

«Розпізнавання об'єктів із використанням нейронних мереж»

Вінницький національний технічний університет

Анотація

Дана наукова стаття розглядає використання різних архітектур нейронних мереж, таких як Convolutional Neural Networks (CNNs), Region-based CNNs (R-CNNs) та Single Shot Multibox Detector (SSD) для задачі розпізнавання об'єктів у просторі. Запропоновані методи вивчаються та порівнюються з точки зору точності, ефективності та придатності для різних застосувань.

Ключові слова: нейронні мережі, розпізнавання об'єктів, комп'ютерний зір, CNNs, R-CNNs, SSD

Abstract

This scientific article considers the use of different neural network architectures, such as Convolutional Neural Networks (CNNs), Region-based CNNs (R-CNNs) and Single Shot Multibox Detector (SSD) for the task of object recognition. objects in space. The proposed methods are studied and compared in terms of accuracy, efficiency and suitability for different applications.

Keywords: neural networks, object recognition, computer vision, CNNs, R-CNNs, SSD

Вступ

Дана наукова стаття розглядає використання різних архітектур нейронних мереж, таких як Convolutional Neural Networks (CNNs), Region-based CNNs (R-CNNs) та Single Shot Multibox Detector (SSD), для задачі розпізнавання об'єктів у просторі. Запропоновані методи вивчаються та порівнюються з точки зору точності, ефективності та придатності для різних застосувань.

Основна частина

Convolutional Neural Networks (CNNs) є основою для багатьох задач комп'ютерного зору, включаючи розпізнавання об'єктів у просторі. Вони виникли з потреби розпізнавання образів в зображеннях та виявлення об'єктів на основі їхньої структури та ознак [1, 2].

Основною особливістю CNNs є використання конволюційних шарів, які дозволяють автоматично виявляти локальні ознаки в зображеннях. Архітектура включає такі ключові елементи, як конволюційні шари, шари підбору, повні з'єднані шари.

Процес навчання включає подачу тренувальних зображень у мережу, визначення втрат та використання алгоритму оптимізації (наприклад, градієнтного спуску) для оновлення ваг та зменшення втрат [3].

У завданнях розпізнавання об'єктів у тривимірному просторі CNNs можуть виявляти та класифікувати об'єкти на зображеннях з використанням технік, таких як слідкування за об'єктами, виявлення об'єктів у реальному часі та оцінка розташування об'єктів.

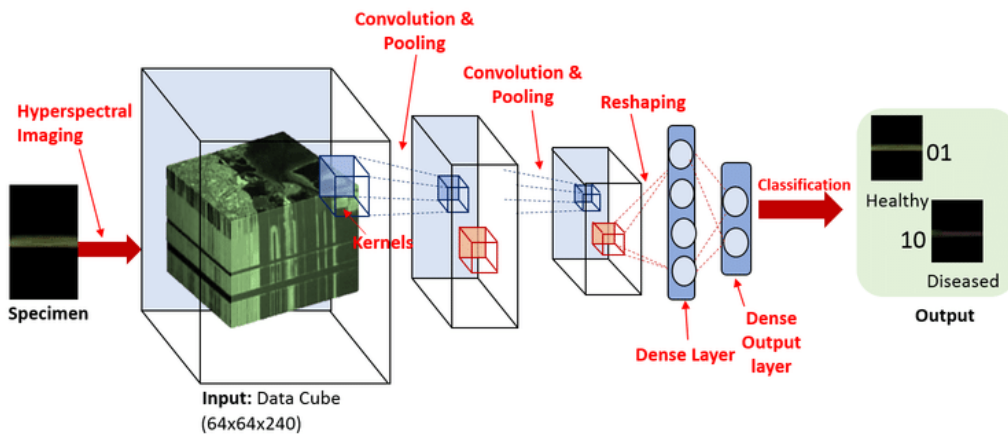


Рисунок 1 – Розпізнавання об'єкту із застосуванням CNN

Region-based Convolutional Neural Networks (R-CNNs) є еволюційним кроком у напрямку розпізнавання об'єктів, особливо в ситуаціях, коли точність локалізації об'єктів є ключовою. R-CNNs вперше були представлені як відповідь на обмежену точність та обробку зображень у попередніх архітектурах [4].

Основна ідея R-CNN полягає у визначенні потенційних регіонів, де можуть знаходитися об'єкти, та подальшому виявленні та класифікації об'єктів у цих регіонах. Зображення подається через області пропозицій, які генеруються заздалегідь за допомогою методів, таких як Selective Search. Ці області потенційно містять об'єкти, і їхній набір передається в мережу для подальшого аналізу [5, 6].

R-CNNs включають дві основні частини - модель виявлення областей та модель класифікації. Ці дві частини навчаються спільно, що дозволяє визначати області та класифікувати об'єкти у цих областях. Під час тренування моделі R-CNN, використовуються набори даних, які включають інформацію про регіони, де мають знаходитися об'єкти, та відповідні мітки класів.

Ці нейронні мережі знаходять застосування в сферах відстеження об'єктів, виявлення обличчя та медичної діагностики до реалізації безпілотних транспортних засобів, де точність та ефективність грають ключову роль в успішному функціонуванні систем розпізнавання об'єктів.

