

文章编号: 2095-2163(2022)08-0161-05

中图分类号: TP391

文献标志码: A

3D-CNN 在肺癌图像识别中的应用研究

李雅迪, 韩佳芳, 马琳琳

(郑州财税金融职业学院 信息技术系, 郑州 450048)

摘要: 肺癌是长期威胁人类健康的恶性疾病之一, 针对传统方法在肺癌 CT 图像分类中的预处理过程复杂、工作量大的问题, 本文提出了基于三维卷积神经网络(3D-CNN)模型的肺部 CT 图像分类方法。该模型以卷积神经网络模型为基础, 并在训练的过程中使用特定顺序输入策略, 还在公开的 Kaggle Data Science Bowl 2017 数据集上进行了实验。实验表明, 该方法对图像的分类准确率达到 76%, 比采用随机顺序的输入策略时有所提升, 能够为肺部病理图像的分类研究提供有价值的参考。

关键词: 肺部 CT 图像分类; 3D-CNN; 特定顺序输入策略

Application of 3D-CNN in lung cancer images classification

LI Yadi, HAN Jiafang, MA Linlin

(Department of Information Technology, Zhengzhou Vocational College of Finance and Taxation, Zhengzhou 450048, China)

[Abstract] Lung cancer is one of the malignant diseases that threaten human health for a long time. Aiming at the problems of complex preprocessing and heavy workload of traditional methods in lung cancer CT images classification, 3D-CNN (Three-dimensional Convolutional Neural Network) model is proposed in this paper. The model is based on the CNN model, combined with a specific order input strategy in the training process, and experiments are carried out on the Data Science Bowl 2017 data set from Kaggle. These experiments show that the classification rate of the model is 76% when the specific order input strategy is used, which is higher than that when the random order input strategy is used. The research can provide a valuable reference for the classification of lung pathological images.

[Key words] lung cancer CT images classification; 3D-CNN; specific order input strategy

0 引言

肺癌是目前严重危害人类健康的恶性疾病之一, 在全世界范围内肺癌的发病率和死亡率都在上升^[1]。虽然现有各种靶向治疗方案, 但晚期肺癌的预后很差, 一旦确诊, 绝大多数的患者会在五年内死亡, 因此肺癌的早发现、早诊断、早治疗在一定程度上能够提高存活率。现有的肺癌诊断主要是基于肺部计算机断层扫描(Computed Tomography, CT)图像的初期筛查, 现有的计算机辅助诊断(Computer Aided Diagnosis, CAD)方法就是对肺部 CT 图像进行预处理、分割、特征提取等过程^[2]。其中, 预处理过程是将肺部可疑组织从复杂的解剖图像中剥离出来, 该步骤操作复杂、工作量巨大^[3]。

随着人工智能的快速发展, 基于深度学习技术的卷积神经网络模型在图像分类领域取得了很大的进展^[4]。将该技术应用到医疗图像的分类识别中,

可以简化图像的预处理过程, 提取到更丰富的特征信息, 而且丰富的数据集和高性能的计算机软硬件也使得深度学习的训练过程更加简便、高效。例如, Google 公司研发的乳腺癌智能检测系统, 检出率高达 92%^[5]; Korbar 在 2017 年设计了精确度高达 93% 的卷积神经网络系统, 通过对肠道染色体来识别带有致癌隐患的肠道息肉等^[6]。由于本文采用的数据集是三维立体的肺部 CT 扫描图像, 而目前的卷积神经网络模型多用于处理二维平面图像, 因此在对肺部 CT 图像分类识别时, 需要将其切割成二维平面图像, 但在提取图像特征信息时会丢失 CT 图像切片之间的特征信息, 对图像的分类识别造成影响。

针对以上问题, 本文以深度学习中卷积神经网络模型为基础, 对网络结构进行调整, 构建了三维卷积神经网络(Three-dimensional Convolutional Neural Network, 3D-CNN)模型, 结合特定顺序的输入策

