Digitizing the Past: Computational Methods for Manuscript Decipherment in Bangladesh

Tahmina Kabir Tresha
Deputy Librarian University of Dhaka Dhaka, Bangladesh

Abstract
This research article focuses on the role of the computational methods in deciphering, prese--rvation and access of Manuscripts in Bangla--desh.The study tackles a contemporary heritage challenge, in which both Bengali and multilingual manuscripts are often fragile, dispersed, hard to read, and poorly linked into searchable digital environments. For the analysis of thirty peer-reviewed studies and academic articles, a structured literature-based research design was used in the following areas:Historical document analysis, Handwritt--en text recognition, Optical character recognition, Document enhancement, Layout analysis, Bangla script recognition, Digital libraries and Digital humanities. The analysis reveals that the four interrelated barriers in the context of Bangladesh are: physical deterioration, complex script features, absence of line-level ground truth data and weak institutional coordination. The article suggests a framework of 'innovation and digitization', connectingcollection survey, informed consent, safe imaging, reversible image enhancement, layout-aware OCR/HTR, human checking, metadata production and sustainable repository governance. The results indicate that global platforms like Transkribus, Kraken and Calamari and transformer-based OCR systems offer valuable technical frameworks but cannot be directly applied to manuscript materials from Bangladesh, unless adequate local training data and paleographic knowledge are available, and ethical access policies are established. The novelty of this article is the link between the computational document analysis and manuscript challenges in Bangladesh as well as providing a practical implementation framework for Universities, Archives, Libraries and community custodians. The article maintains that AI can be used as a "human in the loop" tool, and not a replacement for the human scholar's interpretation.

Keywords:
manuscriptdigitization;Bangladesh;handwritten text recognition; optical character recognition; Bangla script; digital humanities; artificial intelligence; cultural heritage

1.Introduction

Among the most significant among the records of cultural memory in Bangladesh are manuscripts.They maintain religious commen--taries,poetictraditions, legal and administrative documents, genealogy, medical knowledge and folklore, education, farming knowledge and regional history. A manuscript is not just a receptacle for written information. It also records script practices, reading communities, circulationpatterns, ownership marks, marginal commentary and local intellectual history. The materials in these collections are housed in university libraries, public libraries, madarasas, mosques, temples, monasteries, shrines, family collections and local cultural institutions. However, many of these cultural treasures are still
unavailable due to fragility, lack of cataloguing, distribution, or lack of expertise to read the manuscripts.
Environmental and institutional conditions aggravate the preservation problem. Paper and ink can deteriorate quickly if they are subjected to heat, humidity, insects, flooding, or repeated handling or handling in areas with poor ventilation. In some cases the manuscript may perish long before it is catalogued formally, in family trunks or in a small local library. The digitization of the images, therefore, is not only a technical convenience but a preservation strategy. But it's not enough to just be scanned. The digital image of a folio may capture the visual surface of a folio but does not necessarily capture the textual content of the folio, nor does it necessarily make it readable, searchable, analyzable or reusable. The central issue is the decipherment, which

DOI: https://doi.org/10.5281/zenodo.20999737

involves making the difficult handwritten marks reliable and consistent in textual, metadata and scholarly form, and recording evidence of uncertainty. Calculational techniques open new avenues to this task. Manuscript images can be transformed to searchable knowledge resources by technologies such as historical document analysis, image enhancement, layout analysis, optical character recognition (OCR), handwritten text recognition(HTR), and machine learning, as well as by the use of digital repositories. HTR has already transformed archival scholarship in various international settings, enabling the efficient processing and search of large collections of manuscripts when adequate training corpora are available (Muehlberger et al., 2019; Nockels et al., 2022). Transformer-based OCR models like TrOCR have also shown the potential of integrating visual and language modelling for OCR-related tasks (Li et al., 2023). Despite this, historical documents are challenging due to the variability of handwriting and page layouts, and because of degraded surfaces and the difficulty in recognizing handwriting from them (Graves et al., 2009; Stokes, 2020).

Bangladesh is a unique case. Historical Bangla manuscript recognition is less well developed than OCR and HTR for Latin scripts: Bangla is a major world language. In the modern period of Bangla handwriting research, useful resources are available, such as Bangladesh-Isolated which has handwritten samples of Bangla basic characters, numerals and compound characters (Biswas et al., 2017). Convolutional neural network (CNN) has been used in Bangla handwritten character recognition (Chowdhury et al., 2019; Roy et al., 2018). Though these studies indicate the technical feasibility of Bangla character recognition, further studies are needed to recognize the manuscript. Examples of manuscripts contain connected writing, historical spelling, damaged folios, margins and illumination, multilingual vocabulary, compound letters, regional orthographic habits and scribal abbreviations.

This article thus explores the possibility of employing computational methods for manuscript decipherment in Bangladesh. The presentation is made as a research article, using a structured literature-based methodology, rather than as a model training experiment. It is meant to integrate the findings of the document analysis and digital humanities and develop an innovation framework specific to Bangladesh. The article suggests an integration of algorithms, institutions and communities as an essential pillar to successful digitization. While AI can assist with reading, retrieval and preservation, interpretation, correction and responsible access must involve human expertise.

