УДК 004.93'12:004.852
DOI https://doi.org/10.15589/znp2025.2(500).33

ANALYSIS OF EXISTING METHODS FOR HANDWRITTEN TEXT RECOGNITION

АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ

Tetiana A. Vakaliuk
tetianavakaliuk@gmail.com
ORCID: 0000-0001-6825-4697

Denys V. Furikhata
fyrihata.denus@gmail.com
ORCID: 0000-0002-6093-664X

Valentyn M. Yanchuk
v.yanchuk@gmail.com
ORCID: 0000-0002-6715-4667

Mariia O. Medvedieva
medvedeva-masha25@ukr.net
ORCID: 0000-0001-9330-5185

Nazarii V. Leut
leut.naz.2001@gmail.com
ORCID: 0009-0009-4317-0898

Т. А. Вакалюк[1],
докт. пед. наук, професор

Д. В. Фуріхата[1],
старший викладач

В. М. Янчук[1],
канд. техн. наук, доцент

М. О. Медведєва[2],
канд. пед. наук, доцент

Н. В. Леут[1],
студент

[1]Zhytomyr Polytechnic State University, Zhytomyr
[2]Pavlo Tychyna Uman State Pedagogical University, Uman
[1]Державний університет «Житомирська політехніка», м. Житомир
[2]Уманський державний педагогічний університет імені Павла Тичини, м. Умань

Abstract. The study *aims* to comprehensively analyse existing methods and technologies for handwritten text recognition to justify the optimal approach to developing an effective recognition system for mobile applications. *Methodology.* The work uses a systematic literature review method to analyse scientific publications and technical documentation in handwritten text recognition. A comparative analysis was used to evaluate the advantages and disadvantages of different approaches, from traditional methods (pattern recognition, feature extraction) to modern solutions based on machine learning (k-NN, SVM) and deep neural networks (CNN, RNN). An analytical method was used to justify the optimal solution's choice theoretically, and a synthetic method was used to form a hybrid approach. *Results.* A detailed analysis of eight categories of handwritten text recognition methods was conducted. It was found that traditional methods have limited effectiveness due to their low adaptability to variations in handwriting. Machine learning algorithms show better results but depend on the quality of pre-processing. Deep learning methods demonstrate the highest efficiency due to automatic feature extraction and modelling of complex dependencies. A combined CNN + RNN + CTC approach is justified as the optimal solution for handwritten text recognition. *Scientific novelty.* A comprehensive systematisation of handwritten text recognition methods has been carried out, considering the specifics of mobile platforms. The advantages of a hybrid architecture combining convolutional and recurrent neural networks with the CTC function for recognition tasks without prior segmentation have been scientifically substantiated. A methodological approach to selecting optimal algorithms based on criteria of accuracy, speed, and adaptability to mobile constraints has been developed. *Practical significance.* The research results provide a theoretical basis for developing practical mobile applications for handwritten text recognition. The proposed approach ensures high recognition accuracy, adaptability to different handwriting styles, and optimisation for mobile devices. Developers can use the systematisation of methods to make informed decisions about technology selection. The research opens up prospects for creating Ukrainian-language recognition systems and further developing the field of digital processing of handwritten documents.
Key words: handwritten text recognition; machine learning; convolutional neural networks; recurrent neural networks; deep learning.

