Product Recommendation using Deep Learning in Computer Vision

Sharvinteraan A/L C. Mogan
Faculty of Computing
Universiti Malaysia Pahang Al-Sultan Abdullah
26600 Pekan, Pahang, Malaysia
sharvinmogan05@gmail.com

Zuriani Mustaffa*
Faculty of Computing
Universiti Malaysia Pahang Al-Sultan Abdullah
26600 Pekan, Pahang, Malaysia
zuriani@ump.edu.my

Mohd Herwan Sulaiman
Faculty of Electrical and Electronics Engineering technology
Universiti Malaysia Pahang Al-Sultan Abdullah
26600 Pekan, Pahang, Malaysia
herwan@ump.edu.my

Ferda Ernawan
Faculty of Computing
Universiti Malaysia Pahang Al-Sultan Abdullah
26600 Pekan, Pahang, Malaysia
ferda@ump.edu.my

2023 IEEE 8th International Conference On Software Engineering and Computer Systems (ICSECS) | 979-8-3503-1093-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSECS58457.2023.10256332

***Abstract*— Recently, recommendation models have gained popularity due to their effectiveness in improving customer satisfaction and deriving sales. However, current product recommendation models have a drawback: they lack personalized and targeted advertisements for individual users. Consequently, the recommendations provided are random and not tailored to users' preferences. This limitation negatively impacts the system's ability to deliver relevant and personalized advertisements, leading to reduced user engagement and potentially lower conversion rates. Moreover, the absence of personalized advertisements can result in user dissatisfaction as they may receive recommendations that are irrelevant or not aligned with their interests and needs. To address these challenges, this study proposed a targeted product recommendation model using Deep Learning (DL) techniques in computer vision. The study utilizes the dataset of human images obtained from the Kaggle website, which includes details such as gender, class, and age. Findings of the study demonstrated a high level of accuracy in product recommendations, indicating the potential for significant improvements in addressing the issues. In conclusion, the proposed method achieves good accuracy in predicting the gender and age, and provides appropriate product recommendations based on these features.**

Keywords—Deep Learning, Computer vision, machine learning, Multi-task Cascaded Convolutional Networks, Targeted recommendation.

I. INTRODUCTION

For over two decades, recommendation systems have been in existence and have found applications across various domains. These systems have successfully provided users with recommendations on a wide range of information and items, including movies, books, food, music, news, and more. The advancement of information and item recommendations is closely tied to the research and development of core technology recommendation algorithms within recommendation systems. Both domestic and international researchers have been continuously exploring the field of recommendation systems since the 1990s, demonstrating the ongoing efforts and interest in this area [1].

As the digital economy continue to flourish, personalized recommendations have undergone a remarkable evolution, establishing themselves as an indispensable component of the online shopping journey. By leveraging advanced algorithms and user data, e-commerce providers seamlessly integrate personalized recommendations into their platform, enhancing the overall shopping experience for customers. These tailored suggestions not only help users discover relevant products or services but also foster a sense of trust and loyalty, driving customer satisfaction and ultimately contributing to the success of these online businesses [2][3].

Advertising or product recommendation is a significant marketing strategy employed to introduce or promote events, products, and brands to the public. However, conventional advertisements used in shopping stores often consist of static posters or digital displays that are generic and not tailored to specific customers. In malls, digital advertisement screens typically operate in a slideshow mode, where various products and brands are displayed in a looping sequence. This approach lacks personalization and can lead to inappropriate advertisements being shown to viewers, such as displaying makeup products to a 10-year-old child. Consequently, the effectiveness of these advertisements is limited unless viewers patiently wait for items of interest to be displayed, resulting in inefficiency. To capture the attention of customers and optimize advertising impact, it is crucial to tailor the advertisements according to the viewer's age and gender, displaying products and brands that are relevant and appealing to them.

