

# 基于小猿搜题的视觉应用落地

朱珊珊

猿辅导公司 基础研究团队研究员

 猿辅导 [zhuss@fenbi.com](mailto:zhuss@fenbi.com)



# 极客时间

重拾极客精神·提升技术认知

## 下载极客时间App

获取有声IT新闻、技术产品专栏，每日更新



扫一扫下载极客时间App



# 人工智能基础课

“通俗易懂的人工智能入门课”

王天一  
博士 副教授



扫一扫，免费试读

# AI技术内参

你的360度人工智能信息助理

洪亮劫  
Etsy 数据科学主管



扫一扫，免费试读





## 关注落地技术，探寻AI应用场景

- 14万AI领域垂直用户
- 8000+社群技术交流人员，不乏行业内顶级技术专家
- 每周一节干货技术分享课
- AI一线领军人物的访谈
- AI大会的专家干货演讲整理
- 《AI前线》月刊
- AI技能图谱
- 线下沙龙



扫码关注带你涨姿势



# QCon

## 全球软件开发大会

# 成为软件技术专家 的必经之路

### [北京站] 2018

会议：2018年4月20-22日 / 培训：2018年4月18-19日

北京·国际会议中心

# 8折

购票中, 每张立减1360元

团购享受更多优惠



识别二维码了解更多



# ArchSummit

## 全球架构师峰会

2018 · 深圳站

从2012年开始算起，InfoQ已经举办了9场ArchSummit全球架构师峰会，有来自Microsoft、Google、Facebook、Twitter、LinkedIn、阿里巴巴、腾讯、百度等技术专家分享过他们的实践经验，至今累计已经为中国技术人奉上了近千场精彩演讲。

限时**7折**报名中，名额有限，速速报名吧！

● 2012.08.10-12 深圳站



**2018.07.06-09 深圳站**

会议：07.06-07.07

培训：07.08-07.09



# TABLE OF CONTENTES

---

场景

解决方案

再看一次场景

解决方案的落地

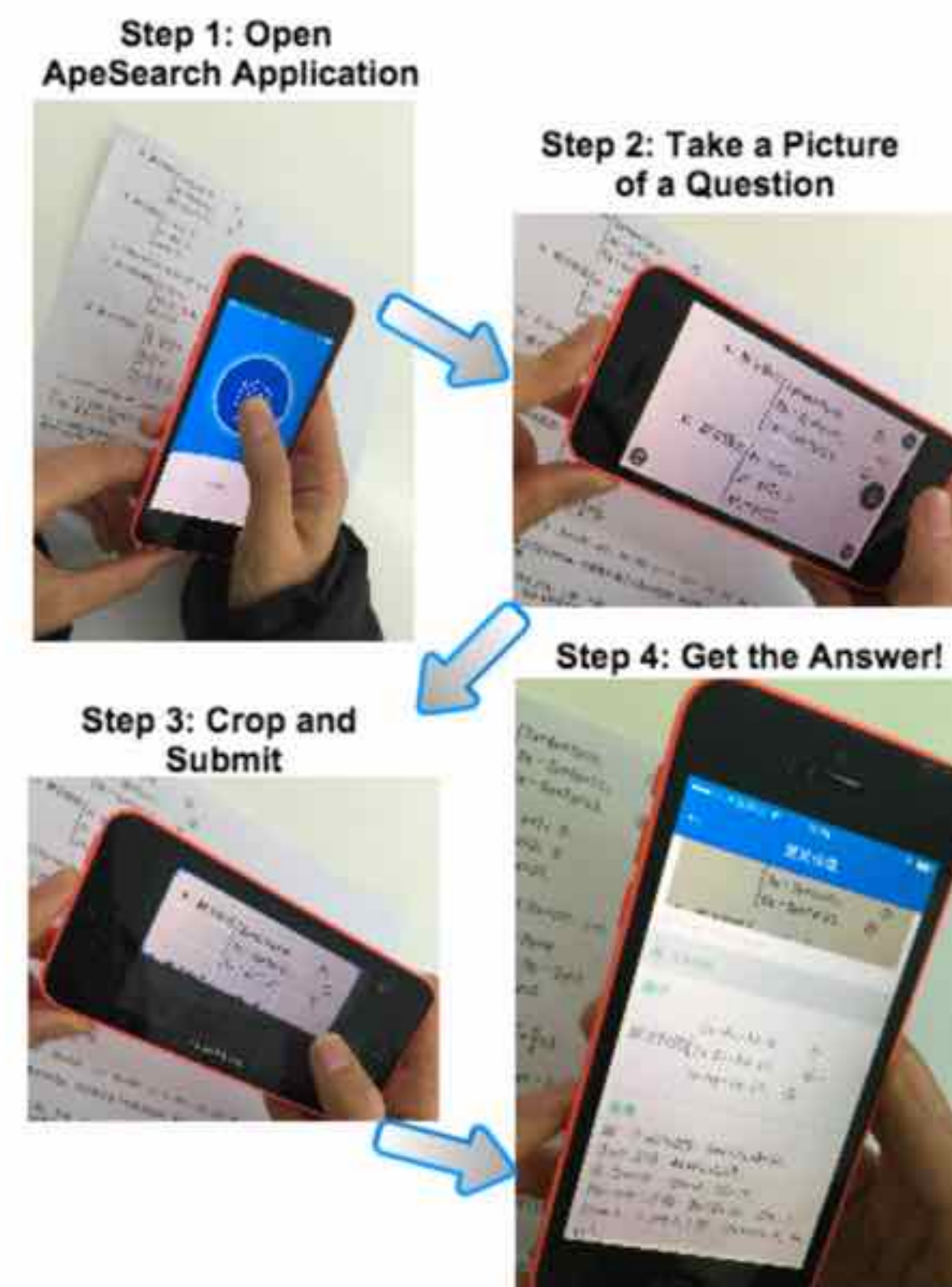
回顾



# 场景

- 小猿搜题

➡ 小猿搜题





# 场景

- 小猿搜题

- ➡ 小猿搜题

- ➡ 英语作文拍照批改

- ➡ 口算检查

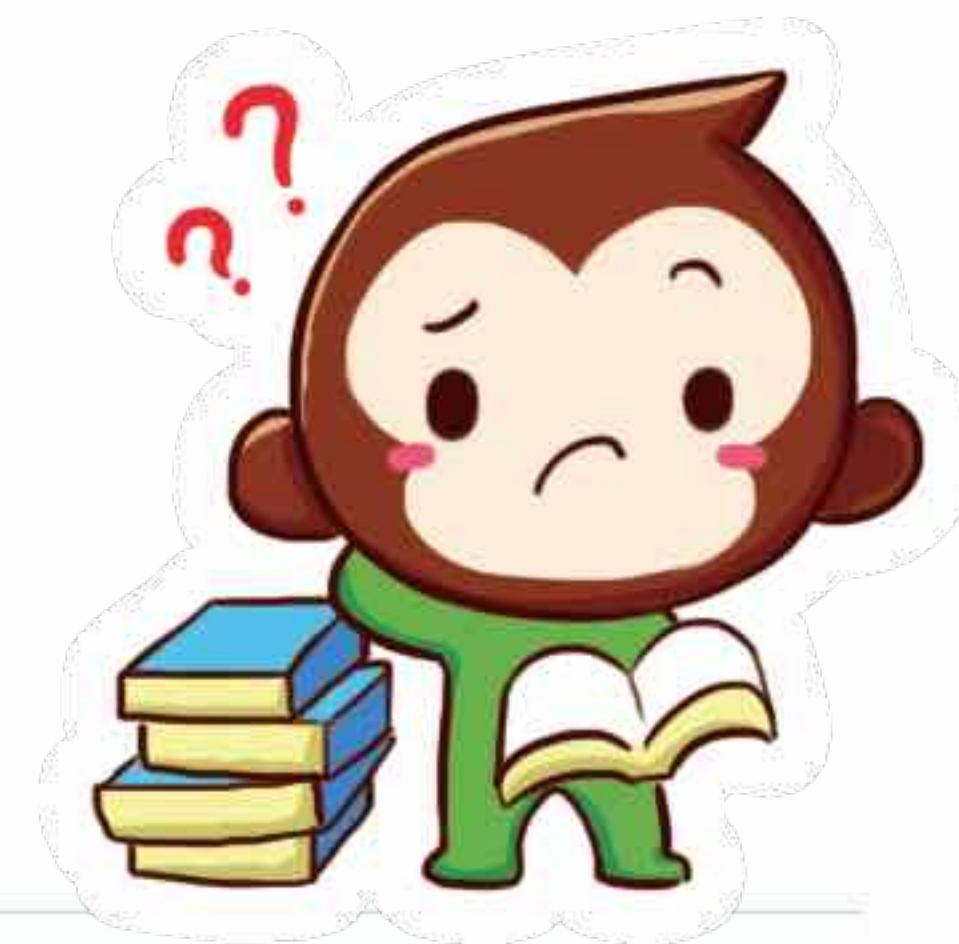
- ➡ 古诗文背诵

- ➡ .....

