

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

课程论文

COURSE PAPER



论文题目: 彩色眼底视网膜图像预处理

学生姓名: 姜燕炜、曲博文、刘学楷、王飞、王梓烨

学生学号: 518030910{274、279、309、311、313}

课程名称: 数字图像处理

指导教师: 王利生教授

学院(系): 电子信息与电气工程

目 录

第一章 作业要求	1
1.1 作业目的	1
1.2 作业题目	1
1.2.1 空域对齐	1
1.2.2 颜色归一化	1
1.2.3 血管检测与填充	1
1.3 目标结果	1
第二章 空域对齐	3
2.1 中心点检测	3
2.1.1 问题分析	3
2.1.2 方法原理	3
2.1.3 训练模型	4
2.2 空域对齐	5
2.3 结果	6
第三章 颜色归一化	7
3.1 问题分析	7
3.2 方法原理	7
3.3 结果	7
第四章 血管检测与填充	9
4.1 血管检测	9
4.1.1 问题分析	9
4.1.2 方法原理	9
4.2 血管填充	10
4.3 结果	10
全文总结	12
参考文献	14

第一章 作业要求

1.1 作业目的

本次作业是课程数字图像处理的第二次大作业，内容为对彩色眼底视网膜图像进行预处理，实现尺寸、位置、颜色的归一化，用于以后的病灶识别任务，主要目的为熟悉并融会贯通本学期所学内容的知识点，并了解人工智能领域在数字图像处理中的应用。

1.2 作业题目

1.2.1 空域对齐

检测视盘中心和黄斑中心，将所有图像的视盘中心和黄斑中心对齐。

1.2.2 颜色归一化

由于不同图像是在不同光照下采集的，本文需要对每幅图像做颜色归一化处理，保证算法可适应不同成像条件下的眼底视网膜图像。

1.2.3 血管检测与填充

为了消除不同视网膜图像中血管的个性化差异带来的干扰，自动检测每幅图像的血管，并根据血管区域的上下文信息对血管区域进行修复。

1.3 目标结果

目标结果如图 1-1所示。

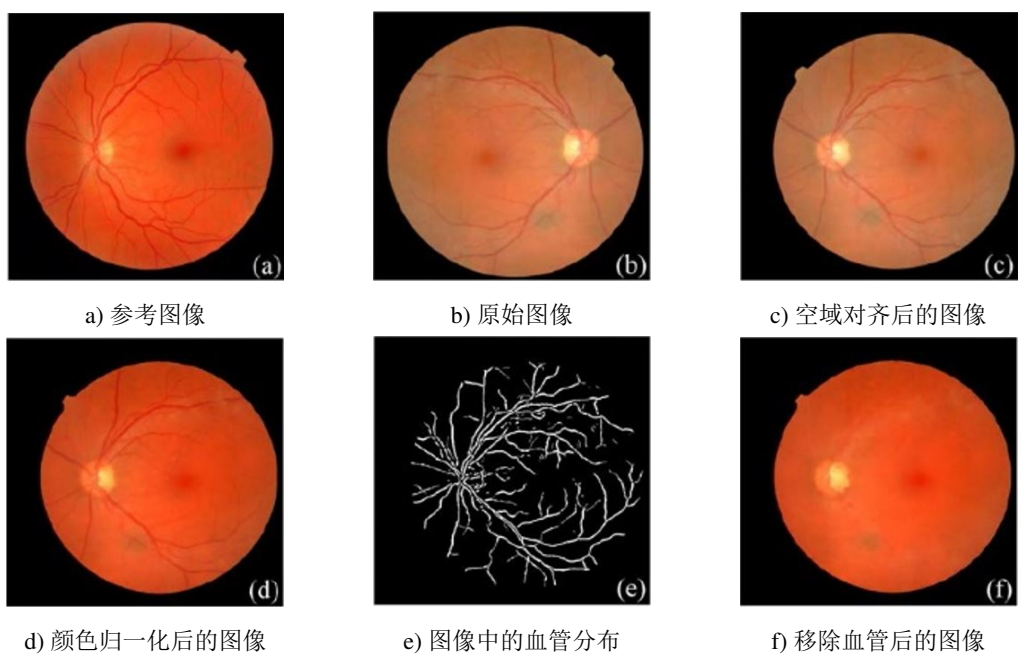


图 1-1 目标结果

第二章 空域对齐

2.1 中心点检测

2.1.1 问题分析

为了更好地识别病灶，我们需要对照片做预处理，其中第一个步骤就是要先把照片放到相同的位置上，也就是空域对齐。空域相当于归一化处理，处理之后方便后面的血管提取，颜色归一化等操作。

