

附件3

AI工程实践活动案例

AI 智能体是指能够感知环境、进行决策和执行动作的智能实体。人机共智体现了人类与 AI 智能体高效协作、共同解决现实问题的过程。“AI 工程实践”活动以“人机共智”为主题，鼓励学生结合生产生活中的真实问题，并以 AI 技术为核心，通过电子设计、机械创新等多元技术手段开展创新项目实践，完成具备展示功能的创意智能作品。活动旨在培养学生从实践角度深入理解人工智能，提升学生运用 AI 技术解决复杂工程问题的能力。

构建有效的 AI 智能体是实现人机共智的关键。学生需积极观察与发现生活中的问题，运用所学知识与技能设计出能够有效解决现实问题的 AI 智能体。AI 智能体的设计可以指向日常琐碎任务的优化，也可以指向更加复杂的现实问题解决，例如利用传感器数据和机器学习算法缓解校园交通拥堵，或是应用 AI 技术分析校园气候数据，预测植物生长的趋势等。

一、案例1：不离不弃的风扇

该案例针对家中风扇不能自动调整吹风方向的不足，结合人体关键点检测技术开发了一款能够自动聚焦用户头部并自动调整吹风方向的智能风扇，提升了用户使用风扇的便利性和舒适度。项目紧扣自动调整吹风方向这一实际需求，在硬件结构部分设计、制作和搭建了风扇的主体结构，并选择 Arduino 作为控制板，完成了作品的外观造

型。在软件工程方面，案例通过程序设计识别人体关键点，并根据识别结果控制电机与舵机，实现作品功能。（见附录 3-1）

二、案例2：基于Mediapipe的儿童窗保护

该案例针对儿童高空坠落现象，基于 AI+行空板系统边端部署设计了一款智能窗户，能够判断儿童是否进入窗户区域，并控制窗户的自动开合，以达到安全防护的效果。其中，行空板外接摄像头和舵机，负责采集视频数据和控制舵机运动；PC 端标记窗户的位置，分析行空板采集的视频数据，并判断人体与窗户的位置关系，反馈给行空板。基于项目设计，学生掌握了行空板、摄像头、舵机等硬件的使用方法，了解了人体关键点技术、局域网视频传输和激光切割技术的应用方式，并感受到了人工智能解决现实问题的优势。（见附录 3-2）

三、案例3：基于大语言模型的英语学习助手

该案例针对学生高效学习英语的需求，使用 LangChain 应用开发框架和大语言模型开发了一款个性化学习助手“慧语”，实现了错题讲解、习题布置等功能。学生使用“慧语”学习助手，首先需要拍摄题目图片，利用 OCR 识别技术将图片中的题目转换为文本，并选择题目类型和题号，完成题目的识别；接着，利用大语言模型对题目进行分析，获取题目的正确答案和详细讲解过程，还可以通过知识点梳理、错题集整理、新题布置等方式提升学生的学习效率。开发者对“慧语”学习助手进行了实验验证，实验结果表明“慧语”学习助手简单易用，具有较好的识别效果。（见附录 3-3）

附录 3-1

案例 1：不离不弃的风扇

项目基本信息：

- 项目难度：☆☆
- 推荐人群：具备 python 编程能力的学生
- 软硬件环境：python 编程环境、Arduino UNO 相关套件
- 项目来源：2023 年创新实践活动
- 项目关键字：数据处理，人体关键点，目标跟踪

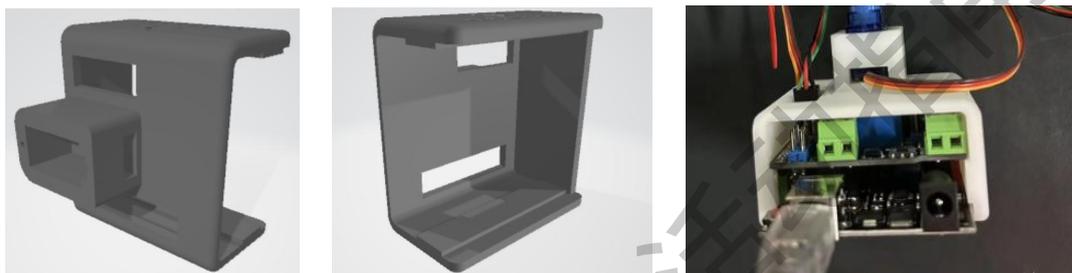
案例编写人：王继征、邱奕盛

作者所在单位：上海人工智能实验室

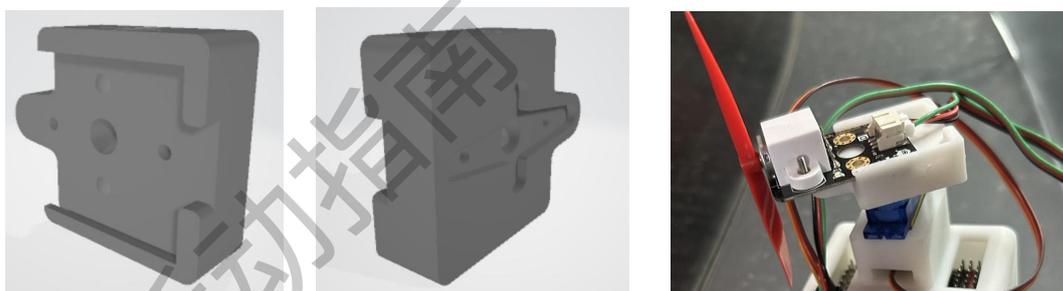
1. 项目简介

随着人工智能技术的迅猛发展，智慧家居已经成为了现代生活的一个重要组成部分。本项目从生活中真实需求出发，发现家中风扇只有规律转头和固定位置吹风，难以满足在工作时，随着人的移动而跟随吹风的需求。通过硬件结构复现真实场景，结合 AI 能力解决生活中发现的实际问题。不离不弃的风扇项目致力于开发一款能够自动聚焦用户头部并自动调整吹风方向的智能风扇，集成了人体关键点检测技术与传统风扇。本项目旨在提升用户在使用风扇时的舒适度和便利性，同时探索智能家居设备的新可能。





舵机上方需要与风扇进行连接，风扇部分需要设计一个底座与舵机连接，并且能够保证风扇的稳定，根据风扇的选型，进行了如下设计：



2.2.2 风扇结构搭建

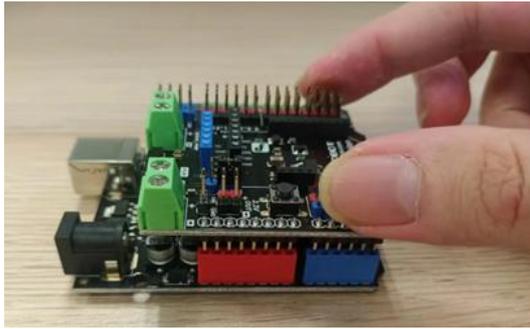
摇头风扇部分搭建在搭建工程中，应先将舵机与 3D 打印件部分进行搭建。



将风扇部分与底座进行连接，这里将风扇的马达使用卡扣进行了固定。



下面进行了控制器的组装，这里我们选择了 Arduino 控制板、IO 扩展板，所以先将这两块板子进行了连接，然后将这两块板子同时插入 3D 打印的外壳进行位置固定，最后将舵机与风扇的线路进行连接。



整体结构搭建完成后，与电脑进行连接。

2.3 任务二：完成人体关键点检测

本项目主要采用了人体关键点检测技术完成风扇的控制。人体关键点检测技术是基于人体检测的基础上完成的进一步检测。计算机需从图像或视频中识别出人体的存在及其位置，而人体关键点检测则更进一步，定位人体的特定部位，如头、手、膝盖等。



人体关键点是指人体骨骼结构中对于识别和分析人体姿态、动作特别重要的

点。这些关键点通常包括头部、颈部、肩膀、肘部、手腕、腰部、膝盖、脚踝和脚尖等部位的位置。在计算机视觉和图像处理领域，人体关键点特指通过图像处理算法能够自动检测出的这些重要位置的坐标。

pose_body17 模型能够实时逐帧检测人体的 17 个关键点，以下程序使用了 pose_body17 模型，关键点对应关系如图所示，左肩和右肩两处的关键点序号分别是 5 和 6。



Python

```
keypoints,img = pose.inference(data=img,img_type='cv2',bbox=i)
print('左肩坐标: ',keypoints[5])
print('右肩坐标: ',keypoints[6])
```

