

ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ ԻՆՍՏԻՏՈՒՏ

Այունգ Հրաչ Յուրիի

ՊԱՏԿԵՐՆԵՐԻ ՄՇԱԿՄԱՆ ՄԵԹՈԴՆԵՐԻ ՕՊՏԻՄԻԶԱՑԻՄ ԵՎ
ԿԻՐԱՌՈՒԹՅՈՒՆՆԵՐ

Ե.13.05 - «Մաթեմատիկական մոդելավորում, թվային մեթոդներ և ծրագրերի համալիրներ»
մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի համար

ՍԵՂՄԱԳԻՐ

Երևան 2025

INSTITUTE FOR INFORMATICS AND AUTOMATION PROBLEMS OF THE NAS RA

Ayunts Hrach

OPTIMIZING IMAGE PROCESSING METHODS WITH APPLICATIONS

SYNOPSIS

of the dissertation for obtaining a Ph.D. degree in Technical Sciences on specialty 05.13.05
“Mathematical modelling, numerical methods and program complexes”

Yerevan 2025

Ատենախոսության թեման հաստատվել է ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում:

Գիտական ղեկավար՝ ֆիզ. մաթ. գիտ. դոկտոր Ս.Ս. Աղայան
Պաշտոնական ընդդիմախոսներ՝ տեխ. գիտ. դոկտոր Դ.Գ. Ասատրյան
տեխ. գիտ. թեկնածու Վ.Կ. Ավետիսյան
Առաջատար կազմակերպություն՝ Հայ-ռուսական համալսարան

Ատենախոսության պաշտպանությունը կայանալու է 2025թ. հուլիսի 8-ին, ժամը 14:00-ին՝ ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում գործող 037 «Ինֆորմատիկա» մասնագիտական խորհրդի նիստում հետևյալ հասցեով՝ Երևան, 0014, Պ. Սևակ փող. 1:

Ատենախոսությանը կարելի է ծանոթանալ ՀՀ ԳԱԱ ԻԱՊԻ գրադարանում:

Սեղմագիրն առաքվել է 2025թ. -ի հունիսի 7-ին:

Մասնագիտական խորհրդի գիտական
քարտուղար ֆիզ. մաթ. գիտ. դոկտոր՝



Մ.Ե. Հարությունյան

The topic of the dissertation was approved at the Institute for Informatics and Automation Problems of NAS RA.

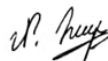
Scientific supervisor: S.S. Agaian, D.Ph.M.S.
Official opponents: D.G. Asatryan, D.Tech.S.
V.K. Avetisyan, Ph.D.
Leading organization: Russian-Armenian University

The Defense will take place on 8 July 2025, at 14:00, at the Specialized Council 037 “Informatics” at the Institute for Informatics and Automation Problems of NAS RA.
Address: Yerevan, 0014, P. Sevak 1.

The Dissertation is available at the library of IIAP NAS RA.

The abstract is delivered on 7 June, 2025.

Scientific Secretary of the Specialized Council, D.Ph.M.S.



M.E. Haroutunian

General Description of the Work

Relevance of the Research

The rapid adoption of solar photovoltaic (PV) technology as a key enabler of global renewable energy strategies has created a pressing need for reliable, large-scale fault detection. Micro-cracks, interconnect corrosion, delamination, and thermal hotspots in PV modules reduce energy output and long-term system performance. Thermal imaging offers a non-contact means of identifying these issues through the detection of infrared radiation, but its use presents several challenges. These include low contrast, high noise levels, and limited spatial resolution, especially in imagery captured by embedded devices and unmanned aerial vehicles (UAVs). Conventional image processing tools developed for the visible spectrum are typically ill-suited for thermal data, resulting in poor generalization and limited diagnostic accuracy. Manual inspection remains labor-intensive and impractical for large PV arrays. These challenges underscore the need for efficient, thermal-specific image analysis tools that are scalable, robust, and compatible with real-time deployment in operational environments.

The absence of objective and perceptually consistent criteria for evaluating grayscale conversions complicates the use of thermal data in automated pipelines. Without reliable quality assessment tools, preprocessing decisions are often based on heuristics, undermining downstream analysis. Furthermore, thermal images often suffer from inconsistent visibility and uneven contrast due to environmental factors, making adaptive enhancement a non-trivial requirement for meaningful interpretation.

In addition to image quality issues, the scarcity and imbalance of labeled thermal datasets restrict the performance of learning-based methods, especially in edge-case fault scenarios. Standard augmentation strategies fail to reflect the statistical and perceptual characteristics of thermal data, limiting their ability to support robust model training. Finally, many existing deep learning models are over-parameterized for embedded applications, where memory and computation budgets are tightly constrained, yet high classification accuracy remains critical for operational viability.

Aim of the Work and Key Objectives

The aim of the work is to develop an end-to-end image processing framework that improves the accuracy and reliability of thermal defect detection, particularly in PV systems, while remaining suitable for deployment on low-power and embedded platforms.

To achieve this goal, the research is structured around the following objectives:

1. Define perceptually meaningful metrics for evaluating the quality of grayscale image conversion, enabling objective assessment without reliance on ground-truth references.
2. Develop techniques for thermal contrast enhancement that improve visibility of structural details and accommodate image-specific variations in quality.
3. Improve training data quality through augmentation strategies that generate representative, high-contrast samples, with special emphasis on underrepresented fault types.
4. Design lightweight neural network architectures optimized for the classification of low-resolution thermal images acquired in real-world conditions.
5. Integrate and validate all components of the framework on diverse thermal datasets representative of solar, industrial, and aerial inspection domains.

Research Objects / Subject of the Research

The object of this research is thermal and color image data used in imaging-based monitoring

systems across renewable energy and industrial domains. This includes visual and thermal imagery captured from photovoltaic modules, wind turbine blades, and electrical equipment such as transformers and motors, typically obtained via UAV platforms or embedded sensing devices.

The subject of the research is the study and development of image processing methods aimed at improving the quality and interpretability of thermal and color image data for downstream tasks such as enhancement, analysis, and classification. Emphasis is placed on the creation of efficient, scalable, and interpretable processing pipelines that are compatible with resource-constrained environments and capable of handling image imperfections common in real-world deployments.

Research Methods

The methodological approach integrates theoretical analysis, algorithmic modeling, and empirical validation. It involves the study of image quality assessment, enhancement strategies, and classification techniques tailored to thermal imaging conditions. The research utilizes both analytical and data-driven methods to evaluate visual quality, optimize image preprocessing, and design compact machine learning models suitable for edge deployment.

Experimental studies are conducted using image datasets from various thermal imaging scenarios, and the proposed methods are validated using standard performance metrics such as classification accuracy, precision, recall, specificity, and computational efficiency. The methods are assessed for their ability to generalize across diverse operational settings and contribute to robust image-based diagnostics.

Scientific Novelty of the Work

The scientific contributions of this dissertation include:

- The introduction of two novel no-reference quality metrics, TIA and WTIA, for perceptual evaluation of grayscale conversion without access to reference images.
- The proposal of BIE, a thermal-specific entropy metric that integrates global and local image properties to improve visibility and guide enhancement.
- The development of SlantNet, a lightweight convolutional neural network incorporating harmonic slant convolutions for efficient classification on embedded devices.
- The formulation of image enhancement and decolorization as optimization problems, enabling automated parameter selection using perceptual metrics and metaheuristic algorithms.

