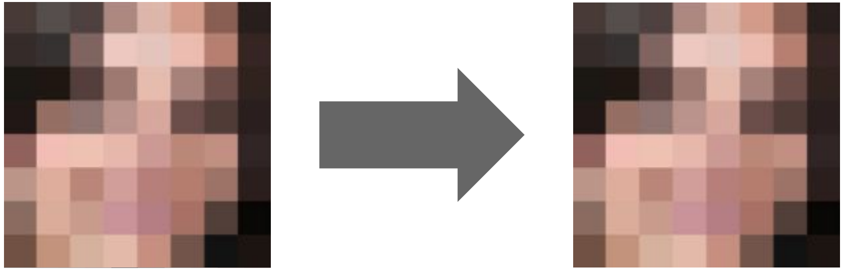
Video Super Resolution

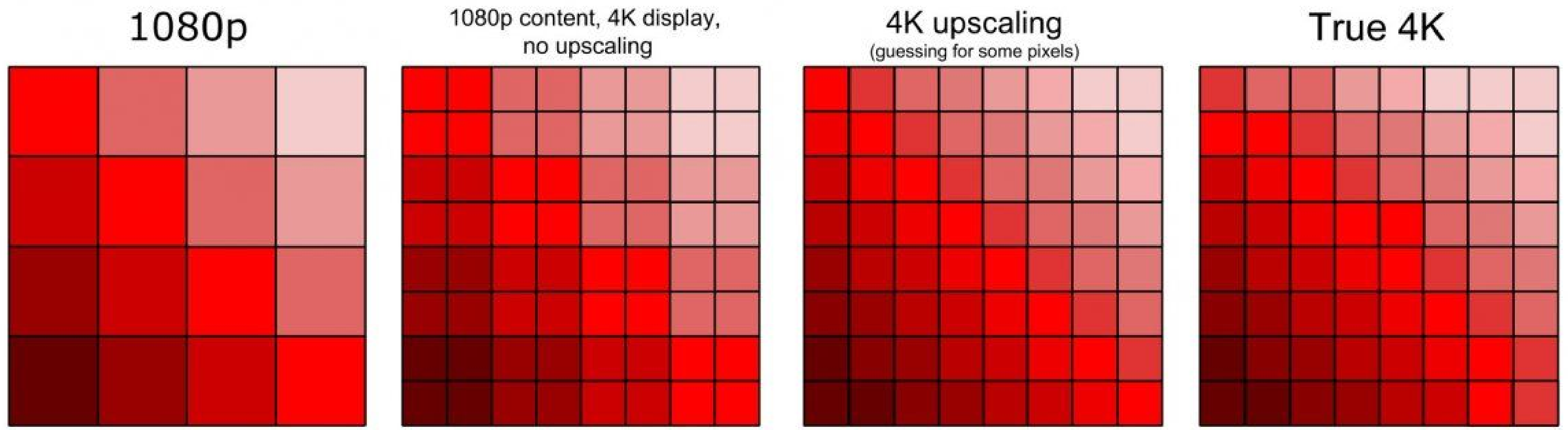
Enhancing videos with the power of Deep Learning


CAN YOU
ENHANCE THAT

No, you can not enhance that

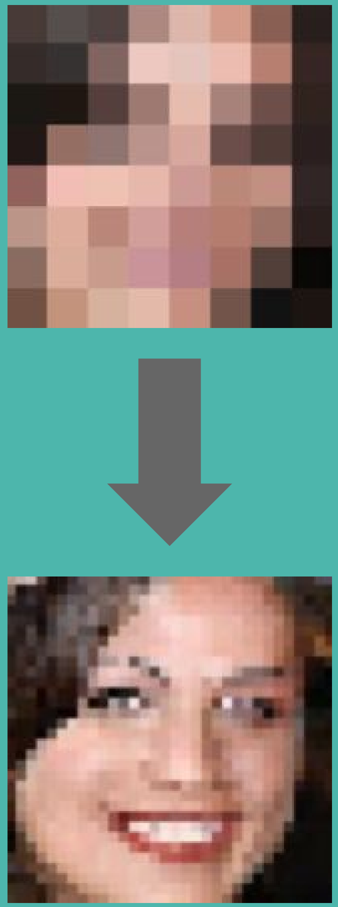


What about upscaling?

