A hybrid framework for multimodal feature learning in chromatic aberration detection

Jarosław Bernacki
Department of Computer Science and Systems Engineering
Wrocław University of Science and Technology
Wrocław, Poland
jaroslaw.bernacki@pwr.edu.pl

Rafał Scherer
Faculty of Computer Science
AGH University of Kraków
Kraków, Poland
rscherer@agh.edu.pl
Department of Artificial Intelligence
Częstochowa University of Technology
Częstochowa, Poland

***Abstract*—Chromatic aberration (CA) remains a common optical artifact in digital imaging, causing spatial misalignments between color channels and resulting in visible color fringing. In this paper, we propose TriadNet, a novel hybrid convolutional-transformer architecture that formulates CA detection as a multimodal feature learning problem for accurate analysis of real-world photographs. TriadNet integrates three complementary modality-specific pathways: an edge-centric pathway that captures structural, high-frequency information where CA manifests most prominently, a channel-aware attention mechanism that models chromatic inter-channel discrepancies, and an edge-modulated feature processing strategy that facilitates effective multimodal feature fusion. Experimental evaluation on a custom dataset of real-world images demonstrates that TriadNet outperforms several baselines, achieving F_1-scores exceeding 0.98 and robust generalization across varying aberration strengths. The proposed method offers a lightweight yet effective solution for automatic CA detection within modern image processing pipelines.**

***Index Terms*—digital forensics, chromatic aberration, transformers, deep learning**

I. INTRODUCTION

Digital forensics is a research field focused on the acquisition, analysis, and interpretation of digital data in order to determine its origin, authenticity, and integrity. In the context of digital images, digital forensics (often referred to as digital image forensics) aims to identify characteristic traces introduced at different stages of the image formation pipeline, including image acquisition, in-camera processing, and post-processing. The main research areas within this domain include source camera identification, detection of image manipulations and forgeries, integrity and consistency analysis, as well as the exploitation of device- and optics-related artifacts as forensic traces. Such artifacts may originate from sensor imperfections, color filter array interpolation, compression algorithms, lens distortion, or optical aberrations [1], including chromatic aberration [2]. The main goal of digital image forensics is to provide reliable and objective tools for assessing the credibility of visual content, which increasingly serves as evidential material in forensic investigations and judicial proceedings [3]–[5].

Chromatic aberration (CA) is an optical phenomenon caused by the wavelength-dependent refraction of light in a lens system, which results in different color components of the same scene point being focused at different positions on the image sensor. As a consequence, color misalignment and characteristic color fringes, typically visible along high contrast edges, appear in the captured image. Chromatic aberration can be classified into longitudinal (axial) [6] and lateral (transverse) [7] aberration, depending on whether the color displacement occurs along the optical axis or in the image plane. The magnitude and spatial distribution of CA depend on the optical design of the lens, the materials used, focal length, aperture settings, and imaging conditions. Since these distortions are inherently linked to the physical properties of the optical system and remain relatively consistent for a given camera–lens configuration, CA constitutes a stable, camera-dependent artifact that can be observed and analyzed in digital images [8]–[10].

In this paper, we propose TriadNet, a novel hybrid convolutional-transformer architecture that formulates CA detection as a multimodal feature learning problem from digital images. The approach integrates complementary modality-specific pathways, including an edge-centric pathway that prioritizes high-frequency structural regions prone to color fringing, a channel-aware attention mechanism that explicitly models chromatic inter-channel spatial misalignments, and an edge-modulated feature processing strategy for effective multimodal feature fusion. As a result, TriadNet effectively captures the characteristic signatures of CA while remaining computationally lightweight. The model is trained on a custom dataset and demonstrates improved performance compared to several baselines in terms of precision, recall, and F_1-score.

A. Contribution

The main contribution of this work is the following. We introduce a novel hybrid convolutional-transformer model that explicitly combines edge-centric processing, channel-aware attention, and edge-modulated feature processing to detect chromatic aberration more effectively than generic architectures. We demonstrate the effectiveness of the proposed method through extensive comparison against baselines, achieving higher precision, recall, and F_1-score on a real-world dataset.

