



DOI:10.12404/j.issn.1671-1815.2402339

引用格式:王逸蓓,王芳.基于胶囊网络的皮肤癌图像分类[J].科学技术与工程,2025,25(8):3280-3287.

Wang Yibei, Wang Fang. Skin cancer image classification based on capsule network[J]. Science Technology and Engineering, 2025, 25(8): 3280-3287.

自动化技术、计算机技术

## 基于胶囊网络的皮肤癌图像分类

王逸蓓,王芳\*

(燕山大学理学院,秦皇岛 066004)

**摘要** 胶囊网络皮肤癌对图像特征的性质和空间关系进行编码,克服卷积神经网络池化过程信息丢失的缺点。针对胶囊网络中只能提取浅层特征和压缩函数收敛性能问题,提出了ResNeXt与胶囊网络级联的Rs-Capsnet网络。首先采用ResNeXt网络中学习图像的复杂特征,利用Inception模块和残差连接提取深层特征;通过CBAM注意力模块调整特征图权重并将其输送到胶囊模块;然后使用改进的压缩函数的胶囊网络完成分类,最后将改进后的网络同主流模型进行对比。结果表明:Rs-Capsnet在皮肤癌图像分类上表现了更佳的性能。

**关键词** 皮肤癌;胶囊网络;ResNeXt;压缩函数;Rs-Capsnet

**中图分类号** TP181; **文献标志码** A

### Skin Cancer Image Classification Based on Capsule Network

WANG Yi-bei, WANG Fang\*

(College of Science, Yanshan University, Qinhuangdao 066004, China)

**[Abstract]** Capsule networks can encode the properties and spatial relationships of skin cancer image features, thereby overcoming the disadvantage of information loss in the pooling process of convolutional neural networks. Aiming at the problem that only shallow features can be extracted and the convergence performance of the squash function in capsule networks, a ResNeXt cascaded with capsule networks was proposed for Rs-Capsnet networks. Firstly, the complex features of the image were learned using the ResNeXt network. The Inception module and the residual connection were used to extract the deep features, and the weights of the feature map were adjusted and delivered to the capsule module through the CBAM attention module. Then, an improved squash function capsule network was used to complete the classification. Finally, the improved network was compared with mainstream models. The results show that Rs-Capsnet exhibits better performance in skin cancer image classification.

**[Keywords]** skin cancer; capsule network; ResNeXt; squash function; Rs-Capsnet

皮肤癌是世界公认的重大健康难题,其发病率以可怕的速度在增长。常见的皮肤恶性肿瘤主要包括基底细胞瘤、鳞状细胞瘤和黑色素瘤,其中基底细胞瘤和鳞状细胞瘤占有皮肤恶性肿瘤的95%。虽然黑色素瘤相对罕见,但它恶性程度极高,往往在早期进行转移。据世界卫生组织统计,2020年全球诊断出 $2\,020\,325 \times 10^4$ 黑色素瘤病例。癌症的早期发现对于在早期阶段治疗和提高患者生存率至关重要<sup>[1]</sup>。皮肤组织的活检检测是发现皮肤癌的主要手段,这种传统的方式依赖于人类视觉检查,很容易延误医疗诊断。

近年来,计算机视觉不断发展,在图像分类、图像分割和目标检测等领域取得了显著成果。卷积神经网络(convolutional neural network, CNN)为骨干网络,从深度和宽度方面不断优化网络。Zhang等<sup>[2]</sup>提出了一种用于皮肤癌图像分类的注意力残差卷积神经网络(attention residual convolutional neural network, ARL-CNN),堆叠多个注意力残差块(ARL)进行特征学习。Gomathi等<sup>[3]</sup>提出了一种双重优化的深度学习模型(dual optimization deep learning, DODL-CNN),通过细菌觅食算法和粒子群优化算法从分割图像中提取特征,建立部分特征与

收稿日期:2024-04-02; 修订日期:2024-12-17

基金项目:国家自然科学基金(62073234);河北省自然科学基金(F2020203105);河北省高等学校科学技术研究项目(ZD2022012)

第一作者:王逸蓓(1999—),女,汉族,河北石家庄人,硕士研究生。研究方向:图像分类。E-mail:1175210416@qq.com。

\*通信作者:王芳(1984—),女,汉族,安徽淮北人,博士,副教授。研究方向:多智能体控制。E-mail:wangfang@ysu.edu.cn。

投稿网址:www.stae.com.cn

全局的关系。Akilandasowmya 等<sup>[4]</sup>利用沙猫群优化算法在已知特征中提取深层隐藏特征,通过 Resnet50对皮肤癌图像分类。轻量级神经网络在 CNN 的发展中逐渐显现它的优势,它在保证准确率的前提下,从体积和速度对网络性能进行优化。Karar 等<sup>[5]</sup>研究了轻量级网络 EfficientNet B0-B7 在皮肤癌图像分类中的表现,通过对比实验, EfficientNet B4 表现出更高的模型泛化能力,准确率达到 87.91%。

