Electronics and Information Technologies, 2026, 34, 5-30
http://publications.lnu.edu.ua/collections/index.php/electronics/index

UDC: 004.932

IMAGE FUSION TECHNIQUES USING POLARIZATION INFORMATION AND INFRARED INTENSITY

Vitalii Artym*, Oleh Krupych
Department of Optoelectronics and Information Technologies
Faculty of Electronics and Computer Technologies
Ivan Franko National University of Lviv,
107 Tarnavsky Str., Lviv, 79017, Ukraine

Artym, V.I., Krupych, O.M. Image Fusion Techniques Using Polarization Information and Infrared Intensity. *Electronics and Information Technologies*, 34, 5–30. https://doi.org/10.30970/eli.34.1

ABSTRACT

Background. Fusion of polarization images with conventional thermal infrared images combines complementary physical information to improve scene interpretation. Infrared intensity imaging captures thermal emission and emissivity patterns. In contrast, polarization parameters, including the degree and angle of linear polarization, encode surface microstructure, material composition, and geometric orientation that are inaccessible to intensity-only sensing. Their fusion enhances spatial detail, target discrimination, and robustness under degraded visibility conditions, reflecting a broader shift toward multimodal perception.

Materials and Methods. This study provides an analytical review of polarization–infrared fusion techniques. The development of methods is examined from classical linear transforms, multiresolution decomposition, and energy-based decision rules to adaptive hybrid frameworks. Particular attention is given to deep learning architectures and physics-informed models that incorporate polarization formation principles and radiometric constraints. The review situates fusion progress within advances in sensor technologies and computational imaging.

Results and Discussion. The analysis indicates that the incorporation of polarization information substantially enhances fused image quality compared with conventional intensity-based techniques. Classical algorithms provide stable and interpretable performance but exhibit limited adaptability to complex environments. Learning-based and hybrid strategies demonstrate superior contrast preservation, structural detail enhancement, and target detectability, especially in thermally heterogeneous or visually degraded scenarios. Persistent challenges include limited annotated datasets, polarization-induced noise sensitivity, and reduced physical interpretability of deep neural representations.

Conclusion. Polarization–infrared fusion represents a transition from mono-modality to multi-modal observation with the system-oriented perception integrating physical modeling with data-driven learning. Future research should emphasize physically reasoned neural design, standardized evaluation protocols, and improved interpretability to support reliable real-world deployment.

Keywords: image fusion, polarization imaging, infrared imaging, multimodal data, deep learning.

INTRODUCTION

Image fusion – understood as the synthesis of heterogeneous sensory information into a unified perceptual or computational representation – has become a foundational topic in optical engineering and computer-vision research. Among the most powerful combinations is the fusion of infrared (IR) intensity and polarization imaging, two modalities that offer inherently coupled complementary physical descriptions of a scene. Infrared imaging provides thermal or emissive contrast independent of visible illumination, making it highly effective under low-light or adverse weather conditions [1]. Polarization imaging, by contrast, encodes the orientation and shape of the electromagnetic field vector oscillations, revealing surface texture, material properties, and scattering phenomena that are inaccessible to conventional intensity images [2].

Historically, fusion methods have evolved from simple arithmetic combinations toward sophisticated adaptive and learning-based schemes. In the 1990s and early 2000s, pyramid and wavelet transforms dominated, reflecting the influence of multiresolution analysis on computer vision. With the advent of polarimetric sensors and computational optics, the possibility of integrating polarization cues into traditional intensity images emerged. The initial goal was enhancement – improving contrast and edge visibility under adverse conditions such as fog, haze, or camouflage [3].

Subsequent decades witnessed the expansion of this field toward autonomous sensing, surveillance, and biomedical diagnostics. Yet the fusion of polarization and infrared modalities also signifies a conceptual evolution: it marks a transition from passive observation to active reconstruction of the scene's physical reality. Each modality expresses a distinct physical ontology – thermal emission versus electromagnetic wave polarization– whose synthesis mirrors the epistemological shift from single-perspective to multimodal cognition.

The following sections of this paper trace this evolution in both technical and conceptual terms. Section 1 outlines the formation of polarimetric and infrared imaging paradigms; Section 2 reviews classical, hybrid, and deep-learning fusion techniques; Section 3 examines applications and benchmarking practices; Section 4 discusses epistemological and methodological dimensions; and Section 5 concludes with perspectives on future integration.

1. POLARIMETRIC AND INFRARED IMAGING: PHYSICAL BACKGROUND

1.1. Historical Evolution of Polarization and Infrared Imaging

The convergence of polarization and infrared imaging is not merely a technological synthesis but the outcome of a long historical process in which distinct scientific traditions – wave optics, thermodynamics, radiometry, sensor engineering, and computational imaging – gradually intersected. A historiographical perspective reveals that polarization–infrared fusion emerged through successive epistemic phases: classical theoretical formalization, instrumental stabilization, digital transformation, and algorithmic integration. Each phase redefined both what could be measured and how measurements were interpreted.

Classical Optics and the Conceptual Birth of Polarization

The origins of polarization science lie in nineteenth-century wave optics. The decisive theoretical advance was the formulation of reflection and refraction laws for polarized light by Augustin-Jean Fresnel [4]. Fresnel's equations demonstrated that electromagnetic waves interacting with material boundaries exhibit polarization-dependent reflection coefficients. This discovery transformed polarization from a curious optical phenomenon into a quantitatively predictable property of wave – matter interaction.

The formal mathematical framework was followed with the introduction of the Stokes parameters by George Gabriel Stokes in 1852 [5]. The Stokes vector representation provided a complete description of partially polarized light through four measurable quantities (S_0, S_1, S_2, S_3). Importantly, this formulation unified intensity and polarization into a single algebraic structure, enabling subsequent generations of researchers to treat polarization as a vectorial extension of radiometric measurement.

Throughout the late nineteenth and early twentieth centuries, polarization research developed largely within laboratory physics and astronomy [6]. Applications included mineralogy, crystallography, and atmospheric scattering studies. However, imaging was not yet central; measurements were typically pointwise and instrument-based, relying on rotating analyzers and photometric detection. The absence of array detectors constrained polarization analysis to experimental optics rather than scene-based imaging.

Historiographically, this period may be characterized as conceptual consolidation: the physical and mathematical principles of polarization were firmly established, but technological limitations prevented their integration into practical imaging systems.

Infrared Radiation: From Discovery to Radiometric Formalization

Infrared science followed a parallel yet distinct trajectory. The discovery of infrared radiation by William Herschel in 1800 [7] extended the electromagnetic spectrum beyond visible light. However, it was not until the development of thermodynamic theory that infrared radiation acquired quantitative meaning.

The formulation of blackbody radiation laws by Max Planck at the turn of the twentieth century [8] provided the theoretical basis for thermal radiometry. Planck's law established a direct relationship between temperature and spectral radiance, enabling the use of infrared measurements as thermometric indicators.

The early twentieth century saw incremental advances in infrared detectors, including bolometers and thermocouples [9]. These devices were sensitive but slow, limiting spatial resolution and dynamic imaging capability. Military applications during World War II accelerated infrared technology development, leading to scanning thermal imagers for target detection and night vision.

By the 1970s and 1980s, focal plane arrays (FPAs) based on materials such as HgCdTe and InSb enabled two-dimensional infrared imaging [10]. Long-wave infrared (LWIR) systems (8–14 μm) became especially prominent due to their compatibility with ambient-temperature thermal emission. The introduction of uncooled microbolometers in the 1990s democratized infrared imaging by reducing system size and cost [11].

From a historiographical standpoint, infrared imaging transitioned from laboratory thermometry to operational sensing through the stabilization of detector technologies. Unlike polarization science, which matured theoretically early but technologically late, infrared imaging experienced rapid technological advance driven by defense and industrial demand.

Digital Imaging and the Sensor Revolution

The late twentieth century introduced a transformative element common to both domains: digital imaging. The development of CCD and CMOS sensor arrays enabled spatially resolved acquisition of optical intensity [12]. For infrared systems, hybrid FPAs provided pixel-level radiometric sampling; for visible-spectrum systems, CMOS sensors allowed integration of micro-optical elements.

Polarization imaging underwent a comparable revolution with the invention of micro-polarizer arrays deposited directly onto sensor pixels [13]. These arrays consisted of sub-pixel polarization filters oriented at discrete angles (e.g., 0°, 45°, 90°, 135°), enabling snapshot acquisition of multiple polarization states without mechanical rotation.

This innovation resolved a long-standing instrumental limitation: polarization could now be measured spatially and temporally in dynamic scenes. Consequently, polarimetry transitioned from complicated laboratory photometry to real-time imaging.

The historical significance of this stage lies in the instrumental convergence of modalities. For the first time, polarization and infrared imaging shared comparable spatial resolution and digital representation, making computational integration feasible.

Early Multisensor Fusion: Transform-Based Integration

Before the specific integration of polarization and infrared modalities, multisensor fusion emerged as a broader research area in the 1980s and 1990s. Theoretical tools from multiresolution analysis played a decisive role.

The Laplacian pyramid method introduced by Peter J. Burt and Edward H. Adelson [14] provided a systematic framework for decomposing images into frequency bands. Shortly thereafter, the multiresolution wavelet formalism developed by Stéphane Mallat [15] established a mathematically rigorous basis for hierarchical image representation.

Fusion algorithms employed coefficient-selection rules such as maximum absolute value, local energy, or weighted averaging [16]. These approaches treated input modalities as statistically comparable signals, abstracting away their physical origins.

