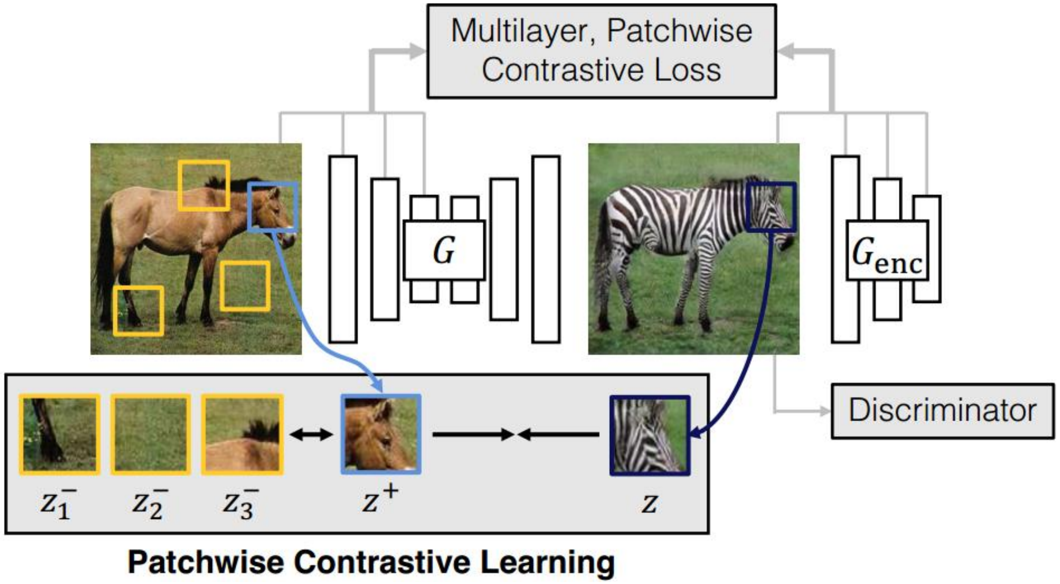


1.把对比学习应用到Image-to-Image中。
2.只需要训练一个生成器和判别器。
3.使用最大化输入输出图像块的**互信息**（mutual information）来替代循环一致性损失，使用infoNCEloss作为对比损失函数， 来学会一个生成器。

4.应用生成器（由编码器和解码器组成)中的编码器将对应的图像块之间相互联系起来，通过MultilayerPatchwiseContrastiveLoss使得编码器专注于两个域之间共性的部分如形状，而忽略两个域之间的差异性部分如纹理。

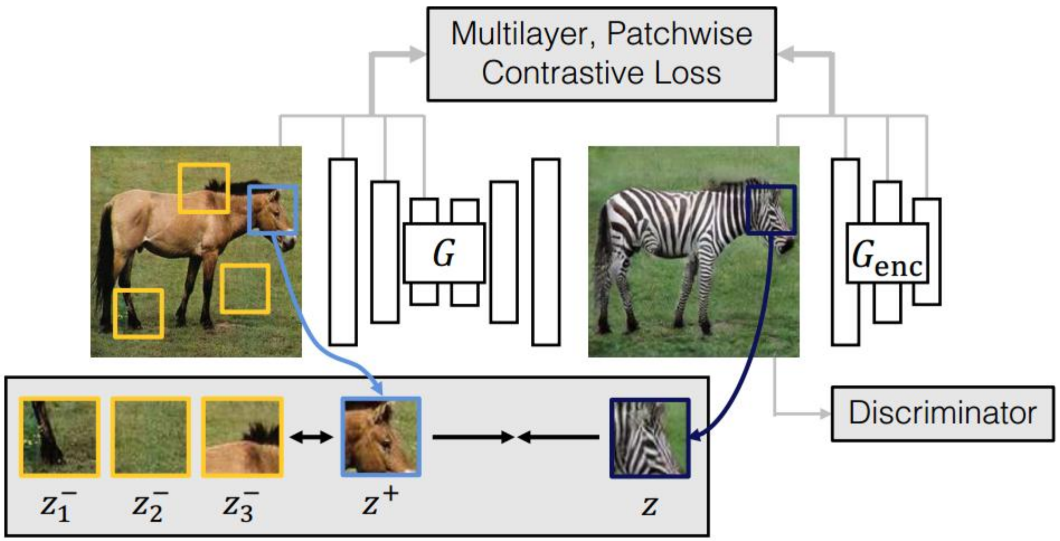
5.CUT从单张图像本身中提取负样本图像块的效果要好于从整个数据集中其他的图像中提取，因此CUT甚至可以在单张图像上实现图像转换。

Contrastive Learning for Unpaired Image-to-Image Translation ECCV'20

Adversarial loss. We use an adversarial loss [21], to encourage the output to be visually similar to images from the target domain, as follows:

$$\mathcal{L}_{\text{GAN}}(G, D, X, Y) = \mathbb{E}_{\boldsymbol{y}\sim Y} \log D(\boldsymbol{y}) + \mathbb{E}_{\boldsymbol{x}\sim X} \log(1 - D(G(\boldsymbol{x}))). \quad (1)$$

