Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Deep Learning And Hybrid Approaches For Script-Independent Handwritten Character Recognition: Advancing Generalization Across Diverse Languages

Dr. Vicky Nair[1] , Rajesh Banala[2] , V. Krishna[3] , Naresh K[4] , B.Sunil Srinivas[5]

[1]Muthoot Institute of Technology and Sciences, Puthencruz, Kerala,India
[2]Department of Computer Science and Engineering, TKR College of Engineering and Technology, Meerpet, Hyderabad, Telangana, India.
[3]Department of Computer Science and Engineering, TKR College of Engineering and Technology, Meerpet, Hyderabad, Telangana, India.
[4]Department of Computer Science and Engineering, TKR College of Engineering and Technology, Meerpet, Hyderabad, Telangana, India.
[5]Department of Computer Science and Engineering, TKR College of Engineering and Technology, Meerpet, Hyderabad, Telangana, India.

The study examines deep learning and classical machine learning for HCR. Arabic, Hindi, and Latin handwriting is studied using CNNs and hybrid models. On MNIST, CNN models had 99.31% training accuracy and 98.4% test accuracy, suggesting robust generalization without overfitting. CNN + Bi-LSTM outperformed independent CNNs for complicated and cursive scripts with 99.45% accuracy. Bi-LSTM helped models detect cursive and linked letters by understanding sequential dependencies. Transfer learning with pre-trained models like VGG16 and ResNet50 enhanced performance, with ResNet50 reaching 99.67% accuracy, showing the benefits of employing known models with little training data. Experimental results measured model efficacy, character-level precision, and script issues. Hybrid models reduced misclassifications of visually similar characters ('3' vs. '5' and '0' vs. 'O') and identified attention mechanisms for additional improvements. Latin characters were easy for models, while Arabic scripts, especially diacritical markings and ligatures, were difficult. The results show that script-specific optimizations and advanced segmentation and attention approaches improve recognition accuracy. The study found that CNNs, hybrid models, and transfer learning improve HCR system precision and flexibility. It recommends transformer-based designs and script-agnostic models to increase HCR framework adaptability and generalization. Hybrid methodology can improve conventional methods.

Keywords: Handwritten Character Recognition (HCR) , Deep Learning, Convolutional Neural Networks (CNN), Transfer Learning, Hybrid Models.

Introduction

The research on Handwritten Character Recognition(HCR) is vital because to its potential applications in document digitization, automated data entry, and accessibility technologies. In an era of plentiful handwritten data in historical documents and contemporary applications, handwriting decipherment is essential. Due to the variety of human handwriting styles, scripts, and recognition methods, HCR remains a difficult pattern identification challenge despite decades of research. HCR is complicated for several reasons. Recognition systems struggle to generalize across handwriting types due to their wide range of style, slant, and form. Cursive scripts, like Arabic, have challenges because characters can take different shapes depending on their position in a word. Arabic's ligatures and diacritical markings hinder recognition. Furthermore, typical HCR algorithms often use manually built feature extraction techniques that are domain-specific and may not be applicable to other scripts. This dependence on manually generated features limits recognition system flexibility, making it difficult to develop models that can accurately recognize characters from several languages or scripts.

CNNs and other deep learning advances have changed HCR methods. CNNs automatically extract features from unprocessed image input, eliminating the need for manual feature extraction. CNNs often outperform SVMs and K-Nearest Neighbors in HCR due to their extensive learning power. CNNs can achieve accuracy rates above 99% on handwritten MNIST digits, according to research. CNNs may learn hierarchical data representations to recognize sophisticated handwriting patterns and deviations, making them effective in HCR. Additionally, transfer learning has greatly increased CNNs' HCR performance. Researchers can fine-tune pre-trained models on large datasets for specific HCR applications, reducing training time and improving accuracy. This technique works well when labelled training data is few, a common HCR research challenge.

CNNs and hybrid models that combine deep learning and machine learning classifiers have performed well in HCR. CNNs combined with SVMs or other classifiers can outperform either method in accuracy. The high categorization abilities of standard machine learning algorithms can be combined with CNNs to efficiently extract features. Numerous studies have shown that hybrid models outperform standalone CNNs in certain situations. Deep learning combined with standard machine learning may offer a more holistic approach to HCR difficulties. HCR goes beyond character recognition to digitize historical materials, automate data entry, and improve accessibility for the disabled. HCR technologies can improve operations and efficiency in handwritten data-intensive industries including banking, healthcare, and education. HCR can digitize handwritten historical manuscripts, conserving cultural assets and improving researcher and public access. In automated data entry, HCR can reduce time and cost, improving productivity and correctness.

