

桂林电子科技大学 信息与通信学院

校级答辩

1800800222 刘舒豪



毕设题目： 基于深度学习的手写报告智能评分系统

主要功能：

通过网页客户端将手写报告拍摄图片进行上传

通过服务器实现图片的接收和保存

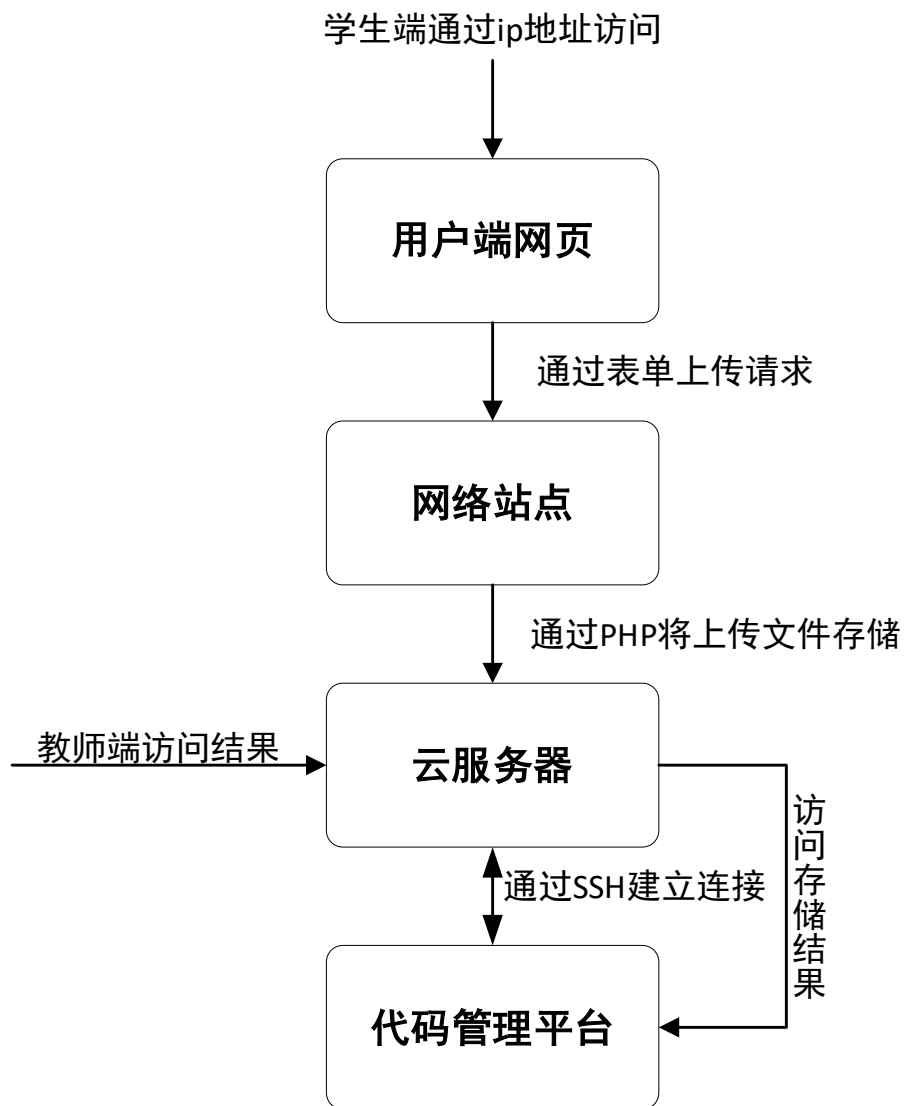
通过代码管理平台程序实现

(1)将手写报告图片进行版面检测，分离图片，表格，文字内容

(2)将手写文字内容进行检测并识别

(3)将识别到的文字与标准答案对比，在语法判断的基础上对其语义进行评分。

系统流程



手写报告智能评分系统

1800800222刘舒豪

WEB

网页前端部分介绍

基于Dreamweaver的网页前端设计

OCR

OCR部分介绍

基于Paddle的手写文本图像识别

NLP

NLP部分介绍

基于BERT的自然语言处理

本网页由1800800222刘舒豪制作

网页首页



图片上传界面

代码运行模块

上传需要评分的报告文件
上传图片: [选择文件] baogao666.png 提交报告
插入报告图片预览:

| 桂林电子科技大学 实验报告 | |
|------------------|----------------------------------|
| 学号 | 1800800222 |
| 姓名 | 刘舒豪 |
| 实验目的 | 学习单片机使用 掌握芯片的输入输出方法 如地址口用途 |
| 实验内容 | 掌握软件操作方法 测试记录程序运行结果 |
| 分析总结 | 懂得软件使用软件 |

