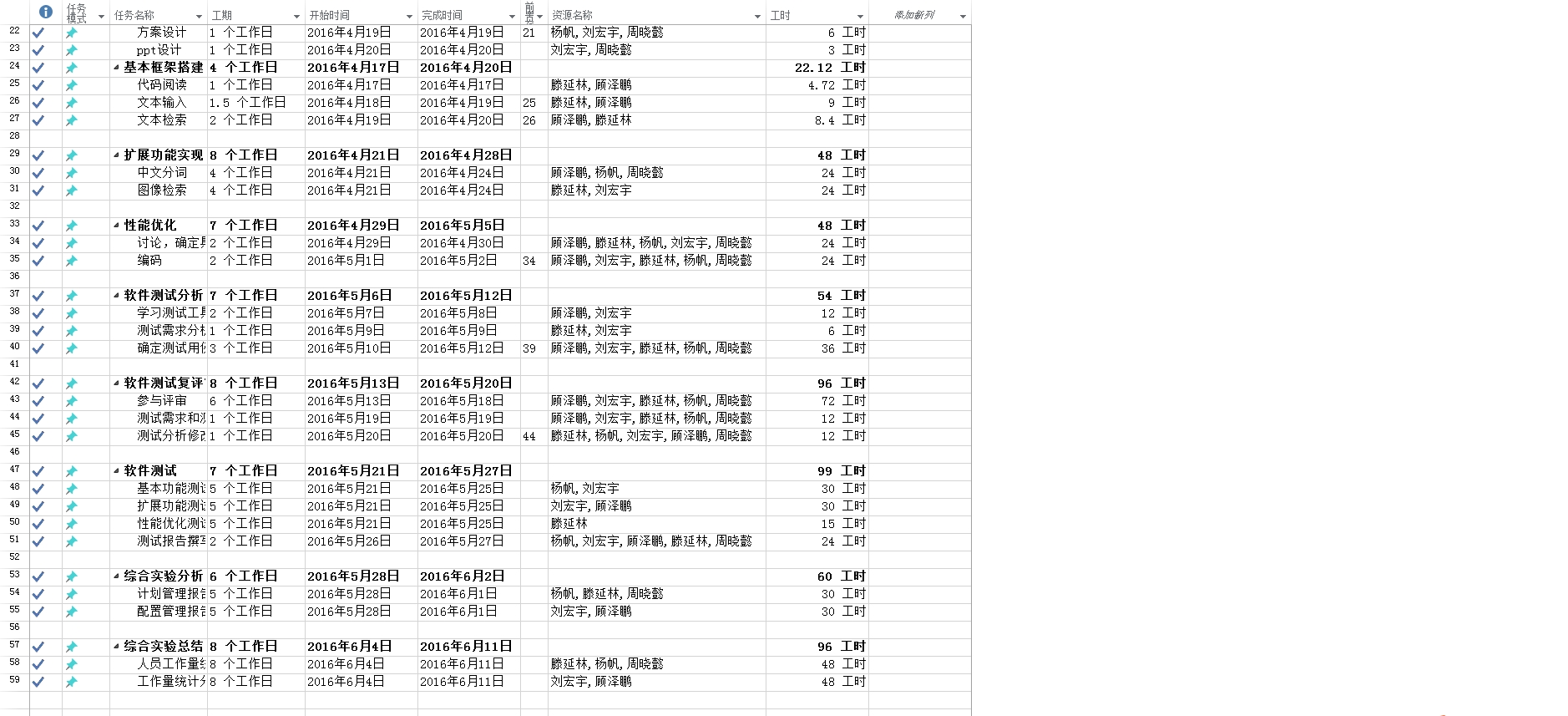
