Modi Lipi Translation Using OCR and Deep Learning

Piyush R. Kulkarni[1], Gayatri R. Gangode[2], Tejal R. Jadhav[3], Sanyogeeta S. Gavhane[4], Divya S. Barkale[5]
[1,2,34,5]*Department of Computer Engineering, Guru Gobind Singh College of Engineering & Research Centre, Nashik, Maharashtra, India*

***Abstract*—Modi Lipi served for centuries as the administrative hand of Marathi and today guards a vast archival, cultural, and legal legacy that few modern readers can unlock, since fluency in the script is now rare. Recognition technologies optical character recognition (OCR), handwritten text recognition, and neural transliteration promise automated access to this heritage. Twenty studies are examined here, spanning Modi character recognition, Modi-to-Devanagari transliteration, OCR for low-resource scripts, historical document processing, and the deep-learning machinery of sequence recognition: CNN and CRNN encoders, Connectionist Temporal Classification (CTC), encoder–decoder learning, attention, Transformers, and augmentation. The surveyed literature follows a clear arc isolated character classification, then line-level sequence recognition, and most recently direct image-to-Devanagari transliteration through vision-language models yet persistent obstacles include thin training sets, writer variability, damaged pages, ligated characters, curved lines, and contextual ambiguity. Notably, the strongest reported Modi transliteration reaches 88.14%-character agreement but only 61% exact-word agreement, evidence that errors compound across words. From these findings an integrated methodology is assembled document preprocessing, line segmentation, deep visual feature extraction, sequence recognition, transliteration, contextual normalization, and confidence estimation and the unresolved problems that a user-facing Modi translator must master on authentic historical documents are identified.**

***Index Terms*—CNN; CRNN; CTC; deep learning; Devanagari; historical manuscripts; Modi Lipi; OCR; optical character recognition; transliteration.**

I. INTRODUCTION

Archival documents preserve first-hand evidence about governance, society, culture, and regional history. Across Maharashtra and neighbouring regions, a considerable share of that record was penned in Modi Lipi, the administrative hand of Marathi for several centuries. With Devanagari now the standard script for Marathi, historical Modi material sits behind an accessibility barrier for contemporary readers.

Two obstacles block automatic processing of such material. Damaged pages present faded inks, stains, uneven backgrounds, warped paper, and fractured strokes, while scribes differ widely in style. Beyond recognition, the recovered text must also be rendered in Devanagari without losing linguistic or contextual nuance. Working on handwritten Modi-to-Devanagari transliteration with a CRNN pipeline built on Calamari, Joseph and George [15] obtained 88.14% agreement at character level but merely 61% at whole-word level, demonstrating at once the tractability and the difficulty of the task.

Early Modi research concentrated on individual characters. Varpe and Sakhare [2] benchmarked ResNet9, ResNet50, and InceptionNet V3 on handwritten Modi characters, establishing deep CNNs as a dependable character-level baseline. The centre of gravity has since moved toward transliteration: Kausadikar et al. [1] contributed MoDeTrans 2,043 Modi document images aligned with Devanagari transliterations and MoScNet, a distilled vision-language model for direct image-to-text conversion.

Work on comparable low-resource scripts offers transferable insight. VOLTAGE [3] couples contrastive learning with automatic glyph feature harvesting, while CURIO [4] confronts curved lines in historical manuscripts through rectification, geometry-matched synthesis, and a compact CNN-Transformer decoder trained with CTC. Modi inherits both problems label poverty and hostile imaging so these designs carry over directly.

This review therefore aims beyond a catalogue: it seeks to weld prior findings into an engineering blueprint for a usable Modi Lipi Translator. Section II lists the objectives. Section III documents the search

and selection procedure. Sections IV and V present and compare the surveyed work. Findings, gaps, and the problem statement appear in Sections VI to VIII. Section IX details the proposed pipeline, Section X the expected contributions, and Section XI the open challenges; Section XII concludes.

