Image-Denoising-with-Deep-CNNs UDNCNN

Written by LinWenjie[1 *]

[1]StudentNo:23320211154233
1020743989@qq.com

Abstract

A deepling learning approach to blind denoising of images with out complete knowledge of the noise statistics is considered. Dncnn,which is a deep convolutional neural network(CNN)consisting if several residual blocks(ResBlocks).With cascade training,UDNCNN with u-net is more accurate and more computationally efficient than the state of art denoiing networks.An edge-aware loss function is further utilized in training UDNCNN,so that the denoising results have better perceptive quality compared to converntionally effucuebt than the state if art denoising networks.An edge-aware loss function.Next,we introduce the depthwise separaable DN ResNet utilizing the proposed Deptwise Seperable ResBlock instead of standard ResBlock,which has much less computational cost.We propose the cascade evolution of UDNCNN from DNCNN by incrementally converting ResBlocks to DS-ResBlocks.While establishing Previous training. Therefore, high precision and good calculation Efficiency is achieved at the same time. And the state of art in the past is very deep The learning method focuses on denoising Gaussian or Poisson damaged images, we think denoising images are more practical The same is true for Poisson with additive Gaussian noise. The results show that DN ResNets is better than The current state-of-the-art deep learning method, as well as a popular variant of the BM3D algorithm, in the case of blind and non-blind denoising Images corrupted by Poisson, Gaussian, or Poisson-Gaussian noise.

Keywords: Denoising CNN DnRESNET U-NET UDNCNN

Introduction

Denoising is an active topic in image processing because it is a key step in many image processing. Practical applications, such as image and video capture. It aims to generate a Clean image X from a given noisy image Y, after the image is degraded Model Y = D(X). For the widely used additive Gaussian noise (AWGN) model, The ith observed pixel is $y_i = D(X_i) = x_i + n_i$ where n_i It is i.i.d Gaussian noise with zero mean and variance. AWGN has been used to simulate signal-independent thermal noise and other system defects. Degraded due to insufficient light shot noise signal dependent and has often been modeled using Poisson noise =$y_i = D(X_i) = p_i, p_i\ P(x_i)$ where $P(X_i)$ is a Possion randon variable with x.Recently, the state of art denoising accuracy is achieved by deep neural net works.which construct a mapping between the noisy image and clean image. Unfortunately, most of existing denoising networks cannot be executed inreal-time due to their large network size. In addition, it is relatively difficult to set the hyperparameters when learning a very deep network, such as the weightinitialization, the learning rate, and the weight decay rate. With inappropriateparameters, the training might fall into local minimum or not converge at all.Denoising Residual Network (DN-ResNet) which is more efficient and accurate than prior art. DN-ResNet consists of residualblocks (ResBlock) which are gradually inserted into the network stage by stage during the training.This training strategy not only allows the resulting DNResNet to converge faster, but also allows it to be more computationally efficient than prior art denoising networks. Even better perceptual quality have been observed by using the proposed edge-aware loss function instead of the conventional mean square error (MSE).In addition, we introduce the depthwise separable ResBlock (DS-ResBlock) into DNCNN to construct the depthwiseseparable UDNCNN. UDNCNN is generated by the proposedincremental evolution from DNCNN, where the ResBlocks in DN-ResNet are replaced by DS-ResBlocks stage by stage.As a result, we may obtain a 2.5 times. As a result, we may obtain a 2.5 times complexity reduction for DN-ResNet, with less than 0.1 dB PSNR loss. To ourknowledge, DN-ResNet is the first unified deep CNN trained for the problemof blind denoising of images corrupted by multiple type of noises. By cascading only 5 ResBlocks, DN-ResNet and DS-DN-ResNet achieve the state of artperformance on all three denoising problems, Gaussian, Poisson, and Poisson Gaussian, for both cases of non-blind denoising (known noise level for noisyinput) and blind denoising (unknown noise level for noisy input).