Анотація. Метою дослідження є комплексний аналіз існуючих методів та технологій розпізнавання рукописного тексту для обґрунтування оптимального підходу до розробки ефективної системи розпізнавання, орієнтованої на реалізацію в мобільних додатках. *Методика.* У роботі використано метод систематичного огляду літератури для аналізу наукових публікацій та технічної документації в галузі розпізнавання рукописного тексту. Застосовано порівняльний аналіз для оцінки переваг та недоліків різних підходів: від традиційних методів (шаблонне розпізнавання, виділення ознак) до сучасних рішень на основі машинного навчання (k-NN, SVM) та глибоких нейронних мереж (CNN, RNN). Використано аналітичний метод для теоретичного обґрунтування вибору оптимального рішення та синтетичний метод для формування гібридного підходу. *Результати.* Проведено детальний аналіз восьми основних категорій методів розпізнавання рукописного тексту. Виявлено, що традиційні методи мають обмежену ефективність через низьку адаптивність до варіацій рукопису. Алгоритми машинного навчання показують кращі результати, але залежать від якості попередньої обробки даних. Методи глибокого навчання демонструють найвищу ефективність завдяки автоматичному виділенню ознак та моделюванню складних залежностей. Обґрунтовано вибір комбінованого підходу CNN + RNN + CTC як оптимального рішення для розпізнавання рукописного тексту. *Наукова новизна.* Здійснено комплексну систематизацію методів розпізнавання рукописного тексту з урахуванням специфіки мобільних платформ. Науково обґрунтовано переваги гібридної архітектури, що поєднує згорткові та рекурентні нейронні мережі з функцією CTC для задач розпізнавання без попередньої сегментації. Розроблено методологічний підхід до вибору оптимальних алгоритмів на основі критеріїв точності, швидкодії та адаптивності до мобільних обмежень. *Практична значимість.* Результати дослідження створюють теоретичну основу для розробки ефективних мобільних додатків розпізнавання рукописного тексту. Запропонований підхід забезпечує високу точність розпізнавання, адаптивність до різних стилів рукопису та оптимізацію для мобільних пристроїв. Систематизація методів може використовуватися розробниками для прийняття обґрунтованих рішень щодо вибору технології. Дослідження відкриває перспективи для створення українськомовних систем розпізнавання та подальшого розвитку галузі цифрової обробки рукописних документів.

Ключові слова: розпізнавання рукописного тексту; машинне навчання; згорткові нейронні мережі; рекурентні нейронні мережі; глибоке навчання.

ПОСТАНОВКА ЗАДАЧІ

Поширеність рукописного тексту, який незважаючи на тенденції діджиталізації, залишається одним із найпростіших способів фіксації інформації, крім того чимало інформації вже зберігається у рукописному форматі. Розробка відповідного алгоритму та додатку націлена на вирішення задачі перетворення користувачами власних рукописів в електронний формат для подальшої роботи з цифровими даними. Популярність мобільних пристроїв, які є невід'ємною частиною повсякденного життя людей вказує на актуальність цієї платформи для реалізації застосунку розпізнавання рукописного тексту. Розробка необхідного алгоритму включає в себе використання штучного інтелекту та машинного навчання. Аналіз та покращення існуючих алгоритмів допоможе підвищити рівень точності розпізнавання та оптимізувати процес. Зростання конкуренції на ринку мобільних додатків стимулює розробників до пошуку інноваційних рішень.

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Неіл Т. у своїй роботі [1] розглядає ключові принципи проектування інтерфейсів для мобільних додатків. Автор детально аналізує найбільш ефективні шаблони дизайну, які забезпечують зручність використання та інтуїтивність взаємодії користувача з додатком. Хеллман Е. досліджує передові методи програмування для платформи Android [2]. Автор акцентує увагу на оптимізації продуктивності додатків та ефективному використанні ресурсів мобільного пристрою.

Колектив авторів [6] представляють систематизацію моделей представлення знань. Хоча робота не безпосередньо стосується розпізнавання тексту, вона надає важливі теоретичні основи для розуміння методів структурування та обробки інформації. Крамер О. [7] надає фундаментальний аналіз методу k-найближчих сусідів. Автор детально розглядає принципи роботи алгоритму, його переваги у простоті реалізації та недоліки у вигляді високих обчислювальних витрат.

Саркер І. [9] систематизує сучасні досягнення в галузі глибокого навчання. Автор детально аналізує різні архітектури нейронних мереж, їх застосування та перспективи розвитку. Дослідження є особливо цінним для розуміння тенденцій та майбутніх напрямків розвитку технологій розпізнавання. Небаур Ц. [10] одним з перших досліджує застосування згорткових нейронних мереж для задач комп'ютерного зору. Автор демонструє ефективність CNN у розпізнаванні візуальних образів та закладає теоретичні основи для подальшого розвитку цієї технології.