B. Organization of the paper

The paper is organized as follows. In Section II, the problem formulation and the proposed method are described. Section III presents the results of classification compared with the state-of-the-art methods. The section IV discusses previous and related works. The final section concludes this work. In formal descriptions, bold font denotes matrices or vectors.

II. PROPOSED HYBRID METHOD

We propose TriadNet, a hybrid convolutional-transformer architecture that addresses CA detection through multimodal feature learning from digital images. The method integrates three complementary modality-specific principles: (1) edge-centric processing, which captures a structural modality by prioritizing high-frequency image regions where lateral CA manifests most prominently as color fringing; (2) channel-aware feature interaction, which models a chromatic modality by explicitly learning cross-channel dependencies between the R, G, and B components to capture dispersion-induced spatial shifts; and (3) physics-guided attention modulation, which facilitates multimodal interaction through an edge-derived weighting mechanism that emphasizes feature representations in regions exhibiting strong gradient discontinuities.

The core of TriadNet is based on ConvNeXt-Tiny, providing robust extraction of local, modality-agnostic features. These features are subsequently refined through a lightweight multi-head self-attention module operating along the channel dimension (channel-aware mixer), enabling effective interaction between chromatic feature modalities. In parallel, a learnable edge-extraction pathway, implemented using trainable Sobel-like filters followed by sigmoid-activated magnitude estimation, produces a per-pixel structural modality in the form of an edge map. This map is spatially downsampled to match the feature resolution and serves as a soft attention signal that multiplicatively modulates intermediate representations, resulting in explicit multimodal feature fusion. The final global representation is obtained via adaptive average pooling of the modulated features, concatenated with a scalar global edge-strength descriptor, and processed by a lightweight MLP classification head that outputs a single logit indicating the presence of CA.

The model is trained end-to-end using binary cross-entropy loss. We employ a physics-based data augmentation strategy that synthetically introduces lateral chromatic dispersion into aberration-free images. This process is implemented via differentiable grid sampling with radially increasing, oppositely directed shifts applied to the red and blue channels. The augmentation enhances dataset diversity while reinforcing the learning of CA patterns grounded in physical dispersion characteristics rather than dataset-specific artifacts. As a result, the proposed multimodal architecture achieves high classification performance while remaining computationally lightweight and interpretable through its explicit structural and chromatic feature pathways.

A. Preliminaries and problem formulation

a) Chromatic aberration model: Formally, let $\mathbf{I} \in \mathbb{R}^{H\times W\times 3}$ denote the actual image captured by the camera. In an ideal, aberration-free system, the red ($\mathbf{I}_R$), green ($\mathbf{I}_G$), and blue ($\mathbf{I}_B$) channels are perfectly aligned. In the presence of lateral chromatic aberration, the observed image can be modeled as the result of shifting the red and blue channels relative to the green channel (taken as reference):

$$\mathbf{I}(x,y) = \begin{bmatrix} \mathbf{I}_R\big(x+\Delta x_R(x,y),\ y+\Delta y_R(x,y)\big) \\ \mathbf{I}_G(x,y) \\ \mathbf{I}_B\big(x+\Delta x_B(x,y),\ y+\Delta y_B(x,y)\big) \end{bmatrix} \tag{1}$$

where $(\Delta x_R, \Delta y_R)$ and $(\Delta x_B, \Delta y_B)$ represent the local displacement fields for the red and blue channels, respectively. In the simplest case, these displacements are modeled as radial functions dependent on the distance from the optical center:

$$\begin{aligned} \Delta x_c(x,y) &= k_c \cdot (x-x_0)\cdot r(x,y)^2, \\ \Delta y_c(x,y) &= k_c \cdot (y-y_0)\cdot r(x,y)^2, \end{aligned} \tag{2}$$

where $r(x,y) = \sqrt{(x-x_0)^2+(y-y_0)^2}$ is the radial distance from the optical center (x_0, y_0), and $k_R > 0$ and $k_B < 0$ are the aberration coefficients for the red and blue channels, respectively. In practice, real aberrations are more complex and depend on lens design, aperture, focus distance, and wavelength. In this work, we treat chromatic aberration as a binary detection problem: an image $\mathbf{I}$ belongs to the class "with chromatic aberration" ($y = 1$) if the shifts Δx_c, Δy_c exceed a minimum detectable threshold across a significant portion of the frame. The goal of the proposed method is to automatically estimate the probability $p(y = 1 \mid \mathbf{I})$ based on the analysis of color channel discrepancies and their spatial patterns.

