Aerial Object Tracking System on Micro Quadrotor Drone for Crowd Detection in Small-Scale Area

Musthafa Dimas Bagaskoro Pranoto[1], Muhammad Ikhsan Sani[1], Marlindia Ike Sari[1]

[1]Computer Technology, School of Applied Science, Telkom University, Bandung, Indonesia
dimaspranoto@student.telkomuniversity.ac.id

Abstract. Remembering when COVID-19 had such a huge impact in Indonesia. For pandemic prevention andcontrol, it is important to use all available technologies for mitigation. There are reports of attempts to use drone technology in various settings. Drones can help control pandemics in a variety of ways. Recent researchhas revealed the benefits of drones, especially the type of quadrotor. The development of the quadrotor dronesystem as a flying platform for air surveillance has received a lot of attention. Drone-based surveillance uses cameras to capture images and videos to obtain information about a particular geographical area or topography.The reason for the analysis of air images is the need to predict changes in the areas of interest, such as to countthe crowd, remembering from the lessons we can learn in the efforts to prevent the COVID-19 pandemic of that time. The main challenge is to figure out how to create a system that can provide accurate counting of thecrowds of the air, which is important for fighting pandemics. The aim of this paper is to show the application of drones from the air object tracking system for crowd detection in a small-scale area. The report also proposes alternative solutions based on the use of microquadrotor drones equipped with cameras on board. In addition,images are transmitted and processed in mobile applications to provide more information about the object. Several experiments show that the system has been effectively installed and provides data collection for furtherresearch. The results showed that the system could be used for small-scale air monitoring.

Keywords: COVID-19, pandemic, drone, aerial, quadrotor, object tracking, surveillance, monitoring, crowdcounting, small-scale.

1 Introduction

The In recent years, there has been a growing interest in the development of efficient and reliable surveillance systems capable of monitoring crowded areas, particularly in the context of public health crises such as the COVID- 19 pandemic (Fadzil et al., 2021). The COVID-19 pandemic has underscored the need for innovative approaches to crowd management and monitoring. Traditional methods of crowd control, such as physical barriers and manual surveillance, often prove insufficient in densely populated areas, making it challenging to implement and enforce social distancing protocols effectively (Al-

D. V. Ferezagia et al. (eds.), *Proceedings of the International Conference on Vocational Education Applied Science and Technology (ICVEAST 2023)*, Advances in Social Science, Education and Humanities Research 783,
https://doi.org/10.2991/978-2-38476-132-6_84

Sa'd et al., 2022). Furthermore, the contagious nature of the virus necessitates prompt identification and response to potential outbreaks, requiring systems capable of detecting and tracking individuals in real-time (Yadav et al., 2022).

The proposed Aerial Object Tracking System leverages the advantages of micro quadrotor drones, which are small, lightweight, and agile aerial platforms equipped with high-resolution cameras and sensors. These drones can be deployed in small-scale areas, such as parks, markets, or public transportation hubs, to capture comprehensive visualdata from an elevated perspective (Wu et al., 2017; C. Zhang et al., 2022). By employing advanced computer vision algorithms and machine learning techniques, the system can identify and track individuals, detecting instances of overcrowding or violations of social distancing guidelines (Wang et al., 2022). One key advantage of using micro quadrotor drones for crowd detection is their ability to navigate complex environments and rapidly respond to changing situations (Castellano et al., 2023). These drones can autonomously adapt their flight paths, providing dynamic coverage of different areas of interest. Moreover, the integration of real-time video streaming and analysis allows for immediate decision-making and response by authorities responsible for crowd control, enabling timely interventions and ensuring the safety and well-being of the public (Alzahrani et al., 2022).

By focusing on small-scale areas, the Aerial Object Tracking System offers a localized approach to crowd detection, enabling targeted interventions to mitigate the spread of COVID-19 (C. Zhang et al., 2022). It provides valuable data for authorities to identify hotspots, monitor compliance with health guidelines, and optimize resource allocation. Additionally, the system can support contact tracing efforts by providing visual records of individual movement patterns, aiding in identifying potential transmission chains and reducing the risk of outbreaks.

