

DOI:10.16136/j.joel.2022.06.0648

基于图卷积神经网络的胸部放射影像疾病分类方法

赵佳雷¹, 黄青松^{2*}, 刘利军³, 黄 冕⁴

(1. 昆明理工大学 信息工程与自动化学院, 云南 昆明 650500; 2. 昆明理工大学 云南省计算机技术应用重点实验室, 云南 昆明 650500; 3. 云南大学 信息学院, 云南 昆明 650091; 4. 云南国土资源职业学院信息中心, 云南 昆明 652501)

摘要:医学 X 射线作为胸部疾病的常规检查手段,可以对早期不明显的胸部疾病进行诊断,并且观察出病变部位。但是,同一张放射影像上呈现出多种疾病特征,对分类任务而言是一个挑战。此外,疾病标签之间存在着不同的对应关系,进一步导致了分类任务的困难。针对以上问题,本文将图卷积神经网络(graph convolutional neural network, GCN)与传统卷积神经网络(convolutional neural network, CNN)相结合,提出了一种将标签特征与图像特征融合的多标签胸部放射影像疾病分类方法。该方法利用图卷积神经网络对标签的全局相关性进行建模,即在疾病标签上构建有向关系图,有向图中每个节点表示一种标签类别,再将该图输入图卷积神经网络以提取标签特征,最后与图像特征融合以进行分类。本文所提出的方法在 ChestX-ray14 数据集上的实验结果显示对 14 种胸部疾病的平均 AUC 达到了 0.843,与目前 3 种经典方法以及先进方法进行比较,本文方法能够有效提高分类性能。

关键词:图卷积神经网络; 胸部放射影像; 疾病诊断; 医学图像处理

中图分类号:TP391.4 **文献标识码:**A **文章编号:**1005-0086(2022)06-0667-06

Classification of chest radiographic image diseases based on graph convolutional neural network

ZHAO Jiale¹, HUANG Qingsong^{2*}, LIU Lijun³, HUANG Mian⁴

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650500, China; 2. Yunnan Key Laboratory of Computer Technology Applications, Kunming University of Science and Technology, Kunming, Yunnan 650500, China; 3. School of Information, Yunnan University, Kunming, Yunnan 650091, China; 4. Yunnan Land and Resources Vocational College Information Center, Kunming, Yunnan 652501, China)

Abstract: Medical X-rays, as a routine examination method for chest diseases, can diagnose early and un-obvious chest diseases and observe the lesions. However, the characteristics of multiple diseases on the same radiographic image are a challenge to the classification problem. In addition, there are different correspondences between disease labels, which further leads to the difficulty of classification tasks. In response to the above problems, this paper combines the graph convolutional neural network (GCN) with the traditional convolutional neural network (CNN), and proposes a multi-label chest radiographic image disease classification method that combines label features with image features. This method uses the graph convolutional neural network to model the global correlation of the labels, that is, constructs a directed relationship graph on the disease label, each node in the directed graph represents a label category, and then inputs the graph into the graph convolutional neural network to extract the label features, and finally merges with the image features to sort. The experimental results of the method proposed in

* E-mail: ynkmlhqs@sina.com

收稿日期: 2021-09-14 修订日期: 2021-11-25

基金项目: 国家自然科学基金(81860318, 81560296)和重点实验室开放基金重点项目(2020106)资助项目

this paper on the ChestX-ray14 dataset show that the average AUC of 14 chest diseases reaches 0.843. Compared with the current three classic methods and advanced methods, the method in this paper can effectively improve the classification performance.

Key words: graph convolutional neural network (GCN); chest radiographic image; disease diagnosis; medical image processing

1 引言

胸部疾病是威胁人类健康的一大难题,全世界每年有数百万人受到胸部疾病的困扰,在不发达的地区或者国家更为明显^[1]。胸部X射线光片(chest X-ray radiographs, CXR)是目前临床上使用最广泛的医学成像技术^[2],对肺结核、肺炎、气胸等各种胸部疾病的早期发现具有十分重要的作用。然而胸部放射影像的解读需要专业的放射科医生来执行,每天解读成百上千的放射影像是一项繁琐的任务,并且长时间大量的解读容易造成误诊和漏诊^[3,4],因此,可靠的放射影像计算机辅助诊断系统具有十分重要的意义。

