

基于卷积神经网络的CT弦图学习与身体部位识别*

陈诗琳, 李淑龙, 马建华

南方医科大学生物医学工程学院/广东省医学图像处理重点实验室, 广东 广州 510515

摘要: 基于医学影像的身体部位识别旨在确定特定医学影像所属的身体部位, 是许多医学影像分析任务必不可少的预处理步骤。目前, 计算机断层扫描(CT, Computed Tomography)技术是临床中最常用的医学影像技术之一, 而基于CT图像的医学影像分析算法(例如, 病灶检测、器官分割等)同样需要首先确定CT图像所属身体部位以获取先验知识, 从而保证算法的速度及鲁棒性。然而CT图像是由被称为CT弦图(CT Sinogram)的CT原始数据重建得到, 而图像重建过程有可能导致信息丢失。因此, 相较于CT图像, CT弦图中应该包含更多适用于身体部位识别的有效信息。然而, 目前基于CT弦图的身体部位识别研究仍比较少。鉴于此, 本研究使用基于卷积神经网络(CNN, Convolutional Neural Network)的深度学习对CT弦图进行特征学习, 并验证其在身体部位识别任务中的可用性。实验采用了1个公开数据集(DeepLesion)和来自著名医学机构的3个临床数据集来验证本研究提出方法的性能。具体而言, 本研究通过Radon变换理论对CT图像进行数据仿真得到CT弦图, 并以CT弦图作为输入, 构造基于CNN的分类器(Sino-Net), 从而对5个最常见的身体部位(头部、颈部、胸部、上腹部以及骨盆)进行识别。实验结果表明, 使用CT弦图进行身体部位识别可以达到与使用CT图像进行身体部位识别相似的性能, 甚至优于基于CT图像识别的结果。

关键词: CT弦图; 神经网络; 身体部位; 多分类

中图分类号: R742; R445.2 **文献标志码:** A **文章编号:** 0529-6579(2021)04-0154-10

Bodypart recognition with CT sinogram based on convolutional neural network

CHEN Shilin, LI Shulong, MA Jianhua

School of Biomedical Engineering, Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou 510515, China

Abstract: Medical-image-based human bodypart recognition, which aims to accurately locate the bodypart of a specific medical image, is an essential preprocessing step for many medical image analysis tasks. Currently, computed tomography (CT) is one of the most available medical imaging techniques in clinic. Many CT-based medical image analysis algorithms (such as lesion detection, organ segmentation, etc.) need to first identify the bodypart information contained in the CT image to obtain prior knowledge, so as to ensure the speed and robustness of the algorithms. However, CT images are reconstructed from CT raw data, which is also known as CT sinogram. And the image reconstruction process may cause information loss. Therefore, compared with CT images, CT sinogram may contain more effective information suitable

* 收稿日期: 2020-11-04 录用日期: 2021-01-19 网络首发日期: 2021-06-02

基金项目: 国家自然科学基金(11771456, U1708261)

作者简介: 陈诗琳(1996年生), 女; 研究方向: 医学图像分析、深度学习; E-mail: shelena.chen@foxmail.com

通信作者: 李淑龙(1981年生), 女; 研究方向: 医学图像分析、影像组学; E-mail: shulong@smu.edu.cn

马建华(1975年生), 男; 研究方向: 医学图像重建、机器学习; E-mail: jhma@smu.edu.cn

(李淑龙、马建华为共同通信作者)

for bodypart recognition tasks. However, there are still relatively few researches on bodypart recognition based on CT sinogram. Therefore, the deep learning based convolutional neural network (CNN) technique is used to train on CT sinogram and its usability in bodypart recognition tasks is verified. A public dataset (i. e., DeepLesion) and three clinical datasets from well-known medical institutions are adopted to verify the performance of our proposed method. Specifically, the Radon transform is used to perform data simulation on CT images to obtain CT sinogram, which is served as input to train a CNN-based classifier (Sino-Net) to recognize the five most common bodyparts (i. e., head, neck, chest, upper abdomen and pelvis). The experimental results show that the use of CT sinogram for bodypart recognition can achieve similar performance to the use of CT images, and sometimes even better than the results based on CT images.

Key words: CT sinogram; convolutional neural network; bodypart recognition; multi-class

基于医学影像的身体部位识别 (BPR, Bodypart Recognition) 旨在准确确定医学影像所属身体部位 (例如, 头部、胸部或腹部等), 是许多医学影像分析算法的预处理步骤^[1-2]。近几十年来, 学者们研发了大量医学影像分析算法以辅助医生进行临床诊断和治疗决策^[3-4], 例如, 病灶检测算法^[5]、器官分割算法^[6]等。不同的医学影像分析算法通常需要结合不同解剖结构的先验知识 (例如, 器官形状) 进行设计^[7-8]。因此, 在执行特定的医学影像分析算法之前, 首先需要识别医学影像中包含的身体部位信息以获取先验知识。此外, 给定具体的身体部位信息可以减小分类、分割等医学影像分析算法的搜索范围, 从而提高算法的速度以及鲁棒性^[9]。再者, 放射科医生希望在阅片前能针对特定部位减少冗余信息以加快阅片速度。因此, 能自动筛选感兴趣部位的身体部位识别算法尤为重要。然而, 目前身体部位识别仍是一项艰巨且值得研究的任务^[10-11]。

