DID: A Nested Dense in Dense Structure with Variable Local Dense Blocks for Super-Resolution Image Reconstruction

Longxi Li, Hesen Feng, Bing Zheng, *Lihong Ma, Jing Tian
eelhma@scut.edu.cn

Motivation

The key challenge in SR methods is how to produce the information missed when imaging, and one of the feasible approaches may rely on efficiently utilizing the hierarchical features created from intermediate layers of deep neural networks. However, methods for hierarchical feature creation from intermediate layers are still open. To tackle this challenge, we try to improve performance in two ways: (1) match the receptive field of an extractor with a suggested depth varied dense block VLDB, and (2) propose a nested DID structure which could be viewed as a super RDN, a global dense mechanism (GDM) multiplexing VLDB's controllable and hierarchical features in different scales.

Previous Work

The the Residual Dense Network (RDN)[1] proposes a fixed-depth residual dense block (RDB) by merging a residual block with dense blocks. It aggregates features from all hidden-layers by global residual learning, exceeding the memory blocks in MemNet [2]. But RDBs are cascade connected, gradient disappearance/explosion would remain when a network continuously deepens. Also, with fixed depth of dense blocks in DenseNet-like networks, rich details are hard to follow.

On the other hand, attention networks come for SR quality promotion. Typical Residual Channel Attention Network (RCAN)[3] includes the residual in residual (RIR) structure, providing global-long and local-short skips with channel attentions (CA), and focusing on detail learning as well. Though CA is introduced, the residual skip connections in RCAN are those actually important [3]. Whereas the Second-order Attention Network(SAN) calculates covariance matrix of feature maps to give an acute statistical attention for correlation utilizing, also applies skip connections to pass LR information[4].

As mentioned above, we could find that either network deepened/widened or channel attention imposed, the key challenge in SR methods is how to produce the information missed when imaging.

References

[1] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu: *Residual dense network for image super-resolution*, IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, Utah, USA, June 2018, pp. 2472-2481.

[2] Y. Tai, J. Yang, X. Liu, C. Xu: *Memnet: A persistent memory network for image restoration*, IEEE Conf. on Computer Vision (ICCV), Venice, Italy, October 2017, pp. 4549-4557.

[3] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, Y. Fu: *Image Super-Resolution Using Very Deep Residual Channel Attention Networks*, European Conf. on Computer Vision (ECCV), Munich, German, September 2018, pp. 1-16.

[4] Tao Dai, Jianrui Cai, Youngbing Zhang, Shu-Tao Xia: *Second-order Attention Network for Single Image Super-Resolution*. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, June 2019, pp. 11057-11066.

Acknowledgements

We would like to acknowledge supports from China NSFC (No.61471173) and GDNSFs (No.2019A1515011949 No.2017A030311028).

Network Structure

To match the RFs adaptively for DBs in different depths, we suggest the variable local dense blocks (VLDB) with each DB layers assigned in an arithmetic progression. The VLDB framework contains three local parts: local dense connection, local feature fusion and local residual learning.

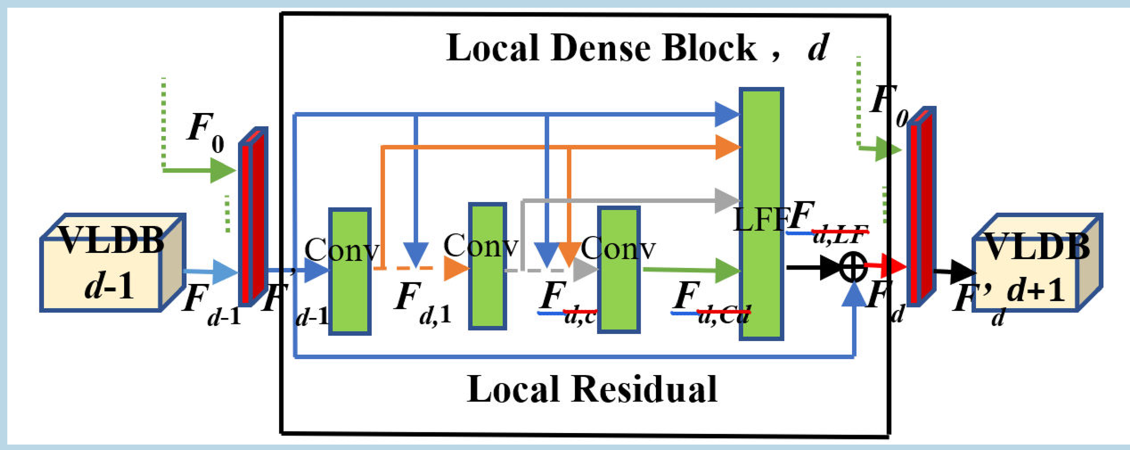

Fig.1.The architecture of variable local dense block (VLDB)

DID network could be viewed as a super RDN, using VLDBs as basic nodes.It also contains three parts: global dense connections, global feature fusion(GFF) and global residual learning(GRL).