Moreover, advertising costs pose a significant financial burden for both small and large businesses. According to the Gartner 2022 CMO Spend and Strategy Survey, marketing spend has increased from 6.4% to 9.5% of company revenue across various industries. In the first quarter of 2022, a net balance of 14.1% of companies increased their marketing budgets [4]. In Malaysia, the total investment in advertising expenditure reached approximately 4.36 billion Malaysian ringgit in 2021 and 4.78 billion Malaysian ringgit in 2022 [5]. Additionally, traditional advertising methods such as static posters, electronic posters, and billboards in Malaysia cost hundreds to thousands of ringgits per day, while digital billboards start at a minimum of RM 1,800 per day. As a result, small companies cannot afford to invest their limited resources in marketing and advertising strategies that do not yield significant benefits to their businesses.

To overcome these challenges, it is essential to develop an enhanced product recommendation model that incorporates targeted advertisements based on user preferences, demographic data, and browsing behavior. This approach ensures personalized and engaging recommendations for each user. By tailoring the advertisements to individual users, the model can foster user engagement, increase the likelihood of conversions, and optimize the overall advertising experience for both users and advertisers. In literature, there are numerous approaches that have been presented for targeted recommendation, particularly using machine learning techniques.

A study in [6] proposed a product recommendation based on Deep Learning (DL) and collaborative filtering. In the study, Artificial Intelligence (AI), Natural language processing (NLP) and machine learning have been combined to implement the concept swiftly and effortlessly discovering suitable items online. The resulting PRS allows for easy access at any given time location, effectively reducing time spent in the process. Meanwhile, in [7] proposed a deep ensemble classifier for recommending product, focusing on fashion goods. When evaluated on a fashion product dataset, the model demonstrates compelling results. It achieves an accuracy of 88.32%. The proposed strategy surpasses existing methods in terms of accuracy, precision, recall, F1-score, and kappa statistics, highlighting its superior performance.

The education field is also not left behind in utilizing product recommendation techniques, as presented in [8]. The study proposed the use of DL and big data for personalized recommendation framework based on Massive Open Online Course (MOOC) system. The experimental findings substantiate that the model proposed in the paper yields favorable recommendation outcomes when compared to alternatives method. On the other hand, personalized recommendation system based on learning clustering representation [9]. In the implemented approach, the combination of RNN and attention mechanism is utilized to design the e-commerce product recommendation system. The effectiveness of this method is demonstrated through extensive experiments conducted.

Study in [10], a DL based intelligent method for Personalized Service System (PSS) is proposed, incorporating a modified customer journey map to uncover user requirements. For that matter, a case study focused on taxi operators is presented to illustrate the method's steps. The primary objective of the presented method is the development of a personalized smart PSS method which could achieve a win-win situation for all players in this method.

In this study, the proposed project aims to address the aforementioned challenges by developing a targeted product recommendation model that utilizes Deep Learning in computer vision. By employing publicly available data on Kaggle website, the model is trained to predict the gender, age and provide product recommendation based on the user's age and gender.

This paper is structured as follow: Section I provides the background study and related works, followed by the description on Multi-task Cascaded Convolutional Networks (MTCNN) in Section II. A The implemented methodology is presented in Section III while the results and discussion are presented in Section IV. Finally, Section V concludes the study.

II. Multi-task Cascaded Convolutional Networks (MTCNN)

A framework known as Multi-task Cascaded Convolutional Networks (MTCNN) was developed to address both face alignments and detection tasks. In this framework, convolutional networks are utilized in three stages to detect faces and identify facial landmarks such as the eyes, nose, and mouth. The cascaded network in MTCNN framework consists of three stages. In the first stage, the images are taken and resized to different scales, generating a pyramid of images.

A. Stage 1 – The Proposal Network (P-Net)

The P-Net provides input to the Refine Network, which includes all the candidate windows. The R-Net utilizes non-maximum suppression (NMS) to merge overlapping candidates, performs bounding box regression for calibration, and further reduces the number of candidates. Depending on whether the input is a face or not, the R-Net produces a four-element vector representing the bounding box coordinates for the face, as well as ten-element vector for localizing facial landmarks.

