Mathematical Problems of Computer Science 62, 112–125, 2024.
doi:10.51408/1963-0125

UDC 004.93

Enhancing Thermal Image Classification with Novel Quality Metric-Based Augmentation Techniques

Hrach Y. Ayunts

Yerevan State University, Yerevan, Armenia
e-mail: hrach.ayunts@ysu.am

Abstract

Thermal image classification is critical in various applications, particularly fault detection and monitoring systems such as photovoltaic (PV) modules. However, a common challenge in these fields is the limited availability of large-scale, labeled thermal image datasets. To address this, color image augmentation is widely adopted in machine learning to artificially increase the size and diversity of training datasets, improving model generalization. Traditional augmentation techniques, such as geometric transformations, provide some benefits but may fail to fully capture the unique characteristics inherent in thermal images, which often have lower contrast and different noise patterns than visible spectrum images. So, we argued we need to develop a novel augmentation technique for thermal imaging, where data collection is costly and time-consuming.

Our research proposes a novel offline augmentation technique guided by quality metrics to enhance the performance of thermal image binary classification models. By leveraging domain-specific quality metrics, such as image clarity, thermal contrast, and noise levels, we optimize the oversampling process for thermal datasets. For example, starting with a dataset of x images, we generate y additional thermal images, resulting in a total of x + y images used to train the deep learning classification framework. Using a dataset of PV module defects, we demonstrate the effectiveness of our quality metric-based oversampling strategy across several state-of-the-art image classification networks. Our approach outperforms traditional augmentation methods regarding classification accuracy and robustness, including geometric transformations and standard image enhancement techniques. The practical implications of our research are significant, as it provides a more effective and efficient way to improve model performance for thermal imaging tasks, mainly when data availability is limited.

Keywords: Dataset augmentation, Image quality assessment, Thermal image classification.

Article info: Received 10 October 2024; sent for review 19 October 2024; accepted 26 November 2024.

Acknowledgment: The research is conducted within the ADVANCE Research Grant provided by the Foundation for Armenian Science and Technology funded by Sarkis and Nune Sepetjians.

I would also like to express my deepest gratitude to my supervisor, Professor Sos Agaian, for his constant guidance, support, and valuable insights throughout this work.

1. Introduction

Thermal image classification plays a pivotal role in various critical applications, including fault detection and monitoring of photovoltaic (PV) modules, building inspections, and medical diagnostics [1, 2, 3, 4]. The unique advantage of thermal imaging lies in its ability to capture infrared radiation, making it highly effective for detecting heat signatures and identifying invisible anomalies to standard visible-light cameras. However, the development of robust deep learning models for thermal image classification remains challenging due to the limited availability of large-scale, labeled thermal image datasets [5]. Acquiring and labeling thermal data is time-consuming and costly, which hinders the creation of models capable of generalizing well across diverse real-world conditions.

This paper aims to develop a novel quality metric-based augmentation technique for fault detection and monitoring systems such as photovoltaic (PV) modules. Our proposed method addresses the challenges of data scarcity and enhances model performance by introducing thermal-specific augmentation strategies.

Data collection is essential when public computer vision datasets are insufficient, as applications such as fault detection and monitoring systems often require more data. However, data collection for computer vision training is both expensive and labor-intensive. Image annotation, which involves creating ground-truth data for model training, requires costly human labor. Building large image datasets is incredibly challenging due to the rarity of events, privacy concerns, the need for industry experts for labeling, and the significant expense and manual effort required to record visual data. These challenges underscore the need for data augmentation in computer vision.

Data augmentation is a set of techniques that enhance the size and quality of machine learning training datasets, enabling the training of better deep learning models. The most difficult challenge is the generalizability of deep learning models, which refers to the performance difference of a model when evaluated on previously seen data (training data) versus data it has never seen before (testing data). Models with poor generalizability have overfitted the training data.

To address the limitations posed by small datasets, data augmentation techniques have become essential in computer vision [6, 7]. Traditional augmentation methods, such as geometric transformations (e.g., rotations, flips, and scaling) and brightness adjustments, artificially increase the size and diversity of training datasets. These techniques expose models to a broader range of variations, enhancing their ability to generalize across unseen data. However, traditional augmentations are unsuitable for thermal images due to their distinct characteristics [8]. Thermal images typically exhibit lower contrast, unique noise patterns, and temperature-specific information that differ from visible-spectrum images. As a result, augmentation strategies for visible-light images may distort critical thermal features like temperature gradients and hot spots, reducing the effectiveness of the models.

