

Система класифікації жанру картин

ДИПЛОМНА РОБОТА
СТУДЕНТА ННК "ІПСА" ГРУПИ КА-44
КОВАЛЕНКА АНДРІЯ ВОЛОДИМИРОВИЧА
НАУКОВИЙ КЕРІВНИК ПРОФ. ДАНИЛОВ В.Я.





Об'єкт дослідження:

методи нейронних мереж для класифікації



Предмет дослідження:

Методи згорткових нейронних мереж для класифікації жанру картин



Мета дослідження:

Проаналізувати і виявити ефективність використання згорткових НМ для розпізнавання та класифікації картин



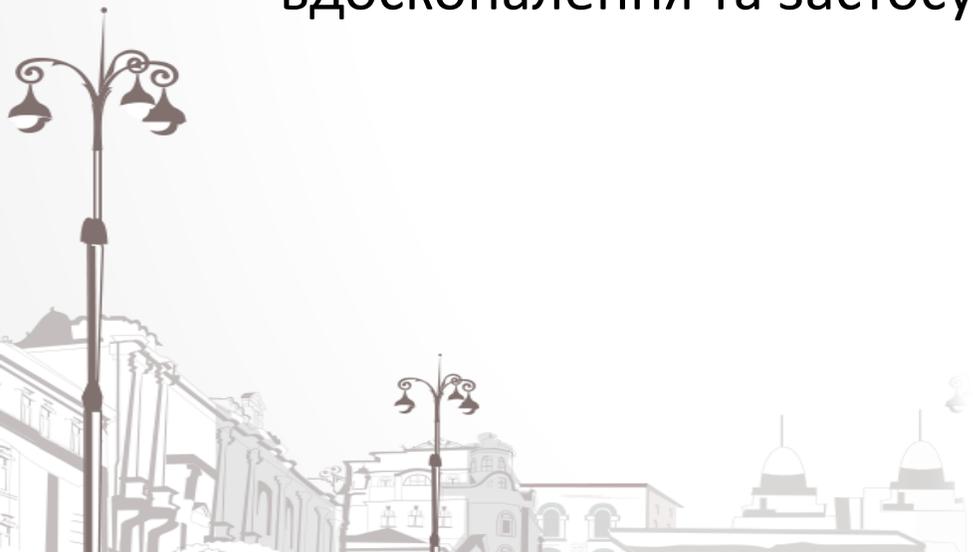
Актуальність роботи

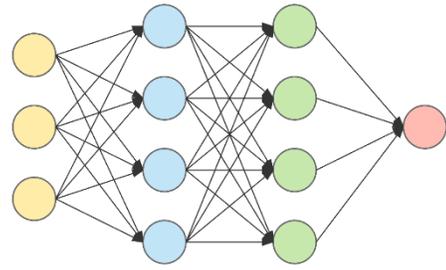
- Дане дослідження є актуальним для великих баз даних художніх картин, де потрібно класифікувати великий обсяг картин, що є трудомісткою та часовитратною задачею для людини.
- Можливість класифікування, як картин так і інших зображень в фільтрах пошукових систем, смартфонах та інших девайсах.



