

УДК 0048:681.3

DOI: 10.26661/2413-6549-2019-1-14

ОПТИМІЗАЦІЯ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ТА ЇХ АНСАМБЛІВ

О. В. Чопорова, А. Г. Кривохата

Запорізький національний університет
krivohata@gmail.com

Ключові слова:

класифікація, згорткові нейронні мережі, ансамбль, генетичні алгоритми.

Штучні нейронні мережі застосовуються до розв'язання великої кількості практичних задач, серед яких можна виділити проблеми класифікації, регресії та кластерного аналізу. Зокрема, для розпізнавання образів ефективно використовуються згорткові нейронні мережі. Така популярність пояснюється структурою згорткових мереж, яка складається зазвичай з двох типів шарів: власне згорткових та повнозв'язних. У шарах згортки відбувається автоматичне визначення суттєвих ознак об'єктів. Цей етап можна інтерпретувати як застосування до даних різноманітних фільтрів. У повнозв'язних шарах відбувається безпосередньо класифікація на базі ознак, вилучених згортковими шарами. Кількість шарів обох типів, тип функцій активації, тип оптимізатора, кількість епох навчання та інші показники відносяться до гіперпараметрів нейромереж. У практичних задачах саме структурна та параметрична оптимізація вимагає суттєвих часових затрат та експертних знань предметної області. Одним з методів, що дозволяє уникнути такого ручного налаштування роботи системи є генетичні алгоритми. Загальна схема застосування генетичних алгоритмів до нейронних мереж полягає у наступному. На першому етапі варто обрати спосіб кодування суттєвих параметрів нейронної мережі у вигляді бінарного або числового вектора. Такі вектори формують деяку базову множину розв'язків, у якій здійснюється пошук оптимального розв'язку. Після цього обирається цільова функція, яка визначатиме найбільш вдалу архітектуру окремих нейронних мереж та їх структуру. Така функція повинна бути близькою за значенням до функції втрат, що використовуються при навчанні нейромереж. Суттєвим етапом генетичних алгоритмів є визначення наступних операторів: селекції, кросовера, мутації та відбору. Вибір цих операторів впливає на збіжність та ефективність методу в цілому. Генетичний алгоритм є прикладом метаевристичних методів. Збіжність таких методів у загальному випадку досить важко довести формально. Однак застосування генетичного алгоритму під час налаштування нейронних мереж дозволяє знизити втручання користувача до мінімуму. Іншим відомим підходом до поліпшення ефективності систем розпізнавання є використання ансамблів – поєднань декількох моделей з генерацією спільного результату. Існує декілька способів побудови такого поєднання. Загальною вимогою є те, що елементи ансамблю повинні розрізнятися або структурно, або параметрично. Також можуть використовуватися подібні моделі, але оптимізовані для різних наборів даних. Використання ансамблевого навчання нейронних мереж пов'язано з проблемою різкого збільшення вимог до обчислювальних ресурсів. Одним з підходів тут може бути застосування так званого Snapshot ансамблю, який дозволяє будувати колекцію нейронних мереж, подібних за структурою, але з різними наборами вагових коефіцієнтів. Час навчання такого ансамблю дорівнює часу навчання однієї мережі. Роботу присвячено застосуванню генетичних алгоритмів для спільного налаштування структури нейронних мереж та параметрів ансамблю.

OPTIMIZATION OF CONVOLUTIONAL NEURAL NETWORKS AND ENSEMBLES

O. V. Choporova, A. G. Kryvokhata

Zaporizhzhia National University
krivohata@gmail.com

Key words:

classification, convolutional neural network, ensemble, genetic algorithm.

