

Learning Disentangled Feature Representation for Hybrid-distorted Image Restoration

Xin Li^[0000-0002-6352-6523], Xin Jin, Jianxin Lin, Sen Liu^[0000-0003-3778-8973],
Yaojun Wu^[0000-0002-8138-4186], Tao Yu^[0000-0003-2550-5008], Wei Zhou, and
Zhibo Chen^[0000-0002-8525-5066] *

CAS Key Laboratory of Technology in Geo-spatial Information Processing and
Application System,
University of Science and Technology of China, Hefei 230027, China
{lixin666, jinxustc, linjx}@mail.ustc.edu.cn, elsen@iat.ustc.edu.cn,
{yaojunwu, yutao666, weichou}@mail.ustc.edu.cn, chenzhibo@ustc.edu.cn

Abstract. Hybrid-distorted image restoration (HD-IR) is dedicated to restore real distorted image that is degraded by multiple distortions. Existing HD-IR approaches usually ignore the inherent interference among hybrid distortions which compromises the restoration performance. To decompose such interference, we introduce the concept of Disentangled Feature Learning to achieve the feature-level divide-and-conquer of hybrid distortions. Specifically, we propose the feature disentanglement module (FDM) to distribute feature representations of different distortions into different channels by revising gain-control-based normalization. We also propose a feature aggregation module (FAM) with channel-wise attention to adaptively filter out the distortion representations and aggregate useful content information from different channels for the construction of raw image. The effectiveness of the proposed scheme is verified by visualizing the correlation matrix of features and channel responses of different distortions. Extensive experimental results also prove superior performance of our approach compared with the latest HD-IR schemes.

Keywords: Hybrid-distorted image restoration, feature disentanglement, feature aggregation.

1 Introduction

Nowadays, Image restoration techniques have been applied in various fields, including streaming media, photo processing, video surveillance, and cloud storage, etc. In the process of image acquisition and transmission, raw images are usually contaminated with various distortions due to capturing devices, high ratio compression, transmission, post-processing, etc. Previous image restoration methods focusing on single distortion have been extensively studied [5, 12, 36, 25, 14, 6, 27, 17, 18, 35] and achieved satisfactory performance on the field of super

* Corresponding Author

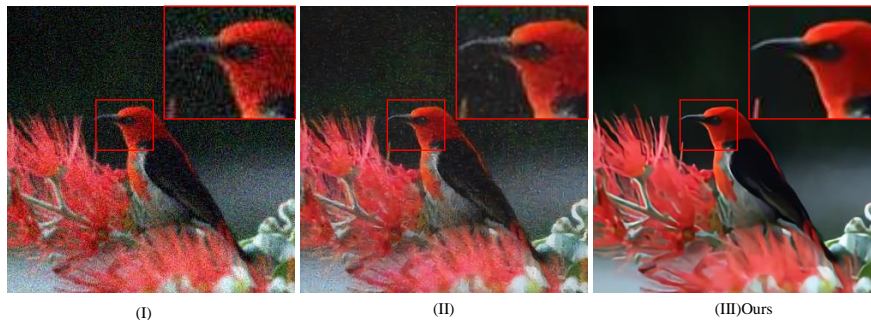


Fig. 1. Examples of hybrid-distorted image restoration. (I) Hybrid-distorted image including noise, blur, and jpeg artifacts. (II) Processed with the cascading single distortion restoration networks including dejpeg, denoise and deblurring. (III) Processed with our FDR-Net.

resolution [22, 7, 21], deblurring [30, 24, 38], denoising [4, 43, 44], deraining [10, 9, 34], dehazing [45, 1, 41] and so on. However, these works are usually designed for solving one specific distortion, which makes them difficult to be applied to real world applications as shown in Fig. 1.

Real world images are typically affected by multiple distortions simultaneously. In addition, different distortions might be interfered with each other, which makes it difficult to restore images. More details about interference between different distortions can be seen in the supplementary material. Recently, there have been proposed some pioneering works for hybrid distortion. For example, Yu *et al.* [37] pre-train several light-weight CNNs as tools for different distortions. Then they utilize the Reinforcement Learning (RL) agent to learn to choose the best tool-chain for unknown hybrid distortions. Then Sukanuma *et al.* [28] propose to use the attention mechanism to implement the adaptive selection of different operations for hybrid distortions. However, these methods are designed regardless of the interference between hybrid distortions.

Previous literature [2] suggests that deep feature representations could be employed to efficiently characterize various image distortions. In other words, the feature representation extracted from hybrid distortions images might be disentangled to characterize different distortions respectively. Based on the above, we propose to implement the feature-level divide-and-conquer of the hybrid distortions by learning disentangled feature representations. On a separate note, Schwartz *et al.* [26] point out that a series of filters and gain-control-based normalization could achieve the decomposition of different filter responses. And the convolution layer of CNN is also composed of a series of basic filters/kernels. Inspired by this, we expand this theory and design a feature disentanglement module (FDM) to implement the channel-wise feature decorrelation. By such feature decorrelation, the feature representations of different distortions could be distributed across different channels respectively as shown in Fig. 2.