2.Aim and Objectives

The purpose of this study is to explore the use of computational techniques to enhance manuscript decipherment, preservation and access in the context of Bangladesh, and to bring forward an innovative digitization plan that can be adopted by the local institutions and community-held collections in Bangladesh.

Specific objectives are: (a) to review major computational techniques used in manuscript digitizationand decipherment, (b) to understand the challenges in Bangladesh related to script complexity,manuscript deterioration, multilingual content and dispersed ownership of manuscripts, (c) to evaluate the relevance and limitations of existing OCR/HTR tools and Bangla handwriting recognition research, (d) to build a framework for a human-in-the-loop approach to digitization that merges technology with paleographic expertise and ethical governance, and (e) to identify outcomes, metrics and future research directions for Bangladesh-based manuscript projects. These goals fall under the cross-disciplinary fields of digital humanities, computer vision, and information science and archival studies. Character error rate, word error rate and model architecture are commonly used to assess systems in computer-science research. Humanities manuscript studies are centered on the meaning of the text, the material evidence, and historical interpretation. The strategy for digitization should be based on both types of knowledge, as it is Bangladesh oriented. The main argument of this research is that manuscripts should NOT be submitted to machines for automatic interpretation. Instead, computational approaches should assist scholars, archivists and communities to read, preserve and share manuscript heritage with greater accuracy and responsibility.

3.Methodology

The method of this study was structured

DOI: https://doi.org/10.5281/zenodo.20999737

literature based. The approach is suitable as Computational Manuscript Decipherment in Bangladesh is a nascent field, and there are only limited number of trained HTR models in the country and also few published case studies on Bangla historical manuscript collection in the country. The study, therefore, creates an evidence-based framework from synthesizing document image analysis, OCR, HTR, Bangla handwriting recognition, digital preservation and digital humanities.

The academic databases and search engines were used to identify relevant literature, such as Google Scholar, Scopus-indexed publisher pages, IEEE Xplore, ACM Digital Library, SpringerLink, Elsevier, Emerald and open-access journal repositories. The terms used for search were related to manuscript digitization, handwritten text recognition, optical character recognition, historical document analysis, document enhancement, layout analysis, Bangla character recognition, digital humanities, digital libraries and cultural heritage preservation.

Peer-reviewed journal articles, conference papers from established document-analysis conferences and other major, scholarly papers published between 1999 and 2025 were given priority. Older foundations studies were kept if they gave an introduction to concepts still much used in the field.

The selection was based on the four principles. The studies had to be directly related to computational document processing, manuscript digitization, Bangla script recognition or digital heritage infrastructure. Secondly, the source was required to offer a deciphering method, a set of data, a platform, a theory or an evaluation issue pertaining to the decipherment of his/her manuscript. Third, both global and evidence from South Asia/Bangla was used to avoid assuming that Bangladesh was an extension of European manuscript environments. Fourth, a study was rejected if it only concerned general image classification, modern office-document processing, commercial OCR, not related to historical or handwritten documents.

The last set included thirty key academic sources and supplemental standards or contextual sources. The themes which were used for analyzing the evidence were preservation and digital libraries, image enhancement, layout analysis, OCR and handwritten text recognition (HTR), Bangla handwriting recognition and governance of digital humanities. This analysis found the results were then utilized to build a practical digitization framework in Bangladesh. The article does not report original CER, WER or recognition accuracy since it is not based on a newly collected manuscript dataset. Rather, it offers an implementation model rooted in research that can be used as a guide for future empirical testing.

Table1 Literature Selection Strategy

Selection item	Application in this study	Purpose
Databases and platforms	GoogleScholar,IEEEXplore, ACMDigitalLibrary, SpringerLink Elsevier Emerald and open-access repositories	To capture both technical and humanities-oriented literature
Time range	1999-2025,witholdersources retained only when foundationa	To combine current AI methods with established digital-library theory
Inclusion focus	OCR, HTR, layout analysis document enhancement, Bangla recognition, digital preservation and digital humanities	To cover the full manuscript decipherment pipeline
Exclusion focus	Generalimageclassification, unrelated AI applications and non-scholarly promotional sources	To improve academic relevance and citation quality
Final evidence base	Thirty core academic sources plus contextual standards	To support a research-based framework for Bangladesh

