Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271

e-ISSN : 2541-2019
p-ISSN : 2541-044X

An End-to-End Balinese Lontar OCR Framework Using Bayesian-Optimized Multiscale Retinex and MobileNetV3s

Putu Ayu Febyanti[1)*], I Putu Agus Eka Darma Udayana[2)], Aniek Suryanti Kusuma[3)]
[1,2,3)] Magister Program of Informatics Institut Bisnis dan Teknologi Indonesia Denpasar,Indonesia
[1)] ayufeby@instiki.ac.id, [2)]agus.ekadarma@gmail.com, [3)]anieksuryanti@instiki.ac.id

Submitted :May 24, 2026 | **Accepted** : June 14, 2026 | **Published** : July 5, 2026

Abstract: The preservation of Balinese lontar manuscripts has become increasingly important due to their cultural, historical, and religious significance, while physical degradation such as uneven illumination, faded ink, texture interference, and manuscript aging continues to reduce readability and complicate digital preservation efforts. This study proposes an end-to-end Optical Character Recognition (OCR) framework for degraded Balinese lontar manuscripts by integrating Bayesian-optimized image enhancement, adaptive preprocessing, morphology-based segmentation, domain-specific augmentation, and lightweight deep learning recognition using MobileNetV3. The proposed enhancement pipeline combines Multiscale Retinex, adaptive gamma correction, edge-preserving filtering, and hybrid binarization to improve character visibility under degraded manuscript conditions. Bayesian Optimization with Optuna and Tree-structured Parzen Estimator (TPE) was employed to automatically optimize enhancement parameters according to manuscript quality characteristics. Experimental results demonstrated substantial improvements in manuscript image quality, where Laplacian Variance increased from 306.7596 to 6685.7641, RMS Contrast improved from 28.976 to 83.9085, Michelson Contrast increased from 0.8238 to 1.0, and Ink Ratio Score improved from 0.6096 to 0.9847. The MobileNetV3-based OCR recognition model achieved a test accuracy of 80.52% and a best validation accuracy of 83.78% across 102 Balinese script classes. The proposed framework demonstrates that adaptive enhancement optimization combined with lightweight OCR recognition can provide robust and computationally efficient recognition performance for degraded historical manuscripts while supporting scalable digital preservation and mobile-oriented cultural heritage applications.

Keywords: Adaptive preprocessing, Balinese script recognition, Bayesian optimization, MobileNetV3, Multiscale Retinex, optical character recognition, palm-leaf manuscript OCR.

INTRODUCTION

The preservation of cultural heritage manuscripts has become increasingly important in the digital era, particularly for historical palm-leaf manuscripts that contain valuable religious, literary, medical, and historical knowledge. Among these cultural artifacts, Balinese lontar manuscripts represent a significant component of Southeast Asian documentary heritage and remain an important medium for preserving traditional knowledge and Balinese script literacy (Kesiman & Pradnyana, 2021) (Siahaan et al., 2022). However, palm-leaf manuscripts are physically vulnerable to deterioration caused by aging, humidity, fungi, handling damage, and environmental exposure, which gradually degrade the readability of manuscript content and complicate long-term preservation efforts (Faizullah et al., 2023). Consequently, digitization initiatives have been widely promoted to preserve manuscript collections through scanned images and digital archives (Siahaan et al., 2022). Nevertheless, digitization alone is insufficient because manually transliterating and indexing manuscript contents remains labor-intensive, time-consuming, and difficult to scale for large collections (Kesiman & Pradnyana, 2021). In this context, Optical Character Recognition (OCR) technology has emerged as an essential computational approach for converting manuscript images into machine-readable text, thereby supporting accessibility, searchability, transliteration, and digital preservation of historical documents (Pratama et al., 2023). In addition, mobile OCR

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

applications have also been proposed to support Balinese script learning and cultural continuity by enabling camera-based recognition systems accessible to wider communities (Indrawan et al., 2022).

Despite the growing interest in manuscript OCR, recognizing text from degraded historical manuscripts remains a highly challenging task. Historical document analysis studies report that OCR performance is significantly affected by physical degradation factors such as non-uniform background, faded ink, stains, bleed-through artifacts, uneven illumination, and intensity variations (Lombardi & Marinai, 2020). Similar challenges are observed in historical handwriting datasets where document aging introduces noise, merged characters, broken strokes, and overlapping components that disrupt recognition performance (Krithiga et al., 2025). Palm-leaf manuscripts introduce additional complexities because the manuscript material itself is susceptible to severe degradation and irregular texture patterns, requiring robust preprocessing, segmentation, and recognition strategies (Faizullah et al., 2023). Specifically, Balinese lontar manuscripts present unique script-related difficulties, including dense character layouts, overlapping symbols, high interclass similarity, and writing rules without spaces between characters, which substantially complicate segmentation and character recognition (Suciati et al., 2022) (Siahaan et al., 2022). Previous studies further reported that character similarity among Balinese glyphs can significantly reduce recognition precision due to ambiguous visual structures (Ruwanmini et al., 2023). Traditional recognition pipelines based on projection-profile segmentation and KNN classification achieved only 52% recognition accuracy on palm-leaf lontar manuscripts compared with 92% on paper-based handwriting, highlighting the severe impact of manuscript degradation and segmentation difficulty on OCR performance (Sutramiani et al., 2022). Similarly, contour-based object detection approaches may incorrectly merge closely spaced characters into single regions and exhibit unstable detection performance under noisy conditions (Manuaba & Indah, 2021).

To address these challenges, previous studies have explored multiple preprocessing and OCR approaches for Balinese manuscripts and related historical scripts. Existing pipelines commonly utilize thresholding, morphology, contour extraction, edge detection, texture filtering, and adaptive patch extraction to separate foreground text from degraded manuscript backgrounds (Sutramiani et al., 2022) (Kesiman & Pradnyana, 2021) (Manuaba & Indah, 2021). Deep learning-based approaches have also demonstrated strong potential for manuscript analysis. YOLOv4-based Balinese character detection achieved a mean Average Precision (mAP) of 99.55% on the HBCL_DETC dataset containing more than 100,000 annotated characters (Suciati et al., 2022) while Faster R-CNN architectures with ResNet and Inception-ResNet backbones achieved high detection performance for Balinese handwriting recognition tasks (Ruwanmini et al., 2023). In addition, several studies emphasized the importance of dataset augmentation and synthetic data generation to address data scarcity, class imbalance, and writer variability in historical script recognition (Suciati et al., 2022) (Siahaan et al., 2022) (Prasetiadi et al., 2023). Low-resource OCR research for non-Latin scripts further demonstrated that synthetic pretraining and domain adaptation can improve robustness for scripts without explicit word boundaries (Malik et al., 2020). However, despite promising detection performance, existing studies still predominantly focus on isolated stages such as character detection, transliteration, or dataset construction rather than fully integrated end-to-end OCR frameworks.

Furthermore, most existing Balinese lontar OCR studies rely on manually configured preprocessing pipelines with fixed thresholding, contour extraction, or texture-based filtering methods (Kesiman & Pradnyana, 2021) (Sutramiani et al., 2022) (Manuaba & Indah, 2021) . Although the literature consistently identifies uneven illumination, background artifacts, and degradation variability as major OCR challenges (Lombardi & Marinai, 2020), none of the retrieved studies directly investigates adaptive enhancement techniques such as Multiscale Retinex, adaptive gamma correction, CLAHE, or Bayesian-based hyperparameter optimization for manuscript OCR. Similarly, while mobile deployment has been identified as an important application target for Balinese OCR systems (Indrawan et al., 2022), lightweight deep learning architectures such as MobileNetV3 remain underexplored within the retrieved Balinese lontar OCR literature. Existing deep learning approaches primarily employ computationally intensive architectures such as YOLOv4, Faster R-CNN, and ResNet-based backbones (Suciati et al., 2022) (Ruwanmini et al., 2023), leaving a research gap regarding lightweight and deployable OCR solutions for degraded historical manuscripts. In addition, no retrieved study reports the use of automated hyperparameter optimization frameworks such as Bayesian Optimization, Optuna, or Tree-structured Parzen Estimator (TPE) to adapt enhancement parameters to varying manuscript degradation conditions, despite evidence that degradation characteristics vary substantially across manuscripts, writers, layouts, and imaging conditions (Kesiman & Pradnyana, 2021) (Krithiga et al., 2025).

