Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 2 (2024)

An Artificial Intelligence Mechanism to Monitor and Manage Crowd

[1]Dr.Ashish Gupta, [2]Shruti Raj, [3]Shivendra Kumar Upadhyay, [4]Vikash Kandu, [5]Shivani Singh

[1] *Dept. Electronics and communication*
Galgotias College of Engineering and Technology
Greater Noida, India
[2] *Dept. Electronics and communication*
Galgotias College of Engineering and Technology
Greater Noida, India
[3] *Dept. Electronics and communication*
Galgotias College of Engineering and Technology
Greater Noida, India
[4] *Dept. Electronics and communication*
Galgotias College of Engineering and Technology
Greater Noida, India
[5] *Dept. Electronics and communication*
Galgotias College of Engineering and Technology
Greater Noida, India

Abstract— Effective crowd control is becoming more and more important for everyone's safety and wellbeing due to the growing number of people living in cities and big events. The accuracy and scalability of existing crowd monitoring and management systems are limited by their manual, labor-intensive nature. AI has shown promise in fixing a number of issues. This work presents a real-time artificial intelligence method for monitoring and controlling crowd behavior. YOLO v8 and image processing are used to detect and count crowds using a variety of processes, including picture acquisition, preprocessing, object recognition, and post-processing. Once the crowds have been identified and counted, the results may be evaluated for accuracy, efficacy, and potential applications. In order to monitor and identify people and assess crowds, the suggested system makes use of deep learning, machine learning, and computer vision techniques.

Keywords— Crowd, Artificial Intelligence, Mechanism.

I.Introduction

The Rise of the Urban Crowd: As the globe grows increasingly urbanized, crowd control is becoming a crucial ability. For a multitude of reasons, cities serve as hubs for social interaction, cultural exchange, and economic activity, drawing millions of people to gather in public spaces. This inflow presents a unique problem for law enforcement and event planners: ensuring the security, comfort, and safety of large crowds while facilitating their mobility and preventing disruptions.

Conventional Methods for Crowd Management: The majority of traditional crowd management tactics are manual ones, such as traffic control, security personnel, and physical barriers. While these methods have their applications, they often suffer from problems related to effectiveness, expandability, and adaptability. While manually monitoring crowds takes time and is subject to human error, deploying a large number of guards may be expensive and resource-intensive.

The Potential of Artificial Intelligence: The advancement of artificial intelligence (AI) offers a realistic route to a ground-breaking method of crowd management. AI-powered systems can Analyzing massive amounts of data from many sources, including social media, sensor networks, and security cameras, to offer real-time insights into the dynamics and behavior of crowds. With this information, one can:

- Crowd density and mobility may be predicted by AI systems through historical data analysis and current pattern recognition. Authorities may use this information to plan ahead, distribute resources proactively, and prepare for any future congestion or bottlenecks.
- Determine unusual or suspicious activity: Systems for anomaly detection based on artificial intelligence (AI) are able to identify odd patterns in crowd behaviour that may indicate potential threats or risks. Security personnel can prevent incidents by acting immediately with this information.
- Align resources as efficiently as possible: By analysing crowd density and movement patterns, AI systems may help determine the ideal locations for traffic signals, security officers, and other resources. This guarantees efficient crowd control with the least amount of disruption.
- Encourage collaboration and communication: AI-powered communication systems can facilitate real-time information sharing between law enforcement, security personnel, and the general public, enabling effective preparation and response in the event of emergencies or unforeseen circumstances.

The Need for More Research: Although AI holds great potential to enhance crowd management, a few problems and limitations still need to be worked out. Among them are:

- Data privacy concerns: Ensuring moral data collection, storage, and usage is essential to winning the public's confidence and adoption of AI-based crowd management systems.
- Fairness and bias: AI programs educated on biased data may generate unjust outcomes and amplify existing inequalities. Developing and deploying AI systems that handle each person in a crowd fairly and impartially is essential.
- Transparency and accountability: The functioning of AI algorithms must be transparent and comprehensible in order to preserve public confidence and responsibility.

This research study proposes a revolutionary artificial intelligence crowd monitoring and management system as a means of overcoming these challenges and limitations. This mechanism will protect individual privacy rights, promote justice and transparency, and improve crowd safety, security, and comfort by applying cutting-edge AI approaches.

