AIGC 技术推进智能视觉创新的影响与前瞻研究

李志良，穆忠伟
张家界航空工业职业技术学院，湖南 张家界 427000

DOI: 10.61369/TACS.2026080054

摘　　要： 人工智能生成内容 (Artificial Intelligence Generated Content, AIGC) 的快速迭代与升级打破了传统智能视觉系统感知与处理能力的局限性，重构了智能视觉领域的技术范式与产业生态。传统智能视觉系统以判别模型作为核心点，存在数据依赖度高、场景泛化能力弱、内容生成与理解割裂等固有缺陷，难以适配复杂动态场景下的现代化应用需求。本文以 AIGC 技术与智能视觉系统的深度融合为研究核心，系统剖析 AIGC 驱动智能视觉系统技术创新的内在机理，从感知能力、内容创作、场景适配、算法架构四个维度阐述技术成效。研究旨在理清 AIGC 与现代化智能视觉系统的融合逻辑，为智能视觉技术创新迭代、产业规模化落地提供理论参考与实践指引。

关 键 词： AIGC; 智能视觉系统 ; 多模态融合 ; 视觉感知

Research on the Impacts and Prospects of AIGC Technology Advancing Innovations in Intelligent Vision

Li Zhiliang, Mu Zhongwei
Zhangjiajie Institute of Aeronautical Engineering, Zhangjiajie, Hunan 427000

Abstract: The rapid iteration and upgrading of Artificial Intelligence Generated Content (AIGC) have broken the limitations of perception and processing capabilities inherent to traditional intelligent vision systems, reshaping the technical paradigm and industrial ecosystem within the intelligent vision field. Centered on discriminative models, traditional intelligent vision systems carry inherent drawbacks including heavy reliance on data, weak scene generalization capacity, and separation between content generation and comprehension, making them ill-suited to meet the demands of modern applications under complex and dynamic scenarios. Focusing on the deep integration of AIGC technology and intelligent vision systems, this paper systematically analyzes the internal mechanism through which AIGC drives technological innovations in intelligent vision systems, and elaborates on technical achievements across four dimensions: perceptual capability, content creation, scene adaptation, and algorithm architecture. This research aims to clarify the integration logic between AIGC and modern intelligent vision systems, and provide theoretical references and practical guidelines for the iterative innovation and large-scale industrial implementation of intelligent vision technologies.

Keywords: AIGC; intelligent vision system; multimodal fusion; visual perception

引言

随着人工智能技术进入生成式发展新阶段，Diffusion 模型、GAN 模型以及 MLL 模型为核心的 AIGC 技术实现了跨越式突破，成为驱动人工智能产业升级的核心动力。智能视觉系统作为人类间接方式感知世界与理解各种应用场景的核心载体，已经广泛应用于工业生产、日常生活等诸多领域，是数字化、智能化社会建设的关键基础设施。传统现代化智能视觉系统依托判别式深度学习算法，聚焦图像分类、目标检测、语义分割、行为识别等感知能力，经过多年发展已形成成熟的技术体系与应用模式，但在复杂动态场景、稀缺样本场景、个性化内容生成等场景中，逐渐暴露出局限性 [1]。

目前 AIGC 在智能视觉系统中的应用已从技术探索阶段迈入规模化创新阶段，催生出了视觉数据智能生成、场景仿真重构、缺陷智能修复、多模态视觉交互、个性化视觉内容创作等一系列新兴技术与应用形态，推动智能视觉系统从“被动识别”向“主动认知、主动创造”转型升级 [2]。但与此同时，在创新应用过程中仍存在算法算力成本过高、生成内容精准度不足、数据版权模糊、伦理风险失控、行业落地标准缺失等诸多问题，制约了技术的深度创新与产业落地。基于此，本文系统研究 AIGC 技术对现代化智能视觉系统技术创新的赋能机制、应用成效与现存困境，精准预判未来技术迭代与产业发展趋势，为推动智能视觉领域高质量技术创新、构建安全可控的产业生态提供理论支撑。

课题项目：湖南省教育厅科学研究项目《AIGC 技术推进现代化智能视觉系统技术创新的影响与前瞻研究》，课题编号：24C1329。
作者简介：李志良（1989.09—），男，硕士研究生，副教授，研究方向：自动控制与图像识别处理。

