

融合文本分类算法的皮肤病辅助诊疗模型

凌天, 诸佳珍, 焦阳, 李露芳

浙江中医药大学图书馆, 浙江 杭州 310053

【摘要】针对当前皮肤病辅助诊断中生物学特征建模规模较小且耗费巨大人工成本,而患者疾病特征的时间序列同样无法准确描述等难点,本研究运用融合文本分类算法,融合常用的文本分类模型TextLSTM、TextCNN、RCNN得到皮肤疾病辅助诊疗模型(TLNN模型),通过提取图像传感器医学特征向量化后进行预处理减少焦块数量以及消除偏差较大的特征信息,提高决策数据精度。在ISIC2018和PH2数据集进行对照实验,TLNN模型的准确率为72.36%,高于其余3种文本分类模型。在与医生主观诊断对比实验中,模型诊断准确率为92%,接近于医生94%的平均准确率,而有效诊断效率(1.17 min/例)明显高于医生人工诊断(4.57 min/例),整体效率提升幅度达290%,结果表明对比传统人工诊断,融合文本分类算法模型能以更短时间获得精确的诊断。TLNN模型可以应用于疾病诊断,辅助医生医疗决策,为患者提供优质便捷的智能诊疗服务。

【关键词】皮肤病;辅助诊断;融合文本分类算法;D-S证据理论;医学特征

【中图分类号】R318;TP520.60

【文献标志码】A

【文章编号】1005-202X(2024)08-1046-07

Skin disease diagnosis and treatment model based on text classification algorithm

LING Tian, ZHU Jiazhen, JIAO Yang, LI Lufang

Library of Zhejiang Chinese Medical University, Hangzhou 310053, China

Abstract: In response to the challenges of small scale and huge labor cost in biomedical feature modeling in current skin disease assisted diagnosis, as well as the inability to accurately describe the time series of patient disease features, a fusion text classification algorithm is used to integrate commonly used text classification models (TextLSTM, TextCNN, and RCNN) to obtain a model based on transfer learning and neural networks (TLNN model). By extracting the medical features of image sensors and quantizing them, the pretreatment reduces the number of foci and eliminates the feature information with large deviations, thus improving the accuracy of decision data. TLNN model achieves an accuracy of 72.36% on ISIC2018 and PH2 datasets, which is higher than those of the other 3 text classification models. The diagnostic accuracy of TLNN model is close to doctor's diagnosis (92% vs 94%), but the effective diagnostic efficiency is significantly higher than doctor's diagnosis (1.17 min/case vs 4.57 min/case), and the overall efficiency is improved by 290%. The results demonstrate that the fusion text classification algorithm model can obtain accurate diagnosis in less time than the traditional manual diagnosis. TLNN model can be applied to disease diagnosis, and assist doctors in medical decision-making, thereby providing patients with high-quality and convenient intelligent diagnosis and treatment services.

Keywords: skin disease; auxiliary diagnosis; fusion text classification algorithm; Dempster-Shafer evidence theory; medical characteristics

前言

【收稿日期】2024-03-20

【基金项目】教育部产学研合作协同育人项目(202102242027, 202102242043)

【作者简介】凌天, 硕士, 研究方向: 中医药文化、医学语料库、信息管理与信息系统、数字图书馆, E-mail: 7645864@qq.com

【通信作者】诸佳珍, 博士, 助理研究员, 研究方向: 中药学、医学管理, E-mail: 80789694@qq.com

皮肤病是一种常见病且种类繁多,不仅让患者承受身体病痛,同时也给患者的心理健康与社会生存带来不少影响。鉴于我国当下医疗资源不均衡^[1]以及医院内部各系统数据来源不匹配^[2],不少皮肤病辅助诊断技术应运而生,但也难免存在瓶颈,如早期皮肤科医生通过计算机辅助诊断系统的临床皮肤影像作出经验式的诊断或通过计算机算法对皮肤病进行辅助诊疗,难免出现非精确结论^[3-7]。因此针对皮肤病的智能辅助诊疗技术的理论探讨有着重大的价值。

