

<https://doi.org/10.31891/2219-9365-2024-77-32>

УДК: 004.89:004.4, 631-047.36:004

СУПРИГАН Олена

Вінницький національний аграрний університет

<https://orcid.org/0009-0001-6129-3383>

e-mail: suprihan@gmail.com

СУПРИГАН Віталій

Вінницький національний аграрний університет

<https://orcid.org/0000-0003-2539-8003>

e-mail: suprihan@gmail.com

ДОСЛІДЖЕННЯ ПРОЦЕСУ СТВОРЕННЯ ОЗНАК КЛАСІВ В ЗОБРАЖЕННЯХ ДЛЯ ПІДВИЩЕННЯ ДОСТОВІРНОСТІ РОБОТИ ЗАГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Стаття описує вирішення задачі попередньої обробки зображень для наступного етапу розпізнавання та класифікації з використанням згорткових нейронних мереж. Детально досліджено виділення характеристичних ознак класів. Науковою новизною статті є пропозиція перетворювати зображення у стан із виділенням ознак з високим пріоритетом і видаленням ознак із низьким пріоритетом та подальшою обробкою.

Ключові слова: згорткова нейронна мережа, логіко-часові функції, оператор впливу.

SUPRYHAN Olena, SUPRYHAN Vitalii

Vinnitsia National Agrarian University

RESEARCH OF THE PROCESS OF CREATING CLASS MARKERS IN IMAGES TO INCREASE THE RELIABILITY OF THE WORK OF CONVERTIBLE NEURAL NETWORKS

Convolutional neural networks (CNNs) are a fundamental tool in the field of artificial intelligence for visual data analysis. Their distinguishing feature is the way they process images, which mimics the vision mechanisms of living organisms, using a multi-level approach to extracting features of objects in an image.

The concept of convolution, a special operation that applies filters or kernels to the input image to highlight important features, such as edges, corners, and textures, is at the heart of SNM. This allows the SNM to focus on key details, ignoring minor information, which simplifies further processing.

After convolution, the resulting information passes through a subsampling layer, where the image is reduced in size to reduce computational complexity and to increase invariance to changes in the input, such as scaling or rotation.

This process is repeated in several layers, where each subsequent layer reveals increasingly complex features based on those identified by previous layers. This hierarchical definition of features allows the network to "understand" complex visual patterns.

At the end of the process, after multi-layer convolutions and subsampling, the obtained characteristics are fed into one or more fully connected layers, where all the studied features are combined to solve the final task - whether it is image classification, object presence detection, or recognition.

ANNs are trained using an error backpropagation method, where the weights of each filter are adjusted based on the discrepancy between the predicted and actual results, thus gradually increasing the accuracy of the network.

ANNs have proven effective in a wide range of applications, from automatic license plate recognition to medical image diagnostics, thanks to their ability to separate important information from noise and process complex visual data.

The article describes the solution to the problem of preliminary image processing for the next stage of recognition and classification using convolutional neural networks. The identification of characteristic features of classes has been thoroughly investigated. The scientific novelty of the article is the proposal to transform the image into a state with the highlighting of high-priority features and the removal of low-priority features, followed by further processing.

Keywords: convolutional neural networks, logic-time functions, influence operator.

Постановка проблеми у загальному вигляді

та її зв'язок із важливими науковими чи практичними завданнями

Згорткові нейронні мережі (ЗНМ) є фундаментальним інструментом в сфері штучного інтелекту для аналізу візуальних даних [1, 2]. Їхнім вирізняльним аспектом є спосіб обробки зображень, який імітує механізми бачення живих організмів, використовуючи багаторівневий підхід до виділення характеристик об'єктів на зображенні.

В основі ЗНМ лежить концепція згортки – спеціальної операції, що застосовує фільтри або ядра до вхідного зображення для виділення важливих особливостей, таких як краї, кути, текстури [3]. Це дозволяє ЗНМ зосередитися на ключових деталях, ігноруючи незначну інформацію, що спрощує подальшу обробку.

Після згортки інформація, яка була отримана, проходить через шар субдискретизації, де зображення зменшується в розмірах для скорочення обчислювальної складності та для підвищення інваріантності до змін у вхідних даних, таких як масштабування чи поворот.

Цей процес повторюється в кілька шарів, де кожен наступний шар виявляє все більш складні особливості на основі тих, що були визначені попередніми шарами. Це ієрархічне визначення характеристик дозволяє мережі «розуміти» складні візуальні патерни.

Наприкінці процесу, після багатошарових згорток та субдискретизацій, отримані характеристики подаються в один або декілька повнозв'язних шарів, де всі вивчені особливості об'єднюються для вирішення кінцевого завдання – чи то класифікація зображення, визначення присутності об'єкта, чи розпізнавання.

