A GAN-based Tunable Image Compression System

Lirong Wu, Kejie Huang, and Haibin Shen

College of Information Science and Electronic Engineering , Zhejiang University, Hangzhou, China

Introduction

Efficient image compression is significant for the storage, transmission, and processing of image information. In recent years, remarkable achievements with DNN-based image compression have been made. However, their compression performance often dramatically drops at low *bpp*. In this paper, we rethinks content-based compression by using Generative Adversarial Network to reconstruct the non-important regions. A tunable compression scheme is also proposed in this paper to compress an image to any specific compression ratio without retraining the model.

Methods

We design a simple network (Masking) to identify the important regions of the image and guide the allocation of bits. (See Fig1)

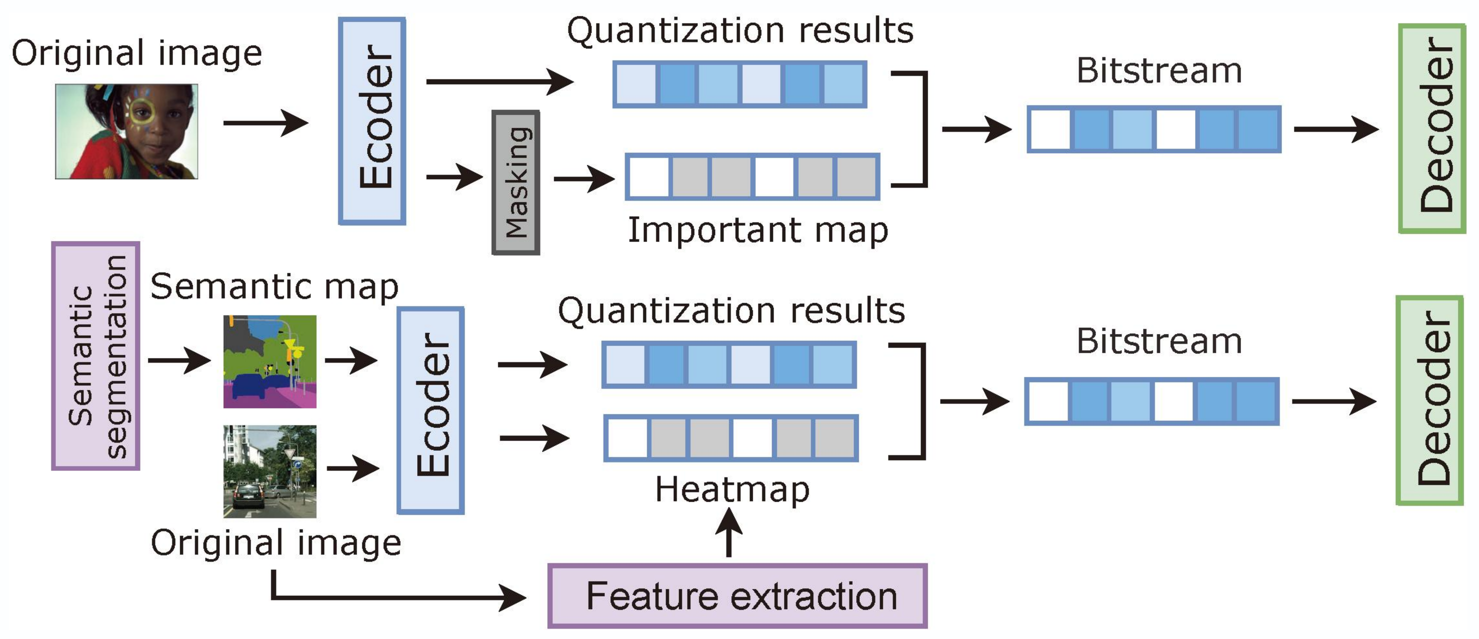

Fig 1

We use the multiscale structure not only in the encoder but also discriminator. The symmetrical multiscale structure makes it more adaptable to different sizes of objects. (See Fig 2)

