УДК 004.932 DOI: https://doi.org/10.30837/IYF.CVSAMM.2024.021

ОБРОБКА РУКОПИСНОГО ТЕКСТУ З ЗОБРАЖЕНЬ

Бутенко П.В.
Науковий керівник – к.т.н., доц. Кобилін О.А.
Харківський національний університет радіоелектроніки, каф. ІНФ,
м. Харків, Україна,
тел.: (057) 702-14-19, e-mail: pavlo.butenko@nure.ua

This study focuses on the efficiency of handwriting processing in images. The main emphasis is placed on the use of image binarization, segmentation, and the Hidden Markov Model algorithm. The study aims to identify key processing steps that help improve the quality of handwriting recognition. The current context of the information society emphasizes the importance of this kind of research, as the transfer of documents into electronic format is still a pressing task.

Основними процесами при обробці зображення є бінаризація, сегментація та використання алгоритму прихованих Марківських моделей. Спочатку виконується бінаризація зображення, щоб перетворити його на двохвимірну бінарну матрицю, яка визначає області тексту та фону. Бінаризація зображення є критичним етапом обробки, оскільки вона перетворює вихідне зображення у двовимірну бінарну матрицю, яка чітко розмежовує області тексту та фону [1]. В даному дослідженні для бінаризації використовується метод порогової бінаризації, який базується на виборі оптимального порогу для розділення пікселів на дві групи:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t), \quad (1)$$

де $\sigma_w^2(t)$ – дисперсія внутрішньокласового розподілу пікселів; $q_1(t)$ – ймовірність того, що піксель належить першому класу (пікселі з інтенсивністю менше від порога); $\sigma_1^2(t)$ – дисперсія інтенсивності пікселів першого класу; $q_2(t)$ – ймовірність того, що піксель належить другому класу (пікселі з інтенсивністю більше за поріг); $\sigma_2^2(t)$ – дисперсія інтенсивності пікселів другого класу.

Після розділення пікселів зображення на чорні та білі, виконується сегментація зображення з метою ефективного виділення текстових областей та підвищення якості бінаризації. Цей процес ґрунтується на методі пошуку зв'язних компонентів у бінаризованому зображенні. Для цього спершу зображення перетворюється у відтінки сірого, що дає змогу спростити подальшу обробку та поліпшити якість бінаризації [2, 3]. Формула перетворення відтінків відображається як:

$$X = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B, \quad (2)$$

де R, G і B – червоний, зелений та синій канали пікселя

Коефіцієнти для перетворення кольорового зображення в відтінки сірого не є випадковими і базуються на спектральній чутливості людського зору до різних кольорів. Зазвичай людина має більшу чутливість до зеленого кольору, меншу до синього та червоного кольорів. Коефіцієнти відповідають середньоваженим чинникам для каналів. Вони допомагають відобразити інтенсивність кожного кольору на зображенні в відтінках сірого з урахуванням їхньої відносної важливості для сприйняття людським зором [4]. Таке перетворення дозволяє зберегти якість зображення та зробити його подальшу обробку більш ефективною.

Наступним кроком в процесі сегментації є порогова фільтрація. Це спеціальний крок, що використовує заданий поріг для розділення пікселів на дві групи: одна група відповідає областям, де інтенсивність більша за поріг, а інша – областям з інтенсивністю нижче порогу.

По завершенню сегментації зображення виникає необхідність в аналізі ряду символів, що були визначені на попередньому етапі. Для цього часто використовують алгоритми прихованих Марківських моделей.

HMM – це статистична модель, яка використовується для опису випадкового процесу зі складним структурним зв’язком між внутрішніми станами системи та спостереженнями. Цей алгоритм використовується на останньому етапі, оскільки він дозволяє ефективно моделювати послідовності символів та здійснювати їх розпізнавання. Даний алгоритм включає такі основні етапи:

– алгоритм прямого поширення. Використовується для обчислення ймовірностей появи кожного можливого спостереження в кожному стані моделі;

– алгоритм зворотного розповсюдження. Використовується для обчислення ймовірностей появи кожного можливого спостереження в кожному стані моделі, враховуючи усі майбутні спостереження;

– алгоритм Вітербі. Використовується для визначення найбільш ймовірного шляху через модель, тобто для визначення послідовності станів моделі, яка найкраще пояснює спостережену послідовність.

Використання алгоритму HMM дозволяє ефективно моделювати та розпізнавати послідовності символів на зображеннях, що в свою чергу відкриває широкі можливості для автоматизованого аналізу та обробки текстової інформації [5].

По проходженню всіх етапів обробки зображення, важливо переглянути його на реальному прикладі, щоб оцінити ефективність методики.

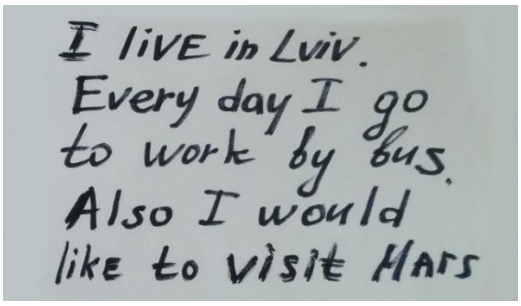

Рисунок 1 – Приклад фото з текстом

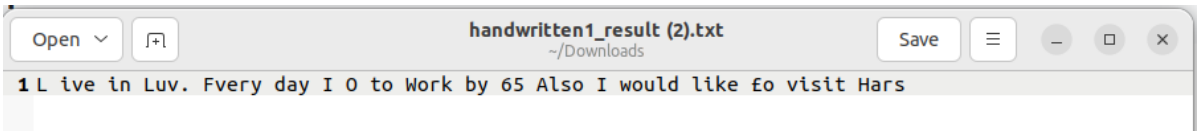

Рисунок 2 – Оброблений текст з фото

Прикладом було обране зображення з інтернету. Загалом, на ньому було присутніх 60 символів. Використання розглянутої послідовності обробки зображень дозволило успішно розпізнати 48 символів, що складає приблизно 80 % від загальної кількості символів. Поточна методика обробки є ефективною, але не ідеальною. Деякі символи залишаються нерозпізнаними, бо текст написаний людиною все ж таки має певну різницю від одного індивіда до іншого. Попри це, вона є потужним інструментом для автоматизованої обробки тексту на зображеннях.

Список використаних джерел:

1. Кобилін О. А., Творошенко І. С. Методи цифрової обробки зображень: навч. посіб. Харків : ХНУРЕ, 2021. 124 с. URL: https://doi.org/10.30837/978-966-659-295-1. (дата звернення: 10.01.2024).

2. Tvoroshenko I., Gorokhovatskyi V., Kobylin O., Tvoroshenko A. Application of deep learning methods for recognizing and classifying culinary dishes in images. International Journal of Academic and Applied Research. 2023. Vol. 7 pp. 57–70.

3. Kobylin O., Gorokhovatskyi V., Tvoroshenko I., and Peredrii O. The application of non-parametric statistics methods in image classifiers based on structural description components. Telecommunications and Radio Engineering. 2020. 79(10), pp. 855–863.

4. Gorokhovatskyi V., Tvoroshenko I., Kobylin O., Vlasenko N. Search for visual objects by request in the form of a cluster representation for the structural image description. Advances in Electrical and Electronic Engineering. 2023. 21(1), 19–27.

5. Yakovleva O., Kovtunenko A., Liubchenko V., Honcharenko V., Kobylin O. Face Detection for Video Surveillance-based Security System (COLINS-2023). In CEUR Workshop Proceedings. 2023. Vol. 3403, pp. 69–86.