Different from high-level tasks such as classification [23, 8], person re-identification [19, 3, 11], the processed feature representations are not clustered into several classes. It is crucial for low-level image restoration to aggregate the useful content information from processed feature representations to reconstruct the image. Therefore, we design an adaptive feature aggregation module (FAM) based on channel-wise attention mechanism and inverse transform of gain-control-based normalization [26].

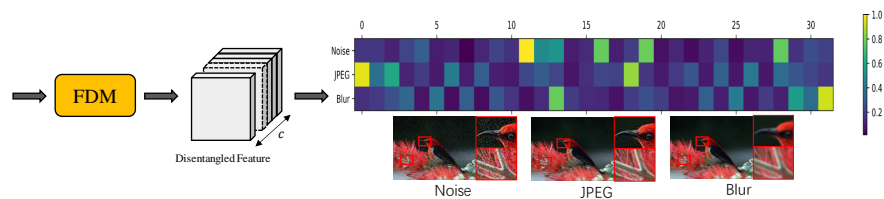


Fig. 2. Illustration of feature disentangle module (FDM) works. The FDM disentangles the input feature representation which characterizes the hybrid-distorted image into disentangled features. The responses of different distortions are distributed across different channels as shown in brighter region. The visualizations of distortions are also displayed.

Extensive experiments are conducted on hybrid distortion datasets and our model achieves the state-of-the-art results. Furthermore, we give a reasonable interpretation for our framework through visualizing the correlation matrix of features and channel response of different distortions. Finally, we verify the robustness of our network on single distortion tasks such as image deblurring and image deraining.

Our main contributions can be summarized as follows:

- We implement the feature-level divide-and-conquer of hybrid-distortion by feature disentanglement. And we design the Feature disentangled module (FDM) by revising gain-control-based normalization to distribute feature representations of different distortions into different channels.
- We propose a Feature Disentanglement Network by incorporating FDM and FAM for hybrid-distorted image restoration named FDR-Net.
- Extensive experiments demonstrate that our FDR-Net has achieved the state-of-the-art results on hybrid-distorted image restoration. We also verify the efficiency of our modules through ablation study and visualization.

2 Related Work

In this section, we briefly summarize the researches on single-distorted image restoration and hybrid-distorted image restoration.

2.1 Image Restoration on Single Distortion

Deep learning has promoted the development of computer vision tasks, especially on image restoration. As the pioneers in image restoration, some early works on special distortion removal have been proposed. Dong *et al.* first employ a deep convolutional neural network (CNN) in image super-resolution in [5]. Zhang *et al.* propose Dncnn [43] to handle noise and JPEG compression distortion. And Sun *et al.* utilize CNN to estimate the motion blur kernels for image deblur in [29]. With the advance of CNN and the improvement of computing resources, the depth and the number of parameters of the neural network have been increased significantly which facilitates a series of excellent works on special tasks such as deblurring [30, 24, 38], super-resolution [22, 7, 21], deraining [10, 9, 34] and so on to improve the image quality.

To further improve the subjective quality of degraded images, the generative adversarial network has been utilized to restore image restoration. SRGAN [16], DeblurGAN [14], ID-CGAN [42] have employed GAN in the tasks of super-resolution, deblurring, and deraining respectively to generate more realistic images. Yuan *et al.* [39] use Cycle-in-Cycle Generative Adversarial Networks to implement unsupervised image super-resolution. Ulyanov *et al.* [31] prove that the structure of generator network is sufficient to capture low-level image statistics prior and then utilize it to implement unsupervised image restoration.

2.2 Image Restoration on Hybrid Distortion

Recently, with the wide applications of image restoration, the special image restoration on single distortion cannot meet the need of real world application. Then, some works on hybrid distortion restoration have been proposed. Among them, RL-Restore [37] trains a policy to select the appropriate tools for single distortion from the pre-trained toolbox to restore the hybrid-distorted image. Then Suganuma *et al.* [28] propose a simple framework which can select the proper operation with attention mechanism. However, these works don't consider the interference between different distortions. In this paper, we propose the FDR-Net for hybrid distortions image restoration by reducing the interference with feature disentanglement, which achieves more superior performance by disentangling the feature representation for hybrid-distorted image restoration.

3 Approach

In this section, we start with introducing the gain control based signal decomposition as primary knowledge for feature disentanglement module (FDM), and then elaborate the components of our FDR-Net and its overall architecture.

