01-Nov-2025

JBHI Ref: JBHI-05124-2025
Reject/Resubmit (major revision and new external review required)

Dear Dr. Ling He,

This letter is to inform you that the peer review process has concluded for manuscript, "DualSeg: Unified multi-scale framework with dual-stage encoder for glomerular segmentation," which you had submitted for possible publication in the IEEE Journal of Biomedical and Health Informatics (J-BHI).

The Associate Editor responsible for your manuscript review has received feedback from independent reviewers and compiled their evaluations. It is the recommendation of the Associate Editor and the Editor-in-Chief that your manuscript requires a MAJOR REVISION before it can be accepted for publication in J-BHI. Please note the J-BHI Editorial Policy that only one major revision is allowed for any submitted manuscript. This means that if the review recommendation for your revised paper triggers another major revision, the paper will be rejected automatically.

Enclosed, please find the comments by the Associate Editor and all the reviewers. I hope that the feedback is helpful for further improving the quality of your manuscript. If you decide to resubmit a revised manuscript, this must be done within 10 weeks (not extendable) from the date of this message. Please quote the above manuscript reference number for all future correspondence.

***IMPORTANT: PLEASE RETURN HERE TO SUBMIT ALL FILES:** https://ieee.atyponrex.com/journal/jbhi-embs

Also note that it is mandatory to enter your replies to the reviewers' questions and indicate how you have dealt with their comments in the revised manuscript. Please include your replies in the authors' response section or a separate file with a point-by-point explanation of all the changes made. Please do NOT include it in the cover letter since this is not accessible by the reviewers. For submitting your revised manuscript, please ensure all changes are highlighted in the manuscript to facilitate the review process.

Citing/including papers suggested by the reviewers or the Associate Editor is up to the authors and only if they add value to your work. Please also report to fotiadis@uoi.gr any suspicious suggestion by the reviewers.

The authors are responsible to follow the J-BHI publication rules for maximum number of pages, quality of figures, etc. as they are mentioned in: https://www.embs.org/jbhi/prepare-and-submit-your-manuscript/
On the opposite your article might need to be returned to you and the review process will be delayed and in case of acceptance no publication is possible.

Please carefully read about how to prepare your manuscript and note the mandatory charge for over-length papers as imposed by IEEE in the website of J-BHI (https://www.embs.org/jbhi/prepare-and-submit-your-manuscript/).

Thank you very much for considering JBHI to publish your research work.

Sincerely,

Prof. Dimitrios I. Fotiadis
Editor-in-Chief

Cc: file

Associate Editor's comments to the authors:
Associate Editor
Comments to the Author:
(There are no comments. Please check to see if comments were included as a file attachment with this e-mail or as an attachment in your Author Center.)

For submitting your revised manuscript, please ensure all changes are clearly highlighted in the manuscript and explained in detail in the rebuttal to facilitate the review process.

Reviewers' comments to the authors:
Reviewer: 1

Comments to the Corresponding Author
Major Concerns

1. Experimental Validation and Efficiency Claims

The paper's central efficiency claim—60% computational reduction via VRWKV—remains entirely unsubstantiated. A core contribution built on efficiency must provide comprehensive benchmarks including FLOPs, memory consumption, and inference latency across varying input sizes. The absence of these fundamental metrics suggests either inadequate experimental rigor or awareness that actual gains may not support the claimed advantages. This gap is particularly damaging given that computational efficiency differentiates DualSeg from existing methods.

2. Dataset Limitations and Clinical Generalizability

The evaluation's restriction to PAS-stained specimens fundamentally limits clinical relevance. Real-world pathology predominantly uses H&E staining with supplementary protocols (Masson's trichrome, Jones silver) for specific diagnoses. Each staining method reveals distinct tissue characteristics—algorithms optimized for PAS often fail catastrophically on H&E due to different contrast patterns and feature visibility. The absence of cross-institutional validation or multi-protocol testing renders clinical applicability claims premature.

3. Incomplete Architectural Comparisons

The baseline selection reveals a critical methodological flaw through systematic omission of contemporary efficient architectures. InceptionNeXt, which achieves transformer-comparable performance with CNN efficiency, directly challenges DualSeg's value proposition yet remains unexamined. Similarly, Mamba-based segmentation models (VM-UNet, U-Mamba) already address the linear complexity challenge DualSeg claims to solve. These omissions appear deliberate rather than oversight, suggesting the authors recognize these comparisons might undermine their architectural superiority claims. Without these essential benchmarks, the true contribution remains indeterminate.

4. Clinical Deployment Analysis

Despite positioning DualSeg for practical application, the manuscript provides no deployment feasibility analysis. Clinical environments impose strict constraints: limited GPU memory (often 8GB), CPU-only workstations in many facilities. The reported 2.73%

mDSC improvement lacks clinical context—pathologists require understanding of whether such margins affect diagnostic confidence, inter-observer agreement, or treatment decisions. Without this translation from metrics to clinical impact, the practical value remains speculative.