*

Related Work

Image Denoising

During the past years, numerous approaches have been exploited for modeling image priors for denoising, such as nonlocal self-similarity (NSS) [8] andsparse coding [5]. The block matching with 3D collaborative filtering

(BM3D)[4] and its variants such as iterative BM3D with variance stabilizing transforms(I+VST+BM3D) [1] and generalized Anscombe variance stabilizing transform with BM3D (GAT-BM3D) [13] are widely used. These methods generally involve a complex optimization problem in the testing stage, which makes the denoising process time-consuming. To improve the efficiency, learning-based methods are proposed to get rid of the iterative optimization procedure, such as the trainable nonlinear reaction diffusion (TNRD) [3], and Gaussian conditional random field [20] for non-blind image deblurring. Unfortunately, the accuracy of these methods is still limited due to the use of specific image prior. It is also difficult to set the handcrafted parameters during the stage-wise learning.Recently, deep neural networks have been deployed for image denoising due totheir significant improvement of the accuracy [2]. Zhang et al. [26] constructeda 20-layer feed-forward denoising convolutional neural networks with residuallearning for Gaussian denoising. Remez et al. trained 20-layer CNNs for each object category respectively and showed good performance for either Gaussiandenoising [15] or Poisson denoising [16]. Zhang et al. [27] proposed FFDNetadopting orthogonal regularization to enhance the generalization ability of Gaussian denoising. Tai et al. designed MemNet [22], where the feature map concatenations and skip connections are utilized to construct a network for image super resolution, Gaussian denoising, and JPEG deblocking. 1 ×1convolutions areadopted to integrate the long-term memorization, which shows significant accuracy improvements. Most of the existing networks are designed for single type of noise only. Due to the high computational cost, they can not be executed inreal-time. In contrast, our DN-ResNet is far more efficient. The same networkarchitecture can be utilized for Gaussian, Poisson, and Poisson-Gaussian noise,as well as other image enhancement tasks.

Deep Learing Based Compressed Image Restoration

Compressed image restoration aims to reduce the artifacts of decoded compressed images, so that the images can be stored or transmitted at low bitrates. Most of existing work design an end-to-end network including both theencoding (compression) and decoding procedure. Toderici et al. [24] presenteda set of full-resolution lossy image compression methods using recurrent neural network based encoder and decoder with entropy coding. Theis et al. [23]constructed the compression network by deep autoencoders with a sub-pixel structure. In these work, although a low bit rate can be achieved, both of theencoding and decoding procedure are replaced by deep neural networks. As aresult, it is difficult to integrate them into real system, where efficient imagecompression algorithms such as JPEG are implemented. In this paper, we consider the compressed image restoration as a 'denoising' problem, where the noisecomes from image compression algorithms. DN-ResNet is trained to refine thequality of decoded compressed image. Since our network can be considered as apost-processing step, it can be applied to any existing image compression algorithms.

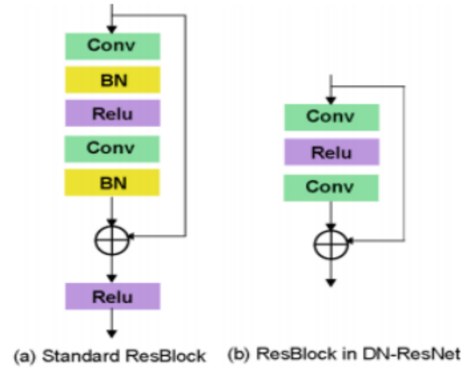

Figure 1: ResBlocks in DN-ResNet. (a) Standard ResBlock (b) ResBlock in DN-ResNet.

.

Proposed Solution

DN-CNN We aim to train a deep convolution neural network for image denoising. Thenetwork takes a noisy image Y as input and predicts a clean image X as its utput. Given a training set Xi, Yi, i = 1,...,N with N samples, our goal is to learn a model S that predicts the clean image $X^i = S(Y_i)$.ResNet [9] has demonstrated considerable performance in computer visionapplications such as image classification. The basic element of our proposeddenoising residual network (DN-ResNet) is a simplified ResBlock, as shown inFig. 1(b). Different from the standard ResBlock in Fig. 1(a), we remove the batchnormalization layers and the ReLU layer after the addition, because removing hese layers will not harm the performance of feature-map based ResNet [12].