→ 外观约束



Patchwise Contrastive Learning

最大化互信息

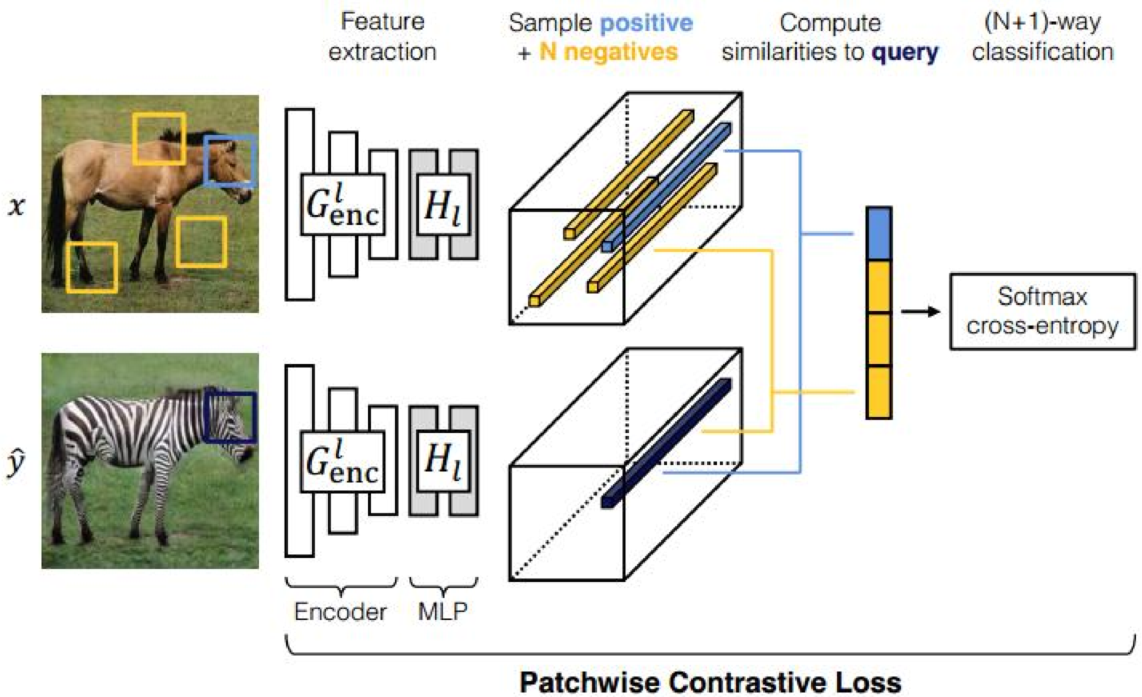
$$\ell(\boldsymbol{v}, \boldsymbol{v}^+, \boldsymbol{v}^-) = -\log\left[\frac{\exp(\boldsymbol{v}\cdot\boldsymbol{v}^+/\tau)}{\exp(\boldsymbol{v}\cdot\boldsymbol{v}^+/\tau) + \sum_{n=1}^{N}\exp(\boldsymbol{v}\cdot\boldsymbol{v}_n^-/\tau)}\right]. \quad (2)$$

→ 内容约束

Adversarial loss. We use an adversarial loss [21], to encourage the output to be visually similar to images from the target domain, as follows:

$$\mathcal{L}_{\mathrm{GAN}}(G, D, X, Y) = \mathbb{E}_{\boldsymbol{y} \sim Y} \log D(\boldsymbol{y}) + \mathbb{E}_{\boldsymbol{x} \sim X} \log(1 - D(G(\boldsymbol{x}))). \tag{1}$$

→ 外观约束

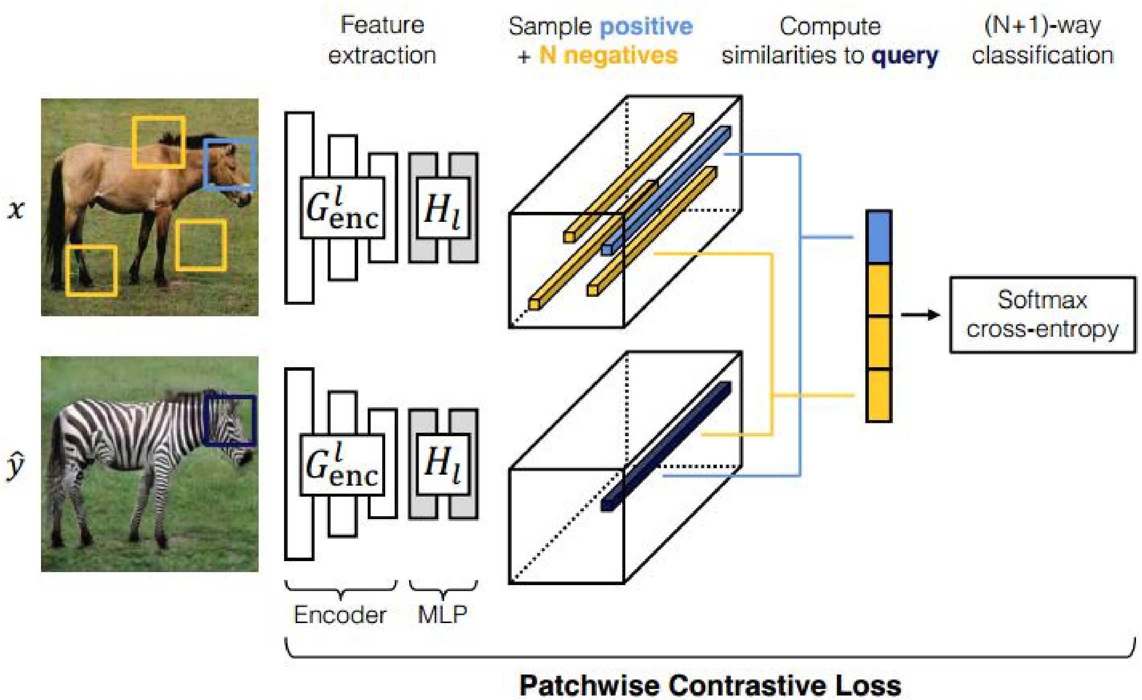


最大化互信息

$$\ell(\boldsymbol{v}, \boldsymbol{v}^+, \boldsymbol{v}^-) = -\log \left[\frac{\exp(\boldsymbol{v} \cdot \boldsymbol{v}^+ / \tau)}{\exp(\boldsymbol{v} \cdot \boldsymbol{v}^+ / \tau) + \sum_{n=1}^{N} \exp(\boldsymbol{v} \cdot \boldsymbol{v}_n^- / \tau)} \right]. \tag{2}$$

→ 内容约束

Multilayer, patchwise contrastive learning.

