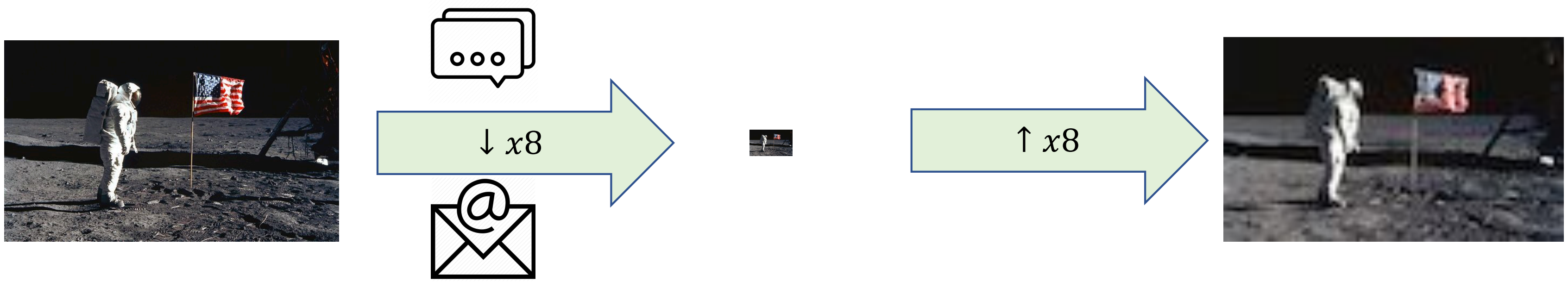
Efficient Task Aware Downscaling for Super-Resolution and Colorization

Improving Super-Resolution and Colorization in Image and Video Domain

Semester Project by Simon Schaefer

Advisor: Dr. Radu Timofte, Dr. Shuhang Gu; Supervisor: Prof. Luc Van Gool

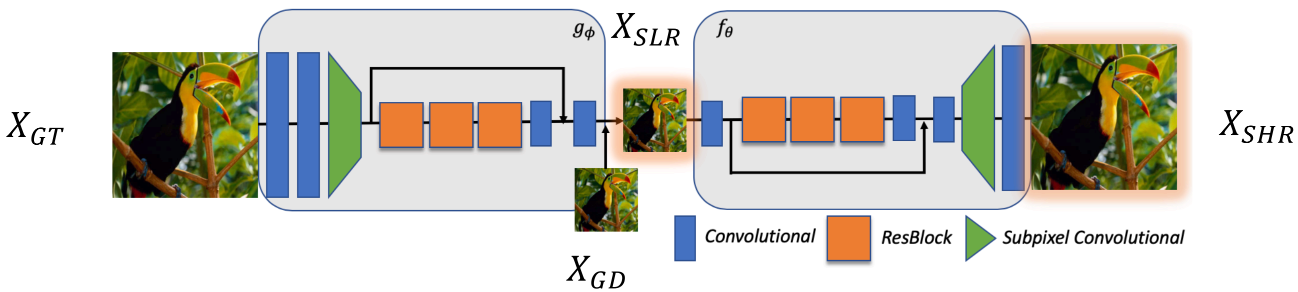
Why do we need Task Aware Downscaling ?



Idea: Downscale image aware of the upscaling process
so that the downscaled image still is human readable
(improve restoration quality with constant compression rate)

Related Work (Kim et al, ECCV'18)

- Fully-Convolutional autoencoder architecture
- End-to-End trainable architecture
- Proof of Concept for SISR and IC problems

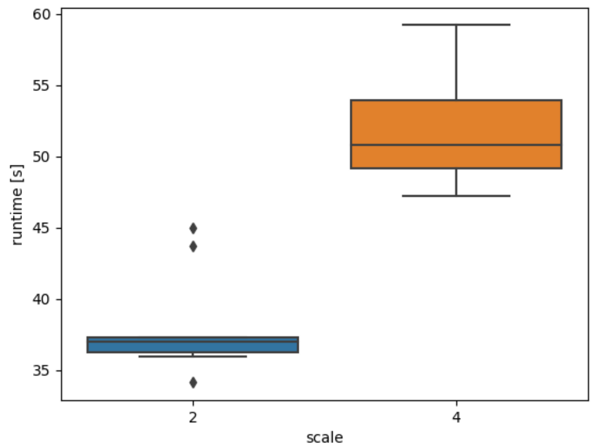


Related Work (Kim et al, ECCV'18) - Problems

White Gaussian noise

$\sigma = 28$

Very vulnerable to perturbations of the SLR image



Runtime on CPU (Mac Pro 2015) per forward pass

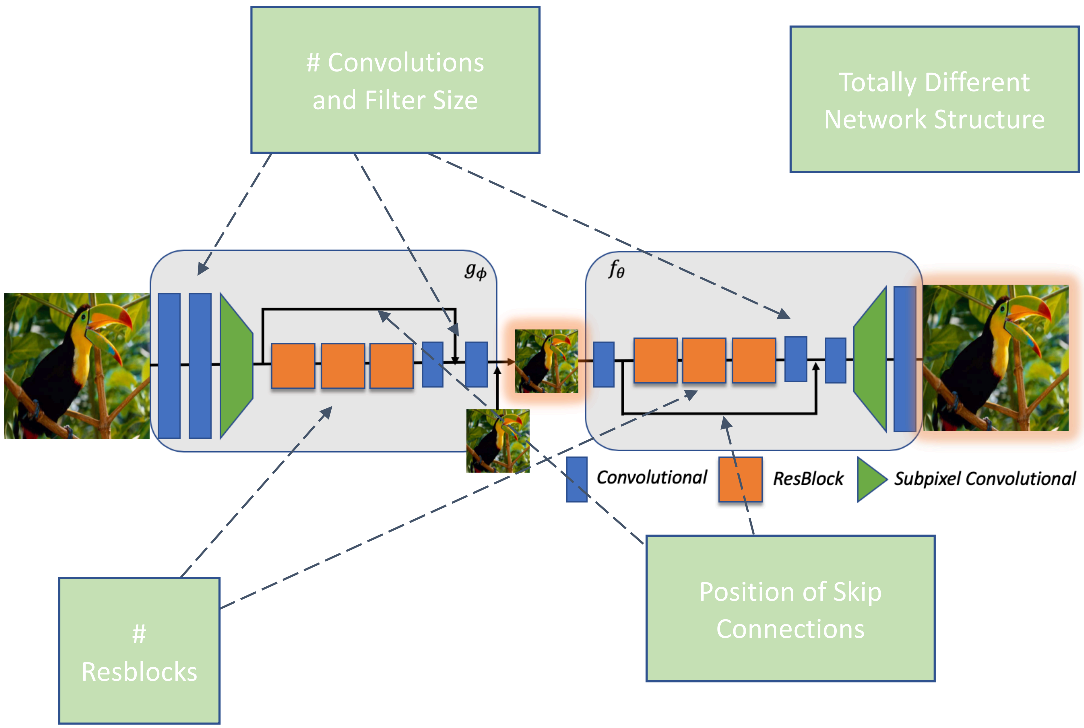
Project Goals

- Improve efficiency for Super Resolution and Colorization
- Increase Robustness against Perturbation
- Extend to Video Domain