已有研究者对皮肤病的智能辅助诊疗技术开展积极探索,如在计算机辅助系统的基础上引入迁移学习和深度学习的图像分割研究。Ge等^[8]和王雪^[9]基于深度卷积神经网络对因癌变致皮肤损伤的数据集深度学习显著性特征描述符,分析皮肤病变的特征;丁相文^[10]利用公开皮肤病数据集(ISIC-2017和PH2)深度学习皮肤图像密集特征,改进编码-解码器的解码能力,展现出良好的图像分割效果,提升皮肤黑色素瘤的诊治效率;陈奕希^[11]和万承北^[12]基于Res Net50卷积神经网络进行皮肤病图像分割,提取多层次的语义信息特征层,通过特征融合实现病变图像准确分类。此外,还有基于计算机算法的图像数据分类研究:基于皮肤轮廓的水平集技术提取临床病理和病变形态学特征数据,从而进行准确的图像分类^[13-16]。侯璐璐^[16]基于元学习的多尺度嵌入网络方法准确地实现皮肤病图像边缘分割以及皮肤镜图像病变区域标记数据,提升分类精度,对临床诊断具有指导意义。有研究者构建痤疮分类诊断模型与知识库,利用语义分析实现对寻常型痤疮图像的自主诊断,提高诊断准确率^[17-20]。陈旭琳^[21]创建循环空间注意力模型,利用空间图像特征和循环神经网络的隐藏状态学习得到皮肤病图像的空间注意力分布,对输入性皮肤病图像进行不一致重采样,提取具有判别性的图像特征数据。上述研究从不同角度阐述运用计算机科学技术进行皮肤病辅助诊断,但也存在一些不足:(1)常规机器学习算法对复杂数据的处理能力受限;(2)应用算法模式前必须完成特征建立,耗费人力;(3)基于生物学特征的机器学习判别方法在反映皮肤症状的时间序列特性方面存在困难,如带状疱疹的典型表现是患者首先身体剧痛,在短时间疾病进程,疼痛部位又产生疱疹,医生很难根据这种特殊病程时间顺序做出准确医疗诊断。

融合文本分类算法^[22]是基于图像分割算法发展而来,提取患者的医学特点,将患者描述成其所拥有医疗特点的集合,而后将这些病例转换为向量数据,如{0,1,-1}。随着神经网络模型的快速发展,融合文本分类算法与神经网络相结合具备如下优势:(1)利用模板可以自动获取文字特征并对特征与分类标签之间的关联性进行训练;(2)当模板表示为重复变量时,可提高人工神经元的层数和各层神经元的总量;(3)利用一个神经网络的文本分类模式可以进行上下文数据考察,在疾病诊断过程中还可以分析患者疾病进展的时间序列^[23-25]。

本研究提出基于融合文本分类算法的皮肤疾病辅助诊疗模型(Model Based on Transfer Learning and Neural Networks, TLNN),通过对患者的医学特征数据集进行D-S方法(Dempster-Shafer证据理论)截断处理,降低数据冗余度,将处理后的数据输入皮肤病诊疗

模型中,得到相应诊断结论。实验结果表明,与现有的皮肤病图像分割或分类算法以及医生人工判断相比,融合文本分类算法模型有着更高的诊断精度与诊疗效果。

1 数据特征提取及预处理

常用医学检验的生物医学图像传感器中节点的决策信息常具有较多的冗余特性,因此必须通过预处理计算来减少焦块数量,并以一致性检验的方式将偏差较大的特征信息消除,从而提高决策数据精度。图像识别系统一般是由特征信息提取和控制部分构成,这两部分都不可避免地要面对不确定性误差数据。采用特征模型的目标识别,首先需要图像传感器获取图像信息,再从这张图像中获取各种很可能具有特定冗余度或互补度的特征信息,最终通过数据预处理程序将这一影响消除或降低,从而实现精确识别目标。通常的预处理步骤包括决策矩阵首次预处理和二次预处理。

首次预处理:通过设置特定条件对识别出的决策矩阵的概率值进行处理,通常根据D-S方法截断矩阵来消除偏离特定条件较大的特征概率所带来的决策误差,消除误差后的矩阵行列再进行向量归一化计算生成待二次预处理的新矩阵。通常设定目标数量值条件 n ,传感器概率值条件 m 等。

二次预处理:传感器感知目标的节点信息具有一定的数据冗余性,所以相邻感知节点信息具有相对接近的概率值。若概率值偏离过大应以消除处理,进一步提高决策结果准确性。

1.1 数据融合数值准备

数据融合是将多传感器图像在空间或时间上的冗余或互补性数据按照一定的算法进行运算处理,得到比任何单一数据更精确、更丰富的信息,以提高归类精度,提高数据的应用价值。传感器可随机生成6个不确定性概率值数值来感知同一目标,分别为M1、M2、M3、M4、M5、M6,可能的目标身份有7个,分别为A1、A2、A3、A4、A5、A6、A7,其基本概率指派如表1所示。

