

DOI: 10.11992/tis.201806015

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180716.1333.014.html>

## 基于深度学习的椎间孔狭窄自动多分级研究

洪雁飞<sup>1,2</sup>, 魏本征<sup>1,2</sup>, 刘川<sup>2</sup>, 韩忠义<sup>1,2</sup>, 李天阳<sup>1,2</sup>

(1. 山东中医药大学 理工学院, 山东 济南 250355; 2. 山东中医药大学 计算医学实验室, 山东 济南 250355)

**摘要:** 椎间孔狭窄症的术前定性分级诊断对临床医生治疗策略的制定和患者健康恢复至关重要, 但目前该方面临床上仍然存在很多问题, 并且缺乏相关的研究和行之有效的方法用于辅助临床医生诊断。因此, 为提高计算机辅助椎间孔狭窄症诊断准确率以及医生工作效率, 本文提出一种基于深度学习的椎间孔狭窄图像自动分级算法。从人体矢状切脊柱磁共振图像中提取脊柱椎间孔图像, 并做图像预处理; 设计一种监督式深度卷积神经网络模型, 用于实现脊柱椎间孔图像数据集的自动多分级; 利用迁移学习方法, 解决深度学习算法在小样本数据集上的过拟合问题。实验结果表明, 本文算法在脊柱椎间孔图像数据集上的分类精确度可达到87.5%以上, 且其具有良好的鲁棒性和泛化能力。

**关键词:** 椎间孔狭窄; 自动分级; 机器学习; 深度学习; 特征提取; 监督训练; 迁移学习; 过拟合

**中图分类号:** TP311 **文献标志码:** A **文章编号:** 1673-4785(2019)04-0708-08

中文引用格式: 洪雁飞, 魏本征, 刘川, 等. 基于深度学习的椎间孔狭窄自动多分级研究[J]. 智能系统学报, 2019, 14(4): 708-715.

英文引用格式: HONG Yanfei, WEI Benzhen, LIU Chuan, et al. Deep learning based automatic multi-classification algorithm for intervertebral foraminal stenosis[J]. CAAI transactions on intelligent systems, 2019, 14(4): 708-715.

## Deep learning based automatic multi-classification algorithm for intervertebral foraminal stenosis

HONG Yanfei<sup>1,2</sup>, WEI Benzhen<sup>1,2</sup>, LIU Chuan<sup>2</sup>, HAN Zhongyi<sup>1,2</sup>, LI Tianyang<sup>1,2</sup>

(1. College of Science and Technology, Shandong University of Traditional Chinese Medicine, Ji'nan 250355, China; 2. Computational Medicine Lab, Shandong University of Traditional Chinese Medicine, Ji'nan 250355, China)

**Abstract:** Preoperative qualitative diagnosis of intervertebral foraminal stenosis is essential for the formulation of clinician treatment strategies and patients' health recovery. However, there are still many clinical challenges in this aspect and a lack of relevant research and proven methods to assist clinicians in diagnosis. Therefore, a deep learning-based automatic classification algorithm is proposed in this study to improve the diagnosis accuracy and the efficiency. First, we extracted the spinal foramen images from the sagittal spine MRI image, and then these images were preprocessed. Second, a supervised deep convolutional neural network model was designed to achieve automatic multi-classification for the datasets of the intervertebral foraminal stenosis. Finally, we used the transfer learning to optimize the overfitting problem of the deep learning algorithm in the small sample dataset. The experimental results show that the classification accuracy of this algorithm on the dataset of spinal foramen was 87.5%, and it has good robustness and generalization performance.

**Keywords:** intervertebral foraminal stenosis; automatic grade; machine learning; deep learning; feature extraction; supervised training; transfer learning; over fitting

收稿日期: 2018-06-05. 网络出版日期: 2018-07-17.

基金项目: 国家自然科学基金项目(U1201258, 61572300); 山东省自然科学基金项目(ZR2015FM010); 山东高等学校科技计划项目(J15LN20); 山东省医药卫生科技发展计划项目(2016WS0577); 山东省中医药科技发展计划项目(2017-001).

通信作者: 魏本征. E-mail: [wzb99@sina.com](mailto:wzb99@sina.com).

