

基于机器学习算法的乳腺超声图像分类研究

汪俊豪¹, 闫士举^{1*}, 魏高峰^{2*}, 董天峰¹, 张涛¹, 邓爽¹

¹上海理工大学健康科学与工程学院, 上海

²海军军医大学海军医学系, 上海

收稿日期: 2023年1月9日; 录用日期: 2023年2月9日; 发布日期: 2023年2月16日

摘要

乳腺癌已成为世界上最常见的癌症, 对适龄女性进行乳腺筛查尤为重要。超声图像筛查在乳腺癌的筛查中占有主要地位。计算机辅助诊断技术可以帮助放射科医生实现快速和有效的乳腺超声图像诊断。因此, 我们提出了一种全流程乳腺超声诊断系统。主要方法如下: 首先基于U-Net网络对超声乳腺图像进行分割, 并根据分割结果将图像划分为不同的病理区域, 然后提取不同病理区域的特征并按照一定的规律得到12组特征。这12个特征组在4个不同的分类器上进行了测试, 最终选出了最佳特征组和最佳分类器。结果表明, 特征组NO. 12在LightGBM分类器上的性能最好。在测试集中, 最终的准确度、敏感性和特异性分别为0.97、0.987和0.938, AUC达到0.962, 这对于放射科医师提高诊断效率和准确性具有重要意义。

关键词

乳腺肿瘤, 超声图像, U-Net, 形状特征, 纹理特征, LightGBM

Classification Study of Breast Ultrasound Images Based on Machine Learning Algorithm

Junhao Wang¹, Shiju Yan^{1*}, Gaofeng Wei^{2*}, Tianfeng Dong¹, Tao Zhang¹, Shuang Deng¹

¹School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai

²Naval Medical Department, Naval Medical University, Shanghai

Received: Jan. 9th, 2023; accepted: Feb. 9th, 2023; published: Feb. 16th, 2023

Abstract

Breast cancer has become the most common cancer in the world. It is particularly important to

*通讯作者。

文章引用: 汪俊豪, 闫士举, 魏高峰, 董天峰, 张涛, 邓爽. 基于机器学习算法的乳腺超声图像分类研究[J]. 软件工程与应用, 2023, 12(1): 30-44. DOI: 10.12677/sea.2023.121004

provide breast screening for women of appropriate age, and the use of ultrasound during general screening is dominant, supplemented by mammography. Computer-aided diagnostic techniques can assist radiologists in achieving rapid and efficient diagnosis of breast ultrasound images. Therefore, we propose a full-flow breast ultrasound diagnostic system. Firstly, the ultrasound breast images are segmented based on the improved U-Net network framework to obtain breast tumor regions, and the images are divided into different pathological regions according to the segmentation results, and then the features of different pathological regions are extracted and the features are combined in different permutations to obtain 12 groups of features. These 12 feature groups were tested on four different classifiers, and the best feature group and the best classifier LightGBM (Light Gradient Boosting Machine) were finally selected. The results show that feature group NO. 12 has the best performance on the LightGBM classifier. In the testing set, the final accuracy, sensitivity and specificity were 0.97, 0.987 and 0.938, respectively. The AUC reached 0.962. This is of significance for radiologists to improve the efficiency and accuracy of diagnosis.

Keywords

Breast Tumor, Ultrasound Image, U-Net, Shape Features, Texture Features, LightGBM

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

根据世界卫生组织国际癌症研究机构(GLOBOCAN)最新的统计显示,乳腺癌已经成为世界上女性群体中最常见的癌症之一。2020年新增乳腺癌病例约200万例, 占有所有癌症病例的11.7%, 但是乳腺癌的死亡率仅占6.9% [1]。乳腺癌的早发现和早治疗, 能够最大限度地减少诊断延误的时间, 这对降低乳腺癌的致死率至关重要。

乳腺癌筛查主要是通过超声筛查和钼靶筛查, 但钼靶是基于X线摄影的对乳腺内的病灶进行检查, X光是根据不同乳腺组织的密度差异进行诊断的[2]。有报告显示, 因为“白色”乳腺组织与乳房X光检查中显示的“白色”乳腺癌组织之间缺乏一定的对比[3], 而超声可检测出乳房X线照片中难以发现的乳腺癌[4], 可提升乳腺组织致密者中的乳腺癌检测的准确率[5], X光在诊断这类患者方面具有一定的局限性[6]。在乳腺组织密度较小的患者中, 超声图像检测也要优于X光[7]。此外超声检测相对于X光检测无辐射, 而且超声检测是实时动态的, 以便于观察图像, 确定病情。

近年来, 计算机辅助诊断技术(Computer Aided Diagnosis, CAD) [8]在医学图像领域取得了飞速发展。他在提高诊断准确性、减少漏诊、提高工作效率等方面发挥了重大作用。

机器学习方法已被用于医学领域的检测和诊断, 其中支持向量机是最常用的机器学习算法之一, 通过提取乳房X线图像ROI区域的灰度共生矩阵(GLCM)纹理特征进行分类, 支持向量机的准确率可达94% [9]。通过提取乳腺超声图像的一阶统计纹理特征进行分类, 支持向量机分类的总体准确率为86.6% [10]。在机器学习中另一个常用的分类器是决策树, 它也是一种基本的分类方法。在某一心脏病识别任务中, 分别使用了三种算法: Decision Tree、Bagging with Decision Tree 和 Bagging with Naive Bayes, 三种模型的准确率分别为78.93%、79.53%和82.50% [11]。梯度提升算法是近几年预测模型中最强大的机器学习算法。文献[12]中使用梯度提升决策树-多参数优化器(Gradient boosting decision tree - Multi-verse optimizer, GBDT-MVO)模型证明了它在乳腺癌诊断中具有更高的精度和较低的方差。

