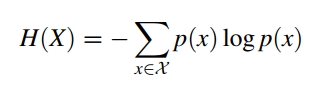
《**Formal Specification for Deep Neural Networks**》

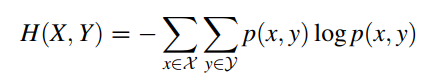
First, we present a semantic classification of properties, based on their meaning and relevance for  
the verification of systems based on deep neural networks. Second, we present  
a trace-theoretic classification, where we take the standard view of properties  
defined using sets of traces, and discuss how the various properties fit into those  
categories.

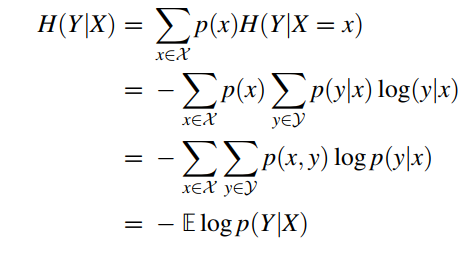
《Weighted pooling for image recognition of deep convolutional neural networks》

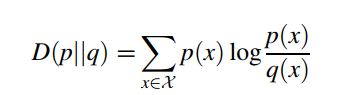
该论文主要是指出了一种新的池化方法。之前的池化方法如max-pooling，average-pooling或者stochastic-pooling都不太好，而weight-pooling采用信息熵的相关知识，来确定池化区域具有代表性的激活值所包含的信息量。

