Journal of Engineering, Project, and Production Management
2026, 16(4), 2025-252

AI-Assisted Automation of Post-Production Compositing Workflows in Film and Television

Qiongfei You[1] and Xiaohang Zhang[2]
[1] Lecturer, Luoyang Vocational College of Science and Technology, Luoyang, 471822, China, E-mail: violet.feng888@outlook.com (corresponding author).
[2] Lecturer, Luoyang Vocational College of Science and Technology, Luoyang, 471822, China

Production Management
Received November 3, 2025; revised April 9, 2026; accepted May 2, 2026
Available online May 29, 2026

Abstract: The purpose of this study is to explore efficiency limitations within the post-production process in terms of compositing in films and TV series through an approach that combines AI with automation to optimize processing and image quality. Mixed methods were used in an experiment that involved 24 professional compositors analyzing 150 video clips using an AI-based algorithm featuring Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and transformers. This algorithm was trained using 50,000 professional sequences of composite images. Tests were carried out using commercial algorithms on a standard workstation. Results were obtained that showed considerable enhancements in terms of efficiency for all measures: time saving of 41.7% (processing time of 78.5±12.3 minutes), Peak Signal-to-Noise Ratio (PSNR) increase of 9.7% resulting in 35.2±0.8 dB, and Structural Similarity Index Measure (SSIM) enhancement by 8.0% to 0. This platform maintained real-time performance at 18.4 frames per second while providing significantly better visual output than any currently available commercial system. Overall user satisfaction stood at 8.4/10, and 92% of all artists involved in the study used the software within a week of the following training. This AI-driven platform proves that the use of automation can improve both technical and artistic performance within professional compositing processes.

Keywords: Artificial intelligence, deep learning, film production, post-production compositing, workflow automation.

DOI 10.32738/JEPPM-2025-252

1. Introduction

The compositing process might include thousands of layers on each composite image (Uddin et al., 2025). This growing complexity has led to several bottlenecks within post-production processes. The task of manual compositing is both cumbersome and prone to mistakes, thus restricting creativity while adding to production expenses (Hemraj, 2025). The issue also becomes more complex due to the business's increasing need for high-resolution content production, shorter time frames, and greater demand for streamed content in mixed formats (Pan et al., 2025).

Machine-learning algorithms proved to be unmatched at automating advanced visualization tasks that had required significant technical skills in humans (Wang, 2025). Advancements in deep learning algorithms have made impressive progress in compositing and real-time rendering technology. Neural compositing systems can integrate digital elements within physical spaces with accurate lighting and shadows (Ma et al., 2021). The practical use of machine learning in animation and visual effects has grown tremendously. Artificial intelligence-powered tools have become popular in the industry, helping increase efficiency and creativity (Botello et al., 2024). The role of AI in shaping the broader design community and the methodologies adopted for problem-solving and optimization in design work has been remarkable (Li et al., 2024).

Foundation work on the theory of AI-automation visual processing has been mostly driven by landmark work stemming from breakthroughs in computer vision and deep learning design. Systematic reviews of new methodologies conceived key usage situations where AI could remediate common vulnerabilities of image processing, perception of objects, and understanding of scenes (Chai et al., 2021). Systematic search on fundamentals of deep learning, i.e., design of Convolutional Neural Networks (CNNs), has established sound blueprints for handling imprecise tasks of vision and unlocking future foundations for research (Alzubaidi et al., 2021). In addition, senior work on compressing videos and images with neural networks has indicated that AI systems could curtail data processing workflows while maintaining high visual quality (Ma et al., 2020). The technological foundations thus created fertile grounds for exploring AI applications specifically for film-making scenarios (Sun, 2024).

Studies have begun to examine the compatibility of automation and creativity in AI-augmented film editing. Such studies analyze the advantages of technology on the one hand and the issues of maintaining artistic integrity on the other (Sora-Domenjó, 2024). Future filmmaking functions using AI technology have not only underscored the revolutionary nature of AI but also required the creation of integration mechanisms (Pradeep et al., 2023). Virtual production with the integration of AI and Extended Reality (XR) appears to be promising. These two technologies offer new avenues through which proven production techniques may be developed and validated (Ewis et al., 2024). Several experiments conducted on the use of AI in animating content production reveal that there have been considerable productivity increases, implying that the same can happen with compositing (Chen et al., 2024). The professional media and recent SIGGRAPH panels dedicated to AI for creative industries further prove this point (Burbano and Reiser, 2024).

Despite the promising research, there is still much to learn regarding optimal use of AI in compositor workflows. Although AI technology has proven its potential in boosting creative innovations and sustainability in design (Agboola, 2024), there are still no clear frameworks for how AI can be used collaboratively with humans during art making (Cetinic and She, 2022). However, the theoretical basis of deep learning methods, which is already fairly well developed when applied to generic cases (Mathew et al., 2020), still requires modification to meet the needs of film compositing, with artistic considerations also taken into account (Mahadevkar et al., 2022). Though promising, computer vision applications for film analysis have mostly been used in an analytical capacity, with little progress toward creative uses (Schmidt et al., 2021). Industrial workflows have used AI technology in a successful manner within the manufacturing process (Sardis and Varvarigou, 2010), and computer vision applications have been successful in many different settings (Salvador et al., 2011). Still, for application within creative fields, substantial methodological advances must be made (Li and Zhang, 2020).

Further problems associated with reproducing and standardizing AI-powered creative processes are to be considered for the successful implementation of such technology in industry settings (Sethi and Gil, 2016). Although machine learning models show universal efficiency in computer vision applications (Khan et al., 2021), the special requirements regarding real-time quality assessment and interaction with other processes of the production cycle need a more focused approach. Modern intelligent quality control systems can be seen as the answer to those challenges through design approaches based on users experience (Lönnqvist 2025).

In response to these gaps, this research designs and prototypes an AI-enhanced automated process for compositing applications in the television and film industry. This study is a contribution to the fields of AI and the creative industries, as it empirically demonstrates AI's capacity to increase efficiency and creativity in commercially relevant compositing environments. Through careful analysis of workflow optimization and quality metrics, as well as user experience factors, this study provides critical insights into integrating AI technology responsibly into creative post-production processes.

2. Methodology

2.1. Research Design and Experimental Setup

This study employs a mixed-method experimental design to compare AI-assisted compositing workflows with traditional manual processes through both quantitative performance metrics and qualitative user experience assessments.

The experiment setup was implemented within the standard industrial hardware settings that include powerful computer stations that have NVIDIA RTX 4090 professional graphics card, 64 GB memory and professional storage. Adobe After Effects and Blackmagic DaVinci Resolve were selected as the main software platforms because of their wide implementation, comprehensive functionality, and advanced Application Programming Interface (API) capabilities to integrate AI modules. AI modules were built within the standard development environment using popular Python frameworks, including TensorFlow and PyTorch. It includes a data set that consists of 150 professionally shot video samples from different genres, including dramatic stories, documentaries, and commercials with diverse levels of visual effects complexity, from simple color correction and compositing to more complex effects that include integration of multiple layers of visual effects.

The experimental workflow consists of a systematic five-step process: input data preparation, standardization, randomization of assignment distribution, workflow execution, and performance assessment. A schematic diagram describing five stages in the workflow of the experiments is represented in Fig. 1 below. There is a direct branching of the workflow at the assignments stage: either repeat video clips are assigned to experimenters by hand, or, using AI-supported compositing workflows, their performance is compared objectively. The schedule for parallel processing permits the assignment and execution of both workflows in parallel within the same experiment environment, and 75 videos are assigned in both directions to 24 professional operators. Assessment of performance includes performance results of both workflows regarding processing efficiency and quality, PSNR and SSIM parameters of images, user satisfaction survey, and error rate analysis.

Selection of participants included having a minimum of two years of experience as a compositor in addition to being actively employed in the industry within 12 months, proficiency with industry-standard software, and a lack of familiarity with the experimental AI model used. The study featured a sample size of 24 professional compositor participants who had experience ranging from two to 15 years and had been divided into three groups: junior, intermediate, and senior artists.

The sample consisted of 14 males and 10 females aged between 24 and 42 years (M = 31.5, SD = 5.2) selected from three key centers producing video content for cinema and TV channels (Los Angeles, London, Vancouver) specializing in different types of video content creation, such as feature films, television, and commercials.

