ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 13 | 18 September

西安電子科技大學
XIDIAN UNIVERSITY

Deep Fusion of RGB and NIR Paired Images Using Convolutional Neural Networks

Lin Mei and Cheolkon Jung

School of Electronic Engineering
Xidian University, China

Problems

- **Why Need fusion of RGB and NIR paired images?**
  RGB Images: noisy and poor
  Feature information: not obvious

- **Research Goal:** Achieve a high quality RGB images

  1) No color distortion
  2) Containing more details
  3) Without excessive contrast enhancement

Applications

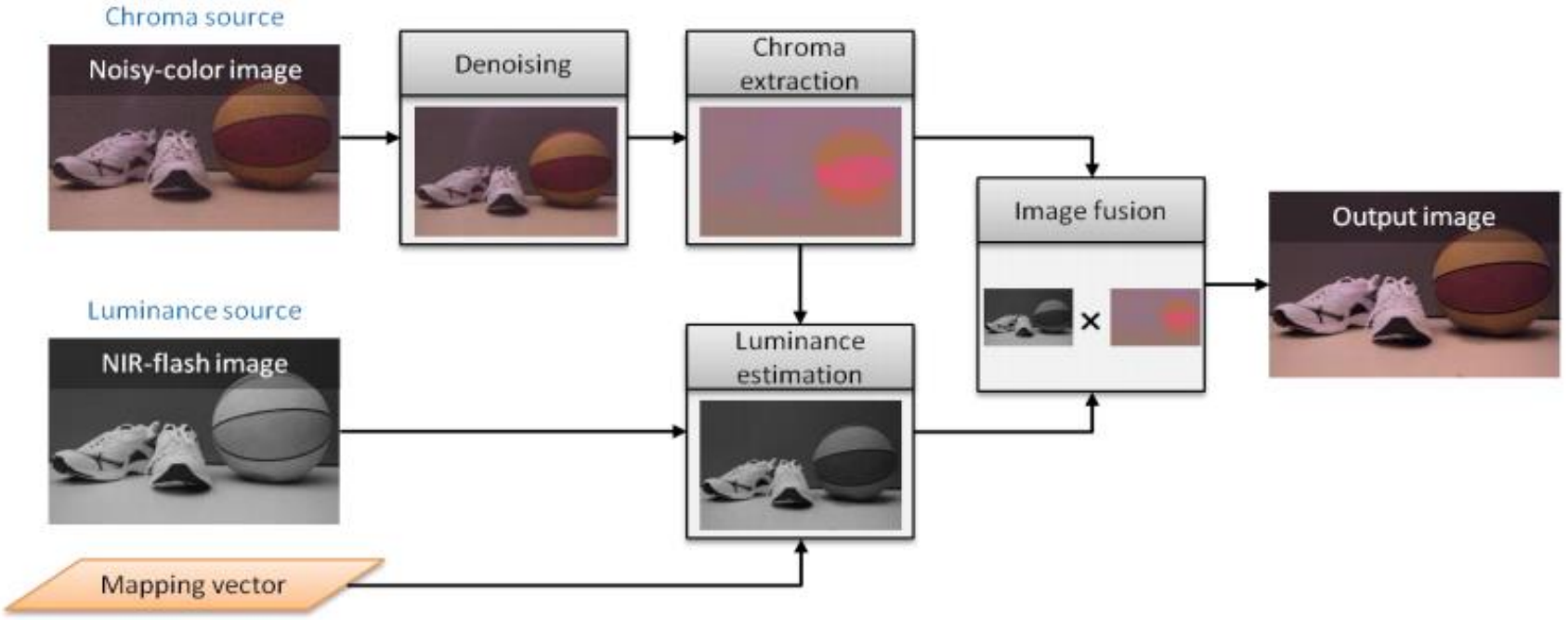
- Computer vision
- Night vision technology
- Medical image processing
- Processing of remote sense images

Night vision equipment

ICCV 2013

Fusion of RGB and NIR images by scale map:

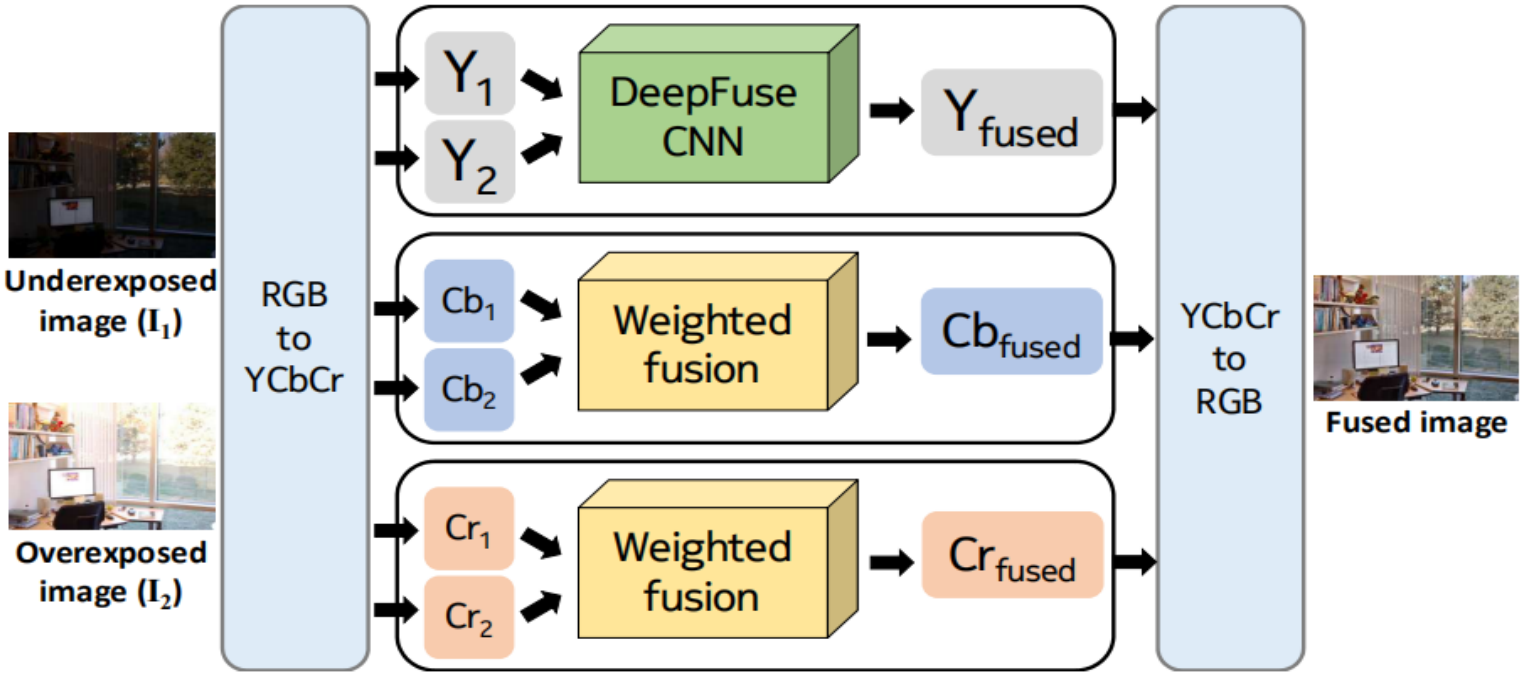
Low-Light Scene Color Imaging Based On Luminance Estimation From Near-Infrared Flash Image.



ICCV 2017

Deep learning-based fusion work (1):

