Uncertainty Learning in Kernel Estimation for Multi-Stage Blind Image Super-Resolution

Zhenxuan Fang[1] Weisheng Dong[1*] Xin Li[2] Jinjian Wu[1] Jinjian Wu[1] Guangming Shi[1]
[1]School of Artificial Intelligence, Xidian University [2]West Virginia University

Motivation

Existing blind SR approaches suffer from two fundamental weaknesses:

- **The lack of robustness**: the performance drops when the estimated degradation is inaccurate.
- **The lack of transparency**: network architectures are heuristic without incorporating domain knowledge.

Contributions

- The blind SR problem is formulated as a joint Maximum a Posteriori (MAP) approach.
- A novel multi-stage SR network is proposed by converting the MAP estimator with a learned LSM prior and estimated kernel into a multi-stage deep network.
- To improve the performance and robustness of kernel estimation, we introduce uncertainty learning to the kernel estimation network.
- A new State-of-the-art method for blind SR.

Uncertainty Learning in Kernel Estimation

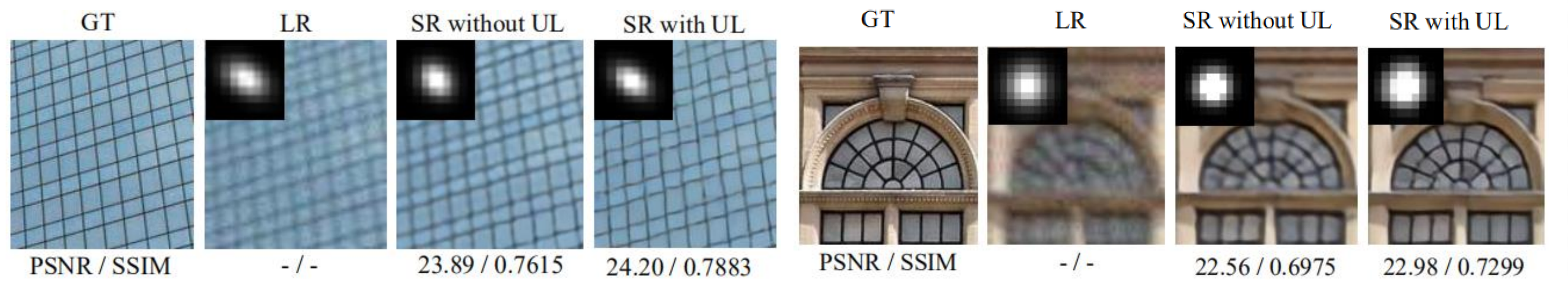
- Uncertainty learning is introduced to the process of blur kernel estimation:

$$p(H \mid y) \sim N(k \mid \mu(y), \sigma^2(y))$$

Both the feature (mean) and uncertainty (variance) of the blur kernel are learned by DCNN.

- An equivalent sampling representation z is generated through re-parameterization method:

$$z = \mu + \varepsilon\sigma, \quad \varepsilon \sim N(0, I)$$



LSM Model for Blind SR

- The blind SR problem can be formulated as a joint Maximum a Posteriori (MAP) problem:

$$\log p(H, x \mid y) \propto \log p(H \mid y) + \log p(y \mid H, x) + \log p(x)$$

- The likelihood term:

$$p(y \mid H, x) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{\| y - DHx \|_2^2}{2\sigma_n^2}\right)$$

- The LSM model is exploited to model $\boldsymbol{x}$:

$$p(x_i \mid \theta_i) = \frac{1}{2\theta_i} \exp\left(-\frac{|x_i - u_i|}{\theta_i}\right)$$