General Improvements - Model

Architecture

Training



Adaptive Learning Rate

Data Augmentation

Loss Weighting

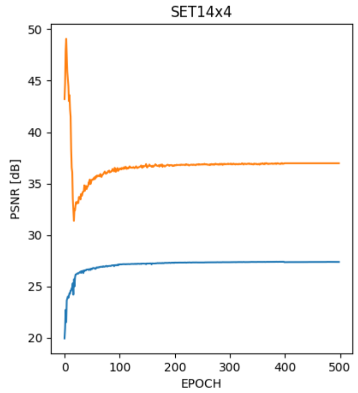
General Improvements - Loss Function

$$L = \alpha L_{TASK} + \beta L_{LATENT}$$

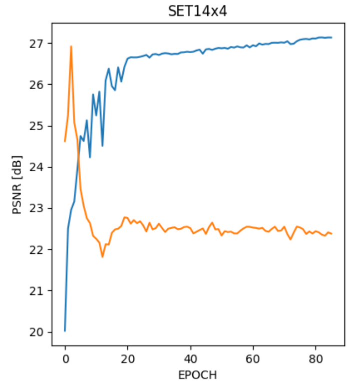
$$L_{TASK} = L1(X_{GT}, X_{SHR})$$

$$L_{LATENT} = \begin{cases} y & \text{if} ||y|| \geq \epsilon \\ 0.0 & \text{otherwise} \end{cases}$$

$$y = L1(X_{GD}, X_{SLR})$$



$\varepsilon = 0$

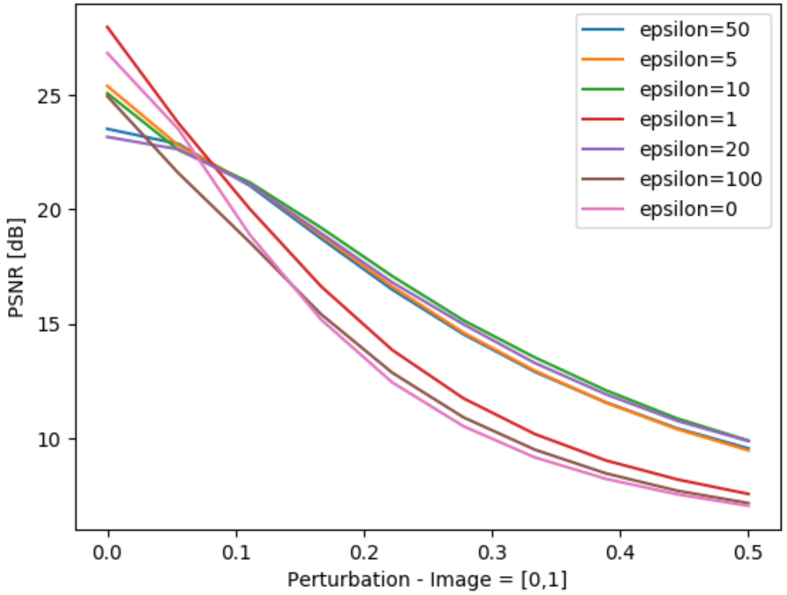


$\varepsilon = 50$

SHR (blue), SLR (orange)

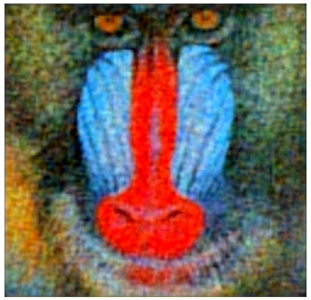
ε-ball avoids overfitting to trivial solution $X_{SLR} = X_{GD}$

Improving Robustness against Perturbation



Validation Data: Set14, Scale: 4

$\varepsilon = 0, \sigma = 0.11$



$\varepsilon = 10, \sigma = 0.11$

σ = std dev. of white Gaussian noise

$\varepsilon = 0, \sigma = 0.22$

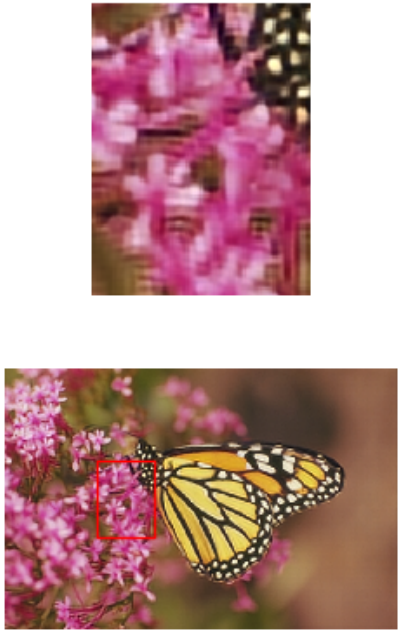
$\varepsilon = 10, \sigma = 0.22$

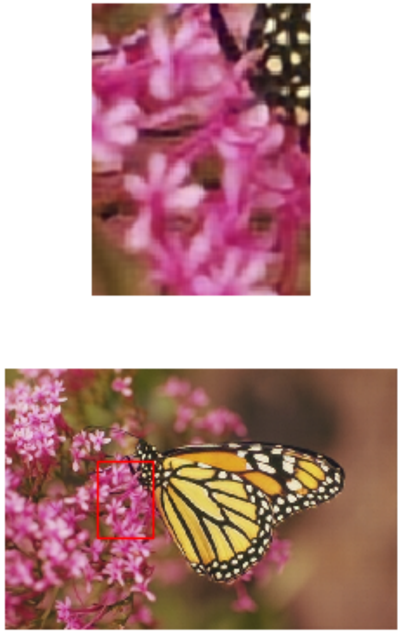
ε = Radius of $l1$-ball of L_{LATENT}

SINGLE IMAGE SUPER RESOLUTION (SISR)

SISR – Upscaling Performance (scale = 4)

