2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986

Baghdad Science Journal

Implementing Real-time Visitor Counter Using Surveillance Video and MobileNet-SSD Object Detection: The Best Practice

***Nasser Al Musalhi**[*1], **Ali Mohammed Al Wahaibi**[2], **Mohammed Abbas**[3]*

[1]Department of Information Technology, University of Technology and Applied Science, Ibra, Oman.
[2]Faculty of Communication, Visual Art and Computing, Universiti Selangor (UNISEL), Bestari Jaya, Malaysia.
[3]Department of Information Technology, University of Technology and Applied Science, Sur, Oman.
*Corresponding Author.
ICCDA2023: International Conference on Computing and Data Analytics 2023.

Received 28/12/2023, Revised 19/04/2024, Accepted 21/04/2024, Published 25/05/2024

Abstract

Counters that keep track of the number of people who enter a building are a useful management tool for keeping everyone who uses it safe and happy. This paper aims to employ the MobileNet-SSD machine learning approach to implement a best practice for visitor counter. The researchers have to build a different scenario test dataset along with the MOT20 dataset to achieve the proposed methodology. Implementing different experiments in single-user, one-one; two-two users; many-two, and multiple users in different walking directions to detect and count shows varied results based on the experiment type. The best achieved by single-user and one-to-one model; both are scored 100% of detecting and calculating for in or out.

Keywords: CNN, Mobilenet-SSD, MobileNet, Object tracking, real-time Object detection, SSD.

Introduction

A visitor count is useful for capacity management, security, event planning, and publicity. It may also be used to track visitor numbers over time, which can inform future expansion plans. Visitor counts can be used as evidence of the attraction's popularity or as a basis for incorporating customer input into future advertising campaigns. In general, tracking the number of visitors is an important tool for controlling and enhancing the experience that visitors have at a location.

When it comes to counting visitors, machine learning may offer several advantages, including increased accuracy, productivity, scalability, and comprehension. Machine learning algorithms may be trained to reliably count visits even in tough circumstances, therefore minimizing mistakes and giving more dependable data[1]. They can handle data swiftly and effectively, making them perfect for conducting visitor counts in real time. They are also capable of being scaled to accommodate a wide variety of distinct locations and sorts of visitors, which makes them appropriate for a wide range of applications.

The computer vision issue of detecting and following objects is crucial and has many practical uses. Autonomous cars, robotics, surveillance, and many more applications rely on fast, precise object detection and tracking. Machine learning methods,

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

including convolutional neural networks (CNNs), have made great strides in recent years towards solving this issue.

When you're monitoring an object you're following its or its owner's movements to locate it or learn more about its trajectory[2]. Many distinct technologies exist, including CCD cameras, night vision equipment like goggles, and thermal imaging cameras [3–6]. The discipline of video surveillance makes extensive use of such gadgets. A vast data set is mined for insights by a smart visual surveillance system. When using many cameras for visual surveillance, not only are things easier to spot but their movements may also be tracked for more insight into their patterns of activity. When there is a need to keep an eye on a high-risk area—one that may be the scene of a theft, an accident, or some other form of crime—a video surveillance system is employed.

Security cameras aimed at the public are now both affordable and widely available. Because of improvements in both surveillance camera technology and computer processing power, it is now possible to create video surveillance systems that can help ensure the public's safety[7].

This research aims to study different scenarios to determine the best possible scenario to be applied for visitor counters using MobileSSD in real-time object detection.

MobielNet model

MobileNet is a popular option for mobile and embedded vision applications, and there are various versions available. The initial version V1, which was released in 2017, makes use of depthwise separable convolutions to lower the amount of computing work required while keeping the same level of accuracy. The second version V2 [8–10], which was released in 2018, makes use of inverted residual blocks and linear bottleneck layers to significantly decrease the amount of processing cost while simultaneously boosting accuracy. The third version V3, which was released in 2019, makes use of attention processes that are channel-wise as well as spatial-wise to enhance accuracy and minimize model size. The V3 is optimized for the Google Edge TPU and employs a quantization-aware training approach to optimize the model for the 8-bit integer operations that are enabled by the Edge TPU [11,12]. MobileNet has become a well-liked solution for mobile and embedded vision applications because it offers a variety of choices for striking a balance between accuracy, performance, and the size of the model. Fig. 1 shows the MobileNet v3 model architecture.

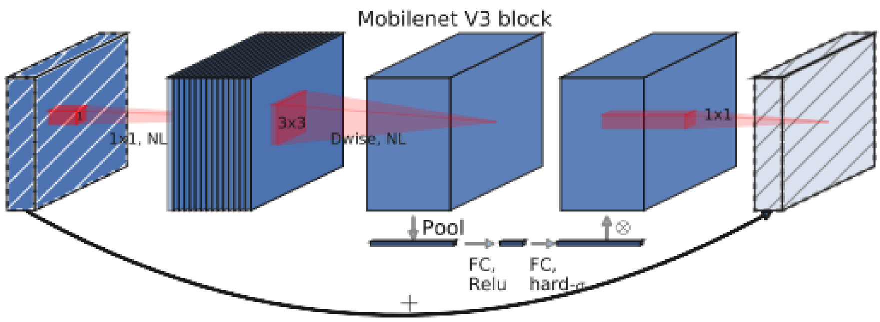


Figure 1. MobileNet v3 model architecture [11]

MobileNet-SSD

SSD allows us to detect many objects in a single shot, while other methods, such as the R-CNN series, require two shots (one to generate region suggestions and another to detect the item of each proposal) from the same input picture. Thus, SSD is significantly quicker than two-shot RPN-based methods [9,11,12].

