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

Indice

Compressione

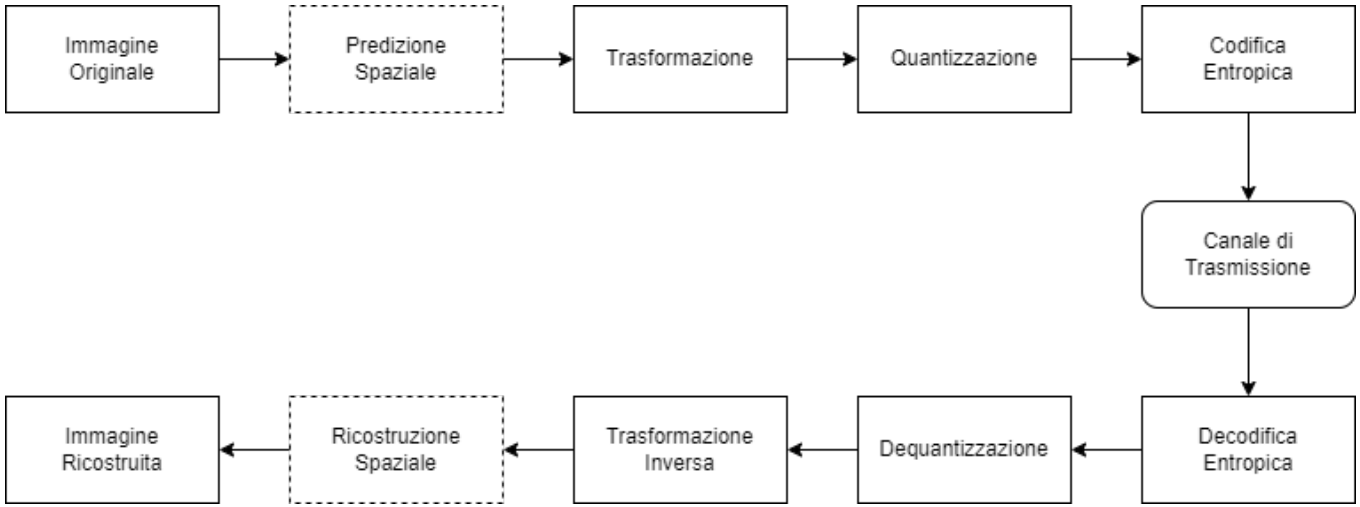

Figure: Blocchi funzionali principali di un framework lossy[1], con l'aggiunta di un quarto blocco per i metodi recenti

[1] 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

Metodi Tradizionali

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

- JPEG[1]
- JPEG 2000[2]
- BPG[3]
- VVC[4]

[1] 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

[2] 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

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

[4] 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

Autoencoder

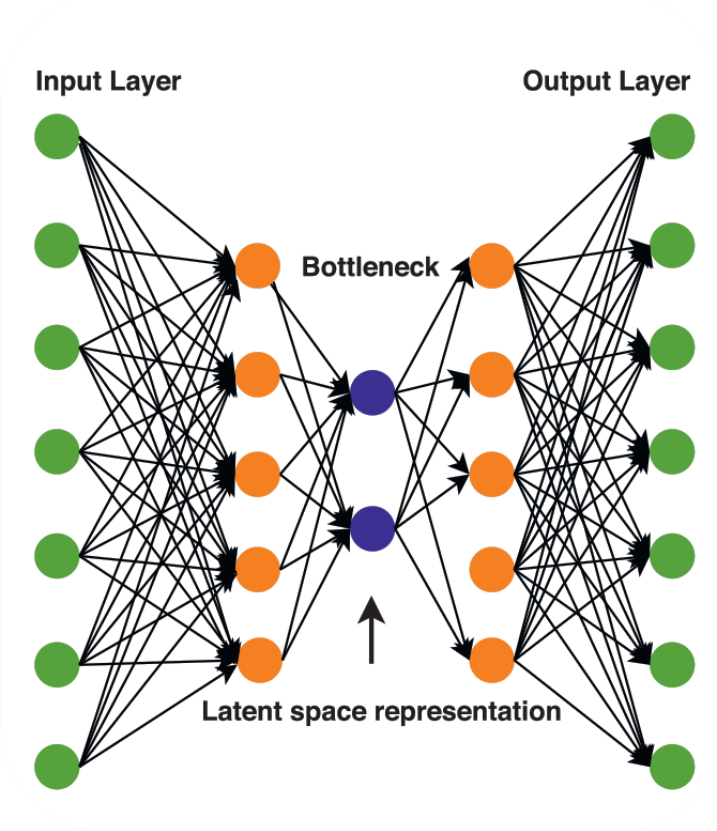

Figure: Schema generico di un autoencoder, immagine presa dal documento *Deep architectures for image compression: A critical review*[1]

