

Received April 2, 2020, accepted May 10, 2020, date of publication May 21, 2020, date of current version June 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2994393

An Extended Hybrid Image Compression Based on Soft-to-Hard Quantification

HAISHENG FU^{ID}, FENG LIANG^{ID}, AND BO LEI

School of Microelectronics, Xi'an Jiaotong University, Xi'an 710049, China

Corresponding author: Feng Liang (fengliang@xjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61474093, and in part by the Natural Science Foundation of Shaanxi Province, China, under Grant 2020JM-006.

ABSTRACT Recently, the deep learning methods have been widely used in lossy compression schemes, greatly improving image compression performance. In this paper, we propose an extended hybrid image compression scheme based on soft-to-hard quantification, which has only two layers. The compact representation of the input image is encoded by the FLIF codec as the base layer. The residual of the input image and the reconstructed image is encoded by the BPG codec as the enhancement layer. The results using the Kodak and Tecnick datasets show that the performance of our proposed methods exceeds some image compression schemes based on deep learning methods and some traditional coding standards including BPG in SSIM metric across a wide range of bit rates, when the images are coded in the RGB444 domain. We explore the issue of bit rates allocation of the base layer and enhancement layer and the impact of enhancement layer codecs. Also, we analyze the limitations of the hybrid coding scheme.

INDEX TERMS Extend hybrid image compression scheme, soft-to-hard quantification, bit rates allocation, enhancement layer codecs.

I. INTRODUCTION

Recently, deep neural networks (DNNs) have been applied in various areas and have made tremendous progress resulting from their superior optimization and representation learning performance. Traditional compression coding standards such as JPEG [21] and JPEG2000 [22] effectively improve image compression quality by adopting different handcrafted transforms. Since they are optimized separately optimization on codecs, the reconstructed image will appear blocking, blurring, or ring-ing artifacts at low bit rates.

Most end-to-end lossy image compression schemes based on deep learning methods [1], [2], [13], [15] were proposed to handle these issues and achieved better compression performance and visual quality than JPEG2000 and the H.265/HEVC-based BPG image codec [18]. In [9], a scheme based on long short-term memory (LSTM)-based recurrent neural networks (RNNs) was adopted to compress thumbnail images. Better SSIM results than JPEG and WebP were reported. In [5], Toderici *et al.* further presented a full-resolution compression scheme for compressing images

The associate editor coordinating the review of this manuscript and approving it for publication was Kan Zheng^{ID}.

of any size by using the progressive encoding and decoding method.

The learning-based image compression methods mainly consist of three operations: transformation, quantization, and entropy coding. The transformation operation can map the input image data to a latent coding space, so as to better handle the problem of spatial statistics among the image data and make use of exploiting aspects of human perception. Early transforms operation [19], [20] are linear and invertible and have a fixed bit rate. The error of the lossy compression framework only comes from the quantization operation. The transforms based on deep neural networks (DNNs) aim for more comprehensible and nonlinear representations to reduce data redundancy. These transforms can be learned to a latent coding space representation of image raw data and this processing is non-invertible. Most works aimed at the performance of the transforms by introducing the new operations. In [13]–[15], the generalized divisive normalization (GDN) operator was adopted into encoding and decoding networks for joint nonlinearity and spatial adaptability. The GDN operator was proven to be very suitable for natural image compression and can improve the performance of transformation.

The end-to-end lossy compression schemes based on deep learning methods require every operation to be derivable, but the quantization operation is a non-differentiable process. In [17], the quantization of the auto-encoder is replaced by a smooth approximation in the back-propagation because the derivative of the rounding function is zero almost everywhere, but the forward pass of the back-propagation still use the quantization to prevent the decoder from learning to invert the smooth approximation. Agustsson *et al.* [1] introduced a soft-to-hard quantification method with an approximate softmax function. This method will give a randomly initialized vector, and the number of elements in the vector represents the quantization level. The elements in this vector are continuously updated during network training, so that the output of the encoder is better quantized to a specific element of the vector.

Existing methods usually use space-invariant bit length allocation, but the content of the image is generally spatially variant. That is, the areas with complex and saliency structures of an image are more important. Therefore, it is reasonable to adopt different numbers of bits to encode different regions in the image. But in the process of image compression, identifying and distinguishing different regions of the image is very complicated. In order to alleviate this issue, many works present the spatially variant bit allocation method to replace the fixed rate allocation of the image. In [11], [24], Li *et al.* introduced an importance map subnet to produce importance map from the input image to guide bit allocation. The importance map can study the complex and salient structures information of the local image content. The complex areas will be automatically allocated more bits and fewer bits to smooth areas under the guidance of the importance map. These methods can generate more detailed information at the same bit rate.

Many works utilized different entropy models to further improve the performance of the compression schemes. In [15], a variational image compression scheme based on a powerful entropy model was presented. The scale parameters of the latent representation were established on a hyperprior learned by a hyper autoencoder. The hyperprior, as side information, is used to capture the spatial dependency. This scheme was further extended in [23]. The hyperpriors were combined with autoregressive priors generated by a single PixelCNN layer. The performance of the scheme outperformed BPG codec on both PSNR and MS-SSIM metrics. In [24], Lee *et al.* presented a joint optimization of image compression and quality enhancement networks to further improve the the coding efficiency. Also, Zhou *et al.* [25] introduced a context-adaptive scheme in which multi-scale masked convolutional networks were used for the autoregressive model.

Recently, Akbari *et al.* [2] and Fu *et al.* [4] combine deep learning and traditional standard coding to form a hybrid image coding framework for improving the compression performance. The hybrid lossy image coding scheme only needs to train a basic model, which can encode

traditional coding quantization factors to achieve different compression bits of the network. In [2], a three-layer hybrid lossy compression framework named DSSLIC based on a deep semantic segmentation was proposed. A segmentation map of the input image and a compact representation were encoded by the FLIF [8] codec as the base layer. The residual image of the original image and the reconstructed image was encoded by the BPG codec as the enhancement layer. The performance of DSSLIC outperforms the BPG codec in the RGB444 domain in both PSNR and MS-SSIM [6] metrics. Fu *et al.* [4] simplified the hybrid compression framework by removing the semantic segmentation layer based on [2]. The convolutional neural networks were adopted to generate the coarse reconstructed image in [4] instead of the conditional GANs [16].

