Universiti Malaysia PAHANG
Engineering • Technology • Creativity

ORIGINAL ARTICLE

Development of Consumers Management System via Face Classification

N. I. Suhardi[1], U. Khairuddin[1*]

[1]Centre for Artificial Intelligence and Robotics i-Kohza, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

ABSTRACT – To improve consumer experience and overall retail management, physical retailers may adapt consumer behaviour management systems using artificial intelligence to imitate the capability of consumer behaviour tracking in online shopping into physical retail. The proposed consumer behaviour management system consists of two parts - face recognition and consumer tracking at an area of interest. Both will be combined to produce a summary of individual customer's visits to the shop. This information can be used to improve consumers experience and optimize retailer's management. The developed system can track consumers' movement inside the shop and summarize their whereabouts according to areas of interest. The face classification system via FaceNet has around 56.67% accuracy with 27.89% mean confidence. The tracking performance shows a consistent performance with a total standard deviation of 4.36 seconds. With the consumers' analysis graph, retailers may pinpoint which area that was always frequented by their customers and take suitable actions with that information.

INTRODUCTION

E-commerce has been seen rapidly growing as recent technology advances, but as e-commerce strives, physical retail is deemed unstable. As of early April 2017, more than 2000 retail stores have been closed down at the same time, and some physical stores strive more such as Target and Walmart [1]. Even when e-commerce became the new shopping norm, physical retail stores still stayed relevant. A review made by Gauri et al. found an amount of evidence that offline and online stores have the same importance as the services are still in demand [2]. Despite the need for physical shops, the current management of physical retail stores is yet to be at par with their online counterpart. Current consumer management of physical stores still lacks compared to their online counterparts, where the latter can track their customer behaviour, optimize management and increase overall revenue.

A solution can be made to the problem faced by physical retailers in analysing their customers' behaviour. A machine vision system consisting of hardware and software components can be developed to detect faces and classify them for consumers' management analysis for the retailers. The system consists of a surveillance camera(s) for face detection and face recognition. It works by first detecting the face of visitors who enter the premises, and then if the visitor is a registered member, it will tag the visitor as a member. The system will detect faces (whether they are a member or first-time customer/non-member) and track their movement. It will record the item they spent time on, how long they spend time checking the item, and how long they spend in-store. Each consumers' individual recordings will be analysed and gathered in a big consumer's behaviour database. The retailers can use this information to strategize and strengthen their sales and marketing. This record will later be summarized into a trend for the retailers to decide their next move such as informing their customer if the item they have shown interest in is restocking or reduced in price. With that, we will need two algorithms, one for customer tracking and one for face recognition.

Thus, in this project we will use Single-Shot Detector (SSD) for customer tracking and FaceNet, a face recognition algorithm for face classification. As the name suggests, SSD only requires one shot to detect multiple objects within one image [3]. The usage of SSD has been implemented in many applications such as object finding by drones [4] and blood cell detection for biomedical research [5]. Meanwhile, FaceNet has been used for face recognition purposes as it learns directly from mapping face images to measure distances that correspond to measure face similarity [6].

RELATED WORK

Similar work has been done to consumer management systems such as Perksy and Just Walk Out technology by Amazon. Perksy, an American-based company, offers a solution to the described problem by asking their customers about their store via the Perksy app equipped with geo-track [7]. How the app works for retailers; they let retailers create custom research questions, select the targeted customers (by demographics, social data, location, and other criteria.), and whenever a customer responds to the survey, retailers can get real-time results [8]. The app has gained many approvals and reviews so far. However, brands can only get the results if customers chose to answer the survey. Even if brands can get a good analysis from the surveys, they can only identify a good percentage of their overall customers.

Just Walk Out technology enables their customers to simply enter the store, grab what they want and need, and just go [9]. This technology is first used in Amazon stores named Amazon Go. This technology automatically detects when products are taken from or returned to the shelves and keeps track of them in a virtual cart. When the customers are done with their shopping, they can leave the store, and the receipt will be charged to their Amazon account [10]. Current optimum opposition-based simulated Kalman filter

METHODOLOGY

The consumer behaviour management system consists of two main components; i) face classification algorithm and ii) customer tracking algorithm.

Face classification algorithm

The face recognition model needs to be trained with images of a named face. In this project, nine datasets are collected with 50 images containing images of faces with assigned names. One dataset named Unknown with 50 random face images varies from gender, skin colour, and facial features. Seven known datasets are images gathered from Google images of popularly known celebrities. Another two datasets are from personal videos that are later broken down into frames and take 50 random images out of the total frames from the video. The program for splitting images from the video was acquired from hernanrazo in GitHub [12]. The images of the datasets do not have a specific image size.

