



字节跳动计算机视觉算法实习生视频面试



HelloWorld

关注他

710 人赞同了该文章

此专栏主要总结各大厂机器学习相关的面试经验，欢迎大家投稿分享面试经验

最近有朋友问我如何准备视觉算法的面试，恰好看到一篇很不错的文章（感谢@辛俊波写的这篇文章），有相关疑惑的朋友可以看一下

如何看待 2020 届校招算法工程师岗位求职人数远大于招聘岗位的现象？

524 赞同 · 30 评论 · 回答

岗位信息如下：

研发

计算机视觉算法实习生

北京 暑期实习

岗位描述：

- 1、利用计算机视觉和人工智能技术，改进头条、抖音等产品的创作和观看体验；
- 2、参与计算机视觉和人工智能等领域的前沿技术研究；
- 3、参与和推动技术落地到头条产品；
- 4、参与研发以计算机视觉和人工智能技术为核心的新产品。

岗位要求：

- 1、2020年取得本科/硕士/博士学位，计算机相关专业；
- 2、有以下领域之一的技术积累：
 - a) 计算机视觉：图像视频分类、视频内容分析、文字检测和识别、场景理解、人脸识别、物体追踪等
 - b) 计算机图形学、增强现实
 - c) 人机交互
 - d) 机器学习
- 3、有一定的工程实现的能力，能够将算法落地到实际的产品中；
- 4、作为主要作者发表过顶级会议论文者优先；
- 5、具备较强的算法与编程能力，ACMICPC, NOI/IOI, Top coder, kaggle比赛获奖者优先；
- 6、沟通合作能力良好，与团队融洽相处。

知乎 @JustDoIT

一面（时间2019.3.14，形式是视频面试） -- 已凉

下午四点面试，晚上七点半收到了凉凉的邮件通知，不慌的，再接再厉，查漏补缺，下面我就详细总结一下面试的经过。

注意：回答不一定正确，仅提供了我个人的回答，答案里有些不正确的地方，我之所以给出我的个人回答只是为了更好地还原面试过程，最近如有时间，我会整理出正确的答案。



致谢：感谢各位朋友和各位大佬给出的宝贵意见

总结：根据各位朋友和各位大佬给出的宝贵意见，我反思总结如下：

- python 和 c++ 基本语法不够扎实，需强化基本语法点

例如对python里面的集合set不太熟悉

- 回答过于简洁，难免给面试官基础知识不扎实的错觉，需讲得更详细
- 深度学习的基础知识不扎实，不系统化，需构建知识体系

面试建议：面试官主要是对着简历进行提问的，然后再扩展一下相关的知识点。面试主要考察了项目、比赛以及一些理论知识和编程语言的语法点。建议大家可以看着岗位信息进行简历的制作并且在简历上出现的内容需要好好准备一下。

下面就开始还原面试过程

1、简单自我介绍

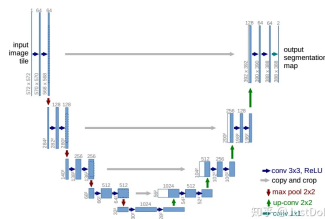
2、kaggle "TGS Salt Identification Challenge"模型的准确度是怎么算的？

我说ap，然后面试官问有没有用miou，我说没有

[TGS Salt Identification Challenge](#)

3、说一下unet的结构

卷积，下采样，上采样，然后低层特征图和高层特征进行融合，最后进行分类



4、熟悉deeplab吗，aspp是怎样的

不太熟悉，目前正在阅读相应的论文和源码，我目前是对FCN，unet，refinenet，mask rcnn比较熟悉

5、说说 智能盘点—钢筋数量AI识别 这个比赛

这个比赛是实现拍照即可完成钢筋点根任务，先使用了YOLO，然后用了faster rcnn还有retinanet

[智能盘点-钢筋数量AI识别 - DF,CCF大数据竞赛平台](#)

6、retinanet的focal loss是解决的什么问题

解决类别不平衡问题，对困难样本的损失分配比较大的权重

既能调整正负样本的权重，又能控制难易分类样本的权重

7、那focal loss有解决背景样本和目标样本的问题吗

我说应该有

8、谈一下比赛中提到k-means算法

我说yolo是根据iou来进行距离度量的

[github.com/AlexeyAB/dar...](https://github.com/AlexeyAB/darknet)

9、那k-means是怎么实现的，k近邻算法呢

这个我用的是官方源码

估计面试官想问的是这个原理，[K-Means聚类算法原理 - 刘建平Pinard - 博客园](#)

10、opencv做了哪些数据增强

直方图，翻转，随机裁剪，高斯模糊，高斯噪声

11、怎么做的模型融合

如果是同个模型，会整合两个模型的检测结果，根据score来排序再做nms操作。

但这里我是用的不同模型，用的yolo和faster rcnn，他们的score没有太大的可比性，所以我没有用score来排序，我是求两个模型的交集，以解决误检的问题

