DEEP JOINT SOURCE-CHANNEL CODING OF IMAGES WITH FEEDBACK

David Burth Kurka and Deniz Gündüz

Department of Electrical and Electronic Engineering, Imperial College London

ABSTRACT

We consider wireless transmission of images in the presence of channel output feedback, by introducing an autoencoder-based deep joint source-channel coding (JSCC) scheme. We achieve impressive results in terms of the end-to-end reconstruction quality for fixed length transmission, and in terms of the average delay for variable length transmission. To the best of our knowledge, this is the first practical JSCC scheme that can fully exploit channel output feedback, demonstrating yet another setting in which modern machine learning techniques can enable the design of new and efficient communication methods that surpass the performance of traditional structured coding schemes.

Index Terms— Deep neural networks, Feedback, Joint source-channel coding, Wireless image transmission

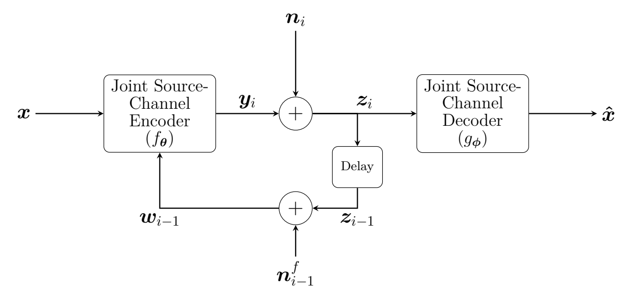


Fig. 1. Communication in the presence of noisy channel output feedback. Source $\boldsymbol{x}$ is transmitted over a noisy forward channel, where after the transmission of each symbol $\boldsymbol{y}_i$, a noisy version of the received signal $\boldsymbol{z}_i$ becomes available to the encoder through a feedback link.

1. INTRODUCTION

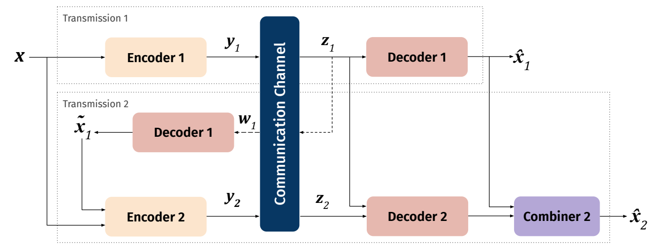
Most of the existing communication systems today are designed based on a two-step process which first compress source samples into bits and then transmit these compressed bits over a channel using a channel code. This is strongly motivated by Shannon's separation theorem [1], one of the fundamental results of information theory, which establishes that even if those steps are optimized separately, there is no loss in the system's general optimality. This principle allows a modular approach in which an application layer (responsible for processing the information source) and a physical layer (responsible for dealing with channel coding and modulation) can be designed independently and combined, for different sources and channel conditions. In fact, current systems employ highly specialized source codes for different types of information sources, e.g., JPEG2000/BPG for images, MPEG-4/WMA for audio, or H.264 for video, and highly optimized channel coding and modulation techniques to be used over different communication channels e.g., Turbo, LDPC, polar codes. Both the compression codecs and the channel codes have been optimized over many decades and gone through many generations of standards.

Despite its huge impact, optimality of separation holds only under unlimited delay and complexity; and, even under these assumptions, it breaks down in multi-user scenarios [2], or non-ergodic source or channel distributions [3, 4]. Still, the lack of powerful and practical joint source-channel coding (JSCC) schemes with reasonable complexity have prevented the adoption of alternative schemes. Although there have been many research efforts on JSCC, they mostly focused either on theoretical analysis under idealistic source and channel distributions [5–7], or the joint optimization of the components of an inherently separate design [8–10]. Another fundamental result in information theory, again due to Shannon [11], states that feedback does not increase the capacity of a memoryless communication channel. It can also be shown that optimality of separation continues to hold in the presence of feedback; therefore, information theoretically feedback does not help the end-to-end performance of source-channel coding either. On the other hand, feedback is known to improve error exponents for channel coding [12, 13], and to significantly simplify the design of joint source-channel coding schemes, at least in some ideal scenarios [14–16]. However, practical schemes have had limited success in converting the theoretical gains of feedback in practice, and only a handful of papers have studied the problem of JSCC with feedback (e.g., [17–19]).

