

# 算法与艺术之争

## ——康德美学视域下AI生成艺术的局限性

裴琮德, 罗盈盈

西安电子科技大学人文学院哲学系, 陕西 西安

收稿日期: 2025年5月13日; 录用日期: 2025年6月6日; 发布日期: 2025年6月17日

### 摘要

AI生成艺术的广泛运用在拓展艺术边界的同时, 也引发了其作品是否具有审美价值、对人类主体性的削弱等一系列问题。康德在《判断力批判》中对审美判断的论述揭示了AI生成艺术在数据学习阶段的功利性局限并缺少先验结构; 天才理论揭示了其在算法演绎阶段的机械性局限与心意精神的匮乏; 崇高理论揭示了其缺少主体理性的自我肯定与自由意志的升华。通过上述分析可以明确, AI生成艺术的根本局限在于其主体性的缺失, 真正的艺术创作是主体的自由表达。分析AI生成艺术的局限并非是为了否定它, 而是为了更好地理解并合理利用这一技术力量。在技术飞速发展的当下, AI与艺术的结合是必然趋势, 但在这一结合过程中, 应维护人类的主体地位, 避免让艺术沦为算法的机械产物。

### 关键词

AI生成艺术, 康德美学, 主体性

# The Algorithm vs. Art Debate

## —The Limits of AI-Generated Art in the Context of Kant's Aesthetics

Congde Pei, Yingying Luo

Department of Philosophy, School of Humanities, Xi'an University of Electronic Science and Technology, Xi'an Shaanxi

Received: May 13<sup>th</sup>, 2025; accepted: Jun. 6<sup>th</sup>, 2025; published: Jun. 17<sup>th</sup>, 2025

### Abstract

The extensive use of AI-generated art expands the boundaries of art while raising a series of questions such as whether its works have aesthetic value and the weakening of human subjectivity. Kant's discussion of aesthetic judgment in *Critique of Judgment* reveals the utilitarian limitation and lack of a priori structure of AI-generated art in the stage of data learning; the genius theory reveals

its mechanical limitation and lack of mindfulness in the stage of algorithmic deduction; and the sublime theory reveals its lack of self-affirmation of subjective rationality and sublimation of free will. Through the above analysis, it is clear that the fundamental limitation of AI-generated art lies in its lack of subjectivity, and true artistic creation is the free expression of the subject. Analyzing the limitations of AI-generated art is not to negate it, but to better understand and reasonably utilize the power of this technology. In the rapid development of technology, the combination of AI and art is an inevitable trend, but in this combination process, the subjective position of human beings should be maintained, to avoid letting the art be reduced to the mechanical product of the algorithm.

## Keywords

AI-Generated Art, Kant's Aesthetics, Subjectivity

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

2022 年，由 Midjourney 生成的图画《Théâtre D'opéra Spatial》在美国的艺术博览会上获得数字艺术类金奖引发轩然大波，更有艺术家声称：“我们正在见证艺术的死亡。”AI 生成艺术在提升效率、降低门槛的同时，也对传统的艺术创作方式及人类艺术家的主体地位提出了重要挑战。康德美学作为近现代美学的奠基性理论，将审美判断与艺术创作定位于人类主体的先验能力。这种人类主体性美学，为分析 AI 生成艺术的局限性提供了理论依据。现有在康德美学视域下对 AI 生成艺术的研究较少，且鲜有结合 AI 技术原理的研究。本文将从技术层面入手，系统性地结合康德美学理论来分析 AI 生成艺术的根本局限。

## 2. AI 生成艺术能否分辨什么是美的

### 2.1. 数据学习与特征归纳

AI 生成艺术的起点是对大量的现有艺术作品进行“学习”。与人类通过观察、模仿前人的艺术作品后再创作不同，AI 通过深度学习技术，将海量艺术作品分解为数据，再归纳其特征。

