

2025

АНАЛІЗ ТА ПОРІВНЯННЯ ЕФЕКТИВНОСТІ АЛГОРИТМІВ У РОЗПІЗНАВАННІ ОБРАЗІВ

Керівник: Картавий М.О.
Виконала: студентка ФІ, ІПЗ
Карбовська З.Ю.

Національний університет
«Києво-Могилянська академія»

Актуальність та мета дослідження

Актуальність:

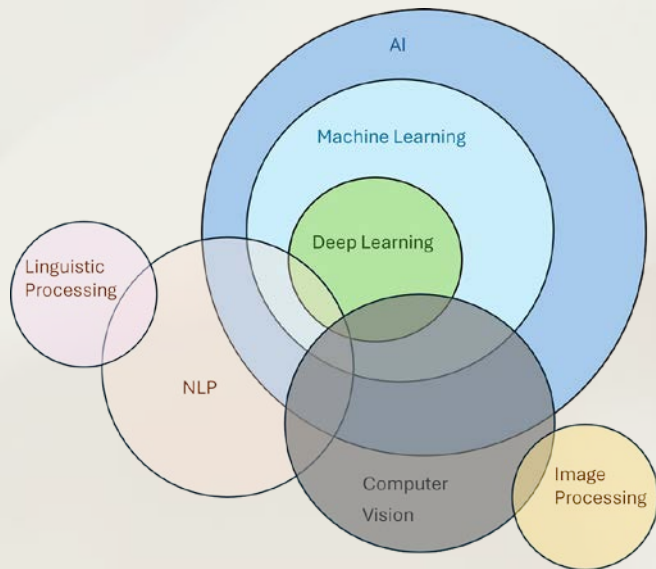
- постійний розвиток ШІ, зокрема напрямку “Комп’ютерного зору”, який працює над завданням “Розпізнавання образів”;
- Застосування всюди: такі непомітні, але важливі камери спостереження

Мета:

Розглянути підходи у розпізнаванні образів: традиційний і на базі згорткових нейронних мереж, їх алгоритми, а також реалізувати вебзастосунок для демонстрації двох підходів.

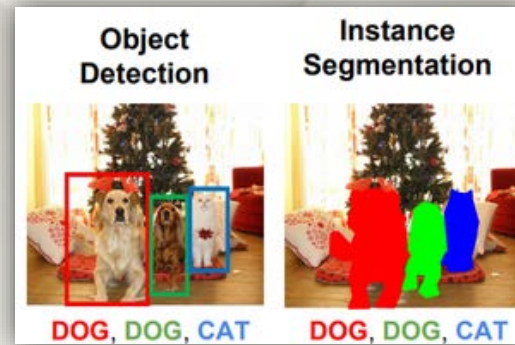
Розпізнавання образів

Computer Vision прагне відповісти на запитання "Який є об'єкт і де він?".



Завдання комп'ютерного зору:

- розпізнавання (класифікація);
- виявлення об'єктів;
- сегментація зображень;

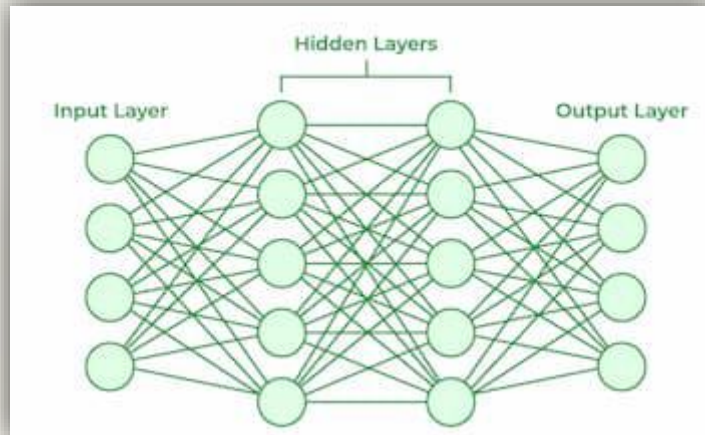


Підходи у розпізнаванні образів

У Computer Vision існує два основні підходи для розпізнавання образів. Із використанням:

- Нейронних мереж (CNN)

- Класичних (традиційних) алгоритмів



Convolutional layers → pooling
→
fully connected

Етапи:

1. Наповнення датасету
2. Витягнення ознак
3. Класифікація

CNN

- Навчаються на великих наборах даних;
- Працюють за принципом, що імітує роботу людського мозку, обробляючи вхідні дані через багат шарові структури;
- Пренавчені моделі

LeNet

ZFNet

VGG

GoogleNet

ResNet

Класичні алгоритми

- переважно потребують два кроки: *витягнення ознак* + навчання моделі для *класифікації*

1. *Метод шаблонного співставлення* — поєднує в собі два кроки. Працює, як «збіг за зразком»;
2. *Методи на основі ключових точок (SIFT, SURF, **ORB**)* — витягують ключові точки
3. *Алгоритми за контурами або геометричними формами (Canny, Sobel/Prewitt, LoG, Hough Transform)* — виділяють краї зображення, потім виявляють геометричні форми

