

**研究生学位论文开题报告**

**Graduate Thesis/Dissertation Proposal**

|  |  |
| --- | --- |
| **学号 Student ID** | 122260910018 |
| **姓名 Name** | 许鑫栋 |
| **学生类别 Degree Program** | 专业型硕士生 Professional Master Student |
| **学习形式 Study Mode** | 全日制 Full-time |
| **导师 Supervisor(s)** | 孟建军 |
| **论文题目 Thesis title** | 基于双光照视觉成像的采血管血样分层检测系统设计及性能验证 |
| **学院 School** | 巴黎卓越工程师学院 |
| **专业 Major** | 机械 |
| **开题日期 Date** | 2025年5月 |
| **开题地点 Venue** | 上海交通大学 |

填 报 说 明

**Instruction**

1. 校本部研究生的开题报告应通过[数字交大](http://my.sjtu.edu.cn/)在线提交申请，填写本表并上传系统。特殊情况下经研究生院事先同意，可不上传系统，并使用《上海交通大学研究生论文开题评审表》完成评审。

The application for thesis/dissertation proposal should be submitted online through [My SJTU](http://my.sjtu.edu.cn/). The student shall fill this form and upload it in the system. Under special circumstance, this form does not need to be uploaded and the review can be proceeded with the review form with prior consent from the graduate school.

1. 开题报告为A4大小，于左侧装订成册。各栏空格不够时，请自行加页。考核前提前一周送交导师、评审专家审阅。

This form should be printed with A4 papers and bound together on the left. If the space left is not enough, please feel free to add extra pages. The print version shall be sent to the supervisor, and the review committee members for review at least one week before the oral presentation.

1. 博士生导师可以根据博士生学位论文选题情况自行确定是否进行开题查新，博士学位论文开题查新报告应由查新工作站提供。

The supervisor should decide, based on the proposed topics, whether a novelty assessment report is needed or not, which should be conducted by an authorized novelty assessment department.

1. 开题报告通过后，定稿版开题报告由研究生、导师各存档一份，无需上传系统。

Upon passing the proposal, the final version of this report shall be archived by the graduate student and his/her supervisors for future reference.

1. 同等学力研究生开题答辩采用会议形式，硕士邀请至少3名相关学科/专业领域具有硕士研究生指导资格的专家。博士邀请5名相关学科/专业领域具有博士研究生指导资格的专家。

The capstone presentation adopts a conference format, and at least three experts with master's degree guidance qualifications in relevant disciplines and professional fields are invited for the master's degree. And five experts with doctoral guidance qualifications in relevant disciplines/professional fields are invited for doctoral guidance.

|  |  |
| --- | --- |
| 论文题目  Proposed Title | **基于双光照视觉成像的采血管血样分层检测系统设计及性能验证** |
| 研究课题来源  Source of Research Project | 请在合适选项前画√ Please select proper options by “√”.  国家自然科学基金课题 NSFC Research Grants  国家社会科学基金 National Social Science Fund of China  国家重大科研专项 National Key Research Projects  其它纵向科研课题 Other Governmental Research Grants  企业横向课题 R&D Projects from Industry  自拟课题 Self-proposed Project  其它 Other |

1. **请综述课题****国内外研究进展、现状、挑战与意义，可分节描述。博士生不少于10,000汉字，硕士生不少于5,000汉字。请在文中标注参考文献。 Please review the frontier, current status, challenges and significance of the research topic. The citations should be marked in the context and listed in order at the end of this section. No less than 8,000 words for doctoral students and 4,000 words for master students if written in English.**
2. 研究背景

血液是人体内分布最广、信息含量最丰富的体液之一，既承担氧气和养分的输送，也携带代谢产物、激素与免疫分子，为疾病筛查、诊疗评估和疗效监测提供了极高的信息密度 [1]。在临床检验科的工作流中，一管血液标本往往需要经过一系列预处理步骤——采集、添加抗凝或促凝试剂、静置或离心、分层判读、移液与分析——才能最终产出可被医生解读的数值或图谱。每一个环节都直接决定检测结论的准确性与可重复性 [2]。

1.1 血液样本分类与抗凝策略

按预处理方式不同，血液标本通常分为全血、血浆、血清和凝胶分层管。全血试管不添加任何试剂，可完整保留红细胞、白细胞与血小板，常用于血型鉴定或流式细胞术。若在采血后立即加入肝素锂或柠檬酸钠等抗凝剂并经离心，可获得含凝血因子的血浆上清，该基质多用于凝血功能、生化电解质和血气分析 [1]。血清试管则加入促凝剂或密度分离胶，静置促使血块形成后再离心，最终获得去除纤维蛋白原的黄色透明血清；血清是现代医院最常用的检测基质，涵盖绝大多数生化与免疫项目 [3]。分离胶在离心过程中会迁移至两相界面并形成物理屏障，既保障运输稳定性，也为自动移液提供清晰的穿刺通道。

