

# 第5章

---



## 图生图： 智能重构与衍生

通过学习本章内容，读者将全面了解图生图技术的核心原理与应用方法，掌握如何通过AI工具对图像进行智能重构、风格迁移、局部编辑及动态化处理。本章将深入讲解不同图生图模式的技术特点、操作流程及优化策略，并结合实战案例，帮助读者熟练运用通义万相、豆包、即梦AI、可灵等工具，实现从静态图像到创意衍生作品的转化，提升设计效率与艺术表现力。

## 5.1 图生图核心技术

图生图是指利用人工智能技术，将输入图像转换为另一张具有特定风格、内容或结构的输出图像的技术。

### 5.1.1 传统修图与AI图生图

传统修图依赖人工操作（如使用Photoshop软件），而AI图生图通过算法自动生成图像。本节从技术原理、操作方式、效率、灵活性及应用场景对二者展开对比，以帮助读者理解技术演进路径。

#### 1. 技术原理

传统修图的核心是基于像素级的手动编辑，依赖专业软件（如Photoshop）提供的工具集，包括图层、蒙版、画笔、滤镜等。用户需要对图像的每个细节进行逐一调整，技术门槛较高，且效果直接取决于操作者的技能水平。

相比之下，AI图生图依托深度学习算法，如生成对抗网络（GAN）和扩散模型（Diffusion Model），通过大量数据训练实现图像的自动生成或修改。用户仅需输入提示词，算法即可根据语义理解生成符合要求的图像，极大地降低了技术门槛。

#### 2. 操作方式

传统修图需手动逐一调整各项参数。例如在调整图像色彩时，要手动设置色相、饱和度、明度等参数，如图5-1所示；处理图像清晰度时，要手动调节锐化程度等。这高度依赖用户的专业技能与经验，操作者需熟悉软件工具的使用，并且具备一定的审美和图像处理知识，才能达到理想效果。

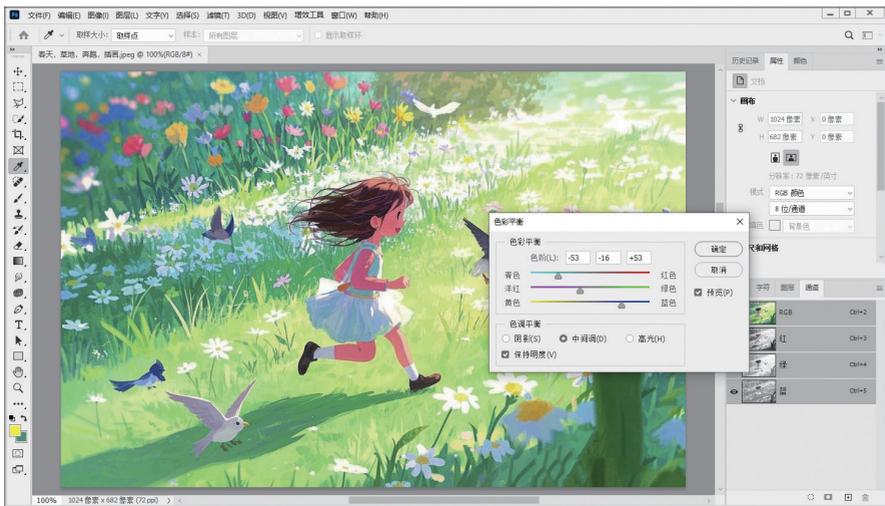


图 5-1

AI图生图则简化了这一过程，用户通过自然语言描述需求，如输入提示词：“将这幅图转换为3D风格”，系统即可快速输出结果，如图5-2所示。操作更加直观，但对提示词的准确性要求较高，且生成结果可能存在随机性。

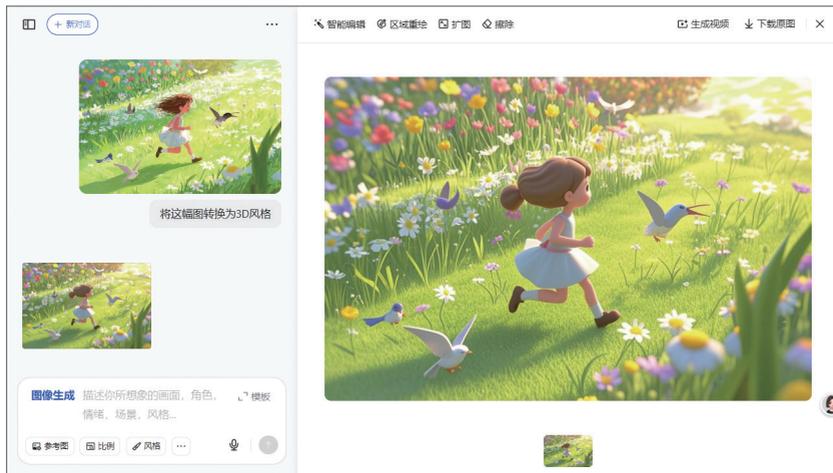


图 5-2

### 3. 效率

在效率方面，传统修图通常需要数小时甚至数天才能完成单张图像的高精度处理，尤其在艺术创作或商业项目中，细节打磨耗时较长。

AI图生图的优势在于其分钟级的生成速度，并能支持批量处理。例如，设计师可通过AI快速生成多套方案草图，显著提升前期构思的效率。然而，若需进一步调整细节，仍需结合传统修图工具。

