

算法岗计算机视觉方向求职经验总结

进入 11 月份，楼主找工作也基本进入尾声了，从 7 月份开始关注牛客网，在求职的过程中学到了很多，感谢牛客提供这样一个平台，让自己的求职历程不再孤单。

先说一下楼主教育背景，本科西部末流 985，研究生调剂到帝都某文科大学。专业都是 CS 专业，求职方向都是计算机视觉算法。有某外企以及二线互联网实习经历，本科虽然 CS 出身，但实际动手能力并不强。研究生的研究方向并不是计算机视觉方向。实习的时候开始接触计算机视觉，自己比较感兴趣，开始转 CV 方向。回想这几个月的求职经历，其中的辛苦只有自己知道。最终拿到了百度 SP，京东 SSP，美团无人驾驶 SP，顺丰科技 SP，拼多多 SP，以及虹软 SP，思科，中电 29 等 offer。

想把我学习与求职路上的一些心得告诉学弟学妹们。

1.一定要有一门自己比较熟悉的语言。

我由于使用 C++ 比较多，所以简历上只写了 C++。C++ 的特性要了解，C++11 要了解一些，还有 STL。面试中常遇到的一些问题，手写代码实现一个 string 类，手写代码实现智能指针类，以及 STL 中的容器的实现机制，多态和继承，构造函数，析构函数等。推荐看一下[网易云课堂翁恺老师](#)的 C++ 的视频以及经典的几本书。

2.一定要刷题

楼主主要刷了[剑指 offer](#)以及 leetcode 上的 easy,middle 的题目。如果编程能力不是很强，推荐可以分类型进行刷题，按照 tag 来刷，对于某一类型的题目，可以先看一下该算法的核心思想，然后再刷题。楼主在求职的过程中，遇到好多跟 leetcode 上类似的题目，刷题的目的不是为了碰见原题，而是为了熟练算法。当然能够碰见原题最好不过啦。

3.机器学习的一些建议

推荐西瓜书，以及李航老师的统计学方法。另外熟悉一种深度学习框架。学习计算机，一定要实战，毕竟只有在实战的过程中，才能懂得更透彻。可以多参加一些比赛，比如 [kaggle](#),[天池](#)，[滴滴的一些比赛](#)。这对找工作的用处很大。

4.能实习就尽量实习。

如果导师是学术大牛，可以带你发顶会的论文，并且自己对方向比较感兴趣，那可以在实验室待着好好搞科研。如果你研究生的研究方向跟你以后的求职方向不一致，建议早点出来实习，找个对口的实习，实习才能发现，实际工作和在学校学习的东西差距比较大。

楼主能不能分享下面试问了哪些视觉的问题啊

问到的问题主要跟我实习做的东西有关，有关于视频拆分的一些算法，以及三维点云的一些问题，传统的图像处理的一些基本操作还是要了解的，比如滤波，边缘检测，以及常用的一些传统的特征，SIFT,SURF，HOG 等。深度学习这方面，我主要做过目标检测，所以问到的就是 rcnn,fast-rcnn,faster-rcnn,yolo,ssd 这些算法。另外，问过一些调参，正则化，Batch normalization,drop out,激活函数的选择。手动推导 BP,LR,SVM，算法题主要有常规的排序，二分查找，BP 相关的题目，还有一些就是关于二叉树的递归和非递归遍历，层次遍历，最近公共祖先等，其余的题目记得不太清楚了

深度学习面经

(1) 代码题 (leetcode 类型)，主要考察数据结构和基础算法，以及代码基本功

虽然这部分跟机器学习，深度学习关系不大，但也是面试的重中之重。基本每家公司的面试都问了大量的算法题和代码题，即使是商汤、face++这样的深度学习公司，考察这部分的时间也占到了我很多轮面试的 60%甚至 70%以上。我去 face++ 面试的时候，面试官是 residual net，shuffle net 的作者；但他们的面试中，写代码题依旧是主要的部分。