b) Chromatic aberration identification task: Formally, let $\mathcal{I}$ denote the space of all possible digital RGB images $\mathbf{I} : \Omega \to [0,1]^3$, where $\Omega \subseteq \mathbb{N}^2$ is a discrete pixel grid, and the three channels correspond to red ($\mathbf{I}_R$), green ($\mathbf{I}_G$), and blue ($\mathbf{I}_B$). We formulate chromatic aberration identification as a binary classification problem: given an image $\mathbf{I} \in \mathcal{I}$, determine whether it exhibits noticeable chromatic aberration. Let $y(\mathbf{I}) \in \{0,1\}$ be the ground-truth label, where

- $y(\mathbf{I}) = 0$ – the image does not contain chromatic aberration (or it is below a detectable threshold),
- $y(\mathbf{I}) = 1$ – the image contains chromatic aberration (lateral and/or axial) at a level significant from the standpoint of image quality.

The objective of the model is to learn a function $f_\theta : \mathcal{I} \to [0,1]$ such that $f_\theta(\mathbf{I}) \approx P(y(\mathbf{I}) = 1 \mid \mathbf{I})$. Binary prediction is then obtained by applying a decision threshold $\tau \in (0,1)$:

$$\hat{y}(\mathbf{I}) = \begin{cases} 1 & \text{if } f_\theta(\mathbf{I}) \geq \tau, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

In this work, we address this classification task in the challenging regime of limited real-world examples of CA, employing physics-based data augmentation grounded in optical models of dispersion.

B. Proposed TriadNet architecture

The proposed TriadNet is a hybrid convolutional-transformer network designed for chromatic aberration detection. As mentioned, the architecture integrates three components. Let $\mathbf{I} \in \mathbb{R}^{H \times W \times 3}$ denote the input RGB image. The model operates as follows:

a) Edge-centric branch: Chromatic aberration is most visible as color fringing along high-contrast edges, which motivates the first component to focus specifically on these regions. An edge map $\mathbf{E} \in [0, 1]^{H \times W}$ is computed using a learnable gradient filter (based on Sobel operators with trainable weights) followed by a sigmoid activation:

$$\mathbf{E} = \sigma\left(4 \cdot \sqrt{G_x^2 + G_y^2 + \varepsilon} - 2\right), \tag{4}$$

where G_x and G_y are convolutional gradient filters, σ is the sigmoid function, and $\varepsilon > 0$ is a small constant for numerical stability. This edge map serves as a soft mask highlighting areas where misalignment between color channels is most likely to occur. Additionally, a scalar global edge strength $e_g = \mathrm{AvgPool}(\mathbf{E}) \in \mathbb{R}$ is extracted, providing a compact summary of overall edge intensity across the entire image.

b) Channel-aware attention mechanism: To capture dispersion-induced spatial shifts between color channels, local features $\mathbf{F} \in \mathbb{R}^{H' \times W' \times C}$ are first extracted using a convolutional backbone (ConvNeXt-Tiny). This backbone delivers rich hierarchical representations that form the basis for subsequent processing. A lightweight multi-head self-attention mechanism is then applied across the channel dimension:

$$\mathbf{F}' = \mathrm{MultiHeadAttention}(\mathbf{F}_{\mathrm{flat}}, \mathbf{F}_{\mathrm{flat}}, \mathbf{F}_{\mathrm{flat}}), \tag{5}$$

where $\mathbf{F}_{\mathrm{flat}}$ is the spatially flattened version of $\mathbf{F}$ (sequence length $H' \times W'$). This step enables each channel to attend to potential misalignments in the others, explicitly modeling the inter-channel discrepancies that are characteristic of chromatic aberration and difficult to capture with convolution alone.

