

基于深度卷积神经网络的脑部图像疾病类别特征库构建*

万艳丽 胡红濮 陈 荃 高 星 王 岩

(中国医学科学院/北京协和医学院医学信息研究所 北京 100020)

〔摘要〕 介绍深度卷积神经网络基本理论, 阐述基于深度卷积神经网络的脑部图像视觉特征提取, 设计一种适用于脑部疾病图像的分类器, 进而实现脑部图像疾病类别特征库构建, 为基于脑部图像疾病类别特征库开展临床辅助决策等应用提供可能。

〔关键词〕 卷积神经网络; 图像视觉特征; 脑部图像

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2018.03.003

Building of Brain Image Disease Category Characteristics Database Based on Deep Convolutional Neural Network WAN Yan-li, HU Hong-pu, CHEN Quan, GAO Xing, WANG Yan, Institute of Medical Information, Chinese Academy of Medical Sciences/Peking Union Medical College, Beijing 100020, China

〔Abstract〕 The paper introduces the basic theory of deep convolutional neural network and expounds on visual feature extraction of brain images based on deep Convolutional Neural Network (CNN), designs a kind of categorizer that is applicable to brain disease image to build brain image disease category characteristics database, making applications like clinical assistant decision-making based on brain image disease category characteristics database possible.

〔Keywords〕 Convolutional Neural Network (CNN); Image visual features; Brain images

1 引言

随着现代医学的发展, 医院的诊疗活动越来越多地依赖医学影像的检查 (如 X 线、CT、MR、超声、内镜、血管造影等), 医院内部所产生的数据

量约 80% 来自于医学图像资料。面对这些日益增长的海量级医学图像资源, 基于现有的影像存储与传输系统 (Picture Archiving and Communication Systems, PACS), 构建更加高效合理的医学图像特征库, 组织和管理好这些来源于各种各样的设备、各种不同的人体组织器官和多种病症的海量医学图像资源, 以便于对某些疑难病症进行归类和分析, 是实现基于医学案例的数据管理及循证医学的关键^[1-3]。

按疾病类别构建医学图像类别特征库可以实现具有类似病理特征的图像快速定位和检索, 对辅助诊断、教学科研和远程医疗等应用领域具有重要的应用价值。本文以脑部图像为例, 基于深度卷积神

〔修回日期〕 2017-06-13

〔作者简介〕 万艳丽, 副研究员, 博士; 通讯作者: 胡红濮。

〔基金项目〕 国家自然科学基金项目 (项目编号: 61703436); 中央级公益性科研院所基本科研业务费项目 (项目编号: 2016ZX330013)。

神经网络设计一种适用于脑部疾病图像的分类器，进而实现脑部图像疾病类别特征库的构建，为基于脑部图像疾病类别特征库开展临床辅助决策等应用提供可能。

2 深度卷积神经网络基本理论

2.1 概念

卷积神经网络（Convolutional Neural Network, CNN）的概念已经存在了几十年，是广泛应用在模式识别、图像处理领域的一种高效识别算法，具有结构简单、训练参数少和适应性强的特点。但采用多个堆叠层训练的深度网络最近几年才受到广泛关注。2012 年 Krizhevsky 等人^[4]提出的深度卷积神经网络架构，也称为 AlexNet，由 5 个卷积层与 3 个全卷积层组成，在 ImageNet 大尺度视觉识别挑战赛中，将一个数据集中 120 万张图像分为 1000 个不同类别，他们设计的模型与其他学者的模型相比在识别性能方面取得了实质性的提升。近年来虽有大量关于用于彩色图像分类的深度卷积神经网络的研究，但关于医学图像分析、处理和识别问题的研究相对较少^[5]。

