Privacy-preserving AI-enabled video surveillance for social distancing: responsible design and deployment for public spaces

Privacy-preserving AI-enabled video surveillance for social distancing: Responsible design and deployment for public spaces

Journal:	*Information Technology & People*
Manuscript ID	ITP-07-2020-0534.R2
Manuscript Type:	Article
Keywords:	Surveillance < Phenomenon, Information management < IT/IS management < Practice, Ethics, Agile computing < Technology, knowledge transfer < Phenomenon, Adoption < Information system development < Practice

Privacy-preserving AI-enabled video surveillance for social distancing: Responsible design and deployment for public spaces

Abstract

Purpose – The paper proposes a privacy-preserving artificial intelligence-enabled video surveillance technology to monitor social distancing in public spaces.

Design/methodology/approach – The paper proposes a new Responsible Artificial Intelligence Implementation Framework to guide the proposed solution's design and development. It defines responsible artificial intelligence criteria that the solution needs to meet and provides checklists to enforce the criteria throughout the process. To preserve data privacy, the proposed system incorporates a federated learning approach to allow computation performed on edge devices to limit sensitive and identifiable data movement and eliminate the dependency of cloud computing at a central server.

Findings – The proposed system is evaluated through a case study of monitoring social distancing at an airport. The results discuss how the system can fully address the case study's requirements in terms of its reliability, its usefulness when deployed to the airport's cameras, and its compliance with responsible artificial intelligence.

Originality – The paper makes three contributions. First, it proposes a real-time social distancing breach detection system on edge that extends from a combination of cutting-edge people detection and tracking algorithms to achieve robust performance. Second, it proposes a design approach to develop responsible artificial intelligence in video surveillance contexts. Third, it presents results and discussion from a comprehensive evaluation in the context of a case study at an airport to demonstrate the proposed system's robust performance and practical usefulness.

Keywords artificial intelligence, responsible AI, data privacy, social distancing, video surveillance technology

Paper type Research paper

1. Introduction

COVID-19 is a global pandemic that spreads quickly through close human contact with infected patients. The virus had infected 10 million people in June 2020 and 38 million people by October 2020 (World Health Organization [WHO], 2020). Despite the economic impact of lockdown, government-imposed social distancing in public spaces is critical for slowing the infection rate (Department of Health, 2020). Monitoring compliance of social distancing in public spaces is usually done by the organization (government or business entities like shopping center management). Manual monitoring through video surveillance to identify breaches is arguably the least invasive method, but tedious and should be automated. Artificial Intelligence (AI) can automate analysis of crowd density and detect social distancing breaches. When crowd density becomes critical and there are too many occurrences of people failing to maintain safe distances, the automatic system would immediately trigger a notification to authorized personnel to take action. By observing long-term patterns of crowd density, movement and spatial distancing, local government and business owners can also gain insights for redesigning their space to better manage social distancing (Boland et al., 2020). AI can enhance the usefulness of existing video surveillance systems already installed in public spaces, and can be integrated as part of a smart city's intelligent infrastructure to maintain safe distancing and prevent COVID-19 spread (Feldstein, 2019).

Video surveillance raises data privacy concerns when humans are the interest of observations. Extensive use of such technology can threaten people's privacy and cause harm to people. Most existing AI-enabled video surveillance systems rely on centralized processing, meaning data are transmitted to a central/cloud machine for video analysis (Kavalionak et al., 2019). Such an approach involves data privacy and data security risks (Sanchez-Iborra and Skarmeta, 2020). Serious data threats such as data theft, eavesdropping or cyber-attacks, can potentially occur during data transmission. Likewise, when AI-enabled video surveillance systems incorporate AI continual learning to adapt the AI models towards constant visual change in the environment, existing approaches require new data to be collected from cameras and retained long term (Sugianto et al., 2019). This also involves data privacy concerns due to potential data misuse. Existing laws in Asia-Pacific countries require data to be de-identified or deleted immediately after use (Kennedys Law, 2018). While many countries have established privacy laws to protect their people (Clifton et al., 2020), such emerging technology still lacks the much-needed people's and community trust as it is often considered technology that invades their privacy. When people's trust is lacking, AI solutions are less likely to be adopted regardless of the intended benefits (Rossi, 2018).

This paper proposes a privacy-preserving AI-enabled video surveillance technology to monitor social distancing in public spaces. To preserve data privacy, the proposed system incorporates a federated learning approach to allow computation performed on edge devices rather than transmitting sensitive and identifiable data to the central machine.

The paper makes three contributions. First, it proposes a real-time social distancing breach detection system on edge that extends from a combination of cutting-edge people detection and tracking algorithms to achieve robust performance. Second, it proposes a design approach to develop responsible AI in video surveillance contexts. The approach is guided by a new Responsible AI Implementation Framework (RAIFF) to enforce responsible AI criteria throughout designing, developing, and deploying the system. Third, it presents results and discussion from a comprehensive evaluation in the context of a case study at an airport to demonstrate the proposed system's robust performance and practical usefulness.

Case Study

The proposed AI solution was designed based on a case study at an airport which has 6.5 million (pre-COVID-19 impacts) passengers annually with 17,000+ passengers daily. Hundreds of integrated surveillance cameras were installed to cover 290,000 square metres with hundreds of shops and 40+ check-ins. To ensure all passengers have a good experience during their visits, the airport employs some trained staff to observe people's activities across all areas through surveillance cameras to ensure that the airport's activities run smoothly, and the airport provides good service to passengers. When a problem is noticed, they will notify other staff to respond immediately. For example, when the number of people in a certain area is above a safe number, in-field staff will be notified and asked to respond immediately to ensure people's safety. Given the complexity of such environments, it is humanly impossible to continuously observe the airport's activity just by relying on human effort. Existing human resources are limited in constant manual observation given the large number of cameras to observe. Human distraction and fatigue may result in missing some important events while observing cameras.

The intelligent AI-based video surveillance system can help airport personnel observe people's activities. The AI solution will provide prompt airport activities analysis to the airport's management in real-time. Authorized staff will receive automatic notifications when there is something urgent or critical and needs immediate follow up by personnel. The use of such intelligent video surveillance technology is proposed as we consider existing available resources at the airport given that surveillance cameras and the associated infrastructure are already well established.

Data privacy is the main concern as the proposed system will constantly observe people's activities. While the airport realizes the benefits of such an intelligent system, they have consistently raised data privacy concerns since the beginning of the project. The proposed system is expected to be able to deliver the benefits while not compromising passenger privacy and obeying existing laws and human ethics. As initial actions, strict agreement has been made to protect any sensitive data such as people's identities. People's faces must be blurred and videos can only be stored up to 7 days. The main goal is to develop an intelligent video surveillance while ensuring data privacy protection.

Adopting AI solutions at the international airport has evolved over time alongside addressing the system's AI responsibilities. Prior experience with the AI solution sets precedence for the airport's technology manager to trust the development of the new capabilities. Table I summarizes the evolution of an AI solution at the airport, starting from people detection and tracking to social distancing breach detection. In 2017, we initially developed automatic people detection and tracking to investigate the feasibility of the proposed intelligent system. Since people are the observation's point of interest at the airport, detecting and tracking people is crucial to establish the foundation of the system prior to further system development. In 2018, we extended the system to automatically count crowd levels to ensure the number of people in certain areas was below safe numbers. During the COVID-19 pandemic in 2020, the Australian government enforced businesses to conduct strict safety responsibilities such as practicing social distancing during business operation. In response, the airport must rearrange its space to ensure passengers have safe distances between them during their visit. Space rearrangement is mostly applied to critical areas, such as waiting areas and queue lines, where people will gather in large numbers. Thus, it motivates us to extend the system further to detect any social distancing breaches and generate a heat map for spatial analysis purpose. By having such information constantly, the airport can understand how well their social distancing practice works, evaluate and plan to rearrange their space to better ensure people's safety.

---Insert Table I here----

2. Related Work

Existing solutions to help to monitor social distancing among people can be categorized into wireless, sensory signals and vision-based tracking devices. Both wireless and sensor technologies can only work if everyone within the observation range uses the same device or tag and cannot present optimum performance without people's cooperation in enabling the signal transmission. Wireless location-based technologies track individuals by picking up signals from their electronic devices, such as mobile phones, to get their location. Apple and Google recently released mobile applications intended for tracing COVID-19 infected people using Bluetooth signals (Romm et al., 2020). Aside from Bluetooth, other signals such as Wi-Fi Received Signal Strength Indicator (RSII), cellular, Ultra-WideBand (UWB),

Global Navigation Satellite System (GNSS), and Zigbee signals are potential solutions for social distancing verification (Al-Suwaidi and Zemerly, 2009, Laoudias et al., 2018, Ledergerber and D'Andrea, 2020, Mazuelas et al., 2009, Niu et al., 2015). Sensor technologies pick up ultrasound, inertial, visible light or thermal frequency signals from specific tags attached to the users (Jovicic and Engineer, 2016, Kok et al., 2015, Szajewska, 2017, Tiemann and Wietfeld, 2017). The signal is then calculated using mathematical models to alert the users when social distancing is breached (Nadikattu et al., 2020). In contrast, vision-based technology uses computer vision to automatically track people's movements without requiring any devices attached to each individual.

2.1. AI Vision Algorithms for Monitoring Social Distancing

The current challenge in existing AI-enabled vision-based people detection and tracking is to produce an accurate yet real-time system in complex and dynamic environments such as public spaces. To cover large areas, the AI-system would need to work with an integrated multi-camera setting, hence it must be robust enough to analyze video streams that have varying characteristics due to the camera's deployment environments.

