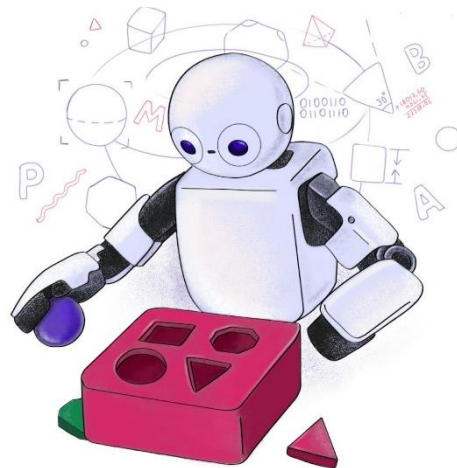


TP558 - Tópicos avançados em
Machine Learning:
*Enhanced Deep Residual
Networks for Single Image
Super-Resolution*



Eylen Jhuliana Mercado Ontiveros
eylen.ontiveros@mtel.inatel.br

Introdução

- A imagem de super-resolução é um problema fundamental no processamento de imagens, em que o objetivo é melhorar a resolução e a qualidade de uma imagem de baixa resolução para torná-la mais clara e detalhada.
- A alta resolução é essencial para uma variedade de aplicações, como telas de alta definição, reconhecimento facial, imagens médicas, entre outras.
- No entanto, frequentemente encontramos imagens de baixa resolução que não apresentam detalhes importantes. É aí que entra a imagem de super-resolução.





+ ? =



EDSR

Enhanced deep
super-resolution network

- Esse algoritmo é usado para aumentar a resolução de uma imagem, um problema que tem sido um desafio no campo da visão computacional.
- O EDSR usa redes neurais convolucionais profundas para resolver esse problema.



$$+ \quad ? \quad =$$



EDSR

- O objetivo da SISR (Image Single Super Resolution) é reconstruir uma imagem de alta resolução I^{SR} a partir de uma única imagem de baixa resolução I^{LR}
- Muitos estudos presumem que a I^{LR} é uma versão com amostragem reduzida bicúbica da I^{HR}
- Mas outros fatores de degradação:
 - Desfoque
 - Decimação
 - Ruído

Deep Neural Networks

- Recentemente, as redes neurais profundas proporcionaram um desempenho significativamente melhor em termos de relação sinal-ruído de pico (PSNR) no problema de SR.

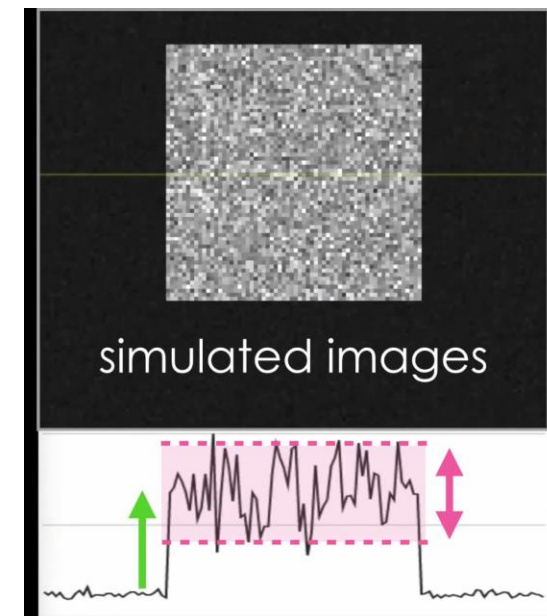
SNR



SNR Limits

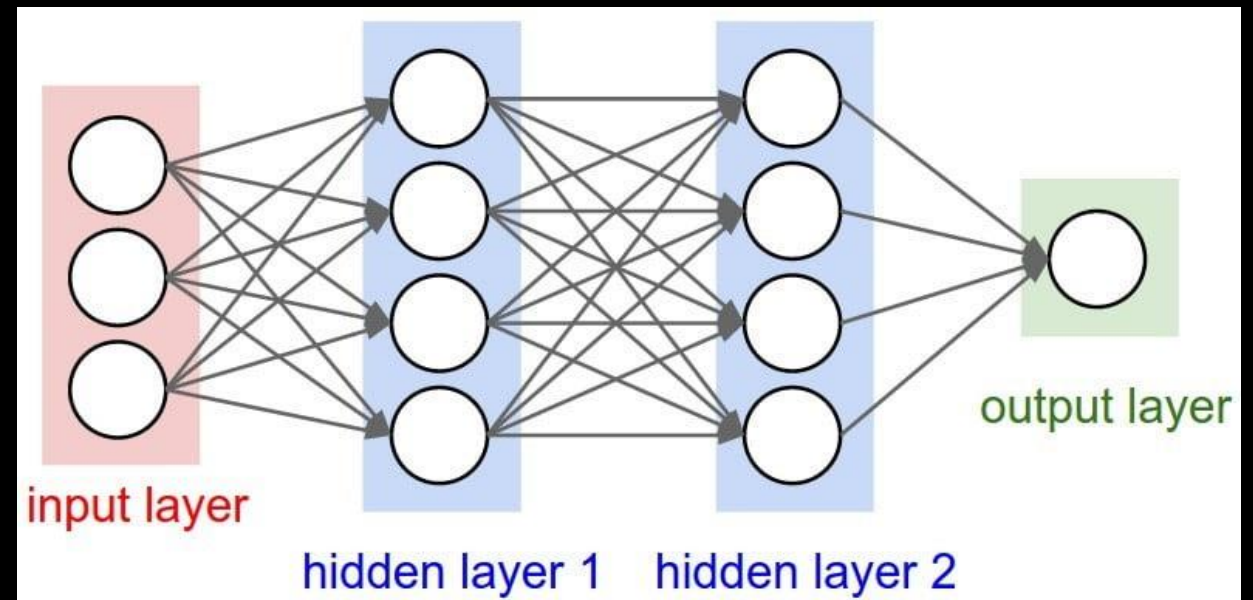
- Detection
- Precision of intensity measurements
- Resolution

$$PSNR = 10 \log_{10}(R^2 / MSE)$$

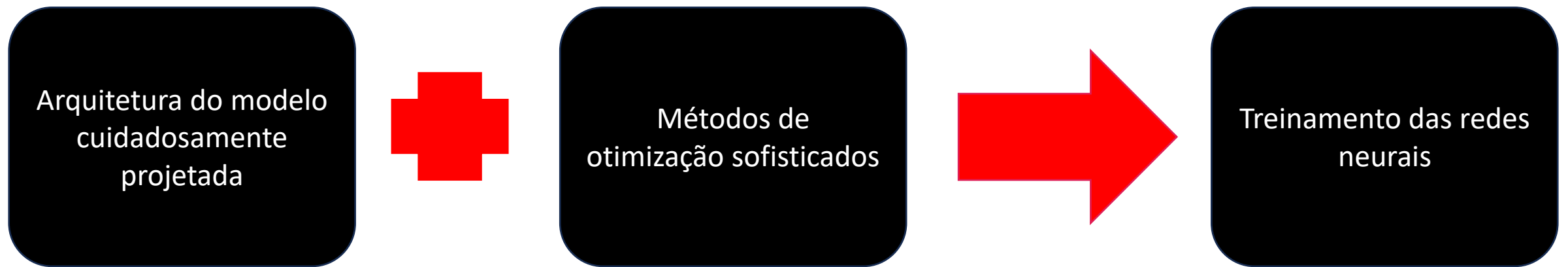


Deep Neural Networks

- Apresentam limitações
 - Desempenho da reconstrução dos modelos de rede neural é sensível a pequenas alterações na arquitetura
 - Mesmo modelo atinge diferentes níveis de desempenho com diferentes técnicas de inicialização e treinamento.



Deep Neural Networks



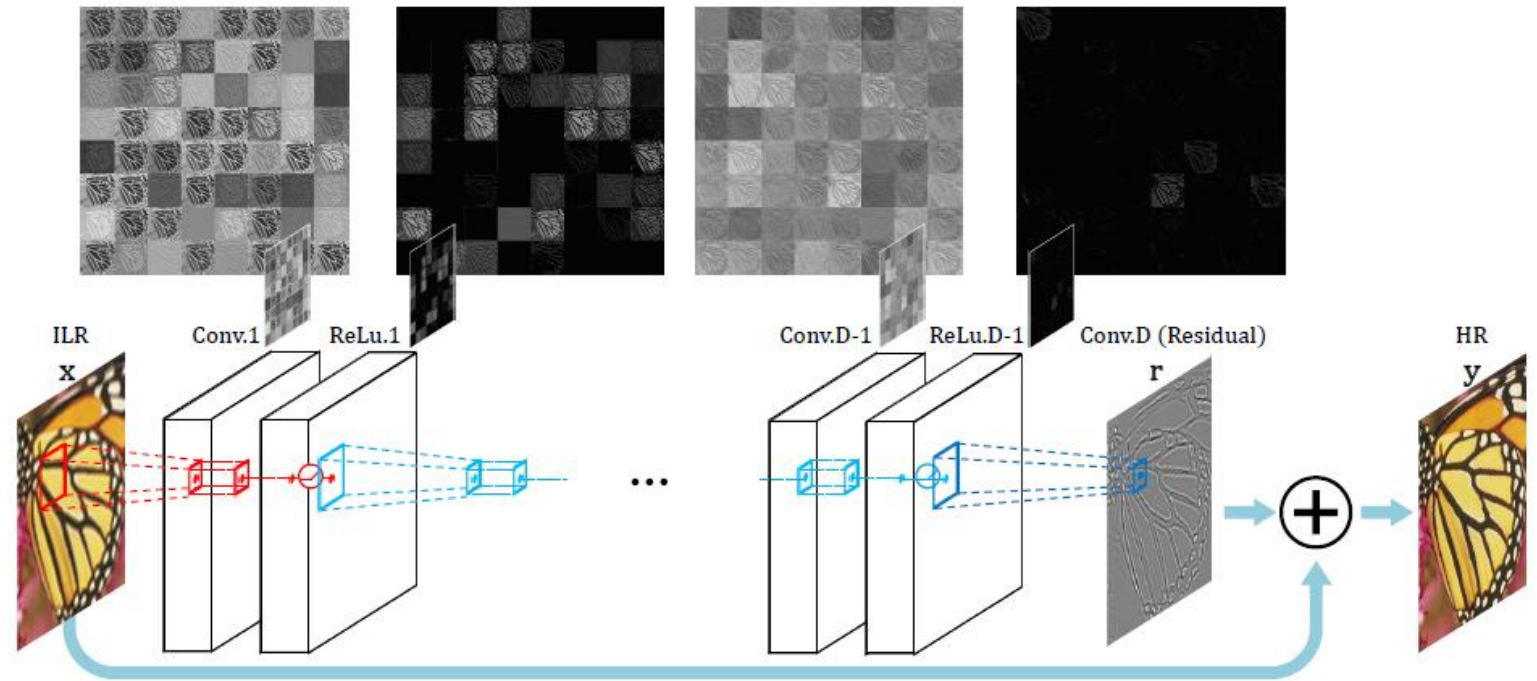


Deep Neural Networks

- Maioria dos algoritmos de SR existentes trata a super-resolução de diferentes fatores de escala como problemas independentes.
 - Sem considerar e utilizar relações mútuas entre diferentes escalas na SR.
- Esses algoritmos exigem muitas redes específicas de escala que precisam ser treinadas de forma independente para lidar com várias escalas.