Класичні алгоритми

Алгоритми класифікації:

SVM (support vector machine) — один з найпопулярніших;

K-NN (k-nearest neighbors) — простий, але повільний при великій вибірці;

Random Forest — стійкий до шуму, добре працює з ознаками;

Naive Bayes – ймовірисний класифікатор, який припускає, що одна властивість (ознака) в моделі є незалежною від інших, але всі вони мають суттєву вагу;

Реалізація

Датасет

Ведмеді (156)

Коти (157)

Собаки (156)

Мавпи (152)

Інше (190)

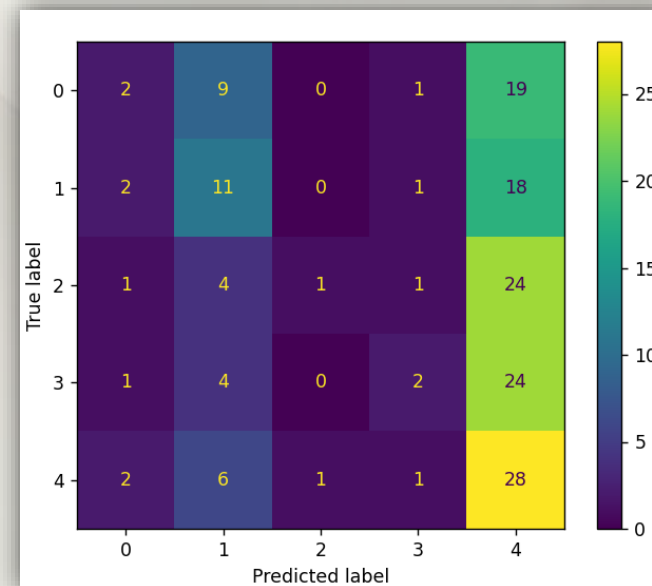
1. *GoogleNet* – претренована модель на датасеті ImageNet (1000 класів);
2. Зібрати датасет зображень → *ORB* → *SVM*

Реалізація класичного алгоритму

	precision	recall	f1-score	support
dogs	0.25	0.06	0.10	31
cats	0.32	0.34	0.33	32
bears	0.50	0.03	0.06	31
monkeys	0.33	0.06	0.11	31
other	0.25	0.74	0.37	38
accuracy			0.27	163
macro avg	0.33	0.25	0.20	163
weighted avg	0.33	0.27	0.20	163

Accuracy: 0.26993865030674846

1. Precision (точність): скільки передбачень певного класу були правильними.
2. Recall (повнота): скільки реальних об'єктів класу модель правильно виявила.
3. F1-score: середнє гармонійне між precision та recall.
4. Support



Вебзастосунок

streamlit


Image Recognition: CNN vs Classical SVM

Upload an image to compare recognition methods

Drag and drop file here
Limit 200MB per file • JPG, JPEG, PNG

Browse files

00000001_012.jpg 303.1KB



Uploaded Image

InceptionV3 (CNN)	SVM + ORB
Prediction: tabby	Prediction: cats
Confidence:	Confidence:
<div style="width: 43%;"><div style="width: 43%;"></div></div> 0.43	<div style="width: 36%;"><div style="width: 36%;"></div></div> 0.36
All Probabilities:	All Probabilities:
<ul style="list-style-type: none">1. tabby : 0.43272. tiger_cat : 0.22763. Egyptian_cat : 0.12324. Persian_cat : 0.02195. lynx : 0.0083	<ul style="list-style-type: none">1. cats: 0.36312. other: 0.21233. dogs: 0.21124. monkeys: 0.14545. bears: 0.0680

Спостереження

1. Дуже важливо коректно зібрати датасет. Він має максимально охоплювати відповідний клас. Якщо датасетів декілька, необхідно, щоб вони були приблизно однакових розмірів;
2. Збільшення кількості зображень у датасеті не означає підвищення точності алгоритму;
3. Зміна датасету однієї категорії впливає на результати інших категорій;
4. Видалення заднього фону може погіршити результат

	precision	recall	f1-score	support
dogs	0.17	0.14	0.15	14
cats	0.44	0.54	0.48	28
other	0.84	1.00	0.92	38
bears	0.57	0.31	0.40	13
monkeys	0.69	0.59	0.63	34
accuracy			0.62	127
macro avg	0.54	0.51	0.52	127
weighted avg	0.61	0.62	0.61	127

Accuracy: 0.6220472440944882



SVM + ORB

Prediction: other

Confidence:

0.31

All Probabilities:

1. other: 0.3082
2. bears: 0.2121
3. cats: 0.2046
4. dogs: 0.1741
5. monkeys: 0.1010

Результати

На основі проведеного аналізу та тестування можна зробити висновок, що CNN залишається найкращим інструментом для задач розпізнавання зображень. Вони демонструють стабільну роботу та забезпечують високу точність. У той же час класичні методи залишаються актуальними для швидкої реалізації базових систем, для яких важливі обмежені ресурси, простота або прозорість роботи алгоритму.

Для майбутнього покращення класичного методу можна використати інший алгоритм для витягнення ознак, а для CNN-моделі — донавчити на своїх даних.

Дякую за увагу!