The SSD is built on VGG as its central network, to which additional convolutional layers have been added. Six feature maps with layers are used to forecast the item class and coordinates, all based on SSD300 with a 300x300 input size. In the end, a non-maximum suppression (NMS) technique is used for detection. This technology outperforms more

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

traditional approaches because it can recognize objects across scales utilizing the feature map of each layer. The shallow layers' feature maps lack substantial high-level semantic information, making it difficult for this model to recognize tiny objects [8,13,14]. Fig. 2 represents the architecture of the SSD.

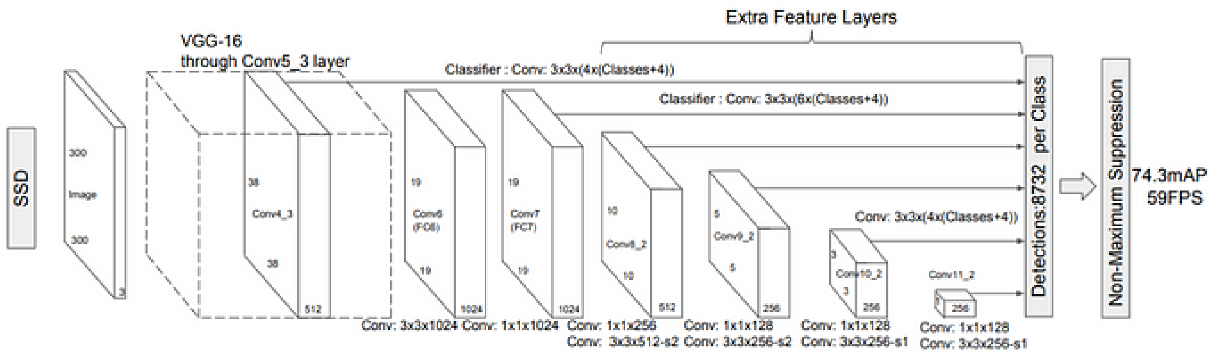


Figure 2.architecture of the improved map block in a single-shot multi-box detector (SSD-EMB)[13]

The MobileNet Single Shot Detector (MobileNetSSD) is a model that has been developed with deep neural networks to detect and track moving objects[8,15,16]. A lightweight and effective model for object recognition on mobile and embedded devices is created by combining MobileNet with SSD. Objects are classified into one of many predefined categories using a multi-layer model developed using convolutional neural networks. Recent developments in deep learning with image processing allow for the use of feature maps with a better resolution to detect such things. Mobilenet SSD is an object detection model that determines the output bounding box and object class based on the input picture. Using Mobilenet as its central network node, the SSD model can accomplish rapid object recognition that is well suited to usage on mobile devices[17].

Single Object Tracking

Single object tracking, or (SOT) refers to a computer vision job that includes monitoring a single item in a video sequence for the course of its duration[18,19]. It finds use in a variety of fields, including autonomous driving, robotics, and surveillance. SOT algorithms often make use of computer vision techniques to track the object. These approaches include object identification, feature extraction, and motion estimation. In the field of computer vision, SOT may be approached from a variety of different angles using techniques such as correlation filters, deep learning, and particle filters. Particle filters use a probabilistic model to track the object by sampling possible object locations and weights based on the similarity between the object's appearance and the observation, whereas correlation filters use feature representations to track the object based on the correlation between the appearance of the object and a template. Deep learning-based SOT methods use neural networks to learn robust feature representations.

Multiple Object Tracking

In the field of machine learning, multiple object tracking (MOT) refers to a computer vision problem that requires tracking several objects in a video stream over time [3,20]. It is difficult because of occlusions, changes in the appearance of things, and the existence of several objects in the same region. Machine learning techniques are utilized to automatically recognize and track objects in video streams to solve these obstacles. Object tracking algorithms are used to connect objects across frames and keep their identities over time, whereas object detection techniques are used to detect objects in each frame of the video stream. Deep learning and reinforcement learning are two examples of machine learning approaches that may be utilized to enhance the accuracy and resilience of MOT systems. In general, the process of tracking many objects at the same time in machine learning is difficult and significant work that has a wide range of applications, including surveillance, robots, and autonomous driving.

