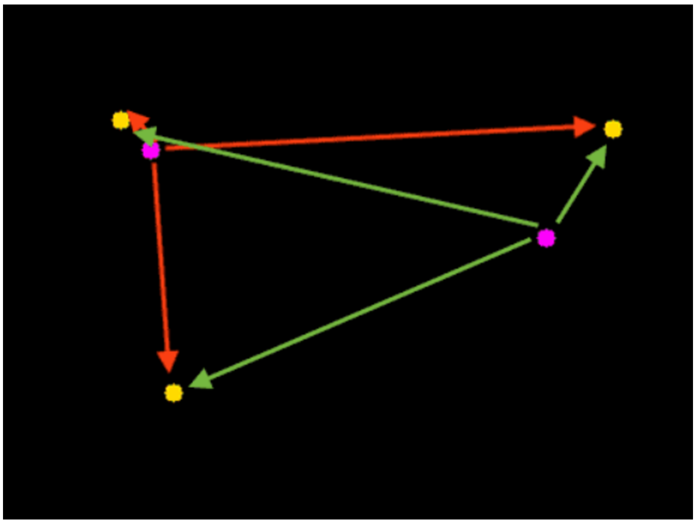


Fig. 2. Computing the distance between the newer and existing centroids

The algorithm works on the assumption that the centroid with minimal Euclidean distance must have the same object ID. On having the Euclidean distances, we attempt to associate object IDs.

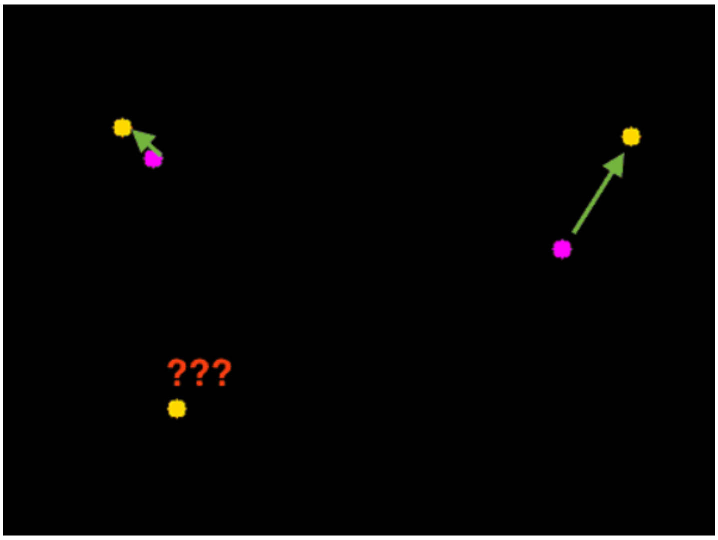


Fig. 3. Grouping centroid points which are in proximity of each other

The object which are not associated (like the bottom left yellow point in Fig. 3.) with any other points are registered as new objects (Fig. 4.). We store the new object IDs along with the centroid of the bounding box coordinates for the new object.

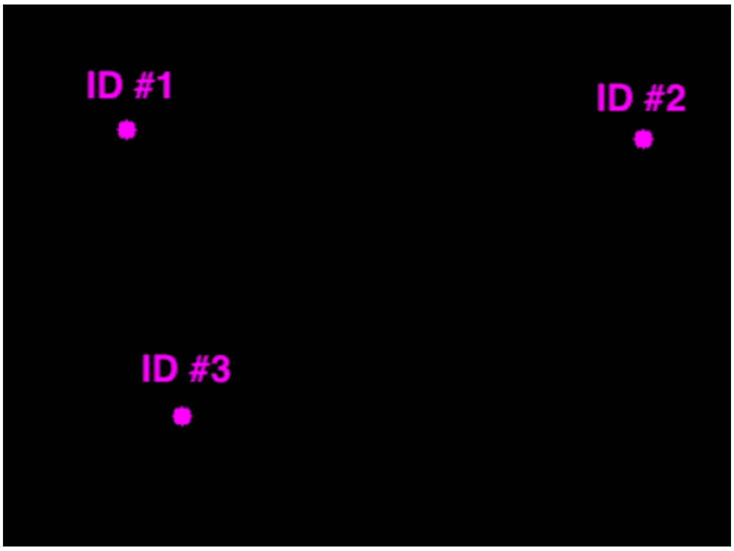


Fig. 4. Assigning IDs to the grouped points

The objects that are lost or are not visible in the field of view are deregistered. The person IDs are deregistered when it could not be matched to any existing person objects for 40 consecutive frames in the video.

