# Enhancing Lightweight IRSR Models via Knowledge Distillation with Structural and Spectral Losses

Seungsoo Han
*Intelligence Research, AI Architecture Team*
*Funzin*
Seoul, Republic of Korea
ethan.han@funzin.co.kr

Dongnyeok Choi
*Intelligence Research, AI Architecture Team*
*Funzin*
Seoul, Republic of Korea
luke.dn.choi@funzin.co.kr

Deukhwa Kim, Jeonghun Kim, Seunghoon Shin
*Intelligence Research*
*Funzin*
Seoul, Republic of Korea
{shirly, marc, kyle.sh.shin}@funzin.co.kr

***Abstract*—For Infrared Image Super-Resolution (IRSR) technology, maintaining performance while reducing model complexity is critical for a wide range of applications. However, existing research on lightweight IRSR has been predominantly limited to modifying model architectures. This study proposes a new methodology that applies Knowledge Distillation, a representative model compression technique from supervised learning, to IRSR models. To this end, we extend the DCKD framework, previously used for RGB image super-resolution, to the IRSR domain and introduce new loss functions designed to maximize the preservation of key structural characteristics in infrared images, namely edge and spectral(Contourlet-domain) information. Through the proposed methodology, a lightweight student model trained with distilled knowledge from a high-performance, complex teacher model consistently achieved superior performance compared to the same architecture trained via standard supervised learning. This study demonstrates that Knowledge Distillation based methodology is effective for developing lightweight IRSR models and is expected to contribute to fields where high-efficiency IRSR is essential, such as real-time military surveillance, disaster response, and nocturnal reconnaissance.**

***Index Terms*—Infrared Image Super Resolution, Super Resolution, Knowledge Distillation**

## I. Introduction

Super-Resolution (SR) is a fundamental task in computer vision focused on reconstructing high-resolution imagery from low-resolution counterparts. By restoring fine-grained details and enhancing visual quality, SR technology has become essential to diverse range of applications, including video restoration, modern surveillance systems, satellite imaging, and medical diagnostics [19]. The field has undergone a significant paradigm shift. Early methods, such as bicubic interpolation, were limited by their inability to recover high-frequency components, often yielding blurry and unsatisfactory results. The advent of deep learning, particularly Convolutional Neural Networks (CNNs) [2], has largely resolved these issues. The subsequent introduction of Generative Adversarial Networks (GANs) [3] in particular dramatically improving restoration quality. In recent years, advanced architectures such as Swin Transformer [4] and Diffusion Models [5] continue to push the boundaries of SR performance.

However, the unique characteristics of infrared images present distinct challenges for SR algorithms. Compared to visible-spectrum imagery, infrared images have less structural details, indistinct edges, and lower contrast [1]. While a number of Infrared Image Super-Resolution (IRSR) models have been proposed to address these issues [8]–[10], [12], research into model lightweighting—a critical factor for deployment in real-time surveillance and on mobile devices—remains relatively insufficient. A few studies have attempted to design lightweight IRSR architectures directly [11], [23], but a critical research gap exists: there is a lack of advanced training methodologies specifically designed for compressing and optimizing these models.

To bridge this gap, we explore Knowledge Distillation (KD)—an effective model compression paradigm where a large, high-capacity "teacher" model transfers its learned knowledge to a smaller, more efficient "student" model. This process enables substantial reductions in computational overhead while preserving model performance. However, prior research on KD for SR has predominantly focused on RGB images, failing to account for the unique statistical properties of the infrared domain.

This paper proposes a novel KD framework specifically engineered for IRSR. Our goal is to enable the effective restoration of high-quality IR images even in computationally constrained environments. To this end, we introduce two specialized loss functions—Sobel Loss and Spectral Loss—that are designed to preserve the edge and spectral information most critical to infrared image structure. This demonstrates

that the effectiveness of the KD methodology in IRSR can be consistently improves solely through the design of loss functions, without the addition of complex modules.