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986

Literature Review

Wei Lu et al research on video object detection using a non-local prior of spatiotemporal context. They have devised a model for detecting objects in videos that uses spatiotemporal attention based on non-local priors. The suggested model can fully utilize the spatiotemporal contextual information retrieved from video sequence pictures, which is an advantage over current attention models. Both the Overhead Contact System (OCS) driving recorder dataset and the OTB50 dataset are used to test and evaluate our models, which are implemented in a standard object detection framework. Their model improves its mAP value more than ours does, demonstrating that ours can improve its performance in a wide range of challenging video sequences [21]. The work by M. R. Marshall et al. provides a real-time object detection and tracking system as well as a unique source-object attribution analysis. For 3-D Object Tracking in Panoramic Video and LiDAR for Radiological Source-Object Attribution and Improved Source Detection, this analysis pipeline has been implemented on a custom-developed system consisting of a stationary 2-inch by 4-inch by 16-inch NaI(Tl) detector colocated with a 64-beam LiDAR and four monocular cameras. They show how to reliably distinguish between objects carrying sources and those without using physics-based models and how to properly connect trajectories from monitored objects to spectroscopic gamma-ray data in real time. Both video and LiDAR are characterized in terms of their source-object attribution skills through a quantitative performance evaluation presented by the researchers. Moreover, the research shows that our contextual-radiological data fusion technique may be used to simultaneously monitor pedestrians and cars in a simulated urban setting, therefore enhancing detection sensitivity and situational awareness [4].

Tiny SSD, introduced by Alexander Wong and colleagues, is a single-shot detection deep convolutional neural network for real-time embedded object detection. It consists of a highly optimized, non-uniform fire subnetwork stack and a highly optimized SSD-based auxiliary convolutional feature layer stack, both of which are tailored to reduce model size without sacrificing object detection performance. Tiny SSD can achieve an mAP of 61.3% on VOC 2007 (4.2% higher than Tiny YOLO) while having a model size of only 2.3MB, which is 26 times smaller. The experimental findings presented here demonstrate the feasibility of designing tiny deep neural network architectures for real-time object identification that are optimal for embedded applications [5].

Also, a group of researchers led by Maha Yaghi wrote a paper entitled "Real-time Contact Tracking." The Use of Several Cameras During a Pandemic to effectively implement pandemic control measures, video object tracking aimed at contact tracing and moving object monitoring is becoming increasingly crucial. A computer vision-based technique for contact tracking using fixed security cameras is proposed in their research. All moving objects in the input videos are tracked, and the space between them is determined thanks to the transformation into a bird's-eye perspective. The technique uses background removal to bring focus to the foreground, morphological procedures to get rid of distracting noise, and blob analysis to locate the interconnected areas in the resultant foreground video. Calculating the Euclidean distance between the objects in the video allows Kalman filters to estimate their velocities, which may then be used to track their interactions. Shopping centers, airports, and schools are just some of the many public spaces that might benefit from this algorithm. As a result, anyone who may have come into contact with the disease can be located, evaluated, and managed. Information for the tests was gathered in a real-world setting, with a mobile phone camera mounted on a tripod taking the place of the traditional stationary camera. The work was put into action and tested, and the outcomes proved the workable and efficient nature of the suggested approach. To estimate the distance between objects in the input video frame from numerous cameras, the system was able to detect them[5].

The goal of Abbas S. B. Sadkhan and coworkers' study, entitled "An Investigation on Moving Item Tracking and Detection in Pictures," was to examine and assess the methods previously used for locating and identifying moving objects in still photographs. They adopted a unique two-stage strategy to find the

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

moving item in the image and follow its path. The primary objective of such systems is to locate the source of motion in a video. When an object has been spotted, the next step is to follow it[22].

Methodology

To achieve the aim and goals of this paper, the researchers have used sequence steps as represented in Fig. 3. Input video from dataset then object detection using MobileNetSSD, tracking the human object only and then counting IN/OUT, compared with other input video scenarios, and finally sending the result to the database. Fig. 3 represents the flowchart of the methodology.

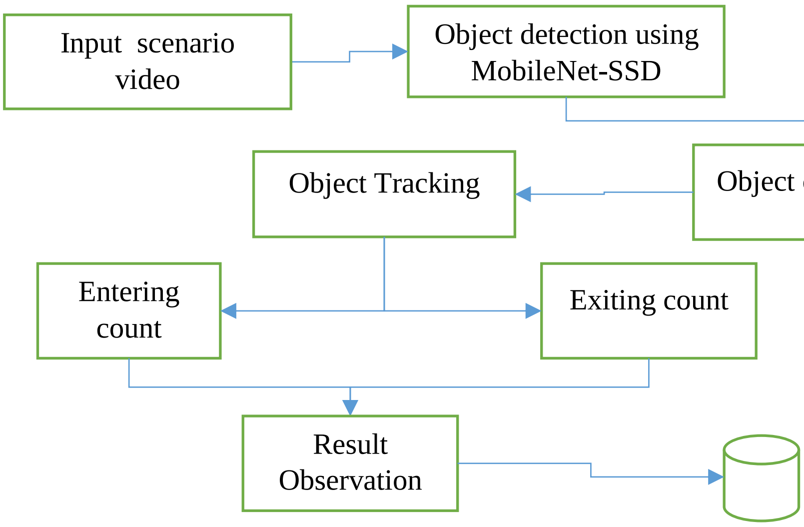


Figure 3. Methodology flowchart

Datasets

Two types of datasets have been used in this paper as the following:

(a) Motion Object detection (MOT20)

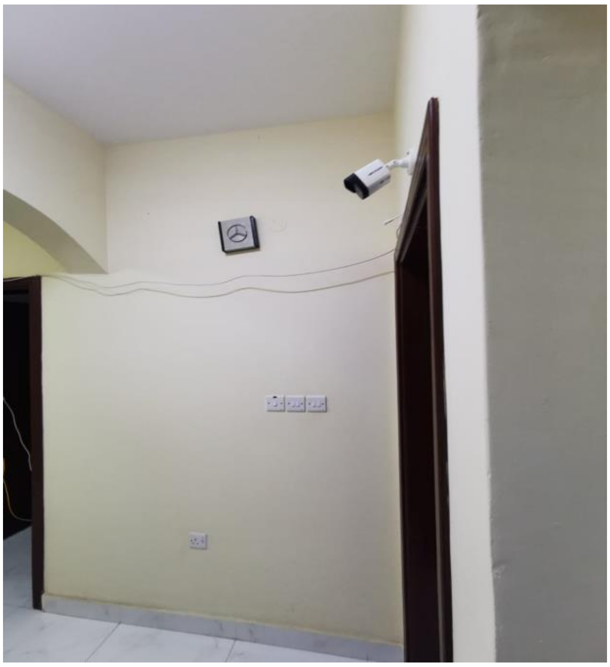
The new benchmark's dataset MOT20 was hand-picked to test trackers and detectors in densely populated environments. Unlike prior tests, some of the new sequences had a pedestrian density of 246 people per frame. MOT20 dataset[14]. accumulated a total of 8 sequences with 25 FPS for each film, both for training and testing purposes. The scenes are shot in three locations. Several takes of each scene are filmed and included in the trial and production versions. Nonetheless, experiments save one scenario for testing purposes to see how well the approaches generalize. The dataset is extended for serials of datasets.

(b) Experimental Lab Dataset

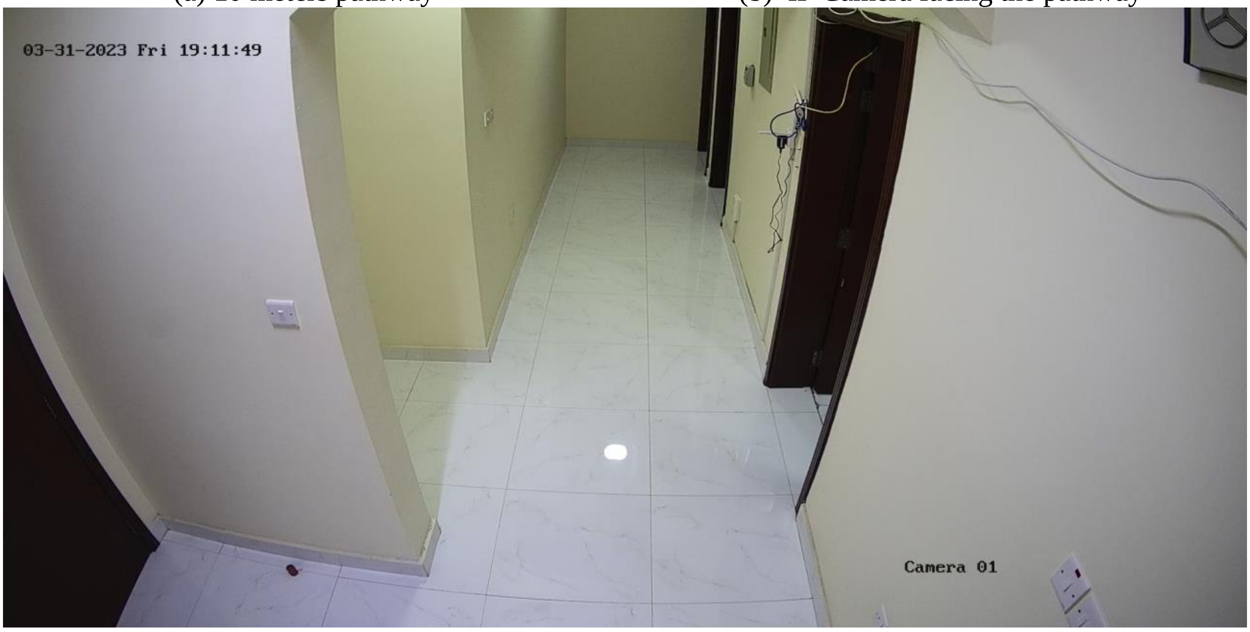
Lab environmental has been set with a 10-meter pathway Fig. 4(a) and a Hikvision IP camera with a 4MP setup with a 65-degree view angle Fig. 4(b) as shown below. Fig. 4(c) shows the camera view over the pathway.

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

(a) 10 meters pathway



(b) IP Camera facing the pathway



(c) Camera view with a 65-degree angle

Figure 4. (a) Passage of 10 meters in length; (b) IP Camera facing passage; (c) Camera view with 65-degree view angle

Video streaming was recorded in different scenarios; a single person walking, multiple persons walking, and multiple persons walking in entering, and exiting the door. Figs. 5 - 9 show the different scenario samples.

