一种基于深度学习的青铜器铭文识别方法

李文英 [1] 曹 斌 [1] 曹春水 [2,3] 黄永祯 [2,3]

摘 要 考古出土的青铜器铭文是非常宝贵的文字材料, 准确、快速地了解其释义和字形演变源流对考古学、历史学和语言学研究均有重要意义. 青铜器铭文的辨识需要综合文字的形、音、义进行研究, 其中第一步也是最重要的一步就是分析文字的形体特征. 本文提出一种基于两阶段特征映射的神经网络模型来提取每个文字的形体特征, 最后对比目前已知的文字研究成果, 如《古文字类编》、《说文解字》, 得出识别的结果. 通过定性和定量的实验分析, 我们发现本文提出的方法可达到较高的识别精度. 特别地, 在前 10 个预测类别中 (Top-10) 准确率达到了 94.2 %, 大幅缩小了考古研究者的搜索推测空间, 提高了青铜铭文识别的效率和准确性.

关键词 模式识别, 青铜器铭文, 文字识别, 深度学习, 深度卷积神经网络

引用格式 李文英, 曹斌, 曹春水, 黄永祯. 一种基于深度学习的青铜器铭文识别方法. 自动化学报, 2018, **44**(11): 2023−2030

DOI 10.16383/j.aas.2018.c180152

A Deep Learning Based Method for Bronze Inscription Recognition

LI Wen-Ying[1] CAO Bin[1] CAO Chun-Shui[2,3] HUANG Yong-Zhen[2,3]

Abstract Bronze inscriptions from archaeology are very valuable text materials. Accurate and rapid understanding of their meaning and shape evolution is important for archeology, history and linguistics. It is necessary to combine characters shape, phonology and meaning for recognition of bronze inscription, wherein the first and also the most important step is to analyze shapes of bronze inscriptions. In this paper, we present a bronze inscription analysis method based on convolutional neural network (CNN) with two-phase feature mapping. We first extract the bronze inscriptions by image acquisition, and then, by comparing with the currently known character research results, e. g., "Ancient Chinese Character Type Series" and "Shuo Wen Jie Zi", we obtain the recognition results. Through qualitative and quantitative experimental analyses, we find that the proposed method achieves high recognition accuracy. Specifically, we achieve 94.2 % accuracy for the Top-10, greatly reducing the space of archaeological search and improving the efficiency and accuracy of bronze inscription recognition.

Key words Pattern recognition, bronze inscription, character recognition, deep learning, convolutional neural network (CNN)

Citation Li Wen-Ying, Cao Bin, Cao Chun-Shui, Huang Yong-Zhen. A deep learning based method for bronze inscription recognition. *Acta Automatica Sinica*, 2018, **44**(11): 2023−2030

青铜器是商周时代社会祭祀和战争的主要工具, 通过对青铜器的造型、工艺、纹饰、铭文等的研究, 可以深入了解古代社会的政治经济文化. 其中, 铸有铭文的青铜器, 因其遗留了古文字材料, 更是古文字学、语言学、考古学及古代史研究的重要研究对象. 目前保存的青铜器中, 铸有铭文的有上万件[1], 其中西周至春秋时期, 可供研究的古文字材料主要都是青铜器铭文, 可见青铜器铭文无论从其本身的古文字学意义, 还是其所著述的内容而言, 都具有极大的研究价值. 释读古文字, 一般先识其形, 通其音, 然后识其义[2]. 古文字形体笔画复杂, 变化繁多, 且古文字是不统一的, 各地方各时代都有不同的写法, 想要释读古文字, 须得熟知古文字最初的形状, 如象形字、形声字、假借字, 以及指事、会意、转注等构字方法[3], 还须参证甲骨文、战国文字, 熟读《说文解字》等这种古文字字典, 难度很大, 有学者甚至将其归之为 "绝学以内". 简而言之, 释读青铜器铭文要求研究者具有广博的知识基础和训练, 是一项极具挑战的任务.

随着计算机技术特别是图像识别技术的高速发

收稿日期 2018-03-19 录用日期 2018-04-04
Manuscript received March 19, 2018; accepted April 4, 2018
国家重点基础研究发展计划 973 计划 (2016YFB1001000), 国家自然科学基金 (61525306, 61633021, 61420106015), 教育部人文社会科学研究青年基金项目 (18YJC780001) 资助
Supported by National Basic Research Program of China (973 Program) (2016YFB1001000), National Natural Science Foundation of China (61525306, 61633021, 61420106015), and Humanities and Social Sciences Research Youth Fund Project, Ministry of Education (18YJC780001)
本文责任编委 刘成林
Recommended by Associate Editor LIU Cheng-Lin
1. 中国人民大学历史学院 北京 100872 2. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190 3. 银河水滴科技 (北京) 有限公司 北京 100190
1. School of History, Renmin University of China, Beijing 100872 2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 3. Watrix Technology Co. Ltd., Beijing 100190

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

展, 以图像识别技术为基础, 结合古文字特别是甲骨文的研究基础, 进行古文字识别已成为可能. 本文提出一种新的青铜器铭文识别研究方法, 应用图像识别技术来识别青铜器铭文, 以期望能更准确、更广泛地识别出青铜器铭文, 或得出铭文的最大相似度文字, 为进一步释读文字提供基础. 其主要方法是: 采集青铜器铭文文字图像, 同时也采集《古文字类编》[4] 中甲骨文、金文以及《说文解字》中相同文字的图像, 提取出每个古文单字, 然后采用深度卷积神经网络 (Convolutional neural network, CNN) 提取图像特征来识别青铜器铭文, 得出识别的结果或者推荐结果.

