

高等數位訊號處理

書面報告

不同的影像壓縮技術以及機器學習上的應用

系所:台灣科技大學 資訊工程研究所 碩士一年級

學號:M11215075

姓名:胡劭

指導教授:丁建均 教授

摘要

本論文探討了卷積神經網路(CNN)在影像壓縮任務中的應用，並將其性能與傳統的 JPEG 標準進行比較。首先介紹了 JPEG 壓縮算法的基本流程，包括色彩子採樣、離散餘弦變換、差分編碼和霍夫曼編碼等技術。接著，論文描述了利用 CNN 進行影像壓縮和解壓縮的架構，並介紹了三種可能的模型架構。

實驗使用 10 張大型貓科動物圖像進行測試。結果表明，儘管 JPEG 在主觀視覺質量方面表現更佳，但基於 CNN 的方法能實現更高的壓縮比率，顯示出更好的壓縮效率。不過 CNN 模型在保持良好圖像品質的同時，還需要進一步優化調整。因此，在硬件成本和緩衝區大小受限的應用場景中，基於深度學習的影像壓縮算法可能會成為傳統技術的有力替代方案，但仍有改進空間。

值得一提的是，實驗中採用的 CNN 模型架構是一種因式分解優先自動編碼器，出自"Variational image compression with a scale hyperprior"這篇論文。該模型通過端到端的訓練，學習從原始影像中提取有效的特徵表示，再對這些特徵進行高效率的編碼，最終達到影像壓縮的目的。儘管如此，與 JPEG 相比，CNN 模型在主觀視覺質量方面的表現仍略遜一籌。

結論未來可以優化神經網路模型架構和訓練策略，在視覺質量和壓縮比率之間取得更好的平衡，為影像壓縮領域提供更先進的解決方案。同時也可以探索更複雜、更強大的 CNN 架構，提升其學習影像特徵的能力，在壓縮率和視覺品質兩個方面都超越傳統算法。

1.Introduction

影像壓縮是影像處理、計算機視覺和信號處理領域中的一個重要課題。隨著物聯網(IoT)設備的快速發展以及訓練資料集規模的日益擴大，高效的影像壓縮算法變得愈發重要。未來的影像壓縮發展可能會集中在硬體成本有限或緩衝區大小受限的情況下如何實現更高的壓縮率。

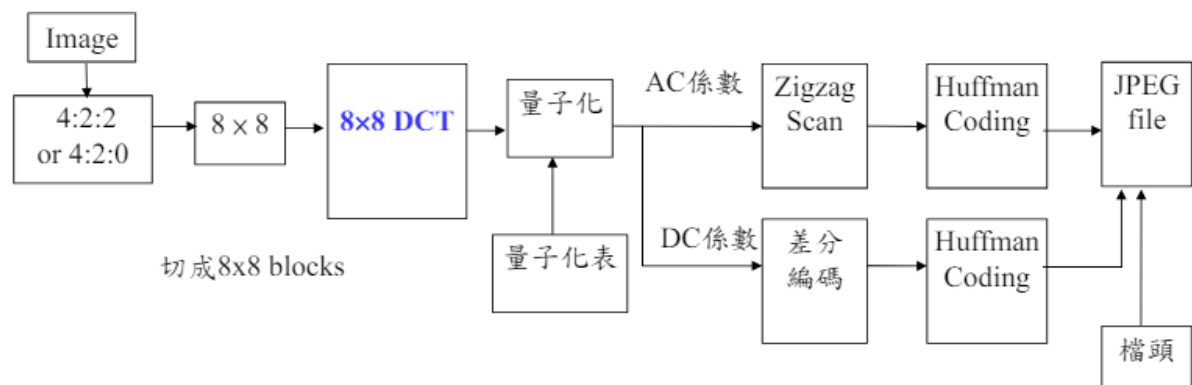
在過去的幾年中，已經開發了許多經典的影像壓縮技術，例如 JPEG、JPG、WebP 和 VVC 等。這些方法已經設定了影像壓縮的標準，但最近的研究表明，深度學習方法在這一領域具有巨大的潛力【1,3】。神經網路，特別是卷積神經網路(CNN)，由於其能夠學習和優化資料中的複雜模式，提供了改善壓縮和解壓任務的有希望的途徑。

值得注意的是，傳統影像壓縮算法往往需要大量的人工設計和領域知識，而基於深度學習的方法則能夠自動從數據中學習最優表示，減少人工干預。此外，神經網路具有良好的並行計算能力，有望在嵌入式系統等硬體資源受限的場景下提供高效的解決方案。

