

Algorithms for Image Processing and Machine Vision

Arslan Munir 

Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431, USA; arslanm@fau.edu

1. Introduction

The field of image processing and machine vision has undergone a remarkable transformation in recent years. From early efforts focused on pixel-level enhancement and geometric transformations [1], research has advanced toward sophisticated algorithms capable of extracting high-level semantic meaning from complex visual data [2]. These developments have been fueled by progress in deep learning architectures, representation learning, and hybrid computational approaches, which together have enabled machines to perceive, interpret, and interact with their environments in ways once limited to human vision. Today, image processing and machine vision are no longer niche domains but essential technologies with applications ranging from safety and healthcare to creative industries and disaster resilience.

This Reprint reflects the dynamism and diversity of ongoing research in image processing and machine vision. The included contributions span multiple layers of the field: foundational algorithmic studies such as adaptive compression and representation learning, advances in detection and recognition that underpin modern machine vision, and cutting-edge explorations of generative artificial intelligence (AI) and diffusion models. Equally important are the works that demonstrate societal relevance, that is, vision algorithms applied to transportation safety, inclusive classroom learning, pharmaceutical and textile quality control, and structural monitoring of civil infrastructure. Together, these works reveal the broad utility of vision algorithms in addressing both theoretical challenges and real-world problems.

Another key theme emerging across the contributions is the balance between efficiency and accuracy. Several papers emphasize methods that achieve state-of-the-art recognition performance while dramatically reducing computational cost, a crucial step toward real-time deployment in safety-critical or resource-constrained environments. Others highlight the integration of generative models with guidance or hybrid optimization techniques, pointing toward new directions where creativity, synthesis, and robustness converge. These innovations underscore how the field is evolving beyond accuracy benchmarks to encompass scalability, interpretability, and application readiness [3,4].

Collectively, the 17 papers included in this Reprint provide a coherent progression—from mathematical and algorithmic foundations to practical deployments across multiple sectors, and from human-centered applications to frontier technologies in generative and vision-based surveillance. The result is a curated collection that captures both the depth of current research and the breadth of its impact, offering valuable insights for researchers, practitioners, and educators seeking to advance the state of image processing and machine vision.

To provide readers with a structured perspective, this Reprint is organized into thematic clusters that reflect both the breadth and progression of the contributions. The first cluster highlights foundational advances in algorithms and compression techniques, establishing the mathematical and computational underpinnings of modern vision systems. The



Received: 12 September 2025

Accepted: 30 October 2025

Published: 28 November 2025

Citation: Munir, A. Algorithms for Image Processing and Machine Vision. *Algorithms* **2025**, *18*, 750. <https://doi.org/10.3390/a18120750>

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

second cluster focuses on **recognition, detection, and scene understanding**, showcasing deep learning-based approaches for action recognition, indoor scene analysis, and rotated object detection. The third cluster emphasizes **safety, healthcare, textile, and inclusive education applications**, where vision algorithms are applied to domains such as traffic safety enforcement, classroom support, and pharmaceutical and textile quality control. The fourth cluster turns to **generative and diffusion-based models**, highlighting new paradigms for image synthesis [5], rotoscoping [6], and hybrid generative adversarial network (GAN)-evolutionary frameworks [7]. The fifth cluster explores **emotion recognition** [8], **anomaly detection** [9], and **surveillance** [10], underscoring the role of vision in human interaction and public safety. Finally, the Reprint closes with contributions in **infrastructure and hazard monitoring**, demonstrating how unmanned aerial vehicle (UAV)-based vision systems can aid in structural health assessment during seismic events. This organization ensures a natural flow—from theoretical foundations to applied innovations and frontier explorations—guiding the reader through the evolving landscape of algorithms for image processing and machine vision.

2. Foundational Algorithmic and Compression Advances

The first cluster emphasizes **mathematical models and efficient transformations**, crucial for compression and representation.

- **Determining Thresholds for Optimal Adaptive Discrete Cosine Transformation** (Khanov et al. [11]) presents a novel method for optimal threshold search in adaptive Discrete Cosine Transform (DCT). By tailoring thresholds to tonal distribution, their algorithm achieves up to 66% higher compression ratios compared to traditional DCT, with applications in traffic-demanding urban video surveillance.
- **Synthetic Face Discrimination via Learned Image Compression** (Iliopoulou et al. [12]) addresses the urgent challenge of distinguishing real from synthetic (GAN/diffusion-generated) faces. Their compression-based forensic approach exploits differences in post-compression quality, achieving efficient and generalized deepfake detection.

