The Innovative Role of Process Mining in building Face Re-identification Trajectory

Amirah Alharbi
Computer Science and Artificial Intelligence Department, Umm Al Qura University, Saudi Arabia
amnharbi@uqu.edu.sa

Received: 24 November 2023 | Revised: 10 December and 14 December 2023 | Accepted: 16 December 2023

ABSTRACT

Face recognition and tracking technology have witnessed significant advancements during the recent decades. These advancements include improved accuracy and speed in identifying and tracking individuals, as well as the ability to recognize faces in various scale and lighting conditions. Leveraging the potential of face recognition and tracking, this article explores the integration of process mining techniques to discover and visualize face's appearing trajectories in crowd scenarios, aiming to enhance crowd security, surveillance, and personal identification. Notably, existing face recognition tools typically focus on bounding box localization, neglecting the utilization of face coordinates to construct trajectory models upon face re-identification. In this paper, full system architecture for building a trajectory model of re-identified faces in a crowd is proposed. This approach significantly helped in building a large database of visitor faces, and the proposed trajectory model resulted in a high rate of true positive face re-identification.

Keywords-process mining; face recognition; face detection; crowd management; surveillance; smart city; tracking

I. INTRODUCTION

Face recognition and tracking technology [1] has seen significant advancements in recent years. These advancements include improved accuracy and speed in identifying and tracking individuals, as well as the ability to recognize faces in various scale and lighting conditions. Detecting faces in crowd is growing rapidly [2]. Authors in [3] highlighted the state of arts techniques that are developed for identifying and verifying faces on large scale events. Face recognition and tracking technology have revolutionized the way we pimpoint and monitor individuals. By harnessing its power, we can now unlock new possibilities in crowd security, surveillance, and personal identification. In this article, we contribute to face recognition and tracking research, by exploring the potential use of process mining for discovering and visualizing trajectory of faces in crowd. Through the latter processes and in conjunction with face recognition techniques it may be possible for crowd behavior to be controlled. To the best of our knowledge, face recognition tools mostly rely on labeling faces by localizing box boundaries of identified faces without further utilization of face coordinates on constructing a trajectory model when a face is re-identified. This work demonstrates the feasibility of process mining principles to model face presence in crowd.

Previous research that tried to address the issue of tracking re-identified faces mostly used active bounding boxes and associated faces with random numbers for single events [4]. Some researchers have used geo-fences information to track faces that appeared in different locations in a specific event [5]. Our proposed approach aims to provide a retrospective graph-based model which shows face appearance detected in several events. The model produced is a lightweight and scalable model derived from the construction process using a textual event log.

The real-world applications of face detection and recognition are needed in an unconstrained environment which generates low resolution image quality. Several research works have addressed this issue. The work presented in [6] was one of the first research attempts that highlighted the challenges of detecting faces in such an environment. Therefore, the large database Labeled Faces in the Wild (LFW), was built for model training purposes. Generally, face recognition technology utilizes algorithms to analyze and single out unique facial features of individuals. It works by capturing an image or video of a person's face and comparing it to a database of known faces. The technology can recognize specific facial characteristics namely the distance between the eyes, the nose shape, and the face contour. This information is then used to create a unique facial template for everyone, allowing for accurate identification and tracking [1].

A. *General Approach of Face Recognition*

A modern face recognition pipeline as suggested by [7, 8] consists of 4 common stages: Detection, Alignment, Representation, and Verification

- The Detection stage aims to localize the key landmarks of a face such as eyes, nose and mouth usually using bound boxing of the detected face area.

- Alignment: The detected faces are horizontally aligned, getting prepared for the next stage. Other image pre-processing techniques such as normalization can be included in the alignment stage as well.
- Representation stage generates a vector representation or feature embeddings for each face. The Representation length depends on the technique used.
- Verification stage compares a pair of faces to test if they match.

B. Crowd Management

The term crowd in the greatest part of literature is referred to a group of people without strict quantification. Crowd density aims to estimate the number of people in square meters as mentioned in [9]. There are different examples of crowds such as political demonstrations, festivals, sports Olympics or religious events like Hajj in Makkah. These kinds of events require professional controlling and management since they include a huge number of people gathered in a specific place. Crowd management can be defined as the process of detecting, monitoring, tracking, and using a robust decision support system to anticipate and prevent possible crowd related risks [3]. Most research addressing crowd-related concerns has predominantly concentrated on one of these areas: crowd modeling, crowd monitoring, and crowd management. Crowd modeling involves the application of various simulation tools, as extensively detailed in this survey [10]. In contrast, crowd monitoring primarily relies on computer vision methods, as outlined in [3, 11]. Crowd modeling and crowd monitoring research share a common overarching objective, which is to enhance crowd management.