c) Edge-modulated feature processing: The edge information is finally used to selectively enhance the attended features in regions where aberration is most pronounced. The edge map $\mathbf{E}$ is downsampled to match the feature resolution, $\mathbf{E}' = \mathrm{Downsample}(\mathbf{E}; H', W')$, and used as a modulator:

$$\mathbf{F}'' = \mathbf{F}' \odot (1 + \alpha \cdot \mathbf{E}'), \tag{6}$$

where $\alpha = 0.6$ is a fixed hyperparameter and $\odot$ denotes element-wise multiplication. This operation constitutes the edge-modulated feature processing component, ensuring that high-frequency details, where chromatic fringing is most perceptible, receive increased emphasis during subsequent classification.

d) Classification head: The processed features are aggregated into a compact global representation. Adaptive average pooling of F'' is concatenated with the global edge strength:

$$z = [\mathrm{AvgPool}(\mathbf{F}''),\ e_g] \in \mathbb{R}^{C+1}. \tag{7}$$

This combined representation is passed through a small MLP to produce a raw output score $s = \mathrm{MLP}(z) \in \mathbb{R}$. The final probability of chromatic aberration presence is then obtained by applying the sigmoid function:

$$p = \sigma(s), \tag{8}$$

where σ is the sigmoid function. This score serves as the basis for binary classification. The model is designed for end-to-end training using binary cross-entropy loss. This formulation allows TriadNet to learn the characteristic patterns of chromatic aberration directly from the data, while remaining suitable for integration into practical image processing pipelines.

III. EXPERIMENTAL EVALUATION

In this section, we present the classification results of the proposed TriadNet method in systematic comparison with three selected baselines: a convolutional network, ResNet-18 pre-trained on ImageNet, and EfficientViT-M0. The evaluation provides a comprehensive assessment of model performance and highlights the contribution of the proposed multimodal feature learning strategy, realized through edge-centric modulation and channel-aware attention mechanisms, in improving detection accuracy over generic convolutional and transformer-based approaches.

A. Experimental setup

a) Dataset: The experiments were conducted on a set of real-world photographs, manually labeled into two classes: images exhibiting noticeable chromatic aberration (CA) and images without it (no CA). It consists of approximately 800 images (at least 400 images per class), capturing a variety of scenes, lighting conditions, and aberration strengths to ensure realistic evaluation of the proposed method. The images come from the IMAGINE dataset [11].

b) Evaluation measures: Since the chromatic aberration identification task is formulated as a binary classification problem, we adopt standard binary classification metrics: precision, recall, F_1-score, and accuracy. Let TP, FP, TN, and FN denote the number of true positives, false positives, true negatives, and false negatives, respectively. The metrics are defined as follows. Accuracy represents the overall proportion of correctly classified instances:

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}. \tag{9}$$

Precision measures the proportion of correctly predicted positive instances among all predicted positives:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}. \tag{10}$$

Recall (also known as sensitivity) measures the proportion of correctly identified positive instances among all actual positives:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{11}$$

The F_1-score is the harmonic mean of precision and recall, providing a balanced measure of both:

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{12}$$

To provide a global, class-independent evaluation of the model, macro-averaged metrics [12] are computed as the arithmetic mean of the per-class (treating all classes equally) values. Specifically, the macro-averaged precision P^{macro}, recall R^{macro}, and F_1-score are defined as follows:

$$P^{\text{macro}} = \frac{P_{\text{(CA)}} + P_{\text{(no CA)}}}{2} \tag{13}$$

$$R^{\text{macro}} = \frac{R_{\text{(CA)}} + R_{\text{(no CA)}}}{2} \tag{14}$$

$$F_1^{\text{macro}} = \frac{F_{1,\text{(CA)}} + F_{1,\text{(no CA)}}}{2} \tag{15}$$

To provide a more complete view of classification performance, we additionally present Precision-Recall curves for all models. These curves illustrate the trade-off between precision and recall across different decision thresholds and enable a direct visual comparison of the proposed TriadNet against the baselines. The area under curve (AUC) depicts the efficacy of the classifier. The AUC closer to 1, the better the classification accuracy, if the AUC = 0.5, then a classifier gives the accuracy equal to a random division.

c) Meta parameters and implementation: The proposed model was trained using an early stopping mechanism with an initial maximum of 100 epochs. The training process was terminated when the loss value did not decrease for 10 consecutive epochs. The learning rate was set to 0.001, and the Adam optimizer was used.

