

DOI:10.16136/j.joel.2022.09.0871

基于多层次特征融合的口腔粘膜性疾病识别方法

张道奥¹, 高 明¹, 刘 青^{2,3*}, 王舒研^{2,3}, 汪苑苑^{2,3}

(1 西安工业大学 光电工程学院, 陕西 西安 710021; 2. 空军军医大学第三附属医院 口腔粘膜科, 陕西 西安 710032; 3. 军事口腔医学国家重点实验室/国家口腔疾病临床医学研究中心/陕西省口腔疾病临床医学研究中心, 陕西 西安 710032)

摘要: 口腔粘膜性疾病的识别主要依靠医生肉眼观察并进行主观判断, 该方法导致疾病识别的准确率低、医生的工作量大。针对以上问题提出一种基于多层次特征融合的口腔粘膜性疾病识别方法。对口腔疾病图像提取深层次特征和浅层次特征共两种, 使用 EfficientNet 模型做深层特征的提取, 使用 HSV、方向梯度直方图 (histogram of oriented gradient, HOG) 和灰度共生矩阵 (gray-level co-occurrence matrix, GLCM) 分别提取口腔疾病的颜色、形状以及纹理的浅层特征, 将特征融合后利用随机森林 (random forest, RF) 算法进行特征选择, 选取特征重要性更大的特征, 降低特征的维度。最后结合多种机器学习分类器进行分类识别。使用收集到的口腔粘膜疾病数据集进行实验验证, 实验结果表明, 该方法能达到准确率 (accuracy, Acc) 92.89%、灵敏度 (sensitivity, Sen) 89.91%、特异性 (specificity, Spe) 96.06% 以及 AUC (area under the curve) 98.09%, 有效地解决识别中误判多、准确率低等问题。

关键词: 口腔粘膜疾病; efficientNet; 多层次特征融合; 随机森林 (RF); 分类器

中图分类号: TP391 **文献标识码:** A **文章编号:** 1005-0086(2022)09-0968-09

Recognition of oral mucosal diseases based on multi-level feature fusion

ZHANG Daoao¹, GAO Ming¹, LIU Qing^{2,3*}, WANG Shuyan^{2,3}, WANG Yuanyuan^{2,3}

(1. College of Optoelectronic Engineering, Xi'an Technological University, Xi'an, Shaanxi 710021, China; 2. Department of Oral Mucosa, the Third Affiliated Hospital of Air Force Military Medical University, Xi'an, Shaanxi 710032, China; 3. State Key Laboratory of Military Stomatology/National Stomatological Clinical Research Center/Shaanxi Stomatological Clinical Research Center, Xi'an, Shaanxi 710032, China)

Abstract: The recognition of oral mucosal diseases mainly depends on doctors' visual observation and subjective judgment. This method leads to low accuracy of disease recognition and heavy workload of doctors. To solve the above problems, an oral mucosal disease recognition method based on multi-level feature fusion is proposed. There are two kinds of deep-level features and shallow features extracted from oral disease images. The efficientNet model is used to extract the deep features. HSV, histogram of oriented gradient (HOG) and gray level co-occurrence matrix (GLCM) are used to extract the shallow features of color, shape and texture of oral diseases respectively. After feature fusion, the random forest (RF) algorithm is used to select the features with greater feature importance, reducing the dimension of the feature. Finally, a variety of machine learning classifiers are combined for classification and recognition. The datasets of oral mucosal diseases collected are used for experiment verification. The experimental results show that the method can achieve the accuracy (Acc) of 92.89%, sensitivity (Sen) of 89.91%, specificity (Spe) of 96.06% and area under the curve (AUC) of 98.09%. It effectively solves the problems of many misjudgments and low accuracy in recognition.

* E-mail: liuqing@fmu.edu.cn

收稿日期: 2021-12-24 修订日期: 2021-01-27

基金项目: 陕西省重点实验室项目 (17JS052) 和军事口腔医学国家重点实验室自主课题 (2019ZA07) 资助项目

Key words: oral mucosal diseases; efficientNet; multi-level feature fusion; random forest (RF); Classifier

1 引言

近年来口腔健康受到越来越多的关注,口腔白斑(oral leukoplakia, OLK)、口腔扁平苔藓(oral lichenplanus, OLP)、口腔癌症(oral cancer, OC)和复发性口腔溃疡(recurrent oral ulcer, ROU)均属于常见的口腔黏膜性疾病。OLK和OLP均属于癌前病变,有转变为OC的机率,ROU若长久不愈也存在恶变的可能,因此及时诊断治疗十分重要。口腔粘膜性疾病的主要症状相似度高,导致了4种疾病难以鉴别。目前,临床诊断方法主要是在大量观察的基础上,进行定性的描述。该方法一方面使医生的工作量大,检验效率低,且医生可能因为长时间工作,注意力下降,导致对疾病的错误诊断,影响患者治疗;另一方面,对病理图像的分析识别易受医师经验和视觉分辨率的限制,产生主观因素,缺乏科学客观的定量依据,有时还会因为需要进一步病理活检而延长检查周期。因此如何提升口腔黏膜性疾病的识别准确率是本文的研究重点,通过本文研究希望减轻医师的工作负担,提高医生的工作效率,减轻病人的痛苦以及降低疾病导致的经济损失。

