

基于图像识别的绳结形成规律的研究

展宝彤 李欣妍 谢逸恒 刘襄南 史旭光（通讯作者）

北京林业大学 理学院，北京，100083；

摘要：本文通过设计实验获得了不同材质、长度绳子所形成的纽结图片。应用多模态大模型和图像识别基础对绳结中交叉点的个数进行了统计。由统计结果发现不同长度的三种材质的线绳随着盒子绕转的圈数的增加，交叉点的数目随之增加。但当圈数很大时，交叉点的数目趋于一个稳定的数值。相较于传统的 OpenCV 图形识别算法，本算法结合多模态大模型，具有轻量化，准确度较高等优点。

关键词：纽结；多模态大模型；二值化；图像识别

DOI：10.69979/3041-0673.25.11.028

引言

一条封闭的曲线在空间扭曲、缠绕形成一个结，在数学上被称为纽结。纽结在物理学和生物学中也有着重要的应用。例如，凝聚态中的磁霍普夫子^[1,2]，光学中的纽结^[3,4,5,6]，量子场论^[7,8]，DNA^[9,10]，量子计算^[11]。在生活中也有很多纽结的例子。例如，放在口袋里的耳机线会随着口袋的运动打结，形成纽结结构等。为了研究绳子打结的规律，人们做了很多的尝试^[12,13,14]。尽管如此，人们对于不同材质、不同长度绳子打结的具体规律依然不是十分清楚。

本文通过设计实验，将不同长度的麻绳、棉线绳、尼龙绳放入翻滚的盒子里形成纽结。然后相机拍照，获得绳结的图片。通过灰度转换、二值化、尺寸调整等预处理图片。选择基于 Kimi 多模态大模型 (MLLM) 的方案对绳结的交叉点数目自动统计。实验结果显示，不同长度的三种材质的线绳随着盒子绕转的圈数的增加，交叉点的数目随之增加。但当圈数很大时，交叉点的数目趋于一个稳定的数值。

1 数据获取

实验中选取常见材质的线绳作为实验对象，分别为麻绳、尼龙绳和棉绳。每种材质的线绳均选取了 2.0 米、1.5 米、1.3 米、1.0 米和 0.7 米。实验装置主要包括：YMS8-D 电机、DM556 驱动器、YDR-24V5A 开关电源电压模块，用于控制电机的旋转。其他设备还包括：铁架台、直径 14 毫米法兰联轴器、米尺等辅助工具。

表 1 不同状态下绳子打结的情况

长度(m)								
材质	麻绳	尼龙绳	棉绳	0.7	1.0	1.3	1.5	2.0
麻绳	0.7	1.0	1.3	1.5	2.0			
尼龙绳	0.7	1.0	1.3	1.5	2.0			
棉绳	0.7	1.0	1.3	1.5	2.0			

选取一根实验用线绳，沿纸箱内壁边缘均匀摆放，确保线绳的开口朝向一致，以减少初始状态对实验结果的影响。封闭纸箱，通过控制面板设置电机的旋转圈数与旋转方向。打开纸箱，使用相机拍照留存，以便后续分析。在同一参数条件下重复实验三次，并改变纸箱旋转方向后再次重复三次，以排除旋转方向对结果的潜在影响。依次调整旋转圈数（2、4、6、8、10、30、50、70、100、150、200 圈），重复上述实验步骤。更换不同材质（麻绳、尼龙绳、棉绳）和长度（0.7 米、1.0 米、1.3 米、1.5 米、2.0 米）的线绳，重复实验过程。

2 基于深度学习的绳结交点识别方法

2.1 算法设计与实现

利用 Kimi API 的强大图像分析能力，结合图像预处理技术，实现了对绳结图像中交叉点数量的自动识别。为提高识别的准确性和效率，需对图像进行了一系列预处理，主要包括：将彩色图像转换为灰度图，减少计算复杂度。使用阈值 128 进行二值化，突出绳子的轮廓。图 1 展示了二值化后绳结的图像。