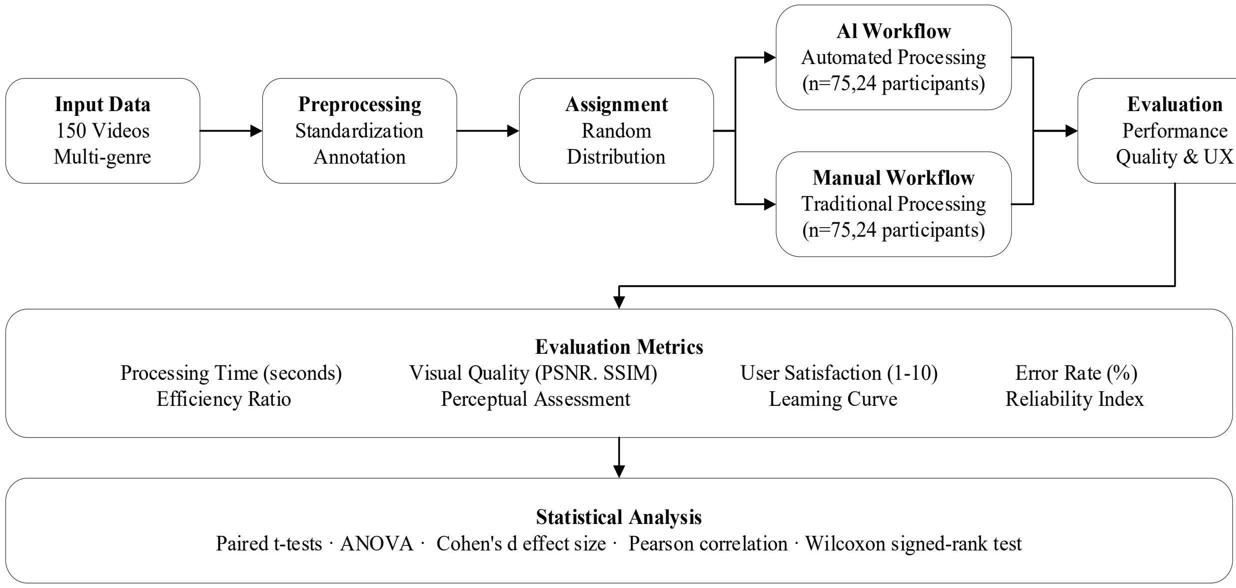


Fig. 1. Experimental workflow design

2.2. AI Model Development and Training

The architecture of the system is designed as a multi-layer neural network design employing CNNs to perform operations on images, along with RNN layers that impose consistency over time, as well as transformers for attention mechanism-based contextualization of compositing demand.

The architectural units play different roles in the process of compositing. In this respect, CNNs extract spatial features within each frame, such as edges, object segmentation, color borders, and texture analysis, to successfully create masks and isolate elements from the background. Recurrent Neural Network (RNN) elements, such as LSTM networks, keep consistency in time across video frames by establishing dependencies between frames, maintaining consistent motions, and predicting sequence patterns for the removal of any flickering effects seen with per-frame analysis. Through attention mechanisms in Transformers, the model has a comprehensive understanding of the context because it takes into account long-term dependencies throughout the whole sequence, allowing the network to identify compositionality between different layers, manage its priorities according to the complexity of the scene, and choose appropriate blending methods based on the entire narrative context. As depicted in Fig. 2, the operations of these modules are conducted through an ordered process where CNN outputs are used as inputs to LSTM, while the attention mechanism of the transformer is responsible for directing the entire composition process, as shown in Eq. (1).

$$Attention(Q,K,V) = softmax(QK^T / \sqrt{d_k})V \tag{1}$$

Where Q, K, and V represent query, key, and value matrices, respectively, and d_k denotes the dimension of the key vectors.

The training set consisted of more than 50,000 professionally composited videos captured from professional post-production studios, which had varying genres and complexities. The database was further enriched using the generation of synthetic data to generate more training samples, which varied in lighting conditions, placements of objects, and other compositing needs. Professional compositing artists provided detailed annotations for layer hierarchies, blending modes, and quality preferences, creating a comprehensive knowledge base that informs the AI system's decision-making processes.

The learning procedure adopted multi-object loss functions, including visual quality, perceptual loss, motion coherence consistency loss over time, and quality-dependent loss of conformity to professional standards, respectively defined in Eq. (2).

$$L = \alpha \cdot L_{perceptual} + \beta \cdot L_{temporal} + \gamma \cdot L_{quality} \tag{2}$$

Optimization of the employed model made use of adaptive learning rates of the Adam optimizer with an initial learning rate of 0.001 and then lowered using a schedule of cosine annealing over 100 epochs. The training took place over communal GPU clusters using systems from NVIDIA A100 with batch sizes of 16 sequences and gradient accumulation steps for reasons of convergence being achieved stably. The optimization target is the expected loss, as shown in Eq. (3).

$$\theta^* = \arg\min_\theta E[L(f_\theta(x), y)] \quad (3)$$

Where f_θ represents the parameterized model and (x, y) denotes input-output pairs.

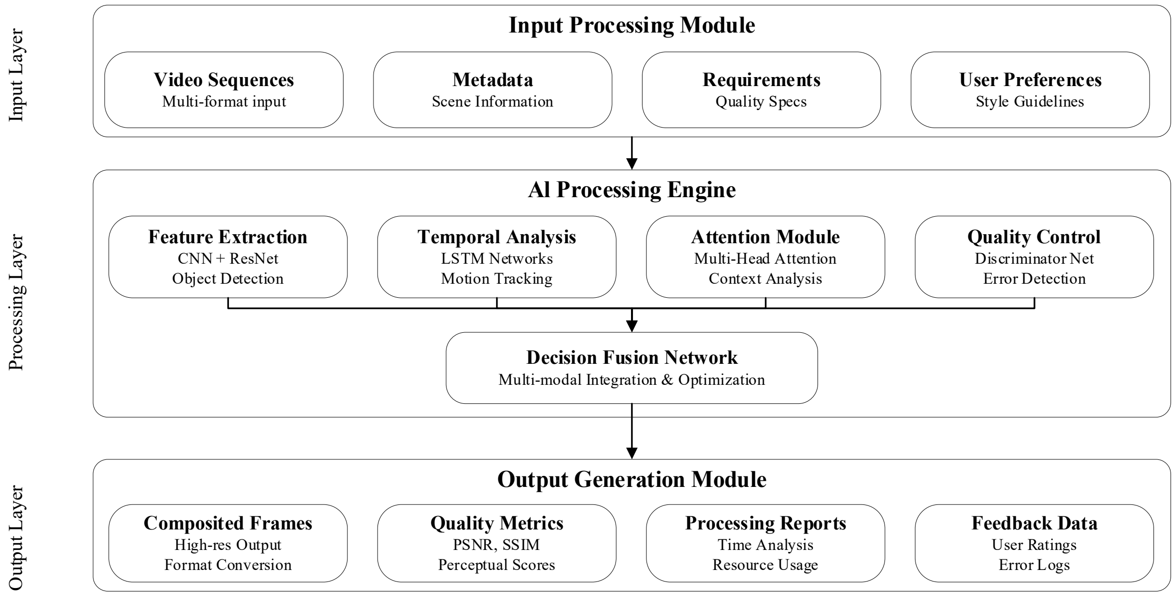


Fig. 2. AI-assisted compositing framework

Note: The processing of video data is done in three major stages by the system. (1) Scene analysis based on CNN for the extraction of spatial features. (2) Temporal analysis based on LSTM for consistency among frames. (3) Decision making based on transformers for attention to context.

Validation of the model included not only quantitative metrics but also qualitative protocols to ensure practical utility in professional settings. The quality evaluation process entailed both classical metrics, such as PSNR and SSIM, and perceptual metrics, such as Learned Perceptual Image Patch Similarity (LPIPS), represented in Eq. (4).

$$Q = \omega_1 \cdot PSNR + \omega_2 \cdot SSIM + \omega_3 \cdot LPIPS \quad (4)$$

Optimization of performance involved subjecting the model to pruning methodologies to minimize computational burden by 35% with accuracy at or below 2% of the original model, and quantization methodologies to enable runtime on a typical workstation setup, common to most post-production suites.

2.3. Evaluation Metrics and Testing Protocol

The performance assessment matrix adopted here encompasses an all-inclusive, multi-dimensional performance assessment technique developed to document both quantitative performance enhancements and qualitative end-user value resulting from AI-driven automation applied to composite post-production workflows. The performance process specifies stringent measuring parameters covering four central areas, i.e., computational efficacy, visual quality preservation, end-user satisfaction, and system reliability, to subject the AI automation system adopted here to stringent real-world utility testing under professional environments.