深度学习的相关算法在医学图像处理任务上也发挥了重大作用。FCN [13]是图像分割的里程碑，具有全监督的全卷积神经网络(FCNN)在肝脏图像的血管分割指数 DICE 中达到 0.79 [14]。U-Net [15]网络的提出是对医学图像语义分割的重大贡献，同时也激发了大量研究人员对 U 型分割网络的思考。受 U-Net 及其变体的启发，所提出的 NAS-UNET 由堆叠在 U-Type 主干上相同数量的 DownSC 和 UpSC 组成，其结果在三种不同类型的医学图像数据集上表现出良好的分割性能[16]。有很多新的分割网络都是在 U-Net 结构的基础上改进的，但正如 nnU-Net [17]文中提到的那样，这些被称为 SOTA 的网络结构大都是基于 U-Net 结构改进而来的，如 nnU-Net，一些网络结构如 SOTA 都是基于 U-Net 结构而针对特定数据集进行特定的调整参数，而不是本质上的改进。因此，U-Net 在医学图像分割领域中仍具有不可动摇的地位。

在本文中，我们基于 U-Net 框架对整个超声图像进行分割，并计算区域的肿瘤形状特征和纹理特征。使用四种不同的机器学习算法对这些功能进行了重组和验证。在验证结果中，选择最好的特征和分类的模型来评估测试集(图 1)。

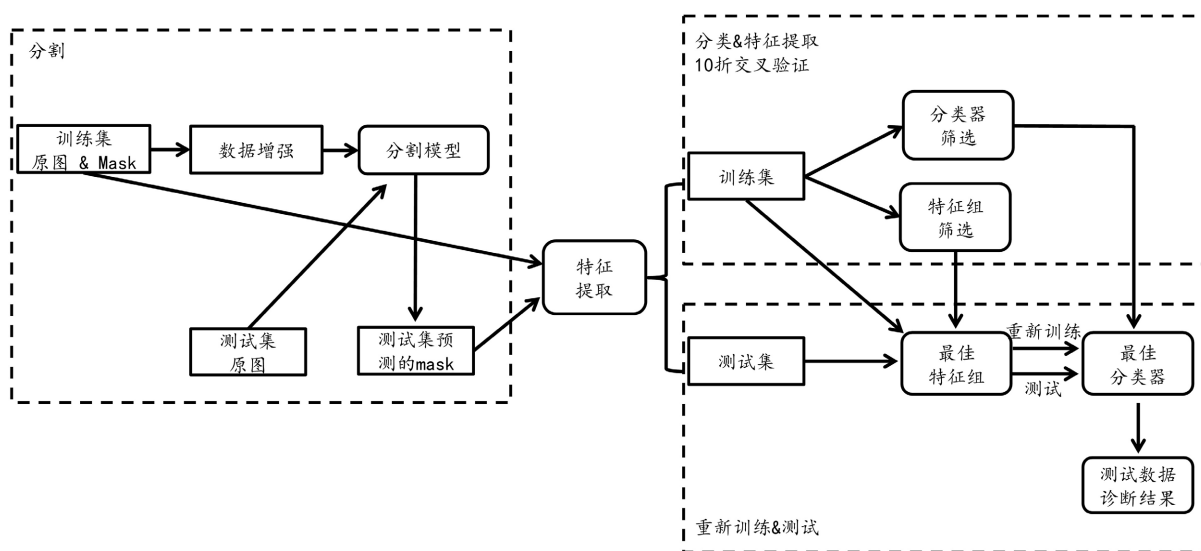


Figure 1. Breast ultrasound diagnostic model

图 1. 乳腺超声诊断模型

2. 研究方法与相关理论

2.1. LightGBM [18] (Light Gradient Boosting Machine)算法

LightGBM 是一种用于机器学习的快速梯度提升算法。它主要用于分类和回归任务。

LightGBM 的优势在于，它可以处理大规模数据集，并且在训练过程中使用多种优化技术，使得它比其他算法[19] [20]更快。它还能够处理稠密数据和稀疏数据，并且具有较高的准确性。

2.2. U-Net 网络

UNet 是一种适用于图像分割的深度学习模型。UNet 的结构非常适合处理医学图像分割任务。它的编码器负责从图像中提取有用的特征，而解码器则负责将这些特征转换为最终的分割图像。UNet 的跳连接层，可帮助网络更好地学习图像中的细节。这些特点使得 UNet 能够在保留图像细节的同时准确地对图像进行分割。此外，UNet 还具有较少的参数，这意味着它需要较少的训练数据来调整模型。这在医学图像领域中尤为重要，因为医学图像数据常常是非常有限的。总的来说，UNet 的准确性、细节保留能力和

较少的参数使它成为医学图像分割的理想选择。

2.3. 形状特征

根据医生的临床经验，肿瘤的形状特征是判断肿瘤是否为恶性的标准之一。因此，以下这些形状特征将被计算用于肿瘤分类。

椭圆相似率[21] (ESR): 最佳拟合椭圆[7]是通过直接最小二乘拟合椭圆算法获得的。如图 2 所示，深蓝线为肿瘤轮廓，黄线为最佳拟合椭圆。椭圆相似度是通过最佳拟合椭圆的周长与原始肿瘤轮廓周长的比率来计算的。如果原始肿瘤轮廓的形状特征规则，则与椭圆相似度的值更接近 1。椭圆相似度(ESR)定义如下， Pe 和 Pt 分别代表最佳拟合椭圆周长和肿瘤轮廓周长。

$$ESR = \frac{Pe}{Pt}$$

偏转角(DA): 肿瘤的偏转角是肿瘤轮廓最佳拟合椭圆长轴与水平方向的偏转角。如图 2 所示， θ 为肿瘤的偏转角。DA 定义如下。

$$DA = \begin{cases} \theta, \theta \in \left[0, \frac{\pi}{2}\right] \\ \pi - \theta, \left[\frac{\pi}{2}, \pi\right] \end{cases}$$

