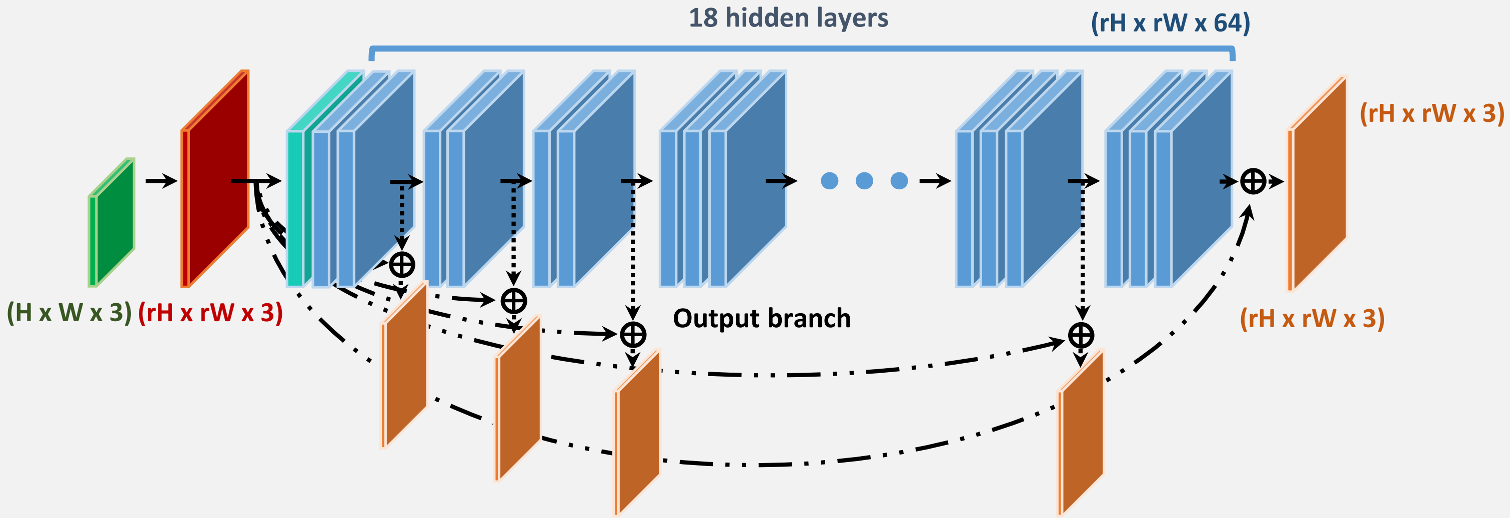
Depth Controllable Very Deep Super-Resolution Neural Network (DCVDSR)

Dohyun Kim (Presenter): Chung-Ang University, Seoul, Korea
E-mail: dohyun.cau@gmail.com
Joongheon Kim: Chung-Ang University, Seoul, Korea
Junseok Kwon: Chung-Ang University, Seoul, Korea
Taehyung Kim: KT AI Tech Center, Seoul, Korea



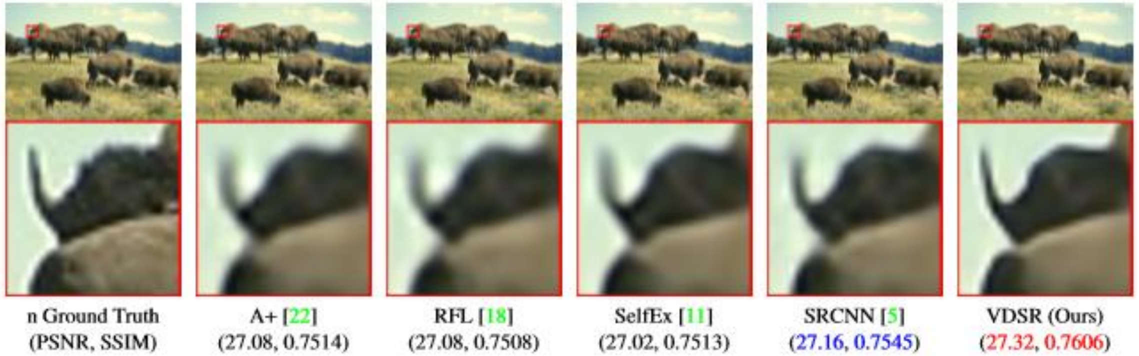
Introduction

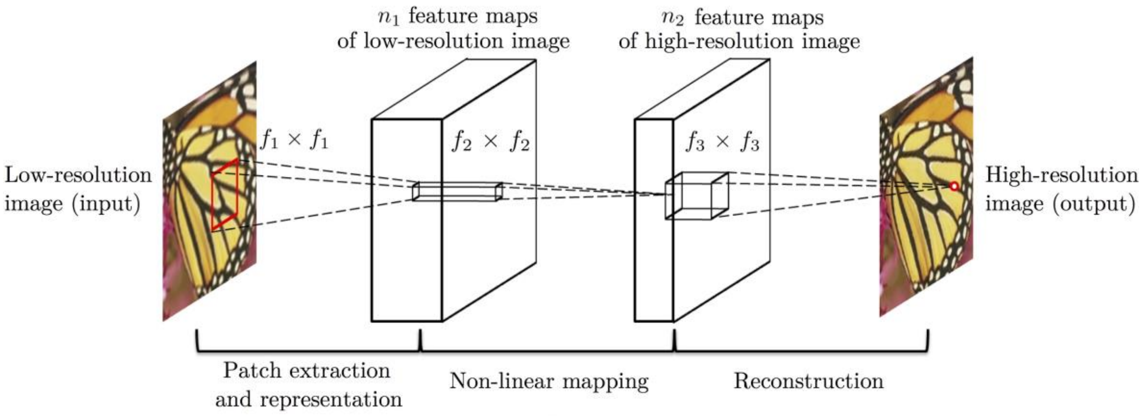
Introduction - Preliminary

Super-resolution (SR)