AI 生成艺术所需要的数据库需要涵盖多个历史时期的艺术风格，涉及多种艺术流派，并且还需要包括多样的艺术形式(如视觉艺术需要包括绘画、雕塑、摄影等)。除此之外，还需要结合创作背景、艺术评论等文本数据，使其建立风格与语境的关联。由于 AI 生成艺术的数据学习本质上是对已有经验的统计归纳，数据库的完善程度便决定了其生成能力。例如，现有主流模型如 DALL·E、Midjourney 等，其数据库主要收录各种经典西式风格，而对东方艺术则涉及较少，这类数据问题会直接限制其作品的多样性。在数据初步收集完成之后，AI 会依靠多层神经网络对这些数据进行标准化处理(数据清洗)。以图片数据为例，AI 会进行调整分辨率、统一色彩空间、去除背景噪声等工作，以确保输入数据的一致性及可计算性。之后，数据增强(Data Augmentation)技术通过对原始数据进行几何变换、颜色调整、噪声注入等操作，来增加训练样本的多样性。但是将艺术作品按照流水线形式标准化，会丧失其中具有艺术价值的很多细节。尽管可以通过数据增强技术进行一定弥补，但这种基于数学变换的增强方式仍局限于既有样本的变形，无法突破数据库的界限。

特征归纳是指 AI 从收集到的海量数据样本中归纳出不同艺术风格的特征。当前最常见的技术是卷积

神经网络(Convolutional Neural Networks), CNN 通过一系列的卷积层和池化层逐步提取图像的低层特征(如边缘、角点)和高层特征(如物体轮廓、结构), 最终形成一种风格编码。特征提取不仅关注作品“是什么”, 也试图判断“像什么”。例如, 通过比较某幅画的笔触密度、色彩布局和形状配置, AI 可以识别出这幅作品更接近印象派还是立体主义风格。处理文本类数据最常用到的是自然语言处理(NLP)模型, 如 Transformer 架构, 来识别关键词、风格描述和评价倾向, 从而将作品融入到更丰富的语境中。但是, 尽管这类技术可以高度模拟人类对图像和语言的分析过程, 但这本质上还是统计意义的“相似性”, 即在大量样本的基础上, 通过误差反向传播和梯度下降算法等, 不断调整参数以最小化预测误差。在这个过程中, 各类艺术风格被还原为可量化的视觉参数或语义向量, 丧失了许多艺术应有的情感、精神。

## 2.2. 功利性与目的导向的局限

康德在《判断力批判》中提出了审美判断的四个契机。其中第一契机“无兴趣的愉悦”和第三契机“无目的的合目的性”共同构成了审美判断的核心特质, 二者之间具有密切的内在关联。第一契机强调审美判断应当是“无利害的”。在传统美学中, 美的价值往往与伦理、宗教或政治等外部因素相联, 而康德则试图在审美领域构建一种超越功利目的的判断模式。康德说, “鉴赏是通过不带任何利害的愉悦或不悦而对一个对象或一个表象方式作评判的能力。一个这样的愉悦的对象就叫作美” [1]。他认为, 真正的美感不应依赖于个体的欲望、利益或特定目标, 应是一种纯粹的愉悦。这种愉悦既非感官上的快感, 也非道德上的满足, 而是主体自身的自由体验。康德通过剥离审美愉悦的经验成分, 将其定位于主体的先验能力, 从而确立了审美判断区别于实践判断或认知判断的独特地位。第三契机“无目的的合目的性”则进一步深化了审美判断的内涵。这一契机想要说明, 美的对象虽然在形式上展现出某种“合目的性”, 但这种合目的性并不指向任何外在的实用目标。康德说: “美是一个对象的合目的的形式, 如果这形式是没有任何目的的表象而在对象身上被知觉到的话” [1]。因此, 审美判断既不同于单纯的感官快感(因为感官快感通常具有生理或心理上的目的性), 也不同于理性的判断(因为理性判断的依据往往在于其实际功效)。在康德看来, 美的对象之所以令人愉悦, 是因为它在形式上体现了一种“自由的和谐性”, 即它虽看似有目的, 但实际上并未指向任何具体的实际目的。

相较于人类的审美判断, AI 生成艺术的“学习”本质上是数据驱动的机械性分析, 其训练有明确的目标, 强调对用户偏好或市场趋势的迎合[2]。这种技术路径不以“美为何物”为出发点, 而是以功利性的评价标准作为分析目的, 从根本上背离了“无利害性”和“无目的的合目的性”。在具体的学习阶段, AI 的训练通常围绕预设目标展开, 例如提升作品的接受度或满足特定审美风格。通过深度神经网络, AI 从大量样本中提取色彩、构图、纹理等视觉特征, 并构建潜在空间中的概率分布, 用于预测性生成。并且在训练过程中, AI 通过优化损失函数, 使数据库中的标签更贴近收集到的用户反馈与市场流行审美。这种建立在“目标拟合”基础上的技术行为, 本质上是一种受功利驱使的“合目的”行为。正因如此, AI 的数据归纳难以满足康德对纯粹美感体验的要求。