3.1 Primary Knowledge

To derive the models of sensory processing, gain-control-based normalization has been proposed to implement the nonlinear decomposition of natural signals

[26]. Given a set of signals \mathbf{X} , the signal decomposition first uses a set of filters $\langle f_1, f_2, \dots, f_n \rangle$ to extract different representations of signals $\langle L_1, L_2, \dots, L_n \rangle$ as

$$L_i = f_i(\mathbf{X}), i = 1, 2, \dots, n, \quad (1)$$

and then uses gain-control-based normalization as Eq. 2 to further eliminate the dependency between $\langle L_1, L_2, \dots, L_n \rangle$.

$$R_i = \frac{L_i^2}{\sum_j w_{ji} L_j^2 + \sigma^2}, \quad (2)$$

where independent response R_i can be generated based on the suitable weights w_{ji} and offsets σ^2 with corresponding inputs L_i . In this way, the input signals \mathbf{X} can be decomposed into R_1, R_2, \dots, R_n according to their statistical property.

3.2 Feature Disentanglement

Previous literature [2] has proved that different distortions have different deep feature representation and could be disentangled at feature representation. Based on this analysis, we design the FDM by expanding the signal decomposition into channel-wise feature disentanglement to reduce the interference between hybrid distortions. As shown in [26], the combination of diverse linear filters and divisive normalization has the capability to decompose the responses of filters. And the feature representation of CNN is also composed of responses from a series of basic filters/kernels. Based on such observation, we implement the adaptive channel-wise feature decomposition by applying such algorithm in learning based framework. Specifically, we use the convolution layer ($C_{in} \times$

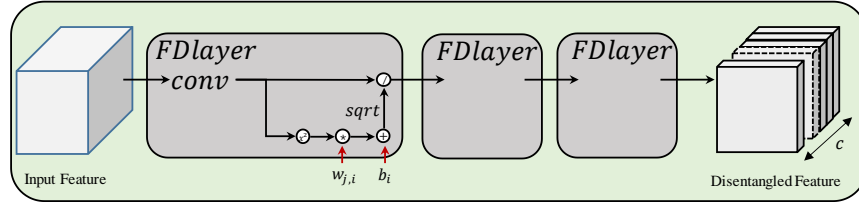


Fig. 3. Feature Disentanglement Module.

$C_{out} \times k \times k$) in neural network to replace traditional filters $\langle f_1, f_2, \dots, f_n \rangle$ in signal decomposition as section 3.1. Here, the number of input channel C_{in} represents the channel dimension of input feature \mathbf{F}_{in} , the number of output channel C_{out} represents the channel dimension of output feature \mathbf{F}_{out} , and k is the kernel size

of filters. In this way, the extraction results $S_1, S_2, \dots, S_{C_{out}}$ of convolution layer will be distributed in different channels as:

$$\langle S_1, S_2, \dots, S_{C_{out}} \rangle = conv(\mathbf{F}_{in}), \quad (3)$$

where S_i represents the i th channel in output feature and $conv$ represents the convolution layer.

To introduce the gain-control-based normalization as Eq. 2 into CNN, we modified the Eq. 2 as Eq. 4.

$$D_i = \frac{S_i}{\sqrt{\sum_j w_{ji} S_j^2 + b_i}}, \quad (4)$$

where w_{ji} and b_i can be learned by gradient descent. S_i and D_i represent the i th channel components of features before and after gain control. In formula 4, we make two major improvements to make it applicable to our task. One improvement is that the denominator and numerator is the square root of the original one in Eq. 2 which makes it is easy to implement the gradient propagation. Another improvement is to replace the response of filters L_i with channel components of features S_i , which is proper for channel-wise feature disentanglement.

In order to guide the study of parameters from convolution layer, w_{ji} and b_i , we introduce the spectral value difference orthogonality regularization (SVDO) from [3] as a loss constraint. As a novel method to reduce feature correlations, SVDO can be expressed as Eq. 5.

$$Loss_F = \|\lambda_1(FF^T) - \lambda_2(FF^T)\|_2^2 \quad (5)$$

where $\lambda_1(FF^T)$ and $\lambda_2(FF^T)$ denote the largest and smallest eigenvalues of FF^T , respectively. F is feature maps and T expresses the transposition.

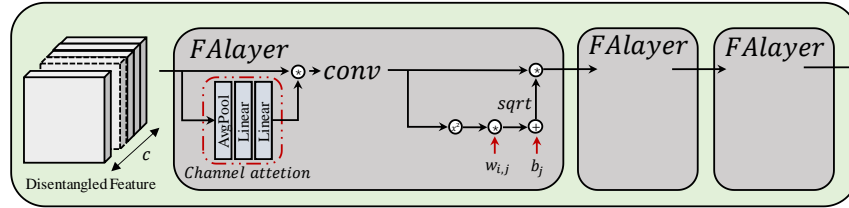


Fig. 4. Feature Aggregation Module.