图像识别的整体过程包括对图像数据增强、图像的分割、图像的特征提取,以及最后的分类。在整个过程中,特征的提取阶段十分重要,好的特征能够很大程度上提升识别的准确率。传统图像的特征识别、提取做法是利用人为设计的特征进行分类识别,也就是提取各种生物学上具有可解释性的、临床上有意义的形状以及基于形态学的特征,包括灰度纹理特征、颜色特征、病损形态特征等浅层次特征。对于口腔粘膜性疾病,传统特征提取的方法能取得较为不错的效果,例如:NAWANDHAR等^[1]从口腔鳞状细胞癌图像中提取了305个特征,包括利用RGB和HSV图像提取的颜色特征、方向梯度直方图(histogram of oriented gradient, HOG)特征、灰度共生矩阵(gray-level co-occurrence matrix, GLCM)纹理特征、梯度特征、几何特征和tamura特征,使用5个不同的分类器结合最大投票法进行分类,最终识别准确率(accuracy, Acc)可以达到95.96%。RAHMAN等^[2]从720幅图像中获取口腔癌图像的颜色、纹理、形状共3种特征,其中形状特征包括基于区域的形状特征(如面积和紧致度)以及基于边界的形状特征(如周长和偏心率),纹理特征包括GLCM

4个方向的特征以及LBP(local binary pattern)特征,颜色特征使用RGB图像提取128维特征,将所有特征经过PCA(principal components analysis)选择后使用多种分类器进行分类,其中支持向量机(support vector machine, SVM)和Logistic回归可以获得100%的Acc。

随着深度学习技术的不断进步,卷积神经网络在医学图像的不同领域持续取得新突破,被广泛应用于多种医学图像分析中。特征提取技术也从原来的人为设计特征变成使用神经网络代替寻找最佳特征,神经网络能够自主从数据集中提取具有分离性的特征,并且具有较强的泛化能力和鲁棒性。近年来,卷积神经网络在口腔疾病的诊断中发挥作用^[3,4]。RAMALINGAM等^[5]利用Inceptionv3和ResNet50提取OC图像的特征,后接随机森林(random forest, RF)和SVM做分类器,对OC的病损程度进行多级分类,结果表明,利用ResNet50提取特征接RF算法做分类Acc最高,在测试集上可以达到92.08%。PALASKAR等^[6]使用经过在ImageNet上预训练的网络模型MobileNet^[7]、ResNet50、InceptionV3对OC进行识别,比较了3个模型在数据增强和微调下的模型性能,发现ResNet50的效果最好,Acc可以达到91.13%,AUC(area under the curve)为87%。DAS等^[8]选择4个预训练的模型Alexnet、VGG-16、VGG-19和Resnet-50做OC的二分类识别,以找到最适合此分类问题的模型,后发现VGG16效果最好,因此基于VGG16网络进行改进,网络加入了批归一化(batch normalization, BN)层和Drop-out层,加快收敛的同时减少过拟合风险,最后Acc达到97.5%。

本文完成4种口腔粘膜性疾病的识别任务,一方面口腔疾病的种类和数量增加导致了病损部位不固定,病症更加复杂多样。另一方面由于四种疾病都与癌症存在关联,导致4种口腔疾病的类内差异较大,类间差异较小。针对以上问题,本文提出一种基于多层次特征融合的口腔粘膜性疾病识别方法。将深层次特征与浅层次特征相结合,得到更加丰富的图像特征信息来提高识别准确率。该方法使用HSV、HOG和GLCM分别提取口腔疾病的颜色、形状以及纹理的浅层次特征,并与EfficientNet模型提取出的深层次特征相融合,利用RF算法计算特征重要性进行特征选择,最后采用SVM、Xgboost等机器学习分类器进行

分类识别。实验结果表明,使用 SVM 分类器能取得最好的分类效果, *Acc* 最高可以达到 92.89%, *Sen* 达到 89.91%、*Spe* 达到 96.06% 以及 *AUC* 值达到 98.09%,可以有效地解决口腔疾病识别过程中误判多和准确率低等问题。

2 口腔数据集建立

2.1 口腔数据集收集

目前基于图像的口腔粘膜性疾病的识别尚处于起步阶段,国内外没有类似的大型公开数据集,很多口腔粘膜性疾病的研究也都是来自于团队从医院收集的数据,本文的工作之一就是收集用于研究的口腔粘膜性疾病数据。本文采集的口腔图像来自第四军医大学口腔医院,为了保证数据标签的绝对准确性,所有数据的标签均是来自于病理活检确认后的结果,因为对于口腔粘膜性疾病,病理活检是绝对的金标准,有时比专业的医生给出的标签要准确得多。因此本文所有图像数据均经过口腔医生的病例确诊后进行拍摄收集。分别采集 OLK、OLP、OC 和 ROU 共 4 种图像。目前在医院采集的图像是专业

医生通过 Canon EOS 60D 相机在白光下拍摄的。从 2020 年 10 月开始到 2021 年 5 月,共采集图像总数 1 125 张,被采集的患者年龄、性别以及健康情况各不相同,保证了采集数据的随机性,同时为了后续进行 ROI(region of interest)提取,寻找了 4 名专业的口腔粘膜科医生手动进行口腔病灶区域的划分。4 种口腔图像数目如表 1 所示,ROU 共采集到 212 张图片,OC 共采集到 255 张图片,OLK 共采集到 271 张图片,OLP 共采集到 387 张图片,所有图像格式均为 JPG,分辨率大小均为 5 184×3 456,4 种疾病的图片数量较为均衡,其中 OLP 数量最多,共有 387 张。图 1 为两种疾病的采集过程与采集到的图像。