[1] 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

Metodi con reti neurali

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

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

[1] 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

[2] 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

[3] 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

Ballé et al[1]. 2018

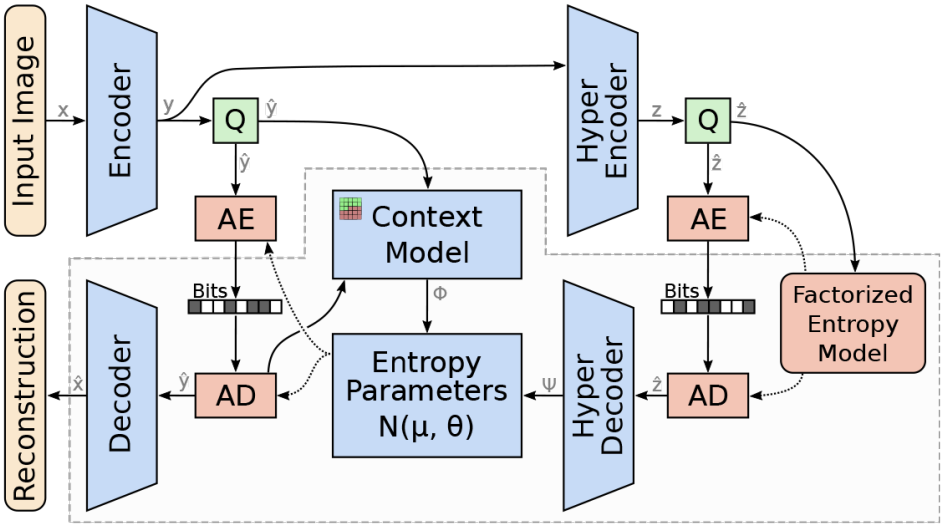

Figure: Diagramma rete Ballé 2018 et al., immagine presa dal documento *Joint autoregressive and hierarchical priors for learned image compression*[1]