本文旨在探索神經網路在影像壓縮中的應用，並將其性能與傳統技術進行比較。具體來說，我們將使用廣泛認可的影像壓縮標準 JPEG 作為比較的基準。研究將採用卷積神經網路(CNN)和 Python 影像處理庫(PIL)來進行神經網路方法與傳統 JPEG 方法的詳細比較。

通過進行這項比較，我們希望能夠突顯深度學習在影像壓縮領域帶來的優勢和潛在改進。我們的最終目標是證明基於神經網路的方法能夠提供更高的壓縮率和品質，使其成為傳統技術在實際應用中的可行替代方案，特別是在資源受限的環境下。同時，這項研究也有望促進影像壓縮領域中基於深度學習的新理論和新算法的發展。

2. JPEG Standard



圖一 Process of JPEG Image Compression

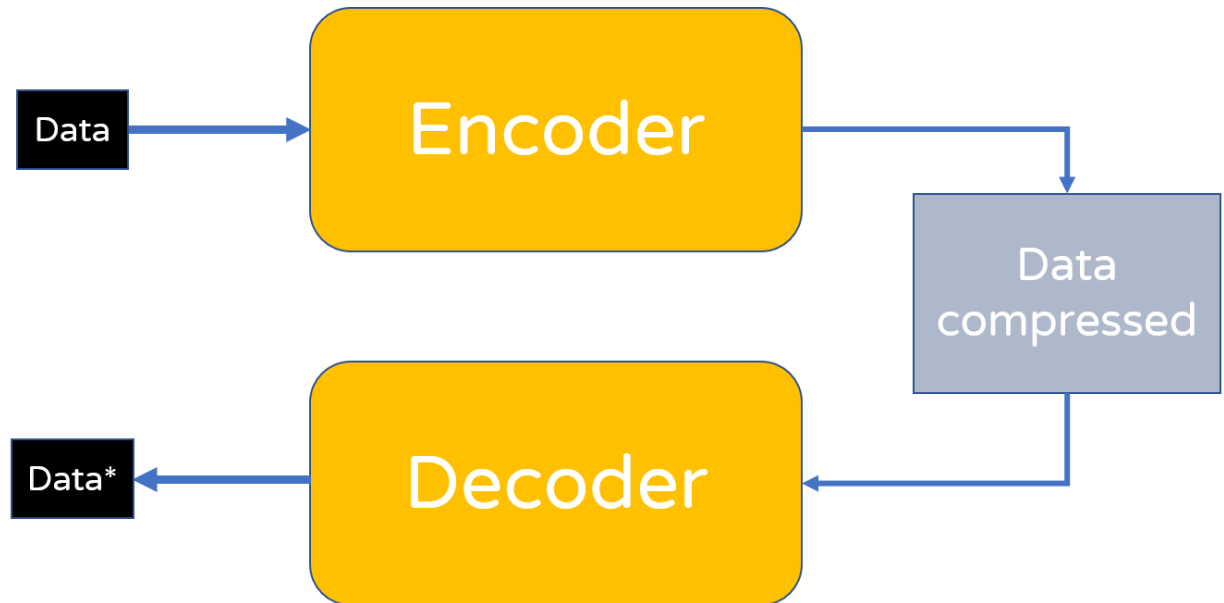
圖一為 JPEG 格式常用的壓縮技術流程圖，主要用到的四個技術:

- (1) 4:2:2 or 4:2:0 (和 space domain 的一致性相關)
- (2) 8 X 8 DCT (和 frequency domain 的一致性相關)
- (3) 差分編碼 (和 space domain 的一致性相關)
- (4) Huffman coding (和 lossless 編碼技術相關)

3. Compression-decompression 方法

此過程的目標是要實現原始 image 和解壓縮影像之間最小的差異，並在壓縮跟解壓後獲得相同的品質，為了比較不同方法的效能，用 SSIM 和 PSNR 指標來測量。

圖二 Compression-decompression 的方法架構



PSNR:

$$10\log_{10}\left(\frac{X_{Max}^2}{\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|y[m,n]-x[m,n]|^2}\right)$$