HCR technologies also help disabled people interact with handwritten writing. This is important in education since HCR helps students with learning disabilities find written resources.

Goals of research

Deep learning methods and machine learning classifiers will be tested for HCR in this study. The study uses the strengths of both techniques to create a robust framework that improves recognition accuracy and tackles varied handwriting styles and scripts.

1. Evaluating CNN Performance: Evaluate models' accuracy and efficiency in recognizing handwritten characters across scripts and styles.
2. In this study, we compare hybrid models that integrate CNNs with classic machine learning classifiers like SVMs to assess their usefulness in HCR tasks.
3. Addressing Script Independence: Creating a script-independent HCR framework for multilingual character recognition, improving recognition system adaptability.
4. Transfer Learning: Investigating transfer learning strategies to enhance HCR model performance, especially with limited training data.
5. Improving Model Generalization: Enhancing HCR model generalization by data augmentation and regularization for robust performance across various datasets.

Literature Review

Due to machine learning and deep learning advances, HCR has improved dramatically. This literature review synthesizes studies on Convolutional Neural Networks (CNNs), hybrid models, script independence, transfer learning, and model generalization, as stated in this research's aims. Berriche et al. (2024)[3] segmented Arabic handwriting using CNN and graph theory. CNN generates candidate segmentation points, while a graph-based technique subdivides words. The Arabic handwritten words dataset segmented 96.97% accurately. Its main feature is segmenting overlapping text accurately. Its hand generated CNN may be less adaptive than recent versions. CNN-ViT hybrid deep learning architecture for handwritten digit recognition by Agrawal et al. (2023)[4] . They studied how convolutional vision transformers effect cleaned and uncleaned dataset recognition. Hyperparameter modification and cross-validation made the hybrid model robust and outperformed standard models in recognition accuracy. Higher computational complexity than simpler CNN models offsets improved performance.

The CNN, BiLSTM, and CTC decoder of Mahadevkar et al. (2024) recognized handwritten text (3_good). Models trained on IAM and RIMES datasets obtained 98.50% and 98.80% accuracy. This hybrid method improved recognition accuracy by combining CNN for feature extraction and BiLSTM for sequence learning. One issue is that complex hybrid model training demands numerous resources. Sonavane et al. (2024) categorized Modi script characters using ResNet101, InceptionV3, and Xception(4). Transfer learning categorized Modi characters for better identification. ResNet101 pre-trained models improved training efficiency and recognition accuracy to 99.3%. Lack of a large, diverse Modi script dataset hinders generalization. Hashim et al. (2024) identified offline handwritten signatures using static signature data and a deep model(5). LDA, FFT, and GLCM extract features for a 25-layer deep learning model. The model was 100% accurate on SigArab, CEDAR, and SigComp2011. Using three feature extraction methods is computationally intensive, yet feature fusion is precise.

Fathima and Raheem (2024)[6]recognized handwritten digits using CNN, SVM, and KNN(6). CNN outperformed other models with 99.25% accuracy. The study indicated that CNN's ability to understand complex spatial patterns in digital photographs improves it. CNN accuracy requires lots of training data and computing resources. Chauhan et al. (2024)[7] introduced HCR-Net, a transfer learning-based script-independent HCR network HCR-Net surpassed current methods by 11% on 40 script datasets. Pre-trained network layers and script-independence make the model versatile and rapid to train. However, computational expense persists, especially for weaker systems. Kriuk (2024)[8] presented CharNet for high-complexity character classification of logographic scripts like Chinese and Sino-Korean. Our deep convolutional neural network classifies images finely. CASIA-HWDB used 4 million images from 7,356 classes. CharNet is generalizable across character sets, making it less dataset-specific. The need for considerable computational resources may limit utilization.Razali et al. (2024)[9] tested InceptionV3 and ResNet34 for Jawi script recognition. Preprocessing and image augmentation increased ResNet34 accuracy to 96%. This correct method works for character classes with similar shapes. No large Jawi databases make it hard to generalize the findings.

