人工智能驱动的数字图像艺术创作：方法与案例分析

王常圣
（世宗大学，韩国 首尔 05006）

摘　要：近年来，人工智能技术的快速发展为数字图像创作领域提供了新的工具。使用AI工具进行数字图像创作逐渐成为一种常见的创作方式。在此背景下，通过访谈、调查与案例分析的方法，分析了人工智能驱动的数字图像艺术创作方法与案例。调查结果显示，大多数受访者认为AI工具对其创作过程产生了显著的影响，尤其是AI图像直出的方法被广泛运用于数字图像艺术创作中。通过对案例进行详细剖析，揭示了基于AI的数字图像艺术创作的典型范式，并将之归纳为“AI艺术创作工作流”。其包括AI图像直出法、AI辅助绘图法和ControlNet精确绘图法。这些方法可以有效地替代数字图像创作中的重复环节并提高出图效率，辅助数字图像的绘制和突破常规出图思路，实现精确可控的数字图像艺术内容表达。

关键词：人工智能艺术；数字图像；生成艺术；AI绘画；绘画生成
中图分类号：TP39
文献标志码：A
doi: 10.11959/j.issn.2096−6652.202333

AI-driven digital image art creation: methods and case analysis

WANG Changsheng
Sejong University, Seoul 05006, Republic of Korea

Abstract: In recent years, the rapid advancement of artificial intelligence technology has introduced novel tools for the realm of digital image creation. The use of AI tools for digital image creation has gradually become a common practice. In this context, this paper analyzed the methods and cases of AI-driven digital image art creation through interviews, surveys, and case analyses. The survey revealed that the majority of respondents believed that AI tools had a significant impact on their creative process, especially the widespread use of direct output from artificial intelligence methods in digital image art creation. By analyzing the cases in detail, this paper revealed the typical paradigms of AI-based digital image art creation and summarized them as the “AI art creation workflow”, which included AI direct output, AI assisted drawing, and ControlNet precision drawing. These methods can effectively replace repetitive steps in digital image creation and improve image output efficiency, assist in digital image drawing and break through the conventional drawing ideas, and achieve precise and controllable expression of digital image art content.

Key words: artificial intelligence art, digital image, generative art, AI painting, painting generation

0　引言

近年来，人工智能艺术广泛应用于人工智能生成内容（AIGC）领域[1]。自Dall-E和Disco Diffusion首次发布以来，这些创新的AI绘画工具已经颠覆了人们对人工智能生成数字图像可能性的认知。随后推出的Midjourney和Stable Diffusion则进一步提升了AI绘画的艺术效果。新一代文本到图像模型利用大型语言模型的进步，通过描述性文本提示能够创建高清真实的图像[2]。随着对这些模型的广泛访问，实践社区（由使用AI绘画技术的艺术家、设计师和研究者组成的一个网络社区）应运而生。人们能够在社区中共享设计、提示词和示例图像，进一步推动了AI绘画技术的发展和迭代[3]。

收稿日期：2023−04−20；**修回日期：**2023−08−09

由于AI表现出强大的图像生成能力，可以大幅提高内容创作效率[4]，AI逐渐成为创作数字艺术图像的流行手段[5-6]。研究者们前瞻性地讨论了AI艺术技术应用的潜力[7-8]。有的研究者探索了人工智能在艺术设计中的应用[9]；还有研究者探讨了人类与人工智能绘画的比较[10]。同时，越来越多的人将AI绘画应用在商业领域中。这也引发了关于人工智能对艺术家和当代艺术影响的讨论[11]。AI艺术家与人类艺术家在艺术市场上的竞争正在成为现实[12]。

在此背景下，自然而然地产生了一些需要解决的问题：人工智能是否能成为数字图像创作中的重要工具，助力创作者达成更高效、更具创新性的创作？在现今使用人工智能工具进行数字图像创作时，存在哪些合适的创作方式？如何运用人工智能工具，直接或间接地解决数字图像创作中的实际需求？

本文采用访谈、调查和案例研究的方法，对相关使用者进行了深入研究，以了解他们常用的AI工具及创作方式。通过案例分析归纳了基于人工智能生成的数字图像创作新范式“AI绘画工作流”。其包含了AI图像直出法、AI辅助绘图法和Contorl-Net精确绘图法3种当下常用的数字图像创作方法。该方法能够有效地替代数字图像创作中的重复环节并提高出图效率，辅助数字图像的绘制和突破常规出图思路，实现精确可控的数字图像艺术内容表达。本文旨在为未来数字图像创作提供新的思路和方法，推动AI艺术的进一步发展和创新。

根据Robert A.Gonsalves使用CLIP模型进行的对比相似性试验和当下常用AI工具，本文主要采用美学质量较好的Midjourney和Stable Diffusion进行后续的案例研究。

