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

Indice

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

Compressione

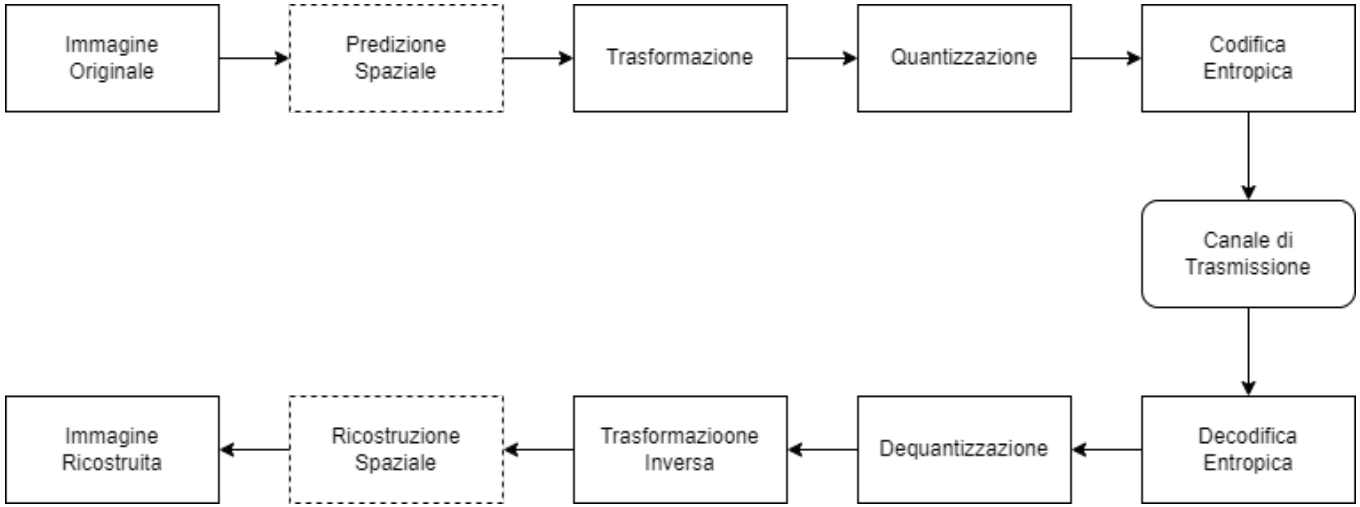

Figure: Blocchi funzionali principali di un framework lossy [1], con l'aggiunta di un quarto blocco per i metodi recenti

Metodi Tradizionali

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

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

Autoencoder

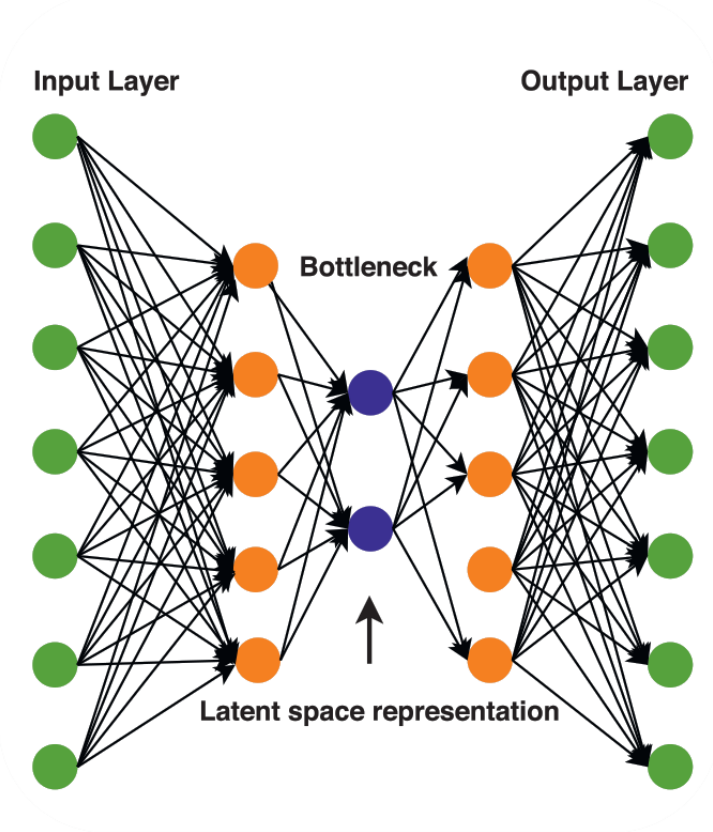

Figure: Schema generico di un autoencoder

Metodi con intelligenza artificiale

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

- Ballé et al. [6]
- Cheng et al. [7]
- Wang et al. [8]

