International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 06 June 2025, pp: 237-245 www.ijaem.net ISSN: 2395-5252

Thermal Image Enhancement and Semantic Multi-Task Learning for Accurate Infrared Colorization

S. Hovhannisyan[1], S. Agaian[2]
[1]*Department of Mathematics and Mechanics, Yerevan State University, Yerevan, 0025, Armenia*
[2] *Department of Computer Science, College of Staten Island, 2800 Victory Blvd. Staten Island, New York, NY 10314, USA*

Date of Submission: 28-05-2025 Date of Acceptance: 08-06-2025

ABSTRACT: Thermal infrared imaging plays a vital role in computer vision applications, especially under difficult conditions such as low light and poor visibility, where traditional RGB cameras struggle. Transforming night-vision thermal images to resemble daytime colors is essential for better scene interpretation and decision-making in surveillance, autonomous navigation, and search-and-rescue operations. Nonetheless, overcoming the significant modality gap between visual and thermal cameras is still a challenge, as they capture fundamentally different physical properties of the same scene, thermal radiation versus reflected light. The spectral mapping from thermal imagery to realistic color imagery presents significant technical challenges due to: (i) the inherent nonlinear relationship between temperature distributions and visible color representations, (ii) the characteristic low signal-to-noise ratio of thermal images, along with their poor contrast and indistinct structural details that significantly hinder effective colorization for real-world applications, and (iii) the lack of pixel-level correspondence between thermal and RGB modalities in existing datasets.

This paper presents an innovative end-to-end neural network architecture pipeline for automatic thermal image colorization, featuring a crucial preprocessing enhancement stage specifically designed to tackle these fundamental challenges. Our approach incorporates specialized enhancement filtering to enhance thermal image quality before colorization. The core network employs a U-Net-based encoder-decoder architecture with skip connections and attention mechanisms to retain spatial details during the thermal-to-RGB translation. Furthermore, we incorporate a segmentation-guided multitask learning framework that explicitly improves semantic consistency during colorization by jointly optimizing both colorization accuracy and semantic segmentation. This dual-objective approach ensures that generated colors align with semantically meaningful regions in the thermal input. Quantitative evaluation using BRISQUE, NIQE, and PIQE metrics, along with perceptual studies, confirms that our preprocessing enhancement significantly improves the visibility and structural clarity of thermal images, enabling more accurate and realistic colorization compared to existing methods. Comprehensive qualitative and quantitative evaluations highlight the superiority of our approach in terms of perceptual realism, structural accuracy, and semantic fidelity, achieving state-of-the-art performance.

KEYWORDS: Image colorization**,** Image enhancement, Object detection.

I. INTRODUCTION

High-quality imagery is crucial for enhancing human perception and enabling the robust performance of automated computer vision systems. Many real-world scenarios involve capturing video or images under challenging conditions, such as low illumination[1], poor visibility[2], structural degradation, and significant noise interference. These adverse factors substantially degrade the accuracy of computer vision tasks, including surveillance, object detection, semantic segmentation, and object tracking[3]. Conventional methods, mainly reliant on visible-spectrum (VIS) cameras, often face limitations in low-light, foggy, or nighttime environments, resulting in diminished accuracy due to low contrast, motion blur, and inherent noise[4, 5].

In response to these limitations, thermal infrared (TIR) imaging has emerged as an invaluable alternative, providing critical advantages in low-light and visibility-compromised situations.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 06 June 2025, pp: 237-245 www.ijaem.net ISSN: 2395-5252

Thermal images are notably resilient to variations in illumination, have strong fog penetration capability, and exhibit high sensitivity to temperature differences. These attributes have made thermal imagery essential in various fields, including wildlife monitoring[6], security and military operations[7, 8], and remote sensing[9].

Figure 1. Example of original thermal image, enhanced image, and colorization results with the proposed pipeline.

Despite these benefits, thermal images pose challenges such as low contrast, lack of color information, and indistinct object boundaries. These limitations hinder both human interpretation and automated analysis. While enhancement techniques can partially address contrast and detail[10, 11, 12, 13, 14], RGB-based computer vision methods still struggle to operate effectively on thermal data due to its fundamentally different characteristics. As a result, methods to convert thermal infrared images into visually interpretable colored images have garnered significant attention.

