

英語論文#3

2023/07/31

M1

建元 了

論文の概要

- タイトル
 - 「**DARN: Distance Attention Residual Network for Lightweight Remote-Sensing Image Superresolution**」
- 執筆者
 - Qingjian Wang , SenWang , Mingfang Chen , and Yang Zhu
- 掲載
 - IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Volume: 16)
- 選択理由
 - リモートセンシングにおいて解像度に関わる改善方法の調査

背景

- リモートセンシング(RS)における **単一画像超解像(SISR)** の応用は非常に重要
 - 低解像度画像を高解像度画像に復元する手法
- **単一画像超解像(SISR)** は、最も代表的な低レベル視覚タスクとして集中的に研究されており、SISRアルゴリズムは高解像度の画像を得ることができるため、様々な分野で大きな貢献をしている

背景

- リモートセンシング画像の解像度が大きいという特性から、一般的なSISRアルゴリズムでは推論速度が極端に遅くなることが多い
- 速度、精度を向上した場合、モデルパラメータの割合が大きくなる

導入

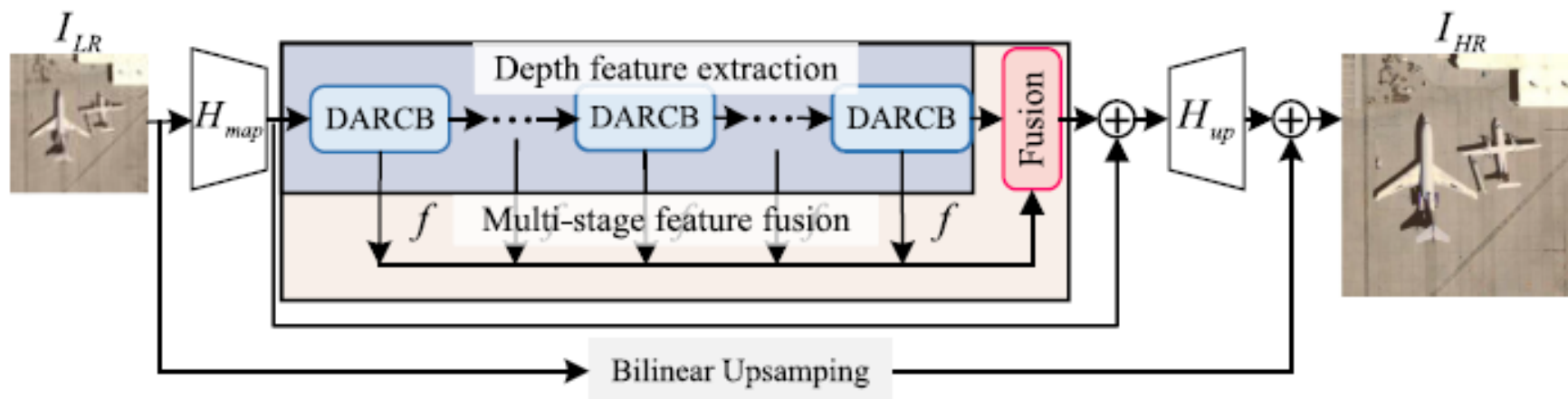
- 特徴表現を強化するために多段階特徴集約を使用する**距離注意残差接続 (DARCB)** コンポーネントを設計することで、単純なCNN連結モジュールより優れた進歩を達成
- **距離注意ブロック (DAB)** モジュールは、浅い特徴量の損失現象を抑制
- 精度と効率の良い妥協点を達成する**軽量画像SR再構築モデル DARN**を提案

提案手法

- **DARN**

- 入力された低画質の衛星リモートセンシング画像ILRに対して、我々の手法は、地上真実IGTに近い高画質画像IHRを再構成する
- 浅い特徴抽出モジュール、深い特徴抽出モジュール、多段階特徴結合ブロック、再構築モジュール

提案手法



提案手法

- Hmapで低次元画像を高次元空間にマッピング
- 複数のDARCBで構成される深層特徴抽出モジュールが、抽出された特徴を徐々に洗練

$$F_0 = H_{\text{map}}(I_{\text{LR}})$$
$$F_K = H_{\text{DARCB}_K}(F_{K-1}), K = 1, \dots, N$$

提案手法

- 軽量モデルの最適化とは、より小さなモデルパラメータ制約の下でモデルの性能と速度を向上させること
 - モデルの異なる深さからの特徴を融合することは、軽量モデルの性能を向上させる効果的な方法
- 融合モジュールは多段階の特徴量を融合することで特徴量の再利用率を向上させる

提案手法

- 合された特徴量 F_{final} を用いて、再構成モジュール H_{up} により高品質なリモートセンシング画像を復元

$$F_{\text{final}} = H_{\text{Fusion}} (f(F_0, F_1, \dots, F_N)) + F_0$$
$$I_{\text{HR}} = H_{\text{up}}(F_{\text{final}}) + I_{\text{LR}}$$

