

doi:10.3969/j.issn.1003-3114.2025.05.012

引用格式:白世强,蔡益人,李丽宏.基于伽马变换与小波卷积的皮肤癌精确检测[J].无线电通信技术,2025,51(5):993-1007.[BAI Shiqiang, CAI Yiren, LI Lihong. Precise Skin Cancer Detection Based on Gamma Transform and Wavelet Convolution [J]. Radio Communications Technology, 2025, 51(5): 993-1007.]

基于伽马变换与小波卷积的皮肤癌精确检测

白世强,蔡益人,李丽宏*

(河北工程大学信息与电气工程学院,河北邯郸056038)

摘要:皮肤癌和黑色素细胞痣的相似性较高,极易被皮肤科医师误诊。为提高早期皮肤癌患者的筛查准确度,基于伽马变换(Gamma Transform, GT)算法和小波变换(Wavelet Transform, WT)算法,分别构建伽马变换模块(Gamma Transform Block, GMTB)与小波卷积模块(Wavelet Convolution Block, WTCB),并在Detection Transformer(DETR)架构基础上创新性地提出用于捕捉皮肤癌细粒度特征的空频变换网络(Space-Frequency Transform Network, SFTNet)。包含SFTNet的皮肤癌筛选系统能够增强样本图像的不同通道,减少模型训练过程中的过拟合效应,进而有效提升疾病检测精度。基于HAM10000数据集的仿真实验结果表明,系统正确率达85.5%,在皮肤癌辅助诊断方面具有重要的临床应用价值。

关键词:深度学习;医学影像处理;皮肤癌检测;伽马变换;小波卷积;空频变换网络

中图分类号:TP181;TN911.7

文献标志码:A

开放科学(资源服务)标识码(OSID):

文章编号:1003-3114(2025)05-0993-15



Precise Skin Cancer Detection Based on Gamma Transform and Wavelet Convolution

BAI Shiqiang, CAI Yiren, LI Lihong*

(College of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, China)

Abstract: Skin cancer and melanocytic nevus share numerous similarities, which can result in a misdiagnosis by dermatologists. To improve the screening accuracy of early skin cancer patients, the Gamma Transform Block (GMTB) based on Gamma Transform (GT) and Wavelet Convolution Block (WTCB) based on Wavelet Transform (WT) are proposed. Furthermore, the Space-Frequency Transform Network (SFTNet) for capturing fine-grained features of skin cancer is innovatively proposed based on the Detection Transformer (DETR) architecture. SFTNet-based skin cancer screening system can effectively improve disease detection accuracy because it enhances the sample image at different channels and reduces over-fitting effect during the model training process. Simulation results on HAM10000 dataset show that the accuracy of this system can reach 85.5%, which underscores the significant clinical value of our approach in skin cancer assisted diagnosis.

Keywords: deep learning; medical image processing; skin cancer detection; GT; wavelet convolution; SFTNet

0 引言

近年来,因紫外线灼伤或化妆品使用不当导致的皮肤癌患者数量不断增加。相较于常见的肺癌和肝

癌,皮肤癌更容易被群众忽视,最终因患者发现不及时而错过最佳治疗期^[1]。相较于普通的体格检查,皮肤镜具有高精度的优点,正在逐步成为皮肤癌筛查的重要方式。由于我国皮肤性病科医师数量不足,与日俱增的皮肤镜检查和海量的结果报告往往让医师捉襟见肘。本文以深度学习领域中的目标检测技术为核心,提出基于GT与小波卷积的皮肤癌检测系统,为皮肤镜检查结果报告不及时的问题提供新的解决方案。

收稿日期:2025-04-22

基金项目:河北省自然科学基金(F2023402011)

Foundation Item: Hebei Provincial Natural Science Foundation of China (F2023402011)

1 相关工作

皮肤性病科的检查项目通常包含体格检查、组织病理学检查、皮肤影像学检查、实验室诊断、变应原检测等^[2]。皮肤镜又称皮表透光显微镜,是一种利用光学技术深入观测皮肤角质层以下结构的仪器。同其他影像学检查一样,皮肤镜检查不会对患者机体产生侵入性损伤,因而在众多方案中被优先考虑。文献[3-6]的研究结果表明,皮肤镜在非黑色素瘤皮肤癌、基底细胞癌、黑素细胞痣、脂溢性角化、日光性角化等常见皮肤疾病的临床诊断中具有重要的应用价值。

皮肤镜检查的结果通常以图像形式呈现。因此,对患者的皮肤镜检查结果进行预筛分析的过程可以建模为一个图像检测系统。在深度学习领域中,现有的目标检测技术可大致分为2类。其中一类是基于卷积神经网络(Convolutional Neural Network, CNN)的目标检测技术,该类的典型代表有单发多框检测(Single Shot Multibox Detector, SSD)模型^[7]、区域卷积神经网络(Region-based Convolutional Neural Network, R-CNN)系列模型^[8-10]以及 You Only Look Once (YOLO)系列模型^[11]。另一类则是基于注意力机制和 Transformer^[12],典型代表为 Facebook AI Research 团队^[13]所提出的 DETR 架构(如图1所示)及其衍生模型。上述各方法的主要优缺点如表1所示。

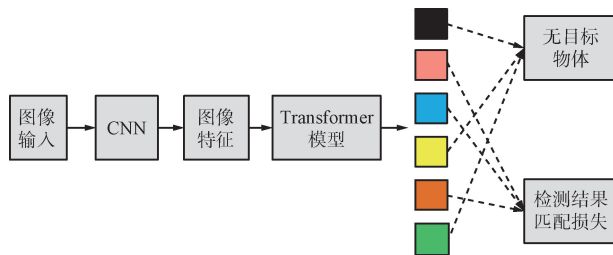


图1 DETR原理

Fig. 1 Principle of DETR



图2 图像的频域、空间域和复频域之间的关系

Fig. 2 Relationship between frequency domain, spatial domain and complex frequency domain of an image

本文主要使用 GT、空间滤波(Spatial Filtering, SF)、WT 三种技术。

2.1 空间域——GT 与 SF

空间域是一幅图像本身所处的表达域,一般

表1 现有目标检测方法的主要优缺点对比

Tab. 1 Comparison of the main advantages and disadvantages of existing target detection methods

技术方案	主要优点	主要缺点
CNN	结构简单	检测性能不佳
SSD	速度与精度的均衡	小目标检测不灵敏
R-CNN	精度较高	训练过程缓慢
YOLO	快速、实时检测	目标数量受限,精度较差
DETR	端到端检测	注意力机制消耗过量成本

当需要检测的每张图像样本中仅考虑一个类别时,可以将其简化为图像分类问题。虽然皮肤镜检查的结果可能显示某患者患有多种皮肤疾病,若仅设定其中一种关键疾病作为检测对象,也能够通过图像分类算法完成既定目标。

现有的智能医学影像检测技术中,多以 CNN 为结构核心。赵嘉晖等^[14]提出基于 CCNet 注意力机制和高维多目标优化算法的小样本皮肤癌检测模型,并在 ISIC2018、Derm7pt 数据集上取得较好效果。王煜坤^[15]在博士学位论文中同样利用 CNN 进行皮肤疾病的智能诊断。与前二者不同, Tschandl 等^[16]采用 CNN 与皮肤科医师协作的方法进行皮肤癌检测,进一步确保诊断结果的可靠性。

虽然上述研究在智能皮肤病识别领域都有明显突破,但均未能有效地解决模型过拟合问题,也未弥补降采样操作所带来的信息损失。因此,为克服现有技术的缺陷,提出 GMTB 和 WTCB,并选取性能更优秀的 DETR 作为基础架构,以对皮肤癌患者进行更精确的筛选。

2 GT 与 WT

一幅图像可以被表示为空间域、频域、复频域3种基本形式,其中的频域和复频域均属于变换域。3种表示方法之间的关系如图2所示。

用于直接对像素进行处理。空间域图像处理技术主要包括直方图增强(Histogram Enhancement, HE)、灰度变换(Grey Level Transform, GLT)、SF 三种^[17]。

HE 主要指通过使用均衡化等方法改变灰度图像原有的灰度值分布规律,以增加或减少灰度图像的对比度。

GLT 是一种通过对灰度图像中像素点的灰度值进行特定的函数运算,以调整并增强图像可视化效果的技术。设一幅灰度图像的空间域函数表达式是 $f(x, y)$, 变换关系函数或变换算子是 $s = T(r)$, 对其进行 GLT 后的表达式为^[17]:

$$g(x, y) = T[f(x, y)]。 \quad (1)$$

GLT 常使用线性函数、对数函数、幂函数 3 种变换关系。基于幂函数的幂变换算法又称 GT, 其函数式为:

$$s = cr^\gamma, \quad (2)$$

式中: r 为输入灰度值, γ 为变换参数, s 为变换后的灰度值, c 为常系数。当伽马参数小于 1 时, 拉伸图像灰度直方图的左端并压缩其右端; 当伽马参数大于 1 时, 压缩图像灰度直方图的左端并拉伸其右端; 当伽马参数等于 1 时, GT 则退化为线性变换^[18]。GLT 的本质是对每个空间像素进行灰度调整。

