



Análise de funções de perda em redes convolucionais para super-resolução de ilustrações

Raphael Nepomuceno e Michel Melo da Silva - {raphael.nepomuceno, michel.m.silva}@ufv.br
Departamento de Informática (DPI-UFV)

Área temática: Inteligência Artificial | Grande área: Ciências Exatas e Tecnológicas | Categoria: Pesquisa

Palavras Chaves: Redes neurais convolucionais, Aprendizado profundo, Super-resolução

Introdução

Super-resolução de imagens é um problema tradicional em visão computacional. Contudo, poucos trabalhos focam em ilustrações. Nosso objetivo é estudar o impacto das funções de perda das redes neurais nesta tarefa no domínio de ilustrações.



Figura 1 – Comparação entre a imagem original (1ª), cuja versão miniaturizada foi então interpolada pelo método bicúbico (2ª) e por uma rede neural (3ª).

Objetivos

- Determinar qual função de perda produz os melhores resultados em super-resolução em imagens ilustradas;
- Otimizar nitidez de bordas e ausência de ruídos;
- Obter fidelidade à imagem original (cores e formas).

Material e Métodos

O trabalho foi conduzido utilizando três conjuntos de dados de imagens ilustradas: Danbooru2020, SYNLA e Manga109.

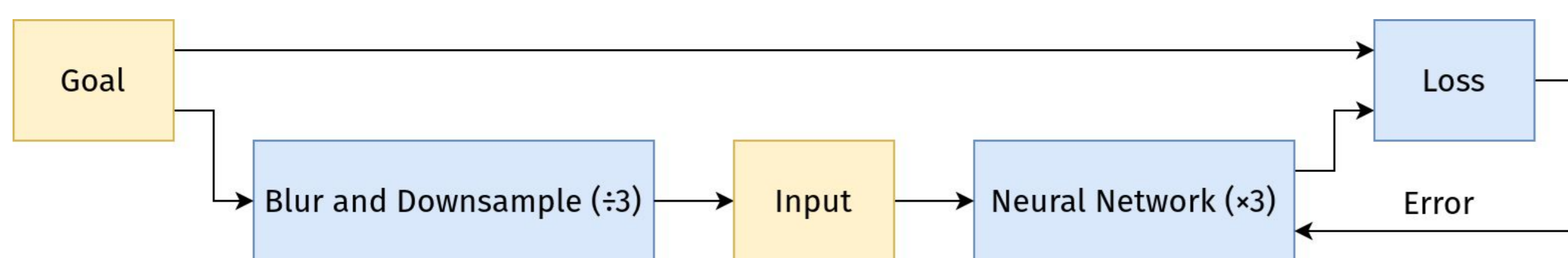


Figura 2 – Processo de treinamento da rede. Fonte: Os autores.