Постановка задачі

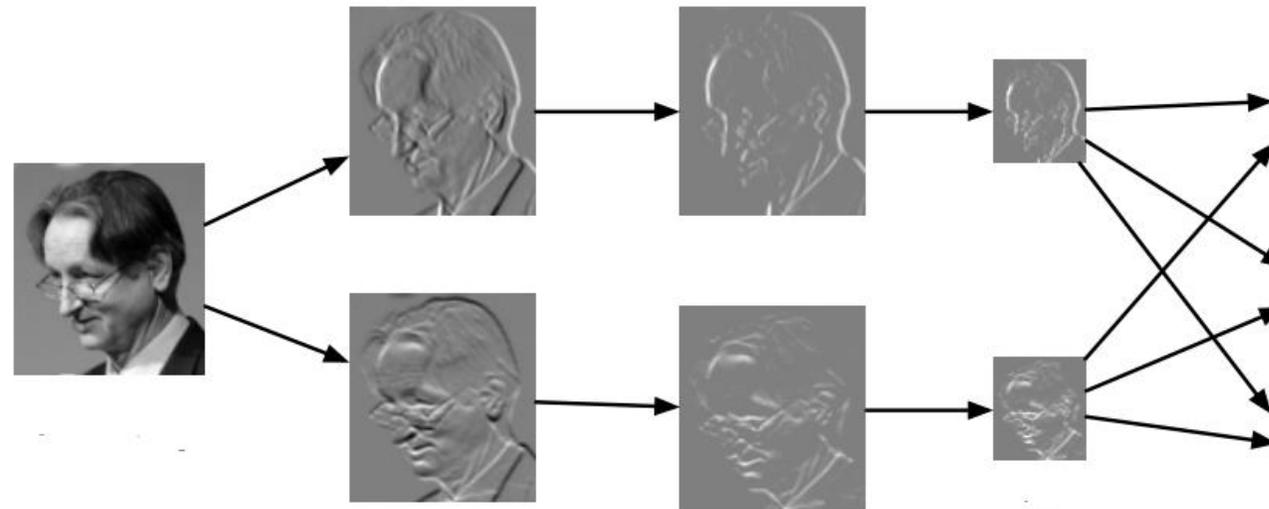
- Виконати аналіз роботи згорткових нейронних мереж
- Дослідити якість роботи мереж на різних вхідних даних.
- Сформулювати рекомендації стосовно можливостей подальшого вдосконалення та застосування даного підходу

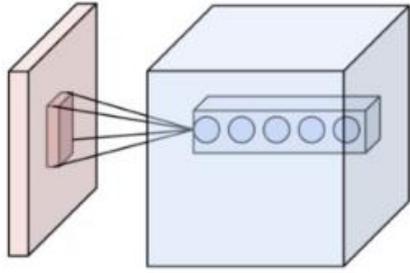




Згорткові нейронні мережі

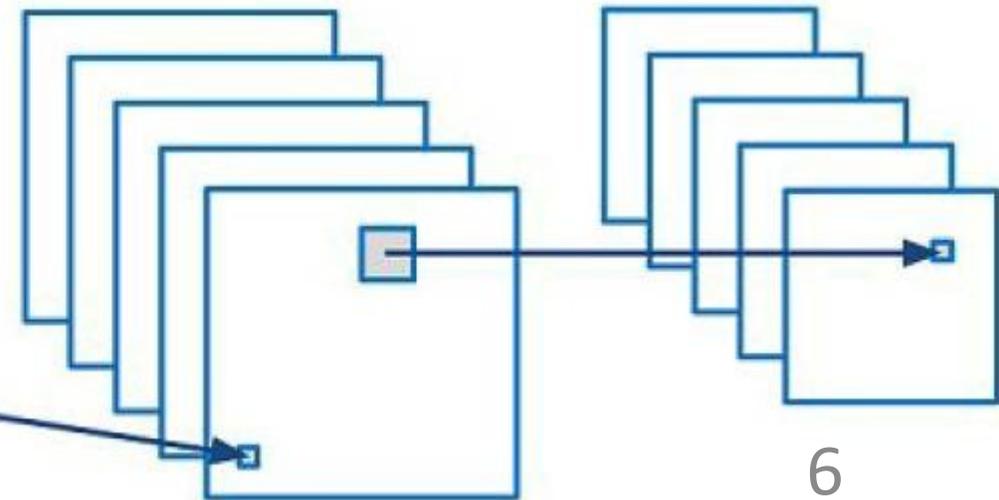
- Найбільш ефективним методом для класифікації зображень є згорткові нейронні мережі.
- Береться зображення, пропускається через серію згорткових, активаційних шарів та шарів пулінгу і генерується виведення. Виведенням може бути клас або ймовірність класів, які краще всього описують зображення.

