

ВПЛИВ ОБРАНОЇ АПАРАТНОЇ ПЛАТФОРМИ НА ЕФЕКТИВНІСТЬ РОБОТИ ЗГОРТАЛЬНОЇ НЕЙРОННОЇ МЕРЕЖІ

Ковтун В'ячеслав, канд. техн. наук, доцент кафедри комп'ютерних систем управління,

Максимов Олексій, студент групи 2АВ-136
Вінницький національний університет, Україна

Вступ

Враховуючи, що більшість інформації людина сприймає у вигляді зображень, їх автоматизований аналіз є природним і дуже інформативним способом опису навколишнього середовища. Так як кількість джерел отримання візуальної інформації стрімко зростає, задача автоматизованого аналізу зображень стає вкрай актуальною. Найбільш ефективними засобами автоматизованого прийняття рішень у подібних задачах класифікації є штучні нейронні мережі, зокрема, згортальні нейронні мережі глибокого навчання [4]. Глибоке навчання (deep learning) – це галузь машинного навчання, яке ґрунтується на підборі алгоритмів, що здійснюються за допомогою нейронних мереж із спеціалізованою архітектурою, про що говоритиметься далі [1].

Метою роботи є підвищення ефективності автоматизованого розпізнавання зображень та дослідження залежності ефективності навчання згортальної нейронної мережі в залежності від обраної апаратної платформи.

Результати дослідження

Для розпізнавання зображень автори застосовують згортальну нейронну мережу, яка складається з впорядкованого набору спеціалізованих шарів, зокрема, шарів згортання (c-layers), шарів підвибірки (s-layers) і повнозв'язних шарів (f-layers) [2]. Для проведення експериментів було створено згортальну нейронну мережу з архітектурою, наведеною на рисунку 1.

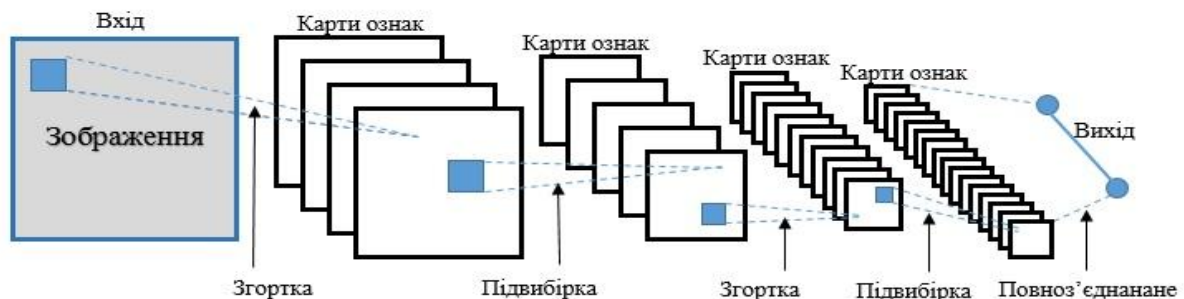


Рис. 1. Архітектура згортальної нейронної мережі

Розглянемо типову структуру згортальної нейронної мережі більш детально. Мережа складається з великої кількості шарів. Після початкового шару (вхідного зображення) сигнал проходить серію згортальних шарів, в яких чергується шар згортки та шар субдискретизації. Чергування шарів дозволяє складати «карти ознак», а з карт ознак, на кожному наступному шарі карта зменшується в розмірі, але збільшується кількість каналів. На практиці це означає здатність розпізнавання складних ієрархій ознак. Зазвичай після проходження декількох шарів карта ознак перетворюється у вектор або навіть скаляр, але таких карт ознак стає сотні. На виході згортальних шарів мережі додатково встановлюють кілька шарів повнозв'язної нейронної мережі, на вхід якого подаються кінцеві карти ознак.

Шар згортки (convolutional layer) - це основний блок згортальної нейронної мережі. Шар згортки включає в себе для кожного каналу свій фільтр, ядро згортки якого обробляє попередній шар за

фрагментами (підсумовуючи результати матричного твору для кожного фрагмента). Вагові коефіцієнти ядра згортки (невеликий матриці) невідомі і встановлюються в процесі навчання.

Особливістю згортального шару є порівняно невелика кількість параметрів, яке встановлюється при навчанні. Так наприклад, якщо вихідне зображення має розмірність 100×100 пікселів по трьом каналам (це значить 30000 вхідних нейронів), а згортальний шар використовує фільтри с ядром 3×3 пікселя з виходом на 6 каналів, тоді в процесі навчання визначається тільки 9 ваг ядра, однак по всім сполученням каналів, тобто $3 \times 3 \times 3 \times 6 = 162$, в такому випадку даний шар вимагає знаходження тільки 162 параметрів, що істотно менше кількості шуканих параметрів повно нейронної мережі. На рисунку 1.11 схематично зображено згортальний шар.

