

基于超参数优化的TransCNN眼底图像分类算法

王小芳¹, 余柯欣², 王张怡², 王剑华², 王静³, 穆楠⁴

1. 吉利学院智能科技学院, 四川 成都 641423; 2. 电子科技大学成都学院计算机学院, 四川 成都 611731; 3. 成都医加医光学有限责任公司, 四川 成都 610031; 4. 四川师范大学计算机科学学院, 四川 成都 610066

【摘要】针对深度学习模型参数多、随机、训练时间长,眼底图像病变处交织重叠、数据集样本不均等问题,提出基于超参数优化的TransCNN(Deho-TransCNN)眼底图像分类算法。该算法以TransCNN网络模型为基础,利用差分进化算法分别对模型网络权重进行初始化寻优和对模型进行超参数组合实现参数自适应优化;最后利用MEB-KSVM对眼底病变图像进行多分类。实验结果表明,改进算法的准确率、敏感性、特异性以及AUC值最优,分别为0.947、0.926、0.937、0.945,与文中9种传统算法比较,分别平均提升5.6%、6.4%、5.1%、7.9%;改进算法检测时间最低,与最佳算法改进CNN相比,平均检测时间降低158.3%。改进算法在一定程度上提升图像多分类效果,降低图像检测时间,对图像多分类处理有一定泛化能力。

【关键词】眼底图像;差分进化算法;TransCNN;超参数优化;MEB-KSVM

【中图分类号】R318;R744.1

【文献标志码】A

【文章编号】1005-202X(2023)06-0672-11

Eye fundus image classification using hyperparameter optimization based TransCNN

WANG Xiaofang¹, YU Kexin², WANG Zhangyi², WANG Jianhua², WANG Jing³, MU Nan⁴

1. School of Intelligence Technology, Geely University of China, Chengdu 641423, China; 2. School of Computing, Chengdu College of University of Electronic Science and Technology of China, Chengdu 611731, China; 3. Chengdu Medical Plus Medical Optics Co., Ltd, Chengdu 610031, China; 4. College of Computer Science, Sichuan Norman University, Chengdu 610066, China

Abstract: A novel algorithm (hyperparameter optimization based TransCNN, Deho-TransCNN) for fundus image classification is proposed to solve the problems of the multiple parameters, randomness and long training time in deep learning model, overlapping lesions in fundus images, and uneven samples in data sets. Based on TransCNN model, the algorithm uses differential evolution algorithm to initialize the weight of model network, combines the hyperparameters to achieve adaptive optimization, and complete the fundus image multi-classification using MEB-KSVM. The experimental results show that the proposed algorithm has an accuracy, sensitivity, specificity and AUC value of 0.947, 0.926, 0.937 and 0.945 for image multi-classification, with average improvements of 5.6%, 6.4%, 5.1% and 7.9% as compared with the other 9 traditional algorithms. The average detection time of the proposed algorithm is 158.3% lower than that of the optimal CNN. The improved algorithm can enhance the performance of image multi-classification, reduce the time required for image detection, and has certain generalization ability for image multi-classification.

Keywords: fundus image; differential evolution algorithm; TransCNN; hyperparameter optimization; MEB-KSVM

前言

视网膜病变是目前眼底疾病的主要病因之一,糖尿病视网膜病变(DR)作为目前增长最快的致盲疾病,致残致死率仅次于心脏血管疾病和癌症^[1]。如何

有效、快速地对眼底图像进行诊断,降低疾病带来的干扰,成为医学眼科的重要方向^[2]。眼科患者急速增长,人工诊断已无法满足需求,利用智能设备辅助人工检测成为智慧医疗发展的核心。检测技术主要包含机器学习方法和深度学习方法。任福龙等^[3]使用融合多列学习与极限学习机方法构建多核图多列学习模型,实现糖尿病图像二分类。张晨曦等^[4]以小波分解图像金字塔模型为基础,通过局部二值模式与灰度梯度共生矩阵算法分别提取眼底图像病灶区纹理特征和细节特征,将提取结果融合得到全局特征进行降维,并使用最小优化模型进行眼底图像分类。

【收稿日期】2022-10-18

【基金项目】国家自然科学基金(62006165);成都市科技局项目(2018-YFYF-00191-SN);四川省民办教育协会项目(MBXH21YB119);吉利学院科研项目(2022xzk004)

【作者简介】王小芳,硕士,讲师,主要研究方向:计算机视觉, E-mail: 939549393@qq.com

Wan等^[5]利用融合迁移学习与超参数优化的卷积神经网络(CNN)模型对眼底图像进行分类。Sunil等^[6]提出基于CNN模型实现糖尿病视网膜眼底图像二分类。Zhou等^[7]采用多单元多任务卷积神经网络(M2CNN)对眼底图像进行分类,已解决梯度消失/爆炸问题,进而提升图像分类效果。Zhang等^[8]将DenseNet121网络与ResNet50网络融合实现对糖尿病病变严重程度的分类。李琼等^[9]提出在AlexNet网络基础上,在其卷积层和全连接层前引入归一化层构建BNnet网络实现特征提取,同时对网络使用迁移学习策略进行预训练,并结合深度分类器实现眼底视网膜图像分类。李轩屹等^[10]提出基于多尺度ResNet网络模型,该模型通过不同分辨率输入、全连接层处理深浅层特征信息以及迁移学习实现参数初始化处理以实现DR眼底图像严重程度分类。孙雨琛等^[11]使用残差思想的DCNN网络模型实现视网膜眼底图像特征自动提取,以实现眼底图像分类。连先峰等^[12]利用多特征融合的改进VGG16提取眼底图像局部特征实现图像分类。许祥丛等^[13]提出融合SEC和Unet的模型实现白内障眼底图像分类,该模型增强局部细节信息同时有效避免Unet网络局部最优问题。以上算法虽能较好地实现对眼底图像分类,但网络结构较为复杂,模型训练时间长,初始参数设置具有随机性,对眼底图像病变处交织重叠处理效果不佳,不适用于不平衡样本分类场景。

为解决上述问题,提出基于超参数优化的改进

TransCNN眼底图像分类算法模型,该模型利用差分进化算法对TransCNN初始参数进行寻优处理,并将处理后的参数送入网络中进行模型训练提取特征信息,在TransCNN模型训练过程中,对模型内部进行超参数优化实现系统训练的自适应迭代,寻找超参数组合最优解,最后使用MEB-KSVM对特征提取结果进行多分类。

1 TransCNN 基础

苏黎世联邦于2021年提出一种新型骨干网络,该网络继承transformer和CNN优点,通过分层级联方式降低计算/空间复杂度,提升图像识别泛化能力^[14]。TransCNN由卷积层、层化层、全连接层构成,网络结构依次通过4次下采样、4层卷积、4层池化以及全局平均池化和全连接实现特征结果输出,如图1所示。由图1可知,TransCNN由连续两个3×3卷积对输入图像降采样至1/4尺度,然后利用金字塔尺度将结果划分为1/4、1/8、1/16、1/32大小,并将H-MASH模块和卷积块融入4个尺度阶段。其中,卷积模块为Inverted Residual Bottleneck(IRB),卷积为深度可分离卷积,IRB架构如图2a所示。激活函数采用非线性激活的SiLU,1/4、1/8、1/16尺度阶段卷积结束后采用二分支降采样块(TDB)处理,以保存更多上下文信息,结构如图2b所示。最后经过全局平均池化与全连接实现图像分类。

