Supporting Information

Acceleration of high-quality Raman imaging via locality enhanced Transformer network

Shizhuang Weng[a,b,*], Rui Zhu[a,b], Yehang Wu[a,b], Cong Wang[a,b], Pan Li[c], Ling Zheng[a,b], Dong Liang[a,b,*]

[a] National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University, Hefei 230601, China.
[b] School of Electronics and Information Engineering, Anhui University, Anhui, Hefei 230601, China.
[c] Institute of Health and Medical Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China.

Contents	**Description**
Figure S1	Qualitative comparison of LETNet and other methods of example 1
Figure S2	Qualitative comparison of LETNet and other methods of example 2
Figure S3	Qualitative comparison of LETNet and other methods of example 3
Figure S4	Grayscale image of ELAN and LETNet
Figure S5	The training and validation loss curves
Figure S6	Visual Comparison of test example in dataset of three-channel Raman images
Figure S7	The differential images of HR and SR images
Table S1	SR performance of different kernel size on hyperspectral Raman image of breast cancer cells
Table S2	SR performance of different kernel size on three-channel Raman image of brain tumor tissues
Supplementary Reference	Reference of supplementary method

Supplementary Figures

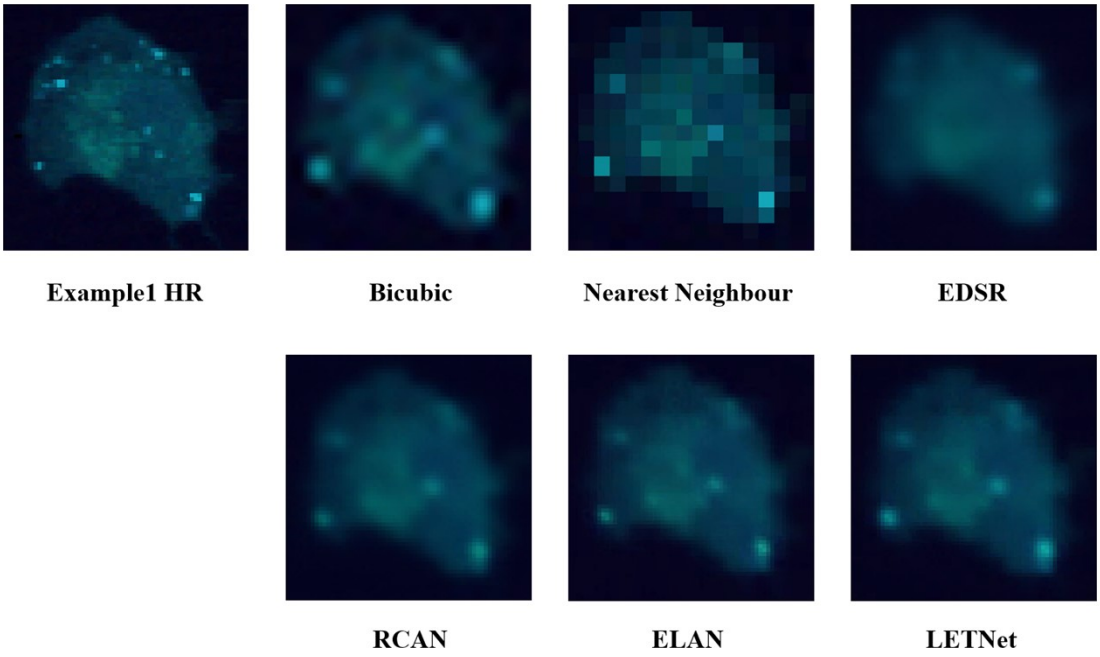

Figure S1: Qualitative comparison of LETNet and bicubic, nearest neighbor, EDSR, RCAN, ELAN on 4× Raman image SR of example 1. The result images are shown through pseudo-color images. The acquisition time of HR image and SR image are 68.25 minutes and 4.27 minutes respectively.

