Thank you very much for taking the time to review this manuscript. I have responded to each of the issues raised by the two reviewers and have revised the manuscript.

Response to Reviewer 2 Comments

Comments 1- Please provide detailed information for different branch in ResSepNet in Section 3.3.1. Has results in Figure 9 been statistically tested? Please include more information on this.

Response: We sincerely thank the reviewer for the valuable comments regarding the structural details and experimental rigor of ResSepNet. In response to your concerns, we have added a detailed description of the design principles and functional roles of each branch in Section 3.3.1, "ResSepNet Branch Ablation Experiment" in the revised manuscript. Specifically, the background branch is designed to extract environmental features and suppress background noise, the icing branch focuses on capturing detailed and textural features of the icing regions on transmission lines, and the global branch employs EfficientNet-B3 to extract global contextual information from the entire image. These multi-scale features are then normalized and fused to enhance the discriminative capability for ice type recognition. In addition, to address the statistical reliability of the results shown in Figure 9, we clarified that the reported results are the average of five independent experimental runs conducted under identical training and testing datasets and hardware environments, ensuring good comparability and consistency. The small variation among repeated experiments further demonstrates the high stability of the proposed model.

Specific modifications are as follows: **(Page 14; Page 15)**

3.3.1 ResSepNet Branch Ablation Experiment

3.3.1 ResSepNet Branch Ablation Experiment

In ResSepNet, the ice type recognition task is achieved through the collaborative efforts of the background branch, icing branch, and global branch. The background branch preprocesses the original image to extract the background subgraph, focusing on capturing environmental features in the image while minimizing the interference of background noise. The icing branch, on the other hand, specializes in extracting features of the icing on the transmission line. It uses a structure similar to that of the background branch but places greater emphasis on capturing the details and texture information of the icing area. The global branch directly inputs the entire image, utilizing EfficientNet-B3 as the backbone network to extract macro features from the full image and capture global context information through a transfer learning model. The design of these three branches aims to capture both local and global information at different spatial scales by performing feature extraction in different regions, thereby effectively reducing the influence of background noise and improving the accuracy of ice type recognition. By normalizing and fusing the features extracted from the different branches, ResSepNet can fully leverage the spatial scale information extracted by each branch, ultimately leading to more accurate ice type recognition results.

Figure 9. Confusion matrix of ice type recognition effect of each branch of ResSepNet.

For the results in Figure 9, five independent experimental runs were conducted and the average values were reported to ensure data stability and reliability. All runs used the same training and test datasets and were performed under identical hardware conditions to guarantee consistency. Thus, the confusion matrix in Figure 9 represents the averaged results with minimal variation, indicating high model stability across repeated experiments. Based on the confusion

Comments 2- The paper currently lacks a dedicated "Discussion" section, which limits the depth and completeness of the research. It is recommended that the authors add this section, focusing on the following three key aspects: Model Advantages, Mechanism Interpretation and Model Limitations.

Response: We sincerely thank the reviewer for pointing out the absence of a dedicated Discussion section in the original manuscript. Following your suggestion, we have newly added and systematically constructed Section 4 Discussion in the revised version, where the proposed method is analyzed in depth from three key aspects: model advantages, mechanism interpretation, and model limitations. Regarding model advantages, we discuss the overall strengths of DTL-IceNet in ice type recognition, ice region segmentation, and ice thickness estimation from the perspectives of multi-task collaboration and multi-source information fusion, highlighting its improvements over existing methods. With respect to mechanism interpretation, we analyze the intrinsic reasons for the strong performance of the model along three main lines, namely multi-branch feature extraction, multi-scale attention mechanisms, and the collaborative modeling of geometric and meteorological constraints. In terms of model limitations, we further discuss the fact that the current dataset does not yet cover extreme weather or complex terrain conditions, that thickness validation is only based on clear-sky observation-site data, and that the dual-view structure has not been fully validated under real side-view imaging conditions, and we objectively analyze their potential impacts on model generalization. In addition, we combine different imaging conditions and meteorological factors to investigate, at the mechanism level, the possible error sources of the model in complex environments. The Discussion section also explicitly outlines future research directions involving extreme weather scenarios, multi-terrain environments, and real dual-view experimental platforms. Overall, the newly added Discussion provides a comprehensive and critical reflection on this work from the perspectives of advantages, mechanisms, limitations, and future developments.

Specific modifications are as follows: **(Page 23-25)**

4 Discussion

4 Discussion

The dual-task learning framework, DTL-IceNet, proposed in this study demonstrates high reliability in both ice type recognition and thickness detection tasks. Its main contribution lies in the unified modeling of three types of information: type, geometry, and meteorology, which effectively enhances the comprehensive sensing ability of icing conditions on transmission lines. This fusion-based design aligns with the view emphasized in the literature that "ice physics, image features, and environmental processes must be considered in a coordinated manner" (Fan et al., 2018; Hao et al., 2023; Chen et al., 2024; Dong et al., 2022; Meng et al., 2025; Han et al., 2024), and it achieves both discriminative ability and physical consistency in typical monitoring scenarios, leading to significant improvements over existing methods.