1　人工智能绘画的发展历程

人工智能在绘画领域探索的早期阶段可以追溯到20世纪70年代。艺术家Harold Cohen尝试着开发名为“AARON”的程序，通过这个程序，他指导机械臂进行绘画[13]。AARON程序基于规则和概率，可以生成复杂的几何形状、颜色和图案。然而，机械臂作画也存在一些问题。首先，机械臂的运动是刚性的，可能无法像人类艺术家那样表达自由流动的线条和手绘的感觉。其次，程序生成的图像比较简单，无法绘制出效果出众的图像作品。尽管如此，机械臂作画还是在一定程度上推动了计算机艺术的发展。

2012年，吴恩达和Jeff Dean利用Google Brain创建的大型神经网络（由1万个CPU构成），生成了一张模糊的猫脸图像。这项工作在同年的国际会议上公布，并引发了广大的关注与讨论[14]。使用神经网络的自动化学习方法，能够在没有预先设置的特征或规则下，自动识别和分类图像中的对象。这表明，对复杂的任务使用神经网络进行自动化学习可能比人类设计的算法更加有效[15]。

2014年，AI理论学界出现了1个重要的深度学习模型：生成对抗网络（generative adversarial network，GAN）。该模型由生成器和判别器组成，以对抗性学习的方式生成数据[16-17]。它能从大量图像数据中学习并复制图像的分布特性，从而生成高质量的逼真图像。这种能力使得GAN可以广泛应用于各类绘画作品的生成，包括风景、人像和抽象画等。

尽管GAN在艺术创作领域展现了优势，但其限制性也不容忽视。GAN生成的图像可能存在一定的失真和噪声，这在生成高分辨率图像时尤为明显。处理复杂场景或绘画作品时，GAN可能会产生不真实或不合理的生成效果，这时就需要人工干预或应用其他算法进行纠正或优化。但GAN很快成为人工智能中最重要的研究领域之一，并且出现了许多原始架构的高级和特定领域的变体[7]，例如CycleGAN[18]、StyleGAN[19]或BigGAN[20]。

OpenAI在2021年推出了对比语言−图像预训练（contrastive language-image pretraining，CLIP）模型，其创新之处在于其同时理解图像和文本内容。通过对大量的图像和描述性文本进行训练，CLIP模型学会了如何将图像和文本映射到相同的向量空间中。在这个空间中，语义上相关的图像和文本彼此接近，而不相关的则相隔较远。这种方法使得模型能够理解自然语言命令，并根据这些命令生成或修改图像[21]。随着扩散模型被证明可以生成高质量的合成图像[22]，首个基于CLIP+扩散模型的实用AI绘画产品—Disco Diffusion在此后得以出现[23]。其相对于之前的AI绘画工具展现出强大的美学表现力、无条件的开放性和更强的可扩展性[11]。

Disco Diffusion在数字插画和传统绘画领域表现相当不错，其鲜艳的色彩和耀眼的造型效果具有很强的视觉冲击力，能够满足大多数人的艺术欣赏

和艺术创作的需求[11]。然而，尽管 Disco Diffusion 在绘画应用中表现出色，但也存在一些明显的局限性。这包括在刻画具体细节方面的能力相对较弱，渲染时间较长，以及生成的画面可能存在割裂和拼凑的现象。近期“Latent Diffusion Models”的研究为这些问题提供了新的解决路径。该方法将原本需要在像素空间花费大量算力进行训练的扩散模型转移到了对算力要求小得多的“Latent Space”上进行训练，同时保留了模型生成图像的质量和灵活性[2]。这使得如 Midjourney 等 AI 绘画工具能够更快速地生成高质量的图像。

审视人工智能绘画的发展历程，我们可以观察到技术的进步如何推动 AI 绘画工具的发展和迭代，并为 AI 艺术的实践打下了坚实的基础。每一次技术的跃进都为我们的创作实践开启了新的可能性。这些技术的进步为我们的研究提供了深厚的技术支持。

2 研究设计

作者采用了访谈、调查和案例分析的方法进行研究。作者向 AI 艺术自媒体创作者、独立艺术家、数字图像创作专业人员以及使用人工智能技术进行艺术创作的研究者和学生发送了一项调查，收集有关他们使用 AI 绘画的信息（使用的 AI 模型工具、创作方法、动机及影响）。作者还进行了 11 次关键信息者访谈（KII）和 4 次深入的半结构访谈（IDI）。参与者包括 15 位高级使用者，目的是探究他们在使用 AI 绘画工具生成数字图像时的创作方式和体验。研究主要包含以下几个方面的内容。

