

3. Zhong Y. Computational intelligence in optical remote sensing image processing / Y. Zhong, A. Ma, Y. S. Ong, Z. Zhu, and L. Zhang // Appl. Soft Comput, Mar. 2018, vol. 64, pp. 75–93.

4. Soulard C. E. Removing rural roads from the national land cover database to create improved urban maps for the United States, 1992 to 2011 / C. E. Soulard, W. Acevedo, and S. V. Stehman // Photogramm. Eng. Remote Sens, Feb. 2018, vol. 84, no. 2, pp. 101–109.

5. Frick A. A framework for the long-term monitoring of urban green volume based on multi-temporal and multi-sensoral remote sensing data / A. Frick and S. Tervooren // J. Geovisualization Spatial Anal, Mar. 2019, vol. 3, no. 1, p. 6.

6. Порогові методи (сегментація) – національна бібліотека ім. Н. Е. Баумана [Електронний ресурс] // Національна бібліотека ім. Н. Е. Баумана. – Режим доступу: [https://ru.bmstu.wiki/Пороговые\\_методы\\_\(Сегментация\)](https://ru.bmstu.wiki/Пороговые_методы_(Сегментация)) (дата звернення: 09.02.2021).

7. Ronneberger O. U-Net: Convolutional networks for biomedical image segmentation / Ronneberger O., Fischer P., and Brox T. // Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent, 2015, vol. 9351, pp. 234–241.

8. Wiedemann C. Empirical evaluation of automatically extracted road axes / C. Wiedemann, C. Heipke, H. Mayer, and O. Jamet // Proc. CVPR Workshop Empirical Eval. Methods Comput. Vis, 1998, pp. 172–187.

УДК 004.93

К.С. Хабарлак<sup>1</sup>

<sup>1</sup>Національний технічний університет «Дніпровська політехніка», Дніпро, Україна

## ПРИСКОРЕНЕ НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ ЗА ДЕКІЛЬКОМА ПРИКЛАДАМИ

**Анотація.** Нейронні мережі мають значну якість вирішення багатьох практично значних задач, але їх недоліком є потреба у великому наборі даних для тренування. Одним із підходів до навчання за малою кількістю прикладів є мета-навчання, однак, цей підхід є обчислювально складним. В даній роботі запропоновано новий спосіб прискорення оптимізаційного мета-навчання.

**Ключові слова:** нейронні мережі, мета-навчання, оптимізація, час виконання.

**Вступ.** Нейронні мережі якісно вирішують багато практичних задач. Насамперед це задачі класифікації зображень, де якість згорткових нейронних мереж має надзвичайно малу помилку. Такі результати досягаються за допомогою навчання на великій розміченій вибірці даних. Однак, для кожної окремої практичної задачі таку вибірку необхідно зібрати та вручну розмітити, що є ресурсозатратним. Одним із підходів до зменшення необхідної

навчальної вибірки є так зване мета-навчання. В даній роботі буде розглянуто найбільш поширений алгоритм оптимізаційного мета-навчання, а саме Model-Agnostic Meta-Learning (MAML), опис якого було наведено в роботі [1]. Однак, цей метод є обчислювально складним, що значно обмежує його практичну застосовність. Отже, метою цієї роботи є прискорення фази адаптації алгоритму MAML.

**Постановка задачі.** Для досягнення поставленої мети в роботі сформовані і вирішені такі завдання:

- розробити власну реалізацію алгоритму MAML для оптимізаційного мета навчання;
- отримати результати щодо точності і часу адаптації для набору даних CIFAR-FS;
- досягти значного пришвидшення роботи фази адаптації алгоритму MAML за допомогою запропонованих  $\Lambda$ -шаблонів.

**Основний зміст роботи.** На відміну від звичайного навчання нейронних мереж, яке оперує на окремих зображеннях, мета навчання вводить поняття «задач». Задача складається із набору «підтримки» та «запитів». Кожна із задач має різні класи (ярлики), що необхідно класифікувати, що значно відрізняє мета-навчання від звичного навчання. За допомогою такої схеми очікується, що алгоритм навчиться швидко пристосовуватися до нових задач та вивчить ознаки, що властиві не тільки декільком фіксованим класам, але і всім класам в цілому. Так, зазвичай, для мета-навчання достатньої 1-10 навчальних прикладів на клас, коли для звичайного необхідно зібрати сотні або тисячі прикладів. MAML є одним із прикладів оптимізаційного мета навчання – це значить, що для того, щоб навчити довільну нейронну мережу за описаною вище процедурою, достатньо змінити оптимізаційну задачу, а архітектура мережі залишиться незмінною. Для MAML алгоритм оптимізації розподіляється на 2 частини: адаптація до нової задачі і навчання мета-параметрів. MAML є одним із найбільш широко використовуваних методів мета-навчання, однак його недоліком є повільний час виконання фази адаптації. В роботі [2] було показано ряд практично-значних задач для яких час виконання нейронної мережі грає ключову роль. В цій роботі буде зменшено час виконання алгоритму MAML.