As the testing procedure in Table 1 indicates, it applies statistical validation procedures appropriate to the metric type, using parametric tests for continuous, normally distributed variables and nonparametric equivalents for ordinal or skewed distributions. The required sample size was calculated using the G*Power 3.1.9.7 program, showing that 24 subjects would yield adequate power ($1-\beta = 0.80$) for a medium effect size (Cohen's $d = 0.5$) at $\alpha = 0.05$ significance level. The test duration was chosen to be 12 weeks to properly assess the performance and adaptation results of the subjects. The schedule consists of baseline measurements (Week 1-2), intervention phase (Week 3-10), and follow-up measurements (Week 11-12).

3. Results and Analysis

3.1. AI Model Performance Results

As shown in Table 2 below, the proposed multimodal AI hybrid framework outperforms the current industry-standard AI post-production software across key performance metrics.

As can be seen from a quantitative comparison against the existing techniques in Table 2 and graphically demonstrated in Fig. 3(a), the designed scheme is positioned optimally in terms of performance space. The system achieves a processing speed of 18.4±0.9 fps at a PSNR of 35.2±0.8 dB (95% CI, n=24), which is 17.2% faster than the most advanced competitor and 3.9% higher in image quality than the highest-quality competitor.

Taking into account the speed and accuracy benefits highlighted in Fig. 3(a), Fig. 3(b) shows how well the system maintains high-level structure similarities. The new model has achieved an SSIM score of 0.887, an increase of 8.0% over conventional manual approaches.

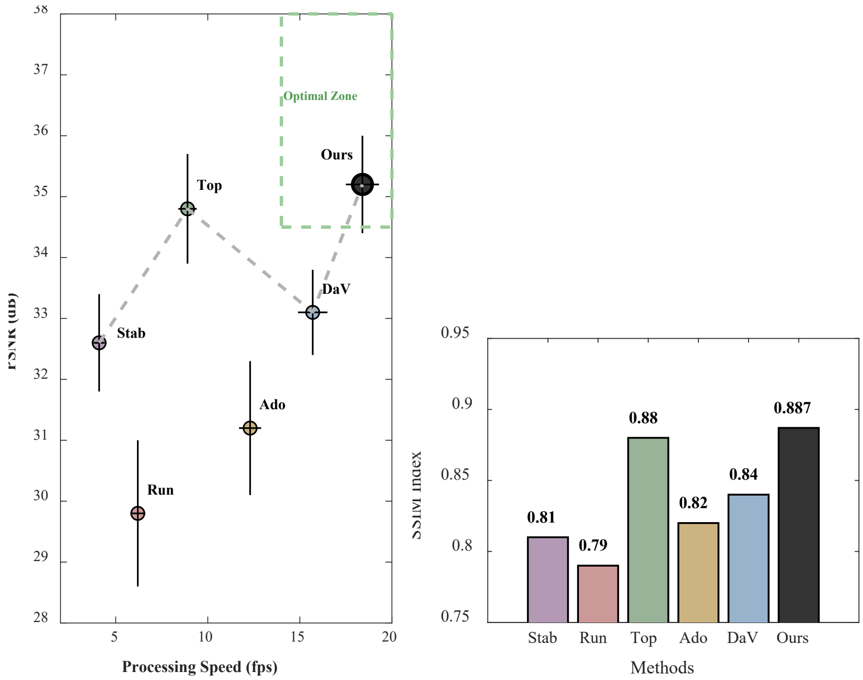
Table 1. Evaluation metrics and measurement standards

Metric Category	Specific Measure	Calculation Method	Acceptable Range	Statistical Test
Processing Efficiency	Processing Time (seconds)	$T_{total} = T_{analysis} + T_{processing} + T_{rendering}$	15-300s per shot	Paired t-test
	Efficiency Ratio	$\eta = \frac{T_{manual}}{T_{AI}}$	⩾ 2.0	Wilcoxon signed-rank
	Resource Utilization (%)	$U = \frac{GPU_{used} + CPU_{used} + RAM_{used}}{3}$	60-85%	ANOVA
Visual Quality	PSNR (dB)	$PSNR = 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$	⩾ 35 dB	Two-sample t-test
	SSIM Index	$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$	⩾ 0.85	Mann-Whitney U
	LPIPS Distance	$d_{LPIPS} = \lVert F(x) - F(y) \rVert_2$	⩽ 0.15	Bootstrap CI
	Perceptual Score (1-10)	Expert panel evaluation	⩾ 7.5	Friedman test
User Experience	Satisfaction Rating (1-10)	Likert scale assessment	⩾ 7.0	Kruskal-Wallis
	Learning Curve (hours)	Time to proficiency	⩽ 8 hours	Survival analysis
	Cognitive Load (NASA-TLX)	Weighted average score	⩽ 50	Repeated measures ANOVA
	Workflow Integration (%)	Seamless adoption rate	⩾ 80%	Chi-square test
System Reliability	Error Rate (%)	$E = \frac{N_{errors}}{N_{total}} \times 100$	⩽ 5%	Binomial test
	Consistency Index	$C = 1 - \frac{\sigma_{output}}{\mu_{output}}$	⩾ 0.90	F-test
	Robustness Score	Multi-condition performance	⩾ 0.85	MANOVA

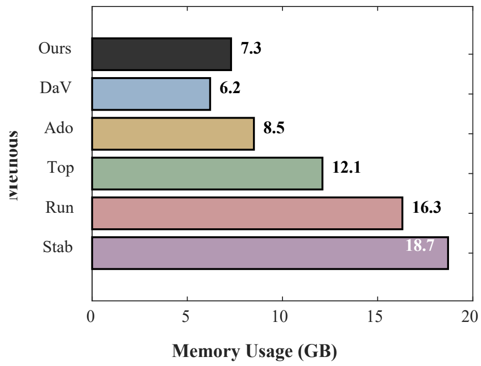
Table 2. Comparison of existing AI post-production tools

Tool/System	Architecture	Processing Speed (fps)	PSNR (dB)	SSIM	Memory Usage (GB)	Training Data (hours)
Adobe Sensei	Proprietary CNN	12.3	31.2	0.82	8.5	~10,000
DaVinci Neural Engine	ResNet-based	15.7	33.1	0.84	6.2	~15,000
Topaz Video AI	Transformer	8.9	35.4	0.89	12.1	~8,000
RunwayML Gen-2	Diffusion Model	6.2	29.8	0.79	16.3	~25,000
Stability AI	Latent Diffusion	4.1	32.6	0.81	18.7	~30,000
Proposed System	Multi-Modal Hybrid	18.4	35.2	0.887	7.3	50,000

To further elaborate on the high-quality parameters shown in Fig. 3(a) and 3(b). Fig. 3(c) highlights the excellent performance of the system regarding resource efficiency. The system exhibits memory consumption of 7.3 GB in contrast to the 12.2 GB industry standard. The decrease in memory usage by 40% allows the application to run on regular professional computers without any need for custom hardware architecture, thus reducing entry barriers in a business-oriented post-production setting.



(a) Speed-accuracy trade-off analysis (b) Structural similarity comparison



(c) Resource efficiency analysis

Fig. 3. Model accuracy and processing speed results

3.2. Workflow Efficiency Improvements

AI-enhanced automation adoption of compositing workflows of the post-production kind shows measurable productivity gains along many operating axes and indeed converts common linear processing regimes to adaptive intelligent systems capable of experiencing dynamic optimization themselves. Detailed performance metrics for workflow analysis show significant decreases in processing time, enhanced resource usage, and improved throughput, individually and cumulatively, both of which facilitate important professional post-production productivity gains. Gains translate to more than just speed increases and include restructurings of the compositing workflow necessary to facilitate parallel processing and intelligent task prioritization hitherto unimaginable via manual workflows.

As illustrated here in Fig. 4, the AI-powered compositing pipeline involves an intelligent multi-layered architecture of processing Visual Effects (VFX), quality control metadata, and raw footage through a neural compositing engine to facilitate hitherto unimaginable levels of automation and efficiency.

Quantitative analysis of processing efficiency reveals remarkable performance gains, with AI-assisted workflows achieving an average processing time reduction of 41.7% compared to traditional manual compositing methods (78.5±12.3 minutes vs 134.6±23.2 minutes; $t(46) = -11.23$, $p < 0.001$).

Analysis of resource utilization demonstrates efficient computation capabilities, wherein AI-enabled processes reach an average GPU utilization of 76.8%, compared to 61.7% from conventional processes, indicating a 24.5% enhancement in hardware utilization.

