

УДК 004.032.26:004.932

ПЕРЕВАГИ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ЯК ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

Соловей І.В.

Науковий керівник – к.т.н., доцент Золотухін О.В.

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

тел. +38 (057) 702-13-37

Image recognition is a crucial area of study in computer vision and machine learning, with numerous applications in various fields such as medical imaging, self-driving cars, security, and entertainment. Convolutional Neural Networks (CNNs) have revolutionized image recognition technology, enabling machines to identify and classify images with high accuracy. This thesis provides an in-depth look at the advantages and challenges of CNNs, as well as their applications in different industries.

Згорткові нейронні мережі (CNN) виявилися потужною технологією для завдань розпізнавання зображень завдяки їхній унікальній здатності вивчати особливості з необроблених даних зображення. CNN – це тип глибоких нейронних мереж, які використовують згорткові шари для вилучення ознак з вхідного зображення, за якими слідує пулінг, який зменшує розмірність карти ознак. Цей процес повторюється кілька разів, створюючи ієрархію функцій, які можна використовувати для класифікації зображень, виявлення об'єктів [2] та інших завдань.

Однією з істотних переваг CNN є їх здатність автоматично вивчати ознаки [3]. Традиційні алгоритми розпізнавання зображень покладаються на створені вручну ознаки, які можуть займати багато часу та не підходити для всіх типів зображень. CNN, з іншого боку, можуть вивчати особливості з необроблених даних зображення, що робить їх більш адаптованими до широкого діапазону. Крім того, ознаки, які вивчаються CNN, ієрархічні, з нижніми шарами, які вивчають прості ознаки, такі як ребра і кути, а верхні шари вивчають більш складні - форми та текстури.

Ще однією перевагою CNN є їх здатність обробляти зміни масштабу та орієнтації зображення. Традиційні алгоритми розпізнавання вимагають, щоб вхідні зображення мали фіксований розмір та орієнтацію, що може бути складним у реальних сценаріях, де знімки можуть бути зроблені з різних кутів або мати різну роздільну здатність. CNN розроблені для інваріантності до трансляцій, що означає, що вони можуть обробляти зображення різних розмірів та орієнтацій. Це робить їх більш стійкими до реальних умов.

CNN показали надзвичайну адаптивність і можуть використовуватися для широкого спектру завдань розпізнавання зображень. Серед інших програм вони використовувалися для виявлення об'єктів, сегментації та розпізнавання обличчя.

CNN можна навчати на великих наборах даних за допомогою сучасних графічних процесорів, що робить їх масштабованою технологією. Це дозволило дослідникам навчати CNN на великих наборах даних, таких як ImageNet, який містить понад 14 мільйонів зображень і 20 000 категорій, що призвело до значного прогресу у розпізнаванні зображень. Можливість навчання на великих наборах даних дозволила розробити високоточні моделі, які можуть розпізнавати зображення з надзвичайною точністю.

CNN мають потенціал розвинути різні сфери, від охорони здоров'я до транспорту. В охороні здоров'я CNN є корисною для аналізу медичних зображень [4], таких як рентгенівські знімки та магнітно-резонансна томографія, що сприяє точності діагностики та лікування. Технологію можна зустріти в безпілотних автомобілях для виявлення та класифікації об'єктів на дорозі [1]. Крім того, CNN можна використовувати в сільському господарстві для виявлення шкідників у культурах, що призводить до підвищення врожайності та зменшення використання пестицидів. Незважаючи на вражаючу продуктивність CNN у задачах розпізнавання зображень, вони не позбавлені своїх обмежень. Однією з головних проблем є потреба у великих обсягах позначених даних для ефективного навчання моделей. Це може бути особливо складно в таких програмах, як аналіз медичних зображень, де отримання мічених даних може бути важким і трудомістким.

Підсумовуючи, CNN – це потужна технологія для розпізнавання зображень, яка пропонує кілька переваг перед традиційними алгоритмами розпізнавання зображень. Вони можуть вивчати особливості з необроблених даних зображення, обробляти варіації в масштабі та орієнтації, є адаптованими і масштабованими та мають потенціал для революції в різних сферах. Зі збільшенням наборів даних і потужних процесорів використання CNN стане більш поширеним у найближчі роки.

Список використаних джерел:

1. Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zhang, X. (2016). End to end learning for self-driving cars. <https://doi.org/10.48550/arXiv.1604.07316>.
2. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).
3. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.
4. Wang, S., Wang, R., Xu, X., Yang, S., Zhang, B., & Shao, L. (2021). Medical image analysis with convolutional neural networks. *IEEE Transactions on Medical Imaging*, 40(4), 998-1014.