Навчання ЗНМ відбувається за допомогою методу зворотного поширення помилки, де ваги кожного фільтру коригуються на основі розбіжності між передбаченим та фактичним результатом, таким чином поступово підвищуючи точність мережі.

ЗНМ показали свою ефективність у широкому спектрі застосувань, від автоматичного розпізнавання номерних знаків до діагностики медичних зображень, завдяки своїй здатності відокремлювати важливу інформацію від шуму та обробляти складні візуальні дані.

Місце ЗНМ – попередня обробка зображень для подальшого аналізу та прийняття рішення у багатошаровій нейронній мережі, відштовхує ідею що об'єктом розпізнавання чи класифікації може бути сигнал чи інше інформаційне повідомлення або навіть аналітичний вираз (функція). Для початку дослідження можливо звузити коло досліджуваних об'єктів до звичного зображення але такого, у якому не частини зображення утримують ознаки, а комбінація значень в усіх точках зображення описує деякий визначник. Аналогією такому джерелу інформації можна співставити голограму де кожен елемент тривимірної структури відповідає кожній точці вхідного зображення [4].

Аналіз досліджень та публікацій

Аналіз джерел та останніх досліджень показав, що ЗНМ вважаються одними із найкращих алгоритмів розпізнавання та класифікації зображень [3]. До переваги слід віднести те, що у порівнянні з повнозв'язною нейронною мережею (такою як перцептрон) ЗНМ має значно меншу кількість регульованих ваг, оскільки одне ядро ваг використовується одразу для всього зображення, замість створення персональних вагових коефіцієнтів для кожного пікселя вхідного зображення. Це стимулює мережу до узагальнення інформації під час навчання, а не до піксельного запам'ятовування кожної картинки в міриадах вагових коефіцієнтів, як це робить перцептрон. Зручне паралельне виконання обчислень, що дає можливість реалізації алгоритмів роботи та навчання мережі на графічних процесорах. Відносна стійкість до повороту та зсуву розпізнаваного зображення. Навчання з використанням класичного методу зворотного поширення помилки.

В якості недоліків слід відмітити, що існує занадто багато змінних параметрів мережі; не очевидно, які налаштування потрібні для конкретного завдання і обчислювальної потужності. До змінних параметрів можна віднести: кількість шарів, розмірність ядра згортки для кожного з шарів, кількість ядер для кожного шару, крок зсуву ядра при обробці шару, необхідність шарів субдискретизації, ступінь їх зменшення розмірності, функція по зменшенню розмірності (вибір максимуму, середнього тощо), передавальна функція нейронів, наявність та параметри вихідної повнозв'язної нейромережі на виході ЗНМ. Всі ці параметри істотно впливають на результат, але вибираються дослідниками емпірично. Існує кілька перевірених та добре працюючих конфігурацій мереж, але бракує рекомендацій, за якими необхідно будувати мережу для нової задачі.

Концептуально описано, що можна виділяти класи ознак та кількісно описувати ознаки, зокрема досліджувати зв'язки та вплив між класами, між ознаками класів та між ознаками різних класів [5]. Дане дослідження вносить додаткових механізм створення матриць взаємного впливу та накладання цих матриць на оригінальне зображення. Така активність може бути описана як додавання метаінформації у зображення але точний опис результату процесу додавання це фактично перетворення початкового зображення у опосередкований носій ознак класів із вилученням інформаційних повідомлень які мають незначний вплив. Алгоритм формування попереднього набору даних включає в себе такі кроки як на рисунку 1.

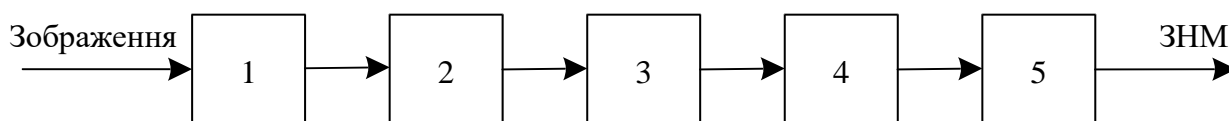


Рис.1. Принцип формування набору даних та подальшої їх обробки [6]: 1 – виділення класів ознак, 2 – утворення ієрархії ознак класів із розгалуженнями у вигляді ознак класів, 3 – виділення характеристичних ознак класів, 4 – формування матриці взаємодії між ознаками та класами об'єктів, 5 – формування набору даних.