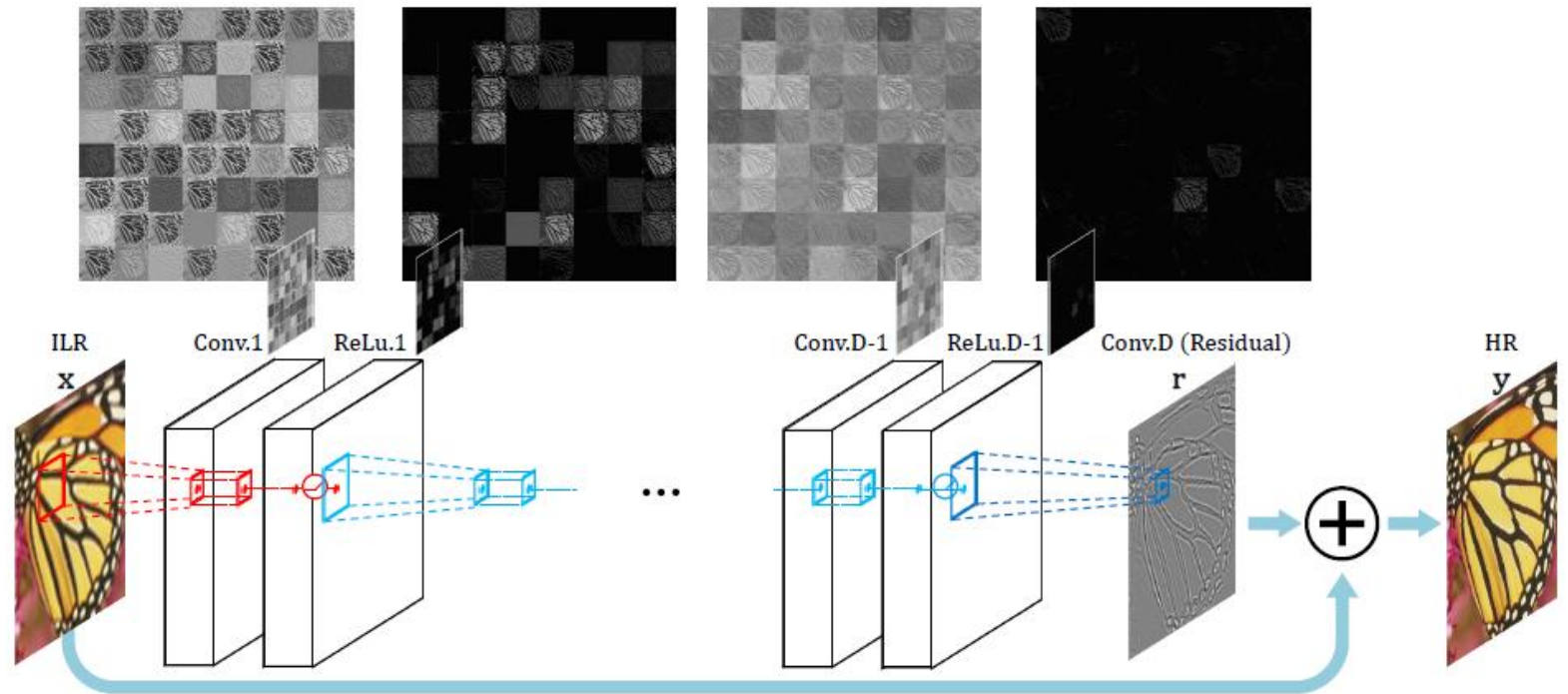
VDSR (Very Deep Super Resolution)

- Pode lidar com a super-resolução de várias escalas em conjunto em uma única rede.
- O treinamento do modelo VDSR com várias escalas aumenta substancialmente o desempenho e supera o treinamento específico da escala
- O que implica a redundância entre os modelos específicos da escala.



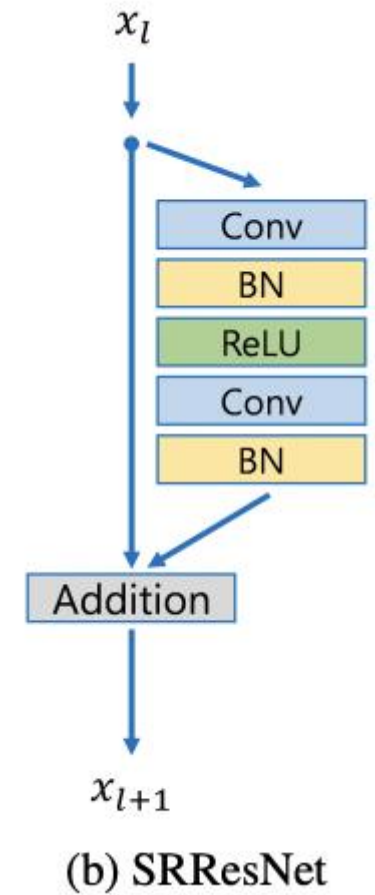
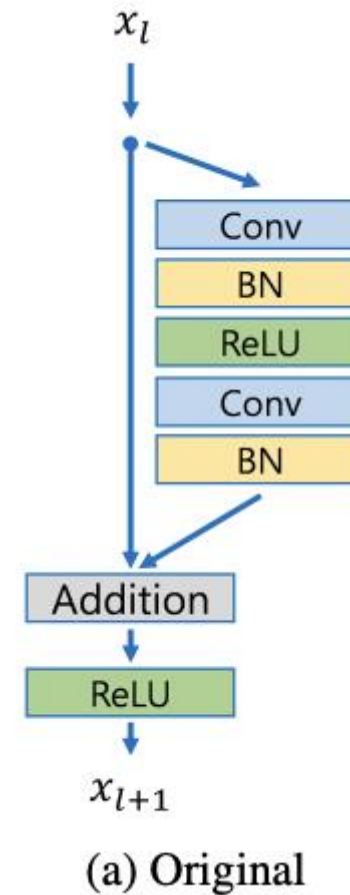
VDSR (Very Deep Super Resolution)

- Requer uma imagem interpolada bicúbica como entrada
- O que leva a um maior tempo de computação e memória em comparação com as arquiteturas com método de sobreamostragem específico de escala.



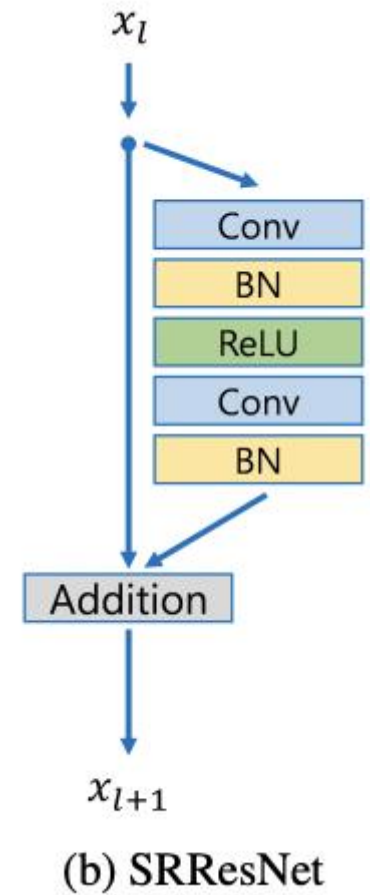
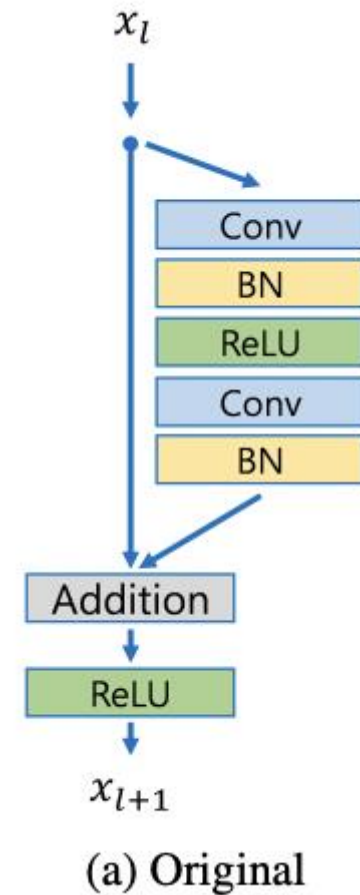
SRResNet

- Resolvido com sucesso esses problemas de tempo e memória com bom desempenho
- Emprega a arquitetura ResNet sem muitas modificações.
 - ResNet original foi proposta para resolver problemas de visão computacional de nível superior, como classificação e detecção de imagens.

