

ЗАСТОСУВАННЯ ГЛИБОКОГО НАВЧАННЯ ДО ВИЯВЛЕННЯ ОБ'ЄКТІВ

Стрельцов О.А.

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

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

тел.: +38(099) 165-42-19, e-mail: oleksandr.streltsov@nure.ua

The paper considers the key aspects of object detection in images in the field of computer vision. It is shown that thanks to the development of deep learning, in particular convolutional neural networks (CNN), modern object detection systems have become extremely effective. Key technologies such as R-CNN, YOLO, SSD used for object detection in images are identified. Important components of training deep learning models, such as data augmentation and transfer learning, are also discussed. Examples of the application of object detection systems in various fields, such as the automotive industry, medicine, security, and cooking, are given.

Виявлення об'єктів на зображеннях є однією з ключових задач у галузі комп'ютерного зору. Початкові методи цього процесу, такі як ковзне вікно, базувалися на простих алгоритмах, що часто виявлялися неефективними. Ці методи вимагали значних зусиль для ручного вибору ознак та були не здатні адекватно справлятися зі складністю реальних зображень.

Завдяки розвитку глибокого навчання (deep learning), виявлення об'єктів пережило справжній прорив. Глибокі нейронні мережі, зокрема згорткові нейронні мережі (CNN), стали основою сучасних систем виявлення об'єктів [1]. Вони забезпечують високу точність та швидкість обробки зображень, що робить їх надзвичайно ефективними для виявлення об'єктів на різних типах зображень.

Ключові технології глибокого навчання в області виявлення об'єктів на зображеннях включають: згорткові нейронні мережі, R-CNN і його еволюції, YOLO та інші. Розглянемо кожен окремо.

Згорткові нейронні мережі (CNN) є важливим компонентом сучасних систем виявлення об'єктів на зображеннях. Вони можуть ефективно аналізувати зображення, автоматично виявляючи важливі ознаки без необхідності ручного налаштування. Ця здатність CNN робить їх незамінними для задач виявлення об'єктів у великих наборах даних, де ручна обробка стає непрактичною [1]. Використання CNN дозволяє створювати моделі з високою точністю і швидкістю обробки, що робить їх важливим інструментом у сферах, де потрібне точне і швидке виявлення об'єктів на зображеннях.

R-CNN (Region-based Convolutional Neural Networks) та його вдосконалені версії, такі як Fast R-CNN та Faster R-CNN, є методами, які використовують згорткові нейронні мережі (CNN) для аналізу регіонів зображен-

ня [2]. Це дозволяє їм здійснювати більш детальний аналіз потенційних локацій об'єктів на зображенні і підвищує точність виявлення. Кожен метод удосконалює попередній, додавши нові функціональності та покращуючи швидкість та ефективність виявлення об'єктів.

YOLO (You Only Look Once) і SSD (Single Shot Detector) є технологіями, спрямованими на досягнення високої швидкості виявлення об'єктів майже в реальному часі. Основна особливість цих методів полягає у тому, що вони використовують лише один прохід через нейронну мережу для класифікації і локалізації об'єктів на зображеннях [2]. Це дозволяє їм забезпечувати швидку та ефективну обробку зображень, що робить їх особливо підходящими для застосування у системах реального часу, таких як автомобільні системи безпеки, системи моніторингу та ін.

Аугментація даних і Transfer Learning є важливими компонентами у тренуванні ефективних моделей глибокого навчання для виявлення об'єктів. Аугментація даних збільшує різноманітність тренувального набору, допомагаючи моделям краще узагальнювати знання, тоді як Transfer Learning дозволяє застосовувати знання, отримані з одного задання, до іншого, зменшуючи потребу в об'ємних датасетах для тренування [3, 4].

Отже, глибоке навчання революціонізувало методи виявлення об'єктів на зображеннях, забезпечуючи високу точність і швидкість обробки. Наприклад, в автомобільній промисловості системи виявлення об'єктів використовуються для розпізнавання дорожніх знаків, пішоходів та ін. У медицині системи виявлення об'єктів можуть допомагати лікарям у ранньому виявленні захворювань на зображеннях. У сфері безпеки вони можуть використовуватися для виявлення підозрілих об'єктів на відеозаписах. У сфері кулінарії вони можуть використовуватися для розпізнавання продуктів харчування, щоб знаходити рецепти страв.

Список використаних джерел:

1. Wang, P., Sun, F., Jiang, W., Li, T., Zhang, H., Wang, C., Liu, Y. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. *IEEE Transactions on Neural Networks and Learning Systems*, 32(5), 1576–1587.
2. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778).
3. Zhuang, F., Zhu, C., Lin, Y., Shi, X., & Hoi, S. C. (2021). Comprehensive survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 33(11), 4583–4608.
4. Тітов С.В., Тітова О.В., Чорна О.С. Опис нескоротних наборів ознак в приблизних множинах з використанням систем числення // *Збірник наукових праць Харківського національного університету Повітряних Сил*. № 1(71) 2022, с. 106–110. <https://doi.org/10.30748/zhups.2022.71.12>