International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
***Article Received:** 25 January 2023 **Revised:** 12 February 2023 **Accepted:** 30 March 2023*

Advancing Hindi Speech and Handwriting Recognition through Deep Learning in Noisy and Diverse Environments

Barkha Sahu
Assistant Professor, Department of Computer Science and Engineering
Institute of Engineering & Science, IPS Academy, Indore, India
barkhasahu@ipsacademy.org

Abstract— Hindi, as the most spoken language in India, requires advanced speech and handwriting recognition systems to enhance human-machine interaction, particularly in noisy and diverse real-world environments. Traditional automatic speech recognition (ASR) techniques like MFCC and PLP struggle with noise and dialect variations. Recent studies demonstrate that Gammatone Frequency Cepstral Coefficients (GFCC) combined with DNN-HMM architectures significantly improve noise-robust Hindi speech recognition. Similarly, handwriting recognition faces challenges due to script complexity, compound characters, and individual writing styles. Hybrid CNN-RNN architectures with spatial transformers and augmented datasets have shown substantial improvements in recognizing handwritten Hindi text. Furthermore, optical character recognition (OCR) and natural language processing (NLP) for Hindi remain underdeveloped, especially when handling diverse scripts and noisy inputs. This research advocates for large-scale, diverse datasets, Transformer-based models, and cross-lingual learning to enhance recognition capabilities. The integration of multimodal systems and advanced deep learning architectures promises to bridge current gaps, making recognition systems more accurate, inclusive, and applicable across educational, social, and digital domains. This paper presents a comprehensive analysis of existing methodologies and proposes future directions to advance Hindi speech and handwriting recognition through deep learning in complex environments.

Keywords- Hindi Speech Recognition, Handwriting Recognition, Deep Learning, GFCC, CNN-RNN, Noisy Environments.

I. INTRODUCTION

Hindi is the most widely spoken language in India, encompassing a rich linguistic heritage and serving as a medium of communication for millions of people across the subcontinent. The growing penetration of digital technologies in India's social, educational, and administrative systems has heightened the need for efficient human-machine interaction systems capable of processing [1] Hindi speech and handwriting. Historically, speech and handwriting recognition technologies were primarily developed for English and other resource-rich languages, leaving a significant technological gap for Indian languages like Hindi. This gap poses challenges for inclusivity in digital access, education, and public services. As India progresses towards a digitally empowered society, the development of robust Hindi recognition systems becomes imperative to ensure equitable technological participation for its large Hindi-speaking population [2].

The early developments in speech recognition began in the 1950s with rudimentary systems that could recognize limited English words. Through the 1980s and 1990s, Hidden Markov Models (HMMs) became the standard for speech recognition, enabling more dynamic modeling of speech patterns [3]. However, these models struggled with the phonetic richness and dialectal variations inherent in Hindi. Conventional feature extraction methods such as Mel Frequency Cepstral Coefficients (MFCC) and Perceptual Linear Prediction (PLP) proved inadequate under noisy conditions, which are common in Indian acoustic environments. Recognizing these limitations, researchers introduced Gammatone Frequency Cepstral Coefficients (GFCC), which better emulate the human auditory system's response to noise. When coupled with deep neural networks and HMMs, GFCC-based systems have demonstrated superior robustness in Hindi ASR, especially when combined with language models like trigrams for improved contextual understanding [4].

Parallelly, handwriting recognition has evolved from template matching techniques to sophisticated machine learning algorithms. In the context of Hindi, the recognition of handwritten text presents unique challenges due to the Devanagari script's complexity, including compound characters, modifiers, and stylistic variations between writers. Early OCR systems for Devanagari were limited [5] in accuracy and heavily dependent on clean, printed text. The advent of deep learning, particularly Convolutional Neural Networks (CNNs), revolutionized image-based recognition tasks. Recent research has employed hybrid CNN-RNN architectures to capture both spatial and sequential information in handwriting, significantly improving recognition rates for handwritten Hindi characters and words. The inclusion of spatial transformer networks has further enhanced these models' ability to manage distortions and varied writing styles.