C. Face Detection and Recognition in Crowd

Face recognition technology has numerous applications in crowded spaces. Face detection is used for counting the number of people in crowded places [12]. They designed an application to easily estimate the head count in a photograph using a Convolutional Neural Network (CNN). The application was tested for preventing crowd in a predefined area during the COVID-19 pandemic. The author in [13] has asserted the importance of face recognition technology in alleviating crowd density overestimation. Authors in [14] highlighted the issue of people entering or leaving the scene or becoming obscured in dynamic environments, which can cause fluctuations in head counts. To address these challenges in video crowd counting, a solution called Locality-constrained Spatial Transformer Network (LSTN) was used, which is a face recognition application for finding missing persons in crowd [5]. A significant application of face recognition is demonstrated in [13] for security purposes where the technique of Local Binary Patterns Histogram (LBPH) is applied with images that were captured by drones. The suggested approach can be followed to enhance surveillance systems by identifying individuals of interest, and thus helps prevent crimes, identify suspects, and enhance public safety. Facial Expression Recognition (FER) can be utilized in analyzing crowd mood to classify people into different groups [16]. Face recognition can be also utilized in access control systems [17], allowing for secure and convenient entry into buildings or gathering events. On the other hand, face tracking and face recognition are two complementary monitoring systems. Face tracking focuses on real-time tracking of individuals within a given area, while face recognition focuses on identifying and matching faces on a database [7]. Face tracking can be implemented to monitor crowd movements, identify suspicious behavior, or track individuals of interest [18] without the need for tracking tools such as wearable techniques, GPS or mobile, and this is known as Tracking By Detection (TBD) [19]. Face recognition, on the other hand, can be used to identify specific individuals, and provide valuable information about their whereabouts and activities. When utilized together, these technologies can provide comprehensive monitoring and surveillance capabilities. Face recognition and tracking technology employment raises several challenges and ethical concerns as discussed elaborately in [20]. One major anxiety is the potential for bias and discrimination, as the technology may not accurately identify individuals from certain racial or ethnic backgrounds. Privacy maintenance is also a significant matter, as the technology collects and stores personal data. Additionally, there are concerns regarding the technology misuse by governments or law enforcement agencies, leading to mass surveillance and violation of civil liberties. It is crucial to address these fears and establish ethical guidelines to ensure responsible and fair technology use [21]. In contrast, face recognition and tracking technology offer several benefits, including enhanced security, improved efficiency, and personalized experiences. It can help law enforcement agencies identify, and track criminals [22]. It can also streamline processes in various industries, such as banking and airports, by providing quick and secure access to services. Advanced face recognition with surveillance systems has upgraded the civil aviation smart security screening in China [23].

D. Challenges of Face Detection in Crowded Places

Detecting faces in a crowd has several challenges [24, 25], which are discussed below:

- Occlusion which refers to the situation where a part of a person's face is obscured or hidden from view, often by an object, e.g. sunglasses, a hat, a scarf, or a mask or another person. Occlusion can pose a significant difficulty to face detection algorithms because the shown part of the face may not contain enough visual information for the algorithm to accurately detect and recognize it.
- Lighting can affect the contrast and brightness of a person's face, making it either easier or more challenging for a face detection system to identify and locate faces accurately. Bad illumination can be caused by low harsh lights or varying lighting conditions, especially in video senses.
- Pose refers to the orientation or position of a person's face in a scene. Face detection systems need to consider variations in pose because faces can appear in different orientations and angles. The pose of a face can be categorized into several aspects. Frontal pose is often considered the easiest for face detection because the facial features are well-aligned and easily recognizable. Another pose is the profile pose where a person's face is turned to the side, making one or both sides of the face not visible to

the camera. The most challenging pose in face detection is the extreme pose which involves faces that are highly tilted, rotated, or otherwise not aligned with the camera. Advanced deep learning models, such as pose-invariant face recognition networks, have been developed to handle a wide range of pose variations during face recognition tasks [26].

- Facial expressions are highly variable and can change rapidly. A person's face can go through numerous expressions in a short period, making it challenging to detect and recognize a face consistently.
- In face size variability, faces can appear in images or scenes at various scales, ranging from close-up portraits to tiny faces in a crowd or distant surveillance footage. Detecting faces at different scales can be tough, especially when a face is very small or low in resolution.

E. Process Mining

Process mining [27] is a discipline within the field of Business Process Management (BPM) that focuses on analyzing and improving business processes using data-driven techniques. It entails the extraction and analysis of data from event logs, which record the activities and actions taken within an organization's IT systems [28]. The three primary goals of process mining are to discover processes automatically, measure the process compliance with business rules and optimize the process flow and performance. Process mining incorporates analyzing and visualizing processes based on event data. This tool can be used to track and understand the movement and interactions of individuals in crowded areas.

