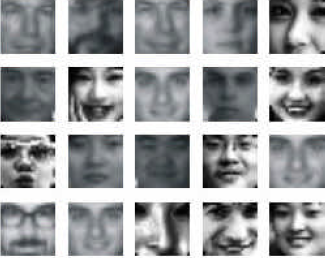
1. **背景介绍**

**目的**：任意一种boost方法实现一个人脸检测程序。

**数据**：MIT Face Data。 数据特点是图像尺寸为19\*19，其中训练样本中人脸数据有2429张，非人脸数据有4548张；测试样本中人脸数据有472张，非人脸数据有23572张。图1则是从MIT数据中随机抽取的人脸样例。



**图1 MIT数据中的人脸样例**

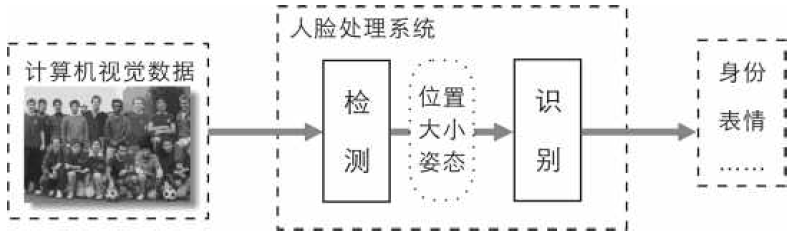
**评价标准：**准确率、召回率。

1. **相关知识概述**

**（1）人脸分析概述**

人脸分析，主要包括人脸检测（face detection）和人脸识别（face recognition）两部分（如下图2所示）。最初人脸分析主要集中在人脸识别领域。与指纹、视网膜、虹膜、基因、掌纹等其它人体生物特征识别系统相比，人脸识别更加直接、友好，其使用者无心理障碍。人脸识别技术应用背景十分广泛，可用于公安系统刑侦破案的罪犯身份识别、身份证及驾驶执照等证件验证、银行及海关的监控、自动门卫系统、视频会议、机器人的智能化研究以及医学等方面。

人脸检测是人脸分析的第一步，指对于任意一幅给定的图像，采用一定的策略对其进行搜索以确定其中是否含有人脸，如果含有则返回人脸的位置、大小和姿态。**本文基于图像的Haar-like 特征入手，通过Adaboost算法实现人脸检测。**



**图2 人脸分析流程**

**（2）AdaBoost概述**

Adaboost是一种重用的统计学习方法，其核心思想是针对同一个训练集训练不同的 弱分类器，然后把这些弱分类器集合起来，构成一个强分类器。其算法本身是通过改变数据分布来实现的，它根据每次训练集之中每个样本的分类是否正确，以及上次的总体分类的准确率，来确定每个样本的权重。将修改过权值的新数据集送给下层分类器进行训练，最后将每次训练得到的分类器最后融合起来，作为最后的决策分类器。AdaBoost一般流程：

(i)数据收集：根据需求选择相应方法进行数据收集

(ii)准备数据：依赖于所使用的弱分类器类型。

(iii)分析数据：根据实验目的进行数据的各方面分析

(iv)训练算法：AdaBoost的大部分时间都用在训练上，分类器将多次在同一数据集上训练弱分类器。

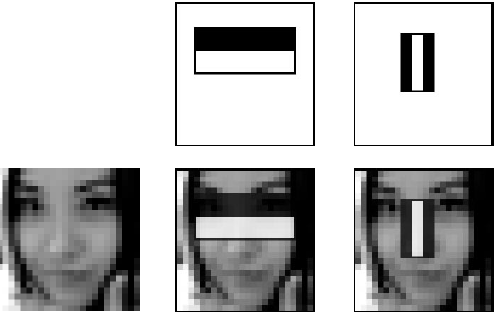
(v)测试算法：计算分类的错误率，如准确率，召回率、F1值等。

(vi)使用算法：同SVM一样，AdaBoost预测二类分类。如果想要将他应用到多个类别的场合，则类似多类SVM做法一样对于AdaBoost进行修改。

1. **Haar-like特征与积分图**

**（1）Haar-like特征**

Haar-like特征最早是由Papageorgiou等应用于人脸表示，其对一些简单的图形结构，比如边缘、线段，比较敏感，但是其只能描述特定走向（水平、垂直、对角）的结构，因此比较粗略。如图2，脸部一些特征能够由矩形特征简单地描绘，例如，通常，眼睛要比脸颊颜色更深；鼻梁两侧要比鼻梁颜色要深；嘴巴要比周围颜色更深。但Haar-like特征只对一些简单的图形结构，如边缘、线段较敏感，所以只能描述特定走向（水平、垂直、对角）的结构。



