International Journal of Computer Applications (0975 - 8887)
Volume 187 - No.51, October 2025

One Cycle Policy-Enhanced Transfer Learning for Robust Bengali Handwritten Character Recognition

Raihan Tanvir
Ahsanullah University of Science and Technology
Dhaka, Bangladesh

ABSTRACT

The recognition of handwritten Bengali characters presents significant challenges due to the script's morphological complexity, including visually similar conjunct characters and subtle distinguishing features such as dots or lines, compounded by variations in handwriting styles. To address these challenges, an efficient transfer learning framework is proposed for Bengali Handwritten Character Recognition (BHCR), leveraging advanced deep learning methodologies. A pretrained ResNet-50 model, originally trained on the ImageNet dataset, is fine-tuned with the One Cycle Policy for cyclic learning rate optimization to expedite convergence. Evaluation on the Ekush dataset, comprising 367,018 isolated handwritten characters, demonstrates that the proposed method achieves an accuracy of 95.78% after 50 epochs, surpassing many comparable techniques in efficiency while minimizing architectural overhead.

General Terms

Handwritten Character Recognition, Transfer Learning, Deep Learning, Computer Vision

Keywords

Bengali Handwritten Character Recognition, Convolutional Neural Networks, ResNet, Ekush Dataset, One Cycle Policy, Cyclic Learning Rate

1. INTRODUCTION

The Bengali language, ranked as the sixth most widely spoken language globally, serves as the primary language of Bangladesh and the second most prevalent in India, with over 200 million native speakers. Its cultural and historical significance is recognized globally, notably through UNESCO's designation of February 21 as International Mother Language Day, commemorating the 1952 language martyrs of Bangladesh. Accurate recognition of handwritten Bengali characters constitutes a critical research area with transformative applications in handwritten character recognition (HCR), optical character recognition (OCR), and automated document processing systems, facilitating enhanced accessibility and digitization of Bengali texts.

Recognizing handwritten Bengali characters poses unique challenges compared to scripts like English due to the complexity of the Bengali writing system. The script encompasses basic vowels and consonants, as well as intricate conjunct-consonant characters formed by combining multiple basic characters, often distinguished by subtle features such as a single dot or a fine stroke. These morphological complexities, combined with diverse handwriting styles, result in significant variability, leading to lower performance for Bengali HCR compared to its English counterpart [6]. Traditional approaches often struggle to generalize across these variations, necessitating advanced methodologies to achieve robust performance.

To address these challenges, a transfer learning-based approach is proposed, leveraging state-of-the-art deep learning techniques. The ResNet-50 architecture [4], pretrained on the ImageNet dataset [3], is utilized to capitalize on its ability to extract generic and task-specific features. The One Cycle Policy [10], a cyclic learning rate strategy, is incorporated to dynamically adjust the learning rate within a single cycle, promoting rapid convergence and improved model performance. This approach minimizes training iterations while maximizing accuracy, ensuring computational efficiency for BHCR tasks.

2. RELATED WORK

Recent advancements in deep learning have significantly advanced handwritten character recognition, particularly through the application of transfer learning and convolutional neural networks (CNNs). These techniques leverage pretrained models to enhance performance on new tasks with minimal retraining, proving highly effective across diverse scripts and languages.

Ciresan et al. [2] investigated transfer learning for character recognition, focusing on transferring knowledge from Latin digits to uppercase Latin letters and extending the approach to Chinese characters. Their findings demonstrated that transfer learning enables near-perfect classification with limited retraining, highlighting its potential for cross-script recognition tasks. This principle is applied to the Bengali script in the proposed methodology.

Tushar et al. [11] explored transfer learning for numeral recognition across Bengali, Hindi, and Urdu scripts. By training a CNN on numerals from one language and fine-tuning it for another, high accuracy was achieved with minimal additional training. These results emphasize the adaptability of transfer learning for scripts with shared characteristics, informing the approach to leveraging pretrained models for Bengali handwritten character recognition (BHCR).