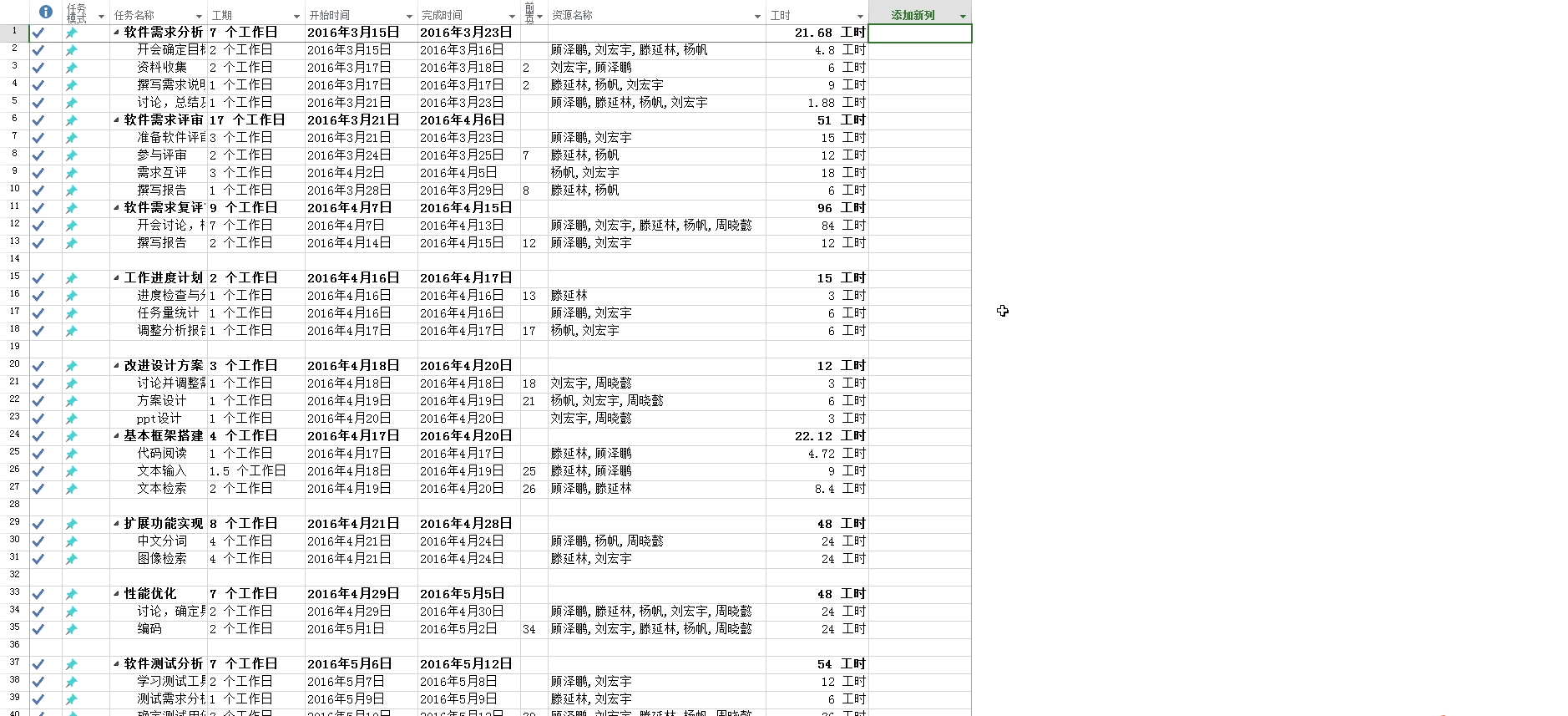
# 人员能力分析