表 1 4 种口腔疾病图片数量
Tab. 1 Number of pictures of four oral diseases

Disease type	Number of pictures/piece
ROU	212
OC	255
OLK	271
OLP	387

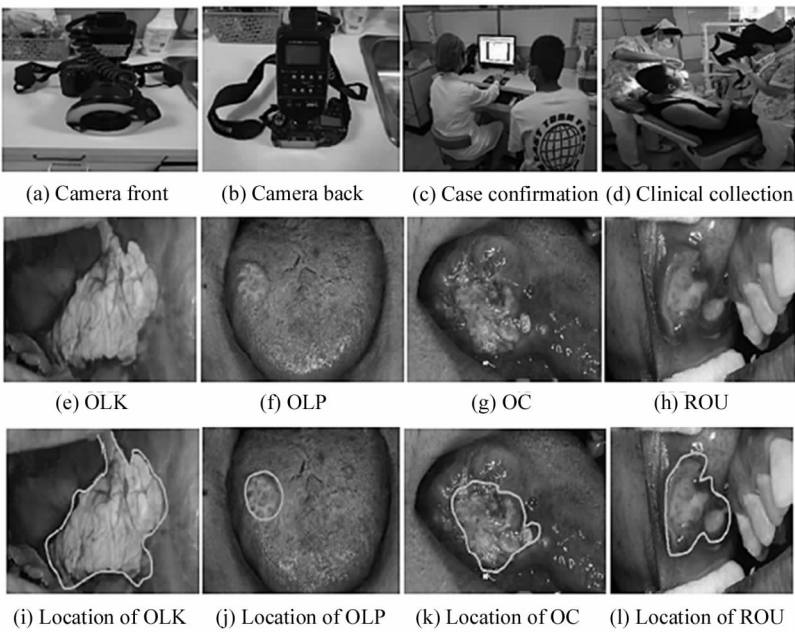


图 1 采集口腔粘膜性疾病图像
Fig. 1 Collected images of oral mucosal diseases

2.2 数据增强

采集到的图像数据总数为 1 125 张,包含共 4 种口腔粘膜性疾病。本文对原始数据进行数据增强,

防止模型过拟合增加泛化能力,同时对数据进行扩充。数据增强方式包括:图像旋转 90°、增加对比度、随机裁剪、水平翻转以及对它们进行随机组合,将扩

充后的图像数据按照 7 : 3 的比例随机划分为训练集和测试集,训练集用于训练各类分类器,测试集用于测试不同分类器的能力。

2.3 数据预处理

对数据进行预处理,根据医生手动勾画的口腔病灶边界进行 ROI 的选取,同时对 ROI 采用自适应中值滤波进行预处理,提高图像的灰度,去除不需要的噪声来增强图像以提高像素清晰度,减少了噪声对口腔病灶区域的影响。

3 口腔疾病图像识别原理及方法

使用上述采集的口腔疾病图像展开实验,OLK、

OLP、OC 和 ROU 4 种疾病的特征相似难以辨别,因此本文提出基于多层次特征融合的口腔疾病识别方法,将深层次特征与浅层次特征结合。首先使用 4 名专业临床医生手动勾画的口腔粘膜疾病的病灶边界作为分割的结果;然后对每个口腔样本使用 HSV、HOG 和 GLCM 分别提取口腔疾病的颜色、形状以及纹理的浅层特征,同时使用 EfficientNet 网络模型做深层特征的提取,将提取的特征融合,利用 RF^[9] 算法进行特征选择,选取与目标值关系更大的特征,后采用 SVM、RF 等传统分类器进行分类识别。4 种口腔疾病的整体识别结构图如图 2 所示。主要包含特征提取、特征选择和分类识别 3 部分。