## 2.3. 普遍必然性的先验缺失

康德认为, 审美判断在主观上具备“无概念的普遍性”与“无概念的必然性”——即第二契机与第四契机所揭示的审美判断的深层结构。二者分别对应“量”与“模态”的范畴, 还构成了一种从心理状态到先验基础的内在关联: 第二契机强调的是审美判断作为一种“人人可有”的感性愉悦, 而第四契机则进一步揭示这种共通性背后的先验结构。康德在第二契机中指出, “凡是那没有概念而普遍令人喜欢的东西就是美的” [1]。真正的审美判断具有“无概念的普遍性”, 也就是说, 当我们对某一对象作出“这是美的”判断时, 虽然这种判断不是建立在客观概念或逻辑推理之上, 但我们却天然地期待他人也能作

出相同的判断。这种期待并不是出于经验上的共识，而是出自一种审美上的共通感(*sensus communis*)，它并非经验概括，而是人类感性与理性结构的一种共鸣。而第四契机则从“模态”出发，强调审美判断的“无概念的必然性”。康德认为，“美是那没有概念而被认作一个必然愉悦的对象的东西”[1]。当我们对某物感到美时，这种愉悦不是偶然的、任意的，而是在主观上具有必然性。也就是说，我们不仅期待他人会赞同这种判断，还会认为这种赞同是理所当然的。这种必然性并非逻辑上的证明，也非经验统计的结果，而是知性与想象力之间的“自由游戏”的结果。因此，第二契机与第四契机共同描绘了审美判断从共通感出发，抵达一种带有理性根据的普遍性必然状态。这一结构完全独立于具体经验或概念，是人类先验审美能力的表现。

尽管 AI 能够通过深度学习技术对海量艺术作品进行模式识别与风格特征提取，但这一过程本质上仍是数据驱动的统计演算。它所得出的“艺术规律”并非来自于对美的直观感受或主观判断，而是基于经验样本的概率分析[3]。这种方式或可生成表面上具有形式美感的作品，却无法涉及先验层面的共通感。在 AI 的学习过程中，艺术作品被简化为数据向量，其审美价值的高低则由高频特征的出现概率来决定。它无法理解为何某些构图、色彩或节奏会引发人类普遍的情感共鸣，也无从把握艺术背后的精神维度。它只会将“共通感”简单地与数据库中统计到的最大概率划等号，认为出现频率最高的元素就具有普遍必然性。这种以概率代替“共通感”的方式，弱化了主体在艺术创作中的价值，更无法生成出真正具有普遍必然性的艺术作品。

### 3. AI 生成艺术能否像人一样创作

#### 3.1. 算法模型的演绎生成

数据库搭建完成后，AI 生成艺术可以通过各类算法模型，依据用户指令，将归纳好的艺术特征转化为新的作品。与传统人类艺术家的创作方式不同，AI 并非从情感、经验或直觉中发掘灵感，而是依托于数据库，通过数学模型生成符合特定风格与内容要求的图像、文本或音频作品。

当前主流的生成模型多采用生成对抗网络(*Generative Adversarial Networks, GANs*)、变分自编码器(*Variational Autoencoders, VAEs*)、扩散模型(*Diffusion Models*)等。以 *GANs* 为例，其核心为生成器与判别器的对抗训练。生成器负责不断优化图像以“骗”过判别器，判别器负责判断该图像是否为“真”艺术作品。经过这两方的多次博弈，最终便能生成高度仿真的艺术作品。但这种优化实质上还是使生成的作品不断趋近于数据库中已有的数字特征，无法跳出这个框架。