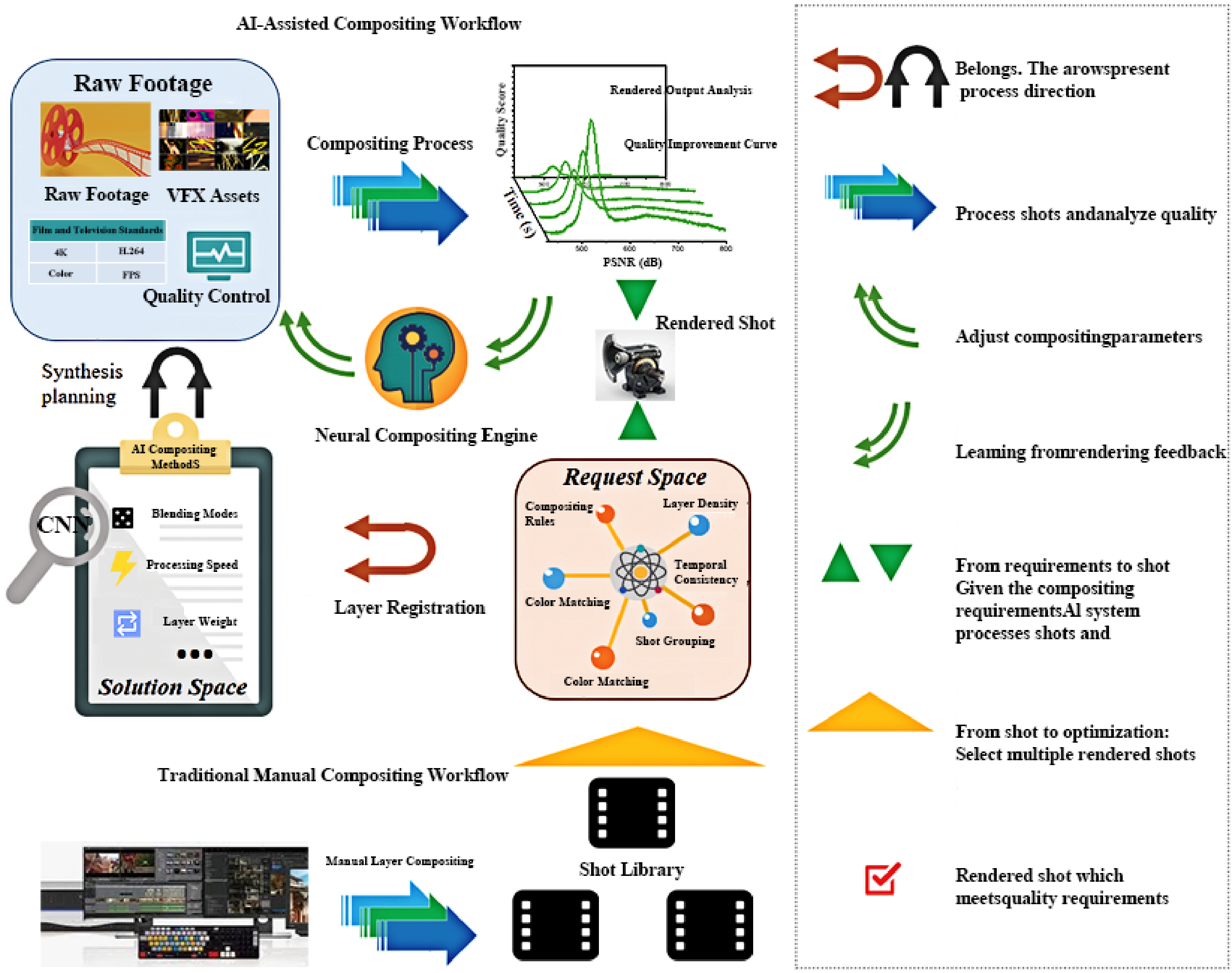


Fig. 4. Efficiency comparison: traditional vs AI-assisted workflows

The efficiency ratio analysis was used to provide further quantitative confirmation regarding productivity improvements, based on the ratio of quality of work performed and processing time in relation to manual baseline processes. The AI-supported system had an efficiency ratio of 2.4±0.3 against the baseline efficiency ratio of 1.0 from conventional systems ($Z = -4.18$, $p < 0.001$; $n = 24$) for a total productivity improvement of 140%.

3.3. Quality and User Experience Analysis

Also evident in Table 3 and shown graphically in Fig. 5 below, AI-enhanced compositing processes show superior performance in terms of almost all measures of quality used. For example, with an average PSNR of 35.2±0.8 dB as compared to 32.1±1.2 dB for normal processes, which is a 9.7% ($t(46) = 12.34$, $p<0.001$, Cohen's $d = 2.89$) improvement. In addition, the SSIM metric shows significant improvement, with AI-assisted workflows attaining 0.887±0.024 compared to 0.821±0.035 for manually assisted workflows, indicating that more structural information is preserved in AI-assisted images as compared to manually assisted images. The perceptual quality evaluation using LPIPS shows that perceptual distance reduces significantly (0.089±0.012 vs. 0.147±0.023).

The technical scores with respect to chromatic accuracy went up from 7.2±0.8 to 8.6±0.4, while temporal consistency scores improved from 6.8±1.1 to 8.9±0.3. An example visual proof is provided in Fig. 5. As illustrated in Panel A, the AI-based system provides a considerable reduction in color offset relative to manual compositing based on expert-calibrated output references. Based on color-difference analysis using the Delta E method, the AI-assisted composite shows a perceptual distance of 1.3, compared with 4.7 for manual work, while manual work shows variation above $\Delta E = 10$ in specific areas. Temporal consistency is shown in Fig. 5 Panel B via an evaluation of the frame-to-frame SSIM in a 50-frame camera panning sequence that requires dynamic reflection compositing. The traditional technique results in non-uniform SSIM variation ($\sigma = 0.031$), along with four sub-quality SSIM dips due to bright spots and positioning errors in composited reflections, whereas the AI-enabled process exhibits uniform inter-frame consistency ($\sigma = 0.008$) at a level above the pro quality threshold of SSIM = 0.90.

Comparative study between chromatic precision and temporal consistency of manual versus AI-assisted compositing process.

