Advanced Techniques in Computer Vision: A Survey on GAM, SPPF, YOLOv11, Super Resolution, Neural Networks, and Image Enhancement

www.surveyx.cn

Abstract

The survey paper presents a comprehensive evaluation of advanced techniques in computer vision, emphasizing the transformative impact of methodologies such as Gated Attention Modules (GAM), Spatial Pyramid Pooling Fast (SPPF), YOLOv11, and Super Resolution. These innovations have significantly enhanced image analysis and processing capabilities, enabling sophisticated applications across diverse domains. The paper highlights the role of neural networks in advancing image super-resolution, particularly through novel architectures like the Matrixed Channel Attention Network (MCAN) and innovative methodologies that leverage multi-frequency component modeling. The integration of attention mechanisms and hybrid architectures has further improved object detection and recognition, exemplified by the enhanced performance of YOLOv11 in real-time applications. Additionally, the survey underscores the effectiveness of advanced algorithms in image enhancement, addressing challenges in edge and noise reduction and improving image clarity in challenging environments. The conclusion emphasizes the importance of continued exploration and innovation in computer vision technologies, with significant implications for enhancing artificial intelligence and machine learning applications. Future research directions include optimizing models for real-time applications, expanding datasets for validation, and exploring novel methodologies for joint image compression and super-resolution. These advancements collectively highlight the potential for future developments in this rapidly evolving field, paving the way for more efficient and effective image processing solutions.

1 Introduction

1.1 Significance of Advanced Techniques in Computer Vision

The introduction of advanced techniques such as Gated Attention Modules (GAM), Spatial Pyramid Pooling Fast (SPPF), YOLOv11, and various image enhancement methodologies has significantly transformed the landscape of computer vision, enabling systems to process and interpret visual data with remarkable accuracy and efficiency. Sophisticated approaches have effectively addressed the challenge of image super-resolution, bridging the gap between low-resolution and high-resolution images [1].

In facial image super-resolution, leveraging facial priors has enhanced detail restoration and identity preservation, even in scenarios involving masked faces during the COVID-19 pandemic. Modern techniques have outperformed traditional machine learning approaches in functional image representation, crucial for maintaining model performance across diverse downstream tasks, especially with large-scale pre-trained datasets [2, 3]. Lightweight super-resolution networks have become essential for real-world applications, providing efficient solutions for image enhancement, particularly in challenging scenarios like crowd counting, where occlusions and density variations are prevalent

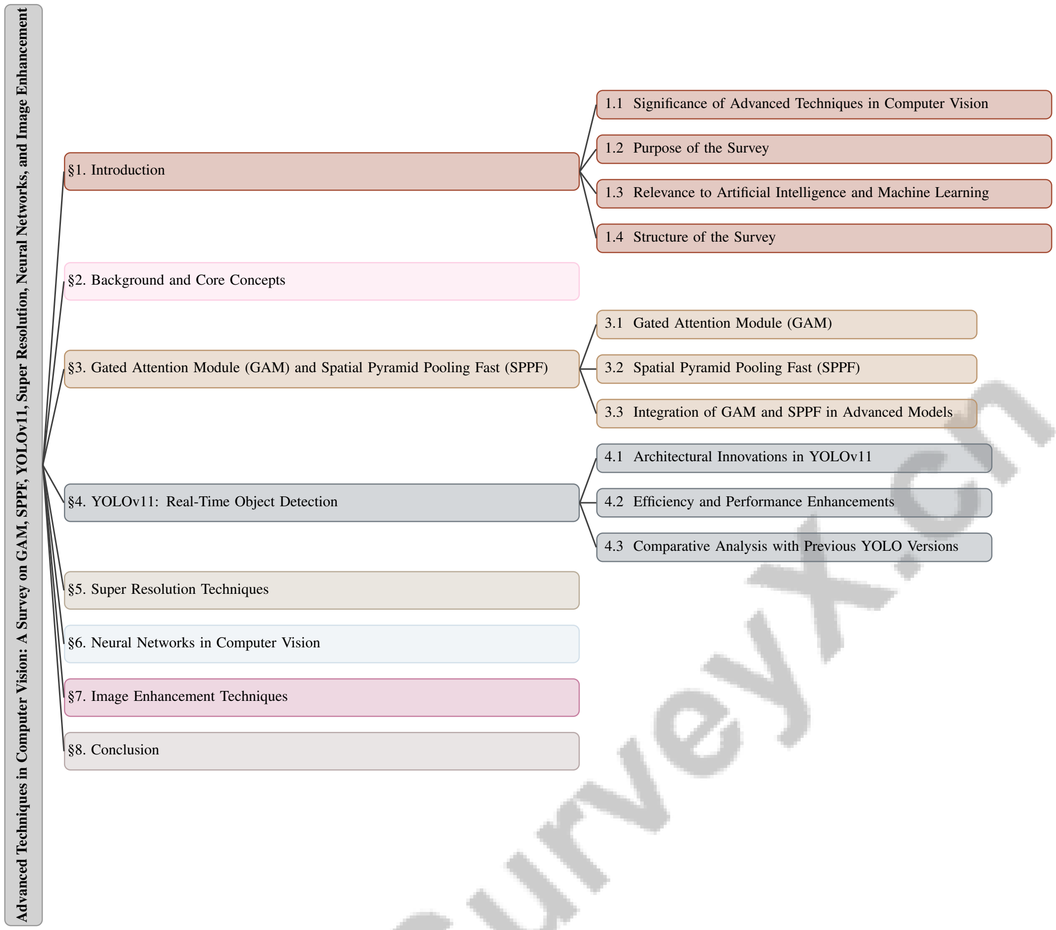


Figure 1: chapter structure

[4, 5]. The analysis of deep learning-based video super-resolution methods further underscores their significance in video processing [6].

In specialized fields such as Synthetic Aperture Radar (SAR) imaging and confocal laser endomicroscopy, advanced techniques have been pivotal in restoring high-resolution images from low-resolution inputs. These methodologies have also been instrumental in underwater imagery, addressing challenges like color distortion and low contrast, thereby facilitating high-level vision tasks [7]. The effectiveness of gaze tracking has improved despite challenges such as motion blur and noise, demonstrating the broad applicability of these advanced methodologies [8]. Furthermore, biologically plausible augmentations in self-supervised learning have shown significant impacts on representation learning, highlighting the transformative potential of these techniques [9].

These methodologies not only overcome the limitations of traditional computer vision approaches but also facilitate groundbreaking innovations that enhance the performance and versatility of computer vision systems across various applications, including autonomous vehicle navigation, security monitoring, and medical image analysis. By leveraging advanced techniques like image enhancement algorithms and deep learning models, these methodologies significantly advance artificial intelligence and machine learning, driving substantial progress in the field [10, 11, 12, 13].

1.2 Purpose of the Survey

This survey provides a comprehensive evaluation of cutting-edge techniques in computer vision, focusing on their role in enhancing image quality and resolution across diverse applications. A key objective is to investigate novel methodologies that improve model performance in visual recognition tasks, such as employing cortical magnification and saccades as augmentations in self-supervised learning, which have shown promise in enhancing representation learning [9].

Additionally, the survey addresses the challenge of efficiently acquiring high-resolution multi-focal plane images without axial scanning, as demonstrated by recent advancements facilitating faster image acquisition [14]. Techniques like the CISRNet model, which enhances the resolution of compressed images while minimizing compression noise, are highlighted for their effectiveness in preserving critical image details [15].

The integration of advanced techniques in specialized applications, exemplified by the DSIR-YOLO method for dual-stained cervical cell detection and the Light-YOLO0 framework for accurate fire detection in complex environments, is another focal point. The survey emphasizes the importance of bridging the gap between image quality and computational cost, particularly in satellite imagery, through the fusion of deep and non-deep methods [16].

In facial image reconstruction, frameworks like GCFSR are introduced, which reconstruct images with faithful identity without additional priors, addressing existing method limitations [17]. The survey also systematically reviews deep learning-based video super-resolution methods and proposes lightweight networks like MPRNet, delivering state-of-the-art performance with computational efficiency.

By concentrating on these objectives, the survey offers a thorough analysis of recent innovations in computer vision, highlighting applications such as autonomous vehicle navigation, security surveillance, medical imaging, and emergency response. It delineates future research directions, emphasizing the transformative potential of advanced techniques like super-resolution and deep learning in enhancing object detection accuracy and real-time analytics across various fields, providing critical insights into how these technologies can reshape industries [18, 19, 13, 10, 20].

