

## 第 3 章

### 多模态教学：让课堂从平面走向立体

教育信息化 2.0 时代推动课堂教学从单一模态向多元模态融合转型。多模态教学理论揭示了人类能通过视觉、听觉等多通道协同处理信息，传统单一文本模式已无法满足数字化学习需求。研究表明，多感官刺激能增强神经突触活性，提高知识留存率。本章基于认知负荷理论，探讨如何运用文生视频、数据可视化等技术构建“感知—体验—交互”一体化的教学场景，为核心素养的培养提供新范式。

#### 3.1 文生视频：制作动画教学短片

文生视频技术正成为突破传统课堂平面化教学的关键技术。基于多媒体学习理论，该技术通过 AI 算法将文本转化为动态影像，实现抽象概念的具象化呈现，构建沉浸式学习情境。本节将从教育传播学角度出发，结合案例分析文生视频如何实现教学内容可视化转译，揭示动画短片在概念理解、情境迁移等方面的认知支架作用，为教师进行数字教学设计时提供理论与实践指导。

##### 3.1.1 AI 实现“贝多芬做自我介绍”

在课堂的审美体验与知识建构中，历史人物的“在场感”缺失始终是引起学生情感共鸣与引导学生深度理解的瓶颈。例如，当贝多芬以虚拟形象在音乐课上开口讲述自己的创作心境时，AI 技术便不再是冰冷的工具，而是搭建起跨越两个世纪的精神桥梁。

下面以 AI 工具豆包与即梦 AI 为例，展示实现“贝多芬以虚拟形象做自我介绍”的具体步骤。

##### 步骤 1 >> 打开豆包图像生成页面

进入豆包首页，单击页面下方的“图像生成”按钮，如图 3-1 所示。进入图像生成页面，如图 3-2 所示。



图 3-1 豆包首页



图 3-2 图像生成页面

### 步骤 2 完成“生成贝多芬历史形象图”的指令

在对话框中输入提示词：“生成贝多芬的历史人物形象图。画面风格是写实油画风格，注重色彩的细腻过渡和光影的真实表现，笔触细腻，能很好地还原人物的真实质感和历史氛围。主体是一位写实油画风格的贝多芬，他有着满头凌乱的卷发、深邃的眼神、高挺的鼻梁，嘴唇紧闭，脸上带着严肃而专注的神情；身着一件黑色的古典长袍，领口系着白色的领巾，站在一间摆满了书籍和乐谱的书房里，背后的书架上摆放着各种与音乐相关的书籍和手稿，书桌上放着笔和未完成的乐谱。正面视角，半身照。”，单击对话框右下角的箭头按钮发送指令，如图 3-3 所示。