In this paper, we further develop an extended lossy hybrid image compression scheme based on [2], [4]. The scheme only has two layers as [4]. The compact representation C of the input image is encoded by the FLIF as the base layer. The residual of the original image and reconstructed image is encoded by the BPG codec as the enhancement layer. Our contributions are as follows. First, we adopt the soft-to-hard [1] quantization method to take the place of the round quantization method in [2], [4] and effectively reduces the number of bits occupied by the base layer. Secondly, we explore the bit allocation of the base layer and the enhancement layer by changing the numbers of the quantization level L and the impact of the entire compression framework. Thirdly, we explore the impact of different codecs on the enhancement layer. Forth, the performance of our proposed method exceeds [2] in MS-SSIM metric.

The organization of the paper is as follows. The proposed scheme, the detailed structure of the compression framework, Quantification and residual coding are introduced in Sec. II. In Sec. III, we evaluate the performance of the proposed scheme by comparing it with some compression framework based on deep learning methods including [2], [4] and some traditional codecs including BPG, JPEG, JPEG2000 and WebP codecs in the RGB444 domain. We also conduct a series of ablation experiments to explore the problem of hybrid image compression schemes. Conclusions are given in Sec. IV.

II. THE PROPOSED ENHANCED HYBRID IMAGE COMPRESSION FRAMEWORK

In this section, we propose a hybrid lossy image coding framework based on the soft-to-hard quantification method. To begin with, we present the detailed hybrid framework including a base layer and an enhancement layer. Then, we introduced in detail the network structure of CompNet and RecNet, the quantification method and residual codecs. Finally, we introduced the related settings of network training, including the setting of the loss function, the network structure and training parameters.

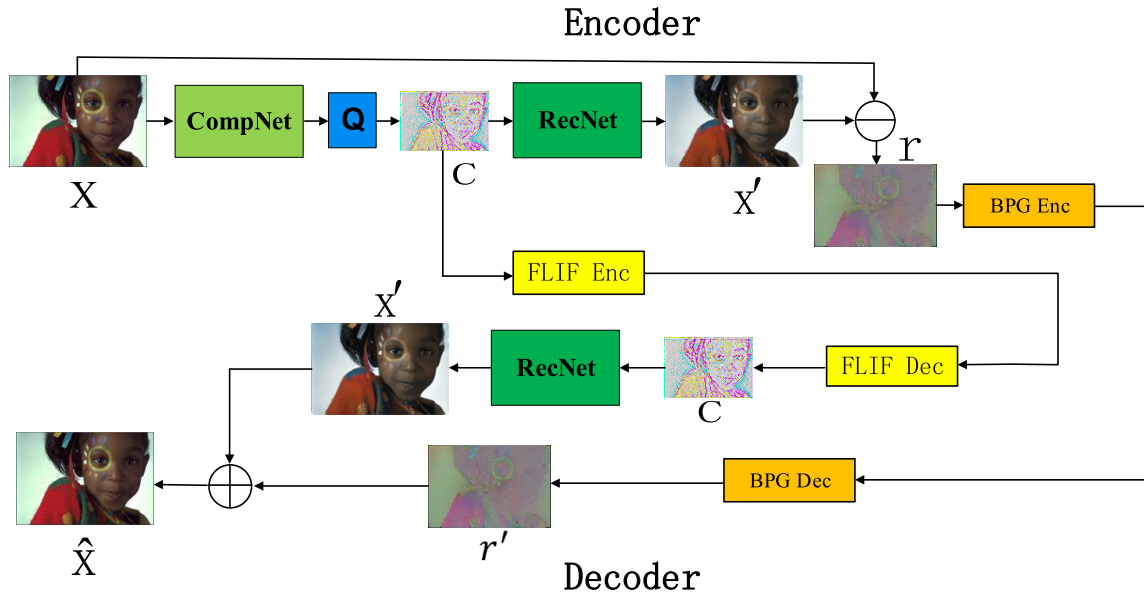


FIGURE 1. A scheme of the proposed improved hybrid image compression network based on soft-to-hard quantification.

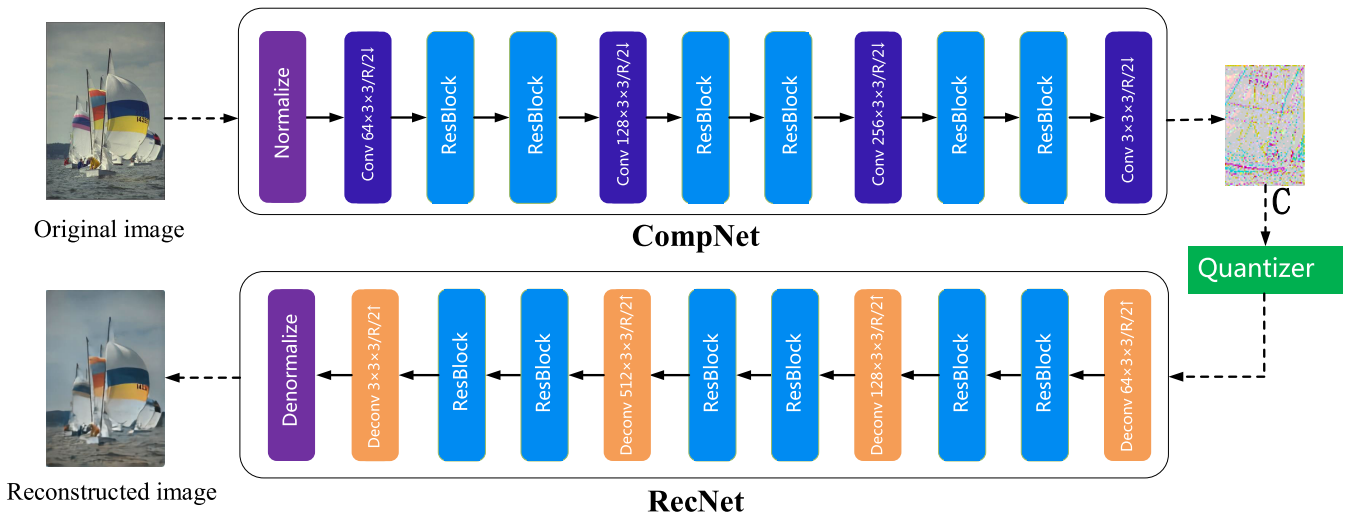


FIGURE 2. The detailed structure of the CompNet and RecNet. R represents rectified linear units (ReLU) activation function.

A. THE STRUCTURE OF THE PROPOSED ENHANCED HYBRID IMAGE COMPRESSION SCHEME

This section will give a detailed description of the proposed scheme illustrated in Fig. 1. In the encoding stage, it includes two layers. A convolutional neural network named CompNet is first adopted to learn the compact representation C of the input image x , which is a multi-layered representation of the original image. Each pixel value of the compact representation C will be quantified into a value in vector L . The compact representation C is encoded by the FLIF codec as the base layer. Next, the compact representation C is used by another deep learning network named RecNet to produce a coarse reconstruction x' of the input image. After that, the residual r between the input image and the coarse reconstruction image is obtained and encoded by the lossy BPG codec.