Our key contributions are as follows:

1) Proposing a new KD-based learning framework for IRSR that considers the unique characteristics of infrared images.
2) Optimize the learning of structural features specific to infrared images by incorporating Sobel Loss and Spectral Loss into the distillation process.
3) Experimentally validate a scalable training methodology for model lightweighting, significantly enhancing the practical applicability of IRSR models.

Through this study, we aim to establish a robust foundation for practical, efficient, and deployable IRSR solutions, particularly suited for critical surveillance applications such as thermal monitoring, emergency response systems, and nighttime reconnaissance.

## II. Related Works

### A. Infrared Image Super Resolution

In recent years, a variety of IRSR models have been proposed. Early studies primarily focused on restoring realistic textures using GANs [7]. For instance, [8] incorporated heterogeneous convolution into a WGAN-based framework to effectively model the complex spatial patterns of infrared imagery. PSRGAN [9] designed its architecture using a knowledge-distillation-based block structure and trained it within a GAN framework, achieving high performance even with limited training data. These studies have substantially contributed to the perceptual realism of reconstructed IR images.

Following the advent of GANs, newer architectures have emerged to capture broader and more complex image features. LKFormer [10], a Transformer-based model, effectively captures long-range dependencies through self-attention mechanisms utilizing large kernels. Meanwhile, signal processing-inspired approaches have also gained attention. CRG [12], for example, leverages the contourlet transform to explicitly model the directional and geometric structures of infrared images, significantly improving contour edge restoration.

Alongside performance improvements, research has also progressed toward lightweight model design for practical deployment. Lite-MTSR [23] combines a lightweight network with meta-transfer learning, enabling fast adaptation and real time super resolution under limited data conditions. Similarly, LISN [11] leverages efficient operations such as channel splitting and shift blocks to strike a balance between restoration quality and runtime efficiency.

While existing works primarily focus on developing novel architectures or modifying model structures for performance gains and model compression, our study takes a different approach. Instead of altering the architecture, we focus on improving the training methodology itself through KD, offering an alternative pathway to lightweight IRSR without sacrificing existing design efficiency.

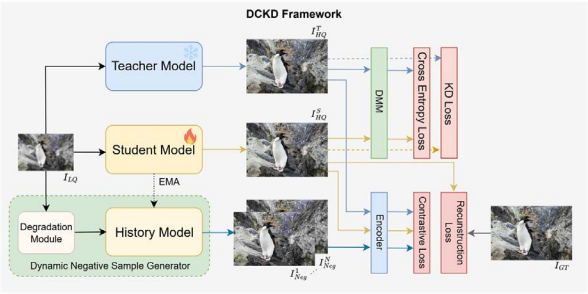



Fig. 1. Illustration of the DCKD Framework. We use this framework as a baseline.

### B. Knowledge Distillation in Super Resolution

KD is a learning framework that transfers the knowledge of a large and complex teacher model to a smaller and more lightweight student model. This enables the student model to achieve high performance comparable to the teacher while maintaining computational efficiency, making it particularly effective for applications requiring real-time processing or deployment on embedded devices. First introduced by [6] in the context of image classification tasks, KD has since been successfully extended to the field of image SR.

Early KD approaches in SR [13] primarily focused on mimicking the feature maps of CNNs. Subsequent studies have evolved to deliver more refined and informative knowledge. For instance, FAKD [14] effectively conveys structural information by leveraging feature affinity between representations, while MiPKD [15] enhances the student's representational capacity by integrating multi-scale priors. Moreover, Data Upcycling KD [16] incorporates data augmentation strategies into the distillation process, offering improved SR performance in data-scarce environments.

More recently, DCKD [17] introduced a novel approach that applies dynamic contrastive regularization in the feature space. This method is architecture-agnostic and demonstrates robust performance across various SR models, suggesting a new level of generalizability in knowledge distillation research.