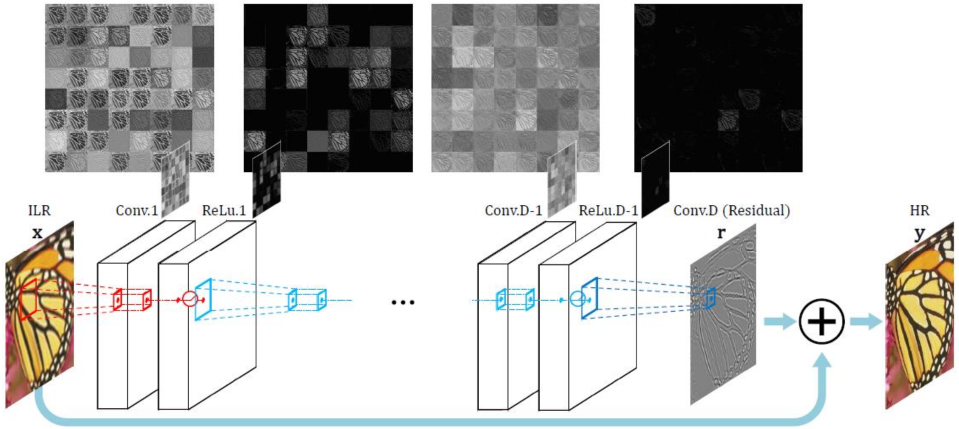
Introduction - Preliminary

Super-resolution with Convolutional Neural Network





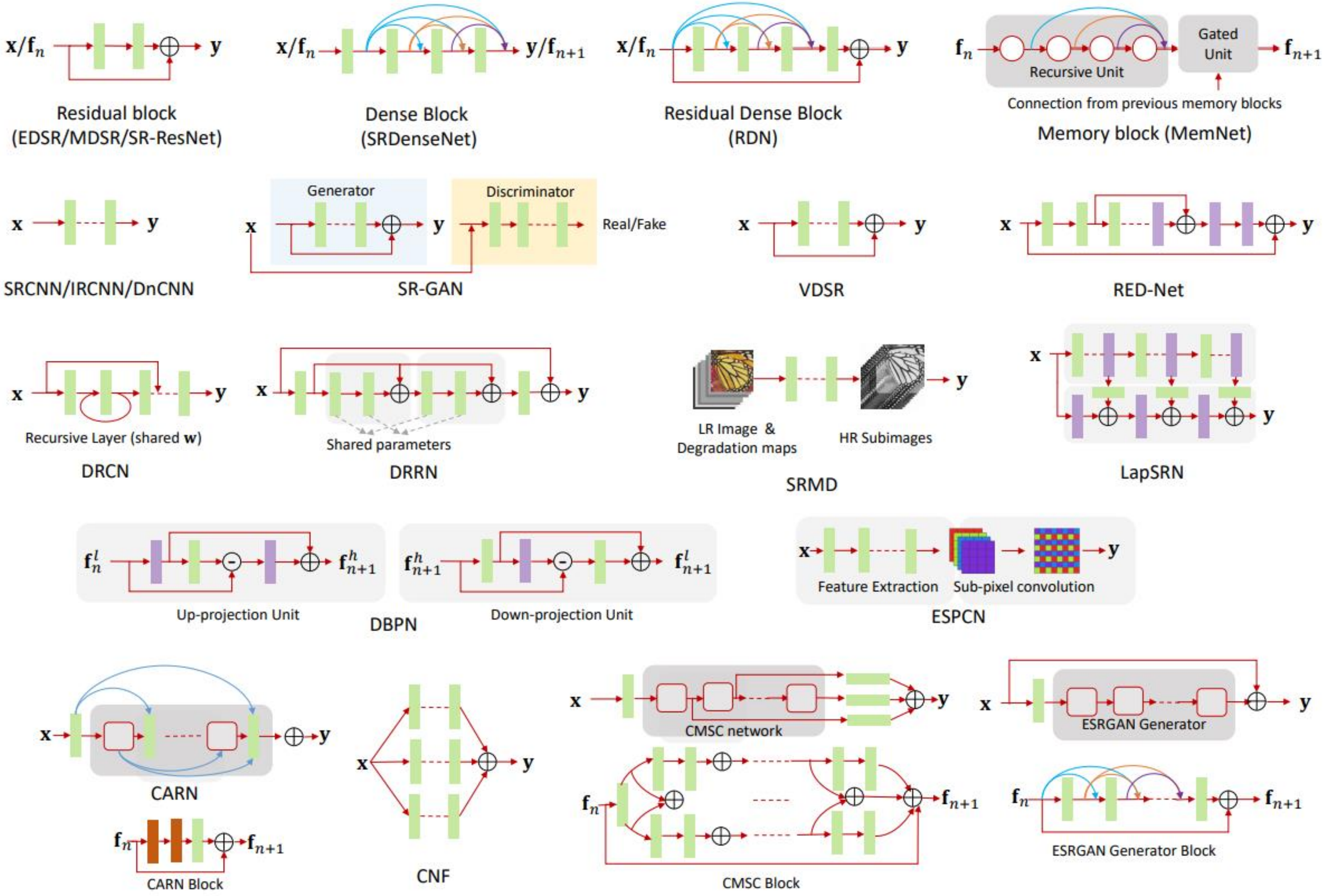
SRCNN



VDSR

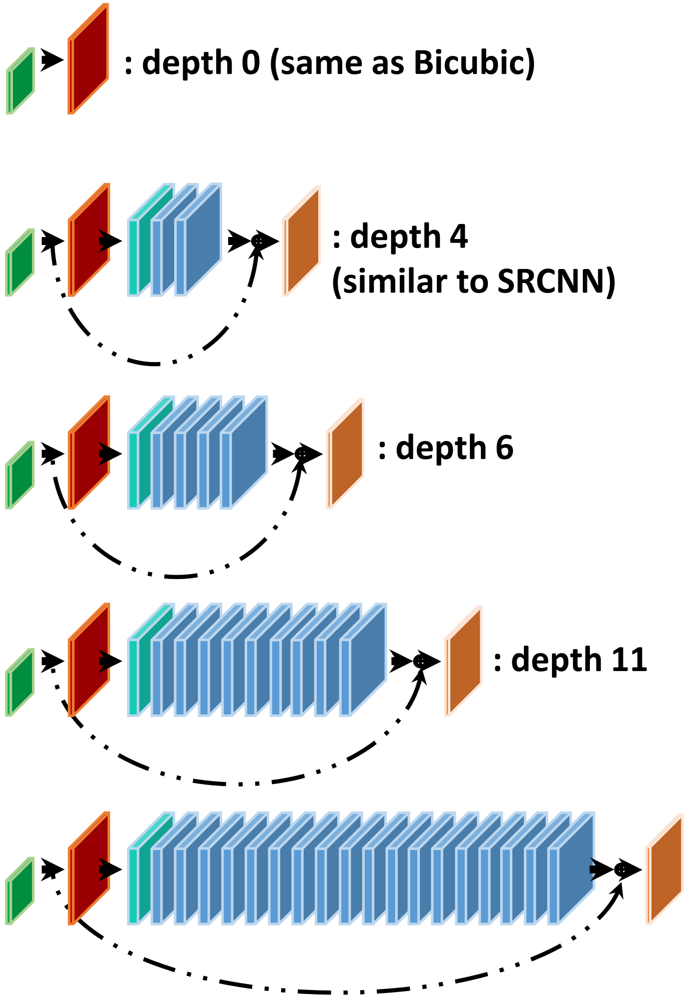
Introduction - Preliminary

Super-resolution with Convolution Neural Network



Introduction

Trade-off relationship between speed and performance over depth



Shallow, Faster, lower performance

Deeper, slower, higher performance

Depth	Bicubic	4	6	8	11	14	17	20
PSRN (dB)	30.4	32.56	33.01	33.229	33.379	33.435	33.495	33.523
SSIM	0.8682	0.91	0.916	0.918	0.92	0.92	0.921	0.922
Processing time (CPU)	0.002	0.321	0.5468	0.7725	0.994	1.317	1.622	1.96
Processing time (GPU)	0.001	0.01	0.012	0.0152	0.0189	0.0224	0.0262	0.0305
# of parameters	0	75K	148K	222K	333K	444K	555K	665K

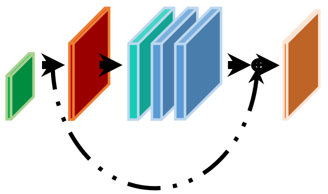
(processing time have measured on butterfly, 512 x 768)

Introduction - Motivation

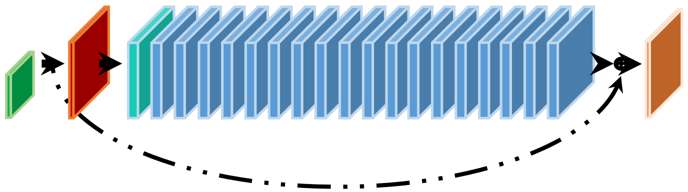
Model selection problem

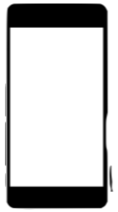
- **Model selection is important for efficiency and stability of the system**

Shallow model



Deeper model





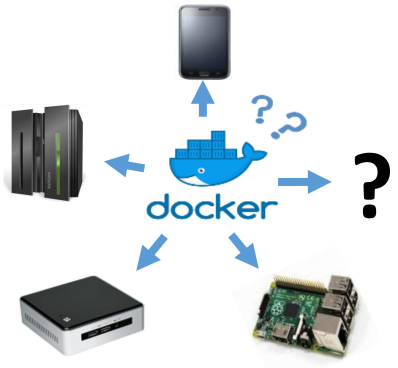
- **Faster, lighter**
- **Lower performance**
- **Suit for mobile, IoT devices**

- **Slower, heavier**
- **Higher performance**
- **Suit for server, station**

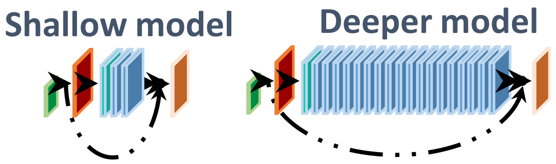
Necessity of Dynamic model adaptation

- However, sometimes it is hard to determine optimal model before running time
- For examples, when

- System requirements change during operating
  - Unfixed input rate (or image size)
  - Unfixed required scale (x4 is heavier than x2)
- Allocable system resource becomes insufficient



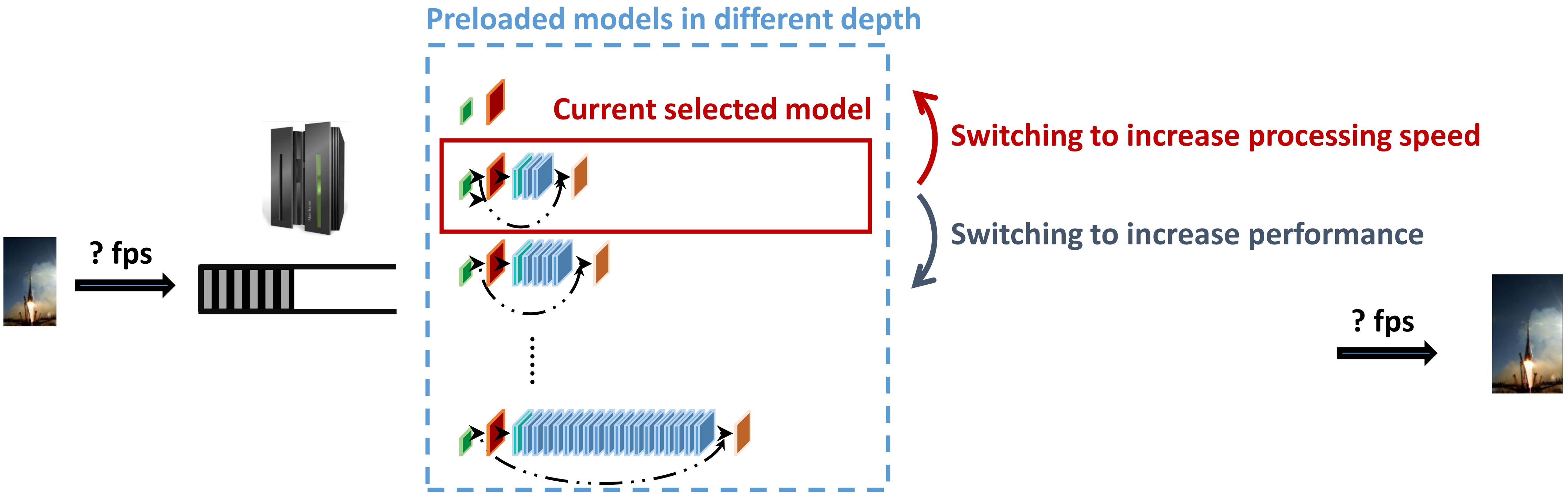
- Service provider do not have any information about clients



Introduction - Motivation

Adaptation with bag of models

- One simple reasonable solution is dynamically selecting SR models

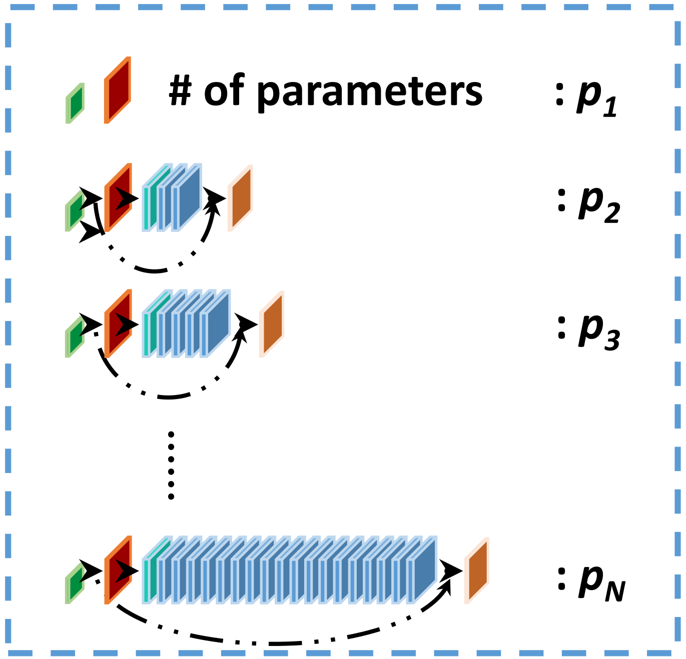


Introduction - Motivation

Limitation of adaptation with bag of models

- **However, It causes additional memory proportion to the number of the models used**

Preloaded models in difference depth



: $\sum_1^N p_n$

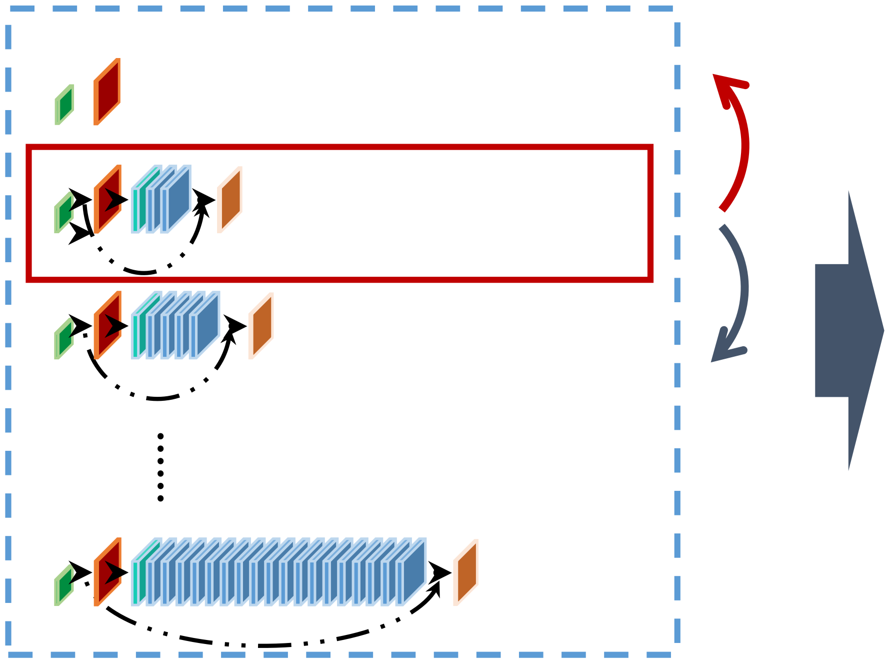
Disadvantages :

