

УДК 004.89

**КЛАСИФІКАЦІЯ ОБ'ЄКТІВ НА ЗОБРАЖЕННІ ЗА ДОПОМОГОЮ НЕЙРОННОЇ  
МЕРЕЖІ**  
**КЛАССИФИКАЦИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ С ПОМОЩЬЮ НЕЙРОННОЙ  
СЕТИ**  
**CLASSIFICATION OF OBJECTS IN IMAGE USING A NEURAL NETWORK**

Науковий керівник - доц. каф. інформаційних систем, канд. техн. наук Шпинковський О.  
А., Шпинковский А. А., Shpinkovski O. A.

Студентка групи AI-171 – То Тхі Ха Мі, То Тхі Ха Ми, То Ткхі Kha Мі.

***Анотація:** У даній роботі розглянута задача класифікації зображень, розміщений короткий опис та проведений огляд структури згорткових нейронних мереж.*

***Ключові слова:** нейронна мережа, глибинне навчання, згортка, зображення класифікація.*

***Аннотация:** В данной работе рассмотрена задача классификации изображений, предложено краткое описание и проведен обзор структуры сверточных нейронных сетей.*

***Ключевые слова:** нейронная сеть, глубинное обучения, свертка, изображение классификация.*

***Abstract:** In this work, the problem of classifying images is considered and a brief description is provided and the structure of convolutional neural networks is reviewed.*

***Keywords:** neural network, deep learning, convolution, image classification.*

Сьогодні штучні нейронні мережі широко використовуються для вирішення різних задач в реальному житті. Такі проблеми, як розпізнавання мови або оптичне розпізнавання символів, тепер вирішують з використанням нейронної мережі з високою точністю, бо технології комп'ютерного зору є дуже поширеними [1-2]. У зв'язку з ростом обчислювальних потужностей і появою великих баз зображень стало можливим більш змістовне вивчення і застосування глибоких нейронних мереж [3-4].

У даній роботі розглядається одна спеціальна архітектура нейронної мережі – згорткова нейронна мережа, а також її структура і застосування для класифікації зображень.

Один з найпоширеніших способів класифікації - спосіб на основі описів об'єктів з використанням характеристик, в якому кожен об'єкт характеризується набором числових або нечислових ознак. Однак, для деяких типів даних, відкриті ознаки не дають хорошою точності класифікації, наприклад, колір точок зображень або цифровий звуковий сигнал. Класифікація зображень, наприклад тварин і літаків є простою для людини, але дуже складною для машини. Причина у тому, що ці типи даних містять «приховані ознаки», які людина може легко сприймати, а машина ні, наприклад, у собаки це морда, очі і ніс, а у літака – крила [5].

Для класифікації зображень використовують згорткові нейронні мережі.(англ. *Convolutional Neural Network – CNN*).

Прорив в побудові моделей для класифікації зображень прийшов з відкриттям, того що згорткову нейронну мережу (*CNN*) можна використовувати для поступового вилучення вищого рівня уявлень вмісту зображення. Замість попередньої обробки даних для отримання таких функцій, як текстури і форми, *CNN* приймає в якості вхідних тільки необроблені піксельні дані зображення і «вчиться» витягувати ці функції і, в кінцевому рахунку, робити висновок, з якого об'єкта вони складаються [6].

Для початку, коли система отримує вхідне зображення, їй вже відомий фіксований набір категорій або міток. Це можуть бути будь-які об'єкти: тварини, одяг, предмети. *CNN* отримує карту вхідних об'єктів: тривимірну матрицю, де розмір перших двох вимірів відповідає довжині і ширині зображень в пікселях. Розмір третього виміру дорівнює 3 (відповідає 3 каналам

кольорового зображення: червоному, зеленому і синьому). *CNN* містить стек модулів, кожен з яких виконує три операції.

#### 1. Згортка

Згортка «витягує» листи вхідної карти об'єктів і застосовує до них фільтри для обчислення нових об'єктів, створення вихідної карти або згорнутих об'єктів (які можуть мати інший розмір і глибину, ніж карта вхідних об'єктів). Згортки визначаються двома параметрами:

##### 1.1. Розмір видобутих «плиток» (зазвичай 3x3 або 5x5 пікселів).

Плитка – це композиція, яка складена з відносно невеликих фрагментів квадратної форми.