2 Literature review

Extensive research has been conducted in the field of computer vision pertaining to crowd counting and crowd density estimation. However, the current emphasis lies in the domain of density estimation. In early studies, person orhead detectors were commonly employed using a sliding window approach to the image. Nevertheless, despite the adoption of cutting-edge object detectors like YOLO (Lan et al., 2018; Molchanov et al., 2017), these methods still yield suboptimal outcomes when confronted with the detection of small objects within highly congested crowds. To address this challenge, regression-based techniques have been introduced, enabling direct learning of the mapping from an image to the overall count of people (Fadzil et al., 2021). Although these methods alleviate the reliance on precise individual positioning within the crowd, a task that is inherently complex, they fail to leverage spatial information, which holds significant predictive value. To circumvent the difficulty of accurately detecting and localizing individuals in the scene while leveraging spatial information, the recent trend involves the acquisition of density maps, thereby directly integrating spatial information into the learning process (Fan et al., 2022). Promising solutions have emerged in this regard. Some methodologies first operate at the patch level and subsequently fuse localfeatures (Zhu et al., 2020). Others incorporate attention mechanisms (G. Zhang et al., 2020), adopt cascade approachesto simultaneously learn people counting and density maps (Sindagi & Patel, 2017). enhance performance through knowledge distillation (Jiang, Lin, & Jane Wang, 2021), or develop frameworks that enable concurrent crowd counting and localization (Jiang, Lin, & Wang, 2021). These successful approaches demonstrate the potential of incorporatingspatial information and employing diverse techniques to advance crowd counting and density estimation tasks (Fadzilet al., 2021).

Despite their effectiveness, the computational demands and stringent requirements imposed by Unmanned Aerial Vehicles (UAVs), such as limited battery capacity and real-time response constraints, pose challenges for the adoptionof existing approaches. The fine-tuning of deep neural architectures to achieve an optimal balance between precisionand performance is an active area of research. To stimulate progress in this domain, the VisDrone Crowd Counting challenge was introduced (Chen et al., 2021). However, the solutions presented by participants in the challenge oftenprioritize effectiveness over efficiency, aiming primarily to minimize counting errors rather than addressing computational constraints. The solution that achieved the lowest counting error in the challenge was TransCrowd (Liang et al., 2022), which builds upon the increasingly popular Vision Transformer (Vaswani et al., 2017). Notably,the proposed method focuses solely on regressing the count of individuals, without providing density maps that couldfacilitate crowd flow detection. Moreover, it is acknowledged that transformer-based solutions are associated with computationally expensive operations, further complicating their practical implementation.

An effective strategy for mitigating these challenges involves the utilization of OpenCV models. An extensive study was conducted on the OpenCV platform and its integrated libraries in order to generate a code that correctly andreliably recognizes the crowd by using modern and powerful hardware. Consequently, OpenCV models offer a potential solution to achieving accurate models while maintaining efficient inference times. In the context of aerial drone imaging, an OpenCV model specifically designed for this purpose was introduced in(Shadakshri et al., 2022). Additionally, a similar approach was proposed in our prior research (Castellano et al., 2020). However, it is crucial tonote that both of these methods primarily focused on crowd detection tasks, specifically discriminating between crowded and uncrowded scenes. Furthermore, the density maps generated by these models were coarse in nature, as they were not trained on individual people labels.

Prior research has explored human tracking methods utilizing RGB cameras or other sensors, employing clusteringor classification models for motion tracking (Gajjar et al., 2017; Xiao et al., 2019; Yan et al., 2020). However, these approaches are typically designed for indoor environments or scenarios involving a limited number of individuals observed from frontal perspectives. In a related work, the authors of (Wen et al., 2021), who also introduced the VisDrone Crowd Counting datasets, proposed a model catering to drone-captured images, addressing density map estimation, localization, and tracking simultaneously. Notably, their model differs from ours, as it involves a complexand costly pipeline specifically designed for tracking individual trajectories. Another recent contribution from different authors involves a periodic crowd tracking method from UAVs, based on a binary linear programming model(Chebil et al., 2022). Nevertheless, this work was conducted using simulated scenarios and did not explicitly tackle the crowd detection challenge from a computer vision perspective. As far as our knowledge extends, there is a lack ofliterature on crowd flow detection in drone videos, a context presenting markedly distinct challenges compared to conventional settings. This paper endeavors to bridge this gap by focusing on tracing centroids that identify groups ofpeople, capitalizing on the spatial information learned and expressed through density maps.

