https://doi.org/10.46335/IJIES.2023.8.3.1 **e-ISSN: 2456-3463**
Vol. 8, No. 3, 2023, PP. 1-6

International Journal of Innovations in Engineering and Science, www.ijies.net

Retail Store Analytics Using Facial Recognition

Nishant Sawant [1] , Akansha Rai [2], Saiprasad Parab [3] , Bansi Ghanva [4]

[1,2,3,4] UG-B.E Student, Information Technology, Shree L.R Tiwari College of Engineering, India

sawantnishant01@gmail.com

***Received on**: 29March, 2023* ***Revised on :** 09 April,2023* ***Published on**: 11 April, 2023*

Abstract—In this paper, we present a system for real-time age, gender, and emotion detection using webcam and machine learning techniques. The system is designed to capture real-time video footage of customers in a retail store and extract demographic and emotional information to perform retail analytics. We used the UTKFace dataset to train our age and gender model and the FER dataset to train our emotion model. The trained models were integrated with OpenCV and TensorFlow to detect faces, predict age, gender, and emotion in real-time. The system stores the collected data in a MySQL database, which is then used to perform various analyses to gain insights into customer behavior. We provided analysis using Flask and built a web interface for getting the insights on our data. The results show that our system is effective in capturing and analyzing customer information in real-time and can provide valuable insights for retailers to make informed decisions. Our system can be extended to include other features such as customer segmentation, heatmaps, and customer journey analysis. Overall, our system provides a powerful tool for retailers to understand customer behavior and improve the shopping experience.

***Keywords—**UTKFace, FER dataset, age, gender, emotion, Flask.*

I. INTRODUCTION

In a democratic country such as India the customer demographics such as age and gender along with the sentiments of a customer experience towards a few products plays a vital role in the sales and business of retailers and small scale vendors. The analyzing part of the customers is an important aspect for the retailers to stay on the top in the business marketing strategies and maintaining a good number of count of costumers to maximize the profits. Now as being the retailer, understanding and working upon the needs and the interests of the existing customers is proven to be a very effective path in order to tackle the battle of knowing what all products needs to be produces and stock up together for the future. This type of analysis would help retailers to ensure the customers finding and what all they are looking for in the market and through this we can make sure a good count of growing customer acquisition in the business approach[1]. Nowadays various approaches are available for the detection of face recognition it can differ at some points in the average precision and in framing the processing speed. Emotional recognition using both audio and video can be used for sales using Mobile phones[3]. Here the main objectives of our work is to experiment and then do an evaluation on how effective the working of CNN (in respect to the processing speed and the accuracy) and also to solve the problems related to any face detection and recognition[4]. We referred the given paper to create UML and get a baseline of our project and get a better understanding of the system to develop[10]. Here comes the presentation of our detection and recognition system which is going to predict about the satisfaction of the customer on the basis of real-time evaluation . This system is built in such a way that it holds the capability to recognize the emotions that are for example being sad, happy, neutral,

calm,surprised, happy, fearful, and disgusted. Only on the basis of relying on the confident predictions only the system has not made any incorrect classification within the premises and also provides a continues prediction and stores selected prediction that happens atleast in every 10 seconds (that is near the real-time). We also referenced few paper that studied the performance of pre-built model for face detection[2]. There are multiple methods that are used in the recognition of the emotion, age and gender of any person which includes the emotion, age and gender detection using the concept of the deep convolution neural networks. This is the solution that has been a trained neural network which helps in recognizing the sentiments that are presented on a human face that has been recognized from the video that is available in the input stream which is present in real time. At the same time it is necessary to implement security to web applications at the beginning stage itself so as to avoid misuse of data by hackers[7]. The numerous requirements of the chunks that are in the process of recognizing the age, gender and sentiments includes the following: The network training of the requires the latest, and huge and different dataset types .The training here requires a very huge power in while computing as we are having a large size of the dataset and quite a high number of parameters which are used for the CNN models.

