基于向量数据库的甲骨文图像检索研究

**摘要：**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古文字研究中的一项基础且重要的任务。然而，甲骨文图像面临数据量庞大、异体字多、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。本文基于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，该方法的平均检索准确率可达到80%。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

# 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，也是汉字的起源。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]。它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。

首先，甲骨文字形图像形态各异，且数量庞大，这使得在进行大规模图像检索时，研究人员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，传统数据库在处理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。

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，高性能向量数据库的出现，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]。

为解决传统甲骨文检索系统的不足，本文提出了一种基于深度学习与向量数据库相结合的创新解决方案。我们采用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。通过这种方法，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。

本文的研究结果表明，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。通过对比传统图像技术和Milvus向量数据库的检索性能，我们验证了后者在处理高维图像数据时的优越性。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数据驱动的考古学研究

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，而向量数据库则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（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，由于甲骨文形态复杂且高度多样，仅依靠低层次特征往往难以捕捉其微妙差别。

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。与Milvus等高性能向量数据库结合，研究者能够在大规模甲骨文数据集中快速进行相似图像检索，从而实现更高效、准确的甲骨文研究。这一技术进展标志着甲骨文研究进入了数据驱动的新时代，显著改善了传统检索方法的局限性。

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HNSW（Hierarchical Navigable Small World）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。

将ResNet-50与Milvus结合使用在甲骨文图像检索中具有重要意义。ResNet-50能够提取甲骨文图像的高维特征，尤其是在捕捉文字细节和复杂形态方面表现出色。而Milvus则通过其高效的向量检索能力，能够在庞大的甲骨文数据库中快速匹配相似图像，极大提高了检索效率和准确性。现有的实验结果表明，该组合在处理甲骨文字形图像数据时，与传统基于特征点的图像检索方法相比，检索性能得到有效的提升。

这一特性对于甲骨文字形图像尤为关键，因为甲骨文的形态复杂多样且细节丰富。ResNet-50提取的高维特征向量能够精确表达甲骨文的形状和纹理信息，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。研究表明，结合深度学习与向量检索的技术方案，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研究方法

## 3.1系统设计背景

甲骨文研究的复杂性要求设计一个高效的图像检索系统，以处理其独特的形态特征和大规模数据集。本系统基于ResNet-50模型和Milvus向量数据库，致力于优化甲骨文字形图像的特征提取与检索过程。ResNet-50模型利用深度卷积网络架构，能够精准地提取甲骨文图像的高层次特征，克服了传统方法在处理复杂形态时的局限性。这些提取的特征向量被存储在Milvus向量数据库中，后者以其高效的向量化存储和实时检索能力，有效提升了大规模数据集的处理速度和准确性。

通过这种设计，系统不仅大幅提高了甲骨文图像检索的效率，还为考古学家提供了一个直观、便捷的工具，支持甲骨文及其他古文字的研究与保护。

## 3.2系统架构设计

甲骨文字形检索系统的处理流程如图1所示。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。系统使用了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]。

在系统中，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。当用户通过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1系统流程图

# 4 系统演示

为了增强用户交互体验，本系统实现了拍照上传、手写识别以及本地上传查询功能。本文将详细介绍本地上传甲骨文图片的过程。系统主界面如图2所示，展示了查询图像“马”的相似结果，并返回了前五个匹配结果。该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。系统将对该查询图像进行处理，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。



图2系统主界面

# 5实验与分析

为了验证系统的优越性和有效性，我们进行了系列性能评估和对比实验。具体而言，我们对比了传统图像检索技术（如局部敏感哈希LSH）与现代向量数据库（如Milvus）在甲骨文图像检索任务中的效率。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，我们评估了各自的检索速度和结果精确率。特别地，我们测试了检索结果的精确率，包括TOP-1、TOP-5和TOP-10的精确率。这些实验旨在揭示向量数据库在处理高维图像特征时的优势，并展示其在实际应用中的效能提升，以期为数字化考古学研究提供更高效的技术支持。

## 5.1数据集

HWOBC(Handwriting Oracle Bone Character Recognition Database)数据集，是一个面向手写甲骨字脱机识别训练的手写字符数据集。手写甲骨字的脱机离线识别是手写甲骨字数据法与手写甲骨学文献数字化必不可少的步骤之一。数据的搜集主要“殷契文渊”中的甲骨字字形库。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标准字形书写并整理而成。数据集共搜集了83245张样本图片，按照字库编码分为3881类[14, 15]。