##### 1.2. Глибина вихідної карти об'єктів, яка відповідає кількості застосованих фільтрів [7].

Під час згортки, фільтри (матриці того ж розміру, що і розмір елемента мозаїчного зображення) ефективно ковзають по сітці вхідної карти об'єктів по горизонталі і вертикалі, по одному пікселю за раз, виділяючи кожен відповідний елемент мозаїчного зображення (див. Рис.1).

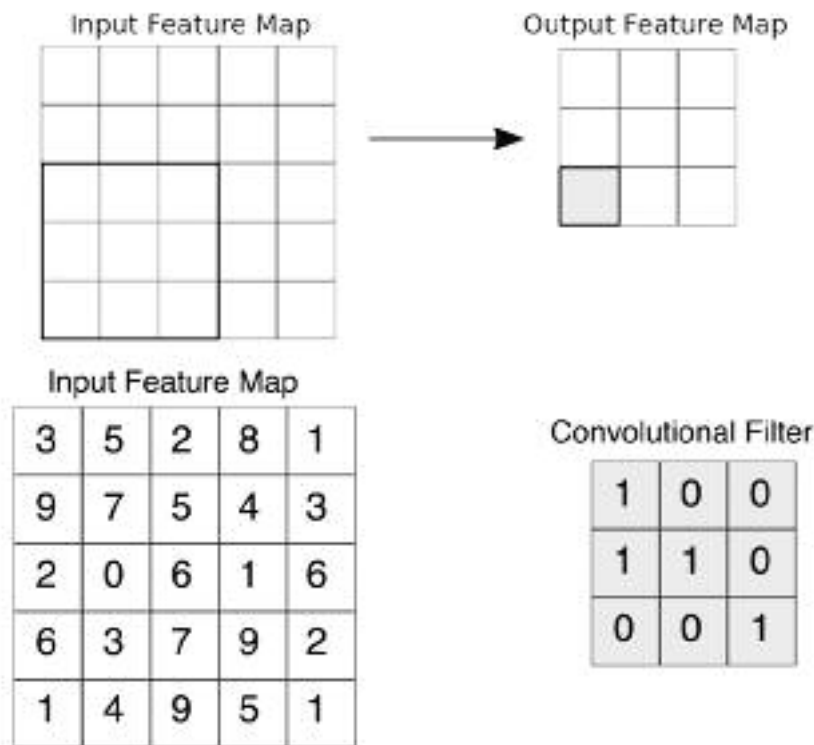


Рис.1 Згортка 3x3, виконана на мапі вхідних об'єктів 5x5.

Для кожної пари елементів мозаїчного зображення *CNN* виконується поелементне множення матриці фільтра і матриці елементів мозаїчного зображення, а потім сумуються всі елементи результуючої матриці, щоб отримати одне значення. Кожне з цих результуючих значень (для кожної пари «фільтр-мозаїка») виводиться у згорнутій матриці ознак (див. Рис. 2).

Під час навчання, *CNN* «запам'ятовує» оптимальні значення для матриць фільтра, які дозволяють йому отримувати значущі елементи (текстури, фігури) з вхідної карти об'єктів. У міру того, як кількість фільтрів (глибина карти вихідних об'єктів), що застосовуються до вхідних даних, збільшується й кількість об'єктів, які може отримати *CNN*. Однак компроміс полягає в тому, що фільтри складають більшу частину ресурсів, які витрачають *CNN*, тому час навчання також збільшується в міру додавання більшої кількості фільтрів. Крім того, кожен фільтр, який додається в мережу, забезпечує менше додаткове значення, ніж попереднє, тому

створюються мережі, що використовують мінімальну кількість фільтрів, необхідних для отримання функцій для точної класифікації зображень.

