

UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

# **Compressione di immagini tramite autoencoder, stato dell'arte e sviluppi futuri**

**Relatore**

Prof. Cagnazzo Marco

**Laureando**

Stella Filippo

ANNO ACCADEMICO 2022-2023

Data di laurea GG/MM/AAAA



*Apes. Together. Strong.*

Testo di prova



# Introduzione

Nell'era moderna la crescente quantità di dati di cui usufruiamo ogni giorno sta ricevendo, dagli esperti del settore, molte attenzioni. In quanto il throughput e la quantità di memoria di cui disponiamo sui nostri dispositivi sono, seppur ad oggi ampiamente sufficienti, comunque limitati. Un'altra ragione di questa attenzione è la crescente diffusione di servizi in tempo reale che quindi richiedono di scambiare quantità di dati considerevoli in pochissimo tempo. Un esempio di tali servizi potrebbe essere l'uso della realtà aumentata. Comprimere i dati è quindi ormai una necessità, che si farà sempre più impellente con il crescere delle dimensioni dei dati che andremo a gestire.

I metodi di compressione delle immagini sono solitamente divisi in due grandi categorie, Lossless e Lossy. I primi comprimono la rappresentazione senza scartare alcuna informazione. I secondi invece permettono di scartare alcune informazioni, senza impattare significativamente sulla qualità, per permettere una rappresentazione più ridotta.[1] Noi ci concentreremo sui metodi Lossy in quanto sono quelli che stanno suscitando più interesse e presentano più difficoltà. Comprimere un'immagine significa ridurre la quantità di bit necessari per rappresentare l'immagine, questo si ottiene riducendo le ridondanze e eliminando i dati irrilevanti. [2] A partire dai primi anni novanta si sono iniziate a sviluppare alcune tecniche di compressione, a cui oggi si fa riferimento come tecniche “tradizionali”. Stiamo parlando ad esempio di JPEG e del suo successore JPEG2000, di più recente sviluppo sono invece i codec BPG e VVC. Più recentemente l'attenzione si è spostata su metodi basati su deep learning. Questi metodi presentano diversi vantaggi rispetto ai metodi “tradizionali”, principalmente la flessibilità e la capacità di individuare con più precisione le regioni più significative delle immagini. In questo documento ci proponiamo di fornire una panoramica dei metodi di compressione tradizionali più usati, e di quelli basati su deep learning che hanno fornito un maggiore contributo allo sviluppo di questi ultimi. Nello specifico ci concentreremo sui metodi che fanno uso di autoencoder, in quanto hanno prodotto i risultati più significativi.

Dopo aver presentato le varie tecniche forniremo una valutazione delle prestazioni ed evidenzieremo potenzialità e difetti. In conclusione vorremmo fornire alcuni spunti per che potrebbero essere oggetto di future ricerche o approfondimenti.



# Indice

<b>1</b>	<b>Metodi tradizionali</b>	<b>1</b>
1.1	JPEG . . . . .	2
1.2	JPEG2000 . . . . .	2
1.3	BPG . . . . .	2
1.4	VVC . . . . .	2
<b>2</b>	<b>Metodi con apprendimento automatico</b>	<b>5</b>
2.1	Ballé 2018 . . . . .	5
2.2	Cheng 2020 . . . . .	5
2.3	Yang 2021 . . . . .	5
2.4	Wang 2022 . . . . .	5
<b>3</b>	<b>Valutazione delle prestazioni</b>	<b>7</b>
3.1	Metriche utilizzate . . . . .	7
3.1.1	BPP . . . . .	7
3.1.2	Tempo di codifica . . . . .	7
3.1.3	PSNR . . . . .	7
3.1.4	MSSIM . . . . .	7
3.1.5	LPIPS . . . . .	7
3.2	Presentazione dei risultati . . . . .	7
<b>4</b>	<b>Sviluppi futuri</b>	<b>9</b>
4.1	Possibili ottimizzazioni . . . . .	9
4.1.1	Spectral ADAM . . . . .	9
4.2	Dispositivi mobili . . . . .	9
4.2.1	Small CAE . . . . .	9
<b>5</b>	<b>Conclusioni</b>	<b>11</b>





# Capitolo 1

## Metodi tradizionali

Se andiamo ad osservare gli algoritmi di codifica Lossy possiamo scomporli tutti in almeno tre blocchi principali. [1]

Un primo blocco si occupa di convertire l'immagine in una rappresentazione latente, tramite l'applicazione di una trasformata, in un altro dominio che permette di rappresentare l'informazione da comprimere in modo più sparso.

Successivamente si ha un blocco che si occupa di quantizzazione, ovvero di mappare i valori in ingresso in un insieme finito di dimensione più piccola rispetto a quello di ingresso. In questo passaggio si realizza la perdita di informazioni caratteristica della codifica Lossy.

Un ultimo blocco si occupa di effettuare la codifica entropica, in modo da comprimere ulteriormente l'informazione mappando i simboli usati più spesso con pochi bit.