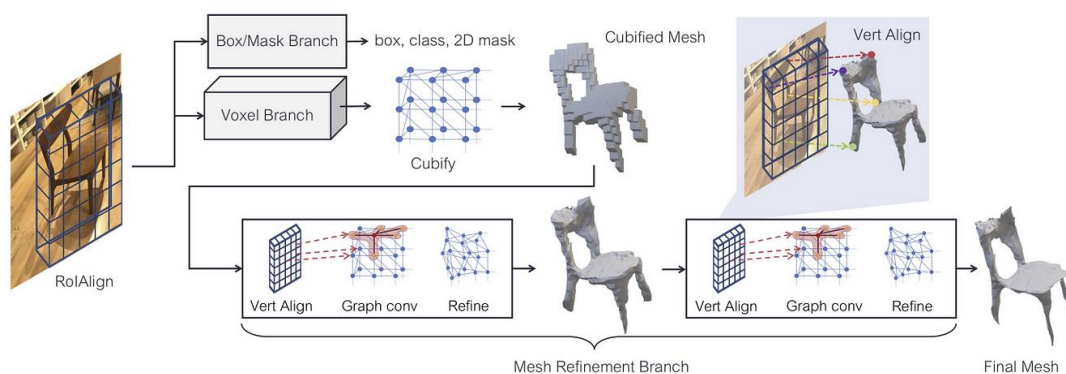


Рисунок 2 – Розпізнавання об'єкту із застосуванням R-CNN

Single Shot Multibox Detector (SSD) є підходом до завдань розпізнавання об'єктів, розробленим з метою ефективності та точності в реальному часі [7].

Основна ідея SSD полягає в тому, щоб в одному проході мережі визначати клас та розташування кількох об'єктів різних розмірів та аспектичних співвідношень. SSD використовує шари виявлення на різних рівнях абстракції, дозволяючи моделі адаптуватися до об'єктів різних розмірів. Кожен шар відповідає за виявлення об'єктів певного розміру [8].

Також SSD включає у себе механізм для оцінювання якості розташування об'єктів та використовує аугментацію для удосконалення точності розташування об'єктів, зокрема,

використовуючи додаткові ознаки, які розширюють область виявлення. Це дозволяє враховувати не тільки факт виявлення, а й точність локалізації. Таким чином SSD є дієвим та ефективним рішенням для задач розпізнавання об'єктів, де важлива швидкість та точність виявлення [9].

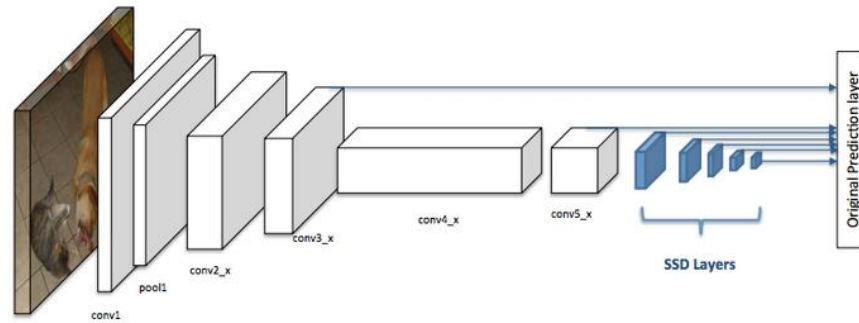


Рисунок 2 – Розпізнавання об'єкту із застосуванням SSD

У порівнянні з CNNs, R-CNNs та SSD вибір залежить від конкретних вимог завдань. CNNs є стандартом з точністю, R-CNNs забезпечують високу точність локалізації, але можуть бути повільні, тоді як SSD найбільш підходить для задач в реальному часі, де важлива швидкість та точність виявлення об'єктів на різних масштабах.

Висновки

У цій статті були розглянуті та порівняні три ключові архітектури для розпізнавання об'єктів у просторі: Convolutional Neural Networks (CNNs), Region-based CNNs (R-CNNs) та Single Shot MultiBox Detector (SSD). Кожен з цих підходів відзначається унікальними особливостями та властивостями, що робить їх придатними для різних застосувань та завдань.

У кожній архітектурі важливо враховувати конкретні вимоги завдань та контекст використання. Обрана архітектура повинна забезпечувати оптимальний баланс між точністю, швидкістю та локалізацією для успішного вирішення конкретної задачі розпізнавання об'єктів у просторі.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Prof. Sujata Bhairnallykar¹, A. P. (2020). Convolutional Neural Network (CNN) for Image Detection. International Research Journal of Engineering and Technology (IRJET), 1239-1243.
2. Baohua Qiang¹, R. C. (2020). Convolutional Neural Networks-Based Object. Sensor, 2-14
3. Szegedy, C. (June 2015). Computer Vision and Pattern Recognition. IEEE Computer Society Conference, 27-30.
4. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99). doi: 10.1109/TPAMI.2015.2437388
5. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks." In Advances in Neural Information Processing Systems, pp. 91-99, 2015.
6. Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 (2015).
7. Shi, Wei et al. Single-shot detector with enriched semantics for PCB tiny defect detection. The Journal of Engineering(2020), 2020 (13):366
8. Weilin Cong, Rana Forsati, Mahmut Kandemir, and Mehrdad Mahdavi. 2020. Minimal Variance Sampling with Provable Guarantees for Fast Training of Graph Neural Networks. 1393–1403.
9. George Karypis and Vipin Kumar. 1998. A fast and high quality multilevel scheme for partitioning irregular graphs. SIAM Journal on Scientific Computing 20 (1998), 359–392.

Максим Андрійович Фурман – студент групи ІКІ-23М, факультет інформаційних технологій і комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, jgnice8@gmail.com.

Крупельницький Леонід Віталійович – канд. техн. наук, доцент, доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, м. Вінниця, krupost@gmail.com.

Furman Maksym A. – student of group 1KI – 23M, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, jgnice8@gmail.com.

Krupelnitskyi, Leonid V. – Cand. Sc. (Eng.), Assistant Professor of the Computer Techniques Chair, Vinnytsia National Technical University, Vinnytsia, krupost@gmail.com.