

DOI:10.16136/j.joel.2023.01.0417

# 基于 CTCNet 对球结膜进行糖尿病视网膜病变分类

黄丝雨<sup>1</sup>, 刘凤连<sup>1</sup>, 汪日伟<sup>2\*</sup>

(1. 天津理工大学 计算机视觉与系统教育部重点实验室和天津市智能计算及软件新技术重点实验室, 天津 300384; 2. 温州理工学院, 浙江 温州 325035)

**摘要:** 糖尿病视网膜病变 (diabetic retinopathy, DR) 是一种糖尿病性微血管病变, 会在球结膜微血管上有所体现, 球结膜血管图像的获取比眼底图像更加便捷, 但微血管的特征变化微小且难以量化。为了能够对患者进行早期辅助诊断, 本文依据球结膜微血管形态与 DR 的关联, 首先对球结膜图像进行预处理, 使用限制对比度自适应直方图均衡 (contrast limited adaptive histogram equalization, CLAHE) 算法进行图像增强, 随机处理使数据增强, 然后结合卷积神经网络 (convolutional neural network, CNN) 和 Transformer 各自的网络优势构建 CTCNet, 对处理后的球结膜血管图像进行 DR 分类, 分类准确率达到了 97.44%, 敏感度 97.69%, 特异性 97.11%, 精确度 97.69%, 通过实验对比 CNN 和 Transformer, CTCNet 网络性能优于其他模型, 能够有效识别 DR。

**关键词:** 卷积神经网络 (CNN); Transformer; 糖尿病视网膜病变 (DR); 球结膜血管  
**中图分类号:** TP311 **文献标识码:** A **文章编号:** 1005-0086(2023)01-0100-07

## Classification of diabetic retinopathy in bulbar conjunctiva based on CTCNet

HUANG Siyu<sup>1</sup>, LIU Fenglian<sup>1</sup>, WANG Riwei<sup>2\*</sup>

(1. Key Laboratory on Computer Vision and System, Ministry of Education of China, Key Laboratory on Intelligence Computing and Novel Software Technology of the City of Tianjin, Tianjin University of Technology, Tianjin 300384, China; 2. Wenzhou University of Technology, Wenzhou, Zhejiang 325035, China)

**Abstract:** Diabetes retinopathy (DR) is a kind of diabetes microvascular disease, and it will be reflected in the bulbar conjunctival microvessels. Images of bulbar conjunctival vessels are easier to obtain than fundus images, but the characteristic changes of microvessels are subtle and difficult to quantify. In order to enable early auxiliary diagnosis of patients, according to the association between the morphology of bulbar conjunctiva microvessels and DR, this paper first preprocesses the bulbar conjunctiva images, enhances the image using the contrast limited adaptive histogram equalization (CLAHE) algorithm, and enhances the data with random processing. Then CTCNet is constructed by combining the advantages of convolutional neural network (CNN) and Transformer, and the processed images of bulbar conjunctival vessels are classified into DR. The classification accuracy reaches 97.44%, sensitivity 97.69%, specificity 97.11%, and accuracy 97.69%. Through experimental comparison between CNN and Transformer, CTCNet has better performance than other models and can effectively identify DR.

**Key words:** convolutional neural network (CNN); Transformer; diabetic retinopathy (DR); bulbar conjunctival vessels

## 0 引言

糖尿病视网膜病变 (diabetic retinopathy, DR) 是糖

尿病的严重并发症之一<sup>[1]</sup>, 也是四大致盲眼病之一, 其对视力功能的损害及其严重且不可逆。现在主流的视网膜病变诊断方法是通过一系列直接或间接的方式观

\* E-mail: wangrw@wzu.edu.cn

收稿日期: 2022-08-28 修订日期: 2022-10-31

基金项目: 国家自然科学基金 (62020106004) 资助项目

察患者的眼底情况,需要专业的医生和复杂的仪器设备,为此许多学者致力于研究计算机辅助诊断视网膜病变的方法,诊断成本低且操作简便。由于眼底视网膜图像中血管形态不明显,需要对眼底视网膜血管图像进行色彩调整、血管增强<sup>[2]</sup>、血管分割<sup>[3]</sup>等操作,在此基础上已有许多深度学习网络模型应用于眼底视网膜图像的分类训练<sup>[4-6]</sup>,并取得了优秀的研究成果。但眼底检查需要复杂且昂贵的设备以及有经验的技术人员,存在眼底照相时的闪光可能会造成患者不适,青光眼和白内障患者的眼底图像不清晰等问题。球结膜是介于眼睛穹隆结膜与角膜缘间部分的结膜<sup>[7]</sup>,覆盖在眼睛巩膜的一种密集的血管化粘液膜,具有直接观察和非侵入性成像的独特优势,许多微血管疾病会反映在球结膜上,获取球结膜血管图像不需要复杂专业的设备,方便快捷且便于观察。

