



泛在计算与智能系统研究中心
Research Center of Ubiquitous Computing and Intelligent Systems



哈尔滨工业大学
Harbin Institute of Technology

HDNet: High-resolution Dual-domain Learning for Spectral Compressive Imaging CVPR2022

李超楠

2024-12-24

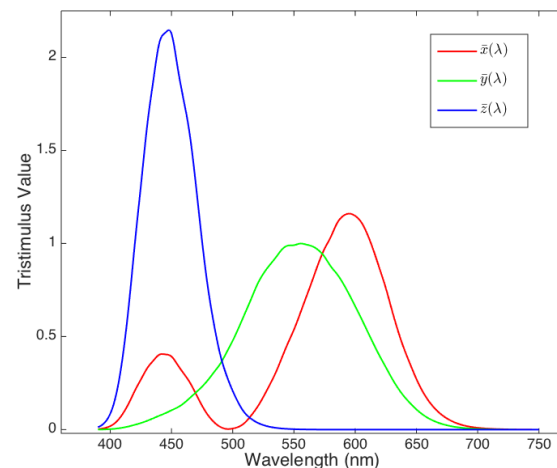
背景

光谱超分辨率（光谱重建）：

高光谱图像(HSI)能够捕获更丰富的场景信息和固定波长成像特性。
HSI到RGB是一个自然过程，HSI到RGB却是需要“凭空”产生更多的通道信息。

但是由于RGB得到的不仅仅是三个单色通道的信息，因为没有无限窄带的滤波器，甚至说，原本用来提取三个通道信息的滤波器本来就是设计为宽带的，RGB图像，其实是将RGB三个宽带滤波器的信息通过隐空间，投射到了三个通道上合成的一张图，本来就含有宽带信息。这个过程反过来，找回原本的宽带信息，然后通过训练找到新的隐空间把宽带信息投射到多个窄带上，形成HSI。本质上还是一个压缩感知。