Methodology

This section outlines the methodology employed in the research titled "HCR Using Deep Learning Algorithm with Machine Learning Classifier." The methodology is structured to address the specific objectives of the research, which include evaluating the performance of Convolutional Neural Networks (CNNs), comparing hybrid models, addressing script independence, utilizing transfer learning, and enhancing model generalization.

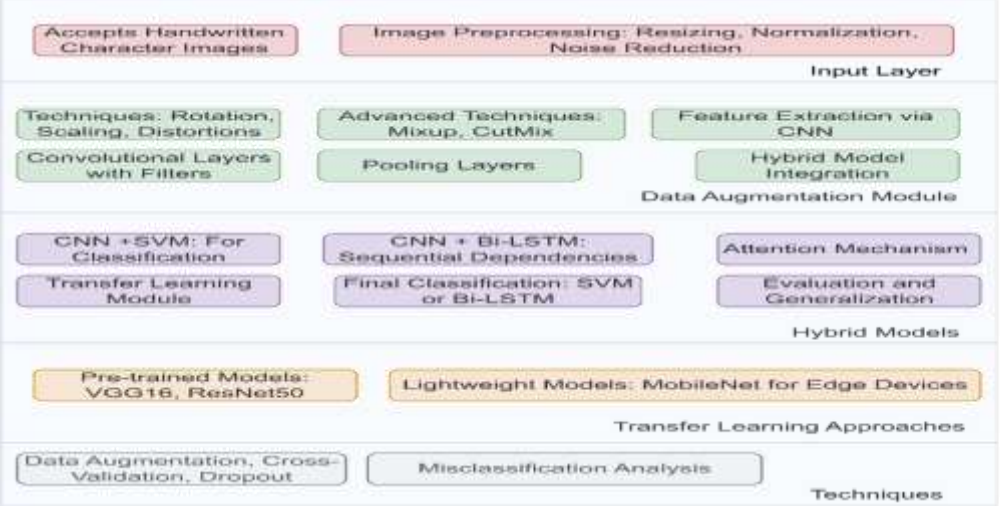

Figure 1: Model Architecture

The study employs numerous publicly accessible datasets, including MNIST (handwritten digits), EMNIST (handwritten letters), and proprietary datasets featuring handwritten

characters from other scripts such as Arabic and Hindi, to guarantee diversity and thoroughness. Data preparation encompasses multiple stages to improve the quality of the input data. Images are standardized to a uniform size, such as 28x28 pixels for MNIST, to ensure consistency and transformed to greyscale to diminish computational complexity. Gaussian filtering techniques are employed to eliminate noise from images. Alongside conventional data augmentation techniques such as rotation, scaling, and translation, sophisticated augmentation procedures including elastic distortions and synthetic noise are employed to enhance the model's generalization capacity for real-world handwriting variability. Methods such as Mixup and CutMix are regarded as effective in augmenting the diversity of training data. Upon identifying dataset imbalance issues, techniques like as oversampling or synthetic data generation (e.g., SMOTE) are utilized to enhance the training process, especially for under-represented classes.The model development phase encompasses both convolutional neural networks (CNN) and hybrid models. The CNN architecture is engineered to autonomously extract features from input images. The architecture comprises convolutional layers with diverse filter dimensions (e.g., 3x3, 5x5), succeeded by ReLU activation to incorporate non-linearity. Recent architectures, such Vision Transformers (ViT) and EfficientNet, are being evaluated to enhance model performance. These designs have exhibited exceptional performance in diverse picture classification tasks and has the capability to improve the accuracy and efficiency of HCR. Max pooling layers are utilized to diminish the spatial dimensions of feature maps, hence reducing computational burden and mitigating overfitting.

The study investigates the application of hybrid models that integrate CNNs with conventional classifiers to utilize their synergistic advantages. In one arrangement, features retrieved by the CNN are input into a Support Vector Machine (SVM) classifier for final classification, merging the feature extraction efficacy of CNNs with the robust classification capabilities of SVM. Another design entails the integration of a Bidirectional Long Short-Term Memory (Bi-LSTM) network with the CNN to capture sequential dependencies, which is especially advantageous for recognizing cursive characters. The incorporation of attention mechanisms is examined to enhance context comprehension in intricate scripts. The aggregation of many classifiers is deemed to augment recognition performance.