尽管 CNN 在图像分类领域发挥了重要作用,但是存在一些缺陷:①CNN 不编码特征之间的空间关系,忽略了姿态和位置信息;②网络容易受到对抗性扰动的影响,鲁棒性较差;③池化时对视点变化的平移不变性,丢失有价值的空间信息。为了克服 CNN 的缺点, Hinton 等<sup>[6]</sup>提出了胶囊网络(Capsnet),将向量代替标量作为神经元,对特征的姿态信息进行编码,在连续两个胶囊层中通过动态路由机制将保留的空间信息从一层传到下一层。Pérez 等<sup>[7]</sup>提出了结合卷积块和胶囊网络的架构,通过复杂的卷积计算模块进行前期的特征提取,对胶囊网络的参数进行调整完成黑色素瘤的分类。Wang 等<sup>[8]</sup>提出了一种同时提取卷积特征和胶囊特征的双通道结构,将两种特征融合到新胶囊,然后通过修改胶囊网络的损失函数以适应路由系数的自适应优化。Lan 等<sup>[9]</sup>提出了基于卷积注意力模块(convolutional block attention module, CBAM)的改进胶囊网络,结果表明, CBAM 中空间注意力和通道注意力的连接会减少由于池化引起的空间信息损失。Goceri<sup>[10]</sup>提出了一种可调节完全卷积胶囊网络(AFConv-Capsnet)来对皮肤癌图像进行分类,在可调节胶囊层中,胶囊的空间关系以恒定的学习偏差进行编码,提高了模型精度,减少了运行时间。当利用胶囊网络对皮肤癌图像进行分类时,由于在进行特征提取时只采取单卷积核,忽略了深层次信息,降低了模型精度,在进行路由过程中,路由的迭代增加了模型训练难度,使训练时间变长。

现提出一种将 ResNeXt 与 Capsnet 级联的深度学习模型 Rs-Capsnet。增加网络层数,以 ResNeXt 网络完成特征提取,将提取的特征输送到胶囊网络中;路由层中对父胶囊的压缩函数以一种收敛性能更快的函数来代替,提高网络运行效率;将 CBAM 注意力机制引入网络架构中,抑制冗余特征。通过改进模型的皮肤癌图像分类,能够在临床诊断中提供快速而准确的辅助支持,减轻医生的工作负担,提高诊断效率。

## 1 胶囊网络

在一般的 CNN 中,池化层用于增加感受野并减少计算次数,同时保持卷积运算的平移不变性。然而,存在的平移不变性失去了对象的相对位置特征,使学习对象在相对位置和空间特征方面表现不佳。CNN 的基本单位是单个神经元<sup>[11]</sup>,而胶囊网络将一组神经元作为基本单位,将这组神经元称为一个胶囊(神经元簇)。以识别人脸为例, CNN 可以很容易地将违反生理结构的人工修改的面部图像识别为常规面部图像。这是因为 CNN 只检测图像中人类的面部特征,而忽略了检测到的特征的空间位置。胶囊网络就可以弥补这个缺点,在进行面部识别时由于将空间位置信息同其他面部信息结合可以将其正确识别为非常规面部图像。胶囊将一个向量看作一个整体,每一个胶囊都表示一种特征,而向量的每一维都表示该特征的特点,向量的模值代表该特征存在的概率。在网络中,每个胶囊代表被处理对象的某个属性。

### 1.1 胶囊网络架构

胶囊网络的总体框架包括输入层、卷积层、主胶囊层、数字胶囊层和全连接解码层。在胶囊网络的输入层中一般是灰色图像或者 RGB 三通道的彩色图像。卷积层由一个卷积核大小为  $9 \times 9$ ,步长为 2,通道数为 256 的卷积构成。主胶囊层由两部分构成:第一部分是单层卷积由卷积核  $9 \times 9$ ,步长为 2,通道数为 256 构成;第二部分是将计算得到的特征图进行 reshape 操作,包装成一组通道维度为 8 的初级胶囊(子胶囊)。数字胶囊层由主胶囊层的胶囊经过路由机制后得到,它是被用来存储代表分类结果的胶囊输出(父胶囊)。三层全连接层是作为胶囊网络的解码器,通过对图像的重构,建立重构图像和原图像的重构损失函数。

### 1.2 路由机制

路由是指在同一层网络中的每一对胶囊都必须有为自身计算的标量权重。动态路由算法是通过测量低级胶囊对高级胶囊的一致性来确定是否应该激活高级胶囊。

如图 1 所示,将卷积层所保留的空间特征信息进行 reshape 操作包装成初级胶囊  $u_i$ ,对初级胶囊进行计算。在动态路由算法中, $i$  为子胶囊个数, $j$  为父胶囊个数。

(1)输入向量的矩阵乘法。对进行卷积 reshape 后的胶囊  $u_i$ ,利用可学习的权重矩阵  $W_{ij}$  作仿射变换,得到预测胶囊  $\hat{u}_{ji}$ 。

$$\hat{\mathbf{u}}_{ji} = \mathbf{W}_{ij} \mathbf{u}_i \quad (1)$$

(2) 输入向量的加权求和。 $c_{ij}$  量化了低级胶囊对高级胶囊的重要程度, 称为耦合系数;  $b_{ij}$  是为了耦合系数设置的中间系数, 在进行路由计算前, 将  $b_{ij}$  进行初始化, 其更新过程见算法 1。通过耦合系数对初级胶囊进行加权并求和, 得到压缩前的高级胶囊  $\mathbf{s}_j$ 。

$$\mathbf{s}_j = \sum_i c_{ij} \hat{\mathbf{u}}_{ji} \quad (2)$$

$$c_{ij} = \frac{e^{-b_{ij}}}{\sum_i e^{-b_{ij}}} \quad (3)$$

(3) 向量非线性化后输出。对向量进行非线性化也就是压缩操作, 设计 squash 函数, 操作后向量的方向不变, 向量长度收缩到 (0,1)。式(4)乘号左边是特征被高级胶囊感知到的概率, 右边是保持向量方向。输出向量  $\mathbf{x}$  的模长表示属于该类结果的概率。 $\mathbf{s}_j$  经过 squash 压缩后得到父胶囊  $\mathbf{v}_j$ 。

$$\text{squash}(\mathbf{x}) = \frac{\|\mathbf{x}\|^2}{1 + \|\mathbf{x}\|^2} \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad (4)$$