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

一、相关概念与技术发展现状

（一）AIGC 模型架构

AIGC 是依托深度学习大模型，基于海量数据训练，自动生成文本、图像、视频、音频、三维模型等多元化内容的人工智能技术体系 [1]。区别于传统人工智能的判别、分析、决策功能，AIGC 的核心特征是创造性生成与多模态交互，能够基于语义理解实现内容的自主创作、优化与迭代，具备强泛化、高适配、低成本的技术优势。当前主流 AIGC 视觉技术架构主要包括生成对抗网络（GAN）、扩散模型（Diffusion Model）、视觉语言大模型（VLM）三大核心体系，分别适配图像重构、高清生成、语义级视觉理解与创作等不同场景需求 [3]。图1为基于 GAN 的 AIGC 模型架构。

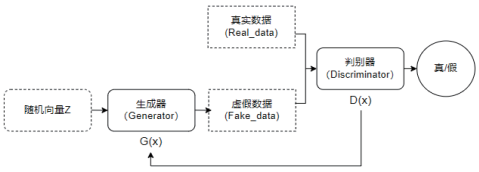


图1 基于 GAN 的 AIGC 模型架构

现代化智能视觉系统是融合机器视觉、深度学习、传感器技术、物联网技术的智能化感知系统，核心是通过图像采集、数据预处理、算法分析、结果输出全流程，实现对客观场景的智能化感知、识别、理解与反馈。相较于传统机器视觉系统，现代化智能视觉系统更强调场景自适应、多模态融合、实时交互、智能决策，不再局限于单一的视觉识别任务，而是实现视觉感知与场景理解、行为预测、内容优化、智能控制的深度融合 [3]。

（二）基于 AIGC 的智能视觉发展

从智能视觉系统发展历程来看，行业历经传统机器视觉、深度学习判别式视觉、生成式智能视觉三个发展阶段。早期传统机器视觉依托固定算法规则，仅能完成简单的图像预处理与特征匹配，无法适配复杂动态场景；深度学习时代的判别式视觉模型，依托卷积神经网络、Transformer 架构，大幅提升了目标检测、语义分割的精度，成为过去十年智能视觉落地的主流技术，但始终未突破“被动感知”的技术边界。

近年来，AIGC 技术的快速成熟推动智能视觉系统进入创新迭代高峰期。在技术层面，扩散模型逐步取代传统 GAN 模型，解决了传统生成模型训练不稳定、生成图像模糊、细节缺失等问题，实现了高清图像、动态视频、三维场景的高精度生成；视觉语言多模态大模型实现了文本、图像、视频的跨模态语义对齐，让智能视觉系统具备了“看懂、读懂、创作”的综合能力 [4]。在产业层面，AIGC 已广泛渗透智能视觉核心应用场景，工业领域实现缺陷样本智能生成与工业视觉检测优化，自动驾驶领域完成复杂路况场景仿真与感知模型迭代，传媒领域实现智能剪辑、视觉内容自动创作，医疗领域完成医学影像修复与辅助诊断，技术落地场景持续拓宽。

整体而言，当前 AIGC 与智能视觉系统的融合已完成技术适配验证，进入规模化创新落地阶段，但技术成熟度、规范体系、产业适配性仍处于完善过程，技术创新潜力尚未完全释放 [4]。

二、AIGC 技术推进智能视觉核心影响

AIGC 技术对现代化智能视觉系统的创新是较为全方位、系统性的重构了传统视觉系统的算法架构、数据体系、感知逻辑与应用模式，从核心技术、功能形态、应用场景、产业模式四个维度推动智能视觉系统实现跨越式创新。

（一）重构视觉数据体系，摆脱大量数据集依赖

AIGC 生成式解决了传统视觉模型样本量不足、数据采集难度大以及模型泛化能力差等问题，依托深度学习模型的自主学习与生成能力，可基于少量真实样本，智能生成海量高仿真、多样化的虚拟视觉数据，涵盖常规场景、极端场景、小众特殊场景等各类数据类型。在工业质检场景中，AIGC 可自主生成各类零部件缺陷图像，包括划痕、变形、裂纹、色差等罕见缺陷样本，弥补真实生产中缺陷样本稀缺的问题；在自动驾驶视觉感知场景中，可生成暴雨、大雾、夜间逆光、道路遮挡等极端路况视觉数据，丰富模型训练数据集。表1为传统智能视觉与 AIGC 智能视觉生成量化对比。