基于球结膜血管形态与 DR 之间的关联,已有许多学者进行研究并得出结论。OWEN C G<sup>[8]</sup>等采用计算机自动算法对数字结膜无红图像进行分析,识别血管轴并量化血管弯曲度,得出糖尿病持续时间越长,总血管弯曲率越低的结论;BANAEI T<sup>[9]</sup>等对 30—60 岁的糖尿病患者和正常人的无红结膜照片进行了拍摄,提出基于 Radon 变换的血管分割算法,分析发现糖尿病与结膜大血管收缩有关,DR 时血管扩张较小;IROSHAN K A<sup>[10]</sup>等利用上球结膜大血管弯曲检测糖尿病,结果显示糖尿病患者的球结膜上大血管弯曲度明显小于非糖尿病组;GUNAY M<sup>[11]</sup>等提出一种自动识别结膜微血管图像的方法,对结膜微血管图像进行精细结构分析,区分非糖尿病组和糖尿病组患者不同阶段的结膜图像,自动化方法对临床和非临床 DR 的识别率分别超过 80% 和 70%,所有自动化的识别率都高于由人类观察者确定的比率。由此可见:球结膜微血管形态与 DR 之间存在关联,且图像识别处理以及计算机自动化方法在球结膜血管分类应用中是有意义的。

为了能够帮助患者早发现、早治疗,本文构建深度学习网络模型,利用科技的手段来辅助诊断 DR。2017 年谷歌提出 Transformer<sup>[12]</sup>模型,在自然语言处理(natural language processing, NLP)领域已有许多成果,其多头注意力机制使得此模型的全局性更强,有利于捕获远距离特征,将 Transformer 的思想应用于计算机视觉成为一种新的尝试方向。现已有学者尝试将 Transformer 应用于医学视觉领域,DAI Y<sup>[13]</sup>等结合卷积神经网络

(convolutional neural network, CNN)与 Transformer 提出 TransMed 方法,用于针对腮腺肿瘤术前诊断的医学影像分析,虽然取得了极具竞争力的结果,但初步的尝试仍然存在许多问题与挑战。CNN 中卷积核的计算方式使 CNN 对于整个图像全局特征的关注是有限的,因此本文将 Transformer 的多头注意力机制融入 CNN 中,提高 CNN 的全局性,优化深度学习分类网络,对球结膜血管图像进行 DR 分类,提高对 DR 的识别率。

## 1 构造球结膜血管数据集

### 1.1 球结膜血管图像的获取

目前没有公开的球结膜数据集,通过眼科医院对健康和 DR 患者进行志愿者征集,采用传统的裂隙灯生物显微镜和单反相机(佳能 EOS7D)搭建而成的数码裂隙灯采集系统,对志愿者的球结膜血管进行拍摄。该系统可以在不损失图像质量的情况下放大 40 倍,所有图像均由同一专业眼科医生在相同环境中拍摄,共征集志愿者包括 DR 患者 143 人,非 DR 患者 20 人,拍摄获得 DR 球结膜血管图像 286 张,非 DR 球结膜血管图像 40 张,数据集数量结构见表 1。

表 1 球结膜血管图像数据集

Tab. 1 Data set of bulbar conjunctival vascular images

Image category	Label	Number of images	Percentage/%
Health	0	20	12.2
DR	1	143	87.3

拍摄前,嘱志愿者闭眼且静坐 5 min,眼科医师调整好下颌托位置后,嘱志愿者向鼻侧转动眼球,并且快速拍摄以减轻眼表干燥和刺激,拍摄过程不会使用闪光灯,不会造成志愿者不适,属于无创成像,拍摄形成的球结膜血管图像如图 1 所示。

