



基于深度残差神经网络迁移学习的 牙形刺图像识别*

柳天滋^{1**} 陈 昕¹ 李 想¹ 樊 茹² 刘逸鹏¹ 李焕菁¹

1 中国农业大学信息与电气工程学院, 北京 100083

2 石油地质实验研究中心, 中国石油勘探开发研究院, 北京 100083

摘要 为了探索基于深度神经网络模型的牙形刺图像智能识别效果, 研究选取奥陶纪 8 种牙形刺作为研究对象, 通过体视显微镜采集牙形刺图像 1188 幅, 收集整理公开发表文献的牙形刺图像 778 幅, 将图像数据集划分为训练集和测试集。通过对训练集图像进行旋转、翻转、滤波增强处理, 解决了训练样本不足的问题。基于 ResNet-18、ResNet-34、ResNet-50、ResNet-101、ResNet-152 五种残差神经网络模型, 采用迁移学习方法, 对网络模型进行训练以获取模型参数, 五种模型测试 Top-1 准确率分别为 85.37%、85.85%、83.90%、81.95%、80.00%, Top-2 准确率分别为 94.63%、94.63%、94.15%、93.17%、93.66%, 模型对牙形刺图像具有较好的识别效果。通过对比研究发现, ResNet-34 识别准确率最高, 说明对于特征简单的牙形刺属种, 增加网络深度并不一定能提升准确率, 而确定合适深度的模型则不仅可以提高识别准确率, 还可以节约计算资源。通过 ResNet-34 模型的迁移学习训练和重新训练效果对比可以看出, 迁移学习不仅可以获得较高的准确率, 而且可以较快获取模型参数, 因而可作为小样本古生物化石图像识别的重要方法。研究还发现, 体视显微镜下牙形刺图像的识别准确率高于扫描电镜下图像识别准确率, 化石完整性和相似性、照相角度以及数据集的大小是影响图像识别准确率的主要原因。

关键词 牙形刺 残差网络 迁移学习 增强处理 图像识别

中文引用 柳天滋, 陈昕, 李想, 樊茹, 刘逸鹏, 李焕菁, 2020. 基于深度残差神经网络迁移学习的牙形刺图像识别. 古生物学报, 59(4): 512–523. DOI: 10.19800/j.cnki.aps.2020.042

英文引用 Liu Tian-zi, Chen Xin, Li Xiang, Fan Ru, Liu Yi-peng, Li Huan-jing, 2020. Conodont image recognition based on transfer learning of deep residual neural network. Acta Palaeontologica Sinica, 59(4): 512–523. DOI: 10.19800/j.cnki.aps.2020.042

Conodont image recognition based on transfer learning of deep residual neural network

LIU Tian-zi¹, CHEN Xin¹, LI Xiang¹, FAN Ru², LIU Yi-peng¹ and LI Huan-jing¹

¹ College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China;

投稿日期: 2020-07-15; 改回日期: 2020-11-12; 录用日期: 2020-11-20

* 北京市大学生创新训练项目“基于卷积神经网络的牙形刺图像分类方法研究”(编号 201910019363)资助。

** 通讯作者: 柳天滋, 大学本科生; e-mail: tianziliu@sina.cn

2 Central Laboratory of Geological Sciences, Research Institute of Petroleum Exploration and Development, Beijing 100083, China

Abstract This paper aims to explore the image recognition effect of conodonts based on deep neural network model. Eight kinds of Ordovician conodonts were selected as the research objects. 1188 photos of conodonts were taken under stereomicroscope, and 778 SEM images of conodonts were collected from published papers. The image data set was divided into training set and test set. In order to solve the problem of insufficient training samples, the methods of data augmentation of rotating, flipping, and filtering are used. Five residual neural network models of resnet-18, resnet-34, resnet-50, resnet-101 and resnet-152, were used to train to obtain the transfer learning model parameters. The top-1 test accuracy for the five models is 85.37%, 85.85%, 83.90%, 81.95%, and 80.00% respectively. The top-2 test accuracy is 94.63%, 94.63%, 94.15%, 93.17%, and 93.66% respectively. It was found that resnet-34 recognition has the highest test accuracy, which indicates that for the simple fossil like conodonts, increasing the depth of network does not improve the test accuracy. Selecting the appropriate depth model can not only improve the recognition accuracy, but also save computing resources. By comparing the transfer learning and retraining results of resnet-34 model, it was found that transfer learning can not only obtain high accuracy, but also quickly obtain model parameters, which is an important method for small sample fossil image recognition. The recognition accuracy of conodont image under stereomicroscope is higher than that under scanning electron microscope. Integrity and similarity of conodont, camera angle, and dataset size are the main factors that affects the image recognition accuracy.

Key words conodont, residual network, transfer learning, data augmentation, image recognition

1 前言

牙形刺是一种从寒武纪中期到三叠纪的微体石生物化石(王志浩等, 2011), 时间跨度达4亿年; 其种类多、多布广、演化快, 在地层划分对比中具有十分重要的意义。传统的牙形刺鉴定方法是专业技术人员在体视显微镜下观察从岩石中分选出的化石特征并与标准化石对比, 从而完成分类、定名(张放等, 2010)。由于牙形刺种类繁多、体形小(多数大小仅0.2–2 mm), 化石鉴定对专业知识和个人经验要求极高, 建立一种智能、准确、高效的牙形刺化石的鉴定方法是古生物学界目前亟需探寻的一个重要方向。

机器学习近年来发展迅速且在图像分类上取得了巨大的进步, 为本文研究智能、高效的牙形刺化石鉴定方法提供了思路。由于传统的机器学习算法通常只能在训练过程中学习到较为低级的特征, 如物体的轮廓或纹理, 这对于需要精确识别大量特征的牙形刺鉴定来说远远不够。同时, 传统的机器学习算法如 SVM, 只有二分类的算法, 而牙形刺的鉴定需要对多个类别进行学习, 深度学习成了最佳的选择。深度学习除了能够在训练过程中学习到低级特征, 也能够学习到目标物体的细微特征, 给牙形刺的鉴定提供了有力保障。

近几年, 随着深度学习模型的提出和完善, 基于卷积神经网络(CNN)模型在人脸识别、医疗分类、

场景分类等领域取得了较好的效果(牛新等, 2016; 魏存超, 2017; 薛迪秀等, 2017; 殷俊、杨万扣, 2018; 马中启等, 2019), 经典的神经网络模型包括 LeNet (Lecun *et al.*, 1998)、ALexNet (Krizhevsky *et al.*, 2012)、VGGNet (Simonyan and Zisserman, 2015)、GoogLeNet (Szegedy *et al.*, 2014)、ResNet (He *et al.*, 2016)、DenseNet (Iandola *et al.*, 2014)等。残差网络模型(ResNet)较好地解决了网络加深带来的学习退化的问题, 赢得了ILSVRC & COCO 2015比赛的图像识别、图像定位、公共图像识别和分类冠军, 该模型目前在各领域取得了广泛应用。

