TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 5, October 2026, pp. 1668~1679
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i5.27607

Improving PaddleOCR optical character recognition for food composition labels using hybrid SymSpell and LayoutLM correction

Ahmad Wahyu Rosyadi[1], Siti Ma'shumah[2], Muhammad Qomaruz Zaman[3]
[1]Department of Informatics Engineering, Universitas Qomaruddin, Gresik, Indonesia
[2]Department of Electrical Engineering, Universitas Qomaruddin, Gresik, Indonesia
[3]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Article Info

Article history:

Received Oct 16, 2025
Revised Aug 19, 2026
Accepted Aug 28, 2026

Keywords:

Food identification
LayoutLM
Optical character recognition
PaddleOCR
SymSpell

ABSTRACT

A high level of text recognition accuracy is crucial for optical character recognition (OCR) applications, particularly those designed to extract information from food composition labels. However, packaging materials often introduce distortions, reflections, or irregular printing, which lead to recognition errors and word concatenation. To mitigate this issue, this study proposes a novel post-processing system that integrates PaddleOCR with the spelling correction (SymSpell) algorithm and layout language model (LayoutLM)-guided bigram-based word segmentation to enhance OCR accuracy. The implementation begins with capturing a composition label image using a smartphone camera, followed by initial text extraction performed by PaddleOCR. Consequently, SymSpell corrects misspelled tokens through a domain-specific food and beverage lexicon, while LayoutLM leverages spatial and contextual information to separate concatenated words. The performance was evaluated using the character error rate (CER) metric across four packaging surface categories: flat, curved, reflective, and textured. The results demonstrate consistent improvements across all categories, achieving an average CER of 0.0991, with the greatest reduction on flat surfaces (from 0.0885 to 0.0659). The experiments show that the system is robust across label conditions, lightweight for real-time use, and domain-aware. Its model-agnostic design adds flexibility, enabling effective post-OCR correction across different engines and improving readability for real-world food label applications.

Corresponding Author:

Ahmad Wahyu Rosyadi
Department of Informatics Engineering, Universitas Qomaruddin
Raya Bungah St. No. 1, Rice Field / Garden Area, Bungah, Gresik, East Java 61152, Indonesia
Email: awrosyadi@gmail.com

1. INTRODUCTION

Information labels on food packaging play a crucial role in providing consumers with transparency regarding product ingredients, helping them make healthier dietary choices [1]. However, consumers may encounter unfamiliar ingredients, prompting them to seek additional information online, which is often inefficient and unreliable. Recent technological advances have enabled intelligent systems to support more effective access and interpretation of information, with artificial intelligence (AI) innovations greatly improving digital accuracy, efficiency, and accessibility, particularly in healthcare [2]-[5]. These developments highlight AI's broader potential to enhance decision-making and expand access to essential information across domains such as food and nutrition [6]-[9]. In line with these developments, several

studies have used smartphone cameras to scan and display product information with minimal user interaction, including Mavani *et al*. [7] who introduced a fuzzy logic–based barcode system that decodes barcodes and retrieves ingredient data from online sources to determine the halal status of packaged products. In reality, many products lack barcodes, thereby reducing their usability. To overcome this issue, Rosyadi *et al.* [10]. introduced a method for scanning composition labels using PaddleOCR, as most products display such labels Although PaddleOCR performs well on flat packaging, its accuracy decreases on curved, textured, and reflective surfaces where distortions and uneven lighting introduce optical character recognition (OCR)-specific challenges, leading to character errors, lost diacritics, and frequent word-splitting or concatenation issues that degrade readability and hinder downstream information retrieval [11]. These limitations demonstrate the need for a post-processing method that is able to correct lexical errors and restore proper word structure in OCR outputs.

To mitigate these issues, lightweight lexical correction methods like SymSpell remain highly effective in specialized domains, while transformer-based models such as layout language model (LayoutLM) are capable of capturing spatial relationships and contextual cues that conventional OCR pipelines overlook. The combination of these complementary strengths forms the foundation of this study's contribution, enabling more robust correction at both the lexical and structural levels in post-OCR processing.

In recent years, OCR post-processing has improved significantly [12], [13]. One algorithm that can be used is symmetric delete spelling correction (SymSpell) [14], [15]. This algorithm can compete with modern language models in some cases, making it a simpler and more affordable solution [16]. Several studies have used SymSpell for OCR post-processing, including the work by Bjerring-Hansen *et al*. [17], the study integrated SymSpell-based spell checking with context-aware selective modifications and safe error corrections, improving word accuracy by up to 73%. Furthermore, Aung *et al.* [18] introduced myOCR, a Myanmar OCR framework that incorporated SymSpell in its post-correction stage, achieving a 0.66% reduction in word error rate. Another comparative study evaluated Hunspell, SymSpell, the Norvig model, and the N-gram model on Azerbaijani text, showing that SymSpell achieved the highest processing speed with 81.4% accuracy [19]. In addition, the integration of SymSpell as a post-processing stage for Tesseract OCR results on passport data was able to increase accuracy by 21.4%, especially in the visual inspection zone with high distortion and noise [20]. The addition of post-processing to the OCR results is expected to reduce spelling errors and thus improve the quality of the extracted text, especially in the domain of food packaging labels that often contain complex visual elements such as decorative fonts, illustrated backgrounds, and curved surfaces. Therefore, a system that is able to validate and repair OCR tokens efficiently is needed to ensure that the extracted information remains accurate and can be interpreted correctly.

In addition to classical OCR post-processing, transformer-based models such as LayoutLM have recently emerged as powerful tools for understanding visually rich documents [21], [22]. By jointly learning textual and spatial features, LayoutLM preserves structural relationships that conventional OCR often loses [23]. This architecture has demonstrated high performance in various applications, including receipt parsing, form understanding, and document classification [24], [25]. Furthermore, recent research has extended the LayoutLM framework towards OCR text correction, such as the study by Xu *et al.* [26] that introduced LayoutLM-Critic. This study leverages multimodal information (text, bounding boxes, and image features) to detect and correct OCR errors more effectively than purely text-based methods. These advances underscore the potential of LayoutLM to act as a complementary component in food label analysis, producing cleaner and semantically coherent token sequences that improve subsequent OCR post-processing.