1.3 Relevance to Artificial Intelligence and Machine Learning

The advanced techniques surveyed, including GAM, SPPF, YOLOv11, and Super Resolution, are pivotal in advancing AI and machine learning, particularly in enhancing image processing capabilities. The integration of deep learning frameworks, such as those employed in hypernetwork functional image representation, highlights the capacity of neural networks to approximate complex functions, facilitating sophisticated AI applications [2]. Super-resolution techniques are vital for improving the visual quality of images, playing a crucial role in AI-driven analytics tasks, including precision mapping in satellite imagery [16].

Innovations in object detection algorithms, such as Light-YOLOv5, which incorporates a separable vision transformer and a lightweight feature pyramid network, demonstrate AI's potential to enhance detection capabilities in challenging scenarios, including fire detection [21]. The significance of super-resolution is further illustrated in applications like gaze prediction, where combining super-resolution and self-supervised learning reduces the need for extensive labeled data, optimizing AI resource utilization [8].

The application of GAM in AI models contributes to transparency and explainability, essential for visual similarity and classification tasks, thereby enhancing the interpretability of machine learning models [22]. Techniques like MWCNN, which improve image restoration, are indispensable for AI applications requiring high-quality visual inputs [23]. Additionally, super-resolution techniques in underwater imagery address environmental challenges, ensuring AI systems perform high-level vision tasks effectively under adverse conditions [7].

In the medical domain, AI-driven methodologies inspired by the need for improved cervical cancer detection leverage advanced techniques to compensate for the shortage of pathologists, demonstrating AI's transformative impact on healthcare [24]. Collectively, these advancements not only enhance the capabilities of AI and ML systems but also pave the way for future innovations, reinforcing the indispensable role of advanced computer vision techniques in the broader AI and ML landscape.

1.4 Structure of the Survey

This survey is methodically structured to provide a comprehensive examination of advanced techniques in computer vision, focusing on their significance, methodologies, and applications. The survey begins with an that emphasizes the transformative role of advanced techniques in computer vision, specifically highlighting Generative Adversarial Models (GAM), Spatial Pyramid Pooling Fast (SPPF), YOLOv11 for real-time object detection, Super Resolution for enhancing image quality, and various neural network architectures. These methodologies improve accuracy in tasks such as autonomous navigation, security monitoring, and medical diagnostics, paving the way for innovative applications like user-driven image and video search, addressing challenges in human memory and retrieval processes [10, 25, 26, 27]. It also outlines the objectives and relevance of the survey to artificial intelligence and machine learning.

In the **Background and Core Concepts** section, the survey delves into the fundamental principles and technologies underpinning these advanced techniques, providing a foundational understanding necessary for appreciating their contributions to the field.

The subsequent sections are dedicated to detailed explorations of specific methodologies: **Gated Attention Module (GAM) and Spatial Pyramid Pooling Fast (SPPF)** are extensively covered to illustrate their roles in enhancing feature extraction and processing capabilities. This is followed by an in-depth analysis of **YOLOv11**, focusing on its architectural innovations, efficiency enhancements, and comparative performance analysis with previous versions.

The survey then addresses **Super Resolution Techniques**, emphasizing the role of neural networks in improving image resolution and quality. This section is complemented by a discussion on **Neural Networks in Computer Vision**, which examines their applications in object detection and recognition, as well as innovative techniques that push the boundaries of current capabilities.

The section on delves into the most recent algorithms and methodologies aimed at improving visual quality in challenging environments. It emphasizes advanced strategies for edge detection and noise reduction, particularly in low-light conditions where high noise levels and poor illumination complicate object detection tasks. The discussion includes a comparison of various image enhancement algorithms, such as Histogram Equalization and Retinex, and their effectiveness when applied to neural network models under adverse conditions, such as foggy or rainy weather. The goal is to enhance feature retrieval for improved performance in computer vision applications, ultimately providing insights for future research directions [13, 11, 28].

The survey concludes with a comprehensive that synthesizes the key findings from recent advancements in artificial intelligence (AI) and machine learning (ML), particularly highlighting their transformative impact on diverse applications such as computer vision, autonomous navigation, security, and medical imaging analysis. It reflects on the implications of these developments for future research directions and practical applications, emphasizing the potential for improved methodologies, such as the innovative use of 2D techniques to efficiently process 3D medical data, which could inspire further exploration in the field [10, 29]. This structured approach ensures a thorough and coherent presentation of the transformative potential of advanced computer vision techniques.The following sections are organized as shown in Figure 1.

2 Background and Core Concepts

2.1 Core Concepts and Technologies

The advancement of computer vision is fundamentally built on technologies such as Gated Attention Modules (GAM), Spatial Pyramid Pooling Fast (SPPF), YOLOv11, Super Resolution techniques, neural networks, and image enhancement methods. These technologies leverage extensive datasets and advanced computational capabilities to significantly enhance the efficiency of computer vision systems, enabling applications across autonomous navigation, security surveillance, and medical imaging [10, 30, 20].

GAM enhances neural network performance by focusing on relevant data areas, refining feature extraction through localized gradient and activation information across multiple layers. This selective attention mechanism is critical for high-precision tasks, improving accuracy and interpretability in visual similarity and classification applications [11, 31, 32, 22, 33].

SPPF addresses varying input sizes by pooling features across multiple scales, enhancing model robustness and adaptability, especially in object detection scenarios [21]. By using a pyramid of convolutional layers, SPPF effectively integrates multi-scale information, improving the model's capability to handle diverse inputs [34].

YOLOv11 represents a significant advancement in the "You Only Look Once" object detection series, introducing architectural innovations that enhance real-time detection capabilities. These enhancements are crucial for applications requiring rapid and accurate object detection, such as autonomous vehicles and surveillance systems. The YOLO series has also been instrumental in agricultural automation, optimizing orchard productivity through early fruit load management [35]. Super-resolution techniques, like B2BDet, enhance aerial image quality before employing a modified YOLOv5 model for effective object detection [36].

Super Resolution techniques aim to reconstruct high-resolution images from low-resolution inputs, focusing on detail restoration and identity preservation. The Residual Bilinear Attention Module (RBAM) improves feature extraction through a novel attention mechanism, enhancing super-resolution outcomes [37]. These techniques are vital for applications like masked face super-resolution, crucial for face recognition [38].

Neural networks are the backbone of advanced computer vision systems, facilitating complex pattern recognition and learning tasks. Their versatility spans applications from image classification to high-resolution multispectral image generation [39]. The integration of neural networks with data-driven methods and variational optimization techniques exemplifies the synergy between deep learning and traditional optimization [40].

Image enhancement techniques, such as Lit-Net, focus on multi-resolution and multi-scale analysis to improve visual quality [7]. These methods are crucial for object detection in challenging environments, such as underwater settings. The modified NL-means algorithm and BPNN are essential for improving image resolution and mitigating challenges posed by speckle noise in SAR images [41].

Generative Adversarial Networks (GANs) enhance these capabilities by providing frameworks for generating high-quality images, crucial in applications requiring image synthesis and enhancement. However, existing GAN-based super-resolution methods face challenges, particularly in reliance on image discriminators, highlighting the need for improved discrimination and visual quality in generated images [42]. Collectively, these core concepts and technologies form the foundation of modern computer vision systems, driving advancements in artificial intelligence and machine learning by facilitating more sophisticated and efficient image processing capabilities.

3 Gated Attention Module (GAM) and Spatial Pyramid Pooling Fast (SPPF)

Advancements in image analysis techniques are pivotal for enhancing neural network performance across diverse applications. The Gated Attention Module (GAM) represents a significant innovation, providing sophisticated mechanisms for improving feature extraction and processing through selective attention. As illustrated in Figure 2, the hierarchical structure of the Gated Attention Module and Spatial Pyramid Pooling Fast (SPPF) is depicted, detailing their methodologies, feature extraction capabilities, and applications in advanced computer vision models. This figure highlights the integration of GAM and SPPF in enhancing model performance across various domains, such as image super-resolution, video processing, and object detection. The contributions of these methodologies are critical in improving accuracy, efficiency, and adaptability in complex visual data analysis. This section explores GAM and SPPF, emphasizing their roles in significantly advancing image analysis and restoration tasks.

3.1 Gated Attention Module (GAM)

The Gated Attention Module (GAM) refines feature extraction in neural networks by employing advanced attention mechanisms. By focusing on relevant data regions, GAM reduces noise and emphasizes critical features, enhancing precision in high-level tasks like visual similarity and classification, especially in image super-resolution [22]. Integrating GAM into convolutional neural networks (CNNs) is crucial for achieving state-of-the-art image restoration outcomes. Attention mechanisms, such as those in the Multi-Scale Super-Resolution Module (MSSRM), guide networks in estimating lost details, improving information for tasks like crowd counting [5]. GAM's application in the Deep Multi-Frame Face Super-Resolution (DMFSR) method enables simultaneous face super-resolution, alignment, and recognition by processing sequences of low-resolution images [43].