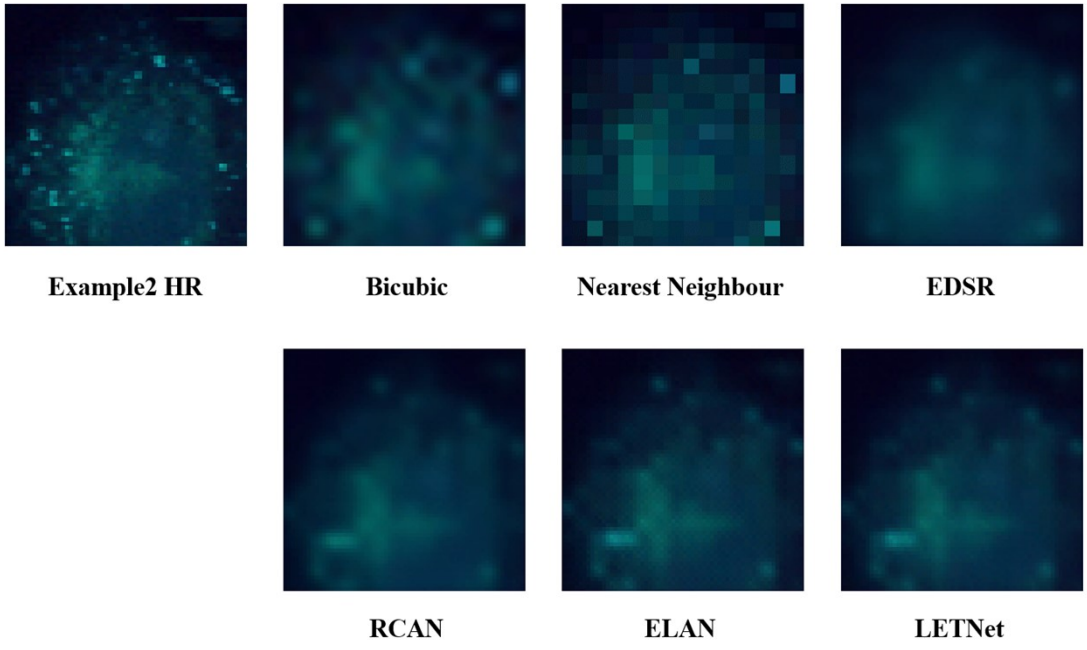

Figure S2: Qualitative comparison of LETNet and bicubic, nearest neighbor, EDSR, RCAN, ELAN on 4× Raman image SR of example 2. The result images are shown through pseudo-color images. The acquisition time of HR image and SR image are 68.25 minutes and 4.27 minutes respectively.