- More overhead when exporting new models to clients

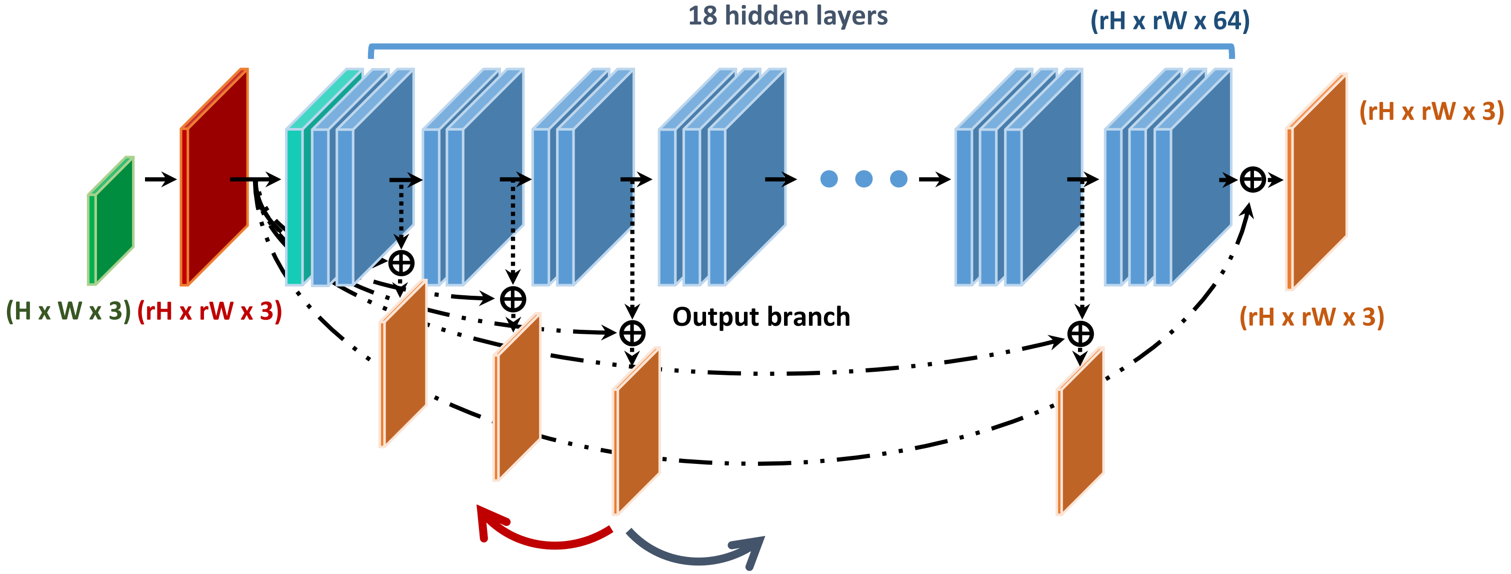
- Difficult FPGA and embedded deployment

Goal of this paper

- <u>Sharing the models</u> to reduce parameter redundancy
- <u>Maintainig comparable performance</u> to its unshared baseline model



