ISO/IEC JTC 1/SC 29/WG 1
(ITU-T SG16)

Coding of Still Pictures

JBIG	**JPEG**
Joint Bi-level Image Experts Group	Joint Photographic Experts Group

TITLE: Report on the State-of-the-art of Learning based Image Coding

SOURCE: Requirements
EDITORS: João Ascenso, Pinar Akayzi

STATUS: Final

REQUESTED ACTION: Distribute

DISTRIBUTION: Public

Contact:
ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch

1. Purpose of This Document

In the last few years, several learning-based image coding solutions have been proposed to take advantage of recent advances in this field. The driving force was the development of improved neural network models for several image processing areas, such as denoising and super-resolution, which have taken advantage of the availability of large image datasets and special hardware, such as the highly parallelizable graphic processing units (GPUs).

This document marks the beginning of the work to be conducted in JPEG and has the following objectives:

- Present a list of learning-based image coding solutions from the literature in an organized way.
- Present a list of software implementations of deep learning-based image codecs public available.
- Present a list of data sets that may be used for exploration studies.

Despite the advances, there are several challenges in this type of coding solutions, namely which encoder/decoder architectures are more promising (e.g. recurrent neural networks vs variational auto encoders), which type of processing layers should be used to compress image data and which type of quality metrics should be used for the optimization. Other relevant problems in this area, is how to minimize the impact of quantization for lossy image coding, what should be the strategy used for learning (e.g. a two neural network architecture as in generative adversarial networks) and how to perform bitrate allocation.

2. Relevant Image Coding Solutions

The main objective of this Section is to identify the most relevant learning-based image compression solutions in the literature that allow to demonstrate the potential of this novel coding paradigm. This area has received many contributions in the recent years and is considered critical for the future of image compression solutions. In this Section a list of the most relevant end-to-end learning-based image coding solutions with a short description is presented:

- *Variable Rate Image Compression with Recurrent Neural Networks* [1]: Toderici et al. earlier work in learning based image compression presents deep fully connected, LSTM and convolutional/deconvolutional residual encoders and was applied to thumbnail images. The dimensionality of the latent representation was highly reduced compared to the input image and was quantized using a direct binarization technique.

- *Full Resolution Image Compression with Recurrent Neural Networks* [2]: Toderici et al. presented a full-resolution lossy image compression method based on RNNs, one of the first solutions that are competitive across a wide range of compression rates and that can be applied to images of arbitrary sizes. The recurrent nature of the proposed solution allows to propagate information between iterations of both encoder and decoder; each iteration is associated with some target bitrate and thus quality.

- *Lossy Image Compression with Compressive Autoencoders* [3]: Theis et al. addresses the problem of lossy image compression with autoencoders. Techniques to optimize AEs due to the non-differentiability of the quantization and entropy coding tools are proposed to obtain a trade-off between quality and bitrate with only a set of models.

- *End-to-end Optimized Image Compression* [4]: Ballé et al. proposed an image compression solution with three stages of linear–nonlinear transformation with convolutions; filter responses are divided by a weighted L2-norm of all other filter responses at that spatial location. This local normalization nonlinearity is inspired by local gain control behaviors observed in biological visual systems. The transformed values are uniformly quantized and converted to a bit stream using an arithmetic entropy coder.

- *Variational Image Compression With a Scale Hyperprior* [5]: Ballé et al. takes advantage of variational autoencoders, proposing an end-to-end optimized deep image compression solution. The proposed image coding architecture includes side information in the form of a hyperprior to capture spatial dependencies in the latent representation space in an efficient way.

- *Real-Time Adaptive Image Compression* [6]: Rippel et al. proposed an image compression architecture based on Generative Adversarial Networks (GANs). The most significant contribution is the multiscale adversarial training framework, which combines an adversarial loss with a reconstruction loss. A reduction of compression artifacts is expected for higher compression ratios (sharper images), since the decoder makes reconstructions from statistics learned with uncompressed images.

