Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

Performance Analysis of an Offline Text Detection System Based on Edge AIA Case Study of DokuScan Pro

Defi Pujianto[1], Kadarsih[2], Sri Hartati[3]
[1,2,3] Universitas Mahakarya Asia
[1] dhelphie85@gmail.com, [2]kadarsih.mail@gmail.com, [3]hartatiakmi1984@gmail.com

ABSTRACT

The growing use of mobile document scanning applications has increased the demand for text detection systems that can operate reliably in offline and on-device environments. Although Edge AI enables local inference without network dependency, system-level empirical evidence regarding its performance under real-world mobile usage conditions remains limited.

This study presents a system-level evaluation of an offline Edge AI–based text detection system for mobile document scanning, using DokuScan Pro as a case study. The evaluation was conducted on 40 document images captured under varying lighting conditions, capture angles, and background characteristics. System performance was assessed using precision, recall, F1-score, and inference time to characterize on-device behavior rather than algorithmic novelty.

Experimental results show that the system achieved a precision of 1.00, a recall of 0.975, and an F1-score of approximately 0.98, with an average inference time of 63.8 ms per image during fully offline execution on mobile devices. These results indicate stable system-level performance under real-world document scanning conditions with controlled computational overhead.

This study provides empirical system-level insights into the feasibility and practical limitations of deploying Edge AI–based text detection in offline mobile document scanning applications, thereby complementing existing model-centric research with evidence from real-world, on-device evaluation.

Keywords: document scanning, Edge AI, mobile processing, text detection, offline inference

INTRODUCTION

The rapid adoption of mobile document scanning applications has transformed document digitization into an essential component of administrative, educational, and public service activities. Recent advances in computer vision and deep learning have enabled text detection systems to achieve high accuracy under controlled evaluation settings, typically relying on cloud-based computation or server-side processing. However, such approaches introduce limitations related to network dependency, latency, and data privacy, which restrict their applicability in offline and resource-constrained environments.

Edge Artificial Intelligence (Edge AI) has emerged as a promising paradigm that enables on-device inference by executing computer vision models directly on mobile devices. In the context of document scanning, on-device text detection offers practical advantages by eliminating network dependency and preserving user data locally. Nevertheless, deploying text detection systems at the edge introduces scientific risks that extend beyond technical implementation challenges. Limited computational resources, hardware variability, and environmental conditions may fundamentally affect system reliability, performance stability, and result reproducibility when compared to conventional server-based evaluation settings.

From a theoretical perspective, the shift from cloud-based inference to on-device Edge AI deployment alters the evaluation paradigm of computer vision systems. Traditional text detection research predominantly emphasizes model-centric evaluation, focusing on accuracy metrics obtained from standardized datasets under controlled conditions. In contrast, on-device deployment requires system-level evaluation that accounts for interactions between models, preprocessing pipelines, hardware constraints, and real-world usage conditions. This shift raises important questions regarding whether conventional evaluation practices remain sufficient to characterize the true behavior of text detection systems operating entirely at the edge.

Despite the growing interest in Edge AI–based document scanning, existing studies largely focus on optimizing individual models or improving algorithmic efficiency. There is still limited empirical evidence regarding whether

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

system-level deployment of Edge AI fundamentally alters the reliability, stability, and evaluation paradigm of document text detection, particularly in fully offline mobile environments. Moreover, a structured methodological framework for evaluating text detection systems at the system level—incorporating performance behavior, inference stability, and resource-related constraints—remains underexplored in the literature.
To address this gap, this study presents a system-level evaluation of an offline Edge AI–based text detection system for mobile document scanning, using DokuScan Pro as a case study. Rather than proposing new detection algorithms, the study aims to empirically characterize system behavior under real-world on-device conditions by analyzing detection performance, inference time characteristics, and resource-related performance tendencies. By focusing on system-level behavior, this work seeks to contribute empirical insights that support a more comprehensive and scientifically grounded evaluation of Edge AI–based text detection systems in offline mobile applications.

