Insights into Deep Learning Paradigms for Text Recognition in Indic Scripts

Akant Mohan Gupta[1], Harmohan Sharma[2]
[1]*Research Scholar, Department of Computer Science, Punjabi University, Patiala.*
[2]*Associate Professor, Department of Computer Science,Multani Mal Modi College, Patiala.*

***Abstract*—**This paper presents a structured study of deep learning (DL) architectures applied to multilingual optical character recognition (OCR), with a particular emphasis on Indic scripts. It analyzes supervised (e.g., CNNs, RNNs, LSTMs), unsupervised (e.g., autoencoders), and hybrid models (e.g., CRNNs, GANs) used in tasks like document analysis, handwriting recognition, and natural language processing. Despite significant progress, challenges persist in recognizing Indic scripts due to their structural complexity and limited annotated datasets. A comparative analysis of existing approaches is provided, along with a categorized summary of DL models and their effectiveness. The paper concludes by outlining key research gaps and offering insights to guide the development of efficient, language-agnostic OCR systems.

Keywords*—Text Identification, Indic Scripts, DL Models*

I. INTRODUCTION

Deep Learning (DL), a prominent field within machine learning, has transformed the way complex patterns and data representations are understood and processed. Motivated by the structure of biological neural networks, deep learning leverages multiple interconnected layers within artificial neural networks (ANNs) to automatically learn abstract and informative patterns from unprocessed data. This contrasts with conventional machine learning techniques, which often depend on manually engineered features for effective performance. DL models have the ability to automatically discover patterns, making them especially effective in handling intricate tasks involving large, unstructured datasets.

The rapid evolution of computational capabilities, coupled with the accessibility of open-source platforms like TensorFlow, PyTorch, and Keras, has significantly accelerated the adoption of deep learning across various domains [1]. These include visual computing, speech analysis, document interpretation, and machine translation. Core architectures—such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) units, and combined hybrid frameworks have shown high effectiveness in classification, pattern identification, and recognition-related applications [2]. Their effectiveness spans both supervised learning scenarios with labeled data and unsupervised contexts where data lacks explicit annotations.

Deep Learning (DL), a vital area within machine learning, is modeled after the structure and functioning of the human brain and has emerged as a key driver in the development of modern artificial intelligence [3]. Its ability to manage complex and high-dimensional information enables DL systems to automatically extract significant features from raw data with minimal human effort. These models are designed to scale effectively with large volumes of data and adapt well across diverse tasks. Due to their high accuracy and strong generalization performance, DL techniques have achieved a great success in different domains, including image analysis, speech interpretation, and language understanding [4].

The widespread adoption of DL stems from its early success in tasks like handwritten text recognition, where models such as Convolutional Neural Networks (CNNs) demonstrated near-human performance. Unlike traditional machine learning, which depends heavily on feature engineering and prior domain knowledge, DL leverages multi-layered artificial neural networks to extract and learn data representations autonomously [5]. Its growing popularity is also driven by robust frameworks like TensorFlow, Keras, Torch, CNTK, and Caffe, which facilitate scalable and efficient implementation.

This review focuses on the application of DL techniques to optical character recognition (OCR), particularly for Indic scripts, which pose unique challenges due to their rich morphology, diverse character sets, and limited annotated data. The paper systematically examines the evolution of DL models in multilingual text recognition, compares their performance across various scripts, and highlights key research gaps. The aim is to support future development of robust, scalable, and script-agnostic OCR systems suited for Indian language technologies [6].

The structure of this review is organized as follows: Section 2 outlines the various categories of Deep Learning models utilized in Optical Character Recognition (OCR). Section 3 provides a structural comparison of these DL models based on their application in text recognition. Section 4 presents an in-depth discussion of previously published studies by different researchers. Section 5 addresses the key challenges encountered in the recognition of Indic scripts. Section 6 highlights the core contributions and novelty of this work. Section 7 outlines potential directions for future research, and finally, Section 8 offers concluding remarks.

II. Deep Learning Models for OCR

Deep Neural Networks (DNNs) consist of several intermediate layers situated between the input and output layers, allowing them to capture intricate and non-linear patterns within the data [7]. These architectures are generally feed-forward, meaning data moves in a single direction that is from the input through hidden layers to the output—without recursive feedback. Modern deep learning development has been greatly simplified by versatile tools such as TensorFlow, Keras, Torch, Caffe, and CNTK, which offer user-friendly interfaces and modular capabilities for building models across diverse domains.

