

# Глибокі Нейронні Мережі для Вирішення Завдань Розпізнавання і Класифікації Зображення

Віктор Синєглазов

Кафедра авіаційних комп'ютерно-інтегрованих комплексів,  
Навчально-науковий інститут інформаційно-діагностичних систем,  
Національний авіаційний університет, Київ, Україна  
svm@nau.edu.ua

Олена Чумаченко

Кафедра технічної кібернетики,  
Національний технічний університет України  
«Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна  
chumachenko@tk.kpi.ua

## Deep Neural Networks for Solution Problems of Recognition and Classification of Images

Victor Sineglazov

Computer-Integrated Complexes Department,  
Educational and Scientific Institute of Information and Diagnostic Systems, National Aviation University  
Kyiv, Ukraine  
svm@nau.edu.ua

Olena Chumachenko

Technical Cybernetic Department  
National Technical University of Ukraine  
“Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine  
chumachenko@tk.kpi.ua

**Анотація**—Розглянуто вирішення завдання обробки зображень використовуючи згорткові нейронні мережі. Показано, що незважаючи на свою ефективність вони мають недоліки, які полягають у відсутності оптимального вибору кількості карт ознак, розмірів вікна, площі перекриття, чергування шарів різних типів. У результаті досліджень поставлено і розв'язано задачу пошуку оптимальних значень вищевказаних параметрів на основі використання гібридного генетичного алгоритму. Синергетичний ефект запропонованого алгоритму реалізується завдяки використанню в ньому елементів алгоритмів найбільш поширених модифікацій генетичних алгоритмів, що реалізують різні схеми призначення придатності і селекції. Розглянуто приклад побудови оптимальної структури загорткової мережі.

**Abstract**— It is considered the problem of image processing using convolutional neural network. It is shown that despite their effectiveness they have disadvantages that are the lack of optimal choice of number of features maps, window size, the area of overlap, alternating layers of different types. Based on the research results it is determined and solved the problem of the above parameters optimal values search, based on hybrid genetic algorithm. It is considered the design of convolutional neural network.

**Ключові слова**—обробка зображень; генетичний алгоритм; синергетичний ефект; згорткові нейронні мережі.

**Keywords**—image processing; genetic algorithm; synergistic effect; convolutional neural network.

### I. ВСТУП ТА ПОСТАНОВКА ЗАВДАННЯ

Постановка задачі класифікації. Нехай  $X$  – множина описань об'єктів,  $Y$  – множина номерів (чи найменувань) класів. Існує невідома цільова залежність – відображення  $y^*$ :  $X \rightarrow Y$ , значення якої відомі тільки на об'єктах кінцевої навчальної вибірки:

$$X^m = \{(x_1, y_1), \dots, (x_m, y_m)\},$$

де  $X^m$  множина елементів навчальної вибірки розмірністю  $m$ .

Потрібно побудувати алгоритм здатний визначити належність довільного об'єкта  $x \in X$  до класу  $y \in Y$ .

### II. ВИКОРИСТАННЯ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ ОБРАЗІВ НЕФОРМАЛІЗОВАНИХ ЕЛЕМЕНТІВ

Застосування багатошарового персептрона з традиційною структурою при вирішенні реальних задач розпізнавання і класифікації зображень викликає певні труднощі. По-перше, зображення, як правило, мають велику розмірність, внаслідок чого зростає число нейронів

і синаптичних зв'язків у мережі. У свою чергу, це вимагає збільшення навчальної вибірки, внаслідок чого збільшується час і обчислювальна складність процесу навчання. По-друге, ігнорується топологія вхідних даних. Компоненти вхідного шару можуть бути представлені в будь-якому порядку, без урахування мети навчання. Однак зображення мають строгу двовимірну структуру, в якій існує залежність між просторово-сусідніми пікселями.