SF 是指在图像的空间域进行矩阵乘积运算, 通过把每个像素点替换为该像素及其邻域像素的函数值来增强或抑制图像中的某个特定成分。设一幅数字图像在空间域的表达式为 $f(x, y)$, 空间滤波器的核为 $k(x, y)$, 则该图像的空间域卷积式为:

$$f(x, y) * k(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b k(s, t) f(x-s, y-t)。 \quad (3)$$

若对卷积核 k 进行反转, 可得到如下 SF 表达式:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b k(s, t) f(x+s, y+t)。 \quad (4)$$

SF 的本质是对图像像素的邻域进行空间卷积运算。

2.2 频率域——WT

图像频率域处理技术的核心思想在于, 使用傅里叶变换或拉普拉斯变换将原始图像由基本的空间域转变至对应的变换域, 在频率域或复频域进行相应的处理后, 再将处理结果反变换至空间域, 以完成既定目标。傅里叶变换^[19]以及后续的余弦变换^[20]、斜变换^[21]、沃尔什-哈达玛变换^[22]、WT^[23]等方法均属于线性变换。

当一维连续函数 $f(t)$ 的傅里叶变换 $F(w)$ 满足条件 $C = \int_{\mathbb{R}} \frac{|F(w)|^2}{|w|} dw < \infty$ 时, 可通过 WT 将已

知信号分解为小基波和尺度函数的线性组合。连续 WT 及其逆变换为:

$$W_f(a, b) = \langle f(t), \varphi_{a,b} \rangle = \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(t) F\left(\frac{t-b}{a}\right) dt, \quad (5)$$

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{a^2} W_f(a, b) \psi\left(\frac{t-b}{a}\right) da db, \quad (6)$$

式中: a 为伸缩因子, b 为平移因子。在此基础上取:

$$a = a_0^j, \quad (7)$$

$$b = ca_0^j b_0 (j \in \mathbb{Z}), \quad (8)$$

可得到函数 $f(t)$ 的一维离散小波变换 (Discrete Wavelet Transform, DWT) 系数。其表达式为:

$$T_\varphi(u, v) = T_\varphi(u+1, n) * h_\varphi(-n) = \sum_n h_\varphi(n-2v) T_\varphi(u+1, n), \quad (9)$$

$$T_\psi(u, v) = T_\psi(u+1, n) * h_\psi(-n) = \sum_n h_\psi(n-2v) T_\psi(u+1, n)。 \quad (10)$$

作为时频分析领域的基础理论之一, WT 在现代信号处理中发挥着不可替代的作用。

3 基于 SFTNet 的皮肤癌筛选系统

3.1 GMTB

已知患者的皮肤镜检查结果为 RGB 格式图像。设该彩色数字图像宽为 W pixel, 高为 H pixel, 像素点在图像中所对应的位置坐标为 (x, y) 。

读取患者的皮肤镜检查结果, 对彩色图像进行通道分解, 得到 R 通道、G 通道、B 通道。设 (x, y) 在 3 个通道中所对应的灰度值大小分别为 $r(x, y)$ 、 $g(x, y)$ 、 $b(x, y)$, 且单个通道中的灰度最大值是 M , 灰度均值和灰度方差分别是 m 和 s , 对 3 个通道分别进行灰度归一化和图像标准化, 可得:

$$r'(x, y) = \frac{\frac{r(x, y)}{M} - m}{s}, \quad (11)$$

$$g'(x, y) = \frac{\frac{g(x, y)}{M} - m}{s}, \quad (12)$$

$$b'(x, y) = \frac{\frac{b(x, y)}{M} - m}{s}。 \quad (13)$$

设定伽马参数, 使用幂函数对标准化后的灰度图像进行逐像素 GT, 依次得到:

$$R(x, y) = C[r'(x, y)]^\gamma, \quad (14)$$

$$G(x, y) = C[g'(x, y)]^\gamma, \quad (15)$$

$$B(x, y) = C[b'(x, y)]^\gamma。 \quad (16)$$

最终将变换后的 3 个通道合并, 作为 GMTB 的输出。

GMTB 的结构如图 3 所示。

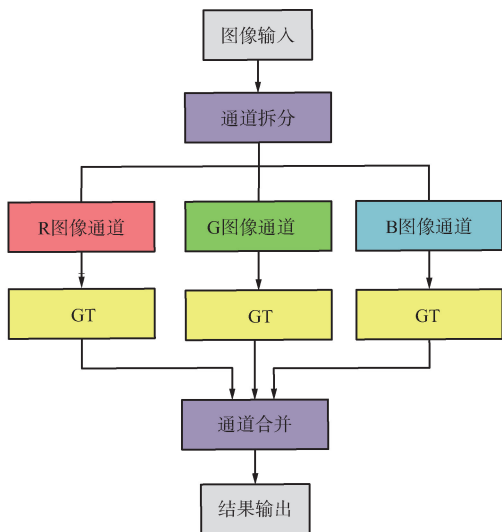


图 3 GMTB 结构示意图
Fig. 3 Structure of GMTB

GMTB 主要由通道拆分、单通道 GT、通道整合 3 部分组成。设模块输入一幅大小是 $w \times h$ 的彩色皮肤镜图像,张量形状为 $3 \times w \times h$ 。在通道拆分过程中,仅使用赋值语句提取张量中的一个维度;在通道整合过程中,只完成张量拼接,并未进行实质性运算。因此,GMTB 的计算复杂度主要由单通道 GT 阶段所决定。

在模块的 GT 阶段,需根据式(2)分别对输入图像的 3 个通道进行逐像素变换计算,GT 阶段的时间复杂度为 $O(w \times h)$ 。同时,该过程会对输入图像与变换输出结果加以保存,故 GT 阶段的空间复杂度同样为 $O(w \times h)$ 。

综上,GMTB 的时间复杂度与空间复杂度均为 $O(w \times h)$ 。

3.2 WTCB

设输入皮肤镜图像中某个通道的灰度函数为 $f(x, y)$ 。通过指定一维函数 φ 和 ψ ,并在 x 轴与 y 轴方向上进行乘积,可得二维尺度函数 $\varphi(x, y)$ 、二维纵向小波 $\Psi^H(x, y)$ 、二维横向小波 $\Psi^V(x, y)$ 、二维对角小波 $\Psi^D(x, y)$ 。

由于上述 4 个二维函数均可分离,故能够采用快速离散小波变换(Fast Discrete Wavelet Transform, FDWT)算法进行求解。具体运算步骤如下:分别对一维函数 φ 和 ψ 构建基函数 h_φ 和 h_ψ ,使用 h_φ 和 h_ψ 对输入图像 $f(x, y)$ 逐行求解一维卷积,并逐列进行降采样,进而得到 2 幅水平方向上具有 1/2 分辨率的子图像。再次使用 h_φ 和 h_ψ 基函数对上述 2 个子图像分别进行逐列一维卷积,并逐行进行降采样计算,最终输出 4 幅长宽均为原图像 1/2 的子图 T_D 、 T_V 、 T_H 、 T 。其中, T_D 包含对角线细节系数, T_V 包含垂直细节系数, T_H 包含水平细节系数, T 包含图像的原始低频信息。二维 FDWT 的计算步骤如图 4 所示。

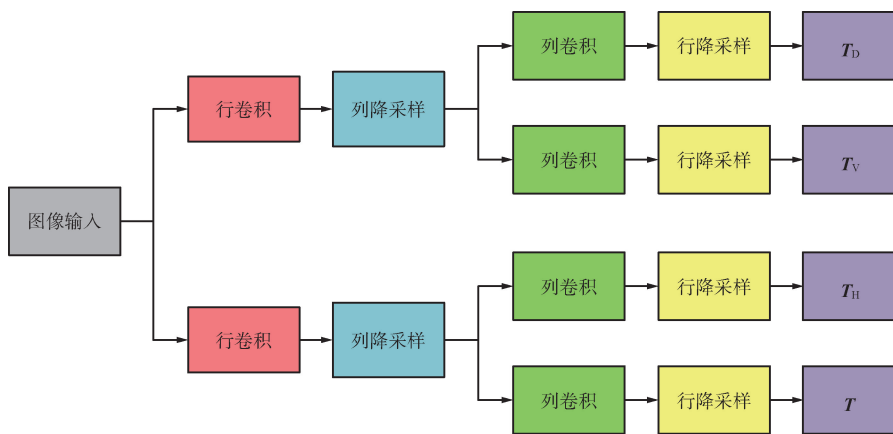


图 4 对图像进行二维 FDWT 的计算方法
Fig. 4 Calculation method of two-dimensional FDWT for image