DOI: https://doi.org/10.5281/zenodo.20999737

4.Literature Review

The computational manuscript decipherment literature can be organized into six related areas. The first one is digital preservation. Research in digital-library projects draws attention to the need to generate structured, discoverable and sustainable resources as products of digitization instead of just collections of images. Compared with the DELOS manifesto, which stated that digital libraries are complex systems, consisting of content, services, users and organisational structures (Candela et al., 2007), the discussion about digital libraries is now no longer confined to the technical aspects of information technology, but extends to topics such as content, services and users. Smith (1999) also noted the role that metadata can play in the management and retrieval of digital collections. These ideas imply that every digital image in a manuscript project should be associated with data concerning title, author, date, script, language, material, condition, provenance and ownership and access rights. Manuscript images are hard to find and to be reused, without metadata. Digital humanities scholarship builds on this, demonstrating that the nature of humanities inquiry is altered when conducted digitally. Digital humanities is not just about technical conversion, it's also about new mechanisms of scholarly access, public engagement and interpretation as claimed by Terras (2016). However, Nockels et al. (2024) also anticipate that HTR can be used as a tool to access the past at scale, and raises questions regarding accuracy, uncertainty, and the involvement of humans in the correction process. It is pertinent to note that the insights of these issues are relevant to Bangladesh, where the digitization of manuscripts should benefit research and public heritage, but should not be extractive in nature, allowing outsiders to demand control from the local custodians.

The second one is image enhancement. Historical documents may be stained, stained through, faded ink, unevenly lit, folded and have damaged surfaces. The multispectral imaging techniques have also been applied to recover text that is hard to read in normal light, demonstrated by the recovery of Archimedes Palimpsest (Easton et al., 2010). Multispectral processing has shown to be useful in revealing hidden and/or degraded text, but care must be taken during processing, as the selection of processing techniques can influence the interpretation (Giacometti et al., 2017). However, degraded/unfocused/binary image problems can be mitigated using deep-learning techniques like DE-GAN, which can enhance the image to improve the readability and downstream OCR performance (Souibgui & Kessentini, 2022). Humidity, insects and informal storage conditions can cause paper and ink deterioration, and enhancement is valuable for Bangladesh.

The third field is layout analysis. Very few manuscripts are beautifully composed as are printed pages today. They could have marginal notes, commentaries, sealing, colophons, decorative headings, uneven line spacing or folio numbers. Before OCR/HTR is applied, the layout analysis points out regions of the page, lines, columns and the reading order of the page. Antonacopoulos et al. (2009) made a contribution in the evaluation of the performance of document layout analysis; whereas, later datasets like DocBank have helped in the development of the layout-aware models (Dutta et al., 2020). Through the LayoutLMv2 and LayoutLMv3 models, Xu et al. (2021) and Huang et al. (2022) have shown that the integration of visual, textual and layout information in document understanding is beneficial. For Bangladeshi manuscripts, layout analysis is crucial since incorrectly separating damaged areas and the marginalia from the main text may cause a misrecognition result.

The fourth field is OCR and HTR. Printed text with a consistent font and a clean background and known language is ideal for traditional OCR. The handwritten manuscripts demand the HTR system which is capable of learning variable handwriting. Graves et al. (2009) incorporated sequence modelling techniques to recognize free handwriting with neural methods. Later, Shi et al. (2016) showed that end-to-end trainable networks are effective for image-based sequence recognition. For historic text recognition, Calamari has emerged as a deep-learning OCR framework that has been applied to the recognition of historical texts (Wick et al., 2020); for non-Latin scripts and humanities projects, Kraken is a flexible OCR recognizer (Kiessling, 2019). One of the more widely used platforms for HTR in archival contexts is Transkribus, and it has facilitated scholarship through training of models from

DOI: https://doi.org/10.5281/zenodo.20999737

manually corrected lines (Muehlberger et al., 2019).

Transformer-based document AI is the fifth field. The transformer encoder-decoder architecture and human labelled data can be used to fine-tune TrOCR. Donut presents an OCR-free document-understanding approach to directly generate structured outputs from images (Kang et al., 2022). These models indicate future possibilities of manuscript work since they can integrate the visual patterns with language information. Their utility for Bangladeshi manuscripts, however, requires training data, and language coverage and careful evaluation. If the model has been trained primarily on modern printed or Latin script data, it will probably not be able to work well without domain adaptation on older Bangla handwritten data.

The sixth field is recognition of Bangla and Indic script. Pal and Chaudhuri (2004) demonstrated that the Indian scripts are particularly hard to recognize because they have large character sets, modifiers, complex spatial arrangement of characters and compound characters. Bangladesh-Isolated is a large Bangla handwritten character dataset from which the basic characters, numerals and compound
characters collected from Bangladesh (Biswas et al., 2017). Chowdhury et al. (2019) and Roy et al. (2018) showed the application of deep learning on the problem of Bangla handwritten character recognition. Although these can form the basis of these studies, decipherment of the manuscripts requires line and page level data as well as individual characters. It also needs historical spelling, scribal variation and multilingual vocabulary to be represented in training data.

The literature indicates both the technical feasibility and the data-dependency of computational manuscript decipherment. While global systems are useful models, it is necessary to put in place models that are locally adapted to Bangladesh. What matters more than just the use of AI is the design of a workflow that links imaging, enhancement, recognition, expert correction, metadata and ethical governance.

