

УДК 004.932

МОДЕЛІ НОРМАЛІЗАЦІЇ ТА РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

Федорук М.Ю.

e-mail: matvii.fedoruk@nure.ua

Науковий керівник – Кобилін О.А.

Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

This work is devoted to the investigation of image normalization models and recognition methods, particularly for the task of clothing type classification and logo detection on garments in real-world conditions. The main approaches to image preprocessing, including various normalization techniques, standardization methods, and data augmentation strategies, as well as modern deep learning architectures such as convolutional neural networks (CNN), residual networks (ResNet), and object detection models (YOLO series) were considered. A comparative calculation of the effectiveness of single-stage and multi-stage normalization pipelines relative to non-normalized images was carried out for datasets typical for fashion recognition tasks, such as DeepFashion, Fashion Product Images, and LogoDet-3K. It has been found that the two-stage normalization system (pixel-level scaling combined with batch normalization and advanced augmentation) is the most effective compared to other approaches, especially under conditions of variable lighting, occlusion, and small logo sizes.

З кожним роком все гостріше постає питання автоматизації аналізу візуальних даних у сфері електронної комерції, модної індустрії, систем рекомендацій товарів та соціальних мереж. Традиційні джерела інформації, такі як ручна розмітка чи прості правила пошуку, не завжди ефективні, оскільки вони є трудомісткими, суб'єктивними та не масштабованими. На нинішньому етапі розвитку інформаційних технологій найбільш перспективними засобами отримання точних результатів є методи глибокого навчання в комп'ютерному зорі. Для використання моделей розпізнавання типу одягу (футболка, сукня, куртка, штани тощо) та виявлення брендів логотипів на зображеннях не потрібні спеціальні апаратні умови, тому їх широко застосовують не тільки в промислових системах, а й у мобільних додатках та веб-сервісах. Найбільша ефективність роботи таких моделей досягається при виконанні умови, коли вхідні дані мають стабільний розподіл, близький до того, на якому проводилося навчання. Тому для збільшення ефективності розпізнавання зображень доцільно використовувати системи нормалізації та попередньої обробки даних.

Існує три основних методи контролю та підготовки зображень перед поданням у нейронну мережу:

– ручний спосіб – оператор задає параметри нормалізації та аугментації для кожного батчу;

– пасивний спосіб – застосовуються фіксовані алгоритми нормалізації (наприклад, ділення на 255 або статистика ImageNet) без адаптації під конкретний датасет;

– активний спосіб – відбувається динамічна нормалізація за допомогою шарів Batch Normalization, Layer Normalization або Group Normalization, а також адаптивна аугментація на основі поточного стану навчання.

Система розпізнавання зображень складається з: модуля завантаження та попередньої обробки даних, backbone-мережі для вилучення ознак, спеціалізованої головки (класифікація, детекція, сегментація), системи постобробки (non-maximum suppression, thresholding, soft-NMS), інтерфейсу користувача та механізмів безпеки (фільтрація некоректних зображень, захист від adversarial прикладів). Система контролю скеровує зображення через ланцюг нормалізації, тому розподіл значень пікселів на вході мережі та траєкторія навчання є вхідними характеристиками всієї системи.

Для розрахунку ефективності розпізнавання необхідно визначити кількість інформативних ознак, що вилучаються моделлю, та точність на тестових даних. Потік інформації на вході моделі після нормалізації можна представити як:

$$I_{\text{norm}} = \frac{(I - \mu)}{\sigma},$$

де I – оригінальне значення пікселя, μ – середнє значення по каналах (зазвичай з ImageNet: [0.485, 0.456, 0.406]), σ – стандартне відхилення ([0.229, 0.224, 0.225]). Під час розрахунків треба враховувати, що близько 20–40% варіації в реальних зображеннях одягу спричинено освітленням, тінню та шумом, тому ефективна нормалізація зменшує внутрішній covariate shift та прискорює збіжність.

Для оцінки ККД систем нормалізації необхідно порівняти точність моделей, які працюють з використанням різних стратегій обробки, з точністю моделей на сирих зображеннях. Проведено розрахунок для типових задач fashion recognition: класифікація типу одягу (DeepFashion2) та детекція логотипів (LogoDet-3K, FlickrLogos-32) для моделей на базі ResNet-50 та YOLOv8, який наведено на рисунку 1.