深度学习模型的训练依赖于大数据集, 但很多领域积累的图像数据并没有达到深度学习训练的要求, 样本的不足会导致过拟合的问题, 所以小样本的学习越来越受到关注。Wang 等(2019)从样本数量、带标签样本数量、先验知识的作用等角度对小样本学习技术进行了讨论和分析。2010年以后, 大量的文献利用语义迁移来解决训练样本不足的问题(Wang *et al.*, 2019)。数据增强和迁移学习在小样本深度学习受到普遍关注(刘颖等, 2020)。

目前, 基于神经网络的牙形刺图像识别研究主要是利用人工神经网络方法对牙形刺鉴定中若干“模糊特征”的数字化研究(黄铮等, 2009); 然而, 基于深度学习的图像识别研究涉及较少, 主要原因是图像数据不足及方法适用性不是很清楚。早期的研究主要针对化石的鉴定, 发表的属种只是少

量典型标本的扫描电镜图像,同一种的牙形刺图像少而分散。体视显微镜通常也只是用于鉴定,很少见有体视显微镜下拍摄的牙形刺图像公开发表。

本文通过以下途径解决基于深度学习的牙形刺图像识别和分类问题:(1)从传统的单一的扫描电镜成像转向低成本高效的更接近人眼观察效果的体视显微镜成像,近年来体视显微镜成像技术快速发展使得这种成像效果大为改善;(2)收集整理大量已发表的文献图片,扩大同一属种牙形刺图像样本数据集;(3)通过图像增强处理扩大训练集的规模;(4)通过迁移学习,将其他领域的图像分类模型应用到小样本牙形刺图像分类研究中。本文采用的深度学习模型是目前在图像分类中取得较好分类效果的残差网络模型。通过本项研究,对比不同深度模型的图像识别效果以及扫描电镜图像和体视显微镜图像测试准确率的差异,探索智能化的古生物化石鉴定新方法。

2 牙形刺图像数据集建立

本文用于深度学习的牙形刺图像数据集的建立包括以下三个步骤:原始图像的获取,数据集的划分以及图像增强处理。

2.1 原始图像获取

本次选取奥陶纪8种牙形刺(图1)为代表进行深度学习研究,分别是:中村圆柱刺 *Teridontus nakamurai*、先祖肿刺 *Cordylodus proavus*、叉开锯颚刺 *Serratognathus diversus*、变异胡安颚刺 *Juanognathus variabilis*、扁平小剑刺 *Scabbardella altipes*、短尾哈玛拉刺 *Hamarodus brevirameus*、角肿刺 *Cordylodus angulatus*、马尼托罗斯刺 *Rossodus manitouensis*。

牙形刺图像通过两种方法获取。一种是来自公开发表的文献(Landing, 1976; Nowlan, 1981; Stouge *et al.*, 1985; Stouge and Rasmussen, 1996; Kaljo *et al.*, 1988; Orndorff, 1988; Nicoll, 1990; Pohler and Orchard, 1990; Taylor *et al.*, 1991; Ji and Barnes, 1994, 1996; Seo *et al.*, 1994; Löfgren, 1996, 1997; Ferretti and Barnes, 1997; Ferretti, 1998; Parsons and Clark, 1999; Zhen *et al.*, 1999, 2011; Johnston and Barnes, 2000; Erlström *et al.*,

2001; Pyle and Barnes, 2001, 2002, 2007; Viira *et al.*, 2002; Landing *et al.*, 2003, 2007, 2010; Bergström *et al.*, 2004; Lee, 2004; Löfgren *et al.*, 2005; Percival and Zhen, 2006; Agematsu *et al.*, 2007; Löfgren and Viira, 2007; Tolmacheva *et al.*, 2008; Zhen and Nicoll, 2009; 王志浩等, 2011),这些图像均为扫描电镜拍摄。不同年代发表的文献图像效果差别较大,早期的图像为黑白图,后期的图像是较高分辨率的灰度图,这类图像共收集整理 778 幅。在不同文献中搜集数据集时,已通过查看文献图版的来源排除了重复引用的图像。另一种图像是笔者通过蔡司 SteREO Discovery V20 体视显微镜拍摄,是一种体视镜下的彩色图像,这类图像共获取 1188 幅(图 1, 表 1),拍摄及图像整理过程中,通过专家进行了鉴定,对不确定种名的图像进行了剔除。

2.2 数据集划分与增强处理

首先将 1966 幅扫描电镜和体视显微镜牙形刺图像按大约 9:1 的比例划分为训练集和测试集,训练集共有 1761 幅图像,测试集共有 205 幅图像(图 2)。

为了扩大训练数据集规模,对训练集图像进行了增强处理。传统的数据增强处理包括旋转、翻转、滤波、裁剪、形变等,因牙形刺种与种之间的差别有时很小,需要依靠某个细微特征进行区分,所以牙形刺的识别重点在让神经网络尽量精确地学习到牙形刺的各个特征。而若采用剪裁、形变等数据增强方法,可能会丢失重要特征,导致数据增强后的图像变得不再可靠。所以,本次实验对训练集牙形刺图像采用三种方法进行了增强处理:旋转、翻转、滤波。旋转是在 $[-30^{\circ}, 30^{\circ}]$ 之间的随机选取,旋转以后边缘空白处采用邻近背景充填;图像翻转采用水平翻转的方式;图像滤波采用高斯滤波方式,经滤波后的图像增加了噪声,较原图像模糊。原训练集 1761 幅图像经三种方式增强处理后,训练集图像数据集扩大到 7044 幅(图 2)。

3 迁移学习网络架构与参数设置

3.1 迁移学习流程

本文选择前人利用 ImageNet 图像数据集训练得到的模型作为基础进行迁移学习。与传统图

表 1 扫描电镜和体视显微镜获取的牙形刺图像数量统计表
Table 1 Dataset of conodont images under SEM and stereomicroscope

编号	中文名称	拉丁文名称	体视显微镜照片		扫描电镜照片		小计
			训练集	测试集	训练集	测试集	
A	中村圆柱刺	<i>Teridontus nakamurai</i>	78	5	89	10	182
B	先祖肿刺	<i>Cordylodus proavus</i>	0	0	178	24	202
C	叉开锯颚刺	<i>Serratognathus diversus</i>	272	31	64	12	379
D	变异胡安颚刺	<i>Juanognathus variabilis</i>	216	21	24	8	269
E	扁平小剑刺	<i>Scabbardella altipes</i>	177	13	38	13	241
F	短尾哈玛拉刺	<i>Hamarodus brevirameus</i>	180	15	62	14	271
G	角肿刺	<i>Cordylodus angulatus</i>	0	0	183	19	202
H	马尼托罗斯刺	<i>Rossodus manitouensis</i>	168	12	32	8	220
总计			1091	97	670	108	1966