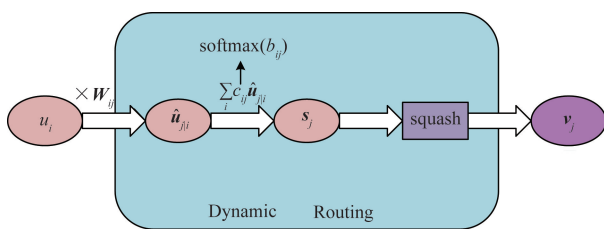


图 1 动态路由过程

Fig. 1 Dynamic routing process

squash 函数是胶囊网络的重要组成部分, 通过该函数压缩胶囊的输出向量, 既可以将向量的长度进行概率归一化也可以保留特征的方向信息。

算法 1 动态路由

```

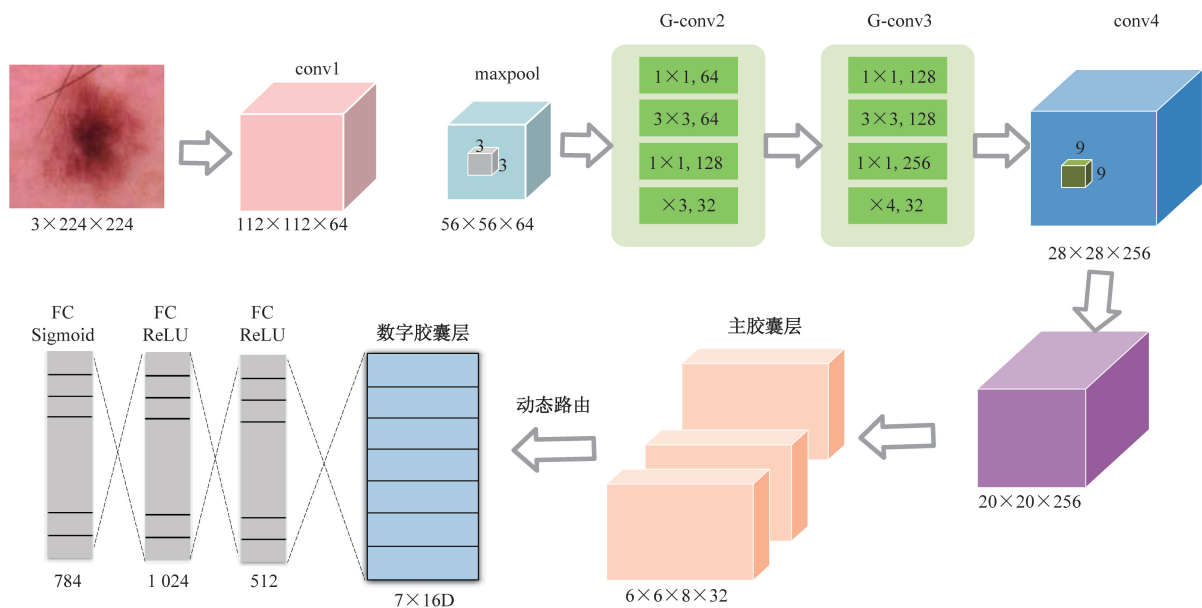
输入:  $\hat{\mathbf{u}}_{ji}, r, L$ , 其中,  $r$  为迭代次数,  $L$  为第  $L$  层。
输出:  $\mathbf{v}_j$ 
for all caps  $i$  in  $L$  layer and all caps  $j$  in  $L+1$  layer:  $b_{ij} \leftarrow 0$ 
for epoch = 1, 2, ..., r
    for all caps  $i$  in  $L$  layer:  $c_i \leftarrow \text{softmax}(b_i)$ 
    for all caps  $j$  in  $L+1$  layer:  $\mathbf{s}_j \leftarrow \sum_{ij} c_{ij} \hat{\mathbf{u}}_{ji}$ 
    for all caps  $j$  in  $L+1$  layer:  $\mathbf{v}_j \leftarrow \text{squash}(\mathbf{s}_j)$ 
for all caps  $i$  in  $L$  layer and all caps  $j$  in  $L+1$  layer:
 $b_{ij} \leftarrow b_{ij} + \hat{\mathbf{u}}_{ji} \mathbf{v}_j$ 
return  $\mathbf{v}_j$ 
    
```

## 2 Rs-Capsnet

搭建 Rs-Capsnet 来完成图像分类任务, 具体如图 2 所示。该网络分为: 输入层、特征提取层、注意力层、胶囊层和全连接层。所提出的模型是使用 ResNext 网络生成初始特征图, 通过注意力层对特征图进行加权, 然后将其传输到胶囊网络中进行分类。

### 2.1 特征提取层

传统胶囊网络只以单层卷积核来进行特征提取, 不足以完全表征所有特征信息, 所以需要额外的卷积层确定输入特征。ResNeXt 网络<sup>[12]</sup>是将 Inception 模块与 Resnet 网络进行结合, 它既保留了



conv 为卷积操作; maxpool 为最大池化; G-conv 为组卷积; FC 为全连接层; Sigmoid 和 ReLU 为激活函数

图 2 Rs-Capsnet 网络总体架构

Fig. 2 The overall architecture of the Rs-Capsnet network

Inception 模块中实现跨通道的交互信息整合的优点又保留了残差连接中保留原始特征的优点。它提出的组卷积是基于普通卷积和深度可分离卷积中间的一种新型卷积方式,组数由变量基数控制。该网络基于在同一层中使用多条并行路径,而不同于经典网络中的单行路径。

