

一种基于ResNet-18的医学影像触觉化系统

王 皓

昭通学院 物理与信息工程学院 云南 昭通 657000

摘 要: 本文针对视障医生与患者对医学影像触觉化呈现的迫切需求展开研究。传统方法在将医学图像转换为触觉格式时存在效率低、成本高及精度不足等问题。为此, 本文提出一种基于深度学习的新颖系统, 能够自动化生成X光医学图像的触觉表示。该系统采用ResNet-18深度学习模型, 生成适用于触觉渲染的高分辨率深度图。系统性能通过准确率等客观指标进行评估, 并结合视障医疗从业者的用户反馈进一步验证。实验结果表明, 该系统在提升医学影像可感知性方面具有可行性, 有望增强诊断精度并推动医疗公平性。后续研究将把系统拓展至其他影像模态, 并优化其在实时临床场景中的应用效果。

关键词: 深度学习; 医学影像可触; 性视障; ResNet-18

1 引言

X射线、计算机断层成像(CT)与磁共振成像(MRI)等医学影像手段在疾病诊断中发挥着关键作用, 当前医疗实践广泛依赖这些技术进行快速病情判断与治疗制定。然而, 其视觉依赖性给视障医务人员及患者获取诊断信息带来了显著困难。虽然已有盲文与触觉图示等辅助手段, 但将复杂医学图像转化为可触摸形式仍面临效率低、成本高和精度不足等瓶颈。为此, 本文致力于设计一种能自动构建X射线图像触觉表示的系统。

当前相关研究存在的主要局限在于缺乏高效、自动化且具备扩展性的方法, 难以将复杂的医学图像准确转换为可用的触觉格式, 多数研究集中于从简单图形生成触觉图像, 未能充分应对医学影像所特有的复杂性^[1,2]。本文提出的系统融合图像分割与深度估计等关键技术, 专门针对X射线影像生成高分辨率深度图, 作为后续触觉图像生成的基础, 旨在为视障医护人员与患者提供内容详实、易于感知的医学信息。

2 系统实现

本文设计了一种基于ResNet-18的医学图像触觉转换系统, 其实现过程包括数据准备、预处理、模型训练与性能评估, 主要用于对胸部X射线图像进行肺炎或正常分类。

2.1 数据集准备

在预处理前, 所有X射线图像均被统一调整为 224×224 像素尺寸, 采用双线性插值方法保留图像细节, 同时保证输入格式与预训练ResNet-18模型的兼容性。本研究基于胸部X光影像(肺炎)数据集^[3]开展, 该数据集包含“正常”(NORMAL)与“肺炎”(PNEUMONIA)两类标注图像。为保障模型训练质量与评估结果可靠性, 数据被划分为三个相互独立的子集。

2.1.1 训练集

训练集共包含5198幅影像。通过引入随机水平翻转与旋转等数据增强手段, 有效增强了模型的泛化能力, 并在一定程度上缓解了过拟合问题。此类增强操作属于图像预处理流程的一部分。

2.1.2 验证集

验证集由24幅影像组成, 其中包括12例“正常”和12例“肺炎”样本, 主要用于训练阶段的超参数调整及性能跟踪。该设置有助于选出泛化性能最优的模型, 对控制过拟合具有重要意义。

2.1.3 测试集

模型的最终性能通过一个独立的测试集进行评估, 该测试集包含618张预留图像。借助这一独立测试机制, 能够对模型在未知数据上的表现进行无偏估计。

2.2 深度图生成

基于预处理后的X射线影像生成深度图是本文的研究核心。该过程通过预训练模型MiDaS (Multi Instance Depth Aggregation System) 实现, 该模型可将二维X射线图像转换为包含深度信息的表征图。MiDaS模型能够从单幅RGB影像中估算相对深度值, 最终生成的深度图均被归一化至 $[0,1]$ 区间, 以确保后续处理阶段的数据一致性。

