

Хабарлак К.С. асистент кафедри Системного аналізу та управління

(Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна)

АНАЛІЗ ШАБЛОНІВ АДАПТАЦІЇ МЕТА-НАВЧАННЯ

Нейронні мережі показують значну якість вирішення багатьох 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 в свою чергу вимагає значного вкладення людських та грошових ресурсів. Для ряду нових проблем збір великої якісно розміченої вибірки і зовсім неможливих – таких даних просто може ще не бути в наявності. Одним із підходів до зменшення необхідної навчальної вибірки є так зване мета-навчання. Одним із ключових алгоритмів є Model-Agnostic Meta-Learning (MAML), опис якого було наведено в роботі [1]. Що ж собою являє мета-навчання? Звичайні мережі для класифікації зображень навчаються на окремих зображеннях. Одна ітерація алгоритму мінімізує помилку на таких зображеннях. В мета-навчанні однією ітерацією є так звана «задача». Задача складається із набору «підтримки» та «запитів». Кожна із задач має різні класи (ярлики), що необхідно класифікувати. Це 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, а 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 ми зменшимо час виконання алгоритму MAML.

Введемо поняття шаблону адаптації Λ . 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. Для такої мережі шаблон Λ 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. Розглянемо всі тривіальні шаблони Λ , тобто такі в яких активним буде лише 1 шар мережі. На рис. 1 зображено залежність часу адаптації мережі для таких шаблонів. За віссю X позначено єдиний активних шар, а за Y час виконання у мілісекундах. Експерименти було проведено для декількох значень кількості кроків адаптації: 1, 3, 5, 10. Така невеличка кількість кроків адаптації є типовою для мета-

навчання. Можна побачити, що 1) чим ближчий активний шар мережі до її кінця, тим швидше проходить адаптація; 2) також є істотна залежність від кількості кроків адаптації. Використовуючи дані спостереження, у подальших експериментах нами буде показано, що завдяки оновленню лише частини шарів нейронної мережі можна досягти 3-кратного зменшення часу адаптації при незначній втраті якості.

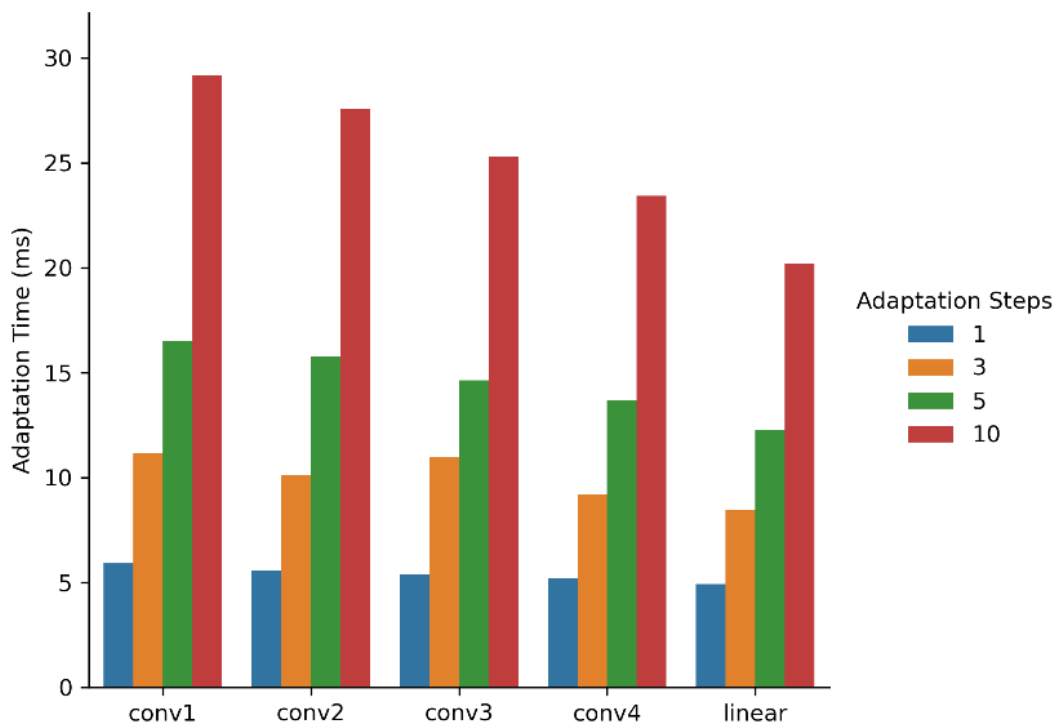


Рисунок 1 – Залежність часу адаптації мета-навчання від поточного активного шару, що відповідає шаблону Λ та кількості кроків адаптації

Висновки. В роботі було запропоновано шаблони адаптації Λ , проаналізовано якість роботи мережі та час її виконання в залежності від обраного шаблону Λ . Завдяки чому вдалося значно зменшити час виконання фази адаптації алгоритму мета навчання MAML. Сподіваємось, що наведена нова модифікація алгоритму дозволить значно поширити застосовність методу в 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 / ред. D. Precup, Y.W. Teh. – 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.