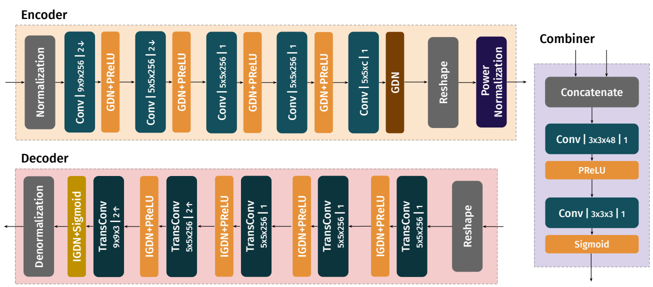
Our goal in this paper is to design practical JSCC scheme that can directly compete with separation-based strategies, by exploiting noisy or noiseless channel output feedback. We build upon the recent success of deep neural networks (DNNs) both for coding and communication problems, such as channel decoding [20, 21], or end-to-end code design [22–24], as well as image compression [25–27]. Most related prior work to the current paper are [28–31], which consider the JSCC problem, and propose autoencoder-based solutions for end-to-end optimization. In [29], we proposed the *deepJSCC* scheme, capable of achieving performance par with state-of-the-art separation-based digital schemes, while also providing *graceful degradation* with the signal-to-noise ratio (SNR). In [30], we also demonstrated that deepJSCC can transmit a source in multiple stages with virtually no losses, being thus almost *successively refinable*.

To the best of our knowledge, this is the first practical implementation of JSCC in the presence of channel output feedback, able to transmit a large content such as an image. We show how a pure machine learning approach can perform,

This work was supported by the European Research Council (ERC) through project BEACON (No. 677854).



(a)



(b)

Fig. 2. (a) Communication scheme used in this work, detailing individual blocks and inputs/outputs for the $L = 2$ case. (b) Architecture of individual blocks with hyperparameters. Convolutions are parametrized by $h \times w \times o \mid s \updownarrow$, where h and w are the filter height, width, o the number of channel outputs and s the stride, that can be either upsampling ($\uparrow$) or downsampling ($\downarrow$) strides.

advancing the state of the art. We present a novel scheme, called *deep-JSCC-f*, that: (1) significantly outperforms ideal separation based schemes (BPG followed by capacity achieving code); (2) enables variable length transmission, greatly reducing the average bandwidth required for a target quality; (3) is robust to noise in the feedback channel and to variations in channel quality (both forward and feedback), providing graceful degradation and analog behavior.

2. PROBLEM FORMULATION

We consider the wireless transmission of images using feedback. An input image with height H, width W and C color channels, represented as a vector of pixel intensities $\boldsymbol{x} \in \mathbb{R}^n$; $n = H \times W \times C$, is to be transmitted in k uses of a noisy communication channel, where k/n is defined as the *bandwidth ratio*. In this work, we consider a complex additive white Gaussian noise (AWGN) channel, modeled as $\boldsymbol{z}_i = \boldsymbol{y}_i + \boldsymbol{n}_i$, where $\boldsymbol{y}_i \in \mathbb{C}$ denotes the complex channel input at time i, $\boldsymbol{z}_i \in \mathbb{C}$ the corresponding complex channel output, and $\boldsymbol{n}_i$ is the independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian noise component with zero mean and variance σ^2. We consider the presence of channel output feedback, through which a noisy version of the channel output at the receiver is made available to the transmitter with a unit delay. Hence, the feedback signal available to the transmitter at time i is denoted by $\boldsymbol{w}_i = \boldsymbol{z}_i + \boldsymbol{n}_i^f$, where $\boldsymbol{n}_i^f$ denotes the i.i.d. additive complex Gaussian noise on the feedback link with zero mean and variance σ_f^2 (see Fig. 1).

