

# 基于复杂网络描述的图像深度卷积分类

洪睿 康晓东\* 郭军 李博 王亚鸽 张秀芳

(天津医科大学医学影像学院, 天津 300203)

(\*通信作者电子邮箱 [hr8229862@163.com](mailto:hr8229862@163.com))

**摘要:** 为了提高在不增加较多计算量的前提下, 提高卷积网络模型用于图像分类的正确率, 提出了一种基于复杂网络模型描述的图像深度卷积分类方法。首先对图像进行复杂网络描述, 得到不同阈值下的复杂网络模型度矩阵; 然后, 在图像度矩阵描述的基础上, 通过深度卷积网络得到特征向量; 最后根据得到的特征向量进行 k 近邻分类。本文在 ILSVRC2014 数据库上进行了验证, 验证结果表明, 本文提出的模型具有较好的正确率和较少的迭代次数。

**关键词:** 复杂网络; 深度卷积神经网络; AlexNet; 度矩阵; 图像分类

**中图分类号:** TN911.73

**文献标志码:** A

## Image deep convolution classification based on complex network description method

Hong Rui Kang Xiaodong\* Guo Jun Li Bo Wang Yage Zhang Xiufang

(Tianjin Medical University School of Medical Image, Tianjin Medical University, Tianjin 300203 China)

**Abstract:** In this article, an image deep convolution classification based on complex network description method was proposed to improve the accuracy of image classification with convolution network model without increasing more computation. First, using image description, to obtain the complex network model degree matrices under different thresholds. Then, obtain the feature vector from degree matrices by deep convolution neural networks based on degree matrices image description method. Finally, the feature vectors are used for image K-Nearest Neighbor classification. The result is verified on ILSVRC2014 database. The results show that the model proposed in this paper has better accuracy and fewer iterations.

**Keywords:** complex network; deep convolution neural networks; AlexNet; degree matrix; image classification

### 0 引言

纹理是非常重要的图像识别特征之一, 对图像纹理进行合理描述是图像分类识别的基础, 图像纹理包含了非常丰富的图像信息, 可以应用于多种任务, 如基于内容的图像检索, 图像分割, 图像融合, 图像分类或者聚类等。图像纹理特征描述的方法主要有基于几何学的特征描述, 基于模型的特征描述, 基于信号处理的特征描述, 以及基于统计的特征描述。基于复杂网络模型的纹理特征描述方法属于基于统计的特征描述的一种, 具有稳定性好, 抗噪声能力强等优点<sup>[1]</sup>。

在计算机视觉领域, 图像分类一直是研究的重点之一。传统基于特征的图像分类方法都是先定义一种特征, 再根据预先定义的特征进行分类识别, 存在着泛化能力弱, 可移植性差, 准确度相对较差的缺点。深度学习作为机器学习的重要发展, 能够逐层地自动学习合适的表示特征, 因而在众多领域中得到了广泛应用, 如人脸特征点检测<sup>[2]</sup>、行人检测<sup>[3]</sup>、人脸识别<sup>[4]</sup>和车牌识别<sup>[5]</sup>等。卷积神经网络作为深度学习的三大重要模型之一, 继承了深度学习技术自动提取特征的优点, 并且通过权值共享大大减少了所需要训练的参数, 使 CNN 能快速处理高维图像, 还具备一定的平移不变性, 因此在图像分类领域取得了较好的效果。其在手写数字识别以及

人脸识别方面都被证明有着非常出色的表现。

20 世纪 80 年代, Fukushima 提出了第一个多层次的神经网络模型 Neocognitron<sup>[6]</sup>。20 世纪 90 年代, Lecun 等构建了 LeNet-5 模型<sup>[7]</sup>, 该模型首次提出了使用梯度反向传播算法(Back Propagation algorithm, BP), 对卷积神经网络进行有监督训练; 当时在 MNIST 数据集上取得最好的识别效果, 其模型采用交替连接的卷积层和池化层, 最后通过全连接层分类, 确立了现代卷积神经网络的基础, 后来的大多数卷积神经网络模型都是以 LeNet-5 模型为基础的。Alex Krizhevsky 等人提出的 AlexNet 模型<sup>[8]</sup>在 2012 年的 ImageNet 图像分类竞赛中大放异彩, 错误率比上一年的冠军下降了十个百分点, 且远远超过当年的第二名; 也因此将深度卷积网络 DCNN (Deep Convolutional Neural Network) 启动了图像分类研究的广泛应用, 对后来的 VGG<sup>[9]</sup>、ResNet<sup>[10]</sup>、Google Net<sup>[11]</sup>等模型都具有非常深刻的影响。