- *Generative Adversarial Networks for Extreme Learned Image Compression* [7]: Agustsson et al. proposes a deep compression solution based on a generative adversarial network for extreme image compression (i.e. for target bitrates below 0.1 dB) still producing visually satisfying high-resolution images. For these bitrates, it is hard to preserve every element in an image and the solution enables the generation of artificial elements to make it more pleasant, exploiting the availability of side information (i.e. semantic information about the image being coded).

Some of the contributions presented in the literature do not propose an entire full image coder but focus on improving different aspects (mainly tools) of an already presented image compression codec. The contributions in this category can be divided on the different parts of an image compression solution:

- *Quantization*: In [10], instead of learning one model for each quantization step (and thus quality) it was proposed to learn a single model and then apply different quantization strategies, such as adaptive (and learned) quantization step sizes (per channel). In [11], a different quantization method is followed, with aggressive pruning of the latent code to encourage sparsity. In [12], it was proposed a non-uniform quantizer procedure which was optimized based on the distribution of the features. The encoder-decoder network is iteratively fine-tuned with the non-uniform quantizer to obtain the best performance. In [13], a unified end-to-end learning framework that jointly optimizes the model parameters, the quantization levels, and the entropy of the resulting symbol stream is proposed. The method relies on soft-assignments of the data to quantization levels which are learned jointly with the model parameters.

- *Non-linear transforms*: In [14], an end-to-end optimization framework for the efficiency of non-linear transforms assuming scalar quantization is proposed, which can support any perceptual metric. This framework is able to find a set of parameters for the generalized divisive normalization transform and the corresponding inverse. Following this work, in [15], this optimized transformation is analyzed, and it was shown that the correlation between transformed components is reduced compared to alternative solutions. In [16], the invertibility of these transforms is analyzed from the perspective of finding the parameters (training algorithms) that ensure a minimization of the loss function and thus ensure maximum reconstruction fidelity.

- *Entropy coding*: In [17], the key idea is to exploit the statistics of the latent representation with a conditional probability model. In such case, an auto-encoder architecture with a 3D-CNN for context modelling is used. During training, the context model is iteratively updated to learn the dependencies between the elements of the latent representation. In [18], spatially local and adaptive entropy coding models are proposed that are shared between encoder and decoder and for which some type of side-information is transmitted. This allows to have entropy coding models adapted to the characteristics of a specific image. In [19], a context-adaptive entropy model framework that exploits different types of contexts is proposed. Two different types of contexts are used and selected to have a more accurate probability of the distribution of the latent code elements, thus exploiting the spatial correlation very efficiently.

3. Organizing the Most Relevant Image Coding Solutions

This Section aims to organize the most relevant image coding solutions in order to propose a meaningful classification taxonomy for deep-learning coding solutions to more easily identify and abstract their differences, commonalities and relationships. With this organization, each image coding solution may be characterized by a classification path along the dimensions of the proposed taxonomy.

3.1. Proposing a Taxonomy

This section presents and defines the classification dimensions for the taxonomy proposed for learning based image coding solutions. First of all, it is important to highlight that only lossy image coding solutions are addressed in this report, especially because solutions targeting lossless image coding with learning-based methods is very preliminary. The main objective is to identify families of coding solutions that share some common characteristics and that may address the requirements of some typical image coding applications in a better way.

After an exhaustive study of the several image coding solutions available in the literature [1]-[37] it was concluded that the most appropriate taxonomy classification dimensions are:

- *Neural network type:* characterizes the coding solution based on the type of neural network that defines its architecture. Naturally, for this class the most relevant solutions are based on autoencoders which are able to extract a compact representation of an image and that can learn encoder-decoder non-linear transforms that are able to efficiently decorrelate the visual data. However, this class doesn't only include this high-level description of the neural network, but also which variant of the autoencoder is used (e.g. variational autoencoder) and the type of layers used at both encoder and decoder (e.g. convolutional, recurrent). Also, the strategy followed during the learning process could be important to characterize the image coding solution, which could be a simple loss function to more advanced strategies with two neural networks applied sequentially, such as generative adversarial network.
- *Coding unit size:* the size of the coding unit is rather important to characterize the image coding solution and has an important impact on the relevance of the system. First of all, since coding solutions do not exploit correlation among different coding units, this allow some type of spatial access. Also, the neural network network architecture is usually tailored for a specific coding unit size, which means that images that have a resolution that is not multiple of this size need to be adapted, e.g. by padding or some type of resolution adaptation, which may have an impact in coding efficiency.
- *Spatial correlation tools:* refers to the way the spatial correlation is exploited to create a more compact representation of the image. There are several approaches in the literature that can be used to exploit spatial dependencies of the image or in some feature space, such as the use of hyperpriors, which correspond to some side information that is transmitted from encoder to decoder that allows to use more accurate entropy coding models. In the same direction, other relevant possibility is the estimation of entropy coding probability models using past decoded information.
- *Bitrate control strategy:* characterizes the coding solution based on the type of bitrate control used, namely if a model is learned per rate-distortion point for a given quantization step size (fixed) or if it provides the flexibility to code a range of qualities with the same model (adaptive). This is an important requirement, since different models for each quantization step size (decoded quality) are necessary, it may have an impact on the number of decoding qualities that can be offered (assuming that no penalty in coding efficiency is allowed) and in the memory requirements of both encoder and decoder (which need to store all models).
- *Quality range:* characterizes the coding solution based on the target quality that the coding solution can offer. While there are solutions that can target a wide range of qualities, there are other solutions that focus on specific quality ranges such as low bitrates. In this case, it is important to highlight extreme image compression solutions that target very low bitrates and that can produce visually appealing high-resolution images from compact global semantic information. However, this comes at the cost that some regions need to be artificially generated and may not faithfully represent the visual scene.

Other important aspects such as coding efficiency, encoding and decoding complexity are also important to characterize learning-based image coding solutions. However, this require a more in-depth detailed study and will be left for future reports of this ad-hoc group.

3.2. Classifying the Coding Solutions

To exercise the proposed classification taxonomy, this section reviews and classifies some relevant and taxonomically representative learning-based image coding solutions available in the literature. Table 1 shows some selected image coding solutions organized in the above classification dimensions.

Paper	Neural network type	Coding unit size	Spatial correlation tools	Bitrate control strategy	Quality range
[1]	1) Fully-connected AE 2) Recurrent AE 3) Convolutional AE	32×32	Residual coding with two-iteration architecture	Fixed (one model per rate)	From low to high quality
[2]	Recurrent AE with convolutional layers	32×32	Recurrent layers with memory between iterations and entropy coding model	Variable rate/quality with one model	From low to high quality
[3]	Convolutional AE	128×128	Skip connections	Three models for low, medium and high rates	From low to high quality
[4]	Convolutional AE with non-linear transformations	256×256	Non-linear transforms and RDO optimization	One model for each rate/quality point	From low to medium qualities
[5]	Variational AE	256×256	Hyperprior to compute entropy coding model	Four models for different rates	From low to high qualities
[6]	Convolutional AE with pyramid decomposition and GAN training	128×128	Adaptive codelength regularization	?	From low to high qualities
[7]	Convolutional AE with GAN training	Entire image ?	Region of interest coding, exploits semantic data	One model for each rate point ?	Low qualities
[8]	Convolutional + JPEG	8×8	Region of interest detection followed by JPEG coding	One model	From low to high quality
[9]	Convolutional AE	128×128	Content-aware bitrate allocation with importance map	?	From low to high quality
[10]	Convolutional AE	256×256	Quantization step size optimization	One model	From low to high quality
[11]	Convolutional AE	128×128	Pruning of the latent code	One model for each rate point	From low to high quality
[12]	Convolutional AE	256×256	Non-uniform quantizer	One model for each rate point	From low to medium qualities
[20]	Wasserstein GAN + Wasserstein AE	64×64	Network architectures for inference an training	?	From low to high quality

AE = Auto-encoder

GAN = Generative Adversarial Networks.