图片上传预览效果

网页设计与编辑

通过Dreamweaver软件设计并制作了网页版面

图片上传与预览技术方法

上传：基于html语言实现表单提交图片进行上传

预览：基于javascript模块实现抓取上传至表单的文件路径，实现预览

代码运行模块

form + 要评分的报告文件

上传图片: 未选择任何文件

插入报告图片预览:

制作单

版面: 刘舒豪

设计: 刘舒豪

图片: 刘舒豪

```
<div class="entry-title">
  <h2>代码运行模块</h2>
</div>
</div>
<div class="art-content">
  <div>上传需要评分的报告文件</div>
  <form action="final.php" name="checklist" method="post" enctype="multipart/form-data">
    上传图片:  <input type="file" name="file" style="width: 200px; height: 20px;"
    onchange="PreviewImage(this)" id="file" />
    <input type="submit" value="提交报告"/>
  </form>
```

458722上传文件名: baogao666.png
文件类型: image/png
文件大小: 447.970703125 kB
文件临时存储的位置: /tmp/phpgdFzH2
文件存储在: filebox/baogao666.png

文件提交响应效果

通过PHP模块对接后台，进行接收保存与响应处理

←

根目录 > www > wwwroot > fileget.com > filebox >

↻

上传

远程下载

新建 ▾

文件查找

收藏夹 ▾

分享列表

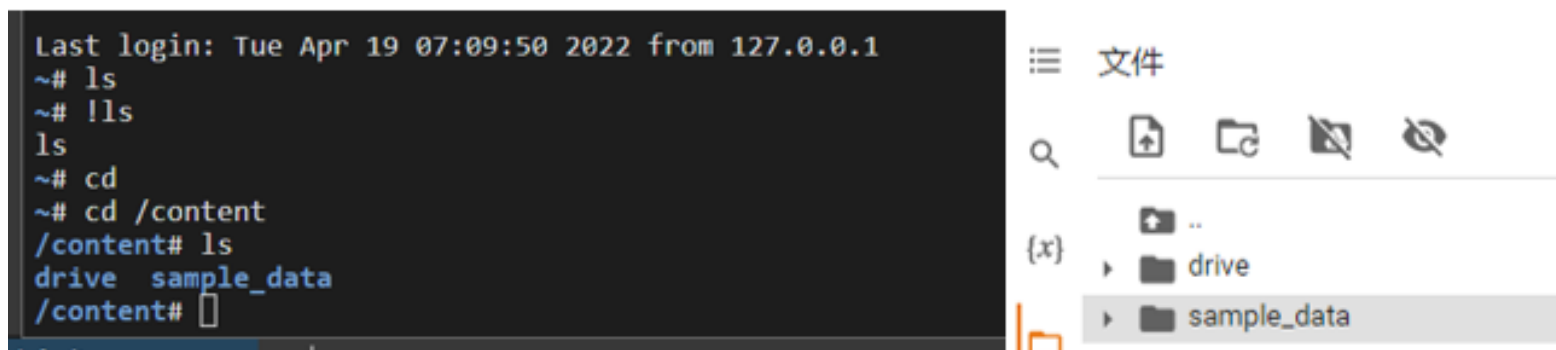
 终端

 /(根目录) (32G)

| <input type="checkbox"/> | 文件名 | 权限 / 所有者 | 大小 | 修改时间 | 备注 |
|--------------------------|--|-----------|-----------|---------------------|----|
| <input type="checkbox"/> | <div> baogao666.png</div> | 644 / www | 447.97 KB | 2022/05/13 12:14:03 | |