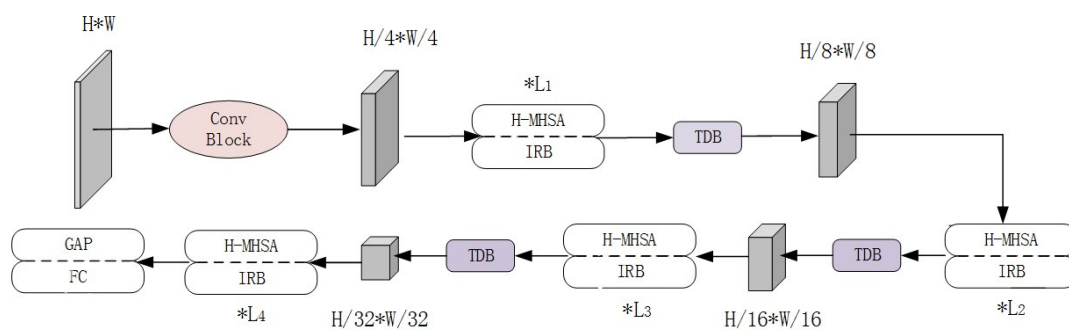


图1 TransCNN网络基本结构

Figure 1 Basic structure of TransCNN

TransCNN模型使用分层级联多头自注意力模块(H-MHSA)替代多头自注意模块(MHSA)建模,实现特征增强,有效解决Transformer中复杂度高产生的效率低问题,该模型可嵌入任何CNN网络结构,实现反向传播训练。MHSA^[14]通过对注意力机制多次执行,生成对应注意力值。每次执行参数不同,各注意力机制间互不干扰,独立存在^[15]。取每个注意力层

的注意力值计算均值得到权重,如式(1)所示:

$$\text{MHSA}(Z) = \text{concat}_{h \in [N_h]} [\text{Self-Attention}_h(Z)] W_{\text{out}} + b_{\text{out}} \quad (1)$$

其中, $W_{\text{out}} \in R^{N_h E_h \times E_{\text{out}}}$, $b_{\text{out}} \in R^{E_{\text{out}}}$ 为偏置(bias), N_h 为注意力层数量,并行运行每个注意力层中不同参数矩阵提取不同特征信息。

H-MHSA^[14]利用MHSA计算,使用局部注意力与分层注意力融合降低空间复杂度,以优化视觉变

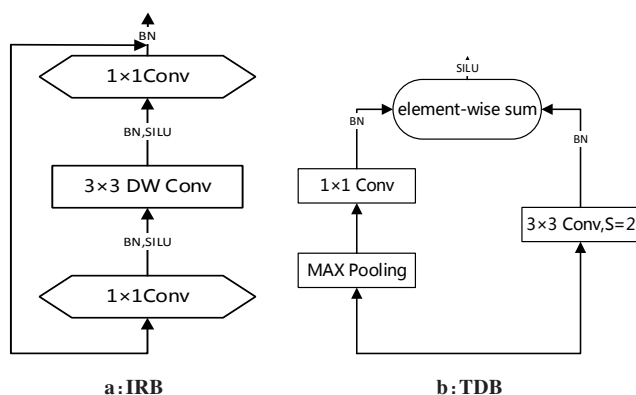


图2 TransCNN中IRB和TDB基本结构
Figure 2 Basic structures of IRB and TDB in TransCNN

换器低效率缺陷。通过将H-MHSA等价于传统MHSA模拟全局,建立全局特征依赖关系,实现分层建模。将标记分步骤分层次处理,不断增加区域网络逐步计算自我注意力,进而保证在不降低性能情况下降低模型复杂度。H-MHSA结构如图3所示。

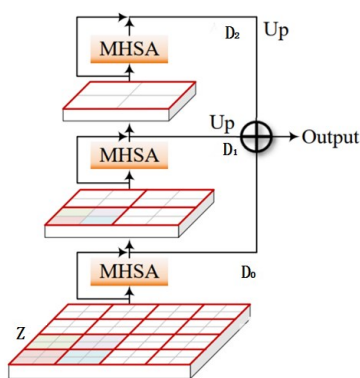


图3 H-MHSA结构
Figure 3 H-MHSA structure

H-MHSA用于特征增强,通过级联方式分层计算,将图片分成独立窗口以patch计算局部注意力,如式(2)所示:

$$J = H_0 \times W_0 \quad (2)$$

其中, J 是标记数量, H_0 和 W_0 分别是DR特征图的高度和宽度。将特征图切分,划分为 $K \times K$ 小网格,如式(3)所示:

$$\begin{aligned} Z \in R^{H_0 \times W_0 \times C} &\rightarrow Z' \in R^{\left(\frac{H_0}{K_0} \times K_0\right) \times \left(\frac{W_0}{K_0} \times K_0\right) \times C} \\ &\rightarrow Z' \in R^{\left(\frac{H_0}{K_0} \times \frac{W_0}{K_0}\right) \times (K_0 \times K_0) \times C} \end{aligned} \quad (3)$$

其中, Z 是特征映射, C 是每个标记的特征维数,满足 $Z \in R^{H_0 \times W_0 \times C}$,将处理所得局部注意力逐步拼接并添加残差连接,使其在反向传播过程中梯度连乘,解决梯度消失问题,如式(4)所示:

$$D_0 = Z + D_0 \quad (4)$$

其中, D_0 为局部注意力机制处理结果,满足 $D_0 \in R^{\left(\frac{H_0}{K_0} \times \frac{W_0}{K_0}\right) \times (K_0 \times K_0) \times C}$, $K_0 \times K_0$ 为第1步处理小网格大小。当 $q^* = Z'W^q$, $k^* = Z'W^k$, $v^* = Z'W^v$ [16],特征图 $Z \rightarrow D_0$ 简化网络。 $W^q \in R^{C \times C}$ 为查询矩阵, $W^k \in R^{C \times C}$ 为关键词矩阵, $W^v \in R^{C \times C}$ 为值矩阵,是线性变换权重矩阵,实现输入特征图的降维。 D_0 特征图简化处理如式(5)所示:

$$\begin{aligned} D_0 \in R^{\left(\frac{H_0}{K_0} \times \frac{W_0}{K_0}\right) \times (K_0 \times K_0) \times C} &\rightarrow D_0 \in R^{\left(\frac{H_0}{K_0} \times K_0\right) \times \left(\frac{W_0}{K_0} \times K_0\right) \times C} \\ &\rightarrow D_0 \in R^{H_0 \times W_0 \times C} \end{aligned} \quad (5)$$

若求 s 步($s>0$)结果,将第 $s-1$ 步结果与网格 $K_{s-1} \times K_{s-1}$ 视为一个标记,对注意力特征进行降采样,如式(6)所示:

$$D'_{s-1} = \text{MaxPool}_{K_{s-1}}(D_{s-1}) + \text{AvePool}_{K_{s-1}}(D_{s-1}) \quad (6)$$

其中, D_{s-1} 为注意力特征, $\text{MaxPool}_{K_{s-1}}(\cdot)$ 和 $\text{AvePool}_{K_{s-1}}(\cdot)$ 分别为最大池化和平均池化,实现内核大小和步长为 K_s 、样本 D_s 降为 K_{s-1} 次。迭代直至 $H_s \times W_s$ 不能split为止,最终得到各层注意力特征图。对各种注意力图进行上采样,实现不同尺寸结果处理,并将处理结果相加,得到多层注意力融合特征结果,如式(7)所示:

$$\text{H-MHSA}(Z) = D_0 + \dots + \text{Upsample}(D_E)W^p + Z \quad (7)$$

其中, $\text{Upsample}(\cdot)$ 是DR注意力特征上采样原始图, E 是最大步数, $W^p \in R^{C \times C}$ 是特征映射权重矩阵。研究使用TransCNN作为网络基础,优化网络参数进行模型训练实现DR图像特征提取。

