

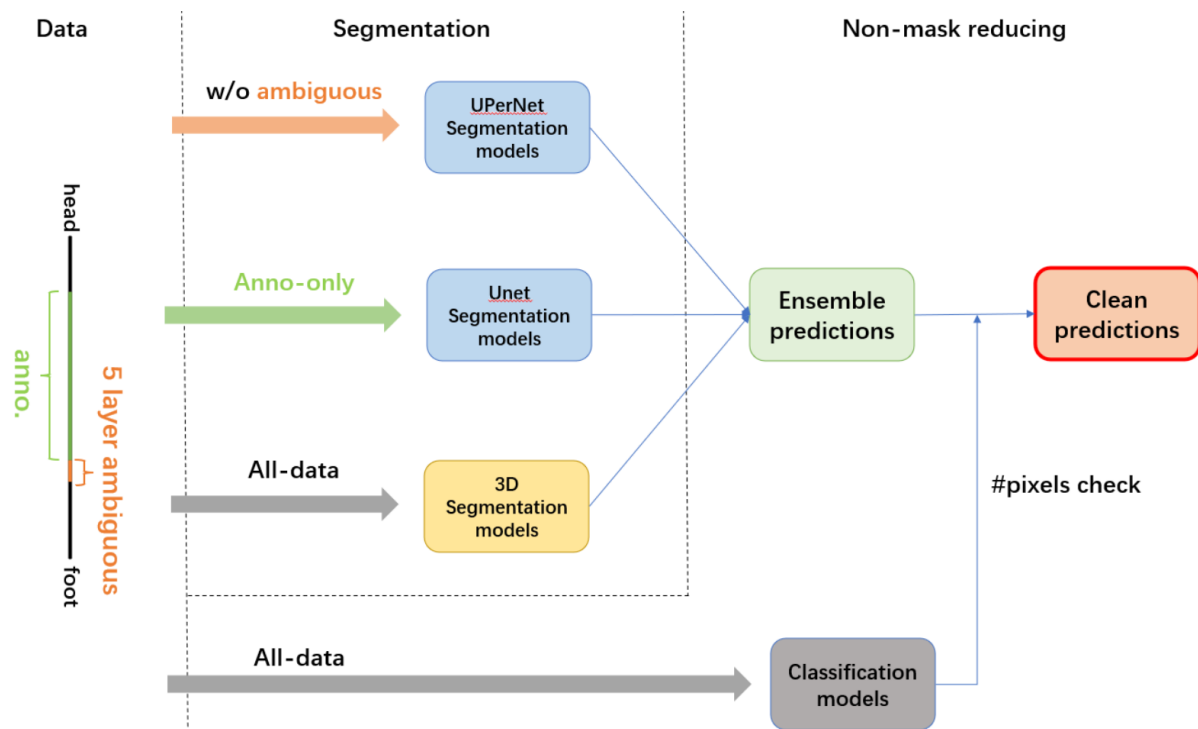
lesson7 比赛复盘

在 discussion 搜索 “place solution” 关键字

1st place solution

链接：<https://www.kaggle.com/competitions/uw-madison-gi-tract-image-segmentation/discussion/337197>

<https://www.kaggle.com/competitions/uw-madison-gi-tract-image-segmentation/discussion/337217>



- 灰色 - CLS Group:
 - train: 4 Unets with efficientnet b4, b5, b6, b7 with All data
 - infer: 只把预测的 slice 区域 > 12 的联通区域保留. - postive slice
- 蓝色 - SEG Group
 - train: 5 Unets with b4, two b5, b6 and b7 & 2 upaernet with convnext-base and convnext-small backbones. with 只有标注的图片.
 - infer: 在灰色切片判断的基础上, 保留 postive slit

- 黄色 - 3D Group

- train: ...
- infer:

Training & test time augmentation

We tried 640*640 and 512*512 image resize resolution. where in the case of 640*640 resolution, we used RandomCrop method to randomly crop down the target of 448*448 resolution during training. We also used random flip, elastic transformation, grid distortion and optical distortion, the specific parameters are

```
dict(type='RandomFlip', direction='horizontal', p=0.5)
dict(type='ElasticTransform', alpha=120, sigma=6.0, alpha_affine=3.6, p=1)
dict(type='GridDistortion', p=1)
dict(type='OpticalDistortion', distort_limit=2, shift_limit=0.5, p=1)
```

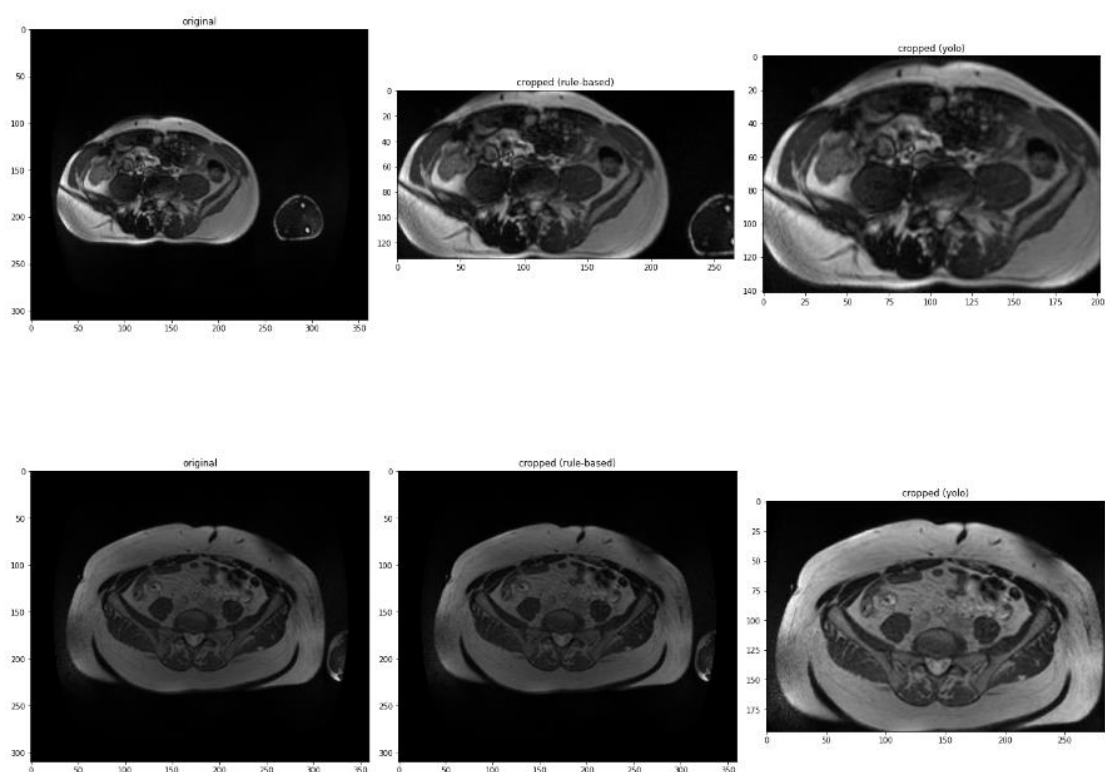
When it comes to test time, we use horizontal flip method to produce a new image and use weighted fusion to get the output mask. TTA can boost our score for 0.001 ~ 0.002 approximately.

整个解决方案亮点：

1. 区域 proposal + 分类 cls 结合模型. 期间采用了 label 有无数数据划分策略.
2. 多模型融合去除置信度低的预测.

2nd place solution

链接：<https://www.kaggle.com/competitions/uw-madison-gi-tract-image-segmentation/discussion/337400>



区域 proposal 提取：yolov5. 由于一些身体运动或者其他因素. 上图 1/2 中的小背景.