Moreover, the development of comprehensive and diverse datasets remains a pressing need. Most existing datasets for Hindi speech and handwriting are limited in size and diversity, restricting the generalization capabilities of deep learning models. Data augmentation techniques and synthetic data [6] generation have shown promise in mitigating this challenge, yet real-world data collection and annotation remain crucial for advancing these systems.

The integration of multimodal approaches combining speech, text, and visual data represents a promising frontier for overcoming the complexities inherent in Hindi recognition tasks. Furthermore, the rise of Transformer-based models like BERT and Vision Transformers (ViTs) has opened new avenues for enhancing both speech and handwriting recognition by capturing deeper contextual and semantic information.

In conclusion, advancing Hindi speech and handwriting recognition through deep learning is not merely a technical endeavor but a critical step toward digital inclusivity and accessibility. By addressing challenges related to noise, script complexity, and data scarcity, future research can enable more robust, accurate, and user-friendly recognition systems. Such advancements will have far-reaching implications across sectors including education, governance, assistive technologies for differently-abled users, and digital communication, ensuring that the benefits of technology are accessible to all segments of India's diverse population.

II. LITERATURE REVIEW

Hindi, being the most spoken language in India, highlights the need for robust communication systems that facilitate human-machine interaction via speech. Traditional automatic speech recognition (ASR) techniques, particularly those based on Mel Frequency Cepstral Coefficients (MFCC) and Perceptual Linear Prediction (PLP), often fail in noisy environments. This study investigates alternative feature extraction methods, finding that Gammatone Frequency Cepstral Coefficients (GFCC) are more noise-resilient. Using a Deep Neural Network-Hidden Markov Model (DNN-HMM) coupled with a trigram language model, the proposed method surpasses traditional approaches like MFCC and PLP, demonstrating significant improvements in Hindi speech recognition accuracy [1].

Modern deep learning advancements have improved handwritten text recognition, especially in complex scripts like Hindi, which presents challenges due to script complexity and freehand variability. The proposed system employs a hybrid CNN-RNN architecture, inspired by residual blocks for deep convolutional layers and incorporates a spatial transformer layer to handle handwriting distortions. The system, trained and tested on the Indic Word Database, evaluates performance using Word Error Rate (WER) and Character Error Rate (CER). Using synthetic data, this approach achieves notable WER reductions, making it a promising method for lexicon-free handwritten Hindi text recognition [2].

Offline handwritten character recognition for Hindi is complicated by its extensive character set, including numerals, compound characters, and modifiers. Individual handwriting variations further complicate recognition tasks. To address these challenges, the study introduces an augmented dataset that enhances data variability by introducing hypothetical changes. Convolutional Neural Networks (CNNs) are utilized due to their suitability for image-based classification tasks. The adoption of deep learning algorithms significantly enhances recognition accuracy, suggesting that augmented datasets coupled with CNN architectures can effectively address the complexities inherent in Hindi character recognition tasks [3].

Social media platforms like Facebook and Twitter are pivotal for digital communication, but they also harbor hate speech. While most research targets English, this study focuses on detecting hate content in Hindi-English code-mixed language, emphasizing sentences that imply hate without explicit abusive terms. Machine learning and deep learning algorithms like CNN, LSTM, and SVM are analyzed for this task, with NLP techniques aiding in text processing. The study finds these models achieve superior F1 scores, proving effective in hate speech detection in code-mixed scenarios and showcasing NLP's utility in understanding social media content [4].

Children with dyslexia face unique challenges in spelling and word analysis. This study introduces a multimodal Hindi language learning system incorporating an eye-gaze-assisted virtual keyboard (VK) to support text input and feedback. The system's performance was tested on 32 children, half with dyslexia, across three input methods: eye tracker, soft switch, and touchscreen. Results showed dyslexic children performed better with multimodal input systems. Classifiers accurately distinguished dyslexic from non-dyslexic children (AUC > 0.9), suggesting the system's effectiveness for early dyslexia detection and inclusive educational technologies [5].

