

МЕТОДИ ОПТИМІЗАЦІЇ SWAP ПОКАЗНИКІВ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИЯВЛЕННЯ ПІДМІНИ ОБЛИЧЬ

SWAP METRICS OPTIMIZATION METHODS FOR MOBILE FACE ANTI-SPOOFING NEURAL NETWORKS

Виявлення підміни обличчя (FAS) має вирішальне значення для захисту систем розпізнавання обличчя, особливо на мобільних пристроях, де обчислювальні ресурси, пам'ять та енергія є обмеженими. Традиційні нейронні мережі FAS часто занадто великі та вимогливі до ресурсів для мобільного розгортання, що робить їхню оптимізацію важливою. Завдання полягає в збереженні високої точності виявлення при дотриманні строгих вимог до розміру, швидкодії та енергоефективності (SWAP) моделі, щоб забезпечити продуктивність у реальному часі та ефективність на мобільних платформах.

Оптимізація моделей FAS для мобільних пристроїв передбачає кілька методів. «Квантування моделі» (Model Quantization) є широко використовуваною технікою, яка знижує точність вагових коефіцієнтів моделі та активації з 32-розрядного числа з плаваючою точкою до нижчого розряду, наприклад 8-розрядного цілого числа. Цей підхід значно зменшує розмір моделі та прискорює роботу з мінімальним впливом на точність. Метод «Виключення ваг» (Pruning) видаляє зайві ваги, щоб мінімізувати використання обчислень і пам'яті. «Дистиляція знань» (Knowledge Distillation) тренує компактну модель учня, щоб імітувати роботу більшої моделі-вчителя, фіксуючи основні шаблони з меншою складністю. Полегшені архітектури, такі як MobileNet, ShuffleNet та EfficientNet, оптимізують дизайн моделі для мобільних платформ. «Пошук нейронної архітектури» автоматизує виявлення ефективних архітектур. «Навчання зі змішаною точністю» поєднує різні числові точності, наприклад, 16-бітні та 32-бітні числа з плаваючою комою, під час навчання та роботи моделі, з ціллю зменшення споживання пам'яті та прискорення обчислення без значної втрати точності.

Аналіз методів оптимізації SWAP показав, що підхід «Дистиляція знань», зокрема «дистиляція карти ознак», ідеально підходить для мобільного розгортання, дозволяючи малій моделі-учню на полегшеній архітектурі відтворювати точність великої моделі-вчителя, яка працює на непридатних для мобільних платформ архітектурах, наприклад ResNet. Такий змішаний підхід дозволяє учню імітувати активації проміжного рівня вчителя, покращуючи його здатність захоплювати дрібні деталі, необхідні для захисту від підробки обличчя. Окрім того, це покращує доменне узагальнення моделі-учня щодо невідомих атак підміни обличчя.

Література

1. Yu Z., Wan J., Li X., Li S., Zhao G. NAS-FAS: Static-Dynamic Central Difference Network Search for Face Anti-Spoofing. URL: <https://doi.org/10.1109/TPAMI.2020.3036338>.
2. Cheng T., Zhang Y., Yin Y., Zimmermann R., Yu Z., Guo B. A Multi-Teacher Assisted Knowledge Distillation Approach for Enhanced Face Image Authentication. URL: <https://doi.org/10.1145/3591106.3592280>.
3. Kong Z., Zhang W., Wang T., Zhang K., Li Y., Tang X., Luo W., Dual Teacher Knowledge Distillation with Domain Alignment for Face Anti-spoofing. URL: <https://arxiv.org/html/2401.01102v1>.