

语义增强的零样本甲骨文字符识别

刘宗昊¹, 彭文杰¹, 代港¹, 黄双萍^{1*}, 刘永革²

(1. 华南理工大学电子与信息学院, 广东广州 510641; 2. 安阳师范学院, 河南安阳 455099)

摘要: 甲骨文识别对于了解中国历史和传承中华文化都有重要的价值。目前, 人工识别甲骨文需要具备丰富的专家经验并耗费大量的时间, 而自动识别甲骨文的方法绝大部分受制于闭集假设, 在甲骨文这种陆续发现新字符的现实场景下适用范围受限。为此, 有研究者提出零样本甲骨文字符识别, 其从视觉匹配的角度出发, 将字模图像作为字符类别参考, 通过拓片图像与字模图像的相似度匹配实现拓片图像的字符识别, 然而其忽略了甲骨文拓片图像样本类内方差大的难点, 仍存在因字形多变而容易匹配错误的不足。本文提出了一种两阶段的语义增强零样本甲骨文字符识别方法。第一阶段为域无关的字符语义学习阶段, 通过提示学习从甲骨文拓片和字模图像中提取字符语义, 解决甲骨文字符缺乏语义的问题。为应对拓片与字模之间的域差异, 我们分别设置可学习的域提示信息 and 字符类别提示信息, 通过解耦两者的语义实现更准确的特征提取。第二阶段为语义增强的字符图像视觉匹配阶段, 模型通过两个分支分别提取类内共享特征和类间差异特征。第一个分支使用对比学习, 将同一字符类别的不同字形视觉特征对齐到字符语义, 引导模型关注类内共享特征; 第二个分支使用损失函数 N-Pair, 增强模型对不同字符类别间差异特征的学习。在测试阶段, 模型无须语义特征, 通过训练中学到的类内相似性和类间差异性特征, 实现更准确的拓片与字模匹配, 提升零样本识别性能。我们在拓片数据集 OBC306 和字模数据集 SOC5519 上进行实验验证, 实验结果表明, 本文提出的方法在零样本甲骨文字符识别准确率比基准方法性能提升超过 25%。

关键词: 甲骨文字符识别; 零样本识别; 视觉匹配; 语义增强; 视觉-语言模型; 对比学习

基金项目: 国家重点研发计划(No.2023YFC3502900); 国家自然科学基金(No.62176093, No.61673182); 广州市重点领域研发计划(No.202206030001); 粤港澳联合创新领域项目(No.2023A0505030016)

中图分类号: TP183

文献标识码: A

文章编号: 0372-2112(2024)10-3347-12

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20240286

Semantic-Enhanced Zero-Shot Oracle Character Recognition

LIU Zong-hao¹, PENG Wen-jie¹, DAI Gang¹, HUANG Shuang-ping^{1*}, LIU Yong-ge²

(1. School of Electronic and Information Engineering, South China University of Technology, Guangzhou, Guangdong 510641, China;

2. Anyang Normal University, Anyang Henan 455099, China)

Abstract: Oracle bone character recognition holds significant value for understanding Chinese history and the inheritance of Chinese culture. Currently, manual recognition of oracle bone character requires extensive expert experience and consumes a great deal of time, while the majority of methods for automatic recognition are constrained by the closed-set assumption. This limitation becomes pronounced in the context of oracle bones, where new characters are continuously discovered. To address this, some researchers achieved zero-shot oracle character recognition by visual matching. This method employs handprinted images as category references, achieving character recognition in scanned images through similarity matching with handprinted references. However, this approach overlooks the challenge of large intra-class variance in oracle bone scanned images, leading to potential mismatches due to the variability in glyphs. This paper proposes a two-stage semantic-enhanced zero-shot oracle character recognition method. The first stage is domain-independent character semantic learning, where the contrastive vision-language pre-training model CLIP is used to extract character semantics from oracle rubbings and template images through prompt learning, addressing the lack of semantic information in oracle characters. To cope with the domain differences between rubbings and templates, we set learnable domain-specific prompts and character category prompts, decoupling their semantics to achieve more accurate feature extraction. The second stage is semantic-enhanced character image visual matching. The model extracts intra-class shared features and inter-class distinctive features

through two branches. The first branch uses contrastive learning to align the visual features of different glyphs within the same character category to the character semantics, guiding the model to focus on intra-class shared features. The second branch employs the loss function N-Pair to enhance the model's ability to learn distinctive features between different character categories. During the testing phase, the model does not require semantic features; instead, it utilizes the intra-class similarity and inter-class distinctiveness learned during training to achieve more accurate matching between rubbings and templates, improving zero-shot recognition performance. Experimental validation on the scanned images dataset OBC306 and the handprinted images dataset SOC5519 demonstrates that our proposed method surpasses the baseline method in zero-shot oracle character recognition accuracy by over 25%.

Key words: oracle character recognition; zero-shot recognition; visual matching; semantic-enhanced; vision language model; contrastive learning

Foundation Item(s): National Key R&D Program of China (No.2023YFC3502900); National Natural Science Foundation of China (No.62176093, No.61673182); Key Realm R&D Program of Guangzhou (No.202206030001); Guangdong-Hong Kong-Macao Joint Innovation Project (No.2023A0505030016)

1 引言

甲骨文是中国历史的文化瑰宝,识别甲骨文对了解中国历史和传承文化具有重要意义.然而,甲骨文识别难度较高,通常需要依赖专家手动辨识^[1],自动化识别可加快甲骨文的研究进程.当前甲骨文图像主要分为拓片图像和字模图像,如图1所示.拓片图像通过在龟甲、兽骨^[2]等文字载体上拓印而得,是甲骨文研究的第一手材料;字模图像由专家对拓片图像临摹得到,修复了残缺的字形,是辅助甲骨文研究的重要工具.本文同样把字模图像作为辅助资料,研究甲骨文拓片图像的零样本识别.

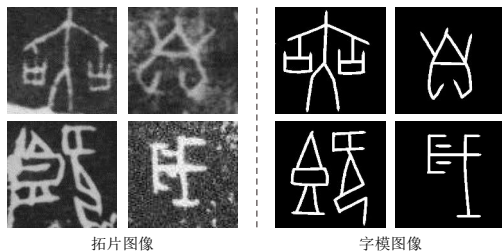


图1 两类甲骨文图像

早期的甲骨文识别方法依赖基于曲率直方图的傅里叶描述符^[3]、无向图^[4]、分块直方图^[5]等人工设计的特征辅助识别,这些方法高度依赖专业知识且需手动构建复杂的集成系统,限制了模型的泛化能力和识别性能.随着深度学习的发展,研究者将其应用于甲骨文识别. Huang 等^[6]贡献了首个大规模的甲骨文拓片数据集,并使用卷积神经网络取得了较好的识别性能.后续方法^[7,8]致力于对甲骨文中数据不平衡的问题进行改进.然而,这些方法都受制于闭集假设,即要求测试集中的类别都要在训练集里出现过,导致模型无法识别新类别字符.在现实场景中,随着更多新甲骨材料的出土,新类别的字符不断涌现,使得闭集甲骨文识别方法

面临挑战.对此,Zhang 等^[9]提出零样本甲骨文识别,通过将拓片图像与字模图像库投影到视觉特征空间进行匹配,找到与拓片图像最相似的字模图像,由此得到拓片图像的分类信息.然而,甲骨文字符类内方差大,即使是同一字符,其局部细节也变化多样,如图2所示,不同书写风格和文字结构的不稳定性导致视觉上的显著差异,这种差异对视觉匹配产生干扰,从而影响识别性能.

