

Single Image Deraining via Multi-scale Gated Feature Enhancement Network

Hao Luo, Hanxiao Luo, Qingbo Wu $^{\boxtimes)},$ King Ngi Ngan, Hongliang Li, Fanman Meng, and Linfeng Xu

University of Electronic Science and Technology of China, Chengdu 611731, China {haoluo,lhx}@std.uestc.edu.cn, {qbwu,knngan,hlli,fmmeng,lfxu}@uestc.edu.cn

Abstract. With the widespread popularity of computer vision applications, single image deraining problem has attracted more and more attentions. Though various deep-learning based algorithms are designed for single image rain steak removal, deraining performance is still limited due to the insufficient utilization of multi-scale features, which either fails to remove rain steaks completely or damages the original image content. In our paper, a novel deraining network called Multi-scale Gated Feature Enhancement Network (MGFE-Net) is proposed to deal with different types of rain streaks meanwhile achieve a satisfied restoration effect. In MGFE-Net, a multi-scale gated module (MGM) is first utilized between the encoder and decoder to extract multi-scale features according to image content and keep the consistence between high-level semantics and low-level detail features. Furthermore, to cope with diverse rain streaks with different representative characteristics, we integrate the receptive field block (RFB) into encoder and decoder branches to enhance extracted rain features. Multi-level outputs of decoder are fused to obtain a final refined result. Extensive deraining results on synthetic and realistic rain-streak datasets present that our MGFE-Net performs better than recent deraining methods.

Keywords: Single image deraining \cdot Gated feature enhancement \cdot Multi-scale features

1 Introduction

Due to the occlusion of rain steaks with various shapes and sizes, the images captured under different rainy conditions usually tend to seriously damage the texture details and loss image contents, which hampers the further applications [3,13,19]. In this way, designing an efficient single-image deraining algorithm is highly necessary, which can remove diverse rain streaks while preserve more image details, especially in the complicated outdoor scenes. In the past few years, deraining researches have drawn considerable attentions, which mainly revolve around rain removal in video and single image [10,23,25]. Compared with video rain removal which exploits the temporary correlations between successive

© Springer Nature Singapore Pte Ltd. 2021

frames to restore clean background video, single image deraining [6,9,12] is more challenging due to the shortage of temporary information.

The widespread popularity of deep learning-based methods in other visual tasks [13,19,20], has promoted convolutional neural networks (CNN) applied into single image deraining. In [5], Fu et al. propose that it is tough to separate background from rainy image by directly using convolution network, thus they adopt CNN to cope with high frequency feature map rather than the original rainy image. Besides, joint detection [26], density estimation [29], and residual learning [11] are also introduced for rain steak detection and removal. Zhang et al. [29] propose a two-stage algorithm, which first predicts rain steak distribution and then removes them from background. Wang et al. [21] utilize an attention network to guide deraining process and generate clear background. Yang et al. [27] focus on hierarchy of local features which influences deraining effect a lot. Although these methods have achieved considerable performance improvements, existing deep learning-based deraining algorithms are still restricted in the details restoration of deraining photos. From the perspective of human visual perception, the restoration effects of some methods are not very satisfactory, which either fail to remove rain steaks completely or have an over-deraining effect resulting in distortion of original image content. For example, some methods tend to blur the background or remove some image details while removing rain steaks, because of the different levels of overlaps between background texture details and rain streaks. Besides, most deep-learning based methods are trained on synthetic data sets that results in limited generalization capability to cope with real-life rainy images well.

To cope with the restrictions of prior frameworks, we propose a novel Multiscale Gated Feature Enhancement Network (MGFE-Net), which is based on a typical framework of encoder and decoder. More specifically, the receptive field block (RFB) [13] is embedded into encoder and decoder to cope with diverse rain steaks removal and clean background restoration. Furthermore, we design the multi-scale gated module (MGM) to control propagation of multi-scale features, which can not only selectively combine multi-scale features acquired from different layers of encoder and decoder, but also keep the consistence between high-level semantics and low-level detail features. At last, several coarse deraining results are obtained by subtracting the feature maps generated by decoder from the original rainy image, and the final refined restored image is obtained by a fusion of these coarse deraining results. The proposed MGFE-Net can remove diverse rain steaks while well preserve the background content details. The comparison results validate that our MGFE-Net achieves best performance among recent designed deraining methods.