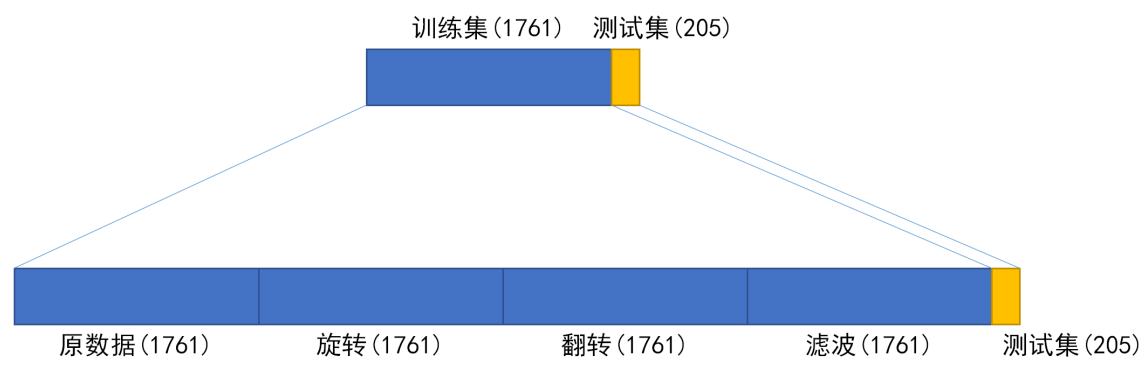


图 2 牙形刺图像数据集划分示意图
Fig. 2 Diagram of data set division of conodont images

数据集的类别数 8。在此基础上，对牙形刺图像识别模型进行训练，利用每一次训练获得的模型对训练集进行识别，得到训练准确率，对测试集牙形刺图像进行识别，得到测试集的准确率(图 3)。

3.2 残差网络结构

为了解决深度学习神经网络层数过深导致梯度消失的退化问题，ResNet 网络模型引进残差单元(He *et al.*, 2016)。残差网络结构使得某一层的输出可以跨过几层作为后面某一层的输入，直接执行恒等映射，从而解决了网络层数过深引起的退化问题。

目前应用较广的残差网络模型有 ResNet-18、ResNet-34、ResNet-50、ResNet-101 和 ResNet-152 五种不同深度的网络结构(He *et al.*, 2016)，其卷积层参数配置如表 2 所示。为了追求更高的准确率，本文将模型的输入在计算机算力允许的情况

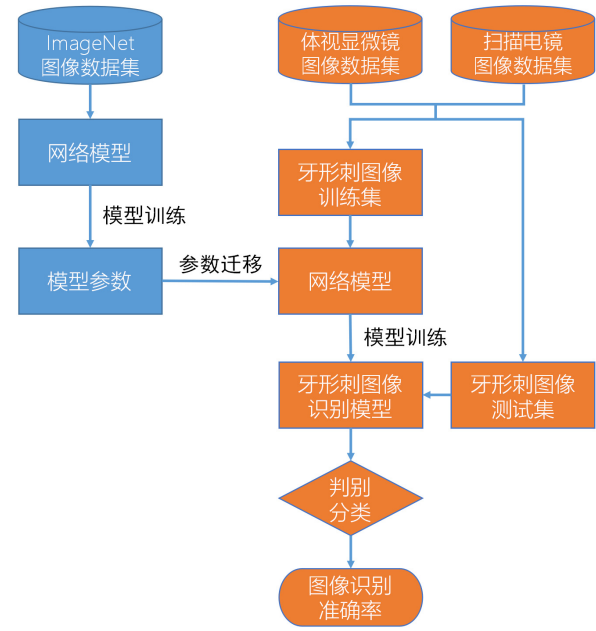


图 3 牙形刺图像分类迁移学习流程
Fig. 3 Transfer learning flow chart of image recognition of conodonts

表 2 不同深度常用残差网络模型结构及参数配置
Table 2 Structure and parameter configuration of residual network models

层名	输出尺寸	模型名称				
		ResNet-18	ResNet-34	ResNet-50	ResNet-101	ResNet-152
conv1	128×128	7×7, 64, 步长为 2				
		3×3 最大池化层, 步长为 2				
conv2_x	64×64	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	32×32	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	16×16	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	8×8	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	平均池化, 8 分类全连接层, softmax 分类层				

下由 112×112 扩大至 128×128。通过对比分析五种网络结构对牙形刺图像识别效果的影响。

3.3 参数设置与优化

读入图像数据后采用批训练的方式进行训练, 每批训练样本 32 个, 即在实验过程中一次读入 32 幅训练集图像数据, 这样可以利用 GPU 的并行计算能力, 加速牙形刺识别模型的训练过程。本次研究中图像尺寸参数设置为 256, 初始学习率设为 0.001, 使用 Adam 优化算法动态调整每个参数的学习率, 激活函数为 Relu 函数, 使用 SoftMax 分类器进行分类, 采用交叉熵损失函数来评估真实值与预测值之间的差距, 训练迭代轮次(epoch, 对训练集全部样本完成一次抽样训练为一个周期)根据试验确定为 100 个 epoch。

为了防止小样本的牙形刺图像训练出现过拟合问题, 采用 L2 范数正则化和丢弃法(dropout)对网络模型进行优化, 旨在减少测试误差, 提高泛化能力。L2 正则化参数取值为 0.0012, dropout 参数设置为 0.05。

本次深度学习在双 GPU 工作站上进行, 工作站硬件配置如下: CPU 为英特尔酷睿第九代 I9 9900K (八核十六线程), 主板为 intel Z390 工作站主板, 内存为海盗船 DDR4 32G 3000 高频内存, 硬盘为三星 PM981 512G NVME 高速固态硬盘, 显卡为 RTX 2080TI 双 GPU 显卡, 操作

系统为 Windows 10。

4 结果与讨论

4.1 评估标准

本文评估标准采用 Top-1 准确率和 Top-2 准确率作为评估标准。Top-1 准确率是指最后输出的概率向量中最大概率所代表的牙形刺类别与正确的牙形刺类别一致的概率, Top-2 准确率是指最后输出的概率向量值最大的 2 个概率所代表的牙形刺类别中包含有正确的牙形刺类别的概率。未特别指定的情况下一般指 Top-1 准确率。