提案手法

- DARNはL 1一般化損失関数を用いて最適化

$$L_1 = \|I_{\text{GT}} - I_{\text{HR}}\|_1.$$

提案手法

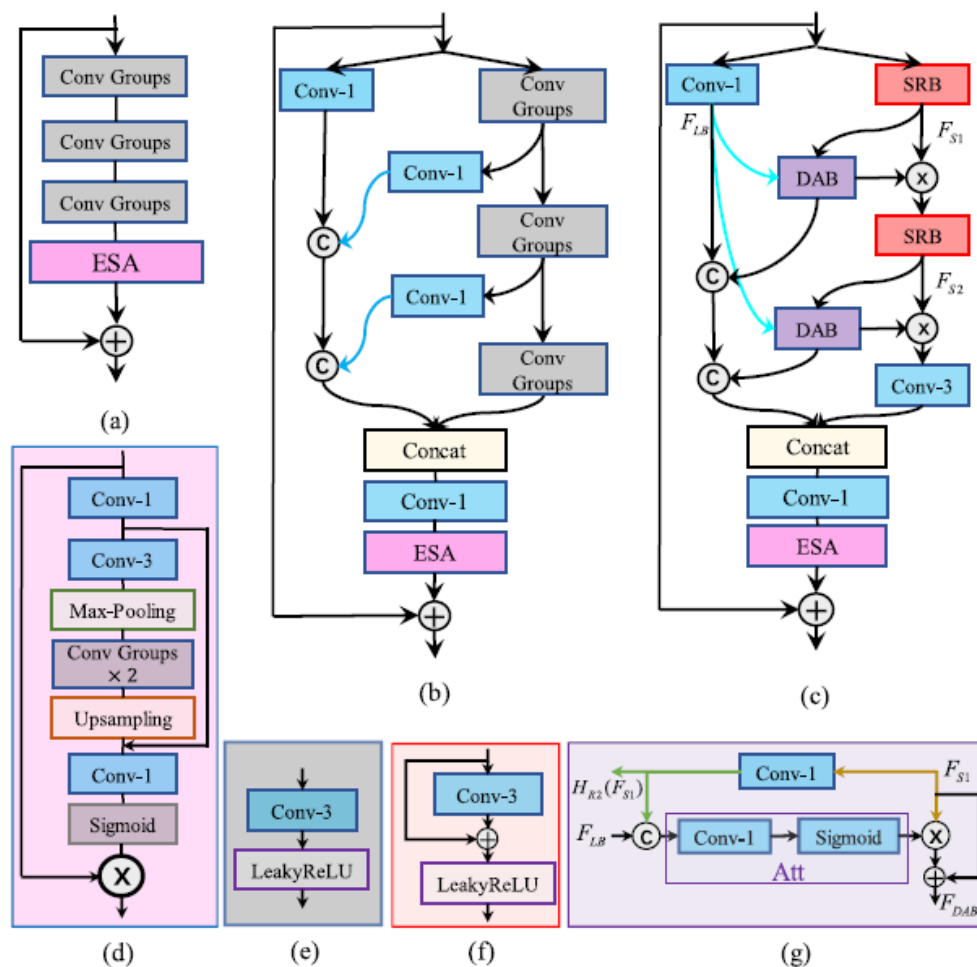


Fig. 3. (a) Simple baseline composed of three convolutions. (b) Feature distillation splicing (FDC) block formed by the simple model baseline transformation. (c) Distance attention residual connection block (DARCB). (d) Enhanced spatial attention (ESA). (e) Conv Groups. (f) Shallow residual block (SRB). (g) Distance attention block (DAB) for guiding CNN depth feature extraction. (a) Basic. (b) FDC. (c) DARCB. (d) ESA. (e) Conv Groups. (f) SRB. (g) DAB.

提案手法

- **軽量構造の考え方**

- 単純な連結によって形成される特徴量の冗長性と蓄積されたランダム誤差は、現在のCNNモデルに共通する問題

- **DABモジュール**を設計

提案手法

- 浅い特徴量には元の画像情報がすべて含まれているが、CNNは深い特徴量を抽出する際に、特徴量の損失、特徴量の冗長性、誤差の蓄積などの問題がある
- **DABモジュール**
 - 入力された精緻な特徴量を用いて、メインブランチCNNモジュール全体の詳細特徴量の損失を抑制
 - 浅い特徴に教師されたCNNは、抽出プロセスにおける誤差の蓄積を効果的に減らすことができる