难点1：手写英文作文的转化

难点2：批改打分的算法

拍照识别？

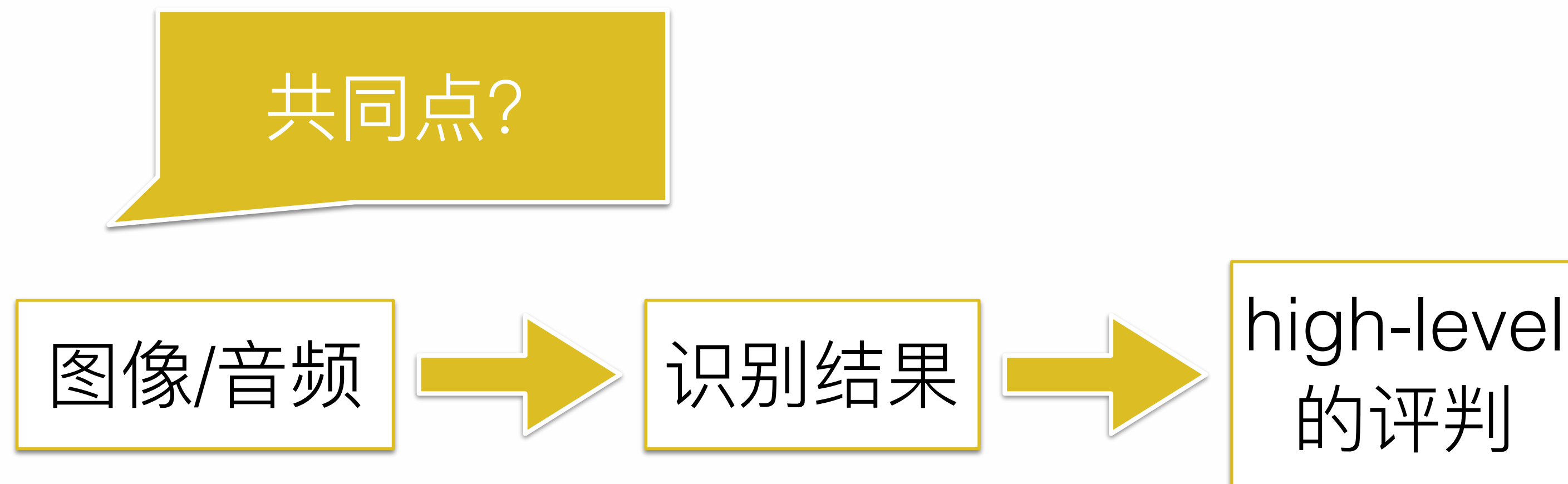




# 场景

- 小猿搜题

- ➡ 小猿搜题
- ➡ 英语作文拍照批改
- ➡ 口算检查
- ➡ 古诗文背诵
- ➡ .....





# 场景

- 小猿搜题

- ➡ 小猿搜题

- ➡ 英语作文拍照批改

← 今天的落地实例！

- ➡ 口算检查

- ➡ 古诗文背诵

- ➡ .....



图像/音频

识别结果

high-level  
的评判

# 英语作文拍照批改

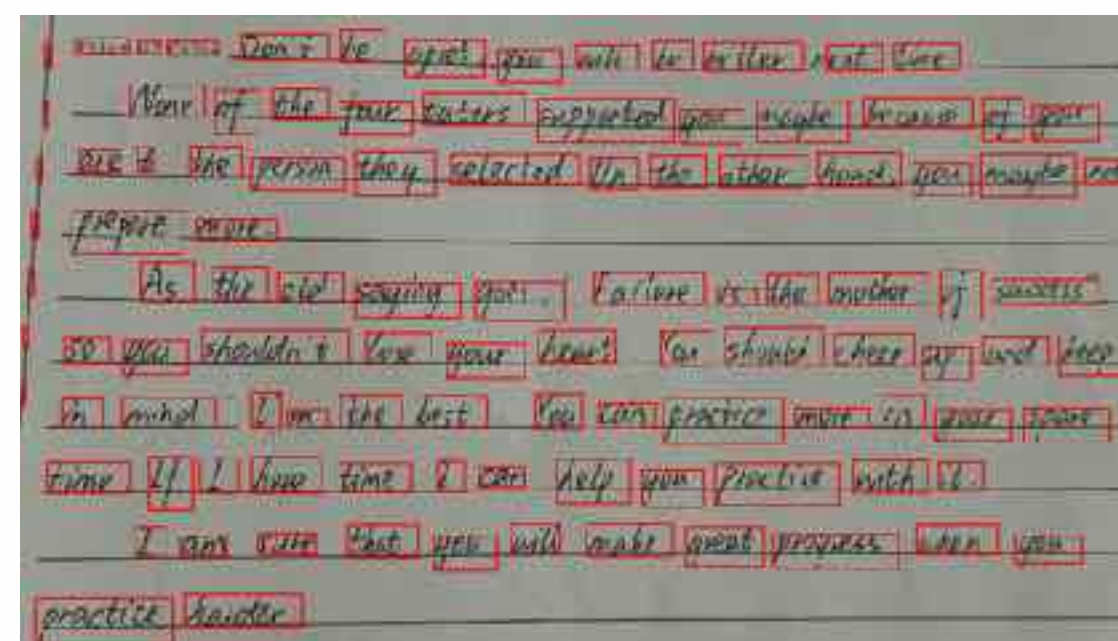
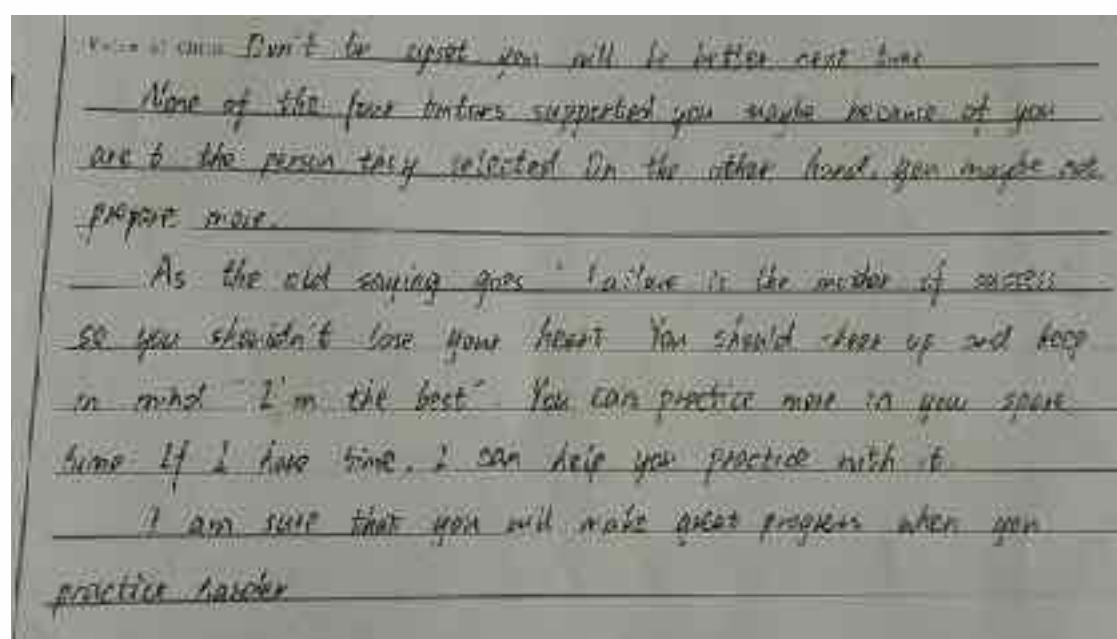
清晰整洁的图片

准确的单词切分

英文字符识别

批改

- ◆ 单词拼写纠错
- ◆ 单词搭配纠错
- ◆ 语法纠错
- ◆ 高级用法加分
- ◆ 评分



Don't be  
upset, you  
will be better  
next time.  
.....



评分: 67.4

评语: 全文结构较为严谨, 再增加词汇量积累就更好了, 高级词汇使用不是很多, 建议增加衔接词汇的积累, 注意不要过多使用简单句。Do not be upset, you will be better next time. None of the four tutors supported you maybe because of you are not the person they selected On the other

[语法错误]: because后面跟句子不需要加of。

As the old saying goes, Failure is the mother of success so you should not lose your heart. You should cheer up and keep in mind, I am the best. You can practice more in your spare time. If I have time, I can help you practice with it. I am sure that you will make great progress when you practice harder

“机器学习落地” ==> “机器学习（中的深度学习）落地”