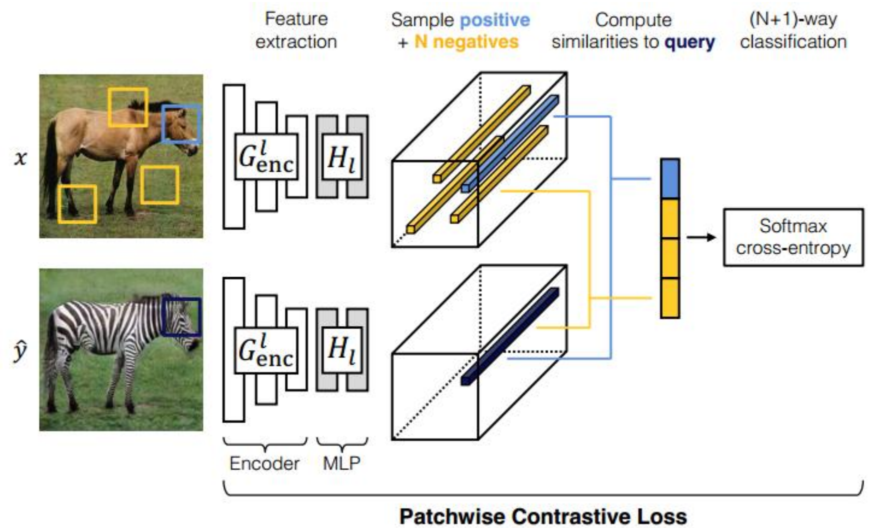


$$\ell(\boldsymbol{v}, \boldsymbol{v}^+, \boldsymbol{v}^-) = -\log\left[\frac{\exp(\boldsymbol{v}\cdot\boldsymbol{v}^+/\tau)}{\exp(\boldsymbol{v}\cdot\boldsymbol{v}^+/\tau) + \sum_{n=1}^{N}\exp(\boldsymbol{v}\cdot\boldsymbol{v}_n^-/\tau)}\right]. \qquad (2)$$

$$\{z_l\}_L = \{H_l(G^l_{\text{enc}}(x)\}_L$$

$$\mathcal{L}_{\text{PatchNCE}}(G, H, X) = \mathbb{E}_{\boldsymbol{x}\sim X}\sum_{l=1}^{L}\sum_{s=1}^{S_l}\ell(\hat{\boldsymbol{z}}_l^s, \boldsymbol{z}_l^s, \boldsymbol{z}_l^{S\setminus s}). \qquad (3)$$

Multilayer, patchwise contrastive learning.



$$\ell(\boldsymbol{v}, \boldsymbol{v}^+, \boldsymbol{v}^-) = -\log \left[\frac{\exp(\boldsymbol{v} \cdot \boldsymbol{v}^+ / \tau)}{\exp(\boldsymbol{v} \cdot \boldsymbol{v}^+ / \tau) + \sum_{n=1}^{N} \exp(\boldsymbol{v} \cdot \boldsymbol{v}_n^- / \tau)} \right]. \quad (2)$$

$$\{z_l\}_L = \{H_l(G^l_{\text{enc}}(x)\}_L$$

$$\mathcal{L}_{\text{PatchNCE}}(G, H, X) = \mathbb{E}_{\boldsymbol{x} \sim X} \sum_{l=1}^{L} \sum_{s=1}^{S_l} \ell(\hat{\boldsymbol{z}}_l^s, \boldsymbol{z}_l^s, \boldsymbol{z}_l^{S \setminus s}). \quad (3)$$

$$\mathcal{L}_{\text{external}}(G, H, X) = \mathbb{E}_{\boldsymbol{x} \sim X, \tilde{\boldsymbol{z}} \sim Z^-} \sum_{l=1}^{L} \sum_{s=1}^{S_l} \ell(\hat{\boldsymbol{z}}_l^s, \boldsymbol{z}_l^s, \tilde{\boldsymbol{z}}_l), \quad (4)$$

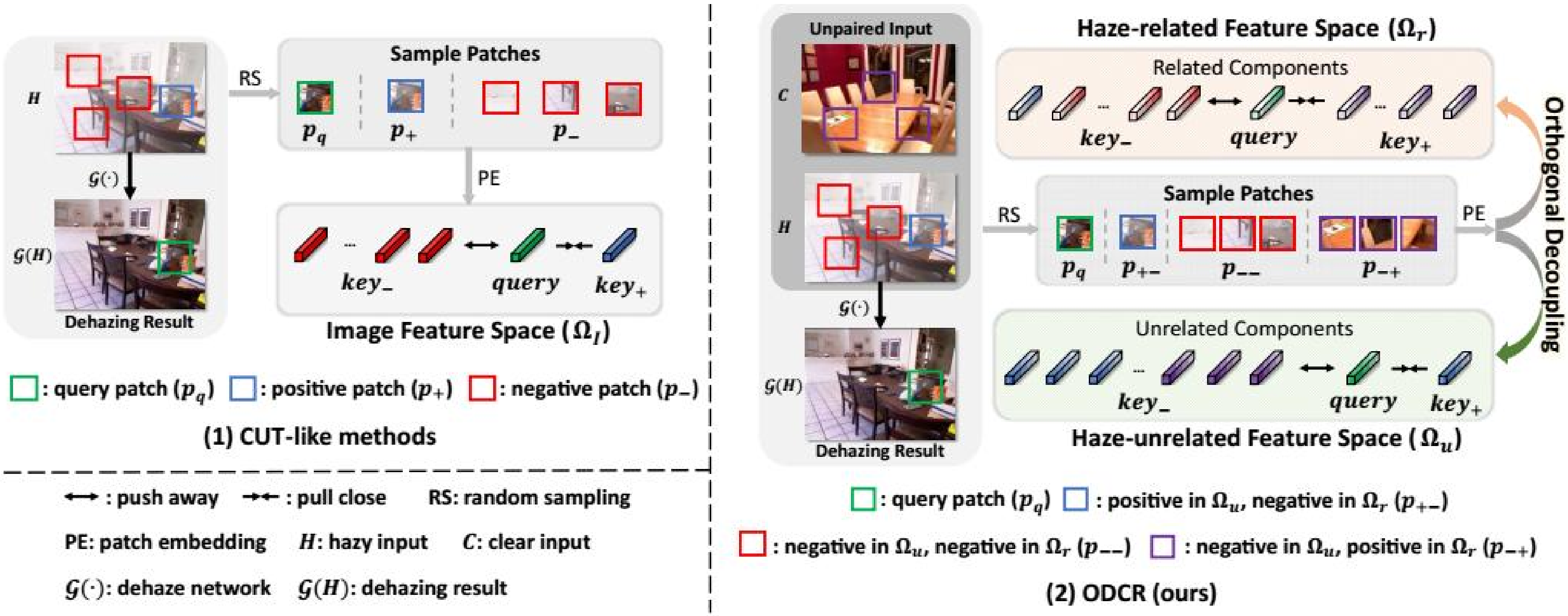
Adversarial loss. We use an adversarial loss [21], to encourage the output to be visually similar to images from the target domain, as follows:

$$\mathcal{L}_{\mathrm{GAN}}(G, D, X, Y) = \mathbb{E}_{\boldsymbol{y} \sim Y} \log D(\boldsymbol{y}) + \mathbb{E}_{\boldsymbol{x} \sim X} \log(1 - D(G(\boldsymbol{x}))). \quad (1)$$

$$\ell(\boldsymbol{v}, \boldsymbol{v}^+, \boldsymbol{v}^-) = -\log \left[\frac{\exp(\boldsymbol{v} \cdot \boldsymbol{v}^+ / \tau)}{\exp(\boldsymbol{v} \cdot \boldsymbol{v}^+ / \tau) + \sum_{n=1}^{N} \exp(\boldsymbol{v} \cdot \boldsymbol{v}_n^- / \tau)} \right]. \quad (2)$$

