

基于虚拟染色的免疫组化图像生成

一、赛题背景

肿瘤微环境（Tumor Microenvironment, TME）的精准解析对于肿瘤进展评估、治疗反应分析和预后判断具有重要意义。多重免疫组化（multiplex immunohistochemistry, mIHC）技术能够在同一组织切片中检测多个生物标记物，帮助识别不同细胞类型及其空间分布关系，但其制备流程复杂、实验周期长、试剂成本高，在临床与科研场景中均存在较高使用门槛。

近年来，基于人工智能的虚拟染色技术逐渐成为数字病理领域的重要研究方向。该类方法通过学习源染色图像与目标染色图像之间的映射关系，在尽量减少额外化学染色实验的前提下生成目标标记图像，从而降低实验成本、缩短分析周期，并提升组织样本处理效率。针对多标记联合建模与跨标记信息利用的研究表明，不同免疫标记之间存在一定关联性，合理利用这类信息有助于提升虚拟染色结果的结构一致性与生成质量。

本赛题围绕虚拟染色这一核心问题展开，采用 DAPI 图像作为输入，面向 HLA-DR、CD45RO、Vimentin、CD68 目标标记图像生成任务开展评测。赛题在保证评测可操作性和可复现性的同时，不限制参赛者只能采用单输出方案；在完成官方阶段任务要求的基础上，鼓励参赛者进一步挑战一对多的生成策略。

二、赛题应用场景

本赛题面向数字病理、智能医学影像分析和肿瘤微环境研究等实际应用场景。通过训练虚拟染色模型，可在已有 DAPI 图像的基础上快速预测目标免疫标记图像，为病理科研、临床辅助分析和后续智能分析任务提供低成本、高效率的技术支撑。

1. 临床辅助分析：在实验条件有限、试剂成本较高或检测周期受限时，算法可基于 DAPI 图像快速生成目标免疫标记图像，为病理专家提供辅助参考。
2. 科研预筛与大规模分析：在肿瘤微环境研究中，研究者往往需要对大量组织切片进行初步筛查。虚拟染色可降低部分实验负担，提高样本分析效率。
3. 下游病理智能分析：生成的目标标记图像可进一步服务于细胞分割、区域识别、表达模式分析、诊断辅助与预后评估等任务。

三、赛题任务

参赛者需利用赛事方提供的训练数据，设计并实现人工智能算法，完成从 DAPI 染色图像到目标 IHC 标记图像的虚拟染色生成。

官方评测采用分阶段设置：初赛、复赛、半决赛和决赛。提交的模型算法至少要完成“一对一”生成任务，可选标记有 HLA-DR、CD45RO、Vimentin、CD68 四类

标记。我们鼓励参赛者挑战一对多生成任务。参赛者提交的算法需满足当阶段规定的目标标记生成要求。

算法输入：DAPI 染色图像 patches。

算法输出：与输入图像空间尺寸一致的目标 IHC 标记虚拟染色图像；若采用联合建模方案，可同时输出多个目标标记生成结果，但须确保官方要求的目标标记结果完整、可独立评测。

参赛者需要完成数据读取、模型训练、图像生成、结果保存和推理脚本编写等工作，提交的算法应能够在赛事方提供的测试集上自动完成虚拟染色图像生成。

四、数据集及数据说明

（一）数据来源与染色图像说明

数据来源于多器官组织的 mIHC 配对数据，包含 DAPI、HLA-DR、CD45RO、Vimentin、CD68 等标记。

DAPI 图像是利用 DAPI 荧光染料对细胞核进行染色后得到的图像，主要反映组织中细胞核的位置、形态和分布模式，作为虚拟染色任务中的源图像输入。HLA-DR、CD45RO、Vimentin、CD68 等目标标记图像则反映不同免疫或基质相关成分在组织中的表达与空间分布，作为虚拟染色任务的监督目标。

赛事以 DAPI 图像作为输入源图像。输出目标可选 HLA-DR、CD45RO、Vimentin、CD68 四类标记。

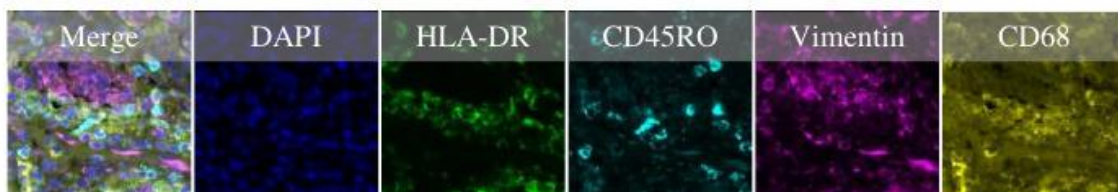


图 1.1 染色图像参考图

（二）数据规模

赛事数据集包含不同器官的 ROI 区域，并将高分辨率 ROI 切分为 256×256 的图像 patch。可参考的规模设置为：colon 数据约 7642 个 patch，liver 数据约 2718 个 patch，stomach 数据约 1836 个 patch。