The Deep Learning Models are represented as in Figure I.

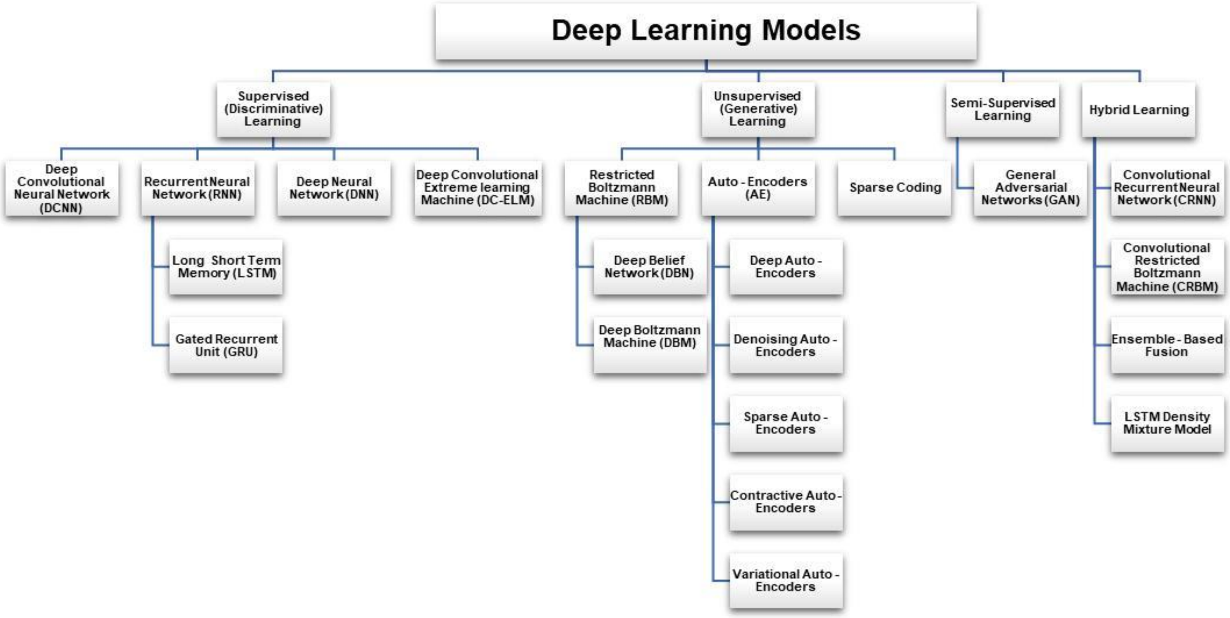

FIGURE I. DEEP LEARNING MODELS FOR TEXT RECOGNITION

Based on their learning strategies and target objectives, deep learning models are typically grouped into four primary categories:

1. Supervised (Discriminative) learning
2. Unsupervised (Generative) learning
3. Semi-Supervised learning
4. Hybrid Deep Networks

These models have achieved significant success in addressing complex pattern recognition and computer vision tasks, and are increasingly being used in multilingual script identification and text recognition systems [8].

A. *Supervised (Discriminative) learning*

Supervised DL models operate on datasets where each input is paired with a corresponding labeled output. These architectures learn to map input features to target labels by minimizing prediction errors during training. Popular supervised models include Convolutional Neural Networks (CNNs) for spatial data like images, and Recurrent Neural Networks (RNNs) for sequential data such as text and speech. These models are widely adopted in tasks like handwriting recognition, object classification, and language translation, owing to their strong generalization capabilities when sufficient annotated data is available [9]. The different discriminative models are:

1) *Deep Convolutional Neural Network (DCNN)*
DCNNs are extensively utilized for analysing two-dimensional inputs such as handwritten characters and digital images. These models employ a sequence of layers including convolution, pooling, and fully connected operations to extract and process spatial features hierarchically [10]. Due to their robust accuracy and rapid learning capabilities, DCNNs are highly effective, provided that sufficient annotated training data is available. Notable model variants include AlexNet, VGGNet, and ResNet, which are widely adopted in pattern recognition and OCR systems.