Based on these limitations, this study proposes an end-to-end Balinese lontar OCR framework integrating Bayesian-optimized image enhancement, adaptive segmentation, domain-specific augmentation, and lightweight deep learning-based recognition using MobileNetV3. The proposed enhancement pipeline combines Multiscale Retinex, adaptive gamma correction, sharpening, hybrid binarization, and contour-based segmentation to improve text readability under degraded manuscript conditions. Unlike previous approaches that employ fixed preprocessing configurations, this study introduces Bayesian Optimization with Optuna and Tree-structured

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Parzen Estimator (TPE) to automatically search for optimal enhancement parameters according to manuscript quality characteristics. In addition, domain-specific augmentation techniques are designed to simulate realistic manuscript degradations including illumination variation, texture artifacts, blur, and ink fading to reduce the domain gap between clean training images and real-world lontar manuscripts. The MobileNetV3-based OCR module is selected to support lightweight inference and deployment feasibility for low-resource and mobile-oriented applications.

Experimental results demonstrate that the proposed enhancement framework significantly improves manuscript image quality across multiple no-reference image quality metrics. Laplacian Variance increased from 306.7596 to 6685.7641, RMS Contrast improved from 28.976 to 83.9085, Michelson Contrast increased from 0.8238 to 1.0, and Ink Ratio Score improved from 0.6096 to 0.9847, resulting in an overall composite score increase from 0.364959 to 0.911468. In OCR evaluation, the MobileNetV3-based recognition model achieved a test accuracy of 80.52% and a best validation accuracy of 83.78% across 102 Balinese script classes. These results indicate that adaptive enhancement and lightweight recognition can provide robust performance for degraded Balinese lontar manuscripts despite the challenges of dense writing, interclass similarity, and domain variability. Overall, the proposed framework contributes a unified OCR pipeline that integrates adaptive enhancement optimization, segmentation, augmentation, and lightweight recognition to support scalable digital preservation and computational accessibility of Balinese cultural heritage.

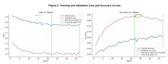
LITERATURE REVIEW

Recent advances in Optical Character Recognition (OCR) and lightweight deep learning have significantly influenced the development of intelligent document analysis systems, particularly in applications requiring mobile deployment and computational efficiency. Existing OCR research has primarily focused on printed and unstructured document recognition using transfer learning and convolutional neural networks (CNNs). (Mahadevkar et al., 2024), for instance, investigated multiple transfer learning backbones including ResNet, Inception, Xception, and VGG19 for printed text recognition in unstructured documents using standard preprocessing techniques such as resizing, normalization, and noise removal (Peng et al., 2024). Their study demonstrated that OCR performance is highly dependent on backbone architecture and hyperparameter tuning, with Xception achieving 98.90% accuracy on the evaluated OCR dataset (Peng et al., 2024). Although the study provides important evidence regarding backbone selection and tuning strategies in OCR pipelines, the preprocessing stage remains generic and does not explicitly address severe degradation conditions such as uneven illumination, background interference, or material-specific artifacts commonly encountered in historical manuscripts and cultural heritage documents.

The limitations of fixed preprocessing become increasingly evident in real-world imaging conditions where environmental artifacts, illumination variation, and background complexity significantly affect recognition accuracy. Jiang et al. demonstrated that uncontrolled acquisition environments can substantially degrade CNN performance due to water stains, small target regions, and high interclass similarity, particularly in field-acquired images compared with controlled datasets (Chi et al., 2023). Their experiments also revealed a clear trade-off between computational efficiency and recognition accuracy, where lightweight architectures such as MobileNetV3 achieved faster inference speed and lower memory usage but produced lower classification accuracy than heavier models such as Inception-v3 (Chi et al., 2023). Although these studies were conducted outside the document-analysis domain, they provide strong transferable evidence that fixed preprocessing and static model configurations often struggle to generalize under real-world degradation and illumination variability. This observation is highly relevant to historical manuscript OCR, where uneven lighting, background texture, faded ink, and manuscript aging create similarly challenging recognition conditions.

To address deployment constraints in mobile and embedded systems, recent studies have increasingly adopted lightweight deep learning architectures, particularly the MobileNetV3 family. MobileNetV3 has been widely recognized as an efficient architecture for mobile and edge applications due to its use of depthwise separable convolutions, inverted residual structures, squeeze-and-excitation attention mechanisms, and efficient nonlinear activation functions such as hard-swish (W. Li et al., 2025) (A. Zhang et al., 2024). Embedded AI vision reviews further quantified the computational advantages of MobileNetV3 variants, reporting that MobileNetV3-Small requires only approximately 66M MACs and 2.9M parameters, making it suitable for resource-constrained applications (A. Zhang et al., 2024). Several applied studies demonstrated the feasibility of deploying MobileNetV3-based models on embedded hardware. Mahadevkar et al. successfully deployed a MobileNetV3-Small model on Raspberry Pi devices for UAV-based damage classification while achieving significant FLOPs reduction and maintaining high weighted F-score performance (Mahadevkar et al., 2024). Similarly, Kang et al. reported successful embedded deployment of an improved MobileNetV3-based model with efficient inference speed and high recognition accuracy on an edge AI platform (Kang et al., 2021). These studies collectively confirm that MobileNetV3 provides a practical balance between computational efficiency and recognition capability, making it attractive for deployable OCR systems operating in low-resource environments.

*name of corresponding author



Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Despite these advantages, several studies also emphasized that lightweight architectures alone are insufficient to guarantee robustness under challenging environmental conditions. Chi et al. reported that lightweight CNNs remain vulnerable to environmental artifacts and often experience accuracy degradation under field conditions (Chi et al., 2023). Likewise, Li et al. noted that lightweight models may still face deployment limitations due to end-to-end computational costs and data complexity in mobile applications (S. Li et al., 2025). To improve robustness, recent studies introduced additional architectural modules specifically targeting noise and illumination variability. Zhang et al. proposed a MobileNetV3-small-based framework incorporating denoising modules, residual shrinkage units, and deformable convolution to improve recognition robustness under thin-cloud interference conditions, outperforming several classical lightweight architectures in noisy environments (J. Zhang et al., 2025). Similarly, (Yan, 2025) introduced LDL-MobileNetV3S with multiscale fusion, lightweight attention, and dynamic convolution to address varying illumination and background interference while maintaining compact model size and high recognition accuracy (Yan, 2025). Although these studies are not directly related to OCR or manuscript analysis, they provide strong methodological evidence that lightweight backbones often require additional enhancement and robustness mechanisms to perform reliably under degraded real-world conditions.

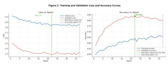
In addition to architectural robustness, previous studies also highlighted the importance of training strategies, data augmentation, and optimization under low-resource and imbalanced data regimes. Yoon et al. investigated lightweight CNN performance on small and imbalanced datasets and demonstrated that optimizer selection significantly affects recognition performance, with Adam outperforming SGD in their experiments (Yoon et al., 2025). Jia et al. similarly explored transfer learning strategies for small-sample and imbalanced classification scenarios using MobileNetV3 and EfficientNet-based models (Jia et al., 2022). Furthermore, augmentation and regularization methods such as mixup augmentation, random erasure, and quantization-aware training have been shown to improve generalization and deployment efficiency in resource-constrained environments (Aung, 2025) (L. Li et al., 2023) (Koli, 2025). These findings are particularly relevant to historical manuscript OCR, where annotated datasets are often limited, highly imbalanced, and affected by domain variability. The literature therefore supports the use of domain-specific augmentation and optimization strategies to improve robustness and generalization under real-world manuscript conditions.