椎间孔狭窄症(intervertebral foraminal stenosis, IFS)是指由于椎间孔狭窄压迫脊髓神经根周围神经系统引发的退行性疾病<sup>[1]</sup>。脊髓神经根作为感觉输入的传递中枢,受压迫后将会引起疼痛、肌肉无力,甚至身体残疾,并且IFS会随着人的自然

衰老过程导致严重的脊柱侧凸和脊椎滑脱,大约80%的老年人深受IFS引起的腰背疼痛折磨<sup>[2]</sup>。临床数据表明,治疗计划随着IFS等级的不同而变化,如对于1级狭窄,通常的物理治疗和运动将是治疗的第一选择,而对于更严重的等级,患者可能需要接受手术治疗或者减压治疗等<sup>[3]</sup>。因此,有效和精确的分级是椎间孔狭窄症诊断和治疗至关重要的一步。但是现有的临床椎间孔狭窄分级主要存在3方面的问题:1)医生对IFS图像的目视检查和人工评分,耗时耗力且低效;2)由于IFS的病症多样性,临床医生常更多关注椎间盘突出、椎管狭窄和侧隐窝狭窄的情况,而容易忽略椎间孔的狭窄情况,主观性强;3)特殊的椎间孔解剖结构及复杂的致狭窄因素,易导致临床难以全面准确判定,误诊、漏诊率高。因此,为提高IFS诊断效率、减少医生主观影响以及降低误诊漏诊率,设计自动化的临床辅助诊断系统具有重要的应用价值<sup>[4]</sup>。

当前,计算机辅助诊断IFS方面的研究已取得了重要进展,按照感兴趣区域的不同可分为两种类型。第一种类型是基于椎间孔周围结构的自动定位分割研究,主要有Alomari等<sup>[5]</sup>使用一种两级模型在核磁共振图像中捕获像素级和对象级特征,以实现椎间盘的局部异常检测;Zhan等<sup>[6]</sup>采用基于CT和MR图像的共同的脊柱几何结构设计分析算法初始化椎骨的标记实现椎体定位;Wang等<sup>[7]</sup>采用回归分割的方法分割出M3脊柱图。第二种类型是直接基于椎间孔的自动分割或分级研究,主要有Ghosh等<sup>[8]</sup>通过对椎体中心线以及椎间孔主要边界点的检测实现分割;Huang等<sup>[9]</sup>采用Adaboost检测算法和迭代归一化分割算法对MRI椎体图像进行检测同时成定位和分割;Klinder等<sup>[10]</sup>采用将全局脊柱形状在连续的局部椎体坐标系中表示,并将个体椎骨建模为三角曲面网络的方法,实现了基于双尺度的脊柱和椎间孔建模与分割;He等<sup>[11]</sup>采用超像素分割法在对椎间孔狭窄图像定位和良恶性分类问题上精确度达到98.52%。

以上研究均取得不同程度的进展,但仍存在一些问题。此类基于部分形态学解剖结构的定位、分割或者比较简单的二级分类的辅助诊断研究,忽略了全局信息,影响诊断精确度。而临床上基于常规仰卧位MRI图像的椎间孔狭窄诊断是由椎间孔形态、神经根周围脂肪或神经根形态改变情况等多种因素决定的<sup>[3]</sup>,如2级狭窄虽然在神经根周围4个方向(垂直和横向)无形态学变化,但椎间盘空间狭窄、韧带黄斑增厚、小关节病

变和椎间盘突出等原因同样会导致神经根受压迫,易导致误诊<sup>[12]</sup>。因此,研究一种高效的算法提取图像全局“病灶特征”并实现自动分级诊断,就成为辅助IFS诊断亟待解决的重要问题。

在图像处理算法方面,深度学习模型能够很好地提取图像深层特征,且可通过改变模型的网络拓扑结构,如适当增加网络的深度和宽度,降低特征冗余度和提高特征相关性。自1995年Lo等<sup>[13]</sup>将卷积神经网络用于医学图像分析起,深度学习模型已经迅速成为研究分析医学图像的重要方法<sup>[14]</sup>,且其在辅助脑、乳腺、肺等器官的临床诊断方面已取得重要进展,但在辅助脊椎脊柱疾病诊断的研究基本处于空白。

针对IFS的自动分级问题,本文采用设计基于椎间孔狭窄自动分级的监督式深度卷积神经网络(supervised deep convolutional neural network based on automatic classification of intervertebral foraminal stenosis, IFS-Net),提取图像纹理特征、颜色特征、形状特征和空间关系特征,并将这些特征融合起来作为不同级别的疾病“症状”特征与IFS之间建立特殊的“病理”联系。另外,由于深度学习算法需要海量的训练数据,而IFS影像数据的数量有限,因此本文选用先进的迁移学习<sup>[15]</sup>和微调<sup>[16]</sup>方法,来克服和避免因IFS影像数据规模不足导致的过拟合等问题,提高算法对椎间孔狭窄图像的识别率,满足高标准的临床要求。

## 1 IFS-Net 算法

### 1.1 算法结构

本文设计的IFS-Net是一个端到端的模式识别模型,该模型可实现IFS的高精度自动分级,且具有网络结构简单、高效的特点。端到端的方法使得模型完成提取图像潜在特征的同时,可以自动将loss值前反馈到每一层,经多次迭代调节模型参数,获得最优解。