II. Problem Statement

Traditionally, majority of enterprises used conventional methods of marketing such as advertisements, customer satisfaction surveys etc. which are time consuming In the consumer base, we find that each individual customer exhibit inclination towards a said commodity so categorizing consumers is necessary to cope up with this problem. The studies suggest that around 20 percent of new businesses survive their first year of operation and half of the small businesses close down in the first 5 years. Analyzing the target audience and finding a way to attract customers of all variations is also a very difficult task.

III. Literature Review

[1] This paper features a work for customer analytics using computer vision with Wide Residual Networks and Xception deep learning models to predict age and gender demographics of the consumers, achieving a 82.9 percent accuracy rate for gender and 70.8 percent accuracy for age-range

[2] It compares the effectiveness of different CNN for face detection and recognition. It shows the comparative study for both face detection and face recognition.

[3] The paper proposed an emotion recognition system for customer satisfaction using both audio and video for recognition using Deep Neural Network of VGG16.

[4] In this paper, the recognition of emotion, gender and age has been attempted in real time video stream. The proposed method of using the general architecture of CNN along with the squeezenet and Xception architecture combined in a hierarchical approach has been shown.

[5] It present a system model to predict the satisfaction of a customer through his emotions. First geometric features was extracted via camera and classification is done via SVM classifier.

[6] This proposes a cost efficient method of deep learning to find popular/hot spots in a retail store with RGB camera and CNN.

IV. Ease Of Use

User-friendly Interface and Scalability: Making Data Analytics Accessible for Retailers The system is designed to be user-friendly and accessible for retailers of all sizes. The system's interface is intuitive and easy to navigate, providing data analytics in a visually appealing format.

The system is easy to run and configure, with minimal hardware and software requirements. It can be integrated with existing security cameras and point-of-sale systems, making it easy for retailers to leverage their existing infrastructure.

The system is also scalable, allowing retailers to easily add additional cameras and locations to the system as their business grows. The system can be customized to meet the specific needs of each retailer, with the ability to define custom categories and metrics for analysis.

The system provides valuable insights into customer behavior, allowing retailers to make informed decisions about how to improve the shopping experience and increase customer satisfaction. The system's ease of use makes it accessible to retailers who may not have the resources or expertise to develop their own data analytics tools.

Overall, the ease of use of the system makes it a valuable tool for retailers who want to better understand their customers and improve their business.

https://doi.org/10.46335/IJIES.2023.x.xx.xxx e-ISSN: 2456-3463
Vol. x , No. xx, 20xx, PP. xx-xx

V. METHODOLOGY AND SYSTEM ARCHITECTURE

To detect and analyze customer demographics, the system uses custom trained models each for age, gender and emotion that predicts the same from facial features. The age and gender model is trained on the UTKFace dataset and integrated with OpenCV and TensorFlow to detect faces and predict age and gender in real-time. Foremost for any software development creating SRS is an important and required step of process and along with security of each step[9]. System containing data of users must have secure structure and needs proper planning[8]. To analyze customer emotions, the system uses a custom trained emotion model that predicts facial features. Similarly, the model is trained on the FER dataset and integrated with OpenCV and TensorFlow to detect facial expressions in real-time.