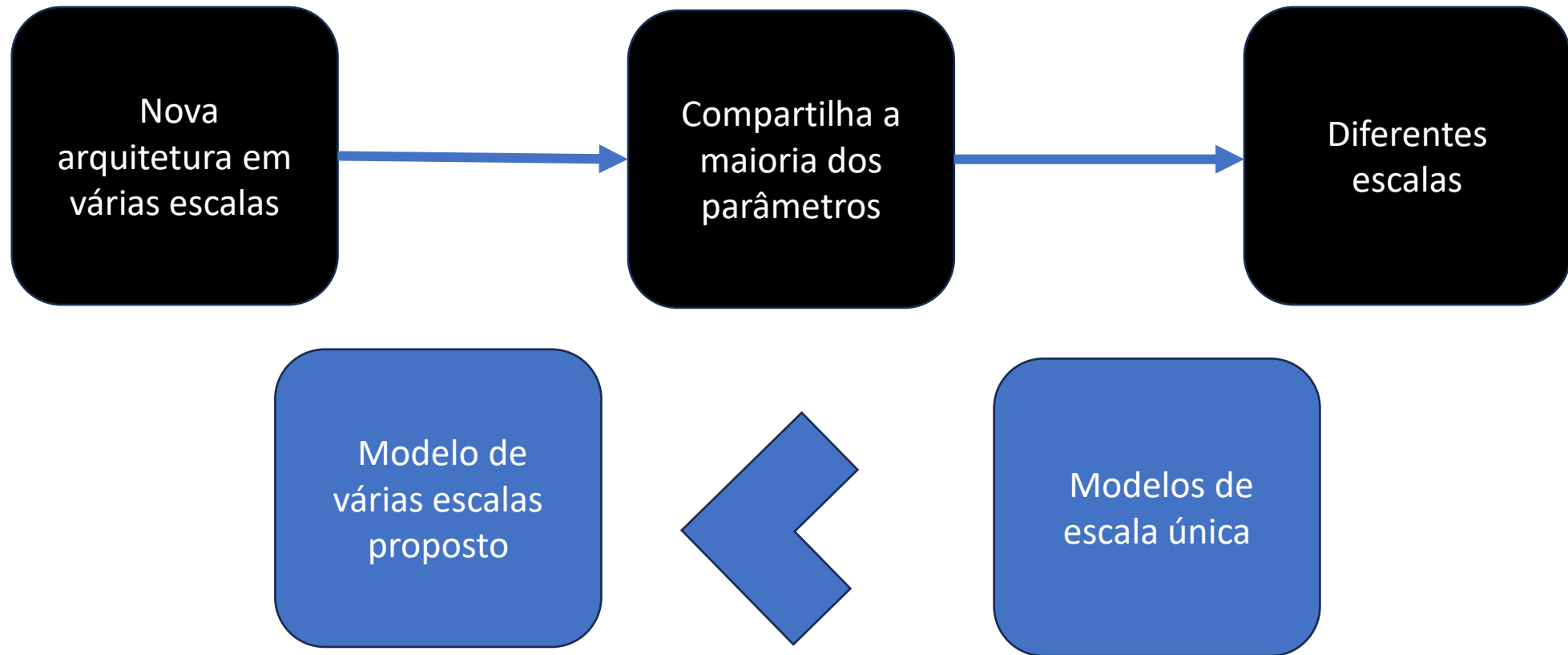


SRResNet

- Otimizamos analisando e removendo módulos desnecessários para simplificar a arquitetura da rede.
- Treinamento:
 - Com uma função de perda apropriada
 - Modificação cuidadosa do modelo durante o treinamento.
- Mostramos experimentalmente que o esquema modificado produz melhores resultados.



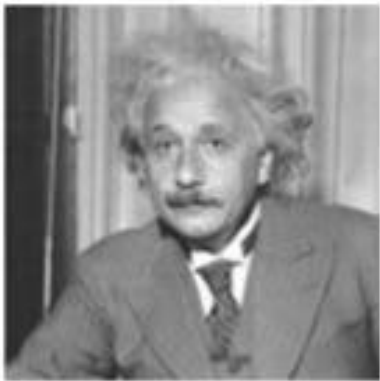
SRResNet



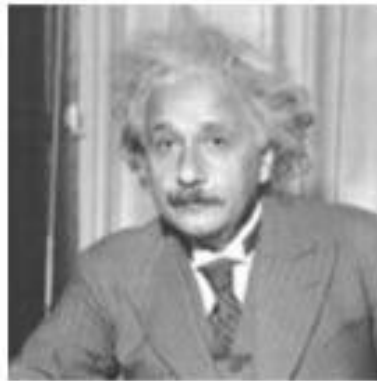
Menos parâmetros mas apresenta desempenho comparável.

SRResNet

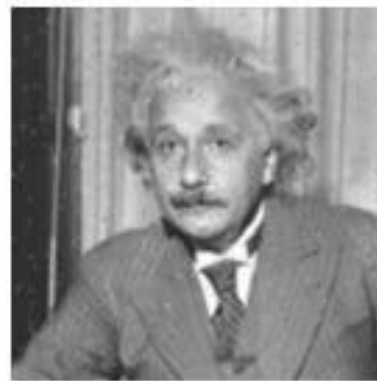
- As redes de super-resolução de escala única e múltipla propostas apresentam o desempenho mais avançado em todos os conjuntos de dados em termos de:
 - PSNR (Peak signal to Noise Ratio)
 - SSIM (Structural Similarity Index Measure)



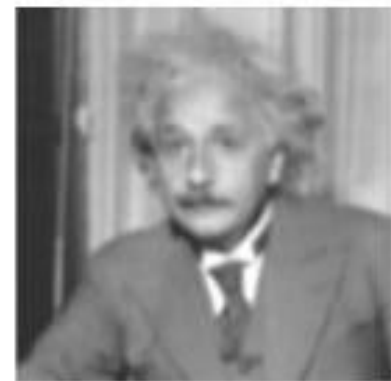
Original
SSIM=1



PSNR=26.547
SSIM=0.988



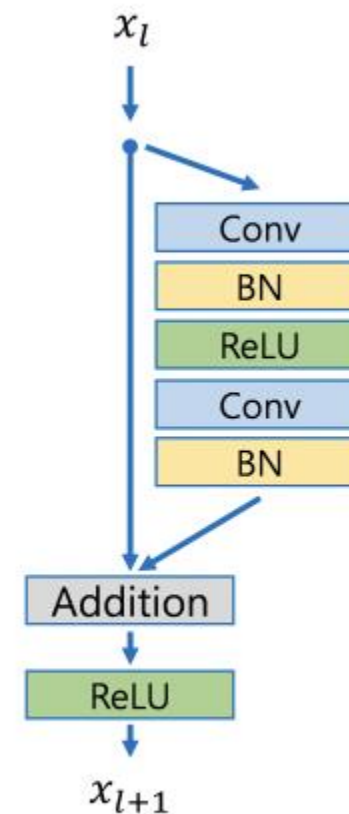
PSNR=26.547
SSIM=0.840



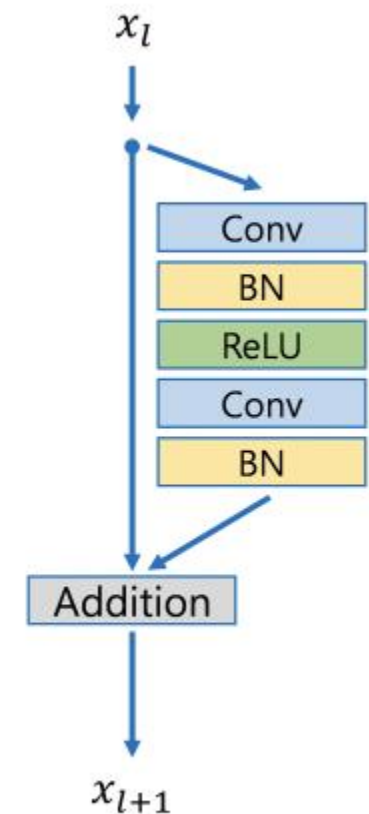
PSNR=26.547
SSIM=0.694

Residual Blocks

- As redes residuais apresentaram excelente desempenho em problemas de visão computacional, desde as tarefas de baixo nível até as de alto nível.
 - Embora Ledig et al. tenham aplicado com sucesso a arquitetura ResNet ao problema de super-resolução com SRResNet



(a) Original



(b) SRResNet

EDSR

- Quanto ao funcionamento interno do EDSR, ele aproveita as conexões residuais para melhorar a qualidade das imagens de saída.
- Durante o treinamento, o modelo aprende a calcular as diferenças entre a versão de baixa resolução da imagem e sua contraparte de alta resolução, usando essas informações para gerar imagens de alta qualidade.



Arquitetura e funcionamento

- A arquitetura EDSR é composta por uma série de blocos residuais, cada um dos quais processa a imagem de entrada e aprende a reconstruir uma versão de alta resolução dela. Esses blocos são projetados para capturar e preservar os detalhes finos da imagem.
- O EDSR tem uma arquitetura específica que inclui várias camadas e parâmetros.
- Ele usa conexões de salto locais e globais para facilitar o aprendizado e realiza o dimensionamento no final da rede, permitindo que imagens de alta resolução de diferentes fatores de dimensionamento sejam reconstruídas em um único modelo.

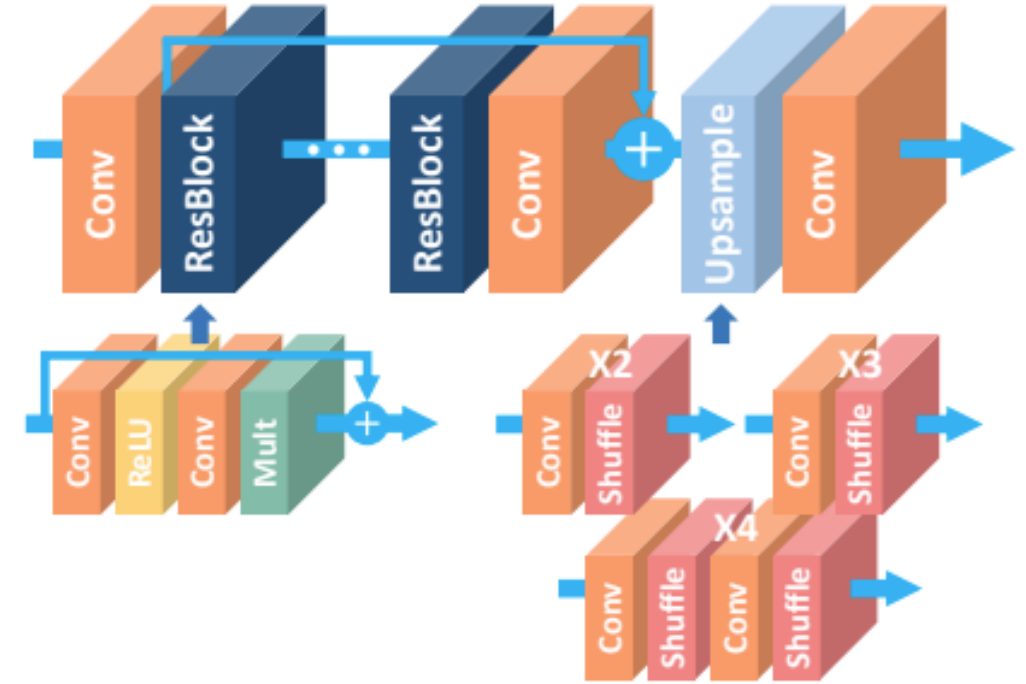


Figure 3: The architecture of the proposed single-scale SR network (EDSR).