Від цих недоліків вільні так звані згорткові нейронні мережі, які являють собою особливий клас багатозарових перцептронів, спеціально створених для розпізнавання двовимірних поверхонь з високим ступенем інваріантності до масштабування, зсуву, повороту, зміни ракурсу та інших просторових спотворень, і глибокі НМ побудовані на базі автоенкодерів, попереднє навчання яких здійснюється на базі обмежених машин Больцмана.

### III. ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ ПІДВИЩЕНОЇ ЕФЕКТИВНОСТІ

Згорткова нейронна мережа (ЗНМ) – тип багатозарової нейронної мережі, яка свою назву «згорткова мережа» отримала за назвою операції – згортка, вона часто використовується для обробки зображень і може бути описана наступною формулою

$$(\mathbf{f} \times \mathbf{g})[m, n] = \sum_{k, l} \mathbf{f}[m - k, n - l] \cdot \mathbf{g}[k, l],$$

де  $\mathbf{f}$  – вихідна матриця зображення;  $\mathbf{g}$  – ядро (матриця) згортки.

Неформально цю операцію можна описати наступним чином – вікном розміру ядра  $\mathbf{g}$  проходимо з заданим кроком (зазвичай 1) все зображення  $\mathbf{f}$  на кожному кроці поелементно множимо вміст вікна на ядро  $\mathbf{g}$ , результат сумується і записується в таблицю результату.

Ідея згорткових нейронних мереж полягає в чергуванні згорткових шарів (англ. convolution layers) і субдискретизуючих шарів (англ. subsampling layers, верств підвибірки). Структура мережі – односпрямована (без зворотних зв'язків), багатозарова (рис. 1).

Модель згорткової мережі складається з трьох типів шарів: згорткові (convolutional) шари, субдискретизуючі (subsampling, підвибірка) верстви і прошарки «звичайної» нейронної мережі – перцептрона.

Архітектура згорткових нейронних мереж реалізує три ідеї, які забезпечують інваріантність мережі до невеликих зрушень, змін масштабу і спотворень:

- кожен нейрон отримує вхідний сигнал від локального рецептивного поля (local receptive fields) у попередньому шарі, що забезпечує локальну двовимірну зв'язність нейронів;
- кожен прихований шар мережі складається з безлічі карт ознак, на яких всі нейрони мають загальні ваги (shared weights), що забезпечує інваріантність до зміщення і скорочення загальної кількості вагових коефіцієнтів мережі;

- за кожним шаром згортки слідує обчислювальний шар, який здійснює локальне усереднення та підвибірку, що забезпечує зменшення розширення для карт ознак.

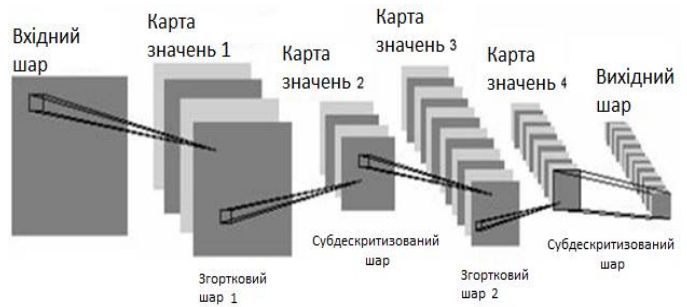


Рис. 1. Структура згорткової мережі.

Робота згорткової нейронної мережі забезпечується двома основними елементами.

- 1) Фільтри (filters) (визначники ознак).
- 2) Карты ознак (feature maps).

Фільтр – це невелика матриця, що представляє ознаку, яку необхідно знайти на вихідному зображенні. За допомогою верхнього фільтра визначаються частини вихідного зображення з вертикальними лініями, нижній фільтр служить для визначення частин зображення з горизонтальними лініями.

Безпосередньо процес визначення заснований на операції згортки фільтром оригінального зображення. Результати згортки, які визначають місце розташування ознак вихідного зображення, називаються картами ознак.