本文的主要贡献有以下几个方面: 1) 首次提出了一种基于 CNN 的青铜器铭文识别模型, 采用两阶段距离约束对应, 相应地设计了两阶段特征映射的神经网络模型来提取每个文字的形体特征, 最后对比目前已知的文字研究成果, 取得了令人满意的效果, 大幅缩小了考古研究者的搜索推测空间, 有效提高了考古过程青铜铭文识别的效率和准确性; 2) 构建了首个数字化的青铜器铭文数据库, 包含 77 个古文字共 2 102 张图片. 该数据库对采用深度学习模型进行青铜铭文识别的技术路线提供了一个很好的试验田, 也为将来采集更大规模数据库奠定了基础.

本文的组织架构如下: 首先, 在第 1 节简单介绍了图像识别在古文字识别方面的研究现状; 然后, 在第 2 节分叙古文字库的搭建和铭文识别模型的设计; 并在第 3 节用定量和定性的实验分析了利用古文字偏旁分析法优化铭文识别的模型效果; 最后, 总结该方法可改进之处以及应用前景.

1 研究现状

青铜器铭文研究悠久, 据史书记载最早的铭文考释者是西汉晚期的张敞, 至北宋即有专书著录, 描录文字并加以考释, 或编为字典. 近代以来, 罗振玉和王国维参证甲骨文字研究铭文, 郭沫若编成《两周金文辞大系》, 综合两周青铜器铭文, 研究两周社会的政治经济文化. 在铭文考释方面, 容庚参照甲骨文字, 编录《金文编》, 杨树达的《积微居金文说》通释铭文, 中国社会科学院考古研究所编纂《殷周金文集成》, 张亚初在《殷周金文集成引得》中提出 12 种铭文考释之法[5] 等. 但商周青铜器金文仍有近四成未辨识, 有待考释和解读的字还有近三千个; 此外, 青铜器铭文与甲骨文的对照研究, 成果也尚属有限[2].

当前, 多数研究仍以单个青铜器或某个考古地出土的青铜器物为对象进行铭文的考释、集释或补释为主要对象, 研究范围较小而局限, 如刘源的 "逨盘铭文考释"、连劭名的 "眉县杨家村窖藏青铜器铭文考述"、祝振雷的《安徽寿县蔡侯墓出土青铜器铭文集释》、胡长春、吴劲松的 "近十年来 (2001–2010) 的殷周青铜器铭文研究综述"、吴劲松的《近十年新出殷周青铜器铭文的整理與研究》等; 在研究方法上仍未有创新, 多数研究者一般先引用王国维、杨树达、唐兰、郭沫若、容庚、陈梦家、李学勤、裘锡圭等前辈学者的研究观点, 而后佐证《说文》、《诗经》、《尔雅》、《吕氏春秋》、《周易》等经史典籍, 最后从释义、语法、构字法上进行简要分析, 以得出自己的解释.

在计算机识别古文字领域, 目前多数的研究论文为甲骨文相关的, 如周新纶的 "甲骨文计算机识别方法研究"[6]、李峰的 "甲骨文自动识别的图论方法"[7]、顾绍通的 "基于拓扑配准的甲骨文字形识别方法"[8]、吕肖庆的 "一种基于图形识别的甲骨文分类方法"[9] 等, 从研究成果来看, 一般为应用了计算机图形、图形拓扑等方法, 有一定识别效果, 但还难以达到满意结果. 在应用最新的图像识别技术方面, 有王嘉梅 "基于图像分割的古彝文字识别系统研究"[10]、孙华的 "基于多特征融合 SVM 的古汉字图像识别研究"[11]、孙莹莹的 "基于混合核 LS-SVM 的古汉字图像识别"[12] 等, 它们应用了图像分割、支持向量机等图像识别技术, 识别古汉字, 其研究对象为古彝文或古籍记录的古文字, 尚不知样本采用规模 (比如文献 [11]) 或者采用的样本规模较小 (文献 [12] 仅用 20 个字共计 140 个样本进行测试), 算法虽有效且准确率不高 (文献 [12] 给出的最高准确率为 73 %), 此外这些研究均未结合古文字研究成果, 未针对特定的考古材料进行研究, 因此研究成果尚无法有效应用于考古.