Table 2 Comparison of Major OCR/HTR Platforms and Methods

Tool or method	Main strength	Main limitation	Relevancefor Bangladesh
Transkribus	Strong HTR workflow and model training for archival documents	Requires ground truth and may involve platform costs	Useful for pilot HTR training if local data are prepared
Kraken/eScriptorium	Open-source workflow, useful for historical and non-Latin scripts	Requires technical setup and trained users	Suitable for universities and local research teams
Calamari	Deep-learning OCR with ensemble possibilities	More suitable when segmentation and training data are available	Possible for printed or semi-printed Bengali materials
TrOCR and transformer models	Strong pre-training and fine-tuning potential	Data-hungry and not automatically reliable for historical Bangla	Promising for future model adaptation
Human verification	Preserves scholarly judgement and cultural context	Time-consuming and requires expertise	Essential for manuscripts with legal, religious or historical value

5.Findings and Analysis

The analysis produces five major findings. The first is that the digitization of manuscripts in Bangladesh should be considered as a conservation work of immediate concern. Many collections are housed in less-than-optimal archival conditions, and are vulnerable to environmental degradation. Digital imaging can minimize the handling of delicate folios for repeated examination and can provide backup

DOI: https://doi.org/10.5281/zenodo.20999737

copies for research and disaster risk management. But, there is a risk that improper scanning will harm manuscripts or affect the evidence. Thus, it is important to have collection assessment, owner consent and non-contact imaging and file-management standards in place prior to large-scale digitization.

The second thing that it is found is that computational recognition is possible but it is not directly transferable. Access to historical documents has been improved in several European and international projects with HTR systems (Muehlberger et al., 2019; Nockels et al., 2022). However, there are certain features that do not match corresponding generic models in Bangladeshi manuscripts. These are Bangla vowel signs, consonant conjuncts, matra structures, historical spellings, Persian-Arabic administrative terms, Sanskritic religious terms, and regional scribal styles. Bangladesh cannot solely depend on OCR engines from outside, therefore. It must have locally trained models and systematic evaluation on manuscript samples of various periods, genres and regions.

The third is about page-layout complexity. Marginal notes, glosses, interlinear additions, seals and uneven line spacing are common for manuscripts. When the layout analysis is unsuccessful, the recognition system can recognize the lines in the wrong order or combine text blocks that are unrelated. The layout-aware systems offer helpful directions

(Dutta et al., 2020; Huang et al., 2022; Xu et al., 2021) and they need to be provided with examples from the surrounding area that are annotated. Projects in Bangladesh should thus annotate text content as well as regions of pages, marginalia, page decorations, folio numbers, and damaged regions. The fourth finding is that it's not just software, it's data. Accurate image-text pairs are required for deep-learning models.

While there are many valuable existing Bangla datasets, most focus on single characters rather than Bangla manuscript lines or pages (Biswas et al., 2017; Roy et al., 2018). A manuscript dataset should contain full-page images, segmentation into lines, an accurate transcription, period and script information, and a label for uncertainty. Much of this data is going to take the collaboration of computer scientists, linguists, historians, paleographers, archivists and students. This is a joint ground truth development which is likely to be the biggest initial investment.

The fifth finding is that manuscript decipherment is human centered. Automated recognition will not result in final scholarship; it is probabilistic. A model may be mistaken for similar letter forms and be inaccurate in the normalization of historical spellings or produce plausible but false readings. This is particularly hazardous for legal, religious and historical documents. Human reviewers need to edit the output, keep on record the uncertain readings and record editorial decision. If there is a private manuscript or culturally sensitive manuscript, community custodians should be engaged. There is a need to enhance the role of shared stewardship through digitization, not relinquish it from the hands of those who own it to those who are far removed.

Table 3 Key Computational Methods and Bangladesh-SpecificApplications

Method	**Purpose**	**Applicationin Bangladesh**	**Main challenge**
High-resolution imaging	Createsdurabledigital surrogates	Digitizes fragile folios in libraries, shrines, family collections and religious institutions	Safe handling and colour consistency
Image enhancement	Improvesvisibilityof degraded pages	Reduces stains, shadows, bleed-through and faded ink problems	Avoidingover-processing
Layout analysis	Detects text regions and reading order	Separates main text, marginalia, seals and folio numbers	Irregular page structures

DOI: https://doi.org/10.5281/zenodo.20999737

OCR/HTR	Convertsimagesinto machine-readable text	SupportsBangla, Sanskritic,Persian-influenced and mixed-script transcription	Lack of line-level ground truth
Human verification	Correctsautomated transcription	Usespaleographers, historians, linguists and local experts	Timeand specialist availability
Metadata and repository	Enablesdiscoveryand preservation	Buildssearchable catalogueswithaccess policies	Institutional coordination

6.Suggested Innovation-Oriented Digitization Framework

The proposed structure includes 6 interdependent phases. The first stage is collection survey and consent. The project team should identify manuscript holders, document ownership, evaluate physical condition, document risks, and secure permission to image and access the manuscripts in the future. This step is necessary as many manuscripts in Bangladesh are not kept in national archives but with the families or the religious institutions. Ethical digitization starts even before the camera is turned on.