3 Materials and Methods

3.1 Dataset processing

The dataset collection phase of this research involved the utilization of the Haar cascade method for face detection.The Haar cascade classifier, implemented in the OpenCV library, was employed to automatically detect and localize faces within the collected images. The Haar cascade classifier is a machine learning-based object detection techniquethat utilizes a set of trained features to identify specific patterns, such as facial features, in an image (Syed Ameer Abbas et al., 2018). Each detected face was considered as an individual instance and counted as one person within thedataset.

During the dataset collection process, emphasis was placed on accurately determining the number of individuals present in the crowd scenes. To achieve this, every face detected by the Haar cascade classifier was counted as one person. This approach ensured that each detected face was treated as an individual entity within the dataset. Additionally, in cases where multiple faces intersected with each other, they were considered as a single group. By grouping intersecting faces together, the dataset captured the collective presence of individuals in close proximity, providing valuable insights into crowd dynamics and density.

The Haar cascade-based dataset collection approach yielded a comprehensive collection of annotated images, where each image contained one or more instances of detected faces (Nadeem et al., 2022). Through this method, thedataset accurately represented the diversity and complexity of crowd scenarios in small-scale areas. The dataset servedas a crucial resource for training and evaluating the Aerial Object Tracking System for Crowd Detection, enabling thedevelopment and validation of robust algorithms for crowd analysis and tracking tasks.

3.2 *Drone control and integration*

The drone control and integration method employed in this research involved a systematic approach to ensure smooth and accurate operation of the drone platform. The DJI Ryze Tello drone was utilized as the aerial platform due to its compact size, maneuverability, and integrated features. The control and integration process encompassed flight control software, sensor integration, and communication interfaces. To enable precise control over the drone's flight parameters, customized flight control software was developed. This software facilitated autonomous flight operations by implementing algorithms for altitude control, position estimation, and trajectory planning (Cocca et al.,2022). Through the flight control software, commands were sent to the drone to adjust its altitude, navigate within the target area, and maintain stable flight. The integration of the flight control software enabled the drone to perform predefined flight patterns and execute maneuvers required for crowd detection and tracking.

In this research, we integrated the DJI Tello drone using the DJITelloPy and Pygame frameworks. At the drone control stage, we connected a laptop to a drone using dji tello drone Wi-Fi. We can transmit control commands like takeoff, landing, and control drone movements like flying forward, flying backward, fly left, and fly right using the DJITelloPy functions. Then, we created an interactive user agent for laptops using Pygame. To make it possible for users to control drones in an intuitive manner, we created a keyboard controller. Keyboard controller in this case refersto the buttons for takeoff, landing, and for directing a drone to various directions. By integrating DJITelloPy and Pygame, we can easily and effectively control the DJI Tello drone. The user can control the drone using the DJITelloPy-supplied controls as well as an application that was created using Pygame.

Furthermore, seamless communication between the drone and the research system was achieved through wireless communication interfaces. The drone's Wi-Fi capabilities enabled the transmission of captured imagery and sensor data in real-time to the research system. This facilitated continuous monitoring and analysis of the crowd scenes fromthe aerial perspective. The integration of wireless communication interfaces ensured a reliable and efficient exchangeof data between the drone platform and the research system, enabling prompt decision-making and adjustments in tracking algorithms.

3.3 Experimental setup

The experimental setup focused on conducting crowd detection experiments within an indoor room. The indoor environment was carefully conditioned to provide a controlled setting for evaluating the performance of the aerial object tracking system. The room dimensions were 29.16 m^2 to accommodate the flight path of the drone and allow for realistic crowd movement simulations. To ensure consistent and reliable results, appropriate measures were takento control the environmental factors. The room was maintained at a constant temperature of 24 °C minimize any potential thermal variations that could affect the drone's flight stability and the accuracy of crowd detection. The roomwas also shielded from external light sources to maintain consistent lighting conditions throughout the experiments.

To mimic real-world indoor scenarios, the room was furnished with objects and structures commonly found in indoor environments. These objects were strategically placed to simulate crowd formations, movement patterns, and potential occlusions. By incorporating these elements, the experimental environment aimed to provide a realistic representation of crowd dynamics in confined spaces. To ensure accurate evaluation of the aerial object tracking system's performance, the room was kept free from any external disturbances that could influence the flight behaviorof the drone or introduce noise in the captured images. Strict control over external factors, such as air drafts and ambient noise, was maintained to minimize their impact on the experimental outcomes. By carefully conditioning the environment, the research aimed to provide a reliable and representative setting to evaluate the performance of the aerial object tracking system in crowd detection within an indoor room.