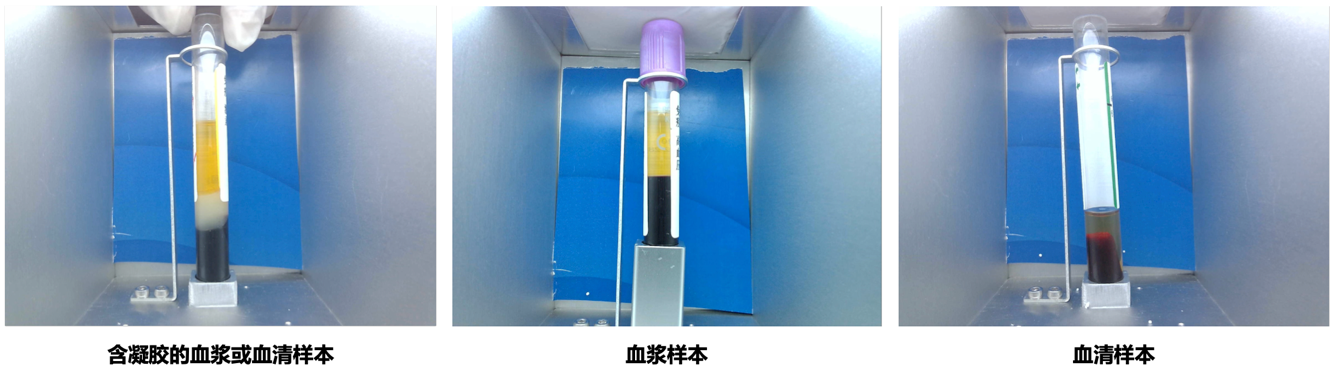


图1. 血液样本的分类

* 1. 血液样本的分层与目标层提取

未经离心的促凝管仅能粗略分为血块与清亮血清，而抗凝管则维持胶体悬浮状态并不适宜直接分析。完成对样本的离心后，促凝管内自上而下可见透明血清、分离胶、血凝块三层；抗凝管内通常分为血浆、白膜层（含白细胞及血小板）和红细胞压积层。白膜层宽度仅 0.5–5 mm，却是提取 cfDNA、PBMC 等高价值组分的关键定位标志，任何分层判读误差都会导致细胞成分交叉污染或目标分子稀释 [4]。

分层后的移液与后续分析通常包含：生化与免疫项目，关注血清或血浆上清，需要无红细胞、无凝块碎片污染；流式细胞或单细胞测序：依赖白膜层的精准穿刺提取，以获得高纯度单核细胞；基因组学与蛋白质组学：cfDNA、外泌体等亦需来源于无细胞成分的上清层。因此，试管内多层界面的准确识别不仅影响初级移液体积控制，还决定后续分析的特异性与信噪比。

* 1. 项目背景

目前大多数检验科仍依赖经验丰富的技师通过目测判断层高与污染情况，再手动设定移液深度。本项目聚焦于医院检验科在庞大医疗数据与高通量样本周转压力下暴露的痛点：传统人工目检血清、血浆与凝胶样本既费时又易出错，难以满足现代检验流水线的效率与精度要求。为此，需要将计算机视觉融入采血管检验分析预处理的环节，期望实现可靠的分层边界和样本类型识别，为后续自动移液、检测等环节提供预处理和结构化数据准备。同时，项目面向整条血液分层检测流水线平台的落地，计划搭建具备精准位移、旋转与夹取能力的机械平台，确保采血管在拍摄、搬运、进出料等环节的连续流转，实现“取—拍—判”一体化的自动化闭环，最终提升实验室处理速度、结果一致性与数据可追溯性。

2　国内外研究进展与现状

计算机视觉在医学影像的早期应用主要聚焦于单模态场景，但随着多光谱成像硬件的普及化以及超声检查与核磁共振检查的影响结果分析的自动化，多模态或多光照图像融合逐渐成为研究热点 [27][28]，为采血管分层、医疗影像检验结果分割等对比度低、透射特性复杂的任务提供了技术契机 [5][6]。

2.1　单模态视觉模型

YOLO系列算法自YOLOv1从2015年提出以来，以端到端、单阶段、速度快的特点在现实生活、工业与医疗各种场景中得到广泛的应用 [7]。最近的YOLOv8–YOLO11在模型骨干网络（backbone）轻量化、空间金字塔池化与注意力机制方面迭代迅速 [8]，但其公开论文不足、模型更新后结构差异难以追溯，给计算机视觉研究领域与社区带来信息不对称 [9]。另外，YOLO系列在各种实际场景的应用仍以单张图像（彩色RGB三通道）为主，对于本项目涉及到的血液检验试管蓝白光图片双光照条件数据集，缺乏通用处理能力。

此外，另有一个国内高校与企业近年在血清质量评估的工作取得的积极进展值得关注。基于 Inception‑ResNet‑V2 的血清干扰分类模型在 30 000 余张标注图像上实现准确率大于 0.98，对溶血、黄疸、乳糜等干扰的识别精度超越与传统视觉模型搭配的阈值限制法 [10]。这一工作显示基于卷积网络架构的视觉模型在血液颜色、边界、材质、浊度等特征提取上的潜力。然而，该工作仍使用单白光图像，无法充分利用蓝光透射的层界增强信息，且该研究聚焦于血清质量分析和血清检验的自动化，未深入探讨血检化验前期预处理工作流程中的实例分割与实时推理场景。

2.2　多模态多光照融合视觉模型

多模态和多光照影像协同分析一直是医学图像检测任务的研究热点 [11]。随着Transformer架构——一种基于注意力机制的深度学习模型组件——在自然语言处理领域（特别是GPT系列模型）中展现出卓越性能，其在计算机视觉领域的潜力也逐渐被挖掘。这种架构通过自注意力机制（Self-Attention）能够高效捕捉输入数据的全局和上下文信息，从而在处理复杂多模态数据时显著提升模型对长程依赖关系的建模能力 [12][13]。

Transformer 进入医学影像分析的标志性工作之一是 TransMed模型，它在核磁共振成像（MRI）多模态肿瘤分类中采用CNN‑Transformer混合架构，显式建模了模态间长程依赖关系，显著提升精度并缓解了小样本过拟合问题 [14]。MicFormer 借助变形跨图像Transformer（Deformable Cross‑attention），通过可学习的局部偏移场完成像素级对齐，在心脏 CT–MRI 语义分割任务上刷新了 DICE 指标 [15]。无监督跨模态对齐亦有新进展，Liu 等提出的“模态协同卷积+注意力机制”框架能处理不配准的多模态对 [16]。这些工作证明了跨模态注意力机制在补偿模态差异、增强语义一致性方面的有效性。不过，上述方法主要聚焦器官影像分割或医学切片分类，特别是MicFormer模型中利用的U型网络架构，对完整大小的图片进行层层递进的特征抽取编码和分割结果生成解码，并最终形成与输入尺寸相同的分割掩码结果，使得该类算法复杂度高、训练难度大、模型参数收敛困难，推理时长延长且依赖大显存 GPU，在实时流水线的边缘端部署仍面临挑战，其模型灵活性差和鲁棒性低不是可以与 YOLO 类架构的模型比较的 [12]。