医学影像种类良多, 其中包括计算机断层扫描 (CT, Computed Tomography)^[12]、磁共振成像 (MRI, Magnetic Resonance Imaging)^[13] 以及正电子发射断层扫描计算机断层扫描 (PET-CT, Positron Emission Tomography-Computed Tomography) 等。由于自身优势, CT 相对其他医学影像技术有更为广泛的临床应用。一般而言, CT 图像的 DICOM 头文件中含有身体部位信息^[14]。对于非全身 CT 扫描, 学者们可以通过对 DICOM 头文件进行文本检索来实现身体部位识别 (例如, 胸部 CT 和腹部 CT 等)。然而, DICOM 头文件中的身体部位信息存在大约 15% 的错误, 从而限制了使用文本检索身体部位的准确性^[15]。此外, DICOM 信息的多语言特性也使得基于文本检索身体部位的研究难以进行^[16]。对于全身 CT 扫描 (例如, PET-CT), 由于其 DICOM 头文件缺少身体部位的信息, 学者们无法基于文本检索进行身体部位识别。为了避免基于文本检索方法的问题, 越来越多的学者转而研究基于 CT 图像进行身体部位识别的算法。与基于文本检索的身体部位识别算法相比, 基于 CT 图像的身体部位识别算法可以有效利用 CT 图像内部的解剖信息, 进而获取更高的身体部位识别精度^[17]。

基于 CT 图像的身体部位识别研究本质上是一个多分类问题。在过去十几年中, 学者们已经提出了许多基于 CT 图像的身体部位识别方法。例如, 田野等^[18] 基于 CT 图像使用 AdaBoost 方法进行身体部位分类识别。Park 等^[19] 提出一种使用小波变换域中的能量信息来确定身体部位的算法。Hong 等^[20] 建立了一个全局参考系去识别身体各部位。该方法从识别头部开始, 确定头部的边界框后, 进而使用不同的算法逐一定位其他身体部位, 包括颈部、胸部、腹部和骨盆。Criminisi 等^[21] 利用回归森林进行解剖部位检测和定位。这些传统分类识别算法大多是先提取预定义的特征, 然后使用特定的分类器进行分类。有时这些传统算法还与特征选择方法结合使用^[22-23]。然而, 这些方法中采用的预定义特征 (也称为人工特征, Handcrafted feature) 通常难以完全反映 CT 图像内部解剖信息^[22], 这极大地限制了身体部位识别准确率的进一步提高。

近年来, 具有强大端到端学习能力的深度学习 (DL, Deep Learning) 技术, 特别是卷积神经网络技术 (CNN, Convolutional Neural Network) 被广泛用于 CT 图像分析任务以及身体部位识别研究中, 取得了一定的成功^[24]。例如, Roth 等^[25] 提出一种基于 CNN 针对医学图像人体解剖识别的特定分类方法。他们

将 2D 轴向 CT 图像作为 CNN 分类器的输入来识别身体的 5 个部分 (颈部、肺部、肝脏、骨盆以及腿部), 其准确率高达 94.1%。Yan 等^[26] 提出了一种用于图像分类的多阶段深度学习框架, 并将其应用于身体部位识别任务。Zhang 等^[27] 利用 3D CT 图像的空间信息作为监督源进行身体部位识别研究。这些基于 CNN 的身体部位识别方法通常是在 CT 图像域中构建, 而不考虑 CT 图像的生成机制。实际上, CT 图像是通过特定的重建算法以及一系列重建步骤从 CT 弦图 (CT Sinogram) 重建生成, 其重建过程势必会导致一些原始数据信息的丢失。虽然, 专家难以从 CT 弦图数据中定义有效的人工特征进行图像分析和疾病诊断, 但基于 CT 弦图的医学影像分析任务在科学和实践上仍具有可行性, 并且具有自主学习能力的深度学习技术使这种可行性得以实现。具体而言, 使用端到端训练的 CNN 框架可以分析和解释 CT 弦图, 从而进行身体部位识别。最近, 越来越多的研究人员试图跳过重建步骤利用深度学习直接对 CT 弦图进行分析, 以获得更好的医学任务性能^[28-29]。Lee 等^[30] 基于深度学习对 CT 弦图数据进行身体部位识别并验证了该方法的有效性。然而, 该研究的实验数据量较小且只局限于同一种 CT 扫描设备所收集的增强 CT 数据。实际上影响身体部位识别性能的因素有很多, 例如增强与非增强 CT 扫描、不同 CT 扫描设备、正常与病变 CT 图像等。因此, 针对更广泛数据类型的身体部位识别方法仍有待进一步研究。

考虑到上述影响因素, 本研究提出利用深度学习对更广泛的 CT 数据类型进行 CT 弦图学习, 并用于身体部位识别。具体而言, 本文方法以 CT 弦图作为 CNN 分类器的输入, 从而构造基于 CNN 的五分类器 (称为 Sino-Net), 对 5 个身体部位 (头部、颈部、胸部、上腹部以及骨盆) 进行识别。为了评估本文提出方法的有效性, 3 种常见的 CNN 结构 (残差网络 (ResNet)^[31]、密集连接网络 (DenseNet)^[32] 以及 Inception 网络^[33]) 将被改进从而分别构建 3 种 CNN 模型, 并使用 1 个公开数据集 (DeepLesion)^[34] 和 3 个来自不同医疗机构的临床数据集来进行实验验证。这些数据集包含了正常和带有病变的 CT 数据、增强和非增强的 CT 数据以及来自于不同扫描设备的 CT 数据, 以尽可能多地考虑更广泛的影响因素。实验结果表明, 基于 CNN 的 CT 弦图学习可以达到与基于 CT 图像进行身体部位识别相似的性能, 甚至优于基于 CT 图像识别的结果。

1 试验材料和仿真

1.1 公开数据集 (DeepLesion)