4. Available Software Implementations

A list with the available software implementations is available in Table 3. Some of the implementations were not done by the authors of the paper and may have different levels of maturity, e.g. cross checking with the paper results could be necessary.

Table 3. Available learning-based image compression implementations.

Paper	Language, Framework and Some Libraries	URL(s)
[2]	Python, Tensorflow	https://github.com/tensorflow/models/tree/master/research/compression [1]
[4]	Pyhton, Tensorflow	https://github.com/tensorflow/compression http://www.cns.nyu.edu/~lcv/iclr2017/
[7]	Python, Tensorflow 1.8, Pandas	https://github.com/Justin-Tan/generative-compression [2]
[8]	Python, Tensorflow, Pandas, PIL, SKimage	https://github.com/iamaaditya/image-compression-cnn
[13]	Python, Tensorflow, Pandas, Pillow, Scikit	https://github.com/fab-jul/imgcomp-cvpr
[9]	Python, Caffe, OpenCV	https://github.com/limuhit/ImageCompression
[10]	Python, Tensorflow	https://www.irisa.fr/temics/demos/visualization_ae/visualizationAE.htm
[11]	Python, Pytorch	https://github.com/JasonZHM/CAE-ADMM
[19]	Python, Tensorflow	https://github.com/JooyoungLeeETRI/CA_Entropy_Model
[20]	Python, Pytorch	https://github.com/mitscha/dplc
[31]	Python, Pytorch Python, Tensorflow	https://github.com/kunalrdeshmukh/End-to-end-compression https://github.com/ppooiiuuyh/ComRecCNN
[3]	Python, Pytorch	https://github.com/alexandru-dinu/cae

[1] Pytorch implementation is available here: https://github.com/1zb/pytorch-image-comp-rnn
[2] Source code will be made available by the authors here: https://data.vision.ee.ethz.ch/aeirikur/extremecompression/

5. Available Image Data Sets

In this Section a list of some available datasets that could be used for training and testing the learning-based image coding solutions are presented (see Table 4). Preference was given to datasets that have a large number of images with almost no distortions or compression artifacts.

Table 4. Available datasets for testing and learning of image coding solutions.

Name	Nº Images	Type	Spatial Resolution	Bit-depth	URL	Description
Kodak	24	Uncompr.	768×512	8 bit	http://r0k.us/graphics/kodak/	Standard set of images widely used as test set in this area.
MCL-JCI[1]	50	Uncompr.	1920×1080	8 bit	http://mcl.usc.edu/mcl-jci-dataset/	The whole set of 50 source images can be classified into ten semantic categories such as people, animals, plants, buildings, water or lake, sky, bridge, transportation vehicles (boats or cars) and indoor. Used for quality evaluation.
Ultra-eye[2]	41	Uncompr.	3840×2160 and 1920× 1080	8 bit	https://mmspg.epfl.ch/downloads/ultra-eye/	Images in the dataset contain small details with large variations, e.g., close ups with variable background, small objects in large landscapes, city skylines, etc. The images also cover a wide variety of scenes, including outdoor and in- door scenes, images