Residual Blocks

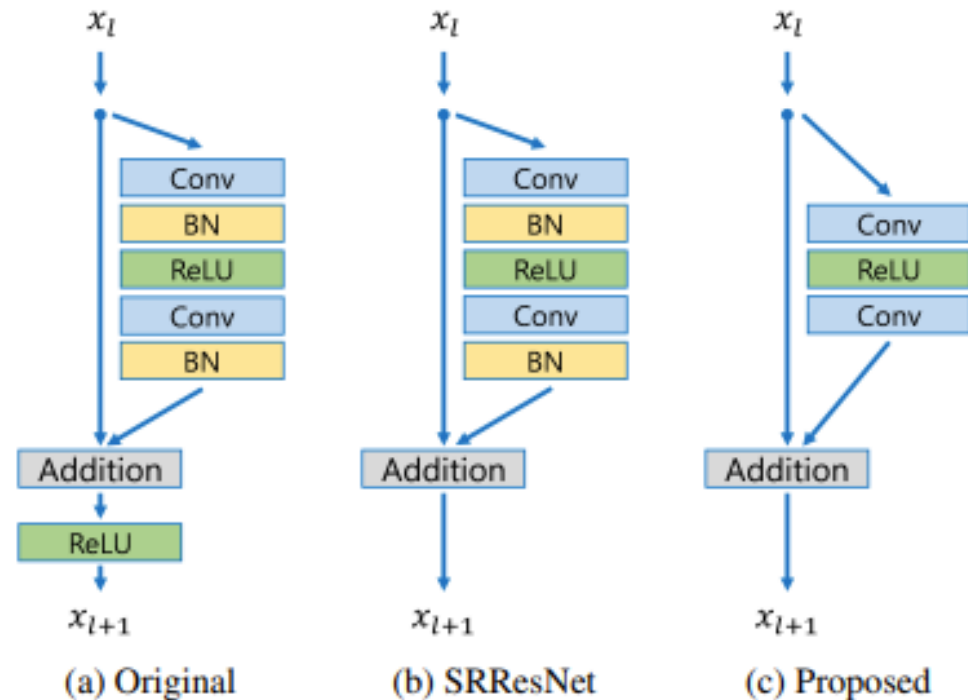


Figure 2: Comparison of residual blocks in original ResNet, SRResNet, and ours.

Eliminação das camadas de normalização.

- Como as camadas de normalização em lote normalizam os recursos, elas eliminam a flexibilidade de alcance das redes ao normalizar os recursos
- O uso da memória da GPU também é suficientemente reduzido
- Economiza aproximadamente 40% do uso de memória durante o treinamento

Podemos criar um modelo maior que tenha melhor desempenho do que a estrutura ResNet convencional com recursos computacionais limitados

Single-scale Model

- A arquitetura geral da CNN com profundidade (o número de camadas) B e largura (o número de canais de recursos) F ocupa aproximadamente $O(BF)$ de memória com parâmetros $O(BF^2)$.
- Portanto, **aumentar F em vez de B** pode maximizar a capacidade do modelo ao considerar recursos computacionais limitados.
- No entanto, aumentar o número de mapas de recursos acima de um determinado nível tornaria o procedimento de treinamento numericamente instável.

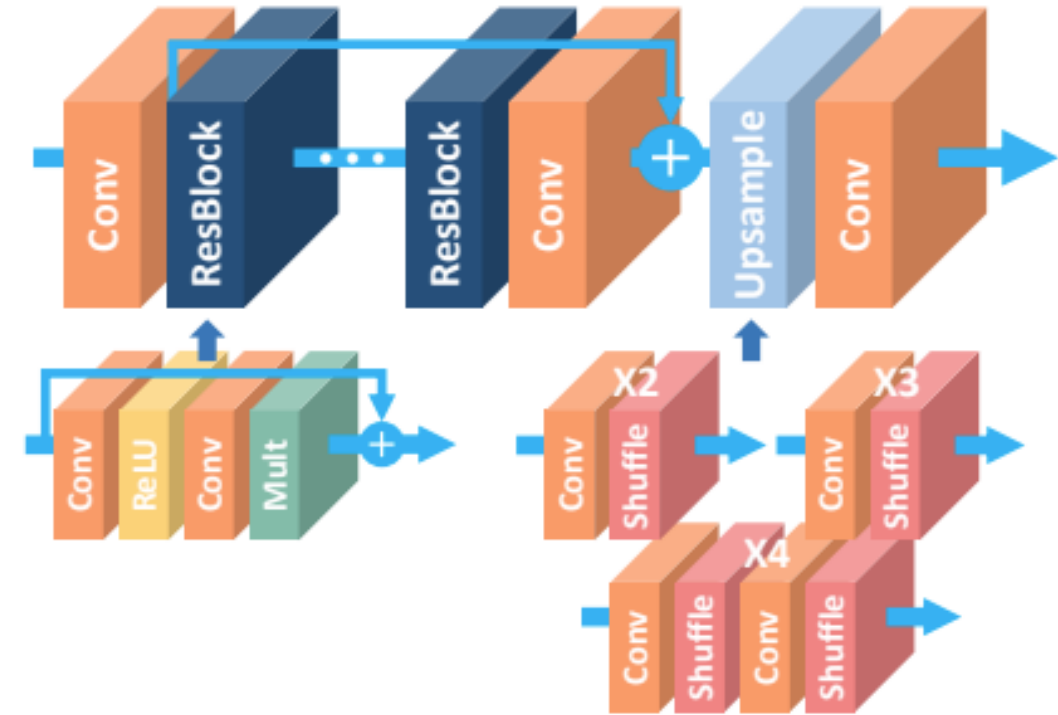


Figure 3: The architecture of the proposed single-scale SR network (EDSR).

Single-scale Model

- O escalonamento residual com fator 0,1 é aplicado no caminho residual antes de adicionar de volta ao caminho convolucional.
- Em cada bloco residual, camadas de escala constante são colocadas após as últimas camadas de convolução. Na fase de teste, esta camada pode ser integrada à camada de convolução anterior para eficiência computacional.
- Não há camadas de ativação ReLU fora dos blocos residuais.
- Modelo de linha de base: Camadas de escala residuais não são usadas, pois apenas 64 feature maps são usados para cada camada de convolução.
- EDSR: O modelo de linha de base é expandido definindo $B = 32$, $F = 256$ com um fator de escala de 0,1.

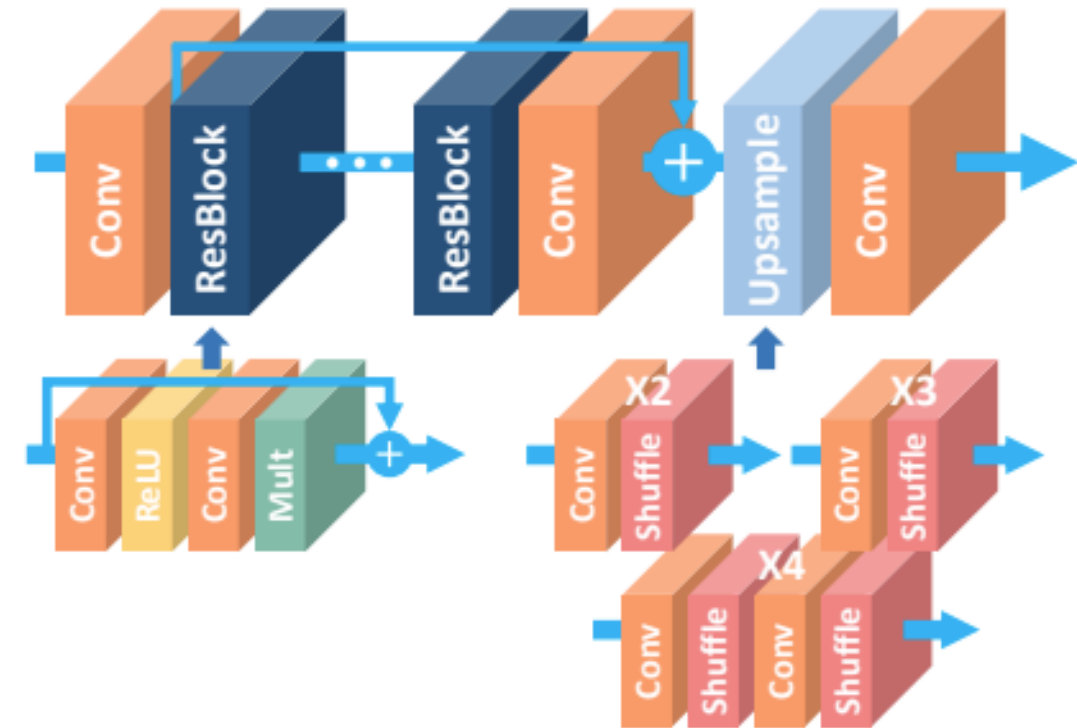
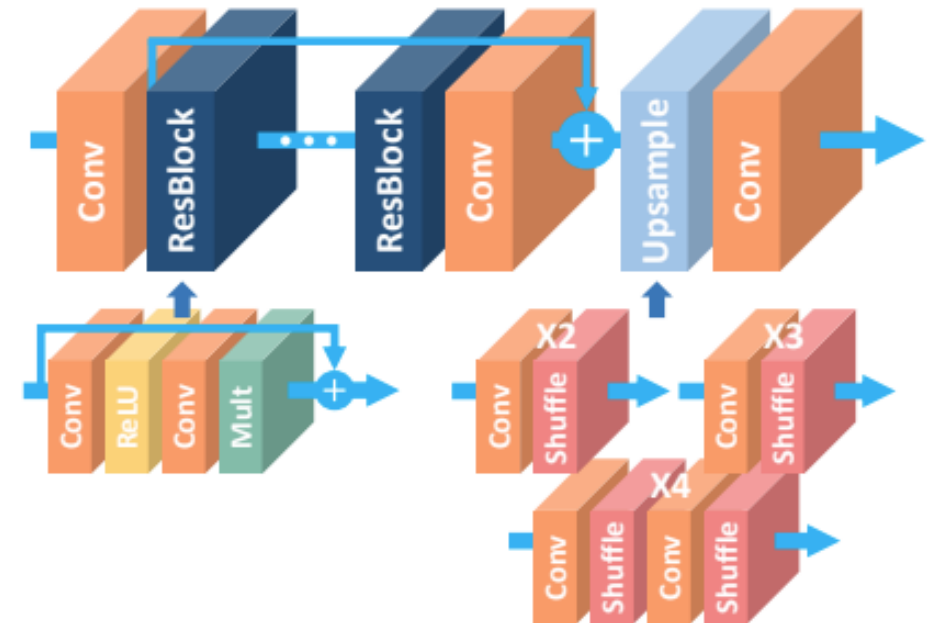
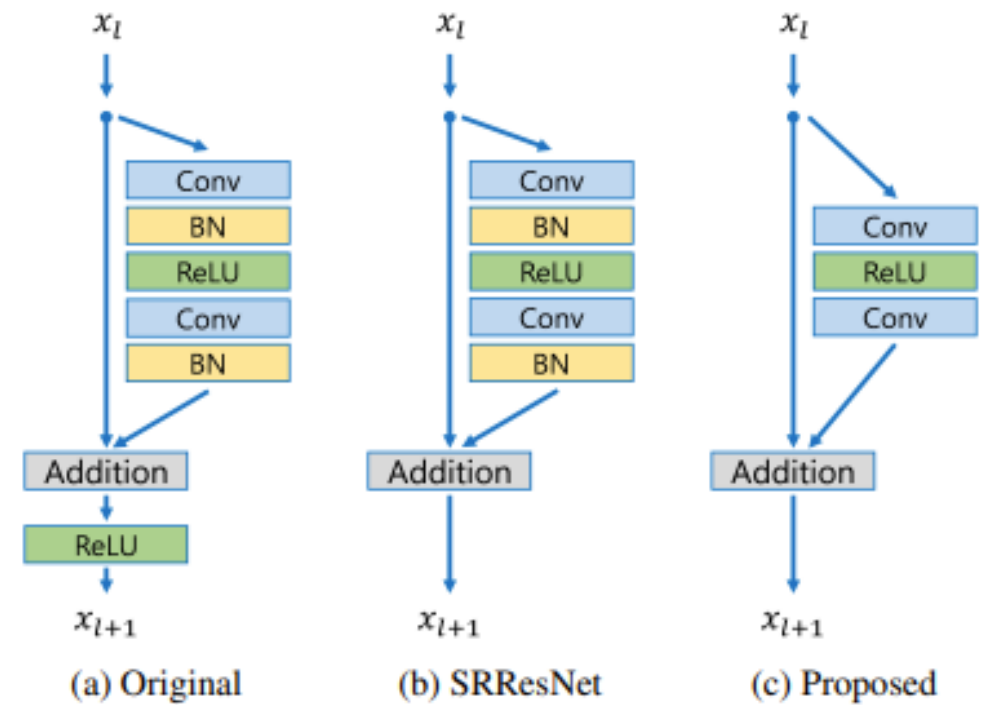


