

GOALS 挑战赛技术报告

AUTOMATE 团队

摘要：百度公司在 MICCAI202 举办了 GOALS 挑战赛，该挑战围绕 OCT 图像设计了环扫 OCT 图像的层分割任务（任务一）和青光眼的自动诊断任务（任务二）。AUTOMATE 团队参加了 GOALS 挑战赛并提交了预测结果。在任务一中使用深度学习全监督语义分割技术，为每个非背景类训练相对应的分割模型，最后对每个类别的预测结果进行叠加融合以得到最终的 OCT 层预测掩膜。在任务二中使用深度学习图像分类技术，训练轻量级图像分类网络实现青光眼判别。由于 GOALS 数据集的少样本特点，在以上方法中增加了在线数据增强、交叉验证、测试时增强和后处理等优化策略。AUTOMATE 团队的方法在 GOALS 挑战赛最终分数上排名第 4，其中任务一在初赛和决赛中的排名分别为 4、7，而任务二在初赛和决赛中均保持与若干队伍的并列第 1。

关键词：数据增强，交叉验证，模型融合，测试时增强

1 方法

1.1 OCT 层分割

我们从全监督语义分割角度来处理 OCT 层分割任务。对于每个非背景类别（共 3 类），都训练一个分割模型用于对该类别进行预测，部分模型预测时进行测试时增强，以提高鲁棒性，最后通过后处理步骤以及各类别结果的融合得到最终分割结果。团队最初的方案为 *One VS Rest*，但赛程结束时提交的预测结果定格在方案衔接处，所以本方法中的各个模型实验细节并不统一。

数据。无离线的数据处理步骤，使用完整图像 RGB 三通道格式作为模型输入。在模型训练期间进行在线数据增强，各个模型采用的在线数据增强方式是以下内容的子集：随机尺度、随机旋转、随机水平翻转、随机亮度、随机对比度、随机饱和度、随机滤波。

数据划分示意图如 Figure 1 所示，有两种策略：① 5 折交叉验证法；② 留出法。两种划分方式在训练集和验证集上均保持青光眼类别分布的均衡，留出法中的验证集会经过水平翻转扩充一倍，以提高单份验证集选出模型的鲁棒性。

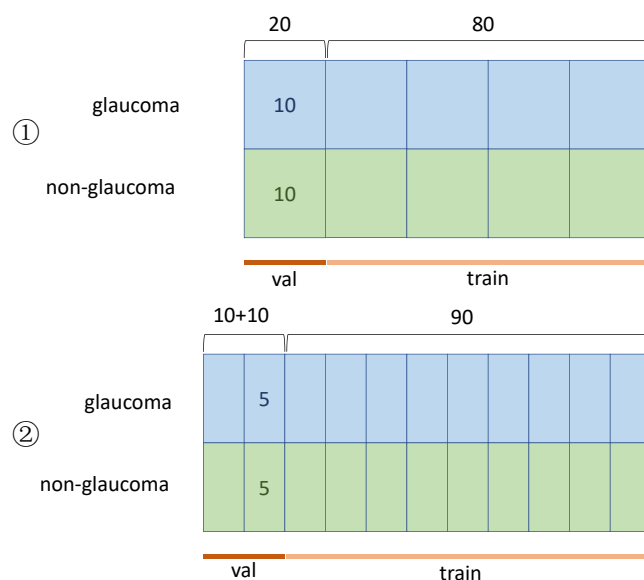


Figure 1 OCT 层分割数据划分

模型。首先定义 OCT 层分割任务中的类别。类别包含：背景、视网膜神经纤维层 RNFL、神经节细胞-内网状层 GC IPL 和脉络膜层 Choroid，并按顺序将以上类别依次命名为类别 0 至类别 3。

FC-HarDNet 70[1]用于分割类别 1，U²-Net[2]用于分割类别 2，以 HRNet-W18[3]为骨干网络的 FCN 用于分割类别 3。其中，U²-Net 的数据划分方式为 5 折交叉验证①，而 HarDNet 和 HRNet 则是留出法②。以上分割模型的网络结构均为标准结构，具体代码实现参考 PaddleSeg[4]。

推理。对于使用留出法的 HarDNet 和 HRNet，在对图像进行预测时，会额外对该图像水平翻转后的图像进行预测，两个预测结果的 Softmax 均值作为此类别的预测结果。对于 U²-Net，5 折交叉验证训练得到的 5 个模型分别对图