以输入通道数是 256 为例,采用 ResNeXt 网络中的 Block 块进行分组卷积,每个分支的拓扑结构都是相同的,在图 3 中结构 1 和结构 2 是等价的。Inception 模块与 Resnet 的结合增加了网络的深度和每层中组数增加了宽度,深度可以捕获更复杂的特征,宽度可以捕获更多样的特征。

在 Rs-Capsnet 中采取 ResNeXt 网络的两个模块来适应胶囊网络的尺寸,使用两个并行路径的组卷积,二者之间通过残差连接。每个单独的残差模块都包含 3 个并行的卷积层,而这些并行的卷积层之间共享输入信息,也就是说进行分组后的卷积具有更小的参数量。最后将它们的输出相加在一起,这样可以达到在分组后不同通道特征信息的交互,而不是通过增加网络的宽度来融合通道信息,该结构在保持计算效率的同时可以给模型带来更好的性能。

### 2.2 注意力层

CBAM 是一种轻量级的注意力模块<sup>[9]</sup>,它沿着两个独立的维度—空间和通道依次推断注意力图,将注意力图与输入特征图相乘以自适应特征

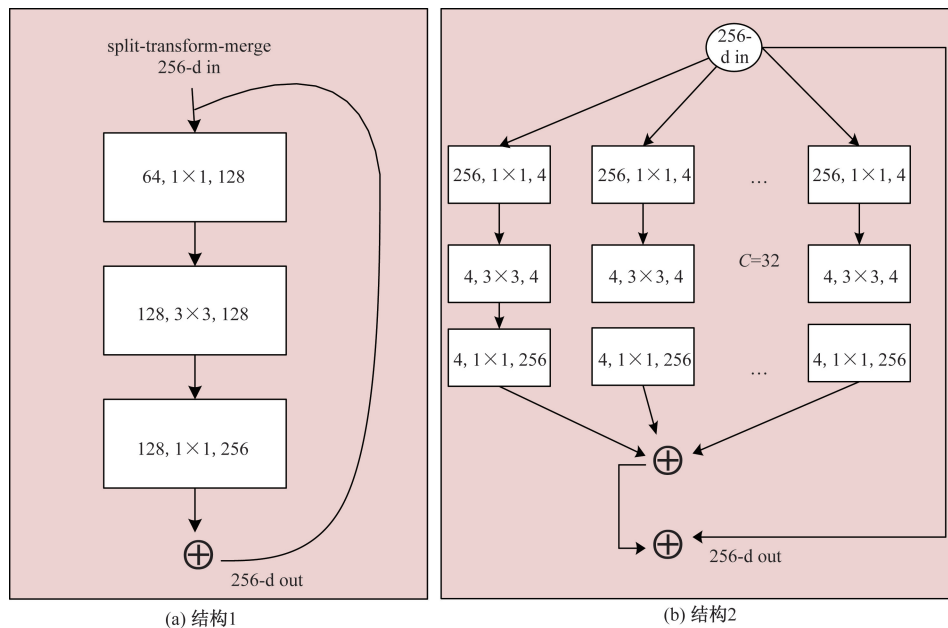
优化。CBAM 是一个即插即用的模块,可以无缝插入深度学习网络中,所造成的计算量可以忽略不计。CBAM 在学习生成空间注意力和通道注意力的过程中,自适应的关注皮肤病变图像中的重要特征信息,在空间注意力图中关注病变区域,在通道注意力图中关注图像的特征。将 CBAM 模块放在完成特征提取后,这样将提取到的特征信息进行筛选,关注重要特征,而对冗余特征进行抑制,提高网络运行效率,增加网络的表征能力。

### 2.3 胶囊层

胶囊层由主胶囊层和数字胶囊层构成<sup>[13]</sup>。两个胶囊层之间的信息通过动态路由来传递。路由中的激活函数确定将子胶囊传递到父胶囊的信息。根据非线性 squash 压缩函数的特点,将其进行改进,使它收敛速度更快,在微小变差中,对输出胶囊的波动影响更大,因此可以优化胶囊网络的运行效率,提高对某类图像的识别准确率。压缩函数值的大小对自变量的变化更敏感并允许各类别之间拥有更大的差距。对 squash 压缩函数改进为  $s_0(\mathbf{x})$  和  $s_i(\mathbf{x})$ , 计算公式为式(5)和式(6),定义一个超参数  $\alpha, \alpha = 1$  时,  $\alpha$  的形式就是 tanh 激活函数,通过对  $\alpha$  的学习,寻求最优的压缩函数。

$$s_0(\mathbf{x}) = \frac{1 - e^{-\|\mathbf{x}\|}}{\|\mathbf{x}\|} \mathbf{x} \tag{5}$$

$$s_i(\mathbf{x}) = \frac{\alpha e^{\|\mathbf{x}\|} - e^{-\|\mathbf{x}\|}}{\alpha e^{\|\mathbf{x}\|} + e^{-\|\mathbf{x}\|}} \frac{\mathbf{x}}{\|\mathbf{x}\|}, \quad i = 1, 2, \dots \tag{6}$$



256-d in 表示输入特征有 256 个通道;经过组卷积后输出为 256-d out,即也有 256 个通道;  
split-transform-merge 为分离输入特征-对特征进行处理-合并输出策略;C 为分组的组数

图 3 ResNeXt 结构

Fig. 3 ResNeXt architecture

新的压缩函数在函数形态上的表现如图 4 所示,对比随着自变量的变化,  $\alpha = i(i = 1, 2, \dots, 10)$  时, squash 原函数(以  $s$  表示)和  $s_0, s_i$  的变化趋势,收敛速度从大到小排列为  $s_0 > s_i > s$ 。但是随着  $\alpha$  增大,整个函数的值开始逐渐趋向于 1,趋向的速度也越来越快,当增加到一定程度,对压缩函数的收敛性起到了相反作用,即  $\alpha$  的增加对胶囊输出的影响反而减小。所以需要选择最佳的  $\alpha$  来优化 squash 函数。