The scientific contributions of this study are articulated as follows:

1. This study formulates a system-level evaluation perspective for offline Edge AI–based document text detection, emphasizing the assessment of integrated system behavior rather than isolated model performance under controlled conditions.
2. Empirical operational evidence is provided to characterize the stability, latency behavior, and operational feasibility of fully offline, on-device text detection systems, based on real-world mobile document scanning scenarios.
3. The study offers deployment-oriented insights into the practical implications of Edge AI adoption for mobile document scanning, highlighting how on-device constraints and environmental variability influence system-level performance characteristics.
4. A reproducibility-oriented evaluation approach is demonstrated through the use of standardized metrics, documented testing scenarios, and repeatable experimental procedures, supporting transparent and verifiable system-level performance assessment for future research.

LITERATURE REVIEW

Research on document text detection has evolved significantly in response to the increasing demand for efficient and autonomous document digitization. Early studies predominantly employed rule-based computer vision approaches, such as stroke width transform and connected component analysis, which offered computational efficiency but exhibited limited robustness under real-world conditions involving illumination variation, perspective distortion, and complex backgrounds (Epshtein et al., 2010; Francis & Sreenath, 2020). These approaches, while foundational, are inherently sensitive to environmental variability, limiting their applicability in mobile document scanning scenarios.
The advancement of deep learning–based text detection models has substantially improved detection accuracy and adaptability to diverse visual conditions. Models such as EAST, CRAFT, and DBNet demonstrate enhanced robustness by learning discriminative features directly from data, enabling more reliable text localization in complex document images (Zhou et al., 2017; Baek et al., 2019; Liao et al., 2020). However, the majority of existing studies remain model-centric, emphasizing algorithmic performance under controlled datasets and focusing primarily on accuracy metrics. Consequently, these evaluations often overlook how such models behave when deployed as components within complete document scanning systems.
Beyond text detection, several studies have explored integrated document analysis pipelines that address layout understanding and text recognition as complementary stages in document processing workflows, highlighting the broader context in which text detection systems operate (Rajan & Devasena, 2025).
In parallel, the computational paradigm for document processing has shifted from cloud-based architectures to on-device Edge AI deployment. Cloud-based approaches provide scalable computational resources but introduce latency, network dependency, and privacy concerns, particularly for mobile document scanning applications (Gunawardena et al., 2022). On-device Edge AI mitigates these limitations by enabling local inference, but it also introduces new constraints related to limited computational resources, hardware heterogeneity, and energy consumption. Recent studies have demonstrated that Edge AI and edge computing can support robust computer vision applications under real-world operational conditions, while emphasizing the importance of system-level deployment considerations and computational efficiency (Gunawardena et al., 2022; MohiEldeen Alabbasy et al., 2023; Peng et al., 2024; N et al., 2025; Fang & Yin, 2025). These constraints may fundamentally affect system reliability, performance stability, and result reproducibility in ways that are not captured by conventional model-level evaluation.

* Corresponding author

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

Despite increasing attention to Edge AI optimization and lightweight model design, existing research largely evaluates individual models rather than complete operational systems. Performance assessments are typically conducted in isolation, without considering interactions among preprocessing pipelines, inference execution, hardware constraints, and real-world usage conditions. As a result, the current literature provides limited guidance on how Edge AI–based text detection systems perform as integrated systems in offline mobile environments.

From a methodological perspective, this gap reflects a broader limitation in evaluation practices for Edge AI–based computer vision systems, where traditional model-centric metrics are insufficient to characterize system-level behavior. There remains a lack of system-level evaluation frameworks that account for operational performance characteristics such as inference stability, latency variability, and resource-related performance tendencies during real-world deployment.

Thus, existing studies insufficiently address how Edge AI behaves as a complete operational system, particularly in the context of fully offline mobile document scanning. This limitation motivates the need for system-level empirical evaluation that bridges the gap between algorithmic performance and real-world deployment behavior.

METHOD

3.1. Research Design and Evaluation Scope

This study adopts a quantitative experimental approach aimed at conducting a system-level performance evaluation of an offline Edge AI–based text detection system deployed on mobile devices. The evaluation is designed as an exploratory system-level study, focusing on empirical characterization of system behavior under real-world, on-device execution constraints rather than statistical generalization or algorithmic benchmarking.