2 超参数优化的TransCNN模型

针对DR分类效果不佳、网络参数初始化随机引起的训练速度慢,眼底图像病变处交织重叠问题,研究使用差分种群优化的TransCNN模型实现图像多分类,其处理整体分为图像预处理、超参数权重初始化、改进TransCNN图像特征提取与分类,算法流程如图4所示。由图4可知,研究核心思路为:(1)对改进TransCNN模型全连接层参数使用差分进化算法进行权值初始化,以网络初始权重寻优;(2)利用差分进化算法监听网络向前传播损失值,使TransCNN训练中网络参数实现自适应迭代,寻找最优超参数组合,优化网络结构;(3)采用MEB-KSVM代替SoftMax进行非均衡DR数据集多分类,以提升分类效果,提高运算效率。

2.1 数据预处理

图像预处理可有效降低图像噪声、避免尺寸差异无法输入以及避免因数据量少影响训练效果。预

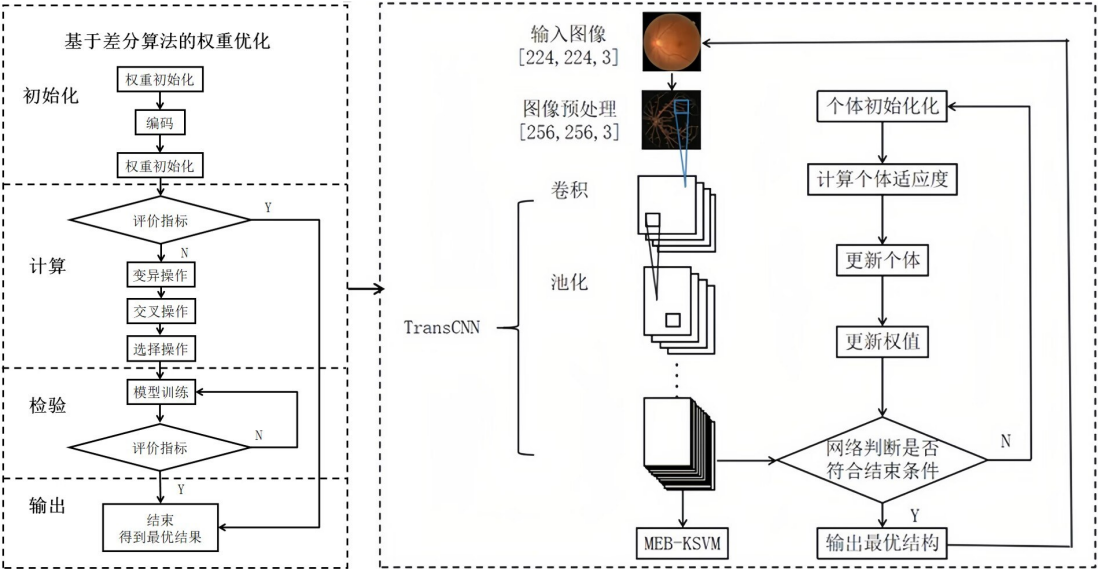


图4 基于差分进化的TransCNN流程图

Figure 4 TransCNN based on differential evolution

处理包含尺寸归一化、图像增广和图像优化3步。其中尺寸归一化根据网络模型特点将图像处理为分辨率256×256结果^[6]。图像增广用于扩充图像数据集,研究使用图像转置、镜像处理增广,将数据集旋转0°、180°以及镜像处理得到扩容数据集,效果如图5

所示。图像增强处理可一定程度降低噪声等带来的干扰,对数据集采用融合高斯滤波的Canny算法进行降噪处理,先使用高斯滤波进行平滑处理,再使用Canny边缘算子强化特征信息。经处理后,可较好地提取眼底图像特征,效果对比如图6所示。

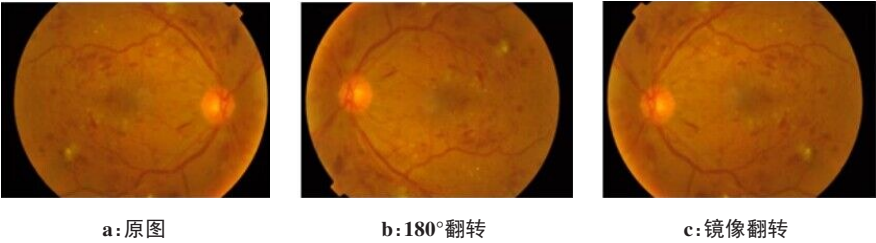


图5 数据增广处理后的视网膜图像

Figure 5 Retinal images after data augmentation

2.2 超参数权重初始化

梯度优化算法在反向处理中,易出现因梯度值相等产生的迭代处理过程模型训练收敛速度过大、权重空间局部最优问题。使用差分进化算法对TransCNN模型的全连接层权重参数进行初始化,以实现全局寻优。差分进化算法是群体自适应全局优化的高效启发式并行搜索策略^[17],算法从多点开始对全局最优值进行搜索,即将种群中任意两个个体向量差与第三个个体求和以产生新个体,再将新个体与当代种群中对应个体进行比较,保留最优个体的不断进化^[18]。

算法整体处理分为初始化、计算、检验和输出4步。

步骤(1):权重初始化、图像编码和种群初始化。网络权重初始化为正态初始,以解决网络因随机生成权重引起无效的模型训练。网络权重初始化需服从正态分布数学期望,概率密度如式(8)所示:

$$f(w) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(w - \mu)^2}{2\sigma^2}\right) \tag{8}$$

其中, w 表示网络权重, μ 表示网络权重参数期望, σ 表示方差,当 $\mu = 0, \sigma^2 = 0.0013$ 时,其初始化参数结果最佳。图像编码用于处理分类器属性数据偏差,使用 one-hot 编码实现输入图像的处理,将图像采用二进制向量进行分类,分别得到样本 $\epsilon(1, 2, 3, 4, 5)$ 对应 DR 严重程度分级,one-hot 编码分别为 $\{[1, 0, 0, 0, 0], [0, 1, 0, 0, 0], [0, 0, 1, 0, 0], [0, 0, 0, 1, 0], [0, 0, 0, 0, 1]\}$ 。