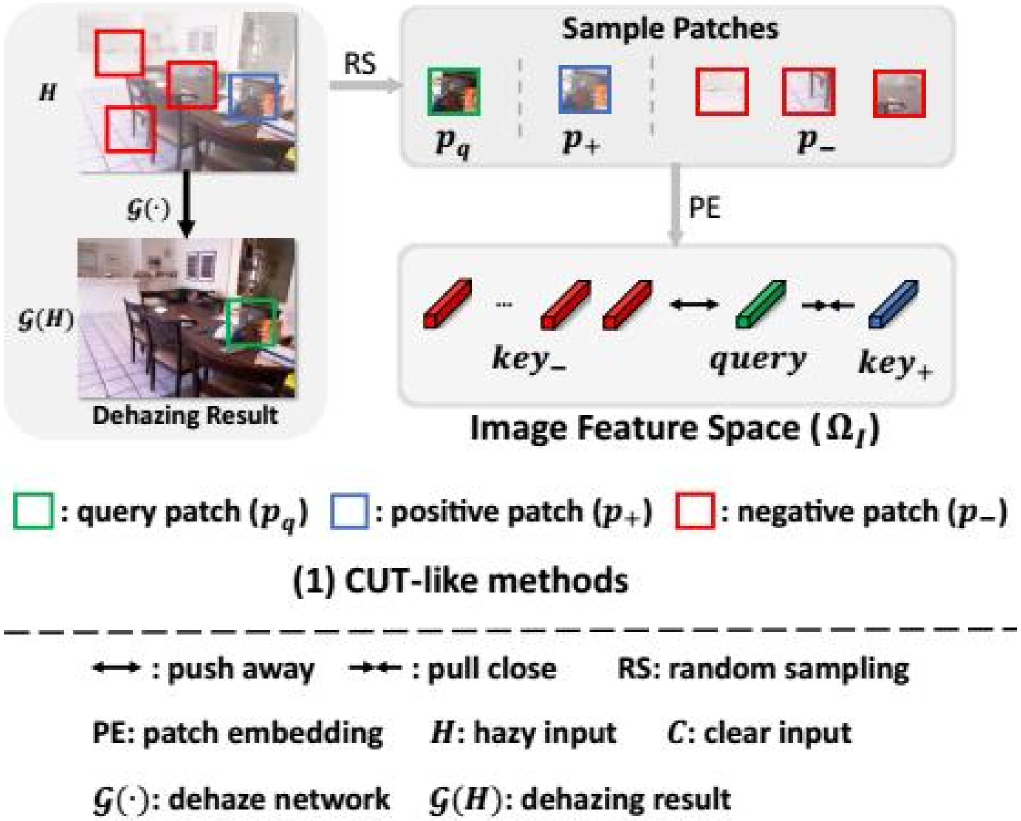
$$\mathcal{L}_{\mathrm{PatchNCE}}(G, H, X) = \mathbb{E}_{\boldsymbol{x} \sim X} \sum_{l=1}^{L} \sum_{s=1}^{S_l} \ell(\hat{\boldsymbol{z}}_l^s, \boldsymbol{z}_l^s, \boldsymbol{z}_l^{S \setminus s}). \quad (3)$$

$$\mathcal{L}_{\mathrm{GAN}}(G, D, X, Y) + \lambda_X \mathcal{L}_{\mathrm{PatchNCE}}(G, H, X) + \lambda_Y \mathcal{L}_{\mathrm{PatchNCE}}(G, H, Y). \quad (5)$$



图像由影响雾霾程度的雾霾相关特征和与雾霾无关的特征（如纹理和语义信息）组成

ODCR: Orthogonal Decoupling Contrastive Regularization for Unpaired Image Dehazing cvpr’24