According to above analysis, we design the corresponding feature disentanglement module (FDM). As seen in the Fig 3, Each FDM have multiple FD-layers and each FD-layer is composed of one convolution layer and a revised

gain-control-based normalization. Experiments demonstrate that three FDlayers for each FDM module are saturated to achieve the best performance as shown in Table. 6.

3.3 Feature aggregation Module

For hybrid-distorted image restoration, we need to restore the image instead of only getting processed feature representation. To further filter out the distortion representation and maintain the raw image content details, we utilize channel-wise attention mechanism to adaptively select useful feature representation from processed disentangled feature as Eq. 6.

$$F_{p.i} = PM(D_i), F_{p.i} = CA(F_{p.i}), \quad (6)$$

where PM represents the process module and CA presents channel attention. D_i is the i th channel of disentangled feature. $F_{p.i}$ represents the i th channel of output feature. To construct the image, we get the inversion formula from Eq. 4 as:

$$F_{c.i} = F_{p.i} \sqrt{\sum_j w_{ji} F_{p.j}^2 + b_i}, \quad (7)$$

where $F_{c.i}$ represents the output feature corresponding to the distribution of clean image. With this module, the processed image information could be aggregated to original feature space, which is proper for reconstructing restored image. Feature aggregation module (FAM) is designed as shown in Fig. 4

3.4 Auxiliary Module

In the processes of feature disentanglement, the mutual information of different channels of features is reduced, which will result in some loss of image information. In order to make up for the weakness of the feature disentanglement branch, we used the existing ResBlock to enhance the transmission of image information in parallel.

3.5 Overview of Whole Framework

To make full use of feature information of different levels, we use multi-phases structure to deal with hybrid distortion as Fig. 5. The whole structure is composed of several same phases. Every phase is composed by two branches, where the lower branch is used for distortion-aware feature disentangle and feature processing, the upper branch is used as auxiliary branch which is used to make up for information lost and enhance the removal of hybrid distortion. The details of lower branch are shown in Fig. 5. In order to ensure the matching between FDlayer and FAlayer, we use skip connection to connect the output of FDlayer and the input of FAlayer. To fuse the information of outputs from two branches, we first concatenate the two outputs and then use channel attention and convolution to fuse the outputs to a feature. Furthermore, we use dual residual to relate the different phases to enhance the mutual utilization of information between different phases.

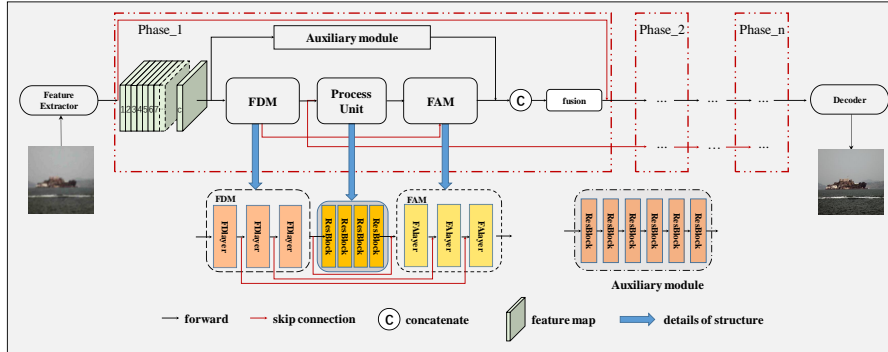


Fig. 5. Architecture of our proposed FDR-Net. It consists of multiple phases and each phase consists of feature disentangle module (FDM) and feature aggregation module (FAM).

3.6 Loss Function.

In this paper, the loss function is composed of L1 loss and feature orthogonality loss with the weight β 0.00001 as Eq. 5. The loss function can be expressed as:

$$Loss = L1loss + \beta loss_F \quad (8)$$

4 Experiments

In this section, we first conduct extensive experiments on DIV2K dataset of hybrid distortions [37] and results show that our FDR-Net achieves superior performance over the state-of-the-art methods. Moreover, in order to increase the difficulty of hybrid-distorted image restoration, we construct a more complex hybrid distortion dataset DID-HY based on DID-MDN dataset [40] and test our approach on it. Then we validate the robustness of our FDR-Net by conducting extensive experiments on single distortion. To make our network more explanatory, the correlation maps of features are visualized. Finally, we conduct a series of ablation study to verify the effectiveness of our modules in FDR-Net.

4.1 Dataset for hybrid-distorted image restoration

DIV2K dataset. The standard dataset on hybrid distortion are produced in [37] based on DIV2K dataset. The first 750 images are used for training and remaining 50 images are used for testing. They first process the dataset by adding multiple types of distortions, including Gaussian noise, Gaussian blur, and JPEG compression artifacts. The standard deviations of Gaussian blur and Gaussian noise are randomly chosen from the range of [0, 10] and [0, 50]. The quality

Table 1. Quantitative results on the DIV2K dataset comparison to the state-of-the-art methods in terms of PSNR and SSIM. OP-Attention represents the operation-wise attention network in [28]. The SSIM is computed in the same way with OP-Attention.

