

РОЗПІЗНАВАННЯ ЕЛЕМЕНТІВ ЗОБРАЖЕННЯ ЗА ДОПОМОГОЮ ТЕХНОЛОГІЇ DEEP LEARNING З ВИКОРИСТАННЯМ GPU ТА CPU

Розпізнавання образів – це розділ теорії штучного інтелекту, що вивчає методи класифікації об’єктів. Об’єкт класифікації називають образом. Образом може бути цифра, буква, фотографія тощо.

Теорія розпізнавання образів включається в більш широку наукову дисципліну – теорію глибинного навчання (deep learning), метою якої є розробка методів побудови алгоритмів, що здатні навчатися.

Однією з типів нейронної мережі, що використовується для розпізнавання елементів зображення являється згорткова нейронна мережа.

Згорткова нейронна мережа (ЗНМ), що була розроблена ґрунтуючись на організації зорової кори тварин, являє собою варіацією багатшарових перцептронів.

Архітектура ЗНМ складається з вхідного та вихідного шару, та багатьох прихованих шарів. Приховані шари складаються із згорткових, підвибіркових або повноз’єднаних шарів.

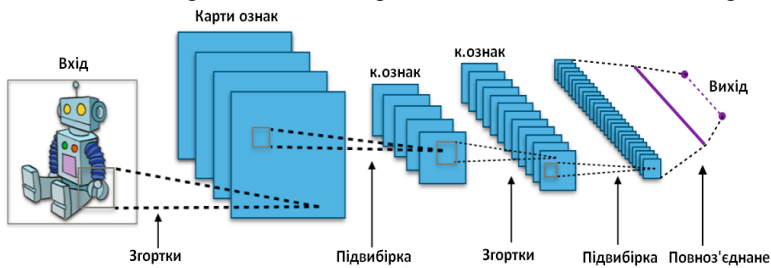


Рис.1. Типова архітектура «згорткової нейронної системи»

Згортковий шар – шар, що застосовує операцію згортки до входу та передає результат до наступного шару. Під час виконання згорткових операцій відбувається застосування фільтрів, що класифікують вхідні дані за певними ознаками.

Підвибірковий шар – шар, основна задача якого полягає у поступовому зменшенні розміру даних для обробки нейронною мережею. Даний шар розділяє зображення на зрізи, та за допомогою функції максимізації вибирає для кожної області максимум. Найчастіше використовується фільтр 2x2, що для кожного проходу зменшує розмір вдвічі.

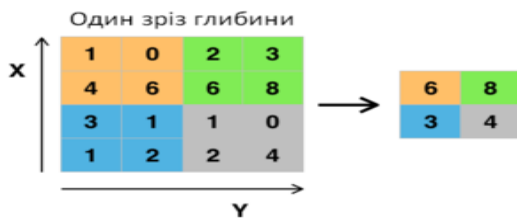


Рис.2. Робота підвибіркового шару із фільтром 2x2

Повноз’єднаний шар – шар, який являє собою звичайний багатшаровий перцептрон, який, на основі зменшеного розміру даних, проводить обробку вхідної інформації та формує цільовий вихідний шар.

Розглянемо швидкість роботи згорткової нейронної мережі в залежності від виду обчислювальної техніки, що використовувалася. Для тесту було взято стандартну модель згорткової мережі фреймворку Tensorflow, тестовий набір CIFAR-10, що містить 60000 зображень розміром 32x32, та обчислювальне обладнання: CPU i7-4930k, GPU: GTX 940MX, GTX 970, GTX 980, GTX 980TI, GTX 1070.

Таблиця 1. Час роботи обчислювального обладнання

Обчислювальне обладнання	Час виводу		Навчання Час / епоху
	1 зображення	128 зображень	
Intel i7-4930K	3 ms	244 ms	231.0 s
GeForce 940MX	4 ms	120 ms	145.6 s
GTX 970	6 ms	32 ms	25.0s-26.3s
GTX 980	3 ms	24 ms	20.5s-21.1s
GTX 980 TI	5 ms	27 ms	22.0s-22.1s
GTX 1070	2 ms	15 ms	14.4s-14.5s

Таким чином, можна зробити висновок, що використання CPU замість GPU, як основного обчислювального обладнання не являється ефективним при роботі зі згортковою нейронною мережею. GPU більш оптимізоване для роботи з великою кількістю зображень. Із результатів можна побачити, що більш нові моделі графічних адаптерів прискорюють роботу нейронної мережі, та можуть значно зменшити час їх навчання.