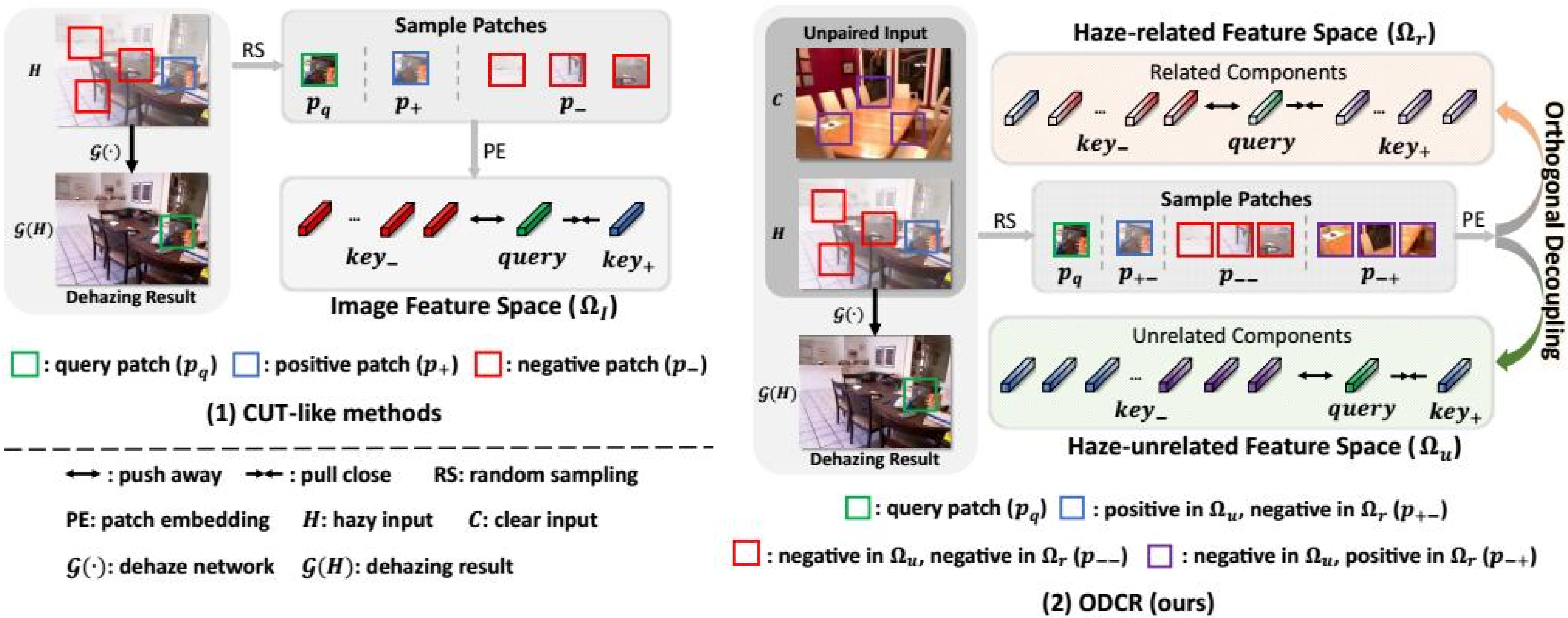
大多数当前的不成对图像去雾（UID）策略采用循环GAN框架

大气散射模型

$$\mathbf{I}(x) = \mathbf{J}(x)t(x) + \mathbf{A}(1 - t(x)) \tag{1}$$

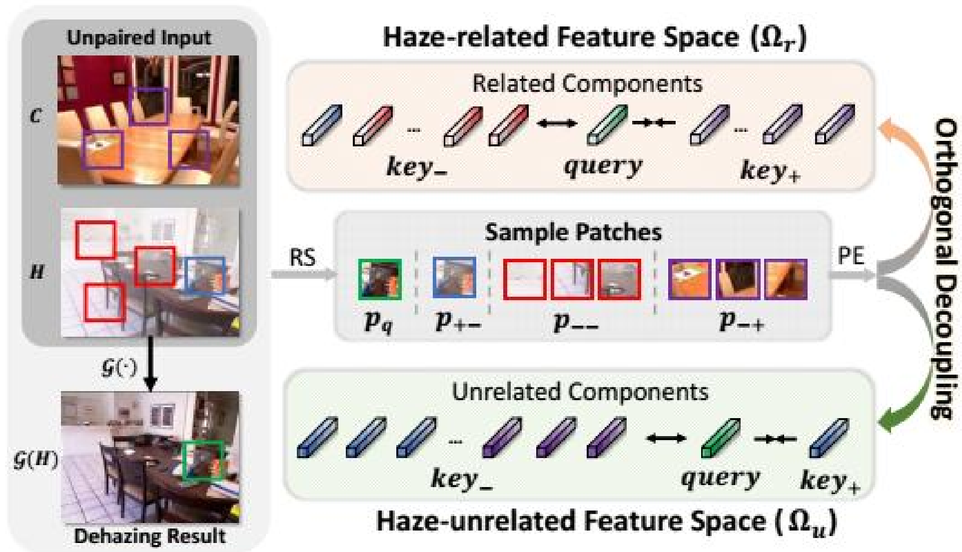
- 如何将图像特征解耦成具有极低关联性的特征
- 如何判断已经解耦的特征是否与雾的污染有关

样本划分



H
RS
Sample Patches
p_q
p_+
p_-
$\mathcal{G}(\cdot)$
PE
$\mathcal{G}(H)$
key_-
$query$
key_+
Dehazing Result
Image Feature Space (Ω_I)
: query patch (p_q)
: positive patch (p_+)
: negative patch (p_-)
(1) CUT-like methods
: push away
: pull close
RS: random sampling
PE: patch embedding
H: hazy input
C: clear input
$\mathcal{G}(\cdot)$: dehaze network
$\mathcal{G}(H)$: dehazing result
Unpaired Input
C
Haze-related Feature Space (Ω_r)
Related Components
key_-
$query$
key_+
Orthogonal Decoupling
Sample Patches
RS
PE
p_q
p_{+-}
p_{--}
p_{-+}
Unrelated Components
key_-
$query$
key_+
Haze-unrelated Feature Space (Ω_u)
Dehazing Result
: query patch (p_q)
: positive in Ω_u, negative in Ω_r (p_{+-})
: negative in Ω_u, negative in Ω_r (p_{--})
: negative in Ω_u, positive in Ω_r (p_{-+})
(2) ODCR (ours)

样本划分



: query patch (p_q) : positive in Ω_u, negative in Ω_r (p_{+-})

: negative in Ω_u, negative in Ω_r (p_{--}) : negative in Ω_u, positive in Ω_r (p_{-+})

(2) ODCR (ours)

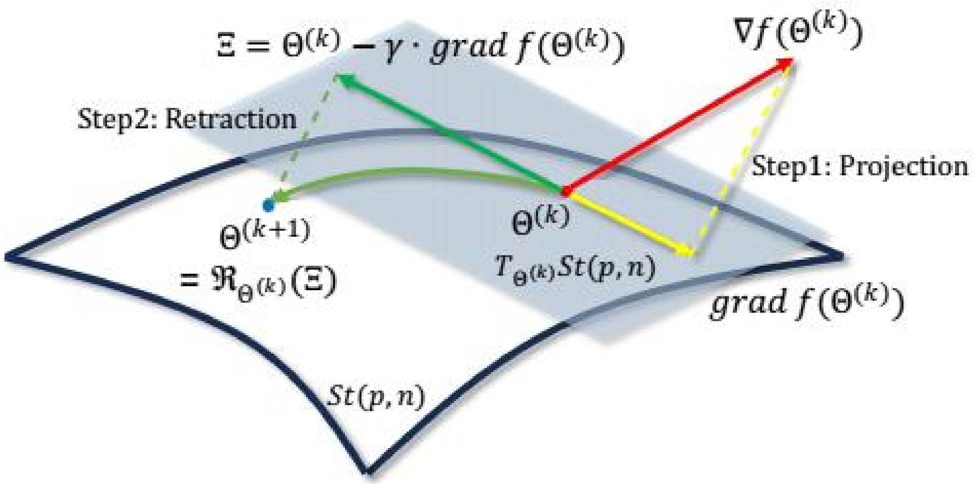
- p_{+-} : the patch in H with the same position as p_q;
- p_{-+} : all patches in C;
- p_{--} : all patches other than p_{+-} in H.