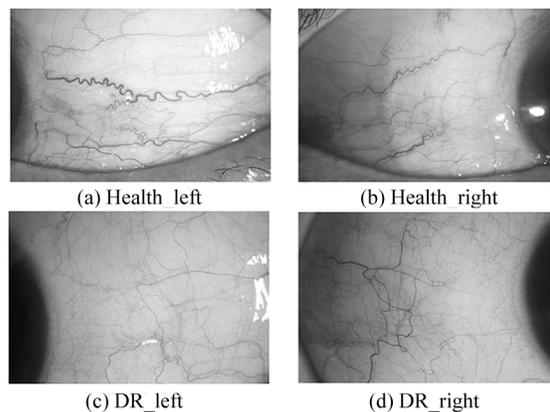


图 1 拍摄形成的球结膜血管图像

Fig. 1 Images of the forming bulbar conjunctival vessels were taken

### 1.2 球结膜血管图像预处理

如图 1 所示,获取的球结膜血管图像存在光照不均匀、光斑、睫毛等影响因素,上述干扰因素可能对训练结果造成一定影响,因此需要对球结膜图像进行预处理,去除干扰因素。处理手段包括图像裁剪、图像增强、数据增强。

首先,对所有球结膜图像进行裁剪,去除较为明显的干扰因素,如睫毛、黑色眼球、眼睑、光斑等,裁剪出结膜中较为清晰的血管图像。裁剪原则为:

- 1) 去除所有干扰因素的前提下,裁剪得到的图像尽可能多地包含所有血管特征;
- 2) 若裁剪所得图像无法包含所有血管特征,可以分块裁剪出两张或两张以上的图像,如图 2 所示。

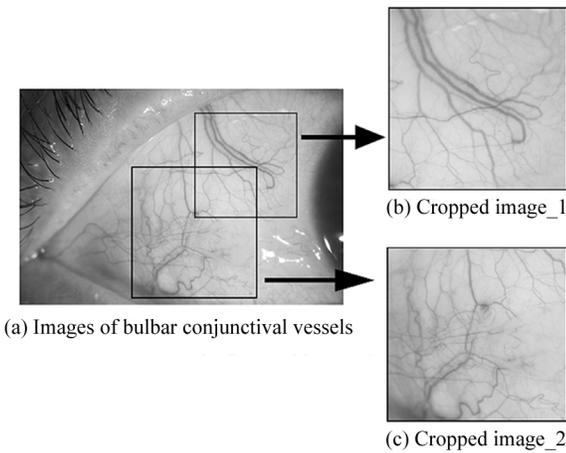


图 2 球结膜血管图像的裁剪方式

Fig. 2 Clipping method of bulbar conjunctiva vascular image

图 2 中(a)无法用一个框包含所有血管特征,若使用一个框进行裁剪可能会框到光斑,因此采取原则 2 分块裁剪的策略,既规避了光斑,又尽可能地裁剪出更多血管特征,此方法还能够在一定程度上达到扩充数据集的效果。

根据已有文献记载,球结膜血管形态与 DR 之间存在一定关联,因此血管特征是图像分类的重点,为了能够突出血管特征,采用限制对比度自适应直方图均衡(contrast limited adaptive histogram equalization,CLAHE)算法对裁剪得到的血管图像进行效果增强,突出血管在图像中的特征。CLAHE 算法是一种限制对比度自适应直方图均衡化方法,主要的作用就是增强图像的对比度并能有效地限制噪声放大,使图像的对比度增强。CLAHE 算法主要应用于医学图像领域,对于图像中的血管有明显的增强效果,使球结膜血管更加明显,增强效果如图 3 所示。

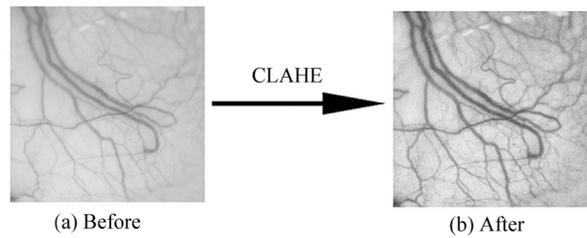


图 3 CLAHE 算法的增强效果

Fig. 3 Enhancement effect of CLAHE algorithm

### 1.3 数据增强

数据集分布为 DR 图像 286 张和非 DR 图像 40 张,经过裁剪后,根据裁剪原则 2 对数据集进行了一定程度上的扩充,得到 DR 图像 325 张和非 DR 图像 52 张,但依然存在两个类别数据集在数量上差距较大的问题,数据集数量较少或不同类别数据之间数量相差较大都会对网络的分类性能产生影响。对于 CNN 而言,随着网络层数的增加,学习的参数也随之增加,数据量少会更容易使网络模型过度拟合训练数据,导致网络没有很好的泛化能力,测试数据时准确率很低。针对数据集数量较少、类别间数据数量差距较大等问题,本实验采取数据增强的方式扩充数据集。