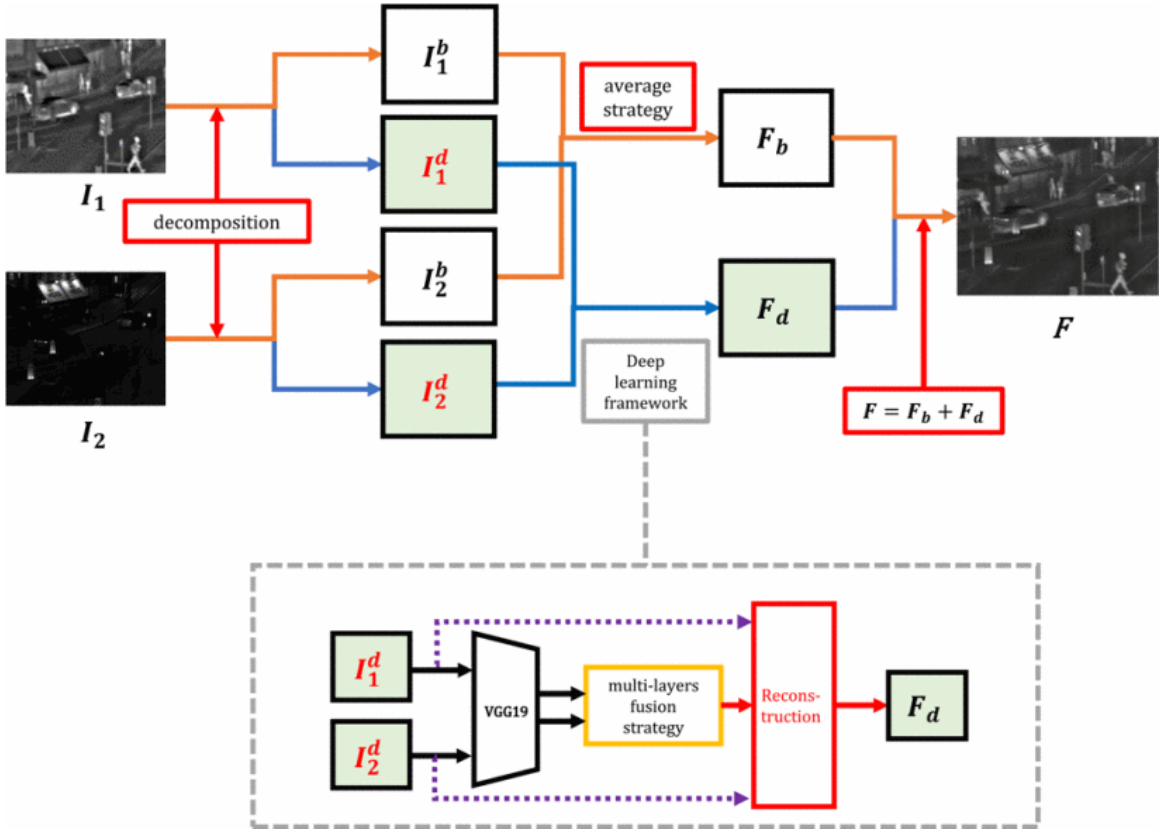
DeepFuse: A Deep Unsupervised Approach for Exposure Fusion with Extreme Exposure Image Pairs.



ICPR 2018

Deep learning-based fusion work (2):

Infrared and Visible Image Fusion using a Deep Learning Framework.

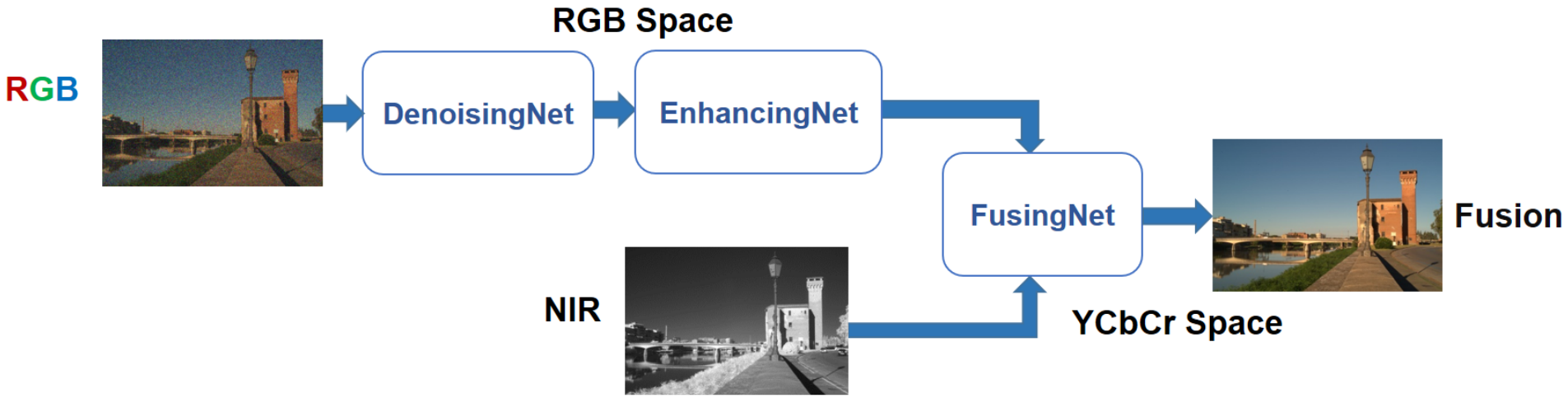


Motivation

- **Why use deep learning for fusion？**
  - Traditional methods: difficult to enhance extremely low-light images.
  - CNN: solve problems, achieves good performance.
  - NIR: sensitive to material properties.