本组人员总体编程能力不错，并且部分成员在课题组的项目中使用过Lucene。大家能力比较接近，在任务分配时，比较灵活，不同成员可以相互替代。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 组内成员 | 专业 | 事先系统了解程度 | 编程能力 | 备注 |
| 刘宏宇 | 软件工程 | B（在工程中使用过Lucene） | B | 组长 |
| 顾泽鹏 | 软件工程 | B（在工程中使用过Lucene） | B |  |
| 滕延林 | 软件工程 | B（在工程中使用过Lucene） | B |  |
| 杨帆 | 软件工程 | C | B |  |
| 周晓懿 | 软件工程 | C | B |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 需求文档变更追踪记录 | | | | |
| 项目名称 | 基于Lucene的分析与应用 | | | |
| 版本号 | 变更项编号 | 内容说明 | 编制人 | 变更量（W-字  P-图） |
| V1.0 | 1 | 最初版 | 顾泽鹏  刘宏宇  滕延林  杨帆 | 3352W |
| V1.01 | 1 | 添加项目整体用例图 | 滕延林 | 2P |
|  | 2 | 根据评审更改内容表述 | 滕延林 | 263W |
| V1.02 | 1 | 更改文档布局，将非功能性需求分离出来作为与系统功能需求对应的独立部分 | 刘宏宇 | 63W |
|  | 2 | 添加了系统运行的数据需求 | 刘宏宇 | 175W |
|  | 3 | 添加了系统的运行需求，包括硬件、软件以及用户界面的需求 | 刘宏宇 | 115W |
| V1.03 | 1 | 调整了实验目的部分的表述 | 顾泽鹏 | 31W |
|  | 2 | 重新确定了用例图与RUCM模型 | 顾泽鹏 | 7P |
|  | 3 | 更正了第三章的业务需求分解的表述 | 顾泽鹏 | 21W |
| V1.04 | 1 | 增加了实现部分的说明：中文分词 | 杨帆  滕延林 | 53W |
|  | 2 | 增加了实现部分的说明：图像特征选择 | 杨帆  滕延林 | 72W |
|  | 3 | 增加了实现部分的说明；图像特征提取 | 杨帆  滕延林 | 126W |
|  | 4 | 增加了实现部分的说明：数据预处理 | 杨帆  滕延林 | 141W |
|  | 5 | 增加了实现部分的说明：距离计算公式的选择 | 杨帆  滕延林 | 165W |
|  | 6 | 增加了实现部分的说明：网站搭建 | 杨帆  滕延林 | 201W |
|  | 7 | 增加了实现部分的说明：Hash排序 | 杨帆  滕延林 | 233W |
| V1.05 | 1 | 在第3章添加对基本需求实现的说明 | 滕延林  杨帆 | 191W |
|  | 2 | 数据需求部分进行修改、添加场景需求说明 | 滕延林  杨帆 | 137W |
|  | 3 | 第6章修改为实现需求和实现方案，并添加部分实现需求方面的内容 | 滕延林  杨帆 | 305W |

