# Real-time People Counting with Deep Learning: A Solution for Crowd Management

**Qudes MB aljelawy** [1]
qudesmb@kus.edu.iq

**Entessar Karim Hanoun**[2]
entessar.karim@kus.edu.iq

**Ghofran Mohammed Ali**
ghufran.m.ali@kus.edu.iq

**Abstract:** Crowd management plays a vital role in ensuring safety, efficiency, and resource optimization in environments such as public events, transportation hubs, shop-ping malls, and workplaces. Traditional people counting methods, including manual observations and sensor-based approaches, often suffer from limitations in accuracy, scalability, and adaptability to dynamic conditions. This research pre-sent a real-time people counting system utilizing deep learning techniques, specifically the YOLO (You Only Look Once) object detection framework. The proposed system processes video streams to accurately detect and count individuals, offering a robust and automated solution for crowd analysis. By leveraging YOLO's high-speed processing and accuracy, the system effectively identifies people within complex and varying environments, addressing challenges such as occlusions, lighting variations, and high-density crowds. Experimental evaluations demonstrate the system's strong performance, achieving an accuracy of 95% in static environments and 87% in dynamic conditions. These results highlight the model's reliability, efficiency, and potential applications in real-time crowd monitoring, security surveillance, and urban planning. Future improvements include integrating advanced tracking algorithms and multi-camera setups to enhance detection accuracy and consistency across different perspectives. This research un-derscores the transformative potential of deep learning in crowd management, paving the way for smarter, safer, and more responsive public spaces.

**Keywords:** Computer vision, YOLO, Deep learning, Python, open Cv

## 1. Introduction

Counting people using computer vision has become an increasingly important area of research, playing a significant role in reducing reliance on human labor and

[1] Assist. lect Al-Karkh university for science
[2] Assist. lect Al-Karkh university for science

physical resources. As urbanization continues to grow and public spaces become more crowded, the need for efficient and accurate people-counting systems has never been more critical. This technology has advanced significantly across various applications, including passenger counting on public transportation, statistical analysis of large gatherings, and monitoring attendance in classrooms.

Beyond these applications, crowd management is essential for ensuring safety and efficiency in densely populated environments such as airports, shopping malls, and stadiums. However, traditional methods of counting people—such as manual observations or infrared sensors often struggle with accuracy, scalability, and adaptability to dynamic environments. Manual counting is labor-intensive and prone to human error, while infrared sensors may not perform well in certain conditions or accurately track individuals in crowded settings.

With rapid advancements in computer vision and deep learning, automated solutions have emerged as more reliable alternatives for monitoring crowd dynamics. This research focuses on developing a real-time people-counting system using the YOLO (You Only Look Once) object detection framework, a cutting-edge approach in com-puter vision. YOLO is known for its high speed and accuracy, enabling the system to identify and count individuals in video streams efficiently. By processing video in real time, the YOLO framework can detect people with remarkable precision, allowing it to adapt seamlessly to different environments and lighting conditions.

Implementing such a system provides a strong foundation for data-driven decision-making in crowd management and resource allocation. With accurate counting data, organizations can optimize operations, enhance safety measures, and improve user experiences in public spaces. Furthermore, insights gained from analyzing crowd behavior can contribute to urban planning and infrastructure development, helping cities become smarter and more responsive to the needs of their inhabitants. As the field of automated people counting continues to evolve, the integration of advanced technologies like YOLO will undoubtedly play a pivotal role in shaping the future of crowd management and enhancing the efficiency of public services [4,5].

## 2. Related work

Recent advancements in deep learning have significantly improved real-time people counting and crowd analysis systems. The YOLO family of models has been at the forefront of these improvements, balancing detection accuracy and processing speed. Bochkovskiy et al. [2] introduced YOLOv4, which optimized both speed and accuracy, making it suitable for real-time object detection applications such as crowd monitoring. This work laid the groundwork for subsequent YOLO versions, including YOLOv8 used in our system.

In the domain of crowd analytics, Luo and Sun [8] explored deep learning techniques tailored for real-time crowd analysis, focusing on challenges such as occlusions, varying lighting conditions, and dynamic crowd behaviors. Their study highlights the importance of robust algorithms capable of adapting to environmental variability, a critical factor in deploying people counting systems in practical scenarios.

