

# 基于深度学习的眼底动静脉交叉压迫征分级

## 1. 摘要

视网膜动静脉交叉压迫征是一种显著而突出的微血管异常，其特点是在动静脉交叉点处的两侧静脉口径的减小。根据最新的研究，视网膜动静脉交叉压迫征对于部分眼科疾病，诸如视网膜分支静脉闭塞等，与部分心血管疾病，诸如中风等，有着强有力的预测功能。

我们提出的基于迁移学习思想的双流多模态卷积神经网络算法能够很好地关注到动静脉交叉点处静脉变窄区域。实验证明，我们提出的模型的分级结果可解释性较好，对动静脉交叉压迫征的严重程度判断具有一定的依据。

## 2. 引言

动静脉交叉压迫征是一种在动脉与静脉交叉处静脉受到硬化的动脉的挤压引发的病理性变化。在正常的情况下，静脉管径不会在动静脉交叉点处有所变化。动静脉交叉压迫征按照严重程度可分为这样四个级别，一级为轻度AV Nicking（静脉一侧受挤压），二级为中度AV Nicking（仍然为静脉一侧受挤压），三级为重度AV Nicking（静脉两侧均受到挤压，因此被视为受挤压程度更加严重）。根据心血管健康相关研究机构统计的数据，世界上动静脉交叉压迫征的患病率为7.7%；根据ARIC动静脉粥样硬化风险研究机构统计的数据，患有动静脉交叉压迫征疾病的群体发生卒中的可能性几率是未患动静脉交叉压迫征疾病的群体的两倍。因此，动静脉交叉压迫征严重程度的正确分级，对于发现心血管疾病高危人群，并对其进行早期有效的诊断具有十分重要的作用。

## 3. 相关工作

Nguyen提出的AV Nicking检测方法主要步骤有两步：

- (1) 血管分割(Vessel Segmentation)，使用多尺度直线检测器与阈值分割器相结合的方法，从图像背景中提取血管用于后续分析。
- (2) 执行动静脉交叉点检测方法(Artery-Vein Crossover Point Detection)检测视网膜图像内的动静脉交叉点位置。

在第一步血管分割任务中，高分辨率视网膜彩色图像的血管分割，存在视网膜血管中央光反射、脉络膜血管化、背景均匀化和其他脉冲噪声等图像伪影等问题。尽管有许多方法曾被提出来用于提取视网膜血管，但它们对类似于DRIVE和IOSTAR这些数据集，即存在血管中心光反射（沿着血管中间的亮条）现象的高分辨率图像，血管检测并不有效。基于这个问题，作者使用了一种多尺度直线检测的技术来高效地提取视网膜血管网络，这种技术将各种尺度的直线检测器结合起来从而产生一种增强后的视网膜图像，再通过设定阈值算子来获得最终的血管分割图，如下图所示。这种方法被证明能够有效地解决血管中心光反射问题从而提供较为精确的血管边界检测。

Roy提出的方法分为两个主要模块：图像处理模块和机器学习模块。在图像处理模块中，完成的内容包括视网膜血管分割(Vessel Segmentation)、AV交叉点检测(AV Crossover Point Detection)、ROI选择(ROI Selection)、动脉静脉分类(Artery Vein Classification)、静脉段分离(Vein Segmentation Separation)、静脉边缘和中心线细化(Vein Edge and Centerline Refinement)和宽度计算(Width Computation)等任务。另一方面，机器学习模块包括利用计算宽度进行的特征提取(Feature extraction from computed widths)、测试和训练数据的创建和分类(test and training data creation and classification)等任务。在此前的方法中，通过产生连续值(continuous value)来代表AV Nicking的严重程度。但是作者难以对当时使用的连续值的实际含义进行合理的解释。此外，宽度在当时的工作中没有得到标准化，因此，衡量AV Nicking现象严重性的得分可能取决于不同病人血管宽度的可变性。在该种方法中，静脉宽度直接从分割图像中计算，可能受到二值化阶段使用的阈值的影响，更精确的不受阈值影响的宽度计算可以产生更精

确的AV Nicking严重程度的测量。故在该文中，作者针对上一种方法的不足与缺陷，提出了一种新的完全自动化的方法，可以直接从彩色视网膜图像中区分AV Nicking严重程度的不同。

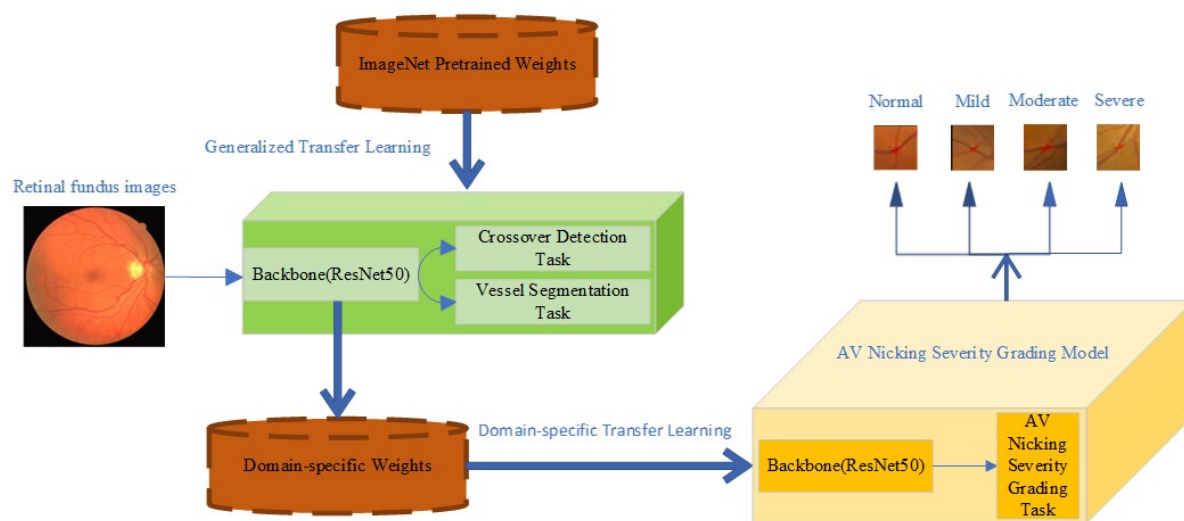
在MDTNET模型中，首先交叉点通过结合血管分割图与血管动静脉分类图的信息来获取，然后通过筛选出合法的交叉点，并将合法的交叉点patch图像输入MDTNET网络中进行分级任务的训练。然而，使用这样一种多模型的方式来进行训练需要去训练每个模型，这是需要巨大的计算资源并且是十分耗时的。