基金项目: 河南省高等学校重点科研项目(22B520032)。

作者简介: 李雅迪(1991-), 女, 硕士, 助教, 主要研究方向: 图像识别、人工智能、物联网应用技术; 韩佳芳(1991-), 女, 硕士, 助教, 主要研究方向: 语音识别、物联网应用技术; 马琳琳(1992-), 女, 硕士, 讲师, 主要研究方向: 计算机应用技术。

通讯作者: 李雅迪 Email: 519303773@qq.com

收稿日期: 2022-04-17

略,并在公开的 Kaggle Data Science Bowl 2017 数据集上进行实验。实验结果表明,本文提出的方法能够对原始图像进行有效分类识别,省略了传统方法中复杂的图像预处理过程,有较好的识别率。

1 本文研究方法

1.1 三维卷积神经网络模型

本文中构建的 3D-CNN 模型架构如图 1 所示。

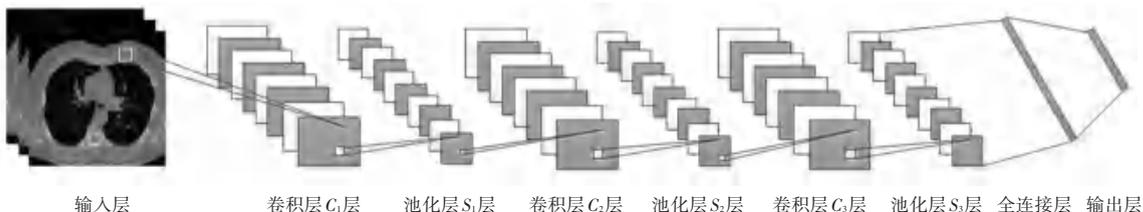


图 1 3D-CNN 架构图

Fig. 1 The structure of 3D-CNN

本文提出的端到端的三维卷积神经网络模型作为肺部 CT 图像的分类模型,其层级结构的具体设置如下:

(1) 3D 卷积层。先设定好卷积核的尺寸、步长、填充选项和卷积核的数量,然后与多个相邻的图像帧做加权求和的卷积计算,来提取目标图像特定区域的特征信息,不同的卷积核应用在图像的不同区域,来提取图像的多种特征信息。第 j 层的第 n 个特征图的 (x, y, z) 处的输出,研究推导出的数学公式可写为:

$$O_n^j(x, y, z) = b_{jn} + \sum_n \sum_{p=0}^{p_j-1} \sum_{q=0}^{q_j-1} \sum_{r=0}^{r_j-1} w_{mn}^j(p, q, r) \cdot I_{im}(x+p, y+q, z+r) \quad (1)$$

其中, b_{jn} 是偏置量; (p_j, q_j, r_j) 分别是卷积核的高度、宽度和深度; w_{mn}^j 是第 i 层的第 m 个特征图与第 j 层的第 n 个特征图之间的权重值; I 为第 i 层的第 m 个特征图的输入。

(2) 激活层。由于在卷积层中做加权求和的线性运算,输出的特征图谱带有明显的线性特征,而线性特征的表现力不足以表达图像的关键信息,需要通过激活函数添加非线性因素,将图像特征保留并映射出来。激活函数主要有非线性、可微性、单调性和在 0 点处近似线性等特性,能够使训练快速收敛,解决梯度弥散的问题。本文中采用了修正线性单元 (Rectified Linear Unit, *ReLU*) 激活函数, *ReLU* 有单边抑制特性,小于 0 的值全部设为 0,大于 0 的值则直接输出,数学公式具体如下:

$$ReLU(x) = \max(0, x) \quad (2)$$

图 1 中,首先对所有的肺部 CT 图像进行简单的预处理,按照先正常人、后肺癌病人的 CT 图像的特定顺序重新组织数据集,预处理后的肺部 CT 图像分别经过卷积层 C_1 层、池化层 S_1 层、卷积层 C_2 层、池化层 S_2 层、卷积层 C_3 层和池化层 S_3 层来提取图像的主要特征信息,全连接层将提取到的分布式特征信息进行整合,判断输入图像所属类别。