As illustrated in Figure 3, GAM's versatility extends to various super-resolution tasks, encompassing image, video, and multimodality image super-resolution. This figure highlights the integration of advanced attention mechanisms, which not only enhance visual similarity but also effectively utilize inter-frame information and enable GAN-based learning for improved image quality [6]. In

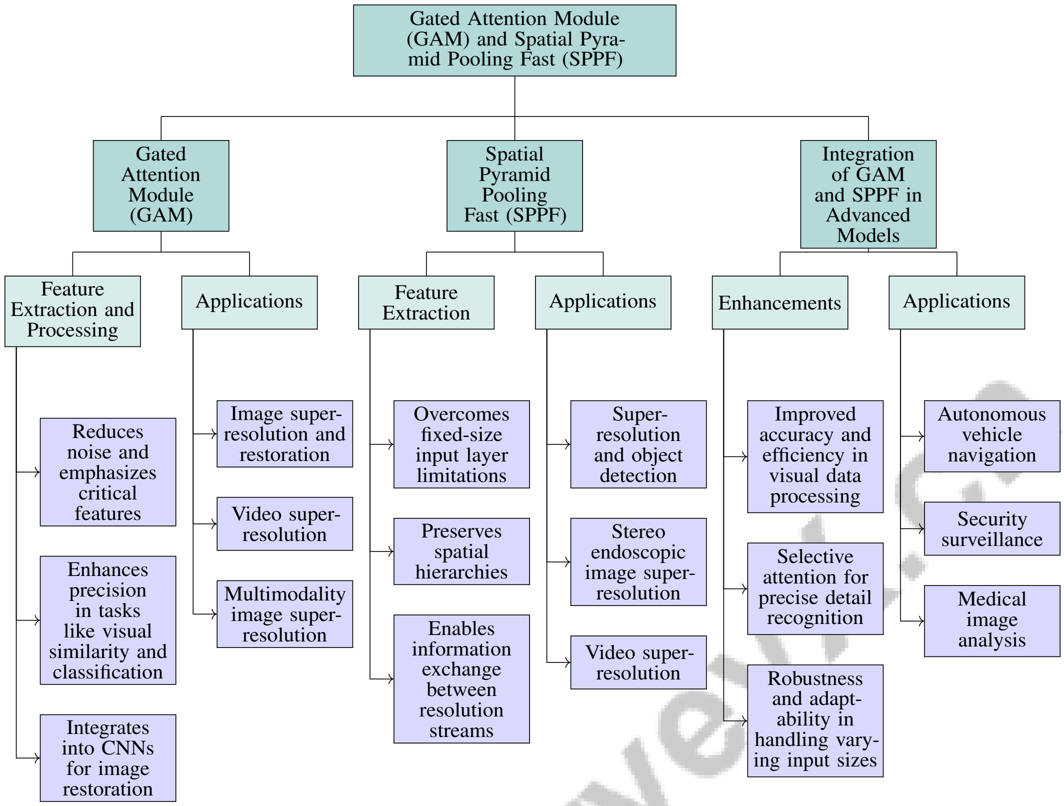


Figure 2: This figure illustrates the hierarchical structure of the Gated Attention Module (GAM) and Spatial Pyramid Pooling Fast (SPPF), detailing their methodologies, feature extraction capabilities, and applications in advanced computer vision models. It highlights the integration of GAM and SPPF in enhancing model performance across various domains, such as image super-resolution, video processing, and object detection, emphasizing their contributions to improving accuracy, efficiency, and adaptability in complex visual data analysis.

image restoration, the MWCNN uses multi-level discrete wavelet transforms for downsampling and inverse transforms for upsampling, refining feature representation and demonstrating GAM's efficacy in enhancing image details [23]. Additionally, GAM optimizes architectures for single image super-resolution tasks, critical for high-precision applications. Multi-resolution and multi-scale attention networks, followed by image reconstruction modules, significantly improve image quality [7]. GAM's capabilities are also evident in multimodality image super-resolution, where GANs learn complex mappings between different image modalities, enabling realistic image generation [38].

GAM's methodologies and applications demonstrate its transformative impact on enhancing the precision and efficiency of computer vision systems. By integrating attention mechanisms, GAM significantly improves neural networks' interpretability and accuracy, particularly in visual similarity and classification tasks, while laying the groundwork for future advancements in artificial intelligence and machine learning applications, including safety-critical environments and crisis management [31, 20, 18, 22].

3.2 Spatial Pyramid Pooling Fast (SPPF)

Spatial Pyramid Pooling Fast (SPPF) enhances image analysis by overcoming the limitations of fixed-size input layers in convolutional neural networks. This technique facilitates robust feature extraction from images of varying sizes through a multi-scale pooling approach, preserving spatial hierarchies and enhancing model adaptability. SPPF employs parallel multi-resolution convolution streams to capture diverse scale features while enabling information exchange between these streams, maintaining high-resolution spatial details and integrating contextual information from lower-resolution representations. Consequently, SPPF improves performance in various image pro-

AI-generated, for reference only.

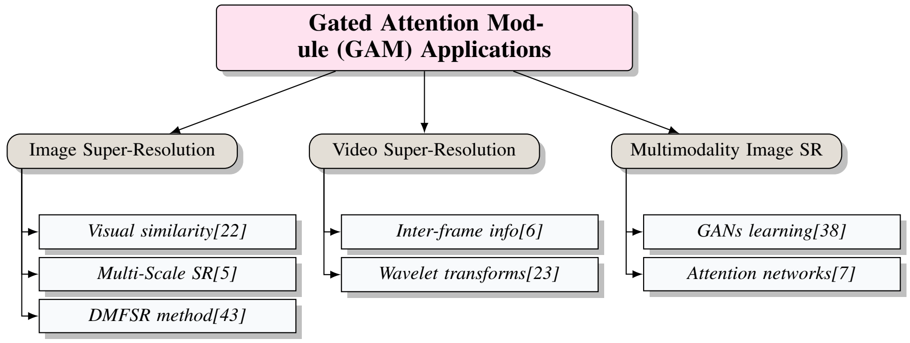


Figure 3: This figure illustrates the applications of the Gated Attention Module (GAM) in various super-resolution tasks, including image, video, and multimodality image super-resolution. It highlights the integration of advanced attention mechanisms to enhance visual similarity, utilize inter-frame information, and enable GAN-based learning for improved image quality.

cessing tasks, including super-resolution, denoising, and enhancement [44, 45]. Its versatility is evident across domains such as super-resolution and object detection.

Integrating SPPF into advanced models significantly enhances feature extraction capabilities. For instance, the SEGSRNet method employs an Atrous Spatial Pyramid Pooling (ASPP) block to enhance feature extraction for stereo endoscopic images, where precise detail recovery is critical [46]. This approach effectively captures multi-scale contextual information essential for reconstructing high-resolution images from low-resolution inputs. In video super-resolution, SPPF addresses challenges like motion estimation inaccuracies and maintaining temporal coherence in reconstructed frames [6]. By leveraging SPPF's multi-scale pooling capabilities, models can better align and integrate temporal information, enhancing the overall quality and coherence of super-resolved videos.

The SSIF method exemplifies the application of continuous representations in both spatial and spectral domains to generate high-resolution images, highlighting SPPF's effectiveness in capturing detailed spatial features [47]. This continuous representation approach facilitates the extraction of fine-grained features critical for high-quality image reconstruction and analysis.

SPPF's ability to pool features at multiple scales and its integration into advanced neural network architectures emphasize its effectiveness in enhancing image analysis. The Spatial Pyramid Pooling Framework (SPPF) enhances computer vision tasks by adeptly managing varying input sizes and capturing multi-scale information. This capability facilitates processing images with diverse resolutions, significantly improving accuracy and efficiency in applications such as autonomous vehicle navigation, security surveillance, and medical image analysis. By leveraging advanced machine learning techniques, SPPF contributes to developing robust vision-based systems, enabling more precise and efficient image processing outcomes across various domains [10, 47].

3.3 Integration of GAM and SPPF in Advanced Models

Integrating Gated Attention Modules (GAM) and Spatial Pyramid Pooling Fast (SPPF) into advanced computer vision models significantly enhances their capacity to process complex visual data with improved accuracy and efficiency. GAM's focus on relevant regions enhances feature extraction and model interpretability, crucial for tasks such as visual similarity and small object localization [22]. This selective attention mechanism is particularly beneficial in scenarios requiring precise detail recognition and classification, directing computational resources towards high-contribution information while suppressing redundant data.