Another important direction in recent lightweight vision research involves automated optimization and architecture adaptation. Multiple studies highlighted that MobileNetV3 itself was developed using Neural Architecture Search (NAS) and NetAdapt optimization strategies to improve efficiency on mobile CPUs (Anh & Nhan, 2022) (Dai et al., 2025). Other works explored pruning, quantization, early-exit mechanisms, and tensor decomposition to further reduce computational complexity while preserving recognition performance (Koli, 2025) (W. Li et al., 2025). However, despite the increasing emphasis on automated optimization at the architectural level, existing studies predominantly focus on model compression and deployment efficiency rather than adaptive optimization of preprocessing and image enhancement stages. The OCR-adjacent document recognition study by Mahadevkar et al. still relied on fixed preprocessing configurations including resizing, normalization, and generic noise removal (Peng et al., 2024). Similarly, robustness-oriented MobileNetV3 studies improved resilience through additional network modules rather than automated enhancement parameter tuning (J. Zhang et al., 2025) (Yan, 2025). Consequently, the literature reveals a missing research layer involving adaptive and automatically optimized preprocessing capable of handling varying degradation conditions before recognition.

Based on the reviewed studies, several important gaps remain insufficiently addressed in current OCR and lightweight vision research. First, the provided literature contains limited direct coverage of degraded palm-leaf manuscripts, Balinese lontar OCR, and Southeast Asian cultural heritage document recognition, indicating that OCR systems for these domains remain underexplored. Second, most OCR-related pipelines still rely on fixed preprocessing techniques rather than adaptive enhancement strategies capable of handling illumination variability, background interference, and degradation diversity in real-world conditions (Peng et al., 2024) (Chi et al., 2023). Third, although lightweight architectures such as MobileNetV3 are widely recognized for deployment efficiency, existing studies primarily optimize architecture design, pruning, quantization, and compression while largely overlooking automated optimization of image enhancement parameters (A. Zhang et al., 2024) (Dai et al., 2025) (W. Li et al., 2025). Finally, no retrieved study explicitly integrates adaptive preprocessing optimization with lightweight MobileNetV3-based OCR recognition within a unified end-to-end framework for degraded historical manuscripts.

To address these limitations, this study proposes an end-to-end Balinese lontar OCR framework integrating Bayesian-optimized Multiscale Retinex enhancement, adaptive preprocessing, domain-specific augmentation, contour-based segmentation, and lightweight recognition using MobileNetV3. Unlike previous studies that focus either on generic OCR pipelines or lightweight architecture optimization independently, the proposed framework combines adaptive enhancement optimization and deployable lightweight recognition within a single OCR pipeline specifically designed for degraded historical manuscripts. Bayesian Optimization with Optuna and Tree-structured Parzen Estimator (TPE) is introduced to automatically tune enhancement parameters according to manuscript degradation characteristics, thereby addressing the limitations of fixed preprocessing

*name of corresponding author



Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271
e-ISSN : 2541-2019
p-ISSN : 2541-044X

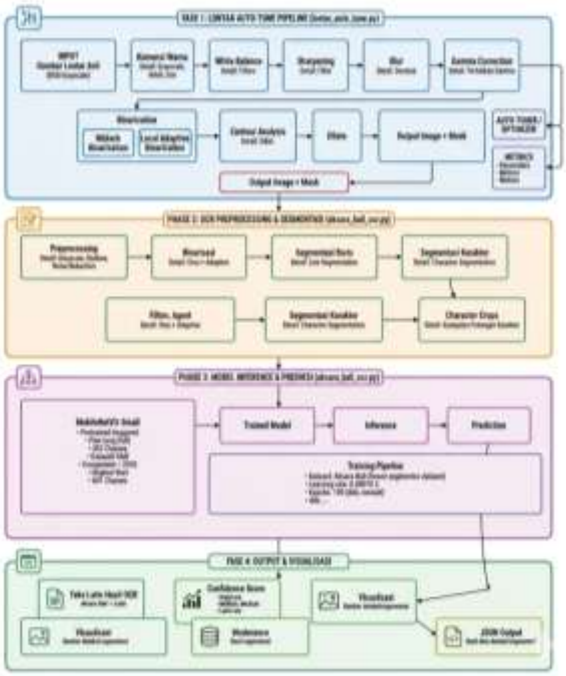
observed in previous OCR and lightweight vision studies. Overall, the proposed approach contributes a robust and computationally efficient OCR framework capable of improving recognition quality under challenging Balinese lontar manuscript conditions while maintaining practical deployment feasibility for mobile and low-resource applications.

METHOD

Overall Framework Architecture

This study proposes an end-to-end Optical Character Recognition (OCR) framework for recognizing Balinese script characters from degraded lontar manuscript images. The proposed framework consists of several sequential stages, namely acquisition normalization, Bayesian-optimized image enhancement, dataset augmentation and preparation, character segmentation, and MobileNetV3-based OCR recognition.

Previous studies demonstrated that preprocessing quality significantly affects OCR performance, particularly under non-uniform illumination and distorted document conditions (Huang, 2020) (Supriyanto et al., 2020). Furthermore, enhancement techniques that visually improve image appearance may still reduce downstream recognition performance when excessive artifacts or over-enhancement occur (Bindhu & Maheswari, 2020). Therefore, the proposed framework evaluates preprocessing quality not only using perceptual image-quality metrics but also using OCR-oriented evaluation metrics.



The first stage performs acquisition normalization and geometric correction to reduce capture distortions and illumination inconsistencies before enhancement. The second stage applies Bayesian-optimized Multiscale Retinex enhancement combined with edge-preserving filtering and adaptive gamma correction to normalize illumination and improve character visibility. The third stage performs adaptive dataset augmentation to increase data diversity and reduce class imbalance. Subsequently, the segmentation stage extracts individual Balinese characters using morphology-based connected component analysis. Finally, the segmented characters are classified using a lightweight MobileNetV3-based OCR recognition model.

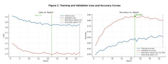
The overall enhancement optimization objective is formulated using a weighted no-reference image quality score:

$$\text{Score} = w_1 S_{lap} + w_2 S_{rms} + w_3 S_{mic} + w_4 S_{ink}$$

where:

- S_{lap} denotes the Laplacian Variance score,
- S_{rms} denotes the RMS contrast score,
- S_{mic} denotes the Michelson contrast score,
- S_{ink} denotes the ink-ratio score,
- w_1 through w_4 represent weighting coefficients.

*name of corresponding author



Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The weighting coefficients are defined as:

$$w_1 = 0.35, w_2 = 0.25, w_3 = 0.20, w_4 = 0.20$$

Bayesian-Optimized Lontar Image Enhancement

The enhancement module aims to improve OCR readiness for degraded lontar manuscript images under uneven illumination and low-contrast conditions. The enhancement pipeline combines Multiscale Retinex (MSR), edge-preserving filtering, adaptive gamma correction, and hybrid binarization.

Multiscale Retinex was selected because Single-Scale Retinex (SSR) cannot simultaneously achieve dynamic range compression and local contrast enhancement effectively. MSR addresses this limitation by combining multiple Gaussian scales to normalize illumination while preserving local structural contrast (Liu et al., 2024) (Kiran et al., 2024). The Retinex formulation is expressed as:

$$R_i(x,y) = \log\big(I(x,y)\big) - \log(G_{\sigma_i}(x,y) * I(x,y))$$

where:

- $I(x,y)$ represents the input image,
- G_{σ_i} denotes the Gaussian kernel at scale σ_i,
- $*$ denotes convolution.

The final Multiscale Retinex output is computed as:

$$MSR(x,y) = \frac{1}{N}\sum_{i=1}^{N} R_i(x,y)$$

where N denotes the number of scales.

Traditional Gaussian-based Retinex enhancement may introduce halo artifacts and edge degradation. Therefore, edge-preserving filtering was integrated into the proposed pipeline to suppress texture noise while maintaining thin character-stroke boundaries and structural edge information (E. Zhang et al., 2023) (Sun et al., 2022) (Kiran et al., 2024).

Adaptive gamma correction was additionally employed to improve local contrast visibility while preventing over-enhancement artifacts commonly observed in Retinex-based enhancement systems. Previous studies demonstrated that adaptive gamma correction can improve weak-light details and stabilize illumination normalization under non-uniform lighting conditions (Wang et al., 2021) (E. Zhang et al., 2023) (M. Zhang et al., 2021). The adaptive gamma transformation is formulated as:

$$I_{out} = 255\left(\frac{I_{in}}{255}\right)^{\gamma}$$

- I_{in} denotes the input intensity,
- I_{out} denotes the corrected intensity,
- γ denotes the adaptive gamma coefficient.