Named Entity Recognition (NER) is crucial in machine learning for identifying names, positions, and organizations in text, enhancing tasks like information retrieval and summarization. However, NER for Hindi is challenging due to linguistic ambiguity, morphological richness, and limited resources. Despite significant efforts, highly accurate NER systems for Hindi remain elusive. This review evaluates current machine learning approaches to Hindi NER, underscoring the complexities and gaps that persist. The need for advanced, resourceful solutions in NER for Hindi is evident, motivating further research into overcoming these linguistic challenges [6].

Handwritten script identification is essential for improving the accuracy of Optical Character Recognition (OCR) systems, especially given the variations in Indian scripts. This research

proposes using Gabor filters for feature extraction to classify scripts into North or South Indic categories before OCR processing. By analyzing documents from various individuals, the system achieved 100% accuracy in distinguishing North vs. South scripts and 92% in biscript classification using the KNN classifier. This pre-processing enhances OCR performance, suggesting further refinement for intra-region script identification within North and South Indian scripts [7].
Natural Language Processing (NLP) applications have evolved significantly, with an increased focus on Hindi language processing. This paper provides a retrospective analysis of NLP research trends in Hindi using fuzzy logic, identifying publication patterns and key research terms. As NLP technologies advance, the research landscape for Hindi is growing, driven by the need for linguistic tools tailored to India's linguistic diversity. The study emphasizes understanding the evolution of NLP research in Hindi, guiding future investigations and development of language-specific NLP applications [8].

Optical Character Recognition (OCR) remains a vital area of research, particularly for languages like Devanagari, used by over 120 regional languages in India. Despite extensive work in English OCR, Indian scripts face challenges in achieving high recognition reliability due to variations in script, font, and handwriting styles. This review discusses ongoing algorithm development aimed at improving OCR accuracy for Indian scripts. The goal is to enhance character recognition systems to handle diverse scripts and styles, ensuring reliable text conversion for both typed and handwritten documents [9].

The rise of generative artificial speech models has enabled the creation of realistic synthetic speech, posing challenges in audio forensics, especially for Hindi. This study proposes methods to differentiate between AI-synthesized and genuine Hindi speech using features like Bicoherence Phase, MFCC, and Delta Cepstral coefficients. Deep learning models, including VGG16 and a custom CNN, achieved accuracies of 99.83% and 99.99%, respectively. The findings highlight the importance of robust detection techniques to prevent misuse of synthetic speech technologies in Hindi, ensuring the integrity of audio communications [10].

Multiword Expressions (MWEs) are vital in understanding contextual meanings in NLP. This research focuses on Hindi MWEs, proposing a hybrid extraction method combining linguistic patterns, word relationships, and context similarity. Using the TDIL dataset, the approach outperforms traditional context-based and association-based methods. Effective MWE extraction enhances various NLP applications like machine translation and semantic analysis, especially in resource-scarce languages like Hindi. The study contributes to more nuanced text understanding by refining techniques for identifying compound expressions in unstructured data [11].

Named Entity Recognition (NER) is foundational for information extraction tasks, including email parsing and conversational data analysis. This research enhances NER for Hindi using BiLSTM models and experiments with embedding techniques like typical and fast text embeddings across various batch sizes. Achieving a precision of 76.13%, recall of 71.49%, and F1-score of 74.26%, the proposed architecture outperforms existing systems like SpaCy, CoreNLP, and NLTK. These improvements demonstrate the model's reliability, advancing the capabilities of Hindi NER systems for practical NLP applications [12].