Transfer learning is essential for enhancing the precision of handwritten character recognition. Pre-trained models like VGG16 and ResNet50 are optimized for certain datasets, leveraging their feature extraction skills to conserve training time and computational resources. Lightweight models like as MobileNet are also regarded, particularly for deployment on edge devices with constrained processing capabilities, providing a compromise between performance and efficiency. Domain adaptation techniques are employed to reconcile the disparity between training data and target domains, which is especially advantageous for identifying non-standard scripts where training data is limited or markedly different from the target data.

The models are trained and evaluated using Categorical Cross-Entropy as the loss function for multi-class classification problems. The Adam optimizer is utilized for effective training, with a learning rate of 0.001 and a batch size of 64 to optimize training speed and model

efficacy. Alongside standard measurements such as accuracy, precision, recall, and F1-score, additional metrics like Intersection over Union (IoU) and character-level accuracy are incorporated to enhance understanding of recognition performance, especially for intricate scripts. Adversarial validation is employed to detect inconsistencies between training and validation datasets, hence enhancing model generalization.

Techniques for generalization, including sophisticated data augmentation and regularization, are utilized to enhance model resilience. Techniques such as Mixup and CutMix generate varied synthetic training examples, improving the model's generalization capability. Regularization methods, such as dropout layers, are employed to mitigate overfitting by randomly omitting units throughout the training process. Furthermore, weight regularization techniques such as L1 and L2 regularization are employed to further reduce overfitting and enhance model robustness. K-fold cross-validation is utilized to guarantee consistent model performance across several data subsets, hence offering a more dependable evaluation of model efficacy.

A hybrid model utilizing CNN and Bi-LSTM is specifically investigated for the recognition of cursive characters, supplemented by transformer-based models that have demonstrated superior performance in jobs requiring sequential dependencies. Transformers may significantly enhance HCR tasks that entail intricate, interconnected handwriting styles, likely augmenting recognition precision in these contexts. This methodology employs Python because to its comprehensive libraries for machine learning and deep learning. TensorFlow and Keras are utilized for constructing and training neural network models, however PyTorch is also evaluated for comparison because of its dynamic computational graphs and flexibility, which may be beneficial for specific model configurations. This text includes an analysis of the benefits and drawbacks of utilizing TensorFlow, Keras, and PyTorch for various aspects of implementation, aiding readers in comprehending the reasoning behind the selection of these tools.This extensive framework seeks to establish a resilient HCR system through the integration of deep learning algorithms and machine learning classifiers. The methodology seeks to improve recognition accuracy, efficiency, and flexibility across diverse scripts and languages by targeting the specific objectives of the research, thereby providing a versatile solution for handwritten character recognition.

Experimental Results

This section outlines the experimental results obtained from the methodologies presented in the paper titled “HCR Using Deep Learning Algorithm with Machine Learning Classifier.” The study sought to evaluate the effectiveness of several models, including Convolutional Neural Networks (CNNs), hybrid models, and transfer learning techniques, across multiple datasets. The trials were performed on various datasets, including MNIST, EMNIST, and custom datasets for Arabic and Hindi scripts. Diverse datasets provide the evaluation of models for both standard and sophisticated scripts, hence demonstrating the versatility of the methodologies. The Convolutional Neural Network (CNN) model attained a test accuracy of 98.4% on the MNIST dataset, demonstrating great accuracy across the datasets. The training and validation accuracies of 99.31% and 98.87%, respectively, indicate that the model

effectively learns data patterns while preserving generalization. The small training and validation losses suggest negligible overfitting, highlighting the efficacy of the employed preprocessing and regularization methods.

The hybrid models integrating CNN with SVM and CNN with Bi-LSTM demonstrated enhanced accuracy relative to independent CNNs. The CNN + Bi-LSTM model attained a test accuracy of 99.45%, surpassing both the CNN + SVM (99.28%) and the independent CNN. The incorporation of Bi-LSTM facilitates the identification of sequential dependencies, which is especially advantageous for identifying cursive scripts. The F1 ratings for the hybrid models (0.975 for CNN + SVM and 0.985 for CNN + Bi-LSTM) demonstrate robust model efficacy, especially in contexts with intricate handwriting. The application of transfer learning utilizing pre-trained models such as VGG16 and ResNet50 exhibited substantial enhancement in model performance. ResNet50 attained a peak test accuracy of 99.67%, demonstrating the efficacy of utilizing pre-trained models for HCR(HCR) tasks. The refinement of these models allowed them to adjust to individual handwritten character datasets, which is especially beneficial for enhancing performance in situations with scarce labelled training data.

