# 题目

# 摘要

甲骨文拓片图像分割是甲骨文数字化工程的基础问题，其目的是利用数字图像处理和计算机视觉技术，在甲骨文原始拓片图像的复杂背景中提取出特征分明且互不交叠的独立文字区域。它是甲骨文字修复、字形复原与建模、文字识别、拓片缀合等处理的技术基础。本文采用了计算机视觉技术和深度学习框架，以期提高识别的准确性和效率。

针对问题一，本研究对附件一的甲骨文图像数据进行数据预处理。在处理甲骨文图像数据的预处理工作中，本研究首先调整图像到模型所需的特定尺寸和长宽比。接着，图像的像素值被缩放到一个确定范围（如0至1或-1至1），以便模型能更有效地进行学习。为了增强模型对新情况的适应能力，本研究采用了多种方法增加数据的变异性，如随机的旋转、裁剪、翻转以及亮度和对比度的调整。此外，还对图像进行了增强处理，比如调整亮度、对比度、锐度和颜色等，来强化图像的特征表现。关于图像中的点状噪声处理，本研究对中值滤波和双边滤波进行了对比分析。中值滤波通过替换每个像素为其周围像素的中值来有效移除噪声，但可能会使图像边缘细节变得模糊。相反，双边滤波在减少噪声的同时保持了边缘的清晰度，因为它同时考虑了像素之间的空间关系和值的相似性，这有助于避免过度模糊。因其在维护图像质量，特别是边缘信息方面的表现更佳，双边滤波特别适合需要精确边缘维护的甲骨文检测。此外，通过应用对比度受限的自适应直方图均衡化技术（CLAHE），图像的局部对比度得到了显著提升，使得细节更加明显，这为后续图像的处理和分析提供了坚实的基础。

针对问题二，本研究基于Yolov5模型和附件2的数据对模型进行了微调，赋予其单字检测和分割能力，并展示了部分分割效果。本研究（需修改）~~采用了一种集成自注意力卷积（Self-Attention Convolution, SAC）的升级版Yolov5检测模型，命名为sac-yolov5，~~该模型旨在更有效地捕获甲骨文图像中的深层特征和全局信息，从而提高文本检测的准确性。采用EIOU损失函数取代传统CIOU损失，该策略改善了模型在处理预测框比真实框大时的性能，大幅增强了框对齐的准确度。鉴于甲骨文图像中背景的复杂性，例如裂缝和噪声可能干扰检测任务，本研究采用了基于ResNet-34的UNet模型对图像进行预处理，以去除这些干扰，为sac-yolov5模型提供清晰的输入图像，大大提升了字符检测的准确率和整体系统效率。

针对问题三，本研究使用改进的u-net-sac-yolov5模型对附件三的测试集进行了精确检测，并把结果存储在了“Test results.xlsx”文件中。

针对问题四，~~本研究首先开发了一个基于Inception\_v3的分类模型，并通过对训练数据集的微调，调整该模型以适应甲骨文的文字识别任务。之后，本研究使用这个调整后的模型对测试数据进行了文字识别，并保存了识别结果。~~此外，本研究还从混合甲骨文数据集中选取了单个甲骨文样本，使用ResNeSt深度学习框架进行识别。ResNeSt框架通过其创新的分割注意力机制，有效增强了模型对图像不同区域的关注度，显著提高了对复杂字符形态的识别准确性。这种技术特别适用于处理甲骨文这类历史文本，因为它能够有效地应对文本的多样性和变异性。利用这种方法，本研究不仅实现了高精度的字符识别，还对结果进行了可视化展示。