12、为什么用交集，不用并集

其实这个我是根据提交结果来决定的，事实上，在这里交集的效果比并集好，交集解决误检问题，并集解决漏检问题

13、安检机物品检测的图片格式是怎样的

x光扫描的RGB图

14、为什么选用yolo而不用unet

比较熟悉yolo

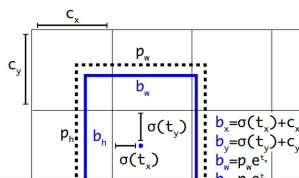
15、yolo是怎么分配检测框的

每个grid分配9个检测框

yolo会输出3种尺寸的特征图，13*13，26*26，52*52，每种尺寸特征图的每个grid预测3个框

16、yolo的预测框是什么值 (x, y, w, h)

是相对grid的相对值



我说欧式距离

实际是smooth l1

The classification loss L_{cls} is log loss over two classes (object vs. not object). For the regression loss, we use $L_{reg}(t_i, t_i^*) = R(t_i - t_i^*)$ where R is the robust loss function (smooth L_1) defined in [2]. The term $p_i^{L_{reg}}$

18、熟悉smooth L1吗

描述了一下函数形式

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise,} \end{cases}$$

19、说一下nms的操作

对每一类进行nms，先根据score进行降序排序，然后计算最高的score和其他框的iou，去掉iou大于阈值的检测框

github.com/rbgirshick/p...

20、那假如一个类有1000个框，怎么计算iou和去掉大于阈值的框

先拿score最大的检测框和其余的检测框进行iou操作，这里用到了numpy数组的广播机制，然后用numpy的筛选操作，把大于一定阈值的检测框的score置0，然后下一次迭代从score不为0的检测框进行上述操作，直到操作完成

21、熟悉哪些基础网络

熟悉VGG、GoogleNet还有resnet这几个提取特征的，还有FPN这个来融合信息的

22、熟悉一些新的网络吗

不太熟悉，最近没关注，最近在研究代码

估计面试官想问的是这些网络Xception、DenseNet、DetNet，在文章开头的链接里汇总了

23、激活函数有哪些

sigmoid还有relu

还有tanh等
[Activation function](#)

24、sigmoid存在什么问题

梯度消失

25、relu的表达式

我说成leaky relu的了，然后面试官说确定吗，不是relu的变形吗，

我说我说错了，然后我纠正了一下

leaky relu

```
Vectorize (out)
#leaky ReLU (out)
#leaky ReLU can be extended to include Gaussian noise, making from noisy ReLU, noisy?
f(x) = max(0, x) + alpha * min(0, -x)
Note that it has been used with some success in restricted Boltzmann machines for computer vision tasks!

#leaky ReLU (out)
Leaky ReLU allows a small positive gradient when the unit is not active!
f(x) = { x, if x > 0;
        alpha * x, otherwise

#leaky ReLU (out)
Parameter ReLU (beta) can also further by making the coefficient of leakage into a parameter that is learned along with the other neural network parameters!
f(x) = { x, if x > 0;
        beta * x, otherwise
Note that for alpha < 1, this is equivalent to
f(x) = max(0, x)
and thus has a relation to "leaky" networks!

ReLU (out)
Experimented ReLU with 10 to make the mean activations closer to zero which speeds up learning. It has been shown that ReLU can obtain higher classification accuracy than ReLU.
f(x) = { x, if x > 0;
        0, otherwise
alpha is a hyper parameter to be learned and alpha < 1, 0 is a constant.
```

知乎 @ JustDoIT

[如何理解ReLU activation function?](#)[请问人工神经网络中的activation function的作用具体是什么? 为什么ReLU要好过于tanh和sigmoid function?](#)

27、熟悉normalization吗

比较熟悉bn和gn

[张俊林：深度学习中的Normalization模型](#)

28、说说bn

bn是解决梯度消失问题的，通过bn能使数据在输入到激活函数之前回到导数较大的位置

#解决问题

解决内部数据分布不稳定问题

#做法

统计mini-batch个样本的期望和方差从而进行归一化，但这样会导致数据表达能力的缺失，所以要引入两个参数从而恢复数据本身的表达能力

#优势

使得网络中每层输入数据的分布相对稳定，加速模型学习速度

允许网络使用饱和性激活函数，缓解梯度消失问题

具有一定的正则化效果

29、bn是做了归一化吗

先做了归一化，例如先减去期望再除以方差，然后再乘以一个参数和加上一个参数，这两个参数是可学习的

30、bn是在哪一维进行操作的

在mini-batch和通道这两维，描述了一下具体是怎么算的。我举了个例子，四维tensor，分别是[n,c,h,w]，然后bn是统计每张特征图的期望和方差，而这个期望和方差是n个样本同一张特征图的期望和方差