Ballé et al. 2018

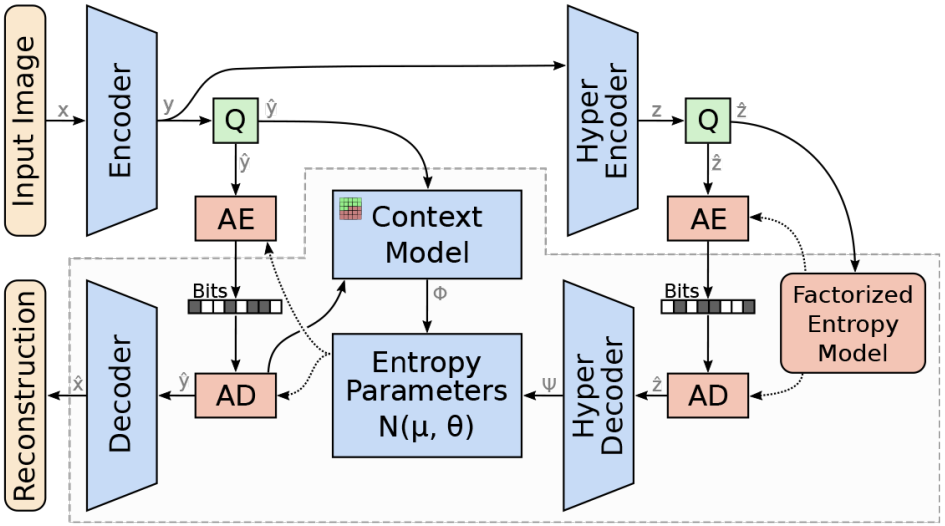

Figure: Diagramma rete Ballè 2018 et al., immagine presa dal documento [6]

Cheng et al. 2020

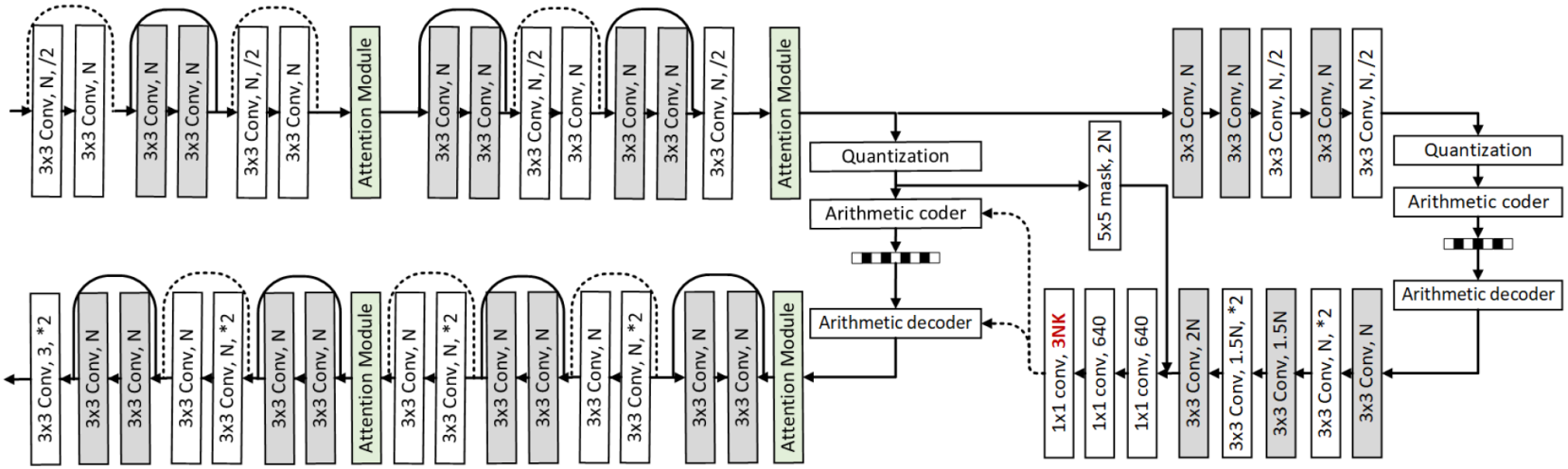

Figure: Diagramma rete Cheng 2020 et al., immagine presa dal documento [7]

