

Article

Not peer-reviewed version

Research on Image Processing and Computer Vision Algorithms and Applications

[Michael Reynolds](#)*

Posted Date: 12 May 2025

doi: 10.20944/preprints202505.0863.v1

Keywords: Image Processing; Computer Vision; Deep Learning; Object Detection; Image Segmentation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

Research on Image Processing and Computer Vision Algorithms and Applications

Michael J. Reynolds

University of Michigan, USA; mjreynolds0203@gmail.com

Abstract: Image processing and computer vision are rapidly advancing technological fields that have been widely applied in medical imaging, autonomous driving, intelligent manufacturing, and other industries. With the rise of deep learning, traditional image processing methods have been gradually replaced by deep learning algorithms, achieving remarkable results in tasks such as object detection, image classification, and segmentation. This paper aims to explore the core algorithms and applications of image processing and computer vision, reviewing classical techniques and algorithms while analyzing advanced methods based on deep learning. By discussing applications across various fields, this paper not only demonstrates the current state of the technology but also highlights its challenges and developmental directions. Finally, it forecasts future research trends in image processing and computer vision, particularly the potential developments under the influence of artificial intelligence and big data.

Keywords: image processing; computer vision; deep learning; object detection; image segmentation

1. Introduction

As important branches of computer science and artificial intelligence, image processing and computer vision have developed rapidly in recent years, finding applications across various industries. From early methods of image enhancement and filtering to today's deep learning-based automated visual recognition, the continuous innovation and progress in these technologies have significantly advanced the development of an intelligent society [1]. Image processing focuses on tasks such as image acquisition, enhancement, analysis, and representation, encompassing a range of methods from noise removal and edge detection to image reconstruction. Computer vision, on the other hand, involves enabling computers to understand and analyze visual information, mimicking the human visual system to perform tasks such as object recognition, motion analysis, and 3D reconstruction [2]. This process requires not only the handling and analysis of large-scale image data but also the development of robust algorithms to address the complexities and uncertainties of real-world scenarios [3]. In recent years, deep learning, particularly convolutional neural networks (CNNs), has brought breakthroughs to the field of image processing and computer vision, significantly improving the accuracy and efficiency of visual tasks [4]. Technologies such as image classification, object detection, and image segmentation, powered by deep learning, have been widely applied in fields like autonomous driving, intelligent security, and medical image analysis, solving problems that were challenging for traditional methods. However, challenges remain, including data annotation difficulties, high computational complexity of algorithms, and stringent real-time requirements. Designing more efficient and accurate algorithms while enhancing their robustness and stability in practical applications continues to be a key focus of research. This paper aims to discuss the latest research findings in image processing and computer vision, particularly under the context of deep learning, by analyzing core algorithms and application scenarios. Through a review of classical image processing methods and essential computer vision technologies, this paper provides readers with a systematic perspective on the advancements, challenges, and future trends in this field. Furthermore, it highlights the development and prospects of image processing and computer vision in various real-world applications, offering insights and references for researchers

and practitioners in the field. Wenqing et al. [5] combined Mamba and Transformer for weather forecast. Tangtang et al. [6] analyzed ARIMA and LSTM for US electricity price forecast. Yimeng et al. [7] classified breast cancer gene data via PCA and XGBoost. Min et al. [8] predicted loan repayment with a DeepFM model. Haosen et al. [9] proposed RPF-ELD for breast cancer recognition in ultrasound images.

