MULTI-CHANNEL MULTI-LOSS DEEP LEARNING BASED COMPRESSION MODEL FOR COLOR IMAGES

Ching-Chun Huang
Dep. of Computer Science
NCTU, Taiwan

Thanh-Phat Nguyen
Dep. of Electrical Engineering
CCU, Taiwan

Chen-Tung Lai
Dep. of Electrical Engineering
CCU, Taiwan

1. INTRODUCTION

Motivation

- The **fixed hand-crafted image transformations** in some traditional image codecs (JPEG, JPEG2000) are **not be** the **optimal transformation** for different variety of images.
- In a **ultra low bit-rate compression** application, the decompressed image may present visible artifacts such as blurring, ringing, blocking and color bias because of information missing.

Original Image | [7], BPP: 0.094 | WebP, BPP: 0.104 | JPEG, BPP: 0.135

Compression artifacts

Goal

- Construct a deep learning-based image compression framework that is **generalized to different variety of images**.
- Learn the compression framework to compress an image at **very low bit rate** while **preserving a pleasant visual quality** of decompressed images.

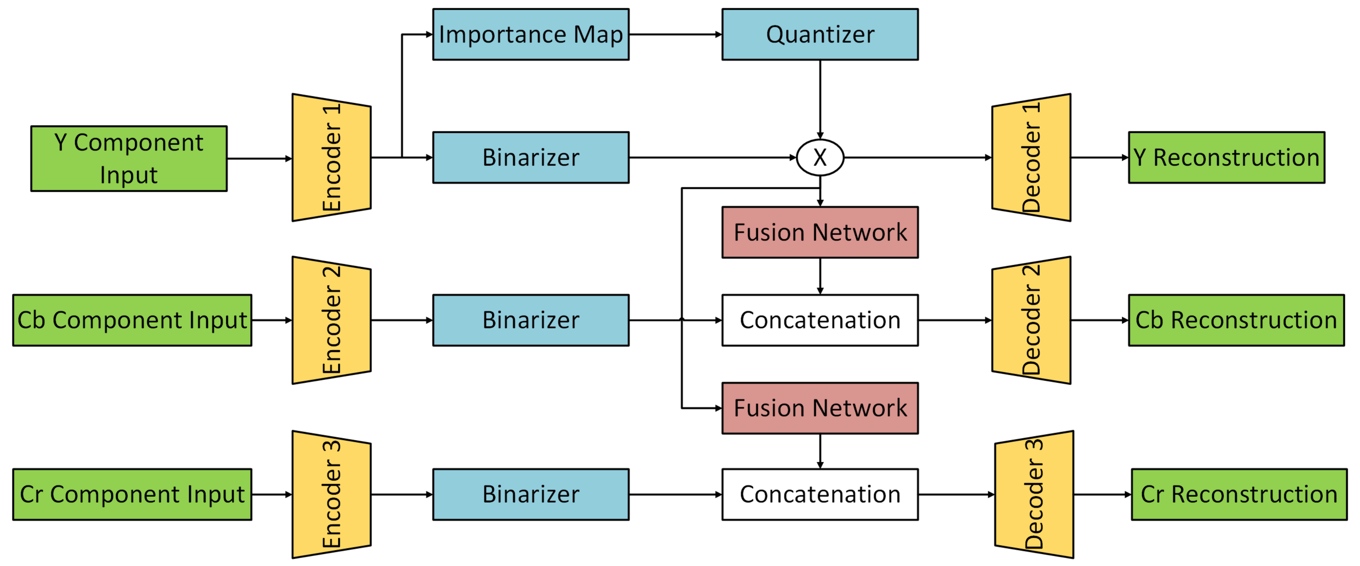
Challenges

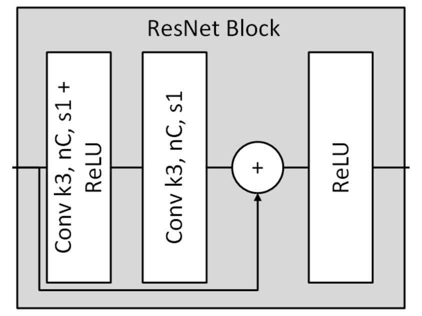
- Reduce color bias issue
- Leverage the dependency between intensity and color components in order to reduce the burden and bit usage in image compression task.

