

人工智能

基础与应用

Fundamentals and Applications of
Artificial Intelligence

主 编 窦煜明 刘兴一



中国劳动社会保障出版社

图书在版编目 (CIP) 数据

人工智能基础与应用 / 窦煜明, 刘兴一主编.

北京: 中国劳动社会保障出版社, 2025. -- ISBN 978-7-5167-7258-4

I. TP18

中国国家版本馆 CIP 数据核字第 2025PX3180 号

人工智能基础与应用

RENGONG ZHINENG JICHU YU YINGYONG

中国劳动社会保障出版社出版发行

(北京市惠新东街 1 号 邮政编码: 100029)

*

北京利丰雅高长城印刷有限公司印刷装订 新华书店经销

787 毫米 × 1092 毫米 16 开本 10.75 印张 154 千字

2025 年 8 月第 1 版 2025 年 8 月第 1 次印刷

定价: 19.90 元

营销中心电话: 400-606-6496

出版社网址: <https://www.class.com.cn>

版权专有 侵权必究

如有印装差错, 请与本社联系调换: (010) 81211666

我社将与版权执法机关配合, 大力打击盗印、销售和使用盗版图书活动, 敬请广大读者协助举报, 经查实将给予举报者奖励。

举报电话: (010) 64954652

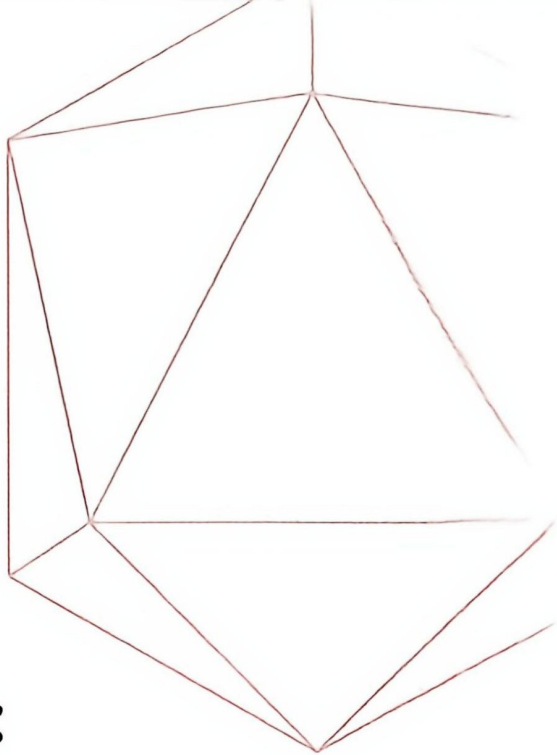
... 目 录

第一章 智慧起航，认识人工智能技术 1

- 模块一 智慧初探：人工智能的奇妙世界
.....2
- 模块二 理论洞察：解析人工智能的本质
..... 15
- 模块三 应用拓展：解锁生成式人工智能
..... 26

第二章 智慧探索，解码生成式人工智能的核心原理33

- 模块一 智创文思：生成式人工智能的
文本创作奥秘34
- ✓ 模块二 智绘视界：生成式人工智能的
图像生成与视觉创新 42
- 模块三 智声灵动：生成式人工智能的
语音生成与交互革命 55
- 模块四 智影生辉：生成式人工智能的
视频创作与动态影像 64



模块二 智绘视界： 生成式人工智能的图像 生成与视觉创新

在这个由像素与代码编织的奇妙世界里，生成式人工智能宛如一位拥有魔法的智能画师，它以数据为颜料、算法为画笔，将人们脑海中的灵感瞬间化为栩栩如生的图像，让每一幅作品都成为超越想象的视觉奇迹。在本模块中，我们将一同探索生成式人工智能是如何用其强大的创造力，重新定义图像创作的边界的。

应用探索

智绘生辉：开启图像创作的奇幻之旅

图像生成技术不仅为艺术创作带来了全新的可能性，也为我们提供了一个充满创意的空间。通过图像生成技术，绘画者可以不受传统绘画的限制，快速实现自己的创意，激发想象力和创造力。