Введемо поняття шаблону адаптації  $\Lambda$  згідно [3]. Слід відмітити, що така ідея застосовна до мережі із будь-якою архітектурою, однак, в даній роботі для більш легкої відтворюваності результатів, ми використаємо згорткову мережу, що було запропоновано в [1]. Така мережа складається із 4 згорткових та 1 повнозв'язного шара, як відображено на рис. 1. Для такої мережі шаблон  $\Lambda$  матиме наступний виклад:  $\Lambda = \{\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5\}$ , де  $\Lambda_i = \{0, 1\}$  – індикаторна функція, що визначає чи необхідно оновлювати  $i$ -й шар мережі під час навчання. У табл. 1 показано залежність часу адаптації мережі від шаблону адаптації та кількості кроків. Експерименти було проведено для декількох значень кількості кроків адаптації: 1, 3, 5, 10. Така невеличка кількість кроків адаптації є типовою для мета-навчання. Використання шаблону адаптації  $\Lambda = \{1, 0, 1, 1, 1\}$  та 3-х кроків адаптації дозволило прискорити роботу алгоритму у 3 рази.

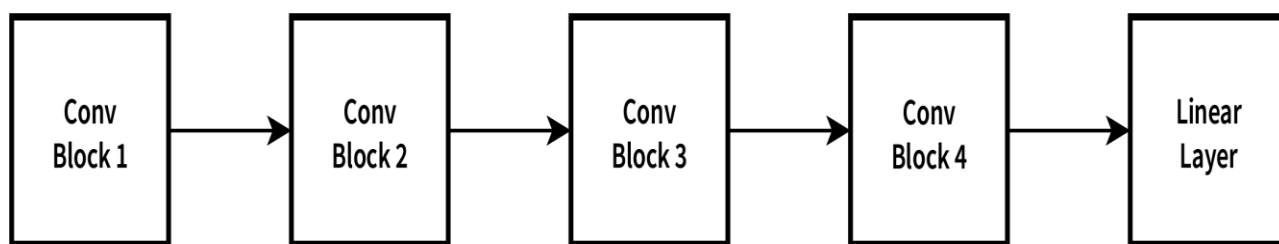


Рис. 1. Схема використаної мережі для мети-навчання

Таблиця 1

**Залежність якості та часу виконання від шаблону адаптації  $\Lambda$  та кількості кроків адаптації**

Кроків адаптації	Шаблон $\Lambda$	1 приклад 2 класи (%)	1 приклад 5 класів (%)		Середній час адаптації (мс)	Відносне прискорення (разів)
3	1,0,1,1,1	76.6	49.3		13.9	3.0
5	0,1,1,1,1	75.2	51.5		20.0	2.1
10	1,0,1,1,1	77.1	51.7		38.6	1.1
10	1,1,1,1,1	77.2	51.7		41.5	1.0

**Наукова новизна** полягає у запропонованому новому підході лямбда-шаблонів для зменшення часу адаптації нейронних мереж для випадку навчання за малою кількістю прикладів. Даний підхід дозволив зменшити час виконання у 3 рази при незначних втратах якості.

**Висновки.** В роботі запропоновано новий спосіб зменшення часу адаптації алгоритмів оптимізаційного мета-навчання. Показано трикратний приріст швидкості роботи алгоритму MAML. Очікуємо, що даний підхід бути у майбутньому застосовним до широкого переліку інших алгоритмів оптимізаційного мета-навчання, що значно збільшить їх застосовність у практичних задачах.

**ПЕРЕЛІК ПОСИЛАНЬ**

1. Finn C. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks / C. Finn, P. Abbeel, S. Levine // Proceedings of the 34th International Conference on Machine Learning: Proceedings of Machine Learning Research. – PMLR, 2017. – Вип. 70. – С. 1126-1135.
2. Khabarlak K. Fast Facial Landmark Detection and Applications: A Survey / K. Khabarlak, L. Koriashkina // arXiv:2101.10808 [cs]. – 2021.
3. Khabarlak K.S. Faster Optimization-Based Meta-Learning Adaptation Phase // Radio Electronics, Computer Science, Control. – 2022. – № 1.