2.4 任务三：完成不离不弃的风扇

2.4.1 电机与舵机的控制

本项目中风扇控制使用电机，风扇转头使用的是舵机。这样的选型是因为电机适用于需要输出功率的场合，如驱动机器人、车辆等。电机通常用于控制机器人的走动、车辆的驱动等，可以输出足够的功率以实现收集、运输、拖拉等任务。而舵机适用于需要精准控制角度的场合，如遥控模型、机器人等。舵机通常用于控制转动舵机、云台、机械臂等，可以实现精确的定位和控制。

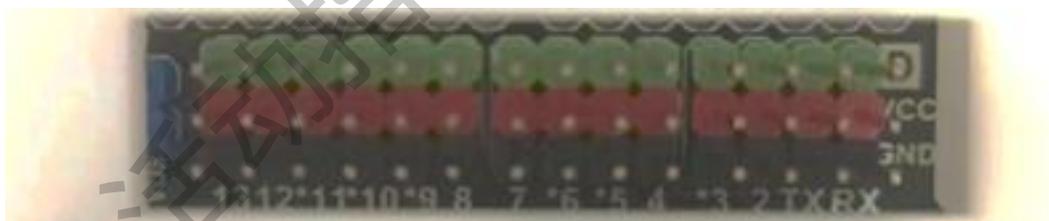


电机



舵机

为了便于在程序端控制电机与舵机，本项目借助于 PinPong 库，PinPong 库是一套控制开源硬件主控板的 Python 库，基于 Firmata 协议并兼容 MicroPython 语法。PinPong 库的名称由“Pin”和“Pong”组成，“Pin”指的是“引脚”，在硬件领域，尤其是在微控制器和开发板中，引脚是连接外部设备的基本单元，负责数据的输入输出。“Pong”寓意着信号或数据的往复传递，与库的功能相呼应，即让数据能够在不同的硬件平台之间流畅地传递和处理。



通过编写程序，能够发现电机与舵机转动是有一定的区别的，这也恰恰印证了电机与舵机在功能需求选择上的差异。在这里应该要注意舵机的角度值应为 0-180 之间的整数。另外，舵机这种转动方式称之为绝对角度转动。

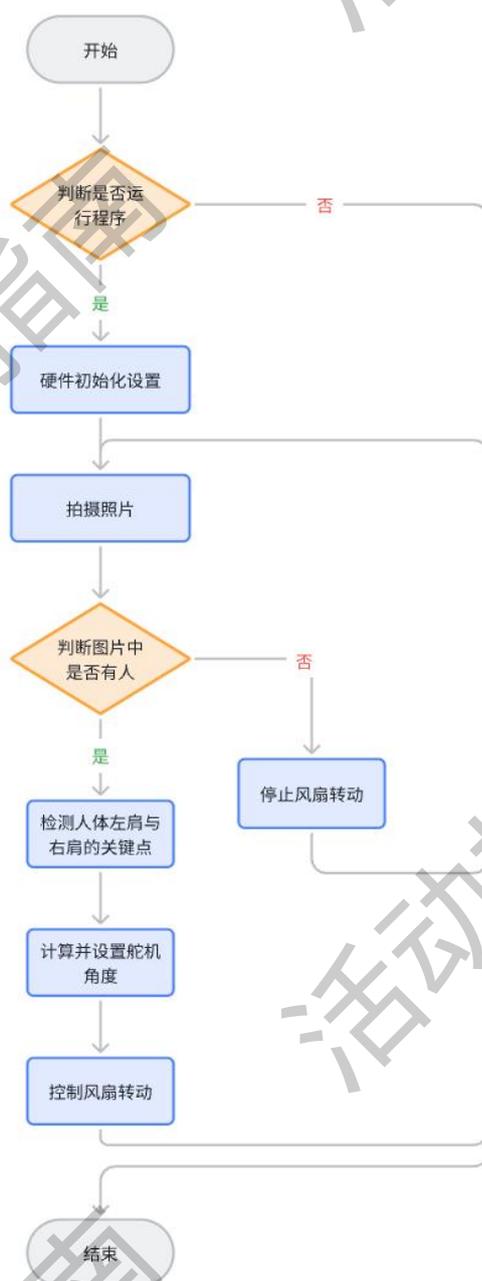
```
Python
# 电机控制
fan = Pin(Pin.D10, Pin.OUT)
fan.write_digital(1)

# 舵机控制
s1 = Servo(Pin(Pin.D9))          # 将 Pin 传入 Servo 中初始化舵机引脚
s1.write_angle(0)                # 控制舵机转到 0 度位置 方法 2
```

2.4.2 完整效果实现

结合舵机转动控制方式与人体关键点检测的内容，编写代码，完成基于摄像头对人体的识别，判断人的位置，控制风扇的开关和朝向的项目效果。

- 当摄像头看到家里没有人出现在画面中时，风扇关闭。
- 当摄像头发现家里有人出现时，开启风扇，并控制风扇下方的舵机转动，实现风扇跟着人走的效果。
- 自行设计一种远程控制开关风扇的逻辑，例如坐下就开风扇，起身离开就关风扇，并实现这套控制风扇开关的效果。



2.5 项目总结

在智能家居和人工智能技术日益普及的今天，智能风扇项目应运而生，在通过集成人体关键点检测技术。该项目不仅体现了技术创新，也展示了对提升用户

生活品质的关注。不离不弃的风扇项目致力于开发一款能够自动聚焦用户头部并自动调整吹风方向的智能风扇，集成了人体关键点检测技术与传统风扇。本项目旨在提升用户在使用风扇时的舒适度和便利性，同时探索智能家居设备的新可能。

(1) 关键技术与方法

- 人体关键点检测：采用 pose_body17 模型，实现对人体关键点的高效准确检测。
- 控制算法：开发 PID 控制算法，根据识别到的人脸位置信息实时调整风扇头的转向。
- 硬件集成：包括风扇主体、控制电路板和电机等，通过精确的硬件配置与集成，实现系统的稳定运行。

(2) 成果与成效

- 技术验证：成功实现了人体关键点检测技术与风扇自动转向的功能，验证了技术方案的可行性。
- 用户体验提升：不离不弃的风扇相比传统风扇在舒适度和便利性上有显著提升。
- 扩展潜力：项目展示了智能家居产品的新方向，为智能家居提供了新的思路。

3. 项目评价

3.1 项目优点

- 项目采用先进的人体关键点检测技术，结合智能控制算法，实现了风扇的智能追踪与风向调节功能。这一创新点不仅提升了风扇的智能化水平，也为智能家居产品开发提供了新的思路。

3.2 项目不足

- 项目完成度不足，本项目仅使用了人体两肩部的关键点，在实际生产生活中，人的动作姿势和幅度是难以预测的，应该加入更多的关键点检测来完善。
- 项目方法创新性不足，项目仅采用一种人体关键点检测的能力，可以考虑通过进一步的技术优化和功能扩展（如加入语音控制、情感识别等），项目有望吸引更多广泛的用户群体，并在智能家居领域占据一席之地。

4. 参考学习资料

https://www.bilibili.com/video/BV14P4y1Q78E/?spm_id_from=333.337.search-card.all.click&vd_source=81dec223793194b6476c9bdb0cb8802c

<https://zhuanlan.zhihu.com/p/510511329>

附录 3-2

案例 2：基于 Mediapipe 的儿童窗保护

项目基本信息：

- 项目难度：☆☆☆
- 推荐人群：具备 Python 编程基础和开源硬件使用经验的爱好者
- 软硬件环境：Python 编程环境、LaserMake、行空板
- 项目来源：2023 年全国青少年人工智能创新实践活动四星卓越作品
- 项目关键字：**Mediapipe**；儿童窗保护；人体关键点技术；局域网视频传输

案例编写人：洪丹妮、陈泓宇

作者所在单位：温州市绣山中学

1. 项目简介

高空坠落是导致儿童意外伤亡的主要原因之一，而儿童高空坠落的事件多数是可以预防的。除了家长加强对孩子们的安全教育外，对窗户进行安全设计也十分重要。因此，我们期望设计一款智能窗户，可以判断儿童是否进入窗户区域而进行开窗或关窗的命令，达到安全防护的效果。

本项目基于 AI+行空板系统边端部署，分为 PC 端和行空板两大核心。行空板外接摄像头和舵机（安装在窗户模型上），负责采集视频数据和控制舵机运动；PC 端标记窗户的位置，分析行空板采集的视频数据，判断人体与窗户的位置关系并反馈给行空板。

通过智能窗户的设计，引导学生利用人体关键点技术、局域网视频传输方式和激光切割技术，设计并打造属于自己的智能窗户，实现对窗户自动开合的控制，模拟实现对儿童高空坠落的安全防护，体验人工智能为美好生活带来的幸福感。

