

*Афанасьєва Д. С., здобувач вищої освіти,
Ніколюк П. К., д-р фіз.-мат. наук, професор,
професор кафедри інформаційних технологій,
Донецький національний університет імені Василя Стуса*

ПОРІВНЯННЯ КЛАСИЧНИХ І НЕЙРОМЕРЕЖЕВИХ МЕТОДІВ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ У РЕАЛЬНОМУ ЧАСІ

Анотація. У роботі проведено порівняльне моделювання ефективності класичних і нейромережових методів розпізнавання об'єктів у реальному часі. Оцінено швидкодію, затримку та стабільність роботи моделей HOG + SVM і Faster R-CNN MobileNetV3 у режимі CPU, визначено напрями подальшої оптимізації гібридних систем.

Ключові слова: комп'ютерний зір, розпізнавання об'єктів, HOG + SVM, Faster R-CNN, реальний час.

Вступ. Системи розпізнавання об'єктів у реальному часі є основою функціональності та характеристик сучасних комп'ютерних технологій для автоматизації процесу розпізнавання візуальної інформації або прийняття рішень щодо її представлення. Вони застосовуються у сфері безпілотного транспорту, робототехніки, відеоаналітики, медицини, промислового контролю та цифрової безпеки. Продуктивність цих систем визначається компромісом між точністю класифікації, обчислювальною складністю, стійкістю до шумів та можливостями обробки в реальному часі (де затримка обробки кадру не перевищує десятків мілісекунд).

Метою дослідження є моделювання та експериментальне порівняння ефективності класичних і нейромережових методів розпізнавання об'єктів у реальному часі, з урахуванням взаємозалежності між точністю, швидкодією та ресурсоемістю алгоритмів. В основі дослідження аналіз двох підходів: традиційного, заснованого на HOG-дескрипторах і SVM-класифікаторі, та сучасного, що базується на згорткових нейронних мережах, зокрема архітектурах типу YOLOv5 і Faster R-CNN.

Основний текст. Ефективність систем розпізнавання об'єктів у реальному часі визначається балансом між точністю та швидкодією. Традиційні алгоритми комп'ютерного зору базуються на створенні ручних дескрипторів, тоді як нейромережові моделі формують ознаки автоматично, навчаючись на великих обсягах даних.

Метод HOG + SVM [1] залишається однією з найвідоміших класичних архітектур. Він виявляє об'єкти на основі гістограм орієнтованих градієнтів, що описують контури і форми у зображенні, а класифікацію виконує за допомогою методу опорних векторів. Такий підхід є обчислювально простим і придатним для роботи на центральному процесорі без GPU. Його стабільність під час зміни освітлення та масштабу робить його ефективним у промислових або мобільних системах. Однак через ручне визначення ознак HOG + SVM втрачає точність у складних сценах з поворотами, шумом і частковими перекриттями.

Розвиток згорткових нейронних мереж (CNN) став поворотним моментом у комп'ютерному зорі. Архітектура Faster R-CNN [2], об'єднала всі етапи: вияв-

лення, побудову ознак і класифікацію – в одну оптимізовану модель. Вона використовує регіональну мережу пропозицій (RPN), що автоматично визначає ймовірні області об'єктів. Подальші розробки призвели до появи компактних версій на базі MobileNetV3 [3], які оптимізовані для мобільних і вбудованих пристроїв, де важлива енергоефективність і швидкість обчислень.

Для емпіричного порівняння обидва підходи – HOG + SVM та Faster R-CNN MobileNetV3 – було реалізовано у середовищі Python із використанням бібліотек OpenCV і TorchVision. Обидві моделі тестувалися у режимі CPU на однаковому відеопотоці для розпізнавання класу «людина». Результати експерименту наведено на рис. 1.

```

=== Summary ===
Model                               Frames    FPS      Latency  AvgDet/Frame
-----
HOG+SVM (CPU)                       400      20.68   48.4 ms  0.00
FRCNN-MobileNetV3 (cpu)             153      5.15   194.2 ms 1.01

```

Рисунок 1 – Порівняння швидкодії та затримки між класичним і нейромережовим підходами

Згідно з отриманими даними, HOG + SVM досягнув середньої швидкодії 20.68 FPS із затримкою 48.4 мс, що забезпечує майже реальний час обробки. Натомість Faster R-CNN MobileNetV3 показав 5.15 FPS із латентністю 194.2 мс, але з вищою середньою кількістю детекцій (1.01 на кадр). Це свідчить про вищу точність нейромережової архітектури, проте значно більшу ресурсоемність і меншу швидкість. Отримані результати узгоджуються з даними досліджень Faster R-CNN [2], які також зазначали, що модель не досягає режиму реального часу без GPU-прискорення.

Висновки. Отже, результати експерименту підтвердили відомий компроміс між точністю та продуктивністю. Класичні методи залишаються оптимальними для систем з обмеженими обчислювальними ресурсами, тоді як нейромережові підходи доцільно застосовувати у сценаріях, де критичною є якість виявлення, а не швидкість обробки. Подальші дослідження доцільно спрямувати на розроблення гібридних систем, які поєднують переваги класичних і глибоких методів, а також на оптимізацію нейромережових архітектур для ефективної роботи на обмежених апаратних ресурсах.

Список використаних джерел

1. Dalal N., Triggs B. Histograms of oriented gradients for human detection, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2005. Vol. 1. DOI: 10.1109/CVPR.2005.177. URL: <https://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf>
2. Faster R-CNN: Towards real-time object detection with region proposal networks / S. Ren, K. He, R. Girshick, J. Sun. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017. Vol. 39, iss. 6. DOI: 10.1109/TPAMI.2016.2577031. URL: <https://shorturl.at/ufNNd>
3. Searching for MobileNetV3 / A. Howard, M. Sandler et al. *IEEE International Conference on Computer Vision (ICCV)*. 2019. DOI: 10.1109/ICCV.2019.00140. URL: <https://shorturl.at/Uiu6B>