A Hybrid CNN-RNN Model for Automated Recognition of Kannada Characters in Ancient Inscriptions

B. K. Rajithkumar
RV College of Engineering, Bengaluru, India | Visvesvaraya Technological University, Belagavi, India
rajith.bkr@rvce.edu.in (corresponding author)

B. V. Uma
RV College of Engineering, Bengaluru, India | Visvesvaraya Technological University, Belagavi, India
umabv@rvce.edu.in

H. S. Mohana
Rajeev Institute of Technology, Hassan, India |Visvesvaraya Technological University, Belagavi, India
mohana@rediffmail.com

Received: 2 August 2024 | Revised: 19 September 2024 | Accepted: 22 September 2024

 DOI: https://doi.org/10.48084/etasr.8602

ABSTRACT

This study presents a novel approach for the automated recognition of Kannada characters in ancient inscriptions using a hybrid Convolutional Neural Network and Recurrent Neural Network (CNN-RNN) model. The unique features of stone inscriptions, such as erosion, uneven surfaces, and varying font styles, pose significant challenges to traditional character recognition systems. The proposed hybrid model leverages the strengths of CNNs for feature extraction and RNNs for sequence prediction, enabling robust recognition of complex and degraded characters. The proposed model was trained and tested on a curated dataset of annotated Kannada inscriptions, achieving an impressive accuracy of 95%. This high accuracy demonstrates the model's effectiveness in deciphering ancient scripts, which is critical for the preservation and study of historical texts. The results highlight the potential of deep learning techniques in advancing the field of epigraphy and cultural heritage preservation.

Keywords-kannada character recognition; image processing; epigraphy

I. INTRODUCTION

Kannada stone inscriptions, dating back several centuries, are invaluable cultural artifacts that offer rich insights into the historical, linguistic, and social heritage of the Indian subcontinent [1]. These inscriptions, often etched into temple walls, pillars, and other stone surfaces, contain a wealth of information, including royal decrees, religious edicts, and records of social customs. However, manual transcription and interpretation of these ancient texts present significant challenges due to the effects of weathering, erosion, and the variability in the writing styles used by different scribes over time. The recognition and digitization of Kannada inscriptions are crucial to preserving this cultural heritage and making it accessible to researchers and the public [2]. Traditional methods, which heavily rely on expert analysis, are time-consuming and prone to errors, especially when dealing with degraded inscriptions. The development of automated recognition systems using advanced machine learning techniques offers a promising solution to these challenges.

This study introduces a novel CNN-RNN model specifically designed for the recognition of Kannada stone inscriptions. The Convolutional Neural Network (CNN) component excels in spatial feature extraction [3], identifying intricate patterns and textures within the inscriptions. Meanwhile, Recurrent Neural Networks (RNNs) are adept at processing sequential information, capturing the flow and continuity of the script despite the irregularities caused by erosion or stylistic variations. The novelty of this research lies in the integration of CNN and RNN architectures to create a hybrid model that is uniquely suited for the complexities of Kannada stone inscriptions. Although CNNs have been widely used for image recognition and RNNs for sequence processing, their combined application in the context of ancient script recognition is relatively unexplored. This approach allows the model to overcome the challenges posed by the degraded and stylistically varied characters that are characteristic of these inscriptions.

In [1], the focus was on the recognition of stone inscriptions that are difficult to identify. In [2], a database of 100,000 words from 600 users was created to train a model to recognize letters and symbols. In [3], a three-level character segmentation technique was proposed, which divided Kannada characters into three regions: the top region, the middle character region, and a post-connection border region. In [4], an area-metric-based feature extraction method was used to divide the image into panels and calculate centroid marks. Finally, in [5], cognitive functions such as binarization, magnification, and pattern analysis were discussed.

II. MODEL DESIGN

The model consists of steps such as segmentation, preprocessing, feature extraction, and model parameter evaluation, as shown in Figure 1.