Given this scope, the evaluation emphasizes operational performance characteristics, including detection reliability, inference time behavior, and resource-related performance tendencies, as recommended in prior studies on Edge AI deployment and mobile computer vision systems.

3.2. Dataset Description and Rationale

The evaluation dataset consists of 40 document images, collected to represent real-world mobile document scanning scenarios. The dataset size is intentionally limited to enable controlled and repeatable system-level analysis under constrained experimental conditions, which is consistent with exploratory evaluations of on-device AI systems.

The dataset includes variations in:

- lighting conditions (uniform, low-light, uneven illumination),
- capture angles (frontal and mild perspective distortion),
- background complexity,
- document layout characteristics.

Rather than aiming for statistical representativeness, the dataset is designed to probe system behavior across common operational variations encountered in mobile document scanning. This approach aligns with system-level evaluation practices where the objective is to observe performance stability and operational feasibility under realistic usage conditions. Thus, existing studies remain largely insufficient in explaining how Edge AI–based text detection behaves as a complete operational system when deployed under real-world, on-device constraints.

3.3. Evaluation Pipeline and Reproducibility

The evaluation pipeline follows a fixed and reproducible sequence consisting of:

1. Image acquisition via mobile device camera,
2. Preprocessing (perspective correction, noise reduction, contrast enhancement),
3. On-device Edge AI inference for text detection,
4. Output evaluation using predefined metrics.

All experiments were conducted using the same pipeline configuration to ensure reproducibility. The evaluation procedure, parameter settings, and execution order were kept constant across all test samples.

3.4. Device Configuration and Testing Environment

System evaluation was conducted on a single, representative Android-based mobile device to reflect realistic deployment conditions for offline document scanning applications. The primary testing device was equipped with a mid-range mobile System-on-Chip (SoC) featuring an octa-core CPU architecture with clock speeds of up to approximately 2.4 GHz, 6 GB of RAM, and sufficient internal storage to ensure stable application

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

execution. The device operated on Android OS version 13.

The evaluated text detection system was executed entirely on-device using CPU-based inference, without hardware acceleration from dedicated Neural Processing Units (NPUs) or mobile GPUs. This configuration was intentionally selected to represent conservative and widely accessible deployment conditions commonly encountered in offline mobile document scanning applications.

During testing, the device was maintained in a normal thermal state, with battery level above 50%, and no background applications actively consuming significant computational resources. All experiments were conducted in foreground execution mode to minimize external interference and ensure consistency of inference time and resource usage measurements. This controlled configuration was applied consistently across all testing sessions to reduce measurement variability and support reproducibility of the reported results.

This device configuration and testing environment are reported to enhance transparency and reproducibility of the system-level evaluation. It is acknowledged that inference time and resource utilization characteristics may vary across different mobile hardware configurations and device specifications, and thus the reported results should be interpreted within the context of the evaluated device.

3.5. Baseline Definition

Instead of comparing against external models or cloud-based systems, this study adopts an internal system baseline approach. The baseline is defined as the default on-device inference configuration of the evaluated system, serving as a reference point for observing performance behavior across different document scanning scenarios. This baseline selection is consistent with the study's focus on system-level operational evaluation rather than model-level performance comparison.

3.6. Performance Metrics

An Intersection over Union (IoU) threshold of 0.3 was adopted to determine detection correctness. This threshold was selected to reflect the system-level objective of the study, which focuses on document-level text region localization rather than fine-grained bounding box accuracy. In mobile document scanning scenarios, moderate spatial misalignment between detected regions and ground truth annotations is common due to variations in capture angle, perspective distortion, and preprocessing transformations. Therefore, a lower IoU threshold is considered appropriate to evaluate whether the system successfully identifies the intended document text region under real-world on-device conditions. Similar threshold relaxation has been employed in system-oriented and deployment-focused evaluations where operational feasibility and detection reliability are prioritized over strict geometric alignment. It is important to note that the selected IoU threshold does not aim to demonstrate algorithmic superiority but rather to characterize system behavior within practical usage constraints.