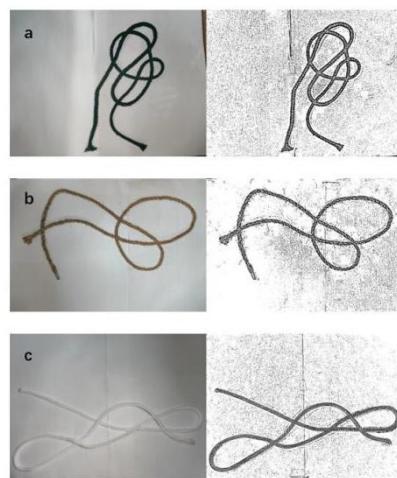


图 1 预处理前后对比图。

左边一列是原始图像，右边一列是二值化后的图像。
a 为棉绳，b 为麻绳，c 为尼龙绳。

选择基于 Kimi 多模态大模型 (MLLM) 的方法对线绳打结的交叉点进行识别。MLLM 的架构通常由三个主要模块组成：模态编码器 (Modality Encoder)、预训练的大型语言模型 (LLM) 和连接这些模态的接口 (Connector)。一些 MLLM 还包括一个生成器 (Generator)，用于输出除文本之外的其他模态^[15, 16]。选择 Kimi API 作为核心识别引擎，是因为其强大的图像分析能力，灵活的 API 接口，便于与现有系统集成，实时处理能力，满足快速识别的需求。

相关的算法流程包括：读取指定路径的绳结图像完成图像输入，应用 3.1 中描述的预处理步骤对图像进行预处理，将预处理后的图像编码为 Base64 格式，创建包含图像数据的结构化请求，向 Kimi 服务发送 HTTP POST 请求完成 API 调用，最后从 API 响应中提取交叉点数量。

2.2 性能评估

选取了 25 张不同复杂度的绳结图像作为测试集，这些图像涵盖了从简单到复杂的各种绳结类型。使用相关的算法对每张图像进行自动识别，记录预测的交叉点数量。由人工数出图像当中的绳结数目，以便评估算法的可靠性。为了全面评估算法性能采用以下指标：平均绝对误差 (MAE)，用来衡量预测值与真实值之间的平均绝对差异。均方根误差 (RMSE) 用来反映预测的整体准确性，对大误差更敏感。相关系数 (R) 用来衡量预测值与真实值之间的线性关系强度。R 方 (R-squared) 用来表示模型解释数据变异性的程度。在允许±1 个交叉点误差的情况下，应用准确率来计算正确预测的比例。

2.3 实验结果

以下是部分组别测试数据的真实值和预测值：

表 2 真实值与预测值对比

麻绳		棉绳		尼龙绳	
真实数量	识别数量	真实数量	识别数量	真实数量	识别数量
2	2	4	4	4	4
1	1	21	8	2	2
6	6	16	11	3	3
9	8	21	8	10	10
6	6	9	6	11	12
1	1	10	10	9	8
3	3	6	7	6	7
2	2	11	9	5	8
3	2	10	9	4	4
4	4	8	8	5	5

由实验结果，获得平均绝对误差 (MAE)：0.56；均

方根误差 (RMSE)：0.87；相关系数 (R)：0.60；R 方 (R-squared)：0.36；准确率 (±1 误差)：0.75。平均绝对误差 (MAE) 为 0.56，表示平均来说，预测值与真实值相差约 0.5–1 个交叉点。均方根误差 (RMSE) 为 0.87，略高于 MAE，说明存在一些较大的误差影响了整体性能。相关系数 (R) 为 0.60，R 方值为 0.36，这表明预测值与实际值之间的线性关系中等强度。这可能是由于某些特殊情况（如图 2）导致的异常预测结果。算法在允许±1 个交叉点误差的情况下，达到了 75% 准确率。数据中存在明显的异常值，例如图 2 所示绳结，由于图像拍摄或处理问题导致算法预测结果严重偏离。这些异常值显著影响了整体性能指标。