像进行预测得到 5 个预测结果，对这些结果进行硬投票得到类别 2 的预测结果。

后处理。在推理阶段得到各个类别的预测掩膜后会进行后处理，以下所列后处理方式是比赛提交方案中曾采用的操作：保留最大连通域、填充空洞、去除小连通域。

完成后处理操作后的各个类别掩膜会进行叠加。由于单类别的预测思路丢失了类别 1 和类别 2 相邻的先验信息，为了防止预测冲突，考虑到 U²-Net 使用交叉验证更具鲁棒性，设定类别 2 的覆盖优先级高于类别 1；在初步合并后，若类别 1 和类别 2 之间的存在背景区域，则将该区域填充为类别 1，得到最终的 OCT 分割预测结果。本方法在模型预测阶段的示意图如 Figure 2 所示。

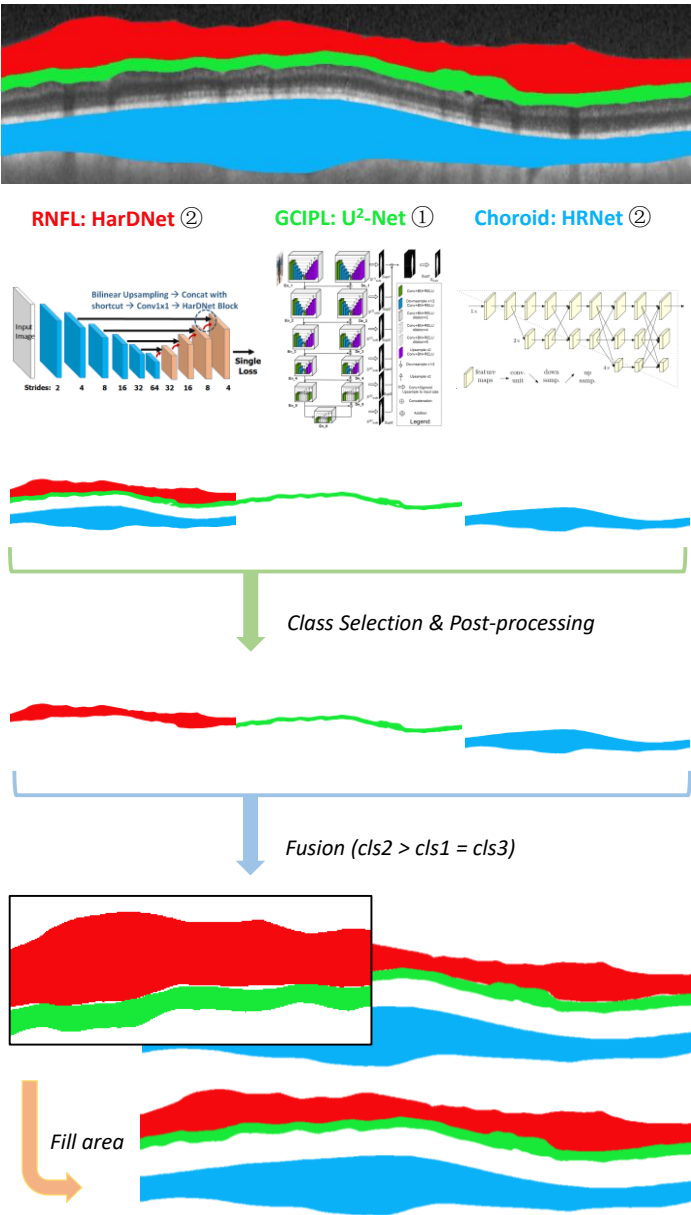


Figure 2 OCT 层分割预测步骤

1.2 青光眼分类

数据。无离线的数据处理步骤，使用完整图像 RGB 三通道格式作为模型输入。在模型训练期间进行在线数据增强，具体包括：随机水平翻转、随机裁剪。

数据划分示意图如 Figure 3 所示。采用 10 折交叉验证，其中训练集、验证集中的类别保持均衡。

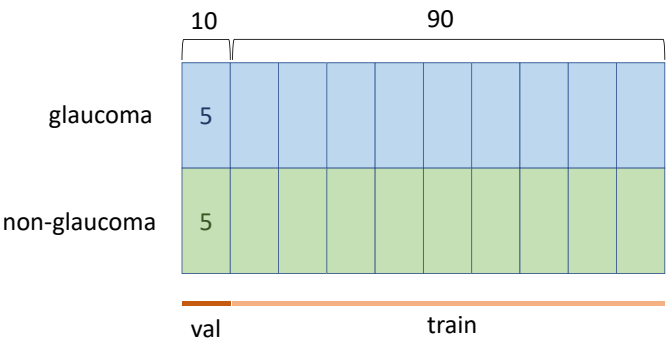


Figure 3 青光眼分类数据划分

模型。轻量级图像分类网络 PP-LCNet [5] 作为青光眼检测模型，该模型的网络结构示意图如 Figure 4 所示。PP-LCNet 是针对 Intel CPU 设备及 MKLDNN 加速库的网络模型，主要由以下优化策略构成：采用更好的激活函数，在合适的位置 {添加通道注意力 SE 模块，采用更大的卷积核}，在全局平均池化层后采用更大的 1×1 卷积层。