在产业侧，美国 PerkinElmer 发布的 JANUS Blood iQ 工作站采用蓝光‑白光双成像，结合深度学习分割模型，实现了血浆与白膜层的自动定位并驱动移液机器人 [17]。系统能在 30 min 内处理 96 条样本，对 cfDNA 提取实验具有显著优势。然而，Blood iQ 的算法细节封闭，采取的仍是较传统的基于聚类的视觉算法，多光源照片分析在空间对齐上多依赖固定光路和试管夹具，对异常样本的鲁棒性与场景迁移能力有限。

此外，资源受限设备上，多模态检测亦出现轻量化趋势。YOLO‑Phantom 通过 Phantom Convolution 将参数量压缩数倍，在 RGB/红外双模态低照数据集 FLIRV2 上保持精度的同时显著提速 [18]。RGB‑D Tea 检测网络则在 YOLOv7 基础上引入深度分支和单向跨模态空间注意力，改善了小目标与重叠目标识别 [19]。在航空玻璃缺陷检测中，正背光双模态成像与旋转包围框（OBB）标注结合，被证明能有效定位微裂纹 [20]。然而，这些方法的跨模态交互大多停留于输入级或者特征级的简单拼接或单向注意力，同时也缺乏针对高度非刚性错位的自适应空间对齐机制。

多模态和多光源医学图像的融合方式大体可归纳为输入层（Input‑level）、特征层（Feature‑level）及决策层（Decision‑level）三种层级，其核心差异体现在“信息保真度—内部关联建模—计算代价”三维度的权衡。以下结合前述文献对比分析。

输入层融合以“多通道堆叠”作为典型范式，即将 CT、MRI 或 RGB‑IR 等模态在通道维度直接拼接，作为单一网络的输入。例如 Phantom‑YOLO 首先将红外与可见光帧在输入端合并，随后用轻量化 Backbone 统一提特征 [18]。此策略能最大限度保留原始像素，但模态间分布差异（如 CT 高频骨性结构与 MRI 软组织灰度）在缺乏像素内容敏感对齐方法的引入时易被卷积滤波器误聚类，引入噪声 [15]。TransMed 的早期版本也曾尝试直接堆叠 MRI 多序列，但在肿瘤边缘存在“模态冲突”现象，后续工作才转向中间层对齐 [14]。

特征层融合则分别为各个输入配置专属特征编码器，在中高层语义空间进行融合交互。MicFormer 采用 CNN 提取 CT、MRI 多尺度特征后，引入 Deformable Cross‑attention 生成可学习偏移场，实现像素级对齐再融合 [15]。这种做法既规避了输入层噪声堆叠，又通过共享注意力权重显式建模了不同输入内部对应关系 [21]；只是模型包含两套以上特征抽取主干网络，显存与推理时延开销明显增大。

决策层融合多见于工业流水线与临床闭源系统。PerkinElmer Blood iQ工作站将蓝光与白光图像各自送入独立 CNN，最终在逻辑回归层做阈值投票 [17]；Inception‑ResNet‑V2 的血清干扰检测亦采用类似思路 [10]。其优势在于可对现有单输入模型做“后期加法”，工程迭代成本低；但因融合仅发生于输出标量，无法捕捉跨模态细粒度对应，一旦多模态或多光照图像出现结构性差异，判定稳定性大打折扣。

综合来看，输入层融合强调信息保真度，特征层融合兼顾内部关联与鲁棒性，决策层融合追求工程可插拔。血液分层场景同时具有“微小界面、非刚性错位、实时约束”的三重特性，本课题因此选择以特征层融合为主，辅以形变对齐机制：一方面通过Cross attention机制保留多光照条件互补优势 [22]，另一方面通过轻量化YOLO架构模型把计算成本控制在流水线可接受区间，从而兼顾精度与时延 [23]。

3　关键挑战

在本项目中，蓝白光双光照图像数据集需要被同时有机地结合应用，并深度使用双光照的特征。蓝光与白光反射成像照片受到光路差异、液柱弯月面与气泡折射的共同影响，同时拍摄过程中的角度以及距离上的细微差异，难以避免地会导致分层边界在两种输入间呈现非刚性、非对称位移。传统卷积的平移不变性不足以吸收该差异，需要引入可学习的形变对齐机制 [24][25]。

其次，待检测目标的物理尺度极其微小：白膜层与血清‑血浆双界面之间的总宽度常不足数毫米 ，在高清照相分辨率下对应的有效像素仅数十行。任何来自候选框定位或掩码边缘抖动的微小误差，都可能在后续自动移液过程中被放大，进而导致离心层高测量失真、移液深度不足或过量等连锁偏差。如何在 YOLO类单阶段检测与高分辨分割之间设计低误差的多光照融合图像算法，既保证端到端推理速度，又维持毫米级量化精度，是当前方法论所未充分讨论的盲区。

另外，流水线级的系统实时性同样是一个挑战，为了对标现有的检验科人工肉眼分类和手动移液操作，医院检验科通常要求采血管单管判读时长小于1秒、单日吞吐1万管，任何模型部署必须在边缘 GPU 或低延迟网络环境的服务器上完成低延时推理。如何在保证精度的同时控制参数量与显存占用，是工程化落地的核心技术挑战。

4　研究意义

本课题拟在双光照采血管血液分层场景下，探索“可学习形变对齐—跨输入 Transformer 图像特征层深交互—YOLO实时检测”三者协同的统一框架，对解决医学影像多模态和多光照条件空间异质性与实时推理矛盾提供新范式，可推广至 CT–MRI、RGB–IR 及 RGB–D 等多种多模态组合领域 [26][27][28]。

