

Learning to Super-Resolve Blurry Images With Events(IEEE 2023)

一、简介

单运动模糊图像的超分辨率问题是由于运动模糊和低空间分辨率的共同退化而导致的严重不适定问题，于是提出利用事件减轻 SRB 的负担，提高模糊图像的空间分辨率，改善图像清晰度。

超分辨是一种图像处理技术，通过从低分辨率图像中生成高分辨率图像来提高图像的空间分辨率。

二、本文工作

1. 提出一种事件增强 SRB(E-SRB)算法，可以从单个 LR 模糊图像中生成一系列清晰的 HR 图像。为了实现，制定了事件增强退化模型，同时考虑低空间分辨率、运动模糊和事件噪声。
2. 构建了一个事件增强稀疏学习网络(eSL-Net++)，事件和强度框架都用稀疏表示建模。
3. 提出了一种事件洗刷合并方案，将单帧 SRB 扩展到序列帧 SRB，无需任何额外的训练过程。

稀疏学习指处理高维数据，其中大多数特征的值为零。稀疏模型的引入有助于减少模型复杂度，提高模型可解释性，提高训练效率。

三、具体方法

3.1 E-SRB

利用事件从模糊的输入中解决 HR 清晰图像，只输入一张模糊的 LR 图像。为此制定了事件增强退化模型如下：

$$Y = \bar{E}(f) \circ I(f) + \varepsilon_Y \quad (1)$$

$$E(f) = \bar{E}(f) + \varepsilon_E \quad (2)$$

$$I(f) = PX(f) \quad (3)$$

Y 为曝光时间间隔内捕获的图像帧， $\varepsilon_Y \sim \mathcal{N}(0, \sigma_Y^2)$ 为标准差为 σ_Y 的高斯噪声。 $E(f)$ 表示强度变化方向的极性，且假设具有加性高斯噪声 $\varepsilon_E \sim \mathcal{N}(0, \sigma^2)$ 。 $\bar{E}(f)$ 表示 $E(f)$ 的估计量， $I(f)$ 为时间 f 的隐像， $X(f)$ 为 HR 图像。

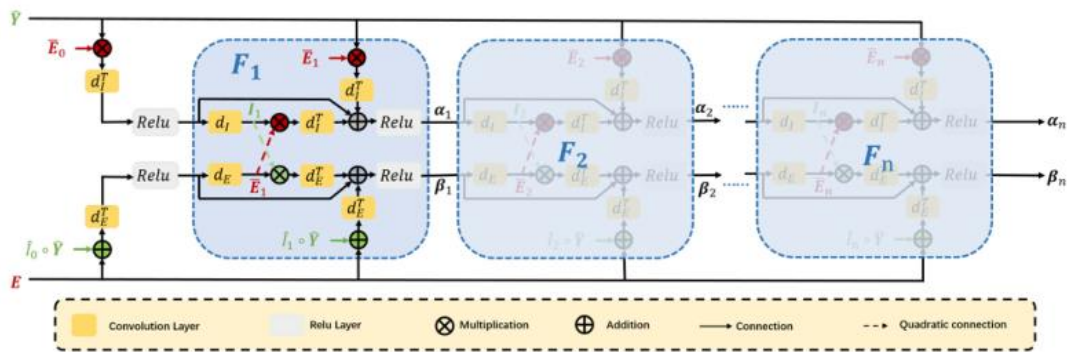
该模型目标即求出 $X(f)$ ，具体来说，给定观察到的图像 Y 和相应的触发事件 $\mathcal{E}_{\mathcal{T}} \triangleq \{e(t), t \in \mathcal{T}\}$ ，重建指定时间 $f \in \mathcal{T}$ 高质量强度图像 X ，即：

$$X(f) = \text{E-SRB}(Y, \mathcal{E}_{\mathcal{T}}, f) \quad (4)$$

3.2 基于事件 SRB 稀疏学习

先通过 DSL 方案对无噪声事件和潜在清晰强度施加稀疏性来制定 E-SRB 问题；然后通过展开稀疏恢复算法的迭代阶段，并使用合成数据集训练 eSL-Net++，构建相应的 eSL-Net++ 网络；最后通过事件洗刷将其从单个 E-SRB 扩展到序列 E-SRB。