去除手臂信号:尤其在腹部 MRI 成像中，由于 RF 场(B1)的不均匀性，手臂周围的信号可能会不自然地高，从而形成热点。当发生这种情况时，图像的最小-最大归一化不能很好地工作，因此手臂被排除在 YOLO 训练的注释阶段之外。

Model

- Backbone
 - stage1 = Efficientnet B4, Swin Base
 - stage2 = Efficientnet L2, ConvNeXt XL, Swin Large
 - For the stage2, increasing the backbone size improves val_dice
- Decoder = UperNet
- CE / CE and Dice Loss (1:1)
- training 20epoch and SWA range(11, 21, 1)

Below are the dice scores for each step in stage2, showing that yolo_crop, large_input, and large_backbone are working.

stage2 (training with gt_positive slices) / Baseline = EfficientNet B4, 256×256

| | Dice @ positive slice |
|---------------|-----------------------|
| Baseline | 0.8011 |
| + YOLOv5 crop | 0.8162 |
| + 512×512 | 0.8277 |
| + backbone=L2 | 0.8349 |

3D model

Model

- stage1 = Unet3d, DynUnet (1000epochs)
- stage2 = Unet3d (500epochs)
- CE and Dice Loss (1:1)
- SWA range(300, 501, 10)

stage2 (training with gt_positive slices)

| | Dice @ positive slice |
|----------------|-----------------------|
| Unet3d + SWA | 0.8271 |
| +flip TTA (HW) | 0.8307 |

Each stage model training

- The stage1 model was trained using all slices of the patient.
- The stage2 model was trained using only the positive slice. That is, for each slice in the train data, `is_positive = np.max(mask) > 0` and only these slices were used.

整个解决方案亮点：

1.2 stage 方案. 先提取整体 propsal. 再进行分割.

3rd Place Solution

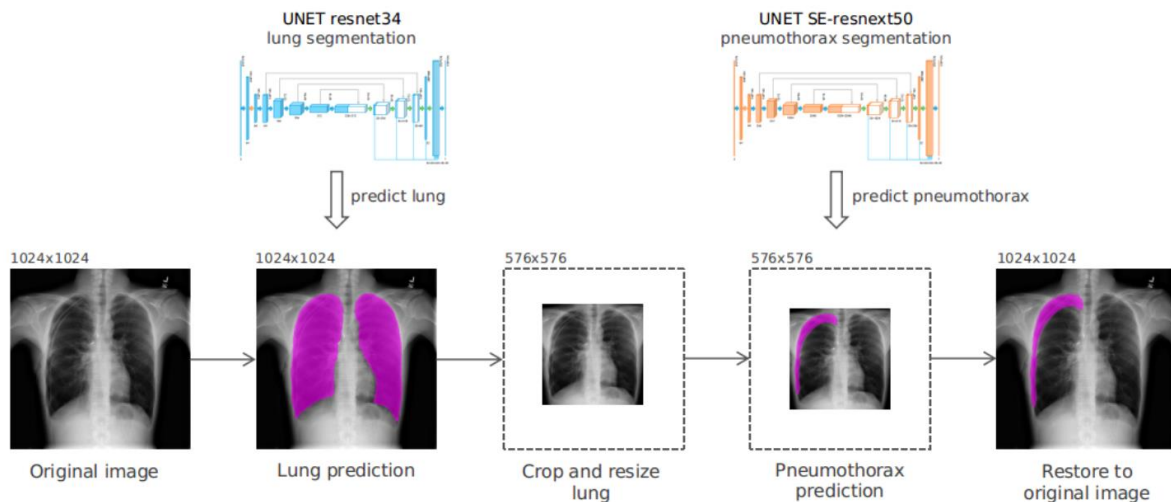
链接：<https://www.kaggle.com/competitions/uw-madison-gi-tract-image-segmentation/discussion/337468>

Work for me

- Add a detection model(It greatly reduces the resources used for training)
- Add some segmentation models for positive sample(0.886→0.890)
- Mixup and cutmix are work for segmentation models
- The swa has a relatively large increase in the single fold, and the score increase is not much different when the 5 folds ensemble.

