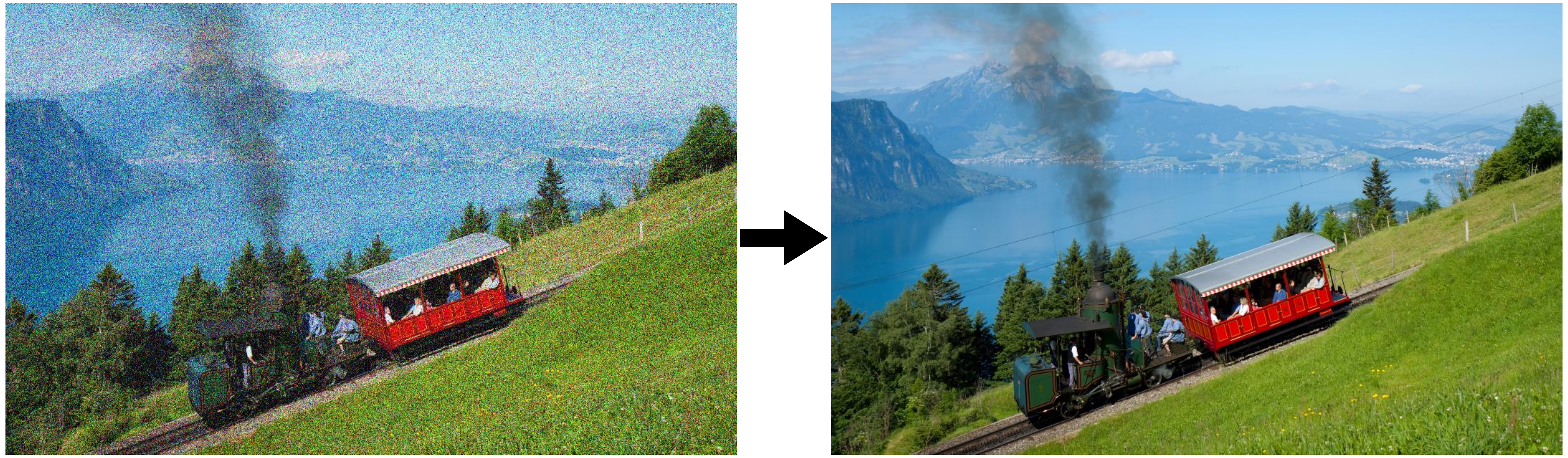
Exploiting Distortion Information for Multi-degraded Image Restoration

CVPR 2022 New Trends in Image Restoration and Enhancement workshop

Wooksu Shin, Namhyuk Ahn, Jeong-Hyeon Moon, Kyung-Ah Sohn

1. Introduction

Image restoration example



Degraded image Clean image

- Conventional Image restoration assumes that the input image is corrupted with **a single and fixed-intensity**
- However, various types of corruption with unknown strength can be applied in real-world applications
- We **integrate the complex perspectives** on the multi-distortion nature and propose a new dataset
- To effectively restore the multi-degraded image, we propose a **distortion information-guided network(DIGNet)**

2. Holistic Multi-Distortion Dataset(HMDD)

Previous Multiple Distortion Dataset

Mixed Distortions Dataset (Yu et al. 2018)[1]

- Applying multiple distortion **sequentially** to entire image.
- restore a single distortion.

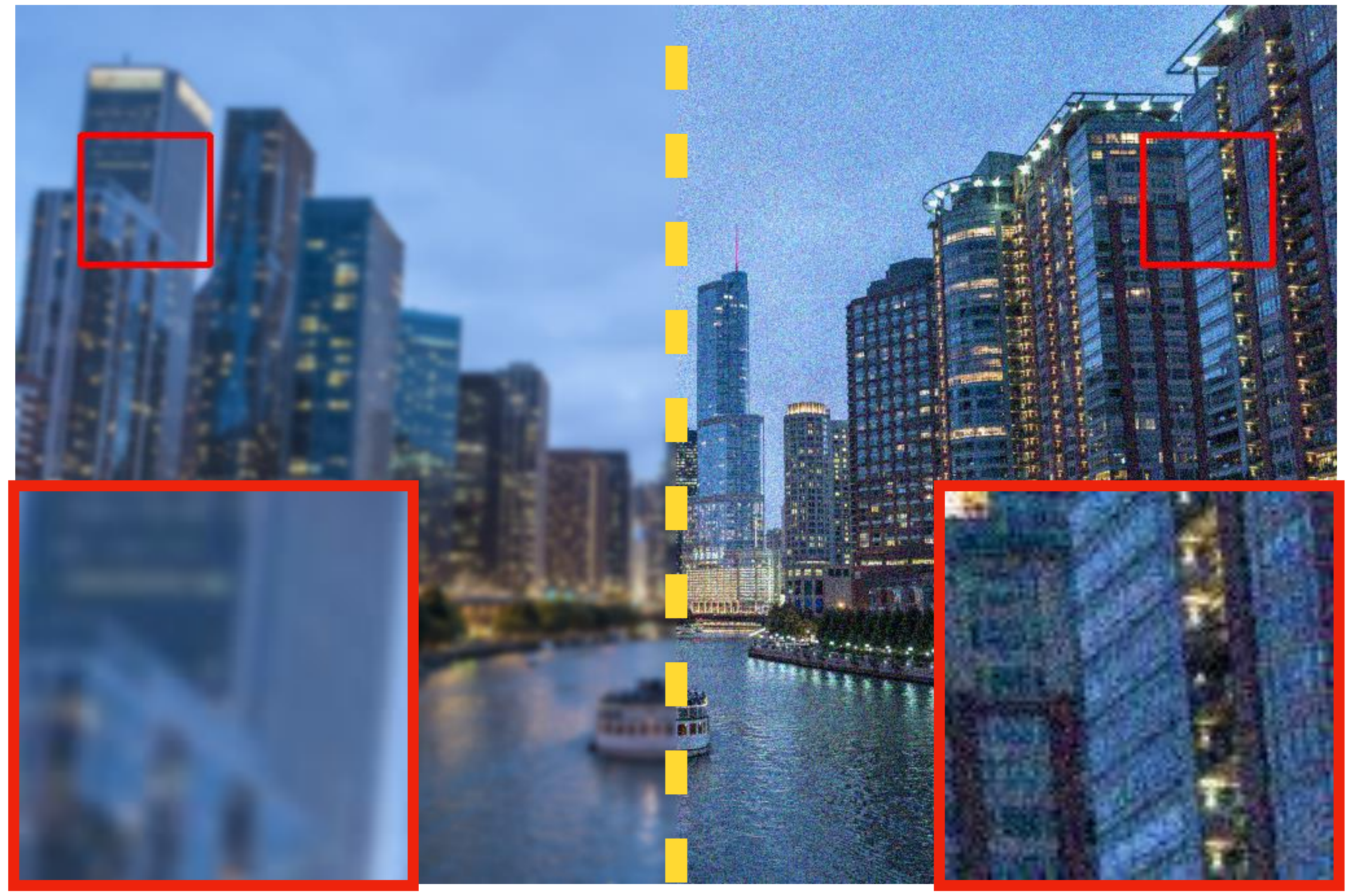
Gaussian noise, Gaussian blur, JPEG compression

Gaussian noise, Gaussian blur, JPEG compression

Mixed distortion image

Spatially-Heterogeneous Distortion Dataset (Sijin et al. 2020)[2]

- Applying **spatially-heterogeneous** distortion.
- restore mixed distortions.