所以光谱超分辨率任务的存在是合理的。



背景

既然是超分辨率的子任务，不仅仅是通道信息，空间信息也很重要。

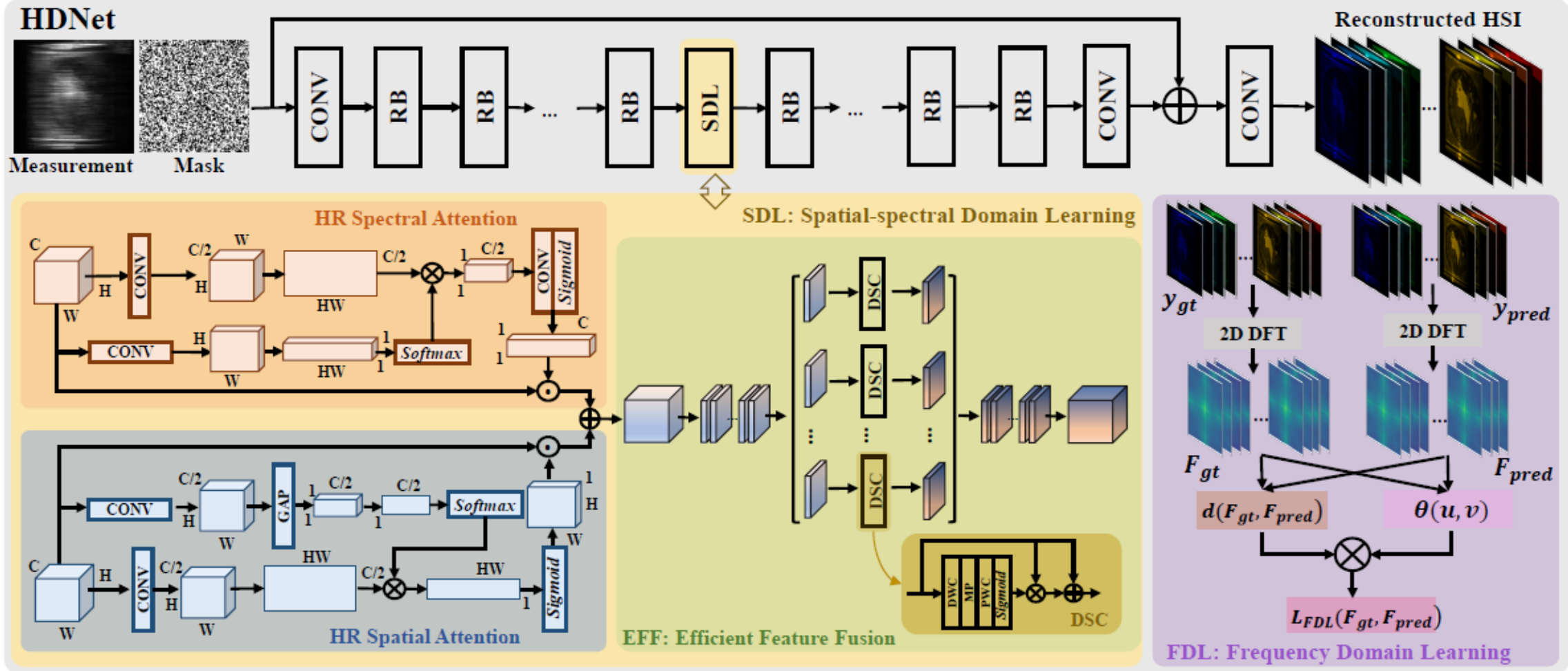
- 卷积能够恢复空间分辨率，但可能忽略了光谱信息的细节，导致高分辨率图像中的光谱信息无法很好地恢复。
- Transformer网络通常为了平衡模型性能和复杂性，牺牲了内部分辨率，导致丢失了细粒度的高分辨率（HR）特征。
- 即使专注于空间-光谱域学习（SDL）的优化收敛到理想解，重建的HSI与真实值之间仍然存在显著的视觉差异。

创新点

这边论文提出了一种用于 HSI 重建的高分辨率双域学习网络 (HDNet)。

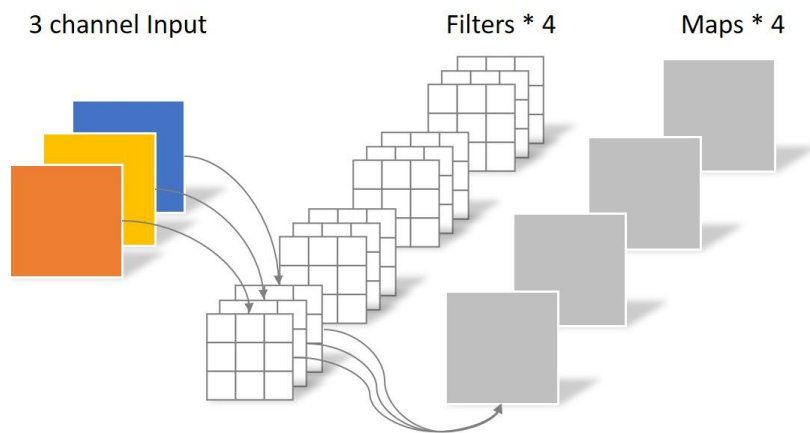
- 提出的高分辨率空间-光谱注意模块及其高效的特征融合提供了连续而精细的像素级特征。
- 频域学习 (FDL) 被引入到 HSI 重建中，以缩小频域差异。动态 FDL 监督迫使模型重建细粒度频率，并对像素级损失造成的过度平滑和失真进行补偿。在 HDNet 中，HR 像素级关注和频率级细化可相互促进 HSI 的感知质量。