1.1 Objectives

1. To survey current work on Modi Lipi character recognition, handwritten OCR, and Modi-to-Devanagari transliteration.
2. To assess CNN, CRNN, CTC, Transformer, and related deep-learning techniques for recognising Modi manuscripts.
3. To expose the difficulties created by thin datasets, writer variability, ligated characters, and damaged historical pages.
4. To isolate the weaknesses of existing Modi OCR and transliteration systems and to pin down a precise research gap.
5. To formulate an integrated methodology for a practical translator that turns manuscript images into readable Devanagari text.

II. REVIEW METHODOLOGY AND STUDY SELECTION

A. Search Strategy

A structured search spanning 2000–2026 was run across Google Scholar, Scopus, IEEE Xplore, the ACM Digital Library, SpringerLink, and arXiv. Script terms and task terms were combined in one Boolean string: ("Modi" OR "Modi Lipi" OR "Modi script") AND (OCR OR "character recognition" OR "handwritten text recognition" OR transliteration OR "deep learning"). Backward and forward citation chaining was then applied to every located Modi study and to the foundational OCR papers it cites, so that relevant work missed by keyword matching would still be captured.

B. Selection Criteria

A study qualified for inclusion if it (i) addressed recognition or transliteration of Modi script, (ii) presented OCR for low-resource or historical scripts, (iii) introduced sequence-recognition methods for text, or (iv) proposed augmentation or practical OCR tooling. English-language publication and full-text access were required. Where the same work appeared in multiple copies several foundations were duplicated in the collected material it was counted once. Editorial items, contributions without methodological substance, and sources unverifiable in any archival database were dropped.

C. Quality Assessment and Categorization

Retained studies were judged on venue credibility, methodological clarity, and reproducibility. They were then grouped into four families: (1) direct Modi recognition and transliteration; (2) OCR for low-resource and historical Indic scripts; (3) sequence-recognition foundations; and (4) augmentation and practical OCR frameworks. The grouping keeps evidence about Modi itself distinct from techniques that merely transfer to it.

D. Selection Outcome

Fig. 1 charts the screening procedure in PRISMA style. Following de-duplication, title and abstract screening, and full-text assessment, 20 studies were retained for their blend of direct Modi relevance and technical foundation.

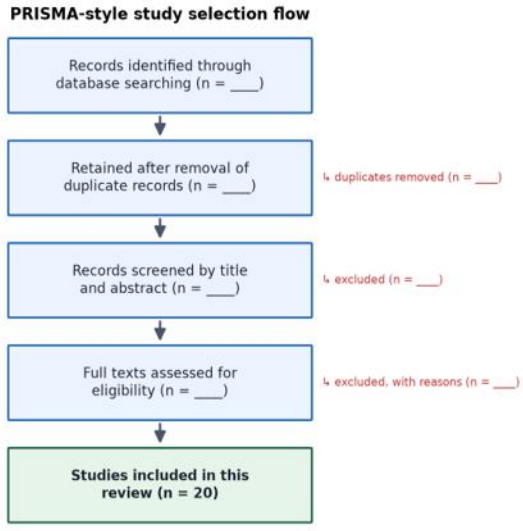

Fig. 1. PRISMA-style flow of the study selection procedure.

III. LITERATURE REVIEW

Fig. 2 charts the trajectory of the surveyed work from isolated characters, through line sequences, to whole-document transliteration. The subsections below follow the same order, opening with studies that tackle Modi script itself.

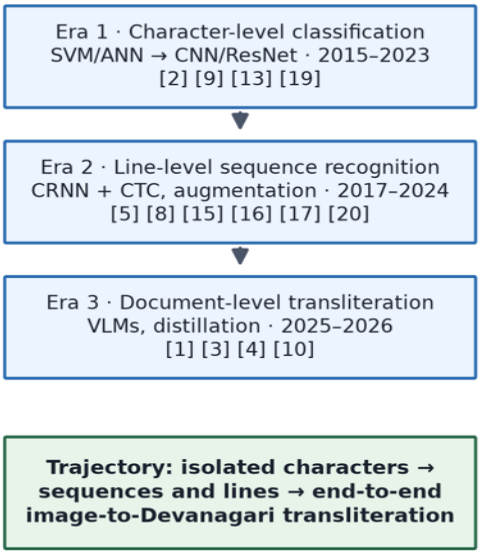

Fig. 2. Evolution of OCR research relevant to Modi Lipi translation.