In general, the following three contributions are included in our paper:

1. We propose a novel network named Multi-scale Gated Feature Enhancement Network (MGFE-Net) based on a typical framework of encoder and decoder to deal with various rain streaks accumulated from different directions with different shapes and sizes while ensure the background content details wellpreserved.

- 2. In MGFE-Net, we introduce receptive field block (RFB) into the encoder and decoder respectively to enhance multi-scale feature extraction. Besides, we design the multi-scale gated module (MGM) to selectively combine multiscale features and keep the consistence between high-level semantics and lowlevel detail features for satisfied rain-free image restoration.
- 3. The comparison results on several benchmark synthetic and realistic datasets indicate that our MGFE-Net can present an excellent deraining performance and generalize well to real-life photos, which significantly improves the deraining effect and human visual perception quality of restored images.

2 The Proposed Method

2.1 Network Architecture

In our paper, the Multi-scale Gated Feature Enhancement Network (MGFE-Net) based on a typical encoder and decoder framework [26] is designed to deal with single image deraining task. Figure 1 presents the overall architecture. First, we embed the receptive field block (RFB) into diverse layers to strengthen the receptive field of filtering kernels in encoder and enhance the extracted deep features of decoder. Then, different from normal skip-connection in U-net framework, we design a gated module to selectively concatenate the shallow features and deep features, which can benefit to keep the consistence between shallow detail content and deep semantics. At last, in order to generate a refined restored rain-free image, a fusion strategy is adopted to integrate coarse deraining results obtained from different outputs of decoder layers.



Fig. 1. Illustration of the MGFE-Net. The designed receptive filed block (RFB) and multi-scale gated module (MGM) are embedded in encoder and decoder. The final deraining result is acquired by fusing several coarse outputs of decoder.



Fig. 2. The schematic illustration of designed modules in our MGFE-Net. (a) The integrated receptive filed block (RFB) for enhancing feature extraction. (b) Our proposed multi-scale gated module (MGM) with well consistence between high-level semantics and low-level details.

2.2 Enhanced Feature Extraction with Receptive Filed Block

The shapes, sizes and extension directions of rain steaks in real life are randomly varied, which makes the single image deraining a challenging problem. The performance of typical single image methods is always restricted, due to the limited receptive filed of simple cascaded convolution filters. To handle this issue, we integrate the RFB to promote model capability of extracting enough information by leveraging multi-scale features between adjacent convolution layers. As illustrated in Fig. 1, RFB is embedded after each layers of encoder and decoder. More specifically, RFB contains multiple forward paths with different kernel sizes, as can be seen in Fig. 2(a). For the input feature map $F_I \in \mathbb{R}^{H \times W \times C}$ from previous layers in encoder or decoder, RFB adopts different filtering kernels followed by diverse dilation rates [3] to effectively extract rain streak features in complex scenes. These feature maps in multiple forward paths are finally concatenated together to obtain the output feature map $f_O \in \mathbb{R}^{H \times W \times C}$.

2.3 Multi-scale Gated Module

Except the inability to completely remove rain steaks, another common shortage of most rain removal methods is over-deraining, which leads to damage original image content and seriously affect the visual perception quality of restored images. Thus we design a gated module to control the propagation of multi-scale features, which can not only selectively combine multi-scale features acquired from different layers of encoder and decoder but also keep the consistence between image semantics in high level and texture details in low level. By adding the gated module between different layers of encoder and decoder, the model can achieve a good deraining effect meanwhile keep background content details well-preserved.

