

# Semi-supervised Multi-view Deep Discriminant Representation Learning (SMDDRL)

## 摘要部分

贡献一：

现有的组合或者说对齐多视图的方法不能同时利用一致性和互补性两个原则来学习视图间共享表示和视图内的特定表示，SMDDRL通过使用共享的和特定的表示学习网络，全面地利用共识和互补的特性，学习共享的和特定的表示

贡献二：

现有的共享-特定的多视图表示学习方法忽略了表示学习中的冗余问题，SMDDRL结合了正交性和对抗相似度约束来减少学习到的表示中的冗余

贡献三：

设计了结合“深度度量学习”和“密度聚类”的半监督学习框架；

## 第一章 引言

描述了半监督多视图学习对于人类活动和科研的现实意义；

阐述了多视图学习的三个特征：

1. 优势：（1）一致性和（2）互补性
2. 缺点：（3）多视图数据冗余

多视图学习的关键：有效利用优势的同时也要处理冗余

### 1.1 动机

半监督和监督的多视图学习可以被分为两种范式：联合表示、对齐表示。其中联合表示是将从不同视图学习到的表示融合在一起，对齐表示是通过对齐使不同视图所得表示的一致性最大化。

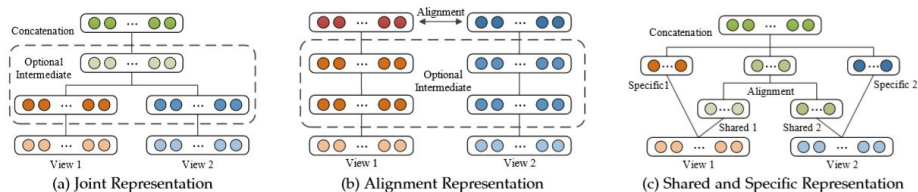
对齐表示有两种方式：

1. 基于CCA：最大化表示之间的相关性
2. 基于距离：最小化表示之间的距离

联合表示优缺点：方法简单、高效，但没有考虑一致性或冗余性，导致特征表示繁琐（特征空间维度高）。

对齐表示优缺点：利用了一致性属性对不同视图进行对齐。但是，特视图内的独有信息会干扰对齐，并在此过程中被消除，从而导致互补性的缺失。

因此基于共享和特有的方法被提出：这种方法倾向于解开共享信息和特有信息，然后在对齐的时候只对共享信息进行对齐，如下是对上述三种对齐方式的图示：



基于共享和特有的方法是一个新的研究方向，几乎没有研究考虑共享和特有信息之间的冗余并主动将这两种信息分离，并且有多项研究工作表明共享和特有信息不会在未施加约束的情况下自动分离，他们会自然地污染特征并且降低学习性能；

阐述为什么使用深度学习和半监督学习并简单描述所提出模型的优势：1. 综合利用了一致性和互补性优势，避免了特征的繁琐和冗余；2. 使用无标记数据提高了学习性能；

## 1.2 本文贡献

1. 研究了多视图表示学习领域，提出了一种新的分类方法。这种分类法填补了现有主流分类法关于出现的共享和特有信息方法的空白。
2. 提出了多视图深度判别表示学习(MDDRL)方法。MDDRL通过设计的正交损失和对抗相似约束，综合利用了多视图数据的一致性和互补性，减少了学习到的表示的冗余性。MDDRL是第一种多视图表示深度学习方法，它考虑了多视图数据的所有一致性、互补性和冗余特征。
3. 深度度量学习和密度聚类相互协作，利用未标记数据提高学习性能，构成了SMDDRL框架；
4. 实验验证了SMDDRL相比于其他多视图表示学习方法的优越性以及实验验证了本框架模型中各个组件的有效性；

