

АНАЛІЗ СУЧАСНИХ МОДЕЛЕЙ НЕЙРОМЕРЕЖЕЙ ДЛЯ РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ

О.В.Корнієнко

Харківський національний університет радіоелектроніки

Україна, 61166, Харків, пр. Науки 14

E-mail: oleksii.korniienko3@nure.ua

Анотація. В роботі розглянуте актуальне питання аналізу сучасних систем розпізнавання рукописного тексту, їх цілі, методи. Їх використання обумовлено спрощенню переходу від паперових документів до цифрових.

Ключові слова: Штучний інтелект, друкований текст, розпізнавання тексту.

ANALYSIS OF MODERN NEURAL NETWORK MODELS FOR HANDWRITTEN TEXT RECOGNITION

O.V. Korniienko

Kharkiv National University of Radio Electronics

Ukraine, 61166, Kharkiv, Nauky Ave. 14

E-mail: oleksii.korniienko3@nure.ua

Abstract. The paper considers the topical issue of the analysis of modern handwritten text recognition systems, their goals, methods. Their use is due to the simplification of the transition from paper documents to digital ones.

Keywords: Artificial intelligence, printed text, text recognition.

Стрімкий розвиток технологій штучного інтелекту та комп'ютерного зору зумовив значний прогрес у галузі розпізнавання образів, зокрема в задачах оптичного розпізнавання символів та рукописного тексту. Незважаючи на глобальну тенденцію до повної комп'ютеризації, колосальний обсяг інформації у світі досі генерується, обробляється або зберігається у вигляді рукописних документів. Це створює об'єктивну потребу в розробці високоефективних автоматизованих систем перетворення фізичних записів у машинозчитуваний формат. Однак розпізнавання рукописного тексту залишається однією з найскладніших задач машинного навчання через високу індивідуальну варіативність почерків, наявність артефактів, дефектів носіїв інформації та структурну складність різних мов.

Необхідно проаналізувати сучасні системи аналізу рукописних текстів з використанням нейромереж, їх задачі, переваги та недоліки. Основні задачі які треба дослідити:

- швидкість розпізнавання;
- якість розпізнавання;
- можливість навчання;
- системні вимоги комп'ютера.

У контексті вирішення задач оптичного розпізнавання символів (OCR) та, зокрема, розпізнавання рукописного тексту (HTR), сучасний ландшафт машинного навчання пропонує низку інструментів, що базуються на різних архітектурних парадигмах. Серед найбільш досліджуваних та імплементованих рішень виділяються фреймворки PaddleOCR та EasyOCR, а також спеціалізована трансформаторна модель TrOCR. Кожна з цих систем демонструє унікальний підхід до екстракції візуальних ознак та лінгвістичного моделювання, що зумовлює їхню різну ефективність при роботі зі складними рукописними даними.

Система PaddleOCR, розроблена на базі фреймворку PaddlePaddle, представляє собою

комплексний багаторівневий конвеєр (pipeline), орієнтований на досягнення балансу між точністю розпізнавання та обчислювальною ефективністю. Сучасні ітерації цієї системи (архітектури серії PP-OCR) використовують алгоритм Differentiable Binarization (DB) для локалізації текстових регіонів та вдосконалені варіації згортково-рекурентних мереж (CRNN) у поєднанні з легковаговими трансформерами (SVTR) для безпосереднього розпізнавання. Головною науковою та практичною перевагою PaddleOCR є застосування методів дистиляції знань, квантування та прунінгу (відсікання нейронних зв'язків), що дозволяє створювати надзвичайно компактні моделі. У задачах розпізнавання рукописного тексту конвеєр PaddleOCR демонструє високу швидкість та хорошу генералізацію на структурованих бланках, проте його ефективність може знижуватися при обробці високозв'язного (курсивного) почерку без проведення додаткового донавчання на специфічних датасетах.

Гібридна парадигма просторового розпізнавання EasyOCR. EasyOCR є репрезентантом класичної, але глибоко модернізованої гібридної архітектури в галузі комп'ютерного зору. Процес виявлення тексту в цій моделі базується на мережі CRAFT (Character Region Awareness for Text Detection), яка аналізує не лише межі текстового блоку, але й просторові зв'язки між окремими символами. Етап розпізнавання реалізовано через зв'язку залишкових нейромереж (ResNet) для виділення візуальних ознак, двонаправлених мереж довгої короткострокової пам'яті (BiLSTM) для аналізу послідовностей та функції втрат Connectionist Temporal Classification (CTC), яка вирішує проблему вирівнювання несегментованих вхідних зображень і вихідних текстових рядків. Хоча EasyOCR забезпечує надійну підтримку понад вісімдесяти мов і є стандартом де-факто для розпізнавання друкованого тексту та вуличних вказівників, класичний CTC-підхід часто виявляється недостатньо гнучким для моделювання складних лінгвістичних залежностей, які є критично важливими для точної інтерпретації нерозбірливого рукописного тексту.

Наскрізна трансформаторна архітектура TrOCR. Модель TrOCR (Transformer-based Optical Character Recognition) представляє фундаментально інший, сучасний підхід до задачі НТР, повністю відмовляючись від традиційних згорткових мереж та алгоритмів CTC на користь наскрізної (end-to-end) архітектури типу Encoder-Decoder. Розроблена дослідниками Microsoft, ця модель використовує попередньо навчений візуальний трансформер (ViT або DeiT) у ролі енкодера для вилучення ознак із зображення на рівні патчів (фрагментів), та попередньо навчену мовну модель (RoBERTa) у ролі декодера для генерації тексту. Ця архітектурна синергія дозволяє TrOCR не лише аналізувати візуальну морфологію літер, але й глибоко розуміти семантичний та синтаксичний контекст написаного. Саме завдяки потужному лінгвістичному моделюванню на стороні декодера, TrOCR демонструє безпрецедентно високу точність у розпізнаванні зв'язного рукописного тексту, здатність виправляти візуальні неоднозначності за рахунок контексту та є на сьогодні одним із найперспективніших напрямів для впровадження у складні системи оцифрування історичних або медичних рукописних архівів.