### 4. 灵活性

传统修图的灵活性体现在对图像细节的绝对控制，用户可精确到像素级别进行修改，适合对精度要求极高的场景，如文物修复或广告大片制作。

AI图生图在风格迁移和创意发散上表现突出，能够轻松实现不同艺术风格的转换。但其局限性在于对特定细节（如人物五官、物体结构）的精准控制能力较弱，生成结果可能不符合预期，需反复进行调试提示词或后期人工修正。

### 5. 应用场景

传统修图广泛应用于商业摄影、高端艺术创作以及需要精细修复的领域，例如杂志封面制作或老照片修复，其成品质量稳定且可控。

AI图生图更适合概念设计、风格化内容生成以及实时需求场景。例如，游戏开发者可用AI快速生成角色原画，或自媒体创作者批量生产风格统一的配图。此外，AI在快速原型设计、虚拟场景构建等方面也展现出了巨大潜力。

## 5.1.2 图生图定义与核心原理

图生图是指利用人工智能技术，将输入图像（如照片、草图、低分辨率图像等）自动转换为符合特定要求的输出图像的过程。其核心目标是学习输入与目标图像之间的映射关系，实现风格转换、内容生成、修复增强等任务。与文生图不同，图生图的输入和输出均为图像，且更强调对现有图像的转换，而非从零生成。

图生图技术主要基于生成对抗网络、变分自编码器和扩散模型三类深度学习模型。

### 1. 生成对抗网络

通过生成器（Generator）和判别器（Discriminator）的对抗训练生成逼真图像，其优势在于生成速度快，适合风格迁移、图像修复等任务。

- **生成器**：学习输入图像的特征，生成与目标图像风格或内容相近的假图像。
- **判别器**：判断输入图像是真（来自数据集）还是假（由生成器生成）。
- **对抗训练**：生成器与判别器相互博弈，生成器不断提高生成图像的真实性，判别器不断提高判别能力，最终生成器能够生成高质量的图像。

### 2. 变分自编码器（VAE）

通过编码器—解码器结构学习图像的潜在表示，实现图像重建或转换。其优势在于生成过程稳定，适合图像去噪、低分辨率增强等任务。

- **编码器**：将输入图像压缩为低维潜在向量，捕捉图像的关键特征。
- **解码器**：从潜在向量中重建图像，生成与输入图像相似的新图像。
- **潜在空间**：通过学习潜在向量的分布，可以在潜在空间中进行插值或采样，生成多样化的图像。

### 3. 扩散模型

通过逐步去噪过程生成高质量图像，在图生图中常结合控制信号（如ControlNet）实现精确编辑。

**注意** 扩散模型的基础原理在文生图部分已详细介绍。

## 5.1.3 常见应用场景

图生图技术凭借其强大的图像转换和生成能力，已在多个领域得到广泛应用。以下是其主要应用场景。

#### 1. 影视后期制作

- **风格化特效**：通过AI算法将实拍画面转换为油画、水墨画等艺术风格，同时支持动态粒子效果渲染，为电影、动画提供极具创新性的视觉表现。
- **动态天气模拟**：能够实时生成带有体积雾粒子的雨雪效果，并且支持风力、密度、持续时间等参数调节，使场景更加真实细腻。
- **老片修复**：采用4K/8K超分辨率重建技术，智能还原经典影片的原始粒子噪点质感，完整保留胶片时代特色。

#### 2. 视频增强与修复

- **超分辨率重建**：基于深度学习算法，将低分辨率监控视频转换为高清格式，显著提升关键细节识别能力。
- **智能修复系统**：自动完成老旧影片的划痕修复、色彩校正和帧率提升，最高支持24~60fps

的流畅度转换。

- **动态补偿技术**：在体育赛事转播中实现动作细节的完美呈现，为观众带来极致流畅的观看体验。

### 3. 广告创意与营销

- **智能素材生成**：可快速产出高质量的产品宣传图、海报及横幅广告，支持用户实时修改画面元素、调整风格，大幅缩短广告制作周期，满足快速投放需求。
- **品牌视觉系统**：依据品牌DNA自动生成系列插画、Logo和IP形象，确保品牌在不同宣传场景下的视觉风格高度统一，强化品牌形象记忆点。
- **社交内容引擎**：为各社交平台定制化生成具有高互动性的内容，结合平台用户特点和流行趋势，有效提升用户的参与度和转化率，助力品牌社交传播。

### 4. 室内与建筑设计

- **虚拟样板间**：将设计草图一键转换为逼真的3D渲染图，支持用户进行多角度展示和实时修改，方便设计师与客户沟通，直观呈现设计方案效果。
- **材料替换模拟**：快速生成不同材质，如木地板、大理石的视觉效果。设计师和客户能够直观地比较不同材质在空间中的呈现效果，从而更科学地作出设计决策，避免实际施工时因材质选择不当而带来的问题。
- **光照模拟系统**：能够准确预测建筑空间在不同时段的光影变化，包括自然光和人工照明效果，辅助设计师做出最优的采光和照明设计决策。

### 5. 电子商务创新

- **虚拟试衣间**：基于人体工学算法，精准模拟服装上身效果，支持用户进行多角度查看，让消费者在线上也能获得接近线下试穿的体验，减少因尺码、款式不合适导致的退货率。
- **3D商品展示**：将静态商品图片转换为可交互的3D模型，支持360°旋转和细节缩放，全方位展示商品外观和细节，增强商品吸引力和消费者购买意愿。
- **智能包装设计**：根据商品特性，如尺寸、重量、形状等，自动生成个性化包装方案，在保证包装保护功能的同时，降低设计成本，减少材料浪费。

