Deep Learning-Based Image and Video Compression: A List of Publications

Last updated on **Feb 19, 2019**
Edited by Dong Liu and Haichuan Ma.

1

End-to-end **auto-encoder** like schemes (for image)

- E. Agustsson, F. Mentzer, M. Tschannen, L. Cavigelli, R. Timofte, L. Benini, L. V. Gool: Soft-to-hard vector quantization for end-to-end learning compressible representations. NIPS 2017: 1141-1151 [NIPS]
- E. Agustsson, M. Tschannen, F. Mentzer, R. Timofte, L. V. Gool: Generative Adversarial Networks for Extreme Learned Image Compression. arXiv:1804.02958
- M. Akbari, J. Liang, J. Han: DSSLIC: Deep Semantic Segmentation-based Layered Image Compression. arXiv:1806.03348
- J. Ballé, V. Laparra, E. P. Simoncelli: End-to-end optimization of nonlinear transform codes for perceptual quality. PCS 2016: 1-5 [DOI]
- J. Ballé, V. Laparra, E. P. Simoncelli: End-to-end optimized image compression. ICLR 2017 [OpenReview]
- J. Ballé, D. Minnen, S. Singh, S. J. Hwang, N. Johnston: Variational image compression with a scale hyperprior. ICLR 2018 [OpenReview]

2

- C. Cai, L. Chen, X. Zhang, Z. Gao: Efficient Variable Rate Image Compression with Multi-scale Decomposition Network. IEEE Transactions on Circuits and Systems for Video Technology [DOI]
- Z. Cheng, H. Sun, M. Takeuchi, J. Katto: Deep Convolutional AutoEncoder-based Lossy Image Compression. PCS 2018 arXiv:1804.09535
- T. Dumas, A. Roumy, C. Guillemot: Image compression with stochastic winner-take-all auto-encoder. ICASSP 2017: 1512-1516 [DOI]
- T. Dumas, A. Roumy, C. Guillemot.: Autoencoder Based Image Compression: Can The Learning Be Quantization Independent? ICASSP 2018 arXiv:1802.09371
- K. Gregor, Y. LeCun: Learning representations by maximizing compression. arXiv:1108.1169
- K. Gregor, F. Besse, D. J. Rezende, I. Danihelka, D. Wierstra: Towards conceptual compression. NIPS 2016: 3549-3557 [NIPS]

3

- N. Johnston, D. Vincent, D. Minnen, M. Covell, S. Singh, T. Chinen, S. J. Hwang, J. Shor, G. Toderici: Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks. arXiv:1703.10114
- A. B. L. Larsen, S. K. Sønderby, H. Larochelle, O. Winther: Autoencoding beyond pixels using a learned similarity metric. ICML 2016: 1558-1566 [PMLR]
- M. Li, W. Zuo, S. Gu, D. Zhao, D. Zhang: Learning Convolutional Networks for Content-weighted Image Compression. CVPR 2018 [CVF]
- J. Lee, S. Cho, S. K. Beack: Context-Adaptive Entropy Model for End-to-End Optimized Image Compression. ICLR 2019 [OpenReview]
- D. Minnen, J. Ballé, G. Toderici. Joint Autoregressive and Hierarchical Priors for Learned Image Compression. arXiv:1809.02736
- F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, L. V. Gool: Conditional Probability Models for Deep Image Compression. CVPR 2018 [CVF]

4

- O. Rippel, L. Bourdev: Real-time adaptive image compression. ICML 2017: 2922-2930 [PMLR]
- S. Santurkar, D. Budden, N. Shavit: Generative compression. arXiv:1703.01467
- L. Theis, W. Shi, A. Cunningham, F. Huszár: Lossy image compression with compressive autoencoders. ICLR 2017 [OpenReview]
- G. Toderici, S. M. O'Malley, S. J. Hwang, D. Vincent, D. Minnen, S. Baluja, M. Covell, R. Sukthankar: Variable rate image compression with recurrent neural networks. ICLR 2016 arXiv:1511.06085
- G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, M. Covell: Full resolution image compression with recurrent neural networks. CVPR 2017: 5306-5314 [CVF]

5

- Y.-H. Tsai, M.-Y. Liu, D. Sun, M.-H. Yang, J. Kautz: Learning binary residual representations for domain-specific video streaming. AAAI 2018 arXiv:1712.05087
- H. Zhao, P. Liao: CAE-ADMM: Implicit Bitrate Optimization via ADMM-based Pruning in Compressive Autoencoders. arXiv:1901.07196
- Z. Zhang, Z. Chen, J. Lin, W. Li: Learned Scalable Image Compression with Bidirectional Context Disentanglement Network. arXiv:1812.09443

