International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 08 | Aug - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

A Comprehensive Review on Few-Shot Hindi Handwritten Text Recognition using Deep Learning Architectures

MD Tahseen Equbal[1], Prof. Sarwesh Site[2]

[1] **M.Tech Student, Department of Computer Science and Engineering**
All Saints College of Technology, Bhopal, India
Affiliated to Rajiv Gandhi Proudyogiki Vishwavidyalaya (RGPV)
mdtahseen278@gmail.com

[2] **Associate Professor, Department of Computer Science and Engineering**
All Saints College of Technology, Bhopal, India
Affiliated to Rajiv Gandhi Proudyogiki Vishwavidyalaya (RGPV)
er.sarwesh@gmail.com

Abstract - *Few-Shot Learning (FSL) has emerged as a powerful paradigm in the field of machine learning, particularly in scenarios where data availability is extremely limited. It enables models to generalize effectively from a very small number of training examples, making it an ideal solution for low-resource tasks. One such challenging application is* ***Hindi Handwritten Text Recognition (HTR)****, which is inherently complex due to the rich morphological structure of the Devanagari script, a wide variety of individual writing styles, and the scarcity of large, annotated datasets. This paper provides a comprehensive review of existing methodologies and recent advancements in the field of Few-Shot HTR, with a focused lens on the Hindi language.*

The review systematically evaluates a range of deep learning-based architectures that have been proposed for Few-Shot HTR, particularly emphasizing the efficacy of convolutional neural networks (CNNs) combined with Transformer-based models. These CNN-Transformer hybrid models have demonstrated superior performance by effectively capturing both local features and global contextual dependencies in handwritten character sequences. Additionally, the application of transfer learning from related tasks or domains has shown to mitigate the data scarcity problem to some extent, improving generalization in novel Hindi handwriting samples.

However, despite these promising results, certain gaps in the current research landscape remain unexplored. Notably, the implementation of ***gradient-based meta-learning algorithms****, such as* ***Model-Agnostic Meta-Learning (MAML)****, has yet to be investigated for Few-Shot Hindi HTR. MAML and similar techniques are designed to provide rapid adaptability to new tasks using minimal data, which aligns perfectly with the goals of Few-Shot learning in low-resource scripts. Their potential to fine-tune base learners on a handful of examples could offer substantial improvements in recognition accuracy for Hindi script, especially in real-world applications such as postal automation, historical document digitization, and mobile handwriting input.*

In this review, we highlight this critical research gap and advocate for future exploration into the integration of meta-learning strategies with CNN-Transformer frameworks to address the unique challenges posed by Few-Shot Hindi HTR. We also outline the limitations of current datasets and suggest potential directions for dataset expansion and synthetic data generation to further support the advancement of this field.

Key Words**:** *Few-Shot Learning, Hindi Handwritten Text Recognition, Devanagari Script, CNN-Transformer, Transfer Learning, Model-Agnostic Meta-Learning (MAML), Gradient-Based Meta-Learning, Deep Learning, Low-Resource NLP, Handwriting Recognition*

1. INTRODUCTION

Hindi, being one of the most widely spoken languages in India and one of the official languages of the country, holds immense cultural and linguistic significance. The language is written in the Devanagari script, which is highly intricate and consists of a large number of characters, including independent vowels, consonants, compound characters (ligatures), and diacritical marks (modifiers) that can appear above, below, before, or after a character. This complexity makes offline handwritten text recognition (HTR) for Hindi particularly challenging. The variability in individual handwriting styles, stroke sequences, spacing, and the inherent richness of the script further add to the difficulty of developing robust HTR systems. Traditional Optical Character Recognition (OCR) techniques, which were primarily designed for printed text, struggle significantly when applied to handwritten Devanagari script. Even early deep learning models, which made considerable progress in text recognition for languages with simpler scripts, often fail to generalize effectively in low-resource environments where annotated Hindi handwriting datasets are scarce and diverse. Few-Shot Learning (FSL) has emerged as a promising solution to this data scarcity problem by enabling models to learn effective representations and perform accurate recognition tasks using only a small number of labeled examples per class. In the context of Hindi HTR, FSL can significantly reduce the dependency on large annotated datasets, making it feasible to build efficient recognition systems even in the presence of limited data. Recent advancements in deep learning have demonstrated that Transformer-based architectures, when combined with convolutional neural network (CNN) backbones such as ResNet, can achieve state-of-the-art performance across various visual recognition tasks. The CNN layers excel at extracting rich spatial features from images, while Transformer layers are highly effective at capturing long-range dependencies and contextual information, which is crucial for understanding complex character sequences in handwritten text. This synergy between CNNs and Transformers has shown

