



基于监督对比学习的小样本甲骨文字识别

毕晓君, 毛亚菲

引用本文:

毕晓君, 毛亚菲. 基于监督对比学习的小样本甲骨文字识别[J]. 智能系统学报, 2024, 19(1): 106–113.

BI Xiaojun, MAO Yafei. Few-shot oracle bone character recognition based on supervised contrastive learning[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(1): 106–113.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202309008>

您可能感兴趣的其他文章

基于孪生变分自编码器的小样本图像分类方法

A small-sample image classification method based on a Siamese variational auto-encoder

智能系统学报. 2021, 16(2): 254–262 <https://dx.doi.org/10.11992/tis.201906022>

非结构化文档敏感数据识别与异常行为分析

Unstructured document sensitive data identification and abnormal behavior analysis

智能系统学报. 2021, 16(5): 932–939 <https://dx.doi.org/10.11992/tis.202104028>

基于Capsule网络的甲骨文构件识别方法

Recognition of Oracle Radical based on the Capsule network

智能系统学报. 2020, 15(2): 243–254 <https://dx.doi.org/10.11992/tis.201904069>

生成对抗网络辅助学习的舰船目标精细识别

Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks

智能系统学报. 2020, 15(2): 296–301 <https://dx.doi.org/10.11992/tis.201901004>

基于增强AlexNet的音乐流派识别研究

Music genre recognition research based on enhanced AlexNet

智能系统学报. 2020, 15(4): 750–757 <https://dx.doi.org/10.11992/tis.201909032>

基于级联宽度学习的多模态材质识别

Cascade broad learning for multi-modal material recognition

智能系统学报. 2020, 15(4): 787–794 <https://dx.doi.org/10.11992/tis.201908021>

DOI: 10.11992/tis.202309008

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20240108.0958.002>

基于监督对比学习的小样本甲骨文字识别

毕晓君^{1,2}, 毛亚菲³

(1. 民族语言智能分析与安全治理教育部重点实验室, 北京 100081; 2. 中央民族大学 信息工程学院, 北京 100081; 3. 哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨 150001)

摘要: 针对由于甲骨文中部分字符的出现频率较低, 直接利用深度神经网络进行识别会产生严重的过拟合现象, 进而导致识别精度较差的问题, 本文提出一种基于监督对比学习的小样本甲骨文字识别方法。选用利用增强样本的 Y 型 (ensemble augmented-shot Y-shaped, EASY) 学习框架作为网络的主干部分, 通过集合数据增强、多骨干网络集成、特征向量投影等训练策略, 直接实现利用少量带标签样本进行识别; 引入监督对比学习, 并提出联合对比损失, 使得特征空间中类内特征向量距离更近, 类间特征向量距离更远, 进一步提高模型性能。实验结果表明: 相比于当前效果最好的 Orc-Bert 模型, 提出的小样本甲骨文字识别模型在 1-shot 任务中的准确率提升了 26.42%, 3-shot 任务的准确率提升了 28.55%, 5-shot 任务的准确率提升了 23.98%, 较好解决了低频率出现的甲骨文字识别精度较差的问题。

关键词: 甲骨文字识别; 小样本; 监督对比学习; 利用增强样本的 Y 型学习框架; 深度学习; 特征空间; 联合对比损失

中图分类号: TP391 文献标志码: A 文章编号: 1673-4785(2024)01-0106-08

中文引用格式: 毕晓君, 毛亚菲. 基于监督对比学习的小样本甲骨文字识别 [J]. 智能系统学报, 2024, 19(1): 106-113.

英文引用格式: BI Xiaojun, MAO Yafei. Few-shot oracle bone character recognition based on supervised contrastive learning [J]. CAAI transactions on intelligent systems, 2024, 19(1): 106-113.

Few-shot oracle bone character recognition based on supervised contrastive learning

BI Xiaojun^{1,2}, MAO Yafei³

(1. Key Laboratory of Ethnic Language Intelligent Analysis and Security Governance of MOE, Beijing 100081, China; 2. Department of Information Engineering, Minzu University of China, Beijing 100081, China; 3. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: Due to low frequency of occurrence of some characters in Oracle, directly using the deep neural network for recognition will produce serious overfitting, which will lead to poor recognition accuracy. To this end, this paper proposes a few-shot oracle bone character recognition method based on supervised contrastive learning. The ensemble augmented-shot Y-shaped (EASY) learning framework is selected as the backbone part of the network. Through training techniques such as collective data enhancement, multi-backbone network integration, and feature vector projection, etc., it is possible to directly use a small number of labeled samples for identification. And then, introducing the supervised contrastive learning and the concept of a joint contrastive loss to make the intra-class feature vectors in the feature space closer and the inter-class feature vectors further apart, thereby the model performance is improved further. The experimental results show that compared with the current best-performing Orc-Bert model, the accuracy of the few-shot oracle recognition model proposed in this paper has increased by 26.42% in the 1-shot task, 28.55% in the 3-shot task, and 23.98% in the 5-shot task, which better solves the problem of poor recognition accuracy of low-frequency oracle bone characters.