## 4. 方法

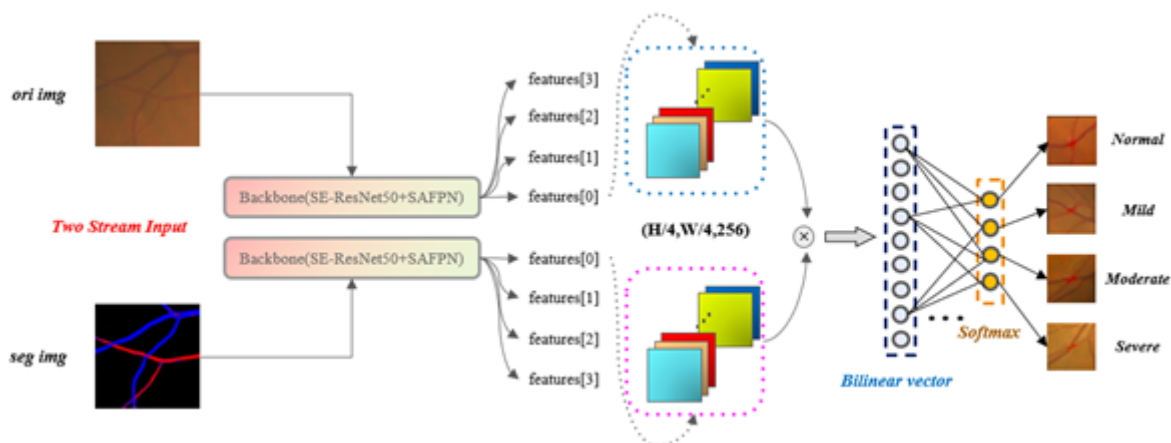
本次研究的视网膜眼底动静脉交叉压迫征现象检测与严重程度分级问题为医疗影像领域内的少样本细粒度图像病变分级的问题，即其难度主要体现在两个方面：患病样本数据稀缺、生理病变区域与正常状态下的同一区域差异不明显。

针对第一个难点，我们采用迁移学习作为核心思想。近些年来不同的深度学习模型被应用在广泛的医疗影像处理实际问题当中，如肺、脑这样的器官病变检测问题，同时也包括各种眼科相关疾病预测分级问题，例如糖尿病视网膜病变、继发黄斑水肿等。与此同时，解决上述问题的模型大多数都使用了迁移学习的手段。在医学影像处理领域中，使用较小的数据集进行深度神经网络的训练时往往会暴露出很多局限性，如难以收敛等。而已经在大型数据集（如ImageNet, CoCo等大型自然图像数据集）上训练过的预训练模型的先验知识会通过迁移学习这一过程迁移到目标模型之中。然而，在所有这些应用中，预训练模型都是基于如ImageNet和CoCo这样的公众数据集进行预训练的，但是没有在特定问题域的数据集中进行预训练。因此，我们采用了一种新的迁移学习策略，即在视网膜眼底彩照这一特定问题域中，选择合适的上游任务进行模型的预训练操作，从而使得得到的预训练模型能够很好地提取视网膜眼底彩照动静脉交叉点的相关特征，而非仅仅能够提取自然图像的一般特征。在获得这样的预训练模型之后，我们便可以将其学习到的先验知识迁移到后面的动静脉交叉压迫征分级模型中，从而使后面的模型能够更好地完成下游的分级任务。

特定域的迁移学习思路图如下图所示：



针对第二个难点，我们在上游的动静脉交叉点目标检测及附近血管实例分割算法中改进了FPN的网络结构，设计了SAFPN模块；同时，在进行下游任务时，在backbone中加入了SENet模块以学习通道注意力，并综合了AV Nicking数据集原图信息与其对应的动静脉多分类信息，设计了一种双流多模态卷积神经网络架构，两支CNN输入流蕴含的特征信息相互补充，从而更好地提取动静脉交叉压迫征病变区域静脉段变窄的细粒度特征。



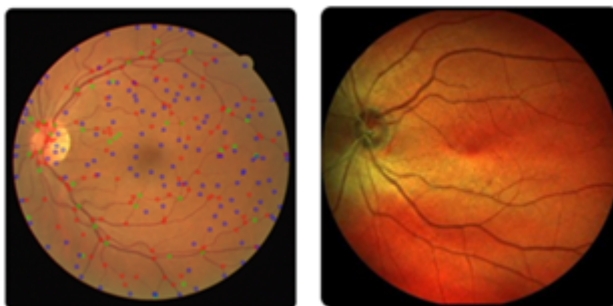
## 5. 数据集介绍与处理

### 5.1 数据集介绍

在进行动静脉交叉点目标检测及附近血管实例分割算法的实现与部署时，我们需要准备合适的数据集，从而满足对应的网络架构的输入数据需求。

DRIVE数字视网膜图像数据集是在2003年发布的，如下图所示。其由40张照片组成，其中7张显示出轻度早期糖尿病视网膜病变迹象。这些图像来自于荷兰的一个糖尿病视网膜病变筛查项目，该项目涵盖了400名年龄在25岁至90岁的糖尿病患者。在这些患者中，随机抽取了40张照片，其中33张没有显示任何糖尿病视网膜病变的迹象，7张显示有轻度早期糖尿病视网膜病变迹象。这些图像是由佳能CR5非滴胶3CCD相机拍摄的，具有45度的视野。每张图像大小为768乘以584像素，每色平面8比特拍摄。每张图像的FOV是圆形的，直径约为540像素。该数据库已经对图像进行了裁剪，以便围绕FOV进行训练和测试。对于每张图像，都提供了一张遮罩图像，以划定FOV。该组40张图像被分为训练集和测试集，每个集都有20张图像。

对于训练图像，可以使用单次手动分割对脉管系统进行训练。对于测试用例，可以使用两种手动分段中的一种作为黄金标准，另一种用于比较计算机生成的分割与独立的人类观察者的分割。此外，每个视网膜图像都有一个掩膜图像，以指示感兴趣的区域。所有手动分割脉管系统的人类观察者都由经验丰富的眼科医生指导和培训。

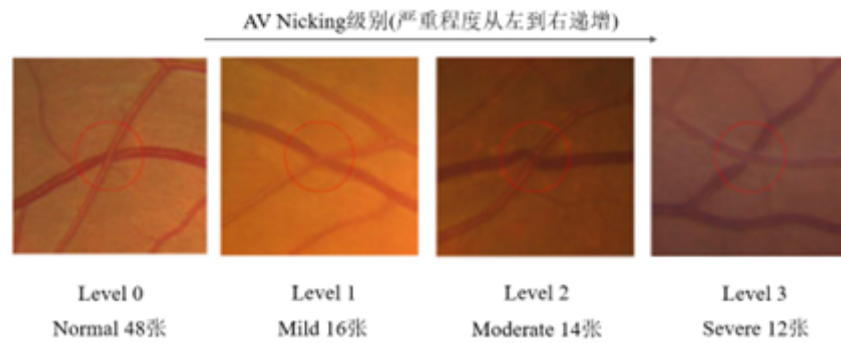


IOSTAR数据集类似于DRIVE数据集，包含原图（彩色图）和分割图和关键点坐标，可以作为点检测（分支点/交叉点/末梢点）或者分割的经典数据集，如上图所示。私有IOSTAR数据库包括24个使用绿色和红外激光的扫描激光检眼镜(SLO)图像。这些图像是用分辨率为1024×1024(~ 14 μm/px)的EasyScan相机拍摄的，并且45°FOV。三个不同的专家对这两个数据集的分叉和交叉进行了注释和校正。