Several recent methods have aimed to tackle thermal image colorization, which can be broadly categorized into supervised and unsupervised approaches. Supervised methods[15, 16] typically minimize pixel-level differences between the generated and ground-truth RGB images. While effective, these approaches often produce blurry outputs with inadequate edge sharpness and unclear semantics.

On the other hand, unsupervised image-to-image translation models have emerged as promising due to their ability to handle unpaired datasets effectively. CycleGAN[17], a foundational unsupervised approach, utilizes adversarial training and cycle-consistency loss; however, it struggles with unrealistic textures and inconsistent semantics. Methods like DRIT++[18] employ disentangled representations for diversified translations, yet they often compromise detail accuracy and semantic integrity. UNIT[19] proposes sharing latent spaces to maintain consistency, but suffers from incomplete translations of smaller objects. U-GAT-IT [20]incorporates attention mechanisms to highlight crucial features, but still faces challenges in achieving entirely realistic object translation. PearlGAN[21] further enhances semantic consistency with guided attention and structured gradient alignment, yet minor differences in thermal imagery remain challenging.

As illustrated in Figure 1, thermal infrared images inherently exhibit poor contrast, noise, and limited visual clarity, which complicate subsequent colorization processes. The original thermal images (left column) reveal significant issues, such as blurred edges and low structural definition, that hinder accurate semantic interpretation. After applying the proposed enhancement method (middle column), notable improvements in visibility and structural details are evident, including more precise edges, reduced noise, and better-defined object boundaries. These enhanced images significantly facilitate more accurate and visually realistic colorization results (right column). Hence, the enhancement step emerges as a critical prerequisite that directly influences the quality of the subsequent colorization and ensures improved interpretability and usability for further vision-related tasks.

Our work addresses these critical gaps by proposing a comprehensive thermal infrared colorization pipeline that systematically tackles the fundamental challenges of thermal-to-RGB domain translation. We introduce several key innovations aimed at overcoming existing limitations in thermal image processing and colorization:

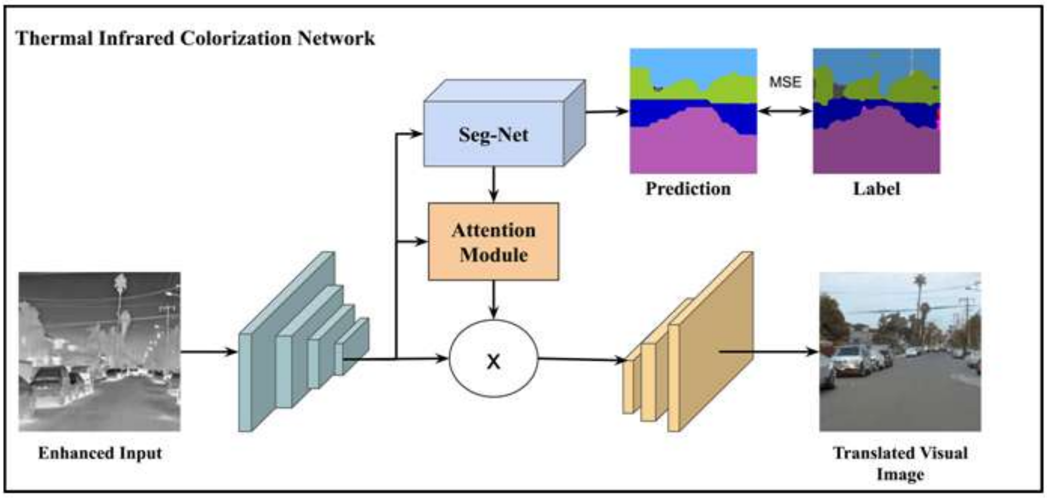

Figure 2. Architecture of the proposed method for thermal image translation.