Keywords: oracle bone character recognition; few-shot; supervised contrastive learning; EASY framework; deep learning; feature space; joint contrastive loss

作为中华文化的瑰宝, 甲骨文字所承载的不仅是古代王朝的兴衰更替的信息, 更是我们中华

五千年文化传承的历史见证^[1]。自甲骨片挖掘工作的展开以来, 国内外学者对于甲骨文的研究就不曾停止。据目前所掌握的甲骨文字数据资料来看, 古代人民对各个甲骨文字的使用度也大小不一, 部分常用字的出现频率可达成百上千乃至上

收稿日期: 2023-09-06. 网络出版日期: 2024-01-08.

基金项目: 国家自然科学基金重点项目 (62236011); 国家社科基金重大项目 (20&ZD279).

通信作者: 毕晓君. E-mail: bixiaojun@hrbeu.edu.cn.

©《智能系统学报》编辑部版权所有

万,但也有不常用字的出现频率仅有十个以内甚至一个。任何一个文字中所包含的信息及意义都将对我们了解历史以及传承文化产生长远的影响,因此,在深度学习长足发展的今天,甲骨文字识别领域中存在的小样本数据情况更是我们关注的重中之重。卷积神经网络优秀的分类表现依赖于充足的样本数据^[2],而甲骨学研究却有着特殊性,随着对甲骨文字的发掘和研究越来越深入,未来发现的甲骨文字的样本数量会有相当一部分存在数据有限的问题^[3]。因此深度神经网络会出现严重的过拟合及较差的泛化性。随着现实世界的强烈需求,小样本学习成为各领域研究者关注的热点,它的提出旨在解决深度学习对大量数据的依赖性^[4]。研究者提出了大量利用少量样本甚至单样本进行识别的优秀方法,并且取得了极其优秀的效果^[5-10],而目前将其成果应用在甲骨文字识别领域的却极少,由此,利用小样本学习方法解决甲骨文字识别中存在的小样本问题则显得更为必要和迫切。

通过结合当前主流小样本图像识别算法对目前存在的小样本甲骨文字识别的研究现状进行分析,存在以下2点问题:

1) Orc-Bert 算法框架^[11]在解决甲骨文字小样本识别问题的过程中需要大量的未标记源数据来学习笔划特征,无法直接利用已有的少量标注样本训练识别模型。

2) 现有方法在小样本的条件下,网络学习到的特征十分有限,从而导致识别效果较差。

针对以上问题,本文提出了一种基于监督对比学习的小样本甲骨文字识别方法。EASY 框架集成了骨干网络训练、数据增强、多骨干网络集成、特征向量投影等训练策略,可以直接使用少

量带标签样本训练识别模型,同时达到先进的效果。在此基础上,将监督对比学习的思想引入到模型训练中来,将输入图片进行数据增强后输入特征提取网络,然后提出联合对比损失来对网络进行优化和参数学习,使特征空间中类内特征向量距离更近,类间特征向量距离更远,骨干网络获得了更加充足的特征,最终实现了识别效果的提升。

1 小样本甲骨文字整体识别框架

小样本学习作为消除深度神经网络对大数据依赖的最佳方法而备受研究者瞩目,尽管目前研究成果颇丰,但应用在甲骨文字识别领域的研究成果整体较少。针对目前提出的小样本甲骨文字识别方法仍然需要在大型数据集上进行预训练的问题,而不能直接利用少量标记样本训练模型,为此本文引入了新的小样本图像识别框架 EASY^[12],它主要汇集了目前领域中常见的骨干网络训练^[13]、数据增强^[14]、多骨干网络集成^[15]、特征向量投影^[16]等优秀训练策略,可以直接使用少量标注样本训练识别模型,同时达到先进的效果。在基于 EASY 的小样本图像识别框架上,针对现有方法小样本的条件下,网络学习到的特征有限,将导致识别效果差的问题,本文引入了监督对比学习的思想^[17],将输入图片进行数据增强后输入特征提取网络,然后提出联合对比损失来对网络进行优化和参数学习,使特征空间中类内特征向量距离更近,类间特征向量距离更远,骨干网络获得了更加充足的特征,从而实现了识别效果的提高。本文提出的小样本甲骨文字识别整体框架如图1所示。