Additionally, Zhang et al. [9] proposed a multi-column convolutional neural network architecture for single-image crowd counting. Their approach addresses issues of crowd density variation and overlapping individuals by leveraging multiple convolutional pathways, enhancing accuracy in static crowd images. This complements real-time detection frameworks by improving performance in challenging visual conditions.

These works collectively underpin the foundation of real-time people counting using deep learn-ing and guide the development of efficient, accurate, and scalable crowd management solutions.

**Traditional Methods**

Historically, people counting systems relied on manual counting, turnstiles, or simple motion sensors. While effective in controlled settings, these methods struggle in complex, dynamic environments, often leading to inaccurate counts and limited scalability.

## 3. Modern Approaches

**3.1 computer vision:** With the advent of machine learning, automated detection and tracking of individuals has become feasible. Detection models like Faster R-CNN and SSD provide high accuracy but typically fall short in real-time applications due to processing delays [6].

**3.2 YOLO framework:** YOLO (You Only Look Once) offers a compelling balance between detection speed and accuracy. It has been successfully deployed in various domains, including traffic surveillance and retail analytics, making it a suitable can-didate for real-time crowd monitoring systems [1][7].

## 4. Challenges Addressed

This study tackles several challenges inherent to people counting in real-time video streams:

Real-time Processing: Ensuring minimal latency during detection

Crowd Dynamics: Handling variable crowd movement and behavior.
Environmental Variability: Adapting to lighting changes, occlusions, and camera angles.
The integration of the YOLO framework aims to overcome these hurdles with high reliability and low computational cost [8]

## 5. Methodology

### 5.1 System Architecture

The proposed system comprises two main components:
**1. Object Detection Module**: Based on the YOLOv8 model, this module detects people in video frames.
**2. Counting and Visualization Module**: Processes detection results to count individuals and visualize the output.

### 5.2 Implementation

This implementation utilizes the YOLOv8 model for real-time detection and counting of people in video streams. The system is designed to process frames from either pre-recorded videos or live camera feeds, detecting individuals with high accuracy and efficiency. Detection involves identifying objects in each frame and filtering them to detect people (class 0) based on a predefined confidence threshold [7–9]. Each detected person is highlighted with a bounding box, and the total count is displayed on the video frame using OpenCV. The count dynamically updates only when changes occur, improving performance and readability. This approach effectively integrates advanced object detection models with real-time video processing frameworks, providing a robust solution for crowd monitoring, security systems, and real-time analytics. The system is adaptable and can be extended to support additional object classes or functionalities such as tracking and behavior analysis.as shown the block diagram of the system in Figure (1).

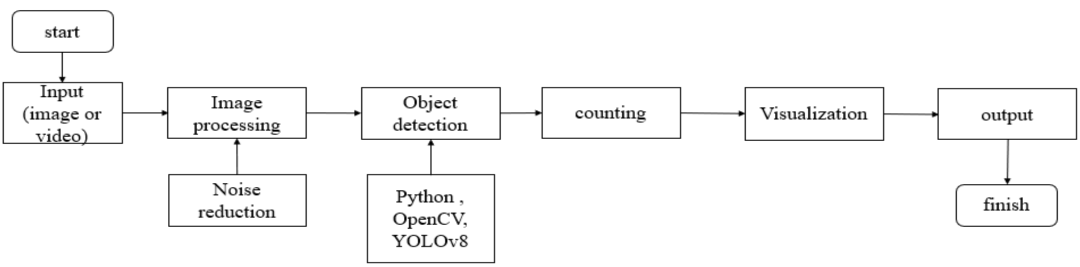

**Figure 1: block diagram of the system**

### 5.3 Tools and Technologies

- **YOLOv8**: Pre-trained on COCO dataset.
- **OpenCV**: For video processing and visualization.
- **Python**: Core programming language

## 6. Results

the system was tested in various scenarios, including: - Using static images with excel-lent accuracy as shown in Figure (2)