When polarization information was first incorporated into fusion pipelines, it was often reduced to scalar derivatives such as maps of *Degree of Linear Polarization* (DoLP), treated analogously to intensity images. The deeper physical implications of polarization – vectorial boundary interactions and anisotropic emission – were not yet fully integrated into algorithm design.

This period may be interpreted as algorithmic abstraction: fusion was understood primarily as signal combination in a transform domain, rather than as physically constrained integration.

Emergence of Polarimetric Thermal Imaging

The true convergence of polarization and infrared modalities required technological advances enabling simultaneous measurement of radiance and polarization state in thermal wavelengths. Polarimetric thermal imagers capable of estimating Stokes vector elements S_0, S_1, and S_2 began to appear in the early twenty-first century [17].

Research demonstrated that thermal emission from surfaces can exhibit partial linear polarization due to angular emissivity dependence and surface anisotropy. This insight revealed that polarization contrast could enhance target detection even when temperature contrast was minimal.

In camouflage scenarios, for example, thermally equilibrated objects might remain indistinguishable in intensity images yet display polarization signatures arising from reflective coatings or geometric discontinuities. Such findings stimulated the development of fusion algorithms specifically designed for polarization–infrared integration.

Unlike earlier multisensor fusion tasks, this integration required reconciliation of two distinct physical models:

- Radiometric emission is governed by Planck's law and emissivity functions.
- Vectorial polarization is governed by Fresnel boundary conditions.

Thus, fusion became a problem of heterogeneous physical synthesis, not merely signal combination.

Learning-Based Integration and the Data-Driven Turn

The 2010s introduced deep learning into image fusion research. Convolutional neural networks (CNNs) demonstrated the capacity to learn hierarchical feature representations directly from data [18]. Unsupervised fusion architectures eliminated the need for handcrafted coefficient rules, instead optimizing reconstruction losses and structural similarity metrics.

The publication of transformer-based architectures by Ashish Vaswani and colleagues [19] further advanced cross-modal feature integration through attention mechanisms. In polarization–infrared fusion, attention modules dynamically weight thermal and polarization features according to contextual saliency.

Recent developments incorporate physics-informed constraints into neural architectures [20]. For example, loss functions may enforce consistency with Stokes parameter relationships or penalize physically implausible polarization amplification. This marks a shift toward hybrid epistemology, in which data-driven inference is constrained by physical law.

Standardization, Benchmarking, and Reproducibility

A crucial development in the maturation of polarization–infrared fusion research has been the emergence of standardized datasets and evaluation metrics. Early studies were limited by proprietary data and inconsistent benchmarking protocols.

Contemporary research increasingly employs multi-criteria evaluation frameworks incorporating entropy, mutual information (MI), structural similarity (SSIM), and visual information fidelity (VIF) [21]. More recently, physically motivated metrics assess whether fused outputs preserve thermal contrast and polarization integrity [22].

This standardization reflects the institutionalization of the field. Fusion research has moved from exploratory experimentation to systematic comparative evaluation, enabling reproducibility and cross-study validation.

Historiographical Evolution of Polarization–Infrared Fusion

From a historiographical perspective, the evolution of polarization–infrared fusion can be interpreted as a sequence of paradigm transitions:

- *Conceptual Phase* – Establishment of wave and radiometric theory (nineteenth–early twentieth century).
- *Instrumental Phase* – Stabilization of detector technologies and array imaging (mid–late twentieth century).
- *Digital Phase* – Emergence of computational imaging and multiresolution analysis (1990s).
- *Integrative Phase* – Physical and algorithmic convergence of polarization and infrared modalities (2000s).
- *Data-Driven Phase* – Deep learning and physics-informed models (2010s–present).

Each phase expanded not only technological capability but also epistemic scope. Polarization imaging transitioned from laboratory optics to dynamic scene analysis; infrared imaging evolved from thermometric measurement to semantic perception; fusion unified these modalities into integrated perceptual systems.

The convergence of polarization and infrared imaging thus exemplifies the broader transformation of optical science into computational multimodal sensing. What began as independent theoretical traditions has become a unified research domain characterized by interdisciplinary integration of physics, engineering, and machine learning.

1.2. Physical and Mathematical Principles

The polarization state of light can be fully described by the *Stokes vector* $\mathbf{S} = [S_0, S_1, S_2, S_3]^{\mathrm{T}}$, where S_0 denotes total intensity, S_1 and S_2 represent intensity differences of linear polarization components at orthogonal orientations ($S_1 = I(0°) - I(90°)$ and $S_2 = I(45°) - I(135°)$), while $S_3 = I_{Right} - I_{Left}$ corresponds to the intensity difference of right- and left-handed circular polarizations. For most passive imaging systems operating in the visible or infrared range, the Stokes parameter S_3 is irrelevant or unattainable; thus, the focus lies on the *Degree of Linear Polarization* (DoLP) and the *Angle of Polarization* (AoP) [23]:

$$\begin{aligned} DoLP &= \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \\ AoP &= \frac{1}{2}\tan^{-1}\left(\frac{S_2}{S_1}\right). \end{aligned} \tag{1}$$

Infrared detectors measure spectral radiance

$$L(\lambda, T) \approx \varepsilon(\lambda) \cdot B(\lambda, T), \tag{2}$$

where $\varepsilon(\lambda)$ is emissivity and $B(\lambda, T)$ is the Planck function.

The polarized component arises from anisotropic emission or reflection and carries information about surface orientation and roughness.

The challenge of fusion lies in reconciling these two data spaces – radiometric (temperature/emissivity) and vectorial (polarization). While IR intensity follows scalar thermodynamic laws, polarization obeys vector optics. Successful fusion therefore, demands both physical modeling and statistical optimization.

2. EVOLUTION OF IMAGE FUSION TECHNIQUES

The historical development of image-fusion techniques can be read as a gradual transition from *analytical* to *synthetic* paradigms of vision: from deterministic transforms to adaptive and data-driven inference. Each period reflects a different understanding of how information from distinct modalities – infrared intensity and polarization – should be interpreted and unified.

2.1. Classical Transform-Based Methods

Early methods were rooted in linear systems theory and multi-resolution signal analysis. Infrared and polarization images were decomposed into spatial–frequency representations using transforms such as the *Laplacian pyramid, discrete wavelet transform* (DWT), or *nonsubsampled contourlet transform* (NSCT). Fusion rules were then applied to corresponding coefficients.

Yue and Li [24] employed an oriented Laplacian pyramid to integrate infrared intensity with polarization information. The fusion rule selected coefficients of higher gradient energy, producing improved edge contrast and object detectability. Zhang et al. [25] later extended this approach using an orthogonality-difference metric that captured the divergence between polarization components, achieving sharper boundaries in long-wave infrared (LWIR) imagery.

Experimental studies on LWIR polarimetric imaging further illustrate the complementary nature of infrared intensity and polarization features. For example, the work mentioned above of Zhang et al. [25] demonstrates how polarization measurements at different analyzer orientations can be combined with infrared intensity to derive physically meaningful polarization descriptors such as the degree of linear polarization (DoLP) and the angle of polarization (AoP). These features highlight surface reflectance and material properties that do not show up in intensity images alone.

Figure 1 is a representative example of long-wave infrared polarization imaging and fusion [25]. Fragments (**Fig. 1 a–d**) show the raw measurements acquired by a polarization camera: the infrared intensity image and three polarization-filtered images captured at analyzer angles of 0°, 60°, and 120°. From these measurements, the polarization parameters DoLP and AoP are computed, as shown in (e) and (f). The final fused image (g) integrates thermal intensity and polarization information, revealing enhanced structural detail and improved discrimination of objects against the background. This example demonstrates how polarization descriptors can provide additional contrast cues that complement thermal information, motivating the development of classical fusion algorithms based on transform-domain representations.

Such approaches are interpretable and computationally efficient, yet they depend heavily on hand-crafted heuristics (e.g., choosing maximum or weighted average coefficients). They also assume stationary statistics across the image, which rarely holds for natural or cluttered scenes. A comparative summary of the main classical transform-

