

# 一种自适应图像增强和 AlexNet 的水书古文字识别算法

杨秀璋<sup>1</sup> 武 帅<sup>1,2</sup> 宋籍文<sup>3</sup> 廖文婧<sup>1</sup> 周既松<sup>1</sup>

YANG Xiuzhang WU Shuai SONG Jiwen LIAO Wenjing ZHOU Jisong

## 摘 要

针对水族古文字受字形变化、年代噪声影响,无法实现高质量数字化提取,且缺乏利用深度学习技术实现对水书古文字的自动化识别问题,提出了一种自适应图像增强和 AlexNet 的水书文字识别算法,通过优化模型识别复杂环境下文字的关键特征,实现水书古文字智能识别以及同汉字的自动翻译。首先,采集具有代表性的水书古文字并进行数据标注处理,通过图像变换算法扩充数据集;其次,计算各种场景的图像噪声阈值,构建自适应图像增强的去噪算法;最后,构建 AlexNet 神经网络模型实现水书古文字的自动识别。实验结果表明,提出的算法能有效识别复杂环境下的水族古文字,其精确率、召回率和  $F_1$  值分别为 0.975 5、0.974 3 和 0.974 3,能为少数民族文字识别、古籍文字自动提取及少数民族古文字与汉字的自动翻译提供有效支撑,具有一定的学术价值和应用前景。

## 关键词

水书古文字;自适应图像增强;AlexNet

doi: 10.3969/j.issn.1672-9528.2023.01.050

## 0 引言

自 2000 年国家开启“中国档案文献遗产工程”<sup>[1]</sup>计划,全国各省、市、区(县)均开展民族档案文献数字化抢救。2003 年至 2021 年,国家先后设立四批《中国档案文献遗产名录》,列举濒危拯救项目。我国十分重视对古籍和电子档案的整理,并亟需实现信息和数字化的文物保护<sup>[2]</sup>。相较于汉族地方志拯救,部分拥有自己语言的少数民族档案文献遗产在拯救过程中,由于数据匮乏、重视度不高、数字化程度低的问题,电子档案进程缓慢<sup>[3]</sup>。以重点民族古籍名录第一位“中华水书”为例,多数古文字以古籍、雕刻等形式保存于民间、博物馆和档案馆,由于年代久远,采集图像存在大量噪声,无法实现高质量数字化处理<sup>[4]</sup>。

基于此,本文针对水族古文字受字形变化、年代噪声影响,无法实现高质量数字化提取,且缺乏利用深度学习技术实现对水书古文字的自动识别等问题,提出一种自适应图像增强和 AlexNet 的水书文字识别算法,通过优化模型识别复

杂环境下文字的关键特征,实现水书古文字智能识别及汉字自动翻译。

## 1 相关研究

水书是水族古文字及其著编典籍的汉译统称,作为世界上“活着”的象形文字,收集整理、研究解读水书对了解水族历史文化、研究少数民族原始宗教和社会信仰、从中窥探中国古文字奥秘等有着重要意义。由于水族古籍和民间各类文献资料没有得到有效梳理和信息共享,一定程度上制约了水族文化研究的深度和广度,也导致国内外学术界对于中国水学研究认识不够清晰。同时,记录水书的载体如刺绣、碑刻、木刻、纸张等,随着时代的发展,大量水书古籍文献资料正在损毁和流失,成为永远不可再生的宝贵资源<sup>[5-8]</sup>。因此,如何利用计算机视觉技术提升水族文字手写体识别效果,是现阶段濒危水族拯救迫切需要解决的难题,利用人工智能和计算机视觉技术构建水书数据库实现对其文本信息数字化处理迫在眉睫。

杨秀璋等人针对水族古文字清晰度不足的问题,提出一种改进的图像增强级识别方法,有效降低图像噪声,准确提取水族文字轮廓。后针对水族文字数字化提取难题,提出一种基于自适应图像增强及区域检测的水族文字提取与分割算法,进一步提升水族文字数字化提取效果。刘杨翎<sup>[9]</sup>详细比较水字和古彝文字,结合量化分析、文献考究、借字及同义比较分析水字和汉字的关系。龙从军等人<sup>[10]</sup>详细分析中国少

