УДК 004.415.2
DOI https://doi.org/10.24919/2308-4863/83-1-29

Юліан БАРАНЕЦЬКИЙ,
orcid.org/0009-0008-5545-0865
аспірант кафедри прикладної лінгвістики
Інституту комп'ютерних наук та інформаційних технологій
Національного університету «Львівська політехніка»
(Львів, Україна) yulian.baranetskyy@gmail.com

Наталія КУНАНЕЦЬ,
orcid.org/0000-0003-3007-2462
професор кафедри інформаційних систем та мереж
Інституту комп'ютерних наук та інформаційних технологій
Національного університету «Львівська політехніка»
(Львів, Україна) nek.lviv@gmail.com

РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ: СУЧАСНІ ПІДХОДИ ТА ВИКЛИКИ

Розпізнавання рукописного тексту є важливою складовою сучасних досліджень у сфері опрацювання текстової інформації. Ця задача охоплює широкий спектр застосувань: від оцифрування історичних документів та автоматизації документообігу до збереження культурної спадщини. Особливу актуальність має опрацювання багатомовних і старовинних текстів, що характеризуються складністю почерків, пошкодженнями та незвичними шрифтами. Основною проблемою в розпізнаванні рукописного тексту є його варіативність, яка залежить від стилю письма, індивідуальних особливостей авторів та мовних особливостей.

Метою статті є аналіз сучасних методів розпізнавання рукописного тексту, заснованих на глибокому навчанні, таких як згорткові нейронні мережі, рекурентні нейронні мережі, багатовимірні мережі, а також трансформери. Особлива увага приділена офлайн методам, комбінованим архітектурам і перспективам застосування трансформерів для багатомовного та історичного контекстів.

У статті узагальнено ключові досягнення у використанні глибоких нейронних мереж для розпізнавання тексту. Розглянуто підходи до попереднього опрацювання зображень, сегментації тексту, виділення ознак і класифікації, що значно покращують продуктивність систем розпізнавання. Зокрема, комбіновані моделі, що поєднують згорткову нейронну мережу і рекурентну нейронну мережу, показали високу точність навіть за умов значної варіативності почерку.

Дослідження трансформерів продемонструвало їхній потенціал у врахуванні глобальних залежностей тексту, що є особливо корисним для багатомовних систем. Водночас виявлено, що ці моделі потребують великих обчислювальних ресурсів і значних обсягів даних для ефективного навчання.

Узагальнені результати підкреслюють необхідність подальших досліджень для створення універсальних адаптивних систем, здатних працювати з текстами різними мовами, стилями і поданих різноманіттям форматів. Це сприятиме розвитку технологій цифрового архівування, автоматизації документообігу та забезпечення доступності інформації в сучасному цифровому середовищі.

Ключові слова: *розпізнавання рукописного тексту, багатомовні документи, глибоке навчання, трансформери, історичні документи, механізм уваги, сегментація тексту.*

Yulian BARANETSKYI,
orcid.org/0009-0008-5545-0865
PhD Student Applied Linguistics
Institute of Computer Science and Information Technologies of Lviv Polytechnic National University
(Lviv, Ukraine) yulian.baranetskyy@gmail.com

Natalia KUNANETS,
orcid.org/0000-0003-3007-2462
Professor at the Department of Information Systems and Networks,
Institute of Computer Science and Information Technologies of Lviv Polytechnic National University
(Lviv, Ukraine) nek.lviv@gmail.com

RECOGNITION OF HANDWRITTEN TEXT: MODERN APPROACHES AND CHALLENGES

Handwritten text recognition is a critical component of modern research in the field of text information processing. This task encompasses a wide range of applications, from digitizing historical documents and automating document workflows to preserving cultural heritage. The processing of multilingual and historical texts is particularly relevant, characterized by challenging handwriting styles, damage, and unconventional scripts. The primary issue in handwritten text recognition lies in its variability, influenced by writing styles, individual characteristics of authors, and linguistic features.

The article aims to analyze modern methods of handwritten text recognition based on deep learning techniques, such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN, LSTM), Multidimensional LSTM (MDLSTM), and Transformers. Special attention is given to offline methods, combined architectures (CNN-RNN), and the prospects of applying Transformers in multilingual and historical contexts.

The article summarizes key achievements in the use of deep neural networks for text recognition. It explores approaches to image preprocessing, text segmentation, feature extraction, and classification that significantly enhance the performance of recognition systems. In particular, combined models like CRNN, which integrate CNN and RNN, have demonstrated high accuracy even under conditions of significant handwriting variability.

Research on Transformers has highlighted their potential for capturing global text dependencies, making them particularly useful for multilingual systems. However, these models require substantial computational resources and large datasets for effective training.

The generalized results emphasize the need for further research to develop universal adaptive systems capable of handling texts in various languages, styles, and formats. This would advance the technologies of digital archiving, document workflow automation, and information accessibility in the modern digital environment.