Формулювання цілей статті

Головною метою статті є опис способу виділення характеристичних ознак класів, якісний аналіз та узагальнення результату досліджень, що в підсумку приводить до того, що така попередня обробка даних підвищує достовірність розпізнавання та класифікації на наступних етапах обробки у ЗНМ та у нейронній мережі.

Виклад основного матеріалу

Розглянемо детальніше послідовність формування набору даних для подальшої обробки. Як було зазначено, що згадана послідовність є лише частиною загального пайплайну у задачі розпізнавання/класифікації даних або зображень. У блоці 1 (Рис.1) виконується кластеризація класів ознак,

виділяються найбільш значні типи опису зображення. Блок 2 рисунка 1 показує, що для кожного класу ознак слід виділити усі ознаки для наступної обробки. У блоці 3 проводиться кількісний опис та числовий розрахунок кореляції бінарних даних між класами ознак, ознаками класів та між ознаками інших класів. Результатом роботи наступного блоку 4 є матриця взаємного впливу коефіцієнтів [5]. Маючи декілька наборів матриць впливу блок 5 вказує, що необхідно застосувати (apply) матрицю впливу до вхідних даних чи зображення.

Зосередимо дослідження на блоці 3, який відповідає за виділення ознак класів (Рис.2) [6].

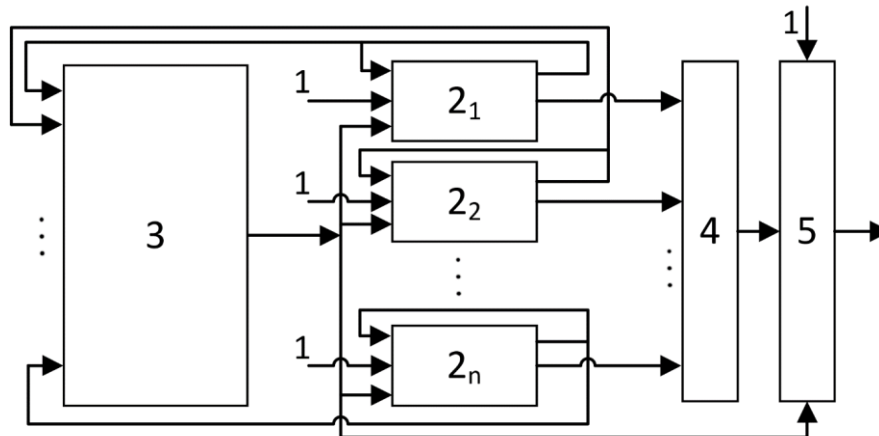


Рис. 2. Виділення ознак класів [6]: 1 – вхідна інформація щодо класу, 2 – блок вилучення найменшої загальної частини від основної інформації, 3 – визначення найменшої загальної частини, 4 – застосування характеристичної функції (apply), 5 – визначення ваги та масштабування ознаки

Опишемо загальний пайплайн процесу визначення характеристик ознак до етапу формування матриці взаємодії (Рис.1, блок 4). Відмітимо що подальший опис стосується обчислень, які відповідають одному класу ознак, наприклад, клас ознак – текстура (ознаки: яскравість, колір, товщина ліній та відносний нахил, крок перетину з іншими елементами тощо). Далі в процесі обробки даних йде етап витягування характеристик для різних категорій об’єктів. Ці ознаки створюються на основі характеристик властивостей окремих фрагментів зображення, які в сукупності надають детальний та змістовний опис об’єкта. Під час цього процесу проводиться детальний розбір кожного сегменту та елемента на зображенні, відсторонено від їхньої якості чи типу. Це досягається завдяки глибокій цифровій обробці інформації. Завдяки цьому, ми маємо змогу розширити кількість врахованих характеристик до їхнього оптимального розмаїття, забезпечуючи тим самим аналіз великою точністю та достовірністю. Для виділення ознак проводиться ітеративне відокремлення найменшої загальної частини. До прикладу як такі обчислення можна було б провести над цілими десятковими числами показано в таблиці 1.

Таблиця 1.

Ітеративне відокремлення найменшої загальної частини

Крок	Цілочислові десяткові дані					Найменша загальна частина	Знайдене значення ознаки
1	3	3	4	5	6	3	3
2	0	0	1	2	3	1	1
3	0	0	0	1	2	1	1
4	0	0	0	0	1	1	1
5	0	0	0	0	0	0	-

Слід відмітити, що на першому кроці ітерації існують тільки вхідні дані, тому відсутній зворотний зв’язок у блоці 2 та на вході у блок 3 (Рис.2). Після першого обчислення найменшої загальної частини, знайдене значення буде інтерпретуватися як ознака, та буде передано на вхід блоків 2, 3 та 4. Ітеративний процес продовжується до межі можливості обчислення найменшої загальної частини. Додатковим наслідком ітеративності є нормування за кількістю ознак у кожному класі ознак, щоб їх було порівну в усіх класах ознак.