- 如何将图像特征解耦成具有极低关联性的特征

$$z_k = \mathcal{H}_\Theta(\mathcal{G}^i_{enc}(p_k)), s.t. \Theta^T\Theta = I \qquad (2)$$

Stiefel流形的几何优化

$$grad\ f(\Theta) = \nabla f(\Theta) - \frac{1}{2}\Theta\Theta^T\nabla f(\Theta) - \frac{1}{2}\Theta\nabla f(\Theta)^T\Theta \qquad (3)$$



SGD、
ADAM ×

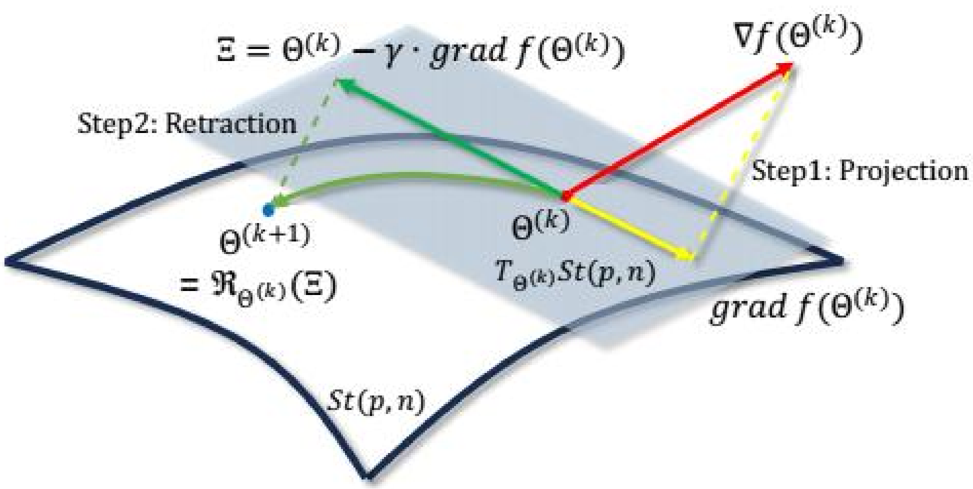
Stiefel流形的几何优化

$$grad\ f(\Theta) = \nabla f(\Theta) - \frac{1}{2}\Theta\Theta^T\nabla f(\Theta) - \frac{1}{2}\Theta\nabla f(\Theta)^T\Theta \tag{3}$$

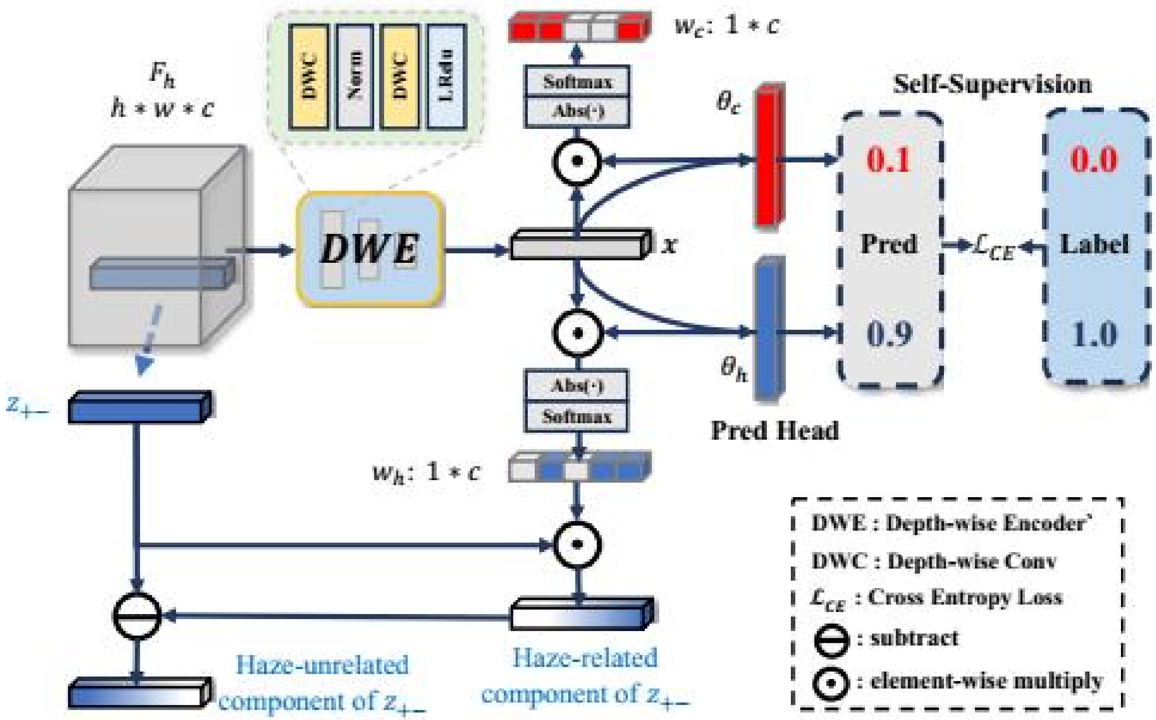
$$\Xi = \Theta^{(k)} - \gamma grad\ f(\Theta^{(k)}) \tag{4}$$

$$\mathfrak{R}_{\Theta^{(k)}}(\Xi) = (\Theta^{(k)} + \Xi)(I + \Xi^T\Xi)^{-\frac{1}{2}} \tag{5}$$

$$\mathfrak{R}_{\Theta^{(k)}}(\Xi)^T\mathfrak{R}_{\Theta^{(k)}}(\Xi) = I \tag{6}$$