F. The Feasibility of using Process Mining as Trajectory Modeling

In this research we aim to demonstrate the potential of employing process mining to visualize traces of groups of persons. This includes tracking individuals' movements over time and across different locations. Some analogies emerge between the modeling of process flows and crowd flows, revealing similarities and differences. A set of fundamental elements is essential to gain insights into the process within the process mining context. These elements comprise case identification (case ID), activity name, and timestamps (time). Conversely, in the context of crowd flow analysis, the notion of case ID takes on a distinct connotation, distinguishing and categorizing the trajectories of individual entities within a crowd, while the activity and its related time have often the same meaning from the process perspective. A comparison between process flow and crowd flow is shown in Table I. Data Elements in process flow encompass case identification (Case ID), activity naming (Activity Name), and time-related data (Time). In contrast, crowd flow primarily involves the tracking of individual entities within a given location, specifying their identity, position (location), and time. Temporal dynamics in process flow typically consists of a sequence of events, each timestamped for analysis, whereas crowd flow involves concurrent, real-time movements of individuals within a space, which often form dynamic patterns. Modeling techniques differ. Process flow employs process models, such as Petri nets or BPMN, whereas crowd flow utilizes spatial models like density maps and simulations to represent the movements of individuals or entities. Metrics and KPIs used for assessment diverge as well—process flow relies on measures of efficiency, cycle time, and compliance, while crowd flow assesses crowd density, flow rates, and safety metrics. Visualization techniques also contrast: process flow is often represented with flowcharts or process models, whereas crowd flow uses density heatmaps and flow diagrams. Both areas find applications in various domains, but application areas differ: process flow is commonly found in manufacturing, IT, and healthcare, whereas crowd flow analysis is pertinent to events, transportation hubs, and urban planning. Challenges are varying: process flow addresses bottlenecks, compliance, and automation, whereas crowd flow considers safety, congestion, and crowd management. Considerations regarding regulatory compliance differ as well, with process flow adhering to industry standards and crowd flow adhering to safety regulations and crowd control measures. Lastly, future directions point towards automation and AI-driven optimization in process flow, while crowd flow explores applications in smart cities and continues to analyze crowd behavior.

TABLE I. ANALOGIES AND DISTINCTIONS BETWEEN PROCESS FLOW AND CROWD FLOW

Aspect	Process flow	Crowd flow
Data elements	CaseID, activity name, time	Individual/entity, location, time
Temporal Dynamics	Sequential, time-stamped events	Concurrent, real-time movement
Modeling techniques	Process models (Petri nets, BPMN)	Spatial models (e.g. density maps, simulations)
Metrics and KPIs	Process efficiency, cycle time, compliance	Crowd density, flow rate, safety
Visualization	Process flowcharts, Gantt charts	Density heatmaps, flow diagrams
Application areas	Manufacturing, IT, healthcare	Events, transportation hubs, urban planning
Challenges	Bottlenecks, compliance, automation	Safety, congestion, crowd management and emergent behavior
Regulatory considerations	Compliance with industry standards	Safety regulations, crowd control measures
Future directions	Automation, AI-driven optimization	Smart cities, crowd behavior analysis

II. CROWD MANAGEMENT AT THE HOLY MOSQUE IN MAKKAH CITY: A CASE STUDY

The holy mosque in Makkah is the largest mosque in the world. It has been expanded multiple times and can accommodate around 3 million visitors [29]. Every year pilgrims visit the holy mosque to perform Ramadan or Hajj rituals, and form a significant mass gathering religious events. Integrating video surveillance with face recognition and tracking technology can provide a powerful monitoring and surveillance solution for this kind of crowd. Moreover, analyzing live video feeds allows identifying and tracking individuals in real-time, providing valuable information to security personnel. This integration can enhance the effectiveness of video surveillance systems by automating the identification and tracking procedures, reducing the reliance on manual monitoring. Our hypothesis in this research states that engaging in frequent visits in a limited timeframe can greatly contribute to the congestion and stampede risks within the Holy Mosque. Consequently, the implementation of a visitor re-identification trajectory can serve as an enhanced tool for crowd monitoring and management.

III. THE PROPOSED SYSTEM ARCHITECTURE

The proposed system architecture is presented in Figure 1, which includes several steps: image collecting and pre-processing, face detection, face recognition, building event log and then visualizing trajectory model, which is the expected outcome of our proposed approach. The system is run on two stages: the first stage includes all steps except the recognition which occurs at the upcoming stages. The Deepface Python package acts as a framework for our experiment since it includes a set of deep learning algorithms that can be utilized for detection and recognition [30].

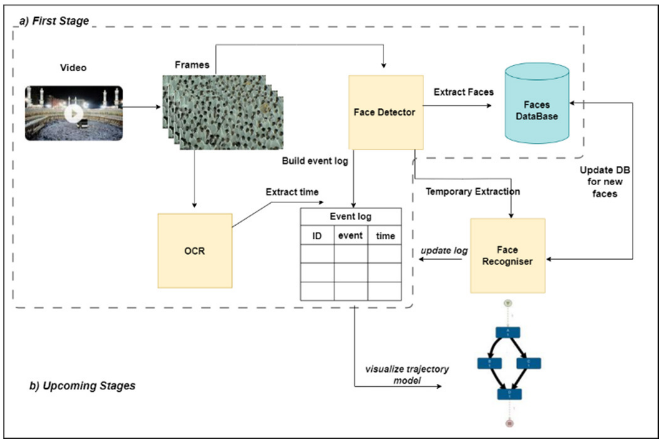

Fig. 1. Overview of the proposed system architecture.