As described in Fig. 2(b), F_i and G_{i+1} denote the corresponding shallow and deep features in encoder and decoder, respectively. We first employ an upsampling layer $UP_{2\times}$ and a 1×1 convolution layer to make the spatial size of F_i same as G_{i+1} . Then the output feature maps are stacked with F_i using a concatenation operator *Conc* along the channel dimension. The cascaded feature maps can be denoted as:

$$U_{i} = \text{Conc}(F_{i}, \text{Conv}_{1 \times 1}(UP_{2 \times}(G_{i+1}))), i = H - 1, \dots, 1$$
(1)

where $U_i \in \mathbb{R}^{H \times W \times C}$ indicates the concatenation result of F_i and G_{i+1} , H indicates the number of convolution layers in total.

As shown in Fig. 2(b), the right branch includes average pooling layer and two fully connected layers, and a weight map for gated feature is generated by selecting sigmoid function after the fully connected network layer. For the left branch, a 1×1 conv-layer and relu activation function [1] are adopted to change channel number of U_i . Then the gated feature is generated by multiply outputs from two branches, which contains consistent low-level detail information and abstract semantic features. The whole process can be denoted as follow:

$$F_{g,i} = (f_{Right}(U_i) \otimes f_{Left}(U_i)) \oplus U_i, i = H - 1, \dots, 1$$

$$(2)$$

where $F_{g,i}$ denotes the gated features in the i^{th} layer, \otimes and \oplus are element-wise product and sum operation, respectively. Before being sent into deeper layer of decoder, the gated feature $F_{q,i}$ is refined by as a dense block:

$$G_{i,1} = f_{\text{Dense}}(F_{g,i}), i = H - 1, \dots, 1$$
 (3)

where $G_{i,1}$ denotes the final output feature in i^{th} layer of decoder and f_{Dense} denotes a dense block (DB) [7], which consists of three consecutive convolution layers with dense connections. The predicted derained image Y_i in the i^{th} layer is obtained by subtracting the decoder output feature maps from the original rain maps,

$$Y_i = I - G_{i,1}, i = H, \dots, 1$$
(4)

In the end, we further fuse the coarse deraining results (i.e., Y_H, \ldots, Y_1) to obtain final refined deraining image \hat{Y} , which can be denoted as follow:

$$\hat{Y} = \operatorname{Conv}_{1 \times 1} \left(\operatorname{Conc} \left(Y_H, \dots, Y_1 \right) \right) \tag{5}$$

2.4 Loss Function

In order to guarantee the satisfied deraining effect and visual perception of restored image, the proposed MGFE-Net is optimized by combining content loss, SSIM loss and gradient loss. Specifically, we first conduct content loss to effectively measure the differences between restored images and corresponding rain-free images by leveraging a L_1 loss, which is formulated as follow:

$$L_1 = \sum_{i=1}^{H} \|Y_i - Y\|_1 + \|\hat{Y} - Y\|_1$$
(6)

where H represents the amount of coarse deraining outputs of decoder, Y is the groundtruth image, Y_i and \hat{Y} denote the restored image obtained from the i^{th} layer of decoder and the final predicted deraining image, respectively.

Besides, the SSIM loss is utilized to evaluate structural similarity between restored images and rain-free images, which can ensure the preservation of content textures and is formulated as follow:

$$L_{ssim} = \sum_{i=1}^{H} (1 - \text{SSIM}(Y_i - Y)) + (1 - \text{SSIM}(\hat{Y} - Y))$$
(7)

Furthermore, inspired by the advantages of sobel operator in edge prediction during image reconstruction [2,24], we compare the derained images with its rain-free images in gradient domain to keep the same gradient distribution. Thus the gradient loss is defined as:

$$L_{grad} = \left\| \nabla_x(\hat{Y}) - \nabla_x(Y) \right\|_1 + \left\| \nabla_y(\hat{Y}) - \nabla_y(Y) \right\|_1 \tag{8}$$

Finally, the total loss function for MGFE-Net is defined as follows:

$$L_{\text{total}} = L_1 + \lambda_g L_{grad} + \lambda_s L_{ssim} \tag{9}$$

where λ_g and λ_s are coefficients to balance different loss items.