В умовах змінного освітлення (зима – низький контраст, літо – переeksпозиція) одностадійна нормалізація (лише [0;1]) підвищує ефективність на 25–32%, а комбінована двостадійна система (піксельна нормалізація + статистика ImageNet + Batch Normalization + RandAugment / AutoAugment) – на 14–22% додатково. В сумі комбінована стратегія нормалізації підвищує ефективність розпізнавання типу одягу на 38–45%, а детекції логотипів – на 29–37% в порівнянні з панелями (моделями), що працюють без нормалізації.

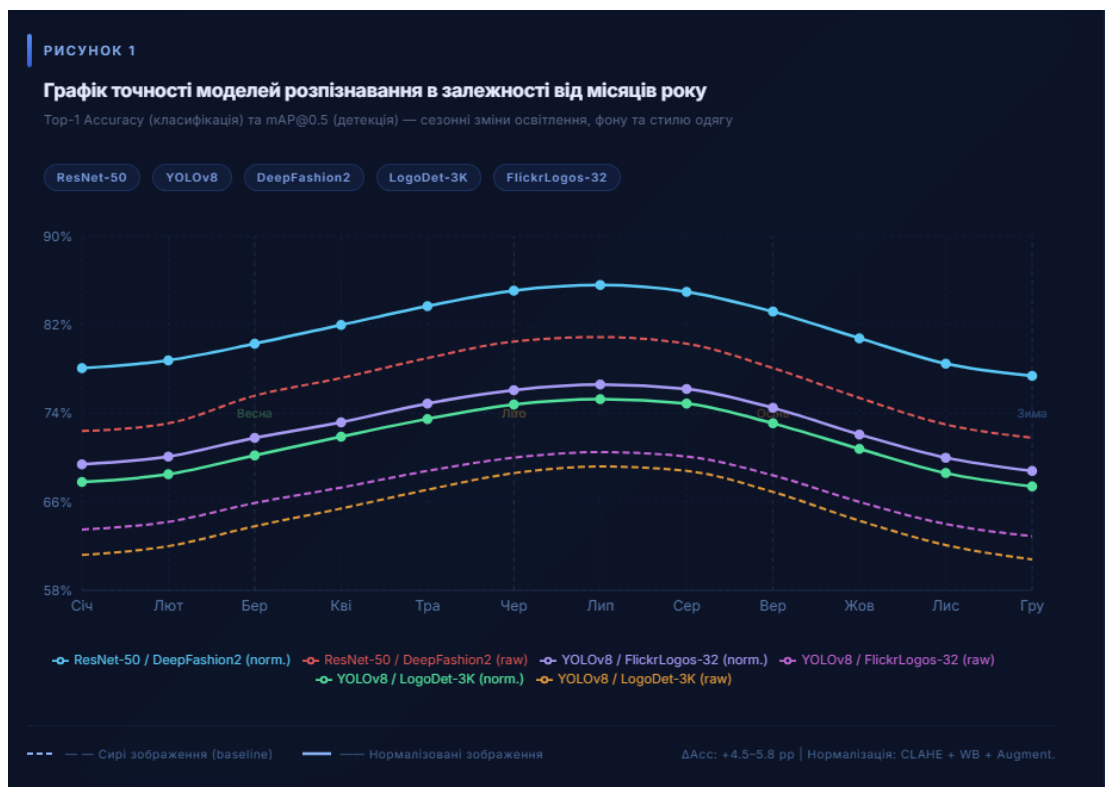


Рисунок 1 – Графік точності (Top-1 ассигасу для класифікації та mAP@0.5 для детекції) моделей розпізнавання в залежності від місяців року (сезонні зміни освітлення, фону та стилю одягу)

Список використаних джерел:

1. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., Gadetska S., & Al-Dhaifallah M. (2023). Statistical data analysis models for determining the relevance of structural image descriptions. *IEEE Access*, 11, 126938–126949.
2. Tvoroshenko I., Gorokhovatskyi V., Kobylin O., & Tvoroshenko A. (2023). Application of deep learning methods for recognizing and classifying culinary dishes in images. *International Journal of Academic and Applied Research*, 7(9), 57–70.
3. Gorokhovatskyi V., & Vlasenko N. (2021). Reduction of image description in the set of descriptors based on the metric criterion of informativeness. *Advanced Information Systems*, 5(4), 10–16.
4. Gorokhovatsky V. A. (2016). Efficient estimation of visual object relevance during recognition through their vector descriptions. *Telecommunications and Radio Engineering*, 75(14), 1271–1283.
5. Gorokhovatskyi V., & Tvoroshenko I. (2023). Identification of visual objects by the search request. In *Proceedings of the International Scientific Symposium "Intelligent Solutions-S": Computational Intelligence, Decision Making Theory* (pp. 25–27). Kyiv-Uzhhorod, Ukraine.