本文采用公开数据集 (DeepLesion) 来训练和五折交叉验证本文提出的方法, 即使用 CNN 分类器对 CT 弦图进行学习, 从而识别 5 个身体部位。DeepLesion 数据集由美国国立卫生研究院临床中心 (NIHCC) 创建, 是目前世界上最大的 CT 影像数据集。该数据集由 4 427 名患者的 CT 图像组成, 涵盖了大多数人体解剖结构^[34]。该数据集数据类型广泛, 其中包含了增强和非增强 CT 图像以及正常和带有病灶的 CT 图像。根据医学解剖对身体部位的划分, 我们从该数据集中选择了 34 416 张身体躯干部位图像, 其图像大小均为 512×512。这些图像均由 2 名放射科医生手动标记分类, 划分为 5 个最常见的身体部位: 头部、颈部、胸部、上腹部和骨盆, 数据分布如表 1 所示。

表 1 DeepLesion 数据集中 5 个身体部位的数量分布
Table 1 Distribution of five bodyparts in the DeepLesion dataset

身体部位	头部	颈部	胸部	上腹部	骨盆	总计
数量	1 675	4 531	15 177	9 626	3 777	34 416

1.2 临床数据集

本文使用的 3 个临床数据集仅用于独立测试以评估所提出方法的性能。(a) 胸部数据集: 该数据集在美国德克萨斯大学西南医学中心 (UTSW) 收集, 由 100 例早期 (IA 和 IB) 非小细胞肺癌 (NSCLC) 患者组成。共包括 579 张胸部 CT 图像, 其图像大小为 512×512, 使用通用电气 (GE) CT 扫描设备扫描。(b) 上腹部数据集: 该数据集由 2010 年 1 月至 2019 年 5 月在南方医科大学附属珠江医院肝胆外科二科就诊的 135 例胰腺疾病患者组成。共包括 1 537 张上腹部 CT 图像, 其图像大小为 512×512, 使用飞利浦 (Philips) CT 扫描设备扫描。(c) 骨盆数据集: 该数据集在美国德克萨斯西南医学中心收集, 由 15 例 IB-IVA 子宫

颈癌患者组成。共包括836张骨盆CT图像, 其图像大小为512×512, 使用通用电气(GE)CT扫描设备扫描。

1.3 数据仿真

由于真实CT弦图为各厂商的商业机密, 难以获取。因此, 本研究通过数据仿真的方式对CT图像进行仿真从而获得弦图数据。此仿真方法已广泛应用于需要用到CT弦图数据的研究^[35]。本研究采用西门子公司Somatom Definition AS+CT扫描仪的成像几何来模拟弦图, 其成像几何参数具体为: X射线射源到旋转中心的距离595 mm; X射线源到探测器的距离1 085.6 mm; 重建尺寸512×512; 体素尺寸0.664 0 mm; 探测器数量736; 2个相邻探测器之间的间距1.094 7 mm。本研究对上述1个公开数据集和3个临床数据集进行前向拉东变换(Radon transform), 在具有泊松(Poisson)噪声和零均值高斯分布随机值的噪声水平下生成CT弦图^[36], 可由如下公式所示

$$b_i = \text{Poisson}\{I_0 e^{-y_i}\} + \text{Normal}(0, \sigma_e^2), \quad (1)$$

其中 b_i 是第 i 个探测器的仿真数据(弦图), y_i 是对应正常剂量条件下的线积分值, σ_e^2 是背景电子噪声方差, σ_e^2 的值被假设用于商用CT扫描设备是稳定的, I_0 为X射线入射光子数量, 本文设置 $I_0 = 3 \times 10^6$ 为参考值, 即正常剂量条件下仿真的数据。

2 方法

2.1 方法概述

本文提出算法的总流程图如图1所示。首先, 对DeepLesion数据集及3个临床数据集根据1.3节所提及的数据仿真方法进行处理, 以获得CT弦图数据。其次, 将DeepLesion数据集仿真得到的CT弦图数据集采用五折交叉验证法训练CNN分类器, 从而识别5个身体部位, 该网络结构被称为Sino-Net。具体而言, 以CT弦图作为Sino-Net的输入, Sino-Net会对CT弦图进行自主特征学习, 在Sino-Net的最后一层将输出一个五维向量 (y_1, y_2, \dots, y_5) , 然后利用Softmax函数可计算出5个身体部位的预测概率 $p_i (i = 1, 2, \dots, 5)$, 计算公式如下

$$p_i = \frac{\exp(y_i)}{\sum_{j=1}^5 \exp(y_j)}, \quad i = 1, 2, \dots, 5. \quad (2)$$

最后, 使用最大概率原则确定Sino-Net预测的类别 L , 即

$$L = \underset{1 \leq i \leq 5}{\operatorname{argmax}} (p_i). \quad (3)$$

当 L 为1, 2, 3, 4或5时, 则Sino-Net预测的身体部位分别为头部、颈部、胸部、上腹部或骨盆。

本实验使用RMSProp算法 (http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf) 对Sino-Net进行优化训练, 损失函数为平均交叉熵(Mean Cross Entropy), 其定义为

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^5 l_i \log p_i, \quad i = 1, 2, \dots, 5, \quad (4)$$