31、那gn有什么用

gn是解决bn对mini-batch过度依赖，gn是在通道处进行分组统计，不依赖mini-batch

32、bn和gn都在哪用

bn和gn我所了解的是在目标检测，当然图像分类也用到bn

33、c++和python是都熟悉吗

是的

35、那引用能重新赋值吗

不能

36、静态变量有什么用

能不在实例化类的情况下进行调用，而且每个实例化后的对象共享这个值

37、静态变量在哪初始化，能在类内初始化吗

在类内类外都行

一般是在类外初始化，如需在类内初始化，要求该静态变量为静态字面值常量类型 [C++ 静态成员的类内初始化 - WanChouchou - 博客园](#)

38、静态函数有什么用

能在不实例化类的情况下进行调用

39、python有什么常用的数据类型

列表、无序字典和有序字典

列表、元组、字典、集合等

40、对多线程了解吗

了解，但没写过

41、那如果让你自己实现pytorch里面的dataloader，你怎么可以使它加载快点

用多线程

42、python的append和extend有什么区别

append是添加单个元素，extend是添加一个列表

43、出一道算法题

算法题如下：给个有序数组，然后求元素平方后不重复的元素个数，例如[-10, -10, -5, 0, 1, 5, 8, 10]

我思想描述对了，然后面试官说有更好的方法吗，我想了一下说没有，然后面试官让我选个语言实现一下，选择了python来实现，用到了字典，然后面试官说用集合会不会更好，我说会的

```
# 这是我的实现，没有用到集合，如果用到集合会更好
data = [-10, -10, -5, 0, 1, 5, 8, 10]
new_list = []
for x in data:
    temp = x*x
    if temp not in new_list:
        new_list.append(temp)
print(len(new_list))

# 其他实现
print(len(set([x**2 for x in data])))
```

我说看岗位信息提到会进行技术的理论研究和技术的落地，如果我去了会负责什么，面试官说因为实习时间比较短，一般是做些技术落地的项目

45、那有师傅带吗，那一般做什么项目

有师傅带，一般做力所能及的项目

参考：

1 U-Net: Convolutional Networks for Biomedical Image Segmentation

2 Focal Loss for Dense Object Detection

3 YOLOv3: An Incremental Improvement

4 Fast R-CNN

5 Faster R-CNN Towards Real-Time Object Detection with Region Proposal Networks

6 [AlexeyAB/darknet](#)

7 [rbgirshick/py-faster-rcnn](#)

8 [Rectifier \(neural networks\)](#)

编辑于 2020-04-11 18:39

[机器学习](#) [计算机视觉](#) [字节跳动 \(ByteDance\)](#)



评论千万条，友善第一条

178 条评论

默认 最新



学控制的雀神健

看你的面经，觉得自己水的一批。实习无望，23333333333333

2019-03-14

回复 47



于归

append和extend那边你好像弄错了，
[1, 2].append([3, 4])=》 [1, 2, [3, 4]]
[1, 2].extend([3, 4])=》 [1, 2, 3, 4]

看了下你的回答，可能感觉你基础掌握的不牢所以把你刷了吧！

2019-03-14

回复 13



HelloWorld 作者

有可能，我仍需更努力，这里面的答案仅我个人所答，我会整理一下的，谢谢指出哈

2019-03-14

回复 3



不懂就多看看

我也是字节头条计算机视觉（AI lab）实习岗，虽然我撑到了三面，结果还是挂了，每轮都要手撕代码，我感觉我挂在代码上，我是不是也应该写个面经给大家

2019-03-20

回复 5



很不错的学生

...

