



基于深度学习技术的古彝文字图像搜集与整理方法

熊铁妞, 邱吉芳, 胡建

引用本文:

熊铁妞, 邱吉芳, 胡建. 基于深度学习技术的古彝文字图像搜集与整理方法[J]. *智能系统学报*, 2025, 20(4): 928-935.

XIONG Tieniu, QIU Jifang, HU Jian. Collection and sorting method of ancient Yi character images based on deep learning technology[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(4): 928-935.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202406036>

您可能感兴趣的其他文章

自步稀疏最优均值主成分分析

Sparse optimal mean principal component analysis based on self-paced learning
智能系统学报. 2021, 16(3): 416-424 <https://dx.doi.org/10.11992/tis.201911028>

多感知兴趣区域特征融合的图像识别方法

Image recognition method based on multi-perceptual interest region feature fusion
智能系统学报. 2021, 16(2): 263-270 <https://dx.doi.org/10.11992/tis.201906032>

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>

面向社群图像的显著区域检测方法

Salient region detection for social images
智能系统学报. 2018, 13(2): 174-181 <https://dx.doi.org/10.11992/tis.201706043>

基于自编码器的特征迁移算法

Feature transfer algorithm based on an auto-encoder
智能系统学报. 2017, 12(6): 894-898 <https://dx.doi.org/10.11992/tis.201706037>

行人重识别研究综述

Survey on pedestrian re-identification research
智能系统学报. 2017, 12(6): 770-780 <https://dx.doi.org/10.11992/tis.201706084>

DOI: 10.11992/tis.202406036

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20250221.0930.008>

基于深度学习技术的古彝文字图像搜集与整理方法

熊铁妞^{1,2}, 邱吉芳³, 胡建^{1,2}

(1. 西南民族大学 计算机系统国家民委重点实验室, 四川 成都 610225; 2. 西南民族大学 计算机与人工智能学院, 四川 成都 610225; 3. 西南民族大学 中国语言文学学院, 四川 成都 610225)

摘要: 古彝文字是中华文化的重要载体之一, 但人工搜集、整理大量古彝文字耗时耗力, 而且能辨识古彝文字的人已非常稀缺且越来越少, 这使得整理工作变得更为困难。对此, 本文提出一种基于深度学习技术的古彝文字图像搜集与整理的新思路。在古彝文字图像搜集方面, 通过目标检测模型得到每个古彝文字在彝文古籍图像中的位置, 据此在彝文古籍图像中截取出古彝文字图像, 实现古彝文字搜集。在古彝文图像整理方面, 首先根据规范彝文来源于古彝文的事实, 采用规范彝文字体文件自动生成彝文字图像用于构建数据集, 并将数据集应用于训练古彝文字图像特征算法, 这有效避免了目前因古彝文字数量庞大、异体字众多、整理尚未完成, 而尚无古彝文字图像数据集的问题; 然后, 通过匹配所搜集的古彝文字图像的特征与现已收录的古彝文字图像的特征的相似性, 判断所搜集的古彝文字图像是否已被收录, 从而整理出未收录的古彝文字图像。实验在多种典型的特征提取算法和相似性计算方式下进行, 实验结果验证了方法的有效性。

关键词: 深度学习; 古彝文字; 古籍; 图像处理; 相似度匹配; 特征提取; 目标检测; 数字化

中图分类号: TP391.4; TP391.1 文献标志码: A 文章编号: 1673-4785(2025)04-0928-08

中文引用格式: 熊铁妞, 邱吉芳, 胡建. 基于深度学习技术的古彝文字图像搜集与整理方法 [J]. 智能系统学报, 2025, 20(4): 928-935.

英文引用格式: XIONG Tieniu, QIU Jifang, HU Jian. Collection and sorting method of ancient Yi character images based on deep learning technology [J]. CAAI transactions on intelligent systems, 2025, 20(4): 928-935.

Collection and sorting method of ancient Yi character images based on deep learning technology

XIONG Tieniu^{1,2}, QIU Jifang³, HU Jian^{1,2}

(1. The Key Laboratory for Computer Systems of State Ethnic Affairs Commission, Southwest Minzu University, Chengdu 610225, China; 2. College of Computer Science and Artificial Intelligence, Southwest Minzu University, Chengdu 610225, China; 3. School of Chinese Language and Literature, Southwest Minzu University, Chengdu 610225, China)

Abstract: The ancient Yi script is one of the important carriers of Chinese culture. However, manually collecting and organizing a large amount of ancient Yi script is time-consuming and labor-intensive. Additionally, very few people can recognize ancient Yi script, and their numbers are dwindling, which makes the task even more difficult. In response to this, this paper proposes a new approach to collecting and organizing images of the ancient Yi script based on deep learning technology. For image collection, the object detection model is used to locate each ancient Yi character in the images of ancient Yi manuscripts, and the characters are extracted from these images accordingly. For image organization, because modern standardized Yi characters are derived from ancient Yi characters, standardized Yi character font files are used to generate images of the Yi characters automatically to construct a dataset. This dataset is then used to train an algorithm for extracting features of ancient Yi script images, which effectively addresses the current lack of an ancient Yi script image dataset due to the large number of characters, many variants, and incomplete organization. Subsequently, matching the features of the collected ancient Yi script images with those of already cataloged images enables determining whether the collected images have been previously recorded and thereby organizing uncatalogued ancient Yi script images. Experiments conducted with various typical feature extraction algorithms and similarity computation methods validate the effectiveness of this approach.