为了能够应对每类字形在局部细节上变化带来的干扰,模型需要更多地关注到类内不同字形下共享的特征.我们从借助语义信息帮助模型学习更好的视觉特征的研究^[10-12]得到新的启发,语义特征关注的是每个类别全局信息,而非视觉上的局部细节.以图2给出的字符样本为例,尽管他们在细节上可能呈现出风格变化、结构差异、字形残缺等特点,但在语义层面都属于“乡”字.

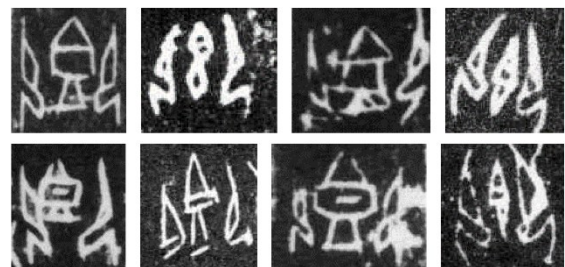


图2 甲骨文字符的类内方差大特点

根据上述启发,我们提出在零样本甲骨文识别任务中引入语义信息,帮助模型更多地关注到相同字符类别下不同字形的共享特征而非局部细节上的差异,从而解决因甲骨文类内方差大影响视觉匹配的问题.本文主要贡献包括:(1)我们提出一种两阶段的语义增强的零样本甲骨文字符识别方法,其包括域无关的字符语义学习阶段和语义增强的字符图像视觉匹配阶段,旨在提取字符的共享特征并增强视觉匹配性能.

(2)在第一阶段,我们引入对比式视觉-语言预训练模型 CLIP^[13],通过提示学习从甲骨文拓片和字模图像中挖掘语义信息,解决甲骨文字符缺乏语义的问题。(3)在第二阶段,我们通过两个分支分别引导模型关注类内共享特征和类间差异特征。第一个分支使用对比学习^[14-17]将不同字形的视觉特征对齐到字符语义,引导模型关注类内共享的关键性特征,第二个分支使用 N-Pair 损失^[18]引导模型关注区分不同类别的差异性特征,进一步增强模型对不同字符类别的区分性。(4)我们在拓片数据集 OBC306 和字模数据集 SOC5519 上展开全面的实验,充分验证了我们模型的有效性和鲁棒性。

2 相关工作

2.1 零样本学习

零样本学习^[19]通过借助辅助信息将知识从已知类别迁移到未知类别,实现对未见过类别的识别,主要分为三类:基于属性^[20-24]、嵌入^[25-29]和生成的方法^[30-32]。基于属性的方法为每个类别标注属性作为辅助信息,预测待识别样本的属性集合,通过属性匹配完成分类。然而,为每个类别标注属性过程复杂,特别是在甲骨文领域,标注属性需要大量专家知识,难以应用。基于嵌入的方法将语义嵌入作为辅助信息,通过在已知类别中学习视觉和语义嵌入的对齐,并将这种对齐能力泛化到未知类别中,从而通过视觉和语义嵌入的匹配实现零样本识别。然而,其面临模态鸿沟等问题,因为视觉和语义侧重表达的信息不同,两者嵌入无法实现绝对对齐,降低了识别性能。基于生成的方法把语义作为辅助信息,以语义为控制条件生成未知类别样本,从而转换为闭集监督学习问题。然而此类方法在识别新类别时需要重新训练,增加了训练成本,不适合甲骨文这种陆续发现新文字的场景。本文方法与基于嵌入的方法相似,利用语义嵌入作为辅助,不同之处在于,本文通过语义增强视觉特征提取能力,再通过拓片和字模图像的视觉匹配来实现零样本识别,避免了模态鸿沟的问题。

在文本识别领域中,当前研究工作者主要研究零样本中文字符识别,包括基于结构和视觉匹配的方法。基于结构的方法^[33-37]以每个字符的结构表示,如偏旁部首^[33-36]、笔画序列^[37]为辅助信息,但获取字符的结构表示需要很强的文字领域知识,且甲骨文作为早期文字还没有稳定的结构,因此不适用。基于视觉匹配的方法^[38-41]则以字符的印刷体模板作为辅助信息,在缺乏字符结构表示时尤为有效,因此成为主流方法。本文方法同样基于视觉匹配,并考虑了甲骨文字符类内方差大的特点。

2.2 甲骨文识别

近年来,随着深度学习的发展,研究者开始引入卷积神经网络进行甲骨文识别。早期方法^[42]使用卷积神经网络提取字符视觉特征,实现对甲骨文字模图像的分类。Huang 等^[6]构建了首个公开甲骨文拓片字符数据集,并使用多种主流卷积神经网络在该数据集上进行验证。后续方法致力于解决甲骨文数据的类别不平衡问题。Zhang 等^[7]通过使用三元组损失并采样难三元组训练模型,来更好地学习少数类的特征。Li 等^[8]使用 Mixup 数据增强^[43]对少数类样本进行增广。Han 等^[44]使用字符的在线轨迹数据以自监督方式预训练 BERT^[45]模型,而后让模型预测被掩盖的轨迹点位置来实现对少数类的增广。还有一些方法^[46,47]引入生成对抗网络,通过生成少数类的样本来应对类别不平衡问题。这些方法虽然取得不错的识别性能,然而都受限于闭集假设,无法识别新类别的甲骨文字符。相比之下,我们把研究重心放在零样本的甲骨文字符识别,更贴合甲骨文这种持续发现新字符的应用场景。

当前零样本甲骨文字符识别的研究较为稀少,Zhang 等^[9]通过视觉匹配实现了这一任务。他们把字模图像作为辅助信息,使用度量学习在特征空间中拉近同类别甲骨文拓片和字模图像之间的距离,同时推远不同类别的距离,构建了一个判别性强的特征空间。在测试时,通过在特征空间中衡量待识别拓片图像与字模库的相似度,选取最相似的字模图像并将其类别作为识别结果。然而,他们忽略了甲骨文字符有类内方差大的问题,导致匹配性能受限。本文方法通过语义引导模型关注全局特征,缓解了类内方差大的影响,提高了识别性能。

3 语义增强的零样本甲骨文字符识别方法

本文方法旨在借助字符语义,提升甲骨文拓片图像与字模图像之间的视觉匹配能力,从而在零样本场景下提高识别性能。假设已知类别和未知类别的字符集分别为 C_b 和 C_n ,其中 $C_b \cap C_n = \emptyset$ 达成零样本的设定。模型通过学习 C_b 中的拓片图像和字模图像的对齐关系,从而当输入 C_n 中的待识别甲骨文拓片图像时,模型能从 C_n 的字模中找到最相似的图像,并输出其类别作为识别结果。

图 3 展示了本文方法的总体框架,训练过程分为两阶段。第一阶段是域无关的字符语义学习。我们根据图像的域和字符类别分别设置可学习的提示信息,并结合固定模板构建提示信息。通过预训练的 CLIP 模型,利用其视觉和语义对齐的特性,从拓片和字模图像中优化提示信息向量。随后去除域提示信息,只保留字符类别提示信息,得到域无关的提示信息。第二阶段是