数据增强的手段包括:旋转、翻转、缩放、裁剪、噪声、色彩抖动、改变色彩饱和度以及改变对比度等,根据本实验对象的特点,每张球结膜图像中都具有重要的血管特征信息,缩放会导致特征丢失,因此不适合进行随机缩放操作。人工拍摄图像时每一张图像的光照、亮度、手法存在细微误差,可能会使拍摄的图像亮度、对比度有所不同,因此加入对图像亮度、对比度的数据增强,会增强网络模型的泛化能力。综上,本文对球结膜图像采取高斯滤波、锐化、改变对比度、改变亮度、翻转 5 种数据增强方式,设置一个随机参数,对每一张图像随机进行增强操作,数据增强效果如图 4 所示。

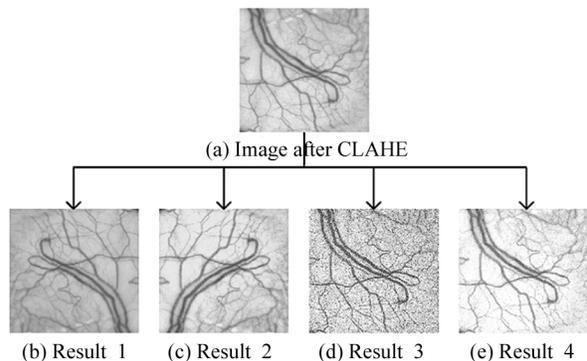


图 4 数据增强

Fig. 4 Data to enhance

由于健康图像与患病图像之间存在的数量差距较大,因此对健康数据进行了 20 倍增强,对患病数据进行了 4 倍增强,最终得到 1 300 张 DR 图像和 1 040 张非 DR 图像,既扩充了数据集数量,又降低了健康图像与患病图像的数量差距,最终得到数据集结构见表 2。

表 2 处理后的球结膜血管图像数据集

Tab. 2 Data set of processed bulbar conjunctival vascular images

Image category	Label	Number of images	Percentage/%
Health	0	1 040	44.4
DR	1	1 300	55.6

## 2 构建 CTCNet 网络模型

CNN 中卷积核的计算方式使 CNN 对于整个图像全局特征的关注的有限的,以往为了提高网络模型的特征提取能力,会通过加深网络模型层数,降低特征图尺度来扩大感受野,从而提升模型特征提取的全局性,但随着网络层数的增加,训练时容易出现梯度消失和梯度爆炸,影响模型的收敛。基于本文研究的球结膜微血管,分析球结膜血管图像特征,血管之间盘根错杂,血管特征不仅有弯曲度、粗细等局部特征,还有长度、血管密度此类全局特征需要考虑,因此球结膜血管的分类训练模型引入全局性特征提取会更加合理。随着 Transformer 在 NLP 领域取得许多优秀的成果,人们意识到其先进性,渐渐尝试将此强大机制应用于计算机视觉, Vision Transformer(ViT)<sup>[14]</sup>网络模型应运而生,其 Self-Attention(SA)机制能够很好地提取图像的全局特征。

基于以上问题,本文将 Transformer 与 CNN 相结合,在 ResNet18<sup>[15]</sup>的基础上加入 Transformer,构建 CTCNet。在网络模型浅层加入多头注意力计算,使模型的全局信息能够早期聚合,将全局特征从较低的层次传播到较深层次,来达到增强捕获特征的目的,再利用卷积网络提取特征。