1. **Adaptive Thermal Image Enhancement Framework:** This enhancement method significantly improves image contrast, reduces artifacts, and enhances edge definition, creating optimal conditions for subsequent processing tasks, including feature extraction, edge detection, and semantic segmentation before colorization. Our enhancement approach is specifically tailored to address the unique characteristics of thermal imagery, including low dynamic range and temperature-dependent noise patterns.
2. **Semantically Guided Multitask Learning Architecture:** We propose an integrated framework that employs a Residual Channel Attention Network (RCAN)-based segmentation module to generate high-quality semantic maps from enhanced thermal inputs. This multitask approach introduces semantically informed cross-attention mechanisms that explicitly guide the colorization process by ensuring color consistency within semantic regions. The joint optimization of colorization and segmentation tasks creates a feedback loop that improves semantic understanding and color accuracy, addressing the critical challenge of maintaining semantic coherence in thermal-to-RGB translation.
3. **Comprehensive Evaluation Framework with Downstream Task Validation:** We establish a rigorous evaluation protocol that extends beyond traditional colorization metrics to include practical applicability assessment. Our evaluation framework employs no-reference image quality assessment metrics—BRISQUE, NIQE, and PIQE—which quantify perceptual distortions by analyzing natural scene statistics without requiring a reference image, alongside downstream task performance validation through semantic segmentation accuracy (mIoU) and object detection performance (mAP) on colorized outputs. This comprehensive evaluation demonstrates the visual quality of our colorization results and their practical utility for computer vision applications, showcasing superior performance and strong real-world applicability compared to existing thermal colorization methods.

By systematically incorporating these innovations, our proposed method achieves marked improvements in visual quality and semantic fidelity, advancing the capabilities of thermal image colorization significantly beyond existing approaches.

II. PROPOSED METHOD

We propose an enhanced thermal infrared colorization network inspired by the PearlGAN [21] architecture, aiming to significantly improve the visual quality and semantic accuracy of translated thermal images, as shown in Figure 2. Our approach

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 06 June 2025, pp: 237-245 www.ijaem.net ISSN: 2395-5252

introduces critical modifications and additions designed to tackle inherent challenges in thermal imagery, such as noise and limited semantic clarity.

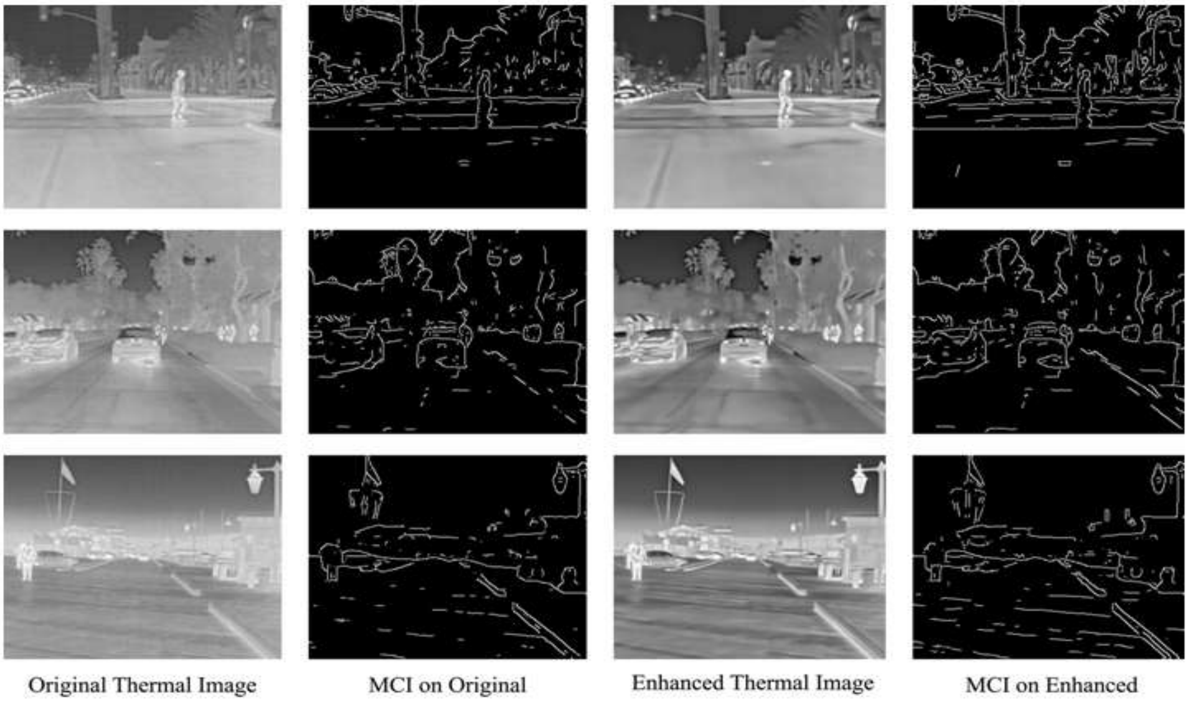

Figure 3. MCI edge detection method results before and after the enhancement method.