表1 传统智能视觉与 AIGC 智能视觉量化指标对比

评估维度	传统判别式智能视觉	AIGC 生成式智能视觉	改善幅度
稀缺样本检测精度	62%	94%	+32%
单项目数据标注成本	100%	38%	−62%
极端场景算法稳定性	41%	87%	+46%
算法研发周期	100%	45%	−55%
场景自主适配能力	3.3	9.0	+5.7

同时，AIGC 可实现现有视觉数据的智能优化，通过图像超分辨率重构、噪声去除、缺陷修复、细节补全技术，提升原始采集图像的质量，解决传统视觉系统图像采集模糊、噪声干扰、细节缺失等问题。此举不仅大幅降低了视觉数据采集与标注的人工成本、时间成本，更完善了视觉数据体系的完整性，从数据源头推动智能视觉算法精度与稳定性的大幅提升，为技术创新奠定核心基础。

（二）算法架构革新，实现感知能力跨越式升级

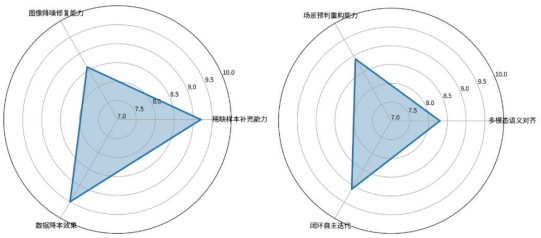
传统智能视觉算法以判别式模型为核心，核心逻辑是“输入图像－提取特征－判别分类”，仅能完成被动的感知与识别任务，无法实现对视觉场景的深度理解、逻辑推理与内容创造，感知维度单一、认知深度不足 [5]。AIGC 技术与视觉算法的融合，重构了智能视觉系统的算法架构，构建起“感知－理解－推理－生成－迭代”的闭环算法体系，实现了视觉认知能力的全方位升级。

首先，实现从“单维感知”到“多模态认知”的升级。传统视觉系统仅能解析图像、视频单一视觉信息，而 AIGC 多模态大模型实现了文本、视觉、语音、三维空间信息的深度融合，可通过文本指令驱动视觉场景重构、通过视觉画面反向生成语义描述，实现跨模态交互理解。

其次，实现从“被动识别”到“主动预测与重构”的升级。传统视觉系统只能对已发生、已存在的视觉场景进行识别分析，不具备预判与重构能力。基于 AIGC 技术的智能视觉系统，可通

Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.

过时序视觉数据学习，预判场景动态变化趋势，同时可对残缺、模糊、遮挡的视觉场景进行智能重构补全。图2为 AIGC 多维度性能分析雷达图。



（a）重构视觉数据指标雷达图　　（b）算法架构革新指标雷达图

图2 AIGC 多维度性能分析雷达图

三、AIGC 赋能智能视觉系统创新发展的现存挑战

尽管 AIGC 技术为现代化智能视觉系统的创新发展提供了强大动力，实现了技术、场景、产业的全方位升级，但当前二者的深度融合仍处于发展初期，在技术、安全、伦理、产业适配等多个层面存在诸多短板与挑战，制约了技术创新的深度推进与规模化落地[6]。

当前 AIGC 视觉技术仍存在明显的技术短板，难以完全满足现代化智能视觉系统高精度、高实时性、高稳定性的应用需求。首先，生成内容精准度不足，现有扩散模型、多模态模型在复杂场景、高精度细节生成中，容易出现纹理错乱、细节失真、语义偏差等问题，在工业精密质检、医疗影像诊断等高精度场景中，生成数据与真实场景存在误差，无法直接用于算法训练与精准检测。其次，实时性短板突出，AIGC 大模型参数规模庞大、算力消耗极高，现有硬件设备难以支撑高密度、高实时性的视觉生成与分析任务，无法适配自动驾驶、实时监控等对延迟要求极高的场景。