Future evaluations may incorporate higher IoU thresholds or multi-threshold analysis to further examine spatial precision under more controlled conditions.

System performance was evaluated using the following metrics:

- Detection performance, measured using precision, recall, and F1-score based on manually annotated ground truth.
- Inference time, measured as the elapsed time between inference start and completion on the mobile device.
- CPU and memory usage, observed during inference execution using system-level monitoring tools provided by the mobile operating system.
- Energy consumption tendency, assessed through monitored battery usage and execution duration during controlled testing sessions.

These metrics were selected to capture both accuracy-related outcomes and operational performance characteristics, reflecting real-world on-device usage constraints.

3.7. Measurement Procedure

Inference time was measured programmatically by recording timestamps before and after the inference process. CPU and memory utilization were monitored using device profiling tools available on the Android platform. Energy-related behavior was assessed by observing battery level changes and execution duration under controlled conditions, providing indicative insights into energy consumption tendencies rather than precise power measurements.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

3.8. Validation Strategy

To enhance result reliability, each evaluation scenario was executed multiple times using the same dataset and configuration. Consistency of detection results and inference time behavior across repeated runs was examined to assess system stability.

This validation approach supports reproducibility-oriented evaluation, ensuring that the observed performance characteristics can be verified under comparable testing conditions.

RESULT

This section presents the results of the performance evaluation of the offline Edge AI–based text detection system implemented in DokuScan Pro. All results are obtained from direct testing on mobile devices in accordance with the scenarios and metrics described in the methodology section. The presentation of results focuses on empirical findings without making claims regarding methodological development or algorithmic superiority.

4.1. Image Preprocessing Output

The preprocessing results indicate an improvement in the visual quality of document images across most test samples. Perspective correction, noise reduction, and contrast enhancement stages produce images with more uniform visual structures, particularly along document edges and text regions.

1. Visually, the preprocessing results demonstrate:
2. Improved text readability;
3. More consistent contrast across document areas;
4. Reduced noise interference in darker regions.

As illustrated in Figure 4.1, the preprocessed images exhibit more stable visual characteristics compared to the original images. This finding is consistent with the observations reported by Lins (2021), which indicate that image quality enhancement contributes to the stability of text detection processes in mobile document scanning systems.

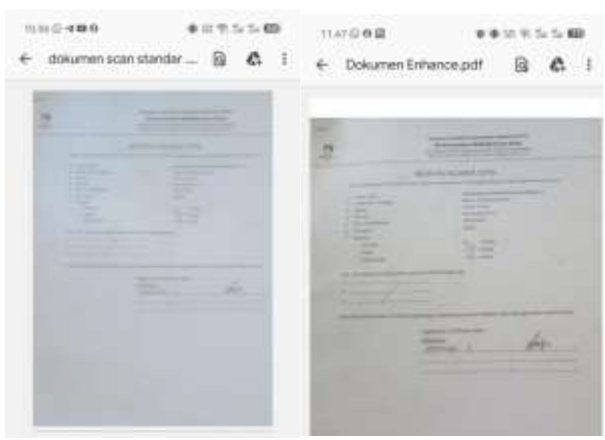

Figure 4.1. Image Preprocessing Output

4.2. Text Detection Evaluation Results

The text detection performance evaluation was conducted on 40 document images captured under varying lighting conditions, capture angles, and background characteristics. Each test image contained a single primary document text region. Ground truth annotations were generated manually, and detection results were evaluated using an Intersection over Union (IoU) threshold of 0.3.

Based on the evaluation results, the system correctly detected text regions in 39 images, with one false negative and no false positive cases. This corresponds to a precision value of 1.00, a recall value of 0.975, and an F1-score of approximately 0.987, as summarized in Table 4.1.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

These results indicate that the evaluated system demonstrates reliable detection performance under the tested offline, on-device conditions. The absence of false positives suggests conservative detection behavior, while the single false negative highlights the presence of occasional detection failure under challenging capture conditions.