赛事面向参赛者仅发布训练数据和测试集输入数据。发布的训练数据已经划分好了训练集和验证集。官方评测所使用的各阶段测试集标签均不公开。示例数据集可在 <https://pan.baidu.com/s/1on5a3BE1zj71944HWn9CnA?pwd=ifwv> 获取，正式数据将在报名后开放下载。

（三）数据格式

我们提供训练集数据，包括 DAPI、HLA-DR、CD45RO、Vimentin、CD68；测试集仅提供 DAPI 输入数据。

赛事数据采用如下组织方式：

Datasets

```
--colon
|  |--train
|  |  |--CD45RO
|  |  |  |--ROI000_00_00.jpg
|  |  |  |--ROI000_00_01.jpg
.....
|  |  |--CD68
.....
|--liver
.....
```

图像统一采用 JPG 格式，单张 patch 尺寸为 256×256 。输入图像为 DAPI 染色图像，输出图像为目标标记灰度图或伪彩色图，具体以赛事正式发布的数据说明为准。

比赛阶段	数据集	说明
初赛	Colon 数据集	仅提供训练集和测试集输入
复赛	Liver 数据集	仅提供训练集和测试集输入
半决赛	Stomach 数据集	仅提供训练集和测试集输入

五、算法设计要求

（一）模型类型

本赛题不限定具体算法类型，鼓励参赛者采用深度学习图像生成方法，包括但不限于 U-Net、ResNet、编码器-解码器、GAN、Diffusion、Transformer、注意力机制网络和基于原型学习的可解释生成模型。

（二）创新性

鼓励参赛者围绕虚拟染色生成开展方法创新，例如提升组织结构保持能力、增强局部染色细节恢复、设计结构一致性损失、引入注意力或原型机制提升可解释性、增强跨器官或跨患者泛化能力等。对于能够有效利用跨标记关联信息的一对多方法，也鼓励在技术报告中展示与分析。

（三）可扩展性

算法应具备良好的可扩展性和复现性，能够在不同规模数据和不同计算设备上稳定训练与推理。模型推理过程应自动化完成，不依赖人工后处理。

六、性能指标要求

本赛题采用图像生成任务常用指标评估参赛者提交结果与真实目标标记图像之间的一致性。

1. SSIM (Structural Similarity)：衡量生成图像与真实目标标记图像在亮度、对比度和结构方面的相似性。SSIM 越高，表示生成图像结构越接近真实图像。

2. PSNR (Peak Signal-to-Noise Ratio)：衡量生成图像与真实图像之间的像素级误差。PSNR 越高，通常表示生成图像的像素保真度越高。

七、功能要求

1. 准确性：生成结果应与真实目标标记图像保持较高一致性，尽量减少结构失真、伪影、虚假染色和局部缺失。

2. 鲁棒性：面对不同器官组织、不同患者样本和不同局部组织形态，算法应能稳定输出可靠结果。

3. 可解释性：鼓励参赛者提供模型关注区域、原型激活图、注意力图或其他可解释性分析，说明模型依据哪些组织形态特征进行虚拟染色生成。

4. 泛化性：算法应在未见患者样本上保持稳定表现，不得通过记忆训练数据或利用测试集信息获得不公平优势。

5. 自动化：提交代码应能够按照指定输入路径自动读取测试数据并生成结果，无需人工逐张处理。

八、开发环境

(一) 编程语言

Python，建议使用 Python3.8 及以上版本，因其具有丰富的科学计算库和深度学习框架支持。

(二) 深度学习框架

推荐使用 TensorFlow2.x 或 PyTorch1.x，这两个框架在深度学习领域广泛应用，具有高效的计算性能和丰富的 API，便于模型的搭建、训练和部署。

(三) 硬件环境

参赛者可使用本地工作站或云端计算平台进行开发和训练。本地工作站需配备 NVIDIA GPU（如 GTX10 系列及以上，或 RTX 系列）以加速深度学习计算；云端平台可选择阿里云天池、腾讯云 TI 平台、百度 AIStudio 等，这些平台提供了多种配置的计算资源，方便参赛者根据需求灵活选择。