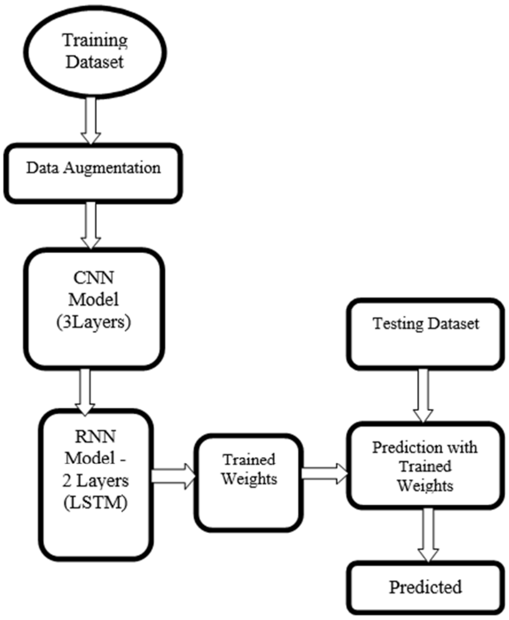

Fig. 1. Proposed model.

A. Dataset

The dataset was created using a 50 megapixel digital camera by visiting historical places in Karnataka, such as the Hampli Bellary and Belur Hassan districts of Karnataka, and the Somnath temple in Mysore district, Karnataka. Approximately ten thousand images were captured and preprocessed using image preprocessing algorithms.

B. Preprocessing

Image preprocessing is an important step to highlight essential features. As a result, the amount of memory required to store image data and the time required to perform further image analysis were drastically reduced.

1) Contrast Normalization

The newMin and newMax are reference values that define the width and the center of the input intensity range. Automatic normalization in image processing software usually normalizes the entire dynamic range of the digital system specified in the image data format [6]. Normalization benefits include faster search, analysis, and test because the table is narrower and more rows can fit in the datasheet, as shown in :

$$I_N = (I - Min)\frac{newMax-newMIN}{1+e^{-\frac{I-\beta}{\alpha}}} + newMin \quad (1)$$

2) Noise Removal

Different types of noise occur in stone inscription images. Noise is caused by errors in the image acquisition process and results in pixel values that are not representative of the real situations [7]. Noise can enter the image in a variety of ways. The algorithm used for denoising is nonlocal, as shown in:

$$x(i) = \sum_{j\epsilon\Omega}^{\infty} \mathrm{w}\, w_{ij}\, y(j) =$$

$$\frac{1}{C(i)}\sum_{j\epsilon\Omega}^{\infty} \mathrm{w}\, exp\left\{-\frac{\|Y_i - Y_j\|_{2,\alpha}^2}{h^2}\right\} Y(j) \quad (2)$$

where h is the filter parameter in the image block centered on i, j, $C(i)$ is the normalized factor, and $\frac{\|Y_i - Y_j\|_{2,\alpha}^2}{h^2}$ is the measure of dissimilarity between similar blocks of the image.

3) Binarization

Binarization is the process of converting a grayscale image. The binarization process begins with determining the grayscale threshold value and then determining whether or not a pixel has a specific gray value. If the gray value of the pixels exceeds the threshold, these pixels are converted to white. Similarly, if a pixel's gray value is less than the threshold, that pixel is converted to black as shown in (3).

$$K = \frac{\sigma^2 B}{\sigma^2 G} \quad (3)$$

The Otsu method automatically calculates the threshold value based on global and between-class variance. K is the threshold value, B is the global variance of the entire image, and G is the variance between classes.

4) Edge Detection and Cropping

Edge detection is a part of preprocessing used to identify the boundaries (edges) of objects or areas in images [8, 9]. First, after binarization, its output is considered as input for edge detection. Second, all pixels are summed as shown in (4). Third, each pixel is compared to the calculated mean to crop out the regions that fall below it (5).

$$avg = \frac{\sum_i^{rows}\sum_j^{cols} aij}{no.of\ rows} \quad (4)$$

$$textarea = \sum_j^{cols} aij - avg \quad (5)$$

5) Segmentation

The segmentation of Kannada stone inscription images is an important step in preprocessing historical documents and involves separating a single word or line from a complex and often distorted background [10, 11]. In line segmentation checks the text and counts the lines that are not visible. The average is taken and the numbers are normalized to this average and the non-zero value regions are marked as text. The green line in the figure represents the beginning of the text, and the red line represents the end of the text, as seen in Figure 2.