In the ice type recognition task, the model explicitly separates the background, conductor, and icing regions through a multi-branch feature extraction structure. Studies by Fan et al. (2018). and Hao et al. (2023) have pointed out that ice recognition is highly sensitive to the environmental background, and deep networks based on a single-path feature extraction often struggle to fully capture the local texture of the conductor in complex backgrounds. The decomposition-based modeling approach of DTL-IceNet significantly enhances the distinction between different types of icing under complex lighting, fog, and noise conditions, providing stronger anti-interference ability compared to single-branch methods.

In the ice region segmentation task, MOMSA-SegNet leverages a multi-scale attention module to improve the representation capability of thin conductors and irregularly shaped icing areas. Existing research has shown that traditional edge detection or low-level feature methods exhibit poor robustness in weak boundaries, low contrast, and nighttime scenarios (Han et al., 2024; Tan & Le, 2019; Hu et al., 2018; Vasu et al., 2023; Li et al., 2016). In contrast, the multi-scale attention mechanism effectively utilizes the contextual structure surrounding the transmission line, allowing the model to maintain stable geometric contour predictions in typical scenes such as sunny, foggy, and nighttime conditions. Its accuracy advantage stems from the targeted utilization of the transmission line image structure rather than relying solely on the depth of the network or the scale of parameters.

For ice thickness estimation, the improvement in model performance is mainly attributed to the synergistic effect of geometric and meteorological constraints. The image geometric information provides the basic trend of ice volume changes, but relying solely on 2D images cannot accurately reflect the true 3D shape of the ice, leading to systematic biases during temperature, humidity, wind speed, and precipitation phases. By introducing the meteorological correction term based on environmental factors, the model performs consistent corrections on the initial thickness estimation according to the basic physical laws of ice growth and melting, effectively compensating for the inherent structural biases in the geometric estimation. Experimental results show that the thickness curves align more closely with the actual distribution across multiple phases, indicating the complementary role of geometric and meteorological information in thickness estimation.

Although DTL-IceNet demonstrates robustness under typical monitoring conditions, its applicability is still limited by the data coverage and experimental conditions. The data used in this study were primarily collected in typical meteorological scenarios and have not yet covered extreme weather conditions such as severe convection or freezing rain coupled with strong winds. Under such conditions, the signal-to-noise ratio of images and the rate of ice morphology change may exceed the training distribution, and the model's robustness needs further validation. Moreover, the current experimental data do not systematically reflect complex spatial environments with significant topographic variations. Existing studies have shown that terrain, especially in valley wind fields, significantly impacts icing distribution, so the model's performance in complex terrain scenarios remains uncertain.

While the segmentation experiments presented in this study systematically show performance differences under various imaging conditions such as sunny days, fog, and nighttime, the thickness validation experiments were limited by the actual conditions of the observation field, which only included real thickness calibration data from sunny scenarios and could not directly quantify thickness estimation accuracy under various weather conditions. To address this logical gap, we will further analyze the potential impact of weather factors on thickness estimation. Mechanistically, weather changes affect thickness results mainly through two paths. First, imaging degradation such as fog and low light reduces the clarity of the ice and conductor boundaries, causing segmentation masks to deviate in geometric details and directly affecting the preliminary thickness calculation based on area and contour inference. Second, meteorological variables such as temperature, humidity, and wind speed determine the growth and melting rates of the ice, strongly driving the temporal evolution of thickness. The meteorological correction module of DTL-IceNet can provide systematic corrections based on this driving pattern, but it cannot completely compensate for geometric deviations caused by severe imaging degradation. Therefore, under real fog or nighttime conditions, the initial thickness errors may increase, while the optimized thickness is expected to show more stable but still limited corrections.

The dual-perspective approach proposed in this study is one of the key innovations of the overall framework. Through joint segmentation of the main and side perspectives, the major and minor axes of the ice-covered cross-section can be theoretically estimated, improving the certainty of geometric parameters. However, it is important to note that the thickness validation experiment in this paper was limited by site constraints and did not actually deploy transmission lines for real side-perspective imaging. Thickness detection relied on a simplified geometric assumption primarily based on the main perspective. This experimental condition means that the final obtained thickness accuracy mainly reflects the performance under the simplified single-perspective system rather than the upper limit of a complete dual-perspective structure. Therefore, while the current results prove the feasibility and potential value of the proposed framework, they do not fully validate the information gain of the dual-perspective structure in real multi-conductor scenarios. This also suggests that the side-perspective segmentation error's impact on thickness calculation has not been fully quantified. Future work will focus on building an experimental platform that truly reflects the dual-perspective structure to systematically evaluate error propagation mechanisms and further optimize geometric fusion methods.

In summary, DTL-IceNet forms a cohesive solution across ice type recognition, geometric structure extraction, and thickness estimation, with its advantages derived from complementary constraints between tasks and explicit incorporation of physical processes.