In order to track and maintain a count of an object in the video stream, we store information such as: - object ID, previous centroids and whether the object has already been counted.

We set a confidence value of 0.4 which is the filter out the weaker detections. We initialize a ‘status’ variable. The possible status of the states include: -

Waiting – This state refers to the phase where we are waiting for people to be detected.

Detecting - This state refers to the phase where people are being detected using MobilenetSSD.

Tracking - This state refers to the phase where people are tracked in the frame and the direction in which they are going (i.e., going up or coming down).

The horizontal visualization is the marker which demarcates which is the outside portion and what constitutes inside portion.

Traffic counting system

Performing DeepSort Object Tracking with YOLOv4 object detection. YOLOv4 is a cutting-edge object detection technology that employs deep convolutional neural networks. To develop a highly accurate object tracker, we can take the output of YOLOv4 and input these object detections into Deep SORT (Simple Online and Realtime Tracking with a Deep Association Metric). It provides each person visible in the frame with a tracker ID along with drawing bounding boxes around that specific person.

Gender Demographic analysis

Gender Detection is performed by the Open cv Face detector. To detect the faces, we have set the confidence value as 0.7. All the values which lie above the threshold would be considered as the bounding box around the detected face. We have made use of the pretrained gender_net caffemodel in order to provide accurate predictions on the gender of the person detected in the box. The resultant prediction is displayed on top of the box.

The Architecture for the proposed methodology is shown in Fig. 5.

