2026 Volume: 02 Issue: 04 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-08

VISION INSIGHT-AN INTELLIGENT PHOTO EDITING AND ANALYSIS SYSTEM

***Pratyusha Valla, Jyothirmayi, B. Manasa, S. Bhanu Prakesh, Mr. Soubhagya Ranjan Nayak**

GMR Institute of Technology.

Article Received: 14 March 2026
Article Revised: 03 April 2026
Published on: 23 April 2026

***Corresponding Author: Pratyusha Valla**
GMR Institute of Technology.

DOI: https://doi-doi.org/101555/ijrpa.8418

1.ABSTRACT

This project presents an advanced AI-powered image editing and analysis platform that seamlessly integrates traditional editing tools with modern intelligent features into a unified and user-friendly system. The application allows users to perform essential image adjustments such as brightness, contrast, saturation, exposure, and color tuning, enabling precise control over image quality. Alongside these basic functionalities, the system offers powerful advanced features such as automatic background removal, which simplifies object isolation, and prompt-based image editing, where users can describe desired modifications using natural language and receive instant, AI-driven results. This significantly reduces manual effort and enhances accessibility for users with different levels of technical expertise.In addition to editing and generation capabilities, the platform incorporates an AI detection module that analyzes uploaded images and provides a percentage-based estimation of whether the content is AI-generated or real. This feature adds an important layer of transparency and helps users assess the authenticity of digital content in an era of increasing AI-generated media. The system is built on a scalable and efficient architecture using REST APIs, which ensure smooth communication between the frontend and backend components. This design enables the platform to handle multiple concurrent user requests while maintaining performance and reliability. Furthermore, the platform emphasizes usability, responsiveness, and real-time processing, offering an interactive editing experience with features such as live previews, edit history tracking, and easy export options. By combining creative editing tools, intelligent automation, and analytical capabilities within a single integrated environment, the project aims to simplify complex image processing tasks,

enhance productivity, and provide a comprehensive solution for modern digital image editing and analysis needs.

KEYWORDS: Photo Editing, Image Processing, Prompt-Based Image Generation, Web-Based Application, Purpose-Based Presets, Image Quality Evaluation, User-Friendly Interface.

2.INTRODUCTION

Digital images play a vital role in modern communication, education, and professional activities. They are widely used in academic projects, business presentations, social media, and creative work to convey information effectively and enhance visual appeal. As digital content continues to grow, the ability to edit and improve images has become an essential skill for students, professionals, and general users. Photo editing enables users to enhance image quality, correct imperfections, and apply creative transformations. Common operations such as cropping, rotating, and adjusting brightness or contrast are frequently required in everyday tasks, from preparing academic reports to creating engaging online content.

Despite its importance, photo editing remains challenging for many users. Existing tools often fall into two extremes: professional software with advanced features but complex interfaces, and simpler tools that lack flexibility and creative control. Applications like Adobe Photoshop and GIMP provide powerful editing capabilities, including layers, filters, and detailed adjustments, but they require technical knowledge and training. For beginners, understanding concepts such as color correction, exposure, and composition can be overwhelming, leading to time-consuming trial-and-error approaches with inconsistent results.

With the rise of web-based applications, image editing has become more accessible, allowing users to work directly in browsers without installing software. However, most of these tools still rely on manual editing and predefined filters, offering limited intelligent assistance or personalization. Recent advancements in artificial intelligence have opened new possibilities in this field. AI-powered tools can analyze images, automate complex editing tasks, and generate transformations based on user input. This evolution creates an opportunity to develop systems that combine ease of use with powerful features, enabling efficient, creative, and user-friendly image editing experiences.

3.LITERATURE REVIEW

Gunaseelan B. et al. (2024) proposed a multimodal-guided image editing system using text-to-image diffusion models. Their approach combined textual prompts and image inputs to perform advanced image transformations and improve editing flexibility. The system demonstrated the potential of diffusion models in generating high-quality edited images with multimodal guidance. However, the model faced challenges in preserving fine-grained details and lacked precise controllability, especially in maintaining temporal consistency across edits.Task-specific image enhancement techniques were explored by researchers (2021) to improve the accuracy of Convolutional Neural Networks (CNNs). The system automatically learned enhancement operations tailored to specific tasks, resulting in improved model performance and better feature extraction. This approach showed significant improvements in classification accuracy. However, it introduced higher computational overhead and faced scalability issues when applied to large datasets or real-time applications.