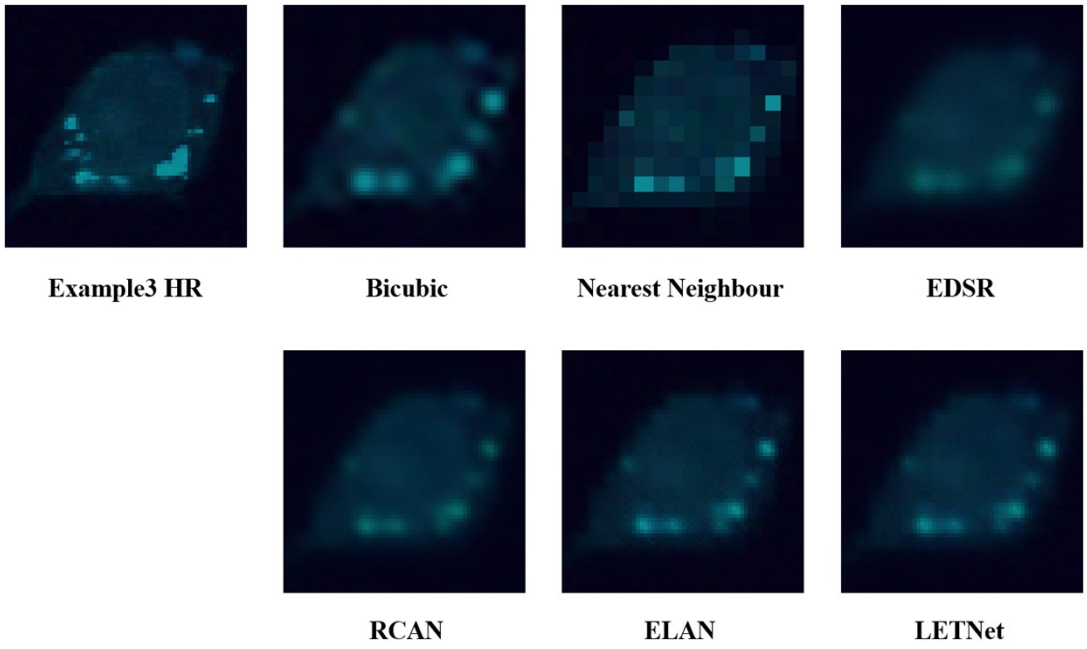

Figure S3: Qualitative comparison of LETNet and bicubic, nearest neighbor, EDSR, RCAN, ELAN on 4× Raman image SR of example 3. The result images are shown through pseudo-color images. The acquisition time of HR image and SR image are 68.25 minutes and 4.27 minutes respectively.

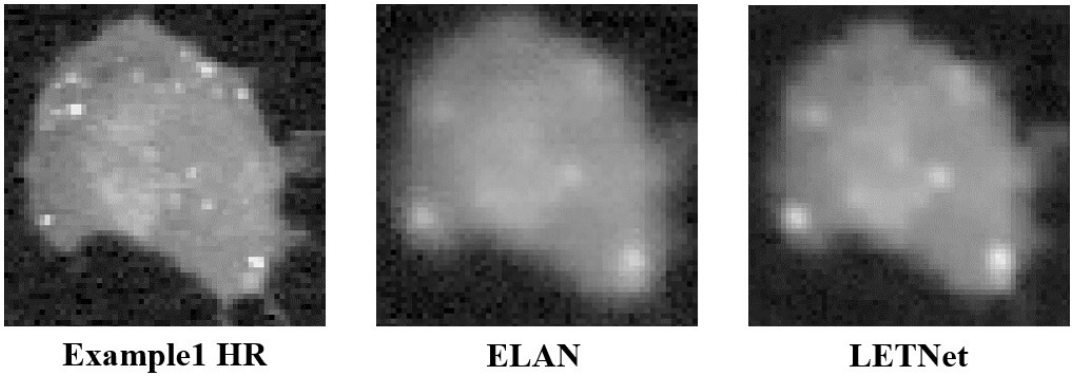

Figure S4: Grayscale image of HR image and SR images of ELAN and LETNet. Raman image at 1440 cm^{-1} (represents lipids of breast cancer cells) was chosen to show difference.

