Response to the referee 1

Thank the referee for taking time to review our manuscript. We really appreciate your time and your review.

Note: The (Page x, Line xx-xx) gives where the explanation or revision locates based on the preprint of our manuscript egusphere-2024-2760.

The revision based on the referee's comments will be uploaded together with the reply to all the reviewers later.

In this manuscript a U-Net model combined with Multi-textRG algorithm has been proposed for fine ice-water classification on SAR imagery from AI4Arctic dataset. Overall the manuscript is clearly written with sufficient details. However, as U-Net and other techniques such as GLCM feature extraction has already been broadly used for sea ice mapping and particularly in AI4Arctic dataset. The manuscript should further emphasize the novelty in terms of methodology in this research, as well as comparison with previous methods as baselines. Detailed comments are listed below.

- **Comment 1:** The authors mentioned in the abstract that the proposed algorithm successfully addresses ice–water classification across all seasons. However, during the evolution of sea ice, the proportion of ice types presented by different seasonal patterns is unstable. In warmer seasons, melting ice surfaces affect the classification results, while in colder seasons, snow cover on sea ice also influences the outcomes. The authors did not evaluate the algorithm's performance under varying environmental conditions, indicating a lack of demonstrated adaptability and effectiveness in different seasonal contexts. This limitation should be clearly emphasized and analyzed.

  Response 1: Yes, we agree with the reviewer's comments about the evolution of sea ice. As for our algorithm's performance under varying environmental conditions, **we can explain it from two respects:**

  1) The SAR scenes provided in the AI4Arctic dataset were selected by the experts from the Technical University of Denmark (DTU), Danish Meteorological Institute (DMI), and Nansen Environmental and Remote Sensing Center (NERSC), considering ice/water type diversity and season balance **(Stokholm et al. 2022)**. In this paper, we used the AI4Arctic dataset in the latest Version 2 released on May 25, 2023. **Figure 1** (a) and (b) in our manuscript respectively show the spatial location and seasonal distribution of the SAR images, i.e., the 512 SAR images have a relative balanced distribution across all seasons with the maximum number in July to September and the minimum number in December to February. Considering the complicate ice conditions within the waters around Greenland and the Canadian Arctic Archipelago, thus, we said the AI4Arctic dataset "features typical classification challenges of ice and water surface characteristics while maintaining a certain feature balance" **(Page 4, Line 105-106)**. please also see the analysis in Section 2.5 **(Page 7, Line 205-213)**.

2) Our algorithm figured out the ice-water classification of these 512 SAR images in high accuracy and high resolution, as the examples shown in **Figure 9, Figure 10, and Figure S2** in the supplement PDF file. Besides, we have made the ice-water classification results of 219 (85+134) SAR images with visual validations to Sentinel-1 and Landsat-8 optical data available at https://zenodo.org/records/13269639, so that the reviewer and other readers can look through the visual result performance.

Except for the focus on method statement, we have added two sentences in the Conclusion section to emphasize the typicality of SAR images provided by the AI4Arctic dataset and the accessible link of our more validation results. The sentences are:

"The 512 Sentinel-1 SAR images provided by the AI4Arctic dataset include abundant ice conditions. The proposed algorithm developed based on this dataset likely addresses the accurate ice-water classification across all seasons."

These words are added before the "Future work …" at **(Page 27, Line 644)**.

- **Comment 2:** The authors stated that the framework is primarily designed for ice-covered regions and wind-driven open water areas. However, wind forcing can also affect the classification accuracy within ice-covered regions. Therefore, the authors should clarify how wind-driven dynamics influence the classification performance across different types of ice-covered areas.

  Response 2: We resolved the identification on most of the wind-driven open water areas from the contiguous thick ice (means large brightness) and densely smooth ice, except for a certain condition, such as the broken/fragmented thin ice and wind-driven open water mixed region. We have given out the explanation in the Discussion section 5.2, referring to **Figure S3** in the supplement PDF file.

- **Comment 3:** In the third paragraph, the authors claimed that the algorithm combining CNN with empirical methods represents the optimal automatic approach for sea ice labeling. However, the evidence supporting this conclusion appears overly strong, as there is insufficient experimental validation to substantiate the claim of optimal performance. Multiple experimental results are needed to support such a conclusion. Furthermore, the proposed algorithm lacks comparative analysis with other sea ice classification methods, whether quantitative or qualitative, which further weakens the assertion of its superiority.

Response 3: Yes, comparative analysis with certain other algorithms is essential. However, due to space limitations, this paper primarily focuses on detailing our methodology and presenting intuitive results, as well as the validations. We acknowledge that our initial phrasing lacked sufficient rigor—without extensive comparisons with other methods, the "optimal" conclusion should be expressed more cautiously.

When referring to our approach as "optimal" in the introduction, we base this claim on theoretical justification of its methodological advantages. That is, the combination of supervised CNN in coarse semantic segmentation and unsupervised empirical method in detailed pixel classification most likely achieve the fine ice-water classification **using coarse labels**, comparing to the collected methods in the third paragraph. This is the subjective conclusion of the review of current research status in the Introduction. We did not claim that our algorithm is the optimal in the Conclusion section.

