

НЕЙРОМЕРЕЖЕВА АРХІТЕКТУРА ДЛЯ ВИЗНАЧЕННЯ СТАТІ ТА ВІКУ ЛЮДИНИ ЗА ЗОБРАЖЕННЯМ ОБЛИЧЧЯ

Вінницький національний технічний університет

Анотація

Запропоновано архітектуру згорткової нейронної мережі, що визначає стать та вік людини як за одним зображенням обличчя, так і в режимі реального часу. Дана мережа може використовуватись для різних цілей, зокрема, аналізу аудиторії у різних підприємствах чи закладах, заборона користувачам веб-сервісів до деяких товарів або обирати контекстну рекламу на вебсайтах, в залежності від віку та статті користувача.

Ключові слова: згорткова нейронна мережа, визначення статі та віку, аналіз аудиторії, контекстна реклама.

Abstract

The architecture of a convolutional neural network that determines a person's gender and age based on a single face image and in real time is proposed. This network can be used for various purposes, in particular, audience analysis in different enterprises or institutions, banning users of web services from certain products or choosing contextual advertising on websites, depending on the age and article of the user.

Keywords: convolutional neural network, determines the sex and age, analyzing the audience, content-targeted advertising.

Вступ

Для вирішення проблеми класифікації зображень було створено велику кількість методів. Ці методи зазвичай спиралися на відмінності у розмірах рис обличчя та дескрипторах обличчя, які не здатні впоратися з різним ступенем варіацій, що спостерігаються в цих складних незалежних умовах. Зображення в цих категоріях мають деякі відмінності у зовнішньому вигляді, шумі, та освітленні, що може вплинути на здатність цих методів комп'ютерного зору точно класифікувати стать і вік обличчя людини на зображенні.

Для прогнозування статі та віку людини дослідники придумали різні алгоритми, використовуючи концепції класифікації та машинного навчання. Більшість алгоритмів примітивного типу використовуються для отримання багатьох вторинних алгоритмів з удосконаленнями. Алгоритми «Fisherfaces» і «Eigenfaces» вважаються примітивними. Крім того, глибокі згорткові нейронні мережі (CNN) є ще одним методом, який можна використовувати [1].

Результати дослідження

Згорткова нейронна мережа — це клас нейронних мереж, які можуть розпізнавати й класифікувати окремі ознаки зображень і широко використовуються для аналізу візуальних зображень. Є дві основні частини архітектури ЗНМ:

1. Інструмент згортки, який відокремлює та ідентифікує різні ознаки зображення для аналізу в процесі, який називається Feature Extraction (вилученням ознак).

2. Повністю зв'язаний шар, який використовує вихідні дані процесу згортки та прогнозує клас зображення на основі ознак, витягнутих на попередніх етапах.

Успіх застосування згорткових нейронних мереж до класифікації зображень привів до безлічі спроб використовувати даний метод до інших місій. Згорткова нейронна мережа, окрім вхідного і вихідного шару, зазвичай являє собою чергування згорткових шарів (convolution layers), субдескриптізуючих шарів (subsampling layers) і при наявності повнозв'язних шарів (fully-connected layer) на виході.

Всі три види шарів можуть чергуватися в довільному порядку [2]. У згортковому шарі нейрони, які використовують одні і ті ж ваги, об'єднуються в карти ознак (feature maps), а кожен нейрон карти ознак пов'язаний з частиною нейронів попереднього шару. При обчисленні мережі виходить, що кожен нейрон виконує згортку деякої області попереднього шару (яка визначається безліччю нейронів, пов'язаних з даними нейроном). Всі три види шарів можуть чергуватися в довільному порядку [3]. У згортковому шарі нейрони, які використовують одні і ті ж ваги, об'єднуються в карти ознак (feature maps), а кожен нейрон карти ознак пов'язаний з частиною нейронів попереднього шару. При

обчисленні мережі виходить, що кожен нейрон виконує згортку деякої області попереднього шару (яка визначається безліччю нейронів, пов'язаних з даними нейроном).

На рис. 1 зображено загальний вигляд архітектури згорткової мережі.