Key words: *Handwritten text recognition, multilingual documents, deep learning, transformers, historical documents, attention mechanism, text segmentation.*

Постановка проблеми. Розпізнавання рукописного тексту є однією з ключових задач у сфері опрацювання текстової інформації. Це важливий напрямок досліджень, який охоплює різні аспекти: збереження історичних рукописів, електронний документообігу, забезпечення доступності інформації в цифрову епоху. Зокрема, розпізнавання рукописного тексту відіграє критичну роль у цифровому архівуванні культурної спадщини, де потрібно опрацьовувати документи різних епох і багатоманіттям мов, що були написані вручну. Це завдання особливо актуальне для історичних документів, багатомовних архівів та текстів, створених на основі різних скриптів.

Однак процес розпізнавання рукописного тексту є досить складним. Він відрізняється від розпізнавання друкованого тексту своєю варіативністю, залежністю від стилю письма, рівня знань автора та його індивідуальних особливостей почерку. Додаткові труднощі виникають при роботі з історичними документами, де текст може бути пошкоджений, вицвілий або написаний незвичними шрифтами. Багатомовність текстів також додає нових викликів, оскільки необхідно розробляти системи, здатні розпізнавати текст різними мовами, враховуючи особливості кожного скрипту.

За останнє десятиліття у цій сфері було досягнуто значного прогресу завдяки використанню методів глибинного навчання, таких як згорткові нейронні мережі, рекурентні нейронні мережі, багатовимірні довготривалі короткочасні пам'яті та трансформери. Ці методи дозволили створити системи, здатні розпізнавати текст із високою точністю навіть у складних умовах.

Аналіз досліджень. Інструменти розпізнавання рукописного тексту поділяється на дві категорії: офлайн та онлайн технології та засоби. Ці підходи суттєво відрізняються один від одного, виходячи з умов отримання тексту та особливостей його опрацювання. Офлайн розпізнавання працює з уже готовими зображеннями тексту, зазвичай отриманими шляхом сканування або фотографування рукописних документів, тоді як онлайн розпізнавання здійснюється в режимі реального часу, захоплюючи динамічну інформацію про рух стилуса або пальця під час написання (Keysers et al., 2016: 1180–1194).

Офлайн розпізнавання є важливим для багатьох застосувань, таких як оцифровування історичних рукописів, автоматизація документообігу в адміністративних установах, збереження культурної спадщини тощо. В офлайн розпізнаванні ключовою проблемою є варіативність стилів почерку, яка впливає на точність роботи систем. Крім того, якість та контрастність зображень можуть суттєво відрізнятися, що додає складності у розпізнаванні тексту (Kozielski et al., 2014).

Сучасні методи офлайн розпізнавання текстів зазвичай поєднують декілька етапів процедури опрацювання: попереднє опрацювання зображення, сегментацію тексту (Kasem et al., 2023) на рівні рядків або символів, виділення ознак (Sahare,

Dhok, 2017) та остаточне розпізнавання. Для розпізнавання символів використовуються різні алгоритми, серед яких найбільш поширеними є глибокі нейронні мережі (Al-Taee et al., 2020), особливо згорткові нейронні мережі (Mishra, 2020) та рекурентні мережі (Hamdani et al., 2014). Офлайн системи, побудовані на основі багатовимірних моделей «Довга короткочасна пам'ять» (Koerich et al., 2003) та комбінованих моделей «Згорткова нейронна мережа – Рекурентна нейронна мережа» (Palehai, Fanany, 2017), продемонстрували високу точність у розпізнаванні тексту різними мовами та скриптами, особливо для таких наборів даних, як Управління ідентифікацією та доступом та Розпізнавання зображень і рукописів для оцінки та систематизації. (Swaileh, Paquet, 2018).

Попереднє опрацювання зображень є ключовим етапом розпізнавання рукописного тексту, оскільки від якості та правильності виконання цього процесу залежить продуктивність і точність подальших результатів. Сюди входять різноманітні підходи та методи для покращення зображення, видалення шумів, нормалізації розміру символів, а також сегментації тексту.

До основних методів попереднього опрацювання належать такі:

Нормалізація зображення, яка забезпечує однорідність вхідних даних. Зображення нормалізуються до певного стандартного розміру та кольорової палітри (зазвичай відтінків сірого). Серед методів нормалізації найчастіше використовуються адаптивне вирівнювання гістограми (Adaptive Histogram Equalization) та алгоритми нормалізації розмірів, які підлаштовують зображення символів під стандартні розміри вхідних даних нейронної мережі.

Видалення шуму, яке передбачає видалення у зображеннях рукописного тексту різних видів шумів, таких як плями, подряпини та інші дефекти, особливо в історичних документах (Shi et al., 2009). Для видалення шуму застосовуються методи фільтрації, зокрема медіанна фільтрація (Median Filtering), яка ефективно видаляє точковий шум, та гаусова фільтрація (Gaussian Filtering), що використовується для згладжування зображень.

Вирівнювання тексту (деслантінг), що передбачає усунення в рукописному тексті нахилу, що ускладнює його розпізнавання, особливо у випадках, коли використовуються методи, чутливі до орієнтації символів. Деслантінг випрямляє символи, що покращує їх розпізнаваність (Scheidl, 2018). Одним із поширених алгоритмів вирівнювання тексту є метод Ланкастера (Lancaster Method), який оцінює нахил символів за допомогою визначення кутів контурів та підлаштовує текст до горизонтальної орієнтації.