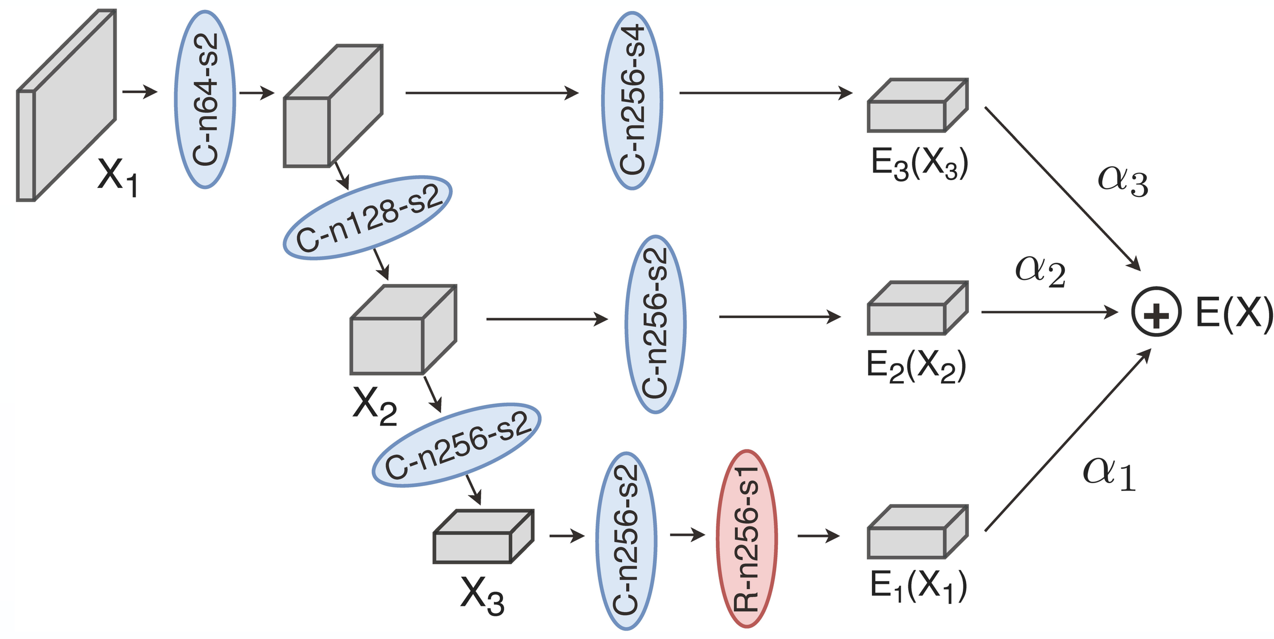

Fig 2

The bit allocation according to the importance of image contents is achieved by constructing a masker. The formula of the normalization procedure is specified as

$$\hat{y}_{i,j} = \frac{y_{i,j} - (\overline{\mu} + n)}{\overline{\sigma}} \qquad (1)$$

where *n* is a user-defined parameter, through which we can achieve different compression ratios without retraining the model. (See Fig 3)

The loss function of ourmodel is composed of adversarial loss and distortion loss.

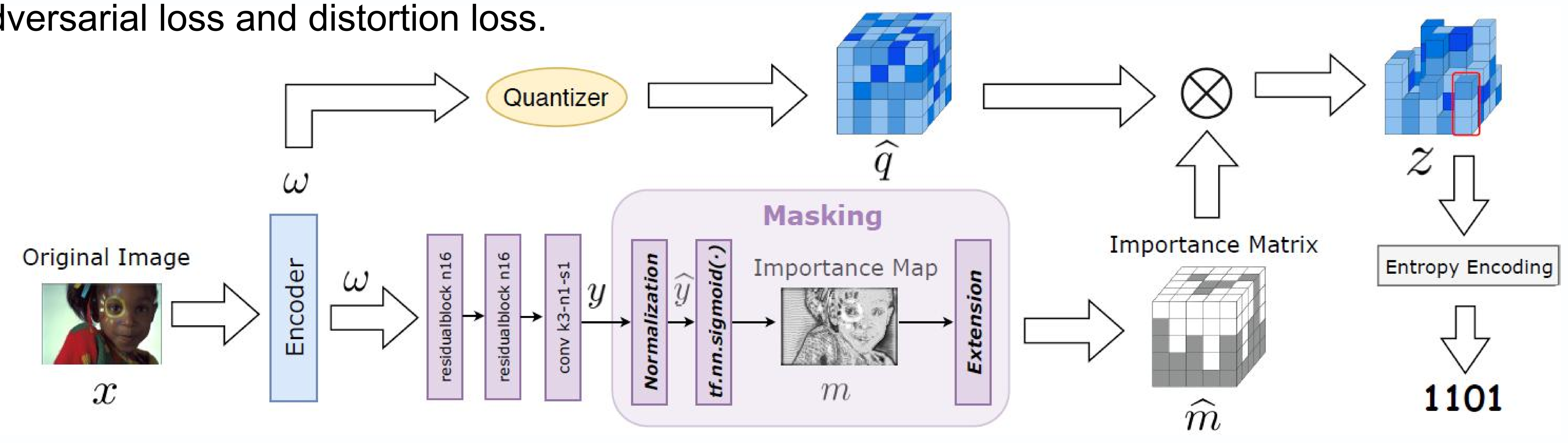

Fig3

- Adversarial Loss

$$\mathcal{L}_A = \sum_{i=1}^{m} \beta_i \Big\{ \mathbb{E}[log D_i(x)] + \mathbb{E}\big[log(1 - D_i(G(x)))\big] \Big\} \qquad (2)$$