Мета процесу згортки – зменшити розмірність карти ознак до такої міри, щоб з повним набором ознак могла працювати мережа прямого поширення (в більшості випадків багатозаровий перцептрон).

Згортковий шар реалізує ідею локальних рецептивних полів, тобто кожен вихідний нейрон з'єднаний тільки з певною (невеликою) областю вхідної матриці і таким чином моделює деякі особливості людського зору.

Недоліками згорткових нейронних мереж (ЗНМ) є:

- висока складність архітектури;
- повнозв'язаність;
- фіксована площа вікна шару згортки.

З метою підвищення ефективності роботи ЗНМ необхідно знайти оптимальні значення наступних параметрів:

- кількість карт ознак;
- щільність зв'язків між картами ознак;
- розмір вікна;
- площа перекриття;
- початкова ініціалізація ваг.

#### IV. ГЕНЕТИЧНІ АЛГОРИТМИ

Генетичні алгоритми належать до класу еволюційних алгоритмів і мають ряд характеристик, які роблять їх кращими, ніж класичні методи оптимізації:

- для пошуку ефективних рішень з використанням генетичних алгоритмів не потрібні специфічні знання про саму задачу і параметри, що входять до неї;
- в генетичних алгоритмах замість детермінованих використовуються стохастичні оператори, які показали себе досить стійкими в умовах зашумленості зовнішнього середовища;
- властивий генетичним алгоритмам паралелізм – одночасне врахування великої кількості індивідів популяції – робить їх менш чутливими до локальних оптимумів і впливу шумів.

#### V. ЗАГАЛЬНА КОНСТРУКЦІЯ ГЕНЕТИЧНОГО АЛГОРИТМУ

Використання генетичних алгоритмів для розв'язання задач багатокритеріальної оптимізації, дозволяє позбутися від основних недоліків класичних методів, так як генетичні алгоритми підходять для задач великої розмірності і здатні захопити Парето-оптимальні точки навіть при одноразовому запуску алгоритму. За допомогою підтримки популяції рішень і застосування концепції Парето-оптимальності, генетичні алгоритми можуть знаходити різні Парето-оптимальні рішення паралельно.

Таким чином, на відміну від більшості класичних підходів до розв'язання задач багатокритеріальної оптимізації, коли для отримання кожної окремої точки необхідно проводити окремий запуск алгоритму пошуку Парето-оптимальних рішень, застосовуючи еволюційний підхід до векторної оптимізації, завдяки закладеному в генетичних алгоритмах паралелізму, можливо отримання різних точок множини Парето при одному прогоні алгоритму. Дана обставина є очевидною перевагою еволюційного підходу до вирішення задач багатокритеріальної оптимізації перед традиційними методами їх вирішення [1] – [4].

За аналогією з природною еволюцією, в генетичних алгоритмах кандидати-рішення називаються *індивідами* або *особинами*, а безліч кандидатів-рішень – *популяцією*. Кожен індивід визначає можливе рішення задачі, при цьому, однак, сам по собі він не є вектором рішень, а скоріше кодує його, ґрунтуючись на відповідній структурі кодування рішення. У генетичних алгоритмах ця структура визначається вектором – вектором бітів або вектором дійсних чисел – набором генів, що утворюють *хромосоми*. Безліч всіх можливих векторів утворює простір індивідів або популяцію.

Якість індивіда під час розв'язання задачі оптимізації визначається скалярним значенням, так званою *функцією пристосованості (fitness function)*.

В процесі *селекції*, який може бути, як стохастичним, так і детермінованим, гірші рішення – непристосовані індивіди – видаляються з популяції, в той час як індивіди

з більшою пристосованістю – найбільш пристосовані – піддаються репродукції. Мета полягає в тому, щоб посилити пошук в певних областях пошукового простору і збільшити середню «якість» всередині популяції.