2.2 原理

深度卷积神经网络实际上是一个多层的感知器，每层由多个二维平面组成，而每个平面由多个独立神经元组成，网络中包含一些简单元和复杂元，分别记为 C 元和 S 元。C 元聚合在一起构成卷积层，卷积层是卷积神经网络的核心层，用它来进行特征提取，每个神经元的输入与前一层的局部感受野相连，提取该局部的特征，通常会使用多层卷积层来得到更深层次的特征图；S 元聚合在一起构成下采样层，也称为池化层，实现对特征图的压缩。图 1 中输入图像通过一组卷积核进行卷积运算，再加上一组偏置项，在 C 层产生 N 个特征图，然后特征图通过一个激活函数（Logistic、Sigmoid、RELU 等函数）得到 S 层的特征图。根据设定的 C 层和 S 层的数量，以上过程依此循环。最终对最尾部的卷积层和输出层进行全连接，然后将输出值送给分类器。

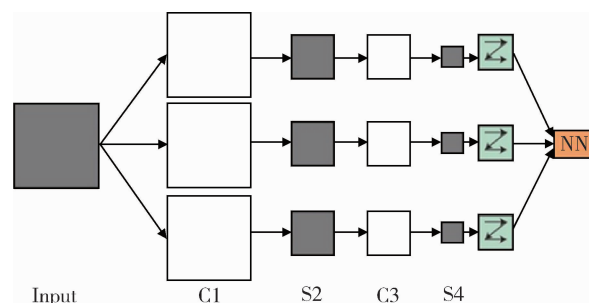


图 1 深度卷积神经网络的原理

3 基于深度卷积神经网络的脑部图像视觉特征提取

3.1 概述

近年来以 ImageNet 的竞赛为标志，深度学习技术在大尺度彩色图像的分类方面取得了长足的进步，各种各样有效的分类器也相继出现，如经典的 AlexNet^[4]，GoogleNet^[6]，ResNet 等。但是这些模型直接应用到医学图像上时，所取得的精度却远远无法与自然场景的彩色图像分类的精度相比。这主要是因为医学图像与彩色自然图像相比，有自己的特殊性，这种特征性导致了两者之间的巨大差异，也对直接应用这些经典的成熟的模型形成了一定阻碍。具体来说，自然场景图像由于场景之间存在较大的差异，而且这种差异更多的表现为一种全局的宏观特征的不同，因此在处理自然场景图像的分类时，学者们所设计和网络都是在逐层的缩小特征图的尺度，从而可以慢慢聚拢一种全局的特征，而这种全局的特征也恰恰是区分各个场景的最有效特征。但是对于医学图像来说，尤其是对于同一器官的不同疾病种类来说，全局特征未必是最有效的特征。如对于大脑的 CT 切片图像，从宏观上来看其结构都是类似的，而疾病往往只是大脑扫描图像中的一个小的区块。要区分不同的疾病种类往往要借助于这些小区块的特征。如果采用像自然场景图像一样的分类网络，这些小区块的信息可能会在池化操作（Pooling）的过程中逐渐的丢失，从而使最终的分类精度降低。因此对于医学图像的特征提取与分类，尤其是同一器官部位的扫描图像的特征提取与分类，往往需要同时考虑全局特征与局部特征的

提取，其中全局特征有利于区分不同方向或不同角度的切片，而局部特征则有利于对同一种切片下的不同疾病种类作出判断。

3.2 基本模块设计

基于医学图像的特点提出一种在多尺度上对图像进行特征得取的网路，即可以提取宏观的全局特征，也可以提取局部的病理特征。最后将这些全局的和局部的特征进一步通过全连接层进行整合，从而得到最终用于分类的视觉特征。所提取的特征兼顾了轮廓信息和局部纹理信息，信息损失比较小，具有足够强的表达能力和区分能力。本文所使用的网路的基本模块，见图 2。这个基本模块是一个残差结构，是著名的 ResNet 的基本模块一个变形。与 ResNet 的基本模块相比，基本模块本身具体多尺度的特征，在这个模块的主分支里，有 3 个连续的卷积，由于每一个卷积都是上一个卷积的基础上进行的运算，所以这 3 个卷积的感受域在逐渐增大，从而可以逐步提取全局上的特征，而 Concat 这一层就可以将这些局部的和全局的特征合并在一起。另外 Add 层的存在也有将不同尺度的特征进行融合的功能，而且还可以加速网路的收敛速度。