Experiments were held on a Gigabyte Aero notebook equipped with an Intel Core i7-13700H CPU with 32 gigabytes of RAM and an Nvidia GeForce RTX 4070 GPU with 8 gigabytes of video memory. Scripts were implemented in Python under the PyTorch framework (with Nvidia CUDA support).

B. Baseline methods

We compare the proposed method against several representative baseline models, including widely adopted state-of-the-art approaches, which we recall below.

a) Classic CNN: The first baseline method is a convolutional neural network (further depicted as classic CNN) consisting of three convolutional layers with increasing channel depth ($32 \to 64 \to 128$), each followed by max-pooling and ReLU activation, equipped with global average pooling and a single fully connected layer producing the final logit. This simple architecture serves as a reference point to evaluate the necessity of deeper feature extraction and specialized mechanisms for capturing CA patterns. Despite its limited capacity, it provides a fair indication of the performance achievable with basic convolutional processing alone [13].

b) ResNet-18: ResNet-18, pre-trained on ImageNet and fine-tuned on our dataset, represents a strong and widely adopted convolutional baseline. It employs residual connections to enable effective training of deeper networks while maintaining computational efficiency. It contains 18 layers and offers robust hierarchical feature extraction, serving as a standard benchmark in many low-level vision tasks. For the CA detection problem, we have replaced the original fully connected layer (originally outputting 1000 ImageNet classes) with a single-output linear layer, and the entire model was fine-tuned end-to-end on our dataset. Its inclusion allows to evaluate whether the proposed task-specific components in TriadNet yield meaningful improvements over a generic, well-established CNN backbone [14].

c) EfficientViT-M0: EfficientViT-M0 is a lightweight vision transformer from the EfficientViT family (2023) [15], optimized for memory efficiency and inference speed while preserving strong representational power. It employs cascaded group attention and efficient patch embedding to reduce quadratic complexity typical of standard transformers and constitutes a modern transformer-based baseline that contrasts with the hybrid convolutional-transformer design of TriadNet. The M0 denotes the smallest variant of EfficientViT, consisting of 3–5 million parameters. This comparison highlights the contribution of the proposed edge-centric modulation and channel-aware attention over a pure attention-driven architecture [16].

C. Results

The quantitative comparison metrics of chromatic aberration detection are presented in Tab. I and II.

TABLE I
PER-CLASS PRECISION, RECALL, AND F_1-SCORE FOR CHROMATIC ABERRATION DETECTION METHODS

Method	Class	Precision	Recall	F_1
Proposed	no CA	0.98	0.98	0.98
	CA	0.98	0.98	0.98
Classic CNN	no CA	0.77	0.75	0.76
	CA	0.75	0.77	0.76
ResNet-18	no CA	0.90	0.88	0.89
	CA	0.88	0.90	0.89
EfficientViT-M0	no CA	0.93	0.92	0.92
	CA	0.92	0.93	0.93

The results demonstrate that the proposed method significantly outperforms the baseline approaches across all evaluation metrics for both classes. It achieves near-perfect precision, recall, and F_1-scores for CA and no CA, indicating robust and well-balanced classification performance. In contrast, the classic CNN exhibits substantially lower metrics, reflecting limited discriminative capability, while ResNet-18 and EfficientViT-M0 show improved but still inferior performance compared to the proposed approach. Overall, the results confirm the effectiveness of the proposed method in accurately detecting

Proposed TriadNet	no CA	CA
no CA	98.0	2.0
CA	2.0	98.0

Classic CNN	no CA	CA
no CA	75.0	25.0
CA	23.0	77.0

ResNet-18	no CA	CA
no CA	88.0	12.0
CA	10.0	90.0

EfficientViT-M0	no CA	CA
no CA	92.0	8.0
CA	7.0	93.0

Fig. 1. Confusion matrices for all evaluated methods. Rows correspond to ground-truth classes, columns to predicted classes, and values are expressed as percentages (%)

chromatic aberration and distinguishing it from non-aberrated images.