Conversely, SPPF contributes to robustness and adaptability by pooling features at multiple scales, essential for handling varying input sizes and capturing multi-scale contextual information. This flexibility is particularly advantageous in object detection and super-resolution tasks, where input dimensions can vary significantly [48]. The integration of SPPF into models like the Hybrid Optimized Deep Convolutional Neural Network (HODCNN) enables robust feature extraction and accurate object classification in complex images, showcasing its effectiveness in enhancing model performance [12].

Advanced models such as PanFlowNet illustrate the potential of incorporating these techniques by extending the vanilla flow model to a probabilistic multi-conditional flow model, addressing previous methods' limitations in tasks like pan-sharpening [39]. This approach allows for nuanced adaptation to the multi-conditionality of complex tasks, improving overall model efficacy.

Moreover, integrating GAM and SPPF into frameworks designed for visual recognition tasks highlights their versatility and applicability across various applications [49]. By combining these techniques with innovations like parameter-free attention networks, which maximize representational power without additional computational overhead, models can achieve enhanced performance in both supervised and self-supervised learning contexts.

The integration of GAM and SPPF into advanced computer vision models not only enhances their feature extraction and processing capabilities but also paves the way for future innovations in artificial intelligence and machine learning. These advanced techniques significantly improve models' ability to tackle intricate visual tasks with enhanced precision and efficiency, transforming various applications such as autonomous vehicle navigation, security surveillance, and medical image analysis. By leveraging vast datasets and sophisticated machine learning methods, these models can achieve human-level or superior accuracy, even in challenging conditions such as poor lighting or adverse weather, revolutionizing the field of computer vision [10, 11].

4 YOLOv11: Real-Time Object Detection

4.1 Architectural Innovations in YOLOv11

YOLOv11 incorporates several architectural innovations to enhance object detection, particularly for real-time applications. A notable advancement is the integration of enhanced attention mechanisms, including the Global Attention Mechanism, which optimizes feature extraction by directing computational resources towards critical regions of the input data. This is complemented by improvements in feature fusion processes that effectively combine multi-scale features, thereby boosting the model's capability to detect objects in diverse input sizes and complex visual scenes [50].

The model's architecture features novel components such as the C3k2 block and C2PSA (Channel and Spatial Attention), which collectively improve accuracy and processing speed. These innovations enable the model to adapt to varying input dimensions and conditions effectively. Additionally, the implementation of ghost convolution layers and C3Ghost modules reduces the number of parameters and floating point operations per second (FLOPs), enhancing efficiency for deployment in resource-constrained environments, including mobile and embedded devices [21].

In specialized applications, YOLOv11 demonstrates versatility through adaptations like the YOLO-based Safety Helmet Detection Framework (YSHDF), which integrates GhostNetv2 and attention mechanisms tailored for specific detection tasks, illustrating the model's adaptability across various detection scenarios [31]. Furthermore, the DAD method, which co-designs optoelectronic systems with YOLOv11, achieves significant reductions in multiply–accumulate (MAC) operations while maintaining comparable top-1 accuracy to electronic-only models, showcasing the potential for hardware-software integration in optimized object detection [51].

The application of super-resolution techniques, exemplified by the BGYolo method, enhances underwater images before applying a tailored YOLOv5 architecture, improving detection capabilities in challenging visual conditions. This reflects a trend in YOLOv11's architecture that incorporates advanced image processing techniques—such as the C3k2 block, SPPF, and C2PSA components—to enhance detection performance, feature extraction, and computational efficiency across various computer vision tasks, including object detection and instance segmentation [25, 52].

As illustrated in Figure 4, the architectural innovations in YOLOv11 not only enhance efficiency and accuracy but also broaden its applicability across domains, from safety monitoring to underwater exploration. These advancements position YOLOv11 as a leading model in real-time object detection, adept at addressing challenges posed by diverse and dynamic environments. The ability to identify matching subnetworks within pre-trained models further enhances the model's transferability to downstream tasks, which is crucial for real-time applications [3].

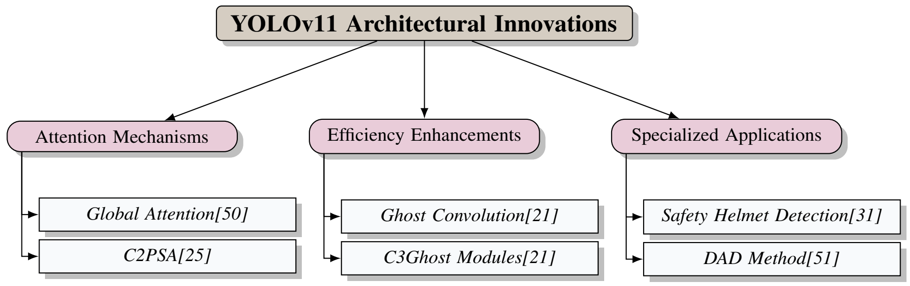


Figure 4: This figure illustrates the key architectural innovations in YOLOv11, highlighting the integration of advanced attention mechanisms, efficiency enhancements, and specialized applications. These innovations include the Global Attention Mechanism and C2PSA for improved feature extraction, Ghost Convolution and C3Ghost Modules for enhanced efficiency, and applications like Safety Helmet Detection and the DAD Method for specialized tasks.

4.2 Efficiency and Performance Enhancements

The efficiency and performance of YOLOv11 have significantly improved through architectural refinements and optimization strategies, making it highly effective for real-time computer vision tasks. The introduction of the G-YOLOv11 detector achieves a mean Average Precision (mAP) of 0.535 at a 0.5 Intersection over Union (IoU) threshold, with an inference time of 2.4 milliseconds on an NVIDIA A10 GPU, establishing a new standard for efficiency, particularly in fracture detection applications [53].

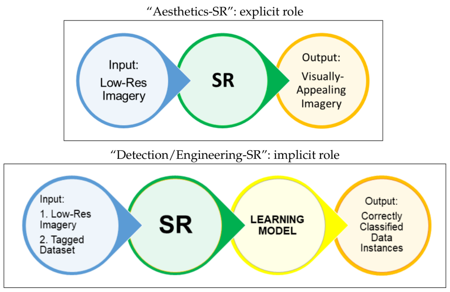
Further evaluations on a helmet detection dataset from Kaggle, utilizing an AMD CPU 5800x and an NVIDIA GeForce RTX4090, underscore YOLOv11's performance in mAP and IoU metrics, illustrating the model's robustness and adaptability across different hardware platforms [31]. Benchmark results reveal that YOLOv11 achieves significant gains in speed and precision, making it invaluable for real-time applications requiring rapid decision-making [25].

Advanced optimization techniques, evidenced in hybrid approaches like HODCNN, further bolster the model's accuracy and efficiency in detecting objects across varied scenarios. These techniques optimize convolutional layers and feature extraction processes, allowing YOLOv11 to maintain superior performance even in complex visual environments [12]. Additionally, the B2BDet method, employing a modified YOLOv5 architecture, demonstrates exceptional performance in detecting small, densely clustered objects in aerial imagery, achieving a 52.5

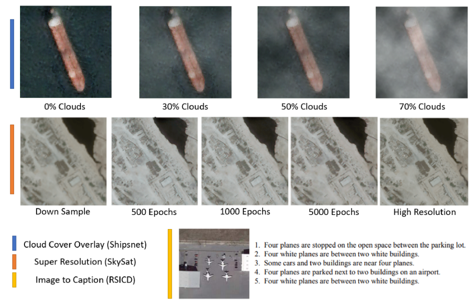
In underwater target detection, the YOLOv7-AC method highlights the effectiveness of YOLOv11's architectural innovations by capturing and utilizing rich feature information while minimizing computational burdens. This balance between accuracy and speed is crucial for applications requiring real-time processing in resource-limited environments [54]. The Light-YOLOv5 variant exemplifies this balance by enhancing feature representation through novel architectural modifications, ensuring computational efficiency without sacrificing detection accuracy [21].

The advancements in YOLOv11's efficiency and performance attest to the model's adaptability and robustness across diverse domains, from safety monitoring to environmental analysis. These enhancements position YOLOv11 as a leading model in real-time object detection, capable of addressing challenges posed by dynamic and varied environments. The model's improvements mirror those seen in the updated YOLOv8, which also demonstrated superior generalization to new datasets, highlighting the ongoing evolution and refinement of the YOLO architecture [50].