Wang et al. 2022

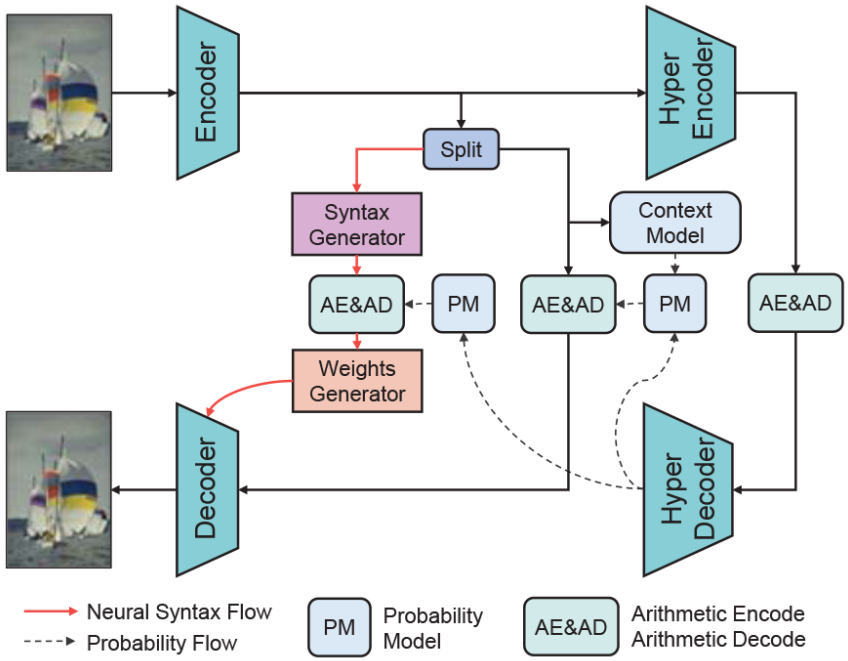

Figure: Diagramma rete Wang 2022 et al., immagine presa dal documento [8]

Esempi Compressione

Esempi di compressione di un immagine del dataset Kodak [9] con le tecniche presentate