所谓相同的位置就是旋转平移等的定位点，用以确定平移的位置。下一节会具体介绍空域对齐的方法，我们先使用黄斑中心和视盘中心作为两个定位点。那么我们识别出黄斑中心和视盘中心就是很有必要的了。根据给出的论文^[1]给出了一种经典方法，这种方法是基于图像的特征的分类法，使用 Knn 聚类对每个像素点逐像素点分类；此外还提到了一种图论方法。但是我们在讨论的过程中，考虑到本任务中视盘和黄斑的图像特征非常明显，尤其是视盘，往往亮度较高；而且视盘和黄斑呈现出空间聚集性，很利于机器学习；再加上我们有很多标注的数据集，图像大小也较小，方便处理，所以我们最终决定不采用经典方法，而是采用基于神经网络的机器学习方法。

2.1.2 方法原理

本任务是一个目标定位的问题，机器学习领域各种目标识别网络层出不穷，例如 RCNN，YOLO，SSD。其中 YOLO，SSD 是 One Stage 方法，而 RCNN 是 Two Stage 方法。这两种方法的差异性体现在速度和精度，One Stage 精度低且快，Two Stage 精度高且慢。具体原因会在下面介绍。出于样本并不复杂和精度考虑，根据查找到的资料，我们锁定了两种网络，RetinaNet 和 FasterRCNN。

2.1.2.1 FasterRCNN

图 2-1 为 FasterRCNN^[2] 的示意图和网络结构。FasterRCNN 的图像特征提取网络是 CNN，也就是一系列的 conv+relu+pooling 经典组合，其中还包含了金字塔形的 FPN 网络。

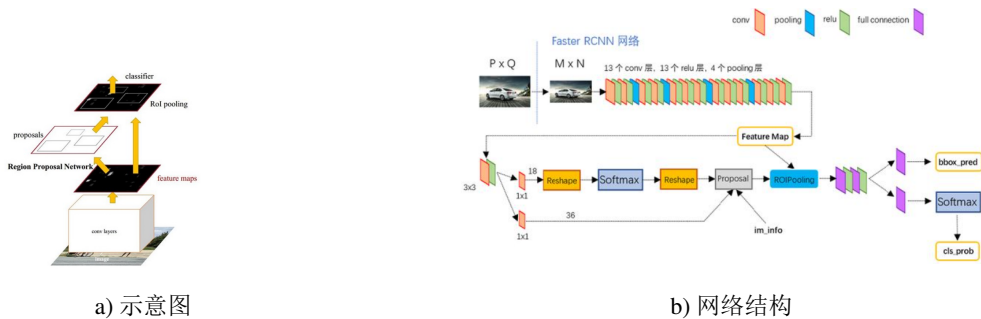


图 2-1 FasterRCNN 网络结构

特征提取完毕后，为了确认每个区域是否是我们要检测的视盘中心或者黄斑中心，对此 FasterRCNN 采用的是 RPN 网络，对图片的大部分区域都是用一个滑窗过一遍。如果在提取

到的特征图上滑动，那么映射回原图的部分就叫做 anchor。对 anchor 使用不同缩放比例，就可以生成几个 proposal（取决于你的先验知识）。对每一个 proposal 采用经典的 softmax 层判断类别。

接下来为了保证输入到最终分类器 boxpredictor 中的图片尺寸合适，使用 roipooling，处理成我们想要的大小，最后送入最终分类器处理。

2.1.2.2 RetinaNet

关于 RetinaNet^[3]，这也是一个基于 CNN 的网络，不过其 backbone 是大名鼎鼎的 Resnet 来提取特征。网络结构如图 2-2

需要注意的是，RetinaNet 采用了 Focal loss，Focal loss 是用来解决样本失衡的问题的。样本失衡的问题，详细来说就是，在机器学习的数据集中，negative 的样本所占的数量太多，导致在 loss 函数中，起作用的基本上都是 negative 的样本，而 positive 所占的权重很小。这样的 loss 对于挖掘 positive 的信息作用很小。这也是一个机器学习领域老生常谈的问题。

Focal loss 在常用的交叉熵损失函数前都加了一个系数： $(1 - p_t)^\gamma$ ，其中 γ 为预先设定好的常数， p_t 是预测为正或者负的概率。在应用了 Focal loss 损失函数之后，数据集中的困难样本得到了重视，所以这个 loss 函数是更科学一点的。