（1）访谈调查：在研究分析过程中，作者对使用 AI 绘画模型进行图像创作的 15 位高级使用者（9 位男性，6 位女性）进行了观察和访谈。通过 40~60 分钟的访谈来探讨他们使用 AI 绘画模型创作的主要工具、方法和体验。参与者是匿名的，但都是具有使用 AI 绘画工具创作的基本能力和半年以上使用 AI 绘画工具经验的人。

（2）参与者：参与者的招募采用了多渠道策略，以确保样本的多样性和代表性。首先，我们通过社交媒体平台及自媒体平台（如微信、抖音、小红书）发布了招募广告，目标是那些在 AI 艺术领域进行图像创作的自媒体创作者。我们还联系了几个与 AI 艺术和数字图像创作相关的在线社区和论坛，请他们在其成员中进行宣传。其次，我们通过与高校教师合作，向天津美术学院等高等艺术院校发送了招募通知，以吸引从事 AI 艺术研究的学者和学生。

（3）访谈结构：该研究由 3 个部分组成。第 1 部分，每个参与者首先被问到他们进行 AI 图像创作时，会选择什么样的工具和方式进行创作，让研究人员观察他们的自然选择。第 2 部分，在这项研究中，参与者被要求反思他们选择这种创作方式/创作工具的理由和使用的创作方法。在访谈的第 3 部分，参与者被要求反思选择不同的 AI 创作方式对创作效率和创作质量的影响。

（4）访谈数据定性分析：两位审查员针对访谈录像内容进行了独立审查和数据抄录。在此过程中，分别对每份抄本进行分析并提取要点。同时对这些研究者在访谈中创作的数据和提示词进行了分析研究，以此形成对抄录数据的解释。针对性地做了关于数据的差异和分歧解释。

（5）数据收集整理与案例分析：为了更好地进行研究，作者从网络上收集了大量基于 AI 创作的数字图像作品。这些作品覆盖了各种艺术风格。同时，在对这些作品进行初步观察和分类后，将其整理成一个包含详细信息的数据库，为后续的案例分析和研究提供便利。

3 调查结果与案例分析

3.1 调查结果

作者收到了 213 份调查回复，经过数据整理和筛选，186 份为有效答卷。在问题“你目前主要在使用哪种人工智能工具？（Midjourney、Stable Diffusion、其他）”的回答中，超过半数的使用者使用 Midjourney 作为 AI 出图工具，使用 Midjourney 和 Stable Diffusion 的受访者占 95.7%，见表 1。

表 1 AI 工具使用情况

AI 工具	使用人数	占比
Midjourney	103	55.4%
Stable Diffusion	75	40.3%
其他	8	4.3%

在问题“AI 工具介入数字图像创作后，对其未来创作方式的影响？（非常有影响、一般影响、没有影响）”的回答中，84% 的受访者认为非常有影响，见表 2。

表 2　　AI 工具的影响力

影响程度	回答人数	占比
非常有影响	162	84.4%
一般影响	23	12.4%
没有影响	6	3.2%

在问题“使用这些 AI 工具进行绘图和创作时，常用哪种方式进行绘制（AI 直接出图、AI 辅助绘图、ControlNet 精确绘图），每种方式写出常用的 AI 工具（Midjourney、Stable Diffusion）并给出理由”中，127 人（68.3%）选择了 AI 直接出图，其中 105 人（82.6%）选择使用的工具为 Midjourney，见表 3。他们表示“AI 直接出图能够替代重复的工作，提高出图的效率。但 AI 直接出图比较困难，需要较高的提示词写作水平、较好的基础模型效果并熟悉自己想要实现的图像风格”。23 人（12.4%）选择了 AI 辅助绘图。其中 19 人（82.6%）选择 Stable Diffusion。他们表示“AI 辅助绘图能够提供更多的绘画及设计灵感，可以在工作中作为较好的辅助工具”。36 人（19.3%）选择了 ControlNet 精确绘图。他们表示“通过 Stable Diffusion 的 ControlNet 插件能够很好地解决工作中的具体问题。Stable Diffusion 直接出图效果不如 Midjourney，但搭配 ControlNet 能够满足工作中复杂多变的商业需求”。

在选择适合特定项目的工作流程时，应综合考虑各种因素，包括生成效果、时间成本和灵活性等。为了更深入地理解这些工作流程的特性和适用性，我们对其优缺点进行了比较分析。在表 4 中，总结了每种工作流的优点及缺陷。

表 3　　AI 工作流的使用状况

AI 工作流	使用人数	占比	Midjourney 使用人数	Midjourney 使用人数占比	Stable Diffusion 使用人数	Stable Diffusion 使用人数占比
AI 图像直出法	127	68.3%	105	82.6%	22	17.4%
AI 辅助绘图法	23	12.4%	4	17.4%	19	82.6%
ControlNet 精确绘图法	36	19.3%	N/A	N/A	36	100%