Table 4.1. System-Level Text Detection Performance Metrics

Metric	Value
True Positive (TP)	39
False Positive (FP)	0
False Negative (FN)	1
Precision	1.00
Recall	0.975
F1-score	0.987

The distribution of IoU values shown in Figures 4.3 and 4.4 indicates that the majority of detection results exhibit a high level of agreement with the ground truth. The relatively narrow distribution suggests consistency in detection performance across different visual conditions.

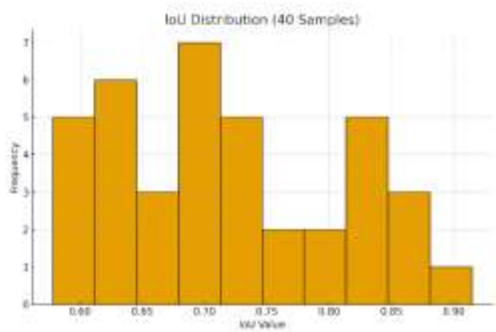

Figure 4.2. Distribution of IoU Values

Figure 4.2. Distribution of Intersection over Union (IoU) values obtained from 40 document images. The distribution shows that most IoU values are concentrated within a moderate to high range, indicating consistent agreement between detected text regions and the corresponding ground truth annotations across the evaluated samples.

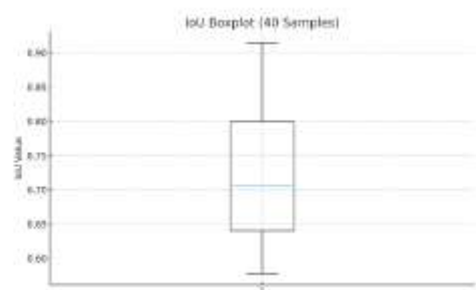

Figure 4.3. Boxplot of IoU Values

Figure 4.3. Boxplot of Intersection over Union (IoU) values for 40 document images. The boxplot illustrates the central tendency and variability of IoU values, with a relatively narrow interquartile range, suggesting stable detection performance across different document capture conditions.

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

4.3. Scenario-Based Evaluation: Inference Time

Inference time measurement was conducted to assess the responsiveness of the text detection system when executed fully on-device. Inference time was recorded on the Android side by measuring the time difference before and after the inference process was executed.

4.3.1. Inference Time Results

The test results show that inference time ranged from 55 to 78 ms, with an average of 63.8 ms across 40 test images. The variation in inference time is relatively small and is associated with differences in image visual characteristics, such as background complexity and perspective angles during image capture.

4.3.2. Statistical Analysis

The statistical summary presented in Table 4.3 shows a standard deviation of 5.87 ms, indicating inference time stability across samples. The inference time distribution shown in Figure 4.5 demonstrates a concentration of values within the 58–68 ms range, with no extreme spikes observed.

Table 4.2. Standard Deviation Values

Parameter	Value
Number of Samples	40 images
Mean	63.8 ms
Minimum	55 ms
Maximum	78 ms
Standard Deviation	5.87 ms

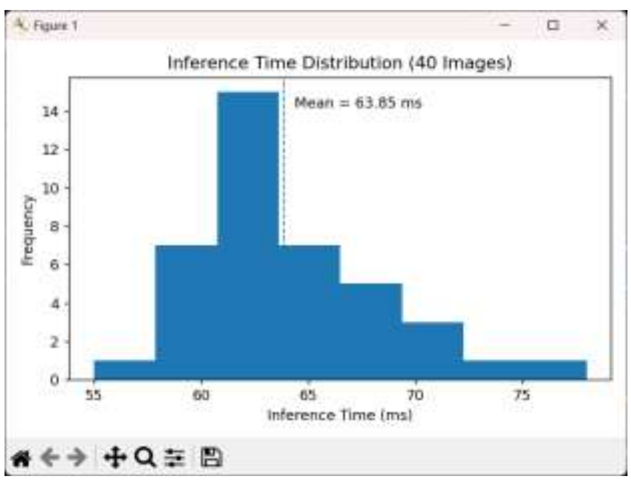

Figure 4.4. Inference Time Distribution

Figure 4.4. Distribution of on-device inference time for 40 document images. The histogram shows the variation of inference time during fully offline execution, with an average inference time of approximately 63.8 ms, indicating consistent execution behavior under the evaluated system configuration.

