

基于向量数据库的甲骨文图像检索研究

董兆安^{1,2}, 王伯勇¹, 王晓凡¹, 黄万丽^{1,2}, 刘智斌^{1,2}, 李光顺^{1,2}, 韩胜伟³, 熊晶^{1,2,3}✉

(1. 曲阜师范大学计算机学院, 山东日照, 276827;

2. 日照曲阜师范大学技术转移中心, 山东日照, 276827;

3. 甲骨文信息处理教育部重点实验室, 河南安阳, 455000)

摘要: 目前甲骨文字形主要以图像形式存在, 因此甲骨文图像检索技术是甲骨文研究中的一项基础且重要的任务。然而, 甲骨文图像面临因数据量大、异体字多、模态多样、差别微妙等特点导致传统数据库难以高效处理的挑战。Milvus 作为高性能开源向量数据库, 凭借其卓越的矢量检索能力和对深度学习平台的良好兼容性, 为甲骨文图像检索提供了新的解决方案。基于 Milvus 和 ResNet-50 模型, 构建了一个甲骨文字形检索系统。用户可通过拍照、手写或上传图像的方式快速检索相似的甲骨文字符图像。实验表明, 该方法的检索平均查准率达到 82.86%。该系统具有良好的扩展性, 可应用于其他古文字图像检索, 为文化遗产数字化保护提供了强大的技术支持。

关键词: 甲骨文; 图像检索; 向量数据库; Milvus

Oracle Bone Inscriptions Image Retrieval Based on Vector Database

DONG Zhaoan^{1,2}, WANG Boyong¹, WANG Xiaofan¹, HUANG Wanli^{1,2}, LIU Zhibin^{1,2}, LI Guangshun^{1,2}, HAN Shengwei³, XIONG Jing^{1,2,3}✉

(1. School of Computer Science, Qufu Normal University, Rizhao, 276827, China;

2. Rizhao-Qufu Normal University Joint Technology Transfer Center, Rizhao, 276827, China;

3. Key Laboratory of Oracle Bone Inscriptions Information Processing, Ministry of Education, Anyang, 455000, China)

Abstract: Oracle Bone Inscriptions (OBI) primarily exist in image form, making oracle bone image retrieval technology a fundamental and crucial task in the study of ancient scripts. However, OBI images face challenges such as massive data volumes, numerous variant characters, and the difficulty of efficiently handling these with traditional databases. Milvus, as a high-performance open-source vector database, offers a new solution for oracle bone image retrieval through its outstanding vector search capabilities and seamless compatibility with deep learning platforms. This paper constructs an OBI glyph retrieval system based on Milvus and the ResNet-50 model. Users can quickly retrieve similar oracle bone character images by taking photos, handwriting, or uploading images. Experiments show that the average retrieval accuracy of this method can reach 82.86%. The system is highly extensible and can be applied to the image retrieval of other ancient scripts, providing strong technical support for the digital preservation of cultural heritage.

Key Words: Oracle Bone Inscriptions; Image retrieval; Vector database; Milvus

1 引言

甲骨文是中国古代文明的瑰宝, 是已知最早的成熟文字系统, 是汉字的起源。作为中华文化的重要组成部分, 甲骨文不仅具有极高的历史和考古价值, 还承载了丰富的文化内涵^[1]。甲骨文记录了古代中国的社会、政治、宗教等各方面的信息, 为研究中华文明的起源与发展提供了宝贵的资料。保护和传承甲骨文及其背后的文化遗产, 对维护中华优秀传统文化、增强文化自信具有重大意义^[2]。然而, 由于甲骨文字形复杂多样, 传统的甲骨文检索和研究工作面临着诸多挑战。

甲骨文字形图像数据量大、异体字多、模态多样、差别微妙, 如图 1 所示。因此, 甲骨文大规模图像检索常常面临效率低下和准确率不足的问题。传统的甲骨文检索系统多依赖人工标注和关键字匹配, 难以应对甲骨文图像的复杂性和多样性^[3]。此外, 随着甲骨文数字化工作的推进, 数据规模呈指数级增长, 传