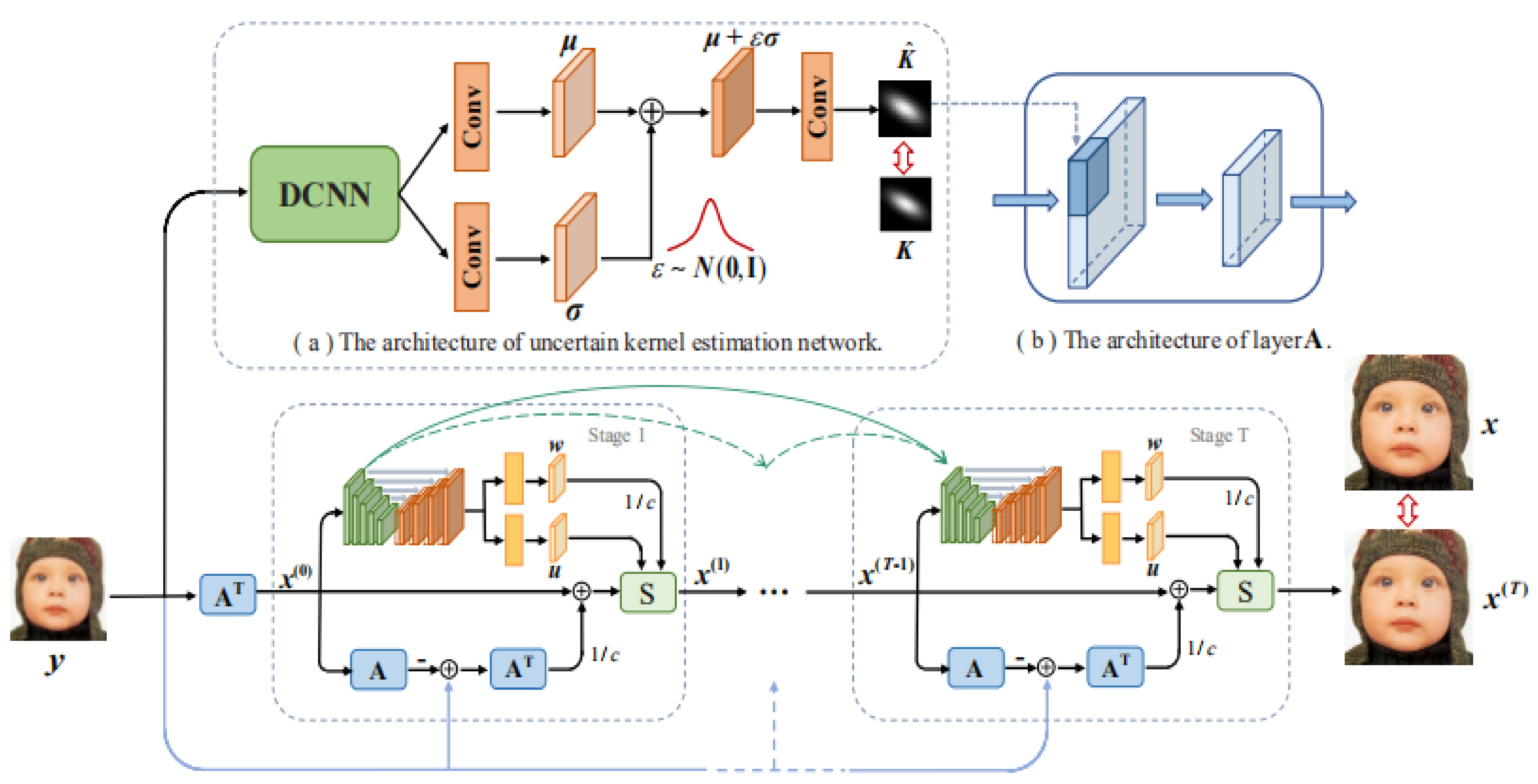
- Jointly estimate $\boldsymbol{x}$ and $\boldsymbol{\theta}$:

$$(x^*, \theta^*) = \arg\max_{x,\theta} \log p(y \mid H, x) + \log p(x \mid \theta) + \log p(\theta)$$

- Optimized by a unified framework

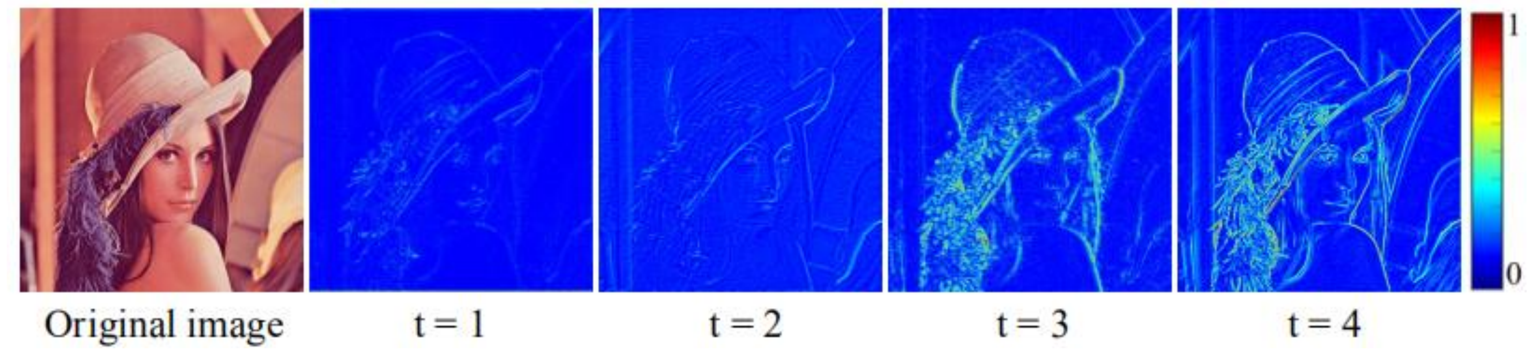
$$x^{(t+1)} = S_{G_w(x^{(t)})/c, G_u(x^{(t)})}\left(x^{(t)} + \frac{1}{c} A^T (y - Ax^{(t)})\right)$$

Multi-Stage Blind SR Network



(a) The architecture of uncertain kernel estimation network. (b) The architecture of layer **A**.