网站后台文件列表显示效果

通过SSH连接的方式建立服务器与代码管理平台的连接



建立连接后预览效果图

| | |
|-------------|--|
| 桂林电子科技大学 | |
| 实验报告 | |
| 学号 | |
| 1800800222 | |
| 姓名 | |
| 刘舒豪 | |
| 实验目的 | |
| 学习单片机使用 | |
| 学会芯片的编入程序方法 | |
| 知晓端口用途 | |
| | |
| | |
| | |
| 实验内容 | |
| 掌握软件操作方法 | |
| 测试记录程序运行结果 | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| 分析总结 | |
| 懂得能如何使用软件 | |
| | |
| | |

- 1: 桂林电子科技大学 1.000
- 2: 实验报告 1.000
- 3: 学号 1.000
- 4: 1800 800222 0.967
- 5: 姓名 1.000
- 6: 刘舒豪 0.988
- 7: 实验目的 1.000
- 8: 学习单片机使用 0.974
- 9: 学会芯片的编入程序方法 0.893
- 10: 知晓端口用途 0.929
- 11: 实验内容 0.999
- 12: 掌握软件操作方法 0.865
- 13: 测试记录程序运行结果 0.889
- 14: 分析总结 0.999
- 15: 懂得能如何使用软件 0.888

文字检测与识别生成结果可视化

手写报告评分

检测识别结果文件

OCRresult.txt ×

```
1 桂林电子科技大学实验报告
2 学号
3 1800800222
4 姓名
5 刘舒豪
6 实验目的
7 掌握单片机使用。
8 实验内容
9 训练模型
10 分析总结
11
```

标准答案相同时评分结果

```
INFO:tensorflow:Loading model.
INFO:tensorflow:BLEURT initialized.
[0.9597343802452087,
 0.8378193974494934,
 0.9673959612846375,
 0.8770206570625305,
 0.9382616877555847,
 0.952293872833252,
 0.9963065981864929,
 0.968858003616333,
 0.9563018083572388,
 0.9370981454849243]
```

当标准答案与检测识别结果文件存在差异时评分结果

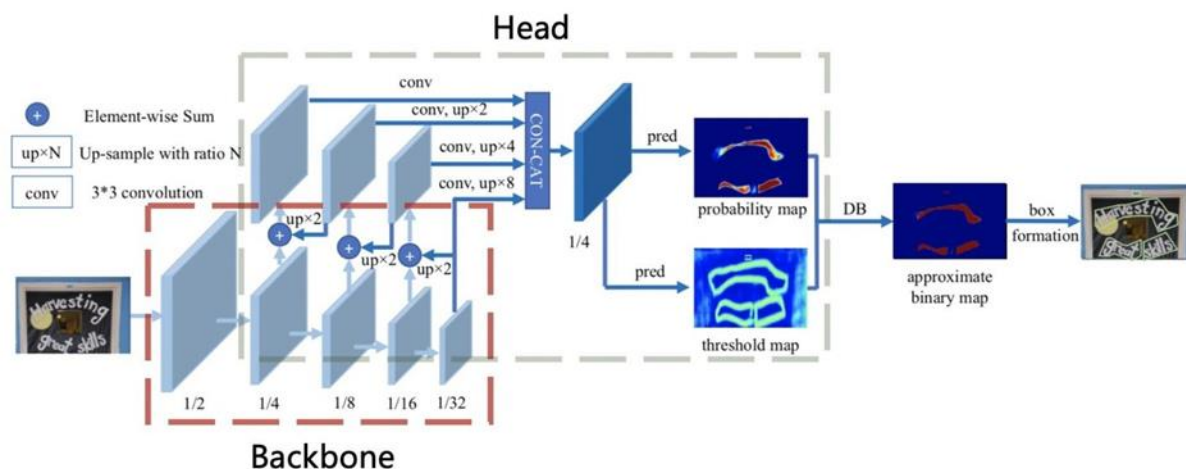
Answerresult.txt ×

```
1 桂林电子科技大学实验报告
2 学号
3 1800800222
4 姓名
5 刘舒豪
6 实验目的
7 掌握单片机使用，学会使用电源模块的手段。
8 实验内容
9 训练模型
10 分析总结
11
```

