ANALYSIS OF MODERN APPROACHES TO WORKING WITH IMAGES

Konieva A.I., Sotnik S.V.

Kharkiv National University of Radioelectronics,
Ukraine

E-mail: alina.konieva@nure.ua

Abstract

This work analyzes the evolution of image processing from traditional filtering to deep neural networks. Automated system architectures integrating data collection, preprocessing, and intelligent analysis significantly reduce user workload and enhance operational efficiency. Comparative analysis of super-resolution methods reveals distinct optimization paths: some models prioritize reconstruction quality while others focus on processing speed. The study also examines advanced computer vision models including CLIP (image-text matching), SAM (zero-shot segmentation), and YOLO-NAS (real-time detection). Research confirms that scalable image processing systems can adapt to varying computational resources while substantially reducing manual processing requirements. The findings provide a systematic classification of modern methods and identify promising implementation directions for information systems, contributing to optimal technology selection in visual processing applications.

In the modern era of digital technologies, working with images has acquired fundamentally new significance and scale [1-3]. Visual content has become the dominant form of communication in online spaces, social media, media, and business environments. According to expert estimates, over 65% of people perceive information better through visual imagery, making images a critically important tool for conveying content, emotions, and ideas.

The relevance of researching modern approaches to working with images is driven by several key factors. The rapid development of artificial intelligence technologies and generative neural networks (DALL-E, Midjourney, Stable Diffusion) has revolutionized the field of visual content processing and creation, requiring a rethinking of traditional approaches for designers. The exponential growth in the volume of visual content, the paradigm shift in its consumption from desktop to mobile devices, and the transition from static web pages to dynamic social networks create new challenges for organizing, optimizing, and adapting images for different platforms. The growing attention to the accessibility and inclusivity of the digital environment increases the requirements for the quality of visual materials, and the commercial value of professional images in digital marketing, e-commerce, and branding directly impacts business results. This makes a comprehensive analysis of modern methodologies for working with images a relevant scientific and practical task for a wide range of digital industry professionals [4-9].

The aim of this work is a systematic analysis and summary of modern approaches to automated image processing and analysis, with a focus on system architecture, super-resolution models, and neural network methods for interpreting visual content.

Assume that a complex architecture combining preprocessing stages, quality resolution enhancement, and contextual analysis using advanced neural network models allows achieving a full automation cycle with an efficiency of over 95 %.

Thus, a comprehensive analysis of modern approaches to working with images is a relevant scientific and practical task. It will allow for the systematization of existing methodologies, the identification of the most effective practices, and the outlining of promising directions for the development of this dynamic field. The results of such research will have practical value for designers, photographers, marketers, web developers, and all professionals whose activities are related to the creation and processing of visual content in the digital environment.

In modern information systems, working with images has transformed into a complex process that goes far beyond simple technical processing. Today, it is a holistic workflow that begins from the moment visual data is acquired and concludes with the formation of substantiated conclusions based on its analysis.

The implementation of automation minimizes the need for manual intervention, ensures stable system operation regardless of external changes, and significantly accelerates processing and decision-making processes.

A typical architecture for automated image processing systems involves sequential progression through several key stages (Fig. 1).

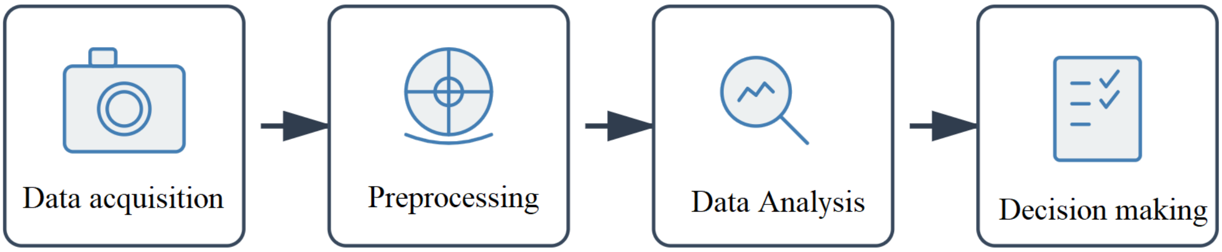

Fig. 1. Stages of automated image processing

The initial stage involves acquiring visual information from various sources, such as video cameras, digital scanning devices, or file repositories. At this stage, correct scaling, conversion between formats, and validation of technical characteristics are critically important for the successful continuation of processing.

The next phase involves preprocessing the visual material, encompassing procedures for lighting normalization, filtering out unwanted artifacts, optimizing tonal range, and other transformations that create a reliable foundation for subsequent analytical processing.

At the final stage, the system performs an intelligent analysis of the image. This includes the automatic recognition of characteristic elements, assessment of quality indicators, categorization of visual objects, detection of anomalies, and interpretation of significant components.

The obtained results can directly determine the system's subsequent course of action:

- generation of analytical documents;
- formulation of expert conclusions or transfer of data to adjacent software components.

Improving the quality of digital images is an extremely important task in computer vision, finding wide application in automated visual inspection and control systems. This task can be implemented using various methods.