IFS-Net模型结构共包含8层网络结构,分别是3个卷积层、3个池化层、1个全连接层和1个输出层,其中的全连接层和输出层是该模型的优势模块。因椎间孔图像目标区域小而集中(如图1所示,图1为椎间孔结构及狭窄症分级示意图),故本文采用全连接的方式将所有模型学到的“分布式特征”映射到样本标记空间<sup>[17]</sup>,以达到最大化利用模型提取的特征信息,进而实现更有效的分类。IFS-Net模型框架如图2所示,主要结构和功能包括:

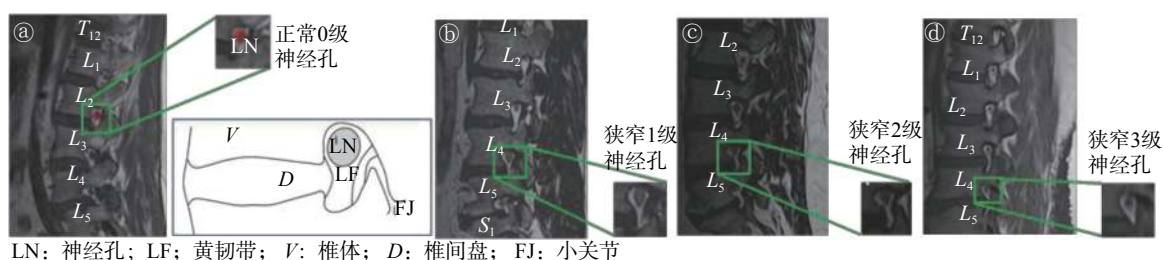


图1 椎间孔结构及狭窄症分级示意

Fig. 1 Intervertebral foramen structure and stenosis classification diagram

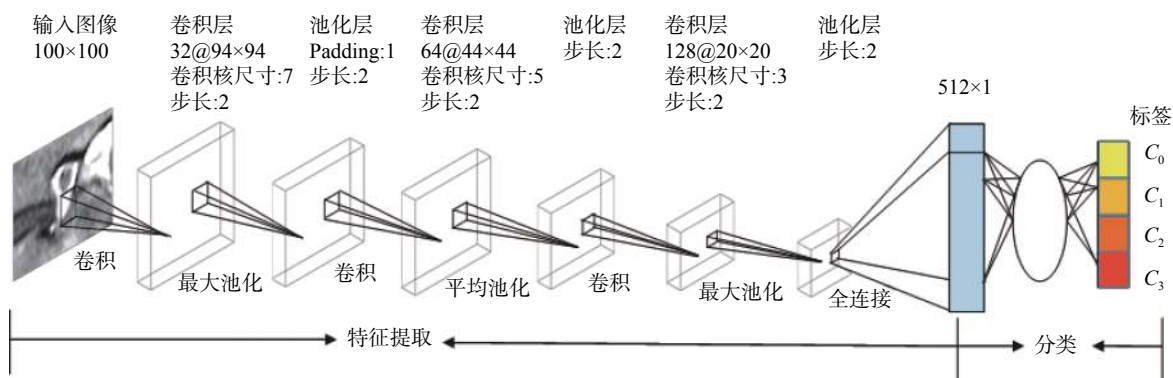


图2 IFS-Net 模型框架

Fig. 2 IFS-Net model framework

1) 输入层: 该层将整幅椎间孔图像导入 IFS-Net 结构, 并产生第 1 个卷积层的输入。

2) 卷积层: 负责特征的学习。该层通过卷积计算连接到输入层或者上一层的局部区域的神经元的输出来提取特征, 每个神经元与前一层的区域连接为稀疏连接。其中输入卷积的权重集为卷积核, 也称作滤波器。一个可学习的卷积核与前一层的若干个 Feature Maps 作卷积, 将所有元素累加后再加上一个偏置, 传给一个非线性激活函数, 本文结构采用 ReLU 函数作为非线性激活函数。另外, 为了减少计算复杂度, 同一层中某些神经元之间的连接权重是共享的, 即相同的。每层滤波器的大小依次递减分别为  $7 \times 7$ 、 $5 \times 5$  和  $3 \times 3$ , 步长设置为 2; 并由高斯分布初始化, 标准差为 0.01。3 个卷积层分别包含 32、64 和 128 个 Feature maps。卷积计算公式如 (1) 所示:

$$x_{K_j}^l = f\left(\sum_{i \in M^{l-1}} x_i^{l-1} * K_{ij}^l + b_j^l\right) \quad (1)$$