在青铜器铭文识别方面, 尚未发现有针对性的相关论文, 仅发现陕西省考古研究院吴镇锋课题组开发的《金文通检》系统, 利用输入文字检索和查阅青铜器铭文资料, 尚未做到自动识别和图像识别.

图像识别技术是人工智能的一个重要领域. 近年来, 深度学习的兴起让图像识别技术越来越实用, 特别地, 卷积神经网络 (CNN) 被广泛使用. CNN 最早可以追溯到 20 世纪 80 年代末, 其创建是基于 20 世纪 60 年代早期的人工神经网络 (Artificial neural network, ANN) 和多层感知机 (Multi-layer perceptron, MLP). 图形处理器 (Graphics processing unit, GPU) 崛起以及 GPU 处理重复性任务的速度使得 CNN 开始大量投入使用. 2012 年, 计算机视觉智能领域的工作取得了重大的飞跃, Krizhevsky[13] 使用 Deep CNN 赢得了 ImageNet[14] 竞赛, 成为了基于深度学习图像识别技术发展的里程碑事件. 经过近几年的研究发展, Deep CNN 在目标识别、目标检测、目标分割等任务上都有非常出色的表现, 在 ImageNet 上, Deep CNN 对普通物体的识别已经超

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

越了人类的表现[15], 特别地, 深度学习在文字识别方面也有了长足的进步[16−17]. 相关研究表明[15, 18−20], Deep CNN 是一个大型的复杂神经网络, 经过海量数据训练后, Deep CNN 内部的神经元学会了表达各种具有抽象语义的概念, 比如各类目标物的局部肢体, 各种各样的表征纹理等, 甚至某些神经元能够表达完整的目标物. 这些抽象概念分布在神经网络不同层中, 并最后形成了对目标物整体的感知. Deep CNN 具有强大的特征归纳能力, 能够抽取不同目标的本质特征并进行准确的目标分类.

2　研究方法

2.1　研究思路

考古学及古文字学的研究发现, 古文字字体的演化过程是有迹可循的, 古文字字体形态和书体风格在各个时期的演化过程中是保留有一些特征和风格的. 如图 1 所示, 我们以 “保” 字的各种演化变体为例子.

图 1　“保” 字的各种演化变体 (包括甲骨文、青铜器铭文、篆书等)

Fig. 1　Various evolutionary shapes of character “保” (including oracle-bone, bronze inscription, seal character, etc.)

从图 1 可以看出, “保” 字的演化变体在形态上不同, 但都保持有一种内在的特征一致性, 尤其是其中的偏旁 “人”, 如图 2 所示, 很明显具有相似的特征.

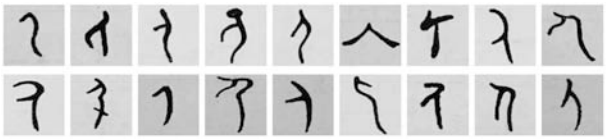

图 2　单人旁的不同形态

Fig. 2　Different shapes of character component “人”

再观察 “女” 字 (图 3) 与 “妇” 字、“好” 字 (图 4), 可以看出作为偏旁部首的 “女”, 在以它为基础构成的字中都具有显著的一致性.

上述研究过程属于古文字偏旁分析法, 受这种研究思路的启发, 我们将 Deep CNN 用于古文字识别的任务: 通过两阶段特征映射让 Deep CNN 去学习古文字整体或部分 (如偏旁) 各种变体的一致性特征, 然后将训练好的 Deep CNN 用于识别模型没有见过的其他变体. 模型设计细节将在后文展开描述.

图 3　“女” 字的不同形态

Fig. 3　Different shapes of character “女”

图 4　“妇” 字和 “好” 字的不同形态

Fig. 4　Different shapes of character “妇” and “好”

2.2　青铜铭文数据库

为了能够量化分析本文所提出的基于 Deep CNN 的古文字推测方法, 我们收集归纳并整理出了一个以青铜器铭文为主的小规模古文字库, 包括商周时代的甲骨文, 商周时代的青铜器铭文, 商周时代的陶器、石刻等文字以及《说文解字》所录的篆书、古籀文. 每一种字体或字形的图片为一个古文字的变体, 文字图片主要来源于高明、涂白奎的《古文字类编》, 部分为青铜器铭文拓片图像. 考虑深度学习模型的需要, 笔者从手头资料中随机选取样本较多、尤其是青铜器铭文较多的字构成字库, 此外考虑青铜器铭文的内容主要涉及分封诸侯、朝觐王室、祭祀典礼、方国征伐、胜利纪功、世官尊容、家族的祭享等, 笔者还挑选在青铜器铭文中常用的字构成字库. 截至目前, 已搭建了包含 77 个古文字的图片库, 一共包含 2 102 张图片, 平均每个字有 27 个变体, 其中青铜器铭文 840 张, 甲骨文 483 张, 陶器石刻等 678 张,《说文》101 张, 部分字库图片如下图 5 所示.

我们将 77 个古文字列举如下图 6 所示.