Operator	Kernel Size	Stride	Input	Output	SE
Conv2D	3 × 3	2	224 ² × 3	112 ² × 16	-
DepthSepConv	3 × 3	1	112 ² × 16	112 ² × 32	-
DepthSepConv	3 × 3	2	112 ² × 32	56 ² × 64	-
DepthSepConv	3 × 3	1	56 ² × 64	56 ² × 64	-
DepthSepConv	3 × 3	2	56 ² × 64	28 ² × 128	-
DepthSepConv	3 × 3	1	28 ² × 128	28 ² × 128	-
DepthSepConv	3 × 3	2	28 ² × 128	14 ² × 256	-
5 × DepthSepConv	5 × 5	1	14 ² × 256	14 ² × 256	-
DepthSepConv	5 × 5	2	14 ² × 256	7 ² × 512	✓
DepthSepConv	5 × 5	1	7 ² × 512	7 ² × 512	✓
GAP	7 × 7	1	7 ² × 512	1 ² × 512	-
Conv2d, NBN	1 × 1	1	1 ² × 512	1 ² × 1280	-

Figure 4 PP-LCNet 网络结构 [5]

本方法使用标准网络 PPLCNet-1x，具体代码实现参考 PaddleX [6]。需要注意的是，在比赛中提交的模型权重为训练时的最后一次保存权重，而不是交叉验证集上的最优权重。

推理。10 折交叉验证训练得到的 10 个 PP-LCNet 青光眼检测模型。每张待检测图像都会由这 10 个检测模型预测得到 10 个类别预测结果，最后对这 10 个类别预测结果进行硬投票，得到该图像的最终预测类别。

2 实验

- 编程语言为 Python 3.7.4，深度学习框架为 PaddlePaddle 2.2.2。
- 任务一采用的开发套件为 PaddleSeg 2.5.0，任务二采用的开发套件为 PaddleX 2.1.0。

2.1 OCT 层分割

所用数据增强的参数配置如 Table 1 所列。其中，随机模糊 RandomBlur 被定义为从 3 种滤波方式中等概率选择一种对输入图像进行操作；归一化 Normalize 有两种系数，一种是初赛所有训练集图像的均值和标准差（100 张），另一种是本地留出法划分出的训练集中的均值和标准差（90 张）。

Table 1 OCT 层分割-数据增强配置表

Operations	HarDNet (RNFL)		U2-Net (GCIPL)		HRNet (Choroid)	
	comments	prob	comments	prob	comments	prob
ResizeStepScaling	/	/	/	/	uniform(min=0.75, max=1.25)	1.0
RandomRotation	max_range=20	-	max_range=20	-	max_range=30	-
RandomHorizontalFlip	-	0.5	-	0.5	-	0.5
RandomBrightness	range=[0.8, 1.2]	0.5	range=[0.7, 1.3]	0.5	range=[0.8, 1.2]	0.5
RandomContrast	range=[0.6, 1.4]	0.5	range=[0.7, 1.3]	0.5	range=[0.6, 1.4]	0.5
RandomSaturation	range=[0.6, 1.4]	0.5	range=[0.7, 1.3]	0.5	range=[0.6, 1.4]	0.5
RandomBlur (randomly select one)	{gaussian, median, mean} filter	0.1	{gaussian, median, mean} filter	0.1	{gaussian, median, mean} filter	0.1
Normalize	mean=0.22894311 std=0.16314624	-	mean=0.2297999174664977 std=0.16316278003205756	-	mean=0.22894311 std=0.16314624	-

模型训练期间的超参数和损失函数配置情况如 Table 2 所列。

Table 2 OCT 层分割-训练超参数表

Parameters	HarDNet	U2-Net	HRNet
iters	5000	4000	5000
batch size	2		
learning rate	0.0001	0.0005	0.0001
optimizer	AdamW (weight decay=0.01)		
LR sche.	cosine (min=0)	warmup (step=200, start=0) & cosine (min=0)	warmup (step=250, start=0) & cosine (min=0)
pretrained	CITYSCAPES	-	CITYSCAPES
loss	Weighted Dice Loss (weight=[1, 1, 2, 1])	Dice Loss	Weighted Dice Loss (weight=[0.95, 1.05])

