

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО”



Методи розпізнавання художнього стилю нейронними мережами

Виконала студентка групи КА-41:

Палій Аліса Олегівна

Науковий керівник:
д.т.н., професор

Данилов Валерій Якович

Актуальність поставленої задачі



- Активний розвиток технологій комп'ютерного зору;
- Необхідність в розробці автоматизованих систем архівування та вилучення даних (через постійно зростаючу кількість відцифрованих картин);
- Необхідність автоматизованого оснащення художніх музеїв, навчальних закладів та виставок.

Постановка задачі



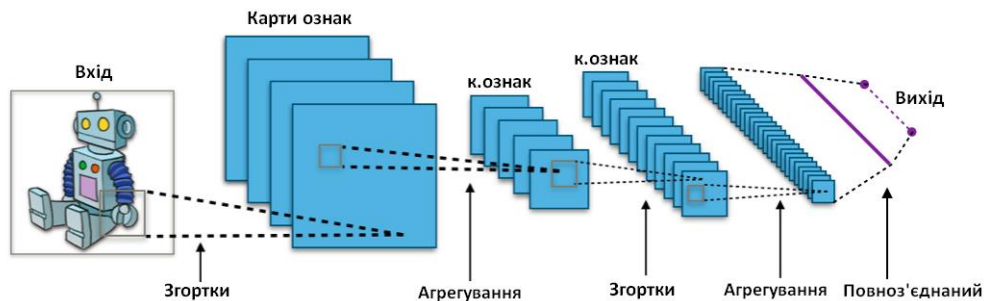
Мета роботи: аналіз існуючих методів класифікації та дослідження нейронних мереж як лідерів у вирішенні даної задачі;

Об'єкт дослідження: розпізнавання стилю картин у живописі та ознаки, що використовуються для класифікації за стилем;

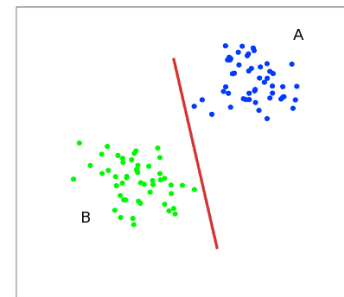
Предмет дослідження: різні види нейронних мереж та їх застосування у вирішенні задач класифікації зображень.

Основні методи класифікації зображень

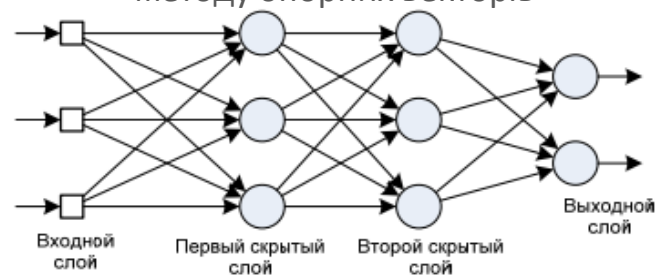
- Метод опорних векторів (Support Vector Machine);
- Багатошаровий перцептрон (Multilayer Perceptron);
- Згорткові нейронні мережі (Convolutional NN);



Приклад архітектури згорткової нейронної мережі



Схематичний приклад роботи методу опорних векторів



Приклад архітектури багатошарового перцептрону

Переваги та недоліки згорткових нейронних мереж

Переваги	Недоліки
- менша кількість параметрів, які необхідно навчити; - інваріантність до поворотів та зміщень у зображенні; - самостійне визначення ознак; - майже нема потреби у попередній обробці зображень; - можливість навчання на графічних процесорах (GPU);	- все одно досить багато змінних параметрів мережі; - немає теоретичного обґрунтування побудови архітектури;