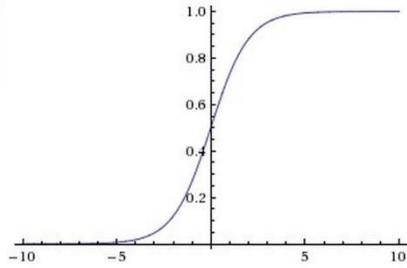




Шар згортки

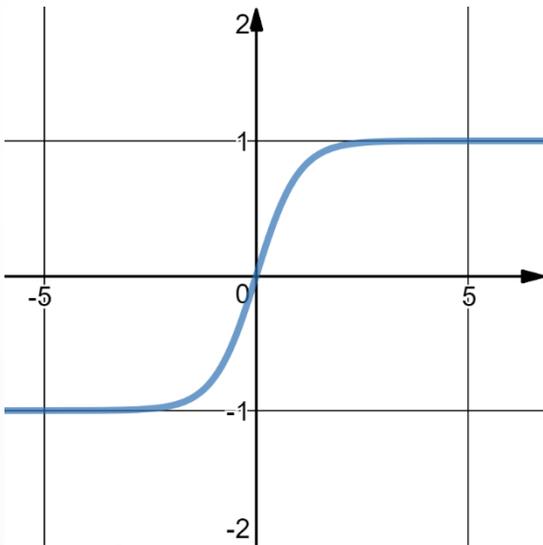
- Шар згортки включає в себе для кожного каналу свій фільтр, ядро згортки якого обробляє попередній шар за фрагментами (підсумовуючи результати матричного добутку для кожного фрагмента).
- Особливістю згорткового шару є порівняно невелика кількість параметрів, яка встановлюється при навчанні.



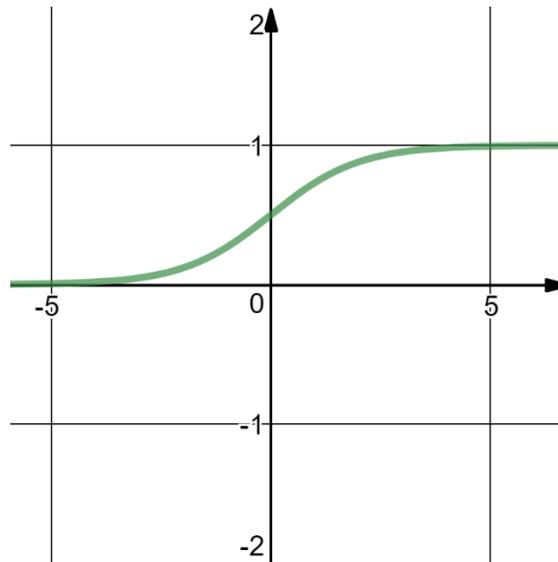


Шар активації

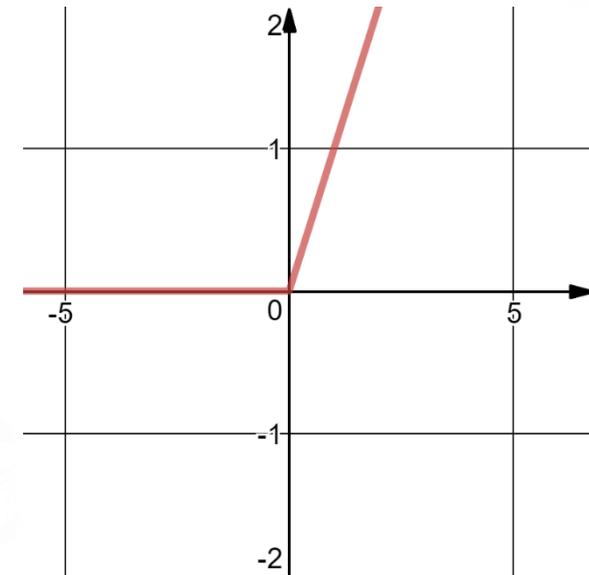
- Скалярний результат кожної згортки потрапляє на функцію активації, яка представляє собою якусь нелінійну функцію. Функція нелінійності може бути будь-якою з вибору дослідника, традиційно для цього використовували функції типу гіперболічного тангенса - $f(x) = \tanh(x)$, сигмоид - $f(x) = \frac{1}{1+e^{-x}}$ або ReLU - $f(x) = \max(0, x)$