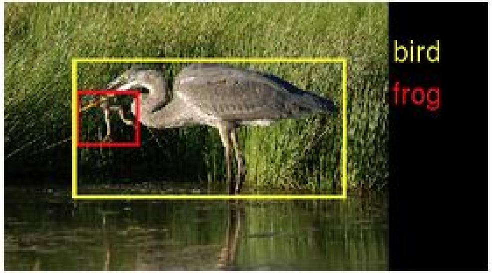


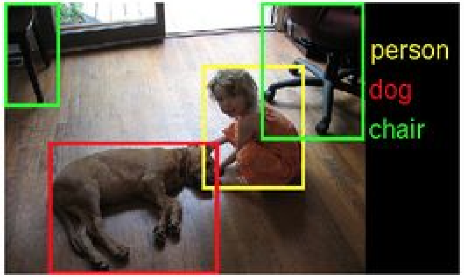
“Super Resolution”

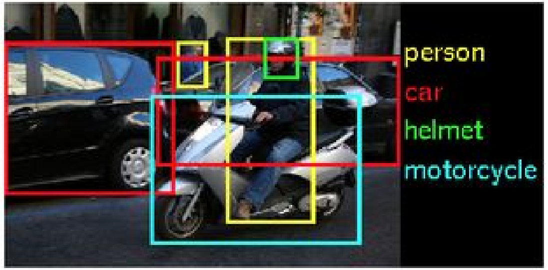
Intelligent Upscaling



Pattern Discovery with Deep Learning





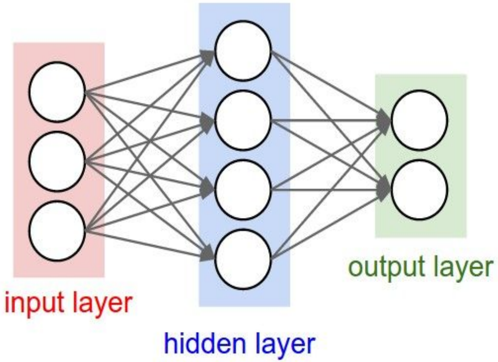


Deep Learning Background

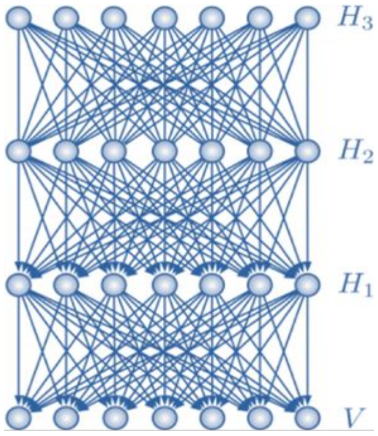