In the decoding side, the compact representation C' is decoded, which is sent to the RecNet to obtain the coarse reconstruction x' . The residual image is decoded by the BPG decoder, denoted by r' . Finally, the coarse reconstruction x' and the decoded residual image r' are added to get the final reconstruction \hat{x} .

B. NETWORK ARCHITECTURE

The whole network architecture mainly includes consists of two networks, i.e., compression network and reconstruction network. They are called CompNet and RecNet respectively and are shown in Fig.2. The size of the input RGB image is $W \times H \times 3$. In the CompNet, the normalization operation is utilized to transform the image pixels from $[0, 255]$ to $[0, 1]$, thereby making the network easier to train and faster

to converge. A convolution layer with a step size of 2 is adopted to downsample the image three times, and the feature channels are 64, 128, and 256 respectively. In the middle of the two downsampling layers, two residual blocks [7] are utilized to further increase the performance of the network. In the last layer of the CompNet network, a convolutional layer is used to change the feature channel back to 3.

The RecNet decodes the code map back to a reconstructed image. This network is the reverse of the CompNet, except that the downsampling convolution layers are replaced by transposed-convolution (upsampling) layers. The channel sizes used in the upsampling convolution layers are 64, 128, 256 respectively. Also, In the last layer of the RecNet network, a convolutional layer is adopted to change the feature channel back to 3. The denormalization layer needs to be used to remap the output value of the feature map back to [0, 255]. All the kernel sizes of the convolutional layers and deconvolutional layers are set to 3×3 . Also, all the activation functions are set to rectified linear units (ReLU).

C. QUANTIZER

Quantization is the main component in lossy image compression. In the hybrid coding scheme, the quantization quality of compact images directly affects the number of bits occupied by the base layer and the quality of reconstructed coarse images. The [2] utilized the round quantization operation to map the output value of the network from a floating-point number in $[-1, 1]$ to an integer value in $[0, 255]$. The detailed mapping function is shown in Eq. 1.

$$x_i = \text{round}(x_f * 128 + 127), \tag{1}$$

where x_f is $(-1, 1]$. The quantized value is then further encoded by FLIF [8] codec to reduce the number of bits occupied by the compact image. However, this quantitative method has the following limitations. First, the value of the compact image needs to be mapped to an integer value in $[0, 255]$. therefore, Each pixel of the compact image is need to adopt at least 8 bits for Representing raw data. This will increase the number of bits occupied by the base layer, and reduce the number of bits occupied by the enhancement layer when the code rate is fixed. Second, in order to enable the network to train, this round quantization ignored the errors caused by the conversion of the floating-point number in $[-1, 1]$ to the integer $[0,255]$ and the errors caused by an integer value in $[0, 255]$ retransformed to a floating-point number in $[-1, 1]$. This would make the performance of the reconstructed coarse image slightly lower.

To effectively solve this problem, we introduce a new quantization method called soft-to-hard [1] in the hybrid image coding scheme. Compared to [2], the output of the CompNet will quantize in a scalar which has a finite floating point number. Given a number of quantization levels $L = \{l_1, l_2, \dots, l_n\} \subset \mathbb{R}$, each pixel of the compact image C need to quantify to a specific value in the scalar, which can be

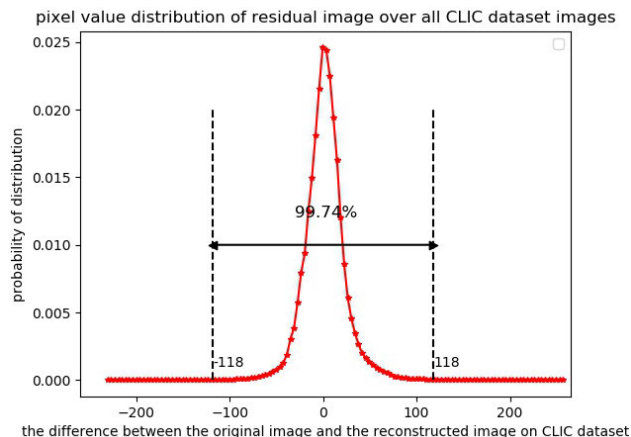


FIGURE 3. Distribution of the residual image pixels over the CLIC dataset images.

updated as the network is trained and be achieved in Eq.3.

$$Q(z_i) = \text{argmin}_j \|z_i - l_j\|. \tag{2}$$

z_i represents the element to be quantified of the compact image C . and argmin represents nearest neighbor assignments algorithm. Since this formula is not differentiable, it can only be used in the process of network forward propagation. Therefore, the formula is deformed to better solve the problem of non-differentiable. The “soft quantization”

$$\tilde{Q}' = \sum_{j=1}^L \frac{\exp(-\sigma_q \|z_i - l_j\|)}{\sum_{k=1}^n \sigma_q \|z_i - l_k\|} l_j \tag{3}$$

can be utilized to calculate the gradient during back-propagation. σ_q represents a hyperparameter and is set to 1. The exp represents exponential function.

D. RESIDUAL CODING

Residual coding as an enhancement layer of our proposed scheme can further improve the performance of the scheme. Most lossy image compression methods need to train different groups of models to meet the needs of different bit rates. They need to change the network structure or change some hyperparameters to control the bit rate. As [4], we analyzed the distribution of our residuals, as shown in Fig.3. Our experimental results show that in most cases the residual image pixel values approximately obey a normal distribution with a mean around zero and very small variance. That is, many residuals are very close to 0, and most residuals are within a small range, hence the adaptive method can preserve more details of the residual, and yields better performance than the simple shift method, the results are shown in Fig.4 and Fig.5. This covers 99.7% of the residuals within in $[-118, 118]$. Therefore, we set the maximum value to 112 and the minimum value to -118 in Eq.7. We will discuss this further in the experiment section.