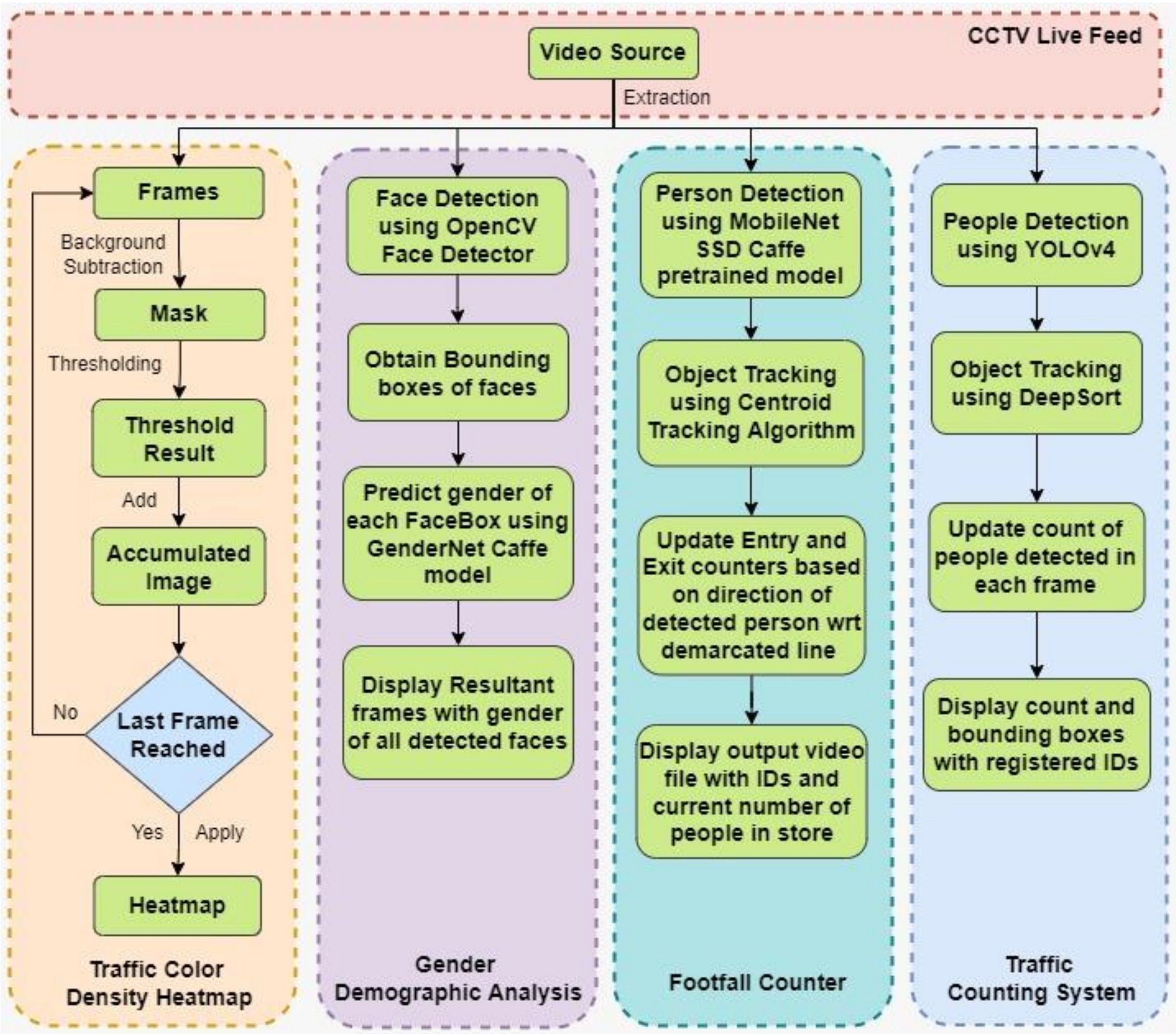


Fig. 5. Proposed Methodology Architecture

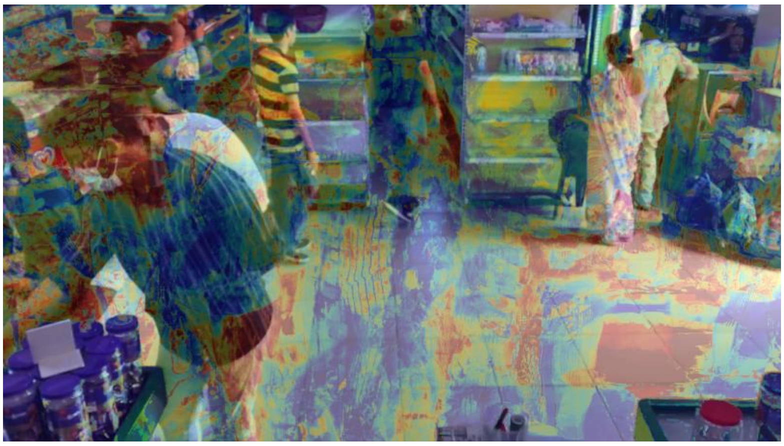
DATASET

We have collected our own dataset from a local store present within the university campus. The videos were captured from smartphones which were placed at an angle comparable to the ones provided by the CCTV cameras to depict the real-world scenario. One of the videos was taken facing towards the shelves and two other videos were captured facing towards the entrance. Each of the video clips were of roughly two minutes duration each.

RESULTS

The following results were observed after analysis of various video footage for Traffic Color Density Heatmap, footfall counter, gender demographic analysis and Person detection. It was observed that the Motion heatmap generation (as in Fig. 6.) was performing successfully as the overlay of colors were visible thereby depicting the density of the traffic across the store.

(a)



(b)

Fig. 6. a) Original Video Clip Frame b) Motion Heatmap Video Clip Frame

The Footfall counter was able to track the individuals by providing them IDs along with providing them the centroid points. It was observed when people crossed the line, based on the movement of the centroid point across the line, the people were considered as entering or exiting the store. The total count of people present in the store was calculated as the difference between the people entering and exiting the store, and displayed as shown in Fig. 7.

Fig. 7. Foot-Fall Counter

For gender detection, the video was zoomed in slightly in order to be detected by face detection algorithm. It was observed that the faces were detected, and the gender prediction was displayed on the top left corner of the bounding box. However, the approach detects only the faces which are sufficiently closer to the camera and are frontal facing. The results could be visualized in Fig. 8.

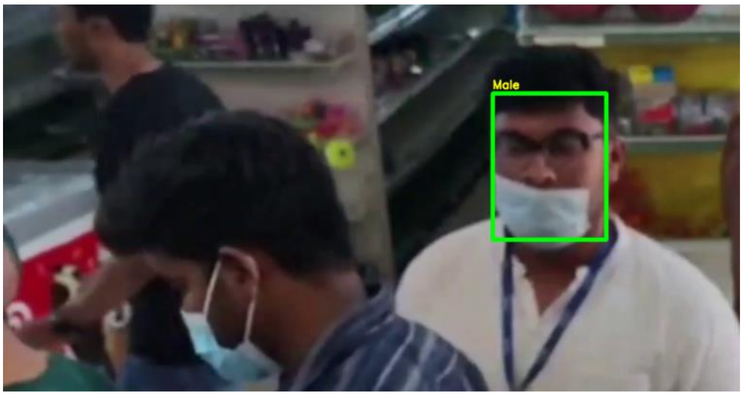


Fig. 8. Gender Demographic analysis

For Traffic counting system YOLOv4 was employed. The model was successfully able to detect the people and enclose them in the bounding boxes even if the person was not entirely visible in the camera. A sample frame is displayed in Fig. 9.

