

галюцинацій та обговорюються практичні наслідки теоретичних висновків для розгортання LLM. Оскільки уникнути галюцинацій неможливо, забезпечення безпеки LLM шляхом ретельного вивчення є вкрай важливим.

#### Література

1. Kees van Deemter. The Pitfalls of Defining Hallucination [Електронний ресурс] / Kees van Deemter // MIT Press Direct. – 2024. – Режим доступу до ресурсу: <https://direct.mit.edu/coli/article/50/2/807/119144/The-Pitfalls-of-Defining-Hallucination>.
2. Ziwei Xu. Hallucination is Inevitable: An Innate Limitation of Large Language Models [Електронний ресурс] / Ziwei Xu, Sanjay Jain, Mohan Kankanhalli. – 2024. – Режим доступу до ресурсу: <https://arxiv.org/abs/2401.11817>.
3. GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models [Електронний ресурс] / Iman Mirzadeh, Keivan Alizadeh, Hooman Shahrokhi та ін.]. – 2024. – Режим доступу до ресурсу: <https://arxiv.org/abs/2410.05229>.

### **ВИРШЕННЯ ЗАДАЧ ОБРОБКИ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ / SOLVING PATTERN RECOGNITION PROBLEMS WITH CONVOLUTIONAL NEURAL NETWORKS**

**Леонов М.С. / Leonov M.S**

Національний університет «Києво-Могилянська академія» / National University of «Kyiv-Mohyla Academy»

61068, Харківська обл., м.Харків, Фесенківський в'їзд, 9, кв.86. тел. +380734424707

Email: [m.leonov@ukma.edu.ua](mailto:m.leonov@ukma.edu.ua)

**Abstract. The given work is devoted to the convolutional neural networks and how they solve common problems faced in image recognition area. It begins with descriptions of common tasks that engineers are dealing with in computer vision field. Then proceeds to frequent pitfalls faced processing images. For each pitfall there will be a description of common solutions. Finally, the paper provides the overview of abstract convolutional neural network and how it resolves issues mentioned before.**

Задачі обробки зображень традиційно є одними з найпоширеніших задач машинного навчання та інтелектуальних систем. До цих задач можна традиційно віднести: класифікацію зображень, пошук об'єктів на зображеннях, визначення переміщення та швидкості об'єкту, ідентифікація особи, розпізнавання тексту написаного від руки, задача сегментації, задача генерації зображень тощо. Кожна з цих задач потребує індивідуального підходу, але існують деякі загальні правила та проблеми, що вирішуються у рамках кожної чи майже кожної системи. Перша проблема, що стає завжди – це дуже висока розмірність даних. Зображення, як правило, представлені у растровій формі, тобто типове кольорове зображення це тривимірна матриця, де перші два виміри позначають точку на зображенні, а третій вимір це кольор. Так невелике зображення розмірністю 224x224 з RGB кольором призводить до того, що матриця зростає до 150 тисяч значень. Очевидно, що обробка кожного елементу даних може коштувати дуже багато. Тому, зазвичай, йде робота на зменшення розмірності даних. Прикладами загальних методів є класична трансформація зображень, feature extraction, та методи проєкцій (PCA, t-SNE). Інша проблема -- це дуже велика чутливість до зміни незначних параметрів на зображенні. Наприклад незначна зміна контрастності, може драматично змінити числове представлення вектору, що вплине на алгоритм навчання. Рішенням цієї проблеми все частіше стає вже знайомі методи feature extraction (у випадку загорткових нейронних мереж, це процес згортки). Інше типове рішення – це знову застосування класичних алгоритмів, наприклад, виділення контурів на зображенні. Згорткові нейронні мережі здатні вирішити усі ці проблеми. Фундаментально, загорткові нейронні мережі типова є звичайною глибокою мережею, але на вхід подаються дані, що пройшли етапи згортки та пулінгу, замість сирих зображень. Згортка дозволяє виділити особливості зображення та дуже часто нормалізувати дані[1].

Згортка може бути представлена формулою:

$$[Y[i, j] = \sum_{m=0}^{H_k-1} \sum_{n=0}^{W_k-1} X[i + m, j + n] \cdot K[m, n]]$$

Де Y – вихід згортки. X – вхід згортки, K – ядро згортки.

Після шару згортки йде шар пулінгу[2]. Пулінг дозволяє відкинути зайвий шум зображення. Існує дві стратегії пулінгу. Перша стратегія це max pooling, де обирається максимальне значення з вікна пулінгу. Інша стратегія – це average pooling, що обчислює середнє значення у вікні. Приклад можна побачити на Fig.1.

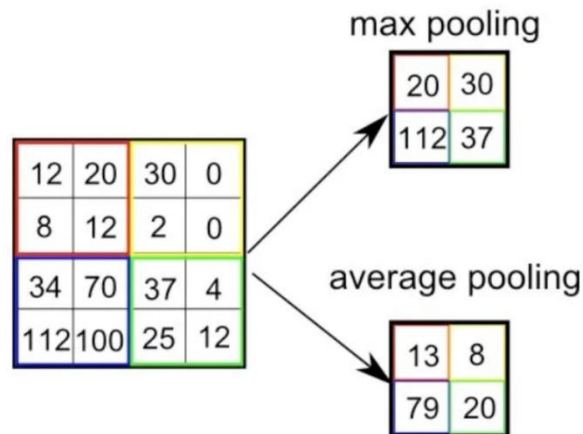


Fig. 1 Зображує різні стратегії пулінгу

Після декількох етапів згортки та пулінгу, трансформований вхід подається до нейронної мережі. Зазвичай це глибока мережа з багатьма скритими шарами, хоча це не обов'язкова умова. Таким чином, мережа навчається вже саме на трансформованих даних.

Отже, задачі обробки зображень є одними з найпоширеніших задач в області інтелектуальних систем та машинного навчання. Незважаючи на існуючі загальні підходи, кожен тип задачі та специфіка даних потребує індивідуального підходу. Використання згорткових нейронних мереж, тим не менш, дає узагальнене рішення для багатьох з цих задач. Водночас, такі мережі потребують подальших досліджень відносно типів згортки при обробці різних зображень та роботи з архітектурою саме глибокої мережі.

#### Література

1. N. Gupta, S. Mujumdar, P. Agarwal, A. Jain and S. Mehta, "Learning Convolutional Neural Networks with Deep Part Embeddings," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, 2019, pp. 2037-2041, doi: 10.1109/ICASSP.2019.8682192.
2. Zafar, Afia & Aamir, Muhammad & Mohd Naw, Nazri & Arshad, Ali & Riaz, Saman & Alruban, Abdulrahman & Dutta, Ashit & Alaybani, Sultan. (2022). A Comparison of Pooling Methods for Convolutional Neural Networks. *Applied Sciences*. 12. 8643. 10.3390/app12178643.