Figure 3: The architecture of the proposed single-scale SR network (EDSR).

EDSR Network

- 32 blocos residuais com 256 canais
- L1 por pixel em vez de L2
- Sem camadas de normalização em lote para manter a flexibilidade da faixa
- Fator de escala de 0,1 para adição residual para estabilizar o treinamento



Single-scale Model

1. Normalize Input by subtracting DIV2K RGB mean
2. Conv2d with 64 filters and kernel size=3
3. Resblock
4. Conv2d
5. Add(ResBlock output and original Input)
6. Upsampling(Conv2d→Pixel Shuffle)

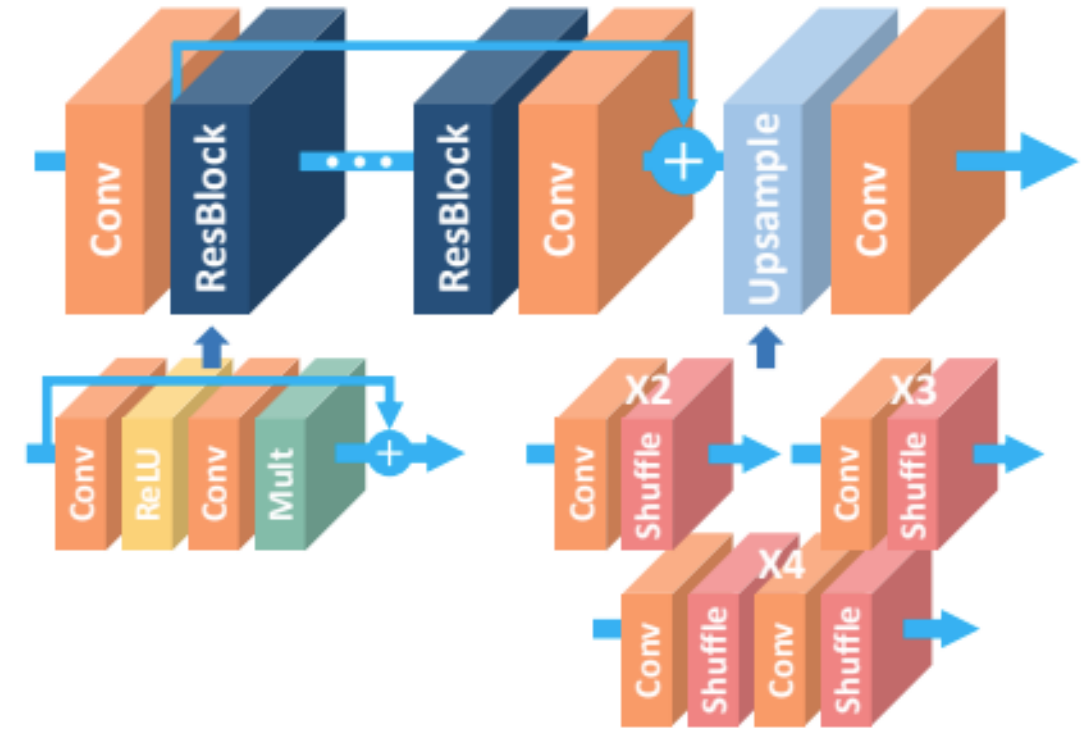


Figure 3: The architecture of the proposed single-scale SR network (EDSR).

Multi-scale Model MDSR

- A super-resolução em múltiplas escalas é uma tarefa inter-relacionada.
- **Modelo de Linha de Base:** Uma única ramificação principal com **B = 16** blocos residuais para que a maioria dos parâmetros sejam compartilhados em diferentes escalas, conforme mostrado.
- Primeiro, os **módulos de pré-processamento** estão localizados no topo das redes para reduzir a variação das imagens de entrada de diferentes escalas.
- Cada módulo de pré-processamento consiste em dois blocos residuais com núcleos 5×5.

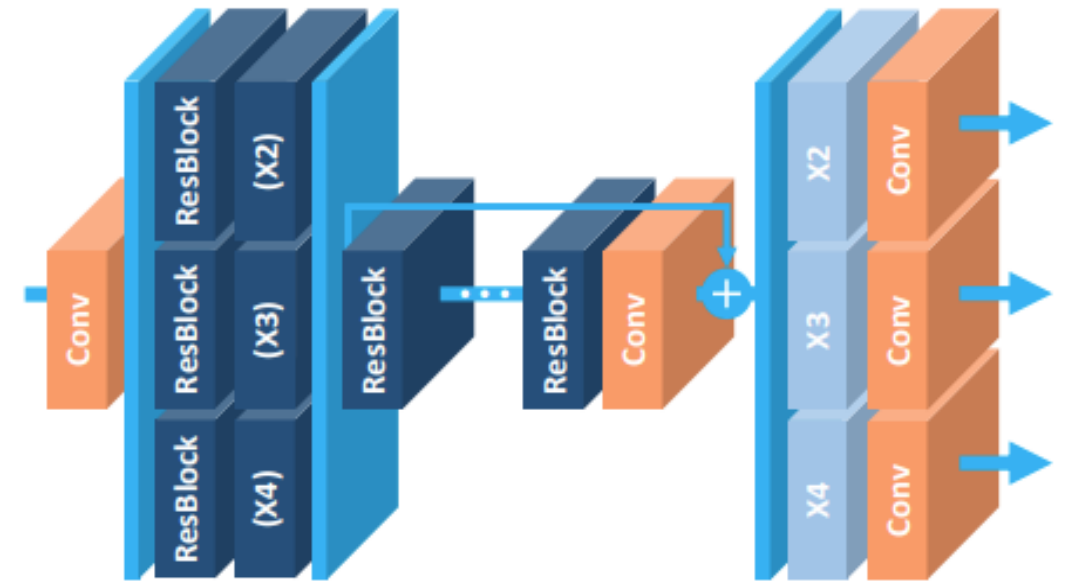


Figure 5: The architecture of the proposed multi-scale SR network (MDSR).

Multi-scale Model MDSR

- **No final** do modelo multiescala, **os módulos de upsampling específicos da escala** são localizados em paralelo para lidar com a reconstrução em várias escalas.
- Os modelos de linha de base de escala única para 3 escalas diferentes têm cerca de 1,5M de parâmetros cada, totalizando 4,5M, o modelo de linha de base multiescala tem apenas **3,2M de parâmetros**, com desempenho comparável aos modelos de escala única.
- **MDSR: $B = 80$ e $F = 64$, aproximadamente 5 vezes mais profundidade em comparação com o modelo de linha de base de várias escalas, apenas 2,5 vezes mais parâmetros** são necessários, pois os blocos residuais são mais leves do que as partes específicas da escala.

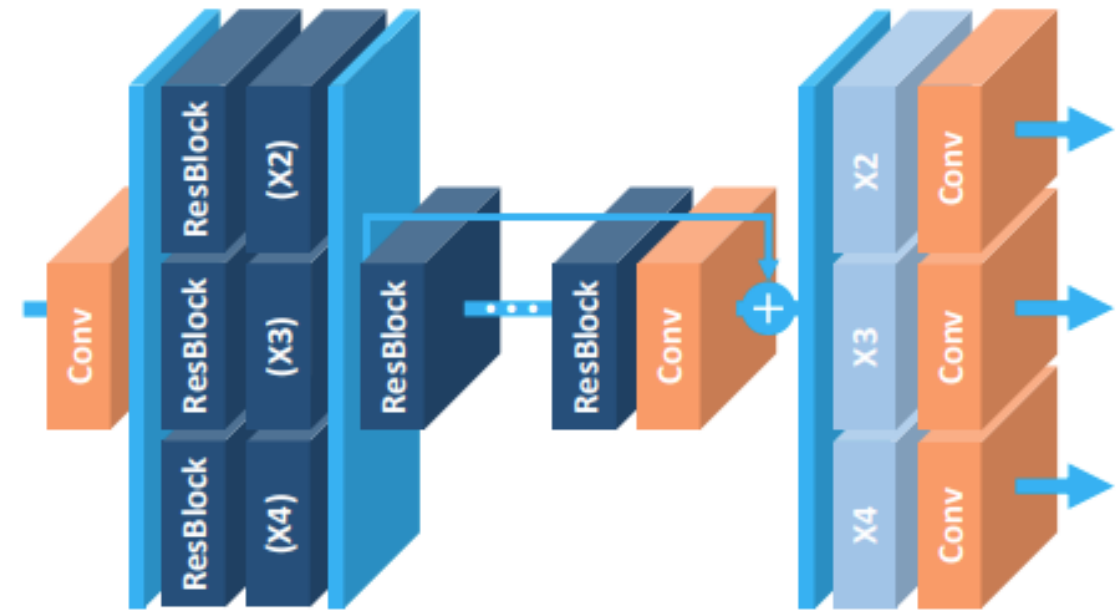


Figure 5: The architecture of the proposed multi-scale SR network (MDSR).

