# 说明文档

* 1. 问题背景

在减少标记需求方面，主动学习方法和迁移学习方法是常用的两种策略，但这两种策略均有相应局限：主动学习对样本标记的需求依旧较大，实际中很难适用于解决小样本分类问题；迁移学习局限于缺乏目标域关键信息以及域间分布的差异性。

* 1. 解决方法

针对以上问题，基于主流的主动学习和迁移学习方法分析，做了两个方面的工作。首先，提出了一种基于单主动学习的域适应算法，将极大极小熵的域适应迁移学习算法和核心集采样的主动学习算法相融合，实现了在小样本条件下对目标域关键样本的选取，大大减少了标记样本所需的成本。其次，为解决无效源域样本问题，本文在源域上采用损失预测主动学习策略，提出了一种基于双主动学习的域适应算法，借助选取出的已知目标域样本，实现了对源域样本的有效筛选，一定程度上消除了负迁移影响。

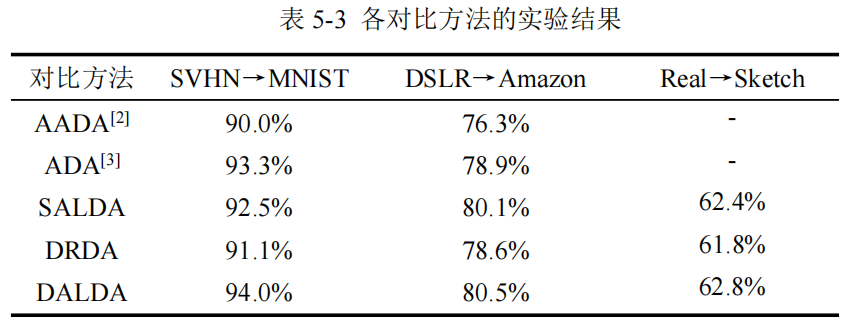
* 1. 实验结果

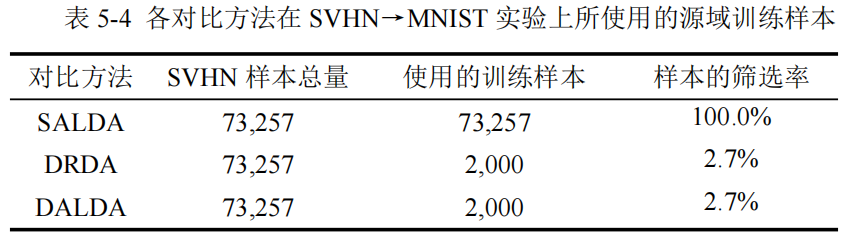
本文将近几年域适应，主动域适应等典型方法与所提的基于单主动学习的域适应算法**SALDA**以及基于双主动学习的域适应算法**DALDA**进行如下对比实验，以验证所提方法的有效性。

1. MME[15]：是一种典型半监督域适应方法，要求直接提供目标域每个分类1个或是3个标记好的样本，随后采用基于直推式的学习方式进行训练，即学习过程中使用的未标记样本恰是待预测样本。
2. MME\*[15]：和MME相比，MME\*是采用纯半监督的学习方式进行训练，其他则与MME保持一致。
3. AADA[2]：是一种典型的主动域适应算法，通过域适应和对抗性统一训练实现在目标域上挑选样本的主动对抗性域适应方法。
4. RADA[2]：和AADA相比，RADA采用随机采样的方式取代主动学习采样，其他则与AADA保持一致。
5. ADA[3]：也是典型的主动域适应算法之一，将不确定性，多样性主动查询策略融合于域适应的算法。
6. RDA[3]：和ADA相比，RDA采用随机采样的方式取代主动学习采用，其他则与ADA保持一致。
7. SALDA：本文所提的基于单主动学习的域适应算法，通过核心集主动学习和MME域适应相结合的方式，弥补了MME对样本标记的高要求问题，以实现目标域的主动挑选。
8. SALDA\*：和SALDA相比，SALDA\*采用MME\*取代MME，其他则与SALDA一致。
9. SRDA：和SALDA相比，SRDA采用随机采样的方式取代主动学习采样，其他则与SLDA保持一致。
10. DALDA：在SALDA的基础上，结合损失预测主动学习，以实现源域的主动筛选。
11. DRDA：和DALDA相比，DRDA采用随机采样的方式取代损失预测主动学习采样，其他则与DALDA保持一致。

### 基于单主动学习的域适应方法有效性验证及分析

### 基于双主动学习的域适应方法有效性验证及分析



* 1. 在投论文

计算机学报

* 1. 代码运行

各组实验运行参数如下表所示

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 实 验 |  |  |  |  |  |  |  | Step | Net |
| SVHN→MNIST | 0.1 | 10 | 50 | 59830 | 73257 | 500 | 2000 | 50000 | LeNet |
| DSLR→Amazon | 0.1 | 50 | 200 | 1585 | 498 | 124 | 248 | 12500 | ResNet |
| Real→Sketch | 0.1 | 126 | 378 | 15632 | 70358 | 16800 | 50400 | 12500 | ResNet |
| Sketch→Painting | 0.1 | 126 | 378 | 20246 | 24582 | - | - | 12500 | ResNet |
| Painting→Clipart | 0.1 | 126 | 378 | 11713 | 31502 | - | - | 12500 | ResNet |
| Clipart→Sketch | 0.1 | 126 | 378 | 15632 | 18703 | - | - | 12500 | ResNet |

执行单主动迁移学习时，运行python s\_ac\_r\_s.py命令，执行双主动迁移学习时，取消r\_s\_ac\_t1.py中的源域主动学习模块注释（P100-P101），再执行python s\_ac\_r\_s.py命令即可。

* 1. 数据/模型/环境

数据集对应data.tar压缩包；模型对应trained\_model.tar压缩包；环境对应ac\_t.tar压缩包。

* 1. 解决方法对应代码位置

单主动迁移学习对应r\_s\_ac\_t0.py文件，涵盖了整个迁移学习，其中目标域主动学习模块对应在第495-501行；双主动迁移学习对应r\_s\_ac\_t1.py文件，其中源域主动学习模块对应在第100-101行。