Сегментація, яка є одним із ключових процесів, оскільки дозволяє розділити текст на окремі структурні одиниці: рядки, слова або символи (Алексєєв та ін., 2002) (Зарубенко, 2014). Основні підходи до сегментації включають методи горизонтального та вертикального сканування, де проекційні профілі використовуються для виявлення меж рядків та символів (Lecolinet, 1996). Більш складні методи, такі як алгоритм «Sliding Window», виконують сегментацію за допомогою ковзного вікна, яке проходить по всьому зображенню та виділяє фрагменти тексту.

Метою статті є огляд сучасних підходів до розпізнавання рукописного тексту, зокрема методів, заснованих на глибинному навчанні, таких як згорткова нейронна мережа, рекурентна нейронна мережа, довга короткочасна пам'ять та трансформери. Особливу увагу приділено аналізу офлайн методів, комбінованих архітектур, а також перспективам застосування трансформерів для роботи з багатомовними історичними документами. У статті визначено основні виклики, що стоять перед системами розпізнавання, включаючи багатомовність, варіативність стилів письма та обмеженість даних. Завданням роботи є узагальнення досягнень у цій галузі, висвітлення ключових тенденцій та підготовка бази для подальших досліджень у напрямку створення універсальних адаптивних моделей.

Виклад основного матеріалу. У розпізнаванні рукописного тексту використовуються різні підходи, що еволюціонували з розвитком обчислювальної техніки та машинного навчання. Основні методи можна розділити на традиційні підходи та сучасні методи глибинного навчання.

Традиційні підходи включають методи, що грунтуються на прихованих моделях Маркова, статистичних моделях, які використовуються для розпізнавання послідовностей символів або слів на основі ймовірнісних розрахунків. Вони широко застосовувалися в системах розпізнавання рукописного тексту до появи сучасних методів глибинного навчання. Ці моделі базуються на припущенні, що процес генерації тексту можна подати як марківський процес зі скінченною кількістю станів, кожен з яких має певну ймовірність переходу до іншого стану та ймовірність емісії певного символу. У контексті розпізнавання рукописного тексту ці стани можуть полягати в окремих символах, слова або навіть цілих рядках (Marti, Bunke, 2001).

З появою глибинного навчання розпізнавання рукописного тексту зробило значний крок вперед. Нейронні мережі дозволяють моделювати складні залежності та структури рукописних символів, не вимагаючи попередньої сегментації тексту на окремі символи. Основними методами глибинного навчання, які знайшли широке застосування у розпізнаванні рукописного тексту, є згорткові нейронні мережі, рекурентні нейронні мережі, довготривала короткочасна пам'ять та трансформери.

Згорткові нейронні мережі є одним із найпотужніших інструментів для опрацювання зображень. Їхня головна перевага полягає у здатності автоматично виділяти ознаки зображень різних рівнів (наприклад, контури, кути, форми), що дозволяє розпізнавати складні структури, такі як рукописні символи. Вони побудовані з шарів згорток, пулінгу та повнозв'язаних шарів, які використовуються для класифікації зображень (Nayab, Arshad, 2021).

У контексті розпізнавання рукописного тексту згорткові нейронні мережі застосовуються для виділення ознак із вхідного зображення тексту. Наприклад, метод «LeNet-5» (Michalski, Plechawska-Wójcik, 2022) для розпізнавання рукописних цифр, запропонований Лекуном та його командою, грунтується на використанні згорткових шарів для виділення ознак, а потім класифікації символів, яка виконується за допомогою повнозв'язаних шарів (Bezerra et al., 2017). Застосовуючи подібний підхід, сучасні згорткові нейронні мережі дозволяють розпізнавати не лише окремі символи, але й цілі слова або рядки тексту (Szegedy et al., 2014).

Рекурентні нейронні мережі – це тип нейронних мереж, призначений для роботи з послідовними даними, такими як текст. Головною їх особливістю є здатність зберігати інформацію про попередні елементи послідовності та використовувати її для розпізнавання наступних символів. У контексті розпізнавання рукописного тексту це означає, що рекурентні нейронні мережі можуть враховувати контекст символів, що оточують поточний символ, що значно підвищує точність розпізнавання.

У розпізнаванні рукописного тексту найчастіше використовуються двонаправлені рекурентні нейронні мережі (Pham et al., 2013), які опрацьовують послідовність як у прямому, так і в зворотному напрямках. Це дозволяє моделі враховувати контекст символів як зліва, так і справа від поточного символу, що є особливо корисним при розпізнаванні слів і рядків тексту. Цей метод описав Теодор Блюше, запропонувавши підхід «end-to-end», що об'єднує сегментацію рядків, розпізнавання кожного рядка та подальше опрацювання в єдиний процес. Використовуючи нейронні мережі, зокрема рекурентні нейронні мережі та механізм з'єднувальної тимчасової класифікації, модель навчається опрацьовувати цілі параграфи тексту без попередньої сегментації на рядки. Цей підхід показав кращу продуктивність у розпізнаванні рукописних документів у порівнянні з традиційними методами.

