Dual-lens Reference Image Super-Resolution

Jing Zhu Wenbo Li Hongxia Jin
Samsung Research America AI Center

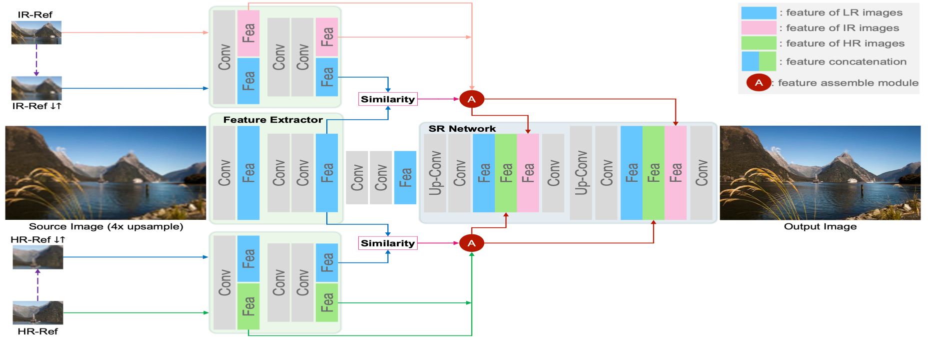
Problem

- Current high-end mobile devices are usually embedded with three lenses of different fields of view, so it is easy to capture three images at a single shot.
- Conventional reference-based image super-resolution (RefSR) methods can only work on one reference image, and most of them are time and storage consuming.

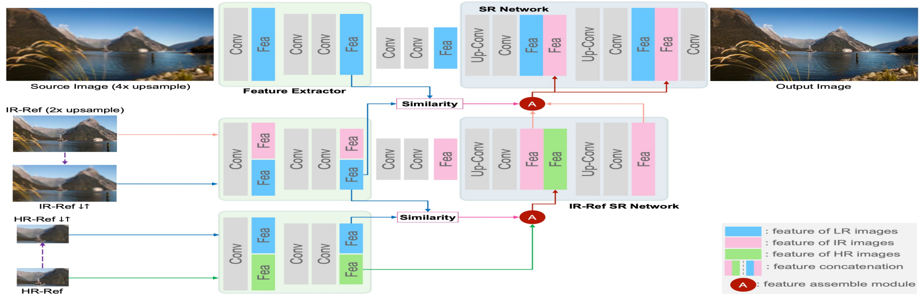
Contribution

- We explore the feasibility of improving the SR performance by taking advantages of the dual-lens reference images of different resolution levels, i.e., IR-Ref and HR-Ref.
- We newly construct seven datasets for evaluation, including a real dataset captured by multi-lens in a phone and six additional datasets that are simulated based on publicly available SR datasets.
- We proposed a base model that outperformed the state-of-the-art RefSR methods (e.g., TTSR and SRNTT) and an enhanced model to further boost the performance by considering the relationship between the dual-lens references.

Approach



- Base model. Shared feature extractors (green bubbles) extract features from the source image and the dual-lens references (HR-Ref and IR-Ref). IR-Ref↓↑ and HR-Ref↓↑ share the same resolution level with the source image, and are used for estimating the patch-wise similarities between the source image and the dual-lens references. Relevant features from IR-Ref and HR-Ref are embedded into the SR network based on the similarities.



- Enhanced model. The feature extractor is the same as the base model, while the feature assemble modules perform the assembling operations progressively and bottom-up. Such a progressive fashion integrates the advantages of the dual-lens references.

Experimental Results

Table 1: Comparisons of (4×) super-resolution with alternative approaches. The best performance has been **bolded** and the second best result is marked in blue. Higher score indicates better performance.