语义增强的字符图像匹配. 我们使用语义编码器将域无关提示信息编码为字符语义, 同时使用视觉编码器

分别对拓片和字模图像进行编码, 获得相应的视觉特征. 模型通过两个分支进行监督:

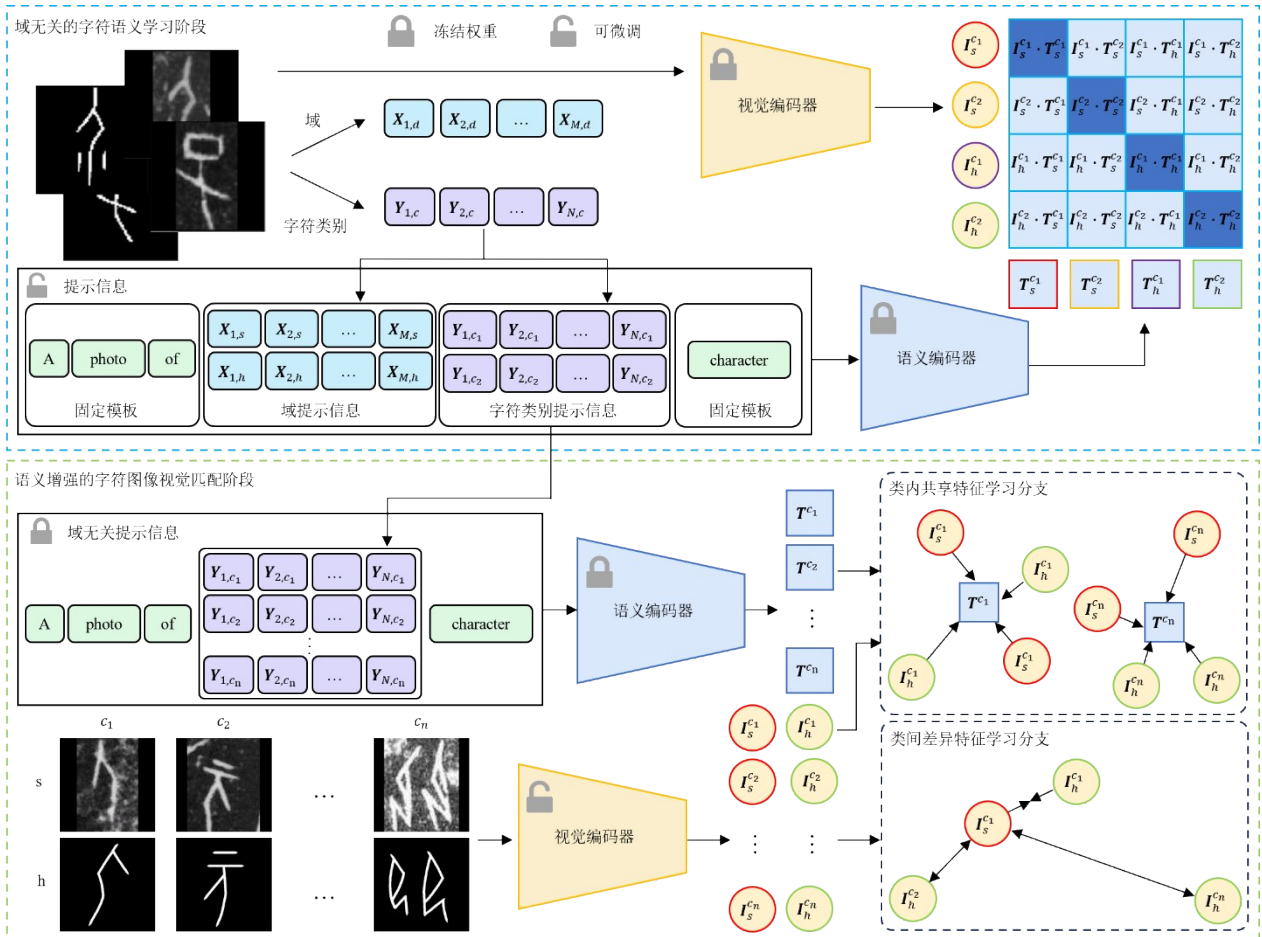


图3 语义增强的零样本甲骨文字符识别方法

一个分支通过对比学习将字形多样的视觉特征与语义特征对齐, 引导模型关注类内共享特征; 另一个分支通过 N-Pair 损失, 拉近同类拓片和字模图像的视觉特征, 推远不同类别的特征, 增强模型对类间差异的区分能力. 在测试阶段, 不再使用这两个分支, 而是通过计算拓片和字模之间的视觉特征相似度, 实现零样本甲骨文识别.

3.1 域无关的字符语义学习

甲骨文与通用物体不同, 其类别信息以索引而非文本描述表示, 如图2中的类别标签为“003059”, 无法直接从标签中提取语义. 受 CLIP 模型启发, 我们利用其对齐视觉和语义的能力, 通过甲骨文拓片和字模图像学习对应字符类别的语义. 在训练阶段, 我们根据域和字符类别为甲骨文图像构造可学习的提示信息, 并基于预训练的 CLIP 模型进行优化, 最终通过去除域提示信息, 获得域无关的提示信息.

3.1.1 CLIP 回顾

本文方法基于对比式视觉-语言预训练模型 CLIP^[13], 该模型由视觉编码器 (ViT 或 ResNet^[48] 架构) 和语义编码器 (Transformer 架构) 组成, 分别用于图像和提示信息的编码. 其中提示信息形如 “A photo of [CLASS]”, 而 [CLASS] 则会被替换成类别标签, 例如 “猫” 等. CLIP 通过对比学习, 最大化正样本对的相似性、最小化负样本对的相似性, 来实现视觉和语义的对齐. 然而, 由于甲骨文字符不像 “猫” 等通用物体拥有具体的类别信息, 无法通过直接构建形如 “A photo of [CLASS]” 的提示信息来得到甲骨文的字符语义. 因此, 我们采用提示学习, 借助 CLIP 的视觉和语义对齐能力, 为甲骨文图像构建可学习的提示信息, 从而挖掘出甲骨文字符的语义.

3.1.2 可学习的提示信息

如图1所示, 甲骨文图像有拓片和字模两类, 他们数据来源和分布都不同, 因此属于两个不同的域. 其中

甲骨文拓片图像因时间久远常伴有裂缝、杂质,背景复杂,笔画破损;字模图像则由专家临摹,背景干净且笔画修复.这种域差异在提取字符语义特征时可能混杂域特征,影响识别性能.为了解耦域与字符类别的语义特征,我们分别为域类别和字符类别设置可学习的提示信息,得到域无关的语义特征.

具体地,对域类别,我们设置可学习的域提示信息 $(X_{1,d}, X_{2,d}, \dots, X_{M,d})$, 其中 $X_{m,d} (m \in \{1, \dots, M\})$ 是与 CLIP 中的词向量相同维度的向量, $d \in \{s, h\}$, s 和 h 分别表示拓片和字模域.类似地,我们为字符类别也设置可学习的字符类别提示信息 $(Y_{1,c}, Y_{2,c}, \dots, Y_{N,c})$, $c \in C_b$.最后,我们根据所在域 d 和所属字符类别 c , 为 C_b 中的每张拓片和字模图像构建可学习的提示信息,其形式为“A photo of $(X_{1,d}, X_{2,d}, \dots, X_{M,d}, Y_{1,c}, Y_{2,c}, \dots, Y_{N,c})$ character”.