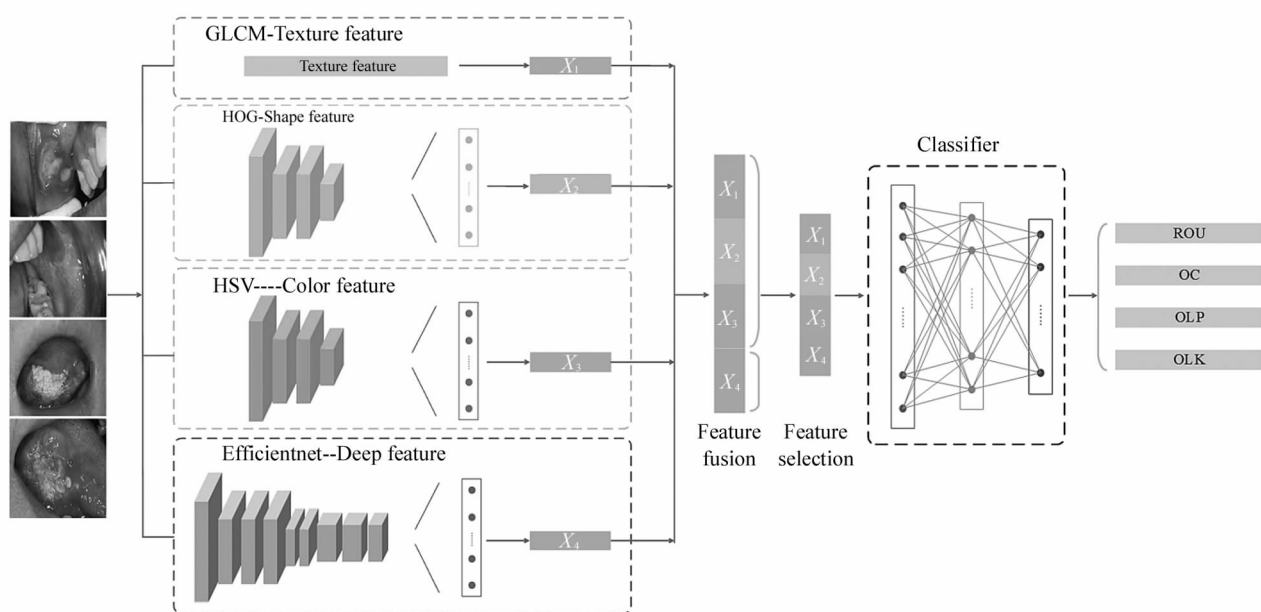


图2 模型整体结构图

Fig. 2 Overall structure diagram of model

3.1 GLCM 纹理特征

GLCM 是常用的纹理特征提取算法^[10],通过计算灰度图像得到它的共生矩阵,然后计算该共生矩阵得到矩阵的部分特征值来分别代表图像的某些纹理特征,能够较好地反映图像灰度关于方向、相邻间隔、变化幅度等综合信息。MAHMOUD 等^[11]将皮肤镜图像划分为 4 个子区域,用 8×8 矩阵重建每个子区域。然后,计算 GLCM 在 $(0^\circ, 45^\circ, 90^\circ, 135^\circ)$ 4 个方向上的特征向量,在每个方向上提取 GLCM 的 22 个特征向量。通过计算 4 个子区域 $\times$ 4 个方向 $\times$ 1 个距离 $\times$ 22 个特征向量,共提取出 1 个维度为 352 的向量。本文将对经过预处理的 ROI 图像同样采用 GLCM 在 $(0^\circ, 45^\circ, 90^\circ, 135^\circ)$ 4 个方向上提取上

述 4 个特征量的纹理能量、对比度、纹理熵和相关性,除此之外还提取包括差异性、同质性、协方差、纹理均值和逆差分矩共 40 维特征。

3.2 HOG 形状特征

本文采用 HOG 来描述 4 种口腔图像的形状特征^[12],因为 HOG 提取的边缘和梯度特征能很好地抓住局部形状的特点,同时能提高模型对光照因素的鲁棒性,通过计算图像局部的 HOG 来表示口腔病灶局部形状特征。HOG 特征在 OC 的识别中使用十分广泛,如 PRAGNA 等^[13]对获得的口腔图像进行自适应中值滤波预处理,从预处理图像中提取 HOG 特征大小 256×256 和 GLCM 特征大小 1×23 等特征。通过特征选择去除冗余特征后使用 SVM 分类

算法对 OC 进行良恶性分类,最后 Acc 可以达到 $\times 4$ 个细胞单元划分为 1 个方向直方图 block,步长设置为 1,把梯度方向平均划分位 9 个区间,采用重叠的方式求出方向直方图的特征。HOG 特征计算式为:

$$G_x(x,y) = H(x+1,y) - H(x-1,y), \quad (1)$$

$$G_y(x,y) = H(x,y+1) - H(x,y-1), \quad (2)$$

式中, $G_x(x,y)$ 代表图像中像素点水平方向梯度, $G_y(x,y)$ 代表图像中像素点垂直方向梯度, $H(x,y)$ 代表图像的像素值。像素点 (x,y) 处的梯度幅值和梯度方向分别如式(3)(4)所示:

$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}, \quad (3)$$

$$\alpha(x,y) = \arctan\left(\frac{G_y(x,y)}{G_x(x,y)}\right). \quad (4)$$

得到的 HOG 特征数量较大,DE 等^[14]在研究中

97%。本文把每 8×8 个像素划分为 1 个细胞单元,4 指出 HOG 特征对于检测胃癌,行人检测以及喉癌等有很好的效果,但唯一的缺点是,产生非常大的特征向量,会增加模型处理时间。为了解决此问题,他们在 HOG 特征的基础上加入神经网络以进一步提取,结果表明,这能够在保留 HOG 特征信息的同时,加快模型运行速度。因此本文将得到的 HOG 特征连接在 2 个卷积层、2 个池化层以及 1 个最终输出特征的全连接层(fully-connected layer, FC)上。卷积核大小分别为 $32 \times 3 \times 3$, $64 \times 3 \times 3$ 。池化层大小分别为 $32 \times 2 \times 2$, $64 \times 2 \times 2$ 。进行 HOG 特征向量的提取,网络结构如图 3 所示,激活层用激活函数 Relu,池化层采用 MaxPooling, FC 的神经元个数为 128 个。最终通过神经网络将得到的 128 维向量作为图像的 HOG 特征。