提取完特征之后就是类似 FasterRCNN 的分类检测结构，不再赘述。当然，RetinaNet 也有 anchor 机制。

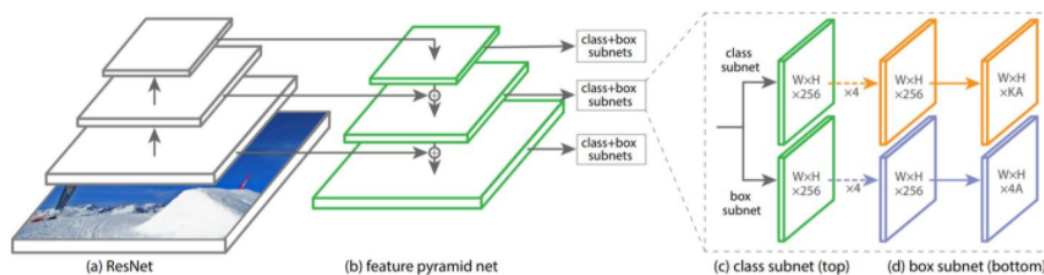


图 2-2 RetinaNet 网络结构

2.1.2.3 模型选择

根据上面的介绍，我们引出 One Stage 和 Two Stage 的最大区别。Two Stage 先进行特征提取，然后对于所有的 anchor 的每一个 proposal 进行过滤，最终经过一次选择再估算剩余样本，训练网络和监测网络都很复杂。但是 One Stage 一般仅是特征提取之后对划分好的 region 进行 softmax 层检测，训练和检测都很简单。

2.1.3 训练模型

需要声明的是，由于 FasterRCNN 与 RetinaNet 都有很多实现的代码，而且从头实现的工作量和挑战程度很大，所以我们并非完全是自己实现。

关于数据集，老师给出的数据集数量不是很大，可能没法训练出效果令人满意的神经网络。因此我们讨论决定，自己上网找更大的数据集。最终我们采用了 ‘IDRiD dataset’，包括 413 张训练集和 103 张测试集。

训练环境是 windows10 下的 Anaconda，硬件为 GTX 1050，cuda9。

两种 models 的训练曲线都比较正常，最终 loss 都收敛。为了验证效果，我们挑选了一些图片用以训练，挑选标准为既挑选黄斑视盘都清晰可见的，也挑选黄斑不明显的。RetinaNet 部分结果如图 2-3 的 a)，b)，c)，可以看到结果不是完全尽如人意；FasterRCNN 部分结果如图 2-3 的 d)，e)，f)，可以看到有的图片因为黄斑太弱检测不到。