# 任务计划分析表



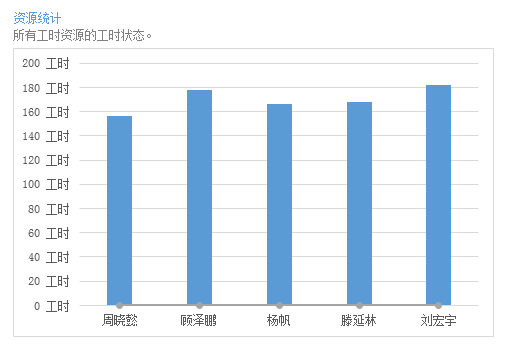


Figure 资源统计表

由图1可见，每个成员的工作量分配比较均匀，均在150-180小时之间，且符合实际工时。

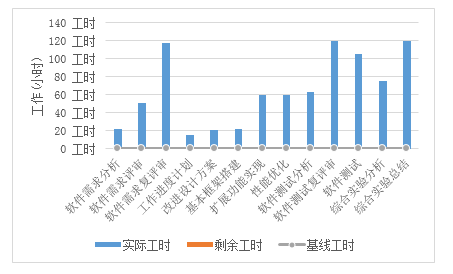


Figure 任务工时图

从图2可见，需求分析、软件测试以及实验分析总结几个实验占比较大。代码方面，前期环境搭建占比较小，主要工作量扩展功能的实现与性能优化方面。

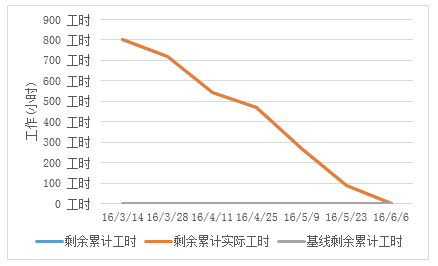


Figure 燃尽图

燃尽图较为平滑，说明项目进行较为流畅。

以下是每个成员的具体工作量说明。



Figure 成员具体工时：刘宏宇



Figure 成员具体工时：滕延林



Figure 成员具体工时：顾泽鹏



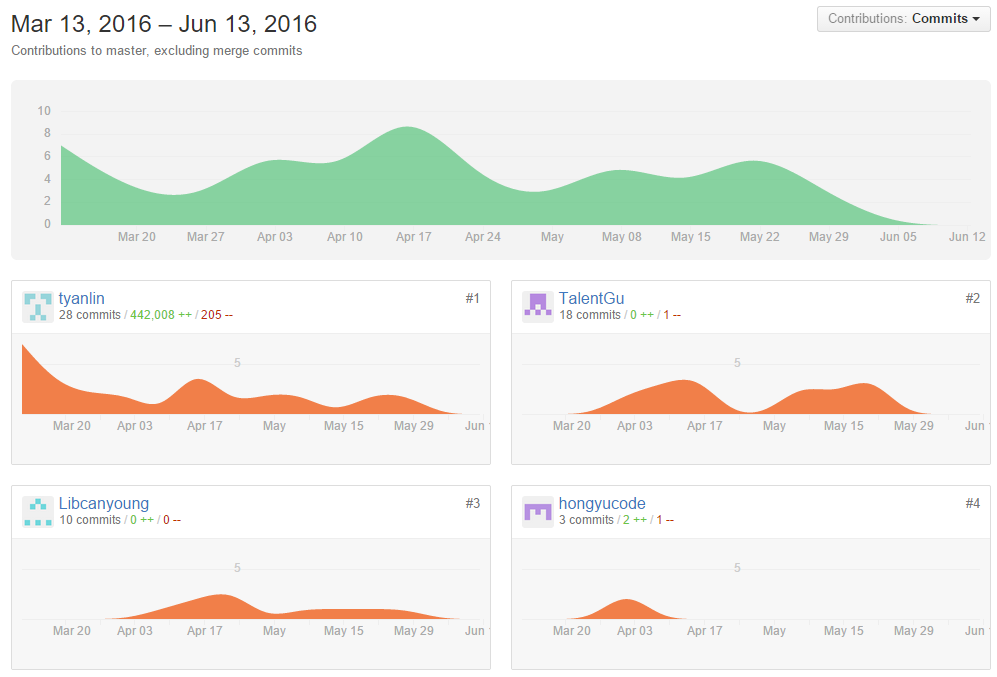
Figure 成员具体工时：周晓懿



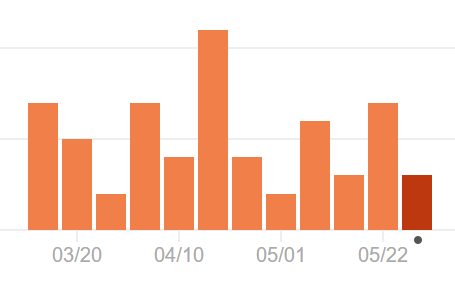
Figure 成员具体工时：杨帆

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 测试文档变更追踪记录 | | | | |
| 项目名称 | 基于Lucene的分析与应用 | | | |
| 版本号 | 变更项编号 | 内容说明 | 编制人 | 变更量（W-字  P-图） |
| V1.0 | 1 | 最初版 | 顾泽鹏  刘宏宇  滕延林  杨帆  周晓懿 | 6723W |
| V1.01 | 1 | 更新文档目录，调整格式。 | 滕延林 | - |
|  | 2 | 修改“性能测试”：针对非功能需求模块的“高效性”需求用例对应的测试用例描述为“性能测试“ | 滕延林 | 72W |
|  | 3 | 在系统部署测试测试用例中测试系统是否正确部署的步骤与部署的步骤混淆了，配置Web应用、导入Lucene包应该属于部署的步骤。 | 刘宏宇 | 107W |
|  | 4 | 测试需求文档中文件添加测试的描述与文本库的管理中的添加文件的描述、以及需求文档中RUCM模型中添加文件的描述不一致 | 顾泽鹏 | 151W |
|  | 5 | 测试员具体分配给个人 | 杨帆 | 21W |
|  | 6 | 添加与其他特征的比较 | 顾泽鹏 | 354W |
|  | 7 | 修改网站搭建测试用例 | 周晓懿 | 43W |
|  | 8 | 重新整理了第一张编码 | 杨帆 | - |