II.Related Work

A framework for assessing and quantifying individuals inside a crowd is one component of the proposed crowd tracking and detection method. The initial step in this process is to remove the backdrop and superfluous pixels from the picture. This method makes counting and identification of crowds easier. A thorough explanation of the database used to evaluate the performance of the proposed system is also given in order to help identify the boundaries of the model. With this method, it will be possible to develop an automated system that can gather useful data from stored videos and do the counting and monitoring in a number of places, such as airports, retail establishments, train stations, universities, and other busy places, without the need for human personal. the recommended system will be able to recognize and keep an eye on crowds in real-time using Frame by Frame analysis, complex algorithms, and image processing methods.

III.Methodlogy

- Video Streaming: We establish the Frames Per Second (FPS) throughput rate for object recognition by using a camera. The accuracy and FPS capabilities are the first two limits to take into account when working on this issue.
- Pre-Processing: This entails reducing the size and converting the colour space to RGB. Common activities related to computer vision and image processing are performed using the OpenCV library. This refers to the process of using Deep neural networks (DNN) to make predictions or inferences based on input data, such as images or video frames, using the OpenCV library. In the realm of computer vision and image processing, the task of capturing and handling output frames will be entirely executed using OpenCV.
- Object Detection: In computer vision and image processing, object detection involves identifying and localizing instances of specific semantic objects within images and videos. This is achieved by defining bounding boxes around the detected objects.

- Object Tracking: "Surveillance, employing the geometric centre tracking method, is used for determining the central point. Enclosed regions (also known as bounding boxes) are utilized. Subsequently, the spatial separation of centroids is defined using straight-line distance based on Euclidean geometry. Additionally, it Avoid entities no longer within scope."

IV.Algorithm

Deep Learning: Deep learning, a subset of machine learning and artificial intelligence, employs artificial neural networks to learn from large volumes of Data. In this subject, understanding and progress are reliant on the computation of mathematical ex-pressions. "Makes use of phoney neural networks" is transformed into "employs artificial neural networks, which imitate human comprehension and reasoning, whereas AI is based on more fundamental assumptions. A neural network is significant in data processing, enabling computers to perceive, comprehend, and react to intricate scenarios more swiftly than humans.

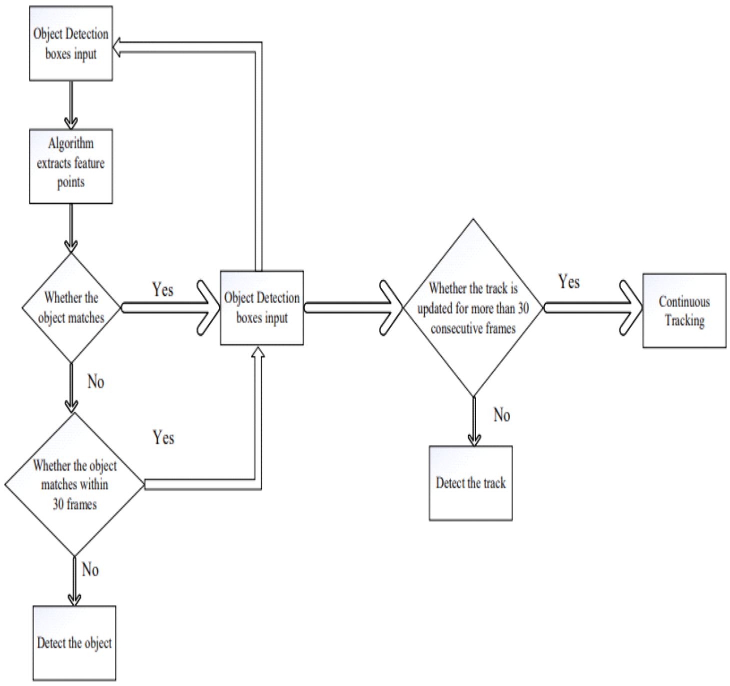
The image, sequence, and language model harnesses classic techniques to enhance computational understanding and response to complex situations understanding and conversion recognition attain all benefited from DNN.

Centroid-based Tracking: Utilizing a centroid-based tracking algorithm, the process of tracking centroids emerges as a highly effective and comprehensible technique. This method hinges on the Euclidean distance, gauging the separation between current and subsequent object centroids within consecutive frames of a video. Termed centroid tracking, it offers a straightforward approach to object tracking, employing {x, y} coordinates for each identified object across frames.