Not work for me

- Make some brightness adjustments to the data(spent a lot of time)
- Hausdorff distance loss function
- External dataset(CT), maybe need some processing with GAN models
- Positive and negative sample balance



问题收集

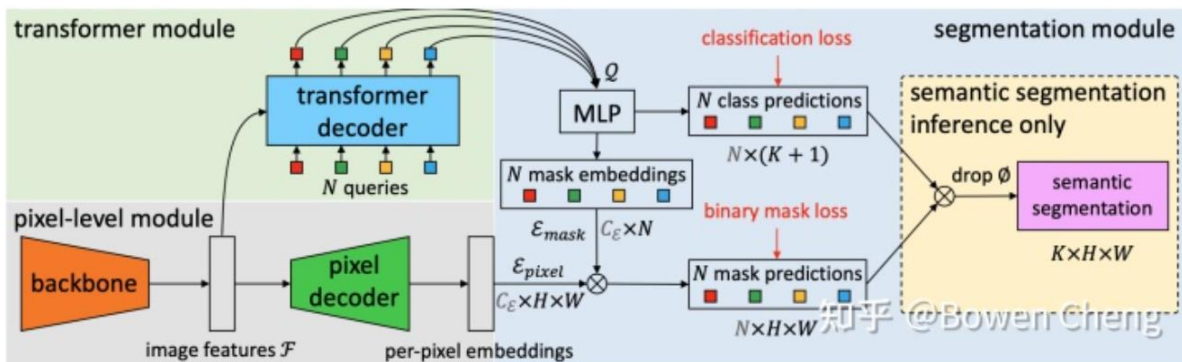
1. 听听 3D 的思路 & 模型还是把图像转换为 3D
2. 对于多器官分割来说，数据前期的处理，模型的学习能力，优化器，损失函数，后处理，各种要素哪个要素是最重要的？要分割的器官种类越多，然后各个因素的作用呢？
3. 此次医学图像是 png 格式，如果是真的医学图像会很大，见过将图像分成小 patch 的，然后再拼接，但是不懂这个拆分和拼接原理，医学图像还是用本身图像会好点，个人医生，所以觉得用 3D 不会丢失信息，希望老师讲下拆分 patch 和合并思路
4. 公开的金牌方案里有哪些值得学习的 idea 以及群里的获奖同学是否可以分享一下自己的经验
5. 可以讲一下 tversky loss 吗？或者可以共享下其他的分割领域的先进的 loss 吗？

A survey of loss functions for semantic segmentation: <https://arxiv.org/pdf/2006.14822.pdf>

TABLE II
TABULAR SUMMARY OF SEMANTIC SEGMENTATION LOSS FUNCTIONS

| Loss Function | Use cases |
|--|--|
| Binary Cross-Entropy | Works best in equal data distribution among classes scenarios Bernoulli distribution based loss function |
| Weighted Cross-Entropy | Widely used with skewed dataset Weights positive examples by β coefficient |
| Balanced Cross-Entropy | Similar to weighted-cross entropy, used widely with skewed dataset weights both positive as well as negative examples by β and $1 - \beta$ respectively |
| Focal Loss | works best with highly-imbalanced dataset down-weight the contribution of easy examples, enabling model to learn hard examples |
| Distance map derived loss penalty term | Variant of Cross-Entropy Used for hard-to-segment boundaries |
| Dice Loss | Inspired from Dice Coefficient, a metric to evaluate segmentation results. As Dice Coefficient is non-convex in nature, it has been modified to make it more tractable. |
| Sensitivity-Specificity Loss | Inspired from Sensitivity and Specificity metrics Used for cases where there is more focus on True Positives. |
| Tversky Loss | Variant of Dice Coefficient Add weight to False positives and False negatives. |
| Focal Tversky Loss | Variant of Tversky loss with focus on hard examples |
| Log-Cosh Dice Loss(ours) | Variant of Dice Loss and inspired regression log-cosh approach for smoothing Variations can be used for skewed dataset |
| Hausdorff Distance loss | Inspired by Hausdorff Distance metric used for evaluation of segmentation Loss tackle the non-convex nature of Distance metric by adding some variations |
| Shape aware loss | Variation of cross-entropy loss by adding a shape based coefficient used in cases of hard-to-segment boundaries. |
| Combo Loss | Combination of Dice Loss and Binary Cross-Entropy used for lightly class imbalanced by leveraging benefits of BCE and Dice Loss |
| Exponential Logarithmic Loss | Combined function of Dice Loss and Binary Cross-Entropy Focuses on less accurately predicted cases |
| Correlation Maximized Structural Similarity Loss | Focuses on Segmentation Structure. Used in cases of structural importance such as medical images. |