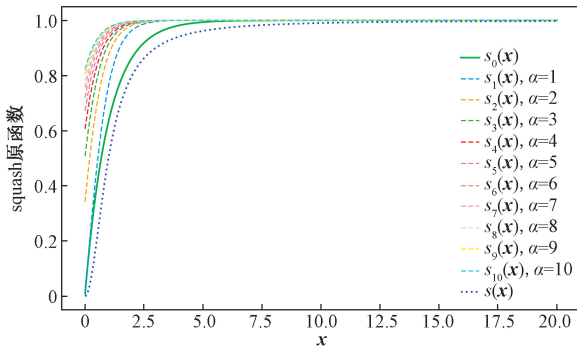


图 4 压缩函数对比图

Fig. 4 Comparison chart of compression functions

### 2.4 全连接层

为了证明胶囊网络能够保留特征的空间信息,三层全连接层作为解码器,反卷积重构输入图像<sup>[14-15]</sup>,使用胶囊的模长来判断局部特征发生的概率。对胶囊网络中出现的每个类别,训练边际损失函数  $L_c$  为

$$L_c = T_c \max(0, m^+ - \|v_c\|)^2 + \lambda(1 - T_c) \max(0, \|v_c\| - m^-)^2 \quad (7)$$

式(7)中:  $v_c$  为第  $c$  类样本的胶囊模长;  $T_c$  为  $c$  类是否存在,若存在  $T_c$  为 1,不存在为 0;  $m^+ = 0.9$ ;  $m^- = 0.1$ ;  $\lambda$  为比例系数,设为 0.5<sup>[16]</sup>。

网络总损失函数为边际损失和重建损失之和,在重建损失中进行参数设定<sup>[17]</sup>,对权重进行分配,以边际损失为主导。

$$L = L_c + 0.0005L_R \quad (8)$$

式(8)中:  $L_R$  为重建损失函数;  $L_c$  为边际损失函数。

## 3 实验设置和结果分析

### 3.1 实验环境和评价指标

本文算法基于 Pytorch 框架在 Python 环境中运行,具体实验环境如表 1 所示,使用 Adam 优化器作为梯度下降算法进行训练<sup>[18-19]</sup>,初始学习率为 0.0001,通过余弦退火调整学习率, batchsize 设置为 128, epoch 设置为 50。在训练过程中,学习率会分阶段降低,从而使网络达到最佳状态。

表 1 实验环境

Table 1 Experimental environment

设施	型号
算力型号	RTX3090
CPU	Xeon(R) Platinum 8255C
Pytorch	1.11.0
Python	3.8
Cuda	11.3
操作系统	Ubuntu20.04

### 3.2 数据集

研究对象是 HAM10000 数据集包含 10 015 张皮肤癌图像<sup>[18]</sup>,包括 7 种皮肤癌类型,分别为光化性角化病(AKIEC)、皮肤纤维瘤(DF)、基底细胞癌(BCC)、黑色素瘤(MEL)、良性角化病(BKL)、血管病变(VASC)和黑色素细胞痣(NV)<sup>[20]</sup>。图 5 为 7 种皮肤癌类型。数据集按 9:1 的比例划分为训练集和测试集。

从图 6 可以看出,黑色素细胞痣在总数据集的占比超过 50%,存在严重的类不平衡问题<sup>[21]</sup>,类样本个数的极度稀少会使得该类样本在网络的准确率下降,从而导致整个网络的表现效果不佳。首先对图片进行预处理,通过 ImageDataGenerator 库进行数据增强<sup>[22]</sup>,增加训练集的少数类样本数,使各类别样本数达到平衡。

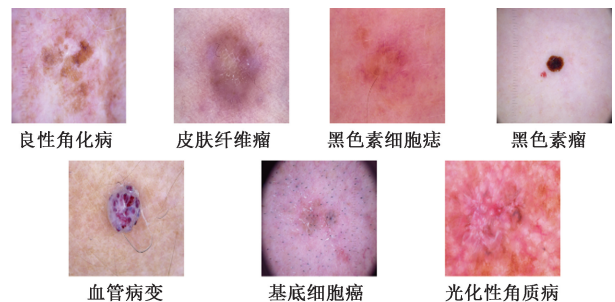


图 5 皮肤癌样本

Fig. 5 Skin cancer sample

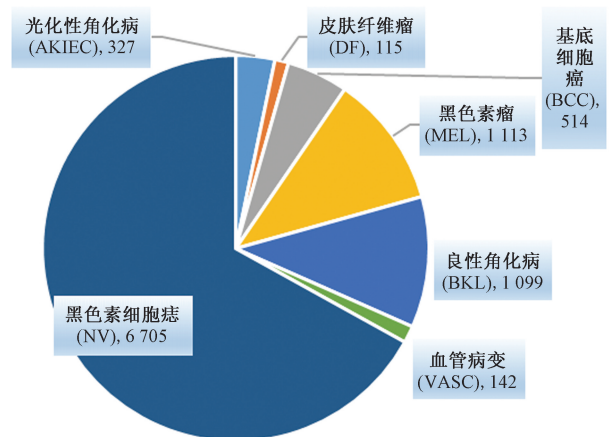


图 6 不同类别皮肤癌样本分布比例

Fig. 6 Proportion of skin cancer samples by category

### 3.3 评价指标

为了全面的评估 Rs-Capsnet 的网络性能, 采用精确度 (precision, 记为  $P$ )、召回率 (recall, 记为  $R$ )、特异性 (specificity, 记为  $S$ ) 和  $F_1$  指标<sup>[23]</sup> 来进行计算。

$$P = \frac{TP}{TP + FP} \times 100\% \quad (9)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (10)$$

$$S = \frac{TN}{FP + TN} \times 100\% \quad (11)$$

$$F_1 = 2 \frac{PR}{P + R} \times 100\% \quad (12)$$

式中: TP 为预测为正样本且预测正确 (真阳性); TN 预测为负样本其预测正确 (真阴性); FP 为正样本且预测错误 (假阳性); FN 为负样本且预测错误 (假阴性)<sup>[24]</sup>。