Artificial neural networks are used to solve a number of practical problems, among them which are the problems of classification, regression and cluster analysis. In particular, convolutional neural networks are effectively used for pattern recognition. Such popularity is explained by the structure of convolution networks, that usually consists of two types of layers: convolution proper and fully connected. Convolution layers automatically detect the essential features of the objects. This step can be interpreted as applying to a variety of filters. The classification of features obtained by convolutional layers occurs in the fully connected layers. The neural network hyperparameters include number of both types' layers, type of activation functions, type of optimizer, number of epochs etc. Both structural and parametric optimizations are time consuming and require domain expert knowledge. Genetic algorithm is one of the methods for avoiding such manual tuning. The general scheme of application of genetic algorithms to neural networks consists of several steps. The first step is to choose the method of encoding the essential parameters of the neural network in the form of a binary or numeric vector. Such vectors form some basic set of solutions in which the optimal solution is searched. The essential stage of genetic method is the definition of following operator's selection, crossover, mutation and selection. These operators affect the convergence and effectiveness of the method as a whole. The genetic algorithm is an example of metaheuristic methods. There is a difficulty in mathematically proving the convergence of these methods. However, applying a genetic algorithm to neural network tuning minimizes user intervention. Another well-known approach to improve the efficiency of recognition systems is ensembles i.e. combinations of several models to generate a common result. There are several ways to build such combination. The general requirement is that the elements of the ensemble must differ either structurally or parametrically. Similar models may also be used but optimized for different datasets. The ensemble training for neural networks is associated with the problem of dramatically increasing demands on computing resources. One approach here might be to use the so-called Snapshot ensemble, that allows us to build a collection of neural networks similar in structure but with different weights matrices. The training time of such ensemble is equal to the learning time of one network. The work is devoted to the application of genetic algorithms for joint tuning of neural network structure and ensemble parameters.

Вступ. Штучні нейронні мережі отримали широке поширення в різноманітних практичних застосуваннях. Сьогодні існує велика кількість різноманітних архітектур нейронних мереж. Зокрема, для задачі розпізнавання образів використовуються такі нейронні мережі прямого поширення, як багатошаровий перцептрон та згорткові нейронні мережі. При цьому останніми роками в задачах розпізнавання перевага надається згортковим нейронним мережам. Це пояснюється здатністю такого типу мереж до автоматичного виділення ознак сигналу, що,

свою чергою, забезпечується наявністю спеціальних типів шарів: згортки та субдискретизації [1]. Завдяки своїй структурі глибокі згорткові нейронні мережі, при вдалому налаштуванні, можуть знаходити закономірності у вихідних даних та використовувати їх для розв'язання задач класифікації. Сучасні архітектури згорткових нейронних мереж можуть використовувати десятки шарів. Так, архітектура VGG (англ. Visual Geometry Group) використовує 19 шарів «згортка-нелінійність-субдискретизація», а

GoogLeNet та DenseNet (англ. Dense Convolutional Network) відповідно 22 та 250 шарів [2]. Кількість шарів сучасних ResNet (англ. Residual Convolutional Network) вже близько тисячі.

Одним із актуальних питань застосувань нейронних мереж є їхня структурна оптимізація, тобто вибір оптимальної кількості шарів, нейронів, функцій активації тощо. Така оптимізація може проводитися як вручну, за умови відносно невеликої кількості параметрів, так і в автоматичному режимі. Генетичні алгоритми досить давно використовуються для задач структурної оптимізації нейронних мереж [3], однак вони менше застосовувались для підбору параметрів ансамблів мереж.

Метою цієї роботи є розробка варіанта генетичного алгоритму структурної оптимізації згорткових нейронних мереж та їх ансамблів в задачах класифікації аудіоданих.

1. Підходи до побудови ансамблю нейронних мереж. Одним із підходів до підвищення якості прогнозування нейронних мереж є побудова ансамблів декількох моделей. Зазвичай при створенні таких систем виникають дві основні задачі: генерація окремих нейронних мереж ансамблю та обчислення спільного результату прогнозування [1].

Окремі нейронні мережі, які входять до ансамблю, повинні розрізнятися між собою за структурою та параметрами. Цього досягають або використанням різних даних на етапі навчання, або диференціюванням параметрів мережі, або залученням різних алгоритмів параметричної оптимізації.

Ансамблевому навчанню нейронних мереж присвячено, зокрема, роботи [4-9]. У статтях [4, 5] детально розглянуто основні етапи побудови ансамблів нейромереж. Пропонується підхід до генерації моделей ансамблю, з яких методом k -середніх проводиться відбір для подальшого навчання.

У роботі [6] пропонується поняття спільного навчання як альтернатива ансамблевому навчанню. Вводяться функції втрат для оптимізації нейронних мереж саме при спільному навчанні.