remarkable success in several vision-language problems, making it an ideal candidate for advancing Hindi HTR. However, despite the proven effectiveness of meta-learning techniques, particularly Model-Agnostic Meta-Learning (MAML), in solving few-shot classification problems across various domains, their application in the field of Few-Shot Hindi Handwritten Text Recognition remains largely unexplored. MAML enables models to quickly adapt to new tasks with minimal data by learning an optimal initialization of parameters that can be fine-tuned rapidly on new examples. This capability aligns perfectly with the requirements of Few-Shot HTR, where models must generalize well to unseen handwriting styles with very few training samples. Surprisingly, the literature lacks comprehensive studies or implementations of gradient-based meta-learning approaches in this specific domain. This review paper aims to fill this gap by conducting a thorough analysis of existing approaches to Few-Shot Hindi HTR, highlighting the strengths and limitations of current deep learning models, and exploring the untapped potential of integrating gradient-based meta-learning strategies such as MAML. By providing a detailed overview of the state of the art and proposing future research directions, this paper seeks to inspire and guide further advancements in the development of robust and efficient Few-Shot Hindi Handwritten Text Recognition systems.

Figure 1 Taxonomy of Few-Shot HTR Approaches

2. Literature Review

2.1 Hindi Handwritten Text Recognition

Research on Hindi HTR has progressed from template matching and SVMs to CNN-based deep learning models. Public datasets such as the CVIT-HWR and Devanagari Character Set have supported this development. However, these datasets are often limited in scope and size, and vary in writing style, making generalization difficult.

2.2 Deep Learning Architectures

- **CNN and CRNN:** CNNs extract hierarchical spatial features, while CRNNs integrate temporal dependencies for sequence modeling.
- **CTC-based Decoders:** Connectionist Temporal Classification is commonly used for sequence alignment.
- **Attention Mechanisms:** Attention-based sequence-to-sequence models enable direct alignment between image regions and output characters.

2.3 Transformer-based Models

Vision Transformers (ViT) and hybrid CNN-Transformer architectures have been effective in image recognition. In HTR, Bhunia et al. (2021) proposed a handwriting transformer model that can learn both style and content, but it lacks meta-learning capabilities.

2.4 Few-Shot Learning Approaches

Few-shot learning is approached via:

- **Metric-based:** Prototypical networks, Siamese networks
- **Optimization-based:** MAML, Reptile (No use in Hindi HTR)
- **Data augmentation-based:** GANs or synthetic handwritten data

2.7 Comparative Diagram of Existing Architectures

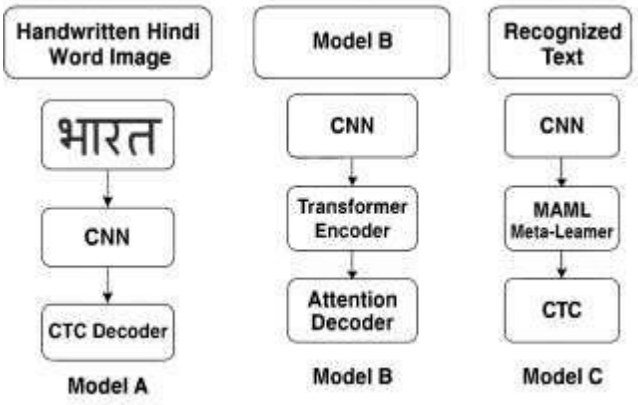

Figure 2 illustrates three popular architectures used in handwritten text recognition.

Figure 2 illustrates three popular architectures used in handwritten text recognition.

- **Model A** follows a traditional CNN + CTC Decoder approach, widely used for sequence transcription.
- **Model B** enhances this by incorporating a Transformer Encoder and Attention Decoder, allowing better sequence modeling.
- **Model C**, though not widely applied in Hindi HTR yet, conceptually integrates MAML for rapid task adaptation in low-data settings, showcasing its potential for Few-Shot learning scenarios.