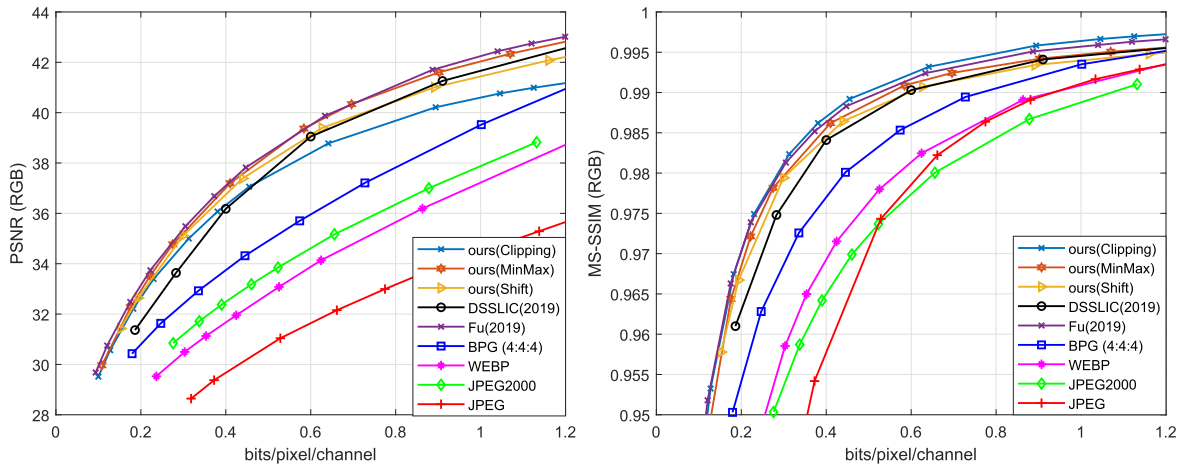


FIGURE 4. Average comparison results on all 24 Kodak images in RGB444 domain in terms of PSNR and MS-SSIM.

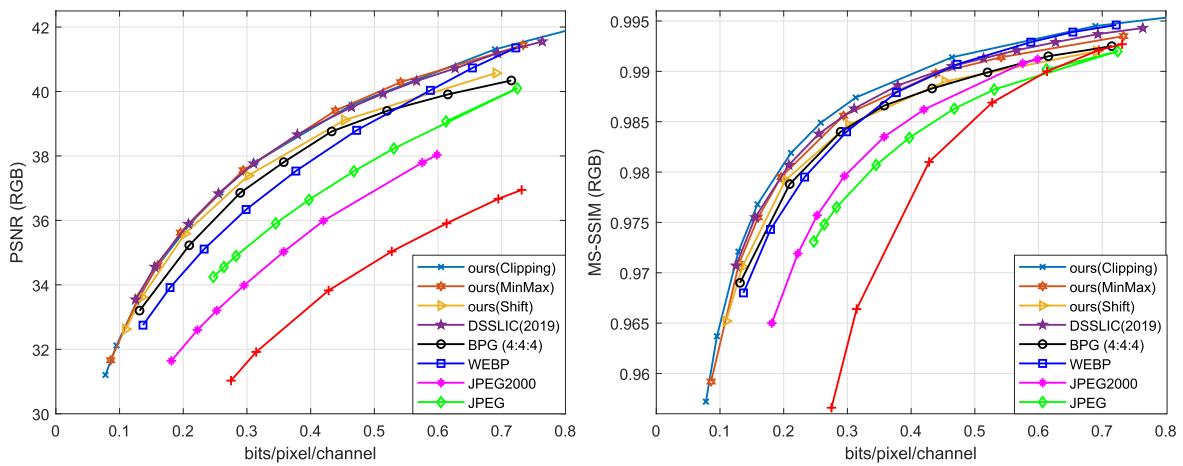


FIGURE 5. Average comparison results on all 100 Tecnick images in RGB444 domain in terms of PSNR and MS-SSIM.

E. LOSS FUNCTION

In this paper, we consider distortion loss and regularization loss as the loss function to train our network.

1) DISTORTION LOSS

The distortion loss function mainly includes peak signal-to-noise ratio (PSNR) and multi-scale structural similarity index (MS-SSIM) metrics. They primarily evaluate the errors between the original image and the reconstructed image. The optimization goal of our scheme is MS-SSIM metric [6], therefore we use such function:

$$L_{MS-SSIM(x, \hat{x})} = \beta(1 - MS-SSIM(x, \hat{x})) \tag{4}$$

to improve the MS-SSIM metric. The β is a hyperparameter that can adjust the proportion of distortion loss function.

2) REGULARIZATION LOSS

In order to better promote the training of the network and prevent the network from overfitting, we adopt the L2 norm to constrain the weight of the network.

3) MODEL OBJECTIVE

We assume that X is a training set and $X_B = \{x^{(1)}, \dots, x^{(B)}\}$ is some minibatches randomly taken from X . We consider the distortion loss function and the regularization loss function together, the final optimization objective function is as follows:

$$L = \frac{1}{B} \sum_{j=1}^B [L_{MS-SSIM(x, \hat{x})} + L_2]. \tag{5}$$

F. TRAINING SET

As discussed before, the RecNet and CompNet are jointly trained to make the reconstructed coarse image which is the output of the RecNet as close as possible to the input image. In the training process, we only need to consider the upper branch of Fig.1. RecNet has the same parameters in the lower and upper branches. Similarly, we do not need to consider the impact of the FLIF codec in our scheme during the training process. The reason is as follows. Since FLIF is a free lossless

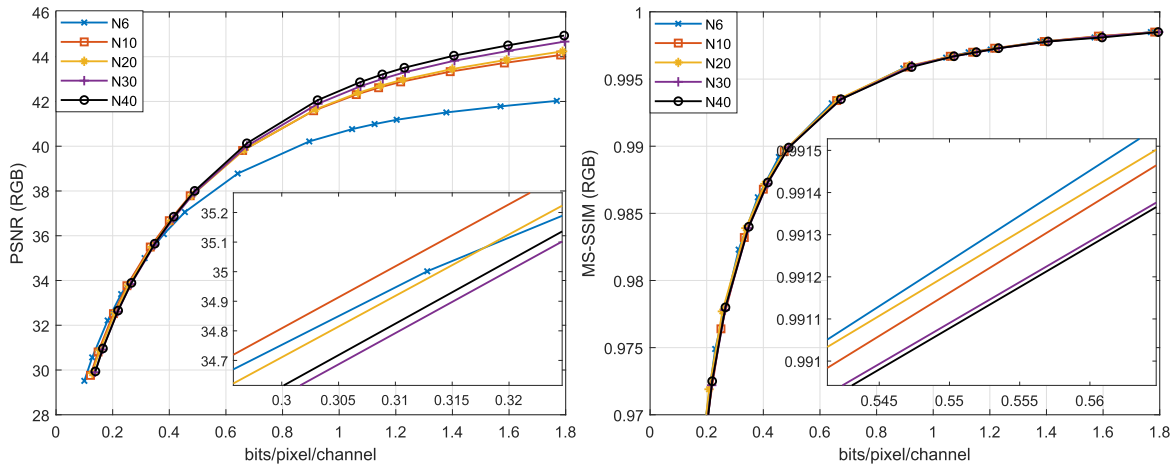


FIGURE 6. Average comparison results on all 24 Kodak images in RGB444 domain in terms of PSNR and MS-SSIM of different numbers of the quantization level L .

