

Класифікація зображень за допомогою мереж глибинного навчання

ДИПЛОМНА РОБОТА
СТУДЕНТА ННК “ІПСА” ГРУПИ КА-35
ЧЕРКЕСА МИКИТИ ВІКТОРОВИЧА

НАУКОВИЙ КЕРІВНИК ПРОФ. ДАНИЛОВ В.Я.

Об'єкт дослідження:

- методи згорткових мереж для класифікації

Предмет дослідження:

- Методи нейронних мереж для класифікації жанру картин

Мета дослідження:

- проаналізувати предмет дослідження та дослідити ефективність застосування нейронних мереж до задач класифікації зображень

Актуальність роботи

- Класифікатор жанру картини буде цікавий художнім музеям, організаціям які колекціонують картини і т.д. Більше не знадобляться фахівці для маркування кожної картини, цим буде займатися машина.
- Крім того, розробка рішення для ефективної класифікації живопису - це перший крок, що дозволяє системам більш високого рівня класифікувати великі сховища зображень або робити художні рекомендації.

Основні підходи до класифікації зображень

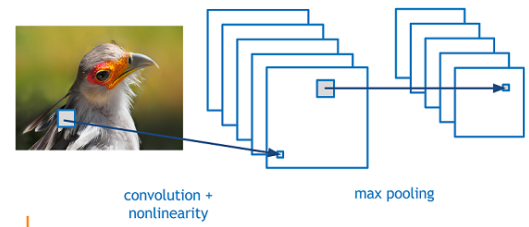
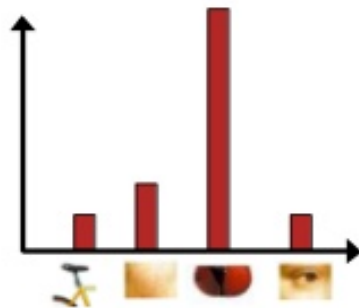
Виділяють два підходи:

- Підхід на підставі аналізу зображень за допомогою Методу Опорних Векторів (SVM)

Собираем множество фрагментов, кластеризируем и строим словарь, считаем “мешки слов” для каждого изображения, обучаем SVM на мешках слов

- Підхід на підставі використання згорткових нейронних мереж

В процессе обучения нейросеть определяет какие визуальные признаки являются более значимыми



Переваги згорткових нейронних мереж

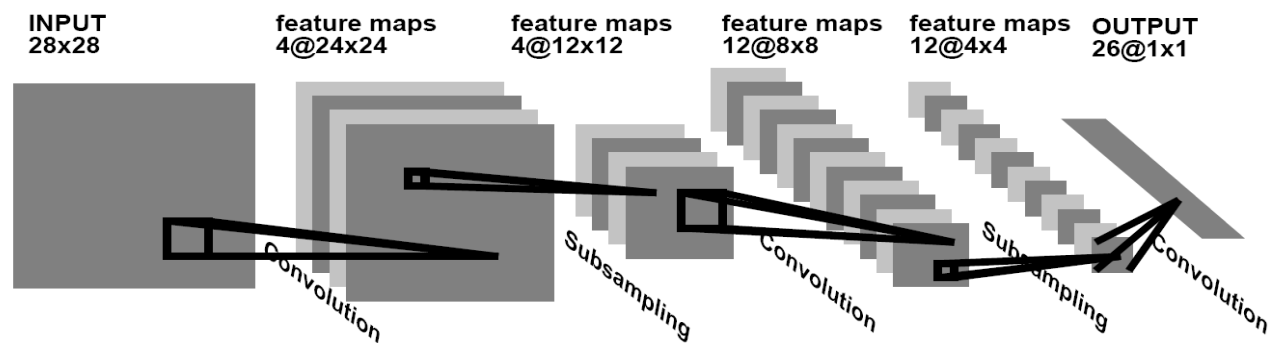
- Використовують мало попередньої предобробки
- Згорткові мережі самостійно визначають ознаки
- Можливість розпізнавання більш складних ієрархій ознак
- Інваріантність до поворотів, зміщення та масштабу зображення

Постановка задачі

- Виконати аналіз роботи згорткової нейронної мережі
- Дослідити якість роботи мережі в залежності від різних параметрах в межах поставленої задачі
- Сформулювати рекомендації стосовно можливостей подальшого вдосконалення та застосування даного підходу