Figure: Originale 11.117bpp

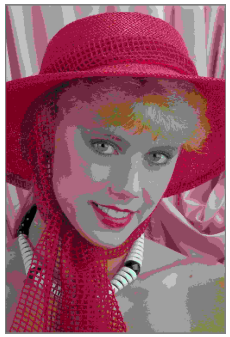
Figure: JPEG 0.167bpp


Figure: JPEG2000 0.171bpp


Figure: BPG 0.156bpp


Figure: VVC 0.144bpp


Figure: Ballé 0.131bpp


Figure: Cheng 0.124bpp

Tempi di Compressione

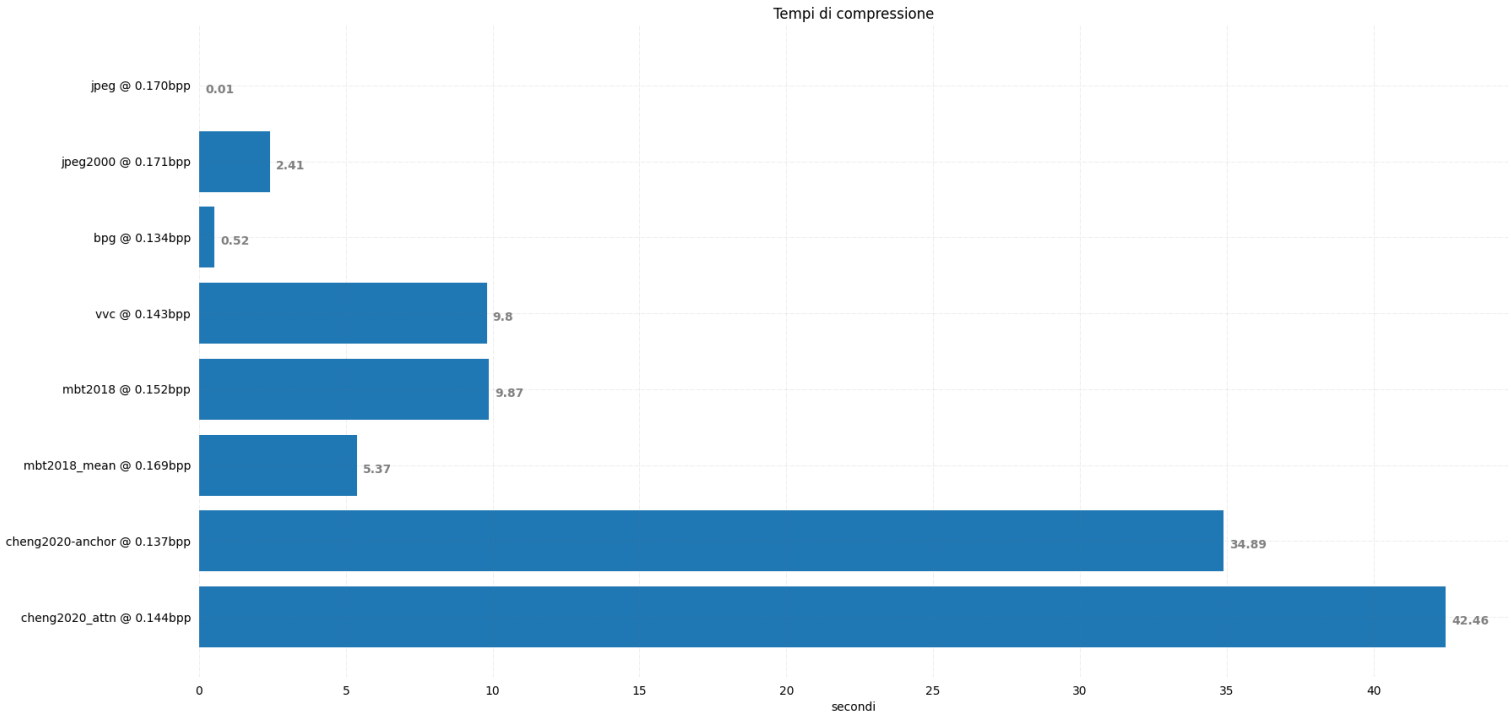