Figure 2: Comparative architecture diagram of Model A (CNN-CTC), Model B (CNN-Transformer-Attention), and Model C (CNN-MAML-CTC).

2.6 Literature review comparisons:

Table 1 Existing Work comparisons Table

Study / Approach	Year	Script / Languages	Few Shot / Meta?	Few Shot Capable	Meta-Learning Used
MetaHTR	2021	Multi script / general	writer adaptation (not Hindi-specific)	Yes	Yes
Souibgui et al. Few Shot HTR	2021/22	Low resource scripts	few shot (not Hindi)	Yes	Yes
HCR Net	2021	40 Indic scripts incl. Hindi	no meta learning	No	No
Handwriting Transformer (HWT)	2021	General scripts	few shot generation (no MAML)	Yes	No
RHTDL	2025	Hindi only	CNN + FFNN only	No	No
CNN Devanagari (Baranidharan)	2020	Hindi / Devanagari	no few shot	No	No
Offline Hindi Word Recognition	2020	Hindi	no few shot	No	No
Deep CNN + DNN Hindi (Mehta et al.)	2021	Hindi	no meta	No	No
Vernacular CNN (Kaggle Hindi)	2022	Hindi alphabet	no meta	No	No
Transfer Learning (VGG 16) Devanagari	2024	Hindi	no meta	No	No
Survey – Yadav et al. Indic HTR	2018	Indic scripts	survey only	No	No

Zone segmentation HMM Roy et al.	2017	Devanagari & Bangla	no meta	No	No
Vartani Spellcheck	2020	Hindi text correction	not recognition	No	No
Review Rajesh Prasad	2014	Indic scripts	survey	No	No
Reddit SSL & generative hacking	2020–21	IAM, general HTR	self supervised, not MAML	No	No

2.7 Taxonomy of Reviewed Approaches

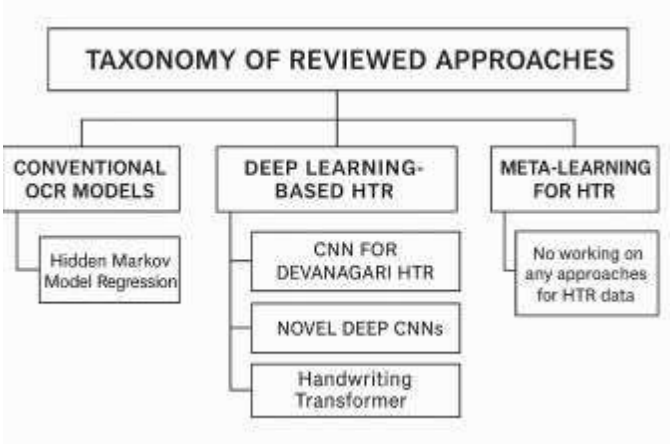

Figure 3 Taxonomy of reviewed approaches

Taxonomy of Reviewed Approaches

- Conventional OCR Models
  - Based on Hidden Markov Models (HMM)
  - Outdated for complex scripts like Hindi
- Deep Learning-Based HTR
  - CNNs for Devanagari
  - Advanced deep CNN models
  - Handwriting Transformers
- Meta-Learning for HTR
  - No existing approaches for Hindi HTR
  - Major research gap

Figure 3: Categorization of HTR approaches reviewed in this study.