式中:  $l$  表示层数;  $K_{ij}$  表示连接第  $l$  层的 Feature Map  $j$  与  $l-1$  层的 Feature Map  $i$  的卷积核;  $M^{l-1}$  表示第  $l-1$  层选择的输入 Feature Maps;  $*$  表示卷积操作;  $b_j^l$  表示  $l$  层中 Map  $j$  的偏置;  $f(\cdot)$  表示非线性激活函数。

3) 池化层: 池化层的目的是通过减少相似特

征点实现 Feature maps 的降维, 除此之外还可以降低噪声和扩大接受域<sup>[15]</sup>。池化层的输出在保持标量不变的基础上减少了参数的数量。本文模型的 3 个池化层分别采用 Max-pooling 和 Mean-pooling 的策略实现特征降维, 步长为 2。池化层拥有与其前端卷积层相等的 Feature map 个数, 分别为 32、64 和 128 个。

4) 全连接层: 该层有 512 个神经单元, 每个神经单元都与上一层进行全连接, 最终得到 512 维特征向量。

5) 输出层: 输出层由欧式径向基函数 (euclidean radial basis function, E-RNF) 单元组成。计算公式为

$$y_j = \sum_j (x_i - w_{ij})^2 \quad (2)$$

式中  $x_i$  和  $w_{ij}$  分别为最后一层卷积层的输入和权重。

此外, 本文模型学习率为 0.000 1, 最大迭代次数为 25 000 次, 动量设置为 0.9, 选取随机梯度下降 (stochastic gradient descent, SGD) 为优化函数。

## 1.2 算法训练策略

### 1.2.1 参数初始化

在 IFS-Net 模型的参数权重初始化方面, 本文采用自训练和迁移学习相融合的参数初始化策略。自训练方法是将网络参数用随机初始化的方



式从头开始训练模型;迁移学习方法则是采用预先训练的参数权重来初始化网络,使得模型预先学到最基本特征(比如色彩、边缘特征等)能够有助于目标数据集分类,并提高算法分类性能。随机初始化网络参数是最常用的深度学习训练方法,但考虑到采用深度学习对医学图像进行处理的一个局限是数据集大小不能得到最大满足。因此,为提高分类精确度并解决深度学习算法在小样本数据集上的过拟合问题,本研究借鉴参考文献[18]的方法,将迁移学习应用于本文算法中。

迁移学习主要原理为:存在一个带有任务  $T_s = \{Y_s, f_s(\cdot)\}$  的源域  $D_s = \{X_s, P_s(X_s)\}$  和一个带有任务  $T_t = \{Y_t, f_t(\cdot)\}$  的目标域  $D_t = \{X_t, P_t(X_t)\}$ , 域和任务之间的条件为  $D_s \neq T_t$  和/或  $T_s \neq T_t$ 。在迁移学习中,  $D_s$  和  $T_s$  获得的“知识”被用来帮助  $D_t$  中预测函数  $f_t(\cdot)$  的学习。其中一个域由特征空间和边缘概率分布两部分组成,  $T_s$  和  $D_s$  为模型训练其他数据集的任务和域,  $T_t$  和  $D_t$  为本文训练任务和域。

本文在 ImageNet 数据集上对模型进行预训练,之后再获得的网络参数迁移到目标数据集上进行训练,从而获得稳定的网络参数。该数据集包含 120 余幅自然图像和 1 000 多种不同的类别。

### 1.2.2 Softmax 分类器

模型采用 Softmax 分类器对模型提取的“分布式特征”进行概率量化, Softmax 是逻辑回归二分类器泛化到多分类的推广。对于训练集  $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$  对应标签为  $y^{(i)} \in \{1, 2, \dots, k\}$ , 本文设定  $k = 4$ 。对于每个输入  $x$  都有对应的每个类的概率  $p(y = i|x)$ , 即 Softmax 函数:

$$P(j) = \frac{\exp(\theta_j^T x)}{\sum_{i=1}^k \exp(\theta_i^T x)} \quad (3)$$

式中:  $\theta_j^T x$  为最够一层向输出层的输入,  $\theta$  为 Softmax 参数。训练优化实质上是逼近最佳  $\theta^T$  的过程。则 Softmax 的损失函数为

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \cdot \log(P(j)) \right] \quad (4)$$

式中:  $1\{y^{(i)} = j\}$  为指示性函数, 取值规则为:  $1\{\text{值为真的表达式}\} = 1$ ,  $1\{\text{值为假的表达式}\} = 0$ 。

## 2 实验及结果分析

### 2.1 数据集

为验证和测试 IFS-Net 模型的性能, 实验选用国际通用的脊柱椎间孔图像 (intervertebral foraminal stenosis znxtb-14-4-hongyanfeis, IFSI) 数据集作为测试数据集。该数据集包含从 110 个临床