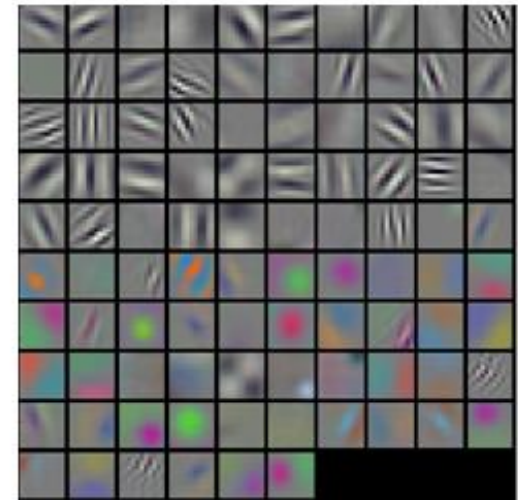
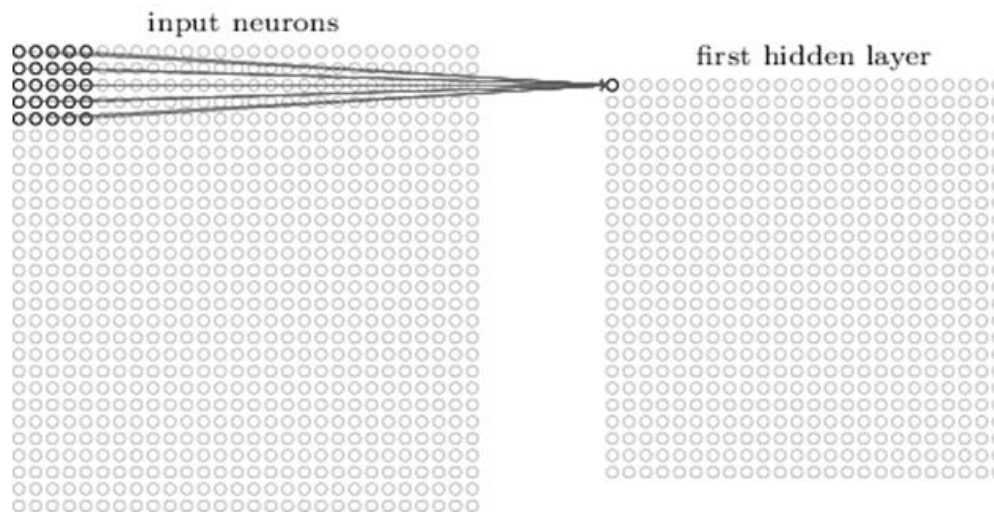
Робота згорткових нейронних мереж

Основною особливістю ЗНМ порівняно із іншими нейронними мережами є чергування загорткових та субдискретизуючих шарів, а також наявності повнозв'язних шарів на виході. Згорткові шари виконують операцію згортки над зображенням з певним ядром, що дозволяє виділити певні ознаки зображення із збереженням його топології. Субдискретизуючі шари зменшують просторову розмірність зображення, що забезпечує інваріантність до масштабу. Чергування шарів дозволяє складати карти ознак з карт ознак, що на практиці означає здатність розпізнавання складних ієрархій ознак.



Архітектура згорткового шару

Шар згортки - це основний блок згорткової нейронної мережі. Шар згортки включає в себе ядро згортки, яке обробляє попередній шар за фрагментами (підсумовуючи результати матричного перемноження для кожного фрагмента). Вагові коефіцієнти ядра згортки невідомі і встановлюються в процесі навчання.