1. 贵州财经大学信息学院 贵州贵阳 550025

2. 南京农业大学信息管理学院 江苏南京 211800

3. 贵州高速公路集团有限公司 贵州贵阳 550027

[基金项目] 贵州省科技计划项目(No.黔科合基础[2019]1041, No.黔科合基础[2020]1Y279); 贵州省教育厅青年科技人才成长项目(No.黔教合 KY 字[2021]135); 贵州财经大学 2021 年度校级项目(No.2021KYQN03)

数民族语言文字信息处理的进展。随着深度学习技术的不断成熟,为水族文字提取带来新的解决方法。丁琮<sup>[11]</sup>率先尝试CNN实现对水书字符识别,为实现水书字符识别系统打下基础。汤敏丽等<sup>[12]</sup>尝试运用Faster-RCNN算法实现对古籍水族文字识别,为数字化识别奠定基础。文献[13]考虑到古文存在字形变化的特征,提出一种改进卷积神经网络的古文字图像识别方法,为手抄本数字化识别奠定基础。

然而,上述方法仅提取水族文字或开展简要的分析,未利用深度学习技术有效识别水族文字类别,缺乏对古籍、碑刻、木刻中的水族文字识别,更忽视了水书古文字和汉字自动翻译的价值。

## 2 模型设计

该部分主要介绍模型构建和实现过程。

### 2.1 总体框架

本文提出一种基于自适应图像增强和 AlexNet 模型的水书古文字识别算法,整个框架如图 1 所示。针对团队先前的工作仅提取水族文字并完成图像分割,本文将在此基础上进一步构建分类模型实现水族古文字释义的识别以及同汉字的翻译。

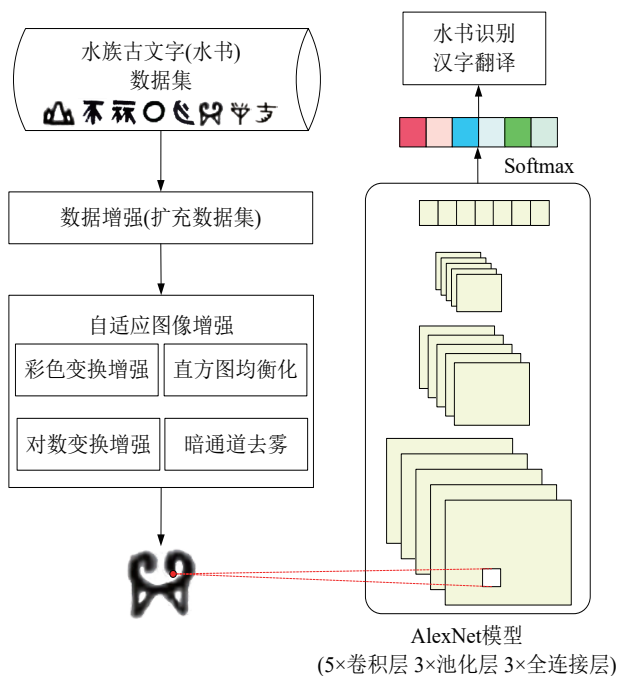


图 1 水书古文字识别整体框架

总体框架包括三个核心步骤。首先,提取并构建 12 个典型水族古文字的图像库,利用多种数据增强技术扩充数据集。其次,提出一种自适应图像增强方法,消除复杂环境下水书的噪声。最后,构建 AlexNet 深度学习模型,并完成水书识别和汉字翻译任务。

### 2.2 图像变换扩充水书古文字数据集

由于深度学习模型需要学习丰富的图像样本,而当今世界仍缺乏一个系统且可用的水书数据集,仅靠人工标注和提取耗时耗力。因此,本文通过先前的技术提取初始样本集,包含 12 个代表性的水族古文字。接着,利用图像变换和数据增强方法扩充水族古文字数据集,如图 2 所示。