表1 基本概率指派
Table 1 Basic probability assignment

传感器	身份						
	A1	A2	A3	A4	A5	A6	A7
M1	0.50	0.20	0.20	0.20	0.03	0.03	0.04
M2	0.40	0.30	0.20	0.10	0.03	0.03	0.03
M3	0.40	0.30	0.30	0.30	0.02	0.04	0.03
M4	0.45	0.25	0.20	0.40	0.01	0.04	0.02
M5	0.20	0.65	0.20	0.20	0.03	0.02	0.03
M6	0.03	0.01	0.04	0.07	0.70	0.03	0.03

1.2 数据融合及结果分析

由基本概率中指派数值转换得到的原始决策矩阵 M 为:

$$M = \begin{pmatrix} 0.50 & 0.20 & 0.20 & 0.20 & 0.03 & 0.03 & 0.04 \\ 0.40 & 0.30 & 0.20 & 0.10 & 0.03 & 0.03 & 0.03 \\ 0.40 & 0.30 & 0.30 & 0.30 & 0.02 & 0.04 & 0.03 \\ 0.45 & 0.25 & 0.20 & 0.40 & 0.01 & 0.04 & 0.02 \\ 0.20 & 0.65 & 0.20 & 0.20 & 0.03 & 0.02 & 0.03 \\ 0.03 & 0.01 & 0.04 & 0.07 & 0.70 & 0.03 & 0.03 \end{pmatrix} \quad (1)$$

原始决策矩阵在首次预处理方式前,可以根据目标数量以及传感器随机生成概率值设置特定条件 $n=4, m \geq 0.05$,再用 D-S 方法进行截断处理,最后可形成决策矩阵 M_1 :

$$M_1 = \begin{pmatrix} 0.50 & 0.20 & 0.20 & 0.20 \\ 0.40 & 0.30 & 0.20 & 0.10 \\ 0.40 & 0.30 & 0.30 & 0.30 \\ 0.45 & 0.25 & 0.20 & 0.40 \\ 0.20 & 0.65 & 0.20 & 0.20 \end{pmatrix} \quad (2)$$

运用 D-S 方法消除部分目标后生成的 M_1 矩阵规模有所减少,这种矩阵向量变化后难免会对决策数据产生影响,因此需进行决策矩阵首次预处理,通过行列进行向量归一化降低影响, M_1 经过归一化可得到决定决策矩阵 M_2 :

$$M_2 = \begin{pmatrix} 0.56 & 0.22 & 0.22 & 0.22 \\ 0.45 & 0.33 & 0.22 & 0.11 \\ 0.45 & 0.33 & 0.33 & 0.33 \\ 0.52 & 0.28 & 0.22 & 0.56 \\ 0.22 & 0.72 & 0.22 & 0.22 \end{pmatrix} \quad (3)$$

通过观察决策矩阵 M_2 中的元素对比每行的矩阵节点,可以明确节点 M_5 的判决结果为 0.22,和另外 4 个节点的判决结果存在较大差异,如果它存在粗大误差,就必须把它剔除。因此,目标仍需进行二次预处理,将归一化后存在较大误差的最大值 A1 的节点 M_5 排除,经处理后生成决策矩阵 M_3 :

$$M_3 = \begin{pmatrix} 0.56 & 0.22 & 0.22 & 0.22 \\ 0.45 & 0.33 & 0.22 & 0.11 \\ 0.45 & 0.33 & 0.33 & 0.33 \\ 0.52 & 0.28 & 0.22 & 0.56 \end{pmatrix} \quad (4)$$

最终通过狄克逊(Dixon)准则检验各阶段决策矩阵。Dixon 准则是一种检验异常值的统计学方法,当目标数量 $n \leq 7$ 时,使用计算公式 $p_{10} = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}}$ 较为准确,其中 $x_{(1)}$ 为最小概率值, $x_{(n)}$ 为最大概率值,最终各矩阵焦元 $M(A)$ 的决策结果如表 2 所示。

