

МЕТОД ДЕТЕКТУВАННЯ ЛЮДЕЙ У НАТОВПАХ РІЗНОЇ ЩІЛЬНОСТІ

Гречишкін Д.С.

Науковий керівник – к.т.н., доц. Яковлева О.В.

Харківський національний університет радіоелектроніки, каф.ІНФ
м. Харків, Україна

e-mail: danylo.hrechyshkin@nure.ua

This work devoted for the analysis of the modern state-of-the-art solutions in crowd counting. Highlighted tendencies and widely used tools of this area. Special attention devoted to the crowd detection with the different density. Methods based on the Convolution Neural Networks analyzed and compared. Provided additional analyze for classic local binary pattern detection methods and brand-new state-of-the-art model YOLOv8, especially for low-dense crowd counting. Various methods compared based on their prospers and concerns.

Прогрес останніх років в вирішенні задач комп'ютерного зору пов'язаний з розвитком архітектури нейронних мереж та покращенням технічних можливостей для їх навчання. Різні моделі можуть бути навчені для розпізнавання, класифікацію, сегментацію, трекінг та інше. Розробники змагаються в проектуванні найкращих та найефективніших моделей, і порівнюють їх між собою.

Вирішення таких задач є актуальним для багатьох сфер життя починаючи від соціальних, закінчуючи стратегічно важливими [1–3]. Аналіз кількості людей в натовпі, наприклад, міг би мати, як соціальне значення, так і економічне. З одного боку продемонстровано інтерес суспільства до певного питання, з іншого – планування можливих заробітків відносно кількості учасників. Навіть такий простий приклад, показує що аналіз зображень може бути вкрай важливим для окремих сфер та може мати свої певні обмеження та специфічні умови.

Дана робота присвячена питанню підрахунку людей у натовпі. Для чого можна використовувати різні підходи, наприклад, на основі аналізу текстур, малювання теплових мап або на основі детектування та класифікації постатей людей. Для того щоб обробляти зображення з натовпом різного типу, або зображення с поодинокими посталями людей однаково добре, треба використовувати додаткову систему класифікації зображень на різні види натовпу, та підготувати моделі, які найкраще підраховують людей для цих умов. Тому в роботі пропонується на першій стадії визначити до якого типу відноситься натовп на зображенні: щільного натовпу, помірного натовпу, або окремих постатей. Для цього була перенавчена модель MobileNetV2 (рис. 1).

Щільний натовп та помірний натовп може сприйматися як текстура, що має певні характеристики, які можна проаналізувати і визначитися із

кількістю людей. Для аналізу текстурних зображень можна використати, наприклад, метод LBP (Local Binary Patterns), або обробляти теплові мапи, що побудують нейроні мережі CSRNet [4], MCNN. В роботі проводилося навчання мереж CSRNet, MCNN на датасеті ShanghaiTech, який складається з двох частин: part_A – щільний натовп, part_B – помірний. Навчання проводилося окремо на part_A, окремо на part_B. Проведені дослідження показали, що підрахунок людей як в щільному натовпі, так й помірному, краще веде модель CSRNet (рис. 2).



Рисунок 1 – Класифікація натовпу

Models	Part A		Part B	
	MAE	MSE	MAE	MSE
MCNN	110,2	173,2	26,4	41,3
CSRNet	68,2	131,3	10,6	16,2

Рисунок 2 – Навчання мереж CSRNet та MCNN

Для підрахунку людей на зображеннях, де присутні окремі постаті було вирішено застосувати методи детекції. До зображень з окремими посталями, наприклад, можна віднести групові фото, або фото людей, що знаходяться в офісі на робочих місцях, в кав'ярні. Якщо потрібно було б рахувати людей тільки за їхніми обличчями, можна було б використовувати метод на основі LBP, каскадів Наар та Віоли-Джонса, або мережі RetinaFace. Але в задачі підрахунку людей необхідно детектувати постаті людей та об-

личчя, які можуть бути не обов'язково фронтальні. Для цього можна використати сучасні потужні нейронні мережі для детектування та класифікації об'єктів, наприклад, YOLOv8 (рис. 3) [5].

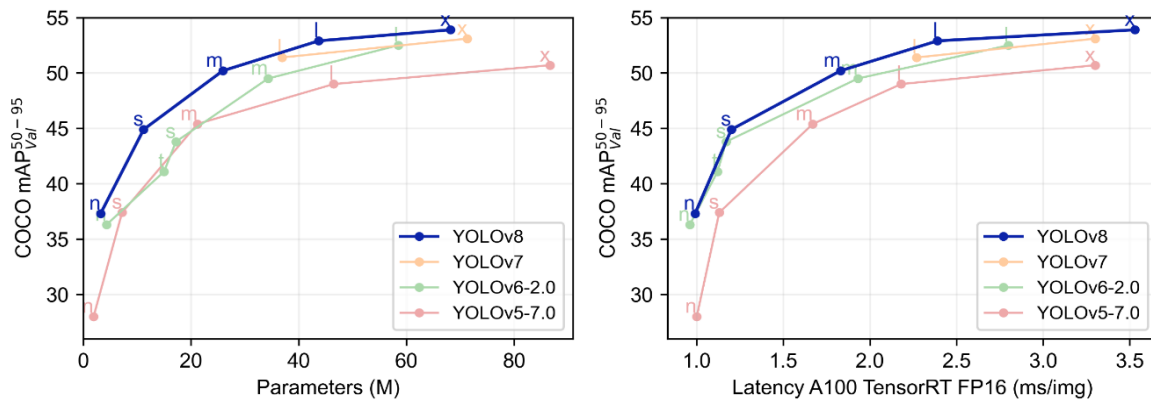


Рисунок 3 – Ефективність YOLOv8 [5]

Таким чином, в роботі досліджено питання підрахунку людей у натовпі, розглянуто різні підходи – від методів, заснованих на аналізі текстур і теплових мап, до методів детектування силуетів. Використання класифікації на першому етапі дозволило використовувати найбільш підходящі методи для натовпу різного типу: для щільного та помірного натовпу – мережу CSRNet, навчену на різних частинах датасету ShanghaiTech; для окремих постатей – модель YOLOv8. У подальшому доцільно було б дослідити використання запропонованого підходу щодо підрахунку людей у відео потоці.

Список використаних джерел:

1. Yakovleva, O., Kovtunenکو, A., Liubchenko, V., Honcharenko, V., & Kobylin, O. (2023). Face Detection for Video Surveillance-based Security System (COLINS-2023). In CEUR Workshop Proceedings (Vol. 3403). pp. 69–86.
2. Gorokhovatskyi V., Tvoroshenko I., and Yakovleva O. (2024) Transforming image descriptions as a set of descriptors to construct classification features, Indonesian Journal of Electrical Engineering and Computer Science, vol. 33, no. 1, pp. 113–125.
3. Ковтуненко, А. Р., Яковлева, О. В., Любченко, В. А., & Янголенко, О. В. (2020). Дослідження сумісного використання математичної морфології та згорткових нейронних мереж для вирішення задачі розпізнавання цінників.
4. Li, Y., Zhang, X., & Chen, D. (2018). CSRNET: Dilated convolutional neural networks for understanding the highly congested scenes. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1091–1100).
5. Ultralytics YOLOv8 is out. (2024). Repository of the project. Retrieved from <https://github.com/ultralytics/ultralytics>.