6.可以详细讲一下 maskformer 中 transformer decoder 的 query 和 output 是什么东西，有什么意义吗？



$$[K, N] * [N, H*W] \Rightarrow [K, H*W]$$

7.对于医学分割比赛的常用的模型集成融合的方法一般怎么去选择和使用？希望老师可以说下代码实现思路。

8.现在确定的这个比赛哪些数据增强是有效的。

同上

9.常见的医学图像数据 CT 是 3D 的，这个也是时间序列吗，这个好像是围绕身一圈体扫描的，

保存成 nii 格式或者 dcm 格式的，这个是一圈一圈扫描出来的还是同时扫描一圈得到的呢？

IJCAI2022: MNet: Rethinking 2D/3D Networks for Anisotropic Medical Image Segmentation

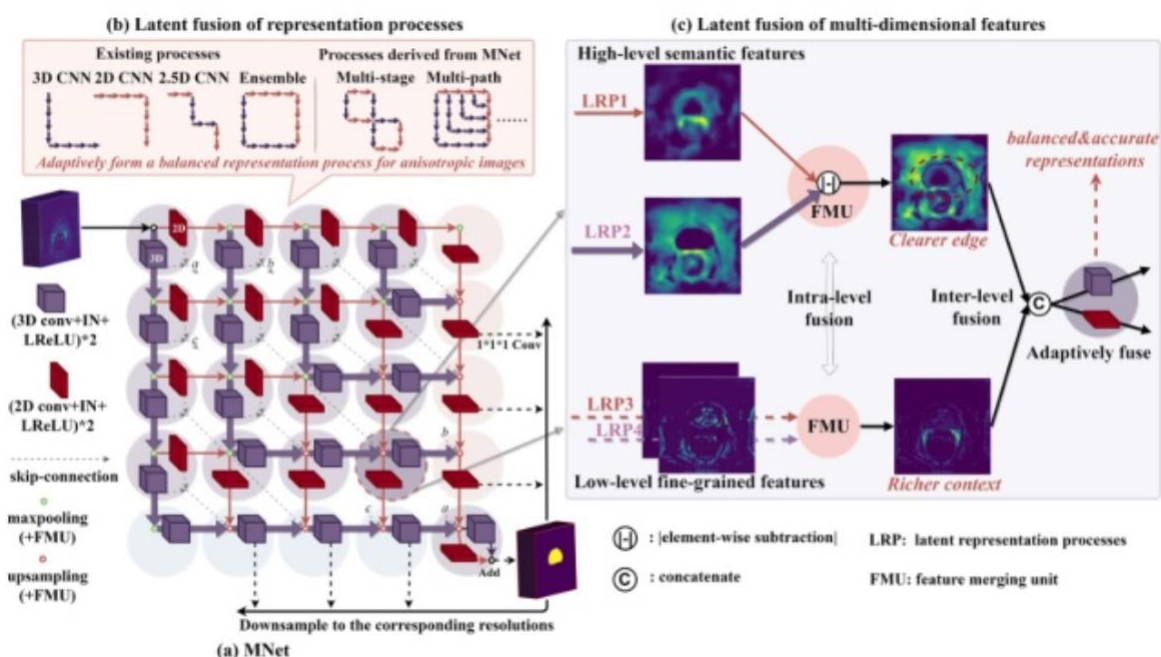