2) *Recurrent Neural Network (RNN)*
RNNs are tailored for sequential data processing, making them ideal for tasks like speech analysis and handwriting recognition. By maintaining internal memory across time steps, these models learn temporal dependencies in input sequences. However, traditional RNNs often face gradient

vanishing or explosion issues, which can hinder long-range learning capabilities.

a) *Long Short-Term Memory (LSTM)*
LSTM models address RNN limitations by incorporating gating mechanisms namely input, forget, and output gates, that regulate information flow. This architecture enables the model to capture long-distance dependencies in sequence data. LSTMs are prominent in applications involving continuous handwriting recognition, language modeling, and speech-to-text tasks.

b) *Gated Recurrent Unit (GRU)*
GRUs offer a more computationally efficient alternative to LSTMs by using only two gates (reset and update) and omitting a separate memory cell. This streamlining enhances training efficiency while maintaining high performance, making GRUs effective for tasks such as real-time language modeling and multimodal recognition systems.

3) *Deep Neural Network (DNN)*
DNNs are structured as multi-layered feedforward networks designed to learn complex, non-linear mappings from input to output. These architectures support both classification and regression and are employed across domains such as document classification, character recognition, and feature extraction. Despite their versatility, DNNs often require longer training times and higher computational capacity.

4) *Deep Convolutional Extreme Learning Machine (DC-ELM)*
DC-ELM is a hybrid model that integrates convolutional feature extraction (as in CNNs) with the fast-learning capacity of Extreme Learning Machines (ELMs). This approach accelerates training by employing random weight initialization and avoiding iterative updates. It has shown promising outcomes in image and text recognition scenarios, particularly where rapid deployment and reduced computational cost are essential.

B. Unsupervised (Generative) learning

Unsupervised DL models are trained using data that lacks explicit labels. These architectures aim to uncover hidden patterns, groupings, or structures within the input data. Models such as Autoencoders, Restricted Boltzmann Machines (RBMs) are commonly used for feature extraction, dimensionality reduction, and data generation. Unsupervised approaches are particularly valuable when labeled datasets are scarce, and they are extensively used in anomaly detection, clustering, and representation learning across various domains [11]. The different generative models are:

1) *Restricted Boltzmann Machine (RBM)*
RBM is a generative neural model designed to discover latent structures within unlabelled datasets. It consists of two layers that is visible and hidden, which are fully connected to each other but have no intra-layer connections, forming a bipartite architecture. RBMs function by learning the probability distribution over inputs and are especially useful in reducing data dimensionality, uncovering hidden features, and collaborative filtering tasks. Common variants include Bernoulli-Bernoulli RBM for binary data and Gaussian-Bernoulli RBM for continuous data.

a) Deep Belief Network (DBN)
DBNs are constructed by layering multiple RBMs and allow unsupervised pre-training followed by fine-tuning in a supervised manner. The model features both directed and undirected connections, enabling robust hierarchical feature extraction. DBNs are extensively applied in domains such as anomaly detection, information retrieval, and classification of text and image patterns.

b) Deep Boltzmann Machine (DBM)
DBMs are another generative deep network that differs from DBNs by utilizing undirected, bidirectional links across all hidden layers. This configuration enables joint learning of complex feature hierarchies from raw, unlabeled data. DBMs are well-suited for data-intensive applications, including speech recognition and visual object detection, though they require substantial computational resources.

2) *Auto-Encoders (AE)*
Autoencoders are unsupervised neural models that aim to learn efficient codings by compressing and reconstructing input data. They comprise an encoder, which maps inputs to a compact latent space, and a decoder, which attempts to reconstruct the original input. These models are widely used for representation learning, noise reduction, and dimensionality compression in fields like OCR and NLP. Several notable AE variants include:

a) *Deep Autoencoders:* Designed for compressing large-scale, high-dimensional datasets and retrieving relevant image/text information.
b) *Denoising Autoencoders:* Trained to reconstruct clean input from intentionally corrupted data, thereby learning robust feature representations.
c) *Sparse Autoencoders:* Encourage only a subset of neurons to activate during

training, ensuring the network learns discriminative and compact encodings.

d) *Contractive Autoencoders:* Incorporate a regularization term to minimize sensitivity to small variations in input, thereby enhancing generalization.

e) *Variational Autoencoders (VAE):* Learn probability distributions over latent variables and are useful in generative tasks and synthetic data creation.