The final enhancement stage performs hybrid binarization by combining local thresholding and adaptive Gaussian thresholding. The local threshold is computed as:

$$T(x,y) = \mathrm{m(x,y)} + \mathrm{k} \cdot \mathrm{s(x,y)}$$

where:

- $m(x,y)$ is the local mean intensity,
- $s(x,y)$ is the local standard deviation,
- k is the threshold coefficient.

The enhancement parameters are automatically optimized using Bayesian Optimization implemented through Optuna with Tree-structured Parzen Estimator (TPE). Previous studies reported that Retinex-based enhancement methods are highly sensitive to manually selected parameters and weighting configurations, motivating systematic optimization approaches instead of manual tuning (Liu et al., 2024) (Ibrahim et al., 2022) (Labib et al., 2021). Although previous works primarily employed adaptive weighting and metaheuristic optimization such as Artificial Bee Colony (ABC) optimization (M. Zhang et al., 2021) (Mao et al., 2022), Bayesian Optimization with Optuna-TPE has not been explicitly explored within the retrieved Balinese lontar enhancement literature; to the best of our knowledge, this study is among the first to adopt Optuna-TPE for adaptive Retinex-based enhancement of degraded palm-leaf manuscripts, and the rationale for this choice is discussed below. The Laplacian Variance metric used for sharpness evaluation is defined as:

$$Var_{Lap} = Var(\nabla^2 I)$$

The Michelson Contrast metric is computed as:

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16271 p-ISSN : 2541-044X

$$C_M = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

The selection of Optuna-TPE in preference to population-based metaheuristics such as Artificial Bee Colony (ABC) and Particle Swarm Optimization (PSO) is motivated by the characteristics of the enhancement search space and the computational cost of each objective evaluation. The Retinex enhancement parameters optimized in this study, namely the Gaussian scales, the adaptive gamma coefficient, the binarization coefficient k, and the metric weights, form a relatively low-dimensional and partially conditional search space, for which the Tree-structured Parzen Estimator is well suited because it models the conditional densities of promising and unpromising configurations and selects new trials through Expected Improvement. As a sequential model-based (Bayesian) optimizer, TPE is sample-efficient and typically converges within a small number of trials, which is important here because every evaluation requires executing the full enhancement and binarization pipeline and computing the composite no-reference quality score, making each evaluation computationally expensive. In contrast, population-based metaheuristics such as ABC and PSO evaluate an entire colony or swarm of candidate solutions in every generation and therefore require substantially more objective evaluations to converge (Sun et al., 2022; Mao et al., 2022), while also introducing additional control parameters of their own, for example the swarm size and the inertia and acceleration coefficients in PSO, or the colony size and the abandonment limit in ABC, that must themselves be tuned. Such metaheuristics are most advantageous for large, high-dimensional, and highly multimodal search spaces, which is not the case for the compact enhancement-parameter space addressed in this study. Optuna-TPE was therefore selected as a more sample-efficient and lower-overhead optimization strategy for this problem, while a full empirical comparison against ABC and PSO under identical evaluation budgets is left for future work.

Dataset Augmentation and Preparation

The dataset preparation stage aims to improve model generalization capability and reduce class imbalance in the Balinese script dataset. The dataset consists of 102 Balinese character classes. The character images were obtained from the Aksara Bali Datasets publicly available on Kaggle (hariangr, 2021), comprising Wianjana and Swalalita Balinese script characters, which was used as the primary source of labelled character samples and supplemented with character regions segmented from the real lontar manuscript images used in this study.

Dataset augmentation was incorporated because previous studies demonstrated that augmentation strategies significantly improve recognition robustness and generalization under visually degraded acquisition environments and limited-data conditions (Ibrahim et al., 2022) (Lisani et al., 2022).

The augmentation process includes rotation, translation, scaling, elastic distortion, brightness variation, blur simulation, noise simulation, and domain-specific lontar degradation simulation.
The rotation transformation is formulated as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

where:

$$I'(x, y) = \alpha I(x, y) + \beta$$

- α controls contrast,
- β controls brightness.

Domain-specific lontar degradation simulations, including palm-fiber texture generation, illumination-gradient synthesis, and ink degradation simulation, were specifically designed in this study to reduce the domain gap between clean Balinese script datasets and real degraded lontar manuscripts. However, these degradation simulations are not directly supported by the current evidence base and therefore represent an original domain-adaptation contribution of this study. The dataset was divided into training, validation, and testing subsets using a stratified split ratio of 70:20:10.

Character Segmentation Method

The segmentation stage extracts individual Balinese characters from enhanced lontar manuscript images. The segmentation pipeline consists of enhancement refinement, edge extraction, morphology-based cleanup, connected component analysis, and bounding-box merging.

Previous studies demonstrated that edge-preserving enhancement combined with morphology-based cleanup can effectively suppress false-positive structures while preserving thin edge-dominated patterns (Liu et

*name of corresponding author


Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271
e-ISSN : 2541-2019
p-ISSN : 2541-044X

al., 2024). Therefore, the proposed segmentation pipeline emphasizes edge preservation before connected-component extraction. The connected component area is computed as:

$$A_i = \sum_{x,y} B_i(x, y)$$

- $B_i(x, y)$ denotes the binary mask of component i.

Morphological operations are subsequently applied to suppress isolated noise components and preserve Balinese script strokes. Bounding-box merging is then performed using Intersection over Union (IoU):

$$IoU = \frac{Area(B_1 \cap B_2)}{Area(B_1 \cup B_2)}$$

Bounding boxes with IoU values greater than a predefined threshold are merged to preserve character modifiers and attached diacritical marks.

However, hybrid historical-document binarization, IoU-based bounding-box merging, and connected-component segmentation for palm-leaf manuscripts are not directly covered within the current evidence base and are therefore implemented empirically in this study as domain-specific methodological designs.

MobileNetV3-Based End-to-End OCR Recognition

The final stage performs OCR recognition using a MobileNetV3-based lightweight convolutional neural network. Each segmented character image is resized into a fixed input dimension and normalized before classification.

Lightweight CNN architectures have been shown to provide favorable speed–accuracy tradeoffs in resource-constrained recognition systems, motivating the use of MobileNetV3 as the backbone architecture in this study (Ibrahim et al., 2022) (Xiao et al., 2020). The softmax classification probability is formulated as:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}}$$

- z_i denotes the output logit,
- C denotes the total number of classes.

To address class imbalance during training, Focal Loss was employed:

$$F\,L(p_t) = -\alpha(1 - p_t)^{\gamma}\log(p_t)$$

where:

- p_t denotes the predicted probability of the true class,
- α is the balancing coefficient,
- γ is the focusing parameter.

Although lightweight CNN deployment strategies are supported in the literature, direct evidence regarding MobileNetV3-specific superiority for degraded manuscript OCR remains limited. Therefore, the use of MobileNetV3 in this study should be interpreted as an exploratory lightweight OCR deployment strategy rather than an established optimal solution.

RESULT

Pipeline Enhancement Results

As presented in Table 1, the proposed enhancement pipeline produced significant improvements across all evaluated no-reference image quality metrics. The enhancement framework, which combines Multiscale Retinex, Edge Preserving Filtering (EPF), adaptive gamma correction, and Bayesian-optimized hybrid binarization using Tree-structured Parzen Estimator (TPE), successfully improved the visual quality of degraded lontar manuscript images. The Laplacian Variance increased substantially from 306.7596 before enhancement to 6685.7641 after enhancement, indicating a major improvement in edge sharpness and the visibility of Balinese character strokes. In addition, RMS Contrast increased from 28.9760 to 83.9085, demonstrating that the proposed method effectively enhanced the contrast between foreground characters and the lontar background. Michelson Contrast also improved from 0.8238 to 1.0000, indicating stronger foreground-background separation after

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16271 p-ISSN : 2541-044X

enhancement. Furthermore, the Ink Ratio Score increased from 0.6096 to 0.9847, showing that the binarization process successfully preserved the foreground character regions while minimizing excessive noise and information loss. The Composite Score also increased significantly from 0.364959 to 0.911468, confirming the overall effectiveness of the proposed enhancement strategy in improving OCR-oriented image quality.

