Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

A Study on Publicly Available Datasets and Preprocessing Techniques for Handwritten Devanagari Character Recognition

Khan Shahajahan[1], Dr. Garima Tyagi[2]

[1]Vice Principal, Dr. B.M.N. College of Home Science, Matunga, Mumbai, India.
[2]Research Supervisor, School of Computer Application & Technology, Career Point University, Kota, Rajasthan, India
khan@bmncollege.com, garima.tyagi@cpur.edu.in

Abstract:

Handwritten character recognition is a fundamental research area in pattern recognition and document image analysis. It has wide-ranging applications such as digitization of handwritten documents, automation of data entry systems, archival of historical manuscripts, and development of assistive technologies. In the Indian context, the Devanagari script holds special importance as it is used by several major languages including Hindi, Marathi, Sanskrit, Nepali, and Konkani. However, the recognition of handwritten Devanagari characters remains a challenging task due to the complexity of character structures, the presence of modifiers, the shirorekha (headline), and the large variability in handwriting styles.

While many recent studies focus on deep learning-based recognition systems, the importance of dataset quality and preprocessing remains fundamental for building any reliable optical character recognition (OCR) system. Poor preprocessing leads to noisy feature representations, reduced classification accuracy, and poor generalization across writers. In many practical scenarios, preprocessing contributes more to system reliability than the choice of classifier itself.

This paper focuses on the first and most essential objective of handwritten Devanagari OCR: the use of publicly available benchmark datasets and their systematic preprocessing to obtain clean, standardized handwritten character samples for training and evaluation. Two widely used datasets, namely the Devanagari Handwritten Character Dataset (DHCD) and the ECO-LAPS dataset, are selected for this study. A complete preprocessing pipeline is designed consisting of normalization, grayscale conversion, binarization, noise removal, morphological processing, shirorekha removal, contour refinement, and size standardization.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
 | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

The study highlights the importance of dataset preparation in reducing intra-class variations and improving recognition reliability. The resulting standardized dataset forms a strong foundation for classical machine learning based OCR systems and ensures reproducibility and fair evaluation.

Keywords: Handwritten Character Recognition, Devanagari Script, DHCD Dataset, ECO-LAPS Dataset, Image Preprocessing, Optical Character Recognition, Pattern Recognition.

I. INTRODUCTION

Optical Character Recognition (OCR) is an important field of research in computer vision and pattern recognition. OCR systems aim to convert printed or handwritten text into a machine-readable form. Such systems are widely used in document digitization, postal automation, bank cheque processing, form processing, and archival of historical documents. With the rapid growth of digital governance, e-learning platforms, and paperless workflows, OCR has become an essential technology for information access and automation.

Handwritten OCR is significantly more difficult than printed text recognition due to variations in writing styles, stroke thickness, slant, spacing, and individual writing habits. Each person writes characters differently, which introduces large intra-class variability even for the same character. This makes handwritten character recognition a challenging pattern recognition problem.

In India, the Devanagari script is one of the most widely used writing systems. It is used by languages such as Hindi, Marathi, Sanskrit, Nepali, and Konkani. Hindi alone is spoken by more than 500 million people worldwide, making Devanagari OCR an important research problem [1]. Large volumes of handwritten records in Devanagari exist in schools, universities, government offices, courts, and historical archives. Digitizing such records would enable easy storage, search, and retrieval.

Despite its importance, handwritten Devanagari character recognition remains a difficult task. Devanagari characters consist of curves, loops, vertical strokes, modifiers, and a horizontal headline known as the shirorekha. The presence of the shirorekha often causes characters in a

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

word to connect, which creates segmentation and recognition challenges. Moreover, handwritten documents often contain noise, smudges, skew, faded ink, and broken strokes.

