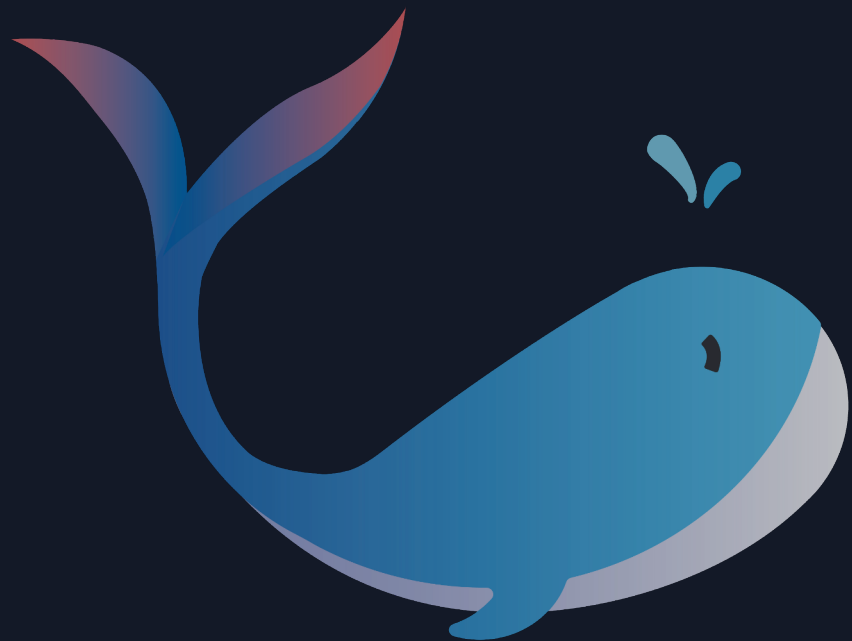


Datawhale零基础入门CV赛事

街景字符编码识别

Baseline和比赛介绍

分享人：阿水



目录

contents

Part 1 比赛介绍

Part 2 baseline

Part 3 比赛知识点

天池新人赛由天池与Datawhale联合发起，并提供学习内容和组织学习：

- ❑ Datawhale是一个专注于数据科学与AI领域的开源组织；
- ❑ CV直播PPT 可关注Datawhale，回复关键词 **CV直播** 下载；
- ❑ 同时可以加入Datawhale数据竞赛交流群，一起组队参赛，交流学习；



个人介绍

阿水, <https://www.zhihu.com/people/finlayliu>

- ✓ 天池数据大神 ;
- ✓ 知乎小V, 全网文章阅读量20W+ ;
- ✓ 数据科学爱好者, 擅长计算机视觉 ;
- ✓ Datawhale成员, 内容学习 & 分享者 ;

数据竞赛开源项目发起人 :