People detection, tracking, and crowd counting originates from classifying and localizing people continuously across the captured image streams, commonly known as object detection (Hu et al., 2018). Current object detection models can be categorized into three types: two-stage, one-stage and mobile-based models. Two-stage models are known for their very accurate detection but slower inference time in applications (Ren et al., 2015, Zhou et al., 2017). One-stage models are designed to improve the inference time of two-stage models while still achieving comparable accuracy (Liu et al., 2016, Redmon and Farhadi, 2018). Mobile-based models are the compact version of the previous models dedicated to specific applications on smaller computing devices such as mobile and edge devices (Wang et al., 2018). Mobile-based models are built on one-stage model with smaller neural networks for faster and cheaper computation (Biswas et al., 2019). People tracking applies the detection model into the image stream using specific tracking algorithms to stitch the detection information. The key in people tracking is to lock onto every targeted object in the video frame, uniquely identify each one of them and track them until they exit the video frame. Spatio-temporal features are essential for this tracking process. Existing work can be categorized into traditional and deep learning methods.

Traditional methods encompass mean-shift, optical flow, and Kalman filter. Mean-shift algorithms examine the distribution mode of features of an object and track the change in the distribution across the frames to track the object (Iswanto and Li, 2017). Optical flow, on the other hand, does not necessarily use the object features but the Spatio-temporal brightness variations of the image frames (Yang et al., 2019b). It applies an equation and some prediction techniques across the frames to track the object. Kalman filter is well known to solve temporal-related vision machine learning problems (Gunjal et al., 2018). It uses the available detections and prior predictions to predict the best possible tracking, while considering the error probability of the task. Despite the usefulness of Kalman filter, it is unsuitable for real-world applications due to occlusions and different viewpoints in the integrated surveillance system.

Prominent deep learning methods for object tracking are deep regression networks, Recurrent YOLO, Simple Real-time Tracker (SORT) and Deep SORT (Bewley et al., 2016, Held et al., 2016, Wojke et al., 2017, Yun and Kim, 2019). Deep regression networks are one of the early deep learning methods. It uses a two-frame Convolutional Neural Network (CNN) architecture on current and previous video frames to regress and localize the targeted objects. Recurrent YOLO uses a one-stage detection model, YOLO, and combines it with LSTM to process both spatial and temporal features of the images providing significant improvements on early tracking methods. SORT is designed to handle object tracking in real-time applications, by leveraging the benefit of CNN for detection with Kalman filter and Hungarian algorithm to improve tracking performance. The pragmatic SORT approach results in comparable accuracy to state-of-the-art methods and faster computation. Deep SORT is an improvement on SORT, where it incorporates deep association metrics pre-training to improve the tracking accuracy during occlusions. The deep metric lowers the chance of false tracking by ensuring the system tracks the correct person despite repeated occlusions.

Social distancing requirements due to COVID-19 have led to some developments in vision-based technology. An initial study compared sensor-based and vision-based systems on pinpointing coordinates of people using AI (Elloumi et al., 2016, Johnson Jr et al., 2020). The vision-based system demonstrated superior accuracy, where it uses images captured from smartphone cameras and applied people detection and Kalman Filter to compute the coordinates. Another idea is to use drones to track people and measure their temperature to determine the likelihood of COVID-19 infection, but this is still under development and not designed to manage social distancing (Daly, 2020). Another proposed approach was to track people's movement from cameras, but privacy issues were not addressed as it used a centralized computation approach (Punn et al., 2020).

2.2. Law and Ethics for Privacy

Data privacy is still an open and challenging issue in AI, especially when dealing with sensitive and personal data. The use of AI for business or by any organization has been deemed a threat to people's privacy as a system's intelligence is reliant on gaining as much customer data as possible (Benson et al., 2015). For instance, many websites have collected data regardless of customers' consent by tracking online behavior and browsing history and sold it as a commodity for marketing purposes (Liyanaarachchi, 2020). While wireless and sensor-based technologies require consent from people before sharing their locations, vision-based technology collects data, including sensitive data such as people's faces, without explicit permission (Australian Government Office of the Australian Information Commissioner, 2019). While collecting more data from cameras to update AI models can improve the models' robustness, it also raises potential serious threats such as data misuse and data theft during transmission.

The increasing number of AI applications has led many countries, specifically Asia-Pacific countries and United States of America (US), to establish privacy acts (Clifton et al., 2020). These regulations indicate the importance of lawful data collection, processing, retention and use of personal data to protect people's privacy. Fair and lawful data collection is regulated across Asia-Pacific (Kennedys Law, 2018), the US (Department of Justice, 2015), and European Union countries (General Data Protection Regulation, 2016), where they require people to know and understand when their data are collected. The acts indicate the importance of upholding data privacy for any collection and processing system, as seen in the inclusion of privacy-by-design principle in the US Privacy Act of 1974. Any organizational body needs to provide notice to people before or during data collection, or in certain countries such as Australia (Office of the Australian Information Commissioner, 2019) and South Korea (Personal Information Protection Commission, 2011), as soon as possible after collection. However, the latter case also requires de-identifying of personal data during collection when possible. The majority of Asia-Pacific countries also require data de-identifying or deletion, specifically after use. The privacy acts also encompass visual surveillance, particularly countries which surveillance systems are well established in public spaces, such as South Korea, Australia and Macau (Office for Personal Data Protection, 2005). Therefore, it is critical and required by law to develop AI-enabled video surveillance systems without breaching data privacy.

A recent study shows that data privacy can be protected using the federated learning approach (Yang et al., 2019a). The problem with the common machine learning approaches in the surveillance system is that data, such as images, need to be sent to the central system's storage and stored for a long time to augment the models' knowledge. The application of intelligent vision application in public spaces will require the system to store and process countless image streams in the central system storage. The federated learning approach enables AI to train an intelligent central model from many smaller branch models by edge computing. This mechanism allows the central model to retrieve the trained AI model (knowledge) from branch models and no longer store the actual image data without compromising model accuracy. Thus, any sensitive data in the image data will be immediately discarded from each camera and not stored in the central system, upholding the law and ethics of data privacy.

2.3. Summary of Current Gaps

Existing research on AI development to help maintain social distancing has not fully considered privacy preservation. Removing person-identifiable information, such as facial data, and training the AI model on edge can potentially support better privacy preservation. The proposed solution adopts on-edge computing to process video on the camera itself, so that no sensitive data is transmitted to the central machine. When processing is performed on edge, the model is limited to learn from its local data and only benefits its device, hence the system still requires aggregation of local model updates across edge devices as well. The federated learning approach enables aggregation of the local model updates without sharing its actual data. For a large site like airports, aggregating local models from multiple cameras can improve the robustness of the overarching AI model.

3. Proposed Framework

The proposed framework consists of three components, namely a design approach for responsible AI, a new AI system designed for monitoring social distancing via video surveillance, and novel algorithms and architectures to support the system.

3.1. Design Approach for Responsible AI-Enabled Video Surveillance

To promote technology integration, our system's design and development was guided by a new Responsible AI Implementation Framework (RAIIF – see Figure 1), which extends the human-centered design framework (Rosenbrock, 2012). The framework emphasizes iterative co-creation processes underpinned by trust establishment and maintenance. It aims to streamline the AI adoption process by involving all stakeholders throughout the implementation process, from planning and design to development and scaling-up. An iterative and agile approach applies to the whole journey of

planning, design, development, deployment, operating, monitoring, and scaling of the AI solution. Each process may be iterated on its own or after the whole journey (from planning to scaling) of the initial AI solution is completed.

The goal of the *planning and designing* process is to achieve AI trustability, which encompasses technological accessibility and impacts consideration in the existing human and technology infrastructures. The initial AI models and algorithms need to determine the business model's viability, technological feasibility, and human desire (to adopt the system). The *development and deployment* process aims to ensure that the AI solution is robust, consistently accurate, meets expectations, and compliant until it reaches the point where the AI solution can be scaled up into full production. The *operating, monitoring, and scaling* process aims to achieve optimized operations, empowered employees, engaged customers, and transformed products. The objective is promoting AI solutions to harmoniously co-exist with businesses and the community while maintaining individual and societal trust in the technology amid perceived risks.

---Insert Figure 1 here---

The framework provide checkpoints to observe the following responsible AI criteria.

Reliable: the AI behaves as expected in every possible set of circumstances, even for novel inputs (unseen inputs not covered in training data). When AI is used in public spaces, the AI solution must work well against any variations such as occlusion, head pose, illumination, gender, or race. Low-reliable AI can result in misleading outcomes and possible huge risks (e.g. loss of human life, finance loss, social risk) especially when deployed in critical and sensitive application domains such as in health. When the change is constantly introduced by the environment where AI is deployed and significantly affects AI performance, continuous learning capability must be considered while developing AI solutions (Parisi et al., 2019). In surveillance systems, continuous learning aims to allow the AI models to learn from new data collected from many cameras.

Non-biased: Biased inferences can affect AI's reliability which can lead to unfairness issues such as racism problems. When AI outcomes may be distorted, the AI users should be explicitly warned (Ghallab, 2019). Bias can result from two sources: training data and algorithms. From the algorithm aspect, the nature of the algorithm used must be thoroughly investigated under any circumstances to ensure its reliability. Although deep learning has shown its high accuracy in classification problems, it is a black box function approximator with limited interpretability as accuracy drops to nearly zero in the presence of small adversarial perturbations (Raghunathan et al., 2018). From the data perspective, unbalanced training data distribution (e.g. age, gender, race) can cause the model to favor over-sampled classes rather than under-sampled classes. In surveillance systems, collecting new data from cameras across different locations can help to balance the dataset distribution issue.