A. *Image Collecting and Prepossessing*

Videos were downloaded from the Holy Mosque streamline public channel. We downloaded 4 videos that represent important religious events, which are the prayers of night (Isha), prayerA1, and late night (Tahajud), prayerB1, of the night of the 27th of the fasting month, and the prayers of night (Isha), prayerA2, and late night (Tahajud), prayerB2, of the night of 29th. This selection is based on our prior knowledge of the rituals. There is a possibility of re-identifying the same visitors as some people prefer to stay at the Holy Mosque to perform two subsequent prayers (Isha and Tahajud, which are prayerA and prayerB respectively), since the time interval between them is expected to be 3 hours only. The downloaded videos entail several thumbnails that do not contain any faces. Therefore, manual filtering is applied on the extracted frames to exclude the ones with no faces.

TABLE II. DATASET USED IN THIS RESEARCH

Video	Length	Frame	Frames containing faces
Video 1	01:18	324	52
Video 2	01:54	157	61
Video 3	01:42	601	73
Video 4	01:26	518	45

The images are tagged manually based on the resolution quality into 3 categories: high, low, and mixed resolution images.

B. *Detection*

There are several face detectors in the literature [31, 32], however, the Retina model [33] is adopted in this paper due to its powerful results of face detection in crowds. It can detect tiny faces due to its algorithm that depends on allocating face landmarks. The Retina model is a pretrained model which uses the WIDER FACE dataset that is a well-known face detection benchmark [34]. It consists of 32,203 carefully selected images derived from the publicly accessible WIDER pedestrian dataset [35]. The WIDER FACE dataset has 393,703 labeled faces with significant variability in terms of scale, pose, and occlusion. The Retina model is a single-stage dense face localization that integrated the existing regression and box classification methods to localize faces in images and then identify 5 face landmarks on top of the existing techniques. It aimed to generate dense 3D face vertices which improved model accuracy in detecting faces in variable scales. The model does not produce a single vector length as an output. Instead, it generates bounding boxes around the detected faces along with the associated facial landmarks. These bounding boxes and landmarks are used to localize and describe the identified faces within the image as pixel coordinates. The model can detect and align faces in different image settings such as blurred, occluded, and partial images. It can also allocate overwritten faces captured from live streaming videos. The 4 downloaded videos are ordered chronologically, hence Video1 represents the first event (prayerA1) and is engaged to run the first stage of the proposed system.

C. *Recognition*

Face recognition comes after face detection. It focuses on generating a face representation or embedding to be compared later with image pair or a database of images for verification or identification. The VGG-Face recognizer is employed in our experiment due to its robustness [36]. It is a model pretrained in the LFW (Labelled Faces in the Wild) dataset that has 2.6 M images with over 2.6K people [6]. The VGGFace model produces a vector representation of 2622 length. This model achieves high recognition rate with optimum computational performance [30].

D. *Building Event Log*

The event log is the key component of process mining. It must have three main elements: Visitor ID, which is a random generated number assigned for a visitor when her/his face is detected. Timestamp is the time when a face is observed in the target place. Event name which is a domain-related element in our experiment. Four video of prayers were used. Each video name represents an event, namely prayerA1, prayerB1, prayerA2, and prayerB2. Other attributes can enrich the event log file, and give more insight into the trajectory model. In our case, the image source is stored to help analyze the trajectory model. Also, face bounding box points can be stored as well to facilitate refereeing to the face location in the source image stored in the database for further analysis. The event log is usually saved as a CSV file and then converted to XES (eXtensible Event Stream) that is the standard format for process mining supported by most process mining tools. A sample of the created event log is presented in Table III. It shows 11 detected faces, where 7 faces are detected from the same image source "img1", 2 faces are detected from "img4", and 1 face is detected from each of "img2" and "img3". The temporal information associated with the observed faces is retrieved using a standard Optical Character Reader (OCR) embedded within the video.

TABLE III. EVENT LOG SAMPLE

Index	Face_id	Date&Time	Event	Source_img	Img_ resolution	Face coordinates
0	2456	20/4/2023 00:33	prayerB1	Img1	high	[100,50,200, 150]
1	7638	20/4/2023 00:33	prayerB1	Img1	high	[20,30,100,8 0]
2	5162	20/4/2023 00:33	prayerB1	Img1	high	[10,10,90,12 0]
3	0918	20/4/2023 00:33	prayerB1	Img1	high	[200,50,300, 250]
4	1632	20/4/2023 00:33	prayerB1	Img1	high	[75,65,155,1 85]
5	8374	20/4/2023 00:33	prayerB1	Img1	high	[30,40,110, 130]
6	6348	20/4/2023 00:33	prayerB1	Img1	high	[90,20,120,1 70]
7	9877	20/4/2023 01:40	prayerB1	Img2	mixed	[15,25,95,85]
8	0914	21/4/2023 09:40	prayerA2	Img3	mixed	[40,50,120,1 40]
9	5412	21/4/2023 09:45	prayerA2	Img4	low	[10,30,170,8 0]
10	7810	21/4/2023 09:45	prayerA2	Img4	low	[25,25,105,1 25]