Most OCR systems follow a general pipeline consisting of:

1. Image acquisition
2. Preprocessing
3. Segmentation
4. Feature extraction
5. Classification

Among these stages, preprocessing is one of the most critical steps. The quality of the input image directly influences the quality of extracted features and classification accuracy. Poor preprocessing can degrade system performance even when powerful classifiers are used. In recent years, many researchers have focused on deep learning-based OCR systems. While such models achieve high accuracy, they require large datasets, long training times, and high computational resources. Classical machine learning based systems, on the other hand, rely heavily on proper dataset preparation and preprocessing for achieving good performance with limited resources [2]. In resource-constrained environments, such as mobile or embedded systems, classical OCR approaches remain highly relevant.

Therefore, this research focuses on the first and most fundamental objective of any OCR system: the use of publicly available datasets and their systematic preprocessing to obtain clean, standardized handwritten Devanagari character samples for training and evaluation.

II. REVIEW OF LITERATURE

Handwritten character recognition has been an active research area in pattern recognition and document image analysis for several decades. With the increasing demand for digitization of handwritten records, several studies have focused on developing OCR systems for Indian scripts, especially Devanagari.

Acharya et al. [3] introduced the Devanagari Handwritten Character Dataset (DHCD), which is one of the largest publicly available datasets for handwritten Devanagari characters. This dataset has been widely used as a benchmark in OCR research and provides a standard platform for evaluating recognition systems.

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

Otsu [4] proposed an automatic threshold selection method for binarization based on gray-level histograms. This method is widely used in OCR preprocessing due to its simplicity and effectiveness in separating foreground text from background.

Deore and Pravin [5] applied Histogram of Oriented Gradients (HOG) features with Support Vector Machine (SVM) classifiers for handwritten Devanagari character recognition. Their study showed that preprocessing plays a critical role in improving classification accuracy.

Gupta and Goyal [6] conducted a comparative study of KNN and SVM classifiers for handwritten Devanagari OCR and highlighted the importance of noise-free and well-prepared datasets.

Patil et al. [7] used zoning and Hu moments for feature extraction and demonstrated that feature stability depends heavily on preprocessing quality. Kumar et al. [8] studied structural feature-based OCR systems for Indic scripts and emphasized the need for robust preprocessing to handle handwriting variability.

Most existing studies focus primarily on feature extraction and classification models, while dataset preparation and preprocessing are often treated as a secondary step. However, poor preprocessing leads to unstable feature representations and reduced recognition performance.

III. RESEARCH GAPS IDENTIFIED

From the literature survey, the following research gaps are identified:

- Most studies focus on improving classifier accuracy, while limited attention is given to dataset preparation and preprocessing quality.
- There is a lack of standardized preprocessing pipelines for handwritten Devanagari datasets.
- Many research works use private or custom datasets, which limits reproducibility and fair comparison.
- Inconsistent preprocessing methods are applied across different datasets, leading to unstable evaluation results.
- Limited research focuses on building clean, standardized benchmark datasets for classical machine learning based OCR systems.

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

These gaps highlight the need for a systematic preprocessing framework using publicly available benchmark datasets.

IV. OBJECTIVES OF RESEARCH

The main objectives of this research are:

1. To study and analyze publicly available handwritten Devanagari character datasets.
2. To identify the challenges present in raw handwritten character images.
3. To design a systematic preprocessing pipeline for handwritten Devanagari OCR.
4. To create a clean and standardized dataset for reliable feature extraction.
5. To improve dataset quality and reproducibility for OCR research.

V. METHODOLOGY

This study follows a systematic and reproducible methodology focused on preparing standardized handwritten Devanagari character datasets using a well-defined preprocessing pipeline. The objective is to transform raw handwritten samples into clean, noise-free, and normalized character images suitable for classical machine learning based OCR systems.
The methodology is designed around three main components: dataset selection, preprocessing pipeline design, and experimental implementation.

V. i Dataset Selection

Two publicly available benchmark datasets are selected for this study:

1. Devanagari Handwritten Character Dataset (DHCD)
2. ECO-LAPS Dataset

These datasets are chosen because they are:

- Widely used in published research
- Properly labeled
- Publicly accessible
- Suitable for reproducible evaluation

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

DHCD provides a large-scale benchmark with 92,000 samples of consonants and numerals, while ECO-LAPS complements it by including vowel characters. Together, they form a comprehensive dataset covering major components of the Devanagari script.