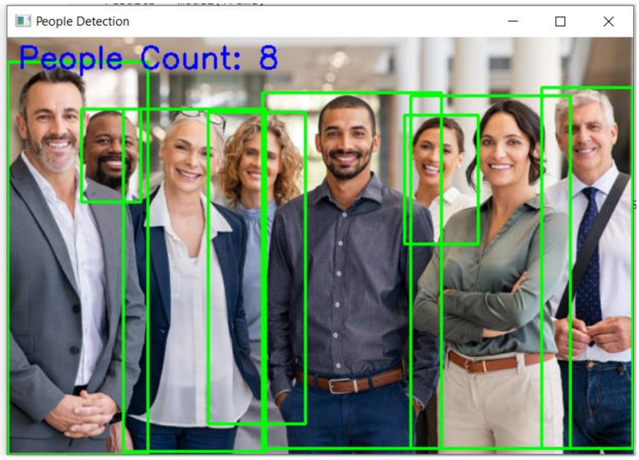


**Figure 2: the result when the system work with static image**

- Using video with only two people walking as shown in Figure (3)

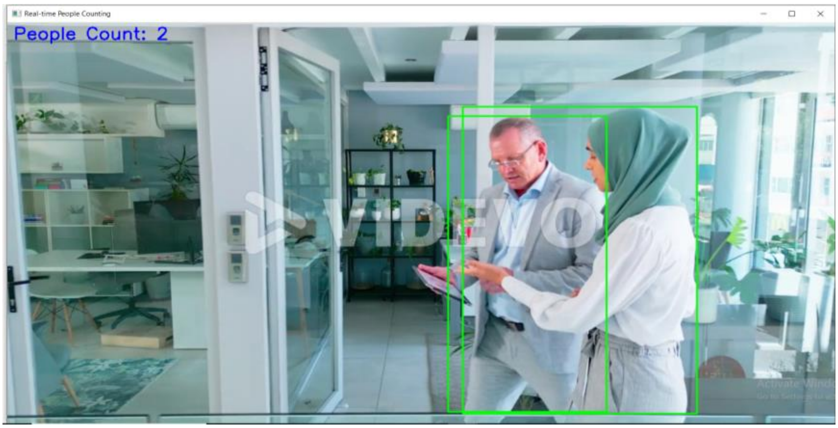

**Figure: 3 the result when the system work with video**

- Using video with more than five people as shown in Figure (4)

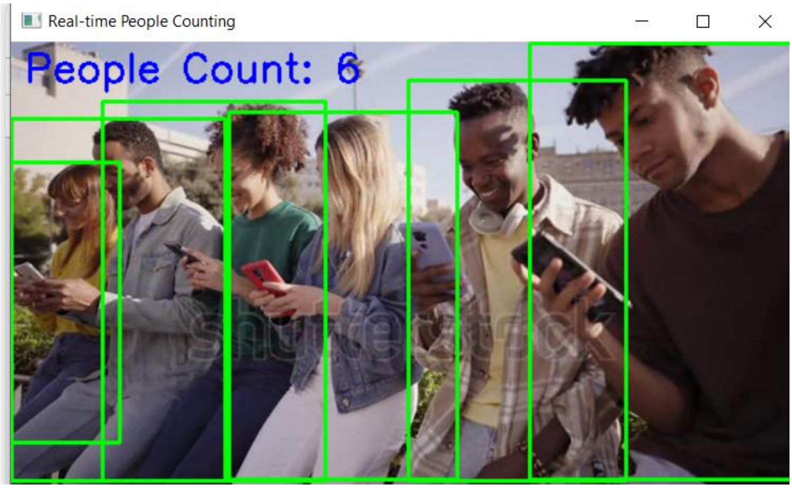


**Figure 4: the result when the system work with video contains more than 5 people**

**Key performance metrics:**
**Detection Accuracy**: 95% for static environments and 87% for dynamic scenes.
**Processing Speed**: Average of 30 frames per second on a GPU-enabled setup.

## 7. Conclusions

The results highlight the system's capability for real-time people counting with high accuracy. Static environments yield near-perfect results, while dynamic conditions pose challenges such as overlapping individuals and rapid movements.
Potential improvements include:
Integration of advanced tracking algorithms to handle occlusions and maintain consistent identities over time.
Enhanced preprocessing techniques to normalize lighting and reduce noise in input frames.
This study demonstrates the effectiveness of the YOLO-based object detection framework for real-time people counting in crowd management applications. The system achieves high accuracy, making it a reliable alternative to traditional methods. Utilizing deep learning and computer vision techniques, the system

provides an automated, efficient approach to monitoring crowd density while addressing com-mon challenges such as occlusions, lighting variations, and movement complexities. The ability to process video feeds in real time enhances its applicability in security surveillance, urban planning, and resource management. Future improvements such as tracking integration and multi-camera setups could enhance accuracy and consistency in dynamic environments.