统数据库在存储和检索这些高维图像数据时，逐渐暴露出实时性和可扩展性方面的不足^[4]。因此，如何构建一个高效、准确且具备扩展性的甲骨文图像检索系统，成为亟待解决的关键问题。



图 1 甲骨文字形图像特征

近年来，随时深度学习和大数据的迅猛发展，图像检索领域获得了许多新的技术手段。卷积神经网络（Convolutional Neural Networks, CNN）模型，特别是 ResNet-50 等深度学习模型，已在图像特征提取和分类任务中得到了广泛应用，能够有效捕捉图像的细节特征^[5]。同时，向量数据库的出现，为处理大规模高维向量检索提供了强有力的支持。Milvus 作为一种开源的高性能向量数据库，因其支持高效的矢量检索和与深度学习平台的无缝集成而备受关注。Milvus 强大的向量检索能力能够快速、准确地处理海量图像数据，并在大规模数据环境下表现出卓越的可扩展性^[6]。

为解决传统甲骨文图像检索的不足，本文提出了一种基于深度学习和向量数据库的创新解决方案。采用 ResNet-50 模型对甲骨文图像进行特征提取，生成高维向量表示，并借助 Milvus 数据库进行高效的相似度检索。基于此，我们构建了一个高效的甲骨文图像检索系统，不仅提高了图像检索的速度和精度，也为研究人员提供了一种便捷的工具，能够快速定位所需的甲骨文字形图像，从而为甲骨文的研究、保护和传承提供了有力支持。

研究表明，结合深度学习与向量数据库的方法能够有效克服传统检索系统在处理大规模图像数据时的种种局限性。通过对比传统图像检索技术和 Milvus 向量数据库的检索性能，我们验证了后者在处理甲骨文图像数据时的优势。该系统不仅能够提高甲骨文研究的效率，还为考古学中其他类似的文字和图像检索任务提供了新的思路。

2 相关工作

2.1 数据驱动的考古学

在考古学领域，数据驱动的研究方法近年来得到了越来越广泛的应用，尤其是在数智时代背景下，数据技术正推动着考古学的理论探索与方法创新^[7]。传统的考古学研究主要依赖于实地发掘、历史文献分析和手工绘图等方式，这些方法虽然在一定程度上有效，但由于考古数据的复杂性和多样性，面临着数据处理效率低、信息提取不全等问题。

随着大数据、人工智能和深度学习等前沿技术的快速发展，考古学家能够以全新的视角处理和分析大量考古数据。通过高效的数据管理和智能分析，考古学研究从依赖经验和定性分析，逐渐转向数据驱动的定量分析。这些技术不仅能提高考古资料的处理速度，还能通过数据挖掘发现隐含的文化和历史信息^[8]。例如，机器学习模型被广泛应用于文物图像分类、特征识别和模式匹配等任务中。向量数据库的出现也为大规模考古数据的检索提供了强有力的技术支撑。

在甲骨文研究领域，数据驱动的技术可以通过处理大量复杂的文字图像，自动提取其关键特征，从而有效解决人工标注的局限性。这种新兴方法不仅提高了数据处理的效率和准确性，还为考古学家的理论探索和研究创新提供了更多可能性。

2.2 甲骨文图像检索

在甲骨文研究领域，图像检索技术的演变为研究者提供了全新的研究手段。早期的甲骨文图像检索依赖于人工分类和注释，研究人员通过手工比对甲骨文图像特征进行分析。这种方法不仅耗时费力，还容易受到主观因素的影响，导致检索效率低下^[9]。随着基于内容的图像检索（Content-Based Image Retrieval, CBIR）技术的兴起，研究者开始利用图像的形状、纹理等低层次特征进行匹配。然而，由于甲骨文的固有特点，仅依靠低层次特征往往难以捕捉其字形的差别。