4.2 不同深度网络模型训练结果

5 种网络模型的训练结果显示, 随着迭代次数的增加, 训练集和测试集准确率呈增长趋势(图 4)。不同模型中训练集 Top-1 准确率分布范围为 97.34%–99.31%; 测试集 Top-1 准确率分布范围为 80.00%–85.85%, Top-2 准确率分布范围为 93.17%–94.63% (表 3), 总体而言, 残差网络模型显示了较好的识别效果。虽然进行了 100 轮次的迭代训练, 但实际上准确率的变化在迭代 30 次基本收敛, 这种现象说明迁移学习具有较好的效果, 此外准确率快速收敛可能与牙形刺的形态等特征简单有关。不同的网络模型训练在后期都不同程度有退化现象, 表现为测试准确率的下降。

表 3 不同深度残差网络模型迁移学习与重新训练结果
Table 3 Results transfer learning and retraining of different depth of ResNet

训练方式	网络模型	模型深度 /层	模型参数文 件大小/M	计算时间 /m	训练准 确率/%	测试 Top-1 准确率/%	测试 Top-2 准确率/%
迁移学习	ResNet-18	18	47.83	265	99.31	85.37	94.63
	ResNet-34	34	87.37	290	99.21	85.85	94.63
	ResNet-50	50	157.70	298	98.81	83.90	94.15
	ResNet-101	101	232.15	402	98.72	81.95	93.17
	ResNet-152	152	293.50	441	97.34	80.00	93.66
重新训练	ResNet-34-N	34	89.35	343	94.88	79.51	90.73

网络模型的深度往往影响到图像特征表达的效果,深度越深,迁移学习后的准确率一般而言会提高(关胤,2019;陈猛夫,2020;樊帅昌等,2020)。ResNet-18、ResNet-34、ResNet-50、ResNet-101、ResNet-152 不同深度网络模型迁移学习 Top-1 准确率分别为 85.37%、85.85%、83.90%、81.95%、80.00%, Top-2 准确率分别为 94.63%、94.63%、94.15%、93.17%、93.66%(表 3)。对比发现,网络深度越深,Top-1 和 Top-2 准确率均有下降趋势,其中 Top-1 准确率下降明显,而 Top-2 准确率下降不大。结果表明,对于特征简单的化石图像进行分类的网络模型训练,所采用的网络模型深度不必追求过深和过复杂,以适用为宜。对比不同的网络模型的训练时间可以看出,ResNet-18 网络模型训练时间只有 265 m,而 ResNet-152 网络模型的训练时间为 441 m,采用简单适用的网络模型可以节省更多的计算资源。

不同模型训练收敛速度差异明显,网络越深,模型收敛越慢(图 4)。ResNet-18 模型 10 轮次训练就能达到收敛,而 ResNet-152 模型训练集经过 50 轮次训练后训练准确率趋于稳定,对应的测试准确率在 20 轮次的训练后增长不明显。此外,较浅的网络模型,如 ResNet-18 和 ResNet-34,预测准确率的跳跃性变化幅度较小,模型预测较稳定;而较深的网络模型,如 ResNet-152,预测准确率的跳跃性变化幅度较大,模型预测稳定性较差。

4.3 迁移学习和重新训练结果对比

为了对比迁移学习和重新训练对牙形刺图像分类的效果,本文针对迁移学习准确率最高的 ResNet-34 模型,采用牙形刺图像数据集重新训

练。结果显示,训练准确率为 94.88%,低于五种迁移学习的训练准确率(表 3)。Top-1 和 Top-2 测试准确率分别为 79.51%和 90.73%,总体低于迁移学习模型测试准确率。模型收敛经历的训练轮次较多,在 40 轮次以后趋于稳定。ResNet-34 模型重新训练 Top-1 测试准确率与 ResNet-152 模型迁移学习 Top-1 测试准确率相当,也说明了对于石生物化石这种简单图像,不必追求过深的神经网络模型。

4.4 不同种牙形刺识别准确率

采用图像识别准确率较高的 ResNet-34 迁移学习网络模型对测试集 205 幅牙形刺图像识别准确率进行分析发现,中村圆柱刺 *Teridontus nakamurai*、先祖肿刺 *Cordylodus proavus*、叉开锯颚刺 *Serratognathus diversus*、变异胡安颚刺 *Juanognathus variabilis*、扁平小剑刺 *Scabbardella altipes*、短尾哈玛拉刺 *Hamarodus brevirameus*、角肿刺 *Cordylodus angulatus*、马尼托罗斯刺 *Rossodus manitouensis* 8 种牙形刺 Top-1 图像识别准确率分别为 73.3%、91.7%、95.3%、86.2%、73.1%、82.8%、84.2%、90.0%(表 4)。

叉开锯颚刺 *Serratognathus diversus* 获得了最高的预测准确率,主要原因是叉开锯颚刺 *Serratognathus diversus* 与其他种相比,特征明显,其独特的“扇形”外观易于识别,其余几种牙形刺则均为特征相似的锥形(图 5,表 5)。此外,叉开锯颚刺 *Serratognathus diversus* 的图像数据最多,也是测试准确率较高的原因。

中村圆柱刺 *Teridontus nakamurai* 和扁平小剑刺 *Scabbardella altipes* Top-1 识别准确率较低,均为 73.1%(表 4)。中村圆柱刺 *Teridontus nakamurai*