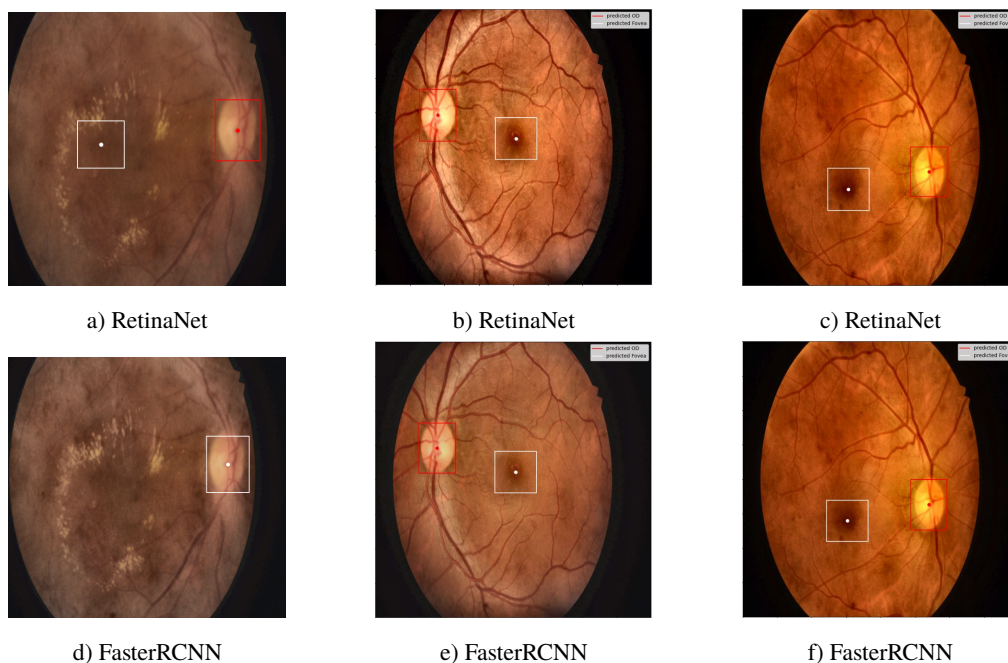


图 2-3 视盘中心和黄斑中心识别结果

总结一下，两种网络的结果都还可以，但无论哪种都存在局限性，在图片质量很差，最典型的是黄斑极弱的时候，可能无法检测到黄斑。如果用肉眼看来和 groundtruth 没有显著偏离的标准衡量，准确率在 90% 以上，在 train 和 test 一共四十张照片里，只有三张因为黄斑不明显检测不出黄斑，其余的都效果很好，和 groundtruth 的差异都在 10 个像素点以内。

分析了一下两种网络在大作业附件中数据集的表现，RetinaNet 同时检测出黄斑和视盘的概率要比 FasterRCNN 稍好，而且最终测试集上的 loss 也是 RetinaNet 要低一些，所以最终选定 RetinaNet 训练出的样本作为最终方案。

检测效果可以见上图所示，另外这一步是和空域对齐绑定的，可以结合下一个部分的空域对齐查看效果。

2.2 空域对齐

然后将我们得到的定位好的结果进行空域对齐，根据多次尝试发现，将视盘中心固定在图片的距左边缘 $3/16$ 处，距上边缘 $1/2$ 处，将黄斑中心固定在图片的距左边缘 $1/2$ 处，距上边缘 $1/2$ 处得到的图像最为完整且大小适当，据此对图像进行平移缩放旋转得到空域对齐的结果。