近年来,随着人工智能(artificial intelligence, AI)技术和计算机硬件的发展,深度学习(deep learning, DL)已经在计算机图像自动识别领域取得了重大的研究进展,科研人员逐渐开始将AI技术应用到胸片计算机辅助诊断系统中^[5,6]。美国国立卫生研究院(National Institutes of Health, NIH)公开发布了胸部放射影像数据集 ChestX-ray14^[7],进一步推动了深度学习算法在胸部放射影像计算机辅助诊断中的发展^[8-11]。WANG等^[7]在4种ImageNet^[12]预训练网络模型上对ChestX-ray14数据集进行分类,有效地利用了ImageNet预训练结果,却未考虑标签的全局相关性,YAO等^[13]加入了长短记忆网络(long-short term memory networks, LSTM),学习标签之间的依赖性关系以提高识别能力,但LSTM以预先定义的顺序去预测标签并不能完全掌握标签的全局信息。MA等^[14]提出一种多注意力机制网络来进行分类,有效地关注到了病灶区域,然而依旧没有学习到标签的全局相关性。

目前大多数研究仅对图像进行处理,而且传统的卷积神经网络(convolutional neural network, CNN)无法对非欧式数据进行卷积。针对上述方法存在的不足,为了更好地解决标签的全局相关性问题,本文提出了一种基于图卷积神经网络(graph convolutional neural network, GCN)的多标签胸部放射影像疾病分类方法,该方法有以下两

个方面的贡献:1)由于在医学影像诊断中疾病的标签属于非欧式空间数据,所以引入图卷积神经网络来提取标签特征。2)由于目前公开的医学影像数据集都没有提供标签的关系数据,因此使用数据挖掘的方式构建了标签相关矩阵,并且将标签用词嵌入方法表示。

2 研究方法

2.1 多标签胸部疾病图卷积神经网络

本文提出了一种多标签胸部疾病图卷积神经网络模型(ML Chest-GCN)。总体框架如图1所示,该模型由GCN与CNN模块构成,其中医学影像从CNN模块输入,经过一系列卷积操作之后提取出胸部医学影像的特征。对于GCN模块而言,首先采用预训练语言模型将医学影像标签转换为GloVe^[15]词嵌入表示,然后采用数据挖掘的方式构建标签关系图矩阵输入到GCN模块,经过两层的卷积操作提取出疾病标签特征。最后,将医学影像特征与疾病标签特征融合来预测多标签结果。

2.2 图像特征提取

如图1(b)所示,本文中的医学影像特征提取采用了DenseNet-121网络的特征提取部分,DenseNet^[16]是由HUANG等提出的一种新的CNN结构,其核心思想是建立所有层之间的互联,将前面所有层的直接连接作为当前层的输入,使信息可以在不同层之间传递。

该网络由4个不同大小的密集网络块DenseBlock1-4组成,其中每个密集网络块由不同个数的卷积单元组成。每个卷积单元包含 1×1 的卷积和 3×3 卷积以及相应的批归一化处理 and Relu 激活函数。密集网络块之间的层是过渡层,其作用是对特征图进行下采样,DenseNet-121共包含3个过渡层,并且为了更好地提取纹路特征,本文去掉DenseNet-121网络的全连接层,替换为最大池化层。

本文的图像输入为 224×224 pixel的医学影像,经过conv5之后的图像特征维度为 $1024 \times 7 \times 7$,最后通过 7×7 全局最大池化层,得到图像的最终特征 x ,其维度为 1024×1 ,如式(1)所示,其中 f_{GMP} 为全局最大池化操作, I 为 224×224 像素的图像输入,

θ_{cnn} 为卷积神经网络参数, f_{cnn} 为图像卷积操作。

$$\mathbf{x} = f_{\text{GMP}}(f_{\text{cnn}}(I; \theta_{\text{cnn}})). \quad (1)$$