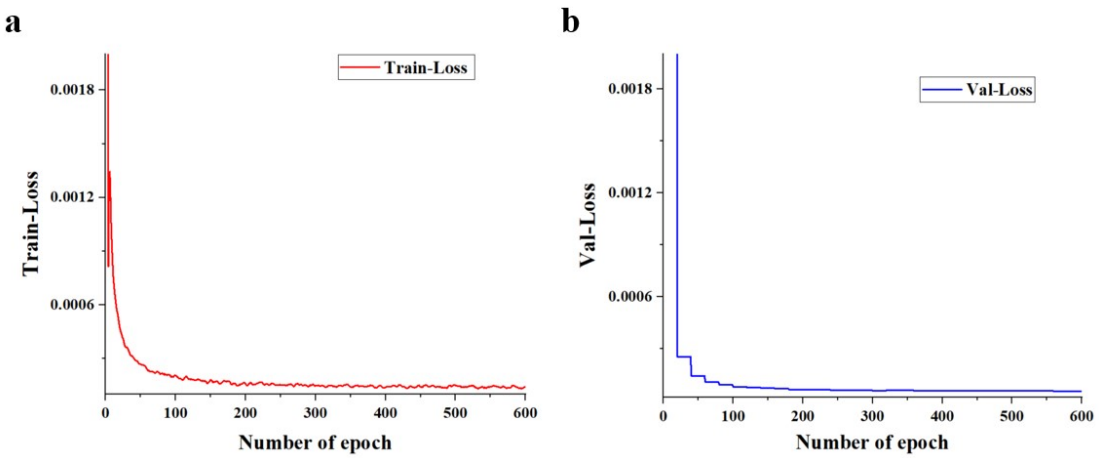

Figure S5: The training loss and validation curves before 600 epochs. (a) The red curve is the loss-epoch curve of training set. (b) The blue curve is the loss-epoch curve of validation set.

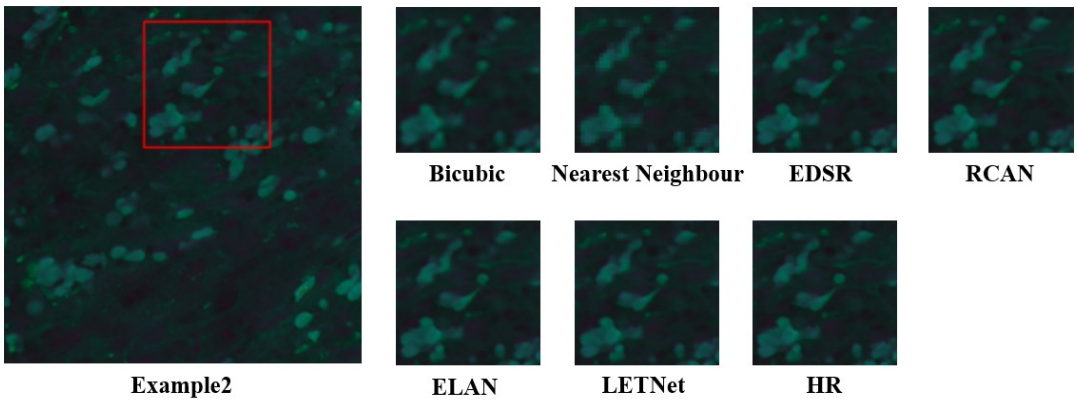

Figure S6: Comparison result of another test example. The first column is corresponding whole HR image, and other columns contain the HR image and SR images within the red rectangle.