Gaussian blur

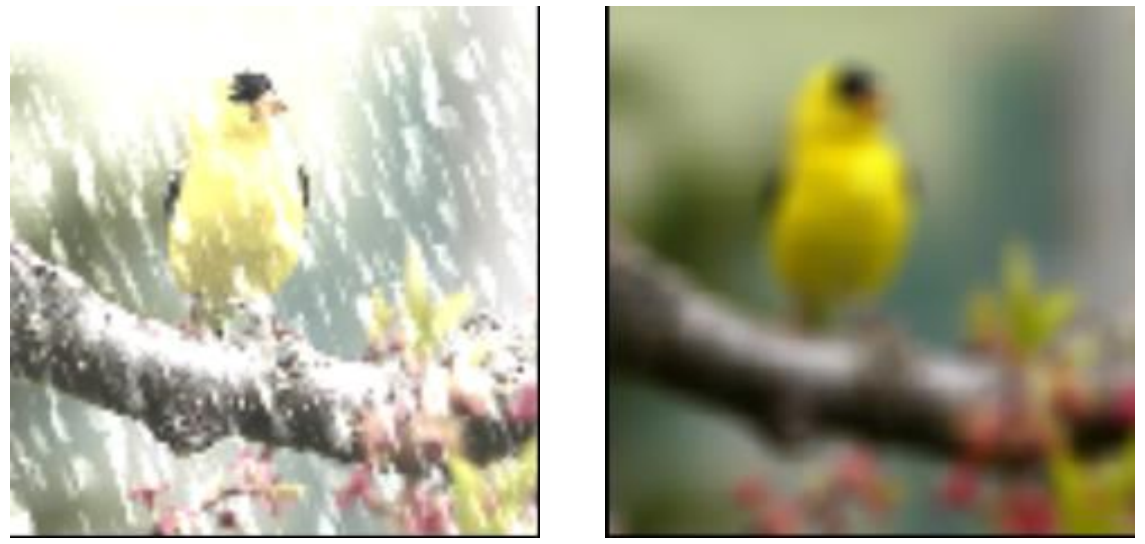
Gaussian noise

Spatially-heterogeneous distortion image

Holistic Multi-Distortion Dataset(HMDD)

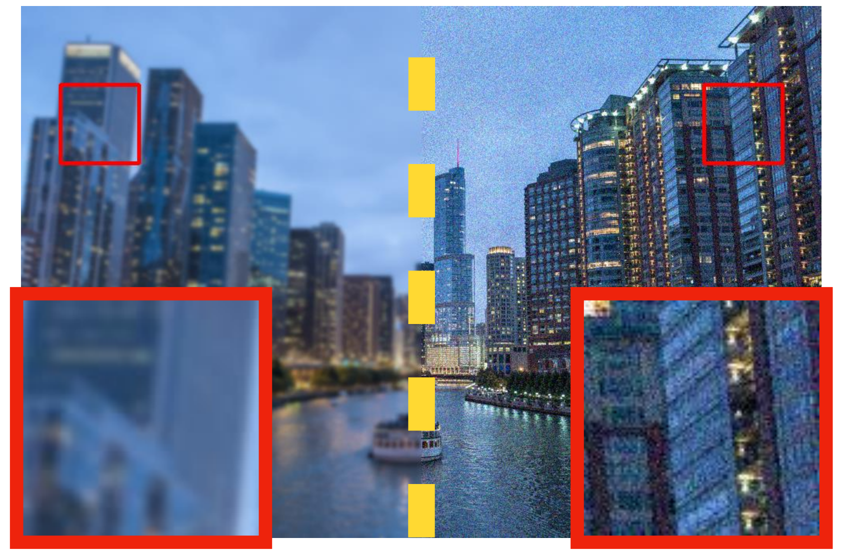
- We integrates both sequential and spatial distortions.
- Also, we make another dataset based on the weather & blur distortions as **HMDD-r**