3. Advances in Recognition, Detection, and Scene Understanding

This section bring together works focused on recognition and detection tasks that form the foundation of machine vision.

- **Human Action Representation Learning Using an Attention-Driven Residual 3DCNN Network** (Ullah and Munir [13]) introduces Dual-Attentional Residual 3D Convolutional Neural Network (DA-R3DCNN), a spatiotemporal attention model delivering 11% improved accuracy and 74× faster inference, ideal for real-time human activity recognition.
- **Indoor Scene Recognition with Transfer Learning and Liquid State Machines** (Suren-dran et al. [14]) combines DenseNet201 features with fuzzy-based selection and liquid state machine classifiers, reaching 96% accuracy on NYU datasets, particularly benefiting assistive technology applications.
- **DDL R-CNN: Dynamic Direction Learning R-CNN for Rotated Object Detection** (Su and Jing [15]) tackles the challenge of detecting arbitrarily oriented targets in remote sensing. Their dynamic direction module pre-extracts orientation features, enabling anchor box alignment and surpassing prior rotation detectors on UCAS-AOD and HRSC2016 datasets.

4. Safety, Healthcare, Textile, and Inclusive Education Applications

Several papers focus on human-centered and safety-critical systems, showcasing the societal impact of vision algorithms.

- **YOLOv7 for Real-Time Car Safety Belt Detection** (Nkuzo et al. [16]) develops a high-performance seatbelt usage detection system with a mean average precision (mAP) of 99.6% and real-time deployment potential.
- **Detecting Motorcyclists' Helmet Violations via YOLOv8 + DCGAN** (Shoman et al. [17]) enhances training data diversity using generative augmentation, improving helmet violation detection F1-scores from 0.91 to 0.96.
- **Continuous Recognition of Teachers' Hand Signals for Students with Attention Deficits** (Chen et al. [18]) applies MediaPipe BlazePose to classroom settings, achieving 88% F1-score and providing real-time non-verbal cues for inclusive education.
- **Improved U2Net-Based Surface Defect Detection for Blister Tablets** (Zhou et al. [19]) advances pharmaceutical quality assurance, combining large-kernel attention with Gaussian–Laplacian loss for defect detection at 99% accuracy within 50 ms per image.
- **Enhanced Curvature-Based Fabric Defect Detection with Gabor Transform and Deep Learning** (Erdogan and Dogan [20]) proposes a curvature algorithm, integrated with the Gabor transform, achieving performance comparable to convolutional neural networks (CNNs) while requiring minimal storage and processing, making it ideal for real-time textile quality control.

5. Generative and Diffusion-Based Models

Another central theme of this Reprint is creative synthesis and generative AI, encompassing diffusion models, Generative Adversarial Network (GAN) hybrids, and animation.

- **Denoising Diffusion Models on Model-Based Latent Space** (Scribano et al. [21]) replaces learned autoencoders with model-based lossy compression schemes, reducing computing cost while preserving generative quality.
- **GDUI: Guided Diffusion for Unlabeled Images** (Xie and Zhao [22]) integrates Contrastive Language–Image Pretraining (CLIP)-based semantic alignment with diffusion, enabling label-free clustering and semantically guided synthesis.
- **Lester: Rotoscope Animation via Segmentation and Tracking** (Tous [23]) leverages Segment Anything Model (SAM) and Decoupling Features in Hierarchical Propagation for Video Object Segmentation (DeAOT) for temporally consistent 2D animation generation, offering deterministic alternatives to diffusion pipelines in creative industries.
- **GAGAN: Hybrid Genetic Algorithm-Optimized DCGANs** (Konstantopoulou et al. [24]) demonstrates how evolutionary strategies can stabilize GAN training and improve image quality, showing the promise of hybrid evolutionary–deep learning models for generative synthesis.

6. Emotion, Anomaly, and Surveillance

This cluster highlights methods tailored for facial, emotional, and surveillance contexts.