图 2 数据增强扩充水书数据集

由图 2 可知,图 2(a)通过几何变换(含镜像、旋转、翻转、缩放、平移)增强数据,图 2(b)通过颜色通道及色彩变换增强数据,图 2(c)利用增加高斯噪声和椒盐噪声来模仿古籍文字的复杂场景,图 2(d)通过亮度变换及怀旧、碑刻模拟来仿造水书真实场景的图像。

### 2.3 自适应图像增强去噪

中华古文字及少数民族文字常于龟甲、兽骨、古籍、石碑、木雕中发现,如甲骨文、水书、藏文。然而,由于年代久远或保存困难,这些古文字或文物中的文字存在大量噪声,传统计算机识别技术不理想。基于此,本文提出一种自适应图像增强去噪算法,通过中值滤波去除噪声,利用对数变换增强图像背景色彩,并构建暗通道先验(dark channel prior, DCP)<sup>[14]</sup>去雾算法将陈旧及复杂环境的水族文字背景增强。

中值滤波通过构建像素点的窗口依次扫描图像,每个像素点所对应的灰度值按照升序或降序排列,然后获取最中间的值来替换该点的灰度值。暗通道先验去雾算法是由何恺明老师提出,旨在利用大气散射模型实现图像去雾。其计算过程如公式(1)~(4)所示。

$$I(x) = J(x)t(x) + A(1-t(x)) \quad (1)$$

$$\frac{I^c(x)}{A^c} = \frac{J^c(x)}{A^c} t(x) + (1-t(x)) \quad (2)$$

$$J^{\text{dark}}(x) = \min_{y \in \Omega(x)} \left( \min_c J^c(y) \right) = 0 \quad (3)$$

$$J(x) = \frac{I(x) - A}{\max(t(x), t_0)} + A \quad (4)$$

大气散射模型按照公式(1)定义,  $I(x)$  表示有雾图像,  $J(x)$  表示要恢复的无雾图像,  $A$  为全球大气光照值,  $t(x)$  表示透射率。结合暗通道去雾算法的先验理论, 无雾图像的非天空区域  $J(x)$  的暗通道趋于 0, 大气散射模型计算如(2)所示, 执行两次最小值计算并将其转换为公式(3)。最后, 设置透射率下限值  $t_0$ , 再将计算的大气光值和透射率, 最终得到暗通道去雾公式(4),  $J(x)$  表示去雾图像。

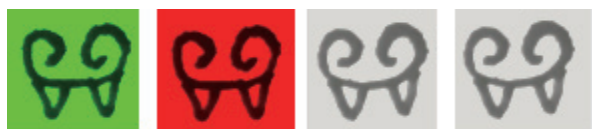
实验结果如图 3 所示, 利用本文提出的算法能有效消除各种类别的噪声并增强图像效果。



(a) 带噪声的原图



(b) 中值滤波去噪



(c) 复杂环境和怀旧场景原图



(d) 对数变换和暗通道去雾增强

图 3 自适应图像增强对比效果

## 2.4 AlexNet 模型识别水书古文字

在计算机视觉和图像分类领域, AlexNet 是经典的神经网络模型, 于 2012 年被 Alex Krizhevsky 等提出<sup>[15]</sup>。该模型的结果如图 4 所示, 由 5 个卷积层和 3 个全连接层组成。卷积层用于提取特征, 池化层用于实现特征降维, 全连接层起到分类作用。本文利用该模型来学习经自适应图像增强的水族古文字数据集, 并实现 12 个水族古文字的分类及汉字翻译。