关键词：yolov5 ~~inception\_v3~~

目录：

# 一. 问题重述

## 1.1问题背景

甲骨文是一种刻在龟甲或兽骨上的古老文字，是中国迄今为止发现的最早的成熟文字系统,是汉字的来源,也是中国优秀传统文化的根源，记载了华夏文明最早期的历史文化：早在西周时期，帝王就会使用龟甲或兽骨进行占卜，并把占卜的结果刻在上面，因此识别和解读甲骨文字对于传承中华文脉极具重要意义，是珍贵的第一手资料[1]，现在甲骨文的研究已经进入信息时代,2019年在安阳市举行的纪念甲骨文发现120周年国际学术研讨会的开幕式上,甲骨文大数据平台——“殷契文渊”正式发布。该平台是甲骨文知识共享平台,对全世界的学者免费开放。

同时，人工智能和大数据技术被应用于甲骨文全息性研究及数字化工程建设，成为甲骨文信息处理领域的研究热点。然而，甲骨拓片图像分割面临着点状噪声、人工纹理和固有纹理等三类干扰元素的严重挑战。甲骨文图像的来源多样，如拓片、照片、扫描或临摹等，不同来源的图像受到这些干扰元素的影响程度各不相同。由于缺乏对甲骨文字及其干扰因素形态的先验知识，目前的通用图像分割技术还无法有效区分甲骨文原始拓片图像中的文字目标与点状噪声、人工纹理、固有纹理等，导致较高的误分割率，对甲骨拓片图像的处理呈现出一定的局限性。如何从复杂且干扰多的背景中准确地分割出独立的文字区域， 仍然是一个亟待解决的具有挑战性的问题。

1.2问题提出

问题 （1）：对于附件1(Pre test 文件夹)给定的三张甲骨文原始拓片图片进行图像预处理，提取图像特征，建立甲骨文图像预处理模型，实现对甲骨文图像干扰元素的初步判别和处理.

问题 （2）：对甲骨文原始拓片图像进行分析，建立一个快速准确的甲骨文图像分割模型，实现对不同的甲骨文原始拓片图像进行自动单字分割，并从不同维度进行模型评估。其中附件2(Train 文件夹)为已标注分割的数据集。

问题 （3）：利用建立的甲骨文图像分割模型对附件3(Test文件夹)中的200 张甲骨文原始拓片图像进行自动单字分割，并将分割结果放在“Test results.xlsx”中。

问题 （4）：基于前三问对甲骨文原始拓片图像的单字分割研究，请采用合适的方法进行甲骨文原始拓片的文字识别，附件4(Recognize 文件夹)中给出了部分已标注的甲骨文字形，请对测试集中的 50 张甲骨文原始拓片图像进行文字自动识别，并以适当结果呈现。

# 问题分析

## 2.1 问题1的分析

**由于古代拓片可能有磨损、裂纹或其他杂质，甲骨拓片图像分割面临着点状噪声、人工纹理和固有纹理等三类干扰元素。为了提升甲骨文图像的质量和清晰度，建立了一个图像预处理模型。噪声污染会严重影响图像特征提取等系列后续处理过程**［2］**采用高斯滤波、中值滤波、双边滤波技术去除图像中的小颗粒和随机噪声。~~接着，将彩色图像转换为灰度图像，简化数据量并降低处理复杂性。通过阈值处理，将灰度图像转换为黑白图像，这一步骤显著提高了文字与背景的对比度，便于进行后续的分割。~~此外，应用直方图均衡化和对比度增强技术进一步提升图像中文字与背景的对比度，从而优化了文字的可识别性。这些预处理步骤共同作用，显著提高了图像处理的效率和文字识别的准确性。**

**2.2 问题2的分析**

**鉴于甲骨文拓片图像的复杂性，传统的图像分割方法难以精确区分文字和干扰元素，需要构建了一个快速而准确的图像分割模型，利用经典和深度学习算法进行有效分割。首先进行了数据预处理，~~包括图像的清洗、标准化尺寸调整以及归一化~~，以确保输入数据的一致性和质量。接着，通过旋转、缩放、裁剪和颜色调整等手段进行数据增强，增加了数据的多样性，从而提升了模型的泛化能力。在模型选择与训练环节，采用了如U-Net和Mask R-CNN这类在图像分割领域表现出色的深度学习模型，它们能够学习复杂的特征并准确地从复杂背景中分割出目标字符。在模型建立后，其性能通过准确率、召回率和F1值等指标进行评估，以验证其有效性和可靠性。**