3) *Sparse Coding*
Sparse coding is a feature learning method where input signals are represented as sparse linear combinations of a learned set of basis functions or "dictionary elements." This approach is advantageous in capturing structured patterns in unlabeled data and is commonly employed for dictionary learning, image clustering, and high-dimensional feature extraction. However, its effectiveness diminishes when applied to binary or non-Gaussian data types.

C. Semi-Supervised learning

Semi-supervised learning leverages a combination of labeled and unlabeled datasets to enhance model accuracy, particularly in scenarios where manually tagging data is expensive or limited in availability [12]. It is widely used in domains like medical imaging, where only a limited subset of data can be labeled by experts.

1) *Generative Adversarial Networks (GANs)*
GANs consist of a pair of neural models that is the generator and the discriminator, that are trained through a competitive process, where one network generates data while the other evaluates its authenticity. Though originally designed for unsupervised learning, GANs can be adapted for semi-supervised settings by using limited labelled data to enhance classification accuracy.
   - *Generative Component:* Produces new data samples that resemble the original dataset.
   - *Adversarial Component:* Ensures both networks improve by competing with each other.
   - *Network Component:* Employs deep neural architectures as the foundational AI models during the training phase.

The various Applications of GANs are that it is used in a range of tasks including text-to-image synthesis, image enhancement, facial aging, inpainting, and 3D object generation.

D. Hybrid Deep Networks

Hybrid deep learning models combine supervised and unsupervised techniques to enhance performance in complex tasks like script, image, and voice recognition. By integrating the strengths of multiple models—especially CNNs—hybrid architectures provide improved accuracy and robustness [13].

1) *Convolutional Recurrent Neural Network (CRNN)*
CRNN integrates CNNs for feature extraction and RNNs for sequence modelling. It outperforms standalone models in handwriting recognition, voice detection, video analysis, and document classification [14].

2) *Convolutional Restricted Boltzmann Machine (CRBM)*
CRBM merges the benefits of CNN and RBM architectures to improve feature learning. It is effective in tasks like dimensionality reduction, regression, and classification, particularly for image and voice processing.

3) *Ensemble-Based Fusion*
This method combines multiple weak classifiers (e.g., CNNs) into a stronger composite model. Used for tasks like handwritten text recognition, face detection, and fake news identification, ensemble models reduce classification errors through classifier diversity.

4) *LSTM Density Mixture Model*
A hybrid model combining LSTM networks with density mixture techniques to enhance recognition in sequences and patterns. It has shown success in fake news detection, speech and handwriting recognition, and long text identification.

The domains and Use Cases of above Deep Learning Models are discussed as follows:

- *DCNN*: Text and image recognition, speech processing, face identification, video analysis, NLP, and healthcare.
- *RNN*: Handwriting and speech sequence modeling.
- *LSTM & GRU*: Long-sequence tasks such as handwriting, speech, gesture recognition, image processing, and text compression.
- *DNN*: Document analysis, writer identification, text and image classification.
- *DC-ELM*: Fast learning for image, gesture, and speech recognition, and text classification.
- *RBM*: Feature extraction, topic modeling, collaborative filtering, and fake news detection.
- *DBN*: Used in predictive systems, information retrieval, and pattern recognition in text and images.
- *DBM*: Hierarchical learning for complex tasks like text, object, and speech recognition.
- *Auto-Encoders*: Dimensionality reduction and feature representation in text and language tasks.

- *Deep Auto-Encoders*: Image search and compression of high-dimensional data.
- *Denoising Auto-Encoders*: Recognition from noisy data in text, speech, and image domains.
- *Sparse Auto-Encoders*: Efficient feature learning for text, image, and NLP tasks.
- *Contractive Auto-Encoders:* Robust feature extraction for text and image processing.
- *Variational Auto-Encoders*: Data generation, document analysis, and text-image understanding.
- *Sparse Coding*: Unsupervised feature extraction for image clustering, signal reconstruction, and dictionary learning.
- *GAN*: Synthetic data generation (e.g., faces, objects), image translation, editing, and 3D modeling.
- *CRNN:* Sequential and spatial learning for handwriting, video, and document recognition.
- *CRBM*: Hybrid modeling for advanced feature learning in visual and audio recognition.
- *Ensemble-Based Fusion*: Combined classifiers for improved accuracy in text, image, and voice identification, including misinformation detection.
- *LSTM Density* Mixture Model: Sequential pattern detection in speech, text, gesture, and long-form content.