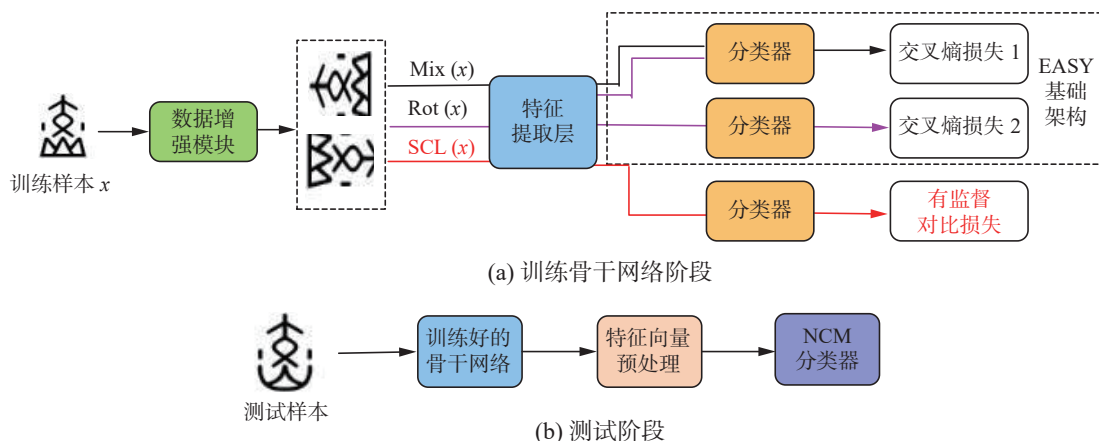


图1 甲骨文字小样本识别模型整体框架

Fig. 1 Overall framework of oracle bone script few-shot recognition model

此识别框架沿用了 EASY 的整体架构,并根据任务的实际情况进行了优化。首先在训练阶段利用训练集对骨干网络进行训练,在这个过程保留了原框架中使用的自监督多样混合(self-supervised manifold mixup, S2M2)训练策略,并且结合了监督对比学习的思想。在输入一张图像之后,经过图像增强模块,会产生 2 张增强后的输入图像,将增强后的数据作为新的输入,在模型学习的过程中除了使用混合^[18]和旋转的训练策略外,还利用多组正负样本对进行样本对比学习,使特征空间中类内特征向量距离更近,类间特征向量距离更远,由此骨干网络获得的特征更加充足。原框架中利用 2 个交叉熵损失函数 L_1 和 L_2 来对训练集分类任务和图像旋转角度预测的辅助任务进行监督,交叉熵损失函数为

$$L = \frac{1}{n} \sum_{i=1}^n \left(- \sum_{j=1}^C y_{i,j} \log p_{i,j} \right) \quad (1)$$

式中: n 为批大小, C 为输入一批数据的类别数, $y_{i,j}$ 为第 i 个样本在第 j 类上的真实标签, $p_{i,j}$ 为第 i 个样本在第 j 类上的预测概率。则原框架中的损失函数可以表示为

$$L_{CE} = 0.5L_1 + 0.5L_2 \quad (2)$$

式中 L_1 和 L_2 的形式同 L 。在加入对比学习的策略之后,本文引入了监督对比损失函数为

$$L_{SC} = \sum_{k \in I} \frac{-1}{|P(k)|} \sum_{p \in P(k)} \log \frac{\exp(z_k \cdot z_p / \tau)}{\sum_{a \in A(k)} \exp(z_k \cdot z_a / \tau)} \quad (3)$$

式中: $k \in I \equiv \{1, 2, \dots, 2n\}$, $P(k)$ 是增强后的样本中所

有的正样本索引集合, $|P(k)|$ 是集合中样本的数量, $A(k)$ 则为增强后除了锚点外所有样本索引的集合, z_k 代表锚点的特征向量, z_p 代表任意正样本的特征向量, z_a 代表增强后样本集合中除锚点外任意样本的特征向量, $\tau \in \mathbf{R}^+$ 为一个标量温度系数。

因此,本文提出的联合对比损失函数为

$$L_{Total} = L_{CE} + L_{SC} \quad (4)$$

本文利用提出的联合对比损失来对网络进行优化和参数学习,然后利用循环余弦退火算法^[19]获得多个优化到局部最小值解的模型,最终将模型进行参数冻结。在对新类进行测试的阶段,首先利用训练好的骨干网络对输入的新类图像进行特征提取,接着将得到的特征向量进行中心化和归一化操作,最终利用最近类平均分类器^[20](nearest class mean classifier, NCM)得到分类结果。