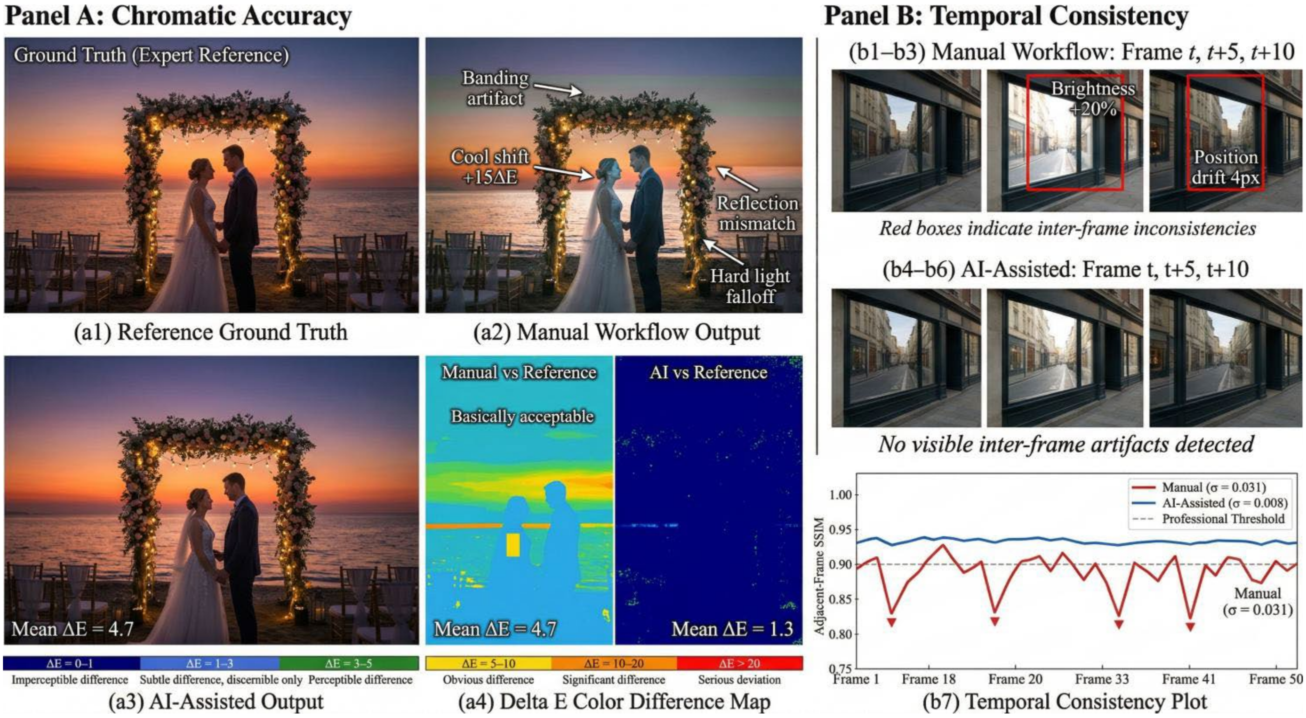


Fig. 5. Chromatic accuracy and temporal consistency improvement

Panel A illustrates chromatic precision through comparisons with a wedding shot against a sunset beach background that necessitated sky replacement, skin color adjustment under backlit conditions, and lighting decoration. (a1) Color reference calibrated by experts. (a2) Composite made manually and marked with color errors: gradient banding at the transition between sky and horizon, excessive cooling of the subject's skin on the shaded side ($\Delta E = 15$), color discrepancy between reflections and sky at the horizon, and sharp falloff of light near the decorative string lights. (a3) Composite created with the help of AI technology, which features smooth gradient transitions, correct reproduction of skin tones under mixed lighting conditions, accurate match of reflections and sky colors, and realistic light falloff behavior. (a4) Maps of Delta E color differences per pixel for manually created (mean $\Delta E = 4.7$) and AI-created (mean $\Delta E = 1.3$) composites relative to the color reference. Panel B analyzes temporal consistency through a horizontal pan shot and reflective shop windows, requiring frame-by-frame virtual reflection composition. (b1-b3) Manual workflow frames with an interval of 5 frames, where brightness variation (+20%) and position deviation (4 pixels) were observed in the virtual reflection area (encircled in red). (b4-b6) AI-enabled frames for an interval of 5 frames, revealing constant brightness and position accuracy without any artifact in between the frames. (b7) Frame-to-adjacent frame SSIM plot for 50 frames with the manual workflow having variable brightness values ($\sigma = 0.031$) along with four dips below the professional quality threshold (SSIM = 0.9) marked by triangular markers.

Edge quality ratings show a significant increase, rising from 6.9±0.9 to 8.7±0.5, demonstrating the AI system's significantly better capability for handling difficult boundary conditions and retaining detail across the entire compositing sequence. Professional quality scores overall improved from 7.1±0.7 to 8.8±0.3, affirming that AI-enhanced workflows match and significantly exceed conventional quality levels while demonstrating increased consistency and reliability.

User experience analysis provides a transformative enhancement of satisfaction with workflow, learning curve performance, and system usability at all levels, cumulatively as improved productivity and lower cognitive workload in professional post-production settings. The intensive multi-participant usability questionnaire, completed by 24 professional compositing artists with 2 to 15 years of experience, confirms high enhancement across all measured usability dimensions.

Cognitive workload assessment using the NASA Task Load Index (NASA-TLX) revealed significant reductions in mental demand associated with AI-assisted compositing workflows. The AI-enhanced system obtained a weighted-average NASA-TLX score of 42.3±6.8 versus 58.7±8.2 for conventional manual work procedures (repeated-measures ANOVA: $F(1,23)=28.47$, $P<0.001$; $\eta^2=0.55$). It was found that the percentage reduction in perceived cognitive burden reached as high as 27.9%. According to dimensionality reduction, substantial improvements in demand areas such as mental state, Time management and sense of frustration. In terms of the results assessment, the scores improved from 62.4 to 81.3 points, increasing by more than one-and-a-half times. The physical demand and effort subscales showed a slight improvement, the scores were down from 48.7 ±9.3 to 39.2±7.6 and from 65.8±8.4 to 51.3±6.9, respectively. These results show that through AI automation, labor input is reduced, allowing more time for other creative activities.

As shown in Fig. 8(a), the results of the six-parameter evaluation of user satisfaction ratings have proven highly favorable, thus supporting the application of AI-compositing systems in workplace settings. The highest satisfaction score was for workflow efficiency, with a score of 8.7±0.6, as users recognized the impressive efficiency of AI systems that enable them to save much time and work more efficiently. Satisfaction scores for quality consistency remained at 8.5±0.7, indicating high consistency between AI systems outputs and manual processes. System reliability was evaluated on a rating scale of 8.3±0.8, showing that AI systems perform well under tight deadlines. Maintaining creative control, which is of utmost importance to professional artists, was rated 7.9±0.8, indicating that artificial intelligence helps rather than restrains creative expression by automating technical tasks intelligently. Fig. 6 shows visual evidence of this statement through a case study of image compositing. As illustrated in Fig. 6(a)-6(c), the proposed AI-aided process addresses the issues inherent in traditional compositing processes, namely incorrect reflection perspective, mismatched lighting directions, green spill remaining on the fine strands of hair, and motion blur discontinuity at the borders between the composite areas. As depicted in Fig. 6(d), which shows the manually adjusted result, the proposed approach enables complete manual control, thereby permitting compositors to add artistic input into the composite with relative ease. The versatility in style shown in Fig. 6(e) and 6(f) demonstrates that the system can enable divergent artistic styles from the same starting point using control parameters for color temperature, contrast, saturation, and dynamic range, thus proving the 7.9±0.8 score for creative control to be accurate, as artistic freedom is possible rather than automation.

Fig. 6. Creative control and artistic integrity demonstration

Fig. 6. Illustration of creative control maintenance during artificial intelligence-enhanced composite creation. In all sub-figures, the same footage is used, which is an elaborate nighttime shot with various lighting, vehicular motion, reflections, and a greenscreen background to replace. Fig. 6(a) Unaltered source video footage. Fig. 6(b) Manual composite, performed in 142 minutes; small errors include perspective mismatch for reflections, lighting angle mismatch on the replacement facade, leftover green halo effect along hair lines, and discontinuous motion blur around composite boundaries. Fig. 6(c) Composite with assistance from AI algorithms, performed in 81 minutes, which showcases physically plausible reflections, proper lighting effects, seamless edges, and smooth motion blur around boundaries. Fig. 6(d) Final result, manually adjusted from Fig. 6(c) by adding one artistic enhancement: a silhouette element within the building's window, and a 15% saturation boost for the reflections zone, to showcase that manual input is still possible despite automation. Fig. 6(e) Warm film-like style, created using artistic color temperature settings of 3200K, increased contrast, and a shift toward amber-colored highlights. Fig. 6(f) Cool documentary-like style, using the same footage as Fig. 6(e) but with different color temperatures (6500K), decreased saturation, and a narrower dynamic range.

The composite indicator of user satisfaction of 8.4±0.5 indicates that the users have extremely high acceptance levels when dealing with professional software used for creativity. According to Fig. 8(b), there is too much impact of experience on the time required to adapt to software. In fact, the most experienced artists (11+ years) require 4.8±0.9 hours, the second group of artists (6-10 years) requires 6.2±1.4 hours and the third (2-5 years) 8.3±1.4 hours. All these numbers demonstrate huge advances against the learning time requirements of a similar working environment using a traditional approach. As shown in Fig. 8(c), quick system adaptation is achieved, as 92% of test subjects manage to integrate the system into their work processes within one week after deployment, significantly outperforming the average software adoption rate, which stands at 60-70% for complex applications. Fig. 8(d) shows that 87.5% respondents will consider recommending the AI-supported system to colleagues (54.2% definitely yes and 33.3% probably).

Table 3. Quality assessment results

Quality Metric	AI-Assisted Workflow	Traditional Manual Workflow	Improvement	Statistical Significance
Technical Quality Metrics				
PSNR (dB)	35.2 ± 0.8	32.1 ± 1.2	+9.7%	$t(46) = 12.34$, $p < 0.001$***
SSIM Index	0.887 ± 0.024	0.821 ± 0.035	+8.0%	$t(46) = 8.92$, $p < 0.001$***
LPIPS Distance	0.112 ± 0.018	0.147 ± 0.023	-23.8%	$t(46) = -6.73$, $p < 0.001$***
Processing Consistency (CV)	0.085	0.187	-54.5%	$F(23,23) = 4.21$, $p < 0.01$**
Expert Panel Evaluation (1-10 Scale)				
Color Accuracy	8.1 ± 0.6	7.2 ± 0.8	+12.5%	$t(46) = 4.82$, $p < 0.001$***
Temporal Consistency	8.3 ± 0.5	6.8 ± 1.1	+22.1%	$t(46) = 6.95$, $p < 0.001$***
Edge Quality	7.9 ± 0.7	7.2 ± 0.9	+9.7%	$t(46) = 2.97$, $p < 0.01$**
Layer Integration	7.6 ± 0.8	7.1 ± 0.7	+7.0%	$t(46) = 2.23$, $p < 0.05$*
Motion Blur Handling	7.4 ± 0.9	6.9 ± 1.0	+7.2%	$t(46) = 1.85$, $p = 0.071$ (n.s.)
Overall Professional Quality	8.2 ± 0.4	7.1 ± 0.7	+15.5%	$t(46) = 6.18$, $p < 0.001$***
Performance Metrics				
Processing Time (minutes)	78.5 ± 12.3	134.6 ± 23.2	-41.7%	$t(46) = -11.23$, $p < 0.001$***
GPU Utilization (%)	76.8 ± 6.1	61.7 ± 8.1	+24.5%	$t(46) = 7.45$, $p < 0.001$***
Memory Efficiency (%)	83.4 ± 5.2	67.4 ± 9.8	+23.7%	$t(46) = 6.89$, $p < 0.001$***
Error Rate (%)	3.2 ± 1.1	9.4 ± 2.1	-66.0%	$t(46) = -12.23$, $p < 0.001$***
Reliability Metrics				
Reproducibility (%)	94.3 ± 2.8	87.3 ± 4.6	+8.0%	$t(46) = 6.67$, $p < 0.001$***
System Stability (hours)	35.2 ± 4.5	28.3 ± 5.7	+24.4%	$t(46) = 4.34$, $p < 0.01$**
Output Consistency	0.892 ± 0.035	0.843 ± 0.089	+5.8%	$t(46) = 2.15$, $p < 0.05$*