近年来，深度学习技术的突破显著提升了甲骨文图像检索的能力。CNN 网络，特别是 ResNet-50 模型，能够自动提取甲骨文图像的高层次特征，大幅提高了检索的准确性和效率。“殷契文渊”甲骨文大数据平台的发布标志着甲骨文研究进入了数据驱动的新时代，显著改善了传统甲骨文研究方法的局限性。

2.3 ResNet-50 和 Milvus

ResNet-50 模型凭借其深层次的卷积网络架构，在图像特征提取中展现出显著优势。通过引入残差学习机制，ResNet-50 能够有效避免深层网络中的梯度消失问题，确保在处理复杂图像时仍能提取到高层次的语义特征。现有研究表明，ResNet-50 在手写文字识别、医疗图像分析等领域中已经表现出了卓越的特征提取能力，尤其在处理复杂形态的图像数据时能够抓取关键特征^[10]。

Milvus 作为一种新兴的高性能开源向量数据库，在大规模图像数据检索中表现卓越。其核心优势在于高效的向量化存储与检索能力，能够处理上亿级别的高维特征向量数据。Milvus 利用分层导航小世界（Hierarchical Navigable Small World, HNSW）等高效算法，能够在大规模数据环境下快速进行相似度计算，实现实时检索^[11]。Milvus 通过计算向量之间的距离来判定它们之间的相似度。若两个嵌入向量相似，则意味着原始数据源也相似^[12]。因此，Milvus 广泛用于内容检索、推荐系统等应用场景。

已有研究表明深度学习与向量检索相结合，能够为文化遗产数字化保护提供创新的技术手段，显著提升了工作效率^[13]。

3 甲骨文图像检索方法

基于上述分析，本文采用 ResNet-50 与 Milvus 结合的方法实现甲骨文的图像检索。甲骨文图像具有规模大、异体字多、字形复杂等特点，ResNet-50 与 Milvus 分别为适应这些特点提供了有力的解决手段。具体而言，ResNet-50 提取的高维特征向量能够精确表达甲骨文的字形和纹理信息，为相似图像的检索提供了坚实的基础。这种精准的特征提取能力显著提升了检索的效率和准确性，尤其是在处理复杂的甲骨文图像数据时，Milvus 的高效检索机制进一步缩短了检索时间。

基于 ResNet-50 与 Milvus 的甲骨文图像检索流程如图 2 所示。其中，用户通过浏览器向服务器发送检索请求，服务器处理请求后，将结果返回给浏览器并呈现给用户。向量数据库存储了甲骨文手写体 HWOBC (Handwriting Oracle Bone Character Recognition Database) 数据集，该数据集专门用于手写甲骨字的离线识别训练，包含了 83,245 张手写甲骨字样本图片，并按照字库编码将这些样本分为 3,881 个不同的类别^[14]。

系统存储时，首先利用 ResNet-50 模型将 HWOBC 数据集中的图片转换为特征向量。这些特征向量随后存储在 Milvus 向量数据库中，以支持高效的检索。系统检索时，用户通过 Web 界面输入待查询的图片时，系统会通过 ResNet-50 模型将该图片转换为特征向量，并在 Milvus 数据库中进行相似度匹配。最终，系统将最相似的甲骨文字形图像通过 Web 界面返回给用户。