大部分题目都不难，基本是 leetcode medium 的难度。但是要求在现场白板编程，思路要流畅，能做到一次性 Bug-free。并且，一般都是要给出时间复杂度和空间复杂度最优的做法。对于少数难度很大的题，也不要慌张。一般也不会一点思路

也没有，尽力给面试官展现自己的思考过程。面试官也会引导你，给一点小提示，沿着提示把题目慢慢做出来也是可以通过面试的。

以下是我所遇到的一些需要当场写出完整代码的题目：

<1> 二分查找。分别实现 C++ 中的 `lower_bound` 和 `upper_bound`。

<2> 排序。手写快速排序，归并排序，堆排序都被问到过。

<3> 给你一个数组，求这个数组的最大子段积

时间复杂度可以到 $O(n)$

<4> 给你一个数组，在这个数组中找出不重合的两段，让这两段的字段和的差的绝对值最大。

时间复杂度可以到 $O(n)$

<5> 给你一个数组，求一个 k 值，使得前 k 个数的方差 + 后面 $n-k$ 个数的方差最小

时间复杂度可以到 $O(n)$

<6> 给你一个只由 0 和 1 组成的字符串，找一个最长的子串，要求这个子串里面 0 和 1 的数目相等。

时间复杂度可以到 $O(n)$

<7> 给你一个数组以及一个数 K ，从这个数组里面选择三个数，使得三个数的和小于等于 K ，问有多少种选择的方法？

时间复杂度可以到 $O(n^2)$

<8> 给你一个只由 0 和 1 组成的矩阵，找出一个最大的子矩阵，要求这个子矩阵是方阵，并且这个子矩阵的所有元素为 1

时间复杂度可以到 $O(n^2)$

<9> 求一个字符串的最长回文子串

时间复杂度可以到 $O(n)$ (Manacher 算法)

<10> 在一个数轴上移动，初始在 0 点，现在要到给定的某一个 x 点，每一步有三种选择，坐标加 1，坐标减 1，坐标乘以 2，请问最少需要多少步从 0 点到 x 点。

<11> 给你一个集合，输出这个集合的所有子集。

<12> 给你一个长度为 n 的数组，以及一个 k 值 ($k < n$) 求出这个数组中每 k 个相邻元素里面的最大值。其实也就是一个一维的 max pooling

时间复杂度可以到 $O(n)$

<13> 写一个程序，在单位球面上随机取点，也就是说保证随机取到的点是均匀的。

<14> 给你一个长度为 n 的字符串 s ，以及 m 个短串（每个短串的长度小于 10），每个字符串都是基因序列，也就是说只含有 A,T,C,G 这四个字母。在字符串中找出所有可以和任何一个短串模糊匹配的子串。模糊匹配的定义，两个字符串长度相等，并且至多有两个字符不一样，那么我们就可以说这两个字符串是模糊匹配的。

<15> 其它一些描述很复杂的题这里就不列了。

（2）数学题或者"智力"题。

不会涉及特别高深的数学知识，一般就是工科数学（微积分，概率论，线性代数）和一些组合数学的问题。

下面是我在面试中被问到过的问题：

<1> 如果一个女生说她集齐了十二个星座的前男友，她前男友数量的期望是多少？

ps: 这道题在知乎上有广泛的讨论，作为知乎重度用户我也看到过。[如果一个女生说，她集齐了十二个星座的前男友，我们应该如何估计她前男友的数量？](#)

<2> 两个人玩游戏。有 n 堆石头，每堆分别有 $a_1, a_2, a_3, \dots, a_n$ 个石头，每次一个游戏者可以从任意一堆石头里拿走至少一个石头，也可以整堆拿走，但不能从多堆石头里面拿。无法拿石头的游戏者输，请问这个游戏是否有先手必胜或者后手必胜的策略？如果有，请说出这个策略，并证明这个策略能保证必胜。