### 3.4 对比实验

根据不断调试参数, 当  $\alpha = 2$  时, 带来的网络增益是最佳的, 因此 Rs-Capsnet 的 squash 函数指的是  $s_2$ 。对 Rs-Capsnet 模型进行训练, 在 50 次迭代中, 训练精度和损失变化分别如图 7 和图 8 所示。

从图 9 的混淆矩阵中可以看出, 每一类样本被正确预测的个数, Rs-Capsnet 模型对不同类别的皮肤癌图像的代表能力也是不同的, 以训练集样本数最多的类别黑色素细胞痣为例, 从图 9 中可以看出, 将其正确分类的样本数为 647, 将 NV 类误判为 BKL 类的样本数为 13, 是 NV 类错判率最高的。根据混淆矩阵可以计算每一类皮肤癌样本的正确率, 正确率最高的黑色素细胞痣, 达到 97.15%。

从表 2 中数据可以看出, Rs-Capsnet 模型取得了较好的实验结果。为了更好地说明所建立模型的有效性, 对其他模型进行对比实验<sup>[25]</sup>。VGG-Capsnet 是以 VGG16 作为特征提取器, Capsnet 作为

分类器建立的深度学习网络模型; CBAM-Capsnet 是在大卷积核特征提取后添加 CBAM 注意力机制后完成分类。根据表 2 中数据可知, 在特征提取层的加深使 Capsnet 的准确率提高了 8.19%, 而在 CBAM-Capsnet 中, 网络精度的提升是显著的。而轻量级网络 EfficientNet 在该数据集上的表现优于 Capsnet, 在准确率上高出 3 个百分点。以 50 层的 ResNeXt 进行分类, 其准确率比提出的模型相差 5.19%。

根据表 3 可知, Capsnet 的引入增加了模型的参数量, 相对于轻量级网络<sup>[26]</sup> 和深层次的传统网络, 胶囊网络所带来的参数的增加是成倍的。

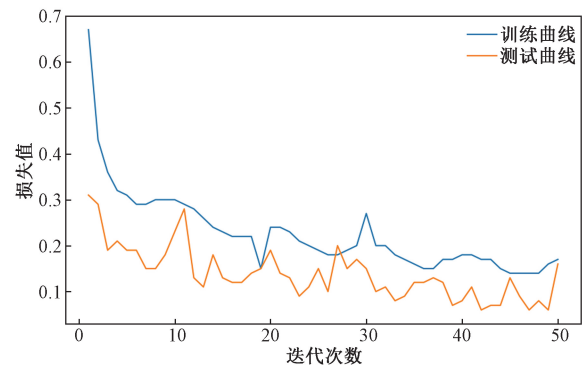


图 8 Loss 曲线

Fig. 8 Loss curve

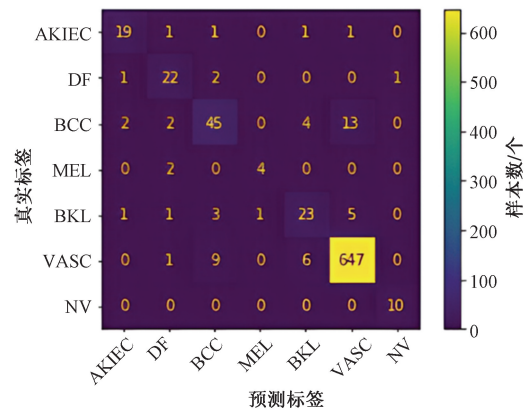


图 9 Rs-Capsnet 的混淆矩阵

Fig. 9 Confusion matrix of Rs-Capsnet

表 2 模型精度对比

Table 2 Comparison of model accuracy

模型	精确度/%	召回率/%	$F_1$ /%
Capsnet <sup>[6]</sup>	84.76	78.48	73.06
ResNeXt-50 <sup>[12]</sup>	87.56	77.25	75.88
VGG-Capsnet <sup>[13]</sup>	85.71	79.70	74.48
CBAM-Capsnet <sup>[9]</sup>	90.45	65.57	76.03
EfficientNet-B4 <sup>[7]</sup>	87.91	78.75	79.86
EfficientNet2 multi-scale <sup>[15]</sup>	91.20	89.71	91.30
SCSO-ResNet50-EHS-CNN <sup>[4]</sup>	91.54	91.44	91.40
Rs-Capsnet	92.95	81.04	81.00

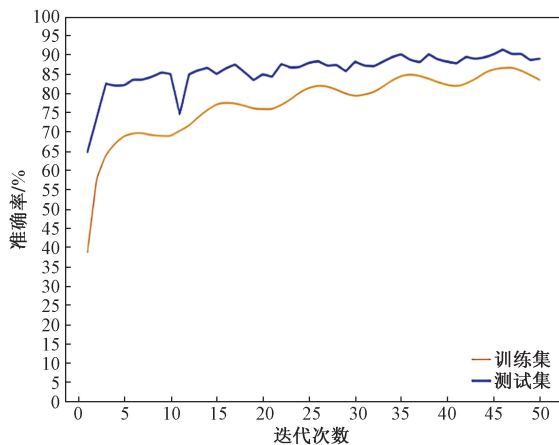


图 7 Accuracy 曲线

Fig. 7 Accuracy curve

表 3 模型参数量

Table 3 Number of model parameters

模型	参数量/M
Capsnet <sup>[6]</sup>	231.925 504
ResNeXt-50 <sup>[12]</sup>	23.522 375
VGG-Capsnet <sup>[13]</sup>	179.505 536
CBAM-Capsnet <sup>[9]</sup>	231.933 795
EfficientNet-B4 <sup>[7]</sup>	17.451 359
Rs-Capsnet	182.546 240