3.1.3 训练算法

为从甲骨文拓片和字模图像中学习最优的域和字符类别提示信息,我们通过提示学习优化可学习提示信息.具体地,使用预训练的 CLIP 视觉编码器对图像进行编码,得到视觉特征 $I_d^c \in \mathbf{R}^{1 \times D}$;同时使用 CLIP 语义编码器对提示信息进行编码,得到语义特征 $T_d^c \in \mathbf{R}^{1 \times D}$, 其中 $d \in \{s, h\}$, s 和 h 分别表示拓片和字模域, $c \in C_b$ 表示字符类别, D 为 CLIP 编码器输出的维度.在训练过程中,视觉和语义编码器被冻结,仅对提示信息进行优化.如图 4 所示,我们使用两个对比学习损失,使甲骨文图像的视觉特征与对应的提示信息语义特征对齐.将同域同类别的视觉特征和语义特征作为正样本对,拉近它们在特征空间中的距离;将不同域或不同类别的特征作为负样本对,推远它们的距离.损失函数如下:

$$L_{\text{stage1}} = \frac{1}{B} (L_{12T} + L_{T21}) \quad (1)$$

其中, L_{12T} 把视觉特征对齐到语义空间, L_{T21} 把语义特征对齐到视觉空间,两者的公式如下:

$$L_{12T} = \frac{-1}{|P(i)|} \sum_{i=1}^B \sum_{p \in P(i)} \log \frac{\exp(\langle I_{d_i}^{y_i}, T_{d_p}^{y_p} \rangle / \tau)}{\sum_{j=1}^B \exp(\langle I_{d_i}^{y_i}, T_{d_j}^{y_j} \rangle / \tau)} \quad (2)$$

$$L_{T21} = \frac{-1}{|P(t)|} \sum_{t=1}^B \sum_{p \in P(t)} \log \frac{\exp(\langle T_{d_t}^{y_t}, I_{d_p}^{y_p} \rangle / \tau)}{\sum_{j=1}^B \exp(\langle T_{d_t}^{y_t}, I_{d_j}^{y_j} \rangle / \tau)} \quad (3)$$

式中 B 是批大小,批中第 i 个视觉特征和第 t 个语义特征对应的字符类别分别为 $y_i, y_t \in \{c_1, c_2, \dots, c_n\}$, 对应的域分别为 $d_i, d_t \in \{s, h\}$, 对应的正样本集合分别为 $P(i)$ 和 $P(t)$, τ 表示温度系数, $\langle \cdot, \cdot \rangle$ 表示余弦相似度.

为更好地解耦域和字符类别特征,并提高训练效率,我们在提示学习中采用了特殊采样策略.其关键在

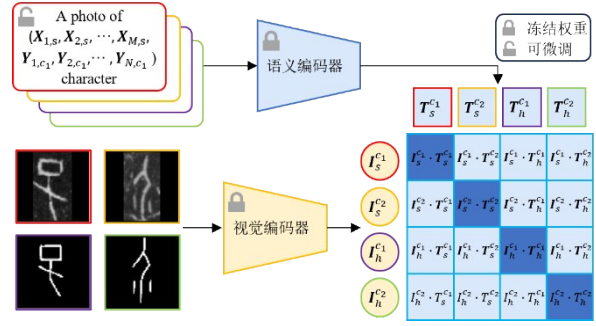


图 4 域无关的字符语义学习阶段

于采样更多不同域同类样本或相同域异类样本作为负样本.这样,当不同域但同类的样本作为负样本时,模型只能通过学习域提示信息区分它们,提升域提示信息的判别性;相同域但异类样本作为负样本时,模型则会学习到更具判别性的字符类别提示信息.相比于随机采样,这种策略避免了由于域样本和字符类别数量不平衡导致的负样本稀缺问题.

具体地,我们首先从 C_b 中随机采样出一个字符类别子集,并为每个类别均匀采样,使得每个类别的样本数量相等;同时在拓片和字模两个域上也均匀采样,保证两个域的样本数量也相等.这样,每个批次中对于任意一个域,有更多来自不同字符类别的负样本;对于任意一个字符类别,有更多来自不同域的负样本.通过增加了上述负样本对的数量,模型能学习到更好的域提示信息和字符类别文本提示信息,从而得到更准确的域无关语义特征.

3.1.4 字符语义的域解耦

前面提到,甲骨文拓片和字模之间的域差异会干扰字符类别语义的提取,因此我们需要解耦域和字符类别,获得域无关的字符语义.

具体地,经过第一阶段训练后,我们得到可学习的提示信息“A photo of $(X_{1,d}, X_{2,d}, \dots, X_{M,d}, Y_{1,c}, Y_{2,c}, \dots, Y_{N,c})$ character”.为消除域对语义特征的干扰,我们通过去除域提示信息 $(X_{1,d}, X_{2,d}, \dots, X_{M,d})$, 只保留字符类别提示信息 $(Y_{1,c}, Y_{2,c}, \dots, Y_{N,c})$, 从而得到域无关的提示信息“A photo of $(Y_{1,c}, Y_{2,c}, \dots, Y_{N,c})$ character”, 用于后续阶段提升模型的视觉特征提取能力.

3.2 语义增强的字符图像视觉匹配

得到字符语义后,我们利用其全局概括性引导视觉编码器提取类内共享特征,并通过度量学习推远异类图像,提取类间差异特征,从而提升零样本甲骨文识别性能.如图 5 所示,在这一阶段,未冻结的视觉编码器对拓片和字模图像编码得到视觉特征 I_d^c .随后,类内共享特征学习分支和类间差异特征学习分支分别引导模型关注类内共享特征和类间差异特征.交叉熵损失 L_{ce} 和 N-Pair 损失 L_{np} 用于监督两个分支的训练,这一阶

段的整体损失函数 L_{stage2} 为 L_{ce} 和 L_{np} 两者之间的加权和,如式(4)所示,其中 λ 是可调节的权重参数.

$$L_{\text{stage2}} = L_{\text{np}} + \lambda L_{\text{ce}} \quad (4)$$

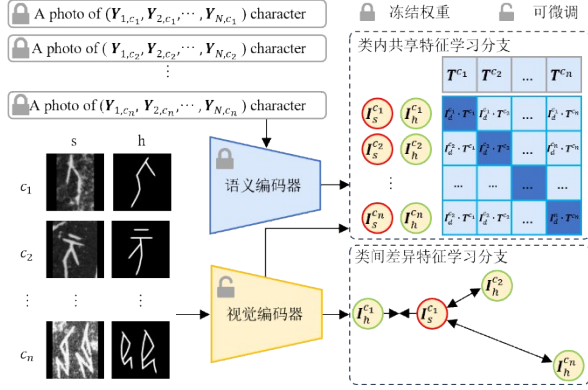


图5 语义增强的字符图像视觉匹配阶段

在测试阶段,我们无须这两个分支,而是依靠训练时视觉编码器学到的聚合类内特征和区分类间特征的能力进行分类.当给定待识别的拓片图像和字模图像集时,视觉编码器将这些图像编码到特征空间,通过最近邻匹配找到与待识别拓片图像最相似的字模图像,并将该字模的字符类别作为拓片图像的识别结果,从而实现零样本甲骨文字符识别.