一、实验目的

通过本次实验，学生将直观领略图像生成技术的独特魅力，从而激发自身对绘画与科技的浓厚兴趣。学生在发挥想象力和创造力的同时，还将掌握图像

生成工具的基本操作方法，进而提升自身的数字绘画技能。

二、实验任务

学生将分组查阅资料，了解图像生成技术的基础知识，选择并了解一款图像生成工具（如 Midjourney、豆包或通义万相等）。

使用选定的图像生成工具完成一系列主题图片创作。首先，通过图像生成工具生成一幅符合描述的油画风格图像，可先感受其艺术效果。随后，将此图像通过图像生成工具转化为另一种风格化的图像，创造出全新的视觉效果。

上传一张照片，利用图像生成工具将其转换为动漫风格的图像，观察图像生成工具如何将真实照片的细节与动漫风格相结合，分析生成的动漫风格的图像与原照片在色彩、线条和整体氛围上的差异。在生成图像的过程中，不断优化提示语，迭代生成图像，比对不同效果，并记录生成过程中的感受与遇到的问题。

最后，每组推选一名学生代表，展示本组生成的图像，并简要介绍生成过程和感受。通过展示和讨论，促进学生之间的交流与互动，引导学生深入思考图像生成技术的优势与局限，以及其对艺术创作的深远影响。

三、思考与讨论

本次实验将引导学生深入思考图像生成技术与传统绘画之间的关系，以及图像生成技术对艺术创作带来的深远影响。讨论一下，图像生成技术的出现为艺术创作带来了哪些机遇和挑战？它是如何改变艺术家的创作方式和观众的审美体验的？未来，人类艺术家与人工智能如何更好地协作，共同推动艺术的发展？

一、图像生成技术的概念

图像生成是一种利用人工智能技术，尤其是深度学习算法，从无到有地创造出图像内容的过程。它通过分析和学习大量的图像数据，生成新的图像，这些图像可以是完全虚构的，也可以是对现有图像的修改或图像效果的增强。图像生成技术在多个领域都有广泛的应用，包括艺术创作、娱乐、广告、医疗、教育等。

知识拓展

解锁图像生成的魔法钥匙——提示词的艺术

在数字艺术的世界里，图像生成技术就像一位神奇的画师，能够将你的文字描述转化为栩栩如生的画面。而这一切的魔法，都始于一个关键的元素——提示词。提示词就像是 you 与人工智能之间的秘密语言，帮助你精准地表达内心的想法，让人工智能为你创造出独一无二的图像。

（一）什么是提示词？

提示词是用户输入给图像生成工具的文字描述，它告诉人工智能你想要生成什么样的图像。提示词的质量直接影响生成图像的效果，好的提示词能够帮助人工智能更准确地理解你的需求，从而生成更符合预期的画面。

（二）如何写出高质量的提示词？

1. 明确具体的目标

尽量用简洁明了的语言描述你想要的图像内容，避免模糊不清的表达。例如，不好的提示词：“美丽的风景”；好的提示词：“一片宁静的山谷，远处有雪山，天空中有彩虹，阳光洒在绿色的草地上”。

2. 尽可能多的细节描述

提供尽可能多的细节，包括场景、主体、颜色、光影等。例如，不好的提示词：“一个女孩”；好的提示词：“一个穿着汉服的少女，站在樱花树下，微风轻拂她的长发，背景是古色古香的庭院，庭院中有许多生机盎然的绿植”。

3. 提供风格和艺术参考

如果你希望生成的图像具有某种特定的艺术风格，可以在提示词中明确给出。例如：“一片梵高风格的星空，夜晚的村庄，明亮的星星闪烁，色彩鲜艳”。

4. 准确的情感和氛围

描述你希望图像传达的情感或营造的氛围，帮助生成器更好地捕捉整体感觉。例如：“一个温馨的家庭场景，父母和孩子在客厅里欢笑，阳光透过窗户洒进来，画面充满温暖和爱”。

5. 避免歧义

尽量避免使用可能引起歧义的词汇或短语，确保生成器能够准确理解你的意图。例如，不好的提示词：“一个大房子”；好的提示词：“一座欧式风格的豪华别墅，有高大的柱子和宽敞的庭院”。