3 layers, and proceeds to 5 layers, 7 layers, etc. Each convolutional layer in theResBlock consists of 32 3 × 3 filters. It ensures a smaller network when goingdeeper. The new layers are inserted just before the last 5×5 layer. The weights ofpre-existing layers are inherited from the previous stage, and the weights of thenew ResBlocks are randomly initialized (Gaussian with 0.001). Hence, onlya few weights of DN-ResNet are randomly initialized at each stage, so the convergence will be relatively easy. We find that using a fixed learning rate 0.0001 or all layers without any decay is feasible. Since new convolutional layers will reduce the size of the feature map, we zero pad 2 pixels in each new 3×3 layer. As a result, all the stages in cascade training have the same size as the output, so that the training samples could be shared.When cascading 5 ResBlocks, the resulting DN-ResNet will have 5 × 2 + 3 = 13convolutional layers. Our experiments show that such DN-ResNet-13 has already achieved the state of art accuracy on all type of noises.

UDNCNN In this section, we propose depthwise separable DN-ResNet (DS-DN-ResNet)to further reduce the network size of DN-ResNet, as well as the computationalcost. In the classification network MobileNet [10,19], the standard convolutionallayer is factorized into a depthwise convolution and a 1 × 1 pointwise convolution, which

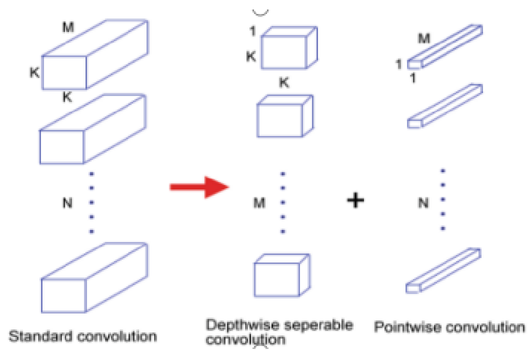

Figure 2: Depthwise separable convolution. The standard convolution (left) is replaced by depthwise convolution (middle) and pointwise convolution (right).

.

achieves significant efficiency gain. As shown in Fig. 2, the standard convolution with M input channels and N K × K filters is replaced by a depthwise convolutional layer with M K × K filters, and a pointwise convolutionallayer with N 1×1 convolutional filters and M input channels. Assume the inputfeature map size is W × H, the number of the multiplications are reduced from M × K × K × N × W × H to M × K × K × W × H + M × N × W × H.

Edge-Aware Loss Function Most of existing denoising networks aim to minimize the Mean Square Error(MSE)$\frac{1}{N}\sum_{i=1}^{n}\|X - X_i\|^2$over the training set. In this paper, we propose an edge-aware loss function, where the pixels in the edges are granted higher weights compared to non-edge pixels and loss:
$\frac{1}{N}\sum_{i=1}^{n}\|X - X_i\|^2 + w * \frac{1}{N}\sum_{i=1}^{n}\left\|X_i M_i - \hat{X}_i M_i\right\|^2$Xi is the ground truth of ith clean image,$hatX_i$ is the ith denoised image, M is an edge map, N is the number of images, and w is a constant to control the trade-off between edge and non-edge pixels.
There are two advantages of applying such edge-aware loss function. Firstly, one of the major challenge in image denoising is that the edges are difficult to be retrieved from a noisy image. Adding a corresponding constraint in the loss function is reasonable. Secondly, the highfrequency information such as edge is very sensitive in human vision. Increasing the denoising accuracy of edge pixels will contribute to the perceptual quality.