Проблемою навчання глибоких нейронних мереж обробки зображень, що використовуються в системі медичної діагностики, є складність отримання навчальної вибірки. Навчальна вибірка виходить на підставі наявної невеликої кількості прикладів шляхом повороту їх на невеликі кути в різних площинах з подальшою обробкою зображень іншим способом.

Задача розпізнавання зображень в такому формулюванні представляє собою досить складну комплексну задачу з низкою обмежень. В роботі пропонується системний підхід до обробки зображень, заснований на структурній декомпозиції (рис. 2).

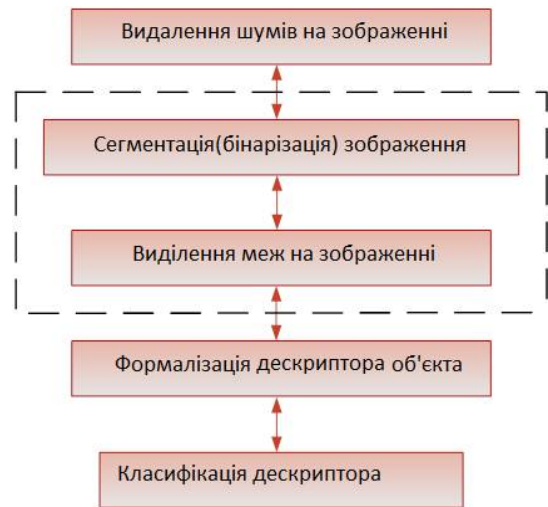


Рис. 2. Алгоритм обробки зображень створення навчальної вибірки для згорткової нейронної мережі.

#### VI. ПРИКЛАД РОЗРАХУНКУ ОПТИМАЛЬНОЇ СТРУКТУРИ ЗГОРТКОВОЇ МЕРЕЖІ

Спробуємо підібрати оптимальну структуру згорткової мережі для розв'язання аналогічної задачі (прогнозування позиції у просторі маніпулятора робота). Розрахунок відрізнятиметься від попереднього прикладу лише способом кодування (і тлумачення) структури згорткової мережі.

Загальну структуру гена для кодування одного шару згорткової мережі показано на рис. 3.

Тип шару (CONV, POOL, FC, ABS)	Параметр 1	Параметр 2	Параметр 3

Рис. 3. Загальна структура гена для кодування одного шару згорткової мережі.

Тоді хромосома, що кодує топологію згорткової мережі буде складатися з  $N$  подібних генів, де  $N$  – задана максимальна кількість шарів. Для демонстрації розглянемо випадок  $N = 4$ .

Оскільки кожен тип шару має свої специфічні (часто несумісні) параметри, то тлумачення кожного параметрів гену залежатиме від типу шару. Призначення кожного з параметрів залежно від типу шару наведено в табл. I.

ТАБЛИЦЯ I. ПРИЗНАЧЕННЯ КОЖНОГО З ПАРАМЕТРІВ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ ЗАЛЕЖНО ВІД ТИПУ ШАРУ

Тип шару (CONV, POOL, FC, ABS)	Параметр 1	Параметр 2	Параметр 3
CONV (00)	Кількість фільтрів (K)	Розмір фільтра (F)	Розмір перекриття (зсув) (S)
POOL (01)	Ігнорується	Розмір фільтра (F)	Розмір перекриття (зсув) (S)
FC (10)	Кількість нейронів у шарі		
ABS (11)	Ігнорується		

Розглянемо приклад кодування детальніше.

**Крок 0. Ініціалізація:** Згенеруємо випадковим чином набір хромосом у вигляді бітових ланцюжків:

- Індивід 1 — 00 11 10 10 00 10 11 01 01 00 10 10 10 101101;
- Індивід 2 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101;
- Індивід 3 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101;
- Індивід 4 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101;
- Індивід 5 — 00 11 10 10 00 10 01 11 01 00 10 11 11 101101.

Кожна згенерована хромосома описує певну структуру згорткової нейронної мережі. Спробуємо «декодувати» хромосому Індивід 1 та побудувати згорткову мережу, що їй відповідає.

