

Restauração de imagens coloridas usando uma rede neural convolucional residual

Guilherme A. Lizarzaburu, Renato Candido e Magno T. M. Silva

Resumo—A restauração de imagens tem por objetivo atenuar distorções causadas no processo de aquisição. Neste artigo, utiliza-se uma rede neural convolucional residual para restauração de imagens coloridas, degradadas por uma função de espalhamento de ponto gaussiana. Como função custo, foram considerados o erro quadrático médio e uma função do índice de similaridade estrutural (*structural similarity* – SSIM). Por meio de simulações, verifica-se que os melhores resultados de restauração foram obtidos com a função custo baseada no SSIM.

Palavras-Chave—Rede neural convolucional residual, restauração de imagens, função de espalhamento de ponto gaussiana, índice de similaridade estrutural.

Abstract—Image restoration aims to attenuate distortions caused in the acquisition process. In this paper, we use a residual convolutional neural network to restore colored images which were degraded by a Gaussian point spread function. As cost function, we consider the mean square error, and a function based on the structural similarity (SSIM) index. By means of simulations, we verify that the best restoration results were achieved with the function based on SSIM.

Keywords—Convolutional neural network with residual skip connections, image restoration, gaussian point spread function, structural similarity index.

I. INTRODUÇÃO E FORMULAÇÃO DO PROBLEMA

No processo de aquisição, as imagens digitais podem sofrer degradações, que em geral são modeladas por uma função de espalhamento de ponto (*point spread function* - PSF) [5]. Um dos modelos de distorção mais usados na literatura é a PSF gaussiana de dimensão $(2N_1 + 1) \times (2N_2 + 1)$ em que

$$\mathbf{H}(n_1, n_2) = (2\pi\sigma^2)^{-1} \exp(-(n_1^2 + n_2^2)/(2\sigma^2)),$$

sendo σ o desvio padrão, $n_1 = -N_1, \dots, -1, 0, 1, \dots, N_1$ e $n_2 = -N_2, \dots, -1, 0, 1, \dots, N_2$. A imagem degradada é obtida a partir da convolução em duas dimensões entre a imagem original e a PSF. No caso de imagens coloridas, compostas por canais de cores vermelho, verde e azul, cada canal da imagem degradada é fruto da convolução entre a PSF e o respectivo canal da imagem original.

Para melhorar a qualidade de uma imagem degradada, o efeito da PSF deve ser atenuado, o que pode ser feito com técnicas de aprendizado de máquina [2], [3], [6], [8]. Neste contexto, a rede neural convolucional (*convolutional neural network* - CNN) tem ganhado destaque na literatura [3]. No entanto, a maior parte dos trabalhos que usam a CNN para restauração de imagens, considera o erro quadrático

médio (*mean square error* – MSE) como função custo no treinamento. O MSE não é uma medida adequada quando se trata de imagens já que ele mede erros absolutos, não levando em conta características do sistema visual humano [7], [8]. No lugar do MSE, é mais adequado usar o índice de similaridade estrutural (*structural similarity* – SSIM), proposto em [7]. O SSIM mede a similaridade entre duas imagens a partir da combinação de três medidas distintas referentes à luminância, ao contraste e à estrutura das imagens [7]. Esse índice assume valores no intervalo $[-1, 1]$, sendo igual a um quando as duas imagens são iguais.

Neste trabalho, considera-se uma CNN residual [3], [4] para restauração de imagens coloridas degradadas por uma PSF gaussiana. Como função custo, considera-se uma função do índice SSIM. Os resultados obtidos são então comparados com a mesma rede treinada para minimizar o MSE.