3.2.1 类内共享特征学习

为缓解甲骨文类内方差大的影响,我们将多样化字形的甲骨文图像的视觉特征对齐到相应的字符类别语义特征,帮助模型学习同一类别不同字形的共享特征.由于字符语义是域无关的,视觉编码器能够专注于字符类别特征的学习,减少拓片和字模之间的域差异干扰.

具体地,得到域无关语义特征 T^c 和拓片图像与字模图像的视觉特征 I_s^c 与 I_h^c 后,我们使用对比学习,以视觉特征为锚点,将相同字符类别的语义特征作为正样本,把 C_b 中的其余类别的语义特征作为负样本,拉近正样本距离并推远负样本距离.在训练过程中,域无关语义特征 T^c 是固定的以提供稳定的语义,因此我们可以预先存储所有字符类别的语义特征.该过程通过交叉熵损失 L_{ce} 实现,其公式如下所示:

$$L_{\text{ce}} = -\frac{1}{B} \sum_{i=1}^B \log \frac{e^{(A_i^d)^T T^p}}{\sum_j e^{(A_j^d)^T T^p}} \quad (5)$$

其中 B 代表批的大小, p 为正样本的索引, $d \in \{s, h\}$.

3.2.2 类间差异特征学习

为增强模型对不同甲骨文字符类别的区分能力,我们在另一个分支中使用 N-Pair 损失进行监督,帮助视觉编码器捕捉字符类别间的差异特征,提升识别性能.

具体地,在训练过程第二阶段,我们按照 Kihyuk^[15] 的批构造策略,每个批次随机采样 n 个字符类别,每个类别采样一张拓片图像和字模图像.通过视觉编码器对这些图像编码后,使用 N-Pair 损失函数 L_{np} 进行监督,如式(6)所示.

$$L_{\text{np}} = \frac{1}{n} \sum_{i=1}^n \log \left(1 + \sum_{j \neq i} \exp \left((I_s^{y_i})^T I_h^{y_j} - (I_s^{y_i})^T I_h^{y_i} \right) \right) \quad (6)$$

4 实验结果及分析

4.1 数据集

OBC306^[6] 是目前最大的甲骨文拓片图像数据集,包含 306 个字符类别的 309 551 张图像,存在明显类别不平衡现象,样本最多的类别有 25 898 个,最少的仅有 1 个.

SOC5519^[49] 是字模图像类别最多的数据集,包含来自 5 491 类的 44 868 张字模图像,也存在类别不平衡,样本最多的类别有 287 个,最少的仅 1 个,值得注意的是, SOC5519 的字模图像可以与 OBC306 中同类别的拓片图像配对使用.

4.2 评估指标

我们使用甲骨文拓片图像的识别准确率 Acc 作为评估模型识别性能好坏的指标,具体如式(7)所示.

$$\text{Acc} = \frac{N_a}{N_t} \quad (7)$$

其中 N_t 为待识别的甲骨文拓片图像总数, N_a 为正确识别的图像数量.值得一提的是,在零样本实验中待识别的甲骨文拓片图像全部来自于未知类别,而在闭集识别实验中则全部来自于已知类别.

4.3 实现细节

在实验中,域和字符类别的可学习提示信息的词符数量 M 和 N 均设为 16, 第一阶段损失函数中对比学习损失中的温度系数 τ 设为 0.01, 第二阶段损失函数中交叉熵损失权重 λ 设为 3. 所有的字模和拓片图像尺寸都调整为 64×64 . 我们使用 CLIP-ResNet50 预训练模型作为特征编码器,并通过线性内插改变其内部的位置编码以适应图像尺寸,最终把图像编码为 $D = 1\,024$ 维的特征向量.在第一个训练阶段,我们使用 Adam 优化器^[50], 批次大小设置为 256, 迭代 20 轮, 初始学习率设置为 3.5×10^{-4} , 并使用余弦衰减策略^[51] 逐步衰减至 1×10^{-6} , 此阶段冻结 CLIP 的视觉编码器和语义编码器, 只更新可学习的提示信息向量. 第二个训练阶段仍使用 Adam 优化器, 批次大小设置为 256, 迭代次数为 15 轮, 初始学习率设置为 5×10^{-4} , 并于第 10 和 13 轮衰减至原来的 1/10, 期间冻结 CLIP 的语义编码器和可学习文本描述, 只更新 CLIP 的视觉编码器.

在零样本实验中,我们参照 Zhang 等^[9]的设定,使用 OBC306 数据集的 241 个字符类别子集进行实验. 这 241 类来自 Wang 等^[52]的公开数据,其中随机选取 200 类拓片和字模图像样本作为训练集,剩余 41 类作为测试集.

在闭集实验中,我们同样参照 Zhang 等^[9]的设定,使用相同的 241 个字符类别进行实验. 对于拓片图像数少于 900 的字符类别,随机选取 2/3 的样本作为训练集;对于拓片图像多于 900 的字符类别,随机选取 600 个样本加入到训练集,其余图像作为测试集. 此外,训练和测试时,均提供 241 类字符类别对应的字模图像作为辅助信息.

4.4 实验结果

4.4.1 零样本甲骨文字符识别结果

在零样本设定下,我们测试了模型对训练集中未出现的新类别甲骨文字的识别准确率,并与代表性的零样本甲骨文字符识别方法^[9]和基于视觉匹配的零样本中文字符方法^[39-41]进行对比. 为适应甲骨文识别场景,我们将手写或场景中文字符图像替换为甲骨文拓片图像,印刷体中文字符图像替换为甲骨文字模图像. 实验结果如表 1 所示,本文方法相比现有方法具有显著优势,尤其相比于零样本甲骨文字符识别方法^[9],性能提升超过 25%. 图 6 展示了拓片-字模图像视觉匹配的可视化结果. 第一列为待识别的拓片图像,其余为各方法中与之最相似的 5 张字模图像,按相似度排序,相同类别用绿框标出,不同类别用红框标出.

可以看到,当前最先进的视觉匹配方法 OpenCCD 容易因局部细节相似而误识别,如第一行中因字符顶部的“尖角”形状相似,导致错误匹配. 而本文方法通过语义引导,能关注同类别不同字形的共享特征,忽略局部干扰,正确进行匹配. 此外,本文方法中最相似的五

个字模图像均来自正确类别,尽管它们与待识别图像存在字形差异,表明方法的稳定性和抗干扰能力较强,识别性能优越.

表 1 不同方法的零样本甲骨文字符识别结果对比 单位:%

方法	识别准确率
文献 ^[9]	62.10
Deep Matching Network ^[39]	73.70
OSOCR ^[40]	76.34
OpenCCD ^[41]	80.99
本文方法	87.76

4.4.2 闭集甲骨文字符识别结果

在零样本甲骨文字符识别中,测试集全部来自新类别,因此仅评估模型对新类别的识别能力. 为验证方法在已知类别上的有效性,我们进行了闭集实验,并与其他零样本甲骨文字符识别方法^[9]及基于视觉匹配的零样本中文字符方法^[39-41]进行比较. 实验中将中文字符图像替换为甲骨文拓片和字模图像. 表 2 结果显示,本文方法在闭集设定下仍具最佳性能,这是因为甲骨文字符的类内方差大在零样本和闭集识别中都存在,通过语义引导模型关注同类字符的共享特征,减轻了这一影响. 注意到闭集识别准确率低于零样本设定,这是由于闭集测试集包含 241 类字符,而零样本测

试集仅包含 41 类,类别规模较大增加了识别难度.