在DRIVE中平均每个图像有100个分叉和30个交叉点，在IOSTAR数据集中有55个分叉和23个交叉点。

在进行动静脉交叉压迫征严重程度分级算法的实现与部署时，我们需要准备的数据集是AV Nicking这份ROI图像的patch数据集，从而满足对应的网络架构的输入数据需求。

AV Nicking数据集共有90张，均为ROI区域的patch图像。AV Nicking数据集来自于两项基于人群的研究the Blue-Mountain-Eye-Study(BMES)和the Singapore-Malay-Eye-Study(SiMES)中收集的47张眼底图像。在这90张patch图像中，动静脉交叉压迫征的严重程度共分为四个级别：正常、轻微、中等、严重，这四种级别的图像数量分别为48张、16张、14张、12张，如下图所示：



## 5.2 规范格式数据集标签制作

由于我们第一步任务，需要检测圆形眼底彩照中的动静脉血管交叉点并对其附近的动静脉血管进行实例分割，而DRIVE和IOSTAR数据集中，只有圆形眼底彩照原图、动静脉血管交叉点横纵坐标信息、圆形眼底彩照对应的黑白二值化分割图等信息，即并没有现成的动静脉交叉位置矩形标注框或动静脉交叉点附近血管分割掩膜信息等，因此我们需要自己根据动静脉血管交叉点的横纵坐标，手动制作目标检测与实例分割过程所需的规范格式的数据集标签。

这里对于目标检测任务，我们制作出包含了动静脉交叉点区域矩形框坐标信息的coco格式的数据集标签，具体操作步骤如下：

- (1) 定义一个categories数组，并在categories数组中加入新类别，即动静脉交叉点，名为“CrossPos”，编号id从1开始（因为在目标检测任务中编号0一般被认为是背景类）；
- (2) 定义一个images数组，并读取圆形眼底彩照数据集的路径，对圆形眼底彩照数据集路径下的每一张图片求出长宽等信息，加入到img\_context字典中，并将加入完后的该图像对应的img\_context信息插入到images数组之中；
- (3) 在标注文件路径中根据与图像文件路径相同的后缀名来读取步骤2中选择的图像对应的标注文件，进而从对应的标注文件中读取该张圆形眼底彩照上所有的动静脉交叉点标注坐标信息；
- (4) 定义一个annotations数组，在遍历步骤3中对应的圆形眼底彩照上的所有动静脉交叉点横纵坐标标注信息时，制作一个bbox\_dict字典，对于每一个动静脉交叉点，以该点为中心，作出一个边长为20的正方形，即boundingbox，为该boundingbox设置一个独立唯一的编号，记录它对应的原图的编号以及该交叉点对应的类别信息（因为这里不考虑血管分支点与血管分岔点，所以所有交叉点的类别信息只有一种，即CrossPos），计算boundingbox的面积，并将所有信息，包括正方形标注框的四个顶点的坐标、正方形标注框的面积、正方形标注框的编号、正方形标注框对应的原图编号、正方形标注框对应的类别信息等，统一加入到刚刚制作的bbox\_dict字典中，再向annotations数组中加入制作好的bbox\_dict字典，然后在继续遍历其它动静脉交叉点的标注信息时循环进行此过程；
- (5) 遍历完一张圆形眼底彩照上的所有交叉点之后，继续遍历相同路径下剩余的圆形眼底彩照，并重复执行步骤1、步骤2、步骤3、步骤4，直到该路径下的所有圆形眼底彩照全部被遍历完毕为止。

对DRIVE和IOSTAR两份圆形眼底彩照数据集均执行上述的算法流程，从而制作出DRIVE和IOSTAR两份圆形眼底彩照数据集对应的coco格式的标注数据txt文件。

而在实例分割的实验部分中，与目标检测任务不同的是，只有动静脉血管交叉点横纵坐标信息是不够的，还需要动静脉血管交叉点附近动静脉血管段的掩膜信息Groundtruth，而这部分信息是DRIVE和IOSTAR数据集所不具有的，这意味着我们需要设法去根据DRIVE和IOSTAR数据集中已有的圆形眼底彩照原图、原图对应的黑白二值化分割图、圆形眼底彩照中动静脉交叉点的横纵坐标信息等，制作出动静脉血管交叉点附近动静脉血管段的掩膜分割图像，我们采用的具体制作步骤如下：

- (1) 设置计数器（初始值为1），遍历圆形眼底彩照对应的黑白二值化分割图路径，记为binary\_mask\_dir，每遍历一张binary\_mask\_dir路径下的黑白二值化分割图时，在标注信息的路径（记为annotation\_dir）下找到当前黑白二值化分割图对应的动静脉血管交叉点标注文件；



(2) 遍历步骤1中圆形眼底彩照（或黑白二值化分割图）对应的动静脉血管交叉点标注文件，每遍历到一个动静脉血管交叉点时，循环遍历以该动静脉血管交叉点为中心的正方形（边长设置为20）内的所有像素，如果遍历到的像素为白色像素，则将该像素的数值赋值为与当前计数器相同的数字，遍历完当前正方形之后，计数器自动加一；

(3) 重复步骤2中的操作，直到将当前的黑白二值化分割图中的所有动静脉血管交叉点遍历完毕为止，此时将计数器清零，并继续遍历黑白二值化分割图路径(binary\_mask\_dir)下的其它黑白二值化分割图像；