[1] 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

Cheng et al[1]. 2020

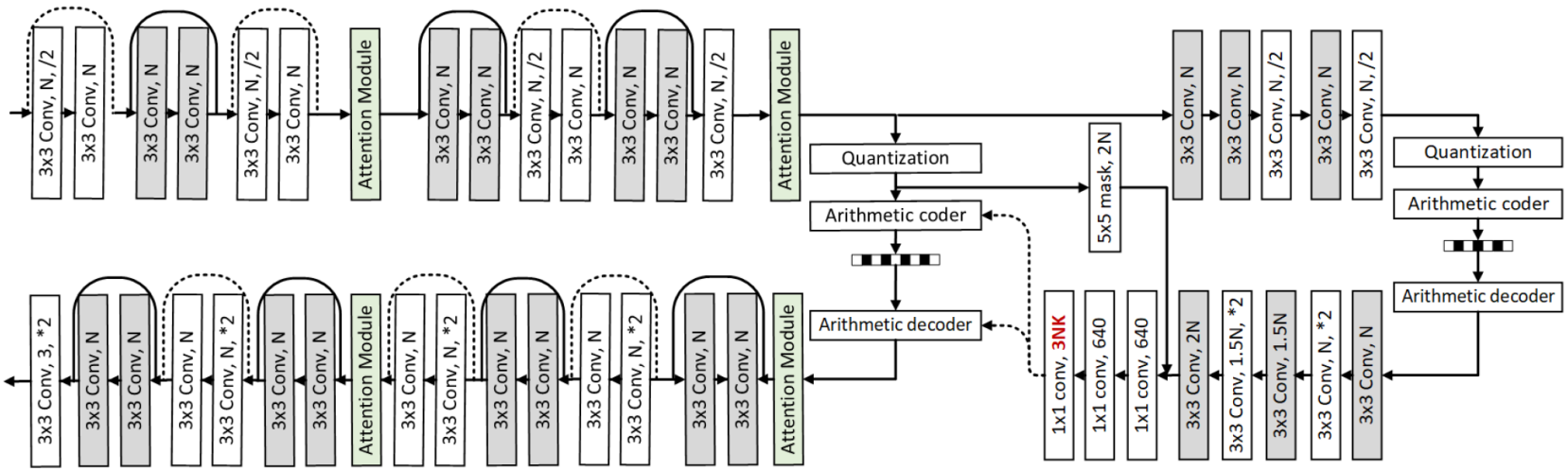

Figure: Diagramma rete Cheng 2020 et al., immagine presa dal documento *Learned image compression with discretized gaussian mixture likelihoods and attention modules*[1]

[1] 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

Wang et al[1]. 2022

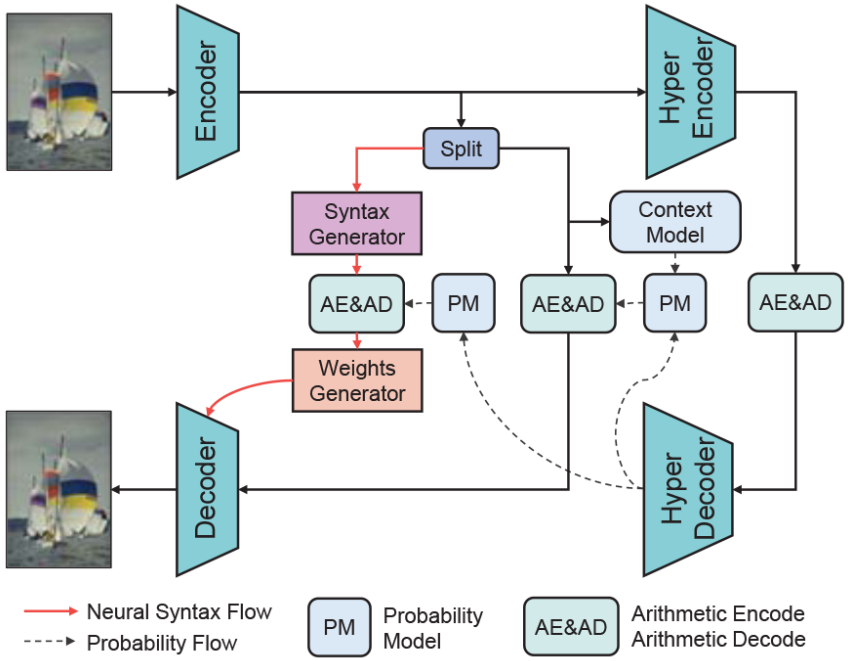

Figure: Diagramma rete Wang 2022 et al., immagine presa dal documento *Neural data-dependent transform for learned image compression*[1]

[1] 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 Compressione

Esempi di compressione di un immagine a 24bpp del dataset Kodak[1] con le tecniche presentate