Lab Experiments

To testify to the best practice implementation, the researches have conducted several lab experiments using lab recorded. The record different scenarios as appears in Figs. 5 - 7. All videos are bi-directional for in and out. detecting threshold set with 50% confidentiality of human detection and classifying;

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986

30 FPS Frame rate skipping in tracking;50% max disappearance in the maximum distance of 7 meters; The model tested in given a single video with no consideration of the length of the video also tested videos given by a huge number of people using the MOT20 dataset as shown in Fig. 4. The accuracy of both experiments is evaluated. Table 1 shows the accuracy result in the MobileNet-SSD model.

(a) Single user tested IN/OUT. This scenario has conducted using a single user for either in or out through the gate. Fig. 5 shows the status of in and out detecting and counting.

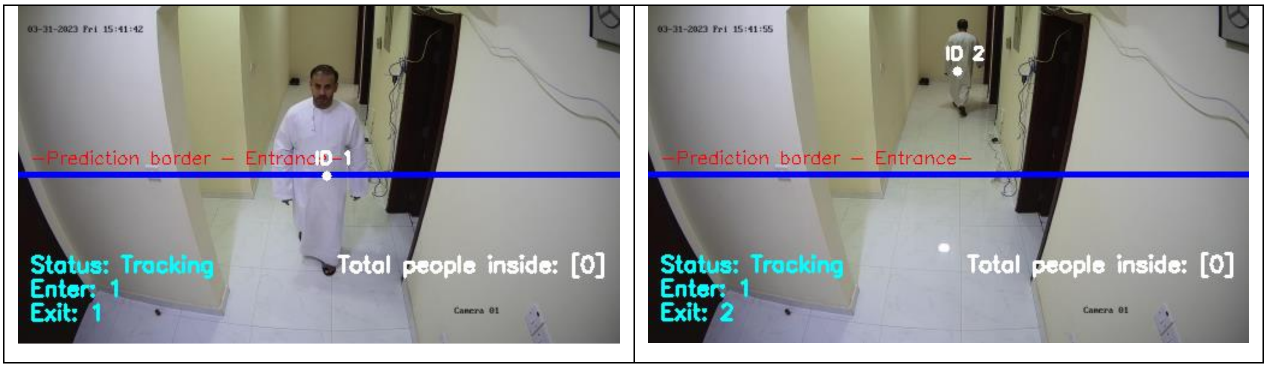


Figure 5. Single user in/out counter

(b) One-One user tested IN and OUT at the same time. This scenario a concurrent users in and out at the same time. Fig. 6 shows the accuracy of in and out detecting and counting.

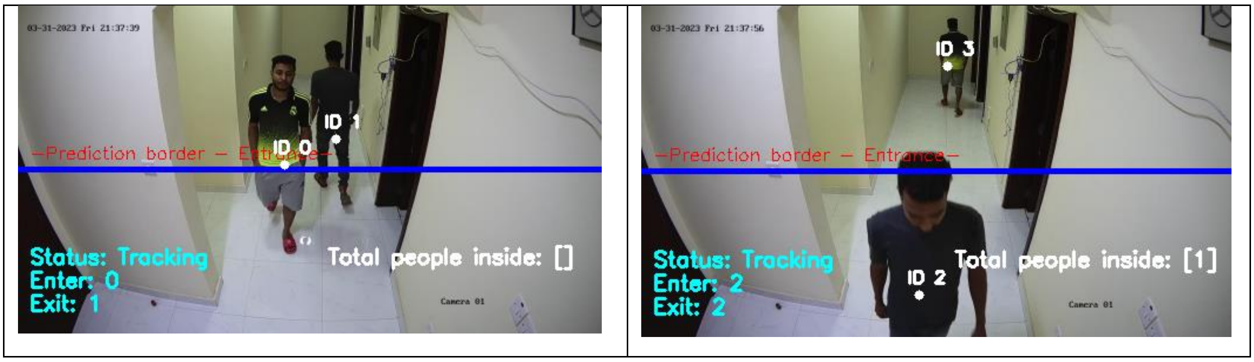


Figure 6. One/One users counter

(c) Two-One users at a time, in this experiment, used two users in one way and a single user in the opposite way. Fig. 7 shows the experiments of Two-One users.