提案手法

- **SRB**

- メインブランチの主要な構成ブロックとして導入することで、ネットワークの軽量化を図る
- 残存情報を柔軟に学習し、モデルをよりロバストにすることができる

提案手法

- 入力特徴量 $F(K-1)$ をメインブランチと特徴量抽出のためのブランチの2つの経路に分け、ブランチでは入力特徴量の元の情報を保持するために特徴量の精密化畳み込みを採用

$$F_{LB} = H_{R1}(F_{K-1}).$$

提案手法

- 2つのRSBモジュールが深い特徴を抽出するために使われ、2つのDABモジュールが深い特徴抽出の効率を高めるために使われる。最後に、深さ特徴 F_{DAB2} を洗練するために畳み込み層を採用

$$F_{S1} = H_{S1}(F_{K-1})$$

$$F_{\text{DAB1}} = H_{\text{Att1}}(f(F_{\text{LB}}, H_{R2}(F_{S1}))) * F_{S1} + F_{S1}$$

$$F_{S2} = H_{S2}(F_{\text{DAB1}})$$

$$F_{\text{DAB2}} = g(H_{\text{Att2}}(f(F_{\text{LB}}, H_{R3}(F_{S2}))) * F_{S2} + F_{S2}).$$

提案手法

- DARCBの多段階特徴量を結合

$$F_f = g(F_{LB}, H_{S2}(F_{S1}), H_{S3}(F_{S2}), F_{DAB2}).$$

- 最終的に、出力は

$$F_K = H_{ESA}(F_f) + F_f.$$

検証

- **DIV2K**

- 広範なSRデータセット
- train:800, validation:100, test:100

- モデルの再構成性能は、FeNet によって提案された2つのリモートセンシングデータRS-T1とRS-T2を用いてテスト



検証

- 追加ベンチマーク
 - Set5
 - Set14
 - BSD100
 - Urban100
 - Manga109

検証

- 平均特徴マップ

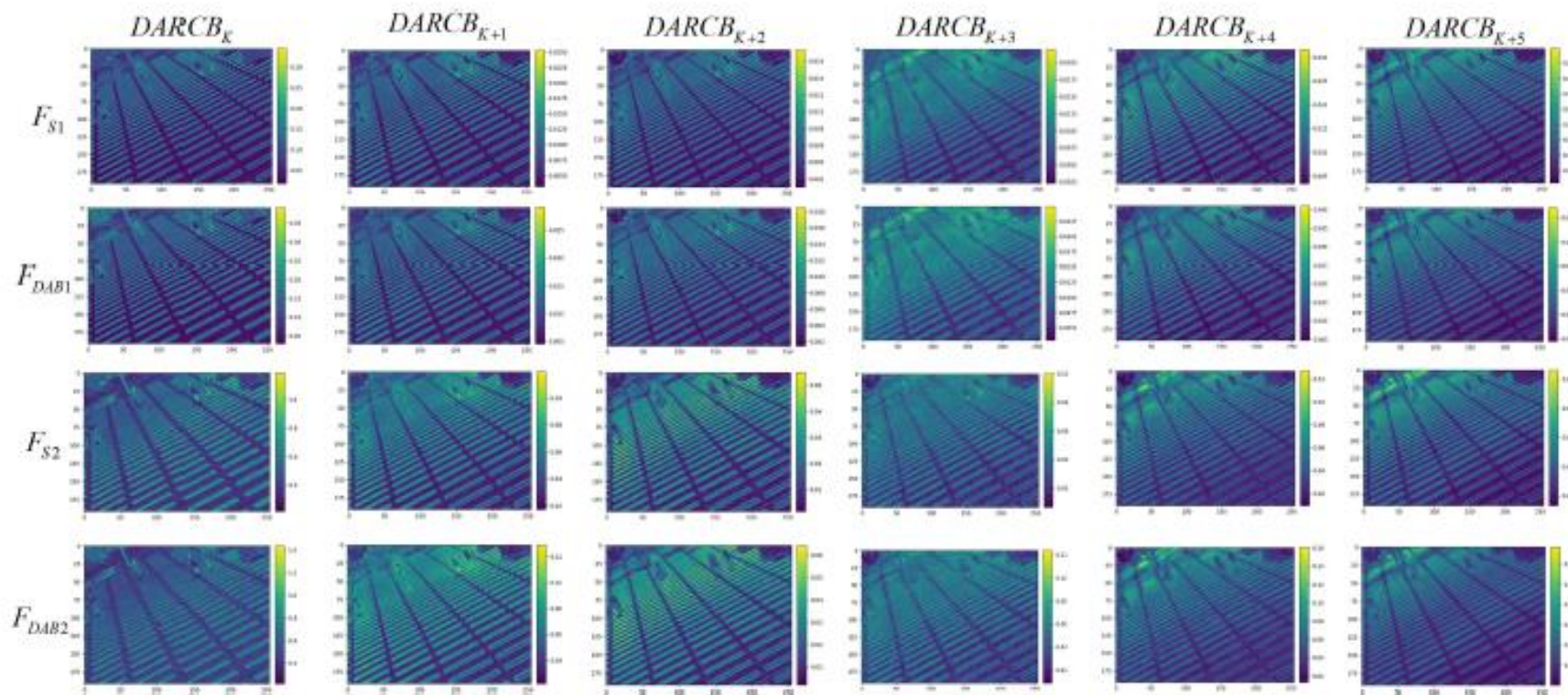


Fig. 4. Average feature maps at various stages of the DARCB.