Recent studies have proposed various methods to improve thermal image classification for PV module fault detection by tailoring augmentation techniques to thermal data. For instance, Korkmaz et al. [5] introduced an efficient fault classification method using transfer learning and a multi-scale convolutional neural network (CNN). They applied geometric transformations and brightness adjustments to enhance their dataset, achieving notable improvements in classification accuracy. Similarly, Pamungkas et al. [8] proposed a novel method using a coupled UDenseNet architecture for efficient solar panel fault classification. Their approach combined geometric transformations with generative adversarial networks

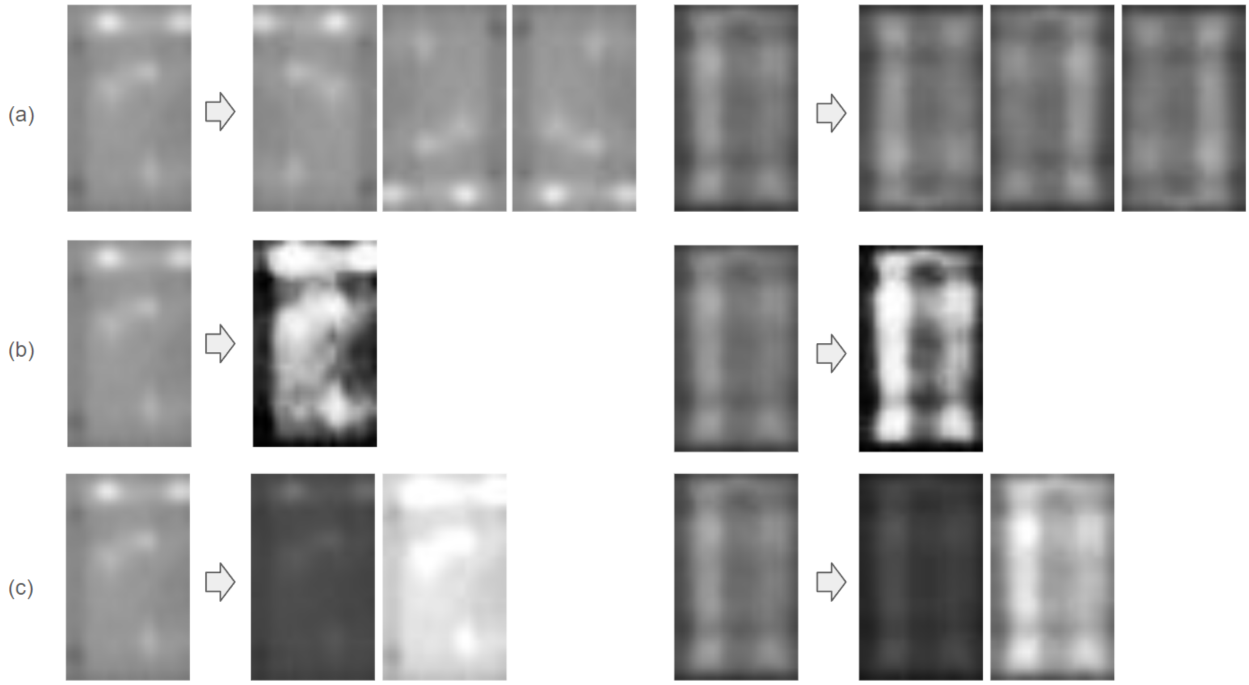

Fig.1. Examples of augmented thermal images: (a) geometric transformations (horizontal, vertical, and both flips), (b) histogram equalization, and (c) brightness adjustment.

(GANs) to generate additional synthetic thermal images, significantly boosting model performance. In another approach, Tang et al. [9] focused on automatic defect identification in PV panels using infrared (IR) images captured by unmanned aircraft. They utilized histogram equalization to enhance the contrast of the thermal images, improving the clarity of defects and anomaly detection. Some of these augmentation techniques, including geometric transformations, brightness adjustment, and histogram equalization, are depicted in Fig. 1.

In this paper, we address the limitations of traditional augmentation methods by proposing a novel offline augmentation technique guided by thermal-specific image quality metrics. These quality metrics, including clarity, thermal contrast, and noise levels, are tailored to the unique characteristics of thermal images. By optimizing the augmentation process based on these metrics, we generate additional samples that more accurately reflect real-world variations in thermal data, improving the robustness and generalization of thermal image classification models.

To validate our approach, we experiment with a PV module defect dataset [10] and apply quality metric-based oversampling to augment the training set. We evaluate our method across several state-of-the-art deep learning architectures, including CNNs like AlexNet [11] and ResNet50 [12], as well as transformer-based models such as Swin Transformer [13]. These networks, described in Table 1, are commonly employed for classification tasks and serve as a benchmark for our proposed augmentation technique. Our experimental results demonstrate that our method not only outperforms traditional augmentation techniques but also significantly improves classification accuracy and model robustness across diverse networks.