表 2 不同方法的闭集甲骨文字符识别结果对比 单位:%

方法	识别准确率
文献 ^[9]	82.10
Deep Matching Network ^[39]	80.41
OSOCR ^[40]	81.17
OpenCCD ^[41]	83.61
本文方法	84.85

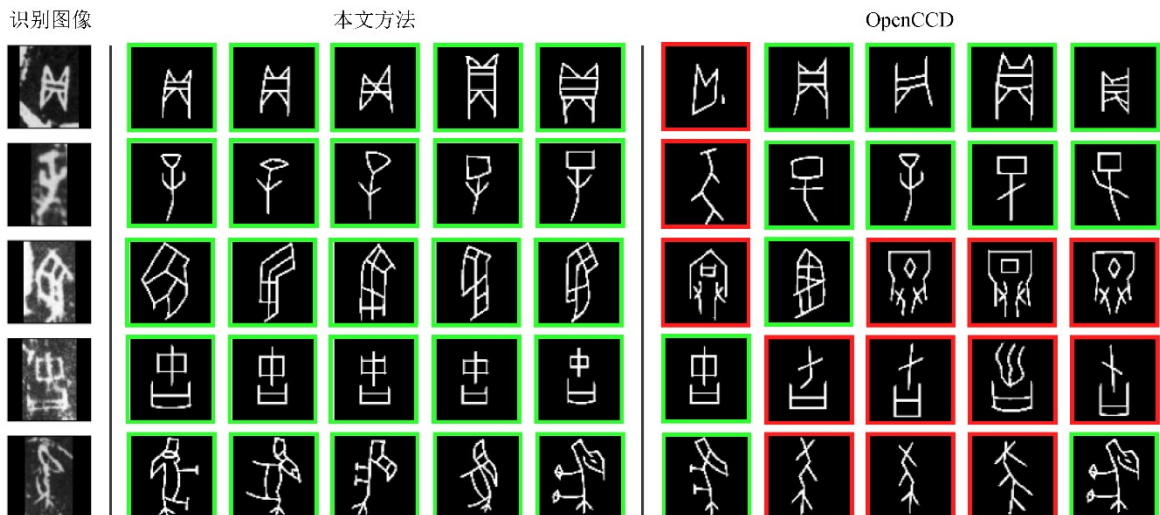


图 6 本文方法定性分析结果

4.4.3 消融实验

本节在零样本设定下进行消融实验,选出最佳的交叉熵损失权重超参数 λ ,以及域和字符类别的可学习提示信息中词符数量超参数 M 和 N ,并验证方法中各个部分的有效性.

(1) 交叉熵损失权重 λ

我们通过调整交叉熵损失权重超参数 λ 来控制语义引导的作用大小,图7显示了 λ 对零样本甲骨文识别准确率的影响.

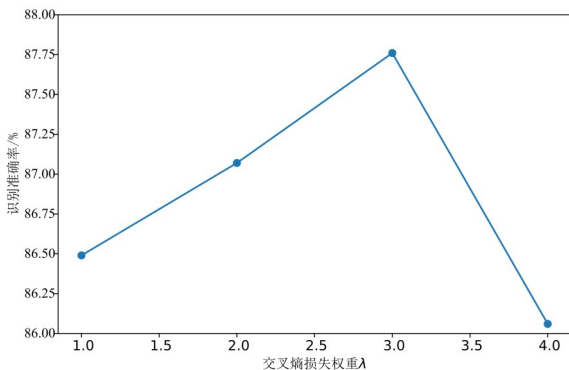


图7 交叉熵损失权重与识别准确率关系曲线图

当 λ 较小时,随着其增大,识别准确率提升,说明语义引导能帮助模型关注类内共享特征,缓解类内方差大的问题.然而,当 λ 过大后,模型过多依赖语义引导,减少了对字符间差异特征的提取,导致识别性能下降.因此本文取适中的 $\lambda=3.0$,以在提取类内共享特征和类间差异特征之间取得平衡.

(2) 域和字符类别的可学习提示信息中词符数量 M 和 N

我们分别为 M 和 N 设置不同的词符长度,以找到最优的参数组合.实验结果如表3所示,当 $M=16$ 且 $N=16$ 时,模型取得了最优的性能,而在其他设定下均出现一定程度的性能衰减.因此本文实验中,采用 $M=16, N=16$.

表3 不同词符数量下的零样本甲骨文识别准确率 单位:%

词符数量	$N=4$	$N=8$	$N=16$
$M=4$	86.52	86.28	87.16
$M=8$	85.60	86.38	85.96
$M=16$	87.01	86.74	87.76

(3) 仅从拓片或字模图像中提取语义

在域无关的字符语义学习阶段,我们使用预训练的CLIP模型从甲骨文拓片和字模图像中提取语义信息.为验证该方法的有效性,我们尝试仅从拓片或字模图像中提取语义,实验结果如表4所示.仅从拓片和字模图像中提取语义的识别准确率分别为85.70%和

85.83%,均低于本文方法.这是因为仅从拓片提取语义时,易受噪声干扰,而仅从字模提取时,样本数量少导致字形多样性不足.此外,单域提取语义会混杂域特征,影响语义增强效果.

(4) 解耦域和字符类别

本文通过为域类别和字符类别分别设置可学习提示信息,实现两者的解耦,从而获得域无关的字符语义.为验证该解耦方法的有效性,我们不单独为域设置可学习的上下文向量,即令第一阶段训练过程中 $M=0$,以验证域解耦方法的作用.结果如表4所示,去除域解耦后,识别性能下降1.83%.这表明域特征与字符类别特征的解耦能够有效减轻拓片与字模图像之间的域差异干扰,提升模型的识别性能.

(5) 特殊采样策略

特殊采样策略旨在应对拓片和字模域样本数量不平衡及字符类别不平衡问题,通过均匀采样将更多字模样本和少数类样本引入批中,帮助模型学习域和字符类别的提示信息.为验证该策略的有效性,采用随机采样方式进行对比实验.结果如表4所示,去除特殊采样策略后,识别准确率从87.76%降至85.66%,表明均匀采样有助于将更多有价值的负样本纳入批中,提升模型对域提示信息和字符类别提示信息的学习效果.

(6) 两个特征学习分支

在训练的第二阶段,本文通过类内共享特征学习分支和类间差异特征学习分支,分别引导模型关注同类字符的共享特征和区分不同字符的差异特征.为验证两个分支的有效性,我们去除其中一个分支进行实验.表4结果显示,去除类内共享特征学习分支后,识别准确率显著下降,表明字形变化带来的视觉差异影响了拓片与字模图像的匹配,而语义引导则帮助模型关注共享特征,减轻类内方差的影响.去除类间差异特征学习分支后,识别准确率也有所降低,验证了N-Pair损失有助于更好地区分不同字符,从而提升识别性能.