图 3-3 输入并发送指令

### 步骤 3 查看生成图并处理

查看豆包根据指令生成的图片，选择一张最合适的图片，如图 3-4 所示。

单击页面上方的“变清晰”按钮，对图片进行高清处理，如图 3-5 所示。



图 3-4 贝多芬历史形象图

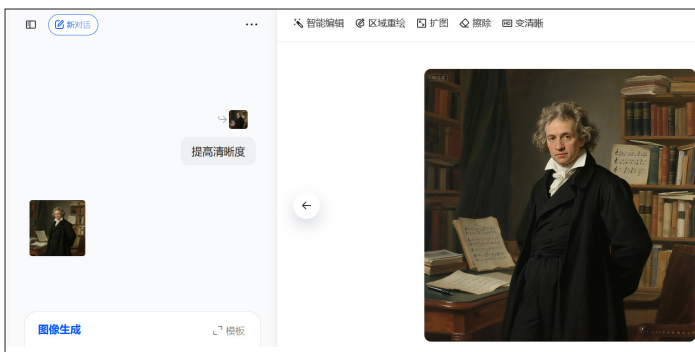


图 3-5 对图片进行高清处理

#### 步骤 4 ➤ 下载图片并保存

单击页面右上角的“下载原图”按钮，将处理好的图片下载并保存，如图 3-6 所示。

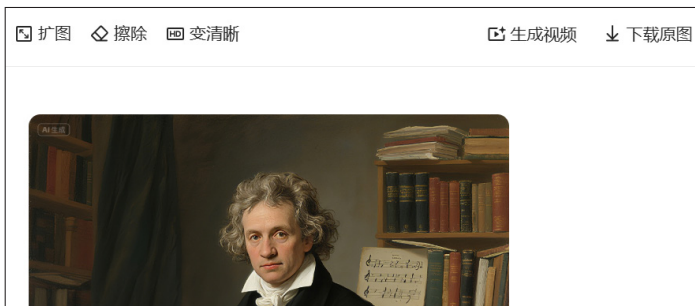


图 3-6 下载并保存原图

**步骤 5** 打开即梦 AI 首页

进入即梦 AI 首页，选择“数字人”功能，如图 3-7 所示。

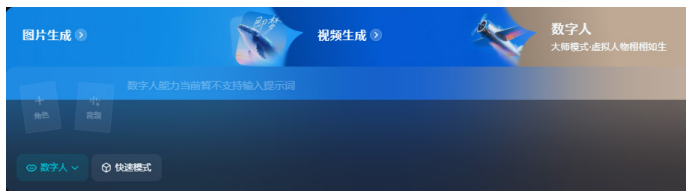


图 3-7 选择“数字人”功能

**步骤 6** 上传图像和音频

单击“角色”，上传在步骤 3 中生成的贝多芬历史形象图，如图 3-8 所示。单击“音频”按钮，选择“生成音频”，如图 3-9 所示。在对话框中输入提示词：“同学们好，我是贝多芬，来自德国波恩，是一名作曲家。”同时，选择音色，单击页面右下角“确认”按钮，生成贝多芬说话音频，如图 3-10 所示。



图 3-8 上传贝多芬形象图



图 3-9 选择“生成音频”



图 3-10 生成音频

**步骤 7** 生成视频并下载

单击对话框右侧的箭头按钮发送指令，生成“贝多芬做自我介绍”视频，如图 3-11 和图 3-12 所示。在视频的右下



图 3-11 发送指令

角编辑栏可以进行重新配音、提升视频分辨率、使用 AI 配乐等操作，最后单击视频右上角的下载按钮进行视频下载，如图 3-13 所示。

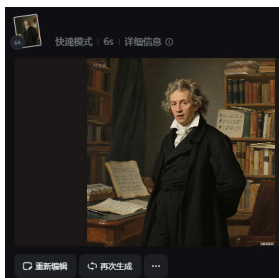


图 3-12 生成视频

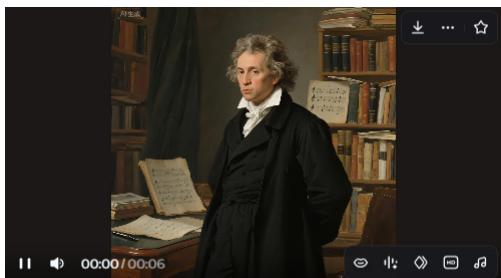


图 3-13 编辑视频并下载视频

### 3.1.2 AI 赋能“课本人物开口说话”

课本人物开口说话技术是通过 AI 语音合成与面部表情生成技术，将静态插图转化为动态形象，实现语言知识的情境化转译。典型应用如李白讲解古诗、爱因斯坦用英语对话等，能够有效提升学生的学习兴趣与理解深度。

