



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Compressione di immagini tramite autoencoder

stato dell'arte e sviluppi futuri

F. Stella M. Cagnazzo

16 Novembre 2023

1. Compressione
2. Metodi tradizionali
3. Metodi con intelligenza artificiale
4. Risultati Sperimentali
5. Sviluppi Futuri
6. Bibliografia

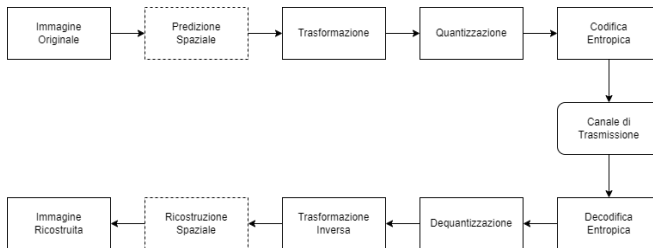


Figure: Blocchi funzionali principali di un framework lossy [1], con l'aggiunta di un quarto blocco per i metodi recenti

I metodi di codifica tradizionale analizzati in questo studio sono i seguenti

- JPEG [2]
- JPEG2000 [3]
- BPG [4]
- VVC [5]

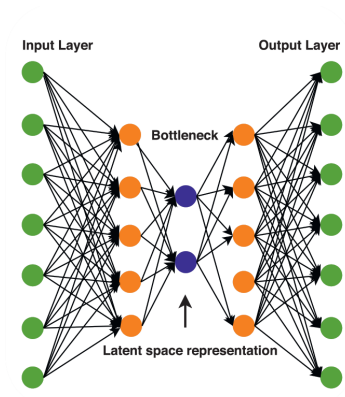


Figure: Schema generico di un autoencoder, immagine presa dal documento [6]

I metodi di codifica con intelligenza artificiale analizzati in questo studio sono i seguenti

- Ballé et al. [7]
- Cheng et al. [8]
- Wang et al. [9]

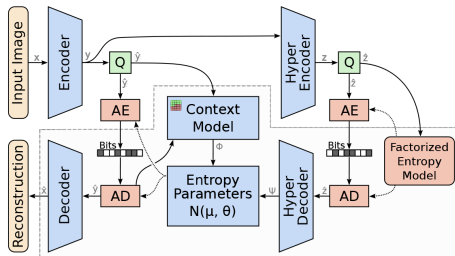


Figure: Diagramma rete Ballé 2018 et al., immagine presa dal documento [7]

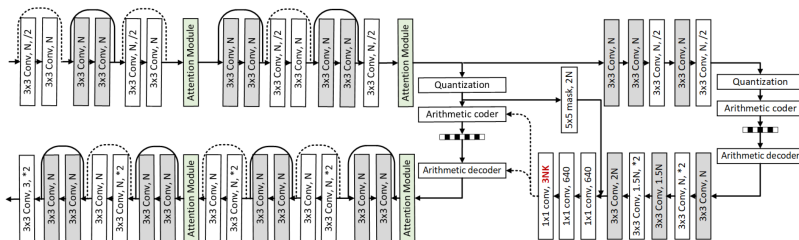


Figure: Diagramma rete Cheng 2020 et al., immagine presa dal documento [8]

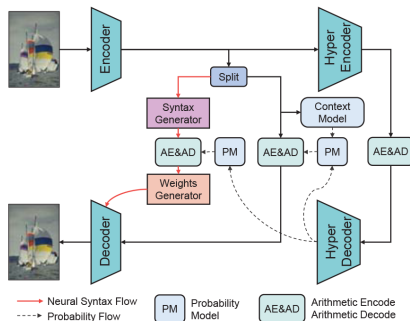


Figure: Diagramma rete Wang 2022 et al., immagine presa dal documento [9]

Esempi di compressione di un immagine del dataset Kodak [10]
con le tecniche presentate



Figure:	Figure:	Figure:	Figure:	Figure:	Figure:	Figure:
Originale	JPEG	JPEG2000	BPG	VVC	Ballé	Cheng
11.117bpp	0.167bpp	0.137bpp	0.103bpp	0.061bpp	0.060bpp	0.056bpp