其中 N 是训练样本的数量, l_i 是相应样本的标签, 满足 $l_i \in \{0, 1\}$ 且 $\sum_{i=1}^5 l_i = 1$ 。

实验使用一个具有24 GB内存容量的NVIDIA Tesla P40图形处理器(GPU)的PyTorch工具包进行实

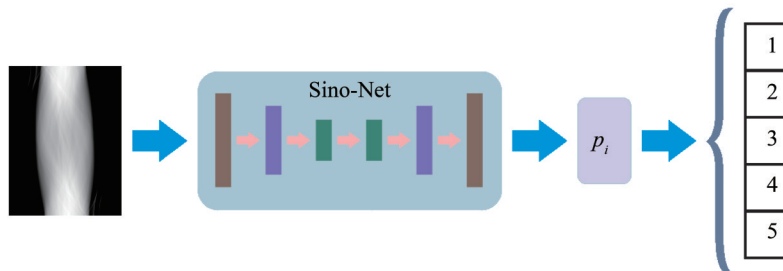


图1 提出方法的总流程图

Fig. 1 Illustration of the proposed method

验。参数动量 (Momentum) 可加速学习过程, 根据实践经验, 我们将其设置为 0.9。批大小 (Batch Size) 为 150, 对所有网络进行了 300 个周期的训练, 学习率为 10^{-5} 。

2.2 CNN 模型结构

为了验证 Sino-Net 对广泛数据类型的 CT 弦图的学习能力和对身体部位的识别性能, 本文对 3 种最常用的 CNN 结构进行修改, 从而执行本文的实验。

2.2.1 基于 ResNet 修改的 CNN 结构 残差网络 (ResNet) 于 2015 年被首次提出, 它在 ImageNet 大规模视觉识别挑战赛 (ILSVRC) 中获得了图像分类的优胜。残差网络易于优化, 并且其内部跳跃连接 (Shortcut) 结构可以缓解由于神经网络深度增加而导致的梯度消失问题。因此, 本文首先基于残差网络的跳跃连接构建了 CNN 模型, 此模型被称为 Res-BPR, 如图 2 所示。它由 8 个卷积 (Convolution) 层, 4 个最大池 (Max-pooling) 层和 1 个全连接 (FC, Fully Connected) 层组成。其中, 卷积核大小为 3×3 , 共有 64 个卷积核, 激活函数为线性整流函数 (ReLU, Rectified Linear Unit), 最大池化大小为 2×2 。除前两个卷积层外, 剩余卷积层之后都添加了批归一化 (BN, Batch Normalization) 操作。

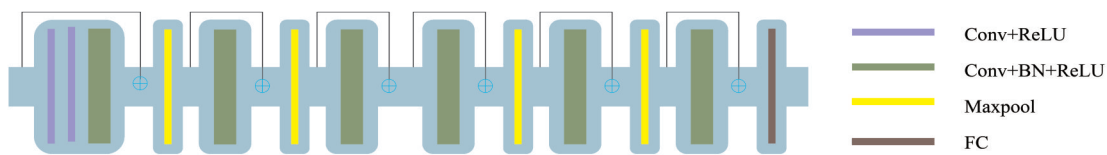


图 2 Res-BPR, 它由 8 个卷积层、4 个最大池化层和 1 个全连接层组成

Fig. 2 Res-BPR, which consists of eight convolution layers, four max-pooling layers and one FC layer

2.2.2 基于 DenseNet 修改的 CNN 结构 密集连接卷积网络 (DenseNet) 于 2017 年被首次提出。自此之后, DenseNet 被广泛用于各种图像分析任务中并取得了上佳的效果。DenseNet 具有较强的泛化能力, 它的成功得益于它内部的密集连接块 (DB, Dense Block) 结构, 该结构能减轻网络训练过程中出现的过拟合问题。鉴于密集连接块的优点, 本文构建了 CNN 网络结构 Dense-BPR, 如图 3 所示。它由 1 个卷积层, 4 个密集连接块, 3 个过渡层 (TL, Transition Layer), 1 个平均池 (Avg-pooling) 层和 1 个全连接层组成。其中卷积核大小为 3×3 , 共有 64 个卷积核。每个密集连接块由不同数量的密集层 (DL, Dense Layer) 组成, 4 个密集连接块的密集层数量依次为 3、6、12 以及 8。每个密集层由两个卷积层组成, 两个卷积层的卷积核大小分别为 1×1 和 3×3 , 卷积核数量分别为 128 和 32。过渡层位于两个相邻的密集连接块之间, 它的作用是改变通道大小, 使得前一个密集连接块输出的特征图大小能与后一个密集连接块的输入相匹配。过渡层由 1 个卷积层和 1 个平均池层组成。其中, 卷积核大小为 1×1 , 卷积数量为 128, 平均池化大小为 2×2 。

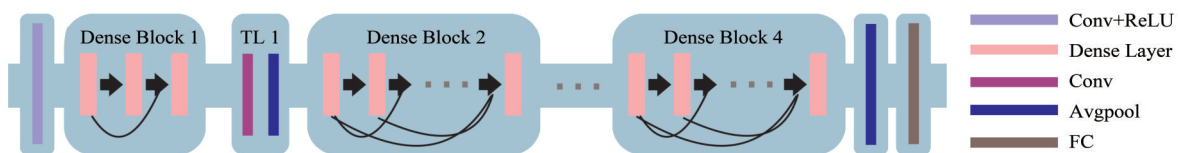


图 3 Dense-BPR, 它由 1 个卷积层、4 个密集连接卷积块、3 个过渡层、1 个平均池化层和 1 个全连接层组成。

2 个相邻的密集连接块由过渡层连接

Fig. 3 Dense-BPR, which consists of one convolution layer, four Dense Blocks, three transition layers (TLs), one average-pooling layer, and one FC layer. Two adjoining Dense Blocks are connected by a TL