Initially, the attention mechanism in PearlGAN, which utilizes a CycleGAN-based[17] architecture, is completely removed. This decision is motivated by our objective to integrate a more semantically robust attention strategy aligned with the segmentation information. Instead, we embed a dedicated segmentation module and leverage multitask training to explicitly incorporate semantic features. The segmentation network employs a Residual Channel Attention Network [22] architecture, known for its effectiveness in capturing detailed semantic representations.

We introduce a simplified convolution-based attention module that effectively generates channel-wise attention maps from segmentation features to leverage semantic information during colorization. Specifically, the encoder features from our primary thermal translation network are multiplied channel-wise with the generated attention maps before feeding them into the decoder. This ensures that semantic context directly guides the image translation, enhancing realism and semantic coherence in the resulting colorized images.

A critical innovation of our methodology is including a pre-colorization thermal image enhancement step using the [23] enhancement method. This enhancement addresses the intrinsic noise and low-quality characteristics of thermal infrared imagery. By improving image visibility and extracting clearer, more informative features, the enhancement substantially aids subsequent processing tasks, particularly segmentation and edge extraction. Enhanced images exhibit fewer artifacts and improved detail retention, which is crucial for accurate semantic labeling and feature extraction.

We adopt an unsupervised semantic segmentation technique described in [24] for segmentation supervision. This method generates high-quality segmentation maps from thermal imagery without manual labeling. Leveraging a memory regularization approach accurately predicts reliable semantic boundaries and classes within thermal images, significantly enhancing the training accuracy of our segmentation module.

We preserve the core PearlGAN training paradigm regarding loss functions, excluding losses directly tied to their original attention mechanism. Specifically, we retain edge-aware loss functions, acknowledging their importance in maintaining structural details. Additionally, our enhanced thermal input significantly reduces artifacts previously observable in edge maps generated by the MCI method [25]initially integrated into PearlGAN, as shown in Figure 3. Visual assessments further highlight that our enhancement

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 06 June 2025, pp: 237-245 www.ijaem.net ISSN: 2395-5252

step markedly improves the fidelity and accuracy of edge delineation.

Figure 4. Qualitative comparison of existing methods.

Our proposed method systematically integrates thermal image enhancement, semantically-informed attention, and segmentation-guided multitask training. This comprehensive approach significantly improves the visual quality and semantic accuracy of colorized thermal infrared images, showing notable advancements compared to baseline methods.

Table 1. Quantitative comparison of the proposed pipeline with existing methods.

	BRISQUE	NIQE	PIQE
CycleGAN	22.17	8.57	12.70
UNIT	25.50	7.81	9.12
UGATIT	26.14	6.73	7.63
DRIT++	24.23	8.56	9.49
PearlGAN	23.43	6.15	9.25
Proposed	**21.61**	**5.94**	**6.82**

III. EXPERIMENTAL RESULTS

This section presents a comprehensive qualitative and quantitative analysis comparing our proposed method against prominent GAN-based thermal-to-visible image translation techniques, such as CycleGAN, DRIT++, U-GAT-IT, UNIT, and PearlGAN. We aim to demonstrate our method's significant improvements regarding realism, detail preservation, and overall structural accuracy.

QUALITATIVE COMPARISON

Figure 4 illustrates qualitative comparisons of translation outcomes across different methodologies on selected thermal images. The translation results highlight several apparent shortcomings in baseline approaches. CycleGAN tends to generate unrealistic textures and frequently oversimplifies structural details, causing essential objects like vehicles and pedestrians to blend into their surroundings or become indistinct. DRIT++, while generally producing brighter images, often leads to overexposure, merging adjacent objects, and losing finer details such as trees and pedestrians. U-GAT-IT achieves improved realism but struggles significantly with clarity, particularly in preserving detailed object boundaries and avoiding object-background fusion. UNIT performs well for larger structures but typically fails to preserve minor details, resulting in blurred representations of distant objects like pedestrians and finer tree branches. PearlGAN, despite demonstrating effectiveness in retaining prominent objects, introduces excessively dark or indistinct regions, reducing the visibility of

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 06 June 2025, pp: 237-245 www.ijaem.net ISSN: 2395-5252

smaller objects and often inaccurately merging closely positioned items.