Comments 3- The connection between the high mIoU of MOMSA-SegNet and the final thickness MAPE is not sufficiently demonstrated. To validate the overall framework, it is essential to provide evidence that the superior segmentation performance is a prerequisite for the low thickness error. Ablation studies comparing the impact of different segmentation qualities on the final MAPE are needed.

Response: We sincerely thank the reviewer for the professional and highly targeted suggestions. In response to your comments, we have added a new Section 3.5.2, entitled “Effect of Different Segmentation Models on Ice Thickness Detection,” in the revised manuscript. Under strictly identical ice-type recognition and thickness optimization procedures, UNet++, SegNet, DySample, and the proposed MOMSA-SegNet are respectively employed as the front-end segmentation module, and comparative thickness estimation experiments are conducted on the same observation-site test dataset. The impacts of different segmentation qualities on both the initial and optimized thickness errors are systematically evaluated. The experimental results (Table 5) demonstrate that better segmentation performance consistently leads to lower MAPE and MSE for both the initial and optimized thickness estimates, and the thickness estimation based on MOMSA-SegNet achieves the best performance across all evaluation metrics. Moreover, even after the introduction of meteorological correction, the error differences among different segmentation models remain evident, indicating that the geometric structural errors introduced at the segmentation stage cannot be fully eliminated by the back-end meteorological correction. These results directly verify, under unified experimental conditions, that high segmentation accuracy is a necessary prerequisite for achieving low thickness estimation error, and quantitatively reveal the decisive influence of segmentation quality on the final thickness estimation performance.

Specific modifications are as follows: **(Page 21-22)**

3.5.2 Effect of Different Segmentation Models on Ice Thickness Detection

3.5.2 Effect of Different Segmentation Models on Ice Thickness Detection

To quantitatively evaluate the impact of the segmentation module on the accuracy of final ice thickness estimation, this study, under the premise of maintaining consistency in the ice type recognition and thickness optimization calculation processes, uses UNet++, SegNet, DySample, and MOMSA-SegNet as the frontend segmentation models. The thickness estimation experiments were conducted on the same observation field test dataset. The ice type recognition module and thickness optimization calculation module were kept unchanged, and only the ice segmentation submodule was replaced with different typical methods. The initial thickness and optimized thickness errors were then computed on the corresponding observation field thickness test data. This approach allows a direct comparison of the effects of different segmentation models on geometric scale calculation and error propagation within a unified framework, providing a clearer insight into the structural relationship between segmentation quality and thickness detection accuracy. The experimental results are shown in Table 5.

Table 5. Comparison of ice thickness detection performance driven by different segmentation models

Modules	Initial ice thickness MAPE(%)	Optimized ice thickness MAPE(%)	Initial ice thickness MSE	Optimized ice thickness MSE
UNet++	69.53	14.58	80.93	2.26
SegNet	65.38	13.71	76.09	2.12
DySample	57.45	12.05	66.86	1.87
MOMSA-SegNet(Ours)	**56.36**	**11.82**	**65.60**	**1.83**

As seen in Table 5, under the same ice type recognition and meteorological correction processes, the performance of ice thickness detection shows a consistent trend with different segmentation models as frontend submodules. The better the segmentation performance, the lower the initial and optimized thickness errors. When MOMSA-SegNet, the segmentation model proposed in this paper, is used, both the MAPE and MSE of the initial thickness are the lowest. After replacing it with UNet++, SegNet, or DySample, both errors increase to varying degrees. This indicates that the segmentation stage directly affects the accuracy of the ice contour and area depiction, which in turn influences the downstream geometric parameter estimation. Segmentation errors accumulate and amplify in the calculation of equivalent ice thickness.

A further comparison of the initial and optimized thickness metrics reveals that after the introduction of meteorological correction, the MAPE and MSE for all models significantly decrease, showing that the environmental-driven correction terms can effectively compensate for the system errors caused by the 2D perspective and geometric simplifications. However, the relative differences between the models still remain even after optimization. Even with meteorological correction, the thickness estimation based on MOMSA-SegNet maintains the lowest MAPE and MSE, while segmentation models with weaker performance still exhibit higher optimization errors. This suggests that the meteorological correction mainly targets global system biases related to environmental processes, and cannot fully counteract the structural errors in the ice region contour and scale caused by segmentation. Thus, it can be concluded that segmentation quality determines the geometric baseline for thickness estimation, while meteorological correction fine-tunes this baseline, forming a hierarchical complementary relationship.

In conclusion, under the condition of complete consistency in the thickness estimation algorithm and test dataset, simply replacing the segmentation module results in a monotonous or nearly monotonous decrease in thickness errors with the improvement of segmentation model performance. This result confirms the substantial contribution of MOMSA-SegNet in ice thickness detection from a data-driven perspective. Its higher segmentation accuracy not only reflects in pixel-level metrics but also significantly reduces the initial errors in downstream thickness estimation, maintaining its advantage even after meteorological correction and effectively transmitting the improved segmentation performance to the final physical quantity estimation results.