Table 1: Popular Networks for Image Classification

Network Description	Number of Parameters
AlexNet [11] is a CNN architecture that won the 2012 ImageNet competition, pioneering the use of deep learning for large-scale image classification. It features five convolutional layers, followed by three fully connected layers. AlexNet's use of ReLU activation and dropout was instrumental in improving performance. *PyTorch model: alexnet*	61 million
ResNet50 [12] introduces residual connections that enable the training of much deeper architectures by addressing the vanishing gradient problem. ResNet50, with its 50 layers, became one of the most influential models in computer vision, especially for transfer learning. *PyTorch model: resnet50*	25.6 million
SqueezeNet [14] is a compact CNN architecture that achieves AlexNet-level accuracy with 50x fewer parameters, thanks to its "fire modules," which consist of a squeeze layer and an expand layer to reduce the number of parameters while maintaining accuracy. *PyTorch model: squeezenet1_1*	1.2 million
ShuffleNetV2 [15] is designed for lightweight mobile and embedded vision applications. It uses group convolutions and a channel shuffle operation to reduce computational complexity while maintaining high accuracy, making it highly efficient for mobile devices. *PyTorch model: shufflenet_v2_x1_0*	2.3 million
MobileNetV3 [16] is the third version of the MobileNet family, designed for high efficiency in mobile and embedded applications. It uses a combination of inverted residuals and squeeze-and-excitation (SE) modules to balance speed and accuracy. *PyTorch model: mobilenet_v3_small*	2.5 million
Swin Transformer [13] is a hierarchical transformer model that applies attention within shifted windows, enabling both local and global information capture. It is scalable and effective across different vision tasks. *PyTorch model: swin_v2_t*	28 million

The key contributions of this work are summarized as follows:

- We introduce a novel augmentation method guided by a thermal-specific quality metric, which enhances the performance of thermal image classification models.
- We demonstrate the effectiveness of our approach using a PV module defect dataset, showing that our method improves classification accuracy and model robustness.
- We evaluate our method on various state-of-the-art deep learning architectures, including both CNN-based and transformer-based models, showcasing its general applicability.

The remainder of the paper is structured as follows. Section 2. presents our proposed quality metric-based augmentation method, followed by experimental setup and results in Section 3. Finally, Section 4. concludes the paper with a discussion of future work and practical implications.

2. Proposed Method

In this study, we propose a domain-specific augmentation technique tailored for thermal images, particularly addressing the challenges presented by the unique characteristics of infrared data. Traditional image augmentation methods, such as geometric transformations and basic image enhancement, often fail to adequately improve thermal image datasets. This is primarily due to the inherently low contrast and distinct noise patterns in thermal images, which differ significantly from those in visible spectrum images. As a result, conventional approaches do not fully capture the structural and qualitative nuances required for effective thermal image classification.

To overcome these limitations, we leverage the Block-wise Image Entropy (BIE) quality metric [17], a no-reference Image Quality Assessment (IQA) technique. No-reference IQA methods are particularly important in scenarios where ground truth image quality ratings are unavailable, as they provide an objective measure of image qualtrongtirakul2022unsupervisedity without the need for reference images [18, 19]. This is crucial in thermal imaging, where acquiring high-quality, well-labeled datasets is often difficult and costly [20]. By incorporating both local and global entropy characteristics, BIE offers a more comprehensive assessment of thermal image quality than traditional methods. It evaluates the image based on entropy-driven criteria, which better aligns with the noise characteristics and structural patterns of thermal data, thus ensuring that the augmented images maintain high levels of interpretability and detail.

The BIE is defined as:

$$BIE(I) = ADP(I) \times \frac{\frac{1}{n}\sum_{k=1}^{n}(\alpha M'(I^k)^{\alpha} \times \ln M'(I^k))}{1 + \frac{1}{n}\sum_{k=1}^{n} E(I^k)} \times \frac{SD(I)}{1 + \frac{1}{n}\sum_{k=1}^{n} SD(I^k)}, \tag{1}$$

where n represents the number of blocks into which the image is divided, $E(I^k)$ is the entropy of each block, and $SD(I^k)$ is the standard deviation of each block. The term $ADP(I)$ denotes the Average Deviation Percentage, and $M'(I)$ is the modified modulation. These are defined as:

$$ADP(I) = 1 - \frac{|A(I) - L/2|}{L/2}, \quad M'(I) = \frac{I_{\max} - I_{\min}}{L}, \tag{2}$$

where $A(I)$ is the average pixel value of the image, $L = 255$ represents the maximum pixel value in an 8-bit image (adjustable depending on the image format), and $I_{\max}$, $I_{\min}$ are the maximum and minimum pixel intensities, respectively. The entropy $E(I)$ is calculated using Shannon entropy [21], which is defined as:

$$E(I) = -\sum_{i=1}^{N} P(i)\log_2 P(i), \tag{3}$$

where $E(I)$ represents the entropy of the image I, N denotes the total number of possible intensity levels, and $P(i)$ is the probability of occurrence of intensity level i within the image. Shannon entropy provides a measure of information content and randomness in the image, making it a crucial component of the BIE metric for evaluating thermal image quality.

Measure-based enhancement techniques have been widely employed in the literature to improve image quality and performance in machine learning tasks, particularly in scenarios