Ще одна група авторів [11] надають сучасний погляд на архітектуру та принципи роботи згорткових мереж. Автори детально аналізують структуру CNN, включаючи згорткові шари, шари об'єднання та повнозв'язні шари. Інша група авторів [12]

систематизують сучасні знання про рекурентні нейронні мережі. Автори детально розглядають різні модифікації RNN, включаючи LSTM та GRU, їх архітектурні особливості та області застосування.

ВІДОКРЕМЛЕННЯ НЕВИРІШЕНИХ РАНІШЕ ЧАСТИН ЗАГАЛЬНОЇ ПРОБЛЕМИ

В сучасному світі мобільні додатки визначають стандарти зручності та ефективності в повсякденному житті. Зокрема, розробка додатків для розпізнавання рукописного тексту стає невід'ємною ланкою цього невпинного технологічного розвитку. Незважаючи на прогресуючу цифровізацію, рукописний текст залишається популярним методом фіксації інформації, а відтак, виникає потреба в зручних та ефективних інструментах для його розпізнавання та обробки.

Метою даного дослідження є аналіз існуючих методів та технологій розпізнавання рукописного тексту.

МЕТОДИ, ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

У даній роботі було використано декілька ключових методів дослідження для проведення аналізу існуючих підходів до розпізнавання рукописного тексту. Метод систематичного огляду літератури став основним інструментом дослідження. Було проведено комплексний аналіз наукових публікацій у напрямку розпізнавання рукописного тексту, що дозволило систематизувати наявні методи. Порівняльний аналіз використовувався для оцінки переваг та недоліків різних методів розпізнавання.

Об'єктом дослідження є розпізнавання рукописного тексту.

Предметом дослідження є методи попередньої обробки зображень, алгоритми машинного навчання для розпізнавання рукописного тексту та методи оцінки якості роботи системи.

ОСНОВНИЙ МАТЕРІАЛ

Мобільні застосунки для розпізнавання рукописного тексту вирішують проблеми користувачів, дозволяючи легко конвертувати нотатки, записи або рукописні документи у цифровий формат. Це надає можливість подальшого редагування, швидкого пошуку та зручного зберігання важливої інформації. Додатки відзначаються високою ефективністю, навіть у випадках нечіткого написання, і надають користувачам різні опції перетворення рукописного тексту, такі як переклад чи конвертація в інші формати. Забезпечення зрозумілого та дружнього інтерфейсу робить використання додатків максимально зручним та доступним [1].

В основі роботи таких додатків лежать потужні алгоритми розпізнавання символів та слів. Деякі з них можуть навіть адаптуватися до стилю рукопису конкретного користувача, стаючи все точнішими у розпізнаванні тексту з плином часу. Це підвищує якість та швидкість роботи додатків, роблячи їх необхідним інструментом для ефективної обробки рукописної інформації в цифровому форматі [2].

Найпоширенішими моделями розпізнавання тексту є OCR (Optical Character Recognition – оптичне розпізнавання символів) та HTR (Handwritten Text Recognition – розпізнавання рукописного тексту). Відзначаючись високою точністю, OCR вирішує завдання розпізнавання друкованих символів без надмірних обчислень. З іншого боку, HTR вимагає застосування штучних нейронних мереж та машинного навчання, відкриваючи нові горизонти для розпізнавання рукописного тексту.

Недавні прориви в галузі штучних нейронних мереж (ШНМ) суттєво трансформують парадигму розпізнавання рукописного тексту. У порівнянні з традиційними методами, ШНМ автоматично навчаються на прикладах, що дозволяє їм адаптуватися до різноманітних варіацій рукопису. Зокрема, згорткові нейронні мережі стали ключовим інструментом для задач розпізнавання зображень, включаючи рукописний текст.