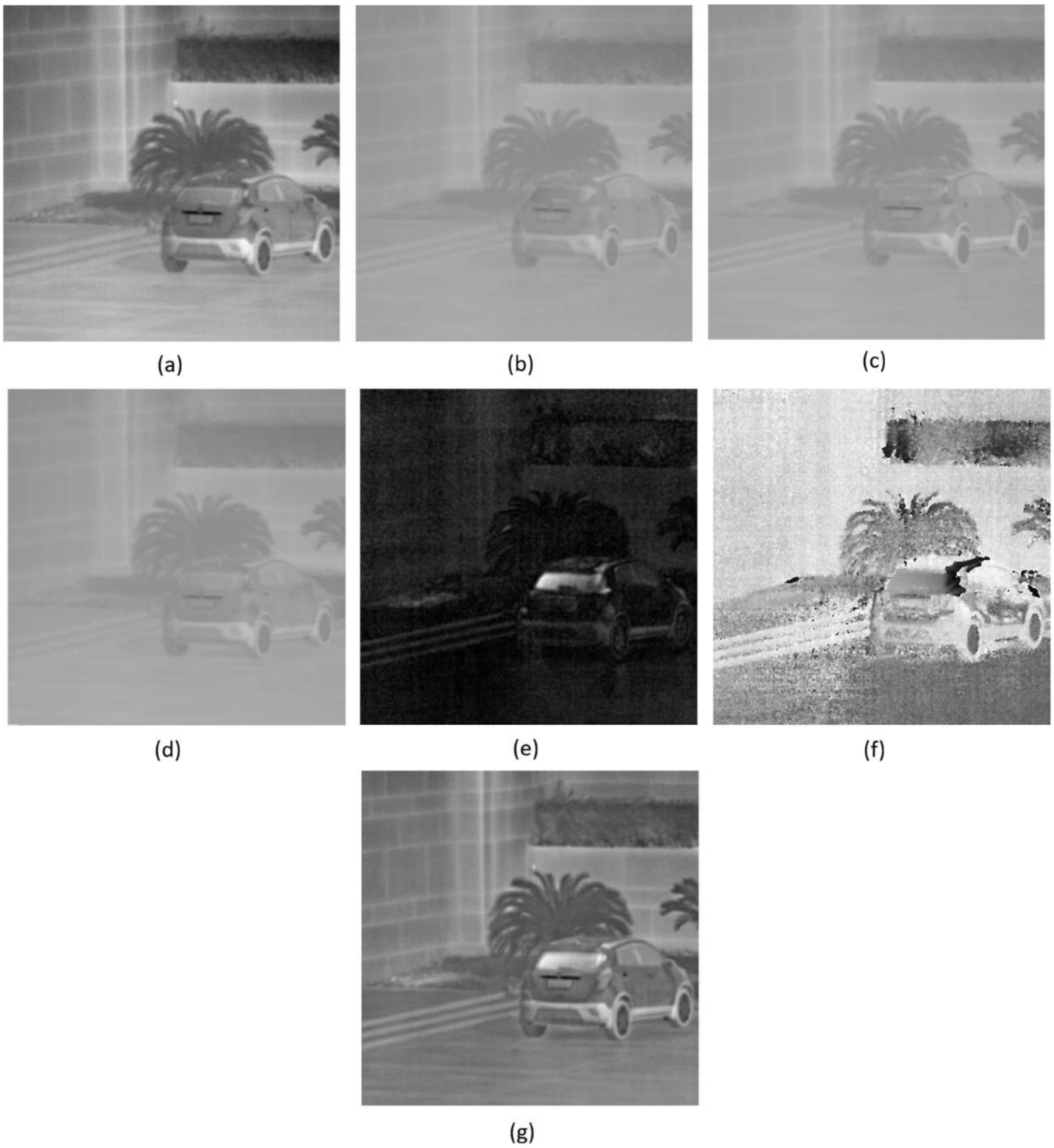

Fig. 1. Example of long-wave infrared polarization image fusion [25]: (a) infrared intensity image, (b) 0° polarization image, (c) 60° polarization image, (d) 120° polarization image, (e) degree of linear polarization (DoLP), (f) angle of polarization (AoP), and (g) fused image combining infrared intensity and polarization information

based fusion methods, including their representative transformations, fusion rules, strengths, and limitations, are provided in **Table 1**.

Laplacian Pyramid

Principle.

The Laplacian pyramid is one of the earliest multiresolution image representations, originally introduced by Burt and Adelson [14]. The method constructs a sequence of progressively smoothed (Gaussian-blurred) images and then computes the difference between consecutive levels, producing a set of band-pass (Laplacian) components. This representation is localized, efficient, and offers a multi-scale decomposition of the image.

Table 1. **Comparison of Classical Fusion Approaches**

Method	Transform Domain	Fusion Rule	Strengths	Limitations
Laplacian Pyramid [1]	Multiscale Spatial	Max Gradient	Simple, fast	Sensitive to noise
Wavelet Fusion [3]	DWT	Weighted Energy	Good localization	Artifacts at edges
NSCT [4]	Contourlet	Regional Entropy	Multi-directional	Computationally heavy

Algorithm (simplified):

1. Construct the Gaussian pyramid:
   $G_0 = I$,
   $G_{k+1} = \mathcal{D}(G_k)$ – smoothing + downsampling.
2. Compute the Laplacian levels:
   $L_k = G_k - \mathcal{U}(G_{k+1})$,
   where $\mathcal{U}$ is upsampling with interpolation.
3. Store all Laplacian levels $L_0, \dots, L_{N-1}$ and the coarsest Gaussian level G_N. Reconstruction is performed by iteratively adding back the upsampled higher-level images.

Fusion rules:

- Maximum absolute value per coefficient.
- Weighted average based on local variance or gradient energy.
- Region-based selection.

Advantages:

- Simple, interpretable, low computational cost.
- Preserves edges and local contrast effectively.

Limitations:

- Sensitive to noise (especially at high-frequency levels).
- No directional representation; cannot capture oriented structures well.

Discrete Wavelet Transform (DWT)

Principle.

The discrete wavelet transform provides an orthogonal or biorthogonal multiresolution decomposition, splitting the image into one low-frequency subband (LL) and three high-frequency directional subbands (LH, HL, HH). The theoretical foundation was introduced by Mallat through the concept of multiresolution approximation [15].

Algorithm:

1. Perform an L-level DWT on each input image to obtain LL, LH, HL, and HH subbands.
2. Fuse the low-frequency LL subbands using weighted averaging, entropy-based measures, or PCA-style rules.
3. Fuse the high-frequency subbands using:
   - Max-absolute-value selection,
   - Energy-based or gradient-based measures,
   - Local saliency criteria.
4. Apply inverse DWT to reconstruct the fused image.

Advantages:

- Strong spatial–frequency localization.
- Widely used in remote sensing, medical imaging, and multisensor fusion.

- Flexible choice of wavelet bases.

Limitations:

- Limited orientation diversity (only three high-frequency directional bands).
- Shift-variant – even small misalignments cause artifacts.
- Prone to aliasing when the sampling geometry changes.

Nonsubsampled Contourlet Transform (NSCT)

Principle.

The nonsubsampled contourlet transform (NSCT), proposed by da Cunha, Zhou, and Do [26], is a shift-invariant, overcomplete, directional multiresolution transform specifically designed to represent edges and contours more effectively than wavelets. Unlike the classical contourlet transform, NSCT avoids downsampling, making it fully translation-invariant.

NSCT consists of two key components:

1. Nonsubsampled Pyramid (NSP): Multiscale decomposition without subsampling.
2. Nonsubsampled Directional Filter Bank (NSDFB): Directional decomposition at each scale, also without subsampling.

Fusion rules:

- High-frequency directional subbands: saliency- or energy-based selection.
- Low-frequency subband: weighted or statistical combination.

Advantages:

- Excellent directional and anisotropic representation of curved structures.
- Shift-invariant (no downsampling), reducing misregistration artifacts.
- Highly effective for preserving edges and textures.

Limitations:

- High computational cost and memory usage due to redundancy.
- Quality depends strongly on filter design.
- Requires parameter tuning (number of scales/orientations).

2.2 Multi-Parameter and Hybrid Polarization Fusion

The next phase introduced multi-parameter models, integrating DoLP, AoP, and sometimes partial Stokes information with IR intensity [17]. The guiding idea was that different polarization parameters represent complementary structural cues.

An important direction within hybrid and physically interpretable fusion methods involves the use of complex-valued representations and matrix formalisms for combining visible, infrared, and polarization data. Khaustov and co-authors proposed several approaches based on complex functions, Jones-matrix formalism, and complex-vector fusion models, demonstrating improved target saliency preservation and adaptive fusion behavior in military sighting systems [27–33].

Unlike single-parameter fusion approaches that rely solely on DoLP, multi-parameter fusion exploits the complementary physical information encoded in different polarization descriptors. Specifically, DoLP highlights material-dependent reflection properties, AoP conveys geometric and orientation-related surface characteristics, while infrared intensity provides robust thermal contrast under low-visibility conditions. Hybrid fusion frameworks typically combine these parameters at different representation levels, including pixel-level weighted fusion, multiresolution transform domains (e.g., wavelet- or contourlet-based fusion), and feature-level strategies guided by saliency or edge measures. More recently, deep learning–based methods have been introduced to adaptively learn fusion rules across multiple polarization parameters and infrared intensity, resulting in improved target discrimination and structural preservation. Nevertheless, multi-parameter polarization fusion remains challenged by polarization noise, parameter normalization, and the physical interpretability of learned fusion mechanisms, motivating ongoing research into physics-informed and noise-aware fusion models.

2.3 Emergence of Learning-Based Paradigms

With the rapid development of machine learning – particularly convolutional neural networks (CNNs) and attention-based architectures – image fusion has entered an adaptive, data-driven era. Unlike classical fusion methods that rely on hand-crafted rules or fixed transform-domain strategies, learning-based approaches automatically learn nonlinear mappings between multimodal inputs and fused outputs through end-to-end optimization.

Unsupervised and Self-Learned CNN Fusion. Li et al. proposed an unsupervised CNN-based fusion framework, known as DenseFuse, in which an encoder–fusion–decoder architecture learns effective fusion representations without requiring ground-truth fused images [18]. The method employs structural similarity–based loss functions to preserve salient features from multiple modalities while maintaining visual consistency. This approach demonstrated that meaningful fusion rules can be learned directly from data, marking a key transition from handcrafted fusion strategies to representation-learning paradigms.

Attention-Driven and Transformer-Based Models. Recent studies have incorporated attention mechanisms to enhance the selective integration of multimodal features. Attention-guided networks dynamically emphasize salient regions – such as thermally prominent targets – while retaining fine structural details derived from polarization cues [34]. These mechanisms improve robustness in complex scenes by adaptively weighting feature contributions across spatial locations and modalities.