Substring search is integral to text processing, and this study validates popular algorithms like Knuth-Morris-Pratt, Boyer-Moore, and Rabin-Karp for Hindi (UTF-8 encoded) text. Utilizing a dataset of 49,000 unique Hindi words, the research adapts these traditionally English-focused algorithms to handle the complexities of Hindi scripts. This experimental validation offers insights into the efficiency and applicability of these algorithms in Hindi text processing, providing a foundation for developing optimized search tools tailored to Indian language data [13].

Braille remains essential for visually impaired communication, and this study enhances English Braille recognition using machine learning. Collecting data from visually impaired students, synthetic data generation expanded the dataset, classified into two groups. Methods like SVM, Decision Trees, and KNN were applied with RICA and PCA-based feature extraction, with RICA yielding superior results. The evaluation used comprehensive metrics, achieving significant accuracy improvements. This research paves the way for more effective Braille-to-text systems, showcasing the potential of advanced feature extraction in assistive technologies [14].

Hypernymy detection, identifying "is-a" relationships (e.g., dog is an animal), is underdeveloped in low-resource languages like Hindi, Bengali, and Amharic. This study introduces gold-standard datasets and utilizes Distributional Thesaurus (DT) embeddings with network embedding methods like DeepWalk and role2vec. Supervised classifiers such as SVM and Random Forest demonstrated promising results. Hybrid approaches combining DT and pre-trained fastText embeddings further improved performance. Validating these methods on English datasets, the study closes a crucial gap, advancing semantic understanding in underrepresented languages [15].

Next word prediction in Hindi enhances typing efficiency by suggesting subsequent words based on prior text. Utilizing BERT and Masked Language Models, this study addresses the digital typing needs of Hindi users, reducing keystrokes and errors. While prevalent in English, such systems are rare for Hindi, despite its widespread use. This predictive technology is crucial for streamlining text input on digital platforms, encouraging further development of intelligent input aids for Indian languages to facilitate faster and more accurate communication [16].

Named Entity Recognition (NER) in NLP is critical for identifying entities like people and places. This study employs a deep learning architecture using recursive BiLSTM with de-noising autoencoders and conditioning logic to improve Hindi

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
***Article Received:** 25 January 2023 **Revised:** 12 February 2023 **Accepted:** 30 March 2023*

NER. Experiments on word embeddings and batch sizes provided insights into optimizing deep learning models. Comparative analysis revealed the proposed architecture's advantages over existing systems, enhancing the precision and reliability of Hindi NER tasks, which are foundational for various linguistic applications in information extraction and data analysis [17].