Rabby et al. [9] developed BornoNet, a lightweight 13-layer CNN designed for recognizing 50 basic Bengali characters (11 vowels and 39 consonants). Utilizing the ADAM optimizer [5] and an automatic learning rate reduction technique, BornoNet achieved validation accuracies of 98%, 96.81%, 95.71%, and 96.40% on the CMATERdb, ISI, BanglaLekha-Isolated [1], and mixed datasets, respectively. This work highlights the efficacy of tailored CNN architectures for BHCR, though the proposed methodology extends this by incorporating transfer learning with a deeper ResNet-50 model to address a broader range of character classes. Earlier work by Pal et al. [7] focused on English handwritten character recognition using multilayer perceptrons (MLPs) with preprocessing techniques such as skeletonization, normalization, and Fourier transforms. While effective for English, this approach underscores the limitations of traditional methods for complex scripts like Bengali, which require more sophisticated deep learning techniques to handle morphological variability.

Collectively, these works demonstrate the transformative potential of transfer learning in character recognition, particularly for morphologically complex scripts. The proposed research builds upon these foundations by integrating transfer learning with ResNet-50 and optimized training strategies to achieve high accuracy in BHCR, addressing the unique challenges posed by the Bengali script.

3. METHODOLOGY

Advancements in deep learning and optimization techniques provide robust solutions for addressing the complexities of Bengali Handwritten Character Recognition (BHCR). The proposed methodology leverages transfer learning, hyperparameter optimization, and data augmentation to develop an efficient and accurate recognition system tailored to the morphological intricacies of the Bengali script.

3.1 Dataset

The experiments utilize the Ekush dataset [8], the most extensive publicly available dataset for Bengali handwritten character recognition, comprising 367,018 isolated handwritten characters from 3,086 unique contributors. The dataset includes vowels, consonants, modifiers, compound characters, and numerals, with each image standardized to a resolution of 28×28 pixels. To ensure diversity, the dataset is balanced across gender and categorized by age groups. Due to computational constraints, the male-contributed subset, containing 180,289 samples, is utilized, split into training and validation sets in an 80:20 ratio (144,232 training images and 36,057 validation images). For testing, 20% of the female-contributed subset is used to evaluate the model's

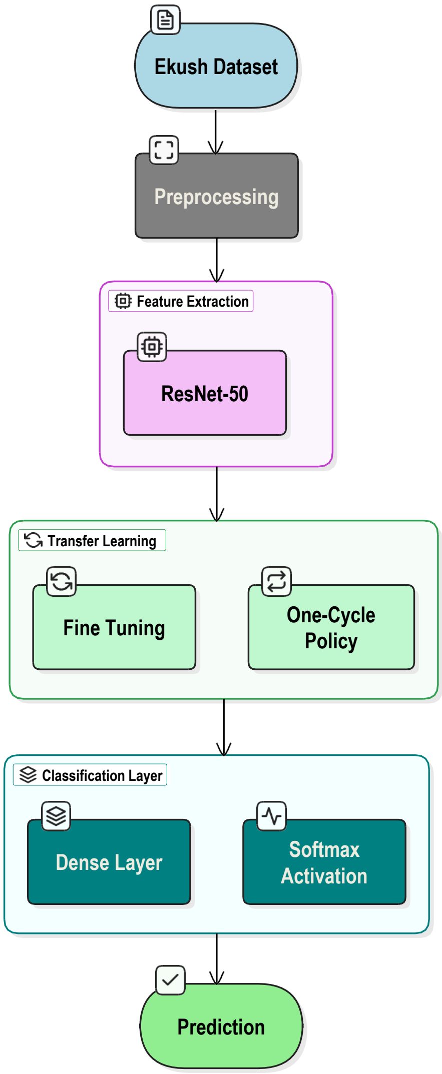

Fig. 1. Architecture diagram of the proposed methodology for BHCR.

generalization to unseen data. A statistical overview of the dataset is provided in Table 1, and representative samples with class labels are shown in Figure 2.