As shown in Figure 5, the advancements in YOLOv11 have marked a significant leap in both efficiency and performance within the realm of real-time object detection, illustrated through various examples of image processing enhancements. The figure presents two distinct scenarios: the roles of Super-Resolution (SR) in image processing and the impact of cloud cover on satellite imagery. The first example highlights the dual roles of SR, contrasting its application in aesthetic enhancement versus its utility in detection and engineering contexts. The second example delves into the challenges posed by cloud cover on satellite imagery, providing a comparative analysis of different techniques used to mitigate these effects. Together, these examples underscore the critical role of innovative



(a) Aesthetics-SR and Detection/Engineering-SR: Different Roles of SR in Image Processing[27]



(b) Effect of Cloud Cover on Satellite Imagery: Comparison of Different Techniques[55]

Figure 5: Examples of Efficiency and Performance Enhancements

technology in overcoming traditional limitations in real-time object detection and image processing [27, 55].

4.3 Comparative Analysis with Previous YOLO Versions

Benchmark	Size	Domain	Task Format	Metric
YOLOv11[25]	1,000,000	Object Detection	Detection	mAP, FPS
ExDark[13]	12,000	Object Detection	Object Detection	mAP
SRCNN[56]	5,000,000	Image Processing	Image Reconstruction	BRISQUE, PSNR
SR-FR[57]	4,500	Face Recognition	Unsupervised Learning	Chi Square
SAM[20]	6,397	Concealed Object Segmentation	Segmentation	F-measure, S
SRGA[58]	6,400	Image Super-Resolution	Generalization Assessment	SRGA
SR-MER[59]	256	Micro-expression Recognition	Classification	PSNR, WAR
GHM[60]	3,000	Image Processing	Single Image Super Resolution	PSNR, SSIM

Table 1: This table provides a comprehensive overview of various benchmarks used in evaluating advanced computer vision models. It includes information on benchmark size, domain, task format, and the metrics employed for performance assessment. The table highlights the diversity and scope of datasets and tasks relevant to cutting-edge research in object detection, image processing, and other related fields.

YOLOv11 represents a significant advancement in the YOLO series, consistently outperforming its predecessors in mean Average Precision (mAP) scores and inference rates across various tasks, demonstrating enhanced capabilities in object detection applications [25]. Evaluations against previous iterations, including YOLOv8, YOLOv9, and YOLOv10, indicate superior accuracy and reduced latency, establishing YOLOv11 as a more efficient model for real-time applications [52].

The advancements in YOLOv11 are further highlighted by its performance on challenging datasets, such as VisDrone2021-DET and SeaDronesSeeV2, characterized by numerous small-scale objects. YOLOv11 demonstrated remarkable improvements over the baseline YOLOv5s, showcasing its ability to accurately detect small objects in complex environments, which is crucial for applications like wildlife monitoring where previous YOLO versions struggled with generalization to unseen datasets [61, 50].

The YOLOv7-AC model, a variant of the YOLO series, displayed superior performance in underwater target detection compared to traditional YOLOv7 and other models, achieving mAP values of 89.6

The comparative analysis of YOLOv11 with its predecessors reveals significant advancements in performance and feature integration, particularly through innovations like the C3k2 block, SPPF, and C2PSA components, which enhance feature extraction and computational efficiency. YOLOv11 demonstrates improved mean Average Precision (mAP) and versatility across various model sizes, catering to applications from edge devices to high-performance computing environments. These enhancements make YOLOv11 a leading choice for real-time object detection tasks, effectively addressing challenges in complex environments and supporting a wide range of computer vision applications, including instance segmentation and pose estimation [25, 52, 35]. Table 1 presents a detailed summary of the benchmarks utilized for evaluating the performance of state-of-the-art models, such as YOLOv11, across different computer vision tasks.

5 Super Resolution Techniques

5.1 Role of Neural Networks in Image Super-Resolution

Neural networks play a pivotal role in image super-resolution, enhancing resolution and quality across multiple domains. Attention mechanisms within these networks significantly improve feature representation and image quality, crucial for tasks like Synthetic Aperture Radar (SAR) image enhancement, where complex mappings between low and high-resolution pixels are learned [37, 41]. The Matrixed Channel Attention Network (MCAN) exemplifies this by utilizing a matrix ensemble of multi-connected channel attention blocks, optimizing feature utilization and image fidelity [62].

In medical imaging, the Deep Network with Spatio-Structural Priors (DNSP) incorporates low-rank and sharpness priors, enhancing super-resolution for MR images, demonstrating the value of structural priors in detail preservation [63]. The Any-time super-Resolution Method (ARM) showcases flexibility by dynamically selecting subnets based on input patch complexity, beneficial in resource-constrained settings [64]. Additionally, integrating super-resolution with image compression, as seen in the Factorized Fields approach, enhances detail recovery in compressed images [65].

These advancements are visually summarized in Figure 6, which illustrates the role of neural networks in image super-resolution, highlighting key advancements in attention mechanisms, medical imaging applications, and integration with image compression techniques. Collectively, these neural network advancements significantly improve image super-resolution, supporting applications in autonomous navigation, security monitoring, and medical diagnostics. They establish a robust foundation for future innovations in AI and machine learning, leveraging extensive datasets and sophisticated algorithms to enhance image enhancement and object detection [11, 66, 29, 13, 10].

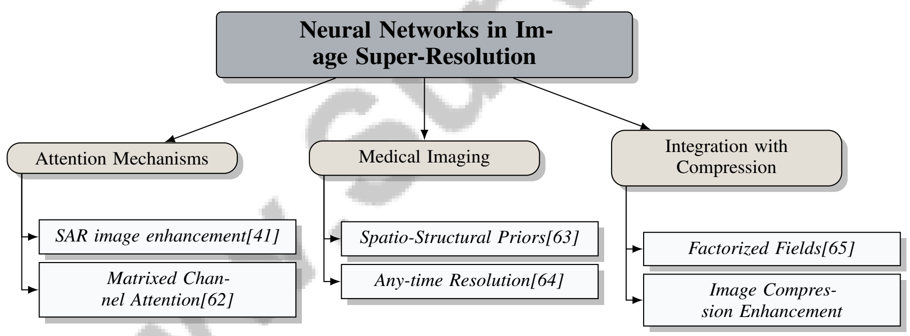


Figure 6: This figure illustrates the role of neural networks in image super-resolution, highlighting key advancements in attention mechanisms, medical imaging applications, and integration with image compression techniques.

5.2 Innovative Architectures and Methodologies

Method Name	Architectural Innovation	Methodological Advancements	Application Scenarios
MCAN[62]	Matrixed Channel Attention	Matrix-in-matrix	-
FIPER[65]	Factorized Fields Approach	Factorized Fields Representation	Environmental Monitoring
SIS[29]	2D Super Images	Super Image Segmentation	Medical Imaging
ASFSR[67]	Matrixed Channel Attention	Dsrnet	Environmental Monitoring

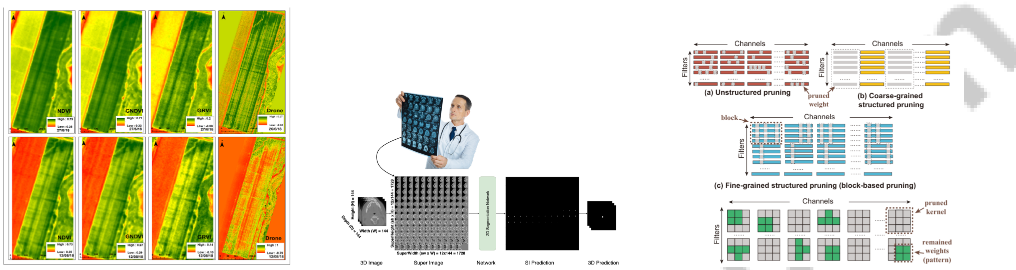
Table 2: A comparative overview of recent advancements in super-resolution methodologies, highlighting architectural innovations and methodological advancements across various application scenarios. The table includes methods such as MCAN, FIPER, SIS, and ASFSR, detailing their unique features and specific domains of application.

Recent developments in super-resolution have led to innovative architectures that significantly enhance image quality. The Matrixed Channel Attention Network (MCAN) employs a matrix-in-matrix structure, leveraging hierarchical features through multi-connected blocks to improve Peak

Signal-to-Noise Ratio (PSNR) performance [62]. The Factorized Fields approach decomposes images into multi-frequency components, facilitating superior detail recovery in super-resolution and image compression tasks, achieving a 204.4

Innovative methodologies, such as DSRNet with specialized enhancement blocks and scale-arbitrary networks with conditional convolutions, further improve image quality and resolution, addressing complex scenes and non-integer scaling challenges. These advancements extend super-resolution applicability to mobile devices and various image enhancement tasks [68, 69, 70, 71]. They enhance the accuracy and reliability of image analysis, paving the way for future AI and machine learning innovations.

Table 2 presents a comprehensive comparison of innovative architectures and methodologies in super-resolution, showcasing the diverse applications and advancements in the field.