1.1 项目展示



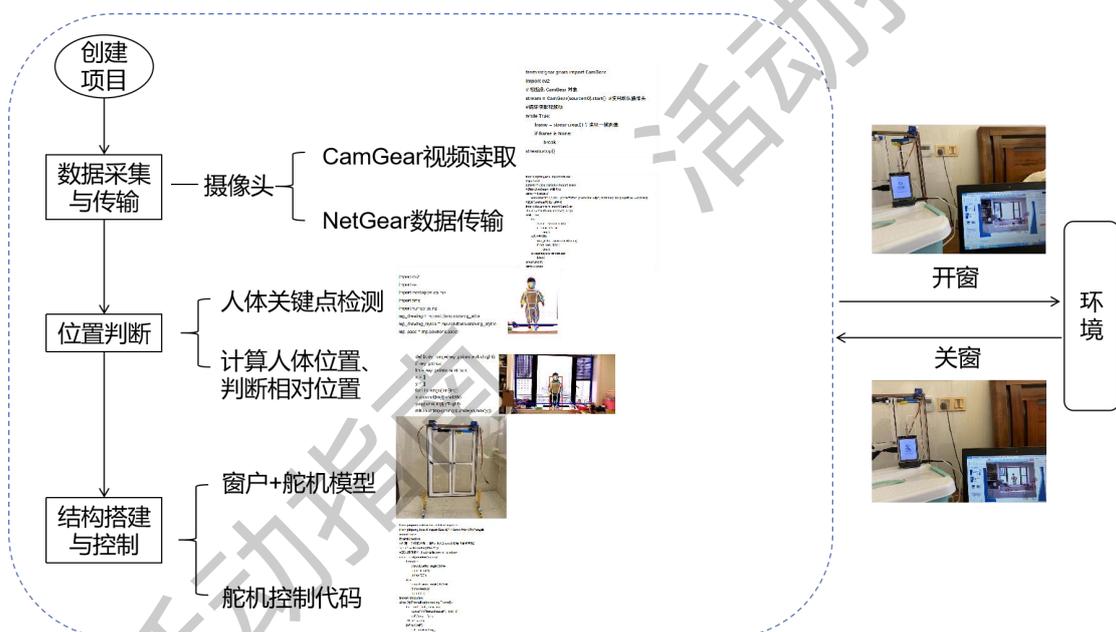
1.2 项目创新点

- 基于人体关键点技术实现了对是否有儿童处于窗户位置的判定；
- 创新性地采用局域网视频传输的方式，实现了借助行空板系统的边端部署，能够控制窗户的自动开合。

2. 项目实现流程

2.1 整体流程图

如下图所示，《基于 Mediapipe 的儿童窗保护项目》可拆分为数据采集与传输、位置判断、结构搭建与控制三个部分。



数据采集与传输部分，外接摄像头的行空板没有足够的算力计算人体与窗户的相对位置关系，因此，可将其作为采集实时视频画面的设备，将数据传送给 PC 端进行处理。借助 VidGear 库的 CamGear 实现视频读取、NetGear 实现数据的双向传输，为后续位置判断和舵机控制搭建通信基础。

位置判断部分，PC 端作为 client 端，负责计算人体与窗户的相对位置关系。首先需要标定固定窗户的位置；然后再借助 Mediapipe 工具库提取人体的 33 个关键点信息进行极值计算，得出极小值点 $A(\min(x), \min(y))$ 和极大值点 $B(\max(x), \max(y))$ ，从而用这两个点形成的矩形代表人体位置；最后通过计算人体矩形与窗户范围是否存在部分重合，来判断人体与窗体的位置关系属于“in danger”还是“out danger”。

结构搭建与控制部分，需要完成窗户的结构搭建和舵机的实时控制。项目使用的是行空板（UniHiker），因此可以选择 Pinpong 库，通过初始化行空板、创建舵机对象以及调用相应的控制方法，方便地设置舵机的转动角度，从而实现对舵机的精确控制。

2.2 任务一：数据采集与传输

2.2.1 算法介绍

行空板作为 server 端，把摄像头采集的图片发送给 PC 端，让 PC 进行下一步人体关键点信息提取、与窗户的位置判断等处理。



VidGear 是一个强大的 Python 视频处理库，它提供了许多有用的工具和类，用于实时视频处理，让用户只需几行代码即可轻松集成和执行复杂的视频处理任务。因此，本项目使用 VidGear，实现行空板与同网段 PC 端之间的视频数据通信。

2.2.2 实现过程

(1) 使用 VidGear 中的 CamGear 实现视频读取

CamGear 类是 VidGear 库中的一个关键部分，用于从摄像头或其他视频源读取视频流。下面代码展示了行空板外接摄像头后，如何使用 CamGear 类从行空板的外接摄像头读取视频流。

```
from vidgear.gears import CamGear
import cv2
# 初始化 CamGear 对象
stream = CamGear(source=0).start() #使用默认摄像头
#循环读取视频帧
while True:
    frame = stream.read() # 读取一帧图像
    if frame is None:
        break
    stream.stop()
```

图 2 应用 CamGear 类读取视频

(2) 使用 VidGear 中的 NetGear 实现数据传输

NetGear 类是 VidGear 库中用于实现视频流传输的组件，允许将摄像头或其他视频源的实时视频流通过网络发送到其他设备或应用程序，通常用于构建分布式视频处理系统，例如远程监控。下面代码展示了行空板连接 PC 后，如何使用 NetGear 来发送视频流。

```
from vidgear.gears import NetGear
import cv2
options = {"bidirectional_mode": True}
#初始化NetGear作为服务器
server = NetGear(
    address="10.1.2.101", port="5454", protocol="udp", pattern=1, logging=True, **options)
#使用CamGear读取的视频帧
from vidgear.gears import CamGear
stream = CamGear(source=0).start()
while True:
    try:
        frame = stream.read()
        if frame is None:
            break
        #发送视频帧
        recv_data = server.send(frame)
        if not recv_data:
            break
    except KeyboardInterrupt:
        break
    stream.stop()
    server.close()
```

```
from vidgear.gears import NetGear
import cv2
options = {"bidirectional_mode": True}
#初始化NetGear作为客户端
client = NetGear(
    address="10.1.2.101", port="5454", protocol="udp", pattern=1, receive_mode=True, logging=
)
in_r=False
while True:
    server_data,image = client.recv(return_data=in_r)
    if image is None:
        break
    cv2.imshow("Received Video", image)
    if cv2.waitKey(5) & 0xFF == 27:
        break
    cv2.destroyAllWindows()
    client.close()
```

图 3 服务器端应用 NetGear 类传输数据 图 4 客户端应用 NetGear 类传输数据

在这个例子中，NetGear 被配置为服务器模式，并监听本地的 TCP 端口 5454。然后，它使用 CamGear 从默认摄像头读取视频帧，并通过 server.send(frame)将每一帧发送出去。

PC 作为接收端，还需要使用另一个 NetGear 实例来接受视频流,代码图 4 所示。

在这个例子中，NetGear 被配置为客户端模式，并连接到之前服务器设置的地址和端口。然后，它循环调用 `client.recv()` 来接收视频帧，并对每一帧进行处理，这里只是简单显示，并将 `in_r` 的值返回给行空板（服务器端）。

2.2.3 实现结果

借助 VidGear 视频处理库，能将连接行空板的摄像头采集的数据发送到 PC 端并显示，实现了图像数据的传输。



图 5 图像数据传输效果

2.3 任务二：人体与窗户相对位置判断

2.3.1 算法介绍

人体与窗户相对位置判断是通过计算人体与窗户的相对位置关系实现的。对于固定的窗户，可以选择窗户左上角和右下角这两点来标定窗户，表示窗户的位置；人体的位置可以通过提取人体关键点进行极值计算，得出极小值点 $A(\min(x), \min(y))$ 和极大值点 $B(\max(x), \max(y))$ ，用这两个点形成的矩形表示，最终通过计算人体矩形与窗户范围是否存在部分重合，来判断人体与窗体的位置关系属于“in danger”还是“out danger”。