Keywords: deep learning; ancient Yi characters; ancient literatures; image processing; similarity matching; feature extraction; object detection; digitalization

收稿日期: 2024-06-21. 网络出版日期: 2025-02-21.

基金项目: 国家社会科学基金重大项目(19ZDA284); 西南民族大学中华民族共同体研究院团队项目(2024-GTT-TD17); 西南民族大学中央高校基本科研业务费专项基金项目(ZYN2023009).

通信作者: 胡建. E-mail: hujian@swun.edu.cn.彝文史称“爨文”“爨字”“爨书”“毕摩文”等, 是一种超方言的音节文字^[1]. 古彝文作为一种重要文字, 距今有 8 000 多年历史, 可与甲骨、苏美尔、埃及、玛雅、哈拉般等 5 种文字并列, 是世界

六大古文字之一,是灿烂中华文化的重要载体之一^[2]。搜集、整理古彝文字是研读彝文古籍、完善规范彝文的基础性工作,这对挖掘中华文化共同性、传承中华优秀传统文化具有重要意义。

目前已有一些古彝文字搜集与整理的成果。例如,《滇川黔桂彝文字集》由滇川黔桂彝文协作组广泛搜集了云南、四川、贵州、广西的彝文古籍文献中的单字,分卷汇集了共 87 000 多个古彝文字;《古彝文常用字典》收录彝语南部方言区峨山古彝文常用字近 4 500 个;《滇南彝文字典》收入彝语南部方言区中使用的古彝文单字 21 970 个;《简明彝汉字典(贵州本)》共搜集、整理贵州常用彝文字约 6 700 个。由于古彝文字繁多,存在众多异体字,每个字通常有多种不同的写法,同时在古籍数字化浪潮和信息化大背景下,一些新的彝文古籍被发现和公开,所以古彝文字搜集与整理工作仍任重道远。但是,目前古彝文字搜集工作往往采用手工方式提取古彝文字图像,这一过程费时费力;同时,古彝文字整理工作需要具备专业的古彝文字基础,但具有此基础的人已非常稀缺且越来越少,使得整理工作变得困难。所以,亟需一种自动化手段来辅助古彝文字搜集与整理。

随着以深度学习技术为核心的人工智能技术的发展,彝文信息处理领域已取得一些成果。例如,贾晓栋^[3]使用信息熵^[4]改进了一种基于密度的聚类算法应用于手写彝文字的标注,结合卷积神经网络实现手写彝文识别;王定旺^[5]基于循环神经网络^[6]实现了彝文联机手写识别;陈善雄等^[7]提出一种基于最大极值稳定区域^[8]和卷积神经网络的彝文古籍文献字符检测方法;陈善雄等^[9]以 Alpha-Beta 散度^[10]作为惩罚项对模型的输出神经元重新进行自编码实现了古彝文手写体的识别。但是这些研究均未解决古彝文字图像搜集

与整理的问题:文献[3]虽实现彝文古籍中的手写彝文字符的识别,但是所使用的数据集只包含 100 个类别,无法用于古彝文字整理;文献[5]未涉及古彝文字;文献[7]虽然涉及古彝文字,且可被用于搜集古彝文字图像,但针对的是字符识别需求下的彝文古籍复杂背景的问题,所采用字符检测方法相对复杂,且未涉及古彝文字图像整理;文献[9]虽然识别古彝文字,但只面向《西南彝志》^[11]中的 2 142 个常用古彝文字,因古彝文字集庞大和训练集构建困难而无法扩大到整个古彝文字范畴,故无法用于古彝文字图像整理。

本文受深度学习技术在图像分类、目标检测、图像检索等计算机视觉任务中应用的启发,提出了一种基于深度学习技术的古彝文字图像搜集与整理的新方法,并实验验证了其有效性。该方法采用目标检测技术实现古彝文字图像搜集,通过图像特征提取和特征相似性匹配实现古彝文字图像整理。这不仅能够为古彝文字图像的搜集和整理提供一种自动化方法,还可以为其他古文字的搜集、整理与研究提供宝贵的参考。

1 基本思路

古彝文图像搜集与整理的基本思路如图 1 所示,分为两个阶段:在古彝文字图像搜集阶段,选取彝文古籍图像作为古彝文字图像来源,通过古彝文字目标检测,得到每个古彝文字在彝文古籍图像中的位置,据此在彝文古籍图像中截取取出古彝文字图像,得到古彝文字图像集合;在古彝文字图像整理阶段,通过特征提取算法计算前述所搜集的古彝文字图像的特征和现已收录的古彝文字图像的特征,再通过计算前者与后者的相似性,判断所搜集的古彝文字图像是否已被收录,从而整理出未收录的古彝文字图像。