2.3　模型设计

在图像识别领域, 神经网络的深度至关重要. 研究表明, 随着网络深度的增加, 神经网络的学习能力和模式表达能力会不断增强. 比如从掀起深度学习热潮的 Alex Net 开始到 VggNet[20], GoogleNet[21] 再到 ResNet[22] 等模型, 网络越来越深, 性能也越来越好. 另外, 我们也注意到在大型图片数据库 ImageNet 领先的团队都利用了很深的模型, 特别地, ResNet 被广泛使用. ResNet 采用了跳跃连接的网络结构, 它有效地解决了超深网络在优化学习时梯

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

图 5 字库图片示例

Fig. 5 Example images of the character database

衆	中	人	伐	侯	元	公	刀	刺	卿
古	右	名	周	喪	大	好	子	尊	市
後	御	會	父	良	長	丘	事	令	何
保	克	兵	初	北	友	可	司	君	唯
国	天	姬	宗	尚	年	得	戒	母	祀
貞	黄	丞	亡	以	余	僕	兒	凡	利
千	取	史	吉	吴	商	夏	女	婦	射
左	弗	徙	文	比	祭	辟			

图 6 77 个古文字库

Fig. 6 Ancient character database with 77 characters

度消失的问题, 并且大幅增加了网络表达图像模式的容量. 这使得采用这一结构的网络深度可以达到 1 000 层, 人工神经网络的强大表达能力被充分的发挥出来. 具体地, ResNet 与传统网络相比较, 其加入了 $y = x$ 层 (恒等映射层), 其主要作用是使得网络随着深度的增加而表达能力不退化, 并且还具有较好的收敛效果. 因此, ResNet 通常可以非常深, 而它取得的识别效果在各大数据库评测中目前也是很好的. 除此之外, ResNet 还具有泛化能力强, 结构简单, 训练收敛迅速等特征.

从我们构建的古文字库可以看出, 每个字平均有 27 种变体, 每个字的形态不固定意味着神经网络要表达的模式不固定, 这就要求神经网络需要有强大的学习容量, 能够细粒度地表达各种不同但本质又存在类似的模式. 进一步, 在我们设定的任务中, 我们希望让网络学习每个字的部分形态特征, 然后期望它对没有见过的新的形态能做出正确的预测. 换而言之, 古文字各种演变体具有 "万变不离其宗" 的特点, 我们希望网络能在 "万变" 中学会求同存异, 归纳推理, 这就要求网络需要强大的抽象映射能力和泛化能力, 使图片像素空间经过不断映射最终走向更抽象更高层的语义空间.

基于以上分析, 我们认为 ResNet 将比较适合完成古文字的识别任务, 同时为了进一步增强网络的泛化能力, 我们精心设计了一套网络模型学习的策略. 下图 7 是我们方法的示意图:

如图 7 所示, 由于总的古文字图片数量较少, 我们采用了主干网络包含 18 个卷积网络层的相对比较轻量级的 ResNet 神经网络, 并且我们进一步限制了卷积层的每一层的通道数最多为 16.

模型设计核心思想: 为了获得更好的模型泛化能力, 我们希望网络对每一类的古汉字提取的特征能够满足以下两个特点: 1) 同一类的汉字特征在特征空间尽可能聚在一起, 模型能够完成这样的映射,