## III. Method

### A. Preliminaries

DCKD [17] is an advanced knowledge distillation technique proposed to solve the information loss and generalization degradation problems caused by the singular focus of existing methods on reducing feature map differences between teacher and student models. The core of this method is the introduction of Contrastive Learning principles to guide the student model toward learning a more robust and distinct feature space, rather than merely mimicking the teacher. To this end, DCKD employs a sophisticated combination of the following four loss functions. The framework of DCKD is illustrated in Fig. 1

*a) Reconstruction Loss:* This loss enforces accurate image reconstruction, which is the primary objective in SR tasks. It typically uses a pixel-wise L1 loss to minimize

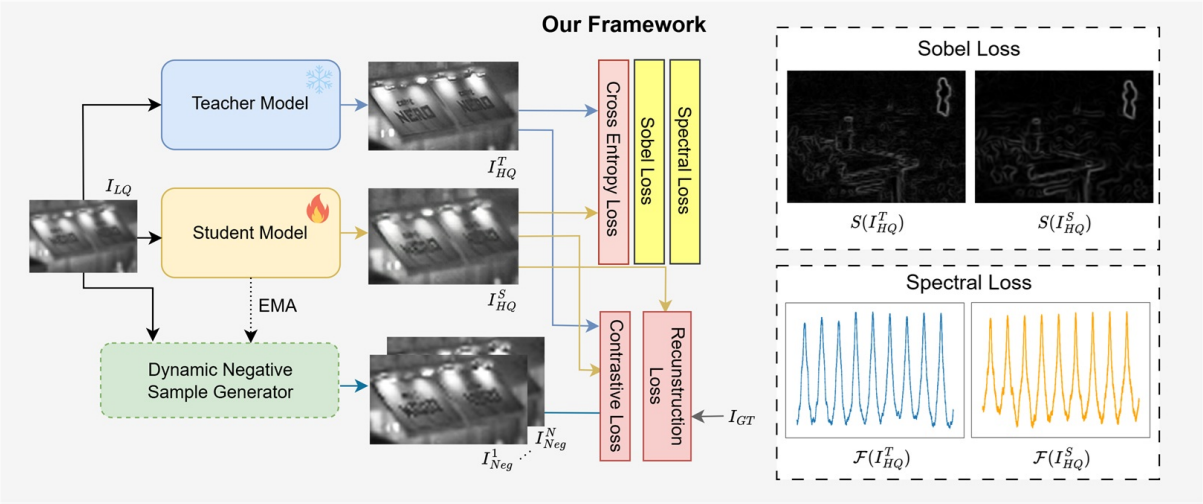



Fig. 2. Illustration of the DCKD Frame work and our losses. The newly proposed loss functions, filled with yellow, are computed between the HR output of the teacher model($I^T_{HQ}$) and student model($I^S_{HQ}$). The right part of figure describe how loss functions work. The Sobel loss is computed based on edge information extracted from the image, and the spectral loss is computed according to the discrepancy between the spectral fidelity between student and teacher model outputs.

the difference between the image generated by the student model($I_{HQ}$) and the ground-truth high resolution image($I_{GT}$).

$$\mathcal{L}_{rec} = \|I_{HQ} - I_{GT}\|_1 \tag{1}$$

*b) Knowledge Distillation Loss:* A fundamental component of conventional KD involves guiding the student model to closely mimic the final outputs of the teacher model. Let $F_S(I_{LQ};\theta_s)$ and $F_T(I_{LQ};\theta_t)$ denote the outputs of the student and teacher models for a low-quality input image $I_{LQ}$, respectively. Through this process, the student is trained to align $F_S$ with $F_T$, effectively inheriting the teacher's capability for generating high-quality outputs.

$$\mathcal{L}_{kd} = \|F_S(I_{LQ};\theta_s) - F_T(I_{LQ};\theta_t)\|_1 \tag{2}$$

*c) Cross-entropy Loss:* This loss function incorporates pixel-wise distributional information that is typically overlooked by conventional L1 loss, enabling more refined pixel-level restoration. To this end, DCKD extracts feature maps ($F^T$, $F^S$)from the teacher and student output images using a pre-trained encoder within a Distribution Mapping Module (DMM). It then calculates the cross-entropy loss between pixel-wise category distributions ($C^T$, $C^S$), which are generated from these feature maps via a VQGAN codebook ($e$). By adopting pixel-level distributions instead of image-level ones, the loss function is better aligned with the requirements of pixel-level image restoration tasks such as SR.