<https://arxiv.org/pdf/2205.04846.pdf>

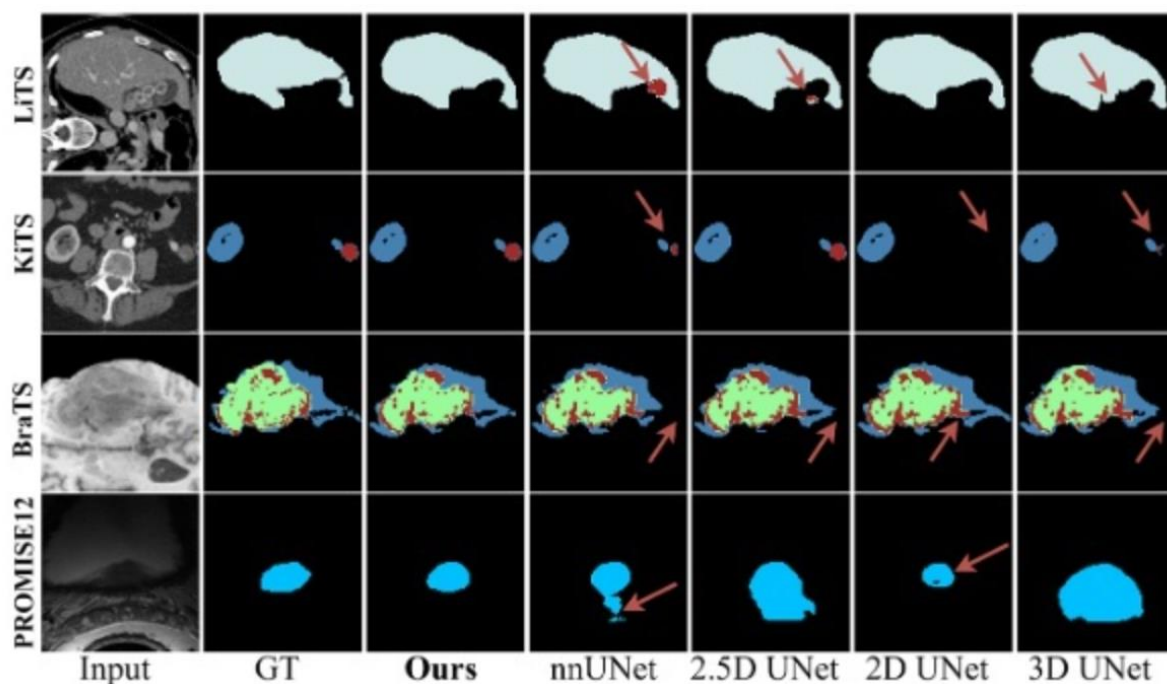
问题：切片扫描的性质导致了三维医学图像的层间不连续性，普通的二维/三维卷积神经网络（CNN）无法平衡地表示稀疏的层间信息和密集(layer)的层内信息，导致切片间的特征拟合不足（对于 2D CNN）和长远距离切片噪声拟合过度（对于 3D CNN）

模型：提出方法：论文提出了一种新的网状网络（MNet），通过学习来平衡不同方向/轴之间的关系。1）MNet 通过将多维卷积深入到基本模块中，实现了大量表示过程的融合，使表示过程的选择更加灵活，从而自适应地平衡稀疏切片间信息和密集切片内信息表示。2）MNet 潜在地融合了每个基本模块内的多维特征，同时利用 2D（2D 视图中容易识别区域的高分割精度）和 3D（3D 器官轮廓的高平滑度）表示的优势，从而获得目标区域更准确的

