# TARGET DRIVEN ADAPTIVE LOSS FOR INFRARED SMALL TARGET DETECTION

## Supplementary Material

**Table 3**: We compared our proposed loss and existing losses with the metrics of IoU (%), $P_d(\%)$, and $F_a(10^{-6})$. The best and second values are highlighted in bold and underline, respectively.

| Method | U-Net [23] | | | UIUNet [11] | | | SCTransNet [13] | | |
|---|---|---|---|---|---|---|---|---|---|
| | IoU↑ | $P_d$↑ | $F_a$↓ | IoU↑ | $P_d$↑ | $F_a$↓ | IoU↑ | $P_d$↑ | $F_a$↓ |
| BCE loss [2] | 64.76 | 88.65 | **10.32** | 64.56 | 88.77 | 10.38 | 67.03 | 89.00 | **6.98** |
| IoU loss [17] | 63.14 | 91.40 | <u>19.58</u> | <u>63.70</u> | 88.43 | 22.88 | 67.38 | 88.60 | 18.97 |
| Dice loss [18] | 62.19 | 91.75 | 12.37 | 62.27 | 90.81 | 22.64 | 67.72 | <u>89.69</u> | 16.64 |
| SLS loss [9] | <u>66.24</u> | 92.43 | 38.10 | <u>65.20</u> | <u>90.81</u> | 25.02 | <u>68.22</u> | 88.65 | 21.18 |
| SLS loss + TDA loss (Ours) | **69.48** | **94.84** | 14.04 | **66.96** | **92.17** | **9.56** | **68.59** | **91.75** | <u>9.41</u> |

The proposed TDA loss can be applied to prior segmentation models developed for IRSTD and enhance their detection performance. To validate this claim, we compared our loss and existing losses using three different networks. We chose U-Net, UIUNet, and SCTransNet as deep-learning-based models [11, 13, 23]. U-Net has a simple encoder and decoder architecture, and it is able to capture both global context and fine-grained features. UIUNet consists of a U-Net embedded with smaller U-Nets. UIUNet effectively extracts multi-scale information using a nested architecture. SCTransNet introduces spatial-channel cross transformer blocks. SCTransNet encodes global context information and improves detection performance for targets with high similarity to the background.

We used Adam as optimizer for three models. We followed the optimization setting used in the open-source Github code (`https://github.com/XinyiYing/BasicIRSTD`) for training U-Net. We set the initial learning rate, the batch size and number of epochs to 0.0005, 16, 400, respectively. We multiplied the learning rate by 0.5 at 200 and 300 epochs. We followed the optimization settings used in the UIUNet paper for training UIUNet. We set the initial learning rate, the batch size and number of epochs to 0.001, 3, 500, respectively. We followed the optimization settings used in the SCTransNet paper for training SCTransNet. We set the initial learning rate, the batch size and number of epochs to 0.001, 16, 1000, respectively. We decreased the learning rate to 0.00001 by cosine annealing strategy.

Table 3 shows the detection performance of our TDA loss and existing losses with three different architectures: UNet, UIUNet, and SCTransNet. We used IRSTD-1k dataset for training and evaluation. The results show that our TDA loss enhances the detection performance for three different networks. These results indicate that the proposed method enables the model to extract target shapes more accurately and improves detection performance for more challenging targets.