

Використання штучних нейронних мереж для розпізнавання тексту на графічних зображеннях

Бандровський Генріх

Кафедра дискретного аналізу та інтелектуальних систем
ЛНУ імені Івана Франка, Львів, Україна
henrikh.bandrovskiy@gmail.com

The problem of natural scene image recognition has recently received much attention especially concerning the usage of neural networks in the context of machine learning. In this paper contemporary approaches towards pattern recognition problem are studied. Of special concern are the disadvantages of these approaches. The usage of convolutional neural network is analyzed.

Ключові слова: наука про дані, машинне навчання, машинний зір, розпізнавання образів, штучні нейронні мережі, штучний інтелект.

ВСТУП

Розпізнавання символів і тексту як частина задачі розпізнавання за допомогою штучних нейронних мереж є однією з провідних задач у сфері науки про дані, а саме в галузях машинного навчання та машинного зору.

Здатність нейронних мереж навчатися є однією з основних можливостей, яка робить їх більш гнучкими і незалежними в порівнянні зі статистичними підходами в розв'язуванні задач класифікації.

Зчитування тексту з цифрових зображень та класифікація об'єктів становлять завдання, де широко застосовуються нейронні мережі. Варто зазначити, що при роботі з цифровими зображеннями існує випадок задачі складних графічних даних, коли окрім розпізнавання тексту необхідно виконати локалізацію і сегментацію тексту та об'єктів на зображенні.

В даній роботі розглянуто основні концепції використання згорткових нейронних мереж для розпізнавання.

МАШИННЕ НАВЧАННЯ ТА РОЗПІЗНАВАННЯ ТЕКСТУ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

Машинне навчання можна виділити як підрозділ комп'ютерних наук, який дає комп'ютерам здатність до навчання, не будучи явно запрограмованими.

Дослідження в галузі машинного навчання присвячені вивченню і побудові алгоритмів, які можуть «навчатися» з даних так званого тренувального набору і виконувати аналіз та робити прогнози.

В контексті задачі розпізнавання варто розглянути один із багатьох алгоритмів машинного навчання – **алгоритм класифікації**.

Даний алгоритм здатний розподіляти дані по різним конкретним групам. Варто відзначити універсальність алгоритмів машинного навчання. Наприклад, один алгоритм, який використовується для розпізнавання символів, написаних як різними шрифтами, так і від руки, може також аналогічно використовуватись для класифікації електронної пошти з метою групування повідомлень на «нормальні» повідомлення і згенерований «спам», чи тенційно класифікувати повідомлення за їх емоційним забарвленням на «позитивні», «нейтральні» та «негативні».

Оглянемо алгоритм класифікації на прикладі розпізнавання цифри від «8» при вхідних даних у формі графічного зображення. Зображення перетворюється у відповідну до потреби мережі форму і до нього застосовується певний алгоритм, який на виході має два варіанти: «це 8» чи «це не 8».

Розглянемо мережу для класифікації написаної від руки цифри «8». Машинне навчання дає точні результати при наявності достатньо великої кількості даних для тренування. Приклад такої вибірки даних зображений на Рис. 1:

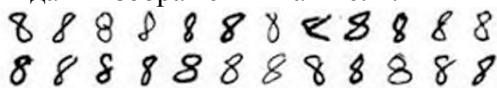
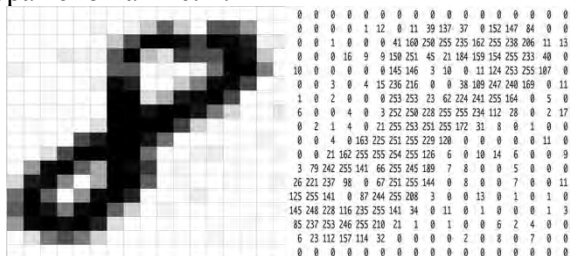


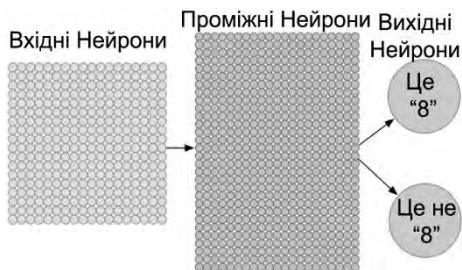
Рис. 1. Вибірка даних

При отриманні вхідного зображення воно переводиться в цифрове представлення, як зображено на Рис. 2.



- Представлення цифри «8» масивом розміру 18x18

Нейронна мережа, яка класифікувала би дану цифру на зображенні, мала би структуру, зображену на Рис. 3.



- Мережа для розпізнавання цифри «8»

Проте дана мережа дає точні результати лише у випадку простих зображень, де цифра знаходиться рівно по центру зображення. У випадку найменшого зсуву цифри на зображенні тотально зменшується точність даної мережі.

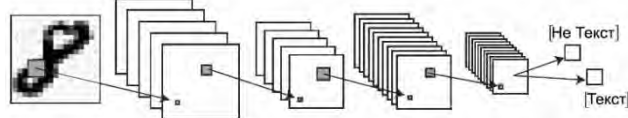
Для збереження точності розпізнавання при використанні даної мережі можна застосувати наступні підходи:

Метод 1. Пошук за допомогою рухомого фрагмента, який буде проходити по всьому зображенню і сканувати кожен ділянку на наявність цифри «8».

Метод 2. При навчанні збільшити кількість вхідних даних методом копіювання та

модифікації вибірки зміщуючи цифру та змінюючи її розмір [2].

Як альтернатива даним методам може бути використана **згорткова нейронна мережа**, яка має спеціальну архітектуру, запропоновану Яном Лекуном і яка була створена для розпізнавання зображень [2]. Ця мережа є варіацією багат шарових перцептронів та складається з різних комбінацій згорткових, підвибіркових та повноз'єднаних шарів. Принцип роботи даної мережі полягає в розбитті зображення на велику кількість фрагментів подібно до пошуку в першому запропонованому методі. Дана мережа зображена на Рис. 4. Кожен фрагмент подається для окремого аналізу. Результати, які потенційно можуть бути текстом для кожного фрагменту, зберігаються в новий масив. Далі за вибраним правилом зменшуємо кількість вибірок результатів.



- Згорткова нейронна мережа

Отриманий масив даних пропускаємо через повноз'єднаний шар мережі і отримуємо прогноз.

ЛІТЕРАТУРА

- [1] Ch. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.
- [2] Y. LeCun and Y. Bengio "Word-level training of a handwritten word recognizer based on convolutional neural networks" in IAPR (Eds), Proc. of the International Conference on Pattern Recognition, II:88-92, IEEE, Jerusalem, October 1994.