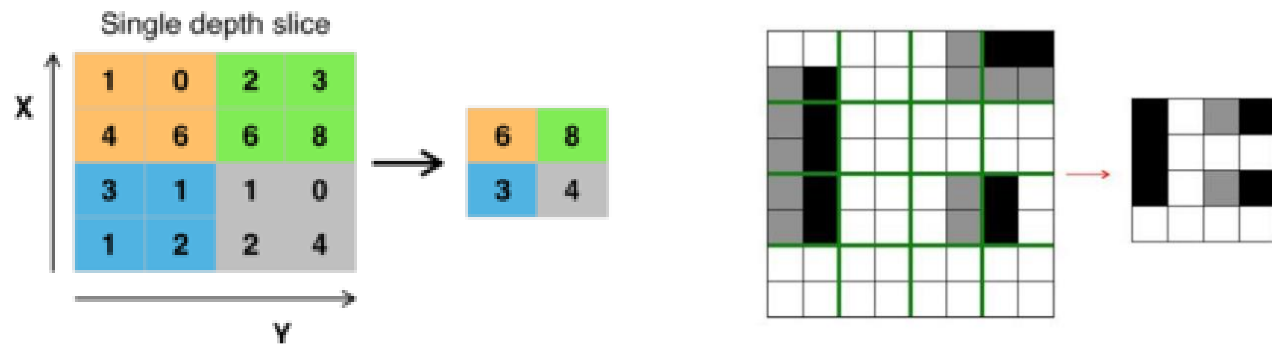
- Слой свёртки — это основной блок свёрточной нейронной сети. Слой свёртки включает в себя ядро свёртки, которое обрабатывает предыдущий слой по фрагментам (суммируя результаты матричного произведения для каждого фрагмента). Весовые коэффициенты ядра свёртки неизвестны и устанавливаются в процессе обучения.

Архітектура шару субдискретизації

Суть роботи шару субдискретизації полягає в розділенні карти ознак на непересічні ділянки і виділення на цих ділянках нейронів з максимальною активністю.

$$F = \max(x)$$

Це дозволяє скоротити число параметрів згорткової нейронної мережі, що в свою чергу дозволяє прискорити процес навчання і знизити можливість перенавчання мережі.



Опис вхідних даних

В якості вхідних даних розглянуто датасет картин на сайті WikiArt.org.:

Навчальні дані містять 3336 зображення, серед яких:

- 1161 зображень картин в жанрі портрет
- 1123 зображень картин в жанрі пейзаж
- 1052 зображень картин в жанрі міфологія

Всі зображення представлені у різних розмірах.

Всього для тренування було обрано по 800 зображень з кожного класу.

Вплив параметрів на роботу мережі

При роботі з нейронною мережею різні параметри по різному впливають на її якість, в даному випадку - якість класифікації. Серед цих параметрів можна вказати:

- Кількість карт в згорткових шарах
- Кількість шарів в нейронній мережі
- Розмір вікна згортки
- Розмір шагу навчання

Початкова архітектура мережі

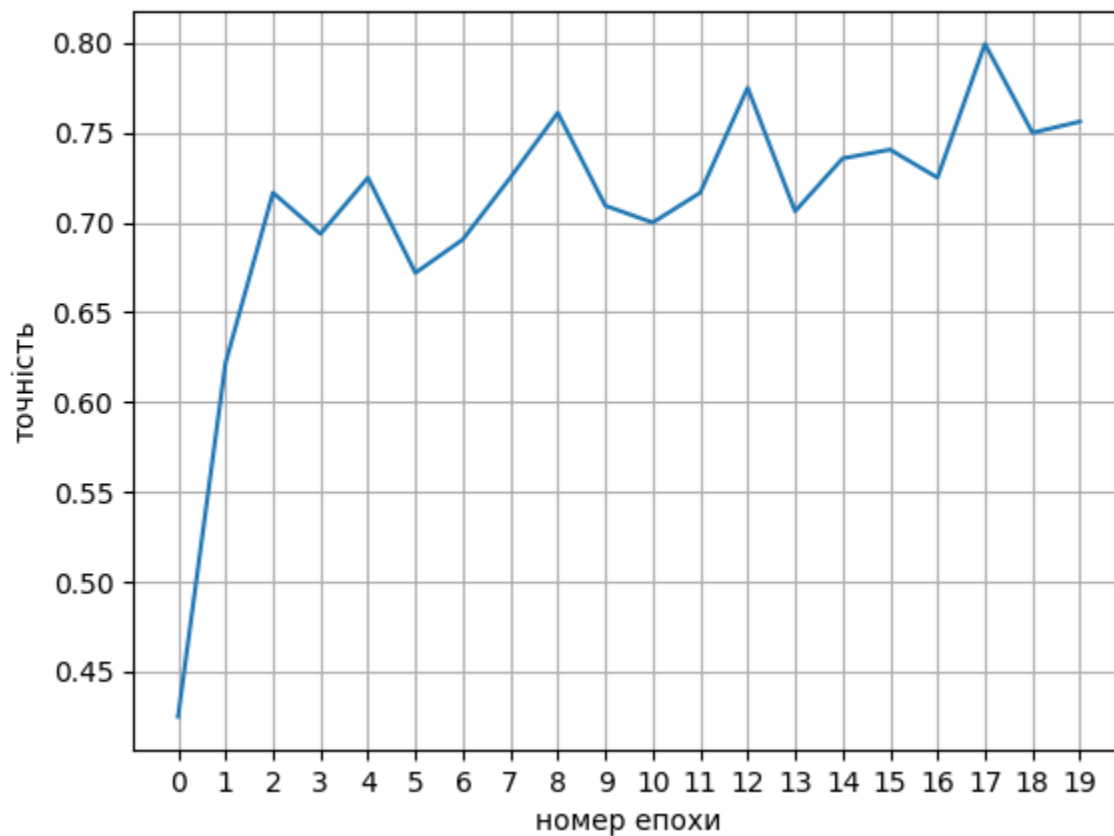
Початкова архітектура згорткової нейронної мережі складалась з 8 шарів, серед яких :