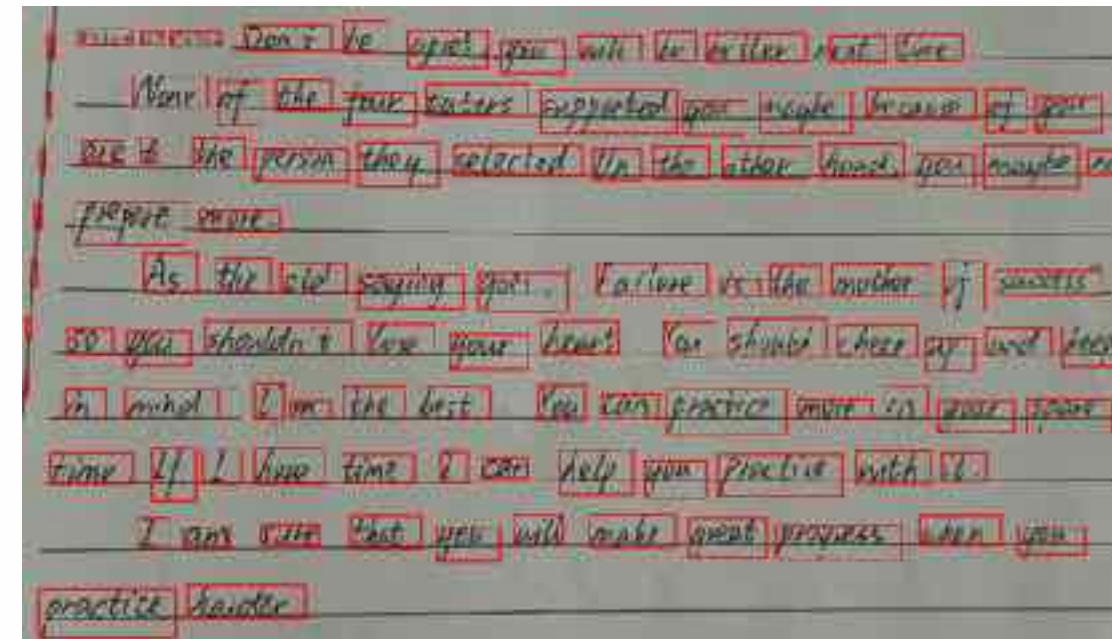
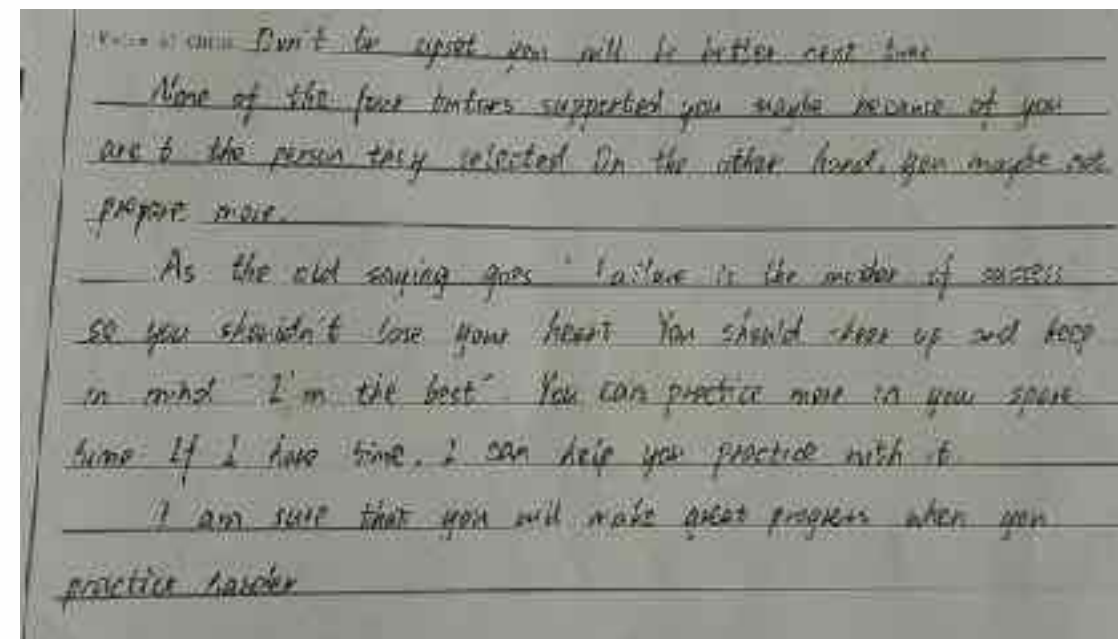
# 场景

图像/音频

识别结果

high-level  
的评判

## What is your solution?



Don't be  
upset, you  
will be better  
next time.  
.....



# TABLE OF CONTENTES

---

场景

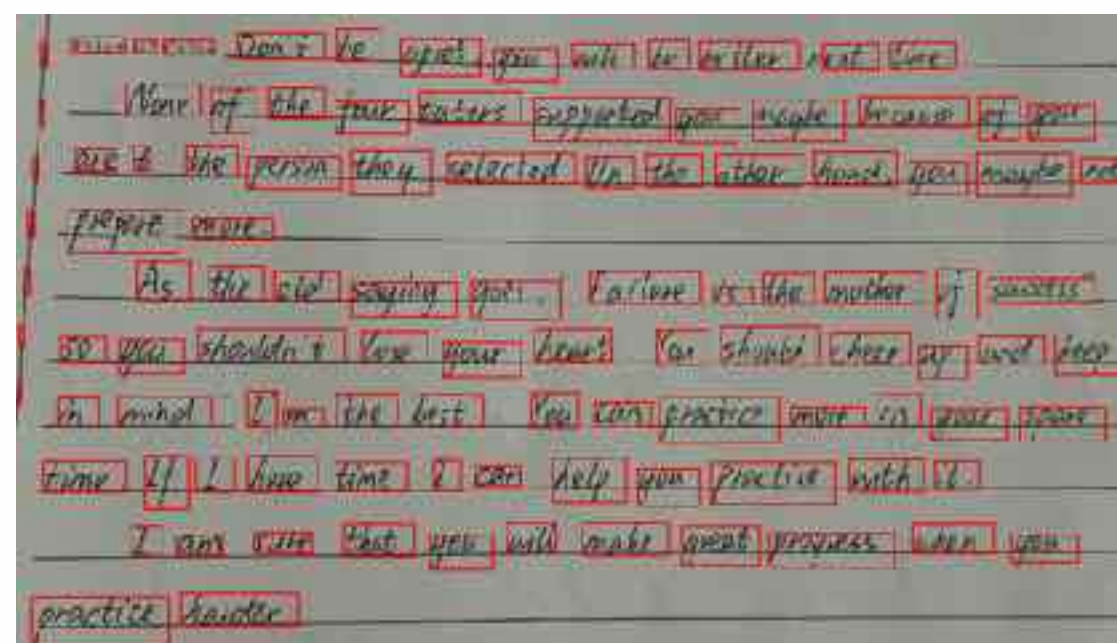
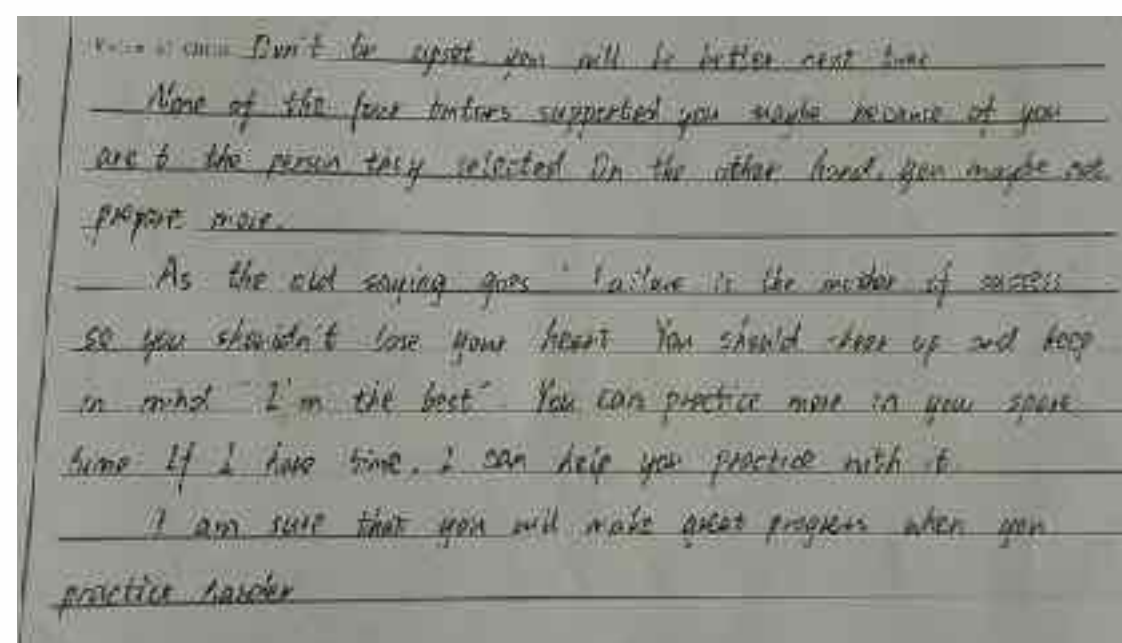
解决方案

再看一次场景

解决方案的落地

回顾

# 解决方案



Don't be  
upset, you  
will be better  
next time.  
.....

充分的单词框标注数据

+

detection模型

?

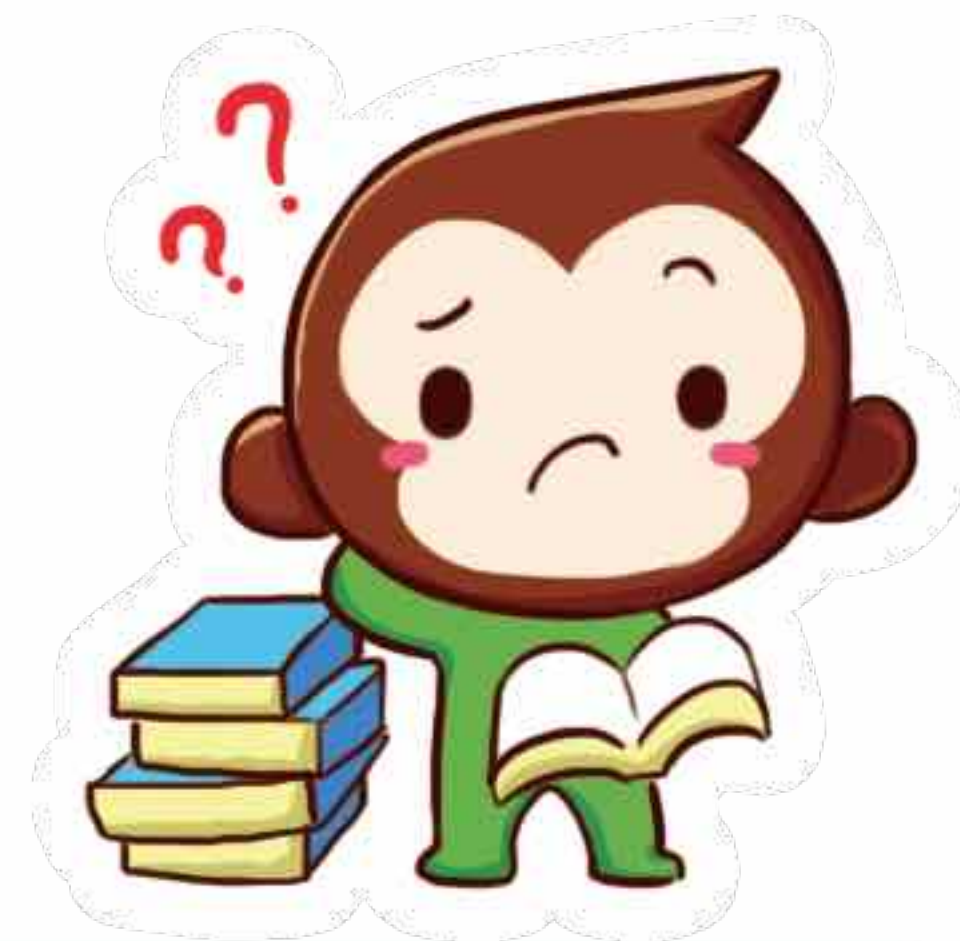
??