Data Augmentation. To enhance model generalization across diverse handwriting styles, data augmentation techniques are applied, including random rotations, flips, and brightness adjustments. These transformations simulate variations in handwriting, such as differences in stroke thickness, orientation, and lighting conditions, improving the model's robustness to real-world variability in the Ekush dataset [8]. By increasing dataset diversity through these augmentations, the risk of overfitting is mitigated, and performance on unseen data is improved, ensuring effective handling of the inherent variability in handwritten Bengali characters.

Table 1. Statistical Overview of the Ekush Dataset.

Basic	Compound	Modifiers	Numerals	Total Samples	Total Classes	Samples per Class
154,824	150,840	30,667	30,687	367,018	122	1507

Fig. 2. Representative images from the Ekush dataset with corresponding class labels.

3.2 Transfer Learning Framework

Transfer learning is employed to leverage knowledge acquired from a pretrained model, enhancing performance on BHCR with limited training data. Unlike traditional isolated learning, which trains models from scratch for specific tasks, transfer learning reuses features and weights from a model trained on a large, general dataset, adapting them to a new, related task. This approach is particularly effective for BHCR due to the script's complex character set, including basic characters, conjuncts, and modifiers, as well as variability in handwriting styles.

The ResNet-50 architecture [4], pretrained on the ImageNet dataset [3], is adopted as the foundation for the model. ResNet-50, a 50-layer convolutional neural network (CNN), incorporates residual connections that mitigate the vanishing gradient problem, enabling the training of deep networks without performance degradation. These connections allow layers to learn residual functions, ensuring that deeper architectures maintain or improve accuracy compared to shallower ones. The architecture is illustrated in Figure 1.

In CNNs, early layers extract generic features, such as edges and textures, while later layers capture task-specific patterns. By fine-tuning the pretrained ResNet-50, generic feature extraction capabilities are preserved in the initial layers, while the final layers are adapted to recognize Bengali-specific character patterns.

3.3 Training Optimization

To enhance training efficiency, the One Cycle Policy [10], a cyclic learning rate scheduling technique, is implemented to optimize the training process. This method dynamically adjusts the learning rate within a single cycle, starting with a low learning rate, gradually increasing it to a maximum value, and then decreasing it again. This approach, associated with super-convergence, accelerates training by enabling the model to explore the loss landscape more effectively, facilitating faster convergence to optimal model parameters with fewer epochs. The One Cycle Policy also incorporates a momentum adjustment, where momentum is inversely proportional to the learning rate, further stabilizing training and improving performance. The learning rate is determined using a learning rate finder, which systematically evaluates the loss landscape by testing a range of learning rates and identifying the optimal rate that minimizes validation loss while maximizing accuracy. This ensures efficient training without requiring extensive hyperparameter tuning, making it suitable for computationally intensive tasks like BHCR.

The training process is conducted in two phases. In the first phase, the earlier layers of ResNet-50 are frozen to retain their pretrained feature extraction capabilities, while the final layers are fine-tuned for 30 epochs to adapt to BHCR-specific features. In the second phase, all layers are unfrozen, and the learning rate finder is used to select an optimal learning rate, followed by retraining for an additional 20 epochs. This two-phase approach ensures efficient adaptation of the model to the target task while leveraging the robustness of the pretrained architecture.

3.4 Evaluation Strategy

Model performance is evaluated using accuracy as the primary metric, supplemented by confusion matrix analysis to assess classification performance across the 122 character classes in the Ekush dataset. Additionally, top misclassified samples (top losses) are identified to pinpoint challenging characters, providing insights for further refinement. This comprehensive methodology, integrating transfer learning, cyclic learning rate scheduling, data augmentation, and robust evaluation, aims to achieve state-of-the-art performance in BHCR while addressing the script's unique challenges.

4. EXPERIMENTS

4.1 Evaluation Metrics

Model performance is evaluated using accuracy, defined as the ratio of correctly classified samples to the total number of samples:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (1)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

Table 2. Loss and Accuracy for Initial Training with Frozen Layers.