表 4　　AI 工作流的比较分析

AI 工作流	优点	缺陷
AI 图像直出法	快速生成高质量图像；方法简单	需要多次生成以获得理想结果；需要较高的提示词写作水平、较好的基础模型效果并熟悉自己想要实现的图像风格
AI 辅助绘图法	拓展创作思路；灵活性高	需要较好的审美筛选生成图像；需要对 AI 技术有一定的理解和操作能力
ControlNet 精确绘图法	极高的美学表达上限；精确控制图像生成	较高的技术门槛；需要更多的时间来调整和优化图像效果

3.2　案例分析

经过对收集的 AI 绘画的数据图库内容的分析、访谈调查和案例分析，作者梳理出当下基于 AI 绘制数字图像的常见创作方式，并将其总结为“AI 绘画工作流”。

第一种常见的工作流程是 AI 图像直出法。受访者调查显示：127 人（68.3%）选择了 AI 直接出图。在使用这种方法的过程中，由于随机性的存在，往往需要多次生成才能得到期望的结果。为了产生高质量的图像，通常需要重新生成多次，找到好的关键词能够比社群流行的关键词产生更好的结果[24]。例如，在图 1 中，两张图像都是使用同一组关键词



图 1　AI 直接生成插画风景作品（图片由作者绘制）

生成的。左边的图像为了得到连贯的天空背景，经过了7次同一组关键词的反复生成才得到。但在构图及内容的完整度上和右图差距较大，因为右边的图像是在左图的基础上“再变化”23次得出的结果。所以好的生成图像也是需要反复“打磨”的。

然而，有时即使进行多次重新生成，也无法得到满意的效果，这时就需要调整提示词的设定。如图2所示，通过不断修改左图的提示词，最终实现了右图丰富的色彩静物油画效果。这个方法非常考验AI使用者的想象力（构建头脑中的画面）、精准的词语表达能力（准确地表达脑海中的画面）、良好的审美观（在不断重新生成的图像中寻找合适的明暗、色彩和构图）以及丰富的经验（理解和应用提示词搭配和艺术家风格）。



图2 AI直接生成静物油画作品（图片由作者绘制）

第二种是AI辅助绘图法，受访者调查显示：23人（12.4%）选择了这种工作方式。这种模式有效地弥补了AI在逻辑上的误差以及在直接生成图像中的细节不足。艺术家“影依望远镜”就使用AI辅助创作了高质量的风景插图（如图3所示）。更值得一提的是，他在如何利用AI辅助图像整合的问题上进行了深入的思考和实践。

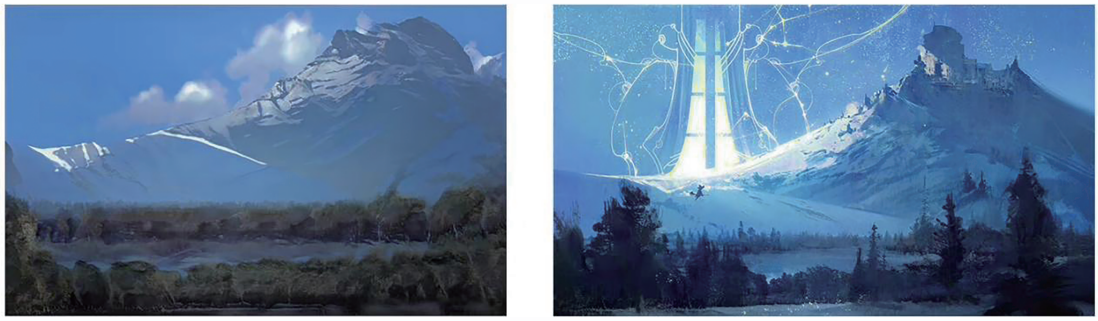


图3 AI辅助绘制风景插画（图片来源：影依望远镜）

如图3所示，左图（AI直接生成的图像）的氛围宁静，色彩以蓝色为主，且构图相当不错。然而，这幅画面缺乏动人的元素，细节部分的节奏需要梳理，同时也需要对构图进行优化。在图3的右图（作者修改后的图像）中，我们可以看到作者的思考和情感的注入。他将“天梯”元素引入画面，强化了图像中的山背光线，使其成为画面的中心（增加了光影的对比），并去除了天空中的云朵，修改了山和前景树木的构图，提升了整个画面的色彩明度和对比度。通过艺术家自身的美学修养、情感以及绘画技巧，原本平淡的AI生成图像具备了更唯美的画面效果。

图3所展示的绘制方法运用了归纳、增减和衍生的手法。另一位艺术家“Nekro”在AI辅助绘图中探索了一种新方法，他试图通过AI生成的零件来组成一幅完整的画面。作为一位才华横溢的数字艺术家，Nekro的绘画风格被训练到了他所使用的人工智能模型中。他运用AI创建了一幅具有“Nekro风格”的插画作品（见图4）。在Midjourney平台生成的“零件”（见图4左图）经过Nekro的筛选和图像裁剪，形成了一系列可以组合的局部元素。随后，他利用数字绘图工具将这些局部元素整合成