6

End-to-end schemes for video

- S. Kim, J. S. Park, C. G. Bampis, J. Lee, M. K. Markey, A. G. Dimakis, A. C. Bovik: Adversarial Video Compression Guided by Soft Edge Detection. arXiv:1811.10673
- S. Lombardo, J. Han, C. Schroers and S. Mandt: Video Compression through Deep Bayesian Learning. NIPSW 2018 [NIPS]
- G. Lu, W. Ouyang, D. Xu, X. Zhang, C. Cai, Z. Gao: DVC: An End-to-end Deep Video Compression Framework. arXiv:1812.00101
- O. Rippel, S. Nair, C. Lew, S. Branson, A. G. Anderson, L. Bourdev: Learned Video Compression. arXiv:1811.06981
- C.-Y. Wu, N. Singhal, P. Krähenbühl: Video compression through image interpolation. arXiv:1804.06919

7

Generative models

- A. Dosovitskiy, T. Brox: Generating images with perceptual similarity metrics based on deep networks. NIPS 2016: 658-666 [NIPS]
- T. Salimans, A. Karpathy, X. Chen, D. P. Kingma: PixelCNN++: Improving the PixelCNN with Discretized Logistic Mixture Likelihood and Other Modifications. ICLR 2017 [OpenReview]
- S. Santurkar, D. Budden, N. Shavit: Generative Compression. PCS 2018 arXiv:1703.01467
- J. Snell, K. Ridgeway, R. Liao, B. D. Roads, M. C. Mozer, R. S. Zemel: Learning to generate images with perceptual similarity metrics. arXiv:1511.06409

8

- L. Theis, M. Bethge: Generative image modeling using spatial LSTMs. NIPS 2015: 1927-1935 [NIPS]
- A. van den Oord, N. Kalchbrenner, K. Kavukcuoglu: Pixel recurrent neural networks. ICML 2016: 1747-1756 [PMLR]
- A. van den Oord, N. Kalchbrenner, O. Vinyals, L. Espeholt, A. Graves, K. Kavukcuoglu: Conditional image generation with PixelCNN decoders. NIPS 2016: 4790-4798 [NIPS]
- F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, L. V. Gool: Practical Full Resolution Learned Lossless Image Compression. arXiv:1811.12817

9

Inter predictive coding

- H. Choi, I. V. Bajic: Deep Frame Prediction for Video Coding. arXiv:1901.00062
- S. Huo, D. Liu, F. Wu, H. Li: Convolutional Neural Network-Based Motion Compensation Refinement for Video Coding. ISCAS 2018 [DOI]
- E. M. Ibrahim, E. Badry, A. M. Abdelsalam, I. L. Abdalla, M. Sayed, H. Shalaby: Neural Networks Based Fractional Pixel Motion Estimation for HEVC. ISM 2018 [DOI]
- S. Jimbo, J. Wang, Y. Yashima: Deep Learning-based Transformation Matrix Estimation for Bidirectional Interframe Prediction. GCCE 2018 [DOI]
- J. Liu, S. Xia, W. Yang, M. Li, D. Liu: One-for-All: Grouped Variation Network-Based Fractional Interpolation in Video Coding. IEEE Transactions on Image Processing [DOI]
- J. K. Lee, N. Kim, S. Cho, and J. Kang: Convolution Neural Network based Video Coding Technique using Reference Video Synthesis. APSIPA 2018 [PDF]

10

- Y. Wang, X. Fan, C. Jia, D. Zhao, W. Gao: Neural Network Based Inter Prediction for HEVC. ICME 2018 [DOI]
- S. Xia, W. Yang, Y. Hu, S. Ma, J. Liu: A Group Variational Transformation Neural Network for Fractional Interpolation of Video Coding. DCC 2018 [DOI]
- N. Yan, D. Liu, H. Li, F. Wu: A convolutional neural network approach for half-pel interpolation in video coding. ISCAS 2017: 1-4 [DOI]
- Z. Zhao, S. Wang, S. Wang, X. Zhang, S. Ma, J. Yang: CNN-Based Bi-Directional Motion Compensation for High Efficiency Video Coding. ISCAS 2018 [DOI]
- Z. Zhao, S. Wang, S. Wang, X. Zhang, S. Ma, J. Yang: Enhanced Bi-prediction with Convolutional Neural Network for High Efficiency Video Coding. IEEE Transactions on Circuits and Systems for Video Technology [DOI]