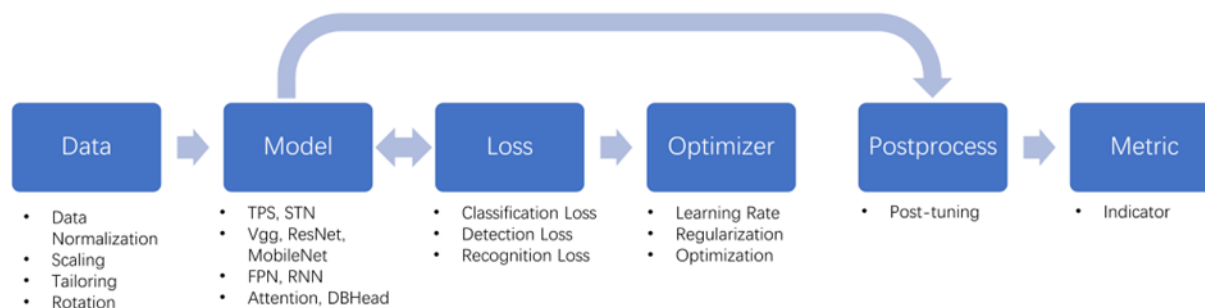
```
INFO:tensorflow:BLEURT initialized.
[0.9597343802452087,
 0.8378193974494934,
 0.9673959612846375,
 0.8770206570625305,
 0.9382616877555847,
 0.952293872833252,
 0.49583059549331665,
 0.968858003616333,
 0.9563018083572388,
 0.9370981454849243]
```

文本检测与识别

基于Paddle开发者工具实现训练网络基本架构



通过调试训练参数，将以CRNN为基础的预训练模型在公共手写文字数据集上按照下图所示的训练思路进行机器学习训练。



数据标注方法

通过GTLC工具能够实现对个人数据集进行标注整理

Off_Line Consecutive Words Annotation System--PAL Group of CASIA

文件(F) 编辑(E) 查看(V) 帮助(H) Text extraction Textline segmentation Character alignment

📄 🖨️ 🔍 🗑️ 🔄 ? + -

S 植在我脑袋最深处可能是回味着刚刚的邂逅也可能是同伴讨论太过开
心，忘记了要将车速放慢我的左后轮爆了在这样鸟不拉屎，鸡不生蛋
S 乌龟不上岸的地方我去哪找补胎的地方啊还好机伶地我在车翻找早
S 被遗忘了的备胎但是我租的这台车太久乏人问津备胎不知道是氧
S 化还是虫BIEK，送手摔车场面，唯一不一样的是他们身上是防摔
S 的皮衣，而我身上是人皮，就这样我进了医院打了全身麻醉，痛

Page annotation information

| Number | Annotation Information |
|--------|---------------------------------------|
| 1 | 文件: 002.bmp 标记行数: 10, 实际行数: 10, |
| State | |