实验对象的脊柱图像中采集的 406 个椎间孔图像。每一个脊柱图像均使用矢状  $T_1$  加权 MRI 扫描, 重复时间 (repetition time, TR) 为 4 000 ms, 回波时间 (echo time, TE) 为 85 ms, 磁场为 1.5 t。平面分辨率为 0.5 mm, 厚度为 1 mm 或 1.6 mm。在 IFSI 数据集中, 每个 IFS 图像 Ground Truth 由拥有 7 年工作经验的专业医生手工标注, 分别是 0 级 (正常) 图像 108 幅、1 级 (轻微) 图像 100 幅、2 级 (显著) 图像 105 幅和 3 级 (严重) 图像 93 幅。IFSI 数据集分级标注遵循的是 Wildermuth 椎间孔狭窄分级诊断系统<sup>[19]</sup>。

### 2.2 评价标准

在医学图像数据集中评估机器学习算法的多分类性能, 主要有基于患者水平和图像水平的两种评价方法。由于 IFSI 数据集是基于图像层面的, 因此本文从图像层面评估算法的识别率。假设  $N_{\text{all}}$  为图像验证集或测试集的图像总数量, 如果  $N_r$  幅图像被正确分类, 那么图像层面的准确率 RR 为

$$\text{RR} = N_r / N_{\text{all}} \quad (5)$$

由于多分类问题中准确率不再是唯一的评价指标, 因此本文引入 F-score 作为评价本文模型的另一个指标。与 F-score 密切相关的两个值分别为精确度  $P_r$  和召回率  $R_c$ :

$$P_r = \text{TP} / (\text{TP} + \text{FP}) \quad (6)$$

$$R_c = \text{TP} / (\text{TP} + \text{FN}) \quad (7)$$

式中: TP 为真阳性记录数; FN 为假阴性记录数; FP 为假阳性记录数。

F-score 作为调和精确度和召回率的一个综合指标, 计算公式如下:

$$F_\beta = (\beta^2 + 1) / \beta^2 P_r + R_c \quad (8)$$

另外,  $\beta$  用来调节两部分的比重。当  $\beta = 1$  时, 式 (8) 退化成简单的调和平均函数, 称为  $F_1$ -score:

$$F_1 = \frac{2P_r R_c}{P_r + R_c} + R_c \quad (9)$$

### 2.3 实验设计

为保证实验的严谨性, 尽量避免特殊个例影响最终结果, 本文采用交叉验证的方式将 IFSI 数据集平均分成 5 份, 每次实验训练集和测试集比例为 4:1, 且确保用于构建训练集的图像不再用于测试。

本文算法采用 Caffe 模型编写基本框架<sup>[20]</sup>。算法在 Lenovo ThinkStation、Intel e3 CPU、NVIDIA Quadro K2200 GPU 上训练, 平均训练时间为 4 min 21 s, 单张测试时间平均为 0.047 s。

## 2.4 对比实验

为了验证本文模型的有效性,本文采用传统机器学习算法对 IFSI 数据集进行分类。传统机器学习算法主要包含两个步骤:特征提取和分类。特征提取部分本文选取 5 种最先进的纹理表示的全局算法,分别是:局部二值模式 (local binary patterns, LBP)<sup>[21]</sup>、局部相位量化 (local phase quantization, LPQ)<sup>[22]</sup>、灰度共生矩阵 (gray-level co-occurrence matrix, GLCM)<sup>[23]</sup>、方向梯度直方图特征 (histogram of oriented gradient, HOG)、ORB (oriented FAST and rotated BRIEF)<sup>[24]</sup>。5 种特征描述子的特征维度如表 1 所示。

表 1 特征描述子及其特征向量维度列表

Table 1 Feature descriptor and feature vector dimension table

描述子	特征向量维度
LBP	59
LPQ	256
GLCM	8
HOG	4 356
ORB	32

本文利用 4 种不同的分类器来评估上述特征集: k 近邻算法 (k-NN)<sup>[25]</sup>、极限学习机 (extreme learning machine, ELM)<sup>[26]</sup>、支持向量机 (support vector machine, SVM)<sup>[27]</sup> 以及随机森林算法 (random forest, RF)<sup>[28]</sup>。

## 2.5 实验结果

### 2.5.1 IFS-Net 与传统机器学习算法性能对比结果

为验证本文算法在 IFSI 数据集上的多分类性能,将其与具有代表性的 4 种传统机器学习分类器算法 (k-NN、ELM、SVM 和 RF) 对比分析,实验结果如表 2 所示。

在测试集准确度和  $F_1$  两个评价指标上,本文设计的 IFS-Net 算法模型的结果明显优于传统机器学习算法所取得的结果。由图 3 传统机器学习算法 (特征描述子+分类器) 分类准确率对比图,可更直观地看出,传统机器学习方法中,选取的特征向量在分类性能上都表现出稳定接近的结果,4 种传统分类器对于每个特征描述子的识别率差距都在 4% 之内。值得一提的是,ORB 虽然为关键点描述符且通常用于对象识别,但其在微观图像分类问题上获得了比传统纹理特征更优的结果。