2.3.2 实现过程

(1) 标定窗户位置

(a) 固定摄像头，使其能完整拍摄到窗户（监视儿童的范围），拍摄并保存图片。

(b) 选取窗户的左上角(177, 81)和右下角(528, 318)两个位置，通过两个点标定并用矩形框出窗户位置。

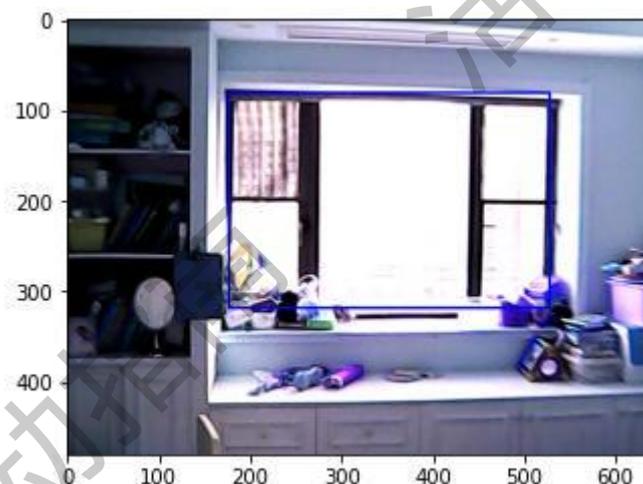


图 6 标定窗户位置

(2) 提取人体关键点

Mediapipe 实际上是一个集成的机器学习视觉算法的工具库，包含了人脸检测、人脸关键点、手势识别、头像分割和姿态识别等各种模型。借助 mediapipe 姿态识别，可提取出 33 个人体关键点信息，只要在模型中输出指定点的数据，就能获取到对应该点的人体位置信息。

本项目应用人体姿态识别提取 33 个人体关键点，代码和效果如图所示。

```
import cv2
import os
import mediapipe as mp
import time
import numpy as np
mp_drawing = mp.solutions.drawing_utils
mp_drawing_styles = mp.solutions.drawing_styles
mp_pose = mp.solutions.pose
```



图 7 人体关键点提取

(3) 计算人体位置

对提取的 33 个人体关键点进行极值计算，能得出所有点中横坐标 X 的最大值和最小值，纵坐标 Y 的最大值和最小值，进一步构造出 2 个新点：左上角点 $A(\min(x),\min(y))$ 、右下角点 $B(\max(x),\max(y))$ ，利用这两个点框出矩形，表示人体范围。

```
def body_range(key_points,width,height):  
    if key_points:  
        lm = key_points.landmark  
        x = []  
        y = []  
        for i in range(len(lm)):  
            x.append(lm[i].x*width)  
            y.append(lm[i].y*height)  
        return (min(x),min(y)),(max(x),max(y))
```



图 8 计算人体位置

(4) 判断相对位置

人体与窗户的相对位置关系，可以分成完全重叠、部分重叠、不存在重叠三种情况。本项目中的人体和窗户位置是用矩形表示的，通过列举两个矩形之间的相对位置关系，我们发现，当人体范围的矩形与窗户范围的矩形存在部分重合时，就可判断人体存在危险。人体与窗户的相对位置关系，如下图所示的 4 种情况。

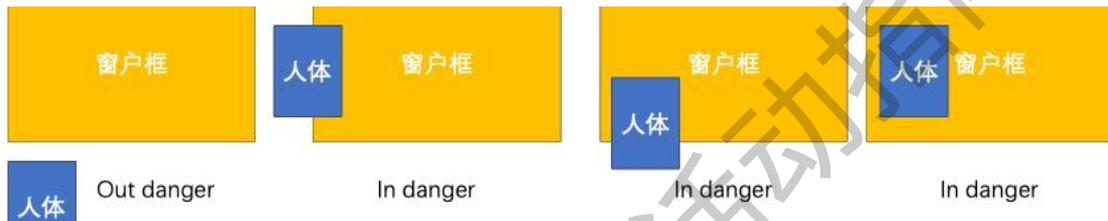


图 9 两个矩形常见的位置关系

- (a) 当人体矩形框完全不在窗户矩形框中，属于 out danger;
- (b) 当人体矩形框仅一部分落在窗户矩形框中，人体矩形框点 $B(\max(x),\max(y))$ 落在内，属于 in danger;
- (c) 当人体矩形框仅一部分落在窗户矩形框中，人体矩形框点 $A(\min(x),\min(y))$ 落在内，属于 in danger。
- (d) 当人体矩形框完全落在窗户矩形框中，属于 in danger;

因此，通过判断点 A(min(x),min(y))、点 B(max(x),max(y))是否在窗户范围内，判断人体与窗体的位置关系属于“in danger”还是“out danger”。

定义函数 in_range () 计算位置关系：

```
#定义函数：计算人体与窗户的相对位置
def in_range(point1,point2,range):
    if point1 is None or point2 is None:
        return False
    if point1[0] > range[0][0] and point1[0] < range[1][0] and point1[1] >range[0][1] and point1[1] < range[1][1]:
        return True
    if point2[0] > range[0][0] and point2[0] < range[1][0] and point2[1] >range[0][1] and point2[1] < range[1][1]:
        return True
    return False
```

图 10 定义 in_range 函数计算位置关系

调用函数 in_range () 输出计算结论：

```
# 调用函数计算人体位置，生成结论
if results.pose_landmarks is not None:
    body_r=body_range(results.pose_landmarks,width,height)
    in_r = in_range(body_r[0],body_r[1],window_location)
    if in_r:
        cv2.putText(image,"in danger",(10,30), cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 0, 255), 2)
    else:
        cv2.putText(image, "out danger", (10, 30), cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 0, 255), 2)
    print(in_r)
```

图 11 调用 in_range 函数计算位置关系

2.3.3 实现结果

当图像中儿童站在窗户边，程序用红框框出人体位置，用蓝框框出窗户区域，人体矩形框(A(min(x),min(y))、点 B(max(x),max(y)))大部分落入窗户矩形框中，屏幕中显示出提示语：in danger。

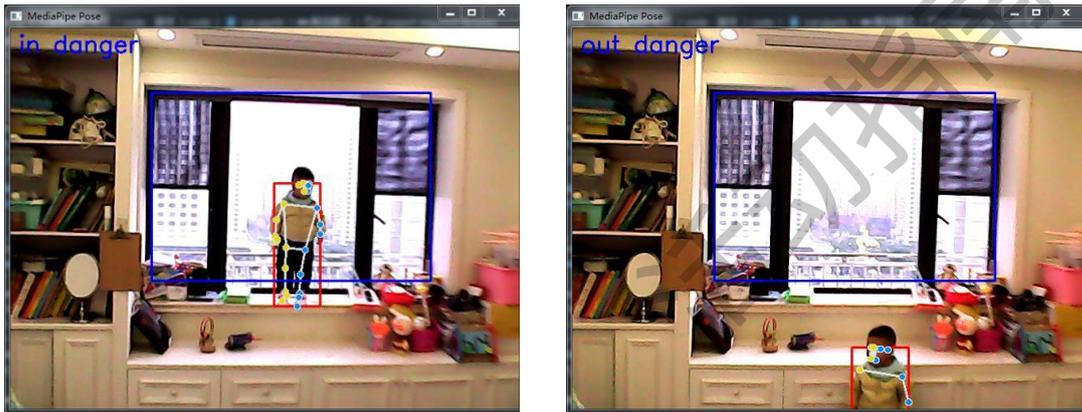


图 12 点 A 落入矩形框：in danger 图 13 点 A、B 都在矩形框外：out danger

当图像中儿童远离窗户，程序用红框框出人体位置，用蓝框框出窗户区域，人体矩形框(A(min(x),min(y))、B(max(x),max(y)))完全不在窗户矩形框中，屏幕中显示出提示语：out danger。

2.4 任务三：结构搭建与控制

2.4.1 算法介绍

行空板作为 server 端，首先要把摄像头采集的图片发送给 PC 端，让 PC 进行人体与窗户的相对位置计算。然后，行空板接收 PC 端的计算结果，根据结论控制舵机运动，从而达到实时检测及反馈的效果。

Pinpong 库作为一套控制开源硬件主控板的 Python 库，能够与行空板完美结合，使得开发者能够直接使用 Python 代码来编程和控制硬件。其原理是给开源硬件烧录一个特定的固件，使开源硬件可以通过串口与电脑通讯。执行各种命令。在控制舵机运动方面，Pinpong 库提供了简洁而强大的 API，通过初始化行空板、创建舵机对象以及调用相应的控制方法，可以方便地设置舵机的转动角度、速度等参数，从而实现对舵机的精确控制。