**图2 脸部特征能由Haar-like特征简单描述**

Viola和Jones在Papageorgiou的基础上，使用3种类型4种形式的特征。Haar特征分为三类：**边缘特征、线性特征、中心特征和对角线特征**，组合成特征模板。特征模板内有白色和黑色两种矩形，并定义该模板的**特征值为白色矩形像素和减去黑色矩形像素和**。图3为4种类型Haar特征。



**图3 4种类型Haar特征。**

对于图3中的A, B和D这类特征，特征数值计算公式为：v=Sum白-Sum黑，而对于C来说，计算公式如下：v=Sum白-2\*Sum黑；之所以将黑色区域像素和乘以2，是为了使两种矩形区域中像素数目一致。

通过改变特征模板的大小和位置，可在图像子窗口中穷举出大量的特征。上图的特征模板称为“特征原型”；特征原型在图像子窗口中扩展（平移伸缩）得到的特征称为“矩形特征”；矩形特征的值称为“特征值”。

矩形特征可位于图像任意位置，大小也可以任意改变，所以矩形特征值是矩形模版类别、矩形位置和矩形大小这三个因素的函数。故类别、大小和位置的变化，使得很小的检测窗口含有非常多的矩形特征，如本文所使用的MIT的19\*19像素大小的检测窗口内矩形特征数量可以达到63960个Haar特征。

**（2）积分图**

上文我们提到过矩形特征的数据量非常巨大，那么如何快速计算那么多的特征呢？这就要说起Haar特征快速计算方法——积分图。

积分图主要的思想是将图像从起点开始到各个点所形成的矩形区域像素之和作为一个数组的元素保存在内存中，当要计算某个区域的像素和时可以直接索引数组的元素，不用重新计算这个区域的像素和，从而加快了计算（这有个相应的称呼，叫做动态规划算法）。积分图能够在多种尺度下，使用相同的时间（常数时间）来计算不同的特征，因此大大提高了检测速度。

积分图是一种能够描述全局信息的矩阵表示方法。积分图的构造方式是位置（i,j）处的值ii(i,j)是原图像(i,j)左上角方向所有像素的和（如图4所示）：



**图4 积分图构造方式**

**积分图构建算法：**

1）用s(i,j)表示行方向的累加和，初始化s(i,-1)=0;

2）用ii(i,j)表示一个积分图像，初始化ii(-1,i)=0；

3）逐行扫描图像，递归计算每个像素(i,j)行方向的累加和s(i,j)和积分图像ii(i,j)的值

s(i,j)=s(i,j-1)+f(i,j)

ii(i,j)=ii(i-1,j)+s(i,j)

4）扫描图像一遍，当到达图像右下角像素时，积分图像ii就构造好了。

举例说明（如图5所示）,区域 D 的像素值,可以利用 1、2、3、4 点的积分图来计算。在图5中：

ii1 =区域A的像素值

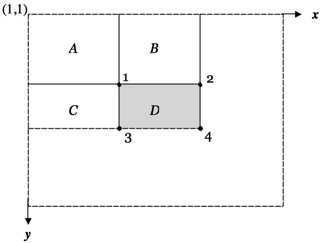
ii2 =区域A的像素值+区域B的像素值

ii3 =区域A的像素值+区域C的像素值

ii4 =区域A的像素值+区域B的像素值+区域C的像素值+区域D的像素值

所以解上述方程就可以得到:

区域D的像素值= ii4 + ii1 − (ii2 + ii3)



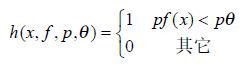
**图5 区域 D 的像素和的积分图计算方式：ii4 + ii1 − (ii2 + ii3 )**

1. **AdaBoost算法训练详解**

**（1）弱分类器**

最初的弱分类器可能只是一个最基本的Haar-like特征，计算输入图像的Haar-like特征值，和最初的弱分类器的特征值比较，以此来判断输入图像是不是人脸，然而这个弱分类器太简陋了，可能并不比随机判断的效果好。对弱分类器的学习(learning)就是训练弱分类器成为最优弱分类器，注意这里的最优不是指强分类器，只是选出的弱分类器对于样本分为的误差最小。训练弱分类器实际上是对分类器设置阈值的过程。至于如何设置分类器，设置什么，我们首先分别看下弱分类器的结构。