<3> 一个一维数轴，起始点在原点。每次向左或者向右走一步，概率都是 0.5。请问回到原点的步数期望是多少？

<4> 一条长度为 1 的线段，随机剪两刀，求有一根大于 0.5 的概率。

<5> 讲一下你理解的矩阵的秩。低秩矩阵有什么特点？在图像处理领域，这些特点有什么应用？

<6> 讲一下你理解的特征值和特征向量。

<7> 为什么负梯度方向是使函数值下降最快的方向？简单数学推导一下

（3）机器学习基础

这部分建议参考周志华老师的《机器学习》。

下面是在面试中被问到过的问题：

<1> 逻辑回归和线性回归对比有什么优点？

<2> 逻辑回归可以处理非线性问题吗？

<3> 分类问题有哪些评价指标？每种适用场景。

<4> 讲一下正则化，L1 和 L2 正则化各自的特点和适用场景。

<5> 讲一下常用的损失函数以及各自的适用场景。

<6> 讲一下决策树和随机森林

<7> 讲一下 GBDT 的细节，写出 GBDT 的目标函数。GBDT 和 Adaboost 的区别与联系

<8> 手推 softmax loss 公式

<9> 讲一下 SVM, SVM 与 LR 有什么联系。

<10>讲一下 PCA 的步骤。PCA 和 SVD 的区别和联系

<11> 讲一下 ensemble

<12> 偏差和方差的区别。ensemble 的方法中哪些是降低偏差，哪些是降低方差？

..... 这部分问得太琐碎了，我能记起来的问题就这么多了。我的感觉，这部分问题大多数不是问得很深，所以不至于被问得哑口无言，总有得扯；但是要想给出一个特别深刻的回答，还是需要对机器学习的基础算法了解比较透彻。

（4）深度学习基础

这部分的准备，我推荐花书（Bengio 的 Deep learning）和[@魏秀参](#) 学长的[《解析卷积神经网络-深度学习实践手册》](#)

下面是在面试中被问到过的问题：

<1> 手推 BP

<2> 手推 RNN 和 LSTM 结构

<3> LSTM 中每个 gate 的作用是什么，为什么跟 RNN 比起来，LSTM 可以防止梯度消失

<4> 讲一下 pooling 的作用，为什么 max pooling 要更常用？哪些情况下，average pooling 比 max pooling 更合适？

<5> 梯度消失和梯度爆炸的原因是什么？ 有哪些解决方法？

<6> CNN 和 RNN 的梯度消失是一样的吗？

<6> 有哪些防止过拟合的方法？

<7> 讲一下激活函数 sigmoid, tanh, relu. 各自的优点和适用场景？

<8> relu 的负半轴导数都是 0，这部分产生的梯度消失怎么办？

<9> batch size 对收敛速度的影响。

<10> 讲一下 batch normalization

<11> CNN 做卷积运算的复杂度。如果一个 CNN 网络的输入 channel 数目和卷积核数目都减半，总的计算量变为原来的多少？

<12> 讲一下 AlexNet 的具体结构，每层的作用

<13> 讲一下你怎么理解 dropout，分别从 bagging 和正则化的角度

<14> data augmentation 有哪些技巧？

<15> 讲一下你了解的优化方法，sgd,momentum,rmsprop, adam 的区别和联系

<16> 如果训练的神经网络不收敛，可能有哪些原因？

<17> 说一下你理解的卷积核， 1x1 的卷积核有什么作用？

.....

同上，这部分的很多问题也是每个人都或多或少能回答一点，但要答得很好还是需要功底的。

（5）科研上的开放性问题

这部分的问题没有固定答案，也没法很好地针对性准备。功在平时，多读 paper 多思考，注意培养自己的 insight 和 intuition

下面是我在面试中被问到过的问题：

<1> 选一个计算机视觉、深度学习、机器学习的子领域，讲一下这个领域的发展脉络，重点讲出各种新方法提出时的 motivation，以及谈谈这个领域以后会怎么发展。

<2> 讲一下你最近看的印象比较深的 paper

<3> 讲一下经典的几种网络结构， AlexNet， VGG， GoogleNet， Residual Net 等等，它们各自最重要的 contribution

<4> 你看过最近很火的 XXX paper 吗？你对这个有什么看法？

.....