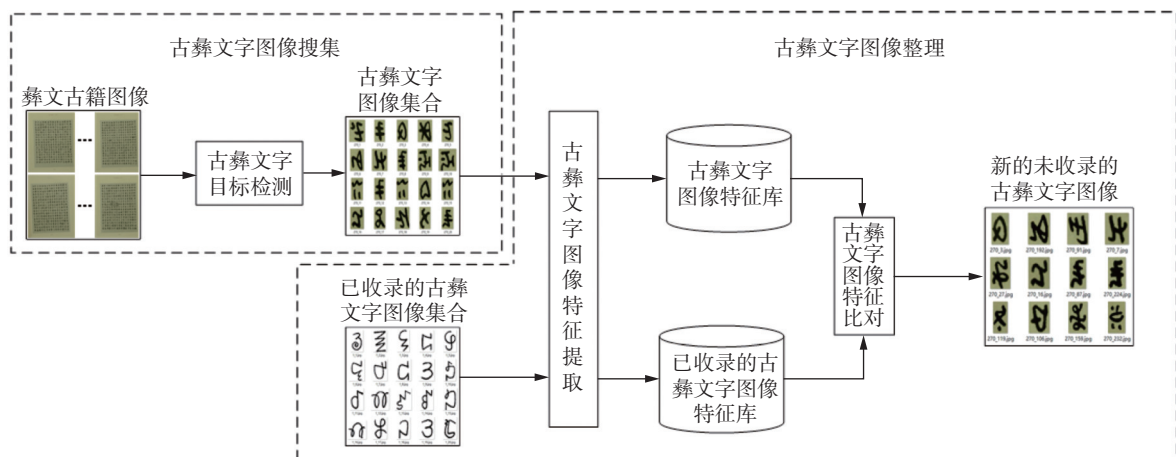


图 1 古彝文图像搜集与整理的基本思路

Fig. 1 Overall process framework for collecting and organizing Ancient Yi character images

具体地,本文在古彝文字图像搜集阶段,选取《彝文典籍集成》^[12]作为古彝文字图像来源,它是由国家出版基金资助的、总规模有 160 辑的、迄今为止最大的彝文古籍出版项目,其古彝文字丰富且图像质量高。在古彝文字图像整理阶段,已收录的古彝文字图像来自《滇川黔桂彝文字集》^[13],其由滇川黔桂彝文协作组历时 10 年完成,分卷汇集 4 个省区的共 87 000 多个古彝文字,并与古彝文字图像搜集阶段一样,进行古彝文字目标检测和截图,得到已收录的古彝文字图像集合。

2 古彝文字目标检测

目标检测是计算机视觉任务中的重要方向,旨在从图像中检测出目标对象的类别和位置^[14]。随着深度学习技术的发展,目标检测方法可分为传统方法和基于深度学习的方法两大类。在传统方法中,需要人工设计特征模板,通过特征模板对待检测图像进行特征提取,进行目标的分类和定位。相比之下,基于深度学习技术的目标检测算法通过深度学习模型来自适应学习生成检测任务对应的特征,可以有效地避免人工设计特征速度慢、在复杂场景下鲁棒性差的问题,已经取得了很好的应用效果,其中代表性算法有 R-CNN 系列、YOLO 系列等。

相应地,古彝文目标检测旨在从彝文古籍图像中检测出古彝文字类别和位置,其中优先运用基于深度学习的目标检测方法。因为古彝文字存在众多异体字,每个字通常有多种不同的写法,判断这些异体字类别存在困难,所以本文只把古彝文字目标检测中的类别分为彝文古籍图像背景和古彝文字两个类别,即不区分古彝文字类别,而只检测古彝文字位置。

基于深度学习技术的古彝文字目标检测流程如图 2 所示,分为古彝文字目标检测模型的训练和推理阶段。其中,在训练阶段,选取一定数量的、具有代表性的彝文古籍图像,通过在其上标注所有古彝文字的位置来构建古彝文字目标检测数据集。在推理阶段,作为古彝文字搜集来源的彝文古籍图像经古彝文字目标检测模型处理后,得到古彝文字在彝文古籍图像中的位置。该位置信息被用于从彝文古籍图像中截取古彝文字图像,从而实现古彝文字图像的自动搜集。