V. ii Data Acquisition

Both datasets are obtained from their official public repositories. All images are stored in grayscale format and organized class-wise.

Each image represents a single handwritten character written by different individuals using varying writing styles, stroke thickness, and pen pressure. This introduces natural handwriting variability, which is essential for building robust OCR systems.

V. iii Preprocessing Pipeline Design

A complete preprocessing pipeline is designed to convert raw handwritten samples into standardized character images. The pipeline consists of the following stages:

1. Image normalization
2. Grayscale conversion
3. Binarization
4. Noise removal
5. Morphological processing
6. Shirorekha removal
7. Contour refinement
8. Size standardization

Each stage is applied sequentially to every image in both datasets. The same preprocessing steps are applied uniformly across all samples to ensure consistency and fairness.

The preprocessing pipeline aims to:

- Remove background noise and artifacts
- Improve stroke connectivity
- Normalize scale and position

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

- Reduce intra-class variations
- Enhance structural clarity of characters

V. iv Preprocessing Implementation

The preprocessing operations are implemented using standard Python-based image processing libraries:

- OpenCV
- NumPy
- Scikit-image

Each image passes through the complete pipeline automatically. Intermediate outputs are visually verified to ensure that character structure is preserved while noise and distortions are removed.

All preprocessing parameters are kept constant for both datasets to avoid dataset-specific bias.

V. v Standardization and Dataset Preparation

After preprocessing, all character images are:

- Centered within the frame
- Aligned to a common reference
- Resized to 32 × 32 pixels
- Converted into binary format

This produces a clean and standardized dataset that can be directly used for feature extraction and classification in classical OCR systems.

V. vi Evaluation Strategy

The effectiveness of preprocessing is evaluated through:

- Visual inspection of processed samples
- Noise reduction assessment
- Stroke continuity verification

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

- Shape preservation analysis

The objective is not classifier benchmarking, but dataset preparation quality. The final output is a well-structured and reproducible dataset suitable for future OCR research.

Methodology Summary

The proposed methodology ensures:

- Reproducibility through public datasets
- Uniform preprocessing across samples
- Noise-free standardized character images
- Reduced handwriting variability
- Reliable dataset foundation for OCR systems

This methodology establishes a strong base for feature extraction and classification stages in subsequent OCR research.

VI. DEVANAGARI SCRIPT CHARACTERISTICS

The Devanagari script has several unique properties that differentiate it from Latin and other scripts.

VI.i Character Set

The Devanagari script consists of:

- 13 vowels
- 36 consonants
- 10 numerals
- Several compound characters formed by combining consonants

Each character is usually written with a horizontal line at the top known as the shirorekha. When writing words, these shirorekha lines often join, forming a continuous horizontal line across the word.

VI.ii Structural Complexity

Devanagari characters are composed of:

- Curves and loops

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

- Straight vertical and horizontal strokes
- Diacritical marks
- Modifiers attached above, below, left, or right of base characters

Many characters look visually similar and differ only by a small stroke or dot. For example, characters such as 'म' and 'भ', or 'न' and 'व' are often confused in handwriting. This structural similarity makes recognition difficult, especially when the handwriting is cursive or hurried.

VI.iii Handwriting Variability

Handwritten Devanagari text exhibits:

- Different writing speeds
- Different stroke thickness
- Different pen pressures
- Different writing instruments
- Individual writing habits

These variations introduce large intra-class diversity and make recognition challenging. Even the same person may write the same character differently at different times. This variability is one of the main reasons why handwritten OCR remains a difficult problem.

VII. IMPORTANCE OF DATASET SELECTION

The dataset used for training and evaluation is the backbone of any OCR system. A good dataset should be:

- Publicly available
- Large and diverse
- Properly labeled
- Balanced across classes
- Reproducible

Private datasets limit reproducibility and comparison across studies. Therefore, publicly available benchmark datasets are preferred. Using standardized datasets ensures fair evaluation and benchmarking across research works.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

Many research works suffer from the limitation of using custom datasets, which makes it difficult for other researchers to validate results. Public datasets help establish a common evaluation framework. When researchers use the same dataset, improvements can be measured objectively.