2. Fundamentals of Image Processing

2.1. Mathematical Representation and Processing Methods of Images

The primary task of image processing is to transform images from the physical world into digital forms and analyze them through mathematical methods. A digital image is typically represented as a matrix of pixels, where each pixel contains brightness information (for grayscale images) or color information (for color images). In grayscale images, each pixel value is an integer between 0 and 255, representing different levels of brightness. For color images, each pixel generally contains three components (e.g., red, green, and blue, or RGB), each within the same range of 0 to 255. Image processing encompasses various methods, including image enhancement, filtering, and denoising [10]. Image enhancement aims to improve image quality or emphasize certain features for better analysis or observation. By adjusting parameters such as contrast, brightness, or color, the visual effect of an image can be enhanced. For instance, histogram equalization is a commonly used technique that adjusts the pixel distribution of an image, enhancing its contrast and improving the visibility of details [11]. Image filtering involves modifying pixel values through convolution operations to achieve effects such as smoothing or sharpening. Filtering typically relies on convolution kernels, which are small matrices applied to each pixel and its neighboring pixels. During filtering, noise can be removed or details can be enhanced by weighted averaging. For example, mean filtering is used to eliminate fine noise, while high-pass filtering enhances edges and finer details of the image. Denoising is another essential task in image processing, especially since noise introduced during image acquisition can affect quality and subsequent processing [12]. The goal of denoising is to remove unwanted noise while preserving the original structure and details of the image. Common denoising methods include median filtering, Gaussian filtering, and wavelet-based approaches. Median filtering is particularly effective for removing salt-and-pepper noise, as it replaces the current pixel value with the median value of its neighborhood, reducing the impact of extreme noise. Overall, the mathematical representation and processing methods of images form the foundation of image processing technology, providing effective tools for handling and analyzing image data. In practical applications, image processing techniques not only improve image quality but also support subsequent computer vision tasks, such as object detection and image recognition, by providing clearer data.

2.2. Feature Extraction and Representation of Images

Feature extraction and representation are critical tasks in image processing and computer vision, aiming to derive representative information from images for further analysis and recognition. By extracting features from images, the dimensionality of data can be effectively reduced, and the most distinguishable aspects of images can be highlighted [13]. Different types of feature extraction methods are suited to various applications, such as object recognition, image retrieval, and scene understanding. Image features can be classified into low-level and high-level features. Low-level features are directly extracted from pixel data, such as color, texture, and edges, which reflect basic image information. High-level features, on the other hand, are derived by combining low-level features or leveraging models, such as the shape and semantic information of objects. These features provide richer contextual information, making them particularly important for complex image analysis tasks. Common image features include color, texture, and shape. Color features are often represented by histograms, capturing the distribution of colors in an image, and are widely used in image retrieval and classification tasks [14]. Texture features reflect the spatial arrangement and

relationships of pixels in an image, with common extraction methods including Gray Level Co-occurrence Matrix (GLCM) and Local Binary Pattern (LBP). Shape features describe the geometric structure of objects in images, with popular methods including contour analysis, Hu moments, and Zernike moments. In practical applications, feature extraction often combines local and global features. Local feature extraction methods focus on key regions of an image, offering better invariance and robustness. For example, Scale-Invariant Feature Transform (SIFT) and Speeded-Up Robust Features (SURF) are commonly used local feature extraction methods, detecting key points in an image and extracting descriptors from their surroundings to enable image matching and object recognition. These local features remain highly accurate under conditions such as rotation, scale changes, and partial occlusion. In contrast, global features analyze the entire image statistically, capturing its overall structure and content [15]. With the development of deep learning, traditional handcrafted feature extraction methods have been gradually replaced by automated feature extraction based on Convolutional Neural Networks (CNNs). CNNs automatically learn hierarchical feature representations from raw images, ranging from low-level edge and texture features to high-level semantic features. By leveraging multi-layer convolutional operations, CNNs capture various levels of information in images, significantly improving the effectiveness and accuracy of feature extraction. This approach has excelled in tasks such as image classification, object detection, and semantic segmentation, becoming the mainstream technique in computer vision today [16]. In conclusion, methods for extracting and representing image features are crucial for analyzing and understanding images. With continuous technological advancements, especially the rise of deep learning, the precision and efficiency of feature extraction have been greatly enhanced, driving the widespread adoption of computer vision technologies in practical applications.

3. Core Algorithms in Computer Vision

3.1. Image Segmentation and Edge Detection

Image segmentation divides an image into multiple meaningful regions, often based on pixel color, intensity, or texture information. Edge detection is a critical step in image segmentation, identifying object boundaries by detecting changes in pixel intensity [17]. Classical edge detection algorithms include the Sobel operator and Canny edge detection. Canny edge detection employs a multi-stage process to extract edges. It starts by applying Gaussian filtering to smooth the image, calculates gradients and their directions, and finally uses a dual-threshold technique to identify edges. The Canny edge detection formula is as shown in Formula 1:

$$\text{Edge}(x,y) = \begin{cases} 1 & \text{if } |G_{(x,y)}| > T_1 \text{ and } |G_{(x,y)}| < T_2 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $G_{(x,y)}$ represents the gradient at position (x,y) , and T_1 and T_2 are the low and high thresholds, respectively. Pixels with gradient values above T_2 are identified as edges, those below T_1 are non-edges, and those in between are classified based on connectivity rules.