### 2.1 SA

Transformer 的实现依赖于 SA 机制,与传统的 CNN 不同,它的实现与卷积完全不同,是一种通过矩阵点积运算来实现的学习机制,如图 5 所示。

SA 计算需要 3 个可训练的矩阵: $W_Q$ 、 $W_K$ 、 $W_V$ ,分别与输入的 token 进行点积运算得到查询向量( $Q$ )、键向量( $K$ )、值向量( $V$ ), $Q$ 、 $K$  向量相乘后经过 softmax 函数得出注意力得分,再与  $V$  相乘得到经过注意力计算后的加权结果,计算式见式(1):

$$Z' = Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, (1)$$

式中,  $d_k$  为  $K$  向量的维度,除以  $\sqrt{d_k}$  以保持梯度稳定。

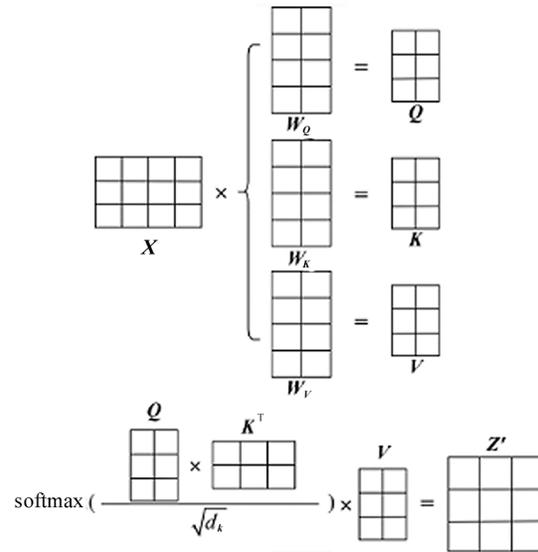


图 5 SA 计算概念图

Fig. 5 Calculation concept diagram of SA

### 2.2 Multi-head self-attention (MSA)

MSA 涉及到多个 SA 的并行操作,称为多个“头”,如图 6 所示。MSA 将  $h$  个 SA 的计算结果拼接起来,再通过一个可学习的映射  $W^0$  矩阵映射回输入维度,保持输入  $X$  与输出  $Z$  维度一致。计算式见式(2):

$$MSA = \text{concat}(SA_1, SA_2, \dots, SA_h)W^0. (2)$$

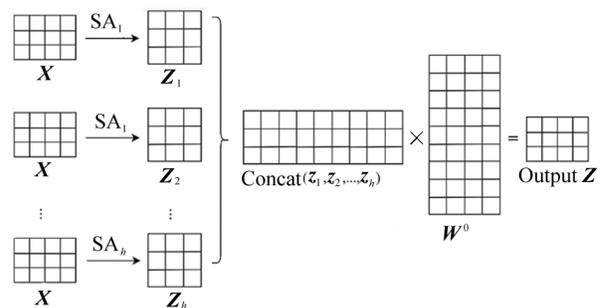


图 6 MSA 计算概念图

Fig. 6 Calculation concept diagram of MSA

ViT 将  $(H, W, C)$  大小的图像 reshape 为  $[N, (P^2 \times C)]$  大小的二维矩阵来作为输入,其中  $(P, P)$  为图像块的大小,  $N$  为图像块的数量  $(N = H \times W /$

$P^2$ )。ViT 将转换的二维矩阵归一化后进行 MSA 操作,最终得到与输入向量相同维度大小的向量,作为 MSA 的输出,再输入全连接层得到预测结果,ViT 的 Transformer 结构如图 7 所示。