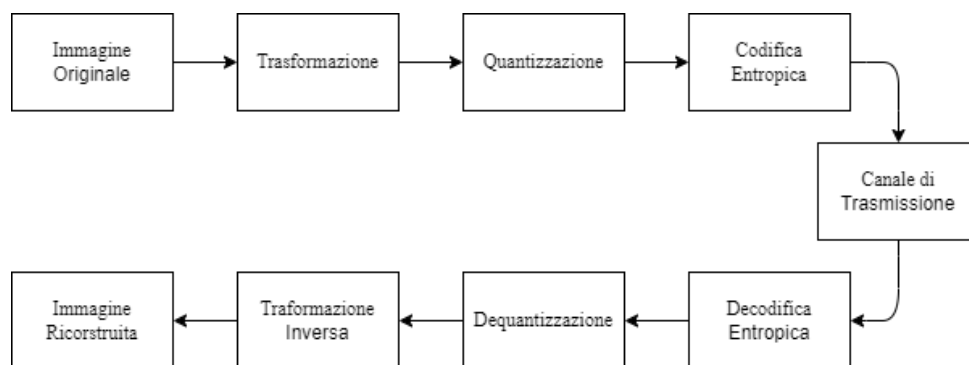


Figura 1.1: Diagramma di compressione Lossy

Gli algoritmi di compressione tradizionali usano trasformate statiche all'interno del primo blocco, messe a punto in numerosi anni di ricerca. Questa staticità non permette a questi metodi di adattarsi dinamicamente a tutti i tipi di contenuti delle immagini. [3] Inoltre rende il processo di sviluppo di un nuovo algoritmo di compressione un processo lungo che richiede anni di studi

e progettazione. [3] Andiamo ora a parlare brevemente dei metodi che andremo a considerare quando valuteremo le prestazioni dei vari algoritmi per poterli confrontare.

## 1.1 JPEG

Questo metodo di compressione sviluppato nel 1992 è diventato in poco tempo ed è attualmente il formato di compressione più diffuso. Nonostante in numerosi tentativi di sostituirlo con formati più moderni, questo è rimasto ancora ad oggi il formato più usato. Il JPEG si basa sull'utilizzo della DCT per realizzare la rappresentazione sparsificata dell'immagine originale. [4]  
IMMAGINE CON JPEG

## 1.2 JPEG2000

Nel 2001 con la crescente diffusione di internet e con l'aumento di dimensione delle immagini e la richiesta di una maggiore qualità da parte degli utenti viene sviluppato questo nuovo formato chiamato appunto JPEG2000 per l'anno in cui è stato sviluppato.

JPEG2000 non utilizza la DCT come il suo predecessore ma viene introdotta una nuova trasformazione la DWT o Discrete Wavelet Transform che si propone di meglio identificare e comprimere i bordi delle figure che compongono le immagini. JPEG2000 quindi voleva essere un formato di qualità superiore con una compressione più efficiente. [5] IMMAGINE CON JPEG2000

## 1.3 BPG

Questo formato è stato sviluppato da Fabrice Bellard nel 2014 come sostituto all'ormai affermato formato JPEG. Questo metodo di codifica si basa sulla codifica intra-frame del codec HEVC o H.265. [6] Bellard voleva realizzare un formato molto leggero che potesse fornire immagini più compresse rispetto a JPEG, ma con una qualità superiore.

IMMAGINE CON BPG

## 1.4 VVC

Versatile Video Coding (VVC) o H.266 è lo standard di codifica video più recente, finalizzato nel luglio 2020. È stato sviluppato dal Joint Video Experts Team (JVET) dell'ITU-T Video Coding Experts Group (VCEG) e dell'ISO/IEC Moving Picture Experts Group (MPEG) per soddisfare la crescente richiesta di una migliore compressione video e per supportare una più ampia gamma di contenuti multimediali attuali e applicazioni emergenti come contenuti HDR, a 360°, per la

Realtà Virtuale (VR) o la Realtà Aumentata (AR).[7]

IMMAGINE CON VVC



## **Capitolo 2**

### **Metodi con apprendimento automatico**

**2.1 Ballé 2018**

**2.2 Cheng 2020**

**2.3 Yang 2021**

**2.4 Wang 2022**



# **Capitolo 3**

## **Valutazione delle prestazioni**

### **3.1 Metriche utilizzate**

#### **3.1.1 BPP**

#### **3.1.2 Tempo di codifica**

#### **3.1.3 PSNR**

#### **3.1.4 MSSIM**

#### **3.1.5 LPIPS**

### **3.2 Presentazione dei risultati**





# **Capitolo 4**

## **Sviluppi futuri**

### **4.1 Possibili ottimizzazioni**

#### **4.1.1 Spectral ADAM**

### **4.2 Dispositivi mobili**

#### **4.2.1 Small CAE**



## **Capitolo 5**

## **Conclusioni**



# Bibliografia

- [1] H. T. Sadeeq, T. H. Hameed, A. S. Abdi e A. N. Abdulfatah, «Image compression using neural networks: a review,» *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, n. 14, pp. 135–153, 2021.
- [2] D. Mishra, S. K. Singh e R. K. Singh, «Deep architectures for image compression: a critical review,» *Signal Processing*, vol. 191, p. 108 346, 2022.
- [3] Z. Cheng, H. Sun, M. Takeuchi e J. Katto, «Deep convolutional autoencoder-based lossy image compression,» in *2018 Picture Coding Symposium (PCS)*, IEEE, 2018, pp. 253–257.
- [4] G. Wallace, «The JPEG still picture compression standard,» *IEEE Transactions on Consumer Electronics*, vol. 38, n. 1, pp. xviii–xxxiv, 1992. DOI: 10.1109/30.125072.
- [5] A. Skodras, C. Christopoulos e T. Ebrahimi, «The JPEG 2000 still image compression standard,» *IEEE Signal Processing Magazine*, vol. 18, n. 5, pp. 36–58, 2001. DOI: 10.1109/79.952804.
- [6] F. Bellard, *BPG Image format*, <https://bellard.org/bpg/>, Consultato: 17-10-2023.
- [7] 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, n. 10, pp. 3736–3764, 2021. DOI: 10.1109/TCSVT.2021.3101953.