A representative example is the attention-guided filtering network proposed by Li et al. [34] for infrared–polarization image fusion. **Figure 2** presents a qualitative comparison of fusion results obtained using several representative algorithms. The attention-guided approach demonstrates improved preservation of structural details and enhanced contrast of thermal targets while suppressing noise in polarization measurements. This illustrates the advantage of attention mechanisms in selectively integrating complementary information from infrared intensity and polarization features.

Physics-Aware Deep Fusion. Beyond purely data-driven learning, recent efforts have explored the integration of physical priorities into deep fusion frameworks. Noise-aware and physics-consistent loss formulations, particularly for polarization–infrared fusion, have been shown to improve stability and interpretability while preserving the advantages of deep learning [34, 35]. Such approaches constrain network behavior in accordance with polarization imaging physics and reduce artifacts caused by noisy polarization measurements.

Overall, learning-based fusion methods demonstrate superior adaptability and robustness compared to classical techniques. However, their performance is strongly dependent on data availability, which remains limited for polarization – infrared imaging. Consequently, current research increasingly focuses on unsupervised, semi-supervised, and physics-informed learning strategies.

3. COMPARATIVE ANALYSIS OF APPROACHES

3.1 Qualitative and Quantitative Evaluation

Evaluating the quality of fusion between polarization and infrared images is a nontrivial task, since for such multimodal systems, there is no universal ground-truth image that simultaneously and fully represents all physical properties of the observed scene. Unlike reconstruction or segmentation problems, where an objectively correct result exists, image fusion quality is determined by the degree of preservation, integration, and coherence of information originating from multiple sources. Consequently, fusion evaluation criteria are inherently multidimensional and encompass informational, structural, physical, and perceptual aspects [36].

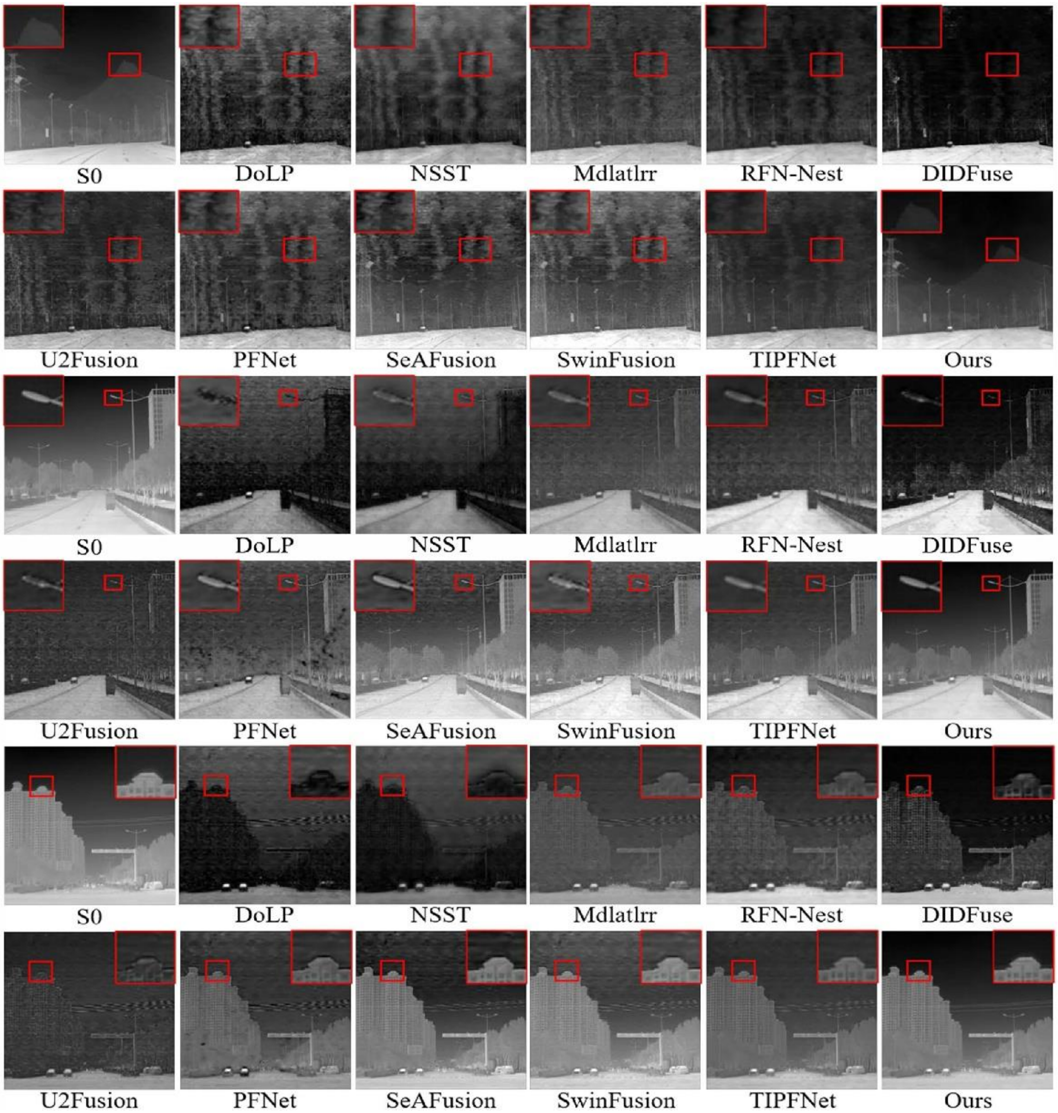

Fig. 2. Visual comparison of polarization–infrared image fusion results produced by different algorithms [34]

Informational Criteria

Informational metrics assess how effectively the fused image preserves the informational content of the source modalities. One of the most commonly used measures is entropy, which characterizes the statistical diversity of intensity levels in the image. In the context of polarization–infrared fusion, increased entropy typically indicates the incorporation of additional textural and structural details, particularly those derived from DoLP or AoP parameters. However, excessive entropy growth may also reflect noise amplification, which is especially pronounced in polarization channels [36].

A more informative measure is mutual information (MI) between the fused image and each input modality. High MI values indicate effective transfer of informative components from both the infrared intensity and polarization parameters. Contemporary studies often employ normalized or summed MI formulations, enabling consistent comparison across different scenes and fusion algorithms [37].

Modern deep-learning-based fusion studies also rely on information-theoretic criteria to validate performance improvements over classical approaches [38].

Structural and Spatial Criteria

Structural metrics are aimed at assessing the preservation of spatial patterns, contours, and local scene features. The average gradient (AG) is a simple yet effective indicator of edge sharpness and fine-detail representation. For polarization–infrared fusion, a high AG value generally correlates with accurate reproduction of geometric structures, which are often emphasized in AoP maps or polarization contrast images.

A more advanced metric is the structural similarity index (SSIM), which evaluates local similarity in terms of luminance, contrast, and structural components [39]. Although SSIM is traditionally used for reference-based comparison, in fusion tasks it is typically computed between the fused image and each source image separately. This approach allows assessment of the balance between thermal feature preservation and polarization-based structural information.

Perceptual Criteria

Perceptual metrics attempt to approximate characteristics of human visual perception. Visual Information Fidelity (VIF) and related measures estimate how much visually meaningful information is conveyed by the fused image [21]. In applications involving human operators, such as surveillance or visual inspection, perceptual criteria often correlate better with subjective image quality than purely statistical metrics.

In addition, subjective expert evaluation remains an important component of fusion assessment, particularly when the fused imagery is intended for direct human interpretation. Such evaluations consider scene readability, target discernibility, and the absence of visually disturbing artifacts [37].

Physically Motivated Criteria

For polarization–infrared fusion, physically motivated criteria are of particular importance. These metrics assess whether the fused image remains consistent with the underlying physical principles governing thermal emission and polarization. This includes preservation of thermal contrast associated with genuine temperature or emissivity differences, as well as faithful representation of polarization characteristics without artificial distortion.

Physically inconsistent enhancement of polarization features may lead to erroneous interpretation of surface material properties, which is unacceptable in scientific, medical, or defense-related applications. Therefore, some studies evaluate consistency with Stokes parameters or derived quantities such as DoLP to ensure physical plausibility [40].

Toward a Comprehensive Evaluation Framework

No single criterion is sufficient to fully characterize fusion quality. As a result, modern research increasingly adopts a multi-criteria evaluation framework, combining informational, structural, perceptual, and physical metrics [35–37]. This comprehensive approach enables more balanced and objective comparison of fusion algorithms, including classical transform-based methods and modern deep learning models.

Ultimately, fusion evaluation criteria serve not only as measurement tools but also as methodological instruments that shape algorithm design and guide future research directions by defining what constitutes “quality” in multimodal image fusion.

3.2 Application Domains

The fusion of polarization and infrared imagery has found increasing relevance across a wide spectrum of application domains, each of which imposes distinct requirements on fusion algorithms. The inherent complementarity of the two modalities – thermal radiometric contrast from IR sensors and structural, material, and surface-orientation cues from polarization measurements – enables enhanced scene understanding, target detection, and material discrimination. The following subsections examine major domains where polarization–IR fusion has demonstrated substantial benefits [36][37].

Remote Sensing and Environmental Monitoring

In remote sensing applications, polarization–infrared fusion is particularly effective for environmental monitoring, terrain classification, and land–water boundary detection. Thermal infrared images provide robust temperature and emissivity information, which is relatively insensitive to changes in visible illumination or cloud cover. Polarization imagery, by contrast, captures surface texture, orientation, and scattering properties, which are essential for distinguishing between vegetation types, soil textures, and water surfaces.