2 小样本甲骨文字识别网络结构

本节主要对提出的小样本甲骨文字识别框架的各个组成部分进行详细介绍。

2.1 EASY 小样本图像识别框架

EASY 是一种简单的小样本图像识别框架,其中组合了许多常见的方法如骨干网络训练、特征向量投影等。虽然结构简单,但是在训练中却很容易达到先进的性能且不会产生巨大的计算消耗。最重要的是, EASY 不需要用大型数据集进行预训练,仅用少量带标签样本即可训练模型。EASY 小样本图像识别框架如图 2 所示。

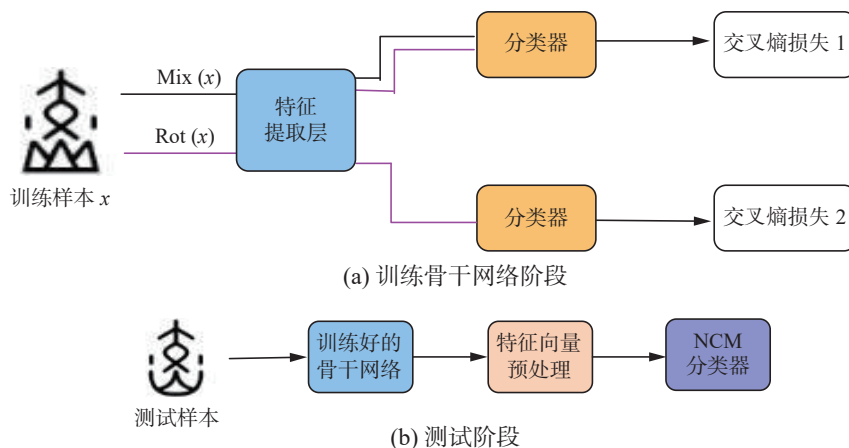


图 2 EASY 基础架构

Fig. 2 Basic structure of EASY

如图 2 所示, EASY 基础架构包含 4 个步骤:

1) 利用训练集训练一个骨干网络集合。在这个过程中,使用循环余弦退火算法,在每一步学习率都在更新。在余弦循环期间,学习率在初始

学习率和 0 之间变化,在循环结束时,重启学习程序,以降低的初始学习率重新开始。初始学习率设置为 0.1,在每个循环中降低 10%,训练过程设置 5 个循环周期,每个周期有 100 个 epoch。训练

骨干网络使用的是 S2M2R 的方法, 其原理是采用标准分类架构 (如 ResNet12^[21]), 并在倒数第 2 层之后分支一个新的逻辑回归分类器, 此外还有一个用于识别样本类别的分类器, 从而整个模型呈 Y 型。这个新的分类器是用来检索 4 种可能的旋转, 即 0° 、 90° 、 180° 、 270° , 哪一种被应用于输入样本。在训练的每一步中使用 2 步前向-后向传播, 第 1 批输入数据与 Mixup 方法结合仅被送到第 1 个分类器, 第 2 批输入数据在进行随机旋转后被同时送到 2 个分类器。训练结束后, 骨干网络参数被冻结, 然后被用于从训练集和测试集中提取特征向量。

2) 新类图像经过第 1) 步产生的骨干网络集合后, 会产生一个相应的平均特征向量集合, 将这些平均向量进行拼接操作, 则获得新类图像的最终特征表示。整个过程如图 3 所示。

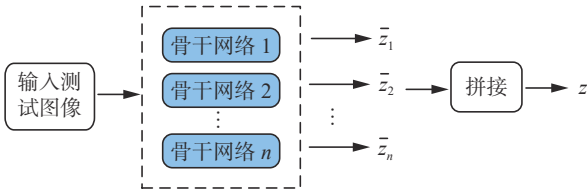


图 3 拼接操作示意

Fig. 3 Schematic diagram of the concat operation

3) 对第 2) 步所获得的终极特征向量进行 2 次预处理。假设 \bar{z} 代表了训练集的平均特征向量, 则第 1 次操作称作中心化:

$$z_C = z - \bar{z} \quad (5)$$

第 2 次操作是在第 1 次操作的基础上进行超球面上的投影:

$$z_{CH} = \frac{z_C}{\|z_C\|_2} \quad (6)$$

4) 使用最近类平均分类器进行分类。通过首次标记样本类重心来获得预测, 类重心计算为

$$\forall i: \bar{c}_i = \frac{1}{|S_i|} \sum_{z \in S_i} z \quad (7)$$

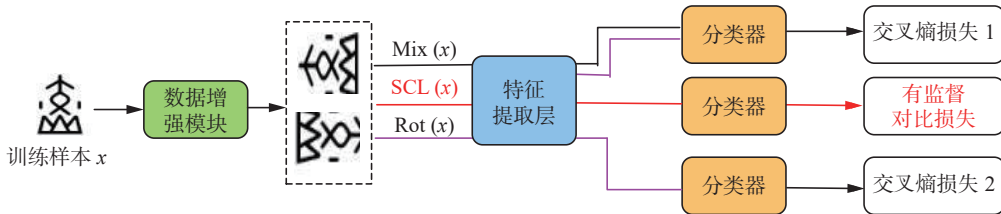


图 4 改进后的训练阶段框架

Fig. 4 Improved framework for the training phase

输入一个训练样本, 经过数据增强模块后, 生成 2 个随机的增强数据。其中数据增强模块中包含随机水平翻转、随机改变图像属性及随机灰度化操作。这些增强数据表示出原始数据不同的视