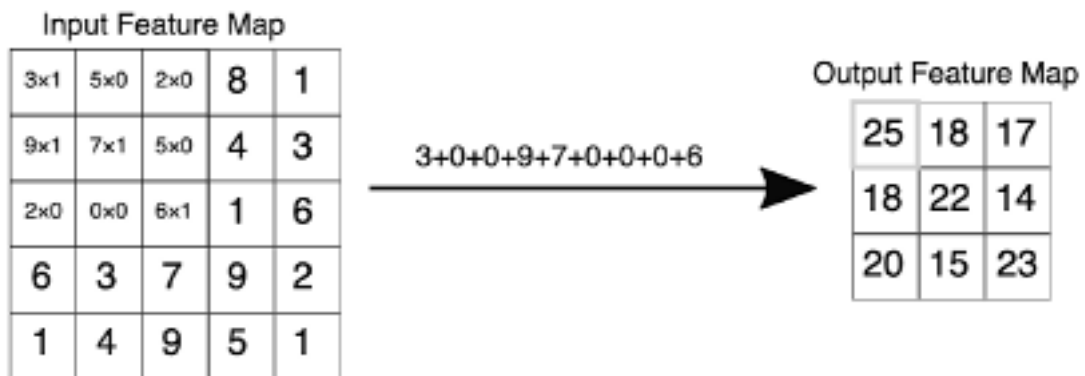


Рис. 2 Розрахунки отриманої згорнутої матриці ознак

### 2. Випрямлена лінійна одиниця *ReLU* (*Rectified Linear Unit*)

Після кожної операції згортки *CNN* застосовує перетворення випрямленої лінійної одиниці (*ReLU*) до згорнутої функції, щоб ввести нелінійність в модель. Функція *ReLU*,  $F(x)=\max(0, x)$  повертає  $x$  для всіх значень  $x > 0$  і повертає 0 для інших значень.

### 3. Субдискретизація

Після *ReLU* настає етап субдискретизації, на якому *CNN* скорочує кількість згорнутих об'єктів (щоб заощадити час обробки), скорочуючи кількість вимірювань карти об'єктів, зберігаючи при цьому найбільш важливу інформацію про функції. Для цього процесу використовують, максимізаційне агрегування (англ. *max pooling*) застосовуючи максимальне значення кожного з кластерів нейронів попереднього шару.

Цей процес виконується аналогічно згортці. Спочатку рухаючись по карті об'єктів витягаються плитки зазначеного розміру. Для кожної плитки максимальне значення виводиться на нову карту об'єктів, а всі інші значення відкидаються. Максимальні операції містять два параметри:

1. Розмір фільтра максимального значення (зазвичай 2x2 пікселя)

2. Відстань в пікселях, що відділяє кожну витягнуту плитку. На відміну від згортки, коли фільтри ковзають по карті об'єктів попіксельно, при максимальному значенні крок визначає місце, де витягується кожна плитка. Для фільтра 2x2 крок 2 вказує, що операція максимального об'єднання буде витягувати всі неперекриті фрагменти 2x2 з карти об'єктів (див. Рис. 3).

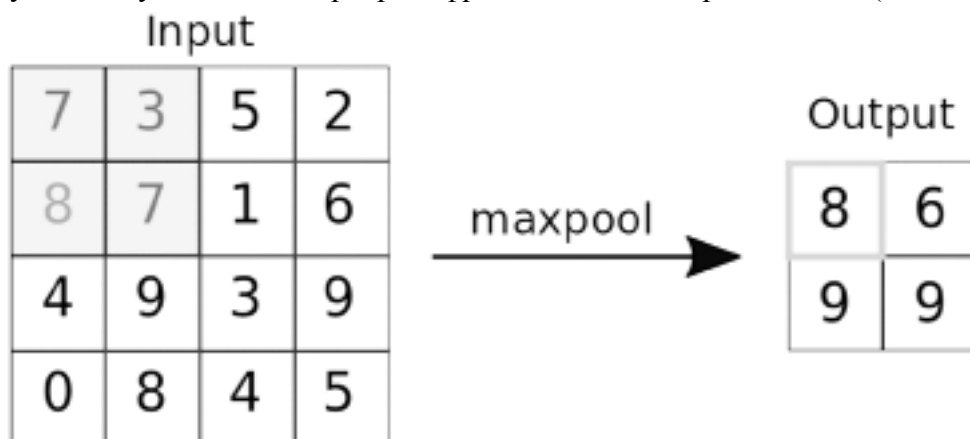


Рис. 3 Максимальне значення для карти об'єктів 4x4 з фільтром 2x2 і кроком 2.