Згорткові нейронні мережі (ЗНМ) володіють здатністю автоматично виділяти важливі ознаки з вхідних зображень, використовуючи фільтри, які адаптуються під час навчання (див. рис. 1). Ці мережі структуровані так, що вони здатні розпізнавати ієрархічні ознаки на різних рівнях абстракції. Застосовані до розпізнавання рукописного тексту, ЗНМ здатні ефективно впоратися з різноманітністю стилів написання та вирізняються високою точністю. За останні роки технологія згорткових нейронних мереж в суттєвий спосіб розвинулася, надаючи практичне застосування в різних галузях, включаючи розпізнавання рукописного тексту.

Головною метою є реалізація власного підходу до розпізнавання рукописного тексту та демонстрація його роботи в розробленому мобільному застосунку. Детально оглянувши основні принципи найбільш ефективних існуючих систем, поставимо перед собою такі задачі, як аналіз аналогічних рішень, виділення сильних і слабких сторін алгоритмів, що використовуються для HTR та на основі отриманих висновків проєктування та реалізація алгоритму розпізнавання, розробка клієнтської частині у форматі мобільного застосунку та імплементація алгоритму в нього.

Розпізнавання рукописного тексту є складною задачею, що включає декілька етапів [3], таких як сегментація, вилучення ознак, класифікація та постобробка. На кожному з цих етапів використовуються різноманітні методи та алгоритми, які спрямовані на досягнення високої точності розпізнавання. Розглянемо основні підходи до розпізнавання рукописного тексту, зокрема традиційні методи, методи на основі машинного навчання та глибокого навчання, принцип їх дії, переваги та недоліки кожного з них.

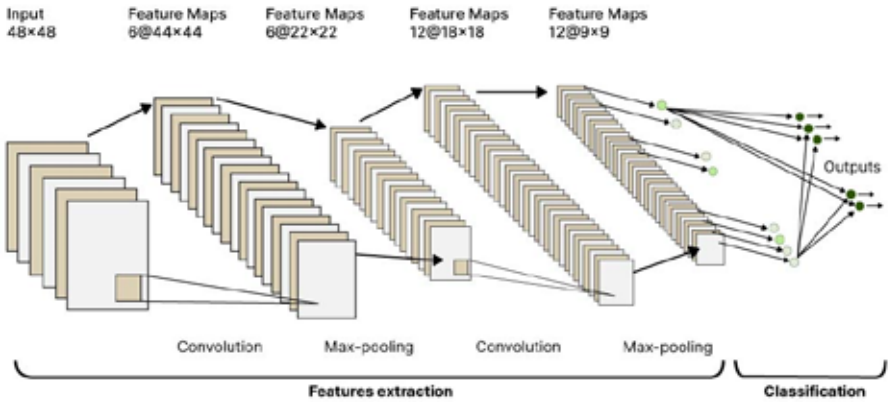

Рис. 1. Структура згорткової мережі

Шаблонне розпізнавання [4] (Template Matching) є одним з найпростіших і базових методів для ідентифікації рукописного тексту. Його принцип роботи полягає в тому, що кожен символ на зображенні порівнюється зі зразками – шаблонами, які заздалегідь збережені в базі даних. При цьому здійснюється пошук найбільш схожого зразка на основі певного методу порівняння, наприклад, кореляції пікселів або мінімізації відстані між зображеннями. Цей підхід широко застосовується для задач, де набір символів обмежений або текст має чітко визначені характеристики, як, наприклад, друковані шрифти.

Головна перевага цього методу полягає у простоті реалізації. Оскільки алгоритм не вимагає складної обробки даних або навчання, його можна швидко впровадити навіть на пристроях із обмеженими обчислювальними ресурсами. Крім того, метод забезпечує високу швидкість обробки, якщо кількість шаблонів у базі даних є невеликою. Це робить його корисним у випадках, коли потрібно розпізнати текст із фіксованим набором символів, наприклад, при роботі з друкованими текстами або стандартними формами рукопису.

