**Changes After Review (Paper ID 0013)**

The following changes were made to the paper in order to address the reviews:

* Spelling and grammar errors were corrected.
* Cropped and improved clarity in Figure 1 and 4.
* **Page 1** **–** **Title**:
  + As suggested by reviewer 3, the title was changed for conciseness and clarity.
* **Page 1 – Abstract and Keywords:**
  + As promised before and discussed during the rebuttal phase, we’ve updated the document with a link to our GitHub page containing our code base and further implementation details required for reproducing the results. We’ve also added that the dataset will be available upon request.
  + The abstract is slightly modified for improved clarity, and cut down for paper length constraints.
  + Several keywords were amended.
* **Pages 1-3 – Introduction:**
  + Following the kind suggestions of all reviewers and program chairs, we amended several paragraphs to the introduction, including 12 new references to relevant prior work in the associated research areas to help situate the contribution better. These references include:
    - Baumgartner, C.F., Koch, L.M., Tezcan, M., et al.: PHISeg: Capturing uncertainty in medical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 119–127. Springer, Cham (2019)
    - Chatterjee, S., Honchar, A., Seibold, M., et al.: PULASki: Learning inter-rater variability using statistical distances to improve probabilistic segmentation. Med. Image Anal. 85, 103623 (2025).
    - Fuchs, M., Gonzalez, C., Mukhopadhyay, A.: Practical uncertainty quantification for brain tumor segmentation. In: Medical Imaging with Deep Learning (MIDL) (2021)
    - Isensee, F., Jaeger, P.F., Kohl, S.A.A., et al.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 18, 203–211 (2021)
    - Kendall, A., Gal, Y.: What uncertainties do we need in Bayesian deep learning for computer vision? In: Advances in Neural Information Processing Systems (NeurIPS), vol. 20 (2017)
    - Kohl, S., Romera-Paredes, B., Meyer, C., et al.: A probabilistic U-Net for segmentation of ambiguous images. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 31 (2018)
    - Lekadir, K., Frangi, A.F., Porras, A.R., Glocker, B., Cintas, C., Langlotz, C.P., et al.: FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ 388:e081554 (2025)
    - Mehta, R., Paunovic, V., Arbel, T.: Propagating uncertainty across cascaded medical imaging tasks for improved deep learning inference. IEEE Trans. Med. Imaging 41(11), 3090–3102 (2022)
    - Monteiro, M., Allken, V., Wang, B., Jacob, M.W.: Stochastic segmentation networks: Modelling spatially correlated aleatoric uncertainty. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 12756–12767 (2020)
    - Nair, T., Chen, L., Yang, C., Precup, D.: Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation. Med. Image Anal. 59, 101557 (2020)
    - Salahuddin, Z., Zhang, T., Wang, Y., Salama, M.S.: Leveraging uncertainty estimation for segmentation of kidney, kidney tumor and kidney cysts. In: International Challenge on Kidney and Kidney Tumor Segmentation, pp. 40–46. Springer, Cham (2023)
    - Sudre, C.H., Dalca, A., Baumgartner, C.F.: Uncertainty for safe utilization of machine learning in medical imaging. In: MICCAI UNSURE Workshop (2022)
  + Cut down (for paper length constraints) on trustworthy bullet points but added citation for further clarity.
* **Pages 8-9 – Conclusions:**
  + Following the suggestions by the reviewers and program chairs, we amended to the conclusions a clear explanation on why standard segmentation networks (U-Net and nnU-Net) were not considered.
  + Following Reviewers 1 and 3, we added a few lines on limitations and discussions.
* **Page 9 – Acknowledgements and Disclosure of Interests**:
  + These sections have been added as required.