充分的字符标注数据

+

识别模型(CNN, RNN,  
CNN+RNN, ...)

?





# 解决方案

为了能快速落地...

传统方法：

1. 规则模型（如，对于图像来说基于连通域的规则）
2. 人工提取feature的判别模型
3. ....

传统方法  
快速出baseline结果



积累标注的Dataset



DL模型  
训练

更准确的结果、  
替代传统模型

补刀

# 解决方案

最后的模型很可能是

传统方法  DL模型  传统方法补刀  DL模型  ...



# TABLE OF CONTENTES

---

场景

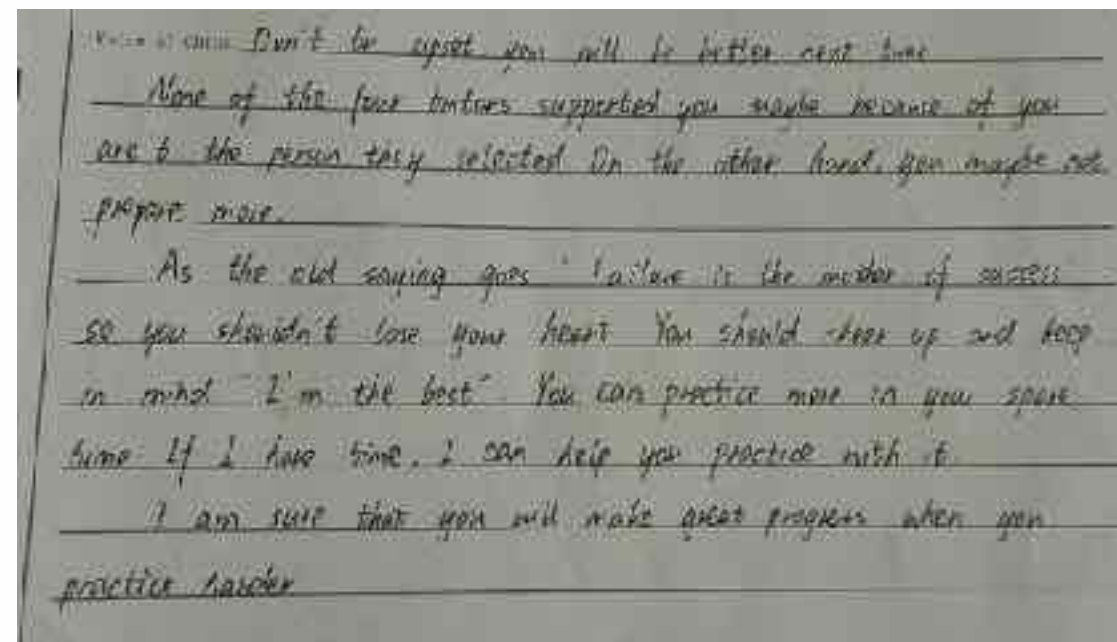
解决方案

再看一次场景

解决方案的落地

回顾

# 再看一次场景 难点!

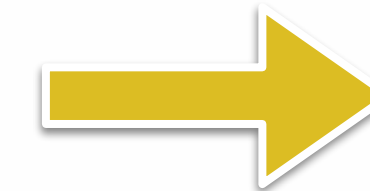


+光照影响  
+模糊  
+倾斜



常见的真实  
图片状况

+印刷体的干扰 (印刷体应不参与评分)  
+批改笔记的干扰  
+删除字符



英语批改任务  
相关的干扰项

(+横格线的干扰)  
(+串行对顺序的干扰)



(注意: 文章中不能出现考生的真实姓名、校名和老师名, 否则不评分)

81. 本周末, 北京友谊商店正在做优惠促销活动, 请根据下面的商品信息为该商店写一广告, 介绍店内的商品名称及价格。

- 男孩的运动鞋每双 20 元, 有各种颜色;
- 绿色短裙每件 14 元;
- 黄色运动袜每双 5 元;
- 白色衬衣每件 20 元, 购三件每件 18 元。

Beijing Friendship Store Sale!

This weekend, our store will have a sale activity. The pair of boy's sports shoes is 20 yuan. It has kinds of color. The green skirts is 14 yuan. The yellow sports socks are 5 yuan. The white shirt is 20 yuan. If you buy three shirts, the price is 18 yuan. Welcome to our store.

第四步: 整理成文

Learning English is hard. I'm glad to know, student Learning English is well. Practise English everyday is important, and in class listen carefully and take note is homework is also important to study English. In addition, listen English program, read newspaper or story are helpful your reading ability. Keep diary in English everyday to is improved your English and some helpful to your time.

My Learning English deficiencies is: composition is not good. Because, I don't take too much time in it. however, I should more read and write to improved the English level.

Learning English is very important. Good luck!

Share and write about your last weekend with your partners. (与同伴分享你上周末的事, 并写下来。)

Last Weekend

Last weekend, I was at home. It was sunny. First, I had fun with friends. Next, I watched TV. After that, I ate some cotton candy. I was really happy.

难点!