For instance, studies have demonstrated that combining DoLP and AoP maps with LWIR images improves the detection of stressed vegetation, subtle changes in water turbidity, and geological features obscured by partial cloud cover [40]. Multi-resolution fusion methods allow analysts to extract both coarse thermal patterns and fine structural details simultaneously [17]. Hybrid methods, which incorporate both transform-based decompositions and adaptive weighting schemes, have been particularly effective in enhancing contrast between land cover classes while preserving spatial resolution [38].

Challenges in this domain include atmospheric interference, variable emissivity of natural materials, and sensor misregistration due to platform motion, which require careful preprocessing and calibration. Nevertheless, the integration of polarization information allows for a more nuanced interpretation of surface properties, enhancing the accuracy of classification and environmental assessment models.

Military and Surveillance Applications

Military and security operations represent one of the most prominent fields for polarization–infrared fusion. Low-visibility environments, including nighttime operations, fog, haze, and camouflage scenarios, often reduce the effectiveness of conventional optical sensors. Infrared imaging provides robust detection of heat-emitting targets, while polarization imaging reveals material properties and geometric structures that may be masked in intensity-only IR images [37].

For example, vehicles or personnel camouflaged against natural backgrounds may exhibit subtle polarization contrasts due to reflective surfaces or surface orientation changes, which are imperceptible in standard thermal imagery. Hybrid and learning-based fusion approaches enhance target detectability by selectively emphasizing salient thermal and polarization features while suppressing background clutter.

In addition, real-time fusion is critical for autonomous surveillance drones and ground vehicles. Lightweight CNN architectures and hybrid transform-learning models are increasingly employed to provide on-board processing that balances detection accuracy with computational efficiency. The combination of physical modeling, such as polarization consistency constraints, with adaptive learning frameworks ensures reliable target recognition in operationally challenging environments.

Biomedical Imaging and Diagnostics

Polarization–infrared fusion has emerged as a promising tool in biomedical imaging, where it contributes to non-invasive diagnostics and tissue characterization. Thermal infrared imaging can detect subtle temperature variations associated with vascular activity, inflammation, or tumor metabolism, while polarization parameters provide information about tissue anisotropy, fibrous structures, and surface morphology [41].

For instance, combining DoLP and AoP maps with thermal images enhances visualization of skin lesions, burns, or abnormal tissue growth by highlighting anisotropic scattering patterns that are invisible in standard IR images. Hybrid fusion techniques that integrate pixel-level weighting, multiresolution decomposition, and feature-level saliency measures have been shown to improve contrast between healthy and pathological tissue regions, potentially supporting early diagnosis of cancer or other conditions.

Challenges include motion artifacts, low polarization signal in soft tissues, and inter-subject variability in tissue properties. Deep learning approaches, particularly physics-informed networks, are increasingly used to address these challenges, learning robust fusion rules that preserve diagnostically relevant features while minimizing noise and artifact amplification.

Autonomous Navigation and Robotics

In autonomous navigation, including self-driving vehicles and robotic systems, reliable perception under adverse environmental conditions is paramount. Fog, dust, low-light, or nighttime scenarios can significantly reduce the effectiveness of standard visible-spectrum cameras. Infrared sensors provide thermal contrast for obstacle detection, while polarization imaging can reveal surface orientation, material properties, and boundaries that are otherwise hidden.

Fused polarization–IR imagery enables more robust obstacle recognition, terrain classification, and scene segmentation, which are critical for path planning and collision avoidance. Multi-level fusion strategies, combining pixel-, feature-, and decision-level integration, allow real-time systems to exploit both modalities efficiently. Transformer-based attention networks have recently been explored to dynamically prioritize salient regions, such as heated objects or high-contrast surfaces, improving the reliability of perception modules in autonomous systems.

Industrial and Cultural Heritage Applications

Beyond defense and biomedical domains, polarization–infrared fusion has applications in industrial inspection and cultural heritage preservation. In industrial settings, fused images can reveal defects, stress patterns, or thermal anomalies in machinery, electronic components, or composite materials. The polarization channel enhances detection of surface cracks, anisotropic textures, and material boundaries that may not manifest as thermal contrasts alone.

In cultural heritage preservation, combining infrared thermography with polarization imagery allows non-invasive examination of artworks, manuscripts, and architectural surfaces. Subsurface features, underdrawings, or layered materials can be revealed more effectively than with either modality alone, supporting restoration decisions and authenticity assessments.

These applications emphasize the versatility of fusion methods across domains with widely differing constraints on spatial resolution, temporal response, and interpretability.

Synthesis

Across all application domains, the complementary nature of polarization and infrared data provides a consistent advantage: the ability to extract thermal and structural/material information simultaneously. The choice of fusion methodology – classical, hybrid, or learning-based – depends on specific operational requirements, including real-time processing, interpretability, robustness to noise, and physical consistency. While classical and hybrid methods remain effective for well-defined and computationally constrained environments [38], deep learning approaches increasingly dominate domains requiring adaptive feature extraction and scene-dependent optimization.

Ultimately, the domain-driven analysis of fusion methods highlights that the design of fusion algorithms cannot be modality-agnostic; instead, it must account for both the physics of the sensors and the operational objectives of the target application.

4. EPISTEMOLOGICAL AND METHODOLOGICAL FOUNDATIONS OF POLARIZATION–INFRARED FUSION

Section 3 demonstrated the algorithmic evolution of polarization–infrared image fusion from transform-domain techniques to deep learning and physics-informed architectures. While those developments describe how fusion is implemented computationally, a deeper question concerns how multimodal physical information is conceptually integrated. This section examines the epistemological and methodological foundations underlying polarization–infrared fusion, providing the theoretical context necessary for understanding the application-driven validation discussed in Section 5.

4.1. Classical Analytical Paradigm: Fusion as Signal Aggregation

Early polarization–infrared fusion methods emerged within a classical signal-processing paradigm [14, 42, 43]. In this framework, each modality is treated as an independent measurement channel. Infrared radiance and polarization descriptors (e.g., DoLP, AoLP, Stokes parameters) are considered separate scalar or vector fields representing distinct physical observables.

Transform-domain approaches – such as Laplacian pyramids, wavelet transforms, and multiscale decomposition – map these modalities into a shared frequency representation, where fusion rules select coefficients based on predefined criteria (e.g., maximum absolute value, local variance, or energy preservation) [16, 41]. The epistemological assumption underlying these methods is additive complementarity: each modality contributes partial but compatible information that can be combined through deterministic rules.

However, infrared and polarization imaging originate from fundamentally different physical processes:

- Infrared radiance obeys Planck's law and thermodynamic emission constraints.
- Polarization contrast arises from boundary conditions in electromagnetic wave reflection and emission governed by Fresnel relations.

The analytical paradigm treats these outputs as statistical signals without explicitly modeling their shared physical causality. Although effective for enhancement tasks, this abstraction limits interpretability and may produce physically inconsistent fused outputs under complex environmental conditions.

4.2. Systems-Oriented Integration: From Independence to Interaction

As discussed in Section 3, limitations of fixed-rule methods motivated adaptive and model-based strategies. These developments reflect a methodological shift from reductionism to systems-oriented integration [44].

In real scenes:

- Surface emissivity affects both thermal intensity and polarization state.
- Viewing geometry modifies DoLP and apparent thermal contrast.
- Atmospheric scattering influences polarization differently from radiometric attenuation.

Thus, modalities are not independent; they are coupled through shared environmental and material parameters.

Systems-oriented fusion models treat polarization and infrared signals as outputs of a unified physical scene rather than independent inputs. Hybrid approaches began incorporating:

1. Scene-adaptive weighting,
2. Region-based feature selection,
3. Physical parameter modeling (e.g., emissivity–angle relations),
4. Statistical dependence analysis.

In this framework, fusion is relational rather than additive. Information emerges from cross-modal interaction under shared constraints. This methodological transformation laid the groundwork for learning-based integration.

4.3. Computational Mediation: Representation in Deep Fusion Models

The introduction of convolutional neural networks (CNNs) and transformer-based architectures marked a significant transition in fusion research [18, 46, 47]. Unlike deterministic transform methods, deep networks construct hierarchical feature representations.

Typical architectures consist of:

1. Modality-specific encoders,
2. Cross-modal fusion modules,
3. Reconstruction or decision layers.

In polarization–infrared fusion, learned intermediate representations capture:

- Spatial gradients,
- Thermal contrast distributions,
- Polarization anisotropy patterns,
- Cross-channel correlations.

From an epistemological standpoint, this introduces computational mediation: knowledge is inferred through learned feature spaces rather than predefined physical rules. The fusion mechanism becomes data-driven and optimization-based.

However, purely empirical learning introduces methodological concerns:

- Latent features lack explicit physical interpretability.
- Generalization across environmental conditions may degrade.
- Physical constraints (e.g., DoLP ≤ 1) are not automatically enforced.

These issues highlight the need to reconcile inductive learning with deductive physical modeling.

4.4. Physics-Informed Fusion: Integrating Law and Learning

Recent research trends, partially introduced in Section 3, address interpretability and robustness through physics-informed modeling [48]. These methods embed physical constraints directly into network architecture or loss functions.

Examples include:

- Enforcing valid Stokes parameter relationships,
- Penalizing nonphysical polarization magnitudes,
- Preserving radiometric consistency,
- Maintaining angular continuity of polarization vectors.