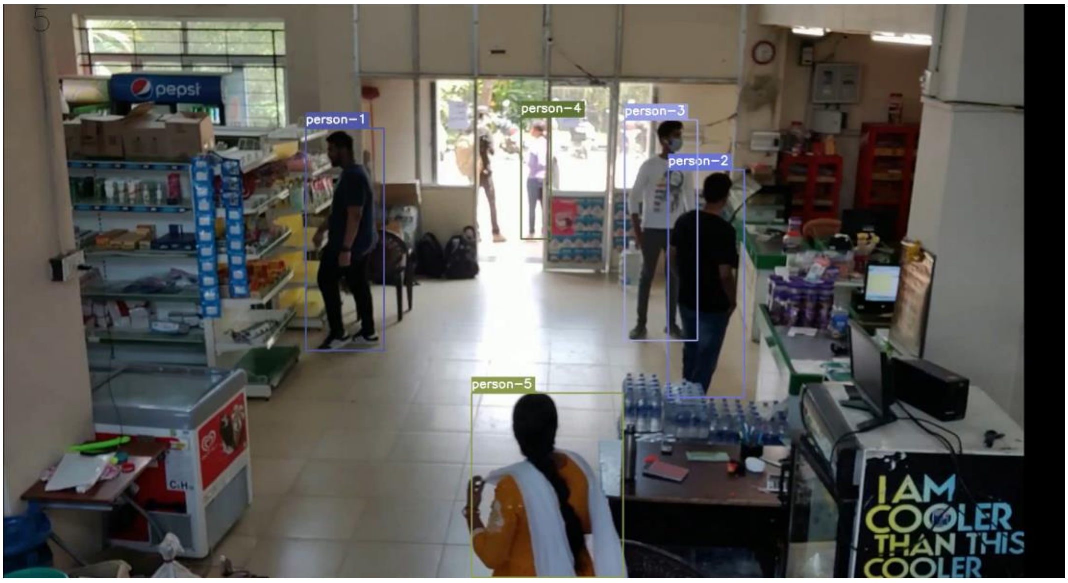


Fig. 9. Person detection

CONCLUSION AND FUTURE WORKS

Surveillance that is intelligent Video analysis has been used in a variety of fields, including transportation, security, and commerce. Object detection, particularly human detection, is the most important aspect of video analysis. Our work proposes a combined retail video analytics system comprising of four modules which cover the important aspects of retail stores such as advertisement placements, targeted marketing of products etc. It was an attempt to provide a low cost deep learning solution which could be integrated with the CCTV camera of the retail stores. The results obtained suggest that our approach tends to work well in the scenario of retail setting.

As future work, we would like to present an age detection module, as detecting age-groups entering or exiting the store provides various useful insights to the retailer. Additionally, we would want to devise strategies in order to remove the presence of salespeople from contributing to the people count and providing accurate insights to the retailer in order to make an analysis on the customer-only flow. It is difficult to detect age and gender detection of people in real time video footage and hence we would be looking forward to making the analysis via gait detection or Openpose in order to efficiently detect and predict the age and gender of the person. We also intend to test deep neural networks that takes into account the problem's temporal coherence, as people's counting habits alter over time. Another future study area would be introducing layers capable of automatically recognising the foreground in an integrated manner into the network architecture.

REFERENCES

[1] Connell, J., Fan, Q., Gabbur, P., Haas, N., Pankanti, S., & Trinh, H. (2013). Retail video analytics: an overview and survey. Video Surveillance and Transportation Imaging Applications, 8663, 260-265.

[2] Senior, A. W., Brown, L., Hampapur, A., Shu, C. F., Zhai, Y., Feris, R. S., ... & Carlson, C. (2007, September). Video analytics for retail. In 2007 IEEE Conference on Advanced Video and Signal Based Surveillance (pp. 423-428). IEEE.