# TABLE OF CONTENTES

---

场景

解决方案

再看一次场景

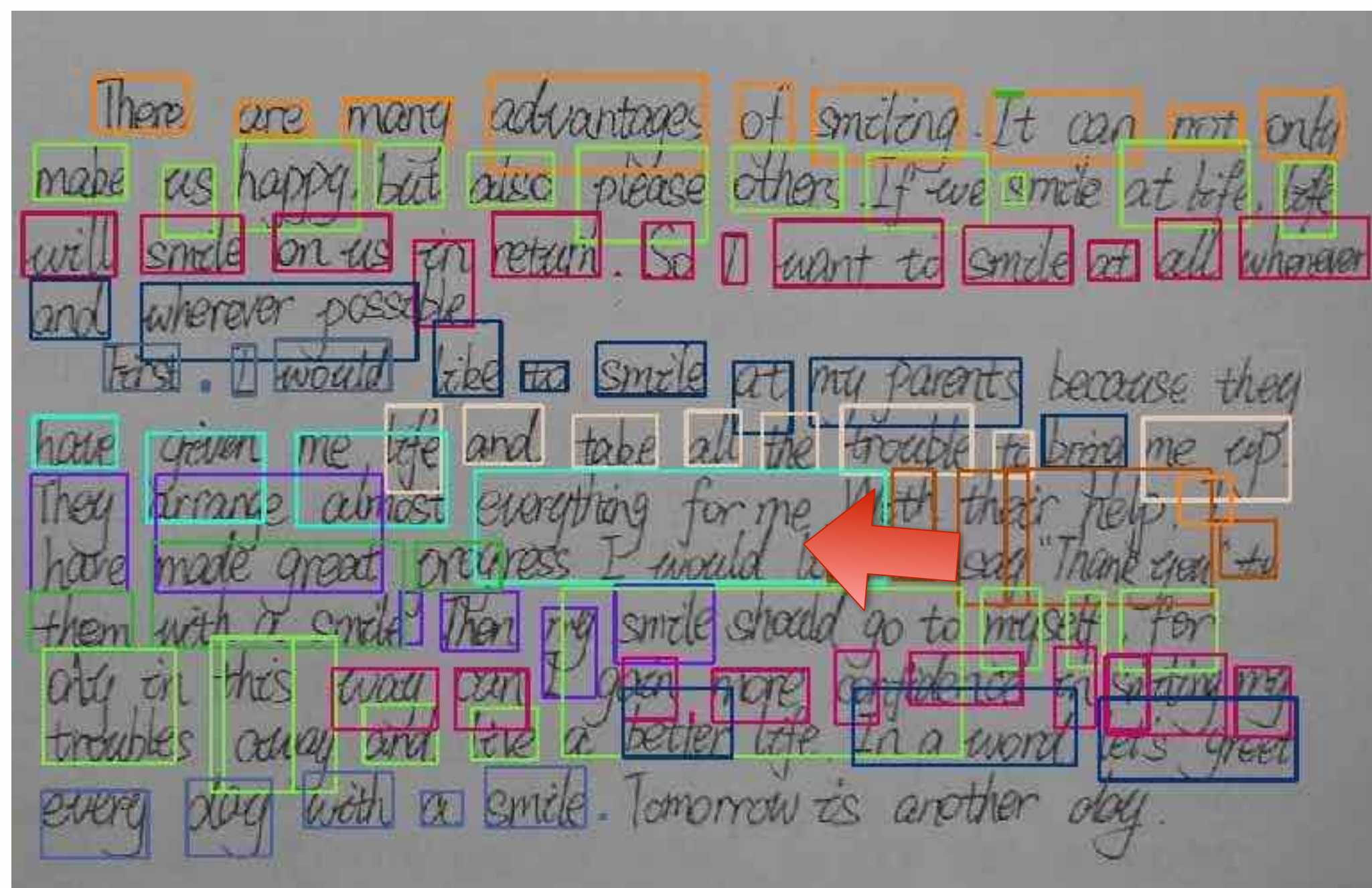
解决方案的落地

回顾



# 传统方法baseline

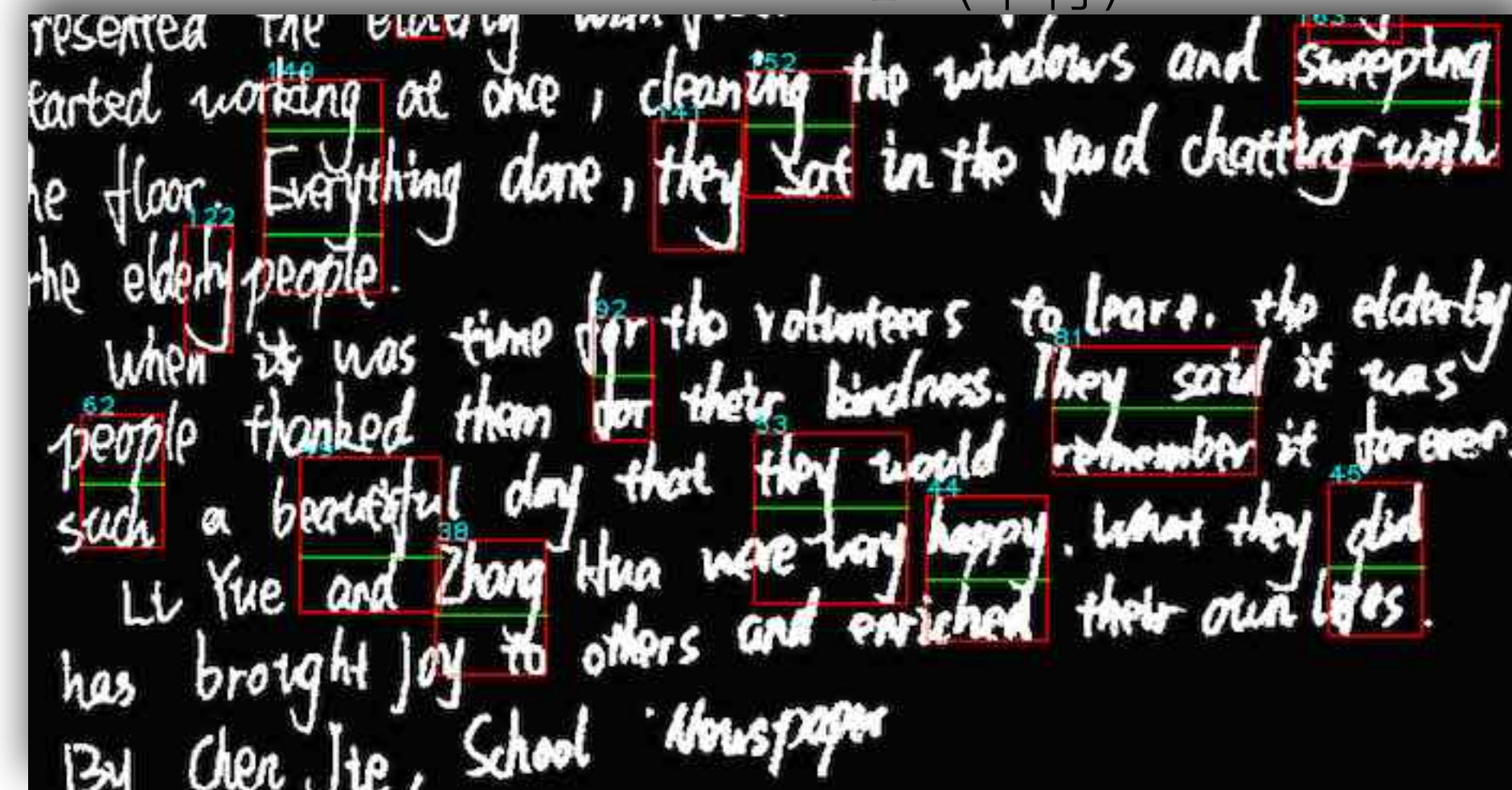
- ☒ 光照影响
- ☐ 模糊
- ☒ 倾斜
- ☐ 印刷体的干扰
- ☐ 批改笔迹的干扰
- ☐ 删除字符
- ☒ (横格线的干扰)
- ☐ (串行)





# 传统方法baseline

- ☑ 光照影响
- ☑ 模糊
- ☑ 倾斜
- ☐ 印刷体的干扰
- ☐ 批改笔迹的干扰
- ☐ 删除字符
- ☑ (横格线的干扰)
- ☑ (串行)





# 传统方法baseline

- ☑ 光照影响
- ☑ 模糊
- ☑ 倾斜
- ☑ 印刷体的干扰
- ☑ 批改笔迹的干扰
- ☑ 删除字符
- ☑ (横格线的干扰)
- ☑ (串行)



Class 1: 印刷体区域  
Class 2: 删除区域  
Class 3: 批改区域  
Class 4: 常规英文手写区域

# 从何处推进DL模型的落地？

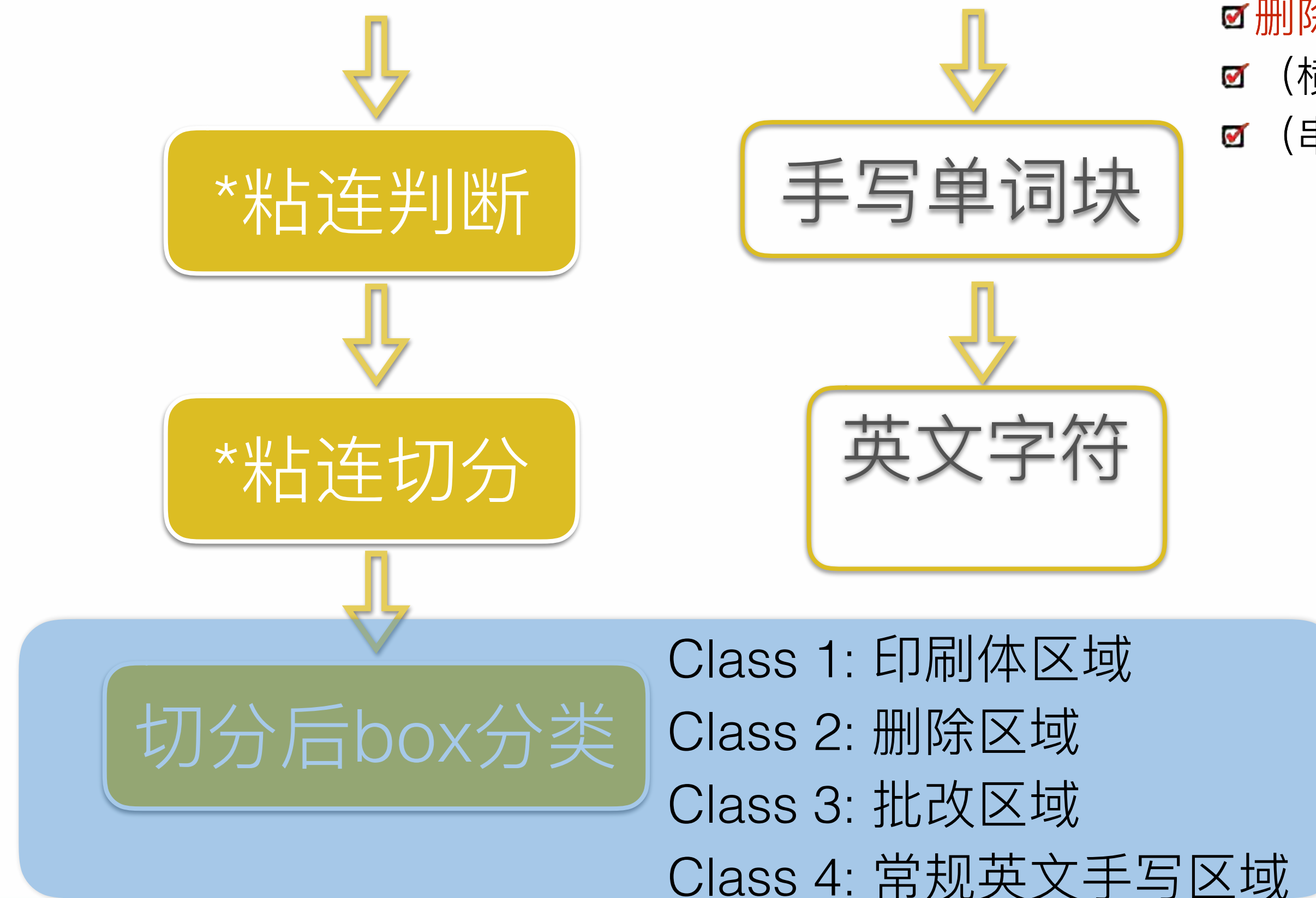
传统模型不够简洁优美？ 实用性上考虑！





# 传统方法baseline

- ☑ 光照影响
- ☑ 模糊
- ☑ 倾斜
- ☑ 印刷体的干扰
- ☑ 批改笔迹的干扰
- ☑ 删除字符
- ☑ (横格线的干扰)
- ☑ (串行)

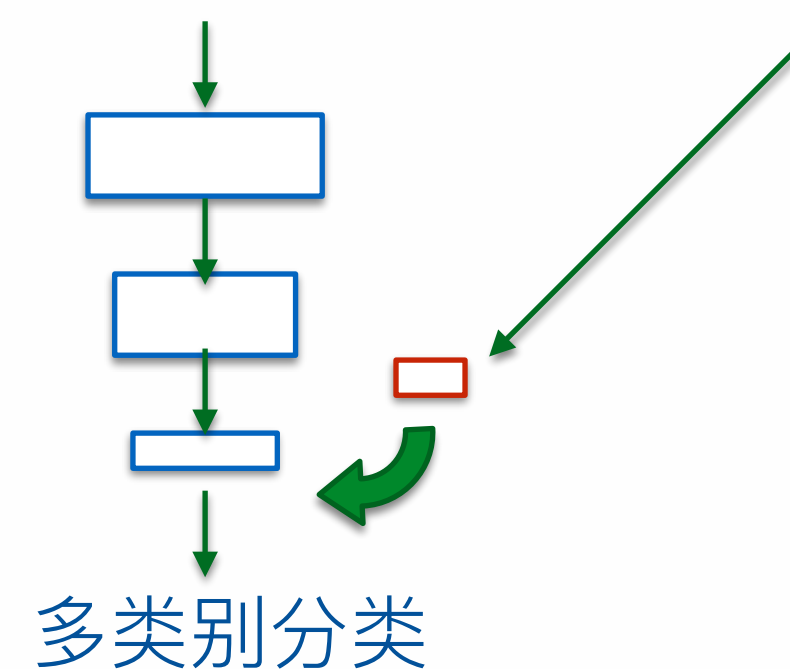
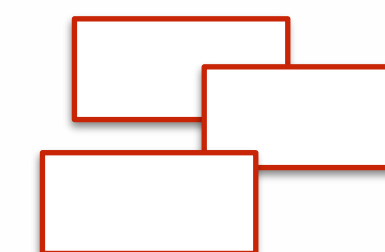
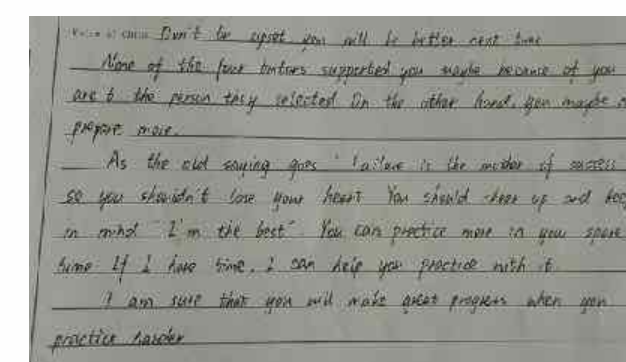