# 提交次数



# 各实验阶段提交项目量统计



# 影响因素分析

**外在因素**

1. 评审意见：包括老师和同学在格式和内容上提出的意见

评审意见贯穿了项目进行的始终，老师和同学们的评审意见开阔了我们的视野，发现了许多我们小组内部讨论时没有想到的地方，帮助我们不断完善我们的项目需求、计划并确保我们在项目进行过程中没有偏离方向。

**内在因素**

1. 对系统的了解程度：

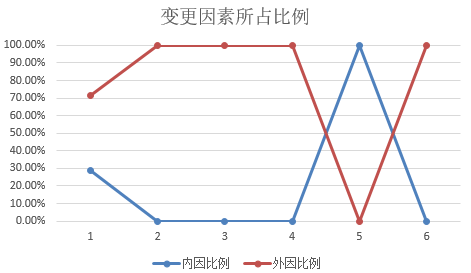
在项目刚开始的时候，对项目的熟悉程度较低，只是大致了解Lucene开源系统是做什么的，有哪些模块，根据这些模块倒推项目的需求分析。

随着对系统了解程度的加深，为了做出实质性的改进，我们对系统具体的使用场景以及数据样例的特点进行了分析，根据不同的使用场景以及数据样例的特点，有针对性地对需求进行了改进，进而变更了相应的需求规格说明书以及任务计划。

**统计分析**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 版本变更记录 | 1 | 2 | 3 | 4 | 5 | 6 |
| 内 因 | 2 | 0 | 0 | 0 | 7 | 0 |
| 外 因 | 5 | 3 | 6 | 2 | 0 | 3 |
| 总 次 数 | 7 | 3 | 6 | 2 | 7 | 3 |
| 内 因 比 例 | 28.57% | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% |
| 外 因 比 例 | 71.43% | 100.00% | 100.00% | 100.00% | 0.00% | 100.00% |

**相关性分析**



为了研究版本变更与影响因素之间的相关性，我们通过分析需求文档的版本变更追踪链进行相关性分析，之所以没有考虑测试规格说明书等是因为这些文件相对变化较少，追踪链较短，无法总结出有效规律。通过变更追踪链分析，我们发现：

1、第一至第四版本变更时，外因占据了主要地位，这是因为这一阶段主要是对需求规格说明书的格式、内容表达及阐述的准确性及合理性、用例图、RUCM等的格式及用法是否恰当进行的讨论及修改，因此是外因占了主导地位；

2、第五次版本变更时，由于涉及到了具体实现过程，项目的具体计划发生了改变，这一阶段内容主要是受到内因及项目组成员对系统的把握程度及其相应计划的影响，因此内因是主导地位；

3、第六次版本变更，是在项目实现工作已经完成的情况下，针对完成内容按照老师给出的指导意见进行规格说明书的增补，如增加“基本需求、实现需求”，将实现内容落实到需求文档中的需求，因此是外因占主导地位；