(3) 3D 池化层。池化层的功能是对上层激活层输出的特征图谱下采样,通过该层将特征图谱关键的特征像素标注出来,在一定程度上简化了参数的运算过程,抑制了过拟合的现象,同时能够提高模型的泛化能力。池化方法主要有均值池化 (Mean Pooling)、最大池化 (Max Pooling)、重叠池化 (Overlapping Pooling)、均方池化 (L_2 Pooling)、归一化池化 (Local Contrast Normalization)、随机池化 (Stochastic Pooling)、形变约束池化 (Def-pooling) 等等。

为了简化各层的计算过程,提取主要特征,本文选取了最大池化方法,取局部接受域中的最大值为该层输出,推得的数学公式见式 (3):

$$O(x, y, z) = \max_{(0 \leq i \leq s_1, 0 \leq j \leq s_2, 0 \leq k \leq s_3)} (I_{x \times p+i, y \times q+j, z \times r+k}) \quad (3)$$

其中, I 为池化层的输入; (p, q, r) 分别是高度、宽度和深度三个维度上的池化尺寸; (i, j, k) 为 3 个维度上的采样步长值。

(4) 全连接层。全连接层 (Fully Connected layers, FC) 通过对上层输出的特征图谱进行卷积计算,将获取到的分布式特征映射到样本标记空间,将这些分布式特征组装成完整的“图”,确定该特征图谱所属分类。该层的优点在于减少特征位置对分类结果的影响,提高整个网络模型的鲁棒性。

(5) Dropout 算法。由于条件的限制,数据集的种类和质量上存在不足之处,比如类型单一、可用的数据量太少等,导致在模型训练的过程中出现分类结果曲线与训练数据曲线重合的情况,即过拟合现

象^[7]。本文采用 *Dropout* 算法, 在模型训练的过程中先将输入数据进行正向地传输, 并将一部分的神经元以概率 p 丢弃掉, 其余神经元以 $q = 1 - p$ 的概率保留, 再将该过程出现的误差反向重传回去。

通过 *Bernolulli* 函数, 实现以概率 p 生成 0 值, 以概率 $q = 1 - p$ 生成 1 值, 相应的数学公式可表示为:

$$r_j^{(l)} = \text{Bernolulli}(p) \quad (4)$$

再将 l 层网络输出的神经元 $y^{(l)}$ 以概率 p 丢弃掉, 生成 $\tilde{y}^{(l)}$, 这里给出的公式表述如下:

$$\tilde{y}^{(l)} = r^{(l)} * y^{(l)} \quad (5)$$

将 $\tilde{y}^{(l)}$ 传输到下一层, 通过权重值为 $w_i^{(l+1)}$ 、偏置量为 $b_i^{(l+1)}$ 的加权求和的卷积计算提取特征信息, 生成的 $z_i^{(l+1)}$, 于是有:

$$z_i^{(l+1)} = w_i^{(l+1)} \tilde{y}^{(l)} + b_i^{(l+1)} \quad (6)$$

将生成的 $z_i^{(l+1)}$ 经过激活函数, 为上层特征信息添加非线性因素, 增强特征表现力, 最终得到 $l + 1$ 层的输出 $y_i^{(l+1)}$, 相应计算公式为:

$$y_i^{(l+1)} = f(z_i^{(l+1)}) \quad (7)$$

(6) *Softmax* 函数。*Softmax* 函数又称归一化指数函数, 是将多分类的结果以概率的形式呈现, 在逻辑回归模型的基础上解决类别大于 1 的分类问题。经过 *Softmax* 函数的计算可得到图像属于某种分类的概率值, 该值可由如下公式计算求出:

$$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_{j=1}^j \exp(x_j)} \quad (8)$$