A. Direct Recognition and Transliteration of Modi Script

Varpe and Sakhare [2] put ResNet9, ResNet50, and InceptionNet V3 to work on handwritten Modi characters and found that deep visual features separate the classes reliably, fixing the character-level benchmark for the field. Character classifiers alone, however, cannot translate documents: running text yields variable-length sequences laced with joins and context-dependent ambiguity.

Joseph and George [15] crossed from recognition to transliteration, driving a CRNN/Calamari pipeline over handwritten Modi lines and reporting 88.14%-character agreement against 61% exact-word agreement. The gap signals that per-character errors compound once predictions are stitched into words.

Kausadikar et al. [1] proposed MoScNet, a vision-language model that reads a Modi page image and emits Devanagari directly. Their MoDeTrans corpus pairs 2,043 document images with transliterations, and knowledge distillation yields a student that beats its teacher with 163× fewer parameters. Folding recognition and transliteration into one image-to-text mapping, MoScNet narrows the traditional gulf between OCR and transliteration an appealing property for deployment.

Kolte and Gupta [13] surveyed Modi character recognition, pinning script complexity, writer variability, and resource poverty as the dominant bottlenecks. Read together, these studies describe a field pivoting from isolated character classification toward sequence- and document-level translation, with robustness and corpus scale still unresolved.

B. Low-Resource and Historical Script OCR

For ultra low-resource scripts, Sharma et al. [3] introduced VOLTAGE, which pairs contrastive learning with automatic glyph harvesting. For Modi the moral is clear: stronger representations can substitute for scarce manual annotation when verified samples are hard to assemble.

Gunda et al. [4] attacked curved text lines in historical manuscripts with CURIO: curvature profiles are extracted, lines are straightened, and synthetic lines are generated along manuscript-derived trajectories to close the synthetic-to-real gap; a slim CNN-Transformer with sparse attention and CTC then transcribes. Modi pages often bend and distort, which makes CURIO a natural methodological anchor.

Kumar et al. [5] probed Hindi OCR robustness with CRNN and ResNet50 under degradation-aware augmentation and recorded lower character and word error rates for CRNN in their setting. Their conclusion that representative degraded data matters more than pristine samples transfers straight to Modi.

Shi et al. [10] charted the digitization of endangered scripts and argued for the preservation value of digital tooling. For Modi, this positions a translator as cultural infrastructure, not merely a predictive model.

C. Sequence Recognition Foundations

Shi, Bai, and Yao [16] reframed text reading as sequence transcription rather than per-character classification: convolutions harvest visual features, recurrent layers model context, and CTC collapses network frames into a variable-length string. Handwritten Modi suits this design because character boundaries never have to be segmented explicitly, and the CTC objective of Graves et al. [8] was built precisely for inputs and targets that lack character-level alignment. Hochreiter and Schmidhuber [14] supplied the LSTM machinery that models long-range dependencies and that OCR pipelines have used ever since.

On the vision side, residual connections [9] and depth studies [19] made deep CNN encoders reliable; such backbones already serve Modi character classification [2]. Simard et al. [18] added that large, distorted training sets sharpen document-analysis CNNs an

argument for aggressive augmentation whenever a script shows heavy visual variation.

D. Sequence-to-Sequence Models and Attention

Encoder–decoder sequence learning [11], soft attention for alignment [12], and the self-attention Transformer [7] together underwrite today's vision-language transliteration, including MoScNet [1], and provide the theory for any learned Modi-to-Devanagari mapping.

E. Augmentation and Practical OCR Frameworks

Shonenkov et al. [17] showed with StackMix and HandWritten Blots that generation-based augmentation lifts handwriting recognition in a ResNet-BiLSTM-CTC stack. Analogous tricks stroke variation, ink breaks, noise, scaling can multiply scarce Modi data. Wick et al. [20] released Calamari, an open-source CNN-LSTM-CTC toolkit that underpins the Modi transliteration of Joseph and George [15]. IndicTrans2 [6], finally, demonstrates what large Indian-language corpora and Transformer modelling can deliver for language-aware design and evaluation, even though Modi-to-Devanagari is script conversion rather than translation.