II. ARQUITETURA DA CNN PROPOSTA

A CNN residual proposta é esquematizada na Figura 1. Ela é composta por 10 camadas convolucionais, tendo a tangente hiperbólica como função de ativação. As dimensões e quantidades dos filtros por camada estão indicadas na figura. Cabe lembrar que cada camada convolucional realiza a operação de convolução entre seus filtros e o tensor de entrada, que por sua vez, passa pelo processo de *zero padding* para adequação das dimensões [3]. A quantidade de filtros por camada na CNN proposta aumenta gradativamente de 3 para 32 da primeira à quinta camada, respectivamente. Isso faz com que o tensor de saída da quinta camada tenha dimensão $256 \times 256 \times 32$, o que possibilita extrair diferentes características da imagem degradada, facilitando o processo de restauração. Da quinta à décima camada, a quantidade de filtros por camada diminui de 32 para 3, respectivamente. Assim, o tensor de saída da última camada é a imagem restaurada com dimensão $256 \times 256 \times 3$, exatamente igual à dimensão da imagem degradada. Além disso, considerou-se uma rede do tipo residual, em que existem atalhos entre camadas não consecutivas a fim de otimizar o processo de aprendizado [4]. No caso da CNN proposta, como esquematizado na Figura 1, a entrada da décima camada é dada pela soma das saídas da nona com a da primeira camada, a entrada da nona camada é dada pela soma das saídas da oitava com a da segunda camada e assim sucessivamente.

Os coeficientes dos filtros de cada camada são adaptados no treinamento utilizando o algoritmo de retropropagação (*backpropagation*) de modo a minimizar uma função custo [3]. Além do MSE, a rede proposta foi treinada para minimizar a função custo $J = 1 - \text{SSIM}(\mathbf{Y}, \mathbf{D})$, em que \mathbf{Y} é a imagem restaurada e \mathbf{D} é a imagem original [8]. Para isso, utilizou-se

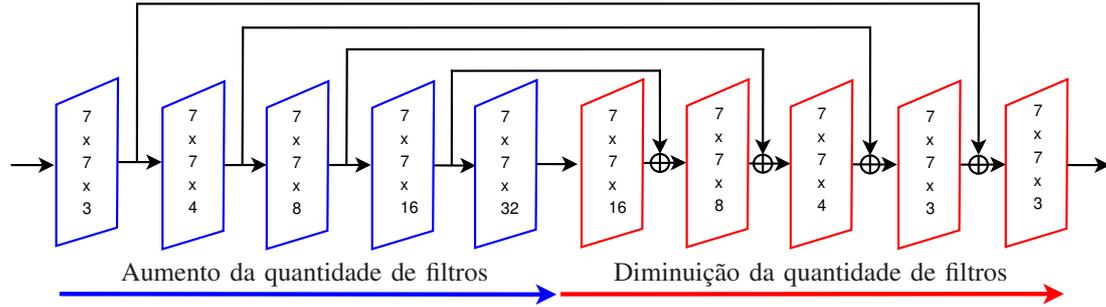


Fig. 1: Esquema da CNN residual utilizada.

a diferenciação automática (*autodiff*), que é um conjunto de técnicas usadas para avaliar derivadas de funções numéricas expressas como programas de computador [1].

III. RESULTADOS DE SIMULAÇÃO

No treinamento da rede, foram utilizadas 29 imagens coloridas degradadas pela PSF gaussiana com $N_1 = N_2 = 3$ e três valores diferentes para σ : 1, 3 e 5. Os pixels de cada canal foram truncados para valores inteiros no intervalo $[0, 255]$.

Depois de treinada por 2×10^4 épocas, a rede foi testada com duas imagens não pertencentes ao conjunto de treinamento e degradadas por PSFs gaussianas com $N_1 = N_2 = 3$ e dois valores diferentes de σ : $\sigma = 5$ considerado nas degradações do conjunto de treinamento e $\sigma = 4$ não utilizado no treinamento. Na Figura 2, são mostrados os resultados de restauração considerando duas versões da CNN residual proposta: uma treinada para minimizar o MSE e outra para minimizar a função do SSIM. É possível notar que a rede que utiliza o SSIM na função custo proporciona uma restauração melhor em termos do SSIM medido. Esse resultado pode ser observado visualmente, com as cores mais precisas no rosto mostrado na Fig. 2-(d) e a menor quantidade de manchas em torno da palmeira mostrada na Fig. 2-(h).

IV. CONCLUSÕES

A CNN residual proposta foi capaz de atenuar as distorções das duas imagens degradadas do teste, independente da função custo considerada no treinamento. No entanto, cabe ressaltar o desempenho superior da rede treinada com a função custo baseada no SSIM quando comparada com a rede treinada com o MSE. Em um trabalho futuro, pretende-se comparar os resultados obtidos com os de uma rede generativa adversária.