表 2 决策结果
Table 2 Decision results

操作	M(A1)	M(A2)	M(A3)	M(A4)	M(A5)	M(A6)	M(A7)
没有预处理	0.726 512	0.236 722	0.007 871	0.183 100	0.000 400	0.000 160	0.000 160
经 D-S 方法截断处理	0.737 742	0.230 051	0.007 871	0.018 200	0.000 030	0	0
首次预处理	0.605 000	0.369 000	0.005 000	0.005 000	0	0	0
二次预处理	0.889 000	0.090 000	0.007 000	0.007 000	0	0	0

通过对决策结果(表 2)进行对比分析发现,采用 D-S 方法截断与两次数据预处理技术相结合能显著降低进行判断的概率值误差,提升判断结论的正确性。与单独采用 D-S 方法比较,通过截断原始矩阵中偏离度较大的 $M(A5)$ 、 $M(A6)$ 、 $M(A7)$,降低参与计算目标参与度,从而降低矩阵规模,也因此减少 D-S 方法与预处理综合计算的复杂性,还能消除所有参与运算的误差节点 m_5 ,从而增加判断结果中对 $M(A1)$ 的信任度,增加 $M(A1)$ 与 $M(A2)$ 的识别性,为对目标身份的判断奠定比较可信的基础。与截断方法基础上进行首次预处理比较,额外二次预处理可以消除 $M5$ 节点,从而排除 $M(A5)$ 参与决策结果计算,再一次降低综合计算的复杂性。

通过对 6 个传感器随机产生概率数值 $M1$ 、 $M2$ 、 $M3$ 、 $M4$ 、 $M5$ 、 $M6$ 进行 D-S 证据理论与两次数据预处理技术融合分析识别目标可能身份 $A1$ 、 $A2$ 、 $A3$ 、 $A4$ 、 $A5$ 、 $A6$ 、 $A7$,其中 $M(A1)$ 最终处理结果为 0.889,远高于其他目标,因此目标最大可能性是 $A1$,而 $M(A5)$ 、 $M(A6)$ 、 $M(A7)$ 的最终值均为 0,因此目标 $A5$ 、 $A6$ 、 $A7$ 可以直接排除。实验结果也表明 D-S 证据理论与数据预处理技术联合进行数据融合可以降低数据高冗余性,并能消除误差等较大数字值的能力。

2 皮肤病辅助诊疗模型构建

为规避其他机器学习方式中无法有效表达疾病发展过程的时间序列、无法建立复杂函数建模系统

和繁琐的特征工程等弊端,本研究以皮肤病数据分析的小样本为例,分析融合文本分类算法的皮肤病辅助治疗方式,提出TLNN模型。

2.1 模型设计与构建

皮肤病辅助诊断模型建立过程如图1所示。首先提取整理皮肤病相关循证医学库的知识点,掌握皮肤病的相关基础知识,然后构建包含多种类皮肤病特征虚拟数据集,即源域数据集;再选取3种对文本浅层特征进行抽取的神经网络文本分类模式(TextLSTM、TextCNN、RCNN)进行虚拟数据集预训练,得到源模型;最终利用迁移机器学习技术,源模型在皮肤病真实数据集中进一步练习,运用直推式迁移方法将循证医学数据库中的知识转移到辅助诊断模型中,同时将3种分类模式平衡融合得到理想的

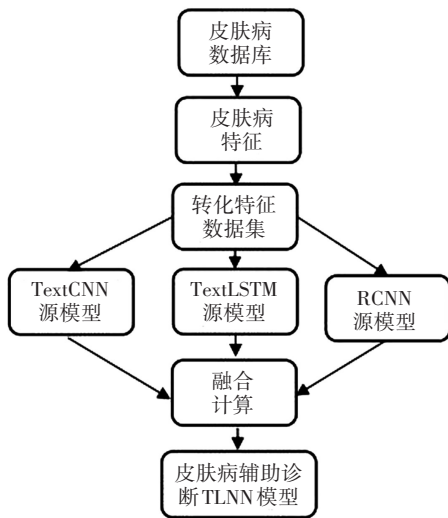


图1 皮肤病辅助诊断模型构建流程

Figure 1 Construction of the model for the auxiliary diagnosis of skin diseases

皮肤病TLNN模型。

基于融合文本分类算法的皮肤病智能辅助诊断模型构建方式将通过融合步骤(1)~(3)进行详细阐述。预训练数据集 D_{pre} 为虚构数据集,资料集中的任一组数字 x 即生成的虚构病灶样本,训练皮肤病学数据集中的 D 为实际值,数据集中的任一组数字 x 为实际病灶数。该方法主要有3个步骤:开始源模型练习、使用真实数据练习、实现与模型相融合。采用迁移和神经网络相结合的方法建设皮肤病诊断模型。输入预训练数据集 $D_{pre} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 其中 $x_i \in X \subseteq R^n$; $y_i \in \{c_1, c_2, \dots, c_m\}$, m 为疾病标签总个数,而训练数据集中的任意 x_i 为虚拟生成的病历样本。真实皮肤病