Machine Learning

X ➡ y

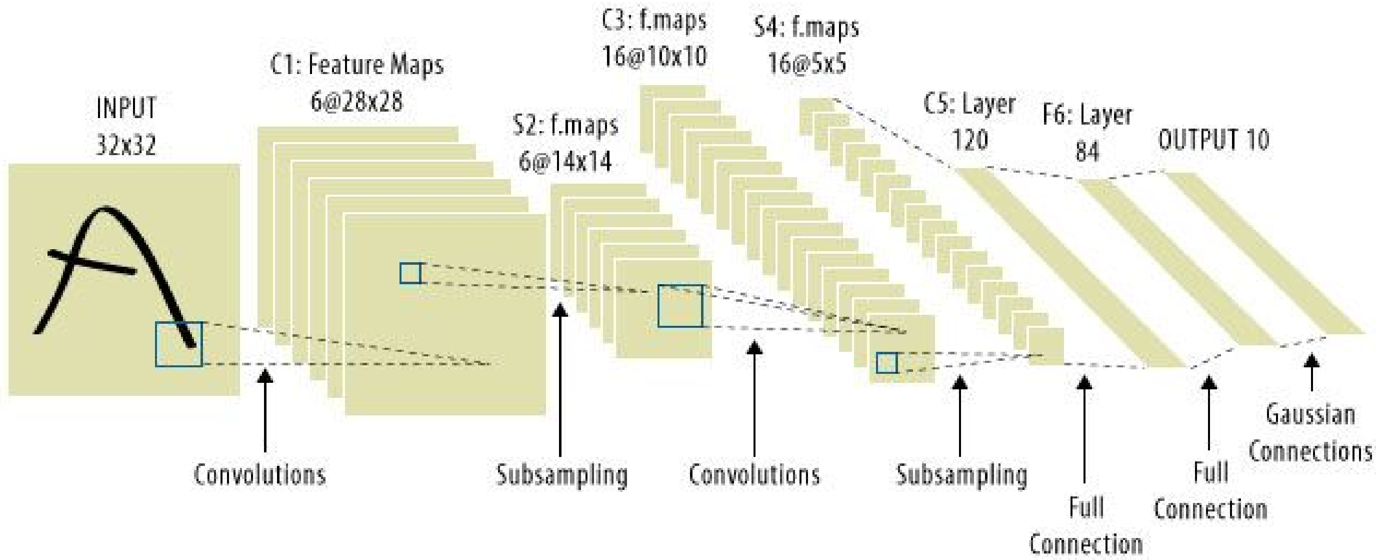
Neural Networks



Deep Learning



Convolutional Neural Networks

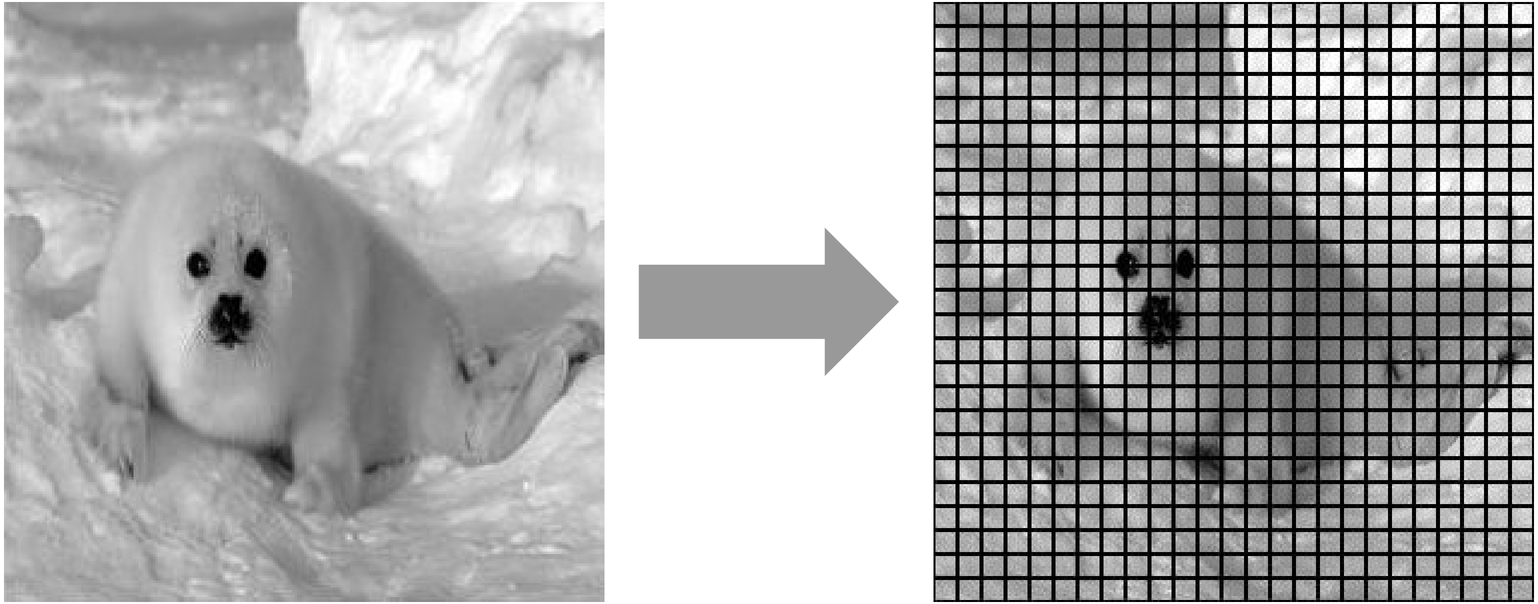


Memory constraints

1. Dataset doesn’t fit into memory
2. Neural network weights don’t fit into memory

Patching

Allows us to support dynamically-sized images!



