



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO  
DI INGEGNERIA  
DELL'INFORMAZIONE

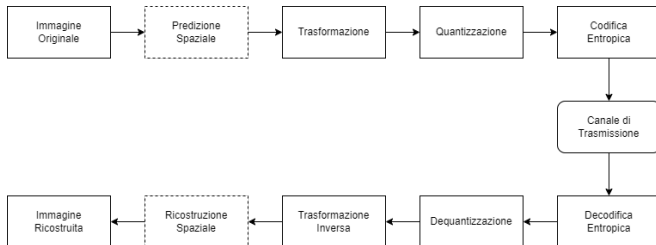
# Compressione di immagini tramite autoencoder

stato dell'arte e sviluppi futuri

Laureando: F. Stella    Relatore: M. Cagnazzo

16 Novembre 2023

1. Compressione
2. Metodi tradizionali
3. Metodi con reti neurali
4. Risultati Sperimentali
5. Sviluppi Futuri



**Figure:** Blocchi funzionali principali di un framework lossy<sup>1</sup>, con l'aggiunta di un quarto blocco per i metodi recenti

<sup>1</sup>H. T. Sadeeq, T. H. Hameed, A. S. Abdi, *et al.*, "Image compression using neural networks: A review," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, no. 14, pp. 135–153, 2021

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

- JPEG<sup>1</sup>
- JPEG 2000<sup>2</sup>
- BPG<sup>3</sup>
- VVC<sup>4</sup>

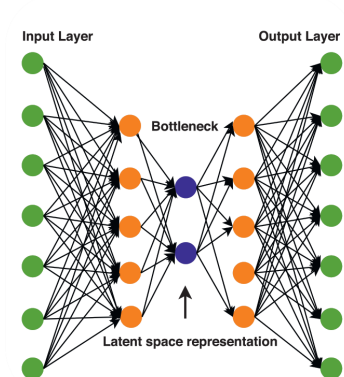
---

<sup>1</sup>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

<sup>2</sup>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

<sup>3</sup>F. Bellard, *BPG image format*, <https://bellard.org/bpg/>, Consultato: 17-10-2023

<sup>4</sup>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



**Figure:** Schema generico di un autoencoder, immagine presa dal documento *Deep architectures for image compression: A critical review*<sup>1</sup>

<sup>1</sup>D. Mishra, S. K. Singh, and R. K. Singh, "Deep architectures for image compression: A critical review," *Signal Processing*, vol. 191, p. 108 346, 2022

I metodi di codifica con reti neurali analizzati in questo studio sono i seguenti

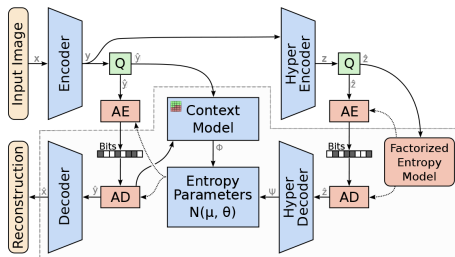
- Variational compression with a scale hyperprior, Ballé et al<sup>1</sup>.
- Discretized gaussian mixture likelihoods, Cheng et al<sup>2</sup>.
- Neural data-dependent transform, Wang et al<sup>3</sup>.

---

<sup>1</sup>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

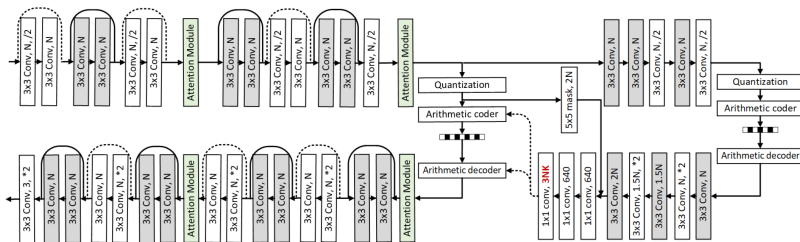
<sup>2</sup>Z. Cheng, H. Sun, M. Takeuchi, et al., "Learned image compression with discretized gaussian mixture likelihoods and attention modules," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 7939–7948

<sup>3</sup>D. Wang, W. Yang, Y. Hu, et al., "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



**Figure:** Diagramma rete Ballé 2018 et al., immagine presa dal documento *Joint autoregressive and hierarchical priors for learned image compression*<sup>1</sup>