B. Stage 2 – The Refine Network (R-Net)

The P-Net supplies all candidate data to the Refine Network. This network is a CNN rather than an FCN, as it includes a dense layer in its final stage of architecture. The R-Net incorporates non-maximum suppression (NMS) to merge overlapping candidates, performs bounding box regression for calibration, and further reduces the number of candidates. Depending on whether the input is a face or not, the R-Net generates a four-element vector that represents the bounding box for the face, as well as a ten-element vector for localizing facial landmarks.

C. Stage 3 – The Output Network (O-Net)

Like the R-Net, this stage aims to analyze the face in more detail and produce the coordinates of the five facial landmarks, such as the eyes, nose, and mouth. MTCNN performs three tasks, including face/non-face classification, bounding box regression, and facial landmark localization.

- Face classification: This task involves solving a cross-entropy loss problem that focuses on binary classification.
- Bounding box regression: The learning objective in this task is to address a regression problem. It involves calculating the offset between a candidate window and the closest ground truth. Euclidean loss is employed as the loss function for this task.
- Facial landmark localization: The task of localizing facial landmarks, such as the right eye,

left eye, nose, right mouth corner, is treated as a regression problem. The Euclidean distance is used as the loss function for this regression task.

III. Methodology

This section is dedicated to discuss the implemented methodology which includes data collection, data pre-processing and data organization. This is followed by model development, training, and testing. The steps end with evaluation of the model prior to real world testing. Fig. 1. shows the model development of targeted recommendation based on MTCNN.

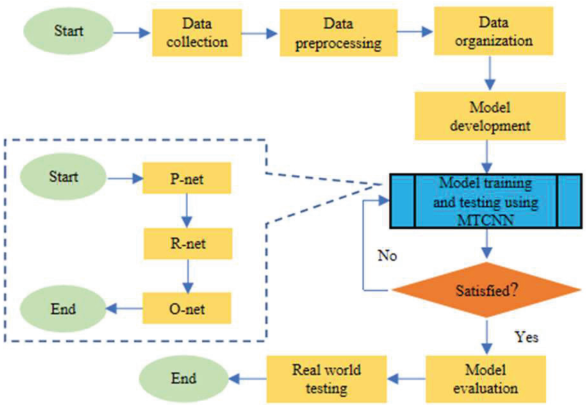


Fig. 1: Model Development of Targeted Recommendation based on MTCNN

A. Data Collection

The dataset of human face images are collected from the Kaggle website, which is a platform that offers diverse datasets for training machine learning models. The dataset will encompass human faces representing various creeds, races, age groups, and profiles. It comprises a total of 7.2K images and is accompanied by an Excel sheet containing additional information about the images, including their ID, gender, class, and age.

TABLE I. SAMPLE OF DATASET

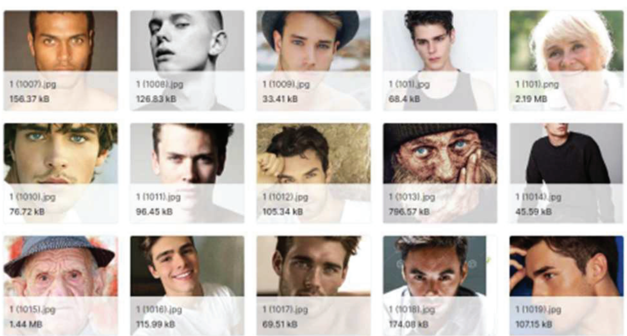


Fig. 2: Sample of dataset

TABLE I. SAMPLE OF IMAGE DETAILS

ID	Gender	Class	Age
377.jpg	Male	MIDDLE	25
17814.jpg	Male	YOUNG	12
21283.jpg	Male	MIDDLE	23
16496.jpg	Female	YOUNG	17
4487.jpg	Male	MIDDLE	26
6283.jpg	Female	MIDDLE	24
23495.jpg	Female	YOUNG	8
7100.jpg	Male	YOUNG	8
6028.jpg	Male	YOUING	12

B. Data Preprocessing

The collected images will undergo two different preprocessing steps. Firstly, the images will be resized to a standard size of 480 x 640 pixels, making it more convenient for loading during model training. Secondly, the images will be converted from colored to grayscale. This conversion reduces the image size and computational complexity during image processing, as color information does not play a significant role in image recognition.