Epoch	Train Loss	Validation Loss	Accuracy
0	3.925877	3.331836	0.237929
1	3.086453	2.695943	0.339130
2	2.320579	1.913326	0.494273
⋮	⋮	⋮	⋮
27	0.438727	0.352028	0.900935
28	0.433550	0.348528	0.901878
29	0.447053	0.348920	0.901240

4.2 Training Procedure

The training process is conducted in two phases to optimize the ResNet-50 model for BHCR. In the first phase, the initial layers of the pretrained ResNet-50 are frozen to preserve their generic feature extraction capabilities, while the final layers are fine-tuned for 30 epochs. Training begins with a high training loss of 3.925877 and an accuracy of 0.2379, which improve significantly to a training loss of 0.447053 and an accuracy of 0.9012 by the end of this phase. The convergence behavior is illustrated in Figure 3, showing the training loss initially exceeding the validation loss but gradually aligning, indicating effective learning without overfitting. Detailed performance metrics for this phase are presented in Table 2.

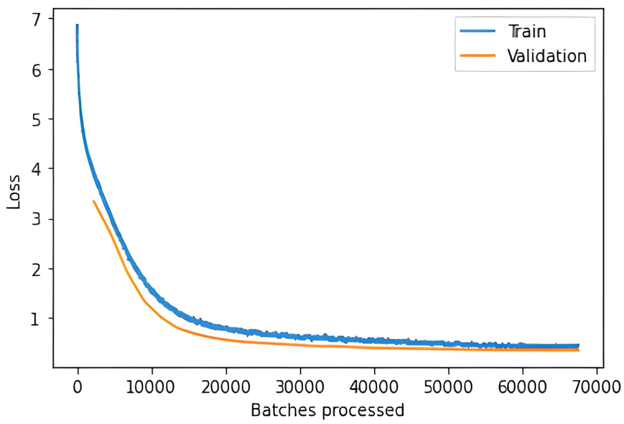

Fig. 3. Loss curves for initial training with frozen layers, showing convergence trends.

In the second phase, all layers are unfrozen, and a learning rate finder is employed to identify an optimal learning rate, as depicted in Figure 4. This graph illustrates the relationship between learning rate and loss, guiding the selection of a rate that minimizes validation loss while maximizing accuracy. The model is retrained for an additional 20 epochs, achieving a final accuracy of 0.957843 and a validation loss of 0.302539. Performance metrics for this phase are detailed in Table 3.

Table 3. Loss and Accuracy for Training with Unfrozen Network.

Epoch	Train Loss	Validation Loss	Accuracy
0	0.442622	0.350906	0.901711
1	0.441207	0.345438	0.902626
2	0.410408	0.344464	0.902210
⋮	⋮	⋮	⋮
17	0.318165	0.306372	0.914164
18	0.362462	0.303969	0.914469
19	0.323438	0.305791	0.913859

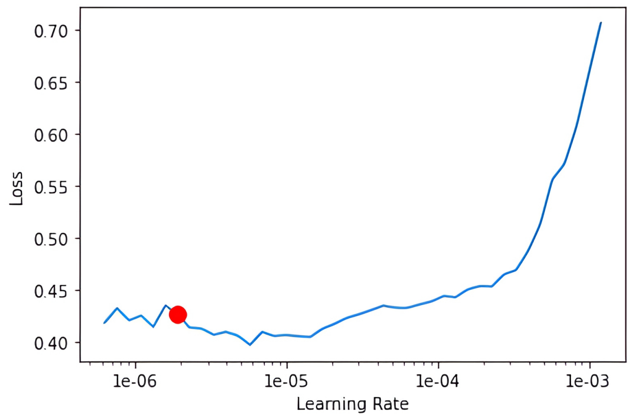

Fig. 4. Learning rate finder results for optimal learning rate selection.