Model switching

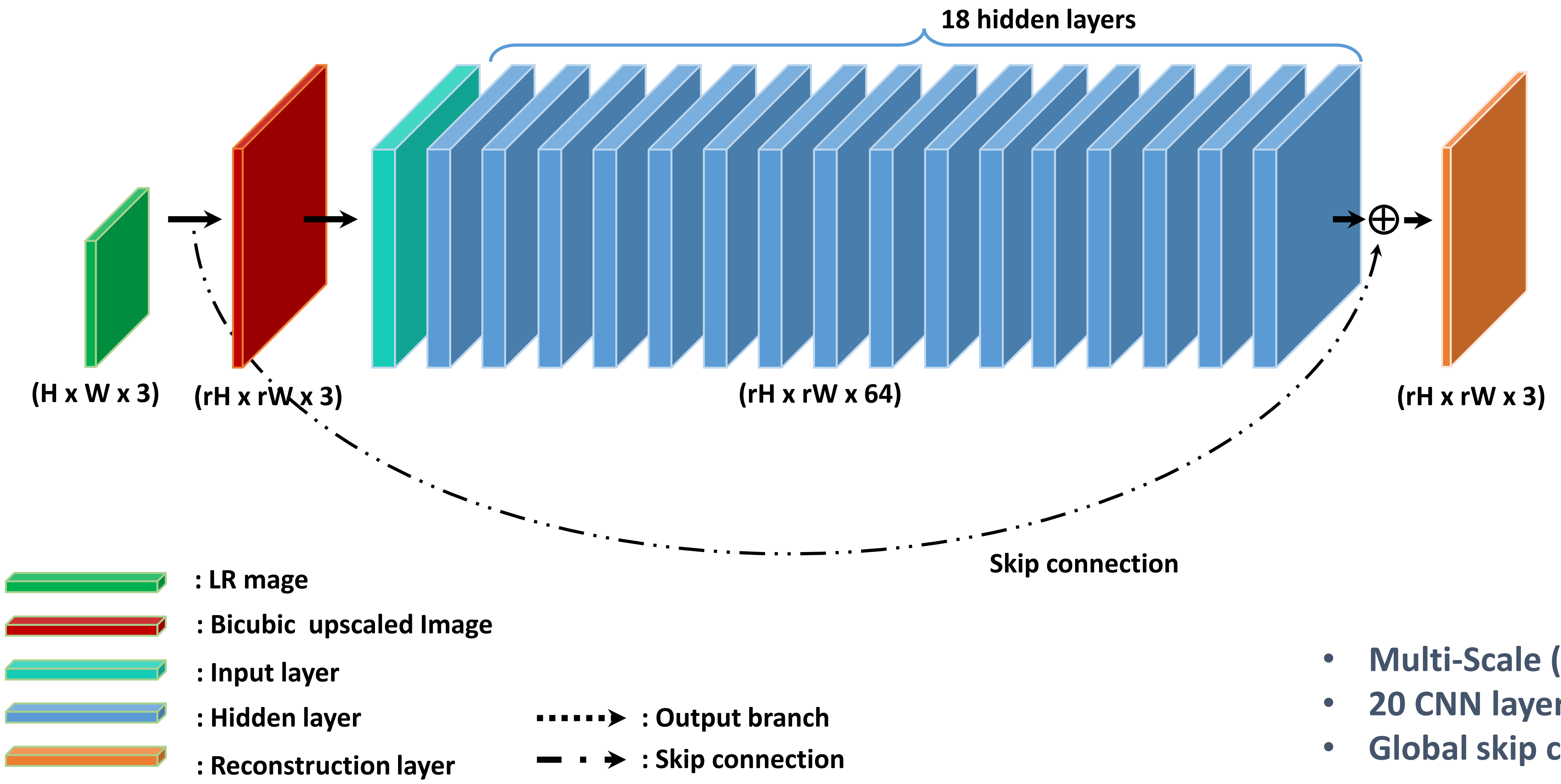


Shared depth controllable network

Depth Controllable SR Network

Introduction - Preliminary

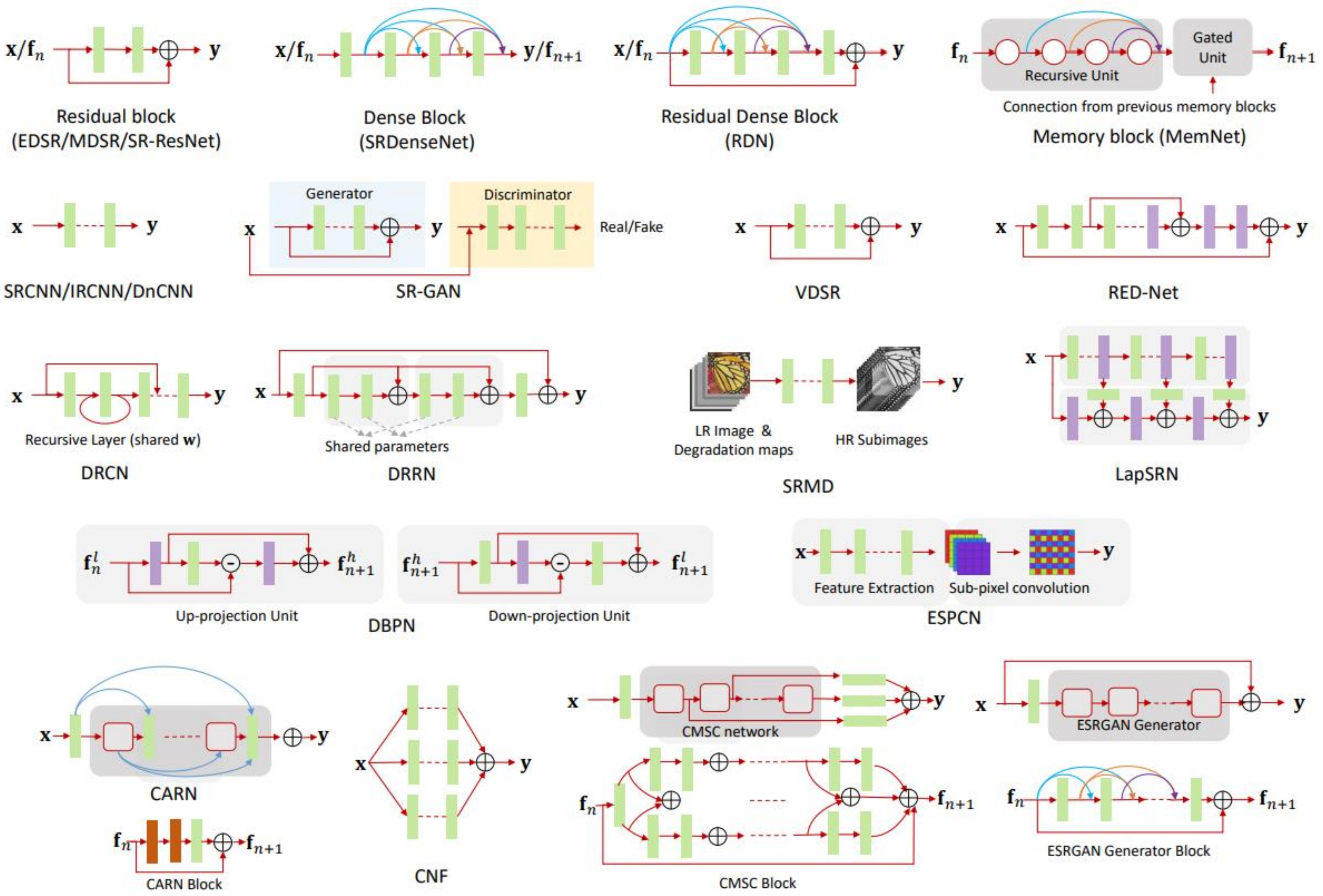
Implemental baseline - VDSR



- Multi-Scale (x2 ~ x4)
- 20 CNN layer
- Global skip connection

Introduction - Preliminary

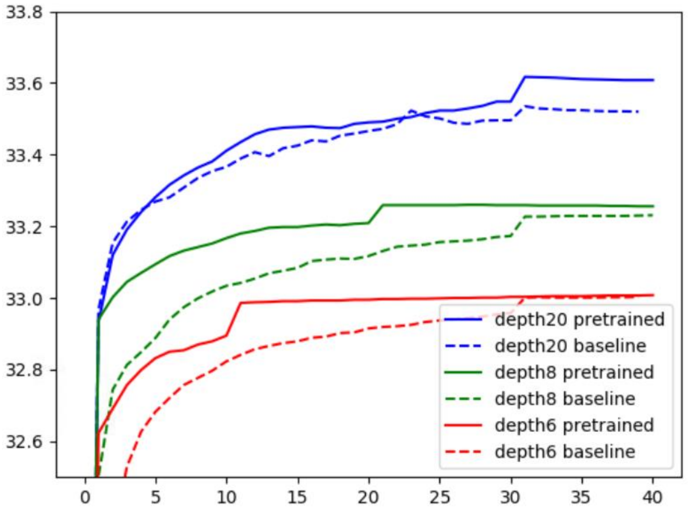
Implemental baseline - VDSR

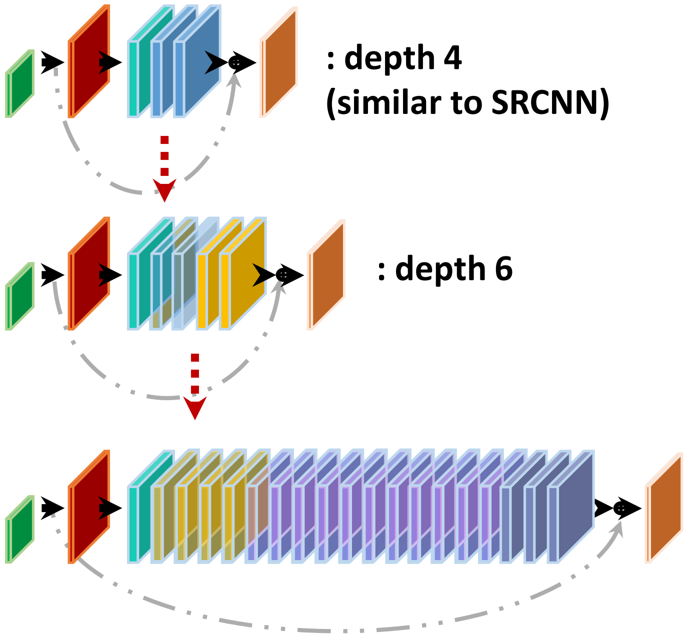


Depth Controllable SR Network (DCSRNet)

Motivation : depth-wise fine-tuning

Depth	Bicubic PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
Paper Baseline	30.4 / 0.8682	-	-	33.6 / 0.921
Implemental Baseline	30.4 / 0.8682	33.01/0.916	33.229/0.918	33.523/0.921
Depth-wise fine-tuning	30.4 / 0.8682	33.08/0.916	**33.256/0.9184**	**33.548/0.9215**



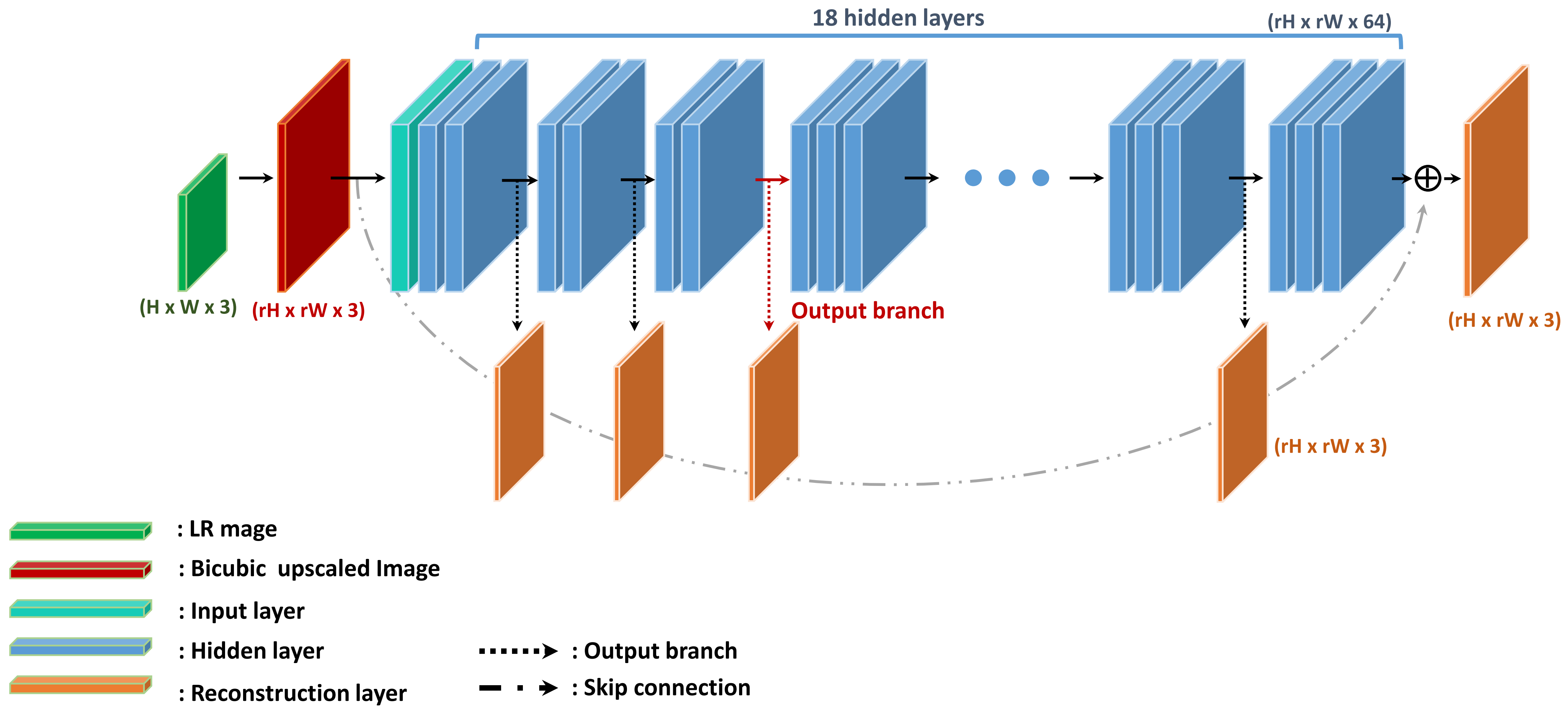


- Training with weights which pretrained at lower depth helps model to converge fast and to be trained stably
- This indirectly present there is close relation between depth (similar observation is addressed in B. Lim *et al.*)*
- Then, Our insight is that **the model can share their weight over depth**

- Enhanced Deep Residual Networks for Single Image Super-Resolution (EDSR) [https://arxiv.org/abs/1707.02921]

Depth Controllable SR Network (DCSRNet)

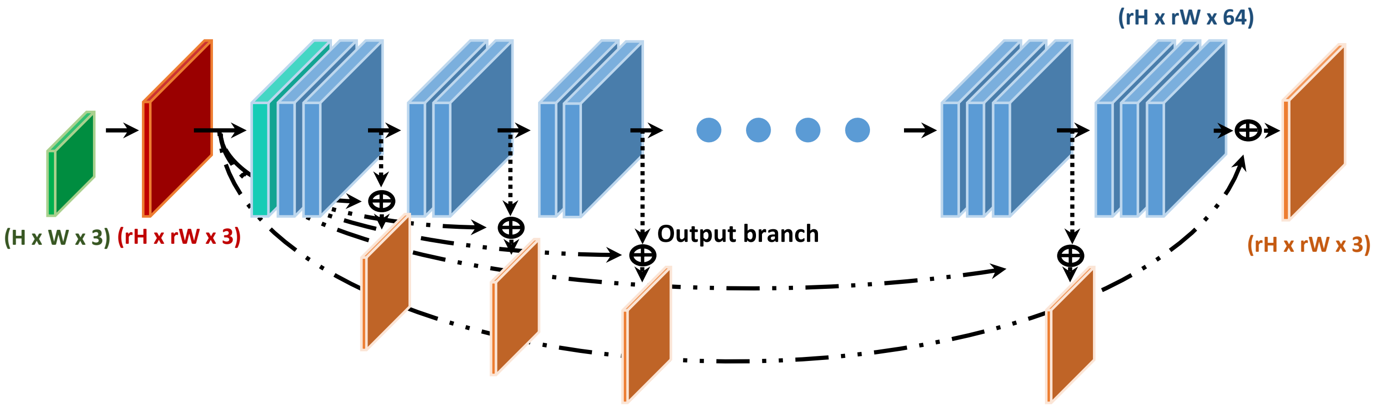
Shared Hidden layer Architecture



Depth Controllable SR Network (DCSRNet)