VIII. PUBLICLY AVAILABLE DATASETS FOR DEVANAGARI OCR

Two major publicly available datasets are used in this study:

1. Devanagari Handwritten Character Dataset (DHCD)
2. ECO-LAPS Dataset

These datasets are widely accepted in the research community and provide a strong benchmark for handwritten Devanagari OCR.

IX. DEVANAGARI HANDWRITTEN CHARACTER DATASET (DHCD)

The DHCD dataset was introduced by Acharya et al. [3] and is one of the largest publicly available datasets for handwritten Devanagari characters.

IX.i Dataset Description

The DHCD dataset contains:

- 92,000 grayscale images
- Image size: 32 × 32 pixels
- 46 character classes
    - 36 consonants
    - 10 numerals

Each class contains approximately 2,000 samples. The dataset is divided into a training set (85%) and a testing set (15%), which makes it suitable for benchmarking.

IX.ii Dataset Structure

TABLE I : DHCD Dataset Overview

Parameter	Description
Total Images	92,000

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
 | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

Image Size	32 × 32
Classes	46
Samples per Class	~2000
Format	Grayscale
Split	Train/Test
Availability	Public

IX.iii Dataset Challenges

Despite its quality, DHCD still contains:

- Noise from scanning
- Broken strokes
- Inconsistent handwriting
- Partial shirorekha
- Background artifacts

Therefore, preprocessing is required before feature extraction.

X. ECO-LAPS DATASET

The ECO-LAPS dataset complements DHCD by including vowel samples.

X.i Dataset Description

The ECO-LAPS dataset contains:

- Approximately 5,800 images
- 58 character classes
    - 36 consonants
    - 12 vowels
    - 10 numerals

Each class contains around 100 samples.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
 | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

X.ii Dataset Structure

TABLE II : ECO-LAPS Dataset Overview

Parameter	Description
Total Images	~5,800
Classes	58
Samples per Class	~100
Includes Vowels	Yes
Format	Grayscale
Availability	Public

X.iii Dataset Challenges

- Smaller dataset size
- More variation in writing styles
- Less standardization compared to DHCD

These factors make preprocessing even more important.

XI. NEED FOR PREPROCESSING

Raw handwritten images contain distortions such as:

- Background noise
- Uneven lighting
- Ink smudges
- Skew and rotation
- Variable image sizes
- Broken strokes

Without preprocessing, these distortions lead to unstable feature extraction and poor classification accuracy.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

Preprocessing aims to transform raw images into clean, standardized representations that highlight essential character structure. This step is crucial for classical machine learning systems where feature quality directly determines classifier performance.

XII. PREPROCESSING PIPELINE

The preprocessing pipeline designed in this research consists of the following stages:

XII.i Image normalization

XII.ii Grayscale conversion

XII.iii Binarization

XII.iv Noise removal

XII.v Morphological processing

XII.vi Shirorekha removal

XII.vii Contour refinement

XII.viii Size standardization

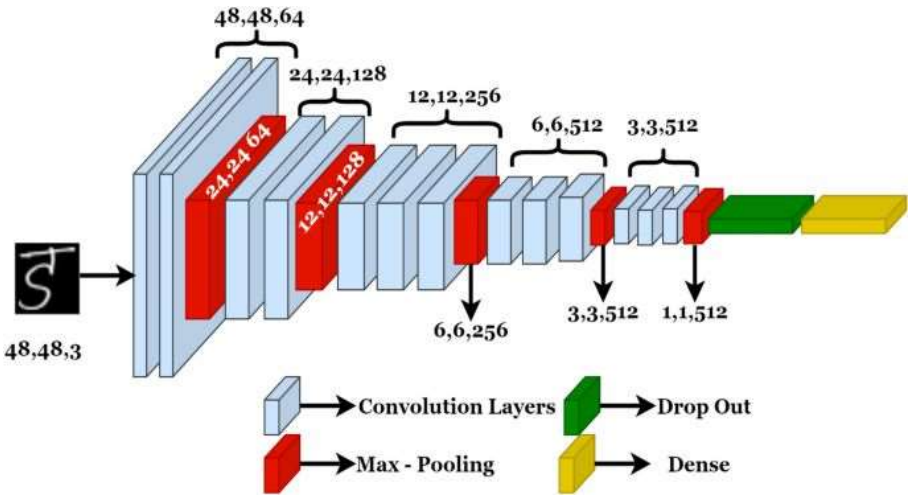


Fig. 1 Commonly used image processing pipeline

Each stage is discussed below.