$$\mathcal{L}_{ce} = -\sum_{i=1}^{M} C_i^T \log C_i^S \tag{3}$$

*d) Dynamic Contrastive Loss:* As the central element of DCKD, this component guides effective feature learning through a dynamic contrastive regularization method. In this process, a 'history model,' which stores a previous state of the student model, functions as a Dynamic Negative Sample Generator, dynamically creating hard negative samples during training. The contrastive loss, therefore, compels the student's features($f_{Anc}$) to be attracted to the teacher's features(positive sample, $f_{Pos}$) and repelled from the Dynamic Lower bound(negative sample,$f_{Neg}$) generated by the history model. Through this process, the student model constructs a more distinct and robust feature space.

$$\mathcal{L}_{dcl} = \sum_{i=1}^{L} \lambda_i \frac{\|f^i_{\text{Anc}} - f^i_{\text{Pos}}\|_1}{\sum_{j=1}^{N} \|f^i_{\text{Anc}} - f^{i,j}_{\text{Neg}}\|_1} \tag{4}$$

## B. Proposed Loss Functions

While DCKD has demonstrated excellent performance in RGB-based SR tasks, its direct application to IRSR faces two primary limitations:

- The inherent domain gap between the structural and textural characteristics of infrared and RGB images.
- The challenge of learning discriminative features with standard distillation losses alone, given the scarcity of infrared training data.

To address these limitations, this study employs two auxiliary loss functions into the existing DCKD framework shown as Fig. 2 Each loss functions are designed to reflect the unique properties of infrared images. The proposed loss functions guide the student model to effectively learn the edge structures and frequency characteristics—core aspects of IRSR—from the teacher model.

*a) Sobel Loss:* The sobel filter is a classic operator widely used for edge detection in images. This study introduces the Sobel loss to enable the student model to effectively learn the edge structure restoration capability from the teacher model. This loss calculates the L1 loss between the edge maps

TABLE I
THE SPECIFICATIONS OF TEACHER AND STUDENT MODELS USED FOR EXPERIMENTS.

| Pair (x2) | Model | Role | Block | #Params |
|---|---|---|---|---|
| #1 | CRG | Teacher | 18 | 755K |
| | PSRGAN | Student | 16 | 423.5K |
| #2 | LISN (6b) | Teacher | 6 | 275K |
| | LISN (4b) | Student | 4 | 199K |

obtained by applying the Sobel filter to the images generated by the teacher and student models, respectively. The formula is as follows:

$$\mathcal{L}_{sobel} = \|S(I^S_{HQ}) - S(I^T_{HQ})\|_1 \tag{5}$$

$S(\cdot)$ denotes the edge map obtained through Sobel filtering [24] while $I^S_{HQ}$ and $I^T_{HQ}$ represent the high-resolution outputs of the student and teacher models, respectively. By enforcing alignment in edge representations, this loss encourages the student model to preserve object boundaries and structural integrity, which are often vulnerable during the super-resolution process.

*b) Spectral Loss:* The spectral loss guides the student model to mimic the frequency-domain characteristics of the teacher model. This method was proposed for IRSR in CRG [12] and is effective at preserving high-frequency details by distilling knowledge about the unique energy distribution of infrared images. The loss function is as follows:

$$\mathcal{L}_{spec} = \|\mathcal{F}(I^S_{HQ}) - \mathcal{F}(I^T_{HQ})\|_1 \tag{6}$$

$\mathcal{F}(\cdot)$ denotes the amplitude spectrum obtained via a 2D Fourier Transform. By maintaining spectral fidelity, this loss contributes to more accurate and visually consistent reconstruction results, particularly in preserving the distinctive frequency components of infrared imagery.