E. *Computational Complexity Analysis of the Proposed System*

- Face Detection step: The computational complexity of face detection depends on the algorithm used. Suppose the detection algorithm processes each image with a time complexity of O(N). If there are M images in total, the complexity for face detection is O(M.N).
- Database Operations step: Adding a face to the database and finding a matching face depend on the data structure utilized for the database. If we assume a basic list structure, adding a face has a time complexity of O(1) on average, while finding a matching face can be O(K), where K is the number of faces in the database. Since these operations are performed for each new face in each new image, the overall complexity is O(M K), where M is the number of new images and K is the total number of faces in the database.
- Event Log step: Appending to the event log has a time complexity of O(1) per entry, and this operation is performed for each detected face in each new image, resulting in a complexity of O(M).
- Visualization step: The complexity of visualizing the trajectory depends on the visualization method, but for simplicity, let us assume it is O(N), where N is the number of entries in the event log. Overall, the total computational complexity of the algorithm is the sum of the complexities for each step:

$$\text{Total Complexity} = O(M\cdot N)+O(M\cdot K)+O(M)+O(N) \quad (1)$$

The dominant terms are typically the ones with the highest growth rate. In many cases, the face detection step (O(M·N)) and the database operations (O(M·K)) dominate the complexity.

IV. RESULTS

The created event log can be imported to any process mining tool for visualizing a trajectory model. In this research, the Python library for process mining PM4Py [4] was used for generating the graph-based trajectory model. The process model discovery algorithm represents an event name in the event log into a node. The edges connecting the nodes have labels that show the number of visitors who are observed in the corresponding events. In Figure 2, there are 1131 visitors who were observed in the first event prayerA1 and were not observed to the other selected events. However, 534 out of 1355 visitors who attended prayerA2, were re-identified later, in the video of the following event prayerB2.

Using the proposed method, a list of shots in which the person of interest appears in each matching video is retrieved. Interestingly, recognition rate was high for faces detected from high resolution images, and thus the model showed the same person appearing in different prayers (Figure 3). On the other hand, recognizing persons captured in low resolution increased the false positive recognition rate as shown in Figure 4. The face id 91017 was given to a person in img 29 0661, who was observed in prayerA2 and then the same id was given to two more persons who were actually different. This false positive is probably due to the very low-resolution of the image. We believe that readjusting camera positions in crowd places can result in better image resolution and consequently, low false positive recognition.

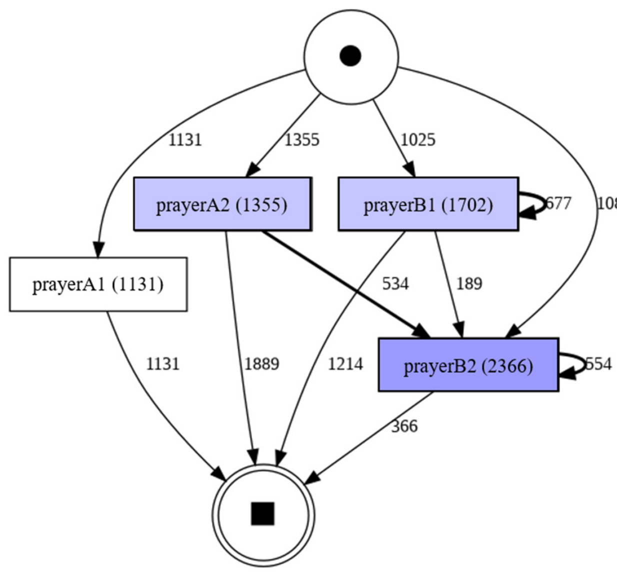

Fig. 2. Trajectory representation of recognized faces.

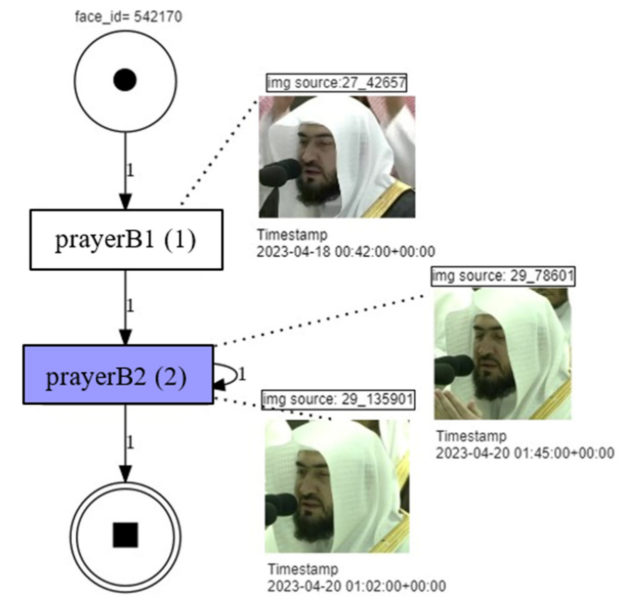

Fig. 3. Trajectory model of true positive face re-identification.