In contrast, our proposed method consistently yields superior results, clearly translating thermal images into visually coherent daytime scenarios. Our approach accurately captures fine structural details, maintains natural illumination without unnatural brightness or darkness, and provides more apparent differentiation among closely positioned objects. The qualitative comparison shows that our method demonstrates enhanced realism, producing translations that accurately reflect scene contents and maintain structural and semantic integrity.

Original Thermal Image | PearlGan | Enhanced Thermal Image | Proposed Method

Figure 5. Qualitative comparison with the baseline method.

Figure 5 further illustrates a detailed comparative evaluation between our method and the baseline PearlGAN, emphasizing the impact of our thermal image enhancement process. The visual results demonstrate noticeable improvements when employing our pre-colorization enhancement step. Enhanced thermal images result in significantly more accurate translations, characterized by sharper edges, improved visibility of structural details, and fewer translation artifacts than those generated directly from the original thermal images. This comparison demonstrates the efficacy and necessity of incorporating the enhancement stage into the colorization pipeline, substantially improving translation quality and ensuring semantic consistency and visual realism.

Overall, the qualitative analysis validates our proposed method's ability to generate visually accurate and structurally precise images, making it highly suitable for practical applications in computer vision and surveillance.

Table 2. Evaluation of object detection accuracy on translated images using various methods (IoU threshold: 0.50).

	Person	Bicycle	Car	mAP
Figure 5. Qualitative comparison with the baseline method.				
CycleGAN	21	3.2	38.4	20.9
UNIT	17.5	11.2	20.5	16.4
UGATIT	11.8	1.1	37.6	16.5
DRIT++	20.6	2.1	48.5	23.8
PearlGAN	59.3	25.1	75.4	53.3
Proposed	**63.8**	**29.8**	**79.1**	**57.5**

QUANTITATIVE COMPARISON

To objectively evaluate the performance of our proposed method, we employed three widely recognized non-reference image quality assessment metrics:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 06 June 2025, pp: 237-245 www.ijaem.net ISSN: 2395-5252

- Natural Image Quality Evaluator (NIQE)[26]: This metric assesses image quality based on statistical patterns found in natural images.
- Perception-based Image Quality Evaluator (PIQE)[27]: This metric evaluates the visual perceptual quality of images by quantifying noticeable distortions.
- BlindImage Spatial Quality Evaluator (BRISQUE)[28]: This method measures image quality relying solely on statistical regularities of spatial features without requiring reference images.

A lower score indicates better image quality, improved realism, and greater perceptual authenticity across all these metrics. The comparative performance results for these metrics are detailed in Table 1. Our proposed method consistently achieves the lowest scores compared to existing techniques, highlighting its effectiveness in producing images with enhanced perceptual realism and natural visual characteristics.

Additionally, to validate the practical utility of our generated images, we assessed object detection performance using the detection model[29]. We computed Average Precision (AP), which measures the precision of detection at different recall thresholds, and used mean Average Precision (mAP) across multiple object classes as a comprehensive performance measure. As detailed in Table 2, our method obtained the highest mAP scores, clearly outperforming all methods across all evaluated object categories.

Table 3. Semantic segmentation results using various image translation methods (i.e., pixel-wise classification into predefined object categories)

	Road	Building	Sky	Person	Car	Truck	Bus	Traffic Sign	Motorcycle	mIoU
CycleGAN	95.6	39.1	90.8	60.8	78.0	0	0	0	0	40.5
UNIT	96.2	60.3	92.1	64.5	71.5	0	0	0.2	14.6	44.4
UGATIT	94.3	18.4	89.0	23.7	59.0	0	0	0	0	31.6
DRIT++	97.3	29.4	78.4	28.0	78.9	0	0	0	0	34.7
PearlGAN	97.7	73.1	93.4	73.2	82.7	0.1	0	0	0	46.7
Proposed	**98.3**	**79.5**	**96.8**	**89.3**	**92.7**	**1.5**	11.8	0	**18.1**	**54.2**

Additionally, semantic coherence was further examined through segmentation analysis, utilizing the segmentation method[30]for color images based on the Intersection over Union (IoU) metric, as shown in Table 3. Our proposed pipeline achieved higher mean IoU values across the semantic classes, demonstrating a strong retention of semantic information within the transformed imagery.

These quantitative assessments demonstrate that our approach significantly enhances image quality and boosts downstream task performance, confirming its suitability and effectiveness for real-world applications.