3.2. Object Detection and Recognition

Object detection is another core task in computer vision, aiming to identify and localize specific objects in images or videos. Traditional methods, such as Haar feature classifiers and Histograms of Oriented Gradients (HOG) combined with Support Vector Machines (SVM), were widely used. However, deep learning, particularly Convolutional Neural Networks (CNNs), has become the dominant approach in object detection [18]. Notable frameworks include R-CNN (Region-based CNN), YOLO (You Only Look Once), and SSD (Single Shot MultiBox Detector). YOLO transforms the object detection problem into a regression problem, allowing a single neural network model to simultaneously predict multiple bounding boxes and class probabilities. The key idea of YOLO is to divide the image into grids, with each grid predicting bounding boxes and corresponding class probabilities, enabling efficient real-time detection. In YOLO, the image is divided into $S \times SS$ times

S grids, and each grid predicts a fixed number of bounding boxes along with their class probabilities and positional coordinates. The output includes predictions for each grid cell's bounding boxes and classification scores.

3.3. Feature Matching and Tracking

Feature matching involves comparing feature points from two images or video frames to establish correspondences. Feature points are often corners, edges, or other prominent regions within an image. Feature matching is widely applied in tasks such as image stitching, 3D reconstruction, and visual SLAM. SIFT (Scale-Invariant Feature Transform) and SURF (Speeded-Up Robust Features) are commonly used feature matching algorithms [19]. These methods extract local invariant features, enabling robust matching under image rotation, scaling, and partial occlusion. SIFT, in particular, provides features that are invariant to scale and rotation, making it effective for handling various image transformations. In dynamic scenarios, object tracking algorithms are used to follow the position of objects across consecutive video frames [20]. Common tracking algorithms include Kalman filters, the Mean-Shift algorithm, and deep learning-based tracking methods. Kalman filters use mathematical modeling to predict object positions based on the relationship between system states and observations, making them suitable for handling noise and uncertainty. The success of computer vision algorithms relies on the effective extraction, representation, and matching of image information [21]. With continuous advancements in deep learning, computer vision algorithms have demonstrated outstanding performance in real-world applications, particularly in object detection, image segmentation, and object tracking, achieving near-human visual system performance.

4. Applications of Image Processing and Computer Vision

Image processing and computer vision technologies have found extensive applications across diverse fields, enhancing efficiency in traditional industries and enabling advancements in emerging areas. In medical imaging analysis, image processing techniques optimize image quality through denoising and contrast enhancement, while computer vision accurately identifies pathological regions, assisting doctors in diagnosis [22]. For instance, CNN-based algorithms have shown exceptional performance in detecting lung nodules and breast cancer, enabling precise segmentation of lesion boundaries and supporting personalized treatment plans. Additionally, Generative Adversarial Networks (GANs) are employed for data augmentation in medical imaging, addressing data scarcity and improving model robustness and generalization. These combined technologies drive precision medicine, making early disease detection and diagnosis more efficient and accurate [23]. In autonomous driving and intelligent transportation, image processing and computer vision serve as core technologies for environmental perception. Autonomous vehicles use cameras and sensors to capture real-time image data, which computer vision processes to identify and classify targets such as pedestrians, vehicles, lane lines, and traffic signs [24]. Deep learning-based object detection algorithms like YOLO and SSD provide real-time localization of targets, ensuring safety in vehicle decision-making. Image segmentation techniques precisely identify lane lines and road regions, ensuring accurate navigation in complex road conditions [25]. Furthermore, multi-sensor fusion integrates visual data with radar and ultrasonic information, constructing a more stable driving system. The integration of these technologies ensures reliable development in intelligent transportation, accelerating the adoption of autonomous driving technology. Industrial automation and quality control represent another vital application domain. Image processing techniques are widely used for defect detection and dimensional measurement in industrial production. High-resolution imaging and edge detection algorithms quickly identify surface cracks or flaws in products [26]. Deep learning-based classification algorithms further enhance automation in defect detection, enabling efficient categorization and handling of defective products. Moreover, in industrial robotics, computer vision guides robots by recognizing object positions and shapes, facilitating precise assembly tasks. Coupled with the Industrial Internet of Things (IIoT), vision technology promotes flexible manufacturing and adaptive production lines, boosting the intelligence of industrial