2.3 触觉表征生成

在深度图的基础上, 生成三种触觉表征方式以探索在获取诊断相关信息中的有效性。

2.3.1 灰度表征

深度图被直接转换为灰度图像, 将深度变化编码为像素强度。较轻的像素表示较近的区域, 而较暗的像素表示较远的区域。

2.3.2 轮廓表征

利用scikit-image图像处理库^[4]中的measure.find_contours函数实现肺部区域轮廓提取。该表征方式通过捕捉影像边缘特征与结构不连续区域,为肺部形态学特征分析提供重要量化依据。

2.3.3 高度图表征

将深度图缩放到8位范围(0-255),以创建高度图,并将深度编码为曲面高程。这提供了一个类似于3D的表示,可以进一步分析纹理特征。

这些表征方法将作为后续特征提取阶段的输入。

2.4 特征提取

在完成触觉表征构建后,将提取系列高级特征集用以捕获各类表征中肺部纹理与结构的独特特征,具体特征包括:

深度标准差:该特征作为触觉表征中深度值的量化参数,通过计算深度数据离散程度反映肺表面不规则性。

深度直方图熵:该特征通过量化深度值分布的随机性与无序性,可有效识别肺炎相关的病变区域。

边缘总强度:采用Sobel算子计算触觉表征中的边缘强度总值,该特征通过强化不同解剖结构间的分界区域,在肺炎检测中实现胸膜识别、血管增强、病灶定位。

局部二值模式(LBP)直方图:通过编码中心像素与邻域像素的灰度关系,LBP直方图可有效捕获触觉数据中的局部纹理模式,并能精确描述触觉数据的微观结构特征。

特征提取后,采用sklearn.preprocessing库中的Standard Scaler函数进行标准化处理,确保各特征满足零均值与单位方差要求,该处理可消除特征量纲差异对模型训练的负面影响。

针对肺部结构与纹理的关键特征,本文实现多维度表征。其中统计特征包括深度标准差(1维)、深度直方图熵(1维),纹理特征包括Sobel算子总边缘强度(1维)、LBP直方图特征(9维,基于8邻域均匀模式),多模态融合包括灰度与高度图表征(12维)、轮廓表征。所有特征经标准化后,重构为ResNet-18分类模型的输入层兼容格式。该架构通过特征重缩放(MinMaxScaler)与维度变换(reshape)实现端到端训练。

2.5 ResNet-18模型训练

本文采用预训练的ResNet-18模型执行“正常”与“肺炎”的二分类任务。其架构包含以下核心组件:

输入层:接收 $3 \times 224 \times 224$ 维张量输入(RGB三通道,224×224像素分辨率),通过双线性插值将原始影像统一缩放至目标尺寸。

初始卷积层:卷积核为 7×7 ,滤波器数量为64,步长

为2。

最大池化层:池化窗口为 3×3 ,步长为2。

如图3所示,ResNet-18主体结构包含四个特征提取阶段(Stage 1-4),各阶段通过残差块堆叠实现层级特征学习。单个残差块的前向传播过程可表述为:

$$y = \mathcal{F}(X, \{W_i\}) + X \quad (1)$$

其中 X 为输入特征图, $\mathcal{F}$ 表示残差函数, W_i 为可训练参数。该设计通过恒等映射有效缓解梯度消失问题。各阶段配置参数如表1所示:

表1 各阶段参数配置

阶段	残差块数量	滤波器数量	特征图尺寸	感受野
Stage1	2	64	56×56	35×35
Stage2	2	128	28×28	91×91
Stage3	2	256	14×14	267×267
Stage4	2	512	7×7	483×483

各阶段的滤波器数量逐级递增:Stage 1配置64个滤波器,Stage 2配置128个,Stage 3配置256个,Stage 4配置512个。这种滤波器的渐进式增加使网络能够在每个阶段学习愈加复杂的特征。

ResNet-18在完成四个阶段处理后,通过全局平均池化层聚合空间信息以降低特征图维度。最终将原预训练模型中包含1000个类别的全连接层调整为适应本文的二分类任务,输出对应“正常”和“肺炎”的两个类别。