6. 使用专业术语

如果你熟悉一些艺术或设计术语，可以使用它们来更精确地描述你的需求。例如：“一幅超现实主义风格的油画，画面中有扭曲的建筑和流动的天空”。

7. 分步细化

逐步优化，如果第一次生成的图像不符合你的预期，不要气馁。根据生成的结果，逐步调整和优化你的提示词。例如，初始提示词：“一场海边的日落”；优化后的提示词：“一场海边的日落，天空中有橙色和紫色的晚霞，海浪轻轻拍打着沙滩，画面风格是油画”。

（三）实践练习

接下来，将通过一个简单的实践练习，帮助你更好地掌握提示词的撰写技巧。首先，选择一个主题，比如“神秘的森林”或“温馨的家庭”。然后，根据之前的建议，尝试撰写一个详细的提示词。接下来，使用图像生成工具，输入你撰写的提示词并生成图像。最后，根据生成的图像效果，调整和优化你的提示词，再次生成图像，直到你对结果感到满意。

例如，对比“一个大房子”和“一座欧式风格的豪华别墅，有高大的柱子和宽敞的庭院”这两个提示词，你会发现它们生成的图像效果截然不同，如图 2-1 所示。通过这个练习，你可以直观地感受到详

的提示词对生成图像效果的重要性。



图 2-1 提示词优化前后图像对比图

二、图像生成技术的原理

图像生成技术的原理是通过计算机算法和模型，根据输入的指令或数据生成新的图像内容。常见的图像生成技术有生成对抗网络、变分自编码器和扩散模型。

（一）生成对抗网络

1. 生成器

生成器的目标是尽可能生成接近真实图像的虚假图像。它从随机噪声开始，通过学习训练数据的分布，逐步生成逼真的图像。

2. 判别器

判别器的目标是区分生成器生成的虚假图像和真实图像。它通过学习真实图像的特征，判断输入的图像是真实的还是生成的。

3. 对抗过程

生成器和判别器相互对抗，生成器不断改进生成图像的质量，判别器则不断提升其识别能力。最终，生成器生成的图像能够“欺骗”判别器，使其无法区分真假图像。

生成对抗网络广泛应用于图像生成、风格迁移、超分辨率等领域。

（二）变分自编码器

1. 编码器

编码器将输入图像压缩在一个低维的潜在空间。这个潜在空间表示一个概率分布，通常是一个高斯分布。

2. 解码器

解码器从潜在空间中采样，将低维表示还原为图像。

3. 变分推断

通过最小化重构误差和正则化潜在空间，使生成的图像尽可能接近真实图像。

变分自编码器主要应用于图像生成、图像修复、数据压缩等领域。

（三）扩散模型

1. 噪声添加过程

从真实图像开始，逐步添加高斯噪声，将图像逐渐“破坏”成噪声。

2. 去噪过程

通过训练一个模型，逐步去除噪声，恢复出原始图像。这个过程类似于“逆向扩散”。

3. 迭代生成

通过多次迭代，逐步从噪声中生成清晰的图像。

扩散模型主要应用于图像生成、从文本到图像的生成、视频生成等领域。

知识拓展

人工智能图像生成工具可以取代摄影师吗？

随着人工智能图像生成技术的进步，很多人开始担心它是否会取代摄影师的工作。今天，我们就来深入解析这个问题。

人工智能图像生成工具可以在短时间内高效地生成各种风格的图像，从写实到卡通，从油画到水彩，无所不能。使用人工智能图像生

成工具生成图像无须昂贵的设备和场地，大幅降低了创作成本。此外，人工智能图像生成工具生成的图像还可以根据需求被快速修改和调整，极大提高了创作的灵活性和效率。

摄影师的工作远不止按下快门，他们需要从创意构思开始，根据客户需求设计独特的拍摄方案。摄影师要在现场精心布置灯光、调整构图并引导客户拍摄，确保拍摄效果。拍摄完成后，他们还要对照片进行后期处理，如色彩调整、剪裁和合成等。此外，与客户的沟通也至关重要，他们需要充分了解客户需求并提供专业建议，以满足客户的需要并创作出高质量的作品。