Practical Significance of the Work

This research enables the creation of scalable and energy-efficient solutions for thermal image analysis in real-world environments. Its practical significance is reflected in the following outcomes:

- Real-time PV fault detection using lightweight, interpretable models compatible with UAVs and edge devices.
- Quality-driven image preprocessing using robust no-reference metrics such as TIA, WTIA, and BIE, improving contrast and structure visibility under noise and resolution constraints.
- Efficient augmentation pipelines that selectively generate high-quality synthetic training samples, reducing annotation needs and improving classifier generalization.
- An integrated image processing pipeline evaluated across Infrared Solar Modules and Thermal Objects datasets, demonstrating robust performance in both enhancement and classification tasks.

- Public release of the augmented PV dataset¹, supporting reproducibility and adoption in academic and industrial settings.

Potential Applications

Although focused on PV systems, the proposed framework is applicable to a broad range of thermal vision tasks, including:

- Pedestrian and vehicle surveillance, where improved thermal contrast aids night-time and low-light detection.
- Industrial equipment monitoring, including motors and transformers, where early thermal anomaly detection supports predictive maintenance.
- Medical thermography, where contrast-sensitive preprocessing improves abnormality detection in diagnostic screening.
- Wind turbine blade inspection using UAV-mounted sensors, enabling early fault detection with minimal human intervention.

Publications

The results of the dissertation have been published in 4 scientific articles, 3 of which are indexed in international databases such as Web of Science and Scopus. The full list of publications is provided at the end of the abstract.

Scope and Structure of the Dissertation

The dissertation comprises 126 pages and includes an introduction, four main chapters, a conclusion, and a reference section with 134 bibliographic entries.

Content of the Dissertation

Introduction is the first chapter of the dissertation and presents the motivation, research context, problem formulation, and overall structure of the work.

In Chapter 2, the proposed threshold-independent quality assessment framework for image decolorization is presented. The chapter introduces novel evaluation metrics designed to address the limitations of existing methods, particularly their reliance on user-defined thresholds and lack of alignment with human perception.

Section 2.1 introduces the problem of image decolorization, highlighting the importance of preserving color contrast and structural content during grayscale conversion. It motivates the need for robust no-reference quality metrics, especially in applications where no ground-truth grayscale reference is available. The challenge lies in designing evaluation methods that account for perceptual contrast loss and structural degradation, without depending on subjective human feedback.

Section 2.2 provides an in-depth review of related work. Traditional grayscale conversion techniques apply fixed linear combinations of RGB values, such as:

$$g = aR + bG + cB, \quad (1)$$

where a , b , and c are fixed weights, e.g., in the Luminosity method $g = 0.21R + 0.72G + 0.07B$. While computationally efficient, these approaches often fail to preserve chromatic contrast and per-

¹<https://github.com/HrachA/augmented-infrared-solar-modules-set/tree/main>

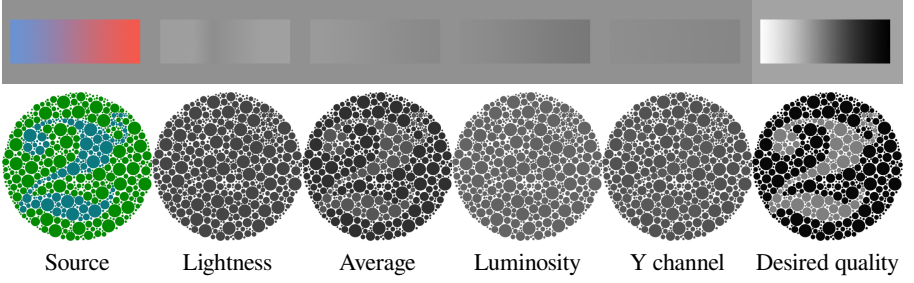


Figure 1: Comparison of linear grayscale conversion methods. Decolorized images can lose the contrast and become hardly visible.

ceptual salience (see Fig. 1).

More sophisticated methods include chrominance-aware techniques and energy minimization approaches, which attempt to retain color edges or adapt the grayscale output based on visual models. Neural network-based methods also emerged, using saliency cues and deep representations to improve decolorization.

As for evaluation, existing no-reference metrics include the Color Contrast Preserving Ratio (CCPR) and Color Content Fidelity Ratio (CCFR). Combined, they form the E-score metric. These metrics depend heavily on a user-defined threshold τ , leading to inconsistent evaluations across different methods and datasets. Additionally, they fail to account for spatial saliency and do not generalize well across varying image content (Table 1).

Section 2.3 presents the proposed quality metrics: Threshold-Independent Area (TIA) and its weighted variant (WTIA). TIA addresses the instability of τ -dependent metrics by analyzing the E-score curve across multiple thresholds ($\tau = 2, \dots, 10$) and computing the area under a fitted regression line:

$$\text{TIA} = \max \left(\frac{2\alpha + \beta}{2}, 0 \right), \quad (2)$$

where α and β are the slope and intercept of the line $y = \alpha + \beta x$ approximating the E-score curve.

To better align with perceptual importance, WTIA incorporates visual attention using weighted E-score components:

$$\text{E-score}_w = \frac{2 \cdot \text{WCCPR} \cdot \text{WCCFR}}{\text{WCCPR} + \text{WCCFR}}, \quad (3)$$

where weights w_x, w_y are derived from saliency maps:

$$\text{WCCPR} = \frac{\sum w_x w_y | (x, y) \in \Omega, |g_x - g_y| \geq \tau}{\sum w_x w_y | (x, y) \in \Omega}, \quad (4)$$

$$\text{WCCFR} = 1 - \frac{\sum w_x w_y | (x, y) \in \Theta, \delta_{x,y} \leq \tau}{\sum w_x w_y | (x, y) \in \Theta}. \quad (5)$$

where $\delta_{x,y}$ is the CIE LAB color difference, g_x is the value of the x pixel after decolorization, Ω is the set of pixel pairs with $\delta_{x,y} \geq \tau$, Θ is the set of pixel pairs with $|g_x - g_y| > \tau$. This modification ensures that regions more important to human perception are emphasized, enhancing metric reliability. When all pixel weights are set to 1, the method behaves identically to the traditional unweighted

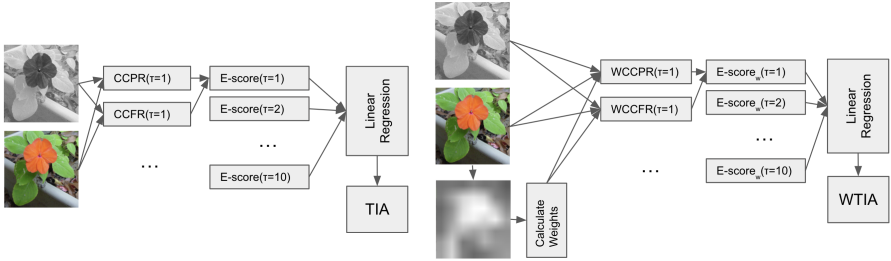


Figure 2: Workflow for proposed quality metrics: An overview of the sequential steps and stages involved in calculating the TIA and WTIA quality metrics

formulation, ensuring consistency and backward compatibility. By assigning continuous pixel-wise weights ranging from 0 to 1, the metric supports more nuanced assessments where perceptually significant details carry greater influence. The full pipeline is illustrated in Fig. 2.