植 在 我 脑 袋 最 深 处 可 能 是 回 味 着 刚 刚 的 邂 逅 也 可 能 是 同 伴 讨 论 太 过 开

0 植在我脑袋最深处可能是回味着刚刚的邂逅也可能是同伴讨论太过开

心 忘 记 了 要 将 车 速 放 慢 我 的 左 后 轮 爆 了 在 这 样 鸟 不 拉 屎 鸡 不 生 蛋

1 心 忘 记 了 要 将 车 速 放 慢 我 的 左 后 轮 爆 了 在 这 样 鸟 不 拉 屎 鸡 不 生 蛋

乌 龟 不 上 岸 的 地 方 我 去 哪 找 补 胎 的 地 方 啊 还 好 机 伶 地 的 我 在 车 翻 找 早

2 乌龟不上岸的地方我去哪找补胎的地方啊还好机伶地我在车翻找早

被 遗 忘 了 的 备 胎 但 是 我 租 的 这 台 车 太 久 乏 人 问 津 备 胎 不 知 道 是 氧

3 被遗忘了的备胎但是我租的这台车太久乏人问津备胎不知道是氧

调参过程

为了达到更好的普适效果，用中科院手写数据集（公共数据集）进行训练和评估，以提升模型的可靠性和有效性。

初始调参后训练epoch: [1/500], iter: 10, lr: 0.000001, loss: 657.583496, acc: 0.000000

调参后训练epoch: [1/4], iter: 10, lr: 0.000001, loss: 667.773621, acc: 0.000000

引入预训练模型

进行调参测试 epoch: [1/100], iter: 10, lr: 0.000001, loss: 34.053398, acc: 0.000000

多次进行训练: epoch: [1/100], iter: 10, lr: 0.000001, loss: 34.053398, acc: 0.000000

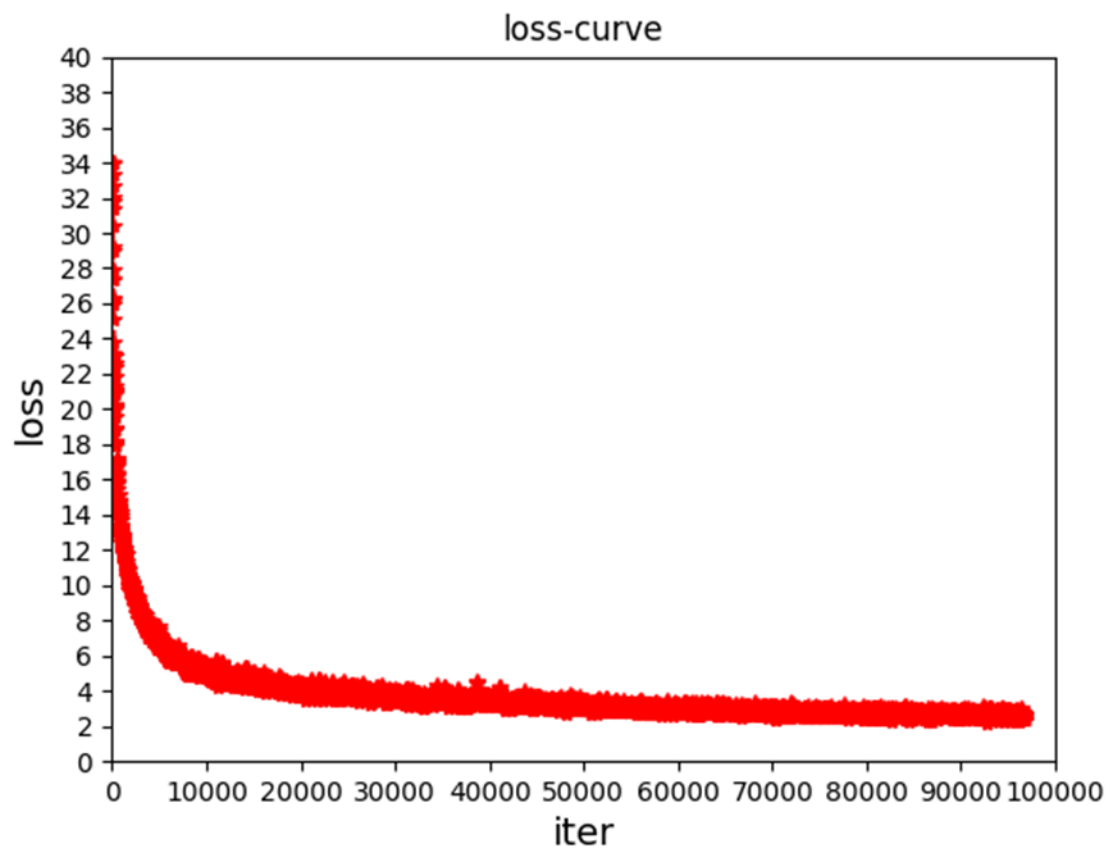
至27次效果 epoch: [27/100], iter: 24650, lr: 0.000892, loss: 4.201280, acc: 0.548826