2.1. DeepJSCC-f Architecture

In terms of architecture, a key innovation of this work is the use of layered autoencoders, which allow us to take advantage of feedback. The proposed design models each layer j, $j = 1, \dots, L$, as a convolutional autoencoder. Motivated by the well-known Schalkwijk-Kailath scheme [14, 15] for the transmission of a single Gaussian source sample over several channel uses in the presence of perfect channel output feedback, we divide the transmission of each image $\boldsymbol{x}$ into L layers, where each layer tries to improve the quality of the receiver's estimation by transmitting additional information about the residual error. Let layer j be allocated k_j channel uses, where $\sum_{j=1}^{L} k_j = k$. For the transmission of layer j, a channel input vector of $\boldsymbol{y}_j \in \mathbb{C}^{k_j}$ is transmitted over the forward channel, resulting in the channel output vector $\boldsymbol{z}_j$.

Each layer of the proposed deepJSCC-f architecture consists of the following components: (a) an encoder, (b) a decoder, and (c) a combiner. Each transmission i contains a copy of all these components (apart from the combiner, that is only used for layers $j \geq 2$). See Fig. 2a for the deepJSCC-f architecture for $L = 2$. The encoder at layer j, $f_j^{\boldsymbol{\theta}_j}$ is modeled as a CNN parametrized by vector $\boldsymbol{\theta}_j$ and receives as input both the source image ($\boldsymbol{x}$) and an estimation of the image reconstructed by the receiver at the previous layer ($\tilde{\boldsymbol{x}}_{j-1}$), i.e., $\boldsymbol{y}_j = f_j^{\boldsymbol{\theta}_j}(\boldsymbol{x}, \tilde{\boldsymbol{x}}_{j-1})$ (when $j = 1$, only $\boldsymbol{x}$ is used as input). The receiver employs a decoder at each layer, which uses all the channel outputs received so far. The decoder of the j-th layer, $g_j^{\boldsymbol{\phi}_j}$, is a CNN parametrized by $\boldsymbol{\phi}_j$, i.e. $\hat{\boldsymbol{u}}_j = g_j^{\boldsymbol{\phi}_j}(\boldsymbol{z}_1, \dots, \boldsymbol{z}_j)$. Then, for $j \geq 2$, a combiner network $c_j^{\boldsymbol{\psi}_j}$ is used to combine the reconstruction of previous layers: $\hat{\boldsymbol{x}}_j = c_j^{\boldsymbol{\psi}_j}(\hat{\boldsymbol{x}}_{j-1}, \hat{\boldsymbol{u}}_j)$, where $\hat{\boldsymbol{x}}_1 = \hat{\boldsymbol{u}}_1$. Finally, we assume that all the trained decoder and combiner parameters $\boldsymbol{\phi}_j, \boldsymbol{\psi}_j$, are known both at the receiver and the transmitter, and are used in the estimation $\tilde{\boldsymbol{x}}_{j-1} = c_j^{\boldsymbol{\psi}_{j-1}}(g_{j-1}^{\phi_{j-1}}(\boldsymbol{w}_{j-1}))$ for $j > 2$, and $\tilde{\boldsymbol{x}}_1 = g_1^{\phi_1}(\boldsymbol{w}_1))$. The source image $\boldsymbol{x}$ and the estimate $\tilde{\boldsymbol{x}}_{j-1}$ are concatenated on the channel axis when sent as input of the next layer.

We assume that the forward and the feedback channels are independent of each other, and the $\boldsymbol{y}_i$ sequences are transmitted over independent realizations of the channel. The specific architecture of each component is given by Fig. 2b, consisting of CNN layers, followed by normalization obtained by the generalized normalization transformations (GDN/IGDN) [32], followed by a parametric ReLU (PReLU) [33] or sigmoid activation. An unit average power constraint is imposed on the the transmission. The parameter c in the encoder's last CNN layer is responsible for defining the dimension of the channel input k_i. Both the forward and feedback channels are modeled as non-trainable layers. The model is trained gradually, layer by layer; each layer aims to minimize the average distance between the input image $\boldsymbol{x}$ and its partial reconstruction $\hat{\boldsymbol{x}}_i$. Upon convergence, layer parameters are fixed and additional layers are trained, using the previous channel output as feedback information.

3. EXPERIMENTAL RESULTS

The performance of deepJSCC-f is evaluated in different scenarios, and compared with benchmarks. Results from this section were trained and evaluated on the CIFAR-10 dataset [34] and all plotted results are average values obtained from 10 realizations of the channel for every image on the evaluation dataset. The model was implemented in Tensorflow and optimized using the Adam algorithm [35]. We used a learning rate of 10^{-4} and a batch size of 128.

