**“MICCAI 2023半监督牙齿分割”IGIP-CBCT队伍比赛介绍文档**

## ****创新思路：****

1. 我们借助相互学习的方法构建了一个全新的抗噪半监督框架。并根据测试集和标记数据集的特性，采用了一种由粗到细的策略来进行分割。但是，在两个网络协同训练的过程中，很容易学习到相似的信息，同时还受到耦合噪声的干扰。因此，我们采用了将数据集划分成两个不同的子数据集的策略，以增强输入网络所接触到的信息的多样性。
2. 我们观察到在学习过程中，模型的预测结果将逐渐变得比初始的噪声伪标签更准确。因此，在训练过程中，除了初始的伪标签，我们还引入了进一步的网络预测结果作为更加可信的伪标签，并逐渐增加这部分伪标签建立的损失约束的权重。
3. 我们注意到标记数据集中有些区域的标注存在误差（这是人工标注很难避免的问题），并且对于未标记的数据，由网络生成的伪标签往往充满噪声。为了应对这种情况，我们引入了噪声转移矩阵来进行损失校正，以有效减少噪声标签的影响。鉴于我们观察到易受错误影响的噪声区域主要位于边缘，我们将噪声转移矩阵应用于这些边缘区域，以避免对内部区域造成不必要的干扰。

## ****开发过程及方法细节：****

**一、训练流程分为三个阶段：**

第一阶段，我们使用了比赛方提供的12例标记数据集，并从未标记的数据集中选取了103个样本用作训练数据。训练模型采用传统的半监督分割方法Mean-Teacher[1]进行半监督训练，以获得初步的全视野分割结果。然后，我们利用已经训练好的网络对所有未标记的数据进行测试，生成这些未标记数据的伪标签。然而，由于标记数据集仅包含部分视野区域，我们观察到在这一阶段训练得到的伪标签质量较差，其中包括将大片背景区域错误地分割成前景区域的问题，即噪声散点问题。因此，我们进行了第二阶段的训练，以改善分割结果。

第二阶段，我们的目标是获得更准确的背景区域分割结果。为了实现这一目标，我们将第一阶段的未标记数据与生成的伪标签合并到标记数据集中，从而创建了一个新的数据集，用于进行第二阶段的训练。具体而言，我们使用了来自标记数据集的12个样本，并从未标记的数据集中选取了186个样本。需要注意的是，由于第一阶段生成的伪标签包含大量散点，因此本阶段采用了查找最大连通域的方法来清除这些散点。具体操作是，一旦网络获得了最大连通区域，就使用包含最大连通区域的边界框来过滤伪标签，仅保留包含在最大连通区域边界框内的伪标签。然后，我们使用经过处理的数据进行第二阶段的训练。接下来，经过训练的网络将对所有未标记的数据集进行测试，以进一步更新未标记数据的伪标签。

第三阶段，为了实现更精细的感兴趣区域分割，我们将第二阶段的未标记数据与生成的伪标签合并到标记数据集中，以创建一个新的数据集，用于进行第三阶段的训练。我们对未标记数据集进行了数据筛选，最终选择了281个未标记数据。在这一阶段，为了实现更加精细的分割，我们仅裁剪感兴趣的区域输入神经网络。

除了第一阶段采用的传统Mean-Teacher架构外，第二和第三阶段都采用了相同的网络架构，即我们新设计的半监督抗噪分割网络。整体的网络架构如图1所示，接下来，将详细介绍这一网络。

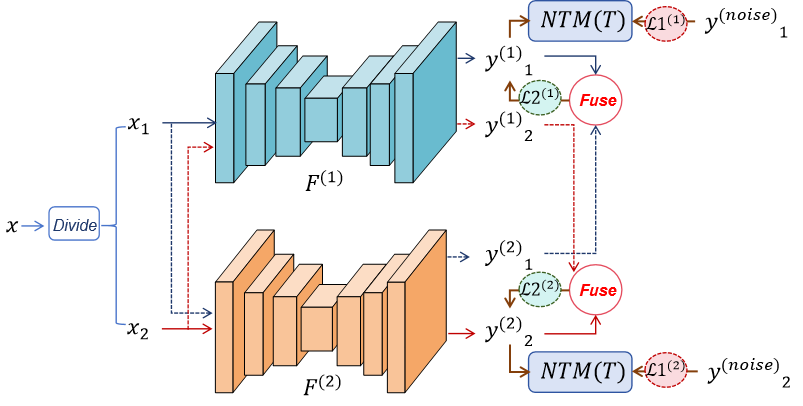


图1 网络架构

首先，我们采用了两个并行的网络和，使它们能够相互学习。但是，在两个网络协同训练的过程中，容易学习到相似的信息，并受到耦合噪声的干扰。为了解决这个问题，我们采用了将数据集划分为两个互斥的子数据集和的方法。将这些不同的子数据集分别输入到两个不同的网络中，以确保网络学习到更加多样性的特征。两个网络都可以采用噪声标签来做初始监督（这里我们把标记数据集的标签以及未标记数据的伪标签统一看做噪声标签）。那么这一过程可以用损失函数来约束：