（6）编程语言、操作系统等方面的一些问题。

C++， Python， 操作系统， Linux 命令等等。这部分问得比较少，但还是有的，不具体列了

（7）针对简历里项目/论文 / 实习的一些问题。

这部分因人而异，我个人的对大家也没参考价值，也不列了。

作者：牛客网

链接：<https://zhuanlan.zhihu.com/p/30212122>

来源：知乎

著作权归作者所有。商业转载请联系作者获得授权，非商业转载请注明出处。

楼主本科，今年前前后后面了一些 BAT 的实习和校招，岗位主要是基础研究，机器学习方面。最后去 T 的技术研究岗。并不是大神，只是在这里总结一下自己的看法和面经。因为面试题主要是看面试官，并不是看公司，所以这里就不分开写了。

这一类的岗位面试大多数是以简历为基础，会根据简历上所提到的一些工具(比如 `caffe`, `cuda`)，算法(比如 `CNN`, 随机森林)和项目进行提问。所以准备面试最重要的就是要深刻的理解这些工具和算法的基础和优劣势，和实际工程上会遇到的问题并且要知道如何解决(比如，`SVD` 在遇到数据特别大的时候会产生一定的问题？如何解决？)。如果自己不太熟或说不清楚的算法或记忆不太清楚的项目最好不要写在简历上。一般之前实习的具体内容不会被问到，面试官感兴趣的主要是学校课程和实验室所做的基础项目。

这类岗位同时也经常会问到一些统计相关的问题，所以基本的统计模型，概率题，矩阵知识得熟悉。也经常会遇到一些很抽象的问题，比如某技术的前景，和一些技术的对比，所以建议大家多看看机器学习相关的公众号啊之类的。

面试时楼主认为不仅是你能回答所有问题，态度等也很重要，要让面试官感受到你的谦虚和热情并且感受到你是一个将来可以融入他们团队的人。面试官更想看到你是个学习能力强并且积极阳光的人，毕竟校招不同于社招，并不会期待你能懂很多很多。

以下是楼主遇到过的问题，这些问题只供参考，不同面试官问的问题差别会很大。所以重点还是放在精通简历内容上。

有任何具体问题可以私信我～祝大家都拿到心仪的 offer！

算法,数据库,数据结构:

1. 快排
2. 归并排序(核心为一个 `merge` 函数，把两个 `sort` 好的 `list` 合并起来)
3. 冒泡排序
4. 两个单向链表相接的一系列问题
5. `reverse` 一个 `linked list`，`while` 和 `for` 两种方法
6. `dynamic programming` (要会做基础的背包问题，`longest substring` 等)
7. `mysql` 索引
8. `B+`树(节点如何分裂的)

9. 数据库底层所用到的 `grace hash join` 和 `merge join`

10. 死锁和解决办法

11. n 个数取 k 个最大

数理统计:

1. 如何求出一个圆的 $\sin 1^\circ$ 度的大小(用 `taylor series`)

2. 有一个不规则硬币，如何用它做一个 1 到 6 的随机数生成器(核心为找六个概率相同的事件)

3. 斗地主摸到王炸和四个二的概率是多少

4. 有一个二位坐标轴，你从原点可以往上下左右走，请问 100 步后走回原点的可能性(题目大概是这样，学过随机过程的可以做做)

5. `confidence interval`

6. 统计中的 P 值和 Alpha 值

7. 三门问题

8. 还有一些其他的概率题不太记得了

9. 矩阵的基本知识得牢记

机器学习方面:

(机器学习方面的问题一般会根据简历来问，简历上写了的算法一定要很清楚如何底层实现，面试官一般都会让你给他讲)