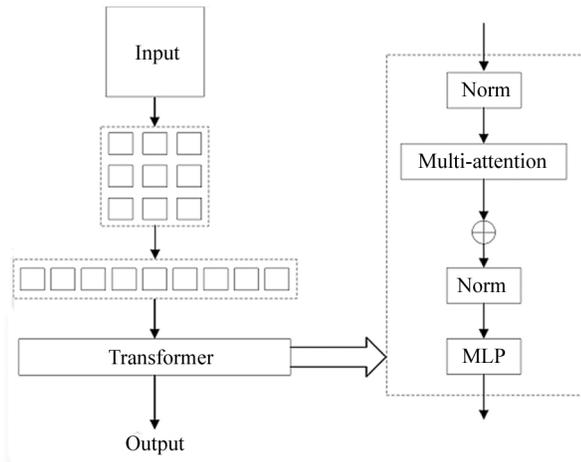


图 7 ViT 网络模型结构  
Fig. 7 ViT network model structure

### 2.3 构建 CTCNet

CTCNet 结合 Transformer 与 CNN 的优势,首先通过卷积得到特征图,提取图像特征,即处理数据集得到  $256 \times 256$  大小的图像输入到模型中,对输入图像进行  $3 \times 3$  卷积以及  $4 \times 4$  的平均池化,得到  $64 \times 64 \times 64$  的特征图,提取特征的同时降低了图像大小,也降低了计算 TransformerBlock 时的计算量。TransformerBlock 首先将每个通道的特征图分为  $8 \times 8$  大小的图像块并转化为一维向量,得到  $64 \times 4096$  大小的向量,加入位置编码后最终得到  $65 \times 4096$  大小的向量,输入 TransformerBlock 进行一个“头”个数为 16,深度为 6 的 MSA 操作,整合特征图像的全局特征,提升特征之间的联系,去除 ViT 末尾的 MLP 重新 reshape 回  $64 \times 64 \times 64$  维度;后续再进行 3 个 CNNBlock 操作,通道数分别为 128、256、512,扩充通道,进一步提取特征。网络结尾对特征图进行一个  $7 \times 7$  的平均池化和全连接神经网络,最终取预测概率较大的索引作为预测的分类标签,CTCNet 网络结构如图 8 所示。

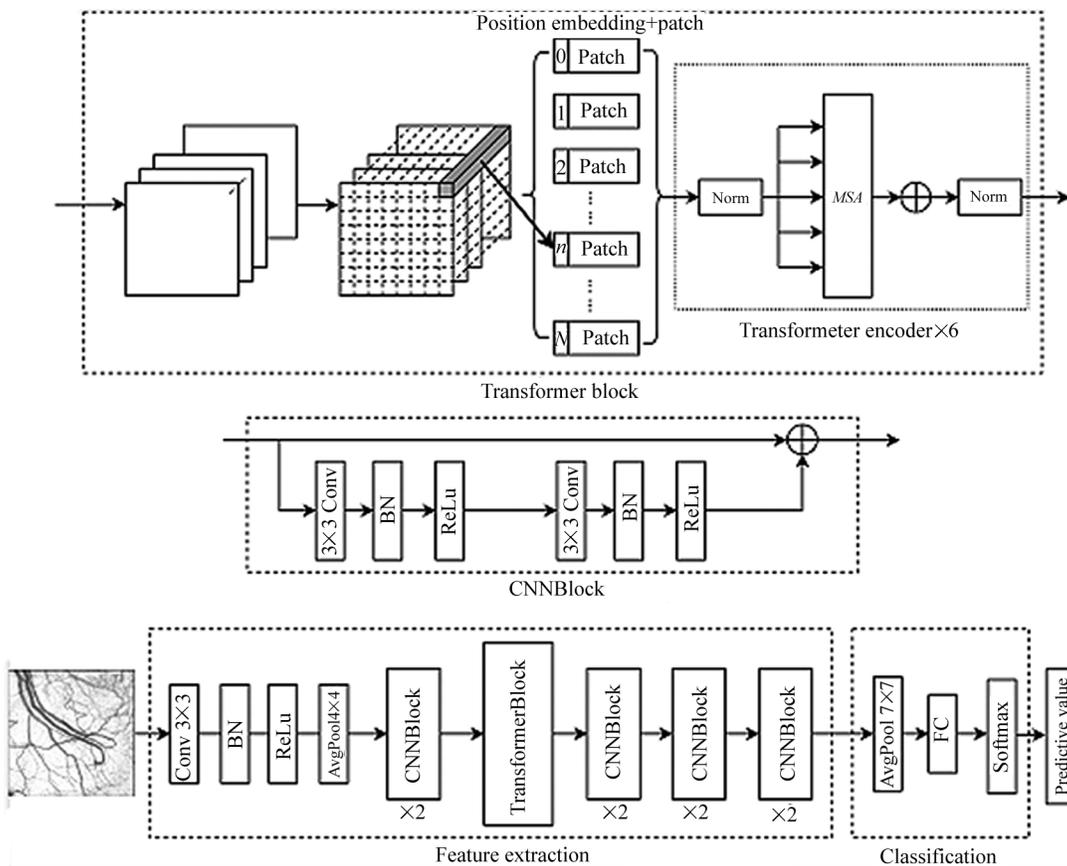


图 8 CTCNet 网络模型结构  
Fig. 8 CTCNet network model structure

### 3 实验结果评价与分析

#### 3.1 实验评价

参考相关领域科研文献,本文通过分类准确率 (*Accuracy*)、灵敏度 (*Sensitivity*)、特异性 (*Specificity*)和精确度 (*Precision*)对模型的性能进行评估。4 个评价指标的计算见式(2)–(5):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (3)$$

$$Sensitivity = \frac{TP}{TP + FN}, \quad (4)$$

$$Specificity = \frac{TN}{TN + FP}, \quad (5)$$

$$Precision = \frac{TP}{TP + FP}. \quad (6)$$

基于二分类模型,将样本识别类型分为真阳性 (*TP*)、假阳性 (*FP*)、真阴性 (*TN*)和假阴性 (*FN*),当 DR 患者的球结膜图像输入模型,被识别为 DR 则为 *TP*,被错误识别为健康则为 *FP*;当健康的球结膜图像输入模型,被识别为健康则为 *TN*,被错误识别为 *FN* 则为假阳性,如图 9 所示。