3.2.1 事件驱动的双稀疏学习及模块



图一 双稀疏学习(DSL)方案

假设 LR 锐利清晰图像 I ，HR 锐利清晰图像 X 和累计事件 \bar{E} 可以用卷积核或字典 d_I, d_X, d_E 稀疏表示，即：

$$\begin{aligned} I &= d_I * \alpha_I, \\ X &= d_X * \alpha_X, \\ \bar{E} &= d_E * \beta, \end{aligned} \quad (5)$$

假设 I 和 X 有相同的稀疏表示, 则给定模糊图像 Y 和相应事件 \mathcal{E}_T 则可求出稀疏系数 α 和 β 。

$$\min_{\alpha, \beta} \frac{1}{2} \|Y - I \circ \bar{E}\|_2^2 + \frac{\lambda_1}{2} \|E - \bar{E}\|_2^2 + \lambda_2 \|\alpha\|_1 + \lambda_3 \|\beta\|_1 \quad (6)$$

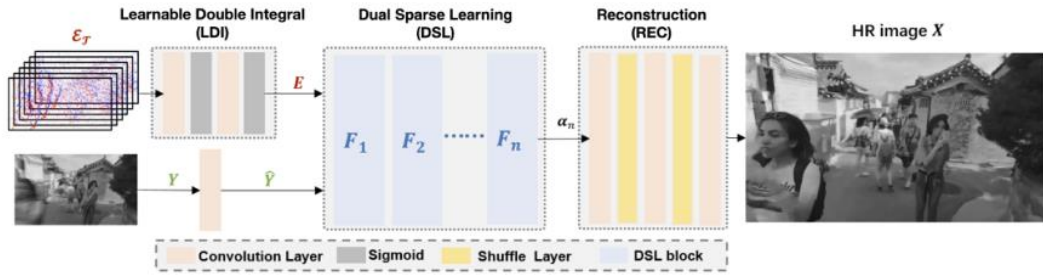
将公式(6)看作 I 和 E 的两个子问题, 每个子问题都是典型的 LASSO 问题, 之后再通过公式(7)求解:

$$\begin{aligned} \alpha^+ &= \Gamma_{\eta\lambda_2}(\alpha - \eta d_I^T * \bar{E} \circ (\bar{E} \circ (d_I * \alpha) - Y)), \\ \beta^+ &= \Gamma_{\eta\lambda_3}(\beta - \eta d_E^T * I \circ (I \circ (d_E * \beta) - Y) \\ &\quad - \eta\lambda_1 d_E^T * (d_E * \beta - E)), \end{aligned} \quad (7)$$

其中, $\Gamma_\theta(l) = \text{sign}(l)\max(|l| - \theta, 0)$ 表示逐元软阈值函数, 在得到稀疏系数 α 的最优解之后, 根据公式(5)恢复 LR 锐利清晰图像 I 和 HR 锐利清晰图像 X 。

公式(7)为 E-SRB 提供了 DSL 方案, 式中统一了噪声抑制和运动去模糊。其中 DSL 方案如图一。

在该模块中, 将图像帧 Y 线性变换到特征域中即 \hat{Y} , 以及 E 输入到 DSL 模块计算稀疏系数 α 及 β 。其中 ReLU 用于实现 Γ_θ 。



图二 eSL-Net++

提出了一个可学习的二重积分(LDI)网络, 采用了两个卷积层和两个 S 形层, 如图二。即:

$$E = \text{LDI}(\mathcal{E}_T) \quad (8)$$

最后通过重建模块 REC 输出图像, 即:

$$E = \text{LDI}(\mathcal{E}_T) \quad (9)$$

E-SRB 总体结构如图二, 总的来说, 就是将 LR 模糊图像帧和曝光时间间隔 T 期间的相应事件输入到 eSL-Net++, 输出 HR 锐利清晰强度图像 X , 即:

$$X, \bar{E} = \text{eSL-Net++}(Y, \mathcal{E}_T) \quad (10)$$

3.2.2 事件洗刷

提出了一种将单帧求解器扩展到序列帧求解器且不需要额外的训练过程的方法。

该方法陈述:

使用特定时间 $f \in \mathcal{T}$ 的地面真实 HR 锐利图像训练 eSL-Net++, 可以得到时间 f 的单帧 E-SRB 任务求解器。我们可以对其进行不同时间 f 的训练从而得到 E-SRB 的序列求解器, 但是过程是繁琐的。于是假设已经训练了 0 时刻 eSL-Net++, 参考原文式(18)。然后将 T 分为两个时间段 $[0, f]$ 及 $[f, T]$, 相应的将事件分为两个

同样的子区间，结合原文式(19)即得式(20)(21)，之后对收集到的时间进行洗牌，则可以在任意时间得到相应的 HR 锐重建 $X(f)$ 。

事件洗牌是指对事件顺序进行重新排列，以帮助模型具有更好的泛化能力。在这里我认为对单帧求解器进行事件洗牌是因为引入了时间维度，将单帧数据通过重新排列的方式组合成序列帧，从而推广至序列帧求解器。

四、实验

4.1 数据集

- 1.合成事件数据集：构建了包含 HR 清晰图像、LR 模糊图像和 Event 流的合成 GoPro 数据集。
- 2.真实世界事件的数据集：来自 HQF 及作者自己构建的 RWS。其中 HQF 数据集包含真实世界的事件和清晰的 LR 真实帧。RWS 只包含真实世界事件和真实世界模糊的 APS 帧，只在该数据集上定性评估 SRB 方法。