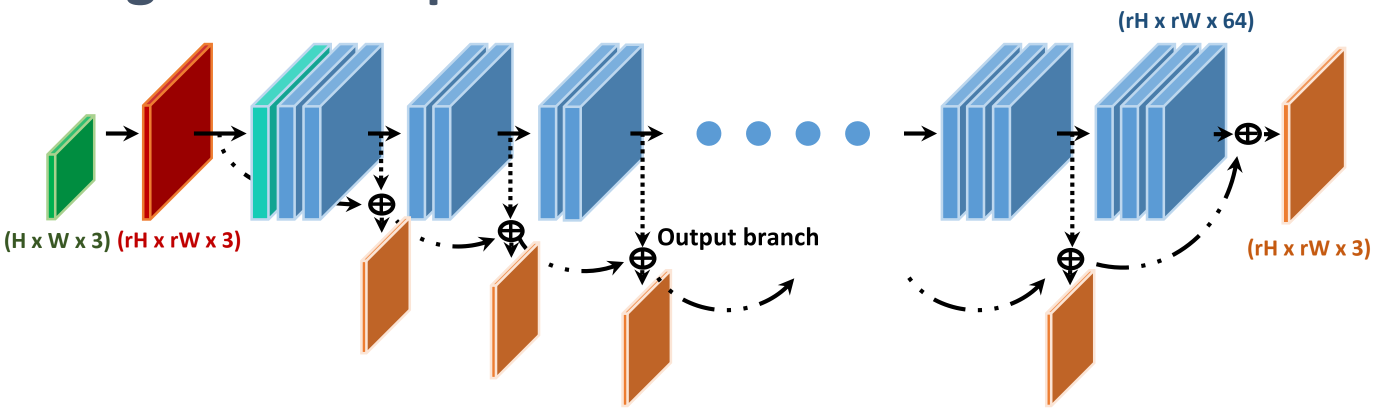
Skip connection strategy

- **Contiguous skip connection**



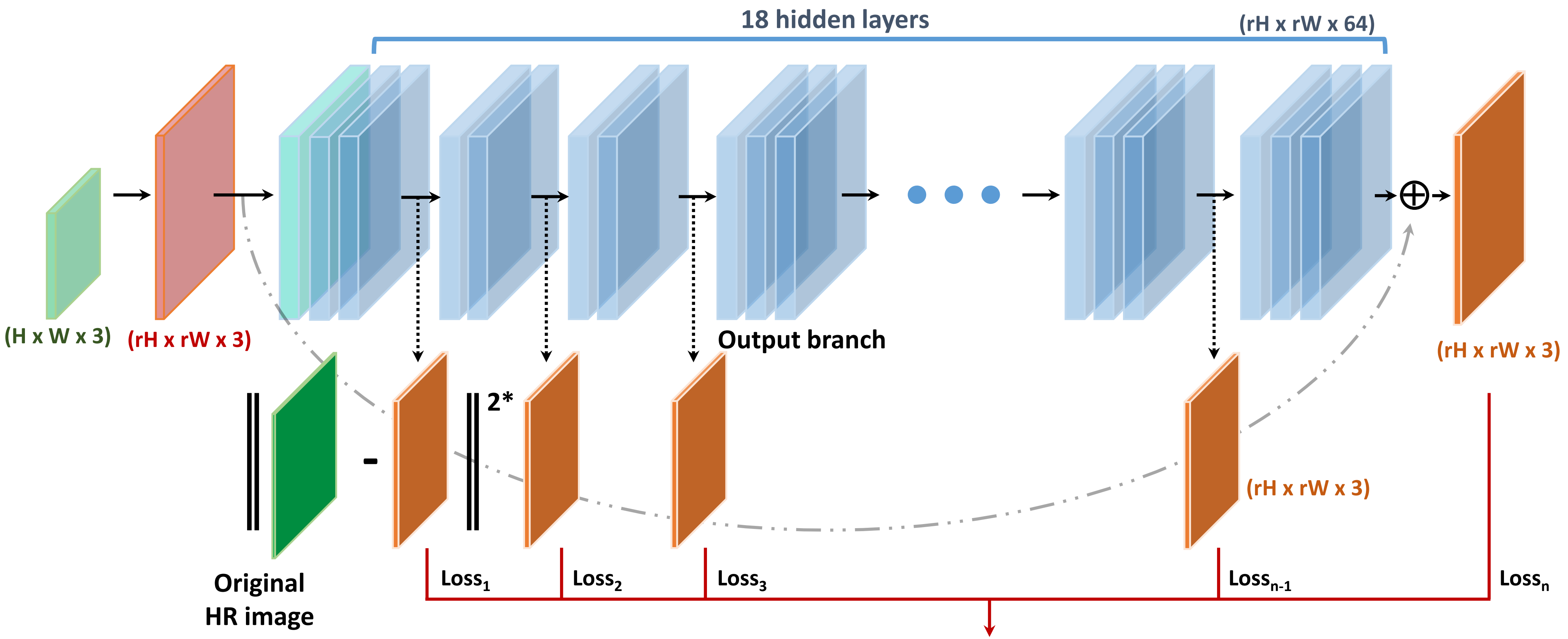
Better performance !

- **Progressive skip connection**



Depth Controllable SR Network (DCSRNet)

Shared Hidden layer Architecture



Backpropagation with averaged loss (multiply 1/n to every loss)

- Special case of auxiliary loss* when weight is 1/n instead of 0.3^n

* L1 loss result higher reconstruction score

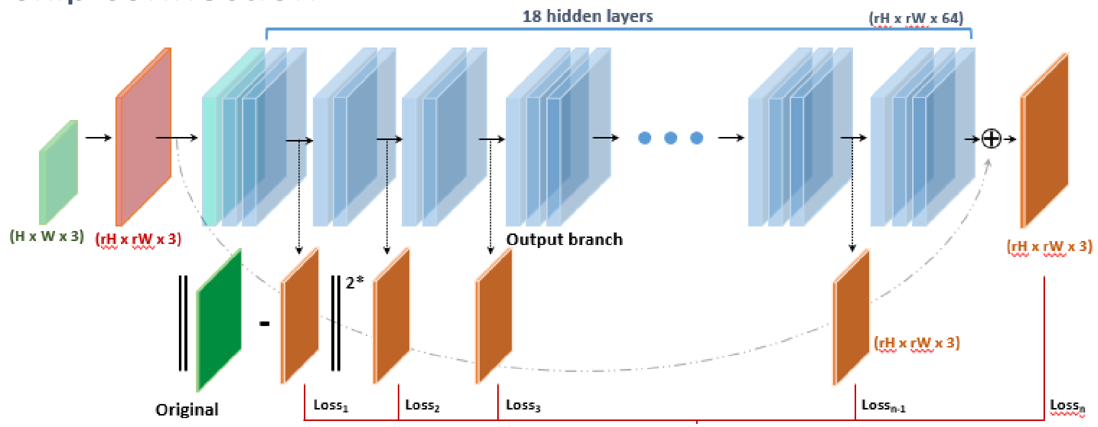
* Rethinking the inception architecture for computer vision (InceptionV2) [https://arxiv.org/pdf/1512.00567v1.pdf]

Depth Controllable SR Network (DCSRNet)

Evaluation result

Depth	**Bicubic PSNR(dB)/SSIM**	**4 (SRCNN) PSNR(dB)/SSIM**	**6 PSNR(dB)/SSIM**	**8 PSNR(dB)/SSIM**	**11 PSNR(dB)/SSIM**	**14 PSNR(dB)/SSIM**	**17 PSNR(dB)/SSIM**	**20 (VDSR) PSNR(dB)/SSIM**
Paper Baseline	30.4 / 0.8682	32.7 / 0.909	-	-	-	-	-	33.6 / 0.921
Implemental Baseline	30.4 / 0.8682	**32.56/0.91**	**33.01/0.916**	**33.229/0.918**	**33.379/0.92**	**33.435/0.92**	**33.46/0.9209**	**33.523/0.921**
Averaged loss (Contiguous skip connection)	30.4 / 0.8682	32.19/0.9058	32.79/0.9128	33.11/0.9164	33.3/0.9187	**33.48/0.9208**	**33.56/0.9218**	**33.59/0.9222**

- Performance increases proportional to depth of hidden layer
- Averaged loss seems to make training stable and be available to leverage performance of deeper model
- Contiguous skip connection shows better result than progressive skip connection

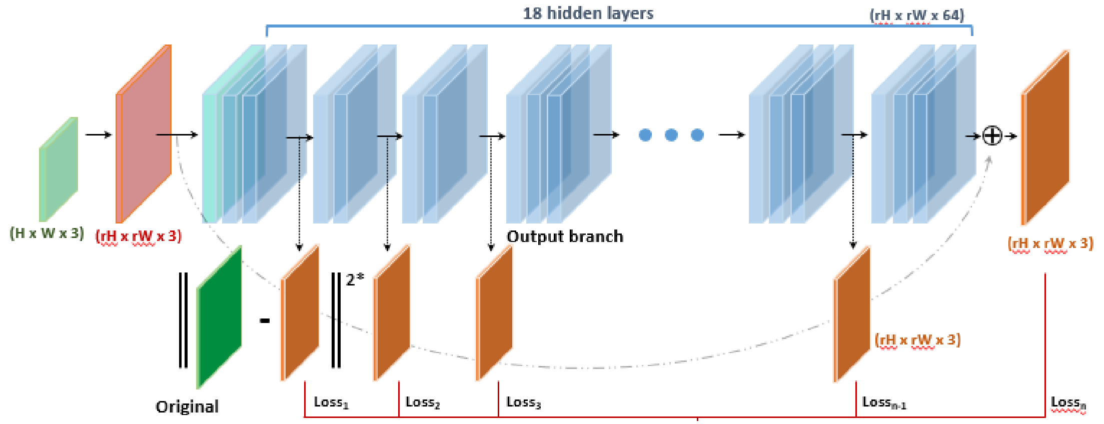


Depth Controllable SR Network (DCSRNet)

Performance bias problem

Depth	Bicubic PSNR(dB)/SSIM	4 (SRCNN) PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	11 PSNR(dB)/SSIM	14 PSNR(dB)/SSIM	17 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
Paper Baseline	30.4 / 0.8682	32.7 / 0.909	-	-	-	-	-	33.6 / 0.921
Implemental Baseline	30.4 / 0.8682	**32.56/0.91**	**33.01/0.916**	**33.229/0.918**	**33.379/0.92**	**33.435/0.92**	**33.46/0.9209**	**33.523/0.921**
Averaged loss (Contiguous skip connection)	30.4 / 0.8682	32.19/0.9058	32.79/0.9128	33.11/0.9164	33.3/0.9187	33.48/0.9208	33.56/0.9218	33.59/0.9222

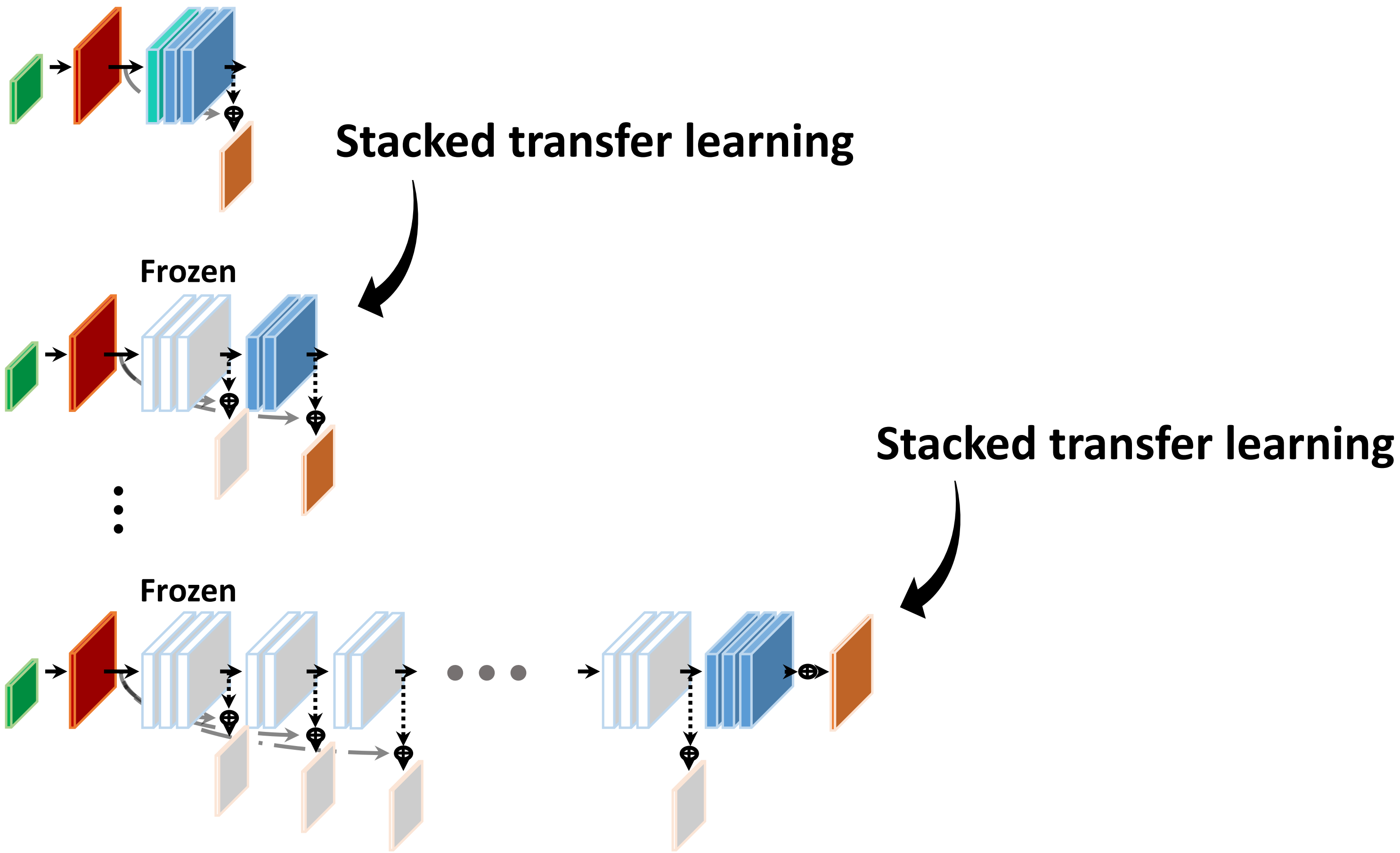
- However, training with averaged loss deteriorates performance at shallow layer
- Moreover, it is hard to predict which depth this deterioration start from



