



Pratya Nuankaew <nuankaew.p@gmail.com>

Manuscript ID: JAIT-20603 - Major Revisions

1 message

Ms. Stella Chan <stella.chan@ejournal.net>
To: Pratya Nuankaew <pratya.nu@up.ac.th>

Tue, Feb 24, 2026 at 8:36 AM

Dear Wongpanya S. Nuankaew, Pathapol Jomsawan, Kuljira S. Nuankaew, Kaewpanya S. Nuankaew, Pratya Nuankaew,

Thank you for submitting your manuscript "A Comparative Evaluation of Text Detection and Recognition Models for Cross-Hospital Generalization of Thai Pharmaceutical Label OCR under Limited Data Conditions" to Journal of Advances in Information Technology.

The editorial team has assessed your submission and feels that it has potential for publication, so we would like to invite you to make major revisions for further review.

You can find your manuscript at the following link:

<https://ojs.ejournal.net/index.php/jait/dashboard/mySubmissions?workflowSubmissionId=20603>

Important notice: Please revise the manuscript according to the reviewers' comments and upload the revised file **within one month. The revisions should be highlighted**, for example using the "Track Changes" function in Microsoft Word, so that changes are easily visible to the editors and reviewers. **Please provide a cover letter** to explain point-by-point the details of the revisions in the manuscript and your responses to the reviewers' comments. (download [author response template](#))

As the reviewer has suggested that your manuscript should **undergo extensive English editing**, please address this during revision. We suggest that you have your manuscript checked by a native English-speaking colleague or use a professional English editing service.

Do not hesitate to contact us if you have any questions regarding the revision of your manuscript.

Ms. Stella Chan

--

Journal of Advances in Information Technology

Website: <http://www.jait.us/>

Email: editor@jait.us

Twitter: [@Jait276068](https://twitter.com/Jait276068)

Indexed in **ESCI (Web of Science, Impact Factor 2024: 1.5)**, **Scopus (CiteScore 2024: 4.8)**

Note: To avoid missing our emails, please add submission@mail-ojs.ejournal.net to your contacts or safe senders list.

Disclaimer: The information and files contained in this message are confidential and intended solely for the use of the individual or entity to whom they are addressed. If you have received this message in error, please notify me and delete this message from your system. You may not copy this message in its entirety or in part, or disclose its contents to anyone.

--

Reviewer A:

The manuscript effectively addresses the challenges of Text Detection and Recognition Models for Cross-Hospital Generalization of Thai Pharmaceutical Label OCR under Limited Data Conditions. The paper's structure and readability ensure that readers can easily follow the work. I would like to offer a few comments for further improvement:

1. In abstract, the result of this work must be described briefly with data. The result of this work is not clear.
2. In Section 1, the authors should further highlight the challenges of handling the proposed problem. Why can the existing work not address this problem well? The authors also need to highlight and supplement the motivation of this work.

3. It will be advisable to place research gap and mention how you addressed the gap.
4. In the Related Work section, it will be advisable to compare the State-of-the-Art with current paper in a tabular form.
5. Some good work could be discussed in intro/literature like: Quantum-based privacy-preserving techniques for secure and trustworthy internet of medical things an extensive analysis
6. The manuscript is strong experimentally, but its core contribution is framed primarily as empirical comparison. Please explicitly clarify whether the novelty lies in:
 - (a) the three-level evaluation protocol,
 - (b) the cross-hospital generalization insights, or
 - (c) the architectural implications for OCR deployment.
7. The flow of the model is not clearly presented. Consider using a graphical representation with numbered steps inside the figure to help readers better understand the process.
8. Some lemma and theorems should be added and the proofs of the theorems to support the new idea of the paper should be added as Appendix. Mathematics modeling to support and analyze the method is not enough. Algorithms should be provided; the theorems and key equations should be embedded into the steps of Algorithms. The cost or complexity analysis of the method or the technology should be added.
9. Model size and inference implications are mentioned qualitatively. Please consider adding: inference latency or FPS comparisons for PaddleOCR vs. TrOCR, and a brief discussion of hardware assumptions.
10. The meaning of variables is not clear. Readers will be confused. To help readers' understanding, the authors should add a notation list.
11. While aggregate metrics are thorough, please consider adding: qualitative error examples for recognition failures in unseen hospitals (e.g., tonal marks, spacing), or a brief taxonomy of common cross-domain OCR errors.
12. Discuss security analysis in Discussion section.
13. Paper contains few grammar mistakes which will be cooperated in final version.
14. Consider including a subsection titled 'Limitations and Future Scope' in the Results and Discussion section.