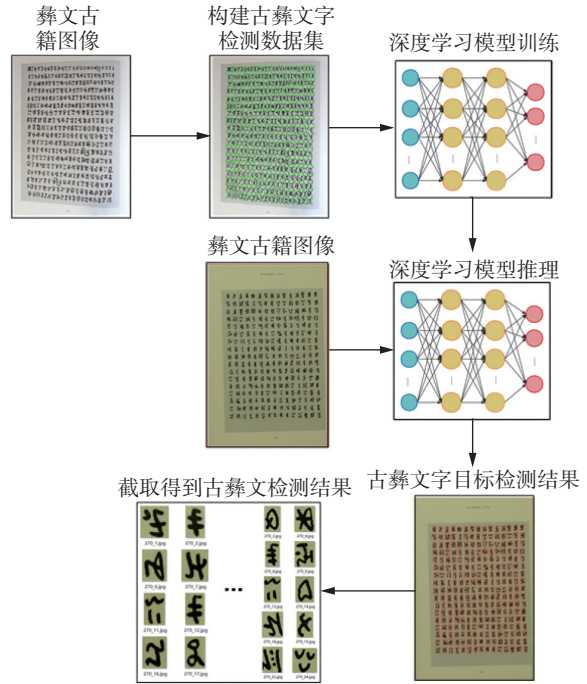


图 2 古彝文字目标检测流程
Fig. 2 Flow chart of object detection for Ancient Yi character

3 古彝文字图像特征提取

目前,深度学习技术已是提取图像特征常用且有效的手段。因此本文将用深度学习技术提取古彝文字图像的特征,并构建用于训练深度学习模型的数据集(下面简称古彝文字图像特征数据集)。理想的古彝文图像特征数据集要包含全部古彝文字的图像数据。然而,因古彝文字数量庞大、异体字众多、整理尚未完成,所以目前尚无全部古彝文字的图像数据,进而无法构建理想的古彝文图像特征数据集。如果只采集部分古彝文字的图像数据,古彝文字之间会因图像数量差异较大,而存在数据不平衡的问题。

为此,本文以解决古彝文字图像特征数据集构建问题为重点,提出如图 3 所示的基于深度学习技术的古彝文字图像特征提取流程。在古彝文字图像特征数据集构建中,基于规范彝文^[15]来源于古彝文的事实,采取了从规范彝文字体文件自动生成的思想,并因古彝文字中并不包含次高调字符和部首字符,将次高调字符和部首字符去掉。其中,从方正仿宋体、方正黑体、方正宋体、方正手写体、方正细黑体、方正圆头体等规范彝文字体文件生成规范彝文图像,并对规范彝文图像进行旋转、等比缩放、宽方向拉伸、腐蚀、膨胀、仿射变换、添加噪声、图像模糊等变换处理。

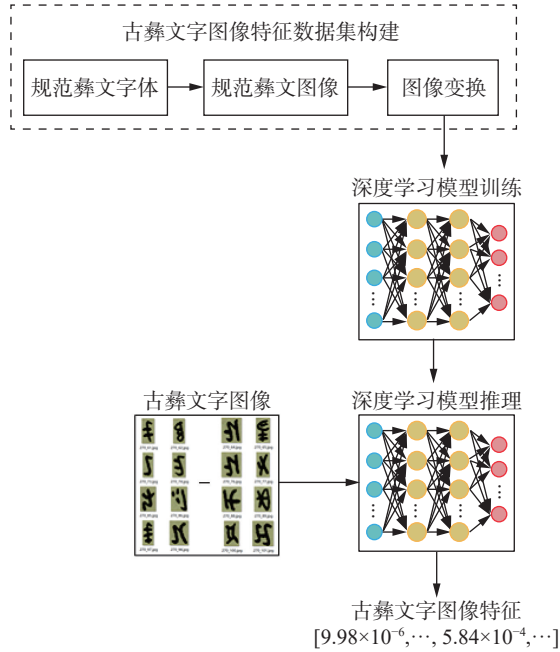


图3 古彝文字图像特征提取流程

Fig. 3 Process diagram for feature extraction of Ancient Yi character images

4 古彝文字图像特征比对

经过特征提取后,古彝文字图像均被特征向量描述。记所搜集的古彝文字图像为序列 $\text{Img} = [\text{img}_1, \text{img}_2, \dots, \text{img}_m, \dots, \text{img}_M]$, 其对应的古彝文字图像特征为序列 $\mathbf{F} = [f_1, f_2, \dots, f_m, \dots, f_M]$, 已收录的古彝文字图像为序列 $\text{Img}' = [\text{img}'_1, \text{img}'_2, \dots, \text{img}'_n, \dots, \text{img}'_N]$, 其对应的古彝文字图像特征为序列 $\mathbf{F}' = [f'_1, f'_2, \dots, f'_n, \dots, f'_N]$ 。其中, M 、 N 分别表示所搜集的和已收录的古彝文字图像总数, img_m 、 f_m 、 img'_n 和 f'_n 分别表示所搜集的第 m 张古彝文字图像及其特征、已收录的第 n 张古彝文字图像及其特征。

古彝文字图像特征比对是在 \mathbf{F} 与 \mathbf{F}' 之间进行比对。比对方法如算法 1 所示, 采用信息检索的思想: 把 \mathbf{F} 中的每个特征作为检索项, 与 \mathbf{F}' 中的所有特征进行相似度计算, 并按相似度从高到低进行排序; 当把 \mathbf{F} 中的所有特征均照此计算后, 得到古彝文图像特征比对结果为 M 组排序结果。其中, $S(f_m, f'_n)$ 表示计算特征向量 f_m 和特征向量 f'_n 之间的相似度。 $\text{Sort}([\text{Sim}_{m,1}, \text{Sim}_{m,2}, \dots, \text{Sim}_{m,N}])$ 表示对序列 $[\text{Sim}_{m,1}, \text{Sim}_{m,2}, \dots, \text{Sim}_{m,N}]$ 中的所有元素按从高到低顺序排序, 并返回如下序列 I_m : I_m 的长度为 N , 其中各元素是该元素在序列 $[\text{Sim}_{m,1}, \text{Sim}_{m,2}, \dots, \text{Sim}_{m,N}]$ 中的序号。 $[\text{Img}'_1[I_1], \text{Img}'_2[I_2], \dots, \text{Img}'_N[I_M]]$ 中的 $\text{Img}'_1[I_1]$ 是 Img' 中所有元素以序列 I_1 中的元素值为序号, 重新排列得到的序列, $\text{Img}'_2[I_2], \dots, \text{Img}'_N[I_M]$ 的含义依此类推。

