



МЕТОДИ ПОПЕРЕДНЬОГО ОБРОБЛЕННЯ БАГАТОМОВНИХ АРХІВНИХ РУКОПИСІВ У ЗАВДАННЯХ РОЗПІЗНАВАННЯ СИМВОЛІВ І ТЕКСТУ

Проаналізовано методи попереднього оброблення зображень архівних рукописних документів у завданнях автоматизованого розпізнавання символів і тексту. Розглянуто еволюцію підходів від класичних глобальних порогових методів до складніших локально-адаптивних, комбінованих і сучасних нейромережових рішень. Систематизовано основні класи методів попереднього оброблення зображень, зокрема – бінаризацію зображень, підсилення контрасту, приглушення шуму та морфологічне оброблення зображень, а також узагальнено принципи їх функціонування та особливості застосування. Проведено порівняльний аналіз переваг та обмежень цих підходів за умов деградованих архівних документів, для яких характерні неоднорідний фон, просвічування, шум оцифрування, плями, вицвітання чорнила, локальні дефекти носія та часткові втрати штрихової структури тексту. Особливу увагу приділено проблемі оброблення багатомовних рукописів, у яких поєднуються різні графічні системи, що істотно ускладнює не тільки сегментацію тексту, а й подальше розпізнавання символів. Показано, що поряд із завданнями відокремлення тексту від фону критично важливим є збереження мікроструктури штриха, контурних характеристик, локальних потовщень, вигинів, а також дрібних графічних відмінностей між символами різних алфавітів. Встановлено, що більшість наявних методів орієнтовані переважно на покращення якості бінарної маски та не забезпечують достатнього збереження інформативних ознак графем, що може негативно впливати на результати подальшого розпізнавання та знижувати точність роботи систем OCR (англ. *Optical Character Recognition*) та HTR (англ. *Handwritten Text Recognition*). Визначено, що найперспективнішими серед класичних підходів є комбіновані методи, які поєднують локально-адаптивну бінаризацію зображень зі структурно-орієнтованим шумозаглушенням і врахуванням особливостей зображення. Обґрунтовано необхідність подальшого розвитку спеціалізованих методів попереднього оброблення зображень, орієнтованих не тільки на сегментацію тексту, а й на збереження структурних характеристик багатомовних архівних рукописів, урахування їх графічної варіативності та підвищення ефективності сучасних систем автоматизованого розпізнавання рукописного тексту.

Ключові слова: архівні рукописи; розпізнавання тексту; оптичне розпізнавання символів; розпізнавання рукописного тексту; деградовані зображення; комп'ютерний зір.

Вступ / Introduction

Цифрове опрацювання архівних рукописних документів є важливим напрямом проведення досліджень у галузях комп'ютерного зору та аналізу документів, оскільки такі джерела містять значний обсяг історичної, лінгвістичної та культурної інформації, що є цінним для досліджень у різних гуманітарних і прикладних галузях. Водночас, автоматизоване розпізнавання документів істотно ускладнюється через деградацію носія, нерівномірний фон, просвічування, наявність шумів, дефектів паперу, а також велику варіативність графічних форм символів, що характерна для різних авторських стилів і періодів.

Проблеми аналізу рукописів досліджено у працях [1, 7, 8], де запропоновано широкий спектр методів попереднього оброблення документів. До них належать бінаризація з глобальними та локально-адаптивними порогами, підсилення контрасту, приглушення шуму, морфологічні операції та комбіновані схеми оброблення зображень. Ці методи дають змогу підвищити чіткість зображень і виділити текст на фоні, однак кожен підхід має свої обмеження і не завжди забезпечує збереження тонких штрихів або деталей графем [1, 8, 13]. Нагадаємо, *графема* – найменша одиниця системи письма, яка дає змогу розрізняти зміст слів. Простими словами, це абстрактна модель графічного знака (літери,

© Copyright 2026 by the author(s)

Лип'яніна-Гончаренко Христина Володимирівна, д-р техн. наук, доцент, кафедра інформаційно-обчислювальних систем і управління. Email: kh.lipianina@wunu.edu.ua; <https://orcid.org/0000-0002-2441-6292>

Івасечко Андрій Вікторович, викладач, кафедра інформаційно-обчислювальних систем і управління.

Email: andrewivasechko@gmail.com; <https://orcid.org/0009-0002-8623-7828>

Папінко Андрій Іванович, доктор phd, викладач, кафедра транспорту і логістики.

Email: andriypapinko@gmail.com; <https://orcid.org/0000-0001-8515-9376>

Гончаренко Олег Олександрович, мол. наук. співробітник, науково-дослідна частини.

Email: goncharenkooleg83@ukr.net; <https://orcid.org/0009-0008-0064-0995>

Цитування за ДСТУ: Лип'яніна-Гончаренко Х. В., Івасечко А. В., Папінко А. І., Гончаренко О. О. Методи попереднього оброблення багатомовних архівних рукописів у завданнях розпізнавання символів і тексту. Науковий вісник НЛТУ України. 2026, т. 36, № 3. С. 114–121.

Citation APA: Lipianina-Honcharenko, Kh. V., Ivasechko, A. V., Papinko, A. I., & Honcharenko, O. O. (2026). Methods for pre-processing multilingual archival manuscripts in character and text recognition tasks. *Scientific Bulletin of UNFU*, 36(3), 114–121.

<https://doi.org/10.36930/40360312>

цифри чи символу), що існує в нашій пам'яті незалежно від того, яким шрифтом чи почерком вона написана.

Водночас, більшість наявних рішень орієнтовані переважно на покращення сегментації тексту "текст-фон", і тільки частково враховують необхідність збереження мікроструктури штриха, що є критично важливим для аналізу багатомовних архівних рукописів та для забезпечення високої точності роботи систем OCR та HTR. Нагадаємо, OCR (англ. *Optical Character Recognition*) та HTR (англ. *Handwritten Text Recognition*) – це технології автоматизованого розпізнавання тексту, які перетворюють зображення (скани, фотографії) на цифровий текст, доступний для редагування та пошуку.