一个完整的画面，调整细节，最终完成了一幅数字插画作品。他在描述这个过程时表示：“我从人工智能中获取我需要的零件，然后构建我想要的东西。”因此，AI辅助图像整合模式可被视为一种新的工作流程，用于辅助数字图像的绘制和设计，预计在未来将成为设计师和插画师们常用的工作流程之一。



图4　AI辅助绘制数字插画（图片来源：Nekro）

在辅助绘图的方法中，还存在一种门槛相对较高的方式，即灵感提供模式。它可以成为艺术家开辟新的创作视野的强大工具和灵感源泉[25]。当需要寻找灵感时，艺术家可以使用AI生成工具，进行多种图像的生成，以探索各种可能性。在探索主题、灵感、氛围、构图、色彩等能够引发共鸣的元素的同时，艺术家可以根据灵感的触动，进行深入的创作。艺术家“乌合麒麟”采用了这种方式，为我们提供了一个参考案例（见图5）。



图5　AI提供灵感绘制插画（图片来源：乌合麒麟）

在微博中，他表达了使用人工智能的动机：“画画一直是谁画的好就学谁，觉得AI画的好，就把它好的部分学过来。”这表明他对有助于自我提升的技术持开放的学习态度。他也通过实际行动将AI生成的图像作为创作灵感，绘制出表现力强烈的图像。在利用AI工具辅助创作过程中，他能够从AI生成的图像中提取氛围、解构构图模式、学习色彩搭配，再将这些元素融入自己的创作中。这种方式有助于创作者突破常规的思维模式。然而，这也要求创作者保持对新事物的敏感度和探索精神，了解领域内的前沿知识，并善于收集相关资料，以增强自己的综合创作能力。

第三种是Controlnet精确绘图法。主要通过搭载在Stable Diffusion上的ControlNet插件来实现精确化绘图及设计。受访者调查显示，36人（19.3%）选择了ControlNet精确绘图。在Zhang等人[26]的研究中，他们介绍了多种ControlNet的实现方式，这些方法涵盖了不同的图像生成任务。主要内容包括：（1）Canny边缘检测，用于提取线稿并上色；（2）HED边界检测，适用于人像线稿提取；（3）MLSD模型，专注于建筑和室内设计的线稿提取；（4）Hough线检测，适用于直线边缘提取；（5）用户涂鸦模型，模拟人类绘制的线稿；（6）基于Openpifpaf和Openpose的人物姿态估计模型，用于生成特定姿态的人物图像；（7）语义分割模型，应用于数据集COCO和ADE20K，生成细粒度的场景图像；（8）基于Midas的深度图模型，用于生成具有空间层次感的图像；（9）法线图模型，用于生成表面法线信息的图像；（10）卡通线稿提取模型，适用于动漫插画的线稿生成[26]。这些模式展示了ControlNet在不同类型的图像生成任务中的广泛适用性和强大的能力，作者选择受访者主要使用的4种模式进行案例的分析与讲解。

（1）基于Canny边缘检测和HED模型提取线稿的能力：Canny边缘检测在处理普通图片时表现出良好的效果，能够捕获图像中的边缘特征并生成相应的线稿。相较之下，HED模型能够生成更柔和的线条，特别适用于人像。通过这两种线稿提取方法，ControlNet可以根据不同类型的图像生成对应的线稿，并在此模式下实现线稿上色的功能。在图6中，我们可以看到一个基于Canny边缘检测的线稿上色案例。这个例子展示了Canny模式下通过对线稿上色来生成具有丰富细节和良好视觉效果的Q版形象插图。

（2）通过提取人物骨骼生成姿势的Openpose是最为常用的功能：通过学习姿态估计方法，ControlNet可以捕捉到人物的关键点，从而生成具有特定姿态的人物图像。这种方法在人物动作设计中具有广泛的应用价值。例如，图7中展示了一个基于



图6　使用Canny模式实现线稿上色（图片来源：网易CFun设计中心）

姿态绑定的应用案例。在这个例子中，根据给定的人物骨骼生成了一个与输入姿态相符的二次元拿着工具的少女形象。这个工具是AI根据人物动态补充上去的，同时保持了人物形象动态的自然和连贯性。这种姿态生成能力为动画制作、游戏角色设计等领域提供了新的可能。

