

# 色情图像识别技术文档 V0.1

目前，对色情图像的识别总结了以下几类解决方案：

## 方案 1：皮肤区域检测法

与正常图片相比，色情图片的最大特征就是一张图像里包含了大量暴露的皮肤区域，因此可初步通过皮肤区域检测法来快速完成对色情图片的识别。主要方法如下：

- 1) 将 RGB 通道的图像转为 HSV 或 YCbCr 通道的图像：
- 2) 分析转换后三通道的图像信息，目前对皮肤区域的定义区间如下：
  - a) YCbCr:

$$\begin{cases} 80 \leq Cb \leq 120 \\ 133 \leq Cr \leq 173 \end{cases}$$

- b) HSV:

$$\begin{cases} 0 < H < 0.25 \\ 0.15 < S < 0.9 \\ 0.2 < V < 0.95 \end{cases}$$

- 3) 皮肤区域统计：

$$\text{skin}_{\text{area}} = \frac{\# \text{skin\_color\_pixels}}{\# \text{image\_pixels\_total}}$$

- 4) 阈值参数选定：  
选定阈值  $t$ ，当  $\text{skin}_{\text{area}} > t$  时，可认为该图片为色情图片。

- 5) 总结：

该方法通过对图像通道的变化分析，可快速检测出嫌疑色情图像，操作简单，但是对于“性感——如女性泳装照”、“正常——如男性裸露上身、裸体婴儿皮肤”等均会误判为色情图像。

## 方案 2：机器学习方法

对于一张图像，我们可以广义地定义为 3 类：“正常”、“性感”、“色情”。因此，可以利用机器学习的方法来对训练集图像进行训练，将训练后的模型用于预测。该方法主要思路如下：

- 1) 分别搜集“正常”、“性感”、“色情”三类大量带有标记的图像样本，保证总数至少为 10 万张，且 3 类样本分布需尽可能均衡，即每类都应为 3 万张左右。
- 2) 提取图像 LBP、HOG、SIFT 特征，尝试 SVM、Adboost、KNN、Random Forest 等分类器模型进行训练，选择分类性能最佳的模型。
- 3) 总结：  
该方法借助机器学习对训练集图像进行训练，因此相比于方案 1，该方法精度会有明显提升；但是需要大量标注样本（不少于 10 万张）。

## 方案 3：深度学习方法

近年来，深度学习在计算机视觉领域超越传统机器学习方法取得了突破性的进展，因

此可借助深度卷积神经网络对图像进行训练，依据数据量的大小设计不同的模型。将其转换为基于深度学习的图像分类或强监督细粒度图像识别（需要关键部位的 bbox 信息）问题。

Github 上有童鞋整理了一份 NSFW 数据集，数据集下载地址：

[https://github.com/alexkimxyz/nsfw\\_data\\_scrapper.git](https://github.com/alexkimxyz/nsfw_data_scrapper.git)

**总结：**该方法准确率和鲁棒性均可以达到最高，但需要海量标记样本。Fine-tune 或许是个不错的思路。

## 方案 4：数据挖掘方法

该方法旨在对平台用户积累的历史行为数据进行深入数据挖掘，少数色情图片上传者与大部分正常用户在行为上往往会有很大的不同。该方法对应于数据挖掘中的“**离群点分析**”。

**总结：**该方法检测效率较高（对数据库的操作效率远胜于二进制图像数据的操作），但需要平台提供海量用户历史行为数据。