Однак зіставлення з шаблоном має значні обмеження, які зменшують його ефективність у більшості задач розпізнавання. Ключовим недоліком є низька варіативність у написанні символів. Рукописний текст різних людей може суттєво відрізнятися за формою, розміром, нахилом і товщиною ліній. Для забезпечення коректного розпізнавання кожного можливого варіанта символу необхідно зберігати великий набір шаблонів, що значно збільшує обсяг бази даних і ускладнює сам процес порівняння. Якість самого розпізнавання сильно залежить від якості вихідного зображення. Шум, розмиття або нерівномірне освітлення суттєво впливають на точність порівняння.

Методи на основі виділення ознак [5, 6] (Feature Extraction Methods) є більш складним та ефективним підходом до розпізнавання рукописного тексту. Їх принцип роботи полягає у виділенні ключових характеристик символів, таких як контури, кути, структури або текстурні особливості. Після цього ці характеристики використовуються як вхідні дані для класифікації символів. Для цього залучаються методи машинного навчання або статистичного аналізу. Завдяки виділенню інформативних ознак зображення стає менш залежним від таких факторів, як нахил або розмір символів, що підвищує точність розпізнавання.

Однією з головних переваг цього підходу є його стійкість до варіацій у написанні символів. Завдяки цьому методи на основі виділення ознак значно перевершують шаблонне розпізнавання у випадках, коли символи написані різними почерками або мають незначні деформації. Використання різних типів ознак, дозволяє адаптувати алгоритми для конкретних задач і покращувати точність розпізнавання навіть у складних умовах.

Складність у виборі та комбінуванні правильних ознак та пошук оптимального набору характеристик вимагає значного досвіду та експериментальної роботи, це вимагає виділення значних ресурсів тому є головним недоліком.

Метод k-найближчих сусідів [7] (k-Nearest Neighbors, k-NN) грунтується на класифікації нового зразка на основі його схожості з іншими зразками з навчальної вибірки. У просторі ознак, який використовується для представлення даних, обчислюється відстань між зразками. Для нового символу визначаються k найближчих сусідів за певною метрикою відстані. Клас символу встановлюється на основі більшості класів серед його найближчих сусідів (див. рис. 2).

Основною перевагою методу k-NN є його простота у реалізації та зрозумілість. Алгоритм не потребує складного етапу навчання, оскільки всі операції класифікації виконуються під час запиту. Гнучкість у виборі функції відстані та кількості сусідів дозволяє адаптувати метод до різних задач і умов розпізнавання. Для якісніших вхідних даних можна використовувати невелике значення параметра k, а для даних із високим рівнем шуму – більші значення k, що допомагає уникнути помилок через одиничні аномалії.

Основним недоліком цього методу є високі обчислювальні витрати. Оскільки для кожного нового

зразка потрібно обчислювати відстань до всіх зразків навчальної вибірки, це є серйозною проблемою при роботі зі значними обсягами даних. Чутливість до вибору параметра k також є важливою проблемою: занадто мале значення може призвести до перенавчання, тоді як занадто велике – до втрати важливої локальної інформації. Ефективність методу може суттєво знижуватися в умовах високої вимірності простору ознак.

Рис. 2. Принцип роботи k-NN

Метод опорних векторів [8] (Support Vector Machines, SVM) знаходить гіперплощину, яка максимально розділяє класи в багатовимірному просторі ознак. Для цього використовується поняття маржі – відстані між гіперплощиною та найближчими точками кожного класу. Необхідно знайти таку гіперплощину, яка забезпечує максимальну маржу, що мінімізує ризик помилок класифікації. У випадках, коли дані не можуть бути лінійно розділені в початковому просторі, алгоритм використовує спеціальні функції ядра (kernel functions), які проектують дані в простір більшої розмірності. У цьому просторі класи стають лінійно роздільними, що дозволяє застосовувати стандартний підхід (див. рис. 3).