III. RESEARCH GAPS

1. **Limited Robust ASR for Noisy Environments in Hindi :** While several feature extraction techniques like GFCC have improved speech recognition in noisy conditions, most Hindi ASR systems still lack robustness against diverse and unpredictable real-world noise scenarios. There's a gap in developing generalized models that maintain high accuracy across varying noise levels, dialects, and speaker variations in Hindi [1].
2. **Scarcity of Lexicon-Free Handwriting Recognition Models for Hindi :** Although CNN-RNN architectures have improved handwritten text recognition, the lack of large, diverse datasets for Hindi handwriting limits model generalization. There's a need for scalable, lexicon-free models that can adapt to varying handwriting styles, including cursive and distorted forms in Hindi scripts [2][3].
3. **Inadequate Detection of Hate Speech in Code-Mixed Languages :** Most hate speech detection models focus on monolingual (mainly English) data. Existing approaches underperform on Hindi-English code-mixed texts, especially when abusive content is implicit. Research is lacking in context-aware models that can accurately interpret sentiment and intent in multilingual and code-mixed scenarios [4].
4. **Deficient Assistive Technologies for Dyslexic Hindi Users :** Current multimodal learning systems for dyslexic children are limited in sample diversity and usability across different age groups and languages. There is a significant gap in developing adaptive, gaze-assisted virtual keyboards and learning aids optimized for Hindi-speaking dyslexic users across varying literacy levels [5].
5. **Low-Resource Constraints in Hindi NER Systems :** NER for Hindi is hindered by morphological complexity and limited annotated datasets. Despite advancements using BiLSTM and embeddings, current models still suffer from lower precision and recall compared to English NER systems, revealing a gap in resource creation and model optimization tailored for Hindi [6][12][17].
6. **Lack of Pre-OCR Script Classification in Multilingual Contexts :** Although Gabor filters enhance script classification, the distinction between scripts beyond North and South Indic types, especially for mixed-script documents, is insufficiently addressed. There is room to develop advanced pre-processing methods to improve OCR accuracy for complex multilingual documents [7].
7. **Underexplored NLP Techniques for Hindi Using Fuzzy Logic :** While retrospective studies highlight NLP trends in Hindi, there remains a methodological gap in integrating fuzzy logic with state-of-the-art NLP models to manage Hindi's linguistic ambiguities and syntactic variations effectively [8].
8. **Inadequate OCR Systems for Devanagari and Other Indian Scripts :** Existing OCR systems fail to achieve high reliability for diverse fonts, styles, and handwritten forms of Devanagari and associated regional scripts. There's a gap in universal OCR models that can seamlessly handle the broad variability in Indian language documents [9].
9. **Insufficient Methods for Detecting Synthetic Hindi Speech :** Despite achieving high accuracy in detecting synthetic Hindi speech, existing methods are not extensively validated across diverse speech conditions and generative models. A gap exists in real-time, scalable forensic tools for synthetic speech detection in Hindi [10].
10. **Limited MWE Extraction Models for Hindi :** Current MWE extraction methods for Hindi, though hybrid, still lag in precision when dealing with complex and idiomatic expressions. There's a need for models that integrate deeper semantic understanding and context for more accurate MWE identification [11].
11. **Incomplete Application of Search Algorithms for Indian Languages :** Though substring search algorithms have been tested for Hindi, they are not yet fully optimized for complex script structures and large-scale applications, highlighting a gap in high-performance, language-specific search optimizations [13].
12. **Underdeveloped Braille Recognition Systems for Indian Scripts :** While effective methods exist for English Braille, equivalent systems for Hindi or other Indian scripts remain unexplored. There's a significant research gap in developing Braille recognition tools tailored to Indian language character sets [14].
13. **Hypernymy Detection for Low-Resource Languages Still Nascent :** Efforts in hypernymy detection for Hindi and other low-resource languages are early-stage. More comprehensive datasets and refined models that mitigate lexical memorization biases are needed for robust semantic relationship extraction in these languages [15].
14. **Scarcity of Next-Word Prediction Models for Hindi :** Next-word prediction is well-developed for English but underrepresented in Hindi. Current models like BERT and ML approaches require further training on large Hindi corpora to ensure accuracy and efficiency in predictive text applications [16].
15. **Optimization Challenges in Deep Learning for Hindi NER :** Although recursive BiLSTM models have improved Hindi NER, optimization regarding batch sizes, embedding types, and model architectures is still underexplored. There's a need for systematic optimization studies to fine-tune NER performance for Hindi across various domains [17].