(四) 依赖库

建议安装 NumPy、OpenCV、scikit-image、Pillow、PyTorch、TorchVision、timm、MONAI 等常用库。具体依赖需在提交说明中给出。

九、成绩评价

本赛题赛事分为初赛、复赛、半决赛和总决赛四个阶段。

（一）初赛阶段

初赛使用测试集 1 进行机器自动评分，主要用于参赛者前期算法验证与调试，成绩不计入总决赛总分。

算法性能得分取参赛者提交结果在测试集 1 上的综合指标得分。

本赛题属于虚拟染色图像生成任务，初赛阶段采用以下指标进行评测：

SSIM (Structural Similarity)：衡量生成图像与真实目标标记图像在亮度、对比度和结构方面的相似性；

PSNR (Peak Signal-to-Noise Ratio)：衡量生成图像与真实图像之间的像素级误差；

综合得分 **Score**：

$$\text{Score} = 70\% \times \text{SSIM} + 30\% \times \text{Normalize(PSNR)}$$

其中，**Normalize(PSNR)** 表示对 PSNR 进行归一化处理。初赛成绩仅用于排行榜展示、模型调试和晋级资格判定。

（二）复赛阶段

复赛使用测试集 2 进行机器自动评分，主要考察参赛者算法在不同测试数据分布下的稳定性与泛化能力，成绩不计入总决赛总分。

算法性能得分取参赛者提交结果在测试集 2 上的综合指标得分。

复赛仍采用 **SSIM**、**PSNR** 和综合得分 **Score** 作为评价指标。仅提交有效成绩的参赛队伍可进入下一阶段。

（三）半决赛阶段

半决赛使用测试集 3，并结合测试集 1、测试集 2 共同构成半决赛综合测试集，由机器自动评分。

半决赛阶段重点考察算法在多器官、多样本来源条件下的鲁棒性与一致性，其成绩计入总决赛总分。

半决赛综合得分按以下方式计算：

$$\text{Semifinal Score} = 10\% \times \text{Score(test1)} + 20\% \times \text{Score(test2)} + 70\% \times \text{Score(test3)}$$

（四）总决赛阶段

总决赛成绩由客观评分与主观评分两部分组成：排行榜得分（70%）由半决赛机器评测得分折算得到；答辩得分（30%）由线下专家评审组根据参赛团队提交的技术报告、方法创新性、模型可解释性、代码复现性、现场答辩表现等进行综合评分。

总决赛最终成绩按以下方式计算：

$$\text{Final Score} = 70\% \times \text{Objective Score} + 30\% \times \text{Subjective Score} + \text{Bonus Score}$$

其中：**Objective Score** 为半决赛机器评测标准化得分；**Subjective Score** 为答辩与材料评审标准化得分；**Bonus Score** 为创新任务附加分。

（五）创新任务附加分

本赛题官方评测以指定测试任务为准，但鼓励参赛者在完成官方要求的基础上，进一步挑战一对多虚拟染色任务。对于采用联合建模策略，并能够证明其对官方任务性能、模型泛化能力或方法创新性具有积极作用的参赛方案，可给予适当附加分。

创新任务附加分规则如下：

一对多联合建模：在完成官方指定目标生成任务的同时，额外实现由同一输入染色图像生成多个目标标记图像，可酌情给予 1-5 分附加分。

附加分判定依据：是否完成官方主任务且结果有效；是否在技术报告中清晰说明联合建模方式；是否提供充分实验结果证明联合建模优于单任务方案；是否具有明确创新性、可解释性和复现性等。

注：为保证比赛公平性，创新附加分仅作为总决赛加分项，不替代官方主任务机器评分，也不影响初赛、复赛、半决赛的晋级规则。

（六）有效成绩说明

若参赛者提交结果低于赛事方公布的基线成绩，赛事方有权将其判定为无效成绩。只有取得有效成绩的参赛队伍，方可进入下一阶段评奖或晋级。

十、解题思路

（一）知识点

- 图像到图像翻译与虚拟染色；
- 医学图像与数字病理图像分析；
- 生成模型与结构一致性约束；
- 注意力机制、原型学习与可解释人工智能；
- SSIM、PSNR 等图像质量评价指标。