(4) 重复上述的步骤1、步骤2、步骤3，直到所有黑白二值化分割图像均被处理完毕为止。

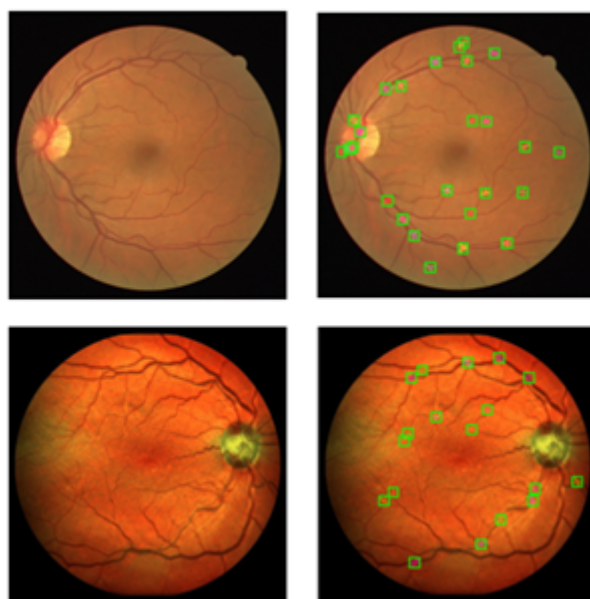
## 6. 实验结果及分析

### 6.1 动静脉交叉点目标检测及其附近血管实例分割算法实验结果及分析

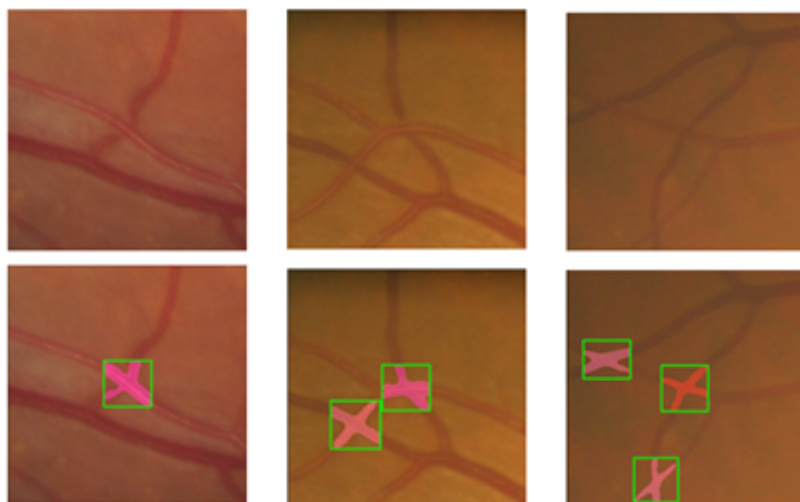
首先我们将展示动静脉交叉点目标检测及其附近血管实例分割算法的实验结果并进行分析。

我们先在圆形眼底彩照上训练了50个epochs，然后在切取的patch图像数据集上微调了10个epochs，第一步训练得到的模型在DRIVE和IOSTAR圆形眼底彩照两份数据集上检测分割的效果如下图所示。

由下图可知，在圆形眼底彩照数据集DRIVE与IOSTAR的图像中，我们第一步训练得到的模型能够较好地检测到动静脉交叉点的位置，并作出以动静脉交叉点为中心的绿色的boundingbox；同时，我们的模型还能够较好地分割出动静脉交叉点中心附近的动静脉血管掩膜，并且绘制不同的动静脉血管实例分割掩膜时所使用的颜色不同，从而达到能够区分不同交叉点及其附近动静脉血管的效果。



第二步训练（微调）得到的模型在AV Nicking的patch数据集上检测的效果如下图所示：



由上图可知，在AV Nicking数据集的patch图像中，我们第二步训练得到的模型能够较好地检测到动静脉交叉点的位置，并作出以动静脉交叉点为中心的绿色的boundingbox；同时，我们的模型还能够较好地分割出动静脉交叉点中心附近的动静脉血管掩膜，并且绘制不同的动静脉血管实例分割掩膜时所使用的颜色不同，从而达到能够区分不同交叉点及其附近动静脉血管的效果。

训练过程中的各指标变化情况如下图所示，其中包括使用调度器不断调整的学习率lr、总loss、分类loss、boundingbox回归loss、分割掩膜loss等指标，各项loss在总体上呈现逐渐稳步下降的态势。

```
Epoch: [9] [ 20/181] eta: 0:04:49 lr: 0.000005 loss: 0.7254 (0.7682) loss_classifier: 0.145
2 (0.1514) loss_box_reg: 0.1787 (0.1859) loss_mask: 0.3423 (0.3447) loss_objectness: 0.0464 (
0.0584) loss_rpn_box_reg: 0.0144 (0.0279) time: 1.7947 data: 0.0119 max mem: 10457
Epoch: [9] [ 30/181] eta: 0:04:31 lr: 0.000005 loss: 0.7544 (0.7700) loss_classifier: 0.146
4 (0.1549) loss_box_reg: 0.1962 (0.1935) loss_mask: 0.3430 (0.3440) loss_objectness: 0.0464 (
0.0545) loss_rpn_box_reg: 0.0140 (0.0232) time: 1.8100 data: 0.0131 max mem: 10457
Epoch: [9] [ 40/181] eta: 0:04:13 lr: 0.000005 loss: 0.7282 (0.7554) loss_classifier: 0.155
4 (0.1515) loss_box_reg: 0.1962 (0.1923) loss_mask: 0.3318 (0.3383) loss_objectness: 0.0385 (
0.0509) loss_rpn_box_reg: 0.0124 (0.0224) time: 1.7984 data: 0.0130 max mem: 10457
Epoch: [9] [ 50/181] eta: 0:03:56 lr: 0.000005 loss: 0.7282 (0.7631) loss_classifier: 0.146
5 (0.1543) loss_box_reg: 0.1929 (0.1958) loss_mask: 0.3254 (0.3416) loss_objectness: 0.0380 (
0.0503) loss_rpn_box_reg: 0.0118 (0.0210) time: 1.8039 data: 0.0130 max mem: 10457
Epoch: [9] [ 60/181] eta: 0:03:39 lr: 0.000005 loss: 0.7378 (0.7580) loss_classifier: 0.145
2 (0.1517) loss_box_reg: 0.1929 (0.1947) loss_mask: 0.3458 (0.3431) loss_objectness: 0.0376 (
0.0490) loss_rpn_box_reg: 0.0082 (0.0195) time: 1.8444 data: 0.0127 max mem: 11097
Epoch: [9] [ 70/181] eta: 0:03:21 lr: 0.000005 loss: 0.7640 (0.7613) loss_classifier: 0.147
0 (0.1524) loss_box_reg: 0.2066 (0.1971) loss_mask: 0.3431 (0.3434) loss_objectness: 0.0439 (
0.0488) loss_rpn_box_reg: 0.0113 (0.0196) time: 1.8565 data: 0.0114 max mem: 11097
Epoch: [9] [ 80/181] eta: 0:03:03 lr: 0.000005 loss: 0.7671 (0.7606) loss_classifier: 0.156
2 (0.1520) loss_box_reg: 0.1933 (0.1951) loss_mask: 0.3383 (0.3423) loss_objectness: 0.0471 (
0.0501) loss_rpn_box_reg: 0.0179 (0.0211) time: 1.8164 data: 0.0122 max mem: 11097
Epoch: [9] [ 90/181] eta: 0:02:44 lr: 0.000005 loss: 0.7555 (0.7613) loss_classifier: 0.154
7 (0.1519) loss_box_reg: 0.1893 (0.1951) loss_mask: 0.3375 (0.3434) loss_objectness: 0.0475 (
0.0500) loss_rpn_box_reg: 0.0143 (0.0209) time: 1.7969 data: 0.0129 max mem: 11097
Epoch: [9] [100/181] eta: 0:02:26 lr: 0.000005 loss: 0.7725 (0.7614) loss_classifier: 0.155
9 (0.1532) loss_box_reg: 0.2116 (0.1969) loss_mask: 0.3333 (0.3415) loss_objectness: 0.0468 (
0.0494) loss_rpn_box_reg: 0.0143 (0.0204) time: 1.7924 data: 0.0122 max mem: 11097
Epoch: [9] [110/181] eta: 0:02:08 lr: 0.000005 loss: 0.7263 (0.7580) loss_classifier: 0.147
4 (0.1516) loss_box_reg: 0.1998 (0.1952) loss_mask: 0.3167 (0.3425) loss_objectness: 0.0398 (
0.0487) loss_rpn_box_reg: 0.0138 (0.0199) time: 1.8084 data: 0.0115 max mem: 11097
Epoch: [9] [120/181] eta: 0:01:50 lr: 0.000005 loss: 0.7093 (0.7594) loss_classifier: 0.133
0 (0.1519) loss_box_reg: 0.1869 (0.1960) loss_mask: 0.3300 (0.3440) loss_objectness: 0.0390 (
0.0481) loss_rpn_box_reg: 0.0120 (0.0194) time: 1.8222 data: 0.0121 max mem: 11097
Epoch: [9] [130/181] eta: 0:01:32 lr: 0.000005 loss: 0.7482 (0.7582) loss_classifier: 0.134
6 (0.1513) loss_box_reg: 0.1980 (0.1957) loss_mask: 0.3470 (0.3452) loss_objectness: 0.0363 (
0.0472) loss_rpn_box_reg: 0.0094 (0.0187) time: 1.8062 data: 0.0125 max mem: 11097
Epoch: [9] [140/181] eta: 0:01:14 lr: 0.000005 loss: 0.7709 (0.7606) loss_classifier: 0.152
```