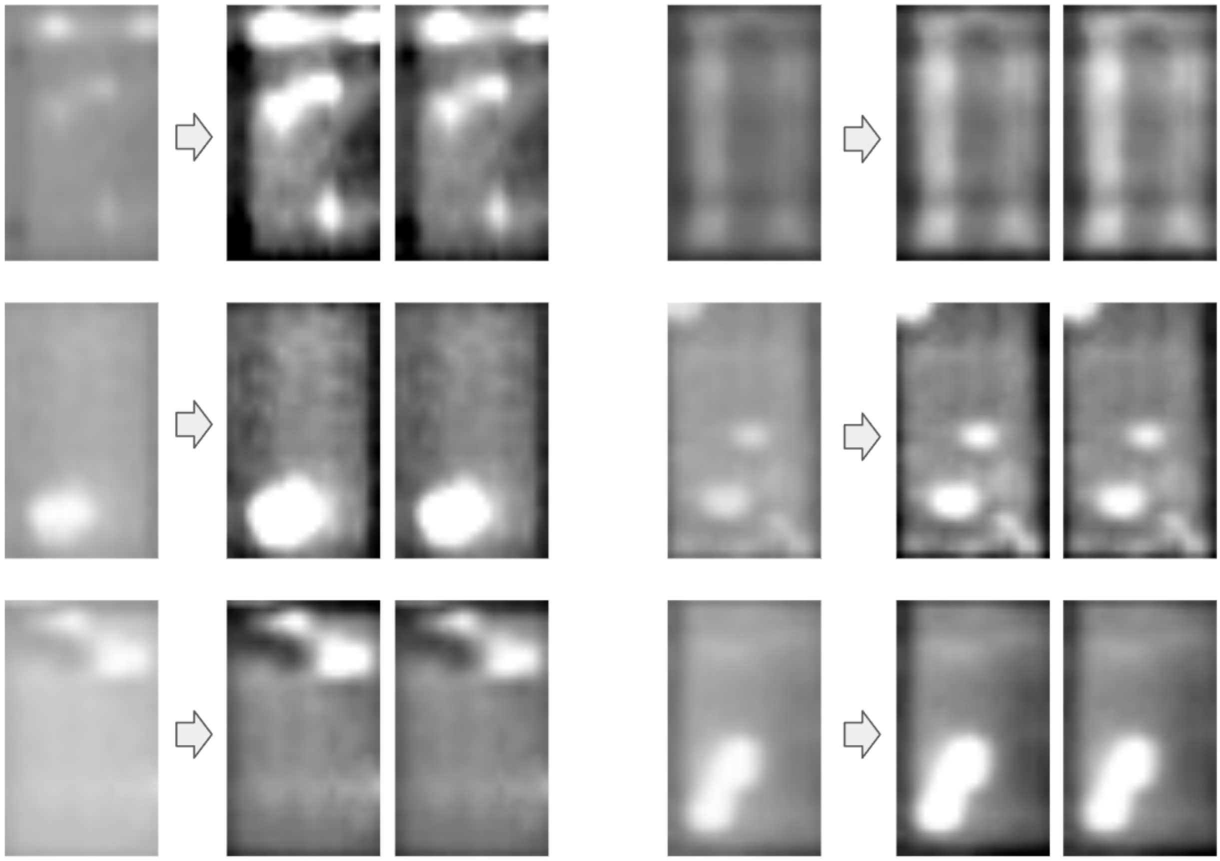

Fig. 2. Example of BIE-based contrast enhancement on thermal images of defective PV modules.

where image contrast and structural detail are crucial [22, 23, 24]. To augment the thermal images, we apply a multi-step process based on contrast enhancement using the BIE quality metric:

1. **Contrast Enhancement:** Each thermal image undergoes parametric contrast stretching to enhance the overall contrast. The stretching parameters are selected from predefined ranges.

2. **Optimization Using BIE:** The low and high stretching parameters are optimized within the ranges $[0, 150]$ and $[150, 255]$, respectively, using the BIE quality measure. The goal is to maximize the BIE value for each image.

3. **Augmentation with Best and Second-Best Images:** The images with the highest BIE value and the second-highest BIE value are selected and added to the augmented dataset, alongside the original image.

This method generates augmented versions of thermal images that have improved contrast and higher entropy, which are better suited for training deep learning models.

Fig. 2. illustrates the results of our BIE-based contrast enhancement for a defective thermal image from the Infrared Solar Modules dataset [10]. The original image, along with the images corresponding to the best and second-best BIE values, are shown.

Table 2: Description of augmentation setups and sets

Augmentation	Description	Training set	Validation set	Test set
A0	No augmentation	16,000	2,000	2,000
A1	Geometric Transformations (GT)	64,000	8,000	8,000
A2	GT + Histogram Equalization	80,000	10,000	10,000
A3	GT + Brightness Adjustment	96,000	12,000	12,000
A4	GT + BIE-based Oversampling	80,000	10,000	10,000

The results highlight how BIE-guided contrast enhancement significantly improves image quality by enhancing local contrast and emphasizing key structural features in the thermal image.

3. Results

In this section, we present the performance results of our proposed augmentation technique and the deep learning models evaluated on the PV module defect classification task. The dataset consists of 20,000 infrared images of solar modules, evenly divided between non-defective (No-Anomaly) and defective modules. The defective class includes a variety of faults, such as Hot-Spots, Soiling, Diode failures, and Cell anomalies [10]. The deep learning models used in our experiments, including CNN-based architectures and a transformer-based model, are described in Table 1. As detailed in Table 3., The augmentation setups were applied to the original dataset, which was randomly split into 80% training, 10% validation, and 10% test sets. The number of samples in each set after augmentation is also listed in Table 3..