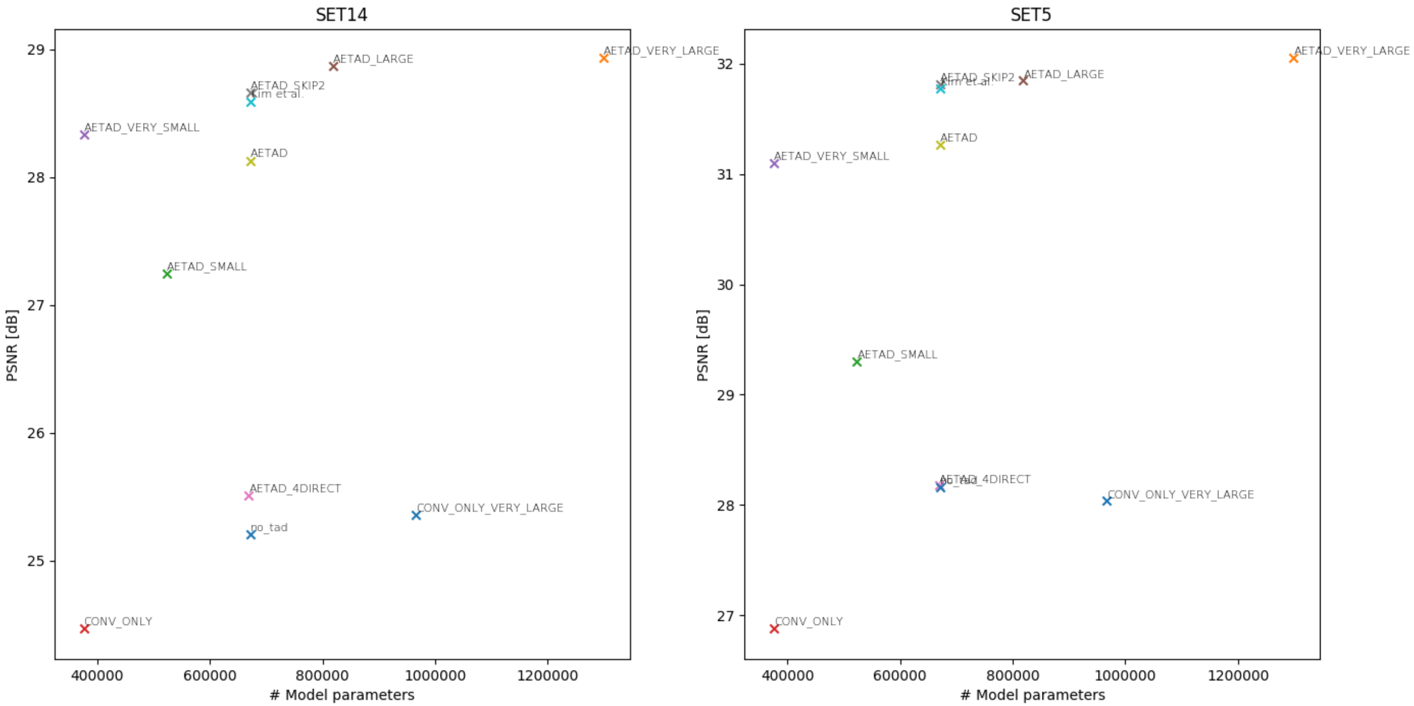
Upscale Bilinear
Interpolated Image
(LR -> SHRB)



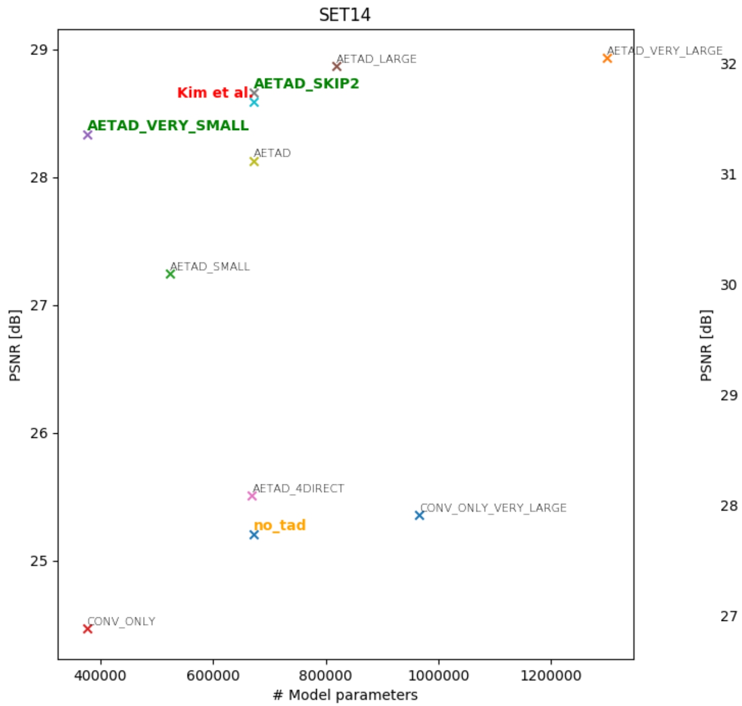


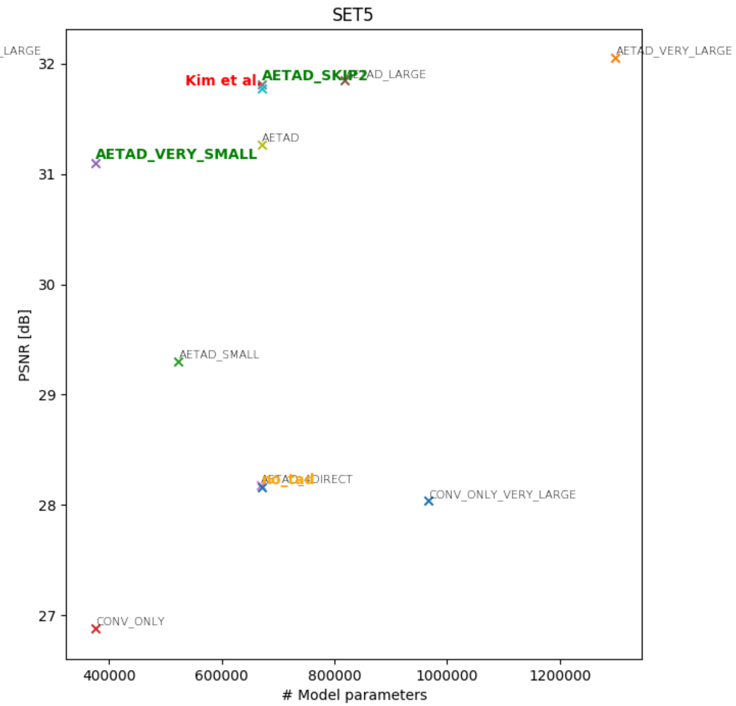
Upscaling
Task-Aware Image
(SLR -> SHRT)