Experiments

Experiment Setting For image denoising, we use the BSDS300 dataset [6]. We follow the same training and testing split as [15], 1,000 testing images are used to evaluate the performance of the proposed DN-ResNet, while the remaining images are used for training. Random Gaussian/Poisson/Poisson-Gaussian noisy images are generated with different noise levels.
3 layers, and proceeds to 5 layers, 7 layers, etc. Each convolutional layer in theResBlock consists of 32 3 × 3 filters. It ensures a smaller network when goingdeeper. The new layers are inserted just before the last 5×5 layer. The weights ofpre-existing layers are inherited from the previous stage, and the weights of thenew ResBlocks are randomly initialized (Gaussian with 0.001). Hence, onlya few weights of DN-ResNet are randomly initialized at each stage, so the convergence will be relatively easy. We find that using a fixed learning rate 0.0001 or all layers without any decay is feasible. Since new convolutional layers will reduce the size of the feature map, we zero pad 2 pixels in each new 3×3 layer. As a result, all the stages in cascade training have the same size as the output, so that the training samples could be shared.When cascading 5 ResBlocks, the resulting DN-ResNet will have 5 × 2 + 3 = 13convolutional layers. Our experiments show that such DN-ResNet-13 has already achieved the state of art accuracy on all type of noises.
DnCNN with all different D have residual with non-zero values now. This is because He's initialization activates half of the neurons which avoid vanishing gradient in the ReLU. Now, since the output is different from the input, the gradient and weight update are different, enabling the network to learn properly

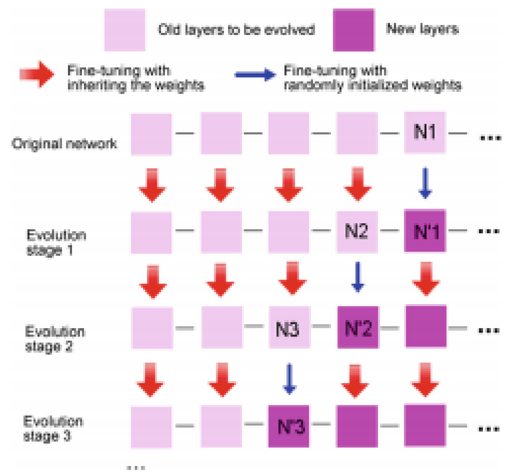

Figure 3: Incremental evolution from DN-CNN to DDCNN.

.

Experiments on Image Denoising We first test the DN-ResNets up to 13 layers on non-blind Gaussian, Poisson, and Poisson-Gaussian denoising. These DN-ResNets are trained by cascading the ResBlocks in Fig. 1(b). The conventional MSE loss is utilized for all networks. In Table 1, we find that for all the above three denoising problems, the PSNR consistently increases along with using more layers. Although the deepest network we show is 13-layer DN-ResNet, the accuracy could still be further improved by cascading more layers. This is consistent with 'the deeper, the better'. We also compare the cascade training versus one-shot training ('13-layer-os' in Table 1), where an end-to-end 13-layer DN-ResNet is trained from unsupervised weight

Next, we test the DN-CNN trained by edge-aware loss function described in Sect. 3.3, as well as utilizing

model	DnCnn	UdnCnn	DudnCnn
PSNR	28.9662	28.3076	29.1349

Table 1: Performance comparison among DnCnn.UdnCnn.DudnCnn

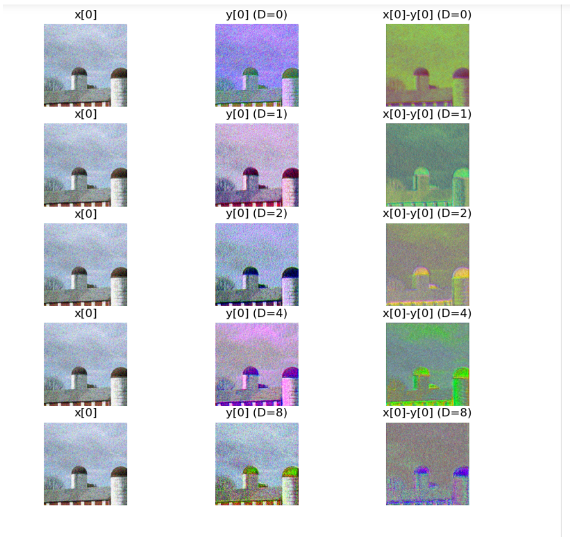

Figure 4: DNCNN with different d

.