计算二维快速离散小波逆变换(Fast Inverse Discrete Wavelet Transform, FIDWT)时,与二维 FDWT 的流程反向对应。其具体计算步骤为:对输入的 T_D 、 T_V 、 T_H 、 T 四幅图像逐行进行升采样计算,分别使用 h_φ 、 h_ψ 、 h_φ 、 h_ψ 基函数对 4 个升采

样结果子图逐列计算卷积,将结果分 2 组并各自相加。对相加后的 2 个图像分别按列进行升采样计算,并使用 h_φ 、 h_ψ 基函数逐行卷积,最终将卷积后的结果相加并输出。二维 FIDWT 的计算步骤如图 5 所示。

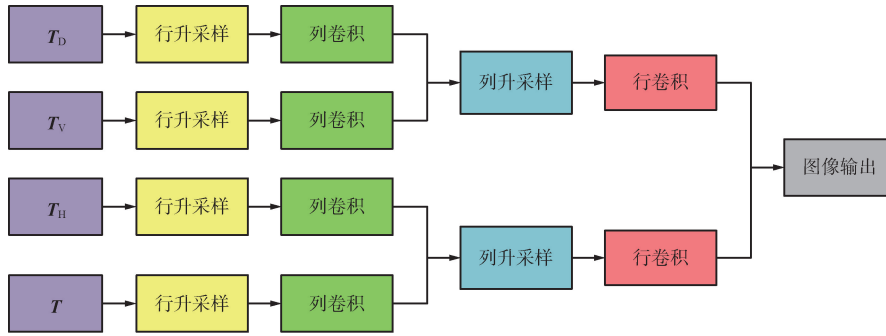


图5 对图像进行二维 FIDWT 的计算方法

Fig. 5 Calculation method of two-dimensional FIDWT for image

WTCB 主要使用小波卷积和 SF 两种技术。在 WTCB 之前,先对输入的皮肤镜彩色图像进行通道分解,得到 R、G、B 三个通道。选取单个通道作为原始输入,并对输入图像计算二维 FDWT,得到低频-低频图像(LL)、低频-高频图像(LH)、高频-低频图像(HL)、高频-高频图像(HH)共 4 幅子图。在此基础上,分别对 LL、LH、HL、HH 四幅子图进行 SF,并对滤波结果计算二维 FIDWT,以获得包含丰富空频信息的图像特征。最终将小波逆变换结果与原始输入图像的卷积结果相加,作为 WTCB 的输出。WTCB 的结构如图 6 所示。

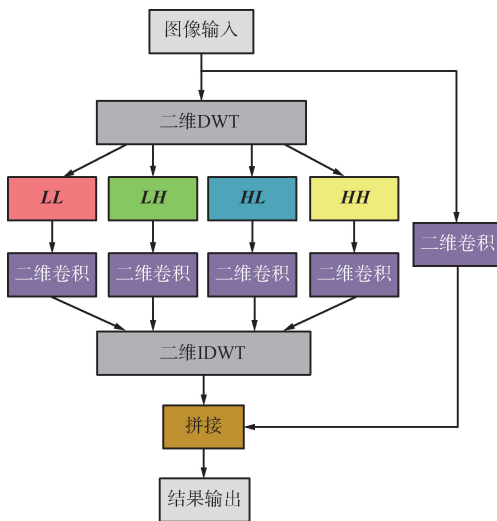


图6 WTCB 的结构示意

Fig. 6 Structure of WTCB

WTCB 主要由二维 DWT、二维空间域卷积、二维 IDWT 三部分组成。设输入为一幅 $w \times h$ 大小的彩色皮肤镜图像,其张量封装形状是 $3 \times w \times h$ 。其中,对该图像计算二维 DWT 或二维 IDWT 时,完成一层小波分解所需要的时间复杂度可表示为 $O(w \times h)$ 。若二维卷积操作的卷积核大小是 $k \times k$,对 DWT 所输出

的 4 幅子图分别进行单通道卷积的计算量可表示为 $k \times k \times \frac{w}{2} \times \frac{h}{2}$;对原始输入图像进行单通道卷积的计算量可表示为 $k \times k \times w \times h$ 。将 IDWT 的输出图像和原图像的空间卷积结果相加,该步骤的复杂度同为 $O(w \times h)$ 。因此,WTCB 的时间复杂度可表示为 $O(k \times k \times w \times h)$ 。

由于二维 DWT 会令输出图像的长宽均缩小至原图像大小的 $1/2$,故 LL、LH、HL、HH 四幅小波子图所占用的存储空间是 $\frac{w}{2} \times \frac{h}{2}$ 。除此之外,原始输入图像、IDWT 输出图像、原图像的空间卷积输出以及模块最终结果的矩阵大小均为 $w \times h$,即 WTCB 的空间复杂度是 $O(w \times h)$ 。

综上所述,WTCB 的时间复杂度是 $O(k \times k \times w \times h)$,而空间复杂度是 $O(w \times h)$ 。

3.3 智能皮肤镜图像处理系统

选取 DETR 作为皮肤镜图像处理系统的基础架构,并在网络中使用 GMTB 和 WTCB 进行特征提取。该系统的图像处理流程表述如下:

① 获取 HAM10000 皮肤镜数据集,将其划分为训练集、验证集、测试集 3 个部分,并依次读取至计算机内存中。

② 将皮肤镜影像输入系统,对图像进行滤波、裁剪、标准化等操作,并封装为张量格式。

③ 使用 GMTB 分别调整 R、G、B 三个通道的灰度,并利用包含 WTCB 的主干网络对图像特征加以提取。

④ 利用图像的高度参数 h 和宽度参数 w 对输出特征进行位置编码,把原始图像与编码结果输入至改进后的 Transformer 结构。

⑤ 把 Transformer 的输出传至 2 个线性层,分别预测该患者的皮肤病类型和位置锚框。

⑥ 以可视化方式输出系统的检测结果。

完整的皮肤镜图像处理系统模型如图 7 所示。

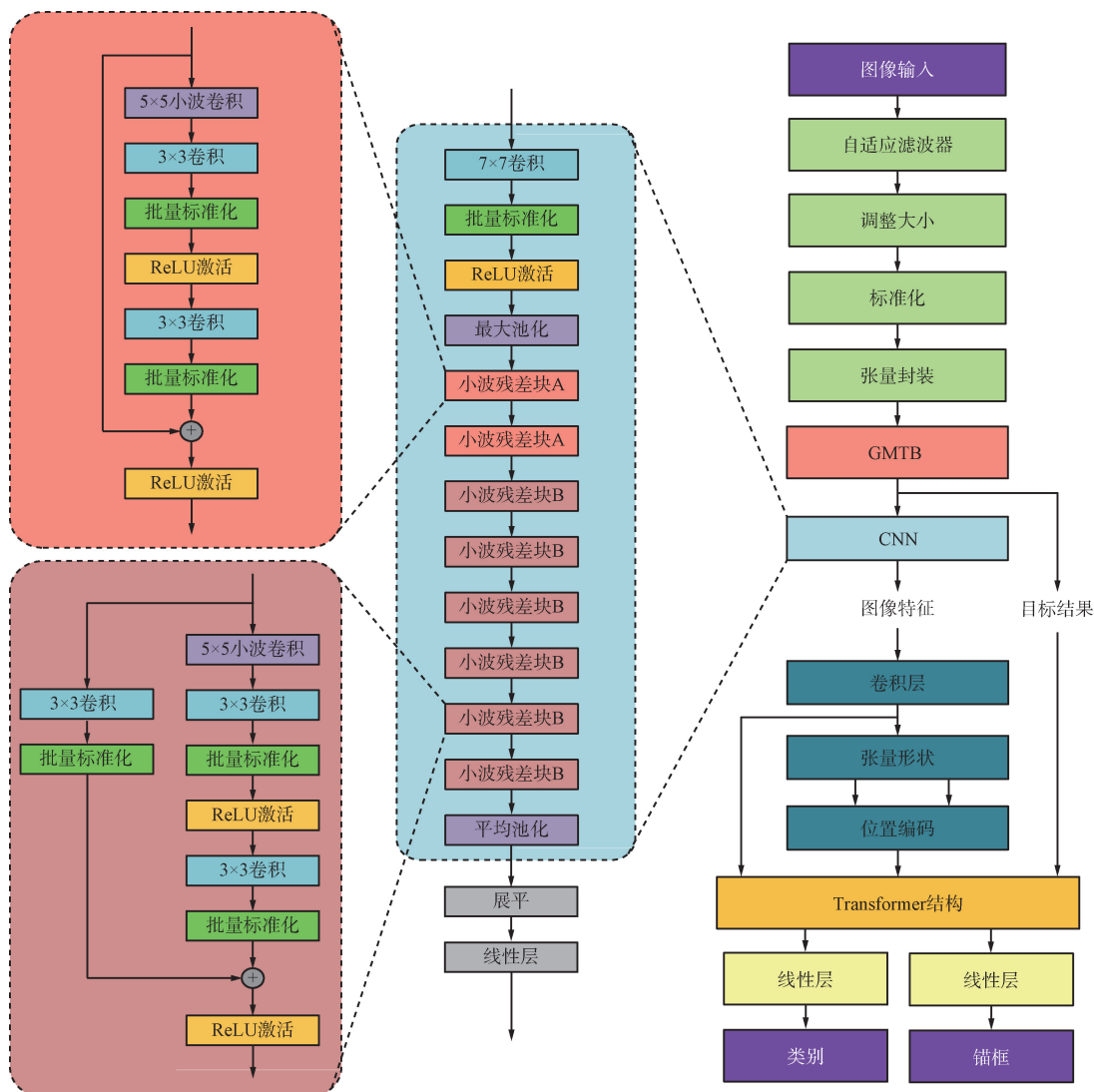


图 7 皮肤镜图像处理系统模型结构

Fig. 7 Model structure of dermoscopy image processing system

3.4 皮肤癌患者筛选流程

皮肤癌患者的筛选流程表述如下^[24]：

① 嘱咐患者暴露并清洗检查区域,确保皮肤表面无污垢等残留物。

② 将皮肤镜放置在患者的皮肤表面,调节镜头的焦距和放大倍数,从不同距离、不同角度观察患者皮肤的毛孔、血管等表面结构及皮下结构。

③ 使用皮肤镜对相应部位拍照或录像,记录患者皮肤的病变情况。

④ 将图像或视频数据保存至计算机中,利用智能皮肤镜图像处理系统初步判断该患者的疾病类型。

⑤ 反馈系统判别结果至诊间计算机中,辅助皮肤科医师诊断疾病。

⑥ 将筛查结果告知患者,并做出相应处理。

4 实验与讨论

4.1 实验环境与数据集

实验程序基于 PyTorch 架构编写,并在具有 RTX3080 显卡和 64 GB 内存的人工智能服务器上运行与仿真。为保证实验结果可复现,使用公开的 HAM10000 皮肤镜数据集^[25]进行训练、验证、测试。该数据集共由 10 015 张皮肤镜图片组成,其中训练集包含 7 007 张图片,占比 70%;验证集包含

1 003 张图片,占比 10%;测试集包含 2 005 张图片,占比 20%。HAM10000 皮肤镜数据集的划分方法如图 8 所示。

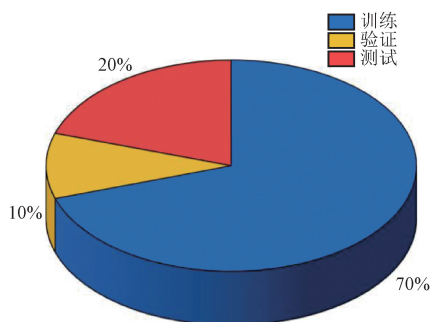
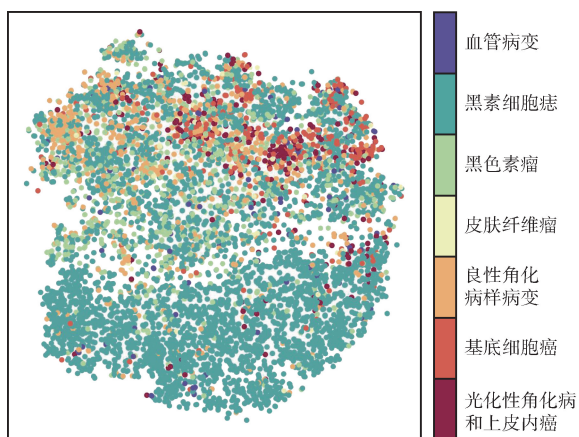
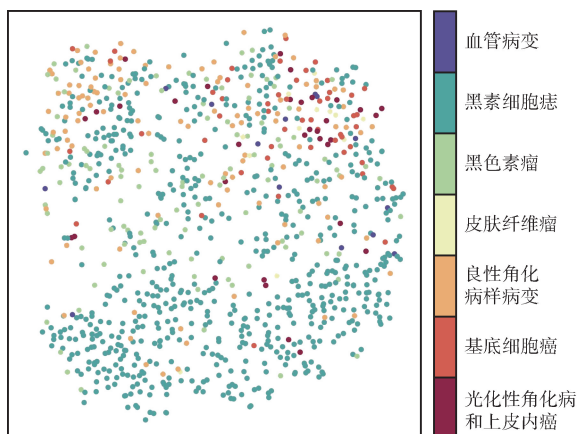


图 8 HAM10000 皮肤镜数据集的划分方法
Fig. 8 HAM10000 dermoscopy dataset division

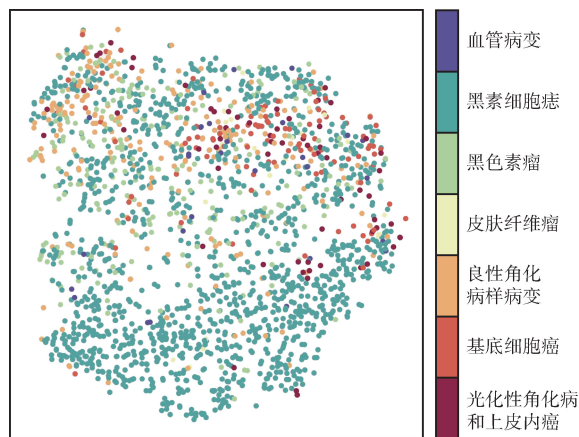
使用 t-SNE 降维聚类算法^[26]分别绘制训练集、验证集、测试集的二维散点图,结果如图 9 所示。



(a) 训练集的降维散点



(b) 验证集的降维散点



(c) 测试集的降维散点

图 9 训练集、验证集、测试集的降维散点

Fig. 9 Dimensionality reduction scatter plots of training dataset, validation dataset and test dataset

由图 9 可知:在 HAM10000 数据集中,黑色素细胞痣(图中浅蓝色点)的样本数量最多,且在整个平面内均有分布;暖色系所代表的光化性角化病和上皮内癌、基底细胞癌、良性角化病变的样本点则普遍集中于 $y > 0$ 平面区域,且与黑色素细胞痣的样本点相互混杂。上述情况表明,在二维情况下,癌症图像样本与细胞痣图像样本之间相似性较大,这极大地增加了识别皮肤疾病的难度。

4.2 理论定性分析

人类皮肤由表皮、真皮和皮下组织 3 个部分构成,皮肤中黑色素的含量和分布直接决定皮肤的颜色^[2]。在皮肤性病科的临床实践中,皮肤颜色会因患者种族、年龄以及发病部位的不同而产生差异,但皮肤镜检查结果主要呈现为白色、粉色、红色和黑色。因此,可以直接对皮肤镜影像使用彩色图像处理技术,或将彩色图像拆为 R、G、B 三个通道分别进行灰度图像处理。

疾病的病理学形态是进行临床诊断的重要依据。黑色素细胞痣具有清晰的边界,多呈现条状或巢状;基底细胞癌的边界不规则,同样呈现为条状或巢状;黑色素瘤的边界不清晰,且细胞形态多样,具有显著差异。因此,通过仔细观察病变部位的边界与纹理,可有效区分黑色素细胞痣、基底细胞癌、黑色素瘤。

GT 是一种灰度值变换算法,通过设定不同的伽马参数值,能够对图像中明暗部分进行增强或抑制,从而提升图像整体的对比度。由于患病部位与正常部位的颜色相近,仅使用肉眼难以区分颜色的细微