						of nature, people, animals, and historical scenes depicted in paintings.
JPEG-XL	68	Uncompr.	From HD to 4K	8-12 bit	?	Different sets of images used for the evaluation of the new JPEG-XL standard with different resolutions, bit-depths, color space, color subsampling, etc.
RAISE[3]	8156	Uncompr.	4288×2848	12-14 bit raw	http://loki.disi.unitn.it/RAISE/	Collected from 4 photographers over 3 years (2011- 2014). Different scenes and moments in over 80 places in Europe. Used for forensic analysis.
LIVE in the Wild[4]	1162	Distorted	Acquired from different mobile devices cameras	?	http://live.ece.utexas.edu/research/ChallengeDB/index.html	Images of faces, people, animals, close-up shots, wide-angle shots, nature scenes, man-made objects, images with distinct foreground/background configurations and without any object of interest. Used for quality evaluation.
Waterloo Exploration Database[5]	4744	Pristine	?	?	https://ece.uwaterloo.ca/~k29ma/exploration/	Images from 7 categories: human, animal, plant, landscape, cityscape, still-life and transportation.
UCID[6]	1338	Uncompr.	512×384	8 bit	http://video.minelab.tw/DETS/	Standard dataset used for research in image retrieval, namely to study the effect of compression artifacts and compressed domain techniques. Nowadays used for several tasks.
Openimages	9 Million	Distorted	Different resolutions	8 bit	https://github.com/openimages	dataset of ~9 million URLs to images that have been annotated with image-level labels and bounding boxes spanning thousands of classes.
Imagenet	14 Million	Distorted	469×387 in average	8 bit	http://image-net.org/index	Image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images.
McMaster Dataset	8	Uncompr.	?	?	https://www4.comp.polyu.edu.hk/~cslzhang/CDM_Dataset.htm	Color demoisaicking dataset
DIVerse 2K	1000	Pristine	At least 2K	8 bit	https://data.vision.ee.ethz.ch/cvl/DIV2K/	Dataset for example-based single image super-resolution

https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/

[1]L. Jin, J. Y. Lin, S. Hu, H. Wang, P. Wang, I. Katsavounidis, A. Aaron and C.-C. Jay Kuo, “Statistical Study on Perceived JPEG Image Quality via MCL-JCI Dataset Construction and Analysis.”, Electronic Imaging, 2016, San Jose, California, United States.

[2]H. Nemoto, P. Hanhart, P. Korshunov, T. Ebrahimi, “Ultra-Eye: UHD and HD images eye tracking dataset. Sixth International Workshop on Quality of Multimedia Experience (QoMEX), Singapore, September 2014.

[3]D.-T. Dang-Nguyen, C. Pasquini, V. Conotter, G. Boato, “RAISE – A Raw Images Dataset for Digital Image Forensics”, ACM Multimedia Systems, Portland, USA, March 2015.

[4]D. Ghadiyaram, A.C. Bovik, "Massive Online Crowdsourced Study of Subjective and Objective Picture Quality,” IEEE Transactions on Image Processing, vol. 25, no. 1, pp. 372-387, January 2016.

[5]K. Ma and Z. Duanmu, Q. Wu, Z. Wang, H. Yong, H. Li, L. Zhang, “Waterloo Exploration Database: New Challenges for Image Quality Assessment Models”, IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 1004-1016, February 2017.

[6]G. Schaefer, M. Stich, “UCID: an uncompressed color image database”, Electronic Imaging, 2004, San Jose, California, United States.

[7]Krasin I., Duerig T., Alldrin N., Ferrari V., Abu-El-Haija S., Kuznetsova A., Rom H., Uijlings J., Popov S., Veit A., Belongie S., Gomes V., Gupta A., Sun C., Chechik G., Cai D., Feng Z., Narayanan D., Murphy K. OpenImages: A public dataset for large-scale multi-label and multi-class image classification, 2017.

Other possibility is to design a dataset using images obtained from the internet (see Table 5). This solution is also possible since nowadays there are a lot of websites with royalty-free images that can be used for research. The main disadvantage is that images are typically not uncompressed and thus, may have some coding artifacts, which means that a careful selection of images must be made.

Table 5. Websites with images that can be used to create a dataset for learning and evaluation of image coding solutions.