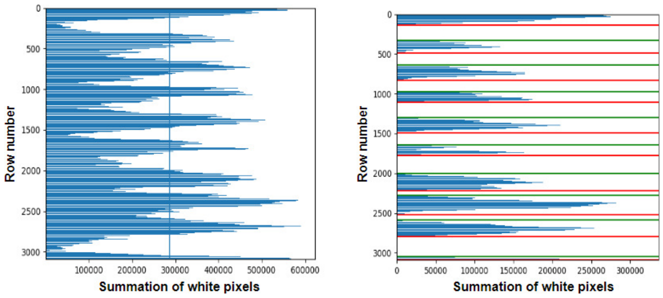

Fig. 2. Graphical analysis.

6) *Data Augmentation*

By applying several changes to current data, data augmentation is used to improve the quantity and diversity of training data [12, 13]. The generated dataset is used to train models after the data has been enhanced. This is especially helpful in situations where few training data are available. Data augmentation can help prevent overfitting, which occurs when the model is too closely suited to the training data and performs poorly on new data. Overfitting can be avoided by developing new variations of existing data. When there is a lack of training data, this can aid in enhancing the efficiency and accuracy of machine learning models [14, 15].

- A rotation range of up to 25 provides a list of possible random rotation angles for the image. The image can be randomly rotated up to 25 degrees in a clockwise or anticlockwise direction.
- The width shift range is 0.1 and the height shift range is 0.1. The range of random shifts that can be applied to the image is determined by these parameters. A value of 0.1 means that the image can be shifted by a maximum of 10% of its width and height, respectively.
- A zoom range of 0.2 specifies how much random zoom can be used on the image. A value of 0.2 means that the image can be zoomed in or out by a maximum of 20%.
- Horizontal flip specifies whether horizontal flipping should be performed on the image. When set to True, the image will be horizontally flipped.
- Fill mode = Nearest determines the strategy used to fill in newly created pixels after applying transformations.

C. *CNN–RNN Hybrid Model*

1) *Data Preparation*

The dataset consists of images of Kannada stone inscriptions, preprocessed to a size of 52×52 pixels and normalized to the range [0, 1]. The dataset is divided into training, validation, and test sets[16-18]. To train and evaluate the CNN-RNN model on Kannada stone inscriptions, the dataset was divided as follows:

- Training set: 70% of the total dataset.
- Validation set: 15% of the total dataset.
- Testing set: 15% of the total dataset.

2) *CNN Architecture*

The CNN comprises three convolutional layers:

- 1st Convolutional Layer: 32 filters, 3×3 kernel, ReLU activation, followed by max-pooling.
- 2nd Convolutional Layer: 64 filters, 3×3 kernel, ReLU activation, followed by max-pooling.
- 3rd Convolutional Layer: 128 filters, 3×3 kernel, ReLU activation, followed by max-pooling [17-20].

The output of the CNN is flattened and reshaped to fit the input requirements of the RNN [19, 20].

3) *RNN Architecture*

The RNN includes two layers of Long Short-Term Memory (LSTM) units:

- 1st LSTM Layer: 128 units, returning sequences.
- 2nd LSTM Layer: 128 units.

4) *Fully Connected Layer*

A dense layer with a softmax activation function is added to the output of the RNN to produce the final classification.

III. RESULTS

The results obtained from the CNN-RNN model were derived through a systematic process, beginning with the capture of the stone inscription images. The captured image, as shown in Figure 3, underwent preprocessing to enhance quality before being segmented into individual characters. Grayscale conversion or gray scaling of the input image is shown in Figure 4. Edge detection seeks to locate the borders or edges of several objects or regions within an image. The Sobel filter was used to find the gradient of intensity at each pixel in an image, which identifies the direction with the highest transition from light to dark and measures the magnitude of that transition. Figure 5 shows the Sobel edge detection [21].

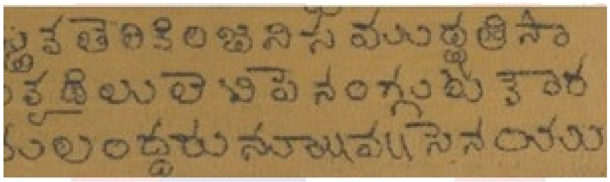

Fig. 3. Input image example.

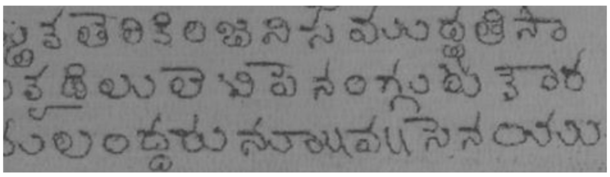

Fig. 4. Gray scaled image.