变化,这给皮肤科医师带来巨大困扰。GT在有效增加二者对比度的同时,也更加突出病变区域的边缘,故采用GMTB更有利于识别边界不规则的基底细胞癌及黑色素瘤。

当指数位置的伽马参数满足条件 $\gamma < 1$ 时,可以提高图像中较暗区域的对比度,同时降低较亮区域的对比度,主要用于增大灰度级偏暗图像的灰度动态范围;当伽马参数满足条件 $\gamma > 1$ 时,能够提高图像中较亮区域的对比度,降低较暗区域的对比度,进而增加整体偏亮图像的灰度动态范围^[18]。皮肤镜图像的关键区域是黑色的细胞痣或细胞癌,因此设定 $\gamma = 0.9$ 有利于改善图像中目标与背景的对比如果。与此同时,图像产生的颜色变化也可认为是对原数据集进行数据增广,以减小模型训练过程中的过拟合效应。

传统的CNN均采用空域卷积技术,以提取图像的空间域特征。WT则延续傅里叶变换和沃尔什-哈达玛变换的思想,其使用的小波函数在空间域和频率域均具有集中性与平滑性的特点,故WT在处理皮肤镜影像等非平稳信号时优势明显^[18]。此外,WT能够根据二维图像自适应地调整空频窗口,并获得可变的空域分辨率和频域分辨率,从而捕捉到

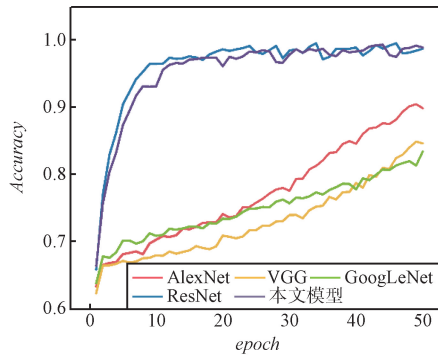
病变部位的纹理细节。在骨干网络中添加WT后,于传统的空域卷积基础之上补充了图像的频域特征,同时也能够增大卷积核感受野,提升模型的细粒度分辨能力。因此,充分利用WT的多分辨率特性可以有效地识别黑色素瘤及其他疾病。

作为目标检测领域的一个重要分支,DETR主要由CNN、Transformer、目标检测头3个部分构成。其中,包含降采样的CNN可以有效获得图像的局部特征,并减少数据量;而Transformer具有优良的全局特征提取能力,很大程度上弥补了CNN在视野受限方面的不足。因此,具有混合结构的SFTNet模型比YOLO系列模型更具性能优势。

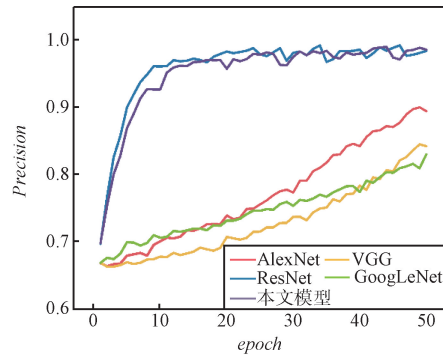
4.3 消融实验分析

4.3.1 采用不同结构的CNN

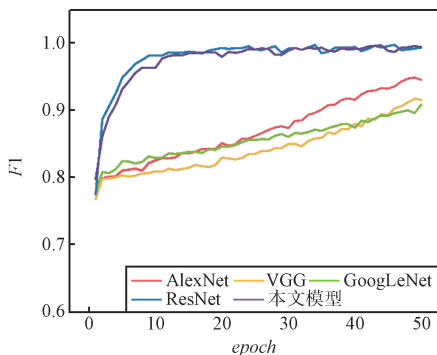
在图像处理系统中,骨干网络的关键作用是提取图像特征。而选取不当的CNN可能给系统带来严重的性能瓶颈。由于第一代CNN的LeNet-5^[27]模型无法有效地提取复杂图像特征,因此选取AlexNet^[28]、VGG^[29]、GoogLeNet^[30]、ResNet^[31]四种经典CNN作为本次消融实验的基础结构。训练与验证过程中,Accuracy、Precision、F1、AP四种模型评价指标的变化曲线如图10所示。



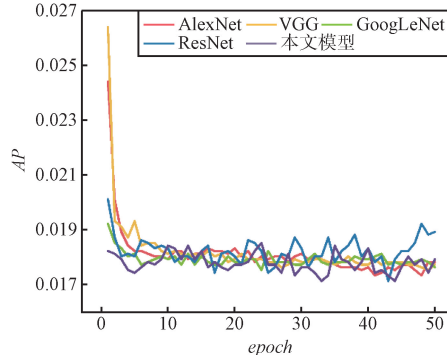
(a) 训练过程中的 Accuracy 变化曲线



(b) 训练过程中的 Precision 变化曲线



(c) 训练过程中的 F1 变化曲线



(d) 训练过程中的 AP 变化曲线

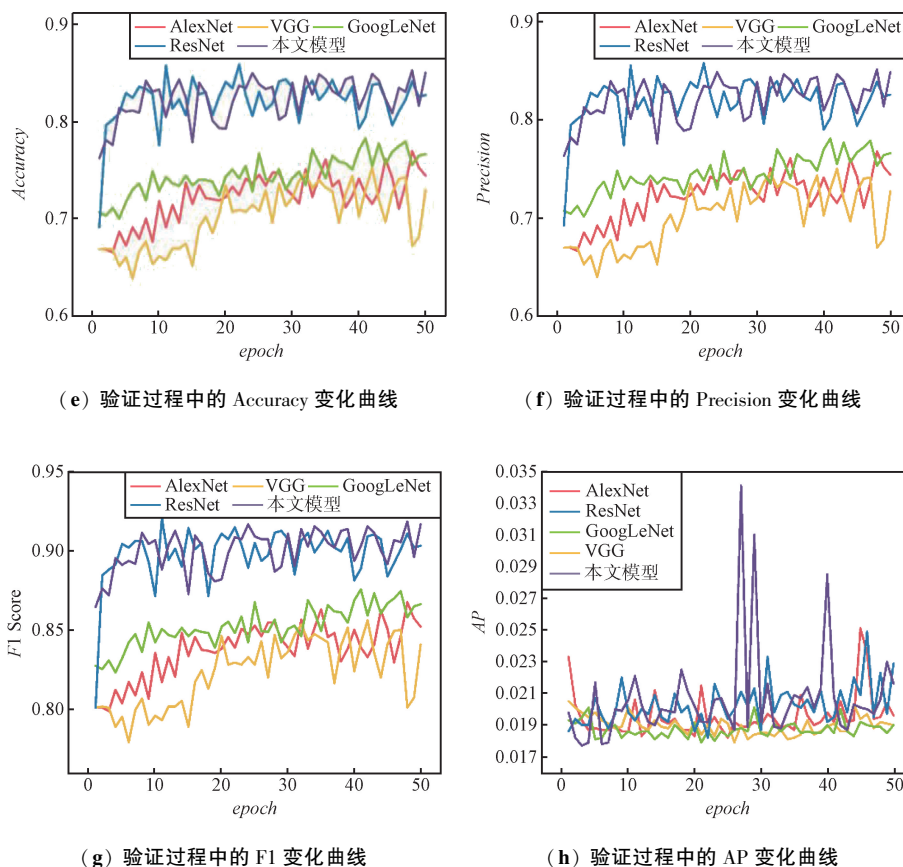


图 10 训练过程与验证过程中评价指标的变化曲线
 Fig. 10 Change curve of evaluation index in training process and verification process

如图 10(a)~图 10(c)和图 10(e)~图 10(g)所示,选取 AlexNet、VGG 和 GoogLeNet 作为系统骨干网络时,三者在训练和验证过程中的 Accuracy 曲线、Precision 曲线、F1 曲线相近,且数值均远低于以 ResNet 为骨干网络时的曲线。由此可知,当数据集与基础参数保持相同时,AlexNet、VGG、GoogLeNet 的特征提取效果相似,且 ResNet 模型的特征提取能力优于前述三者,即以 ResNet 为基础的骨干网络更适合于该系统的特征提取工作。当训练轮数 $epoch < 5$ 时,图 10(d)和图 10(h)中的 AP 曲线快速下降,但在 $epoch > 5$ 后彼此相互重叠。这表明,不论使用何种 CNN 作为基础模型,训练或验证过程中的 AP 曲线变化趋势均无明显差异。

对比 4 幅训练曲线和 4 幅验证曲线可知,相较于平缓的训练过程曲线,验证过程中的 Accuracy、Precision、F1 以及 AP 指标均产生大范围波动。产生该现象的根本原因在于,小批量随机梯度下降算法令训练损失函数的优化方向随 epoch 的变化而变化,进而导致验证过程中各评价指标的数值不稳定。