- Перші шість шарів – це попарі згортковий та субдискретизаційний шари. Розмір згорткового вікна у кожному згортковому шарі – 3x3. Розмір вікна у субдискретизаційних шарах – 2x2.
Активаційна функція – ReLU: $F(x) = \max(0, x)$
- Сьомий шар – повнозв'язний шар, що складається з 64 нейронів.
- Восьмий шар – повнозв'язний шар, що складається з 3 нейронів (кількість класів класифікації)

Крок навчання = 0.01

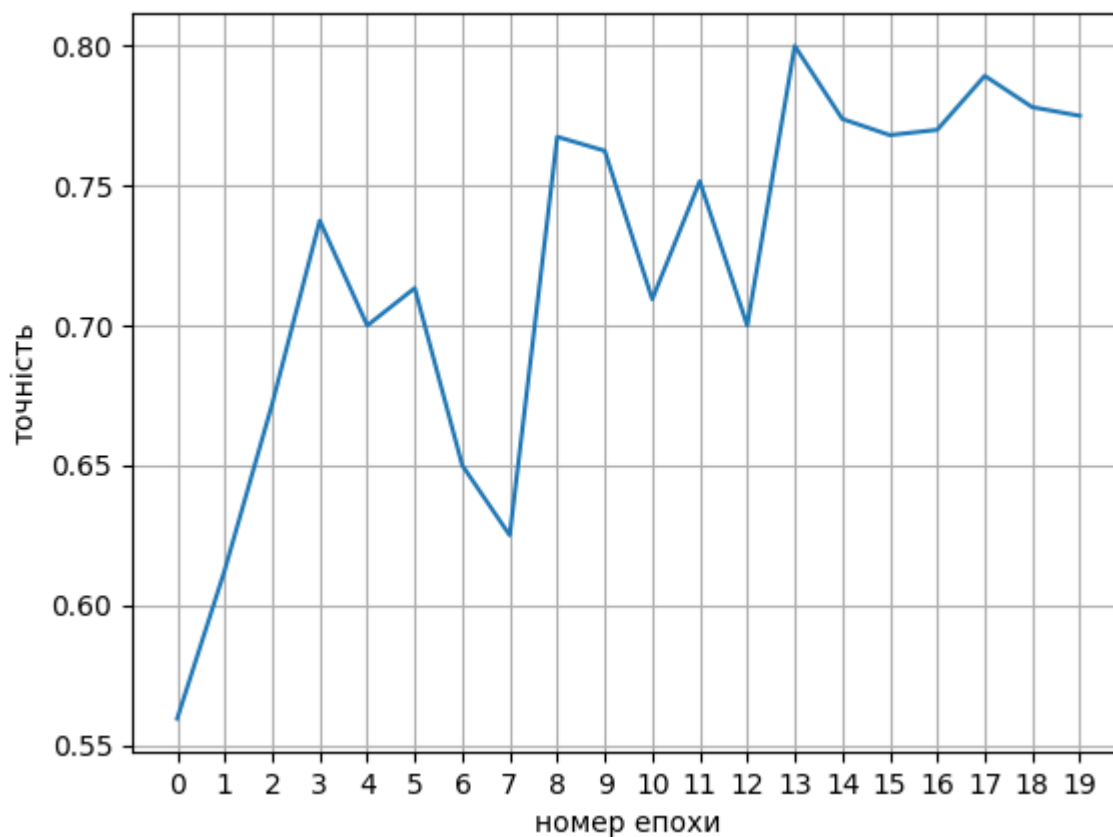
Точність початкової конфігурації мережі

Час навчання: 36 хв.