4.4. Resource Consumption Evaluation

Resource-related performance evaluation was conducted to observe the operational behavior of the evaluated system during fully on-device inference. Rather than performing precise power or utilization measurements, this evaluation focuses on execution characteristics and stability indicators that reflect the system's feasibility for offline mobile document scanning.

The system processed the complete test dataset of 40 document images with an average inference time of 63.8 ms per image, resulting in a total execution duration of approximately 2.54 seconds under foreground execution

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

mode. During inference execution, no abnormal latency spikes or execution interruptions were observed across repeated runs.

Qualitative observation using platform-level profiling tools indicated stable execution behavior during inference, with no noticeable performance degradation or abrupt increases in execution latency. These observations suggest that the evaluated system operates within practical computational bounds for on-device deployment, without imposing excessive operational overhead under the tested conditions.

It is important to note that the reported observations represent indicative resource-related behavior rather than precise quantitative measurements of CPU utilization or energy consumption. Consequently, the findings are intended to characterize operational feasibility and execution stability, rather than to provide absolute resource consumption metrics

Table 4.3. Summary of System Resource Consumption Based on Inference Time

Parameter	Value
Number of test samples	40 document images
Average inference time per image	63.8 ms
Minimum inference time	55 ms
Maximum inference time	78 ms
Total inference time	± 2.54 seconds
Execution mode	On-device (Edge AI)
Application state	Foreground
CPU activity during inference	Stable
Energy consumption indication	Low–moderate

The observation results indicate that during the inference process, the system does not experience significant fluctuations in computational load. The relatively short and stable inference time suggests that the system can operate without imposing excessive pressure on device resources.

The summary of resource consumption evaluation is presented in Table 4.4, showing that the system is able to process the entire test dataset within a total time of approximately 2.54 seconds while maintaining controlled resource utilization characteristics.

4.5. Validation and Reproducibility

Testing was conducted repeatedly using the same datasets and scenarios to evaluate result consistency. The detection metrics and inference time values exhibit stable patterns across multiple runs, both in terms of detection performance and processing efficiency. The evaluation protocol in this study follows reproducibility-oriented practices commonly recommended in machine learning research, including repeated testing under consistent conditions and clearly documented processing pipelines (Pineau et al., 2021). This approach supports the reproducibility of the results, given that all experiments were performed using a documented processing pipeline and evaluation metrics commonly adopted in Edge AI research on mobile devices.

4.6. Summary of Key Findings

1. Based on the evaluation results, the following key findings can be summarized:
2. The text detection system demonstrates a high level of accuracy on the test data, with limited detection errors.
3. Inference time remains within a relatively stable and consistent range across test samples.
4. The short inference time characteristics correlate with controlled device resource utilization.
5. The evaluation results are consistent and reproducible under comparable testing scenarios.
6. This summary provides the basis for further discussion in the Discussion section, particularly regarding result implications and system limitations.

DISCUSSIONS

The results of this study provide insights into the behavior of an offline Edge AI–based document scanning system when deployed entirely on mobile devices. Rather than emphasizing isolated detection accuracy, the findings highlight

* Corresponding author

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

how system-level performance characteristics emerge from the interaction between on-device inference, preprocessing pipelines, and mobile hardware constraints under real-world usage conditions.

Theoretical Implications

From a theoretical perspective, the results support the notion that Edge AI deployment fundamentally shifts the evaluation paradigm of document text detection systems. Traditional computer vision research primarily adopts a model-centric evaluation approach, focusing on accuracy metrics obtained from standardized datasets. In contrast, the observed system behavior in this study indicates that operational stability, inference latency, and execution consistency are equally critical dimensions for understanding real-world performance. This suggests that evaluation frameworks for Edge AI–based document scanning systems should extend beyond model accuracy to incorporate system-level behavioral characteristics.