2. PROPOSED METHOD

Compression Framework

- The whole framework is composed of the deep convolutional autoencoder based on ResNet, the compression module, and the fusion network.

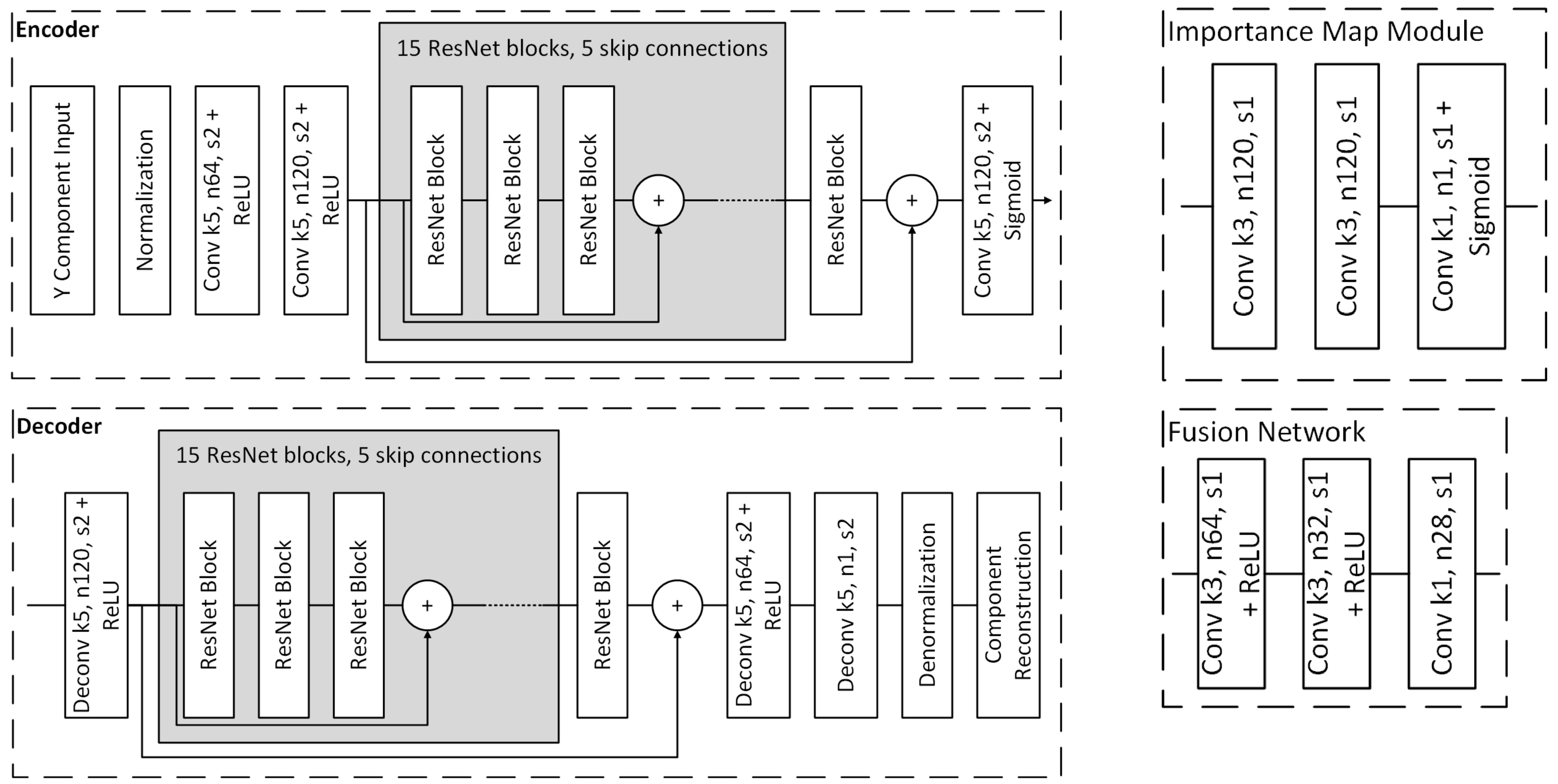




Conv: Convolutional layer
Deconv: Deconvolutional layer
k: kernel size (k x k)
n: number of feature maps
s: stride size(s x s)

Deep Convolutional Autoencoder

- To **preserve the color components better**, separate networks are used to compress intensity and color channels.
- Each encoder and decoder share the similar architecture, which has 16 ResNet blocks with 5 skip connection layers.