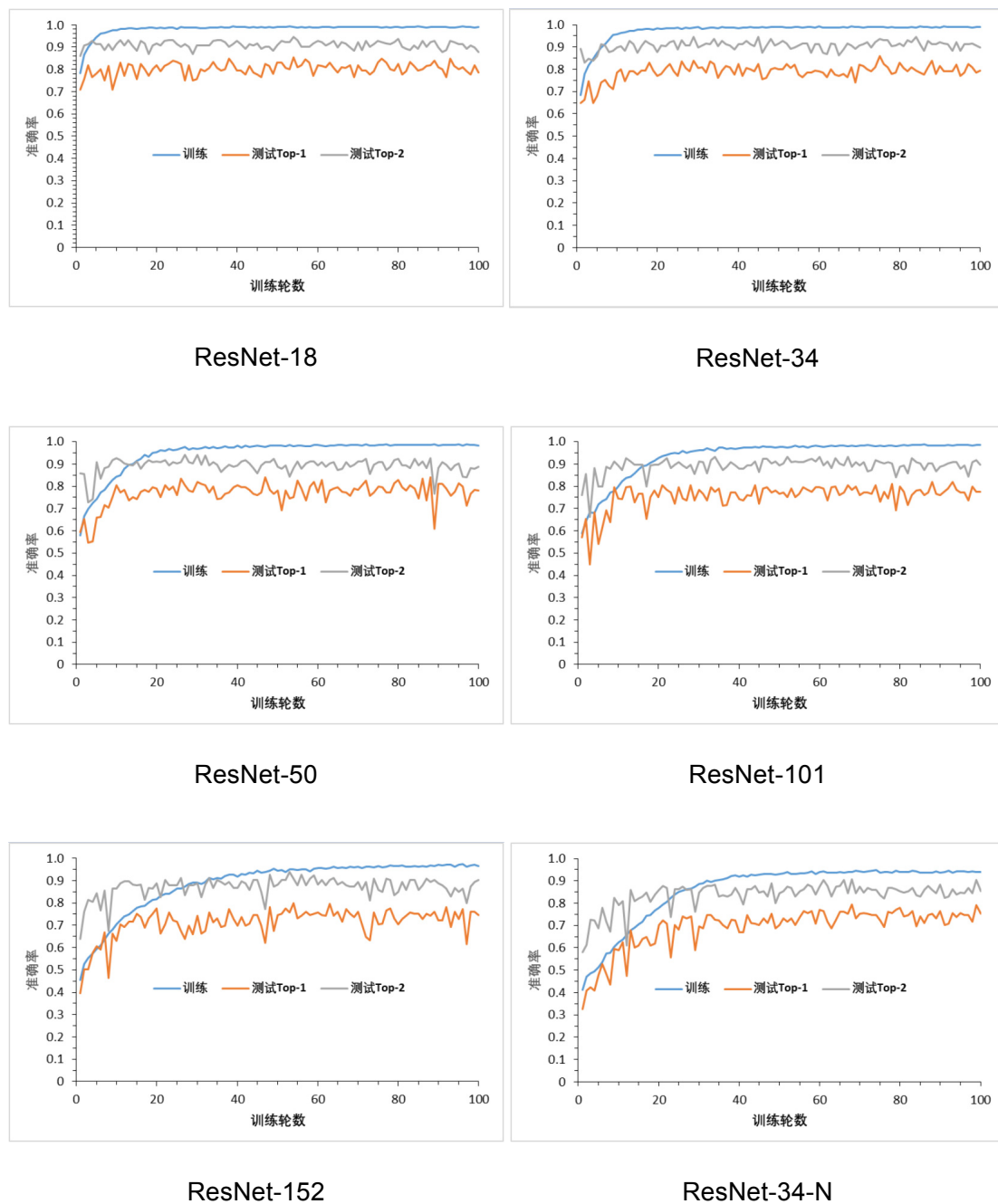


图 4 5 种网络模型迁移学习和重新训练结果

Fig. 4 Transfer learning and retraining results of five network models

图中 ResNet-18、ResNet-34、ResNet-50、ResNet-101、ResNet-152 为迁移学习的结果, ResNet-34-N 为重新训练的结果。

ResNet-18, ResNet-34, ResNet-50, ResNet-101 and ResNet-152 are the results of transfer learning, ResNet-34-N are the results of retraining.

被错误判识为先祖仲刺 *Cordylodus proavus* 的占 13.3%, 误判率最高, 主要原因是训练集中部分先祖仲刺 *Cordylodus proavus* 形貌不全, 与中村圆柱刺 *Teridontus nakamurai* 具有相似性(图 5, 表 5)。

中村圆柱刺 *Teridontus nakamurai* 被错判为扁平小剑刺 *Scabbardella altipes* 和马尼托罗斯刺 *Rossodus manitouensis* 各占 6.7%, 主要原因是三者照相角度造成的图像具有较大程度的相似性。

表 4 不同种牙形刺 ResNet-34 网络模型迁移学习识别结果(%)
Table 4 Image recognition results of different species of conodonts under transfer learning of ResNet-34 model

种名	识别结果(%)									
	A	B	C	D	E	F	G	H	SC	MC
A	73.1	13.3	0.0	0.0	6.7	0.0	0.0	6.7	70.0	80.0
B	0.0	91.7	0.0	0.0	0.0	0.0	8.3	0.0	91.7	
C	0.0	2.3	95.3	2.3	0.0	0.0	0.0	0.0	91.7	96.8
D	6.9	3.4	3.4	86.2	0.0	0.0	0.0	0.0	62.5	95.2
E	3.8	3.8	0.0	0.0	73.1	19.2	0.0	0.0	84.6	61.5
F	0.0	6.9	0.0	3.4	3.4	82.8	0.0	3.4	78.6	86.7
G	0.0	15.8	0.0	0.0	0.0	0.0	84.2	0.0	84.2	
H	5.0	0.0	0.0	0.0	5.0	0.0	0.0	90.0	87.5	91.7

注: A、B、C、D、E、F、G、H 代表牙形刺种名代号, 意义同表 1, SC 为扫描电镜图像测试准确率, MC 为体视显微镜图像测试准确率。
Note: A, B, C, D, E, F, G and H represent the names and codes of the species of conodonts, the meaning of which is the same as that in Table 1. SC is the accuracy rate of scanning electron microscope image test, while MC is the accuracy rate of stereoscopic microscope image test.