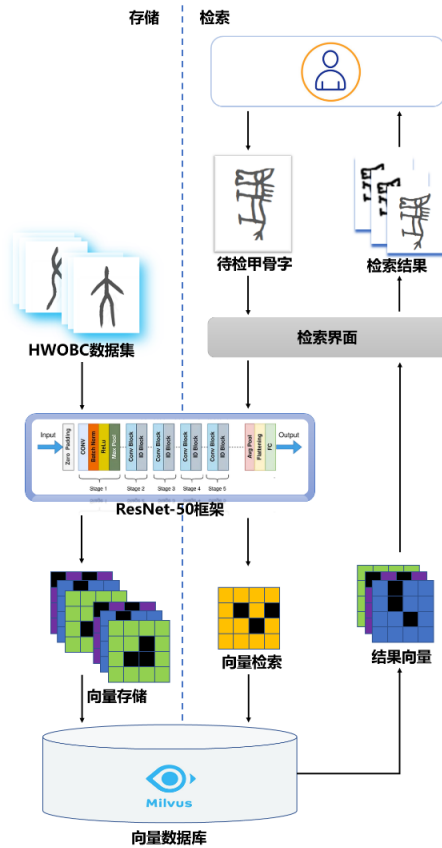


图 2 甲骨文图像检索流程

4 系统实现

针对甲骨文图像检索的实际需求，本文实现了一个基于 Milvus 的甲骨文图像检索系统。为了增强用户交互体验，该系统提供拍照上传、手写识别以及本地上传查询功能。系统主界面如图 3 所示。

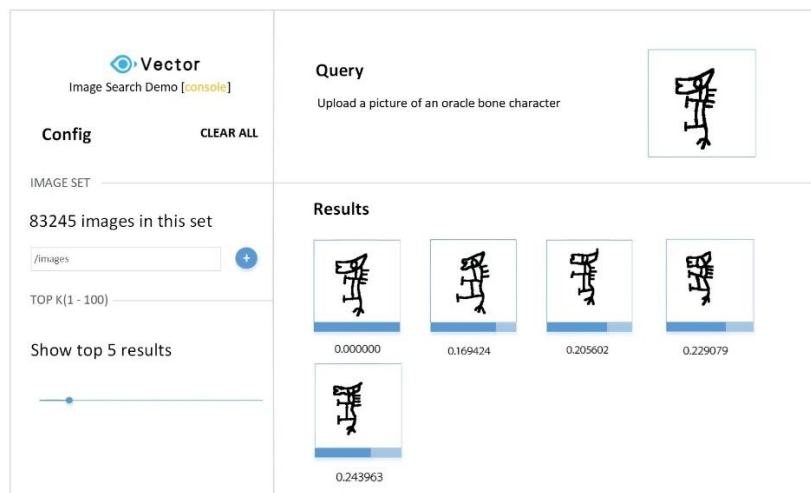


图 3 甲骨文图像检索系统主界面

图 3 中，以检索汉字“马”对应的甲骨文图像为例，展示了 Milvus 返回的与“马”最相似的前 5 个匹配结果。系统主界面分为三个主要板块：配置区、查询区和结果区。在左侧的配置区，用户可以上传甲骨文图片数据集，并通过滚动条设置显示的 TOP-K 结果数量。右上方的查询区允许用户上传一张甲骨文字形图像作为查询输入，系统将对查询图像进行处理，并在右下角的结果区即时显示排名前五的匹配结果及其相似度得分。

5 实验与分析

为验证所提方法的有效性，本文进行了对比实验。选择传统图像检索技术-局部敏感哈希（Locality-Sensitive Hashing, LSH）和 Milvus 向量数据库，比较两者在甲骨文图像检索任务中的性能。实验通过使用 ResNet50 模型提取图像特征，并在这两种数据库中存储和检索甲骨文字形图像特征，以检索时间和检索精度来评价两者的性能。

5.1 数据集

数据集来源于 HWOBC，这是一个面向手写甲骨字脱机识别训练的手写字符数据集[14,15]。数据的搜集方法是根据“殷契文渊”中的甲骨字字形库，由 22 位来自不同专业（文字、书法、考古、历史、计算机等）的甲骨学研究者，通过手写甲骨字搜集软件比照甲骨字字形库书写并整理而成。HWOBC 共搜集了 83245 张甲骨文样本图片，按照字库编码分为 3881 类。

5.2 实验设计

实验使用 HWOBC 数据集，并统一采用 ResNet-50 模型进行图像特征提取，使用 ResNet-50 的卷积层输出作为特征提取的结果。传统图像检索方法采用基于哈希编码的相似性搜索，而本实验则使用 Milvus 向量数据库进行图像检索。