4.3 Results

The two-phase training approach, combining transfer learning with optimized learning rate scheduling, yields a final accuracy of 95.78% on the Ekush dataset, demonstrating significant improvements in BHCR performance. Data augmentation enhances model robustness, enabling effective generalization to unseen handwriting styles. The confusion matrix and top loss analysis reveal specific characters prone to misclassification, often due to visual similarities, suggesting targeted areas for future refinement. Figure 5 illustrates sample outputs from the model, showcasing correct predictions for a selection of handwritten Bengali characters from the test set. This methodology underscores the efficacy of integrating transfer learning, cyclic learning rates, and robust evaluation strategies for tackling the complexities of BHCR. Given the large number of classes, top misclassified samples (top losses) are examined to identify characters that pose significant challenges, guiding potential improvements. The most frequently misclassified characters are highlighted in Figure 6, providing insights into areas for further optimization.

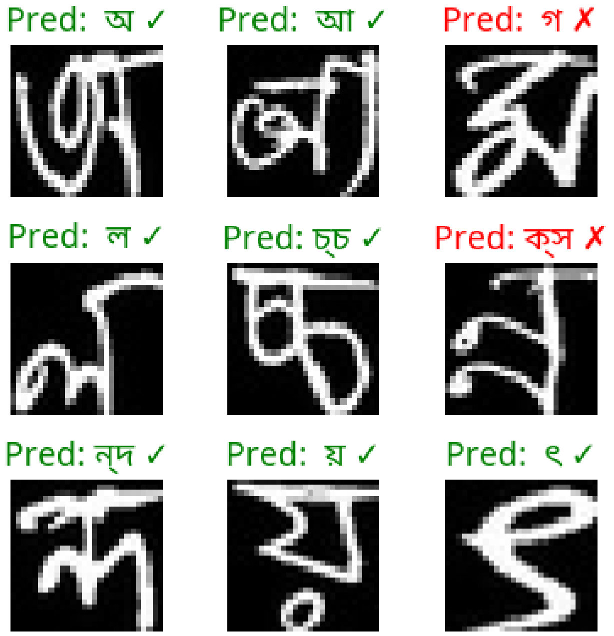

Fig. 5. Sample outputs from the model, showing predicted characters from the test set.

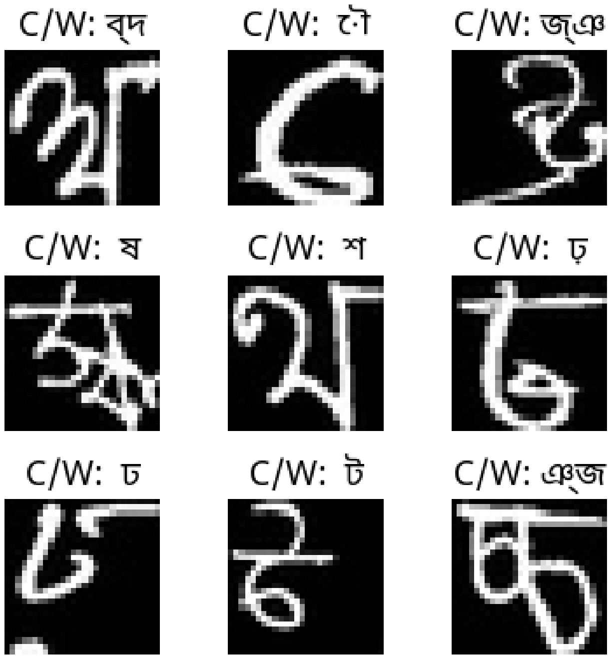

Fig. 6. Top misclassified characters, highlighting challenging predictions. C/W indicates the confused with character.