Актуальність дослідження полягає у необхідності вдосконалення методів попереднього оброблення зображень архівних рукописів у завданнях розпізнавання символів і тексту з урахуванням збереження інформативних ознак графем для підвищення точності роботи систем OCR і HTR.

Об'єкт дослідження – процес попереднього оброблення зображень архівних рукописних документів у завданнях автоматизованого розпізнавання тексту.

Предмет дослідження – методи та засоби попереднього оброблення зображень архівних рукописних документів, спрямовані на покращення якості виділення текстових структур і збереження інформативних ознак графем, що впливають на точність автоматизованого розпізнавання рукописного тексту.

Мета роботи – проаналізувати та узагальнити методи попереднього оброблення зображень архівних рукописних документів у завданнях автоматизованого розпізнавання символів і тексту, що дасть змогу визначити їх придатність для оброблення багатомовних архівних документів і збереження штрихової структури графем.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- проаналізувати наявні підходи до попереднього оброблення архівних рукописів, зокрема – глобальні, локально-адаптивні, комбіновані та нейромережеві методи, що дасть змогу визначити їх придатність для різних умов деградації документів;
- систематизувати основні класи методів попереднього оброблення зображень (бінаризація, підсилення контрасту, приглушення шуму, морфологічне оброблення) та узагальнити принципи їх роботи, що сприятиме формуванню цілісного уявлення про їхні функціональні можливості;
- оцінити переваги та обмеження проаналізованих методів у контексті деградованих архівних документів, що дасть змогу виявити їх практичну ефективність;
- дослідити особливості оброблення багатомовних рукописів з урахуванням необхідності збереження мікроструктури штрихів і відмінностей між графемами різних алфавітів, що забезпечить підвищення якості розпізнавання тексту;
- визначити найперспективніші підходи до попереднього оброблення зображень для завдань OCR/HTR, що сприятиме підвищенню точності роботи автоматизованих систем;
- виявити проблеми та обмеження наявних методів й окреслити напрями подальших етапів дослідження, що дасть змогу сформувати основу для розроблення нових ефективних рішень.

Аналіз останніх досліджень та публікацій. У дослідженні [7] автор представив глобальний пороговий метод бінаризації зображень, який відзначається прос-

тотою реалізації та автоматичним визначенням оптимального порога. Дослідження показало, що метод добре працює на однорідних зображеннях, проте у разі історичних рукописів з неоднорідним фоном, локальними плямами, плямами від вицвітання чорнила та іншими дефектами його застосування призводить до втрати слабких штрихів або до включення надлишкового фоновому шуму. Це обмежує придатність методу Otsu для оброблення складних архівних документів.

У дослідженні [1] автори проаналізували методи бінаризації зображень документів. Автори підкреслюють важливість бінаризації для підвищення розбірливості тексту, видалення фону та зменшення шуму на зображеннях документів, що робить їх придатнішими для подальшого автоматизованого аналізу та OCR. Робота охоплює як традиційні порогові методи, так і сучасні підходи на основі глибокого навчання, а також розглядає еталонні набори даних і метрики оцінювання продуктивності, пропонуючи інтегративний підхід до аналізу бінаризації зображень.

У роботі [8] запропоновано удосконалений локально-адаптивний підхід, який модифікує поріг із врахуванням локальної середньої яскравості та стандартного відхилення. Автори продемонстрували, що цей метод краще справляється з деградованими документами та підвищує точність виділення символів порівняно з класичним методом, описаним у роботі [1], однак він також залишається чутливим до розміру вікна та статистичних коливань локальних фрагментів, що може впливати на цілісність тонких штрихів.

У роботі [14] досліджено метод CLAHE (англ. *Contrast Limited Adaptive Histogram Equalization*) для локального підсилення контрасту зображень. Автор показав, що метод ефективно виділяє слабкі штрихи рукописного тексту та покращує видимість деталей на тлі пошкодженого паперу. Водночас, дослідження зазначає, що CLAHE може підсилювати текстuru паперу, плями та випадкові артефакти, що потребує додаткових кроків для приглушення шуму після підсилення контрасту.

У роботі [11] наведено метод bilateral filtering, який дає змогу ефективно приглушувати шум на зображенні, одночасно зберігаючи контури символів. Дослідження показало, що цей підхід забезпечує стабільніші результати роботи під час оброблення деградованих документів, зменшуючи втрату деталей у символах.

У дослідженні [2] запропоновано метод Non-Local Means (NLM), який здійснює шумозаглушення на основі подібності локальних фрагментів по всьому зображенню. Автори вказали, що NLM дає змогу значно покращити якість бінаризації деградованих документів, зберігаючи контури символів і зменшуючи ефект випадкового шуму, що робить його ефективним для старих архівних рукописів.

У роботі [4] запропоновано багатокрокову адаптивну схему оброблення історичних документів, яка охоплює попереднє фільтрування, оцінювання фону, бінаризацію зображень та постоброблення зображень. Дослідження показало, що така комбінована стратегія дає змогу підвищити якість сегментації тексту "екст-фон", однак автори відзначили, що основна увага приділяється побудові якісної бінарної маски, а не збереженню дрібних відмінностей між символами різних алфавітів.

Узагальнення проаналізованих підходів показує, що класичні методи попереднього оброблення зображень

відрізняються не тільки математичним принципом роботи, а й впливом на штрихову структуру символів. Глобальні порогові методи прості та швидкі, проте нестійкі до складного фонового шуму. Локально-адаптивні схеми чутливіші до неоднорідності документа, але можуть фрагментувати тонкі штрихи або підсилювати шум. Методи приглушення шуму покращують якість бінаризації, але не вирішують проблему подібності між символами. Комбіновані схеми гнучкіші, проте підвищують ризик деформації графем. Загалом більшість рішень оптимізує сегментацію "екст-фон", тоді як для багатомовних рукописів важливе збереження дрібних штрихових відмінностей. Тому класичні методи потрібно розглядати як базовий рівень, а подальші дослідження спрямовувати на спеціалізовані структурно-орієнтовані підходи.