4 Results and Discussion

The conducted experiments revealed several important findings regarding the performance and limitations of theaerial object tracking system and the controller system. These findings contribute to a better understanding of the system's capabilities and provide insights for potential improvements in future implementations. The low resolution of the drone camera emerged as a significant challenge during the experiments. It was observed that the drone's limitedresolution adversely affected the quality of the captured images, thereby impacting the subsequent image processing and analysis. Higher-resolution cameras may be required to capture more detailed and informative images for accuratecrowd detection and tracking. In the other hand, the usage time of the drone was found to influence the overheating of the drone's engine. As the drone operated for prolonged periods, the temperature of the engine increased, leading to performance issues. This overheating phenomenon caused the captured images to exhibit blinking and graphical distortions, compromising the accuracy and reliability of the image processing algorithms. Implementing efficient cooling mechanisms or utilizing drones with better heat dissipation capabilities could address this issue.

Figure 1. The drone controlled by built-in control system.

Figure 1 demonstrates how the djitellopy and pygame control systems can be used by users to control the DJI Tello drone's movement from a computer using the keyboard as an interface. Djitellopy, which enables keyboard inputprocessing in PyGame, makes it simple for users to connect to and fly Tello drones. With this configuration, users areable to quickly control Tello drone operations like takeoff, landing, climbing, descending, twisting, and bending through the keyboard's buttons. The control system offers a straightforward but efficient control experience when moving the DJI Tello drone using a computer and a keyboard.

Figure 2. The system display that has installed PysimpleGUI as its GUI.

Figure 2 shows the display image processing system using OpenCV has been installed in PySimpleGUI and hasbeen equipped with the "Save" button feature to save images. Whenever the button is pressed, the displayed image will be saved with an incremented name, allowing the user to easily track and manage the image result of processing.

Figure 3. The flashing display and graphic distortion due to drone overheat.

Figure 3 displays the output from the drone's image processing overheating camera. This happens as a result of using a drone for more than 7 minutes. The camera drone output would reportedly exhibit flashing displays and graphic distortion as a result of overheating.

Figures 4. The image stored using the features on the GUI.

Figure 4 displays the image processing outcomes that were stored using the user interface's "Save" button. (GUI).Users can save the outcomes for later use after processing photos using various approaches including filtration, segmentation, or other processing. The processed image will be saved in the appropriate format, such as JPEG, PNG,or TIFF, by pressing the "Save" button. With the aid of this procedure, the user is able to record the quality and detailthat have been attained through earlier stages of image processing.

The experiments indicated that the accurate detection range of the system was limited to approximately 1.2 to 5.4meters. Beyond this range, the system's ability to detect and track objects declined, resulting in decreased accuracy. Extending the detection range could be achieved through advancements in camera technology, such as higher zoom capabilities or the integration of additional sensors to enhance the system's perception capabilities. Furthermore, the experiments revealed that the connection range between the drone and the laptop was limited to approximately 10 meters. This restricted range could potentially hinder the system's mobility and operational flexibility, especially in scenarios requiring larger coverage areas. Exploring communication protocols with longer ranges or implementing signal amplification techniques could extend the operational range between the drone and the laptop.

Table 1: Distance test results.

No	Distance (m)	Number of People	Accuracy	Maximum Speed (km/hour)	Time of Kontrol Response (Second)	Detection response time from beginning (Second)
1	1.2	3	100%	22.14	0.83	1.54
2	2.4	3	100%	21.48	0.94	1.47
3	3.6	3	56%	22.64	0.91	3.89
4	4.8	3	32%	20.21	1.14	4.13
5	5.4	3	0%	21.6	0.98	0