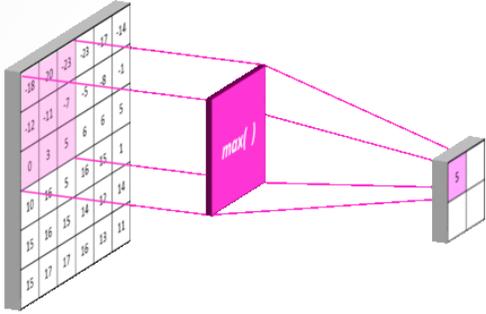
$$f(x) = \tanh(x)$$



$$f(x) = \frac{1}{1+e^{-x}}$$

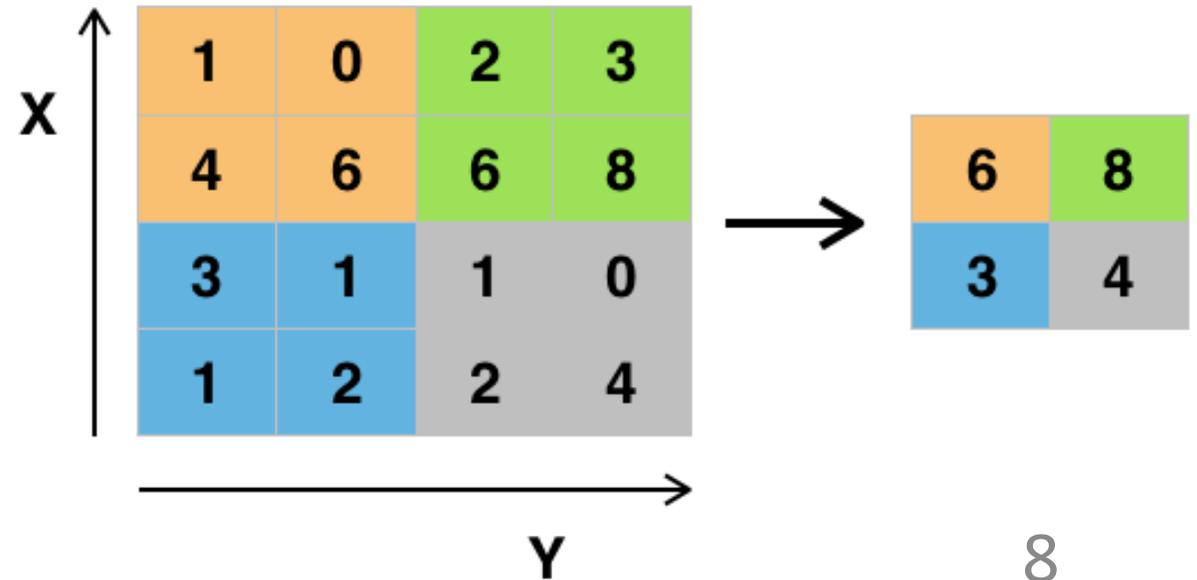


$$f(x) = \max(0, x)$$



Пулінг або шар субдискретизації

Шар пулінг (інакше шар підвибірки) являє собою нелінійне ущільнення карти ознак, при цьому група пікселів (зазвичай розміру 2×2) ущільнюється до одного пікселя, проходячи нелінійне перетворення.





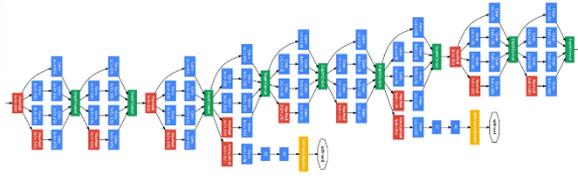
Переваги та недоліки

Переваги:

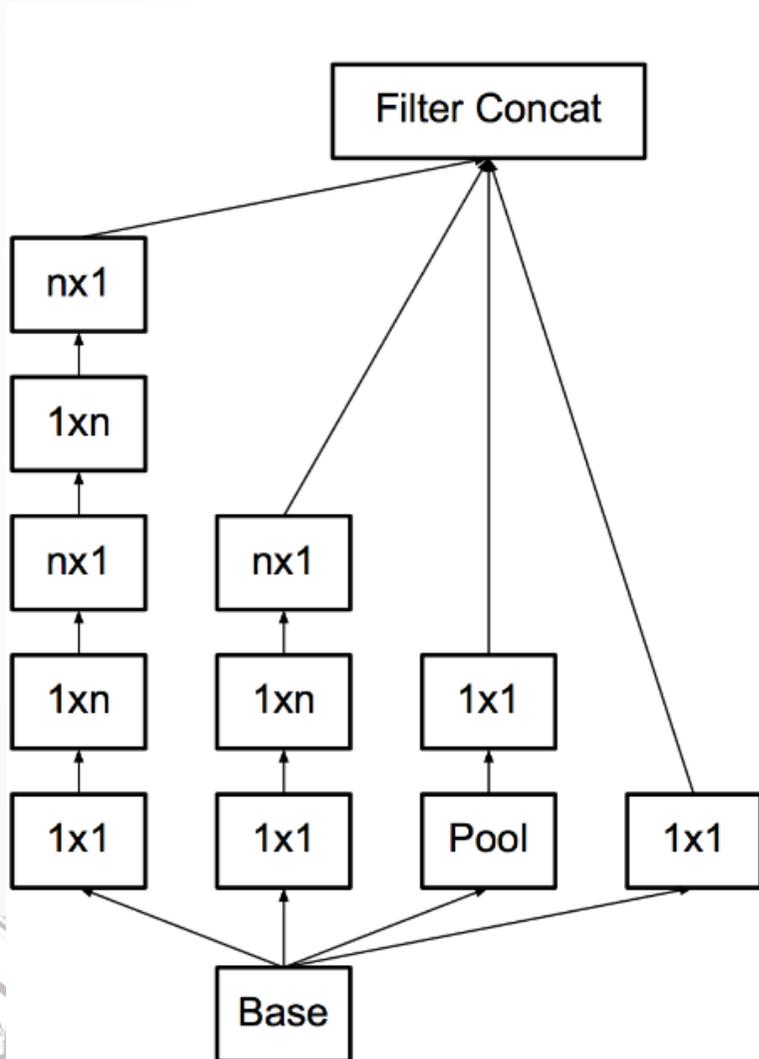
- Зручне розпаралелювання обчислень, а отже, можливість реалізації алгоритмів роботи і навчання мережі на графічних процесорах.
- Відносна стійкість до повороту і зміщення розпізнаваного зображення.
- Навчання за допомогою класичного методу зворотного поширення похибки.

Недоліки:

- Надто багато варійованих параметрів мережі.
- Складність роботи з зображеннями, що містить об'єкти з декількох класів.
- Довгий час навчання на довгих вибірках і великих зображеннях.



Модель Inception v3



Основні принципи:

- Дуже багато сигналів близькі один до одного в просторі, і цим можна користуватися, щоб робити згортку меншого розміру.
- Для ефективного використання ресурсів потрібно збільшувати і ширину, і глибину мережі.
- Погано мати різкі зменшення розмірності, тобто шари з різким зменшенням параметрів, особливо на початку.
- "Широкі" шари швидше навчаються, що особливо важливо на високих рівнях



Результати роботи

Приклад вдалого розпізнавання зображення. Програма з ймовірністю 98% стверджує, що на картині зображено портрет.

C:\Users\User\Desktop\dip\Incep\Data\Portrait\129795638-24942c.jp

Load



portrait 98.6 %

Predict

82.3%

Average percentage of recognition



Результати роботи

- З ймовірністю 99% на картині зображено архітектуру. Загальний відсоток правильно розпізнаних картин на навчальний вибірці становить 82%.

C:\Users\User\Desktop\dip\Incep\Data\Landscapes\227860800-31978

Load



landscapes 99.1 %

Predict

82.3%

Average percentage of recognition



Результати роботи

Тип картини	Кількість зображень в вибірці	Кількість правильно розпізнаних картин
Architecture	1028	82,6%
Birds	1006	84,9%
Cars	598	75,7%
Flowers	777	87,7%
Landscapes	916	89,1%
Portrait	1035	73,5%
Space	757	82,3%



Висновок

- Розкрита проблема класифікації картин та зображень
- Проаналізована структура побудованої моделі згорткової нейронної мережі
- Виконано аналіз моделі InceptionV3, який показав, що вона є ефективним та потужним інструментом класифікації та розпізнавання картин
- Продемонстровано роботу програми при класифікації картин по жанрам, та отримано загальну точність класифікації на тестовій вибірці 82.3 %



Дякую за увагу!