Роботи [7, 8] присвячено особливостям використання генетичних алгоритмів при

створенні ансамблів. Автори пропонують підхід, у якому нейронні мережі та структура ансамблю поєднуються в одну популяцію. По-перше, створюється набір нейронних мереж з високим ступенем різноманітності. Для цього використовуються різні набори навчальних даних для кожної моделі, а також варіюється архітектура шляхом зміни кількості прихованих нейронів, функцій активації та ініціалізації ваг. По-друге, генетичний алгоритм використовується для вибору як найбільш ефективної підмножини синтезованих нейронних мереж, так і оптимальної комбінаційної стратегії для забезпечення точності та надійності ансамблю. Така схема застосування генетичних алгоритмів до нейронних мереж та їх ансамблів є типовою.

Одним із недоліків використання ансамблів нейронних мереж є висока вимогливість до часових та просторових ресурсів, оскільки час навчання збільшується пропорційно до кількості нейронних мереж в ансамблі. Особливо цей недолік стосується глибинних нейронних мереж.

У роботі [9] пропонується замість навчання M нейронних мереж навчати однієї мережі. Основна ідея полягає в тому, що при застосуванні методу стохастичного градієнтного спуску запам'ятовуються значення вагової матриці при потраплянні у M точок локального мінімуму. Після цього для кожної з M вагових матриць генерується відповідна нейронна мережа. Таким чином, час навчання ансамблю (авторська назва – SparseShot ансамбль) майже не відрізняється від часу навчання однієї нейронної мережі.

2. Застосування генетичних алгоритмів для оптимізації згорткових нейронних мереж та їх ансамблів. Теорію еволюційних, зокрема генетичних, алгоритмів на сьогодні добре розроблено, наприклад у роботах [3, 10] формалізовано основні поняття та приклади застосувань до задач дискретної оптимізації. Оскільки генетичні алгоритми є метаевристичними, досягнення глобального оптимуму не гарантується, однак ці підходи добре зарекомендували себе у низці задач дискретної математики.

Загальна схема застосування генетичних алгоритмів до нейронних мереж полягає

у наступному. На першому етапі варто обрати спосіб кодування суттєвих параметрів нейронної мережі у вигляді бінарного або числового вектора. Такі вектори формують деяку базову множину розв'язків X , у якій здійснюється пошук оптимального розв'язку. Скінченні непорожні множини X називаються популяціями [10].

Наступним кроком є вибір цільової функції $f: X \rightarrow R^1$, яка визначатиме найбільш вдалу архітектуру окремих нейронних мереж та структуру ансамблю загалом. Така функція повинна бути близькою за значенням до функції втрат, що використовуються при навчанні нейромереж.

Далі визначаються стандартні для генетичних алгоритмів оператори: селекції, кросовера, мутації та відбору. Докладно зміст та призначення цих операторів викладено в монографіях [3, 10].

Оскільки саме згорткові нейронні мережі відіграють суттєву роль у задачах розпізнавання образів, розглянемо далі останні публікації із застосування генетичних алгоритмів до структурної та параметричної оптимізації цього типу мереж.

Огляд підходів до використання генетичних алгоритмів для оптимізації згорткових нейронних мереж розглянуто в [11-21]. У практичних застосуваннях генетичних алгоритмів до нейронних мереж розв'язуються задачі вибору найбільш оптимальної структури шарів, оптимізація гіперпараметрів та вибір функції втрат. Основними задачами в цьому випадку є вибір методу кодування можливих розв'язків та виду генетичних операторів. Варто також зазначити, що кількість потенційних мережевих структур експоненційно збільшується з кількістю шарів у мережі [12, 13], що робить генетичні алгоритми ефективним підходом для пошуку в такому великому просторі можливих розв'язків.

Робота [14] присвячена застосуванню методів генетичного програмування для оптимізації архітектури згорткової нейронної мережі, яка, своєю чергою, представлена направленим ациклічним графом. Перевагою цього методу подання є його гнучкість, оскільки він може представляти мережеві архітектури змінної довжини. Параметрична оп-

тимізація нейронних мереж базової множини виконується методом стохастичного градієнтного спуску. Функція втрат при навчанні мережі є цільовою функцією генетичного програмування.