- **Histogram Equalization with VGG Models for Facial Emotion Recognition** (Chowdhury et al. [25]) achieves perfect classification on CK+ and near-perfect results on KDEF datasets, showing that simple preprocessing dramatically boosts deep model performance.
- **Video Anomaly Detection with Vision Transformers and Spatiotemporal Attention** (Habeab et al. [26]) develops a hybrid deep learning model that combines a vision transformer (ViT) with a convolutional spatiotemporal relationship (STR) attention block, achieving 95.6% area under the receiver operating characteristic curve (AUC ROC) on UCSD-Ped2 dataset and scaling effectively to very large datasets, such as ShanghaiTech and CHAD.

7. Vision for Infrastructure, Hazards, and the Built Environment

The final cluster demonstrates the impact of vision algorithms in civil infrastructure and hazard resilience.

- **Application Framework for UAV-Based Earthquake-Induced Structural Displacement Monitoring** (Ji et al. [27]) introduces an unmanned aerial vehicle (UAV) vision-based system validated on a six-story mass timber building under shake-table tests, achieving millimeter-level displacement accuracy compared with ground-truth sensors.

8. Conclusions

The contributions in this Reprint collectively chart a rich landscape for algorithms in image processing and machine vision. Several overarching insights emerge:

- **Algorithmic efficiency** is being redefined by hybrid architectures, lightweight preprocessing, and compression-aware methods.
- **Societal impact** is evident in transportation safety, inclusive education, pharmaceutical and textile quality control, and disaster resilience.
- **Generative frontiers** are rapidly advancing, with GAN-evolutionary hybrids and diffusion-guided methods opening creative and practical applications.

As Guest Editor, I thank all authors and reviewers for their rigorous contributions and the MDPI team for their support. This Reprint of the Special Issue thus serves not only as a curated collection of research but also as a roadmap for future inquiry, spanning foundational algorithms, human-centered vision, and generative intelligence.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Golan, I.; El-Yaniv, R. Deep anomaly detection using geometric transformations. In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18), Montréal, QC, Canada, 3–8 December 2018; Curran Associates Inc.: Red Hook, NY, USA, 2018; pp. 9781–9791.
2. Velastegui, R.; Tatarchenko, M.; Karaoglu, S.; Gevers, T. Image semantic segmentation of indoor scenes: A survey. *Comput. Vis. Image Underst.* **2024**, *248*, 104102. [[CrossRef](#)]
3. del Olmo, J.J.L.; López-de-Teruel, P.E.; Ruiz, A.; García-Clemente, F.J. Computer Vision on the Edge: A Scalable Auto-ID Solution for Industrial Logistics. *Procedia Comput. Sci.* **2025**, *265*, 276–284. [[CrossRef](#)]
4. Cheng, Z.; Wu, Y.; Li, Y.; Cai, L.; Ihnaini, B. A Comprehensive Review of Explainable Artificial Intelligence (XAI) in Computer Vision. *Sensors* **2025**, *25*, 4166. [[CrossRef](#)]
5. Wang, T.-C.; Liu, M.-Y.; Zhu, J.-Y.; Tao, A.; Kautz, J.; Catanzaro, B. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8798–8807. [[CrossRef](#)]
6. Bermudez, L.; Dabby, N.; Lin, Y.A.; Hilmarsdottir, S.; Sundararajan, N.; Kar, S. A Learning-Based Approach to Parametric Rotoscoping of Multi-Shape Systems. In Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 3–8 January 2021; pp. 776–785. [[CrossRef](#)]
7. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS 2014): Annual Conference on Neural Information Processing Systems 2014, Montréal, QC, Canada, 8–13 December 2014; Volume 27.
8. Khare, S.K.; Blanes-Vidal, V.; Nadimi, E.S.; Acharya, U.R. Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Inf. Fusion* **2024**, *102*, 102019. [[CrossRef](#)]
9. Nassif, A.B.; Talib, M.A.; Nasir, Q.; Dakalbab, F.M. Machine Learning for Anomaly Detection: A Systematic Review. *IEEE Access* **2021**, *9*, 78658–78700. [[CrossRef](#)]
10. Munir, A.; Kwon, J.; Lee, J.H.; Kong, J.; Blasch, E.; Aved, A.; Muhammad, K. FogSurv: A Fog-Assisted Architecture for Urban Surveillance Using Artificial Intelligence and Data Fusion. *IEEE Access* **2021**, *9*, 111938–111959. [[CrossRef](#)]
11. Khanov, A.; Shulzhenko, A.; Voroshilova, A.; Zubarev, A.; Karimov, T.; Fahmi, S. Determining Thresholds for Optimal Adaptive Discrete Cosine Transformation. *Algorithms* **2024**, *17*, 366. [[CrossRef](#)]