- **Main idea for deep fusion:**
  - Consists of three modules.
  - DenoisingNet: remove noise.
  - EnhancingNet: improve the brightness.
  - FusingNet: fuse the Y channel of RGB image and NIR image .

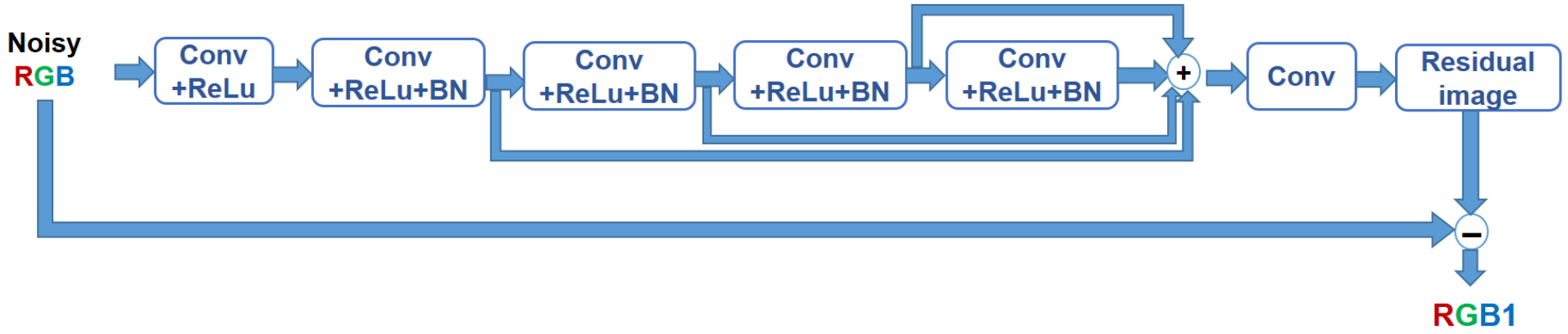
Proposed Method



- **DenoisingNet:** DnCNN (TIP 2017)
- **EnhancingNet:** MCRN (ECCVW 2018)
- **FusingNet:** DeepFuse (ICCV 2017)

Proposed Method

DenoisingNet by DnCNN[1]:



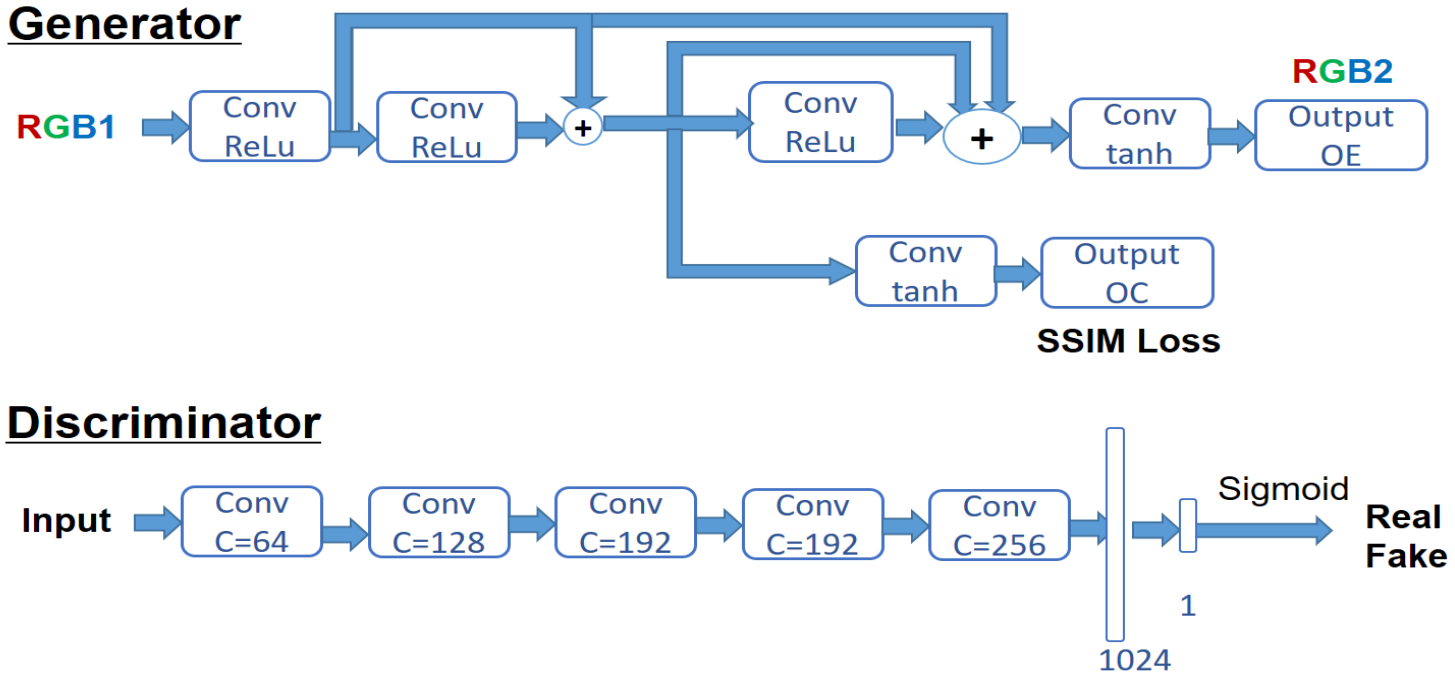
- Residual learning
- batch normalization