OpenCV: Centroid-guided tracing provides a direct and potent methodology. Through computation of the Euclidean span amidst current and fresh object Centers throughout sequential frames in cinematic reels, this technique is traditionally denoted as centroid pursuit. Using (x, y) coordinates for each detected object in each frame, the centroid tracking approach uses the assumption that certain sets of the bounding box are traded. Bounding box calculations are needed for each frame of the video or each object that the camera detects. These figures can be used to remove red eyes from pictures taken with stripes, track moving objects, Craft 3D point arrays through audio-enhanced camera frameworks, discern and identify visages, perceive tangible entities, illustrate human Endeavors in daily life, trail ocular advancements, survey vistas, and pinpoint indicators for augmented reality superpositions, all facilitated by OpenCV. OpenCV states that it has over 18 million downloads and over 47,000 users.

The Single Shot Detector (SSD) Algorithm is one of the methods used for object detection. Every deep learning model file that has been pretrained is stored on the Mobile-net SSD. There are two categories for SSD: 1) Remove the feature maps first. 2) Convolution filters are employed in object identification. The Single Shot Detector (SSD) may work on top of VGG, YOLO, and Mobile-Net in addition to other base networks because it is not dependent on any specific one of them. Portable Network was incorporated into the Flash Drive scaffold with the intent of addressing the difficulty of executing neural networks in real time on low-end devices that have high power and resource consumption.

V.Proposed System



The primary idea is to recognize crowd scenes in images by counting and recognizing human bodies in the image. Several procedures are used to process the input picture from the video in the first stage. Humans were then identified using the YOLO v5 algorithm, which was trained on the Crowd Human data set. determination of the quantity of heads in an image. If there are more heads than there are human bodies, a box is carved around each head. Part cut accuracy increased when the recommended edge enhancement approach was used. Finally, the improved image is re-examined using previous techniques for crowd detection

Pre-processing: The footage is processed in this stage by being split up into several frames. Each frame is a picture that has been resized to fit the YOLO v8 algorithm's input. In order to eliminate noise and blur, the image is further improved by using a Gaussian filter and sharpening. Two-dimensional Gaussian filters are often used in multi-scale edge detection methods for three primary reasons. It is important to consider this is because when combined with a Laplacian operator, only 2D Gaussian filters do not generate erroneous edges as the scale increases. Second, with proper application, the optimal combination of spatial and frequency localization is provided by Gaussian filters. Thirdly, Spatial convolution with Gaussian filters is very efficient as only rotationally invariant 2D filters can be separated in both horizontal and vertical directions.

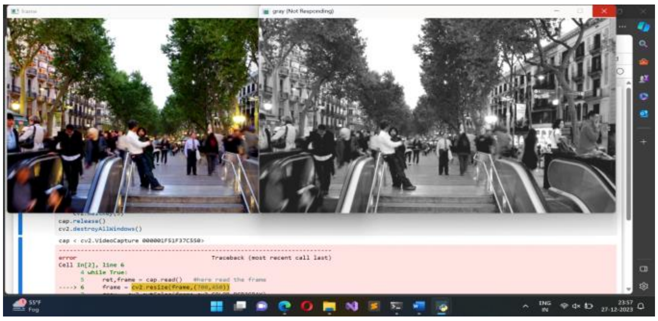
Human head and body detection by YOLO v8: With the YOLO v8 technique, an object detection network based on regions is used to function as a single stage detector. With its quick processing, the YOLO recast object detection as a regression issue. Real-time person searches and vision systems have recently made use of YOLO v8. The three primary components of YOLO v8 are the spine, the brain, and the detection. A Convolutional Neural Network (CNN) which collects and processes picture characteristics at different granularities is the central component of the system. By leveraging the Crowd person data set, the person detecting body in the YOLO v8 model was improved. when using the test model technique. Follow these procedures to train YOLO v8 on these data sets: Preprocessing the data to identify the "head" (0) and "person" (1) classes; the "person" class in the original "Crowd Human" annotations refers to "complete body" (including body portions that are obscured). Still, we have to organize our dataset's structure before we can train the Yolo-v5 model. In order to accomplish this, we must compile all of the label values (x, y, width, and height) for a single image file for each image file (.jpg,.png, etc.) and the related pictures in a text file format (.txt). Second, copy over the required files to download the pre-trained weights. This file has to be modified after it is saved in order for Yolo-v8 to work with it. You must change the *.yaml file for the Yolo-v8 model. To match the number in the model, all we have to do is adjust the number of classes in the YAML file for our model. When it comes to a head and accomplish body, two classes were used (head and body). In the earlier stage, the YOLO v8 algorithm detected the head and human bodies. The number of persons and the number of head detectors are compared in this stage. This model computes and verifies if there are more heads than bodies in order to improve the picture. Secondly, transfer the necessary files to obtain the pre-trained weights. To make this file compatible with Yolo-v8, it must be altered once it is saved. For the Yolo-v8 model, you need to modify the *.yaml file. All we need to do is change how many classes our model's YAML file number and the model number must coincide. Two classifications (head and body) were employed in relation to a head and body. The head and human bodies were identified in the previous stage using the YOLO v8 algorithm. This stage involves comparing the number of individuals and head detectors. This model computes and finds out if there are more heads than bodies in order to enhance the image.