Web-site	Description	URL	Copyright
ISO-Republic	Stock photos	https://isorepublic.com/	Creative Commons Zero license
Pexels	Stock photos	https://www.pexels.com/royalty-free-images/	Creative Commons Zero license
Unsplash	Stock photos	https://unsplash.com/public-domain-images	https://unsplash.com/license
Pixabay	Stock photos	https://pixabay.com/	https://pixabay.com/service/license/
Public domain	Stock photos	publicdomainpictures.net	Creative Commons Zero license
Europeana	Cultural heritage photos	https://www.europeana.eu/portal/en	Mostly free to use
Wikimedia Commons	Wikimedia foundation project	https://commons.wikimedia.org/wiki/Main_Page	Mostly free to use

6. References

[1] 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," *International Conference on Learning Representations*, San Juan, Puerto Rico, May 2016.

[2] G. Toderici, D. Vincent, N. Johnston, S.J. Hwang, D. Minnen, J. Shor, M. Covell, "Full Resolution Image Compression with Recurrent Neural Networks," *IEEE Conference on Computer Vision and Pattern Recognition*, Aug. 2016.

[3] L. Theis, W. Shi, A. Cunningham, F. Huszár, "Lossy Image Compression with Compressive Autoencoders," *International Conference on Learning Representations*, Toulon, France, April 2017.

[4] J. Ballé, V. Laparra, E. P. Simoncelli, "End-to-end Optimized Image Compression," *International Conference on Learning Representations*, Toulon, France, April 2017.

[5] J. Ballé, D. Minnen, S. Singh, S. J. Hwang, N. Johnston, "Variational Image Compression With a Scale Hyperprior," *International Conference on Learning Representations*, Vancouver, Canada, April 2018.

[6] O. Rippel and L. Bourdev, "Real-Time Adaptive Image Compression," *International Conference on Machine Learning*, Sydney, Australia, May 2017.

[7] E. Agustsson, M. Tschannen, F. Mentzer, R. Timofte, L. Van Gool, "Generative Adversarial Networks for Extreme Learned Image Compression," *arXiv:1804.02958*, October 2018.

[8] A. Prakash, N. Moran, S. Garber, A. Dilillo, J. Storer, "Semantic Perceptual Image Compression Using Deep Convolution Networks," *Data Compression Conference*, Salt Lake City, USA, March 2017.

[9] M. Li, W. Zuo, S. Gu, D. Zhao, D. Zhang, "Learning Convolutional Networks for Content-weighted Image Compression," *IEEE International Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, June 2018.

[10] T. Dumas, A. Roumy, C. Guillemot, “Autoencoder based Image Compression: can the Learning be Quantization Independent?,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, Calgary, Canada, April 2018.

[11] H. Zhao, P. Liao, “CAE-ADMM: Implicit Bitrate Optimization via ADMM-Based Pruning in Compressive Autoencoders”, https://export.arxiv.org/abs/1901.07196

[12] J. Cai and L. Zhang, “Deep Image Compression With Iterative Non-Uniform Quantization,” *IEEE Internatioonal Conference on Image Processing*, Athens, Greece, October 2018.

[13] 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,” *Neural Information Processing Systems*, Long Beach, USA, December 2017.

[14] J. Ballé, V. Laparra, E. P. Simoncelli, “End-to-end Optimization of Nonlinear Transform Codes for Perceptual Quality,” *Picture Coding Symposium*, Nuremberg, Germany, December 2016.

[15] J. Ballé, V. Laparra, E. P. Simoncelli, “Density Modeling Of Images using a Generalized Normalization Transformation,” *International Conference on Learning Representations*, Toulon, France, April 2017.

[16] J. Ballé, “Efficient Nonlinear Transforms for Lossy Image Compression,” *Picture Coding Symposium*, San Francisco, USA, June 2018.

[17] F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, L. Van Gool, “Conditional Probability Models for Deep Image Compression,” *IEEE International Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, June 2018.

[18] D. Minnen, G. Toderici, S. Singh, S. J. Hwang, M. Covell, “Image-Dependent Local Entropy Models for Learned Image Compression,” *International Conference on Image Processing*, Athens, Greece, October 2018.

[19] J. Lee, S. Cho, S.-K. Beack, “Context-adaptive Entropy Model for End-to-end Optimized Image Compression”, International Conference on Learning Representations, New Orleans, USA, May 2019.

