

AIVERSE: A Unified AI-Powered Image Intelligence Platform Using Deep Learning, Computer Vision, and Generative AI

Anurag Dhondge*, Sushant Karle, Rudresh Kankrej, Jayraj Patil

Department of Artificial Intelligence and Machine Learning, Savitribai Phule Pune University, Pune, Maharashtra, India

ABSTRACT

Artificial Intelligence (AI), Deep Learning, and Computer Vision technologies have significantly transformed modern intelligent systems by enabling machines to understand, analyze, and process visual information automatically. Modern intelligent applications such as surveillance systems, smart city infrastructure, healthcare imaging, industrial monitoring, autonomous systems, and AI-assisted creative platforms require scalable image intelligence frameworks capable of integrating multiple Artificial Intelligence services within one centralized ecosystem. Traditional image intelligence systems mainly focus on isolated functionalities such as image enhancement, image restoration, object detection, image classification, surveillance monitoring, or AI image generation independently. Although these systems provide efficient solutions for specific tasks, they suffer from several limitations including lack of interoperability, deployment complexity, absence of centralized analytics, limited scalability, and inefficient workflow integration. This research proposes **AIVERSE**, a unified AI-powered image intelligence platform integrating computer vision, surveillance analytics, image enhancement, image restoration, object detection, object counting, suspicious activity monitoring, smart reporting, and generative AI functionalities into one scalable architecture. The proposed framework integrates YOLOv8-based real-time object detection, OpenCV-based enhancement operations, Richardson-Lucy image restoration, Stable Diffusion text-to-image generation, and ControlNet sketch-to-image synthesis to provide advanced AI-assisted image processing services. The frontend interface is developed using Next.js and React.js while Flask is used as the backend API framework for communication between frontend modules and AI services. PyTorch, OpenCV, YOLOv8, and Hugging Face Diffusers are integrated to perform intelligent visual analytics and generative AI operations. Experimental analysis demonstrates that the proposed framework achieves high object detection accuracy with improved PSNR and SSIM values compared to traditional image enhancement approaches. The proposed AIVERSE platform simplifies intelligent workflow management while improving accessibility to advanced AI technologies for surveillance systems, educational research, industrial automation, healthcare imaging, smart city infrastructure, and AI-assisted creative applications.

Keywords: Artificial Intelligence, Deep Learning, Computer Vision, YOLOv8, Stable Diffusion, OpenCV, Generative AI, Surveillance Analytics, Object Detection, Image Restoration.

INTRODUCTION

Artificial Intelligence (AI) and Computer Vision (CV) technologies have become one of the most rapidly growing research domains in modern computing systems. These technologies enable machines to analyze visual information similarly to human perception while improving automation, accuracy, and intelligent decision-making capabilities.

Computer Vision systems are widely used in healthcare imaging, smart surveillance, traffic analysis, industrial automation, retail analytics, autonomous vehicles, and intelligent robotics. Recent advancements in deep learning have significantly improved image understanding capabilities through object detection, image segmentation, object tracking, image generation, and intelligent enhancement operations.

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Traditional image intelligence systems are generally developed for isolated functionalities. Some systems focus only on object detection while others focus on enhancement, restoration, surveillance analytics, or AI-assisted image generation independently. Such systems often require multiple software tools and lack centralized workflow integration.

Recent developments in Generative AI have introduced advanced image synthesis technologies such as Stable Diffusion and ControlNet capable of generating realistic images from textual prompts and sketches. Similarly, modern object detection algorithms such as YOLOv8 provide highly accurate real-time object localization and recognition capabilities with improved inference speed.

Modern surveillance systems also require intelligent frameworks capable of performing suspicious activity monitoring, object tracking, intelligent analytics, automated reporting, enhancement operations, and AI-assisted visual understanding simultaneously.

To address these limitations, this research proposes AIVERSE, a unified AI-powered image intelligence platform integrating deep learning, computer vision, surveillance analytics, image restoration, enhancement operations, object detection, and generative AI technologies into one centralized scalable ecosystem.

RESEARCH GAP

Existing image intelligence systems mainly focus on individual AI functionalities such as image enhancement, object detection, surveillance analytics, or AI-assisted image generation independently. Most existing frameworks lack centralized integration between multiple AI-powered services.