表 2 不同方法的准确率和  $F_1$  统计表

Table 2 Accuracy of different methods and  $F_1$  statistics

传统分类器	方法	验证集 准确率	测试集 准确率	$F_1$
1-NN	LBP	49.4 ± 2.1	51.8 ± 2.5	0.712
	LPQ	63.5 ± 2.7	53.7 ± 2.7	0.517
	GLCM	65.3 ± 1.7	53.6 ± 1.6	0.725
	ORB	66.1 ± 3.5	63.7 ± 3.4	0.764
	HOG	51.7 ± 1.3	61.3 ± 0.7	0.768
ELM	LBP	69.4 ± 1.0	51.3 ± 2.1	0.742
	LPQ	61.6 ± 1.8	57.6 ± 2.3	0.76
	GLCM	73.6 ± 2.9	57.3 ± 3.0	0.786
	ORB	67.2 ± 1.9	65.4 ± 3.8	0.73
	HOG	63.5 ± 2.7	61.4 ± 1.8	0.697
RF	LBP	55.8 ± 4.7	53.4 ± 2.8	0.711
	LPQ	61.4 ± 0.4	55.1 ± 1.5	0.723
	GLCM	74.3 ± 3.6	56.6 ± 2.8	0.731
	ORB	70.1 ± 2.5	67.6 ± 2.4	0.693
	HOG	73.8 ± 5.0	62.0 ± 1.5	0.665
SVM	LBP	50.2 ± 1.9	52.4 ± 0.4	0.728
	LPQ	71.1 ± 6.4	56.2 ± 2.6	0.694
	GLCM	69.3 ± 1.0	54.7 ± 3.8	0.737
	ORB	63.5 ± 2.2	64.1 ± 2.4	0.807
	HOG	73.4 ± 0.8	64.7 ± 1.1	0.81
深度学习 模型	LeNet-5	79.375% ± 1.3	78.5 ± 2.1	0.614
	Alex-Net	65.2% ± 3.8	63.7 ± 0.7	0.739
	GoogLeNet	91.6% ± 1.1	46.3 ± 1.4	0.51
	IFS-Net	82.15% ± 1.7	81.2 ± 1.6	0.763
IFS-Net+TL*		87.5% ± 2.3	86.375 ± 1.8	0.857

注: TL\*为Transfer Learning

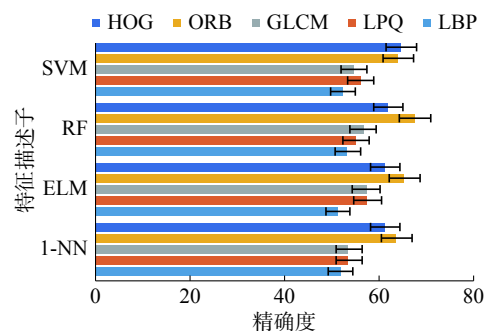


图 3 传统机器学习算法 (特征描述子+分类器) 分类准确率对比

Fig. 3 The classification accuracy of traditional machine learning algorithm (feature descriptor + classifier)

### 2.5.2 IFS-Net 与其他深度学习算法性能对比结果

本文算法与其他典型深度学习算法在 IFSI 数据集上的分类性能对比实验结果,如图 4 所示。实验中,IFS-Net 取得了 87.5% 的分类精度结果,明显优于其他深度学习算法。该图还直观地显示出本文算法的验证集和测试集几乎具有相同的精确度,这表明 IFS-Net 模型具有较强的泛化和避免过拟合的能力。而由于数据集的限制,层数过深的卷积神经网络在 IFSI 数据集上没有表现出层级优势,如 GoogLe-Net 模型出现的严重的过拟合现象,就是由于其过于复杂的结构造成的,这也是本文采用较为简单网络结构的一个原因。另外,采用迁移学习+微调训练策略的分类精度高于自训练 2.5%。如图 5 不同模型 loss 曲线拟合度对比所示,从 IFS-Net 和 IFS-Net + TL 两个模型的 train-loss 和 val-loss 曲线图可以看出,采用迁移学习方法能够降低模型的过拟合程度。

