

这篇论文描述了一个能够快速检测人脸的新型架构。该架构主要有三个贡献：

- 使用哈尔特征并采用积分图进行快速计算
- 利用 AdaBoost 算法训练多个弱分类器并根据分类效果赋予不同权重，最终组合成一个更加准确高效的强分类器。
- 将多个强分类器级联，以逐层过滤的方式从简单特征到复杂特征评估整张图像，从而筛选出最终的结果。

论文中不断改进和强调的就是其快速计算的能力以及架构本身的简洁性。我试就针对这几点讨论一下其主要想法及我的个人观点。

首先，为了达到快速计算的目的，论文中提出了三点必要的措施。其一就是利用积分图以常量时间计算哈尔特征。论文中使用了三种哈尔特征，但都是水平或者垂直相邻的矩形组成的特征。这些特征对边缘和线段比较敏感，可以描述一个区域内的灰度变化情况（在文中所有的人脸都简化为灰度图）。这也是论文中采用哈尔特征的一个重要原因，若直接把像素点引入计算，那么就很难将图片的一个区域的特征表现出来。然而，这几种特征只适合描述那些形状比较规整，走向特定的物体。为了进一步提高识别的准确度，我认为有必要引入其他类型的哈尔特征。比如 45 度倾斜的哈尔特征，提高该模型对其他倾斜边缘和形状特征的敏感度。这样对于那些不规整的物体特征也能够比较良好地识别出来。

其二，采用 AdaBoost 算法在数以万计的哈尔特征中选取最为关键的特征组成强分类器。这样就简化了识别时的计算流程。

其三，将多个不同复杂度的强分类器级联成瀑布模型。据论文中说，在这个级联模型的前期就能将大多数的非人脸区间拒绝。在训练的过程中，随着权重的不断调整以及样本的不断向下迁移，级联模型中越靠后的分类器面对的区间越难以处理。这里就不得不提及这个算法本身的要求：只有当这级联起来的多个分类器相互独立，即对于同一张图像，它们拒绝的概率相互独立时，这样组合起来的级联结构才能够有效互补，形成最终更加强大的分类器。我想只要让训练集中包含足够多的情况，这个级联模型中的每一部分依赖性就会大大降低。

另一方面，级联结构的出现也在几个方面简化了系统的实现。第一，简化了级联各个部分的设计。各个部分可以单独训练并在需要的时候自由组合。在增强系统灵活性的同时也提高了拓展能力。另外，在检测图像时，由于大部分的区域能够在最初几个分类器就被拒绝，因此整个级联结构的平均计算代价能够大大降低。