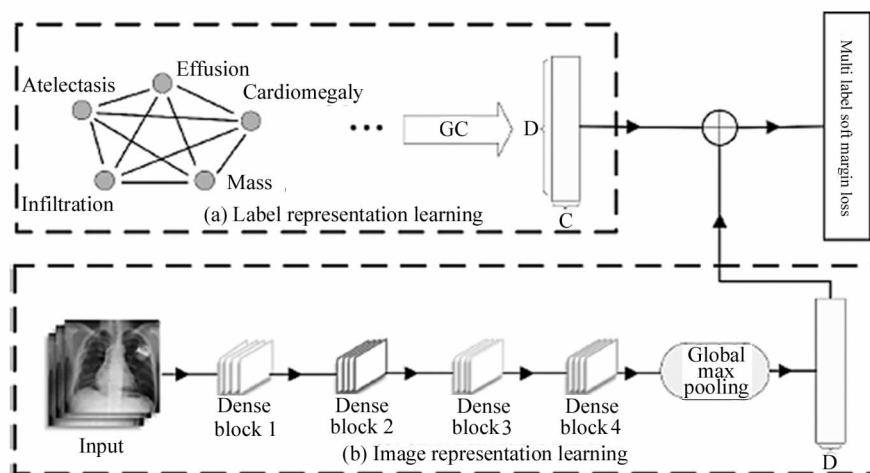


图1 胸部疾病诊断总体框架图

Fig. 1 The framework of the diagnosis of chest diseases

2.3 标签特征提取

如图1(a)所示,本文标签特征提取使用的是GCN^[17],GCN与传统的CNN不同,GCN是通过图中节点之间的信息传递来更新节点特征表示形式。本文中每种疾病标签表示为单独一个节点,GCN设置为2层,输入为特征表示矩阵 $\mathbf{H}^l \in \mathbf{R}^{n \times d}$ (其中 n 是标签的种类, d 是标签的词嵌入维度)和标签相关关系矩阵 $\mathbf{A} \in \mathbf{R}^{n \times n}$,目标是在图 g 上学习一个函数 $f(\cdot, \cdot)$,使新节点更新表示为 $\mathbf{H}^{l+1} \in \mathbf{R}^{n \times d}$,这样GCN的每一层都可以表示为一个非线性激活函数,即:

$$\mathbf{H}^{l+1} = f(\mathbf{H}^l, \mathbf{A}). \quad (2)$$

本文的图卷积运算使用的 $f(\cdot, \cdot)$ 表示形式如下所示:

$$\mathbf{H}^{l+1} = h(\mathbf{A}\mathbf{H}^l\mathbf{W}^l), \quad (3)$$

式中, $h(\cdot)$ 表示非线性操作,本文使用 LeakyReLU 激活函数, $\mathbf{W}^l \in \mathbf{R}^{d \times d}$ 表示要学习的转换矩阵,这样就可以通过GCN学习图 g 中各个节点之间的复杂关系。

本文中GCN的输入为标签特征表示矩阵 $\mathbf{H}$, 维度为 14×300 ,以及标签相关关系矩阵 $\mathbf{A}$, 维度为 14×14 , (在 $\mathbf{H}$ 中14为标签的种类,300为标签的词嵌入维度)。经过最后一层图卷积操作之后得到标签的最终特征 $\mathbf{y}$, 其维度为 14×1024 。

2.4 标签全局关系矩阵构建

GCN是通过节点之间的信息传递及更新节点特征来进行卷积的,因此如何构建标签相关关系矩阵 $\mathbf{A}$ 是能否有效利用GCN模型的关键。由于传统的医

学图像数据集并没有提供标签之间的关系矩阵,本文通过数据挖掘的方式来定义标签之间的关系矩阵。

首先统计出所有疾病类别的总数,然后通过数据挖掘找出每种疾病发生的情况下其他疾病发生的数量,即以条件概率的形式构建关系矩阵,本文定义 $P(L_a | L_b)$ 表示 L_a 标签出现的情况下 L_b 发生的概率。如图2所示,假设 L_a 为气胸(Pneumothorax), L_b 为肺气肿(Emphysema),气胸出现的情况下肺气肿出现的概率为0.3,而肺气肿出现的情况下气胸出现的概率为0.1,两种胸部疾病发生的条件概率不对等,因此关系矩阵 $\mathbf{A}$ 为非对称矩阵。本文所使用的医学影像数据集疾病标签类别为14种,所以最终构建的标签相关关系矩阵为 14×14 的二维矩阵。