Test set	Mild(unseen)		Moderate		Severe(unseen)	
Metric	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DnCNN [43]	27.51	0.7315	26.50	0.6650	25.26	0.5974
RL-Restore[37]	28.04	0.7313	26.45	0.6557	25.20	0.5915
OP-Attention[28]	28.33	0.7455	27.07	0.6787	25.88	0.6167
Ours	28.66	0.7568	27.24	0.6846	26.05	0.6218

of JPEG compression is chosen from [10, 100]. According to the degradation level, the dataset is divided into three parts, respectively are mild, moderate, and severe which represent that the degradation level from low to high. Then they crop the patches of size 63×63 from these processed images as training set and testing set. The training set consists of 249344 patches of moderate class and testing sets consists of 3584 patches of three classes.

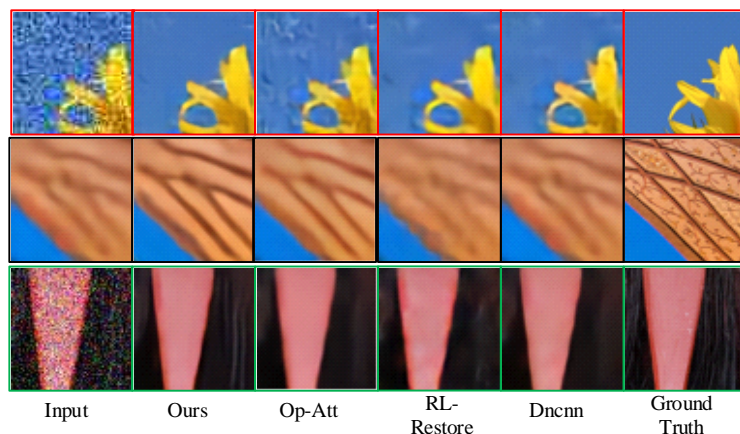


Fig. 6. The performance comparison of our algorithm with the stat-of-the-art methods performed on DIV2K dataset. The three groups are corresponding to severe, moderate and mild from top to bottom.

DID-HY dataset. Since the resolution of test image in DIV2K is too low, which is not practical in real world, we design a more complicated dataset which contains four kinds of distortions and higher resolution. The new dataset is built by adding Gaussian blur, Gaussian noise and JPEG compression artifacts based

Table 2. Quantization results on DID-HY dataset compared with the state-of-the-art methods. The SSIM is computed in the same way with OP-Attention.

Method	Dncnn[43]	OP-Attention [28]	ours
PSNR	22.1315	23.1768	24.9700
SSIM	0.5953	0.6269	0.6799

on DID-MDN dataset [40]. The training set contains 12,000 distortion/clean image pairs, which have resolutions of 512×512 . And the testing set contains 1,200 distorted/clean image pairs.

4.2 Datasets for single distortion restoration

Gopro dataset for Deblurring. As the standard dataset for image deblurring, GOPRO dataset is produced by [14], which contains 3214 blurry/clean image pairs. The training set and testing set consists of 2103 pairs and 1111 pairs respectively.

DID-MDN Dataset for Deraining. The DID-MDN dataset is produced by [40]. Its training dataset consists of 12000 rain/clean image pairs which are classified to three level (light, medium, and heavy) based on its rain density. There are 4000 image pairs per level in the dataset. The testing dataset consists of 1200 images, which contains rain streaks with different orientations and scales.

4.3 Implementation Details

The implement of the proposed framework is based on PyTorch framework with one NVIDIA 1080Ti GPU. In the process of training, we adopt Adam optimizer [13] with a learning rate of 0.0001. The learning rate will decay by a factor valued 0.8 every 9000 iterations and the number of mini-batches is 28. It takes about 24 hours to train the network and get the best results.

Table 3. Quantitative results on Gopro dataset compared with the state-of-the-art methods for image deblur. The SSIM is computed in the same way with SRN [30].