Weather & Blur noise



Mixed distortion image

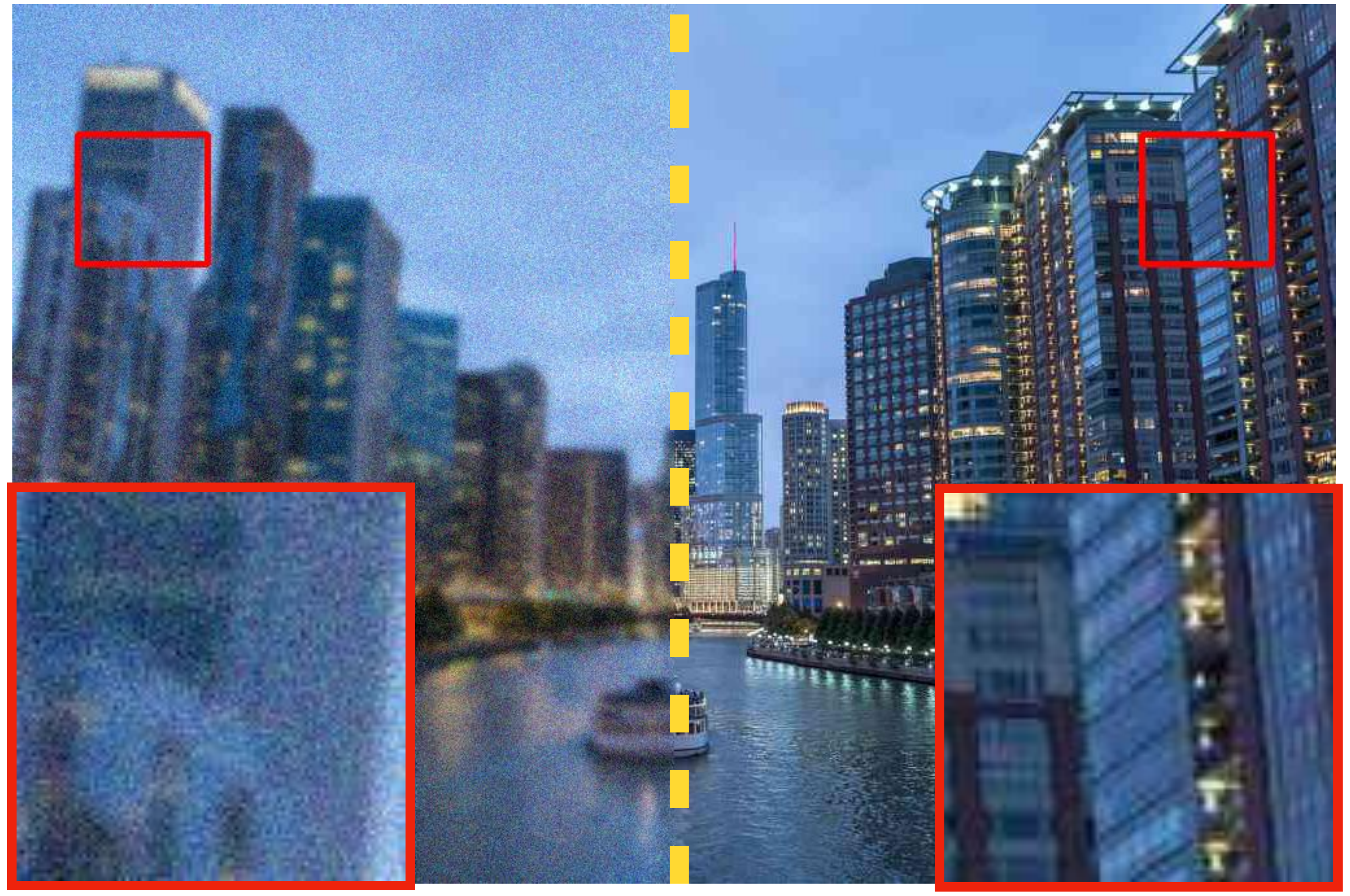
+



Spatially-heterogeneous distortion image

=

Gaussian noise,
Gaussian blur,
JPEG compression



JPEG
compression

Holistic Multi-distortion image

HMDD

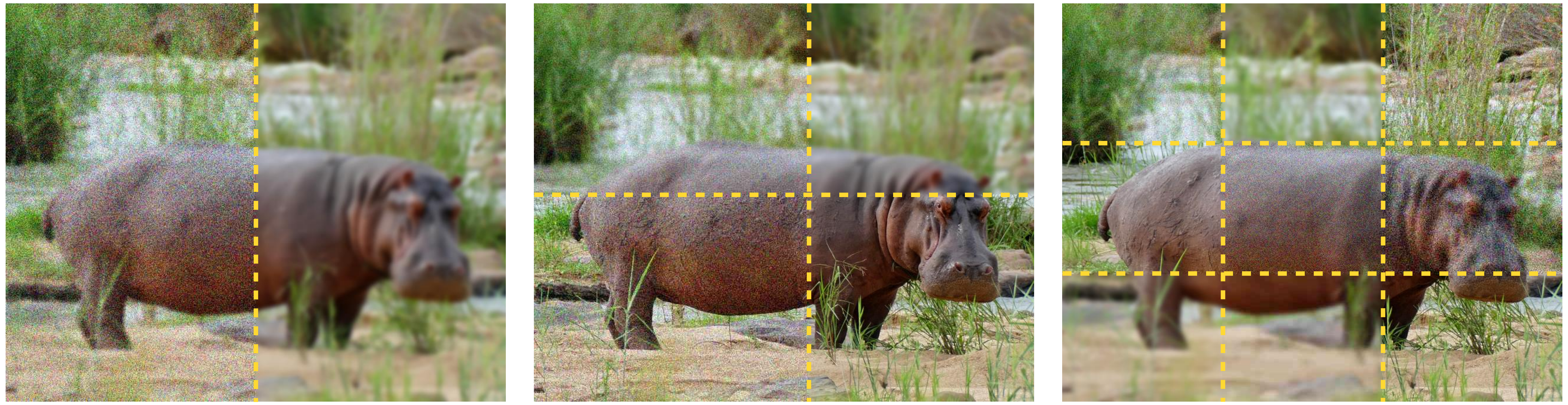
- Based on DIV2K (Agustsson et al. 2017)
- HMDD
  - Gaussian blur, noise, JPEG
- HMDD-r
  - Snow, F-noise, Defocused blur

Distortion	Values
Gaussian blur	$\sigma_b \in \{0.5, 1., 1.5, 2., 2.5, 3., 3.5, 4., 4.5\}$
Gaussian noise	$\sigma_n \in \{5, 10, 15, 20, 25, 30, 35, 40, 45\}$
JPEG quality	$q \in \{80, 60, 50, 40, 35, 30, 25, 20, 15\}$
Snow	$\mu_s \in \{0.85, 0.8, 0.75, 0.7, 0.65, 0.6, 0.55\}$ $\sigma_s \in \{4, 4, 5, 5, 5, 5, 6, 6, 6\}$
F-noise	$\alpha_f \in \{500, 250, 150, 100, 80, 60, 40, 25, 15\}$
Defocused blur	$\sigma_d \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 10\}$

HMDD Process

- First, split the clean image I_{gt} into k pieces, where $k \in [2,4,9]$

$$I_{gt} \longrightarrow \{I_{gt}^{1}, \dots, I_{gt}^{k}\}$$



$k = 2$ $\quad$ $k = 4$ $\quad$ $k = 9$

HMDD Process

- Second, apply distortions
  - Randomly choose from the given distortions.
  - Randomly select the intensity of the distortion
- And, $I_{dis}^{i} = D^{i}(I_{gt}^{i}); D^{i} = D_{b}^{i} \circ D_{n}^{i} \circ D_{j}^{i}$, *for* $i = 1, \dots, k$ where $\circ$ denotes function composition.
- $I_{dis} \longleftarrow \{I_{dis}^{1}, \dots, I_{dis}^{k}\}$

$$D_b(\sigma_b) = \begin{cases} \text{Gaussian blur}(\sigma_b) & \text{if } p_b \geq 0.5 \\ \text{Identity} & \text{if } p_b < 0.5 \end{cases}$$

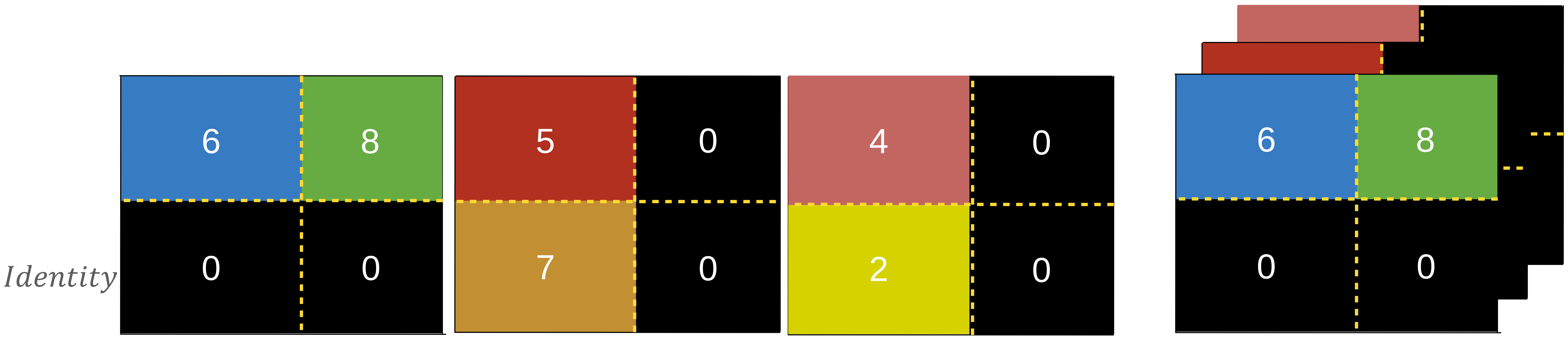
$$D_n(\sigma_n) = \begin{cases} \text{Gaussian noise}(\sigma_n) & \text{if } p_n \geq 0.5 \\ \text{Identity} & \text{if } p_n < 0.5 \end{cases}$$

$$D_j(Q) = \begin{cases} \text{JPEG compression}(Q) & \text{if } p_j \geq 0.5 \\ \text{Identity} & \text{if } p_j < 0.5 \end{cases}$$

p_b, p_n, p_j : arbitrary value between 0 and 1.

Distortion label

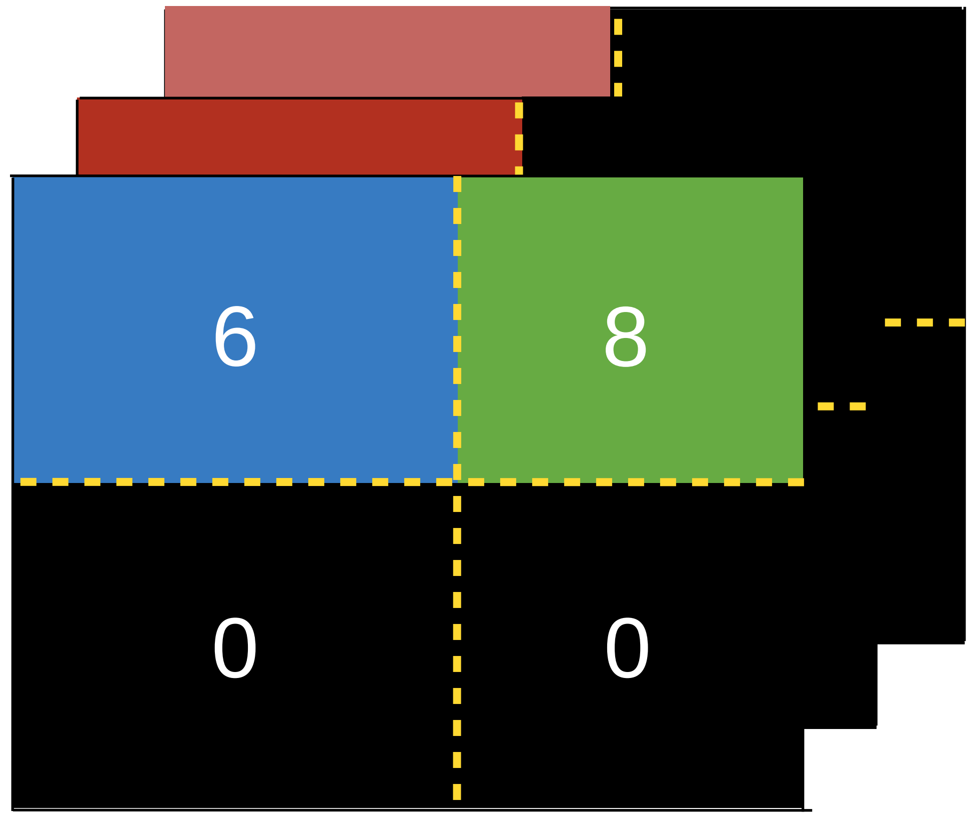
- Creating a spatial distortion label M for learning the Recognition network from distortion image I_d.
- The pixel-wise distortion label M includes the multi-hot embedding for the presence of the distortion along with its strength value.