数据集 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 其中 $x_i \in X \subseteq R^n$; $y_i \in \{c_1, c_2, \dots, c_m\}$, m 为疾病标签总个数,此数据集中的任意 x_i 为真实病灶样本。

(1)原模型训练:使用基于神经网络的文本分类模型在预训练数据集 D_{pre} 上训练皮肤病诊断模型:

$$\text{Model}_{pre}^{(i)}(x), i \in \{\text{TextLSTM}, \text{TextCNN}, \text{RCNN}\} \quad (5)$$

保存模型 $\text{Model}_{pre}^{(i)}(x)$ 的参数:

$$\text{Model}_{pre}^{(i)}(x) = \text{Train}_{pre}(D_{pre}) \quad (6)$$

其中, Train_{pre} 表示预训练过程。

(2)真实数据训练:使用预训练获得的 $\text{Model}_{pre}^{(i)}(x), i \in \{\text{TextLSTM}, \text{TextCNN}, \text{RCNN}\}$ 的模型参数,在皮肤数据集 D 上训练皮肤病分类预测模型 $\text{Model}^{(i)}(x)$:

$$\text{Model}^{(i)}(x) = \text{Train}(\text{Model}_{pre}^{(i)}(x), D) \quad (7)$$

(3)TLNN融合模型:首先使用TextCNN、TextLSTM和RCNN模型执行步骤(1)的原模型训练和步骤(2)的真实数据训练,先后获得3种皮肤病诊断模型 $F_{\text{TextLSTM}}(x)$ 、 $F_{\text{TextCNN}}(x)$ 以及 $F_{\text{RCNN}}(x)$ 。将多个皮肤病的诊断模型组合为一个皮肤病诊断模型, $F_{\text{TextLSTM}}(x)$ 、 $F_{\text{TextCNN}}(x)$ 以及 $F_{\text{RCNN}}(x)$ 成为模型的向量数据。

$$F(x) = \{F_{\text{TextCNN}}(x), F_{\text{TextLSTM}}(x), F_{\text{RCNN}}(x)\} \quad (8)$$

2.2 模型诊断

给予一种未知皮肤病标记集合中的病例样品,检测其存在的皮肤病标签集,与模型中皮肤病数据集进行比对,最终诊断出已知病名。实际实现方式如下:

(1)待检测样本 x 。

(2)根据 $F_{\text{TextLSTM}}(x)$ 、 $F_{\text{TextCNN}}(x)$ 以及 $F_{\text{RCNN}}(x)$,可以获得3个不同的 m 维向量,预测样本 x 可通过TextCNN、TextLSTM和RCNN源诊断模型计算出诊断结果,每个向量均存在 m 层维,每层维表示预测样本 x 属于每一类皮肤病标签的概率值 $\{P_1, P_2, P_3, \dots, P_m\}$ 。

$$\vec{P}_{\text{TextLSTM}} = F_{\text{TextLSTM}}(x) = \{P_1, P_2, P_3, \dots, P_m\} \quad (9)$$

$$\vec{P}_{\text{TextCNN}} = F_{\text{TextCNN}}(x) = \{P_1, P_2, P_3, \dots, P_m\} \quad (10)$$

$$\vec{P}_{\text{RCNN}} = F_{\text{RCNN}}(x) = \{P_1, P_2, P_3, \dots, P_m\} \quad (11)$$

(3)通过将TextCNN、TextLSTM和RCNN源诊断模型进行均等比例加权融合得到待测试样本 x 最终诊断模型 $F_{\text{TLNN}}(x)$ 。由 $F_{\text{TLNN}}(x)$ 输出代表可能标签的概率向量 $\vec{P}_{x \in c_j}$ 。

$$F_{\text{TLNN}}(x) = \frac{1}{3} \vec{P}_{\text{RCNN}} + \frac{1}{3} \vec{P}_{\text{TextCNN}} + \frac{1}{3} \vec{P}_{\text{TextLSTM}} = \vec{P}_{x \in c_j} \quad (12)$$