# DL模型替代

切分后box分类

手写单词

- Class 1: 印刷体区域
- Class 2: 删除区域
- Class 3: 批改区域
- Class 4: 常规英文手写区域







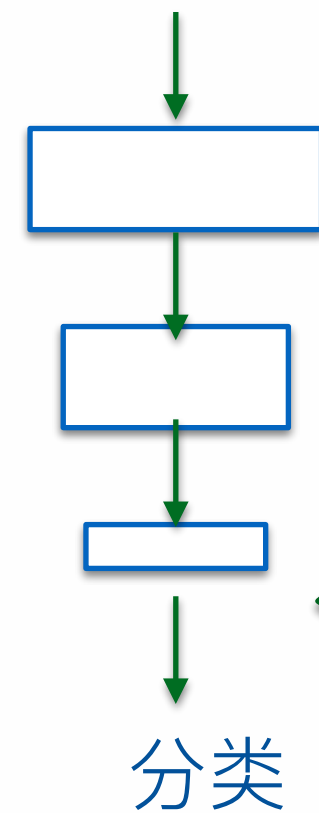
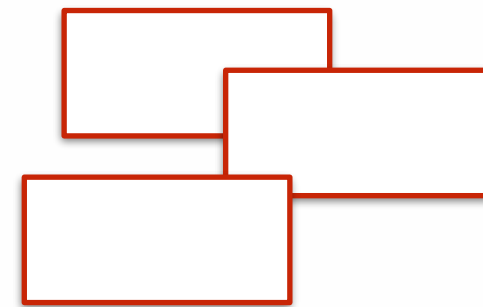
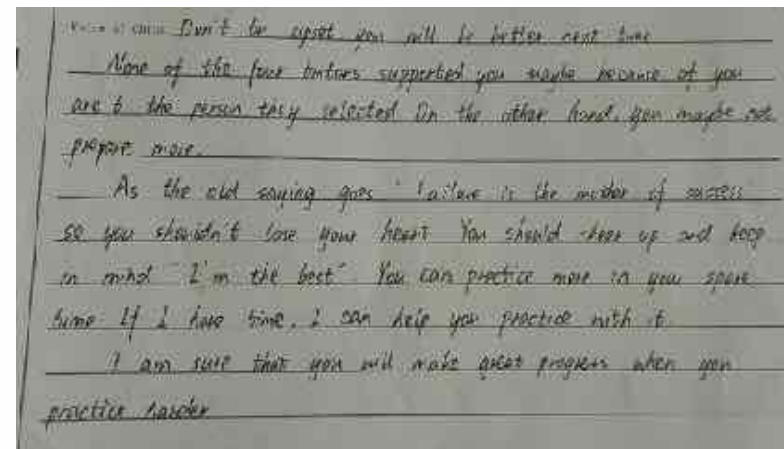
# DL模型替代

- ☑ 光照影响
- ☑ 模糊
- ☑ 倾斜
- ☑ 印刷体的干扰
- ☑ 批改笔迹的干扰
- ☑ 删除字符
- ☑ (横格线的干扰)
- ☑ (串行)



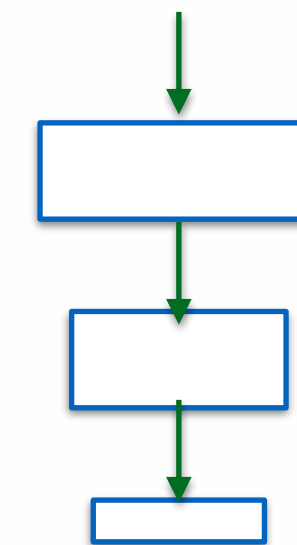
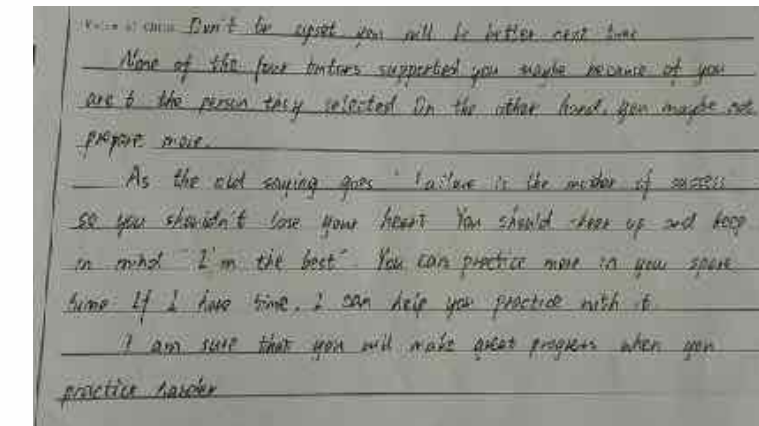