### 6. 文化遗产保护

- **高精度修复**：采用多光谱扫描技术，对老照片和壁画进行无损检测和修复，精准还原其原始色彩与细节，最大程度保护文化遗产的历史信息。
- **数字化存档**：将文物扫描数据转换为高保真3D模型，建立完整的数字资产库，为文物研究、保护和传承提供可靠的数字资料，避免文物因自然因素或人为因素损坏而造成信息丢失。
- **虚拟展览系统**：支持在线沉浸式观展，用户可通过VR设备或网页身临其境般欣赏文物，突破时空限制，实现文化遗产的广泛传播。

### 7. 个性化内容创作

- **社交媒体头像生成**：用户上传照片后，系统可快速生成动漫、水墨画等艺术风格头像，

满足用户在社交平台上展现个性的需求。

- **时光穿梭效果**：通过 AI 算法模拟“老年形象”“童年照片修复”等效果，增强用户互动乐趣，同时唤起用户情感共鸣，创造独特的社交分享内容。
- **家庭相册增强**：自动优化老照片的清晰度与色彩，去除照片中的划痕、污渍等瑕疵，让家庭记忆以更好的状态保存和传承。

## 动手练 使用豆包将人物照片转换为水墨画风格



下面使用豆包将人物照片转为水墨画风格，具体操作步骤如下。

**步骤 01** 打开豆包，进入工作界面，单击“图像生成”按钮，显示如图5-3所示的界面。



图 5-3

**步骤 02** 单击“参考图”按钮，在弹出的“打开”对话框中选择如图5-4所示的图像素材并上传。

**步骤 03** 返回“图像生成”界面，单击“风格”按钮，在弹出的菜单中选择“水墨画”选项，如图5-5所示。



图 5-4



图 5-5

**步骤 04** 继续在输入框中输入提示词：“保留原图像的主体元素”，如图5-6所示。



图 5-6

**步骤 05** 单击“发送”按钮, 系统将根据提示词进行生成, 效果如图5-7所示。



图 5-7

## 5.2 四大应用模式

图生图技术通过不同的输入输出组合（如图像+图像、图像+文本）实现多样化的图像转换需求。下面对图生图中的四大应用模式进行介绍。

### 5.2.1 风格迁移

风格迁移（Style Transfer）是指将原图像的内容（如物体、结构、场景）与目标风格图像的艺术特征（如笔触、色彩、纹理）相结合，生成具有新风格的图像，同时保留原图的核心内容结构。该技术广泛应用于艺术创作、广告设计、影视特效等领域。其技术原理如下。

#### 1. 内容提取

内容提取阶段借助预训练模型（如VGG-19）的深层卷积层（如conv4\_2）提取图像的高阶语义特征，例如物体的轮廓、空间布局等。深层卷积层因其较大的感受野和强大的抽象能力，能够捕捉图像中全局的、语义级别的特征信息，这些特征对于保留原图的核心内容至关重要。

#### 2. 风格提取

风格提取阶段利用同一网络的浅层卷积层来捕捉风格图像的纹理、色彩分布等特征，即通过计算Gram矩阵来量化风格特征。Gram矩阵反映了特征通道之间的相关性，能够有效表示图像的风格信息。浅层卷积层对图像的局部细节更为敏感，因此更适合提取风格特征。

#### 3. 融合生成

融合生成阶段将内容特征与风格特征进行有机结合，然后通过解码器生成目标图像。解码器通常采用反卷积网络（或转置卷积）将融合后的高维特征向量映射回像素空间，从而得到可视化的图像。在融合过程中需要精心设计损失函数。