本文提出了一种基于复杂网络描述的图像深度卷积分类方法。首先, 以复杂网络的图像描述方法, 将图像转化为基于三个不同阈值产生的三组度矩阵; 然后由三个深度卷积神经网络分别学习三组度矩阵的特征; 最后将得到的特征向量进行特征融合。

### 1 深度卷积网络

收稿日期: 2018-05-21; 修回日期: 2018-07-06; 录用日期: 2018-07-11。

基金项目: 天津市重点基金(17JC20J32500)

**作者简介:** 洪睿(1993-), 男, 内蒙古通辽人, 硕士研究生, 主要研究方向: 图像处理; 康晓东(1964-), 天津人, 教授, 博士, CCF 高级会员, 主要研究方向: 医学图像处理、医疗信息系统集成; 郭军(1972-), 男, 四川成都人, 实验师, 本科, 主要研究方向: 实验技术; 李博, 硕士研究生; 王亚鸽, 硕士研究生; 张秀芳, 本科生。

图像分类是通过对图像整体的分析,预测图像类别。图像分类是计算机视觉的挑战领域之一,每年有赛事 Imagenet Large Scale Visual Recognition Challenge (ILSVRC)<sup>[12]</sup>。在 ILSVRC2012 中,来自多伦多大学的 Alex Krizhevsky 团队设计出卷积神经网络 AlexNet,将测试的 top-5 错误率下降到 15.3%,而其他团队做出的最少的 top-5 错误率仍为 26.2%。此后,不断有更加高效、准确的模型出现,如牛津大学的 VGG<sup>[9]</sup>,谷歌研究院的 GoogLeNet<sup>[10]</sup>等。近年来的优秀模型甚至突破了人眼识别的平均错误率,展示了卷积神经网络在图像分类上的强大优势。卷积神经网络需要用带有标签的数据集做有监督的网络训练,目前有许多用于图像分类的带标签数据集,如 ImageNet<sup>[13]</sup>、CIFAR10/100<sup>[14]</sup>、NORB<sup>[15]</sup>、Caltech-101/256<sup>[16]</sup>等。ILSVRC 使用的为 ImageNet 数据集的子集,该子集包含 1000 个分类,128 万张测试图片。

AlexNet 最终的网络结构共有 7 层,包含 5 个卷积层和 2 个全连接层,有约 65 万个神经元和 6000 万个可训练的参数。采用 dropout 和激活函数 ReLu 等技术,训练环境为 2 路 GTX580,对 128 万测试图片的训练周期在 5~6 天,相比于之前的卷积神经网络,其在准确度和训练时间上都有了明显提升。

## 2 复杂网络模型

复杂网络存在大量的静态统计特征描述量,如度、网络直径、平均路径长度、聚类系数、最大度和最大核数等。这些特征量可以用来度量复杂网络的重要属性,如网络的表示、分类和建模等<sup>[17]</sup>。

动态演化是复杂网络的一个重要特征。在动态演化过程中,复杂网络的特征度量是一个时间的函数,在同一种演化方式下,不同时刻得到的两个子网络具有不同的特征。因此,用复杂网络动态演化过程中不同时刻的静态统计特征量来进行网络的分析与分类是十分重要且有效的一种方式<sup>[18]</sup>。

目前,基于复杂网络理论的图像描述方法受到了越来越多研究者的关注<sup>[19]</sup>。Silva 等人<sup>[20]</sup>通过对人脸图像建立复杂网络模型,实现了人脸图像的特征提取与识别。Backes 等人<sup>[21]</sup>利用复杂网络对图像边缘进行建模,实现了图像形状特征提取。汤进等人<sup>[22]</sup>在 Harris 特征点的基础上建立复杂网络模型,实现了图像的形状特征提取。Backes 等人<sup>[23]</sup>通过建立图像的复杂网络平均度矩阵,在平均度矩阵的基础上运行部分自回避确定性游走算法,实现了图像的纹理特征提取。由于基于复杂网络的图像描述是基于复杂网络静态统计特征量的特征描述,相比于传统的图像描述方法,该类方法具有稳定性好,抗噪声能力强等优点。