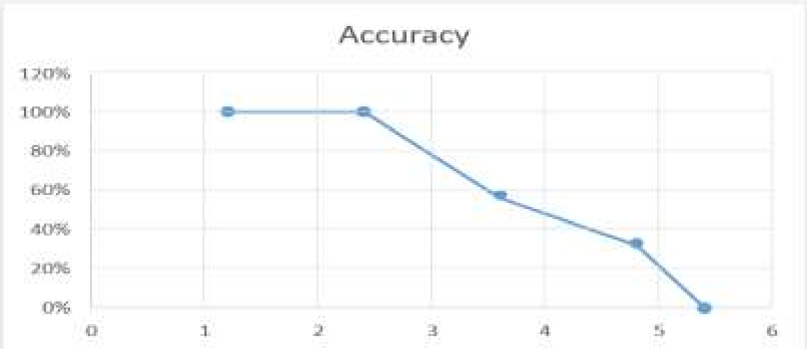

Figure 5. Diagram of accuracy.

Based on table 1 and figure 5, it can be concluded that in image processing, as the distance of object detection increases, the accuracy of detection tends to decrease. This is due to several factors. First, the farther the detection distance, the smaller the size of the object in the image. This leads to the loss of important details and information about the object. Furthermore, the further the detection distance, the more noise or interference that can affect imagequality. Disturbances such as distortion, signaling, or loss of detail in an image can reduce the accuracy of object detection. Furthermore, the farther the detection distance, the smaller the contrast between the object and the background, so the object detection algorithm may have difficulty distinguishing the object from its background. Therefore, it is important to consider the optimal detection distance in image processing to ensure maximum object detection accuracy.

In the context of drone and laptop-based object detection, it's important to understand that the maximum drone control distance and the maximum object detection distance are two different parameters. Maximum drone control ofa laptop typically involves the presence of a wireless communication channel between the drone and the pendali, suchas Wi-Fi or a radio link. As a result of this, there is a greater chance that there will be a failure in communication or a sinister force that can interfere with drone control as the control distance increases. On the other hand, the maximum object detection distance from the laptop depends on how far the object-detection algorithm and image processing system can detect objects. However, because to the loss of detail, increased noise, and decreased contrast, the detectionaccuracy decreases as the object detection distance increases.

Thus, it is important to consider realistic maximum distance limits and allow both in drone control and object detection. Factors such as the environment, communication signals, image quality, and the need for object detection should be considered to a balance between reliable drone control distance and adequate objects detection accuracy.

The result also shows that the experiments also emphasized the significant impact of lighting conditions on the drone camera's performance. The low resolution of the camera combined with suboptimal lighting conditions resulted in reduced image quality and decreased the system's detection capabilities. Enhancing the camera's sensitivity to lightor integrating advanced image processing algorithms that compensate for challenging lighting conditions could improve the system's performance in various lighting environments. Moreover, the experiments highlighted the vulnerability of the graphical output transmitted from the drone to potential disruptions caused by overheating or poornetwork connectivity. These disruptions could lead to the deterioration of the graphical quality, rendering the output unreliable and potentially hindering effective monitoring and analysis. Ensuring proper heat dissipation mechanisms and improving network stability is crucial for maintaining the integrity of the graphical output.

Table 2: Lighting test results.

No	Light	Distance (m)	Number of People	Accuracy
1	Ambient light	1.2	3	100%

2	Ambient light	2.4	3	100%
3	Ambient light	3.6	3	64%
4	Ambient light	4.8	3	36%
5	Ambient light	5.4	3	0%
6	Low light	1.2	3	56%
7	Low light	2.4	3	28%
8	Low light	3.6	3	0%
9	Low light	4.8	3	0%
10	Low light	5.4	3	0%