*c) Overall Loss:* Finally, the overall loss function in this study is constructed by combining the core elements of the original DCKD with the two proposed auxiliary loss functions. We have replaced the vanilla KD loss (2) of the original DCKD with the sobel loss (5) and spectral loss (6), which are more suitable for infrared characteristics. The overall loss function (7) used for training is as follows:

$$\mathcal{L} = \mathcal{L}_{rec} + \mathcal{L}_{dcl} + \mathcal{L}_{ce} + \lambda_{sobel}\mathcal{L}_{sobel} + \lambda_{spec}\mathcal{L}_{spec} \tag{7}$$

$\lambda_{sobel}$ and $\lambda_{spec}$ are weighting parameters that control the importance of each loss. This configuration optimizes the framework's performance on the IRSR task by retaining DCKD's core contrastive learning strategy while tailoring the target of the distillation to the key characteristics of infrared images.

### C. Implementation Details

The training procedure of our study largely follows the original DCKD [17] framework. The Distribution Module for the distribution matching loss was configured identically to that of DCKD, based on the VQGAN [18] encoder. The weighting parameters for the sobel and spectral losses, $\lambda_{sobel}$ and $\lambda_{spec}$, were empirically set to 2, which yielded optimal performance.

For our training data, we used 400 images from the M3FD [20] infrared image dataset. Input images were center-cropped into 256×256 patches. For data augmentation, we randomly applied horizontal flips and rotations of 90°, 180°, and 270°.

All models were optimized using the Adam optimizer($\beta_1 = 0.9$, $\beta_2 = 0.99$). The initial learning rate was set to $1 \times 10^{-4}$ and was progressively decayed by a factor of 0.5 at predefined epochs. Training was conducted for a total of 150 epochs on two NVIDIA A100 80GB GPUs with a batch size of 4. The entire training process took approximately 6 hours.

## IV. EXPERIMENTS

### A. Experiment Details

This section provides a detailed description of the experimental setup, including the models, datasets, and evaluation metrics used.

*a) Evaluation Models:* To validate the performance of the methodology proposed in this study, we conducted evaluations using the following two teacher-student pairings. The first pairing uses CRG [12] as the teacher model and PSRGAN [9] as the student model. CRG is a model that exhibits state-of-the-art IRSR performance, and PSRGAN was deemed suitable as a student model due to its ability to train efficiently even with limited data. The second pairing utilizes the LISN [11] architecture for both the teacher and student models; in this case, the student model is a lightweight version (4 blocks) of the full teacher model (6 blocks). The detailed configurations of the teacher and student models are presented in Table I.

*b) Evaluation Datasets:* The proposed method was evaluated on three benchmark datasets: M3FD, Result-A, and Result-C. The M3FD [20] dataset is divided into Set5, Set15, and Set20 based on the number of test images. In this study, we selected Set20 for our evaluations as it contains a more diverse range of infrared scenes. Result-A [21] and Result-C [22] are major benchmarks in the field of IRSR. Although these two datasets are derived from the same source images, they are constructed using different fusion strategies, thereby providing diverse visual characteristics for robust evaluation.

*c) Evaluation Metrics:* The model's performance is quantitatively evaluated using two standard metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). PSNR measures reconstruction fidelity based on pixel-wise error, while SSIM assesses structural similarity as perceived by the human visual system. For both metrics, higher values indicate better performance. All evaluations were performed exclusively on the Y channel (luminance channel) of the YCbCr color space, which is standard way for

TABLE II
QUANTITATIVE COMPARISON ON THE BENCHMARK DATASETS FOR IRSR. THE BEST PERFORMANCE BETWEEN SCRATCH OR STUDENT MODELS ARE HIGHLIGHTED ON THE BOLD. THE TEACHER MODEL OF PSRGAN IS CRG.

| Scale | Role | Model | Result-A PSNR / SSIM | Result-C PSNR / SSIM | Set20 PSNR / SSIM | Model | Result-A PSNR / SSIM | Result-C PSNR / SSIM | Set20 PSNR / SSIM |
|---|---|---|---|---|---|---|---|---|---|
| ×2 | Teacher | LISN | 38.94/0.9422 | 39.91/0.9532 | 46.94/0.9906 | PSRGAN | 38.76/0.9388 | 39.63/0.9505 | 49.61/0.9917 |
| | Scratch | | 37.59/0.9325 | 38.56/0.9451 | 45.59/**0.9876** | | 37.34/0.9211 | 38.36/0.9344 | 43.34/0.9676 |
| | Student | | **37.61/0.9327** | **38.57/0.9453** | **45.61/0.9876** | | **39.00/0.9432** | **40.00/0.9541** | **47.37/0.9910** |
| ×4 | Teacher | | 32.96/0.8285 | 33.63/0.8480 | 41.50/0.9725 | | 34.60/0.8563 | 35.19/0.8729 | 43.59/0.9756 |
| | Scratch | | 32.93/0.8279 | 33.60/0.8474 | 40.17/0.9649 | | 34.60/0.8563 | 35.19/0.8729 | 43.59/0.9756 |
| | Student | | **32.94/0.8280** | **33.61/0.8475** | **40.18/0.9650** | | **34.27/0.8521** | **34.96/0.8702** | **41.84/0.9733** |