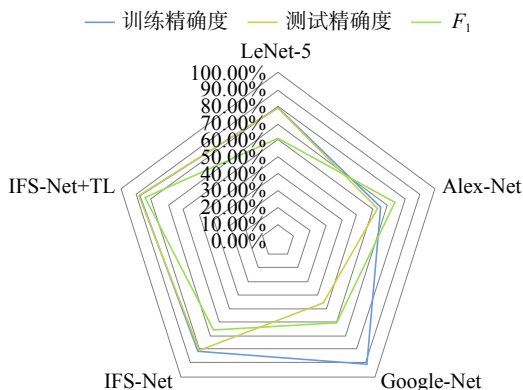


图 4 不同深度学习模型算法性能对比

Fig. 4 Performance comparison of different depth learning model algorithms

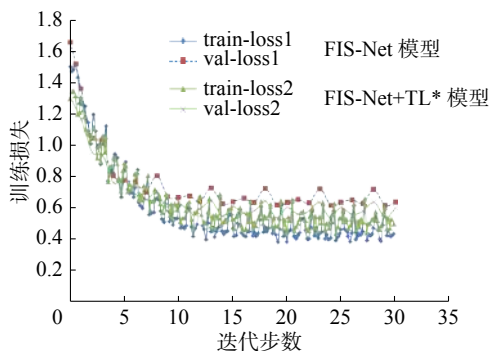


图 5 不同训练策略 loss 曲线拟合度对比

Fig. 5 Different training strategy models loss curve fitting degree comparison chart

### 2.5.3 IFS 图像狭窄度识别分析结果

为进一步测试 IFS-Net 算法在狭窄程度识别

分析中的性能,本研究在 IFSI 数据集上做了不同 IFS 图像狭窄度识别分析测试实验,实验结果如表 3 所示。实验结果表明,IFS-Net 算法对 IFSI 数据集中的 0 级和 3 级的区分度最高,1 级分类结果较差。分析其原因,主要是 1 级图像中的椎间孔是在两个相对方向(垂直或横向)的脂肪闭塞引起的轻度口部狭窄,没有显示神经根的形态学变化<sup>[19]</sup>,因此容易被误分为  $C_0$ (正常类)。另外,虽然 2 级椎间孔也没有表现出形态学上的变化,但其周围结构的变化使得算法可以挖掘出潜在特征并对特征进行相关性建模实现精准分类,说明本文模型具有提取微小病变特征的能力,能够将不同级别的疾病“症状”特征与 IFS 分级之间建立起特殊的“病理”联系。

表 3 IFS-Net 算法狭窄度分类结果混淆矩阵

Table 3 A hybrid matrix of the classification results of IFS-Net algorithm

狭窄度类别	$C_0$	$C_1$	$C_2$	$C_3$	ACC/%
$C_0$	24	0	0	2	92.31
$C_1$	4	16	3	3	71.54
$C_2$	4	0	21	0	84
$C_3$	0	0	3	22	88

## 3 结束语

为提高计算机辅助椎间孔狭窄症诊断准确率及工作效率,本文提出并设计出一种基于深度学习的椎间孔狭窄图像自动多分级模型。并且在基于设计的 IFS-Net 模型基础上,采用迁移学习有效提升了网络分类性能。实验结果表明,本文设计的模型不仅结构简单,且分类性能高效,可为临床辅助诊断及分级提供有效的技术支撑。本研究中迁移学习的特定引入方式以及数据联合训练的应用方法,对于医学影像微观数据集分析及多分类任务具有一定的参考价值。在后续的研究工作中,将针对椎间孔狭窄图像多分级工作,融合特定的形态解剖学特征,研究设计性能更优的深度学习模型算法,进一步提高分类识别的性能,给临床提供更有效的辅助诊断技术。

## 参考文献:

- [1] KANEKO Y, MATSUMOTO M, TAKAISHI H, et al. Morphometric analysis of the lumbar intervertebral foramen in patients with degenerative lumbar scoliosis by mul-