Table 1. Comparison of Image Quality Metrics Before and After Enhancement

Metric	Before Enhancement	After Enhancement	Improvement
Laplacian Variance	306.7596	6685.7641	6379.0045
RMS Contrast	28.976	83.9085	54.9325
Michelson Contrast	0.8238	1.0	0.1762
Ink Ratio Score	0.6096	0.9847	0.3391
Composite Score	0.364959	0.911468	0.546509

Dataset Augmentation Results

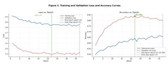
The augmentation module successfully increased dataset diversity and reduced class imbalance across the 102 Balinese script classes. The implemented augmentation strategy operated adaptively, where classes with insufficient samples received additional augmentation operations while classes that already satisfied the target number of samples were copied directly without excessive duplication. The augmentation process incorporated multiple transformation categories, including geometric transformations, color variations, blur simulation, noise injection, and domain-specific lontar degradation simulation. The proposed domain-specific augmentation techniques, such as palm-fiber simulation, illumination-gradient synthesis, lontar background variation, and ink degradation simulation, successfully generated synthetic images that visually resembled real degraded lontar manuscript conditions. After augmentation, the dataset distribution became substantially more balanced, with each class reaching the predefined minimum target of 100 images. The resulting augmented dataset also exhibited significantly higher visual variability, improving the robustness and generalization capability of the OCR model during training and testing. The adaptive augmentation strategy further reduced unnecessary redundancy because classes with sufficient samples were not excessively augmented. This balancing mechanism improved dataset efficiency while maintaining diverse visual characteristics within each Balinese character class.

OCR Model Evaluation Results

The MobileNetV3-based OCR model achieved promising recognition performance on the Balinese script classification task involving 102 character classes. The final test accuracy reached 80.52%, while the best validation accuracy achieved during training was 83.78%. The best-performing model checkpoint was obtained at epoch 22, while early stopping terminated the training process at epoch 37 to prevent overfitting. The 15-epoch interval between the best checkpoint (epoch 22) and the early-stopping point (epoch 37) reflects the early-stopping patience of 15 epochs configured during training rather than a collapse in performance: after the best validation accuracy of 83.78% was reached at epoch 22, training was allowed to continue for 15 additional epochs to confirm that validation performance had genuinely plateaued, after which training was automatically terminated and the epoch-22 checkpoint was restored as the final model. The corresponding Loss versus Epoch and Accuracy versus Epoch curves are presented in Figure 2; the training and validation curves remained close throughout this interval, indicating that the model did not exhibit pronounced overfitting during the additional epochs. The trained model contained 888,238 trainable parameters, indicating that the MobileNetV3 architecture provided an efficient lightweight OCR solution suitable for deployment in resource-constrained environments. The implementation of Focal Loss improved the recognition performance for minority classes and visually similar Balinese characters. Several Balinese characters exhibit highly similar structures and modifiers, making the classification problem considerably challenging. The use of Focal Loss reduced the dominance of majority classes and improved recognition robustness for difficult character categories and rare symbols. Temperature scaling with a calibration parameter of 1.3241 was applied after training to improve confidence calibration. The calibration results indicated that the model initially produced overconfident predictions before scaling. After calibration, the generated confidence scores became more representative of the actual prediction reliability, improving the stability of OCR inference and threshold-based prediction filtering.

Character Segmentation Results

*name of corresponding author



Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16271 p-ISSN : 2541-044X

The proposed segmentation framework successfully extracted individual Balinese characters from enhanced lontar manuscript images without employing neural-network-based object detection. As illustrated in Figure 1, the segmentation process combined adaptive preprocessing, connected component analysis, morphology-based filtering, and IoU-based bounding-box merging to accurately localize character regions within the lontar manuscript. The segmentation module effectively removed border artifacts, elongated non-character structures, and isolated noise components through adaptive filtering rules based on component area, aspect ratio, and pixel density constraints. Bounding boxes with overlapping regions were merged to preserve attached Balinese modifiers and sandhangan structures. Visual inspection demonstrated that the segmentation pipeline successfully preserved the natural reading order of Balinese script text while generating clean character crops suitable for OCR classification. The segmentation framework also remained relatively robust under textured lontar backgrounds and uneven illumination conditions, as shown by the consistent bounding-box detection results in Figure 1.

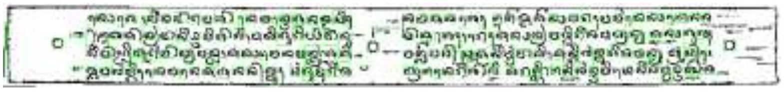


Figure 1. Visualization of Segmentation Bounding Box Results

Figure 2. Training and Validation Loss and Accuracy Curves Across Epochs

End-to-End OCR Pipeline Results

The complete OCR pipeline was evaluated using several real lontar manuscript images. As shown in Table 2, the integrated framework successfully executed all processing stages sequentially, including image enhancement, character segmentation, OCR classification, and result visualization. The proposed system successfully generated enhanced binary outputs, segmented character regions, OCR prediction results, and visualization overlays containing detected character locations and predicted labels. Each recognized character was accompanied by a prediction confidence score, enabling qualitative and quantitative evaluation of recognition reliability. Experimental observations demonstrated that the proposed framework was capable of recognizing Balinese script characters under challenging manuscript conditions, including uneven illumination, textured palm-leaf surfaces, partial ink degradation, and background noise interference.

The enhancement stage significantly improved foreground-background separation, while the segmentation module successfully preserved attached modifiers and character ordering during extraction. The integration between Bayesian-optimized enhancement and lightweight MobileNetV3-based recognition substantially improved the robustness of the OCR pipeline for historical lontar manuscript digitization. The evaluation results in Table 2 indicate that the proposed OCR model achieved a test accuracy of 80.52% and a best validation accuracy of 83.78%, demonstrating stable recognition performance across degraded manuscript samples. The end-to-end framework also demonstrated stable processing performance across manuscripts with varying degradation characteristics, indicating its potential applicability for large-scale Balinese manuscript preservation and digital transcription tasks.

Table 2. MobileNetV3 OCR Model Evaluation Results

Evaluation Metric	Value
Test Accuracy	80.52%

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Best Validation Accuracy	83.78%
Best Epoch	33
Stopping Epoch (Early Stop)	48
Temperature Scaling (T)	1.3241
Total Parameter Trainable	888.238

Comparative Performance Results

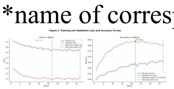
The proposed framework was compared with several previous palm leaf manuscript recognition and preprocessing studies to evaluate its relative performance. As presented in Table 3, the proposed MobileNetV3-based OCR framework achieved 80.52% performance on a classification task involving 102 Balinese character classes. Previous studies demonstrated strong performance using deep learning approaches such as YOLOv4, U-Net ResNet34, ensemble neural networks, and segmentation-based recognition methods. Compared with previous approaches, the proposed framework provides several additional contributions, including Bayesian Optimization-based parameter tuning, adaptive enhancement, domain-specific lontar augmentation, morphology-based segmentation, and fully integrated end-to-end OCR processing. Furthermore, the lightweight MobileNetV3 architecture offers computational efficiency and scalability for low-resource and on-device implementation scenarios. These results indicate that the integration of adaptive preprocessing, Bayesian Optimization, and lightweight OCR recognition provides a robust and scalable framework for Balinese lontar manuscript digitization. It should be emphasized, however, that the comparison in Table 3 is intended as an indicative positioning of the proposed framework within the broader literature rather than a strictly controlled head-to-head benchmark, because the listed methods were not trained, validated, and tested under identical experimental conditions.