RGB=Noisy RGB-Residual image

[1] Zhang et al., "DnCNN: Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," TIP, 2017

Proposed Method

EnhancingNet by MCRN[2]:

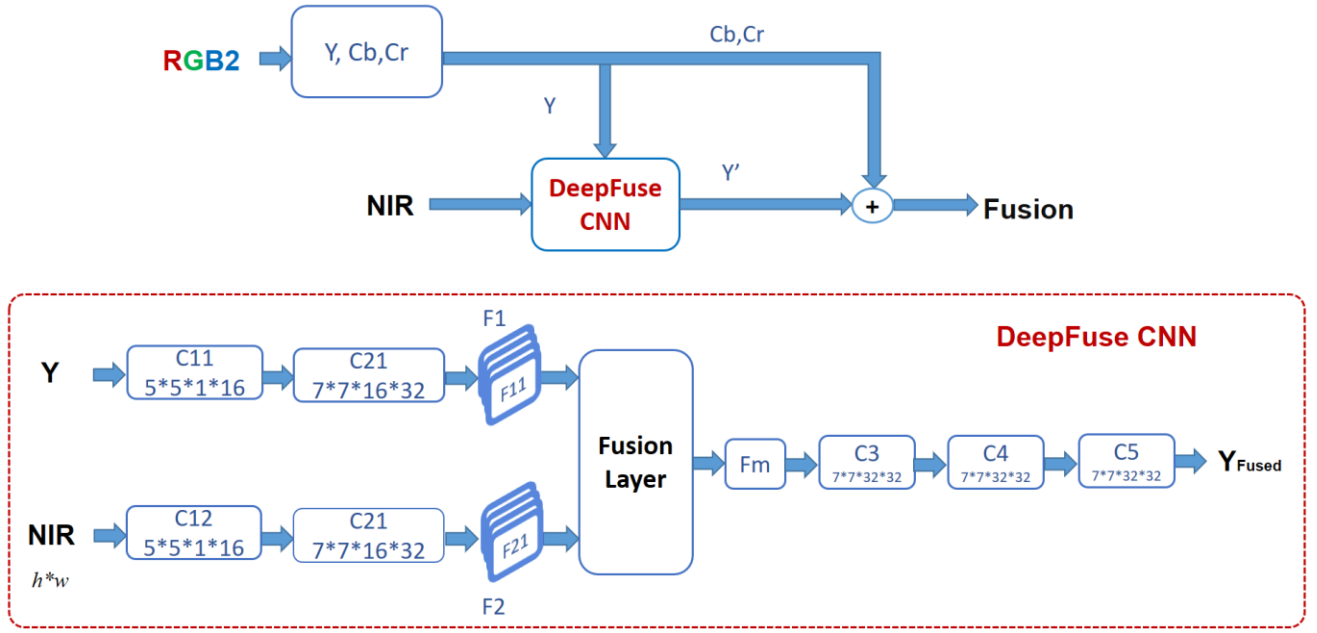


- **Generator:** consists of 5 layers in training and 4 layers in testing.
- **Discriminator:** consists of 5 convolutional layers and 2 fully connected layers.