III. COMPARATIVE ANALYSIS

This section presents a structured comparison of widely used deep learning architectures applied to text recognition tasks. The models are categorized based on their learning paradigms supervised, unsupervised, semi-supervised and hybrid and evaluated in terms of their core characteristics, strengths, and limitations. Each model offers distinct capabilities for handling different aspects of text recognition, such as spatial feature extraction, sequential data processing, and representation learning. The comparison aims to highlight the practical trade-offs and suitability of these models, particularly in the context of OCR systems for complex and diverse scripts, including Indic languages.

TABLE I. STRUCTURAL COMPARISON OF DEEP LEARNING MODELS FOR TEXT RECOGNITION

S.No	Model	Category	Key Characteristics	Strengths	Limitations
1	**DCNN** (Deep Convolutional Neural Network)	Supervised	Extracts spatial features using convolutional filters; ideal for 2D input like images	Fast learning, effective for image-based text classification	Requires large labeled datasets for accurate results
2	**RNN** (Recurrent Neural Network)	Supervised	Captures sequential data dependencies using shared weights across time steps	Excellent for sequence modeling; LSTM, GRU variants improve long-term memory retention	Suffers from vanishing gradients and needs large training data
3	**DNN** (Deep Neural Network)	Supervised	Multiple fully connected layers model non-linear relationships	High accuracy in both classification and regression tasks	Slower training due to backpropagation complexity; sensitive to error propagation
4	**DC-ELM** (Deep Convolutional Extreme Learning Machine)	Supervised	Employs Gaussian functions for sampling local connections	High-speed learning and reduced complexity	Limited information on effectiveness in large-scale OCR tasks
5	**RBM** (Restricted Boltzmann Machine)	Unsupervised	Bipartite structure with visible and hidden layers	Effective in dimensionality reduction and collaborative filtering	Training becomes inefficient for large-scale data
6	**AE** (Autoencoders and Variants)	Unsupervised	Learns compact feature representations; input and output dimensions are equal	No need for labeled data; variants like Denoising and Variational AEs increase robustness	Needs pretraining; performance may suffer from vanishing gradient issues
7	**Sparse Coding**	Unsupervised	Represents data as sparse combinations of basis functions	Useful for feature selection, dictionary learning	Struggles with non-Gaussian/binary data
8	**GAN** (Generative Adversarial Network)	Semi-Supervised	Generator-discriminator setup enabling learning from both labeled and unlabeled data	Effective in data synthesis and augmentation tasks	Unstable training and requires large datasets
9	**CRNN** (Convolutional Recurrent Neural Network)	Hybrid	Combines CNN for spatial and RNN for sequential features	Accurate in handwriting, document analysis, and speech recognition	Higher complexity and computational cost

10	**CRBM** (Convolutional Restricted Boltzmann Machine)	Hybrid	Fuses RBM and CNN architectures for spatial-feature learning	Handles video/image sequences effectively	Computationally intensive
11	**Ensemble-Based Fusion**	Hybrid	Integrates multiple models to form a stronger classifier	Reduces classification error and improves generalization	Requires careful model selection and tuning
12	**LSTM Density Mixture Model**	Hybrid	Combines LSTM with probabilistic mixture components	Excellent for sequence learning and fake news detection	Model complexity and tuning challenges

IV. RELATED WORK

In recent years, significant progress has been observed in the field of offline handwritten text recognition, particularly for Indic scripts. Numerous research efforts have employed deep learning techniques to address the complexities associated with scripts such as Devanagari, Gurmukhi, Bangla, Kannada, and Telugu, with the goal of advancing Optical Character Recognition (OCR) systems tailored for these languages.

Convolutional Neural Networks (CNNs) and Deep Convolutional Neural Networks (DCNNs) have emerged as the most frequently utilized architectures, consistently delivering superior results. Notably, CNN-based models by Reddy et al. achieved an impressive accuracy of 99.85% for Hindi script, while Acharya et al. and Bisht et al. reported over 98% for Devanagari. Similarly, DCNN models demonstrated high accuracy for Bengali (99.50% by Maity et al.) and Malayalam (96.90% by Jino et al.).