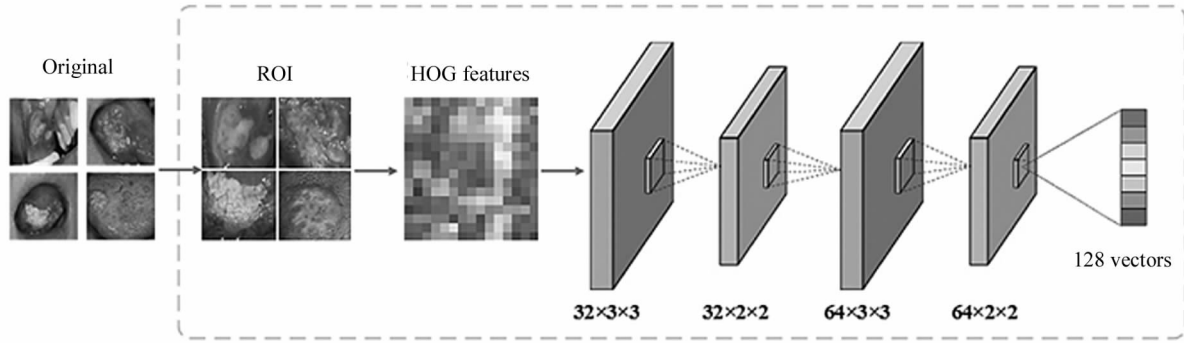


图 3 提取 HOG 特征的模型结构图

Fig. 3 Model structure diagram for extracting HOG features

3.3 HSV 颜色特征

HSV 图像,即代表色调(H)、饱和度(S)、亮度(V),与 RGB 图像相比,HSV 图像更加符合人类对于颜色的直观感受。之所以选择 HSV 图像作为另一种特征,是因为在拍摄口腔图像时,亮度的变化对口腔图像的识别会产生一定程度的影响。因此,本文将口腔图像从 RGB 颜色空间转化为 HSV 颜色空间,转换式如(5)~(7)所示:

$$H =$$

$$\begin{cases} 0^\circ, \text{if } MAX = MIN \\ 60^\circ \times \frac{G-B}{MAX-MIN} + 0^\circ, \text{if } MAX = R \text{ and } G \geq B \\ 60^\circ \times \frac{G-B}{MAX-MIN} + 360^\circ, \text{if } MAX = R \text{ and } G < B \\ 60^\circ \times \frac{B-R}{MAX-MIN} + 120^\circ, \text{if } MAX = G \\ 60^\circ \times \frac{R-G}{MAX-MIN} + 240^\circ, \text{if } MAX = B \end{cases}, \quad (5)$$

$$S = \begin{cases} 0^\circ, \text{if } MAX = 0 \\ \frac{MAX - MIN}{MAX} = 1 - \frac{MIN}{MAX}, \text{otherwise} \end{cases}, \quad (6)$$

$$V = MAX, \quad (7)$$

式中: R, G, B 为红、绿、蓝的颜色值; H, S, V 为色调、饱和度、亮度的值; MAX 为 R, G, B 中的最大值, MIN 为最小值; H 在 $[0, 360^\circ]$ 之间, S 在 $[0, 100^\circ]$ 之间, V 在 $[0, MAX]$ 之间。转换后的图像经过上述自构建的卷积神经网络同样可以得到 128 维颜色特征向量。

3.4 EfficientNet 提取深层特征

深度学习的特征使用 RGB 图像进行提取。RGB 图像,即代表红(R)、绿(G)、蓝(B)3个通道的颜色,通过不同颜色分量来表示彩色图像。本文收集的口腔数据集,即 RGB 图像,直接作为卷积神经网络的输入图像。本文的卷积神经网络选择 EfficientNet 网络作为主干网络, EfficientNet 自诞生起

就在多个大型图像数据集中取得优异的成绩^[15]。

本文的口腔图像数据集总量较少,因此为防止模型过拟合,加快训练速度,使用模型参数较小的 EfficientNetB0 网络做为底层网络对口腔图像进行特征提取,为了使模型更好地适应口腔疾病图像分类实验,去掉原神经网络的顶层,重新构建模型的 FC 进行分类。EfficientNetB0 网络的结构如表 2 所示;EfficientNetB0 网络参考 MobileNet V2 网络,由多个改进的 MBConv 组成,改进的 MBConv 模块先对输入进行 1×1 的逐点卷积并根据扩展比例对输出通道进行升维,接着在 $k \times k$ 的 Depthwise 卷积后加入 SENet^[16] 中的 SE 模块,引入注意力机制使网络关注信息量更大的通道,最终再使用 1×1 的卷积降维恢复原来的通道数目。改进后的 MBConv 模块参数更少,收敛速度更快,加入的注意力模块使网络学

习到更多有用特征。目前大部分医学图像识别分类使用的神经网络是 ResNet50、InceptionV3 等经典神经网络,但是 EfficientNet 能够提取到更加丰富的图像特征,具有更好的模型分类效果,并且它在多个医学图像识别分类中都被证明能取得更好的识别效果^[17,18]。

为了适应本文口腔疾病的识别任务,去掉 EfficientNetB0 的顶层网络,保留特征提取层,重新构建自己的网络顶层。加入一层全局池化层(global average pooling,GAP),一层包含 512 个神经元的 FC,以及最后的输出层使用 Softmax 分类器包含 4 个神经元,模型整体结构如图 4 所示,通过模型训练提取出 512 维特征向量。因为口腔数据集数量较小,模型的总迭代次数 epochs 均设置为 50,批量大小(batchsize)均设置为 64,使用 Adam 算法优化目标函数,损失函数使用 CrossEntropy。

将前面所有特征进行组合,包括 40 维纹理特征,128 维形状特征,128 维颜色特征,共 296 维底层特征,加上由 EfficientNet 提取的 512 维深层特征,组合起来共 808 维特征向量。有效准确的特征向量可以提高口腔疾病的分类精度,但当特征冗余达到一定程度时,不仅会增加整个模型的运算量,还会降低分类的准确率,因此使用 RF 算法计算特征重要性,保留重要性较高的图像特征。经过 RF 选择后共保留 476 维特征向量。