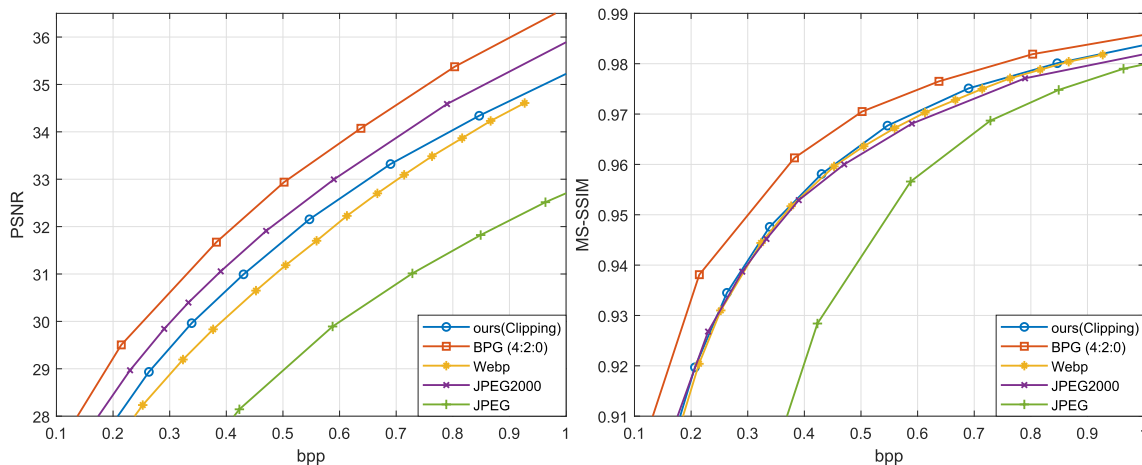


FIGURE 7. Average comparison results on all 24 Kodak images in YUV domain in terms of PSNR and MS-SSIM.

image format. The data does not change before and after the codec, so we can skip this part during the training process.

The CLIC compression dataset [26] which is released by the Computer Vision Lab of ETH Zurich is adopted to train our scheme. In order to better train our proposed image compression scheme, we adopt random rotation and cropped image enhancement techniques to process the training set. The 163300 patches are extracted from 1633 high-quality images. These images are saved as lossless PNGs to train the network better.

III. EXPERIMENT

In this section, we evaluate the performance of our proposed method with traditional codecs including BPG, JPEG2000, WebP, and JPEG and lossy compression schemes based on deep learning methods on both Kodak dataset [27] and Tecnick dataset [28] in RGB444 domain. To be fair, all schemes are compared in the RGB444 domain. Also, we compare the performance of different codecs on both PSNR and MS-SSIM metrics.

Each pixel value of the compact representation which is the output of the RecNet is in the range of $[-1, 1]$. In order to better adopt FLIF codec for encoding and decoding, we need to use a certain mapping function to transform the output value of the compact representation to an integer value of $[0, 255]$. As [4], we present three mapping methods to achieve this process. The first method is called shift method as shown in Eq.6.

$$x_Q = \text{round}[\frac{x_f + 1}{2} \times 255]. \tag{6}$$

x_f represents the output of the CompNet. A better scaling method is to find the maximum and minimum values of the residual images, x'_{max} and x'_{min} , and then adaptively map the range between them to $[0, 255]$. That is:

$$x_Q = \text{round}[\frac{(x_f - x'_{min})}{(x'_{max} - x'_{min})} \times 255]. \tag{7}$$

This map method is called the MinMax method.

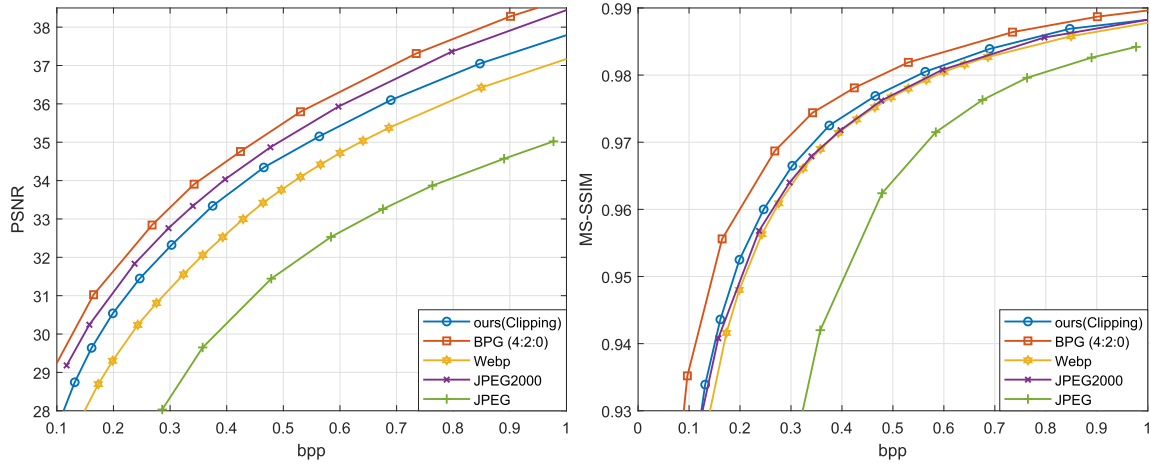


FIGURE 8. Average comparison results on all 100 Tecnick images in YUV domain in terms of PSNR and MS-SSIM.