Data privacy protection: Privacy of personal data is essential when accessing sensitive and personal data, such as facial data (Domingo-Ferrer and Blanco-Justicia, 2020). According to General Data Protection Regulation (GDPR), data privacy breaches happen when collected data are not used for the targeted purposes (Sultan and Jensen, 2020) or accessed by unauthorized people. Such protection must be taken as a proactive approach rather than a reactive approach, meaning it must be anticipated before any individuals are affected or hurt (Sultan and Jensen, 2020). Privacy protection of sensitive data must be applied in any data activities, from collecting (when data are captured by surveillance cameras), processing (when data are processed and analyzed at a processing machine), storing (when data are retained for period of time at a machine) to moving data activities (when data are transferred from cameras to a central or cloud machine). Therefore, surveillance systems must take this responsibility into account when the system is initially designed until fully deployed in operation. A poorly-design surveillance system portrays a false sense of security and violates the fundamental privacy rights of individuals (General Data Protection Regulation, 2016).

Transparent AI decisions: Transparency of how AI works in making decisions can help organizations be aware of risks and mitigate issues of fairness and trust. Transparency of how AI models make automatic decisions means understanding the assumptions, limitations, and criteria. The AI must be able to explain and justify the response generated to the input. Moreover, transparency is also required in the context of data privacy and security (Burt, 2019). How the system deals with any identifiable and sensitive data must be carefully designed and any issues occuring must be highlighted. This must be done in terms that are understandable to the users. More transparency can help reduce people's unfounded fears of AI (Ghallab, 2019). By knowing this, the organization can decide how far the AI system will do the automation. In some cases, a human-in-the-loop approach may be required during the process of decision making as AI progresses until humans cannot add new value to it (Ransbotham, 2017).

Sustainable and scalable for long-term use: Sustainable and scalable AI can contribute to developing trust for users. High investment in AI requires high commitment from stakeholders. Therefore, people demand high assurance that the technology can be scaled up to contribute in wider ranges and longer periods of time. The acceptability of a technology needs to take into account the long term and global constraints (Ghallab, 2019). The technology infrastructure can be used across different levels of organizations.

Table II to Table IV outline the checklists for each process to enforce these criteria.

---Insert Table II to Table IV here---

3.2. System Design

Figure 2 illustrates how the system preserves privacy by ensuring that each camera processes video on edge device, known as on-edge computing, and only uses representative data for calculating social distancing.

---Insert Figure 2 here---

The purpose of on-edge computing is to enable surveillance cameras with computation capability. It aims to allow any computations to be performed at the device hence eliminating the dependency of cloud computing at the central server. Such approaches preserve any sensitive and identifiable data to reside at the device where they are collected. There are two computations performed on edge devices: 1) to detect social distancing breaches and 2) to perform AI continuous learning.

To meet the AI reliability criterion, the system employs AI continuous learning to continuously adapt the people detector model against any changes introduced by the environment. New images are periodically collected from cameras across locations to build new training data that represent any variation and to meet the non-biased inference criterion. To allow continuous learning without transmitting any sensitive and identifiable data, a federated learning approach (Konečný et al., 2016) is employed to meet the data privacy protection criterion.

Federated learning enables smart cameras to continuously update their local models by performing local inference and training within the device and update the central model by aggregating local model updates without sending actual images. An edge device aims to collect and process data from a camera and perform local training to update its local model. The central machine has one global model and aims to aggregate knowledge from edge devices' local models. The central machine does not have direct access to private data from surveillance cameras.

Figure 3 explains the process of detecting social distancing breaches within each smart camera. Initially, people's locations are detected within frames by a pedestrian detector model (Step 1). For any people detected, each person is tracked over frames using an object tracking model (Step 2) and an anchor point is determined based on the person's foot position representing the person's position in 2D projection (Step 3). Any anchor points are then transformed to 2D projection using matrix H (Step 4). Figure 4 illustrates how the camera scene is projected from 3D to 2D scene using the Homography algorithm. The number of people is then calculated based on people being tracked (Step 5) and crowdedness level is calculated (Step 6). Spatial distance between people is then calculated using Euclidean distance (Step 7). Based on the distance, social distance breaches can be determined (Step 8). Any people that are below the safe distance threshold will be drawn in green. Any people that are higher than the safe distance will be drawn in red including the distance. To provide social distancing breach analysis, a heat map is generated based on cumulative distancing breach detection and updated periodically (Step 9). A breaching score is calculated where the incident happens (between 0-1). The closer the distance is, the higher the breaching score. The breaching score will be accumulated throughout the day and will be reset to 0 for the next day. Five levels of breach are used to visualize the area, from low to critical breach level. Breach level is determined by the number of distancing breach occurrences and how close the breaches are. The closer the distancing breach, the higher the multiplier put into the calculation. All spatial information is stored in a JSON file and then transmitted to the server (Step 10). When the breach becomes critical, automatic notifications will be triggered and sent to authorized people via email and phone (Step 11).

---Insert Figures 3 and 4 here---

3.3. On-edge Algorithms and System Architecture

The system employs two cutting-edge algorithms to detect people (Step 1 at Figure 3) and track people (Step 2 at Figure 3) to achieve robust performance. Lightweight computation without compromising accuracy is the main consideration while employing algorithms for on-edge deployment.

People detector model: SSD-MobileNet V2 object detector model is employed to detect people within frames (Sandler et al., 2018). The model is a single-state object detector which detects 90 types of objects from image pixels and returns the detected objects' bounding box coordinates and class confidence. It was trained on MS-COCO dataset, a large-scale object dataset consisting of 300,000 images. For our purpose, we only used the person (or pedestrian) class and disregarded the other classes. The model architecture is based on inverted residual structure where input and output of residual blocks are thin bottleneck layers as opposed to traditional residual models. It refines its initial model by replacing expensive-to-compute convolutional layers with depth-wise separable convolution blocks. Compared to

similar networks, the model has about 9 times less computation while achieving same accuracy with higher FPS, enabling real-time processing.

People tracking model: Simple Online and Realtime Tracking with a deep association metric (Deep SORT) model is employed to track people over frames (Wojke et al., 2017). The purpose of tracking people is to count the same person as one object during its movement over frames. It is an online object tracking approach which is focused on simple yet effective real-time tracking. Improved from its initial version, the algorithm integrates visual appearance information aiming to improve performance by tackling longer periods of occlusion in the real world yet reducing the number of identity switches by 45%.

Figure 5 outlines important steps in how to initiate and perform federated learning. The steps are repeated periodically to keep updating the models based on newly collected data.

---Insert Figure 5 here---

Step 1: At the central machine, a base model with pre-trained weights is adopted as the initial global model. The model is initially trained and fine-tuned using existing people datasets. Then, the global model is deployed to all edge devices as their initial model. This is only executed once. Whenever a new camera is deployed, the most up-to-date global model can be deployed as the initial local model.

Step 2: At each edge device, continuous training is performed periodically (managed by a scheduler) using a fine tuning approach (Käding et al., 2016) to improve the local model's performance by learning from newly collected data. The model update will be then generated once local training is finished and all local images are immediately removed.

Step 3: The central machine will periodically notify all edge devices if they have any local updates and request their participation in sending local updates to update the global model. Edge devices may refuse to participate when they do not have any latest local model updates or for any other reasons.

Step 4: For any edge devices that are willing to participate, they respond to the central machine and send their local model's update obtained from local training. Any data compression and encryption methods can be applied to secure data sent during transmission and avoid data misuse.

Step 5: The central machine updates its global model by aggregating all local model updates using an aggregation algorithm (Wang et al., 2020, Abadi et al., 2016). Once this step is finished, it will return to Step 2 in the next cycle

Edge devices can schedule local training when the device is idle. Ideally, night is best time to perform local training but this depends on the context and the business operating hours. Schedules can be set manually according to each condition but only executed when in idle operation.

4. Results and Evaluation

The proposed system is evaluated in terms of 1) its reliability by validating that the algorithm can achieve state-of-the-art performance in the context of the case study, 2) its usefulness based on its robustness for real deployment in the airport's selected cameras, and 3) its compliance with responsible AI implementation, based on meeting criteria and checklists. It should be noted that the results are presented with people's faces are blurred to protect their identities in accordance with the ethics agreement.

Automatic people detection and tracking

Table V shows the performance comparison of object tracking algorithms evaluated on the Multiple Object Tracking Benchmark 2016 (MOT16 challenge). Several metrics are used to evaluate the model performance. Compared to the initial version, DeepSort could reduce identity switches (ID) by 45%, from 1423 to 781, and increase track fragmentation (FM) slightly. There is also significant increase in the number of mostly tracked objects (MT) and a decrease in mostly lost objects (ML). The incorporation of visual appearance information could maintain identities longer against the occlusion in real-world settings. In terms of run time, half the time was spent on feature generation. However, the use of GPU embedded in the edge device will significantly speed up computation hence achieve real-time processing. A further detailed comparison can be found in the referred paper (Wojke et al., 2017).