Блок 4 (Рис.2) відповідає за формування кореляційної функції бінарних даних для ознак у класі ознак та для ознак між різними класами ознак. Особливості роботи цього блоку потребують додаткових досліджень, оскільки на сучасному етапі створено залежності тільки для логіко-часових функцій [7] але не для бінарних даних.

Блок 5 (Рис.2) відповідає операції застосування аргументу (матриці взаємодії) до вхідного зображення. Особливості функціонування цього блоку також потребують додаткових досліджень.

Такий підхід дозволяє виділяти класи ознак та нормувати відповідні ознаки. Очікуваним результатом роботи виділення характеристичних ознак класів є вихідне зображення або дані із накладеними додатковими метаданими та видаленими ознаками, які мають незначний вплив.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

Запропоновано та описано спосіб попереднього оброблення зображень, зокрема проведено аналіз роботи етапу виділення ознак класів. Виконаний якісний аналіз впливу обчислень на підсумковий результат та показано, що в результаті попередньої обробки зображень буде виділено дані із важливими ознаками та видалено дані із ознаками якими можливо знехтувати. Слід відмітити, що в процесі створення статті знайдені блоки у послідовності обчислень, які потребують додаткових досліджень.

Література

1. LeCun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), 1989. 541–551. DOI: <https://doi.org/10.1162/neco.1989.1.4.541>
2. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient based Learning Applied to Document Recognition [Електронний ресурс] / Режим доступу: <http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf>. – Назва з екрану.
3. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition [Електронний ресурс] / www.arXiv.org. Режим доступу: <https://arxiv.org/pdf/1409.1556.pdf>. – Назва з екрану.
4. Robert Jacob Collier, Christoph B. Burckhardt, Lawrence H. Lin Optical Holography. First Edition, Academic Press, 1971, 605 p.
5. Пат. 67493 А UA, МПК G06G 7/14, G06K 9/00. Спосіб формування оператора впливу [Текст] / В. П. Кожем'яко, О. І. Суприган, Н. В. Сачанюк (Україна). – № 2003109049 ; заявл. 06.10.2003 ; опубл. 15.06.2004, Бюл. № 6. – 2 с. : кресл.
6. Суприган О., Суприган В. Вдосконалення роботи згорткових нейронних мереж для розпізнавання зображень стану ґрунтів у сільському господарстві. *Вимірювальна та обчислювальна техніка в технологічних процесах*, 2023. №4. С.207-215. DOI: <https://doi.org/10.31891/2219-9365-2023-76-28>
7. Сачанюк-Кавецька Н. В., Прозор О. П. Елементи математичного опису логікочасових функцій багатозначної логіки та окремих операцій над ними. *Інформаційні технології та комп'ютерна інженерія*, 2022. №1. С.111-118. DOI: <https://doi.org/10.31649/1999-9941-2022-53-1-111-118>

References

1. LeCun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), 1989. 541–551. <https://doi.org/10.1162/neco.1989.1.4.541>
2. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient based Learning Applied to Document Recognition [Elektronnyi resurs] / Rezhim dostupu: <http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf>. – Nazva z ekrany.
3. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition [Elektronnyi resurs] / www.arXiv.org. Rezhim dostupu: <https://arxiv.org/pdf/1409.1556.pdf>. – Nazva z ekrany.
4. Robert Jacob Collier, Christoph B. Burckhardt, Lawrence H. Lin Optical Holography. First Edition, Academic Press, 1971, 605 p.
5. Pat. 67493 A UA, MPK G06G 7/14, G06K 9/00. Sposib formuvannia operatora vplyvu [Tekst] / V. P. Kozhem'yako, O. I. Supryhan, N. V. Sachaniuk (Ukraine). - № 2003109049 ; zaiavl. 06.10.2003 ; opubl. 15.06.2004, Biul. № 6. - 2 s. : kresl.
6. Supryhan O, Supryhan V. Improving the work of convolutionalneural networks for recognition of soil condition images in agriculture. *Measuring and computing devices in technological processes*, 2023. Issue 4. P.207-215. DOI: <https://doi.org/10.31891/2219-9365-2023-76-28>
7. Sachaniuk-Kavets'ka N., Prozor O. Mathematical description of logic-time functions of multiple-valued logic and some operations over them. *Information technology and computer engineering*, 2022. Vol.53. No.1. P.111-118. DOI: <https://doi.org/10.31649/1999-9941-2022-53-1-111-118>