Beyond CNNs, other deep learning frameworks, including Deep Belief Networks (DBNs), Autoencoders (AEs), and Recurrent Neural Networks (RNNs), have also been explored. These methods were implemented on various datasets using different train-test splits, highlighting their adaptability.

TABLE II. REVIEW-BASED COMPARISON OF TEXT RECOGNITION FRAMEWORKS USING DEEP LEARNING

S.No.	Author(s)	Script/Language	Model	Dataset	Training	Testing	Accuracy (%)
1.	Acharya et al. [15]	Devanagari	DCNN	92,000	78,200	13,800	98.47
2.	Reddy et al. [16]	Hindi	CNN	20,000	20,000	10,000	99.85
3.	Deore et al. [17]	Devanagari	DCNN	5,800	4,640	1,160	96.55
4.	Gurav et al. [18]	Devanagari	DCNN	34,604	23,004	11,600	99.65
5.	Guha et al. [19]	Devanagari	DCNN	Multiple Datasets	78,200+	13,800+	97.29–99.63
6.	Bisht et al. [20]	Devanagari	CNN	92,000	78,200	13,800	98.93
7.	Avadesh et al. [21]	Sanskrit	CNN	11,230	10,106	1,124	93.32
8.	Jindal et al. [22]	Gurmukhi	DCNN	3,500	2,800	700	74.66
9.	Kumar et al. [23]	Gurmukhi	DNN	6,000	3,000	3,000	99.30
10.	Badra [24]	Bangla	DCNN	52,788	28,529	15,859	91.25
11.	Maity et al. [25]	Bengali	DCNN	15,000	12,000	3,000	99.50
12.	Jino et al. [26]	Malayalam	DCNN	29,516	18,840	10,676	96.90
13.	Sujatha et al. [27]	Telugu, Hindi	CNN	44,820	31,374	13,446	-
14.	Kundu et al. [28]	12 Indic Scripts	CNN	18,000	12,000	6,000	96.30
15.	Rani et al. [29]	Kannada	CNN	92,000	31,654	9,401	73.51
16.	Gautam et al. [30]	Brahmi	DCNN	6,475	4,856	1,619	92.47
17.	Shinde et al. [31]	Marathi	CNN	9,360	7,488	1,872	98.00
18.	Chhajro et al. [32]	Urdu	RNN	4,668	3,734	934	97.00
19.	Castro et al. [33]	English	LSTM	IAM-DB	6,482	2,915	84.00
20.	Sazal et al. [34]	Bangla	DBN	36,500	27,900	8,600	90.27
21.	Alom et al. [35]	Bangla	DBN	6,000	5,000	1,000	98.78
22.	Mahapatra et al. [36]	English Devanagari, Kannada	AE	1,31,600	1,12,800	18,800	86.67–95.30
23.	Chandio et al. [37]	Sindhi	ResNet + CNN	6,760	5,408	1,352	94.00

V. CHALLENGES IN INDIC SCRIPT RECOGNITION

Recognition of Indic scripts presents several unique challenges due to their complex structure, large character sets, and rich morphological diversity. Many Indic languages include compound characters, modifiers, and overlapping strokes, making segmentation and feature extraction difficult. The scarcity of high-quality, annotated datasets for handwriting and printed text further limits the development of robust recognition models. In addition, script-specific variations in handwriting styles, regional differences, and multi-script documents introduce inconsistencies that hinder generalization. Unlike widely studied scripts such as Latin or Chinese, Indic script OCR still lacks standardized benchmarks and linguistic resources, posing a barrier to achieving high accuracy [38]. These challenges necessitate the development of script-aware models, better data collection efforts, and language-specific preprocessing techniques to advance recognition performance in this domain.

VI. CONTRIBUTIONS AND NOVELTY

This review highlights the growing impact of deep learning in diverse domains, emphasizing its superior accuracy and adaptability compared to traditional methods. It presents a structured overview of deep learning models applied to text recognition, focusing on their architectural features, merits, and limitations. The paper compiles recent advancements, including the evolution of supervised, unsupervised, and hybrid models, and provides a comparative analysis of their performance. A key contribution lies in its focused examination of Indic script recognition, a relatively underexplored area, identifying existing gaps in data availability, model optimization, and storage challenges. The novelty of this work lies in its comprehensive treatment of deep learning applications for Indic scripts, offering insights to guide future research and dataset development.