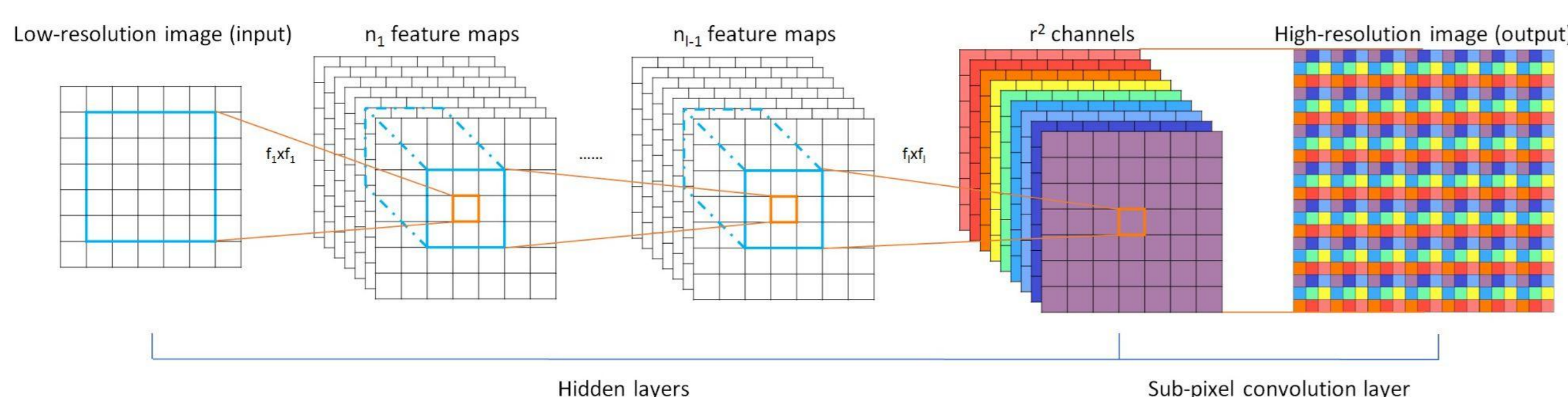


Figura 3 – Arquitetura da rede neural ESPCN. Fonte: Shi et al. (2016)

Bibliografia

- ZHAO, Hang et al. Loss functions for image restoration with neural networks. IEEE Transactions on computational imaging, IEEE, v. 3, n. 1, p. 47–57, 2016.
- SHI, Wenzhe et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: PROCEEDINGS of the IEEE conference on Computer Vision and Pattern Recognition. [S.l.: s.n.], 2016. P. 1874–1883.
- LU, Zhengyang; CHEN, Ying. Single Image Super Resolution based on a Modified U-net with Mixed Gradient Loss. [S.l.: s.n.], 2019. arXiv: 1911.09428 [eess.IV].