DN-ResNet for blind denoising. In Table 1, we observe that utilizing DN-ResNet for blind denoising will not decrease the accuracy much compared to non-blind denoising. This trade-off is valuable since blind denoising does not require a time-consuming noise level estimation. In addition, we show that utilizing edge-aware loss function (blind+'edge-a'/'edgeb') improves the SSIM 0.005–0.01, without degrading the PSNR much. Since the conventional MSE has the same equation as PSNR, the slightly degradation in PSNR of the edge-aware DN-ResNet is reasonable. Using the edge map generated by Sobel gradient magnitude (blind+'edge-a', w = 0.025 in Eq.
Moreover, we evaluate the DN-ResNets constructed by different ResBlocks for the blind denoising problem. In Table 3, We observe that DS-DN-ResNet (DS-DN) has less than 0.1 dB PSNR degradation and less than 0.002 SSIM degradation compared to DN-CNN, but the computational cost (MACs, number of multiplications and accumulations) and the network size are significantly reduced. We also notice that if the UDCNN is constructed by one-shot fine-tuning DN-CDD), both the PSNR and SSIM will decrease a lot. This indicates that the proposed incrementally evolved DS-DN-ResNet is able to improve the efficiency of DN-Cnn.

DnCNN with all different D have residual with non-zero values now. This is because He's initialization activates half of the neurons which avoid vanishing gradient in the ReLU. Now, since the output is different from the input, the gradient and weight update are different, enabling the network to learn properly.

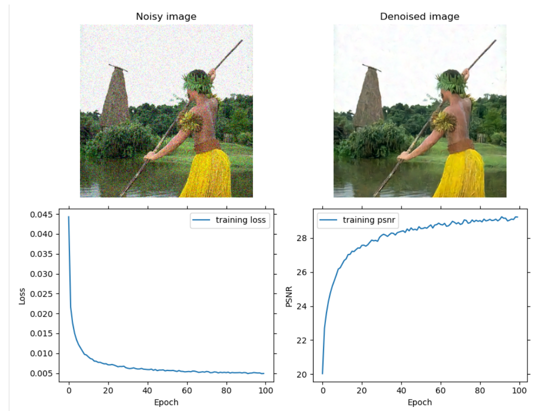

Figure 5: after denoised

.

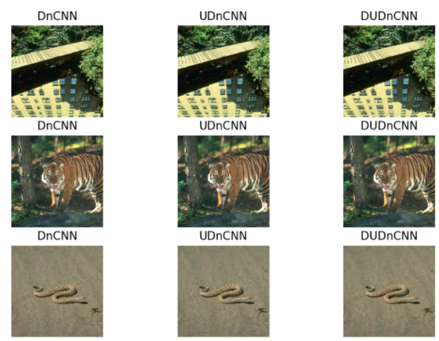

Figure 6: Different CNN

.

Comparison to the State of Art Denoising Algorithms we compare the proposed DN-CNN to the state of art denoising algorithms in PASCAL VOC dataset. For fair comparison, we retrain other networks using the same BSDM dataset. We observe that DN-CNN-13 blind denoising network clearly outperforms other blind and non-blind Gaussian denoisingalgorithms. Compared to the 20-layer DN-CNN-S [26], DenoiseNet [15], and MemNet [22] which contain more than 600K parameters, DN-ResNet achieves competitive performance, but the network size (150K parameters) is 4 times smaller. DN-ResNet takes 15–20 ms to process a 512×512 image on single Titan X GPU, compared to 50–60 ms for DN-CNN and DenoiseNet. DS-DN-ResNet only takes 8–10 ms to process a 512×512 image, with the cost of less than 0.1 dB accuracy loss. These results show the effectiveness of DN-ResNet for Gaussian denoising.

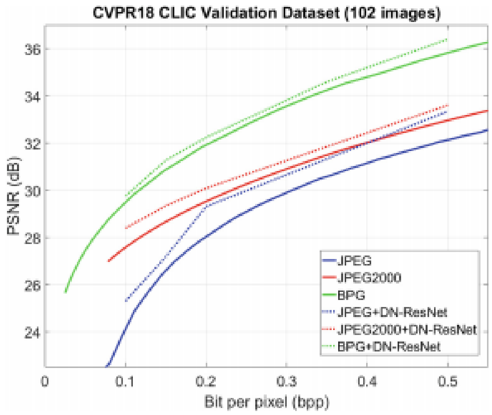

Figure 7: Bit per pixel vs. PSNR in dataset

.