形状紧凑率(FCR): 致密性可以衡量肿瘤边缘的缺损程度。良性肿瘤的形状通常比恶性肿瘤更规则，缺损程度更低。形状紧凑度(FCR)定义如下， Ar 代表肿瘤的面积， C 代表肿瘤轮廓的周长。

$$FCR = 4\pi \frac{Ar}{C^2}$$

横向纵向比(TLR): 肿瘤的纵横比可以通过计算肿瘤的外接矩形的长高比来计算，外接矩形是通过最高点、最低点、最左点和最右点形成的矩形。如图 2 所示，红色虚线矩形区域为肿瘤的外接矩形。肿瘤横向纵向比(TLR)定义如下， T 代表外接矩形的长度， L 代表外接矩形的高度。

$$TLR = \frac{T}{L}$$

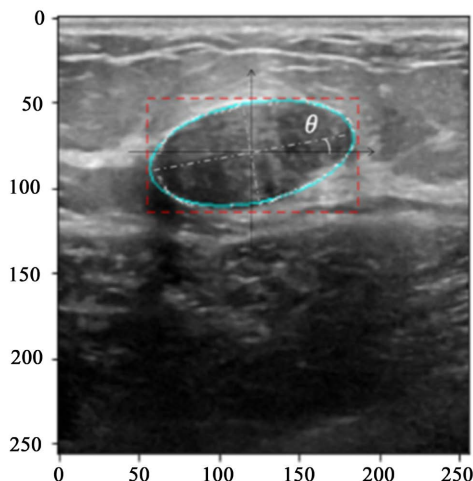


Figure 2. Different contour lines for different shape features
图 2. 不同形状特征的不同轮廓线

2.4. 纹理特征

灰度共生矩阵中的保存的数据量很大,一般不直接用作区分纹理特征,而是将基于灰度共生矩阵构建一些统计量用作纹理分类特征。哈拉利克[22] (Haralick)给出了些基于灰度共生矩阵计算的 14 个统计量:角二阶矩、对比度、相关性、方差、逆差矩、求和平均、求和方差、求和熵、熵、差方差、差熵、相关性 1 的信息度量、相关性 2 的信息度量、最大相关系数。在本研究中,选取 13 个统计量(不包括最大相关系数)作为提取的纹理特征。由于灰度共生矩阵中 θ 有四个方向,可以得到四组统计量,再加上四组统计量的平均值,可以得到五组统计量,共 65 个特征,即被用作本文提取的纹理特征。

3. 乳腺超声诊断模型的方法

3.1. 乳腺超声诊断模型概述

本文乳腺超声诊断模型如图 1 所示,先利用训练集训练一个 U-Net 分割模型,得到每个肿瘤区域的 mask,根据原图与对应的 mask 划分出信息一起进入特征提取模块,再进行特征组筛选和分类器评估,以得到最终的最佳模型。

3.2. 图像数据集

在研究中,数据主要来自数据集 DBUI (Data of Breast Ultrasound Images) [23]。该数据集包括从 600 名女性患者获得的 780 张乳房图像,所有图像的格式均为 PNG 类型。这些图像分为三类:正常、良性和恶性。本此研究选取 399 份良性肿瘤样本和 183 份恶性肿瘤样本作为实验数据。每个图像都有其对应的分割掩码(图 3)。

按照比例随机分为训练集和测试集,其中训练集用于训练分割模型(3 折交叉验证)和特征选择及分类器选择(10 折交叉验证)。在训练超声图像分割时的训练集是基于扩增的训练集数据进行训练,训练集数据还会进行特征选择和分类模型选择,然后使用测试集来对最终整体模型的结果进行评估。

3.3. 分割网络

在本研究中,分割的网络框架基于 U-Net 网络,参考了 nnU-Net 的思想,对原始数据进行了一些预处理。

对训练集使用随机旋转、水平翻转和垂直翻转等技术增强图像,在训练集上使用了 3 折交叉验证,损失函数结合了 Dice 损失和交叉熵损失。

3.4. 特征提取

特征提取流程图详见图 4。

1) 形状特征提取,基于分割网络的得到的肿瘤内部区域 mask 去计算形状特征,分别是椭圆相似率、偏转角、边缘紧凑率、肿瘤纵横比。

2) 纹理特征提取,纹理特征从以下 4 个不同的区域中提取(图 5)。**a** 区域是分割得到的 mask, **b** 区域是 mask 膨胀后减去原始 mask 得到的一个环形区域, **c** 区域是 mask 膨胀算法后的外矩形, **d** 区域是整个图像。这四个区域划定规则是由肿瘤内部由内向外依次扩大范围得到的区域。因为从临床评估上来说肯定是优先关注肿瘤区域再关注逐渐关注肿瘤周围区域,这四个区域的重要性在后续实验中会进一步验证。由于这四个区域的纹理特征均来自同一张超声图像,因此这四个区域的纹理特征将作为同一样本的纹理特征。

3) 降维:从每个超声图像中可以获得 264 (260 + 4)个特征。这些特征包括从四个不同区域获得的 4

个形状特征和4个不同区域获得的260(65×4)个纹理特征。降维特征组是对具有高相关性的冗余特征进行降维得到的。在本文中,将采用LASSO算法[24]对特征降维进行计算。

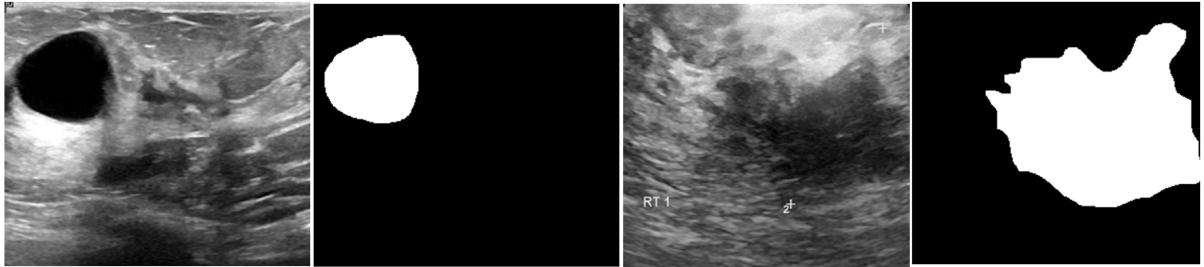


Figure 3. Benign and malignant ultrasound breast tumors and their corresponding masks