Figure: Tempi di compressione a 0.16 bpp

Tempi di Compressione

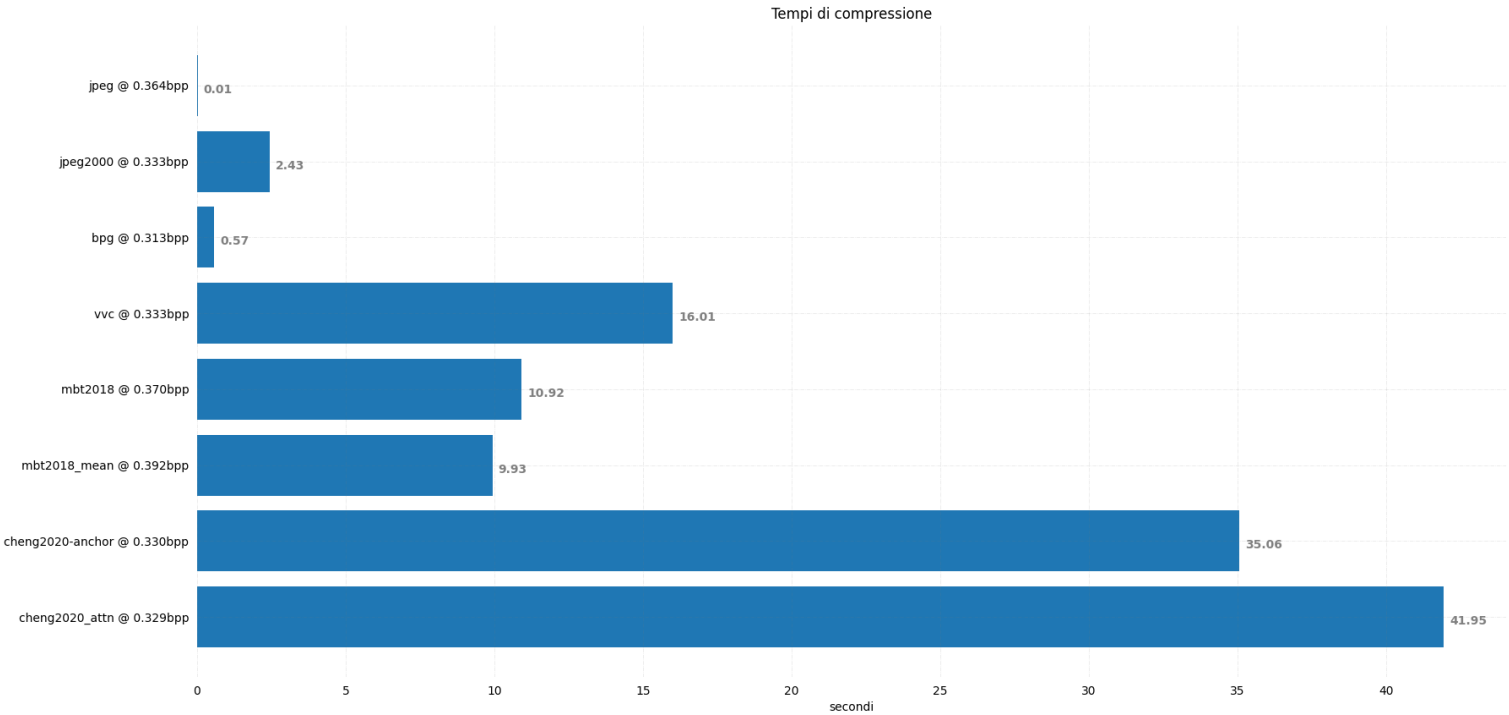

Figure: Tempi di compressione a 0.34 bpp

PSNR

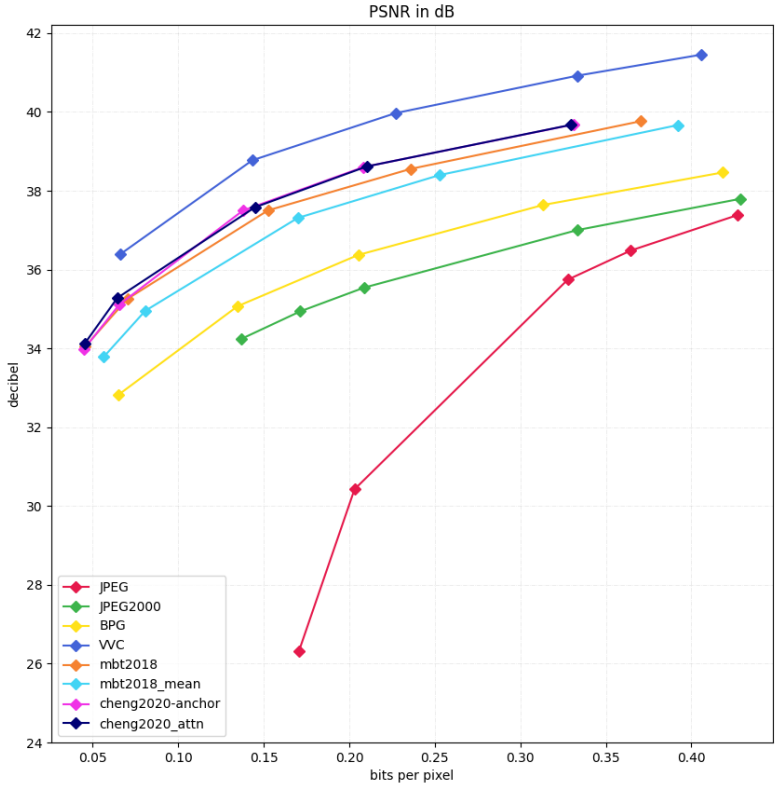

Figure: Grafico del PSNR, punti corrispondenti alla media delle metriche sulle 24 immaigni del dataset