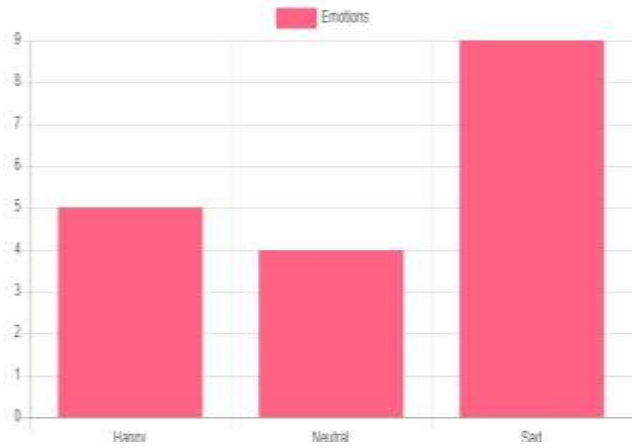

Fig 1. Sample analysis by system for a month

A. UTK Face Dataset

The UTKFace dataset is a large-scale face dataset that was created by analyzing the face images from different IMDb and Wikipedia. The dataset contains over 20,000 face images with annotations for age, gender, and ethnicity. The face images were collected from different sources, including movies, TV shows, and news articles.
The dataset includes faces of different ages, ranging from infants to elderly people, as well as different ethnicities, including Caucasian, African, Asian, and Indian. The dataset also includes gender annotations, with each face image labeled as male or female. The UTKFace dataset is widely used in the computer vision and machine learning communities for age and gender prediction tasks. The dataset provides a large and diverse set of images that can be used to train machine learning models, improving the accuracy and robustness of the models. For the sake of our project requirements and processing we have considered only Age and Gender label of images.

B. FER Dataset

The FER dataset is a facial expression recognition dataset that contains approximate 35,000 images of facial expressions. The dataset includes seven different facial expressions: anger, disgust, fear, happiness, sadness, surprise, and neutral. The face images were collected from different sources, including Google Images and Flickr. The FER dataset is widely used in the computer vision and machine learning communities for facial expression recognition tasks. The dataset provides a large and diverse set of images that can be used to train machine learning models, improving the accuracy and robustness of the models. Facial expression recognition is an important research area in computer vision and machine learning, with applications in human-computer interaction, emotion detection, and psychology. The FER dataset is a valuable resource for researchers and practitioners in this field, providing a benchmark dataset for evaluating facial expression recognition models.

*C. **Tools and Libraries***

We implemented the real-time age, gender, and emotion detection system using Python programming language, with the following tools and libraries:
OpenCV: a popular computer vision library for real-time image and video processing TensorFlow: an open-source machine learning library for building and training neural networks.
NumPy: a library for numerical computing in Python Pandas: a library for data manipulation and analysis in Python.
Matplotlib: a library for data visualization in Python. MySQL: a relational database management system for storing and retrieving customer information. Flask: Flask is a micro web framework used in python to provide tools, libraries for building web application.

D. Face detection

For face detection, we used the Haar-cascade frontal face algorithm, which is a widely used object detection algorithm for detecting faces in images and videos. Haar cascade make use of cascading windows, and it tries to compute features in every feasible window and classify whether that window contains an object.

https://doi.org/10.46335/IJIES.2023.x.xx.xxx e-ISSN: 2456-3463
Vol. x , No. xx, 20xx, PP. xx-xx

Fig. 1- fig shows the haar cascade window traverse

E. Age and gender

For age and gender prediction, we used a convolutional neural network (CNN) that was trained on the UTKFace dataset. The CNN architecture consisted of several convolutional layers followed by max-pooling layers, dropout layers, and fully connected layers. The model was trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. We used confusion matrix to evaluate the performance of the model and get the accuracy of models.

F. Emotion prediction

For emotion prediction, we used a CNN that was trained on the FER dataset. The CNN architecture consisted of several convolutional layers followed by max-pooling layers, dropout layers, and fully connected layers. The model was trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. We used K-fold cross-validation to evaluate the performance of the model, achieving an accuracy of 60 percent for emotion prediction.

G. Experiments and evaluation

To evaluate the performance of the system, we conducted several experiments using different video footage from a retail store. We analyzed the accuracy of the age, gender, and emotion predictions, as well as the processing speed of the system. We also evaluated the effectiveness of the system in capturing customer information and providing insights for retailers. Overall, the system was found to be effective in providing real-time analysis of customer demographics and emotions, with the potential to improve customer experience and drive sales for retailers.