Line segmentation is used to divide an image or document into separate lines of text. Line segmentation is shown in Figure 6.

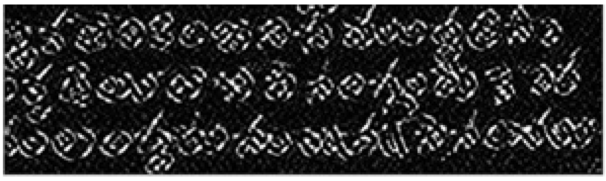

Fig. 5. Edge detected image.

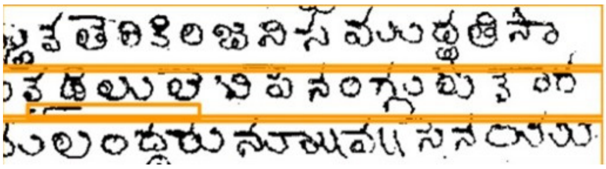

Fig. 6. Line segmentation.

A. Character Segmentation

Character segmentation is shown in Figure 7. The segmented characters are saved as separate images that can be used to feed the recognition model.

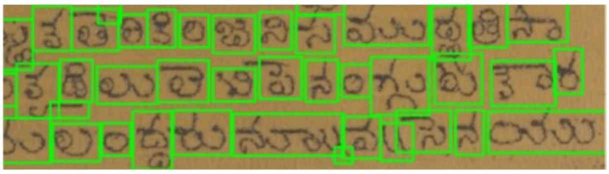

Fig. 7. Character segmentation.

B. Output

By feeding the segmented characters to the recognition model, the segmented output of modern Kannada characters is acquired. The output obtained is shown in Figure 8.

Fig. 8. Output.

C. Recognition of Characters in Stone Inscriptions

1) Input Image

Capturing high-quality images of stone inscriptions is a crucial first step. Figures 9 and 10 represent degraded stone inscription images.

Fig. 9. Degraded Kannada stone inscription image of Hoysala Kingdom.

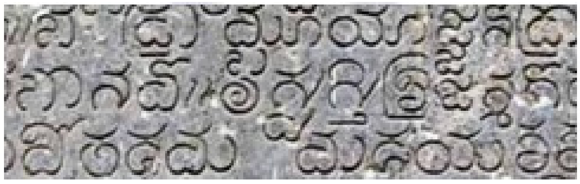

Fig. 10. Degraded Kannada stone inscription image of Vijayanagar Kingdom.

2) Output

Figures 11 and 12 show the resulting digitized outputs.

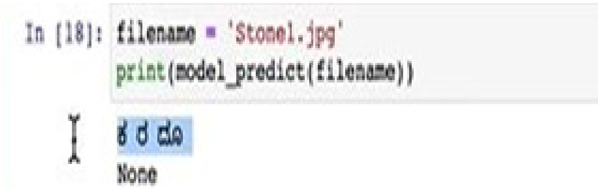

Fig. 11. Output of degraded Kannada stone inscription of Holysala Kingdom.

Fig. 12. Output of degraded Kannada stone inscription of Vijayanagar Kingdom.

D. Model Summary

The CNN-RNN model processes an input image to extract features through CNN layers. Then, the RNN layers interpret sequential information, and finally output class probabilities for each time step. This model takes a 52×52 grayscale image as input, processes it through a series of CNN layers to extract spatial features, reshapes the output for sequential processing by RNN layers, and finally produces a sequence of class probabilities for character recognition. This hybrid model takes advantage of both CNNs (for spatial feature extraction) and RNNs (for sequence modeling) to effectively recognize Kannada stone inscriptions. Table I shows a comparison of the proposed CNN-RNN hybrid model with a CNN model.

TABLE I. MODEL COMPARISON

Parameter	CNN Model (1)	Proposed CNN-RNN hybrid model
Number of layers	8 (4 Conv, 3 Pool, 1 Dense)	10 (3 Conv, 3 Pool, 2 LSTM, 1 Reshape, 1 TimeDist)
Filter sizes	(3×3), (3×3), (3×3), (3×3)	(3×3), (3×3), (3×3)
Number of filters	32, 64, 128, 256	32, 64, 128
Total parameters	~1M	~2M
Accuracy	~85%	~92%
Loss	~0.5	~0.3
Precision	~84%	~91%
Recall	~83%	~92%
F1 score	~83.5%	~91.5%
Inference time	~50ms per image	~100ms per image
Training time	~2 hours for 10 epochs	~4 hours for 10 epochs
Epochs to convergence	10	20

Table II shows a comparison between the proposed and previous approaches.