## 3 基于复杂网络的深度卷积神经网络图像分类

### 3.1 复杂网络度矩阵

本文根据复杂网络的静态统计量建立图像在不同阈值下的度矩阵,通过统计网络节点在每个状态下的度分布完成图像的描述。

文献[17]中提出了一种建立图像复杂网络模型的方法,将图像的每个像素看作是复杂网络的一个节点,并认为每个节点之间都有边相连,边的权值由两个像素之间的距离和灰度差的加权和决定。通过设定一系列边权值的阈值对初始的复杂网络完全图模型进行阈值动态演化,权值高于阈值的边被删去,得到的就是距离较小,且像素值相近的像素之间的边。为了简化复杂网络模型,本文选取节点周围与其距离小于 3 的 28 个节点作为邻域,只有在邻域内的节点可以有边相连。文献[17]中两个节点  $i(x,y)$  与  $j(x',y')$  之间边的权值  $w(v(x,y),v(x',y'))$  为节点之间距离与节点代表的像素的灰度值差的加权和,为了使节点的度分布更加均匀,本文采用的权值  $w$  为以上两项的直接加和,如式(1)所示:

$$w(v(x,y),v(x',y')) = \sqrt{(x-x')^2 + (y-y')^2} + |I(x,y) - I(x',y')| \quad (1)$$

度矩阵元素分布如图(1)所示,其中左侧直方图为文献[17]中算法得到度矩阵的元素分布,右侧直方图为本文算法得到度矩阵的元素分布。

在将得到的边权值归一化之后,设定一系列阈值  $t$ ,节点之间边权值高于阈值的边将被删去,得到每个节点与阈值对应的邻域  $\theta(v_i)$  和度  $\deg(v_i)$ 。如式(2)(3)所示。

$$q(vt) = \{v' \in V \mid (v,v') \in E \text{ 且 } w(v,v') \leq t\} \quad (2)$$

$$\deg(v) = |q(v)| \quad (3)$$

度矩阵可以作为衡量像素与邻域的相似度的标准,同时也包含大量的纹理信息。如图(2)所示,图(2)中 a 为原始图像,图 b、c 和 d 分别为阈值  $t=0.8$ 、0.7 和 0.6 时生成的度矩阵图像。

### 3.2 深度卷积网络

DCNN 一般由卷积层、池化层、全连接层三种神经网络层以及一个输出层(Softmax 等分类器)组成,每层由多个二维平面块组成,每个平面块由多个独立神经元组成,图 3 为 AlexNet 模型的结构示意图。

卷积层通过局部感受域与上一层神经元实现部分连接,在同一局部感受域内的神经元与图像区域中对应像素有固定二维平面编码信息关联,迫使神经元提取局部特征,在每层的各个位置分布着许多组不同的神经元,每组神经元有一组输入权值,这些权值与前一层神经网络矩形块中的神经元关联,即共享权值,减少了权值数量,降低了网络模型的复杂度。卷积层在 DCNN 中起着至关重要的特征提取的功能,通过局部感受域方法获取的观测特征与平移、缩放和旋转无关,其权值共享结构减少了权值数量,进一步降低了网络模型的复杂度。

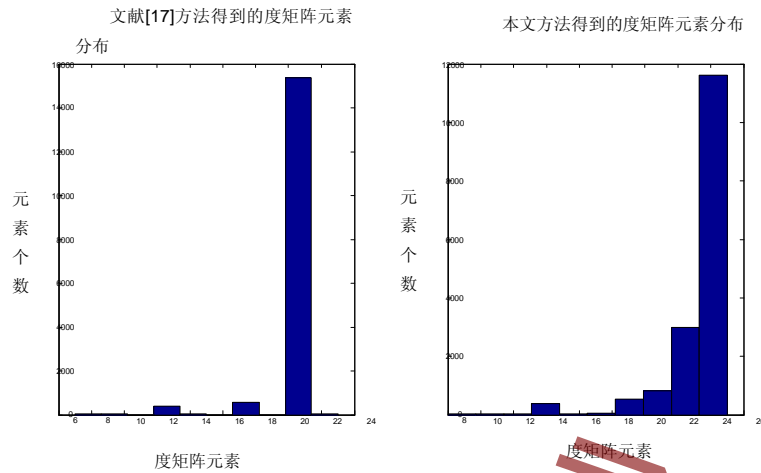


图1 文献[17]与本文应用算法得到的同一图像同一阈值的度矩阵元素分布直方图

Figure 1 The histograms of the degree matrix element distribution of the same image with the same threshold acquired by the algorithm we represent and the algorithm in paper[17]