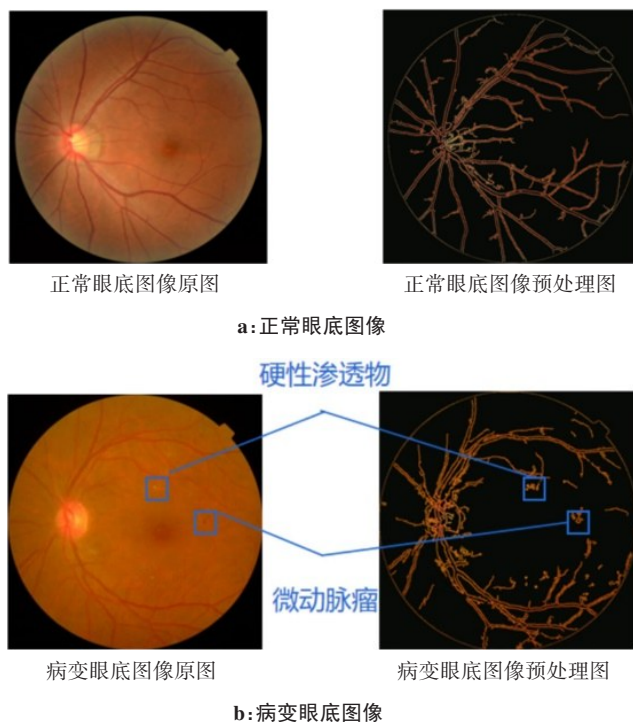


图6 数据预处理对比

Figure 6 Data preprocessing comparison

对编码后数据进行种群初始化,将标签编码后图像输入模型进行训练,记录每次TransCNN向前传播权重值 w_a ,处理如下所示:

$$b = e_1 \times e_1 + e_2 + e_2 \times e_3 + e_3 \quad (9)$$

$$P_a(g) = \{w_1(g), w_2(g), \dots, w_b(g)\} \quad (10)$$

其中, e_1 为全连接层输入层神经元个数, e_2 为隐藏层神经元个数, e_3 为输出层神经元个数,用于计算种群内个体规模, w_b 为种群中单个个体, a 是种群编号,种群规模理性区间一般为 $3b \sim 8b$ ^[19]最佳,经实验文章种群规模为 $5b$ 。对差分权重优化而言,一组完整神经网络权重即为差分的一个种群。

步骤(2):种群计算,包含计算评价、变异、交叉和选择。计算评价用于评价种群优劣,分为计算适应度和计算阈值两个操作。其中,计算适应度记录每次Deho-TransCNN训练后的预测值,利用预测值得到每次模型训练后损失值 $suit$,公式如式(11)所示:

$$suit = \sum_{a,b=1}^{\beta} (-\sum_{a=1}^b \sum_{b=1}^{\sigma} T_{a,b} \ln pre_{a,b}) \quad (11)$$

其中, β 是待检测DR图像程度类数, σ 是每组图像标签的维数, pre 是网络预测值, T 是标签真实值。将计算所得损失值 l 作为适应度,评价DR分类中每个种群优劣程度以及判断阈值。阈值判断以本次适应度与上次迭代适应度差的绝对值进行评判。计算如式(12)、式(13)所示:

$$abs(suit_{a-1} - suit_a) > o_1 \text{ and } abs(suit_{a-1} - suit_a) > o_2 \quad (12)$$

$$o_1 = 280 + 4 \times l_{a-1}, o_2 = 280 + 7 \times l_a \quad (13)$$

其中, o_1 与 o_2 代表阈值, $suit_{a-1}$ 代表第 $a-1$ 次迭代适应度, $suit_a$ 代表第 a 次迭代适应度,若满足绝对值差小于阈值,寻优结束,本次迭代权重为网络初始化权重,反之执行变异操作,以突破局部范围问题,寻找全局最优权重。

差分变异操作在现有权重组合中引入新的权重空间以突破当前搜索限制,实现更优解求解,计算如式(14)所示:

$$VAR_a(g) = P_{n1}(g) + Fa(P_{n2}(g) - P_{n3}(g)) \quad (14)$$

其中, $VAR_a(g)$ 为变异后种群,取值 $n0, n1, n2, n3 \in \{1, 2, \dots, 40\}$ 且为互不相同整数, Fa 是交叉因子,若过小将降低网络收敛速度产生局部最小值较小,过大则会造成种群难收敛,经实验 $Fa=0.2$ 最佳。

为将优良种群传递给下一代,对变异种群进行交叉处理,原理为将两组实用度高的父母种群权重组合的个体进行交叉处理,形成子权重组合。其中,父母种群中群众组合采用轮盘赌算法进行采样,计算如式(15)所示:

$$pr_a = \frac{suit_a}{suit_{all}} \quad (15)$$

其中, $suit_a$ 为种群 $P_{n1}(g)$ 的适应度, $suit_{all}$ 为种群适应度的和, pr_a 为当前超参数组合被选择的概率。对采样后结果进行交叉操作,如式(16)所示:

$$CRO_a(g) = \begin{cases} VAR_a(g), & \text{if } rand[0,1] \leq pr_a \text{ or } b = b_{rand} \\ P_a(g), & \text{if otherwise} \end{cases} \quad (16)$$

其中, $CRO(g)$ 为交叉后的种群, $rand[0,1]$ 为属于0到1的随机数。为确保 $VAR_a(g)$ 的每个种群至少存在一个个体遗传给下一代,则限定第一个交叉操作的基因是 $VAR_a(g)$ 中第 b_{rand} 位基因作为交叉后 $CRO_a(g)$ 的第 b_{rand} 位等位基因。后续交叉操作通过 pr 选取等位基因。

为保留优良种群,以提高种群适应度,将变异和交叉操作后新种群与原种群进行比较,以择优,计算如式(17)所示:

$$P_a(g+1) = \begin{cases} P_a(g), & \text{if } suit(P_a(x)) < suit(CRO_a(x)) \\ CRO_a(g), & \text{otherwise} \end{cases} \quad (17)$$

为判断差分计算是否满足寻优条件,对计算结果进行差分终止检验。将适应度最高的种群迭代至Deho-TransCNN模型进行训练,根据所计算的损失值判断本次迭代的损失值与上一次迭代结果的绝对值差是否小于阈值,若小于则寻优结束,输出最优权重结果,反之,则重复循环过程,直到满足条件。经过差分进化算法处理后,得到Deho-TransCNN全局最优的全连接参数初始化信息。

2.3 Deho-TransCNN 模型训练

为优化训练时间、降低模型计算复杂度,研究采用 Deho-TransCNN 进行 DR 眼底图像训练。Deho-TransCNN 以 TransCNN 为基础,利用差分算法对 TransCNN 中网络结构超参数进行组合优化,提取网络特征,并构建 MEB-KSVM 分类器替代 TransCNN 中的 SoftMax 函数实现不平衡数据集的多分类,完成 DR 图像的病变严重程度分类。

2.3.1 Deho-TransCNN 网络架构 Deho-TransCNN 以 TransCNN-small 为基础网络架构,网络配置如表 1 所示。将预处理后图像送入 TransCNN-small,模型以两个顺序的 vanilla 3×3、步长为 2 卷积将图像下采样至 1/4,并利用多尺度金字塔池化将网络划分为 1/4、1/8、1/16、1/32,每个尺度利用 H-MHSA 块与卷积块交替叠加使用,然后通过局部最大池化计算每个块最大值,实现由多到少的池化处理压缩特征图,扩大原始图像感受野范围。在金字塔池化处理的前 3 个阶段中使用 TDB 解决图像降采样中的失真问题。TDB 中一个分支用于控制步进大小,提取特征,由 vanilla 3×3 卷积构成;另一个分支用于特征选择与降维,由一个最大池化层和一个 1×1 卷积构成。降采样中两个分支以按元素求和方式融合,实现空间信息融合与下采样处理,以保留更多上下文信息,有效保证图像质量、保留特征信息。

表 1 TransCNN-small 网络配置
Table1 TransCNN-small network configuration

阶段	输入尺寸	算子	TansCNN-small
1	256×256	3×3 卷积	C=16, S=2 C=64, S=2
2	64×64	H-MHSA IRB	$\begin{bmatrix} C = 64 \\ K = 5 \\ E = 4 \end{bmatrix} \times 2$
3	32×32	H-MHSA IRB	$\begin{bmatrix} C = 128 \\ K = 3 \\ E = 4 \end{bmatrix} \times 2$
4	16×16	H-MHSA IRB	$\begin{bmatrix} C = 256 \\ K = 5 \\ E = 6 \end{bmatrix} \times 2$
5	8×8	H-MHSA IRB	$\begin{bmatrix} C = 512 \\ K = 3 \\ E = 5 \end{bmatrix} \times 2$
	1×1	-	全局平均池化, 1000-d 全连接层, MEB-KSVM

