

## Reviewer cv8f

1. Added a dedicated Limitations and Future Work section at the end of the paper. This section addresses the need for full ablation studies of key components such as CBAM and FABlocks, the impact of pretraining, experiments with different backbone architectures, and comparisons with other cited segmentation-guided and anatomy-aware methods.
2. Included Fold Variance Results for all evaluation metrics in Section 4.2 to provide a more comprehensive assessment of model performance stability.
3. Incorporated and discussed the recent relevant work by Jaus et al. in the Introduction to contextualize our approach within current anatomy-pathology segmentation research: *Jaus, A., Seibold, C., Reiß, S., Heine, L., Schily, A., Kim, M., Bahnsen, F.H., Herrmann, K., Stiefelhagen, R., Kleesiek, J.: Anatomy-guided Pathology Segmentation. In: MICCAI 2024, LNCS, vol. 15008, pp. 3–13. Springer, Switzerland (2024).*
4. Improved overall writing clarity throughout the manuscript for enhanced readability and precision.

## Reviewer uxyS

1. Added abbreviations and explanations for key terms and components, including AO/OTA, PXR, CBAM, SGD, Mix Transformer B0, Grad-CAM, and DRR20, to improve clarity for readers unfamiliar with these terms.
2. Improved Grad-CAM visualizations by adjusting the heatmap opacity to better preserve the visibility of the underlying X-ray structures.
3. Expanded the description of the segmentation decoder architecture, providing additional details on the Mix Transformer B0 encoder and citing the code library used for implementation to enhance reproducibility.
4. Clarified and added to the future work section a planned ablation comparing binary masks versus segmented bone inputs to investigate the impact of anatomical detail on classification performance.

## Reviewer gjvx

1. Clarified the segmentation and classification validation process, specifying that 5-fold cross-validation was used for classification and 2-fold cross-validation for segmentation. These were conducted independently, ensuring no data leakage.
2. In Section 4.3, explicitly clarified that all models were trained solely on the Visible (VIS) fracture set and evaluated on both Visible (VIS) and Invisible (INVIS) subsets.
3. Provided clarification on the pretraining datasets used for both the classification and segmentation tasks.
4. Noted that sigmoid activation was applied only at the final layer of the segmentation model; this is now explicitly stated.
5. Abbreviations and mathematical notations have been standardized and corrected throughout the paper for clarity and consistency.

6. Table 1 was repositioned closer to the corresponding discussion in the main text to improve readability and narrative flow.
7. Adopted consistent use of VIS and INVIS terminology for visible and invisible subsets, and reflected this across the entire manuscript.
8. Figure 2 was updated to include skip connections and a visual diagram of the CBAM module. Additionally, figures were re-exported in higher quality for better visual clarity.
9. All results were standardized to XX.XX% format for consistency across tables and text.
10. Repetitive content in Sections 2.1 and 2.2 was removed, and additional detail was added regarding the Mix Transformer B0 and U-Net segmentation architecture.
11. Clarified the augmentation parameters in Section 3.3, especially those previously shown in parentheses, to avoid ambiguity.
12. Corrected metric naming: segmentation F1 Score was renamed to Dice Score, as appropriate.
13. Improved Figure 3 caption for better comprehension.
14. Overall, enhanced the writing clarity and comprehension throughout the paper for improved readability.