Methodological Implications

Methodologically, this study demonstrates the importance of system-level evaluation for on-device AI applications. The use of a fixed evaluation pipeline, repeated testing scenarios, and consistent execution conditions provides a reproducibility-oriented approach that complements conventional model-level benchmarking. The findings indicate that small-scale, exploratory datasets can still yield meaningful insights into system behavior when the evaluation objective is explicitly defined as operational characterization rather than statistical generalization. This has implications for future studies seeking to evaluate Edge AI systems under practical deployment constraints.

System Limitations

Despite the reported performance stability, several limitations of the evaluated system must be acknowledged. First, device dependency represents a significant constraint, as the observed performance characteristics are influenced by the specific hardware configurations and operating conditions of the tested mobile devices. Second, the limited dataset size restricts the ability to generalize the findings across broader document types and capture environments. Third, the evaluation focuses on document-level detection performance, which may not fully reflect the behavior of text detection at finer granularities, such as individual text regions or characters.

Threats To Validity

Several threats to validity should be considered when interpreting the results. Dataset bias may arise from the limited diversity of document samples and capture conditions, potentially affecting the observed detection performance. Model generalization risk is also present, as the evaluated system employs a specific model configuration optimized for on-device execution, which may behave differently when applied to unseen document characteristics. Additionally, evaluation bias may be introduced through manual annotation and the selection of evaluation thresholds, which can influence reported performance metrics.

Future Research Directions

Future research should address these limitations by expanding evaluation datasets to include a wider range of document types, capture devices, and environmental conditions. Incorporating quantitative power measurement tools would enable more precise analysis of energy consumption behavior. Comparative system-level evaluations across different device classes and Edge AI configurations would further enhance understanding of deployment trade-offs. Additionally, extending evaluation frameworks to integrate downstream tasks such as optical character recognition (OCR) could provide a more comprehensive assessment of end-to-end document scanning performance.

These findings are specific to the evaluated system configuration, dataset characteristics, and testing conditions, and should not be interpreted as generalized performance guarantees across all mobile devices or document types.

CONCLUSION

This study presents a system-level evaluation of an offline Edge AI–based document scanning system, emphasizing on-device performance behavior rather than isolated model accuracy. The findings demonstrate that evaluating document text detection as an integrated operational system reveals performance characteristics that are not fully captured by conventional model-centric evaluation approaches.

From a scientific perspective, this work highlights that Edge AI deployment alters the evaluation paradigm of document text detection systems, shifting the focus from algorithmic performance under controlled conditions to system-level behavior under real-world constraints. This perspective contributes to a more comprehensive understanding of how on-device execution, hardware limitations, and environmental variability collectively influence detection reliability and operational stability.

The results matter because they provide empirical evidence that system-level performance characteristics—such as inference stability and execution feasibility—are critical for assessing the practical viability of offline document

* Corresponding author

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

scanning systems. By characterizing these aspects, the study bridges the gap between model-centric research and real-world deployment considerations, supporting more informed evaluation practices in Edge AI research.
For future research, this study suggests the need for broader and more diverse system-level evaluations, including larger datasets, multi-device benchmarking, and precise energy measurement methodologies. Extending the evaluation framework to incorporate end-to-end document processing tasks, such as optical character recognition and post-processing quality assessment, would further strengthen understanding of offline document scanning system behavior. Researchers are encouraged to adopt system-level, reproducibility-oriented evaluation strategies when studying Edge AI applications deployed in resource-constrained environments.

ACKNOWLEDGMENT

The authors would like to express their appreciation to all parties who provided support during the conduct of this research. Gratitude is extended to Universitas Mahakarya Asia for providing facilities and an academic environment that supported the system evaluation process.
The authors also acknowledge the contributions of colleagues who offered valuable input and engaged in scholarly discussions throughout the research and manuscript preparation. In addition, technical support in the form of access to mobile devices used for system testing contributed to the successful completion of this study.