Stage two is safe imaging. Photographs or scans should be used to preserve manuscripts. Shooting should be done with the same resolution, color calibration, folio numbers, lighting and file naming. Master file should not be compressed or highly compressed and access copy can be optimized for Web use. Recommendations should be made for the handling of such brittle paper, bindings and marginal annotations. The third stage is computational enhancement. Images should be straight, properly lit, free from stains, bleed-through and low contrast. The enhancement should be reversible and documented. The project should contain an original image, enhanced image and a recognition-ready image.

This means the original page is preserved in its evidential value without compromising computational processing.

The fourth stage is layout analysis and OCR/HTR. Page regions, lines, marginal notes and non-textual elements should be annotated prior to recognition. Printed or semi-printed Bengali texts can be processed by OCR, but handwritten manuscripts need to be hand transcribed using HTR. Local training data should be used to enhance the recognition accuracy, and the confidence scores should be stored to priorities the words that are not recognized with high confidence to be reviewed by experts.

The fifth stage is human verification and scholarly enrichment. Reviewers are responsible for correcting recognition output, maintaining unclear readings, identifying abbreviations and making editorial decisions. Representing the corrected transcriptions back to the training set should form a feedback loop between scholarship and machine learning. This is where innovation turns from being machine-driven to human-in-the-loop.

The last stage is repository publication and long-term management. Manuscripts should be archived with metadata, persistent identifiers, access and preservation backups. The design of the repository should assist in searching, browsing, referring by citation, comparing and future correction of a transcript. Ownership, sacred value, copyright and community preferences should be acknowledged within access policies.

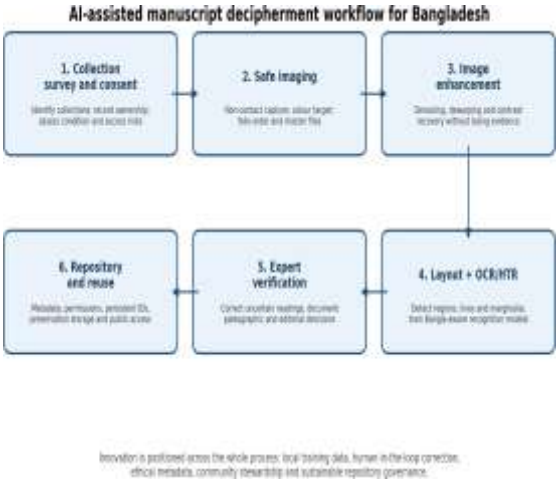

Figure 1. Innovation-oriented framework for AI-assisted manuscript digitization and decipherment in Bangladesh. Source: Developed by the author from the reviewed literature.

DOI: https://doi.org/10.5281/zenodo.20999737

7.Expected Outcomes

Several outcomes are possible for a project that involves computational manuscripts, specifically focused on Bangladesh. The first is the preservation. Digital surrogates can decrease the handling of sensitive originals and serve as backup copies in the event of a fire, flood or physical deterioration. The second is enhanced scholarly access. Useful searchable transcriptions and structured metadata can render manuscripts available to researchers of literature, history, linguistics, religious studies, law, anthropology and information science.

The third outcome is to develop a Bangla manuscript data set. Systematic image and verified transcriptions can be used for future research in OCR, HTR, language modelling, script classification, historical linguistics and digital humanities. Capacity building is the fourthoutcome.Imaging,metadata, paleography, annotation and AI-assisted transcription can be taught to students and researchers. This will help create a new interdisciplinary skill set in Bangladesh.

Public engagement is the fifth outcome. Selected manuscripts may be displayed online, and can be translated and released in educational materials. This can link younger generation to the textual heritage of Bangladesh. The sixth outcome is international co-operation. The creation of a high quality manuscript digitization programmed can link local collections with wider digital humanities and document-analysis networks in Bangladesh, enhancing visibility of its heritage and ensuring ethical governance.

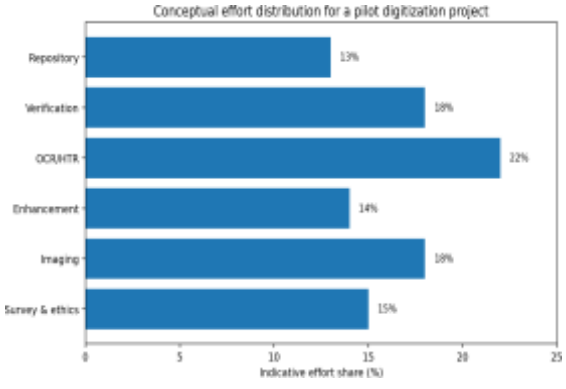

Figure 2. Conceptual effort distribution for a pilot manuscript digitization project. Values are planning weights developed from the proposed framework, not survey results.

8.Originality and Value

The novelty of this article is the synthesis of computational manuscript decipherment which is unique in Bangladesh. A few studies exist on Bangla handwriting in the isolated character recognition domain and most studies in the global handwriting recognition domain have beentested on European manuscript collections. This article seeks to fill that gap by examining the impact of the complexity of local scripts and cultural ownership, as well as institutional realities, on the viability of AI-supported manuscript activities in Bangladesh.