[20] M. Tschannen, E. Agustsson, M. Lucic, “Deep Generative Models for Distribution-Preserving Lossy Compression”, Advances in Neural Information Processing Systems 31, Montreal, Canada, September 2018.

[21] T. Dumas, A. Roumy, C. Guillemot, “Image Compression With Stochastic Winner-Take-All Auto-Encoder,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, New Orleans, USA, March 2017.

[22] M. Akbari, J. Liang, J. Han, “DSSLIC: Deep Semantic Segmentation-based Layered Image Compression,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, May 2019.

[23] D. Minnen, J. Ballé, G. Toderici, “Joint Autoregressive and Hierarchical Priors for Learned Image Compression,” *Advances in Neural Information Processing Systems*, no. 31, 2018.

[24] L. Zhou, C. Cai, Y. Gao, S. Su, J. Wu, “Variational Autoencoder for Low Bit-rate Image Compression,” *CVPR Workshop and Challenge On Learned Image Compression*, Salt Lake City, June 2018.

[25] M. H. Baig, V. Koltun, L. Torresani, “Learning to Inpaint for Image Compression,” *Advances in Neural Information Processing Systems*, Long Beach, CA, USA.

[26] S. Luo, Y. Yang, Y. Yin, C. Shen, Y. Zhao, M. Song, “DeepSIC: Deep Semantic Image Compression,” *International Conference on Neural Information Processing*, Siem Reap, Cambodia, December 2018.

[27] A. Mousavi, G. Dasarathy, R. G. Baraniuk, “DeepCodec: Adaptive Sensing and Recovery via Deep Convolutional Neural Networks,” *Annual Allerton Conference on Communication, Control, and Computing*, Illinois, USA, October 2017.

[28] 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,” *International Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, June 2018.

[29] M. Covell, N. Johnston, D. Minnen, S. J. Hwang, J. Shor, S. Singh, D. Vincent, G. Toderici, "Target-Quality Image Compression with Recurrent, Convolutional Neural Networks," *arXiv:1705.06687*, May 2017.

[30] D. Minnen, G. Toderici, M. Covell, T. Chinen, N. Johnston, J. Shor, S. J. Hwang, D. Vincent, S. Singh, "Spatially Adaptive Image Compression using a Tiled Deep Network," *International Conference on Image Processing*, Beijng, China, September 2017.

[31] F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo, D. Zhao, "An End-To-End Compression Framework Based on Convolutional Neural Networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, August 2017.

[32] Z. Cheng, H. Sun, M. Takeuchi, J. Katto, "Deep Convolutional AutoEncoder-based Lossy Image Compression," *Picture Coding Symposium*, San Francisco, USA, June 2018.

[33] A. G. Ororbia, A. Mali, J. Wu, S. O'Connell, D. Miller, C. L. Giles, "Learned Iterative Decoding for Lossy Image Compression Systems," *arXiv:1803.05863*, November 2018.

[34] F. Hussain and J. Jeong, "Efficient Deep Neural Network for Digital Image Compression Employing Rectified Linear Neurons," Journal of Sensors, vol. 2016, no. 3184840, 2016.

[35] S. Santurkar, D. Budden, N. Shavit, "Generative Compression," *Picture Coding Symposium*, San Francisco, USA, June 2018.

[36] Z. Cheng, H. Sun, M. Takeuchi, J. Katto, "Performance Comparison of Convolutional AutoEncoders, Generative Adversarial Networks and Super-Resolution for Image Compression,", *CVPR Workshop and Challenge On Learned Image Compression*, Salt Lake City, June 2018.

[37] C. Aytekin, X. Ni, F. Cricri, J. Lainema, E. Aksu, M. Hannuksela, "Block-optimized Variable Bit Rate Neural Image Compression," *CVPR Workshop and Challenge On Learned Image Compression*, Salt Lake City, June 2018.