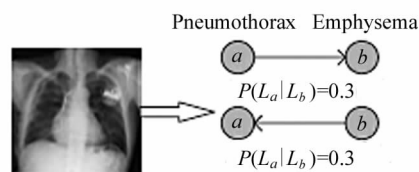


图2 不同疾病之间的关系对照图

Fig. 2 Comparison chart of the relationship between different diseases

2.5 特征融合

本文采取矩阵乘积的方式进行特征融合,具体操作如式(4)所示,其中 $\mathbf{x}$ 为图像特征, $\mathbf{y}$ 为标签特征:

$$\hat{\mathbf{y}} = \mathbf{y}\mathbf{x}. \quad (4)$$

得到总体特征 $\hat{\mathbf{y}}$, 然后放入损失函数求出损失

$loss$, 本文的损失函数采用多标签分类损失函数, 其中 $\sigma(\cdot)$ 为 Sigmoid 函数, 具体式如下:

$$loss = \sum_{c=1}^C y^c \log(\sigma(\hat{y}^c)) + (1 - y^c) \log(1 - \sigma(\hat{y}^c)). \quad (5)$$

2.6 实验数据

实验数据使用由 NIH 整理并公开的大型多标签胸部 X 射线数据集 ChestX-ray14, 该数据集是迄今为止可用的最大的 ChestX-ray 数据集。ChestX-ray 数据集包括 112 120 张 X 射线的正视图, 患者数量为 30 805 人。原始图像大小为 1024×1024 pixel, 所有图像均以 PNG 格式保存, 具有 8 位灰度值。每张图像都标有 14 种胸部疾病, 标签是从放射影像报告中通过自然语言处理(natural language processing, NLP)提取的。在 112 120 张胸部 X 射线图像中, 51 708 张图像标记为一种或多种疾病, 其余 60 412 张图像被认为是正常的。

3 结果与分析

3.1 数据预处理与实验参数设置

在 ChestX-ray14 数据集中同一病人的不同胸片样本也会有不同的标签, 故本文采取常见划分方式, 随机选取 70% 的数据集作为训练集, 20% 的数据集作为测试集, 10% 的数据集作为验证集, 且保证不同集合之间的胸片样本不交叉。为了减少计算量并且匹配网络模型, 将图像全部下采样至 256×256 pixels, 并且将其转换为 RGB 彩色 3 通道格式, 再裁剪至 224×224 大小进行随机水平翻转以实现数据增强。最后将处理过的胸片图像转化为向量格式并进

行归一化处理。

所有实验均在 pytorch 深度学习框架中实现, 均在带有 NVIDIA RTX2080 的 GPU 上运行, 使用 Adam 优化器进行前期优化, 学习率设定为 0.01, 在损失数值处于较小值波动时换成 SGD 优化器进行微调, 设定损失连续 5 轮没有减少就将学习率乘以 0.1, 待损失趋于稳定时停止训练。

3.2 性能评价指标

为了客观地对实验进行评价, 本实验使用 Area under the ROC curves (AUC) 作为评价指标, 该指标被广泛用于生物统计学, 以用来检测真实标签和错误标签的比例。使用该指标首先需要定义假正例率(false positive rate, FPR)和真正例率(true positive rate, TPR)。 FPR 表示所有负样本中分类器错判为正样本的概率, TPR 表示所有正样本中分类器能正确判为正例的概率, 具体计算式如下:

$$FPR = \frac{FP}{FP + TN}, \quad (6)$$

$$TPR = \frac{TP}{TP + FN}, \quad (7)$$

式中, FP 为假正例, TN 为真负例, TP 为真正例, FN 为假负例。ROC 曲线以所有样本的分类结果在不同阈值下的一系列假正例率和真正例率为横、纵坐标进行构建。AUC 为 ROC 曲线下方的面积, 如果面积越大, AUC 越接近于 1, 表明实验分类效果越好。

3.3 实验结果

本文对比了前人在 ChestX-ray14 上的实验结果, 以 AUC 为评价指标, 结果如表 1 所示, 本文方法

表 1 不同胸部多标签分类方法在 Chest-ray14 上的比较

Tab. 1 Comparison of different chest multi-label classification methods on Chest-ray14