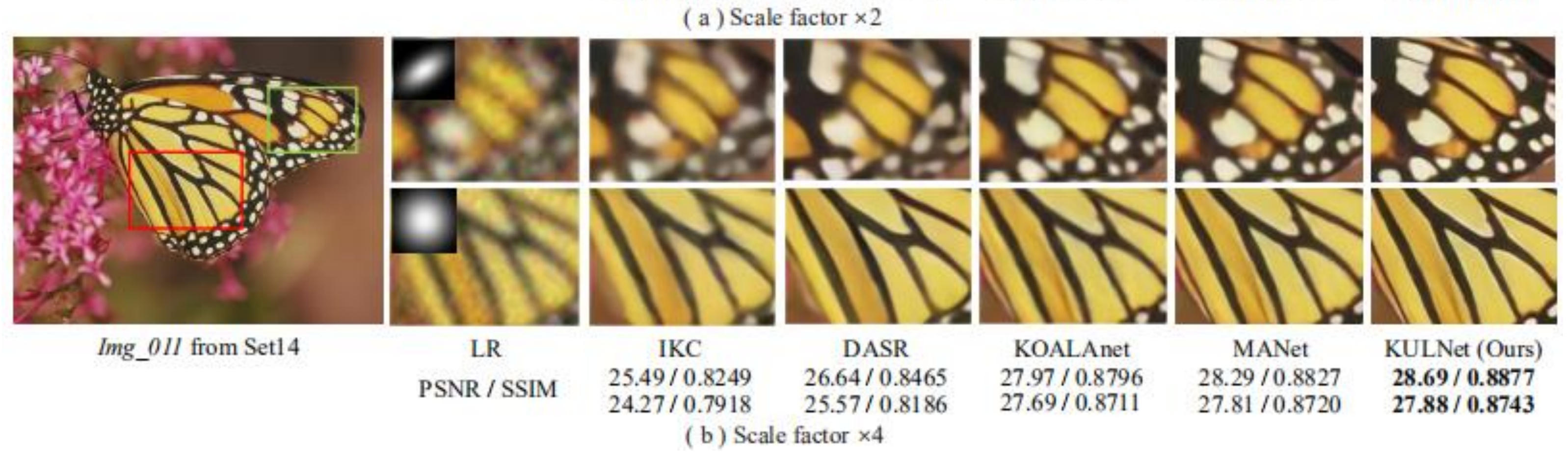
Visualization of the Learned Variance



Experimental Results

Method	Scale	Set5			Set14			BSD100			Urban100		
Noise		0	10	20	0	10	20	0	10	20	0	10	20
KernelGAN [4]+ZSSR [42]	×2	26.94	-	-	23.96	-	-	23.17	-	-	21.69	-	-
IKC [15]	×2	27.89	27.62	26.86	26.29	25.90	25.26	26.03	25.67	25.12	23.84	23.35	22.82
DASR [46]	×2	29.89	28.13	27.18	27.25	26.06	25.41	26.97	25.78	25.21	24.58	23.54	22.98
KOALAnet [27]	×2	33.96	30.59	29.05	30.53	27.98	26.87	29.77	27.23	26.28	27.56	25.59	24.52
MANet [32]	×2	33.99	30.77	29.28	30.61	28.22	27.11	29.85	27.48	26.49	27.64	25.71	24.76
KULNet (Ours)	×2	**34.36**	**31.03**	**29.54**	**30.93**	**28.42**	**27.27**	**30.43**	**27.69**	**26.62**	**28.16**	**26.10**	**25.08**
IKC [15]	×3	28.40	27.01	26.00	26.42	25.32	24.54	26.20	25.20	24.56	23.59	23.24	22.35
DASR [46]	×3	29.40	27.54	26.43	26.92	25.68	24.89	26.65	25.42	24.76	24.23	23.45	22.57
MANet [32]	×3	31.78	29.65	28.10	28.50	27.22	26.18	27.79	26.64	25.74	25.42	24.62	23.85
KULNet (Ours)	×3	**32.27**	**29.91**	**28.34**	**28.92**	**27.53**	**26.45**	**28.17**	**26.83**	**25.88**	**26.26**	**25.19**	**24.31**
KernelGAN [4]+ZSSR [42]	×4	23.85	-	-	22.55	-	-	21.37	-	-	19.12	-	-
IKC [15]	×4	27.91	26.51	25.52	26.06	24.92	24.19	25.80	24.83	24.21	23.26	22.47	21.90
DASR [46]	×4	30.33	27.29	25.94	27.31	25.48	24.54	26.77	25.16	24.42	24.34	22.98	22.28
KOALAnet [27]	×4	30.36	28.56	27.13	27.35	26.19	25.33	26.72	25.73	24.95	24.37	23.76	23.01
MANet [32]	×4	30.38	28.73	27.31	27.41	26.46	25.50	26.78	25.96	25.16	24.49	23.91	23.23
KULNet (Ours)	×4	**30.79**	**29.07**	**27.56**	**27.81**	**26.76**	**25.74**	**27.02**	**26.12**	**25.27**	**25.07**	**24.36**	**23.62**

(a) Scale factor ×2



(b) Scale factor ×4