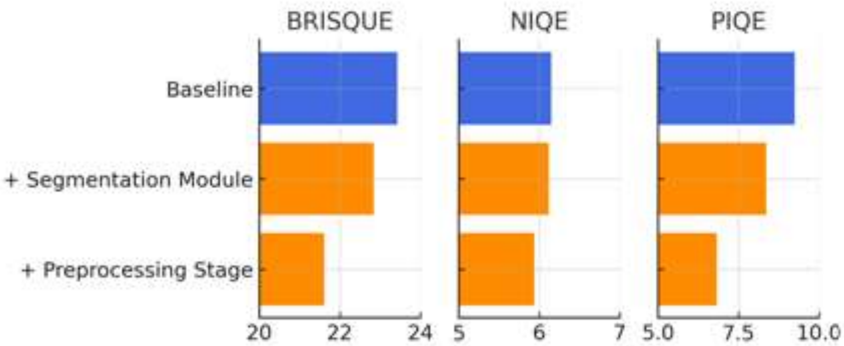

Figure 6. Effectiveness of each module in the proposed pipeline.

IV. ABLATION STUDY

We conducted ablation experiments to thoroughly investigate the individual contributions of the image enhancement preprocessing step and the segmentation-guided multitask learning strategy within our proposed pipeline. Specifically, we systematically modified our network by independently removing each component, allowing for a clear assessment of their specific impacts on overall translation performance. The quantitative results of these ablation tests are summarized in

Figure 6. These findings clearly illustrate each component's unique and significant contributions toward enhancing the pipeline's effectiveness. The enhancement preprocessing step notably improves translated images' clarity and structural detail, while the segmentation multitask learning mechanism significantly enriches semantic accuracy and realism. These results confirm that both elements are essential for achieving the highest quality in thermal-to-visible image translation.

V. CONCLUSION

This paper presents a novel thermal infrared image colorization method integrating a crucial thermal enhancement preprocessing step and a segmentation-guided multitask learning strategy. This preprocessing step significantly reduces inherent noise and enhances structural clarity, critical for accurate colorization. Our method effectively captures detailed semantic representations by employing an RCAN-based semantic segmentation module combined with semantically informed attention mechanisms, further improving the quality and realism of the translated images. Comprehensive experiments demonstrate the clear advantages of the proposed pipeline over existing methods, as indicated by superior performance across established quality metrics and downstream computer vision tasks, including object detection and semantic segmentation. The ablation studies confirm the essential contributions of preprocessing enhancement and multitask learning components. Overall, our approach significantly advances the field of thermal image colorization, offering robust, realistic, and highly applicable visual outputs suitable for practical use in diverse vision-based applications.

REFERENCES

[1]. H. Gasparyan, S. Hovhannisyan, S. Babayan and S. Agaian, “Iterative Retinex-based decomposition framework for low light visibility restoration,” IEEE Access, vol. 11, pp. 40298--40313, 2023.

[2]. S. Hovhannisyan, H. Gasparyan, S. Agaian and A. Ghazaryan, “AED-Net: A single image dehazing,” IEEE access, vol. 10, pp. 12465--12474, 2022.

[3]. S. Hovhannisyan, H. Gasparyan and S. Agaian, “EOD-Net: enhancing object detection in challenging weather conditions using an innovative end-to-end dehazing network,” in 2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA), 2023.

[4]. Z. Liu, Y. He, C. Wang and R. Song, “Analysis of the influence of foggy weather environment on the detection effect of machine vision obstacles,” Sensors, vol. 20, no. 2, p. 349, 2020.

[5]. V. Frants, S. Agaian and K. Panetta, “QCNN-H: Single-Image Dehazing Using Quaternion Neural Networks,” IEEE Transactions on Cybernetics, vol. 53, no. 9, pp. 5448--5458, 2023.

[6]. E. Yoshida, S. Kobayashi, H. Mukai, N. Uthumphirat, Y. Nakamura, N. Ogino and T. Yokotani, “Proposal and prototyping on wildlife tracking system using infrared sensors,” in 2022 International Conference on Information Networking (ICOIN), 2022.

[7]. S. Kim, W. Kim, J. Park and K. Yeo, “Human detection in infrared image using daytime model-based transfer learning for military surveillance system,” in 2023 14th International Conference on Information and Communication Technology Convergence (ICTC), 2023.