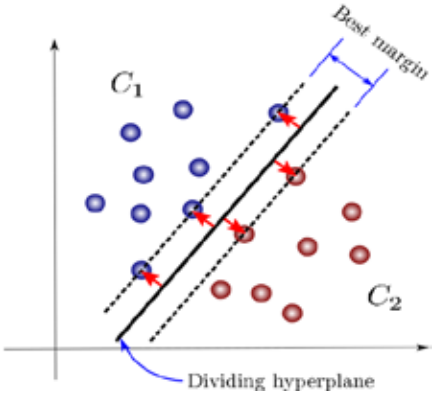

Рис. 3. Принцип роботи SVM

Метод опорних векторів дуже ефективний для задач із малими та середніми наборами даних, оскільки використовує лише дані, що лежать на межі між класами для побудови гіперплощини. Для цього процесу не знадобиться великої кількості ресурсів. SVM добре працює з багатовимірними даними, що робить його універсальним інструментом для різноманітних задач класифікації, стійкий до перенавчання за умови вдалого налаштування параметрів. Однак висока вимогливість до обчислювальних ресурсів при роботі з великими наборами даних є суттєвим недоліком, оскільки обчислення відстаней і побудова гіперплощини потребує значних ресурсів. Крім того, вибір правильного ядра та його параметрів вимагає експериментального підходу. Ще однією проблемою є чутливість до шуму в даних.

Нейронні мережі [10] (Neural Networks) складаються з великої кількості взаємопов'язаних нейронів, що утворюють шари. Процес обробки інформації полягає у проходженні вхідних даних через нейрони, кожен із яких застосовує вагові коефіцієнти, виконує обчислення за допомогою активаційної функції та передає результат на наступний шар. Особливістю нейронних мереж є їх здатність автоматично виділяти ознаки з вхідних даних без необхідності їх ручного визначення.

Нейронні мережі універсальні і здатні адаптуватися до варіацій у вхідних даних. Вони демонструють високу точність на великих наборах даних, ефективно знаходячи складні закономірності. Недоліками цього підходу є висока обчислювальна складність і необхідність великої кількості даних для навчання, підбір параметрів потребує досвіду чи експериментування.

Згорткові нейронні мережі [10, 11] (Convolutional Neural Networks, CNN) є спеціалізованим типом нейронних мереж, які широко використовуються для обробки зображень. Основним елементом їхньої архітектури є шари згортки, які виконують автоматичне виділення ознак зображення. Згорткові шари застосовують фільтри, що проходять через зображення, дозволяючи виявляти локальні ознаки, такі як контури, кути, текстури. Після шарів згортки слідують шари об’єднання (pooling), метою яких є зменшення розмірності даних, що зменшує кількість параметрів і обчислювальну складність, робить модель стійкішою до варіацій, таких як зміщення або масштабування символів. В кінці мережі зазвичай використовуються повнозв'язні шари, які виконують класифікацію на основі ознак, виділених попередніми шарами. Використання шарів підсумовування зменшує ризик перенавчання за рахунок зменшення розмірності даних.

Серед недоліків – CNN вимагають значних обчислювальних ресурсів і пам’яті для навчання. Налаштування архітектури, вибір кількості і розмірів фільтрів, типів активаційних функцій та параметрів навчання є складним і вимагає експериментального підходу. Згорткові нейронні мережі ефективно працюють лише за умови наявності великої кількості навчальних даних.

Рекурентні нейронні мережі [12] (Recurrent Neural Networks, RNN) використовуються для

обробки послідовних даних, зокрема тексту. Головною особливістю цих мереж є наявність зворотних зв'язків, які дозволяють зберігати інформацію про попередні символи в послідовності. Завдяки цьому RNN можуть моделювати залежності між елементами послідовності. Модифікації, такі як LSTM (Long Short-Term Memory) і GRU (Gated Recurrent Unit) вирішують проблему зникнення градієнта і дозволяють зберігати довготривалі залежності та покращувати ефективність навчання.