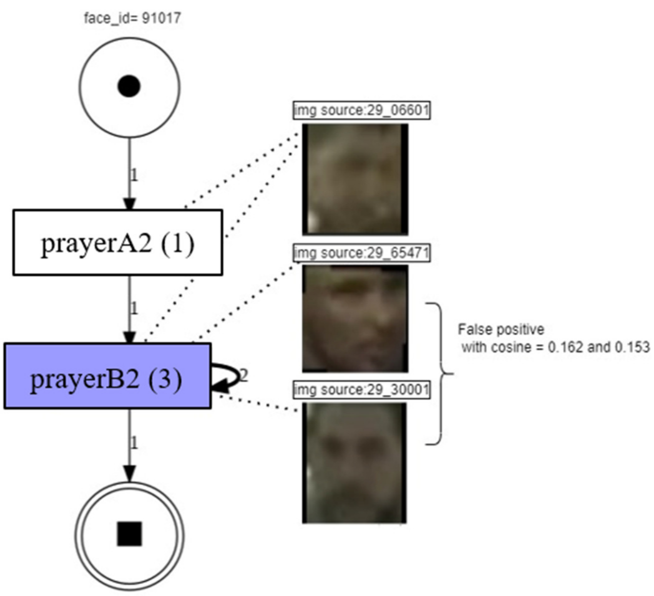

Fig. 4. Trajectory model of true positive face re-identification.

V. DISCUSSION

Utilizing process mining techniques is proved to be advantageous in constructing trajectory models. Nevertheless, the model effectiveness is significantly reliant on the availability of data elements within the event log. Therefore, the process of generating the event log and integrating diverse data sources is of utmost importance. It is essential to be cautious when using face recognition models on unlabeled identities, as the absence of multiple shots per identity makes it more challenging to quantify performance accurately. It is also crucial to consider the potential ethical implications of employing face recognition technologies in scenarios involving unlabeled identities to ensure fairness and privacy. We recommend applying image pre-processing for enhancing face resolution before the recognition phase [37]. Given the potential privacy and ethical concerns associated with face recognition and tracking technology, it is essential to establish regulations and guidelines for its ethical use. This includes ensuring transparency in how the technology is implemented, obtaining informed consent, and protecting personal data. Governments and regulatory bodies should work together with technology developers, privacy advocates, and other stakeholders to develop and enforce these regulations. By doing so, we can ensure that the technology is used responsibly and in a manner that respects individuals' privacy rights. It should be noted that, several countries and jurisdictions have started implementing legal frameworks and guidelines for face recognition technology usage in public spaces. These frameworks aim to balance the need for public safety with the protection of individuals' privacy rights. They outline the conditions under which the technology can be utilized, the limitations on data collection and storage, and the rights of individuals regarding their personal data. It is crucial for governments to continuously review and update these frameworks to keep pace with the technological advancements, and address emerging challenges. One of the major concerns with face recognition systems is the potential for bias and inaccuracy, particularly when it comes to identifying individuals from certain racial or ethnic backgrounds. This raises concerns about discrimination and the potential for false identifications. To address these concerns, it is important to ensure diversity in the datasets engaged to train the algorithms and to continuously evaluate and eliminate the bias. Hence, we aim to use the extracted faces to retrain the model and improve feature extraction. Face recognition and tracking technology can significantly enhance public safety. By identifying and tracking individuals of interest, law enforcement agencies can prevent crimes, locate suspects, and respond more effectively to incidents. The technology can be integrated with existing surveillance systems to provide real-time alerts and notifications when a person of interest is detected.

Another essential finding of this research is that, to build a robust human recognition system in unconstrained environment, the system should include more elements than plain face recognition, e.g. body features and movement should be integrated as well. Despite the advancements in face recognition and tracking technology, there are still limitations that need to be overcome. One limitation is the technology accuracy, particularly in challenging lighting conditions, or when faces are partially obscured, which can result in misidentifications. To overcome such limitations, researchers and developers are working on ameliorating the algorithms employed in the technology, as well as enhancing the hardware and software components.

VI. CONCLUSION

The future of face recognition technology holds great promise. Advancements in artificial intelligence, machine learning, and computer vision are expected to further improve the accuracy, speed, and capabilities of face recognition and tracking systems. This includes advancements in facial expression recognition, emotion detection, and the ability to identify individuals in challenging conditions. Additionally, process mining technique has shown its capabilities in modelling the trajectory of faces in crowds. Further research and development in technology to enable automatic texture face trajectory will help analyze crowd behavior. These advancements will continue to expand the applications of face recognition technology in various industries, and contribute to safer and better crowd monitoring systems. There are also serious concerns regarding privacy and ethical use. It is crucial to establish clear regulations and guidelines to ensure responsible and fair implementation of these technologies.

ACKNOWLEDGMENT

The author would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work under the Grant 23UQU43101400DSR002. She also would like to express her gratitude for the support of this research ID: 4401095348.