**2.3问题3的分析**

**基于问题2的模型，需要对测试集中的甲骨文原始拓片图像进行自动单字分割。首先，对“Test”文件夹中的200张图像执行了与训练阶段相同的预处理步骤，~~如去噪、灰度化、二值化和大小调整，~~以保证输入数据的一致性。随后，这些预处理后的图像被输入到已训练的模型中进行自动分割，同时实时监控分割过程以便及时检测并调整任何明显的分割错误或性能问题。分割后，从模型输出中提取每张图像的详细分割结果，包括每个字符的位置和可能的类别信息，并使用Python的pandas库将这些数据格式化并记录，如记录字符的边界框坐标（x, y, width, height）。最终，所有数据被清晰地保存至“Test results.xlsx”文件中，以便上传至竞赛平台进行评估。**

**2.4问题4的分析**

**基于前三个问题的甲骨文图像分割研究，需要开发一个自动识别甲骨文图像中文字的模型。这可能包括应用光学字符识别（OCR）技术或考虑使用深度学习方法，例如循环神经网络（RNN）来进行文字的序列识别。模型建立之后，将对其进行详尽的测试，并通过验证和评估识别结果来确保模型的准确性和可靠性，从而保证模型在处理实际甲骨文图像时的效果和稳定性。**

1. **模型准备**

**3.1 模型假设**

**基于题目的背景，本研究做以下基本假设。**

1. **.甲骨文图像预处理模型假设：**

**· 假设甲骨文拓片图像中的干扰元素，如点状噪声、人工纹理和固有纹理，具有特定特征。**

**· 假设通过适当的图像预处理操作，包括~~去噪、灰度化和二值化~~，采用高斯滤波、中值滤波、双边滤波技术去除图像中的小颗粒和随机噪声可以显著减少这些干扰元素对图像分割的影响，从而提高图像的质量和清晰度。**

**· 假设采用特征提取和判别方法能够有效地将文字区域与干扰元素区分开来，为后续的图像分割任务奠定基础。**

1. **.甲骨文原始拓片图像分割模型假设：**

**· 假设甲骨文图像中的文字具备一些形态特征，如连续性和紧密性等。**

**· 假设所建立的图像分割模型能够准确识别并有效分割出这些文字区域，并且能够处理来自类别的甲骨文图像。**

**· 假设该图像分割模型展现出一定的泛化能力，即它能在未曾接触过的数据集上也实现良好的分割效果。**

1. **.甲骨文原始拓片文字识别模型假设：**