4.3.2 添加不同类型的特征提取模块

模块消融实验共分 3 组。第 1 组和第 2 组分别使用 ResNet18、ResNet34、ResNet50^[31]作为系统骨干网络,第 3 组实验中加入 ResNet101^[31]模型进行对比。3 组消融实验的各项评价指标如表 2 所示。

表 2 3 组消融实验中的各项评价指标

Tab. 2 Evaluation indices of three groups of ablation experiments

分组	模型	TP	TN	FP	Accuracy	Precision	F1	AP
1	ResNet18	1 444	28	561	0.724 1	0.720 2	0.837 3	0.020 9
	ResNet34	1 493	28	512	0.748 2	0.744 6	0.853 6	0.020 5
	ResNet50	1 341	0	664	0.668 8	0.668 8	0.801 6	0.027 5

续表

分组	模型	TP	TN	FP	Accuracy	Precision	F1	AP
2	WTCB-R18	1 495	19	510	0.748 0	0.745 6	0.854 0	0.019 0
	WTCB-R34	1 506	29	499	0.754 7	0.751 1	0.858 0	0.020 0
	WTCB-R50	1 341	0	664	0.668 8	0.668 8	0.802 0	0.021 0
3	SFTNet-R18	1 651	36	354	0.826 6	0.823 4	0.903 2	0.051 7
	SFTNet-R34	1 708	44	297	0.855 1	0.851 9	0.920 0	0.021 2
	SFTNet-R50	1 683	47	322	0.843 1	0.839 4	0.912 7	0.019 9
	SFTNet-R101	1 617	42	388	0.810 5	0.806 5	0.892 9	0.021 1

在上述3组消融实验中,以ResNet34为骨干网络时模型的各项评价指标均为最优,ResNet18次之。而以ResNet50为骨干网络时,模型的各项指标均为最差。分析原因在于,ResNet18结构的卷积层数量较少,无法有效地提取深层次特征;ResNet50结构的卷积层数量过多,导致模型出现轻度过拟合现象;第3组实验中使用的ResNet101结构则展现出更为严重的过拟合效应。因此,针对HAM10000皮肤镜图像数据集,卷积深度居于二者之间的ResNet34模型可以拥有比其他模型更优秀的特征提取效果。

相较于未做出任何改进的第1组实验,加入WTCB后,WTCB-R18和WTCB-R34模型在TP、Accuracy、Precision、F1四个关键指标上均有提升。以ResNet18为骨干网络时,4项指标的提升幅度分别达到3.53%、3.3%、3.53%、2.03%;以ResNet34为骨干网络时,模型指标的提升幅度分别为0.87%、0.87%、0.87%、0.50%;而TN、FP、AP三项指标的数值则略有下降。除了ResNet50结构出现过拟合与训练瓶颈的现象,导致WTCB-R50模型并未发生实质性改变外,在原有ResNet结构的基础之上添加WTCB可带来明显的性能提升。

在上述实验的基础之上加入GMTB,并采用迁移学习策略,进行第3组消融实验。由第3组实验结果可得,模型的TP、Accuracy、Precision、F1四项指标均有大幅提升。其中,SFTNet-R18模型的指标提升幅度为10.43%、10.51%、10.43%、5.72%;SFTNet-R34模型的提升幅度为13.41%、13.30%、13.42%、7.24%;SFTNet-R50模型的提升幅度达到25.5%、26.06%、25.51%、13.86%。对比第2组和第3组消融实验可知,在WTCB-ResNet模型中添加GMTB后,模型性能提升显著。模型性能获得有效提升的重要原因在于,GMTB改变了原图像中R、G、B三个通道的灰度值,进而增加图像对比度。

4.4 仿真实验分析

测试集混淆矩阵的二维热力图如图11所示。

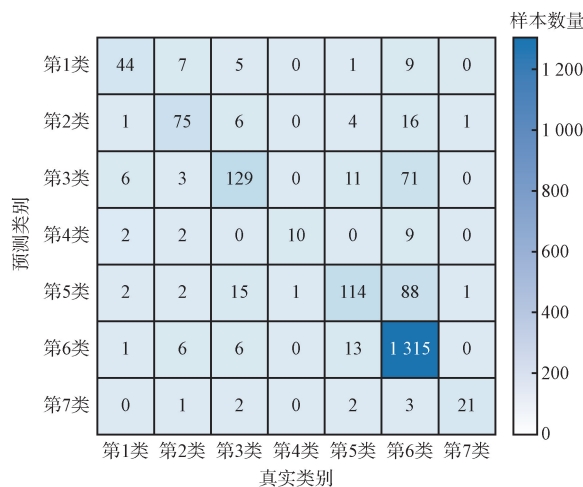


图11 测试集混淆矩阵的二维热力图

Fig. 11 Two-dimensional heat map drawn by confusion matrix of test dataset

图11中,混淆矩阵的行表示由皮肤科医师标注的真实疾病类别标签,混淆矩阵的列表示系统的诊断结果。阿拉伯数字1~7分别代表日光性角化病和原位癌、基底细胞癌、良性角化病变、皮肤纤维瘤、黑色素瘤、黑色素细胞痣、血管病变7种常见皮肤疾病。

由图11可知,混淆矩阵中各行的最大值均位于主对角线区域,表明系统的检测输出与专业皮肤科医师的标准诊断结果基本保持一致。除本类之外,各行中次高的元素值均分布在第6列,说明系统将部分患者的皮肤镜图像错误识别为黑色素细胞痣。

对各行数据分别求和并计算正确率可知,智能皮肤癌筛选系统对日光性角化病和原位癌、基底细胞癌、良性角化病变、皮肤纤维瘤、黑色素瘤、黑色素细胞痣、血管病变7种疾病的诊断正确率分别达到0.6667、0.7282、0.586、0.4348、0.5112、0.9806、0.7241。上述数据表明,系统对黑色素细胞痣的检测正确率远超其他疾病,但皮肤纤维瘤的诊断精度较低。分析原因有二:一方面,黑色素细胞痣患者的样本数量过多,且其他6种疾病患者的样本数量较少,这导致系统在训练过程中更倾向于学习皮肤痣样本的

图像特征;另一方面,由于基底细胞癌、黑色素瘤的皮肤镜图像与黑素细胞痣的皮肤镜图像高度相似,而系统模型未能很好地分辨这3个疾病的细节。

4.5 对比实验分析

本文模型与数据集基准模型的测试集正确率如表3所示。

表3 本文模型与数据集基准模型的测试集正确率对比表

Tab. 3 Comparison table of test dataset accuracy between SFTNet and the dataset benchmark model

来源	ResNet18	ResNet50	Auto-sklearn	Auto-keras	Google AutoML Vision	SFTNet(ours)
文献[32]	0.727	0.719	0.734	0.756	0.766	—
文献[33]	0.754	0.731	0.719	0.749	0.768	—
本文	0.724	0.669	—	—	—	0.855

由于MedMNIST数据集是HAM10000皮肤病数据集的延伸,因此可将二者的结果对比评价。在MedMNIST-v1^[32]和MedMNIST-v2^[33]两个数据集的原始论文中,作者均使用ResNet18、ResNet50、Auto-sklearn、Autokeras、Google AutoML Vision五种模型进行实验。其中,ResNet18和ResNet50为经典CNN,2个模型的平均正确率是0.733,与本文4.4节中消融实验的平均正确率0.696相接近。Auto-sklearn^[34]、Auto-keras^[35]、Google AutoML Vision均为自动调参模型,三者的平均正确率达到0.749,略超

ResNet系列。本文在DETR架构的基础上加入GMTB与WTCB,模型正确率可达0.855。相较于文献[32-33],SFTNet模型的正确率指标平均值提升15.38%,改进效果明显。

不同深度学习模型在HAM10000皮肤病测试数据集上的评价指标参见表4。如表4所示,使用LeNet^[27]、VGG^[29]、GoogLeNet^[30]、ResNet^[31]、BoTNet^[35-36]、GhostNet^[37]、MobileNet^[38]、MobileNet-v2^[39]、MobileNet-v3^[40]、MobileViT-v3^[41]共10种基础模型及其衍生版本进行性能对比实验。

表4 不同模型在HAM10000皮肤病测试数据集上的评价指标

Tab. 4 Test evaluation indexes of different models on HAM10000 skin disease test dataset