TABLE II
THE OVERALL ACCURACY AND MACRO-AVERAGED PRECISION, RECALL, F_1-SCORE

Method	Accuracy	P^{macro}	R^{macro}	F_1^{macro}
Proposed	0.98	0.98	0.98	0.98
Classic CNN	0.76	0.76	0.76	0.76
ResNet-18	0.89	0.89	0.89	0.89
EfficientViT-M0	0.93	0.93	0.93	0.93

The results in Table II clearly demonstrate the advantage of the TriadNet model compared to the baseline models. It achieves the highest scores across all macro-averaged metrics, with a macro F_1-score of 0.98 and overall accuracy of 0.98, indicating both precise and balanced performance across classes. Among the baselines, EfficientViT-M0 performs best, followed by ResNet-18, while the classic CNN shows the lowest performance, particularly in macro precision. Overall, these results confirm that the proposed approach is highly effective for chromatic aberration detection, providing consistent and accurate predictions for both classes. Figure 1 illustrates the confusion matrices for all evaluated methods.

The proposed method achieves highly balanced results, with 98% of samples correctly classified for both the CA and no CA classes, and only marginal misclassification rates of 2%, indicating strong discriminative capability and robustness. In contrast, the classic CNN exhibits substantial confusion between the two classes, particularly for no CA samples, where 25% are incorrectly labeled as CA, reflecting limited generalization ability. ResNet-18 significantly improves class separability, reducing misclassification rates to 10–12%, while EfficientViT-M0 further enhances performance, achieving over 92% correct classification for both classes. Overall, the confusion matrices confirm that the proposed approach achieves the highest accuracy and also obtains consistent and reliable performance across classes, which is critical for accurate chromatic aberration detection. The Fig. 2 presents the Precision-Recall curve illustrating the performance of the proposed method compared with baselines.

These results also confirm the advantage of the TriadNet method over the state-of-the-art models. The proposed model achieved almost 0.99 AUC, while the other methods obtained

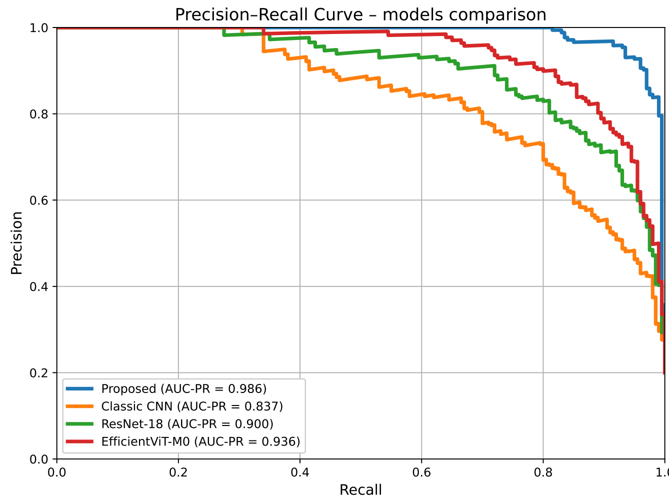

Fig. 2. The Precision-Recall curve of all considered methods

lower values, for example, the EfficientViT-M0 achieved 0.94 and ResNet-18 0.90. The lowest results were reported in the case of the classic CNN, which resulted in an AUC equal to 0.84. Overall, the chart confirms the TriadNet's reliable accuracy.

IV. PREVIOUS AND RELATED WORK

Recent works in low-level vision have explored multi-branch and hybrid convolutional–transformer architectures that integrate complementary feature pathways to improve performance in tasks such as image denoising, super-resolution, and artifact detection. These architectures can be interpreted as learning modality-specific feature representations, which inspires our multimodal formulation for chromatic aberration detection [17]. Attention mechanisms, particularly channel-aware and spatially modulated attention, have been successfully applied to guide the network to focus on task-relevant features [18]. A comprehensive review on multimodal learning is discussed in [19].

