# 破解音频隐写术：结合机器学习

原创

唠嗑！ 已于 2022-04-11 14:30:55 修改 4771 收藏 2
分类专栏： 多媒体编码安全 文章标签： 网络安全 深度学习 语音识别 音视频 视频编解码
于 2022-04-08 14:24:02 首次发布



多媒体编码安全 专栏收录该内容
3 篇文章 0 订阅
订阅专栏

目录



---

## 前言

**专用隐写分析方法**是针对特定算法而设计的，所以通用性很弱；相反地，**通用隐写分析**比专用隐写分析具有更强的普适性，可以检测出多种隐写方法。传统的通用隐写分析模型又被称之为**盲隐写分析模型**，主要包含**隐写特征设计**和**分类器**的选取，其中分类器包含信息分类器和集成分类器。

随着机器学习的发展，进而出现了基于**深度学习(Deep Learning)**的隐写分析，此种方法就不需要预先设计特征，可以直接利用神经网络自动完成特征学习。部分分析特征可以同时应用于多种**嵌入域**，但其适用范围和检测准确性可能有很大不同。

备注：此处的机器学习与深度学习的应用是不同的。

## 一. 马尔科夫特征集

马尔科夫(Markov)特征集是最简单也是最有效的隐写分析特征之一。假定隐写嵌入域是一种服从**马尔科夫链**的随机过程，基于隐写操作就会改变马尔科夫链的**状态转移概率矩阵**。马尔科夫特征集计算简单，并且对于某些隐写算法在**低嵌入率**下也具有较好的检测效果。但是，由于变量的**状态数**和**样本分量数量**都很大，因此马尔科夫特征的**维度**都很高。

在实际的应用中，通常需要对某些状态值、特征分量来进行筛选和优化来降维，去除某些对检测影响很小的分量和训练中**过拟合**的分量，从而满足实际计算代价和检测率的要求。马尔科夫特征集在音频隐写分析中的应用形式非常多，包含**时域**和**压缩域**等等。

以下，以MP3音频为例来说明马尔科夫特征的计算过程步骤。

## 1.1 提取QMDCT系数矩阵

提取待测MP3音频的QMDCT系数矩阵$C_Q$为如下：

$$C_Q = (c_{mn})_{M\times N} = \begin{pmatrix} c_{11} & \cdots & c_{1N} \\ \vdots & \ddots & \vdots \\ c_{M1} & \cdots & c_{MN} \end{pmatrix}$$

上式子中，$C_Q$的每个行向量表示MP3帧中的一个**颗粒**，由此M就代表音频文件的总颗粒数，它可以控制单次检测的粒度，M值越小检测粒度越细。根据MP3的编码标准，N的值大多为576，当然根据实际应用，M和N的值可以灵活选取。从每个颗粒中QMDCT系数的分布来讲（也就是每一行），大约位于后$\frac{1}{3}$的QMDCT系数都属于**零值区**，对特征的计算没有实质作用，由此就可以减小N的取值（大约300~400取值最佳）。

## 1.2 计算差分矩阵

计算$C_Q$的**差分矩阵**$D_Q$和**绝对值差分矩阵**$D_{AQ}$如下：

$$D_Q = (d_{mn})_{(M-1)\times N} = \begin{pmatrix} d_{11} & \cdots & d_{1N} \\ \vdots & \ddots & \vdots \\ d_{M-1,1} & \cdots & d_{M-1,N} \end{pmatrix}$$

$$D_{AQ} = (d'_{mn})_{(M-1)\times N} = \begin{pmatrix} d'_{11} & \cdots & d'_{1N} \\ \vdots & \ddots & \vdots \\ d'_{M-1,1} & \cdots & d'_{M-1,N} \end{pmatrix}$$