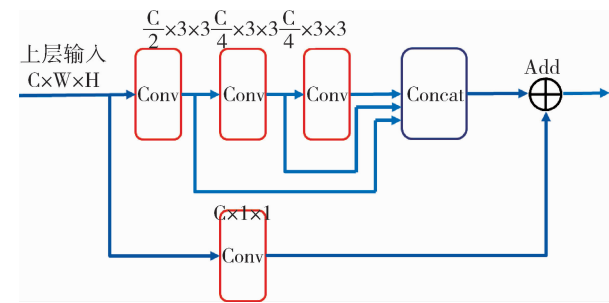


图 2 基本模块

3.3 整体卷积神经网络结构设计

除了上面所提到在基本块里进行多尺度的特征提取之外，在网路中加入了池化层，将多个基本块与池化层有机的组合，进而实现更有效的多尺度特征提取。完整的网路结构图，见图 3。图 3 中所示的 BaseModule 就是对应图 2 中的基本块。Mx 或 Nx 是指本层的特征图的通道数，W 和 H 是指本层特征图的宽高值。没有特殊指明的情况下，所有的卷积层都使用 3 × 3 大小的核。本文所使用的网路可以

在原尺度、1/2 尺度和 1/4 尺度 3 个尺度上对图像的特征进行提取与合并，有效利用了全局特征与局部特征，使得特征的区分能力和表达能力更强。

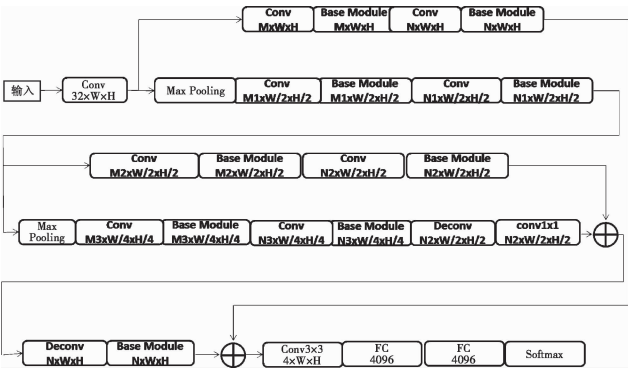


图 3 所使用的完整网络

3.4 分类函数选择

由于 Logistic 函数只能被使用在二分类问题中，但是它的多项式回归，即 Softmax 函数，可以解决多分类问题，当然，使用 Logistic 函数建立 k 个独立的二元分类器，也可以实现多分类问题，但是由于脑部图像疾病类别的划分是一种互斥的关系，因此本文更适于选择 Softmax 函数作为分类函数。Softmax 函数其实就是一个归一化的指数函数，其定义如下：

$$P(i) = \frac{\exp(\theta_i^T x)}{\sum_{k=1}^K \exp(\theta_k^T x)} \tag{1}$$

其中 θ_i 和 x 是列向量， $\theta_i^T x$ 可能被换成函数关于 x 的函数 $f_i(x)$ 。通过 softmax 函数，可以使得 $P(i)$ 的范围在 $[0, 1]$ 之间。在回归和分类问题中，通常 θ 是待求参数，通过寻找使得 $P(i)$ 最大的 θ_i 作为最佳参数。

4 脑部图像疾病类别特征库构建

根据上述深度神经网络结构，提取脑部图像类别特征，并建立训练视觉特征到疾病类别语义特征的映射模型，建立视觉特征与疾病类别语义的关联，构建脑部图像疾病类别特征库，为高效检索应用提供保障。本文采用的脑部图像特征库构建与病理图像检索流程，见图 4。图中所示为本文所设计的检索系统的主要流程，分 3 个部分。其中红色箭