This section also describes simulation studies validating the effectiveness of TIA and WTIA. Using Ćadík and COLOR250 datasets, correlation with human ratings was assessed using the Kendall rank coefficient R :

$$R = \frac{\#\{\text{concordant pair}\} - \#\{\text{discordant pair}\}}{\frac{1}{2}n(n-1)}, \quad (6)$$

TIA and WTIA showed significantly higher R values than E-score and its components, validating their perceptual alignment. As we can see in Table 1, the proposed metrics achieve the highest correlation with both accuracy and preference scores, confirming their superiority in reflecting human judgment. Furthermore, a genetic algorithm was applied to solve:

$$\max_{a,b,c} F(a, b, c), \quad (7)$$

where F is TIA or WTIA, and a, b, c are grayscale weights. This optimization yielded content-adaptive grayscale conversions that outperformed traditional fixed-weight approaches. The genetic algorithm is a population-based optimization method inspired by the process of natural selection. It begins with a randomly initialized population of candidate grayscale weight triplets (a, b, c) and iteratively evolves them to maximize the TIA or WTIA score. Each candidate’s fitness is evaluated by converting the image using its weights and measuring the resulting quality. The algorithm applies crossover and mutation to generate new candidates and selects the most promising ones for the next generation. The best-performing parameters are returned after a fixed number of iterations. The full procedure is described in Algorithm 1.

Section 2.4 presents the quantitative and visual evaluation of the proposed WTIA-based decolorization approach, benchmarking it against both classical and state-of-the-art methods using two standard datasets. The analysis highlights the limitations of traditional grayscale conversions when faced with perceptually challenging images, while demonstrating the robustness of the proposed metric-guided optimization. Figure 3 provides a visual comparison, showcasing how our method consistently preserves perceptual contrast better than other techniques.

Section 2.5 concludes the chapter by summarizing the key contributions. TIA and WTIA pro-

Table 1: Average Kendall correlation rank between metrics and user scores on Ćadík’s dataset (C) and the subset of it (C’)

Metric	Accuracy		Preference	
	C	C’	C	C’
CCPR _{$\tau=4$}	0.2341	0.2971	0.2222	0.2698
CCPR _{$\tau=5$}	0.2341	0.2925	0.2222	0.2562
CCPR _{$\tau=6$}	0.2222	0.2834	0.2183	0.2472
CCFR _{$\tau=4$}	0.2430	0.2210	0.2953	0.2763
CCFR _{$\tau=5$}	0.1950	0.2025	0.2626	0.2479
CCFR _{$\tau=6$}	0.2586	0.2616	0.3180	0.2977
E-score _{$\tau=3$}	0.4167	0.4376	0.4603	0.4558
E-score _{$\tau=4$}	0.4405	0.4603	0.4762	0.4785
E-score _{$\tau=5$}	0.4365	0.4512	0.4563	0.4603
E-score _{$\tau=6$}	0.4206	0.4376	0.4484	0.4467
E-score _{$\tau=7$}	0.4206	0.4376	0.4563	0.4558
TIS	0.1905	0.2517	0.2024	0.2245
TIA	0.4563	0.4785	0.4841	0.4875
WTIA	0.4802	0.5011	0.4921	0.5011

Algorithm 1 Optimal decolorization using Genetic Algorithm

Inputs: I_s = source image, I_w = weights image

Initialization: population = n , maximum number of iterations = N , $t = 0$

Function objective(a, b, c)

$I_g = \text{ConvertToGray}(a, b, c)$

WTIA = CalculateWTIA(I_s, I_g, I_w)

return WTIA

EndFunction

Generate the initial number of n chromosomes

Compute the fitness of each chromosome using the objective function

while $t < N$ **do**

 Select a pair of chromosomes based on fitness

 Apply crossover on selected pair

 Apply mutation operation

 Replace old population with newly generated one

$t \leftarrow t + 1$

end while

Return parameters with the best fitness

Output: a, b, c parameters

vide robust, threshold-free, and perceptually aligned evaluation tools for grayscale conversion. Their integration into optimization frameworks enables high-quality, adaptive decolorization across diverse applications, overcoming limitations of prior methods and advancing the field of image quality assess-

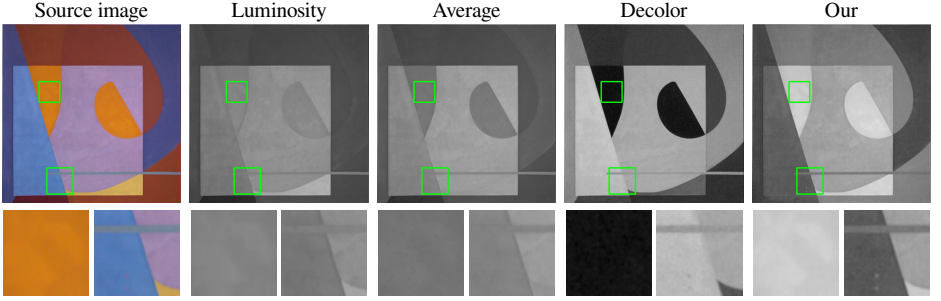


Figure 3: Results of different parametric decolorization methods. Our method preserves the small details and contrast.

ment.

In Chapter 3, the focus shifts to thermal imaging, where a novel entropy-based no-reference Image Quality Assessment (IQA) metric is proposed, aimed at addressing the limitations of existing enhancement and uncertainty quantification methods in infrared images. This chapter introduces Block-wise Image Entropy (BIE), a hybrid metric that combines local structural analysis with global contrast cues to evaluate and optimize the quality of thermal images under challenging conditions.

Section 3.1 introduces the role of thermal imaging across fields such as medicine, building diagnostics, and industrial maintenance, emphasizing the difficulty of processing noisy and low-contrast infrared images. It discusses how uncertainties, stemming from sensor noise, environmental influences, and the complex physics of heat transfer, lead to image artifacts and analysis errors. Despite the broad application of thermal imaging, traditional tools for uncertainty quantification remain underdeveloped. Quality metrics, particularly entropy-based approaches, are central to image evaluation and enhancement. However, conventional formulations fall short in thermal contexts. These limitations motivate the need for a new formulation that can more reliably assess the informational content of thermal images.

Section 3.2 reviews existing entropy-based and block-wise metrics. Shannon entropy,

$$E(I) = - \sum_{i=1}^N P(i) \log_2 P(i), \quad (8)$$

measures the global uncertainty of pixel intensity distribution of the I image, where $P(i)$ is the probability of the i -th intensity level, and N is the number of possible intensity levels. Rényi entropy,

$$R_\alpha(I) = \frac{1}{1-\alpha} \log_2 \left(\sum_{i=1}^N P(i)^\alpha \right), \quad (9)$$

generalizes this with a parameter α that adjusts sensitivity to pixel probability concentrations. Yet both fail to capture spatial structure and are prone to noise. Block-based metrics like EME and AME

offer localized evaluations:

$$EME(I) = \frac{1}{n} \sum_{k=1}^n \left(20 \ln \frac{I_{\max}^k}{I_{\min}^k + c} \right), \quad (10)$$

$$AME(I) = \frac{1}{n} \sum_{k=1}^n \alpha M(I^k)^\alpha \ln M(I^k), \quad M(I^k) = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min} + c}, \quad (11)$$

where n is the number of blocks, $I_{\max}^k$ and $I_{\min}^k$ are maximum and minimum intensities in block k , $\alpha = 1$ is a parameter, and c is a small constant to prevent division by zero. These metrics are often misled by noise, overstating quality in degraded images.