数据是 AIGC 视觉技术创新的核心载体，当前数据层面的版权争议与安全风险成为行业发展的重要阻碍。一方面，AIGC 视觉模型的训练依赖海量互联网图像、视频数据，其中大量数据涉及原创版权，但现有技术与法律体系尚未明确 AIGC 训练数据的版权使用边界，模型生成的视觉内容也存在权属界定模糊的问题，容易引发版权纠纷，制约商业化落地。另一方面，数据安全与隐私泄露风险突出。智能视觉系统广泛采集公共场景、个人场景的图像、视频数据，部分 AIGC 模型在训练与生成过程中，可能无意识复刻训练数据中的隐私信息，导致人脸信息、场景隐私、商业涉密视觉数据泄露。图3为 AIGC 推动智能视觉现存挑战饼状图。

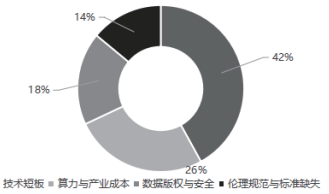


图3 基于 AIGC 智能视觉现存挑战饼状图

四、AIGC 赋能智能视觉系统技术展望

结合当前技术发展现状与现存挑战，未来 AIGC 与现代化智能视觉系统的融合创新，将朝着技术高精度化、应用垂直化、生态规范化、产业普惠化的方向迭代，持续突破技术瓶颈、完善产业生态、规范行业发展[7]。

未来智能视觉 AIGC 技术将聚焦核心技术短板，实现全方位优化升级。一是推进模型轻量化迭代，通过模型蒸馏、量化压缩、结构优化等技术，降低 AIGC 视觉模型的算力消耗，提升模型推理速度，解决实时性不足的问题。二是提升生成精度与细节表现力，优化扩散模型与多模态大模型的算法逻辑，强化复杂场景细节生成能力、语义理解精准度，消除画面失真、语义偏差等问题。三是强化小样本、零样本学习能力，优化生成式模型的训练逻辑，降低对海量标注数据的依赖，进一步解决小众场景、稀缺样本场景的视觉感知难题，全面提升模型泛化能力。

五、总结

AIGC 技术的快速发展与迭代，为现代化智能视觉系统的技术创新提供了颠覆性动力，彻底打破了传统智能视觉系统数据依赖强、感知能力弱、场景适配差、功能形态单一的技术局限，从数据体系、算法架构、应用场景、产业模式四个维度推动智能视觉领域实现全方位革新。通过生成式数据建模、多模态认知融合、主动场景重构、智能内容创作，AIGC 技术有效提升了智能视觉系统的感知精度、泛化能力与场景适配性，大幅拓展了技术应用边界，推动智能视觉产业降本增效、转型升级。

总体而言，AIGC 赋能智能视觉技术创新是人工智能产业发展的必然趋势，随着技术的持续迭代、规范体系的不断完善、产业生态的逐步成熟，AIGC 将全面重构现代化智能视觉产业格局，推动智能视觉技术深度赋能各行各业，为数字经济、智能社会建设提供坚实的技术支撑。

参考文献

[1] 张倩 . 基于 AIGC 技术的计算机视觉系统图像生成方法研究 [J]. 信息记录材料 ,2025(05):89-91.

[2] 王冬雪 . 生成式人工智能在视觉图像设计中的应用研究 [D]. 兰州：兰州财经大学 ,2025.

[3] 中国图象图形学报编辑部 . AIGC 视觉内容生成与溯源研究进展 [J]. 中国图象图形学报 ,2026(02):210-225.

[4] 李浩然，张敏 . 基于扩散模型的视觉图像视频生成技术综述 [J]. 计算机辅助设计与图形学学报 ,2024(11):1980-1992.

[5] 陈宇轩 . AIGC 驱动下智能视觉感知系统的技术变革与应用创新 [J]. 人工智能与大数据 ,2025(03):45-52.

[6] 刘安安，苏育挺，王岚君 . AIGC 视觉内容生成与溯源研究进展 [J]. 中国图象图形学报 ,2024,29 (06)：1535 - 1554.

[7] 孙文倩 . 人工智能技术辅助下的图像生成艺术创作研究 [D]. 上海：华东师范大学 ,2023.

 Copyright © This Work is Licensed under A Commons Attribution-Non Commercial 4.0 International License.