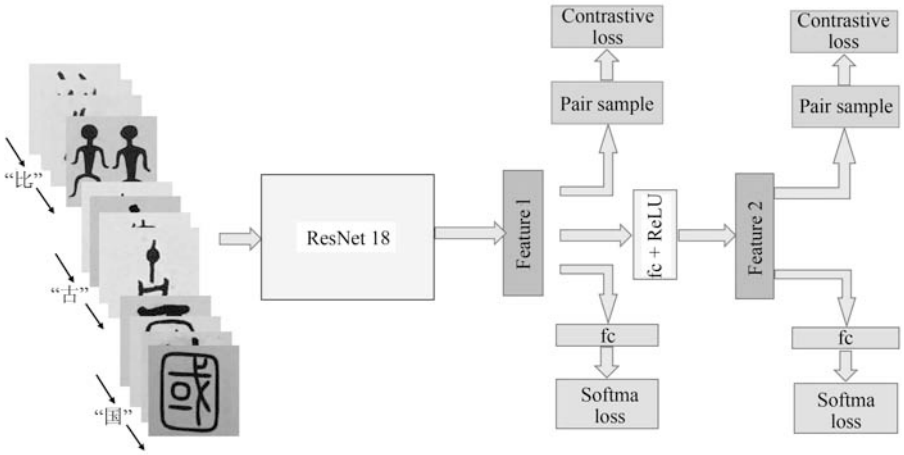

图 7 基于 18 层 ResNet 的古文字识别模型示意

Fig. 7 Pipeline of ancient character recognition based on 18-level ResNet

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

需要尽可能提取同一类汉字的内部共通性; 2) 不同类的汉字特征在特征空间的聚簇彼此尽可能离得远一些, 这样不同汉字之间能留有足够宽的空间给分类器, 使得分类器对于新来的未见过的汉字具有更高的概率将其分对. 基于以上动机, 我们设计了如图 7 模型右侧两阶段特征映射的结构. 训练以 Batch 的方式进行. 首先我们把 ResNet18 最后一层的特征 Feature1 作为基础进行第一阶段的特征映射: 以 Feature1 为输入, 一方面我们通过一层全连接层 (fc) 结合 Softmax loss 训练第一个分类器, 另一方面我们根据一个 Batch 内的汉字类别标签来抽取特征对, 结合 Contrastive loss 实现距离约束, 比如一个 Batch 内取了 3 个汉字: "比"、"古"、"国", 每个汉字有 4 张不同的图片, 则每个汉字内部可以采样 6 个正对, 共 24 个正对, 而负对的数量要多很多, 我们采取困难负对挖掘的做法使得正负对比例为 1:1.5, 以此来实现训练. 为了进一步达到映射同类特征的目的, 我们更进一步做第二阶段的特征映射, 做法是将 Feature1 通过一层简单的非线性全连接映射 (fc + ReLU) 转换到 Feature2, 然后对 Feature2 同样做类似 Feature1 上的分类和距离 Loss 约束. 4 个损失同时优化, 来实现特征表达的映射再映射, 如图 8 所示.

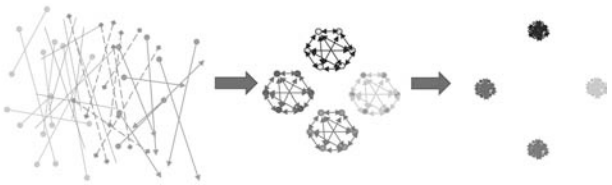

图 8　两阶段映射示意 (第一个 Loss 有能力把杂乱的原始数据聚类得比较好; 第二个 Loss 进一步聚类数据)

Fig. 8　Demonstration of two-stage mapping (The first loss has the ability to originally cluster the messy raw data and the second further clusters the data.)

3　实验分析

3.1　实验设置

我们采用图 7 的模型结构进行训练, 基于 mxnet 平台使用梯度下降法进行模型参数学习. 我们将所有文字的图片尺寸归一化到宽高都是 128 个像素, 并且由彩色图片转换成灰度图片, 我们同时计算了所有图片的均值图片. 一张图片减去均值图片后送入神经网络进行特征表达和学习.

而对于训练和评测, 将每个文字的变体图片中的 60% 作为训练集, 共 1 290 张图片, 而剩余的 40% 作为测试集, 共 812 张. 即每个文字用 60% 的变体图片进行学习, 让模型归纳总结出每个文字的内在一致性特征, 然后让模型对没有见过的 40% 测试集进行识别与预测. 我们模型在所建数据库上迭代学习了 100 轮, 然后分析识别结果.

由于训练图片数量较少, 为了避免过拟合, 我们采取了如下措施: 1) 精简 ResNet18 卷积核的个数, 使得任何一层不能超过 16 个卷积核. 同时中间层全连接层的节点数设置为 64 或者 128; 2) 基础 ResNet18 网络结构的预训练. 我们收集了 8 000 个印刷体简体汉字的 80 000 张图片以分类任务为目标来预训练我们的 ResNet18, 在我们的古汉字数据集上以 0.0001 的学习率微调, ResNet18 外的其他可学习层的学习率设置为 0.001; 3) 我们同时采用 Dropout 和较大的 Weight decay 来进一步降低过拟合.

3.2　实验结果

模型学习完成后, 我们在测试集上对它进行识别准确率的测试. 如前所述, 测试集一共 870 张图片, 77 个字. 我们将每个字设为一个类别, 共 77 个类别, 给定一张图片, 经模型测试, 会给出这张图片属于每个类的准确率, 定义 k 为类别数, 一张图片的真实类别在前 k 个预测类别中的准确率为 Top-k. 我们在表 1 中报告 Top-1、Top-3、Top-5、Top-8、Top-10 这 5 种情况下基准分类器、分类器 I 和分类器 II 的识别准确率. 其中基准分类器指的是: 保留图 7 中的第一阶段分类器, 删除两阶段距离约束和第二阶段分类器. 分类器 I 指的是: 保留图 7 中的第一阶段分类器和第一阶段距离约束. 分类器 II 指的是: 保留图 7 中的第二阶段分类器和两阶段距离约束.

表 1　测试集的识别准确率

Table 1　Recognition accuracy in the testing dataset

分类器结果对比	Top-1	Top-3	Top-5	Top-8	Top-10
基准分类器	57.1%	73.7%	85.8%	89.6%	92.7%
分类器 I	57.7%	74.9%	86.2%	90.5%	93.6%
分类器 II	58.3%	76.1%	87.1%	91.4%	94.2%

从基准分类器和分类器 I 的实验结果对比可以得出: 引入样本对之间的距离约束是有效的, 能够提高模型的泛化能力. 从基准分类器 I 和基准分类器 II 的实验结果对比可以看出: 第二个分类器的分类效果明显好于第一个分类器的分类效果, 说明我们提出的两阶段映射是有效的, 进一步的特征映射有助于提升模型的泛化能力.