operations. In security and surveillance, image processing and computer vision are extensively applied, particularly in intelligent monitoring and facial recognition [27]. Facial recognition technology detects, aligns, and extracts features from faces, enabling efficient identity verification in access control, airport security, and online payments. Deep learning-based algorithms like FaceNet and ArcFace maintain high accuracy even in complex environments, providing robust support for intelligent security systems. Additionally, computer vision is used for behavior analysis, enabling real-time detection of abnormal activities such as crowd gatherings, falls, or unattended objects through surveillance videos. These technologies significantly improve monitoring efficiency and response speed, ensuring public safety. Overall, image processing and computer vision technologies have demonstrated tremendous potential across various fields, from medical imaging analysis to autonomous driving, industrial production, and public security. As deep learning and multimodal fusion technologies advance, these innovations will extend their reach to broader scenarios, contributing to building a smarter society. In the future, the applications of image processing and computer vision will further integrate into all aspects of human life, delivering more intelligent solutions and services to society [28].

5. Research Challenges and Development Trends

Despite significant advancements in image processing and computer vision technologies in recent years, their development still faces numerous challenges, particularly in terms of robustness, efficiency, and ethical considerations in real-world applications [29]. First, data quality and annotation costs remain major bottlenecks in current research. High-quality data is critical for training models, but acquiring large-scale, high-quality, and well-annotated image datasets is often expensive, especially in fields such as healthcare and industry. Even with sufficient annotated data, the diversity of data sources may lead to significant performance variations across different scenarios, making it difficult to ensure model generalization [30]. Moreover, privacy concerns are increasingly important in many application scenarios, especially when dealing with sensitive information, such as medical images or facial recognition data [31]. Striking a balance between data privacy and model performance remains a pressing issue. Another prominent challenge is the trade-off between computational complexity and real-time requirements. Many computer vision algorithms, especially deep learning-based models, demand substantial computational resources for training and inference [32]. This high computational cost limits their deployment in edge devices or resource-constrained environments, such as autonomous vehicles or smart home devices. Optimizing algorithms or leveraging hardware acceleration to enhance model efficiency while maintaining accuracy is a key research direction. Furthermore, robustness in complex scenarios poses an additional challenge. In practical applications, images may be affected by lighting changes, occlusions, noise, and varying perspectives. Existing models often struggle to perform consistently under such conditions [33]. Enhancing model robustness, particularly in handling performance degradation under non-ideal conditions, is a critical focus for both academia and industry. Alongside these challenges, the field of image processing and computer vision also exhibits several exciting development trends [34]. Multimodal fusion is one of the current research hotspots, combining visual information with other sensory data (e.g., speech, text, or sensor readings) to provide richer semantic information and decision-making support. For instance, in the medical domain, integrating imaging data with electronic health records can offer more comprehensive diagnostic insights. In autonomous driving, fusing data from cameras, radar, and LiDAR significantly enhances system perception and safety. As multimodal learning methods continue to mature, this approach is expected to further drive the application of visual technologies in complex scenarios [35]. The development of lightweight models and edge computing also opens new possibilities for the widespread deployment of image processing and computer vision technologies. Lightweight models, achieved through techniques such as network pruning, quantization, and knowledge distillation, greatly reduce computational complexity and storage requirements, enabling deployment on resource-constrained devices [36]. Edge computing shifts part of the computation from the cloud to local devices, reducing data

transmission latency while enhancing privacy protection. The combination of these technologies brings breakthroughs to high real-time demand scenarios, such as drone vision and mobile image processing. Finally, with the ongoing research into cutting-edge technologies like Generative Adversarial Networks (GANs), self-supervised learning, and reinforcement learning, the future of image processing and computer vision is full of potential [37–39]. For example, GANs have significantly improved model performance in data-scarce scenarios through applications in data generation and image enhancement. Self-supervised learning reduces dependence on large-scale annotated data, offering new approaches for building general-purpose vision models. Reinforcement learning demonstrates great potential in dynamic visual decision-making scenarios, particularly in robotics navigation and autonomous driving. In summary, while image processing and computer vision face challenges related to data, computational complexity, and robustness, their development trends highlight immense potential [40]. Through multimodal fusion, lightweight models with edge computing, and the integration of advanced learning techniques, this field will continue to drive technological innovation, expanding its depth and breadth in practical applications and providing robust support for the advancement of intelligent society.