Depth Controllable SR Network (DCSRNet)

Attempts to improve : Stacked transfer learning

- **Gradually stacking additional layers to the previous model**

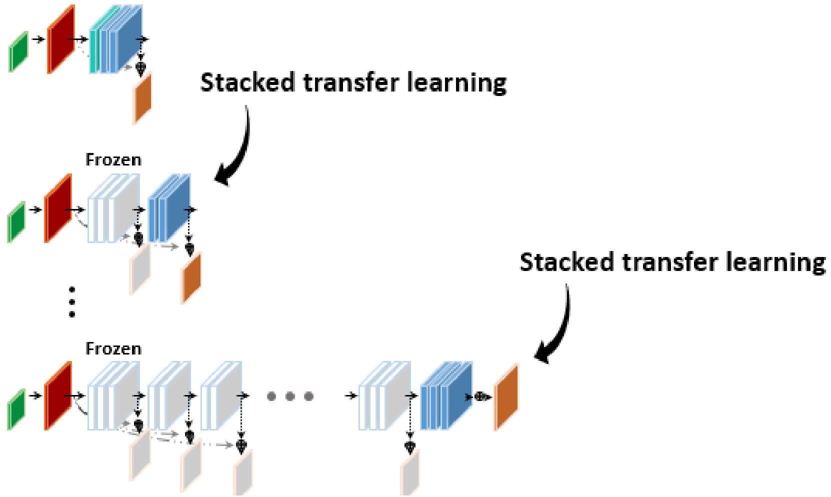


Depth Controllable SR Network (DCSRNet)

Limitation of Stacked transfer learning

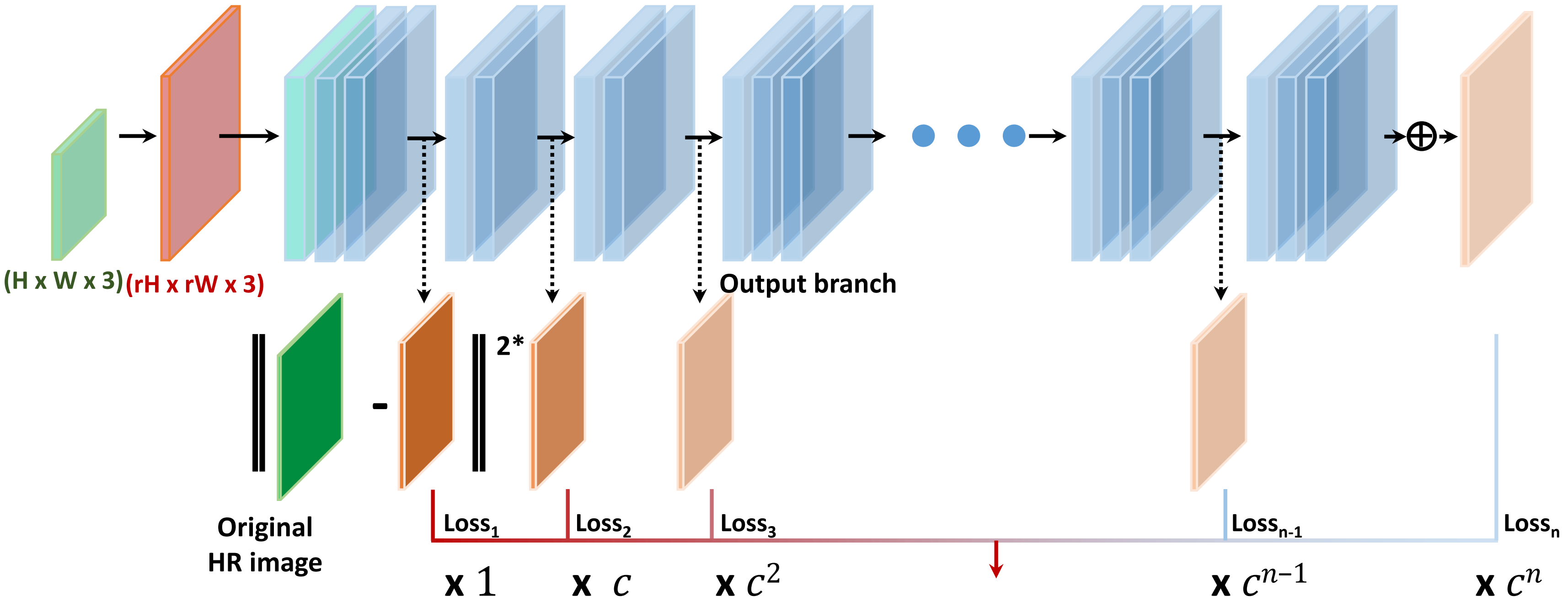
Depth	Bicubic PSNR(dB)/SSIM	4 (SRCNN) PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	11 PSNR(dB)/SSIM	14 PSNR(dB)/SSIM	17 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
Paper Baseline	30.4 / 0.8682	32.7 / 0.909	-	-	-	-	-	33.6 / 0.921
Implemental Baseline	30.4 / 0.8682	32.56/0.91	33.01/0.916	33.229/0.918	33.379/0.92	33.435/0.92	33.46/0.9209	33.523/0.921
Averaged loss	30.4 / 0.8682	**32.19/0.9058**	**32.79/0.9128**	**33.11/0.9164**	**33.3/0.9187**	**33.48/0.9208**	**33.56/0.9218**	**33.59/0.9222**
Staked transfer learning (soft pretraining)	30.4 / 0.8682	32.47/0.91	**32.86/0.9140**	32.952/0.9148	32.959/0.9149	32.96/0.9149	32.96/0.9149	32.965/0.9144

- Shallow layers can be trained explicitly
- However, the performance of deeper layer are deteriorated since there is no consideration of sharing parameters with deep layers



Depth Controllable SR Network (DCSRNet)

Compromise plan : Stacked transfer learning with inversed auxiliary loss



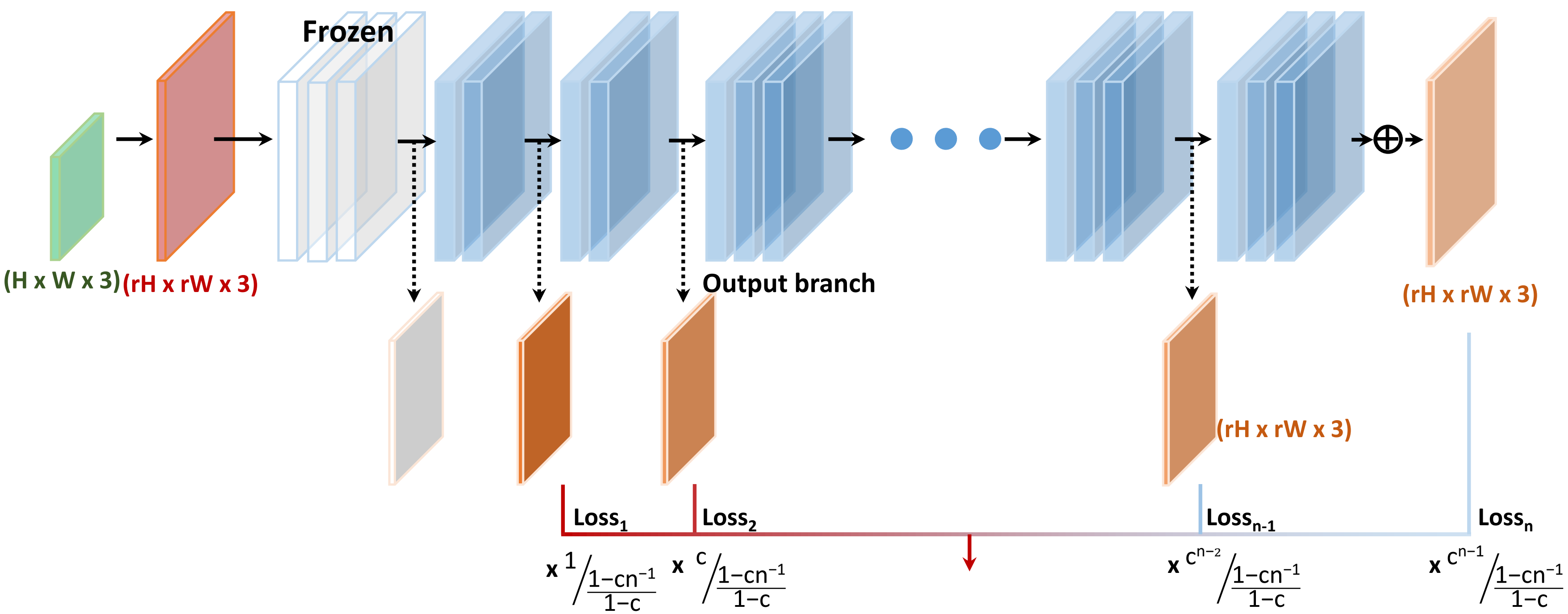
Backpropagation with inversed auxiliary loss
(multiply $c^{n-1} / \frac{1-cn^{-1}}{1-c}$ to every loss) [c=0.5]

- stacking with multiplying progressively smaller weight to loss against depth of layers
- The closer c is to 1, the closer it is to the averaged loss, so that the more it considers parameter sharing with deep layers, vise versa.