## 第二章 相关工作

### 2.1 经典的多视图特征学习工作

概述了几个典型的多视图相关的工作

### 2.2 深度多视图表示学习

深度多视图表示学习也遵循上述三种对齐范式

跨视图分类的深度多视图网络：

这个网络首先通过子网络提取特定视图的特性，然后串联并将这些特性输入到一个公共网络中，该公共网络的设计目的是将它们投射到一个统一的空间中。

还介绍了两项研究工作，它们都遵循：对单个视图使用单独的网络，再将提取到的单视图的特征输入一个统一的网络中；

本文提出的框架是在采用了深度神经网络来处理视图的特征学习，并且没有抛弃多视图学习的三个特性（一致性、互补性和冗余视图）

## 2.3 半监督学习工作

### 2.3.1 通用的半监督学习

半监督学习的优势：解决缺少标记数据的问题

协同训练、伪标签标记、相似图、基于本地平滑正则化的方法；不理解后两种的思想

### 2.3.2 半监督多视图学习

基于协同训练、基于图、基于特征对齐，第一种和第三种较为常用；

介绍了几种基于上述方式的半监督多视图学习工作；

与所有这些方法不同的是，本文通过结合深度度量学习和密度聚类，提出了一个伪并行的SSL框架。这个框架为使用未标记数据提供了一种新的选择。

如何理解这里的伪并行？

## 第三章 半监督多视图深度判别表示学习(SMDDRL)

重述多视图表示学习的关键：恰当处理冗余的同时，充分利用互补性和一致性；

SMDDRL包含如下两部分：

1. 多视图深度判别表示学习
2. 基于半监督的深度度量和密度聚类学习框架

### 3.1 多视图深度判别表示学习

以基于表示学习的分类骨干为基础，增加了深度度量学习和正交对抗损失来提高性能

损失函数如下：

$$L = L_C + \lambda_1(L_{diff} + L_{Adv}) + \lambda_2 L_{Con},$$

第一项是分类损失的加权和，第二项是约束损失（正交对抗损失），第三项是深度度量学习的对比损失

#### 3.1.1 共享-特有表示学习

交叉熵的意义和推导以及多分类交叉熵：[交叉熵](#)

描述了简单的基于共享-特有的表示学习：多个视图的共有信息取平均、特有信息直接连接；

### 3.1.2 正交和对抗相似损失

3.1.1 所描述的方法倾向于分别处理组间共享和组内特有的信息，但是这两种信息不是自动（或者说主动）分离的；

正交约束和对抗相似约束：人工分离共享信息和特有信息，减少两种信息之间的冗余（尽量让两种信息没有交集）

其中：

1. 正交性约束保证来自同一个视图的两种信息的分离
2. 相似对抗约束保证来自不同视图的共享信息的相似性

正交性约束损失

$$L_{diff} = ||S^{k^T} H^k||_F^2$$

$S^{k^T}$ 、 $H^k$  分别为来自视图  $k$  的共享和特有信息；

采用Frobenius范数来度量两个信息矩阵的相似性，这个相似性越大，那么正交性越弱，正交损失越大；

对抗相似性约束损失（想法来自于生成对抗网络）

$$P_i^k = D(G_k(X_i^k))$$

$$L_{Adv} = \min_{\theta_G} \max_{\theta_D} \left( \sum_{i=1}^N \sum_{k=1}^M l_i^k \log D(G_k(x_i^k)) \right),$$

M: 视图数量

N: 样本数

G: 作为生成器，其生成结果是来自该视图的共享信息

从样本生成一个欺骗共享信息的具体做法？

D: 判别器，判断某一个共享信息来自于视图  $k$  的概率（是否来自  $k$ ），对于多分类，判别器给出的是  $k$  维的概率向量，最终通过乘以真实标签，来作为去视图识别的正确率