Pathology	Method			
	Wang et al. (2017)	Yao et al. (2018)	Ma et al. (2019)	Ours
Atelectasis	0.716	0.772	0.7627	0.832
Cardiomegaly	0.807	0.904	0.8835	0.907
Effusion	0.784	0.859	0.8159	0.886
Infiltration	0.609	0.695	0.6786	0.714
Mass	0.706	0.792	0.8012	0.860
Nodule	0.671	0.717	0.7293	0.789
Pneumonia	0.633	0.713	0.7097	0.778
Pneumothorax	0.806	0.841	0.8377	0.879
Consolidation	0.708	0.788	0.7443	0.817
Edema	0.835	0.882	0.8414	0.894
Emphysema	0.815	0.829	0.8836	0.932
Fibrosis	0.769	0.767	0.8007	0.816
Pleural Thickening	0.708	0.765	0.7536	0.764
Hernia	0.767	0.914	0.8763	0.932
Average	0.738	0.803	0.7941	0.843

对 14 种疾病预测的平均 AUC 值均高于对比方法,平均 AUC 达到了 0.843,相较于 Ma 等人的最优方法提升了 4%。肺不张(atelectasis)、肺炎(pneumonia)和肿块(mass)较 Yao 等人提升了 6%,其中胸膜增厚(pleural thickening)AUC 值为 0.764,与前人最优模型基本保持一致,这是由于词嵌入语言模型没有对该词进行充分的表达所致。从总体来看,本文所用方法对 AUC 有很大的提升,能够很

好地识别疾病种类特征,总体分类性能良好。

3.4 消融实验

为了对比评估本模型中 GCN 模块的有效性,本文在 ChestX-ray14 数据集上进行了消融实验,完整移除了 GCN 模块之后测得每组的 AUC 值和平均 AUC 值,实验结果如表 2 所示。

由表 2 可以看出,加入 GCN 模块之后平均 AUC 值提升了 3.1 个百分点,证明该模块确实对总

表 2 消融 GCN 模块对比表

Tab. 2 Comparison table of ablation GCN modules

Pathology	Remove GCN	Ours
Atelectasis	0.800	0.832
Cardiomegaly	0.875	0.907
Effusion	0.870	0.886
Infiltration	0.697	0.714
Mass	0.798	0.860
Nodule	0.744	0.789
Pneumonia	0.741	0.778
Pneumothorax	0.859	0.879
Consolidation	0.799	0.817
Edema	0.885	0.894
Emphysema	0.887	0.932
Fibrosis	0.810	0.816
Pleural Thickening	0.760	0.764
Hernia	0.847	0.932
Average	0.812	0.843

体模型性能有提升效果。

3.5 可视化分析

为了更好地对模型的疾病预测方法进行解释,对模型定量分析之后,本文还使用类激活映射(class activation mapping, CAM)生成了热图,用于可视化

分析模型在判别疾病类别时所依据的病灶区域。如图 3 所示,图 3(b)中疾病标签为心脏扩大,图 3(d)中疾病标签为积液、心脏扩大、肿块,这些热图都能很好地符合病灶区域,验证了模型在特征选取方面是准确有效的。

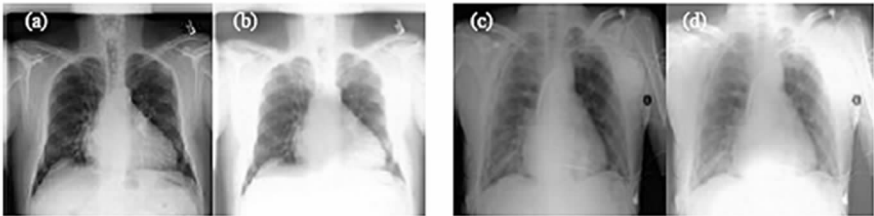


图 3 不同胸片的病灶区域热图:(a)(b) 疾病标签为心脏扩大;(c)(d) 疾病标签为积液、心脏扩大、肿块
Fig. 3 Heatmaps of lesion area on different chest radiographs:(a)(b) Disease is labeled as cardiomegaly;
(c)(d) Disease is labeled as effusion,cardiomegaly and mass