ВИСНОВКИ. Узагальнюючи проведений аналіз сучасних систем аналізу рукописного тексту, можна зробити висновок про швидкий розвиток сучасних технологій з використанням штучного інтелекту для різних задач, зокрема розпізнавання рукописного тексту. Це прискорює та здешевує цифровізацію сучасних підприємств, дозволяючи перейти від паперових документів до цифрових.

ЛІТЕРАТУРА

1. Інформація про PaddleOCR. - Режим доступу: <https://github.com/PADDLEPADDLE/PADDLEOCR> (Дата доступу: 28.04.2026)
2. Інформація про EasyOCR. - Режим доступу: <https://github.com/jaidedai/easyocr> (Дата

доступу: 28.04.2026)

3. Інформація про TrOCR. - Режим доступу: <https://github.com/rsommerfeld/trocr> (Дата доступу: 28.04.2026)

4. Industry 5.0 та колаборативна робототехніка: динамічний опис навколишнього середовища роботів-маніпуляторів з використанням мови Python: монографія / І. Ш. Невлюдов, В. В. Євсєєв, С. С. Максимова. – Харків : Видавництво Іванченка І. С., 2026. – 279 с.

5. Yevsieiev V. Using Historical Data in the NNARX Model to Improve the Accuracy of Microclimate Parameter Forecasting / V. Yevsieiev, I. Holod // Інтелектуальні технології цивільної безпеки та робототехнічні системи аварійно-рятувальних робіт 2026 : матеріали І-ї Всеукраїнської конфер.12-13 лютого 2026 р. - Харків: [електронний друк], 2026. – С. 44-48.

6. Model with Neural Network Component for Adaptive Manipulator Control under Variable Load / Amer Abu-Jassar, Mohammad Hamdan, Nowfal Aweisi, Mahmoud Howaidi, V. Yevsieiev, V. Lyashenko // International Journal of Intelligent Engineering and Systems. –19(1). – 2026. – P. 855-868. <https://doi.org/10.32403/10.22266/ijies2026.0131.51>

7. Nevludov , I. ., Yevsieiev , V. ., Maksymova , S. ., Gopejenko , V. ., & Kosenko , V. . (2025). DEVELOPMENT OF MATHEMATICAL SUPPORT FOR ADAPTIVE CONTROL FOR THE INTELLIGENT GRIPPER OF THE COLLABORATIVE ROBOT MANIPULATOR. Advanced Information Systems, 9(3), 57–65. <https://doi.org/10.20998/2522-9052.2025.3.07>

8. Yevsieiev, V., & Holod, I. (2026). HARDWARE-SOFTWARE MODULE FOR INTELLIGENT MICROCLIMATE CONTROL IN INDUSTRIAL FACILITIES. Scientific Papers of Donetsk National Technical University. Series: “Computer Engineering and Automation”, 4(6(38)), 7–17. [https://doi.org/10.32782/2786-9024/v4i6\(38\).359113](https://doi.org/10.32782/2786-9024/v4i6(38).359113)

9. Using Quantum Computings During Collaborative Mobile Robot Trajectory Constructing / V. V. Yevsieiev, S. S. Maksymova, M. G. Starodubcev, N. P. Demska // Вчені записки ТНУ імені В.І. Вернадського. - Серія: Технічні науки. - 2025. - Т. 36 (75), № 6, частина 2. - P. 111-118. - DOI : <https://doi.org/10.32782/2663-5941/2025.6.2/16>.

10. Development of a Method for Planning the Movement of a Gripping Device for a 3-Link Collaborative Robot Manipulator / I. Nevludov, V. Yevsieiev, S. Maksymova, N. Demska // Journal of natural sciences and technologies. – 2025. - 4(1). - P. 403-413.

11. Maksymova, S., Yevsieiev, V., Chala, O., & Ababneh, J. (2025). DECISION-MAKING MODEL FOR CONTROLLING A COLLABORATIVE ROBOTMANIPULATOR BASED ON THE SENSOR FUSION METHOD AND THE RULES OF RULE-BASED SYSTEMS. Multidisciplinary Journal of Science and Technology, 5(6), 526-538.

12. Nevludov I. Sh. Mathematical Model of Block Process Planning in Systems of Allocation of Task Between People and Collaborative Robots in the Framework of Industries 5.0 / I. Sh. Nevludov, V. V. Yevsieiev, D. V. Gurin // Visnyk of Kherson National Technical University. – 2025. - Vol. 1, № 1(92). - P. 157-163.

13. Development of a model for recognizing various objects and tools in a collaborative robot workspace / V. Yevsieiev, Amer Abu-Jassar, S. Maksymova, N. Demska // Acumen: International Journal of Multidisciplinary Research. – 2025. – Vol. 2(1). – 224-239.

14. Adaptive Regulation of the Manipulator's Movement Speed Depending on the Distance to the Person and the Level of Load on the Actuator / V. Yevsieiev, S. Maksymova, S. Starykova, Jafar Ababneh // The Multidisciplinary Journal of Science and Technoloe. - Volume 5, Issue 9. - P. 9-29.

Науковий керівник: Євсєєв Владислав В'ячеславович, д.т.н., професор, професор кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії (КІТАРБІ), Харківського національного університету радіоелектроніки