**项目创意书**

|  |  |  |  |
| --- | --- | --- | --- |
| 项目基本简介 | | | |
| 设计理念及创新点：基于paddleocr开发的文字识别软件随着文字识别框架的简化提高了场景文本识别的速度与准确度，可**在最快的时间内帮助我们把识别**图像中的文字转换成文本格式的软件，帮助我们“准确，快速，便捷，”地获得所需文字信息。 | | | |
| 项目参赛工具 | | 项目应用场景 | |
| 已选工具（必选） | 飞桨深度学习平台 | 技术方向 | 图像检测，文字识别 |
| 已选工具（可选） | EasyEdge端计算模型生成平台 | 应用行业 | 生活，工业 |
| 已选硬件（可选） | EdgeBorad终端端计算加速套件 | 数据来源 | 飞桨深度学习平台 |
| 其他 |  | 核心突破点 | 图像文字识别准确度与速度 |
| 项目研发来源 | | 预期效果 | 提供多场景、多语种、高精度的OCR文字识别服务。支持上传文件夹，或拖拽上传一张或多张图片文件进行批量识别，且单张识别速度在3秒以内，OCR识别准确率平均高达90%以上，支持10+语种识别，包括英语、韩语、日语、法语、西班牙语等，支持旋转图像、裁剪、倾斜校正等功能，让文字识别更快速准确，针对图片模糊、倾斜、反光、叠字等复杂场景可自动进行优化，上传的识别文件，服务器用后即删，满足用户数据安全的需求。 |
| 研发来源 | * 国家/省部级/地方科研项目 |
| * 企业/其它横向合作项目 |
| ☑独立研发 |
| * 以上均不是（请描述） |
| 项目其他合作机构 |  | | |

**（一）选题定位**

**1、作品的创意与独创性：**

随着我国信息技术领域技术的突破以及资本市场的发展，OCR文字识别技术已经应用在多个领域并产生了不小的价值。市场上也出现了较为成熟的文字识别软件，但在实际应对复杂图文场景中，图像检测不完整，文字识别准确度不高等问题层出不穷。

本项目为基于百度飞桨开发的文字识别软件，提供多场景、多语种、高精度的OCR文字识别服务，针对图片模糊、倾斜、反光、叠字等复杂场景可自动进行优化，使文字识别的准确度大大提高，

1. **作品的落地转化可行性：**

**组织可行性**：

项目负责人以及各位成员对项目开发有着认真谨慎的态度，有相关技术开发经验，项目成员分工明确，项目开发进程秩序明确；学校相关教授定期检查项目进度并给予思路引导以及技术支持。

**技术可行性：**

ocr文字识别技术成熟；使用百度飞桨深度学习框架完成模型构建节省编写大量底层代码的精力：省去了部署和适配环境的烦恼；网络相关技术资源丰富，与相关专业人士交流密切。团队成员积极研究，对复杂图像场景的图像检测与文字识别的准确度，完整度取得巨大进步。

**（二）社会价值**

**1、用户需求贴合度：**

该文字识别软件贴合生活，服务大众。主要针对用户在应对复杂图文场景的文字识别给予帮助，在生活中常见的应用有：1仪器仪表上的数据读取，如居民生活中得到水表、气表、电表的数值读取；2、产品的日期、批号读取，获得产品的过期信息及可追溯信息；3、智能交通的车牌识别，获取违章车辆的信息。

**2、作品所带来的效率提升的明确表现：**

在一些简单环境下OCR的准确度已经比较高了（比如电子文档），但是在一些复杂环境下的字符识别，往往显得力不从心。团队致力于研究场景文本识别（文字检测+文字识别）的速度与准确性，开发了基于百度飞桨的文字识别软件。

**在日常生活工作中，我们难免会遇到一些问题，比如自己辛辛苦苦写完的资料，好不容易打印出来却发现源文件丢了；收集了一些名片，却要一个一个地录入信息，很麻烦；快递公司的业务越来越好，但每天需要花费很多时间登记录入运单，效率非常的低。利用该软件可以准确地对复杂繁琐的图像文字进行检测，高效地进行识别并予以导出，以便后续对文字的整理。满足了用户急切获取具体文本信息的需求，提高了用户的生活效率。**

**3、作品的市场价值及推广性：**

目前国内文本识别的市场较大，就目前从行业需求来看，金融、保险、税务、工商、电子商务等行业对信息识别的需求已经越来越广泛，促进了识别技术的大规模应用。而个人消费者对资料电子化、手写识别技术等各方面需求则拓展了OCR识别技术在这一领域的应用，另一方面，网络时代的高速发展使个人资料电子化、商务办公自动化等需求也变得越来越多。

从调查来看，文本识别市场的的确确存在，而且随着信息化时代不断发展，这种需求会变得越来越大，而本项目可以较好的满足市场需求，提供优良的服务，该项目也会随市场发展对于文本识别的需求变化，不断更新迭代以适应市场，因此本项目有较好的市场价值与推广性。

**（三）技术方案**

该文字识别软件的AI技术方向主要为为计算机视觉CV（文本检测）以及自然语言处理NLP（文本识别）。项目主要技术平台是飞桨paddlepaddle深度学习平台，项目的开发过程大致可以分为两部分：算法，软件

**1、算法：**

**Ocr算法**

算法主要使用PaddleOCR的develop分支，整个文字识别过程包括。1.文字检测，获取位置坐标；2. 把检测的文字抠出来送入文字识别模型。

针对文字检测：目前文字检测主要有基于回归的方式和基于分割的方式，该项目使用的是DBNet的文字检测算法。

**针对文字识别**：使用的是使用广泛的CRNN

**文字识别模型**：在PaddleOCR工程中提供了各种预训练模型，根据里面的使用方法很容易搭建一个文字识别系统。针对个人场景使用预训练模型在自己的数据集上进行微调。该项目检测使用的是backbone为resnet18的DBNet，识别采用的是backbone为mobilenet的CRNN。对于检测和识别模型，团队针对项目的使用场景进行了微调。

**Ocr算法服务**：

在文字识别中，考虑到软件支持多种语言的识别。在具体实现的时候，选择共用一个文字检测器，接多个不同语言的文字识别模型，目前识别语言支持：汉语、日语、韩语等。由于模型比较多，在部署的时候并没有使用飞浆自带的部署工具，而是选择在github上开源flask项目来部署OCR服务。

**软件：**

操作界面利用PyQt5进行编写，利用Qt Designer新建一个界面，之后创建主窗口以及使用各种控件画好所需界面，将其保存ui文件转化为py文件获取界面的python代码，再进行代码测试的编写以及逻辑功能模块的实现

**数据获取及处理、任务开发流程等策略规划：**

项目开发过程所用模型均为百度飞桨开源模型，利用github以及paddleocr等开源资源进行数据的获取应用。对用任务的开发流程首先要确定软件系统特点与适用场景，并认真调查用户需求；之后是利用paddleocr进行ocr算法的编写，数据库的设计以及模型的训练；利用github开源flask项目进行ocr服务的部署；利用PyQt5进行操作界面的编写；之后负责软件的细节完善与品质保障

1. **排期规划**

**作品的设计开发进度规划：**

首先要确定软件系统特点与适用场景，并认真调查用户需求；之后是利用paddleocr进行ocr算法的编写，数据库的设计以及模型的训练；利用github开源flask项目进行ocr服务的部署；利用PyQt5进行操作界面的编写；在软件的细节完善后，进行软件的测试与修改迭代。