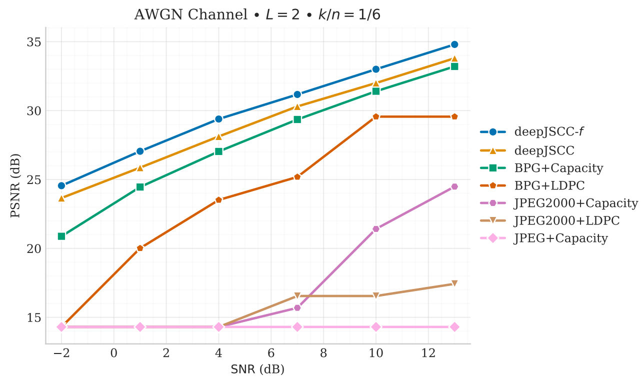


Fig. 3. PSNR vs channel SNR comparison for $k/n = 1/6$.

Models were trained until the performance of a validation set (distinct from the training and test datasets) stops decreasing. As loss function, we considered the mean squared error between the input ($\boldsymbol{x}$) and outputs ($\hat{\boldsymbol{x}}_i$). The performance is measured with the peak signal-to-noise ratio (PSNR), defined as $\text{PSNR} = 10 \log_{10} \frac{255^2}{||\boldsymbol{x}-\hat{\boldsymbol{x}}_i||^2}(dB)$. To measure the quality of a channel in which communication is performed, we consider the average signal-to-noise ratio (SNR) given by $\text{SNR} = 10 \log_{10} \frac{1}{\sigma^2}(dB)$.

3.1. deepJSCC-f with Two Layers ($L = 2$)

We first consider perfect channel output feedback, and assume that the source image $\boldsymbol{x}$ is transmitted in two layers: first a base layer $\boldsymbol{y}_1$ with bandwidth k_1 is sent over the channel; then, using the channel output corresponding to $\boldsymbol{y}_1$, $\boldsymbol{w}_1 = \boldsymbol{z}_1$, a second message $\boldsymbol{y}_2$, of length k_2 is transmitted. For simplicity, we set $k_1 = k_2$. Fig. 3 shows the results for compression rate $k/n = 1/6$, performed in two stages with $k_1/n = k_2/n = 1/12$. The performance of models subject to different forward channel SNRs is presented in the plot, where a separate set of encoders, decoders, and combiners is trained for each SNR.

We first compare our scheme's performance with the state of the art JSCC scheme for image transmission, *deepJSCC* [29], which does not exploit the feedback. We consider an improved version of deepJSCC, by employing the same architecture presented in Fig. 2b. This model is equivalent to our scheme with $L = 1$; that is, a single transmission with a bandwidth ratio $k/n = 1/6$. We show that the use of feedback brings considerable performance improvement, outperforming the deepJSCC by a wide margin, especially in the low SNR regime, when the feedback information is more relevant. We also compare the performance with separation-based digital schemes, where images are first compressed using state of the art compression codecs, and then encoded by a channel code. For the source code, we consider well established image compression algorithms: JPEG, JPEG2000 and BPG. For fair comparison, we remove the header information from the compressed files, so we only consider the communication of compressed bits. For the channel code, we consider a practical scheme – low-density parity-check code (LDPC) followed by quadrature amplitude modulation (QAM) – and a theoretical bound – channel capacity. For LDPC+QAM, we evaluate the performance of different code rates and modulation combinations, presenting here the envelope of the best performing configurations. The capacity achieving code would be a hypothetical channel code, achieving the same rate as the underlying channel capacity, representing an upper bound.

We note here that, since feedback does not increase the capacity of the channel, this upper bound is the same with or without feedback. We see in Fig. 3 that JPEG, currently the most popular and most widely employed image compression codec, presents the worst performance, not being able to compress with enough quality in low SNRs; while BPG is the best performing algorithm. We note that deepJSCC-f surpass even the best performing separation-based scheme highlighting the improvement from JSCC.

