MICCAI GOALS2022-技术报告

参赛队伍：VisionWise（ocb00999 的队伍）

编程语言：Python，框架：Pytorch

任务 1：青光眼识别

数据集：使用大赛提供的 GOALS-training set 及基于 GOALS-validation set 生成的伪标签（pseudo label）[1]。

网络结构：网络结构使用了两种不同骨架， InceptionV3 [3]和 Resnet-50 [2]。网络结构的实现基于 torchvision 中实现，仅修改最后输出分类数。骨架使用 Imagenet 上预训练的权重进行初始化。训练中 batch_size 为 4

输入数据：网络输入为三通道图片，训练时使用随机上下偏移、随机 Gamma（默认）、随机亮度，Resize 成 512x512 大小输入网络。

输出格式：我们把输出替换为两通道的二分类任务

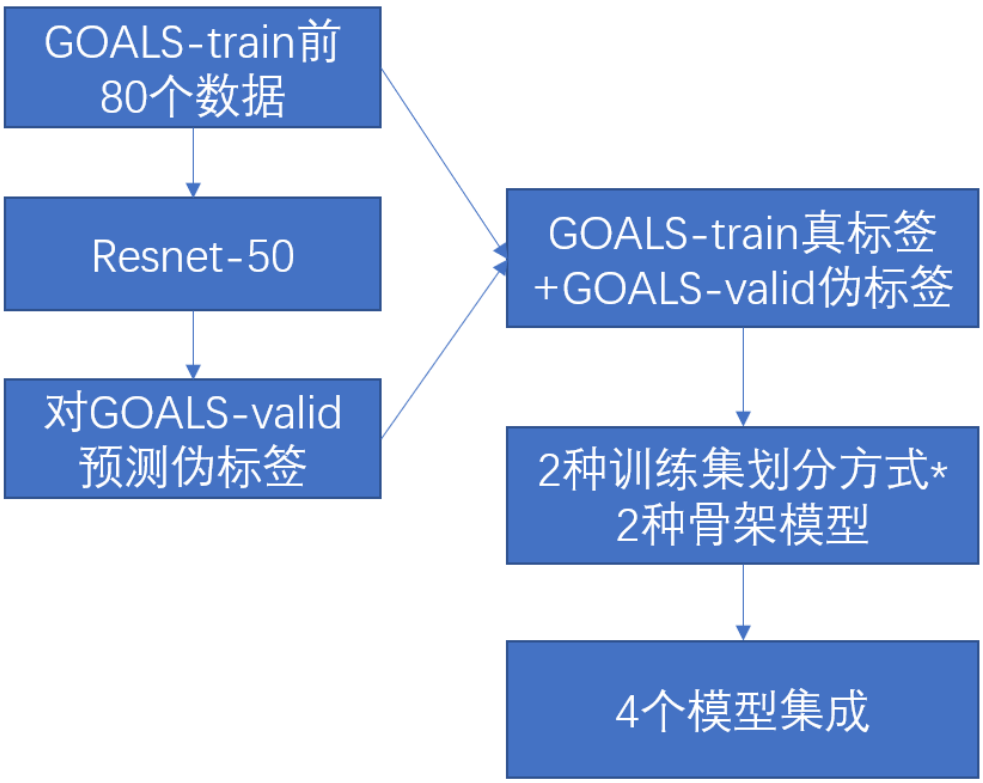
损失函数：CrossEntropyLoss

学习率：采用 Adam 优化器，采用自适应学习率，初始学习率为 0.0003，当连续三个 Epoch training loss 无下降，则把学习率*0.3。训练 100 个 Epoch，选取验

证集上 valid loss 最小的模型。

训练流程

我们把训练集以 8：2 的比例划分成训练集和验证集，我们首先采用 ResNet50 模型，发现已经能在验证集上实现 100%准确率。我们用该模型预测并得到在 GOALS-Validation 上的伪标签。我们重新划分训练集和验证集，第一种划分为使用前 160 个数据为训练集，后 40 个数据为测试集。第二种划分为使用前 40 个数据为验证集，后 160 个数据为测试集。对每一种划分，我们均训练两个模型，分别为 InceptionV3 和 Resnet-50，选取验证集上 valid loss 最小的 checkpoint。最后获得四个模型，对四个模型的预测概率取平均进行集成，得到在 GOALS-validation set 上的最后结果。下图为训练流程示意图：



任务 2：分割

数据集：使用大赛提供的 GOALS-training set 及基于 GOALS-validation set 生成的伪标签（pseudo label）。

网络结构：网络结构为 Encoder-Decoder 的 U-Net 结构，使用了多种不同骨架，包括 InceptionV4 [4] 、ResNest101e [6]、InceptionResnetV2 [4]、Res2Next-50 [5]。网络结构的实现基于 repo: GitHub - qubvel/segmentation_models.pytorch: Segmentation models with pretrained backbones. PyTorch.，仅修改最后输出分类数。骨架使用 Imagenet 上预训练的权重进行初始化。训练中为了减少显存消耗，我们使用了 torch 的混合精度训练，batch_size 为 4-8（根据不同的模型，尽量选取 24G 显存能容纳的最大的 batch size）

输入数据：网络输入为三通道图片，训练时使用随机左右翻转、随机缩放（+/-0.3 范围）、随机对比（albumentation 默认参数）、随机平移（+/-0.1 范围）、随机旋转（+/-20 度）、随机 Gamma（默认）、随机饱和（默认）、随机 Sharpen/Blur/MotionBlur/高斯噪声及 CLAHE 校正，随机裁剪为 768*768 的 Patch 输入，并进行归一化，归一化的参数参照对应使用的骨架。推理时候输入为 1100*800 大小的归一化后的图像。