6. Conclusion

As a critical branch of artificial intelligence, image processing and computer vision technologies have achieved remarkable results across various fields, from medical imaging analysis to autonomous driving, industrial production, and public security, demonstrating immense value. However, the field still faces challenges such as data quality, computational complexity, and model robustness, requiring continuous optimization and innovation. In the future, advancements in multimodal fusion, lightweight models, edge computing, and self-supervised learning will further deepen the applications of image processing and computer vision, providing strong support for societal intelligence and technological progress.

References

1. Tan C, Li X, Wang X, et al. Real-time Video Target Tracking Algorithm Utilizing Convolutional Neural Networks (CNN)[C]//2024 4th International Conference on Electronic Information Engineering and Computer (EIECT). IEEE, 2024: 847-851.
2. Khan, Asharul Islam, and Salim Al-Habsi. "Machine learning in computer vision." *Procedia Computer Science* 167 (2020): 1444-1451.
3. Zhang J, Xiang A, Cheng Y, et al. Research on Detection of Floating Objects in River and Lake Based on AI Image Recognition[J]. *Journal of Artificial Intelligence Practice*, 2024, 7(2): 97-106.
4. Wu Z. Deep learning with improved metaheuristic optimization for traffic flow prediction[J]. *Journal of Computer Science and Technology Studies*, 2024, 6(4): 47-53.
5. Zhang W, Huang J, Wang R, et al. Integration of Mamba and Transformer--MAT for Long-Short Range Time Series Forecasting with Application to Weather Dynamics[J]. *arXiv preprint arXiv:2409.08530*, 2024.
6. Wang T, Cai X, Xu Q. Energy Market Price Forecasting and Financial Technology Risk Management Based on Generative AI[J]. *Applied and Computational Engineering*, 2024, 100: 29-34.
7. Wu, X., Sun, Y., & Liu, X. (2024). Multi-Class Classification of Breast Cancer Gene Expression Using PCA and XGBoost. Preprints. <https://doi.org/10.20944/preprints202410.1775.v1>
8. Min, Liu, et al. "Financial Prediction Using DeepFM: Loan Repayment with Attention and Hybrid Loss." 2024 5th International Conference on Machine Learning and Computer Application (ICMLCA). IEEE, 2024.
9. Wang, G. Zhang, Y. Zhao, F. Lai, W. Cui, J. Xue, Q. Wang, H. Zhang, and Y. Lin, "Rpf-eld: Regional prior fusion using early and late distillation for breast cancer recognition in ultrasound images," in 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2024, pp. 2605–2612.