5. Architectural Design Justification

The dual-stage encoder's sequential arrangement (Wave-Swin → VRWKV) lacks theoretical foundation. The manuscript presents this as optimal without exploring alternative configurations (parallel processing, reversed ordering, or hybrid fusion strategies). The dynamic window sizing, acknowledged as "manually defined based on empirical observations," reveals methodological weakness—critical design parameters derived through trial-and-error rather than principled analysis. This empirical approach, while sometimes necessary, requires thorough justification absent here.

Minor Concerns

Ablation scope: Component analysis limited to mDSC ignores computational overhead—does each module justify its complexity?
Presentation imbalance: Technical density overshadows clinical motivation, limiting accessibility.
Visualization gaps: Comparison figures lack error maps or uncertainty quantification essential for understanding performance differences.

Reviewer: 2

Comments to the Corresponding Author
This manuscript, entitled “DualSeg: Unified multi-scale framework with dual-stage encoder for glomerular segmentation,” presents a dual-stage hybrid segmentation model that combines convolutional and recurrent attention mechanisms (Wave-Swin and VRWKV) to improve glomerular segmentation performance in kidney histopathology images. The topic is timely and relevant to renal pathology and computational histology. The model demonstrates good quantitative results on both mouse and human datasets.
However, despite these merits, the work still contains several limitations in both methodology and presentation, and a major revision is required before it can be considered for publication.

1) the claimed novelty is not entirely convincing. The proposed framework, while integrating CNN and VRWKV modules, appears conceptually similar to previously published hybrid architectures such as TransUNet, H2Former, and DA-TransUNet. The paper does not provide sufficient theoretical or empirical evidence to show that DualSeg fundamentally differs from these approaches. The authors should clearly articulate the unique contribution of their dual-stage design, beyond incremental improvements or architectural recombination. It would be helpful to include visual or quantitative analysis (e.g., feature map visualization, attention distribution comparison) to demonstrate how the proposed design contributes to performance beyond existing hybrid models.

2) the generalization capability of the model remains insufficiently evaluated. The experiments are limited to PAS-stained mouse and human datasets that share similar imaging conditions. Without cross-stain or cross-center validation, the robustness of the model under real-world variations in staining or scanner parameters cannot be confirmed. The authors should include additional experiments or at least discuss the expected behavior of the model under stain variability. It is also recommended to cite and discuss two closely related works—"Unsupervised stain augmentation enhanced glomerular instance segmentation on pathology images" and "Identifying and matching 12-level multistained glomeruli via deep learning for diagnosis of glomerular diseases"—to better position this study within the current research landscape.

3) the clinical relevance and interpretability of the segmentation results are not adequately discussed. Although the paper emphasizes segmentation accuracy, it does not explain how this improvement translates into clinical or diagnostic benefits, such as better lesion quantification, assessment of glomerulosclerosis, or prediction of disease progression. Including examples or quantitative analyses connecting segmentation quality to potential diagnostic metrics would significantly strengthen the manuscript. A brief discussion of the model's interpretability from a pathologist's perspective would also be valuable.

4) the methodological presentation is unnecessarily complicated. The mathematical description of the Wave-Swin and VRWKV blocks is long and dense, making it difficult for readers from the biomedical community to grasp the main idea. Some mathematical symbols, such as Θ and W^T_jk, are not clearly defined, and equation formatting is occasionally inconsistent. The authors should simplify the explanation of equations and focus on the conceptual understanding of each component's role in the overall framework, leaving detailed derivations to supplementary materials if necessary.

5) the experimental analysis requires stronger statistical and methodological support. Although the authors conduct ablation studies, the reported improvements are small, and the absence of variance analysis or statistical testing makes it unclear whether the

gains are significant. The paper would benefit from a more comprehensive evaluation, including inference time, parameter count, and performance on challenging subtypes such as sclerotic or crescentic glomeruli. Additional qualitative examples demonstrating both the strengths and failure cases of DualSeg would help provide a balanced assessment of the model's robustness.

Finally, several minor issues should be corrected:

1) Figure fonts are too small to read, and color schemes in Figures 3–6 are inconsistent. Figure 5 panels are misaligned, and the arrows indicating regions of interest are too faint.

2) Table I has several misaligned columns and overlapping text.

3) There are typographical errors such as "uqualitative" (should be "qualitative") and repeated line-break hyphenations such as "glomeru- lar," which should be corrected.

4) Ensure consistent capitalization of dataset names (e.g., HuBMAP, KPIs) and verb tense consistency throughout the text. References should be carefully checked for completeness and formatting.