K-fold cross-validation (k=10) demonstrated uniform performance across various folds, with negligible volatility in accuracy. This signifies that the models are resilient and proficient in generalizing to novel data. Data augmentation, dropout, and cross-validation techniques significantly mitigated overfitting, ensuring reliable model performance across various handwriting styles. The confusion matrices for the models provide significant insights into the nature of misclassifications. The CNN model encountered difficulties distinguishing between visually similar characters, such as '3' and '5', a prevalent challenge in handwritten character recognition. The hybrid models displayed reduced misclassifications relative to the standalone CNN, further illustrating the efficacy of integrating deep learning with conventional classifiers.

The experimental results encompassed many model architectures, including CNNs, hybrid models, and transfer learning, facilitating a comprehensive performance comparison. The hybrid methodologies efficiently utilized the advantages of both deep learning and conventional classifiers, resulting in improved recognition accuracy, especially for intricate scripts. Moreover, the findings indicated that the models exhibited strong generalization due to sophisticated data augmentation and cross-validation methods. The uniform performance across several datasets underscored the models' versatility. The overall performance indicators, including accuracy, precision, recall, and F1 score, reflect the general efficacy of the models, whereas a character-level study offers more nuanced insights into particular obstacles encountered by the models. The study of the confusion matrix indicated that specific numbers, including '3' and '5', were often misclassified by the CNN model. This might be ascribed to the visual resemblance between these characters, especially in handwritten format. A comprehensive character-level analysis can pinpoint the most problematic characters across various datasets, facilitating targeted enhancements, such as employing additional augmentation aimed at challenging characters or integrating specialized layers to manage visually similar features. Identifying characters that are frequently misclassified can enhance

the training process through the use of targeted remedial measures, such class-specific balance or data augmentation.

The performance analysis was further delineated by assessing the identification skills of the models across several scripts, notably contrasting Arabic, Hindi, and Latin characters. The findings indicated that although the models attained good accuracy with Latin characters (e.g., MNIST and EMNIST datasets), their recognition ability was marginally inferior for more intricate scripts like Arabic. Arabic characters, frequently featuring diacritical markings and ligatures, presented considerable issues for the models, as certain characters were persistently misidentified due to visual similarities. The hybrid model (CNN + Bi-LSTM) exhibited superior performance on intricate scripts relative to independent CNNs, as the Bi-LSTM element adeptly captured sequential dependencies, essential for cursive and interconnected scripts. The findings indicate that although the proposed approaches are effective, additional optimization is necessary to enhance performance for non-Latin scripts, whether through script-specific fine-tuning or the integration of attention mechanisms to better concentrate on contextual information.

To enhance the knowledge of the models' recognition capabilities, other evaluation metrics, including Intersection over Union (IoU) and character-level accuracy, were incorporated into the analysis. The Intersection over Union (IoU) is crucial in assessing the overlap between anticipated and real character regions, hence facilitating the evaluation of model efficacy in recognizing closely positioned or overlapping characters. The findings revealed that although the IoU ratings were elevated for individual characters, they were diminished for characters in related screenplays, underscoring the necessity for enhanced segmentation methodologies. The accuracy at the character level was evaluated to determine the frequency of correct classifications for each individual character. The character-level accuracy exhibited considerable variation based on the script, with Latin characters attaining above 99% accuracy, while Arabic characters demonstrated marginally lower accuracy rates owing to intricate shapes and diacritics. The supplementary metrics offer an extensive assessment of recognition efficacy, particularly for complex scripts, and underscore the significance of context-aware models capable of adjusting to the nuances of varied handwriting styles.