Figure: Lossless 11.117bpp

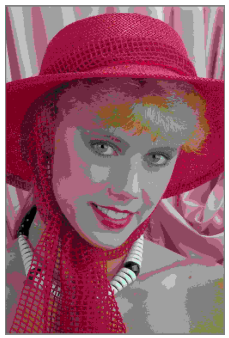
Figure: JPEG 0.167bpp


Figure: JPEG2000 0.137bpp


Figure: BPG 0.103bpp


Figure: VVC 0.061bpp


Figure: Ballé 0.060bpp


Figure: Cheng 0.056bpp

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

Tempi di Compressione

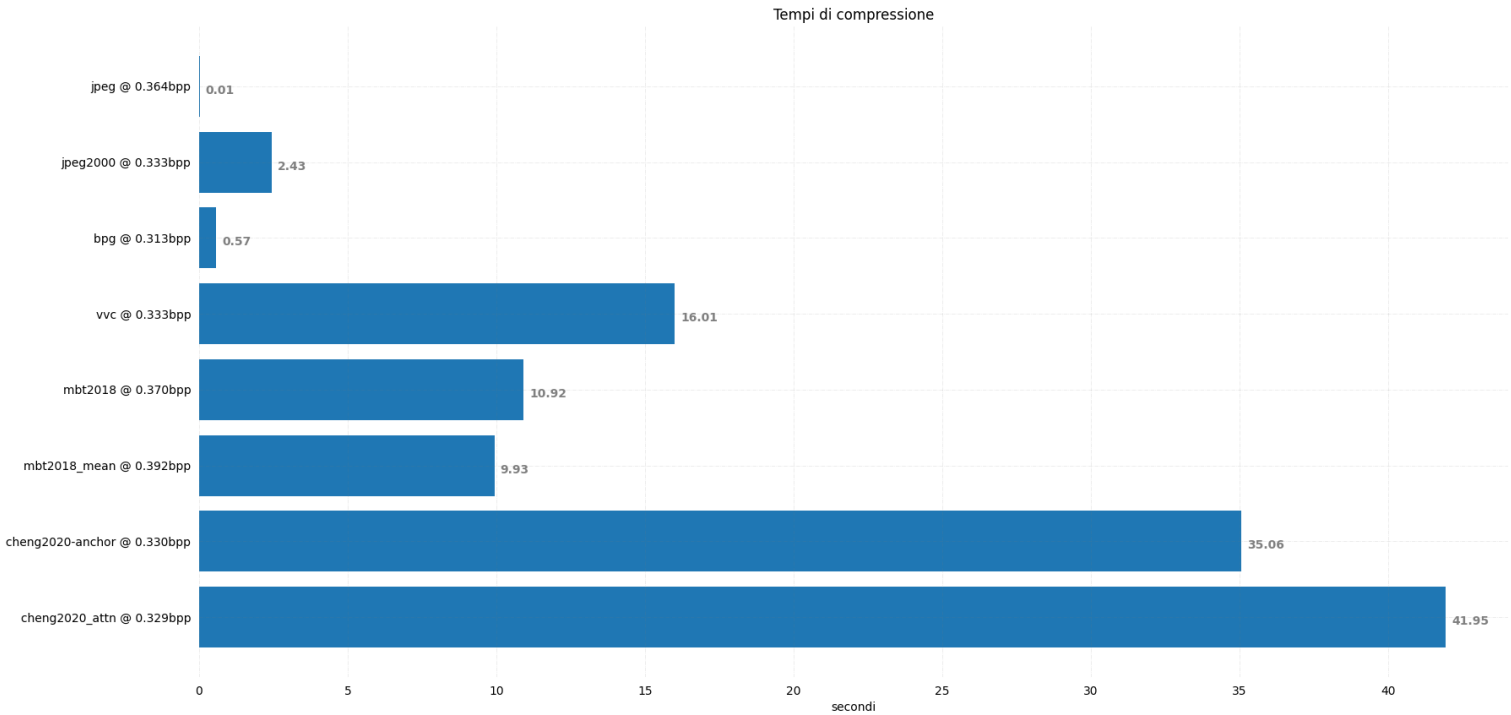

Figure: Tempi di compressione a 0.34 bpp

PSNR

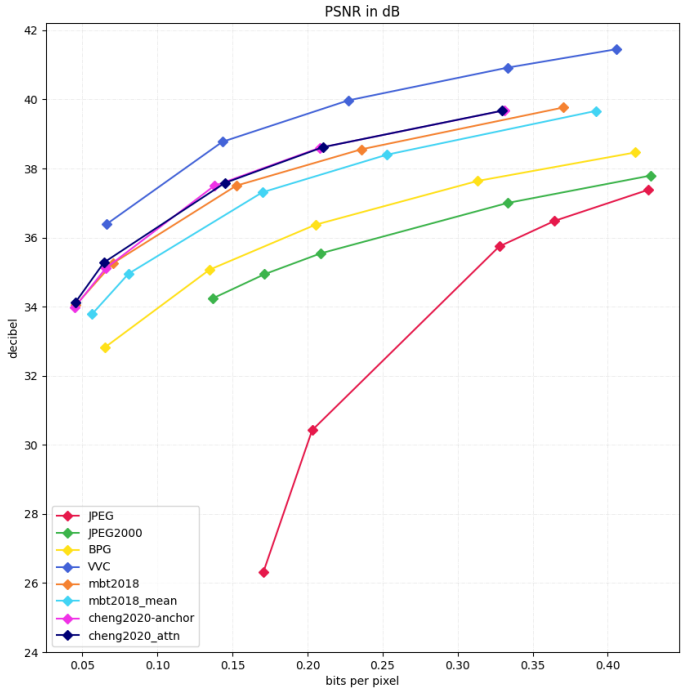

Figure: Grafico del PSNR, punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