IV. SOLUTIONS TO IDENTIFIED RESEARCH GAPS

1. **Enhancing Hindi ASR in Noisy Environments :** To address the limitations of Hindi ASR systems in noisy conditions, future research should integrate advanced noise-robust features like GFCC with deep learning architectures such as CNN-LSTM hybrids. Employing data augmentation techniques to simulate various noise scenarios can enhance model generalization. Additionally, incorporating transfer learning from well-resourced languages and domain adaptation techniques can further improve recognition performance across diverse speaker profiles and dialects, ensuring more accurate and reliable speech recognition in real-world settings.
2. **Developing Scalable Handwriting Recognition Models for Hindi :** The challenge of lexicon-free handwriting recognition for Hindi can be tackled by curating expansive, diverse datasets that capture different handwriting styles, cursive scripts, and distortions. Implementing attention-based sequence-to-sequence models and Transformer-based architectures could significantly improve recognition accuracy. Synthetic data generation methods can further enrich datasets, while unsupervised and semi-supervised learning can mitigate data scarcity issues. These approaches will help create more adaptable and accurate models for handwritten Hindi text recognition.
3. **Advanced Hate Speech Detection in Code-Mixed Languages :** To improve hate speech detection in Hindi-English code-mixed texts, researchers should develop context-aware models leveraging Transformer-based architectures like BERT and multilingual models like mBERT or IndicBERT. These models can be fine-tuned on specialized datasets that include implicit hate speech cases. Additionally, sentiment analysis combined with semantic role labeling can help in understanding nuanced expressions of hate. Incorporating contextual embeddings that capture both linguistic and cultural nuances can further enhance detection accuracy in code-mixed scenarios.
4. **Designing Assistive Technologies for Dyslexic Hindi Users :** The development of adaptive, multimodal learning systems for dyslexic users can be enhanced by integrating eye-tracking, speech input, and haptic feedback mechanisms. Creating larger and more varied datasets of dyslexic Hindi users will enable the training of more accurate and inclusive models. Personalized learning interfaces, powered by AI, can adjust input methods based on the user's proficiency and cognitive needs. Such systems can significantly improve educational outcomes and user engagement for dyslexic learners in the Hindi language.
5. **Resource-Enriched NER for Hindi Language :** To overcome the limitations in Hindi NER, it is essential to develop comprehensive annotated corpora covering diverse domains and linguistic constructs. Using multilingual pre-trained models like XLM-R and fine-tuning them on Hindi-specific datasets can improve entity recognition performance. Additionally, incorporating morphological analyzers and syntactic parsers tailored to Hindi will help resolve linguistic ambiguities. Crowdsourcing and active learning techniques can also accelerate the expansion of labeled datasets, enhancing the robustness of Hindi NER models.
6. **Refining Script Classification for Multilingual OCR :** Improving pre-OCR script classification can be achieved by developing deep learning models that not only distinguish North and South Indic scripts but also classify individual script types within these categories. Employing ensemble learning approaches and leveraging convolutional and recurrent neural networks can improve classification accuracy. Furthermore, integrating metadata such as regional language usage statistics can assist in more precise script identification, facilitating better OCR performance in multilingual documents.
7. **Integrating Fuzzy Logic with NLP for Hindi :** To bridge the gap in NLP methodologies for Hindi, integrating fuzzy logic with deep learning models can help manage linguistic ambiguities effectively. Fuzzy logic can aid in handling uncertainties and imprecise data common in Hindi syntax and semantics. Combining this with Transformer models can create more robust systems for tasks like text summarization, translation, and sentiment analysis. This hybrid approach can enhance the interpretability and reliability of NLP applications in Hindi.
8. **Improving OCR Accuracy for Indian Scripts :** Enhancing OCR systems for Indian scripts requires developing models that can adapt to the diverse visual characteristics of Devanagari and other scripts. Incorporating generative adversarial networks (GANs) for data augmentation can help simulate diverse handwriting styles and printed fonts. Using transfer learning from pre-trained vision models like EfficientNet or Vision Transformers (ViTs) can also boost accuracy. Custom loss functions tailored for script-specific features can further refine recognition performance.
9. **Robust Detection of Synthetic Hindi Speech :** To strengthen synthetic speech detection in Hindi, future research should focus on multimodal approaches combining audio, spectral features, and phonetic cues. Developing real-time detection systems using lightweight deep learning models will enable practical applications in security and forensics. Continuous updating of detection models with newer generative techniques, like those from voice cloning technologies, will keep the detection mechanisms resilient against evolving AI synthesis methods.
10. **Enhanced MWE Extraction for Hindi :** To improve multiword expression (MWE) extraction, hybrid models that combine linguistic rule-based systems with deep learning frameworks can be developed. Incorporating contextual embeddings like those from BERT models trained on Hindi text can enhance semantic understanding. Leveraging graph-based approaches to model word relationships and context can further

improve extraction accuracy. This combination of methodologies will enable more precise identification of complex and idiomatic expressions in Hindi.