(4)最终得到包含 m 维的概率向量 $\vec{P}_x = \{P_1, P_2, P_3, \dots, P_m\}$, $P_i, j \in \{c_1, c_2, \dots, c_m\}$ 表示样本 x 属于标签的最终概率。将概率最高的皮肤病标签作为样本 x 的诊断结果。

3 实验与结果分析

为探究构建的融合文本分类算法的智能辅助诊疗模型精确度与实际应用价值,本研究从国际公开ISIC 2018和PH2皮肤镜图像分割数据集的2 000张JPEG格式的病变图像中挑选50例图像作为病灶诊断样本。主要疾病分类为:普通痣(Common Nevus, CN)、非典型痣(Atypical Nevus, AN)、黑色素瘤、非痣样良性黑素上皮瘤。此4种分类在研究公开数据集中黑色素瘤的发病率高且具有一定的代表性,其可以转化为向量数据的医学特征主要有:色素网、斑点/球状物、对称性、条纹、退化区域以及临床显著颜色等。样本数据集表格如表3所示。

表3 皮肤病样本数据集
Table 3 Sample datasets of skin diseases

类标签号	数据集名	样本数量	正样本数量
1	CN	50	19
2	AN	50	13
3	黑色素瘤	50	8
4	非痣样良性黑素上皮瘤	50	10

3.1 融合模型 TLNN 与 RCNN、TextCNN、TextLSTM 对比客观分析

本组实验采用精度评定各分类方法的分类性能,包含3个指标:精确率(Precision)、召回率

(Recall)和准确率(Accuracy):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{13}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{14}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{15}$$

其中,样本中的正负样本可用标签值进行划分。TP代表将正类预测为正类的样本数,FP代表将负类预测为正类的样本数,FN代表将正类预测为负类的样本数,TN代表将负类预测为负类的样本数。

通过计算得出样本数据集精确率和召回率如表4所示。通过对比分析,在精确率层面,TLNN模型在AN与黑色素瘤的精确率最高,分别达到66.67%与72.09%,另外在非痣样良性黑素上皮瘤种类上的精确率接近于RCNN与TextCNN。在召回率层面,TLNN模型在AN和非痣样良性黑素上皮瘤召回率最高,分别为74.16%与76.04%,CN种类的召回率接近RCNN与TextLSTM,黑色素瘤的召回率接近RCNN与TextCNN。在数据集准确率层面,RCNN、TextCNN、TextLSTM的准确率值分别为63.95%、71.12%、42.22%,而TLNN模型的样本数据集准确率为72.36%,均优于以上3种数据文本分类方法。综合以上结论,TLNN模型在平衡了RCNN、TextCNN、TextLSTM在不同类型皮肤病分类效果的基础上,减少数据冗余以及偏差,从而带来更高的诊断精度与诊疗效果。

表4 精确率比较和召回率比较(%)
Table 4 Precision and recall rate comparisons (%)

模型	CN		AN		黑色素瘤		非痣样良性黑素上皮瘤	
	精确率	召回率	精确率	召回率	精确率	召回率	精确率	召回率
RCNN	21.00	20.97	41.46	74.15	50.28	67.42	66.76	74.50
TextCNN	8.31	7.51	20.85	29.70	69.02	80.06	68.60	64.23
TextLSTM	13.01	26.02	22.12	37.94	11.46	19.66	6.55	15.73
TLNN	8.40	19.51	66.67	74.16	72.09	63.90	65.77	76.04

3.2 融合模型与皮肤科医生人工诊断对比主观分析

皮肤病辅助诊疗模型最终将应用到临床研究前需与临床诊断医生作对照研究,从国际公开ISIC 2018和PH2皮肤镜图像分割数据集选择50例样本作为实验组的病例数目。具体实验步骤如下:A组为邀请4名主任级皮肤科医生依据个人丰富的临床经验

进行病变图像审阅后给出诊断报告,并嘱咐相关实验要求。第二组将数据集的病变图像的元数据转换成特征数值后代入智能辅助诊疗模型后自主诊断。最终实验结果见表5。4名医生由于爱主观能动性与时时间精力等等主观因素影响,实验结果呈现一定差异性,取其平均值与B组比较。经对照发现诊疗模型