[8]. K. Alheeti and K. McDonald-Maier, “An intelligent security system for autonomous cars based on infrared sensors,” in 2017 23rd International Conference on Automation and Computing (ICAC), 2017.

[9]. Z. Wang, S. Li and K. Huang, “Cross-Modal Adaptation for Object Detection in Thermal InfraRed Remote Sensing Imagery,” IEEE Geoscience and Remote Sensing Letters, 2025.

[10]. S. S. Agaian, K. Panetta and A. M. Grigoryan, “Transform-based image enhancement algorithms with performance measure,” IEEE Transactions on image processing, vol. 10, no. 3, pp. 367--382, 2001.

[11]. S. S. Agaian, B. Silver and K. A. Panetta, “Transform coefficient histogram-based image enhancement algorithms using contrast entropy,” IEEE transactions on image processing, vol. 16, no. 3, pp. 741--758, 2007.

[12]. B. Silver, S. Agaian and K. Panetta, “Contrast entropy based image enhancement and logarithmic transform coefficient histogram shifting,” in Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005., 2005.

[13]. S. S. Agaian, Visual morphology, SPIE, 1999.

[14]. V. Voronin, S. Tokareva, E. Semenishchev and S. Agaian, “Thermal image enhancement algorithm using local and global logarithmic

transform histogram matching with spatial equalization," in 2018 IEEE southwest symposium on image analysis and interpretation (SSIAI), 2018.

[15]. U. Qayynm, Q. Ahsan, Z. Mahmood and M. A. Chcmdary, "Thermal colorization using deep neural network," in 2018 15th International Bhurban Conference on Applied Sciences and Technology (IBCAST), 2018.

[16]. T. Wang, T. Zhang and B. C. Lovell, "EBIT: Weakly-supervised image translation with edge and boundary enhancement," Pattern Recognition Letters, vol. 138, pp. 534--539, 2020.

[17]. J.-Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in Proceedings of the IEEE international conference on computer vision, 2017.

[18]. H.-Y. Lee, H.-Y. Tseng, J.-B. Huang, M. Singh and M.-H. Yang, "Diverse image-to-image translation via disentangled representations," in Proceedings of the European conference on computer vision (ECCV), 2018.

[19]. M.-Y. Liu, T. Breuel and J. Kautz, "Unsupervised image-to-image translation networks," Advances in neural information processing systems, p. 30, 2017.

[20]. H.-Y. Lee, Y.-H. Li, T.-H. Lee and M. S. Aslam, "Progressively unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation," Sensors, vol. 23, no. 15, p. 6858, 2023.

[21]. F. Luo, Y. Li, G. Zeng, P. Peng, G. Wang and Y. Li, "Thermal infrared image colorization for nighttime driving scenes with top-down guided attention," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 9, pp. 15808--15823, 2022.

[22]. Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in Proceedings of the European conference on computer vision (ECCV), 2018.

[23]. S. Hovhannisyan, S. Agaian, K. Panetta and A. Grigoryan, "Thermal Video Enhancement Mamba: A Novel Approach to Thermal Video Enhancement for Real-World Applications," Information, vol. 16, no. 2, p. 125, 2025.

[24]. Z. Zheng and Y. Yang, "Unsupervised scene adaptation with memory regularization in vivo," arXiv preprint arXiv:1912.11164, 2019.

[25]. K.-F. Yang, C.-Y. Li and Y.-J. Li, "Multifeature-based surround inhibition improves contour detection in natural images," IEEE Transactions on Image Processing, vol. 23, no. 12, pp. 5020--5032, 2014.

[26]. A. Mittal, R. Soundararajan and A. C. Bovik, "Making a "completely blind" image quality analyzer," IEEE Signal processing letters, vol. 20, no. 3, pp. 209--212, 2012.

[27]. N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya and S. S. Medasani, "Blind image quality evaluation using perception based features," in 2015 twenty first national conference on communications (NCC), 2015.

[28]. A. Mittal, A. K. Moorthy and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Transactions on image processing, vol. 21, no. 12, pp. 4695--4708, 2012.

[29]. A. Mittal, A. K. Moorthy and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Transactions on image processing, vol. 21, no. 12, pp. 4695--4708, 2012.

[30]. A. Tao, K. Sapra and B. Catanzaro, "Hierarchical multi-scale attention for semantic segmentation," arXiv preprint arXiv:2005.10821, 2020.