模型中,每阶段卷积采用 IRB^[16]处理。IRB 搜索所有合理计算顺序并选择最小化的张量,构建有向无环超图 $M(\Gamma)$ ^[20],如下所示:

$$M(\Gamma) = \min_{\pi \in \sum(\Gamma)} \max_{i_1 \in 1 \dots n} \left[\sum_{A \in R(i_1, \pi, \Gamma)} |L| + \text{size}(\pi_{i_1}) \right] \quad (18)$$

其中, $R(i_1, \pi, E)$ 是节点连接的中间张量表, $|L|$ 是张量大小, $\text{size}(\pi_{i_1})$ 是内部所需内存总量, i_1 为操作次数。

IRB 使用 3×3 深度可分离卷积^[21],由深度卷积和逐点卷积构成。DW 将输入特征图按 3 个通道进行分组并对输入层的每个通道独立进行卷积运算;PW 将经过卷积后结果的所有通道串联得到特征图,解决不同通道在相同空间位置上特征信息无法有效利用问题。其计算成本 c^* 如式 (19) 所示:

$$c^* = A_K \times A_K \times U \times H_{out} \times W_{out} + U \times V \times H_{out} \times W_{out} \quad (19)$$

其中, $H_{in} \times W_{in} \times U$ 为输入特征图大小, U 为输入通道, $H_{out} \times W_{out} \times V$ 为输出特征图大小, V 为输出通道, $A_K \times A_K$ 为卷积核大小。

为避免特征提取过程中空间信息破坏,利用全局平均池化层进行通道整合。全局平均池化层将输出的每个通道特征图中所有像素计算平均值,得到通道图,如式 (20) 所示:

$$o = \frac{\sum_{i_2=1}^{n_0} o_{i_2}}{n_0} \quad (20)$$

其中, o 为处理后的 DR 特征图像素值, $\sum_{i_2=1}^{n_0} o_{i_2}$ 为 DR 特征图所有像素值和, n_0 为每个特征图像素值总数。利用全连接层进行特征筛选并将结果转换为一维向量,提供给 MEB-KSVM 进行分类处理。

2.3.2 网络模型超参数优化 为解决传统卷积网络模型中人工超参数调优的经验主义,研究采用差分进化算法对模型的超参数组合进行自适应优化^[22]。使用差分进化算法对超参数寻优,以确定最佳参数组合,达到模型训练最佳效果^[18]。

超参数组合包含固定组合与自适应组合,其中固定参数为卷积核大小初始迭代次数、激活函数以及卷积滑动步长;自适应候选参数为学习率、学习衰减率、批量大小、全连接层的节点数、全连接层的 Dropout,其候选值^[23]与固定值如表 2 所示。

采用差分进化算法进行候选参数自适应,处理为:步骤(1):超参数种群初始化。根据第 2.2 节处理结果确定初始种群数为 40,每个种群包含 5 个个体,每个个体对应一个超参数。步骤(2):计算。根据阈值计算超参数种群适应度,实现种群优劣程度评判。通过 Deho-TransCNN 全连接计算预测值,并通过损失函数计算损失值确定适应度。对损失值与阈值进行判断,若迭代损失值与前一次损失值差的绝对值大于阈值,本组结果为超参数组合,反之执行变异、交叉、选择操作更新种群个体,变异满足表 2 候选值

表2 TransCNN超参数组合表
Table 2 TransCNN hyperparameter combination

参数自适应候选值		参数固定值	
超参数	取值范围	超参数	取值
学习率	{0.1, 0.01, 0.001, 0.0001}	卷积核尺寸	3×3
学习衰减率	{1e-3, 1e-4, 1e-5, 1e-6, 1e-7}	激活函数	ReLU
批量大小	{4, 8, 16, 32, 64, 128}	初始迭代次数	200
全连接层的节点数	{128, 256, 512, 1024}	卷积滑动步长	(2, 2)
全连接层的Dropout	{0.4, 0.5, 0.6, 0.7, 0.8}	-	-

范围。步骤(3):变异、交叉、处理。利用第2.2节中所述方式进行种群变异、交叉及选择处理,处理后形成子超参数组合,并将其与原种群超参数组合进行比较、择优。步骤(4):终止检验。对择优后超参数组合终止检验,终止检验将种群按照适应度进行降序排列,检验最高适应度种群是否满足收敛条件。满足则代表找到最优超参数组合,差分寻优结束,反之,进行新一轮种群初始化。初始化采用浅层模型到深层模型的权重转移方法,解决模型规模相差较大时的评估问题。新迭代中,部分权重继承最佳模型,剩余部分使用差分进化算法步骤(1)处理。新的迭代中只生成size(P)-n组新权重种群进行变异与交叉,并与保留的n组权重组合进行选择,完成一次新迭代。步骤(5):重复执行步骤(2)至步骤(5),直至满足收敛条件为止。其中,经实验测试n=0.05, size(P)=2时,效果最佳。

为解决因超参数组合较少引起重复组合,产生大量无用计算,影响寻优效率,对差分过程设置重复名单。将所有生成的种群保存至Repeat表,根据种群优化不断更新列表。在处理过程中,将新生成种群与Repeat表比较,若存在相同组合,舍弃,重新生成新种群。经差分进化算法处理后确定 Deho-TransCNN超参数配置组合进行模型训练。

2.3.3 MEB-KSVM分类 为解决多分类中样本不平衡及非线性可分问题,避免样本随机导致算术平均质心位置差异,研究使用MEB-KSVM对DR特征提取结果进行多分类。算法通过MEB计算特征提取向量的异常敏感样本,然后利用KSVM进一步分类构造DR病变严重情况分类器,得到DR的5种等级分类结果,依次分别为正常、轻度、中度、严重和增殖性DR。

MEB通过确定球心和半径划分球体,判断样本点是否归类于球体。划分过程中,为优化离群点,防止过拟合,引入非负参数 ξ_l 作为松弛参数,B作为惩罚参数对MEB约束^[24],优化函数如式(21)所示:

$$\begin{cases} \min_{c, R} (r^2 + B \sum_{l=1}^m \xi_l), B > 0 \\ \|\varphi(x_l) - c\|^2 \leq r^2 + \xi_l, \xi_l > 0, l = 1, 2, \dots, m \end{cases} \quad (21)$$

其中,c为最小包围球的球心,r为最小包围球的半径,m为训练样本总数, $\varphi(x_l)$ 为样本输入空间至特征空间的映射。

为降低映射到高维空间存储与计算负担,研究使用高斯函数为核函数,如式(22)所示:

$$K(x, z) = \varphi(x)^T \varphi(z) = \exp(-\gamma \|x - z\|^2), x \in X, z \in X \quad (22)$$

其中, γ 为高斯核函数参数。利用拉格朗日数乘法对r、 ξ_l 、B求导,如式(23)所示:

$$\begin{cases} \min_{\alpha} \sum_{l=1}^m \alpha_l K(x_l, x_l) - \sum_{l=1}^m \sum_{j=1}^m (\alpha_l \alpha_j K(x_l, x_j)) \\ 0 < \alpha_l < B, \sum_{l=1}^m \alpha_l = 1, l = 1, 2, \dots, m \end{cases} \quad (23)$$