4.2 结果对比

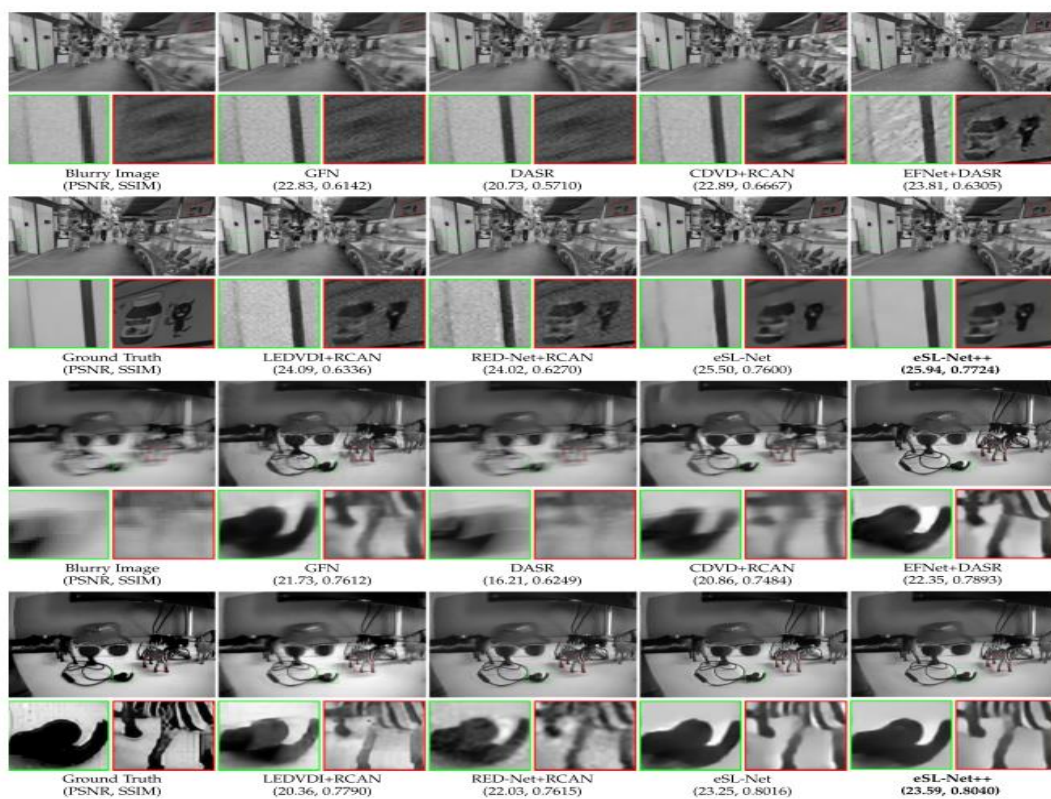
4.2.1 单帧 SRB 结果

将本文方法与最先进的可以实现 SRB 的方法在 GoPro 及 HQF 数据集上进行了定量评估。结果如表一，验证了模型从合成场景推广到真实场景的能力。

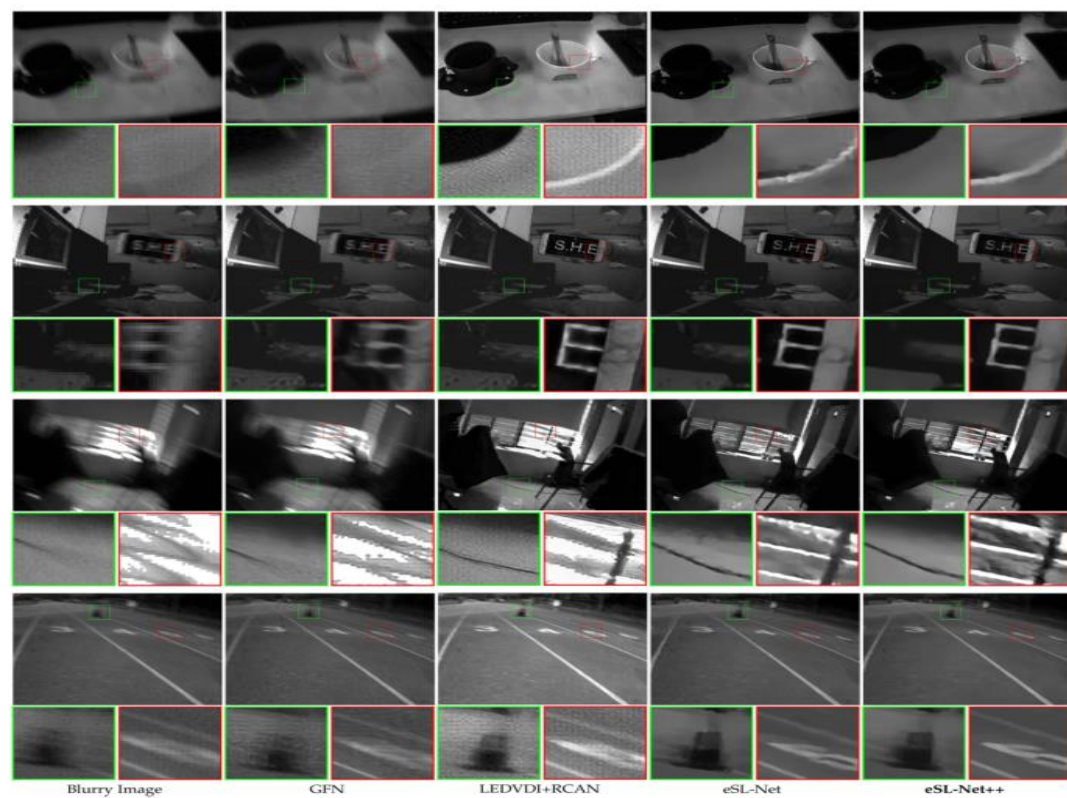
Method	Inputs		Single Frame Reconstruction				Sequence Reconstruction				Params
			Synthetic events		Real events		Synthetic events		Real events		
	Events	Image	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
GFN	✗	✓	22.43	0.5295	21.10	0.7507	/	/	/	/	12.2M
DASR	✗	✓	20.87	0.5124	17.65	0.7128	/	/	/	/	5.97M
FKP	✗	✓	19.40	0.4717	17.25	0.6646	/	/	/	/	0.59M
MANet	✗	✓	20.01	0.4993	17.48	0.6786	/	/	/	/	9.89M
SRN+RCAN _{4×}	✗	✓	22.86	0.5601	20.91	0.7562	/	/	/	/	34.9M
CDVD+RCAN _{4×}	✗	✓	22.20	0.5967	20.95	0.7577	/	/	/	/	47.3M
E2VID+RCAN _{4×}	✓	✗	12.15	0.3385	9.88	0.5392	12.14	0.3369	9.86	0.5382	41.8M
E2SRI+RCAN _{2×}	✓	✗	11.90	0.4183	11.28	0.5832	11.85	0.4162	11.34	0.5853	40.9M
EDI+RCAN _{4×}	✓	✓	22.48	0.6196	21.34	0.7689	21.77	0.6021	19.79	0.7473	31.1M
LEDVDI+RCAN _{4×}	✓	✓	23.75	0.5676	21.81	0.7942	23.30	0.5670	21.34	0.7914	36.1M
RED-Net+RCAN _{4×}	✓	✓	23.54	0.5608	22.78	0.7644	22.82	0.5443	21.64	0.7475	25.3M
EFNet+DASR _{4×}	✓	✓	23.85	0.5556	22.73	0.7996	/	/	/	/	14.4M
LEDVDI+RealBasicVSR	✓	✓	22.76	0.6286	18.11	0.6995	22.23	0.6150	16.97	0.6841	26.9M
RED-Net+RealBasicVSR	✓	✓	23.41	0.5999	19.80	0.7273	20.49	0.5262	21.19	0.7450	16.0M
eSL-Net (Ours)	✓	✓	<u>25.32</u>	<u>0.6705</u>	<u>23.91</u>	<u>0.8075</u>	<u>23.80</u>	<u>0.6455</u>	<u>22.16</u>	<u>0.7790</u>	<u>1.32M</u>
eSL-Net++ (Ours)	✓	✓	25.73	0.6824	23.99	0.8087	24.69	0.6602	22.99	<u>0.7913</u>	1.41M

