

Deep Learning Enhanced Mobile-Phone Microscopy 阅读报告

报告人：宋明辉

April 7, 2018

报告内容

- 1 目的
- 2 成像设备对比
- 3 提出的算法
- 4 分析与总结

1 目的

2 成像设备对比

3 提出的算法

- 数据预处理
- 神经网络结构
- 评价指标以及实验结果

4 分析与总结

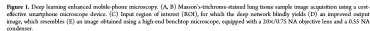


Fig. 1: 示意图

注意：输入图像为通过手机相机得到的图像，输出图像特指经神经网络处理得到的图像，目标图像为通过显微镜得到的图像。

目的：

- 借助深度学习将低成本手机相机的图像增强到与昂贵显微镜的图像相似的分辨率
- 消除输入图像中的噪声、色彩失真等
- 作为一种标准化手段，便于其它医学应用

总而言之，利用低成本、便捷的成像设备，如手机等，代替昂贵、繁杂的专业显微镜。

1 目的

2 成像设备对比

3 提出的算法

- 数据预处理
- 神经网络结构
- 评价指标以及实验结果

4 分析与总结

成像设备对比

参数	手机相机 ¹	专业显微镜 ²
数值孔径 (NA)	较低	0.75
Pixel Size(μm)	1.12	7.4
FOV ³ (mm^2)	~ 1	0.57
曝光时间 (s)	1/49 \sim 1/13	未知

因此，相比于显微镜，手机图像具有以下缺点：

- 分辨率较低
- 敏感性较低, 易引入噪声、色彩失真

¹Nokia Lumia 1020²Olympus IX83

³Field of View

- 1 目的
- 2 成像设备对比
- 3 提出的算法
 - 数据预处理
 - 神经网络结构
 - 评价指标以及实验结果
- 4 分析与总结

预处理的内容

根据实验目的、上述对成像设备的成像特点分析，预处理需要完成以下内容：

- FOV 匹配

即需要保证观察的内容是一致的，否则，计算得到的增强算法就是错误的。

通过 **计算单应矩阵**完成。

- 图像配准

需要保证在更细小结构上的对应。神经网络作用于配准后的图像，这样可以保证学习**输入图像与目标图像**之间的失真关系。

这部分基于 **局部矫正**完成。

- 上采样，这部分由神经网络的最后一层卷积层完成。便于网络的训练、以及保证网络结构的精简。

计算单应矩阵

根据单应矩阵的定义，主要消除因旋转、形变、平移等因素引起的不匹配，表现形式为 3×3 的矩阵，对输入图像进行线性变换。这一步开始之前，需要将输入图像转换成 TIFF 格式或 JPEG 格式。

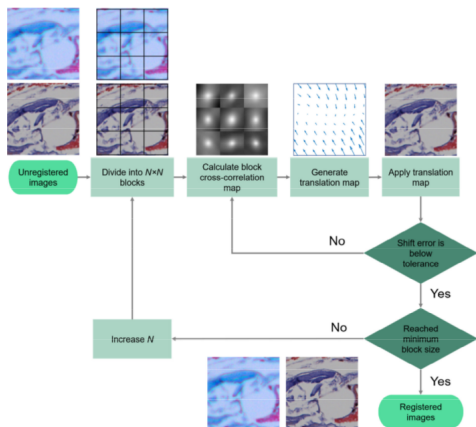
主要计算流程 (经典实现方式):

- 计算 SIFT 特征 (目前来说，应该最优的特征检测算法，或许可以换成计算效率更高的算法)
- 基于 RANSAC 算法计算特征的之间的匹配关系
- 基于匹配的特征点计算单应矩阵

这一步完成后，得到全局匹配 (Globally Matched) 的图像对。

局部配准

基于 Pyramid Elastic Registration Algorithm 完成



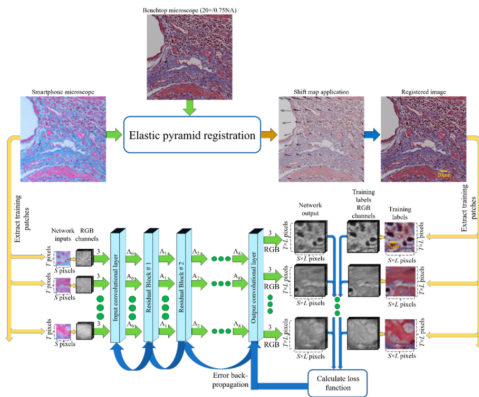
主要步骤:

- 将输入图像、目标图像迭代地分割成 $N * N$ 大小的 Block
- 计算对应 Block 的 Cross-Correlation(互相关)
- 计算转换图 (Translation Map)
- 根据转换图对 Block 进行处理 (线性变换)

Figure S2. Pyramid elastic registration algorithm.