我们先在圆形眼底彩照上训练了50个epochs，然后在切取的patch图像数据集上微调了10个epochs，两步训练均完毕后我们在验证集上进行评估，分别统计这两步训练得到的模型在验证集上的各项指标，如下图所示：

IoU metric: bbox		IoU metric: bbox
Average Precision (AP) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.272		Average Precision (AP) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.497
Average Precision (AP) @ IoU=0.50   area= all   maxDets=100 ] = 0.505		Average Precision (AP) @ IoU=0.50   area= all   maxDets=100 ] = 0.686
Average Precision (AP) @ IoU=0.75   area= all   maxDets=100 ] = 0.250		Average Precision (AP) @ IoU=0.75   area= all   maxDets=100 ] = 0.499
Average Precision (AP) @ IoU=0.50:0.95   area= small   maxDets=100 ] = 0.272		Average Precision (AP) @ IoU=0.50:0.95   area= small   maxDets=100 ] = 0.497
Average Precision (AP) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000		Average Precision (AP) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000
Average Precision (AP) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000		Average Precision (AP) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000
Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 1 ] = 0.169		Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 1 ] = 0.550
Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 10 ] = 0.395		Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 10 ] = 0.589
Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.450		Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.612
Average Recall (AR) @ IoU=0.50:0.95   area= small   maxDets=100 ] = -1.000		Average Recall (AR) @ IoU=0.50:0.95   area= small   maxDets=100 ] = -1.000
Average Recall (AR) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000		Average Recall (AR) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000
Average Recall (AR) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000		Average Recall (AR) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000
IoU metric: segm		IoU metric: segm
Average Precision (AP) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.106		Average Precision (AP) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.147
Average Precision (AP) @ IoU=0.50   area= all   maxDets=100 ] = 0.347		Average Precision (AP) @ IoU=0.50   area= all   maxDets=100 ] = 0.473
Average Precision (AP) @ IoU=0.75   area= all   maxDets=100 ] = 0.027		Average Precision (AP) @ IoU=0.75   area= all   maxDets=100 ] = 0.022
Average Precision (AP) @ IoU=0.50:0.95   area= small   maxDets=100 ] = 0.106		Average Precision (AP) @ IoU=0.50:0.95   area= small   maxDets=100 ] = 0.147
Average Precision (AP) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000		Average Precision (AP) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000
Average Precision (AP) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000		Average Precision (AP) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000
Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 1 ] = 0.080		Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 1 ] = 0.202
Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 10 ] = 0.179		Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets= 10 ] = 0.220
Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.211		Average Recall (AR) @ IoU=0.50:0.95   area= all   maxDets=100 ] = 0.241
Average Recall (AR) @ IoU=0.50:0.95   area= small   maxDets=100 ] = 0.211		Average Recall (AR) @ IoU=0.50:0.95   area= small   maxDets=100 ] = 0.241
Average Recall (AR) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000		Average Recall (AR) @ IoU=0.50:0.95   area= medium   maxDets=100 ] = -1.000
Average Recall (AR) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000		Average Recall (AR) @ IoU=0.50:0.95   area= large   maxDets=100 ] = -1.000

上图中，small、medium、large各表示小、中、大三种不同尺度物体的AP值，其中第一步训练的目标检测总AP值为0.505，实例分割总AP值为0.347；第二步训练（微调）的目标检测总AP值为0.686，实例分割总AP值为0.473。