近年来兴起的扩散模型(如 *Stable Diffusion*)则采用先加噪再去噪的方式，在训练阶段，AI 学习如何把清晰图像“污染”成噪声图；而在生成阶段，再一步步“去除噪声”，直到还原出一幅清晰的艺术作品。这个过程类似于画家在模糊的轮廓中不断勾勒细节，逐渐构建出完整画面。相比于生成对抗网络，扩散模型更稳定、效果更好，在生成细节、融合风格等方面也表现出更强的能力。然而其计算成本较高，推理速度慢，可控性仍然有限。

变分自编码器(*VAEs*)采用编码器—解码器结构，将图像压缩到一个低维潜在空间，并通过解码器重建图像。在这个过程中，VAE 引入概率建模的方式，使编码带有一定的随机性，从而增加生成图像的多样性。但尽管它能在潜在空间中采样新内容，但其生成仍依赖于训练数据，无法脱离既有界限，并且生成的图像往往较为模糊，细节表现不如 *GANs* 和扩散模型。

#### 3.2. 规则内的机械执行

康德说：“天才就是：一个主体在自由运用其诸认识能力方面的禀赋的典范式的独创性”[1]。独创性和典范性是天才的两大基本特征。天才的独创性有两层含义，第一层是强调天才创造的艺术作品是不

可重复的、独一无二的,是“空前而绝后”的发明。第二层含义则强调,天才的创作并非依循既定法则,而是超越法则自成一格。康德明确指出:“天才是与模仿的精神完全对立的”[1]。但康德同时认为,即使是独创性的作品,有时也可能是毫无价值的,甚至是怪诞离奇的东西。“所以天才的作品同时又必须是典范,即必须是有示范作用的;因而它们本身不是通过模仿而产生的,但却必须被别人用来模仿,即用作评判的准绳或规则”[1]。这就是典范性。典范性也包括两方面的内容:一是强调天才的作品对其他天才的“触发”或“唤醒”作用;二是为艺术立法,成为普通人评判艺术作品的范例。

AI生成艺术本质上是一种高度程序化的演绎逻辑。虽然它表面上展现出了各类丰富的变化,但其行为仍是在规则系统内部进行拟合操作[4]。天才之所以“独创”,在于其作品不是经验归纳的产物,而是天才超越法则、自我立法的结果。相较而言,AI的“创作”不过是在既定规则内部进行最优演绎,其生成过程主要是对训练数据中的高频样本进行拆分重组。这种生成并非出于“自由游戏”,而是受限于模型内部的目标函数与经验数据。它不能主动提出问题、开创风格,也就失去了成为“天才”所需的重要特征。AI更无法“为艺术立法”。康德所说的典范性不仅是美的体现,更重要的是它能为后来的艺术创作提供标准与方向,唤起他人的创造力。而AI所生成的作品,由于本身是在模仿已有范式的基础上进行重组,它的“风格多样”不是新风格的生成,而是既有风格的聚合,因此并不具备成为某种艺术的典范、成为评判标准的可能。AI生成艺术注定只能是再生生产者,而非真正的创造者。

### 3.3. 心意精神的匮乏

康德从天才的功能入手来论述其心理机制,认为天才的主要功能在于赋予艺术作品以精神或灵魂。他认为若缺乏“精神”,艺术作品便难以激发人们内心深处的共鸣。“精神,在审美的意义上,就是指内心的鼓舞生动的原则”[1],并强调艺术作品真正的生命力源自于创作者内心灌注的生气和情感。正是这种“心意能力”,使得艺术作品不只是形式与技巧的堆砌,而是承载着一种无法复制的人类情感和审美体验。艺术作品的精神就在于它能否唤起观者的内心的共鸣,使之产生一种超越感官享受的心灵体验。

人工智能在生成艺术作品时,缺乏人类创作者那种自发的、具有情感温度和主观体验的内在驱动。艺术中的精神源于创作者心灵的表达和情感的流露,而这种自由的心意活动能够在艺术作品中注入生气,使之超越形式上的美感。相比之下,AI在生成过程中只是通过参数调整和概率采样完成图像、文字或音乐的组合,其输出仅仅是对大量已有数据的统计重组,根本没有内在的情感体验和主观意向。这种“缺乏心意”的创作方式,使得AI艺术始终难以摆脱一种机械化、公式化的特征,其作品往往缺少那种能够触及人心、引发深层次情感共鸣的精神内核。此外,艺术创作中的心意精神还体现在对情感、意蕴和生命体验的独到理解与再现上。人类艺术家在创作时不仅借助理性和技巧,更依赖于对自身经历、情感体验及对世界的独特洞察,从而赋予作品独特的生命力。而AI在创作时,既无生命体验,也无主观情感,因而无法从内心深处发起对“生命精神”的呼唤。它只能根据既定数据与模式“模拟”情感表达,而这种情感模拟缺乏真正的主体体验,难以形成具有深刻精神内涵的艺术范式。