Method	DeepDeblur[32]	DeblurGan [14]	SRN [30]	Inception_ResNet[15]	ours
PSNR	29.23	28.70	30.10	29.55	30.55
SSIM	0.916	0.927	0.932	0.934	0.938

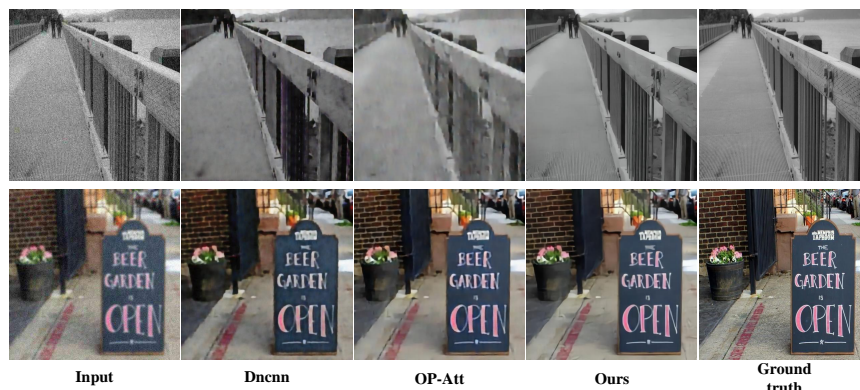


Fig. 7. Examples of hybrid distortion removal on DID-HY dataset.

4.4 Comparison with State-of-the-Arts

In this section, we compare the performance of our FDR-Net on the DIV2K dataset with state-of-the-art methods, including RL-Restore [37] and operation-wise attention network [28]. Since DnCNN [43] was used as baseline in previous work, we compared with it too. Table. 1 have demonstrated the superiority of our method. As shown in Table. 1, our algorithm outperforms previous work in mild, moderate and severe level of hybrid distortion just training with moderate level dataset. And the comparison of subjective quality is shown in the Fig. 6.

Since the size of the image in above dataset is just 63×63 and contains fewer distortion types, which is not practical in real world, we synthesize a more complex hybrid distortion dataset named DID-HY based on DID-MDN dataset. This dataset contains four kinds of distortions including blur, noise, compression artifacts and rain, in which the resolution of image is 512×512 . Then we retrained the DnCNN[43], operation-wise attention network[28] and our FDRNet. The results in Table. 2 have demonstrated that the Dncnn[43], operation-wise attention network [28] can't work well for more complicated hybrid distortion with higher resolution. However, our FDRNet can obtain the best results in objective quality as shown in Table. 2 and subjective quality as shown in Fig. 7.

4.5 Interpretative Experiment

To visualize the channel-wise feature disentanglement, we study the correlation matrix between the channel of features before and after FDM. The feature before FDM reveal high correlations(0.272 in average) as shown in Fig. 8 (a). By disentangling the feature with our FDM, the feature correlations are suppressed as Fig. 8 (b) (0.164 in average). Moreover, we visualize the channel response of different distortion types after FDM in Fig.8 (c). We take three kinds of distortions including blur, noise, and jpeg artifacts. And Each kind of distortion contains ten kinds of distortion levels. The higher level indicates the larger distortion. As

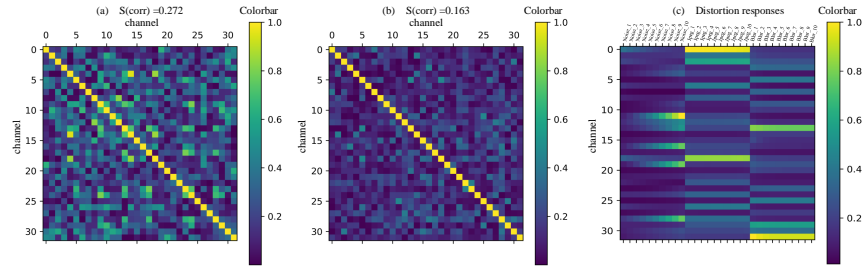


Fig. 8. (a) Visualization of correlation matrix between channels from the feature before FDM. (b) Visualization of correlation matrix between channels from feature after FDM. (c) Channels responses corresponding to different distortion types after FDM. As shown in (a)(b), FDM reduces the channel-wise correlations by disentangling the feature. From (c), the different distortions are distributed in different channels regardless of the levels of distortion by feature disentanglement.

shown in Fig.8(c), different kinds of distortions have different channel responses regardless of the distortion levels, which means that the distortions have been divided across different channels by feature disentanglement. Besides, the different distortion levels only bring the changes of response strength.

4.6 Experiments on Single Distortion

In order to verify the robustness of our method, we separately apply our FDR-Net on image deblurring and image deraining tasks and achieved the state-of-the-art performance.



Fig. 9. Visual comparisons for Gopro-test dataset between our FDR-Net and the state-of-the-art methods for image deblurring.

Comparison to image deblurring. On the task of image deblur, we compared our FDR-Net with DeepDeblur[32], Deblurgan [14], SRN [30] and Deblurganv2 [15] with Inception ResNet respectively. The table. 3 have demonstrate

Table 4. Quantitative results on DID-MDN dataset compared with the state-of-the-art methods for image derain. The SSIM is computed in the same way with SPANET [33].

Method	DID-MDN [40]	RESCAN[20]	SPANET [33]	ours
PSNR	27.95	29.95	30.05	32.96
SSIM	0.909	0.884	0.934	0.951

that our FDR-Net surpasses previous works for image deblurring. As seen in the Fig. 9, our algorithm is better than other state-of-the-art methods in generating rich texture.