- **内容损失（Content Loss）**：确保生成图像与原图像在深层特征上保持一致。

- **风格损失 (Style Loss)**：最小化生成图像与风格图像在Gram矩阵上的差异。

通过调整内容损失和风格损失的权重，可以平衡内容保留与风格迁移的程度，最终生成既具有目标风格艺术特色，又能保留原图核心内容的图像。

## 动手练 使用通义万相将风景照转为油画风格

下面使用通义万相将风景照转为油画风格，具体操作步骤如下。

- 步骤 01** 打开通义万相进入工作界面，单击“应用广场”按钮，显示如图5-8所示的界面。

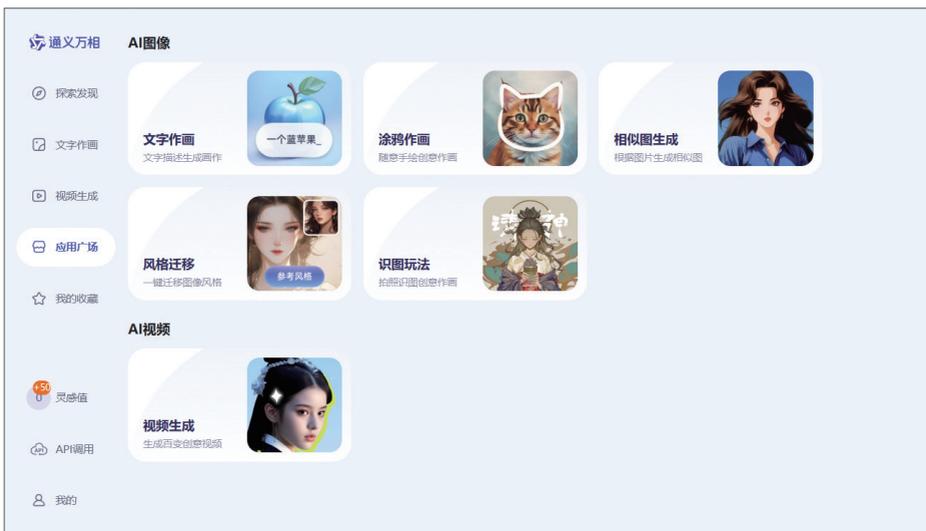


图 5-8

- 步骤 02** 单击“风格迁移”按钮，跳转至如图5-9所示的界面。

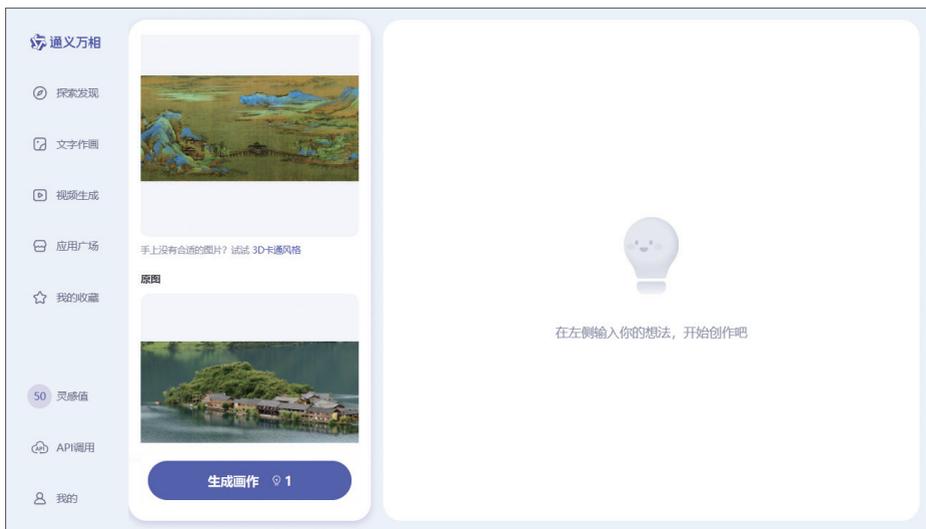


图 5-9

- 步骤 03** 在“风格图”处单击添加如图5-10所示的图像。在“原图”处单击，添加如图5-11所示的图像。

- 步骤 04** 单击“生成画作”按钮，效果如图5-12所示。

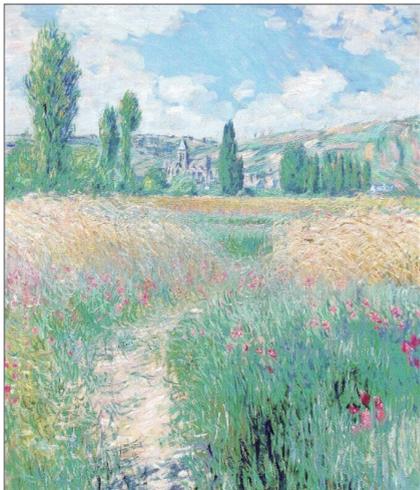


图 5-10



图 5-11

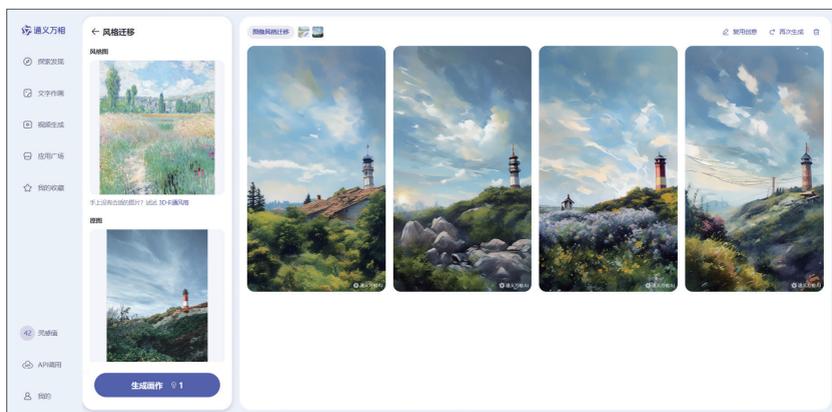


图 5-12

## 5.2.2 图像扩展

图像扩展是在保持图像原有内容完整性的基础上，对图像的尺寸或场景进行拓展，增加新的视觉元素或空间，使图像内容更加丰富。该技术常用于图像修复、虚拟场景生成等场景。其技术原理如下。

### 1. 特征分析

利用扩散模型（Diffusion Model）或生成对抗网络对图像的全局语义和局部纹理进行学习。扩散模型通过逐步添加噪声并学习去噪过程来捕捉图像的特征分布；生成对抗网络则由生成器和判别器相互对抗，生成器生成图像以欺骗判别器，判别器判断图像真伪，两者在博弈过程中学习到图像的特征。通过这些模型，可以获取图像的深层语义信息和丰富的纹理特征，为后续的图像扩展提供基础。

### 2. 边缘检测

采用Canny算法定位图像边界。Canny算法是一种多级边缘检测算法，它通过高斯滤波去噪、计算图像梯度、非极大值抑制以及双阈值检测等步骤，精确地确定图像中的边缘位置，从

而确定扩展区域的位置。

### 3. 上下文推理

利用卷积神经网络（如U-Net）或Transformer模型，根据未扩展区域的内容推断扩展区域的像素值。

- **U-Net**: 采用编码器—解码器结构，通过跳跃连接保留多尺度上下文信息，特别适合处理局部细节的生成任务。
- **Transformer模型**: 基于自注意力机制，能够建模图像中的全局依赖关系，确保生成内容在语义和风格上与原始图像保持一致。