Patch Parameters

Dynamically generated

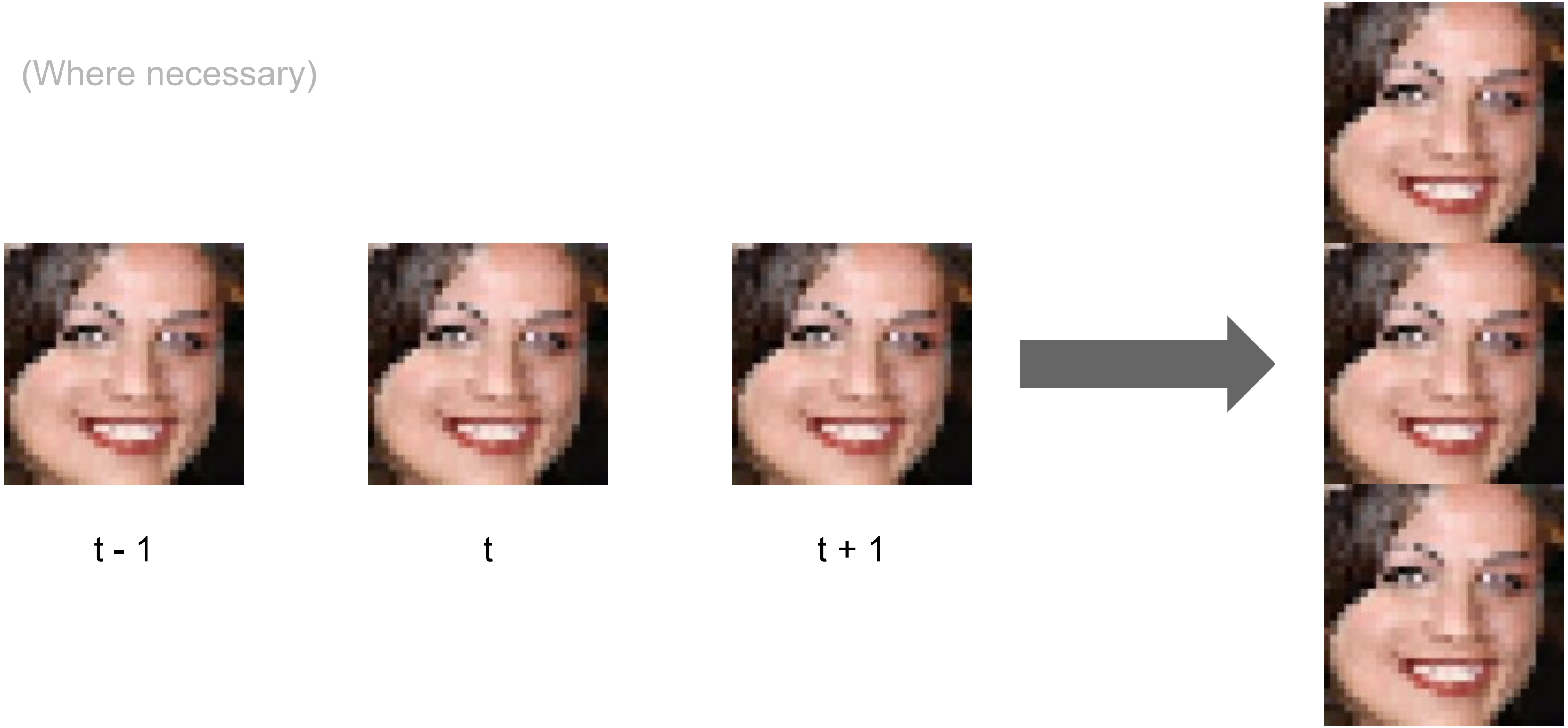
32x32 pixel patches

Random extraction / no stride

50 per frame

Patch concatenation

(Where necessary)