**· 假设甲骨文图像中的文字具备明显的形态和结构特征，例如笔画的组成和顺序。**

**· 假设开发的文字识别模型能够准确识别出甲骨文图像中的文字，并正确输出文字序列。**

**· 假设建立的文字识别模型能够适应各种字体、字形和大小的文字，显示出良好的鲁棒性和泛化能力。**

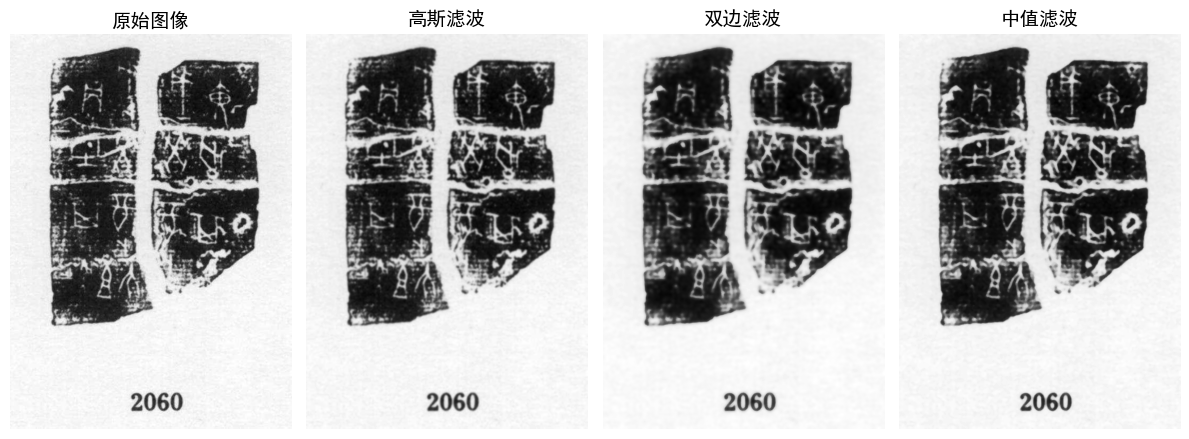
1. **符号说明**
2. **问题一模型的建立与求解**

**在从甲骨文的原始拓片图像中分离出清晰的独立文字区域时，主要的挑战之一就是复杂的背景。这些背景常含有点状噪声、人工添加的纹理以及甲骨文自身的纹理等干扰因素，这些元素显著地降低了文字的清晰度和分割的精确度。因此，在开始文字提取工作之前，对图像进行预处理是事关重大的一步。**

**5.1图像降噪**

**点状噪声通常是由图像捕捉过程中的传感器缺陷或外部环境干扰引起的。这种噪声在图像中以随机分布的亮点或暗点的形式出现，严重影响图像质量和分析的准确度。在观察甲骨文原始拓片图像时，本研究发现了大量点状噪声，这对图像分割的效果极其不利。**

**因此，本研究尝试采用了高斯滤波、中值滤波和双边滤波这三种算法对甲骨文图像进行优化，以确定更适合的处理方法。中值滤波特别适用于消除图像中的“盐和胡椒”噪声，它通过将每个像素点的值替换为其周围像素值的中位数来去除噪声，这种方式能够在不过度模糊图像的前提下清除随机噪点。高斯滤波通过对图像中每个像素点的值应用高斯函数权重的平均来实现。这种方法的核心在于使用正态分布的权重模式，权重中心最高，随着距离中心像素点的距离增加，权重逐渐减少。这种加权平均的方式能够有效平滑图像，减少噪声。而双边滤波是一种更为复杂的技术，它在降噪的同时保持图像边缘的清晰，考虑像素间的强度差异，有效地在平滑图像的同时保留了重要的边缘信息。这对于后续对甲骨文的精确识别和分析至关重要。通过在附件 1 中的图像上应用这三种滤波技术，得到如图5-1所示的可视化结果。**

****

**图5-1**

**虽然三种技术均能有效降低噪点，但双边滤波在处理文字密集的图像中表现更佳，它不仅减少了噪声，同时也更好地保持了文字边缘的清晰度，这对细节的识别尤为关键。**

**相比之下，中值滤波虽然在去除随机噪点方面表现良好，但由于其通过替换像素点为邻域中的中值来执行，某些情况下可能导致图像细节的部分丢失，特别是在文字细节丰富或复杂的区域。在处理轻至中度噪声时，高斯滤波是一个理想的选择，但适用于不希望过度失真的情况下平滑图像。因此，出于能够在最大程度上保持清晰度且尽量减少噪声的考虑下，本研究最终选择了双边滤波方法对甲骨文的原始拓片图像进行降噪操作。**

