# **会议纪要**

## **一、会议基本信息**

* ****会议主题****：图像处理评估体系建立及大模型应用优化讨论
* ****参与人员****：yilongju、Charles、樊亚辉等
* ****会议形式****：在线会议
* 会议时间：2025/07/25 09:00-10:00

## **二、会议主要讨论内容**

### **（一）评估体系建立的必要性**

1. 核心问题：当前缺乏端到端的量化评估框架，无法衡量图像处理（如大模型输出）的具体效果。即使使用算法或大模型处理图像，也无法通过 “输入 100 张图后输出平均分” 等量化指标判断表现。
2. 关键观点：
   * 4000 个数据与 10 个数据的价值差异在于能否量化评估，仅靠人工抽查无法覆盖全部结果，必须通过量化指标实现系统性评价。
   * 评估体系是迭代优化的基础，只有通过量化指标才能排序问题优先级（如优先解决平均分 10-20 分的案例，暂放 80-90 分的案例）。

### **（二）量化评估框架设计**

1. 打分标准：
   * 以 0-100 分为打分范围，定义 “治理后符合标准的图” 为 100 分，“原始未处理图” 为 0 分，中间结果按贴近程度打分。
   * 基于十来条与美观性相关的规则（如建筑物平齐、无乱序元素等），规则可设置不同权重，符合高权重规则则多得分。
2. 实施方式：
   * 对每张处理后的图像，通过人工或自动化工具对照规则打分，最终计算平均分（如 100 张图或 4000 张图的平均分）。
   * 打分结果需记录（包括图像编号、最高分、最低分等），作为迭代优化的依据。

### **（三）数据标注相关讨论**

1. 标注对象：  
   需标注与美观性相关的元素，包括但不限于：
   * 基础地理元素：建筑物、河流、道路；
   * 电网设备：变压器、干塔、分支箱、表箱、电缆连接线等。
2. 标注注意事项：
   * 需考虑设备覆盖、交叉、压叠等标注难点，确保标注与最终美观性评价相关。
   * 标注需服务于量化指标（即对图像处理结果的 0-100 分打分），无量化指标的标注无实际意义。
3. 标注实施：  
   先组织人员人工标注 100 张（或 50 张）图像作为基础数据，确保大模型可识别图像元素，为后续处理奠定基础。

### **（四）大模型应用及流程优化**

1. 大模型处理流程：
   * 单次输入尽量仅 1 张图，通过 prompt 提供 5 个正面案例、5 个反面案例及规则提示，确保输入简洁明确。
   * 编写程序实现自动化处理：将 100 张图放入循环语句，自动输出 100 个处理结果。
2. 结果反馈与迭代：
   * 对比大模型输出结果与人工处理结果，人工打分并记录，重点收集经典错误案例（如低分图的共同问题），反馈给大模型进行训练。
   * 每次迭代聚焦 1-2 类错误，通过修改数据处理步骤、大模型输出处理等方式优化，再通过平均分变化判断改进效果。

### **（五）工作流与自动化要求**

1. 工作流设计：  
   建立 “输入图像→大模型处理→自动化 / 人工打分→记录结果→反馈优化” 的完整流程，通过代码实现打分、数据记录等环节的自动化，减少人工参与。
2. 量化指标延伸：  
   除处理效果打分（0-100 分）外，需记录处理时间，形成 “效果 - 时间” 权衡曲线，为实际应用提供参考。

## **三、下一步行动计划**

| **任务内容** | **具体要求** | **优先级** |
| --- | --- | --- |
| 建立评价体系 | 明确 10 条左右核心规则及权重，制定 0-100 分量化打分标准 | 高 |
| 人工标注 | 组织人员标注 100 张图像（含建筑物、电网设备等关键元素） | 高 |
| 大模型处理与打分 | 运行大模型处理 4000 张图（先试点 100 张），人工对比结果并打分，记录分数及编号 | 高 |
| 自动化流程开发 | 编写代码实现图像自动输入、处理结果自动输出、打分自动化（如循环处理 100 张图） | 中 |
| 结果反馈与迭代 | 收集低分案例的共同错误，反馈给大模型进行训练，优化处理步骤并验证效果 | 中 |

## **四、会议备注**

* 每日 / 每两日需同步工作进展（如代码修改、实验结果等），避免信息空白。
* 所有规则修改、优化方案需记录（如规则一、规则二），确保可追溯性。