11

Intra predictive coding

- M. U. A. Ayoobkhan, E. Chikkannan, K. Ramakrishnan, S. B. Balasubramanian: Prediction-Based Lossless Image Compression. ISMAC 2018 [DOI]
- R. Birman, Y. Segal, A. D. Malka, O. Hadar: Intra prediction with deep learning. SPIE Optical Engineering + Applications 2018 [DOI]
- W. Cui, T. Zhang, S. Zhang, F. Jiang, W. Zuo, Z. Wan, D. Zhao: Convolutional Neural Networks Based Intra Prediction for HEVC. DCC 2017 [DOI]
- Y. Hu, W. Yang, M. Li, J. Liu: Progressive Spatial Recurrent Neural Network for Intra Prediction. arXiv:1807.02232
- J. Li, B. Li, J. Xu, R. Xiong: Intra prediction using fully connected network for video coding. ICIP 2017 [DOI]

12

- J. Pfaff, P. Helle, D. Maniry, S. Kaltenstadler, W. Samek, H. Schwarz, D. Marpe, T. Wiegand: Neural network based intra prediction for video coding. SPIE Optical Engineering + Applications 2018 [DOI]
- I. Schiopu, Y. Liu, A. Munteanu: CNN-based Prediction for Lossless Coding of Photographic Images. PCS 2018 [DOI]
- **Transform coding**
- D. Liu, H. Ma, Z. Xiong, F. Wu: CNN-based DCT-like transform for image compression. MMM 2018: 61-72 [DOI]
- **Entropy coding**
- R. Song, D. Liu, H. Li, F. Wu: Neural network-based arithmetic coding of intra prediction modes in HEVC. VCIP 2017 [DOI]

13

Down/up-sampling-based coding

- M. Afonso, F. Zhang, D. R. Bull: Video Compression based on Spatio-Temporal Resolution Adaptation. IEEE Trans. Circuits and Systems for Video Technology (online) [DOI]
- F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo, D. Zhao: An end-to-end compression framework based on convolutional neural networks. IEEE Trans. Circuits and Systems for Video Technology (online) [DOI]
- M. Jenab, I. Amer, B. Ivanovic, M. Saeedi, Y. Liu, G. Sines, S. Shirani: Content-Adaptive Resolution Control To Improve Video Coding Efficiency. ICMEW 2018 [DOI]

14

- Y. Li, D. Liu, H. Li, L. Li, F. Wu, H. Zhang, H. Yang: Convolutional neural network-based block up-sampling for intra frame coding. IEEE Trans. Circuits and Systems for Video Technology (online) [DOI]
- J. Lin, D. Liu, H. Yang, H. Li, F. Wu: Convolutional Neural Network-Based Block Up-Sampling for HEVC. IEEE Trans. Circuits and Systems for Video Technology [DOI]
- Z. Liu, C. Cui: A New Low Bit-Rate Coding Scheme for Ultra High Definition Video Based on Super-Resolution Reconstruction. CCET 2018 [DOI]

15

- **Saliency-based coding**
- A. Prakash, N. Moran, S. Garber, A. DiLillo, J. Storer: Semantic perceptual image compression using deep convolution networks. DCC 2017: 250-259 [DOI]
- A. Selimović, B. Meden, P. Peer, A. Hladnik: Analysis of Content-Aware Image Compression with VGG16. IWOBI 2018 [DOI]
- **Colorization-based coding**
- M. H. Baig, L. Torresani: Multiple hypothesis colorization and its application to image compression. Computer Vision and Image Understanding 164: 111-123 (2017) [DOI]
- **Inpainting-based coding**
- M. H. Baig, V. Koltun, L. Torresani: Learning to inpaint for image compression. NIPS 2017: 1246-1255 [NIPS]

16

Out-of-loop filtering

- L. Cavigelli, P. Hager, L. Benini: CAS-CNN: A Deep Convolutional Neural Network for Image Compression Artifact Suppression. arXiv:1611.07233
- Y. Dai, D. Liu, F. Wu: A convolutional neural network approach for post-processing in HEVC intra coding. MMM 2017: 28-39 [DOI]
- X. He, Q. Hu, X. Zhang, C. Zhang, W. Lin, X. Han: Enhancing HEVC Compressed Videos with a Partition-Masked Convolutional Neural Network. ICIP 2018 [DOI]
- Z. Jin, P. An, C. Yang, L. Shen: Quality Enhancement for Intra Frame Coding Via Cnns: An Adversarial Approach. ICASSP 2018 [DOI]
- C. Li, L. Song, R. Xie, W. Zhang: CNN based post-processing to improve HEVC. ICIP 2017 [DOI]
- X. Meng, X. Deng, S. Zhu, S. Liu, C. Wang, C. Chen, B. Zeng: MGANet: A Robust Model for Quality Enhancement of Compressed Video. arXiv:1811.09150