IV. COMPARATIVE ANALYSIS

Table I profiles the studies. "Relevance" is descriptive, not a ranking: Direct marks work on Modi itself, High marks transferable OCR or manuscript methods, and Supporting marks foundations invoked to justify a pipeline component. Numbers appear only where the source published a headline metric within the scope of this review.

Table I. Key findings from research papers

Ref.	Study	Year	Method	Task, data, and key reported result	Rel.
[1]	Kausadikar et al.	2025	MoScNet (VLM + knowledge distillation)	MoDeTrans: 2,043 Modi document images paired with Devanagari transliterations; first direct image-to-Devanagari transliteration; distilled student surpasses the teacher with 163× fewer parameters	Direct
[2]	Varpe and Sakhare	2024	ResNet9, ResNet50, InceptionNet V3	Handwritten Modi character images; deep CNNs give strong character-level recognition	Direct
[3]	Sharma et al.	2024	VOLTAGE: contrastive learning + auto glyph features	Ultra low-resource script OCR; representation learning reduces dependence on large annotated sets	High
[4]	Gunda et al.	2026	CURIO: rectification + CNN-Transformer + CTC	Historical manuscripts with curved lines; curvature-aligned synthesis narrows the synthetic-to-real gap	High
[5]	Kumar et al.	2026	CRNN and ResNet50 with augmentation	Devanagari data with degraded samples; CRNN gives lower CER and WER than ResNet50 in the reported setup	High
[6]	Gala et al.	2023	IndicTrans2 (Transformer)	22 scheduled Indian languages; open high-quality translation resources	Supp.
[7]	Vaswani et al.	2017	Transformer	General sequence modelling; self-attention replaces recurrence	Supp.
[8]	Graves et al.	2006	CTC	Unsegmented sequence labelling; objective for unaligned input–target pairs	High
[9]	He et al.	2016	ResNet	Image classification; residual connections enable very deep CNNs	High
[10]	Shi et al.	2026	Heritage review	Endangered-script digitization; consolidates digitalization efforts and preservation impact	High
[11]	Sutskever et al.	2014	Seq2seq	Sequence transduction; encoder–decoder learning of input–output mappings	Supp.
[12]	Bahdanau et al.	2015	Attention	Neural machine translation; soft alignment improves long-sequence mapping	Supp.

[13]	Kolte and Gupta	2024	Survey	Modi character recognition; consolidates script complexity, writer variability, resource scarcity	Direct
[14]	Hochreiter and Schmidhuber	1997	LSTM	Sequence modelling; gated recurrence captures long-term dependencies	Supp.
[15]	Joseph and George	2024	CRNN + Calamari + CTC	Handwritten Modi lines paired with Devanagari; 88.14% character and 61% exact-word matching	Direct
[16]	Shi, Bai, and Yao	2017	CRNN	Scene and document text; end-to-end sequence recognition without character segmentation	High
[17]	Shonenkov et al.	2022	StackMix + HandWritten Blots	Handwritten text-line augmentation (ResNet-BiLSTM-CTC); improved robustness	High
[18]	Simard et al.	2003	CNN + distortion	Visual document analysis; large distorted training sets improve document CNNs	High
[19]	Simonyan and Zisserman	2015	VGG	Image classification; depth improves CNN visual representations	Supp.
[20]	Wick et al.	2020	Calamari (CNN-LSTM-CTC)	Historical print OCR; open-source, high-performance OCR framework	High

Three patterns stand out. The directly relevant line of work climbs from character classification [2] through line-level transliteration [15] to document-level translation [1]. Corpora remain tiny: MoDeTrans [1] tops out at 2,043 images, orders of magnitude under mainstream OCR sets. And the techniques most worth adopting CTC sequence decoding [8], [16], curvature-aware rectification [4], degradation-aware augmentation [5], [17] are exactly those that historical Modi pages demand.

V. SYNTHESIS OF FINDINGS

A. Evolution from Characters to Documents

The literature moves visibly from per-character labelling toward sequence OCR and image-to-text translation. CNN classifiers [2] prove that Modi glyph features are learnable; CRNN and CTC systems [8], [16] drop character segmentation altogether; and the newest Modi work [1] translates a full-page image into Devanagari in one step.