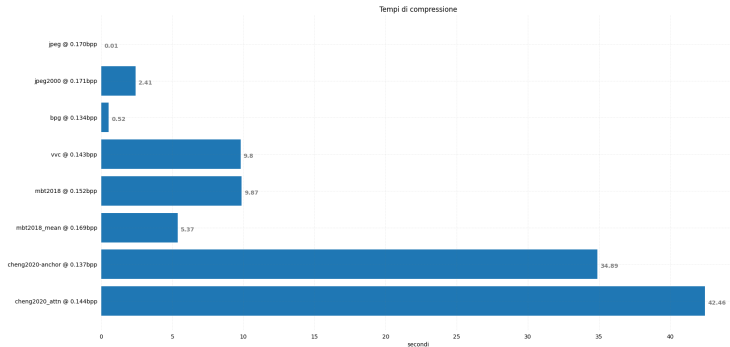


Figure: Tempi di compressione a 0.16 bpp

Tempi di Compressione

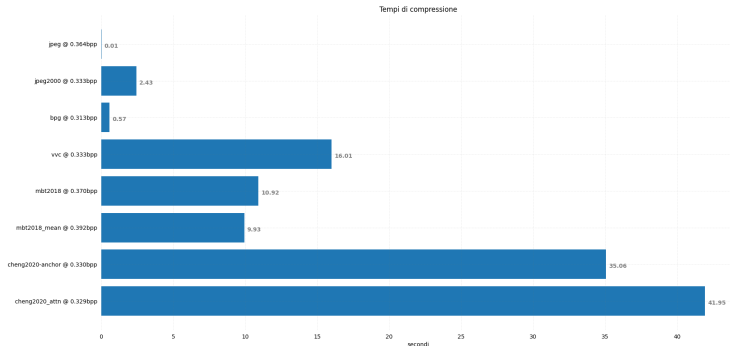


Figure: Tempi di compressione a 0.34 bpp

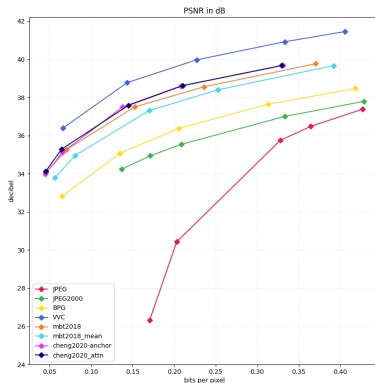


Figure: Grafico del PSNR, punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

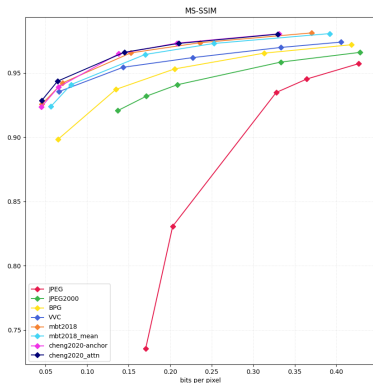


Figure: Grafico dell'MS-SSIM [11], punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

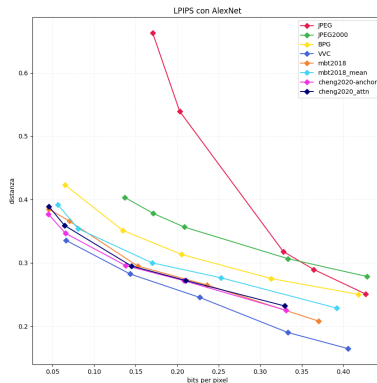


Figure: Grafico LPIPS [12] con AlexNet, punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

Durante la ricerca delle informazioni per la stesura di questa tesi ci siamo imbattuti in due lavori molto interessanti

- StructuralADAM [13]
- SmallCAE [14]

- [1] H. T. Sadeeq, T. H. Hameed, A. S. Abdi, and A. N. Abdulfatah, "Image compression using neural networks: A review," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, no. 14, pp. 135–153, 2021.
- [2] G. Wallace, "The jpeg still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992. DOI: 10.1109/30.125072.
- [3] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The jpeg 2000 still image compression standard," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 36–58, 2001. DOI: 10.1109/79.952804.

- [4] F. Bellard, *Bpg image format*,
<https://bellard.org/bpg/>, Consultato: 17-10-2023.
- [5] B. Bross, Y.-K. Wang, Y. Ye, *et al.*, "Overview of the versatile video coding (vvc) standard and its applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 10, pp. 3736–3764, 2021. DOI: 10.1109/TCSVT.2021.3101953.
- [6] D. Mishra, S. K. Singh, and R. K. Singh, "Deep architectures for image compression: A critical review," *Signal Processing*, vol. 191, p. 108346, 2022.
- [7] D. Minnen, J. Ballé, and G. D. Toderici, "Joint autoregressive and hierarchical priors for learned image compression," *Advances in neural information processing systems*, vol. 31, 2018.

- [8] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, “Learned image compression with discretized gaussian mixture likelihoods and attention modules,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 7939–7948.
- [9] D. Wang, W. Yang, Y. Hu, and J. Liu, “Neural data-dependent transform for learned image compression,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 17 379–17 388.
- [10] E. K. Company, *True color kodak images*, <https://r0k.us/graphics/kodak/>, Consultato: 17-10-2023.

- [11] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, IEEE, vol. 2, 2003, pp. 1398–1402.
- [12] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586–595.
- [13] J. Ballé, "Efficient nonlinear transforms for lossy image compression," in *2018 Picture Coding Symposium (PCS)*, IEEE, 2018, pp. 248–252.

- [14] F. Yang, L. Herranz, Y. Cheng, and M. G. Mozerov, "Slimmable compressive autoencoders for practical neural image compression," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 4998–5007.



Grazie per la vostra attenzione