У кінці згорткової нейронної мережі знаходяться один або кілька повністю пов'язаних рівнів (коли два рівня «повністю пов'язані», кожен вузол першого рівня з'єднується з кожним вузлом другого рівня). Їх робота полягає у виконанні класифікації на основі ознак, витягнутих із згорток. Як правило, останній повністю підключений шар містить функцію активації softmax, яка виводить значення ймовірності від 0 до 1 для кожної мітки класифікації, яку модель намагається передбачити.

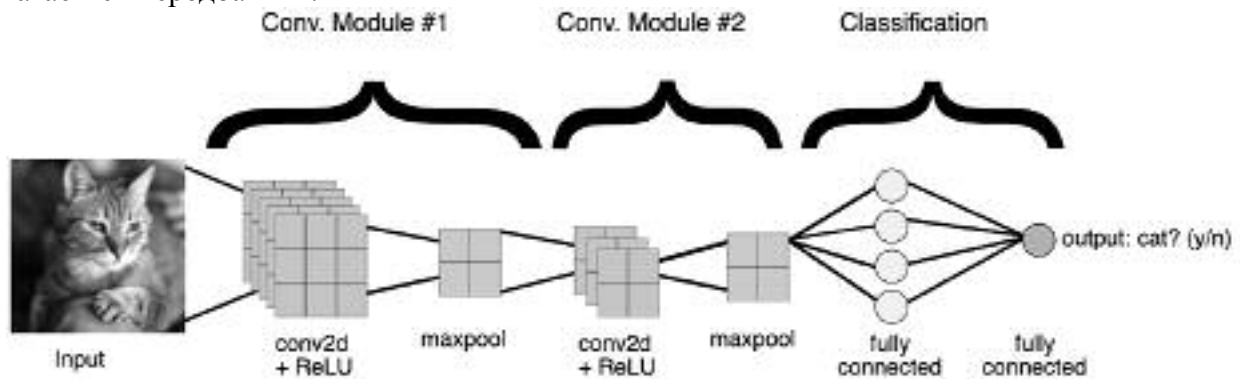


Рис. 4 Наскрізна структура згорткової нейронної мережі

*CNN*, показаний на прикладі (рис.4), містить два модулі згортки (згортка + *ReLU* + max pooling) для вилучення ознак і два повністю пов'язаних рівня для класифікації. Інші *CNN* можуть містити більшу або меншу кількість згортальних модулів або пов'язаних шарів. Доцільно проведення експериментів, щоб з'ясувати конфігурацію, яка дає найкращі результати для моделі [8].

Переваги у використанні згорткової нейронної мережі наступні:

1. Зменшення кількості початкових параметрів та підвищення швидкості навчання порівняно з повно-нейронною мережею.
2. Можливість розпаралелювання обчислень і реалізації алгоритмів навчання мережі на графічних процесорах (*GPU*).
3. Стійкість до зсуву позиції об'єкта у вхідних даних. При навчанні згорткова нейронна мережа зсувається по частинах об'єкта. Тому навчальні ознаки, які залежать від позиції «важливих частин», наприклад голова собаки (див. рис. 3), згорткова нейронна мережа виробляє однакові ознаки для двох зображень, хоча позиції собак в цих зображеннях різні). Таким чином, ця властивість згорткової нейронної мережі допомагає підвищувати якість класифікації [9].

## ВИСНОВКИ

Отже, згорткову нейронну мережу можна застосовувати для вирішення різних типів завдань: розпізнавання, класифікації зображень, мови, текстів тощо. При використанні згорткової нейронної мережі можна істотно зменшувати кількість навчальних параметрів та отримувати високу якість класифікації.

## ЛІТЕРАТУРА

1. Шпинковський, О., Шпинковська, М., Голобородько, В. Інформаційна система для допомоги фінансовим установам у визначенні кредитоспроможності клієнтів // Автоматизація технологічних та бізнес процесів. 11, 3 (Лист. 2019), 14-22.
2. Голобородько В. В., Шпинковська М.І. Рішення задачі бінарної класифікації за допомогою нейронної мережі. XII міжнародна науково-практична конференція «Інформаційні технології і автоматизація – 2019. С. 98-100.