头所指示的流程为训练过程，输入是一批经过标注的训练图像。这些训练图像经过深度卷积神经网络的计算提取到当前参数下的特征，然后送入 Softmax 分类器进行预设，将预测结果与标注的标签进行对比从而计算出分类损失，然后经过反向传导算法调整网络的参数。经过多轮迭代之后，就可以得到训练好的参数。本文借助于一个分类任务来训练网络并提取特征，主要是因为分类与检索存在语义上的相似性。紫色所指示的流程为特征库构建过程，输入是医学图像库中的脑部扫描图像。将这些图像依次通过训练好的网络可以计算得到其对应的特征，这个特征可以是原始的网络计算得到的特征，也可以根据需要进行二次加工，如进行哈希变换等。将这些特征存放到数据库中就建成了一个脑部图像特征数据库。绿色箭头所指示的过程为图像检索过程，将待处理的病理图像通过网络计算得到其特征，然后与特征库中的特征进行对比，从而检索到相应的病理图像作为辅助诊断的参考。

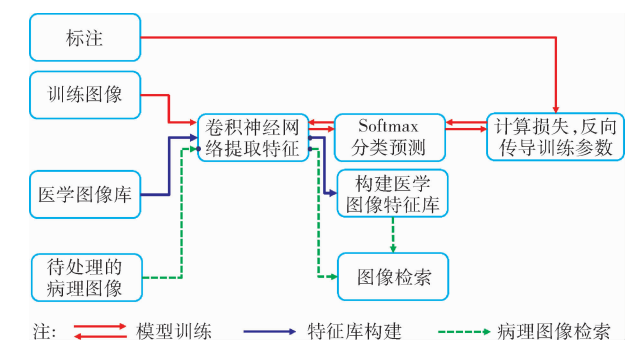


图 4 脑部图像特征库构建与病理图像检索流程

5 检索实验与结果分析

5.1 实验数据

训练数据采集了包含 3 种脑部疾病的图像数据。一是从某医院影像科采集的包括脑出血、垂体瘤两种病种各 200 例病案的图像数据，其中图像模态包括 CT 和 MR 两种。同时，从美国国家癌症研究所（National Cancer Institute, NCI）创建的开源开放的癌症图像存档系统（The Cancer Imaging Archive, TCIA）中采集了 199 例胶质瘤患者的 CT 和

MR 两种模态的图像数据。利用这些训练数据构建脑部图像对应的疾病类别特征库。

5.2 相似性测度

检索测试时，将提取每幅待检索图像的视觉特征向量，将其与特征库中的特征向量进行相似性计算。本文网络所生成的原始特征为 4 096 维的一个向量，因此采用向量的欧式距离为作其相似度的度量公式，公式如下：

$$S_{v,v_0} = v - v_{02} \tag{2}$$

式中为当前待检索图像计算所得到的特征向量，是特征库中的某一个特征向量。将相似性距离按照从小到大的顺序排序，输出相应图像作为当前检索图像的参考。

5.3 评价方法

查准率（Precision）和查全率（Recall）是目前公认的衡量图像检索效果的相对合理的重要的评价指标。公式如下：

$$P = a / (a + b), R = a / (a + c) \tag{3}$$

其中 P 表示查准率， R 表示查全率， a 为检索返回的满足条件的图像数， b 为检索返回的不满足条件的图像数。 c 为系统中满足条件但未被检索出的图像数。精确度和检索率越高，表明该检索系统的效果越好。通常，检索率和精确度是一对矛盾，当要求精度较高时，检索率较低，反之亦然。因此，一般的检索系统只要求在这两者之间达到一个最优的平衡点，就认为实现了较好的检索性能。查全 - 查准率适用于反映整个检索出的结果上的总体性能，但对检索结果中不同的排序不敏感，为此，本文采用一组查询用例上的平均准确率（Mean Average Precision, MAP）和平均查全率作为测度，一方面提高排序的敏感度，另一方面降低查询用例引起的评测的不稳定性。其中 MAP 计算公式如下：

$$AP(q) = \frac{1}{N_R} \sum_{n=1}^{N_R} P_q(R_n)$$
$$MAP = \frac{1}{|Q|} \sum_{q \in Q} AP(q) \tag{4}$$