This research presents a post-OCR correction method that integrates the dictionary-based SymSpell algorithm with the LayoutLM model to address spelling errors and word segmentation issues in OCR outputs. SymSpell applies a domain-specific food and beverage dictionary to ensure lexical accuracy, while LayoutLM produces semantically consistent token sequences that are exploited as bigram references to detect and separate compound words in PaddleOCR results. This hybrid approach reduces misspellings, improves word boundary recovery, and produces more accurate and interpretable text aligned with the intended meaning of food composition labels. In summary, this study aims to build a hybrid post-OCR correction framework that integrates SymSpell's lexical correction with LayoutLM's layout-aware segmentation to enhance text recognition accuracy on food composition labels.

The structure of this paper consists of four main parts. The current part provides the introduction. The second part presents a comprehensive description of the proposed approach. The third part reports and analyzes the experimental results, highlighting system performance and effectiveness. And the final part provides the conclusion, summarizing the key conclusions and outlining potential directions for further study.

2. METHOD

The post-processing stage of OCR results consists of several steps to improve text accuracy as depicted in Figure 1. To illustrate the effect of the proposed methodology, Figure 2 presents an example of OCR results showing Figure 2(a) the input image, Figure 2(b) the baseline PaddleOCR output, and Figure 2(c) the post-OCR result. Each image was processed by PaddleOCR to extract text and by LayoutLM to generate context-aware bigrams. PaddleOCR outputs were normalized, tokenized, and corrected using SymSpell with a domain-specific dictionary, while LayoutLM outputs provided semantic cues for detecting and splitting concatenated words. The final text was assessed using the character error rate (CER) against manually transcribed ground-truth labels.

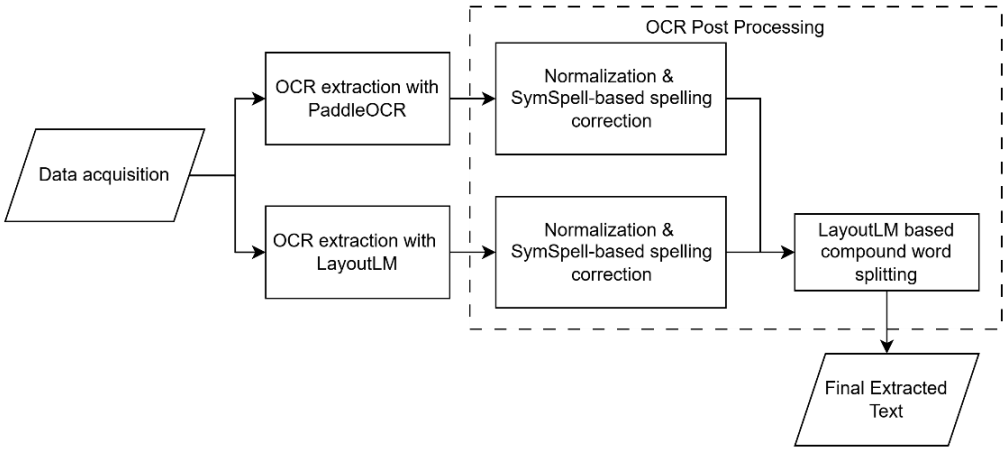

Figure 1. Proposed system design for post-OCR correction using SymSpell and LayoutLM-based compound-word segmentation

(a)

Komposisi: Santan Kelapa 90%).AirPenstabil Nabati, Omega 3 (mengandung Antioksidan Tokoferol). Tanpa pengawet dan pemanis buatan.

(b)

Komposisi: Santan Kelapa (90%), Air, Penstabil Nabati, Omega 3 (mengandung Antioksidan Tokoferol). Tanpa pengawet dan pemanis buatan.

(c)

Figure 2. Example of OCR: (a) input image, (b) baseline PaddleOCR output, and (c) post-OCR result

2.1. Data acquisition

The input images consist of smartphone-scanned and cropped composition labels. As illustrated in Figure 3, the dataset includes four packaging surface categories to represent diverse testing conditions: Figure 3(a) flat surfaces with minimal distortion, Figure 3(b) curved surfaces with text deformation, Figure 3(c) reflective surfaces affected by glare and contrast loss, and Figure 3(d) textured surfaces with background noise that interferes with text detection. Each image was manually verified and transcribed to create the ground truth reference for evaluation.

2.2. OCR extraction

PaddleOCR was selected as the primary OCR engine for three reasons. First, PaddleOCR achieved a substantially lower CER (0.1634) than LayoutLM (0.4531) across all surface categories as shown in Table 1. Second, PaddleOCR provides a lightweight, practical pipeline suitable for smartphone deployment, supporting multiple languages and flexible cross-platform deployment [27], [28]. Third, this study focuses on

post-OCR correction, requiring a reliable upstream character-level engine, while LayoutLM serves as a complementary model for spatial and contextual guidance. To verify the model-agnostic nature of our approach, the same pipeline was also evaluated on Tesseract, yielding a statistically significant though smaller CER reduction. The text output generated by PaddleOCR, after undergoing error correction during post-processing, is presented as the final system output.

In parallel, LayoutLM was applied to capture both textual and spatial relationships within the same image region. While PaddleOCR focuses on character recognition, LayoutLM encodes layout and visual cues, allowing it to generate semantically meaningful text representations that are later used to guide post-processing. The outputs from both models were then passed to the next stage for normalization and error correction.

Figure 3. Sample images illustrating four surface categories, consisting of (a) flat, (b) curved, (c) reflective, and (d) textured

Table 1. Baseline CER comparison of PaddleOCR and LayoutLM across surface categories

Surface type	Number of images	Mean value of CER	
		Base PaddleOCR	Base LayoutLM
Flat	198	0.0885	0.3069
Curved	44	0.1508	0.3879
Reflective	24	0.2113	0.7003
Textured	29	0.2031	0.4171
Average		0.1634	0.4531

2.3. Normalization and SymSpell-based spelling correction

Both OCR outputs underwent a normalization stage to remove irregular characters (such as artifacts introduced by the OCR engine [29]) and correct spacing inconsistencies, particularly those that occur after punctuation marks immediately followed by letters. The normalized text was then segmented into individual tokens. Each token containing alphabetic characters was verified against the SymSpell general dictionary, and tokens not found were corrected using a domain-specific lexicon of food and beverage terms.