Treinamento e testes

- O conjunto de dados **DIV2K** é um conjunto de dados de imagens de alta qualidade (resolução 2K)
- Consiste em 800 imagens de treinamento, 100 imagens de validação e 100 imagens de teste.
- Quatro conjuntos de dados de referência padrão também são comparados: **Set5**, **Set14**, **B100** e **Urban100**



Treinamento

- **Augmentação de Dados:** As imagens de entrada foram patches RGB de tamanho 48x48 extraídos das imagens LR. Para aumentar o conjunto de dados de treinamento, foram aplicadas rotações aleatórias de 90 graus e reflexões horizontais aleatórias.
- **Pré-processamento:** Todas as imagens foram pré-processadas subtraindo o valor médio RGB do conjunto de dados DIV2K.

Scale	SRResNet (L2 loss)	SRResNet (L1 loss)	Our baseline (Single-scale)	Our baseline (Multi-scale)	EDSR (Ours)	MDSR (Ours)	EDSR+ (Ours)	MDSR+ (Ours)
×2	34.40 / 0.9662	34.44 / 0.9665	34.55 / 0.9671	34.60 / 0.9673	35.03 / 0.9695	34.96 / 0.9692	35.12 / 0.9699	35.05 / 0.9696
×3	30.82 / 0.9288	30.85 / 0.9292	30.90 / 0.9298	30.91 / 0.9298	31.26 / 0.9340	31.25 / 0.9338	31.39 / 0.9351	31.36 / 0.9346
×4	28.92 / 0.8960	28.92 / 0.8961	28.94 / 0.8963	28.95 / 0.8962	29.25 / 0.9017	29.26 / 0.9016	29.38 / 0.9032	29.36 / 0.9029

Table 2: Performance comparison between architectures on the DIV2K validation set (PSNR(dB) / SSIM). Red indicates the best performance and blue indicates the second best. EDSR+ and MDSR+ denote self-ensemble versions of EDSR and MDSR.

Treinamento

- Para o MDSR, para cada atualização, o minibatch com uma escala selecionada aleatoriamente entre $\times 2$, $\times 3$ e $\times 4$. Somente os módulos que correspondem à escala selecionada são ativados e atualizados
- **As redes são treinadas usando a perda L1 em vez de L2.** A minimização de L2 é geralmente preferida, pois maximiza a PSNR. No entanto, verificou-se empiricamente que a perda L1 proporciona melhor convergência do que a L2.

Scale	SRResNet (L2 loss)	SRResNet (L1 loss)	Our baseline (Single-scale)	Our baseline (Multi-scale)	EDSR (Ours)	MDSR (Ours)	EDSR+ (Ours)	MDSR+ (Ours)
$\times 2$	34.40 / 0.9662	34.44 / 0.9665	34.55 / 0.9671	34.60 / 0.9673	35.03 / 0.9695	34.96 / 0.9692	35.12 / 0.9699	35.05 / 0.9696
$\times 3$	30.82 / 0.9288	30.85 / 0.9292	30.90 / 0.9298	30.91 / 0.9298	31.26 / 0.9340	31.25 / 0.9338	31.39 / 0.9351	31.36 / 0.9346
$\times 4$	28.92 / 0.8960	28.92 / 0.8961	28.94 / 0.8963	28.95 / 0.8962	29.25 / 0.9017	29.26 / 0.9016	29.38 / 0.9032	29.36 / 0.9029

Table 2: Performance comparison between architectures on the DIV2K validation set (PSNR(dB) / SSIM). Red indicates the best performance and blue indicates the second best. EDSR+ and MDSR+ denote self-ensemble versions of EDSR and MDSR.

Geometric Self-ensemble During Testing

- A imagem de entrada ILR para gerar sete entradas aumentadas:

$$I_{n,i}^{LR} = T_i (I_n^{LR})$$

- onde T_i é a 8 transformação geométrica.
- Após a super-resolução, todas as imagens são transformadas inversamente para a geometria original:

$$\tilde{I}_{n,i}^{SR} = T_i^{-1} (I_{n,i}^{SR})$$

Geometric Self-ensemble During Testing

- The transformed outputs are averaged all together to make the self-ensemble result:

$$I_n^{SR} = \frac{1}{8} \sum_{i=1}^8 \tilde{I}_{n,i}^{SR}.$$

- Esse método de auto-montagem tem uma vantagem sobre outros conjuntos, pois **não requer treinamento adicional de modelos separados**. É vantajoso principalmente quando o tamanho do modelo ou o tempo de treinamento são importantes.
- Ele proporciona aproximadamente o mesmo ganho de desempenho em comparação com o método de conjunto de modelos convencionais que requer modelos treinados individualmente.
- **Os métodos que usam** o autoensemble são indicados pela adição do prefixo "+" ao nome do método, ou seja, **EDSR+/MDSR+**.

Resultados

- A avaliação é realizada nas 10 imagens do conjunto de validação DIV2K, com os critérios PSNR e SSIM.
- Podemos ver que há uma margem de melhorias para EDSR, MDSR, EDSR+ e MDSR+.
- Os modelos exigem muito menos memória da GPU, pois não têm camadas de normalização em lote.

Scale	SRResNet (L2 loss)	SRResNet (L1 loss)	Our baseline (Single-scale)	Our baseline (Multi-scale)	EDSR (Ours)	MDSR (Ours)	EDSR+ (Ours)	MDSR+ (Ours)
× 2	34.40 / 0.9662	34.44 / 0.9665	34.55 / 0.9671	34.60 / 0.9673	35.03 / 0.9695	34.96 / 0.9692	35.12 / 0.9699	35.05 / 0.9696
× 3	30.82 / 0.9288	30.85 / 0.9292	30.90 / 0.9298	30.91 / 0.9298	31.26 / 0.9340	31.25 / 0.9338	31.39 / 0.9351	31.36 / 0.9346
× 4	28.92 / 0.8960	28.92 / 0.8961	28.94 / 0.8963	28.95 / 0.8962	29.25 / 0.9017	29.26 / 0.9016	29.38 / 0.9032	29.36 / 0.9029

Set5, Set14, B100, Urban100, & DIV2K

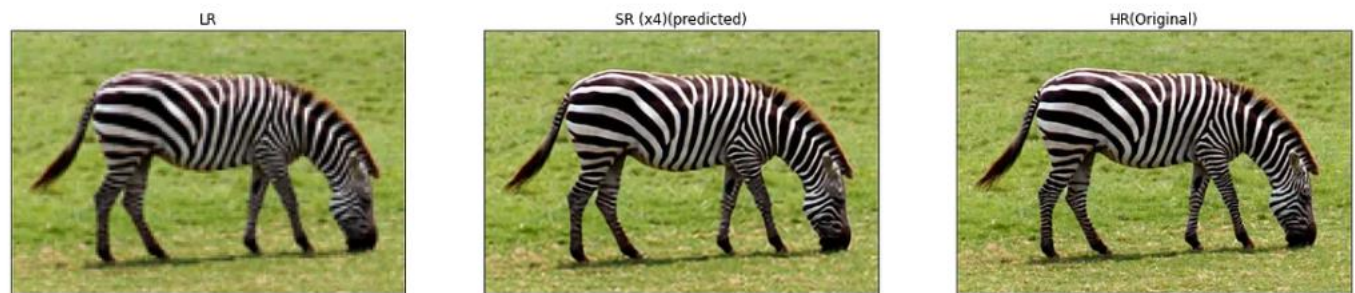
- Os modelos apresentam uma melhoria significativa em comparação com outros métodos, como SRCNN, VDSR e SRResNet.
- As lacunas aumentam ainda mais após a execução do autoensemble.

Dataset	Scale	Bicubic	A+ [27]	SRCNN [4]	VDSR [11]	SRResNet [14]	EDSR (Ours)	MDSR (Ours)	EDSR+ (Ours)	MDSR+ (Ours)
Set5	×2	33.66 / 0.9299	36.54 / 0.9544	36.66 / 0.9542	37.53 / 0.9587	- / -	38.11 / 0.9601	38.11 / 0.9602	38.20 / 0.9606	38.17 / 0.9605
	×3	30.39 / 0.8682	32.58 / 0.9088	32.75 / 0.9090	33.66 / 0.9213	- / -	34.65 / 0.9282	34.66 / 0.9280	34.76 / 0.9290	34.77 / 0.9288
	×4	28.42 / 0.8104	30.28 / 0.8603	30.48 / 0.8628	31.35 / 0.8838	32.05 / 0.8910	32.46 / 0.8968	32.50 / 0.8973	32.62 / 0.8984	32.60 / 0.8982
Set14	×2	30.24 / 0.8688	32.28 / 0.9056	32.42 / 0.9063	33.03 / 0.9124	- / -	33.92 / 0.9195	33.85 / 0.9198	34.02 / 0.9204	33.92 / 0.9203
	×3	27.55 / 0.7742	29.13 / 0.8188	29.28 / 0.8209	29.77 / 0.8314	- / -	30.52 / 0.8462	30.44 / 0.8452	30.66 / 0.8481	30.53 / 0.8465
	×4	26.00 / 0.7027	27.32 / 0.7491	27.49 / 0.7503	28.01 / 0.7674	28.53 / 0.7804	28.80 / 0.7876	28.72 / 0.7857	28.94 / 0.7901	28.82 / 0.7876
B100	×2	29.56 / 0.8431	31.21 / 0.8863	31.36 / 0.8879	31.90 / 0.8960	- / -	32.32 / 0.9013	32.29 / 0.9007	32.37 / 0.9018	32.34 / 0.9014
	×3	27.21 / 0.7385	28.29 / 0.7835	28.41 / 0.7863	28.82 / 0.7976	- / -	29.25 / 0.8093	29.25 / 0.8091	29.32 / 0.8104	29.30 / 0.8101
	×4	25.96 / 0.6675	26.82 / 0.7087	26.90 / 0.7101	27.29 / 0.7251	27.57 / 0.7354	27.71 / 0.7420	27.72 / 0.7418	27.79 / 0.7437	27.78 / 0.7425
Urban100	×2	26.88 / 0.8403	29.20 / 0.8938	29.50 / 0.8946	30.76 / 0.9140	- / -	32.93 / 0.9351	32.84 / 0.9347	33.10 / 0.9363	33.03 / 0.9362
	×3	24.46 / 0.7349	26.03 / 0.7973	26.24 / 0.7989	27.14 / 0.8279	- / -	28.80 / 0.8653	28.79 / 0.8655	29.02 / 0.8685	28.99 / 0.8683
	×4	23.14 / 0.6577	24.32 / 0.7183	24.52 / 0.7221	25.18 / 0.7524	26.07 / 0.7839	26.64 / 0.8033	26.67 / 0.8041	26.86 / 0.8080	26.86 / 0.8082
DIV2K validation	×2	31.01 / 0.9393	32.89 / 0.9570	33.05 / 0.9581	33.66 / 0.9625	- / -	35.03 / 0.9695	34.96 / 0.9692	35.12 / 0.9699	35.05 / 0.9696
	×3	28.22 / 0.8906	29.50 / 0.9116	29.64 / 0.9138	30.09 / 0.9208	- / -	31.26 / 0.9340	31.25 / 0.9338	31.39 / 0.9351	31.36 / 0.9346
	×4	26.66 / 0.8521	27.70 / 0.8736	27.78 / 0.8753	28.17 / 0.8841	- / -	29.25 / 0.9017	29.26 / 0.9016	29.38 / 0.9032	29.36 / 0.9029

