

人脸检测算法Faster R-CNN

人脸检测算法

基于Adaboost的人脸检测方法

- Adaboost 算法主要是通过反复训练样本数据集中的正负样本，不断地调整权重值，进而得到多个不同的弱分类器，接着将得到的多个弱分类器进行加权叠加，最后得到一个强分类器，用来进行人脸检测。理论研究明：只要弱分类器检测的准确率高于50%，通过对弱分类器进行级联组成强分类器，当级联的个数趋向于无穷多个时，强分类器的检测准确率将趋于1。
但是，复杂环境会对Adaboost 人脸检测算法产生较大的影响，从而导致人脸检测结果不稳定，误检率较高。

基于PCA（主成分分析）的人脸检测方法

- PCA 是模式识别中的有效方法之一，通过特定的方法，将样本线性映射到高维空间，使样本的类内散布程度达到最小的同时，而类间散布程度达到最大，进而提取图像数据的主要特征。一般来说，PCA 可以将以下两个参数构成检测特征向量：1) 人脸基准点的相对比率，2) 其他描述人脸部特征的轮廓参数或类别参数等，通过这种方式，使基于人脸整体的检测不仅对人脸不同部件之间的相对位置信息进行了保留，也对各部件本身的信息进行了保留。

基于可变形组件模型的人脸检测方法

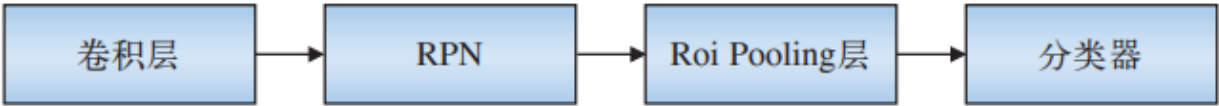
- DPM 是传统检测的主流方法，是一种基于组件的检测方法，通过使用一组混合的 DPM 获取人脸在不同姿势、表情下的参数，进而得到人脸位置和关键点信息。此方法可以有效地使用较少的数据集进行训练，并且因为没有模型的变形，具有较好的泛化性能，对扭曲、多姿态和多角度的人脸有较好的检测效果。但是该模型过于复杂，判断时计算复杂，很难满足实时性的要求。

基于人工神经网络的人脸检测方法

- 通过对生物神经网络的结构和功能进行模仿，构建的神经网络模型叫做人工神经网络。此方法一般需要大量的数据样本进行训练学习，在训练的过程中，对各个层的权重进行校正而创建模型的过程，经常通过反向传播算法对网络的权重值进行验证，最终对模型进行优化，达到较高的检测率。