Згорткові нейронні мережі. Основні поняття

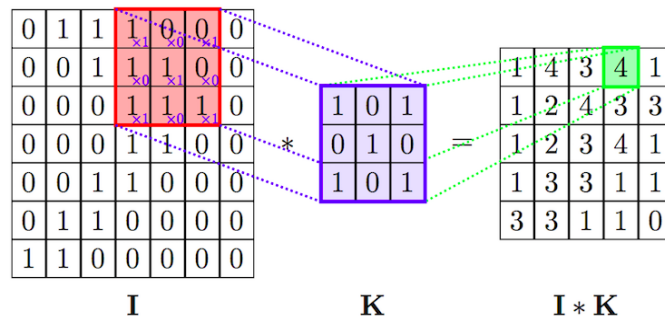
Згорткові шари виконують операцію згортки (з певним ядром) над зображенням;

$$(I * K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{ij} \times I_{x+i-1, y+j-1}$$

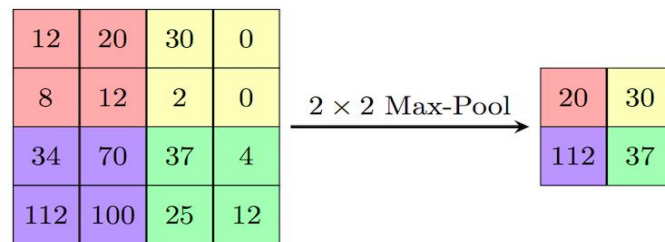
Агрегувальний шар - зменшує розмірність карт минулого шару.

Найчастіше використовується Max-Pool - вибір максимального значення.

Повнозв'язний шар - шар звичайного багатозарового персептрону і відповідає за класифікацію.



Приклад роботи згорткового шару



Приклад агрегувального шару

Огляд художніх напрямів

- **Ренесанс:** увага приділяється перспективі та об'ємності зображень завдяки техніці світлотіні;
- **Імпресіонізм:** техніка ледве помітних, тонких мазків; акцентування уваги на зміні освітлення;
- **Реалізм:** прагнення максимально правдоподібно зобразити навколишнє середовище;
- **Експресіонізм:** характерні зсуви, використання яскравих фарб і гострих контурів;
- **Кубізм:** характеризується комбінованими, ламаними предметами, що мають абстрактну форму.

В якості вхідних даних розглянуто картини датасету **Pandora18k**.

У кожному класі по 3600 зображень.

Усім зображенням надано розмір 150x150px.



Кубізм



Імпресіонізм



Експресіонізм



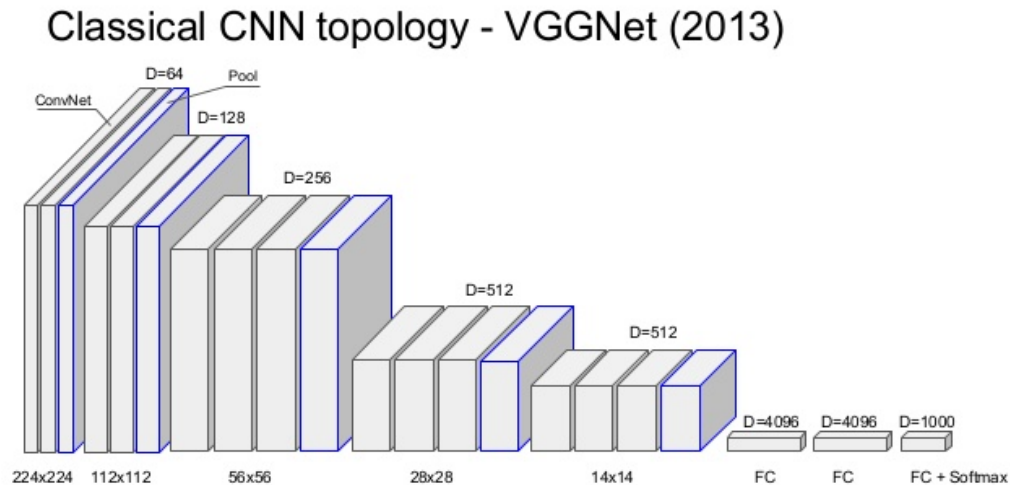
Реалізм



Ренесанс

Розробка програмного продукту

- Мова програмування - **Python**;
- Основна бібліотека - **Keras**;
- Для перенесення навчання обрано натреновану мережу - **VGG16-Net** (містить 16 згорткових шарів)