## 5.2 实验设置

本实验使用了相同的甲骨文图像数据集，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哈希编码和标签匹配等技术，基于字段信息进行检索。而本实验则使用Milvus向量数据库进行图像检索。

## 5.3性能评估

本实验采用检索时间（系统响应速度）、检索精度（TOP-N准确率）和系统负载这三种指标来对本实验进行评估。

本文选取了7张甲骨文字形图片作为结果示例，如表1所示。每张图片的下方展示了与其最相似的前十个甲骨文字形图片。结果中所展示的字体分别为：允、羊、牛、老、马、鼠、虎。

表1 甲骨文字形图片查询结果示意图

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 待检索甲骨字 | 检索结果排序（表中数字为向量距离） | | | | | | | | | |
| 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 |
| 60858_0 | 60858_00.000000 | 0.250785 | 0.270434 | 0.280049 | 0.295217 | 0.311368 | 0.319858 | 0.323494 | 0.342961 | 0.345251 |
| 6003C_0 | 6003C_0  0.000000 | 0.226292 | 0.229278 | 0.241223 | 0.241920 | 0.245087 | 0.254695 | 0.262623 | 0.268814 | 0.272802 |
| 608C1_0 | 608C1_0  0.000000 | 0.169424 | 0.205602 | 0.229079 | 0.243963 | 0.248934 | 0.251772 | 0.251802 | 0.252350 | 0.254874 |
| 608F8_0 | 608F8_0  0.000000 | 2  0.165301 | 0.170600 | 0.200060 | 0.201394 | 0.206687 | 0.207691 | 0.213495 | 0.214552 | 0.215527 |
| 608FC_0 | 608FC_0 |  |  |  |  |  |  |  |  |  |
|  | 0.000000 | 0.206185 | 0.215868 | 0.243649 | 0.268874 | 0.282170 | 0.291061 | 0.297753 | 0.307188 | 0.316846 |

本文通过提取图像特征并在Milvus中进行向量搜索，计算了检索结果的TOP@1、TOP@5和TOP@10精确率。结果如表1。

|  |  |  |  |
| --- | --- | --- | --- |
|  | TOP@1 | TOP@5 | TOP@10 |
|  | 100% | 100% | 90% |
|  | 100% | 100% | 100% |
| 60858_0 | 100% | 100% | 80% |
| 6003C_0 | 100% | 100% | 80% |
| 608C1_0 | 100% | 100% | 70% |
| 608F8_0 | 100% | 80% | 70% |
| 608FC_0 | 100% | 100% | 90% |

表1 Milvus 图像检索精确率统计（部分）

进一步验证Milvus向量数据库在检索性能上的优越性，本实验通过传统图像检索技术和采用Milvus向量数据库检索技术的在时间效率和准确率上对比试验。

|  |  |  |  |
| --- | --- | --- | --- |
| 检索方法 | TOP-1 | TOP-5 | TOP-10 |
| LSH | 120.43 | 141.20 | 165.79 |
| Milvus | 0.06 | 0.02 | 0.01 |

表2 检索时间的对比

|  |  |  |  |
| --- | --- | --- | --- |
| 检索方法 | TOP-1 | TOP-5 | TOP-10 |
| LSH | 98% | 82% | 70% |
| Milvus | 100% | 97% | 92% |

表3 准确率的对比

实验表明，传统的图像检索技术在处理大规模甲骨文字形图像数据时，检索速度随着数据量的增加显著下降，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低的响应时间。

在检索精度方面，Milvus展现了明显的优势。由于Milvus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图像之间的细微差异，从而保证TOP-N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。

在系统负载性能方面，传统图像检索技术在处理大规模数据时消耗更多资源，且随着数据集的增大，性能显著下降。Milvus凭借其优化的向量索引机制，能够在保持检索速度的同时，较好地控制资源消耗，展现出优越的扩展性。

# 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TOP-1、TOP-5和TOP-10精确率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。

**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 Wenyuan. https://jgw.aynu.edu.cn/, last accessed 2024/6/5.