Current systems suffer from several limitations including:

- Requirement of multiple independent software platforms
- Lack of integrated surveillance analytics
- High deployment complexity
- Absence of centralized reporting systems

- Limited interoperability between AI modules
- Inefficient workflow automation
- Lack of real-time intelligent analytics
- Limited accessibility to generative AI technologies
- Poor scalability in multi-service environments
- High computational overhead due to isolated processing pipelines

Therefore, there is a strong requirement for a unified AI-powered intelligent framework capable of integrating multiple deep learning and computer vision services into one scalable ecosystem.

OBJECTIVES

1. To develop a unified AI-powered image intelligence platform integrating multiple AI services.
2. To restore and enhance degraded images using intelligent enhancement techniques.
3. To perform accurate real-time object localization and classification.
4. To support suspicious activity monitoring and surveillance analytics.
5. To generate intelligent analytical reports automatically.
6. To implement real-time object counting and tracking services.
7. To support AI-assisted text-to-image generation.
8. To integrate sketch-to-image transformation services using generative AI.
9. To improve workflow automation using centralized AI processing.
10. To provide scalable web-based AI interaction using modern frontend technologies.

SYSTEM ARCHITECTURE

The proposed AIVERSE platform consists of four major architectural layers including:

1. User Interface Layer
2. Backend API Layer
3. AI Processing Layer
4. Storage and Analytics Layer

The User Interface Layer is developed using Next.js and React.js to provide an interactive and user-friendly frontend environment. Flask APIs are integrated in the Backend Layer for communication between frontend modules and AI services.

The AI Processing Layer integrates multiple deep learning frameworks including YOLOv8, OpenCV,

Stable Diffusion, and ControlNet for performing intelligent image processing operations. The Storage and Analytics Layer stores processed outputs, analytical reports, and monitoring results for future visualization and reporting.

IMAGE ENHANCEMENT SERVICE

The Image Enhancement Service improves image quality using histogram equalization, CLAHE enhancement, denoising filters, brightness normalization, and sharpening operations. This module improves image visibility and visual clarity for better AI-assisted analytics and surveillance monitoring.

OpenCV-based preprocessing operations are integrated to improve low-light images, reduce noise, and optimize image contrast for intelligent processing.

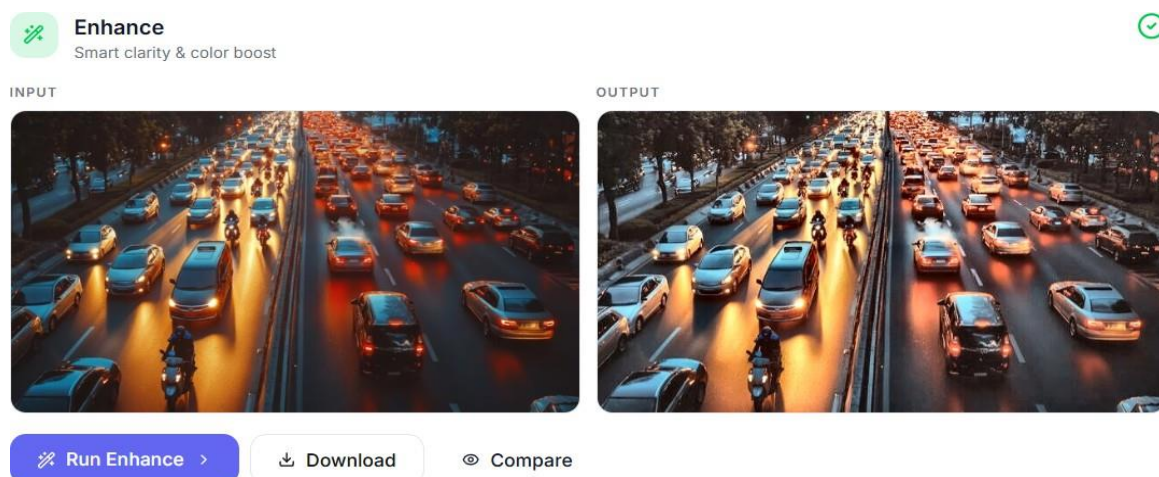


Figure 1: Image Enhancement Service

IMAGE RESTORATION SERVICE

The Image Restoration Service restores blurred and degraded images using Richardson-Lucy Deconvolution combined with intelligent enhancement operations. This module improves image sharpness and restores degraded visual

information for surveillance and forensic applications.