検証

検証

- Basic、FDC、DARCBをコアコンポーネントとするモデルの性能

TABLE I
QUANTITATIVE COMPARISON OF DIFFERENT CORE COMPONENTS

Methods	Params	Set5		Set10		BSD100		Urban100		Manga109	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Basic	550K	32.10	0.8944	28.53	0.7802	27.54	0.7363	25.96	0.7818	30.27	0.9053
FDC	590K	32.15	0.8950	28.57	0.7814	27.56	0.7369	26.04	0.7840	30.36	0.9068
DARCB	606K	32.19	0.8952	28.58	0.7814	27.58	0.7373	26.07	0.7851	30.44	0.9078

DARCB is the core component proposed in this article.

The bold values represents the optimal performance under the corresponding indicator.

検証

- アブレーション研究
 - DABとESAモジュールの有効性を実証するため、アブレーション実験

TABLE II
ABLATION STUDY OF ESA AND DAB

Methods	Params	Set5		Set10		BSD100		Urban100		Manga109	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DARN-woESA	559K	32.06	0.8943	28.51	0.7804	27.53	0.7357	25.94	0.7807	30.27	0.9056
DARN-woDAB	590K	32.14	0.8952	28.56	0.7814	27.56	0.7368	26.02	0.7837	30.40	0.9073
DARN	606K	32.19	0.8952	28.58	0.7814	27.58	0.7373	26.07	0.7851	30.44	0.9078

The bold values represents the optimal performance under the corresponding indicator.

検証

- 特徴絞り込み率
 - 最適な絞り込み率を得るため、比較実験

TABLE III

FEATURE REFINEMENT RATE OPTIMAL SOLUTION IS INVESTIGATED ON THE BENCHMARK DATASET WITH A SCALE FACTOR OF $4\times$

Refinement rate	Params	Set5		Set10		BSD100		Urban100		Manga109	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.25	606K	32.19	0.8952	28.58	0.7814	27.58	0.7373	26.07	0.7851	30.44	0.9078
0.5	646K	32.12	0.8946	28.57	0.7813	27.58	0.7370	26.05	0.7839	30.41	0.9072
0.75	686K	32.19	0.8951	28.59	0.7817	27.58	0.7375	26.10	0.7858	30.48	0.9082

The bold values represents the optimal performance under the corresponding indicator.

検証

- モデルの効率

TABLE IV
QUANTIFY HOW LIGHTWEIGHT THE MODEL IS ON A BENCHMARK DATASET
WITH A SCALING FACTOR OF $4\times$

Methods	Params	Mutil-adds	Times	Manga109	
				PSNR	SSIM
LESRCNN [20]	774K	241.6G	33.27ms	29.94	0.9002
FeNet [22]	352K	20.4G	13.49ms	29.85	0.8992
DARN-S	350K	19.7G	9.86ms	30.48	0.9084
DARN	605K	32.9G	10.70ms	30.65	0.9105

The best and next best results are red and blue, respectively. Computing multiple addition and inference speed corresponds to a 1280×720 HR image.