XII.i Image Normalization

Images from different sources may have different sizes, aspect ratios, and resolutions. Normalization ensures consistency across the dataset.

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
 | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

All images are resized to a fixed dimension of 32 × 32 pixels using bilinear interpolation. This ensures uniformity in feature vector dimensions. Normalization reduces scale variation and improves feature stability.

XII.ii Grayscale Conversion

RGB images are converted to grayscale using the luminance formula:

Gray = 0.299R + 0.587G + 0.114B

This reduces computational complexity while preserving stroke information.

XII.iii Binarization

Binarization converts grayscale images into black and white images. This step separates foreground (ink) from background.

Otsu's thresholding method is used for binarization [4]. It automatically selects an optimal threshold.

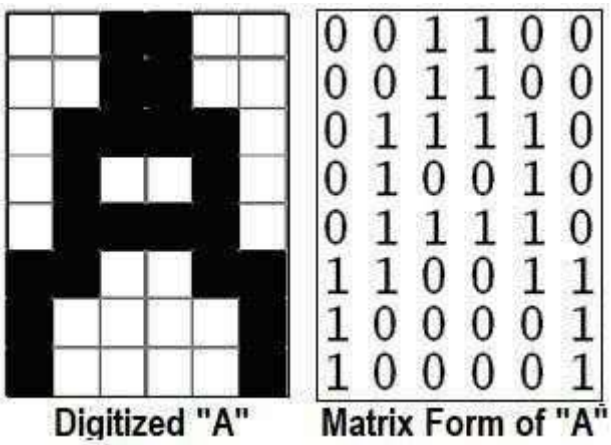


Fig. 2 Sample Grayscale and Binarized Image

XII.iv Noise Removal

Noise can arise due to scanning artifacts, paper texture, ink bleeding, and dust particles.

Noise removal is performed using:

- Median filtering
- Gaussian smoothing

Proceedings of International Conference on "Recent Innovations in Engineering, Technology & Science for Sustainable Development (RIETSSD-2026)"
Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

- Morphological opening

These filters remove small unwanted pixels while preserving character structure.

XII.v Morphological Operations

Morphological operations improve the connectivity and shape of characters.

Common operations include:

- Dilation
- Erosion
- Opening
- Closing

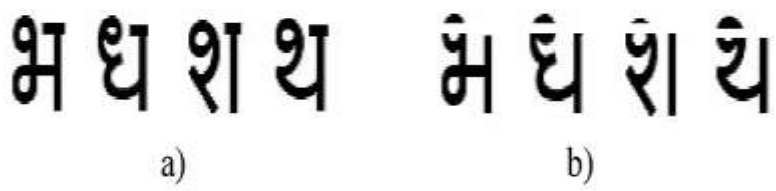
These operations help restore broken strokes and remove small noise.

XII.vi Shirorekha Removal

The shirorekha is detected using horizontal projection profiles.

The general approach is:

- Detect horizontal lines
- Identify the longest horizontal stroke
- Remove or suppress the detected line



a) b)

a) Devanagari Words with Shirorekha b) Devanagari Words After Removing Shirorekha using horizontal Projection Profile Method

Fig.3 Shirorekha Detection and Removal

XII.vii Contour Refinement

Contour refinement smooths jagged edges and removes irregular boundaries.

Techniques used include:

- Contour smoothing

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

- Skeletonization
- Edge refinement

Skeletonization reduces characters to a one-pixel-wide representation.