下面以 AI 工具即梦 AI 为例，展示实现“课本人物开口说话”的具体操作步骤。

#### 步骤 1 预设课本人物图

预设一张课本封面图，如图 3-14 所示。将其中的人物图单独截取，保存至文件夹内备用，如图 3-15 所示。



图 3-14 预设的课本封面图



图 3-15 课本人物图

**步骤 2** 打开即梦 AI 上传图片

进入即梦 AI 首页，选择“图片生成”功能，如图 3-16 所示。上传课本人物图，选择“智能参考”维度，单击“保存”按钮，如图 3-17 所示。



图 3-16 选择“图片生成”功能

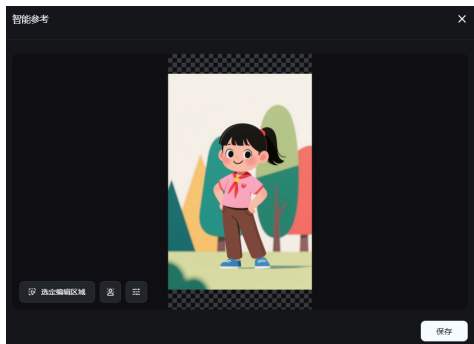


图 3-17 上传课本人物图

**步骤 3** 完成“对图片进行 3D 渲染”的指令

输入指令：“帮我生成图片，风格为 3D 渲染图，原比例，背景为纯白色，保持原图片人物形象。”，单击对话框右侧的箭头按钮发送指令，如图 3-18 所示。生成效果如图 3-19 所示。



图 3-18 输入指令并发送



图 3-19 3D 渲染效果

#### 步骤 4 选择图片并处理

选择一张最适合的图，单击页面右侧“超清”按钮对图片进行高清处理，如图 3-20 所示。高清处理完毕后单击图片右上角的下载按钮，保存图片，如图 3-21 所示。



图 3-20 单击“超清”按钮

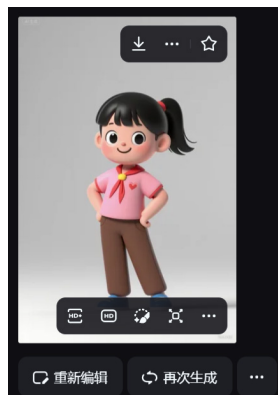


图 3-21 保存图片

#### 步骤 5 使用即梦 AI “数字人”功能

进入“数字人”功能页面，上传 3D 渲染后的课本人物形象图，如图 3-22 所示。单击“音频”按钮，选择“生成音频”，如图 3-23 所示。在文本朗读下方对话框中输入小女孩的说话内容：“同学们好，很高兴认识大家，这学期我们一起学习数学课。”，选择适合角色的音色（如儿童、小女孩），单击“确认”按钮，如图 3-24 所示。



图 3-22 上传图片



图 3-23 选择“生成音频”



图 3-24 生成音频

### 步骤 6 生成视频并下载

单击对话框右侧的箭头按钮发送指令,生成“课本人物开口说话”的视频,如图 3-25 和图 3-26 所示。在视频的右下角编辑栏可以进行重新配音、提升视频分辨率、使用 AI 配乐等操作,最后单击视频右上角的下载按钮进行视频下载,如图 3-27 所示。