Resultados e Discussão

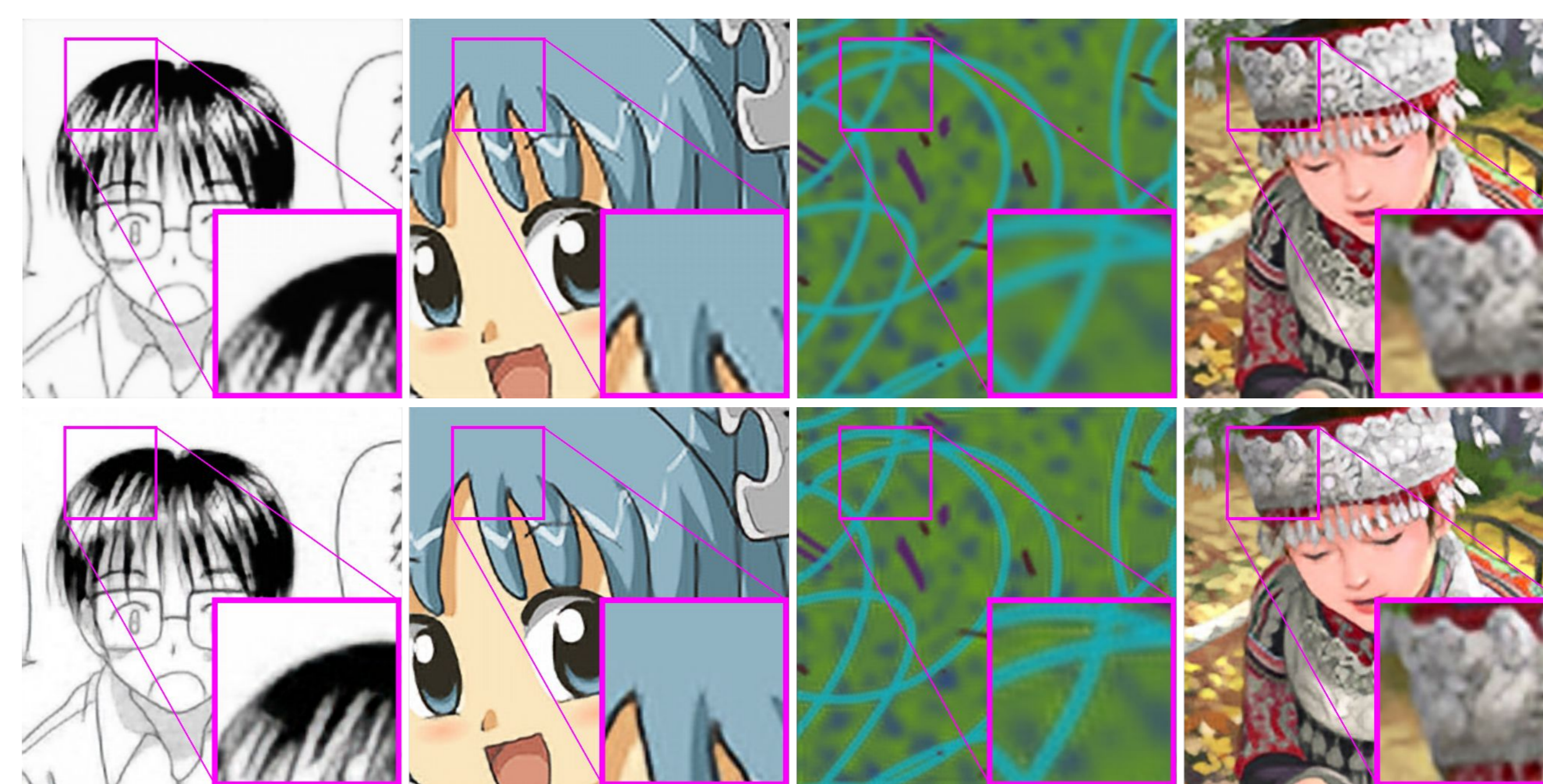


Figura 4 – Comparação entre o resultado do baseline e o melhor resultado encontrado. Primeira linha: *Baseline*. Segunda linha: Método proposto usando MixGE.

A função de perda utilizando detecção de bordas MixGE obteve os melhores resultados na maioria das métricas, vide Tabela 1. Contudo, em imagens desfocadas, houve grande incidência de ruído, sendo um caso patológico em que seu desempenho foi pior que o *Baseline* (Figura 4, coluna 3).

Methods	Danbooru2020		Manga109		SYNLA (Color)	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bilinear	21.42	0.7806	17.52	0.6922	21.19	0.7475
Bicubic	21.97	0.8026	18.01	0.7190	22.29	0.7977
Baseline	23.08	0.8274	19.31	0.7726	22.50	0.7638
Baseline-RGB	22.47	0.7878	18.73	0.7328	22.39	0.7660
Ours (MSE)	23.99	0.8514	20.12	0.8040	23.14	0.7816
Ours (MAE)	23.14	0.8402	19.07	0.7751	22.58	0.7831
Ours (DSSIM)	23.12	0.8600	19.34	0.7949	22.37	0.8013
Ours (MixGE, $\lambda_G = 0.01$)	24.61	0.8708	20.62	0.8235	23.30	0.7864
Ours (MixGE, $\lambda_G = 0.10$)	24.62	0.8700	20.65	0.8216	23.22	0.7858
Ours (MixGE, $\lambda_G = 1.00$)	24.66	0.8707	20.63	0.8231	23.24	0.7839

Tabela 1 – Avaliação quantitativa das funções de perda. Valores maiores são melhores

Conclusões

- Observamos melhorias ao aplicar conhecimentos do domínio de interesse na função de perda do modelo.
- **Trabalhos Futuros:** verificar se as características de qualidade e as relações das funções de perda são mantidas se usadas em uma arquitetura de rede maior que a da Fig 2.

Agradecimentos

Os autores agradecem ao CNPq, CAPES e FAPEMIG.

Fontes das imagens da Figura 4 (Modificadas; acesso em 23 de Setembro de 2021):

1. Manga109, © Ken Akamatsu.
2. https://en.wikipedia.org/wiki/File:Wikipet-an_face.svg
3. <https://github.com/bloc97/SYNLA-Dataset>
4. <https://git.io/JzVV0>