Матеріали та методи дослідження. У дослідженні використано архівні рукописні зображення з типовими ознаками деградації, зокрема – шумом, нерівномірним фоном і просвічуванням. Для аналізу застосовано класичні методи попереднього оброблення зображень, що містять глобальну та локально-адаптивну бінаризацію зображень, підсилення контрасту, приглушення шуму та морфологічне оброблення зображень, а також їх комбіновані варіанти. Ефективність підходів оцінено шляхом порівняльного аналізу їх впливу на якість відокремлення тексту від фону та збереження штрихової структури символів з урахуванням особливостей багатомовних рукописів.

Результати дослідження та їх обговорення / Research results and their discussion

Попереднє оброблення зображень архівних рукописних документів є проміжною, але критично важливою ланкою між етапом оцифрування документа та подальшим розпізнаванням символів або текстових послідовностей. На відміну від сучасних друкованих документів, історичним рукописам характерні неоднорідний фон, просвічування, сліди старіння паперу, знижений контраст, локальні втрати чорнила та широка варіативність штрихової структури графем. Саме тому попереднє оброблення має не тільки відокремлювати текст від фону, а й зберігати інформативні ознаки штриха, що є визначальними для подальшого розпізнавання багатомовних символів. У межах цього огляду розглянуто сім схем попереднього оброблення зображень, які охоплюють основні класи класичних рішень: від відсутності перетворень і глобального порогоування до комбінацій локального підсилення контрасту, адаптивної бінаризації, приглушення шуму та морфологічного очищення [5, 6].

Нехай $I(x, y)$ – напівтонове зображення рукописного документа, де x, y – просторові координати, а значення інтенсивності нормовані в інтервалі $[0,1]$. Загальне завдання попереднього оброблення зображень полягає у побудові перетвореного зображення $I(x, y)$, що покращує видимість текстових структур, та/або бінарної маски $B(x, y) \in \{0,1\}$, де одиниця відповідає пікселям тексту, а нуль – фону. Формально це можна подати як відображення [5, 6]:

$$\Phi: I(x, y) \rightarrow \{J(x, y), B(x, y)\}, \quad (1)$$

де оператор Φ може містити фільтрацію, локальне або глобальне перетворення контрасту, порогоування і морфологічні операції. У практичних системах НТР саме вибір оператора Φ значною мірою визначає, чи буде

збережено тонкі штрихи, переходи товщини лінії та графічні відмінності між близькими символами різних алфавітів [5, 6].

Базовий варіант розпізнавання рукописного тексту без попереднього оброблення зображення. Базовий варіант гав передбачає відсутність будь-яких попередніх перетворень зображення, тобто далі воно передається без змін [7]:

$$I(x, y) = I(x, y). \quad (2)$$

Такий підхід використовують не як самостійний засіб покращення якості документа, а як контрольний сценарій, що дає змогу оцінити реальний внесок кожного наступного методу. У контексті аналізу рукописів цей варіант є важливим, оскільки дає змогу перевірити, чи не призводить агресивна бінаризація або контрастування до втрати корисної інформації, яка ще присутня в оригінальному grayscale-зображенні. Саме порівняння з гав дає підстави говорити не тільки про візуальне очищення, а про реальний вплив препроцесингу на подальше розпізнавання рукописного тексту [5, 6].

Розмивання Гауса з глобальною бінаризацією Otsu. Метод Otsu (англ. *Otsu's method*) – алгоритм комп'ютерного зору, який використовують для автоматичного визначення оптимального порогу бінаризації зображення (перетворення на чорно-біле). Він автоматично знаходить поріг, що мінімізує дисперсію всередині класів (пікселі фону та об'єкта). Водночас, схема gaussian_otсу належить до класу глобальних порогових підходів. На першому кроці виконується згладжування зображення за допомогою гаусового ядра [7]:

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad (3)$$

після чого формується проміжне зображення:

$$I_G(x, y) = I(x, y) * G_{\sigma}(x, y), \quad (4)$$

де: x, y – координати; σ – згладжування; $I_G(x, y)$ – згладжене зображення після використання фільтра Гауса; $*$ – операція згортки; $G_{\sigma}(x, y)$ – гаусове ядро. Далі застосовують метод Otsu, який обирає глобальний поріг t^* шляхом максимізації міжкласової дисперсії:

$$t = \arg \max_t \sigma_B^2(t), \quad (5)$$

$$\sigma_B^2(t) = \omega_0(t)\omega_1(t)(\mu_0(t) - \mu_1(t))^2, \quad (6)$$

де: $\omega_0(t), \omega_1(t)$ – частки двох класів; $\mu_0(t), \mu_1(t)$ – середні інтенсивності двох класів. Бінарне зображення отримують як:

$$B(x, y) = \begin{cases} 1, & I_G(x, y) \leq t^*, \\ 0, & I_G(x, y) > t^*. \end{cases} \quad (7)$$

Перевагою цього підходу є простота реалізації, низька обчислювальна складність й ефективна робота на зображеннях із відносно однорідним фоном. Водночас, для архівних рукописів глобальний поріг часто виявляється нестійким через локальну змінність фону, а попереднє гаусове згладжування може послаблювати тонкі штрихи й дрібні елементи графем [5, 6, 7].

Локальне підсилення контрасту CLAHE з наступною бінаризацією Otsu. Метод clahe_otсу поєднує локальне підсилення контрасту та глобальне порогоування. У його основі знаходиться Contrast Limited Adaptive Histogram Equalization, де зображення поділяють на малі блоки Ω_k , для кожного з яких будується локальна гістограма $H_k(i)$. Після обрізання піків гістограми на рівні

clip-limit β виконується локальне перетворення яскравості [1]:

$$I_k(x, y) = T_k(I(x, y)), \quad (8)$$

де T_k – локальне монотонне відображення, сформоване на основі вирівнювання гістограм в межах блока Ω_k . Після інтерполяції пікселів між сусідніми блоками формується глобально посилене зображення $I_{CLAHF}(x, y)$, до якого далі застосовують глобальний поріг Otsu. Теоретично такий підхід має підвищувати видимість блідих або низькоконтрастних штрихів, однак на історичних документах CLAHF нерідко посилює не тільки текстові елементи, а й паперову фактуру документа, плями та шум. Унаслідок цього остаточна бінаризація зображення може давати надлишкові артефакти або штучне потовщення контурів тексту [1, 5, 7].