表 2 EfficientNetB0 的网络结构
Tab. 2 Network structure of EfficientNetB0

Stage	Structure	Resolution	Channels	Layers
1	Conv3×3	224×224	32	1
2	MBConv1,k3×3	112×112	16	1
3	MBConv6,k3×3	112×112	24	2
4	MBConv6,k5×5	56×56	40	2
5	MBConv6,k3×3	28×28	80	3
6	MBConv6,k5×5	14×14	112	3
7	MBConv6,k5×5	14×14	192	4
8	MBConv6,k3×3	7×7	320	1

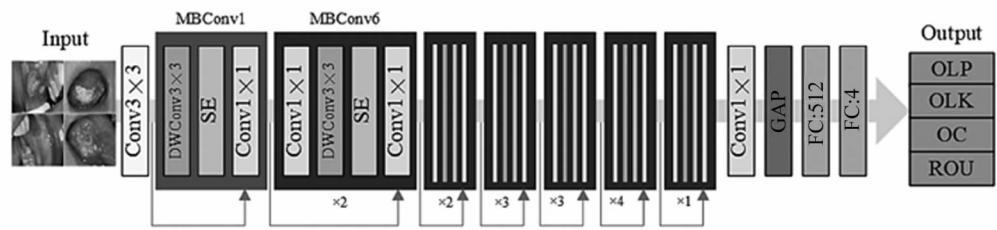


图 4 EfficientNet 网络整体结构图
Fig. 4 Overall structure diagram of efficientNet network

4 实验结果与分析

本文实验环境是基于 Ubuntu16.04 系统,采用 TensorFlow1.15 和 keras2.4.3 深度学习框架实现口腔疾病图像的分类识别。CPU 使用 Intel Xeon (R) E5-2630 v3,2.40 GHz,64 GB 内存,GPU 使用 2 块 GTX2080TI,实验语言为 Python3.7。为确保实验的准确可靠,所有分类器都被设置为默认参数

进行实验。

4.1 评价指标

所有的实验均使用相同的口腔数据集。实验的评价指标采用包括 Acc、灵敏度(sensitivity, Sen)、特异性(specificity, Spe)和 AUC 共 4 个不同的性能指标对本文的模型进行整体评估。各个指标的算式如式(8—10)所示:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \quad (8)$$

$$Sen = \frac{TP}{TP + FP}, \quad (9)$$

$$Spe = \frac{TN}{TN + FP}, \quad (10)$$

式中, TP (true positives)表示被正确识别为正例的病例个数, FP (false positive)表示被错误识别为正例的病例个数, TN (true negative)表示被正确识别为负例的病例个数, FN (false negative)表示被错误识别为负例的病例个数。AUC 计算的是 ROC 曲线下的面积, AUC 的值越接近 1, 表明模型诊断准确性越高, 模型的整体性能越好。

4.2 传统手工设计特征识别实验

表 3 为传统手工设计特征实验的分类结果, 根据本文提取的传统特征, GLCM 特征、HOG 特征和颜色特征, 使用不同的分类器进行分类识别实验, 以下实验结果为不同分类器在测试集上的 Acc , 其中 KNN 为 K 最近邻分类器, SVM 为支持向量机, RF

为随机森林, GBDT 为梯度提升树。实验对 3 种特征进行随机组合, 共产生 7 种不同的组合模型。

从实验结果看, 基于多特征融合的口腔粘膜性疾病分类模型 4、模型 5、模型 6 和模型 7 的 Acc 大体上高于只有单一特征的分类模型。因为特征维度越高, 图像的语义信息就越丰富, 能够从多维度进行决策分类。在单一特征模型中, 表示纹理的 GLCM 特征取得了最好的分类效果, 在 RF 中, Acc 可以达到 82.52%。在 2 个特征融合的模型中, 模型 4、模型 5 和模型 6 的 Acc 都得到了不同程度的提升, 其中模型 5 (GLCM 特征+颜色特征) 中 Xgboost 可以达到 87.03%。多特征融合模型中, 模型 7 (GLCM 特征+HOG 特征+颜色特征) 3 个特征融合的 Xgboost 模型分类效果最好, 可以达到 89.72%。 Acc 相较单一特征提升了 7.2%, 较 2 个特征融合模型提高了 2.69%。总体来说, 模型 7 代表的 3 个特征的融合分类

表 3 传统手工设计特征识别结果

Tab. 3 Feature recognition results of traditional manual design

Model	Feature			$Acc/\%$				
	GLCM	HOG	Color	KNN	SVM	RF	GBDT	Xgboost
1	✓	—	—	81.57	79.64	82.52	78.34	80.42
2	—	✓	—	66.89	68.12	69.38	67.46	69.89
3	—	—	✓	75.77	78.92	79.04	74.36	77.86
4	✓	✓	—	83.14	82.78	82.93	79.79	81.63
5	✓	—	✓	85.96	83.91	86.45	82.47	87.03
6	—	✓	✓	74.16	76.87	80.45	78.25	80.93
7	✓	✓	✓	87.48	88.65	89.53	87.58	89.72

模型识别效果最好。

4.3 结合深度学习特征与传统图像特征的识别实验

在 4.2 节中对传统特征的分类效果进行了实验, 本小节中加入 CNN 模型提取的深层特征进一步展开研究。表 4 为实验的分类结果, Traditional fea-

tures 为上一小节的模型 7 (GLCM 特征+HOG 特征+颜色特征) 3 个特征的融合, 共 296 维。Fusion features 是将传统特征与深度学习特征融合并经过特征选择的特征。从实验结果看, 融合特征在 SVM 分类器中表现最出色, Acc 可以达到 92.89%, Sen 为 89.91%, Spe 为 96.06%, AUC 面积可以达到