图3. 良性和恶性超声乳腺肿瘤及其相应掩码

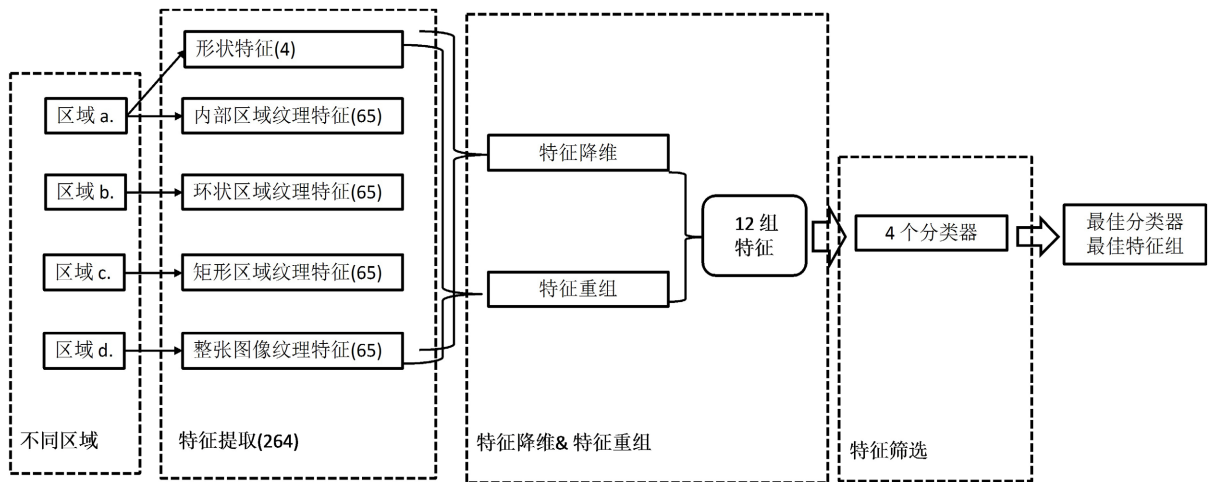


Figure 4. Feature extraction flowchart

图4. 特征提取流程图

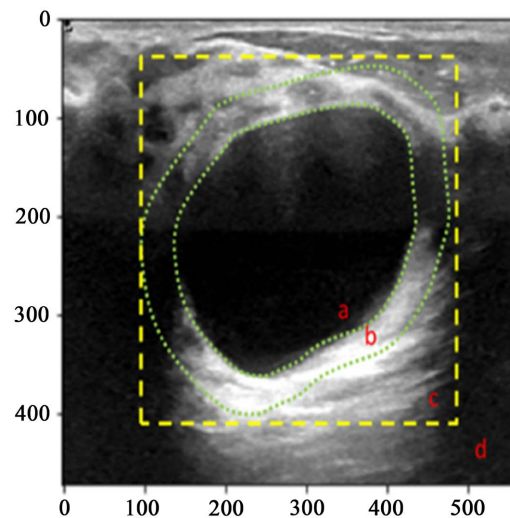


Figure 5. Different areas of breast ultrasound images

图5. 乳腺超声图像的不同区域

3.5. 特征重组

根据不同的区域和不同的类型对特征进行重新组合，得到不同的组合特征集，然后由分类器选择最佳特征集。

3.6. 分类器

在正如 2.1 节提到的我们在数据集上训练和测试 LightGBM (Light Gradient Boosting Machine) 分类器。我们还在同一数据集上训练和测试了其他 3 个分类器，将它们的性能与 LightGBM 分类器进行比较。3 种算法包括 K-最近邻[25] (K-Nearest Neighbor)，支持向量机(Support Vector Machines)，随机森林[26] (Random Forest)。

3.7. 评价指标

为了更全面地评估每个模型的性能，我们使用了如下几个模型评估指标。在分割任务中，利用 DSC (Dice Similarity Coefficient)、IOU (Intersection Over Union)、Recall、Precision、Global Accuracy 和 AUC 来评估分割性能。

使用 AUC、ACC、Sensitivity 和 Specificity 来评估分类性能。

4. 实验结果分析

4.1. 分割结果

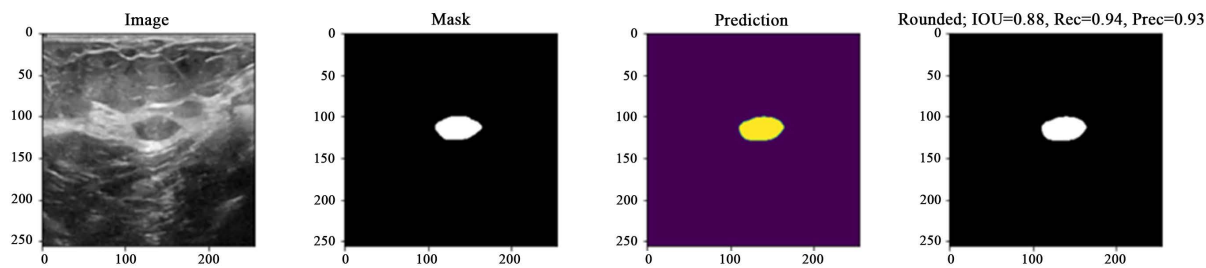
从数据集中总共选择了 582 张图像，并将它们分为良性和恶性两类，作为分割数据。其中随机划分 482 张训练集图像和 100 张测试集图像。为了达到更好的网络训练效果，我们通过水平翻转、垂直翻转、左旋转和右旋转来增加训练样本以得到 4338 个训练样本，并将它们与 U-Net 模型进行 3 折交叉验证。将性能最好的分割网络用于在测试集上进行测试。最终结果如下(表 1，图 6)。