検証

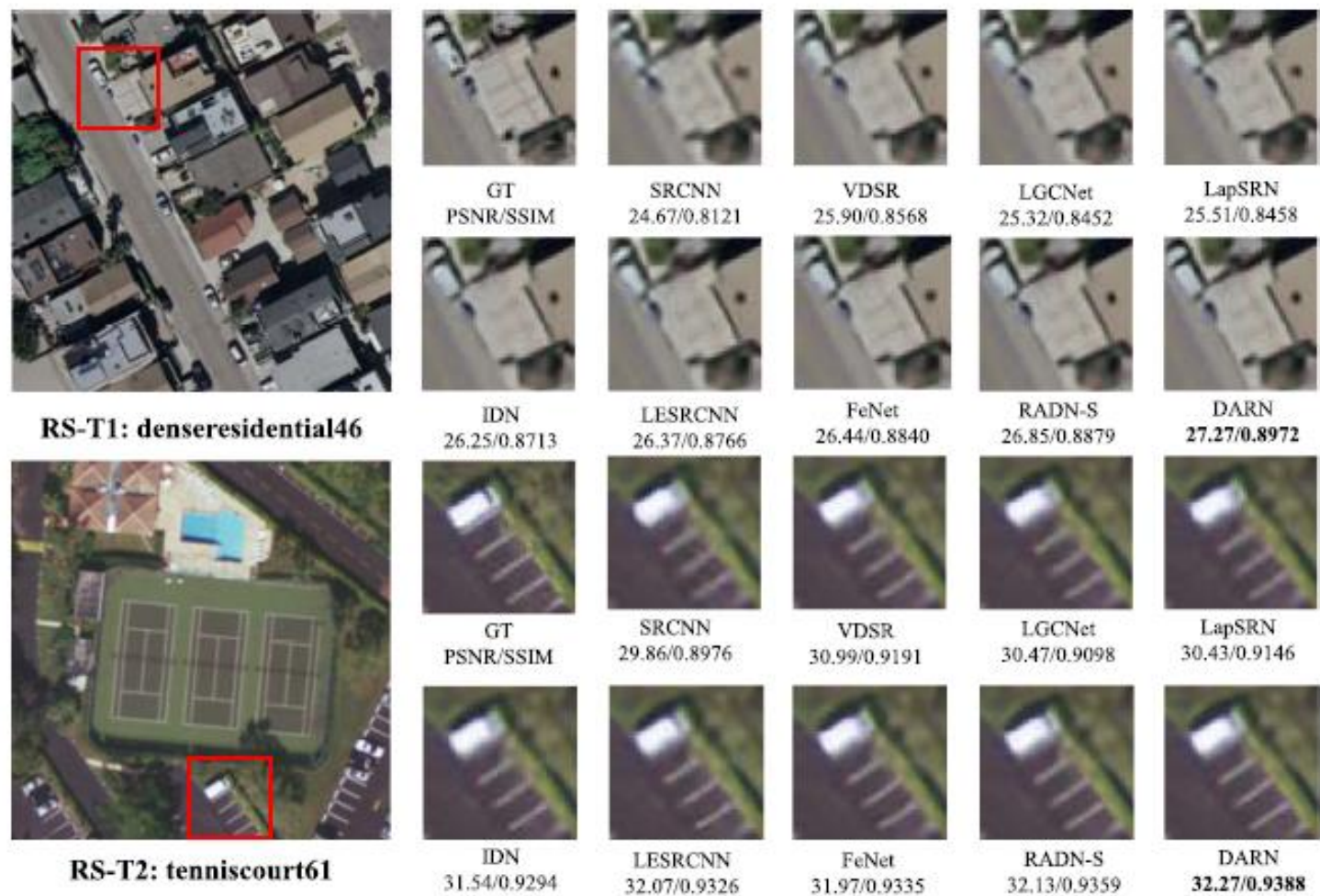
• RSモデルでの検証

TABLE V
QUANTITATIVE COMPARISON RESULTS OF REMOTE SENSING TEST IMAGES
RS-T1 AND RS-T2

Methods	Scale	Params	RS-T1	RS-T2
			PSNR/SSIM	PSNR/SSIM
Bicubic	×2	-	33.25/0.8934	30.64/0.8837
SRCNN [5]		57K	35.18/0.9243	32.87/0.9209
VDSR [8]		666K	35.85/0.9312	33.86/0.9312
LGCNet [33]		193K	35.65/0.9298	33.47/0.9281
LapSRN [30]		251K	35.69/0.9304	33.57/0.9286
IDN [19]		553K	36.13/0.9339	34.07/0.9329
LESRCNN [20]		626K	36.04/0.9328	34.00/0.9320
CARN-M [42]		412K	35.77/0.9314	33.84/0.9315
FeNet [22]		351K	36.23/0.9341	34.22/0.9337
DARN-S		350K	36.31/0.9347	34.35/0.9348
DARN		589K	36.38/0.9348	34.42/0.9357
Bicubic	×3	-	29.73/0.7818	27.23/0.7697
SRCNN [5]		57K	30.95/0.8228	28.59/0.8180
VDSR [8]		666K	31.55/0.9352	29.40/0.8391
LGCNet [33]		193K	31.30/0.8314	29.03/0.8312
LapSRN [30]		290K	31.47/0.8338	29.22/0.8352
IDN [19]		553K	31.73/0.8430	29.59/0.8450
LESRCNN [20]		810K	31.68/0.8398	29.65/0.8444
CARN-M [42]		412K	31.72/0.8426	29.62/0.8452
FeNet [22]		357K	31.89/0.8432	29.80/0.8481
DARN-S		350K	32.00/0.8483	29.98/0.8518
DARN		589K	32.08/0.8470	30.05/0.8537
Bicubic	×4	-	27.91/0.6968	25.40/0.6770
SRCNN [5]		57K	28.87/0.7382	26.46/0.7296
VDSR [8]		666K	29.33/0.7546	27.03/0.7525
LGCNet [33]		193K	29.13/0.7481	26.76/0.7426
LapSRN [30]		543K	29.51/0.7614	27.24/0.7600
IDN [19]		553K	29.56/0.7623	27.31/0.7627
LESRCNN [20]		774K	29.62/0.7625	27.41/0.7646
CARN-M [42]		412K	29.57/0.7624	27.37/0.7647
FeNet [22]		366K	29.70/0.7688	27.45/0.7672
DARN-S		350K	29.78/0.7682	27.59/0.7732
DARN		589K	29.85/0.7710	27.67/0.7758

“—” indicates that the result is unknown. The best and next best results are red and blue, respectively.