2.3 结果

图 2-4是空域对齐的结果，可以看到我们完成了视盘中心，黄斑中心的检测且完成了对齐。

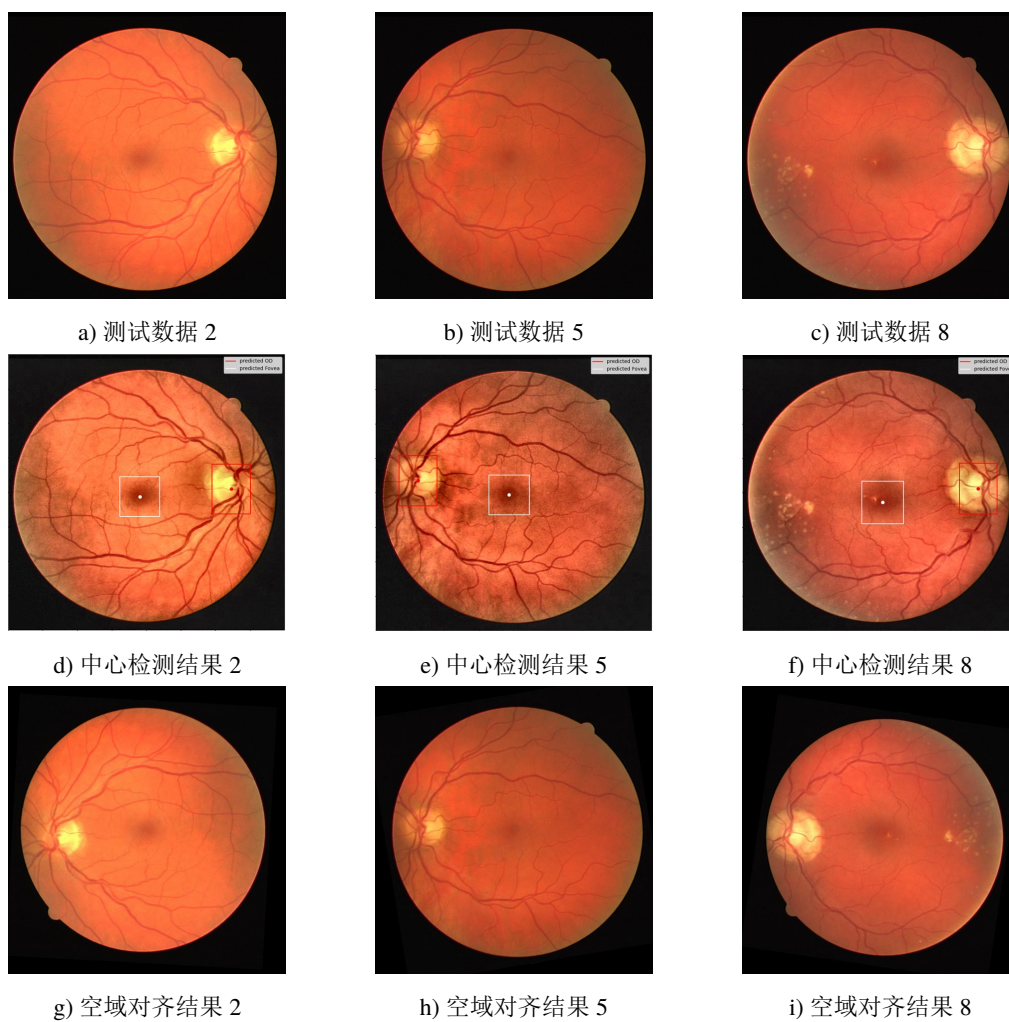


图 2-4 视盘中心和黄斑中心识别结果

第三章 颜色归一化

3.1 问题分析

由于不同图像是在不同光照下采集的，因此需要对每幅图像做颜色归一化处理，保证算法可适应不同成像条件下的眼底视网膜图像。

我们使用传统方法对图像进行色彩归一化。

3.2 方法原理

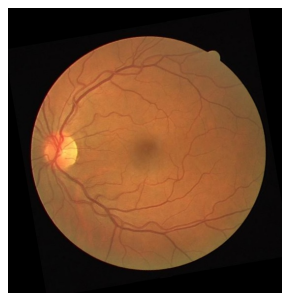
设 Ψ 为需要进行颜色归一化的图像的集合。分别计算参考图像和 Ψ 中的每幅图像的均值和方差，用 σ_1 和 σ_2 分别表示参考图像和 Ψ 中每幅图像的方差， μ_1 和 μ_2 分别用于表示参考图像和 Ψ 中每幅图像的均值。用 F_2 表示一幅来自亚的视网膜图像，而用 F_1 表示对 F_2 做颜色归一化后的图像，可用如下公式来归一化中的每幅图像：

$$F_1 = \frac{\sigma_1}{\sigma_2}(F_2 - \mu_2) + \mu_1$$

上述归一化公式使得 Ψ 中每幅图像和参考图像具有相同的均值和方差，使得每幅图像在整体上具有相似的颜色对比度。如下图所示，待处理图像中的 3 幅视网膜图像颜色和整体亮度与参考图像相差较大，但是通过亮度归一化之后，这些归一化后的图像的颜色和对比度与参考图像基本一致。通过颜色归一化处理，可大大减小由于不同视网膜光照变化带来的干扰。

3.3 结果

图 3-1 是根据我们选择的参考图像进行颜色归一化的结果，可以看到我们完成了颜色归一化。



a) 参考图像



b) 测试数据 2



c) 测试数据 5



d) 测试数据 8



e) 归一化结果 2



f) 归一化结果 5



g) 归一化结果 8

图 3-1 颜色归一化结果

第四章 血管检测与填充

4.1 血管检测

4.1.1 问题分析

为了消除不同视网膜图像中血管的个性化差异带来的干扰，自动从 $\Psi \cup \{F\}$ 中检测每幅图像的血管，并根据血管区域的上下文信息对血管区域进行修复。

将血管从眼底图片中分割出来，这是属于语义分割的问题。虽然，传统的数字图像处理方法也有例如区域生长，阈值分割，分离合并等方法，但是这些方法都是基于较强的假设，很难去界定相似性的问题，在实际应用的效果并不是太好。当神经网络应用于图像处理，语义分割的效果得到极大的提高。但是常常神经网络的训练需要大量的标注数据，但是在医学图像邻域，标注数据难以收集，所以我们要选用对小数据依旧能够表现较好的 U-net 作为分割网络。