PSNR A very classical (but controversial) way to compare the quality of restoration techniques is to use the PSNR (Peak Signal-to-Noise-Ratio) defined for images ranging in[-1,1] as:
$PSNR = 10log_{10}\frac{4n}{\|y-d\|^2}$d is the desired ideal imagethe number of elements y is the estimate obtained from in the tensor. The PSNR measures in decibels (dB) the quality of the restoration: the higher the better.

Applications to Other Image Enhancement Tasks We emphasize that our proposed architecture can be trained for other image enhancement tasks and provide state of art performance on these tasks, with relatively low complexity. Image Restoration: We evaluate the proposed DN-ResNet on compressed image restoration. the curves of compression ratio (bpp, bit per pixel) versus PSNR of the decoded compressed image and restored image are given. We can find that DN-ResNet is able to improve the quality of the decoded images for all compression methods. 1–2 dB, 0.5–0.7 dB, and 0.3–0.4 dB gain can be observed for JPEG, JPEG 2000, and BPG respectively. Figure 7 shows some restored images at 0.15 bit per pixel, where DN-ResNet clearly improves the perceptual quality of the decoded compressed images. Image Super Resolution: We cascade our DN-CNN to 19 layers and apply it for image super resolution [17]. The low-resolution images are considered as noisy input, and the high-resolution images are considered as clean image. Our DN-ResNet achieved state of art SR performance with much less computational complexity. For example, it achieves 0.5, 0.3, 0.2 dB PSNR gain and 0.003, 0.001, 0.001 better SSIM for the SR scales 2, 3, and 4 in Set 14, while having only 1/3rd of the network size compared to existing networks such as MemNet networks such as MemNet [22] or DRRN [21]

Conclusion

In this paper, we presented the DN-CNN for image denoising achieving both high accuracy and efficiency. We show that cascade training is effective in training efficient deep ResNets. The perceptual quality can be enhanced by using edge-aware loss function. We further propose the depthwise separable ResBlock and incrementally evolve the DN-CNN to DN-CNN withU-NET, which reduced the computational cost of DN-ResNet 2.5 times with less than 0.1 dB degradation in PSNR. For both cases of blind and non-blind denoising, our experimental results on benchmark datasets show that the proposed DN-ResNet achieves better accuracy and efficiency compared to the state of art denoising networks on all types of noises, including Gaussian, Poisson, and Poisson-Gaussian. The same network architecture can be utilized for other image enhancement applications as well.

References

1. Azzari, L., Foi, A.: Variance stabilization for Noisy+Estimate combination in iterative Poisson denoising. IEEE Signal Process. Lett. 23(8), 1086–1090 (2016)
2. Burger, H.C., Schuler, C.J., Harmeling, S.: Image denoising: can plain neural networks compete with BM3D? In: IEEE Conference on Computer Vision and Pattern Recognition (2012)
3. Chen, Y., Pock, T.: Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration. IEEE Trans. Pattern Anal. Mach. Intell. 39(6), 1256–1272 (2017)
4. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3- D transform-domain collaborative filtering. IEEE Trans. Image Process. 16(8), 2080–2095 (2007)
5. Dong, W., Zhang, L., Shi, G., Li, X.: Nonlocally centralized sparse representation for image restoration. IEEE Trans. Image Process. 22(4), 1620–1630 (2013)
6. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The PASCAL visual object classes (VOC) challenge. Int. J. Comput. Vis. 88(2), 303– 338 (2010)
7. Foi, A., Trimeche, M., Katkovnik, V., Egiazarian, K.: Practical PoissonianGaussian noise modeling and fitting for single-image raw-data. IEEE Trans. Image Process. 17(10), 1737–1754 (2008)
8. Gu, S., Zhang, L., Zuo, W., Feng, X.: Weighted nuclear norm minimization with application to image denoising. In: IEEE Conference on Computer Vision and Pattern Recognition (2014)
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (2016) 230 H. Ren et al.
10. Howard, A.G., et al.: MobileNets: efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861 (2017)
11. Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M.: Enhanced deep residual networks for single image super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (2017)
12. Makitalo, M., Foi, A.: Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise. IEEE Trans. Image Process. 22(1), 91–103 (2013)
13. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database