# DL的推进



对象：每个  
现有框

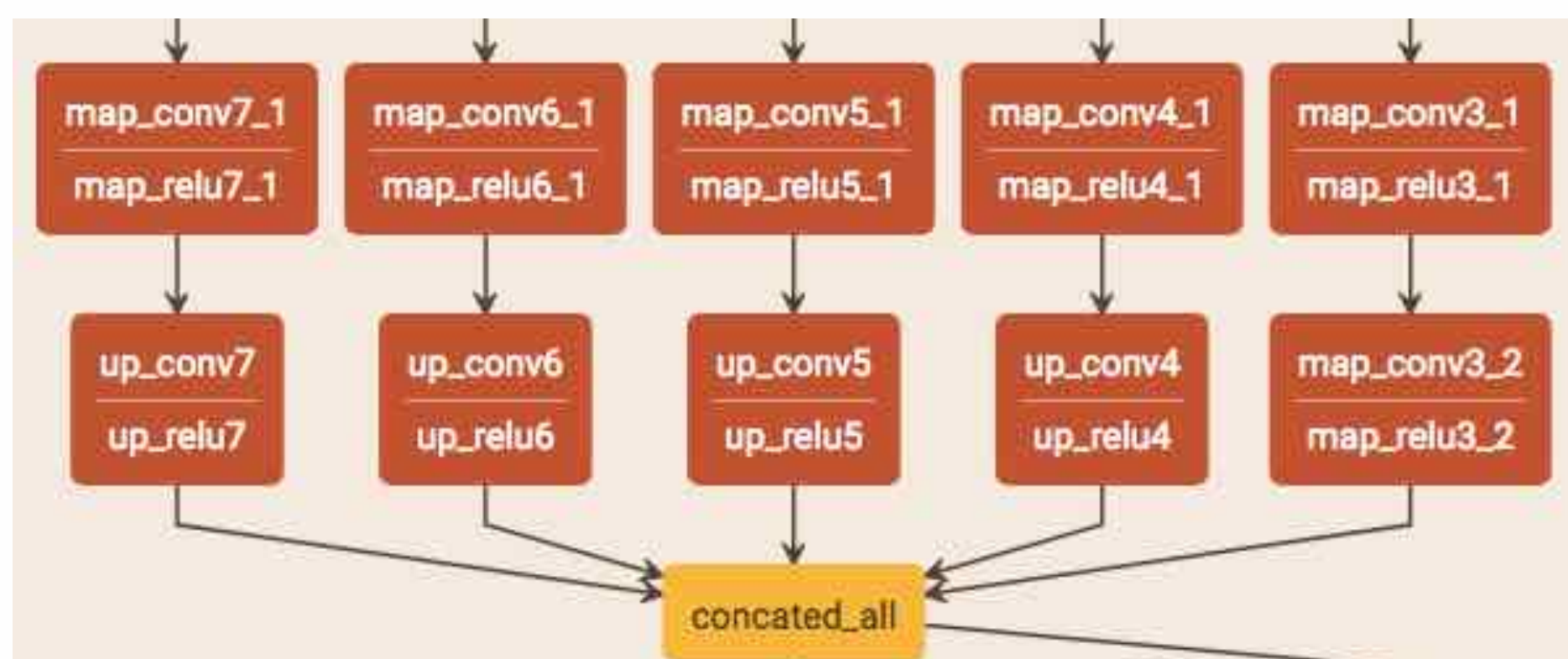
标注数据集



feature map上每一点  
(原图8x8区域) 的  
“是/不是英文手写”的  
分类

对象：原图每8x8  
像素块





```
# Mapping 7
n.map_conv7_1, n.map_relu7_1 = conv_relu(n.pool7, num_output=512)
n.up_conv7, n.up_relu7 = deconv_relu(n.map_relu7_1, num_output=512,
    kernel_size=16, stride=16)

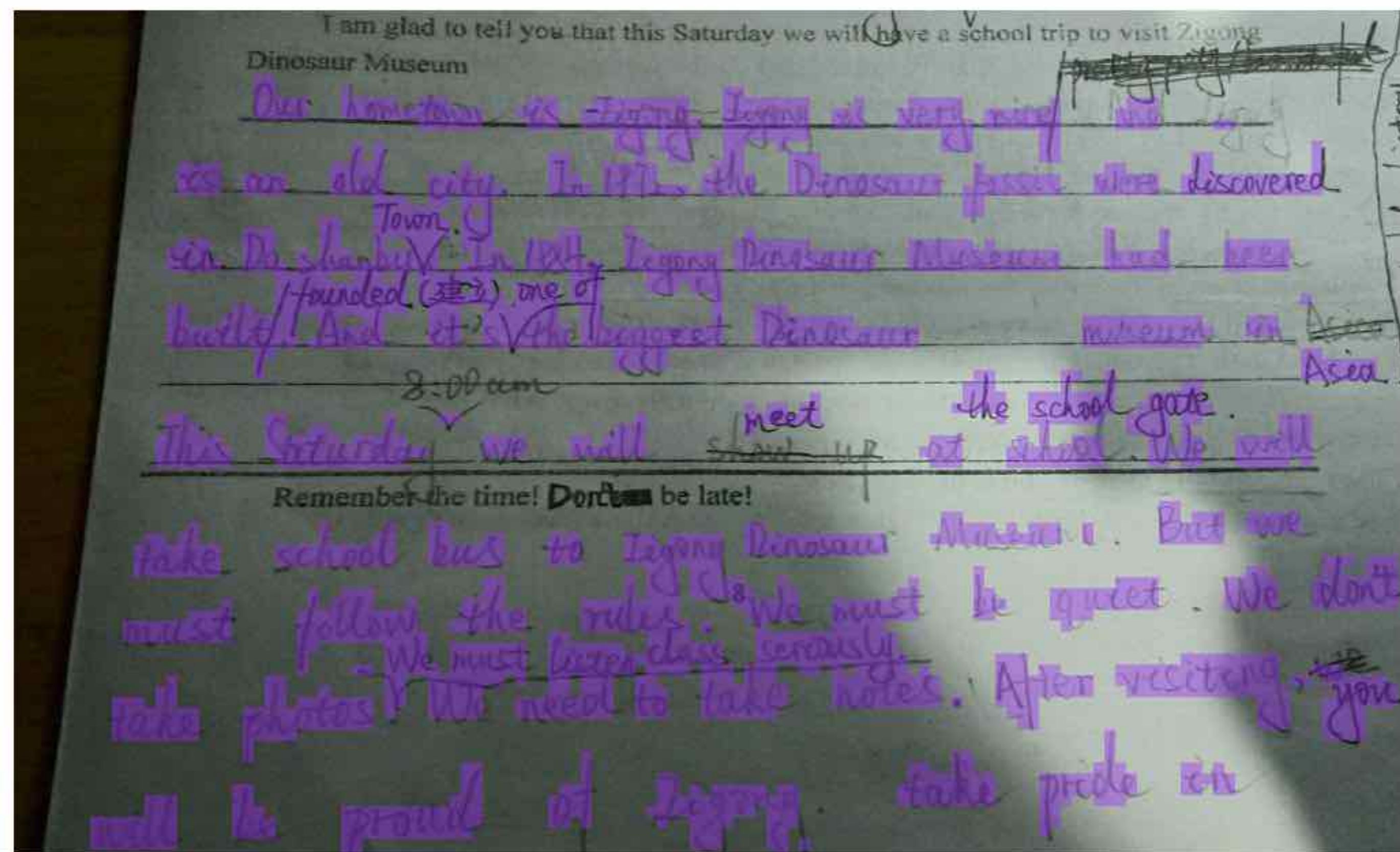
# Mapping 6
n.map_conv6_1, n.map_relu6_1 = conv_relu(n.pool6, num_output=512)
n.up_conv6, n.up_relu6 = deconv_relu(n.map_relu6_1, num_output=512,
    kernel_size=8, stride=8)

# Mapping 5
n.map_conv5_1, n.map_relu5_1 = conv_relu(n.pool5, num_output=512)
n.up_conv5, n.up_relu5 = deconv_relu(n.map_relu5_1, num_output=512,
    kernel_size=4, stride=4)

# Mapping 4
n.map_conv4_1, n.map_relu4_1 = conv_relu(n.pool4, num_output=256)
n.up_conv4, n.up_relu4 = deconv_relu(n.map_relu4_1, num_output=256)

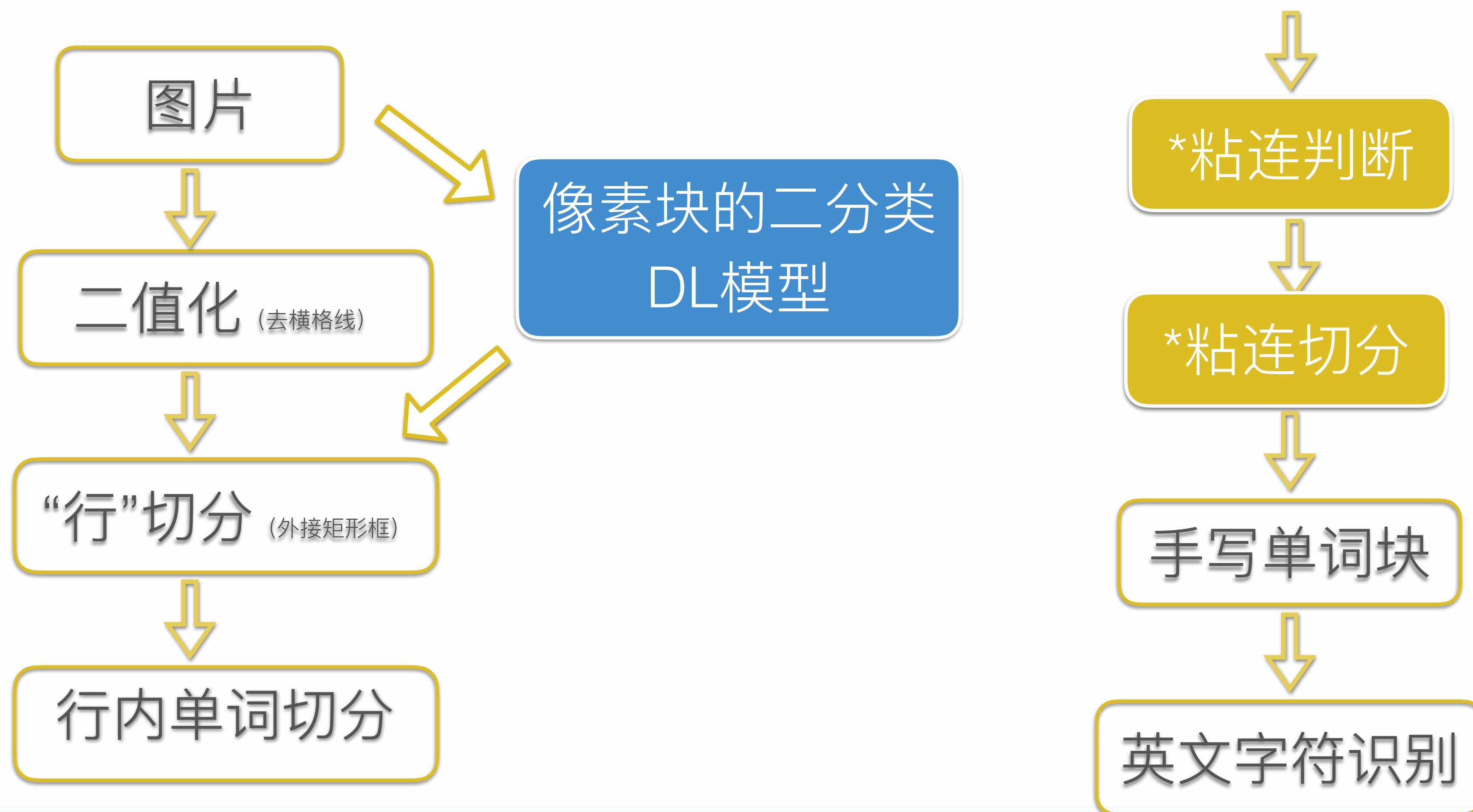
# Mapping 3
n.map_conv3_1, n.map_relu3_1 = conv_relu(n.pool3, num_output=128)
n.map_conv3_2, n.map_relu3_2 = conv_relu(n.map_relu3_1, num_output=128)
n.concated_all = L.Concat(n.map_relu3_2, n.up_relu4, n.up_relu5, n.up_relu6, n.up_relu7)

# fc 1
n.fc_conv1_1, n.fc_relu1_1 = conv_relu(n.concated_all, num_output=1024)
n.fc_conv1_2, n.fc_relu1_2 = conv_relu(n.fc_relu1_1, num_output=1024)
n.fc_conv1_3, n.fc_relu1_3 = conv_relu(n.fc_relu1_2, num_output=4)
```