发现训练迭代和评估周期过于密集，调整评估和iter包数，并测试至30次，发现训练运行效果正常，模型发展效果良好。

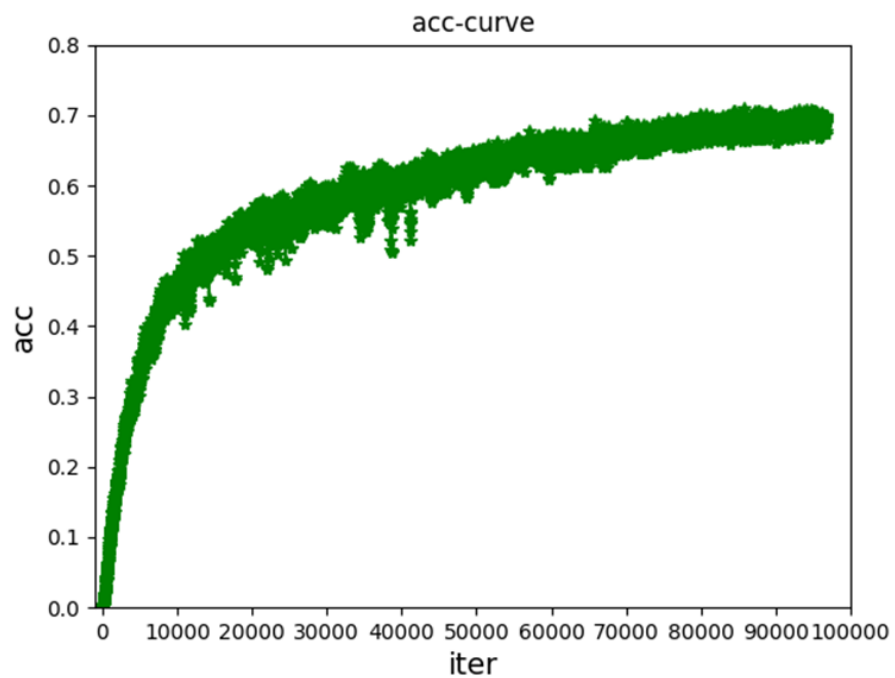
epoch: [30/100], iter: 27330, lr: 0.000862, loss: 4.071111 acc: 0.578123

参照上次参数运行至完成整个训练

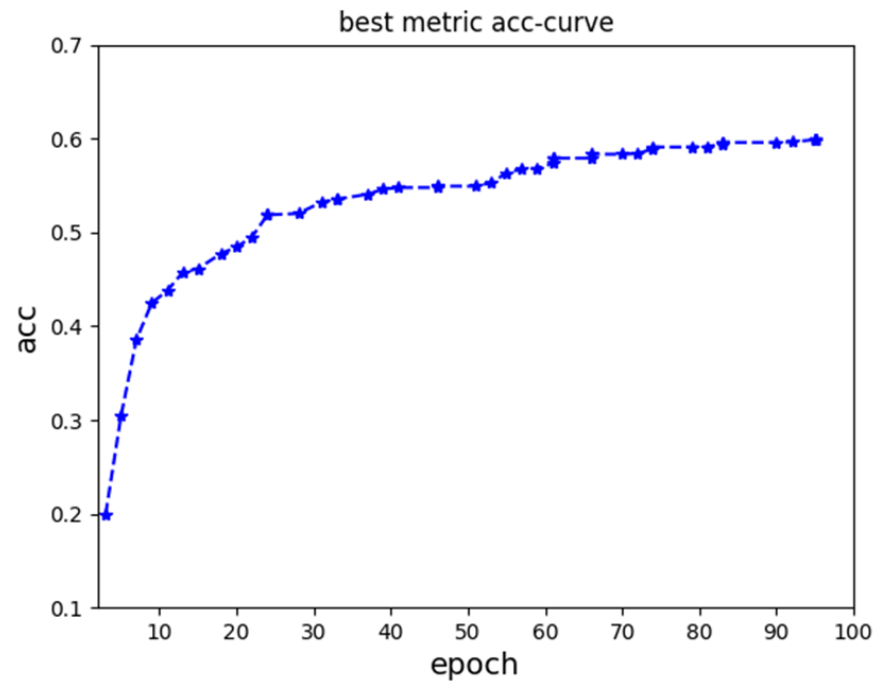
epoch: [100/100], iter: 7487, lr: 0.000013, loss: 2.631539, acc: 0.699216