(a) Comparison of NDVI, GNDVI, and GRVI with Drone Imagery[27]

(b) Medical Image Analysis and Segmentation[29]

(c) Unstructured, Coarse-Grained, and Fine-Grained Structured Pruning of Convolutional Neural Networks[67]

Figure 7: Examples of Innovative Architectures and Methodologies

As illustrated in Figure 7, super-resolution techniques are advancing through innovative architectures and methodologies. For example, comparing vegetation indices with drone imagery aids environmental monitoring and agricultural management. In medical imaging, advanced networks facilitate precise 3D segmentation, enhancing diagnostic accuracy. Pruning techniques in CNNs demonstrate potential for optimizing network performance by reducing complexity without sacrificing accuracy. These examples highlight the diverse applications and transformative impact of super-resolution techniques, emphasizing the importance of ongoing research and innovation in this area [27, 29, 67].

6 Neural Networks in Computer Vision

6.1 Neural Networks in Object Detection and Recognition

Neural networks have revolutionized object detection and recognition by offering robust frameworks that enhance image processing across diverse environments. A critical advancement is the integration of attention mechanisms, optimizing object detection by focusing computational resources on salient image regions. The Gated Attention Module (GAM) exemplifies this by generating high-resolution, class-discriminative saliency maps, enhancing small object recognition and improving visual similarity and classification tasks [22]. Hybrid architectures, such as the Hybrid Optimized Deep Convolutional Neural Network (HODCNN), merge convolutional layers with advanced feature extraction techniques, excelling in tasks like optical character recognition (OCR), where detail recognition is crucial [12].

In image super-resolution, models like the Super-Resolution Convolutional Neural Network (SRCNN) and the Super-Resolution Generative Adversarial Network (SRGAN) have achieved significant improvements in image quality and resolution [72]. These models effectively reconstruct high-resolution images from low-resolution inputs, addressing high-frequency information loss during enlargement [73]. The end-to-end adaptive Monte Carlo framework exemplifies the integration of super-resolution and denoising tasks, enhancing rendering speed and quality by improving information flow and feature learning [74]. The one-shot image restoration approach further underscores neural networks' role in improving sample efficiency and generalization in pattern detection and recognition [75].

Innovative neural network architectures have transcended local window attention limitations by effectively modeling long-range dependencies and utilizing external prior information, leading to superior super-resolution outcomes without increasing model complexity [76]. These advancements are crucial for enhancing object detection and recognition systems in complex visual scenarios. By incorporating advanced attention mechanisms and optimizing architectural designs, neural networks provide robust solutions that enhance the accuracy and efficiency of computer vision systems across various applications. The progress in neural network methodologies, exemplified by models like the Hybrid Optimized Dense Convolutional Neural Network with a detection accuracy of 0.9864, significantly enhances applications in autonomous navigation, security surveillance, and medical diagnostics, paving the way for future innovations in artificial intelligence and machine learning as we explore the integration of these systems with multimodal and context-aware technologies [52, 11, 13, 12, 10].

6.2 Innovative Neural Network Techniques in Computer Vision

Innovative neural network techniques have significantly advanced image processing in computer vision, offering sophisticated solutions for complex tasks. Methods like GPRM and GIDB efficiently distill and aggregate information, optimizing computational resources while maintaining high performance, highlighting the importance of grouped information distillation in enhancing neural network efficiency for intricate visual tasks [77]. The application of detail priors in GAN-based architectures, such as the DSRGAN model, markedly improves the restoration of realistic details in perceptual single-image super-resolution tasks, emphasizing detail enhancement's role in achieving high-quality image outputs [78].

Exploring inter-frame information in video super-resolution has been pivotal, leveraging temporal coherence to enhance reconstructed video quality. Deep learning methods have advanced this area, demonstrating neural networks' potential to utilize temporal data for improved video processing [6]. Additionally, the lottery ticket hypothesis in supervised learning contexts facilitates significant model size reductions while preserving performance, illustrating neural networks' capacity to achieve efficient architectures without compromising accuracy, enhancing their applicability in resource-constrained environments [3].

Innovative neural network techniques are advancing computer vision by providing enhanced methodologies for image quality improvement and processing capabilities. These advancements are vital for applications like autonomous vehicle navigation, where neural networks must adeptly handle adverse conditions such as darkness, overexposure, and fog. Studies analyzing image enhancement algorithms like Histogram Equalization and Retinex reveal that preprocessing images can significantly improve established neural network models' performance, such as ResNet, GoogleLeNet, YOLO, and Vision Transformers, in challenging environments. The evolution of machine learning has enabled more accurate multimedia information processing, facilitating applications in security, healthcare diagnostics, and potentially transforming user-driven image search capabilities [10, 11]. As research advances, developing methods that efficiently utilize computational resources while maintaining high performance remains a crucial exploration area.

7 Image Enhancement Techniques

7.1 Advanced Algorithms for Image Enhancement

Recent advancements in image enhancement involve sophisticated algorithms that leverage high-resolution spatial details and contextual information to significantly improve visual quality. The MIRNet framework exemplifies this trend by combining spatial features with contextual data, achieving state-of-the-art performance across various image restoration tasks and enhancing clarity and quality for diverse applications [79]. In challenging conditions, techniques like Histogram Equalization (HE) and RX algorithms enhance model performance by adjusting contrast and improving feature visibility, especially when environmental factors compromise image quality. These techniques' effectiveness varies based on model type and conditions, necessitating tailored enhancement approaches [11].

Video enhancement research focuses on using trained filters to repair low-quality outputs, addressing issues like noise and blurring, thus restoring quality and enhancing viewer experience [28]. The

AI-generated, for reference only.

Selective Deep Super-Resolution (SDSR) framework further illustrates advancements by intelligently selecting regions for deep super-resolution based on information content, optimizing resource allocation and enhancing image quality in a targeted manner [16]. These algorithms reflect the ongoing evolution of image enhancement techniques, providing robust solutions that improve image quality across various applications. The integration of high-resolution detail extraction and contextual analysis is anticipated to drive significant innovations in fields such as medical imaging, where super-resolution methods enhance volumetric data interpretation, and sports analytics, where advanced image processing techniques improve object detection accuracy [29, 27, 19, 59, 80].

7.2 Edge and Noise Reduction Techniques

Edge and noise reduction techniques are essential for enhancing visual quality, particularly where clarity and detail are vital. These techniques suppress noise while preserving edge features, resulting in clear and informative images. Hierarchical residual attention networks effectively reduce noise by focusing on relevant image regions, enhancing edge clarity [81]. The Multi-Level Wavelet Convolutional Neural Network (MWCNN) improves edge preservation and noise reduction using discrete wavelet transforms with convolutional network blocks, enhancing feature representation and reducing noise [23].

Advanced filtering techniques, like the Guided Edge-Aware Smoothing and Sharpening Filter (GEASSF), address noise reduction and edge enhancement by smoothing noise while sharpening edges, improving overall clarity [48]. This approach benefits scenarios where edge details are critical, such as medical imaging and high-resolution satellite imagery. Neural networks in edge and noise reduction tasks have led to solutions like the Parameter-Free Attention Network (PFAN), enhancing edge clarity without additional computational overhead [82]. These techniques significantly improve image quality in challenging conditions, enhancing image analysis reliability across diverse applications, including autonomous vehicle navigation, medical diagnostics, and security surveillance. Methods like histogram equalization and deep neural networks demonstrate that these enhancements improve feature retrieval and object detection model performance, facilitating more effective computer vision systems [13, 11, 10, 66].

7.3 Enhancement in Challenging Environments

Enhancing images captured under challenging environmental conditions requires addressing issues like low light, noise, and poor contrast, which degrade image quality. Advanced methodologies leverage both traditional image processing techniques and modern neural network architectures to tackle these challenges. The MIRNet framework excels in this domain by combining detailed spatial features with rich contextual information, enhancing clarity across multiple restoration tasks, as evidenced by its performance on benchmark datasets like DND and SIDD for image denoising, RealSR for super-resolution, and LoL and MIT-Adobe FiveK for image enhancement [79].

In low-light environments, algorithms focused on histogram equalization and contrast enhancement improve visibility and feature recognition, employing methods like HE and Retinex for effective object detection [13, 11]. Integrating these methods with neural network approaches, such as hierarchical residual attention networks, further enhances performance by directing computational resources to critical image regions. Specialized filters like the Guided Edge-Aware Smoothing and Sharpening Filter (GEASSF) address noise reduction and edge enhancement through a novel approach integrating guided edge-aware smoothing and sharpening techniques, ensuring high clarity and detail retention even under challenging conditions like low light or motion blur [48, 66, 28, 83]. This is particularly beneficial in applications like medical imaging and high-resolution satellite imagery, where precision is critical.