Ref: IJCAI 2022 | MNET：医学图像分割新模型 - 咚咚的文章 - 知乎

<https://zhuanlan.zhihu.com/p/518195503>

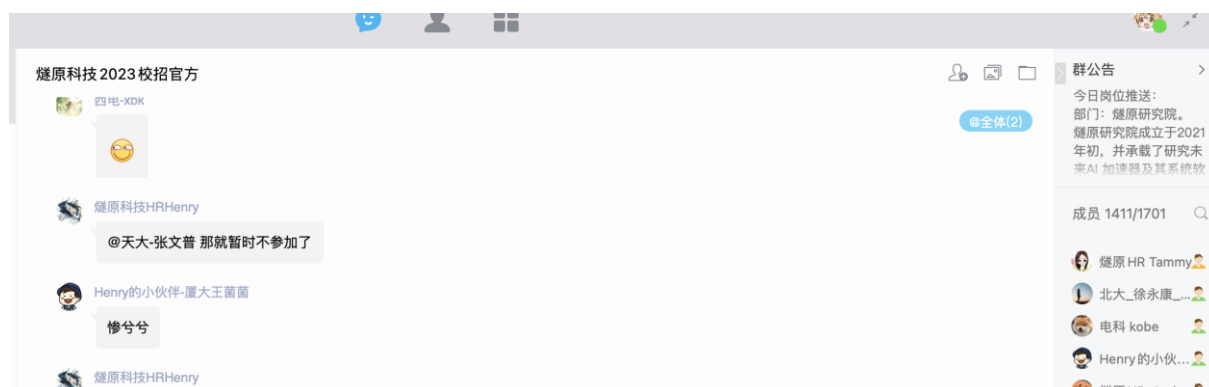




10.各个论文的代码风格不一样，看到新的代码经常就懵了，需要总结一套自己的流程吗，比如预处理总结为一部分，网络一部分等等，这样是否下次看见代码就能往自己的框架里填充呢？代码部分应该怎么提高呢？

11.接下来还有别的比赛吗，想跟着老师继续学习打比赛？

其他



科大讯飞 飞星计划 计算机视觉 二面凉



牛客752414324号

发布于 2022-07-16 12:13:59

APP内打开

赞 4 | 收藏 13 | 回复 20 | 浏览 1662

一面体验感佳，二面体验感差。一面使用科大讯飞自己的软件面试，不能共享屏幕感觉是一大缺点。

一面1个小时，和提前安排的一致。大概自我介绍10分钟，面试官很有条理，自我介绍完就说从四个方面开始面试。1，项目 2，八股 3，工程 4，综合。

项目就主要针对实习项目进行提问，一个是3D检测项目，一个是车道线检测项目，一个是点云语义分割项目。主要聊了聊3D检测项目，以及实习过程中各种项目是怎样负责的。八股除了从项目引申出来的以外，问了相机内参外参（没学过），candy算子等（很久之前学过，面试时忘了，不过讲了一堆图像算法的基本流程，卷积核，梯度，纹理提取啥的，后来知道综合起来也是candy算子了）。工程方面主要问了实习期间遇到的工程问题。综合问家乡，行业规划，平时怎样跟进最新进展等。反问，科大讯飞计算机视觉相关业务，是否在广东有Hc。

二面体验感差，提前约的为pc面试，到时间来了电话面试，首先问成绩如何，考研还是保研，考研多少。开始自我介绍，没有介绍完，介绍了一半左右，介绍到负责的自动驾驶实习项目时就开始提问。（从这开始体验感较差，没有介绍完）。首先详细问了3D检测项目，事先研究挺多回答的很详细。从车道线检测开始就不行了，首先表现为车道线算法中有一个模块只负责拟合作用，并没有参与实例分割部分，一直询问该模块怎样实现实例分割的？？？解释了两三遍没有这个作用，估计从这开始面试官态度也不好了。之后问，这些项目是自己复现修改还是直接源码调参的，回答：实习分配时间很紧，所以都是用源码调参，但是上一个实习，城市级大规模点云语义分割投影优化部分是我自己提出并实现的。问，不了解其他车道线检测算法，3D检测算法。答，我只知道大多数都是以xxx思路实现，但不了解具体其他方案，毕竟不是研究生课题，时间精力也有限，要是确定方向，我会跟进。之后问八股，一周没看八股了，答得不好。问了如何解决正负样本不平衡，loss函数有哪些，项目中用到了哪些函数，nms流程等。最后问实验室课题，答 课题是xxx，已经发表了一篇论文。问，没写在简历上对吧。我？？答，应该写了的吧，我记得写了的，可能忘了。面试官估计也不想问了，直接反问了。反问 反馈时间，广东那边有无hc。