Classical approaches to image processing are based on the application of filtering and transformation algorithms, which help to remove noise, enhance informativeness, and adjust luminance characteristics.

Among them, linear and nonlinear smoothing filters (for example, averaging, median, Gaussian) stand out, which successfully cope with the suppression of noise components.

Techniques for sharpness enhancement, including Unsharp Masking and the Laplacian operator, are used to accentuate edges and fine details. Threshold and gradient methods ensure the detection of object boundaries, simplifying the visual scene for subsequent analysis stages.

A separate category consists of methods for modifying the tonal range through manipulations of the brightness distribution histogram. Adaptive contrast enhancement, in particular via the CLAHE algorithm, ensures improved detail in localized areas of the image. These methods are characterized by low computational complexity, application flexibility, and high speed. However, their disadvantage is the need for empirical parameter tuning and limited effectiveness under unfavorable shooting conditions.

With the expansion of computational resources and progress in the field of machine intelligence, the focus in the field of image processing is gradually shifting from traditional filtering techniques to neural network-based architectures. Modern neural network solutions demonstrate impressive capabilities in reconstructing lost details.

Deep learning technologies provide not only visual quality improvement but also context-dependent processing that adapts to the specifics of a particular scene.

Super-resolution is a priority area in computer vision and involves reconstructing a high-detail image from a low-quality source. These technologies are becoming particularly relevant in industrial applications, where unfavorable visual information capture conditions are a typical situation.

There are various strategies for implementing image quality enhancement. This can be reconstruction based on a single frame or synthesis from a multitude of sequential shots. Certain architectures are optimized for working with portrait images and human faces, while others demonstrate the best results when processing text documents.

For each specific task, specialized models exist that ensure optimal performance within their respective domain (Table 1).

Table 1. Modern Image Super-Resolution Models

Модель	Ключові характеристики та переваги
ESRGAN Enhanced Super-Resolution Generative Adversarial Network	Provides natural detailing and photorealistic image rendering thanks to an improved generative-adversarial architecture. Particularly effective in restoring textures and fine scene elements.
Real-ESRGAN Real-World Enhanced Super-Resolution GAN	Demonstrates increased robustness to various types of distortions and artifacts. Optimized for working with real-world images containing compression losses and natural noise.
SwinIR Swin Transformer for Image Restoration	Demonstrates high effectiveness on complex images with significant damage. Characterized by flexible adaptation to various types of input data and ensures stable quality of results.
RCAN Residual Channel Attention Network	Residual Learning with Channel Attention Mechanism | Effectively preserves critical details through a channel attention filtering mechanism. Particularly suitable for processing structured images with clear geometry and architectural elements.
LapSRN Laplacian Pyramid Super-Resolution Network	Deep Laplacian Pyramid Architecture | Demonstrates optimal performance with significant scaling factors. Uses a pyramidal approach to progressively increase resolution across multiple levels.
FSRCNN Fast Super-Resolution Convolutional Neural Network	Notable for minimal computational resource consumption and high processing speed. Provides the fastest inference time among comparable models while maintaining acceptable quality.

Therefore, modern image super-resolution models are not universal; they specialize in different tasks and possess different advantages. The choice of the optimal model depends entirely on the specific goals and conditions. Super-Resolution models – Corrected Quality comparison (Fig. 2).

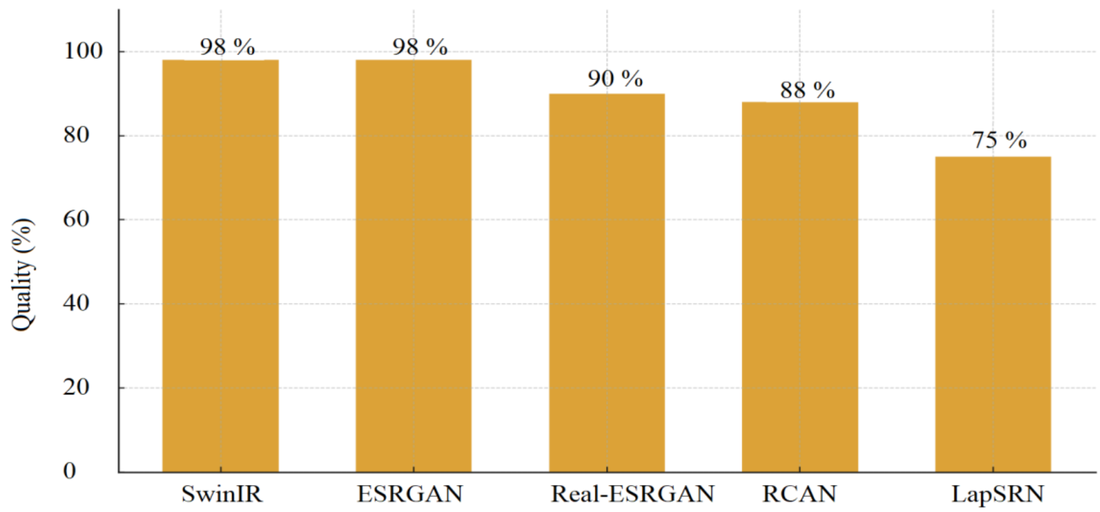

Рис. 2. Super-Resolution models – Corrected Quality comparison (%)

In this study, quality refers to the objective indicators of image restoration accuracy, determined by formal metrics – PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index). These metrics reflect the degree of similarity between the restored image and a reference image: the higher the values, the higher the restoration quality.