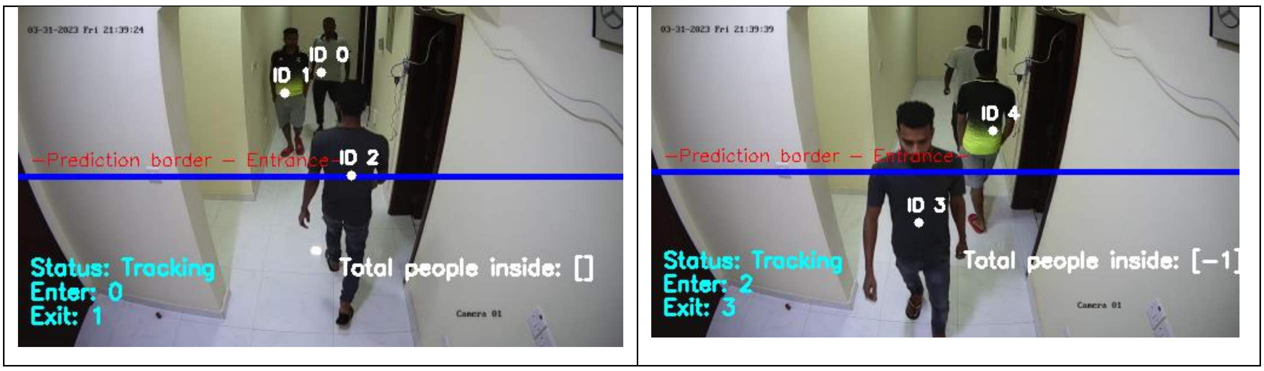


Figure 7. Two users in/out and one opposite direction user's counter

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

(d) Two-Two users at a time. In this experiment, two users in each way at the same time. Fig. 8 shows the results.

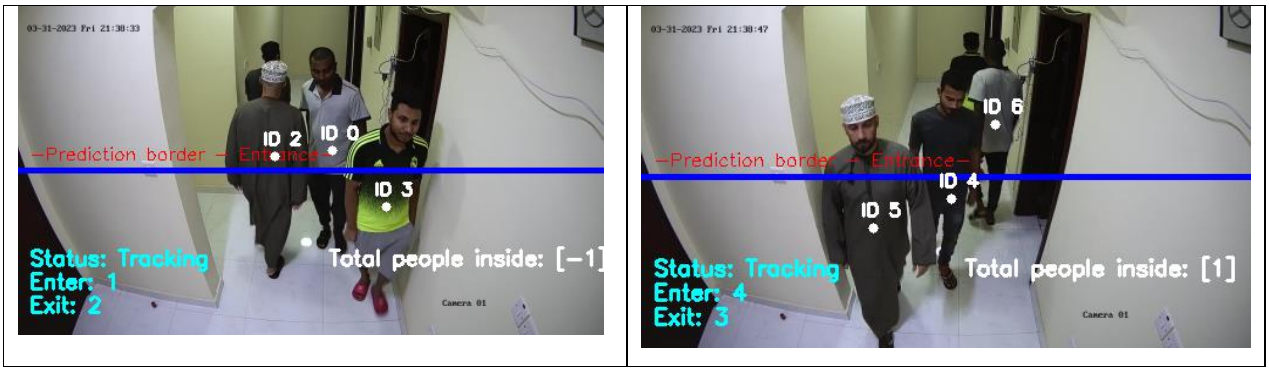


Figure 8. Two users in and two uses out counter

(e) Multiple-Two users at a time. In this experiment, multiple uses used for in and out at the same time with two users in the opposite way. Fig. 9 shows the results.

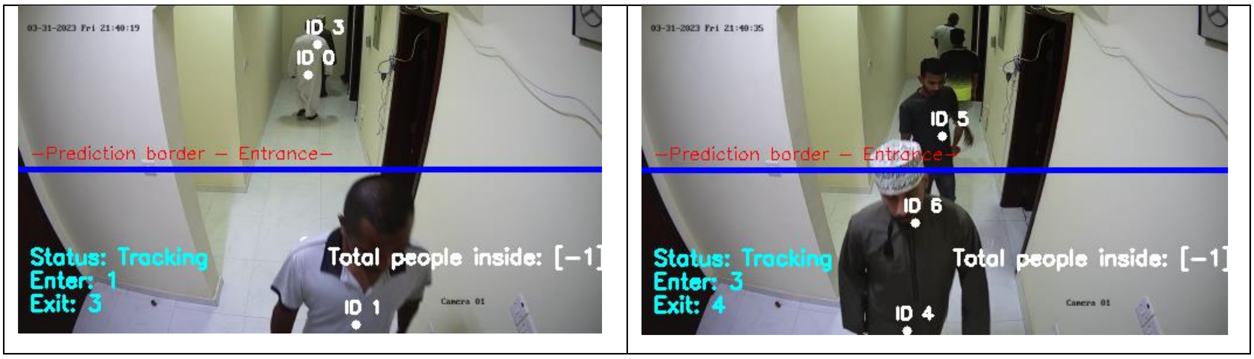


Figure 9. Multiple-Two users in/out counter

(f) Multiple-Multiple users at a time. In this experiment multiple users uses with a wider and longer walking pathway using the MOT20 dataset[14] for in and out at the same time. Fig. 10 shows the results.