模型架构



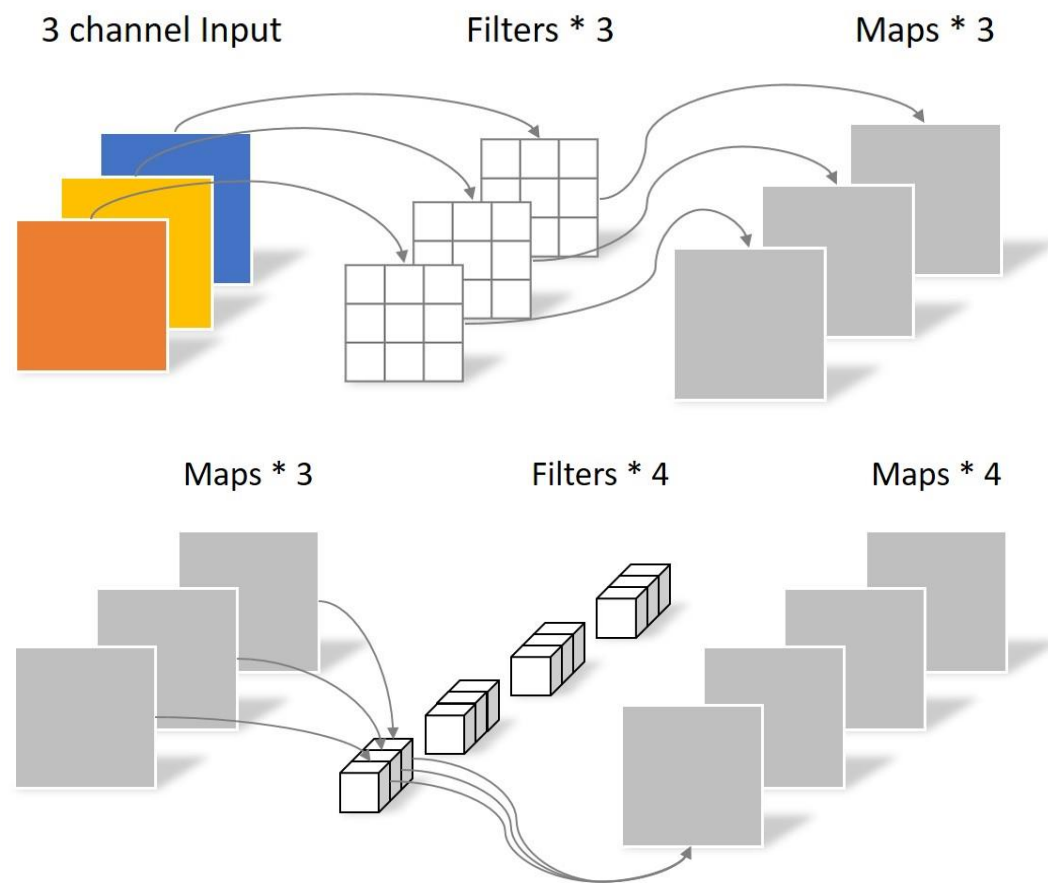
深度可分离卷积(DSC)