evaluating grayscale and infrared images, which are defined primarily by their luminance information.

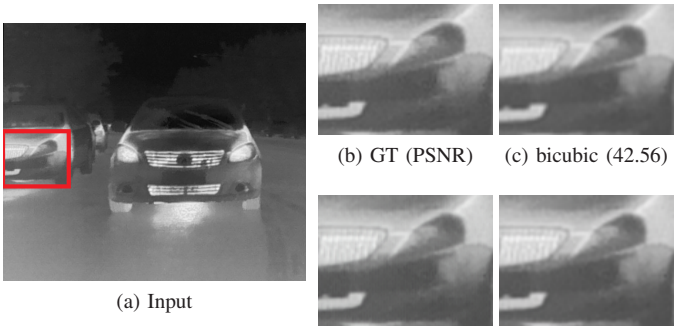
### B. Results on Infrared Image Super Resolution

To evaluate the effectiveness of the proposed methodology for IRSR, we compared the performance of each student model pairing against two baselines: the original teacher model, and the same student model trained from scratch, without other training strategy. All evaluations were conducted at ×2 and ×4 scaling factors.

The quantitative evaluation results are summarized in Table II. Although the teacher model achieved the highest performance in most cases, the student model trained with our proposed method consistently outperformed the scratch model. Notably, for the LISN-based model pairing, the student trained with our method achieved approximately 0.01-0.02 dB higher PSNR scores than the scratch model across all datasets. For example, on the Result-A dataset at ×2 scale, the scratch model achieved 37.59 dB, whereas our proposed student model reached 37.61 dB. Similarly, on the Set20 dataset, the PSNR improved from 45.59 dB to 45.61 dB.

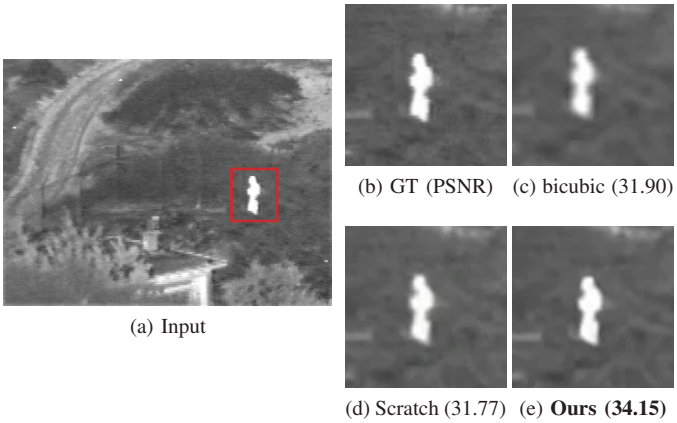
A more pronounced performance improvement was observed in the CRG–PSRGAN combination. The PSRGAN student model trained with our method showed substantial improvements over the scratch-trained counterpart (e.g., a 1.14 dB PSNR improvement on Result-C at ×4) and, in some cases, even surpassed the teacher model performance, a noteworthy and uncommon outcome in KD settings. Specifically, at the ×2 scale, the student model achieved PSNR scores of 39.00 dB and 40.00 dB on the Result-A and Result-C datasets, respectively, outperforming the teacher model by 0.34 dB and 0.37 dB. These results demonstrate that the proposed framework can effectively transfer knowledge from a teacher model to a lightweight IRSR model while maintaining a simple structure. In particular, achieving performance superior to the teacher model with the relatively low-capacity PSRGAN underscores the practicality and scalability of the proposed method.

