可控文生图 主要流程

成员概况（python学习情况）：仅python循环/条件

任务

1.看懂ControlNet算法

2.选择一个现成算法(Openpose, 边缘检测（Canny）, Depth等)，为其添加一个控制条件（做一个小插件）

3.把算法集成到SD上

4.创建可视化界面

公共基础（2个月左右）：

1.Python进阶：函数，类，模块导入，异常处理，（装饰器，上下文管理器）

2.Git版本控制（较容易）：clone, pull, commit, push

3.环境与包管理 pip conda venv

4.深度学习基本概念 B站“深度学习入门”

5.SD+ControlNet概念 理解文生图原理，ControlNet控制条件的作用

**各任务要求**

任务1：看懂ControlNet算法

· 前置知识：公共基础 + PyTorch基础 + 扩散模型原理

· 专属技能：

· PyTorch的nn.Module、张量操作

· ControlNet论文精读（零卷积、条件注入）

· 官方代码结构分析

· 难度：（难）

任务2：添加控制条件插件

· 前置知识：任务1完成 + OpenCV图像处理

· 专属技能：

· 图像预处理技术

· 现有控制器代码分析（如Canny）

· 简单控制条件设计（如颜色块、笔画）

· 难度：（中等到难）

任务3：集成到SD系统

· 前置知识：任务2完成 + Diffusers库使用

· 专属技能：

· StableDiffusionControlNetPipeline使用

· 模型加载和配置

· 参数调优

· 难度：（相对简单）

任务4：创建可视化界面

· 前置知识：公共基础 + 基础Web概念

· 专属技能：

· Gradio（或Streamlit）框架

· 图片上传、参数调节组件

· 函数调用封装（会调用别人写好的函数）

· 难度：（最容易上手）

关键实施建议

1. 立即行动：从公共基础开始，团队先集中学习

2. 并行推进：基础学完后，界面开发与算法理解并行

3. 定期同步：每周开短会，展示进度，解决阻塞问题

4. 先易后难：界面同学可以先基于现有ControlNet做demo，不等算法

5. 里程碑检查：

·学完基础：所有人能运行官方ControlNet demo

·界面原型完成，算法同学理解代码结构

·自定义控制器原型，界面与算法初步集成

· 选择最简单的控制器：如颜色块或简单笔画，而不是复杂的人体姿态

· 利用现有工具：Gradio让界面开发变得极其简单

· 先集成再优化：先让系统跑起来，再考虑性能优化

人员分工：

算法A：写插件的主要负责人 深入了解算法

算法B：也要能辅助写插件，中期转向学习如何集成到SD

界面C：负责界面开发，同时主要负责文档整理

时间安排（暂定）

1. 到明年1月初：学习公共基础知识
2. 1月初 到 3月：C学习Gradio，A、B同时学习任务1

目标：C基本能使用Gradio，A、B能基本看懂ControlNet

1. 3月 到 5月：C继续学习Gradio，A开始研究添加新条件，B能将ControlNet接入SD

目标：C能够用别人现成的demo进行可视化展示，A基本了解如何写新条件，B基本了解如何集成

1. 5月 到 7月：A开始写新条件，B研究完集成后辅助写新条件，C完善可视化界面

目标：A输出模型接口，B调用算法接口并集成，C能够对接并展示我们的插件、优化交互

1. 7月 到 10月结题：优化项目

虽然任务挑战性比较大，但是我们拥有长达1年的时间！虽然大家平时都很忙，但是只要愿意抽点时间投入，一定可以圆满完成项目！