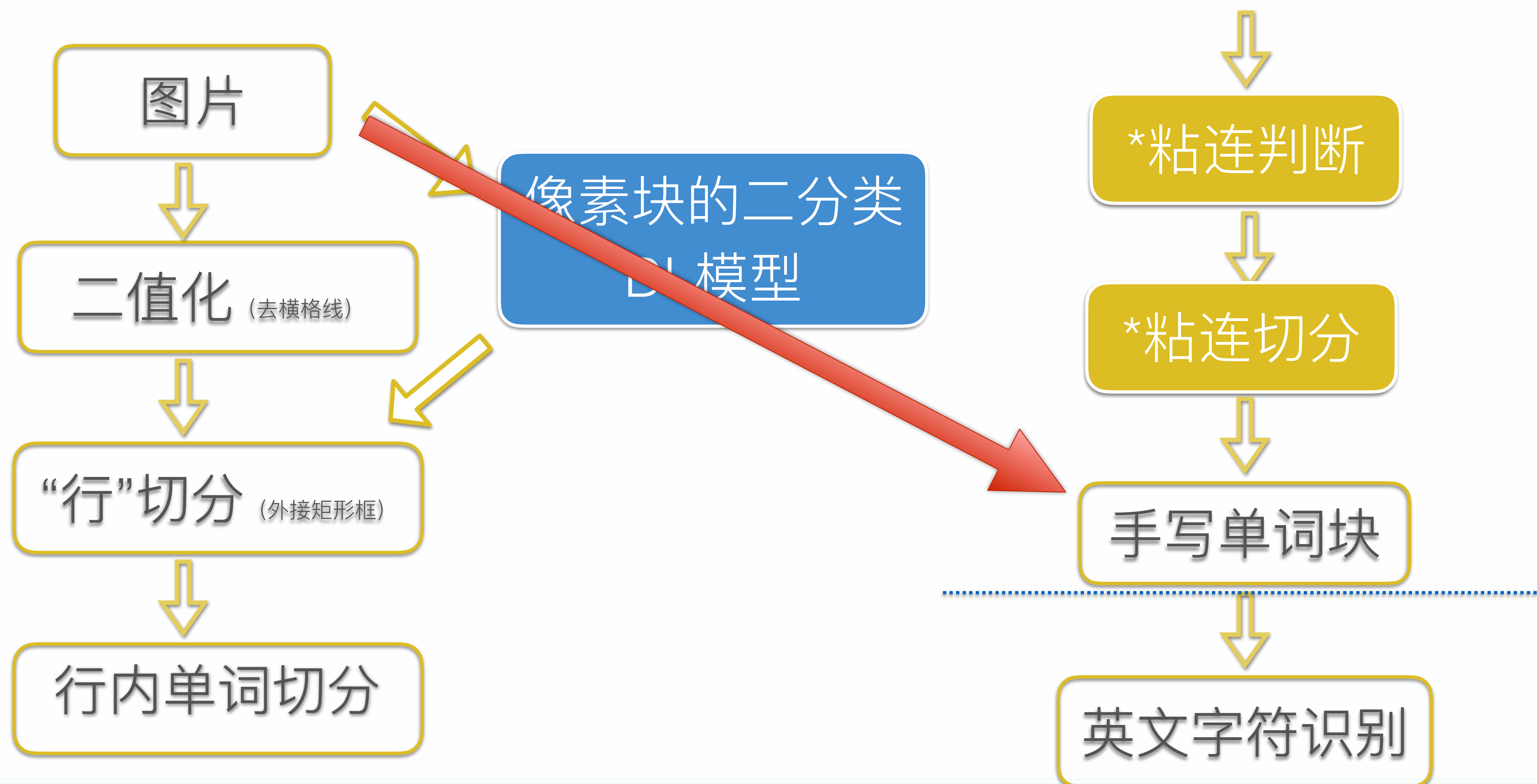




# DL的推进

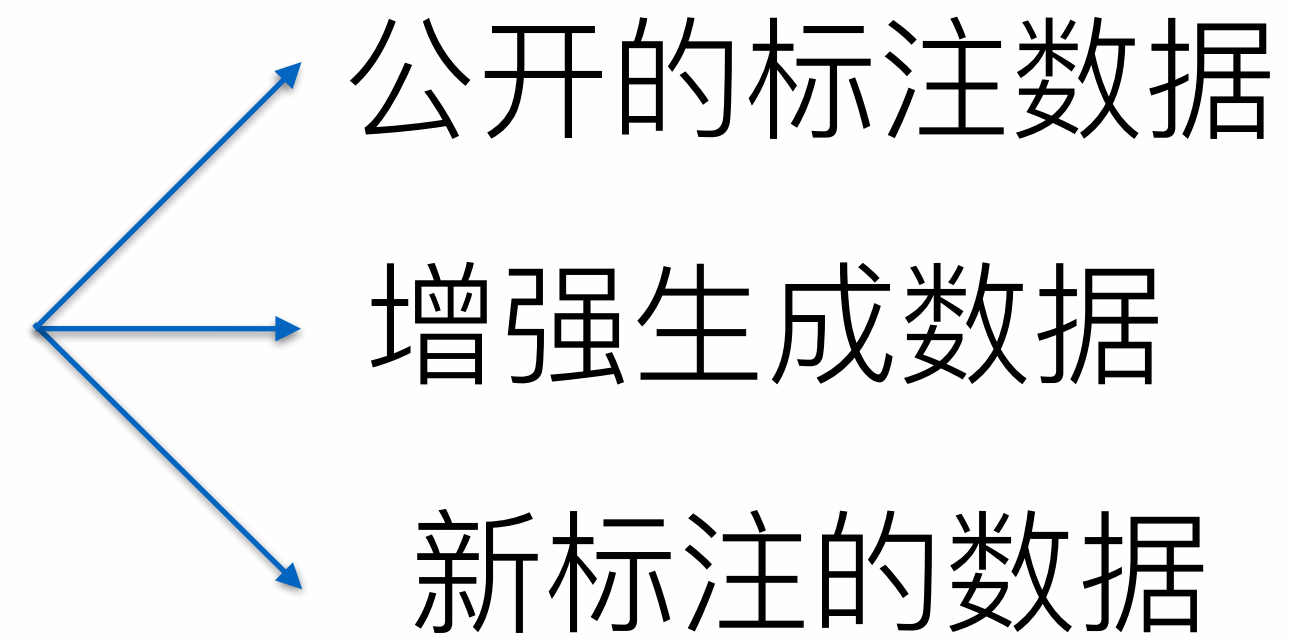


# 进一步推进?





# DL识别模型



中国学生的英文手写vs.  
外国人手写  
【要符合你面对的真实数据的  
特点和分布】

- CNN
- RNN
- CNN+RNN
- CNN+RNN+CTC
- .....

# TABLE OF CONTENTES

---

场景

解决方案

再看一次场景

解决方案的落地

回顾



# 回顾



从何处推进DL模型的落地？传统模型不够简洁优美？实用性上考虑！

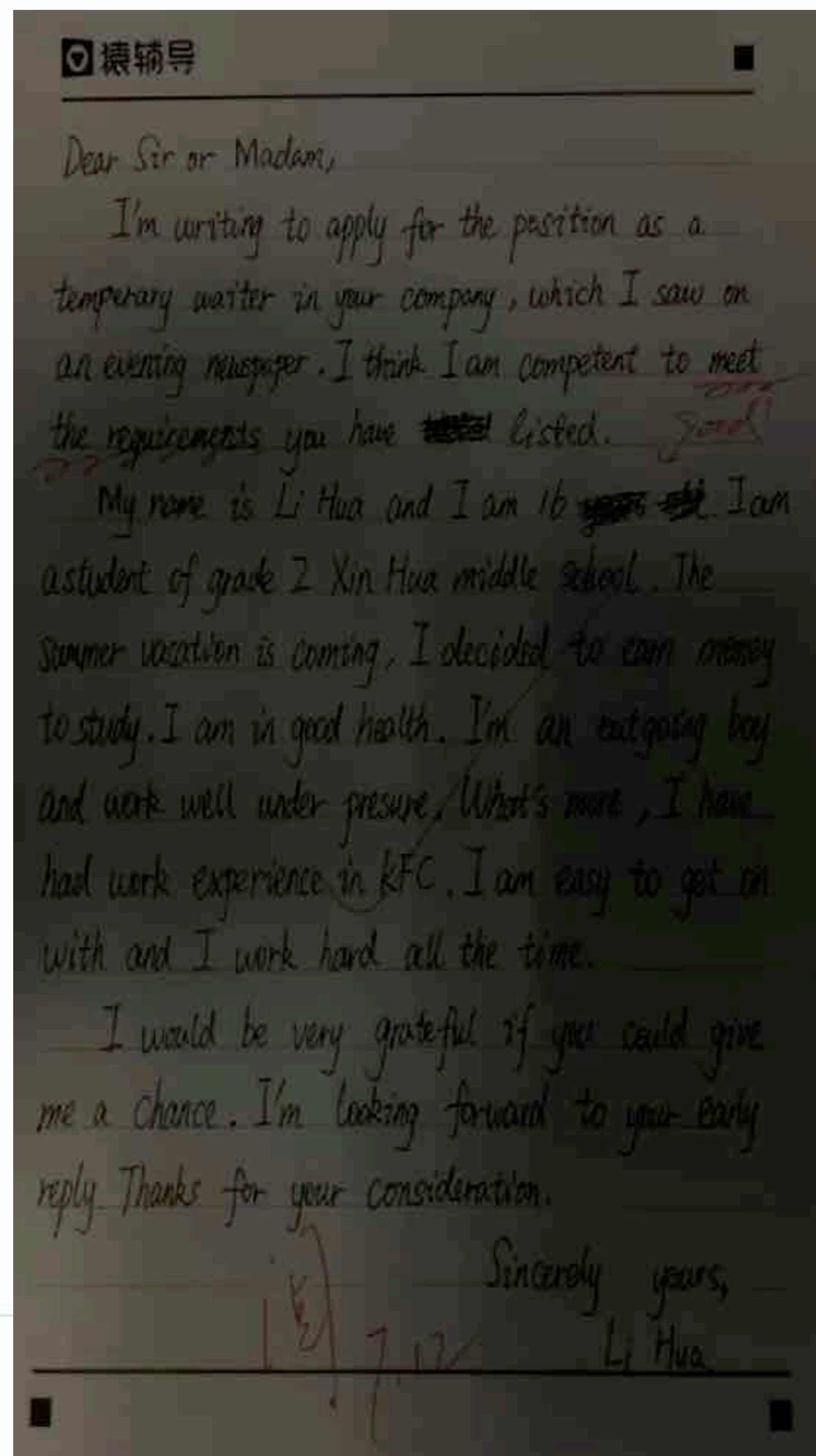


(传统方法) + DL模型 + (传统方法补刀) + DL模型 + ...

落地



欢迎大家的  
提问和交流！



落地





# Thanks!