**5.2增强图像对比度**

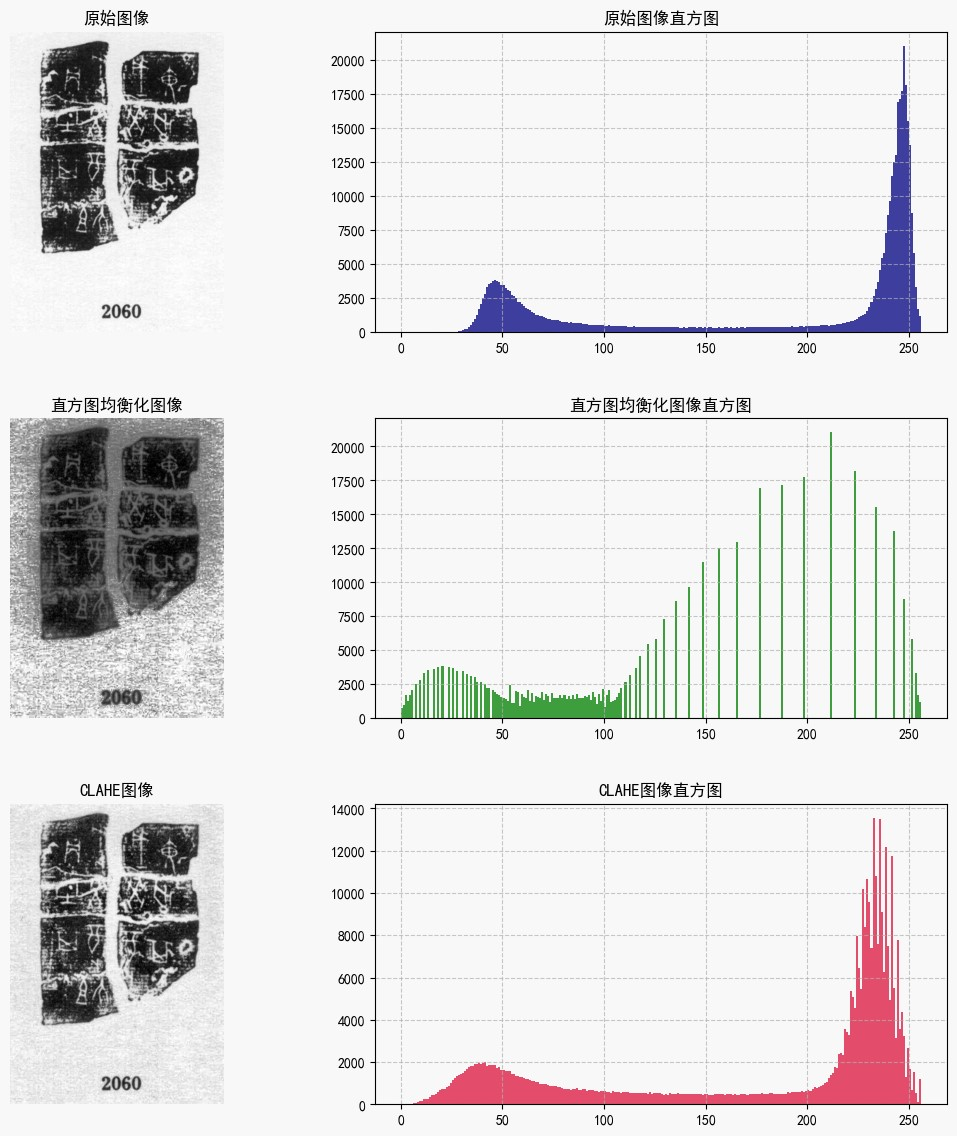
**在甲骨文研究领域，原始拓片的图像处理面临着由于长时间的历史侵蚀，以及拓片在制作过程中可能引入的人为及固有纹理所造成的挑战。这些复杂的纹理问题可能导致部分文字难以识别。**

**为了改善文字与背景之间的可辨识性，本研究采用了图像对比度增强技术，特别是自适应直方图均衡化（CLAHE）。CLAHE 是一种改进的图像对比度增强技术，基于局部对比度的显著性提升而设计。不同于传统的直方图均衡化技术，其对每个小区块进行独立的直方图均衡化处理，从而在不牺牲细节的前提下增强局部对比度，避免了全局处理可能导致的过度增强问题。CLAHE 的数学表达可描述为**