算法 1 古彝文图像特征比对方法

输入 古彝文字图像特征库 \mathbf{F} 和 \mathbf{F}'

输出 古彝文字图像特征库 \mathbf{F} 和 \mathbf{F}' 比对结果

1) For $m = 1 \rightarrow M$ do

2) For $n = 1 \rightarrow N$ do

3) $\text{Sim}_{m,n} \leftarrow S(f_m, f'_n)$

4) $I_m = \text{Sort}([\text{Sim}_{m,1}, \text{Sim}_{m,2}, \dots, \text{Sim}_{m,N}])$

5) return $[\text{Img}'_1[I_1], \text{Img}'_2[I_2], \dots, \text{Img}'_N[I_M]]$

至此, 通过古彝文字图像特征比对, 可为所搜集的每个古彝文字图像, 返回按相似度从高到低顺序排列的已收录的古彝文字图像序列。如果所搜集的该古彝文字图像已在返回结果的前列, 则可迅速地判断其已被收录, 而无需人工遍历整个已收录的古彝文字图像集合才判断其是否已被收录。这将为人工整理古彝文字图像提供有益辅助。

5 实验仿真与结果分析

分为古彝文字图像搜集和古彝文字图像整理两个实验, 下面分别介绍。

5.1 古彝文字图像搜集实验

5.1.1 实验设置

实验所采用的古彝文字目标检测数据集构建于《彝文典籍集成·四川卷》(教育 2-1) 中的 30 张图像, 共有 8 052 个古彝文字目标, 训练集、验证集和测试集之间的比例为 0.81:0.09:0.1。其中, 为了增加彝文古籍图像的多样性, 这 30 张图像选自多个篇章, 其页码分别为 40—46、52—54、130—136、286、287、289、290、292、294—301。

实验所采用的古彝文字目标检测算法涵盖经典方法与最新兴起的算法, 具体包括双阶段经典算法 Faster R-CNN^[16]、单阶段 YOLO 系列的 YOLOv3 算法^[17]、YOLOv5 算法^[18]、YOLOv10 算法^[19]、以及无锚框单阶段的 FCOS 算法^[20]。其中, Faster R-CNN 作为经典的双阶段算法, 通过区域建议网络生成候选框, 并结合 RoI 池化与分类回归步骤, 实现了较高的检测精度。YOLO 系列作为单阶段目标检测算法的代表, YOLOv3 通过多尺度预测和 Darknet-53 主干网络实现速度与精度的平衡; YOLOv5 采用 CSPDarknet 架构进行模型轻量化设计, 配合自适应锚框调整策略, 在保证小目标检测性能的同时提升推理效率。YOLOv10 通过去除非极大值抑制 (NMS), 结合空间-通道解耦检测头等创新设计, 将实时检测速度大幅度提升, 同时维持高精度, 显著优于传统两阶段方法。FCOS 则采用全卷积无锚框设计, 通过逐像素预测直接回归目标边界框, 避免了锚框参数敏

感性问题,为单阶段目标检测算法提供了更简洁高效的解决方案。设置训练轮数为 200,训练批次大小为 2,输入图像大小为 832,优化器为 Adam, beta1 为 0.9,权重衰减系数为 0,初始学习率为 0.01,最终学习率为 0.0001。

5.1.2 实验结果与讨论

各算法在古彝文字目标检测上的 mAP(mean average precision)值如图 4 所示。

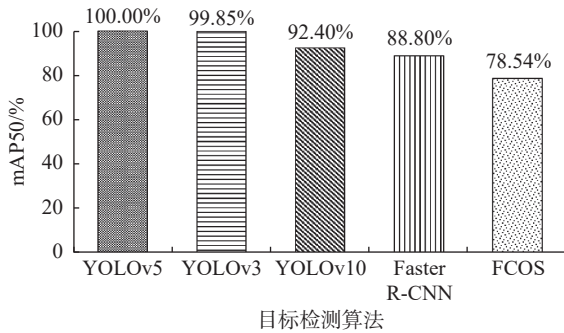


图 4 各算法在古彝文字目标检测上的 mAP 值

Fig. 4 MAPs of each algorithm in Ancient Yi character object detection

实验结果表明, YOLOv5 和 YOLOv3 表现最优, mAP50 分别达到了 100.00% 和 99.85%, 显示出在古彝文字目标检测任务中能很好地将彝文古籍图像中的古彝文字图像检测出并截取, 极大地提升了古彝文字图像搜集的效率和准确性。YOLOv10 紧随其后, 性能稍逊于 YOLOv3 但优于其他算法。相比之下, Faster R-CNN 和 FCOS 的性能较低, 分别仅为 88.80% 和 78.54%, 表明其在该检测任务中的适用性有限。图 5 是古彝文字目标检测结果样例和搜集的古彝文字图像样例。目标检测算法在古彝文字图像上的检测效果明显高于其在自然图像上的效果, 这是因为在所选彝文古籍中, 古彝文字目标之间几乎没有重叠、大小较为均匀、色彩较为单一, 而且没有区分古彝文字类别, 较自然图像中目标检测而言更为简单。