实验环境：处理器为 Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz（8 核），8G 内存，2G 显存。开发语言为 python 3.8，集成开发环境为 PyCharm 2023.2.1，深度学习框架选用 TensorFlow 2.8 和 Keras。图像进行预处理与特征提取采用 ResNet-50；向量数据库采用 Milvus 2.2，并通过 pymilvus 库与其进行交互。

5.3 性能评估

本文采用检索时间、检索精确度来验证方法的性能。针对检索时间，分别取查询结果数量为 1 个、5 个、10 个时所耗时间（如图 3 显示了前 5 个检索结果），即 TOP-K 结果的检索时间。传统图像检索技术 LSH 和 Milvus 向量数据库检索技术的检索时间对比如图 4 所示。

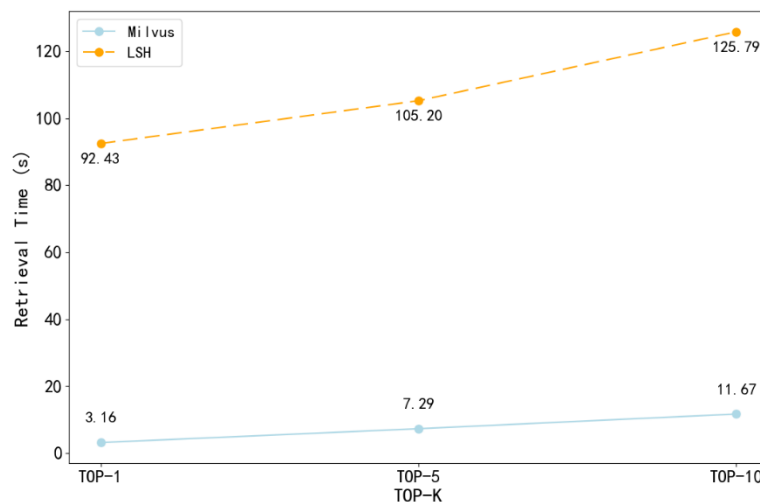


图 4 Milvus 与 LSH 在不同 TOP-K 下的检索时间对比

由如图 4 可知，LSH 在处理大规模甲骨文字形图像数据时，检索时间随着数据量的增加明显上升，尤其在数万级数据集上，传统图像特征检索技术的特征匹配效率低下。相比之下，Milvus 向量数据库通过高效的向量索引机制，即使在数据量较大的情况下，也能保持较快的检索速度和较少的响应时间。

针对检索精度，本文采用 Precision@10、Recall@10、AP@10 和 Precision-Recall@10 曲线等指标进行验证，以全面评估检索系统在不同条件下的表现，从而确保结果的准确性和有效性。

Precision@10 表示前 10 个返回结果中有多少字形是正确的，它表征的是查准率。公式如式（1）所示。

$$Precision@10 = \frac{\text{前 10 个返回结果中正确字形的数量}}{R} \quad (1)$$

其中，R 表示返回的结果数，该实验中取 R=10。

Recall@10 表示前 10 个返回的结果中正确字形数量占全库中该字形总数的比例，它表征的是查全率。

公式如式（2）所示。

$$Recall@10 = \frac{\text{前 10 个返回结果中正确字形的数量}}{S} \quad (2)$$

其中，S 表示该字形在全库中的总数，该实验中取 S=22。

AP@10 即 Average Precision@10，表示返回结果取不同 Top-k ($k \in [1,10]$)时精度累计的平均值，用于衡量检索结果的排序质量。公式如式（3）所示。

$$AP@10 = \frac{1}{N} \sum_{k=1}^{10} P@k \cdot rel(k) \quad (3)$$

其中 N 表示返回正确字形的数量，P@k 表示 Top-k 结果的检索精度， $rel(k)$ 是一个二值函数，如果第 k 个结果是正确的，则 $rel(k)=1$ ，否则 $rel(k)=0$ 。