検証



検証

TABLE VI
QUANTITATIVE COMPARISON WITH STATE-OF-THE-ART METHODS ON SR BENCHMARK DATASETS

Methods	Scale	Params	Multi-Adds	Set5	Set14	BSD100	Urban100	Manga109
				PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic	×2	-	-	33.66/0.9299	30.24/0.8688	29.56/0.8431	26.88/0.8403	30.80/0.9339
SRCNN [5]		8K	52.7G	36.66/0.9542	32.45/0.9067	31.36/0.8879	29.50/0.8946	35.60/0.9663
FSRCNN [6]		13K	6.0G	37.00/0.9558	32.63/0.9088	31.53/0.8920	29.88/0.9020	36.67/0.9710
VDSR [8]		666K	612.6G	37.53/0.9587	33.03/0.9124	31.90/0.8960	30.76/0.9140	37.22/0.9750
LGCNet [33]		193K	178.1G	37.31/0.9580	32.94/0.9120	31.74/0.8939	30.53/0.9112	-
DRCN [16]		1774K	-	37.63/0.9588	33.04/0.9118	31.85/0.8942	30.75/0.9133	37.55/0.9732
LapSRN [30]		251K	29.9G	37.52/0.9591	32.99/0.9124	31.80/0.8952	30.41/0.9103	37.27/0.9740
DRRN [17]		298K	6796.9G	37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188	37.88/0.9749
MenNet [43]		678K	2662.4G	37.78/0.9597	33.28/0.9142	32.08/0.8978	31.31/0.9195	37.72/0.9740
IDN [19]		553K	124.6G	37.83/0.9600	33.30/0.9148	32.08/0.8985	31.27/0.9196	38.01/0.9749
LESRCNN [20]		626K	281.5G	37.65/0.9586	33.32/0.9148	31.95/0.8964	31.45/0.9206	37.89/0.9746
MADNet [21]		878K	187.1G	37.85/0.9600	33.39/0.9161	32.05/0.8981	31.59/0.9234	-
CARN-M [42]		412K	91.2G	37.53/0.9583	33.26/0.9141	31.92/0.8960	31.23/0.9193	37.58/0.9741
FeNet [22]		351K	77.9G	37.90/0.9602	33.45/0.9162	32.09/0.8985	31.75/0.9245	38.18/0.9752
DARN-S		350K	78.9G	37.97/0.9609	33.54/0.9172	32.19/0.9005	32.14/0.9284	38.74/0.9773
DARN		589K	131.6G	38.04/0.9610	33.63/0.9186	32.25/0.9012	32.40/0.9305	38.87/0.9776
Bicubic	×3	-	-	30.39/0.8682	27.55/0.7742	27.21/0.7385	24.46/0.7349	26.95/0.8556
SRCNN [5]		8K	52.7G	32.75/0.9090	29.30/0.8215	28.41/0.7863	26.43/0.7989	30.48/0.9117
FSRCNN [6]		13K	5.0G	33.18/0.9140	29.37/0.8240	28.53/0.7910	26.43/0.8080	31.10/0.9210
VDSR [8]		666K	612.6G	33.66/0.9213	29.77/0.8314	28.82/0.7976	27.14/0.8279	32.01/0.9340
LGCNet [33]		193K	79.0G	33.32/0.9172	29.67/0.8289	28.63/0.7923	26.77/0.8180	-
DRCN [16]		1774K	-	33.82/0.9226	29.76/0.8311	28.80/0.7963	27.15/0.8276	32.24/0.9343
LapSRN [30]		502K	115.2G	33.81/0.9220	29.79/0.8325	28.82/0.7980	27.07/0.8275	32.21/0.9350
DRRN [17]		298K	6796.9G	34.03/0.9244	29.96/0.8349	28.95/0.8004	27.53/0.8378	32.71/0.9379
MenNet [43]		678K	2662.4G	34.09/0.9248	30.00/0.8350	28.96/0.8001	27.56/0.8376	32.51/0.9369
IDN [19]		553K	124.6G	34.11/0.9253	29.99/0.8354	28.95/0.8013	27.42/0.8359	32.71/0.9381
LESRCNN [20]		810K	238.9G	33.93/0.9231	30.12/0.8380	28.91/0.8005	27.70/0.8415	32.76/0.9389
MADNet [21]		930K	88.4G	34.14/0.9251	30.20/0.8395	28.98/0.8023	27.78/0.8439	-
CARN-M [42]		412K	46.1G	33.99/0.9236	30.08/0.8367	28.91/0.8000	27.55/0.8385	32.78/0.9384
FeNet [22]		357K	35.2G	34.21/0.9256	30.15/0.8383	28.98/0.8020	27.82/0.8447	32.99/0.9394
DARN-S		355K	35.0	34.35/0.9274	30.34/0.8428	29.09/0.8065	28.17/0.8528	33.59/0.9446
DARN		596K	58.4G	34.48/0.9286	30.41/0.8443	29.15/0.8076	28.38/0.8570	33.76/0.9457
Bicubic	×4	-	-	28.42/0.8104	26.00/0.7027	25.96/0.6675	23.14/0.6577	24.89/0.7866
SRCNN [5]		8K	52.7G	30.48/0.8626	27.50/0.7513	26.90/0.7101	24.52/0.7221	27.58/0.8555
FSRCNN [6]		13K	4.6G	30.72/0.8660	27.61/0.7550	26.98/0.7150	24.62/0.7280	27.90/0.8610
VDSR [8]		666K	612.6	31.35/0.8838	28.01/0.7674	27.29/0.7251	25.18/0.7524	28.83/0.8870
LGCNet [33]		193K	44.5G	30.87/0.8746	27.82/0.7630	27.08/0.7186	24.82/0.7399	-
DRCN [16]		1774K	-	31.53/0.8854	28.02/0.7670	27.23/0.7233	25.14/0.7510	28.93/0.8854
LapSRN [30]		502K	149.4G	31.54/0.8852	28.09/0.7700	27.32/0.7275	25.21/0.7562	29.09/0.8900
DRRN [17]		298K	6796.9	31.68/0.8888	28.21/0.7720	27.38/0.7284	25.44/0.7638	29.45/0.8946
MenNet [43]		678K	2662.4G	31.74/0.8893	28.26/0.7723	27.40/0.7281	25.50/0.7630	29.42/0.8942
IDN [19]		553K	32.3G	31.82/0.8903	28.25/0.7730	27.41/0.7297	25.41/0.7632	29.41/0.8942
LESRCNN [20]		774K	241.6G	31.88/0.8903	28.44/0.7772	27.45/0.7313	25.77/0.7732	29.94/0.9002
MADNet [21]		1002K	54.1G	32.01/0.8925	28.45/0.7781	27.47/0.7327	25.77/0.7751	-
CARN-M [42]		412K	32.5G	31.92/0.8903	28.42/0.7762	27.44/0.7304	25.63/0.7688	29.80/0.8989
FeNet [22]		366K	20.4G	32.02/0.8919	28.38/0.7764	27.47/0.7319	25.75/0.7747	29.85/0.8992
DARN-S		363K	19.7	32.16/0.8951	28.58/0.7817	27.57/0.7374	26.08/0.7859	30.48/0.9084
DARN		606K	32.9G	32.24/0.8963	28.64/0.7830	27.61/0.7390	26.25/0.7913	30.65/0.9105