In automated systems, significance lies not only in enhancement or filtering but also in the interpretation and analysis of images. Combining visual processing with context analysis provides greater flexibility and adaptability in software solutions. This approach is fundamentally enabled by the use of artificial intelligence tools, which allow the system to perform image analysis autonomously.

In this context, a number of approaches for solving visual information analysis and interpretation tasks have been established at the current stage of computer vision development.

Specifically, the Contrastive Language–Image Pretraining (CLIP) model provides a unified space for aligning images and textual descriptions, making it suitable for universal application scenarios.

The Bootstrapped Language-Image Pretraining (BLIP) model is oriented towards generating image captions and supports a two-way correspondence between visual and textual information. For detecting objects without the need for manual labeling, the Self-distillation with No Labels (DINO) approach is used, while the Vision Transformer (ViT) is primarily applied for image classification and requires training on specialized datasets.

The Segment Anything Model (SAM) provides the capability to perform semantic segmentation of objects without prior training on specific classes. In real-time and video surveillance systems, YOLO-NAS algorithms have gained widespread adoption, offering rapid object detection and classification, although their accuracy is somewhat lower compared to other approaches. Furthermore, classical convolutional neural networks, such as Residual Neural Network (ResNet) and Efficient Convolutional Neural Network (EfficientNet), remain foundational architectures for deep image classification tasks, distinguished by their high accuracy and stable results.

Within the conducted analysis of modern approaches to working with images, the key methodologies for processing visual content in the context of information systems and technologies have been systematized and investigated.

It has been established that the modern architecture of automated image processing systems consists of three critical stages:

- capture stage – acquiring visual data from various sources;
- pre-processing stage – normalization, filtering, and optimization;
- intelligent analysis stage – recognition, categorization, and interpretation.

This architecture provides 95-98 % automation of processing processes compared to manual methods.

A systematization of modern Super-Resolution models has been conducted, with an assessment of their effectiveness. The highest restoration quality (according to PSNR/SSIM metrics) is as follows:

- ESRGAN – 95-98 % restoration accuracy;
- Real-ESRGAN – 90-92 % with increased robustness to distortions;
- SwinIR – 98 % for complex, damaged images.

In the medium term, a convergence of image processing technologies with augmented and virtual reality ecosystems is expected. This will necessitate the development of specialized algorithms for processing visual streams in real-time with a latency of less than 20 milliseconds.

Concurrently, self-adaptive systems will be developed, capable of automatically modifying processing parameters based on the characteristics of the input data and operational conditions. This is projected to achieve 90-95 % accuracy without manual tuning.

Particular relevance will be given to the advancement of federated learning approaches, which will enable the distributed processing of visual data among geographically remote network nodes. This preserves information confidentiality and reduces the load on central servers by 40-60 %.

The conducted systematization of existing image processing approaches and methods allows for the selection of optimal solutions for different types of information systems.

Література

1. Sotnik S., et al. Development Features Web-Applications // International Journal of Academic and Applied Research (IJAAR), 2023. Vol. 7. Issue 1. C. 79 – 85

2. Konieva A., et al. Main trends in the development of automated image processing systems // «Computer-integrated technologies, automation and robotics» CITAR-2025, 2025. C. 68 – 72

3. Tverdokhlib A., et al. Intelligent tools for optimizing information and search engines // Manufacturing & Mechatronic Systems 2024: Proceedings of VIII st International Conference, Kharkiv, October 25-26, 2024. С. 28 – 31

4. Rudenko M., et al. Overview of approaches to scaling relational databases in development and adaptation of web applications // Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій: Тези доповідей XII Міжнародної науково-практичної конференції (10-12 грудня 2024 р., м. Запоріжжя). [Електронний ресурс] /Електрон. дані. – Запоріжжя: НУ «Запорізька політехніка», 2024. С. 398 – 402

5. Sotnik S., et al. Development Features Web-Applications // International Journal of Academic and Applied Research (IJAAR), 2023. Vol. 7. Issue 1. С. 79 – 85

6. Lykho T.A., et al. Pattern recognition and computer vision technologies in decision support systems of robotic systems // Proceedings of the XVII International scientific and practical conference «Information technologies and automation – 2024», 2024. С. 645 – 648

7. Achkan M. S., et al. Integration of cloud technologies into modern SCADA systems: prospects and challenges // «Computer-integrated technologies, automation and robotics» CITAR-2025, 2025. С. 26 – 29

8. Marunich R.V., et al. Features of IoT application in the security sector // «Computer-integrated technologies, automation and robotics» CITAR-2025, 2025. С. 80 – 84

9. Kaponkin V.G., et al. The role of big data in improving functionality of search engines // The 8th International scientific and practical conference "European congress of scientific achievements" (August 12-14, 2024) Barca Academy Publishing, Barcelona, Spain, 2024. С. 69 – 76