11. **Optimizing Search Algorithms for Hindi Scripts :** Optimizing substring search algorithms for Hindi can involve adapting existing algorithms to accommodate UTF-8 encoding complexities. Leveraging parallel processing and GPU acceleration can improve performance for large datasets. Additionally, developing language-specific indexing methods that consider the unique phonetic and script characteristics of Hindi will enhance search efficiency. Integrating machine learning-based optimization techniques can also refine pattern matching in large-scale Hindi text applications.
12. **Developing Braille Recognition Systems for Hindi :** To address the gap in Braille recognition for Hindi, creating dedicated datasets for Hindi Braille is crucial. Adapting the machine learning techniques used for English Braille, such as RICA-based feature extraction, to Hindi scripts can improve accuracy. Developing multimodal recognition systems that combine tactile sensors with visual inputs can further enhance recognition capabilities. This would facilitate better educational and communication tools for visually impaired Hindi users.
13. **Expanding Hypernymy Detection in Low-Resource Languages :** Improving hypernymy detection for Hindi requires expanding gold-standard datasets and refining distributional thesaurus embeddings with context-rich data. Incorporating multilingual embeddings and cross-lingual transfer learning can boost performance in resource-constrained languages. Utilizing advanced graph-based neural networks can also capture hierarchical relationships more effectively. These strategies will enable more accurate semantic relationship modeling in Hindi and other low-resource languages.
14. **Advancing Next-Word Prediction for Hindi :** Developing more accurate next-word prediction systems for Hindi requires training large-scale models like BERT on extensive Hindi corpora. Data augmentation and transfer learning from multilingual models can help overcome dataset limitations. Implementing attention mechanisms to better capture context and sequential dependencies can improve prediction accuracy. Customizing these models for mobile and low-resource devices will also ensure widespread adoption and usability.
15. **Optimizing Deep Learning for Hindi NER :** To optimize Hindi NER, systematic experiments on embedding strategies, batch sizes, and model architectures like recursive BiLSTM and Transformer models are needed. Hyperparameter tuning and the use of automated machine learning (AutoML) can streamline model optimization. Incorporating domain-specific embeddings and contextual information can further enhance model accuracy. Establishing benchmark datasets and evaluation protocols specific to Hindi will standardize performance assessments and drive improvements in NER tasks.

V. CONCLUSION AND FUTURE SCOPE

5.1 Conclusion : The exploration of Hindi speech and handwriting recognition through deep learning has shown significant promise, particularly in addressing the unique challenges posed by noisy environments and diverse script variability. Techniques such as Gammatone Frequency Cepstral Coefficients (GFCC) combined with DNN-HMM architectures have demonstrated improved robustness in speech recognition under adverse conditions. Similarly, the integration of CNN-RNN hybrids, augmented datasets, and attention-based models has substantially enhanced handwriting recognition for Hindi, effectively handling script complexities and individual writing variances. Despite these advances, the current systems still face limitations in real-world applications where unpredictable noise, dialectal differences, and handwriting diversity persist. Moreover, the absence of comprehensive datasets for Hindi further constrains model generalization and scalability.

5.2 Future Scope : Future research can focus on developing large-scale, diverse, and publicly available datasets for both Hindi speech and handwriting to support more generalized and adaptable models. Incorporating Transformer-based architectures, such as Wav2Vec for speech and Vision Transformers (ViTs) for handwriting, could further advance recognition accuracy and efficiency. Additionally, implementing cross-lingual transfer learning can help leverage insights from resource-rich languages to improve Hindi models. Exploring multimodal approaches that combine audio, text, and visual features may also yield more holistic and context-aware recognition systems, making them practical for applications in education, governance, and digital accessibility in India.