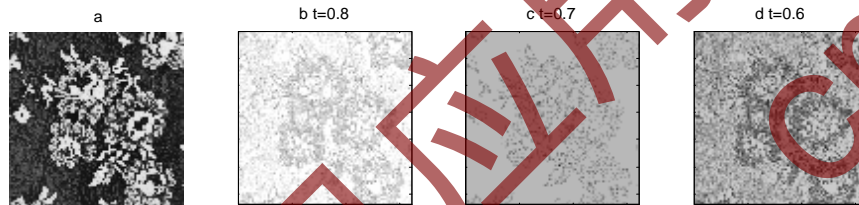


图2 原始图像及其度矩阵图像

Figure 2 The original image and its degree matrix image

池化层是特征映射层,选择卷积特征图中不同的连续范围的作为池化区域,然后取特征的最大值或平均值作为池化区域的特征,从而减少特征向量维度,实现局部平均和抽样,使特征映射输出对平移、旋转、比例缩放等形式变换的敏感度下降。池化层通常跟在卷积层之后,这样便构成了一个两次特征提取的结构,从而在对输入样本识别时,网络有很好的畸变容忍能力。

全连接层是本层神经元与上层神经元两两连接但本层神经元之间不连接的结构,相当于多层感知器(MLP)中的隐含层,局部特征信息作为输出层(Softmax 等分类器)的输入,其后不再接卷积层,因通过全连接层之后,图像特征已由二维信息降为一维信息,已无法进行二维卷积运算。

图4为本文提出的基于复杂网络描述的图像深度卷积分类方法结构图。如图所示,对于一张彩色图像,其RGB三个维度上的分量均可视为一张图像,由每个分量图像可得到三个复杂网络模型度矩阵,即为三组(九个)度矩阵;将每个分量得到的度矩阵作为 AlexNet 的输入,学习得到三组图像的特征;再将特征融合,由分类器得到分类结果并输出。

#### 4 实验结果与分析

为了验证本文算法的有效性,本文对 ILSVRC2014 数据库的样本进行了分类仿真实验。实验环境为联想 80SH, 4GB 内存, Windows10 系统。选取 10 类图像中的 100 张,共 1000 张图像作为分类的训练集合;选取这 10 类图像另外 200 张作为样本集合;进行 K 近邻分类,分类原则为  $K=1$ ,分类判别距离为欧氏距离。复杂网络描述中,阈值选取在 0.5 到 0.9

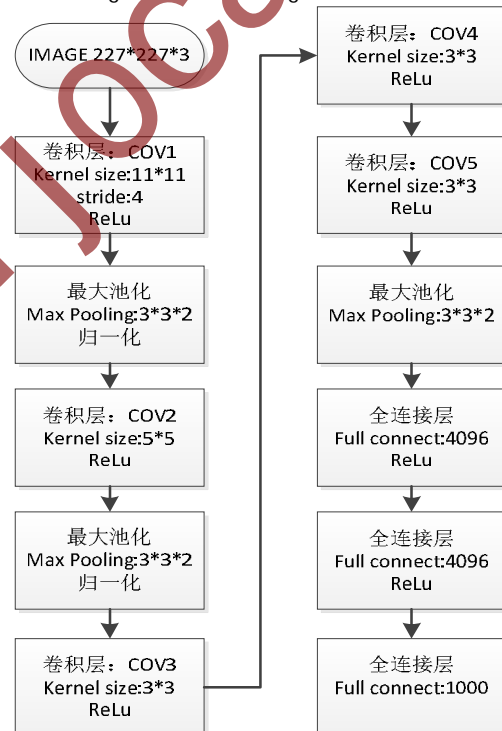


图3 AlexNet 结构示意图

Figure3 The structure schematic diagram of AlexNet model

之间,实验证明阈值在 0.5 到 0.9 之间,能保留更多的信息,保证分类实验较高的正确率;当阈值小于 0.5 时,随着阈值下降,产生的度矩阵能保留的原图信息非常少。共设置 3 个



阈值。由三组 AlexNet 生成 3000 维的特征向量。并对比模型 AlexNet、VGGNet 和在 VGG 模型的基础上使用复杂网络描述的神经网络模型,最大迭代次数设为 500 次,每 50 次迭代进行一次验证并输出识别率。如图 5 所示,当迭代次数为 50 时,AlexNet 的识别率时 24%,最终收敛于 85%;PM(Proposed Method)即为本文方法,迭代次数为 50 时识别率为 50%,最