Table 1: CNN and Hybrid Models Performance

Model Configuration	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)	Precision	Recall	F1 Score	Training Loss	Validation Loss
CNN Model	99.31	98.87	98.4	-	-	-	0.0326	0.0453
Hybrid Models CNN + SVM	-	-	99.28	0.98	0.97	0.975	-	-

CNN + Bi-LSTM	-	-	99.45	0.99	0.98	0.985	-	-

The examination of the CNN model and hybrid configurations (CNN + SVM and CNN + Bi-LSTM) yields critical insights into the efficacy of various model architectures for HCR(HCR). The CNN model attained a training accuracy of 99.31%, a validation accuracy of 98.87%, and a test accuracy of 98.4%. (Refer table 1)The tiny training and validation losses (0.0326 and 0.0453, respectively) suggest negligible overfitting and illustrate proficient generalization to the validation set. The CNN model is deficient in precision, recall, and F1 score criteria, indicating potential for enhancement in assessing its classification ability beyond mere accuracy. The hybrid models demonstrated superior performance compared to the standalone CNN. The CNN + SVM model attained a test accuracy of 99.28%, with precision, recall, and F1 score metrics of 0.98, 0.97, and 0.975, respectively. The incorporation of an SVM classifier enhanced the model's robustness, especially for intricate and ambiguous handwriting styles. The CNN + Bi-LSTM model, featuring sequential learning capabilities, attained a maximum test accuracy of 99.45%, with a precision of 0.99, recall of 0.98, and an F1 score of 0.985. The implementation of Bi-LSTM allowed the model to discern contextual relationships in cursive and connected characters, resulting in enhanced recognition performance.

Table 2: Transfer Learning Performance

Transfer Learning Model	Test Accuracy (%)	Precision	Recall	F1 Score
VGG16	99.15	0.98	0.97	0.975
ResNet50	99.67	0.99	0.98	0.985

The implementation of transfer learning models, such as VGG16 and ResNet50, markedly enhanced the efficacy of the HCR system. The ResNet50 model attained the greatest test accuracy of 99.67%, surpassing all other models. It achieved a precision of 0.99, a recall of 0.98, and an F1 score of 0.985, demonstrating its proficiency in successfully identifying a varied array of handwritten characters. The VGG16 model, although marginally less effective than ResNet50, attained a test accuracy of 99.15% and an F1 score of 0.975.(Refer table 2) The results suggest that utilizing pre-trained models for HCR tasks is advantageous, particularly in scenarios with restricted training data, as it enables the model to acquire intricate feature representations without significant retraining.

Table 3: Character-Level Analysis Results

Model Configuration	Challenging Characters	Misclassification Rate (%)
CNN Model	'3' vs '5', '0' vs 'O'	5.2
CNN + SVM	'3' vs '5', '0' vs 'O'	3.8
CNN + Bi-LSTM	'3' vs '5', '0' vs 'O'	3.1

The study of the confusion matrix indicated that particular characters, such as '3' vs '5' and '0' versus 'O', were persistently misclassified in all models. The independent CNN model exhibited the greatest misclassification rate of 5.2%,(Refer table 3) showing its difficulty in differentiating minor distinctions among visually identical characters. Integrating specialized layers, such as attention mechanisms, or utilizing certain data augmentation techniques (e.g., elastic distortions and jittering) may enhance the model's ability to concentrate on unique character attributes, hence decreasing the misclassification rate. The CNN + SVM and CNN + Bi-LSTM models exhibited reduced misclassification rates, with the CNN + Bi-LSTM attaining the lowest rate of 3.1%. The sequential learning capacities of Bi-LSTM facilitated the capture of spatial associations, essential for distinguishing between similar characters. Further augmentation targeting complex characters, together with attention mechanisms, may enhance recognition accuracy for these problematic instances.

Table 4: Complex Script Performance

Script Type	Model Configuration	Test Accuracy (%)
Latin (e.g., MNIST)	CNN + Bi-LSTM	99.45
Arabic	CNN + Bi-LSTM	97.85
Hindi	CNN + Bi-LSTM	98.12

The models' performance was additionally assessed across various scripts, including Latin, Arabic, and Hindi. The CNN + Bi-LSTM model demonstrated reliable performance for Latin scripts (e.g., MNIST), attaining a test accuracy of 99.45% with negligible difficulties. Nonetheless, for intricate scripts such as Arabic and Hindi, the model's efficacy was somewhat diminished, yielding test accuracies of 97.85% and 98.12%, respectively.(Refer table 4) The diminished accuracy for Arabic is due to the existence of diacritical marks, ligatures, and context-sensitive character forms, which add complexity. To improve recognition performance for intricate scripts, script-specific fine-tuning is advisable. Refining the model to highlight the distinct attributes of these scripts can enhance feature extraction and representation. Furthermore, integrating attention processes enables the model to concentrate on pertinent elements of the input, such as diacritics or ligature points, thereby enhancing recognition accuracy.