Data-Driven Adaptive History for image editing was introduced by researchers (2016), focusing on segmenting editing history into meaningful phases. This approach enabled intelligent undo operations and improved user interaction with editing workflows. The system was particularly effective in portrait retouching tasks, where editing steps are repetitive and structured. However, the model had limited generalization capabilities and struggled to adapt to diverse editing scenarios beyond portrait-based applications.The PyTorch Image Quality framework (2021) provided a standardized set of metrics for image quality assessment, including full-reference (FR), no-reference (NR), and distribution-based (DB) methods. This framework supported researchers in evaluating image quality across various applications and ensured consistency in performance measurement. However, the system focused only on evaluation and did not provide intelligent recommendations or editing support for improving image quality.An intelligent recommendation system based on image processing was proposed in 2020 to evaluate aesthetic quality, generate image descriptions, and provide recommendations. The system used image analysis techniques to guide users in selecting better visual outputs. While it improved decision-making, it did not support actual image editing functionalities and lacked real-time visual feedback. Additionally, the system was domain-specific and could not be generalized for broader applications.In 2024, researchers developed optimization techniques for image enhancement in AI-assisted artificial retinas. The system aimed to enhance primary visual objects and improve clarity using deep learning models, contributing to better visual perception for users. Although effective in medical

applications, the system was highly specialized and did not support user-interactive editing or prompt-based transformations, limiting its usability in general-purpose image editing.

EditEase (2023) introduced an AI-powered web-based image editing platform designed for non-technical users. The system provided accessible editing tools and leveraged AI for automating certain tasks, improving user experience and ease of use. However, it relied heavily on cloud-based APIs, leading to concerns regarding scalability, latency, and real-time performance, especially for high-resolution image processing.

Prompt-to-prompt image editing with cross-attention control (2022) enabled users to edit images using textual prompts while controlling specific regions of the image. This approach improved editing precision and allowed targeted transformations. Despite its effectiveness, the method required high computational resources, particularly for high-resolution images, and sometimes failed to accurately interpret complex user intentions expressed through prompts.

Multimodal generative models for text-to-image synthesis (2023) focused on generating realistic images from textual descriptions. The system improved the quality and realism of generated images and demonstrated the potential of generative AI in creative applications. However, it struggled with abstract or ambiguous prompts and was evaluated only on small-scale datasets, limiting its practical applicability.

Diffusion models for generative AI (2023) provided a comprehensive understanding of denoising diffusion probabilistic models (DDPMs) for realistic image generation. These models achieved high-quality outputs and became a foundation for modern image generation systems. However, they required significant computational resources, including high GPU memory and long training times, making deployment costly and challenging.