Reviewer B:

The manuscript presents a systematic and well-structured empirical evaluation of cross-hospital OCR generalization for Thai pharmaceutical labels. The three-level evaluation protocol is particularly valuable and clearly distinguishes between in-domain fitting and real cross-domain robustness. However, several issues must be addressed before publication

- 1 - The definition of ACC as $1 - \text{CER}$ is mathematically incorrect when ACC refers to sentence-level exact match accuracy. This must be clarified and corrected.
- 2 - The dataset size is relatively small for strong generalization claims; the authors should justify its adequacy and discuss limitations.
- 3 - Statistical significance analysis (e.g., bootstrap confidence intervals for CER/NED) would strengthen technical soundness.
- 4 - Claims regarding Transformer overfitting should be softened or supported with additional analysis.
- 5 - The introduction contains redundancy and should be streamlined.

Overall: the paper provides practical insights and makes a meaningful contribution, particularly regarding the importance of pre-training diversity over model size. More details about the identified issues are described below:

Redundancy in the Introduction (Section I)

Location: Introduction, first two paragraphs.

Issue:

The first and second paragraphs repeat the same ideas regarding:

- The complexity of Thai script (tonal markers, character combinations, context-dependent rendering)
- Domain shift across hospitals (layout, fonts, printing systems)

This creates conceptual redundancy and weakens the impact of the introduction.

Suggestion:

Merge the first two paragraphs into a single, concise paragraph that:

- Introduces Thai script complexity
- Introduces cross-hospital domain shift
- Clearly states the problem in one unified flow

Grammatical and Stylistic Errors

Parallel Structure

Location: Motivation and Problem Statement section.

Original:

"Both hybrid CNN–Transformer [2], [3], [4] and pure Transformer models [5] achieve strong benchmark results..."

Issue:

Lack of proper structural parallelism.

Suggested correction:

"Both hybrid CNN–Transformer architectures [2–4] and pure Transformer architectures [5] achieve strong benchmark results..."

Incorrect Word Usage

Location: Evaluation Metrics – F1 definition.

Original:

“The F1-Score: The harmonic meaning of precision and recall.”

Issue:

“Meaning” is incorrect.

Suggested correction:

“The F1-score is the harmonic mean of precision and recall.”

Redundant Verb in IoU Definition

Location: Detection Metrics section.

Original:

“where θ is the IoU threshold is (typically 0.50).”

Issue:

Duplicated verb “is”.

Suggested correction:

“where θ is the IoU threshold (typically 0.50).”

Inconsistent Capitalization

Location: Recognition Metrics section.

Original:

“We Compute:”

Issue:

Inconsistent capitalization.

Suggested correction:

“We compute:”

Metric Definition Inconsistency (Major Issue)

Location: Recognition Metrics section.

The manuscript defines:

$ACC = 1 - CER$

Later it states:

“In this study, ACC refers to sentence-level exact match accuracy...”

This is mathematically inconsistent.

- CER is a character-level metric.
- Sentence-level exact match accuracy is a different metric.
- They are not complementary measures.

Suggested correction:

- Separate clearly:

o Character-level accuracy ($1 - CER$)

o Sentence-level exact match accuracy

- Remove the equation $ACC = 1 - CER$ if ACC refers to sentence-level accuracy.

- Use distinct notation (e.g., ACC_{char} and ACC_{sent}).

This is the most important technical writing issue in the manuscript.

Inconsistent Terminology and Formatting

Fine-tuning Notation

Throughout the manuscript, different forms are used:

- “Fine-tuned”
- “FT”
- “Finetuned”
- “FT lr=0.001”

Issue:

Inconsistent terminology reduces formal quality.

Suggestion:

Standardize as:

“Fine-tuned (FT)”

Then use “FT” consistently in tables.

Learning Rate Formatting

Different formats are used:

- 0.00001
- 1×10^{-5}

Suggestion:

Use scientific notation consistently:

1×10^{-5}

This improves academic consistency.

Figure Numbering Inconsistency

In some sections, references such as “Figure 4” do not match previous numbering order.

Suggestion:

Verify all figure references and ensure numbering consistency throughout the manuscript.

Table IV – Identical Test and Unseen Values

Location: Text Detection Results (Table IV).

Pretrained and fine-tuned detection models show identical Test and Unseen results.

Issue:

This may raise reviewer suspicion if not explicitly explained.

Suggestion:

Add a short clarification paragraph explaining:

- Whether the same checkpoint was used
- Whether rounding caused apparent equality
- Whether domain adaptation truly had no measurable effect

Without clarification, reviewers may suspect reporting or experimental inconsistencies.

Minor Style Improvements

- Replace informal phrasing such as:

“This setup provides a straightforward interpretation...”

With:

“This setup enables clear interpretation...”

- Avoid strong absolute statements such as:

“Transformer models fail...”

Replace with:

“Transformer models exhibit higher sensitivity to domain shift...”

This improves academic neutrality.