Academic value is interdisciplinary. On the digital humanities front, the study demonstrates the potential of computational tools in enhancing access and maintaining human interpretation. For computer science, it marks Bangla manuscript decipherment as an under-resourced recognition domain, but at the same time culturally important. It provides a useful approach to metadata, library and archival repository governance, and planning for preservation for library and archival studies. It offers a phased approach that can start with small pilot initiatives and not be a national scale investment for heritage institutions.

Equally vital is the social value. The preservation of many manuscripts is due to the efforts of families, religious teachers and local custodians who safeguarded them over many generations. These custodians should be seen as partners in a responsible digitization programmed. Ethical access, local capacity and long-term preservation should be as important as algorithmic performance as measures of innovation.

9.Discussion

The results show the promising and challen-
-ging character of manuscripts decipherment through computer in Bangladesh. Technology is available but local data infrastructure is not at scale. This makes a strategic decision. Either Bangladesh can continue to use OCR tools which are available in general but cannot be used effectively for historical manuscripts or invest in a gradual and steady development of locally created data, model adaptation and expert correction. The second method is harder to do, but more successful in getting a more consistent result.

One of the most important conflict is speed and accuracy. Often many pages can be processed quickly by an automated system, but manuscript scholarship will involve careful

DOI: https://doi.org/10.5281/zenodo.20999737

reading. A recognition model could generate fluent, grammatically correct text with errors which are masked. Such errors in religious, legal or historical manuscripts can impact interpretation. The suggested process is thus one of minimizing risk through the use of confidence scores and human verification. Low-confidence regions should be investigated first, and uncertain readings should be kept and not assumed to be normal. Standardization is another concern. When the image formats, metadata fields, file names and transcription conventions differ between institutions, integration into the future will be challenging. There is a need of a common guideline for manuscript digitization in Bangladesh. This guideline should include image resolution, color targets, file naming, metadata, transcription principles, levels of access and preservation storage. Where one wishes to deliver an image, interoperability standards like IIIF can help this process, and where one needs a structured transcription, where more advanced scholarly editions are required, TEI-style principles can guide this process.

The funding is also crucial. This can take resources in terms of cameras, storage, trained staff, annotation platforms, maintenance and many more. The project doesn't have to start at the national level, though. Feasibility can be proved with a small, well-designed pilot of 300-500 pages from a variety of collections, and it can help to train staff and secure bigger funding. If the large-scale scanning project is not addressing the right needs, it can result in thousands of unstructured images, whereas a pilot can produce knowledge infrastructure that is reusable.

Last but not least, community trust needs to be embeddedwithin the system. While digitization can make something more accessible, it can also raise fear of control, misuse, and ownership. From the start consent forms and access policies, along with benefit sharing and the acknowledgement of custodians. A manuscript digitization programmed can fail even if technologically advanced, if it disregards the values of the community.

10.Implementation Roadmap and Evaluation Metrics

The roadmap should start with a pilot project for the practical. The pilot can choose 300 - 500 pages from two or three different collections that are in contrast, for example, a literary manuscript, a religious commentary and an administrative or land-related document. This diversity would enable the project team to experiment with various page settings, writing techniques, word structures and layout treatments. The pilot should start with a collection register, with basic provenance data, physical condition assessment and ownership documentation.

The second phase should be used to prepare ground truth. It is important that trained reviewers manually transcribe a limited number of pages. Historical spellings should not be normalized and forced to modern spelling unless an extra normalized field is established. Each manuscript image ideally should be aligned with a verified line-by-line text. This will take time, but will be essential as the accuracy of the HTR relies on the quality of training data (Muehlberger et al., 2019; Nockels et al., 2022).

In phase 3, the tools should be compared. A baseline model could be trained in Transcribes, Kraken/eScriptorium or any other appropriate platform. Another experiment can be the recognition using transformer, e.g., TrOCR, with fine-tuning from local data (Li et al., 2023). The accuracy is not the only factor to consider while comparing tools. It should also take into account the cost, transparency, ease of training, language support, ownership of data, sustainability of Bangladeshi institutions.

The evaluation should take quantitative and qualitative measures into account. Quantitative measures can be character error rate, word error rate, layout-region accuracy, line-order accuracy and completeness of metadata. Historic spelling should be maintained, marginalia should be indicated, ambiguous readings should be noted and culturally inappropriate materials should be safeguarded by means of a qualitative evaluation. Only if a project yields readable images, reliable metadata, searchable transcription, documentation of uncertainty, and a long-term preservation plan could it be considered successful.