（3）MLSD功能在建筑设计方面的应用：MLSD

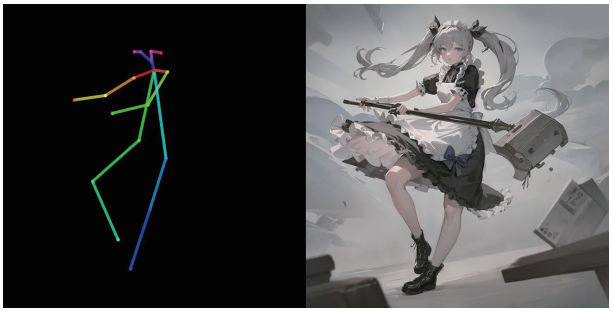


图7　使用Openpose模式实现骨骼姿势生成图像（图片来源：@hoge-hogeo）

是一种特殊的线稿提取方法，非常适用于建筑和室内设计。通过捕捉建筑物和室内空间的几何结构，MLSD能够生成具有清晰线条和明确轮廓的线稿。这种方法为建筑师和设计师提供了一个直观的视觉参考，有助于更好地生成建筑物数字图像。图8展示了MLSD提取的线稿生成具有高质量和真实感的建筑设计图。画面在线稿结构的基础上补充了更多的氛围和细节，从而体现了MLSD功能在建筑设计及效果图领域的潜力和表现力。

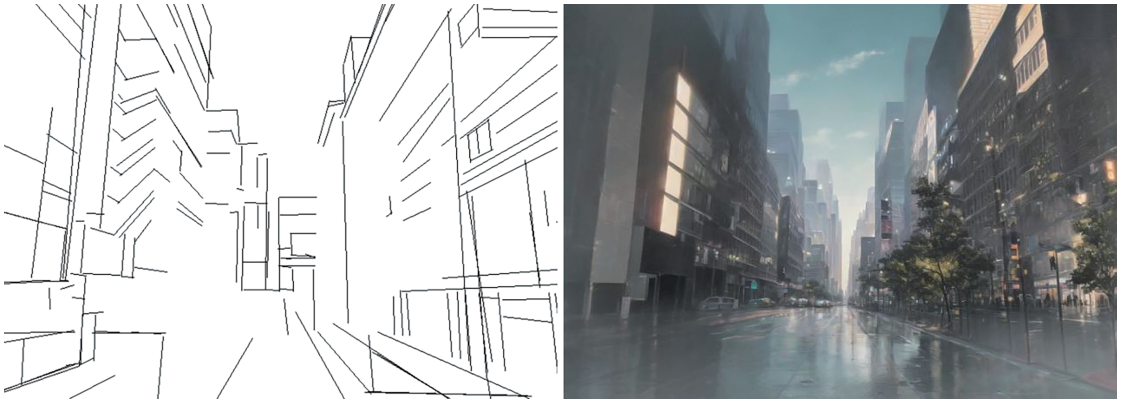


图8　使用MLSD模式生成建筑效果图（图片来源：@AlternaAssets）

（4）深度图控制（Depth）的能力：在深度图的模式中，可以更好地理解图像的空间结构和层次关系，从而生成具有空间感和真实感的场景。这种方法在室内设计、建筑领域、游戏及插画场景设计等需要空间层次的图像中具有较大的潜力。例如，图9展示了一个基于深度图控制生成的全景空间场景。在这个例子中，ControlNet根据输入的深度图生成了一个具有明确空间层次的场景，展现了其在处理复杂场景时的强大能力。这种深度图控制技术可以帮助设计师更直观地呈现空间层次和效果。除了以上ControlNet常用的模型之外，还有其他适合于精确化控制图像生成的功能。通过对ControlNet的研究和应用，我们可以预见，这种基于精确化功能控制的框架将在艺术与设计的众多图像处理领域发挥更大的作用，并在商业化图像处理上提供可靠的帮助。

