

# ACPL: Anti-curriculum Pseudo-labelling for Semi-supervised Medical Image Classification

## 摘要

医学图像分析（MIA）中的有效半监督学习（SSL）必须解决两个挑战：1）有效地处理多类（如病变分类）和多标签（如多种疾病诊断）问题，以及2）处理不平衡学习（由于疾病流行率的高方差）[1]。SSL MIA中要探讨的一种策略是基于伪标签策略，但它有一些缺点。伪标签一般比一致性学习的精度低，它不是专门为多类和多标签问题设计的，并且它可能受到不平衡学习的挑战。在本文中，提出了一种新的SSL算法，称为反课程伪标签（ACPL），该算法引入了新的技术来选择提供有用信息的未标记样本，改善了训练平衡，并允许模型同时适用于多标签和多类问题，并且通过分类器的精确集成来估计伪标签（提高伪标签精度）。**关键词：**有效半监督学习（SSL），反课程伪标签（ACPL）

## 1 引言

**选题背景：**深度学习在医学图像分析（MIA）中显示了出色的结果。与计算机视觉相比，医学专家对MIA训练集进行标记的成本明显更高，导致标记图像的可用性较低，但来自诊所和医院数据库，可用于深度学习分类器建模未标记图像的使用率较高。此外，与倾向于多类别和平衡的计算机视觉问题不同，MIA具有许多多类别（例如，单个类别的病变图像）和多标签（例如，来自患者的图像可以包含多种疾病）问题，其中由于疾病的可变流行率，这两个问题通常包含严重的类别不平衡。因此，除了处理不平衡学习之外，MIA半监督学习（SSL）方法还需要足够灵活，以处理多标签和多类问题。[2]

**选题依据及选题意义：**该论文解决了医学图像分析（MIA）中的半监督学习可能难获取不平衡的多类和多标签的准确标签问题。提出了一种新的半监督学习算法，不需要估计分类阈值就能选择样本，可以适应多类和多标签问题，有助于平衡训练。有效地满足了医学分析中对半监督学习的需求。

## 2 相关工作

### 2.1 基于一致性的SSL方法

优化了标记图像的分类预测，并最小化了未标记图像的不同视图的预测输出，其中这些视图是从不同类型的图像扰动中获得的，例如空间/时间、对抗性或数据增强。基于一致性的

方法的性能可以通过自我监督的预训练进一步提高。尽管基于一致性的 SSL 方法在许多基准测试中显示了 SOTA 结果，但它们依赖于需要领域知识的扰动函数的仔细设计，并且需要适应每种新型医学成像。此外，伪标签 SSL 方法对于多标签问题更准确。[3]

## 2.2 伪标记SSL方法

使用可用的标记数据训练模型，估计以高置信度分类的未标记样本的伪标签，然后取这些伪标记样本重新训练模型。伪标签 SSL 方法可能会使分类偏向于不平衡问题中的大多数类，不能无缝地适用于多类和多标签问题，并且导致确认偏差。伪标签SSL方法的改进取决于选择信息丰富的未标记样本以解决大多数类偏差和适应多类和多标记问题，以及处理确认偏差的精确伪标签机制。

### 3 本文方法

### 3.1 本文方法概述

本文提出了反课程伪标签（Anti-curriculum Pseudo-labelling: ACPL），它解决了多类和多标签不平衡学习SSL MIA问题。介绍了我们的ACPL优化，该优化按照反课程策略为未标记样本生成准确的伪标签，其中，在每个训练阶段选择信息量大的未标记样本进行伪标记。使用提出的信息混合（IM）方法生成信息性未标记样本的伪标签，该方法混合了模型的预测和使用锚集的Knearest neighbor（KNN）分类器的结果。在每个训练阶段结束时，使用锚集净化（ASP）方法更新锚集，该方法仅保持根据CDSI标准，伪标签样本的信息量最大的子集。

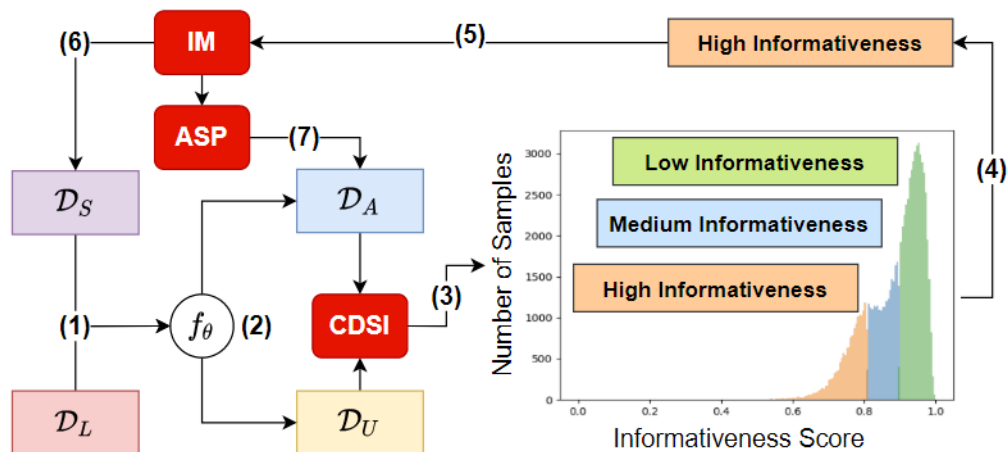


图 1. 方法示意图

### 3.2 交叉分布样本信息量 (Cross Distribution Sample Informativeness: CDSI)

估计未标记样本是否具有高信息含量的函数定义如下:

$$h(f_{\theta}(\mathbf{x}), \mathcal{D}_A) = \begin{cases} 1, & p_{\gamma}(\zeta = \text{high}|\mathbf{x}, \mathcal{D}_A) > \tau, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

输出未标记样本与标记的锚点样本集合的接近程度（锚点样本是高信息性的已标记样本）

### 3.3 信息混合（Informative Mixup: IM）

信息混合：将模型的结果与使用锚点的K近邻分类器融合。使用信息混合生成信息性未标记样本的伪标签。每个训练结束，锚点集都使用锚点集合纯化方法 ASP 更新——根据 CDSI 准则，每次只保留伪标记样本中最具信息量的子集

## 4 复现细节

### 4.1 与已有开源代码对比

参考了本篇论文的源代码：<https://github.com/FBLADL/ACPL>。使用了代码来实现 ACPL 算法，在后续的实验基于 Chest X-Ray14 数据集的胸部疾病分类结果实验中使用到了此源码。但是此源码并不完整不足以支撑全部的实验，而且并不能很好地运行，我的工作首先是对源码的部分进行修改，使其能正常运行。之后实现了能进行消融实验的代码，例如，比较 CDSI 是否对准确率有影响的实验，源码只能选择高信息量的无标签数据，无法进行消融实验对比。复现代码时实现了对低信息量无标签数据的选择，以此来复现消融实验，

### 4.2 实验环境搭建

python3.7，训练模型时使用是两个 NVIDIA RTX 2080ti

### 4.3 界面分析与使用说明

启动程序的命令如图2，可以自己设置一系列的参数，例如，迭代的轮次可以通过 `--epochs` 来进行设置，后续消融实验所需要的参数可以通过设置 `--sel` 等等。

```
CUDA_VISIBLE_DEVICES=0,1 python main.py --data /fuxian/ACPL/dataset --task cx14 --resize 512 --batch-size 8 --epochs 20 --dist-url 'tcp://localhost:10001' --multiprocessing-distributed --world-size 1 --rank 0 --desc ACPL_2_2_k50 --num-workers 4 --eval-interval 100 --reinit --label_ratio 20 --runtime 2 --topk 50 --pl-epochs 10 --ds-mixup --sel 2 --num-gmm-sets 3 --lr 0.03
```

图 2. 启动程序命令

### 4.4 创新点

新的信息含量度量方法——交叉分布样本信息性：选择用于伪标记的最具信息量的未标记数据，有效的学习课程必须关注于离标记样本分布尽可能远的信息性未标记样本。选择这些样本，将均衡训练并消除估计类分类阈值的需要。未标记样本的信息含量度量，将由我们提出的交叉分布样本信息性 CDSI 计算，输出未标记样本与标记的锚点样本集合的接近程度。信息混合 IM：提出了一种信息混合方法，该方法构造了伪标签函数，将模型对无标签数据的

预测和KNN方法的预测线性混合。ASP（锚集净化）：在估计信息性未标记样本的伪标签后，我们的目标是用信息性伪标签样本更新锚集。然而，添加所有伪标签样本将导致锚集尺寸过大，并增加超参数敏感性。因此，我们建议锚集纯化（ASP）模块选择要插入锚集中的连接最少的伪标签样本

## 5 实验结果分析

本部分对实验所得结果进行分析，详细对实验内容进行说明，实验结果进行描述并分析。对于这两个数据集，我们使用DensNet-121作为主干模型。对于Chest X-Ray14，数据集预处理包括将图像大小调整为512×512，以加快处理速度。对于优化，使用Adam优化器、批量大小16和学习率0.05。在训练期间，使用基于随机裁剪和调整大小以及随机水平翻转的数据增强。对于实验中使用到的KNN，K设置为50。实验开始时首先在初始标记子集上训练20个时间点，以预热用于特征提取的模型。然后我们训练50个时期，每10个时期，我们用ASP更新锚集。

实验首先对比不同方法在有标签数据占比不同的数据集中的表现。比较不同标记集训练百分比的Chest X-Ray14疾病类别的平均AUC测试集结果，可以看出本文的方法都比之前的方法表现要好。