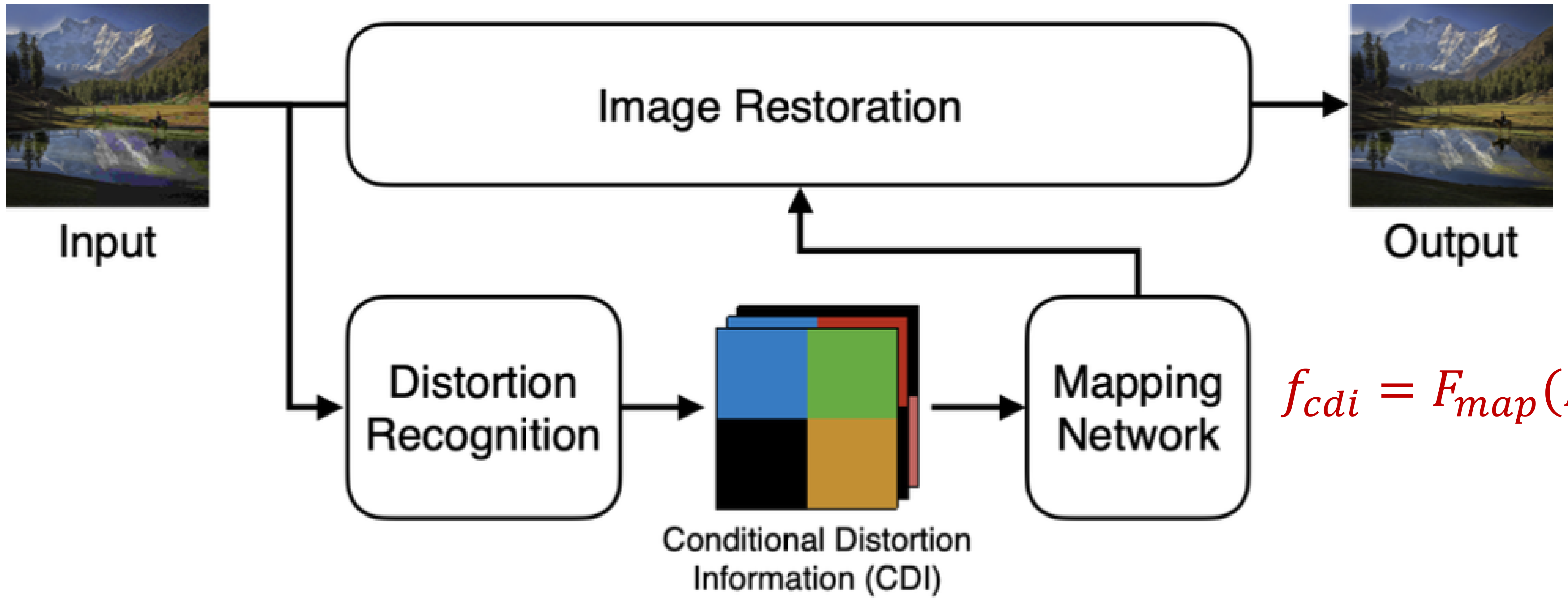
Gaussian blur Gaussian noise JPEG compression Spatial distortion label M

3. Distortion Information-guided network (DIGNet)

Distortion Information-Guided Network



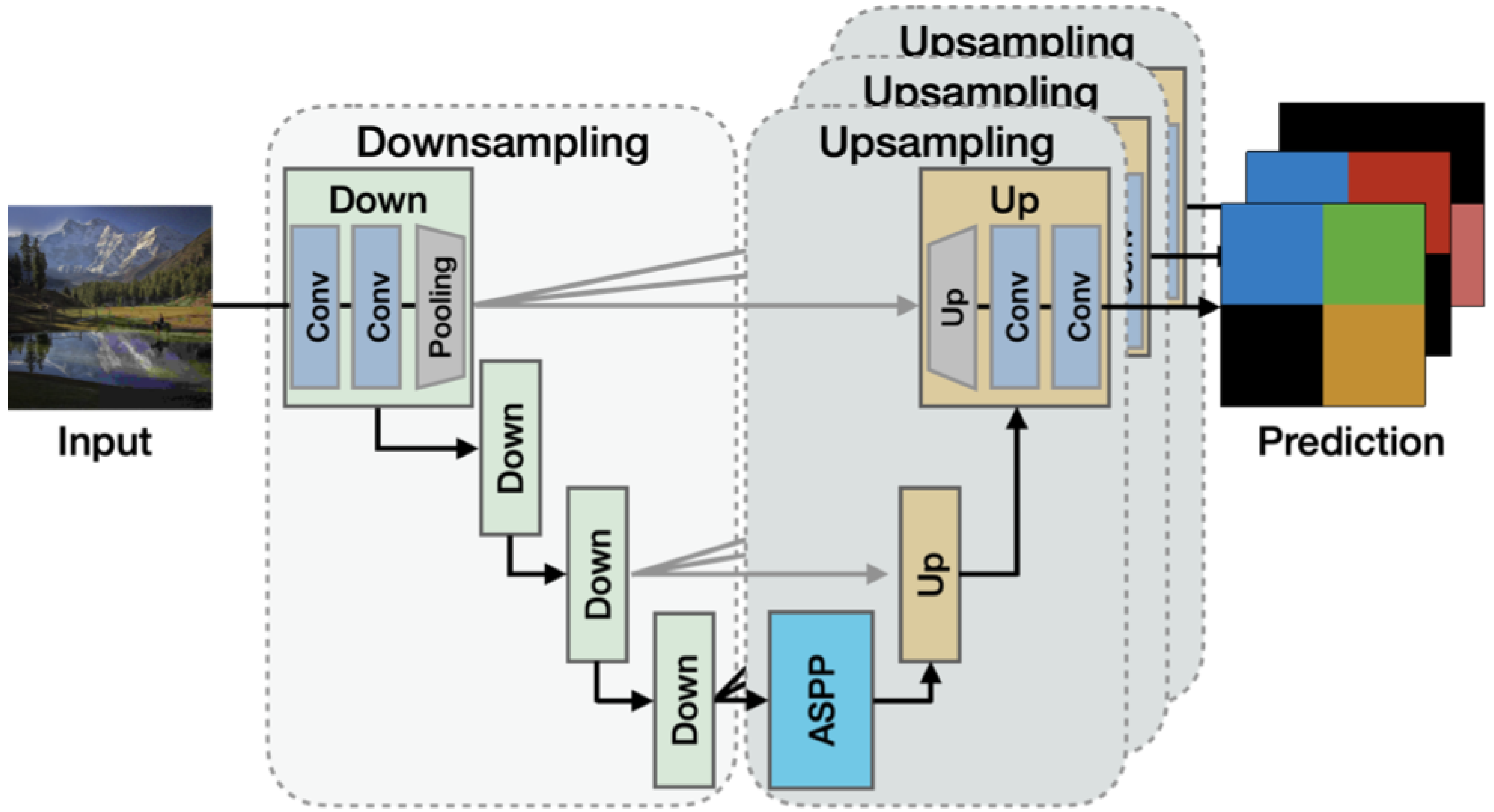
Pixel-wise distortion labels M



$f_{cdi} = F_{map}(\widehat{M})$

- Our proposed restoration framework, DIGNet, is composed of the recognition and the restoration module.
- The recognition module generates the Conditional Distortion Information(CDI) from the predicted distortion.
- The restoration part reconstructs the corrupted image by the distortion guidance from the recognition module