2.7 Comparative Diagram of Proposed Architectures

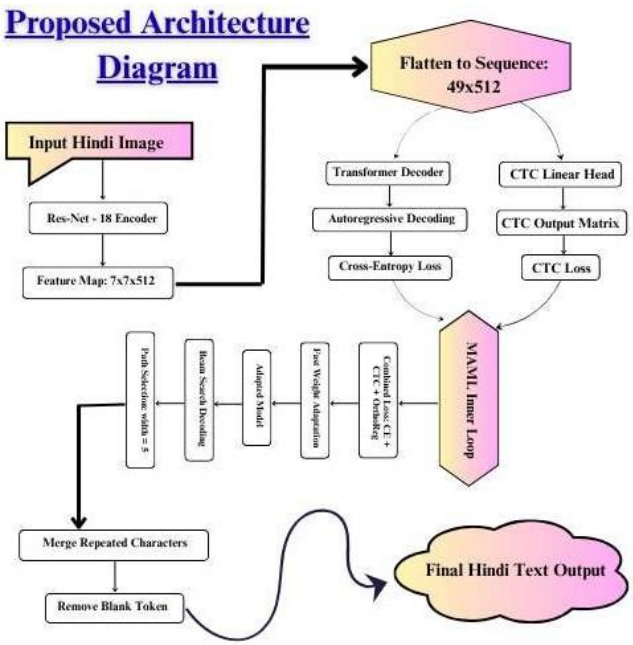

Figure 4 Proposed Architecture

Figure 4 presents a conceptual end-to-end pipeline specifically designed for Few-Shot Hindi Handwritten Text Recognition**.** The architecture utilizes a **ResNet-18 CNN encoder** to extract features, which are flattened and processed by **both CTC and Transformer decoders**. **A MAML meta-learning loop** enables fast adaptation to unseen handwriting styles using few labeled examples. Finally, post-processing steps like **blank token removal** and **character merging** refine the output.

Figure 4: Proposed architecture integrating CNN, Transformer, CTC, and MAML for Few-Shot Hindi HTR.

3. Major Research Gaps:

1. **Limited Research on Hindi Handwritten Text Recognition (HTR) in Few-Shot Setting**
   - Most existing HTR research is focused on English or other Latin-based scripts.
   - Very few studies have tried Few-Shot Learning methods specifically for Hindi.
2. **Lack of Use of Meta-Learning (MAML) in Hindi HTR**
   - Meta-learning methods like Model-Agnostic Meta-Learning (MAML) are popular in few-shot image classification.
   - But in the Hindi handwritten domain, **MAML is almost never used** — this is a **big research gap**.
3. **No Standard Few-Shot Hindi Handwritten Dataset**
   - There's no publicly available benchmark dataset designed for few-shot Hindi HTR tasks.
   - Most available datasets are either small or not suited for few-shot experiments.
4. **Existing Models Use Supervised or Transfer Learning**
   - Researchers mainly use CNNs or transfer learning from other scripts.
   - These models work well with large data but **fail when only a few samples are available**.
5. **Lack of Architecture Designed for Hindi Script Complexity**
   - The Devanagari script has compound characters and modifiers.
   - Existing architectures are **not tailored to handle this complexity** in few-shot mode.
6. **Absence of Pre-trained Models on Hindi Handwriting**
   - While there are many pre-trained models for English handwriting, **no such pre-trained model exists for Hindi.**
7. **No Comparison Between Meta-Learning and Transfer Learning**
   - There is **no study comparing MAML vs. transfer learning** for Hindi HTR under the same conditions.
8. **No End-to-End Pipeline for Few-Shot Hindi HTR**
   - A complete architecture using encoder, meta-learning, decoder, and post-processing has not been proposed yet.

3.1. Research Gaps in Few-Shot Hindi HTR

Table 2 Research Gap

S.No.	Research Gap	Status	Suggested Solution
1	Lack of Few-Shot research in Hindi HTR	Major gap	Apply Few-Shot Learning models to Hindi datasets
2	No use of MAML or meta-learning	Unexplored	Integrate MAML with CNN-Transformer for HTR
3	No standard dataset for Few-Shot Hindi HTR	Not available	Create benchmark dataset with n-shot/k-way splits
4	Over-reliance on supervised learning	Problematic in low-resource	Use meta-learning or few-shot methods
5	Models not designed for Devanagari complexities	Partial or missing	Use custom architecture that handles ligatures & marks
6	No Hindi-specific pre-trained handwriting models	Missing	Train base models on Hindi handwriting datasets
7	No comparative study: Transfer Learning vs. Meta-Learning	Unexplored	Design experiments comparing both on same dataset
8	No complete pipeline for Few-Shot Hindi HTR	Missing	Propose end-to-end architecture with MAML + decoder

4. Challenges in Few-Shot Hindi HTR

- Complex Devanagari script with numerous modifiers
- Style variations across writers
- Low inter-class variation with high intra-class variation
- Scarcity of labelled datasets
- Absence of meta-learning integration