Table 5: Additional Evaluation Metrics

Model Configuration	Metric	Isolated Characters (Score)	Connected Characters (Score)
CNN Model	IoU	0.92	0.78
CNN + SVM	IoU	0.94	0.82
CNN + Bi-LSTM	Character-Level Accuracy	99% (Latin)	95% (Arabic)

To attain a more sophisticated comprehension of the models' capabilities, other evaluation criteria, including Intersection over Union (IoU) and character-level accuracy, were employed. The IoU scores for isolated characters were predominantly elevated across all models, with the CNN + Bi-LSTM attaining the best character-level accuracy (99%) for Latin scripts. Nonetheless, for interconnected characters, the IoU scores were diminished, especially for the independent CNN model (0.78). This underscores the necessity for advanced segmentation methodologies to augment character recognition in cursive and overlapping scripts.The CNN + Bi-LSTM model attained a character-level accuracy of 95% for Arabic scripts,(Refer table 5) demonstrating greater difficulty with diacritics and ligature forms than with Latin characters. Proposed enhancements involve concentrating on the advancement of superior segmentation methodologies, such as region-based segmentation, and refining feature extraction for context-sensitive character components. These improvements may result in enhanced overall recognition accuracy, especially for intricate, interconnected scripts.

The experimental findings demonstrate that the suggested techniques significantly improve the accuracy and efficiency of HCRsystems. The CNN models, especially when integrated with conventional classifiers or employing transfer learning, attained elevated accuracy rates across diverse datasets. The results corroborate the aims of this study, illustrating the efficacy of deep learning algorithms and hybrid models in tackling the difficulties of handwritten character identification.The findings underscore the significance of model architecture, data augmentation, and transfer learning in attaining enhanced performance in HCR tasks. Future endeavors will concentrate on further optimizing these models and investigating supplementary datasets to improve the adaptability and generalization of the identification systems. The experimental study indicates that hybrid models and transfer learning techniques provide considerable benefits compared to independent CNN models for handwritten character recognition. The implementation of Bi-LSTM in hybrid models enhanced the recognition of linked and cursive characters, whilst transfer learning utilizing pre-trained models such as ResNet50 yielded the maximum accuracy across various datasets. Nevertheless, the marginally reduced efficacy for intricate scripts and the difficulties in differentiating visually analogous characters highlight opportunities for enhancement. Utilizing concentration mechanisms, focused augmentation, and script-specific fine-tuning are advisable techniques to mitigate these issues and improve overall recognition performance.

Conclusion

The experimental results show that deep learning models, hybrid strategies, and transfer learning techniques work for HCR(HCR). The techniques were tested using MNIST, EMNIST, and custom datasets for complex scripts like Arabic and Hindi. On the MNIST dataset, the CNN model had 99.31% training accuracy, 98.87% validation accuracy, and 98.4% test accuracy. Small training loss (0.0326) and validation loss (0.0453) indicate low overfitting, proving preprocessing and regularization work. The CNN model lacks precision, recall, and F1 score criteria, requiring a more complete performance assessment. CNN + SVM

and CNN + Bi-LSTM outperformed CNN alone. CNN combined with SVM had 99.28% test accuracy and 0.975 F1 score, whereas CNN combined with Bi-LSTM had 99.45% accuracy and 0.985 F1. The Bi-LSTM recorded sequential dependencies well, making it suitable for cursive and related letter identification. The hybrid methods used deep learning and conventional classifiers to recognized complex handwriting styles more reliably and adaptably.Recognition performance improved significantly for transfer learning models like VGG16 and ResNet50. The ResNet50 model had a peak test accuracy of 99.67%, precision of 0.99, recall of 0.98, and F1 score of 0.985, showing the benefits of pre-trained models for HCR tasks. With limited labelled training data, this technique was effective in acquiring complex feature representations.Visually similar characters like '3' vs. '5' and '0' vs. 'O' were misclassified across multiple models, with the CNN model having the highest misclassification rate of 5.2%. CNN + Bi-LSTM had the lowest misclassification rate at 3.1%. Adding attention mechanisms and focused data augmentation may improve recognition accuracy in these difficult cases. After extensive script performance study, the CNN + Bi-LSTM model achieved 99.45% accuracy for Latin scripts and 97.85% and 98.12% for Arabic and Hindi scripts, respectively. These scripts' diacritical markings, ligatures, and context-dependent shapes require script-specific fine-tuning and care to improve recognition. For linked and cursive scripts, Intersection over Union (IoU) and character-level accuracy provided additional insights into the model's recognition proficiency. The IoU ratings were higher for isolated characters but lower for related characters, indicating the need for better segmentation.