Set5, Set14, B100, and Urban100

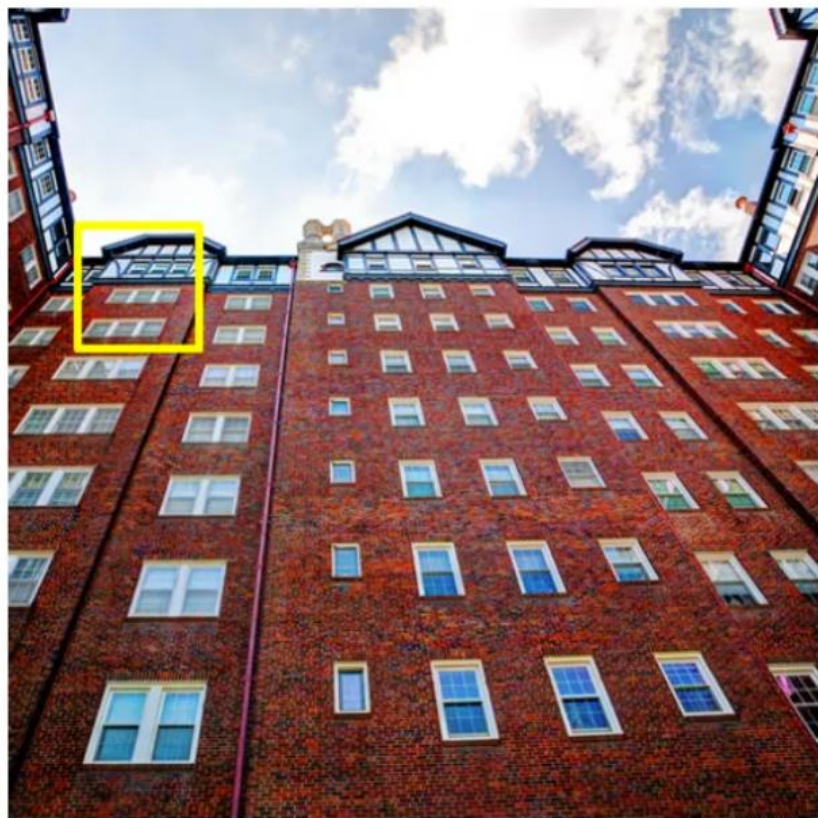


set 5 bicubic_x4



set 14 bicubic_x4

Dataset	Scale	Bicubic	A+ [27]	SRCNN [4]	VDSR [11]	SRResNet [14]	EDSR (Ours)	MDSR (Ours)	EDSR+ (Ours)	MDSR+ (Ours)
Set5	×2	33.66 / 0.9299	36.54 / 0.9544	36.66 / 0.9542	37.53 / 0.9587	- / -	38.11 / 0.9601	38.11 / 0.9602	38.20 / 0.9606	38.17 / 0.9605
	×3	30.39 / 0.8682	32.58 / 0.9088	32.75 / 0.9090	33.66 / 0.9213	- / -	34.65 / 0.9282	34.66 / 0.9280	34.76 / 0.9290	34.77 / 0.9288
	×4	28.42 / 0.8104	30.28 / 0.8603	30.48 / 0.8628	31.35 / 0.8838	32.05 / 0.8910	32.46 / 0.8968	32.50 / 0.8973	32.62 / 0.8984	32.60 / 0.8982
Set14	×2	30.24 / 0.8688	32.28 / 0.9056	32.42 / 0.9063	33.03 / 0.9124	- / -	33.92 / 0.9195	33.85 / 0.9198	34.02 / 0.9204	33.92 / 0.9203
	×3	27.55 / 0.7742	29.13 / 0.8188	29.28 / 0.8209	29.77 / 0.8314	- / -	30.52 / 0.8462	30.44 / 0.8452	30.66 / 0.8481	30.53 / 0.8465
	×4	26.00 / 0.7027	27.32 / 0.7491	27.49 / 0.7503	28.01 / 0.7674	28.53 / 0.7804	28.80 / 0.7876	28.72 / 0.7857	28.94 / 0.7901	28.82 / 0.7876



img034 from Urban100 [10]



HR
(PSNR / SSIM)



Bicubic
(21.41 dB / 0.4810)



A+ [27]
(22.21 dB / 0.5408)



SRCNN [4]
(22.33 dB / 0.5461)



VDSR [11]
(22.62 dB / 0.5657)



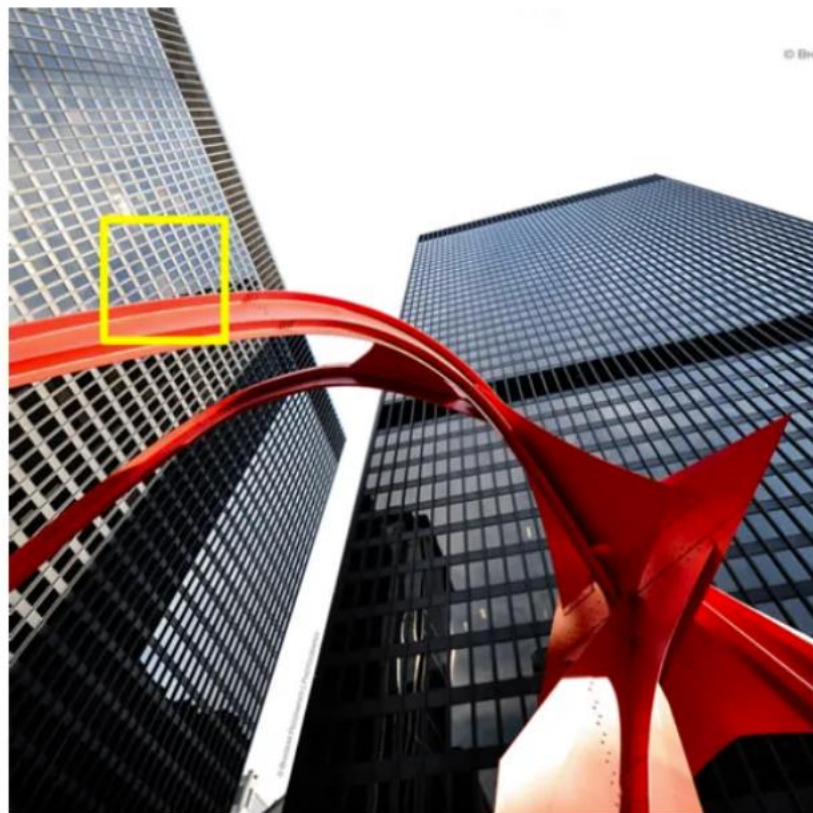
SRResNet [14]
(23.14 dB / 0.5891)



EDSR+ (Ours)
(23.48 dB / 0.6048)



MDSR+ (Ours)
(23.46 dB / 0.6039)



img062 from Urban100 [10]



HR
(PSNR / SSIM)



Bicubic
(19.82 dB / 0.6471)



A+ [27]
(20.43 dB / 0.7145)



SRCNN [4]
(20.61 dB / 0.7218)



VDSR [11]
(20.75 dB / 0.7504)



SRResNet [14]
(21.70 dB / 0.8054)



EDSR+ (Ours)
(22.70 dB / 0.8537)



MDSR+ (Ours)
(22.66 dB / 0.8508)



0869 from DIV2K [26]



HR
(PSNR / SSIM)



Bicubic
(22.66 dB / 0.8025)



A+ [27]
(23.10 dB / 0.8251)



SRCNN [4]
(23.14 dB / 0.8280)



VDSR [11]
(23.36 dB / 0.8365)



SRResNet [14]
(23.71 dB / 0.8485)



EDSR+ (Ours)
(23.89 dB / 0.8563)



MDSR+ (Ours)
(23.90 dB / 0.8558)

Desafio SR NTIRE2017

- O desafio tem como objetivo desenvolver um sistema de super-resolução de imagem única com o mais alto PSNR.
- Existem duas trilhas para diferentes degradadores (bicúbico, desconhecido) com três escalas de downsample ($\times 2$, 3, 4) cada. As imagens de entrada não são apenas reduzidas, mas também sofrem com um grande desfoque.

Method	Track1: bicubic downscailing						Track2: unknown downscailing					
	$\times 2$		$\times 3$		$\times 4$		$\times 2$		$\times 3$		$\times 4$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
EDSR+ (Ours)	34.93	0.948	31.13	0.889	29.09	0.837	34.00	0.934	30.78	0.881	28.77	0.826
MDSR+ (Ours)	34.83	0.947	31.04	0.888	29.04	0.836	33.86	0.932	30.67	0.879	28.62	0.821
3rd method	34.47	0.944	30.77	0.882	28.82	0.830	33.67	0.930	30.51	0.876	28.54	0.819
4th method	34.66	0.946	30.83	0.884	28.83	0.830	32.92	0.921	30.31	0.871	28.14	0.807
5th method	34.29	0.948	30.52	0.889	28.55	0.752	-	-	-	-	-	-

- O EDSR+ e o MDSR+ conquistaram o primeiro e o segundo lugares, respectivamente, com desempenhos excepcionais, conforme mostrado acima.