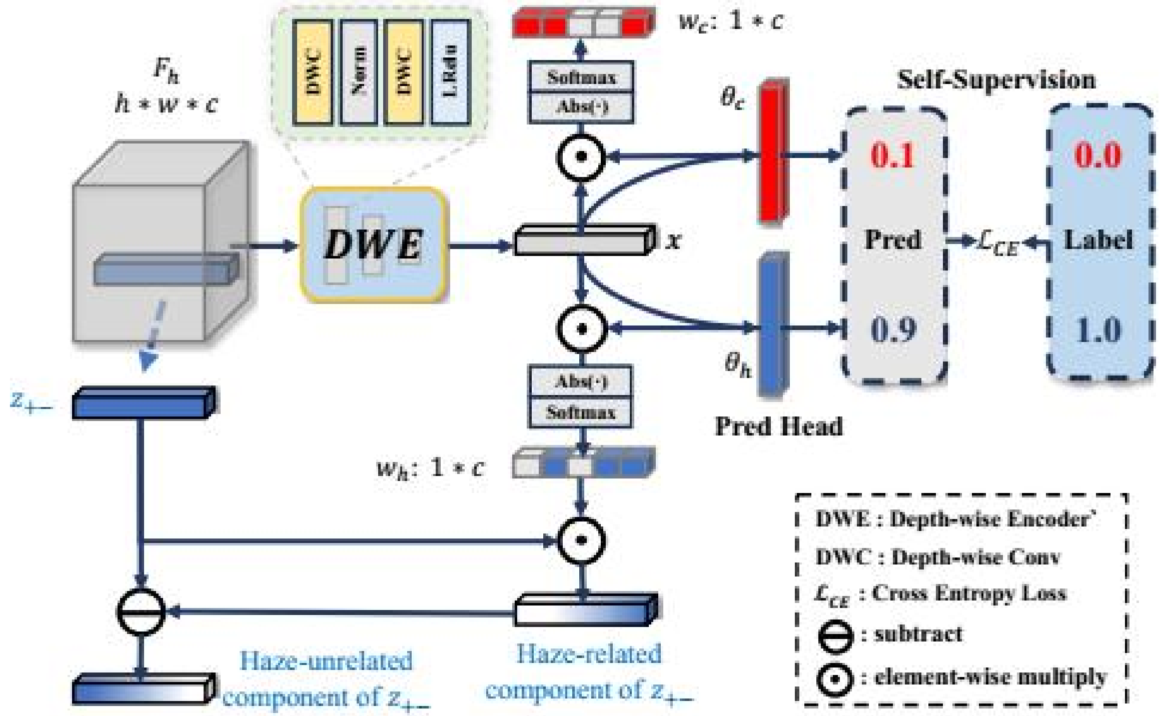


- 如何判断已经解耦的特征是否与雾的污染有关



$$
\begin{aligned}
\mathcal{L}_{CE} = & y_h log(\theta_h^T x) + (1 - y_h) log(1 - \theta_h^T x) \\
& + y_c log(\theta_c^T x) + (1 - y_c) log(1 - \theta_c^T x)
\end{aligned} \tag{8}
$$

- 如何判断已经解耦的特征是否与雾的污染有关



$$w_h = softmax(abs(\theta_h \odot x)) \quad (9)$$

$$w_c = softmax(abs(\theta_c \odot x)) \quad (10)$$

$$l(w, z_1, z_2) = exp(w \odot {z_2}^T \times z_1/\tau) \tag{11}$$

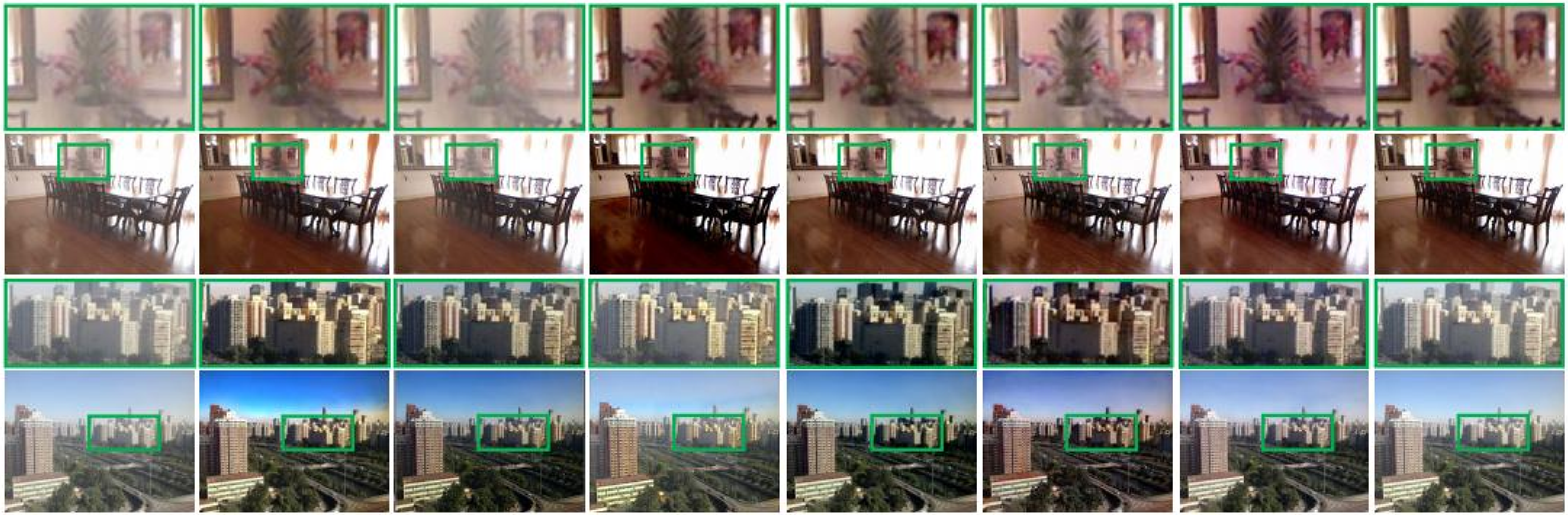
$$\mathcal{P} = l(w_h, z, z_{+-}) + \sum_{n=1}^{N_{-+}} l(w_c, z, {z_{-+}}^n) \tag{12}$$

$$\begin{aligned} \mathcal{N} =& l((\mathbf{1} - w_h), z, z_{+-}) + \sum_{n=1}^{N_{--}} l(\mathbf{1}, z, {z_{--}}^n) \\ &+ \sum_{n=1}^{N_{-+}} l((\mathbf{1} - w_c), z, {z_{-+}}^n) \end{aligned} \tag{13}$$