[2] Liu and Jung, "Multiple Connected Residual Network for Image Enhancement on Smartphones," Proc. ECCV Workshop, 2018

Proposed Method

FusingNet by DeepFuse[3]:



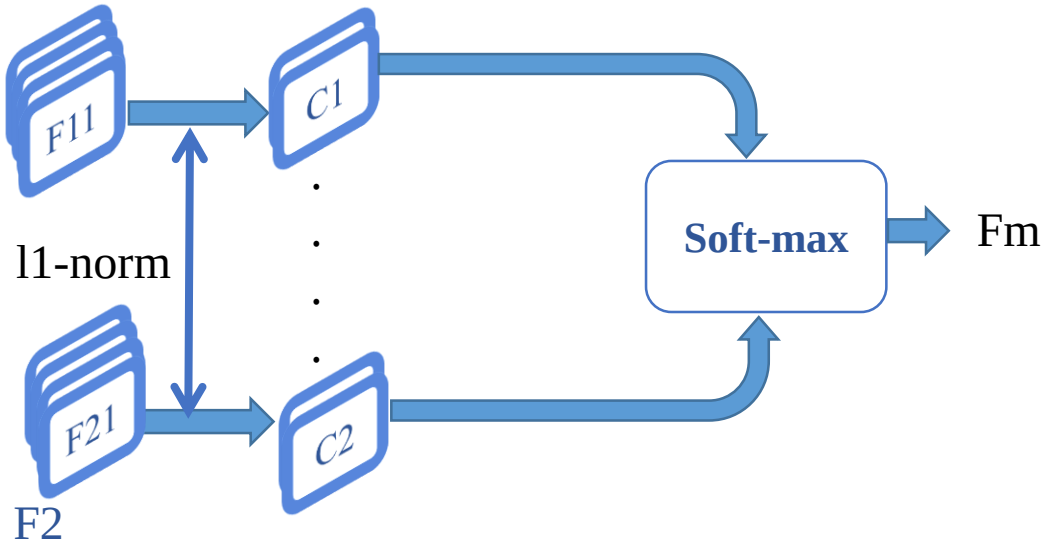
- Input: Y and NIR.
- The pre-fusion layers.
- The fusion layer.
- The reconstruction layers.

[3] Prabhakar et al., "DeepFuse: A Deep Unsupervised Approach for Exposure Fusion with Extreme Exposure Image Pairs," Proc. ICCV, 2017

Proposed Method

FusingNet by DeepFuse[3]:

- **Fusion layer:**



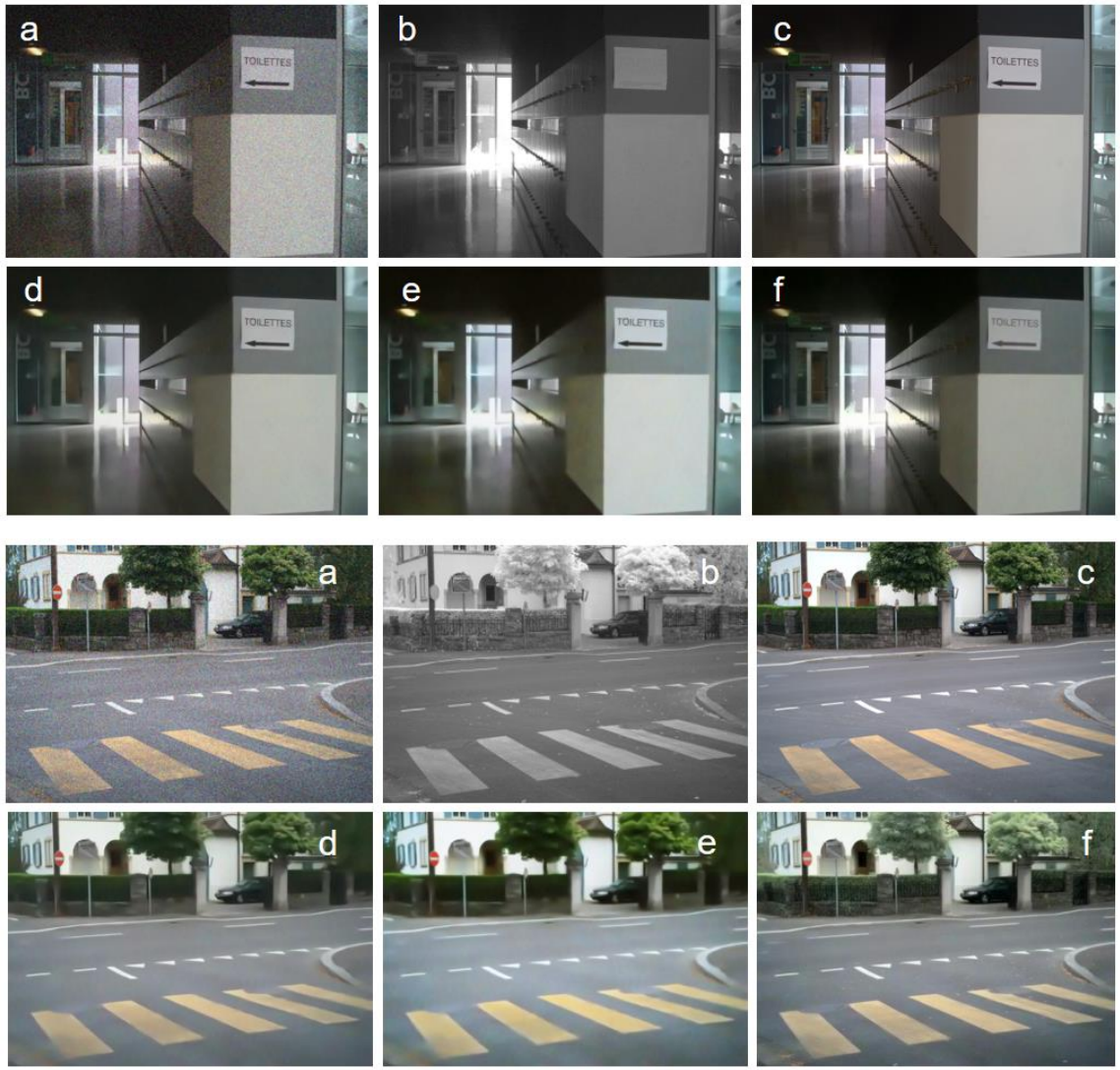
L1-norm.

Soft-max operation.