3 Experiment

3.1 Experiment Setup

Implementation Details. The MGFE-Net is applied on the deep learningbased PyTorch [17] framework. The training image samples are cropped into patches with size of 256×256 and we further horizontally flip these patches in a probability of 0.5. The Adam optimizer is utilized with a batch size of 10 while the learning rate is 2×10^{-4} at the beginning stage and then decreased to 1×10^{-5} after 50,000 training iterations. During testing, these input rainy images keep original sizes without any data augmentations.

Datasets. In our paper, we compare MGFE-Net with other recent deraining algorithms on three synthetic benchmark datasets and a real world rainy image set. For specific, Rain1200 [29] contains a total of 24,000 pairs of rainy/rain-free images, of which 12,000 pairs are in training/testing image set, respectively. Besides the pairs in Rain1200 are conducted with three levels of rainy density.

Rain1400 [5] collects 1,000 clean images and each of them is transformed into 14 different rainy images. There are 12,600/1400 sample pairs for training/testing set. Rain1000 [21], covering a wide range of realistic scenes, is the largest single image deraining dataset including 28,500/1,000 pairs for training/testing set respectively. In addition, we collect 146 realistic rainy photos from [21,26], in which rain steaks vary in content, intensity, and orientation.

Evaluation Metrics. We adopt two typical measures, PSNR [8] and SSIM [22], to compare the performance of our MGFE-Net with recent methods. For real-world set which lacks corresponding ground truth, we use another two quantitative indicators, NIQE [15] and BRISQUE [14], to evaluate the visual quality of deraining photos. Smaller values of NIQE and BRISQUE mean better restoration effect and better visual perceptual quality. The recent deraining models we compared with are Clear [4], JORDER [26], DID-MDN [29], DualCNN [16], RESCAN [11], SPANet [21], UMRL [28] with cycle spinning, and PReNet [18].

3.2 Comparison with the State-of-the-Art Methods

Comparison Results on Synthetic Datasets. Table 1 summarizes the quantitative comparison results of different single image deraining methods where our MGFE-Net outperforms previous methods on all the benchmark datasets. Note that the performance of PReNet is very close to our MGFE-Net, the possible main reason we consider is that PReNet [18] adopts frequent image cropping to expands dataset by several times. Specifically, on Rain1200, Rain1400, and Rain1000 datasets, our method promotes the PSNR values by an average of 0.28 db, 0.68 db, 2.78 db compared with the second best results of each dataset. It is a remarkable fact that our method has an excellent performance on Rain1000, which collects images in kinds of natural scenes and contains lots of real rain steaks.

We then qualitatively compare our MGFE-Net with other methods by demonstrating details of the restored deraining images. As shown in Fig. 3, our MGFE-Net is the only model to successfully handle with different rainy situations. For the first two rows in Fig. 3 where the rain steaks are very densely distributed or different significantly in shapes, we can observe that the recent three methods cannot removal rain steaks completely while our method generates a clean deraining result. For the last two rows, other methods either leave obvious artifacts in restored images or blur the original background, while our method obtain a better visual effect and keep the content details well preserved.

Comparison Results on Real-World Dataset. Considering most deraining models are trained with synthetic rainy images, it is necessary to evaluate the generalization ability of deraining methods on realistic rainy photos. As shown in Table 2, it is obvious that our MGFE-Net performs better than previous methods according to NIQE and BRISQUE for reference-free evaluation. We also present

Methods	Rain1200		Rain1400		Rain1000	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Clear [4]	22.02	0.7889	24.73	0.8448	31.14	0.9248
JORDER [26]	30.29	0.8782	29.03	0.8881	37.45	0.9680
DID-MDN [29]	27.98	0.8626	26.96	0.8642	26.44	0.9052
DualCNN [16]	27.79	0.8210	26.10	0.8403	32.57	0.9272
RESCAN [11]	32.12	0.8998	30.92	0.9095	36.37	0.9578
SPANet [21]	26.97	0.8091	28.46	0.8801	38.51	0.9750
PReNet [18]	32.30	0.9136	30.73	0.9172	38.97	0.9782
UMRL [28]	30.49	0.8917	27.82	0.8924	28.10	0.9169
MGFE-Net(Ours)	32.58	0.9140	31.60	0.9230	41.75	0.9837

Table 1. Comparison results in PSNR and SSIM between our MGFE-Net and otherrecent methods for single image deraining on three synthetic datasets.