综合来看, 模型有效学习到了各个文字变体内在的一致性特征. 在 Top-1 测试条件下, 模型判别的最高正确率为 58.3% (随机判别 Top-1 的正确率仅为 $1/77 = 1.3\%$), 而在 Top-10 的测试条件下, 准确率可以做到 94.2% (随机判别 Top-10 的正确率

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

仅为 10/77 = 13.0 %), 这表明, 当古文字研究者碰到新的文字变体时, 本文提出的方法能够以很高的概率将该文字的识别结果限定在大约 10 个已知文字内, 这将大幅缩减研究者的搜索推测空间, 有效提高效率和准确性.

3.3 实例说明

为了形象地理解整个方法的过程, 我们举一个例子加以说明. 图 9 展示了“母”字的训练和识别过程.

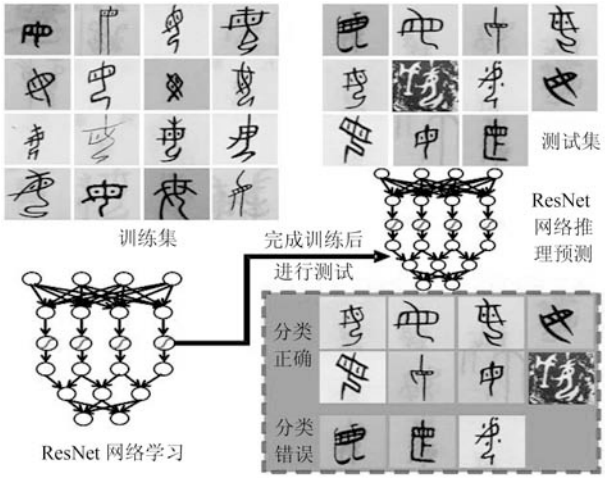

图 9 “母”字的网络学习与预测过程示意图

Fig. 9 Illustration of learning and prediction of character “母”

在我们收集的数据集中,“母”字一共有 27 张图片, 如图 9 所示, 按照我们的实验设置, 随机抽选其中 16 张作为训练图片提供给模型学习, 而剩余 11 图片将留为测试集. 当模型完成学习后, 我们将对每一张测试图片进行预测, 11 张中成功预测了其中 8 张, 错误识别其中 3 张. 分类正确的图片大多在训练集中能找到形体类似的例子, 而识别错误的 3 个字分别来源于《上海博物馆藏战国楚竹书》中《民之父母》的“母”字、《古玺汇编》所收的“母”字、《甲骨文合集》所收的“母”字, 如下图 10 所示.

图 10 识别错误的 3 个“母”字

Fig. 10 Three cases of wrong recognition of character “母”

这 3 个字均不是来源于青铜器铭文, 而分是竹书、玉玺和甲骨文, 直观上看, 与其他“母”字的特征不相似甚至不一致. 如确认此 3 字仍为“母”字的话, 可以看出: 1) 楚国竹书的书写风格、字体构造确与中原国家不同, 如果见字如形的话, 这个“母”字没有其他母字那样婀娜的姿态, 没有女性特征的形态, 也不着裙装, 基本没有相似特征, 难以识别; 2) 在玉玺上篆刻文字, 与青铜铭文或者甲骨钻文, 在书写方式上很不一样, 可能会囿于玉玺形状对文字进行变形, 造成特征在此字上表现非常弱, 故较难识别; 3) 最后一个甲骨文“母”字, 4 个点在类似女性胸部形体的外侧, 与所有其他母字的特征完全不一致, 确实难以识别为同一个字.

从这个例子中我们看到网络具有较强的抽象概括能力, 能够从给定的图片中求同存异地抽取一些本质特征, 从而实现从训练集到测试集的迁移. 这对于古文字识别具有重要价值.

3.4 讨论

在分析实验结果并对比古文字库的过程中, 我们发现有一种现象较为严重地影响了识别结果, 那就是鸟文的存在. 鸟文又称“鸟书”, 是出现于春秋中期, 盛行于春秋晚期至战国中期的一种书法体, 通常会添加鸟纹来修饰文字, 类似现在的装饰字、艺术字. 图 11 中举了 3 个例子.

图 11 “子”、“吉”、“名”3 个字的甲骨文、金文和鸟文的对比

Fig. 11 The comparison of oracle-bone, bronze inscriptions and bird-writing for character “子”, “吉” and “名”

从图 11 可以看出, 每个字的甲骨文和青铜器铭文具有某种程度上的神似, 但是鸟文却差别非常大, 鸟文很大程度上是字与画的结合, 因此我们的模型很难从类似甲骨文和青铜器铭文的形态中推断出一个鸟文古文字的真实含义.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

4 未来工作

本文在搭建古文字库时, 受手头资料所限, 除录入了青铜器铭文外, 也大量录入了陶器、石刻、竹简、帛书、玺印等文字, 同时青铜器铭文中还录入了不少鸟文铭文, 从实验结果看, 一定程度上造成了模型识别古文字的难度增加, 今后在进一步优化模型时, 将首先调整古文字库的收字范围, 尽可能以青铜器铭文为主, 辅以甲骨文, 暂时去除地域特征过于明显的鸟文铭文等文字. 此外, 结合深度学习技术的特点, 还应扩大字库收录字数, 尽可能多地录入字频较高的古文字, 帮助模型更好地学习古文字特征及演变规律.