Experimental Results

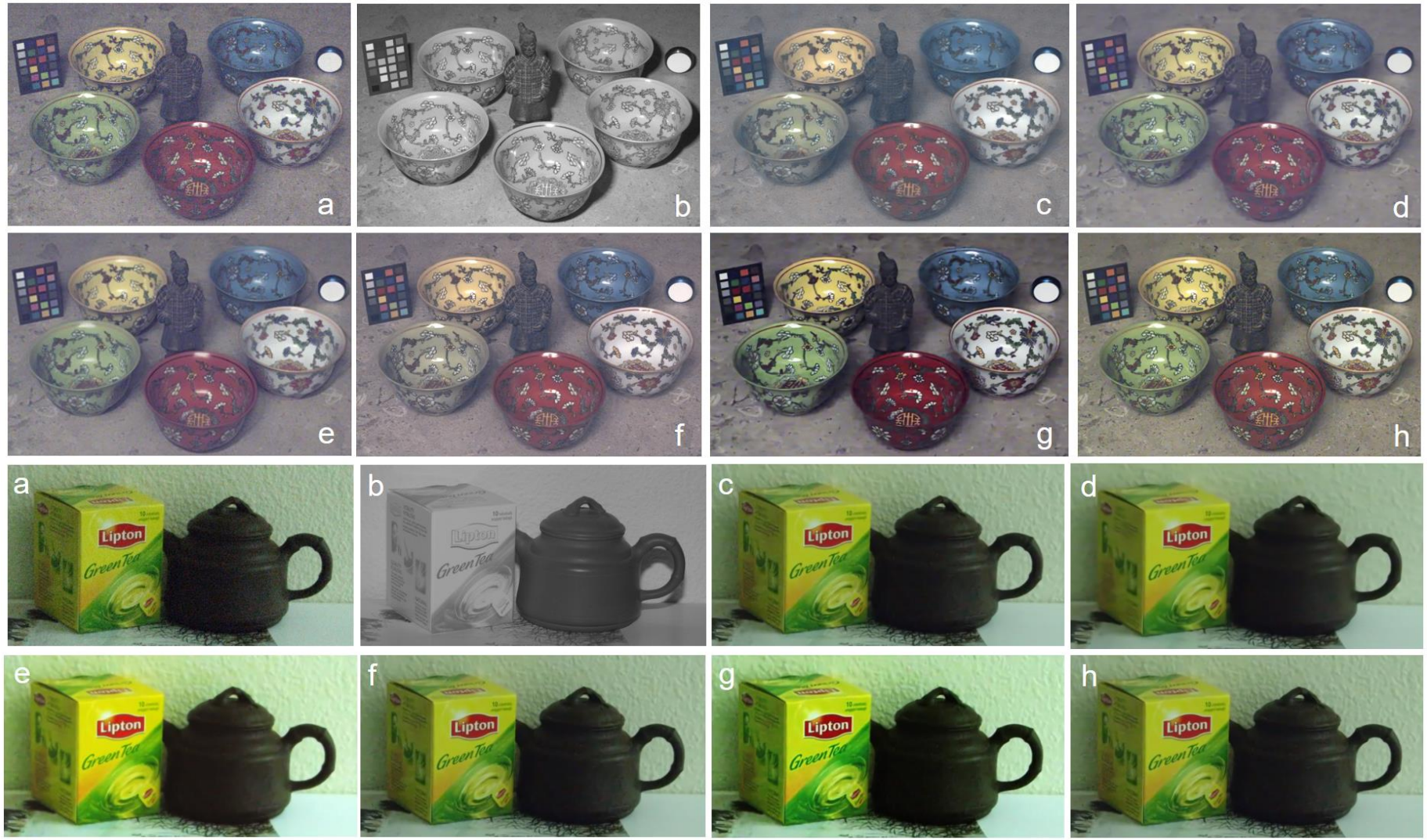
- **Experimental Environment**
  - Hardware: computer i7-7700.
  - Software: Linux, python 3.5
  - Dataset: NIR and VIS image pairs.
- **Evaluation metrics**
  - Visual Comparison
  - BIQE (blind image quality measures)
  - Runtime
- **Comparing methods**
  - [1] BM3D (TIP 2007)
  - [2] WLS (ICIP 2010)
  - [3] LD (TIP 2016)
  - Proposed-1 (DF): DenoisingNet+FusingNet
  - Proposed-2 (DFE): DenoisingNet+FusingNet+EnhancingNet
  - Proposed-3 (DEF): DenoisingNet+EnhancingNet+FusingNet

Experimental Results



Testing results of RGB-NIR scene.
(a) Input RGB.
(b) Input NIR.
(c) GT
(d) Denoising result.
(e) Enhancing result.
(f) Fusion result.

Experimental Results



Performance comparison between different methods. (a) Input RGB, (b) Input NIR, (c) BM3D, (d) WLS, (e) LD, (f) DF, (g) DFE, (h) DEF.

Experimental Results

Method	BM3D	WLS	LD	DF	DFE	DEF
bowls	29.996	26.645	27.248	27.325	26.768	26.591
teapot	29.452	27.963	28.812	26.132	28.481	25.903

BIQE of the fusion result obtained by different methods.

Method	BM3D	WLS	LD	DF	DFE	DEF
Runtime (sec/pair)	4.62	3.71	4.32	2.58	3.11	3.07

The average runtime. The unit is sec/pair while the image size is 1920 * 1080.

Method	BM3D	WLS	LD	DF	DFE	DEF
bowls	10.482	12.304	15.331	16.092	17.296	17.563
teapot	10.231	16.993	17.656	16.832	18.121	19.662

To further evaluate the performance, we also test SF.

Conclusions

- We propose deep fusion of RGB and NIR paired images.

- The proposed method consists of three subnetworks:

  DenoisingNet & EnhancingNet & FusingNet

- Proposed method removes noise while preserving details and recovering color.

- Future work: Network optimization to extract much details and reproduce good color.

THANK YOU!