Fusion Network

- Borrow the image content information from the intensity stream to improve the color component reconstruction.
- The number of the color features can be then reduced before binarization, hence the total bitrate can be reduced.

Compression Module

- Importance map based quantizer and the binarizer are leveraged from [7].

Loss function

- Distortion loss function :

$$L_D(x,\tilde{x}) = \|x_Y - \tilde{x}_Y\|_1 + \gamma.\|x_{Cb} - \tilde{x}_{Cb}\|_1 + \gamma.\|x_{Cr} - \tilde{x}_{Cr}\|_1$$

- Bitrate loss function:

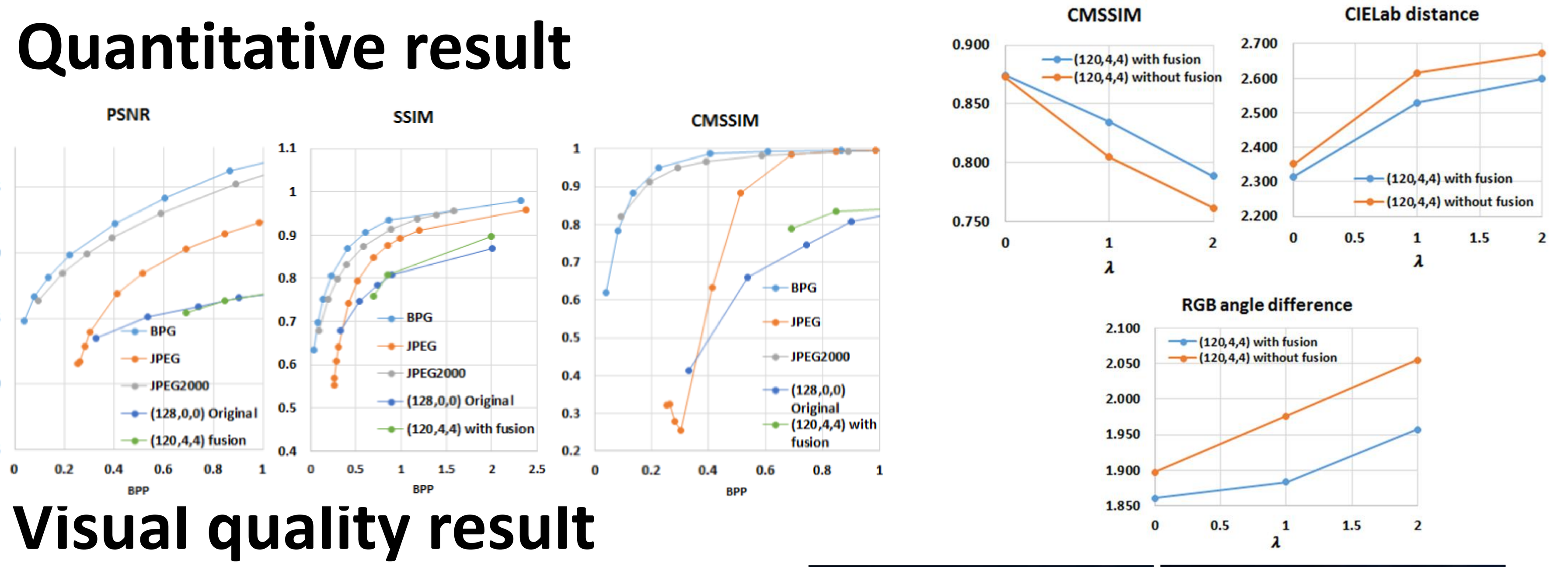
$$L_R(x) = \sum_{i,j} P(E(x_Y))_{ij}\,,$$ $P(E(x_Y))$ is importance map

- Overall loss function:

$$L = \sum_{x\in X} L_D(x,\tilde{x}) + \lambda L_R(x).$$

3. EXPERIMENTAL RESULTS

Quantitative result



Visual quality result

4. CONCLUSION

- By leveraging the intensity information, the network would focus on **disentangling the color-specific features** and allow using **fewer feature maps** to encode the entire image color information.

SELECTED REFERENCE

[7] M. Li, W. Zuo, S. Gu, D. Zhao, and D. Zhang, “Learning convolutional networks for content-weighted image compression,” in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2018