通过提出针对血清‑白膜层‑血浆三层的像素级对齐与跨输入融合机制，预计可显著提升微小分层界面检测精度，为后续样本提取、血液分析提供高置信度位置信息。本课题将视觉算法与机械执行、条码扫描追溯入库及处理结果数据库储存深度耦合，期望构建部署于服务器或轻量级边缘算力平台的开放式 API；相比现有闭源工作站和血液检验系统，具备更高的灵活性、可维护性与质控透明度。

在社会与经济价值的角度上，在大规模体检、血库样本预处理及多中心队列研究中，自动化、高一致性的血液分层技术可降低较多的人工成本以上、极大地缩短周转时间，并有效减少主观误判导致的样本重测，为医院降本增效、提升患者诊疗质量提供重要支撑。

参考文献 Reference：

1. Holland, Paul V. "Overview: diagnostic tests for viral infections transmitted by blood." Nuclear medicine and biology 21, no. 3 (1994): 407-417.
2. 王永安,王家民.血清分离胶的理论基础及应用[J].上海医学检验杂志,1995,(04):234-236.
3. 段洪云,段玲.检验标本采集与分析前的质量控制[J].国际检验医学杂志,2009,30(04):366+369.
4. 朱晶,赵瀛,王蓓丽,等.不合格血液标本的原因分析及对策[J].检验医学,2014,29(03):288-292.
5. 孔令军,王茜雯,包云超,等.基于深度学习的医疗图像分割综述[J].无线电通信技术,2021,47(02):121-130.
6. 朱翌,李秀.医学图像描述综述：编码、解码及最新进展[J].中国图象图形学报,2023,28(07):1990-2010.
7. Vijayakumar, Ajantha, and Subramaniyaswamy Vairavasundaram. "Yolo-based object detection models: A review and its applications." Multimedia Tools and Applications 83, no. 35 (2024): 83535-83574.
8. Jegham, Nidhal, Chan Young Koh, Marwan Abdelatti, and Abdeltawab Hendawi. "Evaluating the evolution of yolo (you only look once) models: A comprehensive benchmark study of yolo11 and its predecessors." arXiv preprint arXiv:2411.00201 (2024).
9. Hidayatullah, Priyanto, Nurjannah Syakrani, Muhammad Rizqi Sholahuddin, Trisna Gelar, and Refdinal Tubagus. "YOLOv8 to YOLO11: A Comprehensive Architecture In-depth Comparative Review." arXiv preprint arXiv:2501.13400 (2025).
10. Yang, Chao, Dongling Li, Dehua Sun, Shaofen Zhang, Peng Zhang, Yufeng Xiong, Minghai Zhao, Tao Qi, Bo Situ, and Lei Zheng. "A deep learning-based system for assessment of serum quality using sample images." Clinica Chimica Acta 531 (2022): 254-260.
11. Harting, M. T., J. M. DeWees, K. M. Vela, and R. T. Khirallah. "Medical photography: current technology, evolving issues and legal perspectives." International journal of clinical practice 69, no. 4 (2015): 401-409.
12. Zhang, Zixiao, Xiaoqiang Lu, Guojin Cao, Yuting Yang, Licheng Jiao, and Fang Liu. "ViT-YOLO: Transformer-based YOLO for object detection." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 2799-2808. 2021.
13. 朱张莉,饶元,吴渊,等.注意力机制在深度学习中的研究进展[J].中文信息学报,2019,33(06):1-11.
14. Dai, Yin, Yifan Gao, and Fayu Liu. "Transmed: Transformers advance multi-modal medical image classification." Diagnostics 11, no. 8 (2021): 1384.
15. Fan, Xinxin, Lin Liu, and Haoran Zhang. "Multimodal information interaction for medical image segmentation." arXiv preprint arXiv:2404.16371 (2024).
16. Liu, Hong, Yuzhou Zhuang, Enmin Song, Xiangyang Xu, Guangzhi Ma, Coskun Cetinkaya, and Chih‐Cheng Hung. "A modality‐collaborative convolution and transformer hybrid network for unpaired multi‐modal medical image segmentation with limited annotations." Medical physics 50, no. 9 (2023): 5460-5478.
17. JANUS® G3 Blood iQ™ Workstation Report
18. Mukherjee, Shubhabrata, Cory Beard, and Zhu Li. "MODIPHY: Multimodal Obscured Detection for IoT using PHantom Convolution-Enabled Faster YOLO." In 2024 IEEE International Conference on Image Processing (ICIP), pp. 2613-2619. IEEE, 2024.
19. Wu, Yanxu, Jianneng Chen, Shunkai Wu, Hui Li, Leiying He, Runmao Zhao, and Chuanyu Wu. "An improved YOLOv7 network using RGB-D multi-modal feature fusion for tea shoots detection." Computers and Electronics in Agriculture 216 (2024): 108541.
20. Li, Zijian, Yong Yao, Runyuan Wen, and Qiyang Liu. "Dual-Modal Illumination System for Defect Detection of Aircraft Glass Canopies." Sensors 24, no. 20 (2024): 6717.
21. Zhu, X., W. Su, L. Lu, B. Li, X. Wang, J. Dai, and Deformable DETR. "Deformable transformers for end-to-end object detection, arXiv preprint,(2020)." arXiv preprint arXiv:2010.04159.
22. Doersch, Carl, Ankush Gupta, and Andrew Zisserman. "Crosstransformers: spatially-aware few-shot transfer." Advances in Neural Information Processing Systems 33 (2020): 21981-21993.
23. INTHIYAZ, SYED, SK HASANE AHAMMAD, A. Krishna, V. Bhargavi, D. Govardhan, and V. Rajesh. "YOLO (YOU ONLY LOOK ONCE) Making Object detection work in Medical Imaging on Convolution detection System." International Journal of Pharmaceutical Research (09752366) 12, no. 2 (2020).
24. 李思进.深度卷积网络图像分类平移不变性的特征可视化[D].长安大学,2024.DOI:10.26976/d.cnki.gchau.2024.000999.
25. 曹张哲.基于放缩与旋转的复合等变性卷积神经网络研究[D].电子科技大学,2021.DOI:10.27005/d.cnki.gdzku.2021.001097.
26. Xue, Yongjie, Zhiyong Ju, Yuming Li, and Wenxin Zhang. "MAF-YOLO: Multi-modal attention fusion based YOLO for pedestrian detection." Infrared Physics & Technology 118 (2021): 103906.
27. Pemasiri, Akila, Kien Nguyen, Sridha Sridharan, and Clinton Fookes. "Multi-modal semantic image segmentation." Computer Vision and Image Understanding 202 (2021): 103085.
28. Cui, Chenhang, Jinyu Xie, and Yechenhao Yang. "Bright channel prior attention for multispectral pedestrian detection." arXiv preprint arXiv:2305.12845 (2023).
29. **课题研究目标、主要研究内容和拟解决的关键问题。 Research objectives, main contents and key issues to be solved.**

