

# Bibliography

## (Learning-based Multimedia Compression)

1. L. Helming, A. Djelouah, M. Gross, and C. Schroers, "Lossy Image Compression with Normalizing Flows," in *International Conference on Learning Representations Workshop*, 2021.
2. J. Li, B. Li, and Y. Lu, "Deep contextual video compression," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2021.
3. J. Lin, D. Liu, H. Li, and F. Wu, "M-lvc: multiple frames prediction for learned video compression," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 3546–3554.
4. Y.-H. Ho, C.-C. Chan, W.-H. Peng, H.-M. Hang, and M. Domański, "Anfic: Image compression using augmented normalizing flows," *IEEE Open Journal of Circuits and Systems*, vol. 2, pp. 613–626, 2021.
5. M. Mathieu, C. Couprie, and Y. LeCun, "Deep Multi-scale Video Prediction beyond Mean Square Error," in *Proc. International Conference on Learning Representations (ICLR)*, 2016.
6. R. Villegas, J. Yang, S. Hong, X. Lin, and H. Lee, "Decomposing Motion and Content for Natural Video Sequence Prediction," in *Proc. International Conference on Learning Representations (ICLR)*, 2017.
7. Z. Liu, R. A. Yeh, X. Tang, Y. Liu, and A. Agarwala, "Video Frame Synthesis using Deep Voxel Flow," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2017.
8. G. Lu, W. Ouyang, D. Xu, X. Zhang, C. Cai, and Z. Gao, "Dvc: An end-to-end deep video compression framework," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11 006–11 015.
9. G. Lu, X. Zhang, W. Ouyang, L. Chen, Z. Gao, and D. Xu, "An end-to-end learning framework for video compression," *IEEE transactions on Pattern Analysis and Machine Intelligence*, 2020.
10. Z. Hu, Z. Chen, D. Xu, G. Lu, W. Ouyang, and S. Gu, "Improving deep video compression by resolution-adaptive flow coding," in *European Conference on Computer Vision*. Springer, 2020, pp. 193–209.

11. O. Rippel, A. G. Anderson, K. Tatwawadi, S. Nair, C. Lytle, and L. Bourdev, "Elf-vc: Efficient learned flexible-rate video coding," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2021, pp. 14 479–14 488.
12. T. Ladune, P. Philippe, W. Hamidouche, L. Zhang, and O. Déforges, "Optical flow and mode selection for learning-based video coding," in *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2020, pp. 1–6.
13. J. H. at al., "Reinforcement learning for HEVC/H.265 intra-frame rate control," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2018, pp. 1–5.
14. M. Z. et al., "Rate control method based on deep reinforcement learning for dynamic video sequences in HEVC," *IEEE Transactions on Multimedia*, 2020.
15. C. C.etal., "HEVC/H.265 coding unit split decision using deep reinforcement learning," in *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. IEEE, 2017, pp. 570–575.
16. J. Shi and Z.-B. Chen, "Reinforced bit allocation under task-driven semantic distortion metrics," in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2020.
17. L. et al., "Continuous control with deep reinforcement learning," in *Proc. International Conference on Learning Representations (ICLR)*, 2015.
18. M. Zhou, X. Wei, S. Kwong, W. Jia, and B. Fang, "Rate control method based on deep reinforcement learning for dynamic video sequences in HEVC," *IEEE Transactions on Multimedia*, 2020.
19. R. Guangjie, L. Zizheng, C. Zhenzhong, and L. Shan, "Reinforcement learning based ROI bit allocation for gaming video coding in VVC," in *IEEE VCIP*, 2021.
20. F. Jun, H. Chen, and C. Zhibo, "360HRL: Hierarchical reinforcement learning based rate adaptation for 360-degree video streaming," in *IEEE VCIP*, 2021.
21. Y.-H. Ho, G.-L. Jin, Y. Liang, W.-H. Peng, and X. Li, "A dual-critic reinforcement learning framework for frame-level bit allocation in HEVC/H.265," in *DCC*, 2021.
22. J.-L. Lin, W. Hung, S.-H. Yang, P.-C. Hsieh, and X. Liu, "Escaping from zero gradient: Revisiting action-constrained reinforcement learning via Frank-Wolfe policy optimization," *UAI*, 2021.