3.2. deepJSCC-f with Multiple Layers ($L \geq 2$)

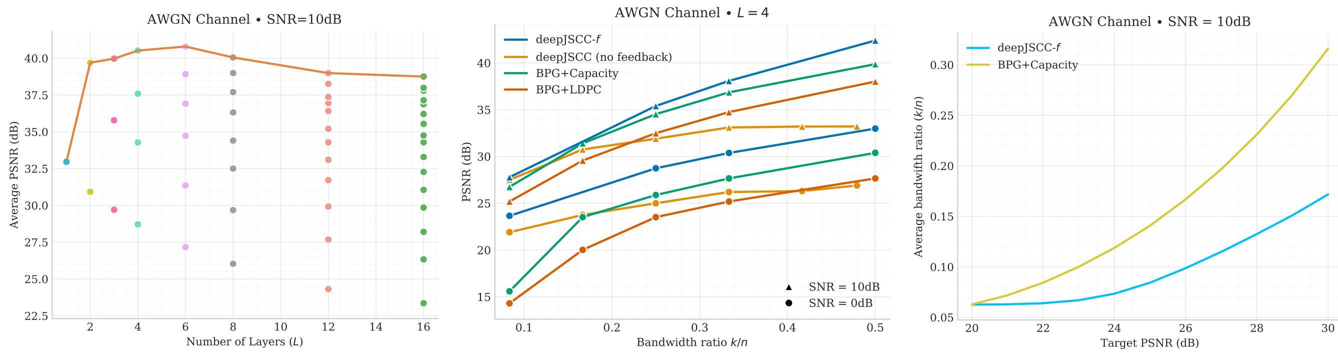
Next, we investigate the impact of increasing the number of channel output feedback uses, by increasing the number of layers (L). Fig. 4a shows the performance for a fixed bandwidth ratio of $k/n = 1/2$, transmitted in different numbers of layers L, with regular size $k_i/n = 1/2L$. We also mark the average PSNR achieved by each intermediate layer.

We see that the use of more layers initially increases the performance. However, as more layers are introduced, the performance stabilizes and even declines for $L > 6$. This trade-off can be explained by the fact that, the increase of layers reduces the size of the partial code, hampering the capacity of the encoder in transmitting relevant information regarding the whole image. Also, increasing L directly increases the complexity of the model, as we need to train a separate set of neural networks (encoder, decoder and a combiner) for each layer. Our simulation results suggest that $L = 4$ layers typically provides a reasonable performance trade-off. In Fig. 4b we present $L = 4$ results for a wider range of compression ratios and two different channel SNRs. We observe that deepJSCC outperforms all other benchmarks, even the separation based bound (which becomes looser as the bandwidth ratio gets smaller), at all settings.

3.3. Variable Length Transmission

We can reformulate the JSCC problem by setting a certain quality target for the delivery of each image, and aim at minimizing the required channel bandwidth. In this case, the perfect channel output feedback provides the transmitter with the knowledge of the stopping time. It is shown in [6] through theoretical analysis that allowing variable rate coding leads to a significant improvement in the delay-distortion trade-off. Since, in our model, the receiver reconstructs the image at each layer, the encoder knows exactly whether it needs to send further information, or it can stop transmission when the distortion target is met.

We experiment this setting by considering the transmission with $L = 8$ layers, and computing the average bandwidth needed for achieving a target PSNR. We compare this to a digital scheme transmitting headerless BPG with an ideal capacity-achieving code. For the digital scheme, we compress each image to the minimum amount of bits that meet the target PSNR value, and find how many channel uses is needed to transmit so many bits over the channel, assuming a capacity-achieving channel code. We should again highlight the fact that this bound is particularly loose when the image can be



(a) (b) (c)

Fig. 4. (a) Impact of feedback channel uses in the performance. (b) Performance comparison for different bandwidth ratios, for $L = 4$. (c) The average bandwidth ratio to achieve a specific target PSNR.

transmitted with only a few layers as this would correspond to a very short blocklength. In Fig. 4c we plot the average bandwidth ratio required to meet different target SNR values. Significant gap between the two curves confirm the theoretical results in this practical setting. We also observe that the gap increases with the PSNR target. For a PSNR target of 30 dB, the average bandwidth ratio is almost half that is required by the bound on the separation-based schemes.