Recognition module

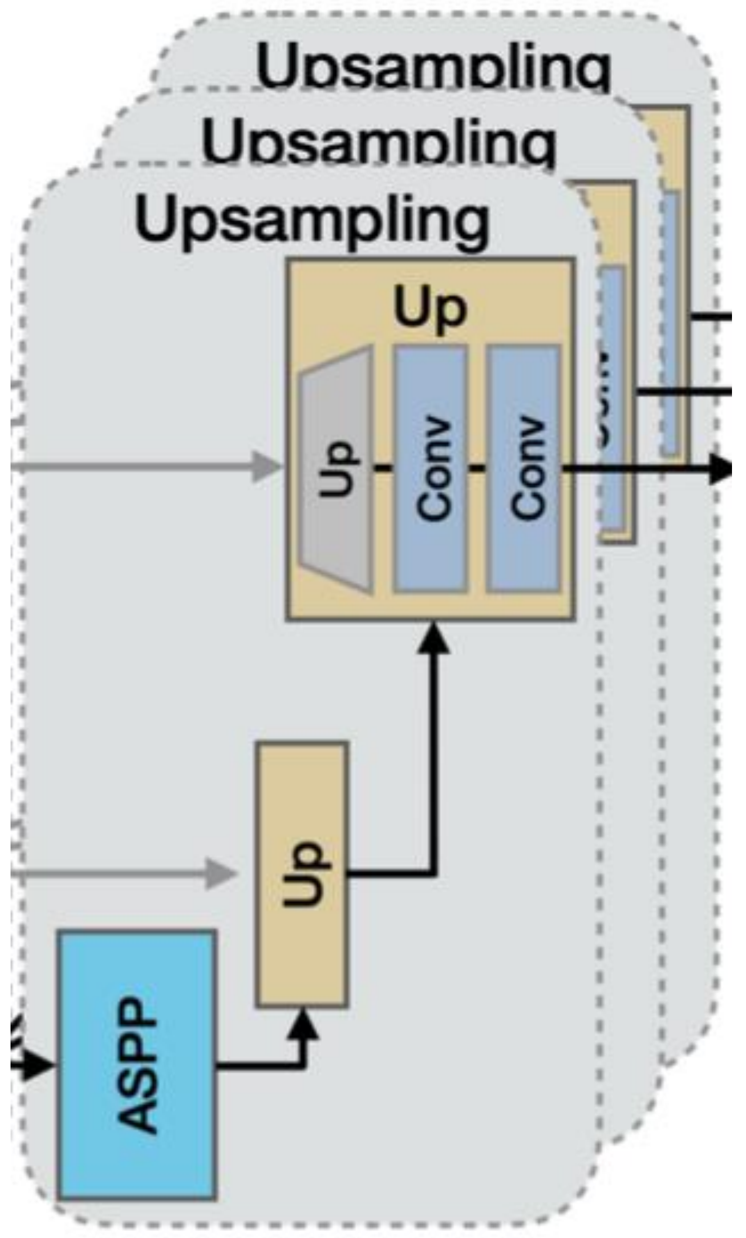


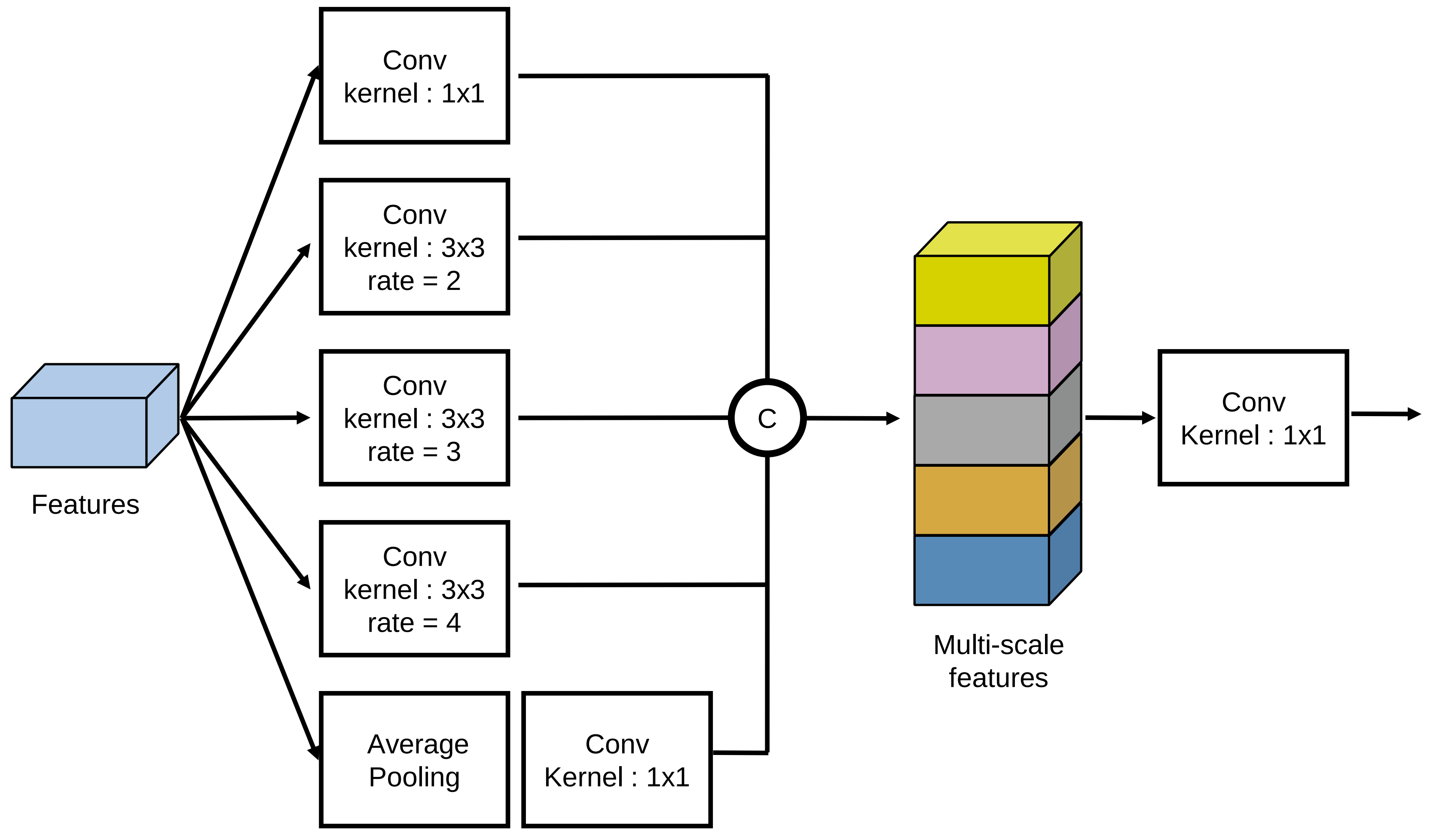
$$F_0 = f_{enc}(X), \quad \widehat{M}^k = f_{dec}^t(F_0)$$

$$\widehat{M} = [\widehat{M}^t], \quad \text{for } t = 1, \ldots, T$$

- The goal of the recognition module is to recognize the distortion types and intensity
- Since our framework assumes that different distortions can be applied for each region, we formulate this as the segmentation tasks.
- Shared encoder & distortion specific representations