C. Data Organization

A total of 400 pre-processed human images will be randomly selected, representing a mix of young, middle-aged, and old-age faces. These images will then be organized and placed in a separate folder dedicated to image training. Additionally, the accompanying Excel sheet data will be sorted and filtered based on the chosen images to ensure alignment and relevance during the training process.

D. Model Development

The model development will utilize Visual code as the integrated development (IDE), with Python being the programming language of choice for coding the model. The model will utilize locally stored data and will incorporate the laptop's webcam for real-time testing purposes.

E. Model Training

The model training process involves splitting the dataset into 80% for training and the remaining 20% for testing. The developed model will then be executed using a dataset containing human images, specifically prepared for gender and age classification. The results of the classification readings will be recorded and stored in a new Excel file for each training iteration. The model will be fine-tuned until the prediction accuracy reaches 0.8. Additionally, the accuracy of the product recommendation results will be recorded and verified. The training process will be repeated until the desired accuracy target is achieved.

F. Model Testing

After completing the model training, it will undergo testing using the laptop's webcam by integrating it with the OpenCV library, which offers image reading capabilities for further processing. The model will be applied to real human faces captured by the camera, and the resulting predictions for gender, age, and product recommendations will be recorded in an Excel file for evaluation purposes.

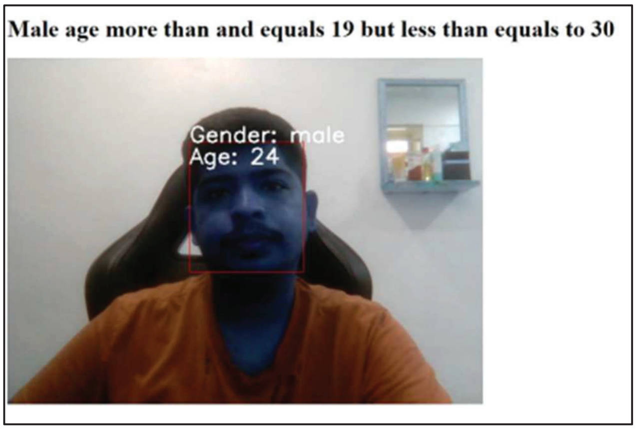


Fig. 3: Model testing using webcam.

G. Model Evaluation

For evaluation purposes, the Mean Absolute Error (MAE) is calculated by finding the absolute difference between each predicted age value and its corresponding actual age value. These differences are averaged across the dataset and expressed as a percentage. The formula for MAE is as follows:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{1}$$

where y_i is the prediction value, x_i denotes the actual value and n is the total number of data points.

H. Real World Testing

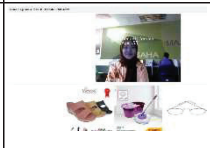
Following proper evaluation and testing, the model will be deployed in a real-world scenario, specifically targeting an area where it can analyze people's faces. This deployment aims to assess the accuracy and effectiveness of the model in predicting, recommending, and attracting customers. The data collected during this phase will be recorded to calculate the model's effectiveness and evaluate its marketability.

IV. Results and Discussion

Table 2 shows the testing result of targeted product recommendation. The inputs are age and gender while the outputs are the recommended products. Based on the table, the predicted age and gender closely matched the actual age and gender provided by the users, as indicated by a low MAE, indicating minimal deviation between predicted and actual ages. Additionally, the model achieved 100% accuracy in gender prediction. The product recommendation based on the predicted age and gender were consistent across all users, demonstrating the reliability of the model. The product recommendation accuracy was measured at 87.5%, accurately recommending products to 7 out of 8 users tested.

These findings highlight the robustness and precision of our product recommendation model, allowing it to deliver personalized recommendations tailored to users' characteristics. As a result, user satisfaction and engagement are expected to be enhanced.