表 4 传统特征与深度学习特征的口腔识别比较

Tab. 4 Comparison of oral recognition between traditional features and deep learning features

Method	$Acc/\%$	$Sen/\%$	$Spe/\%$	AUC/%
Traditional features+GBDT	87.58	84.25	89.24	93.23
Traditional features+RF	89.53	87.42	91.16	94.89
Traditional features+KNN	87.48	88.76	92.03	95.04
Traditional features+Xgboost	89.72	86.22	92.18	95.25
Traditional features+SVM	88.65	86.67	92.40	96.02
EfficientNet	89.35	87.48	89.79	96.79
Fusion features+GBDT	90.35	87.89	92.20	94.66
Fusion features+RF	91.68	90.12	91.35	96.78
Fusion features+KNN	89.81	91.87	92.58	95.54
Fusion features+Xgboost	90.88	87.58	89.47	95.39
Fusion features+SVM	92.89	89.91	96.06	98.09

98.09%。与传统特征分类相比,融合了 CNN 特征的分类方法的各评价指标均有显著提高,证明 EfficientNet 模型提取了能有效反映病灶区域信息的深层次特征,对传统手工设计的浅层特征进行了有效的补充,从而进一步提高了 4 种口腔疾病的分类精度。同时,融合特征的分类器准确率均比 EfficientNet 要高,表明传统特征同样对 CNN 特征进行了有效补充,提高了网络的识别准确率。值得注意的是,使用融合特征连接 SVM 分类器的方法在大部分指标中均取得了最好的成绩,但是在 *Sen* 指标中, KNN 取得了最高的 91.87%,比 SVM 分类器高出 1.96%。而 SVM 的其他各项指标均比 KNN 高,在 *Acc*、*Spe* 和 *AUC* 中分别比 KNN 高 3.08%、3.48% 和 2.55%,因此从整体来说 SVM 的效果更好。

4.4 经典卷积神经网络对比实验

为了进一步评估所提方法的有效性,在相同数据集划分情况下,使用 ResNet50 和 DenseNet121 进行实验对比,与 EfficientNet 网络相同,对顶层进行相同重构,利用不同的神经网络提取 512 维深层特征,结合底层特征进行分类实验,实验中,组合特征所用分类器均为在上一实验中效果最好的 SVM 分类器。图 5 为不同神经网络的分类准确率和 Loss。

从图 5 可以看出, EfficientNet 获得了最高的 *Acc* 89.35%,分别比 ResNet50 和 DenseNet121 高出 1.53% 和 3.86%,在 Loss 上, EfficientNet 同样下降速度是最快的,模型最快达到收敛。以上实验验证了 EfficientNet 网络的分类能力更出色。结合底层特征后实验结果如表 5 所示。

表 5 不同卷积神经网络的口腔识别

Tab. 5 Oral recognition of different convolutional neural networks

Method	<i>Acc</i> /%	<i>Sen</i> /%	<i>Spe</i> /%	<i>AUC</i> /%
DenseNet	85.49	87.45	91.98	95.43
DenseNet+features	89.02	88.59	92.76	96.21
ResNet	87.82	85.89	92.89	96.06
ResNet+features	90.42	88.27	90.25	96.93
EfficientNet	89.35	87.48	89.79	96.79
EfficientNet+features	92.89	89.91	96.06	98.09

由表 5 可得,不同网络组合底层特征后各项指标均有提升, EfficientNet 仍然在各项指标中取得了最高值。实验再次验证 EfficientNet 网络的分类能力以及提取特征的能力十分出色,同时也验证了传统特征与深度学习特征融合识别方法的有效性。