Проте стандартні рекурентні нейронні мережі мають обмежену здатність до запам'ятовування тривалих залежностей через проблему зникання градієнта. Для розв'язання цієї проблеми було запропоновано вдосконалений варіант – довготривала короткочасна пам'ять.

Довготривала короткочасна пам'ять є спеціалізованим типом рекурентних нейронних мереж, який розроблено для подолання проблеми зникання градієнта. Довготривала короткочасна пам'ять має спеціальні «комірки пам'яті» та механізми керування потоком інформації (гейти), які дозволяють їм запам'ятовувати інформацію на довготривалих відрізках послідовності. Це робить Довготривала короткочасна пам'ять є ідеальною для задач, де контекст відіграє важливу роль, наприклад, при розпізнаванні тексту (Hassan et al., 2023).

У розпізнаванні рукописного тексту вона використовується для опрацювання послідовності символів або їхніх ознак. Наприклад, в архітектурі згорткової рекурентної нейронної мережі, яка поєднує обидва проаналізовані мережі, згорткові шари виділяють ознаки зі зображення тексту (Ramjon et al., 2023), а довготривала короткочасна пам'ять опрацьовує ці ознаки в послідовності, розпізнаючи цілі слова або рядки. Цей підхід є ефективним для опрацювання текстів із різними стилями письма, включаючи багатомовні історичні документи.

Також популярною модифікацією є двонаправлені нейронні мережі, які опрацьовують послідовність як зліва направо, так і справа наліво. Це дозволяє враховувати контекст символів як попередніх, так і наступних, що підвищує точність розпізнавання.

Комбіновані підходи, такі як поєднання згорткових нейронних мереж та рекурентних нейронних мереж, стали одними з найбільш успішних методів розпізнавання рукописного тексту. Основна ідея комбінованих моделей полягає у використанні переваг кожного з цих типів нейронних мереж: Згорткові нейронні мережі добре виді-

ляють просторові ознаки з зображень, тоді як рекурентні нейронні мережі можуть враховувати послідовності та контекст символів у тексті. Така комбінація дозволяє створити потужні та універсальні системи розпізнавання, здатні працювати з різноманітними наборами даних (Markou et al., 2021).

Комбіновані підходи, які об'єднують обидва види мереж, використовуються для розпізнавання тексту без необхідності попередньої сегментації на окремі символи. У цьому підході згорткова нейронна мережа опрацьовує вхідне зображення, виділяючи високорівневі ознаки, що містять інформацію про форми символів, контури та текстуру. Ці ознаки потім передаються в рекурентні нейронні мережі (наприклад, Довга короткочасна пам'ять або рекурентний блок із гейтовим управлінням), які опрацьовують їх як послідовність, враховуючи залежності між різними частинами тексту. У такий спосіб модель може одночасно враховувати як просторові характеристики символів, так і їхню послідовність, що суттєво покращує точність розпізнавання.

Однією з найпопулярніших архітектур комбінованих моделей є згорткова рекурентна нейронна мережа. Вона складається з трьох основних компонентів: згорткових шарів, рекурентних шарів та механізму класифікації, такого як з'єднувальна тимчасова класифікація.

Згортковий шар є першим компонентом згорткової рекурентної нейронної мережі, який застосовуються для опрацювання вхідного зображення тексту. Ці шари виділяють різноманітні ознаки, що описують локальні структури символів, такі як контури, кути, форми та інші характерні риси. Зазвичай використовуються кілька згорткових шарів із шарами підвибірки, які зменшують розмірність вхідного зображення та виділяють найбільш релевантні ознаки. Результатом роботи згорткової нейронної мережі є матриця ознак, яка містить інформацію про кожен фрагмент вхідного зображення.

Матриця ознак із згорткових шарів передається до рекурентних шарів, які опрацьовують її як послідовність. Найчастіше в комбінованих моделях використовуються довгу короткочасну пам'ять або двонаправлену довгу короткочасну пам'ять. Ці шари враховують залежності між ознаками, моделюючи послідовність символів та слів у тексті. Наприклад, якщо певна частина зображення відповідає літері «Т», рекурентний шар опрацьовує її, враховуючи контекст сусідніх фрагментів, щоб вирішити, чи належить ця літера до слова «THE» або «TREE». Двонаправлена довга короткочасна пам'ять дозволяє опрацьовувати послідовність як у прямому, так і зворотному напрямках, що є особливо корисним для врахування контексту з обох боків (Bluche, 2017).

Останнім компонентом згорткової рекурентної нейронної мережі є механізм класифікації – з'єднувальна тимчасова класифікація, який дозволяє моделі визначати послідовність символів, не вимагаючи попередньої сегментації тексту. Він використовує ймовірнісний підхід для зіставлення виділених ознак із символами. Це особливо корисно при розпізнаванні рукописного тексту, оскільки дозволяє моделі працювати з різною довжиною слів та символів. Завдяки цьому механізму модель може вирішити, які фрагменти ознак відповідають певним символам, та обрати найбільш ймовірну послідовність розпізнавання (Swaileh et al., 2017).