<https://github.com/datawhalechina/competition-baseline>

Part 1 比赛介绍

本次新人赛是Datawhale与天池联合发起的零基础入门系列赛事第二场：零基础入门CV赛事之街景字符识别。

零基础入门系列赛事：

- ❑ 比赛为个人赛，但可以组队学习；
- ❑ Datawhale提供组队学习社群，组队打卡学习；
- ❑ 全程360度分享，包括baseline和比赛知识点；

零基础入门数据挖掘之二手车交易价格预测大赛

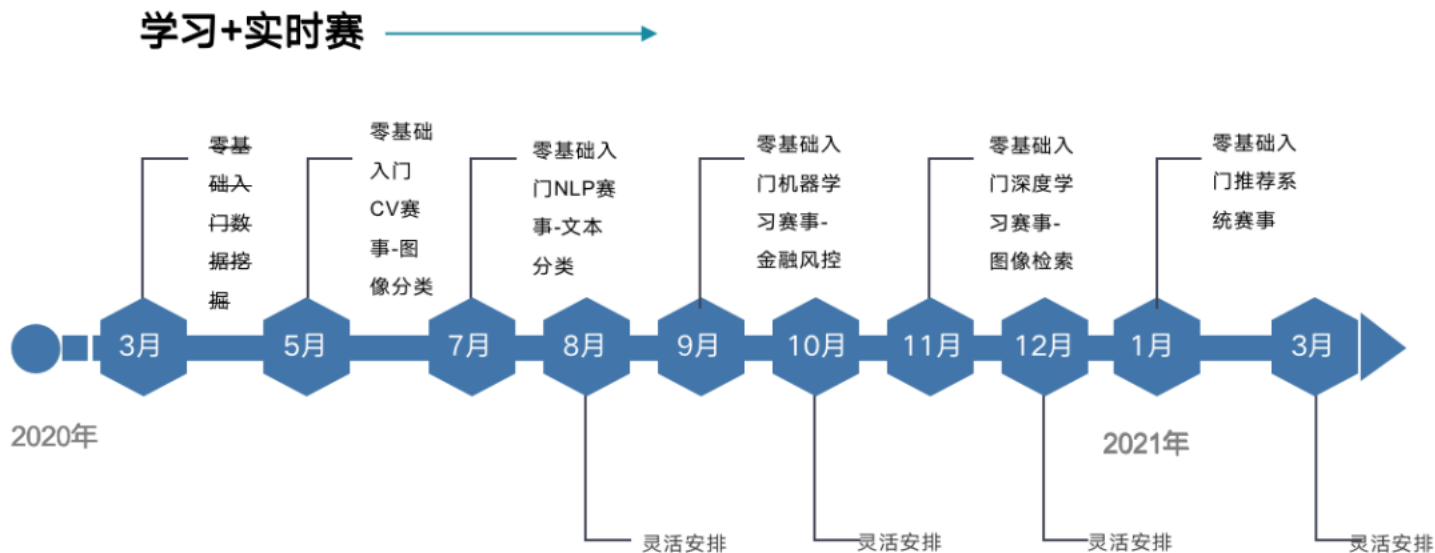
第一场：<https://tianchi.aliyun.com/competition/entrance/231784/introduction>

零基础入门CV赛事之街景字符识别

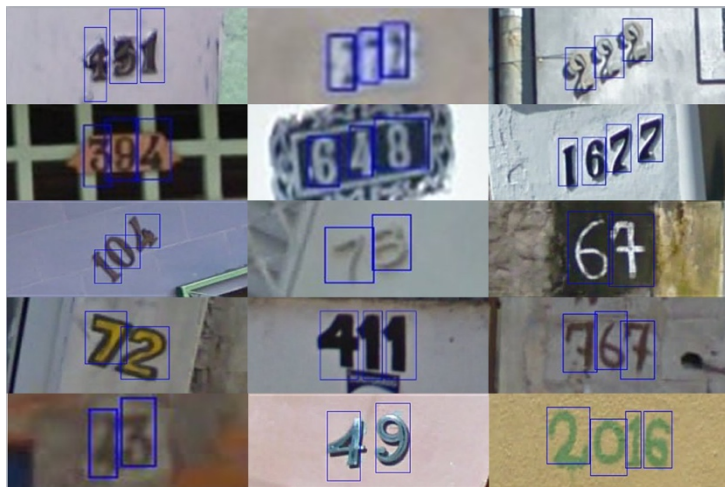
第二场：<https://tianchi.aliyun.com/competition/entrance/531795/introduction>

比赛介绍

本次竞赛是Datawhale联合天池的系列学习（数据挖掘、CV、NLP、机器学习、深度学习、推荐系统）第二场。



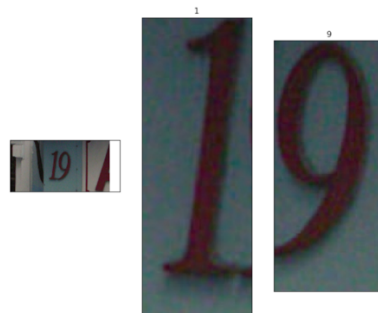
赛题主题：赛题以计算机视觉中字符识别为背景，要求选手预测真实场景下的字符识别，这是一个典型的字符识别问题。