23. G. Lu, W. Ouyang, D. Xu, X. Zhang, C. Cai, and Z. Gao, "DVC: An End-To-End Deep Video Compression Framework," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
24. D. Minnen, J. Ballé, and G. Toderici, "Joint Autoregressive and Hierarchical Priors for Learned Image Compression," in *Neural Information Processing Systems*, 2018.
25. Y. Hu, W. Yang, and J. Liu, "Coarse-to-Fine Hyper-Prior Modeling for Learned Image Compression," in *AAAI Conference on Artificial Intelligence*, 2020.
26. T. Chen, H. Liu, Z. Ma, Q. Shen, X. Cao, and Y. Wang, "End-to-End Learnt Image Compression via Non-Local Attention Optimization and Improved Context Modeling," *IEEE Transactions on Image Processing*, vol. 30, pp. 3179–3191, 2021.
27. Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Learned Image Compression with Discretized Gaussian Mixture Likelihoods and Attention Modules," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
28. J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-end optimized image compression," in *International Conference on Learning Representations*, 2017.
29. C. Lin, J. Yao, F. Chen, and L. Wang, "A Spatial RNN Codec for End-to-End Image Compression," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
30. J. Ballé, D. Minnen, S. Singh, S. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior," in *International Conference on Learning Representations*, 2018.
31. J. Lee, S. Cho, and S. Beack, "Context-adaptive Entropy Model for End-to-end Optimized Image Compression," in *International Conference on Learning Representations*, 2019.
32. Y. Hu, W. Yang, Z. Ma, and J. Liu, "Learning End-to-End Lossy Image Compression: A Benchmark," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.
33. H. Ma, D. Liu, N. Yan, H. Li, and F. Wu, "End-to-End Optimized Versatile Image Compression with Wavelet-Like Transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
34. L. Dinh, J. Sohl-Dickstein, and S. Bengio, "Density estimation using real NVP," in *International Conference on Learning Representations*. OpenReview.net, 2017. [Online]. Available: <https://openreview.net/forum?id=HkpbnH9lx>

35. C. Huang, L. Dinh, and A. C. Courville, “Augmented Normalizing Flows: Bridging the Gap Between Generative Flows and Latent Variable Models.” *CoRR*, 2020.
36. Y. Choi, M. El-Khamy, and J. Lee, “Variable Rate Deep Image Compression with a Conditional Autoencoder,” in *2019 IEEE/CVF International Conference on Computer Vision*, 2019.
37. D. P. Kingma and M. Welling, “Auto-Encoding Variational Bayes,” 2014.
38. J. Ballé, V. Laparra, and E. Simoncelli, “End-to-end Optimized Image Compression,” in *International Conference on Learning Representations*, 2017.
39. C. Ma, Z. Wang, R. Liao, and Y. Ye, “A cross channel context model for latents in deep image compression.” arXiv, 2021.
40. Z. Guo, Z. Zhang, R. Feng, and Z. Chen, “Causal contextual prediction for learned image compression,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 4, pp. 2329–2341, 2022.
41. A. V. Oord, N. Kalchbrenner, and K. Kavukcuoglu, “Pixel Recurrent Neural Networks,” in *Proceedings of International Conference on Machine Learning*, 2016.
42. Y. Wu, X. Li, Z. Zhang, X. Jin, and Z. Chen, “Learned block-based hybrid image compression,” *IEEE Transactions on Circuits and Systems for Video Technology*, 2021.
43. G. Toderici, W. Shi, R. Timofte, L. Theis, J. Balle, E. Agustsson, N. Johnston, and F. Mentzer, “Workshop and Challenge on Learned Image Compression (CLIC2020),” 2020. [Online]. Available: <http://www.compression.cc>
44. B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan, and J.-R. Ohm, “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.
45. 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, pp. 10 771–10 780, 2018.
46. 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.
47. Z. Hu, G. Lu, and D. Xu, “Fvc: A new framework towards deep video compression in feature space,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 1502–1511.