1. 随机森林(注意要了解随机树上的每个节点上的 `feature` 是怎么选择的)

2. 线性回归(最优解表达式要牢记)

3. `logistic` 回归

4. `logistic` 函数和 `relu` 函数的图像是什么样的

5. 反向传播神经网络的讲解

6. PCA 和 SVD(数学方面的分解最好能比较精通)
7. SVD 在遇到数据特别多的时候会产生一定的问题? 如何解决?
8. cuda 可以通过哪些方法对神经网络进行加速
9. 有一张统计 qq 在线时间分布图, 基于一些统计的 test 提问
10. 推荐系统(简单的可以用 PCA 和 SVD 做)
11. RNN 和 CNN 的对比
12. 如何做矩阵的卷积。

语言:

(C++,java 的这里就不写了, 现在这类岗位比较喜欢问 python 相关的问题)

1. python2 和 python3 有哪些区别
2. 用哪些库(比较基础的是 numpy 和 scikit-learn)
3. python 如何继承其他的 class
4. super()的用法
5. python 较 java 相比的优劣势
6. AngularJS, ReactJS, html 的比较(写过前端的同学可以看看)
7. unix 里的一些指令, 比如 cat, mkdir
8. git 的一些指令, 比如 fetch 和 pull 的区别
9. SQL 一系列问题

2017 秋招面试总结-计算机视觉/深度学习算法

作者: 拾荒者 000

链接: <https://www.nowcoder.com/discuss/66114>

来源: 牛客网

楼主秋招过程，也没面几家（8月低到9月中旬面试安排比较多），9月15号之前把公司定了。月底就把三分寄出去了，找的比较随意。国庆节之后，一家没面，把所有的面试都推了，进入二面的也推了。

现在把我面试的公司的面试情况总结一下（**以下仅代表个人经验和观点**）：

- 1、简历上的项目非常重要，对一些细节一定要熟悉，尤其是项目中用到的算法。
- 2、基础很重要！！

图谱科技面试

- 1、 梯度下降：为什么多元函数在负梯度方向下降最快？

数学证明题，和专业无关-->设多元函数及多元函数的单位向量，再求多元函数在单位向量上的方向导数

① 设多元函数

$$f(x) = f(x_1, x_2, \dots, x_n)$$

② 梯度下降公式:

$$x_i^* = x_i - \alpha \left[\frac{\partial f}{\partial x_i} \right] \quad (i=1, \dots, n)$$

α 为学习率, $\frac{\partial f}{\partial x_i}$ 为 $f(x)$ 对 x_i 的梯度

③ 现在取 $v = (\cos \theta_1, \cos \theta_2, \dots, \cos \theta_n)$ 为 $f(x)$ 的一个单位方向向量.

则 f 沿 v 的导数为 (方向导数)

$$\left[\frac{\partial f}{\partial v} \right] = \frac{\partial f}{\partial x_1} \cos \theta_1 + \frac{\partial f}{\partial x_2} \cos \theta_2 + \dots + \frac{\partial f}{\partial x_n} \cos \theta_n$$

$$= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) \cdot \begin{bmatrix} \cos \theta_1 \\ \vdots \\ \cos \theta_n \end{bmatrix}$$

$$= \nabla f \cdot v$$

$$= \|\nabla f\| \cdot \|v\| \cdot \cos(\nabla f, v)$$

$$= \|\nabla f\| \cdot \cos(\nabla f, v)$$

可推出 当 $\cos(\nabla f, v) = -1$ 时, $\frac{\partial f}{\partial v}$ 取最小值
即此时负梯度方向下降最快

2、 Sigmoid 激活函数为什么会出现梯度消失? Sigmoid 函数导数的最大值出现在哪个值? $\rightarrow (x=0)$ 处

ReLU 激活函数为什么能解决梯度消失问题?

3、 Softmax 是和什么 loss function 配合使用? \rightarrow 多项式回归 loss

该 loss function 的公式?