## 动手练 使用豆包扩展人物场景



下面使用豆包为人物扩展场景，具体操作步骤如下。

**步骤 01** 打开豆包，进入工作界面，单击“图像生成”按钮，继续单击“扩图”按钮并上传图像，在新的界面中设置图片比例为4:3，如图5-13所示。

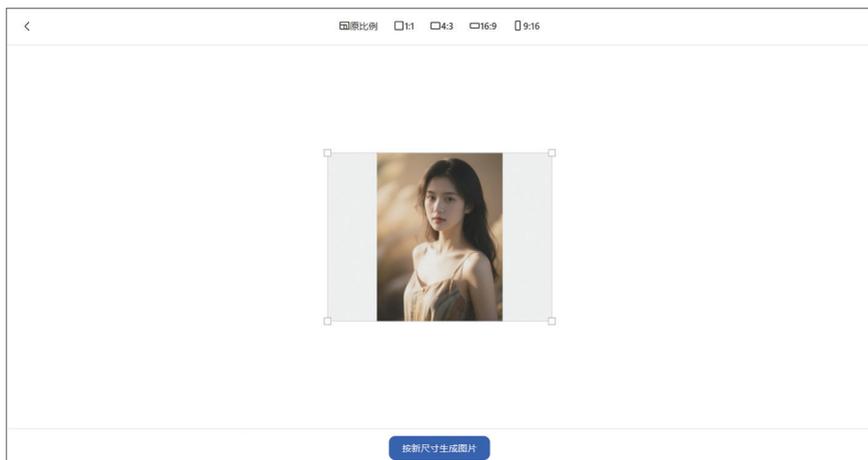


图 5-13

**步骤 02** 单击“按新尺寸生成图片”按钮，效果如图5-14所示。



图 5-14

## 5.2.3 局部编辑

图生图的局部编辑是指仅对图像的特定区域进行修改，而保持其他部分不变的技术。它允许用户通过文本提示或交互式工具（如涂鸦、蒙版）精确控制编辑范围，实现更精细的图像生成。该技术广泛应用于电商商品图编辑、虚拟试衣等领域。其核心功能如下。

### 1. 区域精准控制

用户可以利用输入的蒙版或涂鸦来指定编辑区域。蒙版是一种二值图像，其中白色区域表示需要编辑的部分，黑色区域表示保持不变的部分；涂鸦则是用户直接在图像上绘制的大致轮廓，模型会根据涂鸦的形状和位置来确定编辑区域。通过这种方式，用户能够精确地定义想要修改的图像部分，避免对其他无关区域的干扰。

### 2. 上下文感知生成

模型通过以下机制保证编辑内容与原始图像的自然融合。

- **特征提取**：分析未编辑区域的色彩分布、纹理特征和光照条件。
- **风格迁移**：使生成内容与原始图像保持视觉一致性。
- **结构保持**：维持透视关系和几何结构的合理性。

例如，在人物照片编辑中，模型会根据原始肤色和光线自动调整新生成衣物的色彩和阴影，确保整体协调自然。

### 3. 多模态输入

用户可以通过多种方式指定编辑内容，包括文本提示和上传图像素材。

#### （1）文本提示

用户可以输入简洁的文本指令，如“将眼镜换成墨镜”“把头发染成红色”等。模型会根据文本提示理解用户的意图，并生成相应的编辑结果。文本提示的优点是操作简单、方便快捷，用户只需用文字描述想要的效果即可。

#### （2）上传图像素材

用户还可以上传新的图像素材，如新的图案、物品图片等，指定将其应用到图像的特定区域。模型会将上传的素材与原图进行融合，生成符合要求的编辑图像。例如，在电商商品图编辑中，商家可以上传新的产品图案，将其应用到商品的包装上，以展示不同的设计效果。