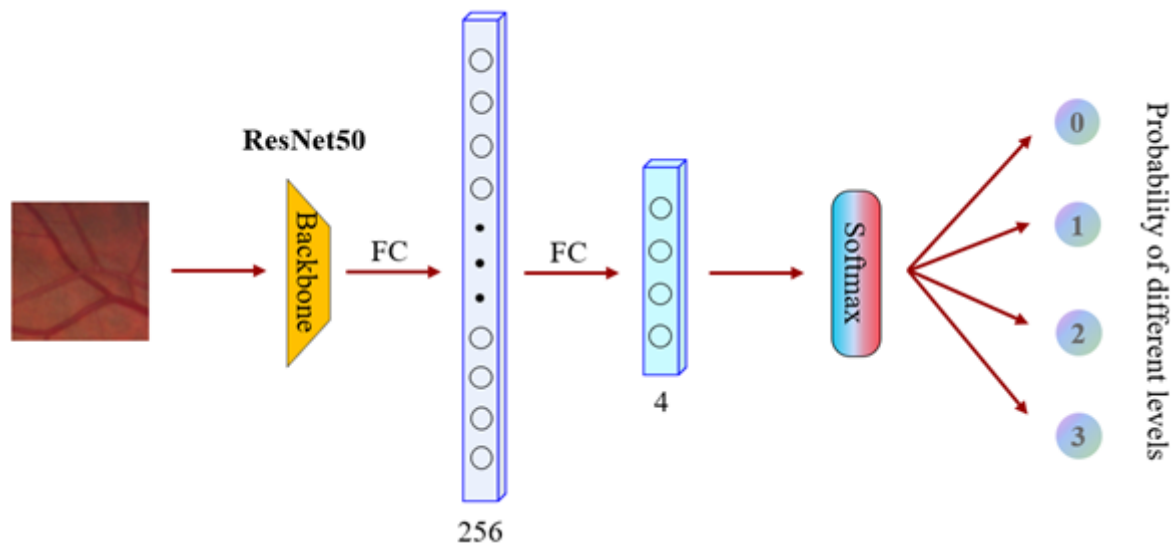
## 6.2 动静脉交叉压迫征现象检测与严重程度分级算法实验结果及分析

接下来我们将展示动静脉交叉压迫征现象检测与严重程度分级算法的实验结果并进行分析。

由于在之前使用深度学习解决本课题的研究工作相对较少，而这些使用深度学习解决方案使用的数据集均为相关医疗机构内部研发的保密数据集，不对外公开，我们无法获取，故我们使用的是目前仅有90张图像的对外公开的AV Nicking数据集。为了验证我们设计的模型及解决方案是可行的，我们采用了将本人设计的模型与baseline实验模型进行比较的思路，通过计算模型训练与评估时的一系列指标，作出相关图像并绘制相关图表，相互比较从而达到实验结果分析的目的。

### 6.2.1 baseline实验

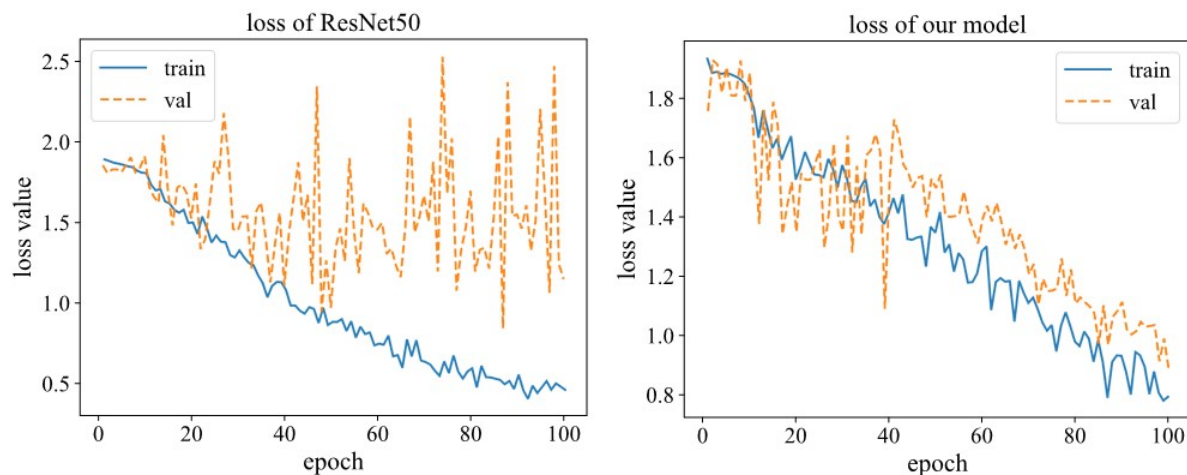
baseline实验需要根据最原始的模型得到一系列性能度量指标，相当于一个基准。在baseline实验中，我们选用ResNet50作为模型提取AV Nicking patch图像特征的backbone，即一个最基本的baseline，并针对baseline和我们设计的模型计算一系列性能度量指标，进而通过对比的方式来分析实验结果并得出结论。baseline网络架构如下图所示：



### 6.2.2 对比分析

在对比实验一节中，我们将分别对上一小节中提出的baseline模型与我们设计的模型在一系列指标上进行对比，从而实现对实验结果的分析与对模型可行性的验证。

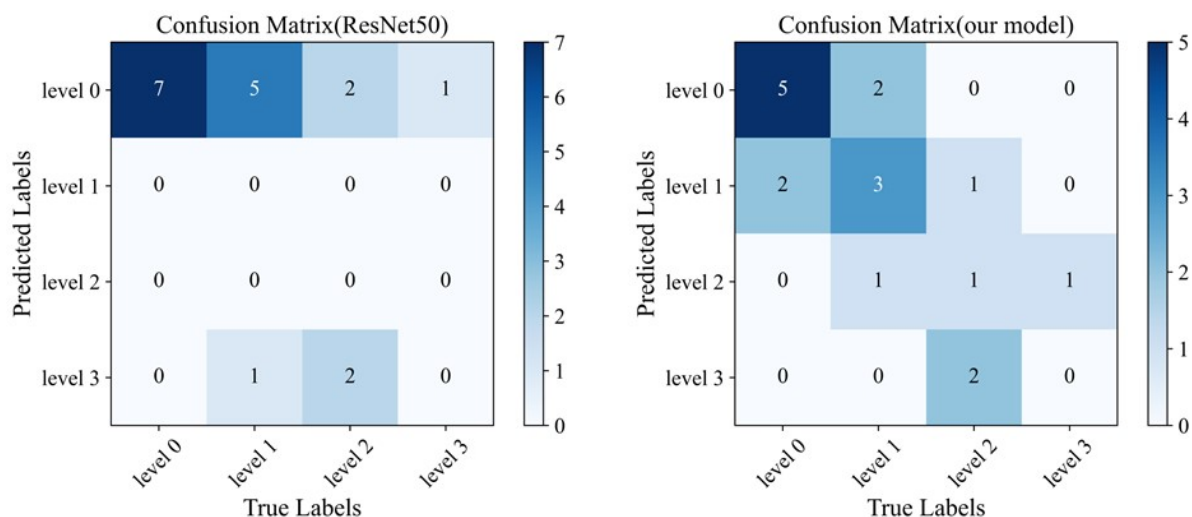
对于baseline实验与我们设计的模型，我们控制数据集划分、数据增强等因素完全相同，我们将模型训练时每个epoch中训练集上产生的loss与验证集上产生的loss保存并绘制成曲线，如下图所示：



从loss曲线中可以看出baseline的ResNet50模型虽然在训练时其loss逐渐下降，但是在验证集上验证loss波动极大，难以收敛；我们提出的模型虽然在训练时loss下降速度相对baseline较慢，但是在约40个epochs后其验证集上的loss能够稳步下降，这一定程度上说明了本模型在训练及验证过程中相对于baseline是更加优越的。