But now, the technological advancements continually redefine the state-of-the-art. For instance, recent work in **(Zhang et al. 2025)** addresses the same image classification challenges in high-resolution remote sensing images, i.e., **"**Ultra High Resolution (UHR) remote sensing imagery (RSI) (e.g. 10,000 × 10,000 pixels) poses a significant challenge for current Remote Sensing Vision Language Models (RSVLMs). If choose to resize the UHR image to standard input image size, the extensive spatial and contextual information that UHR images contain will be neglected. Otherwise, the original size of these images often exceeds the token limits of standard RSVLMs, making it difficult to process the entire image and capture long-range dependencies to answer the query based on the abundant visual context.**"**, consistent with our review discussion in the fourth paragraph **(Page 27, Line 84-88)**.

Zhang et al. (2025) achieves a balance between high-resolution detail and large-context learning through an advanced Remote Sensing Multimodal Large Language Model (RSMLLM), i.e., the ImageRAG model, enabling purely deep learning-based high-precision image classification. In contrast, our method combines deep learning with traditional classification method. However, the ImageRAG model uses the manually annotated MME-RealWorld dataset as benchmark, whereas our algorithm framework just uses the coarsely labelled ice charts, which significantly enhances its practicality for operational applications.

- **Comment 4:** Line 105, the related publication concerning the AutoIce Challenge should be mentioned here (doi: 10.5194/tc-18-3471-2024), which would facilitate readers to refer to this particular challenge and its details.

  Response 4: Thank you for your suggestion, we will add this reference (Stokholm et al. 2024) and (Stokholm et al. 2022) in the revised manuscript.

- **Comment 5:** Line 120, the U-Net-based model has already been further improved by the AutoIce participants using a bunch of techniques and achieved relatively high accuracy in the AI4Arctic dataset (illustrated in doi: 10.5194/tc-18-3471-2024 and doi: 10.5194/tc-18-1621-2024). According to Fig. 2, it seems that the U-Net used in this research has the same architecture as the one used in the challenge. Therefore, it is necessary to illustrate how the U-Net-based method proposed in this research different from the previous ones. It is also necessary to implement those U-Net-based models as benchmarks to compare with the proposed method in the manuscript.

  Response 5: Yes, we employed the same U-Net architecture provided by the AI4Arctic competition. However, our goal is the ice-water classification. Based on coarse SIC labels in ice charts, the original U-Net model yields SIC classification results with rough ice boundaries as depicted in **Figure 7**. Our algorithm further achieves refined identification of ice pixels, as illustrated in **Figure 9**. **Figure S2** presents five additional example images for reference.

  Although the U-Net model remains unchanged, the data preprocessing methodology applied in this study was optimized through extensive experimental validation. Notably, the input SAR images have significant thermal noise, to which the U-Net performance is highly sensitive, as demonstrated in previous research **(Stokholm et al. 2022)**. In the initial manuscript submitted to the TC AI4Arctic special issue, we utilized the denoised and normalized Sentinel-1 SAR data provided by the AI4Arctic competition. However, the classification results still have a higher false-negative (FN) ratio for wetted ice surfaces. In this manuscript, we also conducted substantial sensitivity investigations but not presented and we reprocessed raw Sentinel-1 SAR images by implementing an alternative denoising algorithm combining that proposed by (Sun and Li 2021) and by (Park et al. 2020).

Consequently, the key distinctions between our algorithm and the original U-Net model are: 1) different image preprocessing methods leading to different feature inputs; and 2) superior spatial resolution of the output in our results.

References

[1] Park, J.-W., Korosov, A.A., Babiker, M., Won, J.-S., Hansen, M.W., & Kim, H.-C. (2020). Classification of sea ice types in Sentinel-1 synthetic aperture radar images. The Cryosphere, 14, 2629-2645

[2] Stokholm, A., Buus-Hinkler, J., Wulf, T., Korosov, A., Saldo, R., Pedersen, L.T., Arthurs, D., Dragan, I., Modica, I., Pedro, J., Debien, A., Chen, X., Patel, M., Cantu, F.J.P., Turnes, J.N., Park, J., Xu, L., Scott, K.A., Clausi, D.A., Fang, Y., Jiang, M., Taleghanidoozdoozan, S., Brubacher, N.C., Soleymani, A., Gousseau, Z., Smaczny, M., Kowalski, P., Komorowski, J., Rijlaarsdam, D., van Rijn, J.N., Jakobsen, J., Rogers, M.S.J., Hughes, N., Zagon, T., Solberg, R., Longépé, N., & Kreiner, M.B. (2024). The AutoICE Challenge. The Cryosphere, 18, 3471-3494

[3] Stokholm, A., Wulf, T., Kucik, A., Saldo, R., Buus-Hinkler, J., & Hvidegaard, S.M. (2022). AI4SeaIce: Toward Solving Ambiguous SAR Textures in Convolutional Neural Networks for Automatic Sea Ice Concentration Charting. Ieee Transactions On Geoscience and Remote Sensing, 60, 1-13

[4] Sun, Y., & Li, X.-M. (2021). Denoising Sentinel-1 Extra-Wide Mode Cross-Polarization Images Over Sea Ice. Ieee Transactions On Geoscience and Remote Sensing, 59, 2116-2131

[5] Zhang, Z., Shen, H., Zhao, T., Guan, Z., Chen, B., Wang, Y., Jia, X., Cai, Y., Shang, Y., & Yin, J. (2025). Enhancing Ultra High Resolution Remote Sensing Imagery Analysis with ImageRAG. IEEE Geoscience and Remote Sensing Magazine, 14