图 2 绳结较复杂的情况，出现异常

2.4 误差分析

为了直观地展示算法的误差，创建了以下图表：

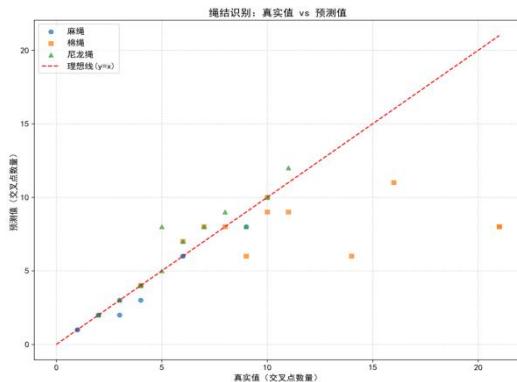


图 3 真实值 vs 预测值的散点图

图 3 展示了预测值与实际值的关系，理想情况下，所有点应该落在对角线上。

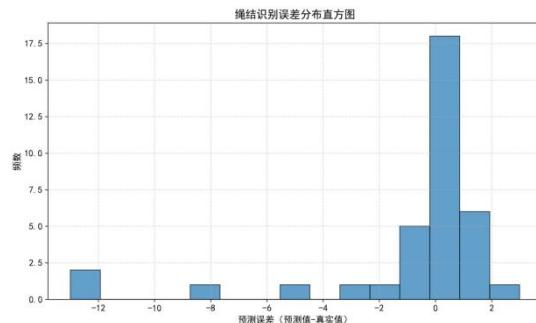


图 4 误差分布直方图

图4显示了误差的分布情况，集中在0附近的误差表示预测较为准确。

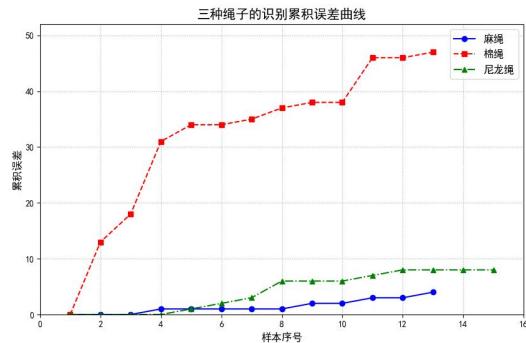


图5 累积误差曲线

图5展示了误差的累积分布，曲线越陡峭表示大部分误差较小。

2.5 关于算法的改进的讨论

算法仍有显著的改进空间。例如，优化图像预处理，改进二值化算法，以更好地处理复杂背景或光线条件不佳的图像。增强数据多样性，扩大训练集，包含更多样化的绳结图像，特别是那些容易导致误判的复杂情况。优化异常检测机制，实施鲁棒的异常检测算法，以识别和处理可能的错误预测，提高整体准确性。后处理优化，在模型输出的基础上，增加基于规则的后处理步骤。通过实施这些改进措施，有望显著提高算法的性能和可靠性，使其在实际应用中更具价值。

3 线绳结实验结论

表1、表2表明不同长度的麻绳、尼龙绳打结数量随长度增加而增加。但棉绳打结数量随长度的增加先增加后减少。其原因在于棉绳的柔性和纤维结构在较长长度下更容易松散。而麻绳和尼龙绳的硬度较大，更易维持扭结形状。

