

[人工智能普及教育]

借助 XEduHub 快速开发人工智能体验作品

文_邱奕盛 / 华东师范大学教育信息技术学系，上海数字化教育装备工程技术研究中心 谢作如 / 温州科技高级中学

问题提出

学习人工智能要从体验开始，青少年只有体验过人工智能，才可能产生兴趣。为此，学校人工智能学习空间（实验室、功能教室等）的展示区域或科普走廊，应部署一些有趣的人工智能体验产品，不仅让学习空间充满人工智能味，也让学生在交互中感受人工智能的魅丽。有条件的学校还可选择采购和订制企业开发的人工智能体验作品，以追求更好的展示效果。

中小學生能否与教师一起设计这类交互性强的人工智能体验作品呢？设计作品的过程就是学习的过程，也是解决真实问题的过程。仔细观察企业开发的人工智能体验作品，会发现背后的技术并不复杂，一般都是输入的信息经过模型推理，再通过屏幕输出。常见的人工智能技术，比如目标检测、物体分类、关键点检测等，学生也能做到。

工具推荐：XEduHub

XEduHub 是上海人工智能实验室浦育团队开发的一个通用人工智能模型推理工具，既是 XEdu 项目的核心组成部分，又是一个可以独立安装的 Python 库（库名为 XEdu-python）。XEduHub 不仅支持 XEdu 系列工具训练得到的模型，还内置了常见的人工智能模型，比如人体关键点检查、80 种物体目标检测、文字识别（OCR）等，只要选择某一任务，即可自动从云端下载相应的模型。

XEduHub 设计了“工作流”（Workflow）的方式实现模型推理。每一次应用人工智能模型推理，数据都经历了输入模型—预测推理—输出结果 3 个阶段，就像物料（数据）经过生产流水线加工后产出半成品一样。在 XEduHub 中，一次推理流只需要 2 行代码，第 1 行代码通过指定 task 名称确定使用哪种模型构建工作流，第 2 行代码为输入数据并得到推理结果输出。如果要处理多组数据，只需灵活使用第 2 行代码即可。

```
Python
body = wf(task='pose_body')
```

```
keypoints,img_keypoints = body.inference(data='data/body.jpg',img_type='pil')
```

在使用 XEduHub 处理较为复杂的问题时，通常需要多个模型联动，工作流可以视为不同模型之间的数据流动。比如，当画面中有多个人像时，应先进行多人的目标检测，以确定每个人所在的位置，然后将检测到的数据传入关键点识别模型，对每个人像进行关键点识别，最终得到每个人的关键点坐标数据。

XEduHub 就像一个装满人工智能玩具的箱子，里面有很多已经集成好的人工智能模型，可以直接被用于完成不同的任务，或根据自身需求，组建属于自己的人工智能体验作品。

项目设计：经典人工智能体验作品的组成

经典的人工智能体验作品，都会有交互效果的呈现。这类作品往往会实时采集现场数据，经过人工智能推理后，再输出呈现在终端上。作品的组成可以分为感知输入、人工智能推理、输出呈现 3 部分。感知输入的常用设备包括鼠标、键盘、摄像头、麦克风、雷达扫描仪，以及各类传感器，人工智能推理可以用 XEduHub 实现，而输出呈现可以借助显示器、扬声器等实现，还可以用开源硬件的 LED、蜂鸣器、舵机等执行器实现。

本项目以摄像头作为数据采集设备，大屏幕显示器作为输出设备，展示一个经典人工智能体验作品“马赛克风格的世界”的设计和实现。当人站在一个大屏幕前，即可看到变化（马赛克）后的画面（图 1）。

```
import cv2
from XEdu.hub import Workflow as wf
cap = cv2.VideoCapture(0)
style = wf(task='gen_style',style='mosaic')
while cap.isOpened():
    ret, frame = cap.read()
    # 进行模型推理
    result, img = style.inference(data=frame, img_type='cv2')
    cv2.imshow('video', img)
    # 按q键退出程序
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break
cap.release()
cv2.destroyAllWindows()
```