一个弱分类器h(x,f,p,)（图6所示）由一个特征f,阈值和指示不等号方向的p组成：



**图6 弱分类器结构**

由上述的分析,阈值的含义就清晰可见了。而方向指示符p用以改变不等号的方向。

训练一个弱分类器(特征f)就是在当前权重分布的情况下,确定f的最优阈值,使得这个弱分类器(特征 f)对所有训练样本的分类误差最低。具体弱分类器的学习方法如下：

对于每个特征f,计算所有训练样本在该特征下的的特征值,并将其排序。通过扫描一遍排好序的特征值,可以为这个特征确定一个最优的阈值,从而训练成一个弱分类器。具体来说,对排好序的表中的每个元素 ,计算下面四个值:

1) 全部人脸样本的权重的和T+;

2) 全部非人脸样本的权重的和T-;

3) 在此元素之前的人脸样本的权重的和S+;

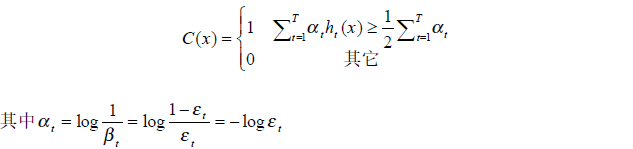
4) 在此元素之前的非人脸样本的权重的和S-;

这样,当选取当前元素的特征值为阈值时,所得到的弱分类器就在当前元素处把样本分开——也就是说这个阈值对应的弱分类器将当前元素前的所有元素分类为人脸(或非人脸),而把当前元素后(含)的所有元素分类为非人脸(或人脸)。 可以认为这个阈值所带来的分类误差为:



**（2）强分类器**

经过 T 次迭代后，获得了T个最佳弱分类器，可以按照下面的方式组合成一个强分类器：



那么，这个强分类器对待一幅待检测图像时，相当于让所有弱分类器投票，再对投票结果按照弱分类器的错误率加权求和，将投票加权求和的结果与平均投票结果比较得出最终的结果。

**（3）AdaBoost训练算法描述**

AdaBoost 训练强分类器的算法描述如下：

A. 给定一系列训练样本 ( x 1 , y 1 ), ( x 2 , y 2 ), K , ( x n , y n ) ,其中 y i = 0 表示其为负样本(非人脸), y i = 1 表示其为正样本(人脸)。n 为一共的训练样本数量。

B. 初始化权重 4 w 1,i = D (i ) ;

C. 对 t = 1, K, T :

1. 归一化权重:



2. 对每个特征 f,训练 .. 一个弱分类器 h ( x, f , p, θ ) ;计算对应所有特征的弱分类器的加权( q t )错误率 ε f :



3. 选取最佳的弱分类器 h t (x) (拥有最小错误率 ε t ):



4. 按照这个最佳弱分类器,调整权重:



5. 其中 e i = 0 表示 x i 被正确地分类, e i = 1 表示 x i 被错误地分类;