Автори [15] досліджують застосування генетичних алгоритмів до оптимізації глибоких нейронних мереж. Множина можливих розв'язків ділиться на підмножини параметрів для згорткових шарів та повнозв'язних шарів. Виділяються такі параметри, як розміри ядра згортки у шарах згортки, типи функцій активації, розмір вікна шару субдискретизації, кількість нейронів у повнозв'язних шарах. Генетичний алгоритм використовується у комбінації з методом прогнозування ефективності згенерованої структури мережі. Передбачена за допомогою нейронної мережі ефективність визначає перспективність використання архітектури, згенерованої генетичним алгоритмом.

У роботі [16] пропонується спосіб кодування можливих розв'язків, при якому кількість шарів нейронної мережі є необмеженою, пропонується використання блоку, з пропусками з'єднань між нейронами, який може використовуватися для уникнення проблеми зникання градієнта.

Підходи до оптимізації гіперпараметрів згорткових нейронних мереж розглянуто у роботах [17, 18]. Пропонується підхід вибору найкращої конфігурації. У роботі [18] замість стандартного для згорткових мереж ядра згортки використовується ядро Габора, яке відповідно застосовує перетворення Габора до певної підмножини даних. Відзначається, що використання ядер Габора зменшує час навчання нейронних мереж.

У роботах [19, 20] зазначається, що у практичних застосуваннях тільки частина гіперпараметрів нейронної мережі суттєво впливає на точність прогнозування, а для різних наборів вхідних даних оптимальними можуть бути різні набори гіперпараметрів. Розглядається задача класифікації зображень та досліджується зв'язок між характеристиками зображень, які поступають на вхід нейронної мережі, та її гіперпараметрами. Особливістю роботи є те, що використовуючи збільшення роздільної здатності

зображень під час процесу оптимізації, досягається підвищення ефективності системи загалом.

Для визначення генетичного алгоритму необхідно задати спосіб кодування векторів множини можливих розв'язків та описати оператори селекції, кросовера, мутації та відбору, а також задати правило припинення обчислень. Далі розглядатимемо випадок, коли відома загальна архітектура нейронної мережі – згортова. Така архітектура обумовлена природою задачі та даних, оскільки необхідно розробити адаптивну систему класифікації аудіофайлів. Отже, найбільш суттєвими гіперпараметрами нейронної мережі будуть наступні: кількість шарів згортки та повнозв'язних шарів, кількість епох навчання, кількість мереж у Snapshot ансамблі. При цьому елементами множини можливих розв'язків будуть цілочисельні вектори, де кожна позиція вектора відповідає одному з параметрів, що перераховані вище.

У статті [21] описано підхід до побудови системи розпізнавання з використанням Snapshot ансамблю, який поєднує оптимізовані генетичним алгоритмом нейронні мережі.

Отже, метою цієї роботи є розширення досліджень [21] у частині оптимізації параметрів Snapshot ансамблю, наприклад кількості елементів ансамблю.

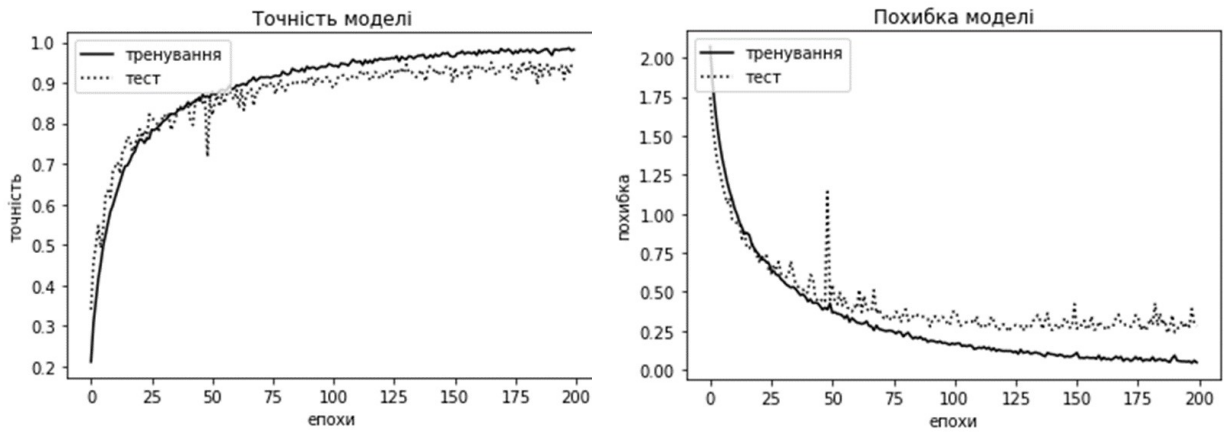
3. Обчислювальні експерименти. Обчислювальні експерименти проводяться з набором даних, що містить 3449 звукових файлів у форматі .wav для навчання та тестування системи. Вибірка для навчання містить звукові файли з 9 категорій. Здебільшого це вуличні міські звуки: шум транспорту, автомобільні сирени, дитячий сміх тощо. Мінімальна кількість файлів в одній категорії – 94, максимальна – 300. Тривалість звукових файлів переважно становить 4 с.