对比来源	核心模型或主要方法	Accuracy	Precision	Recall	F1
文献[42]	LeNet	0.780	0.414	0.414	0.414
文献[43]	VGG	0.850	0.850	0.850	0.850
文献[44]	GoogLeNet	0.805	0.685	0.581	0.594
文献[45]	ResNet	0.852	0.849	0.975	0.908
文献[46]	BoTNet	0.815	0.726	0.686	0.692
文献[47]	GhostNet	0.837	0.837	0.832	0.833
文献[48]	MobileNet	0.793	0.680	0.800	0.763
文献[49]	MobileNet	0.832	0.830	0.830	0.754
文献[44]	MobileNet-v2	0.819	0.714	0.659	0.664
文献[44]	MobileNet-v2 & GoogLeNet-v2	0.835	0.766	0.656	0.672
文献[46]	MobileNet-v3	0.843	0.798	0.671	0.693
文献[47]	MobileNet-v3	0.838	0.839	0.834	0.833
文献[47]	MobileNet-v3 & CA	0.847	0.848	0.845	0.846
文献[47]	MobileNet-v3 & SE	0.848	0.846	0.849	0.846
文献[47]	Attention MobileNet-v3	0.850	0.849	0.842	0.841
文献[46]	MobileViT-v3	0.851	0.802	0.722	0.738
本文	SFTNet	0.855	0.852	0.851	0.851

由表4可知,Nugroho等^[42]对LeNet架构进行改进,该模型的Accuracy、Precision、Recall和F1指标分别为0.780、0.414、0.414、0.414,在所有模型中表现

最差;而本文提出的SFTNet在HAM10000测试数据集上表现优秀,其Accuracy、Precision、Recall和F1分别达到0.855、0.852、0.851、0.851,在17个模型中最

具优势。但 Patil 等^[45]所使用的 ResNet 模型在该数据集上的 Recall 和 F1 为 0.975 和 0.908,2 项指标远高于其他模型。出现异常数值的原因可能是作者对测试数据集进行调整,或实验数据记录有误。

在均衡性层面, VGG、ResNet、GhostNet、MobileNet、MobileViT-v3、MobileNet-v3、SFTNet 模型的 4 种评价指标之间方差较小,而 LeNet、GoogLeNet、BoTNet、MobileNet 四种模型的评价指标之间方差较大。这表明,前述 7 种模型的性能比后 4 种更加均衡。

进一步计算表 4 数据可得, MobileNet 系列模型的 4 个评价指标均值分别是 0.813、0.755、0.815、

0.759; MobileNet-v2 系列的 4 个评价指标均值分别是 0.827、0.740、0.658、0.668; MobileNet-v3 系列的均值则是 0.845、0.836、0.808、0.812。由上述结果可知,从 MobileNet 到 MobileNet-v2 再到 MobileNet-v3,正确率指标的数值逐步提升。这表明,随着 MobileNet 系列架构的迭代升级,其模型性能也有所提高。除 MobileNet-v3 模型的召回率略小于 MobileNet 模型外,其 Accuracy、Precision 和 F1 三个评价指标均为最优,而 MobileNet-v2 模型在上述 3 个评价指标中表现均为最差。

不同深度学习模型在同一指标下的数据对比如图 12 所示。

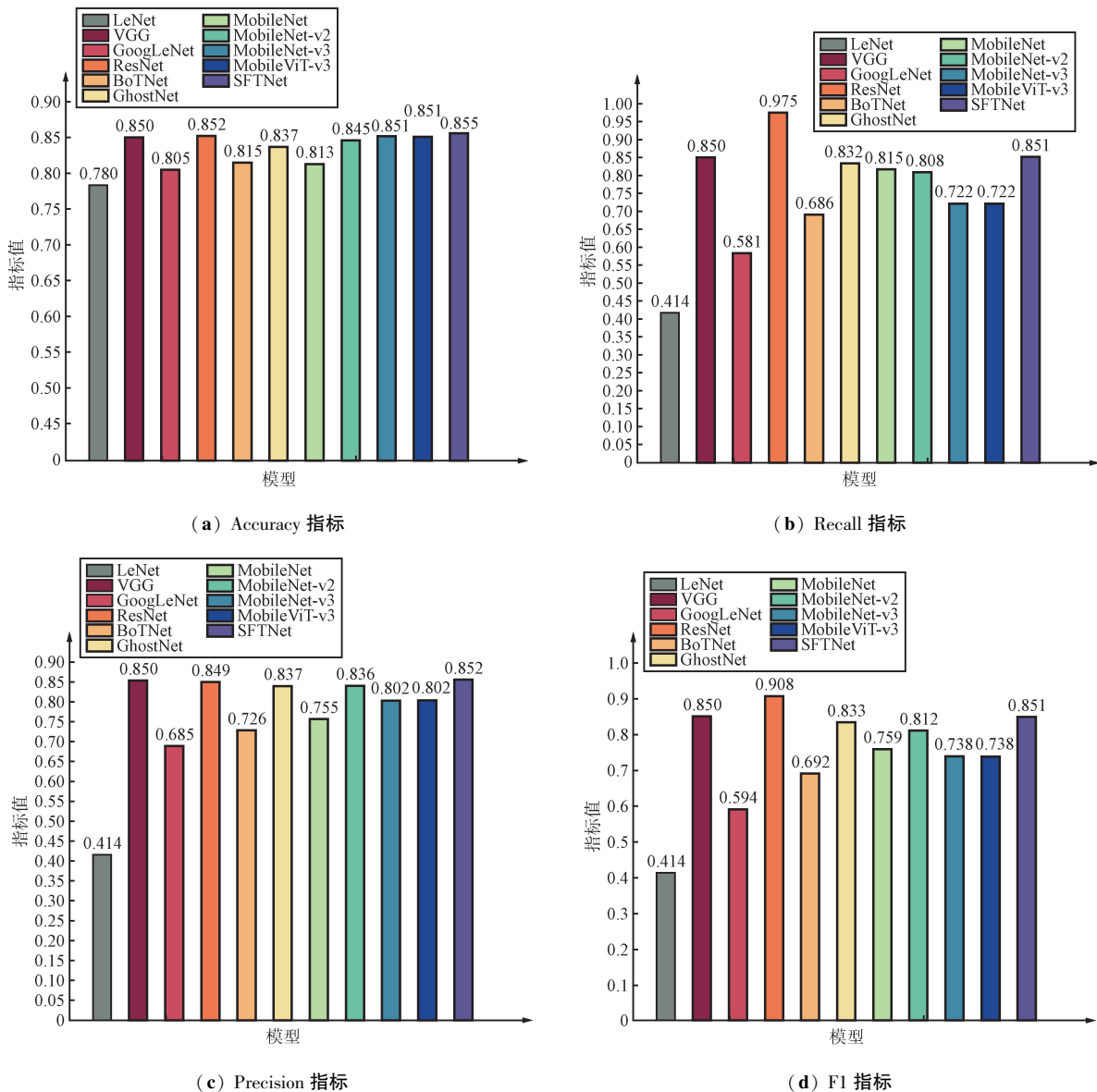


图 12 不同模型在同一指标下的数据对比

Fig. 12 Data comparison of different models under the same index

由图 12 可知,在 Recall 与 F1 方面,MobileViT-v3 模型与 BoTNet、MobileNet-v2、GoogLeNet、LeNet 模型之间出现显著断层。针对正确率指标,11 类模型的性能排序为 SFTNet>ResNet>MobileViT-v3>VGG>MobileNet-v3>GhostNet>MobileNet-v2>BoTNet>MobileNet>GoogLeNet>LeNet;而针对精确率指标,11 类模型的性能排序为 SFTNet>VGG>ResNet>GhostNet>MobileNet-v3>MobileViT-v3>MobileNet>MobileNet-v2>BoTNet>GoogLeNet>LeNet。上述结果显示,不同模型在正确率和精确率层面上互有优劣。总体而言,SFTNet、ResNet、MobileNet-v3 三类模型的性能较强,而 GoogLeNet、LeNet 的性能偏弱。模型的正确率过低暗示其极易出现误诊现象,并不适合应用在皮肤病诊断的临床场景。

5 结束语

本文创新性地提出基于 GT 和小波卷积的图像处理模块,并于 DETR 架构的基础上加以改进,在皮肤病检测方面获得较佳效果。仿真实验表明,GMTB 能够提升皮肤镜图像的对比度,减少模型训练过程中的过拟合效应;WTCB 在原有的空域信息基础上补充频域信息,增大卷积核感受野的同时捕捉更多图像细节^[50]。该研究的不足之处在于,在提升系统整体检测性能的同时,GMTB 和 WTCB 也增加了模型的训练成本,且系统针对皮肤纤维瘤和黑色素瘤的诊断正确率较低,即系统仍有一定改进空间。

由此可见,基于 SFTNet 的智能皮肤癌预筛系统不仅能够成功识别患者的皮肤疾病,也有效地减轻了皮肤性病科医师的工作负担。该系统可以对皮肤癌患者进行快速筛查和预警,适合在各级医疗单位推广与应用。