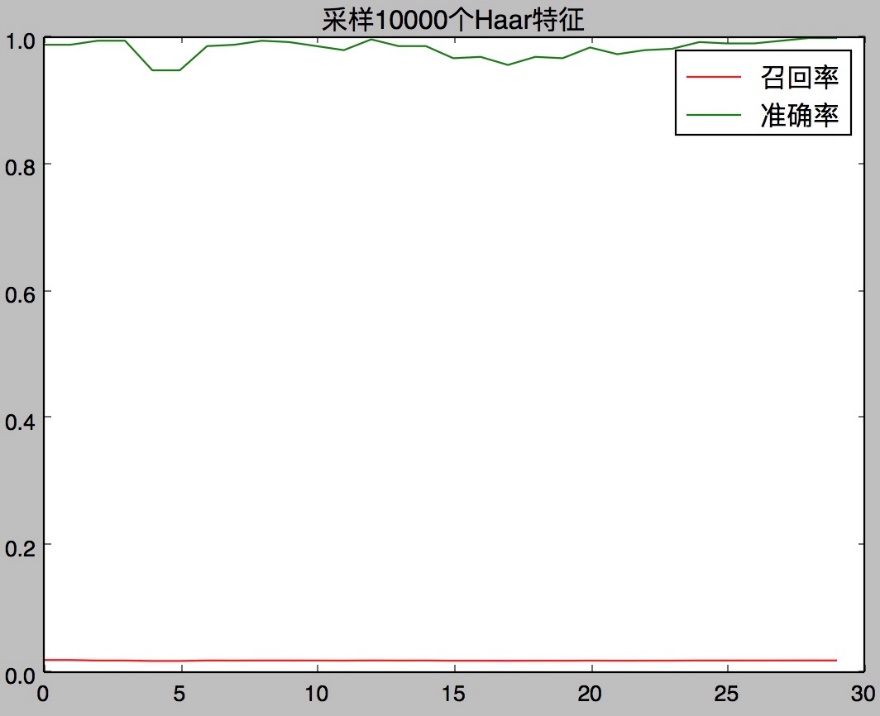
D. 最后的强分类器为：



1. **实验结果分析**

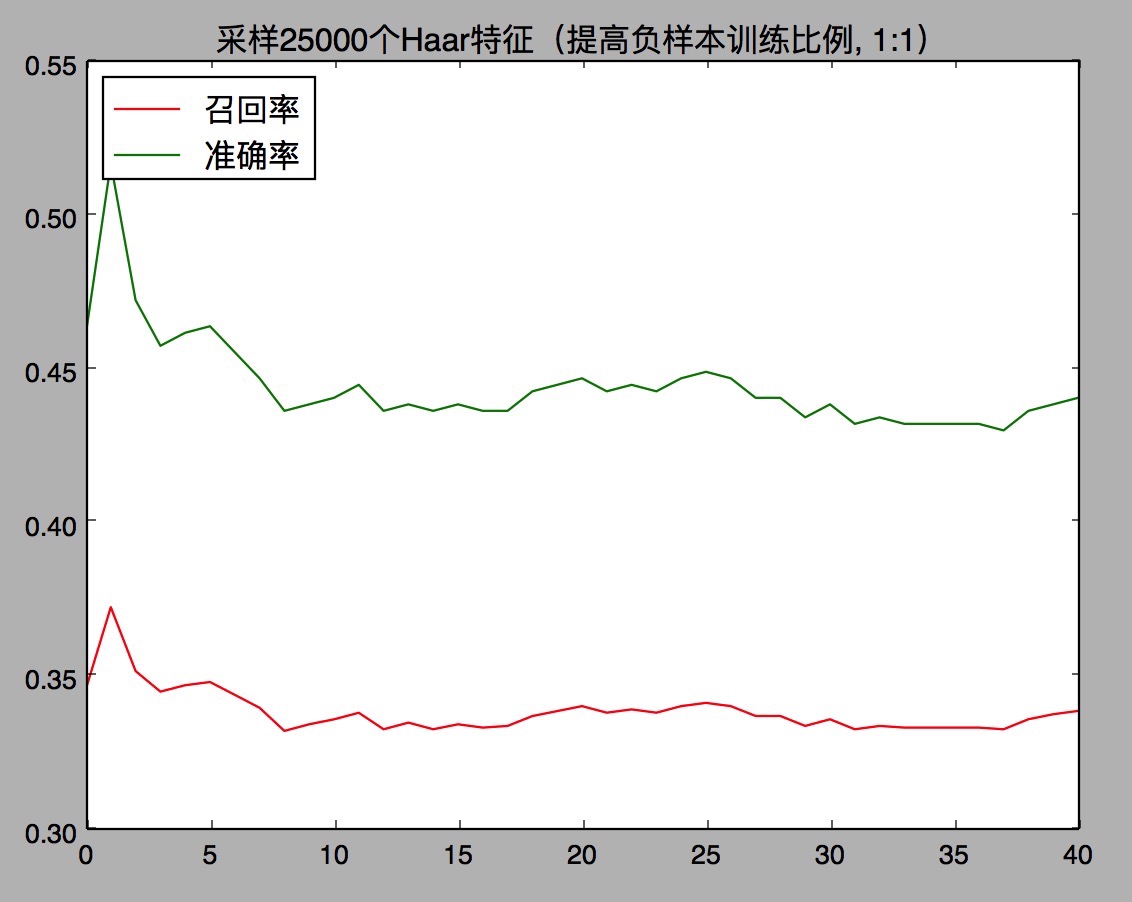
**（1）实验结果**

通过上文的论述可知一张图片的Haar特征太多了，所以本文实验中随机选取了2500个样本，构建了5\*2500=5000个Haar特征，训练了30轮。从图6可以看出，训练出来的强分类器的的准确率逼近于1，而召回率却非常低。