Для ілюстрації роботи генетичного алгоритму наводяться значення згаданих вище гіперпараметрів нейронної мережі та значення відповідних метрик точності. При цьому комбінації параметрів обираються з різних популяцій для демонстрації збіжності генетичного алгоритму. Результати обчислень наведено в таблиці 1. Використано наступні позначення: Conv2D – шар згортки; Dense – повнозв'язний шар.

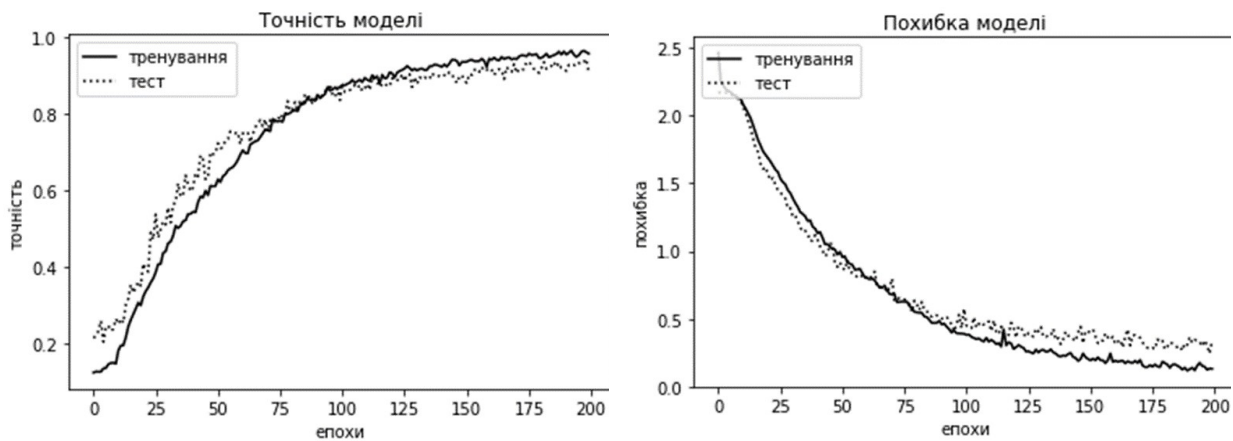
Табл. 1. Етапи роботи генетичного алгоритму

Номер популяції	Кількість шарів Conv2D	Кількість шарів Dense	Кількість епох	Кількість мереж у Snapshot ансамблі	Точність
2	1	4	50	100	0,876963
3	2	4	100	100	0,908377
4	2	2	100	100	0,939791
5	1	3	150	100	0,942408
6	1	2	200	100	0,954026