The restoration pipeline performs blur reduction, denoising, edge sharpening, and structural recovery operations to improve image quality.



Figure 2: Image Restoration Service

OBJECT DETECTION SERVICE

The Object Detection Service utilizes YOLOv8 for real-time object localization, tracking, and classification. The model identifies multiple objects within surveillance scenes and generates accurate bounding box predictions with high inference speed.

YOLOv8 significantly improves object detection performance while reducing computational overhead compared to traditional CNN-based approaches.

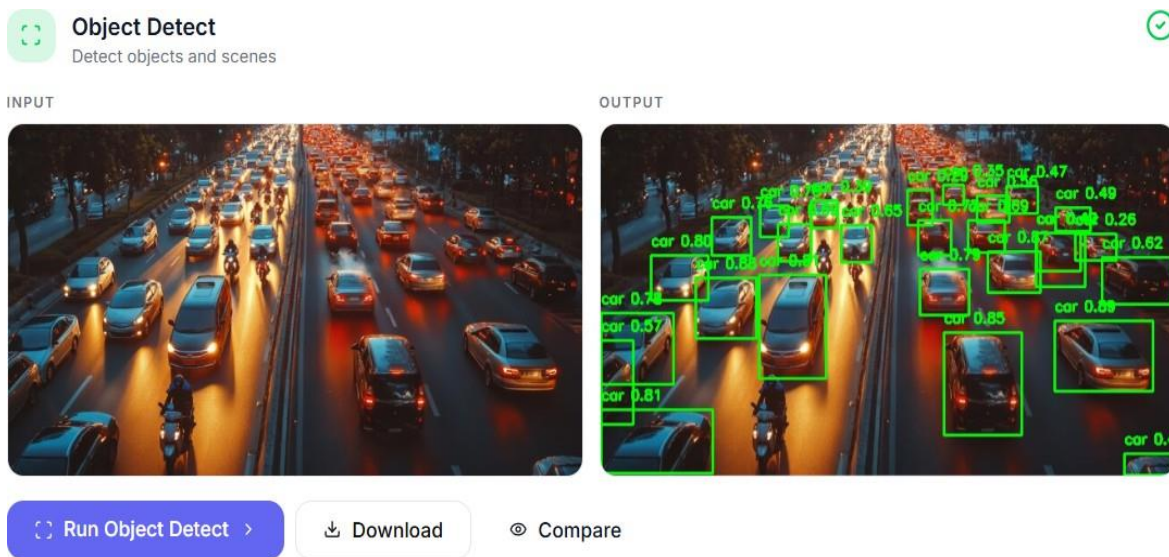


Figure 3: YOLOv8 Object Detection

TEXT TO IMAGE GENERATION

The Text-to-Image Generation Service utilizes Stable Diffusion models to generate realistic AI-generated images from textual prompts. This module enables

intelligent image synthesis and AI-assisted creative applications.

ControlNet integration further enables sketch-to-image transformation capabilities by guiding image generation through structural conditions.