SISR - Complexity/PSNR/Robustness Tradeoff



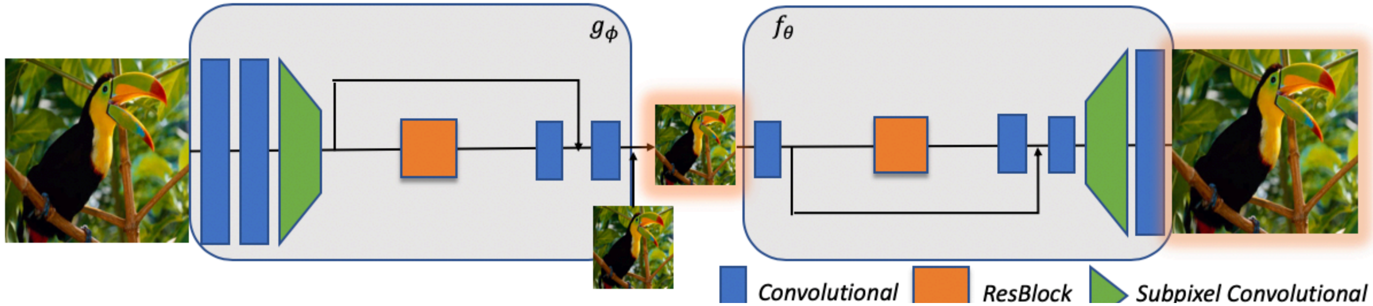
SISR - Complexity/PSNR/Robustness Tradeoff