TABLE II. COMPARATIVE ANALYSIS OF KANNADA CHARACTER RECOGNITION

No.	Techniques	Total datasets	Accuracy (%)
1	Neural Network [22]	2450 broken character dataset synthetically generated	98.9
2	FDA (Fit Discriminant Analysis) [23]	250 real datasets from historical, 150×49=7350 Synthetically generated dataset	99.38
3	FLD (Fisher Linear Discriminant Analysis) [24]	21560 Kannada and English Characters	98.2
4	End point algorithm [25]	Degraded 100 Kannada characters	89
5	CNN-RNN (Proposed method).	Degraded 16100 Kannada characters	95

The experimental results in Table II show that the proposed method has a higher level of recognition accuracy than the previous methods that are already in use.

IV. CONCLUSION

The proposed CNN-RNN model introduces innovative elements specifically tailored for recognizing Kannada stone inscriptions, marking a significant leap in historical document analysis. By combining CNNs for spatial feature extraction with RNNs for sequential processing, the model effectively addresses the challenges of eroded and stylistically varied characters. Comprehensive data augmentation further enhanced its robustness, allowing it to generalize across different conditions of degradation. This tailored approach achieved 95% accuracy, showing improved performance compared to traditional methods, advancing both technological capabilities in document recognition and the preservation of cultural heritage. By automating the recognition of ancient Kannada scripts, this study makes these cultural artifacts more accessible and aids in their preservation. Future work should focus on refining the model, exploring advanced variants such as transformers, extending applicability to other ancient scripts, and developing more sophisticated data augmentation techniques to further improve accuracy and robustness.

REFERENCES

[1] B. K. Rajithkumar, B. V. Uma, and H. S. Mohana, "Recognition of Historical Kannada Manuscripts using Convolution Neural Network," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 3, pp. 1138–1147, Mar. 2024.

[2] B. K. Rajithkumar and H. S. Mohana, "Template Matching Method for Recognition of Stone Inscripted Kannada Characters of Different Time Frames Based on Correlation Analysis," *International Journal of Electrical and Computer Engineering*, vol. 4, no. 5, pp. 719–729, Oct. 2014, https://doi.org/10.11591/ijece.v4i5.6333.

[3] T. N. Sridevi, *et al.*, "Deep Convolution Neural Network for Degraded Printed Kannada Character Recognitions," *Indian Journal of Computer Science and Engineering*, vol. 4, no. 3, May 2021.

[4] H. S. Mohana, B. K. Rajithkumar, B. R. Sujatha, and K. Navya, "Extraction of stone in-scripted Kannada characters using sift algorithm based image mosaic," *International Journal of Electronics & Communication Technology*, vol. 5, no. 2, 2014.

[5] H. S. Mohana and B. K. Rajithkumar, "Era identification and recognition of Ganga and Hoysala phase Kannada stone inscriptions characters using advance recognition algorithm," in *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, Kanyakumari District, India, Jul. 2014, pp. 659–665, https://doi.org/10.1109/ICCICCT.2014.6993043.

[6] M. Bhalekar and M. Bedekar, "D-CNN: A New model for Generating Image Captions with Text Extraction Using Deep Learning for Visually Challenged Individuals," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8366–8373, Apr. 2022, https://doi.org/10.48084/etasr.4772.

[7] S. R. Basha, J. K. Rani, and J. J. C. P. Yadav, "A Novel Summarization-based Approach for Feature Reduction Enhancing Text Classification Accuracy," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 5001–5005, Dec. 2019, https://doi.org/10.48084/etasr.3173.

[8] R. Prabavathi and R. Karthikeyan, "Eighth Century Tamil Character Recognition," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, no. 2, pp. 707–712, 2015.

[9] H. T. Chandrakala and G. Thippeswamy, "Deep Convolutional Neural Networks for Recognition of Historical Handwritten Kannada Characters," in *Frontiers in Intelligent Computing: Theory and Applications*, 2020, pp. 69–77, https://doi.org/10.1007/978-981-13-9920-6_7.