REFERÊNCIAS

- [1] A. G. Baydin *et al.*, “Automatic differentiation in machine learning: a survey,” *J. Machine Learning Research*, v. 18, p. 1–43, 2018.
- [2] S. A. Bigdeli and M. Zwicker, “Image restoration using autoencoding priors,” in *13th Int. Joint Conf. Comp. Vision, Imaging and Comp. Graphics Theory and Appl. (VISIGRAPP)*, p. 33–44, 2018.
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [5] D. Kundur and D. Hatzinakos, “Blind image deconvolution,” *IEEE Signal Process. Mag.*, v. 13, p. 43–46, 1996.
- [6] X. Mao, S. Chunhua, and Y. Yu-Bin, “Image restoration using very deep convolutional enc.-dec. networks w/ symmetric skip connections,” *Adv. Neural Inf. Process. Syst.*, p. 2802–2810, 2016.
- [7] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.*, v. 13, p. 600–612, 2004.
- [8] H. Zhao *et al.*, “Loss functions for image restoration with neural networks,” *IEEE Trans. Computational Imaging*, v. 3, p. 47–57, 2016.

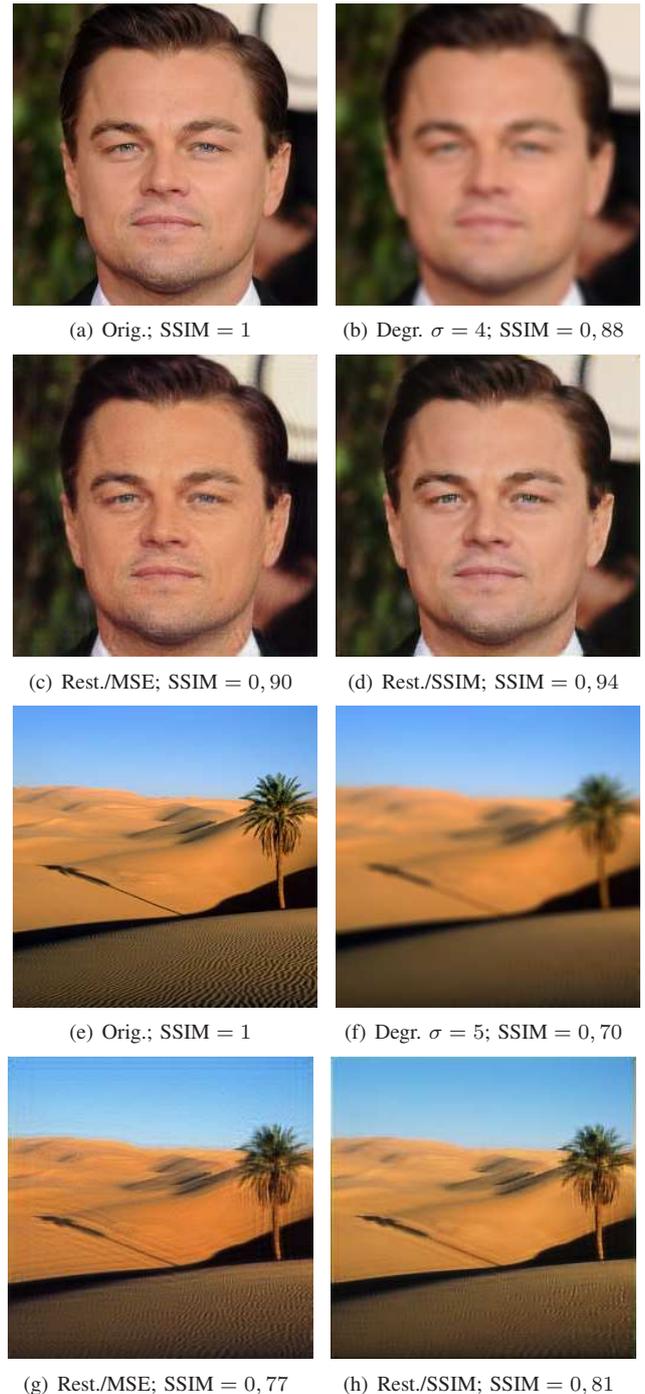


Fig. 2: Exemplos de duas imagens restauradas pela CNN residual proposta usando MSE ou função do SSIM como função custo; SSIM calculado em relação à imagem original.