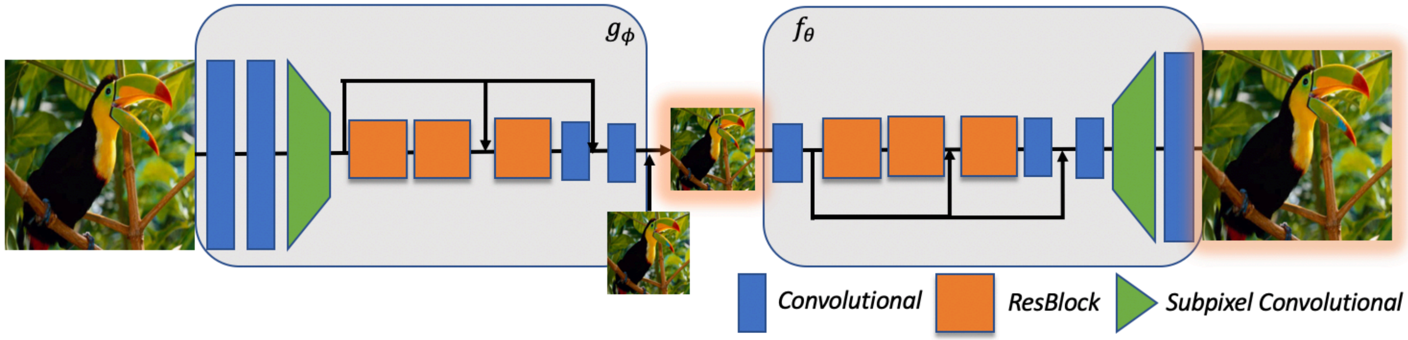


SISR - Final Model

AETAD_VERY_SMALL



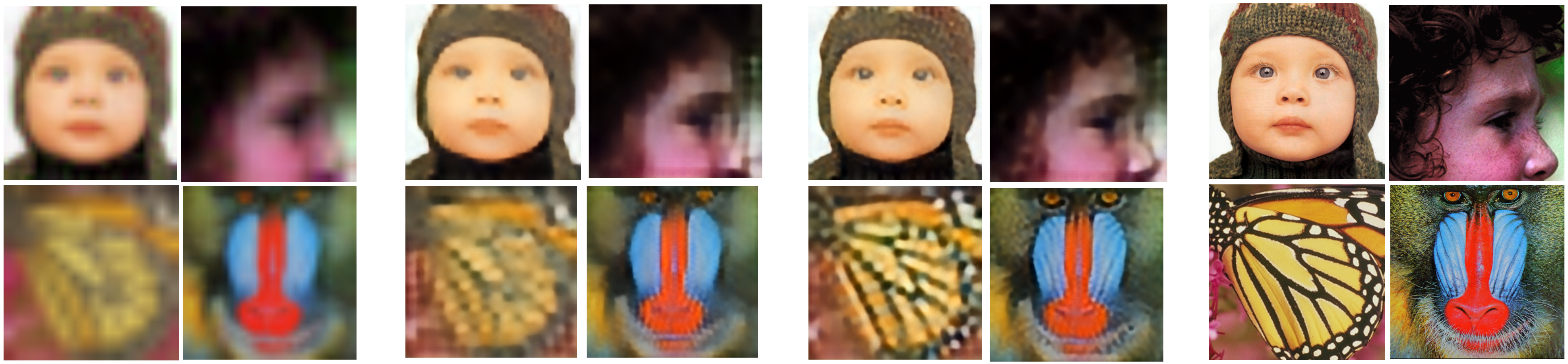
AETAD_SKIP2



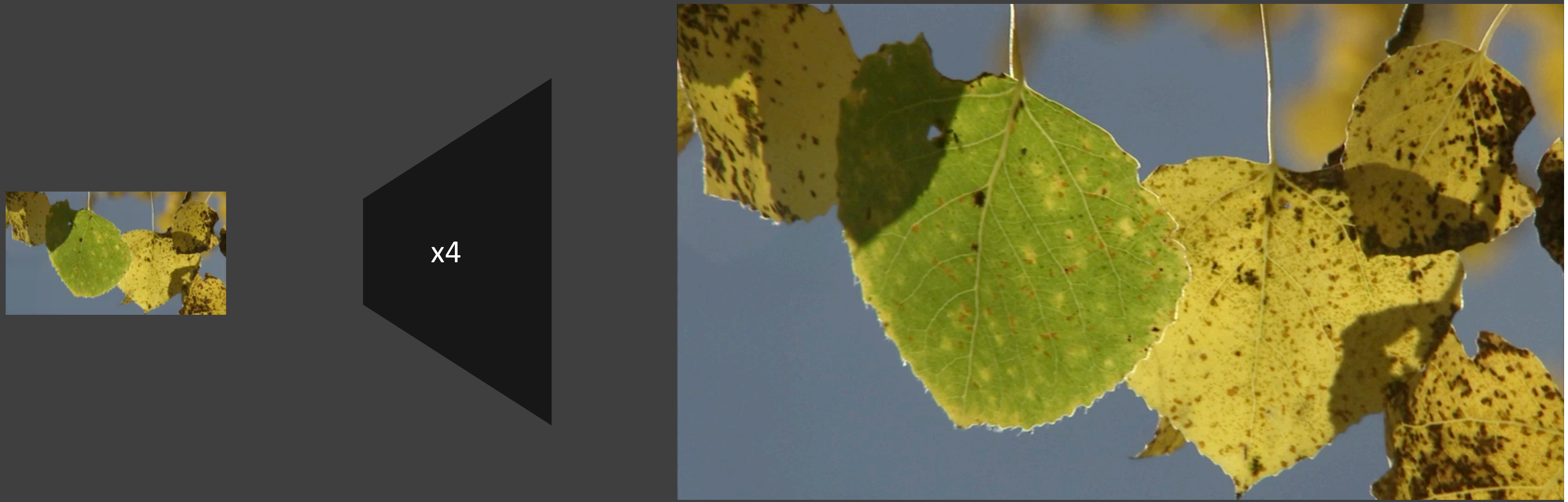
SISR - Final Model Performance

dataset	PSNR (Kim et al.)	PSNR (*aetad_skip2*)	# model param. gain (*aetad_skip2*)	PSNR (*aetad_very_small*)	# model param. gain (*aetad_very_small*)
SET5	31.81	31.814	0 %	31.102	-44 %
SET14	28.63	28.665	0 %	28.334	-44 %
URBAN100	26.63	24.156	0 %	23.084	-44 %
BSDS100	28.51	28.601	0 %	25.719	-44 %

SISR – Upscaling Performance (scale = 16)



LR Bicubic HR TAD HR Groundtruth HR

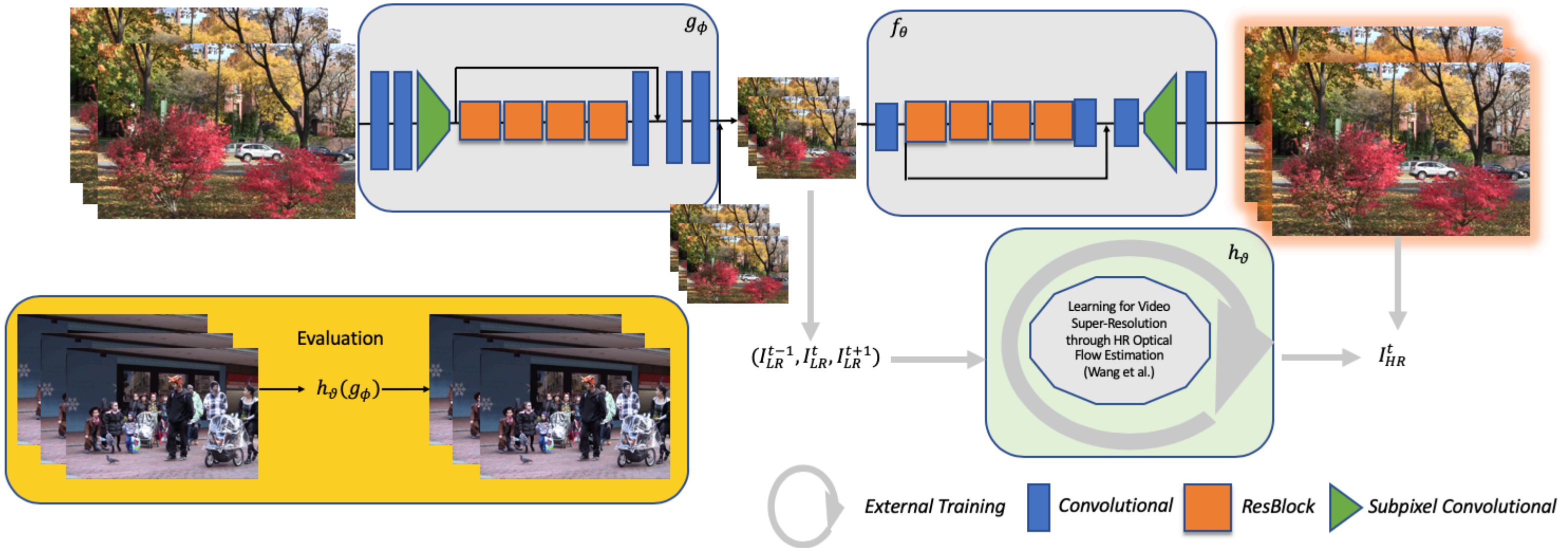


VIDEO SUPER RESOLUTION (VSR)

VSR – External Approach

Training:

1. TAD basic training (DIV2K)
2. TAD video training (NTIAASPEN)
3. SOFVSR training (NTIAASPEN_SR)



VSR - Performance

scale	dataset	Bicubic	non task aware SOFVSR	task aware SOFVSR
x2	WALK	24.224	26.215	30.433
x2	FOLIAGE	21.771	25.122	26.620
x4	CALENDAR	18.537	18.573	19.190
x4	CITY	23.483	24.191	24.677

VSR – Upscaling Performance (scale = 2)

Bilinear

TAD

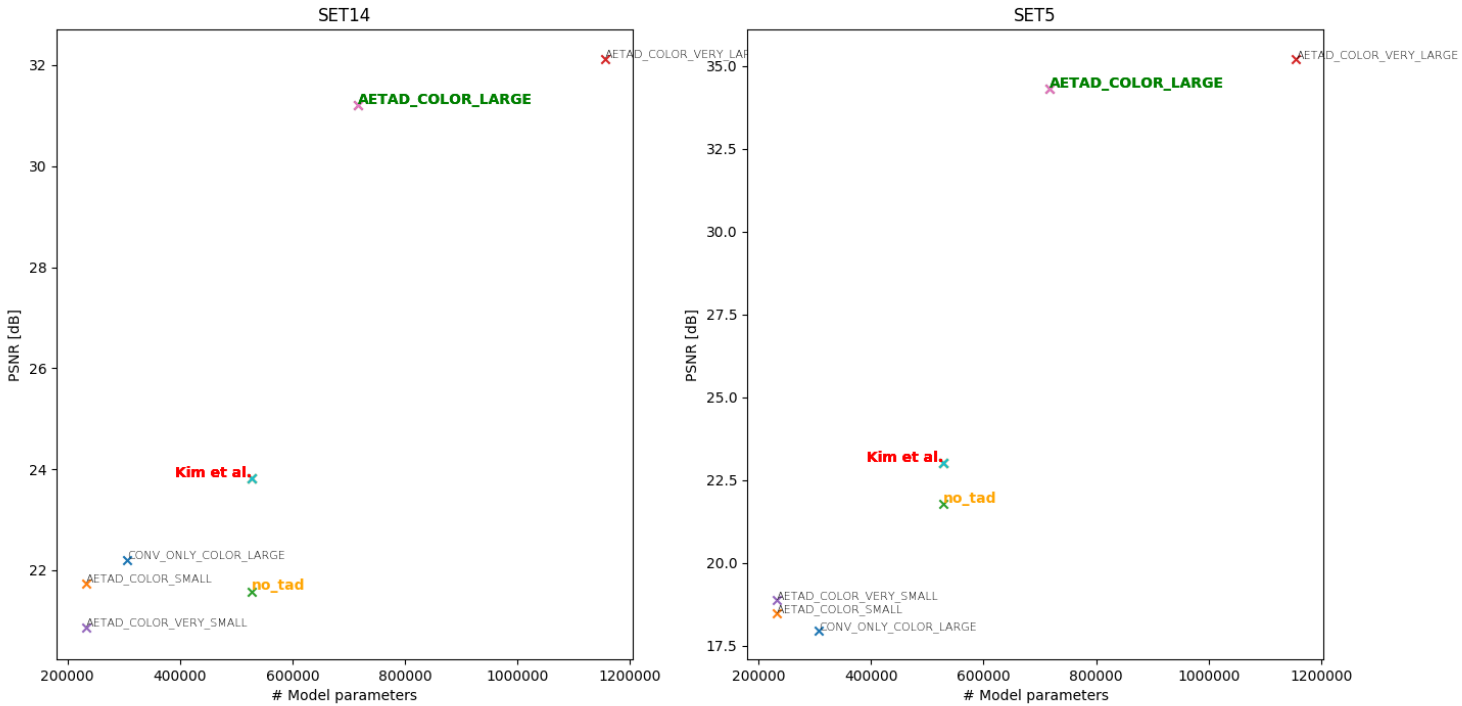
VSR – Upscaling Performance (scale = 4)