式中 $P_q(R_n)$ 表示对应查全率为 R_n 时的查准率 P_q 的值， Q 为查询用例集合， $|Q|$ 为查询用例数量。

5.4 结果分析

从本文采集的 3 种脑部疾病的图像中，每种疾病图像选取 50 幅作为测试数据，即查询用例用于检索测试。图 5 显示了采用本文的检索方法对胶质

瘤患者图像检索测试所得到的两个结果示例。其中最左侧的图像为检索图像，右面的 5 幅图像为从图像库中检索得到的结果。可以看出检索得到的结果与给定的检索图像非常相似，说明本文所提出的网络在医学特征提取方面十分有效。

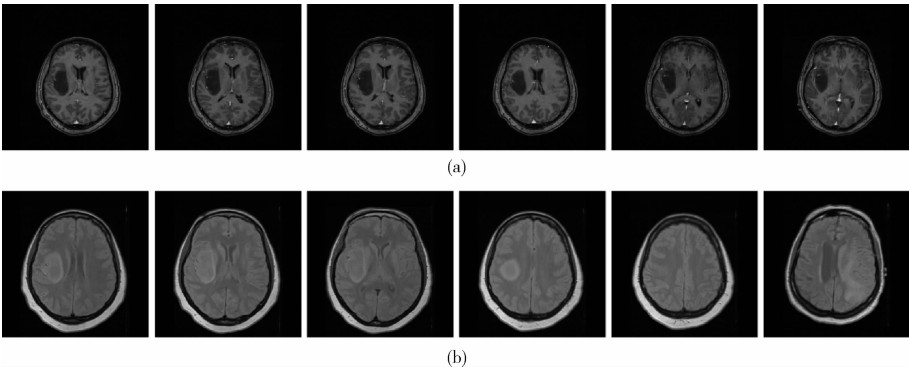


图 5 检索结果示例

表 1 所示为 3 种疾病各选取 50 幅作为查询用例检索得到的平均查全率和查准率。

表 1 平均查准率和查全率（%）

含病灶的脑部图像	平均查准率	平均查全率
脑出血图像	72.5	62.8
垂体瘤图像	84.8	76.6
胶质瘤图像	78.3	66.2

6 结语

在临床实际应用中，医生除了希望从图像库中检索出具有整体相似性的脑部肿瘤图像外，更希望检索出病灶位置相同、肿瘤性质一致的图像，为图像检测任务提出了更高的要求。为此，本文提出了一种基于卷积神经网络的脑部图像疾病类别特征库构建方法，所采用的卷积神经网络可以在多尺度图像上提取医学图像对应的视觉特征，基于构建的视觉特征库，从而可以较快速精准地检索出与查询用例图像具有相似病例特征的脑部图像，为临床医生的诊断提供参考。今后的工作重点将结合临床知识，考虑肿瘤的空间位置和肿瘤性质特点，进一步提高检索的平均查准率和查全率，以满足临床医师需求。

参考文献

1 Akgül C B, Rubin D L, Napel S, et al. Content – based Image Retrieval in Radiology: Current Status and Future Directions [J]. Journal of Digital Imaging, 2010, 24 (2): 208 – 222.

2 Welter P, Fischer B, Günther R W, et al. Generic Integration of Content – based Image Retrieval in Computer – aided Diagnosis [J]. Computer Methods & Programs in Biomedicine, 2012, 108 (2): 589 – 599.

3 Mourão A, Martins F, Magalhães J. Multimodal Medical Information Retrieval with Unsupervised Rank Fusion [J]. Computerized Medical Imaging & Graphics the Official Journal of the Computerized Medical Imaging Society, 2015, (39): 35 – 45.

4 Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks [C]. International Conference on Neural Information Processing Systems. Curran Associates Inc, 2012: 1097 – 1105.

5 Avni U, Greenspan H, Konen E, et al. X – ray Categorization and Retrieval on the Organ and Pathology Level, Using Patch – based Visual Words [J]. IEEE Transactions on Medical Imaging, 2010, 30 (3): 733 – 746.

6 Ng Y H, Yang F, Davis L S. Exploiting Local Features from Deep Networks for Image Retrieval [C]. Computer Vision and Pattern Recognition Workshops. IEEE, 2015: 53 – 61.