1 Faster R-CNN人脸检测算法

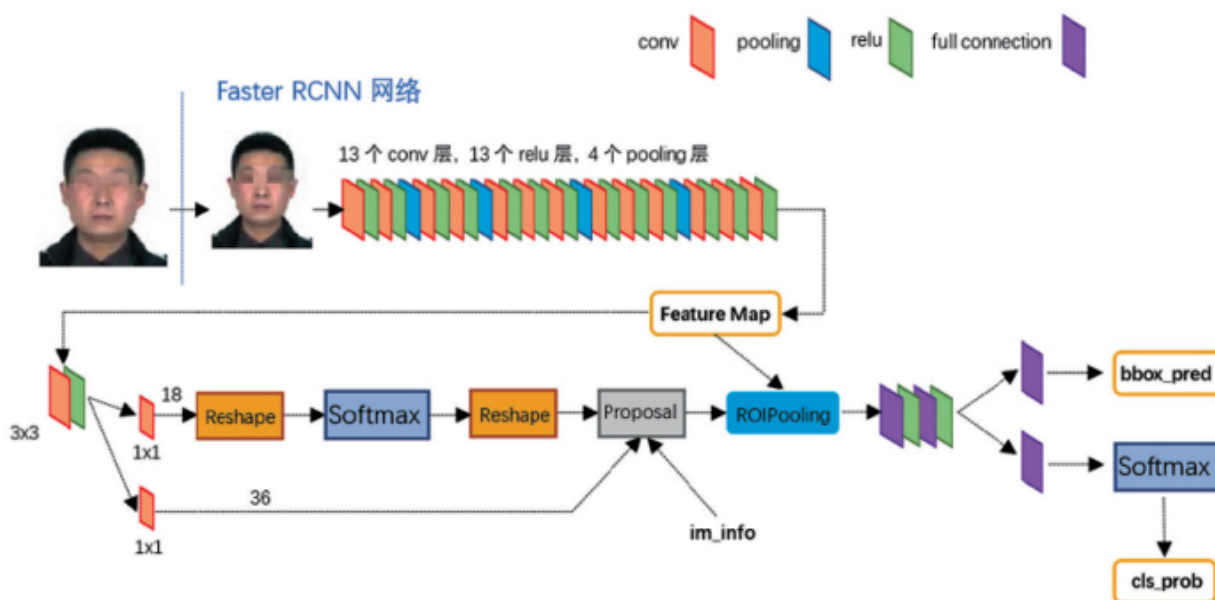
1.1人脸检测算法设计



Faster R-CNN的基本结构

- 卷积层：**作为一种 CNN 网络目标检测方法，Faster R-CNN 中的卷积层使用 ResNet为特征提取 网络，该层中网络提取了图像的特征图，该特征图被共享用于后续 RPN 层和全连接层。在训练过程中 神经网络可以自动学习提取人脸特征的卷积核。随 着层数的加深，卷积核能表示更高级的人脸特征。 其中，低级的人脸特征为颜色、纹理、边缘等基础特征；随着特征层的加深，逐渐可以学习人脸器官，包括鼻子、眼睛和嘴巴等的特征；由这些特征组合成整张人脸的特征。

- **RPN层**：用于生成目标候选区域，该层通过softmax 判断目标候选区域属于目标或背景，再利用目标坐标框位置回归修正目标候选区域获得精确的目标候选区域。其中，坐标位置回归网络使用目标的左上角坐标和长宽4 个参数来表示一个目标的位置。
- **Roi Pooling层**：该层收集输入的特征图和目标候选区域，综合这些信息后提取候选目标区域的特征图，送入后续全连接层判定目标类别。
- **分类层**：利用候选目标区域的特征图计算候选目标区域的类别，再次通过目标的坐标框位置回归获得目标最终的精确位置。



Faster R-CNN网络结构

- 数据集：名人人脸属性数据集CelebA
- **卷积层**的卷积运算可以使原图像信号特征增强，并降低噪音。卷积层相当于对图像进行滤波，抽象出来图像的局部信息，局部信息是它能够通过较小的卷积核在图像不同的局部位置上扫描得到。
- **池化**可以根据图像局部相关性的原理，对图像进行子采样减少计算量，同时保持图像旋转的不变性。主要作用包括降低特征图的分辨率，从而减少计算量，以及增强网络的鲁棒性，降低数据维度。池化的方式一般有最大值池化和均值池化，而空间金字塔池化，使得任意大小的特征图都可以转换为固定大小的特征向量。
- **下采样层**主要是为了减少过拟合的问题，减少不同参数之间的耦合性。
- **全连接层**在一定程度上泛化了Dropout，也是一种缓解拟合的技术。而Dropout只在全连接层使用，随机的将全连接层的某些神经元的输出置0。
- **Softmax层**，得到的激活值就是卷积神经网络提取到的图像特征。

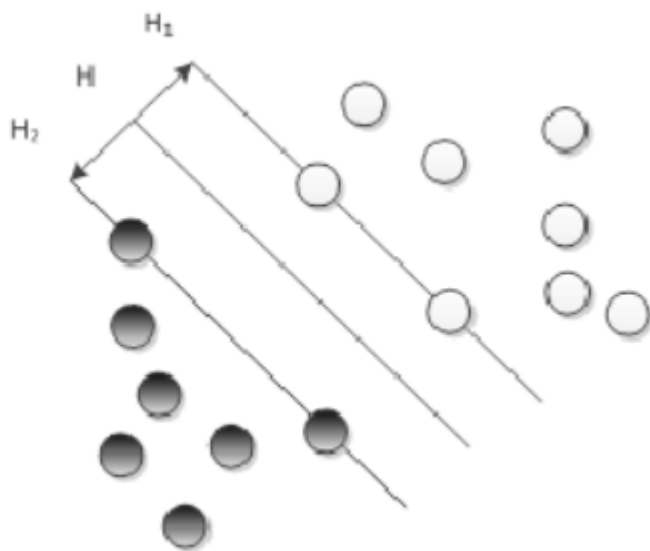
1.2 VGG网络模型

- VGG网络模型一共有19层，其中，有13个卷积层，5个池化层，3个全连接层，分别是两个图像特征层和一个分类特征层。具有较小的卷积核和较小的跨步。

1.3 分类器

- **Softmax分类器**主要是针对多分类问题，类别之间是相互排斥的，即每一个输入样本只能被归为一类。
- **逻辑回归 (LR) 分类器**主要用于二分类问题，多个LR分类器可以进行多分类，但是输出的类别之间不是相互排斥的。