为方便直观展示，选取 7 张甲骨文字形图片作为样例，利用 Milvus 检索得到的结果如表 1 所示。

表 1 甲骨文字形图片查询结果





















































































待检索 甲骨字	检索结果排序（表中数字为向量间的距离，标红者为错误检索结果）									
	1	2	3	4	5	6	7	8	9	10
	 0.000000	 0.139918	 0.150854	 0.184490	 0.187116	 0.189549	 0.202251	 0.212236	 0.213200	 0.225720
	 0.000000	 0.213609	 0.234897	 0.236730	 0.248233	 0.260127	 0.265513	 0.271772	 0.272594	 0.298729
	 0.000000	 0.250785	 0.270434	 0.280049	 0.295217	 0.311368	 0.319858	 0.323494	 0.342961	 0.345251
	 0.000000	 0.226292	 0.229278	 0.241223	 0.241920	 0.245087	 0.254695	 0.262623	 0.268814	 0.272802
	 0.000000	 0.169424	 0.205602	 0.229079	 0.243963	 0.248934	 0.251772	 0.251802	 0.252350	 0.254874
	 0.000000	 0.165301	 0.170600	 0.200060	 0.201394	 0.206687	 0.207691	 0.213495	 0.214552	 0.215527
	 0.000000	 0.206185	 0.215868	 0.243649	 0.268874	 0.282170	 0.291061	 0.297753	 0.307188	 0.316846

表 1 中，由上而下的甲骨字分别对应的现代汉字为：允、羊、牛、老、马、鼠、虎。每张甲骨文图片都展示了 Milvus 返回的前 10 个最相似的甲骨文字形图片。检索结果的 Precision@10，Recall@10 和

AP@10 及其平均值(AVG) 指标如表 2 所示。

表 2 Milvus 图像检索精确率

	Precision@10	Recall@10	AP@10
	90.00%	37.50%	100.00%
	100.00%	45.45%	100.00%
	80.00%	36.36%	100.00%
	80.00%	33.33%	84.55%
	70.00%	31.82%	96.17%
	70.00%	31.82%	80.43%
	90.00%	40.91%	86.54%
AVG	82.86%	36.74%	95.38%

Milvus 与 LSH 两者的 Precision-Recall@10 曲线比较结果如图 5 所示。Milvus 在 Precision-Recall@10 曲线上表现更为平衡和稳定，其精确率总体较高，即使在较高的召回率下也能保持相对高的精确度。相反，LSH 的精确率在召回率增加时下降较快，表现出较大的波动。这表明，Milvus 在甲骨文图像检索中更具优势，而 LSH 虽然在一些召回率水平下接近 Milvus，但在高召回率下的精确率较低。