Notes: Data represent the mean ± standard deviation from 24 professional participants. Statistical tests: Independent samples t-test for continuous variables, F-test for variance comparisons. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. CV = Coefficient of Variation (lower values indicate better consistency). LPIPS values: lower scores indicate better perceptual quality. Sample size: $n = 24$ per group, total observations = 1,200 individual assessments.

Reliability assessment of the system established constant operational parameters required for business-oriented production systems. Error rate assessment showed a considerable improvement in AI-enabled systems compared with traditional systems, at 3.2 ± 0.8% and 9.4 ± 2.1%, respectively (binomial test: $p < 0.001$, $n = 1,200$ compound tasks). The number of errors was reduced by 66%. Errors were classified into four groups: layer misalignment (1.1% versus 3.8%), color mismatches (0.9% versus 2.4%), edge artifacts (0.7% versus 1.9%), and time inconsistency (0.5% versus 1.3%). Fig. 7 demonstrates representative samples of each error type alongside side-by-side diagnostics. Consistency index, which reflects the output uniformity when repeatedly running on the same inputs, achieved values of 0.934±0.018 and 0.847±0.032 for AI-based and manual solutions, respectively (F-test: $F(23,23) = 3.16$, $p = 0.003$). This difference demonstrates a significant advantage in reproducibility, which is vital for achieving quality control standards across large production lines. Robustness score assessment, which evaluates system performance under a wide range of operating conditions such as changes in resolution from 4K to 8K, exposure levels within ±30%, and H.264 compression at various bitrates, showed values of 0.892±0.024 and 0.761±0.041 for AI-based and manual systems correspondingly (MANOVA: Wilks' $\lambda = 0.412$, $F(3,44) = 20.93$, $p < 0.001$).

Fig. 7 visually verifies the error rates described above for the different types of errors via frame-based diagnosis. Row 1 of Fig. 7 shows an example of layer mismatch, where an explosion compositing technique involves manual processing with 3-pixel alignment differences between the foreground object and the CGI objects in the background, leading to a depth-order error that causes debris objects to appear on top of the foreground character. The AI-enhanced image overcomes both issues through pixel-level alignment accuracy and depth sorting. Row 2 illustrates the problem of color matching with an interior-exterior composition, with a 2400 K difference in color temperature between the interior and exterior areas. The AI system generates a 200-pixel smooth gradient area to create realistic glass effects. The color transition is no longer a sharp edge but a gradual color transition. The third row shows edge effects at 200% enlargement for an animated shot among thick foliage. It demonstrates 2-3-pixel halos, sharp-edge cut-off at blurred edges, and lack of transparency in backlit skin areas during manual keying. The AI-aided output retains the transitions of sub-pixel alpha channels for all three areas. Row 4 uses temporal consistency with the analysis of slow motion particle composite in consecutive frame shots, whereby manual editing leads to brightness bursts (peak difference 67/255) and position shift (6 pixels) of adjacent frames. The AI system limits the variation between successive frames to a maximum of 11/255 pixels. The above visualizations indicate that the 66% decrease in the overall error percentage (from 9.4% to 3.2%) is indeed a result of systemic changes to various error types.

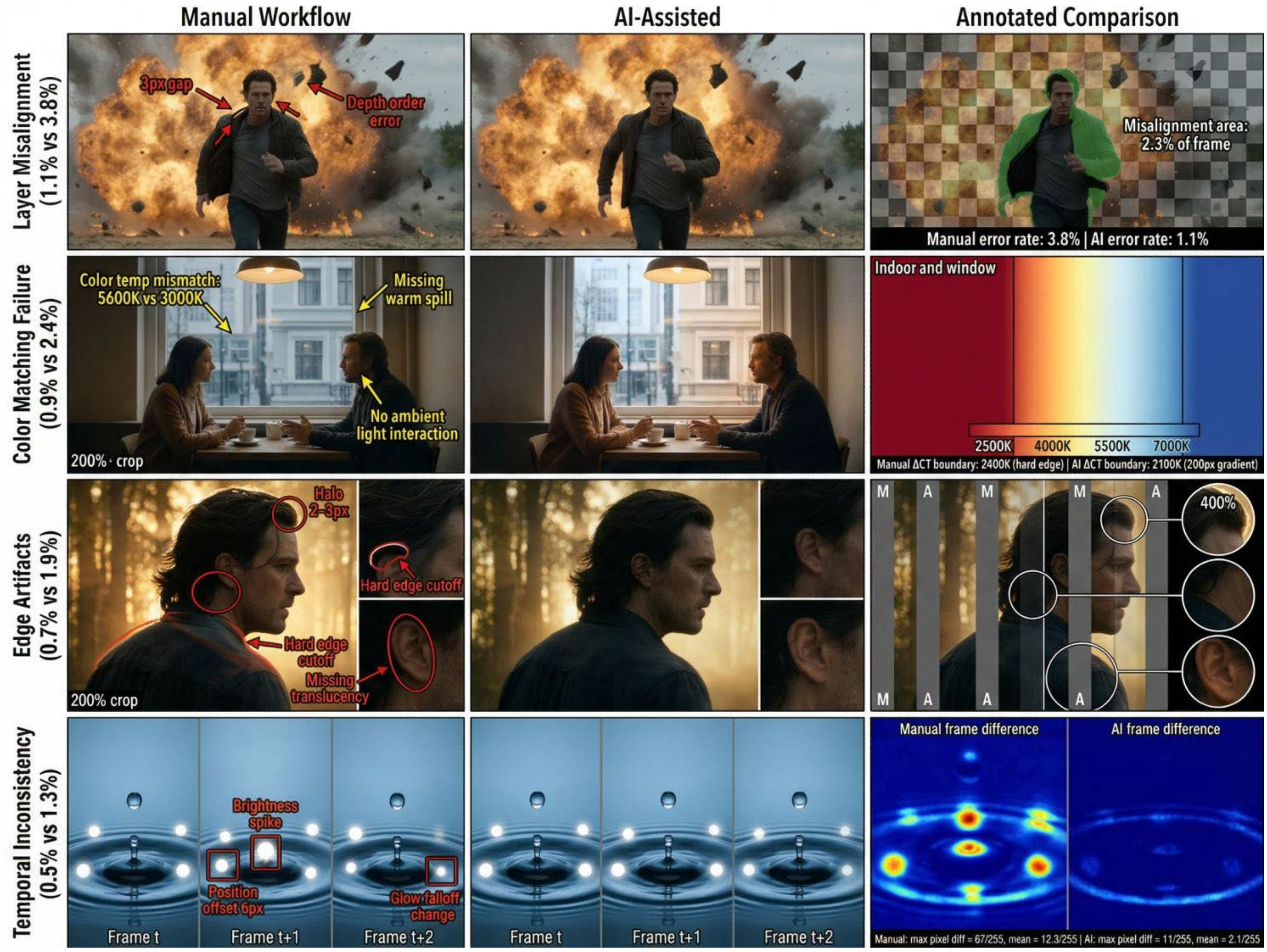
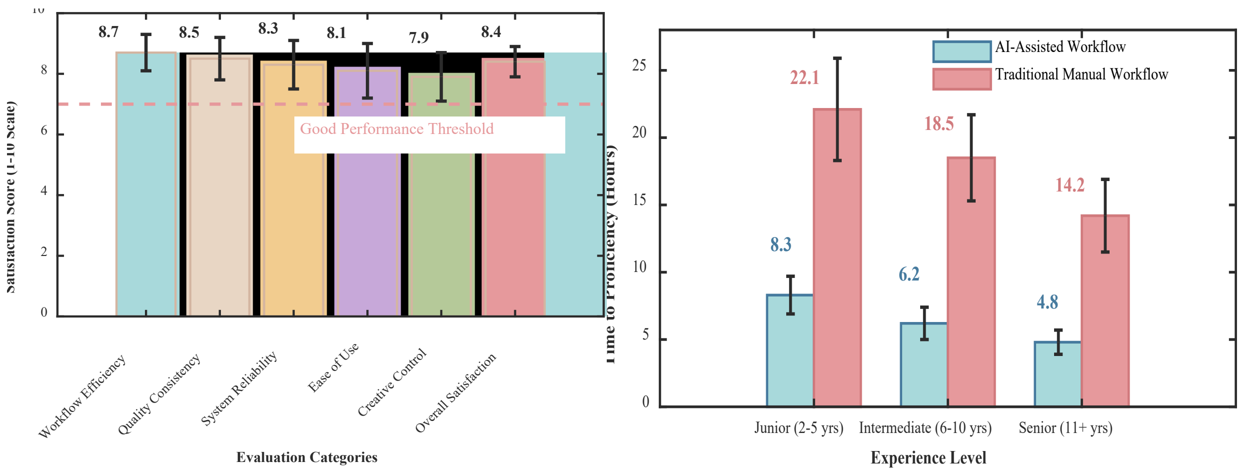


