

Restaurando imagens com uma rede neural convolucional

Guilherme A. Lizarzaburu¹

Renato Candido²

Magno T. M. Silva³

Escola Politécnica da USP, São Paulo, SP

Soluções para restauração de imagens com técnicas de aprendizado de máquina têm sido propostas na literatura [2,6,7,9]. Neste trabalho, apresentam-se resultados de restauração de imagens obtidos com uma rede neural convolucional (*convolutional neural network* - CNN) residual [3,4].

No experimento, foram consideradas imagens em 256 níveis de cinza com resolução 256×256 . As imagens foram degradadas por uma função de espalhamento de ponto (*point spread function* - PSF) gaussiana [5], definida pela matriz $\mathbf{H}(x, y) = (2\pi\sigma^2)^{-1} \exp[-(x^2 + y^2)/(2\sigma^2)]$, em que σ^2 representa a variância da PSF, e x e y são as distâncias horizontal e vertical do centro da matriz, respectivamente. A imagem degradada é obtida a partir da convolução em duas dimensões entre a imagem original e a PSF. No treinamento da rede, foram utilizadas 20 imagens degradadas pela PSF gaussiana de dimensão 7×7 com $\sigma = 1$ e 20 imagens degradadas pela PSF gaussiana de dimensão 7×7 com $\sigma = 3$. Para emular uma situação realista, os *pixels* das imagens degradadas foram arredondados para valores inteiros no intervalo $[0, 255]$.

Para medir a similaridade entre as imagens, considerou-se o índice de similaridade estrutural (*Structural Similarity* - SSIM) [8,9], que assume valores no intervalo $[-1, 1]$, sendo igual a um quando as duas imagens são iguais. Assim, utilizou-se a função custo $J = 1 - \text{SSIM}(\mathbf{Y}, \mathbf{D})$ no treinamento da rede, em que \mathbf{Y} é a imagem restaurada e \mathbf{D} é a imagem original. Para isso, utilizou-se a diferenciação automática (*autodiff*), que é um conjunto de técnicas usadas para avaliar derivadas de funções expressas por programas de computador. Independentemente da complexidade da função, o sistema de diferenciação automática converte o programa em uma sequência de operações primitivas (adição, subtração, multiplicação, divisão, entre outras) e funções elementares (seno, cosseno, exponencial natural, entre outras), que possuem rotinas predefinidas para o cálculo de suas respectivas derivadas e sobre as quais é aplicada a regra da cadeia para o cálculo da derivada da função do programa [1].

A CNN proposta tem as dimensões de sua saída idênticas às dimensões de sua entrada, em que a entrada é a imagem degradada e a saída a imagem restaurada. A rede testada está esquematizada na Figura 1. Ela é composta por 5 camadas convolucionais com quantidades crescentes de filtros seguidas por 5 camadas convolucionais com quantidades decrescentes de filtros, tendo a tangente hiperbólica como função de ativação. O tensor de saída de cada camada passa pelo processo de *zero padding* para manter a resolução 256×256 [3]. As dimensões e quantidades de filtros por camada estão indicadas na figura. Além disso, considerou-se uma rede do tipo residual, que possui ligações entre camadas não consecutivas a fim de otimizar o processo de aprendizado [4]. No caso, soma-se a saída da primeira camada à entrada da última camada, a saída da segunda camada à entrada da penúltima camada e assim por diante.

Na Figura 2, é mostrada uma restauração realizada pela rede após 20000 épocas de treinamento. A imagem considerada foi degradada pela PSF gaussiana de dimensão 7×7 e com $\sigma = 2$. Tanto

¹guilherme.lizarzaburu@usp.br; bolsista FAPESP (2019/08636-8)

²renatocan@lps.usp.br

³magno.silva@usp.br; FAPESP (2017/20378-9), CNPq (304715/2017-4) e CAPES (código de Financiamento 001).