Evaluations are conducted for single frame and sequence reconstructions, respectively.

进一步在 GoPro、HQF 和 RWS 数据集上对性能进行了定性评估。如图三图四。



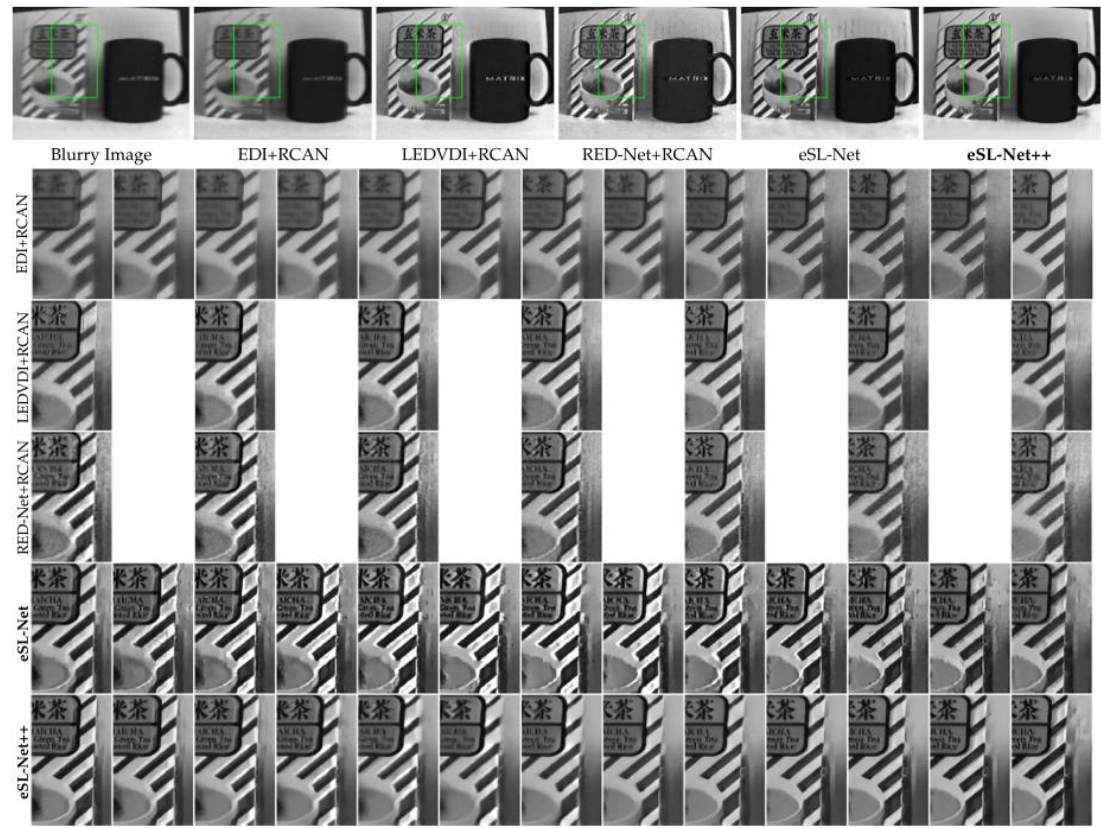
图三 合成运动模糊的 GoPro(前两行)和 HQF(后两行)数据集上的定性结果