Трансформери – це один із найсучасніших та найпотужніших підходів у галузі опрацювання природної мови та розпізнавання рукописного тексту. Вперше запропоновані Васвані та його колегами у 2017 році, трансформери стали основою багатьох передових систем розпізнавання завдяки своїй здатності ефективно опрацьовувати послідовні дані та враховувати глобальні залежності у тексті (He et al., 2016) (Porwal et al., 2022). Головною особливістю трансформерів є механізм уваги (Attention Mechanism), який дозволяє моделі зосереджуватися на найбільш важливих частинах вхідної послідовності при опрацюванні кожного символу або слова.

Архітектура трансформера складається з двох основних компонентів: енкодера та декодера. Кожен із цих компонентів містить кілька шарів самоуваги та повнозв'язаних шарів, що дозволяє моделі ефективно опрацьовувати послідовності будь-якої довжини.

Енкодер приймає вхідні дані (наприклад, зображення тексту або послідовність символів) та перетворює їх у високорівневі ознаки. Він складається з кількох шарів, кожен із яких містить самоувагу та нормалізацію. Самоувага дозволяє моделі визначати, які частини вхідної послідовності є найбільш важливими для поточного символу, враховуючи глобальний контекст. Це дає можливість енкодеру вивчати складні залежності між символами та словами у всьому тексті.

Декодер відповідає за генерування вихідної послідовності (наприклад, розпізнаного тексту). Він приймає високоуровневі ознаки, які виходять з енкодера, та застосовує до них багатошаровий механізм уваги. На відміну від енкодера, декодер має додатковий механізм уваги, що працює з вхід-

ними ознаками з енкодера, дозволяючи враховувати як локальний, так і глобальний контекст під час генерації кожного символу.

Ключовим компонентом трансформерів є механізм уваги. На відміну від рекурентних мереж, які опрацьовують послідовності послідовно (крок за кроком), трансформери дозволяють моделі «бачити» всю послідовність відразу. Це дозволяє опрацьовувати дані паралельно, суттєво підвищуючи швидкість опрацювання великих обсягів тексту. Механізм уваги дозволяє кожному символу або слову в тексті впливати на опрацювання інших символів, оцінюючи вагу кожного елемента послідовності.

Математично, механізм уваги працює шляхом обчислення ваги важливості для кожного елемента послідовності. Ці ваги визначаються шляхом зіставлення вхідних ознак, що дозволяє кожному символу взаємодіяти з усіма іншими символами в тексті. У результаті модель може сфокусуватися на тих частинах послідовності, які найбільш релевантні для поточного завдання.

Оскільки трансформери не мають внутрішньої послідовності, як це властиво рекурентним мережам, для врахування порядку символів використовується метод позиційного кодування. У цьому методі до кожного символу або ознаки додаються значення, які вказують на їхню позицію в послідовності.

Завдяки своїй здатності опрацьовувати складні залежності між символами та словами, трансформери стали одним із найефективніших методів розпізнавання рукописного тексту. Вони дозволяють враховувати як локальні (форму символів), так і глобальні (контекст слова або фрази) залежності, що особливо корисно для багатомовного розпізнавання та обробки текстів із неоднорідною структурою.

Одним із прикладів успішного застосування трансформерів у розпізнаванні рукописного тексту є модель розпізнавання рукописного тексту за допомогою трансформерів, яка була розроблена для розпізнавання арабського рукописного тексту. Вона використовує багатошаровий механізм уваги для обробки послідовності символів та генерує вихідний текст, враховуючи глобальні контексти. Цей підхід дозволяє моделі адаптуватися до різних стилів письма та враховувати контекст, підвищуючи точність розпізнавання (Chan et al., 2024).

Інший підхід включає поєднання трансформерів із функцією втрат з'єднувальної тимчасової класифікації, яка дозволяє моделі зіставляти вхідні ознаки з вихідними символами без попередньої сегментації тексту. У цьому випадку трансформери застосовуються для виділення ознак та моделювання послідовності, а з'єднувальної тимчасової класифікації використовується для визначення оптимального шляху розпізнавання. Цей підхід використовується для обробки як рукописного, так і друкованого тексту, досягаючи високої точності у великих наборах даних (Carbune et al., 2019).

Трансформери мають ряд переваг перед іншими моделями глибокого навчання. По-перше, вони можуть враховувати глобальні залежності в тексті, що дозволяє їм розпізнавати складні структури та патерни, характерні для різних мов та стилів письма. По-друге, завдяки паралельній обробці послідовностей, трансформери працюють з великими текстами швидше, ніж рекурентна нейронна мережа або довга короткочасна пам'ять.

Однак трансформери мають і деякі обмеження. Найголовніше з них – це їхня потреба у великій кількості обчислювальних ресурсів та даних для ефективного навчання. Крім того, хоча трансформери можуть враховувати глобальні залежності, іноді вони недостатньо ефективні для опрацювання дрібних деталей символів, що характерно для рукописного тексту з високою варіативністю.

