

ЗГОРТКОВА НЕЙРОННА МЕРЕЖА ДЛЯ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

Згорткова нейронна мережа (CNN) – це особливий вид нейронної мережі, який значно зменшує кількість параметрів у глибокій нейронній мережі з багатьма одиницями, не надто втрачаючи якість моделі. CNN застосовується в обробці зображень і тексту, та це одна з основних категорій для розпізнавання зображень, класифікації зображень, виявлення об'єктів, розпізнавання обличчя тощо.

За рахунок застосування спеціальної операції, власне згортки, згорткова нейронна мережа дозволяє водночас зменшити кількість інформації, що зберігається в пам'яті інформації, за рахунок чого краще справляється з картинками більш високої роздільної здатності, і виділити опорні ознаки зображення, такі як ребра, контури або грані. На наступному рівні обробки з цих ребер і граней можна розпізнати повторювані фрагменти текстур, які далі можуть скластися в фрагменти зображення.

Архітектура згорткової нейронної мережі.

CNN зазвичай має три рівні: згортковий рівень, рівень об'єднання та повністю зв'язаний рівень (рис 1.).

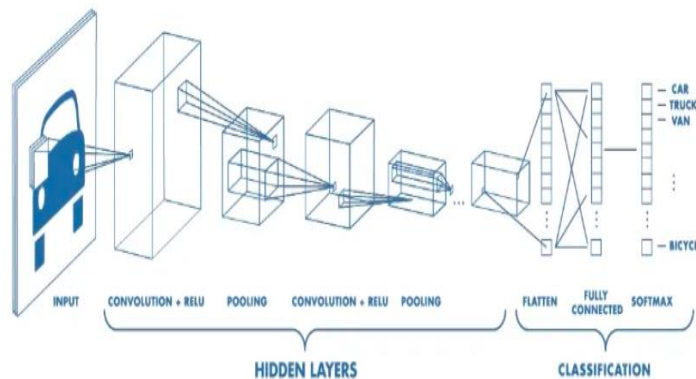


Рис. 1. Типова структура роботи. Архітектура CNN

На сьогоднішній час можна виділити наступні згорткові нейронні мережі, які використовуються:

1. Виявлення об'єктів. Завдяки CNN тепер є такі складні моделі, як R-CNN, Fast R-CNN і Faster R-CNN, які є основною системою для багатьох моделей виявлення об'єктів, що застосовуються в автономних транспортних засобах, розпізнавання обличчя тощо.
2. Семантична сегментація – це мережа для глибокого аналізу на основі CNN.
3. Субтитри до зображень: CNN використовуються з рекурентними нейронними мережами для написання титрів до зображень і відео. Це можна використовувати для багатьох програм, наприклад для розпізнавання активності або опису відео та зображень для людей із вадами зору.

Існують різні архітектури CNN, які відіграють ключову роль у побудові алгоритмів, які забезпечують і будуть підтримувати ШІ в цілому в майбутньому. Деякі з них: LeNet, AlexNet, VGGNet, GoogLeNet, ResNet, ZFNet.

Класична згорткова нейронна мережа:

1. Згортковий шар.
2. Операція активації після кожного згорткового шару.
3. Рівень об'єднання, особливо рівень максимального об'єднання, а також інші на основі вимог.
4. Нарешті повністю підключений рівень.

Виявлення об'єктів

CNN може виявляти об'єкти на зображеннях у реальному часі, а також маркувати та класифікувати їх. Тому ця техніка широко використовується в автоматизованих транспортних засобах [1]. Це також дозволяє розумним будинкам і пішоходам розпізнавати обличчя власника автомобіля. Він також використовується в системах спостереження на основі ШІ для виявлення та позначення об'єктів.

Виявлені області (рис. 2) об'єднуються за різними проміжками кольорів та показниками подібності. Кожна пропозиція відеопотоку змінюється так, щоб відповідати входу CNN, з якого отримується векторні ознаки розміру 4096 bit. Вектор ознак надходить в декілька класифікаторів для отримання імовірностей, що належать до кожного класу.

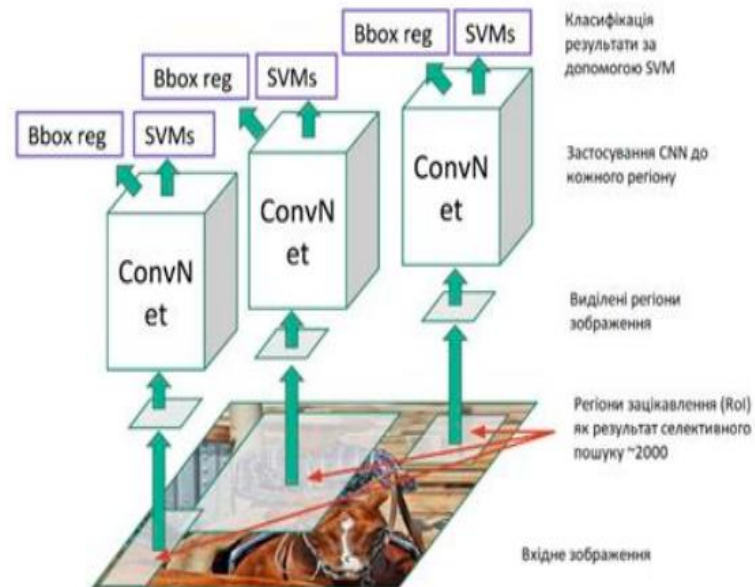


Рис. 2. Типова структура роботи R-CNN: на основі [2]

CNN на сьогодні широко використовується в області нейронних мереж. Це відбувається переважно для вирішення завдань комп'ютерного зору, хоча може застосовуватися також для роботи з аудіо і будь-якими даними, які можна представити у вигляді матриць.

Список використаних джерел

1. A Comprehensive Guide to Convolutional Neural Networks-the ELI5 way. URL :<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-theeli5-way-3bd2b1164a53>.
2. R. Girshick, «Fast R-CNN» 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 2015, pp. 1440-1448.