Method	Reference	Params (M)	Set14	BSDS100	Urban100	Manga109	Real104
Bicubic	×	–	26.10 / .702	25.96 / .667	23.15 / .657	24.92 / .789	23.53 / .590
SRResNet [18]	×	1.6	28.49 / .782	27.61 / .736	26.09 / .783	30.70 / .908	25.11 / .700
SRGAN [18]	×	1.6	26.60 / .718	25.74 / .666	24.50 / .736	27.79 / .856	22.65 / .609
LapSRN [17]	×	0.9	28.09 / .770	27.31 / .727	25.21 / .756	29.09 / .890	24.73 / .670
SRDenseNet [30]	×	2.0	28.50 / .778	27.53 / .733	26.05 / .781	29.49 / .899	25.04 / .680
EDSR [19]	×	43.1	28.81 / .787	27.72 / .742	26.64 / .803	31.03 / .915	25.25 / .697
RCAN [37]	×	15.6	28.85 / .788	27.74 / .743	26.74 / .806	31.19 / .917	25.30 / .700
SAN [3]	×	15.9	28.92 / .788	27.79 / .743	26.79 / .806	31.19 / .917	25.35 / .700
DRN [9]	×	4.8	28.93 / .790	27.78 / .744	26.84 / .807	31.52 / .919	25.37 / .701
SRNTT-*rec* [38]	IR-Ref	4.4	27.66 / .767	27.13 / .721	25.50 / .783	28.95 / .885	24.65 / .662
SRNTT [38]	IR-Ref	4.4	26.73 / .732	25.88 / .683	24.40 / .725	27.61 / .860	23.77 / .621
TTSR-*rec* [33]	IR-Ref	6.7	28.07 / .778	27.47 / .733	25.56 / .771	29.90 / .904	24.86 / .675
TTSR [33]	IR-Ref	6.7	26.98 / .750	26.42 / .704	24.76 / .753	28.09 / .879	23.61 / .641
MASA-*rec* [22]	IR-Ref	4.1	28.26 / .783	27.63 / .737	25.69 / .776	28.07 / .761	25.07 / .682
MASA [22]	IR-Ref	4.1	25.00 / .644	23.92 / .583	22.59 / .662	24.43 / .606	21.52 / .533
C^2-Matching-*rec* [13]	IR-Ref	8.9	26.82 / .780	26.28 / .739	24.83 / .792	28.27 / .901	23.77 / .683
C^2-Matching [13]	IR-Ref	8.9	26.31 / .758	25.80 / .713	24.29 / .769	27.51 / .883	23.42 / .661
SRNTT-*rec* [38]	HR-Ref	4.4	27.80 / .771	27.32 / .729	25.39 / .763	29.03 / .888	24.74 / .666
SRNTT [38]	HR-Ref	4.4	26.77 / .737	26.30 / .693	24.47 / .731	27.70 / .866	23.84 / .628
TTSR-*rec* [33]	HR-Ref	6.7	28.18 / .785	27.84 / .749	25.89 / .784	29.76 / .900	24.97 / .678
TTSR [33]	HR-Ref	6.7	27.08 / .754	26.68 / .718	24.87 / .761	28.68 / .888	23.71 / .647
MASA-*rec* [22]	HR-Ref	4.1	28.20 / .781	27.59 / .735	25.71 / .779	28.03 / .903	25.03 / .683
MASA [22]	HR-Ref	4.1	25.98 / .684	25.13 / .632	23.60 / .699	26.06 / .681	22.74 / .585
C^2-Matching-*rec* [13]	HR-Ref	8.9	26.49 / .789	26.59 / .751	25.16 / .807	28.58 / .911	24.10 / .702
C^2-Matching[13]	HR-Ref	8.9	26.01 / .762	26.13 / .727	24.62 / .785	27.82 / .891	23.72 / .679
Base Model	IR-Ref+HR-Ref	4.8	29.02 / .795	28.06 / .763	27.02 / .813	31.61 / .919	25.41 / .709
Enhanced Model	IR-Ref+HR-Ref	4.8	**29.46 / .810**	**28.75 / .775**	**27.58 / .824**	**31.76 / .931**	**25.93 / .718**

Table 3: The average inference time comparisons.

	SRNTT [38]	TTSR [33]	MASA [22]	C^2-Mat [13]	Base	Enhanced
Inference Time (s)	1.9563	0.5023	0.3298	0.5882	0.2288	0.3946

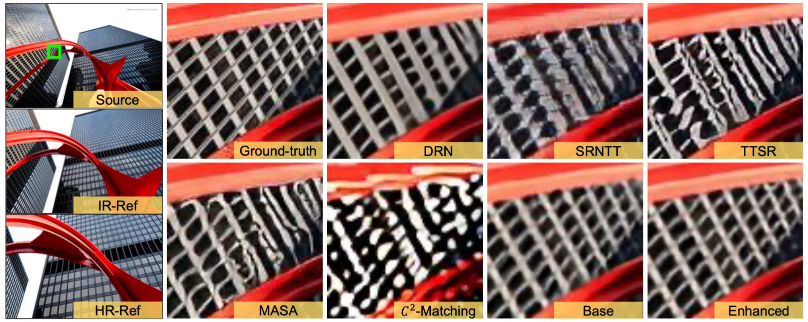


Figure 4: Qualitative comparison of the (4×) SR results. More results are put in the supplementary.