These datasets are later than trained in FaceNet. A video was developed with nine face images to test this detection along with customer tracking, following the number of IDs detected in person tracking algorithm. In the video developed, five were Unknown (random face images) and four registered people. The video was developed using Sony Vegas Pro 17.0, and the duration of the video is equal to the duration of the CCTV video tested for the person tracking algorithm.

Customer tracking algorithm

In this process, a CCTV video feed was used. The video feed was acquired from YouTube [11]. The tracking aims to track customers' behaviour in the store in which to track are identified. For example, in Figure 1, the Cold drinks aisle is labelled as Area (1) and Magazine's aisle is labelled as Area (2). The xy coordinates of these areas are then collected by drawing the box on the video frame. These areas are the Region-of-interest (ROI).

Figure 1. Areas that are identified in CCTV video feed

The video with labelled ROIs was then tested with the SSD along with a Centroid Tracker. When SSD detects an object, it created a bounding box. It is easier to track with a single-point tracking To ensure a person enters the distinctive ROI. Centroid Tracker creates one-point of the bounding boxes, and when the single point crosses the ROI, a timer will start, measuring the time taken for the person inside the area. The timer stops when the person leaves.

The information on how long the customer spent time at the specific aisle and how many customers visited the aisle; were stored under the generated ID of the person. In this video, nine IDs were generated.

Overall system

The system runs two algorithms simultaneously. The proposed flowchart is both true for testing and the proposed system using real-time CCTV positioned high in a retail store and one camera located at the cashier for face recognition.
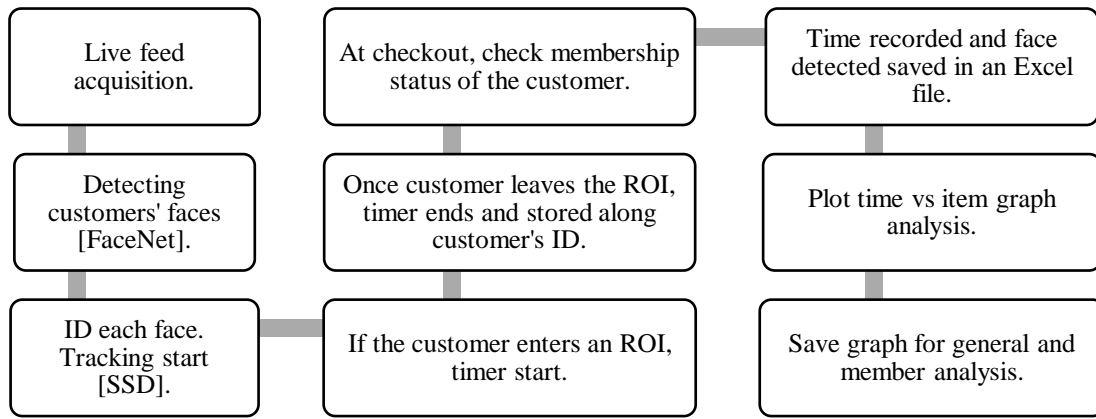
Figure 2. Flowchart of the proposed system for a consumer management system with face classification

RESULTS AND DISCUSSIONS

Table 1 shows the performance of the face classification via FaceNet. The performances observed were the output and label matching (5.1 true matches out of 9) and the highest confidence percentage of 39.99%, which is below average. However, the margins between each person's confidence percentage are quite high (26.6% to 73.7%), which means there is potential for improvement for the system if more iterations were added in the FaceNet algorithm.

Table 1. Face classification result

Trials	Tested Face Confidence (%) | Face labelled correctly (√/X)																	
	1		2		3		4		5		6		7		8		9	
1	29.8	√	28.5	X	26.1	X	83.3	√	55.3	√	22.6	√	33.1	√	35.0	√	48.7	X
2	30.6	√	29.6	X	26.1	X	72.8	√	55.0	√	24.4	√	32.6	X	38.7	√	54.5	X
3	30.9	√	29.6	X	27.2	X	72.6	√	55.8	√	24.4	√	29.8	X	36.4	√	50.1	X
4	29.6	√	30.0	X	27.1	X	72.4	√	56.0	√	24.5	√	29.4	X	39.5	√	54.9	X
5	29.8	√	29.6	X	27.1	X	72.5	√	55.9	√	24.3	√	29.8	X	36.3	√	54.9	X
6	30.5	√	29.6	X	27.1	X	72.8	√	55.8	√	24.2	√	32.1	X	36.4	√	45.5	X
7	30.6	√	29.7	X	26.1	X	72.6	√	56.0	√	24.4	√	29.4	X	36.3	√	54.7	X
8	30.6	√	29.7	X	26.1	X	72.6	√	55.1	√	24.1	√	29.6	X	36.3	√	55.0	X
9	30.6	√	29.7	X	26.1	X	72.4	√	55.1	√	24.1	√	29.6	X	36.3	√	55.0	X
10	29.5	√	29.6	X	27.2	X	72.7	√	55.0	√	24.2	√	29.8	X	38.4	√	54.6	X
Mean | No. of correct faces	30.25	10	29.56	0	26.62	0	73.67	10	55.5	10	24.12	10	30.52	1	36.96	10	52.79	0