Recognition module - ASPP

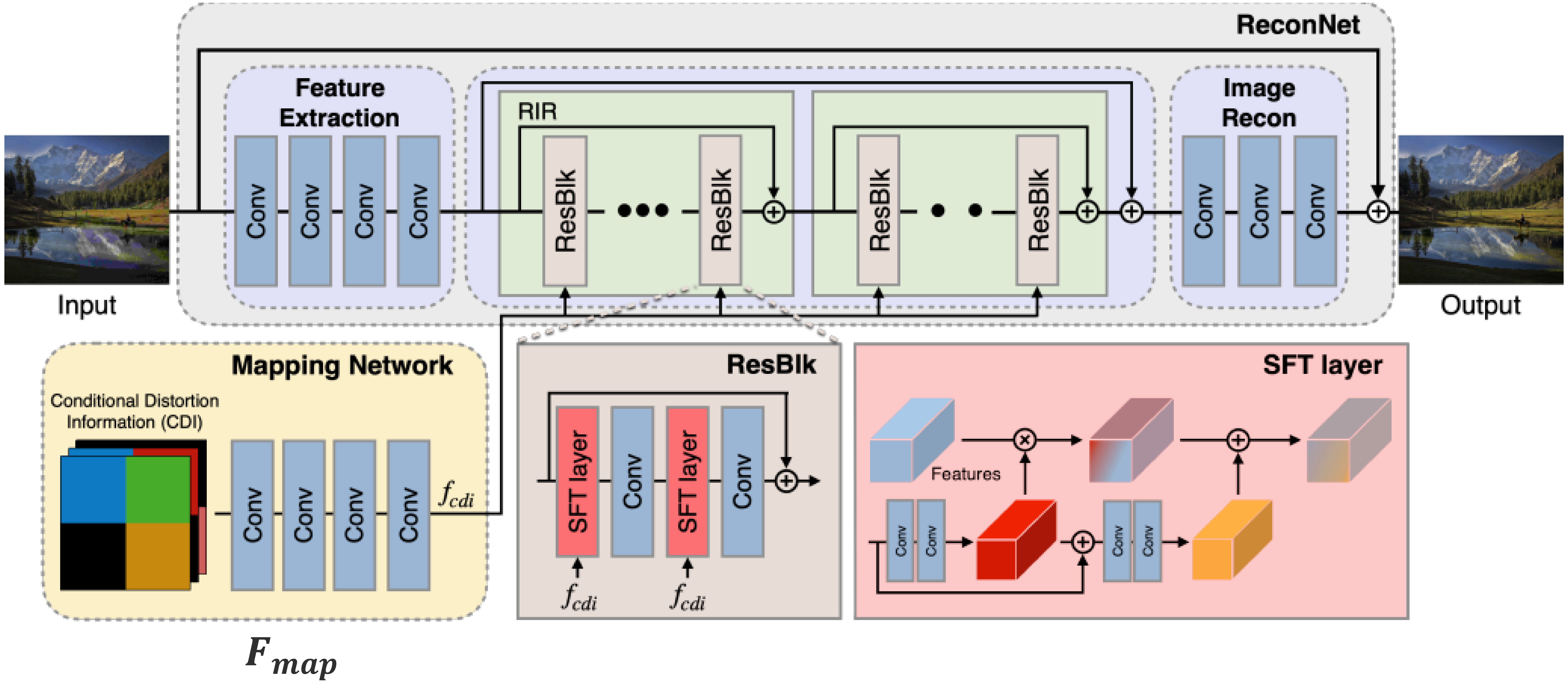




ASPP block

- Our UNet based recognition module contains ASPP block at the first layer of the decoder.
- ASPP is proposed to capture the multi-scale contextual information by overlapping Atrous pooling layers.

Restoration Module

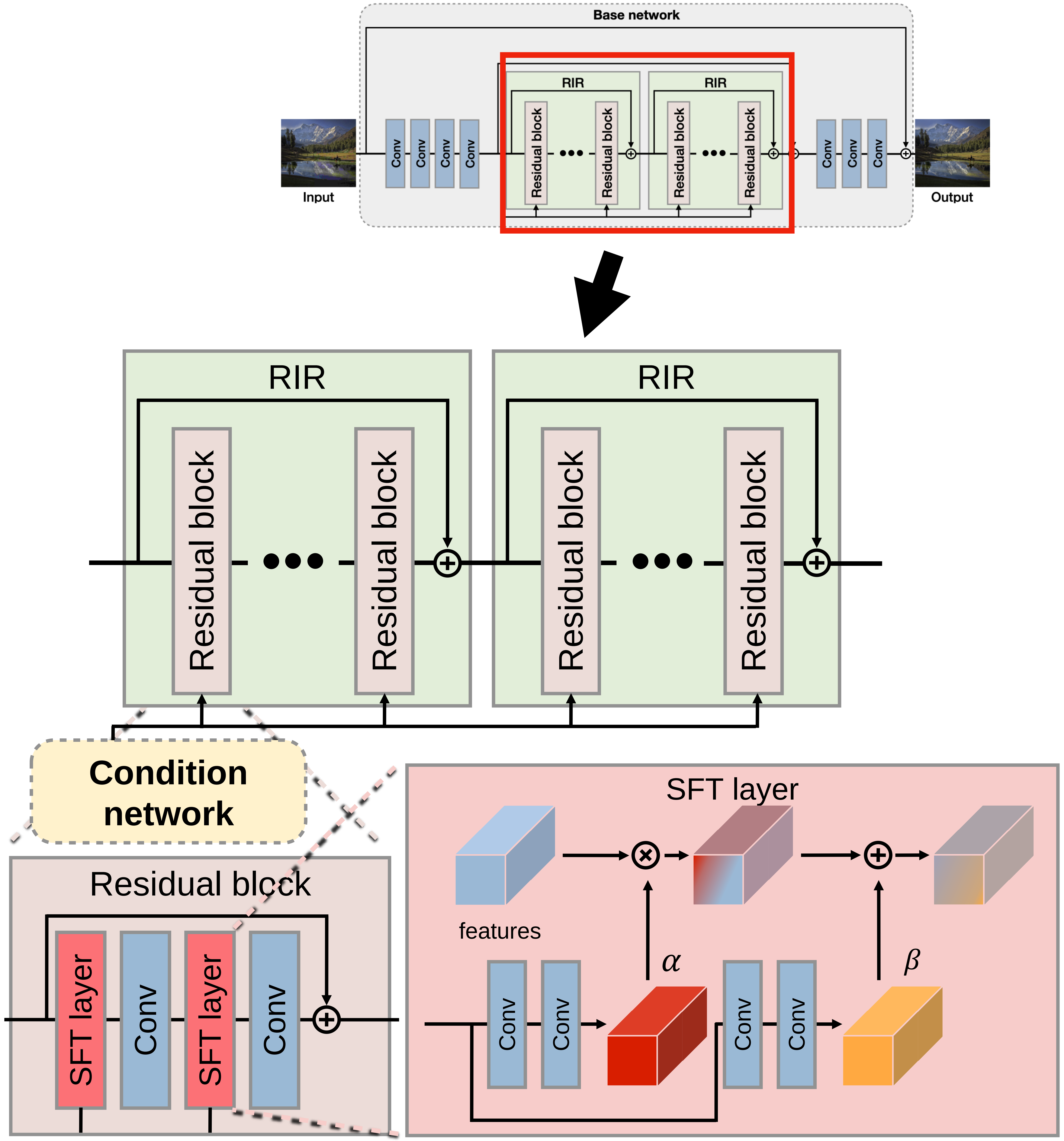


- For restoration with CDI, the mapping network F_{map} produce feature f_{cdi} as follows: $f_{cdi} = F_{map}(\widehat{M})$