SymSpell is a spell-checking algorithm that extends Peter Norvig's approach based on Bayes' Theorem [30], [31]. It operates by leveraging a dictionary of valid words along with their frequencies. When a misspelled token is encountered, the algorithm calculates the edit distance between the token and the dictionary entries, then searches for candidate words with the smallest edit distance. If multiple candidates are identified, SymSpell selects the word with the highest frequency as the most likely correction [32].

A correction is applied when the similarity score between the detected token and its candidate replacement is greater than or equal to 0.75. SymSpell was selected as the core SymSpell algorithm because it combines computational efficiency with robust performance in constrained vocabulary domains. Although prior studies (e.g., Gasim [19]) have shown that SymSpell's accuracy decreases in morphologically rich languages, this limitation is mitigated in the food and beverage domain, where the vocabulary is relatively fixed. By leveraging a domain-specific lexicon, SymSpell effectively corrected misspelled tokens, reducing substitution and deletion errors while producing cleaner, more readable text for subsequent processing.

2.4. LayoutLM based compound word splitting

After spelling correction, the PaddleOCR output underwent a compound word segmentation stage guided by bigram references derived from LayoutLM. In this pipeline, LayoutLM was not used as a full OCR engine, but rather as a contextual bigram guide that provides semantic and spatial cues for detecting and

separating concatenated words. For example, a merged token such as "AirPenstabil" was successfully segmented into "Air" and "Penstabil."

This stage is essential because word concatenation can alter the semantic meaning of ingredients and mislead users in interpreting nutritional information. By leveraging LayoutLM's transformer-based embeddings which jointly encode textual, visual, and positional features, the system was able to detect boundary inconsistencies that traditional text-only methods overlook [21]-[34]. The combination of lexical correction from SymSpell and structural correction from LayoutLM produced text that was both syntactically and semantically coherent.

Technically, each bigram b_i is constructed from a pair of consecutive LayoutLM tokens drawn from an output sequence of length n. For every adjacent pair (l_i, l_{i+1}), a bigram is formed by concatenating the two tokens as defined in (1). These bigrams are then compared against each PaddleOCR token p_j from the OCR output sequence of length m. If the bigram exactly matches a token, the original token is automatically replaced with the bigram-derived suggestion $\hat{p}_j$ which inserts a space between l_i and l_{i+1} as defined in (2) and operationalized through the replacement rule that produces the corrected token p_j^* in (3). The final corrected text sequence P^* is obtained by applying this replacement rule across all m tokens as formalized in (4). This process not only resolves concatenation errors but also ensures that all non-alphabetic characters such as punctuation, numerals, and symbols remain preserved and properly integrated into the corrected sequence, resulting in text that is syntactically coherent and semantically more accurate.

$$b_i = l_i || l_{i+1},\ i = 1,2,\dots,n-1 \tag{1}$$

$$\hat{p}_j = l_i || " " || l_{i+1} \tag{2}$$

$$p_j^* = \begin{cases} \hat{p}_j, & if\ b_i = p_j \\ p_j, & otherwise \end{cases} \tag{3}$$

$$P^* = \{p_1^*, p_2^*, \dots, p_m^*\} \tag{4}$$

The original PaddleOCR token is replaced with the resulting pair, yielding a corrected token sequence that restores missing word boundaries. This mechanism enables the system to recover merged or tightly concatenated words. The output examples of each processing stage (baseline OCR, spelling correction, word segmentation, and post-OCR results) are summarized in Table 2.

Table 2. Examples of OCR outputs across different stages of the post-processing pipeline

Base OCR	Corrected mispelled	Segmented word	Post-OCR
komposisi:olesan kacang hazeldengan cokelat (81.68%) (gula,lemaknabati,kacang hazel(13%susu skim bubuk (8.7%kaka0bubuk(7.4%,pengemulsilesitin (kedelai) perisa sintetik vanila),tepung terigu (16.34%).ragi roti, ekstrak barley malt,_garam, susu skim bubuk pengemulsi:lesitin (kedelai), protein gandum, protein whey.total unsur susu: 7%.tanpa pengawet dan pewarna	kedela→kedelai	komposisi:olesan → komposisi: olesan, hazeldengan → hazel dengan, lemaknabati, → lemak nabati, pengemulsi:lesitin → pengemulsi: lesitin, lesitin(kedelai → lesitin (kedelai	komposisi: olesan kacang hazel dengan cokelat (81.68%) (gula, lemak nabati, kacang hazel(13%susu skim bubuk (8.7%kaka0bubuk(7.4%, pengemulsilesitin (kedelai) perisa sintetik vanila), tepung terigu (16.34%). ragi roti, ekstrak barley malt, _garam, susu skim bubuk pengemulsi: lesitin (kedelai), protein gandum, protein whey. total unsur susu: 7%. tanpa pengawet dan pewarna

2.5. Evaluation using CER

The text generated after post-processing is measured in terms of the CER, which measures the proportion of misrecognized characters produced by the OCR model [35]-[37]. The CER obtained from the post-processing results is then compared with the baseline CER of PaddleOCR to determine the effect of the proposed post-processing stage on improving the OCR output.

3. RESULTS AND DISCUSSION

This part provides the experimental results acquired from evaluating the proposed post-processing method on OCR outputs of food composition labels. The analysis centers on CER-based accuracy

measurements that cover substitutions, insertions, and deletions by comparing baseline PaddleOCR recognition with outputs produced after applying the SymSpell and LayoutLM hybrid correction pipeline. Evaluations were also conducted across flat, curved, reflective, and textured surfaces to assess the robustness of the method under different visual conditions. To ensure the reliability of the improvement, additional statistical validation using both parametric and non-parametric tests was performed. The flexibility of the proposed approach was further examined by applying the same pipeline to Tesseract, and its performance was also benchmarked against a BERT-based masked language model spell correction baseline. Finally, the computational cost and latency of the hybrid pipeline were analyzed to evaluate its practicality for food-label scanning applications.