Physics-informed loss functions constrain optimization within physically plausible solution spaces. This restores theoretical grounding without sacrificing adaptive capability.

Methodologically, physics-informed fusion represents a synthesis of two epistemic traditions:

- Deductive reasoning based on electromagnetic and thermodynamic theory.
- Inductive inference through data-driven optimization.

Rather than replacing analytical modeling, deep learning becomes a flexible approximator operating within physical law constraints. In this context, Jones-formalism-based models demonstrate how classical polarization optics can be integrated with modern computational fusion strategies, establishing a bridge between physically grounded analytical representations and adaptive learning-based architectures [30].

4.5. Comparative Conceptual Framework

To clarify the methodological progression described above, a comparative conceptual framework is proposed in **Table 2**.

The diagram visually demonstrates the transition from additive signal combination to constraint-aware integrative systems.

Table 2. **Comparative Methodological Evolution of Polarization–Infrared Fusion**

Component	Classical Analytical Model	Data-Driven Deep Fusion	Physics-Informed Hybrid Model
Inputs	Infrared intensity Polarization parameters	Raw infrared image Raw polarization image	Infrared radiance Stokes polarization parameters Physical priors
Processing Approach	Transform-domain decomposition (wavelet, pyramid, or multiscale transforms) Fixed fusion rule (e.g., coefficient selection or weighted averaging)	CNN or Transformer encoders for feature extraction Feature-level fusion layers Reconstruction or decoding network	Neural feature extraction module Physics-constrained fusion mechanism Regularized loss function incorporating physical priors
Output	Enhanced composite image	Optimized fused image	Physically consistent fused representation
Adaptivity	Low – relies on predefined rules	High – model learns fusion strategy from data	High – adaptive learning guided by physical constraints
Interpre-tability	High – processing steps mathematically transparent	Limited – internal representations often opaque	Moderate to high – interpretability supported by physical modeling
Physical Consistency	Implicit – physical relationships not explicitly modeled	Low – fusion driven primarily by data patterns	High–fusion constrained by radiometric and polarization physics
Performance Characte-ristics	Stable and computationally efficient, but limited in complex scenes	High performance in diverse scenarios, but potentially sensitive to dataset bias	Balanced performance combining generalization capability with physical reliability

4.6. Evaluation as an Epistemic Criterion

As fusion methodologies evolved, evaluation strategies also transformed. Metrics such as entropy, mutual information, and the structural similarity index measure (SSIM) have traditionally been used to quantify information preservation and structural coherence.

However, these metrics encode implicit normative assumptions:

- Entropy prioritizes diversity.
- Mutual information prioritizes cross-modal retention.
- SSIM prioritizes spatial structure.

For polarization–infrared fusion, a physically meaningful evaluation must additionally consider:

- Preservation of genuine thermal gradients,
- Integrity of polarization signatures,
- Absence of artificial enhancement artifacts,
- Radiometric consistency.

Thus, evaluation functions not merely as a performance measurement but as an epistemological filter shaping methodological development. Algorithms optimized under specific metrics implicitly define what constitutes a “good” fused image.

4.7. Interdisciplinary Context

Polarization–infrared fusion integrates principles from:

- Physical optics,
- Electromagnetic theory,
- Signal processing,
- Statistical learning,
- Systems engineering.

Each discipline contributes constraints and modeling strategies. Effective fusion requires harmonizing theoretical rigor with computational adaptability. Overreliance on purely empirical optimization risks sacrificing interpretability, while strict analytical modeling may limit robustness in complex scenes.

This interdisciplinary synthesis prepares the transition to Section 5, where operational performance and application-driven validation are examined.

4.8. Transition to Applications

The methodological developments outlined above directly influence practical performance in:

Target detection under low contrast,
Camouflage discrimination,
Maritime and aerial surveillance,
Autonomous navigation,
Biomedical imaging diagnostics.

Physics-informed and system-oriented models demonstrate improved robustness under environmental variability and noise – factors critical in real-world deployment.

Section 5 therefore, evaluates fusion algorithms not solely by quantitative metrics but by operational reliability and application-specific effectiveness. The epistemological framework developed in this section provides the theoretical foundation for interpreting those results.

5. APPLICATION-DRIVEN VALIDATION AND OPERATIONAL PERFORMANCE OF POLARIZATION–INFRARED FUSION

Section 3 analyzed algorithmic evolution, and Section 4 established the epistemological and methodological foundations of polarization–infrared image fusion. The present section evaluates how these theoretical and computational developments translate into operational performance across real-world application domains. Fusion quality must ultimately be validated not only through numerical metrics but through task-oriented reliability under environmental variability, sensor noise, and scene complexity.

Polarization–infrared fusion is particularly valuable in conditions where visible imaging fails or provides insufficient discriminative information. Its principal operational strength lies in combining:

- Thermal contrast (temperature-dependent radiative emission),
- Polarimetric contrast (surface geometry and material-dependent polarization signatures).

This dual sensitivity enhances target discrimination, material classification, and structural interpretation.

5.1. Target Detection and Surveillance

One of the primary application domains of infrared imaging is target detection under low illumination or adverse weather conditions. However, thermal contrast alone may be insufficient when temperature differences between target and background are small or dynamically varying [48].

Polarization imaging provides complementary information, particularly for detecting:

- Man-made objects against natural backgrounds,
- Surface discontinuities,
- Camouflaged structures,
- Specular reflections from engineered materials [10].

Studies have shown that man-made objects often exhibit stronger or more structured polarization signatures compared to natural vegetation or soil [49]. When fused with infrared radiance, polarization features improve:

- Signal-to-clutter ratio,
- Edge clarity,
- False alarm reduction.

Deep fusion architectures further enhance performance by learning discriminative cross-modal features optimized for detection objectives [17].

Operational validation typically involves:

- Receiver operating characteristic (ROC) analysis,
- Probability of detection vs. false alarm rate,
- Target-background separability metrics.

Physics-informed fusion models demonstrate improved robustness in cases where polarization contrast varies with viewing geometry, as physical constraints stabilize representation [18].

5.2. Camouflage and Concealment Discrimination

Camouflage detection represents a critical use case where fusion provides clear advantages. Thermal camouflage attempts to match background temperature, reducing infrared contrast. However, matching polarization signatures is substantially more difficult due to dependence on:

- Surface microstructure,
- Reflectance anisotropy,
- Material refractive index.

Polarimetric contrast remains detectable even when thermal gradients are minimized. Fusion strategies amplify subtle inconsistencies between thermal and polarization responses.

Empirical studies demonstrate that fused imagery increases classification accuracy in camouflage scenarios by integrating:

- Local polarization anisotropy,
- Thermal spatial gradients,
- Texture features.

Machine learning-based fusion frameworks outperform fixed-rule approaches when background complexity increases. However, physics-informed regularization prevents over-amplification of noise in polarization channels, which are often more sensitive to environmental variability.

5.3. Maritime and Aerial Remote Sensing

Maritime environments present challenging conditions:

- Low thermal contrast between vessels and water,
- High specular reflection,
- Dynamic atmospheric scattering.

Polarization is particularly effective in water–air boundary detection and glare suppression. Infrared imaging enhances the detection of thermal signatures from engines or exhaust systems.

Fusion in maritime surveillance improves:

- Vessel boundary delineation,
- Wake detection,

- Oil spill discrimination,
- Floating object identification [48].

In aerial remote sensing, polarization assists in distinguishing man-made infrastructure from natural terrain. Infrared channels contribute to temperature-based material separation.

Operational studies show that multiscale fusion enhances spatial resolution and reduces clutter in coastal monitoring scenarios [45]. Transformer-based models have demonstrated improved generalization across environmental conditions compared to conventional CNN-based architectures.

5.4. Autonomous Navigation and Robotics

Autonomous systems require robust perception under variable illumination, weather, and terrain conditions. Infrared imaging ensures functionality in low-light scenarios, while polarization assists in:

- Road surface detection,
- Ice or water hazard identification,
- Surface orientation estimation [46].

Fusion contributes to improved edge detection and structural recognition in outdoor navigation tasks. Cross-modal attention mechanisms allow dynamic weighting depending on environmental context.

In robotics, fused polarization–infrared inputs improve object segmentation and obstacle detection, particularly in low-contrast environments [50].

Real-time constraints impose additional requirements:

- Computational efficiency,
- Hardware compatibility,
- Low latency.

Physics-informed lightweight networks demonstrate promising trade-offs between interpretability and speed for embedded systems.

5.5. Biomedical Imaging Applications

Beyond defense and surveillance, polarization–infrared fusion shows potential in biomedical diagnostics. Infrared imaging captures temperature variations related to inflammation or vascular abnormalities. Polarization imaging reveals structural changes in tissue microarchitecture [47]

Applications include:

- Skin lesion detection,
- Burn assessment,
- Tumor boundary visualization,
- Vascular mapping.

Fusion enhances diagnostic reliability by integrating thermal anomalies with polarization-sensitive structural contrast. Clinical studies suggest improved sensitivity in detecting subsurface irregularities compared to single-modality imaging [51].

However, biomedical validation requires strict radiometric calibration and careful handling of polarization noise, as biological tissues introduce complex scattering behavior.

5.6. Evaluation Metrics and Operational Criteria

While classical metrics such as entropy, mutual information, and SSIM remain widely used, application-driven validation emphasizes task-specific criteria:

- Detection accuracy,
- Segmentation precision,
- Classification reliability,
- Robustness to noise,

- Stability across viewing angles.
- Target visibility [31].