Figure 2. Detected faces result. The image on the right side shows the correct face detection of 30.11% while the image on the left side shows incorrect detection with a confidence percentage of 28.73%

In Table 2, the consumer tracking algorithm via SSD tracks nine customer's whereabouts at each area of interest. The grey boxes indicate that the ID did not cross the ROI during the tracking process.

Table 2. Consumer Tracking result via Single Shot Detection (SSD)

Variable Customer	Trial	Recorded Data								
		1	2	3	4	5	6	7	8	9
Other areas	1	67.43	360.51	28.02	107.59	21.51	25.63	122.9	55.28	50.98
	2	61.13	333.41	24.71	93.71	22.93	22.28	101.7	43.81	44.57
	3	55.24	337.17	23.97	91.59	21.87	23.56	95.57	42.6	42.52
	4	54.49	308.21	23.05	88.25	22.54	23.6	101.18	42.19	42.1
	5	54.44	343.89	23.1	88.5	22.62	21.32	92.14	43.77	39.57
	6	63.6	297.87	25.59	88.15	19.38	22.21	92.61	39.39	39.68
	7	58.92	359.66	22.33	102.2	29.52	65.35	261.99	108.81	140.94
	8	49.36	269.88	20.27	77.19	20.31	20.58	87.39	35.15	34.6
	9	58.16	280.82	21.54	75.24	21.35	23.63	98.22	49.12	46.6
	10	68.22	351.88	24.86	76.95	19.12	18.61	78.65	34.61	34.21
	Mean	**59.1**	**324.33**	**24.04**	**88.94**	**22.12**	**26.68**	**113.24**	**49.47**	**51.58**
Area 1	1	14.3	44.68		8.63		106.6	0.37	62.56	11.37
	2	12.61	39.62		6.8		86.1	0.2	54.47	9.74
	3	11.75	36.75		6.41		82.09	0.15	51.82	9.35
	4	11.2	36.3		6.64		93.08	0.21	51.27	9.5
	5	12.04	36.18		5.56		80.35	0.18	48.38	9.01
	6	14.46	37.18		6.45		76.89	0.17	48.7	9.08
	7	11.9	43.35		7.61		193.9	0.53	204.69	40.27
	8	11.17	32.2		6.14		80.23	0.22	42.62	7.92
	9	12.39	38.13		5.77		80.44	0.15	56.49	10.12
	10	14.29	42.2		5.5		66.75	0.14	41.95	7.72
	Mean	**12.61**	**38.66**		**6.55**		**94.64**	**0.23**	**66.3**	**12.41**
Area 2	1	1.13	87.52		114.41	22.17	0.56	2.58	8.55	1.34
	2	1.12	78.7		99.59	29.56	0.53	2.08	6.3	1.3
	3	0.94	75.64		109.62	31.17	0.69	1.98	6.0	1.19
	4	0.85	73.97		94.57	29.01	0.7	3.01	6.21	1.14
	5	0.89	74.67		82.39	29.67	0.47	2.11	6.01	1.19
	6	1.11	77.51		96.79	29.94	0.46	1.88	5.63	1.09
	7	1.08	78.99		107.32	35.41	1.34	3.79	21.29	3.94
	8	0.82	65.03		82.37	25.21	0.63	2.65	4.99	0.97
	9	1.73	73.24		81.41	25.6	0.62	1.7	6.58	1.37
	10	1.26	82.23		81.99	26.6	0.61	1.66	5.03	0.96
	Mean	**1.10**	**76.75**		**95.05**	**28.33**	**0.66**	**2.34**	**7.66**	**1.45**