Methodologies for enhancing images in challenging environments are rapidly advancing, driven by innovations in traditional processing techniques and modern neural network architectures. Recent studies demonstrate these enhancements' effectiveness in improving neural network performance in adverse conditions like low light, fog, and overexposure. These advancements address the limitations of networks trained on ideal datasets, providing viable alternatives to retraining models for real-world applications, particularly in autonomous vehicle navigation [13, 11]. Such innovations enhance both visual quality and analysis reliability across a broad range of applications, ensuring critical details are preserved even in demanding conditions.

8 Conclusion

The survey illustrates the profound influence of advanced methodologies on computer vision, spotlighting innovations such as Gated Attention Modules (GAM), Spatial Pyramid Pooling Fast (SPPF), YOLOv11, Super Resolution, and neural networks. These techniques have markedly improved image analysis and processing capabilities, facilitating sophisticated applications across diverse domains. The enhancement of high-resolution image reconstruction in facial recognition exemplifies significant advancements in accuracy and visual quality, showcasing the potential of advanced super-resolution techniques. The CS-NL attention module's utilization of cross-scale feature similarities further boosts single-image super-resolution (SISR) performance.

Architectural innovations like MPRNet surpass existing lightweight models, achieving an ideal balance between model size and reconstruction quality, which is crucial for practical applications. The Multi-Level Wavelet Convolutional Neural Network (MWCNN) highlights progress in image restoration, offering a better tradeoff between efficiency and performance through wavelet transforms. Furthermore, the integration of advanced methodologies in video super-resolution GAN architectures has resulted in notable improvements in PSNR and SSIM metrics, underscoring GANs' potential in enhancing video quality.

The conclusion emphasizes hypernetworks' effectiveness in generating functional image representations, achieving outcomes comparable to specialized super-resolution models, thus demonstrating their relevance in complex systems. Future research should prioritize optimizing models like YOLOv11 to enhance speed while maintaining or improving accuracy, and exploring their application in intricate detection scenarios. Expanding datasets for comprehensive validation and optimizing models for real-time applications are also crucial areas for future investigation.

These advancements collectively highlight the necessity for continuous exploration and innovation in computer vision technologies, with significant implications for advancing artificial intelligence and machine learning applications. As research evolves, the integration of advanced methodologies and innovative architectures is poised to drive further innovations, paving the way for more efficient and effective image processing solutions across various fields. The focus on reducing inference time while maintaining competitive performance, as demonstrated by existing approaches, underscores the potential for future developments in this rapidly advancing domain. The proposed Factorized Fields representation effectively addresses challenges in joint image compression and super-resolution, achieving state-of-the-art performance in both tasks through innovative modeling of multi-frequency components.

References

[1] Marco Pesavento, Marco Volino, and Adrian Hilton. Attention-based multi-reference learning for image super-resolution, 2021.

[2] Sylwester Klocek, Łukasz Maziarka, Maciej Wołczyk, Jacek Tabor, Jakub Nowak, and Marek Śmieja. Hypernetwork functional image representation, 2019.

[3] Tianlong Chen, Jonathan Frankle, Shiyu Chang, Sijia Liu, Yang Zhang, Michael Carbin, and Zhangyang Wang. The lottery tickets hypothesis for supervised and self-supervised pre-training in computer vision models, 2021.

[4] Armin Mehri, Parichehr B. Ardakani, and Angel D. Sappa. Mprnet: Multi-path residual network for lightweight image super resolution, 2020.

[5] Jiahao Xie, Wei Xu, Dingkang Liang, Zhanyu Ma, Kongming Liang, Weidong Liu, Rui Wang, and Ling Jin. Super-resolution information enhancement for crowd counting, 2023.

[6] Hongying Liu, Zhubo Ruan, Peng Zhao, Chao Dong, Fanhua Shang, Yuanyuan Liu, Linlin Yang, and Radu Timofte. Video super resolution based on deep learning: A comprehensive survey, 2022.

[7] Alik Pramanick, Arijit Sur, and V. Vijaya Saradhi. Harnessing multi-resolution and multi-scale attention for underwater image restoration, 2024.

[8] Galen O'Shea and Majid Komeili. Toward super-resolution for appearance-based gaze estimation, 2023.

[9] Binxu Wang, David Mayo, Arturo Deza, Andrei Barbu, and Colin Conwell. On the use of cortical magnification and saccades as biological proxies for data augmentation, 2021.

[10] Alan F. Smeaton. Computer vision for supporting image search, 2021.

[11] Jonathan Sanderson and Syed Rafay Hasan. A case study of image enhancement algorithms' effectiveness of improving neural networks' performance on adverse images, 2023.

[12] Venkata Beri. Hybrid optimized deep convolution neural network based learning model for object detection, 2022.

[13] Winston Chen and Tejas Shah. Exploring low-light object detection techniques, 2021.

[14] Jiabo Ma, Sibo Liu, Shenghua Cheng, Xiuli Liu, Li Cheng, and Shaoqun Zeng. Reconstruct high-resolution multi-focal plane images from a single 2d wide field image, 2020.

[15] Agus Gunawan and Sultan Rizky Hikmawan Madjid. Cisrnet: Compressed image super-resolution network, 2022.

[16] Gaurav Kumar Nayak, Saksham Jain, R Venkatesh Babu, and Anirban Chakraborty. Fusion of deep and non-deep methods for fast super-resolution of satellite images, 2020.

[17] Jingwen He, Wu Shi, Kai Chen, Lean Fu, and Chao Dong. Gcfsr: a generative and controllable face super resolution method without facial and gan priors, 2022.

[18] Dat Tien Nguyen, Firoj Alam, Ferda Ofli, and Muhammad Imran. Automatic image filtering on social networks using deep learning and perceptual hashing during crises, 2017.

[19] Karolina Seweryn, Gabriel Chęć, Szymon Łukasik, and Anna Wróblewska. Improving object detection quality in football through super-resolution techniques, 2024.

[20] Ge-Peng Ji, Deng-Ping Fan, Peng Xu, Ming-Ming Cheng, Bowen Zhou, and Luc Van Gool. Sam struggles in concealed scenes – empirical study on segment anything, 2023.

[21] Light-yolov5: A lightweight algorithm for improved yolov5 in complex fire scenarios.

[22] Oren Barkan, Omri Armstrong, Amir Hertz, Avi Caciularu, Ori Katz, Itzik Malkiel, and Noam Koenigstein. Gam: Explainable visual similarity and classification via gradient activation maps, 2021.

[23] Pengju Liu, Hongzhi Zhang, Kai Zhang, Liang Lin, and Wangmeng Zuo. Multi-level wavelet-cnn for image restoration, 2018.

[24] Xiao-Jun Wu, Cai-Jun Zhao, Chun Meng, and Hang Wang. Research on cervical cancer p16/ki-67 immunohistochemical dual-staining image recognition algorithm based on yolo, 2024.

[25] Rahima Khanam and Muhammad Hussain. Yolov11: An overview of the key architectural enhancements, 2024.

[26] Zhihao Wang, Jian Chen, and Steven C. H. Hoi. Deep learning for image super-resolution: A survey, 2020.

[27] Martin Feder, Michal Horovitz, Assaf Chen, Raphael Linker, and Ofer M. Shir. Lessons learned report: Super-resolution for detection tasks in engineering problem-solving, 2023.

[28] Lijun Wang and Ling Shao. An algorithm for repairing low-quality video enhancement techniques based on trained filter, 2011.

[29] Ikboljon Sobirov, Numan Saeed, and Mohammad Yaqub. Super images – a new 2d perspective on 3d medical imaging analysis, 2023.

[30] Chengxing Xie, Xiaoming Zhang, Linze Li, Yuqian Fu, Biao Gong, Tianrui Li, and Kai Zhang. Mat: Multi-range attention transformer for efficient image super-resolution, 2024.

[31] Shuqi Shen and Junjie Yang. Better yolo with attention-augmented network and enhanced generalization performance for safety helmet detection, 2024.

[32] Qilei Li, Zhen Li, Lu Lu, Gwanggil Jeon, Kai Liu, and Xiaomin Yang. Gated multiple feedback network for image super-resolution, 2019.

[33] Zilong Chen, Yaling Liang, and Minghui Du. Attention based broadly self-guided network for low light image enhancement, 2021.

[34] Xiaoyong Shen, Ying-Cong Chen, Xin Tao, and Jiaya Jia. Convolutional neural pyramid for image processing, 2017.

[35] Ranjan Sapkota, Zhichao Meng, Martin Churuvija, Xiaoqiang Du, Zenghong Ma, and Manoj Karkee. Comprehensive performance evaluation of yolov12, yolo11, yolov10, yolov9 and yolov8 on detecting and counting fruitlet in complex orchard environments, 2025.