图四 RWS 数据集的模糊图像与真实世界的运动模糊和事件的 SR 结果

4.2.2 序列帧 SRB 结果

定量结果如表一。为了对序列帧 SRB 进行定性比较，在 RWS 数据集的一张模糊图像上展示了四种方法的序列帧重建结果。如图五。

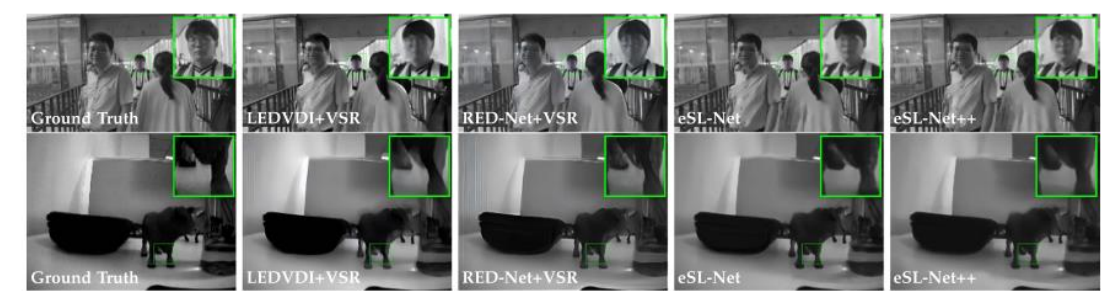


图五 RWS 数据集上定性结果

得益于双稀疏学习模块和严格的事件洗牌合并模块，eSL-Net++可以减轻事件噪声和晕波伪影的影响，从而在序列重建中获得最佳的可视化性能。

4.2.3 视频 SR 的比较

将我们的 eSL-Net 和 eSL-Net++与 Deblur-thenVSR 方法进行了比较，定量结果如表一，定性结果如图六。



图六 在 GoPro(上)和 HQF(下)数据集上与 Deblur-then-VSR 方法进行定性比较

五、总结

在本文中，为 E-SRB 提出了一种名为 eSL-Net++ 的新型网络。从制定事件增强的 d 生成模型(EDM)开始，同时考虑到由噪声、运动模糊和下采样引起的图像退化。通过假设潜在图像和事件的稀疏性，提出了一种双稀疏学习方案(Dual Sparse Learning scheme, DSL)，然后通过展开 DSL 迭代构建了一个深度神经网络，即 eSL-Net++。与之前的版本(即 eSL-Net)相比，eSL-Net++ 进一步考虑了事件噪声，并通过严格的事件 shuffle-and-merge 方案扩展到序列帧 SRB，从而获得了优于 eSL-Net 的性能。在合成数据和实际数据上的大量实验证明了我们的 eSL-Net++ 的有效性和优越性。