For training the models, we used the following hyperparameters. Each model was trained for 30 epochs using the Cross-entropy loss function. We optimized the models using the stochastic gradient descent (SGD) optimizer with an initial learning rate of 0.001 and momentum set to 0.9. The batch size for all experiments was set to 32. To enhance training efficiency and mitigate overfitting, we employed a learning rate scheduler, specifically the StepLR, with a step size of 10 and a gamma of 0.5 to reduce the learning rate after every 10 epochs.

Additionally, the input images were resized to meet the requirements of each network before training and validation. For instance, images were resized to 227x227 for AlexNet and 224x224 for ResNet50.

To evaluate the performance of the models, we used four standard classification metrics: accuracy (Acc), precision (Pr), recall (Rec), and specificity (Sp). These metrics were calculated based on the following equations:

$$\text{Accuracy (Acc)} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$\text{Precision (Pr)} = \frac{TP}{TP + FP} \tag{5}$$

$$\text{Recall (Rec)} = \frac{TP}{TP + FN} \tag{6}$$

Table 3: Quantitative results for each deep learning model trained on five different augmentation techniques. The table presents accuracy, precision, recall, and specificity for both the test and validation sets.

	Dataset	Test				Validation			
	Aug Method	Acc	Pr	Rec	Sp	Acc	Pr	Rec	Sp
AlexNet	A0	86.85	88.83	86.90	93.31	85.80	88.26	85.81	92.79
	A1	90.88	91.85	90.91	95.37	89.28	90.75	89.28	94.57
	A2	90.39	91.34	90.42	95.11	89.62	90.90	89.63	94.73
	A3	88.32	89.61	88.35	94.02	87.83	89.52	87.84	93.81
	A4	93.21	92.62	94.05	96.56	92.53	92.02	93.54	96.22
ResNet50	A0	89.70	89.92	89.72	94.61	89.30	89.89	89.31	94.45
	A1	92.20	92.28	92.21	95.95	91.41	91.56	91.42	95.54
	A2	91.99	92.01	91.99	95.83	91.28	91.36	91.28	95.45
	A3	88.33	88.48	88.35	93.83	87.38	87.64	87.38	93.31
	A4	93.66	93.06	94.12	96.75	93.16	92.57	93.76	96.51
SqueezeNet	A0	88.10	88.98	88.13	93.83	88.30	89.30	88.31	93.96
	A1	88.42	88.73	88.41	93.91	88.66	88.80	88.66	94.01
	A2	88.41	88.47	88.40	93.86	88.06	88.12	88.06	93.66
	A3	86.17	86.46	86.18	92.63	85.71	86.06	85.71	92.37
	A4	91.34	90.89	91.09	95.47	91.71	91.22	91.63	95.68
ShuffleNetV2	A0	89.95	90.19	89.97	94.75	90.30	90.77	90.31	94.98
	A1	92.35	92.39	92.36	96.03	91.75	91.90	91.75	95.72
	A2	91.44	91.57	91.45	95.55	90.54	90.73	90.54	95.07
	A3	90.69	90.83	90.70	95.14	89.96	90.20	89.96	94.75
	A4	93.72	93.19	93.90	96.77	93.07	92.51	93.36	96.43
MobileNetV3	A0	92.65	92.66	92.65	96.19	91.20	91.26	91.20	95.41
	A1	93.04	93.05	93.03	96.39	92.26	92.27	92.26	95.98
	A2	92.40	92.41	92.40	96.05	92.06	92.09	92.06	95.87
	A3	91.72	91.80	91.73	95.70	91.42	91.52	91.43	95.54
	A4	94.59	94.18	94.64	97.23	94.08	93.59	94.27	96.96
Swin	A0	88.85	88.91	88.84	94.10	88.90	88.90	88.90	94.12
	A1	91.69	91.82	91.68	95.68	91.38	91.41	91.37	95.50
	A2	91.60	91.74	91.59	95.64	90.96	90.98	90.96	95.27
	A3	89.59	89.63	89.60	94.52	89.41	89.54	89.41	94.43
	A4	93.85	93.38	93.95	96.84	93.67	93.13	93.94	96.75

$$\text{Specificity (Sp)} = \frac{TN}{TN + FP} \tag{7}$$

In these equations, TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives.

The results for each deep learning model, trained using five different augmentation techniques, are summarized in Table 3. The confusion matrices for AlexNet and Swin Transformer are shown in Fig. 3.