2.2.3 基于 Inception 网络修改的 CNN 结构 自 GoogLeNet 于 2014 年在 ILSVRC 中获得第一名以来, Inception 模块引起了学者们的广泛关注。Inception 模块使用多个小卷积核代替大卷积核, 从而提高了参数的利用率并加快了网络的计算速度。如图 4 所示, Inception 模块也被用来构建用于身体部位识别的网络结构, 称为 Incept-BPR。该网络结构由 3 个相同的 Inception 模块、1 个卷积层和 1 个全连接层组成。每个 Inception 模块由 2 个不同的卷积层、1 个最大池化层以及“四条分支”组成。其中, 2 个卷积层的卷积核大小为

3×3, 卷积核数量分别为16和32。最大池层紧接其后, 最大池化大小为2×2。4条分支中的3条分支为卷积层, 卷积核大小分别为3×3、5×5以及1×1。剩余1条分支为平均池层, 池化大小为2×2。由于1×1卷积核运算可以限制通道数并降低计算成本, 因此在卷积核大小为5×5、3×3的卷积层前和平均池层后额外添加了卷积核大小为1×1的卷积层。

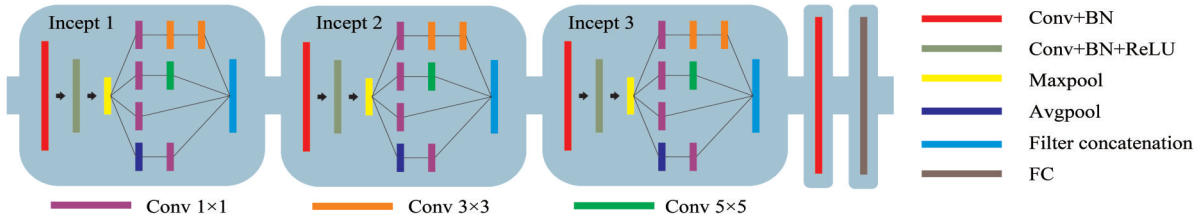


图4 Incept-BPR, 它由3个Inception模块、1个卷积层和1个全连接层组成

Fig. 4 Incept-BPR, which consists of three Inception modules, one convolution layer and one FC layer

2.3 实验设置

整体实验设置如图5所示, 本文使用DeepLesion数据集训练Sino-Net, 并采用五折交叉验证法以验证该算法的有效性, 3个临床数据集将进一步用来独立测试该算法的性能。此外, 与CT弦图相对应的CT图像作为以上3个CNN网络的输入, 训练了相应的CNN网络(称为Img-Net), 将用来和Sino-Net进行比较, 从而进一步验证Sino-Net对身体部位识别的性能。Sino-Net的结果还将与7种传统分类方法的结果进行比较。7种传统分类方法分别为: 逻辑回归(LR, Logistic Regression)^[37]; 线性判别分析(LDA, Linear Discriminant Analysis)^[38]; K近邻(KNN, K-Nearest Neighbor)^[39]; 分类和回归树(CART, Classification And Regression Tree)^[40]; 随机森林(RF, Random Forest)^[41]; 朴素贝叶斯(NB, Naive Bayesian)^[42]以及支持向量机(SVM, Support Vector Machine)^[43]。为了保证实验的公平性, 采用同样数据划分的五折交叉验证法训练7种传统分类方法。具体地, 首先对每张CT图像进行人工特征提取, 共包括532个手工特征, 其中, 7个为形状特征, 13个为纹理特征, 512个为直方图特征。然后, 将这些特征作为7种传统分类器的输入, 以获得最终的结果。

实验的评估准则包括准确率(Acc, Accuracy), 宏F1(macro-F1), 接收者操作特征曲线(Receiver Operating Characteristic Curve, ROC曲线)^[44], ROC曲线下面积(Area Under Curve, AUC)^[45]。用于身体部位识别的Acc计算公式

$$\text{Acc} = \frac{\sum_{i=1}^5 Q_i}{N}, \quad (5)$$

其中 N 代表测试样本的数量, Q_i 为测试样本中第 i 类预测正确的样本个数。宏F1公式如下

$$\text{macro-F1} = \frac{2 \times P_{\text{macro}} \times R_{\text{macro}}}{P_{\text{macro}} + R_{\text{macro}}}, \quad (6)$$

其中 R_{macro} 代表宏查准率: $R_{\text{macro}} = \frac{1}{5} \sum_{i=1}^5 \frac{TP_i}{TP_i + FP_i}$, P_{macro} 代表宏查全率: $P_{\text{macro}} = \frac{1}{5} \sum_{i=1}^5 \frac{TP_i}{TP_i + FN_i}$, TP_i 为真阳性率, FP_i 为假阳性率, FN_i 为假阴性率。 i 代表类别个数, 在本实验中, 共有5个类别, 分别为头部、颈部、胸部、上腹部以及骨盆。

3 结果

表2总结了针对DeepLesion数据集使用3个CNN分类器进行身体部位识别的准确率和宏F1结果。由表2可以观察到基于3种常用CNN改进的模型(Res-BPR、Dense-BPR、Incept-BPR), Sino-Net的性能都比Img-Net的性能更好。Sino-Net的准确率和宏F1在Dense-BPR模型下取得最佳结果, 分别为99.77%和99.76%。实验结果表明, 基于CT弦图进行身体部位识别具有可行性和有效性, 并且与基于CT图像的身体部位识别算法相比, 结果均有所提高。