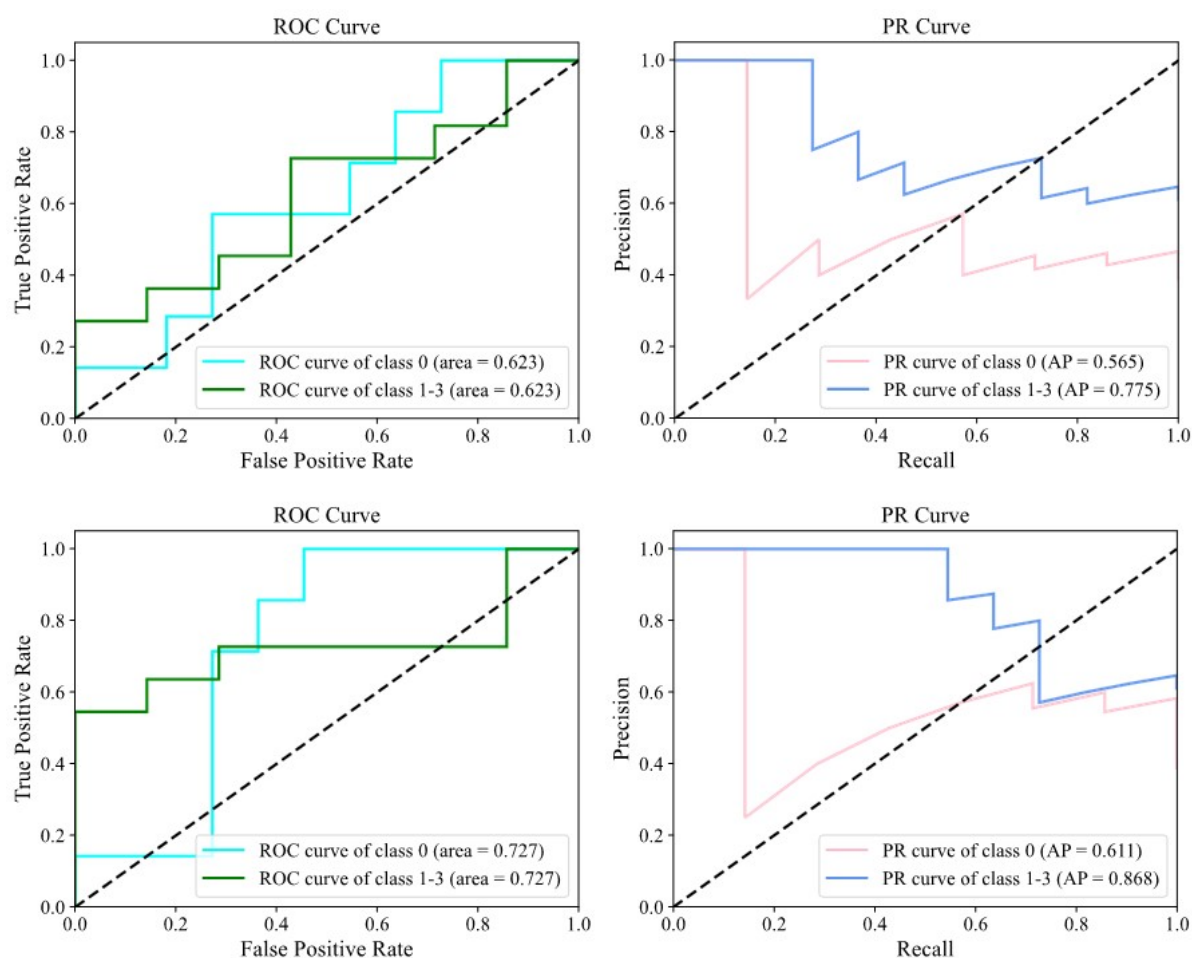
在模型训练完毕之后，我们将得到的模型在测试集上进行测试，并绘制测试结果的混淆矩阵，如下图所示：





由上图我们可以看出，baseline在测试集上的准确率较低，且易将图像误分级为level 0；而我们提出的模型在accuracy、kappa系数、quadratic\_weighted\_kappa系数等指标上，相对baseline都有一定的提升。通过对混淆矩阵进行计算，我们可以得到baseline的accuracy为0.389，kappa系数为0.083，quadratic\_weighted\_kappa系数为0.317。而我们提出的模型的accuracy为0.500，相对于baseline提升了0.111；kappa系数为0.280，相对于baseline提升了0.197；quadratic\_weighted\_kappa系数为0.727，相对于baseline提升了0.410。

由于本实验旨在检测出动静脉交叉压迫征，level 0对应的级别是normal，同时level 0的图像数量较多使得其特征信息更加丰富，因此我们将level 0视为负例，将level 1-3视为正例，并统计了level 0、level 1-3对应的AP等信息，分别绘制了baseline与我们的模型在level 0与level 1-3上的ROC曲线与PR曲线，如下图所示：



由上图可知，我们提出的模型对于level 0和level 1-3而言相对于baseline的ROC曲线更凸向左上方，PR曲线更凸向右上方，体现出其分级效果相对更加优越。

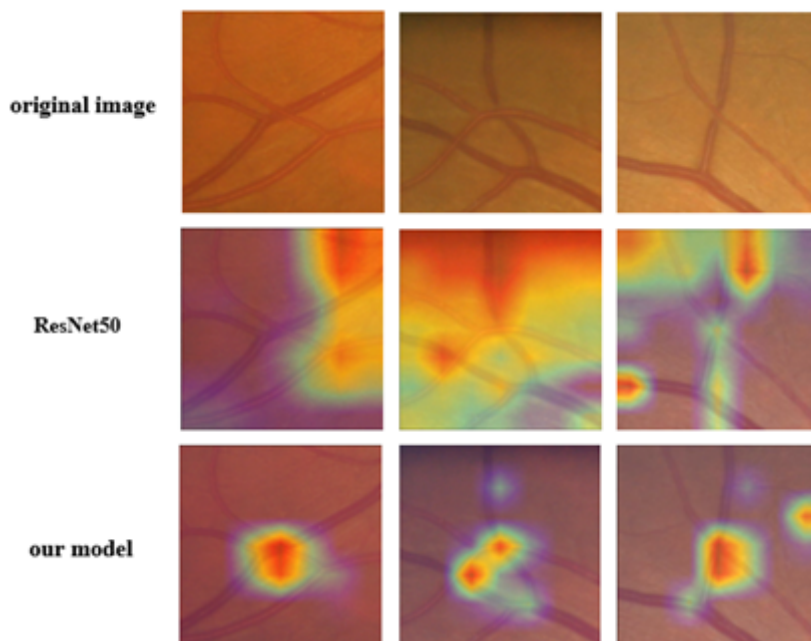


### 6.2.3 可解释性分析

为了证明我们提出的模型对不同级别图像的分级是有依据的，即模型确实关注到了动静脉交叉点处静脉变窄严重程度的不同而作出了对应的判断，在这里我们使用GradCAM热力图来可视化模型对图像分级时的注意力所在区域。