上式子中，$d_{mn} = c_{m+1,n} - c_{mn}$，$d'_{mn} = |c_{m+1,n}| - |c_{mn}|$

此步骤集中于行差分，实际上也可以修改为列差分计算。

## 1.3 计算转移概率

根据马尔科夫链的状态转移原理，可以分别计算$C_Q, D_Q, D_{AQ}$的一阶转移概率。

行方向的**帧间转移概率**如下：

$$P_{C_Q\_Inter}=\frac{\sum_{i=1}^{M-1}\sum_{j=1}^{N}\delta(c_{ij}=x,\ c_{i+1,j}=y)}{\sum_{i=1}^{M-1}\sum_{j=1}^{N}\delta(c_{ij}=x)}$$

$$P_{D_Q\_Inter}=\frac{\sum_{i=1}^{M-2}\sum_{j=1}^{N}\delta(d_{ij}=x,\ d_{i+1,j}=y)}{\sum_{i=1}^{M-2}\sum_{j=1}^{N}\delta(d_{ij}=x)}$$

$$P_{D_{AQ}\_Inter}=\frac{\sum_{i=1}^{M-2}\sum_{j=1}^{N}\delta(d'_{ij}=x,d'_{i+1,j}=y)}{\sum_{i=1}^{M-2}\sum_{j=1}^{N}\delta(d'_{ij}=x)}$$

列方向上的**帧内转移概率**如下：

$$P_{C_Q\_Intra}=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N-1}\delta(c_{ij}=x,\ c_{i,j+1}=y)}{\sum_{i=1}^{M}\sum_{j=1}^{N-1}\delta(c_{ij}=x)}$$

$$P_{D_Q\_Intra}=\frac{\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\delta(d_{ij}=x,\ d_{i,j+1}=y)}{\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\delta(d_{ij}=x)}$$

$$P_{D_{AQ}\_Intra}=\frac{\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\delta(d'_{ij}=x,d'_{i,j+1}=y)}{\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\delta(d'_{ij}=x)}$$

上式子中T代表状态量阈值，且 $\delta(X=x,Y=y)=\begin{cases}1, & X=x,Y=y\\ 0, & others\end{cases}$ ，此式子中$x,y\in[-T,T]$。最常用的T值为15。

## 1.4 特征优化

为了降低**特征向量维度**，可以通过调节**阈值T**，同时对以上式子特征进行优化选择。

$P_{D_{AQ-Inter}}$特征优化条件如下：

$$\begin{cases}\left|\mu\left(\left\{P^{S_i}_{D_{AQ}\_Inter}(x,y)\right\}_{i=1}^{I}\right)-\mu\left(\left\{P^{C_i}_{D_{AQ}\_Inter}(x,y)\right\}_{i=1}^{I}\right)\right|>\varepsilon_{Inter\_mean}\\ \sigma\left(\left\{P^{S_i}_{D_{AQ}\_Inter}(x,y)\right\}_{i=1}^{I}\right)<\varepsilon_{Inter\_std}\end{cases}$$

$P_{DAQ-Intra}$特征优化条件如下：

$$\begin{cases}\left|\mu\left(\left\{P^{S_i}_{D_{AQ}\_Intra}(x,y)\right\}_{i=1}^{I}\right)-\mu\left(\left\{P^{C_i}_{D_{AQ}\_Intra}(x,y)\right\}_{i=1}^{I}\right)\right|>\varepsilon_{Intra\_mean}\\ \sigma\left(\left\{P^{S_i}_{D_{AQ}\_Intra}(x,y)\right\}_{i=1}^{I}\right)<\varepsilon_{Intra\_std}\end{cases}$$

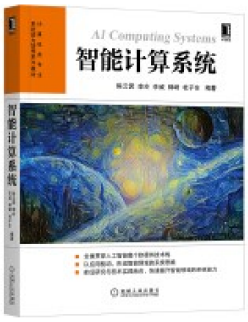
上式子中，$\mu()$和$\sigma()$分别代表**均值函数**和**标准差函数**；$S_i$和$C_i$分别表示第i个隐写音频样本和载体音频样本；I是样本对总数；$\epsilon_{Inter-mean},\epsilon_{Inter-std},\epsilon_{Intra-mean},\epsilon_{Intra-std}$是对应的阈值；$x,y\in[-T',T'](T'<T)$。