汉字大多数为合体字, 即多数字由独体字以及固定的偏旁部首构成, 因此, 在利用模型分析古文字的形体特征时, 可搭建单独的独体字库并设计模型先识别独体字, 待模型训练完成后, 测试合体字识别效果, 同时进一步深入研究独体字特征对于识别古文字的作用, 优化模型的识别效果.

此外, 本文提出的方法是构建在单个文字特征识别的基础上的. 文字的识别除形体识别之外, 还需借助其他研究方法, 综合相关因素进行分析. 单个青铜器铭文的识别与推测除了可以依赖它本身的特征之外, 还可以依据偏旁构造分析、字音、字义、语法、上下文、青铜器器物类别、分期与出土地点等进行综合释读. 对于其中尤其重要上下文文字间的关联信息, 深度学习领域另外一个分支: 递归神经网络 (Recurrent neural network, RNN[23]), 具有巨大的潜力, 它能同时综合单个文字变体本身的特征和其使用环境的上下文文字特征, 使得推断的准确率能大幅提高.

5 结论

本文提出了一种基于图像识别技术进行青铜铭文识别的方法, 是人工智能与青铜铭文考古的一次有意义的交叉学科研究. 受古文字偏旁分析法所启发, 我们设计了基于两阶段特征映射的深度学习模型. 该方法利用大量字例样本, 有效吸收消化古文字研究, 特别是甲骨文及青铜器铭文的研究成果, 准确、快速地识别未知古文字, 有效地提高古文字研究者的研究效率, 很好地辅助新发掘古文字的推断释读与解义.

我们认为, 基于深度学习的图像识别技术在古文字识别中具有重要应用价值. 实验证明, 它能够有效地识别未知古文字, 并给出准确率很高的推测识别结果, 能够极大地提升古文字识别的准确性和识别效率. 有效扩大古文字图片数据库, 进一步优化模型, 并结合考古学及古文字学的已有研究成果, 可以积极地应用于新出土的考古青铜器铭文识别, 推动现场考古研究的发展, 也可以帮助普通大众去识别了解陈列在博物馆的青铜器铭文释义, 甚至判别识别有争议的古文字或疑难字词, 帮助识别未知的古文字. 在将来的工作中, 我们会考虑扩大数据库或者从深度学习小样本学习的角度来探索青铜铭文识别.

References

1 Ma Cheng-Yuan. *Ancient Chinese Bronzes* (Second Edition). Shanghai: Shanghai Renmin Press, 2016. 9−41
(马承源. 中国古代青铜器. 第 2 版. 上海: 上海人民出版社, 2016. 9−41)

2 Li Xue-Qin. *Primary Chinese Paleography* (Second Edition). Beijing: Zhonghua Book Company, 2006. 9−46
(李学勤. 古文字学初阶. 第 2 版. 北京: 中华书局, 2006. 9−46)

3 Gao Ming. *General Theory of Chinese Paleography*. Beijing: Beijing University Press, 1996. 56−170
(高明. 中国古文字学通论. 北京: 北京大学出版社, 1996. 56−170)

4 Gao Ming, Tu Bai-Kui. *Ancient Chinese Character Type Series*. Shanghai: Shanghai Classics Publishing House, 2014. 1−1427
(高明, 涂白奎. 古文字类编. 上海: 上海古籍出版社, 2014. 1−1427)

5 Zhang Ya-Chu. *Shang and Zhou Dynasties Bronze Inscriptions Integration Index*. Beijing: Zhonghua Book Company, 2001. 1−225
(张亚初. 殷周金文集成引得. 北京: 中华书局, 2001. 1−225)

6 Zhou Xin-Lun, Li Feng, Hua Xing-Cheng, Wei Jian. A method of Jia Gu Wen recognition based on a two-level classification. *Journal of Fudan University (Natural Science)*, 1996, **35**(5): 481−486
(周新伦, 李锋, 华星城, 韦剑. 甲骨文计算机识别方法研究. 复旦学报 (自然科学版), 1996, **35**(5): 481−486)

7 Li Feng, Zhou Xin-Lun. Recohnition of Jia Gu Wen based on graph theory. *Journal of Electronics & Information Technology*, 1996, **18**(S1): 41−47
(李峰, 周新伦. 甲骨文自动识别的图论方法. 电子科学学刊, 1996, **18**(S1): 41−47)