然后将最近的重心与查询集进行关联:

$$\forall z \in Q: C_{ind}(z, [\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n]) = \underset{i}{\operatorname{argmin}} \|z - \bar{c}_i\|_2 \quad (8)$$

式中: $S_i (i \in \{1, 2, \dots, n\})$ 表示支持集中第 i 类经过第 2) 步预处理之后的特征向量集合, Q 代表经过预处理的查询集的特征向量集合。

2.2 监督对比学习

虽然利用 EASY 框架解决了小样本甲骨文字识别过程中无法使用少量标注数据直接训练分类模型的问题, 但是在研究 EASY 框架的骨干网络训练部分时, 发现在输入一批训练样本时, 由于样本量较小, 在交叉熵损失函数的指导优化下, 骨干网络获得的特征十分有限。因此利用此种设置下训练好的网络在进行测试的时候, 在特征空间中, 利用已学习到的特征不足以区分文字的各个类别, 从而影响了识别准确率的提升。

为了解决上述问题, 本文引入了监督对比学习的相关思想^[11]。在监督对比学习的思想提出之前, 对比学习领域中也出现过许多优秀的相关工作^[22-24]。这些工作的共同思想是在特征空间中将锚点和正样本拉得更近, 而将负样本从锚点处推得更远。这里要提到 2 种经典的对比损失: 三元组损失和 N 对损失。三元组损失对于每个锚点仅使用一个正样本和一个负样本, 相当于使用了一个正负样本对; 而 N 对损失则是对每个锚点使用一个正样本和多个负样本, 因此可以以一个正样本构建多个正负样本对。不同于前面 2 个损失, 监督对比学习在考虑多个负样本的同时考虑多个正样本, 这样就更有效地利用标签信息, 通过对比更好地使特征空间中类内特征向量距离更近, 类间特征向量距离更远, 从而使骨干网络学习到更为充足的特征, 有助于更好的识别精度的提升。引入监督对比学习的 EASY 框架的训练部分如图 4 所示。

角, 并包含了原始数据的部分信息。然后将增强的 2 个数据分别输入同一个特征提取网络, 然后在原有损失基础上加入监督对比损失, 即构造出新的联合对比损失来对网络进行优化和参数学习。

3 仿真实验及结果分析

3.1 数据集的选取与处理

本文所选用的数据集有 Oracle-FS^[25] 和 HWOBC-FS, 其数据示例如图 5 所示。Oracle-FS 是于 2020 年由 Han 等^[11] 提出的用来进行小样本甲骨文字识别的公开数据集。Oracle-FS 中包含了 200 个甲骨文字, 对于小样本的设置分为 3 种。 k -shot 任务所对应的训练集中每个类包含了 k 个样本, 如对于 1-shot 任务, 对应的训练集中仅有 1 个训练样本。而 3 种设置下的测试集中均包含 20 个样本。在本文中, 设置 k 的取值为 1、3、5。Oracle-FS 的具体情况如表 1 所示。HWOBC 是一个公开的手写甲骨文字数据集, 由 22 位来自不同专业的甲骨学研究者, 通过手写甲骨字搜集软件比照甲骨文标准字形书写并整理而成, 数据集共包含 3881 类, 共包含样本图片 83245 张, 其中每类包含样本 19~25 张不等。数据集 HWOBC-FS 是基于上述数据集产生的, 数据集的制作设置以 Oracle-FS 作为参照。本文从 HWOBC 中随机挑选 200 类, 在 k -shot 任务中, 训练集每个类包含了 k 个样本, 测试集均包含了 15 个样本。在 HWOBC-FS 数据集中, 训练集和测试集中的样本在随机划分阶段保证了互不相交。HWOBC-FS 的具体情况如表 2 所示。