综上1.1~1.4，马尔科夫特征的维度与T值选取和特征优选有关系，如果不考虑特征优选，则特征维度计算为$6(2T+1)^2$维。x,y的计算分别有三种，总共6种；对单个x或y，其取值从-T到T，一共2T+1个值，由此可计算特征维度为$6(2T+1)^2$。

# 二. 深度学习的应用

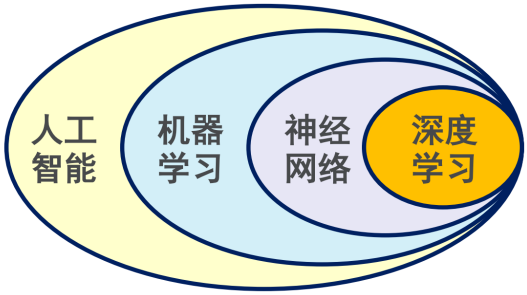
## 2.1 概述

深度学习（Deep Learning,DP）在近些年受到了世界各国研究人员的热捧，成为了一个很有潜力的研究方向。深度学习已被广泛应用于**图像分类，语音识别，目标跟踪**以及**语义分割**等众多领域。其中最具有代表性的深度学习网络包括LeNet,AlexNet,VGG,GoogLeNet,ResNet和DenseNet。列举相关的书籍如下：