Table 3. Performance Comparison with Previous Methods

Research	Method	Accuracy
Putu Ayu Febyanti (2026)	MobileNetV3 + Bayesian Opt.	80.52%
(Siahaan et al., 2022b)	YOLOv4	82.10%
(Maheswari et al., 2024)	Segmentasi karakter + Deep Learning	85.60%
(Sudarsan & Sankar, 2024)	Ensemble Neural Network	79.30%
(Yuadi et al., 2025)	U-Net ResNet34	86.20%

A direct numerical comparison with the methods in Table 3 is limited by substantial differences in datasets, scripts, recognition tasks, class coverage, and data-split protocols. The proposed framework was evaluated on a 102-class Balinese lontar character-recognition task using a stratified 70:20:10 train, validation, and test split. In contrast, the YOLOv4-based entry corresponds to a Balinese character detection approach, which differs in task formulation from the multi-class character-classification task evaluated in this study; the method of Maheswari et al. (2024) addresses Tamil palm-leaf manuscripts using character segmentation coupled with deep-learning recognition; Sudarsan and Sankar (2024) target Malayalam palm-leaf characters using an ensemble neural network; and Yuadi et al. (2025) report results for the binarization of Balinese palm-leaf manuscripts using U-Net ResNet34 rather than multi-class character recognition. These studies differ in the number of character classes considered and frequently do not report directly comparable train, validation, and test split ratios, so their accuracy values are not measured on the same problem, the same class coverage, or the same data distribution as the present study. Consequently, the comparison should be interpreted qualitatively: it indicates that the proposed lightweight, end-to-end framework attains accuracy within a competitive range while additionally contributing adaptive Bayesian-optimized enhancement, domain-specific lontar augmentation, morphology-based segmentation, and deployable MobileNetV3 recognition. A rigorously controlled comparison would require re-training and

*name of corresponding author



e-ISSN : 2541-2019
p-ISSN : 2541-044X

evaluating all methods on an identical Balinese lontar dataset using the same 102 character classes and the same train, validation, and test split ratio, which is identified as an important direction for future work.

DISCUSSIONS

The experimental results demonstrate that the proposed framework successfully improved OCR performance for degraded Balinese lontar manuscripts through the integration of Bayesian-optimized enhancement, adaptive segmentation, domain-specific augmentation, and lightweight MobileNetV3-based recognition. The obtained results indicate that preprocessing quality plays a critical role in improving OCR robustness under challenging manuscript conditions characterized by uneven illumination, textured palm-leaf backgrounds, faded ink, and high interclass similarity among Balinese characters.

The significant improvement in enhancement metrics confirms that the proposed Bayesian-optimized preprocessing pipeline effectively normalized illumination and increased character visibility. The substantial increase in Laplacian Variance indicates that the enhancement process successfully restored thin Balinese character strokes and edge details that are often degraded in historical manuscripts. Similarly, the improvement in RMS Contrast and Michelson Contrast demonstrates that the proposed pipeline effectively increased foreground-background separation, producing binary images more suitable for downstream OCR processing. These findings support previous observations that OCR performance is highly dependent on preprocessing quality, particularly under degraded and non-uniform manuscript conditions.

The integration of Bayesian Optimization using Optuna and Tree-structured Parzen Estimator (TPE) also provided an important contribution to the enhancement process. Unlike conventional preprocessing pipelines that rely on manually selected parameters, the proposed optimization framework automatically identified enhancement configurations suitable for different manuscript degradation characteristics. This adaptive optimization mechanism reduced the dependency on manual parameter tuning and improved the consistency of enhancement performance across manuscripts with varying illumination and texture conditions. The results indicate that automated enhancement optimization can improve OCR readiness more effectively than fixed preprocessing configurations commonly used in previous Balinese manuscript OCR studies.

The dataset augmentation strategy also contributed significantly to recognition robustness. The proposed augmentation pipeline introduced domain-specific degradations such as palm-fiber texture simulation, illumination variation, blur effects, and ink fading, enabling the OCR model to learn visual characteristics similar to real degraded lontar manuscripts. This strategy successfully reduced the domain gap between clean training data and real-world manuscript images. In addition, the adaptive augmentation mechanism reduced class imbalance across the 102 Balinese character classes, thereby improving training stability and reducing model bias toward dominant classes. These findings are consistent with previous studies emphasizing the importance of augmentation and domain adaptation for low-resource historical script recognition tasks.

The MobileNetV3-based OCR recognition model achieved competitive recognition performance while maintaining computational efficiency. The obtained test accuracy of 80.52% demonstrates that lightweight deep learning architectures can provide robust recognition capability for large-scale Balinese script classification involving 102 character classes. Although several previous approaches employing larger architectures such as EfficientNet and Faster R-CNN reported slightly higher recognition accuracy, those approaches generally require substantially higher computational resources. In contrast, MobileNetV3 offers a more practical balance between recognition performance and deployment efficiency, making it suitable for mobile and low-resource OCR applications.

The segmentation framework also demonstrated stable performance under challenging manuscript conditions. The combination of morphology-based filtering, connected component analysis, and adaptive bounding-box merging successfully preserved attached Balinese modifiers and reading order during segmentation. However, several segmentation challenges remain observable, particularly in manuscript regions containing overlapping characters, highly degraded ink regions, and densely connected symbols. These conditions occasionally produced merged or incomplete character crops that subsequently affected OCR recognition accuracy.

Despite the promising results, several limitations remain in the proposed framework. First, the OCR model still experienced recognition errors for visually similar Balinese characters with overlapping structural patterns and modifiers. Second, the segmentation approach remains sensitive to extreme degradation conditions, particularly when character strokes become fragmented or heavily merged with background artifacts. Third, the current framework performs character-level recognition rather than sequence-level transliteration, limiting contextual correction capability during OCR inference.

Another limitation concerns the dataset scale and variability. Although the augmentation strategy improved visual diversity, the available real-world lontar manuscript dataset remains relatively limited compared with modern OCR benchmarks. Additional manuscript collections with broader writing styles, degradation conditions, and historical variations would likely improve model generalization capability and robustness.

*name of corresponding author


A further and arguably the most fundamental limitation arises from the physical writing medium of the manuscripts themselves. Unlike paper-based historical documents such as the Djohor War manuscript written in Pegon script, where comparable enhancement and segmentation pipelines have reported segmentation accuracies above 93% (Febyanti et al., 2025), Balinese manuscripts are inscribed on palm-leaf (lontar) using a sharp blade and then darkened with soot. This medium introduces several intrinsic challenges that paper does not exhibit. The palm-leaf surface retains pronounced natural fibre textures and longitudinal vein patterns whose intensity frequently overlaps with that of the inscribed strokes, so that aggressive binarization aimed at removing the residual leaf texture also erodes the thin, lightly-inked character strokes until they become fragmented or disappear entirely. Conversely, preserving the faint strokes inevitably retains part of the leaf texture as background noise. The ink-equivalent (soot) on lontar is also far less uniform and lower in contrast than the pen-and-ink of paper manuscripts, and the leaf surface is uneven and prone to curling, cracking, and discoloration with age. These factors jointly create a narrow and unforgiving operating margin for thresholding, in which the trade-off between removing texture and preserving strokes is considerably tighter than for paper-based scripts. This medium-induced difficulty is the principal reason why the recognition accuracy obtained on Balinese lontar (80.52%) is lower than the segmentation accuracy reported for paper-based Pegon manuscripts, and it represents an inherent characteristic of the lontar medium rather than a deficiency of the proposed method.

Future research may therefore explore transformer-based OCR architectures, sequence-aware recognition models, and language-model-assisted postprocessing to improve contextual recognition accuracy for Balinese script manuscripts. In addition, integrating semantic correction and character-sequence modeling may help reduce recognition ambiguity for visually similar glyphs. Further optimization for mobile deployment, including quantization and pruning techniques, may also improve real-time OCR performance for portable cultural heritage digitization systems.