Fig. 3 and 4 present a visual comparison of the results from the CRG-PSRGAN pairing, based on the different training methodologies. Compared to the scratch model, the student model trained with the proposed method can be observed to effectively restore finer textures at the ×2 scale (Fig. 3), while restoring object contours and edge structures more sharply at the ×4 scale (Fig. 4).



(a) Input (b) GT (PSNR) (c) bicubic (42.56) (d) Scratch (41.97) (e) **Ours (46.18)**

Fig. 3. Comparison of IRSR performance between different training methods on an image from Set20 dataset with 2× upscaling.



(a) Input (b) GT (PSNR) (c) bicubic (31.90) (d) Scratch (31.77) (e) **Ours (34.15)**

Fig. 4. Comparison of IRSR performance between different training methods on an image from Result-C dataset with 4× upscaling.

### C. Loss Function Comparison

We conducted an ablation study to analyze the individual contribution of each proposed loss function. The experiment used the CRG-PSRGAN pairing at ×4 upscaling factor with M3FD Set20 dataset, where we measured performance changes by sequentially adding our proposed loss functions to the baseline DCKD framework. The results, presented in Table III, clearly show that the two proposed loss functions

TABLE III
ABLATION STUDY ON THE LOSS FUNCTIONS.

| x4 | Original | + Spectral | + Sobel | + Spectral + Sobel |
|---|---|---|---|---|
| PSNR | 41.73 | 41.75 | 41.81 | **41.84** |
| SSIM | 0.9732 | 0.9731 | **0.9733** | **0.9733** |

contribute to IRSR performance improvement both individually and in a complementary manner.

Specifically, adding only the spectral loss to the baseline DCKD framework resulted in a 0.02 dB improvement in PSNR, while adding only the Sobel loss led to a larger improvement of 0.08 dB. The Sobel loss, in particular, is observed to help guide the student model to effectively mimic the teacher's boundary representations, owing to its sensitivity to edge structures.

When both losses were applied simultaneously, a synergistic effect was observed, achieving the highest performance with a PSNR improvement of more than 0.1 dB over the baseline. These results demonstrate that our study's approach of simultaneously optimizing for structural features and spectral-domain fidelity is highly effective for knowledge distillation for IRSR.

## V. CONCLUSION

This study proposed a novel KD methodology that effectively lightweights IRSR models by integrating Sobel Loss and Spectral Loss into the existing DCKD framework. Through our experiments, we have demonstrated that the proposed methodology can enhance the performance of lightweight IRSR models more effectively than standard supervised learning approaches. In particular, the two loss functions, designed to reflect the unique characteristics of infrared images, played a critical role in preserving structural information such as edges and maintaining consistency in the frequency domain during the super-resolution process. This highlights the potential for generating high-quality results even with low-complexity models by enabling a more sophisticated transfer of knowledge from the teacher model. Future work on more efficient architectures, diverse distillation strategies, and advanced multi-teacher or self-distillation frameworks will be crucial for bridging the gap between IRSR performance and practical application, thereby advancing the technological horizons of the field.

## ACKNOWLEDGMENT

This research was supported by the Korea Research Institute for Defense Technology planning and advancement(KRIT) grant funded by the Korean government(DAPA) (R240203)

## REFERENCES

[1] Huang, Y., Miyazaki, T., Liu, X., Omachi, S. Infrared image super-resolution: Systematic review, and future trends. arXiv preprint arXiv:2212.12322. (2022).

[2] Dong, C., Loy, C. C., He, K., Tang, X. (2015). Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence, 38(2), 295-307.

[3] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4681-4690).

[4] Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R. (2021). Swinir: Image restoration using swin transformer. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 1833-1844).

[5] Xiao, Y., Yuan, Q., Jiang, K., He, J., Jin, X., Zhang, L. (2023). EDiffSR: An efficient diffusion probabilistic model for remote sensing image super-resolution. IEEE Transactions on Geoscience and Remote Sensing, 62, 1-14.