B. Dataset Scarcity Is a Central Constraint

Modi is data-poor by OCR standards, so corpus engineering dominates the primary studies. MoDeTrans offers 2,043 paired items [1], Joseph and George [15] assembled their own line set, and synthesis plus augmentation measurably offsets the shortfall [3], [4], [5], [17].

C. Historical Image Quality Is a Core Problem

Clean character crops breed overconfidence: real manuscripts bring curvature, fading, noise, uneven lighting, broken strokes, and cluttered backgrounds. CURIO [4] treats curvature and the synthesis gap head-on; the Hindi study [5] shows what degradation-aware augmentation buys. Preprocessing and robust training therefore belong inside the recognition system, not in front of it.

D. Character Accuracy Does Not Equal Translation Accuracy

Joseph and George's 88.14%-character agreement collapsing to 61% exact-word agreement [15] is the sharpest warning in the survey: solid characters still yield shaky words. Evaluation must therefore report OCR accuracy and transliteration quality as separate, parallel numbers.

VI. RESEARCH GAP

The surveyed studies reveal interrelated gaps that motivate the proposed Modi Lipi Translator.

- Paired data scarcity: Modi corpora shrink beside mainstream OCR sets, restricting generalization across hands and capture conditions.
- Characters versus documents: much work stops at single characters, while deployment demands words, lines, pages, and multi-page volumes.
- Historical damage: fading, stains, curvature, skew, and background clutter erode accuracy on genuine archival sheets.
- Writer variability: glyph shapes drift across hands and historical eras, requiring diverse samples and augmentation.

- Contextual transliteration: right characters can still spell wrong words; sequence context and language knowledge must correct them.
- Unified workflow: upload, preprocessing, OCR, transliteration, confidence scoring, and human verification belong in one tool.
- Honest evaluation: CER, WER, exact-word accuracy, and transliteration quality should be reported side by side.

VII. PROBLEM STATEMENT

Large holdings of Marathi heritage are trapped in Modi Lipi, a script modern readers can neither read nor search as knowledge of it recedes. Manual transcription into Devanagari is slow and expert-bound. Deep learning has proven that Modi characters can be recognized and that CRNN systems can transliterate them, yet dependable document-level recognition remains blocked by thin labelled data, writer variability, ligated and cursive forms, faded or noisy pages, page curvature, and contextual ambiguity. What is needed is an integrated pipeline that accepts a manuscript image, preprocesses and segments it, recognizes the handwritten Modi, transliterates it into Devanagari, and returns text with confidence scores for expert verification.

VIII. PROPOSED METHODOLOGY

The literature argues for a modular pipeline (Fig. 3), whose virtue is that every component can be trained and scored in isolation before assembly.

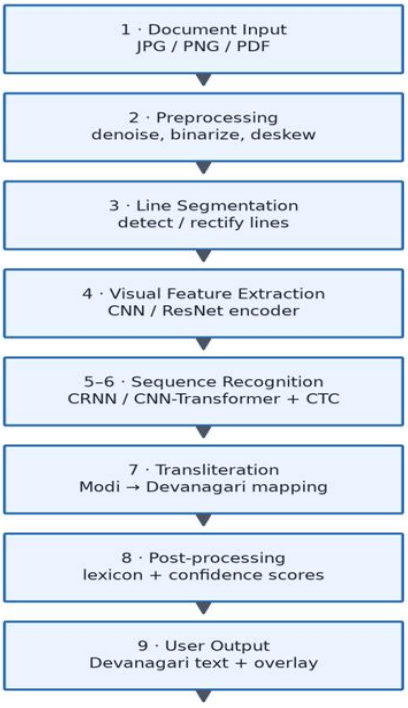

Fig. 3. Proposed modular pipeline of the Modi Lipi Translator.

