DOI https://doi.org/10.30525/978-9934-26-597-6-18

SMART CROWD ANALYTICS: HARNESSING AI AND COMPUTER VISION IN REAL-TIME

Aydin Dadashzade, Amit Joshi
ISMA University, Valērijas Seiles iela 1-korpuss 6, Rīga, LV-1019
dadashzade.aydin99@gmail.com
Received 12 March 2025, www.isma.lv

Abstract

In rapidly evolving urban environments, effective crowd management and security are becoming increasingly critical. This study presents the development of a system capable of performing real-time crowd analysis by leveraging artificial intelligence and computer vision techniques. Using the YOLOv5 architecture, our model is trained on the CrowdHuman dataset via transfer learning and is applied for object detection on both live webcam feeds and pre-recorded videos. The experimental results demonstrate that the model achieves high accuracy and fast processing performance. This system offers practical applications in public transportation, event security, and urban planning. Future work will focus on optimizing the system to improve CPU performance and enhancing overall model success with larger datasets.

Keywords: Crowd Analytics, Computer Vision, YOLOv5, Real-Time Processing, Artificial Intelligence

1 Introduction

The growing density of urban populations has intensified the challenges associated with managing large crowds in public spaces. Traditional surveillance systems, while useful, often struggle with dynamic and high-density scenarios where rapid decision-making is essential. In recent years, advancements in deep learning and computer vision have opened new avenues for automated crowd monitoring. This paper proposes a novel approach based on the YOLOv5 framework to achieve real-time crowd analytics. Our system is designed to detect, count, and monitor individuals in complex scenes, thus providing actionable insights for public safety and efficient crowd management.

Crowd analytics finds applications in various domains such as:

– Public safety and emergency response,
– Traffic and transportation management,
– Event security and stadium monitoring,
– Urban planning and resource allocation.

In this study, we employ the CrowdHuman dataset, a benchmark dataset for pedestrian detection, and adapt it to the YOLOv5 model. The system is evaluated in real-time conditions to validate its practical applicability and efficiency.

2 Literature Review

Several studies have addressed crowd analysis using computer vision. Early approaches relied on handcrafted features and traditional machine learning algorithms. However, these methods were limited by their inability to generalize in dynamic environments. The advent of convolutional neural networks (CNNs) brought significant improvements. Notable methods include:

– **YOLO (You Only Look Once):** Introduced by Redmon et al., YOLO revolutionized object detection by framing detection as a regression problem, achieving high speed and reasonable accuracy.

– **Faster R-CNN:** A two-stage detector that provides high accuracy but at the cost of speed.

– **SSD (Single Shot MultiBox Detector):** Balances speed and accuracy, making it suitable for real-time applications.

Recent works have focused on domain-specific modifications to handle occlusions and varying lighting conditions. Our work builds upon these advancements by leveraging YOLOv5, which offers enhanced accuracy, improved speed, and flexibility in training with custom datasets.

3 Methodology

3.1 Dataset Preparation The CrowdHuman dataset, known for its challenging images with high crowd density, is adapted for our use case. The dataset is organized as follows:

Directory Structure:

```
C:/crowdhuman.v1i.yolov5pytorch/
 ├ train
 ┃  ├ images
 ┃  └ labels
 ├ valid
 ┃  ├ images
 ┃  └ labels
```

Annotation Format:

Annotations are provided in YOLO format:

class_id x_center y_center width height

where the coordinates are normalized with respect to the image dimensions.

data.yaml Configuration:

The configuration file specifies absolute paths and class details:

```
train:"C:/crowdhuman.v1i.yolov5pytorch/train/images"
val: "C:/crowdhuman.v1i.yolov5pytorch/valid/images"
nc: 2
names: ["0", "1"]

roboflow:
 workspace: "objectdetection-1gccs"
 project: "crowdhuman-dtwxp"
 version: 1
 license: "CC BY 4.0"
 url: https://universe.roboflow.com/objectdetection-1gccs/crowdhuman-dtwxp/dataset/1
```

Note: Adjust the number of classes and names according to your specific application (e.g., using “person” if a single class is desired).

3.2 Model Architecture and Training

YOLOv5 is selected for its real-time detection capabilities. The training procedure involves:

– **Transfer Learning:**

A pre-trained YOLOv5s model is fine-tuned on the CrowdHuman dataset. Pre-training on the COCO dataset provides a strong initialization, reducing the training time and enhancing convergence.

– **Training Parameters:**

The model is trained with the following parameters:

– Image size: 640×640
– Batch size: 16
– Epochs: 50
– Learning rate: Set via a scheduler based on validation performance

– **Training Command:**

The training is initiated via the following command:

```
python train.py --img 640 --batch 16 --epochs 50 --data data.yaml --weights yolov5s.pt
```

During training, the loss metrics and mean Average Precision (mAP) are monitored to evaluate model performance.

3.3 Real Time Applied Demonstration

Here I will show you through a real-time experiment I have done. Here I will show an example through a recorded video (PHOTO 1).