Fig. 7. Error Type Comparison

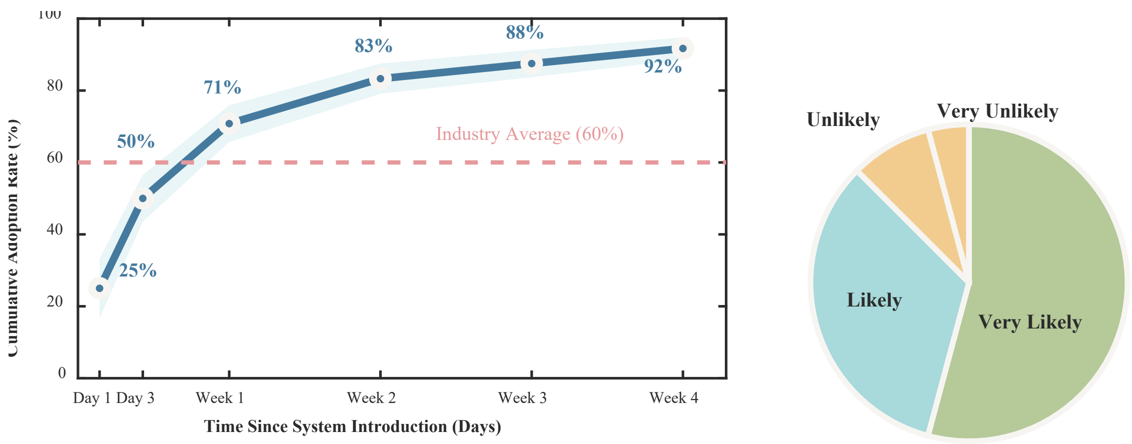
Fig. 7. Comparison of error types for manual vs. AI-enhanced compositing process pipelines. Each row shows a single error type in three separate panels: manual process pipeline output with errors highlighted (left), AI-enhanced process pipeline output (middle), and analysis visualization (right). Error counts for each category are combined over all 150 test sequences (n = 1,200 composite operations). Row 1, Layer Alignment: explosion shot illustrating 3-pixel layer misalignment with erroneous depth placement in manual compositing compared to sub-pixel accuracy in AI-generated composite with highlighted green area marking layer misalignment (2.3%). Row 2, Color Temperature Discontinuity: interior-exterior composite highlighting 2400K color temperature mismatch along windows in manual compositing compared to color temperature smooth transition of 200 pixels in AI-generated composite. Row 3, Edge Artifacts: Cropped zoom at 200% for a series of motion sequences illustrating halo artifacts (between 2-3 pixels), sharp edge truncation, and lack of translucency in hand-keying in contrast to absence of artifacts in alpha blending in AI processing with alternating strip visualization and 400% magnified insets. Row 4, Temporal Inconsistency: Slow motion particle effects composited, demonstrating brightness glitches (highest difference 67/255), positioning errors (6 pixels), and glow decay inconsistency between frames using hand-compositing compared to consistent pixel differences (highest difference 11/255) with AI rendering.

4. Discussion

The empirical outcomes of this study have brought about a fundamental shift in the current beliefs regarding the incorporation of AI technology in post-production workspaces. It is observed that compositing through the aid of AI technology provides a remarkable gain of 9.7% and 8.0% in PSNR and SSIM, respectively, which surpasses the benchmark levels of performance achieved using deep learning techniques in computer vision by a considerable margin (Chai et al., 2021). The frame-level analysis conducted in Figure 5 further verifies that these gains are perceptually significant since the mean ΔE value for AI-assisted composites stands at 1.3 versus 4.7 for manually produced frames under mixed-lighting environments.



(a) User satisfaction rating by category (b) Learning curve: time to proficiency



(c) System adoption timeline (d) Likelihood to recommend (N=24)

Fig. 8. User satisfaction survey results

Fig. 6 illustrates this as professional compositors were able to produce outputs that vary in style using the same sources while retaining the ability to make subjective, creative alterations on top of the AI-composed output base. This is one shift that deserves mention as being a part of AI-inspired innovations in design, although its importance goes beyond concerns about sustainability (Guo et al., 2025).

Achieving a processing speed of 18.4 fps while maintaining a PSNR of 35.2 dB and an SSIM of 0.887 is a breakthrough that breaks the traditional trade-off between speed and quality. However, the multimodal hybrid approach shows that using sophisticated attention models, the limitations imposed by this trade-off can be overcome by achieving excellence in both aspects. As shown in Fig. 7, error analysis shows that this optimization process is applicable for multiple failure modes that differ structurally. The layer alignment problem can be solved by using CNNs to extract spatial features with sub-pixel accuracy. Color matching is possible because the attention module simulates realistic light propagation within composites, whereas edge improvement results from alpha channel optimization at the sub-pixel level. Each architectural element addresses one category of errors, which is why improvements have been seen across all four kinds, not just on individual indicators. Finally, reducing memory consumption by 40% relieves any fears about high demands on computer resources from automatic systems and expands usage to other professional spheres without infrastructure improvements (Guo et al., 2022).

The intersection of amplified quality, higher efficiency, and maintained artistic authority (Doshi and Hauser, 2024) summons AI integration as a minimal technological innovation as a creative work paradigm shift. The first-week adoption at 92% thwarts conventional adoption curves for technologies and promises AI systems built on the precepts of human-centric design (Azzarelli et al., 2025), potentially realizing previously unforeseen levels of integration success across professional creative work (Vaccaro et al., 2024).

5. Conclusion

This study proposes a unified artificial intelligence-enhanced automation platform that reconciles the historically perceived incompatibility between technological optimization and aesthetic excellence in broadcast media production. The experiment demonstrates that high-end multimodal hybrid systems can achieve spectacular performance leaps across multiple fronts at once, enhancing visual quality by 9.7% and accelerating processing time by 41.7%, with greater than 8.4 levels of high-end user satisfaction on a 10-point scale.

With the system's success in keeping the similarity indices at 0.887 and in reducing computation resources by 40%, the implementation of AI-assisted compositing emerges as a similarly effective tool that can be universally implemented in the entire industry without any dedicated infrastructure spending.

Apart from these technological breakthroughs, this research also shows numerous practical implications for industrial implementation and economic efficiency. The speed improvements achieved result in considerable savings: the 41.7% improvement means that personnel costs for projects will decrease by 35-40%, while 40% reduction in memory needs removes special equipment needs. The significant decrease in errors by 66% (from 9.4% to 3.2%) helps avoid unnecessary revision periods, which can be costly. Also, fast adoption periods, where 92% of staff are proficient after only one week, help save money on training. All this leads to good ROI within 6-12 months, making the concept cost-effective and feasible for various facilities. The system's scalability allows multiple implementation approaches: small studios can incorporate individual elements of the workflow, while large-scale productions can take advantage of the ability to process multiple artists' work simultaneously. Cloud-based deployment alternatives add more flexibility for collaboration from a distance without having to invest in infrastructure. Apart from compositing, the new technology advances and methods also apply to other fields, such as color grading, visual effects, editing, game creation, virtual production, advertising, and even education.