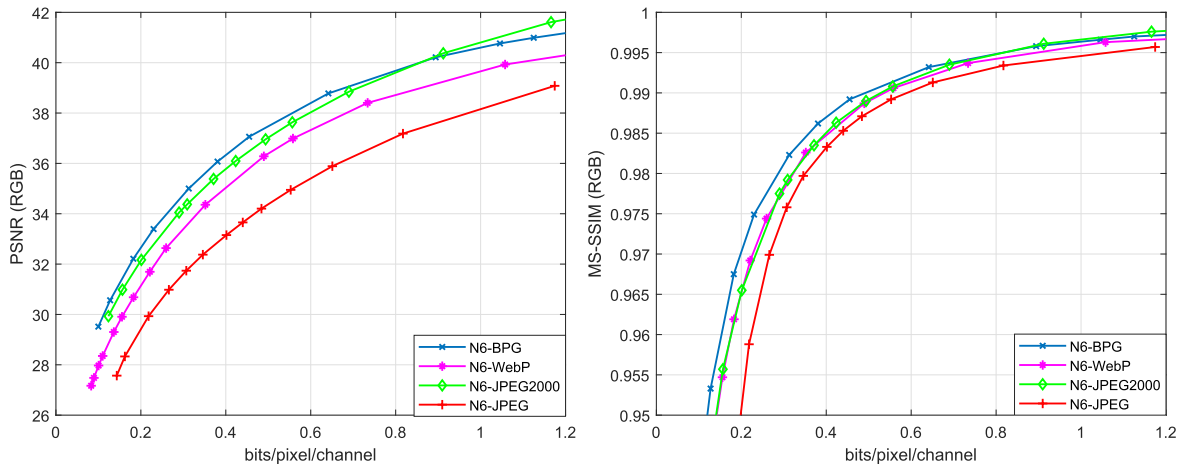


FIGURE 9. Average comparison results on all 24 Kodak images of different codecs in RGB444 domain in terms of PSNR and MS-SSIM.

The third method called clipping method is further extended in the Eq.7, and the optimal decomposition point is determined according to the distribution of the pixels of the residual image, that is, the maximum and minimum values of Eq.7. As discussed before, the maximum and minimum values are set to 118 and -118 respectively. In our scheme, we adopt these three mapping methods to encode the residual images to achieve three different performance curves in fig.4. The quantization level L is set to 6. that is, there are six point numbers in vector L .

The results are shown as Fig.4. Our three scaling methods are better than traditional codecs including BPG, JPEG2000, Webp and JPEG in RGB444 domain on both PSNR and MS-SSIM metrics. And the Clipping method exceeds some hybrid lossy image compression schemes including DSSLIC(2019) [2] and Fu(2019) [4] across a wide range bit rate in RGB444 domain in MS-SSIM metric. Our MinMax almost achieves the same performance with [4] at low bit rates.

In order to better compare the compression quality of the image, the two typical images are chosen from the Kodak

dataset [27] for comparison in Fig.10 and Fig.11. We zoom in the same part of the image to better compare the different results of the different codes. By analyzing the two samples, the following conclusions are as follows. Compared to other codecs, the JPEG has the worst performance due to the existence of the blocking artifacts. Also, the JPEG2000 also has some artifacts in the processing of the image compression. The BPG codec can make images smoother, but it would lose some detailed information in the processing of the image compression at low bit rates. Our proposed method is better than the BPG codec and [4] in MS-SSIM [6] metric. It can display more detailed information at low bit rates.

The main reasons why our proposed scheme is better than [2], [4] schemes are as followed. First, we adopt the soft-to-hard quantization method to take the place of the rounded quantization method. References [2], [4] adopt rounded quantization methods to deal with the compact image C , and they ignore the quantization error which slightly reduces the performance of the coarse reconstructed images. Secondly, they adopt at least 8 bits to represent each raw pixel of the compact image C , which increases the number of bits occupied by the



FIGURE 10. An example from the Kodak dataset(bits/pixel/channel, PSNR, MS-SSIM).

base layer. We can control the number of bits in the base layer by changing the number of quantization levels L . When the bit rate is fixed, we can reduce the number of bits occupied by the base layer and allocate more bits to the enhancement layer. This is why our proposed can exceed the [2], [4] in SSIM metric across a wide range of bit rates.

A. BIT RATE ALLOCATION

The key point of the hybrid scheme is the problem of rate allocation of the base layer and enhancement layer. When quantifying the number of levels, the number of bits occupied by the base layer is almost a fixed value. The FLIF codec

will slightly reduce the number of bits in the base layer. The scheme mainly adopts the BPG codec to encode the residual to adjust the bit rate of the entire scheme, so as to achieve compression at different bit rates. We also can adjust the bit rate of the base layer by changing the number of quantization levels L , so as to explore the impact of the bit ratios of the base layer and the enhancement layer on the experimental scheme. The result is shown as Fig.6. In order to better illustrate the impact of the numbers of quantization level L on compact images, we also show the visualization of the compact image C in different numbers of quantized levels L .