3.4. Noisy Feedback

Lastly, we consider the impact of noisy channel output feedback, i.e., $\sigma_f^2>0$. This is a particularly challenging setting as the encoder cannot track the quality of the receiver's reconstruction accurately, and hence, cannot steer the decoder to the right decision as efficiently as possible. Indeed it is known that the known schemes with theoretical performance guarantees [14,15] break down even with a slightly noisy feedback.

Fig. 5 shows the performance for different feedback channel SNRs and the single layer model (no feedback). The SNR of the feedback channel, SNR^{fb}, is measured in terms of the channel input, i.e., $1/\sigma_f^2$. We can see that our model is robust to noise in the feedback channel. When $\text{SNR}^{fb} = 20$ dB, the performance is only slightly below the noiseless feedback transmission. As SNR decreases, the performance degrades, but overall remains quite high and competitive. The network can learn to make good use of the feedback even with $SNR^{fb} = 10$ dB. When the feedback channel is very noisy ($SNR^{fb} = 0$ dB), the transmission of additional layers still positively contributes to the refinement of the reconstruction, but the overall performance is below than what can be achieved by a single transmission, i.e., $L = 1$. However, a hybrid approach could be proposed, in which the encoder having access to both the original source image and the previous reconstructions, can decide either to use or not the channel output feedback information at each transmission. An implementation of such hybrid solution is left for future work.

Finally, we also highlight that deepJSCC-f is robust to perturbations in both the feedback and the forward channel. It was observed that, in situations in which a model is trained at a specific channel SNR (forward or feedback), but evaluated at a lower SNR, our model can still operate apart from gradual decline in performance, thus presenting analog behavior. This *graceful degradation*, already demonstrated for

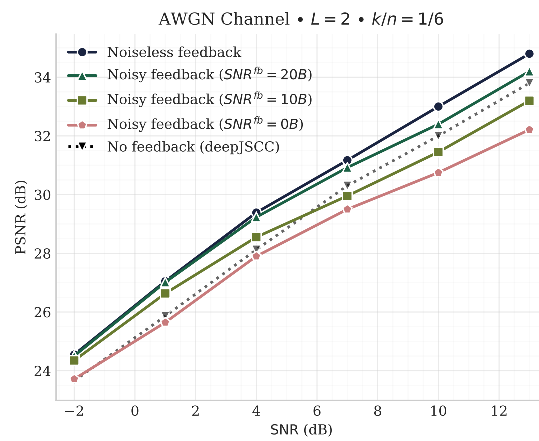


Fig. 5. Model performance when feedback channel has noise.

the forward channel in [29], contrasts with the typical *cliff effect* present in separation-based digital schemes.

4. CONCLUSION

In summary, this work presented, to the best of the authors' knowledge, the first practical implementation of a DNN-aided JSCC scheme with feedback for images, deepJSCC-f. We have shown that deepJSCC-f can achieve considerable gains in performance, compared to (a) JSCC without feedback; (b) state-of-the-art image compression codecs followed by practical and high performing channel codes; and (c) ideal capacity achieving channel codes. Our experiments reveal that the use of the feedback channel improves the quality of the transmission, justifying the adoption of a multi-step strategy for image transmission, in which a source is sent over multiple layers, exploiting the feedback between transmissions. Moreover, a flexible variable length coding scheme is also presented, allowing a considerable economy of resources when a target quality goal is set. Apart from the direct benefits of exploiting the feedback information, we also show that deepJSCC-f has other advantages, such as exhibiting analog behavior and graceful degradation in case of variations in the system, being able to adapt to either forward or feedback channel variations.

5. REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[2] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, Berkeley, CA, 1961, vol. 1, pp. 611–644.

[3] S. Vembu, S. Verdu, and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Transactions on Information Theory*, vol. 41, no. 1, pp. 44–54, Jan 1995.

[4] D. Gunduz and E. Erkip, "Joint source-channel codes for MIMO block-fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 116–134, Jan 2008.

[5] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[6] V. Kostina, Y. Polyanskiy, and S. Verdú, "Joint source-channel coding with feedback," *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 3502–3515, June 2017.

[7] F. Hekland, P. A. Floor, and T. A. Ramstad, "Shannon-Kotel'nikov mappings in joint source-channel coding," *IEEE Transactions on Communications*, vol. 57, no. 1, pp. 94–105, January 2009.