The text detection and recognition component of this system adopts the same design as GiziLens, utilizing PaddleOCR PP-OCRv4 with three pre-trained models, each serving a specific function: (i) en_PP-OCRv3_det_infer, which detects image regions containing English text; (ii) en_PP-OCRv4_rec_infer, which recognizes and converts the detected regions into digital text; and (iii) ch_ppocr_mobile_v2.0_cls_infer, which classifies text images to differentiate between content types such as titles, tables, and paragraphs [10]. Before being presented to the user, the detected text is refined through a post-processing stage that incorporates SymSpell and LayoutLM to correct errors and improve accuracy. The text recognition performance of PaddleOCR enhanced with post-processing was evaluated using 295 OCR input images. The ground truth data were acquired by cropping the composition label areas from product packaging and manually transcribing them to ensure the reliability of the reference dataset.

Several stages were carried out to evaluate the proposed system. First, scanned and cropped images of composition labels from product packaging were input into the system. These images were subsequently processed with PaddleOCR to extract the corresponding digital text, which was stored as the raw OCR output. Subsequently, the extracted text was refined through a post-processing stage using the SymSpell and LayoutLM-based algorithm, and the results were stored as the final post-processed output. Finally, the post-processed text was evaluated against the ground truth using the CER metric to assess recognition accuracy.

3.1. Performance

Table 3 presents the effect of SymSpell's similarity threshold on correction performance, evaluated using the CER metric and its detailed error components: substitutions, insertions, and deletions. The similarity threshold in SymSpell is a critical parameter that determines whether a suggested correction is accepted or rejected. A lower threshold allows for more corrections to be applied but also increases the risk of introducing inappropriate substitutions. Conversely, a higher threshold applies corrections more selectively, which can reduce the number of valid corrections captured. Therefore, determining the optimal threshold is crucial to balance accuracy and reliability in post-processing.

Table 3. Effect of SymSpell similarity thresholds on OCR text correction performance across CER, substitutions, insertions, and deletions

Threshold of SymSpell similarity	Mean value of			
	CER	*s*	*i*	*d*
0.25	0.1103	3.878	1.3695	26.0237
0.50	0.1082	3.4576	1.3627	25.7864
0.75	0.0991	3.5322	1.3254	22.8508

At a threshold of 0.25, the system produced a CER of 0.1103 with errors predominantly in the form of deletions (an average of 26.0237 characters per document). A low threshold makes SymSpell less selective, as it accepts and applies corrections even when the similarity score is very low (e.g., 0.3). As a result, many tokens that are already correct or only slightly misspelled are unnecessarily replaced with irrelevant or incorrect words, leading to erroneous substitutions or the removal of important character sequences as shown in Table 4. Such corrections not only fail to improve the text but may also degrade overall accuracy, as reflected in the relatively high CER and the large number of deletions observed at this threshold.

When the threshold is increased to 0.50, the CER decreases slightly to 0.1082. At this value, SymSpell becomes more cautious in accepting correction suggestions because only corrections with a moderate to high similarity are applied. This helps reduce incorrect corrections, although the OCR error correction is not fully optimized as shown in Table 4.

A threshold of 0.75 performed best, with the lowest CER of 0.0991 and a significant decrease in the average deletion rate to 22.8508 characters. At this level, SymSpell becomes more selective, accepting only correction suggestions with very high similarity. Although the number of substitutions increased slightly

compared to the 0.50 threshold, this is likely because the corrections applied were truly relevant and accurate. The decrease in deletions indicates that previously deleted or unrecognized words are now being more successfully corrected or retained.

Overall, the findings underscore the critical role of choosing an appropriate similarity threshold to maximize the effectiveness of SymSpell. A threshold set too low leads to overly aggressive corrections that often introduce errors, whereas a threshold set too high may overlook valid corrections. In this study, a threshold of 0.75 achieved the best balance between correction capability and text accuracy, as indicated by the lowest CER among the evaluated settings. These results are consistent with prior OCR post-processing studies. Bjerring-Hansen *et al*. [17] reported up to 73% word accuracy improvement using SymSpell with context-aware modifications on historical texts. Had *et al*. [20] achieved a 21.4% accuracy increase on passport data, while Aung *et al*. [18] reported a 0.66% word error rate reduction in Myanmar OCR. These findings confirm that SymSpell provides an efficient and effective solution for lexical correction in constrained vocabulary domains such as food labels, particularly when combined with a domain-specific lexicon as also employed by Santos *et al*. [15].

Furthermore, integrating LayoutLM into the post-processing stage improved OCR performance by detecting and separating concatenated words caused by segmentation errors in PaddleOCR. The CER decreased from 0.1009 with SymSpell alone to 0.0991 with the combined method, confirming LayoutLM's complementary role in enhancing segmentation. This aligns with recent multimodal document understanding studies which demonstrate that LayoutLM's joint learning of textual and spatial features effectively addresses concatenation errors that dictionary-based methods alone cannot resolve [21], [22]. However, as noted by Lamott *et al.* [33], layout-aware model performance remains highly dependent on image quality. On noisy labels (curved, reflective, and textured surfaces), LayoutLM detects far fewer valid text regions, limiting its segmentation contribution. Nevertheless, even small CER reductions particularly those involving critical characters can meaningfully reduce the likelihood of misinterpretation in food safety and allergen contexts. Future improvements may be achieved through enhanced pre-processing such as geometric rectification, contrast enhancement, and fine-tuning LayoutLM on domain-specific food-label data.

This improvement can be attributed to the complementary roles of the two methods. SymSpell performs SymSpell by relying on the proximity of tokens to dictionary entries, but it cannot identify cases where two words are incorrectly merged into a single unit. In contrast, LayoutLM leverages spatial context and token distribution to detect word boundaries more effectively. By combining these approaches, the system can reduce errors at both the spelling and word segmentation levels, resulting in text outputs that are more accurate, readable, and informative for end users.