Fig. 3. Qualitative comparison of SPANet [21], UMRL [28], PReNet [18] and our proposed MGFE-Net on three synthetic datasets.

several restored images for qualitative comparison in real-world rainy situation in Fig. 4. Whether it is in the case of heavy rain with dense rain steaks or spare rain steak distribution with complicated shapes, our MGFE-Net has a better generalization ability to remove rain steaks in the realistic scenes than other methods.

3.3 Ablation Study

To verify the effectiveness of designed modules in MGFE-Net, we conduct four different experimental settings and evaluate their performances on Rain1200 [29]. As shown in Table 3, the four experimental settings are used to present the

Method	DualCNN [16]	SPANet $[21]$	UMRL [28]	$PReNet \ [18]$	MGEF-Net
NIQE	5.425	5.105	5.047	5.641	4.477
BRISQUE	34.74	30.92	27.58	33.42	26.44

 $\label{eq:Table 2. Comparison results in NIQE and BRISQUE on the real-world photos.$



Fig. 4. Comparison results in real-world rainy situations of SPANet [21], DualCNN [16], UMRL [28], PReNet [18] and our proposed MGFE-Net. Intuitively, our MGFE-Net performs better than recent deraining methods.

effectiveness of receptive filed block (RFB), multi-scale gated module (MGM) and gradient loss (GL), respectively. Note that, the Backbone means the simple encoder and decoder framework under the only optimization of L_1 loss and L_{ssim} loss. It can be seen obviously that supervising rain-free image generation by adding gradient loss does have a great effect on performance improvements. By integrating RFB and MGM into the experimental setting M_b sequentially, the fourth setting M_d (i.e., our proposed MGFE-Net) enhances the extracted deep features in larger receptive fields and effectively utilizes the multi-scale features, which could further promote the model capability and generate rain-free images with best visual effects.

Settings	Backbone	RFB	MGM	GL	PSNR/SSIM
M_a	\checkmark				31.61/0.8963
M_b	\checkmark			\checkmark	32.08/0.9071
M_c	\checkmark	\checkmark		\checkmark	32.45/0.9127
M_d	\checkmark	\checkmark	\checkmark		32.53 / 0.9151

Table 3. Ablation study on four experimental settings of MGFE-Net. Performancesare evaluated on Rain1200 [29] dataset.

3.4 Conclusion

In our paper, a novel multi-scale gated feature enhancement network (MGFE-Net) is proposed to solve single image deraining task. In MGFE-Net, we leverage the receptive field block (RFB) to strengthen the efficient extraction of multi-scale features and use the multi-scale gated module (MGM) to selectively combine multi-scale features and keep the consistence between image semantics in high level and texture detail information in low level. By embedding the two modules into typical framework of encoder and decoder, the proposed MGFE-Net can not only generate a clean deraining image but also keep the background content well preserved. Sufficient comparison results demonstrate that our MGFE-Net not only presents an excellent performance but also generalize well to real-life photos, which significantly improves the deraining effect and enhance human visual perception quality of derained images.

Acknowledgement. This work was supported in part by the National Natural Science Foundation of China under Grant 61971095, Grant 61871078, Grant 61831005, and Grant 61871087.

References

- Agarap, A.F.: Deep learning using rectified linear units (ReLU). arXiv preprint arXiv:1803.08375 (2018)
- Barbosa, W.V., Amaral, H.G., Rocha, T.L., Nascimento, E.R.: Visual-qualitydriven learning for underwater vision enhancement. In: ICIP, pp. 3933–3937. IEEE (2018)
- Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE TPAMI 40(4), 834–848 (2017)
- Fu, X., Huang, J., Ding, X., Liao, Y., Paisley, J.: Clearing the skies: a deep network architecture for single-image rain removal. IEEE TIP 26(6), 2944–2956 (2017)
- Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., Paisley, J.: Removing rain from single images via a deep detail network. In: CVPR, pp. 3855–3863 (2017)
- Huang, D.A., Kang, L.W., Wang, Y.C.F., Lin, C.W.: Self-learning based image decomposition with applications to single image denoising. IEEE TMM 16(1), 83–93 (2013)

 Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: CVPR, pp. 4700–4708 (2017)

83

- Huynh-Thu, Q., Ghanbari, M.: Scope of validity of PSNR in image/video quality assessment. Electron. Lett. 44(13), 800–801 (2008)
- Kang, L.W., Lin, C.W., Fu, Y.H.: Automatic single-image-based rain streaks removal via image decomposition. IEEE TIP 21(4), 1742–1755 (2011)
- Kim, J.H., Sim, J.Y., Kim, C.S.: Video deraining and desnowing using temporal correlation and low-rank matrix completion. IEEE TIP 24(9), 2658–2670 (2015)
- Li, X., Wu, J., Lin, Z., Liu, H., Zha, H.: Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 262–277. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_16
- Li, Y., Tan, R.T., Guo, X., Lu, J., Brown, M.S.: Rain streak removal using layer priors. In: CVPR, pp. 2736–2744 (2016)
- Liu, S., Huang, D., Wang, Y.: Receptive field block net for accurate and fast object detection. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11215, pp. 404–419. Springer, Cham (2018). https://doi.org/10. 1007/978-3-030-01252-6_24
- Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE TIP 21(12), 4695–4708 (2012)
- Mittal, A., Soundararajan, R., Bovik, A.C.: Making a "completely blind" image quality analyzer. IEEE SPL 20(3), 209–212 (2012)
- Pan, J., et al.: Learning dual convolutional neural networks for low-level vision. In: CVPR, pp. 3070–3079 (2018)
- Paszke, A., et al.: PyTorch: an imperative style, high-performance deep learning library. In: NeurIPS, pp. 8026–8037 (2019)
- Ren, D., Zuo, W., Hu, Q., Zhu, P., Meng, D.: Progressive image deraining networks: a better and simpler baseline. In: CVPR, pp. 3937–3946 (2019)
- Saleh, F.S., Aliakbarian, M.S., Salzmann, M., Petersson, L., Alvarez, J.M.: Effective use of synthetic data for urban scene semantic segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11206, pp. 86–103. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01216-8_6
- Song, T., Cai, J., Zhang, T., Gao, C., Meng, F., Wu, Q.: Semi-supervised manifoldembedded hashing with joint feature representation and classifier learning. Pattern Recogn. 68, 99–110 (2017)
- 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: CVPR, pp. 12270– 12279 (2019)
- 22. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE TIP **13**(4), 600–612 (2004)
- Wei, W., Yi, L., Xie, Q., Zhao, Q., Meng, D., Xu, Z.: Should we encode rain streaks in video as deterministic or stochastic? In: ICCV, pp. 2516–2525 (2017)
- Wen, Q., Tan, Y., Qin, J., Liu, W., Han, G., He, S.: Single image reflection removal beyond linearity. In: CVPR, pp. 3771–3779 (2019)
- Yang, W., Liu, J., Feng, J.: Frame-consistent recurrent video deraining with duallevel flow. In: CVPR, pp. 1661–1670 (2019)
- Yang, W., Tan, R.T., Feng, J., Liu, J., Guo, Z., Yan, S.: Deep joint rain detection and removal from a single image. In: CVPR, pp. 1357–1366 (2017)
- Yang, Y., Lu, H.: Single image deraining via recurrent hierarchy enhancement network. In: ACM MM, pp. 1814–1822 (2019)

H. Luo et al.

- 28. Yasarla, R., Patel, V.M.: Uncertainty guided multi-scale residual learning-using a cycle spinning CNN for single image de-raining. In: CVPR, pp. 8405–8414 (2019)
- 29. Zhang, H., Patel, V.M.: Density-aware single image de-raining using a multi-stream dense network. In: CVPR, pp. 695–704 (2018)