- **SVM分类器**是一种二分类模型，通过确定一个分类的超平面，来使特征空间上的间隔最大化，只需要少数的样本信息就可以确定分类超平面。



1.4 Fast R-CNN

- 由 5个卷积层， 1个ROI Pooling层， 4个全连接层 和 1个Softmax层 组成。

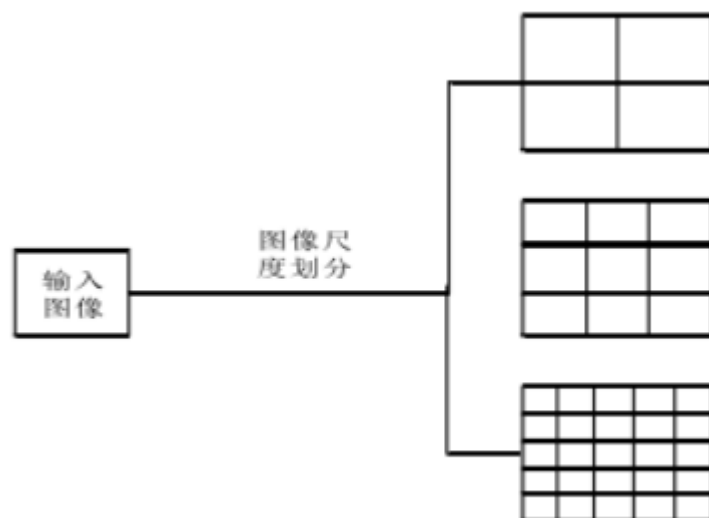
1.5 ReLU激活函数

- ReLU本质上是分段线性模型，前向计算非常简单，无需指数之类的操作；
- ReLU的偏导也很简单，反向传播梯度，无需指数或者除法之类操作；
- ReLU不容易发生梯度发散问题，Tanh和Logistic激活函数在两端的时候倒数容易趋近于0，多级连乘后梯度更加约等于0；
- ReLU关闭了右边，从而会使得很多的隐藏层输出为0，即使网络变得稀疏，起到了和正则化相似的作用，在一定程度上可以缓解过拟合问题。
- ReLU是一个非线性操作，可使得网络的非线性表达更加丰富。

$$f(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$$

1.6 空间金字塔池化

- 空间金字塔池化（SPP）使得任意大小的特征图都可以转换成固定大小的特征向量。具体流程：
 - 输入任意大小的一张图像，假设大小为（W,H），输出神经元的个数是38个，希望特征向量是38维。



SPP示意图

- 对于输入的图像，采用大小不同的刻度对其进行划分，分别为（2x2, 3x3, 5x5），从而得到4+9+25=38个特征点，从这38个不同特征点抽取一个特征，从而就得到一个38维的特征向量。

1.7 算法评价标准

表3.1 分类器易混淆名词矩阵

		真实类别	
		正类别	负类别
预测类	正类	真阳性(TP)	假阳性(FP)
	负类	假阴性(FN)	真阴性(TN)

正类样本： $T = TP + FN$ ，负类样本： $N = FP + TN$

- 重叠度IoU:

$$IoU = \frac{\text{Region Proposal} \cap \text{Ground Truth}}{\text{Region Proposal} \cup \text{Ground Truth}}$$

$IoU > 0.7$ ：物体被检测到，反之物体没有被检测到。

- 精确率Precision:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- 准确率Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{P + N}$$

- 错误率Error:

$$\text{Error} = \frac{\text{FP} + \text{FN}}{\text{P} + \text{N}}$$

- 召回率Recall:

$$\text{Error} = \frac{\text{FP} + \text{FN}}{\text{P} + \text{N}}$$