- Distortion Loss

$$\mathcal{L}_D = E[d(x, \hat{x})] \qquad (3)$$

- Overall Loss

$$\mathcal{L}_{G,D,E,B} = \frac{1}{B}\sum_{j=1}^{B} \Big\langle \eta \sum_{i=1}^{m} \beta_i \Big\{ \mathbb{E}\big[log(1 - D_i(G(x^j)))\big] + \mathbb{E}[log D_i(x^j)] \Big\} + \kappa E\big[d(x^j, \hat{x}^j)\big] \Big\rangle \qquad (4)$$

Results

- Qantitative Results (See Fig 4)

Our proposed method improves MS-SSIM by more than 10.3% compared to the recently reported GAN-based method [1] to achieve the same low *bpp* (0.05) on the Kodak dataset.

- Qualitative Results (See Fig 5)

The details of the image, such as the window of the house and holes in the women's hat are well preserved due to the use of important map.

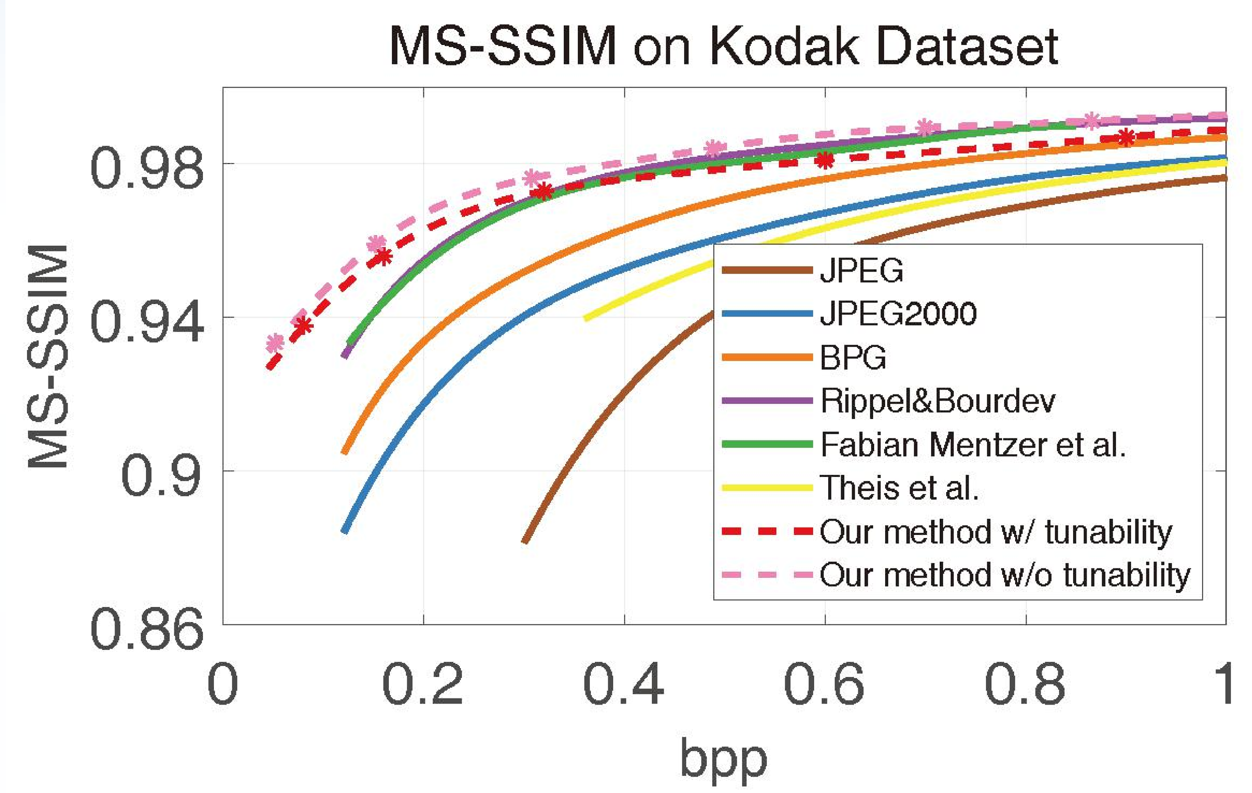

Fig 4

(a) *bpp* / MS-SSIM: Ours 0.039 / 0.927, Agustsson et al. 0.030 / 0.824

(b) *bpp* / MS-SSIM: Ours 0.058 / 0.921, Agustsson et al. 0.065 / 0.845

Fig 5

References

[1] Agustsson et al. Generative adversarial networks for extreme learned image compression. arXiv preprint arXiv:1804.02958, 2018.

[2] Mentzer et al. Conditional probability models for deep image compression.CVPR, pages 4394–4402, 2018.

[3] Rippel and Bourdev. Real-time adaptive image compression. ICML, pages 2922–2930. JMLR. org, 2017.