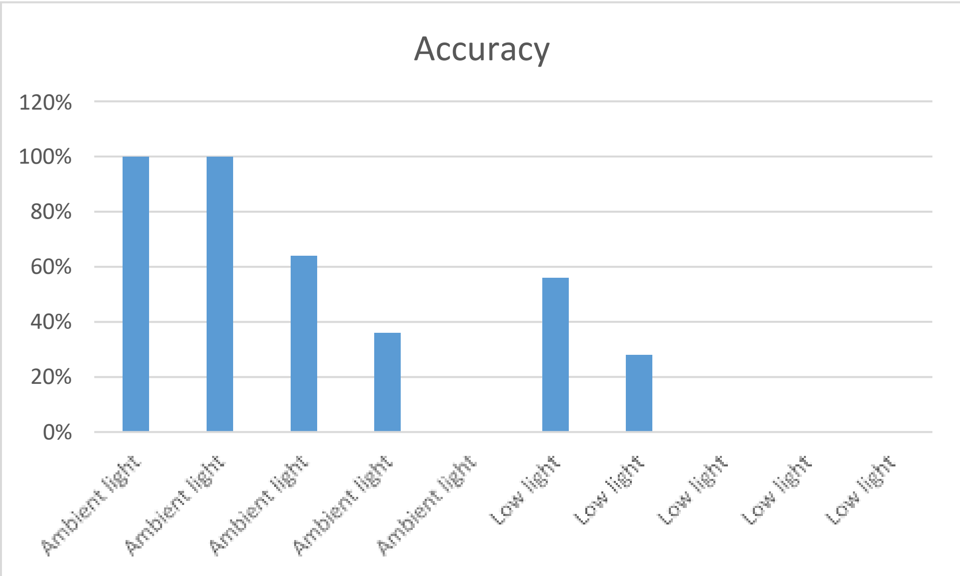

Figure 6. Accuracy diagram based on lighting condition.

Based on table 2 and figure 6, The contrast between ambient light and low light plays a crucial role in object detection in image processing. When the ambient light is sufficiently bright, the resulting image has high contrast, sharp details, and more discernible objects. This enables the object detection system to perform more accurately and produce better results. However, difficulties arise when working in low light circumstances. The resulting photos aredark and often have a high level of noise. This can result in object detail loss, low contrast, and fuzzy images. As a result, object detection algorithms may struggle to differentiate objects from backgrounds with small brightness changes.

At the same time, the experiments demonstrated the impact of adjusting the "detectMultiscale" function'sparameters in OpenCV on the accuracy of image processing.

Specifically, reducing the scaleFactor parameter resultedin lower accuracy in detecting relevant objects while also increasing the chances of false detections. Fine-tuning these parameters is crucial to strike a balance between detection accuracy and false positive rates, ensuring the system's optimal performance.

Table 3: scaleFactor Accuracy test results.

No	scaleFactor	Distance (m)	Number of People	Accuracy
1	1.1	2.4	3	100%
2	1.2	2.4	3	100%
3	1.3	2.4	3	56%
4	1.4	2.4	3	32%
5	1.5	2.4	3	0%
6	1.1	3.6	3	84%
7	1.2	3.6	3	64%
8	1.3	3.6	3	56%
9	1.4	3.6	3	0%
10	1.5	3.6	3	0%
11	1.1	4.8	3	24%
12	1.2	4.8	3	0%
13	1.3	4.8	3	0%
14	1.4	4.8	3	0%
15	1.5	4.8	3	0%

Based on table 3, scaleFactor parameters in object detection using play an important role in controlling image scalesize for the detection process. However, it should be remembered that the larger the scaleFactor value, the more inaccurate the detection of the object. If the scaleFactor value is too small, near-distance objects may be difficult to detect due to the relatively small size of the object on an image that is not scaled. On the other hand, if the scaleFactorvalue is too large, the object at a distance can become too small and lose detail in the image that is scaled massively. Therefore, it is important to choose the right scaleFactor to optimal object detection at various distances.

5 Conclusion

In conclusion, the conducted experiments provided valuable insights into the performance and limitations of theaerial object tracking system. The low resolution of the drone camera was identified as a challenge, impacting imagequality and subsequent image processing. Higher-resolution cameras are recommended to improve the system's ability to capture detailed images for accurate crowd detection. Overheating was observed as the drone operated for longer durations, leading to image distortions. Implementing effective cooling mechanisms or using drones with better heat dissipation capabilities can mitigate this issue. The experiments also highlighted the limited detection and communication ranges of the system, emphasizing the need for advancements in camera technology and communication protocols to extend these ranges. Lighting conditions were found to significantly affect image qualityand detection capabilities, suggesting the importance of enhancing camera sensitivity and employing advanced imageprocessing algorithms. Furthermore, adjusting the parameters of the image processing function in OpenCV was crucialto optimize detection accuracy and minimize false positive rates.

6 Grant information/Funding:

This research was funded by Penelitian Dasar dan Terapan (PDT) PPM Telkom University. This research was conducted on Embedded and Network System (ENS) Research Laboratory and Robotic SAS Research Laboratory, Faculty of Applied Science, Telkom University.