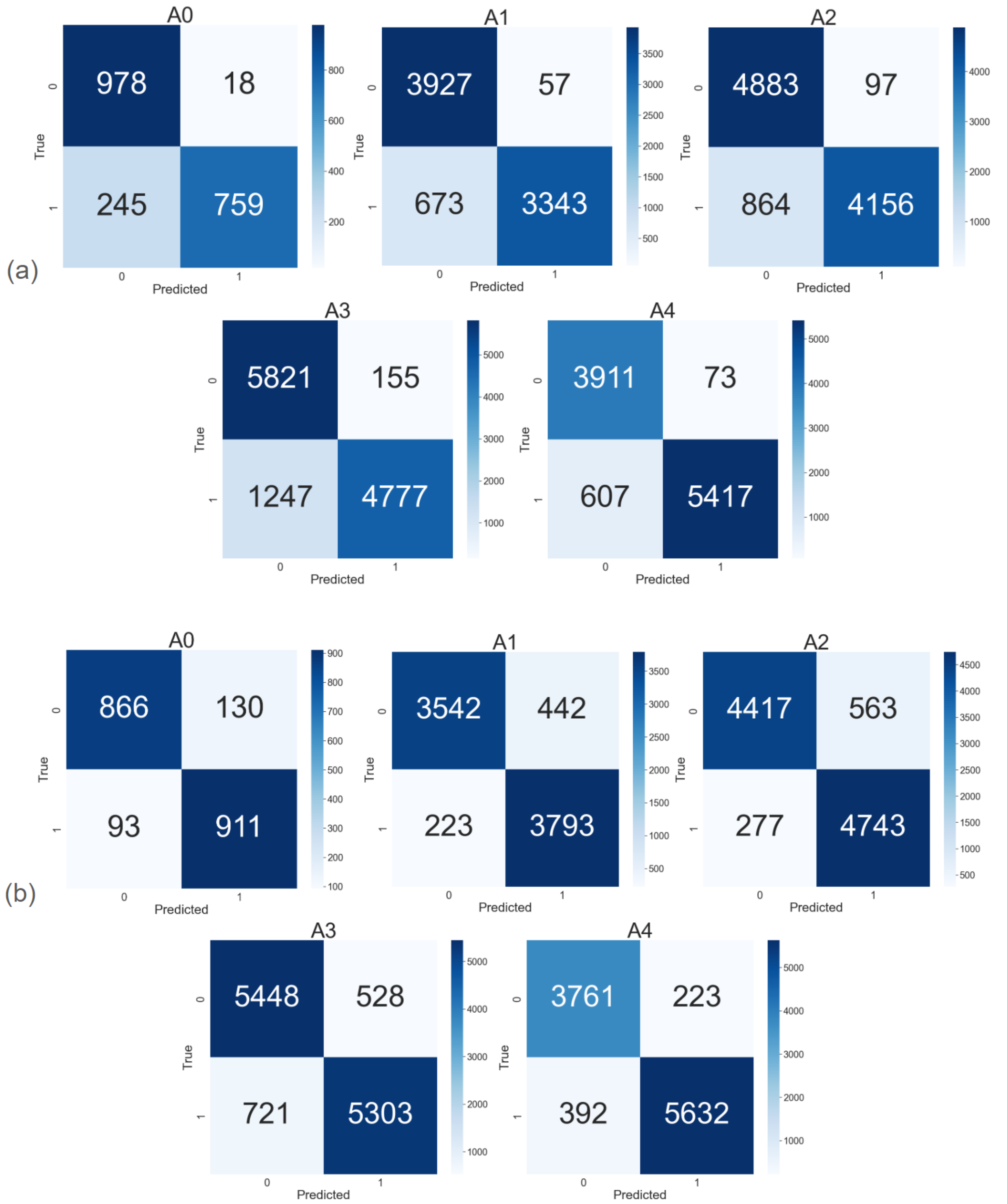

Fig. 3. Confusion matrices for binary classification using (a) AlexNet and (b) Swin Transformer, showcasing performance across different augmentation types on the test set.

As observed, the setups without augmentation and those using only brightness adjustment yielded the lowest performance metrics across all models. The setups employing geometric transformations and histogram equalization achieved moderate scores, indicating some improvement but still lacking compared to other methods.

However, our proposed contrast enhancement augmentation, combined with geometric transformations, outperformed all other approaches. This setup achieved the best performance, particularly on MobileNetV3, with an accuracy of 94.59%, precision of 94.18%, recall of 94.64%, and a specificity of 97.23% on the test set. Swin Transformer also performed well with our augmentation technique, though its results were less impressive than those of other augmentation setups.

In conclusion, the proposed contrast enhancement method not only improves model performance but demonstrates its effectiveness in enhancing the robustness and classification accuracy of thermal image datasets, particularly when data availability is limited.

4. Conclusion

This paper presented a novel augmentation technique for thermal image classification, particularly for fault detection in photovoltaic modules. Given the challenges of limited thermal image datasets, traditional augmentation methods like geometric transformations and simple enhancements are often inadequate. To address this, we proposed a new augmentation method that uses thermal quality assessment-based contrast enhancement to enrich the dataset. Our approach significantly improves the diversity of the training dataset, leading to enhanced classification accuracy and robustness across several state-of-the-art models. The results demonstrate the effectiveness of the new augmentation over conventional methods, making it a valuable tool for thermal imaging applications with constrained data.

Future work will focus on improving classification accuracy by exploring thermal-specific deep-learning models and further refining our augmentation technique to optimize performance.

References

[1] M. W. Akram, G. Li, Y. Jin, X. Chen, C. Zhu, X. Zhao, M. Aleem, and A. Ahmad, “Improved outdoor thermography and processing of infrared images for defect detection in pv modules,” *Solar Energy*, vol. 190, pp. 549–560, 2019.

[2] R. H. F. Alves, G. A. de Deus Junior, E. G. Marra, and R. P. Lemos, “Automatic fault classification in photovoltaic modules using convolutional neural networks,” *Renewable Energy*, vol. 179, pp. 502–516, 2021.