TABLE II. Testing Results

U	A	G	PA	PG	R_products	Output
A	23	M	24	M, YA	Sling bag, hoodie, sneaker	
B	23	M	24	M, YA	Sling bag, hoodie, sneaker	
C	23	M	24	M, YA	Sling bag, hoodie, sneaker	
D	23	M	24	M, YA	Sling bag, hoodie, sneaker	
E	25	F	33	F, OA	Sandal, Mr. DIY mop, Old fashioned spectacle	
F	20	F	21	F, YA	Dress, Handbag, pink shoes	
G	24	F	27	F, YA	Dress, Handbag, pink shoes	
H	36	M	36	M, OA	Mr. DIY hardware storage, Mr. DIY cordless screwdriver	

**U=User, A=Actual age, G=Actual gender, PA=Predicted age, PG=Predicted gender, R_products=Recommended products, M=Male, F=Female, YA=Young adult, OA=Old adult*

TABLE III. Age Prediction Results

User	Actual Age	Predicted Age	Absolute Difference
A	23	24	1
B	23	24	1
C	23	24	1
D	23	24	1
E	25	33	8
F	20	21	1
G	24	27	3
H	36	36	0

V. Conclusion

This paper presents a personalized recommendation method that integrates DL in computer vision. By using the freely available dataset in Kaggle, the DL is employed to train the data. Evaluated based on MAE, the experimental results have demonstrated that the proposed recommendation method is efficient. The proposed method was able to provide good accuracy in predicting the gender and age, and recommend appropriate products based on the two features. For future works, the large-scale experiment analysis will be considered to verify the capability of the presented model.

ACKNOWLEDGMENT

This research was supported by UMPSA Grant, #RDU220379.

REFERENCES

[1] Z. Wang, A. Maalla, and M. Liang, "Research on E-Commerce Personalized Recommendation System based on Big Data Technology," in *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, 2021, vol. 2, pp. 909–913, doi: 10.1109/ICIBA52610.2021.9687955.

[2] Z. Zhan and B. Xu, "Analyzing review sentiments and product images by parallel deep nets for personalized recommendation," *Inf. Process. Manag.*, vol. 60, no. 1, p. 103166, 2023, doi: https://doi.org/10.1016/j.ipm.2022.103166.

[3] S. Basu, "Personalized product recommendations and firm performance," *Electron. Commer. Res. Appl.*, vol. 48, p. 101074, 2021, doi: https://doi.org/10.1016/j.elerap.2021.101074.

[4] J. Turner, "What Marketing Budegts Look Like in 2022," *Gartner*, 2022. https://www.gartner.com/en/articles/what- marketing-budgets-look-like-in-2022.

[5] "Total investiment in advertising expenditure in Malaysia from 2012 to 2020 with a forecast for 2021 and 2022," *Statistica Research Department*, 2022. statistica.com.

[6] M. D. Bhagat and P. N. Chatur, "A Study on Product Recommendation System based on Deep Learning and Collaborative Filtering," in *2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, 2023, pp. 1–5, doi: 10.1109/ICAECT57570.2023.10117628.

[7] B. Suvarna and S. Balakrishna, "A Novel deep ensemble classifier for recommending fashion products," in *2022 3rd International Conference on Communication, Computing and Industry 4.0 (C2I4)*, 2022, pp. 1–6, doi: 10.1109/C2I456876.2022.10051256.

[8] B. Li *et al.*, "A personalized recommendation framework based on MOOC system integrating deep learning and big data," *Comput. Electr. Eng.*, vol. 106, p. 108571, 2023, doi: https://doi.org/10.1016/j.compeleceng.2022.10857 1.

[9] K. Wang, T. Zhang, T. Xue, Y. Lu, and S.-G. Na, "E-commerce personalized recommendation analysis by deeply-learned clustering," *J. Vis. Commun. Image Represent.*, vol. 71, p. 102735, 2020, doi: https://doi.org/10.1016/j.jvcir.2019.102735.

[10] M.-C. Chiu, J.-H. Huang, S. Gupta, and G. Akman, "Developing a personalized recommendation system in a smart product service system based on unsupervised learning model," *Comput. Ind.*, vol. 128, p. 103421, 2021, doi: https://doi.org/10.1016/j.compind.2021.103421.