[3] Choudhary, A., & Chaudhury, S. (2016). Video analytics revisited. IET Computer Vision, 10(4), 237-247.

[4] Rai, H. G., Jonna, K., & Krishna, P. R. (2011, August). Video analytics solution for tracking customer locations in retail shopping malls. In Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 773-776).

[5] Musalem, A., Olivares, M., & Schilkrut, A. (2016). Retail in high definition: Monitoring customer assistance through video analytics. Columbia Business School Research Paper, (15-73).

[6] Musaev, A., Wang, J., Zhu, L., Li, C., Chen, Y., Liu, J., ... & Wang, D. (2020). Towards in-store multi-person tracking using head detection and track heatmaps. arXiv preprint arXiv:2005.08009.

[7] Rai, H. G., Jonna, K., & Krishna, P. R. (2011, August). Video analytics solution for tracking customer locations in retail shopping malls. In Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 773-776).

[8] Singh, H. (2018). Applications of intelligent video analytics in the field of retail management: A study. In Supply Chain Management Strategies and Risk Assessment in Retail Environments (pp. 42-59). IGI Global.

[9] Kaur, J., Arora, V., & Bali, S. (2020). Influence of technological advances and change in marketing strategies using analytics in retail industry. International journal of system assurance engineering and management, 11(5), 953-961.

[10] Olatunji, I. E., & Cheng, C. H. (2019). Video analytics for visual surveillance and applications: An overview and survey. Machine Learning Paradigms, 475-515.

[11] Ananthanarayanan, G., Bahl, V., Cox, L., Crown, A., Nogbahi, S., & Shu, Y. (2019, June). Video analytics-killer app for edge computing. In Proceedings of the 17th annual international conference on mobile systems, applications, and services (pp. 695-696)

[12] Gagvani, N. (2009). Challenges in video analytics. In Embedded Computer Vision (pp. 237-256). Springer, London.

[13] Ganesan, V., Ji, Y., & Patel, M. (2016). Video meets the Internet of Things. McKinesy Quarterly (December).

[14] Zhang, H., Shen, M., Huang, Y., Wen, Y., Luo, Y., Gao, G., & Guan, K. (2021). A Serverless Cloud-Fog Platform for DNN-Based Video Analytics with Incremental Learning. arXiv preprint arXiv:2102.03012.

[15] Zalud, B. (2012). Video analytics: big retail deal with big data. Integrated Solution: Video.

[16] Battiato, G. E., Cavallaro, A., & Distante, C. (2016). Special issue on" Video analytics for audience measurement in retail and digital signage". Pattern Recognition Letters, 81, 1-2.

[17] Larsen, N. M., Sigurdsson, V., & Breivik, J. (2017). The use of observational technology to study in-store behavior: Consumer choice, video surveillance, and retail analytics. The Behavior Analyst, 40(2), 343-371.

[18] Prabhakaran, B., & KALVA, H. (2018). Guest Editorial Special Section on Video Analytics: Challenges, Algorithms, and Applications. IEEE Transactions on Multimedia, 20(5), 1037.

[19] Gibbon, D., & Begeja, L. (2013). Distributed processing for big data video analytics. E-LETTER, 15(5).

[20] Distante, C., Battiato, S., & Cavallaro, A. (2014). Video analytics for audience measurement. Stockholm, Sweden: Springer.

[21] Yu, T., Zhou, B., Li, Q., Liu, R., Wang, W., & Chang, C. (2009, November). The design of distributed real-time video analytic system. In Proceedings of the first international workshop on Cloud data management (pp. 49-52).

[22] Guevara, E. C. (2015). Multiple camera video surveillance for customer behaviour analysis in a retail store.

[23] Dailey, M. N., Esichaikul, V., & Ekpanyapong, M. (2017). Android Application for Customer Service Video Analytic Systems.

[24] SÜMER, S. I., ÖZGÜRBÜZ, G., & SÜMER, E. (2018). EXTRACTION OF CUSTOMER DEMOGRAPHIC CHARACTERISTICS IN SUPERMARKETS BASED ON IMAGE PROCESSING TECHNIQUES. The Online Journal of Science and Technology-October, 8(4).

[25] Katanyukul, T., & Ponsawat, J. (2017). Customer analysis via video analytics: customer detection with multiple cues. Acta Polytechnica Hungarica, 14(3), 187-207.