Збільшення кількості шарів

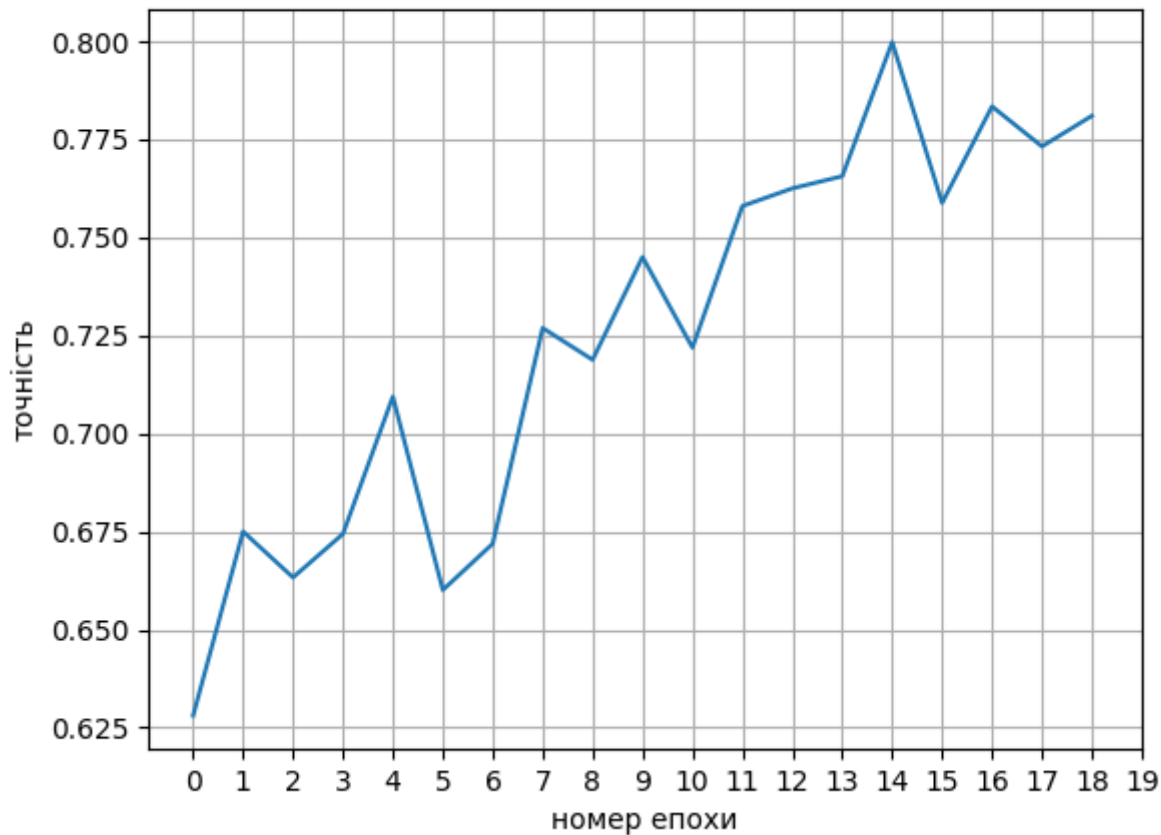
При додаванні ще одної пари згорткового та субдискретизаційного шара відбулися наступні зміни в точності класифікації:



Час навчання: 59 хв.