总体感觉，二面体验感较差，八股背的不好，大概就面了半个小时。

面试完就感觉凉了，过了一周多，问hr，已凉

#科大讯飞# #计算机视觉# #提前批#

视觉技术部：

零面（1h）：

1. 讲论文
2. 手写NMS，手写旋转bbox的NMS，面试官鼓励用C++实现。

一面（1.5h）：

1. 讲论文
2. 基础知识：
 - a. 过拟合
 - b. 量化/剪枝
3. Coding: 前缀和的一个题目

二面（2.5h）：

1. 讲论文
2. 讨论Transformer
3. 讨论分割精度提高技巧

秋招以来最长的一次面试，收获也很多，和面试官在很多技术领域上探讨了很多问题。

三面（20min）

1. 讲论文
2. 聊天

四面（40min）

1. 对分割的看法
2. 对方向选择的看法
3. 职业规划

爱奇艺

在生客上面试

一面 (1.5h) :

1. 讲论文
2. 交叉熵公式 (这题我被问了四五遍了)
3. Coding: 模拟三体运动, 给定3个物体的质量、初始位置、速度和模拟的时间间隔 T , 假定 T 时间内都进行匀速直线运动, 模拟段时间内的物体运动轨迹。

二面 (1h) :

1. 讲项目细节
2. 有没有模型加速的经验

三面 (20min) :

1. 主要是讲项目
2. 简单聊一下就结束了

四面 (30min) :

1. 自我介绍
2. 项目经历
3. 超分辨率的研究内容
4. 分割的研究内容
5. 分割面临的问题, 前沿的发展方向
6. 模型速度优化
7. 职业规划

地平线

腾讯会议面试。

一面 (1.5h) :

1. 简单聊了下项目
2. 知不知道Clip这个工作
3. 介绍语义分割的发展
4. 介绍 Deeplab v3+, 这个模型存在什么问题, 你会怎么改进?
5. 超分和语义分割如何结合?
6. Pooling的作用: 1) 提高感受野 2) 降低计算成本 3) 多尺度融合
7. Pooling反向传播的过程
8. Dropout的原理以及反向传播过程
9. 神经网络怎么初始化好, 各种初始化方法有什么区别?
10. 为什么分类会使用softmax, 和CE Loss在一起用有什么好处?
11. 神经网络训练的时候出现了很多神经元都为0的情况是什么原因?
12. 做题, 最小代价路径, 一个二维dp
13. 反问

二面 (2h) :

1. 项目
2. 论文
3. 度量学习中难样本挖掘策略
4. 开放问题: 分割的边缘如何更精确?
5. 开放问题: 标注标签如果比较粗糙, 有什么办法可以训练精细的分割结果?
6. 开放问题: 无类别的多前景分割怎么做?
7. 开放问题: 无有考虑自监督的方法做分割吗?
8. Coding: 分割的miou, 有n张尺寸为h*w的图片, 已知类别为c, gt和pred的labelmap, 求这批数据的miou
9. 反问

三面 (1h)

1. 主要还是聊了下论文, 问了我做的技术在自动驾驶里面有没有可能的应用
2. 面试官介绍了组里面的做的事情

中兴

腾讯会议上面试，每一面都有好多面试官在，n对1。

一面（1h）（两三个面试官）：

1. 自我介绍
2. 项目+论文+实习经历

二面（40min）：

1. 自我介绍
2. 项目+论文+实习经历
3. 讲在审的论文
4. 图像分割的发展方向