### 3.5 消融实验

为了验证 CBAM 模块和改进 squash 函数对 Rs-Capsnet 的有效性, 设置表 4 所示的消融实验, 在进行实验时, 基础网络架构由 ResNeXt 做特征提取器, 将提取到的特征输入到胶囊网络中。表 4 中模型为分别插入 CBAM 模块和改进的 squash 函数计算各个模块在网络中的表现。

从表 4 中可以看出, CBAM 模块和 squash 函数对模型的性能有着不同的贡献, 只采用 CBAM 模块进行实验时, 模型的精确度减少了 2.5%, 而召回率大幅下降, 减少了 15.47%; 当模型只使用改进的 squash 函数时, 模型的召回率和特异性也有了明显的下降。从实验结果可以看出, CBAM 和改进的 squash 函数的引入对于该模型来说具有可行性。

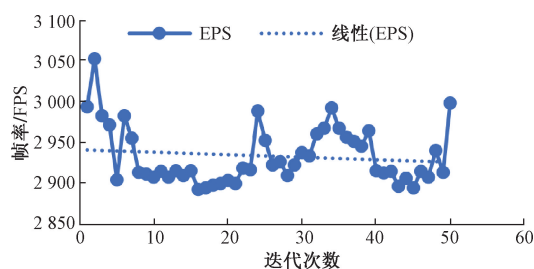
为了进一步验证该模型的实时性, 计算 Rs-Capsnet 在测试集中处理皮肤病变图像的帧率, 结果如图 10 所示, 可以看出, 在整个迭代中, 帧率在 2 940 FPS, 说明本文模型在进行图像处理和分类时有很好的实时性。

在胶囊网络中, 生成重建图像可以直观地看到模型学习到的特征和表示, 尤其是在皮肤病变图像

表 4 消融实验

Table 4 Ablation experiments

模型	精确度/%	召回率/%	特异性/%	$F_1$ /%
Rs-Capsnet	92.95	81.04	79.45	81.00
+ CBAM	90.45	65.57	77.15	76.03
+ $s_0$	88.16	62.27	63.26	67.13
+ $s_2$	91.18	76.79	69.77	77.41



虚线表示帧率的线性预测

图 10 帧率变化图

Fig. 10 Frame rate change graph

中, 捕捉到病变区域对分类结果至关重要<sup>[27-28]</sup>, 所以选择可视化重建图像, 以此来反映模型是否可以学习到输入图像的重要特征。在图 11 中可以看出, 两种类别的皮肤病变图像, 通过模型训练在重建图像种更关注于病变区域, 进一步证明了该模型的实用性很好。

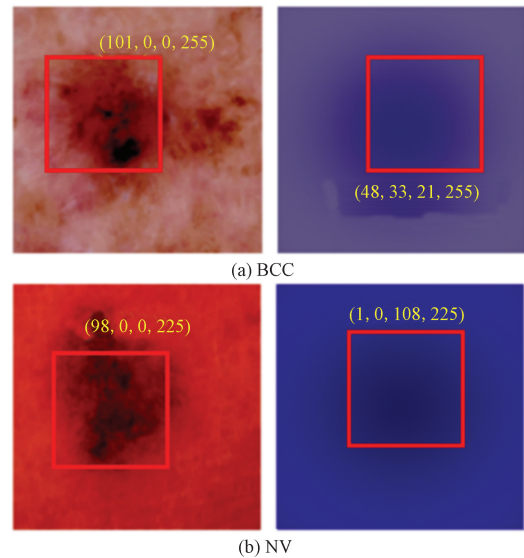


图 11 皮肤病变重建图像

Fig. 11 Skin lesion reconstruction image

## 4 结论

构建一种新的胶囊网络 Rs-Capsnet, 该模型将 ResNeXt 与 Capsnet 进行级联, 不同于传统的 Capsnet 网络架构, 有效提高了对 HAM10000 数据集的分类准确率。ResNeXt 结合残差网络 (Resnet) 和 Inception 模块的优势, 可以从深度和宽度两个方向同时学习特征的有用信息。Capsnet 可以捕捉特征的空间位置关系, 缓解 CNN 中由于池化丢失信息的缺点。二者的结合更充分的提取了特征信息, 优化了网络分类性能。在 HAM10000 数据集上 Rs-Capsnet 与传统的 Capsnet 相比, 精确度, 召回率和  $F_1$  分别提高了 8%、3% 和 8%, 表现了更佳的性能。在未来的工作中该网络对其他数据集的表现能力还有待验证, 同时还需要考虑如何减少胶囊网络路由过程的损耗。

### 参考文献

- [1] Hu W, Fang L, Ni R, et al. Changing trends in the disease burden of non-melanoma skin cancer globally from 1990 to 2019 and its predicted level in 25 years[J]. BMC Cancer, 2022, 22(1): 836-847.
- [2] Zhang J, Xie Y, Xia Y, et al. Attention residual learning for skin lesion classification[J]. IEEE Transactions on Medical Imaging,