2.1 研究目标

本课题旨在构建一套面向临床高通量场景的蓝光-白光双光照条件血液分层视觉-机械臂协同系统。具体目标包括：

在视觉算法理论层面上提出一种融合可学习形变对齐与跨输入 Transformer 深交互的轻量检测框架，系统的识别错筛率和漏筛率均小于1%，最小可识别单层高度0.5mm，并保持小于一秒的端到端视觉算法推理时延，且模型需要具有迁移和进一步训练的能力。

工程和数据库搭建层面上，设计并实现具备毫米定位精度的直线模组-旋转夹持机械平台，与视觉算法通过统一消息总线耦合，形成“取-拍-判-取”闭环，并构建可保存图片、分层高度信息、条码信息、样品状态等信息的可交互数据库。

2.2 研究内容与拟解决问题

（1）双光照条件数据采集与精细标注，基于我们布置的实验平台中的蓝光和白光光源与相机，开发同步触发装置以获取高一致时空的配对图像；采用roboflow平台对620组白光-蓝光双通道图像数据进行半自动的像素级人机共创标注，对血清、白膜、血浆界面、气泡与白膜厚度进行精密标注，形成规模化数据集。

（2）对跨输入融合算法研究主要包含，YOLO架构类的双主干网络结构，使用两个并行的特征提取网络，分别处理白光图像和蓝光图像；随后在不同尺度的特征上应用跨输入Attention机制进行图像特征融合和堆叠，进行多尺度语义互补；最后使用共享的检测头部分，处理融合后的特征，生成最终的检测和分割结果；基于YOLO的网络整体架构将会极大程度上方便我们最终实现嵌入式 GPU 低显存推理。

（3）对于机电协同抓取平台的设计，本课题基于现有的工作已经可以实现一套具有两自由度直线模组和一自由度旋转夹持末端的机械臂平台，实现从试管架上精准夹取和试管的高速搬运与 180° 翻转。我们预计继续研究基于STM32F103与Jetson Nano的协同光源控制与照相系统的开发。通过STM32驱动WS2812灯带实现多模式照明（UART控制），Jetson Nano作为上位机协调Modbus协议电机运动与图像采集流程，构建完整的双光色轮流采集机制。

（4）质控与数据治理框架，搭建 PostgreSQL + pgvector 存储体系，将原始图像、推断结果、机械日志与最终检验数据关联入库。实现一个可交互并可以向外主动输出试管图片、分层结果、条形码等信息的数据库。

（5）在某医院检验科流水线部署原型系统进行系统评估与临床验证，评估界面检测准确度、移液体积误差、日处理节拍与系统稳定性。

1. **拟采取的研究方法、****研究方案及其可行性分析。Research methods and research scheme to be adopted and feasibility analysis.**

本课题将在“数据—算法—系统”三级并行推进。首先，在数据层，我们已部署蓝光–白光双光路采集装置并完成 620 组试管图像的半自动像素级标注，如图2；下一阶段将在合作医院连续采集，滚动扩充至 1 000 组，以囊括不同管径、分离胶类型和离心转速下的层厚差异。为弥补中长尾样本，计划在 Roboflow 平台进行五类增广：随机旋转、随机拉伸，光照抖动、局部模糊和通道噪声注入，从而放大数据多样性而不过度偏离真实分布。采集与增广后的图像将实时写入 PostgreSQL + pgvector 数据库，连同真实分层标注与试管条码形成唯一主键，保证后续训练与质控可追溯。

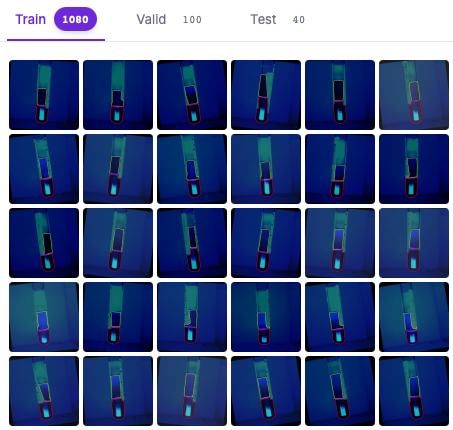


图2. 数据集标注构建

算法侧以 PyTorch为框架实现双-Backbone 主干网络、双光照形变对齐网络、跨输入Transformer与共享Neck/Head四个模块组成。首先，以 PyTorch 搭建两条轻量主干，参数初始化甚至可以直接迁移我们训练完毕的蓝光 YOLO-11，以便快速收敛并保留原有学习到的特征抽取能力，如图3；在每个主干的 P3、P4、P5 三尺度特征图上串接卷积式偏移场预测器与Cross Transformer，实现内容感知的像素级配准；对齐后特征在同尺度送入 Deformable Cross Attention，采用“蓝光 Query/白光 Key-Value”与“白光 Query/蓝光 Key-Value”双向交互，并插入 MLP 细化语义，如图4；经过三尺度融合后的特征进入共享的类YOLO架构的Neck与Head模块，并最终输出分层界面分界线，通过解码后得到掩码与移液深度回归值。整个网络先在 RTX 3090 上以 FP16 预训练 100 epoch，再对 Backbone 与注意力层进行渐进式剪枝，确保参数量符合边缘部署上限；随后以 TensorRT 完成 INT8 量化并转移至 Jetson Nano，验证端到端推理延迟。