4 结 论

随着深度学习的发展,越来越多的网络模型被应用到医学领域。针对目前的胸部医学图像标签类别不均衡、识别精度不够等问题,本文提出了一种将标签特征与图像特征融合的医学多标签胸部疾病分类方法,解决了非欧式数据标签的使用问题,通过消融实验和可视化分析进一步表明了本文模型的有效性。在 ChestX-ray14 数据集上与目前较为经典的方法比较,本方法均能够明显提升分类效果。

由于本实验所采用的数据集不能提供患者的病史资料,无法融入历史疾病信息,所以会降低诊断性能。如何有效利用患者病史信息,有待进一步研究。在后续的工作中将考虑把同一患者不同视图的胸片加入训练,更加全面的提高模型对患者的诊断能力。

参考文献:

- [1] BAEMANI M J, MONADJEMI A, MOALLEM P. Detection of respiratory abnormalities using artificial neural networks [J]. *Journal of Computer Science*, 2008, 4(8): 663.
- [2] RAOOF S, FEIGIN D, SUNG A, et al. Interpretation of plain chest roentgenogram [J]. *Chest*, 2012, 141(2): 545-558.
- [3] TUDOR G R, FINLAY D, TAUB N. An assessment of inter-observer agreement and accuracy when reporting plain radiographs [J]. *Clinical Radiology*, 1997, 52(3): 235-238.
- [4] QUEKEL L G B A, KESSELS A G H, GOEI R, et al. Miss rate of lung cancer on the chest radiograph in clinical practice [J]. *Chest*, 1999, 115(3): 720-724.
- [5] CICERO M, BILBILY A, COLAK E, et al. Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs [J]. *Investigative Radiology*, 2017, 52(5): 281-287.
- [6] LEE M Z, CAI W, SONG Y, et al. Fully automated scoring of chest radiographs in cystic fibrosis [C]//2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 3-7, 2013, Osaka, Japan. New York: IEEE, 2013: 3965-3968.
- [7] WANG X, PENG Y, LU L, et al. Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2097-2106.
- [8] GUENDEL S, GRBIC S, GEORGESCU B, et al. Learning to recognize abnormalities in chest x-rays with location-aware dense networks [EB/OL]. (2018-03-12) [2021-09-14]. <https://arxiv.org/abs/1803.04565>.
- [9] KUMAR P, GREWAL M, SRIVASTAVA M M. Boosted cascaded convnets for multilabel classification of thoracic diseases in chest radiographs [EB/OL]. (2017-11-23) [2021-09-14]. <https://arxiv.org/abs/1711.08760>.
- [10] LIU J, ZHAO G, FEI Y, et al. Align, attend and locate: chest x-ray diagnosis via contrast induced attention network with limited supervision [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October, 27-November 02, 2019, Seoul, Korea (South). New York: IEEE, 2019: 10632-10641.
- [11] RAJPURKAR P, IRVIN J, ZHU K, et al. Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning [EB/OL]. (2017-11-14) [2021-09-14]. <https://arxiv.org/abs/1711.05225>.
- [12] DENG J, DONG W, SOCHER R, et al. Imagenet: a large-scale hierarchical image database [C]//2009 IEEE Conference on Computer Vision and Pattern Recognition, June 20-25, 2009, Miami, FL, USA. New York: IEEE, 2009: 248-255.
- [13] YAO L, POBLENZ E, DAGUNTS D, et al. Learning to diagnose from scratch by exploiting dependencies among labels [EB/OL]. (2017-10-28) [2021-09-14]. <https://arxiv.org/abs/1710.10501v2>.
- [14] MA Y, ZHOU Q, CHEN X, et al. Multi-attention network for thoracic disease classification and localization [C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 12-17, 2019, Brighton, UK, New York: IEEE, 2019: 1378-1382.
- [15] PENNINGTON J, SOCHER R, MANNING C D. Glove: global vectors for word representation [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), October 25-29, 2014, Doha, Qatar, Stroudsburg PA, USA: ALC, 2014.
- [16] HUANG G, LIU Z, VAN D M L, et al. Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 4700-4708.
- [17] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks [EB/OL]. (2016-09-09) [2021-09-14]. <https://arxiv.org/abs/1609.02907>.

作者简介:

黄青松 (1962—),男,硕士,教授,硕士生导师,主要从事智能信息系统方面的研究。