2.4.2 实现过程

(1) 搭建结构

(a) 设计与切割。使用 LaserMaker 软件设计窗户模型，利用激光切割机切割出窗户窗框及窗户架子。

(b) 搭建与安装。利用胶水粘连窗框和架子，每扇窗户上下各安装一只合页，用于转动窗户。选择乐高零件搭建支架，固定窗户，方便演示。



图 14 模拟窗户实物

(2) 控制舵机

连接舵机与行空板，用行空板（UniHiker）和 Pinpong 库来控制舵机的运动，从而模拟一个窗户的开关功能，以实现“开窗”与“关窗”效果。

下面这段代码定义了函数 `control_window()` 和多线程类 `MyThreading()`。其中，`control_window()` 函数用于控制舵机运动，根据传入的 `mode` 参数决定舵机的角度。若 `mode` 为 `True`，舵机转到 0 度（关闭窗户状态）；若 `mode` 为 `False`，舵机转到 90 度（打开窗户状态）；`MyThreading()` 用于创建多线程，以便并行执行 `control_window()` 函数。

```

from pinpong.extension.unihiker import *
from pinpong.board import Board,Pin,Servo #导入PinPong库
import time
Board().begin()
#创建一个舵机对象，将Pin传入Servo中初始化舵机引脚
servo1 = Servo(Pin(Pin.P2))
#定义控制窗户开关的函数control_window
def control_window(mode):
    if mode:
        servo1.write_angle(0)#关
        time.sleep(1)
        print("90")
    else:
        servo1.write_angle(90)#开
        time.sleep(1)
        print("0")
import threading
class MyThreading(threading.Thread):
    def __init__(self, func, arg):
        super(MyThreading,self).__init__()
        self.func = func
        self.arg = arg
    def run(self):
        self.func(self.arg)

```

图 15 utils.py

```

from vidgear.gears import NetGear
from vidgear.gears import CamGear
import cv2
import time
from utils import *
stream = CamGear(source=0).start()
options = {"bidirectional_mode": True}
server = NetGear(
    address="10.1.2.101", port="5454", protocol="udp", pattern=1, logging=True,
    **options
)
while True:
    try:
        frame = stream.read()
        if frame is None:
            break
        recv_data = server.send(frame)
        if not (recv_data is None):
            obj = MyThreading(control_window, recv_data)
            obj.start()
    except KeyboardInterrupt:
        break
stream.stop()
server.close()

```

图 16 run.py

图 16 的代码中，主要作用是建立一个基于视频流的实时监控系統，并通过网络传输视频帧数据，然后基于接收到的数据来触发一个控制舵机的线程，模拟

窗户的开关。具体包括视频捕获、网络传输、数据处理与线程控制和异常处理四个部分，其中数据处理与线程控制对应本环节的舵机控制。

在无限循环中，行空板作为服务器，不断读取摄像头捕获的每一帧（frame）。然后，通过 `server.send(frame)` 将帧发送到作为客户端的 PC 端。如果服务器接收到了返回数据（`recv_data`），则根据这个数据创建一个新的线程（`MyThreading` 的实例），并传入 `control_window` 函数和接收到的数据 `recv_data` 作为参数。这个线程将负责根据 `recv_data` 的值来调用 `control_window` 函数，从而控制舵机的角度（即模拟窗户的开关）。

2.4.3 实现结果

借助 `Pinpong` 库，实现了对舵机的控制。

当屏幕中的孩子处于危险状态（见下图），行空板接收到 PC 端 `in_r` 为 `True` 的反馈，控制舵机关闭窗口，PC 端屏幕显示 `in danger`。



图 17 危险区域-关闭窗口



图 18 安全区域-开窗

如图 18 所示，当屏幕中的孩子处于安全状态。行空板接收到 PC 端 `in_r` 为 `False` 的反馈，控制舵机打开窗户，PC 端屏幕显示 `out danger`。

2.5 项目总结

本项目旨在通过智能窗户的设计，实现对儿童高空坠落的安全防护。项目分为 PC 端和行空板两大核心，具体包括窗户机械结构的搭建、系统初始化、PC 端人体检测和行空板数据采集及处理等步骤。

通过利用人体关键点技术和局域网视频传输方式，实现了对窗户自动开合的控制。实验结果表明，在单一场景和多场景的测试中，系统的准确率达到 80% 以上，证明了该系统在多场景和单场景的识别和控制方面具有较高的准确率。

总的来说，本项目在儿童安全方面有着重要的应用价值。通过智能窗户的设计和人体关键点技术的应用，可以有效地预防儿童高空坠落事故的发生。未来，可以进一步完善系统功能，例如检测大人和小孩是否同时出现等，提高系统的安全性和实用性。通过这个项目，我们深刻认识到了安全教育的重要性，并希望通过科技的力量，能够改善儿童高空坠落的频发现状，让人工智能助力生活更加美好。

3. 项目评价

从**主题契合度**看，《儿童保护窗》这个项目，非常的切合生活，也借助了AI技术打造了一种切实可行的方案，并进行了实践与实施。

从**创新性**角度看，项目的选题具有一定新意，但是略显不足；采用人体关键技术的方法进行人体位置的判定，获取的信息更加丰富，使得判定更加的准确，但本项目中并未完全体现关键技术优势。图像数据进行无线传输在一定程度上，解决了行空板算力不足的问题，能比较好地提升系统性能。

从**系统性**角度看，项目完成度较高，借助摄像头进行感知，并做出是否需要关窗的判断，从而控制执行机构进行开关窗，比较好的完成了既定任务。但是系统鲁棒性有待考证，未提供相关的实验数据。

从**AI技术应用**角度看，虽然项目中没有涉及到模型的训练，但是通过AI技术来监测儿童的安全状态，并在必要时自动关闭窗户，防止儿童发生意外。正是AI智能体感知周围环境并依据所收集的信息做出决策以实现既定目标的体现。当AI智能体检测到儿童处于危险区域时，它会自动关闭窗户；反之，则会开启窗户，为儿童提供舒适的室内环境。

4. 参考学习资料

[1] 钟声标.窗户安全装置设计研究[J].门窗,2013,(12):259-260.

[2] 望运丹,胡德英,孙瑞.综合医院住院病人经窗户坠楼自杀的现况分析及安全防护管理 [J].护理研究,2018,32(10):1601-1603.

[3] 袁梅.防坠落安全窗[J].发明与创新(中学生),2014,(08):19-20.

[4]https://blog.csdn.net/qq_64605223/article/details/125606507

[5][常用工具] Python 视频处理库 VidGear 使用指北-CSDN 博客

附录 3-3

案例 3：基于大语言模型的英语学习助手

项目基本信息：

- 项目难度：☆☆☆☆☆
- 推荐人群：有较强 python 编程基础，了解大语言模型与计算机视觉基本知识
- 项目来源：2023 年创新实践活动 5 星卓越作品
- 项目关键字：**ocr**，**提示词工程**，**大语言模型调用**

案例编写人：高凯^[1]、王海涛^[2]

作者所在单位：

[1]北京二中

[2]上海人工智能实验室

1. 项目简介：

许多学生在英语学习过程中面临各种挑战和问题，如何让学生更高效的学习英语是一个重要的课题。随着生成式人工智能的迅速发展，大语言模型展现出了出色的英语表达与理解能力。以此为基础，本项目开发一个针对学生英语学习需求的个性化学习助手“慧语”，结合英语教育知识库和直观的用户界面，可以对话式的讲解错题、出练习题等。

首先，慧语利用 OCR 识别技术将图片中的题目转换为文本，学生只需拍摄题目图片并选择相关类型和题号，系统即可识别题目。接着，我们利用大语言模型对题目进行分析，慧语不仅可以告诉学生正确答案，还能详细讲解题目，出新的题目（举一反三）以及整理错题集和相关知识点，从而提升学生的学习效率。我们使用 LangChain 应用开发框架作为工具，调用大语言模型，并开发了易用的用户界面。最后，我们对所开发的系统进行了实验验证，表明该慧语具有较高的识别效果，易用性好。

1.1 项目展示



1.2 项目创新点

在这个项目中，我们利用语言大模型开发了一个名为“慧语”的英语学习工具，旨在个性化的帮助学生高效学习英语。“慧语”可以像老师一样根据学生的问题，提供个性化的答案和详解，并可以生成知识点相近的新题目帮助学生进行巩固，实验表明“慧语”答复的正确率较高，讲解也很详细有针对性。

项目的创新点包括几个方面：

创新点一：“慧语”通过设计并优化提示，以 LangChain 作为开发框架，将大语言模型的英语能力转化为一个可以解题、讲题并出题的英语学习助手，并在实践中取得了优异的效果。

创新点二：“慧语”可以针对学生的错题和提问进行回答，并可以出新题帮助学生巩固知识，相比之前依赖固定题库和讲解的工具，“慧语”个性化讲解有利于提升学生的学习效率。

创新点三：“慧语”采用多种技术提升效率和用户体验，包括使用 OCR 识别技术来解决手动输入题目过程繁琐、效率低下的问题，使用 Gradio 开发学生友好的用户界面。

