**Кромкач Владислав Олександрович**

Аспірант, спеціальність 123 «Комп’ютерна інженерія»

Національний університет «Львівська політехніка», м.Львів

**ОГЛЯД МЕТОДІВ АНАЛІЗУ СТАТИЧНИХ ЗОРОВИХ СЦЕН**

Візуальне відстеження є однією з найбільш класичних проблем комп’ютерного зору, і його історія сягає корінням у початок 80-х років, коли були розроблені класичні концепції, такі як трекер Лукаса-Канаде та узгоджені фільтри [1].

Відстеження об'єктів - відповідальне і складне завдання. Таким чином, відстеження об'єктів стало популярною темою вивчення в останні роки. Існує багато застосувань для візуального відстеження об'єктів, таких як візуальне спостереження [2], відеорозуміння [3], робототехніка [4] та інші.

Статичні сцени характеризуються високим рівнем структурної однорідності, де основні об'єкти часто взаємодіють з навколишнім середовищем сталого чинника. Досягнення високої точності та ефективності в аналізі статичних сцен важливо для широкого кола застосувань, включаючи медичні дослідження, автоматизовані системи моніторингу та технології штучного інтелекту.

Для вирішення завдань виділення об'єктів на статичних зорових сценах може бути застосований підхід на основі «ковзного вікна» спільно з шаблонами [5], дескрипторами локальних особливостей, такими як HOG, LBP, SIFT, SURF [6], колірними ознаками, методами контурного аналізу. Основними недоліками таких методів, є необхідність перебору досить великої кількості областей, необхідних для виділення об'єктів, а також специфічність ознак для «ковзного вікна».

Інша група методів заснована на сегментуванні зображень [7]. Такі методи є евристично ненавченими методами і не залежать від специфіки розв'язуваного завдання, крім того, вони вимагають значних обчислень і навіть із використанням сучасних обчислювальних засобів не дають змогу реалізувати режим реального часу.

Методи генерації гіпотез про розташування об'єктів на зображенні, засновані на нейромережевому підході [8], дозволяють усунути зазначені недоліки.

Для вирішення завдань розпізнавання об'єктів на статичних зорових сценах найчастіше використовуються методи, засновані на застосуванні математичної статистики та машинного навчання. Серед методів машинного навчання досить добре себе зарекомендували штучні нейронні мережі, у тому числі, згорткові нейронні мережі [9].

Такі нейронні мережі характеризуються значним збільшенням точності розпізнавання, порівняно з класичними методами. Застосування «глибоких» згорткових нейронних мереж дозволило зменшити середню помилку розпізнавання приблизно в півтора рази в порівнянні з одним із кращих методів, що вирішує задачу класифікації зображень із використанням векторів Фішера та SIFT, запропонованої в [10].

Отже, для вирішення завдань детектування об'єктів на статичних зорових сценах доцільним є спільне використання методів виділення та розпізнавання об'єктів. Однак для зменшення обчислювальної складності ефективніше використовувати методи, що вирішують обидва ці завдання одночасно. Такі методи (наприклад, Faster R-CNN [11]) є досить точними, але, як правило, не дозволяють детектувати об'єкти на зорових сценах в режимі реального часу. Інші методи (наприклад, YOLO [12]) працюють у режимі реального часу, але не забезпечують необхідну точність. Ще одна група методів (наприклад, SSD [13]) забезпечує компроміс між точністю та швидкістю виділення та розпізнавання об'єктів. Така модель виконує детектування зображень у різних масштабах з використанням кількох детекторів.

Таким чином, найбільш перспективними методами аналізу статичних зорових сцен на сьогоднішній день є методи, засновані на «глибокому» навчанні та згорткових нейронних мережах, оскільки саме вони дозволяють отримати високу точність розпізнавання, порівняно з іншими методами.

**Література**

1. Felsberg M. Visual tracking: Tracking in scenes containing multiple moving objects. Advanced Methods and Deep Learning in Computer Vision. Elsevier, 2022. P. 305—336
2. Xing J, Ai H, Lao S. Multiple Human Tracking Based on Multi-view Upper-Body Detection and Discriminative Learning. In: 2010 20th International Conference on Pattern Recognition. IEEE; 2010.
3. Renoust B, Le DD, Satoh S. Visual Analytics of Political Networks From Face-Tracking of News Video. IEEE Transactions on Multimedia. 2016;18(11):2184–2195.
4. L Liu HA J Xing, Ruan X. Hand posture recognition using finger geometric feature. IEEE. 2012.
5. Brunelli, R. Template Matching Techniques in Computer Vision: Theory and Practice // Wiley. – 2009.
6. Ahonen T., Hadid A., Pietikainen M. Face Recognition with Local Binary Patterns // Proc. 8th European Conference on Computer Vision (ECCV). – 2004. P. 469–481.
7. Endres I., Hoiem D. Category Independent Object Proposals // Proc. 11th European Conference on Computer Vision (ECCV). – 2010. P. 575–588.
8. Erhan D. Scalable Object Detection using Deep Neural Networks // Computer Vision and Pattern Recognition. – 2014. P. 2155-2162.
9. LeCun Y., Boser B., Denker J.S. Handwritten Digit Recognition with a BackPropagation Network // Proceedings of the Neural Information Processing Systems conference, NIPS. – 1989. P. 396–404.
10. Sánchez J., Perronnin F. High-dimensional signature compression for largescale image classification // Computer Vision and Pattern Recognition, CVPR. – 2011. – P. 1665-1672.
11. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks //Proceedings of the Neural Information Processing Systems conference, NIPS. – 2015.
12. Redmon, J., Divvala, S., Girshick, R., Farhadi, A. You only look once: Unified, real-time object detection //Computer Vision and Pattern Recognition, CVPR, – 2016.
13. . Liu W., Anguelov D., Erhan D., Szegedy C., Reed S. E. SSD: single shot multibox detector // CoRR. – 2015.