4、（以 xx loss function ）推导 BP 算法？

5、CNN 中，卷积层的输入为 $\text{df} \times \text{df} \times \text{M}$ ($\text{weight, height, channel}$)，输出为 $\text{df} \times \text{df} \times \text{N}$ （或输出为 $\text{df} \times \text{df} \times \text{M}$ ），卷积核大小为 $\text{dk} \times \text{dk}$ 时，请问由输入得到输出的计算量为多少？题中默认 $\text{stride}=1$

计算量-->浮点数计算量： 49HWC^2 ， 27HWC^2 -->会把滤波器用到每个像素上，即为长 \times 宽 \times 可学习的参数个数

6、说一下 dropout 的原理？若在训练时，以 p 的概率关闭神经元，则在预测（测试）的时候概率 $1-p$ 怎么使用？<https://yq.aliyun.com/articles/68901>

测试时，被 dropout 作用的层，每个神经元的输出都要乘以 $(1-p)$ 使训练和测试时的输出匹配

7、传统机器学习是否了解？

8、说一下作项目时遇到的困难？

9、表达式为 $\max(x, y)$ 的激活函数，反向传播时， x 、 y 上的梯度如何计算？

答：较大的输入的梯度为 1，较小输入的梯度为 0；即较小的输入对输出没有影响；另一个值较大，它通过最大值运算门输出，所以最后只会得到较大输入值的梯度。这也是最大值门是梯度路由的原因。

前向传播时，最大值往前传播；反向传播时，会把梯度分配给输入值最大的线路，这就是一个梯度路由。

地平线机器人面试

1、检测框架 faster rcnn 是怎样的一个框架？这里回答了 faster rcnn 的过程

2、Faster rcnn 中，ROI pooling 具体如何工作（怎么把不同大小的框，pooling 到同样的大小）？

RoIPool 首先将浮点数值的 ROI 量化成离散颗粒的特征图，然后将量化的 ROI 分成几个空间的小块（spatial bins），最后对每个小块进行 max pooling 操作生成最后的结果。

3、优化代码的方法：多线程等多线程比单线程快

3、深度学习那个项目做的方法没有创新点；深度学习项目，数据集要自己做，检测方法要创新自己制作数据集并添加新层（新的激活函数 maxout）

- 4、 每个项目的衡量指标；如：（1）双目追踪能检测的目标最小是多大？能检测的最小目标是根据实时图像中最大的目标而定的，设定目标面积小于最大的目标的面积的 1/5 是不能检测的。

（2）深度学习中的指标 mAP 等（衡量模型好坏的指标？）平均精度（mAP）如何计算的？<http://blog.csdn.net/timeflyhigh/article/details/52015163>

<http://blog.csdn.net/Relocy/article/details/51453950>

目标检测的指标：识别精度，识别速度，定位精度

- A、目标检测中衡量**识别精度**的指标是 mAP（mean average precision）。多个类别物体检测中，每一个类别都可以根据 recall 和 precision 绘制一条曲线，AP 就是该曲线下的面积，mAP 是多个类别 AP 的平均值。
- B、目标检测评价体系中衡量**定位精度**的指标是 IoU, IoU 就是算法预测的目标窗口和真实的目标窗口的交叠（两个窗口面积上的交集和并集比值），Pascal VOC 中，这个值是 0.5（已被证明相对宽松）。

机器学习中评价指标：Accuracy、Precision、Recall

- 6、熟悉基本的图像处理算法和图像处理方法（如图像矫正）
- 5、Caffe 中具有哪些层，如 data layer、image data layer、softmaxwithloss（还有其它 loss）
- 6、训练网络时，如果要每个 batch 中每种类别的图数量固定（按自己定的取），则该怎么做？（训练时，每个 batch 都是随机从数据集中抽取一定数量的样本）。
- 7、立体匹配有哪些方法？收藏的链接
- 8、混合高斯模型（GMM）是怎么样的？原理和公式