10. Zhao Y, Hu B, Wang S. Prediction of brent crude oil price based on lstm model under the background of low-carbon transition[J]. arXiv preprint arXiv:2409.12376, 2024.
11. Yu Q, Wang S, Tao Y. Enhancing anti-money laundering detection with self-attention graph neural networks[C]//SHS Web of Conferences. EDP Sciences, 2025, 213: 01016.
12. Mo K, Chu L, Zhang X, et al. Dral: Deep reinforcement adaptive learning for multi-uavs navigation in unknown indoor environment[J]. arXiv preprint arXiv:2409.03930, 2024.
13. Ma D, Wang M, Xiang A, et al. Transformer-Based Classification Outcome Prediction for Multimodal Stroke Treatment[J]. arXiv preprint arXiv:2404.12634, 2024.
14. Li X, Cao H, Zhang Z, et al. Artistic Neural Style Transfer Algorithms with Activation Smoothing[J]. arXiv preprint arXiv:2411.08014, 2024.
15. Guo H, Zhang Y, Chen L, et al. Research on vehicle detection based on improved YOLOv8 network[J]. arXiv preprint arXiv:2501.00300, 2024.
16. Diao, Su, et al. "Ventilator pressure prediction using recurrent neural network." arXiv preprint arXiv:2410.06552 (2024).
17. Cheng Y, Yang Q, Wang L, et al. Research on Credit Risk Early Warning Model of Commercial Banks Based on Neural Network Algorithm[J]. arXiv preprint arXiv:2405.10762, 2024.
18. Xiang A, Qi Z, Wang H, et al. A Multimodal Fusion Network For Student Emotion Recognition Based on Transformer and Tensor Product[J]. arXiv preprint arXiv:2403.08511, 2024.
19. Tang, Xirui, et al. "Research on heterogeneous computation resource allocation based on data-driven method." 2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS). IEEE, 2024.
20. Tan C, Zhang W, Qi Z, et al. Generating Multimodal Images with GAN: Integrating Text, Image, and Style[J]. arXiv preprint arXiv:2501.02167, 2025.
21. Yan, Hao, et al. "Research on image generation optimization based deep learning." Proceedings of the International Conference on Machine Learning, Pattern Recognition and Automation Engineering, 2024.
22. Yang H, Wang L, Zhang J, et al. Research on Edge Detection of LiDAR Images Based on Artificial Intelligence Technology[J]. arXiv preprint arXiv:2406.09773, 2024.
23. Xiang A, Zhang J, Yang Q, et al. Research on splicing image detection algorithms based on natural image statistical characteristics[J]. arXiv preprint arXiv:2404.16296, 2024.
24. Paneru, Suman, and Idris Jeelani. "Computer vision applications in construction: Current state, opportunities & challenges." Automation in Construction 132 (2021): 103940.
25. Chouhan, Siddharth Singh, Uday Pratap Singh, and Sanjeev Jain. "Applications of computer vision in plant pathology: a survey." Archives of computational methods in engineering 27.2 (2020): 611-632.
26. Xiang A, Huang B, Guo X, et al. A neural matrix decomposition recommender system model based on the multimodal large language model[J]. arXiv preprint arXiv:2407.08942, 2024.
27. Shih K, Han Y, Tan L. Recommendation System in Advertising and Streaming Media: Unsupervised Data Enhancement Sequence Suggestions[J]. arXiv preprint arXiv:2504.08740, 2025.
28. Wu Z, Wang X, Huang S, et al. Research on prediction recommendation system based on improved markov model[J]. Advances in Computer, Signals and Systems, 2024, 8(5): 87-97.
29. Shi X, Tao Y, Lin S C. Deep Neural Network-Based Prediction of B-Cell Epitopes for SARS-CoV and SARS-CoV-2: Enhancing Vaccine Design through Machine Learning[J]. arXiv preprint arXiv:2412.00109, 2024.
30. Zhao R, Hao Y, Li X. Business Analysis: User Attitude Evaluation and Prediction Based on Hotel User Reviews and Text Mining[J]. arXiv preprint arXiv:2412.16744, 2024.

31. Ziang H, Zhang J, Li L. Framework for lung CT image segmentation based on UNet++[J]. arXiv preprint arXiv:2501.02428, 2025.
32. Gao, Dawei, et al. "Synaptic resistor circuits based on Al oxide and Ti silicide for concurrent learning and signal processing in artificial intelligence systems." *Advanced Materials* 35.15 (2023): 2210484.
33. Wu Z. Mpgaam: Effective and efficient heterogeneous information network classification[J]. *Journal of Computer Science and Technology Studies*, 2024, 6(4): 8-16.
34. Wang L, Cheng Y, Xiang A, et al. Application of Natural Language Processing in Financial Risk Detection[J]. arXiv preprint arXiv:2406.09765, 2024.
35. Rakhimov, Bakhtiyar Saidovich, et al. "Review And Analysis Of Computer Vision Algorithms." *The American Journal of Applied sciences* 3.5 (2021): 245-250.
36. Fernandes, Arthur Francisco Araújo, João Ricardo Rebouças Dórea, and Guilherme Jordão de Magalhães Rosa. "Image analysis and computer vision applications in animal sciences: an overview." *Frontiers in Veterinary Science* 7 (2020): 551269.
37. Oliveira, Dario Augusto Borges, et al. "A review of deep learning algorithms for computer vision systems in livestock." *Livestock Science* 253 (2021): 104700.
38. Afif, Mouna, Yahia Said, and Mohamed Atri. "Computer vision algorithms acceleration using graphic processors NVIDIA CUDA." *Cluster Computing* 23.4 (2020): 3335-3347.
39. Desai, Brishaman, et al. "Image filtering-techniques algorithms and applications." *Applied GIS* 7.11 (2020): 970-975.
40. Huang B, Lu Q, Huang S, et al. Multi-modal clothing recommendation model based on large model and VAE enhancement[J]. arXiv preprint arXiv:2410.02219, 2024.

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.