Section 3.3 introduces the Block-wise Image Entropy (BIE) metric, which integrates global contrast, block-wise entropy, and structural consistency. It is defined as:

$$BIE(I) = ADP(I) \times \frac{\frac{1}{n} \sum_{k=1}^n (\alpha M'(I^k)^\alpha \ln M'(I^k))}{1 + \frac{1}{n} \sum_{k=1}^n E(I_k)} \times \frac{SD(I)}{1 + \frac{1}{n} \sum_{k=1}^n SD(I_k)}, \quad (12)$$

where $M'(I^k)$ is the normalized modulation of block I^k , $E(I_k)$ is its Shannon entropy, and $SD(I)$ is standard deviation. The term $ADP(I)$ captures average deviation percentage:

$$ADP(I) = 1 - \frac{|A(I) - L/2|}{L/2}, \quad M'(I^k) = \frac{I_{\max} - I_{\min}}{L}, \quad (13)$$

where $A(I)$ is the image mean and L the dynamic range (typically 255). BIE penalizes uniform and noisy images while rewarding balanced contrast with perceptually meaningful variation.

Table 2 presents a thermal image alongside two distorted versions with identical histograms, along with their corresponding entropy-based metric values. While global metrics such as E , R_2 , and SD remain unchanged across all versions, the block-based BIE metric successfully detects the structural distortions. In contrast, EME and AME tend to increase in noisy cases, indicating their higher sensitivity to noise and reduced robustness in distinguishing perceptual quality.

Section 3.4 evaluates the BIE metric across several thermal datasets. Computer simulation results show that BIE yields consistent rankings for enhancement methods and correlates better with visual quality than AME or Shannon entropy. The section also introduces optimization frameworks using Genetic Algorithms (GA) and the Bat Algorithm (BA), with BIE as the objective function:

$$\max_{p_1, p_2, \dots, p_n} BIE(F(I_s, p_1, \dots, p_n)), \quad (14)$$

where F denotes the image enhancement function applied to the source image I_s , and $p_1, \dots, p_n$ are the tunable parameters of the enhancement method. The goal is to find the parameter set that maximizes the Block-wise Image Entropy (BIE), yielding optimal visual quality. For example, Figure 4 shows the optimization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm, which enhances image contrast by applying localized histogram equalization while limiting noise amplification. The parameters tuned include clip limit (CL) in the range $[1, 60]$ and grid size (GS) in $[4, 40]$. In the first case, both default settings and Shannon entropy-based optimization result in over-enhanced, noisy images. In contrast, BIE selects optimal parameters ($CL = 6$, $GS = 4$),

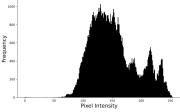
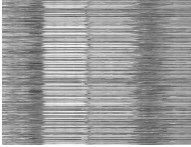
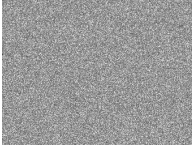
Histogram	Source Image	Shuffled Rows	Shuffled Pixels
			
E	7.202	7.202	7.202
R ₂	4.868	4.868	4.868
SD	39.89	39.89	39.89
EME*	9.892	18.12	28.63
AME*	0.289	0.345	0.315
BIE*	0.114	0.045	0.027

Table 2: Visual and quantitative comparison of a thermal image and its distorted versions. The histogram (leftmost column) is identical across all images, but entropy-based quality metrics distinguish the visual integrity. All block-based measures are calculated using *block_size* = 15.

producing visually superior results.

Finally, the chapter introduces a BIE-weighted image fusion model, which serves as an effective application of the thermal image quality measure:

$$I_f = \frac{\sum_{i=1}^N m_i I_i}{\sum_{i=1}^N m_i}, \quad (15)$$

where I_i are enhanced images and m_i their BIE scores. Figure 5 illustrates the results of this fusion process, showing the BIE scores for each input and the resulting fused image.

Section 3.5 concludes the chapter by summarizing the key findings: the BIE metric addresses limitations in traditional entropy and block-wise measures, offering a perceptually consistent and noise-robust quality criterion for thermal images. Its integration into parameter tuning and image fusion pipelines demonstrates utility across multiple image enhancement frameworks and datasets. BIE is shown to facilitate reliable thermal image assessment and optimization, advancing uncertainty quantification and visual clarity in critical infrared imaging tasks.

In Chapter 4, a novel thermal-specific augmentation strategy is introduced to address data scarcity challenges in fault classification tasks, particularly for photovoltaic (PV) modules. The chapter begins by discussing the limitations of conventional augmentation techniques when applied to thermal data, highlighting how unique infrared characteristics demand tailored strategies. It then presents the use of BIE, a no-reference thermal image quality metric, as the foundation for a metric-driven augmentation pipeline.

Section 4.1 provides background on thermal image classification and outlines the motivation for quality metric-based augmentation. It details the shortcomings of traditional augmentation methods, such as flips, brightness adjustments, and histogram equalization, when applied to thermal imagery, and reviews recent works that incorporate deep networks or GANs for thermal data expansion.

Section 4.2 introduces the proposed method in detail. It begins by defining the BIE metric, which incorporates both global and local image characteristics, making it well-suited for evaluating thermal

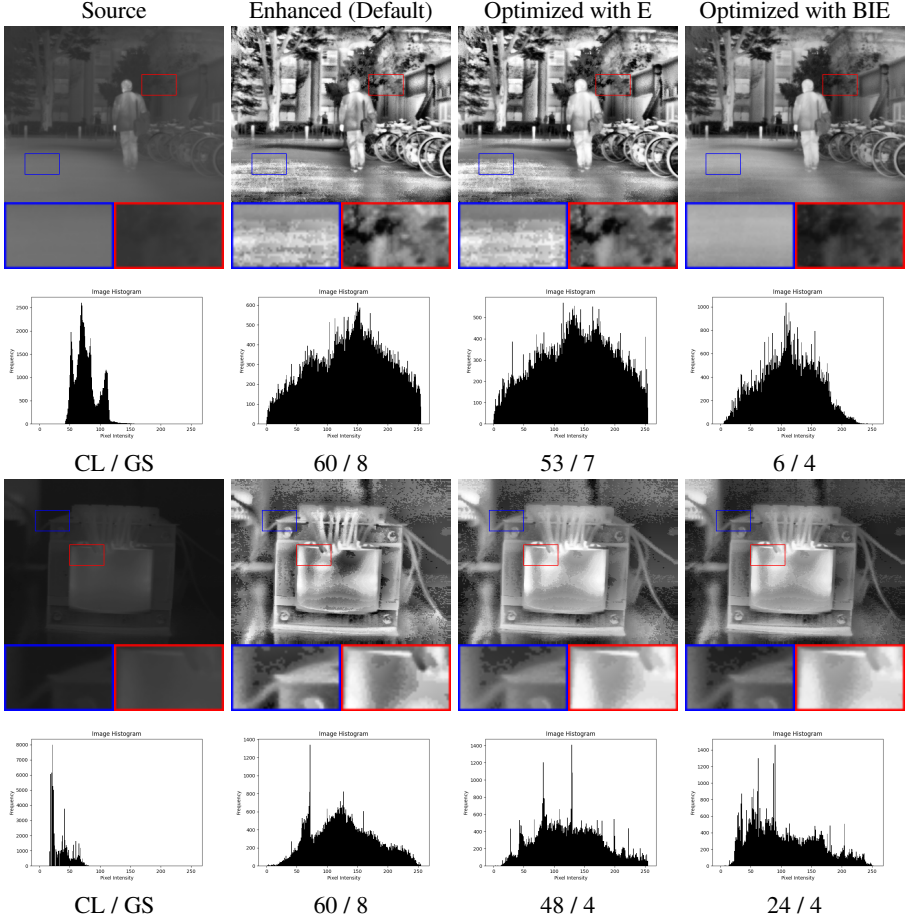


Figure 4: Results of the optimization of CLAHE algorithm (clip limit (CL) and grid size (GS) parameters) using E and BIE metrics.