准确率92%接近于A组94%的平均准确率,而有效诊断效率,1.17 min/例,明显高于医生人工诊断的4.57 min/例,整体效率提升幅度达290%。

表5 对照实验结果
Table 5 Controlled trial results

对照实验		时间/min	有效诊断数/例	准确率%	效率/min·例 ⁻¹
A组	医生 A	207	45	90	4.60
	医生 B	218	48	96	4.54
	医生 C	192	46	92	4.17
	医生 D	243	49	98	4.96
	平均值	215	47	94	4.57
B组	诊疗模型	53.8	46	92	1.17

因此,在医疗资源集中的大型医院,医生可以根据临床经验作出的诊断与智能辅助诊疗模型的自主诊断相互佐证,从而提高诊疗效率。而在医疗资源相对缺乏的小型医院、社区医院,医生可以借鉴智能辅助诊疗模型的诊断结果进行分析验证,从而有效地为患者作出合理的治疗决策。

4 讨论与结论

针对当前医学皮肤病运用计算机辅助诊断技术无法精确决策的难点,本研究运用融合文本分类算法,提取影像传感器医学特征向量化后进行预处理,减少焦块数量以及消除偏差较大的特征信息,从而提供决策数据精度。对照实验结果表明,融合文本分类算法模型相比传统人工诊疗以更短的时间获得精确的诊断。该模型可以应用于疾病诊断,辅助医生医疗决策,降低医疗风险和医生的工作时间成本,更好地为患者提供优质便捷的智能诊疗服务。

【参考文献】

[1] 周爱萍. 城乡医疗资源不均衡配置及改善研究[J]. 合作经济与科技, 2020(22): 172-176.
Zhou AP. Study on unbalanced allocation and improvement of urban and rural medical resources[J]. CO-Operative Economy & Science, 2020(22): 172-176.

[2] 廖杨春, 李杰. 医院信息系统数据共享[C]//中华医学会第二十次全国医学信息学术会议论文集. 重庆: 中华医学会, 2014: 511.
Liao YC, Li J. Hospital information system data sharing [C]// Proceedings of the 20th National Medical Information Academic Conference of the Chinese Medical Association. Chongqing: Chinese Medical Association, 2014: 511.

[3] 李昂, 崔勇. 人工智能在皮肤科中的应用[J]. 中国皮肤性病学杂志, 2022, 36(8): 872-876.
Li A, Cui Y. Application progress of artificial intelligence in dermatology[J]. The Chinese Journal of Dermatovenereology, 2022, 36(8): 872-876.

[4] 孟如松. 皮肤影像学概述与新进展[C]//第九届全国生物医学体视学学术会议、第十二届全军军事病理学学术会议暨第八届全军定

量病理学学术会议论文集. 呼和浩特: 中国体视学学会生物医学分会, 2014: 51-57.

Meng RS. Overview and new progress of skin imaging [C]// Proceedings of the 9th National Biomedical Stereology Academic Conference, the 12th Military Pathology Academic Conference, and the 8th Military Quantitative Pathology Academic Conference. Hohhot: Biomedical Branch of the Chinese Society of Stereology, 2014: 51-57.

[5] 商显震. 皮肤病辅助诊断关键技术研究与应用[D]. 银川: 北方民族大学, 2019.
Shang XZ. Disease research and application of auxiliary diagnostic methods for skin[D]. Yinchuan: North Minzu University, 2019.

[6] Peruch F. (SEMI)-automated analysis of melanocytic lesions[D]. Padova: Università degli Studi di Padova, 2015.

[7] Hameed N, Ruskin A, Abu Hassan K, et al. A comprehensive survey on image-based computer aided diagnosis systems for skin cancer[C]// 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA). Piscataway, NJ, USA: IEEE, 2016: 205-214.

[8] Ge ZY, Demyanov S, Chakravorty R, et al. Skin disease recognition using deep saliency features and multimodal learning of dermoscopy and clinical images[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2017. Cham: Springer International Publishing, 2017: 250-258.

[9] 王雪. 基于神经网络的医学图像分割算法研究[D]. 长春: 吉林大学, 2022.
Wang X. Research on medical image segmentation algorithms based on neural network[D]. Changchun: Jilin University, 2022.

[10] 丁相文. 基于深度学习和特征融合的皮肤病变图像分割研究[D]. 长春: 吉林大学, 2022.
Ding XW. Research on skin lesion image segmentation based on deep learning and feature fusion[D]. Changchun: Jilin University, 2022.