4.1.2 方法原理

4.1.2.1 U-net

U-net^[4] 是一种比较年轻的神经网络，顾名思义，其结构为 U 形，在卷积层一层一层提取特征的同时将首尾对称的两层联系起来，结构图如图 4-1 所示。

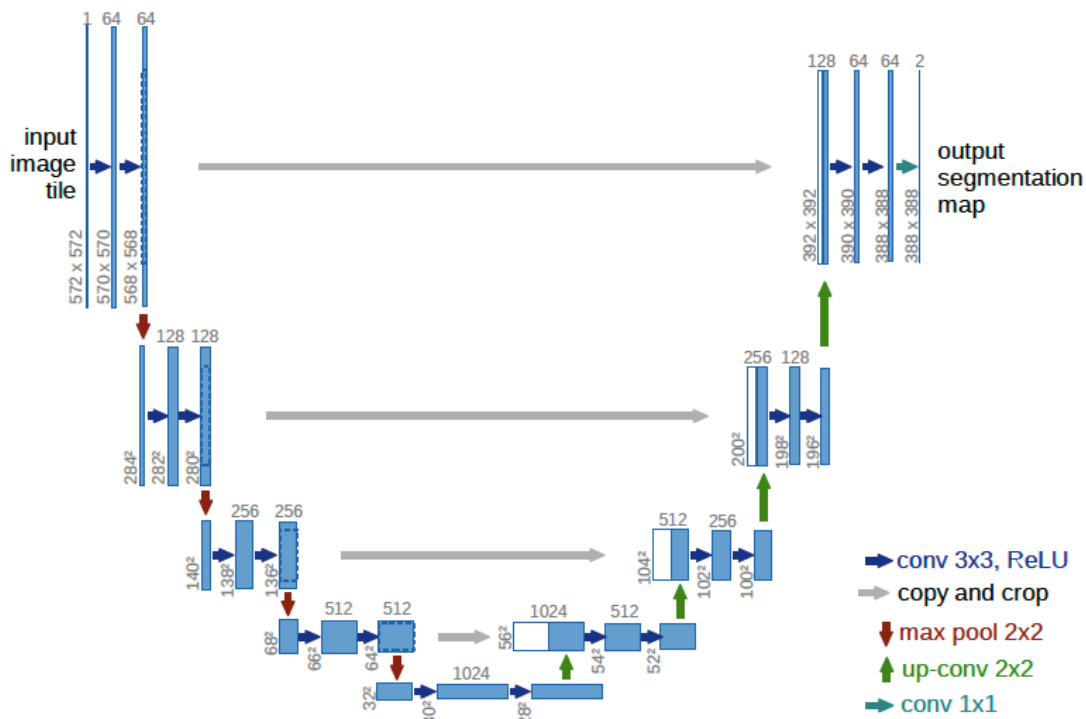


图 4-1 FasterRCNN 示意图

U-net 的 encoder 下采样 4 次，一共下采样 16 倍，对称地，其 decoder 也相应上采样 4 次，将 encoder 得到的高级语义特征图恢复到原图片的分辨率。除输出层外，U-net 每一层以三

层卷积构成，层间采用池化或上采样的方法实现特征的提取和整合，最后一层将之前所有提取出来的特征做一个二分类，实现语义分割。U-net 共进行了 4 次上采样，并在同一个阶使用了 skip connection 操作，而不是直接在高级语义特征上进行监督和损失反传，这样就保证了最后恢复出来的特征图融合了更多的低阶的特征，也使得不同阶层的特征得到了融合，从而可以进行多尺度预测和深度监督。4 次上采样也使得分割图恢复边缘等信息更加精细。

4.1.2.2 模型优化

我们发现在 U-net 在 max pool 过程中会让图片变小，这种变小是在上卷积的过程中无法恢复的，考虑到数据集图像分辨率本来就不高，而且手工标注图像和眼底图像的尺寸是一致的，如果修改手工标注图像的尺寸，会导致标注图像信息损失，有些细小血管会在改变尺寸过程中断开，所以在 max pool 过程中，我们对边缘进行了 padding，保证图片分割过程中尺寸不改变。

4.1.2.3 数据增强

首先训练一个神经网络最重要的就是数据集，所以我们又在网络上下载了 STARE 数据集手工标注部分，混和组成了当前使用的数据集。

同时由于数据集数量少，我们也对数据集进行了数据增强，在训练过程中逐渐增大训练图片颜色，对比度，饱和度等变化，希望能够增强网络对于光照，对比度，饱和度等变化的不变性。

同时我们希望网络能够在细节表现更好，所以在网络对完整图像训练几乎收敛之后，我们采用随机剪裁的方法，剪裁了不同大小的图片块进行训练。

4.2 血管填充

检测出血管后，使用快速修复算法 (Fast Marching Method, FMM) 对检测出的血管区域进行修复。快速修复算法的原理是通过对像素点逐一操作来达到修复的目的。在本实验中，我们首先将分割出的血管使用 3×3 的十字模板进行膨胀，然后将膨胀出来的边框作为边界。根据边界，我们将像素点分成三类并用 flag 表示：BAND：在边界上的像素点；KNOWN：边界外不需要修复的像素点；INSIDE：边界内待修复的像素点。另外，每个像素点还需要存储两个值：T（该像素离到边界的距离）；I（灰度值）。将 BAND 和 KNOWN 中的像素点的 T 设置为 0，将 INSIDE 中的像素点的 T 设为 10^6 （表示无穷大）

