

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

Ringraziamenti

Sommario

Questo elaborato esplora il campo emergente della compressione delle immagini utilizzando autoencoder. L'obiettivo principale della ricerca è studiare e analizzare le tecniche esistenti per la compressione delle immagini tramite autoencoder.

Nel primo capitolo, la tesi fornisce una panoramica dello stato dell'arte della compressione delle immagini, discutendo le tecniche tradizionali e come gli autoencoder si inseriscono in questo panorama. Viene fornita una spiegazione dettagliata del funzionamento degli autoencoder e dei loro vantaggi rispetto ai metodi di compressione delle immagini tradizionali.

Il secondo capitolo si concentra sull'analisi delle tecniche esistenti per la compressione delle immagini tramite autoencoder. Vengono presentati vari studi di caso e vengono discusse le prestazioni di ciascuno in termini di qualità dell'immagine e rapporto di compressione.

Nel terzo e ultimo capitolo, la tesi esplora i possibili sviluppi futuri nel campo della compressione delle immagini tramite autoencoder. Si discute di come le tecniche attuali potrebbero essere migliorate e si ipotizzano nuove direzioni di ricerca.

In conclusione, questa tesi fornisce una panoramica completa dello stato dell'arte della compressione delle immagini tramite autoencoder e offre spunti preziosi per future ricerche in questo campo.

Indice

Introduzione	1
1 Metodi tradizionali	3
1.1 JPEG	3
1.2 JPEG2000	3
1.3 BPG	4
1.4 VVC	4
2 Metodi con apprendimento automatico	5
2.1 Ballé 2018	5
2.2 Cheng 2020	5
2.3 Yang 2021	5
2.4 Wang 2022	5
3 Valutazione delle prestazioni	7
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
4 Sviluppi futuri	9
4.1 Possibili ottimizzazioni	9
4.1.1 Spectral ADAM	9
4.2 Dispositivi mobili	9
4.2.1 Small CAE	9

5 Conclusioni	11
Bibliografia	13

Elenco delle figure

1	Diagramma di compressione Lossy	2
---	---	---

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 limitate. 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 in ambito medico o di ricerca. 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. A partire dai primi anni novanta infatti si sono iniziate a sviluppare alcune tecniche, a cui oggi si fa riferimento come tecniche “tradizionali”, per comprimere le immagini. 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 dei ricercatori si è spostata su metodi basati su deep learning. Questi metodi presentano diversi vantaggi rispetto ai metodi “tradizionali”, infatti molte volte permettono di ottenere performance migliori rispetto ai metodi classici. 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. Dopo aver presentato le varie tecniche vogliamo fornire una valutazione delle prestazioni in modo da poterli comprare ed evidenziare potenzialità e difetti. La compressione è un processo che mira a minimizzare il numero di bit utilizzati per rappresentare una certa informazione senza intaccarne drasticamente la qualità. Questo obiettivo viene raggiunto riducendo le ridondanze e eliminando i dati irrilevanti. Tutti i framework di compressione consistono di una coppia codificatore-decodificatore. Data un'immagine da comprimere x , il codificatore è composto da una trasformata ϵ e una funzione di quantizzazione Q può essere espresso nella seguente formulazione

$$y = Q(\epsilon(x; \theta_\epsilon)) \quad (1)$$

Dove θ_ϵ denota i parametri del codificatore.

Per riottenere la rappresentazione dell'immagine il decodificatore ricostruisce l'immagine \hat{x} dal codice y nel seguente modo.

$$\hat{y} = D(y; \theta_D) = D(Q(\epsilon(x; \theta_\epsilon)); \theta_D) \quad (2)$$

Dove θ_D denota i parametri del decodificatore. [1] ognuno. In conclusione vorremmo fornire alcuni spunti per che potrebbero favorire l'adozione di queste tecniche più recenti nella nostra vita digitale di tutti i giorni. Tutti gli algoritmi di compressione possono essere raggruppati in due macro categorie. Gli algoritmi Lossy o con perdita e quelli Lossless o senza perdita. Come si può già intuire dal nome gli algoritmi Lossless comprimono le informazioni senza scartare informazioni, si limitano quindi ad applicare delle trasformate per ottenere delle nuove rappresentazioni più efficienti. Gli algoritmi Lossy invece ammettono la possibilità di scartare delle informazioni superflue che non vanno ad intaccare drasticamente la qualità percepita dell'immagine, in modo da poter comprimere ulteriormente. Se andiamo ad osservare gli algoritmi di codifica Lossy possiamo scomporli tutti in almeno tre blocchi principali. [2]

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.

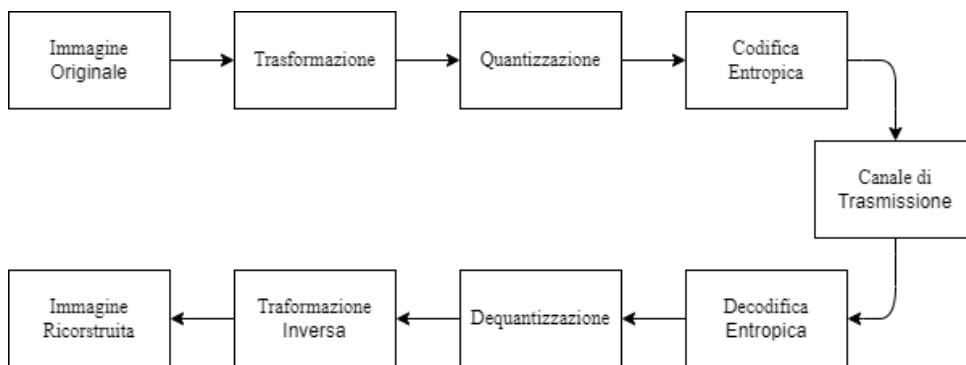


Figura 1: Diagramma di compressione Lossy

Capitolo 1

Metodi tradizionali

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 trasformata 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] Y. Hu, W. Yang, Z. Ma e J. Liu, «Learning end-to-end lossy image compression: A benchmark,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, n. 8, pp. 4194–4211, 2021.
- [2] 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.
- [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.