Enhancement segmentation: Image processing frequently uses segmentation as a technique. Once a digital image is split apart into many zones, each with its own distinctive label, the visual qualities of Each zone's pixels often have a similar appearance. For the purpose of segmenting images, a variety of techniques and methodologies can be used, such as region-based approaches, border monitoring, clustering, and thresholding. The augmentation of picture segmentation based on edges is the main emphasis of this research. When applying the YOLO v8 approach in the previous step to compare the number of heads and bodies, particularly in large groups, there are fewer people than there are heads. To improve person identification, a method to accentuate the body characteristics in the image has been provided. The operation is taking the coordinates of the head and its centre and subtracting a square whose coordinates are found by processing them by the code.

VI.Results And Discussion

YOLO v8 and image processing are used to detect and count crowds. The procedure entails a few steps: pretreatment, object recognition, post-processing, and image capture. The findings may be evaluated in terms of precision, effectiveness, and possible uses after the crowds have been identified and tallied. This is a summary of the project's discussion and outcome sections.

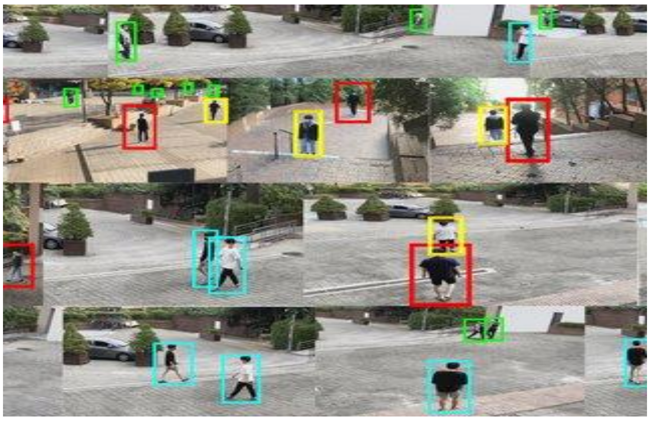
A testbed for people detectors in congested areas is the Crowd Human dataset. It is divided into three sections: testing (5000 images), verification (4370 photos), and training (15000 photos). From the training and verification subsets, there are 470k human examples, with 22:6 persons per image and additional opacities [19]. Every human example contains three limiting boxes: one for the face, one for the visible portion of the human body, and one for the entire body.



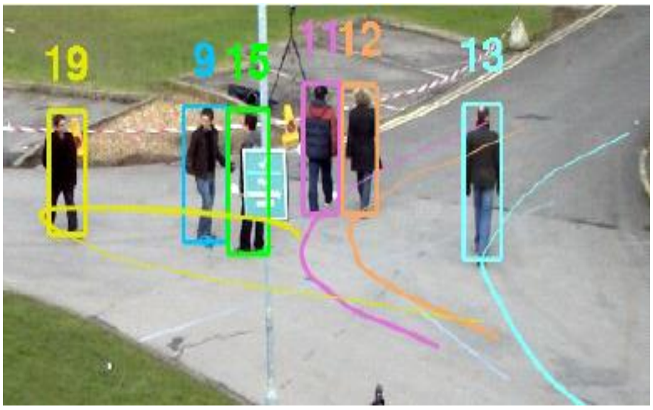
Yolov8 was evaluated on the validation dataset and trained on the Crowd Human dataset because the testing dataset's online assessment site is currently down. In the process of training, the input image is enlarged so that the long edges are limited to 1333 pixels and the short edges measure 800 pixels.