替换传统卷积，减少参数量，并加速。



传统卷积: $4 \times 3 \times 3 \times 3 = 108$

DSC: $3 \times 3 \times 3 + 4 \times 3 \times 1 \times 1 = 39$



频域学习

什么是频域？

以信号为例，信号在时域下的图形可以显示信号如何随着时间变化，而信号在频域下的图形（一般称为频谱）可以显示信号分布在哪些频率及其比例。

空间信息也可以转为频域信息。

空间转频域有什么用？

图像高频：亮度或灰度变化激烈的部分，如边缘和轮廓

图像低频：变化不大或变化较为平缓的地方，如大块的结构

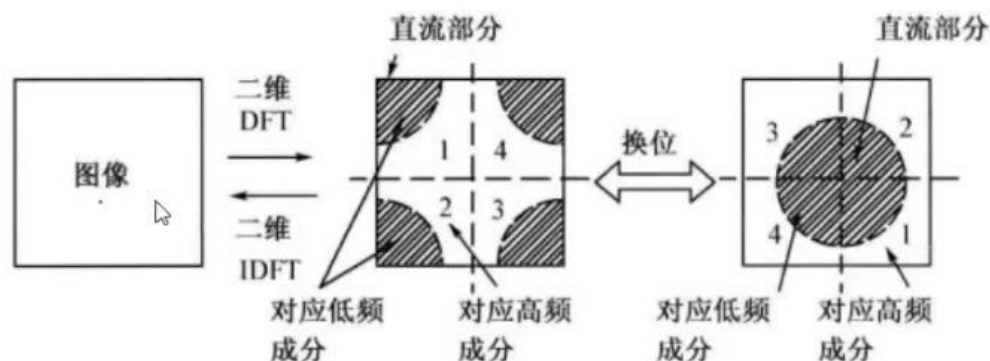
该论文的另一个创新点：一些像素级损失函数在对一些高频区域的复原效果不好，会出现失真，过于平滑的效果，因为他们更关注全局的信息。而转为频域后就对这些高频信息提高权重，低频信息减小权重，让其更关注高频信息的复原。

二维离散傅里叶变换(DFT)

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)}$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)}$$

$$e^{j\theta} = \cos(\theta) + j \sin(\theta)$$



可知变换之后的图像shape不变，但元素变为复数。

再进行中心化，最后元素取模即可生成可视化的频谱图

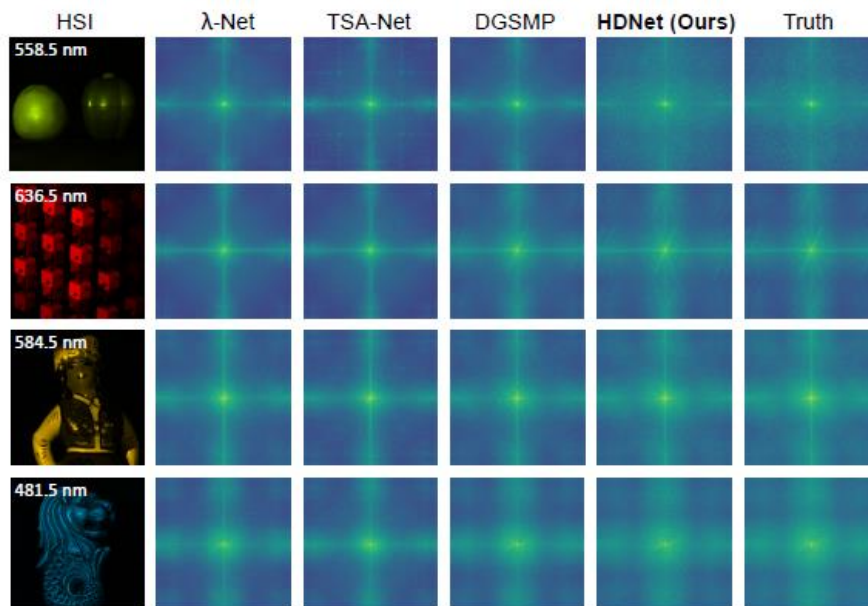


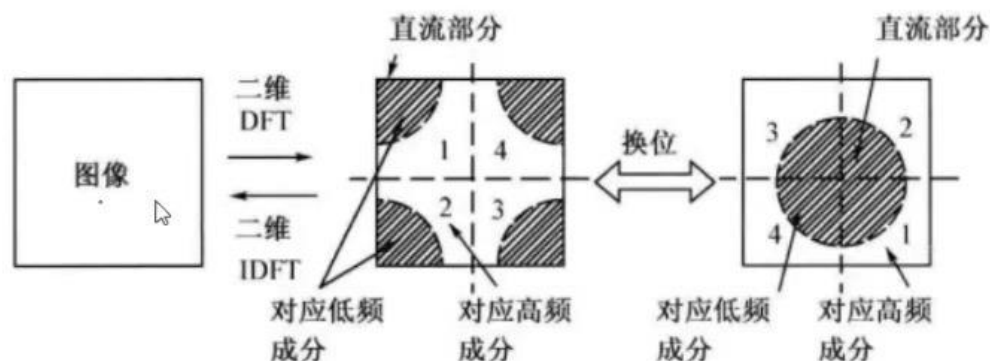
Figure 1. Frequency spectra visualization of different methods.