Bilinear

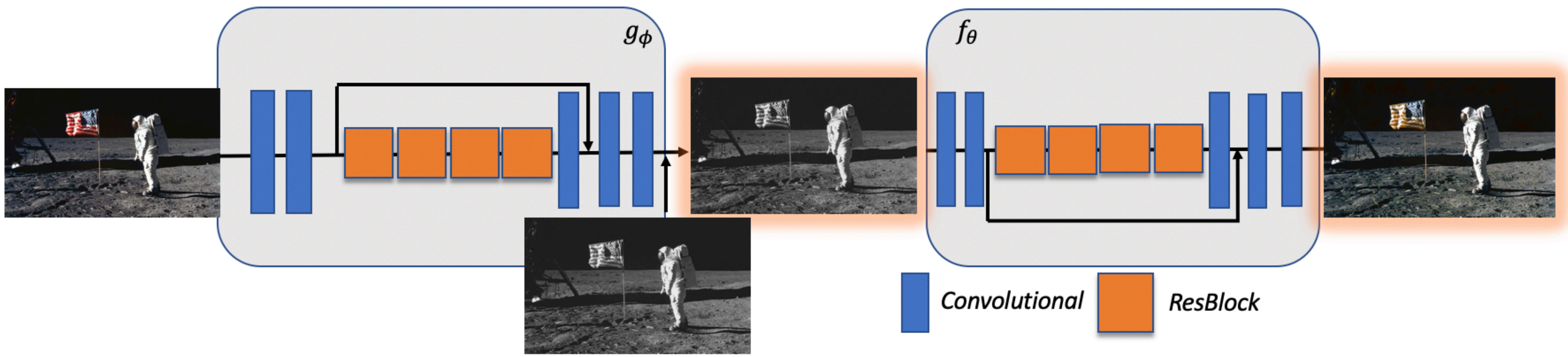
TAD

IMAGE COLORIZATION (IC)

IC - Complexity/PSNR/Robustness Tradeoff



IC - Final Model



AETAD_COLOR_LARGE

IC - Final Model Performance

dataset	PSNR (Kim et al.)	PSNR (*aetad_color_large*)	# model param. gain
SET5	22.56	34.416	35.7 %
SET14	23.78	31.262	35.7 %
URBAN100	33.68	33.604	35.7 %
BSDS100	36.14	36.786	35.7 %

IC – Colorization Performance

Gray

TAD Color

Groundtruth Color

Does TAD qualitatively outperform "classic" approaches ?


Gray


NO-TAD Color

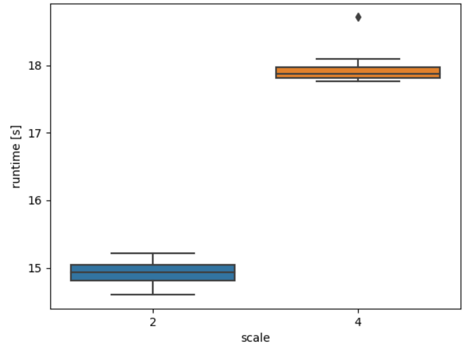

TAD Color


Groundtruth Color

TAD can restore images that are impossible to restore from grayscale !

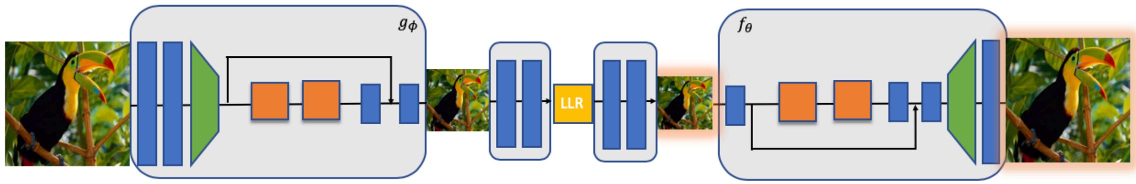
Conclusions

- ✓ Increase robustness of TAD
- ✓ Extended TAD approach to video super-resolution
- ✓ Improved model runtime – performance tradeoff for single image super-resolution
- ✓ Improved model runtime – performance tradeoff for image colorization



Future Work

- Apply purposed video super-resolution approach to other models
- Deploy video super-resolution approach to even larger scale factors (> 16)
- Improve compression rate of TAD



References

- Heewon Kim, Myungsub Choi, Bee Lim, and Kyoung Mu Lee. Task-aware image downscaling. In *ECCV*, 2018.
- Longguang Wang, Yulan Guo, Zaiping Lin, Xinpu Deng, and Wei An. Learning for video super-resolution through HR optical flow estimation. *CoRR*, abs/1809.08573, 2018.
- Radu Timofte, Shuhang Gu, Jiqing Wu, Luc Van Gool, Lei Zhang, Ming-Hsuan Yang, Muham- mad Haris, et al. Ntire 2018 challenge on single image super-resolution: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018.
- Dong C., Loy C.C., He K., and Tang X. Learning a deep convolutional network for image super resolution. *ECCV 2014*, 2014.
- S.Schulter,C.Leistner,andH.Bischof.Fast and accurate image upscaling with super-resolution forests. pages 3791–3799, June 2015.