Selvaraju提出的Grad-CAM模型能够帮助我们分析网络对于某个类别的关注区域。因为Grad-CAM(Gradient-weighted Class Activation Mapping)在没有attention的情况下，也能对图像位置进行判别，可以对任意结构的CNN进行可视化，不需要修改网络结构或者重新训练，所以我们使用Grad-CAM对模型所做的决策在视觉上进行可视化的解释。Grad-CAM使用流入CNN的最后一个卷积层的梯度信息为每个神经元分配重要值，以进行特定的关注决策，得到类似于热力图的class activation mapping，可用于定位图像中与类别相关的区域。

我们以AV Nicking数据集集中的三张patch图像为例作出GradCAM热力图，如下图所示：

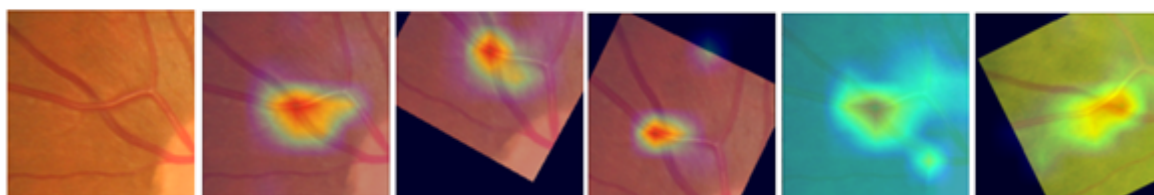


对于Grad-CAM可视化的结果，判定为某种类别c时对应的热力图中偏深红色的地方表示对该区域关注程度比较大，且由于该区域的观察结果将图像类别预测为c，偏深蓝色的地方表示对该区域关注程度比较小，而模型最终预测的类别由模型关注不同类别对应区域的概率值决定，即模型认为如果关注某一类别对应区域的概率值最大，那么就可以将结果预测为该类别。

由上图可知，baseline的ResNet50并不能正确地关注到动静脉交叉点处静脉变窄的区域，这也是其分级准确性与quadratic\_weighted\_kappa等系数不高的原因；而我们提出的模型基本上能够正确地关注到图像中动静脉交叉点处静脉变窄的区域，故其分级的依据是可靠的，其分级的结果相对于baseline也自然就更有说服力。

同时，为了证明我们所作的数据增强是有效的，下面我们选取AV Nicking数据集中的一张图像，进行旋转、平移、亮度饱和度对比度色调的变化，并对经过不同方式数据增强得到的新图像进行模型分级的GradCAM热力图可视化操作，如下图所示：

rotation、translation、brightness、saturation、contrast、hue...



由上图可知，在对原始图像进行旋转、平移、亮度饱和度对比度色调的变化这些不同的数据增强操作之后，模型在新得到的所有图像上均能够正确地关注到动静脉交叉点处静脉变窄的区域，这说明我们所作的数据增强操作成功地解决了AV Nicking的patch数据集数据量少、交叉点位置单一等问题，使模型能够真正地关注到动静脉交叉点这一形态学特征而不被其它无关因素所干扰（如并没有真正地学习到交叉点特征而是一概重点关注图像中心位置区域），从而能够作出合理的分级判断。

## 7. 总结与展望

### 7.1 总结

动静脉交叉压迫征是预测心血管中风、缺血性中风、动脉粥样硬化和高血压的最重要的生物学讯息，其严重程度的正确分级对于发现心血管疾病高危人群并进行早期有效的诊断具有十分重要的作用。

关于动静脉交叉压迫征严重程度正确分级的相关研究也是层出不穷，传统的数字图像处理方法需要手工提取图像特征，过程繁琐，且精度极大地依赖于血管分割这一前置任务的精度。随着深度学习技术的不断发展，其在智能医疗影像处理中发挥着越来越重要的作用。近年以来出现了一些深度学习解决AV Nicking严重程度分级的方案，虽然能提升一定的精度，但是不具备良好的可解释性，而且需要消耗大量的计算资源与时间。

对此，我们提出了一种基于迁移学习思想的双流多模态卷积神经网络算法，该算法使用迁移学习的思想，先训练上游动静脉交叉点目标检测及其附近血管实例分割的任务，再将上游任务学习到的特定域的先验知识迁移到下游动静脉交叉压迫征严重程度分级任务当中，并综合使用原图与对应的动静脉多分类图信息，构建信息互补的两支CNN输入流，最终解决AV Nicking严重程度分级问题。

对于数据量稀缺的问题，我们通过各种数据增强方式以及使用半监督算法作为优化策略来丰富数据集的特征分布；对于数据集中不同级别样本比例分布不均衡的问题，我们通过合理控制训练过程中不同级别样本被抽样的概率来缓解；对于交叉点处静脉变窄这一细粒度特征难以提取的问题，我们通过采用迁移学习的思想，让模型在进行AV Nicking严重程度分级任务之前，先熟悉特定域的眼底彩照动静脉交叉点的图像特征，在Backbone的FPN部分中添加了设计的SAFPN模块，以对高分辨率更小的交叉点与低分辨率更大的交叉点这两种容易误判的情况进行正确的分级，在Backbone的ResNet50部分中添加了SENet模块以学习特征图的通道注意力，使用原图与动静脉多分类图作为两支输入流以互补信息，使用双线性卷积操作以提高模型提取细粒度特征的能力。

为了验证我们提出的模型相对于设定的baseline更加优越，我们将提出的模型与baseline进行对比实验并统计各项指标，结果我们的模型在众多指标上表现的更加优秀；同时，我们用热力图对模型的注意力进行了可视化，一定程度上证明了我们的模型对AV Nicking严重程度的分级依据是合理的。

### 7.2 展望

我们提出的模型尽管在一些方面取得了性能上的提升，但是还存在这样一些问题：

- (1) 原先的提取动静脉交叉点附近的动静脉血管掩膜标注信息的方式不够精确，存在误差，这可能会对后面的实例分割模型以及AV Nicking严重程度分级模型的性能造成一定的负面影响。
- (2) 我们提出的模型对相邻级别图像仍存在易误分级的问题，例如易将级别2误分级为级别1、级别3误分级为级别4等。
- (3) 我们设计的动静脉血管分类算法存在易将部分动脉误分类为静脉和将部分静脉误分类为动脉的问题，这可能会对后面的双流多模态网络的性能造成一定程度的负面影响，故其在性能上尚有待提升。

对于这些问题，我们将继续探索并寻求解决方案。