Figure 10. Multiple users in/out counter

2024, 21(5 Special Issue): 1775-1785
https://doi.org/10.21123/bsj.2024.10540
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Results and Discussion

In machine learning, accuracy is one of the most important metrics used to assess a model's efficacy[14]. The accuracy rate is how well a model predicts for a specific dataset. Nevertheless, accuracy may not always give a whole view of a model's performance, especially when the dataset is unbalanced or when the cost of false positives and false negatives is different. Eq. 1 was used to calculate the accuracy of the model as given below:

$$Accuracy = \frac{\sum TP\ t + TNt}{Nt} \qquad 1$$

Where N is the total number of predictions; t is the current frame; TP is true Positive; TN is True negative; FP is false positive, and FN is false negative

As a result of different experimenters conducted in this paper, researchers found that the model is performing and resulting as shown in Table 1. The accuracy is calculated based on the number of humans appearing in the video and how the model is detected and calculated. Table 1 represents the accuracy of records in each experiment.

Table 1. Accuracy Result of Performance

Experiment	Result
Single user in/out	**100%**
One-One users in/out	100%
Two-One user in/out	**95%**
Two-Two users in/out	87.5%
Multiple-Two users	**90%**
Multiple-Multiple users (MOT20)	72.5%

The accuracy results for various user interaction scenarios show that the system achieves perfect accuracy (100%) when dealing with a Single user in/out or one-in and one-out interaction. As the complexity of user interactions increases, such as involving multiple users, the accuracy decreases, with 95% accuracy for two users, 87.5% accuracy for two users in a more complex scenario, and 90% accuracy for multiple users. In the most complex scenario, which involves multiple users, the accuracy drops to 72.5%. These results highlight the system's decreasing accuracy as the complexity of user interactions grows, emphasizing the importance of context and user-specific requirements when implementing or improving such systems.

Conclusion

In conclusion, this paper has tested different scenarios of gate entrance counting using pre-trained Mobilenet-SSD for real-time detection and tracking of the objects and found that the best accuracy performance among the different tested scenarios has been scored for a single user and one-one user scenarios, which scored 100%. Two-one users in both directions scored 95%. The two-two users scenario implementations recorded 87.5%. The Multiple-Two users scenario performs better than Two-Two users, which can detect users in different spaces and scored 90%, while the Multiple-Multiple scenario scored 72.5%. Therefore, the best option to implement is the one-one scenario, which can detect and count both sides at the same time and appears to have the highest accuracy.

Acknowledgment

The researchers would like to express their appreciation to Mr. Ali Al Wahaibi and his team for their contributions to the experimental design with full ethical acknowledgment to use and publish the recorded videos.

Authors' Declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore, any Figures and images, that are not ours, have been included with the necessary permission for re-publication, which is attached to the manuscript.
- The authors signed an animal welfare statement.
- Authors sign an ethical consideration's approval.

- Ethical Clearance: The project was approved by the local ethical committee at University of Technology and Applied Science, Oman.

Authors' Contribution Statement

N.A. and A.M.A. designed the study. and built the lab experiments, M.A performed lab experiments coding. A. performed proof reading. N. presented the paper in conference.