二维离散傅里叶变换(DFT)

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)}$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)}$$

$$e^{j\theta} = \cos(\theta) + j \sin(\theta)$$



可知变换之后的图像shape不变，但元素变为复数。

再进行中心化，最后元素取模即可生成可视化的频谱图

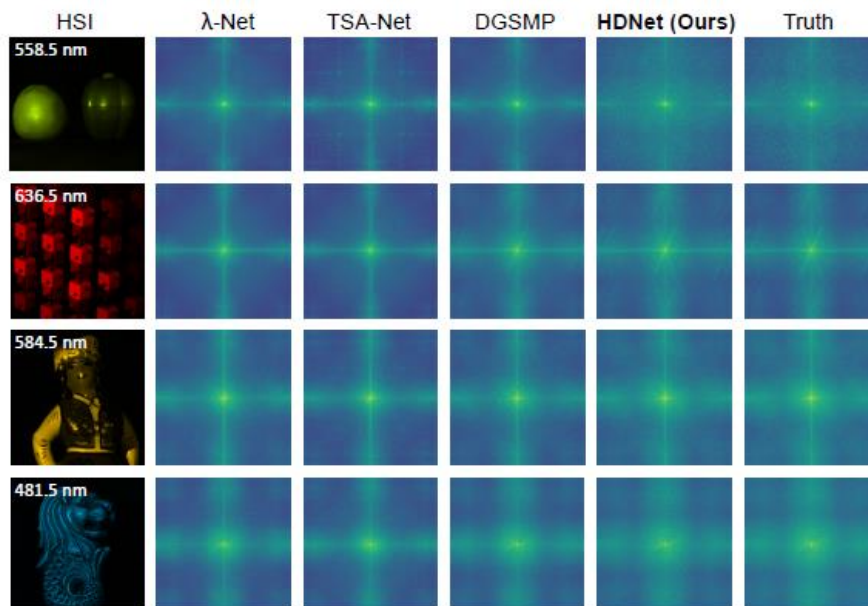


Figure 1. Frequency spectra visualization of different methods.

L_FDL

$$d^k(u, v) = \|\mathbf{F}_{gt}^k(u, v) - \mathbf{F}_{pred}^k(u, v)\|^\alpha.$$

由此建立了L_FDL损失函数。

$$d(\mathbf{F}_{gt}^k, \mathbf{F}_{pred}^k) = \frac{1}{HW} \sum_{u=0}^{H-1} \sum_{v=0}^{W-1} \theta^k(u, v) d^k(u, v),$$

$$L_{FDL}(\mathbf{F}_{gt}, \mathbf{F}_{pred}) = \sum_{k=0}^{C-1} d(\mathbf{F}_{gt}^k, \mathbf{F}_{pred}^k).$$

$$L_{total} = L_{SDL}(y_{gt}, y_{pred}) + \lambda L_{FDL}(\mathbf{F}_{gt}, \mathbf{F}_{pred}).$$

实验

Method	Params.	CR	SR	Complexity	PSNR/SSIM
λ -Net [24]	62.64M	1	$[H,W]$	$C^2WH + CW^2H^2$	30.85 / 0.9062
TSA-Net [23]	44.25M	$C/4$	$[H,W]$	CWH	32.68 / 0.9267
HDNet (Ours)	2.37M	$C/2$	$[H,W]$	CWH	34.34 / 0.9572

λ	0	0.1	0.3	0.5	0.7	0.9	1
SSIM \uparrow	0.9093	0.9369	0.9498	0.9538	0.9572	0.9425	0.9399
PSNR \uparrow	31.91	33.27	33.86	34.05	34.34	33.75	33.52

Table 3. Performance comparisons of different loss weight factor.