2. 项目实现流程

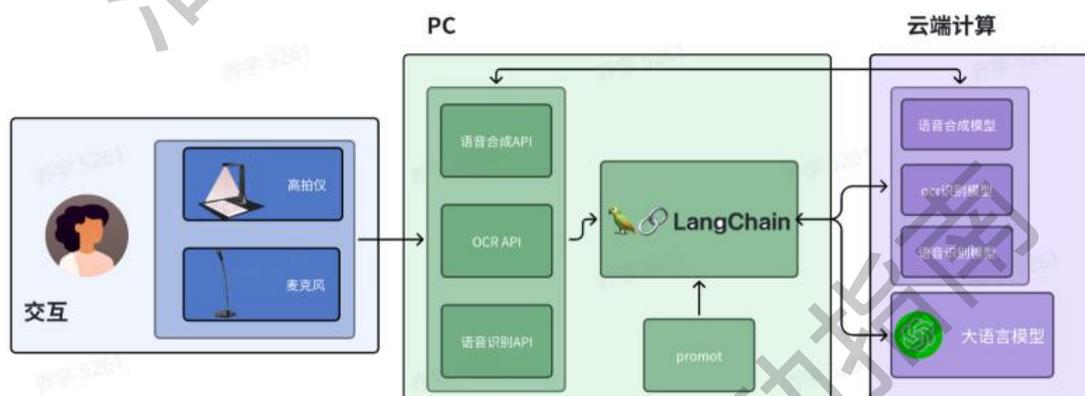
本项目研究目的是利用大语言模型开发一个可以教学生学英语的 AI 助手“慧语”，针对学生的个性化情况进行教学。这个研究中，我们重点针对中学生学习中的试题讲解，错题纠正薄弱环节巩固等需求进行开发。满足能够对题目进行解答与讲解，以及进一步给出一道新题。并能将历史学习数据进行整理记录，形成个人的数据库、错题集。

概括起来，“慧语”英语学习助手有以下特点：

- (1) 能够对题目进行解答与讲解；
- (2) 能够根据学生学习情况出新的题目，进行个性化练习；
- (3) 友好的用户交互界面。

2.1 系统整体架构

如下图所示，“慧语”系统架构主要分为用户交互、边端部署和云端计算三大部分。本项目采用 LangChain 构建语言大模型应用，将语言模型连接到其他数据源，并与外部环境进行交互。以图像交互方式为例：首先利用交互端的高拍仪扫描课本或试卷等文字材料，获得输入图像；第二步通过边端的 OCR API 读取输入图像并调用云端的 OCR 识别模型解析图中的文字，从而得到题目文字信息；第三步通过将提示信息（prompt）、题目信息和指令信息分别输入云端的语言大模型，语言模型根据输入信息生成文字答案，然后将文字信号传输至边端设备；最后在交互端的屏幕上显示语言大模型生成的文字内容。



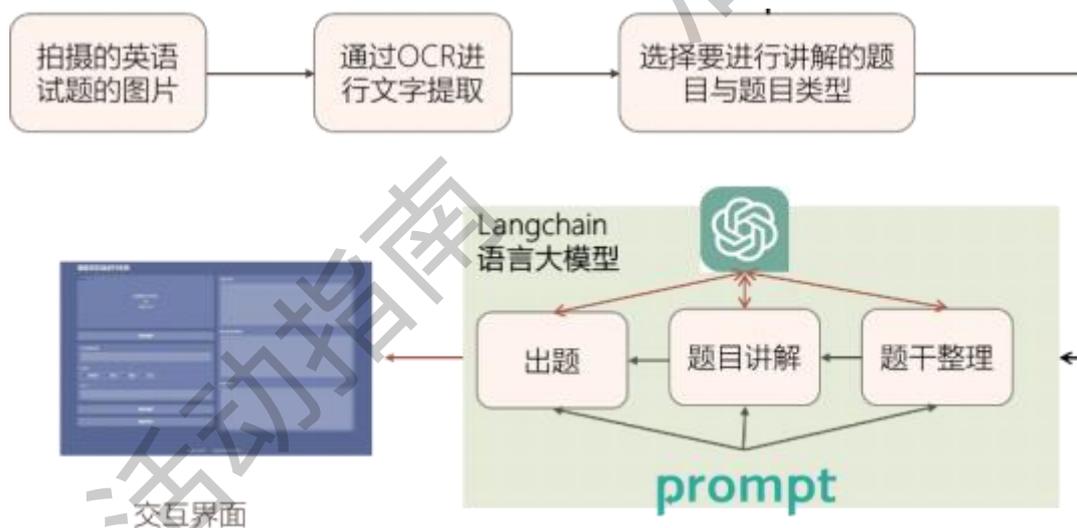
2.2 实现路线图

从下面的技术路线图，可以看出，本项目的核心内容是采用 LangChain 方法的进行英语题目的讲解与相应知识点的出题。

首先，由于手动输入题目效率低、用户体验感不好。为简化这一过程，利用 OCR 提取图片中文字，学生只需拍摄一张较清晰的题目图片，便可转换为文本。其次，由于英语题目种类较多，为使准确度提高，学生需要选择题目类型。而一张图片也可能包含多道题目，为避免回答目标不正确，学生也需输入相对应的题号。

然后大语言模型虽具有很强的英语能力，但仍会出现不理解题目要求、回答

错误、需要反复提示的情况。为提高它回复的准确性，需设计提示词。提示词共包含三个部分，首先，OCR 识别技术会出现细微的识别错误和格式问题。因此第一步要求它按固定格式进行处理。然后仿照给出示例的解析思路、方式回答学生输入的问题。最后，为检测学生是否真的掌握此知识点，出几道知识点类似的题目，以此进行巩固。并把结果输出至界面。



2.3 用户交互界面设计

采用简介的界面设计，其他的内容均可以通过直接与大语言模型进行对话实现：如进一步的问题讲解，出一套阶段性试题等。



界面初始画面



界面运行界面与介绍

2.4 OCR 识别实现

由于手动输入题目过程繁琐、效率低下。所以使用 OCR 识别技术自动提升题目输入的效率。OCR (Optical Character Recognition, 光学字符识别) 是指利用人工智能算法将图像或照片中的字符识别成计算机中文本格式的过程，识别的结果供软件进一步编辑加工的技术。

2.5 cnocr 的本地方式实现

首先我们尝试了，CnOcr Python 库来进行识别，因本项目主题为英语学习助手，所以具体选用了 CnOcr 的英文识别模型 en_PP-OCrv3。

具体的讲，我们用这个模型主要完成两个部分的功能：

- (1) 文本检测：检测出图片中文字所在的位置；
- (2) 文字识别：识别包含文字的图片局部，预测具体的文字。

实现代码如下：

```
from cnocr import CnOcr  
  
ocr1= CnOcr(det_model_name='densenet_lite_136-fc',rec_model_name='en_PP-OCrv3')  
  
out = ocr1.ocr("3.png")
```

使用该模型后，成功地将文字转化为文本，对于清晰的图片识别效果很好，基本无误。但是尝试直接使用高拍仪进行拍摄，或是手机照片拍摄的时候，会受到光线的影晌比较大。

2.6 云端方式实现

微软 Azure OCR 是微软 Azure 云平台提供的一项计算机视觉服务，旨在将印刷或手写文本从图像或扫描文档中自动识别和提取出来。它使用深度学习和自然语言处理技术，能够将图像中的文字转化为可编辑的文本数据，使得文档数字化、文本搜索和信息提取变得更加容易和高效。

微软 Azure 实现 OCR 是借助云端的算力进行运算后将结果返回的处理方法。实现的步骤包含三步：

- (1) 创建 Microsoft Azure 订阅账户：Azure 官方网站 <https://azure.com> 创建一个免费的试用账户，并订阅计算机视觉服务，并获取 Azure OCR API 密钥。
- (2) 安装必要的库。
- (3) 基于参考教程进行代码修改，主要为代码的封装，实现模块化调用该功能，方便主程序中对该部分功能的调用。

2.7 两种方法的对比

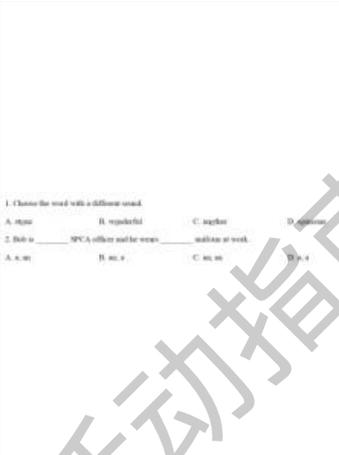
- (1) 优缺点对比

实现方式	优点	不足
基于 cnocr 的本地方式	无需联网，跨平台，免费。	需要本地有一定算力，识别效果比较差

基于 Azure 的云端方式	识别效果比较好，无算力要求。	受网络影响，需要付费，跨平台能力比较差。
----------------	----------------	----------------------