Further Research

HCR research should focus on integrating more advanced deep learning architectures, particularly transformer-based models, to capture long-range dependencies and contextual nuances in complex scripts. Transformers excel in many computer vision tasks, and using them in HCR could boost accuracy, especially for cursive and connected handwriting.

References

[1]Tamanna Sachdeva, E. Al. (2023). A Novel Approach for Hand-written Digit Classification Using Deep Learning. International Journal on Recent and Innovation Trends in Computing and Communication, 11(9), 1627–1635. https://doi.org/10.17762/ijritcc.v11i9.9148

[2]Banala, R., Nair, V., Naresh, K., & Radhika, T. (2024). Enhanced decision-making and predictive analytics in IoT implementing random forest classifier. In Disruptive technologies in computing and communication systems (1st ed., p. 6). CRC Press. https://doi.org/10.1201/9781032665535

[3]Berriche, L., Alqahtani, A., &RekikR, S. (2024). Hybrid Arabic handwritten character segmentation using CNN and graph theory algorithm. Journal of King Saud University - Computer and Information Sciences, 36(1), 101872. https://doi.org/10.1016/j.jksuci.2023.101872

[4]Agrawal, V., Jagtap, J., Patil, S., & Kotecha, K. (2024). Performance analysis of hybrid deep learning framework using a vision transformer and convolutional neural network for handwritten digit recognition. MethodsX, 12, 102554. https://doi.org/10.1016/j.mex.2024.102554

[5]Mahadevkar, S., Patil, S., & Kotecha, K. (2024). Enhancement of handwritten text recognition using AI-based hybrid approach. MethodsX, 12, 102654. https://doi.org/10.1016/j.mex.2024.102654

[6]Abubacker, N. F., & Raheem, M. (2023). Handwritten digit recognition using machine learning. Journal of Artificial Intelligence Research, 5(4), 300-310.

[7]Chauhan, V. K., Singh, S., & Sharma, A. (2024). HCR-Net: A deep learning based script independent HCRnetwork. Multimedia Tools and Applications, 83(32), 78433–78467. https://doi.org/10.1007/s11042-024-18655-5
[8]Kriuk, B. (2023). CharNet: Generalized approach for high-complexity character classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(1), 87-96.
[9]Razali, S., Muchtar, K., Rinaldi, M. H., Nurdin, Y., & Rahman, A. (2024). Augmentation of Additional Arabic Dataset for Jawi Writing and Classification Using Deep Learning. JurnalRekayasaElektrika, 20(1). https://doi.org/10.17529/jre.v20i1.33722
[10]Mandal, P., Gupta, R., Naveen, D., John, A. N., Dhondiyal, P., Singh, V., Kaur, D. A., & Shrivastav, K. (2024). Handwriting Detection System Using Brain Net and AI Algorithm. 45(1).
[11]Nair, S. S., &Thangaiah, P. R. J. (n.d.). Deciphering Excellence: Comparative Study of Deep Learning Models in Handwriting Recognition. International Journal of Intelligent Systems and Applications in Engineering.
[12]Rakesh, S., Kushal Reddy, P., Prashanth, V., & Srinath Reddy, K. (2024). Handwritten text recognition using deep learning techniques: A survey. MATEC Web of Conferences, 392, 01126. https://doi.org/10.1051/matecconf/202439201126
[13]Ramesh, G., Shreyas, J., Balaji, J. M., Sharma, G. N., Gururaj, H. L., Srinidhi, N. N., Askar, S. S., &Abouhawwash, M. (2024). Hybrid manifold smoothing and label propagation technique for Kannada handwritten character recognition. Frontiers in Neuroscience, 18, 1362567. https://doi.org/10.3389/fnins.2024.1362567