[10] F. Lombardi and S. Marinai, "Deep Learning for Historical Document Analysis and Recognition—A Survey," *Journal of Imaging*, vol. 6, no. 10, Oct. 2020, Art. no. 110, https://doi.org/10.3390/jimaging6100110.

[11] A. S. Rao, S. Nayak, A. K. Arpitha, C. Nayak, and M .S. Nayak, "Exploring Deep Learning Techniques for Kannada Handwritten Character Recognition: A Boon for Digitization," *International Journal of Advanced Science and Technology*, vol. 29, no. 5, pp. 11078–11093, 2020.

[12] H. T. Chandrakala and G. Thippeswamy, "A comprehensive survey on OCR techniques for kannada script," *International Journal of Scientific Research*, vol. 5, no. 4, pp. 35–39, 2016.

[13] M. R. Murty, J. V. R. Murthy, and P. V. G. D. Prasad Reddy, "Text document classification based-on least square support vector machines with singular value decomposition," *International Journal of Computer Applications*, vol. 27, no. 7, pp. 21–26, 2011.

[14] S. Roy, N. Das, M. Kundu, and M. Nasipuri, "Handwritten isolated Bangla compound character recognition: A new benchmark using a novel deep learning approach," *Pattern Recognition Letters*, vol. 90, pp. 15–21, Apr. 2017, https://doi.org/10.1016/j.patrec.2017.03.004.

[15] J. Jomy, K. Balakrishnan, and K. V. Pramod, "A System for Offline Recognition of Handwritten Characters in Malayalam Script," *International Journal of Image, Graphics and Signal Processing*, vol. 5, no. 4, pp. 53–59, Apr. 2013, https://doi.org/10.5815/ijigsp.2013.04.07.

[16] A. B. Shinde and Y. H. Dandawate, "Shirorekha extraction in Character Segmentation for printed devanagri text in Document Image Processing," in *2014 Annual IEEE India Conference (INDICON)*, Pune, India, Dec. 2014, pp. 1–7, https://doi.org/10.1109/INDICON.2014.7030535.

[17] S. Sharma, A. Sasi, and A. N. Cheeran, "A SVM based character recognition system," in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, India, May 2017, pp. 1703–1707, https://doi.org/10.1109/RTEICT.2017.8256890.

[18] D. K. Gowda and V. Kanchana, "Kannada Handwritten Character Recognition and Classification Through OCR Using Hybrid Machine Learning Techniques," in *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, Jul. 2022, pp. 1–6, https://doi.org/10.1109/ICDSIS55133.2022.9915906.

[19] S. Puri and S. P. Singh, "An efficient Devanagari character classification in printed and handwritten documents using SVM," *Procedia Computer Science*, vol. 152, pp. 111–121, Jan. 2019, https://doi.org/10.1016/j.procs.2019.05.033.

[20] N. Sarika, N. Sirisala, and M. S. Velpuru, "CNN based Optical Character Recognition and Applications," in *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, Jan. 2021, pp. 666–672, https://doi.org/10.1109/ICICT50816.2021.9358735.

[21] N. Sandhya and R. Krishnan, "Broken kannada character recognition — A neural network based approach," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, India, Mar. 2016, pp. 2047–2050, https://doi.org/10.1109/ICEEOT.2016.7755047.

[22] N. Sandhya, R. Krishnan, D. R. R. Babu, and N. B. Rao, "An efficient approach for handling degradation in character recognition," *International Journal of Advanced Intelligence Paradigms*, vol. 14, no. 1–2, pp. 14–29, Jan. 2019, https://doi.org/10.1504/IJAIP.2019.102960.

[23] V. N. M. Aradhya, G. H. Kumar, S. Noushath, and P. Shivakumara, "Fisher Linear Discriminant Analysis based Technique Useful for Efficient Character Recognition," in *2006 Fourth International Conference on Intelligent Sensing and Information Processing*, Bangalore, India, Dec. 2006, pp. 49–52, https://doi.org/10.1109/ICISIP.2006.4286060.

[24] N. Sandhya, R. Krishnan, and D. R. Ramesh Babu, "A novel local enhancement technique for rebuilding Broken characters in a degraded Kannada script," in *2015 IEEE International Advance Computing Conference (IACC)*, Banglore, Jun. 2015, pp. 176–179, https://doi.org/10.1109/IADCC.2015.7154693.