

УДК 004.852

М. Михайлів

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ПОПЕРЕДНЯ ОБРОБКА ВІДЕОЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ

UDC 004.852

М. Mykhayliv

PRE-PROCESSING OF VIDEO IMAGES USING NEURAL NETWORKS

Відео, отримані в реальних умовах зйомки, можуть містити різні артефакти, що виникають внаслідок шумів обладнання або через складні погодні умови, тому якість роботи нейронних мереж може суттєво погіршуватися. Для того, щоб отримати правильні прогнози із застосуванням нейронних мереж глибокого навчання, потрібно попередньо обробити дані, усунувши ці шуми. Попередня обробка зображень включає широкий спектр операцій, їх призначення – спрощення чи поліпшення наступного аналізу. Попередня обробка підвищує якість зображення, усуваючи ненавмисні спотворення або покращуючи деякі функції зображення, які важливі для подальшої обробки та створення більш відповідного зображення, ніж оригінал для конкретного завдання. Дії, що можуть бути виконані на цифрових зображеннях, включають точкові, локальні чи «сусідні» та глобальні операції. Точкові операції перетворюють пікселі незалежно від сусідніх пікселів. Значення сірого кольору для вихідного зображення у конкретному пікселі залежить тільки від значення сірого для того ж пікселя у вхідному зображенні. Вони відтворюють пікселі в одному зображенні, щоб сформувати інше, використовуючи одну функцію відображення.

Точкові операції не враховують просторову організацію зображення. Приклади цих операцій включають розтягування контрасту, сегментацію на основі значення сірого та вирівнювання гістограми, також середнє віднімання та масштабування за деякими факторами. У контексті глибокого навчання та класифікації відео розглянемо такі методи попередньої обробки: середнє віднімання; масштабування.

OpenCV – бібліотека з відкритим вихідним кодом, що має програмну ліцензію університету Берклі, яка включає кілька сотень алгоритмів комп'ютерного зору. Нейронна мережа застосовує модуль OpenCV, містить дві функції, які можна використовувати для попередньої обробки кадрів та підготовки їх до класифікації за допомогою моделей глибокого навчання. Тому середнє віднімання розглядається як метод, що використовується для покращення роботи згорткових нейронних мереж. Перш ніж приступити до тренування нейронних мереж глибокого навчання, необхідно обчислити середню інтенсивність кольору для кожного з червоного, зеленого та синього каналів за всіма зображеннями в навчальному наборі. Зазвичай результуючі значення являють собою три кортежі, що складаються із середнього значення червоного, зеленого та синього каналів відповідно. Наприклад, середні значення для навчального набору ImageNet є такими: $R = 103,93$, $G = 116,77$ та $B = 123,68$. Для набору даних, котрий застосовується у роботі, використовуються такі значення: $R = 114.7748$, $G = 107.7354$, $B = 99.4750$.

Однак у деяких випадках середні значення кольорів можуть бути розраховані по каналах, а не за кольорами, що призводить до матриці $M \times N$. У цьому випадку обчислюється матриця $M \times N$ для кожного каналу, потім віднімається з вхідного зображення під час навчання або тестування. Обидва методи є формами середнього віднімання, але версія, у якій застосовується інтенсивність кольору, використовується частіше, особливо для великих наборів даних.