- 2019, 38(9): 2092-2103.
- [3] Gomathi E, Jayasheela M, Thamarai M, et al. Skin cancer detection using dual optimization based deep learning network[J]. *Biomedical Signal Processing and Control*, 2023, 18(84): 104968-104976.
- [4] Akilandasowmya G, Nirmaladevi G, Suganthi S U, et al. Skin cancer diagnosis: leveraging deep hidden features and ensemble classifiers for early detection and classification[J]. *Biomedical Signal Processing and Control*, 2024, 88: 105306-105315.
- [5] Karar A, Shaikh Z A, Khan A A, et al. Multiclass skin cancer classification using EfficientNets: a first step towards preventing skin cancer[J]. *Neuroscience Informatics*, 2022, 2(4): 100034-100038.
- [6] Hinton G E, Sabour S, Frosst N. Dynamic routing between capsules[J]. *Advances in Neural Information Processing Systems*, 2017, 30(1): 3856-3866.
- [7] Pérez E, Ventura S. Melanoma recognition by fusing convolutional blocks and dynamic routing between capsules[J]. *Cancers*, 2021, 13(19): 4974-4991.
- [8] Wang P, Wang J, Li Y, et al. Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing[J]. *Biomedical Signal Processing and Control*, 2021, 65(1): 102341-102348.
- [9] Lan Z, Cai S, He X, et al. FixCaps: an improved capsules network for diagnosis of skin cancer[J]. *IEEE Access*, 2022, 10(8): 76261-76267.
- [10] Gocer E. Classification of skin cancer using adjustable and fully convolutional capsule layers[J]. *Biomedical Signal Processing and Control*, 2023, 18(85): 104949-104951.
- [11] Zeng R, Qin Y S, Song Y. A non-iterative capsule network with interdependent agreement routing[J]. *Expert Systems with Applications*, 2024, 34(238): 122284.
- [12] Xie S, Girshick R, Dollar P, et al. Aggregated residual transformations for deep neural networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 1492.
- [13] Bushara A R, Kumar R S V, Kumar S S. An ensemble method for the detection and classification of lung cancer using computed tomography images utilizing a capsule network with Visual Geometry Group[J]. *Biomedical Signal Processing and Control*, 2023, 18(85): 104930-104935.
- [14] Swetha R N, Shrivastava V K, Parvathi K. Multiclass skin lesion classification using image augmentation technique and transfer learning models[J]. *International Journal of Intelligent Unmanned Systems*, 2021, 36(8): 2047-2055.
- [15] Wei M, Wu Q, Ji H, et al. A skin disease classification model based on DenseNet and ConvNeXt fusion[J]. *Electronics*, 2023, 12(2): 438-458.
- [16] 刘光辉, 占华, 孟月波, 等. 多尺度显著特征双线注意力细粒度分类方法[J]. *计算机辅助设计与图形学学报*, 2023, 35(11): 1683-1691.  
Liu Guanghui, Zhan Hua, Meng Yuebo, et al. Fine-grained classification method of two-line attention for multi-scale salient features[J]. *Journal of Computer-Aided Design and Graphics*, 2023, 35(11): 1683-1691.
- [17] 蔡松柏. 基于 FixCaps 的皮肤癌辅助诊断方法研究[D]. 重庆: 重庆交通大学, 2023.  
Cai Songbai. Research on auxiliary diagnosis method of skin cancer based on FixCaps[D]. Chongqing: Chongqing Jiaotong University, 2023.
- [18] Swetha R N, Shrivastava V K, Parvathi K. Multiclass skin lesion classification using image augmentation technique and transfer learning models[J]. *International Journal of Intelligent Unmanned Systems*, 2024, 12(2): 220-228.
- [19] 朱子豪, 宋燕. 融合注意力和胶囊池化的轻量型胶囊网络[J]. *电子科技*, 2024, 37(5): 1-8, 31.  
Zhu Zihao, Song Yan. Lightweight capsule network fusing attention and capsule pooling[J]. *Electronic Science and Technology*, 2024, 37(5): 1-8, 31.
- [20] Freitas N R, Vieira P M, Cordeiro A, et al. Detection of bladder cancer with feature fusion, transfer learning and capsnet[J]. *Artificial Intelligence in Medicine*, 2022, 33(126): 102275.
- [21] Satapathy S C, Cruz M, Namburu A, et al. Skin cancer classification using convolutional capsule network (CapsNet)[J]. *Journal of Scientific & Industrial Research*, 2020, 79(11): 994-1001.
- [22] Magdy A, Hussein H, Abdel-Kader R F, et al. Performance enhancement of skin cancer classification using computer vision[J]. *IEEE Access*, 2023, 11: 72120-72133.
- [23] Deepika J, Rajan C, Senthil T. Improved CAPSNET model with modified loss function for medical image classification[J]. *Signal, Image and Video Processing*, 2022, 16(8): 2269-2277.
- [24] Lei Y, Wu Z, Li Z, et al. BP-CapsNet: an image-based deep learning method for medical diagnosis[J]. *Applied Soft Computing*, 2023, 23(146): 110683-110690.
- [25] Zhang J, Yu X, Lei X, et al. A novel CapsNet neural network based on MobileNetV2 structure for robot image classification[J]. *Frontiers in Neurorobotics*, 2022, 16(9): 1007939-1007949.
- [26] Wang K, He R, Wang S, et al. The efficient-CapsNet model for facial expression recognition[J]. *Applied Intelligence*, 2023, 53(13): 16367-16380.
- [27] Jia X, Li J, Zhao B, et al. Res-CapsNet: residual capsule network for data classification[J]. *Neural Processing Letters*, 2022, 54(5): 4229-4245.
- [28] Yadav S, Dhage S. TE-CapsNet: time efficient capsule network for automatic disease classification from medical images[J]. *Multimedia Tools and Applications*, 2023, 29(11): 1-30.