5. CONCLUSION AND FUTURE WORK

A robust transfer learning-based approach is presented to address the intricate challenges of Bengali Handwritten Character Recognition (BHCR), achieving an accuracy of 95.78% in 50 epochs using a pretrained ResNet-50 model fine-tuned with the One Cycle Policy. The high accuracy, attained despite training on a subset of the Ekush dataset due to computational constraints, underscores the efficacy of leveraging pretrained convolutional neural networks (CNNs) and advanced optimization techniques to tackle the morphological complexity and variability inherent in the Bengali script. Data augmentation techniques, such as random rotations and brightness adjustments, enhance the model's ability to generalize across diverse handwriting styles, while the One Cycle Policy ensures rapid convergence with minimal computational overhead. The confusion matrix and top loss analysis provide valuable insights into misclassified characters, often due to visual similarities, highlighting specific areas for refinement. These results demonstrate the power of transfer learning in addressing complex computer vision tasks, particularly for under-resourced scripts like Bengali, where robust recognition systems can significantly advance digitization efforts and accessibility in applications such as optical character recognition (OCR) and automated document processing.

Future work can further enhance the performance and applicability of this approach. Training the model on the full Ekush dataset could improve accuracy and robustness by capturing a broader range of handwriting variations across all contributors. Exploring alternative deep learning architectures, such as Vision Transformers or EfficientNet, may yield improved feature extraction capabilities for BHCR, potentially addressing the misclassification of visually similar conjunct characters. Incorporating advanced data augmentation techniques, such as generative adversarial networks (GANs) to synthesize additional handwritten samples, could further mitigate overfitting and enhance generalization. Extending the framework to recognize connected or cursive Bengali text could broaden its applicability to real-world scenarios like handwritten document transcription. Integrating the model into practical applications, such as mobile-based OCR systems for Bengali, could facilitate real-time text recognition and contribute to the preservation and accessibility of Bengali cultural heritage. This work lays a strong foundation for advancing Bengali HCR and paves the way for addressing real-world challenges in automated text recognition for complex scripts.

6. REFERENCES

[1] Mithun Biswas, Rafiqul Islam, Gautam Kumar Shom, Md. Shopon, Nabeel Mohammed, Sifat Momen, and Anowarul Abedin. Banglalekha-isolated: A multi-purpose comprehensive dataset of handwritten bangla isolated characters. *Data in Brief*, 12:103 – 107, 2017.

[2] D. C. Cireşan, U. Meier, and J. Schmidhuber. Transfer learning for latin and chinese characters with deep neural networks. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6, 2012.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.

[5] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[6] Farjana Omee, Shiam Himel, and Md. Abu Naser Bikas. A complete workflow for development of bangla ocr. *International Journal of Computer Applications*, 21:1–6, 05 2011.

[7] Anita Pal and Dayashankar Singh. Handwritten english character recognition using neural network. *International Journal of Computer Science & Communication*, 1(2):141–144, 2010.

[8] AKM Shahariar Azad Rabby, Sadeka Haque, Md. Sanzidul Islam, Sheikh Abujar, and Syed Akhter Hossain. Ekush: A multipurpose and multitype comprehensive database for online off-line bangla handwritten characters. In K. C. Santosh and Ravindra S. Hegadi, editors, *Recent Trends in Image Processing and Pattern Recognition*, pages 149–158, Singapore, 2019. Springer Singapore.

[9] Akm Shahariar Azad Rabby, Sadeka Haque, Sanzidul Islam, Sheikh Abujar, and Syed Akhter Hossain. Bornonet: Bangla handwritten characters recognition using convolutional neural network. *Procedia Computer Science*, 143:528 – 535, 2018. 8th International Conference on Advances in Computing & Communications (ICACC-2018).

[10] Leslie N. Smith. A disciplined approach to neural network hyper-parameters: Part 1 - learning rate, batch size, momentum, and weight decay. *CoRR*, abs/1803.09820, 2018.

[11] Abdul Kawsar Tushar, Akm Ashiquzzaman, Afia Afrin, and Md. Rashedul Islam. A novel transfer learning approach upon hindi, arabic, and bangla numerals using convolutional neural networks. In D. Jude Hemanth and S. Smys, editors, *Computational Vision and Bio Inspired Computing*, pages 972–981, Cham, 2018. Springer International Publishing.