VI. Accuracy And Evaluation

$$Accuracy = \frac{TP + TN}{TP+TN+FP+FN}$$

...eqn 1

Where, TP = True positive

TN = True Negative

FP = False Positve

FN = False Negative

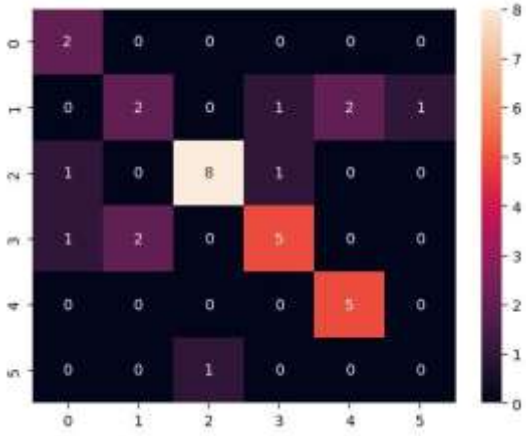

Fig. 2- Matrix for emotion model

Fig. 3- Confusion matrix for Age model

We created a confusion matrix based on the result of test dataset and got 86% accuracy for gender and 68.75% for emotion.

VII. Ethical Considerations

As the system collects personal information such as age, gender, and emotions, it is important to discuss ethical considerations such as privacy and data protection. It is important to ensure that the collected data is used ethically and transparently. One way to do this is to inform customers that their data is being collected. Our system doesn't store any crucial data specific to the customer that may identify them, instead only data of age, gender and emotion is stored anonymously. Additionally, it is important to ensure that

https://doi.org/10.46335/IJIES.2023.x.xx.xxx e-ISSN: 2456-3463
Vol. x , No. xx, 20xx, PP. xx-xx

the data collected is stored securely and only accessed by authorized personnel as web application are vulnerable to many cyber crimes[7].

A. Improved customer experience

The proposed system can help retailers gain a better understanding of customer demographics and behavior, allowing them to personalize the shopping experience and provide better customer service. For example, if the system detects that a customer is feeling unhappy or frustrated, store employees can intervene and address their concerns in real-time, improving their overall experience and increasing the likelihood of repeat visits.

B. Increased sales

By analyzing customer behavior and preferences, the proposed system can help retailers optimize their inventory and product placement, leading to increased sales. For example, if the system detects that a particular product is popular among a certain demographic, retailers can stock more of that product and place it in a prominent location to increase visibility and sales.

C. Optimized operations

he proposed system can also help retailers optimize their operations by providing insights into staffing levels, store layout, and promotional campaigns. For example, by analyzing customer traffic patterns, the system can help retailers determine the most efficient staffing levels for different times of day, reducing labor costs while maintaining optimal service levels.

VIII. COMPARISON TO OTHER METHODS

Some of the methods that can be compared to the proposed system include traditional survey-based methods, manual observation, and other automated systems such as RFID-based tracking systems and Wi-Fi tracking systems.

Survey-based methods rely on collecting data through questionnaires or interviews with customers. While these methods can provide valuable insights into customer opinions and preferences, they are often limited by response bias, small sample sizes, and the time and effort required to conduct surveys.

Manual observation involves trained observers monitoring customer behavior and recording data manually. While this method can provide detailed insights into customer behavior, it is limited by the subjectivity of the observers and the time and cost required to conduct the observations.

RFID-based and Wi-Fi tracking systems use sensors to detect customer movement and track their behavior. While these methods can provide accurate data, they are often limited by the need for specialized hardware, privacy concerns, and the cost of implementation.

Compared to these methods, the proposed system offers several advantages. It is automated and can provide quick data analysis, reducing the time and effort required to collect and analyze data. It is also non-invasive and does not require specialized hardware, making it easy to implement and less intrusive to customers. Additionally, the system can provide insights into customer demographics, emotions, and behavior, providing a more complete view of customer experience.

IX. LIMITATIONS

One of the main limitations of the system is that the accuracy of the age and gender detection may be affected by factors such as lighting conditions and camera angle. In order to improve the accuracy of the age and gender detection, it may be necessary to use multiple cameras with different angles and lighting conditions. Additionally, the emotion detection may not always accurately reflect the true emotional state of the customer. For example, a customer may appear to be angry or frustrated, but in reality, they may just be having a bad day. Therefore, it is important to ensure that the data collected is interpreted carefully and not solely relied upon to make decisions.

X. RESULT AND CONCLUSION

Sentiment Analysis, Face and Gender Detection, individually have numerous use-cases in today's world. Emotion, Age and Gender detection will be an important topic in the field of computer vision and artificial intelligence due to its crucial educational and commercial potential. While we obtained a satisfactory result for gender and emotion, the age parameter is difficult for the machine to predict and gives a prediction with an error of ±10.