According to reference [20], the Crowd Human dataset's anchor aspect ratios are 0:5, 1, and 2.This system can function in real time on a regular CPU. Talk about how well YOLO v8 was able to identify crowds. Metrics like accuracy, recall, and F1-score can be used to assess this. Describe the quantitative findings, including the quantity of false positives, false negatives, and genuine positives.

The accuracy of human detection is improved by using deep learning object detectors. In order to improve tracking precision, it further utilizes two other object tracking algorithms, Centroid tracking and correlation filters. These enable it to identify new individuals and retrieve those who may have become "stuck" in the monitoring procedure.



Vehicle traffic counts may also be performed with this technique. A trained YOLO (You Only Look Once) model was used to count the number of people in busy and complicated settings. Experiments were carried out in a lab and an outside environment. A camera was positioned at a fixed location to capture video of persons walking on the road for the outside experiment. The camera was positioned high in the lab setup to capture a video of individuals moving through a congested environment. Show off the crowd counting algorithm's performance. Incorporate measures like processing speed and counting accuracy. Talk about the algorithm's efficiency at counting crowds in situations that happen in real time.



To identify and tally the persons in the films, the trained YOLO model was applied. Furthermore, conventional computer vision methods like blob detection and background removal were employed to tally the individuals in the films. The outcomes from conventional computer vision methods and the trained YOLO model were then contrasted. Additionally, by altering the object detection confidence threshold, the trained YOLO model's performance was assessed.

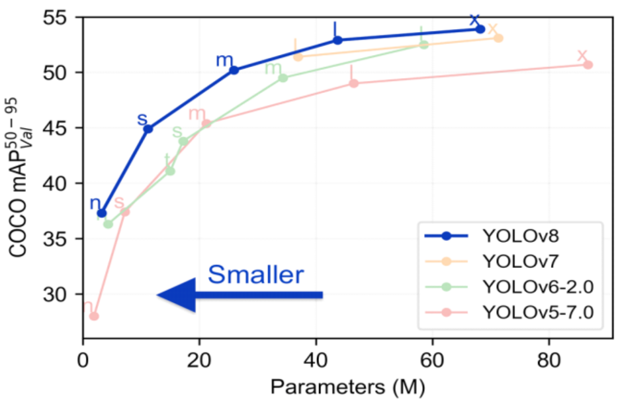
The performance of the YOLOv8 and its modified variants is displayed in the training results. includes information on adjusted network layers, settings, and outcomes. After 100 epochs, the results are assessed using the validation set. The pre-trained weight model performs better than the suggested models.

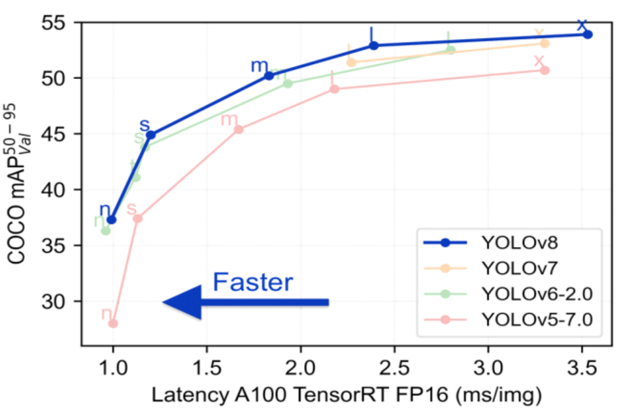
Finding the object in the image is the first problem of object detection; the second is determining if the thing exists at all. The majority of classes in a typical dataset will also have nonuniform distributions. Because it's important to estimate the risk of inappropriate categorization, bias will emerge with a simple precision-oriented metric. As a result, each Box was identified, and a final score was given to assess the model with different levels of confidence. As the evaluation standard for this work, the Mean Average Precision (mAP) [21] was used. To get the mAP score, the mAP is multiplied by all classes and all Intersection over Union (IoU) criteria.

VII.Comparison

Table I compares the performance of the three techniques using Yolo's data set training. Yolov5 with training on the Crowd Human dataset has a mAP of 0.771, which is 0.166 less than the suggested methods. Conversely, Yolov5's map of mAP is least in the Coco dataset. at 0.041. The suggested methods therefore offer a better trade-off between speed and accuracy in algorithm creation. The models in question were ones for counting people in a crowd and for detecting human bodies.

A little quicker inference speed is possessed by EL YOLOv8s, as per the network specifications. Analyze the performance of the suggested method against other cutting-edge crowd identification and counting techniques, if relevant, by comparing it to baselines. Point out each approach's advantages and disadvantages.