其中, x_i 为第 i 个节点的输出值, j 为输出节点的个数, 即分类的类别数。

(7) 交叉熵损失函数。采用的交叉熵损失函数 (Cross Entropy Loss Function), 通过计算真实概率分布与预测概率分布之间的损失值来评价卷积神经网络的性能优劣, 值越小, 说明模型的分类识别效果越好。此时需用到的数学公式为:

$$\text{Loss} = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (9)$$

其中, n 为测试集中的数据个数; \hat{y}_i 为癌症患者的预测概率; y_i 为每个图像的标签, 癌症患者标签为 1, 正常人标签为 0。

1.2 特定顺序输入策略

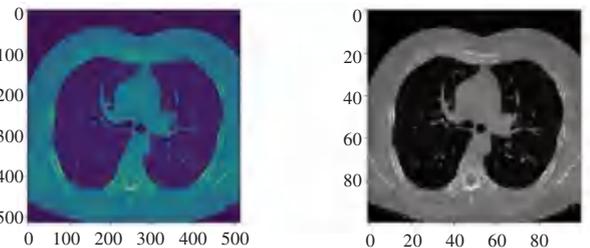
本文采用了一种特定顺序的输入策略来指导模

型的训练过程。将正常人的肺部 CT 扫描图像放在一起, 后面拼接上肺癌高危患者的肺部 CT 扫描图像, 形成一个整体的数据集, 再将其输入到 3D-CNN 模型中进行训练。使用该策略能提高模型学习的速度和质量, 优化卷积神经网络模型的训练过程。

2 实验与分析

2.1 数据集

本文的数据来自自由 Kaggle 主办的数据科学竞赛 Data Science Bowl 2017 的数据集。该数据集是由 1 397 个 .dcm 格式的肺部 CT 扫描图像组成, 其中 1 036 个图像为正常人的肺部图像, 361 个图像为肺癌高危患者肺部图像。先按照 CT 图像中图像位置属性对切片进行排序, 并实现三维渲染; 将 512×512 大小的图片缩放为 100×100 , 并对其进行灰度化处理; 最后, 将三维 CT 图像合并成数量为 20 的切片组合, 得到 $100 \times 100 \times 20$ 的三维灰度图像。这里以一个肺癌患者的 CT 图像预处理为例, 单个切片的原始图像和预处理后的图像如图 2 所示。



(a) 原始图像

(b) 预处理后图像

图 2 单个切片的原始图像和预处理图像

Fig. 2 Raw and pre-processed images of a single slice

2.2 实验过程

首先, 按照预定的输入顺序重新组织训练样本, 即将常规的肺部 CT 扫描图像放在一起, 拼接上患癌的肺部 CT 扫描图像, 形成一个整体的训练数据集, 将其输入到 3D-CNN 模型中进行训练。

该 3D-CNN 模型在对重组后的数据集进行训练后, 就要对未知肺部 CT 扫描图像来做分类识别。其中, 第一层卷积层 C_1 层 32 个大小为 $5 \times 5 \times 5$ 不同的卷积核过滤, 生成 32 个大小为 $100 \times 100 \times 20$ 的特征图谱, 该层产生的变量总数为 $32 \times 100 \times 100 \times 20$; 再将这 32 个特征图谱输入进第一层池化层 S_1 层进行最大池化操作, 该层选用的池化步长为 $2 \times 2 \times 2$, 输出 32 个大小为 $50 \times 50 \times 10$ 的特征图像, 产生的变量总数为 $32 \times 50 \times 50 \times 10$; 第二层卷积层 C_2 层采用 64 个大小为 $3 \times 3 \times 3$ 的不同的卷积核进行填充, 得到的

64个大小为 $50 \times 50 \times 10$ 特征图谱,产生的变量总数为 $64 \times 50 \times 50 \times 10$;再进行第二层池化层 S_2 , S_2 层的参数设置参照 S_1 层,由此得到64个大小为 $25 \times 25 \times 5$ 的特征图谱,而变量总数为 $64 \times 25 \times 25 \times 5$;最后将输出的特征图谱进行第三层的卷积操作,第三层 C_3 选择128个大小为 $3 \times 3 \times 3$ 的不同的卷积核,经过该