## 4. AI生成艺术能否理解崇高

### 4.1. 数学崇高的算法局限

康德关于数学崇高的理论强调,在量的层面上,崇高表现为绝对的大,即一切与之较量的对象都显得微不足道。数的崇高要求把无限纳入整体直观,但想象力只能把握有限事物,当面对超出感性尺度的量时,想象力即便竭尽全力,也难以完成直观综合。理性却有能力将这一无限整体内在化为一种主体机能,从而“以自身为尺度去估量无限的整体”。“想象力和感性的无力和局限,恰好反映和证明了超感性的理性能力的高超和优越,感性层次上的不合目的性,恰好显示了理性层次上的合目的性,并且是更高

的合目的性”[5]。正是这种理性对感性局限的超越,使主体体验到一种优越感和尊严感。这种超越感并非来源于感官直觉的宏大,而是一种内在的理性升腾,是理性与自由意志对“无限”的把握和自我肯定。

崇高作为一种内在体验,首先必须被主体感受到。而人工智能在其生成机制中完全缺乏“感受”的能力。AI并没有“惊叹”“敬畏”“震撼”等情绪机制,更无法意识到某种量的无限性与自身处理能力的冲突,因此也无法在经验之限与理性之超越之间建立起崇高的心理张力。它对“巨大”的判断只是一种对像素规模、比例、结构复杂性的数值识别,是从技术角度的模拟,而非心理意义上的“超越”。并且,即使假设AI可以从数据中“识别”出具有崇高特征的图像,它也无法将这种感受转化为富有精神内涵的艺术作品。在康德看来,真正的艺术作品必须融合心意精神,是创作者自由意志的体现。而AI生成的图像,仅是对数据中“宏大”元素的统计重组。例如通过分形算法、高维插值或图像拼接实现“宏伟壮观”的图像输出,但这种输出的背后并无主体性驱动、无自我意识与审美升华,无法像人类艺术家那样将崇高作为精神力量注入作品之中。

## 4.2. 力学崇高的算法局限

康德所定义的力学崇高,强调自然之“力”的震撼性场面如何唤起人的理性尊严。康德指出:“自然界当它在审美判断中被看作强力,而又对我们没有强制力时,就是力学的崇高”[1]。当面对自然界浩瀚无边的力量时,人们初感恐惧,因为这种力量对感性而言是压倒性的。但正是在这种恐惧之中,主体的理性力量得以显现,当我们处于安全状态下,面对更为可怕的景象时,内心反而涌起一种超越平常的精神力量和自豪感。这种精神上的抵抗能力,构成了人类尊严与自信的根基,也使得力学崇高成为一种对自身本性和使命的肯定。

AI作为非感性存在,其生成过程并不包含对自然威力的主观体验。人类之所以能在面对自然巨力时产生崇高感,是因为他们感知到了“可以压倒自身的外部力量”,进而在对比中唤醒内在的理性自由与道德尊严。而AI并不会感到“恐惧”或“威胁”,也没有“反抗”的能力,它无法经历从“害怕”到“超越”的心理跃迁。对AI而言,自然的力量只是数据与模式的集合,无法引发心灵震撼,也无法产生康德意义上的理性自觉。即便AI试图通过算法模拟自然景象,呈现如风暴、地震、火山爆发等壮观场面,其生成也仅限于形式上的宏大重现。AI依据训练数据,通过图像合成、特效增强或分形模拟等技术生成看似“震撼”的视觉图像,但这种图像只是对自然“威力外形”的再加工,缺乏主观体验注入后所产生的精神升华。它无法将“对力量的感知”转化为“对人的自由”的尊重与赞颂,也就不能真正唤起力学崇高所蕴含的道德精神。