Table 4. Examples of OCR token corrections produced by SymSpell at a similarity threshold of 0.25, 0.5, and 0.75

PaddleOCR results	Tokens repaired by SymSpell $t = 0.25$	Tokens repaired by SymSpell $t = 0.5$	Tokens repaired by SymSpell $t = 0.75$
Komposisi:Tepung Terigu, Gula, Minyak Nabati, Cokelat Bubuk (5.979%,) SusuBubuk 4.950% Glukosa, EkstrakMalt,Garam, Pengembang (Amonium Bikarbonat, dan Natrium Bikarbonat),Pengemulsi Nabati (Lcsitin Kedclai, Mono dan Digliserida Asam Lemak), Perisa Sintetik (Cokelat,Susu)	Susububuk → bubuk, ekstrakmalt → ekstrak, lcsitin → lesitin, kedclai → kedelai, ronan → dengan, ie → de, seam → cream, cota → total, arate → anato, pentel → mentega, ned → mede, rcm → krim, rset → see, cld → bold,	Susububuk → bubuk, ekstrakmalt → ekstrak, lcsitin → lesitin, kedclai → kedelai	lcsitin → lesitin, kedclai → kedelai

To further validate these improvements, the post-processing results showed a clear and statistically significant reduction in ΔCER as shown in Table 5. Both parametric and non-parametric tests confirmed that this improvement was not due to random variation. The confidence interval (CI) around the estimated reduction was also narrow, indicating that the improvement is consistent and that the true effect is unlikely to differ much from the value observed in this study. Furthermore, the enhancement remained significant across all surface categories including flat, curved, reflective, and textured labels, even after applying a correction for multiple statistical comparisons, indicating that the improvement is genuine and not a result of chance due to the number of tests performed.

After validating the improvements on PaddleOCR, we applied the same pipeline to Tesseract to examine its generalizability. The pipeline produced a measurable CER reduction on Tesseract, confirming model-agnostic compatibility though the improvement was modest compared to PaddleOCR where the reduction was substantially larger. This suggests that the approach is more effective when the underlying OCR engine provides reliable character-level predictions. Our pipeline achieved an average CER of 0.0991, representing a 39% relative improvement over the PaddleOCR baseline of 0.1634. For context, Had *et al*. [20] achieved a 21.4% accuracy increase using SymSpell alone, while Aung *et al*. reported a 0.66% word error rate reduction [18]. Hemmer *et al*. [12] also noted that post-OCR denoising complexity varies across document types, supporting our hybrid lexical-structural approach.

In addition, we compared the proposed method with a BERT-based spell-correction baseline as shown in Table 6. Although widely used for general text cleaning, the BERT-based approach frequently introduced inappropriate substitutions such as replacing correct terms with words from an unrelated domain which increased CER across all categories. This failure reflects a critical challenge in OCR post-processing where general-purpose language models often lack domain awareness. Xu *et al*. [26] noted that purely text-based correction methods frequently fail in visually rich documents, while de Araújo *et al*. [16] showed that byte-level language models can introduce errors without domain-specific constraints. In contrast, our hybrid approach overcomes this limitation by combining dictionary-based lexical correction with layout-aware segmentation, ensuring corrections that are both lexically valid and semantically coherent within the food-label domain.

Table 5. Comparison of post-OCR improvements on PaddleOCR and Tesseract

OCR engine	Mean ΔCER	95% CI of ΔCER	Paired t-test	Wilcoxon signed-rank	Effect size (Cohen's dz)	Significance across image types (Holm–Bonferroni)
PaddleOCR	−0.019944	[−0.021598, −0.018291]	$t(295) = -23.7338, p = 2.21 \times 10^{-70}$	$W = 348, p = 3.27 \times 10^{-45}$	−1.3795 (large)	Significant across all types
Tesseract	−0.003716	[−0.004480, −0.002951]	$t(293) = -9.5652, p = 4.84 \times 10^{-19}$	$W = 684, p = 5.86 \times 10^{-20}$	−0.5579 (medium)	Significant across all types

Table 6. Comparative performance of the BERT-based spell corrector and the proposed hybrid post-OCR pipeline

Method	Avg ΔCER (post − base)	Failure modes	Latency	Notes
BERT-based spell corrector (MLM)	ΔCER ≈ +0.012 (CER increases)	Replaces correct words with unrelated tokens (e.g., mengandung → tergantung, ekstrak → ekstrim, total → tolak) Not domain-aware Applies substitutions even with very low prediction probabilities	≈ 6–9 seconds per label	Degrades OCR output; overly aggressive; unsuitable for nutritional-label text. Stable in execution but systematically incorrect.
Proposed hybrid method (SymSpell + LayoutLM)	ΔCER ≈ −0.0199 (CER significantly decreases)	Limited gains on severely distorted labels (curved + reflective surfaces) Cannot fix text that PaddleOCR fails to detect	≈ 0.3 seconds per label	Stable, domain-robust, and semantically safe; performs localized corrections without altering meaning; consistently improves OCR across all categories.

In summary, the combined SymSpell–LayoutLM method improves both spelling accuracy and word segmentation by addressing character-level errors and detecting merged tokens. Its overall effectiveness, however, remains constrained by the quality of the initial OCR outputs, particularly on visually complex labels where severe distortions limit recoverable text. Despite these limitations, the post-OCR pipeline consistently enhances recognition accuracy and provides a more reliable basis for downstream analysis.

3.2. Discussion

The evaluation results indicate that the SymSpell algorithm, combined with bigram references from LayoutLM, successfully reduces the CER across all image categories as shown in Table 7. This aligns with prior OCR correction studies which demonstrated that character-level errors are most effectively corrected when the baseline OCR output has a relatively low initial error rate, and that packaging distortions often lead to recognition errors and word concatenation [11], [12], [35]. The CER decreased from 0.0885 to 0.0659 for flat surfaces accompanied by fewer substitutions and a marked decrease in deletions. Although insertions

increased slightly due to over-segmentation, the overall error rate declined. This superior performance on flat surfaces is consistent with observations by Rosyadi *et al*. [10], who noted that PaddleOCR achieves its highest accuracy on flat packaging where minimal distortion occurs, and is also supported by de Lima and da Silva who demonstrated that image preprocessing techniques significantly improve OCR accuracy by reducing visual noise [13]. The stable visual characteristics of flat surfaces enable higher initial accuracy, providing an advantage for post-processing where SymSpell corrects spelling errors based on word similarity and LayoutLM assists in separating concatenated words, resulting in more consistent improvements compared to categories with greater visual complexity.