表4 方法中各模块的有效性验证实验 单位:%

方法	识别准确率
本文方法	87.76
仅从拓片图像中提取语义	85.70
仅从字模图像中提取语义	85.83
去除域解耦	85.93
去除特殊采样策略	85.66
去除类内共享特征学习分支	83.42
去除类间差异特征学习分支	86.11

4.4.4 探究训练字符类别数量的影响

为了探究训练字符类别数量对模型识别新类别能力的影响,我们固定41类作为测试字符,并从剩余200类中随机采样50、100、150、200类字符作为训练集进行

零样本甲骨文字符识别实验. 结果如图 8 所示, 本文方法即使在训练字符类别较少时, 仍能达到 71.48% 的识别准确率, 显著优于 Zhang 等人^[9]的方法. 此外, 随着训练类别数量增加, 模型的识别准确率逐渐提升, 证明更多类别的语义特征有助于进一步提高模型性能.

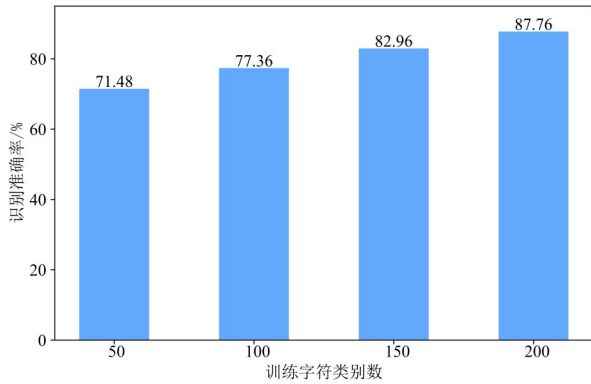


图 8 训练字符类别数与识别准确率关系曲线图

5 结论

本文提出了一种语义增强的零样本甲骨文字符识别方法, 其能够在不需要新类别拓片样本的情况下, 通过提供专家临摹的新类别字模图像, 实现对甲骨文拓片图像的识别. 该方法适用于甲骨文识别这种陆续发现新字符类别的场景. 我们借助 CLIP 模型, 从甲骨文拓片和字模图像中学习字符的语义特征, 并通过语义引导模型关注同类字符不同字形下的共享特征, 从而缓解甲骨文字符类内方差大的问题. 此外, 方法中引入了类间差异特征学习分支, 在视觉特征空间中将拓片图像与多个异类字模图像的特征推远, 从而提高字符类别的区分性. 多种实验设置下的结果证明了该方法的优越性和鲁棒性.

参考文献

- [1] 史先进, 曹爽, 张重生, 等. 基于锚点的字符级甲骨图像自动标注算法研究[J]. 电子学报, 2021, 49(10): 2020-2031.
SHI X J, CAO S, ZHANG C S, et al. Research on automatic annotation algorithm for character-level oracle-bone images based on anchor points[J]. Acta Electronica Sinica, 2021, 49(10): 2020-2031. (in Chinese)
- [2] 张重生, 王斌. 基于序列相似性计算的甲骨残片缀合算法[J]. 电子学报, 2023, 51(4): 860-869.
ZHANG C S, WANG B. Oracle bone fragments conjugation based on sequence matching[J]. Acta Electronica Sinica, 2023, 51(4): 860-869. (in Chinese)
- [3] 祁友杰, 朱恩. 一种非闭合曲线的傅里叶描述新算法[J]. 东南大学学报(自然科学版), 2014, 44(5): 886-890.
QI Y J, ZHU E. New Fourier description of non-closed curve[J]. Journal of Southeast University (Natural Science Edition), 2014, 44(5): 886-890. (in Chinese)
- [4] 顾绍通. 基于拓扑配准的甲骨文字形识别方法[J]. 计算机与数字工程, 2016, 44(10): 2001-2006.
GU S T. Identification of oracle-bone script fonts based on topological registration[J]. Computer & Digital Engineering, 2016, 44(10): 2001-2006. (in Chinese)
- [5] 刘永革, 刘国英. 基于 SVM 的甲骨文字识别[J]. 安阳师范学院学报, 2017(2): 54-56.
LIU Y G, LIU G Y. Oracle bone inscription recognition based on SVM[J]. Journal of Anyang Normal University, 2017(2): 54-56. (in Chinese)
- [6] HUANG S P, WANG H B, LIU Y G, et al. OBC306: A large-scale oracle bone character recognition dataset[C]//2019 International Conference on Document Analysis and Recognition (ICDAR). Piscataway: IEEE, 2019: 681-688.
- [7] ZHANG Y K, ZHANG H, LIU Y G, et al. Oracle character recognition by nearest neighbor classification with deep metric learning[C]//2019 International Conference on Document Analysis and Recognition (ICDAR). Piscataway: IEEE, 2019: 309-314.
- [8] LI J, WANG Q F, ZHANG R, et al. Mix-up augmentation for oracle character recognition with imbalanced data distribution[C]//Document Analysis and Recognition — ICDAR 2021. Cham: Springer International Publishing, 2021: 237-251.
- [9] 张颐康, 张恒, 刘永革, 等. 基于跨模态深度度量学习的甲骨文字识别[J]. 自动化学报, 2021, 47(4): 791-800.
ZHANG Y K, ZHANG H, LIU Y G, et al. Oracle character recognition based on cross-modal deep metric learning[J]. Acta Automatica Sinica, 2021, 47(4): 791-800. (in Chinese)
- [10] CAI S T, QIU L P, CHEN X J, et al. Semantic-enhanced image clustering[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(6): 6869-6878.
- [11] 黄俊场, 陈宏辉, 王嘉宝, 等. 多域字符距离感知的场景文本图像超分辨率重建[J]. 电子学报, 2024, 52(7): 2262-2270.
HUANG J Y, CHEN H H, WANG J B, et al. Scene text