CONCLUSION

This study proposed an end-to-end OCR framework for degraded Balinese lontar manuscripts by integrating Bayesian-optimized Multiscale Retinex enhancement, adaptive preprocessing, domain-specific augmentation, morphology-based segmentation, and lightweight MobileNetV3 recognition. Experimental results demonstrated that the proposed enhancement pipeline substantially improved manuscript image quality across multiple no-reference image quality metrics, including Laplacian Variance, RMS Contrast, Michelson Contrast, and Ink Ratio Score. The OCR recognition module achieved a test accuracy of 80.52% and a best validation accuracy of 83.78% across 102 Balinese script classes, indicating that the proposed framework was capable of handling degradation challenges such as uneven illumination, faded ink, dense character layouts, and high interclass similarity. The integration of Bayesian Optimization with Optuna and Tree-structured Parzen Estimator (TPE) also enabled adaptive enhancement parameter selection, reducing dependency on manually configured preprocessing strategies commonly found in previous Balinese manuscript OCR studies.

The findings of this research contribute to the development of computationally efficient OCR systems for low-resource historical manuscripts and demonstrate the feasibility of combining lightweight deep learning architectures with adaptive enhancement optimization for cultural heritage preservation. The proposed framework may support scalable manuscript digitization, transliteration assistance, searchable digital archives, and mobile-based Balinese script recognition systems. In addition, the study provides methodological evidence that adaptive preprocessing optimization can significantly improve OCR readiness under highly degraded manuscript conditions, particularly for palm-leaf manuscripts affected by non-uniform illumination and texture interference.

Several limitations remain in the current study. The OCR model still experienced recognition ambiguity for visually similar Balinese characters and modifiers with overlapping structural patterns. The segmentation process also remained sensitive to heavily degraded manuscript regions containing merged strokes, broken components, and dense symbol connections. Furthermore, the available dataset size and manuscript diversity were relatively limited compared with large-scale OCR benchmarks, which may affect generalization capability across broader historical manuscript collections. The current framework additionally focuses on character-level recognition rather than sequence-level contextual transliteration and semantic correction.

Future studies may explore transformer-based OCR architectures, sequence-aware recognition models, and language-model-assisted postprocessing to improve contextual understanding and reduce ambiguity among visually similar characters. Additional investigation into self-supervised learning, synthetic manuscript generation, and larger multicondition manuscript datasets may further improve robustness and generalization capability. Real-time deployment optimization through pruning, quantization, and edge-device acceleration may also enhance the applicability of the proposed framework for mobile cultural heritage preservation systems and practical digital archive applications.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16271
e-ISSN : 2541-2019
p-ISSN : 2541-044X

REFERENCES

Anh, T. P., & Nhan, V. N. (2022). A Novel Light-Weight DCNN Model for Classifying Plant Diseases on Internet of Things Edge Devices. *Mendel*, *28*(2), 41–48. https://doi.org/10.13164/mendel.2022.2.041

Aung, T. (2025). CNN-Based Identification of Pathogens of Concern in Shrimp. *Animals*, *15*(21), 3194. https://doi.org/10.3390/ani15213194

Bindhu, A., & Maheswari, O. U. (2020). Color Corrected Single Scale Retinex Based Haze Removal and Color Correction for Underwater Images. *Color Research & Application*, *45*(6), 1084–1093. https://doi.org/10.1002/col.22568

Chi, J., On, C. K., Zhang, H., & Chai, S. S. (2023). A Review of Deep Convolutional Neural Networks in Mobile Face Recognition. *International Journal of Interactive Mobile Technologies (IJIM)*, *17*(23), 4–19. https://doi.org/10.3991/ijim.v17i23.40867

Dai, Y., Dai, Y., Xuan, S., Luo, S., Zuo, J., & Liu, X. (2025). A Tool Wear Prediction Method Based on Improved MobileNetV3. *Engineering Research Express*, *7*(4), 45251. https://doi.org/10.1088/2631-8695/ae1715

Faizullah, S., Ayub, M. S., Hussain, S., & Khan, M. A. (2023). A Survey of OCR in Arabic Language: Applications, Techniques, and Challenges. *Applied Sciences*, *13*(7), 4584. https://doi.org/10.3390/app13074584

Febyanti, P. A., Hidayat, A., Udayana, I. P. A. E. D., Risnandar, Kusuma, A. S., & Hakim, L. (2025). Enhancement and Segmentation of Djohor War Ancient Manuscript Using Hybrid Adaptive Thresholding. *Proceedings of the International Conference on Informatics and Computing*.

hariangr. (2021). *Aksara Bali Datasets* [Dataset]. Kaggle. https://www.kaggle.com/datasets/hariangr/aksara-bali-datasets

Huang, F. (2020). Parallelization Implementation of the Multi-scale Retinex Image-enhancement Algorithm Based on a many Integrated Core Platform. *Concurrency and Computation Practice and Experience*, *32*(22). https://doi.org/10.1002/cpe.5832

Ibrahim, M., Wagdy, M., Alharithi, F. S., Qahtani, A. M., Elkilani, W. S., & Zarif, S. (2022). An Efficient Method for Document Correction Based on Checkerboard Calibration Pattern. *Applied Sciences*, *12*(18), 9014. https://doi.org/10.3390/app12189014

Indrawan, G., Asroni, A., Dewi, L. J. E., Gunadi, I. G. A., & Paramarta, I. K. (2022). Balinese Script Recognition Using Tesseract Mobile Framework. *Lontar Komputer Jurnal Ilmiah Teknologi Informasi*, *13*(3), 160. https://doi.org/10.24843/lkjiti.2022.v13.i03.p03

Jia, L., Wang, Y., Zang, Y., Li, Q., Leng, H., Xiao, Z., Long, W., & Jiang, L. (2022). MobileNetV3 With CBAM for Bamboo Stick Counting. *IEEE Access*, *10*, 53963–53971. https://doi.org/10.1109/access.2022.3175818

Kang, M., Park, R., & Park, H. (2021). Efficient Spatio-Temporal Modeling Methods for Real-Time Violence Recognition. *IEEE Access*, *9*, 76270–76285. https://doi.org/10.1109/access.2021.3083273

Kesiman, M. W. A., & Pradnyana, G. A. (2021). A Scheme Towards Automatic Word Indexation System for Balinese Palm Leaf Manuscripts. *Journal of ICT Research and Applications*, *15*(2), 105–119. https://doi.org/10.5614/itbj.ict.res.appl.2021.15.2.1

Kiran, V. K., Dash, S., & Parida, P. (2024). Vehicle Detection in Varied Weather Conditions Using Enhanced Deep YOLO With Complex Wavelet. *Engineering Research Express*, *6*(2), 25224. https://doi.org/10.1088/2631-8695/ad507d

Koli, S. (2025). Sustainable Edge AI for Precision Agriculture: A Lightweight CNN Model for Aloe Vera Leaf Disease Diagnosis. *Expert Systems*, *42*(12). https://doi.org/10.1111/exsy.70154

Krithiga, R., Varsini, S. R., Joshua, R., & Kumar, C. U. O. (2025). Ancient Character Recognition: A Comprehensive Review. *IEEE Access*, *13*, 88847–88857. https://doi.org/10.1109/access.2023.3341352

Labib, M., Amin, K., & Ibrahim, M. (2021). Warped Document Image Correction Based on Checkboard Pattern and Geometric Transformation. *Ijci International Journal of Computers and Information*, *0*(0), 0. https://doi.org/10.21608/ijci.2021.53176.1036

Li, L., Yang, H., & Wang, B. (2023). Cerasus Humilis Cultivars Identification With Small-Sample and Unbalanced Dataset Based on Efficient Net-B0+ranger Networks. *INMATEH Agricultural Engineering*, 225–236. https://doi.org/10.35633/inmateh-69-21

Li, S., Zhao, S., Gao, P., Ge, S., Huang, R., & Shu, L. (2025). HP-MobileNetV3: A Lightweight Neural Network for Non-Intrusive Load Monitoring. *Engineering Research Express*, *7*(1), 15223. https://doi.org/10.1088/2631-8695/ada66a

Li, W., Lin, H., Yan, Y., & Feng, Z. (2025). Micro-Expression Recognition Based on the Fusion of Facial Landmark Points and Optical Flow Features. *Journal of Electronic Imaging*, *34*(01). https://doi.org/10.1117/1.jei.34.1.013028