REFERENCES

[1] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face Recognition Systems: A Survey," *Sensors*, vol. 20, no. 2, Jan. 2020, Art. no. 342, https://doi.org/10.3390/s20020342.

[2] D. V. Becker and H. Rheem, "Searching for a face in the crowd: Pitfalls and unexplored possibilities," *Attention, Perception, & Psychophysics*, vol. 82, no. 2, pp. 626–636, Feb. 2020, https://doi.org/10.3758/s13414-020-01975-7.

[3] U. Singh, J.-F. Determe, F. Horlin, and P. De Doncker, "Crowd Monitoring: State-of-the-Art and Future Directions," *IETE Technical Review*, vol. 38, no. 6, pp. 578–594, Nov. 2021, https://doi.org/10.1080/02564602.2020.1803152.

[4] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, "YOLO-face: a real-time face detector," *The Visual Computer*, vol. 37, no. 4, pp. 805–813, Apr. 2021, https://doi.org/10.1007/s00371-020-01831-7.

[5] A. Nadeem *et al.*, "Tracking Missing Person in Large Crowd Gathering Using Intelligent Video Surveillance," *Sensors*, vol. 22, no. 14, Jan. 2022, Art. no. 5270, https://doi.org/10.3390/s22145270.

[6] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database forStudying Face Recognition in Unconstrained Environments," presented at the Workshop on Faces in "Real-Life" Images: Detection, Alignment, and Recognition, Marseille, France, Oct. 2008.

[7] M. Wang and W. Deng, "Deep face recognition: A survey," *Neurocomputing*, vol. 429, pp. 215–244, Mar. 2021, https://doi.org/10.1016/j.neucom.2020.10.081.

[8] Y. Said, M. Barr, and H. E. Ahmed, "Design of a Face Recognition System based on Convolutional Neural Network (CNN)," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5608–5612, Jun. 2020, https://doi.org/10.48084/etasr.3490.

[9] M. S. Zitouni, H. Bhaskar, J. Dias, and M. E. Al-Mualla, "Advances and trends in visual crowd analysis: A systematic survey and evaluation of crowd modelling techniques," *Neurocomputing*, vol. 186, pp. 139–159, Apr. 2016, https://doi.org/10.1016/j.neucom.2015.12.070.

[10] A. Owaidah, D. Olaru, M. Bennamoun, F. Sohel, and N. Khan, "Review of Modelling and Simulating Crowds at Mass Gathering Events: Hajj as a Case Study," *Journal of Artificial Societies and Social Simulation*, vol. 22, no. 2, 2019, Art. no. 9, https://doi.org/10.18564/jasss.3997.

[11] A. B. Altamimi and H. Ullah, "Panic Detection in Crowded Scenes," *Engineering, Technology & Applied Science Research*, vol. 10, no. 2, pp. 5412–5418, Apr. 2020, https://doi.org/10.48084/etasr.3347.

[12] C.-T. Lu, C.-J. Ou, and Y.-Y. Lu, "A Practical App for Quickly Calculating the Number of People Using Machine Learning and Convolutional Neural Networks," *Applied Sciences*, vol. 12, no. 12, Jan. 2022, Art. no. 6239, https://doi.org/10.3390/app12126239.

[13] L. Wang and A. A. Siddique, "Facial recognition system using LBPH face recognizer for anti-theft and surveillance application based on drone technology," *Measurement and Control*, vol. 53, no. 7–8, pp. 1070–1077, Aug. 2020, https://doi.org/10.1177/0020294020932344.

[14] Y. Fang, B. Zhan, W. Cai, S. Gao, and B. Hu, "Locality-Constrained Spatial Transformer Network for Video Crowd Counting," in *International Conference on Multimedia and Expo*, Shanghai, China, Jul. 2019, pp. 814–819, https://doi.org/10.1109/ICME.2019.00145.

[15] G. Keith Still, "Crowd Science and Crowd Counting," *Impact*, vol. 2019, no. 1, pp. 19–23, Jan. 2019, https://doi.org/10.1080/2058802X.2019.1594138.

[16] R. P. Holder and J.-R. Tapamo, "Using Facial Expression Recognition for Crowd Monitoring," in *8th Pacific Rim Symposium on Image and Video Technology*, Wuhan, China, Nov. 2017, pp. 463–476, https://doi.org/10.1007/978-3-319-75786-5_37.

[17] J. Nelson, "Access control and biometrics," in *Handbook of Loss Prevention and Crime Prevention*, 6th ed., L. J. Fennelly, Ed. Oxford, UK: Butterworth-Heinemann, 2020, pp. 239–249.

[18] K. B. S. Reddy, O. Loke, S. Jani, and K. Dabre, "Tracking People In Real Time Video Footage Using Facial Recognition," in *International Conference on Smart City and Emerging Technology*, Mumbai, India, Jan. 2018, pp. 1–6, https://doi.org/10.1109/ICSCET.2018.8537296.