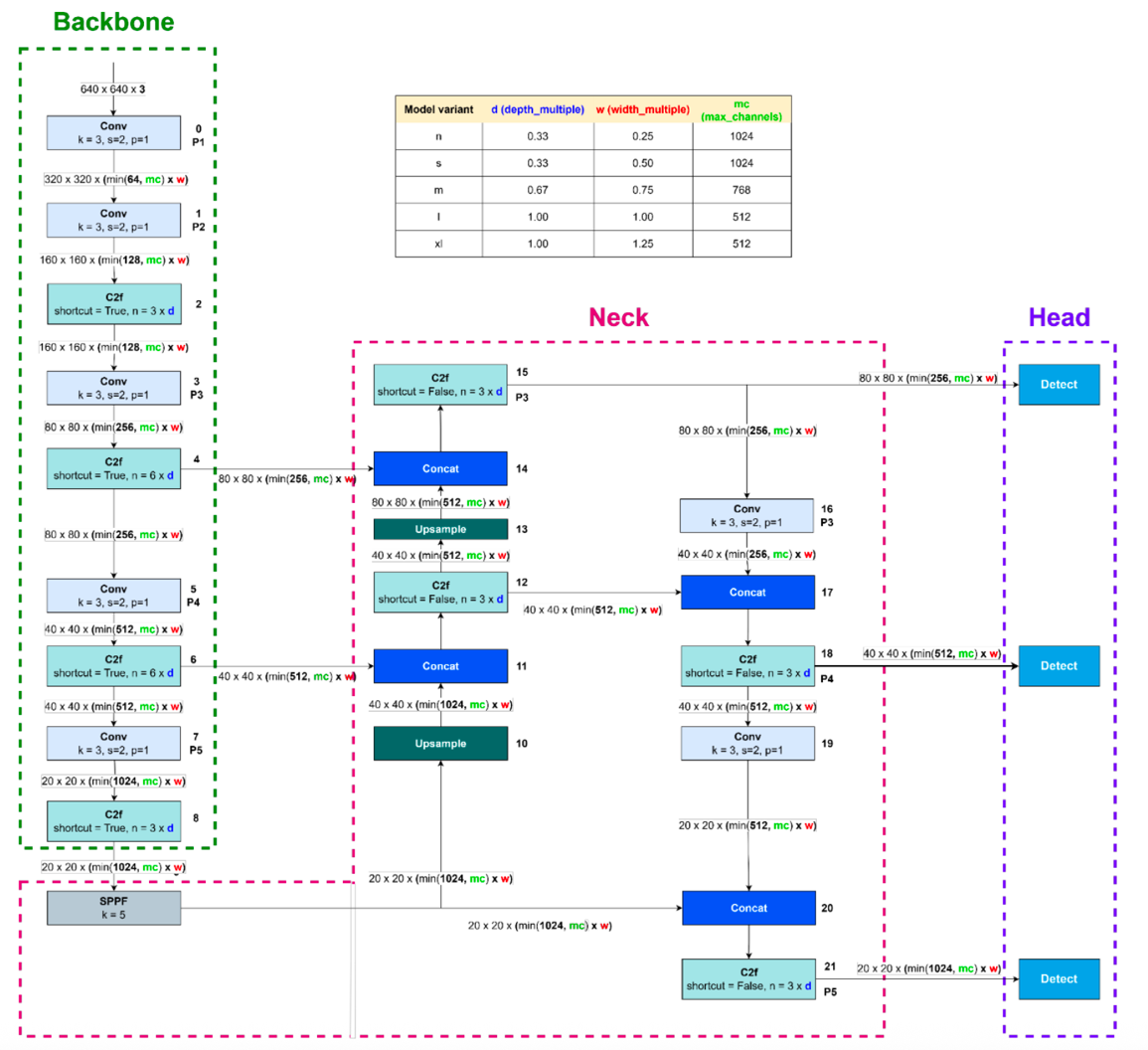


图3. YOLO11骨干网络（backbone）颈部网络（neck）架构

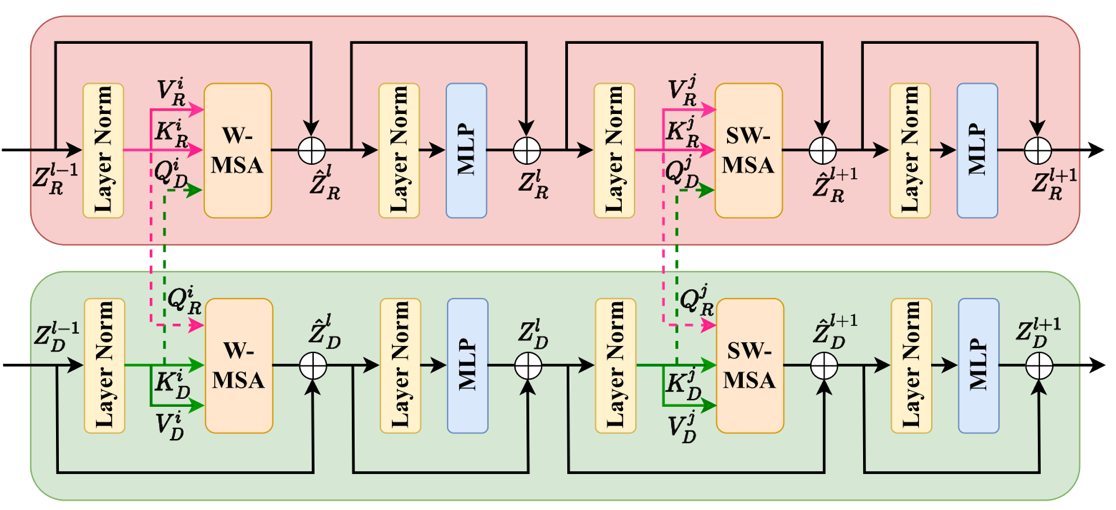


图4. 跨输入Attention模块

在系统层，方案延续已调通的 ROS2 架构：视觉节点运行于 Jetson Nano，周期性触发相机采集双模态图像并推理，如图5；运动控制节点基于 Modbus 指令驱动双直线模组与旋转夹爪，其轨迹点由视觉节点实时发布；光源调度节点借助 STM32F103 串口控制 WS2812 灯带，在蓝/白光模式间轮换，如图5与6，且所有时序信号写回 pgvector 数据库，实现算法输出与机械轨迹的闭环追溯。

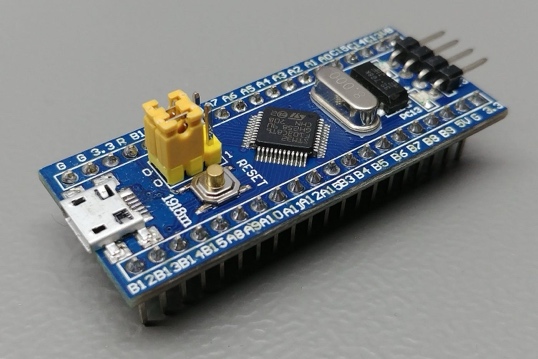
 

图5. STM32F103与Nvidia Jetson Nano算力板

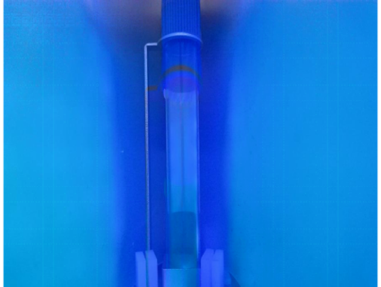
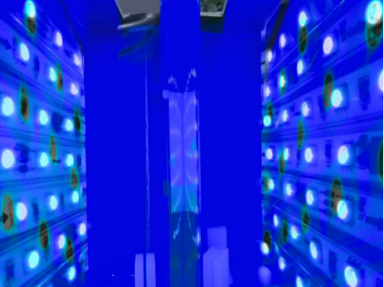


图6. 灯带与柔光板配合下的蓝光光照摄像系统

我们认为该研究方案具有较高可行性。一方面，团队已搭建蓝-白双光路采集装置并完成 620 组像素级标注，且数据集正按既定流程扩充至 1 000 组，结合常规数据增强即可覆盖不同管型与层厚的视觉差异；同时，基于这些数据训练的初步 YOLO-11 模型已能在单蓝光输入下可靠辨识层界，为后续引入白光信息和形变对齐提供了可量化的性能基线。算法环节选择的双 Backbone＋Deformable Cross Transformer 设计与现有单模态网络高度兼容，只需在 PyTorch 框架中增添形变采样和跨光照条件输入的注意力算子即可迭代开发，RTX 3090 的显存与算力足以支持多尺度预训练，而 Jetson Nano 经 INT8 量化后亦已在同类 YOLO 模型上验证过实时推理能力。系统层面，ROS2-驱动的视觉、运动与光源三节点总线已完成实机调试；现有 STM32F103-Jetson 协同平台可实现无明显偏差的试管抓取与搬运，说明机械侧鲁棒性达标且留有算法时延冗余。因此，数据资源、算法框架与机电平台三者均已具备良好基础，整体方案在技术和工程上均可行