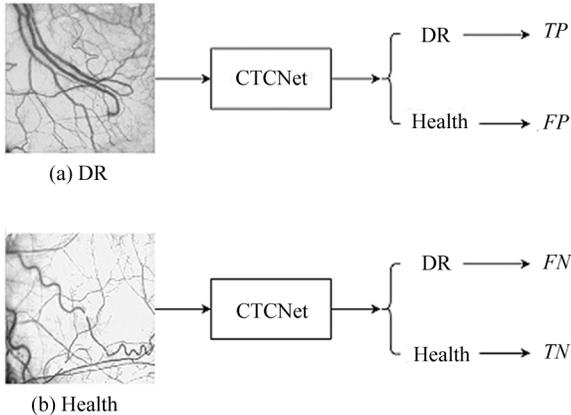


图 9 球结膜图像的识别类型

Fig. 9 Recognition types of bulbar conjunctival images

#### 3.2 实验结果分析

本文实验在 Ubuntu16.04,64 位系统下,硬件方面采用 TITAN Xp 型号显卡(11 GB 显存),基于深度学习框架 Pytorch 完成实验。数据集方面,将数据集预处理后,对实验数据进行划分,其中 80%作为训练集,用于网络训练权重参数,20%作为测试集,对模型的泛化能力进行评估。将处理好的数据集输入 CTCNet 中进行训练测试,记录训练过程中训练集和测试集的准确率以及损失值,由图 10 可以看到准确

率能够很好地收敛,损失值也在降低。

为了对本文的 CTCNet 分类模型的性能进行评估,将本文实验与经典 CNN ResNet18、ResNet50 和结合 Transformer 的网络分类模型 ViT、TransMed 进行比较,使用相同处理后的球结膜血管图像训练集和测试集,依次进行分类训练,在相同的设备上训练 50 个 epoch,实验结果比较如表 3 所示。

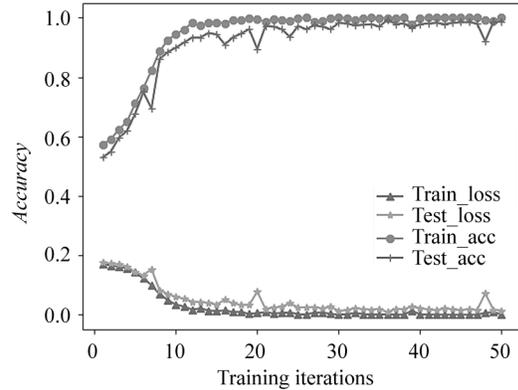


图 10 Acc-loss 变化曲线图

Fig. 10 Acc-loss change curve

表 3 实验结果

Tab. 3 Experimental result

Model	Accuracy /%	Sensitivity /%	Specificity /%	Precision /%
RestNet18	83.33	88.07	77.4	82.97
RestNet50	75.00	76.92	72.59	77.82%
ViT	86.54	88.07	84.61	87.73
TransMed	96.37	95.38	97.59	98.02
CTCNet	<b>97.44</b>	<b>97.69</b>	<b>97.11</b>	<b>97.69</b>

由表 3 可以看出:本文的 CTCNet 网络对球结膜图像进行 DR 分类的准确率为 97.44%,敏感度为 97.69%,特异性为 97.11%,精确度为 97.69,平均水平高于其他网络模型,验证了可以通过球结膜血管图像对 DR 进行分类,也验证了此模型在球结膜图像上识别是有效的。

### 4 结论

本文针对球结膜血管图像对 DR 进行分类,首先对数据集进行裁剪、图像增强、数据增强,结合 CNN 与 Transformer 的优势提出 CTCNet,改善传统 CNN 的局部局限性,使用 Transformer 增强模型对于图像的全局注意力,更好地提取球结膜血管图像的特征,提高分类的准确率。为了验证本文提出的模型性能,设置了几组对比实验,对比网络模型分别为 ResNet18、ResNet50、ViT、TransMed,实验结果