---Insert Table V here---

The performance was evaluated on four cameras by observing how people are detected and tracked by the system compared to ground truth. We were unable to evaluate the automatic people detection and tracking since there is no available annotation of the airport's cameras. Therefore, performance evaluation will be further investigated in automatic crowd counting and social distancing breach detection (discussed in the next two sections) since people

detection and tracking performance will greatly affect crowd counting and social distancing breach detection performance.

Through observation and experiments, we found that camera angle plays as an important role in obtaining good people detection. About 45 to 60 degree is a good camera angle to capture, detect and track people's movement in a public area. Detection accuracy dropped significantly when tested on side-placed cameras or too-close cameras especially in busy times (e.g. departing time at departure gate) as seen in Figure 6. That number of people will immediately create high numbers of occlusions; hence cameras cannot detect people independently. Therefore, detection accuracy can be maintained when the camera is placed far enough away with good angle (bird-eye) to reduce occlusion.

---Insert Figure 6 here---

Automatic crowd counting

The performance was evaluated on two cameras by validating the number of people counted by the system compared to ground truth. Several different algorithms were used for comparison. For detecting people, three pedestrian detector models (YoloV3, SSD300, and MobileV2 SSD) are employed. For tracking people, three object tracker models (KCF, SORT, and Deep SORT) are employed. The evaluation was tested on 30-minute video feed from each camera in a standard busy hour. For ground truth data, the number of people is counted by manual observation. Table VI presents the accuracy on two cameras using four combinations of algorithms. Figure 7 shows some samples of crowd counting in several cameras.

---Insert Table VI here---

---Insert Figure 7 here---

When tested on CPU, YoloV3 + DeepSort has the best accuracy but requires the most expensive computation (nearly 60 min computation time for processing all frames). YoloV3 + Sort has good accuracy but cannot have real-time computation as it requires less time (nearly 55 min) to process all frames. SSD300 + KCF has fastest computation time (up to 10 min) but has the worst accuracy (Figure 8).

---Insert Figure 8 here---

When tested on Google Coral Dev Board, MobileV2 + DeepSort can detect and track people on all frames achieving real-time computation. It was tested on videos with dimensions of 1,920 x 1,080 pixels and 12 fps. Pre-processing is performed to reduce unnecessary computation. Video is downsized to about half size where people objects are still identifiable and detection accuracy is not affected. As a result, the device can achieve 9 fps, enabling real-time computation on edge devices. In addition, the model is initially converted to TPU model so inference can work much faster compared to CPU computation.

Automatic social distancing breach detection

Performance was evaluated in busy areas by validating social distancing breach detection by the system compared to ground truth. Three cameras were used for evaluation: check-in area, food court area, and waiting area where people may gather in large numbers. Since it is difficult to annotate ground truth computationally, manual observation was performed by humans through video feeds to ensure if social distancing breaches occurred. Two people were employed to carefully observe and compare the actual video and analysis results to determine if people marked as red were breaching safe distances alongside the heat maps and automatic notifications generated. Figure 9 shows samples of social distancing breach detection including heat maps at a particular time on check-in and food courts cameras, respectively. Figure 10 shows how people detection model detects walking and sitting people on the departure waiting area camera. It confirms the importance of continuous learning capability in the proposed system to periodically re-adapt the AI models.

---Insert Figures 9 and 10 here---

Checklist for meeting the criteria for Responsible AI

The proposed technology was developed while involving the stakeholders that can help to ensure responsibility towards building trustworthy AI.

Reliable: Federated learning supports continuous learning aiming to improve AI reliability by learning from periodically collected new data without compromising data privacy and model accuracy.

Non-biased: Through federated learning, data bias can be gradually reduced by collecting more data for under-sample classes while respecting data privacy. Human involvement (e.g. experts in observing surveillance) can reduce AI bias by periodically evaluating the system's output. Automatic notification heat maps also can help the airport to evaluate bias e.g. manual in-field observation/checking to double check. Continuous benchmarking in different datasets can help understand and measure fairness (e.g. evaluate the model using many different types of camera, different datasets). Any previously uncaptured bias/unintentional bias can be discussed through regular meetings so airport staff can understand and be aware of the biases. Involving third parties can help to determine if the algorithm works as expected (e.g. publishing peer-reviewed papers, presenting at scientific conferences).

Data privacy protection: Through federated learning, privacy and use of personal data are managed through all data activities. From collecting, processing, storing activities, all sensitive data and personal data are kept secure in local devices where the data are obtained. For moving data, only non-sensitive data are sent to the central data, not actual images. Data compression and aggregation can be applied to ensure data are more secure. Local training requires shorter training time and limited movement of sensitive data. Hence, it shortens the duration of data retention as images are immediately erased after training and minimizes the risks of data treats.

Transparent AI decisions: Human-in-the-loop approach (i.e. integrating airport staff and the AI system to work together in making decisions) will allow humans to double check results or make decisions. For example, auto notification can ask airport personnel to validate how the system works. Working closely with airport management through regular meetings helps build transparency in AI models.

Sustainable and scalable for long-term use: The incorporation of AI continual learning capability enables system to scale up in the future such as adding more cameras with verified camera deployment characteristics. Furthermore, the system is not limited to monitoring social distancing during the COVID-19 pandemic, but can be adjusted for other purposes such as to observe people's social interactivity at home or in the workplace.

Table VII to Table IX demonstrate how trust is established and maintained throughout the process by following the checklists provided in the framework.

---Insert Tables VII, VIII, and IX here---

5. Conclusion and Future Work

This paper proposes on-edge privacy-preserving AI-enabled surveillance technology to detect social distancing breaches in public spaces to respect people's privacy. Comprehensive evaluation in the context of a case study at an airport demonstrates how the proposed system can fully address the case study's requirements in terms of its reliability, usefulness when deployed to the airport's cameras, and compliance with responsible AI. The results confirm that limiting sensitive data movement via on-edge computing is effective to protect data privacy without compromising the performance. This paper also contributes to the study on addressing data privacy issue in AI-enabled surveillance applications. Existing work on automated social distancing monitoring has not fully considered the need for addressing privacy-preserving AI. While the proposed system has shown the benefits for monitoring social distancing during pandemics, it can be adjusted for other purposes in the future, such as observing people's social interactivity at home or in workplaces that respects people's privacy. The proposed system and its design approach can be also useful for informing future work in developing AI-enabled video surveillance applications, in which data privacy can be protected using the federated learning approach, while ensuring its compliance with responsible AI principles.

References

Al-Suwaidi, G. B. & Zemerly, M. J. Locating friends and family using mobile phones with global positioning system (GPS). 2009 IEEE/ACS International Conference on Computer Systems and Applications, 2009. IEEE, 555-558.

Australian Government Office of the Australian Information Commissioner. 2019. *Security cameras* [Online]. Australian Government Office of the Australian Information Commissioner. Available: https://www.oaic.gov.au/privacy/your-privacy-rights/surveillance-and-monitoring/security-cameras/ [Accessed].

Benson, V., Saridakis, G. & Tennakoon, H. 2015. Information disclosure of social media users. *Information Technology & People*.

Bewley, A., Ge, Z., Ott, L., Ramos, F. & Upcroft, B. Simple online and realtime tracking. Image Processing (ICIP), 2016 IEEE International Conference on, 2016. IEEE, 3464-3468.

Biswas, D., Su, H., Wang, C., Stevanovic, A. & Wang, W. 2019. An automatic traffic density estimation using Single Shot Detection (SSD) and MobileNet-SSD. *Physics and Chemistry of the Earth, Parts A/B/C,* 110**,** 176-184.

Boland, B., De Smet, A., Palter, R. & Sanghvi, A. 2020. Reimagining the office and work life after COVID-19.

Burt, A. 2019. The AI transparency paradox. *Harv. Bus. Rev.(Dec. 13, 2019), https://bit. ly/369LKvq*.

Clifton, J., Glasmeier, A. & Gray, M. 2020. When machines think for us: the consequences for work and place. Oxford University Press UK.

Daly, N. 2020. *A 'pandemic drone' and other technology could help limit the spread of coronavirus and ease restrictions sooner, but at what cost?* [Online]. ABC News. Available: https://www.abc.net.au/news/2020-05-01/new-surveillance-technology-could-beat-coronavirus-but-at-a-cost/12201552 [Accessed 26 Oct 2020 2020].

Department of Health, A. G. 2020. *Physical distancing for coronavirus (COVID-19)* [Online]. Australia Government Department of Health. Available: https://www.health.gov.au/news/health-alerts/novel-coronavirus-2019-ncov-health-alert/how-to-protect-yourself-and-others-from-coronavirus-covid-19/physical-distancing-for-coronavirus-covid-19 [Accessed].