(a) 目标检测结果样例

(b) 搜集的古彝文字图像样例

图 5 古彝文字图像搜集结果

Fig. 5 Results of Ancient Yi character image collection

5.2 古彝文字图像整理实验

5.2.1 实验设置

已收录的古彝文字图像来自《滇川黔桂彝文字集·四川卷》, 其经古彝文字目标检测和截图后, 共有 20 525 张古彝文字字符图像。待整理的古彝文字图像是在《彝文典籍集成·四川卷》的教育 2-1 第 270 页上运用上述古彝文字图像搜集方法所搜集的 255 张古彝文字图像。经人工查找, 在这些待整理的古彝文字图像中, 共有 215 张被《滇川黔桂彝文字集》收录, 40 张未被收录。

目前尚没有可直接用于对古彝文字图像整理方法的性能评价指标。鉴于古彝文字图像整理方法的目的是提高整理古彝文字图像的自动化水平, 减少人工参与程度, 故定义为

$$\alpha = \frac{\sum_{m=1}^M a_m}{M}$$

式中 M 为待整理的古彝文字图像总数 (即上述 255)。如果第 m 张待整理的古彝文字图像的最相似的 K (其最大值为 20 525) 张已收录的古彝文字图像, 包含该张待整理的古彝文字图像, 则 a_m 为 1, 如果不包含, 则 a_m 为 0。 α 越大表示古彝文字图像整理方法在待整理的古彝文字图像中自动找到已被收录的古彝文字图像越多, 需要人工在已收录的古彝文字图像中去遍历式查找才能判断待整理的古彝文字图像是否被收录的情形就越少。

在古彝文字图像特征提取时, 为了缓解待整理的古彝文字图像与已收录的古彝文字图像之间的跨域问题, 本实验对这些古彝文字字符图像均作了预处理: 首先作二值化处理, 然后采用文献 [21] 提出的算法作骨架化处理。

为了对比古彝文字图像特征提取算法的影响, 选用属于深度学习技术的 VGG19^[22] 和 ResNet152^[23] 两种算法, 以及被广泛应用的、基于像素值统计的 Histogram^[24] 算法。VGG19 具有深度网络和小卷积核的特点, 能够提取丰富的细节特征, 适用于分析复杂的形状和纹理; ResNet152 通过残差连接解决了深层网络的梯度消失问题, 能够学习到更复杂和抽象的特征; 而 Histogram 算法则通过统计像素值分布提取特征, 计算简单且高效。其中, VGG19 和 ResNet152 的训练过程采用文献 [25] 中的预训练模型, 设置训练轮数大小为 30, 训练批次大小为 32, 学习率为 1×10^{-5} , 权重衰减系数为 5×10^{-4} , 选择交叉熵损失函数和 Adam 优化器; Histogram 算法的直方图箱数为 2, 计算区域直方图, 图像沿宽/高方向平均分割的子图

数为 8。

为了对比相似度算法的影响, 选用被广泛应用的曼哈顿距离^[26]、欧氏距离^[27]的平方以及余弦距离^[28]3 种。它们的计算式分别为

$$D_M = \sum_{i=1}^n |x_i - y_i|$$

$$D_E = \sum_{i=1}^n (x_i - y_i)^2$$

$$D_{\cos} = 1 - \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

式中 (x_1, x_2, \dots, x_n) 与 (y_1, y_2, \dots, y_n) 表示待求相似度的两个 n 维向量。若两个古彝文字图像特征之间的距离越小, 则这两个古彝文字图像的相似度越高。

5.2.2 实验结果及讨论

图 6 是在 ResNet152、VGG19 和 Histogram 等

3 种特征提取算法和曼哈顿距离、欧氏距离平方以及余弦距离等 3 种距离计算式下的 α 随 K 的变化曲线。表 1 为在 K 分别取 1、10、50、100、200、500 和 700 的情况下的 α 值, 其中最优结果被加粗显示。