REFERENCES

Baek, Y., Lee, B., Han, D., Yun, S., & Lee, H. (2019). *Character Region Awareness for Text Detection*. http://arxiv.org/abs/1904.01941

Epshtein, B., Ofek, E., & Wexler, Y. (2010). Detecting text in natural scenes with stroke width transform. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2963–2970. https://doi.org/10.1109/CVPR.2010.5540041

Fang, Z., & Yin, B. (2025). A low functional redundancy-based network slimming method for accelerating deep neural networks. *Alexandria Engineering Journal*, *119*, 437–450. https://doi.org/10.1016/j.aej.2024.12.118

Francis, L. M., & Sreenath, N. (2020). TEDLESS – Text detection using least-square SVM from natural scene. *Journal of King Saud University - Computer and Information Sciences*, *32*(3), 287–299. https://doi.org/10.1016/j.jksuci.2017.09.001

Gunawardena, N., Ginige, J. A., Javadi, B., & Lui, G. (2022). Performance Analysis of CNN Models for Mobile Device Eye Tracking with Edge Computing. *Procedia Computer Science*, *207*, 2291–2300. https://doi.org/10.1016/j.procs.2022.09.288

Liao, M., Wan, Z., Yao, C., Chen, K., & Bai, X. (2020). Real-Time Scene Text Detection with Differentiable Binarization. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(07), 11474–11481. https://doi.org/10.1609/aaai.v34i07.6812

Lins, R. D. , C. G. D. C. , & S. J. D. (2021). Image enhancement techniques for mobile document scanning systems. *Pattern Recognition Letters*.

Mohi Eldeen Alabbasy, F., Abohamama, A. S., & Alrahmawy, M. F. (2023). Compressing medical deep neural network models for edge devices using knowledge distillation. *Journal of King Saud University - Computer and Information Sciences*, *35*(7). https://doi.org/10.1016/j.jksuci.2023.101616

N, K., Choudhury, A. R., G S, R., G, B., & Sanchi, C. (2025). On-Device Deep Learning for Retrieving System and User Timestamps from Noisy Chat Images. *Procedia Computer Science*, *258*, 3760–3770. https://doi.org/10.1016/j.procs.2025.04.631

Peng, Z., Li, J., Hao, H., & Zhong, Y. (2024). Smart structural health monitoring using computer vision and edge computing. *Engineering Structures*, *319*. https://doi.org/10.1016/j.engstruct.2024.118809

Pineau, J., Vincent-Lamarre, P., Sinha, K., Larivì, V., Beygelzimer, A., D'alché-Buc, F., Paris, T., Larochelle, H., Florence D'alché-Buc, E., & Fox, H. L. (2021). Improving Reproducibility in Machine Learning Research (A Report from the NeurIPS 2019 Reproducibility Program). In *Journal of Machine Learning Research* (Vol. 22).

Rajan, R., & Devasena, M. S. G. (2025). Deep learning based optimization model for document layout and text recognition. *Ain Shams Engineering Journal*, *16*(10), 103587. https://doi.org/10.1016/j.asej.2025.103587

Wang, P., Ma, Z., Dong, B., Liu, X., Ding, J., Sun, K., & Chen, Y. (2024). Generative data augmentation by conditional inpainting for multi-class object detection in infrared images. *Pattern Recognition*, *153*. https://doi.org/10.1016/j.patcog.2024.110501

Yıldız, S. (2024). Turkish scene text recognition: Introducing extensive real and synthetic datasets and a novel

* Corresponding author

Journal of Computer Networks, Architecture and High Performance Computing
Volume 08, Number 1, Januari 2026
https://doi.org/10.47709/cnapc.v8i1.7854

Submitted: 2026-01-19
Accepted : 2026-02-01
Published : 2026-02-01

recognition model. *Engineering Science and Technology, an International Journal*, *60*, 101881. https://doi.org/10.1016/j.jestch.2024.101881

Zhou, X. , Y. C. , W. H. , W. Y. , Z. S. , H. W. , & L. J. (2017). EAST: An efficient and accurate scene text detector. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

* Corresponding author