表明,本文的 CTCNet 分类效果优于其他分类模型。

### 参考文献:

- [1] XIE X Q. Related risk factors and prevention strategies for the development of diabetes retinopathy[J]. Chinese Journal of Metallurgical Industry Medicine, 2019, 36(2): 212-213.  
谢小青. 糖尿病视网膜病变发展的相关危险因素及防治策略[J]. 中国冶金工业医学杂志, 2019, 36(2): 212-213.
- [2] LIU S P, CHEN J. Illumination equalization and adaptive vascular enhancement algorithm for fluorescent retinal images[J]. Journal of Optoelectronics · Laser, 2011, 22(5): 793-797.  
刘尚平,陈骥. 荧光视网膜图像的照度均衡及自适应血管增强算法[J]. 光电子·激光, 2011, 22(5): 793-797.
- [3] LIANG L M, ZHOU L S, YU J, et al. Double U-shaped retinal segmentation algorithm based on multi-scale feature fusion [J]. Journal of Optoelectronics · Laser, 2022, 33(3): 272-282.  
梁礼明,周珑颂,余洁,等. 多尺度特征融合双U型视网膜分割算法[J]. 光电子·激光, 2022, 33(3): 272-282.
- [4] LIAN X F, LIU Z Y, ZHANG L, et al. A retinopathy image recognition method based on deep learning [J]. Computer Applications and Software, 2021, 38(1): 179-185.  
连先峰,刘志勇,张琳,等. 一种基于深度学习的视网膜病变图像识别方法[J]. 计算机应用与软件, 2021, 38(1): 179-185.
- [5] SARANYA P, PRABAKARAN S, KUMAR R, et al. Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning [J]. The Visual Computer, 2021, 38(3): 977-992.
- [6] WANG W, PU Y W. Classification method of hypertensive retinopathy based on regional feature fusion[J]. Computer Engineering and Application, 2022, 58(8): 230-236.  
王伟,浦一雯. 区域特征融合的高血压视网膜病变分类方法[J]. 计算机工程与应用, 2022, 58(8): 230-236.
- [7] ZHU K, CHEN X X, JI L G. Vascular recognition of bulbar conjunctiva image[J]. Chinese Journal of Medical Physics, 2010, 27(6): 2268-2270.  
朱坤,陈新兴,季立贵. 球结膜图像的血管识别[J]. 中国医学物理学杂志, 2010, 27(6): 2268-2270.
- [8] OWEN C G, NEWSOM R S, RUDNICKA A R, et al. Diabetes and the tortuosity of vessels of the bulbar conjunctiva[J]. Ophthalmology, 2008, 115(6): 27-32.
- [9] BANAEI T, POURREZA H, DOOSTI H, et al. Distribution of different sized ocular surface vessels in diabetics and normal individuals[J]. Journal of Ophthalmic & Vision Research, 2017, 12(4): 361-367.
- [10] IROSHAN K A, DE ZOYSA A D N, WARNAPURA C L, et al. Detection of diabetes by macrovascular tortuosity of superior bulbar conjunctiva[C]//Annual International Conference of the IEEE Engineering in Medicine and Biology Society, July 18-21, 2018, Honolulu, HI, USA. New York: IEEE, 2018: 1-4.
- [11] GUNAY M, GOCERI E, DANISMAN T, et al. Automated detection of adenoviral conjunctivitis disease from facial images using machine learning[C]//IEEE 14th International Conference on Machine Learning and Applications (ICMLA), December 09-11, 2015, Miami, FL, USA. New York: IEEE, 2015: 1204-1209.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is all you need[EB/OL]. (2017-12-06)[2022-08-28]. <https://arxiv.org/abs/1706.03762>.
- [13] DAI Y, GAO Y. TransMed: Transformers advance multi-modal medical image classification[EB/OL]. (2021-03-10)[2022-08-28]. <https://arxiv.org/abs/2103.05940>.
- [14] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale[EB/OL]. (2021-06-03)[2022-08-28]. <https://arxiv.org/abs/2010.11929>.
- [15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.

### 作者简介:

汪日伟 (1973—),男,工学博士,温州理工学院教授,硕士生导师,主要研究方向为计算视觉、人工智能、虚拟现实技术。