

УДК 004.89

Яцишин В.В. канд. техн. наук, доцент, Щербаков О.О., Лова М.Р.

Тернопільський національний технічний університет імені Івана Пулюя

АНАЛІЗ БАЗ ДАНИХ ЗОБРАЖЕНЬ У ГАЛУЗІ КОМП'ЮТЕРНОГО ЗОРУ

Yatsyshyn V.V. PhD, Assoc. Prof., Scherbakov O.O., Lova M.R.

ANALYSIS OF IMAGE DATA SETS IN THE AREA OF COMPUTER VISION

Комп'ютерний зір відноситься до області штучного інтелекту, який швидко розвивається. Для того, щоб полегшити роботу при проектуванні таких систем, зокрема, моделей та їхніх архітектур, деякі фірми поділились сформованими навчальними вибірками.

Одним з таких наборів даних є модифікована база даних рукописних цифр Національного інституту стандартів і технологій (MNIST). Цей набір представляє собою елементарну колекцію даних для комп'ютерного зору, що містить 70 тисяч зображень рукописних цифр для кожного числа, тобто 0 – 9, які відформатовані у форматі сірого 28x28. Дані попередньо розділені на навчальну вибірку (60 тисяч) і тестовий набір (10 тисяч). Усі цифри розміщені у центрі зображення. Даний набір є корисним для фундаментального проекту комп'ютерного зору, коли необхідно автоматично оцифрувати рукописний текст

Набори даних CIFAR-10 і CIFAR-100 підготовлені Канадським інститутом перспективних досліджень. CIFAR-10 складається з 60 тисяч зображень розподілених за 10-ма класами. Дана колекція включає зображення літаків, автомобілів, птахів, котів, оленів, собак, жаб, коней, кораблів і вантажівок. До складу CIFAR-100 входять 60 тисяч зображень, але тепер вже передбачено 100 класів, тобто 600 зображень на кожен клас. Колекції CIFAR-10 і CIFAR-100 є зручними для використання, оскільки всі вони відформатовані у форматі 32x32 пікселі та попередньо розділені на навчальний набір із 50 тисяч зображень та тестовий набір із 10 тисяч зображень з рівною часткою даних з усіх класів.

Набір даних IMDB-Wiki містить 520 тисяч зображень обличчя, виділених з IMBD та Вікіпедії. Дані містять важливу метаінформацію, наприклад розташування обличчя на зображенні, ім'я, дата народження та стать людини на фотографіях. Цей набір даних зазвичай використовується для завдань прогнозування статі та віку.

ImageNet набір даних створено спільно Стенфордським університетом та Принстонським університетом для типового змагання з комп'ютерного зору під назвою ImageNet Large Scale Visual Recognition Challenge (ILSVRC), де командам-учасникам було запропоновано 5 основних завдань, тобто класифікація об'єктів, локалізація об'єктів, виявлення об'єктів, виявлення та розпізнавання об'єктів з відео за допомогою набору даних ImageNet. Цей набір даних побудований на основі ієрархії WordNet (лексична база даних для англійської мови), де містяться лише іменники. В середньому на кожен вузол ієрархії припадає понад 500 зображень. Всього є понад 1,4 мільйона зображень понад 220 тисяч класів. Наразі це найбільший доступний набір зображень, який є відкритим і загальнодоступним.

Набір PASCAL VOC даних було відкрито дослідницьким інститутом PASCAL, що фінансується Європейським Союзом. Дана колекція містить зображення за 4-ма основними темами і включає транспортні засоби, домашнє господарство, тварин та людей.

При реалізації комп'ютеризованої системи виявлення справжності ID-картки доцільним є використання наведених наборів даних, що дозволить забезпечити виявлення об'єктів у різних предметних областях.