8 Gu Shao-Tong. Identification of oracle-bone script fonts based on topological registration. *Computer & Digital Engineering*, 2016, **44**(10): 2001−2006
(顾绍通. 基于拓扑配准的甲骨文字形识别方法. 计算机与数字工程, 2016, **44**(10): 2001−2006)

9 Lv Xiao-Qing, Li Mo-Nan, Cai Kai-Wei, Wang Xiao, Tang Ying-Min. A graphic-based method for Chinese oracle-bone classification. *Journal of Beijing Information Science & Technology University*, 2010, **25**(S2): 92−96
(吕肖庆, 李沐楠, 蔡凯伟, 王晓, 唐英敏. 一种基于图形识别的甲骨文分类方法. 北京信息科技大学学报, 2010, **25**(S2): 92−96)

10 Wang Jia-Mei, Wen Yong-Hua, Li Yan-Qing, Gao Ya-Li. The recognition system of old-Yi character based on the image segmentation. *Journal of Yunnan Nationalities University (Natural Sciences Edition)*, 2008, **17**(1): 76−79
(王嘉梅, 文永华, 李燕青, 高雅莉. 基于图像分割的古彝文字识别系统研究. 云南民族大学学报 (自然科学版), 2008, **17**(1): 76−79)

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

11 Sun Hua. Study of Ancient Chinese Character based on Multi-feature SVM Image Recognition Method [Master thesis], Central South University, China, 2010
(孙华. 基于多特征融合 SVM 的古汉字图像识别研究 [硕士学位论文], 中南大学, 中国, 2010)

12 Sun Ying-Ying. Recognition of Ancient Chinese Characters Based on Hybrid Kernel LS-SVM [Master thesis], Anhui University, China, 2015
(孙莹莹. 基于混合核 LS-SVM 的古汉字图像识别 [硕士学位论文], 安徽大学, 中国, 2015)

13 Krizhevsky K, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: ACM, 2012. 1097−1105

14 Deng J, Dong W, Socher R, Li L J, Li K, Li F F. ImageNet: a large-scale hierarchical image database. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 248−255

15 Cao C S, Liu X M, Yang Y, Yu Y N, Wang J, Wang Z L, et al. Look and think twice: capturing top-down visual attention with feedback convolutional neural networks. In: Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015. 2956−2964

16 Zhang X Y, Bengio Y, Liu C L. Online and offline handwritten Chinese character recognition: a comprehensive study and new benchmark. *Pattern Recognition*, 2017, **61**: 348−360

17 Wu Y C, Yin F, Liu C L. Improving handwritten Chinese text recognition using neural network language models and convolutional neural network shape models. *Pattern Recognition*, 2017, **65**: 251−264

18 Zeiler M D, Fergus R. Visualizing and understanding convolutional networks. In: Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014. 818−833

19 Zhou B L, Khosla A, Lapedriza A, Oliva A, Torralba A. Object detectors emerge in deep scene CNNs. In: Proceedings of the 2015 International Conference on Learning Representations. San Diego, USA: ICLR, 2015.

20 Simonyan K, Zisserman A. Very deep convolutional networks for large-Scale image recognition. In: Proceedings of the 2015 International Conference on Learning Representations. San Diego, USA: ICLR, 2015.

21 Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 1−9

22 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 770−778

23 Mnih V, Heess N, Graves A, Kavukcuoglu K. Recurrent models of visual attention. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: ACM, 2014. 2204−2212

李文英 华中科技大学控制科学与工程系硕士研究生. 中国人民大学历史学院考古文博系硕士研究生. 主要研究方向为基于模式识别方法的古文字识别, 基于计算机视觉的考古学研究.
E-mail: freemin77@126.com
(**LI Wen-Ying** Master student in Department of Control Science and Engineering, HuaZhong University of Science and Technology. Master student in the Archaeology and Museology Department, School of History, Renmin University of China. Her research interest covers ancient character recognition based on pattern recognition, and archaeological research based on computer vision.)

曹 斌 中国人民大学历史学院考古文博系副教授. 主要研究方向为商周考古、青铜器与金文研究. 本文通信作者.
E-mail: caobin@ruc.edu.cn
(**CAO Bin** Associate professor at Archaeology and Museology Department, School of History, Renmin University of China. His research interest covers Shang and Zhou archaeology (archaeology on the Chinese bronze age), bronze and inscriptions research. Corresponding author of this paper.)

曹春水 中国科学技术大学自动化系与中国科学院自动化研究所模式识别国家重点实验室联合培养的博士研究生. 主要研究方向是深度学习与计算机视觉.
E-mail: ccs@mail.ustc.edu.cn
(**CAO Chun-Shui** Ph. D. candidate in the Department of Automation, University of Science and Technology of China and now studying in the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. His research interest covers artificial intelligence, machine learning, and computer vision.)

黄永祯 中科院自动化所模式识别国家重点实验室副研究员. 主要研究方向为模式识别, 计算机视觉.
E-mail: yzhuang@nlpr.ia.ac.cn
(**HUANG Yong-Zhen** Associate professor at National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences. His research interest covers computer vision and pattern recognition.)

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net