(2) 识别效果

准确度的判定采用的是 difflib 库的 SequenceMatcher 方法。

数据来源	数据图片	cnocr 识别结果	Azure 识别结果	cnocr 准确度	Azure 准确度
电脑截图		1. Choose the word with a different sound A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work A. a; an B. an; an C. an; an D. aa	1. Choose the word with a different sound. A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work. A. a; an B. an; a C. an; an D. a; a	89.02%	99.71%
手机拍照		1. Choose the word with a different sound A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work A. aan B. ana C. anan D. aa	1. Choose the word with a different sound. A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work. A. a; an B. an; a C. an; an D. a; a	90.06%	99.71%
摄像头拍照		1. Choose the word with a different sound A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work A. a; an B. an; a C. an; an D. a; a	1. Choose the word with a different sound. A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work. A. a; an B. an; a C. an; an D. a; a	83.10%	96.27%
夜间高拍仪拍		1. Choose the word with a different sound A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work A. a; an B. an; a C. an; an D. a; a	1. Choose the word with a different sound. A. stone B. wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform at work. A. a; an B. an; a C. an; an D. a; a	75.88%	93.49%

照		D. someone SPCA officer and he wears uniform work. An Bm Cann D.	wonderful C. another D. someone 2. Bob is SPCA officer and he wears uniform ut work. A & an C. an; an D. R. a		
---	--	--	---	--	--

文字的识别将直接影响后面答题的效果，考虑到最终实际的使用场景中很多存在光线等问题，我们最终选择了云端方式实现。

2.8 基于 LangChain 实现英语学习助手的搭建

LangChain 是语言模型的应用程序设计和开发的架构。LangChain 可以将 LLM 模型、向量数据库、交互层 Prompt、外部知识、外部工具整合到一起，进而可以自由构建 LLM 应用。借助 LLM 我们能够更轻量的开发程序。

2.9 初期尝试

2.9.1 直接调用大模型

虽然网页版调用大语言模型时它能够表现出一定的记忆能力。但是直接调用大模型接口进行对话，其并不具备记忆能力。从下图可以看出，虽然大语言模型能够基本上解答了第一个关于解答题目的问题，但是在第二个出题的任务上，发生了灾难性的错误，基本上和所问的题目没有任何的相关性。

```

text = "28. Put the rug, the TV set and the sofa, please.e', 'A. on', 'B. above', 'C. betw
prompt = "讲一下下面这道题目: 28. Put the rug, 'the TV set and the sofa, please.e', 'A. on',
print(get_completion(prompt))
prompt = "出一道类似考点的题目"
print(get_completion(prompt))

```

15.1s Python

这道题目是要求将地毯、电视机和沙发放在一起，然后选择正确的介词来描述它们的位置。

选项A. on表示将物体放在另一个物体的表面上，例如将地毯放在电视机或沙发的表面上。

选项B. above表示将物体放在另一个物体的上方，例如将地毯悬挂在电视机或沙发的上方。

选项C. between表示将物体放在两个物体之间，例如将地毯放在电视机和沙发之间。

选项D. among表示将物体放在多个物体之间，例如将地毯放在电视机和沙发之间，但同时还有其他物体存在。

根据题目的描述，我们需要将地毯、电视机和沙发放在一起，而不是将地毯放在其他物体的表面上或上方，也不是将地毯放在多个物体之间。

题目: 某公司的年度销售额数据如下:

2016年: 100万
2017年: 120万
2018年: 150万
2019年: 180万

2.9.2 借助 LangChain 的 memory 功能

首先简单设计了提示词，内容为:

```
template = """你是一名英语老师，你需要回你你学生的提问。
```

```
{chat_history}
```

```
student: {human_input}
```

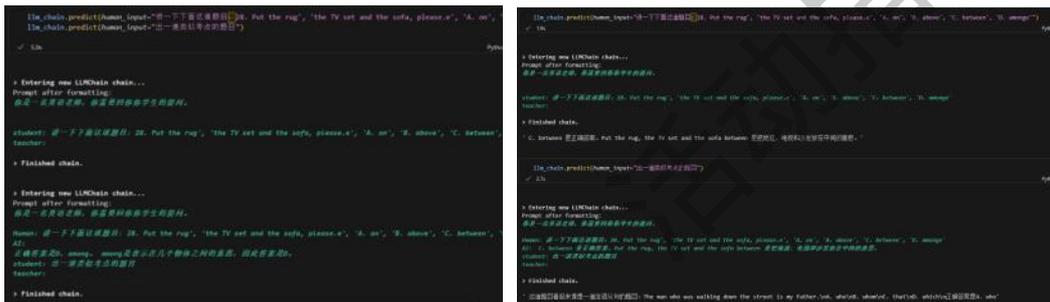
```
teacher:"""
```

然后对大语言模型进行提问了两次相同的问题。为了能够将测试结果更清晰的呈现，我将测试的结果呈现在下面这个表中。

	问题内容	第一次回答	第二次回答
问题一	讲一下下面这道题目： 28. Put the rug, 'the TV set and the sofa, please.e', 'A. on', 'B. above', 'C. between', 'D. among'"	正确答案是 D,among。 among 是表示在几个物体之间的意思，因此答案是 D。	AI: C. between 是正确答案。 Put the rug, the TV set and the sofa between 是把地岳、电视和沙发放在中间的意思。
问题二	出一道类似考点的题目	这是一道填空题： John bought a new dress, a hat and a pair of shoes. Put the dress, the hat and the shoes _____. \nA. on\nB. above\nC. between\nD. among\n 正确答案是 D: among, 因为要求把新买的衣服、帽子和鞋子放在一起，而不是放在一起之上，也不是放在一起之下，所以正确答案是 D: among. '	这道题目看起来像是一道定语从句的题目： The man who was walking down the street is my father.\nA. who\nB. whom\nC. that\nD. which\n 正确答案是 A. who

仅从两个例子就可以看出，采用本方法，能够比较有效的解决，遗忘的问题，但是也存在题目回答不准确，出的题目相关性不足，以及问题回答的一致性不足的问题。

具体程序运行的截图为：



(a) 基于 LangChain memory 的示例一 (b) 基于 LangChain memory 的示例二

基于 LangChainmemory 方法的解题示例

2.9.3 基于 LangChain 顺序链方法

Langchain 一个核心的功能在于提供了链 (chain)。可以把 Chain 理解为任务。一个 Chain 就是一个任务，当然也可以像链条一样，一个一个的执行多个链。Langchain 中提供了顺序链 (SequentialChain)。顺序链被定义为一系列链，

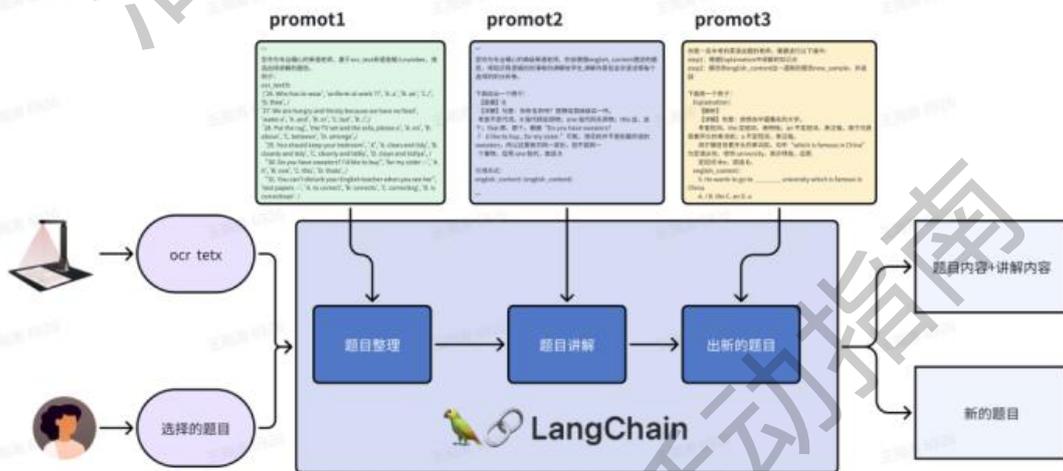
按确定的顺序调用。

在我们英语的题目的讲解的过程中，看似简单的任务，其实需要拆分成三个步骤：

- (1) 题目整理：将识别拍摄图片中的题目，整理为标准格式；
- (2) 题目讲解：讲解该题目中的核心知识点与对每个选项进行解释；
- (3) 出新的题目：出一道类似、知识点相近的题目。

顺序链刚好能够满足英语讲解过程的逐步任务。

具体的实现过程如下图。首先 OCR 识别的结果与用户输入的需要讲解的题目作为链的输入，针对每一个任务设计提示词（Prompt）。我们发现提示词对于大语言模型答复的准确程度有很大影响。为此，我们针对上述三个步骤，分别设计了不同的提示词，并对其进行了优化和改进。具体的提示词内容见附录。最后链输出“题目讲解内容”作为交互界面题目讲解的文本框内容。出新的题目作为交互界面中新的题目中的内容。