训练损失随迭代次数变化曲线图



模型准确率随迭代次数变化曲线图



训练最佳模型准确率图表

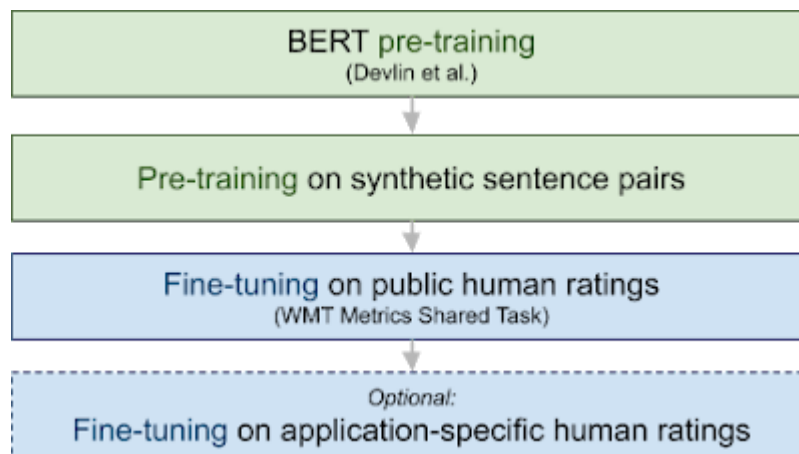
训练前后单行识别效果对比如下

| 模型列表 | 识别文本结果 | |
|-----------|---|-----------------|
| 初始轻量模型 |  | 1: 心有的 0.760 |
| 预训练中文识别模型 |  | 1: 心1中有山海 0.685 |
| 训练后中文识别模型 |  | 1: 心中有山海 0.998 |

整页文本报告识别对比结果

| 识别图片 | 模型 | 识别结果及参考准确率 |
|------|-----------|---|
| | 初始轻量模型 | 1. 桂林电子科技大学 (0.848) 2. 实验报告 (0.951) 3. 学号 (0.837) 4. 1800800 工 2 (0.730) 5. 姓名 (0.893) 6. 实验目的 (0.887) 7. 掌握单片机起德团 (0.612) 8. 实验内容 (0.933) 9. 分析总结 (0.996) |
| | 预训练中识别模型 | 1. 桂林电子科技大学 (0.999) 2. 实验报告 (0.999) 3. 学号 (0.996) 4. 1800800222 (0.976) 5. 姓名 (0.819) 6. 刘子豪 (0.942) 7. 实验目的 (0.998) 8. 掌握单片机使用 (0.718) 9. 实验内容 (0.997) 10. 训练模型 (0.872) 11. 分析总结 (0.918) |
| | 训练后中文识别模型 | 1. 桂林电子科技大学 (1.000) 2. 实验报告 (1.000) 3. 学号 (1.00) 4. 1800800222 (0.988) 5. 姓名 (1.000) 6. 刘舒豪 (0.889) 7. 实验目的 (1.000) 8. 掌握单片机使用 (0.919) 9. 实验内容 (1.00) 10. 训练模型 (0.997) 11. 分析总结 (1.000) |

自然语言处理



基于BERT语言模型利用上述框架发展的BLEURT文本评估模型进行评分

本次所应用的BLEURT评估方法是一种新颖的、基于机器学习的自动度量，可以捕获句子之间的重要语义相似性。它在公共评级集合（WMT指标共享任务数据集）以及用户提供的其他评级上进行训练。Bleurt模型是一个自然语言综合组织的评估模型。他通过导入一组对应的语句，一个对照文件和一个基准文件，他返回一个分数来显示出引用文件的文本流畅和正确水平，同时表示出这个引用文件表达出的基准文件的内容意义的程度。

自然语言评估实现效果

当参考文本为 “掌握使用电源模块的方法”

| 序号↵ | 评估文本↵ | 评估结果↵ |
|-----|--------------|--|
| 1↵ | 掌握使用电源模块的方法↵ | INFO:tensorflow:BLEURT initialized. [1.0012092590332031]↵ |
| 2↵ | 掌握使用单片机的方法↵ | INFO:tensorflow:BLEURT initialized. [0.5198370814323425]↵ |
| 3↵ | 复习模电知识↵ | INFO:tensorflow:BLEURT initialized. [0.3324030041694641]↵ |
| 4↵ | 学会使用电源模块的手段↵ | INFO:tensorflow:BLEURT initialized. [0.7354535460472107]↵ |

当参考文本为 “掌握使用电源模块的方法，并能自主进行通电测试。”
评估文本设置为 “学会使用电源模块的手段，同时能自己进行通电测试。”
评估结果如下：

```
INFO:tensorflow:BLEURT initialized.  
[0.7701483368873596]
```