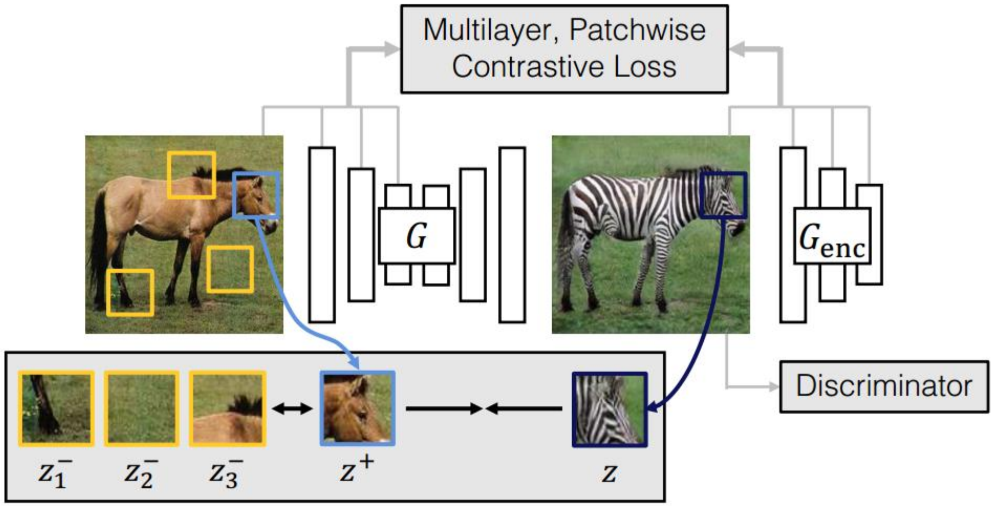
$$\mathcal{L}_{WPNCE} = -log(\frac{\mathcal{P}}{\mathcal{P} + \mathcal{N}}) \tag{14}$$

$$\mathcal{L} = \mathcal{L}_{GAN} + \mathcal{L}_{WPNCE} + \mathcal{L}_{CE} + \mathcal{L}_{idt} \tag{15}$$

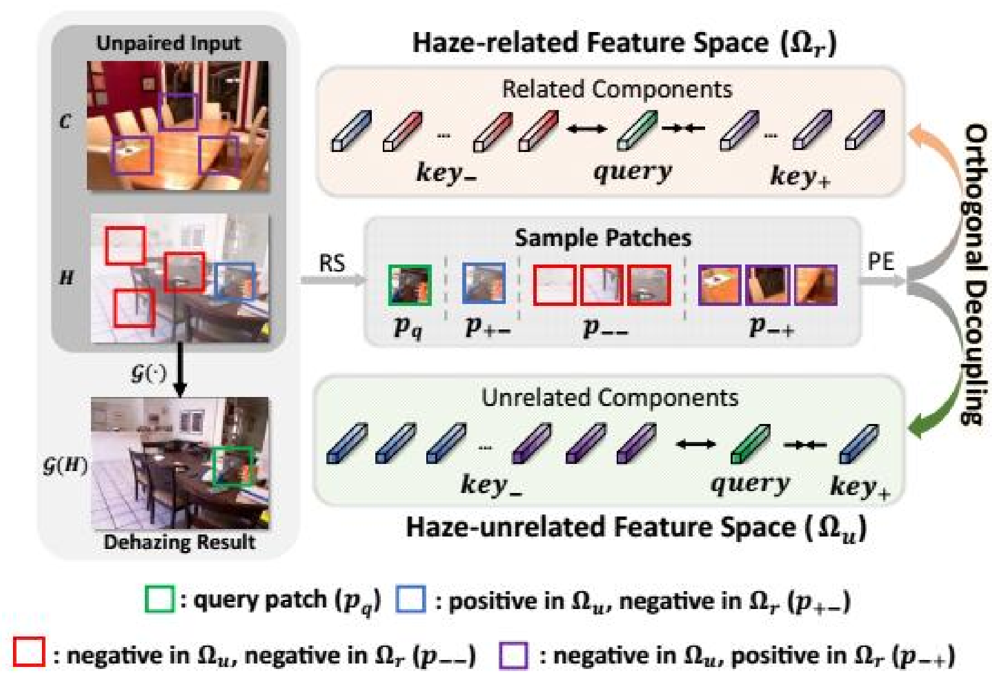
	Method	SOTS-indoor [25]		SOTS-outdoor [25]		NH-HAZE 2 [1]		Overhead	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	#Param (M)	Latency (ms)
Paired	DehazeNet [6]	19.82	0.818	24.75	0.927	10.62	0.521	0.009	0.919
	AOD-Net [24]	20.51	0.816	24.14	0.920	12.33	0.631	0.002	0.390
	MSCNN [37]	19.84	0.833	14.62	0.908	11.74	0.566	0.008	0.619
	GDN [31]	**32.16**	**0.983**	17.69	0.841	12.04	0.557	0.956	9.905
Unpaired	DCP [18]	13.10	0.699	19.13	0.815	14.90	0.668	-	-
	CycleGAN [57]	21.34	0.898	20.55	0.856	13.95	0.689	11.38	10.22
	CycleDehaze [15]	20.11	0.854	21.31	0.899	14.12	0.701	11.38	10.19
	YOLY [26]	15.84	0.819	14.75	0.857	13.38	0.595	32.00	-
	USID-Net [28]	21.41	0.894	23.89	0.919	15.62	0.740	3.780	31.01
	RefineDNet [54]	24.36	0.939	19.84	0.853	14.20	0.754	65.80	248.5
	D^4 [47]	25.42	0.932	<u>25.83</u>	<u>0.956</u>	14.52	0.709	10.70	28.08
	CUT [35]	24.30	0.911	23.67	0.904	<u>15.92</u>	<u>0.758</u>	11.38	10.06
	ODCR (ours)	<u>26.32</u>	<u>0.945</u>	**26.16**	**0.960**	**17.56**	**0.766**	11.38	10.14



(a) hazy (b) DCP (c) AOD-Net (d) YOLY (e) D^4 (f) CUT (g) Ours (h) GT



Patchwise Contrastive Learning



(2) ODCR (ours)