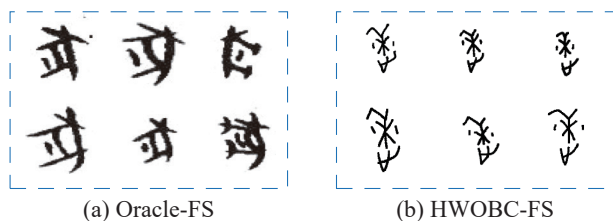


图 5 Oracle-FS 和 HWOBC-FS 的数据示例

Fig. 5 Data example for Oracle-FS and HWOBC-FS

表 1 数据集 Oracle-FS 的具体情况

Table 1 Specific situation of the dataset Oracle-FS

数据集	k -shot	每个类中的样本数量		甲骨文字类别数
		训练集	测试集	
Oracle-FS	1	1	20	200
	3	3	20	
	5	5	20	

表 2 数据集 HWOBC-FS 的具体情况

Table 2 Specific situation of the dataset HWOBC-FS

数据集	k -shot	每个类中的样本数量		甲骨文字类别数
		训练集	测试集	
HWOBC-FS	1	1	15	200
	3	3	15	
	5	5	15	

3.2 实验环境及参数设置

本文所有实验的硬件配置为: TITAN X GPU 用于计算加速, 内存 16 GB, Ubuntu 16.04 操作系统, 编程框架为 Pytorch 1.7.1。本文算法训练时设置训练批次大小为 64, 迭代次数为 100 个 epoch。初始学习率设置为 0.1, 并且在此使用了循环余弦退火算法, 在 100 个 epoch 之内, 学习率在初始学习率和 0 之间变化。超参数 n -way 设置为 5, 超参数 n -runs 设置为 10000。

3.3 评价指标

为了验证提出的基于监督对比学习的小样本甲骨文字识别方法的分类性能, 本文使用 k -shot 分类准确率 (k -shot acc) 来作为小样本甲骨文字识别的分类指标。 k -shot 分类准确率是用来定量评估本文提出的方法在 k -shot 任务中进行多类别分类的能力。首先在训练集中随机抽取 N 类, 每类中包含 k 张甲骨文字样本构成一个支持集。然后输入一张待测试的甲骨文字图片进行所属类别测试, 其中测试样本所属类别包含在支持集的 N 类中, 称其为一次 k -shot 任务。定义 P_i 为一次 k -shot 任务所得到的结果, 如果识别正确, 则值为 1, 错误则为 0。假设 M 为进行 k -shot 任务的次数, 则 k -shot 分类准确率可以表示为

$$k\text{-shot acc} = \frac{\sum_{i=1}^M P_i}{M} \quad (9)$$

3.4 模型的有效性验证

为了验证本文提出的基于监督对比学习的小样本甲骨文字识别方法的有效性, 本文选用了小样本学习研究中常见的 3 种骨干网络, 在上述提到的 2 个公开数据集上分别对基本框架 EASY 和引入监督对比学习的框架 EASY+SCL 进行了实验验证及对比, 在数据集 Oracle-FS 上的实验验证结果如表 3 所示, 在 HWOBC-FS 上的实验验证结果如表 4 所示。

表 3 在数据集 Oracle-FS 上的实验验证结果

Table 3 Experimental verification results on dataset Oracle-FS

k -shot	骨干网络	EASY	EASY+SupCon	%
1	ResNet18	52.70	58.32	
	ResNet20	55.71	59.65	
	ResNet34	49.93	53.18	
	ResNet18	82.57	85.75	
3	ResNet20	83.10	85.20	
	ResNet34	80.91	81.35	
	ResNet18	90.05	92.18	
5	ResNet20	91.07	91.54	
	ResNet34	91.06	92.54	

表4 在数据集 HWOBC-FS 上的实验验证结果
Table 4 Experimental verification results on dataset HWOBC-FS %

<i>k</i> -shot	骨干网络	EASY	EASY+SupCon
1-shot	ResNet18	69.34	80.84
	ResNet20	73.22	84.36
	ResNet34	61.06	76.80
	ResNet18	96.32	99.06
3-shot	ResNet20	98.24	98.97
	ResNet34	95.85	98.75
	ResNet18	99.54	99.58
5-shot	ResNet20	99.44	99.60
	ResNet34	99.51	99.57

从表3可以看出,在骨干网络为 ResNet18 时,引入监督对比学习后,1-shot、3-shot、5-shot 任务的准确率相较于原框架分别提升了 5.62%、3.18% 和 2.13%;在骨干网络为 ResNet20 时,1-shot、3-shot、5-shot 任务的准确率相较于原框架分别提升了 3.94%、2.10% 和 0.47%;在骨干网络为 ResNet34 时,1-shot、3-shot、5-shot 任务的准确率相较于原框架分别提升了 3.25%、0.44% 和 1.48%。