Training pipeline

Downscale Patch → Upscale Patch → Neural Network

Size reduction

Our ground truth patches are reduced from their original size of 32x32 to 16x16

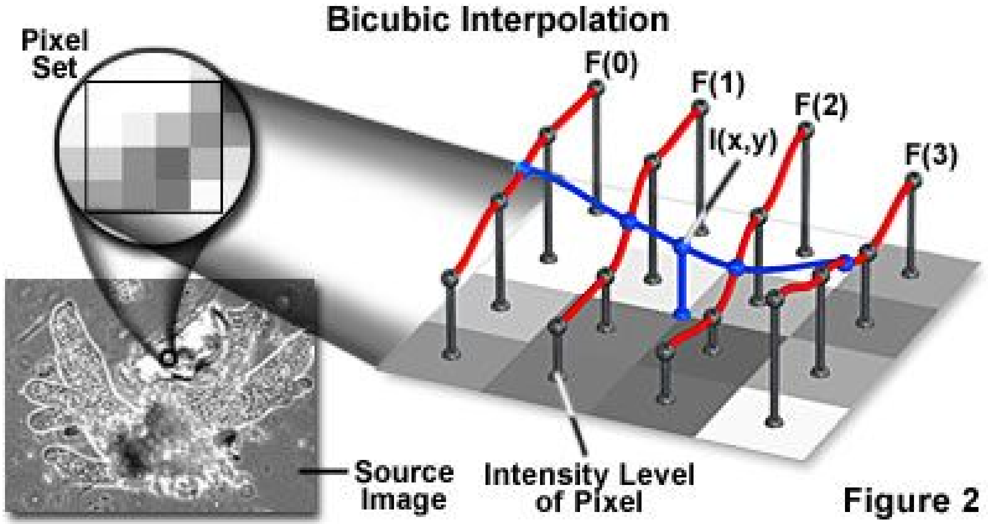
Size increase

Our reduced samples are brought back to 32x32 size by interpolation methods