- image super-resolution reconstruction based on perceiving multi-domain character distance[J]. *Acta Electronica Sinica*, 2024, 52(7): 2262-2270. (in Chinese)
- [12] WANG G R, TANG Y S, LIN L, et al. Semantic-aware auto-encoders for self-supervised representation learning [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 9654-9665.
- [13] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[EB/OL]. (2021-02-26) [2024-04-02]. <https://arxiv.org/abs/2103.00020v1>.
- [14] KHOSLA P, TETERWAK P, WANG C, et al. Supervised contrastive learning[J]. *Advances in Neural Information Processing Systems*, 2020, 33: 18661-18673.
- [15] DAI G, ZHANG Y F, WANG Q F, et al. Disentangling writer and character styles for handwriting generation[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 5977-5986.
- [16] LIN Z H, LI J R, DAI G, et al. Contrastive representation enhancement and learning for handwritten mathematical expression recognition[J]. *Pattern Recognition Letters*, 2024, 186:14-20
- [17] DAI G, ZHANG Y F, KE Q H, et al. One-shot diffusion mimicker for handwritten text generation[EB/OL]. (2024-09-26)[2024-04-02]. <https://arxiv.org/abs/2409.04004v2>.
- [18] KHOSLA P, TETERWAK P, WANG C, et al. Supervised contrastive learning[J]. *Advances in neural information processing systems*, 2020, 33: 18661-18673.
- [19] 张鲁宁, 左信, 刘建伟. 零样本学习研究进展[J]. *自动化学报*, 2020, 46(1): 1-23.
- ZHANG L N, ZUO X, LIU J W. Research and development on zero-shot learning[J]. *Acta Automatica Sinica*, 2020, 46(1): 1-23. (in Chinese)
- [20] LAMPERT C H, NICKISCH H, HARMELING S. Learning to detect unseen object classes by between-class attribute transfer[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2009: 951-958.
- [21] LAMPERT C H, NICKISCH H, HARMELING S. Attribute-based classification for zero-shot visual object categorization[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(3): 453-465.
- [22] JAYARAMAN D, GRAUMAN K. Zero-shot recognition with unreliable attributes[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 3464-3472.
- [23] FU Y W, HOSPEDALES T M, XIANG T, et al. Learning multimodal latent attributes[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(2): 303-316.
- [24] HUANG S, ELHOSEINY M, ELGAMMAL A, et al. Learning hypergraph-regularized attribute predictors[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015: 409-417.
- [25] FROME A., CORRADO G. S., SHLENS J., et al. DeViSE: A deep visual-semantic embedding model[C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2013: 2121-2129.
- [26] SOCHER R, GANJOO M, MANNING C D, et al. Zero-shot learning through cross-modal transfer[C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2013: 935-943.
- [27] SHIGETO Y, SUZUKI I, HARA K, et al. Ridge regression, hubness, and zero-shot learning[C]//Machine Learning and Knowledge Discovery in Databases. Cham: Springer International Publishing, 2015: 135-151.
- [28] YANG Y X, HOSPEDALES T M. A unified perspective on multi-domain and multi-task learning[EB/OL]. (2014-11-23)[2024-04-02]. <https://arxiv.org/abs/1412.7489v3>.
- [29] ZHANG L, XIANG T, GONG S. Learning a deep embedding model for zero-shot learning[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 3010-3019.
- [30] KODIROV E, XIANG T, GONG S G. Semantic Autoencoder for zero-shot learning[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 4447-4456.
- [31] XIAN Y Q, LORENZ T, SCHIELE B, et al. Feature generating networks for zero-shot learning[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition

- dition. Piscataway: IEEE, 2018: 5542-5551
- [32] SHEN Y M, QIN J, HUANG L, et al. Invertible zero-shot recognition flows[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020: 614-631.
- [33] WANG W C, ZHANG J S, DU J, et al. Denseran for offline handwritten Chinese character recognition[C]//2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). Piscataway: IEEE, 2018: 104-109.
- [34] GAN J, CHEN Y, HU B, et al. Characters as graphs: Interpretable handwritten Chinese character recognition via Pyramid Graph Transformer[J]. Pattern Recognition, 2023, 137: 109317.
- [35] WANG T W, XIE Z C, LI Z, et al. Radical aggregation network for few-shot offline handwritten Chinese character recognition[J]. Pattern Recognition Letters, 2019, 125: 821-827.
- [36] ZU X Y, YU H Y, LI B, et al. Chinese character recognition with augmented character profile matching[C]//Proceedings of the 30th ACM International Conference on Multimedia. New York: ACM, 2022: 6094-6102
- [37] CHEN J Y, LI B, XUE X Y. Zero-shot Chinese character recognition with stroke-level decomposition[C]//ZHOU Z H. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. California: International Joint Conferences on Artificial Intelligence Organization, 2021: 615-621.
- [38] AO X, ZHANG X Y, YANG H M, et al. Cross-modal prototype learning for zero-shot handwriting recognition [C]//2019 International Conference on Document Analysis and Recognition (ICDAR). Piscataway: IEEE, 2019: 589-594.
- [39] LI Z Y, WU Q, XIAO Y, et al. Deep matching network for handwritten Chinese character recognition[J]. Pattern Recognition, 2020, 107: 107471.
- [40] LIU C, YANG C, QIN H B, et al. Towards open-set text recognition via label-to-prototype learning[J]. Pattern Recognition, 2023, 134: 109109.
- [41] LIU C, YANG C, YIN X C. Open-set text recognition via character-context decoupling[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 4523-4532.
- [42] GUO J, WANG C H, ROMAN-RANGEL E, et al. Building hierarchical representations for oracle character and sketch recognition[J]. IEEE Transactions on Image Processing, 2016, 25(1): 104-118.
- [43] ZHANG H Y, CISSE M, DAUPHIN Y N, et al. Mixup: Beyond empirical risk minimization[EB/OL]. (2017-10-25)[2024-04-02]. <https://arxiv.org/abs/1710.09412v2>.
- [44] HAN W H, REN X L, LIN H Y, et al. Self-supervised learning of Orc-BERT augmentor for recognizing few-shot Oracle characters[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021: 652-668.
- [45] DEVLIN J, CHANG MW, LEE K, et al. BERT: Pre-training of deep bidirectional Transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis: Association for Computational Linguistics, 2019: 4171-4186.
- [46] YUE X B, LI H Y, FUJIKAWA Y, et al. Dynamic dataset augmentation for deep learning-based oracle bone inscriptions recognition[J]. ACM Journal on Computing and Cultural Heritage, 2022, 15(4): 1-20.
- [47] LI J, WANG Q F, HUANG K Z, et al. Towards better long-tailed oracle character recognition with adversarial data augmentation[J]. Pattern Recognition, 2023, 140: 109534.
- [48] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [49] HUANG H X, YANG D H, DAI G, et al. A2gan: Unpaired image translation for photographic ancient character generation[C]//Proceedings of the 30th ACM International Conference on Multimedia. New York: ACM 2022: 5456-5467.
- [50] ZHUANG Z Z, LIU Z H, LAM K M, et al. A new semi-automatic annotation model via semantic boundary estimation for scene text detection[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021: 257-273.

- [51] LOSHCHILOV I, HUTTER F. SGDR: Stochastic gradient descent with warm restarts[EB/OL]. (2016-08-13) [2024-04-02]. <https://arxiv.org/abs/1608.03983v5>.
- [52] WANG M, DENG W H, LIU C L. Unsupervised structure-texture separation network for oracle character recognition[J]. IEEE Transactions on Image Processing, 2022, 31: 3137-3150.

作者简介



刘宗昊 男,1999年出生,广东清远人.华南理工大学电子与信息学院硕士研究生.主要研究方向为计算机视觉、零样本学习、开集识别.
E-mail: zonghaoliu_work@foxmail.com



彭文杰 男,1998年出生,广东肇庆人.华南理工大学电子与信息学院博士研究生.主要研究方向为计算机视觉、度量学习、零样本学习.
E-mail: eepwj@mail.scut.edu.cn



代 港 男,1997年出生,湖北广水人.华南理工大学电子与信息学院博士研究生.主要研究方向为图像生成、手写体风格化、光学字符识别.
E-mail: eedaigang@mail.scut.edu.cn



黄双萍 女,1972年出生,湖南邵阳人.华南理工大学电子与信息学院教授.主要研究方向为计算机视觉、自然语言处理、多模态学习、人工智能生成内容、联邦学习.
E-mail: eehsp@scut.edu.cn



刘永革 男,1966年出生,安阳师范学院教授.安阳师范学院计算机与信息工程学院教授.主要研究方向为甲骨文信息处理、多媒体分析.
E-mail: ay_liuyongge@163.com