Висновки. Розпізнавання рукописного тексту залишається однією з ключових задач у сфері опрацювання текстової інформації, яка поєднує виклики варіативності почерків, багатомовності та обмеженості доступних даних. Сучасні методи глибинного навчання, такі як згорткові нейронні мережі, рекурентні мережі та трансформери, продемонстрували значний прогрес у вирішенні цих проблем. Зокрема, комбіновані архітектури згорткова нейронна мережа – рекурентна нейронна мережа та згорткова рекурентна нейронна мережа, а також трансформери із механізмом з'єднувальної тимчасової класифікації забезпечують високу точність навіть для складних багатомовних і історичних текстів.

Попри це, існують невирішені виклики, пов'язані із залежністю моделей від великих обсягів даних і обчислювальних ресурсів, а також з потребою у вдосконаленні опрацювання дрібних деталей символів. Для подальшого розвитку цієї галузі важливими є дослідження універсальних адаптивних моделей, здатних враховувати локальні та глобальні контексти тексту, а також інтеграція сучасних методів збереження і оцифрування культурної спадщини.

Таким чином, розпізнавання рукописного тексту є перспективним напрямком, що має значний потенціал для розв'язання завдань цифрового архівування та забезпечення доступності інформації у сучасному світі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Multi-Language Online Handwriting Recognition / Keysers D. та ін. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2016. Vol. 39, No. 6. P. 1180–1194.
2. Multilingual Off-line Handwriting Recognition in Real-world Images / Kozielski M., Doetsch P., Hamdani M., Ney H. *14th International Conference on Frontiers in Handwriting Recognition (ICFHR).* 2014. P. 121–125, doi: 10.1109/DAS.2014.8.
3. Kasem M. S., Mahmoud M., Kang H. S. Advancements and Challenges in Arabic Optical Character Recognition: A Comprehensive Survey *arXiv preprint*. 2023. doi:arXiv:2312.11812.
4. Sahare P., Dhok S. B. Script Identification Algorithms: A Survey *International Journal of Multimedia Information Retrieval*. 2017. Vol. 6, No. 3. P. 211–232.
5. Al-Taee M. M., Neji S. B. H., Frikha M. Handwritten Recognition: A Survey *Proceedings of the 2020 IEEE 4th International Conference*. 2020. P. 199–205. doi: 10.1109/IPAS50080.2020.9334936.
6. Mishra M. Convolutional Neural Networks, Explained *Towards Data Science*. URL: https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939 (дата звернення: 28.12.2024).
7. The RWTH Large Vocabulary Arabic Handwriting Recognition System / Hamdani M. and others. *Proceedings – 11th IAPR International Workshop on Document Analysis Systems,* DAS 2014. doi: 10.1109/DAS.2014.61.
8. Koerich A. L., Sabourin R., Suen C. Y. Large Vocabulary Off-line Handwriting Recognition: A Survey *Pattern Analysis & Applications*. 2003. Vol. 6. P. 97–121.
9. Palehai D., Fanany M. I. Handwriting Recognition on Form Document Using Convolutional Neural Network and Support Vector Machines (CNN-SVM) *5th International Conference on Information and Communication Technology (ICoIC7)*, Melaka, Malaysia, 2017, pp. 1–6, doi: 10.1109/ICoICT.2017.8074699.
10. Swaileh W., Paquet T. A Syllable Based Model for Handwriting Recognition *arXiv preprint.* 2018. doi: *arXiv*. 1808.07277.
11. Shi Z., Setlur S., Govindaraju V. Text Extraction from Gray Scale Historical Document Images Using Adaptive Local Connectivity Map *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR)*. 2009. P. 1375–1379.
12. Scheidl H. Handwritten Text Recognition in Historical Documents. [Diploma Thesis. Technische Universität Wien]. 2018. 80 p. reposiTUm. doi: 10.34726/hss.2018.43931
13. Алексєєв А., Заяць В., Іванов Д. Алгоритм розпізнавання символів на основі структурного підходу. *Вісник Нац. ун-ту «Львівська політехніка» «Комп'ютерна інженерія та інформаційні технології»*.– 2002. – № 468. – С. 129–133.
14. Зарубенко О. О. Побудова алгоритму аналізу рукописного тексту *Технологічний аудит та резерви виробництва*. 2014. № 6(4). С. 28–31.
15. Lecolinet E. A Survey of Methods and Strategies in Character Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July 1996, 1996. vol. 18, no. 7. pp. 690–706. doi: 10.1109/34.506792.
16. Marti U.-V., Bunke H. Using a Statistical Language Model to Improve the Performance of an HMM-Based Cursive Handwriting Recognition System, *International Journal of Pattern Recognition and Artificial Intelligence*. 2001. Vol. 15, No. 1. P. 65–90.
17. Nayab A., Arshad A. A Survey of Image Classification Methods. *ILMA Journal of Technology & Software Management*. 2021. Volume 1 Issue. Pp. 115–24.
18. Michalski B., Plechawska-Wójcik M. Comparison of LeNet-5, AlexNet and GoogLeNet Models in Handwriting Recognition. *Journal of Computer Sciences Institute*. 2022. Vol. 23, Pp. 145–151. doi: 10.35784/jcsi.2919
19. Handwriting Recognition, Development, and Analysis. / Bezerra B. L. D., Zanchettin C., Toselli A. H., Pirlo G. (Eds.). Hauppauge: Nova Science Publishers, 2017. 395 p.
20. Going deeper with convolutions. / Szegedy C. and others. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2014. pp. 1–9. doi: 10.1109/CVPR.2015.7298594.
21. Pham V., Kermorvant C., Louradour J. Dropout Improves Recurrent Neural Networks for Handwriting Recognition *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR).* 2014. Pp. 285–290. doi: 10.1109/ICFHR.2014.55
22. Hassan N., Miah A. S. M., Shin J. A Deep Bidirectional LSTM Model Enhanced by Transfer-Learning-Based Feature Extraction for Dynamic Human Activity Recognition *Applied Sciences.* 2024, Vol. 14, Pp. 603–605. doi: 10.3390/app14020603
23. Ramjon A., Bwatiramba A., Venkataraman S. Handwritten Character Recognition Using Deep Learning (Convolutional Neural Network) *Computer Engineering and Intelligent Systems*. 2023. Vol. 14, No. 1. Pp. 34–54. doi:10.7176/CEIS/14-1-05.
24. A Convolutional Recurrent Neural Network for the Handwritten Text Recognition of Historical Greek Manuscripts / Markou K. and others. *Proceedings of the 17th International Conference on Document Analysis and Recognition (ICDAR).* 2021. doi: 10.1007/978-3-030-68787-8_18.
25. Bluche T. Joint Line Segmentation and Transcription for End-to-End Handwritten Paragraph Recognition *Advances in Neural Information Processing Systems (NeurIPS).* arXiv:1604.08352. 2017. doi: 10.48550/arXiv.1604.08352
26. Handwriting Recognition with Multigrams / Swaileh W., Paquet T., Soullard Y., Tranouez P. *Proceedings of the 14th International Conference on Document Analysis and Recognition (ICDAR)*. Nov. 9–15, 2017. Kyoto, Japan 137–142 doi: 10.1109/ICDAR.2017.31