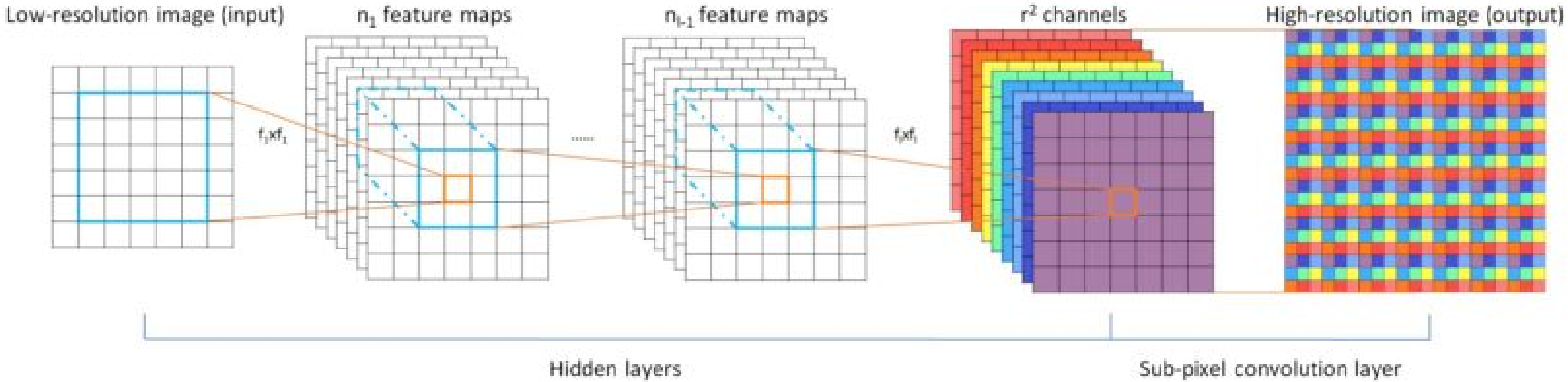
Evaluation

The neural network “upscales” the 32x32 patch to a higher quality 32x32 patch

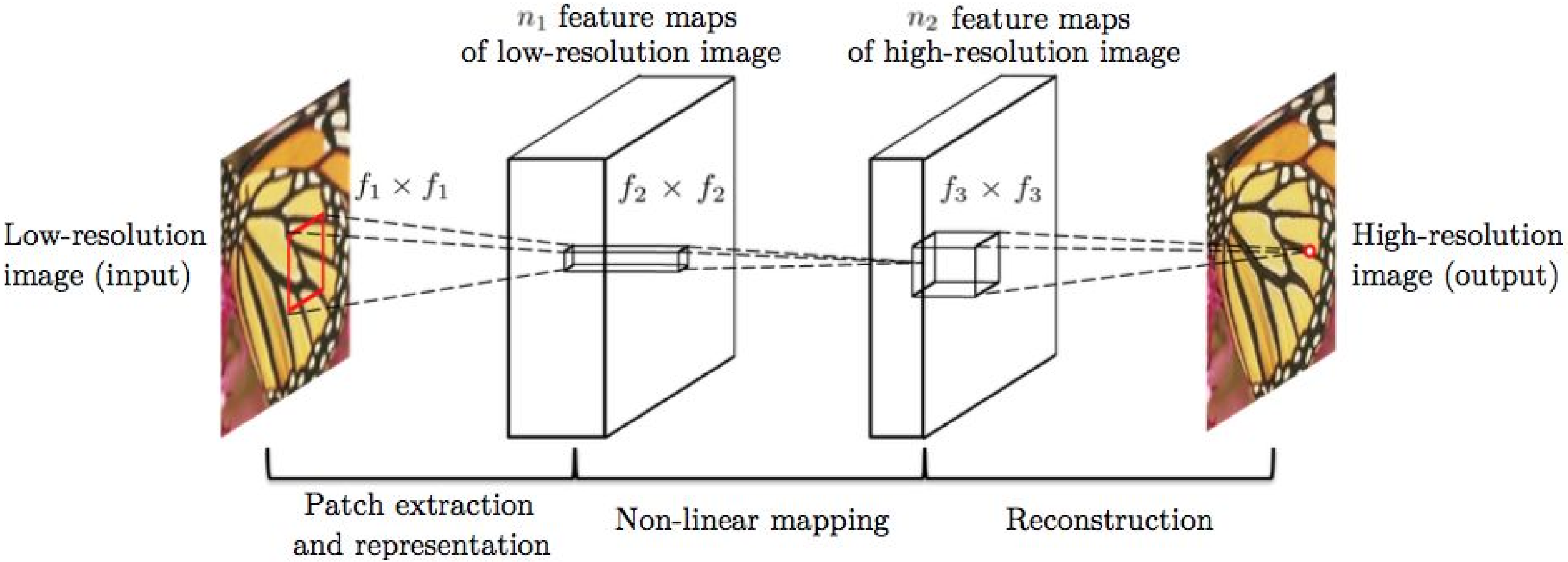
Traditional upscaling



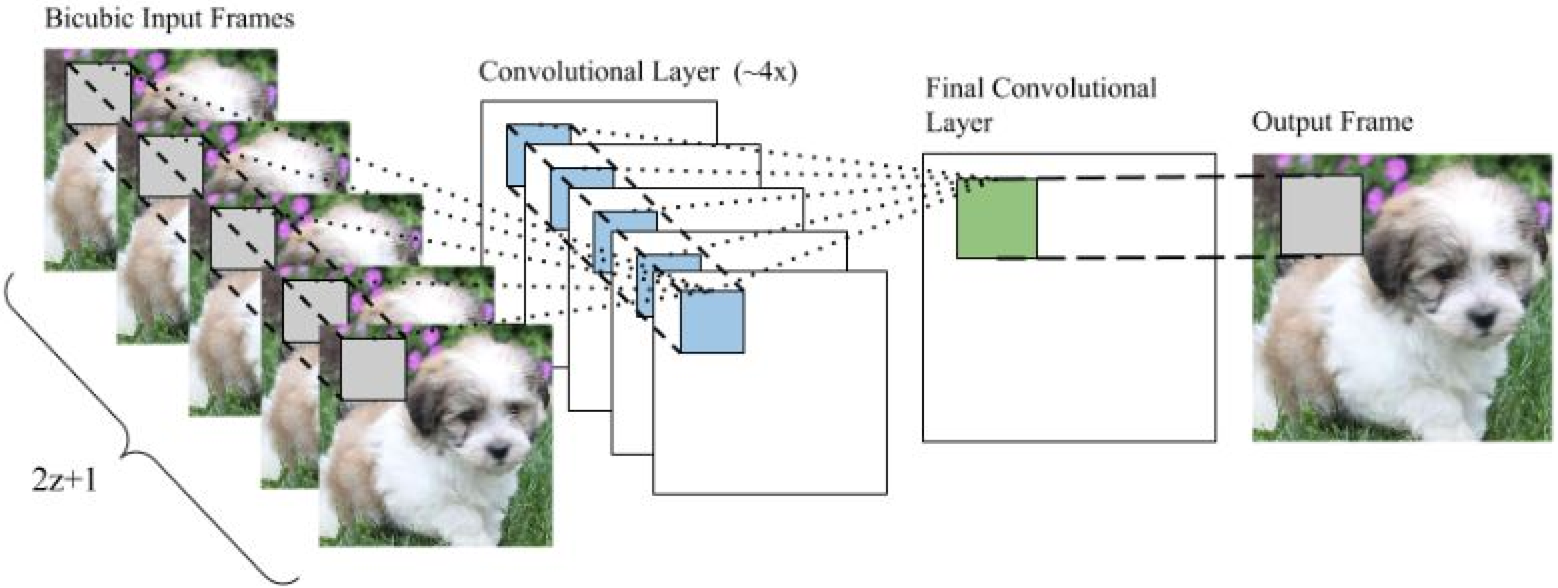
True upscaling