## 5. 结论

回顾前文对AI生成艺术各个维度的分析,都表明它无法突破数据的框架与算法的限制。如果将其与康德所批判的经验论美学与唯理论美学进行类比,就会发现它与这两种近代美学范式具有高度相似的结构困境。经验论美学强调感官经验,把美的标准简化为愉悦或快感的刺激。在这种框架中,艺术不过是感性材料的排列与组合,美感只是源于被动的感觉愉悦。AI生成艺术的图像训练与内容输出同样建立在大量感性数据的收集与分析之上,通过识别人类审美偏好的图像特征,不断优化其生成结果,以最大程度触发观众的视觉愉悦,但其本质不过是对“快感模式”的技术复制。相对地,唯理论美学则试图借助理性的逻辑推演给出美的普遍法则,将美作为可以被归纳、分析和定义的理性对象。AI生成艺术的背后正是这类逻辑结构的延续:算法在构建图像时依赖数学模型、损失函数与规则约束,其“创作”更像是对设定审美公式的参数化执行。无论是形式比例的对称平衡,还是色彩的协调计算,最终都服从于技术理性框架下的最优化操作。

但正如康德所指出, 真正的审美判断并非建立在经验的感性刺激, 也非归结于理性的法则推演, 而是主体在自由状态下无目的地体验。这一判断不仅是一种超越实用功能的自由游戏, 更是一种唤起想象力与理性协同作用的精神活动。AI 的“生成”过程则恰恰缺乏这种自由判断力: 它不具有主体意识, 也无法从内部经验出发进行自我目的的设定与超越, 因此, 其创作始终停留在形式的表象之上, 无法触及康德意义上“自由心灵”的内在维度[6]。康德所强调的“人本主义的主体性美学”本质上就是这种自由精神的确立。艺术不只是信息的组合与表达, 更是自由主体对世界的再创造, 是情感、意志、理性三者统一的精神生产。正是在这种意义上, 艺术才不仅是“做出来的”, 更是“活出来的”。而 AI 生成艺术作为一种没有生命体验、没有意志意图、没有自由判断的工具性技术, 其所谓“创作”不过是被动地执行人类预设的逻辑指令与数据算法, 是对作品形式的生产而非具有意义的创造[7]。

因此, AI 生成艺术的根本局限不在于它的技术尚未成熟, 而在于其主体性的缺失, 它在对客观经验的鉴赏学习和生成艺术作品这两个角度都无法成为像人类艺术家一样的审美主体, 它只能被视为一种工具。但与传统的艺术创作工具不同, 艺术家的主体身份在使用 AI 生成艺术时会从“创作者”逐渐转变为“策展人”, 主体性的创造活动被压缩为一系列输入与筛选的技术操作。这不仅改变了艺术创作的方式, 还潜移默化地影响着人类艺术家对自身创作能力的认知, 过度依赖于 AI 生成艺术更会削弱人类自身的主体性。

AI 的发展与应用是时代进步的必然趋势, 分析其局限性不是要否定它, 而是为了更好地理解、掌握并合理地运用这种技术手段。在智能时代, 美学的真正挑战, 不在于 AI 技术本身, 而在于人类如何在与技术的协作中守护自身的主体地位, 并发挥艺术的真正价值。

## 基金项目

中西哲学核心命题的当代诠释与比较研究(项目编号: YJSJ25015)。

## 参考文献

- [1] 康德. 判断力批判[M]. 邓晓芒, 译. 北京: 人民出版社, 2017.
- [2] 杨毅凌. 人工智能的作品存在与艺术市场研究[J]. 中国市场, 2024(9): 119-122.
- [3] 王赠怡. 大数据思维的方法论意义与审美认识的转向——以美本质追问的困惑为例[J]. 天府新论, 2016(4): 127-133.
- [4] 张爱军, 郭镇毓. 生成式人工智能的技术神话、祛魅与认知进路——以 Sora 模型为例[J]. 理论建设, 2025, 41(1): 74-89.
- [5] 曹俊峰. 康德美学引论[M]. 天津: 天津教育出版社, 2001.
- [6] 陈海静. 人工智能能否成为审美主体——基于康德美学的一种扩展性探讨[J]. 学术研究, 2022(7): 149-157.
- [7] 卢卫红, 杨新福. 人工智能与人的主体性反思[J]. 重庆邮电大学学报(社会科学版), 2023, 35(2): 85-92.