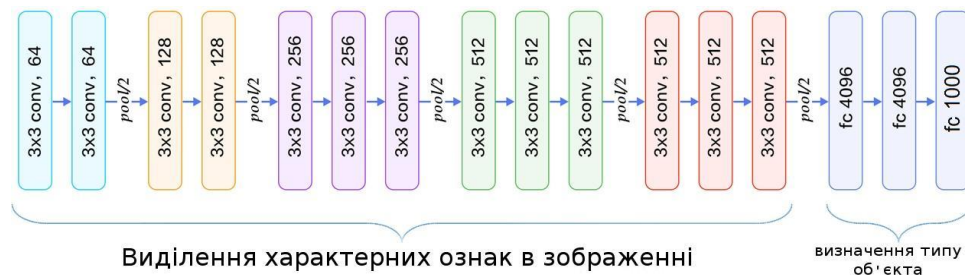


Архітектура мережі VGG16

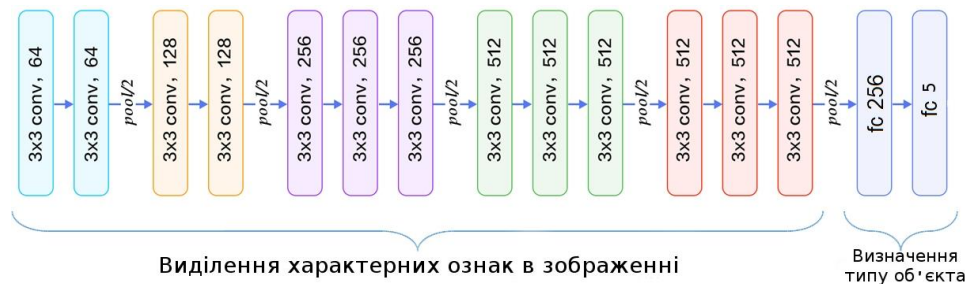
Процедура перенесення навчання

Послідовність дій:

1. Прибрано 3 повнозв'язні шари з мережі VGG16;
2. Додано 2 власні повнозв'язні шари;
3. Заблоковано навчання згорткових шарів;
4. Проведено навчання комбінованої мережі на нових даних



Архітектура нейронної мережі VGG16



Архітектура видозміненої нейронної мережі VGG16

Процес навчання мережі



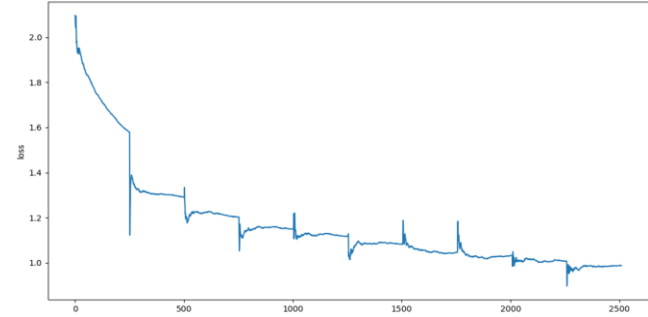
Процес зайняв приблизно 26 годин.

Точність під час навчання сягла 64.11% (від початкових 20%).

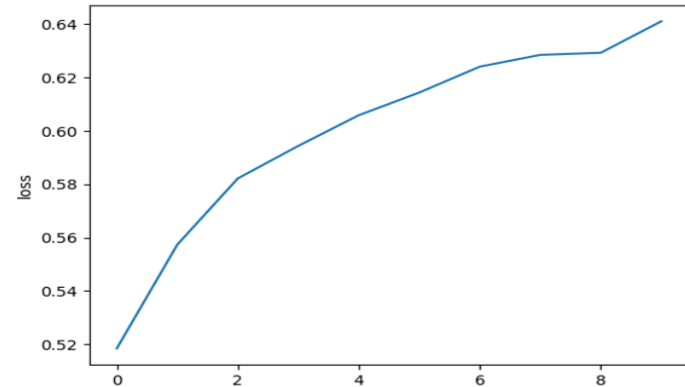
Точність на тестовій вибірці - 56.44%.