混合高斯模型是无监督学习，可用于聚类

<http://www.cnblogs.com/mindpuzzle/archive/2013/04/24/3036447.html>

<http://blog.csdn.net/wqvbjhc/article/details/5485242>

混合高斯模型使用 K（基本为 3 到 5 个）个高斯模型来表征图像中各个像素点的特征，在新一帧图像获得后更新混合高斯模型，用当前图像中的每个像素点与混合高斯模型匹配，如果成功则判定该点为背景点，将其归入该模型中，并对该模型根据新的像素值进行更新，若不匹配，则以该像素建立一个高斯模型，初始化参数，代理原有模型中最不可能的模型。最后选择前面几个最有可能的模型作为背景模型，为背景目标提取做铺垫。

9、 光流法？

<http://blog.csdn.net/carson2005/article/details/7581642>

<http://www.cnblogs.com/xingma0910/archive/2013/03/29/2989209.html>

光流法，通过求解偏微分方程求的图像序列的光流场，从而预测摄像机的运动状态

10、 Kalman 滤波器的原理？

Kalman 滤波器最重要的两个过程：预测和校正

<http://blog.csdn.net/carson2005/article/details/7367135>

11、 需要熟悉简历上项目写的每个算法的具体过程甚至公式；；； 以及是否对算法进行改进，即修改 OpenCV 的源码

12、 编程：合并两个单调不减的链表，并保证合并后的链表也是单调不减的？

好未来

1、 LeetCode 第一题“TwoSum”

2、 通过简单示例，详细解释 ROC 曲线，要求给出必要的公式推导。

3、 给出 LR（逻辑回归）算法的 cost function 公式的推导过程。

4、 目标检测时，输入的是视频时，如何进行检测？视频中有很多无用的帧（不包含要检测的目标等）-->人工分割视频、每隔一定数量的帧进行检测

5、 项目。

美团

一面：

1、 faster rcnn 中 ROI pooling 是不是第一次用？-->第一次用是在 fast rcnn 中

2、 在检测物体时，物体只有少部分在图像中时，是否检测？

系统检测的最小目标为 16×16 ；当部分在图像中时，也对其进行检测（这里有一个阈值，当目标的面积占图像的面积比小于 $1/5$ 时，不检测；否则就检测（这个思路是从单路分类那来的））

- 3、 双目视觉中，立体校正如何进行？-->立体标定得出俩摄像机的旋转和平移矩阵，然后在左右图像进行校正，使其行对齐。
- 4、 Kalman 滤波器是否有运动方程？没建立运动方程，直接将物体轮廓的外接矩形的中心点作为初始化追踪点进行后续追踪。
- 5、 双目视觉中，光流法用的哪一种？L-K 光流是稠密的还是稀疏的？

金字塔 L-k 光流，其计算的是稀疏特征集的光流。

非金字塔 L-k 光流(原始的 L-k 光流)计算的是稠密光流。

http://blog.sina.com.cn/s/blog_15f0112800102wjai.html

二面：

- 1、 Fatser rcnn 与 rcnn 的不同？-->fatser rcnn 是端到端；rcnn 不是端到端
- 2、 Rcn、fatse rcnn、fatser rcnn、mask rcnn 的原理？

mask rcnn-->在 fatser rcnn 的基础上对 ROI 添加一个分割的分支，预测 ROI 当中元素所属分类，使用 FCN 进行预测；

具体步骤：使用 fatser rcnn 中的 rpn 网络产生 region proposal (ROI)，将 ROI 分两个分支：（1）fatser rcnn 操作，即经过 ROI pooling 输入 fc 进行分类和回归；（2）mask 操作，即通过 **ROIAlign** 校正经过 ROI Pooling 之后的相同大小的 ROI，然后在用 fc 进行预测（分割）。