来自视图  $k$  的概率的具体计算方法？

回答上述问题：

体现在训练中的做法应该是：

step1: 将样本输入生成器网络  $G$ ，该网络可以是任何提取共享信息的网络，然后输出的共享信息可以作为“欺骗共享信息”

step2: 将“欺骗共享信息”作为判别器网络D的输入，该网络是一个完成多分类任务的网络，判断输入的“欺骗共享信息”来自某个视图的概率，根据视图标签的真实值可以将概率向量转化为标量，D的目标是最大化这个概率值；这一步重复多次，以更新D的参数；

step3: 再次将样本输入生成对抗网络D(G)，但是这一次是通过最小化D(G)的判别损失，来优化G的欺骗性；这一步重复多次，以更新G的参数；

多次训练后，D已经无法区分不同视图产生的共享信息，G将可以生成极具欺骗性的共享信息，换句话说G提取的不同视图之间的共享信息已经非常相似了，即G已经是一个优化的共享信息提取器；

### 3.1.3 深度度量学习的孪生网络

孪生网络是度量两个样本相似度的通用框架

对于两个输入样本  $x_i$  和  $x_j$ ，将其表示为向量  $h_i$  和  $h_j$ ，然后投影到低维的共享子空间后，它们被表示为  $code_i$  和  $code_j$ ，我们通过平方欧氏距离来度量他们的相似度：

$$d(x_i, x_j) = ||code_i - code_j||^2$$

对比损失定义为如下：

$$L = \frac{1}{2N} \sum_{n=1}^N [y_n d_n^2 + (1 - y_n) \max^2(\text{Margin} - d_n, 0)],$$

其中当两个样本来自同一个类的时候， $y_n$  取 1，来自不同类的时候取 0；

显然当两个样本来自同一个类的时候，他们距离越大，则相似度越小，损失越大，当两个样本来自不同类的时候，他们距离越小相似度越大，损失越大；

基于样本随机选取的缺陷，对上述对比损失进行改进，采用 mini-batch 进行损失计算，损失函数如下：

$$L_{Con} = \frac{1}{2m} \sum_{i=1}^m \left\{ \begin{aligned} &(x_i - \mu_{same})^2 + \\ &\max^2[\text{Margin} - (x_i - \mu_{diff}), 0] \end{aligned} \right\}$$

$\mu_{same}$  为与  $x_i$  具有相同标签的样本集合的中心， $x_i$  与该中心距离越大，则损失越大；

$\mu_{diff}$  为与  $x_i$  具有不同标签的样本集合的中心， $x_i$  与该中心的距离越小，则损失越大；

优点：mini-batch 可以增加结果的稳定性，“中心”的思想可以加快收敛速度；

## 3.2 基于半监督学习框架的深度度量学习-密度聚类

为什么这样组合？

半监督学习最近发展很快，同时伪标签风格的半监督学习因为已经提供了标签，所以更加适用于深度学习；

深度度量学习可以将属于同一类的样本拉近（在特征空间中），将属于不同类的样本拉远，也就是说可以使不同类之间的类边界更加清晰（类边界处的样本密度更低）；而密度聚类的假设便是：类边界处的样本密度更低；

因此，使用深度度量学习-密度聚类是一种自然的算法结合（使用了密度峰聚类算法）；

### 3.2.1 半监督密度聚类过程

step1: 使用已标记数据训练深度度量学习器MDDRL

step2: 使用MDDRL对已标记和未标记数据进行编码

step3: 使用密度聚类对所有样本进行聚类

step4: 根据聚类结果对未标记数据进行标记

其中，在step4中，使用一个分类网络对样本进行分类，只信赖分类和聚类标签一致的结果（只标记这一部份），然后将所有样本作为数据再次训练；

直到验证精度不再改善或者所有数据得到标记则算法停止；

### 3.2.2 实现细节

1. 深度度量学习、密度聚类、分类网络使用同一个特征空间，其距离度量方式均为平方欧氏距离
2. 原始标记数据保持不变

## 后续

对比实验：

与当前最优算法相比；

测试各个组件功能的实验；

最优超参数的实验；

以及研究前景