References

1. Al-Sa'd, M., Kiranyaz, S., Ahmad, I., Sundell, C., Vakkuri, M., & Gabbouj, M. (2022). A Social Distance Estimationand Crowd Monitoring System for Surveillance Cameras. *Sensors*, *22*(2). https://doi.org/10.3390/s22020418
2. Alzahrani, B., Barnawi, A., Irshad, A., Alhothali, A., Alotaibi, R., & Shafiq, M. (2022). A secure key agreement scheme for unmanned aerial vehicles-based crowd monitoring system. *Computers, Materials and Continua*, *70*(3). https://doi.org/10.32604/cmc.2022.020774
3. Castellano, G., Castiello, C., Mencar, C., & Vessio, G. (2020). Crowd Detection in Aerial Images Using Spatial Graphs and Fully-Convolutional Neural Networks. *IEEE Access*, *8*. https://doi.org/10.1109/ACCESS.2020.2984768
4. Castellano, G., Cotardo, E., Mencar, C., & Vessio, G. (2023). Density-based clustering with fully-convolutional networks for crowd flow detection from drones. *Neurocomputing*, *526*. https://doi.org/10.1016/j.neucom.2023.01.059
5. Chebil, K., Htiouech, S., & Khemakhem, M. (2022). Toward Optimal Periodic Crowd Tracking via Unmanned Aerial Vehicles. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4100367
6. Chen, G., Wang, W., He, Z., Wang, L., Yuan, Y., Zhang, D., Zhang, J., Zhu, P., Van Gool, L., Han, J., Hoi, S., Hu,
7. Q., Liu, M., Sciarrone, A., Sun, C., Garibotto, C., Tran, D. N. N., Lavagetto, F., Haleem, H.,

… Luo, Z. (2021). VisDrone-MOT2021: The Vision Meets Drone Multiple Object Tracking Challenge Results. *Proceedings of theIEEE International Conference on Computer Vision*, *2021-October*.https://doi.org/10.1109/ICCVW54120.2021.00318

8. Cocca, U., Giusti, A., & Roveda, L. (2022). Hand Gestures-Based Smooth 3D Trajectories Computation Applied to Real-Time Drone Control by Tracking 2D Hand Landmarks. *International Conference on Electrical, Computer,Communications and Mechatronics Engineering, ICECCME 2022*. https://doi.org/10.1109/ICECCME55909.2022.9987798
9. Fadzil, N., Abu Bakar, N. H., Idrus, S. M., Azmi, A. I., Mohammad, S. H., & Ali, N. (2021). Recent Development ofCrowd Monitoring Technology Solution for Covid-19 Prevention at Airport Terminal. In *International Journalof Nanoelectronics and Materials* (Vol. 14, Issue Special Issue InCAPE).
10. Fan, Z., Zhang, H., Zhang, Z., Lu, G., Zhang, Y., & Wang, Y. (2022). A survey of crowd counting and density estimation based on convolutional neural network. *Neurocomputing*, *472*. https://doi.org/10.1016/j.neucom.2021.02.103
11. Gajjar, V., Khandhediya, Y., & Gurnani, A. (2017). Human detection and tracking for video surveillance: A cognitive science approach. *Proceedings - 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017*, *2018-January*. https://doi.org/10.1109/ICCVW.2017.330
12. Jiang, M., Lin, J., & Jane Wang, Z. (2021). SHUFFLECOUNT: TASK-SPECIFIC KNOWLEDGE DISTILLATION FOR CROWD COUNTING. *Proceedings - International Conference on Image Processing, ICIP*, *2021- September*. https://doi.org/10.1109/ICIP42928.2021.9506698
13. Jiang, M., Lin, J., & Wang, Z. J. (2021). A smartly simple way for joint crowd counting and localization. *Neurocomputing*, *459*. https://doi.org/10.1016/j.neucom.2021.06.055
14. Lan, W., Dang, J., Wang, Y., & Wang, S. (2018). Pedestrian detection based on yolo network model. *Proceedings of2018 IEEE International Conference on Mechatronics and Automation, ICMA 2018*. https://doi.org/10.1109/ICMA.2018.8484698
15. Liang, D., Chen, X., Xu, W., Zhou, Y., & Bai, X. (2022). TransCrowd: weakly-supervised crowd counting with transformers. *Science China Information Sciences*, *65*(6). https://doi.org/10.1007/s11432-021-3445-y
16. Molchanov, V. V., Vishnyakov, B. V., Vizilter, Y. V., Vishnyakova, O. V., & Knyaz, V. A. (2017). Pedestrian detection in video surveillance using fully convolutional YOLO neural network. *Automated Visual Inspection and Machine Vision II*, *10334*. https://doi.org/10.1117/12.2270326
17. Nadeem, A., Ashraf, M., Qadeer, N., Rizwan, K., Mehmood, A., AlZahrani, A., Noor, F., & Abbasi, Q. H. (2022). Tracking Missing Person in Large Crowd Gathering Using Intelligent Video Surveillance. *Sensors*, *22*(14). https://doi.org/10.3390/s22145270
18. Shadakshri, V. H. C., Veena, M. B., & Dev, V. K. R. G. (2022). OpenCV Implementation of Grid-based Vertical Safe Landing for UAV using YOLOv5. *International Journal of Advanced Computer Science and Applications*, *13*(9). https://doi.org/10.14569/IJACSA.2022.0130957
19. Sindagi, V. A., & Patel, V. M. (2017). CNN-Based cascaded multi-task learning of high-level prior and density estimation for crowd counting. *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2017*. https://doi.org/10.1109/AVSS.2017.8078491