Department of Justice, T. U. S. 2015. Privacy Act of 1974. Department of Justice of The United States
Domingo-Ferrer, J. & Blanco-Justicia, A. 2020. Privacy-Preserving Technologies. *The Ethics of Cybersecurity.* Springer, Cham.
Elloumi, W., Latoui, A., Canals, R., Chetouani, A. & Treuillet, S. 2016. Indoor pedestrian localization with a smartphone: A comparison of inertial and vision-based methods. *IEEE Sensors Journal,* 16**,** 5376-5388.
Feldstein, S. 2019. *The global expansion of AI surveillance*, Carnegie Endowment for International Peace Washington, DC.
General Data Protection Regulation. 2016. *REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016* [Online]. Official Journal of the European Union. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 [Accessed].
Ghallab, M. 2019. Responsible AI: requirements and challenges. *AI Perspectives,* 1**,** 3.
Gunjal, P. R., Gunjal, B. R., Shinde, H. A., Vanam, S. M. & Aher, S. S. Moving Object Tracking Using Kalman Filter. 2018 International Conference On Advances in Communication and Computing Technology (ICACCT), 2018. IEEE, 544-547.
Held, D., Thrun, S. & Savarese, S. Learning to track at 100 fps with deep regression networks. European Conference on Computer Vision, 2016. Springer, 749-765.
Hu, H., Gu, J., Zhang, Z., Dai, J. & Wei, Y. Relation networks for object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018. 3588-3597.
Iswanto, I. A. & Li, B. 2017. Visual object tracking based on mean-shift and particle-Kalman filter. *Procedia computer science,* 116**,** 587-595.
Johnson Jr, J., Hasan, S., Lee, D., Hluchan, C. & Ahmed, N. 2020. Social-Distancing Monitoring Using Portable Electronic Devices.
Jovicic, A. & Engineer, P. Qualcomm ® Lumicast TM : A high accuracy indoor positioning system based on visible light communication April 2016. 2016.
Kavalionak, H., Gennaro, C., Amato, G., Vairo, C., Perciante, C., Meghini, C. & Falchi, F. 2019. Distributed video surveillance using smart cameras. *Journal of Grid Computing,* 17**,** 59-77.
Kennedys Law. 2018. *Summary of Data Privacy Laws in Asia Pacific* [Online]. Available: https://kennedyslaw.com/media/3154/asia-pacific-data-principles.pdf [Accessed 26 Oct 2020 2020].
Kok, M., Hol, J. D. & Schön, T. B. 2015. Indoor positioning using ultrawideband and inertial measurements. *IEEE Transactions on Vehicular Technology,* 64**,** 1293-1303.
Konečný, J., Mcmahan, H. B., Ramage, D. & Richtárik, P. 2016. Federated optimization: Distributed machine learning for on-device intelligence. *arXiv preprint arXiv:1610.02527*.
Laoudias, C., Moreira, A., Kim, S., Lee, S., Wirola, L. & Fischione, C. 2018. A survey of enabling technologies for network localization, tracking, and navigation. *IEEE Communications Surveys & Tutorials,* 20**,** 3607-3644.
Ledergerber, A. & D'andrea, R. 2020. A Multi-Static Radar Network with Ultra-Wideband Radio-Equipped Devices. *Sensors,* 20**,** 1599.
Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. & Berg, A. C. Ssd: Single shot multibox detector. European conference on computer vision, 2016. Springer, 21-37.
Liyanaarachchi, G. 2020. Online privacy as an integral component of strategy: allaying customer fears and building loyalty. *Journal of Business Strategy*.
Mazuelas, S., Bahillo, A., Lorenzo, R. M., Fernandez, P., Lago, F. A., Garcia, E., Blas, J. & Abril, E. J. 2009. Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks. *IEEE Journal of selected topics in signal processing,* 3**,** 821-831.
Nadikattu, R. R., Mohammad, S. M. & Whig, D. 2020. Novel Economical Social Distancing Smart Device for COVID19. *International Journal of Electrical Engineering and Technology,* 11.
Niu, J., Wang, B., Shu, L., Duong, T. Q. & Chen, Y. 2015. ZIL: An energy-efficient indoor localization system using ZigBee radio to detect WiFi fingerprints. *IEEE Journal on Selected Areas in Communications,* 33**,** 1431-1442.
Office for Personal Data Protection 2005. Personal Data Protection Act. *In:* REGION, O. F. P. D. P. O. T. M. S. A. (ed.). Office for Personal Data Protection.
Office of the Australian Information Commissioner 2019. Australian Privacy Principle Guidelines. *In:* COMMISSIONER, O. O. T. A. I. (ed.). Office of the Australian Information Commissioner.
Parisi, G. I., Kemker, R., Part, J. L., Kanan, C. & Wermter, S. 2019. Continual lifelong learning with neural networks: A review. *Neural Networks,* 113**,** 54-71.
Personal Information Protection Commission 2011. Personal Information Protection Act. Ministy of Government Legislation of Korea.
Punn, N. S., Sonbhadra, S. K. & Agarwal, S. 2020. Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques. *arXiv preprint arXiv:2005.01385*.
Raghunathan, A., Steinhardt, J. & Liang, P. 2018. Certified defenses against adversarial examples. *arXiv preprint arXiv:1801.09344*.
Ransbotham, S. 2017. *Justifying Human Involvement in the AI Decision-Making Loop* [Online]. Available: https://sloanreview.mit.edu/article/justifying-human-involvement-in-the-ai-decision-making-loop/ [Accessed].
Redmon, J. & Farhadi, A. 2018. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
Ren, S., He, K., Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 2015. 91-99.
Romm, T., Harwell, D., Dwoskin, E. & Timberg, C. 2020. *Apple, Google debut major effort to help people track if they've come in contact with coronavirus* [Online]. The Washington Post. [Accessed 27 July 2020 2020].
Rosenbrock, H. H. 2012. *Designing human-centred technology: a cross-disciplinary project in computer-aided manufacturing*, Springer Science & Business Media.
Rossi, F. 2018. Building trust in artificial intelligence. *Journal of international affairs,* 72**,** 127-134.
Sanchez-Iborra, R. & Skarmeta, A. F. 2020. TinyML-Enabled Frugal Smart Objects: Challenges and Opportunities. *IEEE Circuits and Systems Magazine,* 20**,** 4-18.
Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. & Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018. 4510-4520.
Sugianto, N., Tjondronegoro, D., Sorwar, G., Chakraborty, P. & Yuwono, E. I. 2019. Continuous Learning without Forgetting for Person Re-Identification. *In:* ABADI, M. (ed.) *16th IEEE International Conference on Advanced Video and Signal-based Surveillance.* Taipei, Taiwan: IEEE.
Sultan, S. & Jensen, C. D. Privacy-preserving measures in smart city video surveillance systems. 6th International Conference on Information Systems Security and Privacy, 2020. SciTePress, 506-514.
Szajewska, A. 2017. Development of the Thermal Imaging Camera (TIC) Technology. *Procedia Engineering,* 172**,** 1067-1072.
Tiemann, J. & Wietfeld, C. Scalable and precise multi-UAV indoor navigation using TDOA-based UWB localization. 2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2017. IEEE, 1-7.
Wang, R. J., Li, X. & Ling, C. X. Pelee: A real-time object detection system on mobile devices. Advances in Neural Information Processing Systems, 2018. 1963-1972.

World Health Organization. 2020. *Coronavirus disease (COVID-19) pandemic* [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [Accessed 19 Oct 2020].

Wojke, N., Bewley, A. & Paulus, D. Simple online and realtime tracking with a deep association metric. Image Processing (ICIP), 2017 IEEE International Conference on, 2017. IEEE, 3645-3649.

Yang, Q., Liu, Y., Chen, T. & Tong, Y. 2019a. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST),* 10**,** 12.

Yang, T., Cappelle, C., Ruichek, Y. & El Bagdouri, M. 2019b. Online multi-object tracking combining optical flow and compressive tracking in Markov decision process. *Journal of Visual Communication and Image Representation,* 58**,** 178-186.

Yun, S. & Kim, S. Recurrent YOLO and LSTM-based IR single pedestrian tracking. 2019 19th International Conference on Control, Automation and Systems (ICCAS), 2019. IEEE, 94-96.

Zhou, X., Gong, W., Fu, W. & Du, F. Application of deep learning in object detection. 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), 2017. IEEE, 631-634.