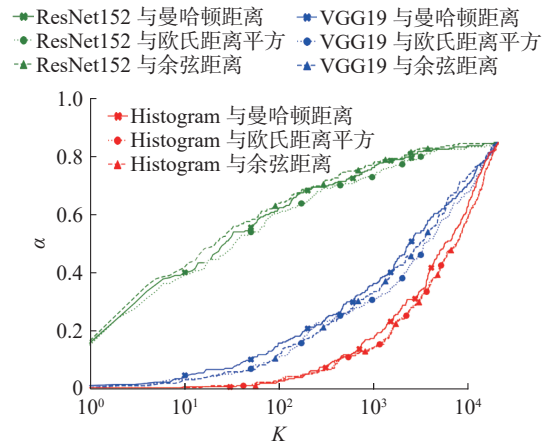


图 6 α 随 K 的变化曲线

Fig. 6 Curves of α with K

表 1 随 K 变化的 α 值
Table 1 Value of α as a function of K

特征提取算法 与相似度计算方法	$K=1$	$K=10$	$K=50$	$K=100$	$K=200$	$K=500$	$K=700$
ResNet152与曼哈顿距离	15.7	39.6	54.5	62.0	67.8	69.0	71.0
ResNet152与欧氏距离平方	14.9	38.8	53.3	60.4	64.3	67.1	70.2
ResNet152与余弦距离	16.5	41.6	57.3	63.5	68.6	69.4	74.1
VGG19与曼哈顿距离	1.2	3.9	9.8	15.7	20.8	22.0	29.0
VGG19与欧氏距离平方	0.4	3.1	7.1	11.0	18.8	22.7	25.5
VGG19与余弦距离	0.4	3.1	6.3	11.8	17.3	19.2	25.9
Histogram与曼哈顿距离	0.4	0.8	1.2	2.7	5.1	6.3	11.8
Histogram与欧氏距离平方	0.4	0.4	1.2	3.1	5.1	6.3	11.0
Histogram与余弦距离	0.4	0.4	1.2	2.7	5.1	5.5	11.0

实验结果表明, 深度学习算法明显优于传统的 Histogram 算法, 且在深度学习算法中, ResNet152 较 VGG19 表现更为出色。这是因为古彝文字图像存在手写字体多样且大小不一、纸张纹理多样、纸张破损、污渍褶皱等复杂因素, 基于像素值统计的 Histogram 算法无法较好地提取古彝文字图像特征。相较于 VGG19, ResNet152 使用了残差网络, 允许信息在网络层之间直接跳跃, 这使得非常深层的网络能够避免梯度消失问题, 解决了神经网络结构加深出现的退化现象, 并更有效地学习古彝文字图像中复杂、细微的特征。此外, ResNet152 更强的表征能力可以帮助模型在处理不同风格的古彝文字图像时, 表现出更强的

泛化能力。

从实验结果中可发现, 在 K 为 200 时, 采用 ResNet152 和余弦距离, 近 70% 的待整理的彝文字图像能够被自动判断为已被收录, 只有约 30% 的需要人工判断是否被收录, 这将有效辅助人工整理古彝文字图像。

6 结束语

古彝文字是中华文化载体的重要组成部分, 搜集整理古彝文字图像对研读彝文古籍、挖掘中华文化共同性、传承中华优秀传统文化具有重要意义。本文针对人工搜集、整理数量繁多的古彝文字存在耗时耗力的问题, 提出了一种基于深度

学习技术的古彝文字图像搜集与整理方法:利用目标检测技术实现古彝文字图像搜集,通过图像特征提取和特征相似性匹配实现古彝文字图像整理。实验表明该方法能够有效辅助人工搜集与整理古彝文字图像。值得注意的是,本文并不是改进图像特征提取方法和图像相似度计算方法,而是把一些经典方法组合起来创新性地应用在古彝文字领域,为古彝文字图像搜集和整理难题提供一套简单可行的技术路线。

虽然本文在古彝文字图像自动搜集与整理上做了一些工作,然而由于古彝文异体字众多、四省区的古彝文字的差异性、彝文古籍的丰富性等,对自动搜集与整理古彝文字图像的方法研究仍然是开放的,例如:进一步对古彝文字图像进行聚类分析,使古彝文字图像的自动整理更加精细;对云南、四川、贵州、广西古彝文字异同进行辅助分析;在更大范围的,尤其是在复杂背景下的彝文古籍上,进行验证;与自然语言处理技术相结合,把前后字符关系、语言知识引入古彝文字图像搜集与整理;扩展应用到其他文种的古文字搜集与整理。

参考文献:

- [1] 孔祥卿. 彝文的源流[M]. 北京: 民族出版社, 2005.
- [2] 韩旭. 彝文古籍字符检测和识别的研究与实现[D]. 重庆: 西南大学, 2020.
HAN Xu. Research and implementation of character detection and recognition in Yi ancient books[D]. Chongqing: Southwest University, 2020.
- [3] 贾晓栋. 基于深度学习的手写彝文识别技术应用研究[D]. 北京: 中央民族大学, 2017.
JIA Xiaodong. Research on the application of handwritten Yi recognition technology based on deep learning[D]. Beijing: Central University for Nationalities, 2017.
- [4] 胡峰, 李路正, 代劲, 等. 结合聚类边界采样的主动学习[J]. 智能系统学报, 2024, 19(2): 482–492.
HU Feng, LI Luzheng, DAI Jin, et al. Active learning combined with clustering boundary sampling[J]. *CAAI transactions on intelligent systems*, 2024, 19(2): 482–492.
- [5] 王定旺. 彝文联机手写体识别的研究与应用[D]. 重庆: 西南大学, 2021.
WANG Dingwang. Research and application of online handwriting recognition in Yi language[D]. Chongqing: Southwest University, 2021.
- [6] WANG Chongguang, EVANS K, HARTLEY D, et al. A systematic review of artificial neural network techniques for analysis of foot plantar pressure[J]. *Biocybernetics and biomedical engineering*, 2024, 44(1): 197–208.
- [7] 陈善雄, 韩旭, 林小渝, 等. 基于 MSER 和 CNN 的彝文古籍文献的字符检测方法[J]. 华南理工大学学报(自然科学版), 2020, 48(6): 123–133.
CHEN Shanxiong, HAN Xu, LIN Xiaoyu, et al. MSER and CNN-based method for character detection in ancient Yi books[J]. *Journal of South China University of Technology (natural science edition)*, 2020, 48(6): 123–133.
- [8] CHEN Huizhong, TSAI S S, SCHROTH G, et al. Robust text detection in natural images with edge-enhanced Maximally Stable Extremal Regions[C]//2011 18th IEEE International Conference on Image Processing. Brussels: IEEE, 2011: 2609–2612.
- [9] 陈善雄, 王小龙, 韩旭, 等. 一种基于深度学习的古彝文识别方法[J]. 浙江大学学报(理学版), 2019, 46(3): 261–269.
CHEN Shanxiong, WANG Xiaolong, HAN Xu, et al. A recognition method of ancient Yi character based on deep learning[J]. *Journal of Zhejiang University (science edition)*, 2019, 46(3): 261–269.
- [10] CICHOCKI A, CRUCES S, AMARI S I. Generalized alpha-beta divergences and their application to robust non-negative matrix factorization[J]. *Entropy*, 2011, 13(1): 134–170.
- [11] 贵州省彝学研究会. 西南彝志[M]. 贵阳: 贵州民族出版社, 2015.
- [12] 《彝文典籍集成》编委会. 彝文典籍集成·四川卷·教育[M]. 成都: 四川民族出版社, 2014.
- [13] 滇川黔桂彝文协作组编. 滇川黔桂彝文字集[M]. 昆明: 云南民族出版社, 2004.
- [14] JIANG Peiyuan, ERGU Daji, LIU Fangyao, et al. A review of yolo algorithm developments[J]. *Procedia computer science*, 2022, 199: 1066–1073.
- [15] 沙马拉毅. 《规范彝文方案》推行 30 年实践效果述评[J]. 西南民族大学学报(人文社科版), 2010, 31(8): 28–31.
SHA M. A review of the practice effect of standardizing Yi language program for 30 years[J]. *Journal of Southwest University for Nationalities (humanities and social science edition)*, 2010, 31(8): 28–31.
- [16] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(6): 1137–1149.
- [17] FARHADI A, REDMON J. Yolov3: An incremental improvement[C]//Computer vision and pattern recognition. Berlin: Springer, 2018, 1804: 1–6.
- [18] JANI M, FAYYAD J, AL-YOUNES Y, et al. Model compression methods for YOLOv5: a review[EB/OL].

- (2023-07-21)[2024-06-21]. <https://arxiv.org/abs/2307.11904v1>.
- [19] WANG Ao, CHEN Hui, LIU Lihao, et al. YOLOv10: real-time end-to-end object detection[EB/OL]. (2024-05-23)[2024-06-21]. <https://arxiv.org/abs/2405.14458v2>.
- [20] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: fully convolutional one-stage object detection[EB/OL]. (2019-04-02)[2024-06-21]. <https://arxiv.org/abs/1904.01355v5>.
- [21] ZHANG T Y, SUEN C Y. A fast parallel algorithm for thinning digital patterns[J]. *Communications of the acm*, 1984, 27(3): 236-239.
- [22] YADAV S, SAWALE M D. A review on image classification using deep learning[J]. *World journal of advanced research and reviews*, 2023, 17(1): 480-482.
- [23] ZHAO Xia, WANG Limin, ZHANG Yufei, et al. A review of convolutional neural networks in computer vision[J]. *Artificial intelligence review*, 2024, 57(4): 99.
- [24] SRIVASTAVA D, WADHVANI R, GYANCHANDANI M, et al. A review: color feature extraction methods for content based image retrieval[J]. *International journal of computational engineering & management*, 2015, 18(3): 9-13.
- [25] DENG Jia, DONG Wei, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE, 2009: 248-255.
- [26] 曾婷, 唐孝, 谭阳, 等. 相似度三支决策模糊粗糙集模型的决策代价研究[J]. *智能系统学报*, 2020, 15(6): 1068-1078.
- ZENG Ting, TANG Xiao, TAN Yang, et al. Decision costs of the similarity three-way decision-theoretic fuzzy rough set model[J]. *CAAI transactions on intelligent systems*, 2020, 15(6): 1068-1078.
- [27] 卞则康, 王士同. 基于混合距离学习的鲁棒的模糊 C 均值聚类算法[J]. *智能系统学报*, 2017, 12(4): 450-458.
- BIAN Zekang, WANG Shitong. Robust FCM clustering algorithm based on hybrid-distance learning[J]. *CAAI transactions on intelligent systems*, 2017, 12(4): 450-458.
- [28] LEI Shiye, TAO Dacheng. A comprehensive survey of dataset distillation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2024, 46(1): 17-32.

作者简介:



熊铁妞, 硕士研究生, 主要研究方向为深度学习、图像处理、古彝文字数字化。E-mail: xiongtieniu@stu.swun.edu.cn。



邱吉芳, 本科生, 主要学习方向彝语语言学、彝语方言学。E-mail: 18384496920@163.com。



胡建, 教授, 博士, 主要研究方向为计算机视觉、群体智能、文献数字化。E-mail: hujian@swun.edu.cn。