image quality without the need for a reference image. This section explains how each thermal image is enhanced using parametric contrast stretching, with stretching limits optimized to maximize the BIE score. For each original image, the enhancement parameters that yield the top two BIE values are used to create two new augmented samples. These enhanced images, along with the original, are then used to expand the dataset. The process ensures that augmented samples are not arbitrary but are perceptually and structurally meaningful according to the thermal quality metric. Figure 6 illustrates this process by showing examples of an original image, its best BIE-enhanced version, and the second-best result. The section highlights how this augmentation technique generates thermally diverse, high-quality training data that improves downstream model performance while preserving essential fault-related features in PV modules.

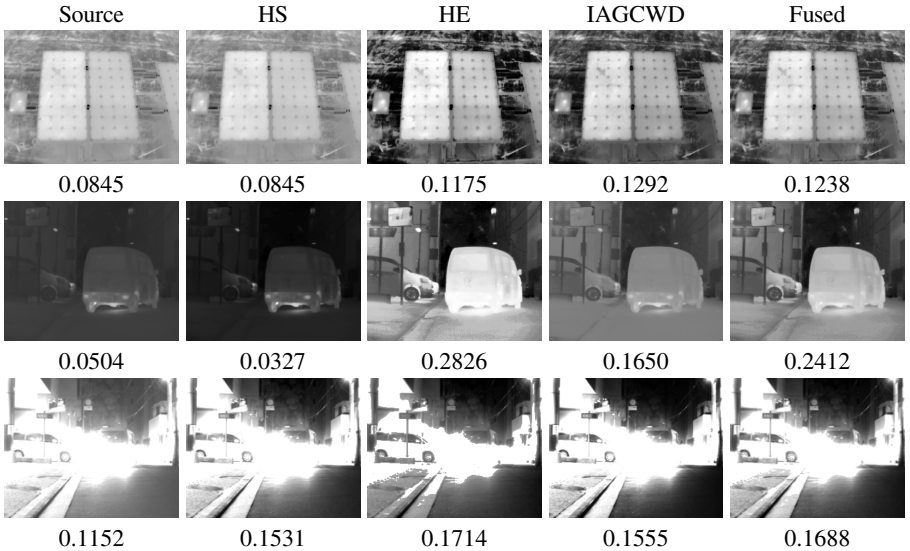


Figure 5: Metric-based fusion of image enhancement algorithms. The BIE metric values of each image are shown below the corresponding image. The enhancement methods include HS (Histogram Stretching), HE (Histogram Equalization), and IAGCWD (Improved Adaptive Gamma Correction with Weighted Distribution).

Section 4.3 presents the experimental setup, including datasets, training configurations, and the neural network architectures employed for evaluation, ranging from AlexNet to Swin Transformer. It compares performance metrics across several augmentation schemes, such as geometric, brightness-based, and BIE-based strategies, reporting improvements in accuracy, precision, recall, and specificity. The section notes particularly strong gains on lightweight networks like MobileNetV3, where contrast-aware augmentation significantly boosts generalization.

Section 4.4 concludes the chapter by reaffirming the practicality and effectiveness of the proposed technique. It outlines future directions, including the development of thermal-specific deep architectures and the refinement of augmentation policies for broader thermal imaging applications.

In Chapter 5, a lightweight neural network architecture called SlantNet is proposed for efficient and accurate classification of faults in thermal images of photovoltaic (PV) systems. The chapter introduces the need for computationally efficient models in large-scale solar installations and presents SlantNet as a solution combining Slant Convolutional layers with thermal-specific data augmentation to enable real-time inference and robust fault identification.

Section 5.1 introduces the motivation and challenges associated with PV system fault detection using thermal imagery. It highlights the limitations of manual inspection and traditional electrical testing, emphasizing the need for automated, scalable approaches. The section also outlines the opportunity to enhance classification models with directional and spectral feature sensitivity.

Section 5.2 provides technical background on image transforms, particularly the Slant Transform (SLT), and its relevance to thermal imaging. It explains how SLT is well-suited for encoding

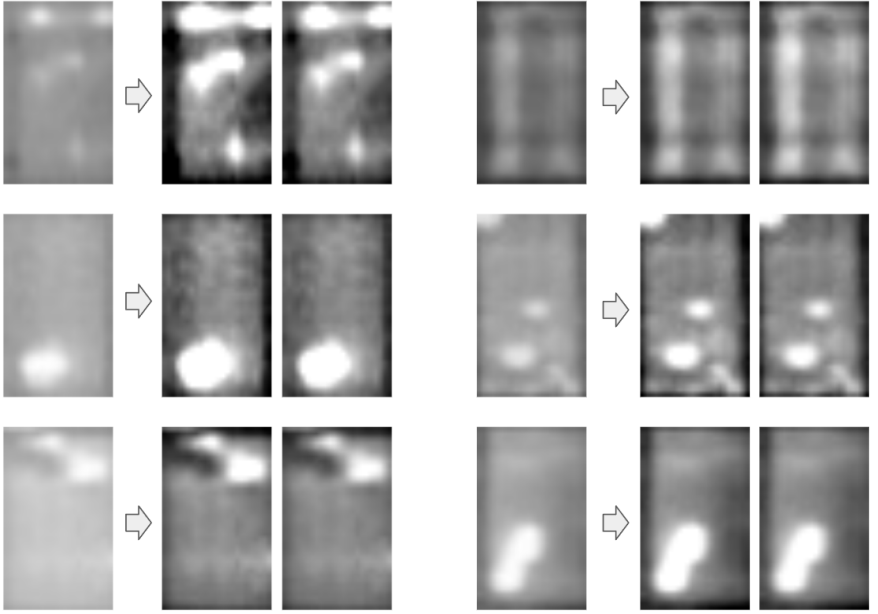


Figure 6: Example of BIE-based contrast enhancement on thermal images of defective PV modules.

linear brightness gradients and piecewise structures, making it ideal for capturing fault-relevant patterns in low-resolution thermal data. The section reviews related work on harmonic convolutions and lightweight CNN models.

Section 5.3 presents the proposed method in depth. It introduces Slant Convolution (SC) as a replacement for traditional learnable filters. These SC layers use fixed SLT basis functions, enhanced by trainable weights (α, γ) that modulate the frequency response using a logarithmic transformation. The section describes how the SC layer improves interpretability and efficiency by leveraging structured directional features. The architecture of SlantNet is then detailed, comprising two SC blocks followed by max-pooling, fully connected layers, and dropout regularization. Input images are processed at a resolution of 40×40 , making the network suitable for mobile and embedded devices. Furthermore, the augmentation pipeline includes geometric flips, contrast enhancement based on the BIE metric, and optimal decolorization guided by the TIA metric, targeting class imbalance and thermal feature preservation. This combination yields a highly optimized training set with improved visual discriminability of rare fault types.