Table II. Stages of the proposed modular pipeline

Stage	Module	Purpose	Implementation notes
1	Document Input	Accept a scanned page or photograph of a Modi manuscript	JPG / PNG / PDF upload
2	Preprocessing	Suppress noise while preserving strokes	Grayscale, contrast enhancement, denoising, adaptive binarization, deskew, resize
3	Line Segmentation	Detect, crop, and where needed rectify individual text lines	Projection profiles or a learned line detector; CURIO-style rectification for curved lines [4]
4	Visual Feature Extraction	Encode line images into deep visual features	CNN / ResNet backbone [9] or a lightweight visual encoder
5	Sequence Recognition	Transcribe feature sequences into Modi tokens	CRNN or CNN-Transformer with CTC decoding [8], [16]
6	Modi Text Output	Emit the predicted Modi character/token sequence	Raw model output retained for audit
7	Transliteration	Convert the Modi sequence into Devanagari text	Learned seq2seq model [11], [12], explicit mapping, or a hybrid
8	Post-processing	Normalize language and score confidence	Historical Marathi/Modi lexicon, Unicode normalization, implausibility filtering
9	User Output	Present readable Devanagari with verification support	Devanagari text, optional annotated image, low-confidence words flagged for review

A. Image Preprocessing

Inputs may be scans or photographs. Preprocessing should suppress noise while protecting strokes: grayscale conversion, contrast enhancement, denoising, adaptive binarization where suitable, deskew, and resizing. Curved lines may warrant a CURIO-style rectification stage [4]; faint but genuine strokes must survive the process.

B. Line Segmentation

Recognising at line level is preferable to character segmentation for a first build. Projection profiles or a learned detector can locate horizontal text bands, and polygon- or scribble-guided extraction following [4] can straighten strongly curved lines.

C. Recognition Model

CRNN with CTC is the natural baseline convolutional features, sequence modelling, and variable-length decoding [8], [16] with residual blocks in the encoder [9]. Later iterations can swap the recurrent stage for a slim CNN-Transformer in the spirit of CURIO [4] or benchmark against a MoScNet-style vision-language model [1].

D. Transliteration Layer

Conversion to Devanagari can rest on a learned transliteration model, an explicit mapping table, or a hybrid that maps deterministically where correspondences are reliable and corrects contextually where they are not. Attention and encoder–decoder training [11], [12] justify the learned variant.

E. Post-processing and Confidence

Post-processing can fold in a historical Marathi/Modi lexicon, Unicode normalization, implausibility filtering, and confidence scoring. Raw model output must remain visible so predictions can be checked against the page, and low-confidence words should be routed to human reviewers instead of being silently edited.

F. Evaluation

Scoring should operate at several levels CER for characters, WER for words, and exact-word accuracy for whole items supplemented by transliteration accuracy and judgement from a competent Marathi/Modi reader. Results should be split between clean and degraded images wherever possible.

IX. EXPECTED CONTRIBUTIONS

- A preprocessing pipeline tailored to scanned or photographed Modi manuscripts.
- A CRNN and CTC baseline for line-level handwritten Modi OCR.
- A Modi-to-Devanagari transliteration module with contextual correction.
- A small, carefully verified paired dataset or evaluation set, subject to archival and copyright clearance.
- Robustness studies built on realistic augmentation and degraded historical imagery.
- Confidence-aware output that exposes uncertain predictions to the user.
- A lightweight interface for upload, extraction, and Devanagari export.
- A side-by-side comparison of character-level, sequence-level, and end-to-end approaches.

X. OPEN CHALLENGES AND FUTURE DIRECTIONS

Several directions remain open beyond the proposed pipeline.

- Vision- and large-language models: distilled VLMs already transliterate page images [1]; fine-tuning strategies and hallucination control for Modi deserve systematic study.
- Benchmarks and datasets: shared splits, writer-independent partitions, and standardized CER, WER, and transliteration metrics are missing; MoDeTrans [1] starts the job, but corpora must grow.
- Document-level evaluation: today's numbers are character- or line-scoped; multi-page layout, orientation, and cross-line context are untested.
- Human-in-the-loop transcription: confidence routing and active learning could make archival-scale transcription economically feasible.
- Heritage deployment: searchable archives need metadata, indexing, and faceted search on top of transliterated text [10].
- Low-resource learning: contrastive pretraining [3], manuscript-matched synthesis [4], and handwriting augmentation [5], [17] should be combined and ablated for Modi.