## 动手练 使用豆包为模特换装

下面使用豆包为模特换装，具体操作步骤如下。

**步骤 01** 打开豆包，进入工作界面，单击“图像生成”按钮，显示如图5-15所示的界面。

**步骤 02** 继续单击“区域重绘”按钮并上传图像，在新的界面中使用画笔涂抹旗袍部分，释放鼠标左键后在文本框中输入提示词：“**替换旗袍的样式为新中式国风**”，如图5-16所示。

**步骤 03** 单击“发送”按钮，系统将根据提示词进行生成，效果如图5-17所示。



图 5-15

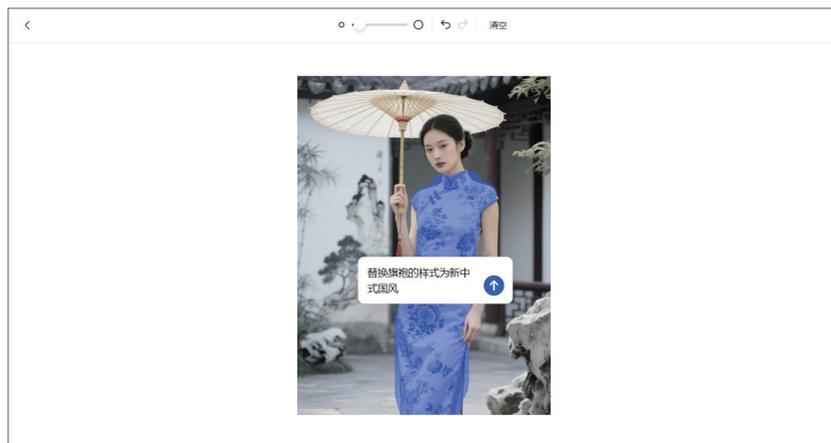


图 5-16

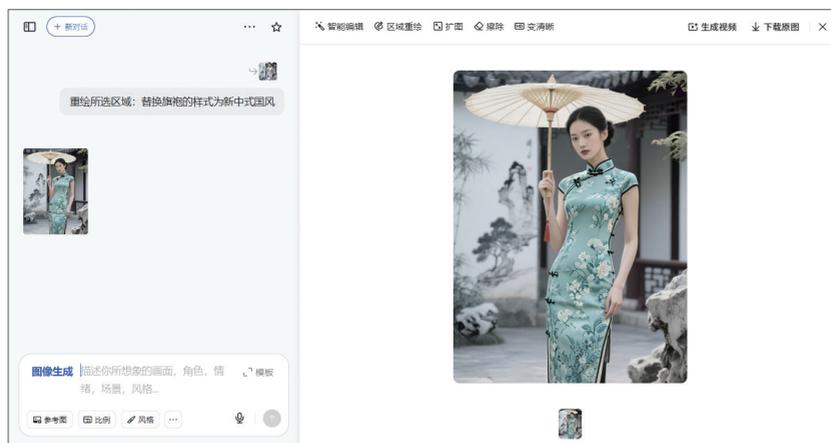


图 5-17

## 5.2.4 图像生成控制

图生图的图像生成控制（如ControlNet）是通过额外输入，如草图、深度图、姿态关键点等，精确控制生成图像的构图、结构或动态效果。该技术广泛应用于游戏角色设计、工业设计、草图渲染等领域。其技术原理如下。

## 1. 多条件输入

- **草图控制**：用户输入简单的线稿草图，ControlNet 能依据草图的轮廓和构图，生成细节丰富、内容完整的图像，常用于建筑效果图、产品设计图等创作场景。
- **姿态关键点**：通过指定人体或生物的动作姿态关键点，系统可生成对应姿态的角色形象，在动画设计、游戏角色动作制作等方面发挥重要作用。
- **深度图**：输入场景或物体的深度图信息，可精准控制生成图像的空间层次和透视关系，用于虚拟场景搭建、3D模型渲染等。

## 2. 模型架构

ControlNet在扩散模型的基础上进行改进，引入额外的控制分支。该控制分支可以将额外输入条件（如草图、深度图等）与扩散模型的生成过程紧密结合，在图像生成过程中，强制模型严格遵循输入条件，从而实现对生成内容的精准控制，使输出图像在满足语义需求的同时，更符合指定的结构、构图或动态要求。



### 动手练 使用即梦AI优化草图

下面使用即梦AI优化草图，具体操作步骤如下。

**步骤 01** 打开即梦AI，在“图片生成”界面的输入框下方单击“导入参考图”按钮。在弹出的“打开”对话框中选择素材图像，跳转至图5-18所示的界面。



图 5-18

**步骤 02** 单击“保存”按钮后，在输入框中输入提示词：“为该照片进行上色处理”，单击“立即生成”按钮，系统将生成四张不同的图像，单击缩览图即可查看效果，如图5-19所示。

**步骤 03** 单击右下角的“细节修复”按钮，效果如图5-20所示。



图 5-19



图 5-20

## 5.3 基于图像的AI视频生成

在图像生成技术日益成熟的今天，AI的创造力已不再局限于单帧画面的转换。基于图像的AI视频生成将静态图像作为输入，通过深度学习模型赋予其动态表现力，实现从“瞬间定格”到“连续叙事”的跨越。

### 5.3.1 图文内容动态化

图文内容动态化技术能够将静态的平面设计、插画或照片转化为生动的动态内容，适用于社交媒体广告、品牌宣传片、教育动画等场景。其核心技术如下。

#### 1. 运动轨迹预测技术

运动轨迹预测技术基于光流（Optical Flow）轨迹或时空注意力机制，预测图像中物体的运动方向和速度。例如，通过分析静态插画中云彩的纹理，生成自然飘动的动画效果。

## 2. 分层动画技术

分层动画技术将图像分解为前景、背景、角色等不同层次，分别赋予动态属性。例如，在一张城市风景图中，可单独为车辆、行人、光影等元素添加动画，实现“画中画”式的动态叙事。

## 3. 商业应用案例

- **社交媒体广告**：将产品海报转化为动态视频，提升用户互动率。
- **教育动画**：将科学插图转化为动态演示，辅助知识讲解。
- **虚拟展厅**：为静态艺术作品添加光影变化，增强沉浸感。

## 动手练 使用可灵生成火山爆发效果

下面使用可灵生成火山爆发效果，具体操作步骤如下。

**步骤 01** 打开可灵，单击“视频生成”按钮进入“视频生成”界面，继续单击“文生视频”按钮，在输入框上方单击DeepSeek按钮，如图5-21所示。



图 5-21

**步骤 02** 激活“DeepSeek-R1 灵感版”功能，输入提示词：“**动态化火山喷发插图，岩浆层用黏稠流体模拟向上涌动，烟雾粒子按风速场扩散，地壳板块缓慢开裂，5秒，扁平化演示**”，单击“发送”按钮，系统自动深度思考。

### DeepSeek

好的，我需要处理用户关于生成视频提示词请求。用户这次的需求是“动态化火山喷发插图，岩浆层用黏稠流体模拟向上涌动，烟雾粒子按风速场扩散，地壳板块缓慢开裂，5秒，扁平化演示”。首先，我需要按照之前设定的规则来分析这个请求。首先，用户提到了几个关键元素：火山喷发的动态插图、黏稠流体模拟的岩浆、烟雾粒子按风速场扩散、地壳板块开裂，以及5秒时长和扁平化演示风格。这些都是必须保留的关键元素，不能遗漏。