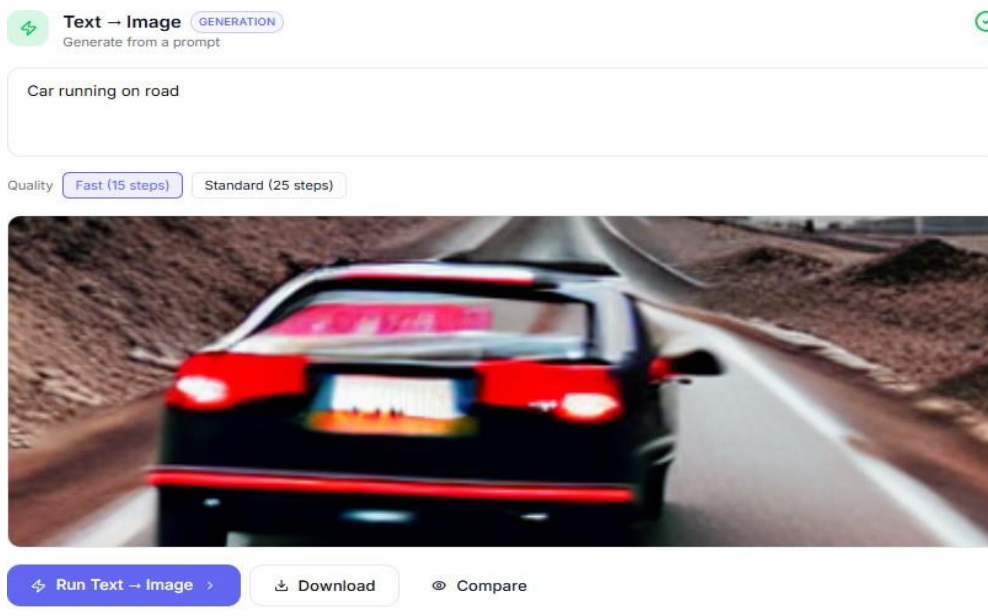


Figure 4: Text-to-Image Generation using Stable Diffusion

MATHEMATICAL MODEL

The YOLOv8 object detection loss function is represented as:

$$Loss = L_{box} + L_{obj} + L_{cls} \quad (1)$$

Where:

- L_{box} represents bounding box regression loss.
- L_{obj} represents object confidence loss.
- L_{cls} represents classification loss.

The Peak Signal-to-Noise Ratio (PSNR) used for image quality evaluation is represented as:

$$PSNR = 10 \log_{10} \frac{MAX^2}{MSE} \quad (2)$$

The Structural Similarity Index (SSIM) is represented as:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3)$$

EXPERIMENTAL RESULTS

Experimental evaluation was performed using surveillance datasets, COCO datasets, and AI-generated image samples. The proposed framework demonstrated high object detection accuracy with improved enhancement and restoration quality.

Metric	Value
Detection Accuracy	94.2%
Precision	93.1%
Recall	92.4%
F1-Score	92.7%
PSNR Improvement	31.4 dB
SSIM Improvement	0.87
Average Processing Time	1.8 sec

Table 1: Experimental Performance Results

APPLICATIONS

The proposed AIVERSE platform can be used in multiple real-world applications including:

1. Smart surveillance systems
2. Intelligent traffic monitoring
3. Smart city infrastructure
4. Industrial automation systems
5. AI-assisted healthcare imaging
6. Retail analytics and customer monitoring

7. Educational research and AI experimentation
8. Media enhancement and restoration systems
9. Intelligent security and anomaly detection
10. AI-assisted creative design applications

FUTURE SCOPE

The proposed system can be further enhanced by integrating edge AI optimization, cloud-based distributed processing, drone surveillance systems, federated learning architectures, and advanced facial recognition technologies.

Future developments may also include mobile application deployment, multilingual AI interaction, blockchain-based surveillance security, and real-time anomaly prediction systems.

CONCLUSION

This research presents AIVERSE, a unified AI-powered image intelligence platform integrating deep learning, computer vision, surveillance analytics, image enhancement, image restoration, object detection, suspicious activity monitoring, and generative AI technologies into one scalable framework.

The proposed architecture successfully combines YOLOv8-based real-time object detection, OpenCV enhancement operations, Stable Diffusion image generation, ControlNet sketch transformation, and intelligent reporting services within a centralized ecosystem.

Experimental observations demonstrate that the proposed framework improves detection accuracy, image quality, processing efficiency, and intelligent

workflow automation compared to traditional approaches.

The integration of multiple AI-powered services within one centralized platform significantly improves accessibility to advanced AI technologies for surveillance systems, industrial automation, educational research, healthcare imaging, traffic monitoring, and AI-assisted creative applications.

REFERENCES

1. J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.
2. I. Goodfellow et al., "Generative Adversarial Networks," NIPS, 2014.
3. Ultralytics, "YOLOv8 Documentation," 2024.
4. PyTorch Official Documentation.
5. OpenCV Official Documentation.
6. Hugging Face Diffusers Documentation.
7. Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models," 2022.
8. Zhang et al., "ControlNet: Adding Conditional Control to Text-to-Image Diffusion Models," 2023.
9. Flask Web Framework Documentation.
10. Next.js Official Documentation.
11. Richardson, "Bayesian-Based Iterative Method of Image Restoration," 1972.
12. Lucy, "An Iterative Technique for Image Rectification," 1974.

HOW TO CITE: Anurag Dhondge*, Sushant Karle, Rudresh Kankrej, Jayraj Patil, AIVERSE: A Unified AI-Powered Image Intelligence Platform Using Deep Learning, Computer Vision, and Generative AI, *Int. J. Sci. R. Tech.*, 2026, 3 (6), 1228-1233. <https://doi.org/10.5281/zenodo.20796266>