**XXX**

**其中 代表灰度级数， 和 分别代表原始和输出的像素值， pr​(r) 是原始图像的概率密度函数。通过该技术，图像中的每个细节，即使在纹理复杂的区域，也可以得到精细的处理，从而显著提高清晰度和识别度。**

**在上述的图像处理实践中，CLAHE 技术的应用显著增强了甲骨文的可读性。如图5-2可视化结果所示，原始甲骨文图像中的像素值主要集中在较窄的亮度范围内，这在直方图上表现为集中的峰值。这种分布限制了图像的整体对比度，从而影响了细节的清晰度。应用CLAHE之后，可以看到直方图中像素值的分布变得更加广泛和平坦，这表明亮度的动态范围得到了扩展，整体对比度得到了显著的提升。经过对比度增强的甲骨文图像，在清晰度和辨识度方面均得到了显著改进，边缘及内部特征被更为清晰地展示出来，为甲骨文的进一步图像解析和学术研究提供了高质量的图像资源。**

****

**图5-2**

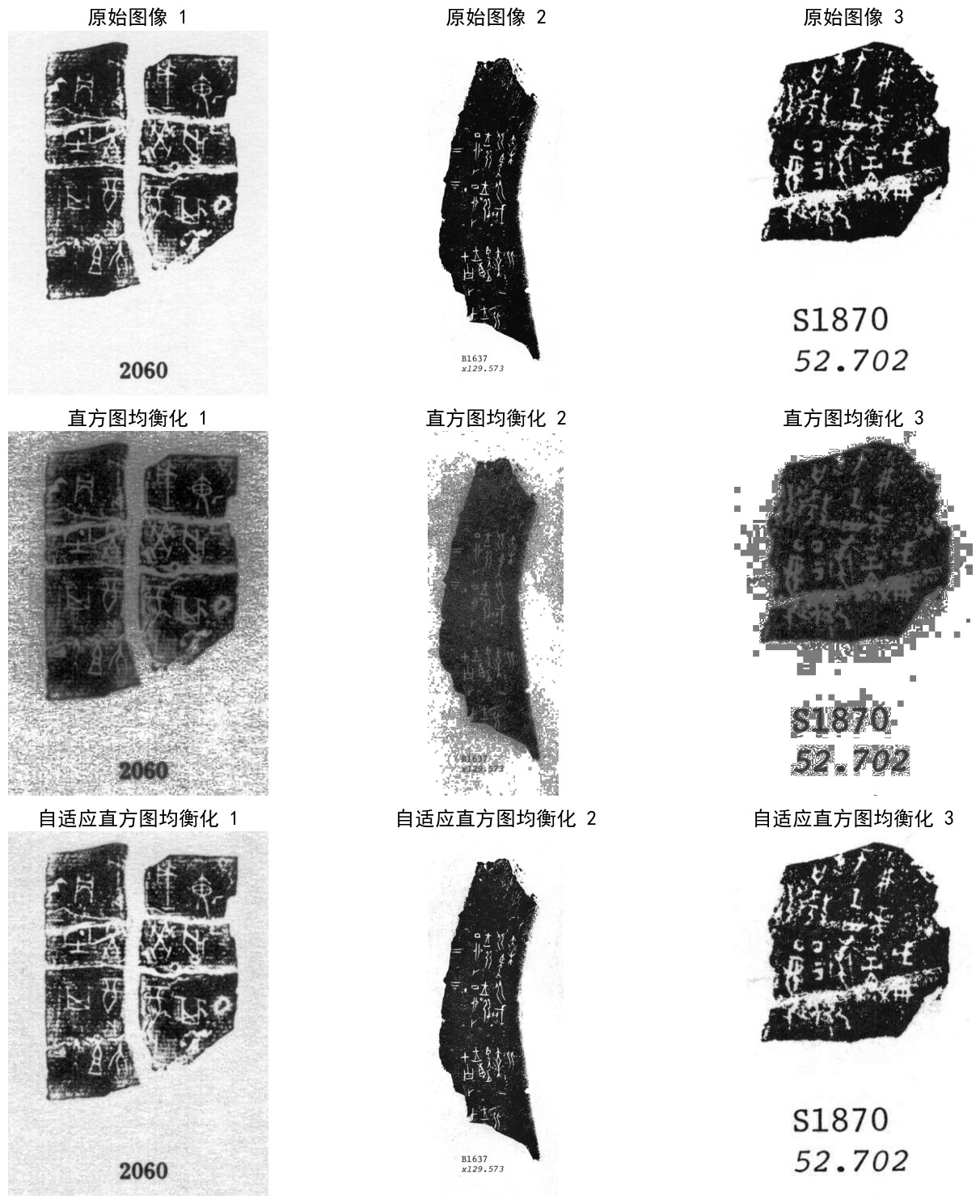
**直方图均衡化是一种在图像处理中广泛使用的技术，其目的是改善图像的全局对比度，尤其是在图像的对比度较低的情况下。这种技术的核心思想是对图像的直方图进行变换，使得结果图像的直方图分布均匀，从而增加像素值的动态范围，让图像的细节变得更加清晰。直方图均衡化的基本数学表达式如上所述，其中 代表灰度级数， 和 分别代表原始和输出的像素值， 是原始图像的概率密度函数。**