通过激活值贡献来消除预测的不确定性。通常通过选择具有代表性的激活值，其权值往往优于激活值。

信息熵的定义

联合概率分布

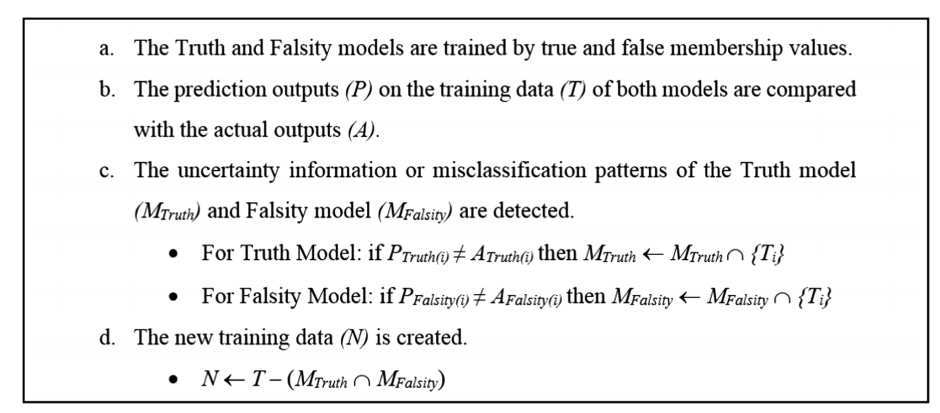
条件熵

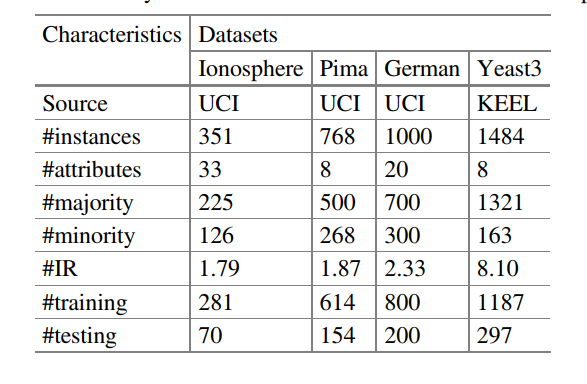
相对熵

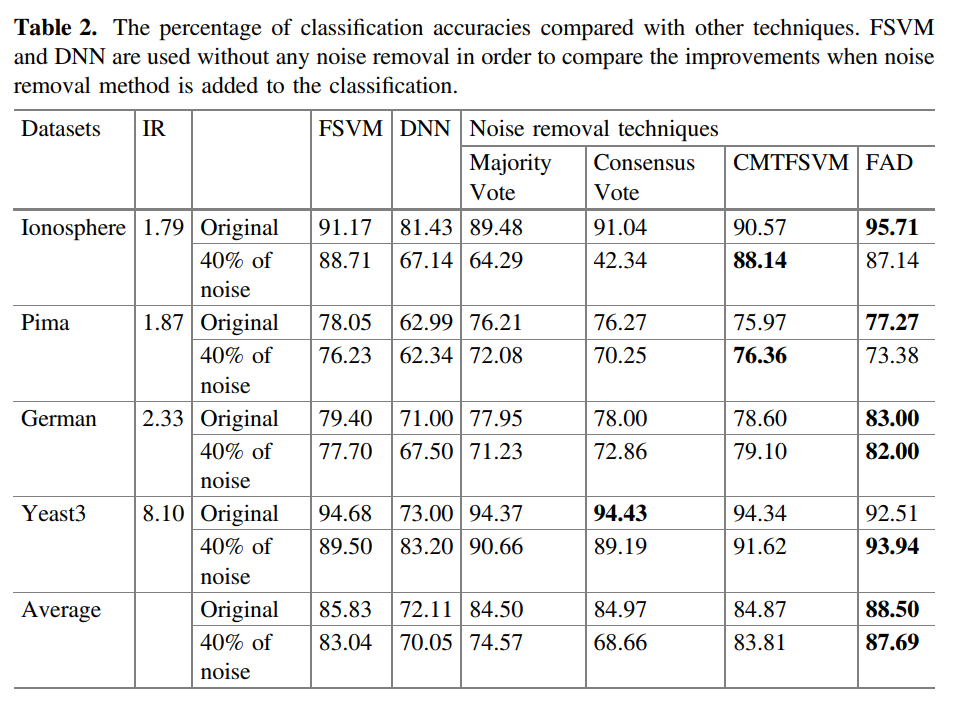
总之，通过一系列公式推导得到权重池化方式由于其他池化方式，并通过实验证明了其有效性和性能的提升。

《The Fuzzy Misclassification Analysis with Deep Neural Network for Handling Class Noise Problem》

本文主要提出了一种模糊错误分类检测方法。其核心思想是以二分类为例，训练一个预测真模型和一个预测假模型，然后将真模型预测错误样本Mtruth和假模型预测错误样本Mfalse样本的交集从训练数据集中减去，用剩下的样本去训练模型，从而实现去除错误样本。