[3] T. Trongtirakul, S. Agaian, and A. Oulefki, “Automated tumor segmentation in thermographic breast images,” *Mathematical Biosciences and Engineering*, vol. 20, no. 9, pp. 16786–16806, 2023.

[4] M. d. F. O. Baffa and L. G. Lattari, “Convolutional neural networks for static and dynamic breast infrared imaging classification,” in *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 174–181, IEEE, 2018.

[5] D. Korkmaz and H. Acikgoz, “An efficient fault classification method in solar photovoltaic modules using transfer learning and multi-scale convolutional neural network,” *Engineering Applications of Artificial Intelligence*, vol. 113, p. 104959, 2022.

[6] Z. Yang, R. O. Sinnott, J. Bailey, and Q. Ke, "A survey of automated data augmentation algorithms for deep learning-based image classification tasks," *Knowledge and Information Systems*, vol. 65, no. 7, pp. 2805–2861, 2023.

[7] W. Li, C. Chen, M. Zhang, H. Li, and Q. Du, "Data augmentation for hyperspectral image classification with deep cnn," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 4, pp. 593–597, 2018.

[8] R. F. Pamungkas, I. B. K. Y. Utama, and Y. M. Jang, "A novel approach for efficient solar panel fault classification using coupled udensenet," *Sensors*, vol. 23, no. 10, p. 4918, 2023.

[9] C. Tang, H. Ren, J. Xia, F. Wang, and J. Lu, "Automatic defect identification of pv panels with ir images through unmanned aircraft," *IET Renewable Power Generation*, vol. 17, no. 12, pp. 3108–3119, 2023.

[10] M. Millendorf, E. Obropta, and N. Vadhavkar, "Infrared solar module dataset for anomaly detection," in *Proc. Int. Conf. Learn. Represent*, 2020.

[11] A. Krizhevsky, "One weird trick for parallelizing convolutional neural networks," *arXiv preprint arXiv:1404.5997*, 2014.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[13] Z. Liu, H. Hu, Y. Lin, Z. Yao, Z. Xie, Y. Wei, J. Ning, Y. Cao, Z. Zhang, L. Dong, *et al.*, "Swin transformer v2: Scaling up capacity and resolution," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12009–12019, 2022.

[14] F. N. Iandola, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.

[15] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 116–131, 2018.

[16] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, *et al.*, "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 1314–1324, 2019.

[17] H. Ayunts, A. Grigoryan, and S. Agaian, "Novel entropy for enhanced thermal imaging and uncertainty quantification," *Entropy*, vol. 26, no. 5, p. 374, 2024.

[18] K. Panetta, C. Gao, and S. Agaian, "No reference color image contrast and quality measures," *IEEE transactions on Consumer Electronics*, vol. 59, no. 3, pp. 643–651, 2013.

[19] H. Ayunts and S. Agaian, "No-reference quality metrics for image decolorization," *IEEE Transactions on Consumer Electronics*, vol. 69, no. 4, pp. 1177–1185, 2023.

[19] H. Ayunts and S. Agaian, "No-reference quality metrics for image decolorization," *IEEE Transactions on Consumer Electronics*, vol. 69, no. 4, pp. 1177–1185, 2023.

[20] T. Trongtirakul and S. Agaian, "Unsupervised and optimized thermal image quality enhancement and visual surveillance applications," *Signal Processing: Image Communication*, vol. 105, p. 116714, 2022.

[21] P. Bromiley, N. Thacker, and E. Bouhova-Thacker, "Shannon entropy, renyi entropy, and information," *Statistics and Inf. Series (2004-004)*, vol. 9, no. 2004, pp. 2–8, 2004.

[22] 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.

[23] A. M. Grigoryan and S. S. Agaian, "Image processing contrast enhancement," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–22, 1999.

[24] A. M. Grigoryan, A. John, and S. S. Agaian, "Color image enhancement of medical images using alpha-rooting and zonal alpha-rooting methods on 2d qdft," in *Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment*, vol. 10136, pp. 325–333, SPIE, 2017.

Ջերմային պատկերների դասակարգման բարելավում որակի չափորոշիչների վրա հիմնված տվյալների ավելացման նոր մեթոդով