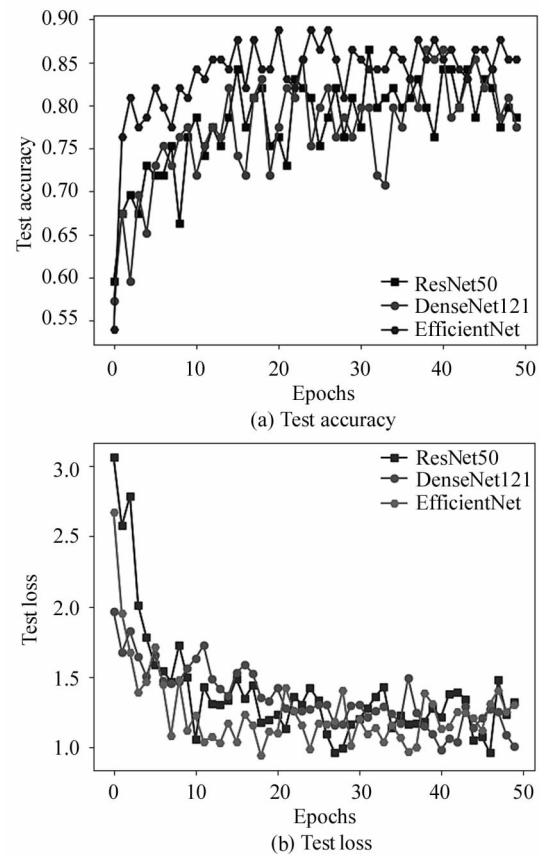


图 5 不同神经网络的分类准确率和 Loss

Fig. 5 Classification accuracy and Loss of different neural networks

5 结 论

本文提出一种基于多层次特征融合的口腔粘膜性疾病识别方法,并使用第四军医大学口腔医院收集的口腔粘膜性疾病数据进行验证。所提方法首先提取了纹理、形状和颜色的底层特征共 296 维,使用效果较好的 EfficientNet 网络提取深层特征 512 维。将两类特征融合后使用 RF 算法进行特征选择,筛选出特征重要性高的特征,减少了大量冗余特征,并且使用多种分类器进行实验验证。实验结果表明, SVM 分类器效果最好,在 *Acc*、*Sen*、*Spe* 和 *AUC* 指标中分别取得 92.89%、89.91%、96.06% 以及 98.09%。并且与多种经典神经网络展开对比实验,验证了对于口腔粘膜性疾病, EfficientNet 模型的分类效果最好,提取特征的能力最出色,同时再次验证本文所提方法的有效性。通过实验证明将深层次特征和浅层次特征进行融合使用大大丰富了口腔特征的代表性,一方面特征融合既解决了传统特征不能充分反映病灶区域信息的问题,又对神经网络提取的特征进行了有效补充,得到了分类能力更强的特

征集。另一方面对当下小样本医学数据的识别提供了新的思路,弥补了数据量不充足的情况下模型分类效果不好的劣势。

未来的研究工作主要是加入更多类别的口腔疾病,不断扩充口腔图像数据,进一步探究方法的可行性与泛化能力。在特征提取阶段,尝试加入更多的传统特征,如 LBP 特征、面积、周长和偏心率等几何特征,提高口腔疾病的识别准确率。在特征选择阶段,尝试 PCA 或者基于相关性的特征子集筛选算法。在分类阶段,本实验所用的分类器均设置为默认参数,在未来需要进一步探究默认参数对实验结果的影响。

参考文献:

- [1] NAWANDHAR A, KUMAR N, VEENA R, et al. Stratified squamous epithelial biopsy image classifier using machine learning and neighborhood feature selection[J]. Biomedical Signal Processing and Control, 2020, 55: 101671.
- [2] RAHMAN T Y, MAHANTA L B, DAS A K, et al. Automated oral squamous cell carcinoma identification using shape, texture and color features of whole image strips[J]. Tissue and Cell, 2020, 63: 101322.
- [3] SULTAN A S, ELGHARIB M A, TAVARES T, et al. The use of artificial intelligence, machine learning and deep learning in oncologic histopathology[J]. Journal of Oral Pathology & Medicine, 2020, 49(9): 849-856.
- [4] HWANG J J, JUNG Y H, CHO B H, et al. An overview of deep learning in the field of dentistry[J]. Imaging Science in Dentistry, 2019, 49(1): 1-7.
- [5] RAMLINGAM A, AURCHANA P, DHANALAKSHMI P, et al. Analysis of oral squamous cell carcinoma into various stages using pre-trained convolutional neural networks[J]. IOP Conference Series: Materials Science and Engineering, 2020, 993(1): 012058.
- [6] PALASKAR R, VYAS R, KHEDEKAR V, et al. Transfer learning for oral cancer detection using microscopic images[EB/OL]. (2020-11-23)[2021-12-24]. <https://arxiv.org/abs/2011.11610>.
- [7] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2021-12-24]. <https://arxiv.org/abs/1704.04861>.
- [8] DAS N, HUSSAIN E, MAHANTA L B. Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network[J]. Neural Networks, 2020, 128: 47-60.
- [9] IWENDI C, BASHIR A K, PESHKAR A, et al. COVID-19 patient health prediction using boosted random forest algorithm[J]. Frontiers in Public Health, 2020, 8: 357.
- [10] GARG M, DHIMAN G. A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants[J]. Neural Computing and Applications, 2021, 33: 1311-1328.
- [11] MAHMOUD H, ABDEL-NASSER M, OMER O A. Computer aided diagnosis system for skin lesions detection using texture analysis methods[C]//2018 International Conference on Innovative Trends in Computer Engineering (IT-CE), February 19-21, 2018, Aswan, Egypt. New York: IEEE, 2018: 140-144.
- [12] ZHANG T, ZHANG X, KE X, et al. HOG-ShipCLSNet: A novel deep learning network with HOG feature fusion for SAR ship classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-22.
- [13] PRAGNA D P, DANDU S, MEENAKZHI M, et al. Health alert system to detect oral cancer[C]//2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), March 10-11, 2017, Coimbatore, India. New York: IEEE, 2017: 258-262.
- [14] DE JESUS L C M, MITAL M E G, LAUGUICO S C, et al. Throat detection and health classification using neural network[C]//2019 International Conference on Contemporary Computing and Informatics (IC3I), December 12-14, 2019, Singapore, New York: IEEE, 2019: 38-43.
- [15] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[EB/OL]. (2020-09-11)[2021-12-24]. <https://arxiv.org/abs/1905.11946>.
- [16] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7132-7141.
- [17] WANG J, LIU Q, XIE H, et al. Boosted efficientnet: detection of lymph node metastases in breast cancer using convolutional neural networks[J]. Cancers, 2021, 13(4): 661.
- [18] MARQUES G, AGARWAL D, DE LA TORRE D I. Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network[J]. Applied Soft Computing, 2020, 96: 106691.

作者简介:

刘 青 (1968—),男,博士,教授,硕士生导师,主要从事医学图像处理、口腔粘膜疾病诊断方面的研究。