X_{Max} : the maximal possible value of $x[m,n]$

In image processing, $X_{Max} = 255$

for color image: $10\log_{10}\left(\frac{X_{Max}^2}{\frac{1}{3MN}\sum_{R,G,B}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|y_{color}[m,n]-x_{color}[m,n]|^2}\right)$

color = R, G, or B

圖三 PSNR 的公式

SSIM:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + (c_1L)^2)}{(\mu_x^2 + \mu_y^2 + (c_1L)^2)} \frac{(2\sigma_{xy} + (c_2L)^2)}{(\sigma_x^2 + \sigma_y^2 + (c_2L)^2)}$$

$$DSSIM(x, y) = 1 - SSIM(x, y)$$

μ_x, μ_y : means of x and y σ_x^2, σ_y^2 : variances of x and y

σ_{xy} : covariance of x and y c_1, c_2 : adjustable constants

L : the maximal possible value of x – the minimal possible value of x

圖四 SSIM 的公式

其他指標:

1.Compression efficiency/compression coefficient:壓縮後的 image size 與

原始 image size 之間的比率

$N_compression = size(壓縮過的 image) / size (未壓縮的 image)。$

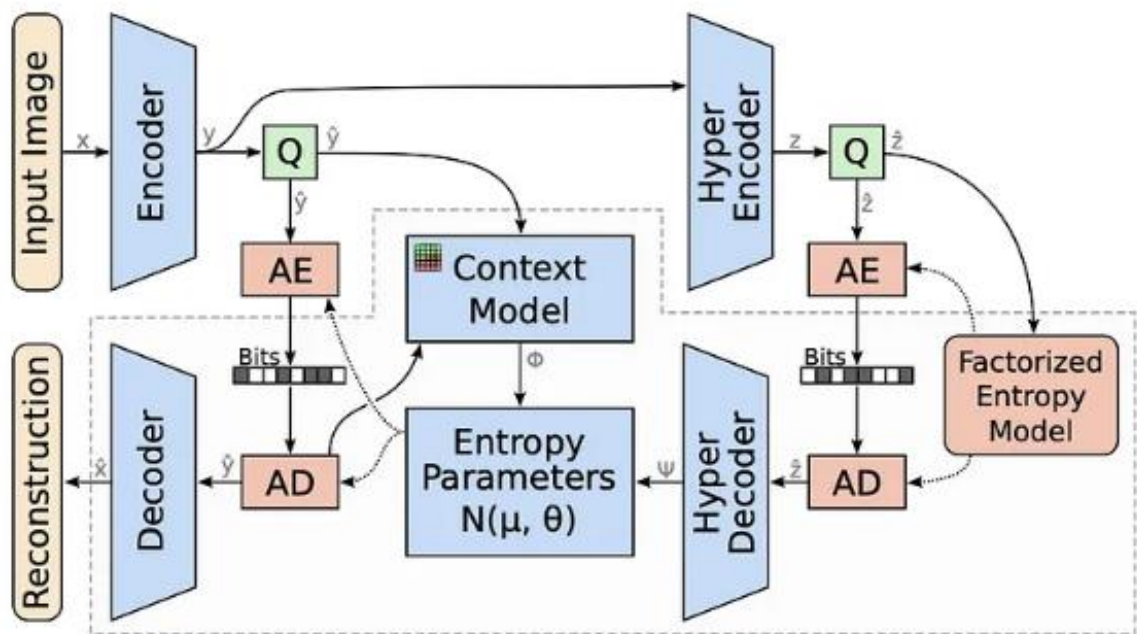
Size 為 $height * width * channel$

Image quality:原始 image 和壓縮/解壓之間的差異

比較的機器學習架構，根據一些 paper 與前人的經驗選擇了最少算力的

model，其中有幾個參考的模型:

Model 3: 具有非零均值高斯條件的超先驗模型



圖七 具有非零均值高斯條件的超先驗模型架構

雖然我介紹了三個 model，但本次 project 只使用第一個 model，主因是其他 model 的 code 有點丟失，在時間不允許的情況下，很難有大量的時間可以把剩餘的 model 的程式碼給手刻出來，再三考量之下決定以 model1 來跟 JPEG 做比較。

2.實驗

數聚集:用 10 個圖像來比較和測試壓縮任務，每張圖片有一隻大型貓科動物。





圖八 資料集

實驗結果:

圖九為使用 JPEG 的壓縮方法，把 bmp->jpeg，SSIM 大約為 0.8945

1	Image	SSIM	PSNR	N Compression
2	01.png	0.9439	26.353	0.523958333
3	02.png	0.9202	28.649	0.364305556
4	03.png	0.875	27.162	0.388680556
5	04.png	0.9166	26.532	0.4871875
6	05.png	0.8272	31.087	0.256111111
7	06.png	0.8911	27.995	0.377013889
8	07.png	0.9242	28.015	0.392395833
9	08.png	0.9312	28.401	0.370486111
10	09.png	0.8176	28.899	0.280763889
11	10.png	0.8979	31.91	0.262326389
12	Average	0.8945	28.5	0.370322917