Збільшення кількості карт в згорткових шарах

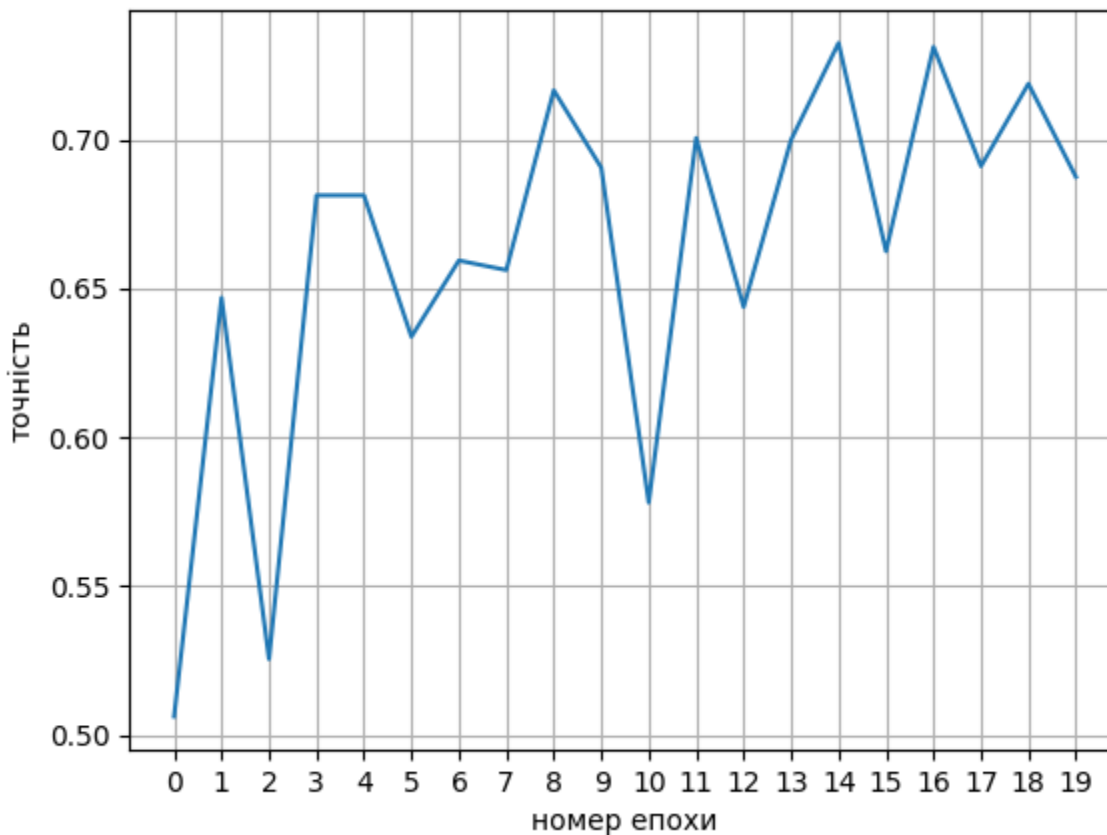
При збільшенні кількості карт у першому та другому згорткових шарах в 2 рази відбулися наступні зміни:



Час навчання:
1 год. 42 хв.