赛题目标：通过这道赛题可以引导大家走入计算机视觉的世界，主要针对竞赛选手上手视觉赛题，提高对数据建模能力。

赛题数据：赛题需要选手识别图片中所有的字符，为了降低比赛难度，我们提供了训练集、验证集和测试集中字符的位置框。

Field	Description
top	左上角坐标X
height	字符高度
left	左上角最表Y
width	字符宽度
label	字符编码



注：数据集来源自SVHN，网页链接<http://ufldl.stanford.edu/housenumbers/>，并进行匿名处理和噪音处理，请各位选手使用比赛给定的数据集完成训练。

Part 2 baseline

本次赛题：赛题本质是分类问题，需要对图片的字符进行识别。但赛题给定的数据图片中不同图片中包含的字符数量不等，如下图所示。有的图片的字符个数为2，有的图片字符个数为3，有的图片字符个数为4。

字符：42 字符个数：2	
字符：241 字符个数：3	

思路1：可以将赛题抽象为一个**定长字符识别**问题，在赛题数据集中大部分图像中字符个数为2-4个，最多的字符个数为6个。

因此可以对于所有的图像都抽象为6个字符的识别问题，字符23填充为23XXXX，字符231填充为231XXX。

23XXXX

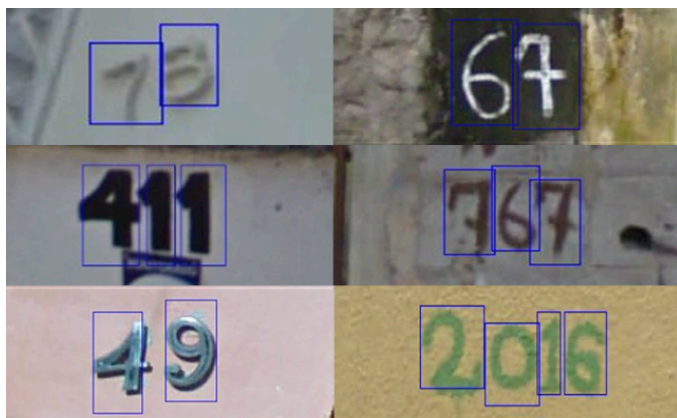
241XXX

思路2：在字符识别研究中，有特定的方法来解决此种不定长的字符识别问题，比较典型的有CRNN字符识别模型。

在本次赛题中给定的图像数据都比较规整，可以视为一个单词或者一个句子。



思路3：在赛题数据中已经给出了训练集、验证集中所有图片中字符的位置，因此可以首先将字符的位置进行识别，利用物体检测的思路完成。



思路1实现思路：

- ❑ 构建读取数据集；
- ❑ 构建多分类模型；
- ❑ 训练+验证；
- ❑ 提交；

思路1 notebook：

<https://tianchi.aliyun.com/notebook-ai/detail?spm=5176.12586969.1002.6.2ce832bcmeL9bN&postId=108342>

Part 3 知识点

Baseline中使用的知识点：

- ❑ 使用Pytorch读取数据集；
- ❑ 构建多分类模型；
- ❑ 使用ImageNet预训练模型进行微调；
- ❑ 构建训练集和验证；
- ❑ 模型保存于加载；

Baseline如何继续深入，提高精度？

- ❑ 1- 首先提高单模型在20个Epoch以内的精度
单纯baseline思路，是榜单0.7+左右的分数；
- ❑ 2- 尝试加入其它数据扩增方法，如模糊和像素噪音；
- ❑ 3- 尝试进行多折预测结果集成；

- ❑ 4- 尝试其它模型，或者改进模型的多字符训练过程；
假设修改多字符分类的损失函数；
- ❑ 5- 尝试其它模型，对结果进行修正；
训练一个字符字数预测模型，对结果进行修正；

其它尝试和思考：

- 这个赛题需要使用物体/字符检测模型么？
- 如何复用YOLO、SSD或Faster-RCNN模型？

问答

- ❑ 你对比赛有什么问题？
- ❑ 你对学习有什么问题？
- ❑ 你对PPT内容有什么问题？



一个专注于AI领域的开源组织