Recent research increasingly incorporates:

- Task-driven loss functions,
- End-to-end detection pipelines,
- Real-time benchmarking.

Physics-consistent fusion models exhibit superior stability under environmental perturbations compared to unconstrained networks, particularly when polarization signal-to-noise ratio is low [52].

Additionally, uncertainty quantification methods are emerging to estimate confidence in fused outputs – an important factor for safety-critical systems [53].

5.7. Limitations and Practical Challenges

Despite demonstrated advantages, polarization–infrared fusion faces several practical limitations:

1. Sensor Calibration Complexity

Accurate alignment of polarization channels and infrared radiometry requires precise calibration.

2. Polarization Noise Sensitivity

Polarimetric measurements are sensitive to atmospheric scattering and sensor misalignment.

3. Dataset Scarcity

Public datasets containing synchronized, calibrated polarization–infrared pairs remain limited.

4. Ground Truth Ambiguity

Defining objective ground truth for fused imagery is inherently challenging.

Addressing these challenges requires standardized benchmarks and shared evaluation protocols.

5.8. Synthesis and Transition

Application-driven validation confirms that the methodological principles described in Sections 3 and 4 directly influence operational robustness. Systems-oriented and physics-informed approaches demonstrate enhanced stability, interpretability, and cross-domain generalization.

Fusion has evolved from visual enhancement toward operational perception support. In contemporary systems, fused imagery often feeds higher-level modules:

- Object detection networks,
- Tracking systems,
- Decision-support algorithms.

Thus, polarization–infrared fusion becomes an integral component of multimodal perception architectures rather than an isolated preprocessing step.

CONCLUSION

The review of image-fusion techniques using polarization and infrared intensity reveals a trajectory that mirrors the broader evolution of scientific thought: from analytical separation to systemic integration. Technically, this progression moves from transform-based and energy-driven methods toward adaptive, learning-based architectures that embody physical and cognitive principles simultaneously. Conceptually, it demonstrates the convergence of physics, computation, and epistemology in the modern understanding of perception and information.

Future work will require:

- development of standardized multi-modal datasets;

- creation of interpretable, physics-aware neural architectures;
- and exploration of cognitive models that explain how fusion contributes to understanding rather than mere visualization.

Ultimately, polarization–infrared image fusion exemplifies the post-non-classical vision of science – where observation, computation, and cognition become integrated aspects of a single evolving system of knowledge.

ACKNOWLEDGMENTS AND FUNDING SOURCES

The authors received no financial support for the research, writing, and publication of this article.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that the research was conducted in the absence of any conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, [V.A., O.K.]; methodology, [V.A., O.K.]; writing – original draft preparation, [V.A.]; writing – review and editing, [O.K.]; visualization, [V.A.], supervision, [O.K.];

All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Ma, W., Wang, K., Li, J., Yang, S. X., Song, L., & Li, Q. (2023). Infrared and visible image fusion technology and application: A review. *Sensors*, *23*(2), 599. https://doi.org/10.3390/s23020599

[2] Kolobrodov, V. G. (2024). Energy model of optical system of polarimetric thermal imager. *Visnyk NTUU KPI Seriia - Radiotekhnika Radioaparatobuduvannia*, *95*, 47–53 (in Ukrainian). https://doi.org/10.20535/RADAP.2024.95.47-53

[3] Guo, F., Zhu, J., Huang, L., Li, F., Zhang, N., Deng, J., Li, H., Zhang, X., Zhao, Y., & Jiang, H. (2024). Multi-dimensional fusion of spectral and polarimetric images followed by pseudo-color algorithm integration and mapping in HSI space. *Remote Sensing*, *16*(7), 1119. https://doi.org/10.3390/rs16071119

[4] A.-J. Fresnel (1821). Mémoire sur les modifications que la réflexion imprime à la lumière polarisée. *Annales de Chimie et de Physique*, *17*, 45-176. https://doi.org/10.5281/ZENODO.5060346

[5] Stokes, G. G. (1851). On the composition and resolution of streams of polarized light from different sources. *Transactions of the Cambridge Philosophical Society*, *9*, 399. https://doi.org/10.1017/CBO9780511702266.010

[6] Born M, Wolf E, Bhatia AB, et al. (1999). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. 7th ed. Cambridge University Press. https://doi.org/10.1017/CBO9781139644181

[7] Herschel, W. (1800). II. Experiments on the refrangibility of the invisible rays of the sun. *The Philosophical Magazine*, *8*(29), 9-15. https://doi.org/10.1080/14786440008562602

[8] Planck, M. (1901). On the law of distribution of energy in the normal spectrum. *Annalen der physik*, *4*, 553-563. https://doi.org/10.1002/andp.19013090310

[9] Jones, R. C. (1941). A new calculus for the treatment of optical systems I. Description and discussion of the calculus. *Journal of the Optical Society of America*, *31*(7), 488-493. https://doi.org/10.1364/JOSA.31.000488

[10] Rogalski, A. (2003). Infrared detectors: status and trends. *Progress in quantum electronics*, *27*(2-3), 59-210. https://doi.org/10.1016/S0079-6727(02)00024-1

[11] Rogalski, A. (2010). *Infrared Detectors* (2nd ed.). CRC Press. https://doi.org/10.1201/b10319

[12] Fossum, E. R. (1997). CMOS image sensors: Electronic camera-on-a-chip. *IEEE transactions on electron devices*, *44*(10), 1689-1698. https://doi.org/10.1109/16.628824

[13] Gruev, V., Ortu, A., Yang, Z., Van der Spiegel, J., & Engheta, N. (2007, September). High-resolution integrated image sensor with polymer micropolarization array. In *Frontiers in Optics* (p. FTuS6). Optica Publishing Group. https://doi.org/10.1364/FIO.2007.FTuS6

[14] Burt, P. J., & Adelson, E. H. (1983). The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, *31*(4), 532-540. https://doi.org/10.1109/TCOM.1983.1095851

[15] Mallat, S. G. (2002). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, *11*(7), 674-693. https://doi.org/10.1109/34.192463

[16] Piella, G. (2003). A general framework for multiresolution image fusion: from pixels to regions. *Information fusion*, *4*(4), 259-280. https://doi.org/10.1016/S1566-2535(03)00046-0

[17] Tyo, J. S., Goldstein, D. L., Chenault, D. B., & Shaw, J. A. (2006). Review of passive imaging polarimetry for remote sensing applications. *Applied optics*, *45*(22), 5453-5469. https://doi.org/10.1364/AO.45.005453

[18] Li, H., & Wu, X. J. (2018). DenseFuse: A fusion approach to infrared and visible images. *IEEE transactions on image processing*, *28*(5), 2614-2623. https://doi.org/10.1109/TIP.2018.2887342

[19] aswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*, 5998-6008. https://doi.org/10.48550/arXiv.1706.03762

[20] Yang, J., Fu, X., Hu, Y., Huang, Y., Ding, X., & Paisley, J. (2017). PanNet: A deep network architecture for pan-sharpening. In *Proceedings of the IEEE international conference on computer vision* (pp. 5449-5457). https://doi.org/10.1109/ICCV.2017.193

[21] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600-612. https://doi.org/10.1109/TIP.2003.819861

[22] Zhang, Q., & Guo, B. L. (2009). Multifocus image fusion using the nonsubsampled contourlet transform. *Signal processing*, *89*(7), 1334-1346. https://doi.org/10.1016/j.sigpro.2009.01.012

[23] Pratt, W. K. (2007). *Digital image processing: PIKS Scientific inside* (Vol. 4). Hoboken, New Jersey: Wiley-Interscience. https://doi.org/10.1002/0470097434

[24] Yue, Z., & Li, F. M. (2014, February). An infrared polarization image fusion algorithm based on oriented Laplacian pyramid. In *Selected Papers from Conferences of the Photoelectronic Technology Committee of the Chinese Society of Astronautics: Optical Imaging, Remote Sensing, and Laser-Matter Interaction 2013* (Vol. 9142, pp. 60-70). SPIE. https://doi.org/10.1117/12.2054074

[25] Zhang, J., Zhang, Y., & Shi, Z. (2018). Long-wave infrared polarization feature extraction and image fusion based on the orthogonality difference method. *Journal of Electronic Imaging*, *27*(2), 023021-023021. https://doi.org/10.1117/1.JEI.27.2.023021

[26] Cunha, A. L., Zhou, J., & Do, M. N. (2006). The nonsubsampled contourlet transform: Theory, design, and applications. *IEEE Transactions on Image Processing*, *15*(10), 3089–3101. https://doi.org/10.1109/TIP.2006.877507

[27] Khaustov, Ya. Ye., Khaustov, D. Ye., Lychkovskyy, E., Ryzhov, Ye., & Nastishin, Yu. A. (2019). Image fusion for a target sightseeing system of armored vehicles. *Milit. Techn. Collect.*, *21*, 28-37. https://doi.org/10.33577/2312-4458.21.2019.28-37

[28] Khaustov, Ya. Ye., Khaustov, D. Ye., Ryzhov, Ye., Lychkovskyy, E., & Nastishin, Yu. A. (2020). Fusion of visible and infrared images via complex function. *Milit. Techn. Collect.*, *22*, 20-31. https://doi.org/10.33577/2312-4458.22.2020.20-31