Thank you for Your Attention !

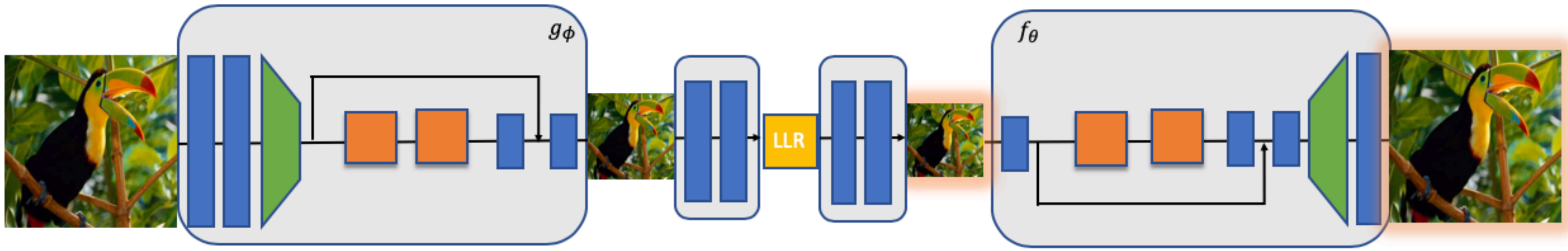
Code available on https://github.com/simon-schaefer/tar

Backup - Normalized Epsilon Ball

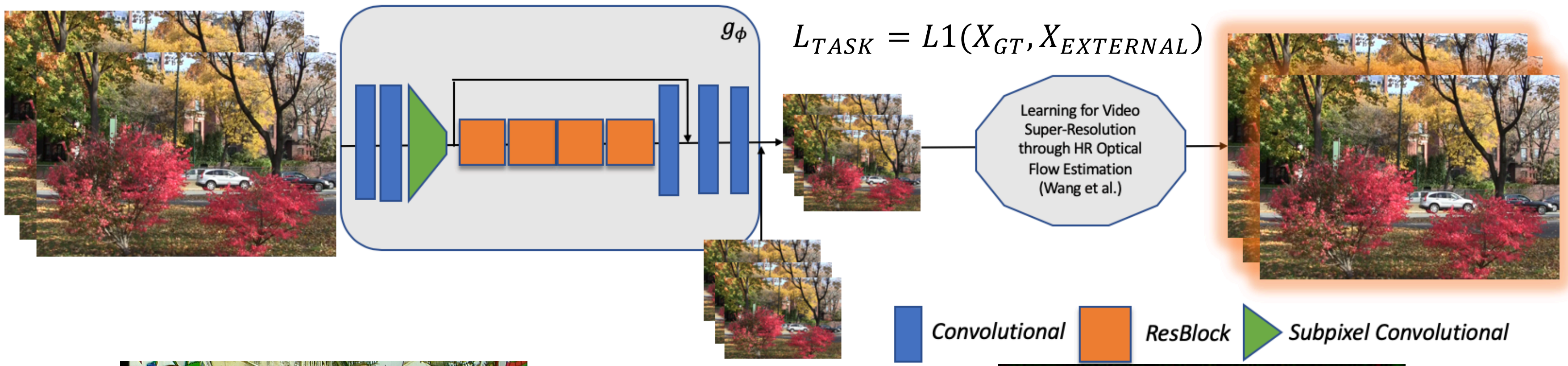
$$normalized\ loss = \frac{L1(X_{GD}, X_{SLR})}{(p_{max} - p_{min}) * p_{num}} * 100$$

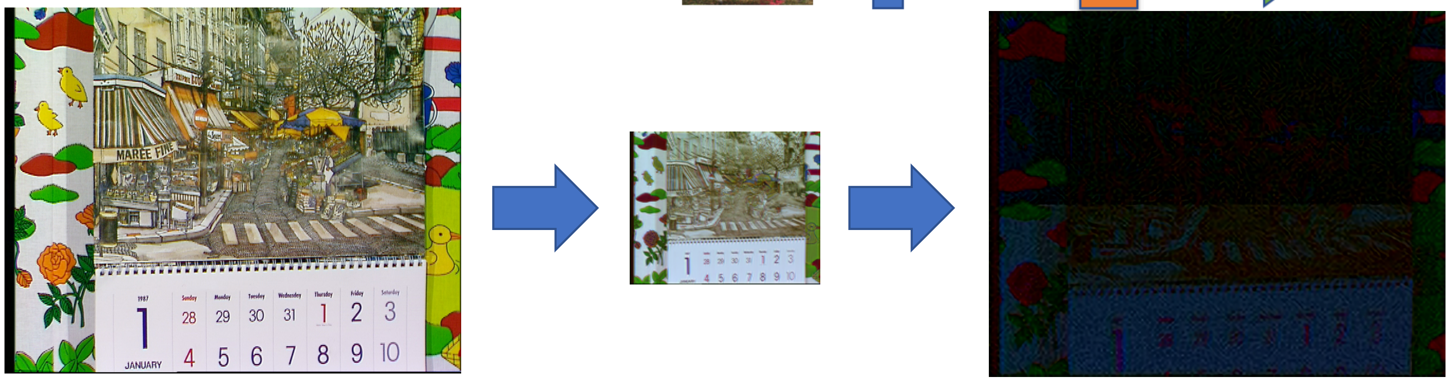
p_{max} = maximal pixel value
p_{min} = minimal pixel value
p_{num} = number of pixels