模型预测时设置及后处理的方法如 Table 3 所列。其中，后处理方法在初赛和决赛中有所不同：初赛中对各类别掩膜保留最大连通域，而决赛中则是去除类别掩膜中的小区域和孔洞。

Table 3 OCT 层分割-模型预测与后处理的配置表

Schedule	HardNet	U2-Net	HRNet	Post-processing
Preliminary	TTA	5 Fold Cross Validation	TTA	Keep the largest connected domain.
Finals	(Flip Horizontal)	(Hard-Voting)	(Flip Horizontal)	Remove small objects (threshold=128) & Remove small holes (threshold=860)

2.2 青光眼分类

数据读取与在线增强操作的配置情况如 Table 4 所列。在训练阶段，输入图像会经过随机水平翻转和随机裁剪。其中，随机裁剪的计算逻辑是：随机在原图像 S 上裁剪出一块区域 s（s 的面积除以 S 的面积 $\in [0.3, 1.0]$ 、s 的宽高比 $\in [0.5, 2.0]$ ），并将区域 s 缩放到固定尺度 224 \times 224。在推理阶段，输入图像会被缩放至 352 \times 256，然后在已缩放图像的中央裁剪出固定尺寸 224 \times 224 作为输入。图像归一化系数采用 IMAGENET 数据集系数。

Table 4 青光眼分类-数据处理配置表

Mode	Operations	Comments
Train	RandomHorizontalFlip	prob=0.5
	RandomCrop	crop_size=224 aspect_ratio=[0.5, 2.0] scaling=[0.3, 1.0]
	Normalize	mean=[0.485, 0.456, 0.406] std=[0.229, 0.224, 0.225]
Infer	ResizeByShort	short_size=256
	CenterCrop	crop_size=224
	Normalize	mean=[0.485, 0.456, 0.406] std=[0.229, 0.224, 0.225]

模型训练相关的超参数及配置情况如 Table 5 所列。

Table 5 青光眼分类-模型训练相关参数配置表

Parameters	Configs
epoch	50
batch size	10
learning rate	0.0005
optimizer	SGD (momentum=0.9, weight decay=3e-5)
LR sche.	warm-up (epoch=5, start=0) & cosine (min=0)
pretrained	IMAGENET
loss	Cross Entropy Loss

3 评估

AUTOMATE 团队方法在 GOALS 挑战赛中各项评估指标的表现如 Table 6 所列。

Table 6 GOALS 挑战赛评估结果表

Tasks	Index [7]	Preliminary		Finals		GOALS
		Details	Score (Rank)	Details	Score (Rank)	Score (Rank)
OCT Layer Segmentaion	RNFL_DICE	0.9481	6.83262 (4th)	0.9561	6.86407 (7th)	8.85463 (4th)
	RNFL_ED	1.0922		1.1014		
	GCIPL_DICE	0.8909		0.8966		
	GCIPL_ED	1.2546		1.2848		
	Choroid_DICE	0.9454		0.9569		
	Choroid_ED	1.7114		1.6767		
Glaucoma Classification	AUC	1	2 (1st)	1	2 (1st)	
	F1	1		1		
	ACC	1		1		
	SEN	1		1		
	SPE	1		1		

4 参考文献

[1] Chao, Ping, et al. "Hardnet: A low memory traffic network." Proceedings of the IEEE/CVF international conference on computer vision. 2019.

[2] Qin, Xuebin, et al. "U2-Net: Going deeper with nested U-structure for salient object detection." Pattern recognition 106 (2020): 107404.

[3] Wang, Jingdong, et al. "Deep high-resolution representation learning for visual recognition." IEEE transactions on pattern analysis and machine intelligence 43.10 (2020): 3349-3364.

[4] "GitHub - PaddlePaddle/PaddleSeg at release/2.5." <https://github.com/PaddlePaddle/PaddleSeg/tree/release/2.5>.

[5] Cui, Cheng, et al. "PP-LCNet: A Lightweight CPU Convolutional Neural Network." arXiv preprint arXiv:2109.15099 (2021).

[6] "GitHub - PaddlePaddle/PaddleX at release/2.1." <https://github.com/PaddlePaddle/PaddleX/tree/release/2.1>.

[7] Fang, Huihui, et al. "Dataset and Evaluation algorithm design for GOALS Challenge." International Workshop on Ophthalmic Medical Image Analysis. Springer, Cham, 2022.