将 BAND 中的像素点按 T 的大小组成队列，然后优先处理 T 最小的像素点，将其 flag 改为 KNOWN，随后依次检测与之相邻的四个像素点的 flag，如果为 INSIDE，若是则重新计算 I，修复该点，并更新其 T 值，修改该点类型为 BAND，根据 T 值大小加入队列；如果不为 flag 则跳过。依次进行，每次处理的都是队列中 T 最小的像素点，直到队列中没有像素。

4.3 结果

由图 4-2 是进行血管检测与填充的结果，可以看到血管区域被很好地检测与修复，取而代之的是平滑的过渡区域。



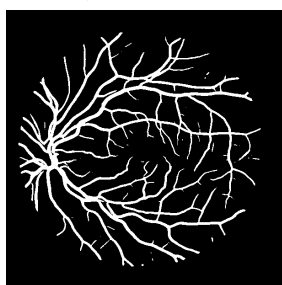
a) 测试数据 2



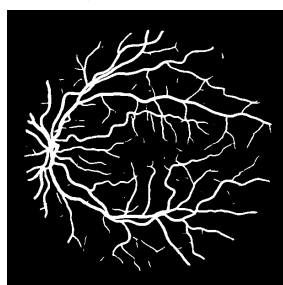
b) 测试数据 5



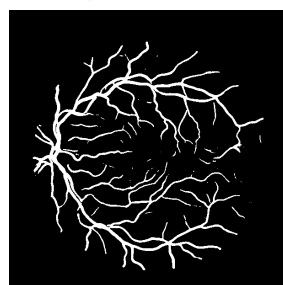
c) 测试数据 8



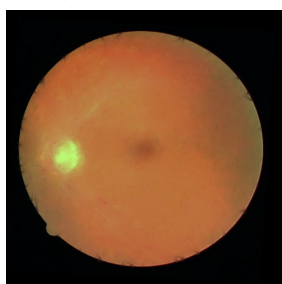
d) 血管检测结果 2



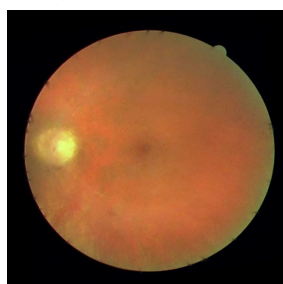
e) 血管检测结果 5



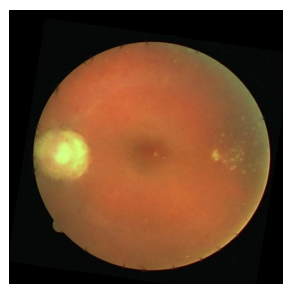
f) 血管检测结果 8



g) 修复结果 2



h) 修复结果 5



i) 修复结果 8

图 4-2 血管检测与填充结果

全文总结

本次作业是对彩色眼底视网膜图像进行预处理，实现尺寸、位置、颜色的归一化，用于以后的病灶识别任务，主要目的为熟悉并融会贯通本学期所学内容的知识点，并了解人工智能领域在数字图像处理中的应用。将数字图像处理课堂知识与实际内容相结合，并熟悉该领域的前沿知识。围绕完成题目要求的几个问题来熟悉处理图像的过程。

首先进行视盘中心检测和黄斑中心检测，在经典方法和深度学习方法中选择了深度学习 RetinaNet 方法在 IDRiD 数据集上进行训练得到结果。其次进行空域对齐，将每幅图片进行平移旋转放缩来对两个中心进行对齐。再次进行颜色归一化，依据第一张测试图像将所有测试数据的颜色直方图进行均值方差的归一化。然后进行血管的检测，在经典方法和深度学习方法中选择了深度学习 U-net 方法在 STARE 数据集先进行增强再进行训练得到结果。最后进行血管的填充，使用 FMM 算法对检测出的血管进行修复。由于篇幅原因图 4-3 展示测试集中的第二、五、八张图片处理过程和结果。

本次实验我们学习到了如何将数字图像处理的内容应用于实际，同时也在与同学的合作上学习到了很多的知识，老师和助教也在学习方面给予我们很大的帮助。我们也学会了在遇到问题时从其他角度想出替代的解决方法。感谢老师和助教在我们学习这节课以来给我们带来的帮助和指导。