Table 4 Proposed Evaluation Metrics for a Bangladesh Manuscript Pilot

DOI: https://doi.org/10.5281/zenodo.20999737

Evaluation area	Metric	Target use	Reason
Recognition accuracy	Character error rate and word error rate	CompareOCR/HTR models	Measurestranscription reliability
Layout quality	Correct text-region and line-order detection	Assesspage segmentation	Prevents misreading of marginaliaand commentary
Metadata quality	Completeness of title, date, script, language, condition and ownership fields	Evaluatecatalogue usefulness	Supports discovery and citation
Human verification	Percentage of uncertain words reviewed	Control scholarly risk	Reduces false confidence
Access governance	Consent record and access level assigned	Protectcustodiansand sensitive materials	Ensuresethical digitization
Sustainability	Backup, version control andrepository maintenance plan	Supportlong-term preservation	Prevents digital loss

11.Limitations and Future Research

There are limitations to this article. It does not report OCR/HTR original recognition accuracy as it doesn't train a new OCR or HTR model on Bangladeshi manuscript images. It is also not a comprehensive listing of the collections of manuscript materials in Bangladesh. It would need institutional co-operation and field work to produce such an inventory. The article should, therefore, be interpreted as a research based structured approach and not as a fully-fledged technical experiment.

Further study is needed to establish a pilot collection of manuscript images and line-level transcriptions of the Bangla language. The list of the periods, genres, scripts, conditions and regions should be different. These baseline HTR models should then be trained and tested with CER, WER and expert review. For the local application, comparative tests of Transkribus, Kraken/eScriptorium, Calamari, and transformer-based systems would be helpful in determining which systems are most feasible.

The study should also be extended to explore language modelling for historical Bangla, named-entity recognition for persons and places and automated metadata extraction and ethical repository design. In universities and archives, trainings should be done in the field of manuscript imaging, manuscript paleography and AI-assisted transcription. In addition, the questions of ownership, copyright, sacred value, restricted access and benefit sharing need to be addressed in future research.

12.Conclusion

Bangladesh has vast manuscript heritage, which is disorganized, fragile and unused. Computational techniques offer a good opportunity to maintain and interpret this heritage, but they need to be used judiciously. The digitization process cannot be limited to merely scanning, and AI cannot replace subject-matter expertise. Safe imaging, image enhancement, layout analysis, OCR/HTR, human verification, metadata and ethical governance of the repository are all elements that should be combined in a responsible approach.

The review reveals that the existing international research on HTR, OCR and document AI provides valuable building blocks; however, there is a need to adapt these tools for use in Bangladesh. Complexity of Bangla script, traditional manuscript forms, multilingual vocabulary, and ownership of the manuscript's present unique challenges. While existing Bangla handwriting databases are good sources, but the ground-truth data for deciphering manuscript is required in the form of line-level and page-level data

The proposed innovation-oriented framework provides a feasible roadmap to the future research and institutional action. It links technology with community stewardship and preservation. Combined efforts between universities, archives, libraries, computer scientists, historians, linguists and local custodians will lead to computational manus--cript decipherment that will enhance

DOI: https://doi.org/10.5281/zenodo.20999737

access to Bangladesh's cultural memory and help worldwide research studies on under-resourced scripts and digital humanities.

Conflict of interest statement: The author declares no conflict of interest. Funding statement: No specific funding is declared for this manuscript.

Data availability statement: No new dataset was generated for this article. The proposed workflow and evaluation metrics are developed from the reviewed literature.

Ethics statement: This article is based on published literature and does not involve human participants, personal data or newly collected manuscript images.

Declarations

References

Antonacopoulos,A.,Bridson, D., Papadopoulos, C., & Pletschacher, S. (2009). A realistic dataset for performance evaluation of document layout analysis. In Proceedings of the 10th International Conference on Document Analysis and Recognition (pp. 296-300). IEEE.

Banerjee, A., Biswas, S., Lladós, J., & Pal, U. (2023). SwinDocSegmenter: An end-to-end unified domain adaptive transformer for document instance segmentation. In Document Analysis and Recognition - ICDAR 2023 (pp. 307-325). Springer.

Biswas, M., Islam, R., Shom, G. K., Shopon, M., Mohammed, N., Momen, S., & Abedin, A. (2017). BanglaLekha-Isolated: A multi-purpose comprehensive dataset of handwritten Bangla isolated characters. Data in Brief, 12, 103-107.

Bluche, T. (2016). Joint line segmentation and transcription for end-to-end handwritten paragraph recognition. In Advances in Neural Information Processing Systems, 29, 838-846.

Candela, L., Castelli, D., Pagano, P., Thanos, C., Ioannidis, Y., Koutrika, G., Ross, S., Schek, H. J., & Schuldt, H. (2007). Setting the foundations of digital libraries: The DELOS manifesto. D-Lib Magazine, 13(3/4).

Chowdhury, R. R., Rahman, M. S., & Hossain, M. A. (2019). Bangla handwritten character recognition using convolutional neural network. In Proceedings of the International Conference on Bangla Speech and Language Processing (pp. 1-5). IEEE.

Das, N., Sarkar, R., Basu, S., Kundu, M., Nasipuri, M., & Basu, D. K. (2012). A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application. Applied Soft Computing, 12(5), 1592-1606.