因此，摄影师是不可替代的。主要原因在于他们通过镜头捕捉到的不仅是画面，更是情感和故事。人工智能图像生成工具虽然可以生成逼真的图像，但很难捕捉到真实的情感。此外，摄影师需要根据现场情况灵活调整拍摄方案，而人工智能图像生成工具目前还无法做到这一点。摄影师还需要具备专业的摄影技能和艺术素养，这些也是人工智能图像生成工具无法做到的。因此，摄影师在创作过程中所发挥的独特作用，使得他们具有不可替代性。

但是，人工智能图像生成工具可以成为摄影师的好助手。人工智能图像生成工具可以为摄影师提供灵感，帮助摄影师快速生成创意草图，甚至完成一些基础的后期处理工作。这样，摄影师可以将更多的时间和精力放在创意构思和现场拍摄上。摄影师可以利用自己的专业知识和艺术素养，指导人工智能图像生成工具生成更符合要求的图像。

未来，人工智能图像生成工具和摄影师将共同进步，创造更多的可能性。

三、图像生成技术的多样化应用场景

如今，文生图工具和解决方案已广泛应用，成为创意设计和内容创作领域的重要助力。在国内，百度、阿里云、抖音等企业纷纷推出了各具特色的同

类工具，这些工具不仅图像生成速度快，还支持中文提示词，完美契合国内用户的需求，并在多种图像使用场景中表现出色。

(一) 创意艺术

数字绘画借助计算机的强大处理能力、图形输入板的精准操作、手写笔的灵活运用，以及专业软件工具的强大功能，能够逼真地模拟水彩、油画、素描等传统绘画效果，甚至在某些细节上打破了对传统绘画的限制。随着人工智能技术的融入，数字绘画的创作边界被进一步拓展。人工智能不仅能够生成具有独特风格的艺术作品，还能显著提升创作效率。在人工智能图像生成软件中，用户只需输入精准的提示词，即可快速生成符合要求的图像，这大大降低了创作门槛，让更多人能够轻松参与到艺术创作中来，见表 2-1。

表 2-1 生成绘画作品效果

提示词	图像
雪地背景，一张可爱小猫的照片，小猫从雪地探出脑袋，萌萌的大眼睛，以超现实主义风格绘制（制作工具：豆包，极速创作模型，文生图生成方式，人像摄影风格开启）。	
国画，山水，小桥，船，飞鸟，写意风格（制作工具：豆包，绘画模型 3.0，极速创作模型，文生图生成方式，中国风风格开启）。	

(二) 虚拟形象

人工智能文字生成虚拟形象的工具借助先进的人工智能算法，通过用户输入的文字描述或上传的图片，能够快速生成个性化的虚拟形象，见表 2-2。用户只需简单描述虚拟形象的外貌特征、服饰风格、场景背景等信息，人工智能就能精准理解并生成高度符合要求的虚拟形象。这项技术广泛应用于多个领域：在游戏开发中，它能快速生成丰富多样的游戏角色，提升开发效率；在社交媒体上，用户可以创建独特的虚拟形象并用它进行社交互动，在保护隐私的同时增加趣味性；在影视制作中，人工智能生成的虚拟形象可用于特效制作或动画角色设计，节省时间和成本；在虚拟会议中，它能让参会者以虚拟形象出现，增强沉浸感。


表 2-2 文生图工具生成虚拟形象

提示词	图像
生成一个虚拟形象：直播小能手，才华横溢，能说会道，有亲和力，长发，表情轻松，有幽默感（制作工具：豆包）。	

(三) 广告与营销

人工智能能够根据用户需求自动生成高质量的产品图像，支持个性化营销活动和产品展示。它不仅可以快速生成符合品牌形象的设计方案，还能辅助企业设计和优化品牌形象，提升品牌视觉效果和市场竞争力。表 2-3 展示了豆包制作的虚拟产品照片，生动呈现了人工智能在实际应用中的高效性和创意性。

表 2-3 文生图工具生成产品效果图

提示词	图像
请生成一张用于营销的图片：有未来感的办公桌，风格柔和圆润，配色大胆，45 度角拍摄，奶油色背景，阳光照射，比例 4：3，风格为电影写真（制作工具：豆包）。	

（四）电影与游戏

文生图工具可用于电影特效制作，生成逼真的视觉效果，提升作品观赏性，同时实现低成本与高效率；也可以用于自动生成多样化的游戏角色和场景。表 2-4 是一幅由文生图工具生成的卡通影视图片。