（二）思路引导

参赛者可首先构建 U-Net 或 ResNet 编码器-解码器作为基线模型，使用 L1/L2 损失进行监督训练；随后可加入 SSIM 损失、感知损失、边缘约束或结构一致性损失，以提升组织结构保持能力。对于细节恢复不足的问题，可尝试多尺度特征融合、注意力机制或对抗训练。若希望进一步提升模型可解释性，可引入原型学习、区域响应可视化或中间特征约束等机制，将局部组织形态与生成结果之间的关联显式呈现出来。对于高阶方案，可尝试利用跨标记共享表示，在满足官方阶段任务的同时实现多目标联合生成。

（三）注意事项

1. 参赛者应自行划分本地训练集与验证集，同时避免训练时数据泄漏问题。

2. 生成图像不得通过人工修改或测试集后处理提升成绩。
3. 应关注生成结果的结构真实性，避免只追求像素指标而产生不符合病理逻辑的伪影。
4. 提交代码应保证可复现，环境配置和模型加载方式应描述清楚。

十一、赛题约束条件

（一）算法约束

1. 允许使用开源深度学习框架和公开预训练权重。
2. 不允许调用商业闭源在线 API 直接完成图像生成任务。
3. 不允许人工参与测试集结果修正。
4. 最终提交模型应可复现，并能在官方测试环境中完成推理。

（二）数据使用约束

1. 原则上仅允许使用赛事官方提供的数据。
2. 严禁使用隐藏测试集或其衍生信息进行训练、调参或后处理。
3. 严禁泄露、传播赛事数据或将其用于非赛事相关商业用途。

十二、参考资源

（一）文献资料

1. 《深度学习》（DeepLearning），由 IanGoodfellow、YoshuaBengio 和 Aaron Courville 撰写，系统介绍了深度学习的基础概念、模型架构和训练方法，对理解和应用神经网络有很大帮助。
2. Rivenson Y. et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning.
3. Liu Y. et al. Virtual staining of unlabeled microscopy images using deep learning.
4. Isola P. et al. Image-to-Image Translation with Conditional Adversarial Networks.

（二）在线资源

1. PyTorch 官方文档：<https://pytorch.org/>
2. scikit-image 图像处理文档：<https://scikit-image.org/>
3. MONAI 医学影像工具文档：<https://monai.io/>

十三、提交要求

（一）初赛

提交测试集输入数据 **DAPI** 对应的输出目标图像数据文件并保存在 **results** 文件夹下，输出图像名称应与输入图像保持对应关系。提交格式示例如下：

```
results
|--test
| |--CD68
| | |--ROI025_00_00_fake.jpg
.....
```

（二）复赛、半决赛

1. 提交测试集输入数据 **DAPI** 对应的输出目标图像数据文件并保存在 **results** 文件夹下，输出图像名称应与输入图像保持对应关系。

2. 提交完整的算法代码，包括数据预处理、模型训练、预测推理等各个环节的代码。代码需使用 **Python** 语言编写，具备清晰的注释和文档说明，以便评审人员理解和运行。

3. 提交训练好的模型文件，并提供模型的加载和使用说明，包括所需的运行环境、依赖库等信息。模型文件应能够在指定的测试环境中正常运行并输出预测结果。

4. 提交详细的技术报告，内容包括算法设计思路、模型架构图、实验设置（如训练参数、数据增强方法等）、性能分析（对主要指标和次要指标的详细分析）以及算法的创新点和不足之处。技术报告格式采用 **PDF** 格式，字数不少于[2000]字。

注：如果提交模型或结果文件是多输出，将取多个输出成绩的平均分作为本次提交作品的成绩。最终成绩取最后一次提交作品的成绩。

（三）总决赛

提交内容及具体要求以组委会后续正式通知为准。

十四、其他说明

（一）公平性

严禁任何形式的作弊行为，包括但不限于数据泄露、测试集信息泄漏、模型预训练数据与测试数据重叠、抄袭他人代码、人工修改测试结果等。一经发现，立即取消参赛资格，并追究相关责任。

（二）知识产权

参赛者提交的作品必须为原创，未在其他比赛中获奖或公开发表。比赛主办方有权对参赛作品进行展示、宣传等相关活动，但知识产权仍归参赛者所有。

（三）复现要求

赛事方有权对参赛团队提交的代码、模型和结果进行复现。如使用官方数据和说明无法复现主要结果，赛事方可判定成绩无效。

十六、联系方式

赛题交流 QQ 群: 1084060012

邮 箱: linxue202509@nuaa.edu.cn

报名官网: www.aicomp.cn