Медіанна фільтрація зображення з адаптивною гаусовою бінаризацією. Схема median_adaptive_gaussian спрямована на придушення імпульсного шуму та локальне порогоування з урахуванням просторової структури документа. На першому етапі формується медіанно відфільтроване зображення, а саме:

$$I_M(x, y) = \text{median}\{I(u, v) : (u, v) \in N_{xy}\}, \quad (9)$$

де: $I(u, v)$ – інтенсивність пікселя після перетворення; N_{xy} – локальне вікно навколо точки (x, y) . Такий фільтр належить до порядкових нелінійних операторів і є стійким до поодиноких аномальних пікселів, що особливо важливо за наявності дрібних шумових краплень. Після цього для кожного пікселя обчислюється локальний гаусово-зважений поріг [2, 5, 13]:

$$T_G(x, y) = \sum_{(u, v) \in N_{xy}} w_G(u, v) I_M(x - u, y - v) - C, \quad (10)$$

де: $w_G(u, v)$ – коефіцієнти гаусового ядра; $I_M(x, y)$ – напівтонове зображення маски рукописного документа; C – коригувальна константа. Остаточна бінаризація задається як [2, 5, 13]:

$$B(x, y) = \begin{cases} 1, & I_M(x, y) \leq T_G(x, y), \\ 0, & I_M(x, y) > T_G(x, y). \end{cases} \quad (11)$$

Перевага цього методу полягає в тому, що він краще, ніж глобальні схеми, враховує локальні зміни освітленості та може зменшувати кількість дрібних шумових точок. Водночас, за складного фону або дуже тонких штрихів медіанна фільтрація та локальний поріг можуть спричиняти часткову фрагментацію графеми [2, 5, 13].

Двобічна фільтрація зображення з адаптивною середньою бінаризацією. Метод bilateral_adaptive_mean поєднує двобічну фільтрацію та адаптивну середню бінаризацію зображення. Двобічна фільтрація належить до методів приглушення шуму зі збереженням меж, оскільки враховує не тільки просторову близькість пікселів, а й подібність їх інтенсивностей [9, 11]. У загальному вигляді відфільтроване значення пікселя p визначається як зважене середнє сусідніх пікселів q у локальному вікні $N(p)$:

$$I_B(p) = \frac{1}{W_p} \sum_{q \in N(p)} \text{Exp}\left(\frac{-\|p - q\|^2}{2\sigma_s^2}\right) \text{Exp}\left(\frac{-(I(p) - I(q))^2}{2\sigma_r^2}\right) I(q), \quad (12)$$

де: $I_B(p)$ – значення після двобічної фільтрації; W_p – нормувальний коефіцієнт; p, q – відфільтроване значення відповідного пікселя; σ_s – параметр просторового згладжування; σ_r – параметр згладжування за інтенсивністю. На відміну від звичайного гаусового згладжуван-

ня, такий підхід дає змогу приглушувати шум без істотного розмиття меж штриха, що є важливим для рукописних зображень [9, 11].

Після етапу фільтрації виконується адаптивна середня бінаризація, у межах якої для кожного пікселя обчислюється локальний поріг як середнє значення інтенсивностей у вікні $N(p)$ з поправкою на константу C [1]:

$$T(p) = \frac{1}{(|N(p)|) \sum_{q \in N(p)} I_B(q)} - C. \quad (13)$$

Далі формується бінарне зображення, у якому пікселі належать до тексту або фону залежно від порівняння з локальним порогом [8]:

$$B(p) = \begin{cases} 1, & I_B(p) \leq T(p), \\ 0, & I_B(p) > T(p). \end{cases} \quad (14)$$

Перевагою такого підходу є можливість поєднати приглушення дрібного шуму зі збереженням контурів символів і локально чутливим відокремленням тексту від фону. Це робить метод потенційно корисним для рукописів із розмитими межами, але відносно контрастними межами графем. Водночас, його ефективність істотно залежить від добору параметрів фільтра та розміру локального вікна, а за складного архівного фону він може залишати паразитні структури фонового шуму, або не забезпечувати достатнього відокремлення близьких штрихів [9, 11, 13].

Нелокальне приглушення шуму Fast Non-Local Means з локальною бінаризацією Sauvola. Схема nlm_sauvola є однією з найперспективніших серед класичних комбінованих рішень для деградованих документів. На першому кроці використовують Non-Local Means, де значення пікселя відновлюється як зважене середнє по множині подібних околів [8]:

$$I_{NLM}(p) = \sum_{q \in \Omega} w(p, q) I(q) \sum_{q \in \Omega} w(p, q) = 1. \quad (15)$$

Вагомості $w(p, q)$ визначаються подібністю локальних фрагментів навколо точок p і q , а не тільки їх просторовою близькістю, що робить цей підхід стійкішим до складного шуму [11]. На другому етапі застосовують локальний поріг Sauvola [5]:

$$T_S(x, y) = m(x, y) [1 + k(S(x, y) / R - 1)], \quad (16)$$

де: $m(x, y)$ – локальне середнє; $S(x, y)$ – локальне стандартне відхилення; R – максимальне очікуване значення стандартного відхилення; $k(\cdot)$ – емпіричний коефіцієнт. Бінарна маска формується порівнянням $I_{NLM}(x, y)$ з $T_S(x, y)$. Для архівних рукописів така комбінація є привабливою, оскільки дає змогу приглушувати шум без надмірного руйнування текстових структур і водночас адаптувати поріг до локальної статистики документа. Водночас, метод є обчислювально дорожчим, а за невдалого підбору параметрів NLM може надмірно згладжувати слабкі елементи графеми [5, 8, 11].