从结果中我们发现，经过数据预处理后训练得到的模型的分割性能较之前提高了 10%左右。DSC、IOU 和 Precision 的最终结果分别为 0.778、0.703、0.899。

Table 1. Segmentation result

表 1. 分割结果

数据操作	评价指标					
	DSC	IOU	Recall	Precision	Global Accuracy	ROC AU
未扩增	0.666	0.547	0.687	0.796	0.960	0.904
扩增后	0.778	0.703	0.761	0.899	0.967	0.939



(a)

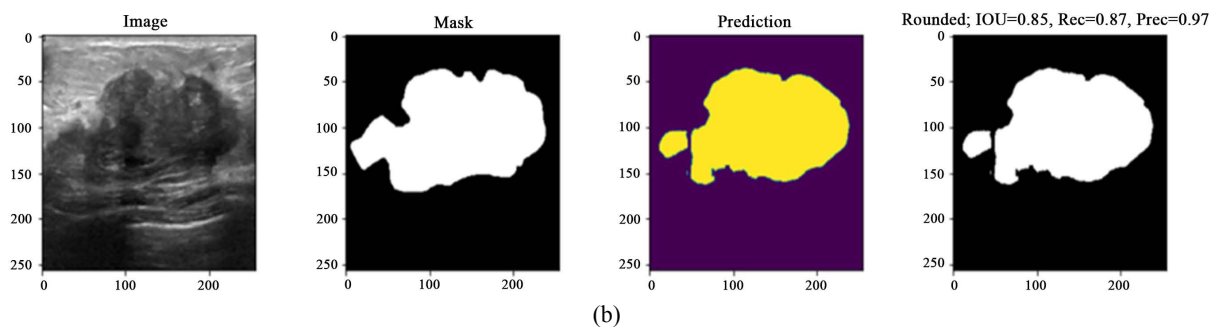


Figure 6. Segmentation results for benign and malignant tumors. (a) Benign tumor segmentation; (b) Benign tumor segmentation

图 6. 良性和恶性肿瘤的分割结果。(a) 良性肿瘤分割结果图示；(b) 恶性肿瘤分割结果图示

4.2. 特征选择结果

4.2.1. 特征重组

重新排列组合方式见表 3, 我们之前特征提取得到的 5 组特征, 就是对应表 3 中的特征组 NO.1-NO.5, 重组规则是按照肿瘤位置区域依次往外扩大区域, 再依次结合形状特征得到 6 组复合特征组就是对应表 3 中的特征组 NO.6-NO.11, 其中特征组 NO.11 是复合了全部 5 组特征。特征组 12 就是对特征组 11 在 LASSO 上降维得到的特征组(NO.12) (见表 2), 因此得到一共 12 组特征(见表 3), 我们并没有直接将降维特征组 NO.12 作为我们的最佳特征组。而是通过后续实验从这 12 特征组(表 3)中进一步验证从而选出最佳的特征组合。

4.2.2. 实验结果

上面 12 个不同的特征集在 4 个分类器中进行了 10 倍交叉验证。图 7 显示了 4 个分类器上 12 个特征组中的 ROC 曲线。表 4 是各特征组和分类器对应的具体指标性能。对 12 个特征组指标可以看出, 形状特征(特征 NO.1)明显优于其他单组特征(特征组 NO.2-NO.4), 而从单从纹理特征组(NO.2-NO.4)可以看出, 单独的内部纹理特征组(区域 a, 特征组 NO.2)性能最好, 环状纹理特征组(区域 b, 特征组 NO.3)性能其次, 另外两组纹理特征(区域 c、d, 特征组 NO.3、NO.4)性能指标可以看出没有明显区分性, 就复合特征组而言(NO.6-NO.11), 随着复合的特征越多维度越高, 对应的指标也是在提升的, 但与之对应的是会有很多冗余特征, 所以对应的经 LASSO 降维后的特征组(NO.12)性能优于其他任何复合组。对四种分类器的比较可以看出, LightGBM 算法性能明显优于其他三种分类器, 因此实验得出结论是最佳特征组为经 LASSO 降维后的特征组(NO.12), 最佳分类器为 LightGBM。

Table 2. Characteristic coefficients screened by LASSO algorithm

表 2. LASSO 算法筛选的特征系数

特征名称	系数
theta	-0.001010
DSR	-0.181094
CMP	-0.294333
DWR	-0.114881
Inner-0 Contrast	0.001756
Inner-0 Correlation	0.014504

Continued

Inner-0 Sum Average	-0.008083
Inner-45 Contrast	0.032217
Inner-45 Information Measure of Correlation 1	-0.009291
Inner-45 Information Measure of Correlation 2	0.003110
Rectangle-0 Sum Variance	-0.052775
Rectangle-45 Inverse Difference Moment	-0.007978
Rectangle-45 Entropy	0.011290
Rectangle-45 Information Measure of Correlation 2	0.017163
Rectangle-90 Entropy	0.012900
Ring-0 Inverse Difference Moment	0.006270
Ring-135 Difference Variance	0.008834
Ring-135 Information Measure of Correlation 2	-0.068414
Whole-0 Contrast	0.034904
Whole-0 Difference Variance	-0.002620
Whole-0 Information Measure of Correlation 1	0.017086
Whole-135 Correlation	0.036530
Whole-135 Information Measure of Correlation 2	0.006758
Total	24

Table 3. Feature recombination information

表 3. 特征重组信息

Features Combination Group NO.	Shape Features (4)	Textural Features (65 × 4)				Dimensionality Reduction Features (24)
		Inner Area (65)	Ring Area (65)	Rectangular Area (65)	Whole Area (65)	
1	√					
2		√				
3			√			
4				√		
5					√	
6		√	√			
7		√	√	√		
8	√	√				
9	√	√	√			
10	√	√	√	√		
11	√	√	√	√	√	
12						√