VII. FUTURE aSPECTS

While substantial progress has been made in recognizing offline handwritten text in non-Indic scripts like Chinese and Arabic, research in Indic script recognition remains limited. There is a growing need for high-quality datasets and efficient models tailored to Indic languages. Additionally, due to the high memory requirements of modern deep learning architectures, future efforts should focus on model compression and optimization to meet storage constraints without compromising performance.

VIII. CONCLUSION

Deep learning has emerged as a powerful and rapidly advancing framework within machine learning, demonstrating high accuracy and versatility across various domains. Its layered architecture and supervised learning capabilities enable effective classification and data management. This paper provides a comprehensive overview of deep learning models applied to text recognition, outlining their strengths, limitations, and real-world applications. A comparative analysis of existing models further highlights their effectiveness in diverse recognition tasks.

REFERENCES

[1]. Ram, S., Gupta, S., & Agarwal, B. (2018). Devanagri character recognition model using deep convolution neural network. Journal of Statistics and Management Systems, 21(4), 593–599.

[2]. Moudgil, A., Singh, S., Gautam, V., Rani, S., & Shah, S. H. (2023). Handwritten devanagari manuscript characters recognition using capsnet. International Journal of Cognitive Computing in Engineering, 4, 47–54.

[3]. Gupta, D., & Bag, S. (2021). CNN-based multilingual handwritten numeral recognition: A fusion-free approach. Expert Systems with Applications, 165, 113784.

[4]. Narang, S. R., Kumar, M., & Jindal, M. K. (2021). Deepnet devanagari: a deep learning model for devanagari ancient character recognition. Multimedia Tools and Applications, 80, 20671–20686.

[5]. Sonawane, P. R., & Shelke, S. (2018). Handwritten devanagari character classification using deep learning. In 2018 International Conference on Information, Communication, Engineering and Technology (ICICET) (pp.1–4).

[6]. Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2020). A survey of deep learning and its applications: a new paradigm to machine learning. Archives of Computational Methods in Engineering, 27, 1071–1092.

[7]. Roy, A. (2020). Akhcrnet: Bengali handwritten character recognition using deep learning. arXiv preprint arXiv:2008.12995.

[8]. Mahto, M. K., Bhatia, K., & Sharma, R. K. (2021). Deep learning-based models for offline gurmukhi handwritten character and numeral recognition. ELCVIA Electronic Letters on Computer Vision and Image Analysis, 20(2), 69–82.

[9]. Jangid, M., & Srivastava, S. (2018). Handwritten devanagari character recognition using layer-wise training of deep convolutional neural networks and adaptive gradient methods. Journal of Imaging, 4(2), 41.

[10]. Mestry, S., Patil, B., Pede, A., & Raul, S. (2023). Devanagari digit and character recognition using convolutional neural network.

[11]. Narang, S. R., Kumar, M., & Jindal, M. K. (2021). DeepNetDevanagari: a deep learning model for Devanagari ancient character recognition. Multimedia Tools and Applications, 80(13), 20671–20686.

[12]. Mosavi, A., Ardabili, S., & Varkonyi-Koczy, A. R. (2019). List of deep learning models. In International conference on global research and education (pp. 202–214). Springer.

[13]. Bhardwaj, H., Das, R., Garg, P., & Kumar, R. (2024). Handwritten Text Recognition Using Deep Learning. In 2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT) (pp. 506–511). IEEE.

[14]. Sonkar, C. K., & Kumar, V. (2025). Review of Artificial Intelligence Methods in Handwriting Identification Using CNN-RNN for Textural Features. In 2025 International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI) (pp. 648-654). IEEE.

[15]. Acharya, S., Pant, A. K., & Gyawali, P. K. (2015). Deep learning based large scale handwritten devanagari character recognition. In 2015 9th International conference on software, knowledge, information management and applications (SKIMA), IEEE (pp. 1–6).

[16]. Reddy, R. V. K., & Babu, U. R. (2019). Handwritten Hindi Character Recognition using Deep Learning Techniques. International Journal of Computer Sciences and Engineering, 7(7).

[17]. Deore, S. P., & Pravin, A. (2020). Devanagari handwritten character recognition using finetuned deep convolutional neural network on trivial dataset. Sādhanā, 45, 1–13.