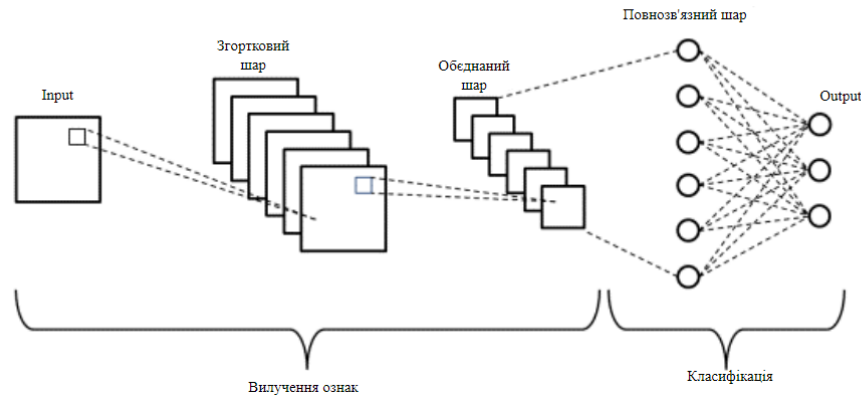


Рис 1. Загальний вигляд архітектури згорткової мережі

Вхідний шар враховує двовимірну топологію зображень і складається з декількох карт (матриць), карта може бути одна, в тому випадку, якщо зображення представлено в відтінках сірого, інакше їх 3, де кожна карта відповідає зображенню з конкретним каналом (червоним, синім і зеленим).

Найбільш популярною є архітектура ЗНМ є AlexNet. Ця архітектура спочатку була розроблений для категоризації рукописних цифр від 0 до 9 набору даних MNIST. Архітектура моделі AlexNet зображена на рисунку. 2.

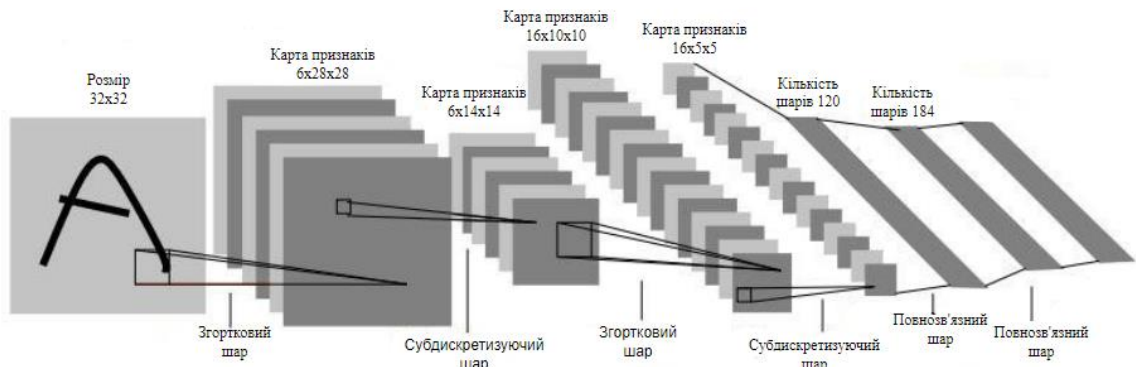


Рис. 2. Архітектура AlexNet згорткової нейронної мережі

Дана архітектура згорткової нейронної мережі містить набагато більше фільтрів, що дозволяє класифікувати набагато більше об'єктів. Складається з 5 згорткових шарів та 3 повноз'язних. Це неймовірно потужна модель, здатна досягати високої точності на дуже складних наборах даних. Однак видалення будь-якого з згорткових шарів різко погіршить продуктивність. AlexNet є провідною архітектурою для будь-якого завдання виявлення об'єктів і може мати величезне застосування в секторі комп'ютерного зору проблем штучного інтелекту. У майбутньому AlexNet може використовуватися більше, ніж CNN для завдань зображення.

Висновки

В результаті роботи, було запропоновано архітектуру AlexNet для побудови згорткової нейронної мережі, зокрема, для визначення статі та віку людини за зображенням обличчя. Встановлено, що запропонована архітектура дозволяє визначати додаткову інформацію про користувача, при мінімальних його діях. Даний аналіз можна використовувати у різних сферах ринкової діяльності.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Jason Brownlee. What is Deep Learning?. MachineLearningMastery.com. URL: <https://machinelearningmastery.com/what-is-deep-learning/> (дата звернення: 15.11.2023).
2. Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick. Mask-RCNN. arXiv:1703.06870v3 [cs.CV] 24 Jan 2018.
3. Neurocomputer architecture based on spiking neural network and its optoelectronic implementation / Oleh K. Kolesnytskyj; Vladislav V. Kutsman; Krzysztof Skorupski; Mukaddas Arshidinova, Proc. SPIE 11176, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2019, 1117609 (6 November 2019); doi: 10.1117/12.2536607.

Гнатенко Роман Іванович – студент групи ЗКН-22м, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, email: hnatenkoroman@gmail.com.

Колесницький Олег Костянтинович – доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, Вінниця, e-mail: kolesnytskyi@vntu.edu.ua.

Hnatenko Roman Ivanovich – Department Intelligent Information Technology and Automation, Vinnytsia National Technical University, Vinnytsia, email: hnatenkoroman@gmail.com.

Kolesnytskyy Oleh Kostiantynovych — Associate Professor of Computer Science Dpt, Vinnitsa National Technical University, Vinnitsa, e-mail: kolesnytskyi@vntu.edu.ua.