**观察直方图均衡化处理后的图像，会发现其直方图在广泛的亮度区间内分布较为均匀。相比于原始图像的直方图，其峰值较为分散，显示了较宽的亮度范围，意味着图像的对比度得到了整体提升。然而，直方图均衡化在增加对比度的同时，也可能会增强图像中的噪声，尤其是当图像包含较大的均匀区域时。**

**与直方图均衡化不同，自适应直方图均衡化（CLAHE）在提高图像对比度的同时，致力于避免过度增强的问题。CLAHE不是对整个图像进行直方图均衡化，而是将图像分割成小块（tiles），并对每一小块独立地进行直方图均衡化。此外，CLAHE引入了对比度限制，防止在直方图的局部区域过度放大对比度，这可以显著降低噪声的放大，同时保持更多的图像细节。因此，CLAHE处理后的直方图通常显示出更加均匀和平滑的亮度变化，更适合细节复杂或对比度不均的图像，如文化遗产图像的处理。**

**直观上，从直方图的形态可以观察到这些差异：原始图像的直方图可能表现为集中的亮度分布，CLAHE处理后的图像直方图分布更广泛，显示出更多的亮度级别和更好的亮度分布。直方图均衡化结果的直方图可能在亮度范围内过于均匀，可能丧失了一些细节信息。通过这些直方图，可以直观地评估图像处理方法对图像数据的影响，这对于图像分析和后续的图像识别工作至关重要。**

**总结来说，直方图均衡化提供了一个全局的对比度增强，适用于对比度较低的图像，而CLAHE提供了一种更精细和控制更好的局部对比度增强方法，特别适用于对细节保留要求较高的情况，如甲骨文这类历史文化图像的分析和解读。**

****

**图5-3**

1. **问题二、问题三模型的建立与求解**
2. **问题四模型的建立与求解**
3. **模型的优缺点**
4. **结论**

甲骨文作为我国最古老的成熟文字系统，在科学研究、文化传承方面具有极其重要的价值。针对甲骨文拓片图像分割

# 参考文献

1. 陈婷珠.殷商甲骨文字形系统再研究[D]. 华东师范大学, 2007.
2. [1]刘勍,温志贤,杨筱平,秦云霞. 现代数字图像噪声滤除技术及其评价[J]. 自动化与仪器仪表,2012,(02):146-148.