**图6 2500个样本分类器的准确率和召回率**

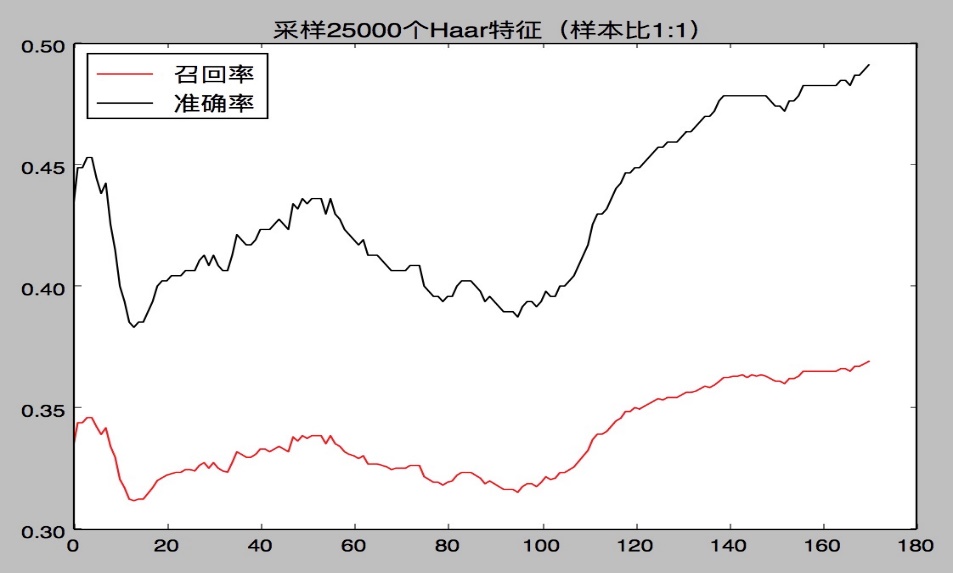
根据相关方面的调研，分析出具体的原因是训练给定的特征样本太少以及给定的非人脸的数据太少。所以根据调研的建议，我们进行了训练非人脸样本的增加，以及迭代次数的增加。故本文将测试样本即分成两部分：一部分用于模型的测试使用，一部分用于训练语料使用，其中两部分数据没有交叠。当然由于模型训练所花费的时间较长，目前只跑完了40轮的迭代，从图7是其训练结果。



**图7 增加负例样本够的训练结果**

从图中可以看到召回率比图6中的提升很多，间接证明增加非人脸的样本会有效果的提升。然而仔细分析准确率还是太低，当然这是因为该种环境下训练分类器太耗时间，训练一个分类器就要50分钟。所以本文只训练了40个弱分类器故效果不是特备明显。

由于上面的参数设定鲜果也不是很明显，所以进一步调研本人选择了5000个样本，共25000个特征。同时选择了50~150个弱分类器组成强分类器。通过训练样本的训练的强分类器的效果图8所示。



**图8 25000个Haar特征的强分类器**

图8 显示的结果虽然也是但是随着迭代次数的增加准确率和召回率一直在增加，所以在训练足够长的时间迭代次数足够多的话便可以得到更好的条件。然而由于时间原因本文没有训练足够，所以没有提高到很高的精度。

除了准确率和召回率之外的评估，本文也用训练出来的分类器对于输入带人脸的照片进行了人脸检测。下面是具体的人脸检测结果。



**图9 人脸检测结果**

**（2）结果分析**

从上面的实验结果可以看出，本文实现的三种分类器的虽然可以对于人脸进行准确是被，但是准确率和召回率的效果不是非常好，这主要原因是本文实现的分类器是随机选择了5000个样本共25000特征，并没有获得最好的人脸识别的特征。导致分类器的分类能力不佳，这也是日后需要进一步实验的部分，即将所有的6万多特征全部进行选择，选出最有的特征。但是这样的训练特别费时间，所以需要很长的时间。