of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: IEEE International Conference on Computer Vision, vol. 2, pp. 416–423 (2001)
14. Remez, T., Litany, O., Giryes, R., Bronstein, A.M.: Deep class-aware image denoising. In: International Conference on Sampling Theory and Applications (2017)
15. Remez, T., Litany, O., Giryes, R., Bronstein, A.M.: Deep convolutional denoising of low-light images. arXiv:1701.01687 (2017)
16. Ren, H., El-Khamy, M., Lee, J.: Image super resolution based on fusing multiple convolution neural networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1050–1057. IEEE (2017)
17. Ren, H., El-Khamy, M., Lee, J.: CT-SRCNN: cascade trained and trimmed deep convolutional neural networks for image super resolution (2018)
18. Toderici, G., et al.: Full resolution image compression with recurrent neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 5435–5443 (2017)

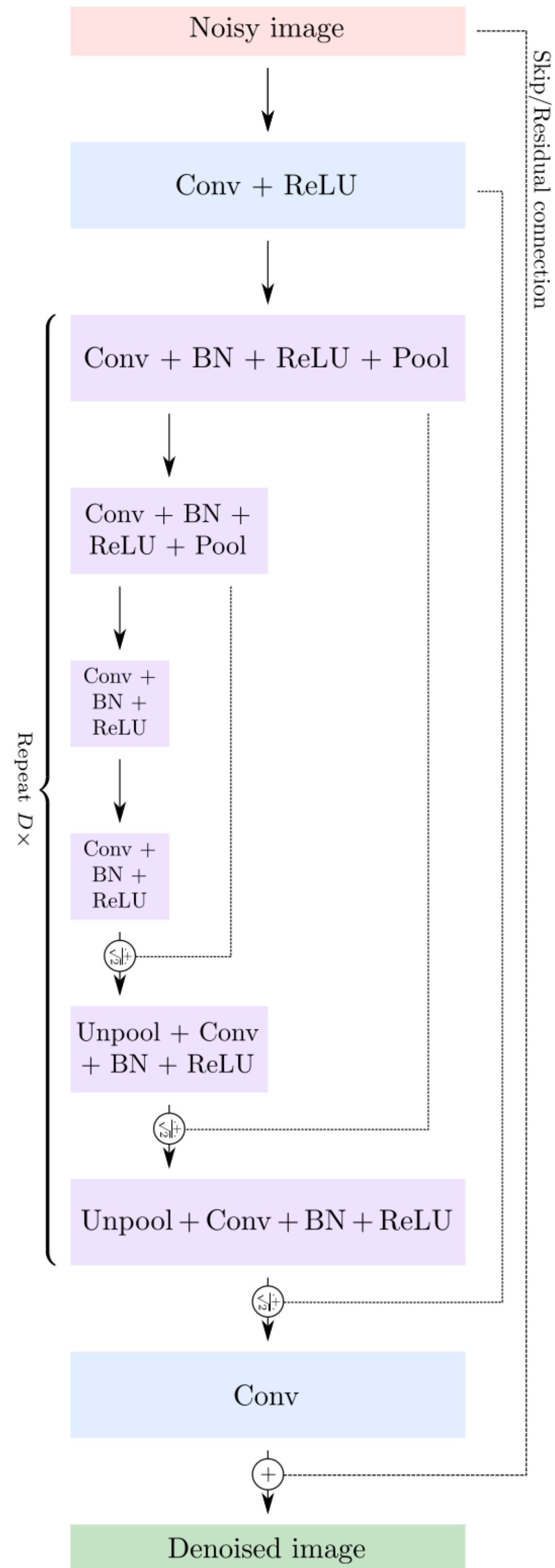

Figure 8: UdnCNN
.