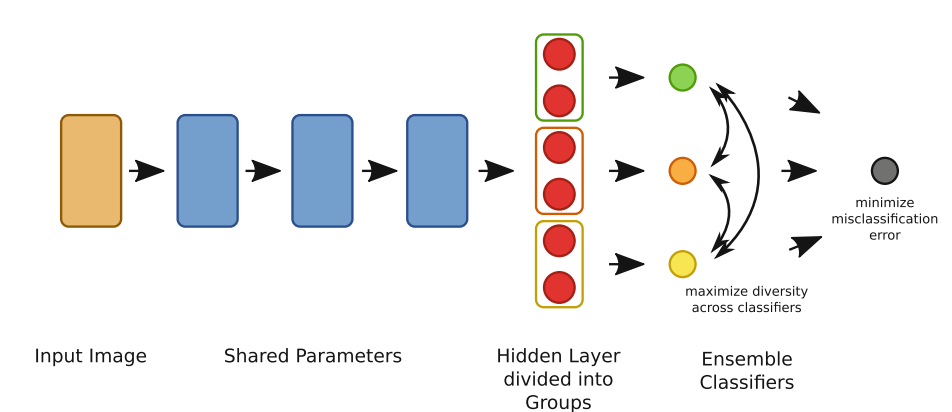


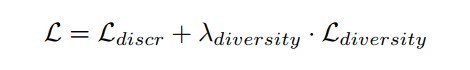


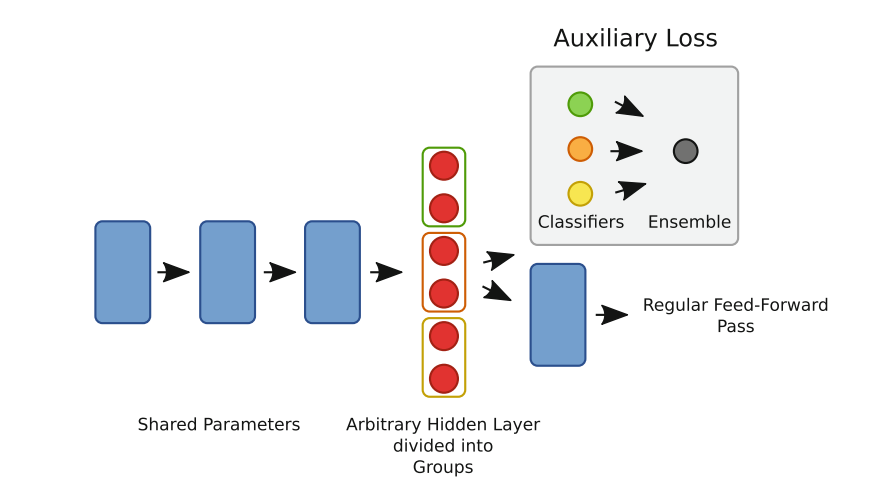


《Efficient Model Averaging for Deep Neural Networks》

传统的网络集成方式有bag或者boots，staking方法，或者通过random dropout来获得多个网络的集成，但是这样的网络差异性不大，而且还要训练多个网络是十分耗费计算能力的，所以提出一种新的网络继承方式，能够扩大网络之间的差异，主要是通过设计新的损失函数来实现的。



损失函数如下：即误分类损失函数加上差异性损失函数。

中间辅助损失函数可以将其放置在其他的隐藏层。通常是放在池化层之上。

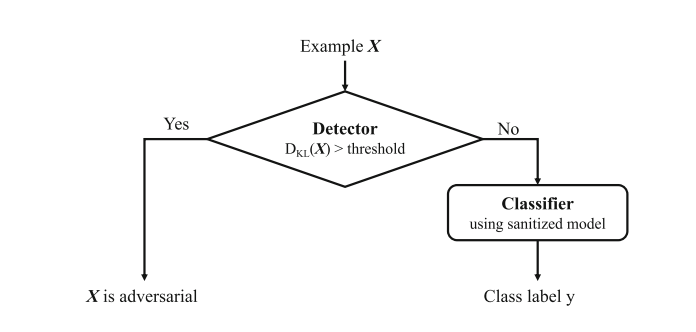
《Less is More: Culling the Training Set to Improve Robustness of Deep Neural Networks》

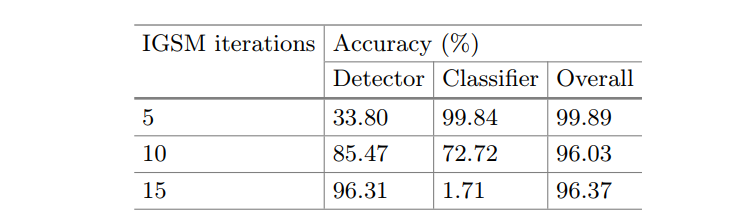
过滤掉噪声因素，即异常，恶意数据。主要有两种消毒方式，1self-sanitization和concanical-sanitization方法。

Concanical-sanitization是使用在有典型数据的时候，用典型数据选出训练集中的异常值并将其剔除，

Self-sanitization是用在没有典型值的时候采用训练集训练一个模型，然后用这个模型将预测结果差的作为异常值直接剔除。

然后用剔除异常值的数据集训练sanitization模型和完整的训练数据集训练origin模型。然后通过两者预测结果的KL散度值来确定测试样本是恶意数据还是真实数据。

加入检测器的框架。

抵抗恶意数据攻击后的准确度。