其中, α_l 为拉格朗日算子,满足 $\alpha_l > 0$ 。处理得到最优解 $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*)^T$,其中, $\alpha_{k_l}^*$ 为 α^* 分量,满足 $0 < \alpha_{k_l}^* < B$,优化参数 r^{2*} 、 b^* ,求解如下所示:

$$r^{2*} = K(x_{k_l}, x_{k_l}) + \sum_{l=1}^m \sum_{j=1}^m (\alpha_l \alpha_j K(x_l, x_j)) - 2 \sum_{l=1}^m (\alpha_l K(x_l, x_{k_l})) \quad (24)$$

$$b^* = \sum_{k=1}^m \alpha_{k_l} \varphi(x_{k_l}) \quad (25)$$

其中, x_{k_l} 为集中支持向量样本点。分类决策函数f(x)如式(26)所示:

$$f(x) = \text{sign}(\sum_{l=1}^m \alpha_l^* [K(x_l, x) - K(x_l, x_{k_l})]) \quad (26)$$

其中,x为输入新样本点,当f(x)>0则与训练样本为一类;反之,为另一类。

MEB处理后,完成第一轮分类,将总样本分为DR病变样本和正常/增殖性DR样本,然后利用KSVM一对多方法进一步构建分类器,实现DR多分类。KSVM参考MEB中超参数C和 γ 进行超平面划分,目标函数^[25]如式(27)所示:

$$\min_{\gamma} P(\gamma) = 1 - z_1(\gamma) + z_2(\gamma) \quad (27)$$

其中,P(γ)为微函数,利用梯度下降法求解。若 γ 取

值 $z_1(\gamma)$ 趋近于1,则判定为同类别,反之,非同类。经实验测试,参数 $B=1.0, \gamma=0.1$ 分类效果最佳。

通过KSVM进一步分类,得到num1、num2、num3、num4、num5,对应为DR图像病变情况的正常、轻度、中度、严重和增殖性DR 5类结果,分别对应0~4级的DR严重程度分级。

3 结果

3.1 数据集与实验环境

研究使用Kaggle中APTOS 2019 Blindness Detection提供的糖尿病眼底图,共3 653张,经数据增广得到数据集,共11 689张,数据包括DR严重程度0~4级,分别对应正常、轻度、中度、严重和增殖性DR。其中,正常视网膜图像无异常特征;轻度白内障视网膜仅可见散在的点状红色的微血管瘤;中度白

内障介于轻度与重度之间的视网膜病变,可合并点、片状或放射状视网膜出血,黄白色的硬性渗出和团状白色的棉绒斑;严重白内障于每个象限视网膜内出血大于等于20个出血点,或至少2个象限的静脉串珠样改变,或至少1个象限视网膜内微血管一场,视网膜内微血管异常位于毛血管无灌注区边缘,呈树墩状或尖端形扩张;增殖性视网膜图像有新生血管形成或玻璃体积血或视网膜前积血,眼底能够看到增殖膜,甚至发生牵拉性视网膜脱落^[26-27]。数据样本如图7所示,数据参数如表3所示。研究以8:2比例划分训练集和测试集,并使用Intel(R) Core(TM) i7-10750H CPU、16 G内存、8 G Inter UHD Grapics GPU 硬件环境,Python3.8、Tensorflow、OpenCV、Numpy、Pytorch、Sklearn软件环境。模型训练初始参数使用表2固定参数信息。

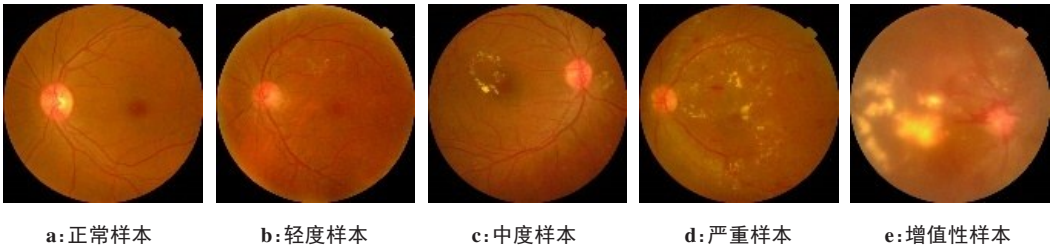


图7 数据集样本
Figure 7 Data set samples

表3 DR严重程度分级表
Table 3 DR severity scale

标签	英文表示	中文表示	训练集/张	测试集/张
0	No_DR	正常	5 773	1 443
1	Mild	轻度	1 254	314
2	Moderate	中度	3 136	784
3	Severe	严重	662	166
4	Proliferate_DR	增殖性	864	216

3.2 实验分析

3.2.1 评价指标 采用准确率(Acc)、AUC、灵敏度(Sensitivity)和特异性(Specicity)客观评价指标对糖尿病图像分类效果进行分析,如下所示:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{28}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{29}$$

$$Specificity = \frac{TN}{TN + FP} \tag{30}$$

其中,TP为预测患有DR的DR眼底图像数量,FP为

预测患有DR的正常眼底图像数量,TN为预测未患DR的正常眼底图像数量,FN为预测未患DR的DR眼底图像数量。其中患有DR的包括4类病变。

3.2.2 超参数选取 研究使用差分进行算法对TransCNN-Small网络进行超参数优化,设置数据集进行30次超参数寻优,结果如图8所示。由图8可知,网络损失值随着差分算法超参数优化不断降低,在迭代至25次出现趋平,即超参数优化在差分算法迭代循环至第25次时,其网络训练性能最佳,损失趋近于0.1。根据图8确定损失值为0.1,进行迭代训练,利用差分进化算法对模型进行超参数计算,得到最优超参数组合,如表4所示。

3.2.3 不同算法性能分析 研究在相同实验环境下,使用Mkmi-Graph^[3]、W-GLBP^[4]、改进CNN^[5]、ResNet50^[8]、M2CNN^[7]、BNnet^[9]、mResNext^[10]、SEC-UNet^[13]、VGG-16算法^[12]以及Deho-TransCNN进行模型训练,使用Sensitivity、Specicity、Acc、AUC和平均检测时间进行分析,如表5所示。由表5可知,Acc、AUC、Sensitivity、Specicity、平均检测时间的结果最好皆为Deho-TransCNN,分别达0.947、0.945、0.926、0.937、1.2 s。改进算

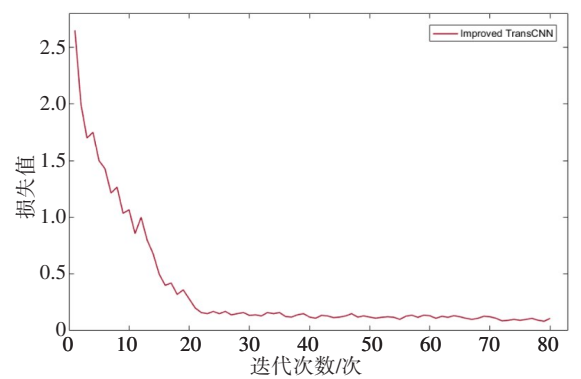


图8 损失值随优化次数变化曲线图
Figure 8 Variation of loss with number of optimizations

表4 TransCNN超参数优化结构表	
Table 4 Optimal hyperparameter combination for TransCNN	
超参数	取值范围
学习率	0.001
学习率的衰减率	1e-6
批量大小	32
全连接层的节点数	512
全连接层的Dropout	0.5