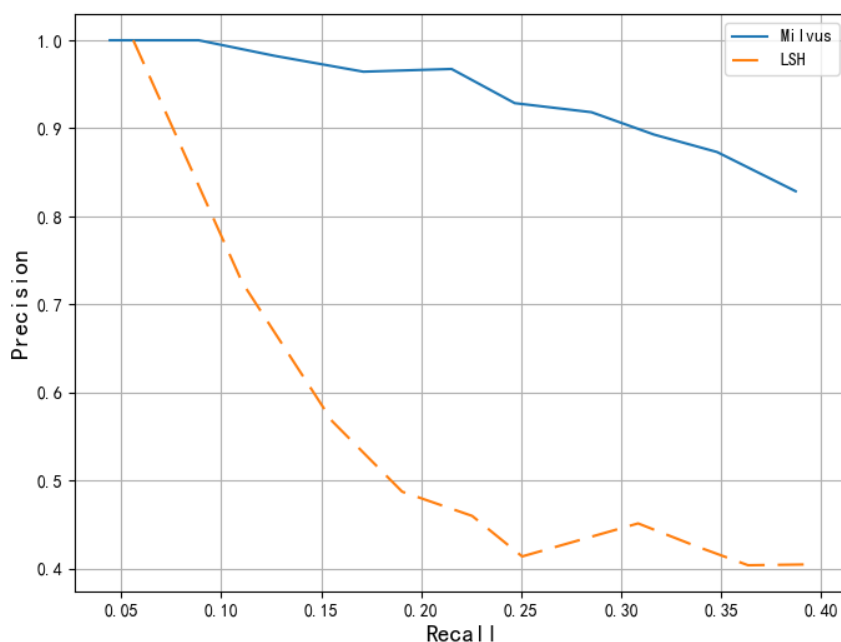


图 5 Precision-Recall@10 曲线对比图

Milvus 与 LSH 的 $AVG_{Precision@10}$ 、 $AVG_{Recall@10}$ 和 $AVG_{AP@10}$ 等指标对比结果如图 6 所示，Milvus 在图像检索中优势明显，其基于图像特征向量进行相似度计算，能够精准捕捉甲骨文字形向量之间的细微差异，从而保证了图像检索结果的高准确率。传统方法依赖于人工定义的标签、类别或其他辅助特征，容易忽略图像细节特征，导致在复杂的甲骨文字形数据集上的检索准确性较低。

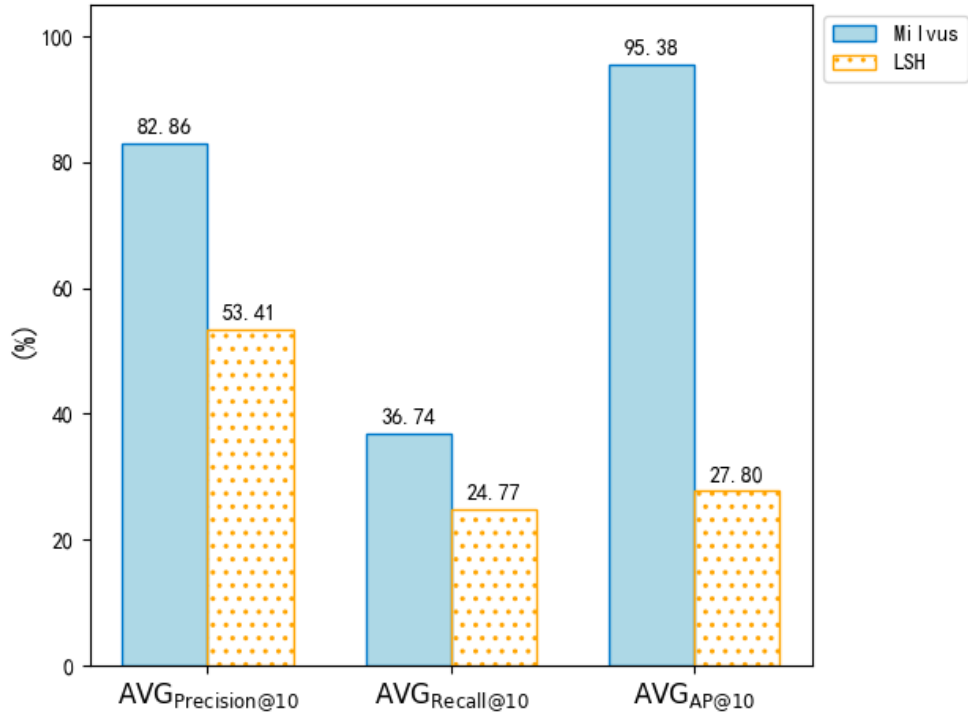


图 6 Milvus 与 LSH 的 $AVG_{Precision@10}$ 、 $AVG_{Recall@10}$ 和 $AVG_{AP@10}$ 等指标对比图