Most existing literature on CA focuses on correction rather than explicit detection [20]–[31]. Approaches include post-processing correction pipelines [24], calibration transfer across cameras [32], edge-based detection [27], statistical models for forgery detection [28], [29], segmentation-based methods [30], and wavelet- or multi-task learning-based detection frameworks [31], [33]. While these studies achieve various degrees of success, they typically either emphasize correction over detection or rely on task-specific feature engineering. Moreover, most approaches focus on single representations of the image, such as edges, color channels, or handcrafted features, without jointly utilizing complementary information. This limitation motivates the use of end-to-end multimodal feature learning strategies that integrate multiple representations to improve detection accuracy.

V. CONCLUSION

In this work, we introduced TriadNet, a hybrid convolutional–transformer architecture for chromatic aberration detection, treating the task as multimodal feature learning. The model captures structural and chromatic inconsistencies using edge-focused processing, channel-aware attention, and edge-modulated fusion. Experimental results on a custom dataset of real-world images demonstrate that TriadNet achieves improved classification performance compared to several baselines, including a classical CNN, ResNet-18, and EfficientViT-M0, with F_1-scores exceeding 0.98 and robust generalization across varying aberration strengths. Results highlight the benefits of task-specific inductive biases for low-level optical artifact detection, especially in low-data settings.

In the future, it is planned to extend the model to estimate local channel shifts, enabling both detection and automatic correction of CA. We also intend to test the approach on larger and more diverse datasets, especially images from modern smartphones and various lens types.

ACKNOWLEDGEMENTS

This research received partial financing from the funds assigned to AGH University of Krakow by Polish Ministry of Science and Higher Education.

REFERENCES

[1] J. Bernacki and R. Scherer, "Lens aberrations detection and digital camera identification with convolutional autoencoders," in *Proceedings of the 22nd International Conference on Security and Cryptography - Volume 1: SECRYPT*, INSTICC. SciTePress, 2025, pp. 84–95.

[2] J. Bernacki, "Digital camera identification based on analysis of optical defects," *Multimedia Tools and Applications*, vol. 79, no. 3, pp. 2945–2963, 2020.

[3] F. Casino, T. K. Dasaklis, G. P. Spathoulas, M. Anagnostopoulos, A. Ghosal, I. Borocz, A. Solanas, M. Conti, and C. Patsakis, "Research trends, challenges, and emerging topics in digital forensics: A review of reviews," *Ieee Access*, vol. 10, pp. 25 464–25 493, 2022.

[4] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.

[5] M. Goljan and J. Fridrich, "Estimation of lens distortion correction from single images," in *Media Watermarking, Security, and Forensics 2014*, vol. 9028. SPIE, 2014, pp. 234–246.

[6] S. Bimurzaev and Z. Sautbekova, "Transmission electron microscope with a mirror objective free of spherical and axial chromatic aberrations," *Ultramicroscopy*, p. 114168, 2025.

[7] C. Chen, X. Cai, B. Yi, Q. Chang, H. Luo, and X. Chen, "Compensation of the transverse chromatic aberration in a color-coded fringe projection profilometry," *Applied Optics*, vol. 64, no. 18, pp. 4949–4955, 2025.

[8] G. N. Manion and T. J. Stokkermans, "Chromatic aberration," in *StatPearls [Internet]*. StatPearls Publishing, 2023.

[9] T. J. Gawne and M. S. Banks, "The role of chromatic aberration in vision," *Annual Review of Vision Science*, vol. 10, 2024.

[10] J. Pereiro-García, M. Caño-García, O. Blanco-Fernández, R. Ramos-Uña, X. Quintana, and M. Geday, "Chromatic aberration compensation using thin, transparent, large aperture, wide focal range, adaptive liquid crystal lens," *Optics & Laser Technology*, vol. 180, p. 111532, 2025.

[11] J. Bernacki and R. Scherer, "Imagine dataset: Digital camera identification image benchmarking dataset." in *SECRYPT*, 2023, pp. 799–804.

[12] P. Hari and M. P. Singh, "A lightweight convolutional neural network for disease detection of fruit leaves," *Neural Computing and Applications*, vol. 35, no. 20, pp. 14 855–14 866, 2023.

[13] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 12, pp. 6999–7019, 2021.

[14] S. Ayyachamy, V. Alex, M. Khened, and G. Krishnamurthi, "Medical image retrieval using resnet-18," in *Medical imaging 2019: imaging informatics for healthcare, research, and applications*, vol. 10954. SPIE, 2019, pp. 233–241.