References

1. Gaudani H, Patel NM. Comparative study of robust feature extraction techniques for ASR for limited resource Hindi language. InProceedings of second international conference on sustainable expert systems: ICSES 2021 2022 Feb 26 (pp. 763-775). Singapore: Springer Nature Singapore.
2. Thomas M, DM BG. CNN-RNN Hybrid model based Hindi Character Recognition. In2022 International Interdisciplinary Humanitarian Conference for Sustainability (IIHC) 2022 Nov 18 (pp. 606-610). IEEE.
3. Indian A, Bhatia K, Kumar K. Offline handwritten hindi character recognition using deep learning with augmented dataset. InCyber Security in Intelligent Computing and Communications 2022 Mar 12 (pp. 129-141). Singapore: Springer Singapore.
4. Pareek K, Choudhary A, Tripathi A, Mishra KK. Comparative Analysis of Social Media Hate Detection over Code Mixed Hindi-English Language. InAdvances in Data and Information Sciences: Proceedings of ICDIS

2021 2022 Feb 8 (pp. 551-561). Singapore: Springer Singapore.

5. Meena YK, Cecotti H, Bhushan B, Dutta A, Prasad G. Detection of dyslexic children using machine learning and multimodal hindi language eye-gaze-assisted learning system. IEEE Transactions on Human-Machine Systems. 2022 Nov 28;53(1):122-31.
6. Shelke R, Vanjale S. Review Based on Named Entity Recognition for Hindi Language Using Machine Learning Approach. InProceedings of Second International Conference in Mechanical and Energy Technology: ICMET 2021, India 2022 Jun 27 (pp. 333-340). Singapore: Springer Nature Singapore.
7. Shreesha SK, Anita HB. Classification of North and South Indian Handwritten Scripts using Gabor Wavelet Features. Indian Journal of Science and Technology. 2022 Apr 29;15(16):712-7.
8. Vij S, Virmani D, Jain A. Retrospective inspection for research in natural language processing of Hindi language using fuzzy logic. InICT Infrastructure and Computing: Proceedings of ICT4SD 2022 2022 Nov 8 (pp. 303-309). Singapore: Springer Nature Singapore.
9. Srivastava S, Verma A, Sharma S. Optical character recognition techniques: A review. In2022 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) 2022 Feb 19 (pp. 1-6). IEEE.
10. Bhatia K, Agrawal A, Singh P, Singh AK. Detection of AI Synthesized Hindi Speech. arXiv preprint arXiv:2203.03706. 2022 Mar 7.
11. Mishra A, Shaikh SH, Sanyal R. Context similarity based hybrid approach for extracting hindi multiword expressions. Journal of Intelligent & Fuzzy Systems. 2022 Sep 22;43(5):5595-605.
12. Shelke R, Vanjale S. Recursive LSTM for the Classification of Named Entity Recognition for Hindi Language. Ingenierie des Systemes d'Information. 2022 Aug 1;27(4):679.
13. Sahu B, Joshi BK. Application of substring search algorithms to Hindi based on pattern recognition. In2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC) 2022 May 9 (pp. 893-895). IEEE.
14. Shokat S, Riaz R, Rizvi SS, Khan I, Paul A. Characterization of English braille patterns using automated tools and RICA based feature extraction methods. Sensors. 2022 Feb 25;22(5):1836.
15. Jana A, Venkatesh G, Yimam SM, Biemann C. Hypernymy Detection for Low-resource Languages: A Study for Hindi, Bengali, and Amharic. Transactions on Asian and Low-Resource Language Information Processing. 2022 Mar 4;21(4):1-21.
16. Agarwal S, Sukritin, Sharma A, Mishra A. Next word prediction using hindi language. InAmbient Communications and Computer Systems: Proceedings of RACCCS 2021 2022 May 8 (pp. 99-108). Singapore: Springer Nature Singapore.
17. Shelke R, Vanjale S. Deep Named Entity Recognition in Hindi Using Neural Networks. Revue d'Intelligence Artificielle. 2022 Aug 1;36(4).