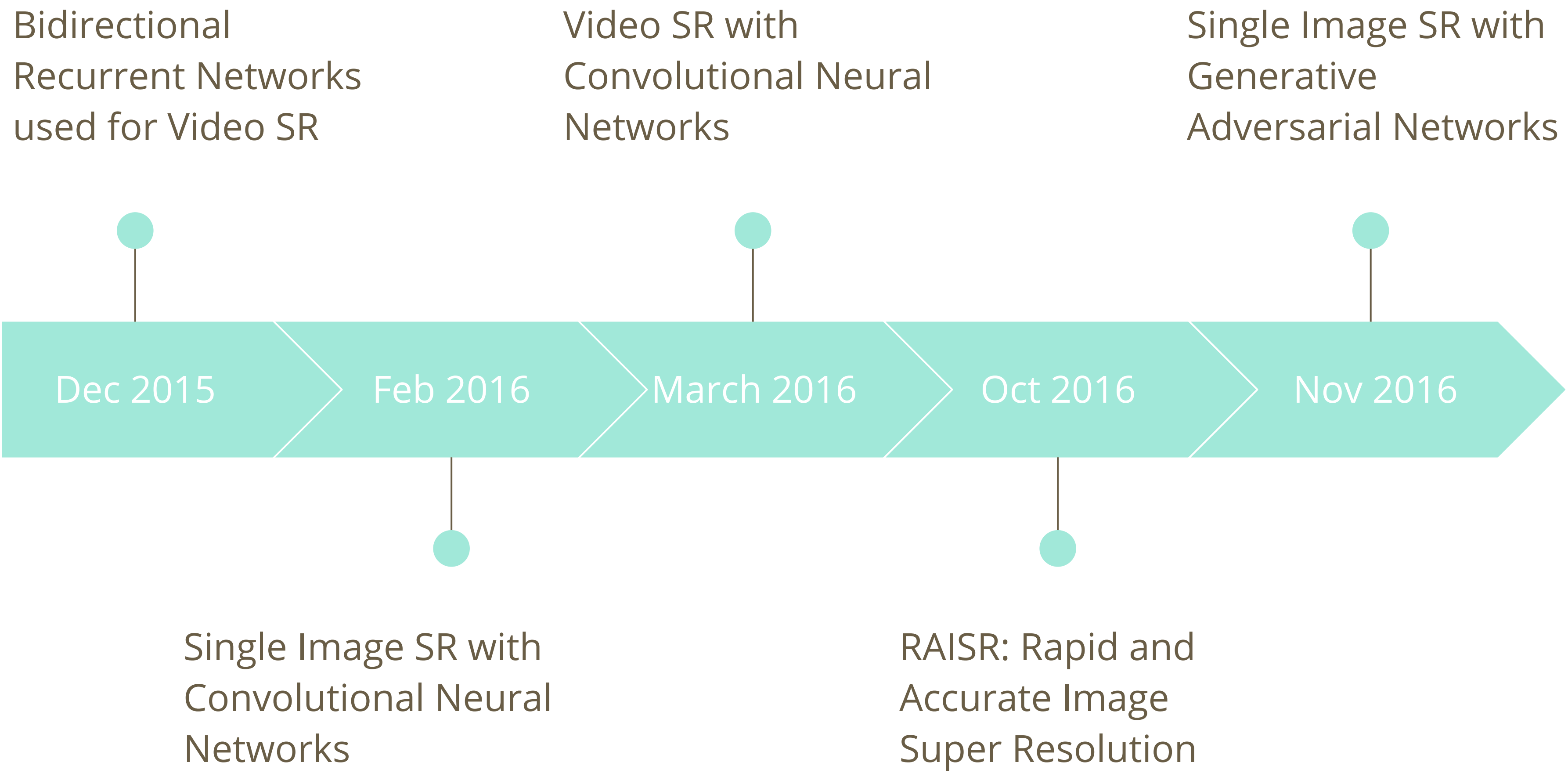


Single Image Super Resolution

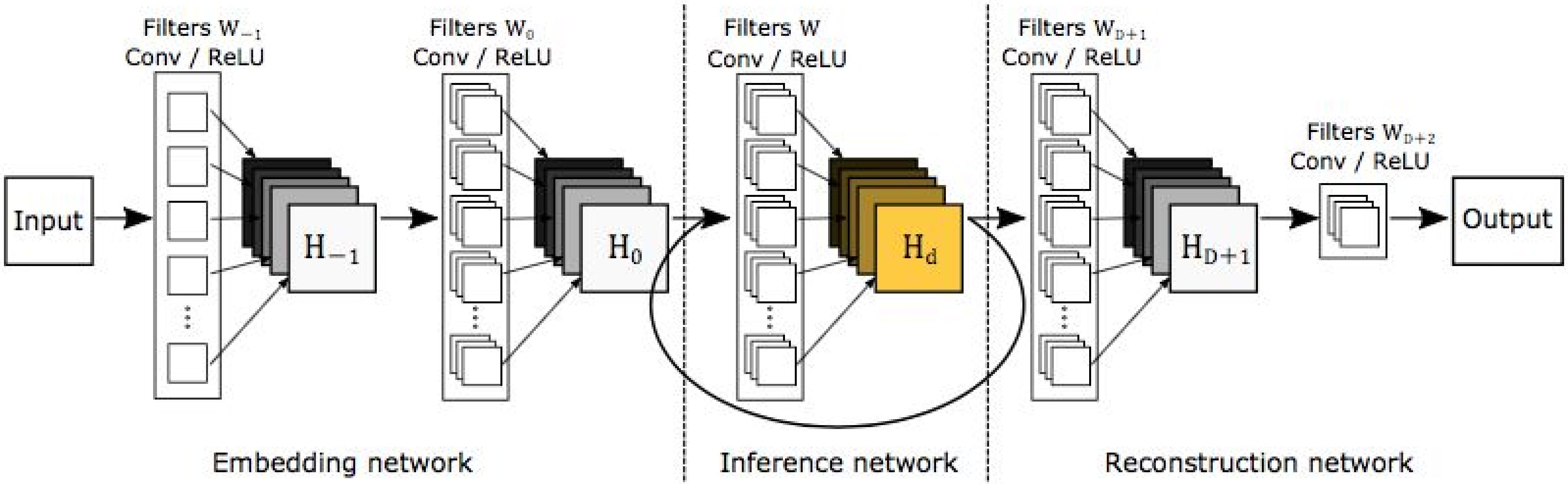


Video Super Resolution

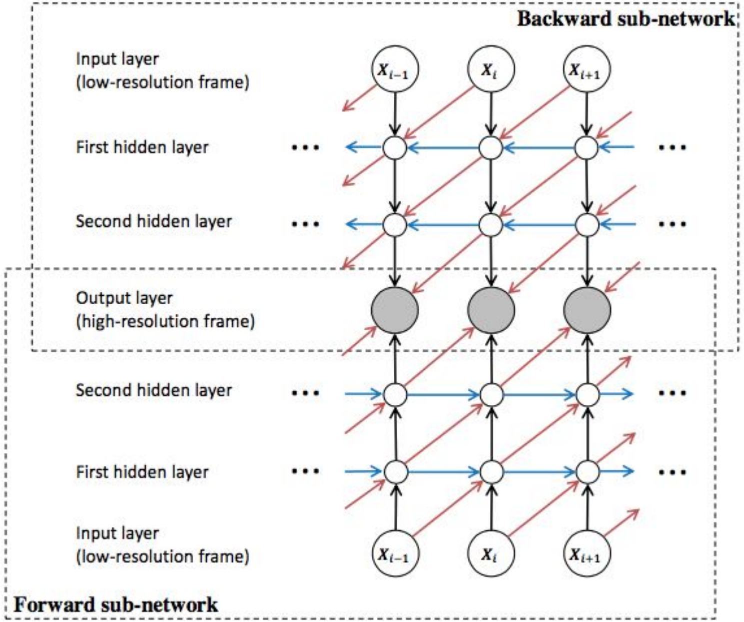



Dec 2015
Bidirectional Recurrent Networks used for Video SR
Feb 2016