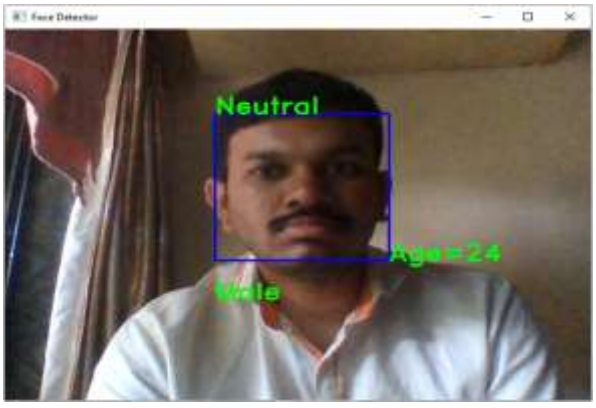

Fig. 4-. Sample result of system detection

https://doi.org/10.46335/IJIES.2023.x.xx.xxx e-ISSN: 2456-3463
Vol. x , No. xx, 20xx, PP. xx-xx

Fig. 5- Sample result of system detection

ACKNOWLEDGMENT

We would like to express our sincere gratitude to our project guide Prof. Madhuri Gedam for her guidance, feedback, and support throughout the project. Her expertise and insights were invaluable in shaping the direction and scope of our research.

We would also like to thank our project coordinator Prof. Vikas Kaul who provided support and encouragement during the project. His feedback and input helped us refine our ideas and improve the quality of our work.

We are grateful to Shree L R Tiwari College of Engineering for providing access to the necessary resources and facilities, including the library, computer labs, and equipment, that were essential to completing this project.

REFERENCES

[1] *E. P. Ijjina, G. Kanahasabai and A. S. Joshi, "Deep Learning based approach to detect Customer Age, Gender and Expression in Surveillance Video," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-6.*

[2] *K. Smelyakov, A. Chupryna, O. Bohomolov and N. Hunko, "The Neural Network Models Effectiveness for Face Detection and Face Recognition," 2021 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2021, pp. 1-7.*

[3] *S. -A. Naas and S. Sigg, "Real-time Emotion Recognition for Sales," 2020 16th International Conference on Mobility, Sensing and Networking (MSN), Tokyo, Japan, 2020, pp. 584-591.*

[4] *P. Aiswarya, Manish and P. Mangalraj, "Emotion recognition by inclusion of age and gender parameters with a novel hierarchical approach using deep learning," 2020 Advanced Communication Technologies and Signal Processing (ACTS), Silchar, India, 2020, pp. 1-6.*

[5] *M. S. Bouzakraoui, A. Sadiq and A. Y. Alaoui, "Appreciation of Customer Satisfaction Through Analysis Facial Expressions and Emotions Recognition," 2019 4th World Conference on Complex Systems (WCCS), Ouarzazate, Morocco, 2019, pp. 1-5.*

[6] *M. S. Bouzakraoui, A. Sadiq and A. Y. Alaoui, "Appreciation of Customer Satisfaction Through Analysis Facial Expressions and Emotions Recognition," 2019 4th World Conference on Complex Systems (WCCS), Ouarzazate, Morocco, 2019, pp. 1-5.*

[7] *Gedam, Madhuri N., and Bandu B. Meshram. "Proposed Secure Content Modeling of Web Software Model." IJETT 6.2 (2019).*

[8] *Gedam, Madhuri N., and B. B. Meshram. "DATABASE PRIVATE SECURITY JURISPRUDENCE: A CASE STUDY USING ORACLE."*

[9] *Gedam, M., and B. Meshram. "Vulnerabilities and attacks in SRS for object-oriented software development." WCES (2019), San Francisco, USA (2019): 22-24.*

[10] *Gedam, Madhuri, and Bandu B. Meshram. "Proposed Secure 3-Use Case Diagram." International Journal of Systems and Software Security and Protection (IJSSSP) 13.1 (2022): 1-18.*