1) Розіб'ємо обрану хромосому на  $N$  ( $N = 4$  для даного прикладу) рівних частин:

Індивід 1 — 00 11 10 10 | 00 10 1 01 | 01 00 10 10 | 10 101101.

2) Кожна з отриманих частин кодує один шар згорткової мережі, а саме: тип шару та значення його параметрів (табл. II–V).

ТАБЛИЦЯ II. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ CONV

	Тип шару	Параметр 1	Параметр 2	Параметр 3
Шар 1	00 <sub>2</sub> =0 <sub>10</sub>	11 <sub>2</sub> =3 <sub>10</sub>	10 <sub>2</sub> =2 <sub>10</sub>	10 <sub>2</sub> =2 <sub>10</sub>
Тлумачення	CONV	Кількість фільтрів (K) = 3	Розмір фільтра (F) = 2	Розмір перекриття (S) = 2

ТАБЛИЦЯ III. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ POOL

Шар 2	00 <sub>2</sub> =0 <sub>10</sub>	10 <sub>2</sub> =2 <sub>10</sub>	11 <sub>2</sub> =3 <sub>10</sub>	01 <sub>2</sub> =1 <sub>10</sub>
Тлумачення	CONV	Кількість фільтрів (K) = 2	Розмір фільтра (F) = 3	Розмір перекриття (S) = 1

ТАБЛИЦЯ IV. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ ABS

Шар 3	01 <sub>2</sub> =1 <sub>10</sub>	00 <sub>2</sub> =0 <sub>10</sub>	10 <sub>2</sub> =2 <sub>10</sub>	10 <sub>2</sub> =2 <sub>10</sub>
Тлумачення	POOL	Ігнорується	Розмір фільтра (F) = 2	Розмір перекриття (S) = 2

ТАБЛИЦЯ V. КОДУВАННЯ ПАРАМЕТРІВ ЗГОРТКОВОГО ШАРУ CONV

Шар 4	10 <sub>2</sub> =2 <sub>10</sub>	101101 <sub>2</sub> =45
Тлумачення	FC	Кількість нейронів у шарі = 45

Таким чином, даний індивід кодує згорткову мережу наступного вигляду:

- вхідний шар – кількість нейронів визначається параметрами задачі;
- згортковий шар CONV – 3 фільтри 2×2 з розміром перекриття 2;
- згортковий шар CONV – 2 фільтри 3×3 з розміром перекриття 1;
- субдискретизуючий шар POOL – розмір фільтра 2×2 з розміром перекриття 2;
- повнозв'язний шар FC – 45 нейронів у шарі.

**Крок 1-8.** Аналогічні прикладу розрахунку оптимальної структури мережі глибокої довіри.

## VII. ВИСНОВКИ

Проаналізовано недоліки використання згорткової нейронної мережі при обробці зображень. Обґрунтовано необхідність розв'язання задачі оптимізації для визначення оптимальних параметрів налаштування нейронної мережі. Запропоновано гібридний генетичний алгоритм для її розв'язання. Наведено приклад, який пояснює отримання оптимальної структури згорткової мережі.

## ЛІТЕРАТУРА REFERENCES

- [1] J. Horn, N. Nafpliotis, and D. E. Goldberg, "A niched Pareto genetic algorithm for multiobjective optimization," *In Proceedings of the First IEEE Conference on Evolutionary Computation*, Piscataway, vol. 1, pp. 82-87, 1994.
- [2] S. Watanabe, T. Hiroyasu, and M. Miki, "NCGA: Neighborhood Cultivation Genetic Algorithm for Multi-Objective Optimization Problems," *Proceedings of the Genetic and Evolutionary Computation Conference*, 2002, pp. 458-465.
- [3] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evolutionary Computation*, vol. 8, pp. 173-195, 2000.
- [4] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257-271, 1999.