[11] 陈奕希. 基于深度学习的皮肤病图像分割与分类研究[D]. 赣州: 江西理工大学, 2022.
Chen YX. Research on skin disease image segmentation and classification based on deep learning[D]. Ganzhou: Jiangxi University of Science and Technology, 2022.

[12] 万承北. 基于卷积神经网络的皮肤病检测方法研究[D]. 南京: 南京邮电大学, 2022.
Wan CB. Research on skin disease detection method based on convolutional neural network[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2022.

[13] Thamizhvani TR, Lakshmanan S, Sivaramakrishnan R. Computer aided diagnosis of skin tumours from dermal images [C]// Computational Vision and Bio Inspired Computing. Cham: Springer International Publishing, 2018: 349-365.

[14] Zhou HY, Schaefer G, Sadka AH, et al. Anisotropic mean shift based fuzzy C-means segmentation of dermoscopy images[J]. IEEE J Sel

- Top Signal Process, 2009, 3(1): 26-34.
- [15] 刘衍禄. 基于深度特征融合的皮肤病分类算法研究与实现[D]. 南昌: 南昌大学, 2022.
Liu YL. Research and implementation of skin disease classification algorithm based on depth feature fusion[D]. Nanchang: Nanchang University, 2022.
- [16] 侯璐璐. 基于深度学习的皮肤病变图像分类算法研究[D]. 兰州: 西北师范大学, 2021.
Hou LL. Research on classification algorithm of skin lesions image based on deep learning[D]. Lanzhou: Northwest Normal University, 2021.
- [17] 杨聚加, 段然, 吴亚光, 等. 数据分类算法智慧诊疗模型在皮肤痤疮诊断中的研究与应用[J]. 重庆医学, 2022, 51(3): 507-511.
Yang JJ, Duan R, Wu YG, et al. Research and application of a data classification algorithm smart treatment model in the diagnosis of cutaneous acne[J]. Chongqing Medicine, 2022, 51(3): 507-511.
- [18] Przystaliski K, Nowak L, Ogorzałek M, et al. Decision support system for skin cancer diagnosis[C]//The Ninth International Symposium on Operations Research and its Applications (ISORA'10), Chengdu, China, 2010: 406-413.
- [19] 吴晶. 皮肤镜图像中皮肤病变分类算法的研究[D]. 武汉: 武汉科技大学, 2021.
Wu J. Research on classification algorithms of skin lesions in dermoscopy images[D]. Wuhan: Wuhan University of Science and Technology, 2021.
- [20] Li MY, Han CY, Fahim F. Skin cancer diagnosis based on support vector machine and a new optimization algorithm[J]. J Med Imaging Health Inform, 2020, 10(2): 356-363.
- [21] 陈旭琳. 皮肤病图像分类算法的研究[D]. 广州: 广东工业大学, 2021.
Chen XL. Research on classification algorithm of skin disease image [D]. Guangzhou: Guangdong University of Technology, 2021.
- [22] 徐沛娟, 李雄飞, 惠玥, 等. 中文文本分类相关算法的研究与实现[J]. 吉林大学学报(理学版), 2009, 47(4): 790-794.
Xu PJ, Li XF, Hui Y, et al. Research and implementation of related algorithm of Chinese text categorization[J]. Journal of Jilin University (Science Edition), 2009, 47(4): 790-794.
- [23] 袁野, 廖薇. 基于双通道神经网络的疾病文本分类方法[J]. 中国医学物理学杂志, 2021, 38(5): 655-660.
Yuan Y, Liao W. Disease text classification model based on two-channel neural network[J]. Chinese Journal of Medical Physics, 2021, 38(5): 655-660.
- [24] 李启行, 廖薇. 基于注意力机制的生物医学文本分类模型[J]. 中国医学物理学杂志, 2022, 39(4): 518-523.
Li QX, Liao W. Biomedical text classification model based on attention mechanism[J]. Chinese Journal of Medical Physics, 2022, 39(4): 518-523.
- [25] 王壮, 胡卫东, 郁文贤, 等. 基于截断型D-S的快速证据组合方法[J]. 电子与信息学报, 2002, 24(12): 1863-1869.
Wang Z, Hu WD, Yu WX, et al. A fast evidential combination method based on truncated dempster-shafer[J]. Journal of Electronics & Information Technology, 2002, 24(12): 1863-1869.

(编辑: 谭斯允)