[29] Khaustov, Ya. Ye., Khaustov, D. Ye., Hryvachevskyi, A. P., Ryzhov, Ye., Lychkovskyy, E., Prudyus, I. N., & Nastishin, Yu. A. (2021). Complex function as a template for image fusion. *Results in Optics*, 2, 100038. https://doi.org/10.1016/j.rio.2020.100038

[30] Khaustov, D. Ye., Khaustov, Ya. Ye., Ryzhov, Ye., Lychkovskyy, E., Vlokh, R., & Nastishin, Yu. A. (2021). Jones formalism for image fusion. *Ukr. J. Phys. Opt.*, *22*(3), 165-180. https://doi.org/10.3116/16091833/22/3/165/2021

[31] Khaustov, D. Ye.,Kyrychuk, O. A., Khaustov, Ya. Ye., Stakh, Ya. Ye., Zhyrna O. V., & Nastishin, Yu. A. (2023). The Measure of Target Saliency for Target-Oriented Image Fusion. *Collection of scientific papers of the State Research Institute for Testing and Certification of Armaments and Military Equipment*, *3*(17), 122-136. https://doi.org/10.37701/dndivsovt.17.2023.15

[32] Khaustov, D., Khaustov, Ya., Ryzhov, Ye., Burashnikov, O., Lychkovskyy, E., & Nastishin, Yu. (2021). Dynamic fusion of images from the visible and infrared channels of sightseeing system by complex matrix formalism. *Milit. Techn. Collect.*, *25*, 29-37. https://doi.org/10.33577/2312-4458.25.2021.29-37

[33] Khaustov D. Ye., Kyrychuk O. A., Stakh T. M., Khaustov Ya. Ye., Burashnikov O. O., Ryzhov Ye., Vlokh R., & Nastishin Yu. A. (2023). Complex-scalar and complex-vector approaches for express target-oriented image fusion. *Ukr. J. Phys. Opt.*, *24*(1), 62-82. https://doi.org/10.3116/16091833/24/1/62/2023

[34] Li, K., Qi, M., Zhuang, S., Liu, Y., & Gao, J. (2023). Noise-aware infrared polarization image fusion based on salient prior with attention-guided filtering network. *Optics Express*, *31*(16), 25781-25796. https://doi.org/10.1364/OE.492954

[35] Duan, J., Zhang, H., Liu, J., Gao, M., Cheng, C., & Chen, G. (2023). A dual-weighted polarization image fusion method based on quality assessment and attention mechanisms. *Frontiers in Physics*, *11*, 1214206. https://doi.org/10.3389/fphy.2023.1214206

[36] Jagalingam, P., & Hegde, A. V. (2015). A review of quality metrics for fused image. *Aquatic Procedia*, *4*, 133-142. https://doi.org/10.1016/j.aqpro.2015.02.019

[37] Kaur, H., Koundal, D., & Kadyan, V. (2021). Image Fusion Techniques: A Survey. *Archives of computational methods in Engineering*, *28*(7), 4425-4447. https://doi.org/10.1007/s11831-021-09540-7

[38] Jin, X., Jiang, Q., Yao, S., Zhou, D., Nie, R., Hai, J., & He, K. (2017). A survey of infrared and visual image fusion methods. *Infrared Physics & Technology*, *85*, 478-501. https://doi.org/10.1016/j.infrared.2017.07.010

[39] Zhang, H., Xu, H., Tian, X., Jiang, J., & Ma, J. (2021). Image fusion meets deep learning: A survey and perspective. *Information Fusion*, *76*, 323-336. https://doi.org/10.1016/j.inffus.2021.06.008

[40] Sheikh, H. R., & Bovik, A. C. (2006). Image information and visual quality. *IEEE Transactions on Image Processing*, *15*(2), 430-444. https://doi.org/10.1109/TIP.2005.859378

[41] Zhang, Z., & Blum, R. S. (1999). A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proceedings of the IEEE*, *87*(8), 1315-1326. https://doi.org/10.1109/5.775414

[42] Wang, L., Dou, J., Qin, P., Lin, S., Gao, Y., Wang, R., & Zhang, J. (2021). Multimodal medical image fusion based on nonsubsampled shearlet transform and convolutional sparse representation. *Multimedia Tools and Applications*, *80*(30), 36401-36421. https://doi.org/10.1007/s11042-021-11379-w

[43] Li, H., Manjunath, B. S., & Mitra, S. K. (1995). Multisensor image fusion using the wavelet transform. *Graphical models and image processing*, *57*(3), 235-245. https://doi.org/10.1006/gmip.1995.1022

[44] Li, S., Kang, X., & Hu, J. (2013). Image fusion with guided filtering. *IEEE Transactions on Image Processing*, *22*(7), 2864-2875. https://doi.org/10.1109/TIP.2013.2244222

[45] Wald, L. (1999). Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(3), 1190-1193. https://doi.org/10.1109/36.763269

[46] Ma, J., Ma, Y., & Li, C. (2019). Infrared and visible image fusion methods and applications: A survey. *Information fusion*, *45*, 153-178. https://doi.org/10.1016/j.inffus.2018.02.004

[47] Xu, H., Ma, J., Jiang, J., Guo, X., & Ling, H. (2020). U2Fusion: A unified unsupervised image fusion network. *IEEE transactions on pattern analysis and machine intelligence*, *44*(1), 502-518. https://doi.org/10.1109/TPAMI.2020.3012548

[48] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, *378*, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

[49] Goldstein, D.H. (2011). *Polarized Light* (3rd ed.). CRC Press. https://doi.org/10.1201/b10436

[50] Geiger, A., Lenz, P., Stiller, C., & Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research*, *32*(11), 1231-1237. https://doi.org/10.1177/0278364913491297

[51] Tuchin, V. V. (2015). *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnostics,* Third Edition. Society of Photo-Optical Instrumentation Engineers (SPIE). https://doi.org/10.1117/3.1003040

[52] Drgoňa, J., Tuor, A. R., Chandan, V., & Vrabie, D. L. (2021). Physics-constrained deep learning of multi-zone building thermal dynamics. *Energy and Buildings*, *243*, 110992. https://doi.org/10.1016/j.enbuild.2021.110992

[53] Kendall, A., & Gal, Y. (2017). What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? *arXiv preprint arXiv:1703.04977*. https://doi.org/10.48550/arXiv.1703.04977

МЕТОДИКИ КОМПЛЕКСУВАННЯ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ ПОЛЯРИЗАЦІЙНОЇ ІНФОРМАЦІЇ ТА ІНТЕНСИВНОСТІ ІНФРАЧЕРВОНОГО ВИПРОМІНЮВАННЯ

Віталій Артим*, Олег Крупич
Кафедра оптоелектроніки та інформаційних технологій,
Львівський національний університет імені Івана Франка,
вул. Ген. Тарнавського 107, Львів, 79017, Україна

АНОТАЦІЯ

Вступ. Комплексування поляризаційних та звичайних теплових інфрачервоних зображень поєднує взаємодоповнювальну фізичну інформацію для покращення інтерпретації сцени. Інфрачервоне інтенсивнісне зображення відображає теплове випромінювання та розподіл коефіцієнта випромінювальної здатності, тоді як параметри поляризації, зокрема ступінь і кут лінійної поляризації, кодують мікроструктуру поверхні, матеріальний склад і геометричну орієнтацію, що є недоступними для сенсорів, заснованих лише на інтенсивності. Їхнє комплексування підвищує просторову деталізацію, розпізнавання об’єктів і стійкість за умов погіршеної видимості, відображаючи ширший перехід до мультимодального сприйняття.

Матеріали та методи. У роботі подано аналітичний огляд методів комплексування поляризаційних та інфрачервоних зображень. Розвиток підходів розглянуто від класичних лінійних перетворень, багатомасштабної декомпозиції та енергетичних правил прийняття рішень до адаптивних гібридних схем. Особливу увагу приділено архітектурам глибокого навчання та фізично-інформованим моделям, що враховують принципи формування поляризаційного сигналу та радіометричні обмеження. Прогрес у сфері комплексування проаналізовано в контексті розвитку сенсорних технологій і обчислювальної візуалізації.

Результати. Аналіз свідчить, що включення поляризаційної складової суттєво підвищує інформативність синтезованих зображень порівняно з традиційними інтенсивнісними методами. Класичні алгоритми забезпечують стабільні та інтерпретовані результати, проте мають обмежену адаптивність у складних умовах. Підходи на основі навчання та гібридні стратегії демонструють кращу збереженість контрасту, покращену структурну деталізацію та вищу виявлюваність цілей, особливо в умовах теплової неоднорідності або візуальної деградації. Серед актуальних проблем – обмеженість анотованих наборів даних, чутливість до поляризаційного шуму та недостатня фізична інтерпретованість глибинних нейронних моделей.

Висновки. Комплексування поляризаційних та інфрачервоних зображень відображає перехід від одномодального до багатомодального спостереження для системно-орієнтованого сприйняття, що поєднує фізичне моделювання з підходами на основі даних. Подальші дослідження мають бути зосереджені на фізично-обґрунтованому проєктуванні нейронних архітектур, стандартизації протоколів оцінювання та підвищенні інтерпретованості для забезпечення надійного практичного впровадження.

Ключові слова: комплексування зображень, поляризаційне зображення, інфрачервоне зображення, мультимодальні дані, глибоке навчання

Received / Одержано	Revised / Доопрацьовано	Accepted / Прийнято	Published / Опубліковано
18 May, 2026	03 June, 2026	05 June, 2026	29 June, 2026