Convolutional Networks for Detecting and Counting People

Trần Thị Minh Nguyệt
Thai Nguyen University of Information and Communication Technology
Thai Nguyen University, Thai Nguyen, Vietnam
student576@ictu.edu.vn

***Abstract*—This undergraduate thesis investigates convolutional-network methods for detecting and counting people in images or video. Automated people counting requires object detection that remains useful across changing scenes and limited computing resources. The student reviews OpenCV, machine learning, deep learning, and YOLOv8, then builds a detection and counting demonstration. The thesis discusses training, sample predictions, the interface, and the constraints of running on lower-specification hardware. The conclusion reports a working research demonstration and describes relatively fast YOLOv8 training and useful detection in the selected examples. It also says the dataset is limited, prediction on a CPU is slow, and the system is not ready for practical deployment. The source does not establish general accuracy across difficult environments or a validated real-time field system. The author's qualitative performance assessment is not treated as an independent benchmark. The contribution is a documented student study or prototype whose stated scope can guide further testing.**

***Index Terms*—people counting, YOLOv8, OpenCV, object detection**

I. Introduction

Automated people counting requires object detection that remains useful across changing scenes and limited computing resources. This thesis investigates convolutional-network methods for detecting and counting people in images or video. It begins from a practical engineering need and explains the technologies and workflow chosen for the case. The objective is to document a student implementation or comparative investigation, together with the evidence available in the manuscript. Broader operational or economic effects require separate measurements and are not assumed from the existence of a prototype [1]. Earlier ICTU student work on related designs is included here for context [2], [3]. Related published studies provide context for the subject and methods [4], [5], [6], [7], [8], [9], [10]. These contextual works were selected for this English summary; they are not asserted to appear in the original thesis bibliography.

II. Methods

The student reviews OpenCV, machine learning, deep learning, and YOLOv8, then builds a detection and counting demonstration. The thesis discusses training, sample predictions, the interface, and the constraints of running on lower-specification hardware. The source is summarized through its stated design, implementation material, examples, and conclusion. Features shown in the manuscript are distinguished from proposals for later development. This review does not add a new experiment, fill missing measurements with estimates, or treat a planned capability as an observed result. The source thesis remains the provenance for the reported work [1].

Supervisor: Trần Quang Quý, Thai Nguyen University of Information and Communication Technology (tqquy@ictu.edu.vn).

III. Results

The conclusion reports a working research demonstration and describes relatively fast YOLOv8 training and useful detection in the selected examples. It also says the dataset is limited, prediction on a CPU is slow, and the system is not ready for practical deployment. The report's screens, configurations, examples, or measurements constitute the evidence for this summary. A capability described only as a design intention is not counted as a validated operational outcome. No additional performance, adoption, financial, or safety figure has been introduced here. The result is interpreted within the scope of the student's documented setup [1].

IV. Discussion

The source does not establish general accuracy across difficult environments or a validated real-time field system. The author's qualitative performance assessment is not treated as an independent benchmark. Further work could collect varied real-world data, report precision and recall on a held-out set, and measure inference time on target devices. These limitations affect how the evidence should be read. The current manuscript can support understanding of the design and reported demonstration, while independent tests would be needed for stronger claims. Any later deployment should document its environment and separate observed outcomes from intended benefits [1].

V. Conclusion

The thesis documents convolutional-network methods for detecting and counting people in images or video. The conclusion reports a working research demonstration and describes relatively fast YOLOv8 training and useful detection in the selected examples. It also says the dataset is limited, prediction on a CPU is slow, and the system is not ready for practical deployment. The source does not establish general accuracy across difficult environments or a validated real-time field system. The author's qualitative performance assessment is

not treated as an independent benchmark. Further work could collect varied real-world data, report precision and recall on a held-out set, and measure inference time on target devices. Accordingly, the present contribution is a source-grounded account of the student work and its tested or illustrated scope. Claims about broader effectiveness should be assessed in later work using explicit evaluation conditions [1].

References

[1] Trần Thị Minh Nguyệt, ”Convolutional Networks for Detecting and Counting People,” internal undergraduate thesis (original in Vietnamese), Thai Nguyen University of Information and Communication Technology, 2024.

[2] Nguyễn Minh Việt, ”System for Mask-Wearing Alert to Prevent COVID-19,” internal undergraduate thesis, Thai Nguyen University of Information and Communication Technology, 2021.

[3] Nguyễn Minh Tuấn, ”Construction of a Face Recognition Program for Family Robots,” internal undergraduate thesis, Thai Nguyen University of Information and Communication Technology, 2021.

[4] Maksat Kanatov et al., ”Deep Convolutional Neural Network based Person Detection and People Counting System,” Advanced Engineering Technology and Application, 2018. doi:10.18576/aeta/070301.

[5] Liping Zhu et al., ”Crowd counting via Multi-Scale Adversarial Convolutional Neural Networks,” Journal of Intelligent Systems, 2020. doi:10.1515/jisys-2019-0157.

[6] Bowen Zhang et al., ”Detecting and Counting People without Mask with Deep Neural Network,” Journal of Signal Processing, 2021. doi:10.2299/jsp.25.251.

[7] Khouloud Ben Ali Hassen et al., ”Convolutional Neural Networks and Heuristic Methods for Crowd Counting: A Systematic Review,” Sensors, 2022. doi:10.3390/s22145286.

[8] Yaocong Hu et al., ”Dense crowd counting from still images with convolutional neural networks,” Journal of Visual Communication and Image Representation, 2016. doi:10.1016/j.jvcir.2016.03.021.

[9] G.Thomas Prathiba, ”Literature survey for people counting and human detection,” IOSR Journal of Engineering, 2013. doi:10.9790/3021-03110510.

[10] Jia Hong Chee et al., ”Development of Automated People Counting System using Object Detection and Tracking,” International Journal of Online and Biomedical Engineering (iJOE), 2023. doi:10.3991/ijoe.v19i06.38515.