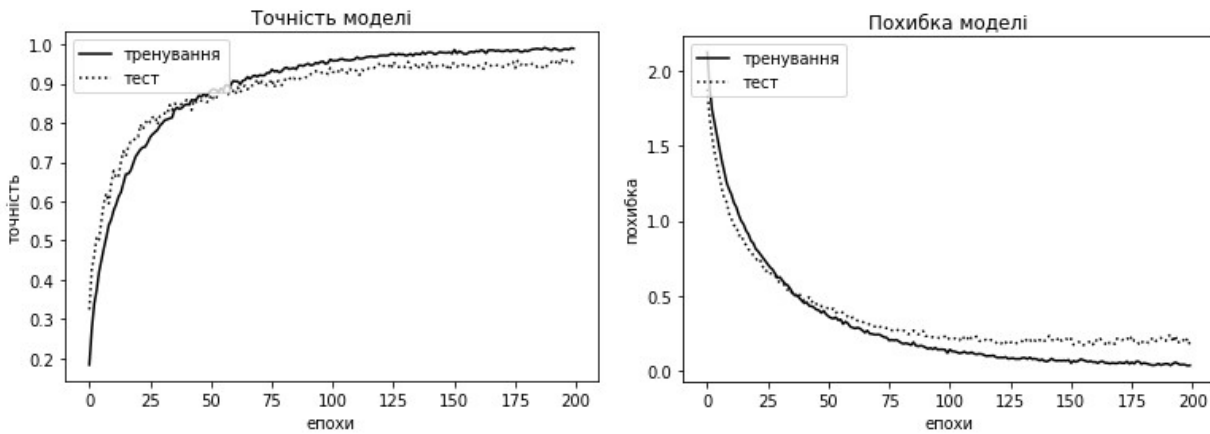
На рисунку 1 наведено вигляд функції залежності точності від кількості епох при різних наборах гіперпараметрів.



а) 2xConv2D, 2xDense



б) 2xConv2D, 4xDense



в) 1xConv2D, 2xDense

Рис. 1. Точність та похибка моделі з різною кількістю повнозв'язних шарів (Dense) та згортки (Conv2D)

Варіанти конфігурацій а) та б), наведені на рисунку 1, свідчать про недостатню стійкість моделей при варіюванні кількості епох, отже, результати, отримані за допомогою цих моделей, можуть бути не достатньо достовірними в разі зменшення кількості епох з метою оптимізації часу розрахунку. Навпаки, модель в) демонструє відносну стійкість після ста епох навчання.

Висновки. З аналізу літературних джерел можна зробити висновок, що задачу параметричної оптимізації ансамблів нейронних мереж розглянуто недостатньо. Зокрема, для автоматичного налаштування параметрів такого методу ансамблювання, як Snapshot ансамбль, не застосовувалися генетичні алгоритми. Водночас метод Snapshot

дозволяє будувати моделі з множиною нейронних мереж без суттєвого збільшення витрат ресурсів обчислень.

У роботі пропонується використання генетичних алгоритмів одночасно і для підбору оптимальних параметрів Snapshot ансамблю, і для оптимізації згорткових нейронних мереж, які входять в ансамбль.

Проведено обчислювальні експерименти та продемонстровано збіжність генетичного алгоритму.

Перспективи подальших досліджень пов'язано з розширенням множини параметрів, за якими проводиться оптимізація, та застосуванням варіантів генетичних алгоритмів, для яких можлива оцінка збіжності, наприклад еволюційно-фрагментарних методів.