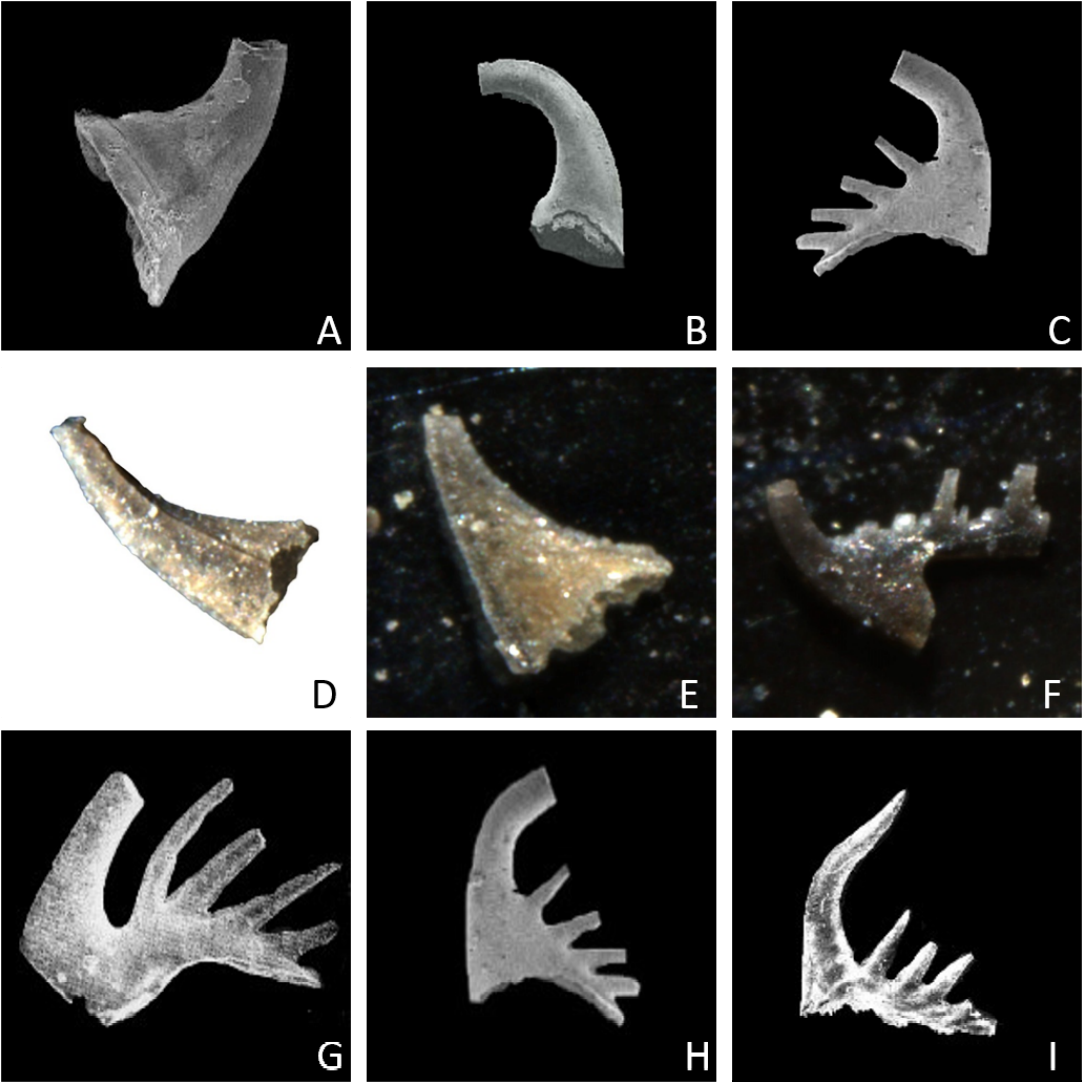


图 5 用于分析不同场景图像识别错误原因的图像(化石代号对应名称见表 5)
Fig. 5 Images used to analyze the cause of image recognition errors in different scenes (conodont names are listed in Table 5)

表 5 不同场景图像识别错误原因分析
Table 5 Analysis on the causes of image recognition errors in different scenes

场景	识别图像	对比图像	判识错误原因
1	图 5-A. 中村圆柱刺 <i>Teridontus nakamurai</i>	图 5-B. 先祖仲刺 <i>Cordylodus proavus</i> (不完整图像)	测试集中中村圆柱刺 <i>Teridontus nakamurai</i> 图像(图 5-A)不完整, 训练集先祖仲刺 <i>Cordylodus proavus</i> 图像不完整(图 5-B), 训练数据集图像较少。
2	图 5-D. 扁平小剑刺 <i>Scabbardella altipes</i>	图 5-E. 短尾哈玛拉刺 <i>Hamarodus brevirameus</i> (不完整图像)	训练集中部分图像不完整(图 5-E)。
3	图 5-G. 角肿刺 <i>Cordylodus angulatus</i>	图 5-H. 先祖仲刺 <i>Cordylodus proavus</i> (较完整图像)	两种化石特征相近, 训练数据集较小。

扁平小剑刺 *Scabbardella altipes* 被错判为短尾哈玛拉刺 *Hamarodus brevirameus* 的比例达 19.2%(表 4)。从牙形刺完整图像上看, 两种化石差异明显, 不应该有如此高的误判率。研究过程中对比了每一幅错误图片的特征, 发现相当一部分短尾哈玛拉刺 *Hamarodus brevirameus* 特征不全, 其局部特征与扁平小剑刺 *Scabbardella altipes* 极为相似(图 5, 表 5)。

另外, 角肿刺 *Cordylodus angulatus* 被误判为先祖仲刺 *Cordylodus proavus* 的比例也较高, 占 15.8%, 主要原因是两者完整化石特征的图像本来就相近, 当残缺不全时, 两者识别的准确率就会大幅降低。实际上虽然先祖仲刺 *Cordylodus proavus* 整体识别准确率较高, 为 91.7%, 但其误判的图像均归为角肿刺 *Cordylodus angulatus*。两者均是互为误判率较高的化石类型(表 4)。

从以上分析看出, 图像误判主要原因有以下几方面: 化石特征相近, 化石不完整, 照相角度没有反映化石典型特征, 训练和测试数据集较小。

4.5 扫描电镜和体视镜图像识别准确率

将 205 幅测试集图像分为体视显微镜和扫描电镜图像两类, 采用通过迁移学习训练获得的 ResNet-34 模型对两类图像进行识别, 扫描电镜和体视显微镜图像 Top-1 识别准确率分别为 70%–91.7%和 61.5%–96.8%, 平均值分别为 83.3%和 88.7%(表 4)。总体而言, 两者差别不是太大, 体视显微镜的图像识别准确率略高于扫描电镜图像。造成这种原因可能与体视显微镜为彩色成像, 图像信息更加丰富, 同时也说明高分辨率的体视显微

镜可以作为微体古生物化石照像的重要手段。

5 结 论

基于 ResNet-18、ResNet-34、ResNet-50、ResNet-101 和 ResNet-152 五种残差神经网络模型, 以 1966 幅体视显微镜和扫描电镜图像为基础进行迁移学习训练, 获得 Top-2 测试准确率分布范围为 93.17%–94.63%, 模型对牙形刺图像具有较好的识别效果。残差网络模型迁移学习训练比重新训练能获得较高的准确率、较短的训练时间, 是小样本古生物化石图像识别的重要途径。

五种残差模型试验结果表明, ResNet-34 识别取得的准确率(85.85%)最高, 说明对于特征简单的牙形刺化石, 增加网络深度并不一定能取得准确率的提升, 确定合适深度的模型不仅可以提高识别准确率, 而且可以节约计算资源。

图像误判主要由化石特征相近、化石不完整、照相角度没有反映化石典型特征、训练和测试数据集较小等原因造成。体视显微镜下牙形刺图像的识别准确率高于扫描电镜下图像识别准确率, 可以作为今后牙形刺和其他化石图像识别数据获取的重要手段。

致谢 评审专家提出了宝贵建议, 特此致谢!

参考文献 (References)

陈猛夫, 2020. 基于迁移学习的暴恐图像自动识别研究. 北京航空航天大学学报, 46: 1–7.
樊帅昌, 易晓梅, 李剑, 惠国华, 郜园园, 2020. 基于深度残差网络与迁移学习的毒覃图像识别. 传感技术学报, 33: 74–83.