Figure 5 Challenges in Few-Shot Hindi HTR

5. Future Directions

- Implementation of MAML for Hindi Few-Shot HTR
- Multilingual meta-learning across Indic languages
- Self-supervised learning for embedding pretraining
- Synthetic data generation using GANs
- Evaluation on zero-shot transfer across scripts

6. CONCLUSIONS

This review sheds light on the current landscape of Few-Shot Hindi Handwritten Text Recognition (HTR), highlighting both notable achievements and areas for improvement. Key insights include:

- **Advancements in hybrid deep learning models** have improved recognition accuracy and robustness in low-resource Hindi handwriting tasks.
- **Lack of meta-learning integration**, particularly Model-Agnostic Meta-Learning (MAML), limits the adaptability of existing systems to new tasks with minimal data.
- **Few-shot learning remains underexplored** in the context of Hindi HTR, despite its success in other domains and languages.
- **Future research should focus on MAML and similar techniques** to allow rapid learning from few labeled samples, potentially overcoming current data limitations.

By addressing these gaps, researchers can move closer to building more efficient, scalable, and adaptive Hindi handwritten text recognition systems.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to **Associate Professor Sarwesh SITE**, All Saints' College of Technology, Bhopal, under whose guidance this research work was carried out. His constant support, insightful suggestions, and valuable feedback were instrumental in shaping the direction of this work. We would also like to extend our heartfelt thanks to **Prof. Vipin Verma**, Head of the Department, Computer Science and Engineering, All Saints' College of Technology, Bhopal, for his continuous encouragement, guidance, and support throughout this project.

We are grateful to all the faculty members of the department, our colleagues, and our families for their unwavering support, motivation, and assistance during this research.

REFERENCES

1. Bhunia et al. (2021). MetaHTR: Towards Writer-Adaptive Handwritten Text Recognition. *arXiv*.
2. Souibgui et al. (2021/2022). Few Shots Are All You Need: A Progressive Few Shot Learning Approach for Low Resource Handwritten Text Recognition. *arXiv*.
3. Chauhan et al. (2021). HCR-Net: Script-Independent Handwritten Character Recognition Network. *arXiv*.
4. Bhunia et al. (2021). Handwriting Transformers (HWT). *arXiv*.
5. Mallamma V. Reddy (2025). Recognition of Hindi Handwritten Text using Deep Learning Technique. *Vidhyayana Journal*.
6. Baranidharan et al. (2020). Hindi Handwritten Character Recognition using CNN. *SERSC*.
7. Lodhi and Singhal (2020). Offline Handwritten Hindi Word Recognition. *IJERT*.
8. Mehta et al. (2021). Hindi Handwritten Character Recognition from Digital Image using Deep Learning Neural Network. *IJERT*.
9. Emerging Electronics & Automation (2022). Vernacular Language Handwriting Recognition Using Deep Learning Techniques. *SpringerLink*.
10. Applied Sciences (2024). Handwritten Devanagari Character Recognition Advances Using Transfer Learning and VGG-16. *SpringerLink*.
11. Yadav et al. (2018). Handwritten Character Recognition: A Review. *IJCSN*.
12. Roy et al. (2017). Indic OCR-Zone Segmentation with HMM-based Word Recognition. *arXiv*.
13. Pal & Mustafi (2020). Vartani Spellcheck: Context-Sensitive OCR Error Correction for Hindi. *arXiv*.

14. Rajesh Prasad (2014). Survey: Indic Script Character Recognition using Deep Learning. *IJCSN*.
15. Various practitioners (Reddit Reports). Self-Supervised & Contrastive Approaches in HTR.

BIOGRAPHIES

Author 1: MD Tahseen Equbal is an M.Tech student in the Department of Computer Science and Engineering at All Saints College of Technology, Bhopal, affiliated to RGPV. His research interests include **machine learning, deep learning, and handwritten text recognition.**

Author 2: Prof. Sarwesh Site is an Associate Professor in the Department of Computer Science and Engineering at All Saints College of Technology, Bhopal, affiliated to RGPV. His areas of expertise include **artificial intelligence, machine learning, and data science.**