(1)

其中表示为第个网络，表示为第个网络。

然后，我们将两个网络的预测结果融合，作为进一步的伪标签，并逐渐增加这部分伪标签的权重。这么做的目的是因为在网络学习过程中，预测结果会逐渐变得比噪声标签更准确，而且两个网络的预测结果集成后产生的伪标签会更加准确。融合网络的结果生成伪标签的过程如下所示：

(2)

并且利用这个伪标签建立的损失函数为：

(3)

那么总的损失函数就被概括为：

(4)

其中,随着迭代次数的增加而增加。而我们采用/2都是使用了常见的交叉熵损失函数和Dice损失函数,因此可以表示为：

/2 = (5)

此外，我们注意到标记数据集的标签存在一些误差，而对于未标记的数据，使用网络预测生成的伪标签无疑充满噪声。因此，我们采用噪声转移矩阵（noise transition matrix NTM）[3][4]进行损失校正，以减少噪声标签的干扰。噪声转移矩阵表示不同类数据点的干净标签转换为噪声标签的概率。我们假设生成的包含类的噪声标签可以通过一个噪声转移矩阵（NTM）被桥接到人工标注的Ground Truth，表示了Ground Truth标签的类翻转到有噪声的标签的第的概率，也就是：

我们观察到牙齿的边缘区域相对于内部更容易出现错误，因此更容易受到噪声标签的影响。在应用噪声矩阵时，将其应用于所有区域都是不合理的，因此我们将噪声转移矩阵设计在边缘区域，以提高噪声转移矩阵的准确性。根据这个过程，公式（1）可以重新设计为：

(6)

+(1-) (7)

其中为边缘处为1，其他区域为0的掩码，因此可以对于边缘处的噪声进行损失校正。

**二、推理过程：**

## 在推理过程中，我们仅使用了第二阶段和第三阶段的网络模型。具体来说，首先使用第二阶段进行粗分割，然后通过感兴趣区域裁剪，将感兴趣的区域输入第三阶段的网络进行更精细的分割。同时，我们还采用了后处理操作，包括去除散点以及沿着横截面填充空洞的操作。

## ****参数设置及训练技巧：****

1、数据处理：我们将输入数据压缩到0-1的范围内。为了应对过度拟合问题，我们采用了数据增强技术，例如随机旋转和翻转。

2、在测试阶段，我们采用了测试时数据增强(TTA)来提高预测结果，即对多次翻转后的结果进行单独测试并进行集成。

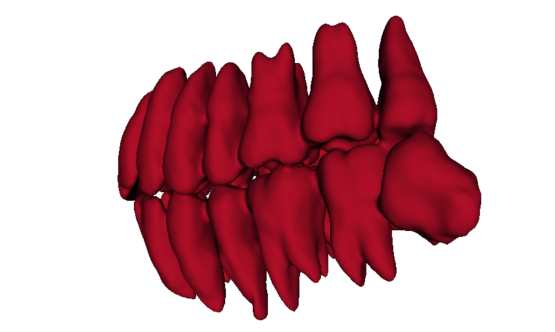
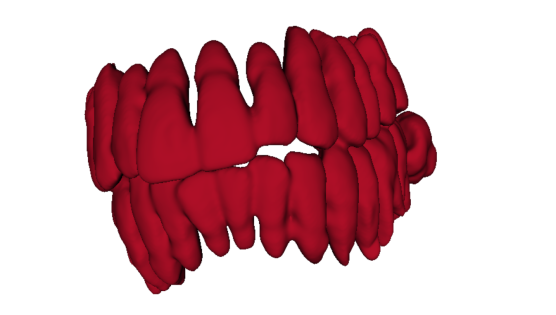
3、我们的优化方法涉及随机梯度下降（SGD），初始学习率为0.1，并采用了用于更新的poly learning rate策略，动量设置为0.9，权重衰减系数设置为0.0001。我们在各个阶段采用的网络结构都为3DUnet[2]。

4、采用了后处理操作，包括去除散点以及沿着横截面填充空洞的操作。

5、我们使用的损失函数主要是传统的交叉熵损失函数和Dice损失函数，具体细节可见公式。

## ****结果展示：****

## ****预测结果预览图****



## ****评价指标结果****

|  |  |  |  |
| --- | --- | --- | --- |
| Score | Dice | Iou | Hausdorff\_distance |
| 0.8237 | 0.8070 | 0.8386 | 0.1689 |

## ****参考文献：****

1. Tarvainen A, Valpola H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results[J]. Advances in neural information processing systems, 2017, 30.
2. Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. Springer International Publishing, 2016: 424-432.
3. Shu J, Zhao Q, Xu Z, et al. Meta transition adaptation for robust deep learning with noisy labels[J]. arXiv preprint arXiv:2006.05697, 2020.
4. Guo X, Yang C, Li B, et al. Metacorrection: Domain-aware meta loss correction for unsupervised domain adaptation in semantic segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 3927-3936.