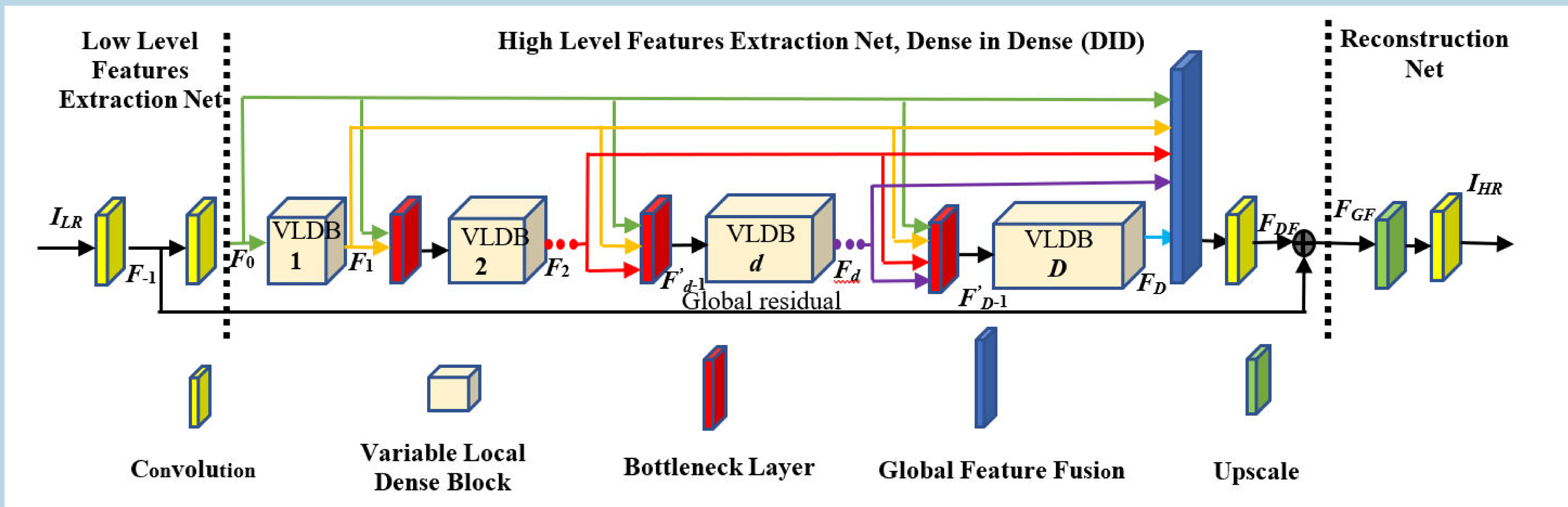

Fig.2.The architecture of our DID network.

Experimental Results

With the same loss function, and the same numbers of dense blocks and Conv. layers in total, Figure 3 displayed loss curves of DID network and RDNs in 200 epochs. DID curves are better in each paired setting. Figure 4 showed that both Model III utilizing VLDBs and Model II utilizing GDM can improve the PSNR values on the basis of Model I. Model-IV, the DID network jointly utilizing VLDBs and GDM, achieves the highest average PSNR 38.04 dB and well explains that the proposed VLDBs and the GDM are consistent.