Високі вимоги до обчислювальних ресурсів є основним недоліком RNN. Крім того, налаштування і тренування таких моделей є складним завданням, що вимагає ретельного вибору параметрів і архітектури. Ще одним обмеженням є потреба в значному обсязі навчальних даних для досягнення високої точності, що може бути проблематичним у задачах із обмеженим доступом до великих датасетів.

Обговорення отриманих результатів. В результаті огляду було вирішено використати комбінацію згорткової нейронної мережі та рекурентної нейронної мережі разом з функцією Connectionist Temporal Classification. Такий підхід обирається через його високу ефективність у задачах розпізнавання тексту, особливо коли структура даних є послідовною і містить часові залежності. Крім цього необхідно знайти великий підходящий набір даних, оскільки обрані методи потребують великої кількості навчальних даних.

CNN здатні автоматично виявляти значущі ознаки на різних рівнях абстракції з вхідного зображення. Шари згортки обробляють локальні патерни, що особливо важливо для обробки зображень символів. Шари агрегування зменшують розмірність простору ознак, що зменшує обчислювальну складність. RNN зберігають інформацію про попередні символи, що важливо для розпізнавання послідовного тексту. Connectionist Temporal Classification (CTC) дозволяє моделі навчатись без потреби в попередній сегментації символів, що значно спрощує процес підготовки даних.

CTC є функцією втрат (loss function), розробленою для навчання моделей, які генерують послідовності з невизначеною довжиною. Вона застосовується в задачах, де вхідні дані мають одну довжину, а вихідні послідовності можуть мати іншу, як, у нашому випадку, розпізнавання рукописного тексту. CTC знаходить усі можливі відповідності між вхідною послідовністю (зображенням тексту) та вихідною послідовністю (розпізнаним текстом). Це дозволяє моделі враховувати різні можливі вирівнювання між вхідними та вихідними символами. Вводиться спеціальний символ «blank», який представляє відсутність символу на даній позиції. Це допомагає моделі коректно розпізнавати випадки, коли деякі позиції у вхідній послідовності не відповідають жодному символу у вихідній. Для кожного можливого вирівнювання між вхідними та вихідними символами обчислюється ймовірність. Ймовірності всіх можливих вирівнювань сумуються, щоб отримати кінцеву ймовірність відповідності вхідної та вихідної послідовностей після чого використовується негативний логарифм ймовірності правильного вирівнювання як функцію втрат. Метою навчання моделі є мінімізація цієї функції втрат.

ВИСНОВКИ

Проведене дослідження методів та алгоритмів розпізнавання рукописного тексту дозволило здійснити комплексний аналіз існуючих підходів та визначити оптимальну стратегію для розробки ефективної системи розпізнавання. Традиційні методи, такі як шаблонне розпізнавання та підходи на основі виділення ознак, хоча і залишаються корисними для специфічних задач з обмеженими наборами символів, демонструють значні обмеження у роботі з варіативним рукописним текстом. Алгоритми машинного навчання показали кращі результати порівняно з традиційними підходами, однак їх ефективність суттєво залежить від якості попередньої обробки даних та правильного вибору ознак.

На основі проведеного аналізу було обґрунтовано вибір комбінованого підходу, що поєднує згорткові нейронні мережі, рекурентні нейронні мережі та функцію Connectionist Temporal Classification (CTC). Ця архітектура максимально використовує переваги кожного компонента: CNN забезпечує ефективне виділення візуальних ознак з зображень символів, RNN моделює послідовний характер тексту та контекстуальні залежності між символами, а CTC дозволяє навчати модель без попередньої сегментації тексту.

Дослідження також виявило ключові виклики, які необхідно враховувати при практичній реалізації, зокрема потребу у великих обсягах навчальних даних та високі обчислювальні вимоги для мобільних пристроїв. Результати роботи створюють надійну теоретичну основу для розробки ефективних систем розпізнавання рукописного тексту та відкривають перспективи для подальших досліджень у галузі оптимізації архітектури мереж для мобільних пристроїв.