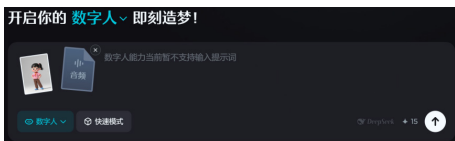


图 3-25 发送指令



图 3-26 生成视频

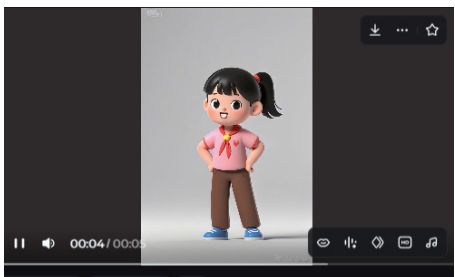


图 3-27 编辑并下载视频

### 步骤 7 打开剪映进行人物及画面的合成制作

打开剪映,将课本封面图和人物视频分别拖动到剪映轨道里,课本封面图作为第一轨道,人物视频作为第二轨道,如图 3-28 所示。

**步骤 8** 调整人物视频大小

选中人物轨道，用鼠标指针拖动人物左上角的原点缩小人物视频画面，如图 3-29 所示。

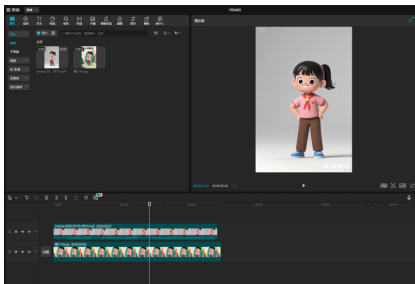


图 3-28 上传课本封面和人物视频



图 3-29 调整人物视频画面

**步骤 9** 对人物视频进行抠像

图 3-30 自定义抠像

单击右侧工具栏中的“抠像”按钮，勾选“自定义抠像”，如图 3-30 所示。选择“智能画笔”，涂抹选中人物，如图 3-31 所示。单击页面右侧的“应用效果”按钮，进行抠图，效果如图 3-32 所示。



图 3-31 选中人物图

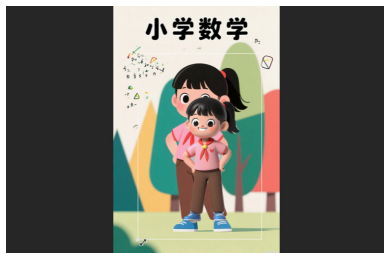


图 3-32 人物抠图效果

**步骤 10** 调整课本封面图及人物视频轨道长度

将课本封面图轨道部分裁剪，使其长度与人物视频轨道长度一致，如

图 3-33 所示。将视频人物的大小调整至课本封面图中原人物的大小，直至完全覆盖，如图 3-34 所示。

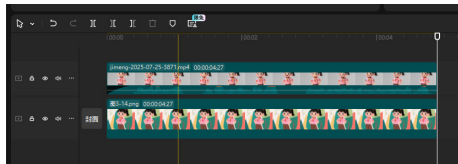



图 3-33 调整轨道长度



图 3-34 调整视频人物的大小至覆盖课本封面原人物形象

### 步骤 11 为视频添加关键帧

选中人物视频，将鼠标指针移动到视频起始位置，单击画面右侧工具栏中的 ，添加关键帧，如图 3-35 所示。

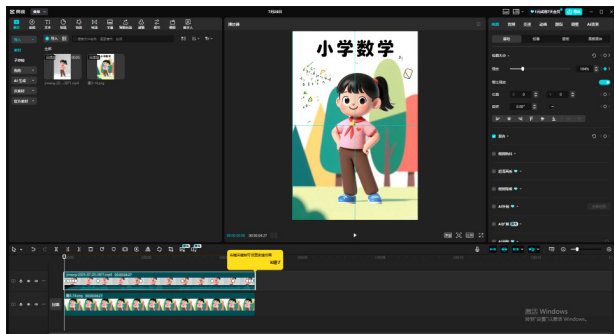


图 3-35 添加关键帧

### 步骤 12 导出视频

单击画面右上角的“导出”按钮，如图 3-36 所示。设置好各参数数值，

再单击“导出”按钮，保存视频，如图 3-37 所示。



图 3-36 单击“导出”按钮



图 3-37 设置参数并保存视频

### 3.1.3 AI 助力“小学生与诸葛亮通话”

要实现“小学生与诸葛亮通话”这一奇幻场景，需要充分发挥天马行空的想象力，通过文字向 AI 描述出小学生与诸葛亮跨越时空对话的场景，并运用文字生成图片技术将这些构想转化为具体的静态视觉素材。

下面以 AI 工具豆包和即梦 AI 为例，展示实现“小学生与诸葛亮通话”的具体操作步骤。

#### 步骤 1 打开豆包并完成“生成视频通话内容”的指令

进入豆包首页，在对话框中输入指令：“请生成四年级小学生与诸葛亮进行视频通话的对话内容。要求：小学生先自我介绍，再向诸葛亮提两个问题，主要是关于他的历史故事；一问一答，语言简洁；输出格式分两列，标注‘小学生’和‘诸葛亮’；视频时长在 30 秒左右。”，如图 3-38 所示。单击对话框