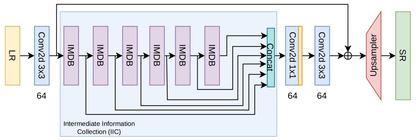
IMDeception: Grouped Information Distilling Super-Resolution Network

Mustafa Ayazoglu
(mayazoglu@aselsan.com.tr)
Aselsan Research, Ankara, Turkey

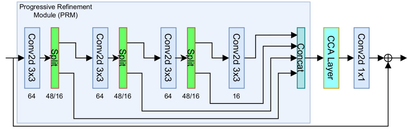
IDEA:

IMDN[1] uses:

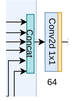
Intermediate Information Collection (IIC) for IMDB feature aggregation



Progressive Refinement With in IMDB



Problems !!!



Requires a lot of parameters

Marginal Contributions (Manga 109)
IIC (0.04dB) < PRM (0.24dB)

IMDB uses standard 2D Convolutions

Proposed Solutions !!!

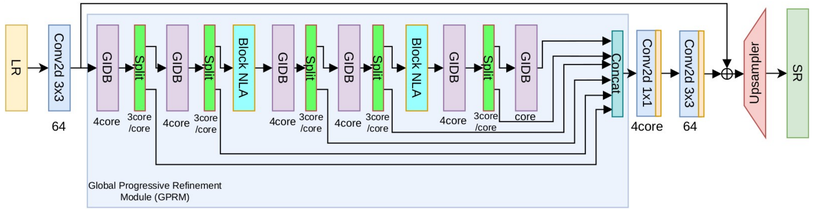
Replace IIC with PRM mechanism for feature aggregation, **Global Progresive Refinement Module, GPRM**

Network-in-Network[3] type structure

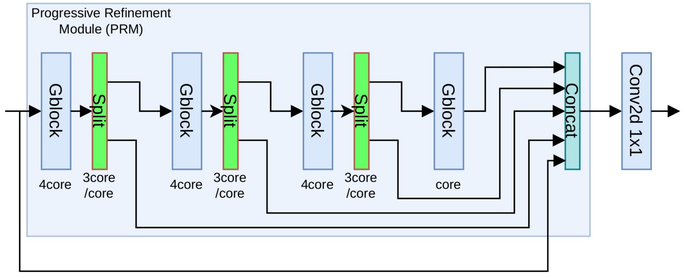
Replace IMDB's standard 2D Conv with **Modified GBlocks** of Mobile AI 2021 SISR winner XLSR [2], using grouped convolutions. **Grouped Information Distilling Blocks, GIDB**

Remove Channel Contrast Aware Attention at the local lave and add Block Non-Local Attention in the main branch

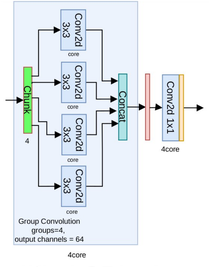
IMDeception



Global Progresive Refinement Module, GPRM



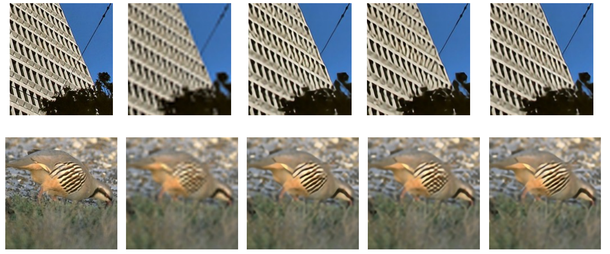
Grouped Information Distilling Block, GIDB



Modified GBlock

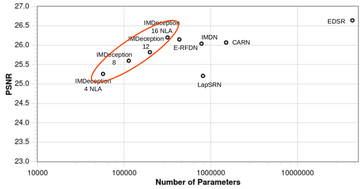
RESULTS:

Original Bicubic CARN[4] IMDN IMDeception



Parameters

IMDeception	$core = 16 + NLA$	$core = 12$	$core = 8$	$core = 4 + NLA$	$core = 4$
#Params	$316K$	$198K$	$113K$	$57K$	$57K$
FLOPS[G]	20.7	12.9	7.4	3.7	3.7
Act.[M]	206	149	103	57	57
Runtime[ms]*	60	45	31	24	21
#Conv2d	62	58	58	62	58
#Div2K Val. (PSNR)	29.02	28.82	28.70	28.48	28.45



Edge Device Comparison

Model\Hardware	RTX 2080 TensorRT (ms)	Jetson Xavier TensorRT (ms)
IMDeception core=16+NLA	17.7	88.9
IMDeception* core=12	19	98.9
IMDeception core=8	9.9	51.1
IMDeception core=4+NLA	9.8	44.7
IMDeception core=4	9.2	41.9

Performance Comparison

Model	Set5	Set14	BSD100	Urban100	Manga109	Div2K (Val)
Bicubic	28.42	26.00	25.96	23.14	24.89	26.66
LapSRN (812K)	31.54	28.19	27.32	25.21	29.09	28.75
IMDN (779K)	32.21	28.58	27.56	26.04	30.45	28.94
EDSR (43M)	32.46	28.80	27.71	26.64	31.02	29.25
CARN [2] (1.5M)	32.14	28.61	27.58	26.07	30.46	28.96
E-RFDN (433K)	32.16	28.65	27.60	26.15	30.59	29.04
IMDeception (316K) core=16 + NLA	32.14	28.64	27.60	26.20	30.67	29.02*
IMDeception(198K) core=12	31.83	28.47	27.50	25.82	30.28	28.83
IMDeception(113K) core=8	31.69	28.35	27.42	25.60	29.89	28.69
IMDeception(57K) core=4	31.35	28.12	27.27	25.23	29.20	28.45
IMDeception(57K) core=4 + NLA	31.42	28.18	27.26	25.26	29.28	28.48

Summary/Conclusion

Very efficient SR network only with **316K** parameters and **20G** FLOPS

Many high performing variants of it exist with little sacrifice of the performance

Can run on edge devices in real-time with inputs up to 512x256

References

[1] Zheng Hui, Xinbo Gao, Yunchu Yang, and Xiumei Wang. Lightweight image super-resolution with information multidistillation network. In Proceedings of the 27th ACM International Conference on Multimedia. ACM, oct 2019

[2] Mustafa Ayazoglu. Extremely lightweight quantization robust real-time single-image super resolution for mobile devices. In IEEE Conf. Comput. Vis. Pattern Recog. Workshops, 2021

[3] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network, 2014

[4] Namhyuk Ahn, Byungkon Kang, and Kyung-Ah Sohn. Fast, accurate, and lightweight super-resolution with cascading residual network. In Eur. Conf. Comput. Vis.2018