表 2-4 文生图工具生成卡通影视图片

提示词	图像
请生成一张卡通影视图片，有逼真的视觉效果，内容为功夫熊猫和功夫狐狸，比例 4：3，风格为中国风（制作工具：豆包）。	

从上面的示例中不难发现，要想获得高质量的图像生成效果，需要准备合适的提示词，通过文生图工具完成图像的创建。有时为了使生成的图像更准

确，也可以采用图生图的方式。具体而言，提供参考图像作为基础，通过提示词指导系统如何修改并生成新的图像。图 2-2、图 2-3 展示了将小鸭的原图重新生成电影风格的效果图。



图 2-2 原图



图 2-3 新图

知识拓展

人工智能制作活动背景图

一、实验目的

本次实验旨在通过人工智能图像生成技术，创作一幅符合“创新未来”主题的背景图。通过明确主题、合理运用色彩搭配、保持简洁大方的设计风格以及融入创意元素，培养学生利用人工智能图像生成技术进行主题性图像创作的能力，同时提升其对科技与艺术融合的理解力，激发创新意识和创造力，为未来科技活动的视觉设计积累实践经验，激发创作灵感。

二、实验任务

（一）任务描述

使用人工智能图像生成技术生成一张以“创新未来”为主题的活动背景图。

（二）任务要求

1. 主题明确

背景图应明确体现“创新未来”这一主题。可以使用象征创新、

进步和科技发展的元素，如创新的图标、电路图、科技设备等。

2. 注重色彩搭配

采用明亮、现代且富有科技感的色彩，如蓝色、紫色、银色或橙色等。这些色彩可以突出活动的未来感和创新性。

3. 风格简洁大方

背景图应简洁明了，避免使用过于复杂的图案和元素。保持整体设计的清晰度和易读性，以便大家能够轻松理解活动的主题。

4. 使用创意元素

鼓励使用创意元素，如抽象的形状、光线效果或动态图形，以突出活动的创新性和独特性。这些元素可以激发大家的创造力和创新意识。

三、实现工具

目前，国内可以根据提示词生成图像的人工智能工具有很多。

（一）豆包

抖音的豆包可通过自然语言描述生成独特图片，支持局部重绘、一键扩图等编辑功能。它擅长东方美学风格，如古风、水墨画等，还能生成带中文的插图和海报，满足多种创作需求，兼具实用性和艺术性。

（二）文心一格

百度的文心一格可以预置二次元、国潮等风格，支持图片拓展和背景替换，适合快速生成社交媒体配图。

（三）通义万相

阿里的通义万相支持文生图、涂鸦作图、相似图生成等功能，适合电商和内容创作者。

四、实现步骤

第一步：打开人工智能工具。

第二步：进入图像生成界面。在对话界面下方找到“图像生成”

功能入口，也可在输入框附近寻找专门的“AI 绘画”按钮，点击进入。

第三步：构思并输入提示词。尽量详细地描述想要的画面，包括主体、细节、环境、风格、构图、画质、比例等要素。

第四步：设置参数。根据需求调整相关参数，如画面比例（常见的有 1:1、16:9、9:16 等）、图片数量（一次可生成 1~4 张）、画质（高清或普通）、风格强度等。

第五步：生成。填好提示词、调好参数后，点击“生成”按钮，等待片刻，人工智能工具就会根据描述生成相应图片。

第六步：优化迭代。若对初次生成的图片不满意，分析问题后修改提示词，再次生成图片，直到得到理想的画面。

五、测试、优化与应用

根据提示内容生成的图像，并结合不同的提示词进行调整优化。



思考题



一、填空题

1. 常见的图像生成技术包括 _____、_____ 和 _____。
2. 生成对抗网络的组成部分包括 _____、_____ 和 _____。
3. 图像生成技术中，_____ 的质量直接影响生成图像的效果。

二、简答题

1. 根据你的了解，有哪些常见的图像生成工具？它们有什么特点？
2. 请结合你的学习、工作，介绍一下图像生成技术的各种应用场景，有哪些应用可以借助图像生成技术协助我们更好地完成任务？