12. Iliopoulou, S.; Tsinganos, P.; Ampeliotis, D.; Skodras, A. Synthetic Face Discrimination via Learned Image Compression. *Algorithms* **2024**, *17*, 375. [\[CrossRef\]](#)
13. Ullah, H.; Munir, A. Human Action Representation Learning Using an Attention-Driven Residual 3DCNN Network. *Algorithms* **2023**, *16*, 369. [\[CrossRef\]](#)
14. Surendran, R.; Chihi, I.; Anitha, J.; Hemanth, D.J. Indoor Scene Recognition: An Attention-Based Approach Using Feature Selection-Based Transfer Learning and Deep Liquid State Machine. *Algorithms* **2023**, *16*, 430. [\[CrossRef\]](#)
15. Su, W.; Jing, D. DDL R-CNN: Dynamic Direction Learning R-CNN for Rotated Object Detection. *Algorithms* **2025**, *18*, 21. [\[CrossRef\]](#)
16. Nkuzo, L.; Sibiya, M.; Markus, E.D. A Comprehensive Analysis of Real-Time Car Safety Belt Detection Using the YOLOv7 Algorithm. *Algorithms* **2023**, *16*, 400. [\[CrossRef\]](#)
17. Shoman, M.; Ghouli, T.; Lanzaro, G.; Alsharif, T.; Gargoum, S.; Sayed, T. Enforcing Traffic Safety: A Deep Learning Approach for Detecting Motorcyclists' Helmet Violations Using YOLOv8 and Deep Convolutional Generative Adversarial Network-Generated Images. *Algorithms* **2024**, *17*, 202. [\[CrossRef\]](#)
18. Chen, I.D.S.; Yang, C.-M.; Wu, S.-S.; Yang, C.-K.; Chen, M.-J.; Yeh, C.-H.; Lin, Y.-H. Continuous Recognition of Teachers' Hand Signals for Students with Attention Deficits. *Algorithms* **2024**, *17*, 300. [\[CrossRef\]](#)
19. Zhou, J.; Huang, J.; Liu, J.; Liu, J. Improved U2Net-Based Surface Defect Detection Method for Blister Tablets. *Algorithms* **2024**, *17*, 429. [\[CrossRef\]](#)
20. Erdogan, M.; Dogan, M. Enhanced Curvature-Based Fabric Defect Detection: A Experimental Study with Gabor Transform and Deep Learning. *Algorithms* **2024**, *17*, 506. [\[CrossRef\]](#)
21. Scribano, C.; Pezzi, D.; Franchini, G.; Prato, M. Denoising Diffusion Models on Model-Based Latent Space. *Algorithms* **2023**, *16*, 501. [\[CrossRef\]](#)
22. Xie, X.; Zhao, J. GDUI: Guided Diffusion Model for Unlabeled Images. *Algorithms* **2024**, *17*, 125. [\[CrossRef\]](#)
23. Tous, R. Lester: Rotoscope Animation through Video Object Segmentation and Tracking. *Algorithms* **2024**, *17*, 330. [\[CrossRef\]](#)
24. Konstantopoulou, D.; Zacharia, P.; Papoutsidakis, M.; Leligou, H.C.; Patrikakis, C. GAGAN: Enhancing Image Generation Through Hybrid Optimization of Genetic Algorithms and Deep Convolutional Generative Adversarial Networks. *Algorithms* **2024**, *17*, 584. [\[CrossRef\]](#)
25. Chowdhury, J.H.; Liu, Q.; Ramanna, S. Simple Histogram Equalization Technique Improves Performance of VGG Models on Facial Emotion Recognition Datasets. *Algorithms* **2024**, *17*, 238. [\[CrossRef\]](#)
26. Habeb, M.H.; Salama, M.; Elrefaei, L.A. Enhancing Video Anomaly Detection Using a Transformer Spatiotemporal Attention Unsupervised Framework for Large Datasets. *Algorithms* **2024**, *17*, 286. [\[CrossRef\]](#)
27. Ji, R.; Sorosh, S.; Lo, E.; Norton, T.J.; Driscoll, J.W.; Kuester, F.; Barbosa, A.R.; Simpson, B.G.; Hutchinson, T.C. Application Framework and Optimal Features for UAV-Based Earthquake-Induced Structural Displacement Monitoring. *Algorithms* **2025**, *18*, 66. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.