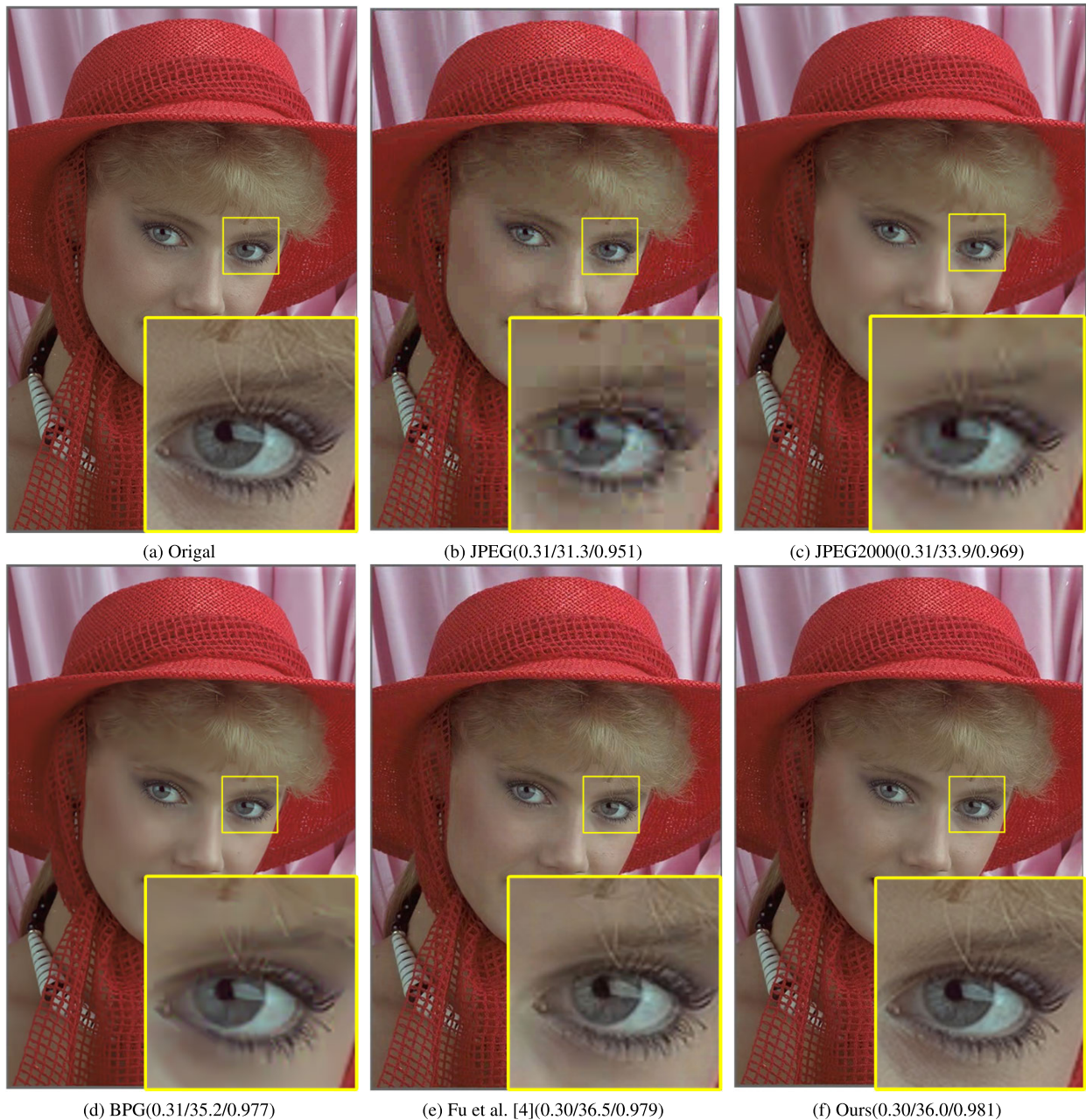


FIGURE 11. An example from the Kodak dataset(bits/pixel/channel, PSNR, MS-SSIM).

The number of quantization levels L directly affects the performance of our proposed scheme. As shown in Fig.12, the compact images can display richer information with the increase of quantization levels L . As shown in Fig.6, compared with the MS-SSIM metric, the PSNR metric is more susceptible to the number of quantization levels L . PSNR performance will gradually increase as the number of quantization levels L increases across a wide range bit rate. The MS-SSIM metric will slightly decrease as the number of quantization levels L . In order to further explore the bit rate allocation of the base layer and the enhancement layer,

we control the bit rate of the entire scheme by controlling the quality factor of the enhancement layer encoder. The result is shown in Table 1. The average uncompressed size of the compact images is 5.08KB. After FLIF encoding, the average size of the compact image C will be reduced from 5.08KB to 3.85KB. For a fixed network, the number of bits occupied by the compact image using FLIF codec is fixed, and the total number of bits can only be adjusted by the coefficients of residual image coding. Since the channel for the compact and residual image is fixed, the fewer the numbers of the quantization levels L , the smaller the number of bits occupied

TABLE 1. Average compressed sizes of the compact images and residual images.

		bits rate	Total bits	Compressed Residual Image	Compressed Compact Image	ratio
Kodak	N7	0.1277	147.11KB	143.26KB	3.85KB	37.21:1
	N7	0.1823	210.01KB	206.16KB	3.85KB	53.55:1
	N7	0.230	264.96KB	261.11KB	3.85KB	67.82:1
	N7	0.3128	360.35KB	356.50KB	3.85KB	92.60:1
	N7	0.4550	524.16KB	520.31KB	3.85KB	135.15:1
	N7	0.6415	89.51KB	739.01KB	3.85KB	190.95:1
Tecnick	N7	0.0952	401.62KB	12.60KB	12.60KB	30.88:1
	N7	0.1289	543.80KB	531.20KB	12.60KB	42.16:1
	N7	0.2115	892.27KB	879.67KB	12.60KB	69.81:1
	N7	0.3131	1320.88KB	1308.29KB	12.60KB	103.83:1
	N7	0.4641	1957.92KB	1929.14KB	12.60KB	154.39:1
	N7	0.6893	2907.98KB	2895.38KB	12.60KB	229.79:1

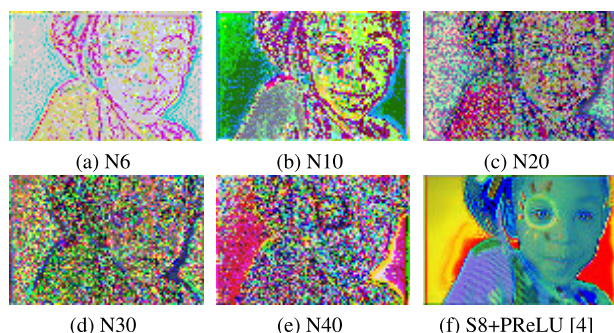


FIGURE 12. Visualization of the compact image C in different numbers of quantized levels L.

by the base layer, and more bits can be allocated to the enhancement layer.

1) YUV DOMAIN

We further evaluate the performance of our proposed method in the YUV domain compared to some traditional codecs including BPG, JPEG2000, WebP, and JPEG. The results are shown in Fig.7 and Fig.8. The proposed scheme exceeds JPEG2000, Webp and JPEG in MS-SSIM metric across a wide bit rate on both Kodak dataset and Tecnick dataset.

The main reasons why our hybrid coding scheme is very limited in the YUV domain are as followed. As discussed before, the performance of the hybrid coding scheme is jointly determined by the base layer and the enhancement layer. As shown in Table 1, although the enhancement layer occupies most of the bits, the base layer is the key factor that determines the effectiveness of the hybrid coding scheme. The base layer determines the initial performance of the hybrid encoding, that is, the performance of the rough reconstructed image. The effect of adding the encoded residual to the coarse reconstructed image to enhance the final reconstructed image is very limited. Since compact images C need to be encoded with FLIF codec, this results in an image with only three channels. Although the performance of mixed image coding can be improved by quantizing the number of levels, the effect is very limited. The Fig.6 can prove this point.

2) ENHANCEMENT LAYER CODEC IMPACT

The purpose of the enhancement layer is to further encode the residual image, and then add the encoded residual image to the rough reconstructed image, thereby further improving the performance of the final reconstructed image. Therefore, the performance of the enhancement layer codec is very important. We also explore the impact of different codecs on the experimental scheme. We adopt JPEG, JPEG2000, Webp, and BPG as enhancement layer encoders to encode the residuals, and then explored the impact of encoder performance on overall solution performance. The result is shown in Fig.9.