Комбінований підхід CLAHF + Sauvola + морфологічне відкриття і закриття. Метод clahe_sauvola_morph поєднує локальне підсилення контрасту, адаптивну бінаризацію зображень та постоброблення бінарної маски морфологічними операціями. На першому етапі формується посилене зображення $I_{CLAHF}(x, y)$. Далі застосовують локальний поріг Sauvola [4]:

$$B_0(x, y) = \begin{cases} 1, & I_{CLAHF}(x, y) \leq T_S(x, y), \\ 0, & I_{CLAHF}(x, y) > T_S(x, y). \end{cases} \quad (17)$$

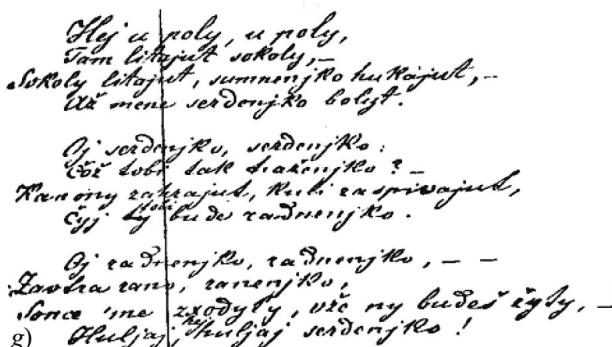
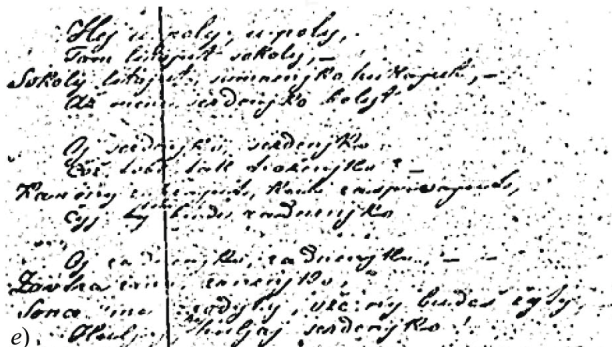
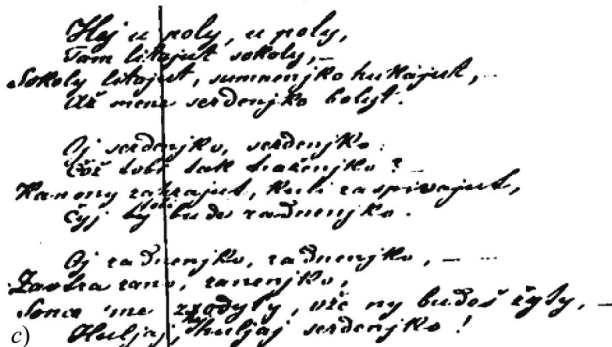
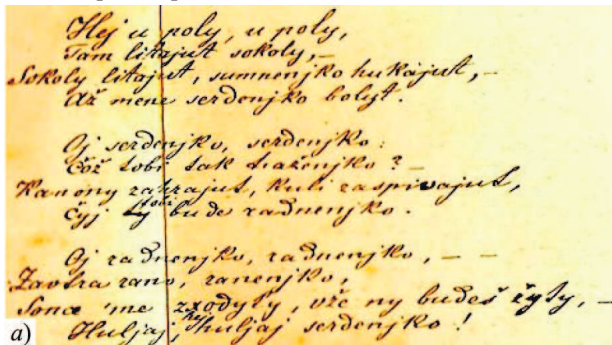
Після цього виконується морфологічне відкриття

$$B_{open} = (B_0 \odot S) \oplus S \quad (18)$$

та морфологічне закриття

$$B_{close}(B_{open} \oplus S) \odot S, \quad (19)$$

де: S – структурний елемент; $\odot$ – ерозія; $\oplus$ – дилатація. Відкриття сприяє видаленню дрібних шумових об'єктів, а закриття – заповненню малих розривів і вирівнюванню контуру [4]. Для архівних рукописів такий підхід може бути корисним у випадках, коли після бінаризації зображення утворюється значна кількість дрібних фонових артефактів. Однак надмірна комбінація CLANE, локального пороговування та морфологічного очищення може призвести до деформації реальної форми символу, злиття сусідніх штрихів або втрати тонких елементів письма [1, 4, 8].



На рисунку показано візуалізацію кожного з методів попереднього оброблення багатомовних архівних рукописів у завданнях розпізнавання символів і тексту. На цьому рисунку наведено такі позначення: А – оригінал; В – базовий варіант розпізнавання рукописного тексту без попереднього оброблення зображення; С – розмивання Гауса з глобальною бінаризацією Otsu.; D – локальне підсилення контрасту CLANE з наступною бінаризацією Otsu; E – медіанна фільтрація зображення з адаптивною гаусовою бінаризацією; F – двобічна фільтрація зображення з адаптивною середньою бінаризацією; G – нелокальне пригнічення шуму Fast Non-Local Means з локальною бінаризацією Sauvola; H – комбінований підхід CLANE + Sauvola + морфологічне відкриття і закриття.