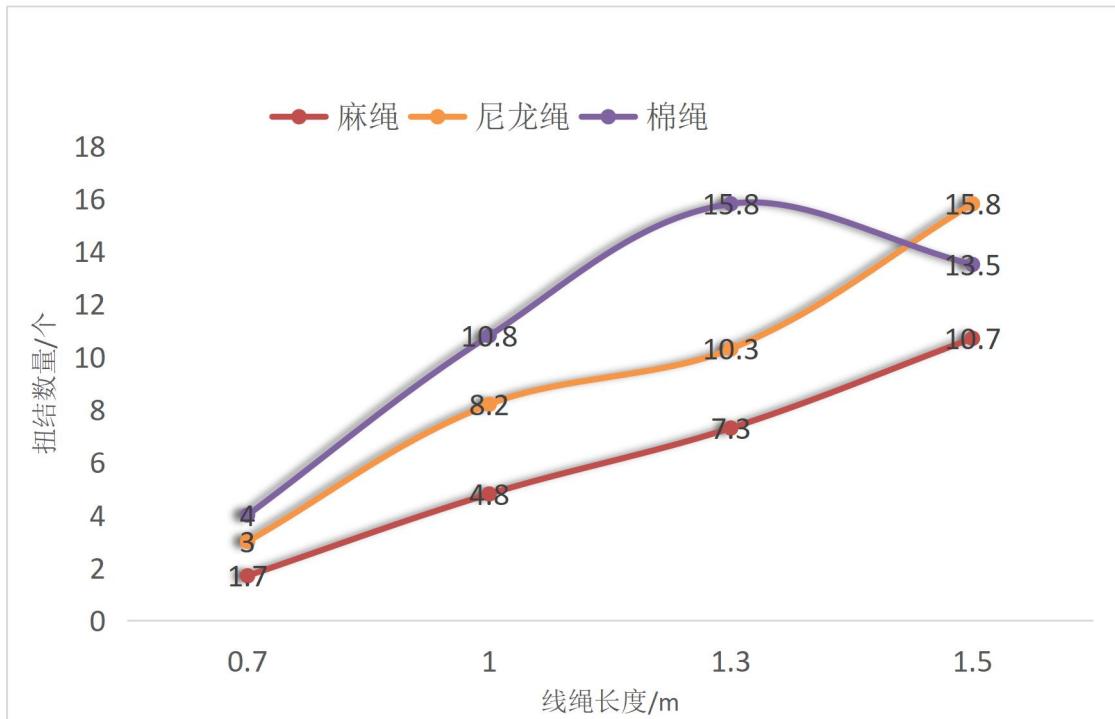


图6 三种材质线绳正向旋转50圈打结平均数量

如图5、图4所示可知麻绳、尼龙绳打结数量随旋转圈数的增加而增加，但当旋转圈数达到一定阈值时，

打结数量趋于饱和，进一步增加旋转圈数对打结数量的影响减弱。

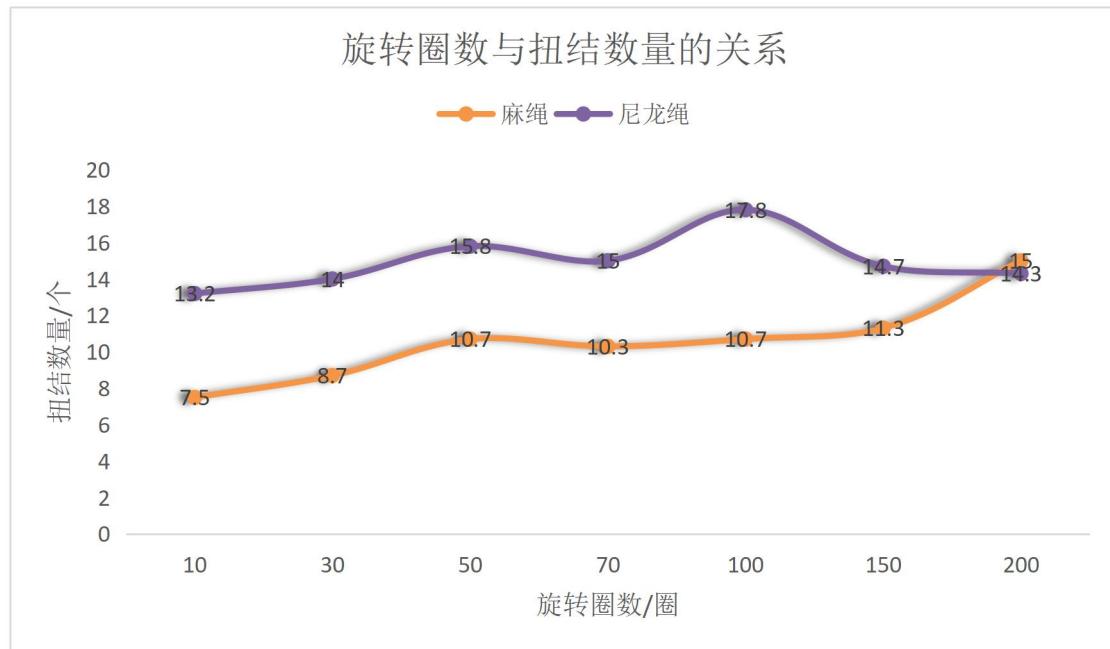


图 7 1.5m 两种材质线绳旋转圈数与打结平均数量的关系

其中棉绳旋转 10 圈及以上圈数后线绳成团状，较难统计交叉点，故不再统计。

参考文献

- [1] Kleckner, D.; Kauffman, L. & Irvine, W. How superfluid vortex knots untie. *Nature Phys.*, 2016, 12: 650 – 655.
- [2] Bisset, R. N. Wang, W. & Ticknor, C.; Robust vortex lines, vortex rings, and hopfions in three-dimensional Bose-Einstein condensates. *Phys. Rev. A*, 2015, 92: 063611.
- [3] Wan, C. Shen, Y. & Chong, A. Scalar optical hopfionseLight, 2022, 2: 22.
- [4] Zhong, J. Wan, C. & Zhan, Q. Optical twisted phase strips. *ACS Photonics*, 2023, 10: 3384 – 3389.
- [5] Zhao, H. Tai, J. S. B.; Wu, J. S. & Li, Y. Liquid crystal defect structures with Möbius strip topology. *Nat. Phys.* 2023, 19: 451 – 459.
- [6] Bauer, T. Banzer, P. & Karimi, E. Observation of optical polarization Möbius strips. *Science*, 2015, 347: 964 – 966.