Depth Controllable SR Network (DCSRNet)

Compromise plan : Stacked transfer learning with inversed auxiliary loss

Extra Step 1

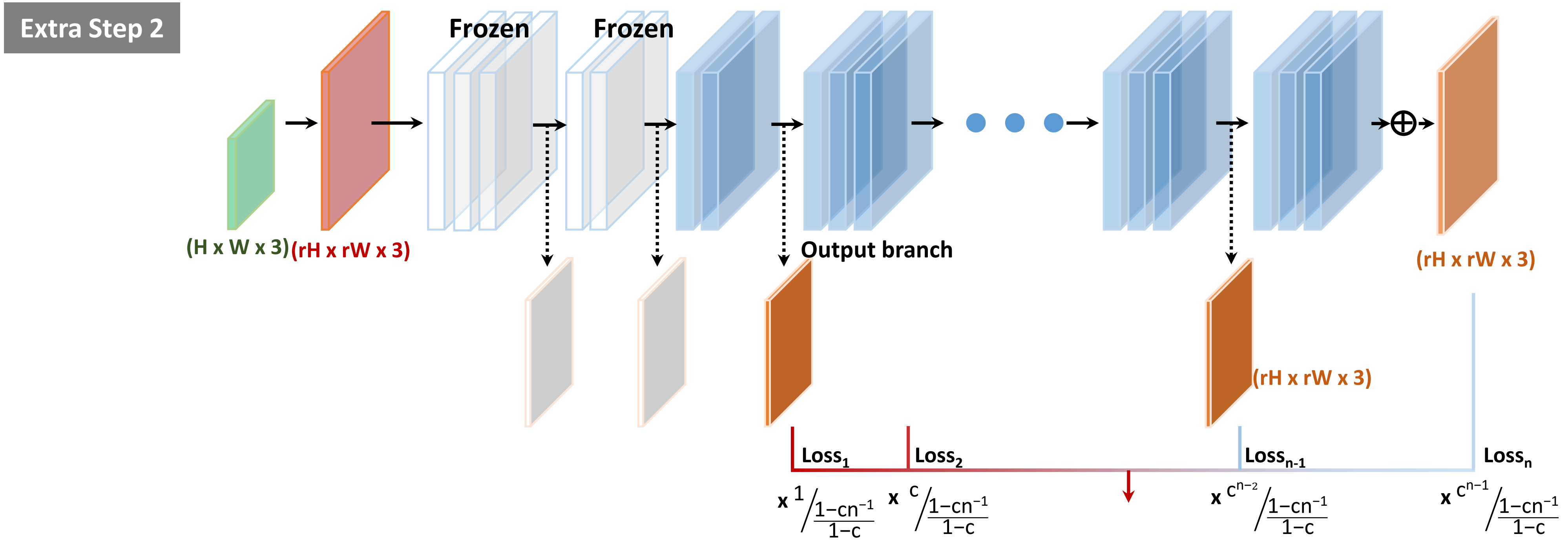


Backpropagation with inversed auxiliary loss
(multiply $c^{n-1} / \frac{1-c^{n-1}}{1-c}$ to every loss) [c=0.5]

- stacking with multiplying progressively smaller weight to loss against depth of layers
- The closer c is to 1, the closer it is to the averaged loss, so that the more it considers parameter sharing with deep layers, vise versa.

Depth Controllable SR Network (DCSRNet)

Compromise plan : Stacked transfer learning with inversed auxiliary loss

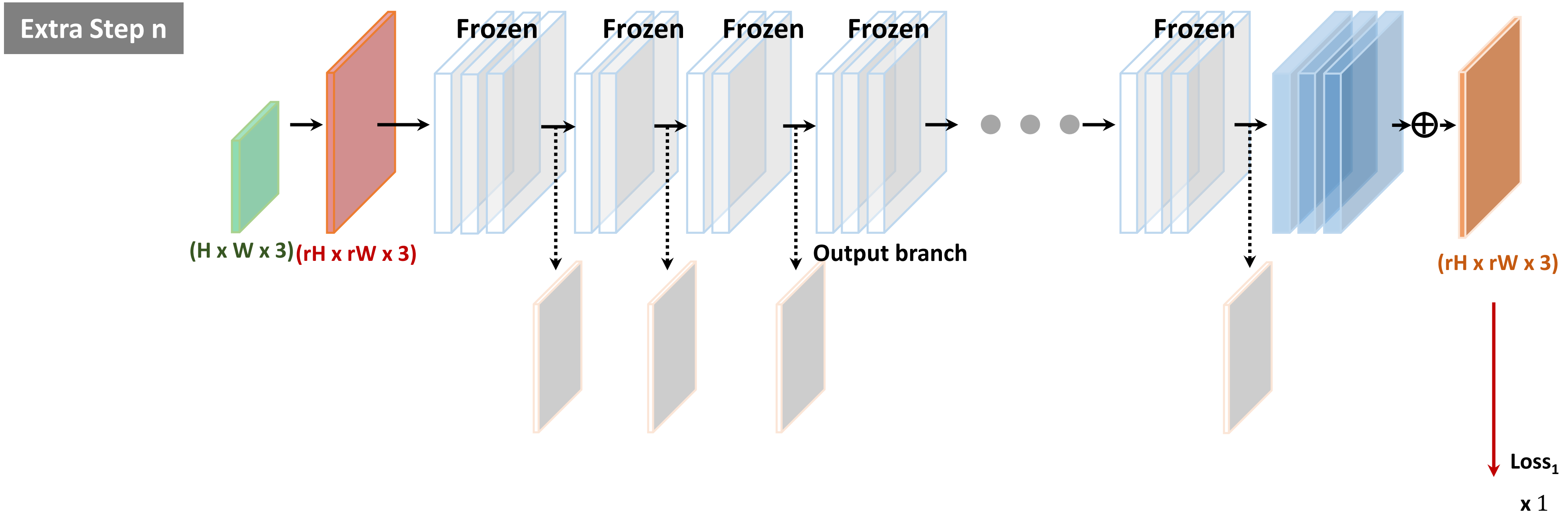


Backpropagation with inversed auxiliary loss
(multiply $c^{n-1} / \frac{1-c^{n-1}}{1-c}$ to every loss) [c=0.5]

- stacking with multiplying progressively smaller weight to loss against depth of layers
- The closer c is to 1, the closer it is to the averaged loss, so that the more it considers parameter sharing with deep layers, vise versa.

Depth Controllable SR Network (DCSRNet)

Compromise plan : Stacked transfer learning with inversed auxiliary loss



- Backpropagation with weighted loss (multiply $c^{n-1} / \frac{1-cn^{-1}}{1-c}$ to every loss) [c = 0.5]
- stacking with multiplying progressively smaller weight to loss against depth of layers
- The closer c is to 1, the closer it is to the averaged loss, so that the more it considers parameter sharing with deep layers, vise versa.

Experiments and Evaluations

Final Evaluation Result

Depth		Bicubic PSNR(dB)/SSIM	4 (SRCNN) PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	11 PSNR(dB)/SSIM	14 PSNR(dB)/SSIM	17 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
Set5									
Paper Baseline	x 2 x 3 x 4	33.66/0.9299 30.40/0.8682 28.42/0.8104	36.66/0.9542 32.7/0.909 30.48/0.8628	-	-	-	-	-	37.53/0.9587 33.6/0.921 31.35/0.8838
Implemental Baseline	x 2 x 3 x 4	**33.66/0.9299 30.40/0.8682 28.42/0.8104**	**36.56/0.9551 32.56/0.91 30.29/0.8649**	**37.05/0.9577 33.01/0.916 30.64/0.8736**	**37.22/0.9584 33.229/0.918 30.84/0.8774**	**37.33/0.9589 33.379/0.92 30.97/0.8795**	**37.36/0.9591 33.435/0.92 31.07/0.8816**	**37.37/0.9592 33.46/0.9209 31.1/0.8834**	**37.4/0.8252 33.523/0.921 31.15/0.8837**
Averaged loss	x 2 x 3 x 4	33.66/0.9299 30.40/0.8682 28.42/0.8104	35.59/0.931 32.19/0.9058 29.81/0.8523	36.82/0.9501 32.79/0.9128 30.45/0.869	37.13/0.9563 33.11/0.9164 30.7/0.8745	37.27/0.9579 33.3/0.9187 30.89/0.8784	**37.39/0.9582 33.48/0.9208 31.12/0.882**	**37.43/0.9594 33.56/0.9218 31.21/0.8837**	**37.45/0.9595 33.59/0.9222 31.26/0.8845**
inversed auxiliary loss 0.5	x 2 x 3 x 4	33.66/0.9299 30.40/0.8682 28.42/0.8104	36.26/0.9537 32.43/0.9093 30.13/0.8612	36.98/0.9572 32.97/0.915 30.58/0.8722	37.18/0.9582 33.18/0.9172 30.79/0.8726	37.31/0.9587 33.38/0.9195 31.01/0.8761	37.38.0.959 33.49/0.9207 31.15/0.8821	37.41/0.9591 33.55/0.9213 31.22/0.8831	37.41.0.9597 33.57/0.9214 31.23/0.8833