1. **课题的创新点 Novelties of the proposed topic.**

变形跨光照条件输入的Transformer的提出是我们算法的核心创新之一，结合了两项关键技术：空间变形网络：使用Deformable Cross Transformer进行像素级的空间对齐，确保不同光照输入间特征的精确对齐；Cross Attention机制：通过跨输入注意力机制进行特征信息的交互，使用一种光照条件的信息对另一个输入中的关键特征进行查找和匹配，深度增强每个输入的表示能力。

基于YOLO架构的多级特征提取，我们预计提供三种不同的多尺度特征融合策略，可根据具体需求选择，且他们的参数都是可训练并对输入的图片具有内容感知的：加权融合，对不同特征层采用不同的权重进行融合；拼接融合，将双光照输入的特征直接拼接，并通过卷积或注意力机制进行融合；Cross Transformer融合，利用Transformer进行多输入之间的特征增强和融合。

1. **计划进度、预期成果 Research schedule, and expected outcomes**

计划进度上，前 1 个月完成采集装置调试与 1 000 组双光照数据标注；2–3 个月内完成形变对齐网络预训练并在 RTX 3090 上取得可发布的算法基线，并将模型量化移植至 Jetson Nano 并与双直线-旋转夹持平台闭环集成，单循环节拍目标为 ≤10 秒；3个月在合作医院进行两周连续运行测试并完成标准的方法学验证。

预期成果包括：一套可在临床演示的原型系统，以及一份蓝-白双光照血液分层数据集与公开的视觉算法，从而在学术、工程与临床应用三个维度形成可量化的综合产出。

1. **与本课题有关的工作积累、****已有的研究工作成绩。Prior experience and accomplished achievements related to the proposed topic.**