 鹿鹿

好巧呀，我也是今天（14号晚）面字节跳动的计算机视觉实习，因为是美东时间，所以比你晚了半天。看完你的面经，我已经开始佛系了

2019-03-14

● 回复 5



HelloWorld 作者

大佬肯定过的

2019-03-14

● 回复 赞



Deeppower

这么强。。都凉了，字节跳动真恐怖。。。

2019-03-14

● 回复 5



HelloWorld 作者

我要思考人生了

2019-03-14

● 回复 4



coreajax

他代码基础好像不行。

2019-03-20

● 回复 2



Serenade

直接po面试题是非常不负责任的行为 建议把面试题修改成几个大的方向 既能给后人提示 又不损害公司利益

2019-03-16

● 回复 4



zchs

我给这个评论点了举报，兄弟们我做的对吗？

2021-07-30

● 回复 15



麦哲伦

这种面试是根据每个人的项目经验来的，换个候选人或者换一个面试官，题目就完全不一样了。如果是发的统一笔试题确实过分，但是答主这种是不会泄露什么公司机密吧

2019-03-31

● 回复 10

 鹿鹿

补充一下刚刚介绍的字节跳动计算机视觉的面试经历，供大家参考。
自己没有编程基础（生物医学工程专业），意料之中的挂在了coding test上。

1 自我介绍

2 挨个项目问用到的算法

项目的难点

BOW(Bag of Word)的过程，建立与分类原理

为什么BOW会用到spatial pyramid pooling，怎么使用的

为什么interest points matching 会用到RANSAC，怎么使用的

图像分割的传统方法，医学图像分割的特殊方法

问了项目中用到的医学图像active contour的原理

有其他用machine learning 处理图像的经验吗

3 coding test, 限时15mins

segmentation label connection 的问题：

输出一个二维list的连通区域的个数（八连通），和最大区域的面积。

eg:

inputs:

00000110001,

11000110000,

11000000010,

00000111100,

00000011000,

00000010000,

outputs:

4, 8

自己是思路对了，bug没调出来，就挂了

2019-03-14

● 回复 10



某湘

这么强都凉了，咱凉透了



吃口鱼冷静一下

2019-03-14

● 回复 9



HelloWorld 作者

不一定，好好加油，做好准备。我仍需努力，我离上岗就业还有点距离

2019-03-14

● 回复 6



timm

该条评论被折叠

2019-03-14



HelloWorld 作者

嗯好的，不过其实那些问题我都清楚的，只不过我简单回答一下，我还以为面试官会问深点，结果没有问很深。总结：我仍需努力，哈哈，不慌不慌

2019-03-14

● 回复 3



余努力

43题，两头扫描：先从绝对值大的那头扫，记录不重复的个数，绝对值小于另一头以后再从另一头扫描。不用多余的空间可解

2019-05-04

● 回复 2



RICHARD

看了答主的回答，补充了很多不会的知识👍

2019-03-14

● 回复 2



HelloWorld 作者

我也是在查漏补缺😁

2019-03-14

● 回复 赞



wiley

年前面的.....也凉了

2019-03-14

● 回复 2



HelloWorld 作者

同凉😓

2019-03-14

● 回复 3



小鹏

跨专业，目前研一，希望十二个月后有你这么牛就好啦~

2019-03-14

● 回复 2



HelloWorld 作者

别笑我了，评论区的各位才是真牛人

2019-03-14

● 回复 2



HelloWorld 作者

加油，一定可以的



● 回复 2

以

2019-05-18

 回复  1**taotaotao**

那道算法题是关键，你python写的太不pythonic了

2019-05-08

 回复  1**HelloWorld** 作者

还是吃了代码写少了的亏

2019-05-08

 回复  赞**lambda**

欢迎大家找我内推头条 ailib? 🤔

2019-05-08

 回复  1**HelloWorld** 作者

优秀的优秀的

2019-05-08

 回复  赞**sanada**

我也投的头条的ailab cv实习，今天刚hr面完，感觉你面的比我难很多啊😂

2019-04-30

 回复  1**HelloWorld** 作者  **sanada**

大佬谦虚了，我连hr的声音都没听到哈哈

2019-04-30

 回复  赞**sanada**  **HelloWorld**

不不不，大佬你肯定没问题的😂

2019-04-30

 回复  赞[展开其他 1 条回复 >](#)**jiazhuamh** 

大佬这么深的功力都凉，吓得我都不敢投简历了.....

2019-04-12

 回复  1**v1cfan**

想起我两个同学靠过硬的工程能力一个去了旷视一个去了MSRA，他们的算法水平倒是不如楼主，感觉还是需要落地吧，工程能力强的话读完论文马上能做出模型也是一种优势

2019-04-09

 回复  1**HelloWorld** 作者

看完论文能复现，并且性能和原作者差不多，那很优秀

2019-04-09

 回复  2[查看被折叠评论 2](#)[点击查看全部评论 >](#)

评论千万条，友善第一条

**机器学习面经**

欢迎大家投稿分享面试经验

推荐阅读



字节跳动计算机视觉算法实习内推

BigBoss

2022计算机视觉算法工程师面经-字节、商汤、海康、快手

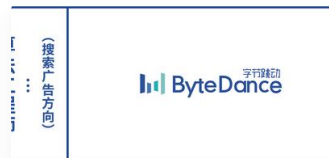
先说下本人情况： - 本科北交大，保研到浙江大学电子信息，科班，目前研一 - 一篇很水的顶会，一篇ccfb，两个都是共一，CV方向的(医学图像) - 一篇北京市优秀本科毕业论文+一段小公司3个月...

zerone



字节跳动-ailab-视觉算法面经(2020届)

AI算法与... 发表于算法岗面经...



AI面试锦囊 | 字节算法工程师(搜索广告方向)三面分享

AI面试指南