- tidetector-row computed tomography[J]. *European spine journal*, 2012, 21(12): 2594–2602.
- [2] RAJAEE S S, BAE H W, KANIM L E, et al. Spinal fusion in the united states: analysis of trends from 1998 to 2008[J]. *Spine*, 2012, 37(1): 67–76.
- [3] LEE S, LEE J W, YEOM J S, et al. A practical MRI grading system for lumbar foraminal stenosis[J]. *American journal of roentgenology*, 2010, 194(4): 1095–1098.
- [4] HAN Zhongyi, WEI Benzhen, LEUNG S, et al. Automated pathogenesis-based diagnosis of lumbar neural foraminal stenosis via deep multiscale multitask learning[J]. *Neuroinformatics*, 2018, 16(3/4): 325–337.
- [5] ALOMARI R S, CORSO J J, CHAUDHARY V. Labeling of lumbar discs using both pixel- and object-level features with a two-level probabilistic model[J]. *IEEE transactions on medical imaging*, 2011, 30(1): 1–10.
- [6] ZHAN Yiqiang, MANEESH D, HARDER M, et al. Robust MR spine detection using hierarchical learning and local articulated model[C]//*Proceedings of the 15th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Nice, France, 2012: 141–148, DOI: [10.1007/978-3-642-33415-3\\_18](https://doi.org/10.1007/978-3-642-33415-3_18).
- [7] WANG Zhijie, ZHEN Xiantong, TAY K, et al. Regression segmentation for M3 spinal images[J]. *IEEE transactions on medical imaging*, 2015, 34(8): 1640–1648.
- [8] GHOSHA S, ALOMARI R S, CHAUDHARY V, et al. Automatic lumbar vertebra segmentation from clinical CT for wedge compression fracture diagnosis[C]//*Proceedings of the SPIE 7963, Medical Imaging 2011: Computer-Aided Diagnosis*. Lake Buena Vista (Orlando), United States, 2011: 796303, DOI: [10.1117/12.878055](https://doi.org/10.1117/12.878055).
- [9] HUANG S H, CHU Yihong, LAI Shanghong, et al. Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI[J]. *IEEE transactions on medical imaging*, 2009, 28(10): 1595–1605.
- [10] KLINDER T, WOLZ R, LORENZ C, et al. Spine segmentation using articulated shape models[C]//*Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention*. New York, USA, 2008: 227–234.
- [11] HE Xiaoxu, YIN Yilong, SHARMA M, et al. Automated diagnosis of neural foraminal stenosis using synchronized superpixels representation[C]//*Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Athens, Greece, 2016: 335–343.
- [12] VERBIEST H. Results of surgical treatment of idiopathic developmental stenosis of the lumbar vertebral canal. A review of twenty-seven years' experience[J]. *The journal of bone and joint surgery*, 1977, 59(2): 181–188.
- [13] LO S C B, LOU S L A, LIN J S, et al. Artificial convolution neural network techniques and applications for lung nodule detection[J]. *IEEE transactions on medical imaging*, 1995, 14(4): 711–718.
- [14] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. *Medical image analysis*, 2017, 42(9): 60–88.
- [15] WANG Xiaosong, PENG Yifan, LU Le, et al. ChestX-Ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases[C]//*Proceeding of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA, 2017: 3462–3471.
- [16] SHEN Dinggang, WU Guorong, SUK H I. Deep learning in medical image analysis[J]. *Annual review of biomedical engineering*, 2017, 19: 221–248.
- [17] BENGIO Y, DELALLEAU O. On the expressive power of deep architectures[C]//*Proceedings of the 22nd International Conference on Algorithmic Learning Theory*. Espoo, Finland, 2011: 18–36.
- [18] PAN S J, YANG Qiang. A survey on transfer learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345–1359.
- [19] WILDERMUTH S, ZANETTI M, DUEWELL S, et al. Magnetic resonance imaging and magnetic resonance myelography in the presurgical diagnosis of lumbar foraminal stenosis [J]. *Spine*, 2007, 32(8): 896–903.
- [20] JIA Yangqing, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding[C]//*Proceedings of the 22nd ACM International Conference on Multimedia*. Orlando, USA, 2014: 675–678.
- [21] OJALA T, PIETIKÄINEN M, MÄENPÄÄ T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2002, 24(7): 971–987.
- [22] GUO Zhenhua, ZHANG Lei, ZHANG D. A completed modeling of local binary pattern operator for texture classification[J]. *IEEE transactions on image processing*, 2010, 19(6): 1657–1663.
- [23] HONG Huichao, ZHENG Lixin, PAN Shuwan. Fast com-



putational technique for gray-level co-occurrence matrix based on graphics process unit in biomedical engineering applications[J]. Journal of medical imaging and health informatics, 2018, 8(2): 309–312.

- [24] RUBLEE E, RABAU V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]//Proceedings of 2011 International Conference on Computer Vision. Barcelona, Spain, 2011: 2564–2571.
- [25] ŞAHAN S, POLAT K, KODAZ H, et al. A new hybrid method based on fuzzy-artificial immune system and k-nn algorithm for breast cancer diagnosis[J]. Computers in biology and medicine, 2007, 37(3): 415–423.
- [26] WONG P K, GAO Xianghui, WONG K I, et al. Online extreme learning machine based modeling and optimization for point-by-point engine calibration[J]. Neurocomputing, 2018, 277: 187–197.
- [27] CORTES C, VAPNIK V. Support-vector networks[J]. Machine learning, 1995, 20(3): 273–297.
- [28] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5–32.

### 作者简介:



洪雁飞, 女, 1994 年生, 硕士研究生, 主要研究方向为医学影像分析和机器学习算法。



魏本征, 男, 1976 年生, 教授, 主要研究方向为医学图像处理、机器学习、计算医学和医学信息工程。发表学术论文 60 余篇。



刘川, 男, 1972 年生, 副教授, 主要研究方向为中医骨伤学和中医药文化社会学。