终收敛域 90%;VGG 模型在迭代次数为 50 时,识别率为 73%,最终收敛于 94%;而在 VGG 模型上使用复杂网络描述的神经网络模型最终的数据结果与 VGG 模型的结果几乎相同。

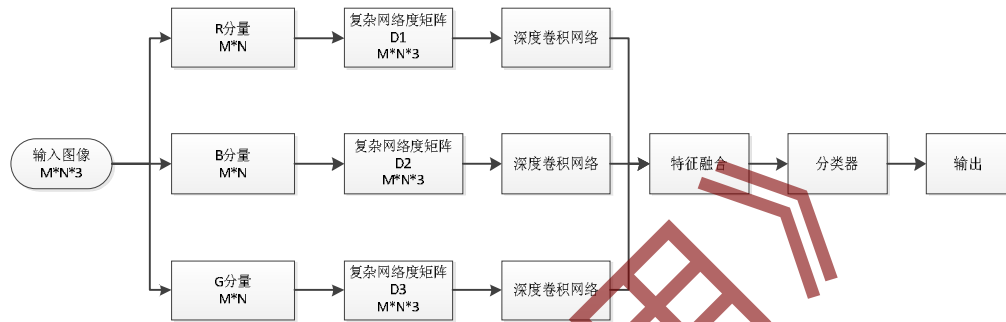


图 4 本文算法流程图

Figure 4 The flow chart of proposed method

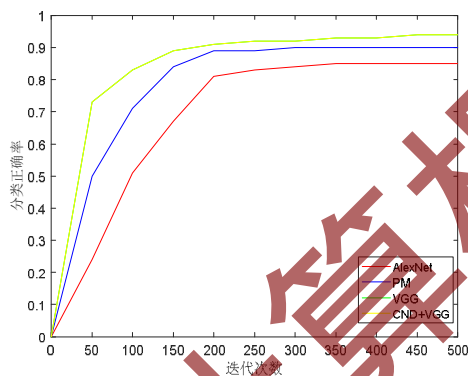


图 5 3 种网络模型随着迭代次数变化的识别率

Fig5 The recognition rate of 3 network models along with the number of iterations

实验表明,本文提出的方法与 AlexNet 相比识别率较高,且较早收敛;与 VGG 相比较差,仍有改进空间;由于 AlexNet 与 VGG 相比,网络结构较为简单,增加一个复杂网络描述层之后,分类正确率有明显的改进,而 VGG 由于其网络层数较多,增加复杂网络描述层之后,结果并没有明显变化。

表 1 复杂网络描述方法及卷积网络模型第一层卷积核大小的影响

Tab 1 The influence from the complex network description method and the kernel size of convolution network model first layer

阈值	第一层卷积核大小	正确率/%	收敛时迭代次数
高阈值	11×11	90	160
	10×10	90	160
	9×9	83	140
	5×5	81	220
低阈值	11×11	58	170
	10×10	67	140
	9×9	64	140
	5×5	60	150

实验为了验证复杂网络描述方法中阈值大小选取以及 AlexNet 网络模型的第一层卷积核大小,对于精准度和收敛时迭代次数的影响,改变第一个卷积核的大小,分别记录了最高正确率和收敛时迭代次数,结果如表 1 所示。其中高阈值为 0.8、0.7、0.6,低阈值为 0.4、0.3、0.2。

实验表明选择复杂网络描述方法的阈值选择和深度卷积网络第一层卷积核大小对正确率和收敛时迭代次数的影响较大。选取较高的阈值能得到较高的正确率,但是收敛时迭代次数较高;选取合适 10×10 或者 9×9 的首层卷积核大小能得到较高的正确率和较少的收敛时迭代次数。

## 5 结语

提出了一种基于复杂网络描述的图像深度卷积分类方法,以复杂网络的描述方法构建度矩阵,然后采用 AlexNet 进行特征学习,完成图像分类的工作。与主流分类模型 VGGNet、AlexNet 的对比试验,本文提出的方法具有较高的图像识别率,迭代次数收敛较早,具有较强的稳定性和一定的提升空间。

## 参考文献

- [1] 康晓东. 医学影像图像处理[M]. 人民卫生出版社, 2009.( Kang X D. Image processing of medical imageology[M]. People's Medical Publishing House, 2009.)
- [2] Sun Y, Wang X, Tang X. Deep Convolutional Network Cascade for Facial Point Detection[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2013:3476-3483.
- [3] 徐超, 闫胜业. 改进的卷积神经网络行人检测方法[J]. 计算机应用, 2017, 37(6):1708-1715.(Xu C, Yan S Y. Improved pedestrian detection