[8] G. Cheung and A. Zakhor, "Bit allocation for joint source/channel coding of scalable video," *IEEE Transactions on Image Processing*, vol. 9, no. 3, pp. 340–356, March 2000.

[9] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 799–809, July 1990.

[10] M. Skoglund, "On channel-constrained vector quantization and index assignment for discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2615–2622, Nov 1999.

[11] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.

[12] M. V. Burnashev, "Data transmission over a discrete channel with feedback: Random transmission time," *Problems of Information Transmission*, vol. 12, no. 4, pp. 250–265, 1976.

[13] J. Schalkwijk, "A coding scheme for additive noise channels with feedback–ii: Band-limited signals," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 183–189, April 1966.

[14] J. Schalkwijk and L. Bluestein, "Transmission of analog waveforms through channels with feedback," *IEEE Transactions on Information Theory*, vol. 13, no. 4, pp. 617–619, October 1967.

[15] T. Kailath, "An application of shannon's rate-distortion theory to analog communication over feedback channels," *Proceedings of the IEEE*, vol. 55, no. 6, pp. 1102–1103, June 1967.

[16] S. Butman, "Rate distortion over band-limited feedback channels," *IEEE Transactions on Information Theory*, vol. 17, no. 1, pp. 110–112, January 1971.

[17] V. Kafedziski, "Joint source channel coding of images over frequency selective fading channels with feedback using dct and multicarrier block pulse amplitude modulation," in *Asilomar Conf. on Signals, Systems and Computers*, Nov 1998, pp. 37–41.

[18] H. Jafarkhani V. Chande and N. Farvardin, "Joint source-channel coding of images for channels with feedback," in *IEEE Information Theory Workshop*, Feb. 1998.

[19] J. Lu, A. Nosratinia, and B. Aazhang, "Progressive joint source-channel coding in feedback channels," in *Data Compression Conference (DCC)*, March 1999, pp. 140–148.

[20] Y. Jiang, S. Kannan, H. Kim, S. Oh, H. Asnani, and P. Viswanath, "Deepturbo: Deep turbo decoder," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019, pp. 1–5.

[21] M. Ebada, S. Cammerer, A. Elkelesh, and S. ten Brink, "Deep learning-based polar code design," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2019, pp. 177–183.

[22] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, Dec 2017.

[23] H. Kim, Y. Jiang, S. Kannan, S. Oh, and P. Viswanath, "Deepcode: Feedback codes via deep learning," in *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., pp. 9436–9446. Curran Associates, Inc., 2018.

[24] Y. Jiang, H. Kim, H. Asnani, S. Kannan, S. Oh, and P. Viswanath, "Turbo autoencoder: Deep learning based channel codes for point-to-point communication channels," in *Advances in Neural Information Processing Systems 32*, 2019.

[25] J. Ballé, D. Minnen, S. Singh, S. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior," in *International Conference on Learning Representations (ICLR)*, 2018.

[26] David Minnen, Johannes Ballé, and George D Toderici, "Joint autoregressive and hierarchical priors for learned image compression," in *Advances in Neural Information Processing Systems*, 2018, pp. 10771–10780.

[27] J. Lee, S. Cho, and S. Beack, "Context-adaptive entropy model for end-to-end optimized image compression," in *International Conference on Learning Representations (ICLR)*, 2019.

[28] N. Farsad, M. Rao, and A. Goldsmith, "Deep learning for joint source-channel coding of text," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018.

[29] E. Bourtsoulatze, D. Burth Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 567–579, Sep. 2019.

[30] D. B. Kurka and D. Gündüz, "Successive refinement of images with deep joint source-channel coding," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2019, pp. 1–5.

[31] R. Zarcone, D. Paiton, A. Anderson, J. Engel, H. S. P. Wong, and B. Olshausen, "Joint source-channel coding with neural networks for analog data compression and storage," in *2018 Data Compression Conference*, March 2018, pp. 147–156.

[32] J. Ballé, V. Laparra, and E. P. Simoncelli, "Density modeling of images using a generalized normalization transformation," *arXiv preprint arXiv:1511.06281*, 2015.

[33] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[34] A. Krizhevsky, "Learning multiple layers of features from tiny images," Tech. Rep., University of Toronto, 2009.

[35] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, 2015.