最终软件运行效果图如下：



下面通过实际的使用案例来介绍本项目中所设计的方法有效性。

选择《上海市第三女子初级中学 2022 学年第一学期七年级英语期中考试试卷》作为测试的内容，拍摄得到的图片：

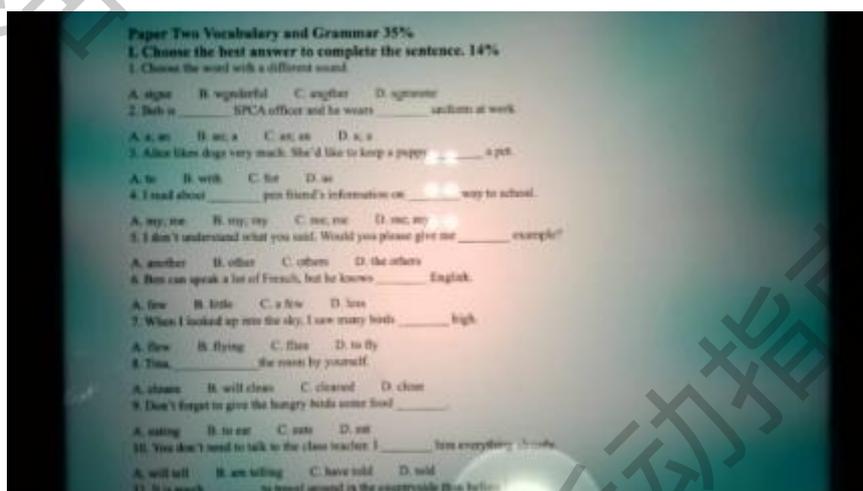


图 11 试题拍照图

图 11 中，OCR 识别的文本为：

['Paper Two Vocabulary and Grammar 35%', '1. Choose the best answer to complete the sentence. 14%', '1. Choose the word with a different sound.', 'A stone B wonderful', 'C. another', 'D. someone', '2. Bob is', 'SPCA officer and he wears', 'uniform ut work.', 'A. ; an', 'B. an: a', 'C.an; an D. R.a', '3. Alice likes dogs very much. She'd like to keep a puppy", 'a pct.', 'A. In', 'B. with C. for D.as', '4. I read about', "pen friend's information on", 'way to school.', 'Amy; me', 'B. my: my', 'C me; me', 'D. me. my', '5. I don't understand what you said. Would you please give me', 'example?', 'A. another', 'B. other', 'C. others', 'D. the others', '6. Ben can speak a lot of French, but he knows', 'English.', 'A. few', 'B. little', 'C. a few', 'D. less', '7. When I looked up into the sky, I saw many birds', 'high', 'A. flew', 'B flying', 'C. flics', 'D. to fly', '# Tina', 'the room by yourself', 'A. cleans', 'B. will clean', 'C. cleaned', 'D. cican', '9. Don't forget to give the hungry birds some food", 'A. cating', 'B to cat', 'C. cats', 'D. cat', '10. You don't need to talk to the class teacher. 1", 'him everything already.', 'A. will tell', 'B. am telling', 'C. have told', 'D. mold', 'D. His much', 'to travel around in the countryside B-a before']

'11. It is much', 'to travel around in the countryside than before']

接下来，我们调用“慧语”分别对本试卷中的第 2，3，6，7 几道题进行提问，“慧语”答复的结果整理如下。可以看出几道题中题目整理的完成度都很高，四道题目“慧语”均做出正确的答复。在题目内容讲解上，能够准确的抓住考察的知识点，甚至要比原试卷中参考答案给的讲解更加详细。但是在出新的题目这一任务上，尚显得不够变通，与原题目相似度、重合度过高，经常会出现只是换了一种说法、几个单词，需要进一步完善和提高。

“需要讲解的题目”	“题目整理”	“题目讲解内容”	“新的题目”	参考答案中的讲解
第二题	2. Bob is ___ SPCA officer and he wears ___ uniform at work.A. a; an B. an; a C. an; an D. a; a	【答案】B【详解】句意：鲍勃是一名 SPCA 的官员，他上班时穿着制服。考查冠词。根据单词 SPCA 的读音（以元音开头），可知应该使用 an 而不是 a。而根据 officer 的读音（以辅音开头），应该使用 a 而不是 an。所以正确的选择是 An SPCA officer。故选 B。	new_sample:Who has to wear ___ uniform at work?A. a B. an C. / D. the	【答案】B【详解】句意：鲍勃是爱护动物协会的官员，他在工作时穿着制服。考查 a 和 an 的区别。a 一，用于辅音音素开头的单词前；an 一，用于元音音素开头的单词前。officer 为单数名词，SPCA 为元音音素开头的单词，因此其前应用不定冠词 an 来限定。uniform 为单数名词，且为辅音音素开头的单词，其前应用不定冠词 a
第 3 题	3. Alice likes dogs very much. She'd like to keep a puppy ___ a pet.A. In B. with C. for D.as	【答案】D【详解】句意：爱丽丝非常喜欢狗。她想养一只小狗作为宠物。考查介词。keep 和 pet 之间是逻辑上的关系，表示“把……养作为……”。根据句意，应该使用介词 as，表示“作为”的意思，故选 D	new_sample:What does Alice want to keep ___ a pet?A. In B. with C. for D. as	【答案】D【详解】句意：Alice 非常喜欢小狗。她想要养一只小狗作为宠物。考查介词辨析。to 到；with 和；for 为了；as 作为。根据“keep a puppy”可知是要把小狗当作宠物来养。故选 D。
讲一下第 6 题	6. Ben can speak a lot of French, but he knows ___ English. A. few B. little C. a few D. less	【答案】B【详解】句意：Ben 会说很多法语，但他对英语了解甚少。考查数量词。few 表示否定概念的数量词。few 表示否定概念的“few”或“a few”，表示很少；little 表示否定概念的“little”，表示很少；a few 表示几个；less 表示更少。根据句意可知，Ben 对英语了解很少，所以应用 little，故选 B。	new_sample:Sarah has ___ friends in this city.A. few B. little C. a few D. less	【答案】B【详解】句意：Ben 会说很多法语，但他几乎不懂英语。考查代词辨析。few 几乎没有，修饰可数名词复数；little 几乎没有，修饰不可数名词；a few 几个，修饰可数名词复数；less 更少，修饰不可数名词。English 表示语言“英语”时，是不可数名词，排除 AC；根据“Ben can speak a lot of French, but he knows...English”可知此处无比较之意，用 little。故选 B。

讲一下第7题	7. When I looked up into the sky, I saw many birds _____ high. A. flew B. flying C. flics D. to fly	【答案】B【详解】句意：当我抬头看向天空时，我看到很多鸟在高空飞翔。考查动词的时态和语态。根据句意，当我抬头看向天空时，鸟已经在高空飞翔了，因此应该使用现在分词形式 flying 来修饰 birds，表示一个正在进行的动作。故选 B。	new_sample: 7. John enjoys _____ tennis with his friends every weekend. A. play B. to play C. playing D. played	【答案】B【详解】句意：当我抬头望向天空时，我看到许多鸟飞得很高。考查非谓语动词。see sb. doing sth.看见某人正在做某事；see sb. do sth.看见某人做了某事（强调过程）。根据“When I looked up into the sky”可知是抬头时，看见鸟正在飞，用结构 see sb. doing sth.。故选 B。
--------	---	---	---	--

程序运行的结果的截图如下。



图 12 试题运行样例

3. 项目评价

3.1 项目优点

(1) 能够借助大语言模型这样的新技术结合学习需求进行结合开发英语学习助手，具有一定新意，也与主题契合；

(2) 对 ocr 实现进行多种方式对比分析，与多种大语言模型应用方法进行尝试与方法提升，最终提升了项目鲁棒性；

(3) 在实现项目的基础上进行多个维度的鲁棒性测试，一定程度上证明系统有效性。

3.2 项目不足

(1) 项目方法创新性不足，项目采用了基本方法实现了一个英语学习助手的雏形，但是没有显著的方法创新性，如没有明显证明其个性化学习；

(2) 项目完成度不足，虽然项目借助了 gradio 实现了功能的完整演示，但是目前项目呈现的案例中更多是选择题，没能完全实现项目目标中学习助手的全部功能。

4. 参考学习资料

(1) prompt engineering:

https://www.bilibili.com/video/BV1Z14y1Z7LJ/?vd_source=c413937b5717c829d7a89d1a42b0f462

(2) LangChain 中文网, : <https://www.langchain.asia/>

(3) <https://azure.microsoft.com/zh-cn/>