Table 4. Performance of 12 sets of features on 4 classifiers
表 4. 12 组特征在 4 个分类器上的性能表现

特征组	模型	AUC	准确率	灵敏度	特异度
NO.1	KNN	0.635	0.717	0.855	0.416
	SVM	0.872	0.902	0.952	0.792
	RF	0.94	0.959	0.99	0.89
	LightGBM	0.956	0.966	0.982	0.929
NO.2	KNN	0.669	0.725	0.82	0.518
	SVM	0.653	0.759	0.937	0.369
	RF	0.869	0.895	0.94	0.797
	LightGBM	0.857	0.885	0.932	0.781
NO.3	KNN	0.613	0.679	0.792	0.434
	SVM	0.59	0.716	0.93	0.251
	RF	0.798	0.845	0.925	0.671
	LightGBM	0.817	0.849	0.902	0.732
NO.4	KNN	0.608	0.682	0.807	0.409
	SVM	0.514	0.681	0.962	0.066
	RF	0.72	0.787	0.9	0.539
	LightGBM	0.742	0.8	0.9	0.583
NO.5	KNN	0.526	0.634	0.817	0.235
	SVM	0.5	0.682	0.99	0.011
	RF	0.63	0.728	0.894	0.365
	LightGBM	0.653	0.73	0.862	0.444
NO.6	KNN	0.709	0.751	0.822	0.596
	SVM	0.759	0.819	0.922	0.595
	RF	0.867	0.899	0.952	0.781
	LightGBM	0.898	0.921	0.96	0.836
NO.7	KNN	0.717	0.765	0.845	0.589
	SVM	0.832	0.866	0.925	0.738
	RF	0.878	0.912	0.97	0.786
	LightGBM	0.905	0.926	0.962	0.847
NO.8	KNN	0.665	0.723	0.822	0.508
	SVM	0.691	0.78	0.93	0.452
	RF	0.911	0.933	0.97	0.852
	LightGBM	0.961	0.971	0.988	0.934
NO.9	KNN	0.709	0.751	0.822	0.596
	SVM	0.782	0.828	0.905	0.659
	RF	0.901	0.928	0.972	0.83
	LightGBM	0.955	0.969	0.992	0.918

Continued

NO.10	KNN	0.717	0.765	0.845	0.589
	SVM	0.793	0.843	0.927	0.659
	RF	0.875	0.911	0.97	0.78
	LightGBM	0.959	0.972	0.995	0.923
NO.11	KNN	0.726	0.768	0.84	0.612
	SVM	0.797	0.839	0.91	0.685
	RF	0.89	0.921	0.972	0.808
	LightGBM	0.961	0.973	0.992	0.929
NO.12	KNN	0.635	0.72	0.865	0.405
	SVM	0.855	0.89	0.95	0.759
	RF	0.916	0.945	0.992	0.84
	LightGBM	0.964	0.973	0.988	0.94

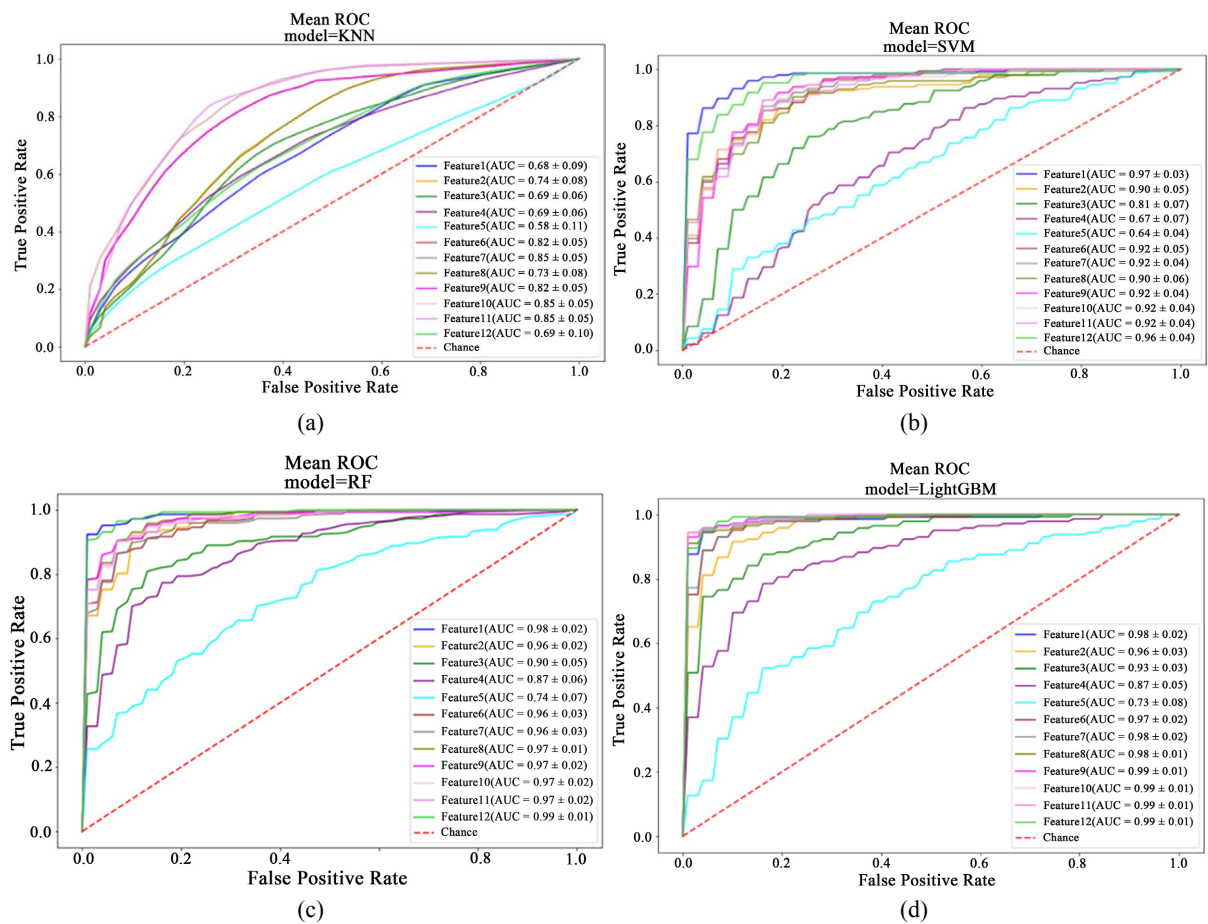


Figure 7. The average ROC curve obtained by 10-fold cross-validation of 12 sets of features on 4 classifiers. (a) KNN ROC curve; (b) SVM ROC curve; (c) RFROC curve; (d) LightGBMROC curve