Experimental results show that the effect of the scheme depends on the performance of the enhancement layer encoder, that is, the stronger the performance of the enhancement layer encoder, the better the overall performance of the scheme. The BPG codec achieves the best performance across a wide bit rate on both PSNR and MS-SSIM metrics. But the JPEG2000 will exceed the BPG codec at high bit rates on both PSNR and MS-SSIM metrics. Also, compare with other codecs, the JPEG codec has the worst performance. For a fixed code rate, since the number of bits occupied by the base layer is fixed, the number of bits allocated to the enhancement layer is also fixed. Different codecs need to adjust the quantization factor to meet this number of bits. Which type of codec loses less information during the adjustment process, the better the performance it gets.

IV. CONCLUSION

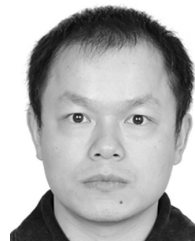
In this paper, we further propose an extended improved hybrid image compression scheme, which adopts the soft-to-hard quantification method to take the place of the rounding quantization method. We also explore the problem of bit rate allocation and the impact of different encoders on the enhancement layer. The performance of our proposed method is better than [2], [4] in MS-SSIM metric across a large range of bit rates. In future work, we will further improve the performance of the hybrid image scheme in the YUV domain by improving the performance of the base layer and introducing the new arithmetic coding algorithm to processing compact images C.

REFERENCES

- [1] E. Agustsson, F. Mentzer, M. Tschannen, L. Cavignoli, R. Timofte, L. Benini, and L. Van Gool, "Soft-to-hard vector quantization for end-to-end learning compressible representations," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1141–1151.
- [2] M. Akbari, J. Liang, and J. Han, "DSSLIC: Deep semantic segmentation-based layered image compression," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 2042–2046.
- [3] A. V. D. Oord, N. Kalchbrenner, O. Vinyals, L. Espeholt, A. Graves, and K. Kavukcuoglu, "Conditional image generation with pixelcnn decoders," in *Proc. Neural Inf. Process. Syst.*, 2016, 4797–4805.
- [4] H. Fu, F. Liang, B. Lei, N. Bian, Q. Zhang, M. Akbari, J. Liang, and C. Tu, "Improved hybrid layered image compression using deep learning and traditional codecs," 2019, *arXiv:1907.06566*. [Online]. Available: <http://arxiv.org/abs/1907.06566>
- [5] G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, and M. Covell, "Full resolution image compression with recurrent neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5435–5443.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [7] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, 2003, pp. 1398–1402.
- [8] J. Sneyers and P. Wuille, "FLIF: Free lossless image format based on MANIAC compression," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 66–70.
- [9] G. Toderici, S. M. O'Malley, S. Jin Hwang, D. Vincent, D. Minnen, S. Baluja, M. Covell, and R. Sukthankar, "Variable rate image compression with recurrent neural networks," 2015, *arXiv:1511.06085*. [Online]. Available: <http://arxiv.org/abs/1511.06085>
- [10] M. Li, W. Zuo, S. Gu, J. You, and D. Zhang, "Learning content-weighted deep image compression," 2019, *arXiv:1904.00664*. [Online]. Available: <http://arxiv.org/abs/1904.00664>
- [11] M. Li, W. Zuo, S. Gu, D. Zhao, and D. Zhang, "Learning convolutional networks for content-weighted image compression," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3214–3223.
- [12] F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, and L. V. Gool, "Conditional probability models for deep image compression," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4394–4402.
- [13] J. Balle, V. Laparra, and E. P. Simoncelli, "End-to-end optimization of nonlinear transform codes for perceptual quality," in *Proc. Picture Coding Symp. (PCS)*, 2016, pp. 1–5.
- [14] C. Cai, L. Chen, X. Zhang, and Z. Gao, "End-to-end optimized ROI image compression," *IEEE Trans. Image Process.*, vol. 29, pp. 3442–3457, 2020, doi: [10.1109/TIP.2019.2960869](https://doi.org/10.1109/TIP.2019.2960869).
- [15] J. Ballé, D. Minnen, S. Singh, S. J. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior," in *Proc. Int. Conf. Learn. Represent.* 2018, pp. 1–23.
- [16] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional GANs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8798–8807.
- [17] L. Theis, W. Shi, A. Cunningham, and F. Huszár, "Lossy image compression with compressive autoencoders," in *Proc. Int. Conf. Learn. Represent.* 2017, pp. 1–19.
- [18] F. Bellard. (2019). *BPG Image Format*. [Online]. Available: <https://bellard.org/bpg/>
- [19] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. C-23, no. 1, pp. 90–93, Jan. 1974.
- [20] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 719–746, Mar. 1992.
- [21] G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 30–44, Apr. 1991.
- [22] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 36–58, Dec. 2001.
- [23] D. Minnen, J. Ballé, and G. D. Toderici, "Joint autoregressive and hierarchical priors for learned image compression," in *Proc. 31th Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 10794–10803.
- [24] J. Lee, S. Cho, and M. Kim, "An End-to-End joint learning scheme of image compression and quality enhancement with improved entropy minimization," 2019, *arXiv:1912.12817*. [Online]. Available: <http://arxiv.org/abs/1912.12817>
- [25] J. Zhou, S. Wen, A. Nakagawa, K. Kazui, and Z. Tan, "Multi-scale and context-adaptive entropy model for image compression," 2019, *arXiv:1910.07844*. [Online]. Available: <http://arxiv.org/abs/1910.07844>
- [26] *Challenge on Learned Image Compression 2018*. Accessed: 2018. [Online]. Available: <http://www.compression.cc/>
- [27] *Kodak Lossless True Color Image Suite*. Accessed: Jun. 2018. [Online]. Available: <http://r0k.us/graphics/kodak/>
- [28] N. Asuni and A. Giachetti, "TESTIMAGES: A large data archive for display and algorithm testing," *J. Graph. Tools*, vol. 17, no. 4, pp. 113–125, 2013.



HAISHENG FU was born in Henan, China, in 1994. He received the B.S. degree in automation engineering from Henan Polytechnic University, China. He is currently pursuing the Ph.D. degree in electronic science and technology with Xi'an Jiaotong University, Xi'an. His research interests include machine learning, deep learning, and VLSI design.



FENG LIANG received the B.E. degree from Zhengzhou University, China, in 1998, and the M.E. and Ph.D. degrees from Xi'an Jiaotong University, China, in 2001 and 2008, respectively. He has been an Associate Professor with the Microelectronics School, Xi'an Jiaotong University, since 2011. His current research interests include signal processing, machine learning, VLSI design, test, and computer architecture.



BO LEI was born in Lanzhou, China. He received the bachelor's degree in electronic materials and devices from Lanzhou University, China, in 2018. He is currently pursuing the master's degree in electronic science and technology with Xi'an Jiaotong University, Xi'an. His research interests include machine learning and deep learning.

...