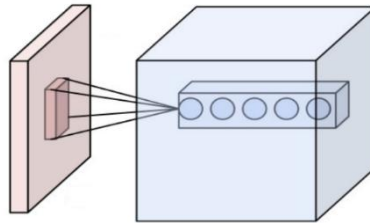


Рис. 2. Шар згортки

Шар субдискретизації (підвибірки) являє собою нелінійне ущільнення карти ознак, при цьому група пікселів ущільнюється до одного пікселя, проходячи нелінійне перетворення. Найбільш споживані при цьому функція максимуму. Перетворення зачіпають непересічні прямокутники або квадрати, кожен з яких скорочується в один піксель, при цьому вибирається піксель, що має максимальне значення. Операція підвибірки дозволяє істотно зменшити просторовий обсяг зображення. Субдискретизація інтерпретується так: якщо на попередній операції згортки вже були виявлені деякі ознаки, то для подальшої обробки настільки детально зображення вже не потрібно, і воно ущільнюється до менш детального. До того ж фільтрація вже непотрібних деталей допомагає не перенавчатися. Шар підвибірки, як правило, вставляється після шару згортки перед шаром наступної згортки. На рисунку 1.11 схематично зображено шар субдискретизації.

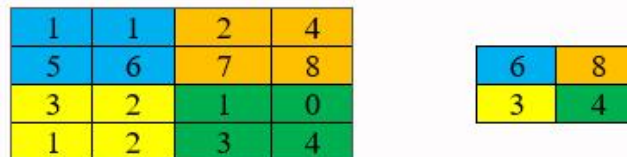


Рис. 3. Шар субдискретизації

Спроектований класифікатор реалізовано у програмному забезпеченні Matlab із використанням Deep learning toolbox. Алгоритми навчання згортальної нейронної мережі, реалізовані у цьому toolbox-і, можуть навчатися використовуючи ресурси процесора (CPU) або відеоадаптера (GPU) з архітектурою від виробника NVidia [3]. Отже, доцільним є виявлення залежності між тривалістю навчання згортальної нейронної мережі та обраною апаратною платформою (CPU/GPU). Проведення досліджень передбачало розпізнавання зображень елементів 3-х класів об'єктів (літаки, кораблі, ноутбуки), об'єднаних у відповідні множини. Кількість зображень в кожному класі дорівнювала 67. Загалом вибірку було розбито на навчальну та тестувальну у пропорції 7/3 [5]. Тривалість навчання наведеної на рисунку 1 згортальної нейронної мережі представлено на рисунку 4.

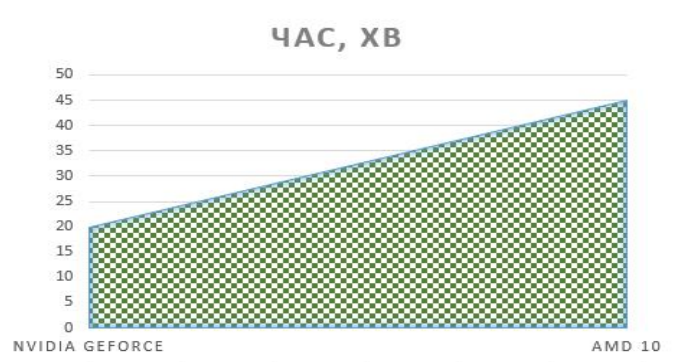


Рис. 4. Залежність тривалості навчання CNN від апаратного забезпечення

Проведені експериментальні дослідження показали високу ефективність згортальної нейронної мережі в задачі розпізнавання зображень, що підтверджується середньою імовірністю правильного розпізнавання елементів відповідних класів зображень у 92%, розрахованою на основі класифікації зображень із тестувальної вибірки.

Висновки

Відповідно до поставленої мети авторами експериментально доведено ефективність застосування згортальної нейронної мережі для розпізнавання зображень, за результатами яких імовірність правильного розпізнавання трьох класів об'єктів склала 92%, при чому тривалість навчання нейронної мережі із застосуванням GPU є у 4 рази меншою порівняно із навчання мережі на CPU при застосуванні алгоритмів з Matlab Deep learning toolbox.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Субботін С.О. Згортальна нейронна мережа як парадигма для реалізації технології глибокого навчання / С.О. Субботін, А.О. Олійник, О.Ю. Благодарьов. — ЗНТУ, 2015. — 1с.
2. Фурман М.С. Дослідження нейромережевого класифікатору цифрових зображень / М.С. Фурман, В.В. Ковтун ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2016. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/10974>.
3. Биков М.М. Метод оптимізації процесу навчання нейромережі в задачі розпізнавання мовців / М.М. Биков, В.В. Ковтун, А. Раїмі ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2015. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/3584> .
4. Биков М.М. Метод нормалізації тривалості звучання парольних фраз для системи розпізнавання мовців / М.М. Биков, В.В. Ковтун, А. Раїмі ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2015. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/3585?locale-attribute=en> .