Method Type	Label Percentage	2%	5%	10%	15%	20%
Consistency based	SRC-MT	66.95	72.29	75.28	77.76	79.23
	NoTeacher	72.60	77.04	77.61	N/A	79.49
	$S^2MTS^2$	74.49	78.96	79.90	80.31	81.06
Pseudo Label	Graph XNet	53.00	58.00	63.00	68.00	78.00
	UPS	65.51	73.18	76.84	78.90	79.92
	Ours	73.54	79.09	80.11	80.88	81.21

图 3. 不同标记集训练百分比的Chest X-Ray14疾病类别的平均AUC

显示了使用SSL方法的20%标记数据的类级性能，这表明本文提出的方法在Chest X-Ray14数据集14个类中的大多数类中实现了最佳结果。本文的方法可以使用ImageNet预训练模型来优于它，从而减轻了计算上昂贵的自监督预训练的需要。

Method Type	Supervised	Consistency based			Pseudo-labelling		
Method	Densenet-121	MT	SRC-MT	$S^2MTS^2$	GraphXNet	UPS	Ours
Atelectasis	75.75	75.12	75.38	78.57	71.89	77.09	79.02
Cardiomegaly	80.71	87.37	87.7	88.08	87.99	85.73	87.02
Effusion	79.87	80.81	81.58	82.87	79.2	81.35	83.44
Infiltration	69.16	70.67	70.4	70.68	72.05	70.82	70.29
Mass	78.40	77.72	78.03	82.57	80.9	81.82	82.21
Nodule	74.49	73.27	73.64	76.60	71.13	76.34	76.92
Pneumonia	69.55	69.17	69.27	72.25	76.64	70.96	73.22
Pneumothorax	84.70	85.63	86.12	86.55	83.7	85.86	84.21
Consolidation	71.85	72.51	73.11	75.47	73.36	74.35	75.21
Edema	81.61	82.72	82.94	84.83	80.2	83.56	84.21
Emphysema	89.75	88.16	88.98	91.88	84.07	91.00	92.61
Fibrosis	79.30	78.24	79.22	81.73	80.34	80.87	80.62
Pleural Thicken	73.36	74.43	75.63	76.86	75.7	75.55	76.31
Hernia	86.05	87.74	87.27	85.98	87.22	85.68	84.98
Mean	78.19	78.83	79.23	81.06	78.00	79.92	80.73

图 4. Chest X-Ray14不同疾病类别的平均AUC

消融实验研究探究采用高信息度的无标签数据进行伪标签对预测结果的影响。使用本文的CDSI对具有不同信息度（低、中、高）的未标记样本进行伪标签。从达到65%AUC的基线分类器DensNet-121开始，我们观察到采用信息含量低的未标记样本进行伪标签产生最差的结果，而选择高信息含量的非标记样本产生最佳结果。高信息含量的未标记样本能够构造较小的标记集（与低或中等信息含量的无标记样本相比），从而允许更有效的训练过程，产生更准确的KNN分类器。而且采用ASP更新锚点集，能够使KNN分类器的准确率上升，而且其标准差降低，使方法具有稳定性。

Information Content	ASP	AUC $\pm$ std
Baseline		65.87 $\pm$ 0.14
Low	$\times$	67.18 $\pm$ 2.40
	$\checkmark$	67.76 $\pm$ 1.05
Medium	$\times$	70.83 $\pm$ 1.49
	$\checkmark$	71.16 $\pm$ 0.51
High	$\times$	73.81 $\pm$ 0.75
	$\checkmark$	74.44 $\pm$ 0.38

图 5. ASP及信息量的消融实验

## 6 总结与展望

本部分对整个文档的内容进行归纳并分析目前实现过程中的不足以及未来可进一步进行研究的方向。反课程伪标记（ACPL）SSL方法。与使用阈值来选择可靠分类样本的传统伪标

签方法不同，ACPL使用了一种新机制来选择用于伪标签的高信息量未标记样本，并使用分类器集合来生成准确的伪标签。这使ACPL能够解决MIA多类和多标签不平衡分类问题。通过对比不同方法在不同数据集上的表现，得出本文提出的方法的性能较为优越，消融实验验证了提出的方法对模型分类精度的提升。目前实现过程中，所使用的KNN分类器的参数K，也就是如何确定合适的最近邻个数是通过先验知识和不断实验得出来的，如果要追求最好的性能，需要花费的时间够久，这点希望在之后的工作中能够解决。进一步的研究工作是将ACPL应用到更加一般的计算机视觉场景中去。

## 参考文献

- [1] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.
- [2] Fengbei Liu, Yu Tian, Yuanhong Chen, Yuyuan Liu, Vasileios Belagiannis, and Gustavo Carneiro. Acpl: Anti-curriculum pseudo-labelling for semi-supervised medical image classification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 20697–20706, 2022.
- [3] Fengbei Liu, Yu Tian, Filipe R Cordeiro, Vasileios Belagiannis, Ian Reid, and Gustavo Carneiro. Self-supervised mean teacher for semi-supervised chest x-ray classification. In *International Workshop on Machine Learning in Medical Imaging*, pages 426–436. Springer, 2021.