- 1 目的
- 2 成像设备对比
- 3 提出的算法
 - 数据预处理
 - 神经网络结构
 - 评价指标以及实验结果
- 4 分析与总结

概覽



有监督学习：输入为手机成像结果，Label 为显微镜成像结果

基本的卷积块：残差块 (Residual Blocks)

网络结构 1

Residual Block:

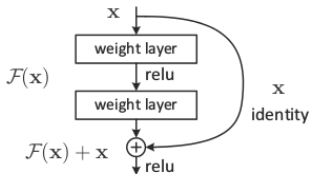


Figure 2. Residual learning: a building block.

数学表达式:

$$X_{k+1} = X_k + ReLU(Conv_{k_2}(ReLU(Conv_{k_1}(X_k))))$$

卷积核大小: 3×3

网络结构 2

各个 Residual Block 的 Feature map 数量:

$$A_k = A_{K_1} + \text{floor}((\alpha \times k)/K + 0.5)$$

其中, $K = 5$, $k = [1 : 5]$, $\alpha = 10$, $A_0 = 32$.

Upsampling:

最后一层卷积层输出 3 个通道, 分别对应 R、G、B 三个通道;
同时完成升采样的过程: 不清楚是通过 Deconvolution 还是其它的方式实现。

损失函数

损失函数为 MSE(均方误差:Mean-Squared Error) 与一个正则项的和:

$$l(\theta) = \frac{1}{3 * S * T * L^2} \left[\sum_{c=1}^3 \sum_{s=1}^{S*L} \sum_{t=1}^{T*L} \| Y_{c,s,t}^{\theta} - Y_{c,s,t}^{Label} \|^2 + \lambda \sum_{c=1}^3 \sum_{s=1}^{S*L} \sum_{t=1}^{T*L} |\nabla Y^{\theta}|_{c,s,t}^2 \right] \quad (1)$$

其中正则项为**输出图像**梯度幅度的平方, $\lambda = 0.001$.

训练: Adam

评价指标、实验结果

评价指标 | Color Distance (CIE 空间) | SSIM⁴

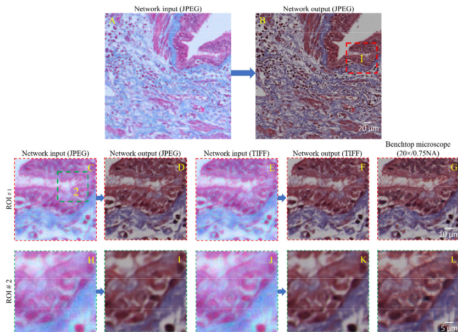


Figure S4. Comparison of the deep network inference performance when trained with lossy compression (JPEG) and lossless compression (TIFF). (A) JPEG-compressed image, and (B) its corresponding deep network output. Zoomed-in versions of (C–G) ROI #1 and (H–L) ROI #2.

- 1 目的
- 2 成像设备对比
- 3 提出的算法
 - 数据预处理
 - 神经网络结构
 - 评价指标以及实验结果
- 4 分析与总结

优势与不足

● 优势

- 可以方便的以迁移学习的方式，部署到其它基于手机的成像系统中
- 输入图像可以为有损压缩后的图像 (如 JPEG)，方便存储以及传输等
- 结果表明，以深度学习为工具可以实现预定目标

● 不足

- 为什么选择以残差块 (ResNet) 形式的神经网络结构 (透明性?)
- 目标函数 (其它的类型是否有效)
- 为了保证网络的精简，一些工作由预处理完成，或许可以设计统一的端到端学习结构

请各位老师、同学批评指正！