接下来，按照生成Prompt的逻辑，我需要确定六个必要成分：主体、主体描述、主体动作、场景、场景描述、其他细节补充。这里的主体显然是火山喷发，主体描述是动态化插图和扁平化风格。主体动作包括岩浆向上涌动、烟雾扩散、地壳开裂。场景可能设定在地质环境，场景描述需要包含扁平化演示的科学质感。其他细节补充需要考虑运镜、景别、光影等。

根据用户提到的“扁平化演示”，我需要确保整体风格符合扁平化设计，可能使用简约背景和几何形状。用户还提到5秒时长，但提示词中不需要直接写时间，而是通过动作描述来体现动态过程。烟雾粒子按风速场扩散可能需要避免使用粒子特效词汇，但用户明确提及，所以必须保留。

接下来检查禁止项，用户提到了黏稠流体、烟雾粒子、地壳开裂，这些属于允许的内容，但需要确保不出现禁止的复杂物理运动或特效词汇。用户的需求中没有提到运镜，但可能需要远景展示火山全貌，使用仰视角度来增加张力。光影方面，暖光和自然光交织可以烘托氛围。

最后，整合所有元素，确保提示词连贯且符合字数要求。需要将六个成分自然连接，避免使用禁止的符号，保持语句通顺。例如：“扁平化风格的动态火山喷发演示，黏稠岩浆以流体形态沿山体向上涌动，地壳板块匀速裂开缝隙，烟雾粒子按虚拟风速轨迹扩散，仰视远景呈现整体地质运动，暖光与自然光交织，简约背景凸显科学演示的几何美感”。

检查是否有遗漏的关键元素，如扁平化、风速场、5秒（通过动态描述暗示），并确保没有违反生成禁止项。最终输出符合格式的要求，没有多余内容，结构正确。

**步骤 03** 根据深度思考，生成以下提示词。

扁平化风格的动态火山喷发演示，黏稠岩浆以流体形态沿山体向上涌动，地壳板块匀速裂开缝隙，烟雾粒子按虚拟风速轨迹扩散，仰视远景呈现整体地质运动，暖光与自然光交织，简约背景凸显科学演示的几何美感。

**步骤 04** 单击提示词后的“使用提示词”按钮，自动将内容填充至输入框内，如图5-22所示。

**步骤 05** 单击“立即生成”按钮，效果如图5-23所示。



图 5-22

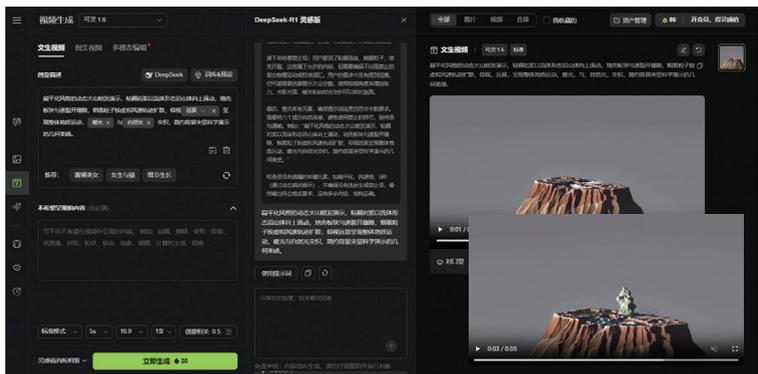


图 5-23

## 5.3.2 人物视频生成

人物视频生成技术专注于将静态人物图像转化为自然流畅的视频序列，适用于影视制作、虚拟主播、个性化视频生成等场景。其核心技术如下。

### 1. 面部动画技术

面部动画技术基于面部关键点检测（如68点模型）和生成对抗网络，实现表情的细微变化（如微笑、眨眼、皱眉）。例如，通过输入一张人物肖像，生成其说话、唱歌或做出特定表情的动态视频。