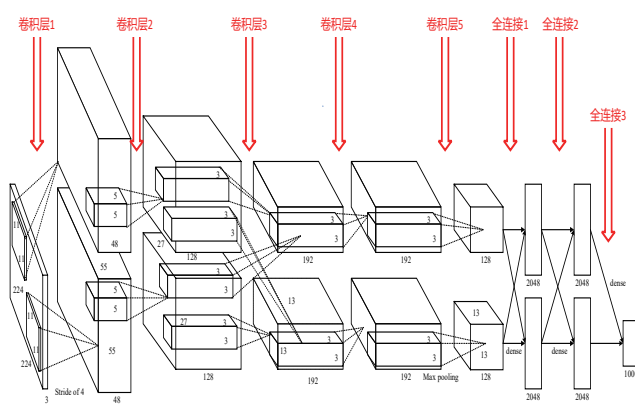


图 4 AlexNet 模型结构

## 3 实验评估

本文利用 Keras、TensorFlow 和 Sklearn 库构建多种模型, 并利用 GTX 1080Ti 加速训练。所有实验均在相同的环境下实现。

### 3.1 数据集

本文构造的水书古文字数据集包括 5280 幅古文字图像, 其中训练集、测试集和验证集分别为 3000、1800 和 480 幅, 完整数据集如表 1 所示。

表 1 水族古文字数据集构造及分布表

水书	汉字	类别	训练	测试	验证
	山	0	250	150	40
	木	1	250	150	40
	林	2	250	150	40
	日	3	250	150	40
	月	4	250	150	40
	弼星	5	250	150	40
	雨	6	250	150	40
	虎	7	250	150	40
	鱼	8	250	150	40
	神	9	250	150	40
	时	10	250	150	40
	卯	11	250	150	40



由表 1 可知,共包含 12 个水族古文字,它们极具代表性,分别表示了山川树木、日月星辰、动物神仙和占卜时辰,象征着水族文化和水学历史。其中,第一列为水书古文字,第二列为对应的汉字释义,后续分别为类别和对应数量。本文利用经典的精确率、召回率和  $F_1$  值指标来评估实验。

3.2 水书古文字识别

本文与经典的机器学习、深度学习进行详细对比实验,并比较了是否利用自适应图像增强算法的前后效果。实验结果如表 2 所示,本文模型的精确率、召回率和  $F_1$  值分别为 0.975 5、0.974 3 和 0.974 3。由表 2 可知,本文所提出的自适应图像增强和 AlexNet 模型能更好地识别水书古文字,并优于其它机器学习和深度学习模型,比其它未增强的三种机器学习模型平均  $F_1$  值提升 48.07%,比其它未增强的四种深度学习模型平均  $F_1$  值提升 4.07%。

表 2 水书古文字识别实验结果

分类	识别模型	是否增强	精确率	召回率	$F_1$ 值	$F_1$ 值提升
机器学习	SVM	否	0.546 7	0.525 3	0.535 8	—
	SVM	是	0.616 1	0.597 2	0.606 5	0.070 7
	KNN	否	0.526 5	0.410 4	0.461 2	—
	KNN	是	0.699 3	0.685 6	0.692 4	0.231 2
	RF	否	0.487 3	0.480 4	0.483 8	—
	RF	是	0.646 5	0.640 8	0.640 7	0.156 9
深度学习	1layer CNN	否	0.926 5	0.920 8	0.921 8	—
	1layer CNN	是	0.942 3	0.939 8	0.940 0	0.018 2
	2layer CNN	否	0.930 2	0.927 6	0.927 6	—
	2layer CNN	是	0.958 1	0.952 6	0.953 3	0.025 7
	3layer CNN	否	0.936 7	0.934 4	0.934 6	—
	3layer CNN	是	0.962 0	0.956 0	0.956 5	0.021 9
	AlexNet	否	0.960 3	0.947 2	0.950 6	—
	本文方法	是	0.975 5	0.974 3	0.974 3	0.023 7

同时,本文详细对比是否经过自适应图像增强的效果,不同模型均有一定程度提升,并且机器学习模型提升更为明显,经过图像增强的 KNN 模型提升 23.12%,本文模型相比于 AlexNet 模型提升 2.37%。