REFERENCES

[1] Neil, T. (2014). *Mobile Design Pattern Gallery: UI Patterns for Smartphone Apps.* O'Reilly Media.

[2] Hellman, E. (2013). *Android Programming: Pushing the Limits.* Wiley & Sons, Incorporated, John.

[3] How to easily do Handwriting Recognition using Machine Learning. URL: https://nanonets.com/blog/handwritten-character-recognition/

[4] Template Matching. URL: https://docs.adaptive-vision.com/4.7/studio/machine_vision_guide/TemplateMatching.html

[5] Feature Extraction. URL: https://www.mathworks.com/discovery/feature-extraction.html

[6] Bentahar, J., Moulin, B. & Bélanger, M. A taxonomy of argumentation models used for knowledge representation. *Artif Intell Rev 33*, 211–259 (2010). https://doi.org/10.1007/s10462-010-9154-1

[7] Kramer, O. (2013). K-Nearest Neighbors. In: Dimensionality Reduction with Unsupervised Nearest Neighbors. *Intelligent Systems Reference Library*, vol 51. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-38652-7_2

[8] Cortes, C., Vapnik, V. Support-vector networks. Mach Learn 20, 273–297 (1995). https://doi.org/10.1007/BF00994018

[9] Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN COMPUT. SCI.* 2, 420 (2021). https://doi.org/10.1007/s42979-021-00815-1

[10] Nebauer, C., Evaluation of convolutional neural networks for visual recognition, in *IEEE Transactions on Neural Networks*, vol. 9, № 4, pp. 685–696, July 1998, doi: 10.1109/72.701181.

[11] Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional neural networks: an overview and application in radiology. *Insights Imaging 9*, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

[12] Mienye, I. D., Swart, T. G., & Obaido, G. (2024). Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications. *Information*, 15(9), 517. https://doi.org/10.3390/info15090517

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] Neil, T. (2014). *Mobile Design Pattern Gallery: UI Patterns for Smartphone Apps.* O'Reilly Media.

[2] Hellman, E. (2013). *Android Programming: Pushing the Limits.* Wiley & Sons, Incorporated, John.

[3] How to easily do Handwriting Recognition using Machine Learning. URL: https://nanonets.com/blog/handwritten-character-recognition/

[4] Template Matching. URL: https://docs.adaptive-vision.com/4.7/studio/machine_vision_guide/TemplateMatching.html

[5] Feature Extraction. URL: https://www.mathworks.com/discovery/feature-extraction.html

[6] Bentahar, J., Moulin, B. & Bélanger, M. A taxonomy of argumentation models used for knowledge representation. *Artif Intell Rev 33*, 211–259 (2010). https://doi.org/10.1007/s10462-010-9154-1

[7] Kramer, O. (2013). K-Nearest Neighbors. In: Dimensionality Reduction with Unsupervised Nearest Neighbors. *Intelligent Systems Reference Library*, vol 51. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-38652-7_2

[8] Cortes, C., Vapnik, V. Support-vector networks. Mach Learn 20, 273–297 (1995). https://doi.org/10.1007/BF00994018

[9] Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN COMPUT. SCI.* 2, 420 (2021). https://doi.org/10.1007/s42979-021-00815-1

[10] Nebauer, C., Evaluation of convolutional neural networks for visual recognition, in *IEEE Transactions on Neural Networks,* vol. 9, № 4, pp. 685–696, July 1998, doi: 10.1109/72.701181.

[11] Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional neural networks: an overview and application in radiology. *Insights Imaging 9*, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

[12] Mienye, I. D., Swart, T. G., & Obaido, G. (2024). Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications. *Information*, 15(9), 517. https://doi.org/10.3390/info15090517

Дата надходження статті до редакції: 05.06.2025
Дата затвердження статті до друку: 14.06.2025