参考文献

- [1] 穆亚楠,陈星会,何璐,等.常见非黑色素瘤皮肤癌新兴治疗的研究进展[J/OL].(2024-11-18)[2025-04-18].
<http://kns.cnki.net/kcms/detail/61.1415.R.20241115.2350.007.html>.
- [2] 崔勇,高兴华.皮肤病性病[M].10版.北京:人民卫生出版社,2024.
- [3] 刘道远.皮肤镜诊断常见非黑色素瘤皮肤癌的临床价值研究[D].昆明:昆明医科大学,2023.
- [4] 邓凡智.皮肤镜在基底细胞癌诊断中的应用价值[D].蚌埠:蚌埠医学院,2021.
- [5] 崔莉,王贵佐,马小萍,等.皮肤镜检查在黑素细胞痣、脂溢性角化病和基底细胞癌中的诊断价值[J].陕西医学杂志,2021,50(4):428-431.
- [6] 江浩波,孙志平,陈海针,等.皮肤镜在脂溢性角化、日光性角化和基底细胞癌诊断中的应用[J].皮肤性病诊疗学杂志,2019,26(4):204-209.
- [7] LIU W, ANGUEVOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector [C]//European Conference on Computer Vision. Cham;Springer,2016:21-37.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus:IEEE,2014:580-587.
- [9] GIRSHICK R. Fast R-CNN [C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago:IEEE,2015:1440-1448.
- [10] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks [J]. Advances in Neural Information Processing Systems. IEEE,2015,28:1137-1149.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-time Object Detection [C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas;IEEE,2016:779-788.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is All You Need [J]. Advances in Neural Information Processing Systems,2017,30:5998-6008.
- [13] CARION N, MASSA F, SYNNAEVE G, et al. End-to-End Object Detection with Transformers [C]//Proceedings of the 2020 European Conference on Computer Vision. Glasgow;Springer,2020:213-229.
- [14] 赵嘉晖,温杰,蔡星娟,等.基于高维多目标优化的皮肤癌检测[J].控制与决策,2024,39(11):3597-3606.
- [15] 王煜坤.皮肤影像及人工智能辅助皮肤疾病精准诊疗新探索[D].北京:北京协和医学院,2024.
- [16] TSCHANDL P, RINNER C, APALLA Z, et al. Human-computer Collaboration for Skin Cancer Recognition [J]. Nature Medicine,2020,26(8):1229-1234.
- [17] GONZALEZ R C, WOODS R E. Digital Image Processing [M].北京:电子工业出版社,2020.
- [18] 禹晶,肖创柏,廖庆敏.数字图像处理[M].北京:清华大学出版社,2022.
- [19] FOURIER J B J. Théorie Analytique de la Chaleur [M]. Paris:Firmin Didot,1822.
- [20] AHMED N, NATARAJAN T, RAO K R. Discrete Cosine Transform [J]. IEEE Transactions on Computers,1974,23(1):90-93.

- [21] DEVANEY A J. The Slant Transform: An Orthogonal Transform for Signal Processing[J]. IEEE Transactions on Audio and Electroacoustics, 1972, 20(2): 140–146.
- [22] WALSH J L. A Closed Set of Normal Orthogonal Functions[J]. Transactions of the American Mathematical Society, 1923, 25(1): 5–24.
- [23] DAUBECHIES I. Ten Lectures on Wavelets[M]. Philadelphia: Society for Industrial and Applied Mathematics, 1992.
- [24] 李艳玲. 皮肤镜在颜面部皮肤肿瘤筛检中的应用[J]. 中国现代医生, 2022, 60(3): 21–23.
- [25] TSCHANDL P, ROSENDAHL C, KITTLER H. The HAM10000 Dataset, a Large Collection of Multi-source Dermatoscopic Images of Common Pigmented Skin Lesions [J]. Scientific Data, 2018, 5: 1–10.
- [26] LAURENS V D M, HINTON G E. Visualizing Data Using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(86): 2579–2605.
- [27] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based Learning Applied to Document Recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278–2324.
- [28] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks [C]//Advances in Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2012: 1097–1105.
- [29] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-scale Image Recognition [C]//Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015). San Diego: OpenReview, 2015: 1–14.
- [30] SZEGEDY C, LIU W, JIA Y Q, et al. Going Deeper with Convolutions [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1–9.
- [31] HE K M, ZHANG X Y, REN S Q, et al. Delving Deep into Rectifiers: Surpassing Human-level Performance on ImageNet Classification [C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 1026–1034.
- [32] YANG J C, SHI R, NI B B. MedMNIST Classification Decathlon: A Lightweight Automl Benchmark for Medical Image Analysis [C]//IEEE 18th International Symposium on Biomedical Imaging (ISBI). Piscataway: IEEE, 2021: 191–195.
- [33] YANG J C, SHI R, WEI D L, et al. MedMNIST v2: A Large-scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification [J]. Scientific Data, 2023, 10(1): 1–41.
- [34] FEURER M, KLEIN A, EGGENSPERGER K, et al. Efficient and Robust Automated Machine Learning [C]//Advances in Neural Information Processing Systems. Montreal: Curran Associates Inc., 2015: 2962–2970.
- [35] JIN H F, SONG Q Q, HU X. Auto-keras: An Efficient Neural Architecture Search System [C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage: ACM, 2018: 1946–1956.
- [36] ARAVIND S, TSUNG-YI L, NIKI P, et al. Bottleneck Transformers for Visual Recognition [EB/OL]. [2025–04–10]. <https://arxiv.org/pdf/2101.11605>.
- [37] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More Features from Cheap Operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 1580–1589.
- [38] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [EB/OL]. (2017–04–17) [2025–04–18]. <https://arxiv.org/pdf/1704.04861>.
- [39] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2018: 4510–4520.
- [40] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). New York: IEEE, 2019: 1314–1324.
- [41] WADEKAR S N, CHAURASIA A. MobileViTv3: Mobile-Friendly Vision Transformer with Simple and Effective Fusion of Local, Global and Input Features [EB/OL]. (2022–09–01) [2025–04–18]. <https://arxiv.org/pdf/2209.15159v2>.
- [42] NUGROHO A A, SLAMET I, SUGIYANTO. Skins Cancer Identification System of HAM10000 Skin Cancer Dataset Using Convolutional Neural Network [EB/OL]. (2019–12–27) [2025–04–18]. <https://doi.org/10.1063/1.5141652>.
- [43] RATUL A R, MOZAFFARI M H, LEE W S, et al. Skin Lesions Classification Using Deep Learning Based on Dilated Convolution [EB/OL]. (2020–06–05) [2025–04–18]. <https://doi.org/10.1101/860700>.
- [44] THURNHOFER H K, LÓPEZ R E, DOMÍNGUEZ E, et al. Skin Lesion Classification by Ensembles of Deep Convolutional Networks and Regularly Spaced Shifting [J]. IEEE Access, 2021, 9: 112193–112205.

- [45] PATIL S N, PATIL H D, ADHIYA K P, et al. Densenet-201 for Skin Melanoma Classification; A Comprehensive Performance Evaluation and Analysis [J]. International Research Journal of Multidisciplinary Scope, 2024, 5(4): 711-721.
- [46] LIU H, DOU Y, WANG K, et al. A Skin Disease Classification Model Based on Multi Scale Combined Efficient Channel Attention Module [J]. Scientific Reports, 2025, 15: 1-12.
- [47] 胡吉红. 基于轻量级网络的皮肤病检测算法研究 [D]. 兰州: 西北师范大学, 2023.
- [48] SAE L W, WETTAYAPRASIT W, AIYARAK P. Convolutional Neural Networks Using MobileNet for Skin Lesion Classification [C] // 2019 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Long Beach: IEEE, 2019: 242-247.
- [49] CHATURVEDI S S, GUPTA K, PRASAD P S. Skin Lesion Analyser: An Efficient Seven-way Multi-class Skin Cancer Classification Using MobileNet [C] // Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020. Singapore: Springer, 2021: 165-176.
- [50] 朱秋慧, 杨靖, 黄若愚, 等. 基于部分卷积的多尺度特征卷积神经网络模型 [J/OL]. (2025-03-11) [2025-04-18]. <http://kns.cnki.net/kcms/detail/13.1099.TN.20250310.1707.012.html>.

作者简介:

白世强 男, (1999—), 硕士研究生。主要研究方向: 智能医学工程、医学影像感知。

蔡益人 男, (2000—), 硕士研究生。主要研究方向: 小样本分类。

(*通信作者) 李丽宏 女, (1973—), 博士, 教授。主要研究方向: 图像处理、计算机视觉。