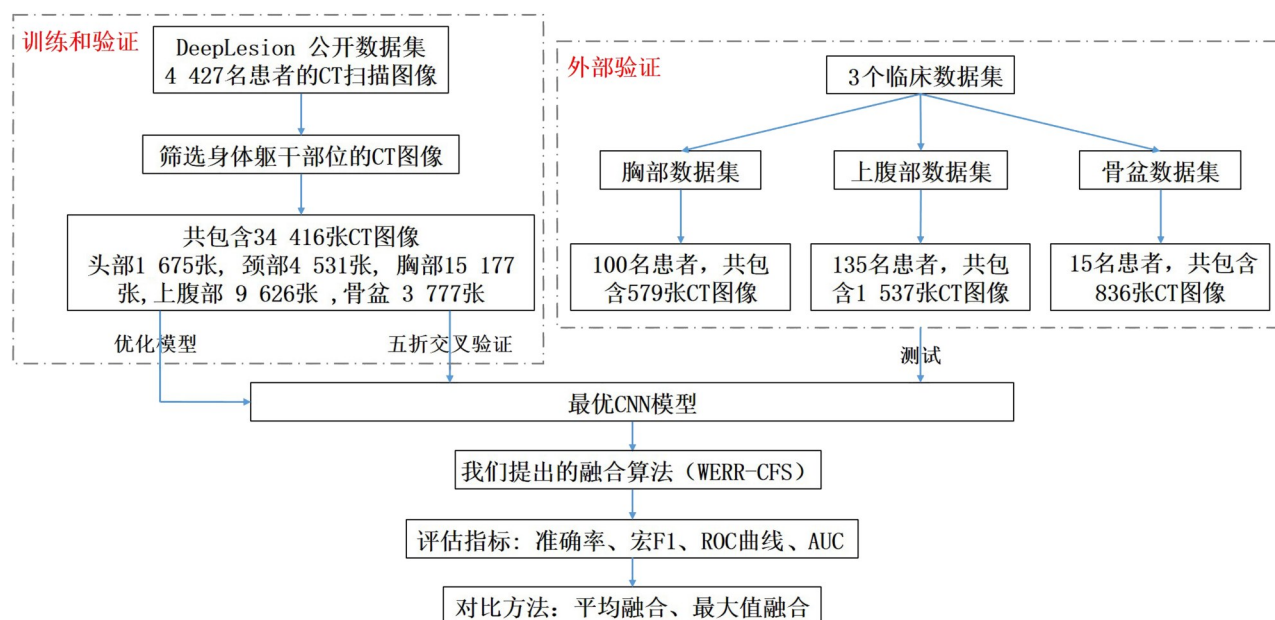


图5 实验设置总流程图

Fig. 5 Illustration of experiment setup

表2 3个CNN网络识别身体部位的准确率和宏F1¹⁾

Table 2 The accuracy and macro-F1 of three CNN networks to recognize bodyparts

%

方法	准确率			宏F1		
	Res-BPR	Dense-BPR	Incept-BPR	Res-BPR	Dense-BPR	Incept-BPR
Img-Net	98.52±0.01	98.34±0.02	98.22±0.04	98.52±0.02	98.35±0.01	98.22±0.04
Sino-Net	99.71±0.01	99.77±0.03	99.06±0.02	99.71±0.01	99.76±0.03	99.07±0.03

1) Img-Net是以CT图像作为输入的CNN分类器, Sino-Net是以CT弦图作为输入的CNN分类器。

表3给出了针对DeepLesion数据集, 本文提出的方法(Sino-Net-Dense-BPR)和7种传统分类方法的准确率比较结果。此实验用在3个Sino-Net模型中表现最好的Dense-BPR模型来与7种传统分类方法进行比较。由表3可以观察到, 7种传统方法中随机森林分类器获得最高的准确率(85.52%), 远低于Sino-Net-Dense-BPR(99.77%)。图6展示了本文提出的方法和7种传统分类方法的ROC曲线, 可以看出, Sino-Net-Dense-BPR取得了最优的结果。综上结果表明, 在身体部位识别算法研究中, 本文提出方法在各方面都优于7种传统方法。

表3 7种传统分类方法和Sino-Net-Dense-BPR的准确率比较

Table 3 The comparison of accuracy between seven traditional classification methods and the Sino-Net-Dense-BPR

%

方法	LR	LDA	KNN	CART
准确率	74.48±0.40	77.57±0.14	82.50±0.35	77.21±0.32
方法	RF	NB	SVM	Sino-Net-Dense-BPR
准确率	85.52±0.42	53.68±2.07	58.45±0.94	99.77±0.03

为了验证本文提出的方法的泛化能力, 来自3个不同机构的临床数据集将用于外部验证, 以独立测试本文提出方法的性能, 结果展示在表4。由表4可以观察到, Sino-Net对3个临床数据集的预测准确率结果都高于Img-Net的预测准确率结果, 并且Sino-Net在3个数据集中都取得了上佳的结果。这些结果表明, 对于来自不同CT扫描设备的数据本文提出的方法具有良好的泛化能力。

