

分工 21.05.30

写在前面：关于项目代码实现阶段的注意事项

项目代码实现阶段从暑假开始。

- **Code Review** : 请与合作方保持近乎即时的代码进度沟通，这需要熟练的使用 Github
- **任务周期**: 单个任务周期为三天，一周以上是红线
- **沟通周期**: 每日进行视频+共享屏幕的工作沟通，当天解决问题，解决不了的立马询问老师
- **善于使用工具**: 先查一查网上有没有现成的代码、包等，有没有现成的 Github 项目，如果没有，再着手自己写。例如，如果项目需要用到 Transformer, 千万不要花一周自己把 Transformer 写一遍，因为 Hugging Face 有 Transformer 的包，直接 import 来用！

鲁棒图像滤镜风格迁移组

王泽坤：理论设计(数理、模型、创新)，进度把控(资源、实验、管理)，论文撰写

- 深入学习机器学习、深度学习理论，到达数理底层的理解
- 广泛调研，了解行业前沿技术与进度，探索可用的技术创新方案
- 分析实验结果，提出模型改进方案以及实现方案
- 参与全部代码工作但不主导

侯亚鲲：模型改进(技巧的加入、热点跟踪)，数据集

- 快速学习项目代码，包括但不限于 RepMLP 等，并应用
- **alignment 实现(主)**: 包括但不限于 Morphing Module 的代码实现等
- 数据集整理：自动+手动

协作方：王泽坤提供新学习代码以及 alignment 设计方案，和袁梓钊共同改进图像生成算法(并行)，和李瑞琦共同开发 alignment(串行)

李瑞琦：语义分割与图像处理

- **alignment 实现**: 包括但不限于语义匹配实现、RoI 抽取等
- 判别器设计：即图像分类与识别的实现
- 一些基本的图像处理比如数据增强、预处理、Mask、图像形态学处理等

协作方：王泽坤提供技术方案，和袁梓钊对接判别器工作(并行+串行)，和侯亚鲲共同开发 alignment(串行)，为全组提供图像基本处理服务(串行+并行)

袁梓钊：风格提取与图像生成

- 滤镜抽取算法的实现
- 图像生成算法的实现

- 图像生成算法的改进(主): 包括但不限于 multi-scale 设计、style-based methods 等

协作方: 王泽坤提供技术方案, 和李瑞琦配合完成滤镜抽取(李瑞琦承包基本图像处理, 并行), 和侯亚鲲合作实现生成算法改进(并行)

知识指导的自由问答系统组

王泽坤: 理论设计(数理、模型、创新), 进度把控(资源、实验、管理), 论文撰写

- 深入学习机器学习、深度学习理论, 到达数理底层的理解
- 广泛调研, 了解行业前沿技术与进度, 探索可用的技术创新方案
- 分析实验结果, 提出模型改进方案以及实现方案
- 参与全部代码工作但不主导

孙肇泽: 知识语言模型

- 语言模型实现: 包括但不限于 Seq2Seq 模型、预训练模型等的实现
- 文本生成实现(主): 包括但不限于 SeqGAN 等文本生成模型实现
- 把知识融入到语言模型中(代码实现)

王泽坤提供技术方案, 与张毅博合作实现知识语言模型(并行), 与侯亚鲲合作处理文本(串行)

张毅博: 知识相关, 模型改进

- 知识表示(主): 可融入到语言模型中的知识表示的实现, 将文本型知识转化为知识图谱的代码实现, 将其和语言模型结合起来的代码实现
- 文本匹配的代码实现: 包括但不限于问答对匹配判别、知识文本与问题文本匹配判别
- 快速学习项目代码, 包括但不限于 ERNIE、KEPLER 等, 并应用

王泽坤提供技术方案, 与孙肇泽合作把知识融入到语言模型中, 与侯亚鲲合作寻求基本文本处理服务

侯亚鲲: 数据集, 文本处理

- 知识检索(主): 使用 Elasticsearch 等从互联网中检索知识, 或者其他 Knowledge Acquisition 算法等的实现
- 一些基本的文本处理如分词、mask、检索、NER 等
- 法学数据集查找、下载与处理, 使用 pandas, spacy 等

王泽坤提供资源, 为全组成员提供数据集保障、知识库保障以及基本文本处理服务