3.3 水书古文字和汉字自动翻译

最后,本文实现了水书古文字和汉字的自动翻译实验,利用 AlexNet 模型预测的分类类别来实现汉字翻译,最终结果如图 5 所示。

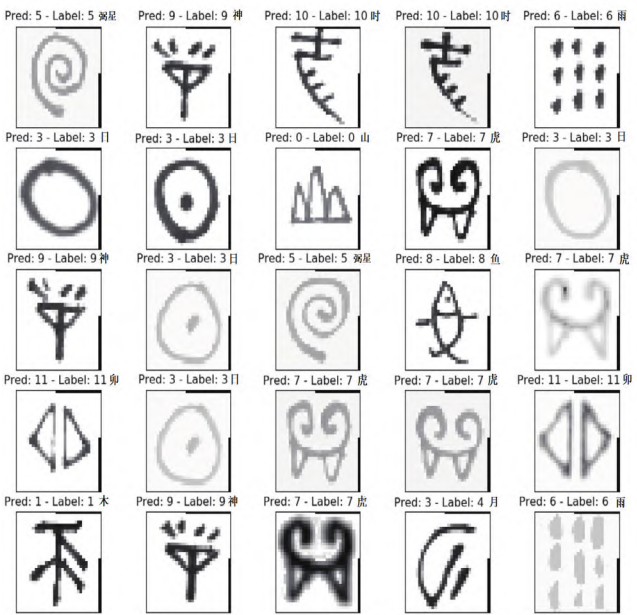


图 5 水书古文字和汉字自动翻译实验

由图 5 可知,随机预测了 25 个水族古文字,输出结果包括预测图像、预测类别、实际类别和自动翻译对应的汉字,比如第一个水族古文字预测的类别为 5,其真实类别也是 5,因此预测正确且将其翻译为“弼星”。类似地,本文实现了“神”“时”“雨”“日”“山”“虎”“鱼”“卯”“木”“月”等汉字与水书古文字的自动翻译。至此,本文有效实现了水书古文字识别和汉字翻译的任务。

4 结语

针对水族古文字受字形变化、年代噪声影响,无法实现高质量数字化处理,且缺乏利用深度学习技术实现对水书古文字的自动化识别。本文提出一种自适应图像增强和 AlexNet 的水书文字识别算法,通过模型识别复杂环境下文字的关键特征,实现水书古文字智能识别以及同汉字的自动翻译。

实验结果表明,本文的方法能更好地识别水族古文字,由于传统的机器学习和深度学习方法,并且自适应图像增强算法能更好地优化图像,导致其识别结果均有一定程度提升。

总而言之,本文的模型和思路能有效促进水书古文字的数字化保护及水族文化的宣传,为少数民族文字识别、古籍文字自动提取及少数民族古文字与汉字的自动翻译提供有效支撑,能为民族学和计算机科学的学科交叉融合提供思路,并具有一定的学术价值和应用前景。未来,本团队将继续深入研究,利用迁移学习和深度学习更好地实现大规模水族古文字提取以及同汉字的自动翻译。