Такий результат може бути зумовлений:

- а) неякісністю вхідних даних;
- б) потенційною схожістю деяких стилей;
- в) замалим обсягом зображень для навчання.



Графік зміни помилки у процесі навчання



Графік зміни точності на кожній епісі навчання

Ефективність класифікації



Назва стилю	Частка правильної класифікації	До якого стилю з якою ймовірністю відносить при помилковій класифікації	
Імпресіонізм	60%	Кубізм 0%	Експрес.11%
		Реалізм 24%	Ренесанс 5%
Кубізм	77%	Експрес.10%	Імпрес. 1%
		Реалізм 4%	Ренесанс 8%
Реалізм	68%	Експрес. 3%	Імпрес. 7%
		Кубізм 3%	Ренесанс 19%
Експресіонізм	67%	Реалізм 2%	Імпрес. 6%
		Кубізм 17%	Ренесанс 8%
Ренесанс	61%	Реалізм 18%	Імпрес. 5%
		Кубізм 4%	Експрес. 12%

Приклад роботи програми



MainWindow

Визначення стилю картини



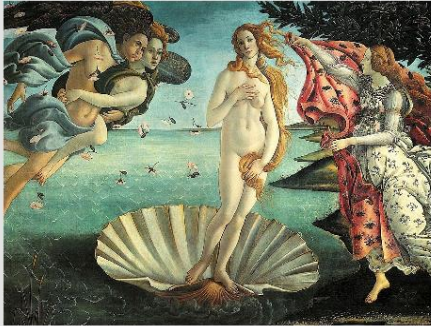
Кубізм: 0.39%
Експресіонізм: 15.87%
Імпресіонізм: 66.89%
Реалізм: 14.46%
Ренесанс: 2.37%

Отже, стиль картини:
Імпресіонізм

Обрати картину
і визначити її стиль

MainWindow

Визначення стилю картини



Кубізм: 10.1%
Експресіонізм: 28.37%
Імпресіонізм: 5.76%
Реалізм: 1.04%
Ренесанс: 54.7%

Отже, стиль картини:
Ренесанс

Обрати картину
і визначити її стиль

Приклади невірної класифікації



MainWindow

Визначення стилю картини



Кубізм: 22.7%
Експресіонізм: 25.7%
Імпресіонізм: 25.37%
Реалізм: 23.17%
Ренесанс: 3.04%

Отже, стиль картини:
Експресіонізм

Обрати картину
і визначити її стиль

MainWindow

Визначення стилю картини



Кубізм: 5.58%
Експресіонізм: 11.02%
Імпресіонізм: 19.64%
Реалізм: 23.85%
Ренесанс: 39.89%

Отже, стиль картини:
Ренесанс

Обрати картину
і визначити її стиль

Порівняння з існуючими реалізаціями задачі

Нейронна мережа	Назва роботи, автор	Точність
VGG16 із перенесенням навчання на даних Pandora18k	Мережа, розроблена у цій дипломній роботі	56,4%
ResNet із перенесенням навчання на даних WikiArt	Recognizing Art Style Automatically in painting with deep learning, Lecoutre Adrian	62,5%
InceptionV3 із перенесенням навчання на даних Pandora18k	Artistic Movement Recognition using Deep CNNs, Yu Yangyang	56,6%
Оригінальна структура ЗНМ із 8 прихованих шарів, навчання на даних з Kaggle.com	Deep Learning Approaches to Art Style Recognition in Digital Images, Johnsen R.H.	79,8%

Висновки



- Проведено аналіз методів класифікації зображень та обґрунтовано вибір згорткових мереж в якості інструменту.
- Виконано перенесення навчання нейронної мережі VGG16 із внесенням певних змін в архітектуру.
- Розроблено програмний додаток, який за допомогою перенавченої згорткової мережі виконує класифікацію зображень за художнім стилем.
- Порівняно ефективність розробленого додатку з існуючими реалізаціями задачі.

Шляхи подальшого розвитку



- Проведення навчання на більшому і якіснішому датасеті;
- Розробка нейронної мережі для класифікації більшої кількості художніх стилів;
- Розробка більш складної архітектури мережі для збільшення точності класифікації;

Дякую за увагу!