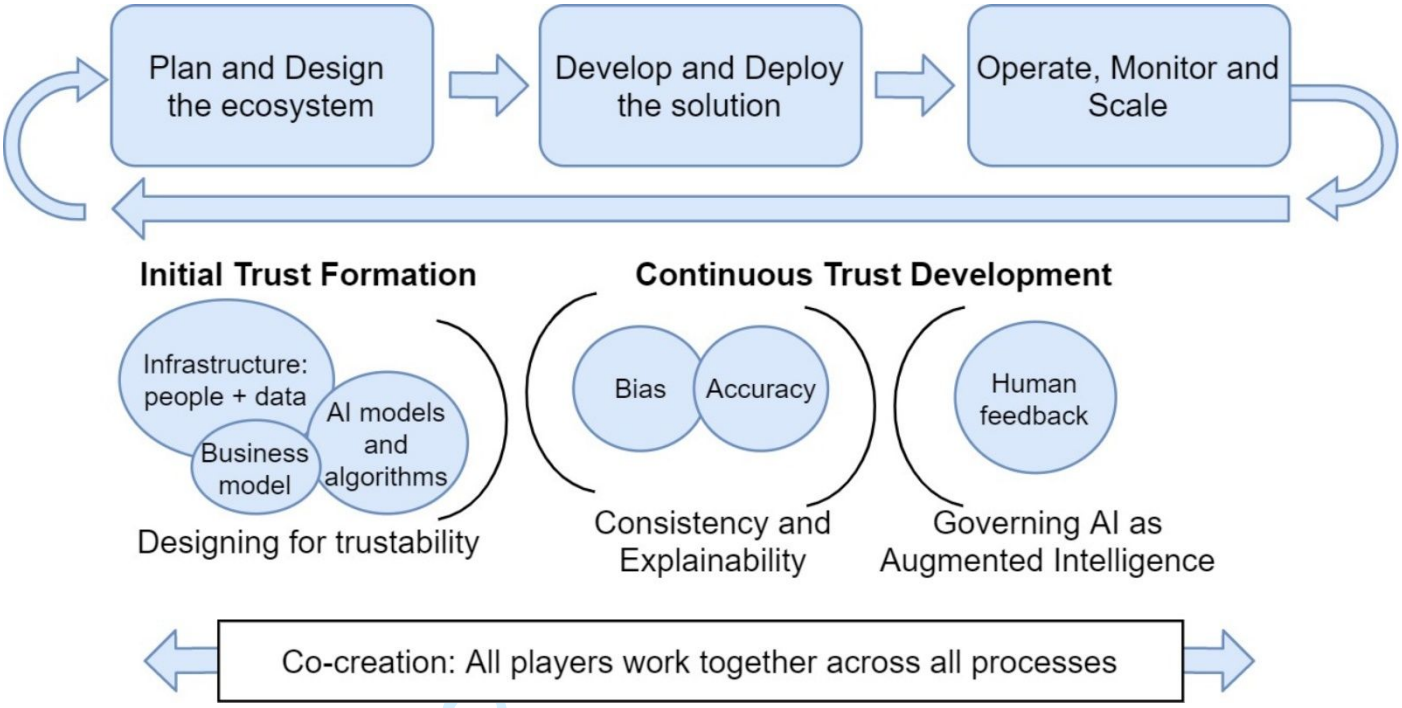

Figure 1. Responsible AI Implementation Framework (RAIIF). The framework is iterative across all processes from planning and designing to development and scaling-up processes.

Figure 2. The proposed system for detecting social distancing breach in public space. Left: People detected on cameras. Right: People detected are represented as dots in a 2D projection view to make them feel less observed. Green identifies a person is at a safe distance whereas red identifies a person at an unsafe distance.

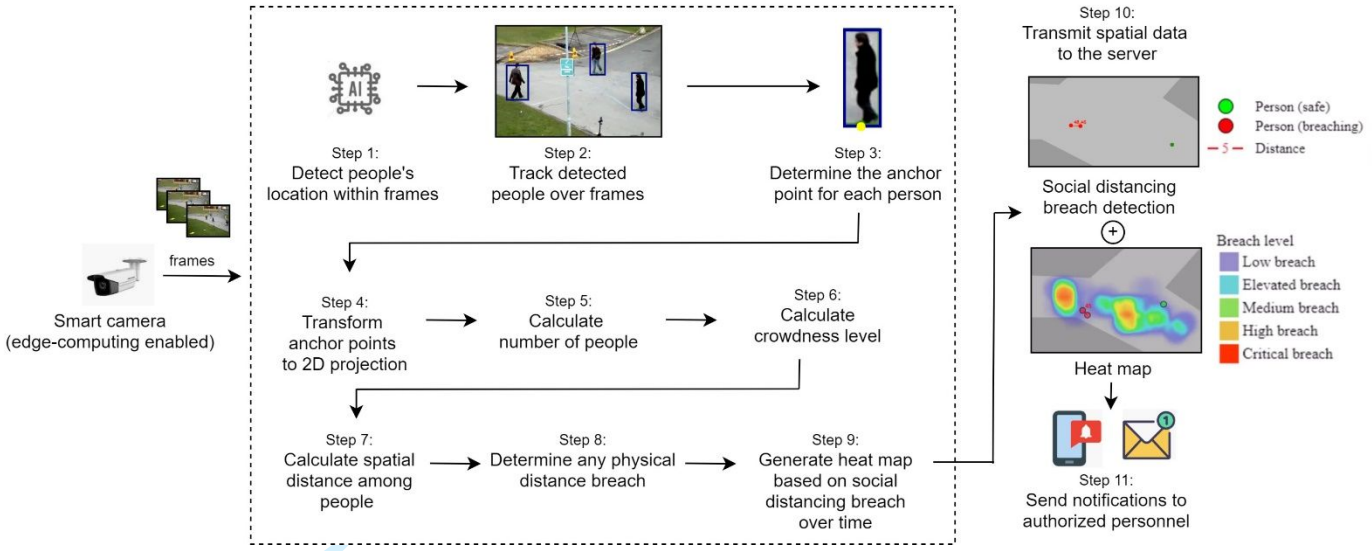

Figure 3. The process of detecting social distancing breaches within each smart camera.

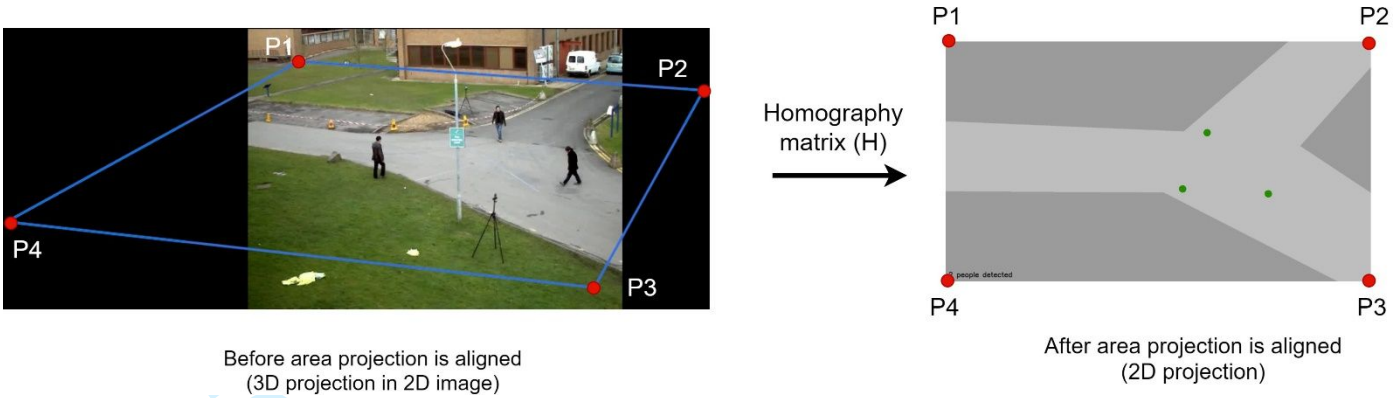

Figure 4. Area projection alignment from 3D to 2D projection using the Homography (H) algorithm. During the initial camera setup, corner points (red dots) are defined to calculate the matrix H.

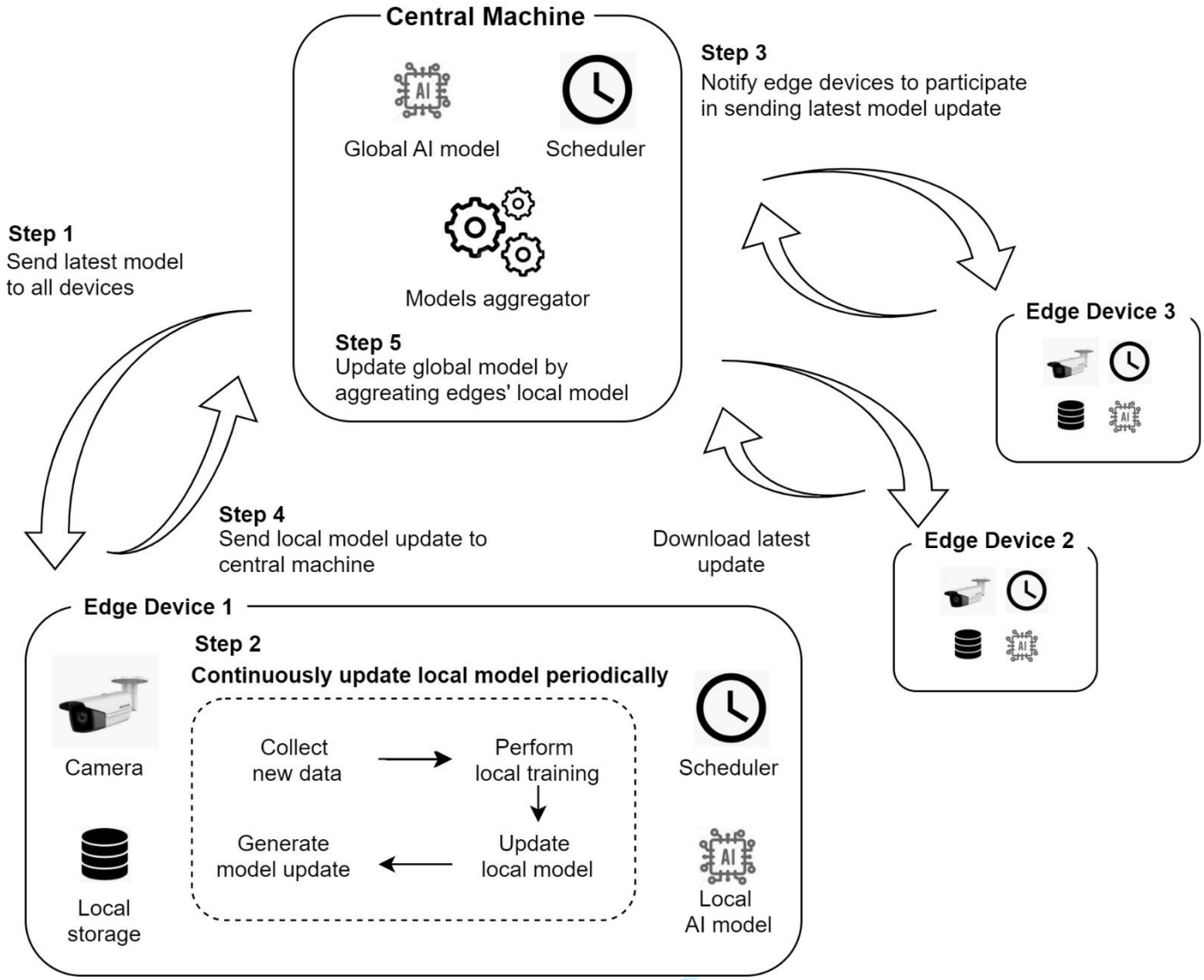

Figure 5. On-edge computing to preserve data privacy

Figure 6. With side-place cameras or too-close cameras, people detection accuracy can be maintained when few people are present at one time (left) but will drop significantly with too many people as occlusion happens (right).