3.3　案例反思与未来挑战

经过对各类AI绘画工作流案例的深入分析和理解，可以从用户需求、内容精准生成和原创问题这3个关键领域反思现有的工作流，并探讨未来的挑战和可能的解决策略。

首先，从用户需求的角度，我们的案例显示，艺术家们根据自身的需求选择使用不同的AI工具和创作方法。例如，选择AI直接出图的艺术家重

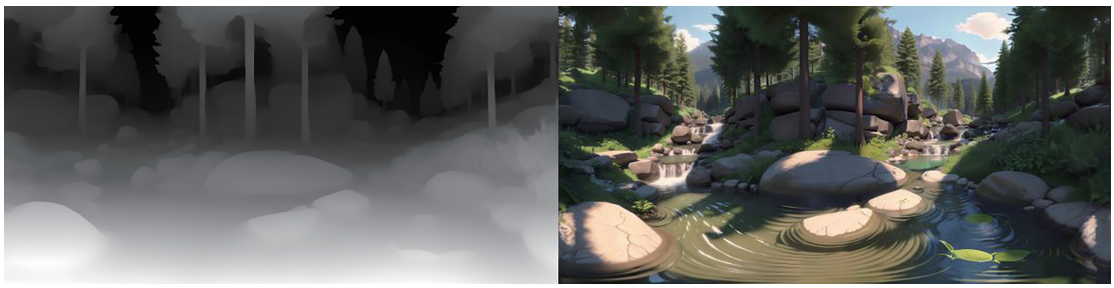


图 9　使用 Depth 模式输出全景图（图片来源：@mickmumpitz）

视该方法在提高出图效率和替代重复工作方面的优势。相反，倾向于使用 AI 辅助绘图和 ControlNet 精确绘图的艺术家则更看重 AI 工具在提供创作灵感和精准控制上的能力。这些反馈揭示了艺术家的需求是多样化的，而且可能会随着他们对 AI 工具熟悉度的提升而变得更为复杂。因此，AI 工具的开发者需要深入理解艺术家的需求，以提供更加贴切的服务。

然后，我们从内容精准生成的角度进行反思。无论是 AI 直接出图还是 AI 辅助绘图，都需要生成具有吸引力的高质量图像。然而，由于 AI 生成的图像具有一定的随机性，艺术家们通常需要进行多次尝试才能得到满意的结果，这无疑增加了创作的难度和时间成本。因此，未来的挑战在于如何进一步提高 AI 工具的生成质量和可控性，以实现更精准的内容生成。

最后，我们需要面对原创问题。在本文的案例中，艺术家们利用 AI 工具创作出了独特和新颖的图像，扩展了艺术的边界。然而，这也引发了关于 AI 生成艺术的原创性的讨论。在现有的工作流中，AI 工具主要作为辅助工具，帮助艺术家生成图像，而真正的创作灵感和决定仍然来自艺术家。因此，我们需要深入理解 AI 在艺术创作中的角色，思考如何评价 AI 生成艺术的原创性，以及如何在尊重和保护艺术家原创权益的同时，推动 AI 在艺术创作中的应用。总的来说，我们需要从多个角度深入理解和反思现有的 AI 驱动的数字图像艺术创作的工作流，以应对未来的挑战，推动这一领域的发展。

4　结论

4.1　研究结论

本文以人工智能技术在数字图像创作领域的应用为切入点，通过访谈、调查问卷和案例分析的方法，分析了当下人工智能驱动的数字艺术图像创作方式及案例特征，并得出以下几点结论。

（1）人工智能技术的快速发展为数字图像创作领域带来了新的可能性，AI 绘画将成为一种具有广泛应用前景和价值的创作方式。在这种情况下，人们会调整他们的创造性实践和语言，以促进与 AI 的有效交互，因为这是日常生活中需要的一项技能[27]。利用 AI 来提升创作效率和质量，将成为艺术设计创作者的“必修课”之一。

（2）根据访谈、调查和案例研究，本文总结当下人工智能驱动的数字艺术图像创作常见方式——“AI 绘画工作流”，包括 AI 图像直出法、AI 辅助绘图法和 ControlNet 精确绘图法。这些方法可以有效提高出图效率，辅助数字图像的绘制，实现精确可控的数字图像艺术内容的表达。

4.2　未来方向

本文在分析人工智能驱动的数字艺术图像创作方式及其应用前景时，尚存在一定的局限性。例如，在实证研究方面，本文主要采用访谈和问卷调查的方式收集数据，可能会受到样本量和受访者背景等因素的影响。在 AI 技术快速发展和迭代的背景下，AI 图像的创作方式亦会因为技术变化而有所不同。

在未来的研究中，我们将继续关注 AI 绘画技术的发展和进步，以期在这一领域取得更深入的理论和实践成果。人工智能技术在数字图像创作领域的应用已经初步显现出巨大的潜力和价值，这将是艺术设计创作者和充满丰富想象力之人最好的时代。AI 绘画技术有助于推动数字艺术创作的创新和发展，并帮助每一个“有脑洞”、有艺术追求、有探索欲望的人在其中获得属于自己的“艺术图像自由”。我们需要关注 AI 技术在艺术领域中的不断突破和进步，同时关注与之相关的伦理、版权等问题，以期在新技术的推动下，实现数字艺术的繁荣

和进步。

致谢：笔者衷心感谢所有参与此研究的受访者，他们的时间和投入对于本研究至关重要。感谢那些通过问卷调查提供信息的个人和组织，他们的贡献提高了笔者的数据质量和丰富性。此外，对所有为本研究提供支持和建议的专家表示深深的谢意。他们的专业知识和反馈对笔者的工作产生了深远的影响。

参考文献：

[1] CAO Y, LI S, LIU Y, et al. A comprehensive survey of AI-generated content (AIGC): a history of generative AI from GAN to ChatGPT [EB]. arXiv preprint, 2023, arXiv: 2303.04226.