27. Deep Residual Learning for Image Recognition / He, Kaiming, X. Zhang, Shaoqing Ren and Jian Sun *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, Las Vegas, 2016. P. 770–778, doi: 10.1109/CVPR.2016.90.

28. Porwal U., Fornés A., Shafait F. Advances in Handwriting Recognition *International Journal on Document Analysis and Recognition (IJDAR)*. 2022. Vol. 25, Pp. 241–243. https://doi.org/10.1007/s10032-022-00421-8

29. HATFormer: Historic Handwritten Arabic Text Recognition with Transformers. / Chan A. and others. *ArXiv, abs/2410.02179*.2024. doi: 10.48550/arXiv.2410.02179

30. Fast Multi-language LSTM-based Online Handwriting Recognition / Carbune V. and others. *International Journal of Document Analysis and Recognition* (IJDAR). 2020. Vol. 23. P.89–102. doi:10.1007/s10032-020-00350-42019.

REFERENCES

1. Keysers, D., Deselaers, T., Rowley, H., Wang, L., Carbune, V. (2016) Multi-Language Online Handwriting Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 6. P. 1180–1194.

2. Kozielski, M., Doetsch, P., Hamdani, M., Ney, H. (2014) Multilingual Off-line Handwriting Recognition in Real-world Images, 14th International Conference on Frontiers in Handwriting Recognition (ICFHR), P. 121–125, doi: 10.1109/DAS.2014.8.

3. Kasem, M. S., Mahmoud, M., Kang, H. S. (2023) Advancements and Challenges in Arabic Optical Character Recognition: A Comprehensive Survey, arXiv preprint, doi:arXiv:2312.11812.

4. Sahare, P., Dhok, S. B. (2017) Script Identification Algorithms: A Survey, International Journal of Multimedia Information Retrieval, Vol. 6, No. 3. P. 211–232.

5. Al-Taee, M. M., Neji, S. B. H., Frikha, M. (2020) Handwritten Recognition: A Survey, Proceedings of the 2020 IEEE 4th International Conference, P. 199–205, doi:10.1109/IPAS50080.2020.9334936.

6. Mishra, M. (2024) Convolutional Neural Networks, Explained, Towards Data Science, URL: https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939

7. Hamdani M., Doetsch P., Kozielski M., El-Desoky Mousa A., Ney H. (2014) The RWTH Large Vocabulary Arabic Handwriting Recognition System, Proceedings – 11th IAPR International Workshop on Document Analysis Systems (DAS), doi: 10.1109/DAS.2014.61.

8. Koerich, A. L., Sabourin, R., Suen, C. Y. (2003) Large Vocabulary Off-line Handwriting Recognition: A Survey. Pattern Analysis & Applications, Vol. 6. P. 97–121.

9. Palehai, D., Fanany, M. I. (2017) Handwriting Recognition on Form Document Using Convolutional Neural Network and Support Vector Machines (CNN-SVM), 5th International Conference on Information and Communication Technology (ICoIC7), pp. 1–6, doi: 10.1109/ICoICT.2017.8074699.