Збільшення розміру вікна згортки

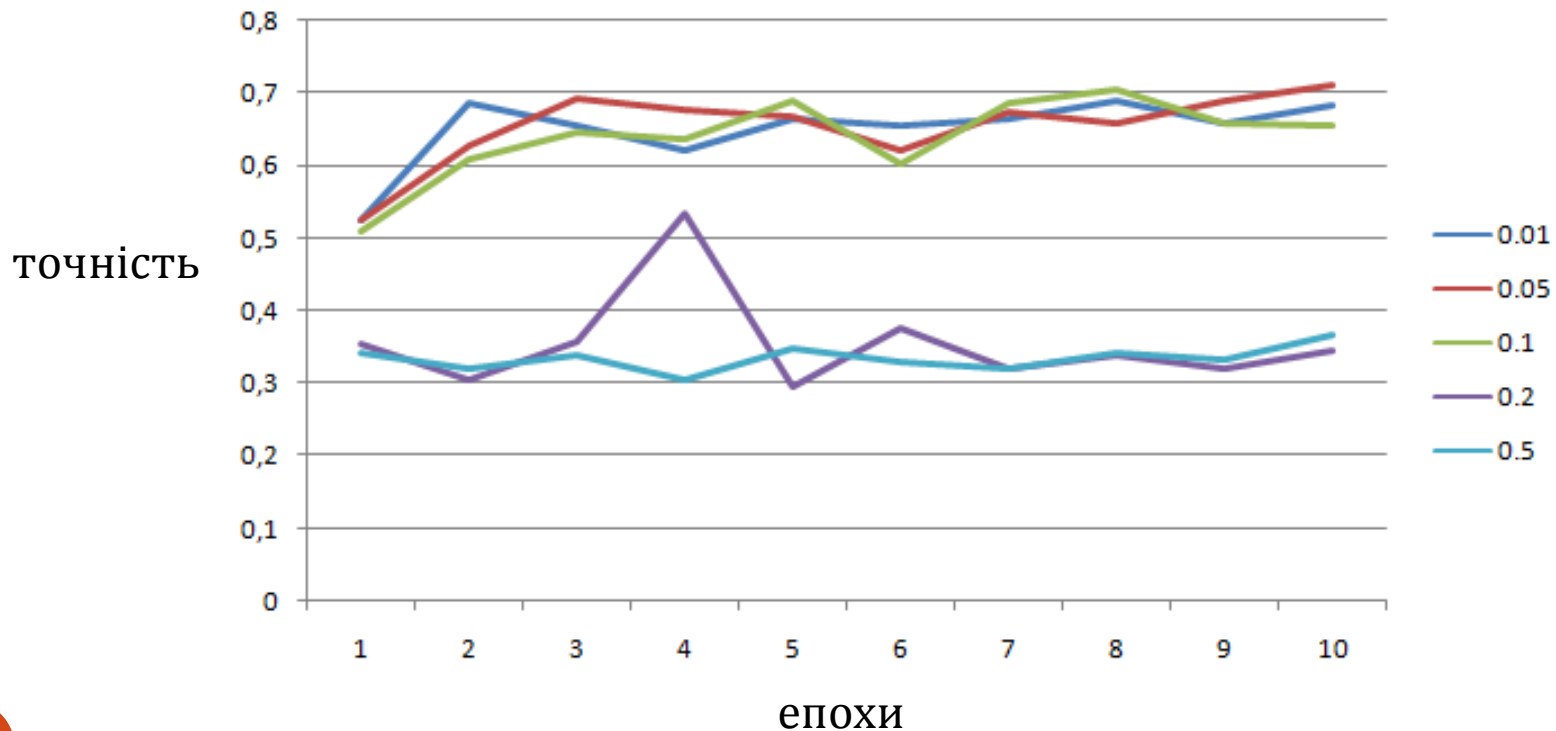
При збільшенні вікна згортки від 3x3 до 5x5 відбуваються наступні зміни:



Час навчання: 31 хв.

Зміна кроку навчання

Мережа навчалася з різними кроками навчання (0.01, 0.05, 0.1, 0.2, 0.5), після чого було отримано наступні точності класифікації:



Кінцева архітектура мережі

В результаті дослідження точності мережі в залежності від параметрів, були обрані наступні:

- Кількість шарів залишилась без змін: 8 шарів
- Кількість карт у першому, другому та третьому згорткових шарах відповідно: 16, 32, 32
- Розмір вікна згортки: 3x3
- Крок навчання мережі: 0.05

Результати роботи програми

При тестуванні навченої мережі на датасеті, що складався з 300 зображень (по 100 зображень на кожний клас), були отримані наступні результати:

	Кількість	Правильне віднесення до класу	Точність
Міфологія	100	74	74%
Портрет	100	78	78%
Пейзаж	100	82	82%

Приклад роботи програми

Пейзаж



```
1/1 [=====] - 0s  
1/1 [=====] - 0s  
[0]  
[[ 1. 0. 0.]]  
Landscape: 74.28  
Mythological: 16.55  
Portrait: 9.17
```

Міфологія



```
1/1 [=====] - 0s  
1/1 [=====] - 0s  
[1]  
[[ 0. 1. 0.]]  
Landscape: 24.72  
Mythological: 64.01  
Portrait: 11.27
```

Як можна побачити, картини жанрів пейзаж та міфологія корелюють між собою більше, ніж з картинами жанру портрет.

Приклади невірної класифікації

- пейзаж, класифікований як міфологія



Висновки

- Розкрита проблема класифікації зображень
- Проаналізована структура побудови згорткової нейронної мережі
- Проаналізовано залежність точності класифікації зображень від параметрів мережі
- Продемонстровано роботу програми при класифікації картин по жанрам, та отримано загальну точність класифікації на тестовій вибірці 79.2 %

Дякую за увагу!