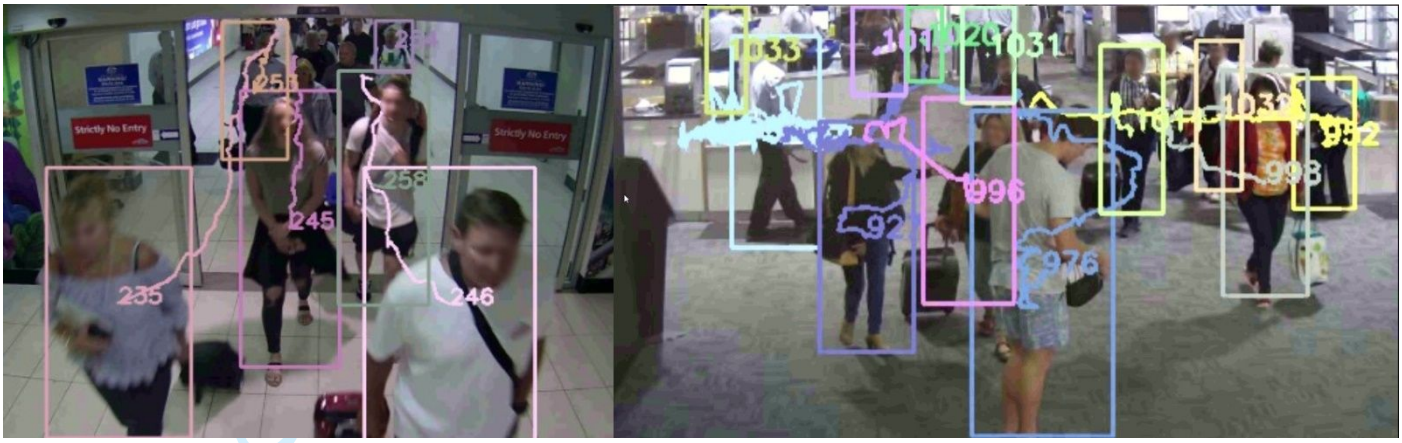

Figure 7. Samples of people detection and tracking by airport cameras. Left: arrival gate camera. Right: after-security check camera

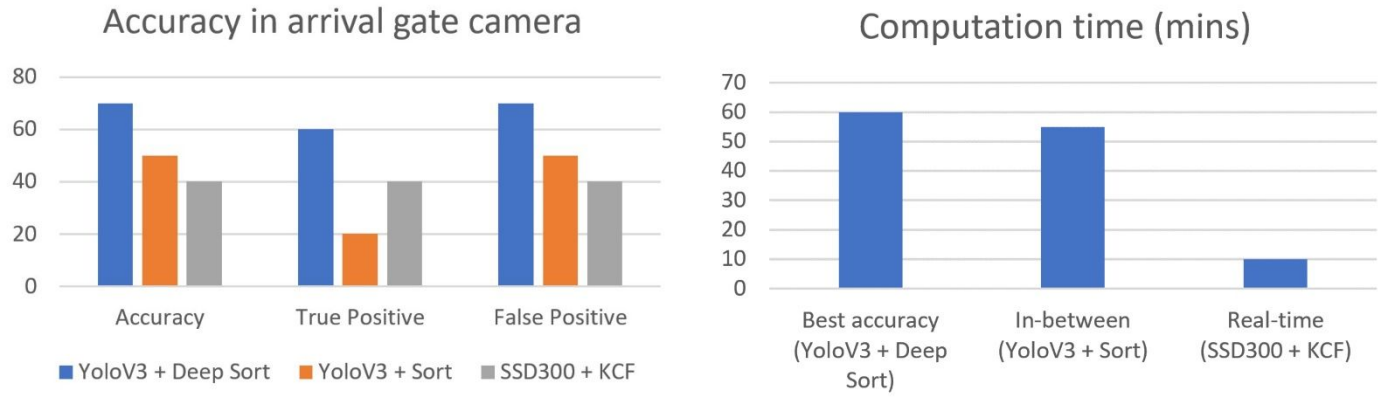

Figure 8. Accuracy (left) and computation time (right) comparison of people counting in arrival gate camera.

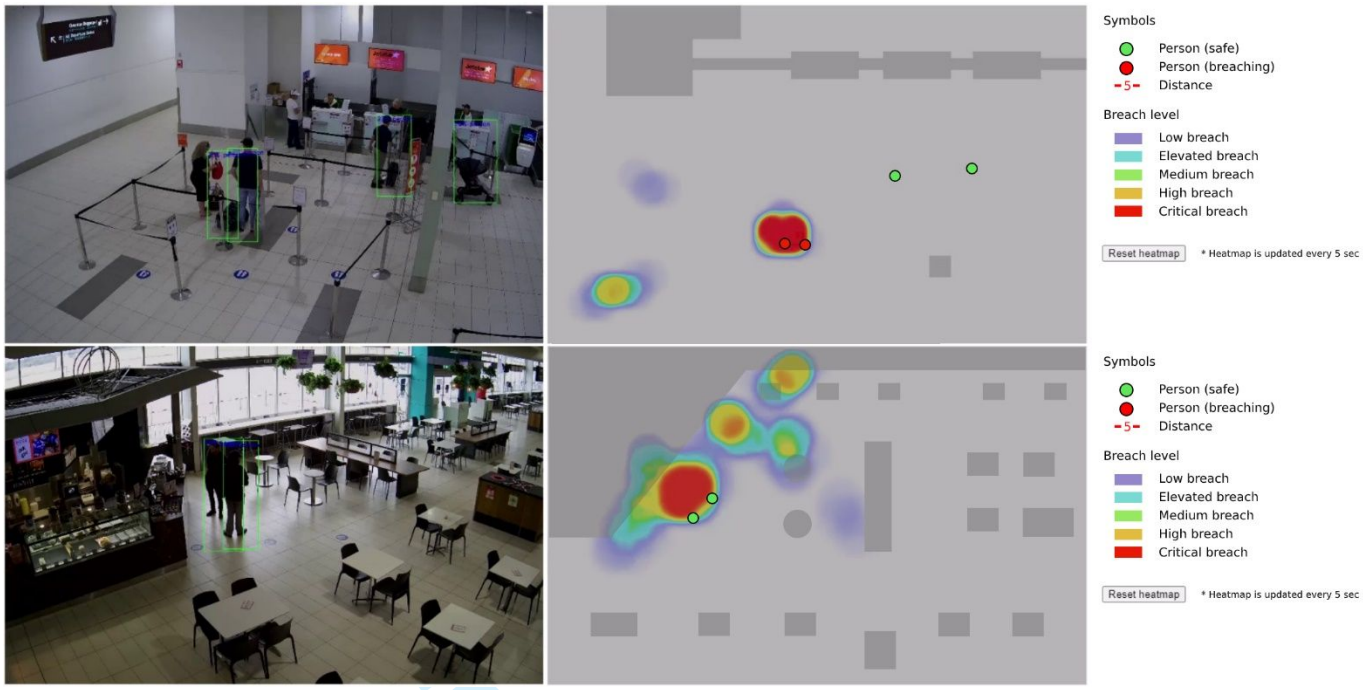

Figure 9. Sample of social distancing breach detection by two cameras. Top: check-in area camera. Bottom: café at departure camera. Images are captured at a particular time and the heat map shows cumulative breaches on that day.

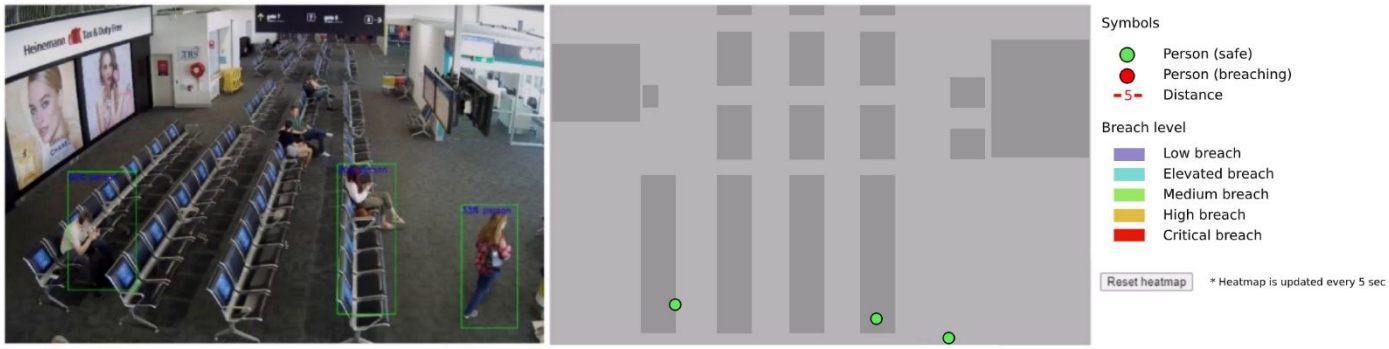

Figure 10. Sample of social distancing breach detection by the departure waiting area camera. Images are captured at a particular time and the heat map shows cumulative breaches on that day.