图 1 “马赛克风格的世界”设计和实现

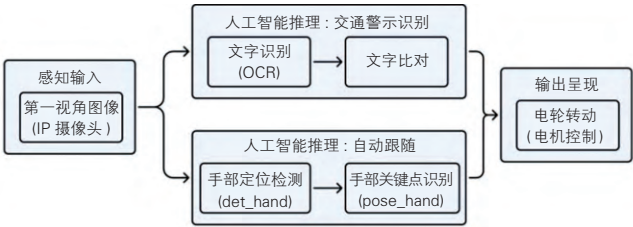


图2 交通指示识别和自动跟随功能实现过程

设计中，通过指定摄像头确定从何处作为数据输入源，接着声明一个推理流任务为“图像风格迁移”（gen_style），并指定风格为“马赛克风”（mosaic）。在循环中读取采集到的每一帧图像，对其进行风格迁移处理，并将结果显示在窗口中。

◎ 作品展示：软硬融合支持无限创意

了解 XEdu 的功能后，结合学生天马行空的想法，将感知输入、人工智能推理、输出呈现三大环节灵活组合，就能设计出各式各样的创意人工智能体验作品。

➤ 创意 1. 智行小车

身边总能看到搭载智能辅助驾驶系统的汽车，我们也能做一款小车模拟无人驾驶。由于智能驾驶涉及的技术很多，图 2 仅列出了交通指示识别和自动跟随两项技术进行模拟。交通指示识别利用了文字识别识别出多种语言的文字，通过字符串比对决定驾驶策略；自动跟随则选择让小车跟随手的移动而移动，定位检测和关键点识别支持多个类型，比如全身检测、脸部检测、手部检测等，可以通过指定不同的任务名称进行切换。核心代码如图 3 所示。

```
ocr = wf(task='ocr')
texts= ocr.inference(data=img)
det = wf(task='det_hand')
pose = wf(task='pose_hand')
# 小车控制
if len(texts)>0:
    if texts[0][0]=='stop':
        control_car('S') # 停止
    elif texts[0][0]=='left':
        control_car('L') # 左转
    elif texts[0][0]=='right':
        control_car('R') # 右转
    elif texts[0][0]=='go':
        control_car('F') # 前进
bboxes = det.inference(data=img)
# 如果检测到手
if len(bboxes):
    keypoints = hand.inference(data=img,bbox=bboxes[0])
    # 获取食指根部'5'和小指根部'17'的坐标点
    x1, y1 = keypoints[5]
    x2, y2 = keypoints[17]
    # 勾股定理计算关键点'5'和'17'之间的距离，并变成整数
    distance = int(math.sqrt((x2-x1)**2 + (y2-y1)**2))
# 小车控制
if 90 > distance > 70:
    control_car('S') # 停止
    print("stop")
elif distance >= 90:
    control_car('F') # 前进
    print("go")
elif distance <= 70:
    control_car('B') # 后退
    print("back")
```

图3 智行小车核心代码

➤ 创意 2. 坐姿检测

坐姿不良会导致腰酸背痛，可以设计一个坐姿检测算法实现及时提醒功能。首先，使用摄像头采集坐姿画面，然后提取人像关键点信息，并将这些信息传入一个训练好的二分类神经网络模型进行判断，根据坐姿良好或坐姿异常的推理结果作出相应的提醒。这个作品使用了多个模型，首先是使用人体关键点检测（pose_body）对人体关键点进行提取（图 4），然后使用一个经过个性化训练的全连接神经网络模型（basenn）进行坐姿分类，由于模型是自己训练的，还需要特别说明模型存放的位置。核心代码如图 5 所示。



图4 人体关键点检测

```
pose = wf(task='pose_body')
nn = wf(task='basenn',
        checkpoint='./姿态分类.onnx')
keypoints = pose.inference(data=img)
result = nn.inference(data=keypoints)
res = nn.format_output()
```

图5 坐姿检测核心代码

➤ 创意 3. 文本问答

此外，也可以用简单的自然语言处理（NLP）模型进行基于文本匹配的问答（nlp_qa）。只需构建一个文本问答的模型推理流，将长文本信息和问题都输入模型，模型就能够给出特定的回答。

◎ 总结与反思

借助 XEduHub，用简洁的语法、相同的代码，快速调用多种模型，经过创意的叠加，就能自由组合出各种好玩的项目。相比于采购企业开发的人工智能体验项目，自行设计更有挑战性，而且学生还能参与到项目制作过程。教师也可将设计交互作品作为人工智能课程的大作业，并将优秀作品直接展示出来。此外，如果学生能使用常见任务模型结合自己训练的模型，组建更复杂的人工智能项目，也能加深学生对人工智能的认识，特别是感知深度学习中模型训练的重要性。

本文系 2022 年度上海市自然科学基金面上项目“基于多模态数据融合的在线学习认知模型及优化研究”（编号：22ZR1421300）阶段性研究成果