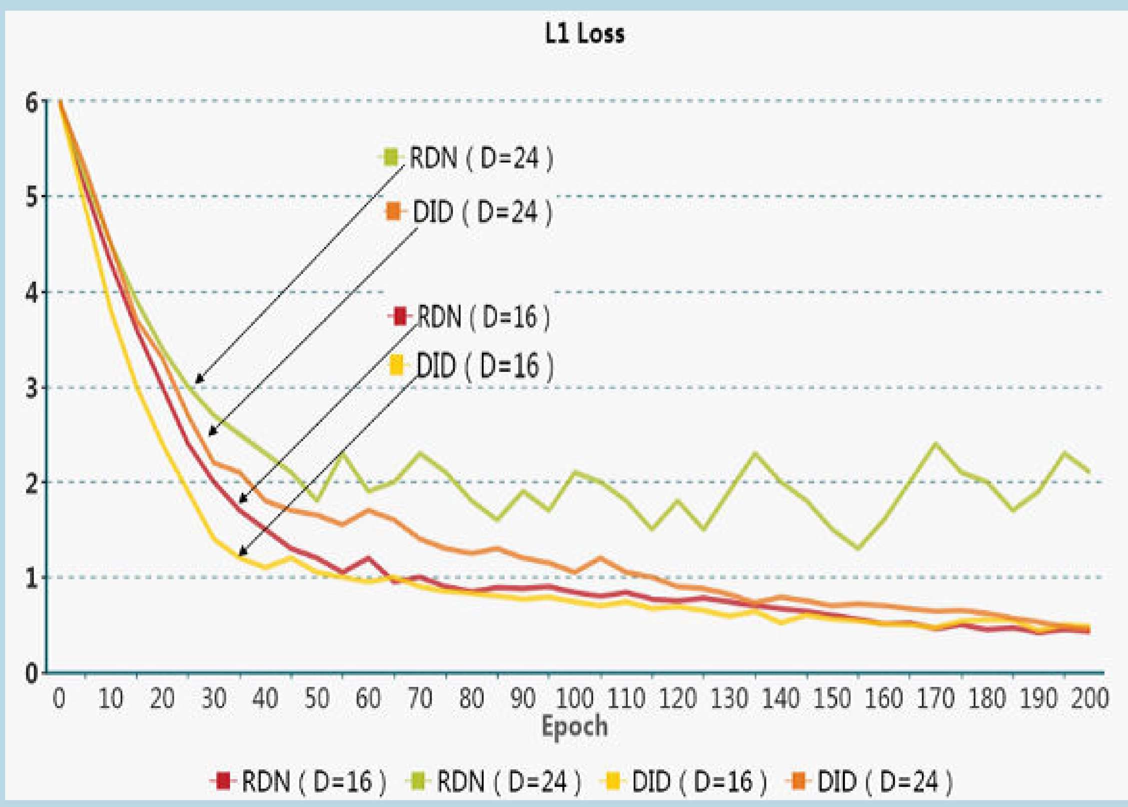

Fig.3.Loss curves of DID network and RDNs.