Table I. Evolution of AI solutions adopted at the local international airport.

Business requirement	**Proposed AI solution**	**Benefit**	**Consideration**
People are detected accurately by different cameras and deployments	Automatic people detection and tracking	The airport can observe passenger flow (e.g. queuing duration)	Accuracy of people detection and tracking
Crowd counting is accurate for different cameras and deployments	Automatic crowd counting	The airport can monitor the level of crowdedness in each area to maintain safe distancing.	Accuracy of crowd counting
Social distancing breach is detected accurately based on government regulations during COVID-19 pandemic (e.g. < 1.5 m)	Automatic social distancing breach detection	The airport can enforce social distancing responsibility and rethink/redesign the space layout to reduce distancing breaches	Accuracy of the social distancing breach detection and spatial analysis

Table II. Checklist for Plan and Design stages

Consideration	Checklist
Technological accessibility	• The organization's existing internal infrastructure can support the AI solution, including people and data. • The organization commits to transform its operational processes to align with the AI solution and its required technologies, including data collection and management.
AI-enabled business refinement	• The initial solution can confirm the AI's feasibility and ability to serve a purpose, such as solving an identified problem or supporting business processes. • The organization's decision-makers understand the business case of the AI solution, in terms of its potential for delivering a return of investment, such as cost-saving, enhanced decisions, and improved business processes' efficiency. • Decision-makers understand the potential impacts on the existing human and technology infrastructure, and risks to the organization's reputation.
Customer satisfaction	• The AI solution brings direct benefits to end-users and improves customer experience. • Users can trust the AI solution and the organization that manages it due to clear manuals, documentations, and transparent risk assessments. • Users can use the AI-enabled services, including the time for learning (how to use), and buy the required equipment and fees to access it.
Operational compliance	• AI solution is compliant with regulations and policies and operates lawfully. • The impacts of operating and maintaining the AI solution can be managed by and compatible with the existing human and technology infrastructure. • AI solution aims to augment human abilities to support actions and decisions while ensuring the end-users' safety and wellbeing. • AI solution is equitable and particularly accessible for those who need the most. • AI solutions can be potentially used on a global scale without a significant shift in governance.

Table III. Checklist for Development and Deployment stages

Consideration	Checklist
Reliable operations	• AI decisions and reasoning can be ascertained and trusted to support operation, and bias can be managed and reduced. • AI solution will function as intended and cannot be hacked and manipulated.
Empowered employees	• AI solution can improve a human's ability to perform operational tasks and make decisions. • AI-based decisions are consistent, reliable, fair, and aligned with social and cultural justice values in the local and global contexts. • If applicable, the change management for transforming the workforce always ensures fair employment and labor practices.
Engaged customers	• AI solution consistently meets performance expectations and increasingly more reliable to support or reduce manual work, make better decisions, or make services more enjoyable and personalized.
Transformed products	• AI solutions can transform the existing norms via the new product(s) that people can enjoy and benefit from. • AI solution (potentially) adheres with the existing ethical- and legal- theories or general principles.

Table IV. Checklist for Operate, Monitor, and Scale-Up stages

Consideration	Checklist
Continuous refinement	• AI solution can adapt and refine its model, improving its accuracy based on human feedback. • AI solution can be continuously tested by research and development (R&D) and quality assurance process to ensure that the outcomes continue to meet the required tasks and performance expectations. • AI solutions can continuously adapt to new operations and socio-economic requirements.
Responsible business practice	• AI solution can be managed by an end-to-end enterprise governance framework to be consistently accountable and transparent. • AI's risks and scope of controls align with the organization's current and newly operationalized ethics.
Individual (user) trust	• Human trust cannot be broken by being disadvantaged, such as being subjected to automated decisions, especially when such decisions have legal ramifications. • Users have the right to access all information needed to fully understand the AI's product and the test results before adoption. • The policy and standard of practice are documented and transparent to ensure that users can trust the use of data and AI.
Societal (community) trust	• An appropriate level of human supervision is maintained to ensure that the AI solution continually aligns with human rights, social norms, and privacy regulations. • Individual and collective AI use is transparent and auditable to ensure compliance with laws and fundamental rights and freedom.

Table V. Performance comparison for object tracking algorithms evaluated on MOT16 challenge.

Algorithm	MOTA ↑	MOTP ↑	IDF1 ↑	MT ↑	ML ↓	ID ↓	FM ↓	FP ↓	FN ↓	Runtime ↑
KCF (Chu et al., 2019)	48.8	-	47.2	-	-	906	-	5875	86567	-
SORT (Bewley et al., 2016)	59.8	79.6	-	25.4%	22.7%	1423	1835	8698	63245	60Hz
Deep Sort (Wojke et al., 2017)	**61.4**	79.1	62.2	32.8%	18.2%	781	2008	12852	56668	40Hz

MOTA = Multi-object tracking accuracy, MOTP = Multi-object tracking precision, IDF1 = ID F1 score, MT = Mostly tracked object, ML = Mostly lost object, ID = Identity switches, FM = Fragmentation. ↑ indicates the higher it is, the better it is whereas ↓ indicates the lower it is, the better it is.

Table VI. People detection and tracking performance in the context of counting people in and out. Tested on 30-min videos on both cameras.

Camera	YoloV3 + DeepSORT	YoloV3 + SORT	SSD300 + KCF	SSD MobileV2 + DeepSORT
Camera 1 (Arrival gate)	70.55%	61.11%	53.91%	58.28%
Camera 2 (After security check)	85.02%	80.05%	71.40%	74.53%

Table VII. List of actions taken to establish and maintain trust in the Plan and Design stages

Checklist for Plan and Design	**Actions/Strategies taken**
Technological accessibility	• Ensured existing IT infrastructure supports the proposed system, e.g. number of cameras installed is sufficient to cover targeted observation area, number of personnel to observe, hardware to process computation • Required the airport to commit to providing cameras' access to develop the proof of concept of the proposed AI solution
AI-enabled business refinement	• Provided a proof of concept to show the benefits of the proposed AI solution • Collaborated with the airport management through regular progress meetings and identify and disclose any impacts and potential risks of the proposed system
Customer satisfaction	• Analyzed existing problems at the airport and identified the business requirements to ensure the proposed system meets the airport's expectations • Assessed and disclosed any potential risks found • Continuously provided any findings to the airport during the development process
Operational compliance	• Worked with the airport's IT and surveillance teams to understand the airport's regulations and policies when accessing the cameras and dealing with the data • Prior to data collection, obtained human ethics clearance from both parties to ensure the proposed AI solution is compliant with existing regulations and policies and operates lawfully

Table VIII. List of actions taken to establish and maintain trust in the Develop and Deploy the solution stages

Checklist for Develop and Deployment	**Actions/Strategies taken**
Reliable operations	• Conducted all experiments and testing in the internal airport's machine • Evaluated the AI models' performance through rigorous benchmarking with state-of-the-art-model, compared it with existing systems and manual observation prior to deployment • Deployed the system within a secure network hosted by the airport and only performed computations in the airport's network. No data transmission outside the airport's network
Empowered employees	• Obtained continuous feedback from the airport's personnel to improve the proposed system's usefulness
Engaged customers	• Tailored the AI solution to the airport's context and requirements. Further fine-tuning model and continuous training were conducted to meet the expected performance • Provided continuous evidence of the benefits of the AI solution. This can be seen from the evolution of AI adoption from crowd counting to social distancing breach detection
Transformed products	• Demonstrated the proposed AI solution to the airport management to introduce a new approach to ensuring passengers' safety and comfort at the airport through surveillance while respecting people's privacy

Table IX. List of actions taken to establish and maintain trust in the Operate, Monitor, and Scale-up stages

Checklist for Operate, Monitor and Scale-up	**Actions/Strategies taken**
Continuous refinement	• Enforced continuous development of AI through rigorous evaluation and benchmarking with the-state-of-art AI model • Evaluated the AI performance with existing systems to ensure accuracy in the decision-making process • Obtained continuous feedback from the airport's personnel to improve the system's accuracy • Published the work in internationally peer-reviewed journals and presented the demo at international events to justify and evaluate the solution, and obtain experts' feedback
Responsible business practice	• Followed the airport's regulation and policies during data collection and experiments • Airport applied strict monitoring to ensure any activities during system development did not breach the agreement
Individual/user trust	• Applied AI solution aims to support the airport's personnel in making well-informed decisions. The AI solution does not make automatic decisions and replace existing systems • Introduced the AI solution to reduce human involvement. Human is still required to make final decisions (automatic notifications will be sent to personnel when prompt responses are needed) • Conducted periodic evaluation through in-person meetings to discuss existing issues • Documented any agreements between both parties and presented regular updates and results to promote transparency • Applied strict data retention policy by keeping video feeds for up to 7 days • Only collected data related to the objectives
Societal (community) trust	• Conducted the research in full accordance with human research ethics committee and legal department's requirements • Ensured that all camera access is approved by the airport's surveillance department prior to access. Any changes were reported and documented • Informed the public by placing signage stating that video observation was in place to maintain passenger awareness