“—” indicates that the result is unknown. The best and next best results are red and blue, respectively. Computing multiple addition corresponds to a 1280×720 HR image.

結果

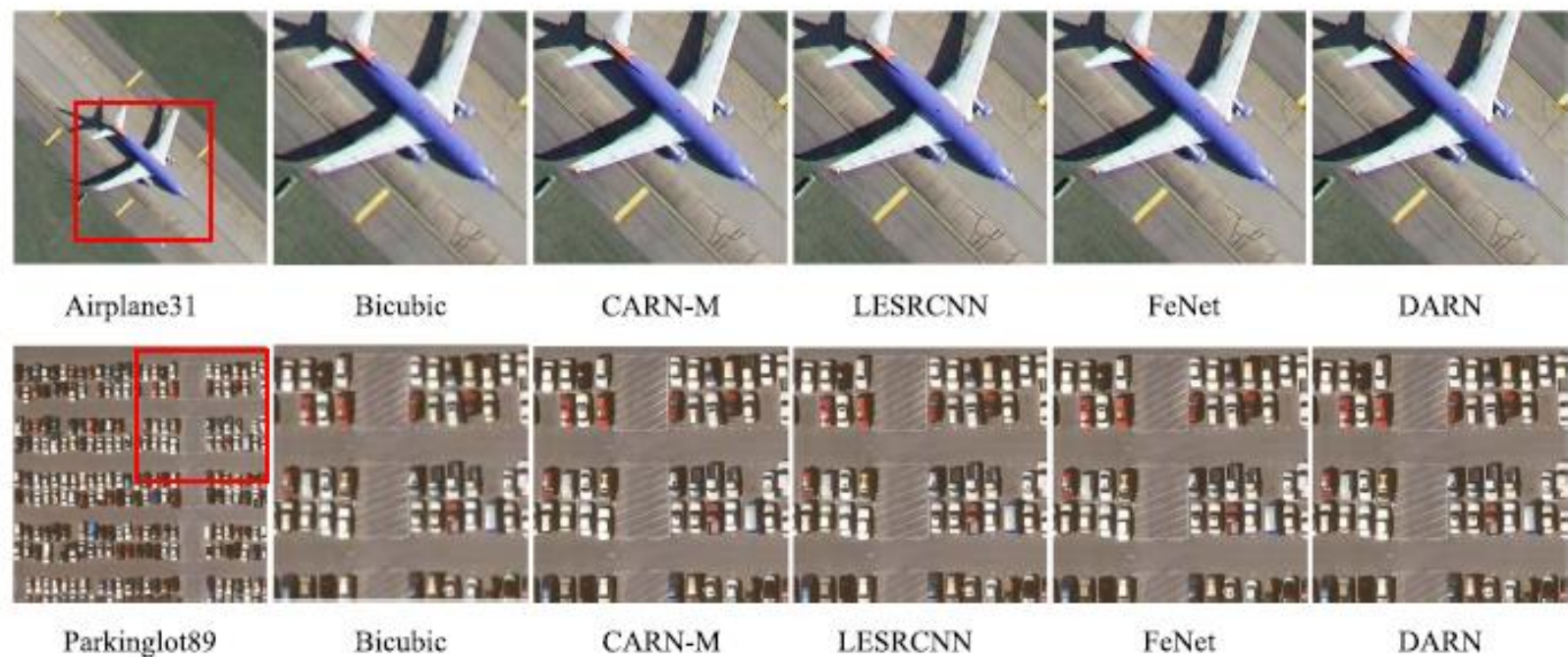


Fig. 6. Visual comparison of the DARN with other SR methods on real remote sensing images. Please zoom in for better visualization.