Backup - Improving Compression Rate



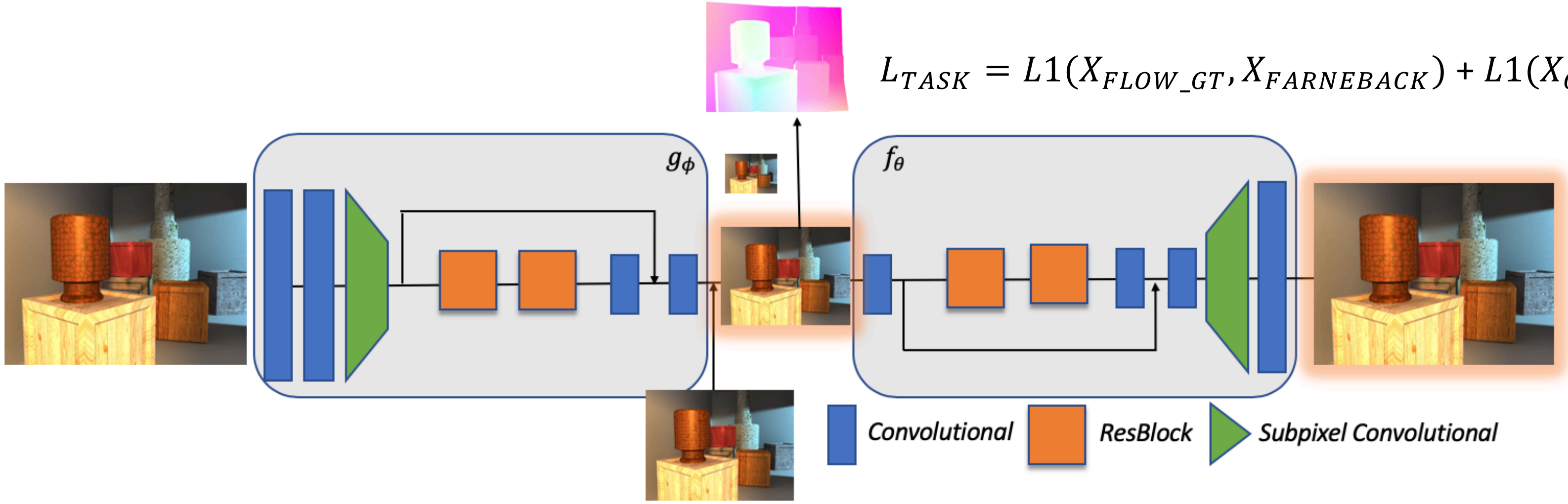
Backup – VSR Direct Approach



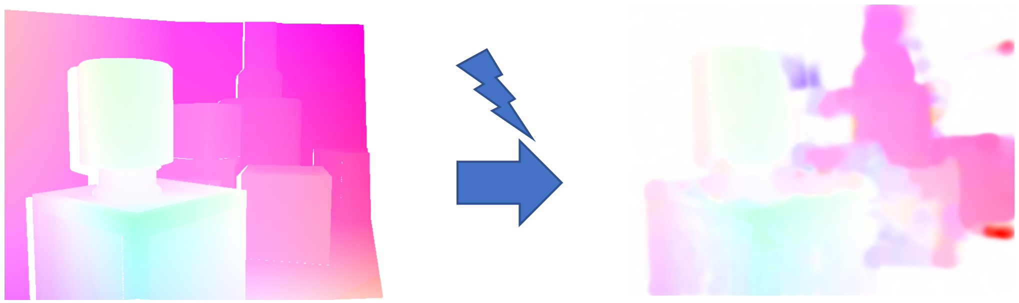


Backup – VSR Flow Approach

$$L_{TASK} = L1(X_{FLOW\_GT}, X_{FARNEBACK}) + L1(X_{GT}, X_{SHR})$$



UCL Optical
Flow Dataset



Backup – SOFVSR Selection

Rank	Method	PSNR	SSIM	MOVIE	Paper Title	Year	Paper	Code
1	**VSR-DUF**	27.31	0.832		Deep Video Super-Resolution Network Using Dynamic Upsampling Filters Without Explicit Motion Compensation	2018		
2	**RBPN/6-PF**	27.12	0.818		Recurrent Back-Projection Network for Video Super-Resolution	2019		
3	**FRVSR**	26.69	0.822		Frame-Recurrent Video Super-Resolution	2018		
4	**SOF-VSR**	26.01	0.771	4.32	Learning for Video Super-Resolution through HR Optical Flow Estimation	2018		
5	**DRDVSR**	25.88	0.774		Detail-revealing Deep Video Super-resolution	2017		
6	**DBPN**	25.37	0.737		Deep Back-Projection Networks For Super-Resolution	2018		
7	**VESPCN**	25.35	0.7557	5.82	Real-Time Video Super-Resolution with Spatio-Temporal Networks and Motion Compensation	2016		
8	**ESPCN**	25.06	0.7394	6.54	Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network	2016		
9	**SRCNN**	24.68	0.7158	6.90	Image Super-Resolution Using Deep Convolutional Networks	2014		
10	**BRCN**	24.43	0.662		Bidirectional Recurrent Convolutional Networks for Multi-Frame Super-Resolution	2015		

According to website PapersWithCode comparing stated results of papers in standard computer vision problems

Performance and Closeness to state-of-the-art

Availability of PyTorch implementation open-source

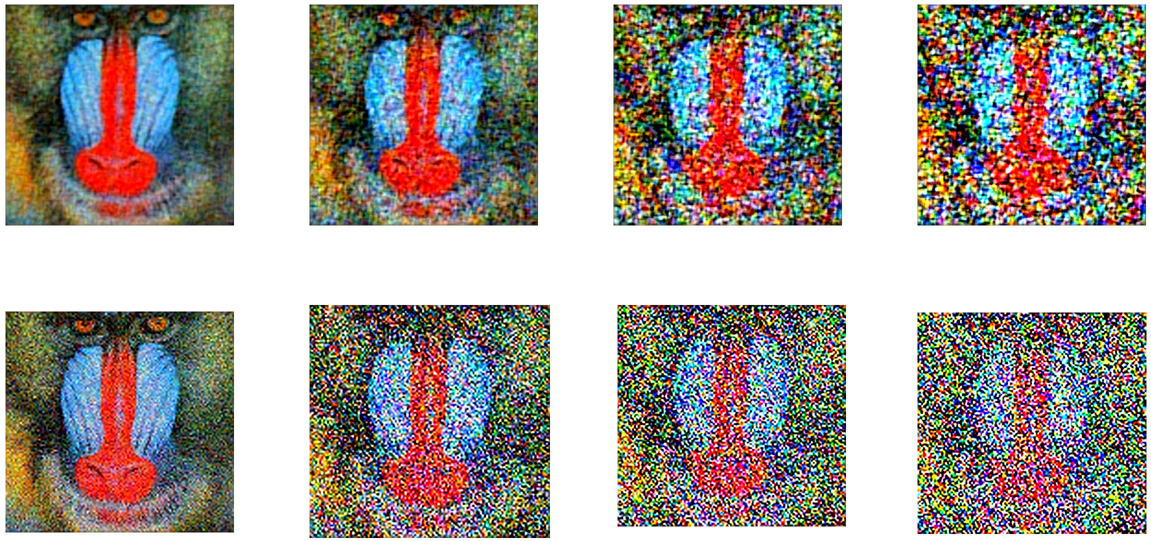
Backup – Limitations of TAD

- Increased computations compared to trival downscaling
- Larger vulnerability against perturbations

Backup – Large Perturbation

$\epsilon = 10$

$\epsilon = 0$



$e = 0.11$ $e = 0.22$ $e = 0.39$ $e = 0.50$

Backup – SISR PSNR for scale = 16

dataset	PSNR (bicubic)	PSNR (*aetad_skip2*)
SET5	19.137	20.982
SET14	18.669	20.038
VDIV2K	20.920	23.209