Comparison to image deraining. In the field of image derain, we compared our FDR-Net with DID-MDN[40], RESCAN[20] and SPANET[33]. The quantization quality as shown in Tab. 4 have demonstrated the superiority of our method for image deraining task. The subjective quality is compared with the state-of-the-art methods in Fig. 10.



Fig. 10. Visual comparisons for DID-MDN dataset between our FDR-Net and the state-of-the-art methods for image deraining.

4.7 Ablation Studies

Effect of FDM. Distortion-aware feature disentangle method is a key factor to the success of FDR-Net for hybrid-distorted image restoration. To verify its importances and effectiveness, we substitute the FDlayer and FAlayer with a series of ResBlocks[8] as our baseline. Then we retrained the baseline network with DID-HY dataset. The parameters of two networks are almost the same. However, the performance of our FDR-Net is far higher than the baseline without FDlayer and FAlayer. The quantitative results are shown in Table. 5.

Table 5. Ablation study on different combinations of FDM and Auxiliary module. Tested on DID-HY dataset.

FDM.	Aux.	Params	PSNR	SSIM
✓	✓	1.66M	24.97	0.680
×	✓	1.62M	23.91	0.646
✓	×	1.21M	24.27	0.653

Effect of multiple FLayers and multiple channels. To investigate the effect of multiple FLayers in FDM, we set the number of FLayers as 1 to 5 respectively. As shown in Table. 6 , three FLayers are saturated to achieve the best performance. Then we set the number of FLayers as 3 and set the number of channels as 16, 32, 48 and 64 respectively. As shown in Table. 6, increasing the number of channels could increase the ability to represent the distortions for Network, which could bring the improvement of performance and 32 is saturated to achieve the best performance.

Table 6. Comparisons between different number of FLayers in each FDM and comparisons between different number of channels with 3 FLayers.

DFlayers	1	2	3	4	5	Channels	16	32	48	64
PSNR	24.63	24.70	24.97	24.98	24.99	PSNR	24.47	24.97	24.98	25.00
SSIM	0.666	0.668	0.680	0.680	0.681	SSIM	0.659	0.680	0.682	0.683

Effect of Auxiliary Module. As discussed in Section 3, auxiliary module would make up for the weakness of the distortion-aware feature disentangle branch. We demonstrate the effectiveness of auxiliary module by removing it. The performance reduction is shown as Table 5.

5 Conclusion

In this paper, we implement the channel-wise feature disentanglement by FDM to reduce the interference between hybrid distortions, which achieves the best performance for hybrid-distorted image restoration. Furthermore, we also validate the effectiveness of our modules with the visualizations of correlation maps and channel responses for different distortions. Extensive experiments demonstrate that our FDR-Net has stronger robustness for single distortion removal.

6 Acknowledgement

This work was supported in part by NSFC under Grant U1908209, 61632001 and the National Key Research and Development Program of China 2018AAA0101400.

References

1. Berman, D., Avidan, S., et al.: Non-local image dehazing. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1674–1682 (2016)
2. Bianco, S., Celona, L., Napoletano, P., Schettini, R.: Disentangling image distortions in deep feature space (2020)
3. Chen, T., Ding, S., Xie, J., Yuan, Y., Chen, W., Yang, Y., Ren, Z., Wang, Z.: Abdnnet: Attentive but diverse person re-identification. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 8351–8361 (2019)
4. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing* **16**(8), 2080–2095 (2007)
5. Dong, C., Loy, C.C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence* **38**(2), 295–307 (2015)
6. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. In: *ACM transactions on graphics (TOG)*. vol. 25, pp. 787–794. ACM (2006)
7. Haris, M., Shakhnarovich, G., Ukita, N.: Deep back-projection networks for super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1664–1673 (2018)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
9. Jin, X., Chen, Z., Lin, J., Chen, J., Zhou, W., Shan, C.: A decomposed dual-cross generative adversarial network for image rain removal. In: *BMVC*. p. 119 (2018)
10. Jin, X., Chen, Z., Lin, J., Chen, Z., Zhou, W.: Unsupervised single image deraining with self-supervised constraints. In: 2019 IEEE International Conference on Image Processing (ICIP). pp. 2761–2765. IEEE (2019)
11. Jin, X., Lan, C., Zeng, W., Chen, Z., Zhang, L.: Style normalization and restitution for generalizable person re-identification. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3143–3152 (2020)
12. Kim, J., Kwon Lee, J., Mu Lee, K.: Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1646–1654 (2016)
13. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
14. Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., Matas, J.: Deblurgan: Blind motion deblurring using conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 8183–8192 (2018)
15. Kupyn, O., Martyniuk, T., Wu, J., Wang, Z.: Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 8878–8887 (2019)
16. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4681–4690 (2017)
17. Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M., Aila, T.: Noise2noise: Learning image restoration without clean data. *arXiv preprint arXiv:1803.04189* (2018)