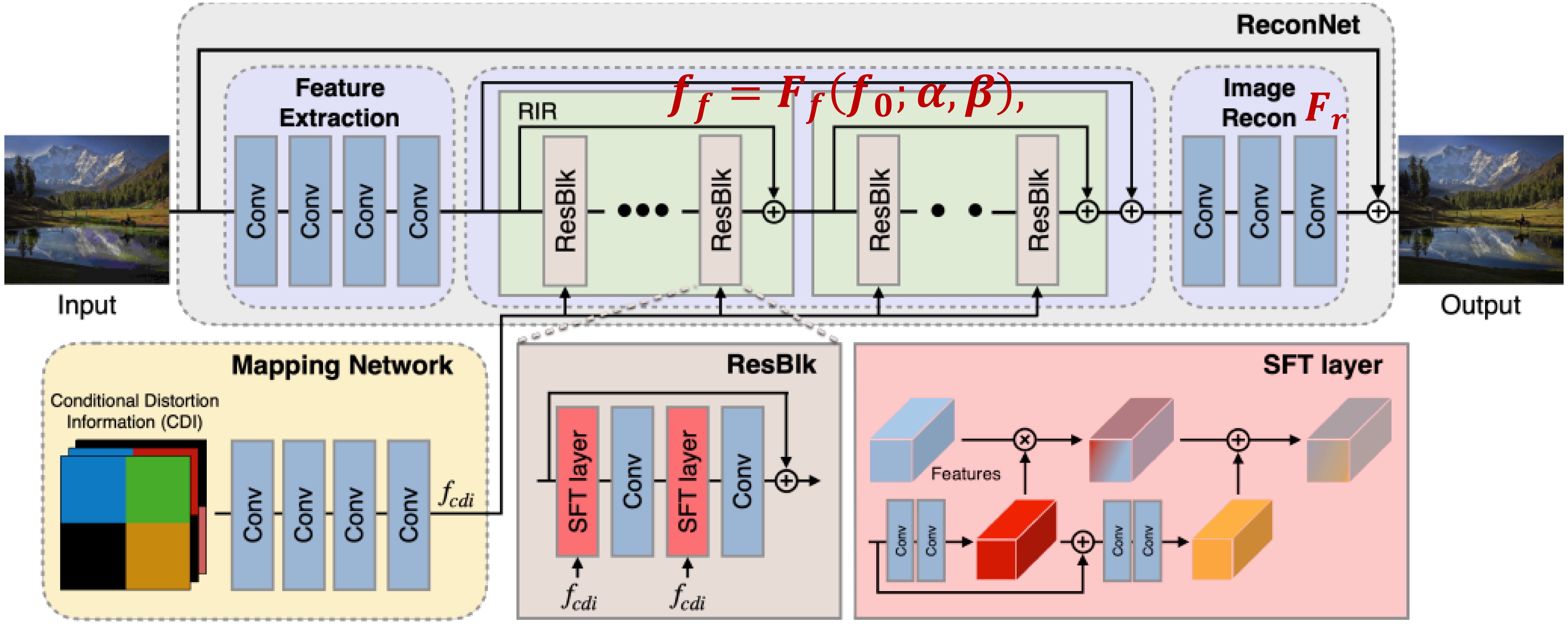
SFT layer

- Spatial Feature Transform (SFT) (Wang et al. 2018)[4] delivers information by scaling and shifting features x through the modulation parameters α, β.

- $SFT(x) = \alpha \otimes x + \beta$
  $\otimes$: element-wise multiplication



Restoration Module



- We now produce intermediate feature f_0 as : $f_f = F_f(f_0; \alpha, \beta), \quad f_0 = f_f + f_0$
- Finally, we generate the cleaned image $\hat{y}$ by using the image reconstruction module F_r as