输出格式：我们发现进行六分类的分割比四分类的分割要效果更好，因此我们把输出替换为 6 通道的六分类分割任务（上部背景、RNFL、GCIPL、GCIPL 与 Choroid 之间的结构、Choroid、下部背景）。

损失函数：Dice Loss + Focal Loss

学习率：采用 Adam 优化器，初始学习率为 0.0003，每 40 个 Epoch 乘以 0.3，训练 170 个 Epoch。

测试增强 TTA：测试时候输入为 1100*800 大小的图像，除了原图，还对其进行左右翻转、Gamma 0.8 处理、Gamma 1.2 处理，最后结果为上述四个输出的概率图的各像素各通道的平均概率值。

模型集成：对经过 TTA 处理后的概率分布图，不同模型间的概率相加后取平均值。

后处理：对经过模型集成或 TTA 后的概率分布图，采用最短路径方法，生成五条分层线。其中第 i 条路径的 cost map 为第 i-1 通道和第 i 通道的相邻行间的概率差

$$cost_i = C_i(x,y+1) - C_{i-1}(x,y)$$

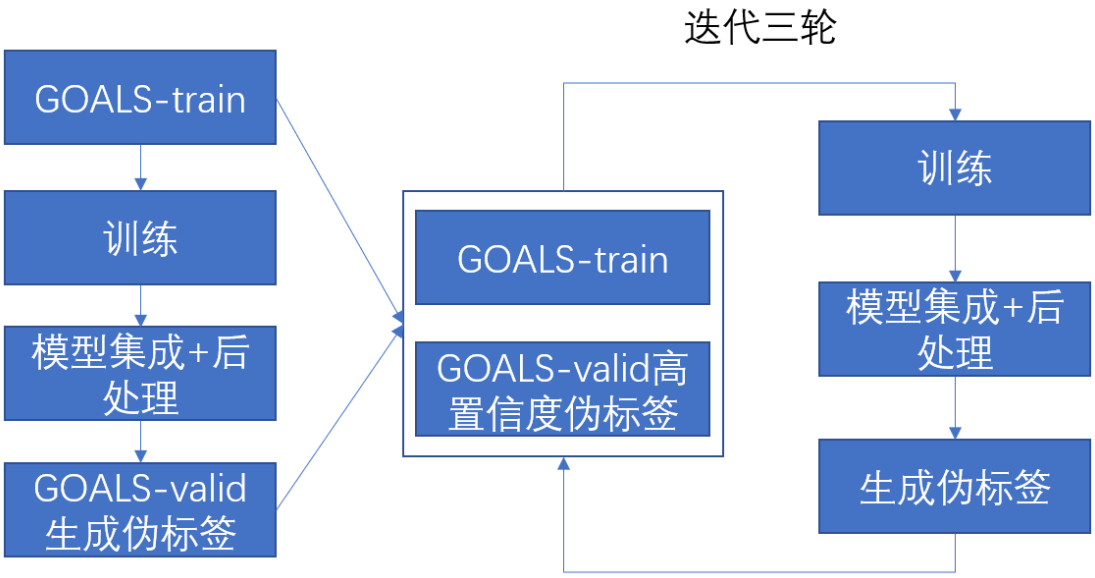
最短路径使用 skimage 中的 route_through_array 实现。各层间路径生成后，重新赋值两条路径间的值为对应的 mask 值（0，80，160，255）。

训练流程

我们把训练集以 9：1 的比例划分成训练集和验证集，在尝试多次交叉验证后，

我们发现通常在 150-170 epoch 之间模型表现较好。我们首先采用 InceptionV4 和 InceptionResNetV2，使用全部 90 张训练集图片训练 170epoch 得到的两个模型，进行上述的模型集成和后处理，预测并得到在 GOALS-Validation 上的伪标签。对于生成的伪标签，我们仅采用高置信度的。我们根据 TTA 和模型集成后得到的概率分布图，对图上的每个位置取各分类（通道）中的最大值为置信度，仅选取图上最低置信度均>0.4 的伪标签。保持验证集不变，我们把选取的伪标签及图加入到原本的训练集中，在训练的每一个 epoch 中，轮流迭代真标签样本和伪标签样本。同样训练 170 个 Epoch，我们采用 InceptionV4 和 InceptionResNetV2 的集成得到一批新的伪标签，同样选取伪标签上置信度高的加入训练集，进行下一轮的真+伪标签混合训练。在第二轮混合训练中，我们采用了 InceptionV4+Res2Next-50 的模型集成，获取新一轮的伪标签。在第三轮混合训练中，我们采用了 InceptionV4 和 ResNest101e 两种骨架。最后，我们选取所有模型中表现较好的模型，包括第一轮混合训练中的 InceptionV4，第二轮混合训练中的 Res2Next-50，第三轮训练中的 ResNest101e，对每个模型进行 TTA，然后对 TTA 后的结果进行模型集成，再进行最短路径的后处理，得到我们最后提交的结果。

训练流程示意图如下：



参考文献：

【1】 Lee D H. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks[C]//Workshop on challenges in representation learning, ICML. 2013, 3(2): 896.

【2】 He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

【3】 Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2818-2826.

【4】 Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]//Thirty-first AAAI conference on artificial intelligence. 2017.

【5】 Gao S H, Cheng M M, Zhao K, et al. Res2net: A new multi-scale backbone architecture[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 43(2): 652-662.

【6】 Zhang H, Wu C, Zhang Z, et al. Resnest: Split-attention networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 2736-2746.