18. Li, G., He, X., Zhang, W., Chang, H., Dong, L., Lin, L.: Non-locally enhanced encoder-decoder network for single image de-raining. arXiv preprint arXiv:1808.01491 (2018)
19. Li, M., Zhu, X., Gong, S.: Unsupervised person re-identification by deep learning tracklet association. In: Proceedings of the European conference on computer vision (ECCV). pp. 737–753 (2018)
20. Li, X., Wu, J., Lin, Z., Liu, H., Zha, H.: Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 254–269 (2018)
21. Li, Z., Yang, J., Liu, Z., Yang, X., Jeon, G., Wu, W.: Feedback network for image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3867–3876 (2019)
22. Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K.: Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp. 136–144 (2017)
23. Liu, Z., Wu, B., Luo, W., Yang, X., Liu, W., Cheng, K.T.: Bi-real net: Enhancing the performance of 1-bit cnns with improved representational capability and advanced training algorithm. In: Proceedings of the European conference on computer vision (ECCV). pp. 722–737 (2018)
24. Lu, B., Chen, J.C., Chellappa, R.: Unsupervised domain-specific deblurring via disentangled representations. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 10225–10234 (2019)
25. Nazeri, K., Ng, E., Joseph, T., Qureshi, F., Ebrahimi, M.: Edgeconnect: Generative image inpainting with adversarial edge learning. arXiv preprint arXiv:1901.00212 (2019)
26. Schwartz, O., Simoncelli, E.P.: Natural signal statistics and sensory gain control. *Nature neuroscience* 4(8), 819 (2001)
27. Shi, Y., Wu, X., Zhu, M.: Low-light image enhancement algorithm based on retinex and generative adversarial network. arXiv preprint arXiv:1906.06027 (2019)
28. Suganuma, M., Liu, X., Okatani, T.: Attention-based adaptive selection of operations for image restoration in the presence of unknown combined distortions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 9039–9048 (2019)
29. Sun, J., Cao, W., Xu, Z., Ponce, J.: Learning a convolutional neural network for non-uniform motion blur removal. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 769–777 (2015)
30. Tao, X., Gao, H., Shen, X., Wang, J., Jia, J.: Scale-recurrent network for deep image deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 8174–8182 (2018)
31. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Deep image prior. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 9446–9454 (2018)
32. Wang, L., Li, Y., Wang, S.: Deepdeblur: Fast one-step blurry face images restoration. arXiv preprint arXiv:1711.09515 (2017)
33. Wang, T., Yang, X., Xu, K., Chen, S., Zhang, Q., Lau, R.W.: Spatial attentive single-image deraining with a high quality real rain dataset. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 12270–12279 (2019)
34. Yang, W., Tan, R.T., Feng, J., Liu, J., Guo, Z., Yan, S.: Deep joint rain detection and removal from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1357–1366 (2017)

35. Yang, X., Xu, Z., Luo, J.: Towards perceptual image dehazing by physics-based disentanglement and adversarial training. In: Thirty-second AAAI conference on artificial intelligence (2018)
36. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., Huang, T.S.: Generative image inpainting with contextual attention. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5505–5514 (2018)
37. Yu, K., Dong, C., Lin, L., Change Loy, C.: Crafting a toolchain for image restoration by deep reinforcement learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2443–2452 (2018)
38. Yuan, Q., Li, J., Zhang, L., Wu, Z., Liu, G.: Blind motion deblurring with cycle generative adversarial networks. arXiv preprint arXiv:1901.01641 (2019)
39. Yuan, Y., Liu, S., Zhang, J., Zhang, Y., Dong, C., Lin, L.: Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 701–710 (2018)
40. Zhang, H., Patel, V.M.: Density-aware single image de-raining using a multi-stream dense network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 695–704 (2018)
41. Zhang, H., Sindagi, V., Patel, V.M.: Joint transmission map estimation and dehazing using deep networks. arXiv preprint arXiv:1708.00581 (2017)
42. Zhang, H., Sindagi, V., Patel, V.M.: Image de-raining using a conditional generative adversarial network. IEEE transactions on circuits and systems for video technology (2019)
43. Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE Transactions on Image Processing **26**(7), 3142–3155 (2017)
44. Zhang, K., Zuo, W., Zhang, L.: Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. IEEE Transactions on Image Processing **27**(9), 4608–4622 (2018)
45. Zhao, J., Zhang, J., Li, Z., Hwang, J.N., Gao, Y., Fang, Z., Jiang, X., Huang, B.: Dd-cyclegan: Unpaired image dehazing via double-discriminator cycle-consistent generative adversarial network. Engineering Applications of Artificial Intelligence **82**, 263–271 (2019)