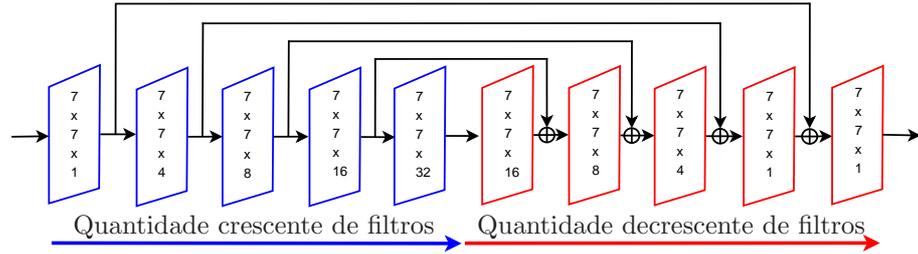


Figura 1: Esquema da CNN residual utilizada; quantidade de filtros 7×7 indicada.

a imagem quando o valor de σ não foram utilizados no treinamento. É possível notar que a rede testada foi capaz de atenuar as distorções da imagem degradada, chegando a um SSIM de 0,92. Em trabalhos futuros, pretende-se considerar uma mistura de especialistas baseada na rede proposta.

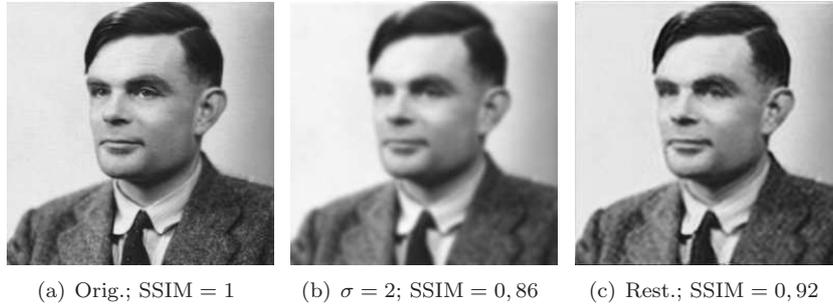


Figura 2: Da esquerda para a direita: imagem original, imagem degradada e imagem restaurada. Os índices SSIM foram calculados em relação à imagem original.

Referências

- [1] Baydin, A. G., Pearlmutter, B. A., Radul, A. A. and Siskind, J. M. Automatic differentiation in machine learning: a survey, *J. Machine Learning Research*, 18: 5595–5637, 2017. DOI: 10.5555/3122009.3242010.
- [2] Bigdeli, S. A. and Zwicker, M. Image restoration using autoencoding priors, *13th Int. Joint Conf. Comp. Vision, Imaging and Comp. Graphics Theory and Appl.*, 5: 33–44, 2018. DOI: 10.5220/0006532100330044.
- [3] Goodfellow, I., Bengio, Y. and Courville, A. *Deep Learning*. MIT Press, 2016.
- [4] He, K., Zhang, X., Ren, S. and Sun, J. Deep residual learning for image recognition, *IEEE Conf. Computer Vision and Pattern Recognition*, 1: 770–778, 2016. DOI: 10.1109/CVPR.2016.90.
- [5] Kundur, D. and Hatzinakos, D. Blind image deconvolution, *IEEE Signal Process. Mag.*, 13: 43–46, 1996. DOI: 10.1109/79.489268.
- [6] Lizarzaburu, G. A., Candido, R. and Silva, M. T. M. Restauração de imagens coloridas usando uma rede neural convolucionar residual, *Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBRT)*, 2020. DOI: 10.14209/SBRT.2020.1570649730.
- [7] Mao, X., Chunhua, S. and Yu-Bin, Y. Image restoration using very deep convolutional enc.-dec. networks w/ symmetric skip connections, *30th Int. Conf. on Neural Info. Processing Systems*, 2810–2818, 2016. DOI: 10.5555/3157382.3157412.
- [8] Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.*, 13: 600–612, 2004. DOI: 10.1109/TIP.2003.819861.
- [9] Zhao, H. Loss functions for image restoration with neural networks. *IEEE Trans. Computational Imaging*, 3: 47–57, 2016. DOI: 10.1109/TCI.2016.2644865.