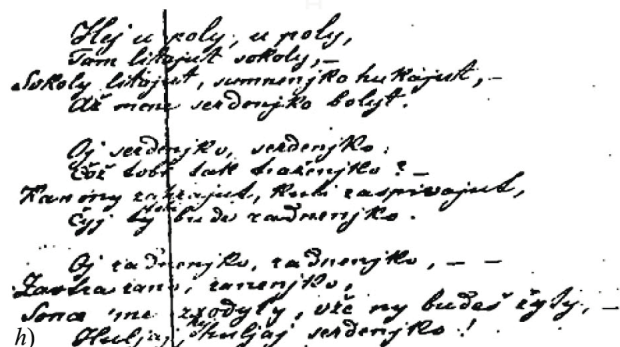
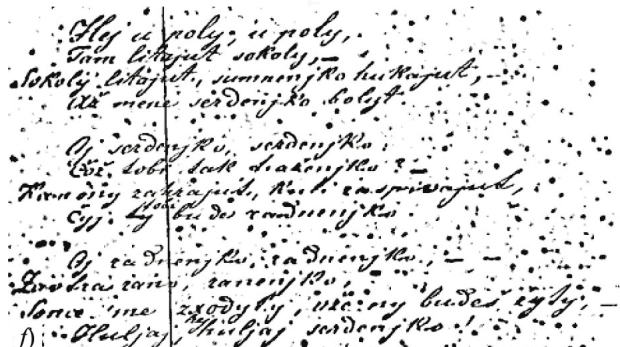
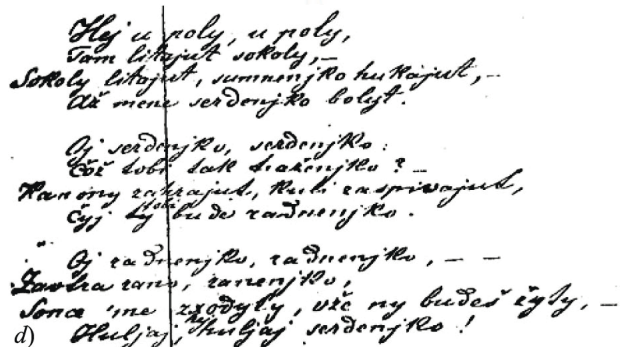
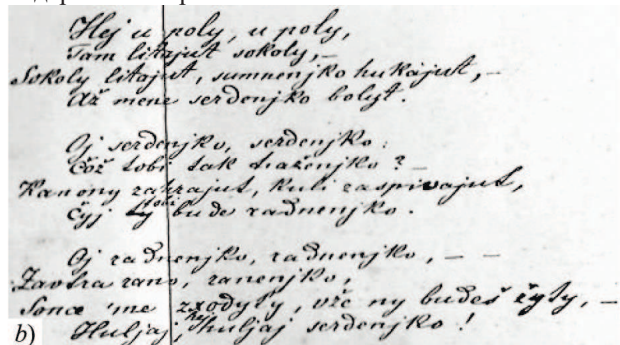


Рисунок. Візуалізація методів попереднього оброблення багатомовних архівних рукописів / Visualisation of methods for the preliminary processing of multilingual archival manuscripts

Таблиця. Узагальнення щодо придатності проаналізованих методів попереднього оброблення зображень /
Generalization of the suitability of the considered image preprocessing methods

Метод	Математична ідея	Переваги	Недоліки	Придатність до багатомовного архівного рукопису
Raw	Відсутність перетворення [5, 6]: $I(x, y) = I(x, y)$.	Зберігає повну первинну інформацію про градації яскравості, товщину та плавність штриха; не вносить штучних спотворень.	Не усуває шум, фонову неоднорідність, просвічування та артефакти оцифрування.	Доцільний як базова контрольна схема; може бути корисним, якщо подальша модель здатна самостійно працювати з grayscale-структурою.
Gaussian + Otsu	Гаусове згладжування $I_G = I \cdot G_\sigma$ глобальний поріг t^* , що максимізує міжкласову дисперсію.	Простота реалізації, низька обчислювальна складність, ефективна робота на відносно чистому фоні.	Глобальний поріг нестійкий до локальних змін фону; можливе згладжування тонких штрихів.	Помірна придатність; ефективний для контрастніших документів, але обмежений для сильно деградованих багатомовних рукописів.
CLAHE + Otsu	Локальне вирівнювання гістограм у блоках + глобальне порогоування.	Підсилює бліді штрихи, локально збільшує контраст слабких символів.	Разом із корисними деталями підсилює паперову текстуру, шум і плями; глобальний поріг не завжди адекватний після CLAHE.	Обмежена придатність; корисний тільки за слабого контрасту, але ризикований для складного архівного фону.
Median + Adaptive Gaussian	Медіанна фільтрація I_M + локальний гаусово-зважений поріг $T_G(x, y)$.	Добре усуває імпульсний шум, краще враховує локальні перепади яскравості, ніж глобальні методи.	Може розривати тонкі штрихи та спрощувати дрібні елементи графем.	Досить придатний для рукописів із локальним шумом, однак не гарантує збереження дрібних міжалфавітних відмінностей.
Bilateral + Adaptive Mean	Edge-preserving filtering з вагами за просторовою та інтенсивнісною близькістю + локальний середній поріг.	Згладжує шум зі збереженням контурів зображення, краще працює з розмитими межами, ніж звичайне згладжування.	Висока чутливість до параметрів; недостатньо ефективний за складного текстурного фону.	Помірна придатність; перспективний для збереження контуру штриха, але недостатній як самостійне рішення для багатомовного архівного матеріалу.
NLM + Sauvola	Нелокальне усереднення за подібністю оточів + локальний статистичний поріг $T_S(x, y)$.	Добре приглушує шум і водночас зберігає текстові структури; адаптується до неоднорідного фону.	Вища обчислювальна складність; за агресивного налаштування можливе згладжування слабких деталей.	Одна з найпридатніших класичних схем для архівного рукопису, особливо за наявності складного фону та слабого контрасту.
CLAHE + Sauvola + Morphology	Локальне підсилення контрасту + локальний поріг + морфологічне відкриття/закриття.	Дає змогу поєднати підсилення штриха, сегментацію тексту та очищення дрібних артефактів.	Підвищує ризик потовщення контуру, злипання сусідніх штрихів або втрати дрібних елементів.	Частково придатний, але потребує дуже обережного налаштування; для багатомовних графем може деформувати критично важливі відмінності.
Спеціалізовані класичні підходи для деградованих документів (Gatos, Hedjam, Howe)	Багатокрокова адаптивна схема, просторово-статистичне моделювання або автоматичне налаштування параметрів.	Краще враховують деградації документа, ніж базові порогові методи; орієнтовані на історичні матеріали.	Переважно оптимізують якість бінарної маски, а не збереження унікальних ознак графем.	Вища придатність, ніж у простих схем, однак усе ще недостатня для завдання точного розрізнення візуально близьких багатомовних символів.
Глибокі нейромережеві методи бінаризації	Навчання нелінійного відображення $I(x, y) \rightarrow B(x, y)$ на основі піксельної класифікації або сегментації тексту.	Висока гнучкість, здатність моделювати складні типи деградацій і фонів.	Залежність від великих розмічених даних, значних обчислювальних ресурсів і доменної узгодженості.	Потенційно найперспективніші, але для багатомовного архівного рукопису потребують спеціалізованих навчальних корпусів й адаптації до конкретної графічної традиції.