S. No	Model mAP	Body of human
1.	Yolov5+coco dataset training	0.771
2.	Yolov5+ CrowdHuman dataset training	0.896
3.	Proposed methods	0.937

Table 1Result from the Detection Model

References

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, ''What will 5G be?'' *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] V. Cisco, ''Cisco visual networking index: Forecast and trends, 2017–2022,'' Cisco, San Jose, CA, USA, White Paper, 2018.

[3] M. Mohammadi and A. Al-Fuqaha, ''Enabling cognitive smart cities using big data and machine learning: Approaches and challenges,'' *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 94–101, Feb. 2018.

[4] H. Hassanein, N. Zorba, S. Han, S. S. Kanhere, and M. Shukair, ''Crowd management,'' *IEEE Commun. Mag.*, vol. 57, no. 4, pp. 18–19, Apr. 2019. [5] U DESA, New York, NY, USA. (2018). *World Urbanization Prospects: The 2018 Revision, Key Facts*. Accessed: Dec. 20, 2018. [Online]. Available: https://population.un.org/wup/Publications/

[6] V. A. Sindagi and V. M. Patel, ''A survey of recent advances in CNN-based single image crowd counting and density estimation,'' *Pattern Recognit. Lett.*, vol. 107, pp. 3–16, May 2018.

[7] A. B. Chan, Z.-S. John Liang, and N. Vasconcelos, ''Privacy preserving crowd monitoring: Counting people without people models or tracking,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–7.

[8] F. Adib and D. Katabi, *See Through Walls With WiFi!*, vol. 43, no. 4. New York, NY, USA: ACM, 2013.

[9] T. Franke, P. Lukowicz, and U. Blanke, ''Smart crowds in smart cities: Real life, city scale deployments of a smartphone based participatory crowd management platform,'' *J. Internet Services Appl.*, vol. 6, no. 1, p. 27, Aug. 2015.

[10] W. Kang and Y. Han, ''SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization,'' *IEEE Sensors J.*, vol. 15, no. 5, pp. 2906–2916, May 2015.

[11] S. Qiu, Z. Wang, H. Zhao, K. Qin, Z. Li, and H. Hu, ''Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion,'' *Inf. Fusion*, vol. 39, pp. 108–119, Jan. 2018.

[12] J. Wang, Y. Wang, D. Zhang, L. Wang, C. Chen, J. W. Lee, and Y. He, ''Real-time and generic queue time estimation based on mobile crowdsensing,'' *Frontiers Comput. Sci.*, vol. 11, no. 1, pp. 49–60, Feb. 2017.

[13] Y. Wang, J. Wang, and X. Zhang, ''QTime: A queuing-time notification system based on participatory sensing data,'' in *Proc. IEEE 37th Annu. Comput. Softw. Appl. Conf.*, Jul. 2013, pp. 770–777.

[14] J. Steenbruggen, E. Tranos, and P. Nijkamp, ''Data from mobile phone operators: A tool for smarter cities?'' *Telecommun. Policy*, vol. 39, nos. 3–4, pp. 335–346, May 2015.

[15] F. Calabrese, L. Ferrari, and V. D. Blondel, ''Urban sensing using mobile phone network data: A survey of research,'' *ACM Comput. Surv.*, vol. 47, no. 2, p. 25, Jan. 2015.

[16] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, ''Position and orientation estimation through millimeterwave MIMO in 5G systems,'' *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.

[17] F. Zafari, A. Gkelias, and K. K. Leung, ''A survey of indoor localization systems and technologies,'' *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2568–2599, 3rd Quart., 2019.

[18] F. Aloul, A. Sagahyroon, A. Al-Shami, I. Al-Midfa, and R. Moutassem, ''Using mobiles for on campus location tracking,'' in *Proc. 7th Int. Conf. Adv. Mobile Comput. Multimedia (MoMM)*. New York, NY, USA: ACM, 2009, pp. 231–235.

[19] F. Meneses and A. Moreira, ''Large scale movement analysis from WiFi basedlocationdata,''in*Proc.Int.Conf.IndoorPositioningIndoorNavigat. (IPIN)*, Nov. 2012, pp. 1–9.

[20] S. Sendra, M. Garcia, C. Turro, and J. Lloret, ''People mobility behaviour study in a university campus using WLANs,'' in *Proc. 3rd Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol.*, Oct. 2009, pp. 124–129.