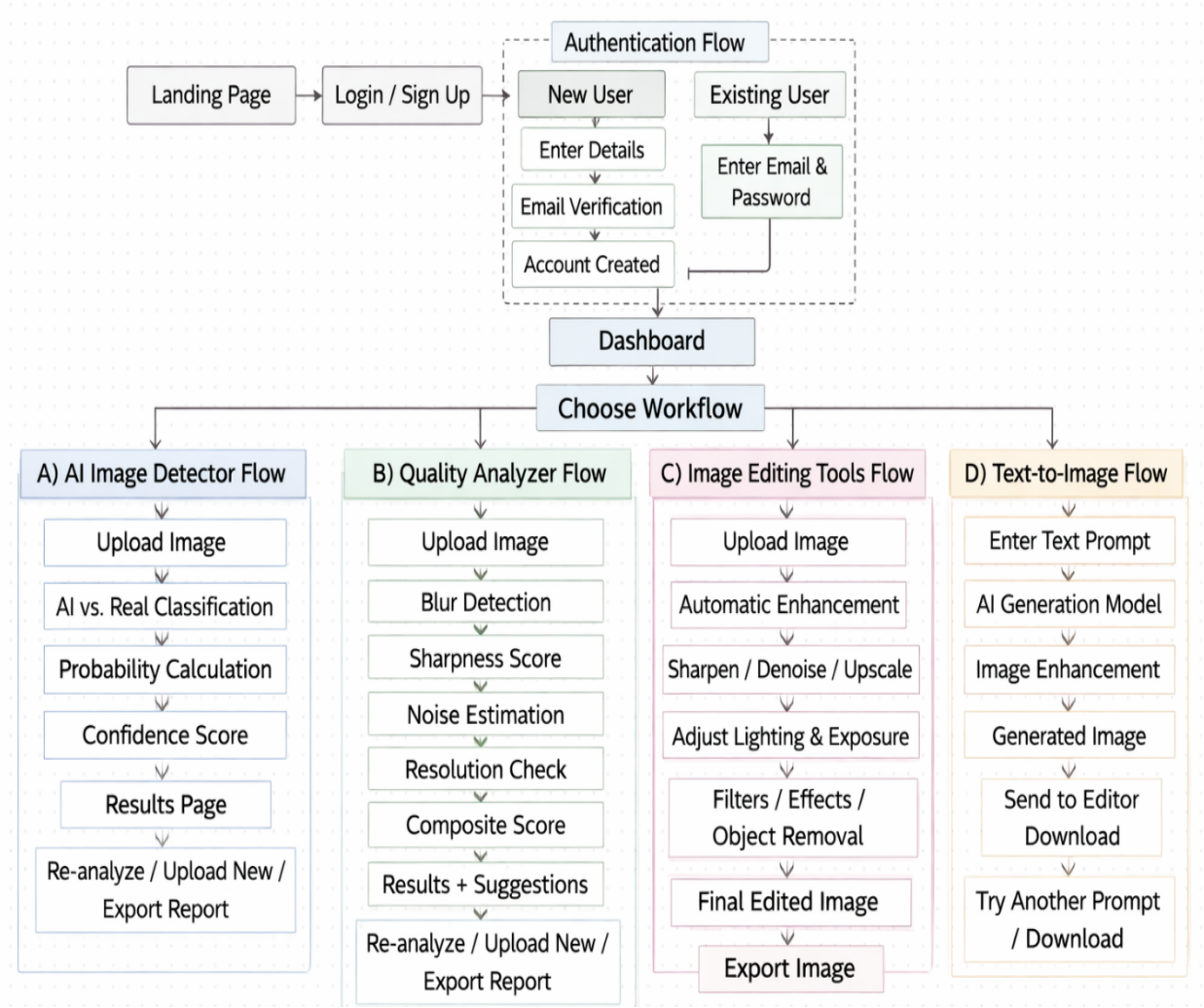
A systematic review on photo editing and body image concerns (2022) analyzed the psychological impact of photo editing tools on users. The study highlighted how excessive editing can influence body image perception and self-esteem. While the research provided valuable insights, it lacked intervention strategies and long-term studies to address these concerns effectively.

Prompt-to-prompt image editing techniques presented at SIGGRAPH (2022) further explored attention control mechanisms for editing images using textual prompts. The system demonstrated improved editing flexibility and user control over generated outputs. However, similar to other diffusion-based methods, it faced performance issues with high-resolution images and had limitations in capturing complex editing intentions accurately.

4.METHODOLOGY

System Workflow

The system follows a simple process where users upload an image, select the required module (editing, background removal, AI detection, or prompt-based editing), and receive the processed output. Users can preview, re-edit, and download the final image.



Main Modules:

Basic Image Editing Features

The system includes common editing functionalities such as brightness, contrast, saturation, cropping, and rotation. These tools help users improve image quality, adjust lighting, and enhance overall appearance easily without requiring technical skills.

Background Removal

This module automatically removes the background from an image using deep learning models like U²-Net. It separates the main subject from the background and provides a clean output with a transparent background, making it useful for design and editing purposes.

Prompt-Based Editing

The system allows users to edit images using natural language. Users can type instructions such as “change background” or “apply sketch effect,” and the AI will understand and apply the changes automatically. This reduces manual effort and makes editing more creative and user-friendly.

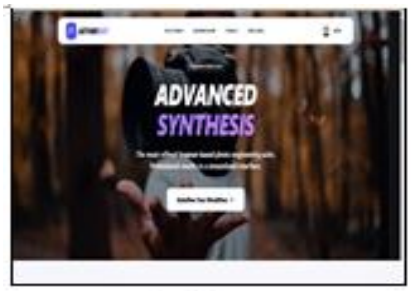
AI Image Detection

The AI detection module analyzes uploaded images and determines whether they are AI-generated or real. It provides a percentage result indicating how much of the image is likely generated by AI, helping users understand the authenticity of the image.

5.RESULTS AND DISCUSSIONS

The Vision Insight system was successfully developed as a comprehensive web-based platform that integrates image editing, analysis, and AI-driven image generation into a single, unified environment. The system was evaluated using a diverse set of images that varied in resolution, quality, lighting conditions, and content complexity. This extensive testing ensured that the system’s performance could be measured across real-world scenarios. The results demonstrate that the system operates reliably and consistently, producing stable outputs regardless of input variations, thereby confirming its robustness and adaptability.