MS-SSIM

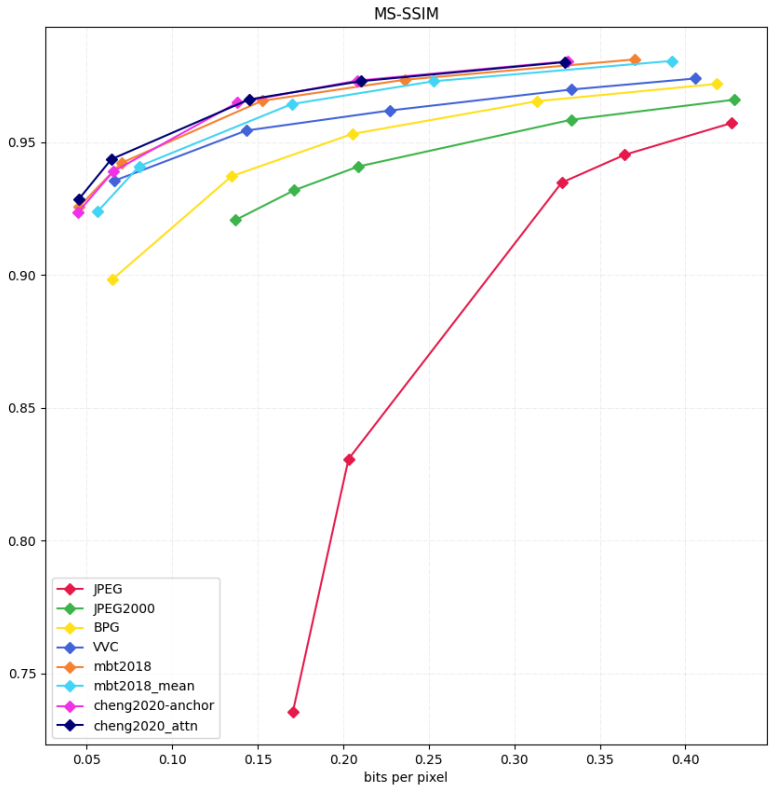

Figure: Grafico dell'MS-SSIM [10], punti corrispondenti alla media delle metriche sulle 24 immaigni del dataset

LPIPS con AlexNet

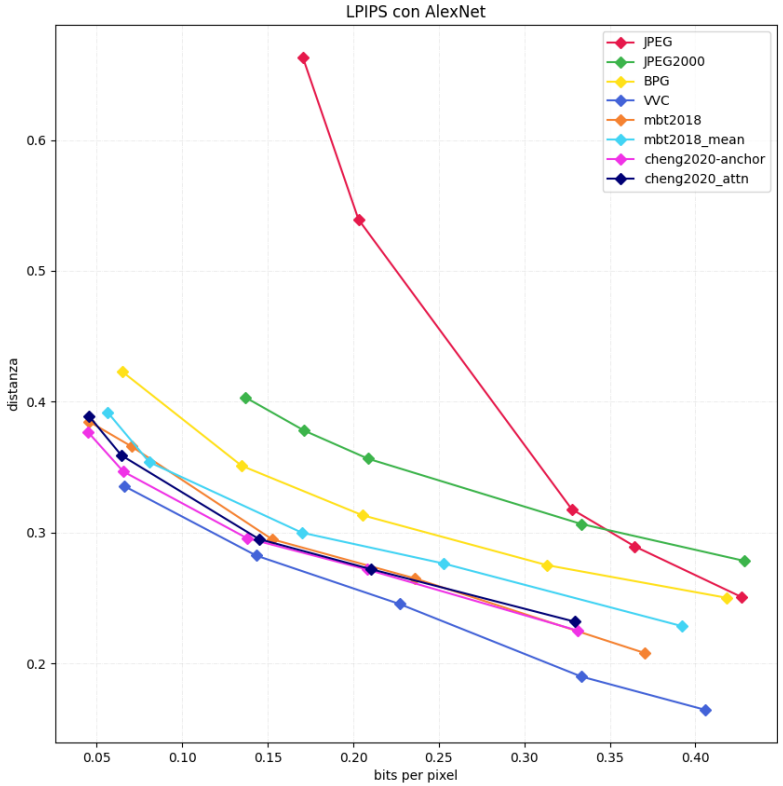

Figure: Grafico LPIPS [11] con AlexNet, punti corrispondenti alla media delle metriche sulle 24 immaigni del dataset

Sviluppi Futuri

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

- StructuralADAM [12]
- SmallCAE [13]

Bibliografia I

[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.

Bibliografia II

[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. 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.

Bibliografia III

[7] 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.

[8] 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.

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

Bibliografia IV

[10] 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.

[11] 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.

[12] J. Ballé, "Efficient nonlinear transforms for lossy image compression," in *2018 Picture Coding Symposium (PCS)*, IEEE, 2018, pp. 248–252.

Bibliografia V

[13] 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