

УДК 004.9

Паничук К.С.¹, Примаков В.В.², Туркевич В.М.³, Степаненко О.О.⁴

¹ студ. гр. КНТ-114м НУ «Запорізька політехніка»

² студ. гр. КНТ-223м НУ «Запорізька політехніка»

³ асп. НУ «Запорізька політехніка»

⁴ канд. техн. наук, доц. НУ «Запорізька політехніка»

ДОСЛІДЖЕННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗОБРАЖЕННІ НА ОСНОВІ СІАМСЬКИХ НЕЙРОННИХ МЕРЕЖ

З розвитком комп'ютерних наук та машинного навчання методи порівняння даних та пошуку схожих об'єктів набувають все більшої актуальності. Традиційні метрики, такі як евклідова та косинусна відстані, коефіцієнти кореляції Пірсона і Спірмена тощо [1], добре працюють при

порівнянні даних однакового типу, однак є неефективними для об'єктів зі складною структурою. Однією з перспективних архітектур для таких задач є сіамські нейронні мережі [1], що дозволяють ефективно визначати подібність між зображеннями, наприклад, у задачах розпізнавання рукописних підписів та біометричної автентифікації. Розроблене дослідження має на меті аналіз ефективності сіамських нейронних мереж у задачах класифікації та пошуку схожих зображень. Для реалізації використовувався фреймворк PyTorch, що забезпечує гнучкість у зміні архітектури моделей та пришвидшує навчання завдяки підтримці обчислень на GPU. Вибіркою для навчання слугував набір даних MNIST, який містить 60 000 зображень рукописних цифр для тренування та 10 000 для тестування. Аналіз предметної області охоплює архітектуру сіамських нейронних мереж, зменшення розмірності даних шляхом навчання інваріантного відображення, функції втрат та методи генерації пар і триплетів онлайн (стратегії Batch Hard, Batch All та Lifted Embedding) [2]-[4]. Зменшення розмірності дозволяє ефективно працювати з високовимірними даними, використовуючи метод DrLIM для навчання інваріантних представлень. Функції втрат, зокрема контрастна та триплетна, допомагають мережам краще розрізнити схожі та відмінні об'єкти. Генетичний алгоритм використовується для оптимізації відступу у функціях втрат, покращуючи здатність моделі до розрізнення навіть в умовах шуму чи зміщення в даних. Генерація пар і триплетів онлайн оптимізує процес навчання, забезпечуючи вибір найскладніших позитивних і негативних прикладів для підвищення якості класифікації.

Для порівняння були створені три варіанти мереж: згорткова нейронна мережа для класифікації, сіамська мережа з сигмоїдальною функцією активації для бінарної класифікації та сіамська мережа для пошуку ембедингів. Використання згорткових шарів дозволило автоматично виділяти суттєві ознаки зображень, що підвищило точність моделей. Архітектури сіамських мереж включали спільні згорткові шари для обох входів, а також повнозв'язні шари для остаточного порівняння ембедингів. Навчання моделей проводилося на 20 епохах із використанням оптимізатора Adam. Згорткова мережа досягла точності 98,26%, сіамська мережа з сигмоїдальною функцією активації – 98,97%, сіамська мережа з контрастною функцією втрат (онлайн та Batch Hard) – 99,56% (99,52%), сіамська мережа з триплетною функцією втрат (онлайн та Batch Hard) – 99,81% (99,83%). Результати зображено на рисунку 1.

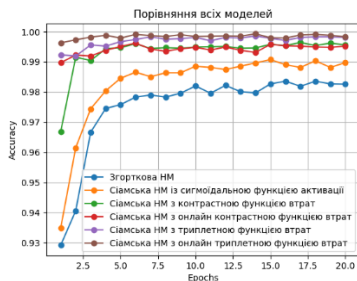


Рисунок 1 – Точність класифікації всіх мереж

Найкращі результати продемонструвала модель з триплетною функцією втрат та стратегією Batch Hard. Крім того, моделі з онлайн стратегіями вибірки образів показали кращі результати, ніж випадкові вибірки образів. Дослідження підтверджує ефективність вдосконалених методів навчання сіамських мереж у задачах розпізнавання та зіставлення об'єктів. Подальші дослідження можуть включати вдосконалення архітектури мереж, застосування більш складних вибірок та оптимізацію функцій втрат для покращення точності та продуктивності моделей.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Chicco, D. Siamese Neural Networks: An Overview / D. Chicco // Artificial Neural Networks. Methods in Molecular Biology. – New York, NY: Humana, 2020. – Vol. 2190. – P. 73–94.
2. Hadsell, R. Dimensionality Reduction by Learning an Invariant Mapping / R. Hadsell, S. Chopra, Y. LeCun // CVPR'06 : 17–22 June 2006 : proceedings. – New York, NY: IEEE, 2006. – Vol. 2. – P. 1735–1742.
3. Hermans, A. In Defense of the Triplet Loss for Person Re-Identification / A. Hermans, L. Beyer, B. Leibe // arXiv: Computer Vision and Pattern Recognition. – 2017. – P. 1–17.
4. Katoch, S. A review on genetic algorithm: past, present, and future / S. Katoch, S. S. Chauhan, V. Kumar // Multimedia Tools and Applications. – 2020. – Vol. 80. – № 5. – P. 8091–8126.