In conclusion, this research highlights deep learning's transformative potential in crowd management, paving the way for smarter, safer, and more efficient public spaces

## 8. References


[1] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement.

[2] Bochkovskiy, A., Wang, C., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection.

[3] Zhang, Z., & Tao, D. (2019). Efficient Human Detection in Crowded Scenes. IEEE Transactions on Image Processing, 28(4), 1748–1760.

[4] OpenCV Documentation. (2023). Retrieved from https://docs.opencv.org/

[5] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context.

[6] COCO Dataset. (2023). Retrieved from https://cocodataset.org/

[7] Jocher, G. et al. (2023). YOLO by Ultralytics. Retrieved from https://ultralytics.com/

[8] Luo, Z., & Sun, X. (2021). Real-time Crowd Analytics using Deep Learning. Journal of Computer Vision Applications, 39(2), 102-115.

[9] Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y. (2016). Single-Image Crowd Counting via Multi-Column,Convolutional Neural Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 589-597

# العدّ الآني للأشخاص باستخدام التعلم العميق
# حل لإدارة الحشود

**غفران محمد علي** [3]
ghufran.m.ali@kus.edu.iq

**م.م انتصار كريم حنون** [2]
entessar.karim@kus.edu.iq

**م.م قدس محمد باقر** [1]
qudesmb@kus.edu.iq

**المستخلص:** تلعب إدارة الحشود دورًا حيويًا في ضمان السلامة والكفاءة وتحسين استخدام الموارد في البيئات التي تشمل الفعاليات العامة، ومراكز النقل، والمجمعات التجارية، وأماكن العمل. وتعاني أساليب عدّ الأشخاص التقليدية، مثل الملاحظة اليدوية والأنظمة المعتمدة على المستشعرات، من محدودية في الدقة، وقابلية التوسع، والتكيف مع الظروف المتغيرة. يقدم هذا البحث نظامًا آنيًا لعدّ الأشخاص يعتمد على تقنيات التعلم العميق، وتحديدًا إطار الكشف عن الأجسام(YOLO (You Only Look Once

يعالج النظام المقترح تدفقات الفيديو لاكتشاف الأفراد وعدّهم بدقة، موفرًا بذلك حلاً آليًا وفعّالًا لتحليل الحشود. ومن خلال الاستفادة من سرعة YOLO العالية ودقته، ينجح النظام في التعرف على الأشخاص حتى في البيئات المعقدة والمتغيرة، متغلبًا على تحديات مثل التداخل بين الأجسام، وتغير الإضاءة، وارتفاع كثافة الحشود. وتُظهر التقييمات التجريبية أداءً قويًا للنظام، حيث حقق دقة بلغت ٩٥٪ في البيئات الثابتة و٨٧٪ في البيئات الديناميكية.

تسلط هذه النتائج الضوء على موثوقية النموذج وكفاءته وإمكانياته التطبيقية في مجالات مراقبة الحشود الآنية، والمراقبة الأمنية، والتخطيط الحضري. وتشمل التحسينات المستقبلية دمج خوارزميات تتبع متقدمة واستخدام أنظمة متعددة الكاميرات لتعزيز دقة الكشف وتناسقه عبر زوايا الرؤية المختلفة. ويؤكد هذا البحث على الإمكانات التحويلية للتعلم العميق في مجال إدارة الحشود، مما يمهّد الطريق نحو فضاءات عامة أكثر ذكاءً وأمانًا واستجابة.

**الكلمات المفتاحية : الرؤية الحاسوبية، YOLO، التعلم العميق، بايثون، OpenCV**

[1] مدرس مساعد -جامعة الكرخ للعلوم – بغداد - العراق
[2] مدرس مساعد – جامعة الكرخ للعلوم – بغداد - العراق
[3] مبرمج اقدم – جامعة الكرخ للعلوم- بغداد - العراق