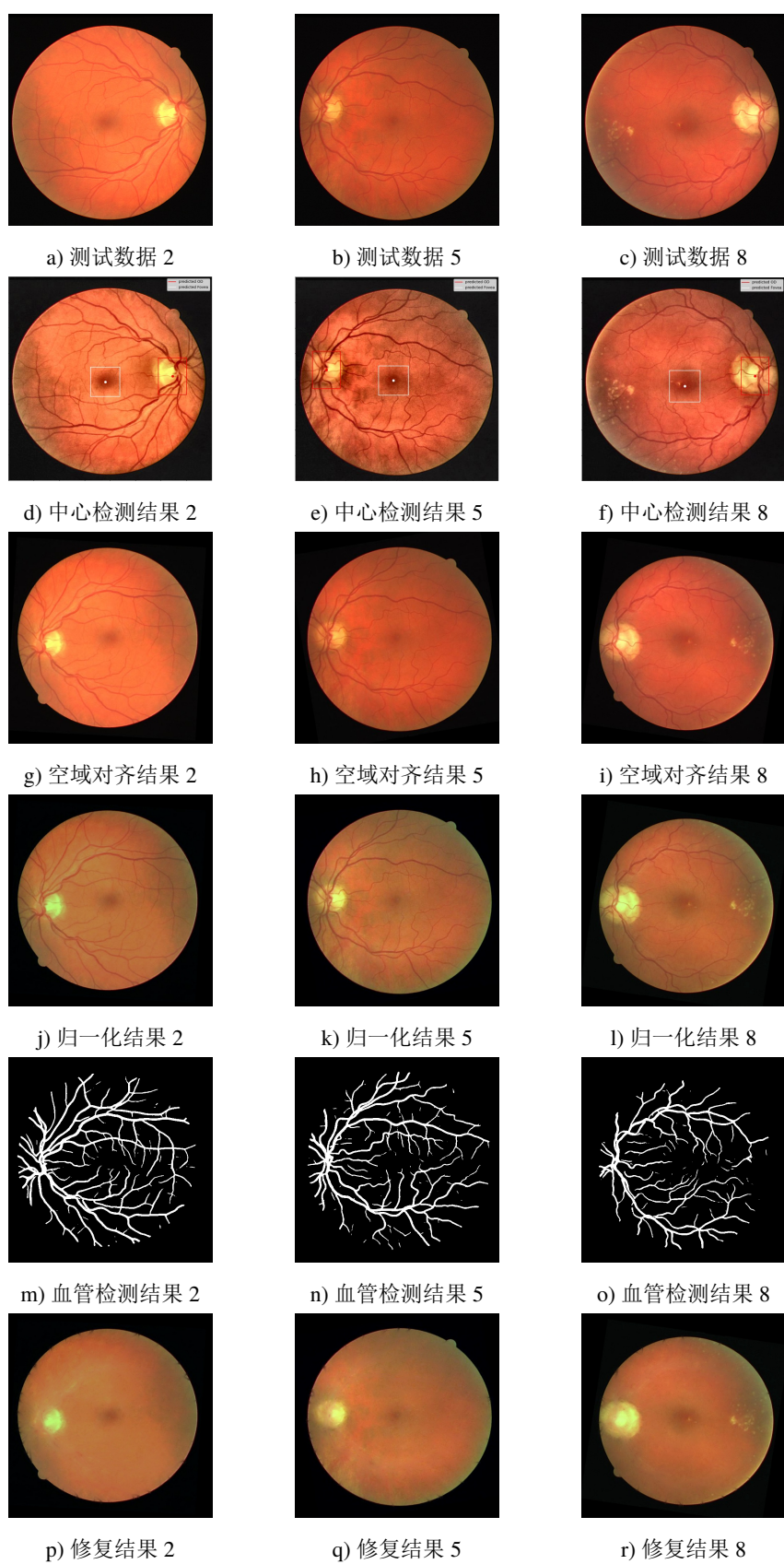


图 4-3 血管检测与填充结果

参考文献

- [1] Abràmoff M D, Garvin M K, Sonka M. Retinal imaging and image analysis[J]. IEEE reviews in biomedical engineering, 2010, 3: 169-208.
- [2] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28: 91-99.
- [3] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]. in: Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.
- [4] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]. in: International Conference on Medical image computing and computer-assisted intervention. 2015: 234-241.