Література

1. Geron A. Hands-On Machine Learning with Scikit-Learn and TensorFlow. Sebastopol: O'Reilly. 2017. 861 p.
2. Николенко С., Кадурын А., Архангельская Е. Глубокое обучение. Санкт-Петербург: Питер, 2018. 480 с.
3. Олійник А. О., Субботін С. О., Олійник О. О. Еволюційні обчислення та програмування. Запоріжжя: ЗНТУ. 2010. 324 с.
4. Li K., Liu W., Zhao K., Shao M., Liu L. A Novel Dynamic Weight Neural Network Ensemble Model. *International Journal of Distributed Sensor Networks*. Vol. 2015. 2015. Article ID 862056, 13 pages. DOI: 10.1155/2015/862056.
5. Tao S. Deep Neural Network Ensembles. URL: <https://arxiv.org/abs/1904.05488>.
6. Webb A.M., Reynolds C., Iliescu D.-A., Reeve H., Lujan M., Brown G. Joint Training of Neural Network Ensembles. URL: <https://arxiv.org/abs/1902.04422>.
7. Sallam H., Regazzoni C.S., Talkhan I., Atiya A. Evolving neural networks ensembles nnes. *IAPR Workshop on Cognitive Information Processing*. P. 142–147.
8. Symone G. Soares, Carlos H. Antunes, Rui Arajo. A Genetic Algorithm for Designing Neural Network Ensembles. *Proceedings of the 14th annual conference on Genetic and evolutionary computation*. P. 681–688.
9. Huang G. et al. Snapshot Ensembles: Train 1, Get M for Free. arXiv, 2017. URL: <http://arxiv.org/abs/1704.00109>.
10. Козин И.В. Эволюционные модели в дискретной оптимизации. Запорожье: ЗНУ. 2019. 204 с.
11. Lone M.A., Islam M. A Brief Overview of Developing Convolutional Neural Network Using Genetic Algorithm. *International Journal of Computer Sciences and Engineering*. Vol. 7, Issue 2, 2019. P. 812-818. DOI: 10.26438/ijcse/v7i2.812818.
12. Lingxi Xie, Alan Yuille. Genetic CNN. <https://arxiv.org/abs/1703.01513>.
13. Alejandro Baldominos Yago Saez, Pedro Isasi. Evolutionary convolutional neural networks: An application to handwriting recognition. *Neurocomputing*. Vol. 283. 2018. P. 38-52. DOI: 10.1016/j.neucom.2017.12.049.
14. Suganuma M., Shirakawa S., Nagao T. A Genetic Programming Approach to Designing Convolutional Neural Network Architectures. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*. P. 5369–5373.
15. Mattioli F., Caetano D., Cardoso A., Naves E., Lamounier E. An Experiment on the Use of Genetic Algorithms for Topology Selection in Deep Learning. *Journal of Electrical and Computer Engineering*. Vol. 2019. 2019. Article ID 3217542, P. 1-12, DOI: 10.1155/2019/3217542.
16. Sun Y., Xue B., Zhang M., Yen G.G. Automatically Designing CNN Architectures Using Genetic Algorithm for Image Classification. URL: <https://arxiv.org/abs/1808.03818>.

17. Loussaief S., Abdelkrim A. Convolutional Neural Network Hyper-Parameters Optimization based on Genetic Algorithms. *International Journal of Advanced Computer Science and Applications*. Vol. 9, No. 10, 2018. P. 252–266.
18. Meng F., Wang X., Shao F., Wang D., Hua X. Energy-Efficient Gabor Kernels in Neural Networks with Genetic Algorithm Training Method. *Electronics*. Vol. 8, №105. 2019. Doi:10.3390/electronics8010105.
19. Hinz T., Navarro-Guerrero N., Magg S., Wermter S. Speeding up the Hyperparameter Optimization of Deep Convolutional Neural Networks. *International Journal of Computational Intelligence and Applications*. Vol. 17, №. 2. 2018. Article ID 1850008. P. 1–15.
20. Molina-Cabello M.A., Accino C., López-Rubio E., Thurnhofer-Hemsi K. Optimization of Convolutional Neural Network Ensemble Classifiers by Genetic Algorithms. In: *Rojas I., Joya G., Catala A. (eds) Advances in Computational Intelligence. IWANN 2019. Lecture Notes in Computer Science*. Vol. 11507. 2019. P. 163–173.
21. Gonzalez J.A., Hurtado L.-F., Pla F. ELiRF-UPV at SemEval-2019 Task 3: Snapshot Ensemble of Hierarchical Convolutional Neural Networks for Contextual Emotion Detection. *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019)*. 2019. P. 195–199.