Metric	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$
LFD \downarrow	14.8633	14.3792	13.9825	13.6571	13.3238	15.0863
SSIM \uparrow	0.9397	0.9428	0.9543	0.9569	0.9572	0.9065
PSNR \uparrow	33.16	33.51	34.14	34.40	34.34	31.89

Table 4. Model performance comparison using different coefficients to calculate the spectrum distances in frequency domain.

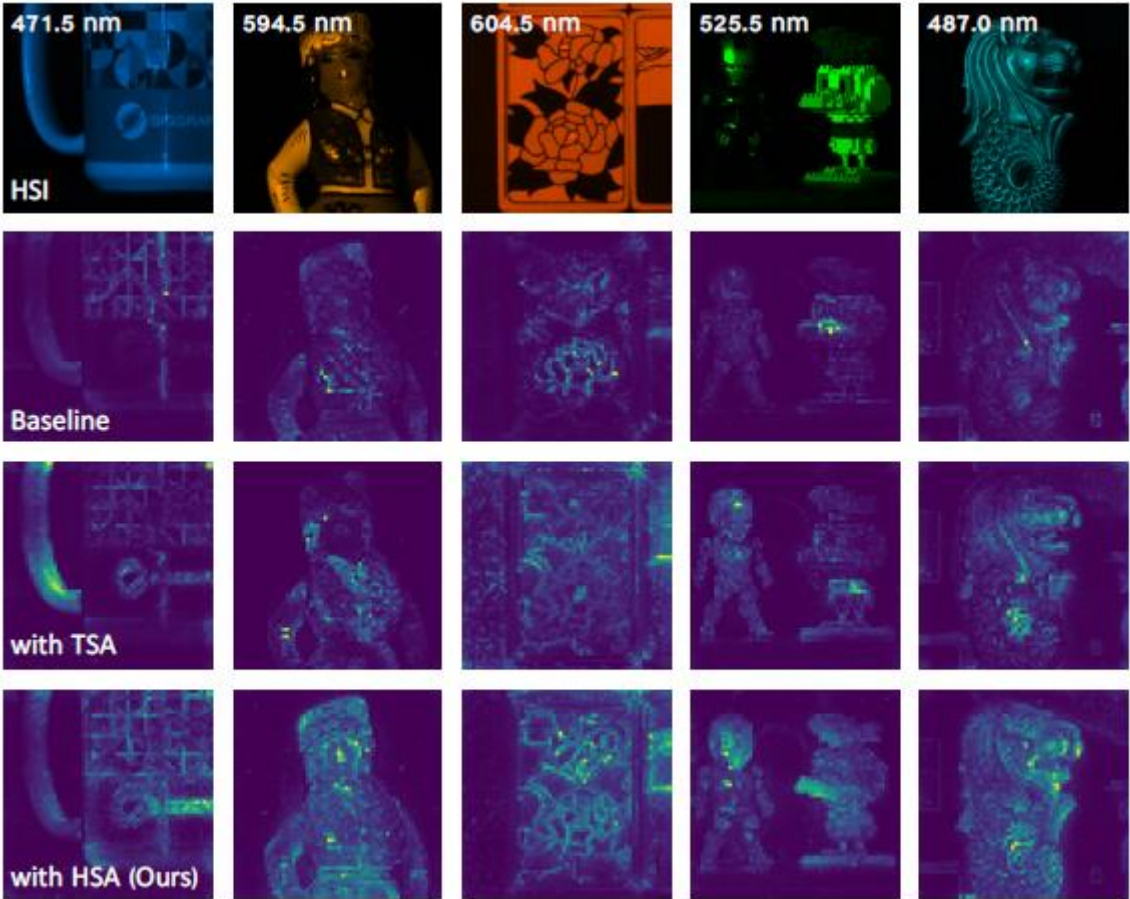


Figure 4. Feature visualization with different attention modules.