48. J. Ballé, D. Minnen, S. Singh, S. J. Hwang, and N. Johnston, “Variational image compression with a scale hyperprior,” in *International Conference on Learning Representations*, 2018.
49. T. Chen, H. Liu, Z. Ma, Q. Shen, X. Cao, and Y. Wang, “End-to-end learnt image compression via non-local attention optimization and improved context modeling,” *IEEE Transactions on Image Processing*, vol. 30, pp. 3179–3191, 2021.
50. T. Ladune, P. Philippe, W. Hamidouche, L. Zhang, and O. Déforges, “Conditional coding for flexible learned video compression,” in *Neural Compression: From Information Theory to Applications–Workshop@ICLR 2021*, 2021.
51. H. Liu, M. Lu, Z. Ma, F. Wang, Z. Xie, X. Cao, and Y. Wang, “Neural video coding using multiscale motion compensation and spatiotemporal context model,” *IEEE Transactions on Circuits and Systems for Video Technology*, 2020.
52. R. Yang, F. Mentzer, L. V. Gool, and R. Timofte, “Learning for video compression with hierarchical quality and recurrent enhancement,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 6628–6637.
53. R. Feng, Z. Guo, Z. Zhang, and Z. Chen, “Versatile learned video compression.” arXiv, 2021.
54. R. Yang, F. Mentzer, L. Van Gool, and R. Timofte, “Learning for video compression with recurrent auto-encoder and recurrent probability model,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 2, pp. 388–401, 2020.
55. A. K. Fabian Brand, Jürgen Seiler, “Generalized difference coder: A novel conditional autoencoder structure for video compression,” *arXiv:2112.08011*, 2021.
56. X. Sheng, J. Li, B. Li, L. Li, D. Liu, and Y. Lu, “Temporal context mining for learned video compression,” 2021.
57. A. Golinski, R. Pourreza, Y. Yang, G. Sautiere, and T. S. Cohen, “Feedback recurrent autoencoder for video compression,” in *Proceedings of the Asian Conference on Computer Vision*, 2020.
58. E. P. S. Johannes Ballé, Valero Laparra, “End-to-end optimized image compression,” *International Conference for Learning Representations*, 2017.
59. W. Byeon, Q. Wang, R. K. Srivastava, and P. Koumoutsakos, “ContextVP: Fully Context-Aware Video Prediction,” in *Proc. European Conference on Computer Vision (ECCV)*, 2018.

60. N. Kalchbrenner, A. vanden Oord, K. Simonyan, I. Danihelka, O. Vinyals, A. Graves, and K. Kavukcuoglu, "Video Pixel Networks," in *Proc. International Conference on Machine Learning (ICML)*, 2017.
61. W. Lotter, G. Kreiman, and D. Cox, "Deep Predictive Coding Networks for Video Prediction and Unsupervised Learning," in *Proc. International Conference on Learning Representations (ICLR)*, 2017.
62. C. Finn, I. Goodfellow, and S. Levine, "Unsupervised Learning for Physical Interaction through Video Prediction," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2016.
63. X. Liang, L. Lee, W. Dai, and E. P. Xing, "Dual Motion GAN for Future-flow Embedded Video Prediction," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2017.
64. Patraucean, Viorica, A. Handa, and R. Cipolla., "Spatio-Temporal Video Autoencoder with Differentiable Memory," in *Proc. The International Conference on Learning Representations (ICLR)*, 2016.
65. F. A. Reda, G. Liu, K. J. Shih, R. Kirby, J. Barker, D. Tarjan, A. Tao, and B. Catanzaro, "SDC-Net: Video Prediction using Spatially-displaced Convolution," in *Proc. European Conference on Computer Vision (ECCV)*, 2018.
66. H. Jiang, D. Sun, V. Jampani, M. Yang, E. G. Learned-Miller, and J. Kautz, "SuperSlomo: High Quality Estimation of Multiple Intermediate Frames for Video Interpolation," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
67. Niklaus, Simon, L. Mai, and F. Liu., "Video Frame Interpolation via Adaptive Separable Convolution," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2017.
68. Y.-W. Chen and W.-H. Peng, "Parametric OBMC for Pixel-adaptive Temporal Prediction on Irregular Motion Sampling Grids," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 1, pp. 113–127, 2012.
69. M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial Transformer Networks," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2015.
70. R. S. Sutton, D. A. McAllester, S. P. Singh, and Y. Mansour, "Policy Gradient Methods for Reinforcement Learning with Function Approximation," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2000.
71. V. Mnih, N. Heess, A. Graves, and K. Kavukcuoglu, "Recurrent Models of Visual Attention," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2014.

72. R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
73. G. Lu, C. Cai, X. Zhang, L. Chen, W. Ouyang, D. Xu, and Z. Gao, "Content adaptive and error propagation aware deep video compression," in *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings*.
74. N. Zou, H. Zhang, F. Cricri, H. R. Tavakoli, J. Lainema, M. Hannuksela, E. Aksu, and E. Rahtu, "L2c -learning to learn to compress," in *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, 2020.
75. T. van Rozendaal, J. Brehmer, Y. Zhang, R. Pourreza, and T. S. Cohen, "Instance-adaptive video compression: Improving neural codecs by training on the test set," in *arXiv*, 2021.
76. Y. Yang, R. Bamler, and S. Mandt, "Improving inference for neural image compression," in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2020.
77. Z. Guo, Z. Zhang, R. Feng, and Z. Chen, "Soft then hard: Rethinking the quantization in neural image compression," in *Proc. International Conference on Machine Learning (ICML)*, 2021.
78. N. Yan, C. Gao, D. Liu, H. Li, L. Li, and F. Wu, "Sssic: Semantics-to-signal scalable image coding with learned structural representations," *IEEE Transactions on Image Processing*, vol. 30, pp. 8939–8954, 2021.