从表4可以看出,在骨干网络为 ResNet18 时,引入监督对比学习后,1-shot、3-shot、5-shot 任务的准确率相较于原框架分别提升了 11.50%、2.74% 和 0.04%;在骨干网络为 ResNet20 时,1-shot、3-shot、5-shot 任务的准确率相较于原框架分别提升了 11.14%、0.73% 和 0.16%;在骨干网络为 ResNet34 时,1-shot、3-shot、5-shot 任务的准确率相较于原框架分别提升了 15.74%、2.90% 和 0.06%。

综上所述,对于 *k*-shot 任务,在引入监督对比学习后,在不同的骨干网络上均有不同程度的识别精度的提升,由此可以证明本文提出的小样本甲骨文字识别方法有效地拉近了特征空间中类内特征向量的距离,推远了类间特征向量之间的距离,骨干网络获得了更加充足的特征,有效地实现了 *k*-shot 任务识别精度的提升,最终证明本文提出的小样本甲骨文字识别方法是有效的。

3.5 模型的先进性验证

为了验证本文提出的基于监督对比学习的小样本甲骨文字识别方法的先进性,将本文提出的方法与现阶段具有代表性的小样本甲骨文字识别

方法 Orc-Bert 进行了实验对比,对比结果如表5所示。

表5 模型先进性验证实验结果
Table 5 Experimental results of model advanced verification %

<i>k</i> -shot	Orc-Bert	本文方法
1-shot	31.9	58.32
3-shot	57.2	85.75
5-shot	68.2	92.18

在本次模型先进性验证实验中,选取与 Orc-Bert 使用的相同的数据集 Oracle-FS,且骨干网络的选取与 Orc-Bert 相同,均为 ResNet18,在此种设置条件下,从表5的实验结果中可以看出,本文提出的基于监督对比学习的小样本甲骨文字识别方法,1-shot、3-shot、5-shot 任务的准确率相较于 Orc-Bert 分别提升了 26.42%、28.55% 和 23.98%。而且,本文提出的小样本甲骨文字识别模型不需要大规模的未标记数据集进行预训练,直接使用现有的少量有标签数据即可实现分类模型的训练,解决了小样本甲骨文字识别任务中训练数据有限的问题。综上所述,本文提出的基于监督对比学习的小样本甲骨文字识别方法具有先进性。

4 结束语

针对当前小样本甲骨文字识别方法需要依赖大规模未标注数据集来学习相应笔划特征,无法利用现有少量标注样本直接进行分类模型训练的问题,本文引入了目前先进的小样本图像识别框架 EASY,其中结合了多种优秀且有效的训练策略,可以直接使用少量带标签样本训练识别模型,达到了较好的效果。并且针对现有方法在小样本的条件下,网络学习到的特征十分有限,从而导致识别效果较差的问题,在 EASY 框架的基础上,将监督对比学习的思想引入到模型训练中来,将输入图片进行数据增强后输入特征提取网络,然后提出联合对比损失来对网络进行优化和参数学习,有效地拉近了特征空间中类内特征向量的距离,推远了类间特征向量之间的距离,骨干网络获得了更加充足的特征,最终实现识别效果的提升。为了验证本文方法的有效性和先进性,在实验部分将本文提出的方法与基础框架在不同数据集和不同的骨干网络上对 *k*-shot 分类任务的精度进行了比较,同时将本文提出的方法

与现阶段具有代表性的小样本甲骨文字识别方法进行了对比, 本文方法均有不同程度的提升, 充分验证了本文提出方法的有效性和先进性, 为今后的小样本甲骨文字识别研究奠定了良好的基础。

参考文献:

- [1] 谢乃和. 从殷墟走向世界的“绝学”甲骨文字研究: 韩国釜山“纪念甲骨文发现 120 周年国际学术研讨会”述评[J]. 管子学刊, 2020(3): 125–128.
XIE Naihe. Flourishing from Yin Ruins to the world—a review of “busan, South Korea international symposium commemorating the 120th anniversary of oracle bone inscription discovery”[J]. Guan zi journal, 2020(3): 125–128.
- [2] GUPTA J, PATHAK S, KUMAR G. Deep learning (CNN) and transfer learning: a review[J]. *Journal of physics: conference series*, 2022, 2273(1): 012029.
- [3] HUANG Shuangping, WANG Haobin, LIU Yongge, et al. OBC306: a large-scale oracle bone character recognition dataset[C]//2019 International Conference on Document Analysis and Recognition. Piscataway IEEE, 2020: 681–688.
- [4] 安胜彪, 郭昱岐, 白宇, 等. 小样本图像分类研究综述[J]. 计算机科学与探索, 2023, 17(3): 511–532.
AN Shengbiao, GUO Yuqi, BAI Yu, et al. Survey of few-shot image classification research[J]. *Journal of frontiers of computer science and technology*, 2023, 17(3): 511–532.
- [5] LI Na, HAO Huizhen, GU Qing, et al. A transfer learning method for automatic identification of sandstone microscopic images[J]. *Computers & geosciences*, 2017, 103: 111–121.
- [6] LIU Wenhe, CHANG Xiaojun, YAN Yan, et al. Few-shot text and image classification via analogical transfer learning[J]. *ACM transactions on intelligent systems and technology*, 9(6): 71.
- [7] LONG M, ZHU H, WANG J, JORDAN MI. Deep transfer learning with joint adaptation networks[C]//International Conference on Machine Learning. PMLR, 2017: 2208–2217.
- [8] LONG Mingsheng, WANG Jianmin, JORDAN M I. Un-supervised domain adaptation with residual transfer networks[EB/OL]. (2017–02–16)[2023–09–06]. <https://arxiv.org/abs/1602.04433.pdf>.
- [9] GE Weifeng, YU Yizhou. Borrowing treasures from the wealthy: deep transfer learning through selective joint fine-tuning[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 10–19.
- [10] XIE Jiangtao, LONG Fei, LV Jiaming, et al. Joint distribution matters: deep Brownian distance covariance for few-shot classification[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 7962–7971.
- [11] HAN Wenhui, REN Xinlin, LIN Hangyu, et al. Self-supervised learning of orc-bert augmentor for recognizing few-shot oracle characters[C]//Ishikawa H, Liu CL, Pajdla T, et al. Asian Conference on Computer Vision. Cham: Springer, 2021: 652–668.
- [12] BENDOU Y, HU Yuqing, LAFARGUE R, et al. EASY: ensemble augmented-shot Y-shaped learning: state-of-the-art few-shot classification with simple ingredients[EB/OL]. (2022–02–07)[2023–09–06]. <https://arxiv.org/abs/2201.09699.pdf>
- [13] MANGLA P, SINGH M, SINHA A, et al. Charting the right manifold: manifold mixup for few-shot learning[C]//2020 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2020: 2207–2216.
- [14] ZHANG Chi, CAI Yujun, LIN Guosheng, et al. Deep-EMD: few-shot image classification with differentiable earth mover’s distance and structured classifiers[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 12200–12210.
- [15] LIU Jialin, CHAO Fei, LIN C M. Task augmentation by rotating for meta-learning[EB/OL]. (2020–02–08)[2023–09–06]. <https://arxiv.org/abs/2003.00804.pdf>.
- [16] WANG Yan, CHAO Weilun, WEINBERGER K Q, et al. SimpleShot: revisiting nearest-neighbor classification for few-shot learning[EB/OL]. (2019–11–16)[2023–09–06]. <https://arxiv.org/abs/1911.04623.pdf>.
- [17] KHOSLA P, TETERWAK P, WANG C. Supervised contrastive learning[J]. *Advances in neural information processing systems*, 2020, 33: 18661–18673.
- [18] ZHANG Hongyi, Cisse M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[EB/OL]. (2017–10–25)[2023–09–06]. <https://arxiv.org/abs/1710.09412.pdf>.

- [19] LOSHCHIOV I, HUTTER F. SGDR: stochastic gradient descent with warm restarts[EB/OL]. (2016-08-13) [2023-09-06]. <https://arxiv.org/abs/1608.03983.pdf>.
- [20] MAI Zheda, LI Ruiwen, KIM H, et al. Supervised contrastive replay: revisiting the nearest class mean classifier in online class-incremental continual learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2021: 3589-3599.
- [21] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [22] WU Zhirong, XIONG Yuanjun, YU S X, et al. Unsupervised feature learning via non-parametric instance discrimination[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3733-3742.
- [23] VAN DEN OORD A, LI Yazhe, VINYALS O. Representation learning with contrastive predictive coding [EB/OL]. (2018-07-10)[2023-09-06]. <https://arxiv.org/abs/1807.03748.pdf>.
- [24] TIAN Yonglong, KRISHNAN D, ISOLA P. Contrastive multiview coding[EB/OL]. (2019-01-13)[2023-09-06]. <https://arxiv.org/abs/1906.05849.pdf>.
- [25] LI Bang, DAI Qianwen, GAO Feng, et al. HWOBC-a handwriting oracle bone character recognition database[J]. *Journal of physics:conference series*, 2020, 1651(1): 012050.

作者简介:



毕晓君, 教授, 博士生导师, 主要研究方向为智能信息处理、图像处理、机器学习。主持和参与国家和省部级重点科研项目 10 余项, 获省部级科学技术一等奖 1 项, 省部级科学技术二等奖 6 项。发表学术论文 175 篇。E-mail: bixiaojun@hrbeu.edu.cn。



毛亚菲, 硕士研究生, 主要研究方向为甲骨文识别、深度学习。E-mail: 643878157@qq.com。