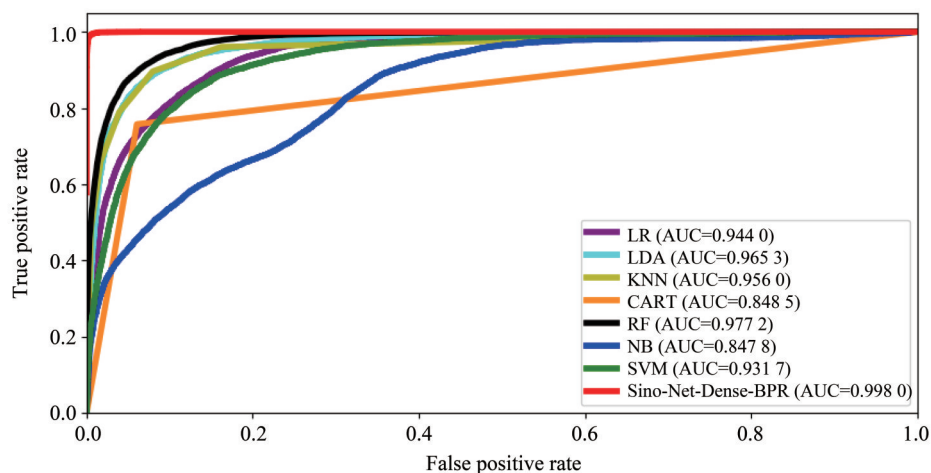


图6 7种传统分类方法和Sino-Net-Dense-BPR的ROC曲线比较

Fig. 6 The comparison of ROC curve between seven traditional classification methods and the Sino-Net-Dense-BPR

表4 3个临床数据集的准确率

Table 4 The accuracy of three clinical datasets

%

方法	胸部数据集	上腹部数据集	骨盆数据集
Img-Net	99.50±0.01	96.36±0.02	95.99±0.03
Sino-Net	100±0.00	97.49±0.01	96.81±0.02

4 讨论和结论

本文使用基于CNN的深度学习对广泛数据类型的CT弦图数据进行学习并应用于身体部位自动识别,通过改进3种常用CNN结构(ResNet、DenseNet和Inception网络)分别构建3个CNN五分类器,并采用公开数据集DeepLesion和3个临床数据分别进行模型训练、五折交叉验证和外部验证。实验结果表明,基于CT弦图的CNN分类器(Sino-Net)在身体部位识别任务中能达到和基于CT图像方法进行识别类似的效果,甚至优于基于CT图像识别的结果。这表明基于CT弦图的医学图像分析任务值得进一步探究,具有潜力。虽然本实验取得了不错的结果,但是本文只对临床CT扫描中5个最常见的身体部位进行分类识别,可能不足以用于临床实践。在将来的研究中,我们将进一步细分身体部位探究本文方法的性能。此外,由于身体部位过渡区域界限的不明确性^[46],仍然存在着一些错分的情况,将来有必要进一步研究解决这个问题。

参考文献:

- [1] 王芹,王然冉,姜述凤,等. 关于人体部位识别的一种模糊算法[J]. 科学技术与工程, 2004, 4(8):687-690.
WANG Q, WANG R R, JIANG S F, et al. A fuzzy algorithm of recognizing human body parts [J]. Science Technology and Engineering, 2004, 4(8):687-690.
- [2] COOTES T F, HILL A, TAYLOR C J, et al. Use of active shape models for locating structures in medical images [J]. Image & Vision Computing, 1994, 12(6): 355-365.
- [3] LI W, YANG Y, ZHANG K, et al. Dense anatomical annotation of slit-lamp images improves the performance of deep learning for the diagnosis of ophthalmic disorders [J]. Nature Biomedical Engineering, 2020, 4(8): 1-11.
- [4] ZHANG K, LIU X, SHEN J, et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography [J]. Cell, 2020, 181(6): 1423-1433.
- [5] ZLOCHA M, DOU Q, GLOCKER B. Improving RetinaNet for CT lesion detection with dense masks from weak RECIST labels [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019: 402-410.