MS-SSIM

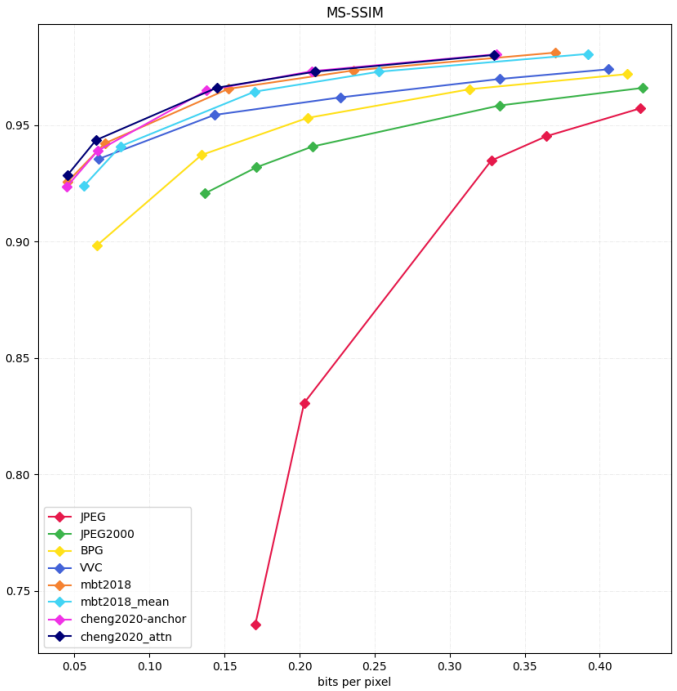

Figure: Grafico dell'MS-SSIM[1], punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

[1] 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

LPIPS con AlexNet

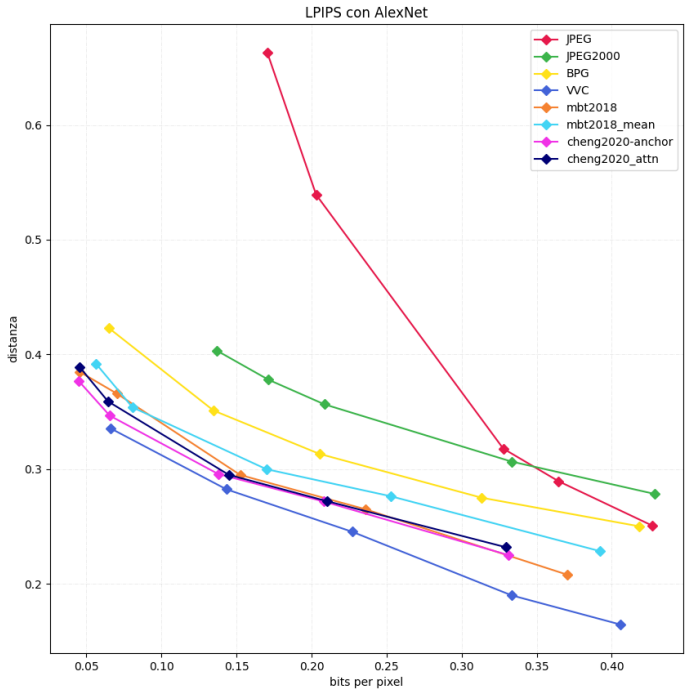

Figure: Grafico LPIPS[1] con AlexNet, punti corrispondenti alla media delle metriche sulle 24 immagini del dataset

[1] 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

Sviluppi Futuri

Durante la ricerca delle informazioni per la stesura di questa tesi ci siamo imbattuti in SlimCAE[1], 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.

[1] 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