**Кромкач Владислав Олександрович**

Аспірант, спеціальність 123 «Комп’ютерна інженерія»

Національний університет «Львівська політехніка», м.Львів

**ДОСЛІДЖЕННЯ МЕТОДІВ АНАЛІЗУ ДИНАМІЧНИХ ЗОРОВИХ СЦЕН**

Динамічні зорові сцени є складними та багатовимірними об'єктами для аналізу. Поєднання руху та об'єктів у зорових сценах вимагає від методів аналізу великої обчислювальної потужності та ефективних алгоритмів.

Глибоке навчання дозволяє обчислювальним моделям, які складаються з кількох рівнів обробки, вивчати представлення даних із кількома рівнями абстракції. Ці методи значно вдосконалили сучасні технології розпізнавання мови, візуального розпізнавання об’єктів, виявлення об’єктів і багатьох інших областей, таких як відкриття ліків і геноміка. Глибоке навчання виявляє складну структуру у великих наборах даних за допомогою алгоритму зворотного поширення, щоб вказати, як машина повинна змінити свої внутрішні параметри, які використовуються для обчислення представлення на кожному рівні з представлення на попередньому рівні. Глибокі згорткові мережі принесли прорив в обробці зображень, відео, мови та аудіо, тоді як рекурентні мережі висвітлили послідовні дані, такі як текст і мова [1].

Для аналізу динамічних зорових сцен найчастіше використовують наступні методи:

1. Метод відстеження об'єктів

Цей метод використовує алгоритми комп'ютерного зору для відстеження руху об'єктів на відео. Він базується на аналізі зміни пікселів у послідовних кадрах та визначенні траєкторій руху об'єктів. Метод відстеження об'єктів широко використовується в системах безпеки, відеоспостереження та віртуальній реальності [2].

Переваги: Ефективний у виявленні та відстеженні об'єктів навіть у складних умовах, таких як зміна освітлення або часткова прихованість об'єкта.

Недоліки: Може бути нестабільним при швидкому русі об'єктів, а також вимагає значних обчислювальних ресурсів для роботи в реальному часі.

2. Метод детекції та класифікації об'єктів

Цей метод полягає у виявленні об'єктів на зоровому відео та їх класифікації за певними ознаками, такими як форма, розмір, кольорові характеристики тощо. Для цього використовуються методи машинного навчання, зокрема нейронні мережі. Цей підхід застосовується у системах розпізнавання облич, автоматичного сортування відеоматеріалів та медичній діагностиці [3].

Процес класифікації об’єктів складається з двох основних етапів: виділення ознак і машинного навчання [4]. Традиційний підхід полягає у сегментуванні об’єктів, а потім застосуванні деяких простих класифікаторів для подальшої класифікації, таких як простий класифікатор Байєса [5] або класифікатор опорних векторних машин (SVM) [6]. З поширеністю глибокого навчання було запропоновано багато методів, заснованих на CNN, для виявлення 3D-об’єктів [7], [8].

Переваги: Може надавати високу точність у виявленні об'єктів різних класів, що робить його корисним у задачах розпізнавання облич та інших завдань, що вимагають високої точності.

Недоліки: Вимагає велику кількість попередньо навчених даних та обчислювальних ресурсів для навчання та інференції моделей.

3. Метод аналізу глибини

Аналіз глибини полягає в оцінці відстаней до об'єктів на зоровому відео. Цей метод може бути реалізований за допомогою стереозору або методів, що базуються на рухомій камері. Він знайшов застосування у робототехніці для навігації та у віртуальній реальності для створення ефекту глибини [9], [10].

Переваги: Забезпечує інформацію про глибину сцени, що корисно для багатьох задач, таких як робототехніка та віртуальна реальність.

Недоліки: Вимагає спеціального обладнання, такого як стереокамера або дорогі алгоритми обробки для визначення глибини.

Загалом, кожен з цих методів має свої переваги та обмеження, і вибір конкретного методу може залежати від вимог застосування та характеристик даних. Дослідження методів аналізу динамічних зорових сцен відкриває широкі можливості їх застосування у різних галузях, що сприяє подальшому прогресу в сферах комп'ютерного зору, робототехніки та безпеки.

**Література**

1. LeCun Yann, Bengio Y., and Hinton Geoffrey. 2015. Deep learning. Nature 521, 05(2015), 436–444.
2. Lucas, B. D., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2, 674-679.
3. He K., Znang X. Deep Residual Learning for Image Recognition // Computer Vision and Pattern Recognition, CVPR. – 2015. – P. 770-778.
4. Tangruamsub S., Takada K., Hasegawa O. A fast online incremental learning method for object detection and pose classification using voting and combined appearance modeling. Signal Processing: Image Communication Volume 27, Issue 1, January 2012, Pages 75-82.
5. I. Rish, An empirical study of the Naïve Bayes classifier, in: IJCAI 2001 Work Empir Methods Artif Intell, vol. 3.
6. HearstM.A.Support vector machines IEEE Intell. Syst. (1998)
7. B. Graham, M. Engelcke, L. van der Maaten, 3D Semantic segmentation with submanifold sparse convolutional networks, in.
8. YanY. et al. SECOND: Sparsely embedded convolutional detection J. Sens. (2018).
9. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
10. Jain, A. K., & Dubes, R. C. (1988). Algorithms for clustering data. Prentice-Hall, Inc.