# 总结

从各个阶段的计划以及实际进度来看，本组的实际进度与计划相符合程度高，在相应的时间节点都有相应的产出。这得益于以下几点：

1. 每周都有频繁的交流，任务分配明确。
2. 课题记录清晰、明确，特别对于时间节点等重要信息记录详实，使得每名成员都会提前计划好自己的任务。

对于每名成员的能力有比较清晰的了解，让每名成员做最擅长的事，达到事半功倍的效果。

# 有效方法总结

## 软件需求分析

遇到的问题:对项目的了解不够深刻，对项目的难度和复杂度的定位不是很明确。

解决方案：为了加深对项目的理解，增加对项目的熟悉度，小组成员先利用网上的技术文档，对项目进行初步理解。项目成员滕延林、刘宏宇、顾泽鹏对该项目有一些初步的了解，曾经利用lucence进行开发，而杨帆、周晓懿对lucence的了解比较少。因此，在项目开始之前，先通过网络对lucence进行初步的学习。并且，每人重点关注一个方向，其中，刘宏宇负责Lucene 索引机制部分，滕延林负责Lucene 搜索机制部分，顾泽鹏负责Lucene 段合并(merge) 机制部分，杨帆负责Lucene 的分词器Analyzer部分。

结果：经过初步的学习，小组成员对项目的了解加深了不少，对项目的难度和复杂度的定位有了初步的认识。

## 软件需求评审

遇到的问题 本小组成员对其他小组的项目了解比较少，因此，在评审时，难以提出有价值的问题。

解决方案：为了解决这个问题，小组成员进行讨论。最后决定首先对其他小组的项目进行学习，特别是我们组之前的两组，即F组和G组。其中，滕延林和顾泽鹏负责学习F组spring项目关于控制反转部分的相关知识，而刘宏宇、杨帆和周晓懿负责学习关乎spark的项目知识。并且，在课堂上，当这两个小组进行讲解时，认真进行学习。

结果：经过努力，我们组较好地完成了需求的评审，提出的问题，得到了被评审组的认可。

## 改进与展示

遇到的问题: 在给lucence添加图像检索功能时，选用图像的哪种特征效果最好。

解决方案：为了解决这个问题，我们组进行了大量的测试。通过对各个类型的图像进行提取多种特征。由于无法进行机器识别，只能采用人工识别与判断。小组5名成员，都进行了选择和测试。对领导人、自然风景、条幅、天安门、水、火等多个数据集进行了测试与分析。

结果：通过对各个数据集的图像的准确率进行统计分析，同时考虑特征的维数以及检索效率的问题，最终我们选择了cedd特征。

## 测试需求分析

遇到的问题：如何对测试用例进行确定的问题。

解决方案：首先对软件需求说明书进行查看，根据软件需求说明书的用例确定测试需求的用例。并根据已经实现的项目，进行测试用例的撰写。其中，杨帆负责基础部分测试用例的编写，滕延林负责实现部分测试用例的编写，顾泽鹏负责非功能需求部分测试用例的编写，刘宏宇负责汇总。

结果：较好地完成了测试需求分析实验。

## 测试需求评审

遇到的问题：由于评审组对我们的项目不是很了解，提出的个别问题与我们的项目不相符

解决方案：从评审组的角度考虑问题，与评审组进行交流，明确评审组提出的问题的具体含义，与评审组共同解决这个问题。

结果：与评审组同学达成一致，对部分进行修改，成功解决这个问题。

## 进度计划与控制

遇到的问题：由于对Microsoft project软件不是很熟悉，导致最开始的时候，使用并不是很顺利。

解决方案：专门制定杨帆对该软件进行深入的学习，并向有经验的同学进行学习。

结果：增加了对Microsoft project软件的认识，完成了该实验的要求。

## 配置管理

遇到的问题：对github软件的使用不是很熟悉，需要同时保存多个版本的文档。

解决方案：文档部分，使用冗余存储的方式，对每个版本的文档都进行了保存。代码部分，只对最新版本的代码进行保存。同时，顾泽鹏对项目的文件夹进行了设计和整理，以方便文件的查找。

结果：在版本管理中没有出现太大的问题，较好地完成了该实验。

## 工作量估计与统计分析

遇到的问题：如何对每个成员的工作量进行准确评估的问题。

解决方案：由于不同的工作难度，性质的不同，难以对每个成员的工作量进行评估，因此，拟采用时间对每个成员的工作量进行统计。每个成员的工作时间由成员自己提交、然后小组成员共同进行审核。

结果：能够比较准确地评估出每个成员的工作量。