- [7] Kobayashi, M. & Nitta, M.; Toroidal domain walls as hopfions. 2013, arXiv: 1304.4737.
- [8] Kobayashi, M. & Nitta, M.; Torus knots as Hopfions. *Phys. Lett. B*, 2014, 728: 314 – 318.
- [9] Yang, H. M.; Shi, X. G.; The Free Energy of Nucleosomal DNA Based on the Landau Model and Topology, *Biomolecules*, 2023, 13: 1686.

[10] Wang, Y. & Shi, X. G.; Knot soliton in DNA and geometric structure of its free-energy density. *J. Biol. Phys.* 2018, 44: 81 – 91.

[11] 王正汉; 结绳与量子计算. *数学通报*, 2010, 49(5): 1 – 5.

[12] 鄢盛丰; 朱皓莹; 朱文燕&刘建波; 纽结理论在绳结模型上的应用与分析. 中国运筹学会, 第十三届中国不确定系统年会暨第九届中国智能计算大会论文集, 2015: 430 – 434.

[13] Raymer, D. M. & Smith, D. E. Spontaneous knotting of an agitated string. *PNAS*, 2007, 104(42): 16432 – 16437.

[14] Alexander, K. Taylor, A. J. & Dennis, M. R. Knotting and weak knotting in confined, open random walks using virtual knots. *J. Phys. A: Math. Theor.* 2017, 50: 405001.

[15] Zhang, Y. Li, X. & Wang, Z. A survey on multimodal large language models. 2023, arXiv: 2306.13549.

[16] 刘一闻; 李彤; 王宇琦&张磊; 开源多模态大语言模型初探. *中国传媒科技*, 2024, 31(8): 82 – 85.

*通讯作者：史旭光。