References

1. Geron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. Sebastopol: O'Reilly.
2. Nikolenko, S., Kadurin, A. & Arhangelskaya, E. (2018). *Deep learning*. Saint Petersburg: Piter (In Russian).
3. Oliinyk, A. O., Subbotin, S. O. & Oliinyk, O. O. (2010). *Evoliutsiini obchyslennia ta prohramuvannia. Zaporizhzhia: ZNTU*. (In Ukrainian).
4. Li, K., Liu, W., Zhao, K., Shao, M. & Liu, L. (2015). A Novel Dynamic Weight Neural Network Ensemble Model. *International Journal of Distributed Sensor Networks*, Vol. 2015, Article ID 862056, 13 pages. DOI: 10.1155/2015/862056.
5. Tao, S. *Deep Neural Network Ensembles*. Retrieved from <https://arxiv.org/abs/1904.05488>.
6. Webb, A. M., Reynolds, C., Iliescu, D.-A., Reeve, H., Lujan, M. & Brown, G. *Joint Training of Neural Network Ensembles*. Retrieved from <https://arxiv.org/abs/1902.04422>.
7. Sallam, H., Regazzoni, C. S., Talkhan, I. & Atiya, A. (2008). Evolving neural networks ensembles. *IAPR Workshop on Cognitive Information Processing*, pp. 142–147.
8. Symone, G., Soares, Carlos, H. Antunes & Rui Arajo. (2012). A Genetic Algorithm for Designing Neural Network Ensembles. *Proceedings of the 14th annual conference on Genetic and evolutionary computation*, pp. 681–688.
9. Huang, G. et al. (2017). *Snapshot Ensembles: Train 1, Get M for Free*. arXiv. Retrieved from <http://arxiv.org/abs/1704.00109>.
10. Kozin, I. V. (2019). *Evolutsionnyie modeli v diskretnoy optimizatsii. Zaporozhe: ZNU*. (in Russian).
11. Lone, M. A. & Islam, M. (2019). A Brief Overview of Developing Convolutional Neural Network Using Genetic Algorithm. *International Journal of Computer Sciences and Engineering*, Vol. 7, Issue 2, pp. 812–818. DOI: 10.26438/ijcse/v7i2.812818.
12. Lingxi, Xie & Alan, Yuille. *Genetic CNN*. Retrieved from <https://arxiv.org/abs/1703.01513>.
13. Alejandro Baldominos Yago Saez & Pedro Isasi. (2019). Evolutionary convolutional neural networks: An application to handwriting recognition. *Neurocomputing*, Vol. 283, pp. 38–52. DOI: 10.1016/j.neucom.2017.12.049.
14. Suganuma, M., Shirakawa, S. & Nagao, T. (2018). A Genetic Programming Approach to Designing Convolutional Neural Network Architectures. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*, pp. 5369–5373.

15. Mattioli, F., Caetano, D., Cardoso, A., Naves, E. & Lamounier, E. (2019). An Experiment on the Use of Genetic Algorithms for Topology Selection in Deep Learning. *Journal of Electrical and Computer Engineering*, Vol. 2019, Article ID 3217542, pp. 1–12. DOI: 10.1155/2019/3217542.
16. Sun, Y., Xue, B., Zhang, M. & Yen, G. G. Automatically Designing CNN Architectures Using Genetic Algorithm for Image Classification. Retrieved from <https://arxiv.org/abs/1808.03818>.
17. Loussaief, S. & Abdelkrim, A. (2018). Convolutional Neural Network Hyper-Parameters Optimization based on Genetic Algorithms. *International Journal of Advanced Computer Science and Applications*, Vol. 9, No. 10, pp. 252–266.
18. Meng, F., Wang, X., Shao, F., Wang, D. & Hua, X. (2019). Energy-Efficient Gabor Kernels in Neural Networks with Genetic Algorithm Training Method. *Electronics*, Vol. 8, No. 105. doi:10.3390/electronics8010105.
19. Hinz, T., Navarro-Guerrero, N., Magg, S. & Wermter, S. (2018). Speeding up the Hyperparameter Optimization of Deep Convolutional Neural Networks. *International Journal of Computational Intelligence and Applications*, Vol. 17, No. 2, Article ID 1850008, pp. 1–15.
20. Molina-Cabello, M. A., Accino, C., López-Rubio, E. & Thurnhofer-Hemsi, K. (2019). Optimization of Convolutional Neural Network Ensemble Classifiers by Genetic Algorithms. In: Rojas I., Joya G., Catala A. (eds) *Advances in Computational Intelligence. IWANN 2019. Lecture Notes in Computer Science*, Vol. 11507, pp. 163–173.
21. Gonzalez, J. A., Hurtado, L.-F. & Pla, F. (2019). ELiRF-UPV at SemEval-2019 Task 3: Snapshot Ensemble of Hierarchical Convolutional Neural Networks for Contextual Emotion Detection. *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019)*, (pp. 195–199).