图 7. 12 组特征在 4 个分类器上通过 10 折交叉验证得到的平均 ROC 曲线。(a) KNN ROC 曲线; (b) SVM ROC 曲线; (c) RF ROC 曲线; (d) LightGBM ROC 曲线

4.3. 训练和测试

我们将数据集按照 7:3 的比例划分为训练集和测试集,通过对训练集进行 10 倍交叉验证得到最佳特征组和最佳分类器(这些工作都是在训练集上完成的)。最终模型需要重新训练和测试以保证客观性。首先,从我们提取的所有特征中选择最佳特征组(降维特征组 12)作为数据特征,用最好的分类器(LightGBM)用训练集数据重新训练得到模型,然后对模型用测试集进行测试,最终测试结果如表 5 所示。最终结果的准确度、灵敏度和特异性分别为 0.966,分别为 0.98、0.938。

4.4. 与深度学习相关算法对比

为了进一步验证我们所提出的模型的性能优势,我们选取了常用的几种深度学习模型进行对比。这几组深度学习初始权重都是在 ImageNet 上进行预训练得到的,在我们的划分的乳腺超声图像训练集上训练得到最佳参数,再在测试集上测得性能指标见表 6。可以看出,我们所提出的方法性能指标明显高于深度学习各个算法。

Table 5. Performance of the model on the test set

表 5. 模型在测试集上的表现

分类器	特征组	权重均衡	AUC	准确率	灵敏度	特异度
LightGBM	NO.12	否	0.959	0.966	0.980	0.938
		是	0.962	0.970	0.987	0.938

Table 6. Compare with the deep learning models

表 6. 与深度学习各个模型对比

模型	ACC	Precision	Recall
ResNet101 [27]	0.920	0.915	0.915
ResNet34 [27]	0.910	0.885	0.905
VGG_16 [28]	0.900	0.900	0.890
VGG_19 [28]	0.870	0.855	0.865
GoogLeNet [29]	0.910	0.900	0.900
SqueezeNet [30]	0.900	0.900	0.865
DenseNet [31]	0.880	0.855	0.890
EfficientNet [32]	0.851	0.852	0.799
SwinTransformer [33]	0.795	0.778	0.732
ViT [34]	0.821	0.857	0.735
本文模型	0.966	0.968	0.980

5. 讨论

在此项中,良性样本与恶性样本的比例接近 2:1,存在数据不平衡的问题。解决分类任务中的数据不平衡问题主要有三种方法[35]: 1) 过采样或欠采样,通过增加少数类样本或减少多数类样本。2) 采用代

价敏感型学习方法该方法根据不同类别的错误分类样本数去重新定义损失函数。3) 集成学习。该方法将集成学习方法与数据重采样技术相结合。在本文中, 我们尝试按照训练集类比例给样本添加权重, 结果表明增加权重后各项指标确实均有一定提升(参考表 5 中均衡权重组)。

与传统算法相比, 深度学习可以更好地学习到样本数据的内在规则。深度学习可能需要更大的数据量以支持系统进行训练模型。然而在医学领域, 很难获得大数据集, 只有小样本数据集可用, 因此在本文中, 我们仍然使用传统的算法(LightGBM)作为分类器并取得了不错的效果。

在医学图像领域中, 前处理和后处理的效果有时会大于网络本身的变换。从不同分类器下不同特征的结果可以发现, 与其他特征组相比, 大部分分类器对于具有单个形状的特征组都具有较好的分类效果。内部区域和环形区域的纹理也表现出一定的分类性能, 但单独不同区域的纹理特征并没有表现出优异的性能, 原因可能是图像的纹理特征与设备的类型有关。从实验结果可以看出, 从医学图像中提取临床病理特征对最终诊断提升是有一定意义和价值的, 如果基于各种图像算法盲目地提取大量特征可能会在一定程度上对结果有一定的增强, 但会大大降低临床可解释性。

6. 结语

在本文中, 我们提出了一种全流程的乳腺超声诊断系统, 结果表明我们所提方法准确率达到 0.966, 灵敏度达到 0.98, 特异度达到 0.93。可以实现乳腺超声图像良恶更高精度的诊断, 帮助医生提高超声乳腺图像诊断的效率和准确性。具有一定的研究价值和临床意义。