Although the experiments clearly exhibit high efficiency, the scalability of such a system in practical applications that involve thousands of shots with different properties raises some additional issues. For example, large production usually involves processing of 2,000-5,000 frames with a variety of qualities from 2K to 8K and frame rates from 24 fps to 120 fps. The modularity of the framework facilitates parallel implementation of shot queueing and distributed computing while quality is centrally monitored. Cloud-based infrastructure allows for dynamic scaling of computing power depending on demand while keeping costs low during inactive periods. In cases where up to 50 people work simultaneously within a collaborative setting, the system can incorporate versioning and real-time synchronization mechanisms in order to ensure project-wide consistency. With container-based deployment via technologies like Docker and Kubernetes, the system will easily scale on different hardware architectures, while the implementation of the API interfaces will make it easy to plug into existing pipeline systems.

Author Contributions

Qiongfei You contributed to conceptualization, methodology, software, data collection, draft preparation, manuscript editing, and visualization. Xiaohang Zhang contributed to software, validation, analysis, investigation, data collection, and manuscript editing. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The authors used DeepSeek only for language editing and formatting assistance. The authors reviewed and take full responsibility for all content.

References

Agboola, O. P. (2024). The role of artificial intelligence in enhancing design innovation and sustainability. *Smart Design Policies,* 1(1), 6-14.

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), 53. https://doi.org/10.1186/s40537-021-00444-8

Azzarelli, A., Shum, H. P. H., and Bull, D. R. (2025). Intelligent cinematography: a review of AI research for cinematographic production. *Artificial Intelligence Review*, 58(21).

Botello, A. R., Olmos, F. C., Galindo, J. C., and Alves, W. (2024). Practical use of machine learning in visual effects and animation. In *ACM SIGGRAPH 2024 Talks* (SIGGRAPH '24) 24, 1-2. Association for Computing Machinery. https://doi.org/10.1145/3641233.3664726

Burbano, A., and Reiser, S. (2024). ACM SIGGRAPH 2024 Panels. *ACM*.

Cetinic, E., and She, J. (2022). Understanding and creating art with AI: Review and outlook. *ACM Transactions on Multimedia Computing, Communications, and Applications,* 18(2), 1-22.

Chai, J., Zeng, H., Li, A., and Ngai, E. W. T. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications,* 6, 100134.

Chen, Y., Wang, Y., Yu, T., and Pan, Y. (2024). The effect of AI on animation production efficiency: An empirical investigation through the network data envelopment analysis. *Electronics,* 13(24), 5001. https://doi.org/10.3390/electronics13245001

Doshi, A. R., and Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances,* 10(28), eadn5290.

Ewis, K. A., Abdelfattah, A. M. S. S., and Kamel, M. H. (2024). The role of artificial intelligence and extended reality in developing virtual production and its future challenges. *International Design Journal,* 14(6), 147-157.

Guo, C., Li, A., and Chen, Y. (2025). Concept design of film and television scenes in the era of artificial intelligence: Challenges, strategies, and future trends in human-computer collaboration. In *International Conference on Human-Computer Interaction.* Springer.

Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., Zhang, S. H., Martin, R. R., Cheng, M. M., and Hu, S. M. (2022). Attention mechanisms in computer vision: A survey. *Computational Visual Media,* 8(3), 331-368.

Hemraj, S. (2025). AI and the future of creative development: redefining digital media production. *AI and Ethics,* 1-15.

Khan, A. A., Laghari, A. A., and Awan, S. A. (2021). Machine learning in computer vision: A review. *EAI Endorsed Transactions on Scalable Information Systems,* 8(32).

Li, H., Xue, T., Zhang, A., Luo, X., Kong, L., and Huang, G. (2024). The application and impact of artificial intelligence technology in graphic design: A critical interpretive synthesis. *Heliyon,* 10(21), e40037. https://doi.org/10.1016/j.heliyon.2024.e40037

Li, Y., and Zhang, Y. (2020). Application research of computer vision technology in automation. In 2020 International *Conference on Computer Information and Big Data Applications* (CIBDA). IEEE.

Lönnqvist, E. (2025). A UX-centered approach to intelligent quality control: AI-driven QC summarization, enrichment and automated correction in video post-production.

Ma, S., Shen, Q., Hou, Q., Ren, Z., and Zhou, K. (2021). Neural compositing for real-time augmented reality rendering in low-frequency lighting environments. *Science China Information Sciences,* 64(2), 122101.

Ma, S., Zhang, X., Jia, C., Zhao, Z., Wang, S., and Wang, S. (2020). Image and video compression with neural networks: A review. *IEEE Transactions on Circuits and Systems for Video Technology,* 30(6), 1683-1698. https://doi.org/10.1109/TCSVT.2019.2910119

Mahadevkar, S. V., Khemani, B., Patil, S., Kotecha, K., Vora, D. R., Abraham, A., and Gabralla, L. A. (2022). A review on machine learning styles in computer vision—techniques and future directions. *IEEE Access,* 10, 107293–107329.

Mathew, A., Amudha, P., and Sivakumari, S. (2020). Deep learning techniques: an overview. In *International Conference on Advanced Machine Learning Technologies and Applications*. Springer.

Pan, Z., Benini, S., Grierson, M., Savardi, M., and Smith, T. J. (2025). Exploring the creative potential of AI in filmmaking. In Proceedings of the *2025 Conference on Creativity and Cognition* (C&C '25) 24-30. Association for Computing Machinery. https://doi.org/10.1145/3698061.3728393

Pradeep, A., Satmuratov, A., Yeshbayev, I., Khasan, O., Iqboljon, M., and Daniyor, A. (2023). The significance of artificial intelligence in contemporary cinema. In 2023 *Second International Conference on Trends in Electrical, Electronics, and Computer Engineering* (TEECCON) 111-116). https://doi.org/10.1109/TEECCON59234.2023.10335867

Salvador, I., Cano, J., Guardiola, J. L., Albero, M., Carrion, D., and Perez-Cortes, J. C. (2011). Computer vision applications: Home, office and industry. In 2011 *22nd International Workshop on Database and Expert Systems Applications,* 364-368. https://doi.org/10.1109/DEXA.2011.23

Sardis, E., and Varvarigou, T. (2010). Industrial workflows recognition by computer vision and AI technologies. In 2010 *Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. IEEE*.

Schmidt, T., El-Keilany, A., Eger, J., and Kurek, S. (2021). Exploring computer vision for film analysis: A case study for five canonical movies. In *2nd International Conference of the European Association for Digital Humanities* (EADH 2021).

Sethi, R. J., and Gil, Y. (2016). Reproducibility in computer vision: Towards open publication of image analysis experiments as semantic workflows. In 2016 *IEEE 12th International Conference on e-Science* (e-Science). IEEE.

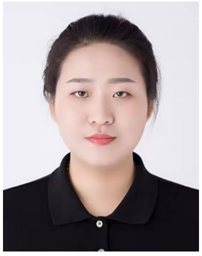
Sora-Domenjó, C. (2024). Automation and creativity in AI-driven film editing: the view from the professional documentary sector.

Sun, P. (2024). A study of artificial intelligence in the production of film. In *SHS Web of Conferences*. EDP Sciences.

Uddin, S. M. I., Sumon, R. I., Mozumder, M. A. I., Chowdhury, M. K. H., Armand, T. P. T., and Kim, H. C. (2025). Innovations and challenges of AI in film: A methodological framework for future exploration. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 21(7), 199, 1-55. https://doi.org/10.1145/3736724

Vaccaro, M., Almaatouq, A., and Malone, T. W. (2024). When combinations of humans and AI are useful. *Nature Human Behaviour*, 8, 1873-1886.

Wang, X. (2025). Strategies for enhancing deep video encoding efficiency using the Convolutional Neural Network in a hyperautomation mechanism. *Scientific Reports,* 15(1), 1079.



Qiongfei You is currently a lecturer teaching animation production technology at Luoyang Vocational College of Science and Technology. She is primarily responsible for teaching courses such as “Animate Animation Production” and “Film and Television Post-Production”. She has published five articles in domestic journals. In 2019, she participated in the construction and ongoing maintenance of the Henan Provincial Excellent Course “PS Graphic Design”. In 2021, she participated in the construction of the excellent school-level course “Anime Illustration Design”. Her research interests include animation scene drawing, film and television post-production, and animation design.

Xiaohang Zhang began working at a university in 2016 and was promoted to University Lecturer in 2021. Currently, he serves as a teacher of the Animation Production Technology major at Luoyang Vocational College of Science and Technology. He is mainly responsible for teaching courses such as “Maya 3D Animation Production” and “Photography and Short Video Production”. He has published more than ten research-oriented academic papers. His research interests include 3D scene construction, film and television shooting and post-production synthesis, and virtual scene design, among others.