[19] P. Bergmann, T. Meinhardt, and L. Leal-Taixe, "Tracking Without Bells and Whistles," in *International Conference on Computer Vision*, Seoul, South Korea, Nov. 2019, pp. 941–951, https://doi.org/10.1109/ICCV.2019.00103.

[20] M. Smith and S. Miller, "The ethical application of biometric facial recognition technology," *AI & Society*, vol. 37, no. 1, pp. 167–175, Mar. 2022, https://doi.org/10.1007/s00146-021-01199-9.

[21] M. A. P. Chamikara, P. Bertok, I. Khalil, D. Liu, and S. Camtepe, "Privacy Preserving Face Recognition Utilizing Differential Privacy," *Computers & Security*, vol. 97, Oct. 2020, Art. no. 101951, https://doi.org/10.1016/j.cose.2020.101951.

[22] A. K. Chaitanya, C. H. Kartheek, and D. Nandan, "Study on Real-Time Face Recognition and Tracking for Criminal Revealing," in *Soft Computing: Theories and Applications*, M. Pant, T. Kumar Sharma, R. Arya, B. C. Sahana, and H. Zolfagharinia, Eds. New York, NY, USA: Springer, 2020, pp. 849–857.

[23] Y. Shi *et al.*, "'One-Time Face Recognition System' Drives Changes in Civil Aviation Smart Security Screening Mode," in *China's e-Science Blue Book 2020*, Singapore, Singapore: Springer, 2021, pp. 399–425.

[24] J. Chen, W. Su, and Z. Wang, "Crowd counting with crowd attention convolutional neural network," *Neurocomputing*, vol. 382, pp. 210–220, Mar. 2020, https://doi.org/10.1016/j.neucom.2019.11.064.

[25] G. Gao, J. Gao, Q. Liu, Q. Wang, and Y. Wang, "CNN-based Density Estimation and Crowd Counting: A Survey." arXiv, Mar. 28, 2020, https://doi.org/10.48550/arXiv.2003.12783.

[26] O. Elharrouss, N. Almaadeed, S. Al-Maadeed, and F. Khelifi, "Pose-invariant face recognition with multitask cascade networks," *Neural Computing and Applications*, vol. 34, no. 8, pp. 6039–6052, Apr. 2022, https://doi.org/10.1007/s00521-021-06690-4.

[27] W. van der Aalst, "Process Mining: Overview and Opportunities," *ACM Transactions on Management Information Systems*, vol. 3, no. 2, Apr. 2012, Art. no. 7, https://doi.org/10.1145/2229156.2229157.

[28] A. B. Lashram, L. Hsairi, and H. A. Ahmadi, "HCLPars: A New Hierarchical Clustering Log Parsing Method," *Engineering, Technology & Applied Science Research*, vol. 13, no. 4, pp. 11130–11138, Aug. 2023, https://doi.org/10.48084/etasr.6013.

[29] "The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research - University President Office | Umm Al-Qura University." https://uqu.edu.sa/en/hajj.

[30] S. I. Serengil and A. Ozpinar, "LightFace: A Hybrid Deep Face Recognition Framework," in *Innovations in Intelligent Systems and Applications Conference*, Istanbul, Turkey, Oct. 2020, pp. 1–5, https://doi.org/10.1109/ASYU50717.2020.9259802.

[31] M. O. Oloyede, G. P. Hancke, and H. C. Myburgh, "A review on face recognition systems: recent approaches and challenges," *Multimedia*

Tools and Applications, vol. 79, no. 37, pp. 27891–27922, Oct. 2020, https://doi.org/10.1007/s11042-020-09261-2.

[32] C. Li, R. Wang, J. Li, and L. Fei, "Face Detection Based on YOLOv3," in *Recent Trends in Intelligent Computing, Communication and Devices*, V. Jain, S. Patnaik, F. Popențiu Vlădicescu, and I. K. Sethi, Eds. New York, NY, USA: Springer, 2020, pp. 277–284.

[33] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, Jun. 2020, pp. 5202–5211, https://doi.org/10.1109/CVPR42600.2020.00525.

[34] S. Yang, P. Luo, C. C. Loy, and X. Tang, "WIDER FACE: A Face Detection Benchmark," in *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, Jun. 2016, pp. 5525–5533, https://doi.org/10.1109/CVPR.2016.596.

[35] Y. Xiong, K. Zhu, D. Lin, and X. Tang, "Recognize complex events from static images by fusing deep channels," in *IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, Jun. 2015, pp. 1600–1609, https://doi.org/10.1109/CVPR.2015.7298768.

[36] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in *Procedings of the British Machine Vision Conference 2015*, Swansea, UK, 2015, pp. 41.1-41.12, https://doi.org/10.5244/C.29.41.

[37] K. Dharavath, F. A. Talukdar, and R. H. Laskar, "Improving Face Recognition Rate with Image Preprocessing," *Indian Journal of Science and Technology*, vol. 7, no. 8, pp. 1170–1175, Aug. 2014, https://doi.org/10.17485/ijst/2014/v7i8.26.