3. Забезпечення ефективності маркетингових компаній за допомогою машинного навчання. Голобородько В.В., Шпинковська М.І. Матеріали XII Всеукраїнської науково-практичної WEB конференції аспірантів, студентів та молодих вчених «Комп'ютерні інтелектуальні системи та мережі» (20-22 березня 2019 р.). – Кривий Ріг: Криворізький національний університет, 2019. – С. 103-104.

4. Bengio Y., Courville A., Vincent P. Representation learning: A review and new perspectives. Pattern Analysis and Machine Intelligence. IEEE Transactions. 2013.

5. Lee H., Grosse R., Ranganath R., Ng A.Y. Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations. Proceedings of the 26th Annual International Conference on Machine Learning. 2009.

6. Bengio Y. Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2009.

7. Machine Learning: Neural Networks [Електронний ресурс] – Режим доступу: <https://developers.google.com/machine-learning/crash-course/introduction-to-neural-networks/video-lecture>

8. ML Practicum: Image Classification. Introducing Convolutional Neural Networks. [Електронний ресурс] – Режим доступу: <https://developers.google.com/machine-learning/practica/image-classification/convolutional-neural-networks>

9. Bengio Y., LeCun Y. Scaling learning algorithms towards AI / eds. L. Bottou, O. Chapelle, D. DeCoste, J. Weston. Large Scale Kernel Machines. MIT Press, 2007.

10. Шпинковський, О. А. ЗАСОБИ РЕКОМЕНДОВАНОГО ПОШУКУ ГРУП КОРИСТУВАЧІВ У СОЦІАЛЬНИХ МЕРЕЖАХ / О. А. Шпинковський, В. М. Ус // Тези доповідей 52-ої наукової конференції молодих дослідників ОНПУ-магістрантів «Сучасні інформаційні технології та телекомунікаційні мережі», Аналіз та синтез інформаційно-аналітичних систем. - Одеса: ОНПУ, 2017, вип. 52, Т. 4. - С. 3

11. Дібров, Д. І. Застосування технології віртуальної машини / Д. І. Дібров ; наук. керівник О. А. Шпинковський // Сучасні інформ. технології та телекомунікаційні мережі : тези доп. 48-ої наук. конф. молодих дослідників ОНПУ-магістрантів. - Одеса, 2013. - С. 3.

12. Шпинковський, О. А. Рекомендаційна система у виборі медіа-контенту / О. А. Шпинковський, А. П. Радіонова // Програмовані логічні інтегр. схеми та мікропроцесорна техніка в освіті і вир-ві : зб. тез Міжнарод. наук.-практ. семінару молодих вчених та студентів, 20-21 квіт. 2018 р. / Луцький нац. техн. ун-т. - Луцьк, 2018. - С. 77.

13. Шпинковський, О. А. РОЗШИРЕНА ІНФОРМАЦІЙНА СИСТЕМА СОЦІАЛЬНОЇ МЕРЕЖІ У ДОПОВНЕННІЙ РЕАЛЬНОСТІ / О. А. Шпинковський, А. А. Балан // Тези доповідей 52-ої наукової конференції молодих дослідників ОНПУ-магістрантів «Сучасні інформаційні технології та телекомунікаційні мережі», Аналіз та синтез інформаційно-аналітичних систем. - Одеса: ОНПУ, 2017, вип. 52, Т. 4. - С. 2

14. Бордан І.С. Сучасні інструменти і технології створення штучного інтелекту / І. С. Бордан, О. А. Шпинковський // Сучасні інформ. технології : матеріали Дев'ятої Міжнарод. наук. конф. студентів та молодих вчених, м. Одеса, 23-24 трав. 2019 р. / Одес. нац. політех. ун-т. – Одеса, 2019. – С. 122–123.

15. Шпинковський, О. А. Інформаційна система визначення уподобань користувачів контенту / О. А. Шпинковський, А. П. Радіонова // Проблеми енергоресурсозбереження в пром. регіоні. Наука і практика : IV Всеукр. наук.-практ. конф. молодих вчених, фахівців, аспірантів : тези доп., м. Маріуполь, 15-17 трав. 2018 р. - Маріуполь, 2018. - С. 129.