法较9种传统算法而言,其Acc、AUC、Sensitivity、Specicity,分别平均提升5.6%、7.9%、6.4%、5.1%。改进算法平均检测时间最佳,为1.2 s,远远低于其他传统模型检测时间。改进算法较最佳算法改进CNN而言,平均检测时间降低158.3%,改进算法在图像实时检测中性能较佳,更能满足实际运用中的实用性要求。改进算法对超参数进行自适应性组合调优,优化网络结构后,能有效地避免过拟合问题,提升算法效率及分类结果的准确率。在图像多分类检测时间上大大降低,更适用于实时图像分类场景。

3.2.4 Deho-TransCNN Area 值分析 使用Deho-TransCNN对模型进行训练,并对5种不同DR眼底病变的AUC-ROC和micro AUC-ROC以及macro AUC-ROC进行分析,结果如图9、图10所示。由图9可知,DR图像5种病变的最佳AUC分别为0.938、0.896、0.938、0.969、0.946。图10可知,其宏观平均值为0.947,微观平均值为0.939,改进算法对DR图像多分类处理中,效果无明显偏差,改进算法对眼底病变检测具有一定泛化性。

3.2.5 消融实验 为验证改进算法对分类检测效果和模型训练效果,研究在相同环境下使用基础TransCNN

表5 不同算法性能分析					
Table 5 Performance analysis of different algorithms					
模型	Acc	AUC	Sensitivity	Specicity	检测时间/s
Mkmi-Graph	0.907	0.861	0.876	0.902	8.9
W-GLBP	0.888	0.884	0.897	0.893	10.3
改进CNN	0.938	0.901	0.884	0.914	3.1
ResNet50	0.803	0.837	0.793	0.894	15.9
M2CNN	0.917	0.845	0.904	0.929	13.8
BNnet	0.861	0.794	0.832	0.785	18.8
mResNext	0.913	0.925	0.891	0.897	10.3
SEC-UNet	0.924	0.916	0.849	0.921	5.9
VGG-16	0.916	0.918	0.905	0.887	15.7
Deho-TransCNN	0.947	0.945	0.926	0.937	1.2

和改进算法进行模型训练,并对使用准确率、损失处理进行消融实验,如图11、图12所示。由图11可知,改进算法和基础TransCNN在训练前期两者皆实现较快收敛。改进算法在迭代25次后准确率趋于平稳,而TransCNN在迭代100次后才趋于平稳。改进算法的准确率优于基础TransCNN,准确率达到0.947,较基础TransCNN提升4.5%。改进算法收敛更快,分类效果更好,改进算法较基础算法有一定优

势。由图12可知,改进算法利用差分算法对权重参数和超参数进行寻优,其损失在迭代25次后趋于平缓,基础TransCNN损失在迭代100次后趋于平缓。差分算法寻优可以为模型提供参数组合最优,进而有效降低训练过程中无意义迭代。此外,改进算法的整体损失值和波动较基础TransCNN更小,寻优中能有效避免局部最优问题,进而有效降低模型训练时间和计算复杂度。

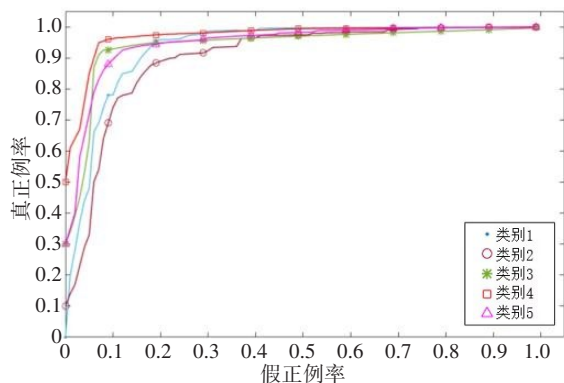


图9 Deho-TransCNN 5种类别 AUC-ROC 分析

Figure 9 Five-category AUC-ROC analyses of Deho-TransCNN

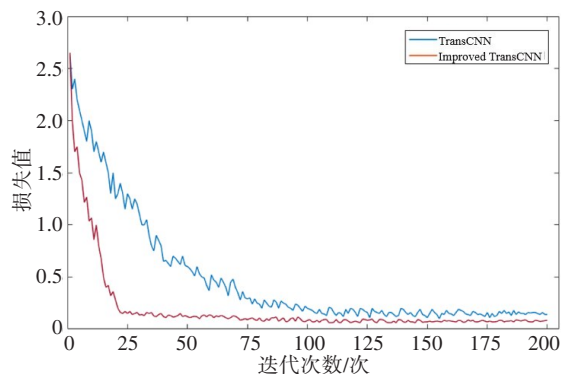


图12 TransCNN与改进算法的损失分析

Figure 12 Loss analyses of TransCNN and the improved algorithm

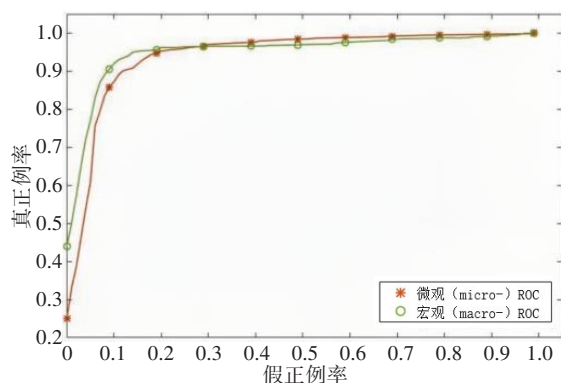


图10 Deho-TransCNN micro AUC-ROC 和 macro AUC-ROC 分析

Figure 10 Analyses of Deho-TransCNN micro AUC-ROC and macro AUC-ROC

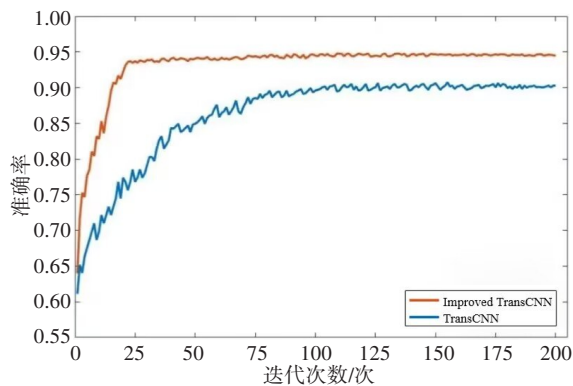


图11 TransCNN与Improved TransCNN的实验对比

Figure 11 Experimental comparison of TransCNN and Improved TransCNN

4 结语

本研究提出一种基于差分进化算法的超参数优化改进 TransCNN 图像多分类算法,算法在 TransCNN 基础上,优化网络参数和初始权重,并使用 MEB-KSVM 算法对特征提取结果进行多分类。使用改进算法与多种深度学习算法进行对比实验发

现,改进算法在准确率、图片平均检测时间上有较大提升;在实际应用中,除了可以降低模型训练所需计算资源外,在实时检测上有一定促进,能在一定程度上辅助专家进行 DR 初步预判,实现患者分流,降低患者与医生间的医患缺口。在后续研究中,可探索将其应用于眼底图像分割与检测中以解决实际问题。