ROIAlign 产生的原因：RoI Pooling 就是实现从原图区域映射到卷积区域最后 pooling 到固定大小的功能，把该区域的尺寸归一化成卷积网络输入的尺寸。在归一化的过程当中，**会存在 ROI 与提取到的特征不对准的现象出现**，由于分类问题对平移问题比较鲁棒，所以影响比较小。但是这在预测像素级精度的掩模时会产生一个非常的大负面影响。作者就提出了这个概念 ROIAlign，使用 ROIAlign 层对提取的特征和输入之间进行校准。

ROIAlign 方法：作者用双线性插值（bilinear interpolation）在每个 RoI 块中 4 个采样位置上计算输入特征的精确值，并将结果聚合（使用 max 或者 average）。

Lmask 为平均二值交叉熵损失。

实例分割的目的是区分每一个像素为不同的分类而不用区别不同的目标。实例分割就是要在每一个像素上都表示出来目标所属的具体类别。

- 3、 介绍 resnet 和 GoogLeNet 中的 inception module 的结构？

ResNet 主要的创新在残差网络，其实这个网络的提出本质上还是要解决层次比较深的时候无法训练的问题。这种借鉴了 **Highway Network** 思想的网络相当于旁边专门开个通道使得输入可以直达输出，而优化的目标由原来的拟合输出 $H(x)$ 变成输出和输入的差 $H(x)-x$ ，其中 $H(x)$ 是某一层原始的期望映射输出， x 是输入。

à 优化后

优化后的结构新结构中的中间 3×3 的卷积层首先在一个降维 1×1 卷积层下减少了计算，然后在另一个 1×1 的卷积层下做了还原，既保持了精度又减少了计算量

Resnet http://blog.csdn.net/mao_feng/article/details/52734438

Inception module:

共有四个版本。网上搜 **inception v4** 就会出现 **v1-v4**

<http://www.voidcn.com/article/p-zglerubc-ty.html> （重点）

<http://blog.csdn.net/u010025211/article/details/51206237>

<http://blog.csdn.net/sunbaigui/article/details/50807418>

inception v1 中用 1×1 的卷积à降维

Inception v2 (BN-inception) 在 **v1** 的基础上增加 **BN** 层，同时将 5×5 的卷积核替换成两个 3×3 的卷积核（降低参数数量，加速计算）

Inception v3 见博客 <http://www.voidcn.com/article/p-zglerubc-ty.html>

v3 一个最重要的改进是分解（Factorization），将 7×7 分解成两个一维的卷积（ $1 \times 7, 7 \times 1$ ）， 3×3 也是一样（ $1 \times 3, 3 \times 1$ ），这第一个样的好处，既可以加速计算（多余的计算能力可以用来加深网络），又可以将 1 个 conv 拆成 2 个 conv，使得网络深度进一步增加，增加了网络的非线性

v3 使用的是 **RMSProp** 优化方法

inception v4→**16 年 imagenet 分类的第一名**

<http://blog.csdn.net/lynnandwei/article/details/53736235>

由链接中的图可以看出，**v4** 包含 **v3** 和 **v2**

4、 yolo 和 ssd?

5、 Fatser rcnn 不能检测小目标的原因？

6、 在训练好的模型数据里， 如何添加识别错误的数据， 在进行训练呢？

方法一：直接往 lmdb 数据里添加，再次重新训练；

方法二：把你的 proto 里 datalayer 改成用 image data layer 然后把需要添加的图像路径写到 list 文件里，然后 fine tune 你的网络

RoKid 机器人

1、 Adaboost 算法？

2、 逻辑回归实现多分类？

3、 Fatser rcnn 中，如何处理背景候选框和含有目标的候选框不平衡问题？

4、 SVM 的核函数的作用？

5、其他就是和项目相关的问题？

数码视讯

1、 Canny 边缘检测算法的过程？

2、 常用的局部特征和全局特征？

3、 LDA 原理？

----- 本文来自 Murray_ 的 CSDN 博客，全文地址请点击：
https://blog.csdn.net/Murray_/article/details/79952076?utm_source=copy