XII.viii Standardization

After preprocessing, all characters are:

- Centered
- Normalized
- Aligned

This ensures scale invariance and position consistency.

XIII. BENEFITS OF PREPROCESSING

Effective preprocessing results in:

- Reduced intra-class variation
- Improved feature stability
- Higher classification accuracy
- Better generalization

Several studies show that good preprocessing improves accuracy by 5–10% [5].

XIV. EXPERIMENTAL SETUP

Python tools used:

- OpenCV
- NumPy
- Scikit-image

Preprocessing applied uniformly on DHCD and ECO-LAPS.

XV. RESULTS AND DISCUSSION

After applying the proposed preprocessing pipeline on both DHCD and ECO-LAPS datasets, the following results were observed.

Background noise and scanning artifacts were effectively removed using median filtering and morphological operations. Broken and disconnected strokes were restored using dilation and

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

closing operations, resulting in improved stroke continuity. Binarization using Otsu's method successfully separated foreground characters from the background.

Shirorekha removal reduced unwanted character connections and improved individual character isolation. Contour refinement and skeletonization enhanced shape clarity and preserved the essential structural features of characters.

Standardization ensured that all character images were centered, aligned, and resized to 32 × 32 pixels. This reduced scale and position variations across samples.

The preprocessing pipeline significantly reduced intra-class variations and improved feature stability. Clean and standardized character samples provide a strong foundation for feature extraction and classification in OCR systems.

The results confirm that preprocessing is a critical stage in handwritten OCR and directly influences recognition reliability and robustness.

XVI. CONCLUSION

This paper presented a comprehensive study on the use of publicly available datasets and preprocessing techniques for handwritten Devanagari character recognition.

The DHCD and ECO-LAPS datasets were selected due to their wide acceptance and reproducibility. A complete preprocessing pipeline was designed including normalization, binarization, noise removal, morphological processing, shirorekha removal, and contour refinement.

The study shows that preprocessing is a critical stage that directly influences recognition accuracy and system reliability. Proper dataset preparation reduces noise, improves consistency, and provides a strong foundation for feature extraction and classification.

XVII. FUTURE SCOPE

The standardized dataset generated in this study can be extended for further OCR research. Future work may include:

- Feature extraction using techniques such as HOG, zoning, and structural descriptors.
- Classification using machine learning models such as SVM, KNN, Random Forest, and Neural Networks.
- Performance comparison with deep learning-based OCR systems.

Career Point International Journal of Research (CPIJR)
©2022 CPIJR | Special Issue | ISSN : 2583-1895
Conference Paper | DOI: https://doi.org/10.5281/zenodo.20470734

- Development of real-time handwritten Devanagari OCR applications for mobile and embedded platforms.
- Extension of the preprocessing pipeline to word-level and document-level recognition.

REFERENCES

1. Census of India. (2011). *Language statistics*. Government of India.
2. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
3. Acharya, S., Pant, A. K., & Gyawali, P. K. (2015). Large scale handwritten Devanagari character dataset. *Proceedings of the IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, 1–5. IEEE.
4. Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076
5. Deore, S. G., & Pravin, A. (2019). Handwritten Devanagari character recognition using HOG features and SVM classifier. *International Journal of Computer Applications*, 178(7), 23–28.
6. Gupta, R., & Goyal, A. (2021). Comparative study of KNN and SVM classifiers for handwritten Devanagari character recognition. *International Journal of Advanced Computer Science and Applications*, 12(3), 455–461.
7. Patil, R., Deshmukh, S., & Kulkarni, V. (2020). Handwritten Devanagari character recognition using zoning and Hu moments. *International Journal of Engineering and Advanced Technology*, 9(3), 2456–2461.
8. Kumar, M., Sharma, R., & Singh, A. (2019). Structural feature based handwritten character recognition for Indic scripts. *Procedia Computer Science*, 152, 380–387. https://doi.org/10.1016/j.procs.2019.05.037
9. Chakrabarti, A., & Ray, S. (2023). Energy efficient OCR systems for handwritten character recognition. *Journal of Intelligent Systems*, 32(1), 421–435.