20. Syed Ameer Abbas, S., Oliver Jayaprakash, P., Anitha, M., & Vinitha Jaini, X. (2018). Crowd detection and management using cascade classifier on ARMv8 and OpenCV-Python. *Proceedings of 2017 International Conference on Innovations in Information, Embedded and Communication Systems, ICIIECS 2017*, *2018- January*. https://doi.org/10.1109/ICIIECS.2017.8275988
21. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). An image is worth 16*16 words: transformers for image recognition at scale. *Advances in Neural Information Processing Systems*, *2017-Decem*.
22. Wang, X., Lv, J., & Yun, Z. (2022). A Real-time improved pedestrian dead reckoning trajectory tracking algorithm. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *46*(3/W1-2022). https://doi.org/10.5194/isprs-archives-XLVI-3-W1-2022-197-2022
23. Wen, L., Du, D., Zhu, P., Hu, Q., Wang, Q., Bo, L., & Lyu, S. (2021). Detection, Tracking, and Counting Meets Drones in Crowds: A Benchmark. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. https://doi.org/10.1109/CVPR46437.2021.00772
24. Wu, Y., Sui, Y., & Wang, G. (2017). Vision-Based Real-Time Aerial Object Localization and Tracking for UAV Sensing System. *IEEE Access*, *5*. https://doi.org/10.1109/ACCESS.2017.2764419
25. Xiao, Y., Kamat, V. R., & Menassa, C. C. (2019). Human tracking from single RGB-D camera using online learning. *Image and Vision Computing*, *88*. https://doi.org/10.1016/j.imavis.2019.05.003
26. Yadav, S., Gulia, P., Gill, N. S., & Chatterjee, J. M. (2022). A Real-Time Crowd Monitoring and Management System for Social Distance Classification and Healthcare Using Deep Learning. *Journal of Healthcare Engineering*, *2022*. https://doi.org/10.1155/2022/2130172
27. Yan, Z., Duckett, T., & Bellotto, N. (2020). Online learning for 3D LiDAR-based human detection: experimental analysis of point cloud clustering and classification methods. *Autonomous Robots*, *44*(2). https://doi.org/10.1007/s10514-019-09883-y
28. Zhang, C., Yang, Z., Liao, L., You, Y., Sui, Y., & Zhu, T. (2022). RPEOD: A Real-Time Pose Estimation and ObjectDetection System for Aerial Robot Target Tracking. *Machines*, *10*(3). https://doi.org/10.3390/machines10030181
29. Zhang, G., Pan, Y., Zhang, L., & Tiong, R. L. K. (2020). Cross-scale generative adversarial network for crowd densityestimation from images. *Engineering Applications of Artificial Intelligence*, *94*. https://doi.org/10.1016/j.engappai.2020.103777
30. Zhu, L., Li, C., Yang, Z., Yuan, K., & Wang, S. (2020). Crowd density estimation based on classification activation map and patch density level. *Neural Computing and Applications*, *32*(9). https://doi.org/10.1007/s00521-018- 3954-7