- 关胤, 2019. 基于残差网络迁移学习的花卉识别系统. 计算机工程与应用, 55: 174–179.
- 黄铮, 白志强, 柴华, 2009. 利用人工神经网络方法对牙形石鉴定中若干“模糊特征”的数字化研究. 地质科技情报, 28: 94–98.
- 刘颖, 雷研博, 范九伦, 王富平, 公衍超, 田奇, 2020. 基于小样本学习的图像分类技术综述. 自动化学报: 1–20 (录用定稿). DOI: 10.16383/j.aas.c190720.
- 马中启, 朱好生, 杨海仕, 王琪, 胡燕, 2019. 基于多特征融合密集残差 CNN 的人脸表情识别. 计算机应用与软件, 36: 197–201.
- 牛新, 窦勇, 张鹏, 曹玉, 2016. 基于深度学习的光学遥感机场与飞行器目标识别技术. 大数据, 5: 54–67.
- 王志浩, 祁玉平, 吴荣昌, 2011. 中国寒武纪和奥陶纪牙形刺. 合肥: 中国科学技术大学出版社, 3: 397–580.
- 魏存超, 2017. 基于卷积神经网络的医学图像分类的研究. 硕士学位论文. 哈尔滨: 哈尔滨工业大学, 37–40.
- 薛迪秀, 2017. 基于卷积神经网络的医学图像癌变识别研究. 博士学位论文. 合肥: 中国科学技术大学, 84–85.
- 殷俊, 杨万扣, 2018. 基于局部表示的分类方法及其人脸识别应用. 计算机工程与科学, 40: 500–506.
- 张放, 熊志东, 马冬晨, 蒋飞虎, 杨静, 2010. 牙形石分析鉴定方法. 北京: 石油工业出版社, 1–18.
- Agematsu S, Sashida K, Salyapongse S, Sardud A, 2007. Ordovician conodonts from the Satun area, southern Peninsular Thailand. *Journal of Paleontology*, 81: 19.
- Bergström S M, Löfgren A, Grahm Y, 2004. The stratigraphy of the Upper Ordovician carbonate mounds in the subsurface of Gotland. *GFF*, 126: 289–296.
- Chen Meng-fu, 2020. Research on automatic recognition model for terrorism related image based on transfer learning. *Journal of Beijing University of Aeronautics and Astronautics*, 46: 1–7 (in Chinese with English abstract).
- Erlström M, Ahlberg P, Löfgren A, 2001. Lower Palaeozoic stratigraphy at Lyby and Tängelsås, central Scania, southern Sweden. *GFF*, 123: 7–14.
- Fan Shuai-chang, Yi Xiao-mei, Li Jian, Hui Guo-hua, Hao Yuan-yuan, 2020. Toadstool image recognition based on deep residual network and transfer learning. *Chinese Journal of Sensors and Actuators*, 33: 74–83 (in Chinese with English abstract).
- Ferretti A, 1998. Late Ordovician Conodonts from the Prague Basin, Bohemia. *Proceedings of the Sixth European Conodont Symposium (ECOS VI)*, *Acta Palaeontologica Polonica*, 58: 123–139.
- Ferretti A, Barnes C, 1997. Upper Ordovician conodonts from the Kalkbank limestone of Thuringia, Germany. *Palaeontology*, 40: 15–42.
- Guan Yin, 2019. Flower species recognition system based on residual network transfer learning. *Computer Engineering and Applications*, 55: 174–179 (in Chinese with English abstract).
- He Kai-ming, Zhang Xiang-yu, Ren Shao-qing, Sun Jian, 2016. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 770–778.
- Huang Zheng, Bai Zhi-qiang, Chai Hua, 2009. Identification of the 'ambiguity features' of the conodont by the artificial neural network. *Geological Science and Technology Information*, 28: 94–98 (in Chinese with English abstract).
- Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K, 2014. DenseNet: Implementing efficient ConvNet descriptor pyramids. Report of University of California, Berkeley, 1–11.
- Ji Zai-liang, Barnes C, 1994. Lower Ordovician conodonts of the St. George Group, Port au Port Peninsula, western Newfoundland, Canada. *Palaeontographica Canadiana*, 11: 1–149.
- Ji Zai-liang, Barnes C, 1996. Uppermost Cambrian and lower Ordovician conodont biostratigraphy of The Survey Peak Formation (Ibexian/Tremadoc), Wilcox Pass, Alberta, Canada. *Journal of Paleontology*, 70: 871–890.
- Johnston D I, Barnes C, 2000. Early and Middle Ordovician (Arenig) conodonts from St. Pauls Inlet and Martin Point, Cow Head Group, western Newfoundland, Canada. 2. Systematic paleontology. *Geologica et Palaeontologica*, 34: 11–87.
- Kaljo D, Heinsalu H, Mens K, Puura I, Viira V, 1988. Cambrian–Ordovician boundary beds at Tõnismägi, Tallinn, North Estonia. *Geological Magazine*, 125: 457–463.
- Krizhevsky A, Sutskever I, Hinton G, 2012. ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25: 1–9.
- Landing E, 1976. Early Ordovician (Arenigian) conodont and graptolite biostratigraphy of the Taconic allochthon, eastern New York. *Journal of Paleontology*, 50: 614–646.
- Landing E, Westrop S, Hernick L, 2003. Uppermost Cambrian–Lower Ordovician faunas and Laurentian Platform sequence stratigraphy, eastern New York and Vermont. *Journal of Paleontology–Journal of Paleontology*, 77: 78–98.
- Landing E D, Westrop S R, Keppie J D, 2007. Terminal Cambrian and lowest Ordovician succession of Mexican West Gondwana: biotas and sequence stratigraphy of the Tiñu Formation. *Geological Magazine*, 144: 909–936.
- Landing E D, Westrop S R, Kröger B, English A M, 2010. Left behind–delayed extinction and a relict trilobite fauna in the Cambrian–Ordovician boundary succession (east Laurentian platform, New York). *Geological Magazine*, 148: 529–557.
- Lecun Y, Bottou L, Bengio Y, Haffner P, 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86: 2278–2324.
- Lee B S, 2004. Late Cambrian conodonts from Cheung-san, Kangwon Province, Korea. *Alcheringa: An Australasian Journal of Palaeontology*, 28: 53–64.
- Liu Ying, Lei Yan-bo, Fan Jiu-lun, Wang Fu-ping, Gong Yan-chao, Tian Qi, 2020. Survey on image classification technology based on small sample learning. *Acta Automatica Sinica*, 1–20 (accepted, in Chinese with English abstract). DOI: 10.16383/j.aas.c190720.
- Löfgren A, 1996. Lower Ordovician conodonts, reworking, and biostratigraphy of the Orreholmen quarry, Västergötland, south-central Sweden. *GFF*, 118: 169–183.
- Löfgren A, 1997. Reinterpretation of the lower Ordovician conodont apparatus *Paroistodus*. *Palaeontology*, 40: 913–929.
- Löfgren A, Viira V, 2007. A *Cordylodus* fauna from the SW Gulf of