[36] Ragib Amin Nihal, Benjamin Yen, Katsutoshi Itoyama, and Kazuhiro Nakadai. From blurry to brilliant detection: Yolov5-based aerial object detection with super resolution, 2024.

[37] Saeed Izadi, Darren Sutton, and Ghassan Hamarneh. Image super resolution via bilinear pooling: Application to confocal endomicroscopy, 2019.

[38] Guangwei Gao, Lei Tang, Fei Wu, Huimin Lu, and Jian Yang. Jdsr-gan: Constructing an efficient joint learning network for masked face super-resolution, 2023.

[39] Gang Yang, Xiangyong Cao, Wenzhe Xiao, Man Zhou, Aiping Liu, Xun chen, and Deyu Meng. Panflownet: A flow-based deep network for pan-sharpening, 2023.

[40] Milena Gazdieva, Oleg Voynov, Alexey Artemov, Youyi Zheng, Luiz Velho, and Evgeny Burnaev. Can we use neural regularization to solve depth super-resolution?, 2021.

[41] Zeling Wu and Haoxiang Wang. Super-resolution reconstruction of sar image based on non-local means denoising combined with bp neural network, 2016.

[42] Xuan Zhu, Yue Cheng, Jinye Peng, Rongzhi Wang, Mingnan Le, and Xin Liu. Super-resolved image perceptual quality improvement via multi-feature discriminators, 2019.

[43] E. Ustinova and V. Lempitsky. Deep multi-frame face super-resolution, 2017.

[44] Ruicheng Feng, Weipeng Guan, Yu Qiao, and Chao Dong. Exploring multi-scale feature propagation and communication for image super resolution, 2020.

[45] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for fast image restoration and enhancement, 2022.

[46] Mansoor Hayat, Supavadee Aramvith, and Titipat Achakulvisut. Segsrnet for stereo-endoscopic image super-resolution and surgical instrument segmentation, 2024.

[47] Gengchen Mai, Ni Lao, Weiwei Sun, Yuchi Ma, Jiaming Song, Chenlin Meng, Hongxu Ma, Jinmeng Rao, Ziyuan Li, and Stefano Ermon. Ssif: Learning continuous image representation for spatial-spectral super-resolution, 2023.

[48] Guang Deng, Fernando J. Galetto, Mukhalad Al-nasrawi, and Waseem Waheed. A guided edge-aware smoothing-sharpening filter based on patch interpolation model and generalized gamma distribution, 2021.

[49] Xutong Ren, Lingxi Xie, Chen Wei, Siyuan Qiao, Chi Su, Jiaying Liu, Qi Tian, Elliot K. Fishman, and Alan L. Yuille. Generalized coarse-to-fine visual recognition with progressive training, 2019.

[50] Aroj Subedi. Improving generalization performance of yolov8 for camera trap object detection, 2024.

[51] Zihan Zang, Haoqiang Wang, and Yunpeng Xu. Dad vision: opto-electronic co-designed computer vision with division adjoint method, 2022.

[52] Ranjan Sapkota, Rizwan Qureshi, Marco Flores Calero, Chetan Badjugar, Upesh Nepal, Alwin Poulose, Peter Zeno, Uday Bhanu Prakash Vaddevolu, Sheheryar Khan, Maged Shoman, Hong Yan, and Manoj Karkee. Yolo11 to its genesis: A decadal and comprehensive review of the you only look once (yolo) series, 2025.

[53] Abdesselam Ferdi. Lightweight g-yolov11: Advancing efficient fracture detection in pediatric wrist x-rays, 2024.

[54] Kaiyue Liu, Qi Sun, Daming Sun, Mengduo Yang, and Nizhuan Wang. Underwater target detection based on improved yolov7, 2023.

[55] Matthew Ciolino, Dominick Hambrick, and David Noever. Enhancing satellite imagery using deep learning for the sensor to shooter timeline, 2022.

[56] Chris M. Ward, Josh Harguess, Brendan Crabb, and Shibin Parameswaran. Image quality assessment for determining efficacy and limitations of super-resolution convolutional neural network (srcnn), 2019.

[57] Ahmed ElSayed, Ausif Mahmood, and Tarek Sobh. Effect of super resolution on high dimensional features for unsupervised face recognition in the wild, 2017.

[58] Yihao Liu, Hengyuan Zhao, Jinjin Gu, Yu Qiao, and Chao Dong. Evaluating the generalization ability of super-resolution networks, 2023.

[59] Ling Zhou, Mingpei Wang, Xiaohua Huang, Wenming Zheng, Qirong Mao, and Guoying Zhao. An empirical study of super-resolution on low-resolution micro-expression recognition, 2023.

[60] Ben Lowe, Hadi Salman, and Justin Zhan. Ghm wavelet transform for deep image super resolution, 2022.

[61] Xiaohui Guo. A novel multi to single module for small object detection, 2023.

[62] Hailong Ma, Xiangxiang Chu, Bo Zhang, Shaohua Wan, and Bo Zhang. A matrix-in-matrix neural network for image super resolution, 2019.

[63] Venkateswararao Cherukuri, Tiantong Guo, Steve. J. Schiff, and Vishal Monga. Deep mr brain image super-resolution using spatio-structural priors, 2019.

[64] Bohong Chen, Mingbao Lin, Kekai Sheng, Mengdan Zhang, Peixian Chen, Ke Li, Liujuan Cao, and Rongrong Ji. Arm: Any-time super-resolution method, 2022.

[65] Yang-Che Sun, Cheng Yu Yeo, Ernie Chu, Jun-Cheng Chen, and Yu-Lun Liu. Fiper: Generalizable factorized fields for joint image compression and super-resolution, 2024.

[66] Parag S. Chandakkar and Baoxin Li. Joint regression and ranking for image enhancement, 2017.

[67] Zheng Zhan, Yifan Gong, Pu Zhao, Geng Yuan, Wei Niu, Yushu Wu, Tianyun Zhang, Malith Jayaweera, David Kaeli, Bin Ren, Xue Lin, and Yanzhi Wang. Achieving on-mobile real-time super-resolution with neural architecture and pruning search, 2023.

[68] Chunwei Tian, Xuanyu Zhang, Qi Zhang, Mingming Yang, and Zhaojie Ju. Image super-resolution via dynamic network, 2024.

[69] Longguang Wang, Yingqian Wang, Zaiping Lin, Jungang Yang, Wei An, and Yulan Guo. Learning a single network for scale-arbitrary super-resolution, 2021.

[70] Saem Park and Nojun Kwak. Image enhancement by recurrently-trained super-resolution network, 2019.

[71] Baran Ataman, Mert Seker, and David Mckee. Single image super-resolution, 2021.

[72] Xiaohui Zhang, Varun A. Kelkar, Jason Granstedt, Hua Li, and Mark A. Anastasio. Impact of deep learning-based image super-resolution on binary signal detection, 2021.

[73] Kwanyoung Kim and Se Young Chun. Sredgenet: Edge enhanced single image super resolution using dense edge detection network and feature merge network, 2018.

[74] Xinyue Wei, Haozhi Huang, Yujin Shi, Hongliang Yuan, Li Shen, and Jue Wang. End-to-end adaptive monte carlo denoising and super-resolution, 2021.

[75] Deborah Pereg. One-shot image restoration, 2024.

[76] Leheng Zhang, Yawei Li, Xingyu Zhou, Xiaorui Zhao, and Shuhang Gu. Transcending the limit of local window: Advanced super-resolution transformer with adaptive token dictionary, 2024.

[77] Mustafa Ayazoglu. Imdeception: Grouped information distilling super-resolution network, 2022.

[78] Ziyang Liu, Zhengguo Li, Xingming Wu, Zhong Liu, and Weihai Chen. Dsrgan: Detail prior-assisted perceptual single image super-resolution via generative adversarial networks, 2021.

[79] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement, 2020.

[80] Ram Krishna Pandey and A G Ramakrishnan. Language independent single document image super-resolution using cnn for improved recognition, 2017.

[81] Parichehr Behjati, Pau Rodriguez, Armin Mehri, Isabelle Hupont, Carles Fernández Tena, and Jordi Gonzalez. Hierarchical residual attention network for single image super-resolution, 2020.

[82] Cheng Wan, Hongyuan Yu, Zhiqi Li, Yihang Chen, Yajun Zou, Yuqing Liu, Xuanwu Yin, and Kunlong Zuo. Swift parameter-free attention network for efficient super-resolution, 2024.

[83] Keke Zhang, Tiesong Zhao, Weiling Chen, Yuzhen Niu, and Jinsong Hu. Spqe: Structure-and-perception-based quality evaluation for image super-resolution, 2022.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.