10. Swaileh, W., Paquet, T. (2018) A Syllable-Based Model for Handwriting Recognition, arXiv preprint, doi: arXiv.1808.07277.

11. Shi, Z., Setlur, S., Govindaraju, V. (2009) Text Extraction from Gray Scale Historical Document Images Using Adaptive Local Connectivity Map, Proceedings of the International Conference on Document Analysis and Recognition (ICDAR), P. 1375–1379.

12. Scheidl, H. (2018) Handwritten Text Recognition in Historical Documents, Technische Universität Wien, 80 p., reposiTUm, doi: 10.34726/hss.2018.43931

13. Aleksieiev A., Zaiats V., Ivanov D. (2002) Alhorytm rozpiznavannia symvoliv na osnovi strukturnoho pidkhodu. [Algorithm for Character Recognition Based on the Structural Approach] Visnyk Nats. un-tu „Lvivska politekhnika" „Kompiuterna inzheneriia ta informatsiini tekhnolohii". № 468, 129–133. [in Ukrainian].

14. Zarubenko, O. O. (2014) Pobudova alhorytmu analizu rukopysnoho tekstu [Development of an Algorithm for Handwritten Text Analysis] Tekhnolohichnyi audyt ta rezervy vyrobnytstva. № 6(4). 28–31. [in Ukrainian].

15. Lecolinet, E. (1996) A Survey of Methods and Strategies in Character Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 7, pp. 690–706, doi: 10.1109/34.506792.

16. Marti, U.-V., Bunke, H. (2001) Using a Statistical Language Model to Improve the Performance of an HMM-Based Cursive Handwriting Recognition System, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 15, No. 1, P. 65–90.

17. Nayab, A., Arshad, A. (2021) A Survey of Image Classification Methods, ILMA Journal of Technology & Software Management, Vol. 1, Issue. Pp. 115–24.

18. Michalski, B., Plechawska-Wójcik, M. (2022) Comparison of LeNet-5, AlexNet and GoogLeNet Models in Handwriting Recognition, Journal of Computer Sciences Institute, Vol. 23, Pp.145–151, doi:10.35784/jcsi.2919

19. Bezerra, B. L. D., Zanchettin, C., Toselli, A. H., Pirlo, G. (Eds.). (2017) Handwriting Recognition, Development, and Analysis, Nova Science Publishers, 395 p.

20. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. (2014) Going Deeper with Convolutions, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

21. Pham, V., Kermorvant, C., Louradour, J. (2014) Dropout Improves Recurrent Neural Networks for Handwriting Recognition, Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR), Pp. 285–290, doi: 10.1109/ICFHR.2014.55

22. Hassan N., Miah A. S. M., Shin J. A (2024) Deep Bidirectional LSTM Model Enhanced by Transfer-Learning-Based Feature Extraction for Dynamic Human Activity Recognition Applied Sciences, Vol. 14, Pp. 603–605, doi: 10.3390/app14020603

23. Ramjon A., Bwatiramba A., Venkataraman S. (2023) Handwritten Character Recognition Using Deep Learning (Convolutional Neural Network), Computer Engineering and Intelligent Systems, Vol. 14, No. 1, Pp. 34–54, doi: 10.7176/CEIS/14-1-05.

24. Markou K., Tsochatzidis L., Zagoris K., Papazoglou A., Karagiannis X., Symeonidis S., Pratikakis I. (2021) A Convolutional Recurrent Neural Network for the Handwritten Text Recognition of Historical Greek Manuscripts, Proceedings of the 17th International Conference on Document Analysis and Recognition (ICDAR), doi: 10.1007/978-3-030-68787-8_18.

25. Bluche T. (2017) Joint Line Segmentation and Transcription for End-to-End Handwritten Paragraph Recognition, Advances in Neural Information Processing Systems (NeurIPS), arXiv:1604.08352, doi: 10.48550/arXiv.1604.08352

26. Swaileh W., Paquet T., Soullard Y., Tranouez P. (2017) Handwriting Recognition with Multigrams, Proceedings of the 14th International Conference on Document Analysis and Recognition (ICDAR), 137–142, doi: 10.1109/ICDAR.2017.31

27. He, Kaiming, X. Zhang, Shaoqing Ren, Jian Sun (2016) Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), P. 770–778, doi: 10.1109/CVPR.2016.90

28. Porwal U., Fornés A., Shafait F. (2022) Advances in Handwriting Recognition, International Journal on Document Analysis and Recognition (IJDAR), Vol. 25, Pp. 241–243, https://doi.org/10.1007/s10032-022-00421-8

29. Chan A., Mijar A., Saeed M., Wong C.-W., Khater A. (2024) HATFormer: Historic Handwritten Arabic Text Recognition with Transformers, ArXiv, abs/2410.02179.2024, doi: 10.48550/arXiv.2410.02179

30. Carbune V., Gonnet P., Deselaers T., Rowley H. A., Daryin A., Calvo M., Wang L.-L., Keysers D., Feuz S., Gervais P. (2019) Fast Multi-language LSTM-based Online Handwriting Recognition, International Journal of Document Analysis and Recognition (IJDAR), Vol. 23, P. 89–102, doi:10.1007/s10032-020-00350-42019