课题组已在双光照条件视觉检测与医学检验自动化领域积累了坚实基础：前期搭建的蓝光 / 白光同步采集装置已稳定运行，并在人机协同标注平台上完成 620 组像素级分层数据，为后续算法扩展提供了高质量样本；基于该数据训练的初步 YOLO-11 模型已能在单蓝光输入下实现可靠的血清-白膜-血浆分界检测，如图7，验证了网络在血液图像场景的迁移可行性；算法侧掌握了形变采样、跨输入注意力与 TensorRT-INT8 量化等关键实现细节，能够在 RTX 3090 与 Jetson Nano 上完成从预训练到边缘部署的全链条闭环；机电侧已开发出双直线模组与旋转夹持末端的采血管搬运平台，并通过 ROS2 + STM32 协同实现毫米级无偏差抓取与高速翻转，如图8，表明系统集成与时序控制方案成熟。以上数据、算法与硬件三方面的先行成果为本课题的进一步深入研究奠定了可靠的技术和工程基础。

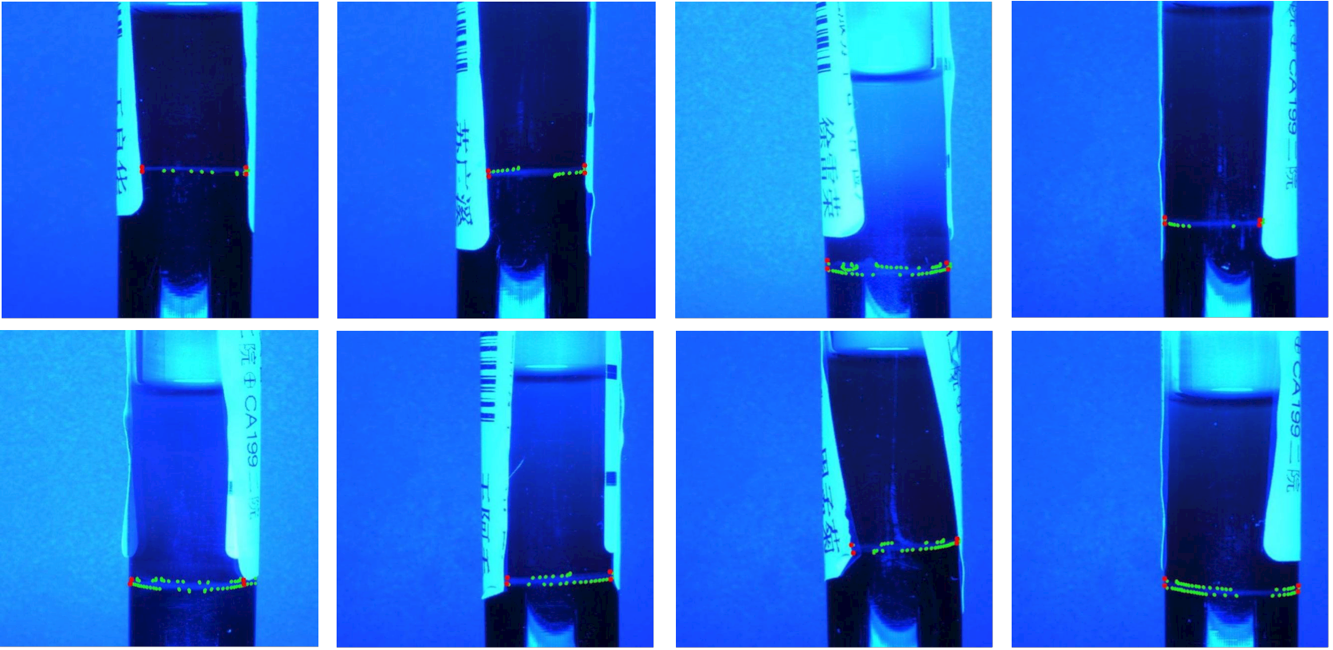


图7. 单蓝光输入下的YOLO11模型分层结果

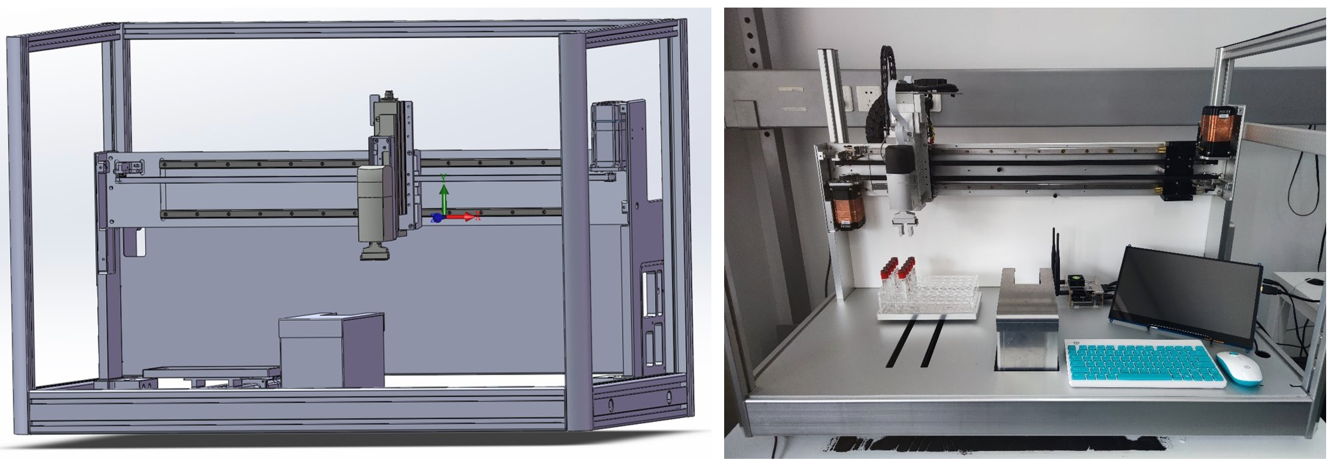


图8. 采血管图像采集和血液分层识别系统的模型​与实物搭建

**本人承诺：开题报告中的内容真实无误，若有不实，愿承担相应的责任和后果。****I hereby declare and confirm that the details provided in this Form are valid and accurate.** **If anything untruthful found, I will bear the corresponding liabilities and consequences.**

**学生签字/Signature of Student： 日期/Date：** 2025-05-29