训练过程采用Adam优化器和交叉熵损失函数。基于验证损失实施的早停法被用于防止过拟合并选择性能最优模型,确保对未见数据具有良好的泛化性能。该流程通过微调预训练的ResNet-18网络,使其专门适配基于提取触觉特征的肺炎检测任务。

2.6 模型评估

最终训练完成的ResNet-18模型在独立测试集上通过综合评估指标进行性能验证,包括准确率、F1值、AUC(曲线下面积)及混淆矩阵。该方法可对模型在未见数据上的表现进行可靠评估,这些指标全面反映了模型在肺炎病例精准分类中的有效性。

3 实验结果与分析

本节基于实验结果与可视化分析,评估提出的基于ResNet-18的X射线肺炎检测系统性能。分析涵盖模型分类性能、触觉表征有效性,以及提升医学影像可及性的潜在价值。

基于触觉表征(灰度图、轮廓图及高度图)特征微调的ResNet-18模型,相比初始结果展现出显著性能提升,如表2所示。在包含150张影像的独立测试集(73例正常,77例肺炎)中,模型达到97.33%的高准确率。

分类（“正常”与“肺炎”）的精确率、召回率及F1值均表现优异，体现其均衡的分类能力。0.9733的F1值与0.9730的ROC曲线下面积（AUC）进一步证实模型对两类样本的区分能力。上述结果表明，该模型有效避免了假阳性与假阴性错误。

表2 模型性能指标

维度	分类正常	分类肺炎	宏平均	加权平均
精准度	0.99	0.96	0.97	0.97
召回率	0.96	0.99	0.97	0.97
F1分数	0.97	0.97	0.97	0.97

本文将提出的系统与基于类似X射线数据训练的其他AI方法进行比较。现有AI方法虽能快速生成高细节输出，但其局限性在于高度依赖大规模数据集进行有效训练且需高性能硬件支持。本方法虽同样需较高算力，但在生成触觉表征方面效率更优，与另一现有AI方法（测试集准确率95.66%、AUC 0.968、F1值0.964）相比展现出竞争优势（本方法测试集准确率97.33%、AUC 0.973、F1值0.972）。

4 结论

本文通过技术创新显著提升了医学影像对视障医疗从业者的可及性，成功验证了将X射线影像自主转换为触觉感知的系统方案。该系统在肺炎检测任务中达到97.33%的准确率，较现有模型性能提升显著。其99%的精确率与96%的召回率（肺炎与正常病例鉴别）凸显了技术普惠价值，使关键医学诊断信息不仅服务于视障医师，更惠及广大患者群体。

该系统具备良好的技术延展性，未来可适配CT扫描、磁共振成像（MRI）等其他医学影像模态，从而扩大其在医疗健康领域的影响范围。随着底层算法的持续优化，系统将实现近实时细节处理能力，为医学诊断提供更稳健的快速分析工具。此项技术突破消除了视觉障碍在获取与提供优质医疗服务中的壁垒，标志着医学可及性领域的重大进步。

参考文献

[1]Barros, C., Lopes, A., & Silva, J. 3D-Printed Tactile Graphics: A Solution for Accessibility in Complex Medical Imaging. May 2023. *Polymers*, vol. 15, no. 9, p2180. DOI: 10.3390/polym15092180

[2]Theurel, A., Witt, A., Claudet, P., Hatwell, Y., & Gentaz, E. Tactile Picture Recognition by Early Blind Children: The Effect of Illustration Technique. June 2013. *Journal of Experimental Psychology: Applied*, vol. 19, no. 2, pp. 233–240. DOI: 10.1037/a0034255

[3]BST 261 2024 - Pneumonia Chest X-ray Classification. (n.d.). Kaggle. Retrieved [Date you accessed Oct 05, 2024], from Kaggle.

[4]Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu and the scikit-image contributors. *scikit-image: Image processing in Python*. *PeerJ* 2:e453 (2014)