[15] H. Cai, J. Li, M. Hu, C. Gan, and S. Han, "Efficientvit: Lightweight multi-scale attention for high-resolution dense prediction," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 17 302–17 313.

[16] Y. Zhang and Q. Zhan, "Weldvit: A lightweight network for online identification of multi-label welding defects," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2025, pp. 447–459.

[17] Y. Qiu, K. Zhang, C. Wang, W. Luo, H. Li, and Z. Jin, "Mb-taylorformer: Multi-branch efficient transformer expanded by taylor formula for image dehazing," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 12 802–12 813.

[18] C. Wang, J. Huang, M. Lv, Y. Wu, and R. Qin, "Dual-branch adaptive convolutional transformer for hyperspectral image classification," *Remote Sensing*, vol. 16, no. 9, p. 1615, 2024.

[19] P. Xu, X. Zhu, and D. A. Clifton, "Multimodal learning with transformers: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 10, pp. 12 113–12 132, 2023.

[20] K. Schmidt, N. Guo, W. Wang, J. Czarske, and N. Koukourakis, "Chromatic aberration correction employing reinforcement learning," *Optics Express*, vol. 31, no. 10, pp. 16 133–16 147, 2023.

[21] J. Chan, X. Zhao, S. Zhong, T. Zhang, and B. Fan, "Chromatic aberration correction in harmonic diffractive lenses based on compressed sensing encoding imaging," *Sensors*, vol. 24, no. 8, p. 2471, 2024.

[22] H. Liu, J. Liu, W. Zhou, B. Xu, D. Xiong, and X. Yang, "Impact of axial chromatic aberration on color-multiplexed differential phase contrast microscopy: A quantitative study," *Optics and Lasers in Engineering*, vol. 184, p. 108660, 2025.

[23] J. Wu, D. Li, A. Cui, J. Gao, K. Zhou, and B. Liu, "Digital infrared chromatic aberration correction algorithm for a membrane diffractive lens based on coherent imaging," *Applied Optics*, vol. 61, no. 34, pp. 10 080–10 085, 2022.

[24] J. T. Korneliussen and K. Hirakawa, "Camera processing with chromatic aberration," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4539–4552, 2014.

[25] M. C. C. Mihaila, N. Laštovičková Streshkova, and M. Kozák, "Light-based chromatic aberration correction of ultrafast electron microscopes," *Physical Review Letters*, vol. 134, no. 20, p. 203802, 2025.

[26] J. Xiong, Z. Wang, Y. Shan, D. Cheng, and Y. Wang, "Eliminating chromatic aberrations of diffractive lenses in varifocal virtual reality displays," *Photonics Research*, vol. 13, no. 12, pp. 3466–3475, 2025.

[27] S.-W. Chung, B.-K. Kim, and W.-J. Song, "Detecting and eliminating chromatic aberration in digital images," in *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2009, pp. 3905–3908.

[28] O. Mayer and M. Stamm, "Improved forgery detection with lateral chromatic aberration," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 2024–2028.

[29] O. Mayer and M. C. Stamm, "Accurate and efficient image forgery detection using lateral chromatic aberration," *IEEE Transactions on information forensics and security*, vol. 13, no. 7, pp. 1762–1777, 2018.

[30] W. Kesornsukhon, P. Visutsak, and S. Ratanasanya, "Chromatic aberration detection based on image segmentation," in *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE, 2018, pp. 1–6.

[31] A. Fryskowska, M. Kedzierski, M. Wojtkowska, and A. Grochala, "A novel method of chromatic aberration detection and correction using wavelet analysis," in *2017 Baltic Geodetic Congress (BGC Geomatics)*. IEEE, 2017, pp. 18–24.

[32] J. Rayas, A. Sicardi-Segade, and A. Martínez-García, "Chromatic aberration compensation of a lens for use in any ccd camera," *Experimental Mechanics*, pp. 1–14, 2025.

[33] S. Yu, H. Yang, H. Cui, Y. Wu, and H. Cao, "A multi-task learning framework for textureaware chromatic aberration detection," in *2024 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*. IEEE, 2024, pp. 206–213.