$$\hat{y} = F_r(f_0) + x$$

4. Experiments

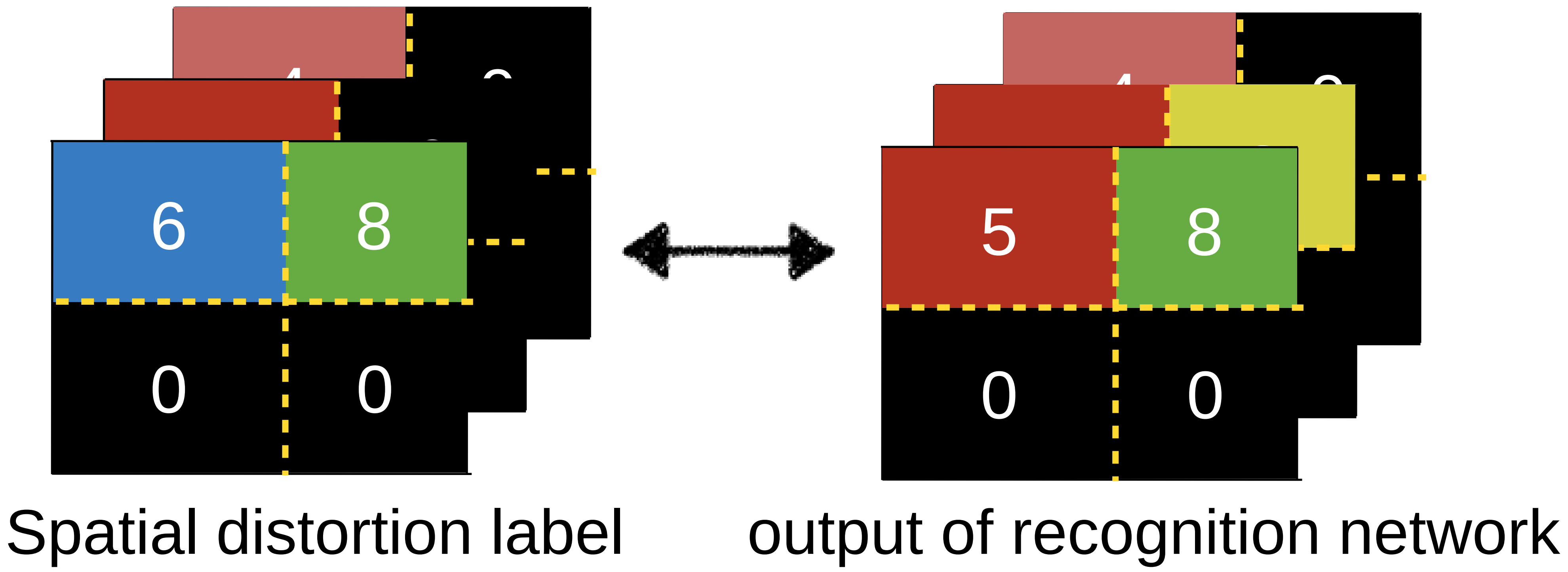
Metrics

1. PSNR, SSIM : pixel driven metrics
   -> To compare the restoration performance

2. Accuracy(Blur, Noise, JPEG)
   : pixel-wise accuracy per channel

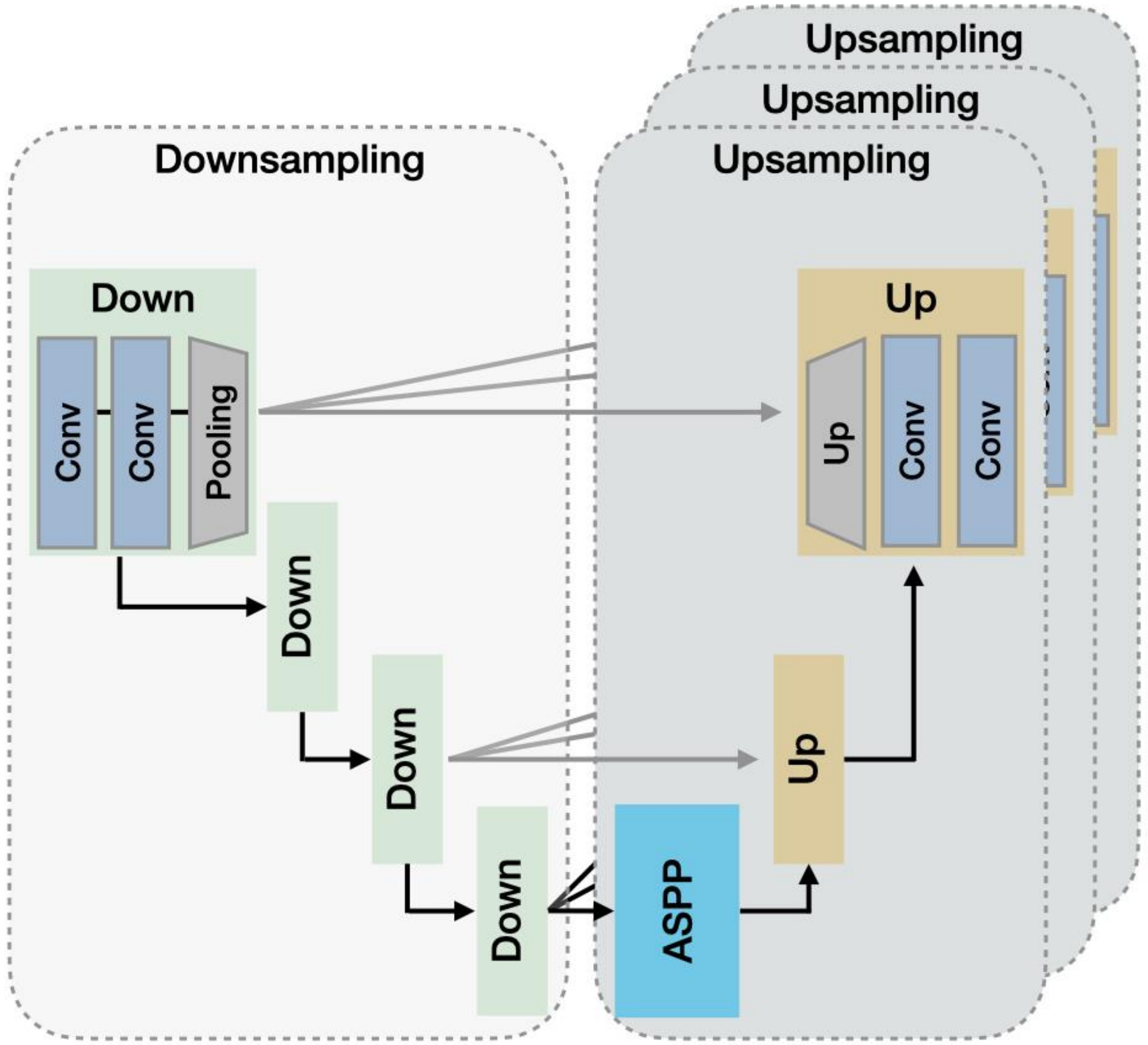
   Accuracy(Pixel)
   : the percentage of pixels that
     match the values of three channels

   -> To compare the performance of recognition module



Analysis of Recognition module

# Down.	# Up.	ASPP	Accuracy (HMDD)				$\rightarrow$ DIGNet	
			G-blur	G-noise	JPEG	Pixel	PSNR	SSIM
2	2		60.76	79.22	87.39	44.59	26.60	0.7559
4	2		71.72	84.26	**94.24**	60.29	26.68	0.7594
4	2	✓	**79.12**	**87.93**	94.01	**64.49**	**26.74**	**0.7634**
4	4	✓	78.77	85.38	93.98	63.01	26.71	0.7628



1. PSNR, SSIM : pixel driven metrics
   -> To compare the restoration performance

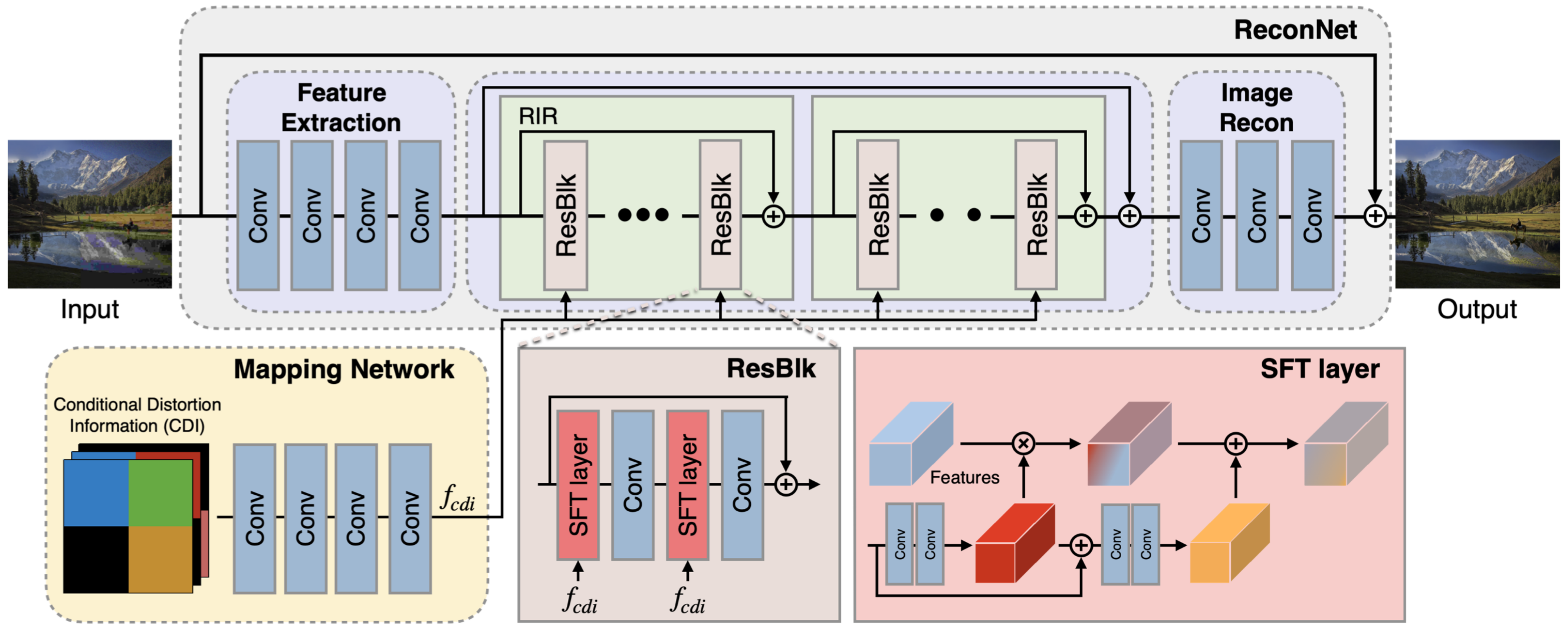
2. Accuracy(Blur, Noise, JPEG)
   : pixel-wise accuracy per channel

   Accuracy(Pixel)
   : the percentage of pixels that
     match the values of three channels

   -> To compare the performance of recognition module

Quantitative Comparisons with other methods

Method	HMDD		HMDD-r	
	PSNR	SSIM	PSNR	SSIM
OWAN [25]	23.52	0.5948	22.25	0.5694
+ CDI	25.96	0.7323	27.13	0.7885
MEPSNet [15]	25.77	0.7257	26.08	0.7757
+ CDI	26.60	0.7606	28.43	0.8270
EDSR [19]	26.25	0.7461	26.70	0.7795
+ CDI	26.63	0.7622	28.56	0.8401
Ours w/o CDI	26.52	0.7528	27.91	0.8177
+ CDI (DIGNet)	**26.74**	**0.7634**	**28.70**	**0.8560**



Qualitative Comparisons with other methods

5. Conclusion

Conclusion

- We introduce **a novel dataset HMDD**, which is a more realistic dataset by fusing two previously proposed dataset generation methods, each focusing on sequential and spatial distortions, respectively.

- Our **DIGNet** outperforms other methods in multi-distortion image reconstruction performance by **predicting spatial distortion information through recognition module** and then **injecting it into the feature map of the reconstruction module.**

- In addition, when the spatial distortion information is applied to the other models, other methods show high performance improvement.

Thank you