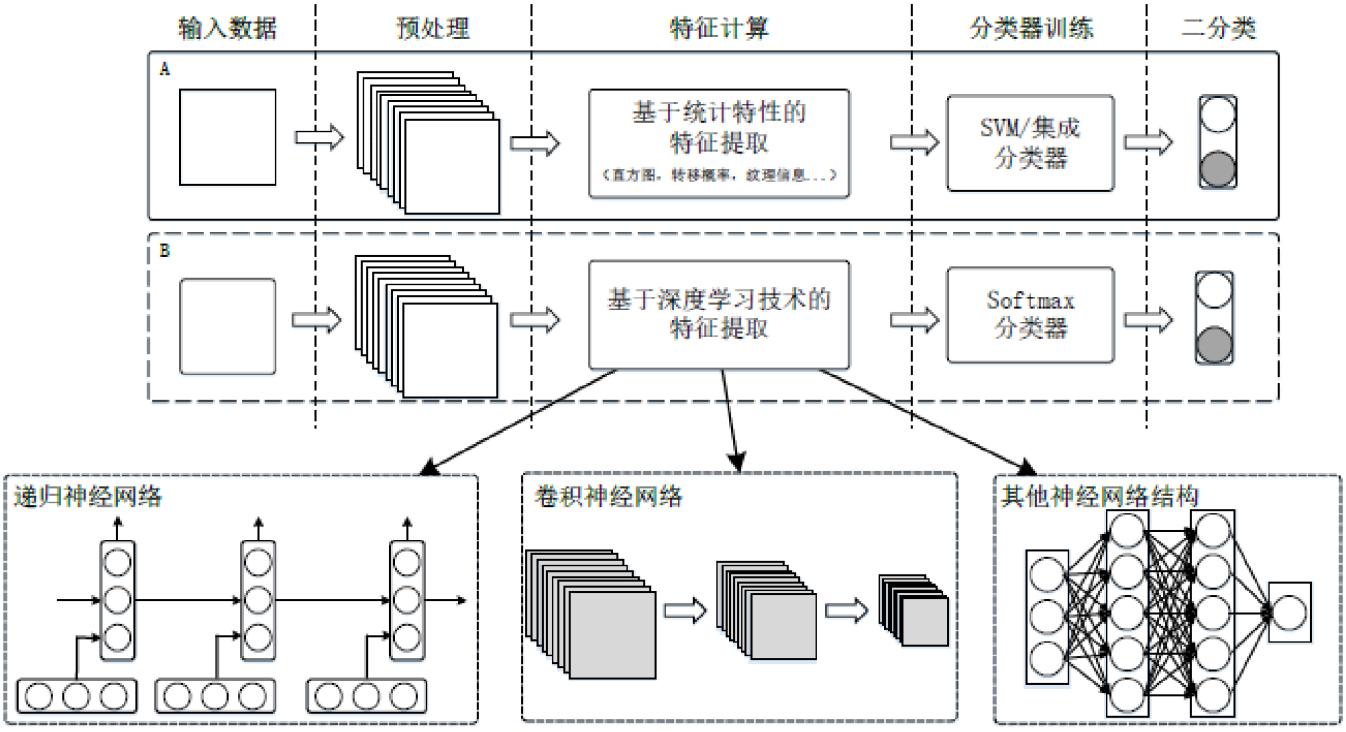
人工智能，机器学习，神经网络，深度学习，四者之间的关系可见如下图：




隐写分析的本质任务是**分类任务**，它的目标就是寻找最优的隐写统计特征表达式，从而更好地区分**正常载体**和**隐写载体**。所以，深度学习技术也是可以应用于隐写分析的。

## 2.2 两种隐写分析框架

基于**卷积神经网络**和基于**手工特征设计**的隐写分析方法，两种框架的基本部件是相似的，其中最本质的区别就在于隐写特征的表达。基于卷积神经网络的隐写分析方法可实现对特征空间的**自动寻优**；基于手工特征设计的隐写分析方法则需要较强的专业知识和设计经验。这两类隐写分析框架的比较，可见如下：



随着**自适应隐写术**等现代隐写技术的发展，基于手工特征设计的隐写分析方法变得愈发困难。实际上，深度学习技术已经在隐写分析中得到了很好的运用，检测效果已经超过了基于手工特征设计的传统隐写分析方法。

基于深度学习的音频隐写分析就是将如何计算隐写分析特征问题，转化为深度学习网络结构的设计问题。设计一个好的深度学习网络模型用于隐写分析检测，其实就是一个基于大样本驱动的**迭代**过程，然后利用反馈结果不断改进网络模型。遗憾的是，当前深度学习模型的构建仍然缺乏完备的理论体系，整个网络构成也依赖于具体的分析对象。

## 2.3 网络模型基础部件

（1）**批量标准化层**。由Google公司提出的一种训练优化方法，目的是缓解输入数据的**协方差偏移**现象，从而降低梯度弥散的风险。如果批次训练样本的分布特性不同，那么神经网络还需要在迭代时重新学习不同的分布，这个过程将影响网络的训练效率。所以我们需要提出多种训练数据的标准化方法。

（2）**激活函数**。 可以引入非线性因素，提升网络对各类函数的表达能力。

（3）**池化层**。对特征图进行**降采样**，一方面可以减少特征图的维度，降低运算复杂度；另一方面可以对特征进行压缩，保留其**鲁棒性**。常用的池化类型包含：均值池化，最大池化，随机池化和卷积池化。

（4）**正则化**。在机器学习中，通过显示控制模型复杂度来避免模型过拟合，确保其泛化能力。正则化在损失函数中引入模型复杂度指标，强化训练数据中的噪声。

## 系列文章

音频隐写术：结合“熵”理解隐写算法的具体步骤_唠嗑！的博客-CSDN博客

音频隐写术：两种具体的实现方法_唠嗑！的博客-CSDN博客

音频隐写术：分析剑桥大学提出的MP3Stego算法_唠嗑！的博客-CSDN博客

隐写术基础_唠嗑！的博客-CSDN博客

音频隐写术总结篇（附隐写软件下载链接）_唠嗑！的博客-CSDN博客