- Bothnia, Sweden. GFF, 129: 117–121.
- Löfgren A, Viira V, Mens K, 2005. Conodont biostratigraphy and sedimentary history in the upper Tremadoc at Uuga, Cape Pakri, NW Estonia. GFF, 127: 283–293.
- Ma Zhong-qi, Zhu Hao-sheng, Yang Hai-shi, Wang Qi, Hu Yan, 2019. Facial expression recognition based on multi-feature fusion dense residual CNN. *Computer Applications and Software*, 36: 197–201 (in Chinese with English abstract).
- Nicoll R S, 1990. The genus *Cordylodus* and a latest Cambrian-earliest Ordovician conodont biostratigraphy. *Journal of Australian Geology & Geophysics*, 11: 529–558.
- Niu Xin, Du Yong, Zhang Peng, Cao Yu, 2016. Airport and flight recognition on optical remote sensing data by deep learning. *Big Data Research*, 5: 54–67 (in Chinese with English abstract).
- Nowlan G S, 1981. Some Ordovician conodont faunules from the Miramichi Anticlinorium, New Brunswick. *Geological Survey of Canada*, 345: 1–35.
- Orndorff R C, 1988. Latest Cambrian and earliest Ordovician conodonts from the Conococheague and Stonehenge limestones of northwestern Virginia, U.S. *Geological survey Bulletin*, 1837: A1–A18.
- Parsons B, Clark D, 1996. Conodonts and the Cambrian-Ordovician boundary in Wisconsin. *Journal of Paleontology*, 70: 871–890.
- Percival I, Zhen Yong-yi, 2006. Late Cambrian-Early Ordovician conodont faunas from the Koonenberry Belt of western New South Wales. *Memoirs of the Association of Australasian Palaeontologists*, 32: 267–285.
- Pohler S M L, Orchard M J, 1990. Ordovician conodonts from the western Canadian Cordillera. *Geological Survey of Canada*, Paper 90–15: 1–37.
- Pyle L J, Barnes C R, 2001. Conodonts from the Kechika Formation and Road River Group (Lower to Upper Ordovician) of the Cassiar Terrane, northern British Columbia. *Canadian Journal of Earth Sciences*, 38: 1387–1401.
- Pyle L J, Barnes C R, 2002. Taxonomy, evolution, and biostratigraphy of conodonts from the Kechika Formation, Skoki Formation and Road River Group (Upper Cambrian to Lower Silurian), northeastern British Columbia. Ottawa: NRC Research Press, 1–227.
- Pyle L J, Barnes C R, 2007, Mcanally L M. Conodont biostratigraphy of the latest Cambrian - Early Ordovician upper McKay Group, southeastern British Columbia. *Canadian Journal of Earth Sciences*, 44: 1713–1740.
- Seo K S, Lee H Y, Ethington R, 1994. Early Ordovician conodonts from the Dumugol Formation in the Baegunsan syncline, eastern Yeongweol and Samcheog areas, Kangweon-Do, Korea. *Journal of Paleontology*, 68: 599–616.
- Simonyan K, Zisserman A, 2015. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations*, 1–11.
- Stouge S, Bagnoli G, Albani R, 1985. Lower Ordovician conodonts from Washington Land, western North Greenland. *Bulletin of the Geological Society of Denmark*, 33: 261–272.
- Stouge S, Rasmussen J, 1996. Upper Ordovician conodonts from Bornholm and possible migration routes in the Palaeotethys Ocean. *Bulletin of the Geological Society of Denmark*, 43: 54–67.
- Szegedy C, Liu Wei, Jia Yang-qing, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A, 2014. Going deeper with convolutions. 1–12.
- Taylor J F, Kennedy D J, Miller J, Repetski J E, 1991. Uppermost Cambrian slope deposits at Highgate Gorge, Vermont: a minor miscorrelation with major consequences for conodont- and trilobite-based chronocorrelation. *Journal of Paleontology*, 65: 855–863.
- Tolmacheva T J, Degtyarev K E, Samuelsson J, Holmer L E, 2008. Middle Cambrian to Lower Ordovician faunas from the Chingiz Mountain Range, central Kazakhstan. *Alcheringa: An Australasian Journal of Palaeontology*, 32: 443–463.
- Viira V, Löfgren A, Mägi S, Wickström J, 2002. An early to middle Ordovician succession of conodont faunas at Mäekalda, northern Estonia. *Geological Magazine*, 138: 699–718.
- Wang Ya-qing, Yao Quan-ming, Kwok J T, Ni L M, 2019. Generalizing from a Few Examples: A Survey on Few-Shot Learning. *ACM Computer Survey*, 1: 1–34.
- Wang Zhi-hao, Qi Yu-ping, Wu Chang-rong, 2011. Cambrian and Ordovician Conodonts in China. Hefei: China University of science and Technology Press. 3, 397–580 (in Chinese with English abstract).
- Wei Cun-chao, 2017. Research on medical image classification method based on convolutional neural network. Master's Thesis, Harbin: Harbin Institute of Technology. 37–40 (in Chinese with English abstract).
- Xue Di-xiu, 2017. Research on cancer image recognition based on convolutional Neural Networks. Doctoral dissertation. Hefei: University of Science and Technology of China, 84–85 (in Chinese with English abstract).
- Yin Jun, Yang Wan-he, 2018. Local representation based classification and its application in face recognition. *Computer Engineering & Science*, 40: 500–506 (in Chinese with English abstract).
- Zhang Fang, Xiong Zhi-dong, Ma Dong-chen, Jiang Fei-hu, Yang Jing, 2010. Analysis and authenticates method for conodonts. Beijing: Petroleum Industry Press. 1–18 (in Chinese).
- Zhen Yong-yi, Nicoll R S, 2009. Biogeographic and biostratigraphic implications of the *Serratognathus bilobatus* fauna (Conodonta) from the Emanuel Formation (Early Ordovician) of the Canning Basin, Western Australia. *Records of the Australian Museum*, 61: 1–30.
- Zhen Yong-yi, Wang Zhi-hao, Zhang Yuan-dong, Bergström, Percival I G, Cheng Jun-feng, 2011. Middle to late Ordovician (Darriwilian-Sandbian) conodonts from the Dawangou section, Kalpin area of the Tarim Basin, northwestern China. *Records of the Australian Museum*, 63: 203–266.
- Zhen Yong-yi, Webby B D, Barnes C R, 1999. Upper Ordovician conodonts from the Bowan Park succession, central New South Wales, Australia. *Geobios*, 32: 73–104.

(责任编辑: 王永栋)