Table 7. Comparative performance of PaddleOCR baseline and post-OCR Correction across image types

Surface type	Number of images	Mean value of CER		Substitutions		Insertions		Deletions	
		Base	Post-OCR	Base	Post-OCR	Base	Post-OCR	Base	Post-OCR
Flat	198	0.0885	**0.0659**	4.0152	**3.4040**	**1.4293**	1.6970	18.6919	**13.0455**
Curved	44	0.1508	**0.1350**	6.0227	**5.1818**	**0.4091**	0.6591	52.8182	**46.7727**
Reflective	24	0.2113	**0.1986**	2.8750	**2.5417**	0.7083	**0.6250**	43.9167	**41.7500**
Textured	29	0.2031	**0.1891**	3.3448	**2.7241**	**0.3448**	0.3793	40.2414	**37.8621**
Average		0.1634	**0.1472**	4.0644	**3.4629**	**0.7229**	0.8401	38.9171	**34.8576**

In the curved category, the CER decreased from 0.1508 to 0.1350 after post-processing though the improvement was smaller than in the flat category. The post-OCR stage effectively corrected misspellings and separated concatenated words such as “lemaknabati” and “pengemulsi:lesitin.” However, deletion errors remained relatively high due to character loss from non-uniform character scaling where characters appear larger in the center and smaller near the edges. This challenge consistent with Hemmer *et al*. [12] who noted that OCR accuracy decreases significantly on non-planar surfaces. This limitation is also supported by Xu *et al.* [26] who observed that purely text-based correction methods frequently fail in visually rich documents. Applying advanced preprocessing techniques such as perspective correction as suggested by recent studies, could reduce these distortions and improve recognition accuracy before the post-processing stage [13].

For reflective surfaces, the CER decreased from 0.2113 to 0.1986 while for textured surfaces it decreased from 0.2031 to 0.1891, demonstrating that post-processing improves accuracy even under complex visual conditions. However, the challenges on these surfaces highlight the inherent limitations of OCR engines in handling extreme visual conditions as noted by Rosyadi *et al*. regarding contrast gradients on reflective surfaces [10] and de Oliveira *et al*. [11] regarding visual noise from background textures. These findings suggest that while post-processing can mitigate some errors, visual preprocessing remains essential for handling severely distorted images.

The proposed OCR pipeline has practical potential for mobile dietary tracking and food safety applications as it can be integrated into smartphone apps to automatically extract nutritional information and ingredient details from product labels, enabling faster food logging and real-time alerts for ingredients of concern. The OCR and post-processing components operate fully offline and can be expanded to support additional safety and nutrition functions. Beyond individual use, the pipeline offers broader applicability in supply chain, retail, and regulatory environments. In supply chain operations, it can assist with supplier verification and traceability by automatically extracting batch or production information from labels. In retail settings, it can support shelf auditing and ensure that displayed product data such as nutritional or allergen information matches the information printed on packaging. From a regulatory standpoint, the pipeline can help automate compliance checks by inspecting mandatory label components including ingredient lists, nutritional declarations, and expiry dates. These capabilities are supported by recent studies showing that multimodal approaches combining text and layout information significantly improve document data extraction accuracy [24], and that LayoutLM-based models achieve superior performance when fine-tuned on domain-specific data [25]. Although commercial OCR services such as Google Vision or AWS Textract may offer strong accuracy, they require continuous connectivity and incur usage costs. In contrast, the offline-capable pipeline developed in this work provides advantages in latency, cost efficiency, and data privacy, with online access needed only for advanced nutritional interpretation.

Finally, it is important to note that the current post-processing modules (SymSpell and LayoutLM) primarily address lexical and contextual errors rather than visual distortions or segmentation issues. Future research should therefore explore visual distortion correction techniques such as geometric rectification, contrast enhancement, or text-rectification networks prior to OCR to generate cleaner input text and further improve recognition accuracy.

3.3. Computational cost and latency

The proposed hybrid pipeline was benchmarked on a Kaggle Linux environment (Intel® Xeon® 2.20 GHz, 4 vCPUs, 31 GB RAM, no GPU). Processing times were measured across 295 label images, covering the entire workflow from input acquisition to final post-processing output. PaddleOCR required 1.11 seconds per image on average (95% CI: 1.01–1.20 seconds), while the LayoutLM-based post-OCR module added 0.43 seconds per image (95% CI: 0.41–0.46 seconds). The full pipeline achieved an end-to-end latency of 1.54 seconds per image (95% CI: 1.43–1.65 seconds), representing a 39% overhead relative to PaddleOCR alone. Despite this increase, the total runtime remains within 1–2 seconds per label, which is acceptable for practical food label scanning scenarios. The post-OCR operations are lightweight and are expected to benefit further from faster CPUs or optional GPU acceleration.

4. CONCLUSION

This study introduced a hybrid post-OCR correction system that combines PaddleOCR, SymSpell, and LayoutLM to enhance text recognition on food composition labels. The method strengthens OCR output by resolving misspellings and word concatenations through domain-aware lexical correction and layout-guided segmentation, resulting in more coherent and readable text. The system also provides a modular and model-agnostic design, offering a practical foundation for downstream applications such as mobile dietary tracking, food safety monitoring, supply chain traceability, retail shelf auditing, and automated regulatory compliance checks. Looking ahead, this approach can be further optimized to reduce latency and enhance stability across varied packaging conditions, particularly on curved or reflective surfaces through improved preprocessing, and can be extended to multilingual OCR to broaden its applicability across diverse labeling environments.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support from Universitas Qomaruddin and its Research and Community Service Institute (LPPM) for facilitating this study. Appreciation is also extended to colleagues for their valuable feedback on the PaddleOCR–SymSpell–LayoutLM post-processing development.

FUNDING INFORMATION

This work was supported by the Penelitian Dosen Pemula (PDP) scheme funded by the Directorate of Research, Technology, and Community Service (DRTPM), Ministry of Education, Culture, Research, and Technology, Republic of Indonesia, under the Fiscal Year 2025 program.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ahmad Wahyu Rosyadi	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
Siti Ma'shumah		✓		✓		✓	✓	✓	✓			✓		✓
Muhammad Qomaruz Zaman		✓		✓		✓				✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

With respect to the content of this paper, the authors confirm that there are no personal, financial, or professional interests that could have influenced this work.

DATA AVAILABILITY

Upon request and with a valid justification, the associated author may provide the research data.

REFERENCES

[1] B. Ozhan and B. Tugrul, "Analysis of Turkish cuisine flavors network," *International Journal of Food Science and Technology*, vol. 59, no. 2, pp. 908–915, 2024, doi: 10.1111/ijfs.16849.

[2] Q. A'yuni, N. Nasaruddin, M. Irhamsyah, M. Azhary, and R. Roslidar, "Intelligent Tuberculosis Detection System with Continuous Learning on X-ray Images," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 1, pp. 130–141, Nov. 2025, doi: 10.35882/jeeemi.v7i1.572.