圖九 JPEG 壓縮方法結果

Image	SSIM	PSNR	N Compression
01.png	0.5175	16.9	0.726
02.png	0.5835	20.37	0.6663
03.png	0.7053	20.61	0.592
04.png	0.4821	18.87	0.6531
05.png	0.7703	25.4	0.5885
06.png	0.7622	21.41	0.6368
07.png	0.6091	19.9	0.6659
08.png	0.7045	20.09	0.6928
09.png	0.6634	24.65	0.5092
10.png	0.7932	31.91	0.629
Average	0.6591	21.26	0.63956

圖十 model 1 壓縮的方法

以實驗的結果來說，JPEG 的壓縮方法 SSIM 比較高約為 0.8945，在 model1 SSIM 產生大幅的下降，不確定是 model 的問題，還是資料集的問題，因為我是用不同的資料集來跑這個實驗，但是明顯的利用 CNN 可以得到比較好的壓縮比。

3.結論

儘管傳統的影像壓縮技術如 JPEG 已經建立了壓縮標準，但最新的研究表明，基於深度學習的神經網路方法在這一領域具有巨大潛力。本文探索了卷積神經網路(CNN)在影像壓縮任務中的應用，並將其性能與 JPEG 進行了比較。實驗結果顯示，儘管 JPEG 在主觀視覺質量方面表現更佳，但基於 CNN 的方法能實現更好的壓縮比率。因此，基於深度學習的影像壓縮算法有望在實際應用中替代傳統技術，特別是在硬件成本和緩衝區大小受限的情況下。

值得強調的是，本文採用的 CNN 模型架構雖然較為簡單，但已經展現出了深度學習方法在影像壓縮領域的潛力。通過設計更加深層和複雜的網路結構，未來可以進一步挖掘神經網路的能力，在壓縮率和視覺品質兩方面都超越傳統算法。同時，更有效的訓練策略、損失函數設計以及硬件加速等技術的應用，都有助於提升深度學習在影像壓縮中的表現。

另一個值得關注的方向是，探索將深度學習影像壓縮技術應用到更廣泛的場景中，比如無線傳輸、雲存儲、視頻壓縮等，發揮其在計算資源和傳輸帶寬有限時的優勢。同時也可以考慮將其與傳統壓縮算法相結合，發揮各自的長處，實現更優的壓縮效果。

總的來說，儘管目前的神經網路模型在某些方面還有待提高，但本研究證實了基於深度學習的影像壓縮算法具有廣闊的發展前景。未來可以進一步優化模型架構和訓練策略，在視覺質量和壓縮比率之間取得更好的平衡，為影像壓縮領域提供更先進、更高效的解決方案，克服傳統技術的局限性。相信隨著算力的不斷提升和深度學習理論的發展，這一領域將會有更多創新性的突破和成果問世。

References

1. J. Balle, V. Laparra, E. P. Simoncelli, "END-TO-END OPTIMIZED IMAGE COMPRESSION", 2017.
2. H. K. Cilingir, Sivaramakrishnan Sankarapandian, M. Tezcan "Image Compression Using Deep Learning", 2017

3. G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, and M. Covell. Full resolution image compression with recurrent neural networks. In CVPR, pages 5435–5443. IEEE, 2017.
4. Haojie Liu, Tong Chen, Qiu Shen, “Deep Image Compression via End-to-End Learning”
5. Johannes Ballé, David Minnen, Saurabh Singh, Sung Jin Hwang, and Nick Johnston. 2018. “Variational image compression with a scale hyperprior.” arXiv:1802.01436 [eess.IV]: bmshj2018-factorized-msssim-[1–8]
6. <https://en.wikipedia.org/wiki/JPEG>
7. <https://github.com/scelesticsiva/Neural-Networks-for-Image-Compression>
8. G. K. Wallace, “The JPEG still picture compression standard,” IEEE Transactions on Consumer Electronics, vol. 38, issue 1, pp.18-34, 1992
9. I. Bocharova, Compression for Multimedia, Cambridge, UK, Cambridge University Press, 2010
10. D. Salomon, Introduction to Data Compression, Springer, 3rd ed., New York, 2004.
11. D. Jian-Jiun, “Advanced Digital Signal Processing”,2023