层后生成128个大小为 $25 \times 25 \times 5$ 的特征图谱,变量总数为 $128 \times 25 \times 25 \times 5$;再将此层生成的图像输入最后一层池化层 S_3 中,其参数设置与 S_1 、 S_2 保持一致,输出128个大小为 $13 \times 13 \times 3$ 的特征图谱,得到的变量总数为 $128 \times 13 \times 13 \times 3 = 64\ 896$ 。3D-CNN数据处理过程如图3所示。

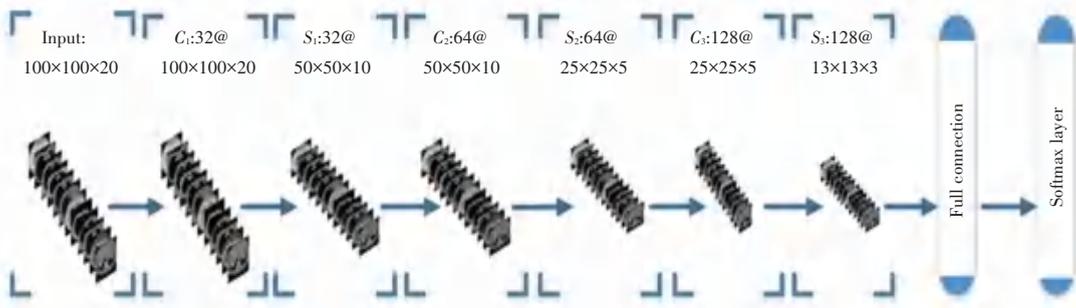


图3 3D-CNN数据处理过程

Fig. 3 3D-CNN data processing

2.3 实验对比分析

将932个正常人的肺部CT图像数据和325个肺癌高危患者的肺部CT图像数据混合,形成一个随机顺序的数据集,将其输入进该3D-CNN模型中进行训练,比较2种方法的性能。根据分类正确的样本个数占所有样本个数的比例来计算准确率,并根据公式(9)计算损失值,实验结果分别见表1和表2。由表1、表2可知,采用特定顺序输入策略的3D-CNN模型的分类准确率最高为76%,高于采用随机顺序输入策略的3D-CNN模型的最高准确率(70%),说明特定顺序的输入策略提升了模型的训练精度,从而有效提高肺癌图像的分类准确率。

表1 3D-CNN模型在特定顺序输入策略的结果

Tab. 1 Results of 3D-CNN model in specific order input strategy

组别	准确率	损失值
1	0.685 714	1.469 389 0
2	0.689 914	1.063 208 5
3	0.592 857	2.151 141 5
4	0.697 857	0.927 307 0
5	0.757 143	1.200 940 8
6	0.757 143	1.567 617 4
7	0.756 613	0.826 516 5
8	0.759 876	0.719 830 1
9	0.762 257	0.873 349 6
10	0.764 431	0.933 137 9

表2 3D-CNN模型在随机顺序输入策略的结果

Tab. 2 Results of 3D-CNN model in random order input strategy

组别	准确率	损失值
1	0.614 28	1.424 743 6
2	0.678 57	0.636 111 5
3	0.692 85	0.599 292 7
4	0.700 00	0.578 136 3
5	0.700 00	0.571 537 0
6	0.707 14	0.574 485 3
7	0.707 14	0.570 952 7
8	0.707 14	0.569 861 9
9	0.707 14	0.568 439 8
10	0.707 14	0.567 721 3

3 结束语

为解决传统方法在肺癌CT图像分类中的预处理过程复杂、工作量大的问题,本文提出了基于3D-CNN的肺部CT图像分类模型。该模型以卷积神经网络为基础,设计了3D卷积层、激活层、3D池化层以及全连接层等层级来获取图像的分类特征,结合特定顺序的输入策略来优化模型的训练过程,并与采用随机顺序输入策略进行了对比分析。结果表明,该模型对肺癌CT图像有良好的分类特性,最高识别准确率达到76%,对于肺癌的早期诊断有一定的实用价值。但本文提出的3D-CNN模型在架

(下转第170页)