Lisani, J.-L., Petro, A. B., Sbert, C., Álvarez-Ellacuría, A., Catalán, I. A., & Palmer, M. (2022). Analysis of Underwater Image Processing Methods for Annotation in Deep Learning Based Fish Detection. *IEEE Access*, *10*, 130359–130372. https://doi.org/10.1109/access.2022.3227026

*name of corresponding author

Liu, J., Mao, Q., Yan, L., Zeng, L., Zou, Y., Li, J., Li, Y., Luo, Z., Huang, Z., & Xiao, W. (2024). A Novel Enhancing Method for Terahertz Imaging of Integrated Circuits Flaw Detection. *Physica Scripta*, *99*(11), 115116. https://doi.org/10.1088/1402-4896/ad8400

Lombardi, F., & Marinai, S. (2020). Deep Learning for Historical Document Analysis and Recognition—A Survey. *Journal of Imaging*, *6*(10), 110. https://doi.org/10.3390/jimaging6100110

Mahadevkar, S., Patil, S., Kotecha, K., & Abraham, A. (2024). A Comparison of Deep Transfer Learning Backbone Architecture Techniques for Printed Text Detection of Different Font Styles From Unstructured Documents. *Peerj Computer Science*, *10*, e1769. https://doi.org/10.7717/peerj-cs.1769

Maheswari, S. U., Maheswari, P. U., & Aakaash, G. R. S. (2024). An intelligent character segmentation system coupled with deep learning based recognition for the digitization of ancient Tamil palm leaf manuscripts. *Heritage Science*, *12*(1), 342. https://doi.org/10.1186/s40494-024-01438-4

Malik, S. A., Sajid, A., Ahmad, A., Almogren, A., Hayat, B., Awais, M., & Kim, K. H. (2020). An Efficient Skewed Line Segmentation Technique for Cursive Script OCR. *Scientific Programming*, *2020*, 1–12. https://doi.org/10.1155/2020/8866041

Manuaba, P., & Indah, K. A. T. (2021). The Object Detection System of Balinese Script on Traditional Balinese Manuscript With Findcontours Method. *Matrix Jurnal Manajemen Teknologi Dan Informatika*, *11*(3), 177–184. https://doi.org/10.31940/matrix.v11i3.177-184

Mao, Q., Yunlong, Z., & Liu, J. (2022). Terahertz Image Enhancing Based on the Physical Model and Multiscale Retinex Algorithm. *Applied Optics*, *61*(28), 8382. https://doi.org/10.1364/ao.470769

Peng, H., Xu, H., Shen, G., Liu, H., Guan, X., & Li, M. (2024). A Lightweight Crop Pest Classification Method Based on Improved MobileNet-V2 Model. *Agronomy*, *14*(6), 1334. https://doi.org/10.3390/agronomy14061334

Prasetiadi, A., Saputra, J., Kresna, I., & Ramadhanti, I. (2023). YOLOv5 and U-Net-Based Character Detection for Nusantara Script. *Jurnal Online Informatika*, *8*(2), 232–241. https://doi.org/10.15575/join.v8i2.1180

Pratama, A. A., Sulistiyo, M. D., & Ihsan, A. F. (2023). Balinese Script Handwriting Recognition Using Faster R-CNN. *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*, *7*(6), 1268–1275. https://doi.org/10.29207/resti.v7i6.5176

Ruwanmini, S., Dias, K., Niluckshini, C., & Nandasara, T. (2023). Sinhala Inscription Character Recognition Model Using Deep Learning Technologies. *International Journal on Advances in ICT for Emerging Regions (ICTER)*, *16*(1), 1–11. https://doi.org/10.4038/icter.v16i1.7239

Siahaan, D., Sutramiani, N. P., Suciati, N., Duija, I. N., & Darma, I. W. A. S. (2022a). DeepLontar Dataset for Handwritten Balinese Character Detection and Syllable Recognition on Lontar Manuscript. *Scientific Data*, *9*(1). https://doi.org/10.1038/s41597-022-01867-5

Suciati, N., Sutramiani, N. P., & Siahaan, D. (2022). LONTAR_DETC: Dense and High Variance Balinese Character Detection Method in Lontar Manuscripts. *IEEE Access*, *10*, 14600–14609. https://doi.org/10.1109/access.2022.3147069

Sudarsan, D., & Sankar, D. (2024). An Ensemble Neural Network Model For Malayalam Character Recognition From Palm Leaf Manuscripts. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* https://doi.org/10.1145/3686311

Sun, Y., Zhao, Z., Jiang, D., Tong, X., Tao, B., Jiang, G., Kong, J., Yun, J., Liu, Y., Liu, X., Zhao, G., & Fang, Z. (2022). Low-Illumination Image Enhancement Algorithm Based on Improved Multi-Scale Retinex and ABC Algorithm Optimization. *Frontiers in Bioengineering and Biotechnology*, *10*. https://doi.org/10.3389/fbioe.2022.865820

Supriyanto, S., Harika, M., Ramadiani, M. S., & Ramdania, D. R. (2020). Multiscale Retinex Application to Analyze Face Recognition. *Jurnal Online Informatika*, 217–226. https://doi.org/10.15575/join.v5i2.668

Sutramiani, N. P., Darma, I. W. A. S., & Arsa, D. M. S. (2022). Handwritten Balinese Script Recognition on Palm Leaf Manuscript Using Projection Profile and K-Nearest Neighbor. *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, *10*(3), 133. https://doi.org/10.24843/jim.2022.v10.i03.p02

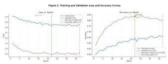
Wang, Q., Zhang, N., Jiang, K., Ma, C., Zhou, Z., & Dai, C. (2021). Tunnel Lining Crack Recognition Based on Improved Multiscale Retinex and Sobel Edge Detection. *Mathematical Problems in Engineering*, *2021*, 1–9. https://doi.org/10.1155/2021/9969464

Xiao, Y., Jiang, A., Ye, J., & Wang, M. (2020). Making of Night Vision: Object Detection Under Low-Illumination. *IEEE Access*, *8*, 123075–123086. https://doi.org/10.1109/access.2020.3007610

Yan, C. (2025). TinyML -Enhanced Cloud-Edge Collaborative Framework for Real-Time Sport Action Recognition. *Internet Technology Letters*, *8*(5). https://doi.org/10.1002/itl2.70100

Yoon, K., Lee, S., Park, J., & Kim, K. G. (2025). CNN-Based Automatic Tablet Classification Using a Vibration-Controlled Bowl Feeder With Spiral Torque Optimization. *Sensors*, *25*(14), 4248. https://doi.org/10.3390/s25144248

*name of corresponding author



Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16271 p-ISSN : 2541-044X

Yuadi, I., Nisa', K., Nazikhah, N. U., Halim, Y. A., Asyhari, A. T., & Hu, C.-C. (2025). A Benchmark Study of Classical and U-Net ResNet34 Methods for Binarization of Balinese Palm Leaf Manuscripts. *Heritage*, *8*(8). https://doi.org/10.3390/heritage8080337

Zhang, A., Li, Y., & Wang, S. (2024). 2ddsru-MobileNet: An End-to-End Cloud-Noise-Robust Lightweight Convolution Neural Network. *Journal of Applied Remote Sensing*, *18*(02). https://doi.org/10.1117/1.jrs.18.024511

Zhang, E., Guo, L., Guo, J., Yan, S., Li, X., & Kong, L. (2023). A Low-Brightness Image Enhancement Algorithm Based on Multi-Scale Fusion. *Applied Sciences*, *13*(18), 10230. https://doi.org/10.3390/app131810230

Zhang, J., Yang, X., Fu, X., Wang, B., & Li, H. (2025). LDL-MobileNetV3S: An Enhanced Lightweight MobileNetV3-small Model for Potato Leaf Disease Diagnosis Through Multi-Module Fusion. *Frontiers in Plant Science*, *16*. https://doi.org/10.3389/fpls.2025.1656731

Zhang, M., Zhang, Y., Jiang, Z., Lv, X., & Guo, C. (2021). Low-Illumination Image Enhancement in the Space Environment Based on the DC-WGAN Algorithm. *Sensors*, *21*(1), 286. https://doi.org/10.3390/s21010286

*name of corresponding author