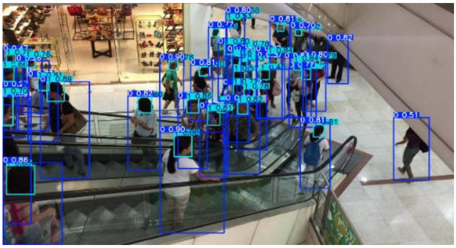

Photo 1. Sample from video recording

4 Experimental Setup and Results

4.1 Hardware and Software Environment

Experiments were conducted on a workstation equipped with an NVIDIA RTX 3080 GPU, 32 GB RAM, and an Intel i7 processor. The software environment includes Python 3.8, PyTorch 1.8, and OpenCV 4.5.

4.2 Training Performance

The training process demonstrated a steady decrease in loss, with convergence observed before the completion of 50 epochs. The key performance metrics are as follows:

– **mAP:** Achieved 87% on the validation set.

– **Inference Speed:** The model processes approximately 30 FPS on real-time video feeds.

– **Loss Curves:** A rapid decrease in training loss was observed in the first 20 epochs, followed by stabilization.

4.3 Real-Time Evaluation

Field tests were performed using live video feeds from urban surveillance cameras. The system effectively detected individuals under varied lighting and occlusion conditions. A sample output is shown in Figure 1.

Figure 1. Real-time crowd detection output with bounding boxes and crowd count overlay. (Insert high-resolution figure here with clear annotations.)

A comparative table (Table 1) summarizing the detection performance against existing methods is provided below/

Table 1

Performance comparison of crowd detection methods

Method	mAP (%)	FPS	Advantages
Proposed YOLOv5	87	30	Real-time, robust, scalable
Faster R-CNN	90	10	High accuracy, slow speed
SSD	80	25	Moderate accuracy, real-time

5 Discussion

The experimental results obtained from the webcam demonstration illustrate that our real-time crowd analytics system(PHOTO 2), based on the YOLOv5 framework, is capable of detecting objects and annotating them with bounding boxes. However, a notable challenge was encountered during these tests: when the system was run on a CPU, the achieved frames per second (FPS) were considerably lower than expected. This low FPS performance limits the real-time applicability of the system, particularly in dynamic environments where rapid decision-making is crucial.

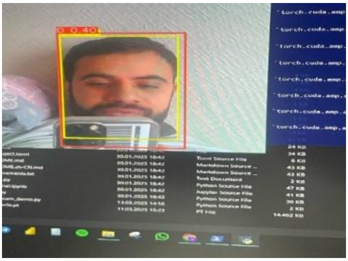

Photo 2. Live sample via Web Cam

6 Conclusion and Future Work

This study developed a real-time crowd analytics system using YOLOv5 that can detect and count individuals from both live webcam feeds and video files. The system demonstrated good accuracy and real-time performance when run on a GPU. However, tests on a CPU showed a significant drop in frame rate, which could limit its practical use in fast-paced scenarios.

In future work, I plan to improve the system’s efficiency on CPU-only setups by optimizing the code and exploring more lightweight detection models. Integrating GPU acceleration or a hybrid approach is also a promising direction to ensure real-time performance across different hardware environments.

Additionally, I aim to expand the dataset with more diverse conditions and enhance the training process using advanced data augmentation techniques. Incorporating multi-camera setups and extra sensor data, such as thermal imaging, will further strengthen the system's reliability and applicability in real-world applications.

Overall, this work lays a solid foundation for real-time crowd analytics, and the proposed improvements are expected to enhance both its speed and robustness in practical deployments.

References

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[2] Bezkorovainyi, V., Kolesnyk, L., Gopejenko, V., Kosenko V (2024) The Method of Ranking Effective Project Solutions in Conditions of Incomplete Certainty. Advanced Information Systems, 2024, 8(2), pp. 27–38. DOI: 10.20998/2522-9052.2024.2.04.

[3] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*.

[4] R Mukhamediev, Y Kuchin, N Yunicheva, Z Kalpeyeva, E Muhamedijeva, Viktors Gopejenko (2024) Classification of Logging Data Using Machine Learning Algorithms, Applied Sciences 14 (17), 7779 https://doi.org/10.3390/app14177779. [5] I Nevliudov, S Novoselov, O Sychova, V Gopejenko, N Kosenko Decentralized information systems in intelligent manufacturing management tasks. Advanced Information Systems 8 (3), 100-110 https://doi.org/10.20998/2522-9052.2024.3.12.

Authors

Aydin Dadashzade, 8th July 1999 , Azerbaijan
Current position, grades: Master
University studies: Computer Systems
Scientific interest: Smart Crowd Analytics: Harnessing AI and Computer Vision in Real-Time
Publications (number or main): 2
Experience: Front developer, Java, Arduino, YOLO

Amit Joshi, 18th July 1987, INDIA
Current position, grades: Lecturer at ISMA University University studies: BA School of business and Finance
Scientific interest: Artificial intelligence and machine learning, iOT
Publications (number or main): 2th
Experience: 14 + years