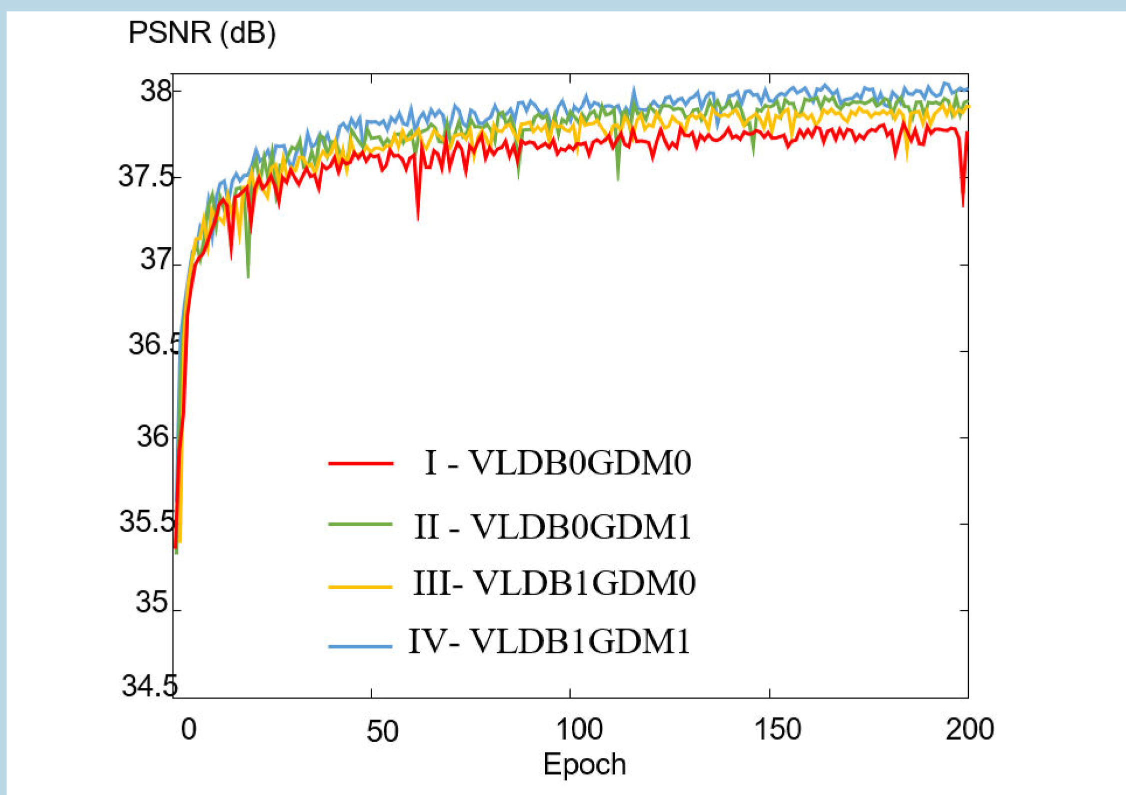

Fig.4.Performance of different combinations of modules.

The related parameters in the proposed DID network are the number of VLDBs (denoted as D) and the number of Conv. layers per VLDB (denoted as C). Figure 5 could reveal effects of D and C, and confirms the variable depths in VLDB is reasonable and feasible.

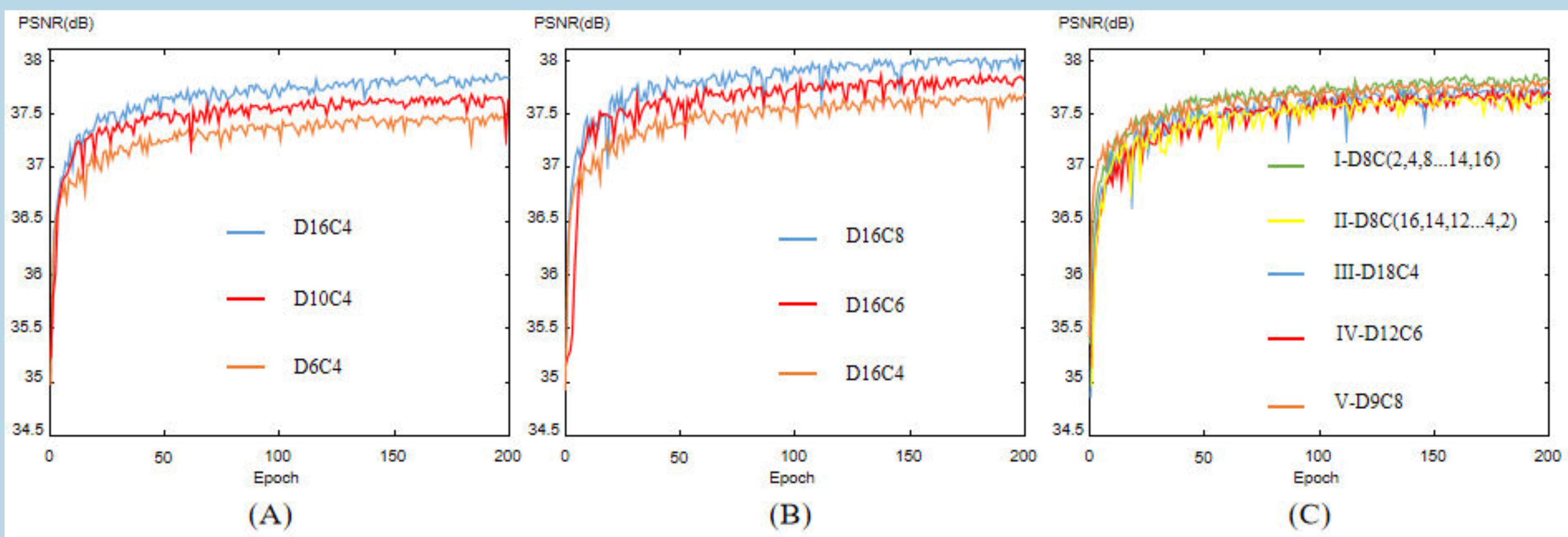

Fig.5.Convergence analysis of DID with different values of D, C.