- [6] TANG Y, GAO R, LEE H H, et al. High-resolution 3D abdominal segmentation with random patch network fusion [J]. *Medical Image Analysis*, 2020, 69(3): 101894.
- [7] TONG N, GOU S, YANG S, et al. Fully automatic multi-organ segmentation for head and neck cancer radiotherapy using shape representation model constrained fully convolutional neural networks [J]. *Medical Physics*, 2018, 45(10): 4558-4567.
- [8] XU X, ZHOU F, LIU B, et al. Efficient multiple organ localization in CT image using 3D region proposal network [J]. *IEEE Transactions on Medical Imaging*, 2019, 38(8): 1885-1898.
- [9] TANG Y, GAO R, HAN S, et al. Body part regression with self-supervision [J]. *IEEE Transactions on Medical Imaging*, 2021. DOI:10.1109/TMI.2021.3058281.
- [10] YAN K, LU L, SUMMERS R M. Unsupervised body part regression via spatially self-ordering convolutional neural networks [C]//*IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018: 1022-1025.
- [11] MEINE H, HERING A. Efficient prealignment of CT scans for registration through a bodypart regressor [J]. arXiv: 1909.08898, 2019.
- [12] KARCAALINCABA M, AKTAS A. Dual-energy CT revisited with multidetector CT: review of principles and clinical applications [J]. *Diagnostic and Interventional Radiology*, 2011, 17(3): 181-194.
- [13] BASSER P J, JONES D K. Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review [J]. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, 2002, 15(7/8): 456-467.
- [14] MILDENBERGER P, EICHELBERG M, MARTIN E. Introduction to the DICOM standard [J]. *European Radiology*, 2002, 12(4): 920-927.
- [15] GUELD M O, KOHNEN M, KEYSERS D, et al. Quality of DICOM header information for image categorization [J]. *Proc Spie Medical Imaging*, 2002, 4685: 280-287.
- [16] MUSTRA M, DELAC K, GRGIC M. Overview of the DICOM standard [C]//*2008 50th International Symposium ELMAR*, IEEE, 2008, 1: 39-44.
- [17] 张乐锋, 郑逸, 傅超. 用改进的深度差分特征识别人体部位[J]. *微型机与应用*, 2015, 34(14): 54-57.
ZHANG L F, ZHENG Y, FU C. Improved depth comparison feature for the recognition of human parts [J]. *Microcomputer & Its Applications*, 2015, 34(14): 54-57.
- [18] 田野, 姜雯, 李强. 基于CT图像的身体部位自动识别方法[J]. *计算机工程与设计*, 2017, 38(1): 247-252.
TIAN Y, JIANG L, LI Q. Automated localization of body part in CT images [J]. *Computer Engineering and Design*, 2017, 38(1): 247-252.
- [19] PARK J, KANG G, PAN S, et al. A novel algorithm for identification of body parts in medical images [C]//*International Conference on Fuzzy Systems and Knowledge Discovery*, 2006: 1148-1158.
- [20] HONG L, HONG S. Methods and apparatus for automatic body part identification and localization [J]. *U S Patent App*, 2008, 11. DOI:US20080112605.
- [21] CRIMINISI A, ROBERTSON D, KONUKOGLU E, et al. Regression forests for efficient anatomy detection and localization in computed tomography scans [J]. *Medical Image Analysis*, 2013, 17(8): 1293-1303.
- [22] LI S, XU P, LI B, et al. Predicting lung nodule malignancies by combining deep convolutional neural network and handcrafted features [J]. *Physics in Medicine & Biology*, 2019, 64(17): 175012.
- [23] ZHOU Z, LI S, HAO H, et al. A multi-objective based feature selection method for lung nodule malignancy classification [C]//*Medical Physics*, NJ USA, 2018, 45:E678.
- [24] 林鹏, 张超, 李竹良, 等. 基于深度图像学习的人体部位识别[J]. *计算机工程*, 2012, 38(16): 185-188.
LIN P, ZHANG C, LI Z L, et al. Human body part recognition based on depth image learning [J]. *Computer Engineering*, 2012, 38(16): 185-188.
- [25] ROTH H R, LEE C T, SHIN H C, et al. Anatomy-specific classification of medical images using deep convolutional nets [C]//*IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015: 101-104.
- [26] YAN Z, ZHAN Y, PENG Z, et al. Multi-instance deep learning: discover discriminative local anatomies for bodypart recognition [J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1332-1343.
- [27] ZHANG P, WANG F, ZHENG Y. Self supervised deep representation learning for fine-grained body part recognition [C]//*2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017: 578-582.

- [28] HE J, WANG Y, MA J. Radon inversion via deep learning [J]. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 2076–2087.
- [29] HE J, YANG Y, WANG Y, et al. Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction [J]. *IEEE Transactions on Medical Imaging*, 2019, 38(2): 371–382.
- [30] LEE H, HUANG C, YUNE S, et al. Machine friendly machine learning: interpretation of computed tomography without image reconstruction [J]. *arXiv*, 2018. [<https://arxiv.org/abs/1812.01068v1>]
- [31] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770–778.
- [32] HUANG G, LIU Z, MAATEN L V D, et al. Densely connected convolutional networks [C]//*IEEE Conference on Computer Vision & Pattern Recognition*, 2017: 2261–2269.
- [33] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 2818–2826.
- [34] KE Y, WANG X, LE L, et al. DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning [J]. *Journal of Medical Imaging*, 2018, 5(3): 036501. DOI:10.1117/1.JMI.5.3.036501.
- [35] YU L, SHIUNG M, JONDAL D, et al. Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols [J]. *Journal of Computer Assisted Tomography*, 2012, 36(4): 477–487.
- [36] ELBAKRI I A, FESSLER J A. Statistical image reconstruction for polyenergetic X-ray computed tomography [J]. *IEEE Transactions on Medical Imaging*, 2002, 21(2): 89–99.
- [37] KEATING K A. Use and interpretation of logistic regression in habitat-selection studies [J]. *Journal of Wildlife Management*, 2011, 68: 774–789.
- [38] YE J, JANARDAN R, QI L, et al. Feature extraction via generalized uncorrelated linear discriminant analysis [C]//*Machine Learning, Proceedings of the Twenty-first International Conference (ICML)*, Banff, Alberta, Canada, 2004.
- [39] GUO G, WANG H, BELL D, et al. KNN model-based approach in classification [C]//*OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*, 2003: 986–996.
- [40] SABARIAH M M K, HANIFA S A, SA'ADAH M S. Early detection of type II Diabetes Mellitus with random forest and classification and regression tree (CART) [C]//*International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA)*, IEEE, 2014: 238–242.
- [41] BREIMAN L. Random forests [J]. *Machine Learning*, 2001, 45(1): 5–32.
- [42] MOŽINA M, DEMŠAR J, KATTAN M, et al. Nomograms for visualization of naive bayesian classifier [C]//*European Conference on Principles of Data Mining and Knowledge Discovery*, Berlin, Heidelberg, 2004: 337–348.
- [43] SELVARAJ H, SELVI S T, SELVATHI D, et al. Brain MRI slices classification using least squares support vector machine [J]. *International Journal of Intelligent Computing in Medical Sciences & Image Processing*, 2007, 1(1): 21–33.
- [44] DAVIS J, GOADRICH M. The relationship between precision-recall and ROC curves [C]//*the 23rd International Conference on Machine Learning*, 2006: 233–240.
- [45] LOBO J M, JIMÉNEZ-VALVERDE A, REAL R. AUC: a misleading measure of the performance of predictive distribution models [J]. *Global Ecology and Biogeography*, 2008, 17(2): 145–151.
- [46] YAN K, WANG X, LU L, et al. Deep lesion graphs in the wild: relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 9261–9270.

(责任编辑 冯兆永)