17

- X. Song, J. Yao, L. Zhou, L. Wang, X. Wu, D. Xie, S. Pu: A Practical Convolutional Neural Network as Loop Filter for Intra Frame. ICIP 2018 [DOI]
- T. Wang, M. Chen, H. Chao: A novel deep learning-based method of improving coding efficiency from the decoder-end for HEVC. DCC 2017: 410-419 [DOI]
- T. Wang, W. Xiao, M. Chen, H. Chao: The Multi-Scale Deep Decoder for the Standard HEVC Bitstreams. DCC 2018 [DOI]
- R. Yang, M. Xu, Z. Wang: Decoder-side HEVC quality enhancement with scalable convolutional neural network. ICME 2017 [DOI]
- R. Yang, M. Xu, T. Liu, Z. Wang, Z. Guan: Enhancing Quality for HEVC Compressed Videos. arXiv:1709.06734
- R. Yang, M. Xu, Z. Wang, T. Li: Multi-Frame Quality Enhancement for Compressed Video. CVPR 2018 [CVF]
- Y. Yuan, J. Sun, M. Wang: Dilated Deep Residual Network for Post-processing in TPG Based Image Coding. IDCS 2018 [DOI]

18

In-loop filtering

- C. Jia, S. Wang, X. Zhang, S. Wang, S. Ma: Spatial-Temporal Residue Network Based In-Loop Filter for Video Coding. VCIP 2017 arXiv:1709.08462
- C. Jia, S. Wang, X. Zhang, S. Wang, J. Liu, S. Pu, S. Ma: Content-Aware Convolutional Neural Network for In-loop Filtering in High Efficiency Video Coding. IEEE Transactions on Image Processing [DOI]
- J. Kang, S. Kim, K. M. Lee: Multi-modal/multi-scale convolutional neural network based in-loop filter design for next generation video codec. ICIP 2017 [DOI]
- L. Ma, Y. Tian, T. Huang: Residual-Based Video Restoration for HEVC Intra Coding. BigMM 2018 [DOI]

19

- X. Meng, C. Chen, S. Zhu, B. Zeng: A New HEVC In-Loop Filter Based on Multi-channel Long-Short-Term Dependency Residual Networks. DCC 2018 [DOI]
- W.-S. Park, M. Kim: CNN-based in-loop filtering for coding efficiency improvement. IEEE Image, Video, and Multidimensional Signal Processing Workshop (IVMSP) 2016: 1-5 [DOI]
- Y. Zhang, T. Shen, X. Ji, Y. Zhang, R. Xiong, Q. Dai: Residual Highway Convolutional Neural Networks for in-loop Filtering in HEVC. IEEE Transactions on Image Processing [DOI]

20

- **Rate control**
- J. Hu, W. Peng, C. Chung: Reinforcement Learning for HEVC/H.265 Intra-Frame Rate Control. ISCAS 2018 [DOI]
- Y. Li, B. Li, D. Liu, Z. Chen: A convolutional neural network-based approach to rate control in HEVC intra coding. VCIP 2017: 1-4 [DOI]
- **Simultaneous compression and retrieval**
- Q. Zhang, D. Liu, H. Li: Deep network-based image coding for simultaneous compression and retrieval. ICIP 2017: 405-409 [DOI]

21

Fast algorithms for encoding

- H. Amer, A. Rashwan, E. Yang: Fully Connected Network for HEVC CU Split Decision equipped with Laplacian Transparent Composite Model. PCS 2018 [DOI]
- K. CHEN, X. ZENG, Y. FAN: CNN Oriented Fast CU Partition Decision and PU Mode Decision for HEVC Intra Encoding. ICSICT 2018 [DOI]
- M. Hassan, T. ShanablehEmail author: Predicting split decisions of coding units in HEVC video compression using machine learning techniques. Multimed Tools Appl (2018) [DOI]
- T. Katayama, T. Song, T. Shimamoto: QP Adaptation Algorithm for Low Complexity HEVC based on a CNN-Generated Header Bits Map. ICCE-Berlin 2018 [DOI]