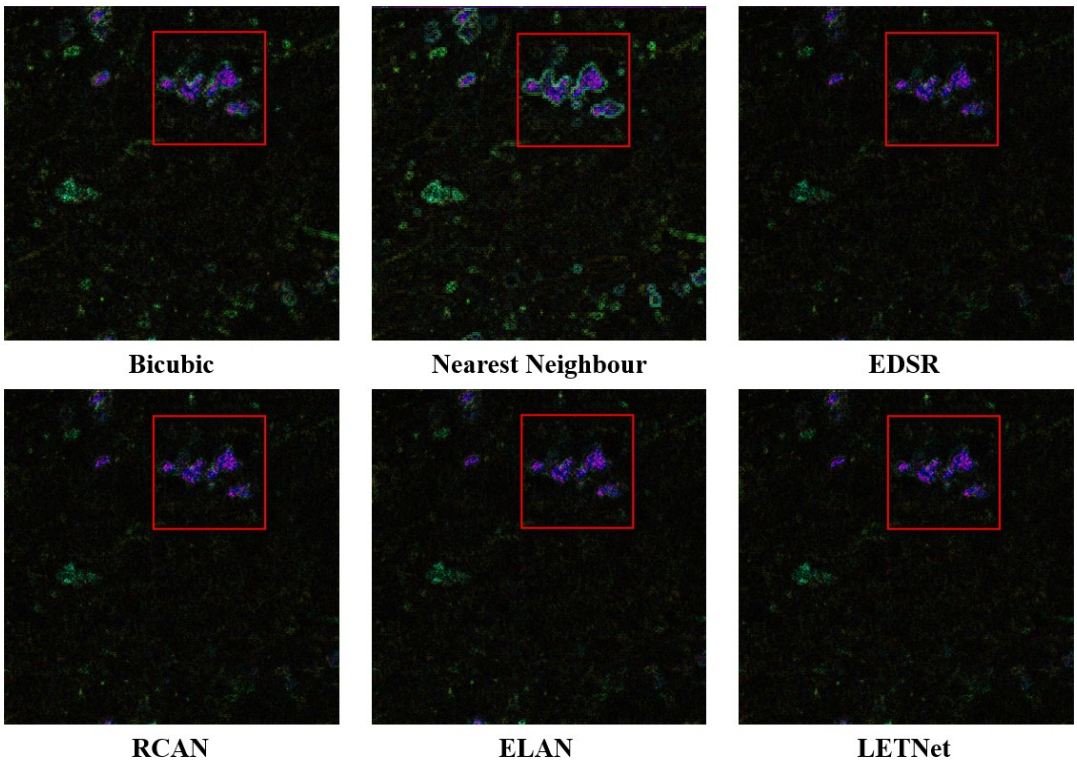

Figure S7: The whole differential images of HR image and SR images of different methods. The brighter pixels in the image represent more errors in SR images.

Supplementary Tables

The performance of LETNet with depth-wise convolution of different kernel size on 2×, 4× and 8×hyperspectral Raman image SR was presented in Table S1. The using of kernel size in this experiment comprehensively considers the SR performance of the model, including PSNR, SSIM and Params. The best results are shown in bold.

Table S1: SR performance of depth-wise convolution with different kernel size on hyperspectral Raman image of breast cancer cells.

Methods	Scale	PSNR	SSIM	Params(K)
LETNet-3×3	2×	44.46	0.9624	**17906**
LETNet-7×7		44.34	0.9623	18146
LETNet-11×11		**44.74**	**0.9666**	18578
LETNet-13×13		44.54	0.9637	18866
LETNet-3×3	4×	42.01	0.9416	**44912**
LETNet-7×7		**42.14**	0.9436	45152
LETNet-11×11		42.05	**0.9448**	45584
LETNet-13×13		41.96	0.9436	45872
LETNet-3×3	8×	40.27	**0.9292**	**152936**
LETNet-7×7		**40.36**	0.9291	153176
LETNet-11×11		40.08	0.9267	153608
LETNet-13×13		39.93	0.9233	153896

The performance of LETNet with depth-wise convolution of different kernel size on 2- 8× three-channel Raman image SR. The highest PSNR and SSIM are achieved when the kernel size is 11, demonstrated that large DW convolution is efficient for SR of this dataset [1]. The best results are shown in bold.

Table S2: SR performance of depth-wise convolution with different kernel size on three-channel Raman image of brain tumor tissues.

Methods	Scale	PSNR	SSIM	Params(K)
LETNet-3×3	2×	47.04	0.9852	**403**
LETNet-7×7		47.24	0.9854	461
LETNet-11×11		**47.26**	**0.9855**	565
LETNet-13×13		47.19	0.9854	634
LETNet-3×3	4×	40.12	0.9449	**404**
LETNet-7×7		40.23	0.9462	462
LETNet-11×11		**40.27**	**0.9464**	566
LETNet-13×13		40.25	0.9463	635
LETNet-3×3	8×	36.06	0.896	**408**
LETNet-7×7		36.10	0.8965	466
LETNet-11×11		**36.13**	**0.8971**	570
LETNet-13×13		36.09	0.8964	639

Supplementary Reference

[1] Ding, X., Zhang, X., Zhou, Y., Han, J., Ding, G., & Sun, J. (2022). Scaling Up Your Kernels to 31×31: Revisiting Large Kernel Design in CNNs. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 11953-11965.