Հրաչ Յու. Այունց

Երևանի պետական համալսարան, Երևան, Հայաստան
e-mail: e-mail: hrach.ayunts@ysu.am

Ամփոփում

Ջերմային պատկերների դասակարգումը կարևոր նշանակություն ունի տարբեր ոլորտներում, մասնավորապես, անսարքությունների հայտնաբերման և մոնիտորինգի համակարգերում, ինչպիսիք են ֆոտոգալվանային (ՖՎ) մոդուլները: Այնուամենայնիվ, այս ոլորտներում ընդհանուր մարտահրավերը լայնածավալ, անոտացված ջերմային պատկերների տվյալների հավաքածուների սահմանափակ հասանելիությունն է: Այս խնդիրը լուծելու համար մեքենայական ուսուցման մեջ լայնորեն ընդունված է պատկերների ավելացումը՝ արհեստականորեն մեծացնելու տվյալների հավաքածուների չափն ու բազմազանությունը՝ բարելավելով մոդելների ընդհանրացումը: Ավանդական մեծացման մեթոդները, ինչպիսիք են երկրաչափական փոխակերպումները, տալիս են որոշ առավելություններ, սակայն կարող են չարտացոլել ջերմային պատկերներին բնորոշ յուրահատուկ հատկանիշները, որոնք հաճախ ունեն ավելի ցածր կոնտրաստ և տարբեր աղմուկի մոդելներ, քան տեսանելի սպեկտրի պատկերները: Այսպիսով, մենք պնդում ենք, որ պետք է մշակվի ջերմային պատկերների ավելացման նոր մեթոդ, քանի որ ջերմային տվյալների հավաքագրումը ծախսատար է և ժամանակատար:

Մեր հետազոտությունն առաջարկում է տվյալների ավելացման նոր տեխնիկա, որն առաջնորդվում է որակի չափորոշիչներով՝ ջերմային պատկերների երկուական դասակարգման մոդելների արդյունավետությունը բարձրացնելու համար: Օգտագործելով տիրույթին հատուկ որակի չափորոշիչներ, ինչպիսիք են պատկերի պարզությունը, ջերմային կոնտրաստը և աղմուկի մակարդակները, մենք օպտիմիզացնում ենք ջերմային տվյալների հավաքածուների հարստացման գործընթացը: Օրինակ, x քանակով պատկերների տվյալների բազայից սկսած՝ մենք ստեղծում ենք y լրացուցիչ ջերմային պատկերներ, ինչի արդյունքում ընդհանուր առմամբ $x + y$ քանակով պատկերներ են օգտագործվում խորը ուսուցման դասակարգման մոդելը ուսուցանելու համար: Օգտագործելով ՖՎ մոդուլի թերությունների հավաքածուն, մենք ցուցադրում ենք մեր որակի չափման վրա հիմնված տվյալների ավելացման մեփոդի արդյունավետությունը մի քանի ժամանակակից պատկերների դասակարգման ցանցերում: Մեր մոտեցումը գերազանցում է ավանդական մեծացման մեթոդներին՝ կապված դասակարգման ճշգրտության և կայունության հետ, ներառյալ երկրաչափական փոխակերպումները և պատկերի բարելավման ստանդարտ տեխնիկաները: Մեր հետազոտության գործնական հետևանքները նշանակալի են, քանի որ այն ապահովում է ավելի արդյունավետ միջոց ջերմային պատկերների խնդիրներում մոդելների կատարողականությունը բարելավելու համար, հիմնականում, երբ տվյալների հասանելիությունը սահմանափակ է:

Բանալի բառեր՝ տվյալների հավաքածուի մեծացում, պատկերի որակի գնահատում, ջերմային պատկերների դասակարգում:

Улучшение классификации тепловых изображений с использованием новых техник аугментации на основе метрик качества

Грач Ю. Аюнц

Ереванский государственный университет, Ереван, Армения
e-mail: hrach.ayunts@ysu.am

Аннотация

Классификация тепловых изображений имеет решающее значение в различных приложениях, особенно в системах обнаружения неисправностей и мониторинга, таких как фотоэлектрические (PV) модули. Однако распространенной проблемой в этих областях является ограниченная доступность крупномасштабных аннотированных наборов данных тепловых изображений. Чтобы решить эту проблему, в машинном обучении широко применяется дополнение изображений для искусственного увеличения размера и разнообразия обучающих наборов данных, что улучшает обобщение моделей. Традиционные методы дополнений, такие как геометрические преобразования, обеспечивают некоторые преимущества, но могут не полностью улавливать уникальные характеристики, присущие тепловизионным изображениям, которые часто имеют более низкую контрастность и другие шумовые паттерны, чем

изображения видимого спектра. Поэтому мы утверждаем, что нам необходимо разработать новый метод дополнений для тепловизионных изображений, где сбор данных является дорогостоящим и отнимает много времени.

Наше исследование предлагает новый метод дополнения, руководствующийся показателями качества, для повышения производительности моделей бинарной классификации тепловых изображений. Используя специфичные для домена показатели качества, такие как четкость изображения, тепловой контраст и уровни шума, мы оптимизируем процесс дополнения для тепловых наборов данных. Например, начиная с набора данных из x изображений, мы генерируем y дополнительных тепловых изображений, в результате чего получается в общей сложности $x + y$ изображений, используемых для обучения фреймворка классификации глубокого обучения. Используя набор данных дефектов фотоэлектрических модулей, мы демонстрируем эффективность нашей стратегии аугментации на основе метрики качества в нескольких современных сетях классификации изображений. Наш подход превосходит традиционные методы аугментации с точки зрения точности и надежности классификации, включая геометрические преобразования и стандартные методы улучшения изображений. Практические последствия нашего исследования значительны, поскольку оно обеспечивает более эффективный и действенный способ улучшения производительности модели для задач тепловидения, в основном, когда доступность данных ограничена

Ключевые слова: аугментация набора данных, оценка качества изображений, классификация тепловых изображений.