- method based on convolutional neural network[J]. Computer application, 2017, 37(6):1708-1715.)
- [4] Moon H M, Chang H S, Pan S B. A face recognition system based on convolution neural network using multiple distance face[J]. Soft Computing, 2017, 21(17):4995-5002.
- [5] 康晓东, 王昊, 郭军,等. 无监督深度学习彩色图像识别方法[J]. 计算机应用, 2015, 35(9):2636-2639.( Kang X D,Wang H,Guo J,et al.Unsupervised deep learning method for color image recognition[J].Computer application,2015, 35(9):2636-2639.)
- [6] Fukushima K, Miyake S, Ito T.Neocognitron: A neural network model for a mechanism of visual pattern recognition[J].IEEE Transactions on Systems, Man, and Cybernetics, 1983 (5) : 826-834.
- [7] LeCun Y, Boser B, Denker J S, et al.Backpropagation applied to handwritten zip code recognition[J].Neural Computation, 1989, 1 (4) : 541-551.
- [8] 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.
- [9] Simonyan K, Zisserman A.Very deep convolutional networks for large-scale image recognition[C]//Proceedings of ICLR 2015, 2015.
- [10] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015:770-778.
- [11] Szegedy C, Liu W, Jia Y, et al.Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
- [12] Berg A, Deng J, Fei-Fei L.Large scale visual recognition challenge 2010[R/OL].(2010)[2017-02-28].http: //www.image-net.org/challenges.
- [13] Deng J, Dong W, Socher R, et al.Imagenet: A large-scale hierarchical image database[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR 2009), 2009: 248-255.
- [14] Krizhevsky A.Learning multiple layers of features from tiny images[D].University of Toronto, 2009.
- [15] LeCun Y, Huang F J, Bottou L.Learning methods for generic object recognition with invariance to pose and lighting[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004), 2004.
- [16] Fei-Fei L, Fergus R, Perona P.Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories[J]. Computer Vision and Image Understanding, 2007, 106 (1) : 59-70.
- [17] Backes A R, Casanova D, Bruno O M. Texture analysis and classification: A complex network-based approach[J]. Information Sciences, 2013, 219(1):168-180.
- [18] Backes A R, Casanova D, Bruno O M. A complex network-based approach for boundary shape analysis[J]. Pattern Recognition, 2009, 42(1):54-67.
- [19]Strogatz S H. Strogatz, S.H.: Exploring Complex Networks. Nature 410, 268[J]. Nature, 2001, 410(6825):268-276.
- [20] Silva J D A, Bruno O M. A rotation invariant face recognition method based on complex network[C]// Iberoamerican Congress Conference on Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. Springer-Verlag, 2010:426-433.
- [21] Backes A R, Casanova D, Bruno O M. A complex network-based approach for boundary shape analysis[J]. Pattern Recognition, 2009, 42(1):54-67.
- [22] 汤进, 陈影, 江波,等. 基于复杂网络的图像建模与特征提取方法[J]. 计算机工程, 2013, 39(5):243-247.(Tang J,Chen Y,Jiang B,et al. Image modeling and feature extraction method based on complex network[J]. Computer Engineering, 2013, 39(5):243-247.)
- [23] Couto L N, Backes A R, Barcelos C A Z. Texture characterization via deterministic walks' direction histogram applied to a complex network-based image transformation[J]. Pattern Recognition Letters, 2017, 97:77-83.

天津市重点基金(17JC20J32500)

洪睿(1993-), 男, 内蒙古通辽人, 硕士研究生, 主要研究方向: 图像处理; 康晓东(1964-), 天津人, 教授, 博士, CCF 高级会员, 主要研究方向: 医学图像处理、医疗信息系统集成; 郭军(1972-), 男, 四川成都人, 实验师, 本科, 主要研究方向: 实验技术; 李博, 硕士研究生; 王亚鸽, 硕士研究生; 张秀芳, 本科生。

**This work is partially supported by the Key Foundation of Tianjin,China(17JC20J32500).**

**HONG Rui**, born in 1993, M. S. candidate. His research interests is image processing.

**KANG Xiaodong**, born in 1964, Ph. D. professor. His research interests include medical image processing and medical information system integration.

**GUO Jun**, born in 1972, B.S. His research interests is laboratory technology