XI. CONCLUSION

Twenty studies were distilled here around the construction of a Modi Lipi Translator. Across OCR research the direction of travel is unmistakable: from isolated characters to end-to-end sequence recognition and image-to-text transliteration. Modi-specific evidence confirms that handwritten characters can be

recognized and converted to Devanagari and MoScNet now pairs a direct vision-language approach with the MoDeTrans corpus.

Imported methods carry their value across: contrastive learning stretches scarce labels,curvature-aware synthesis repairs historical imaging,CRNN with CTC delivers segmentation-free recognition, and augmentation hardens models against writer variation and damage. The open prize is a pipeline that runs on real Modi documents and is judged on full lines and pages via CER, WER, exact-word accuracy, transliteration quality, and expert verification with confidence-aware output fit for archival use. Delivering it would serve both the digitization of Marathi heritage and practical heritage technology.

REFERENCES

[1]. H. Kausadikar, T. Kale, O. Susladkar, and S. Mittal, "Historic scripts to modern vision: A novel dataset and a VLM framework for transliteration of Modi script to Devanagari," arXiv preprint arXiv:2503.13060, 2025.

[2]. K. Varpe and S. Sakhare, "Reading Modi Lipi: A deep learning journey in character recognition," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 9, no. 1, pp. 75–83, 2024.

[3]. P. Sharma, P. Goyal, V. Sharma, and N. Goyal, "VOLTAGE: A versatile contrastive learning-based OCR methodology for ultra-low-resource scripts through auto glyph feature extraction," in Proc. European Chapter of the Association for Computational Linguistics (EACL), 2024, pp. 881–899.

[4]. S. M. Gunda, T. Ghosh, S. S. Sandral, and R. K. Sarvadevabhatla, "CURIO: Curvature-aligned and efficient OCR for low-resource historical manuscripts," in Proc. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2026, pp. 2011–2021.

[5]. S. Kumar, S. Sourabh, J. Pati, and S. Kumar, "Enhancing Hindi OCR robustness with CRNN-ResNet50: A data augmentation approach on the Devanagari dataset," Scientific Reports, 2026.

[6]. J. Gala et al., "IndicTrans2: Towards high-quality and accessible machine translation models for all 22 scheduled Indian languages," arXiv preprint arXiv:2305.16311, 2023.

[7]. A. Vaswani et al., "Attention is all you need," in Advances in Neural Information Processing Systems (NeurIPS), 2017.

[8]. A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks," in Proc. International Conference on Machine Learning (ICML), 2006, pp. 369–376.

[9]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[10]. J.-Q. Shi, F. Tsung, and K. Zhang, "A review of recent efforts in digitalization and application of endangered scripts," npj Heritage Science, 2026.

[11]. I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in Advances in Neural Information Processing Systems (NeurIPS), 2014.

[12]. D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in Proc. International Conference on Learning Representations (ICLR), 2015.

[13]. S. Kolte and B. Gupta, "Bridging heritage and technology: A review of Modi script character recognition," International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 2024.

[14]. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[15]. S. Joseph and J. George, "Machine transliteration of handwritten Modi script to Devanagari using deep neural networks," International Journal of Computing, vol. 23, no. 2, pp. 219–225, 2024.

[16]. B. Shi, X. Bai, and C. Yao, "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, vol. 39, no. 11, pp. 2298–2304, 2017.

[17]. A. Shonenkov, D. Karachev, M. Novopoltsev, M. Potanin, D. Dimitrov, and A. Chertok, "Handwritten text generation and strikethrough characters augmentation," arXiv preprint arXiv:2112.07395, 2022.

[18]. P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks

applied to visual document analysis," in Proc. International Conference on Document Analysis and Recognition (ICDAR), 2003.

[19]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. International Conference on Learning Representations (ICLR), 2015.

[20]. C. Wick, C. Reul, and F. Puppe, "Calamari: A high-performance TensorFlow-based deep learning package for optical character recognition," in Proc. International Conference on Document Analysis and Recognition (ICDAR), vol. 1, 2018, pp. 25–28.