結果

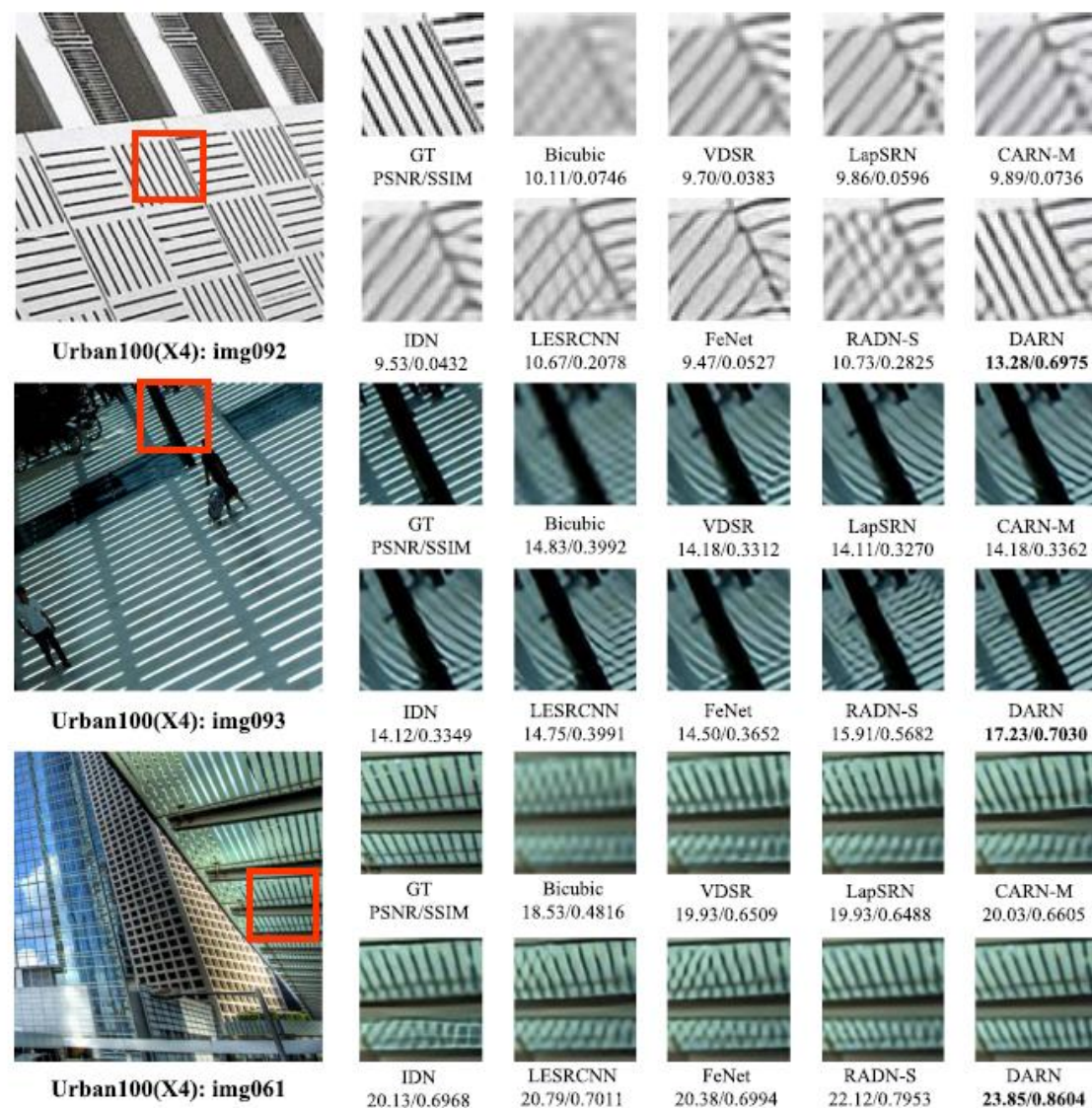


Fig. 7. Visual comparisons of the DARN with other SR methods on Urban100($\times 4$) datasets. Please zoom in for better visualization.

結論

- 既存の軽量モデルと比較し最良の結果を出した
- DARNは主にコアコンポーネントDARNに依存して優れた成績を出す
- DARCBの特徴可視化は、我々のDABにおける事前情報の利用が、深層CNNの詳細な特徴の損失を抑制できることをさらに証明