References

1. Ghosh S, Kumar A, Saha S. Highly Accurate Real time human Counter with Minimum Computation Cost. IEEE Xplore. 2021. p. 1–7. https://doi.org/10.1109/ICAECT49130.2021.9392459
2. Yusro MM, Ali R, Hitam MS. Comparison of Faster R-CNN and YOLOv5 for Overlapping Objects Recognition. Baghdad Sci J. 2022; 20(3). https://doi.org/10.21123/bsj.2022.7243
3. Raghunandan A, Mohana, Raghav P, Aradhya HVR. Object Detection Algorithms for Video Surveillance Applications. IEEE Xplore. 2018. p. 0563–8. https://doi.org/10.1109/ICCSP.2018.8524461
4. Marshall MR, Hellfeld D, Joshi THY, Salathe M, Bandstra MS, Bilton KJ, et al. 3-D Object Tracking in Panoramic Video and LiDAR for Radiological Source–Object Attribution and Improved Source Detection. IEEE Trans Nucl Sci. 2021 Feb; 68(2): 189–202. https://doi.org/10.1109/tns.2020.3047646
5. Womg A, Shafiee MJ, Li F, Chwyl B. Tiny SSD: A Tiny Single-Shot Detection Deep Convolutional Neural Network for Real-Time Embedded Object Detection. IEEE Xplore. 2018. p. 95–101. https://doi.org/10.1109/CRV.2018.00023
6. Tom AJ, George SN. Simultaneous Reconstruction and Moving Object Detection From Compressive Sampled Surveillance Videos. IIEEE Trans Image Process. 2020; 29(3): 7590–602. https://doi.org/10.1109/TIP.2020.3004696
7. Jadhav LH, Momin BF. Detection and identification of unattended/removed objects in video surveillance. IEEE Xplore. 2016. p. 1770–3. https://doi.org/10.1109/RTEICT.2016.7808138
8. Milan A, Leal-Taixé L, Reid ID, Roth S, Schindler K. MOT16: A Benchmark for Multi-Object Tracking. arXiv (Cornell University). 2016 Jan 1. arXiv:1603.00831. https://doi.org/10.48550/arXiv.1603.00831
9. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018 Jun. https://doi.org/10.48550/arXiv.1801.04381
10. Khasoggi B, Ermatita E, Samsuryadi S. Efficient mobilenet architecture as image recognition on mobile and embedded devices. Indones J Electr Eng Comput Sci . 2019 Oct 1; 16(1): 389. https://doi.org/10.11591/ijeecs.v16.i1.pp389-394
11. Howard A, Sandler M, Chen B, Wang W, Chen LC, Tan M, et al. Searching for MobileNetV3. IEEE Xplore. 2019. p. 1314–24. https://doi.org/10.1109/ICCV.2019.00140
12. Prasad K, El-Sharkawy M. Compressed MobileNet V3:A Light Weight Variant for Resource-Constrained Platforms. 2021 IEEE 11th Annual Computing and Communication Workshop and Conference, CCWC 2021. 2021 Jan 27. https://doi.org/10.1109/CCWC51732.2021.9376113
13. Choi HT, Lee HJ, Kang H, Yu S, Park HH. SSD-EMB: An Improved SSD Using Enhanced Feature Map Block for Object Detection. Sensors. 2021 Apr 17; 21(8): 2842. https://doi.org/10.3390/s21082842
14. Dendorfer P, Rezatofighi H, Milan A, Shi J, Cremers D, Reid I, et al. MOT20: A benchmark for multi object tracking in crowded scenes. arXiv. 2020. https://doi.org/10.48550/arXiv.2003.09003
15. Chu P, Wang J, You Q, Ling H, Liu Z. TransMOT: Spatial-Temporal Graph Transformer for Multiple Object Tracking. 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2023 Jan 1. https://doi.org/10.48550/arXiv.2104.00194
16. Wahab F, Ullah I, Shah A, Khan RA, Choi A, Anwar MS. Design and implementation of real-time object detection system based on single-shoot detector and OpenCV. Front psychol. 2022 Nov 2; 13: 1-17. https://doi.org/10.3389/fpsyg.2022.1039645
17. Sabri M A A. Ahmed, Khalifa OO. Vision-Based Detection and Tracking of Moving Target in Video Surveillance. IEEE. Conference: 2014 International Conference on Computer & Communication Engineering (ICCCE). 2014 Sep 1; 14(1). https://doi.org/10.1109/ICCCE.2014.18
18. Dong E, Zhang Y, Du S. An Automatic Object Detection and Tracking Method Based on Video Surveillance. IEEE International Conference on Mechatronics and Automation. 2020 Oct 13; 8(20). https://doi.org/10.1109/ICMA49215.2020.9233627.
19. Verma KK, Kumar P, Tomar A. Analysis of Moving Object Detection and Tracking in Video Surveillance System. 2nd International Conference on Computing for Sustainable Global Development . 2015; 8(15).

20. Jassim OA, Abed MJ, Saied ZH. Indoor/Outdoor Deep Learning Based Image Classification for Object Recognition Applications. Baghdad Sci J. 2023 Dec 5; 20(6(Suppl.)): 2540–0. https://doi.org/10.21123/bsj.2023.8177.
21. Lu W, Xu W, Wu Z, Xu Y, Wei Z. Video Object Detection based on Non-local Prior of Spatiotemporal Context. 8th International Conference on Advanced Cloud and Big Data. 2020 Dec 1.https://doi.org/10.1109/CBD51900.2020.00040
22. Sadkhan A, Talebiyan SR, Farzaneh N. An Investigate on Moving Object Tracking and Detection in Images. Babylon International Conference on Information Technology and Science 2021. 2021 Apr 28;6(21). https://doi.org/10.1109/BICITS51482.2021.9509887

تنفيذ عداد الزوار في الوقت الفعلي باستخدام كاميرات المراقبة والذكاء الصناعي : أفضل الممارسات

ناصر المصلحي[1]،علي محمد الوهيبي[2]،محمد عباس[3]

[1]قسم تقنية المعلومات، جامعة التقنية والعلوم التطبيقية فرع إبرا، إبرا، سلطنة عمان.
[2]كلية الاتصالات والحوسبة، جامعة سلنجور، سلنجور، ماليزيا.
[3]قسم تقنية المعلومات، جامعة التقنية والعلوم التطبيقية فرع صور، صور، سلطنة عمان.

الخلاصة

إن معرفة عدد الأشخاص في المباني والمنشآت يساهم بشكل كبير في المحافظة على سلامة الزائرين وكذلك توفير معلومات مفيدة لإدارة تلك المؤسسات. تهدف هذه الورقة الى استخدام الذكاء الصناعي والاستفادة من كاميرات المراقبة في معرفة عدد الأشخاص المتواجدين في مبنى ما. قام الباحثون ببناء قاعدة بيانات بسيناريوهات مختلفة لتنفيذ تجارب متنوعة لمعرفة افضل ممارسة ممكنة في استخدام هذا النوع من العدادات. ومن خلال تلك التجارب حقق الباحثون نتائج متميزة في السيناريوهات التي تتكون من شخص واحد او شخصين باتجاهين منفصلين حيث وصلة نسبة الدقة الى 100%.

الكلمات المفتاحية: الكشف بنظرة واحدة ، شبكة الموبايل ، مودل الكشف بنظرة واحدة في شكبات الموبايل ، كشف الكائنات في الوقت الفعلي، تتبع الكائنات.