### 2. 全身动作生成技术

全身动作生成技术结合运动捕捉数据和时空卷积网络（3D-CNN），生成符合物理规律的全身动作。例如，将一张模特照片转化为走秀视频，或为历史人物画像添加动态姿势。

### 3. 应用场景

- **影视制作**：快速生成虚拟演员的表演片段。

- **虚拟主播**：通过实时驱动面部和肢体动作，实现24小时不间断直播。
- **个性化视频**：上传照片后，系统自动生成生日祝福、节日贺卡等定制化视频。



## 动手练 使用即梦AI让古画人物动起来

下面使用即梦AI让古画中的人物动起来，具体操作步骤如下。

**步骤 01** 打开即梦AI，在“视频生成”界面单击“图片生视频”按钮，如图5-24所示。

**步骤 02** 单击“上传图片”按钮，在弹出的“打开”对话框中上传素材图像，在输入框中输入提示词：“**请让这幅图中的人物动起来，她们正在演奏乐器。保持背景不变，只让人物进行互动和演奏动作**”，如图5-25所示。



图 5-24



图 5-25

**步骤 03** 单击“立即生成”按钮，系统将根据描述自动生成视频，如图5-26所示。

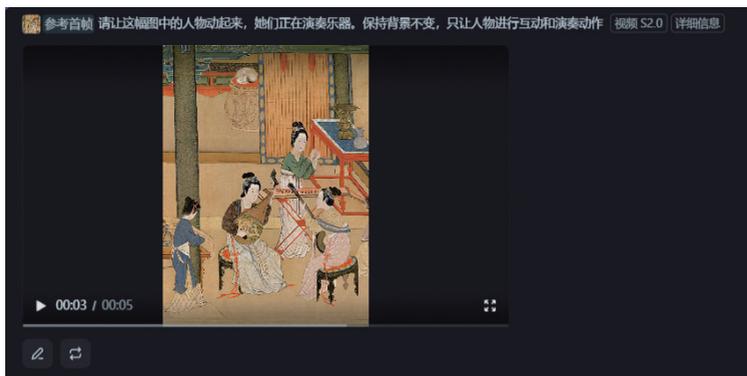


图 5-26

**步骤 04** 效果如图5-27~图5-29所示。



图 5-27



图 5-28



图 5-29

### 5.3.3 商品3D视频生成

商品3D视频生成技术为电子商务和产品展示带来革新，通过将静态商品图转化为可交互的3D动态视频，提升用户购买决策效率。其核心技术如下。

#### 1. 3D建模技术

3D建模技术基于单张或多张商品图像，利用深度学习模型（如NeRF、3D-GAN）自动生成高质量3D模型。例如，输入一张家具照片，系统可快速重建其3D结构并添加材质和纹理。

#### 2. 动态展示功能

- **视角旋转**：用户可使用手势或鼠标拖曳的方式360°查看商品细节。
- **物理模拟**：为服装商品添加布料动态模拟，展示真实穿着效果。
- **光影交互**：实时渲染商品在不同光照条件下的表现（如金属反光、织物褶皱）。

#### 3. 商业价值

- **数据稀缺性**：人物视频生成需大量高质量动作捕捉数据，而商品3D建模依赖精确的CAD模型，数据不足可能导致生成结果失真。
- **计算成本**：实时渲染3D视频对算力要求极高，需优化模型效率或结合边缘计算。
- **伦理与版权**：动态化技术可能被滥用（如伪造视频证据），需建立内容溯源和版权保护机制。

### 动手练 使用即梦AI生成模特展示视频



下面使用即梦AI生成模特展示视频，具体操作步骤如下。

**步骤 01** 打开即梦AI，在“视频生成”界面单击“文本生视频”按钮，输入框下方有“DeepSeek-R1”按钮，如图5-30所示，单击后跳转至图5-31所示界面。



图 5-30



图 5-31

**步骤 02** 在输入框中输入提示词：“**女装、模特、多角度展示**”。向下拖动设置图片比例为9:16，单击“发送”按钮, 经过深度思考后生成以下三种推荐提示词。

模特优雅转身展示驼色风衣立体剪裁，前镜头凸显双排金属扣开合动态，侧身时腰带自然收束勾勒曲线，后摆丝绸蝴蝶结随步伐轻盈飘动，面料在行走间泛起柔光涟漪。

职业装女性自信行走多角度呈现，正面西装领口珠片随步态闪烁，转身时A字裙摆划出流畅弧度，侧面镜头捕捉袖口金线刺绣细节，光影在挺括面料上形成渐变波纹。

时尚高挑模特站立分段展示套装，上半身定格展示领口钉珠细节，下半身定格展示裙摆立体剪裁，最后全身旋转呈现流线型轮廓，冷调灯光凸显面料金属光泽。

**步骤 03** 选择第一种推荐提示词，单击“立即生成”按钮。系统将根据描述自动生成视频，如图5-32所示。



图 5-32

**步骤 04** 效果如图5-33~图5-36所示。



图 5-33



图 5-34



图 5-35



图 5-36

## 5.4 实战演练：古诗意境动态化——《江雪》



下面使用即梦AI中的DeepSeek-R1功能，将古诗《江雪》意象化，然后选择其中的一种提示词，生成图片后将其动态化，具体操作步骤如下。

**步骤 01** 打开即梦AI，在“图片生成”界面单击“DeepSeek-R1”选项按钮，输入提示词：“江雪，孤舟蓑笠翁，独钓寒江雪”。向下拖动设置图片比例为16:9，单击“发送”按钮, 经过深度思考后生成以下四种推荐提示词。

孤舟蓑笠翁持钓竿独坐船头，寒江雪景中只留船体轮廓剪影，宋代水墨写意，俯视镜头，冷墨青色调，枯笔飞白技法。

戴斗笠的老者蜷坐船尾垂钓，被雪雾笼罩的江面延伸至天际线，唐代青绿山水，全景深构图，积雪压枝细节，留白处题写柳宗元诗句。

三角形构图的雪山倒映江中，小舟横向切割画面三分之二处，渔翁背影融入雪幕，马远一角式构图，绢本设色质感，雾凇晶莹渐变处理。



**步骤 05** 单击“立即生成”按钮，系统将自动擦除，效果如图5-40所示。



图 5-40

**步骤 06** 单击右下角的“生成视频”按钮，跳转至“视频生成”界面，在输入框中输入提示词：“大雪纷飞，船和老翁保持不动，鱼线和湖面轻轻泛起涟漪”，选择视频模型为“视频 S2.0Pro”，单击“生成视频”按钮，系统自动生成5秒视频，如图5-41所示。



图 5-41

**步骤 07** 效果如图5-42和图5-43所示。



图 5-42



图 5-43