[3] S. Bala, K. Arora, R. Jeevitha, R. Chowdhury, P. Kumar, and C. Shobana Nageswari, "A Novel Encoder Decoder Architecture with Vision Transformer for Medical Image Segmentation," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 1, pp. 176–186, Jan. 2025, doi: 10.35882/jeeemi.v7i1.571.

[4] Wiharto *et al*., "Model Group Decision Support System Based on Depression Anxiety Stress Scales Using Ordered Weighted Averaging Aggregation Method," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 2, pp. 307–317, Feb. 2025, doi: 10.35882/jeeemi.v7i2.678.

[5] M. Maringer, N. Wisse-Voorwinden, P. V. t. Veer, and A. Geelen, "Food identification by barcode scanning in the Netherlands: A quality assessment of labelled food product databases underlying popular nutrition applications," *Public Health Nutrition*, pp. 1–8, Jul. 2018, doi: 10.1017/s136898001800157x.

[6] R. N. Kadu and S. N. Pawar, "Advanced Bi-CNN for Detection of Knee Osteoarthritis using Joint Space Narrowing Analysis," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 1, pp. 80–90, 2024, doi: 10.35882/jeeemi.v7i1.574.

[7] N. R. Mavani, M. A. Ismail, N. Abd Rahman, and J. Mohd Ali, "Fuzzy logic-based barcode scanning system for food products halal identification," *Food Control*, vol. 168, no. April 2024, p. 110926, 2025, doi: 10.1016/j.foodcont.2024.110926.

[8] S. M. Gondowijoyo, R. Setiawan, and N. F. Hikmah, "Applying artificial neural network on heart rate variability and electroencephalogram signals to determine stress," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 22, no. 4, pp. 910–920, 2024, doi: 10.12928/TELKOMNIKA.v22i4.25832.

[9] T. H. Thuy, H. M. Toan, V. T. Lam, and N. H. Phuong, "SYNDIAG: an expert system for disease syndrome diagnosis of traditional Vietnamese medicine," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 22, no. 5, pp. 1167–1177, 2024, doi: 10.12928/TELKOMNIKA.v22i5.25831.

[10] A. W. Rosyadi, S. Ma'shumah, M. Q. Zaman, and M. R. Fajar, "Ingredients Identification Through Label Scanning Using PaddleOCR and ChatGPT for Information Retrieval," *Jurnal RESTI*, vol. 8, no. 6, pp. 758–767, 2024, doi: 10.29207/resti.v8i6.6119.

[11] L. L. de Oliveira *et al*., "Evaluating and mitigating the impact of OCR errors on information retrieval," *International Journal on Digital Libraries*, vol. 24, no. 1, pp. 45–62, 2023, doi: 10.1007/s00799-023-00345-6.

[12] A. Hemmer, J. Brachat, M. Coustaty, and J. M. Ogier, "Estimating Post-OCR Denoising Complexity on Numerical Texts," *Communications in Computer and Information Science*, vol. 1863 CCIS, pp. 67–79, 2023, doi: 10.1007/978-3-031-42430-4_6.

[13] F. B. de Lima and E. G. da Silva, *Enhancing Text Recognition in OCR Systems Through Image Processing with BSRGAN*, vol. 1, no. 1. Association for Computing Machinery, 2025, doi: 10.5753/sbsi.2025.246551.

[14] M. Ogrodniczuk, "Fine-Tuning OCR Error Detection and Correction in a Polish Corpus of Scientific Abstracts," *Communications in Computer and Information Science*, vol. 1716 CCIS, pp. 450–461, 2022, doi: 10.1007/978-981-19-8234-7_35.

[15] M. K. Santos, G. Bazzo, L. L. de Oliveira, and V. P. Moreira, "ESTER-Pt: An Evaluation Suite for TExt Recognition in Portuguese," in *Lecture Notes in Computer Science*, G. A. Fink, R. Jain, K. Kise, and R. Zanibbi, Eds., Cham: Springer Nature Switzerland, 2023, pp. 366–383, doi: 10.1007/978-3-031-41682-8_23.

[16] S. S. De Araújo, B. L. D. Bezerra, and A. F. De Sousa Neto, "A Proposal of Post-OCR Spelling Correction Using Monolingual Byte-level Language Models," in *DocEng 2025 - Proceedings of the 2025 ACM Symposium on Document Engineering*, in DocEng '25. New York, NY, USA: Association for Computing Machinery, 2025, doi: 10.1145/3704268.3748673.

[17] J. Bjerring-Hansen, R. Deans Kristensen-McLachlan, P. Diderichsen, and D. Haltrup Hansen, "Mending Fractured Texts," *Digital Humanities in the Nordic and Baltic Countries Publications*, vol. 4, no. 1, pp. 177–186, 2022, doi: 10.5617/dhnbpub.11285.

[18] T. Aung, Y. K. Thu, and M. N. Oo, "myOCR: Optical Character Recognition for Myanmar language with Post-OCR Error Correction," *19th International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2024*, no. November, 2024, doi: 10.1109/iSAI-NLP64410.2024.10799448.

[19] I. Gasim, "Evaluating Hunspell, SymSpell, Norvig, and N-gram Spellcheckers for Azerbaijani Text," *International Journal of Innovative Science and Research Technology*, vol. 10, no. 5, pp. 2574–2577, 2025, doi: 10.38124/ijisrt/25may1640.

[20] I. S. Had, W. Maulana Baihaqi, and D. P. N. Kinding, "Improving Tesseract OCR Accuracy Using SymSpell Algorithm on Passport Data," *Sinkron*, vol. 9, no. 1, pp. 374–381, 2025, doi: 10.33395/sinkron.v9i1.14395.

[21] Y. Huang, T. Lv, L. Cui, Y. Lu, and F. Wei, *LayoutLMv3: Pre-training for Document AI with Unified Text and Image Masking*, vol. 1, no. 1. Association for Computing Machinery, 2022, doi: 10.1145/3503161.3548112.

[22] Y. Xu *et al*., "LayoutLMv2: Multi-modal pre-training for visually-rich document understanding," in *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, C. Zong, F. Xia, W. Li, and R. Navigli, Eds., Online: Association for Computational Linguistics, Aug. 2021, pp. 2579–2591, doi: 10.18653/v1/2021.acl-long.201.