[18]. Gurav, Y., Bhagat, P., Jadhav, R., & Sinha, S. (2020). Devanagari handwritten character recognition using convolutional neural networks. In 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), IEEE (pp. 1–6).

[19]. Guha, R., Das, N., Kundu, M., Nasipuri, M., & Santosh, K. (2020). Devnet: an efficient cnn architecture for handwritten devanagari character recognition. International Journal of Pattern Recognition and Artificial Intelligence, 34(12), 2052009.

[20]. Bisht, M., & Gupta, R. (2020). Multiclass recognition of offline handwritten devanagari characters using CNN. International Journal of Mathematical, Engineering and Management Sciences, 5(6), 1429.

[21]. Avadesh, M., & Goyal, N. (2018). Optical character recognition for sanskrit using convolution neural networks. In 2018 13th IAPR International Workshop on Document Analysis Systems (DAS), IEEE (pp. 447–452).

[22]. Jindal, U., Gupta, S., Jain, V., & Paprzycki, M. (2019). Offline handwritten gurumukhi character recognition system using deep learning. In Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals: Proceedings of GUCON, Springer (pp. 121–133).

[23]. Kumar, N., & Gupta, S. (2018). Offline handwritten gurmukhi word recognition using deep neural networks. International Journal of Pure and Applied Mathematics, 119(12), 14749–14767.

[24]. Bhadra, D. (2019). Hand-written bangla character recognition using deep convolutional neural network.

[25]. Maity, S., Dey, A., Chowdhury, A., & Banerjee, A. (2020). Handwritten bengali character recognition using deep convolution neural network. In MIND 2020, Springer (pp. 84–92).

[26]. Jino, P., Balakrishnan, K., & Bhattacharya, U. (2019). Offline handwritten malayalam word recognition using a deep architecture. In SocProS 2017, Volume 1, Springer (pp. 913–925).

[27]. Sujatha, P., & Bhaskari, D. L. (2019). Telugu and hindi script recognition using deep learning techniques. Int. J. Innov. Technol. Explor. Eng, 8, 2278–3075.

[28]. Kundu, S., Paul, S., Singh, P. K., Sarkar, R., & Nasipuri, M. (2020). Understanding nfc-net: a deep learning approach to word-level handwritten indic script recognition. Neural Computing and Applications, 32, 7879–7895.

[29]. Rani, N. S., Subramani, A., Kumar, A., & Pushpa, B. (2020). Deep learning network architecture based kannada handwritten character recognition. In ICIRCA, IEEE (pp. 213–220).

[30]. Gautam, N., Chai, S. S., & Jose, J. (2020). Recognition of brahmi words by using deep convolutional neural network.

[31]. Shinde, A., & Dandawate, Y. (2020). Convolutional neural network based handwritten marathi text recognition. J. Xidian Univ., 14(8), 528–536.

[32]. Chhajro, M. A., Khan, H., Khan, F., Kumar, K., Wagan, A., & Solangi, S. (2020). Handwritten urdu character recognition via images using different machine learning and deep learning techniques. Indian Journal of Science and Technology, 13(17), 1746–1754.

[33]. Castro, D., Bezerra, B. L., & Valença, M. (2018). Boosting the deep multidimensional long-short - term memory network for handwritten recognition systems. In ICFHR, IEEE (pp. 127–132).

[34]. Sazal, M. M. R., Biswas, S. K., Amin, M. F., & Murase, K. (2014). Bangla handwritten character recognition using deep belief network. In EICT, IEEE (pp. 1–5).

[35]. Alom, M. Z., Sidike, P., Taha, T. M., & Asari, V. K. (2017). Handwritten bangla digit recognition using deep learning. arXiv preprint arXiv:1705.02680.

[36]. Mahapatra, D., Choudhury, C., & Karsh, R. K. (2020). Handwritten character recognition using knn and svm based classifier over feature vector from autoencoder. In MIND 2020, Springer (pp. 304–317).

[37]. Chandio, A. A., Leghari, M., Panhwar, A. O., Nizamani, S. Z., & Leghari, M. (2020). Deep learning-based isolated handwritten sindhi character recognition. Indian Journal of Science and Technology, 13(25), 2565–2574.

[38]. Singh, S., Garg, N. K., & Kumar, M. (2023). Feature extraction and classification techniques for handwritten devanagari text recognition: a survey. Multimedia Tools and Applications, 82(1), 747–775.