Figure 7 compares standard and Slant Convolution pipelines. While standard convolution learns spatial filters directly from data, Slant Convolution first projects the input onto a fixed harmonic basis and then modulates it using trainable logarithmic parameters. This structured process enhances the extraction of directional and frequency-dependent features. Figure 8 illustrates the overall structure of the network, comprising two convolutional blocks, max-pooling layers, and a fully connected classifier.

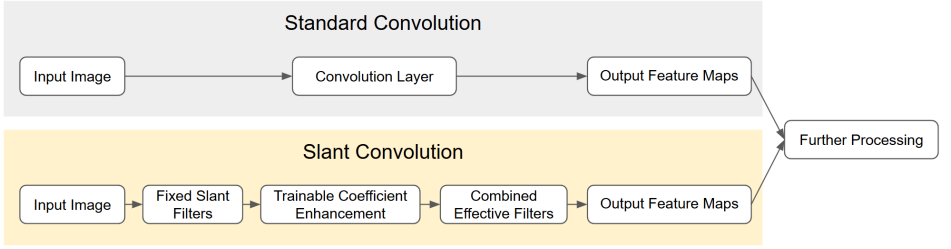


Figure 7: Comparison of Standard Convolution and Slant Convolution. Standard convolution learns arbitrary filters directly from data, while Slant Convolution first decomposes the input using a fixed harmonic basis and then applies trainable logarithmic enhancement to generate effective filters that better capture directional intensity variations.

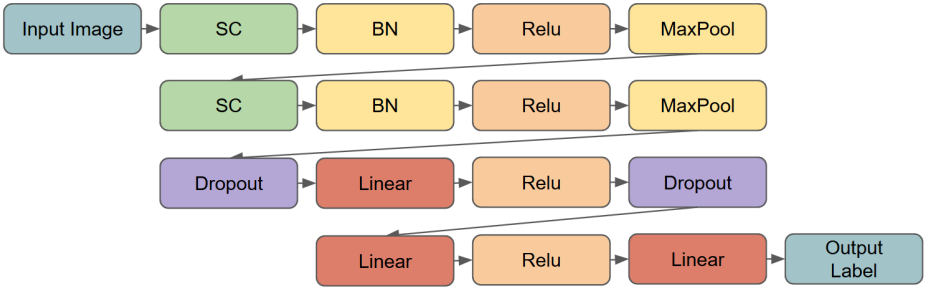


Figure 8: Overall architecture of the proposed SlantNet model incorporating the Slant Convolution (SC) layers.

Section 5.4 reports experimental results and comparisons. The section benchmarks SlantNet against AlexNet, ResNet50, MobileNetV3, EfficientNet, ShuffleNetV2, and Swin Transformer on binary and 12-class PV fault classification tasks. Evaluation metrics include accuracy (Acc), precision (Pr), recall (Rec), and specificity (Sp). SlantNet achieves the highest binary classification accuracy (95.1%) and competitive multiclass accuracy (82.75%), outperforming all evaluated models in classification performance. Full metric results are presented in Tables 3 and 4. In terms of efficiency, SlantNet demonstrates exceptional performance across all metrics. As shown in Table 5, it achieves the lowest FLOPs (3.55 MMac), a compact model size (12.82 MB), and a competitive parameter count (3.36 million), while significantly outperforming all other models in terms of throughput, approximately 55,000 images per second. This makes SlantNet highly suitable for real-time deployment on resource-constrained platforms such as UAVs, embedded systems, and edge devices.

Section 5.5 concludes the chapter by summarizing the contributions of SlantNet in advancing thermal image classification. It reiterates the benefits of integrating spectral transforms with deep learning and highlights the success of metric-based augmentation. Future directions include deployment in drone or IoT systems using TinyML, adaptation to other domains such as wind turbines and medical imaging, and exploration of other fast orthogonal transforms for further efficiency gains.

Table 3: Classification Performance on the Validation and Test Sets for binary classification

Model	Test				Validation			
	Acc	Pr	Rec	Sp	Acc	Pr	Rec	Sp
AlexNet	92.45	93.21	91.63	93.27	92.80	94.31	91.11	94.49
ResNet50	92.65	92.98	92.33	92.97	92.05	92.18	91.91	92.19
SqueezeNet	89.60	92.25	86.55	92.67	88.75	92.08	84.82	92.69
ShuffleNetV2	92.95	93.02	92.93	92.97	92.20	93.07	91.21	93.19
MobileNetV3	93.30	93.07	93.63	92.97	92.95	93.26	92.61	93.29
EfficientNet	93.50	94.87	92.03	94.98	94.05	95.37	92.61	95.50
ViT	88.05	89.80	85.96	90.16	88.40	90.69	85.61	91.19
Swin	91.35	92.27	90.34	92.37	91.75	93.36	89.91	93.59
Proposed	95.10	95.48	94.72	95.48	94.35	95.40	93.21	95.50

Table 4: Classification Performance on the Validation and Test Sets for 12-class classification

Model	Test				Validation			
	Acc	Pr	Rec	Sp	Acc	Pr	Rec	Sp
AlexNet	77.50	61.41	58.50	97.61	77.95	65.40	60.16	97.59
ResNet50	78.75	66.45	62.56	97.68	78.35	67.32	60.94	97.62
SqueezeNet	76.70	62.46	57.41	97.46	77.85	67.20	59.16	97.49
ShuffleNetV2	79.30	66.43	62.59	97.78	80.65	72.32	64.00	97.82
MobileNetV3	82.10	68.11	67.92	98.11	81.60	71.32	64.93	98.05
EfficientNet	82.20	69.37	71.05	98.19	82.55	72.35	69.51	98.18
ViT	74.70	60.60	54.76	97.17	75.65	65.26	57.57	97.16
Swin	80.45	65.93	63.19	97.98	81.55	71.61	66.77	98.02
Proposed	82.75	69.52	66.83	98.15	84.30	74.06	66.67	98.28

Table 5: Comparison of model efficiency metrics, including number of parameters (P), floating-point operations (FLOPs), model size in megabytes (M), and throughput (T).

Model	P (M)	FLOPs (MMac)	M (MB)	T (img/s)
AlexNet	57.01	714.97	217.48	18976
ResNet50	23.51	4130.00	89.68	1591
SqueezeNet	0.7	298.14	2.76	9069
ShuffleNet	1.26	151.36	4.81	7232
MobileNetV3	2.54	60.91	9.69	7759
EfficientNet	4.01	408.92	15.30	3281
ViT	85.80	17610	327.30	535
Swin Transformer	27.58	3120	105.21	828
SlantNet	3.36	3.55	12.82	55431

Chapter 6 concludes the dissertation by synthesizing the key innovations developed across four interrelated studies into a unified thermal image analysis framework tailored for PV fault detection. It reflects on how the proposed no-reference quality metrics, entropy-based enhancement, quality-guided augmentation, and the SlantNet architecture collectively address the challenges of low-resolution, noisy, and imbalanced thermal datasets. These contributions advance both the theoretical foundations and practical implementations of thermal imaging in renewable energy diagnostics. The chapter also outlines the broader impact of this work, demonstrating its scalability, generalizability, and applicability to edge deployment and multi-modal infrastructure monitoring, while highlighting promising future directions in autonomous inspection, cross-modal learning, and scalable AI systems for thermal

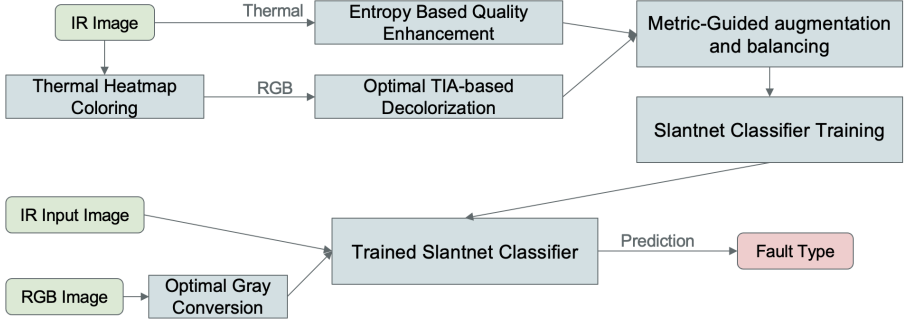


Figure 9: Overview of the proposed pipeline integrating quality assessment, contrast enhancement, data augmentation, and lightweight classification.

diagnostics.

