

REVIEWER'S REPORT

Manuscript No.: IJAR-54192

Title: AI-Assisted Detection of Gastric Intestinal Metaplasia: Design and Validation of the "IntelliMeta" Algorithm

Recommendation:

Accept as it is

Accept after minor revision.....

Accept after major revision yes.....

Do not accept (*Reasons below*)

Rating	Excel.	Good	Fair	Poor
Originality		yes		
Techn. Quality		yes		
Clarity		yes		
Significance		yes		

Reviewer Name: Dr.Shaweta Sachdeva

Date: 8/10/25

Detailed Reviewer's Report

1. Limited Dataset Size and Imbalance

- The dataset (229 slides, 902 patches) is too small and imbalanced (173 normal vs. 56 GIM).
 - *Recommendation:* Expand the dataset with multicentric data collection or public datasets (e.g., TCGA, PAIP) to improve model generalization. Apply techniques like **SMOTE** or **class-weight adjustment** to mitigate imbalance.

2. Moderate Model Accuracy

- The reported accuracy (56.5%) and recall for positive GIM cases (35%) are below acceptable diagnostic standards for clinical AI.
 - *Recommendation:* Optimize hyperparameters, refine data augmentation, and explore more powerful architectures (e.g., EfficientNet, ResNet, or hybrid CNN-transformer models). Include **ROC curves and AUC** to better represent diagnostic performance.

3. Lack of Statistical Significance Testing

- The performance results are presented descriptively without confidence intervals, p-values, or statistical comparisons among models.

REVIEWER'S REPORT

➤ *Recommendation:* Include **statistical validation** (e.g., cross-validation with variance reporting) to ensure the reliability of performance claims.

4. Incomplete Model Evaluation

- Evaluation focuses only on accuracy, sensitivity, and specificity. Metrics like **F1-score**, **Precision-Recall curves**, **IoU**, and **Dice coefficients** are missing.
 - *Recommendation:* Add these metrics for a more comprehensive and comparable performance analysis.

5. Weak Comparative Analysis

- The results section mentions baseline models but lacks **quantitative comparison tables or visual plots** illustrating improvements.
 - *Recommendation:* Present clear comparative charts or tables with baseline and improved results to demonstrate the benefits of the IntelliMeta model.

6. Segmentation Validation Not Quantitatively Assessed

- While segmentation outputs are described, no objective metrics (e.g., IoU, Dice score) are provided to assess segmentation accuracy.
 - *Recommendation:* Quantify segmentation performance using standard metrics and include representative visual examples with ground truth overlays.

7. Insufficient Methodological Detail

- The algorithm's implementation lacks detail (e.g., optimizer type, batch size, learning rate, number of epochs, hardware used).
 - *Recommendation:* Provide full training and implementation details for reproducibility and technical transparency.

8. Limited Discussion of Limitations

- The discussion briefly mentions dataset limitations but fails to critically analyze other potential issues such as overfitting, bias, and generalization.
 - *Recommendation:* Expand the **limitations section** to address these and suggest concrete mitigation strategies.

REVIEWER'S REPORT

9. Need for External or Cross-Center Validation

- The study uses data from a single institution, which limits generalizability.
 - *Recommendation:* Include or propose **multicentric validation** to assess the model's performance on independent datasets.

10. Clinical Integration and Interpretability

- The manuscript briefly mentions a GUI but does not evaluate its usability or interpretability for clinical application.
 - *Recommendation:* Add a **usability evaluation** or discussion on how IntelliMeta can integrate into real-world pathology workflows, including interpretability and decision-support roles.

11. Insufficient Comparison with Related Work

- The literature review does not position IntelliMeta strongly against similar existing studies (e.g., AI-based gastric lesion detection by Iwaya et al., 2023).
 - *Recommendation:* Expand the **related work** section with comparative discussion to highlight the novelty and advantages of the proposed approach.

12. Figure and Table Improvements

- Figures (e.g., segmentation results, GUI screenshots) are not well-labeled or described.
 - *Recommendation:* Improve figure resolution, add legends, and ensure each visual aids comprehension of the findings.

13. Language and Structural Revision

- The manuscript includes minor grammatical inconsistencies and formatting errors (e.g., “—GIM” quotes, inconsistent headings).
 - *Recommendation:* Perform professional **language and formatting editing** for better readability and presentation.