

# 周报

---

冯浩哲

2018.8.26

---

## 周报

本周工作

秋学期工作安排

## 本周工作

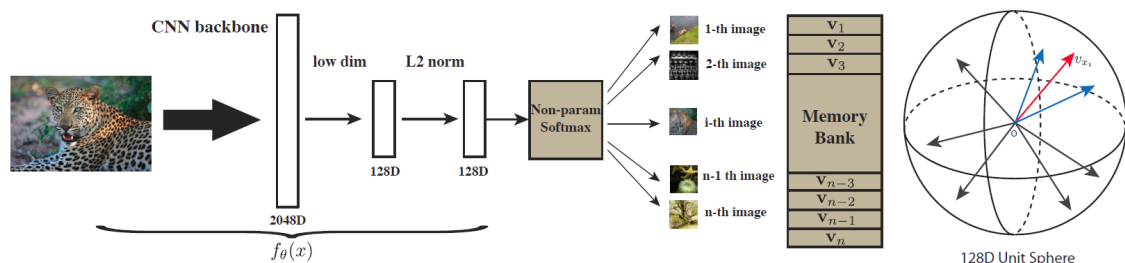
1. 在郝老师指导下进行论文修改任务
2. 将无监督推荐标注系统应用于肺结节良恶性检测任务

本周我们将原来在肺结节分割任务上所进行的由深度自动编码器与聚类算法组成的无监督推荐标注系统迁移到肺结节良恶性检测任务上。

实验结果表明我们的系统可以用40%的数据达到分类精确度为86.6%的结果，比起用100%数据进行训练达到分类精确度为87.2%的结果而言，我们减少了60%的标注量同时仅损失了0.6%的精确度，但是随机选40%的数据达到的分类精确度也达到了85.6%，因此我们的系统仅比随机选策略精确度高1%，这个结果仍然需要改进。

### 3. 将CVPR2018的论文"Unsupervised Feature Learning via Non-Parametric Instance Discrimination"应用于无监督推荐标注系统

本周我们将CVPR2018的论文"Unsupervised Feature Learning via Non-Parametric Instance Discrimination"中阐述的无监督特征提取方法应用在我们的无监督推荐标注系统中。这个特征提取方法主要用于图像相似度的度量，从而构造用于分类的图像关系矩阵，它的基本思想是用深度学习技术直接将输入的n个数据分为n类，从而提取能够对实例进行区分的特征，如果有2个数据所提取的特征很接近，则说明这两个数据很难区分，彼此相似度很大，其基本框架如图所示：



我们用该论文的方法所提取的特征在肺结节数据集上直接进行基于k近邻算法的简单分类，分类精确度达到了82%，而用我们原来采用的深度自动编码器，分类精确度仅达到73%，说明了用CVPR2018论文的方法提取的特征更加有效。

## 秋学期工作安排

我们将在秋学期以应用于肺结节良恶性分类任务的无监督推荐标注算法为核心课题，依次处理以下四个问题：

1. 首先使用CVPR2018文献的特征提取方法代替深度自动编码器提升无监督推荐标注算法的效果。
2. 其次改进聚类算法与代表性样本选取方法，使其与新的特征提取方法相适应。
3. 考虑到CVPR2018文献的特征提取方法所提取的特征在discrimination角度很有效，但是在explainable角度可解释性不强。而自动编码器所提取特征的可解释性很强，因此我们可以尝试将自动编码器的思想与CVPR2018文献的特征提取方法相结合，使得提取的特征既有区分度，又具有一定的可解释性。
4. 我们现在所提出的无监督推荐标注算法需要训练两个网络，一个是无监督特征提取网络，另一个是完成分类任务的网络，我们可以尝试将两个网络合并，如用无监督特征提取网络所提取的特征作为分类网络的输入，从而快速完成分类任务。

同时，我们也将大量阅读与无监督特征提取，图像深度聚类方法以及图像相似度度量相关的文献，将更多方法应用在我们的无监督推荐标注系统中提升其效果。