[6] Hinton, G., Vinyals, O., Dean, J. (2015). Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531.

[7] Huang, Y., Omachi, S. (2024). Infrared Image Super-Resolution via GAN. In Applications of Generative AI (pp. 565-576). Cham: Springer International Publishing.

[8] Huang, Y., Jiang, Z., Wang, Q., Jiang, Q., Pang, G. Infrared image super-resolution via heterogeneous convolutional WGAN. In: Pacific rim international conference on artificial intelligence. Cham: Springer International Publishing, 2021. p. 461-472.

[9] Huang, Y., Jiang, Z., Lan, R., Zhang, S., Pi, K. Infrared image super-resolution via transfer learning and PSRGAN. IEEE Signal Processing Letters, 2021, 28: 982-986.

[10] Qin, F., Yan, K., Wang, C., Ge, R., Peng, Y., Zhang, K. LKFormer: Large kernel transformer for infrared image super-resolution. Multimedia Tools and Applications, 2024, 83.28: 72063-72077.

[11] Liu, S., Yan, K., Qin, F., Wang, C., Ge, R., Zhang, K., et al. Infrared image super-resolution via lightweight information split network. In: International Conference on Intelligent Computing. Singapore: Springer Nature Singapore, 2024. p. 293-304.

[12] Zou, Y., Chen, Z., Zhang, Z., Li, X., Ma, L., Liu, J., et al. Contourlet refinement gate framework for thermal spectrum distribution regularized infrared image super-resolution. arXiv preprint arXiv:2411.12530, 2024.

[13] Gao, Q., Zhao, Y., Li, G., Tong, T. Image super-resolution using knowledge distillation. In: Asian Conference on Computer Vision. Cham: Springer International Publishing, 2018. p. 527-541.

[14] He, Z., Dai, T., Lu, J., Jiang, Y., Xia, S. T. Fakd: Feature-affinity based knowledge distillation for efficient image super-resolution. In: 2020 IEEE international conference on image processing (ICIP). IEEE, 2020. p. 518-522.

[15] Li, S., Zhang, Y., Li, W., Chen, H., Wang, W., Jing, B., Hu, J. Knowledge distillation with multi-granularity mixture of priors for image super-resolution. arXiv preprint arXiv:2404.02573, 2024.

[16] Zhang, Y., Li, W., Li, S., Chen, H., Tu, Z., Wang, W., et al. Data upcycling knowledge distillation for image super-resolution. arXiv preprint arXiv:2309.14162, 2023.

[17] Zhou, Y., Qiao, J., Liao, J., Li, W., Li, S., Xie, J., et al. Dynamic Contrastive Knowledge Distillation for Efficient Image Restoration. In: Proceedings of the AAAI Conference on Artificial Intelligence. 2025. p. 10861-10869.

[18] Esser, P., Rombach, R., Ommer, B. Taming transformers for high-resolution image synthesis. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021. p. 12873-12883.

[19] Wang, Z., Chen, J., Hoi, S. C. Deep learning for image super-resolution: A survey. IEEE transactions on pattern analysis and machine intelligence, 2020, 43.10: 3365-3387.

[20] Liu, J., Fan, X., Huang, Z., Wu, G., Liu, R., Zhong, W., Luo, Z. Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visible for object detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022. p. 5802-5811.

[21] Liu, Y., Chen, X., Cheng, J., Peng, H., Wang, Z. Infrared and visible image fusion with convolutional neural networks. International Journal of Wavelets, Multiresolution and Information Processing, 2018, 16.03: 1850018.

[22] Zhang, Y., Zhang, L., Bai, X., Zhang, L. (2017). Infrared and visual image fusion through infrared feature extraction and visual information preservation. Infrared Physics & Technology, 2017, 83: 227-237.

[23] Wu, W., Wang, T., Wang, Z., Cheng, L., Wu, H. Meta transfer learning-based super-resolution infrared imaging. Digital Signal Processing, 131, 103730. 2022.

[24] Gupta, S. and Mazumdar, S. G. Sobel edge detection algorithm. International journal of computer science and management Research, 2(2):1578–1583. 2013.