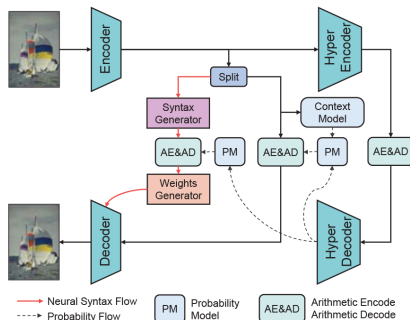
<sup>1</sup>D. Minnen, J. Ballé, and G. D. Toderici, "Joint autoregressive and hierarchical priors for learned image compression," *Advances in neural information processing systems*, vol. 31, 2018



**Figure:** Diagramma rete Cheng 2020 et al., immagine presa dal documento *Learned image compression with discretized gaussian mixture likelihoods and attention modules*<sup>1</sup>

<sup>1</sup>Z. Cheng, H. Sun, M. Takeuchi, et al., "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





**Figure:** Diagramma rete Wang 2022 et al., immagine presa dal documento *Neural data-dependent transform for learned image compression*<sup>1</sup>

<sup>1</sup>D. Wang, W. Yang, Y. Hu, et al., "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

Esempi di compressione di un immagine a 24bpp del dataset Kodak<sup>1</sup> con le tecniche presentate



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

<sup>1</sup>E. K. Company, *True color kodak images*, <https://r0k.us/graphics/kodak/>, Consultato: 17-10-2023

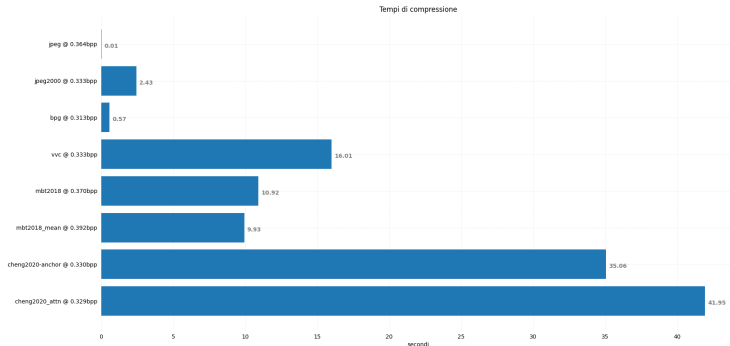
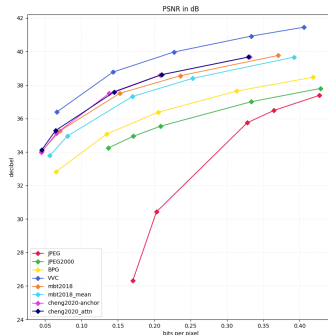
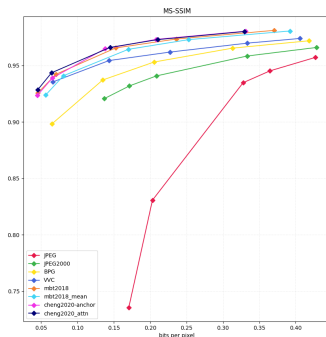


Figure: Tempi di compressione a 0.34 bps

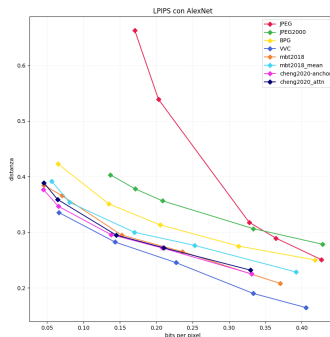


**Figure:** Grafico del PSNR, punti corrispondenti alla media delle metriche sulle 24 immagini del dataset



**Figure:** Grafico dell'MS-SSIM<sup>1</sup>, punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

<sup>1</sup>Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, Ieee, vol. 2, 2003, pp. 1398–1402



**Figure:** Grafico LPIPS<sup>1</sup> con AlexNet, punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

<sup>1</sup>R. Zhang, P. Isola, A. A. Efros, *et al.*, “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

Durante la ricerca delle informazioni per la stesura di questa tesi ci siamo imbattuti in SlimCAE<sup>1</sup>, un lavoro molto interessante del team di Yang et al., che potrebbe facilitare l'utilizzo di metodi con reti neurali sui dispositivi mobili con ridotta potenza di calcolo.

---

<sup>1</sup>F. Yang, L. Herranz, Y. Cheng, *et al.*, "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