Отже, результати узагальнення придатності проаналізованих методів попереднього оброблення зображень свідчать, що серед класичних підходів найперспективнішими для архівного рукопису є комбіновані схеми локально-адаптивної бінаризації та структурно-орієнтованого шумозаглушення, зокрема – підходи на основі Sauvola та Non-Local Means. Водночас, жоден із проаналізованих методів не забезпечує повною мірою збереження тонких штрихових відмінностей між багатомовними графемами за умов складної деградації. Це обґрунтовує доцільність розроблення спеціалізованих методів попереднього оброблення зображень, орієнтованих не тільки на виділення тексту з фону, а й на збереження структурних ознак символів, критично важливих для подальшого автоматизованого розпізнавання рукописного тексту.

Обговорення результатів дослідження. У дослідженні [5] автори запропонували просторово-адаптивний статистичний підхід до бінаризації історичних документів. Вони продемонстрували високу ефективність методу під час роботи з документами, що зазнали значної деградації, включно з нерівномірним освітленням, плямами та плямами від чорнила. Водночас, автори зазначили, що основна орієнтація підходу спрямована на покращення сегментації тексту від фону, тоді як збереження тонких графічних ознак символів, особливо у багатомовних документах, залишає бажати кращого, що може призводити до втрати дрібних штрихів та ускладнювати подальше розпізнавання символів і тексту.

У роботі [6] автор дослідив метод автоматичного налаштування параметрів бінаризації для деградованих

рукописних документів. Результати дослідження показали підвищення стабільності отриманих бінарних зображень у разі зміни параметрів, що значно спрощує оброблення великих наборів архівних зображень. Проте, як і у попередньому випадку, метод не забезпечує достатнього збереження дрібних штрихових елементів, що особливо важливо для документів із складним фоном шумом і різними алфавітами, де мінімальні відмінності між символами визначальні для точності розпізнавання тексту документа.

У роботі [10] автори проаналізували застосування глибоких нейромереж для бінаризації зображень документів, формулюючи завдання як піксельну класифікацію. Дослідження показало, що нейромережеві моделі здатні ефективно справлятися з різноманітними видами деградації, враховувати локальний контекст та складні типи фонів артефактів. Водночас, автори вказали на те, що точність отриманих результатів значною мірою залежить від обсягу та якості навчальних наборів даних, а також від доменної специфіки документів, що обмежує їхнє безпосереднє застосування для архівних багатомовних рукописів без додаткової адаптації.

У дослідженні [9] автори здійснили систематичний аналіз сучасних методів бінаризації деградованих документів. Вони дійшли висновку, що більшість підходів спрямована на поліпшення розділення тексту та фону, при цьому недостатньо уваги приділено збереженню структурних характеристик графем, що є критичним для документів із складними шрифтами, багатомовними текстами та старовинними почерками.

У роботі [13] автори узагальнили сучасні тенденції розвитку методів бінаризації, відзначивши перехід від глобального порогоування до локально-адаптивних, комбінованих та гібридних схем. При цьому вони вказали на те, що питання збереження мікроструктури штрихів та оброблення багатомовних документів досі недостатньо досліджене, а наявні підходи переважно оптимізовані під сегментацію "текст-фон".

Результати дослідження узгоджуються з сучасними науковими уявленнями про обмеженість підходів, орієнтованих виключно на бінаризацію зображень. Встановлено, що для багатомовних архівних рукописів визначальним є не тільки якісне відокремлення тексту від фону, а й збереження мікроструктури штриха та графічної варіативності символів. Це підтверджує доцільність використання комбінованих методів як базового підходу та обґрунтовує необхідність подальшого розвитку спеціалізованих методів попереднього оброблення рукописного тексту, орієнтованих на структурні характеристики графем і підвищення ефективності роботи систем OCR і HTR.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – удосконалено методіку аналізу методів попереднього оброблення зображень архівних рукописних документів, яка, на відміну від наявних методик, систематизує та оцінює зображення з позиції збереження мікроструктури штриха і графічних відмінностей багатомовних символів, що дає змогу підвищити обґрунтованість вибору методів для завдань OCR і HTR задля підвищення точності розпізнавання графем.

Практична значущість результатів дослідження – можна використати для вибору та комбінування методів попереднього оброблення зображень архівних рукописних документів у системах OCR і HTR задля підвищення точності розпізнавання графем.

Висновки / Conclusions

Проаналізовано та узагальнено методи попереднього оброблення зображень архівних рукописних документів у завданнях автоматизованого розпізнавання символів і тексту, що дало змогу визначити їх придатність для оброблення багатомовних архівних документів і збереження штрихової структури графем. За результатами проведеного дослідження можна зробити такі основні висновки.

1. Систематизовано наявні підходи до попереднього оброблення рукописних зображень: від глобальних порогових методів до локально-адаптивних та комбінованих схем, із зазначенням їхніх принципів відмінностей та обмежень. Це дало можливість виявити їхню придатність до оброблення архівних рукописів різної якості, визначити оптимальні сфери застосування кожного підходу та обґрунтувати вибір ефективних методів для подальшого розпізнавання тексту.
2. Проаналізовано ефективність застосування глобальних методів: вони забезпечують швидке оброблення зображень, проте мають обмежену стійкість до локальних дефектів та неоднорідного фону, що впливає на роботу з багатомовними рукописами.
3. Встановлено, що локально-адаптивні методи підвищують чутливість до дрібних деталей рукописного тексту, але можуть спричиняти фрагментацію тонких штрихів та підсилення шуму, що потребує точного налаштування параметрів їх оброблення.
4. З'ясовано, що комбіновані підходи до оброблення зображень забезпечують більш збалансоване оброблення тексту та фону, хоча вони не завжди гарантують повне збереження штрихових відмінностей між близькими символами різних алфавітів. Це дало можливість обґрунтувати доцільність їх використання як базового етапу з подальшим застосуванням спеціалізованих методів уточнення структури символів.
5. Досліджено, що більшість методів виявлення тексту орієнтовані на сегментацію "текст-фон", але не вирішують повною мірою завдання збереження унікальної графічної структури символів, що критично для точного розпізнавання рукописного тексту. Все це дає змогу виявити обмеження наявних підходів та визначити напрями їх удосконалення з урахуванням структурних особливостей графем.
6. Обґрунтовано необхідність розроблення спеціалізованих структурно-орієнтованих методів оброблення зображень, які поєднують виділення тексту з фону та збереження мікроструктури штрихів, а також створення навчальних корпусів для багатомовних рукописів.
7. Проведено узагальнення, згідно з яким класичні методи та їх комбінації забезпечують базове оброблення зображень і слугують надійною відправною точкою для подальшого аналізу, проте реальні потреби автоматизованого розпізнавання тексту потребують нових підходів із увагою до збереження унікальних ознак графем.