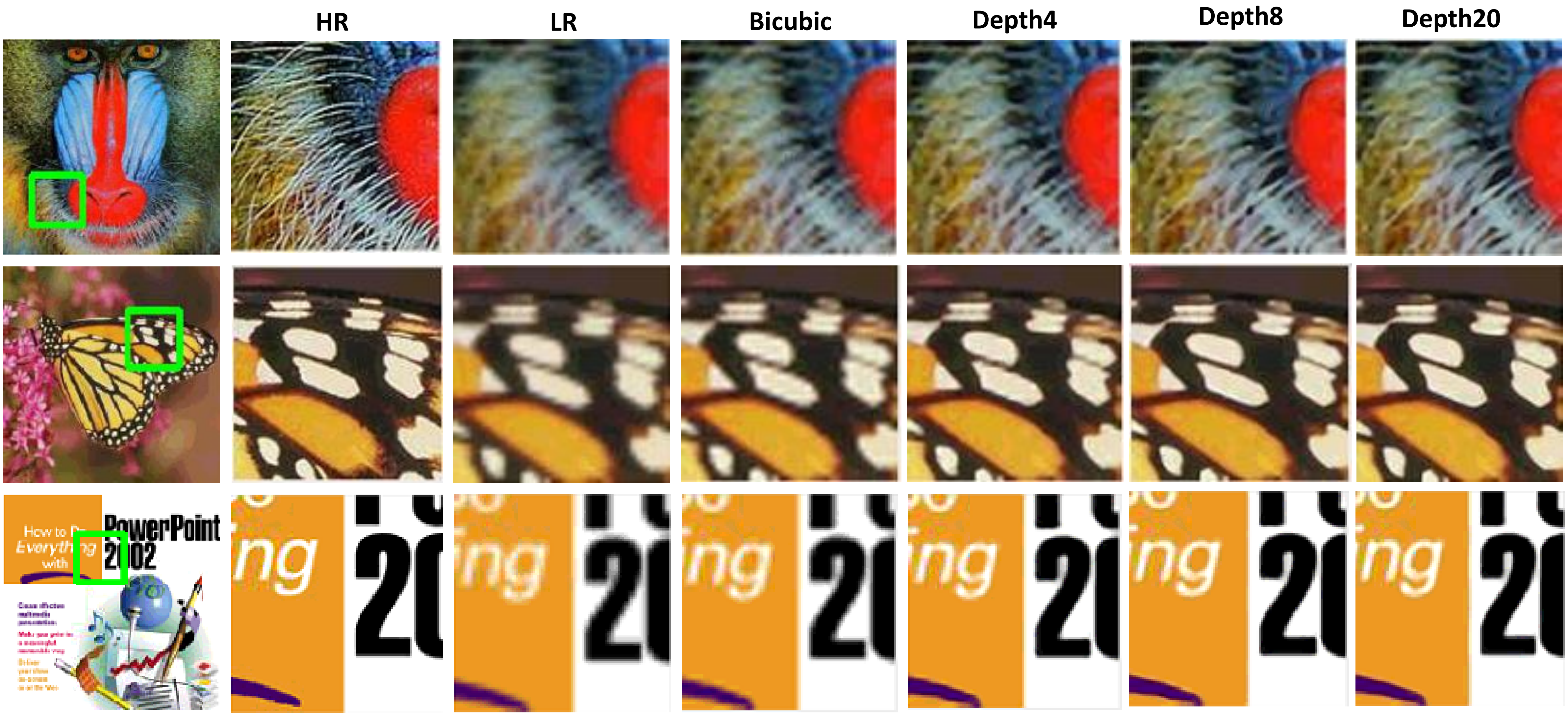
Final Evaluation Result

Depth		Bicubic PSNR(dB)/SSIM	4 (SRCNN) PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	11 PSNR(dB)/SSIM	14 PSNR(dB)/SSIM	17 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
Set14									
Paper Baseline	x 2 x 3 x 4	30.24/0.8688 27.55/0.7742 26.00/0.7027	32.42/0.9063 29.28/0.8209 27.49/0.7503	-	-	-	-	-	33.03/0.9124 29.77/0.8314 28.01/0.7674
Implemental Baseline	x 2 x 3 x 4	30.24/0.8688 27.55/0.7742 26.00/0.7027	**32.46/0.9064** **29.17/0.8195** **27.34/0.7480**	**32.8/0.9099** **29.47/0.8251** **27.58/0.7567**	**32.9/0.9106** **29.56/0.8273** **27.7/0.7593**	**32.99/0.9112** **29.66/0.8285** **27.76/0.7614**	**33.01/0.9115** **29.71/0.8296** **27.8/0.7632**	**33.02/0.9115** **29.71/0.8301** **27.82/0.7639**	**33.03/0.9117** **29.71/0.8301** **27.84/0.7639**
Averaged loss	x 2 x 3 x 4	30.24/0.8688 27.55/0.7742 26.00/0.7027	32.01/0.9022 28.95/0.8147 26.99/0.7356	32.65/0.9082 29.34/0.8222 27.46/0.7525	32.85/0.9101 29.52/0.8259 27.63/0.7573	32.94/0.911 29.63/0.8279 27.74/0.7605	**33.03/0.9117** **29.72/0.8299** 27.84/0.7634	**33.07/0.9122** **29.75/0.8307** **27.89/0.7649**	**33.08/0.9123** **29.76/0.8309** **27.9/0.7656**
inversed auxiliary loss 0.5	x 2 x 3 x 4	30.24/0.8688 27.55/0.7742 26.00/0.7027	32.31/0.9053 29.12/0.8191 27.22/0.7444	32.74/0.9093 29.43/0.8241 27.51/0.7549	32.88/0.9104 29.57/0.8266 27.66/0.7587	32.98/0.9112 29.67/0.8286 27.79/0.762	33.02/0.9116 29.72/0.8295 **27.85/0.7635**	33.04/0.9118 29.74/0.8298 27.88/0.7642	33.04.0.9119 29.74/0.9298 27.88/0.7644

Final Evaluation Result

Depth		Bicubic PSNR(dB)/SSIM	4 (SRCNN) PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	11 PSNR(dB)/SSIM	14 PSNR(dB)/SSIM	17 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
BSD100									
Paper Baseline	x 2 x 3 x 4	29.56/0.8431 27.21/0.7385 25.96/0.6675	31.36/0.8879 28.41/0.7863 26.90/0.7107	-	-	-	-	-	31.90/0.8960 28.82/0.7976 27.89/0.7251
Implemental Baseline	x 2 x 3 x 4	29.56/0.8431 27.21/0.7385 25.96/0.6675	31.30/0.889 28.26/0.7853 26.75/0.7082	31.56/0.893 28.466/0.791 26.92/0.7166	31.65/0.8941 28.56/0.7932 27/0.7188	31.71/0.895 28.61/0.7947 27.05/0.7207	31.76/0.8954 28.65/0.7959 27.08/0.7221	31.76/0.8957 28.66/0.7963 27.09/0.7229	31.76/0.8958 28.66/0.7965 27.1/0.7231
Averaged loss	x 2 x 3 x 4	29.56/0.8431 27.21/0.7385 25.96/0.6675	31/0.8849 28.1/0.7804 26.51/0.6966	31.42/0.8908 28.37/0.7879 26.83/0.7128	31.59/0.8933 28.49/0.7916 26.94/0.7171	31.68/0.8946 28.58/0.7938 27.02/0.7198	**31.75/0.8956** 28.65/0.7961 **27.1.0.7225**	**31.79.0.8962** **28.69/0.797** **27.12/0.7238**	**31.8/0.8963** **28.7/0.7974** **27.14/0.7244**
inversed auxiliary loss 0.5	x 2 x 3 x 4	29.56/0.8431 27.21/0.7385 25.96/0.6675	31.2/0.8879 28.22/0.7851 26.67/0.7051	31.5/0.8922 28.44/0.7898 26.89/0.7151	31.61/0.8937 28.54/0.7925 26.97/0.7179	31.69/0.8947 28.62/0.7947 27.05/0.7209	31.74/0.8952 **28.66/0.7958** 27.09/0.7222	31.75/0.8953 28.67/0.7962 27.1/0.7229	31.75/0.8955 28.67/0.7964 27.1/0.7231

Final Evaluation Results



Further Experiments

- **Comparison on different coefficient of inversed auxiliary loss**

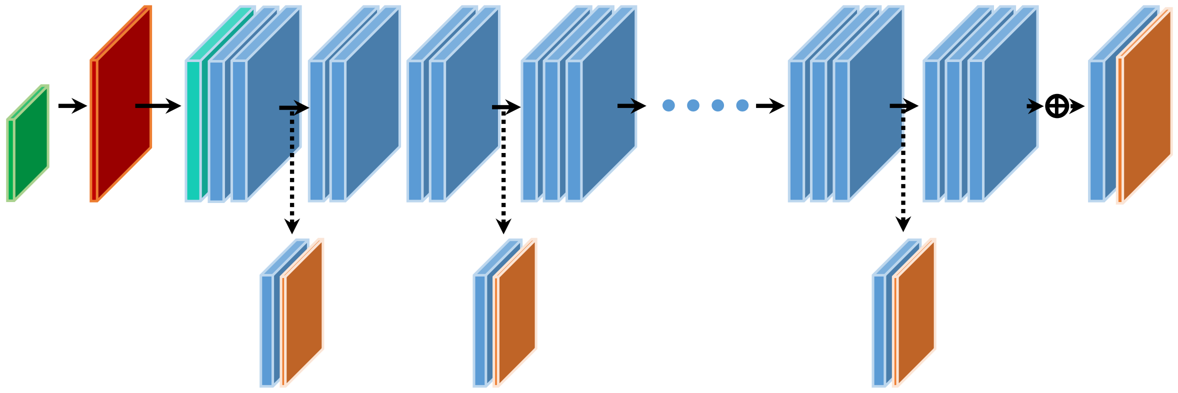
Depth	**Bicubic PSNR(dB)/SSIM**	**4 (SRCNN) PSNR(dB)/SSIM**	**6 PSNR(dB)/SSIM**	**8 PSNR(dB)/SSIM**	**11 PSNR(dB)/SSIM**	**14 PSNR(dB)/SSIM**	**17 PSNR(dB)/SSIM**	**20 (VDSR) PSNR(dB)/SSIM**
Paper Baseline	30.4 / 0.8682	32.7 / 0.909	-	-	-	-	-	33.6 / 0.921
Inversed auxiliary loss 0.7	30.4 / 0.8682	**32.34/0.9079**	**32.9/0.9141**	**33.16/0.917**	33.37/0.919	33.51/0.921	33.56/0.921	33.58/0.922
inversed auxiliary loss 0.5	30.4 / 0.8682	32.43/0.9093	**32.97/0.915**	33.18/0.9172	33.38/0.9195	33.49/0.9207	33.55/0.9213	33.57/0.9214
inversed auxiliary loss 0.3	30.4 / 0.8682	**32.45/0.9092**	32.907/0.9148	**33.75/0.9171**	**33.33/0.9193**	**33.44/0.9203**	**33.47/0.9206**	**33.47/0.9206**

Further Experiments

- **Comparison of different depth of output branch**

Depth	Bicubic PSNR(dB)/SSIM	4 (SRCNN) PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	11 PSNR(dB)/SSIM	14 PSNR(dB)/SSIM	17 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
Paper Baseline	30.4 / 0.8682	32.7 / 0.909	-	-	-	-	-	33.6 / 0.921
Implemental Baseline	30.4 / 0.8682	32.56/0.91	33.01/0.916	33.229/0.918	33.379/0.92	33.435/0.92	33.46/0.9209	33.523/0.921
Averaged loss (branch depth 1)	30.4 / 0.8682	**32.19/0.9058**	**32.79/0.9128**	**33.11/0.9164**	**33.3/0.9187**	**33.48/0.9208**	**33.56/0.9218**	**33.59/0.9222**
branch depth 2	30.4 / 0.8682	**32.446/0.9088**	**33.026/0.9151**	**33.306/0.9185**	**33.487/0.9208**	**33.603/0.9223**	**33.647/0.9228**	**33.666/0.923**

- The deeper output branch shows the higher performance , but requires the more parameters



Further Experiments

- **comparison of interval of output branch**

Depth	Bicubic PSNR(dB)/SSIM	4 (SRCNN) PSNR(dB)/SSIM	6 PSNR(dB)/SSIM	8 PSNR(dB)/SSIM	11 PSNR(dB)/SSIM	14 PSNR(dB)/SSIM	17 PSNR(dB)/SSIM	20 (VDSR) PSNR(dB)/SSIM
Paper Baseline	30.4 / 0.8682	32.7 / 0.909	-	-	-	-	-	33.6 / 0.921
Implemental Baseline	30.4 / 0.8682	32.56/0.91	33.01/0.916	33.229/0.918	33.379/0.92	33.435/0.92	33.46/0.9209	33.523/0.921
Averaged loss	30.4 / 0.8682	**32.19/0.9058**	**32.79/0.9128**	**33.11/0.9164**	33.3/0.9187	**33.48/0.9208**	33.56/0.9218	**33.59/0.9222**

Depth	Bicubic PSNR(dB)	4 (SRCNN) PSNR(dB)	6 PSNR(dB)	8 PSNR(dB)	10 PSNR(dB)	12 PSNR(dB)	14 PSNR(dB)	16 PSNR(dB)	18 PSNR(dB)	20 (VDSR) PSNR(dB)
Paper Baseline	30.4	32.7 / 0.909	-	-	-	-	-	-	-	33.6 / 0.921
Averaged loss (interval 2)	30.4	**32.179 /0.9051**	**32.668 /0.9114**	**33.045 /0.9155**	33.285 /0.9182	33.425 /0.9199	**33.538 /0.9212**	33.618 /0.9220	33.653 /0.9005	**33.666/0.9227**

- The shorter interval of branch shows the higher performance at the deeper layer, but requires the more parameters
- It is useful that training with shorter branch then using only some of them when the deeper layer is more important

Discussion and Q&A

- **Contribution**
  - We proposed **depth controllable** super-resolution network in form of **shared single architecture** while maintaining **comparable performance**

- **Future work**
  - Adapt this work to other SOTA methods

- **Q&A**