Dutta, A., Gupta, A., & Zisserman, A. (2020). DocBank: A benchmark dataset for document layout analysis. In Proceedings of the 28th International Conference on Computational Linguistics (pp. 949-960).

Easton, R. L., Knox, K. T., & Christens-Barry, W. A. (2010). Multispectral imaging of the Archimedes Palimpsest. In Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop (pp. 1-8). IEEE.

Giacometti, A., Campagnolo, A., MacDonald, L., Mahony, S., Robson, S., Weyrich, T., Terras, M., & Gibson, A. (2017). The value of critical destruction: Evaluating multispectral image processing methods for the analysis of primary historical texts. Digital Scholarship in the Humanities, 32(1), 101-122.

Graves, A., Liwicki, M., Fernandez, S., Bertolami, R., Bunke, H., & Schmidhuber, J. (2009). A novel connectionist system for unconstrained handwriting recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(5), 855-868.

Huang, Y., Lv, T., Cui, L., Lu, Y., & Wei, F. (2022). LayoutLMv3: Pre-training for document AI with unified text and image masking. In Proceedings of the 30th ACM International Conference on Multimedia (pp. 4083-4091).

Jaume, G., Ekenel, H. K., & Thiran, J. P. (2019). FUNSD: A dataset for form understanding in noisy scanned documents. In Proceedings of the International Conference on Document Analysis and Recognition Workshops (pp. 1-6). IEEE.

Kang, G., Hong, T. I., Yim, M., Nam, J., Park, J., Yim, J., Hwang, W., Yun, S., Han, D., & Park,

S. (2022). Donut: Document understanding transformer without OCR. In European Conference on Computer Vision (pp. 498-517). Springer.

Kiessling, B. (2019). Kraken: An universal text recognizer for the humanities. In Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage (pp. 13-18). ACM.

Li, M., Lv, T., Chen, J., Cui, L., Lu, Y., Florencio, D., Zhang, C., Li, Z., & Wei, F. (2023). TrOCR: Transformer-based optical character recognition with pre-trained models. Proceedings of the AAAI Conference on

DOI: https://doi.org/10.5281/zenodo.20999737

Artificial Intelligence, 37(11), 13094-13102.
Muehlberger, G., Seaward, L., Terras, M., Ares Oliveira, S., Bosch, V., Bryan, M., Colutto, S., Dejean, H., Diem, M., Fiel, S., Gatos, B., Greinoecker, A., Grüning, T., Hackl, G., Haukkovaara, V., Heyer, G., Hirvonen, L., Hodel, T., Jokinen, M., & Kahle, P. (2019). Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study. Journal of Documentation, 75(5), 954-976.
Nockels, J., Gooding, P., Ames, S., & Terras, M. (2022). Understanding the application of handwritten text recognition technology in heritage contexts: A systematic review of Transkribus in published research. Archival Science, 22, 367-392.
Nockels, J., Gooding, P., & Terras, M. (2024). The implications of handwritten text recognition for accessing the past at scale. Journal of Documentation, 80(4), 1-20.
Pal, U., & Chaudhuri, B. B. (2004). Indian script character recognition: A survey. Pattern Recognition, 37(9), 1887-1899.
documents and automatic text recognition. Journal of Data Mining and Digital Humanities, Special Issue on Historical Documents and Automatic Text Recognition, 1-15.
Roy, S., Das, N., Kundu, M., & Nasipuri, M. (2018). Handwritten isolated Bangla compound character recognition: A new benchmark using a novel deep learning approach. Pattern Recognition Letters, 90, 1-8.
Shi, B., Bai, X., & Yao, C. (2016). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(11), 2298-2304.
Smith, T. R. (1999). The MetaNet digital library: Metadata for managing and searching digital collections. D-Lib Magazine, 5(12).
Souibgui, M. A., & Kessentini, Y. (2022). DE-GAN: A conditional generative adversarial network for document enhancement. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(3), 1180-1191.
Stokes, P. A. (2020). The humanist and the machine: Reflections on handwritten text recognition for manuscript studies. Digital Medievalist, 13, 1-18.
Terras, M. (2016). Crowdsourcing in the digital humanities. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), A new companion to digital humanities (pp. 420-438). Wiley Blackwell.
Wick, C., Reul, C., & Puppe, F. (2020). Calamari: A high-performance Tensorflow-based deep learning package for optical character recognition. Digital Humanities Quarterly, 14(2).
Xu, Y., Xu, Y., Lv, T., Cui, L., Wei, F., Wang, G., Lu, Y., Florencio, D., Zhang, C., Che, W., Zhang, M., & Zhou, L. (2021). LayoutLMv2: Multi-modal pre-training for visually-rich document understanding. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (pp. 2579-2591).
Zhang, P., Li, C., Qiao, L., Cheng, Z., Pu, S., Niu, Y., & Wu, F. (2021). VSR: A unified framework for document layout analysis combining vision, semantics and relations. In Document Analysis and Recognition - ICDAR 2021 (pp. 115-130). Springer.

DOI: https://doi.org/10.5281/zenodo.20999737