利益冲突声明

本文全体作者均声明不存在利益冲突。

参考文献

- [1] Sung, H., Ferlay, J., Siegel, R.L., *et al.* (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, **71**, 209-249. <https://doi.org/10.3322/caac.21660>
- [2] Winkler, N.S., Raza, S., Mackesy, M. and Birdwell, R.L. (2015) Breast Density: Clinical Implications and Assessment Methods. *Radiographics*, **35**, 316-324. <https://doi.org/10.1148/rg.352140134>
- [3] Machida, Y., Tozaki, M., Shimauchi, A. and Yoshida, T. (2015) Breast Density: The Trend in Breast Cancer Screening. *Breast Cancer*, **22**, 253-261. <https://doi.org/10.1007/s12282-015-0602-2>
- [4] Linver, M.N. (2008) 4-19 Mammographic Density and the Risk and Detection of Breast Cancer. *Breast Diseases*, **18**, 364-365. [https://doi.org/10.1016/S1043-321X\(07\)80400-0](https://doi.org/10.1016/S1043-321X(07)80400-0)
- [5] Brem, R.F., Lenihan, M.J., Lieberman, J. and Torrente, J. (2015) Screening Breast Ultrasound: Past, Present, and Future. *American Journal of Roentgenology*, **204**, 234-240. <https://doi.org/10.2214/AJR.13.12072>
- [6] Thigpen, D., Kappler, A. and Brem, R. (2018) The Role of Ultrasound in Screening Dense Breasts—A Review of the Literature and Practical Solutions for Implementation. *Diagnostics*, **8**, 1-14. <https://doi.org/10.3390/diagnostics8010020>
- [7] Chen, H.L., Zhou, J.Q., Chen, Q. and Deng, Y.C. (2021) Comparison of the Sensitivity of Mammography, Ultrasound, Magnetic Resonance Imaging and Combinations of These Imaging Modalities for the Detection of Small (≤ 2 cm) Breast Cancer. *Medicine (Baltimore)*, **100**, e26531. <https://doi.org/10.1097/MD.0000000000026531>
- [8] Halalli, B. and Makandar, A. (2018) Computer Aided Diagnosis—Medical Image Analysis Techniques. In: *Breast Imaging*, InTech, Rijeka, 1854-1858. <https://doi.org/10.5772/intechopen.69792>
- [9] Pharswan, R. and Singh, J. (2020) Performance Analysis of SVM and KNN in Breast Cancer Classification: A Survey. Springer International Publishing, Berlin. https://doi.org/10.1007/978-3-030-39119-5_9
- [10] Jain, S. and Virmani, J. (2020) Texture Ratio Vector Technique for the Classification of Breast Lesions Using SVM. Springer, Singapore. https://doi.org/10.1007/978-981-15-2740-1_14
- [11] Tu, M.C., Shin, D. and Shin, D. (2009) A Comparative Study of Medical Data Classification Methods Based on Decision Tree and Bagging Algorithms. 2009 *Eighth IEEE International Conference on Dependable, Autonomic and Se-*

- cure Computing*, Chengdu, 12-14 December 2009, 183-187.
- [12] Tabrizchi, H., Tabrizchi, M. and Tabrizchi, H. (2020) Breast Cancer Diagnosis Using a Multi-Verse Optimizer-Based Gradient Boosting Decision Tree. *SN Applied Sciences*, **2**, 1-19. <https://doi.org/10.1007/s42452-020-2575-9>
- [13] Shelhamer, E., Long, J. and Darrell, T. (2017) Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 640-651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- [14] Mishra, D., Chaudhury, S., Sarkar, M. and Soin, A.S. (2019) Ultrasound Image Segmentation: A Deeply Supervised Network with Attention to Boundaries. *IEEE Transactions on Biomedical Engineering*, **66**, 1637-1648. <https://doi.org/10.1109/TBME.2018.2877577>
- [15] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, 5-9 October 2015, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [16] Weng, Y., Zhou, T., Li, Y. and Qiu, X. (2019) NAS-Unet: Neural Architecture Search for Medical Image Segmentation. *IEEE Access*, **7**, 44247-44257. <https://doi.org/10.1109/ACCESS.2019.2908991>
- [17] Isensee, F., Jaeger, P.F., Kohl, S.A.A., et al. (2021) nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation. *Nature Methods*, **18**, 203-211. <https://doi.org/10.1038/s41592-020-01008-z>
- [18] Ke, G., Meng, Q., Finley, T., et al. (2017) LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 3147-3155.
- [19] Friedman, J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, **29**, 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- [20] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [21] Fitzgibbon, A.W., Pilu, M. and Fisher, R.B. (1996) Direct Least Squares Fitting of Ellipses. *Proceedings—International Conference on Pattern Recognition*, Vol. 1, 253-257. <https://doi.org/10.1109/ICPR.1996.546029>
- [22] Dinstein, I., Shanmugam, K. and Haralick, R.M. (1973) Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-3**, 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>
- [23] Al-Dhabyani, W., Gomaa, M., Khaled, H. and Fahmy, A. (2020) Dataset of Breast Ultrasound Images. *Data in Brief*, **28**, Article ID: 104863. <https://doi.org/10.1016/j.dib.2019.104863>
- [24] Robert, T. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, **58**, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [25] Peterson, L. (2009) K-Nearest Neighbor. *Scholarpedia*, **4**, 1883. <https://doi.org/10.4249/scholarpedia.1883>
- [26] 王楠, 王远军, 廉朋. 基于影像组学的直肠癌术前 T 分期预测[J]. 波谱学杂志, 2022, 39(1): 43-55.
- [27] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [28] Simonyan, K. and Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015*, San Diego, 7-9 May 2015, 1-14.
- [29] Szegedy, C., et al. (2015) Going Deeper with Convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 1-9.
- [30] Iandola, F.N., Han, S., Moskewicz, M.W., et al. (2016) SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and < 0.5 MB Model Size. <http://arxiv.org/abs/1602.07360>
- [31] Huang, G., Liu, Z., Laurens, V., et al. (2016) Densely Connected Convolutional Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2261-2269. <https://doi.org/10.1109/CVPR.2017.243>
- [32] Tan, M. and Le, Q.V. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *36th International Conference on Machine Learning, ICML 2019*, Long Beach, 10-15 June 2019, 10691-10700.
- [33] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S. and Guo, B. (2021) Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. *Proceedings of the IEEE International Conference on Computer Vision*, Montreal, 11-17 October 2021, 9992-10002. <https://doi.org/10.1109/ICCV48922.2021.00986>
- [34] Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021) An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *International Conference on Learning Representations*, Sydney, 30-31 August 2021, 1-21.

- [35] Jing, X.-Y., Zhang, X., Zhu, X., *et al.* (2021) Multiset Feature Learning for Highly Imbalanced Data Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **43**, 139-156.
<https://doi.org/10.1109/TPAMI.2019.2929166>