Figure 9 illustrates the end-to-end workflow proposed in this dissertation. The process begins with an infrared image, which may be processed in two branches: one for quality enhancement and one for visualization. In the first branch, the raw thermal image undergoes entropy-based enhancement guided by the Block-wise Image Entropy (BIE) metric to improve contrast and suppress noise. In the second branch, the thermal image is colored using heatmaps and then converted to grayscale using the optimal TIA-based decolorization method. Both enhanced and decolorized images feed into a metric-guided augmentation module that selectively generates high-quality training samples, addressing class imbalance and improving generalization. These augmented datasets are then used to train the SlantNet classifier, a lightweight neural network designed for low-resolution thermal input. During inference, the trained SlantNet model receives either an IR or grayscale image and predicts the corresponding fault type. This pipeline enables robust, interpretable, and real-time fault classification, suitable for deployment on embedded systems and UAV platforms.

Main Results of the Research

This dissertation introduces a unified framework for thermal image analysis and classification, with a particular focus on solar photovoltaic (PV) fault classification. The research is grounded in five key contributions:

- Proposed two no-reference image quality metrics, **TIA** and **WTIA**, which assess perceptual fidelity in color-to-grayscale conversion. These metrics are robust, monotonic, and threshold-independent, outperforming classical methods like CCPR and E-score [1].
- Developed a novel entropy-based quality metric (**BIE**) tailored for thermal images. This metric combines block-wise entropy, standard deviation, and average deviation percentage to guide contrast enhancement and quantify uncertainty in noisy, low-resolution thermal imagery [2].
- Introduced an optimization framework for image processing based on nature-inspired meta-heuristic algorithms, including the Genetic Algorithm and the Bat Algorithm. These were demonstrated in the context of optimal decolorization and thermal image enhancement, using the proposed quality metrics as objective functions.

- Proposed a quality-aware augmentation pipeline that selects contrast-enhanced samples based on BIE and TIA/WTIA scores. This technique generates diagnostically meaningful synthetic data and significantly boosts classification performance under class imbalance [3].
- Designed **SlantNet**, a lightweight convolutional neural network architecture that incorporates novel Slant Convolution (SC) layers. These layers enable efficient directional feature extraction, achieving state-of-the-art accuracy and throughput for PV fault classification at reduced computational cost [4].

These contributions constitute a coherent and scalable pipeline for robust, interpretable, and computationally efficient thermal image analysis. The publicly released augmented dataset enhances reproducibility and supports broader research efforts. The proposed methods are applicable not only to PV monitoring but also to a wide range of real-time fault detection tasks, including wind turbine inspection, transformer diagnostics, and industrial visual monitoring. Future directions include integration into drone-based multimodal inspection systems, exploration of label-efficient training through self-supervised and federated learning approaches, and the development of an open-source toolkit to facilitate practical adoption.

Acknowledgment

This research was conducted within the ADVANCE Research Grant, provided by the Foundation for Armenian Science and Technology (FAST) and funded by Sarkis and Nune Sepetjians. The work was also supported by the Higher Education and Science Committee of the Republic of Armenia, under project No. 25FAST-1B001.

List of Published Works

- [1] 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, (Q1). doi: 10.1109/TCE.2023.3325744.
- [2] H. Ayunts, A. Grigoryan, and S. Agaian, “Novel entropy for enhanced thermal imaging and uncertainty quantification”, *Entropy*, vol. 26, no. 5, 2024, (Q2). doi: 10.3390/e26050374. [Online]. Available: <https://www.mdpi.com/1099-4300/26/5/374>.
- [3] H. Y. Ayunts, “Enhancing thermal image classification with novel quality metric-based augmentation techniques”, *Mathematical Problems of Computer Science*, vol. 62, pp. 112–125, 2024.
- [4] H. Ayunts, S. Agaian, and A. Grigoryan, “Slantnet: A lightweight neural network for thermal fault classification in solar pv systems”, *Electronics*, vol. 14, no. 7, 2025, (Q2). doi: 10.3390/electronics14071388. [Online]. Available: <https://www.mdpi.com/2079-9292/14/7/1388>.

ՊԱՏԿԵՐԻ ՄՇԱԿՄԱՆ ՄԵԹՈԴՆԵՐԻ ՕՊՏԻՄԻԶԱՑԻԱ և ԿԻՐԱՌՈՒԹՅՈՒՆՆԵՐ

Ամփոփում

Արևային ֆոտովոլտային համակարգերի զանգվածային ներդրման հետ մեկտեղ աճել է ավտոմատ, ցածր ծախսով մոնիտորինգի մեթոդների պահանջը՝ ուղղված միկրոճաքերի, միացումների կոռոզիայի և ջերմային սխալների հայտնաբերմանը, որոնք նվազեցնում են էներգիայի արդյունաբերումը: Ինֆրակարմիր պատկերների մշակումն արդյունավետ միջոց է այդ թերությունների հայտնաբերման համար, սակայն ցածր կոնտրաստը, փոքր չափսերը և աղմկայնությունը խանգարում են դասական համակարգչային տեսողության ալգորիթմներին:

Մասնավորապես, տեսանելի սպեկտրի մեթոդների վրա հիմնված գործիքները հաճախ չեն տալիս բավարար արդյունքներ ջերմային տվյալների դեպքում: Բարդություններ են առաջանում նաև տվյալների սահմանափակության, դասերի անհավասարակշռության, և ցածր հզորությամբ սարքերում կիրառելիության տեսանկյունից: Այս մարտահրավերները պահանջում են ջերմային տվյալներին հարմարեցված արդյունավետ, ճկուն և բաշխվող լուծումներ, որոնք կարող են գործել իրական ժամանակում:

Աշխատանքի նպատակը և դիտարկված խնդիրները

Աշխատանքի նպատակը թերությունների ճշգրիտ հայտնաբերման համար արդյունավետ խողովակաշար մշակելն է, որն օպտիմալացված է արևային վահանակների ստուգման և ցածր հզորությամբ սարքերում տեղակայման համար:

Խնդիրներն են՝ նախագծել որակի չափիչներ, մշակել ադապտացիաների նկատմամբ կայուն, էնտրոպիայի վրա հիմնված ջերմային պատկերների բարելավում, ստեղծել տվյալների հավաքածուի աուգմենտացիայի մեթոդ, մշակել նեյրոնային ցանցը՝ արդյունավետ դասակարգման համար, և ամբողջ շղթան ինտեգրել միավորված համակարգում՝ փորձարկելով բազմաբնույթ ջերմային տվյալների վրա:

Ստացված արդյունքների կիրառական նշանակությունը

Աշխատանքում ներկայացված գործիքները ապահովում են պատկերների արդյունավետ գնահատում և դասականում իրական ժամանակում, հարմարեցված դրոնների, ներկառուցված սարքերի և եզրային համակարգերի վրա կիրառման համար: Շղթան՝ ներառյալ առանց հղման որակի չափիչները, հարմարվող բարելավումը և SlantNet մոդելը, նվազեցնում է դասակարգման ժամանակը՝ բարձրացնելով արևային վահանակների արդյունավետությունը: Մեթոդները նաև կիրառելի են տեսահսկման, արդյունաբերական և անվտանգության մոնիտորինգի ոլորտներում: Բացի այդ, տվյալների ընդլայնված հավաքածուն հանրայնացվել է GitHub հարթակում՝ հեշտացնելու գիտական վերարտադրելիությունը և խթանելու հետազա հետազոտությունները:

Աշխատանքի ծավալը և կառուցվածքը

Աշխատանքը ընդգրկում է 126 էջ՝ ներառյալ ներածություն, չորս գլուխ և եզրակացություն: Այն պարունակում է 134 գրականության աղբյուր:

Աշխատանքի հիմնական արդյունքները

- Առաջարկվել են ոչ-հղման **TIA** և **WTIA** որակի չափիչներ՝ պատկերների գունաթափման ընթացքում կառուցվածքային և գունային տեղեկատվության պահպանումը գնահատելու համար: Չափիչները մոնոտոն են, կայուն և անկախ են գունային տարբերության պարամետրից՝ գերազանցելով գոյություն ունեցող մեթոդները [1]:
- Մշակվել է ջերմային պատկերների համար հարմարեցված նոր էնտրոպիայով որակի չափիչ՝ **BIE**: Այն համադրում է բլոկային էնտրոպիան, ստանդարտ շեղումը և միջին շեղման տոկոսը՝ թույլ տալով կոնտրաստի օպտիմալացում և անոռոշության քանակականացում աղմկոտ ու ցածր-բարձրության ջերմային պատկերներում [2]:
- Ներկայացվել է պատկերի մշակման օպտիմալացման համակարգ՝ հիմնված բնությունից ներշնչված մետաէվրիստիկ ալգորիթմների վրա (օր.՝ գենետիկ և չդիֆերենցիալ ալգորիթմներ), որտեղ որպես նպատակային ֆունկցիա օգտագործվում են առաջարկված որակի չափիչները՝ գունաթափման և ջերմային բարելավման խնդիրների համար:
- Առաջարկվել է որակային չափիչներով ղեկավարվող ընդլայնման մեթոդ, որը ընտրում է կոնտրաստով հարստացված նմուշները՝ ապահովելով արժեքավոր տվյալների ստեղծում և զգալիորեն բարձր դասակարգման ճշգրտություն դասերի անհավասարակշռության պայմաններում [3]:
- Մշակվել է **SlantNet**՝ թեթև նեյրոնային ցանցային ճարտարապետություն, որը ներառում է նորարարական թեք կոնվոլյուցիաներ (Slant Convolution)՝ ուղղորդված առանձնահատկությունների արդյունավետ վերլուծության համար: Ճարտարապետությունը հասնում է առաջատար ճշգրտության և արագության՝ նվազագույն հաշվարկային բարդության պայմաններում [4]:

Ներդրումները կազմում են ամբողջական համակարգ, որն արդյունավետորեն և վստահելի կերպով իրականացնում է ջերմային պատկերների վերլուծություն՝ ժամանակի պահանջներին համապատասխան:

ОПТИМИЗАЦИЯ МЕТОДОВ ОБРАБОТКИ ИЗОБРАЖЕНИЙ И ИХ ПРИМЕНЕНИЯ

Заключение

Масштабное внедрение солнечных фотоэлектрических (ФЭ) систем усилило потребность в автоматических и недорогих методах инспекции для выявления дефектов, таких как микротрещины, коррозия соединений и термические аномалии, снижающие выработку энергии. Инфракрасная термография позволяет выявлять такие неисправности, однако её применение осложняется низким контрастом, ограниченным разрешением и высокой чувствительностью к шуму.

Стандартные методы обработки изображений, разработанные для видимого спектра, плохо справляются с анализом тепловизионных данных. Кроме того, наблюдается нехватка размеченных тепловизионных данных, сильная дисбалансировка классов и сложность применения существующих моделей в условиях ограниченных вычислительных ресурсов. Эти ограничения требуют создания специализированных подходов к обработке инфракрасных изображений, адаптированных для беспилотников и встроенных систем.

Основная цель работы и задачи

Целью является создание легковесного и завершённого конвейера для точного обнаружения тепловых дефектов, оптимизированного для обследования солнечных панелей и пригодного для размещения на маломощных устройствах.

Ключевые задачи включают: разработку безэталонных метрик качества; создание метода контрастного улучшения на основе энтропии; генерацию синтетических данных для балансировки редких классов; проектирование эффективной модели; интеграцию всех компонентов в единый открытый программный комплекс, протестированный на различных тепловизионных наборах данных.

Практическая значимость полученных результатов

Предложенные в работе инструменты обеспечивают эффективную и доступную тепловизионную инспекцию в реальном времени, адаптированную для использования на дронах, встроенных устройствах и пограничных системах. Интеграция безэталонных метрик качества, адаптивного улучшения и лёгкой модели SlantNet позволяет сократить время проверки и простой оборудования, повышая эффективность фотоэлектрических панелей. Разработанные методы также применимы в видеонаблюдении, промышленной диагностике и системах безопасности. Кроме того, расширенный набор синтетических данных размещён в открытом доступе на GitHub для обеспечения воспроизводимости и поддержки новых исследований.

Объём и структура работы

Диссертация содержит 126 страниц и включает введение, четыре главы и заключение. В работе приведены 134 библиографических источника.

Основные результаты работы

- Разработка двух безэталонных метрик качества изображений — **TIA** и **WTIA**, предназначенных для оценки сохранения хроматических и структурных признаков при преобразовании цветных изображений в оттенки серого. Метрики демонстрируют устойчивость к порогам, монотонность и превосходят классические подходы [1].
- Предложена новая энтропийная метрика для тепловизионных изображений — **BIE**, учитывающая блочную энтропию, стандартное отклонение и показатель средней девиации. Она предназначена для оценки качества, усиления контраста и количественной оценки неопределённости в условиях шума и низкого разрешения [2].
- Разработана схема оптимизации методов обработки изображений на основе метаэвристических алгоритмов, таких как генетический алгоритм (GA) и алгоритм летучих мышей (BA). Предложенные метрики используются в качестве целевых функций для задач деколоризации и улучшения качества тепловизионных изображений.
- Представлен метод генерации синтетических данных, основанный на оценке качества изображений. Он позволяет отбирать контрастно-усиленные и диагностически значимые образцы, что значительно улучшает точность классификации при наличии дисбаланса классов. Также опубликован расширенный термальный датасет [3].
- Разработана нейросетевая архитектура **SlantNet** — лёгкая и эффективная модель на основе наклонных свёрток (Slant Convolution), обеспечивающая направленное извлечение признаков, высокую точность и производительность при минимальных вычислительных затратах [4].

Эти результаты формируют масштабируемую и интерпретируемую систему для анализа тепловизионных изображений в реальном времени.