0791 from DIV2K [26]



HR
(PSNR / SSIM)



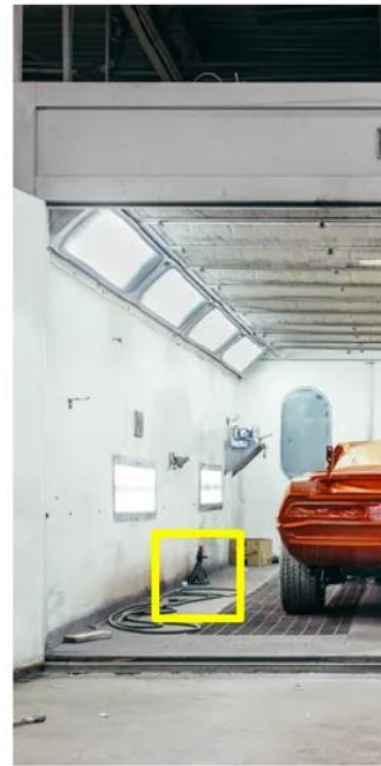
Bicubic
(22.20 dB / 0.7979)



EDSR (Ours)
(29.05 dB / 0.9257)



MDSR (Ours)
(28.96 dB / 0.9244)



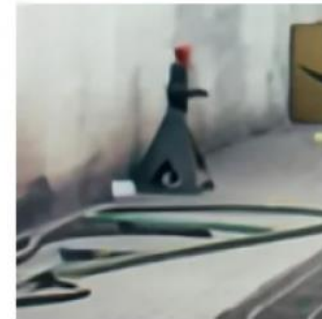
0792 from DIV2K [26]



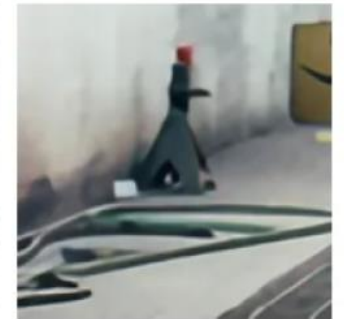
HR
(PSNR / SSIM)



Bicubic
(21.59 dB / 0.6846)



EDSR (Ours)
(27.24 dB / 0.8376)



MDSR (Ours)
(27.14 dB / 0.8356)



0793 from DIV2K [26]



HR
(PSNR / SSIM)



Bicubic
(23.81 dB / 0.8053)



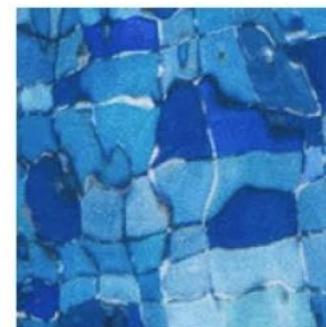
EDSR (Ours)
(30.94 dB / 0.9318)



MDSR (Ours)
(30.81 dB / 0.9301)



0797 from DIV2K [26]



HR
(PSNR / SSIM)



Bicubic
(19.77 dB / 0.8937)



EDSR (Ours)
(25.48 dB / 0.9597)



MDSR (Ours)
(25.38 dB / 0.9590)

Vantagens

- **Alta qualidade de imagem:**

- O EDSR é capaz de gerar imagens de alta resolução com detalhes mais nítidos e claros. Sua arquitetura otimizada permite resultados superiores em termos de qualidade visual.

- **Tempo de processamento eficiente:**

- Apesar de sua profundidade, o EDSR é eficiente em termos de tempo de processamento. Ele pode gerar imagens de alta resolução rapidamente, o que é benéfico para aplicativos em tempo real.

- **Escalabilidade:**

- O modelo EDSR pode se adaptar a diferentes fatores de escala sem perder a qualidade. Isso o torna versátil para várias tarefas de super-resolução.

- **Eliminação de módulos desnecessários:**

- O EDSR otimiza a arquitetura com base na rede residual (ResNet), analisando e eliminando módulos desnecessários. Isso simplifica a rede e melhora seu desempenho.

- **Desempenho superior ao de abordagens anteriores:**

- Em comparação com outros modelos, como SRResNet, VDSR e SRCNN, o EDSR demonstrou desempenho superior. Ele venceu o desafio de super-resolução do NTIRE2017.

- **Auto-montagem:**

- O EDSR usa um método de automontagem que não requer o treinamento de modelos separados. Isso é vantajoso em termos de tamanho do modelo e tempo de treinamento.



Desventajas

- **Necesidad de Grandes Conjuntos de Datos:**

Para obtener resultados óptimos, EDSR generalmente requiere grandes conjuntos de datos de imágenes de alta y baja resolución, lo que puede ser difícil de obtener en algunas aplicaciones.

- **Sensibilidad a la Variación de Hiperparámetros:**

EDSR puede ser sensible a la selección de hiperparámetros durante el entrenamiento, lo que puede requerir un ajuste cuidadoso para obtener los mejores resultados.

- **Dificultad en Escenarios de Texturas Homogéneas:**

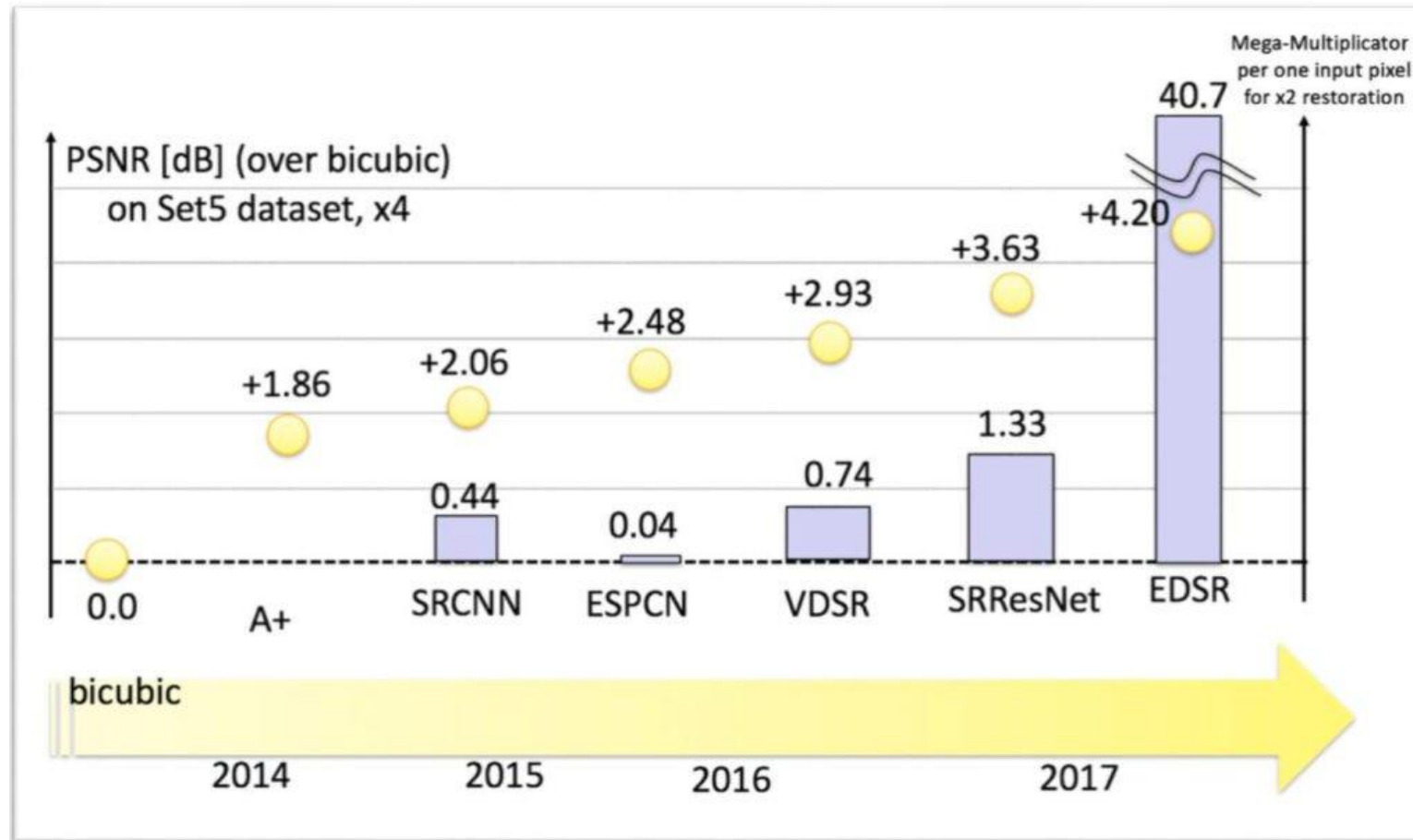
En escenarios donde las texturas son homogéneas o la información visual es limitada, EDSR puede tener dificultades para mejorar la calidad de la imagen de manera significativa.

- **Posible Sobreajuste:**

Como con cualquier modelo de aprendizaje automático profundo, existe el riesgo de sobreajuste al entrenar EDSR en conjuntos de datos limitados o poco representativos



- Ao considerar essas vantagens e desvantagens, é importante avaliar cuidadosamente se a EDSR é a abordagem certa para uma aplicação específica e se seus benefícios superam suas possíveis limitações nesse contexto específico.



Exemplos de aplicação

- Na medicina, a EDSR é usada para melhorar a resolução de imagens médicas, facilitando um diagnóstico mais preciso.
- No campo da fotografia, a EDSR pode restaurar imagens antigas e melhorar a qualidade das fotografias digitais.
- Ela pode ser usada para melhorar a qualidade das imagens em tempo real.
- Imagens de satélite que também exigem que você dê zoom em torno de um objeto específico,
- Câmera de vigilância de vídeo quando o rosto de alguém está desfocado etc.



Perguntas?

Referências

- Lim, B., Son, S., Kim, H., Nah, S., & Lee, K. (n.d.). *Enhanced Deep Residual Networks for Single Image Super-Resolution*. <https://arxiv.org/pdf/1707.02921.pdf>
- C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. arXiv:1609.04802, 2016. 1, 2, 3, 4, 5, 6, 7
- Anwar, S., Khan, S., & Barnes, N. (n.d.). *A Deep Journey into Super-resolution: A Survey*. <https://arxiv.org/pdf/1904.07523.pdf>
- Dong, C., Loy, C., He, K., & Tang, X. (n.d.). *Image Super-Resolution Using Deep Convolutional Networks*. <https://arxiv.org/pdf/1501.00092.pdf>
- Kim, J., Lee, J., & Lee, K. (2016). *Accurate Image Super-Resolution Using Very Deep Convolutional Networks*. <https://arxiv.org/pdf/1511.04587.pdf>
- Dong, C., Loy, C., & Tang, X. (n.d.). *Accelerating the Super-Resolution Convolutional Neural Network*. <https://arxiv.org/pdf/1608.00367.pdf>
- Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A., Bishop, R., Rueckert, D., Wang, Z., & Twitter. (n.d.). *Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network*. <https://arxiv.org/pdf/1609.05158.pdf>
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi Twitter, W. (n.d.). *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*. <https://arxiv.org/pdf/1609.04802.pdf>

QUIZ

<https://forms.gle/aoRDdjuvKPMZVw5B8>

Obrigado!