22

- Z. Liu, X. Yu, Y. Gao, S. Chen, X. Ji, D. Wang: CU partition mode decision for HEVC hardwired intra encoder using convolution neural network. IEEE Trans. Image Processing 25(11): 5088-5103 (2016) [DOI]
- M. Tahir, I. A. Taj, P. A. Assuncao, M. Asif: Fast video encoding based on random forests. Journal of Real-Time Image Processing [DOI]
- C. Wang, L. Yu, S. Wang: Accelerate CU Partition in HEVC using Large-Scale Convolutional Neural Network. arxiv 2018 [arXiv:1809.08617]
- M. Wang, J. Li, X. Xie, Y. Li, H. Jia: Fast Intra CU Size Decision for HEVC Based on Machine Learning. AMCCE 2018 [DOI]
- J. Xu, M. Xu, Y. Wei, Z. Wang, Z. Guan: Fast H.264 to HEVC Transcoding: A Deep Learning Method. IEEE Transactions on Multimedia [DOI]

23

New References, after 2019

References: CVPR 2021 Workshop papers

[1] Shoma Iwai, et al., "Self Texture Transfer Networks for Low Bitrate Image Compression,"
[2] Fabian Brand, et al., "Rate-Distortion Optimized Learning-Based Image Compression using an Adaptive Hierachical Autoencoder with Conditional Hyperprior ,"
[3] Zhengxue Cheng, et al., "Perceptual Image Compression using Relativistic Average Least Squares GANs,"
[4] Chi D. K. Pham, et al., "Deep Learning based Spatial-Temporal In-loop filtering for Versatile Video Coding,"
[5] Yixin Gao, et. al., "erceptual Friendly Variable Rate Image Compression,"
[6] Akifumi Suzuki, et. al., "Learned Image Compression with Super-Resolution Residual Modules and DISTS Optimization,"
[7] Yung-Han Ho, et al., "End-to-End Learned Image Compression with Augmented Normalizing Flows,"
[8] Sudeep Katakol, et al., "DANICE: Domain adaptation without forgetting in neural image compression,"

24

[9] Khawar Islam, et al., "Image Compression with Recurrent Neural Network and Generalized Divisive Normalization,"
[10] Xining Wang,et al., "Subjective Quality Optimized Efficient Image Compression,"
[11] Yuyang Wu, et al., "Deep Image Compression with Latent Optimization and Piece-wise Quantization Approximation,"
[12] Yanding Peng, et al., "Multi-Metric Fusion Network for Image Quality Assessment,"
[13] Nannan Zou, et al., "Learned Video Compression with Intra-Guided Enhancement and Implicit Motion Information,"
[14] Benoit Brummer, et al., "End-to-end optimized image compression with competition of prior distributions,"
[15] Zhimeng Huang, et al., "Beyond VVC: Towards Perceptual Quality Optimized Video Compression Using Multi-Scale Hybrid Approaches,"
[16] Jing Zhao, et al., "A Universal Encoder Rate Distortion Optimization Framework for Learned Compression,"

25

CLIC-2021 papers:

[1] Xing Zeng, et al., "An Adaptive In-Loop Filter based on Neural Network for Video Coding,"
[2] Zongyu Guo, et al., "Improved Neural Image Compression: A New Two-Stage Quantization Strategy,"
[3] Alexander Lytchier, et al., "Perceptually-Guided Lossy Image Compression,"
[4] Ge Gao, et al., "Attentional Multi-scale Back Projection Network for Low Bit-rate Image Compression,"
[5] Shangyin Gao, et al., "Perceptual Learned Image Compression With Continuous Rate Adaptation,"
[6] Ting Fu, et al., "Quality Enhancement of VVC Intra-frame Coding based on HGRDN,"
[7] Chenhao Wu, "Approach of Team HHC: A Method for Low Bitrate Image Compression "
[8] The´o Ladune, et al., "Conditional Coding and Variable Bitrate for Practical Learned Video Coding,"

26

[9] The´o Ladune, et al., "Coding standards as anchors for the CVPR CLIC video track,"
[10] Jianzhao Liu, et al., "Learned Dual Perceptual Image Similarity for Compressed Image ,"
[11] David Alexandre, et al., "Hybrid Video Coding for CLIC 2021,"
[12] David Alexandre, et al., "Learning-based Video Coding for CLIC 2021,"
[13] Jean Be´gaint, et al., "Hierarchical Temporal Structure for End-to-End Neural Network-based Video Compression,"
[14] Yi Ma, et al., "ROI Perceptual Metric,"
[15] Chih-Peng Chang, et al., "Learned Inter-frame Compression for CLIC 2021,"
[16] Yeda Chen, et al., "MCM: Multi-channel Context Model for Entropy in Generative Image Compression,"

Benedetta Tondi , Andrea Costanzo, Dequ Huang, and Bin Li, "Boosting CNN-based Primary **Quantization matrix estimation of Double JPEG Images** via a **Classification-like Architecture**," *EURASIP Journal on Information Security* (2021) 2021:5 https://doi.org/10.1186/s13635-021-00119-0

27