The image editing module performed efficiently across all tested scenarios. Basic editing operations such as cropping, rotation, resizing, and adjustments to brightness, contrast, and saturation were executed with high accuracy and minimal processing delay. The inclusion of real-time preview functionality significantly enhanced the usability of the system. Users were able to instantly observe the effects of applied changes, which reduced the dependency on repetitive trial-and-error methods. This not only improved editing speed but also made the system more intuitive and accessible, particularly for users with limited technical expertise.

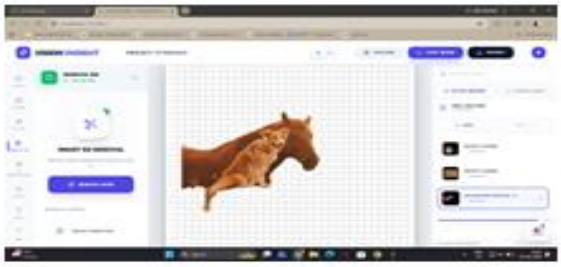


Landing page

Project page

Edit History

Prompt-Based Editing



Background image removal

5.CONCLUSION

The Vision Insight system successfully fulfills its primary objective of developing a comprehensive, user-friendly, web-based platform for intelligent photo editing and analysis. By effectively integrating traditional image processing techniques with modern artificial intelligence methods, the system provides a unified environment that supports both fundamental editing operations and advanced AI-driven transformations. This integration demonstrates how emerging technologies can enhance the capabilities of conventional tools while maintaining simplicity and accessibility for users.

A key contribution of this project is the implementation of text-based image editing and generation, which enables users to modify images through natural language instructions. This feature represents a significant shift from traditional manual editing approaches, as it reduces complexity and eliminates the need for in-depth technical knowledge. Users can perform sophisticated transformations such as style conversion, enhancement, and creative filtering simply by providing descriptive prompts. This not only improves efficiency but also encourages creativity by allowing users to experiment with different visual styles effortlessly.

6.REFERENCES

1. X. Shuai, H. Ding, X. Ma, R. Tu, Y.-G. Jiang, and D. Tao, "Multimodal-Guided Image Editing with Text-to-Image Diffusion Models," *arXiv preprint arXiv:2406.14555*, 2024.
2. N. Mitschke, Y. Ji, and M. Heinzmann, "Task-Specific Image Enhancement for Improving the Accuracy of CNNs," in *Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, 2021, pp. 174–181.
3. S. B. Kang, A. Agarwala, et al., "Data-Driven Adaptive History for Image Editing," *ACM Transactions on Graphics (SIGGRAPH)*, 2016.
4. E. Koutelidakis and D. M. Zarpalas, "PyTorch Image Quality: Metrics for Image Quality Assessment," *IEEE Access*, vol. 9, pp. 153036–153050, 2021.
5. H. Zhang et al., "Intelligent Recommendation System Based on Image Processing," *Journal of Physics: Conference Series*, vol. 1544, 2020.
6. A. Mehmood, J. Ko, H. Kim, and J. Kim, "Optimising Image Enhancement: Feature Engineering for Improved Classification in AI-Assisted Artificial Retinas," *Sensors*, vol. 24, no. 9, p. 2678, 2024.
7. H. Agarwal, A. Bhardwaj, P. Jain, and R. Upadhyay, "EditEase: AI-Powered Web-Based Image Editing for Accessible, Intuitive Photo Transformation," *JECRC University*, 2023.
8. A. Hertz, R. Mokady, J. Tenenbaum, K. Aberman, Y. Pritch, and D. Cohen-Or, "Prompt-to-Prompt Image Editing with Cross-Attention Control," *ACM Transactions on Graphics (SIGGRAPH)*, 2022.
9. N. K. Htwe and W. P. Pa, "Multimodal Generative Model Based Text-to-Image Synthesis," *University of Computer Studies, Yangon*, 2023.
10. C. F. Higham, D. J. Higham, and P. Grindrod, "Diffusion Models for Generative Artificial Intelligence: An Introduction for Applied Mathematicians," *arXiv preprint arXiv:2312.14977*, 2023.
11. P. Ozimek, S. Lainas, H.-W. Bierhoff, and E. Rohmann, "How Photo Editing in Social Media Shapes Self-Perceived Attractiveness and Self-Esteem," *Computers in Human Behavior*, Elsevier, 2023.
12. A. Hertz, R. Mokady, J. Tenenbaum, K. Aberman, Y. Pritch, and D. Cohen-Or, "Prompt-to-Prompt Image Editing with Cross-Attention Control," *ACM Transactions on Graphics (SIGGRAPH)*, 2022.