6 总结与展望

本文提出了一种基于深度学习和向量数据库的甲骨文图像检索系统，旨在弥补传统方法在处理甲骨文图像数据量大和异体字多情况下的不足。该方法结合了 ResNet-50 模型和 Milvus 向量数据库，构建了一个高效且准确的检索系统。ResNet-50 模型用于提取甲骨文图像的高维特征，利用其深层卷积网络和残差学习机制，克服了传统方法的局限。特征向量存储在 Milvus 数据库中，该数据库凭借其高效的向量存储和检索能力，显著提升了检索速度和准确性。实验结果表明，相比传统图像检索技术，基于 Milvus 的向量检索在 Precision、Recall、AP 和 Precision-Recall 等指标及检索效率方面均表现更佳。此外，系统设计了拍照上传、手写识别和本地上传查询功能，以方便甲骨文研究者和爱好者。该方法的应用范围可扩展至其他古文字和文物图像的检索，例如青铜器铭文和古代石刻，推动更广泛的文化遗产数字化保护与研究。

但是，目前向量数据库仅存储了 HWOBC 数据集中的甲骨文图像，还没有存储甲骨拓片、照片、摹本等多模态数据。在下一步的研究中，将扩大甲骨图像的规模和模态，以便更好地服务甲骨文研究。另外，将引入更深层次的卷积神经网络或 Transformer 架构，尝试提高图像特征提取的精度和效率。同时，优化 Milvus 数据库的索引结构和查询算法，以适应更大规模的数据集和实时检索需求。

基金支持：本研究受河南省高校重点科研项目计划基础研究专项(22ZX010)、山东省自然科学基金面上项目(ZR2020MF149, ZR2024MF146)等基金支持。

参考文献：

- [1] Xiong, J., Liu, G., Liu, Y., Liu, M.: Oracle bone inscriptions information processing based on multi-modal knowledge graph[J]. Computers & Electrical Engineering. 2021, 92: 107173.
- [2] N. Wang, Q. Sun, Q. Jiao and J. Ma, Oracle Bone Inscriptions Detection in Rubbings Based on Deep Learning[C]//2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 2020, 1671-1674.
- [3] Li Y, Ma J, Zhang Y. Image retrieval from remote sensing big data: A Survey[J]. Information Fusion, 2021, 67(3): 94-115.

- [4] Johnson, J., Douze, M., & Jégou, H.: Billion-scale similarity search with GPUs[J]. IEEE Transactions on Big Data, 2019, 7(3):535-547.
- [5] 高旭.基于卷积神经网络的甲骨文识别研究与应用[D].长春:吉林大学,2021.
- [6] Wang J, Yi X, Guo R, et al. Milvus: A Purpose-Built Vector Data Management System[C]// SIGMOD '21: Proceedings of the 2021 International Conference on Management of Data, 2021,Association for Computing Machinery, New York, NY, USA, 2614-2627.
- [7] GUAN Ying, ZHOU Zhenyu. Data science in Paleolithic Archaeology[J]. Acta Anthropologica Sinica, 2022, 41(01): 169-179.
- [8] 张兆基.唐蕃古道典型文物知识图谱构建及应用[J].西藏民族大学学报(哲学社会科学版),2024,45(1):80-87.
- [9] 王浩彬.基于深度学习的甲骨文检测与识别研究[D].广州:华南理工大学, 2019.
- [10] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [11] Wang J, Yi X, Guo R, et al. Milvus: A purpose-built vector data management system[C]//Proceedings of the 2021 International Conference on Management of Data. 2021: 2614-2627.
- [12] 许林. 基于 Bert 与 Milvus 的专利智能语义检索系统研究 [J]. 中国发明与专利, 2023, 20 (2): 5-11.
- [13] 沈贵华, 祝孔强.人工智能在文化遗产保护传承中的价值研究[J].博物院, 2023(3):43-49.
- [14] Li, B., Dai, Q., Gao, F., Zhu, W., & Liu, Y.: HWOBC-A handwriting oracle bone character recognition database. Journal of Physics Conference Series[J], 1651, 012050 (2020).
- [15] Yinqi Wen Yuan. <https://jgw.aynu.edu.cn/>, last accessed 2024/9/1.