The Face classification result was then integrated with the consumer tracking algorithm, and a summary of each customers' visit was plotted in the bar graph in Figure 4 below. From the graph, it is observed that customers spent more time in Area 2.
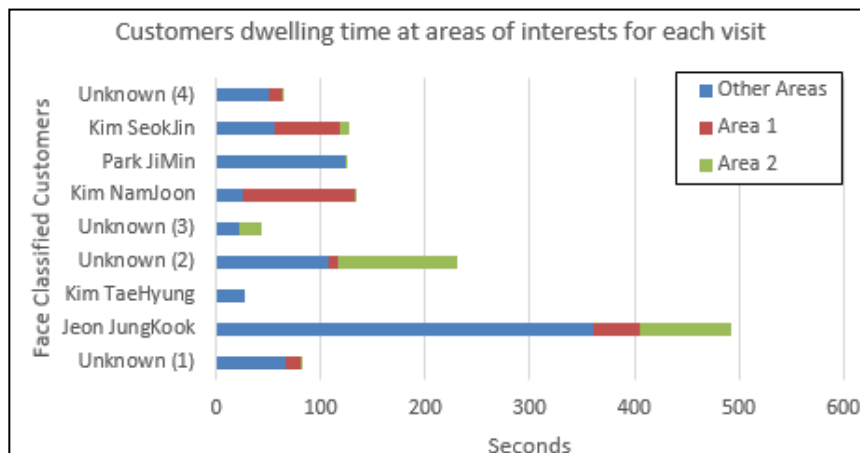

Figure 4. Summary of each customer visit of Trial (1)

With this information, retailers can strategize their product placements inside the store. For example, retailers may want to place magazines where it will require the customer to go through new released products to piqued their customers' interest in buying one before going to their targeted products. This analysis can also help retailers identify their loyal customer wants and needs. A customer may dwell in the aisle to check the prices and consider whether to buy it or not and retailers can give vouchers to their loyal customers to ensure their customers return. Thus, this system can help retailers to optimize their management and increase overall revenue.

CONCLUSION

A consumer behaviour management system was developed using the integration of face classification via FaceNet and customer tracking using a single shot detection algorithm. The system has 56.67% face classification accuracy and manages to track customer's whereabouts and dwelling time inside the store per visit. With the result of the integrated information from face classification and person tracking algorithms, a consumer behaviour analysis can be made by making a timeline graph of each identified customer. From the graph, it is found that the customers spent more time in Area (2) which is the Magazines aisle. With that, the system enables physical retailers to collect similar information as online retailers and the data may help them improve the consumer experience, optimize overall management and increase revenue.

REFERENCES

[1]   B. D. Frew, "There Is Hope for Brick and Mortar Retail," no. october, pp. 105–106, 2017.
[2]   D. K. Gauri et al., "Evolution of retail formats: Past, present, and future," J. Retail., 2020, doi: 10.1016/j.jretai.2020.11.002.
[3]   W. Liu et al., "SSD: Single shot multibox detector," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9905 LNCS, pp. 21–37, 2016, doi: 10.1007/978-3-319-46448-0_2.
[4]   A. Faza and S. Darma, "Implementation of single shot detector for object finding in drone platform," J. Phys. Conf. Ser., vol. 1528, no. 1, 2020, doi: 10.1088/1742-6596/1528/1/012005.
[5]   I. Huh, C. Science, and S. Diego, "Blood Cell detection using Singleshot Multibox Detector."
[6]   F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 815–823, 2015, doi: 10.1109/CVPR.2015.7298682.
[7]   A. Mishra and R. Vishvas, "Retail shopper empowerment : A consumer- centric measure for store performance," IIMB Manag. Rev., vol. 31, no. 1, pp. 20–36, 2019, doi: 10.1016/j.iimb.2018.08.006.
[8]   "5 Key Challenges Facing Retailers Today – And How to Solve Them," 2020. https://etaileast.wbresearch.com/blog/five-key-challenges-for-retailers-how-to-solve-them.
[9]   Amazon.com, "Just Walk Out," Amazon.com. 2020, [Online]. Available: https://justwalkout.com/%0Ahttps://www.justwalkout.com/.
[10]   "Amazon Go." https://www.amazon.com/b?ie=UTF8&node=16008589011 (accessed Jun. 28, 2021).
[11]   iProxCCTV, "iProx CCTV for retail - High definition iProx IP CCTV HD Cameras for retail industry - YouTube," 2014. https://youtu.be/B9xQM8GEiYk.
[12]   hernanrazo, "hernanrazo_split-videos-to-frames_ Python script that splits videos into individual frames," 2019. https://github.com/hernanrazo/split-videos-to-frames/commits/master (accessed Aug. 17, 2020).