[23] B. Uthayasooriyar, A. Ly, F. Vermet, and C. Corro, "Training LayoutLM from Scratch for Efficient Named-Entity Recognition in the Insurance Domain," *Proceedings - International Conference on Computational Linguistics, COLING*, pp. 101–110, 2025, doi: 10.48550/arXiv.2412.09341.

[24] M. Bajrami, N. Ackovska, B. Stojkoska, P. Lameski, and E. Zdravevski, "Deep Dive into Invoice Intelligence: A Benchmark Study of Leading Models for Automated Invoice Data Extraction," in *Lecture Notes in Networks and Systems*, X.-S. Yang, S. Sherratt, N. Dey, and A. Joshi, Eds., Singapore: Springer Nature, 2024, pp. 177–191, doi: 10.1007/978-981-97-3289-0_15.

[25] K. Fathima, A. Athar, and H. C. Kim, "Multi-Class Document Classification using LayoutLMv1 and V2," in *International Conference on Advanced Communication Technology, ICACT*, IEEE, 2024, pp. 286–289, doi: 10.23919/ICACT60172.2024.10471992.

[26] Q. Xu, L. Wang, H. Liu, and N. Liu, *LayoutLM-Critic: Multimodal Language Model for Text Error Correction of Optical*

Character Recognition, vol. 1701 CCIS. Springer Nature Singapore, 2022, doi: 10.1007/978-981-19-7943-9_11.
[27] H. Zhang *et al.*, "Identification of Illegal Outdoor Advertisements Based on CLIP Fine-Tuning and OCR Technology," *IEEE Access*, vol. 12, no. July, pp. 92976–92987, 2024, doi: 10.1109/ACCESS.2024.3424258.
[28] F. Ren *et al.*, "Research on Algorithm for License Plate Detection in Complex Scenarios Based on Artificial Intelligence," in *2023 3rd International Conference on Communication Technology and Information Technology, ICCTIT 2023*, 2023, pp. 166–170, doi: 10.1109/ICCTIT60726.2023.10435878.
[29] D. Wang *et al.*, "BuDDIE: A Business Document Dataset for Multi-task Information Extraction," *Proceedings - International Conference on Computational Linguistics, COLING*, no. Dc, pp. 35–47, 2025.
[30] M. Safitri, E. M. Yuniarno, and R. F. Rachmadi, "Indonesian Sign Language (SIBI) Recognition and Extraction Using Convolutional Neural Networks - Symmetric Deletion Spelling Correction," *Proceedings - International Seminar on Intelligent Technology and its Applications, ISITIA*, no. 2024, pp. 220–225, 2024, doi: 10.1109/ISITIA63062.2024.10667714.
[31] Z. Mansur, N. Omar, S. Tiun, and E. M. Alshari, "A normalization model for repeated letters in social media hate speech text based on rules and spelling correction," *PLoS ONE*, vol. 19, no. 3 March, pp. 1–25, 2024, doi: 10.1371/journal.pone.0299652.
[32] L. D. F. Vantie and S. Rochimah, "Refining Words for Improved BERT Model Performance in User Review Classification," *ICADEIS 2025 - 2025 International Conference on Advancement in Data Science, E-learning and Information System: Integrating Data Science and Information System, Proceeding*, pp. 1–6, 2025, doi: 10.1109/ICADEIS65852.2025.10933336.
[33] M. Lamott, Y. N. Weweler, A. Ulges, F. Shafait, D. Krechel, and D. Obradovic, "LAPDoc: Layout-Aware Prompting for Documents," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14807 LNCS, pp. 142–159, 2024, doi: 10.1007/978-3-031-70546-5_9.
[34] J. Wang, L. Jin, and K. Ding, "LiLT: A Simple yet Effective Language-Independent Layout Transformer for Structured Document Understanding," *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 7747–7757, 2022, doi: 10.18653/v1/2022.acl-long.534.
[35] C. Neudecker, K. Baierer, M. Gerber, C. Christian, A. Apostolos, and P. Stefan, "A survey of OCR evaluation tools and metrics," *ACM International Conference Proceeding Series*, no. October, pp. 13–18, 2021, doi: 10.1145/3476887.3476888.
[36] E. Vidal, A. H. Toselli, A. Ríos-Vila, and J. Calvo-Zaragoza, "End-to-End page-Level assessment of handwritten text recognition," *Pattern Recognition*, vol. 142, p. 109695, 2023, doi: 10.1016/j.patcog.2023.109695.
[37] A. Hemmer, M. Coustaty, N. Bartolo, and J. M. Ogier, "Confidence-Aware Document OCR Error Detection," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14994 LNCS, pp. 213–228, 2024, doi: 10.1007/978-3-031-70442-0_13.

BIOGRAPHIES OF AUTHORS

Ahmad Wahyu Rosyadi is a lecturer and researcher at Qomaruddin University, specializing in Informatics Engineering. He received his Bachelor's degree in Informatics Engineering from Universitas Islam Negeri Maulana Malik Ibrahim Malang in 2014 and his Master's degree in the same field from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, in 2018. His research interests focus on artificial intelligence, digital image processing, computer vision, machine learning, and information retrieval. He can be contacted at email: awrosyadi@gmail.com.

Siti Ma'shumah is a researcher and lecturer at Qomaruddin University specializing in Electrical Engineering. She earned her Bachelor's degree in Electrical Engineering from Institut Teknologi Nurul Jadid in 2015 and her Master's degree in the same field from Institut Teknologi Sepuluh Nopember in 2018. Her research interests include analog electronics, IoT, artificial neural networks and applications related to the electrical field. She can be contacted at email: sitimashumah@uqgresik.ac.id.

Muhammad Qomaruz Zaman earned his Bachelor of Engineering and Master of Engineering degrees in Electrical Engineering from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia, in 2018. He completed his Ph.D. at the National Taipei University of Technology, Taiwan, in 2025. He currently serves as a faculty member in the Department of Electrical Engineering at Institut Teknologi Sepuluh Nopember. His research focuses on robotics and intelligent systems, particularly in the areas of fuzzy inference systems, evolutionary algorithms, and machine learning for autonomous applications. He can be contacted at email: me@zamzaman.com.