References

1. Bataineh, B., Tounsi, M., Zamzami, N., Janbi, J., Abu Ain, W. A., & Elnazer, S. (2025). A comprehensive review on document image binarization. *Journal of Imaging*, 11(5), article number 133. <https://doi.org/10.3390/jimaging1105013>
2. Buades, A., Coll, B., & Morel, J.-M. (2011). Non-local means denoising. *Image Processing On Line*, 1, 208–212. https://doi.org/10.5201/ipol.2011.bcm_nlm

3. Dyak, T. P., & Hrytsiuk, Y. I. (2024). Application of corpus tools to identify keywords of Ukrainian rebel songs as a genre of the folklore discourse. *Scientific Bulletin of UNFU*, 34(7), 60–71. <https://doi.org/10.36930/40340708>
4. Gatos, B., Pratikakis, I., & Perantonis, S. J. (2006). Adaptive degraded document image binarization. *Pattern Recognition*, 39(3), 317–327. <https://doi.org/10.1016/j.patcog.2005.09.010>
5. Hedjam, R., Farrahi Moghaddam, R., & Cheriet, M. (2011). A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images. *Pattern Recognition*, 44(9), 2184–2196. <https://doi.org/10.1016/j.patcog.2011.02.021>
6. Howe, N. R. (2013). Document binarization with automatic parameter tuning. *International Journal on Document Analysis and Recognition*, 16(3), 247–258. <https://doi.org/10.1007/s10032-013-0204-2>
7. Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
8. Sauvola, J., & Pietikäinen, M. (2000). Adaptive document image binarization. *Pattern Recognition*, 33(2), 225–236. [https://doi.org/10.1016/S0031-3203\(99\)00055-2](https://doi.org/10.1016/S0031-3203(99)00055-2)
9. Sulaiman, A., Omar, K., & Nasrudin, M. F. (2019). Degraded historical document binarization: A review on issues, challenges, techniques, and future directions. *Journal of Imaging*, 5(4), article number 48. <https://doi.org/10.3390/jimaging5040048>
10. Tensmeyer, C., & Martinez, T. (2017). Document image binarization with fully convolutional neural networks. In *Proceedings of the International Conference on Document Analysis and Recognition* (pp. 99–104). <https://doi.org/10.1109/ICDAR.2017.28>
11. Tomasi, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Proceedings of the International Conference on Computer Vision* (pp. 839–846). <https://doi.org/10.1109/ICCV.1998.710815>
12. Torskyi, O. I., & Hrytsiuk, Y. I. (2025). Application of machine learning to enhance the efficiency of automated software testing. *Scientific Bulletin of UNFU*, 35(4), 142–149. <https://doi.org/10.36930/40350416>
13. Yang, Z., Zuo, S., Zhou, Y., He, J., & Shi, J. (2024). A review of document binarization: Main techniques, new challenges, and trends. *Electronics*, 13(7), article ID 1394. <https://doi.org/10.3390/electronics13071394>
14. Zuiderveld, K. J. (1994). Contrast limited adaptive histogram equalization. In P. S. Heckbert (Ed.), *Graphics Gems IV* (pp. 474–485). Academic Press. <https://doi.org/10.1016/B978-0-12-336156-1.50061-6>

Kh. V. Lipianina-Honcharenko, A. V. Ivasechko, A. I. Papinko, O. O. Honcharenko

West Ukrainian National University, Ternopil, Ukraine

METHODS FOR PRE-PROCESSING MULTILINGUAL ARCHIVAL MANUSCRIPTS IN CHARACTER AND TEXT RECOGNITION TASKS

The paper presents an analytical review of preprocessing methods for archival handwritten images in automated character and text recognition tasks. The evolution of approaches from classical global thresholding methods to more advanced locally adaptive, combined, and modern neural network-based solutions is examined. The main classes of preprocessing methods are systematized, including binarization, contrast enhancement, noise suppression, and morphological processing, and the principles of their operation are generalized with consideration of their applicability to degraded historical materials. A comparative analysis of the advantages and limitations of these approaches is conducted under conditions of degraded archival documents, which are characterized by non-uniform background, bleed-through, digitization noise, stains, ink fading, local defects, and partial loss of stroke structure. Particular attention is paid to the problem of processing multilingual manuscripts, in which different writing systems coexist, significantly complicating not only text segmentation but also subsequent character recognition. It is shown that, in addition to text-background separation, preserving stroke microstructure, contour characteristics, local thickness variations, curves, and fine graphical distinctions between visually similar characters of different alphabets is critically important. It is established that most existing methods are primarily focused on improving binarization quality and do not sufficiently preserve informative grapheme features, which may negatively affect recognition accuracy in OCR and HTR systems. The analysis demonstrates that preprocessing quality directly influences the stability of feature extraction and classification stages. It is determined that the most promising classical approaches are combined methods that integrate locally adaptive binarization with structure-oriented noise suppression and context-aware filtering. However, even such approaches have limitations when dealing with highly degraded and heterogeneous archival materials. Therefore, the necessity of further development of specialized preprocessing methods aimed at preserving structural characteristics of multilingual archival manuscripts, accounting for their graphical variability, and improving the overall robustness and accuracy of OCR and HTR systems is substantiated.

Keywords: archival manuscripts; text recognition; OCR; HTR; degraded images; computer vision.