## 参考文献:

- [1] 周耀林, 吴化. 数字人文视野下少数民族档案文献遗产数字化保护研究[J]. 档案学研究, 2022(5):123-129.
- [2] 李明华. 关于建立国家电子档案战略备份中心的提案[J]. 中国档案, 2022(3):20.
- [3] 高慧宜. 水族水文和傣族竹书的异体字比较研究[J]. 民族论坛, 2008(3):44-45.
- [4] 刘杨翎. 水字与古彝文比较研究[D]. 上海: 华东师范大学, 2018.
- [5] 杨秀璋, 夏换, 于小民. 一种基于水族濒危文字的图像增强及识别方法[J]. 计算机科学, 2019, 46(S2):324-328.
- [6] 杨秀璋, 武帅, 夏换, 等. 基于自适应图像增强技术的水族文字提取与识别研究[J]. 计算机科学, 2021, 48(S1):74-79.
- [7] 陈思. 水书揭秘[M]. 北京: 光明日报出版社, 2010.
- [8] 杨秀璋. 基于 LDA 模型和文本聚类的水族文献主题挖掘研究[J]. 现代计算机, 2019(5):13-17.
- [9] 刘杨翎. 水字与古彝文比较研究[D]. 上海: 华东师范大学博士学位论文, 2018.
- [10] 龙从军, 安波. 中国少数民族语言文字信息处理的进展[J]. 暨南学报(哲学社会科学版), 2022, 44(9):12-23.
- [11] 丁琮. Matlab 平台下水书文字特征提取与分类方法实现研究[J]. 电子技术与软件工程, 2020(14):155-157.
- [12] 汤敏丽, 谢少敏, 刘向荣. 基于 Faster-RCNN 的水书古籍手写文字的检测与识别[J]. 厦门大学学报(自然科学版), 2022, 61(2):272-277.
- [13] 杨秀璋, 施奕, 李娜, 等. 一种改进卷积神经网络的阿拉伯文字图像识别方法[J]. 信息技术与信息化, 2021(9):6-11.
- [14] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior[J]. IEEE transactions on pattern analysis and machine intelligence, 2011, 33(12): 2341-2353.
- [15] SHRIVASTAVA V K, PRADHAN M K, MINZ S, et al. Rice plant disease classification using transfer learning of deep convolution neural network[J]. International archives of the photogrammetry, remote sensing & spatial information sciences, 2019, 3(6): 631-635.

## 【作者简介】

杨秀璋(1991—), 男, 苗族, 贵州凯里人, 硕士, 助教, 研究方向: 数据挖掘、知识图谱、数据分析。

武帅(1994—), 通讯作者, 男, 江苏淮安人, 硕士, 工程师, 研究方向: 数智赋能、自然语言处理。

(收稿日期: 2022-10-12 修回日期: 2022-11-20)

(上接第 211 页)

- [6] SCHNEIDER F, THONEMANN U W, KLABJAN D. Optimization of battery charging and purchasing at electric vehicle battery swap stations[J]. Transportation science, 2017, 52(5): 1211-1234.
- [7] 马婧, 冯锋. 基于 WSN 的仓储环境监测与预警机制研究[J]. 信息与电脑(理论版), 2021, 33(13): 21-23.
- [8] 付晓. 智能仓储系统中机器人优化配置及动态路径规划研究[D]. 西安: 西安电子科技大学, 2021.
- [9] TAO F, CHENG Y, ZHANG L, et al. Advanced manufacturing systems: socialization characteristics and trends[J]. Journal of intelligent manufacturing, 2017, 28(5): 1079-1094.
- [10] 王海涛, 宋丽华, 向婷婷, 等. 人工智能发展的新方向: 人机物三元融合智能[J]. 计算机科学, 2020, 47(S2): 1-5+22.
- [11] 郭斌. 论智能物联与未来制造: 拥抱人机物融合群智计算时代[J]. 人民论坛·学术前沿, 2020(13): 32-42.
- [12] 韩共乐. 基于边缘计算的制造状态主动感知和协同决策研究[D]. 西安: 长安大学, 2020.
- [13] 翟书颖, 郭斌, 李茹, 等. 信息物理社会融合系统: 一种以数据为中心的框架[J]. 大数据, 2017, 3(6): 85-92.
- [14] 丁凯, 陈东嶸, 王岩, 等. 基于云一边协同的智能工厂工业物联网架构与自治生产管控技术[J]. 计算机集成制造系统, 2019, 25(12): 3127-3138.
- [15] 董津, 王坚, 王兆平. 面向制造领域人机物三元数据融合的本体自动化构建方法[J]. 控制与决策, 2022, 37(5):1251-1257.

## 【作者简介】

楚昌昊(1998—), 男, 河南漯河人, 硕士研究生, 研究方向: 智能制造系统, email: 2020125046@chd.edu.cn.

张富强(1984—), 通信作者(email: fqzhang@chd.edu.cn), 男, 山西运城人, 副教授, 硕士生导师, 研究方向为面向服务的制造过程。

(收稿日期: 2022-07-13 修回日期: 2022-08-18)