[2] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution image synthesis with latent diffusion models[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2022: 10674-10685.

[3] CHANG M, DRUGA S, FIANNACA A J, et al. The prompt artists[C]// Proceedings of the 15th Conference on Creativity and Cognition. New York: ACM, 2023: 75-87.

[4] 武强, 季雪庭, 吕琳媛. 元宇宙中的人工智能技术与应用[J]. 智能科学与技术学报, 2022, 4(3): 324-334.
WU Q, JI X T, LYU L Y. Artificial intelligence technology and applications in the metaverse[J]. Journal of Intelligent Science and Technology, 2022, 4(3): 324-334.

[5] RUNCO M A, JAEGER G J. The standard definition of creativity[J]. Creativity Research Journal, 2012, 24(1): 92-96.

[6] CROWSON K, BIDERMAN S, KORNIS D, et al. VQGAN-CLIP: open domain image generation and editing with natural language guidance[EB]. arXiv preprint, 2022, arXiv: 2204.08583.

[7] CETINIC E, SHE J. Understanding and creating art with AI: review and outlook[J]. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 2022, 18(2): 1-22.

[8] HONG J W, CURRAN N M. Artificial intelligence, artists, and art: attitudes toward artwork produced by humans vs. artificial intelligence[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2019, 15(2): 1-16.

[9] DENG Y. Application of artificial intelligence in art design[C]//Proceedings of 2021 International Conference on Computer Technology and Media Convergence Design (CTMCD). Piscataway: IEEE Press, 2021: 200-203.

[10] SUN Y K, LYU Y R, LIN P H, et al. Comparison of cognitive differences of artworks between artist and artistic style transfer[J]. Applied Sciences, 2022, 12(11): 5525.

[11] LI L. The impact of artificial intelligence painting on contemporary art from disco diffusion's painting creation experiment[C]//Proceedings of 2022 International Conference on Frontiers of Artificial Intelligence and Machine Learning (FAIML). Piscataway: IEEE Press, 2022: 52-56.

[12] TAE H S, KIM S Y. Exploratory review on the convergence of artificial intelligence and art[J]. Official Journal of the Korean Society of Dance Science, 2019, 36(2): 27-42.

[13] MALINA R F, MCCORDUCK P, COHEN H. Aaron's code: meta-art, artificial intelligence and the work of Harold Cohen[J]. Leonardo, 1991, 24(5): 628-629.

[14] LE Q V. Building high-level features using large scale unsupervised learning[C]//Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2013: 8595-8598.

[15] MIIKKULAINEN R, LIANG J, MEYERSON E, et al. Evolving deep neural networks[M]//Artificial Intelligence in the Age of Neural Networks and Brain Computing. Amsterdam: Elsevier, 2019: 293-312.

[16] WANG K F, GOU C, DUAN Y J, et al. Generative adversarial networks: introduction and outlook[J]. IEEE/CAA Journal of Automatica Sinica, 2017, 4(4): 588-598.

[17] LI X R, LI H, WANG C Y, et al. Visual-attention GAN for interior sketch colourisation[J]. IET Image Processing, 2021, 15(4): 997-1007.

[18] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2017: 2242-2251.

[19] KARRAS T, LAINE S, AILA T M. A style-based generator architecture for generative adversarial networks[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2019: 4401-4410.

[20] BROCK A, DONAHUE J, SIMONYAN K. Large scale GAN training for high fidelity natural image synthesis[EB]. arXiv preprint, 2018, arXiv: 1809.11096.

[21] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[EB]. arXiv preprint, 2021, arXiv: 2103.00020.

[22] NICHOL A, DHARIWAL P, RAMESH A, et al. GLIDE: towards photorealistic image generation and editing with text-guided diffusion models [EB]. arXiv preprint, 2021, arXiv: 2112.10741.

[23] DHARIWAL P, NICHOL A. Diffusion models beat GANs on image synthesis[C]// Proceedings of Advances in Neural Information Processing Systems. [S.l.:s.n.], 2021: 8780-8794.

[24] PAVLICHENKO N, USTALOV D. Best prompts for text-to-image models and how to find them[C]//Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2023: 2067-2071.

[25] CHOI S Y. A study on AI art in the post-digital age[J]. The Journal of Aesthetics & Science of Art, 2022, 67: 188-214.

[26] ZHANG L, RAO A, AGRAWALA M. Adding conditional control to text-to-image diffusion models[EB]. arXiv preprint, 2023, arXiv:2302.05543.

[27] OPPENLAENDER J, LINDER R, SILVENNOINEN J. Prompting AI art: an investigation into the creative skill of prompt engineering[EB]. arXiv preprint, 2023, arXiv:2303.13534.

[作者简介]



王常圣（1995− ），男，世宗大学公演、影像、动画系博士生，主要研究方向为人工智能艺术和AI绘画。