【参考文献】

- [1] 李云翠,石亚静,张欢,等. 糖尿病性心脏病的治疗进展[J]. 医学综述, 2016, 22(18): 3625-3628.
Li YC, Shi YJ, Zhang H, et al. The recent advances in the treatment of diabetic heart disease[J]. Medical Recapitulate, 2016, 22(18): 3625-3628.
- [2] 高晔,孙桂波,罗云,等. 糖尿病视网膜病的发病机制及药物干预研究进展[J]. 中国药理学通报, 2020, 36(4): 491-495.
Gao Y, Sun GB, Luo Y, et al. Research progress on the pathogenesis of diabetic retinopathy and pharmacological interventions[J]. Chinese Journal of Pharmacology, 2020, 36(4): 491-495.
- [3] 任福龙,曹鹏,杨金柱,等. 多核多示例学习的糖尿病视网膜病变诊断[J]. 中国图象图形学报, 2018, 23(4): 552-563.
Ren FL, Cao P, Yang JZ, et al. Multi-kernel multi-instance learning based diabetic retinopathy diagnosis[J]. Journal of Image and Graphics, 2018, 23(4): 552-563.
- [4] 张晨曦,陈明惠,高乃璐,等. 基于小波分解金字塔的 OCT 图像特征提取算法[J]. 光学技术, 2021, 47(1): 72-79.
Zhang CX, Chen MH, Gao NJ, et al. Feature extraction algorithm of OCT image based on wavelet decomposition pyramid[J]. Optical Technique, 2021, 47(1): 72-79.
- [5] Wan SH, Liang Y, Zhang Y. Deep convolutional neural networks for diabetic retinopathy detection by image classification[J]. Comput Electr Eng, 2018, 72(7): 274-282.
- [6] Sunil S, Saamil M, Anupam S. An intelligible deep convolution neural network based approach for classification of diabetic retinopathy[J]. Bio-Algorithms Med-S, 2018, 14(2): 13.
- [7] Zhou K, Gu Z, Liu W, et al. Multi-cell multi-task convolutional neural networks for diabetic retinopathy grading[J]. Conf IEEE Eng Med Biol Soc, 2018: 2724-2727.
- [8] Zhang J, Xie BW, Wu X, et al. Classification of diabetic retinopathy severity in fundus images with DenseNet121 and ResNet50[J]. arXiv preprint arXiv: 2108.08473, 2021.
- [9] 李琼,柏正尧,刘莹芳. 糖尿病性视网膜图像的深度学习方法[J]. 中国图象图形学报, 2018, 23(10): 1594-1603.
Li Q, Bai ZY, Liu YF. Automated classification of diabetic retinal images by using deep learning method[J]. Journal of Image and Graphics, 2018, 23(10): 1594-1603.

- [10] 李轩屹, 朱晓军. 多尺度深度CNN的糖尿病视网膜病变分类[J]. 计算机工程与设计, 2020, 41(11): 3229-3234.
Li XY, Zhu XJ. Classification of diabetic retinopathy with multi-scale depth CNN[J]. Computer Engineering and Design, 2020, 41(11): 3229-3234.
- [11] 孙雨琛, 刘宇红, 张达峰, 等. 基于深度学习的糖尿病视网膜病变诊断方法[J]. 激光与光电子学进展, 2020, 57(24): 359-366.
Sun YC, Liu YH, Zhang DF, et al. Diagnosis method of diabetic retinopathy based on deep learning[J]. Laser & Optoelectronics Progress, 2020, 57(24): 359-366.
- [12] 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.
Lian XF, Liu ZY, Zhang L, et al. A retinopathy image recognition method based on deep learning[J]. Computer Applications and Software, 2021, 38(1): 179-185.
- [13] 许祥丛, 陈俊彦, 王雪花, 等. 通过SEC-UNet精准分割糖尿病视网膜病变眼底OCT图像脉络膜层[J/OL]. 生物化学与生物物理进展, 2022, 38(1): 179-185.
Xu XC, Chen JY, Wang XH, et al. Precise segmentation of choroid layer in diabetic retinopathy fundus OCT images by using SEC-UNet[J/OL]. Progress in Biochemistry and Biophysics, 2022, 38(1): 179-185.
- [14] Liu Y, Sun GL, Qiu Y, et al. Transformer in convolutional neural network[J]. arXiv preprint arXiv: 2106.03180, 2021.
- [15] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. arXiv preprint arXiv: 1706.03762, 2017.
- [16] Cordonnier JB, Loukas A, Jaggi M. On the relationship between self-attention and convolutional layers[J]. arXiv preprint arXiv: 1911.03584, 2019, 2020.
- [17] 丁青峰, 尹晓宇. 差分进化算法综述[J]. 智能系统学报, 2017, 12(4): 431-442.
Ding QF, Yin XY. Research survey of differential evolution algorithms[J]. CAAI Transactions on Intelligent Systems, 2017, 12(4): 431-442.
- [18] 邹东池, 白富生. 基于差分进化算法的神经网络优化设计[J]. 重庆师范大学学报(自然科学版), 2022, 39(1): 79-89.
Zou DC, Bai FS. Optimal neural network design based on the differential evolution algorithm[J]. Journal of Chongqing Normal University (Natural Science), 2022, 39(1): 79-89.
- [19] Qing A. A parametric study on differential evolution based on benchmark electromagnetic inverse scattering problem[C]. IEEE CEC, 2007: 293-298.
- [20] Sandler M, Howard A, Zhu ML, et al. MobileNetV2: inverted residuals and linear bottlenecks[J]. CVPR, 2018.
- [21] 许玥, 冯梦如, 皮家甜, 等. 基于深度学习模型的遥感图像分割方法[J]. 计算机应用, 2019, 39(10): 2905-2914.
Xu Y, Feng MR, Pi JT, et al. Remote sensing image segmentation method based on deep learning model[J]. Journal of Computer Applications, 2019, 39(10): 2905-2914.
- [22] Xiao X, Yan M, Basodi S, et al. Efficient hyperparameter optimization in deep learning using a variable length genetic algorithm[J]. arXiv preprint arXiv: 2006.12703, 2020.
- [23] 齐欣, 赵玲玲, 班晓军. 基于Q-learning的神经网络超参数优化[C]//第40届中国控制会议论文集(15). 2021: 759-764.
Qi X, Zhao LL, Ban XJ. Hyper-parameter optimization of neural networks using Q-learning[C]//Proceedings of the 40th Chinese Control Conference. 2021: 759-764.
- [24] 杨迪, 方扬鑫, 周彦. 基于MEB和SVM方法的新类别分类研究[J]. 广西师范大学学报(自然科学版), 2022, 40(1): 57-67.
Yang D, Fang YX, Zhou Y. New category classification research based on MEB and SVM methods[J]. Journal of Guangxi Normal University (Natural Science Edition), 2022, 40(1): 57-67.
- [25] 丁茜, 赵晓东, 吴鑫俊, 等. 基于RBF核的多分类SVM滑坡易发性评价模型[J]. 中国安全科学学报, 2022, 32(3): 194-200.
Ding Q, Zhao XD, Wu XJ, et al. Landslide susceptibility assessment model based on multi-class SVM with RBF kernel[J]. China Safety Science Journal, 2022, 32(3): 194-200.
- [26] 赵学功. 基于眼底图像的糖尿病诱发视网膜病变辅助诊断关键技术研究[D]. 成都: 电子科技大学, 2021.
Zhao XG. Research on key technologies for auxiliary diagnosis of diabetic-induced retinopathy based on fundus images[D]. Chengdu: University of Electronic Science and Technology, 2021.
- [27] Flaxel CJ, Adelman RA, Bailey ST, et al. Diabetic retinopathy preferred practice pattern[J]. Ophthalmology, 2020, 127(1): 66-145.

(编辑:陈丽霞)