Single Image SR with Convolutional Neural Networks
March 2016
Video SR with Convolutional Neural Networks
Oct 2016
RAISR: Rapid and Accurate Image Super Resolution
Nov 2016
Single Image SR with Generative Adversarial Networks

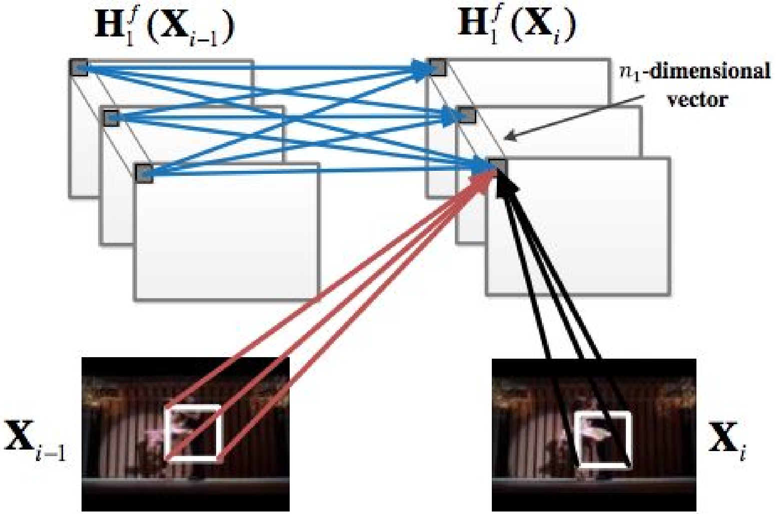
Recursive Neural Network



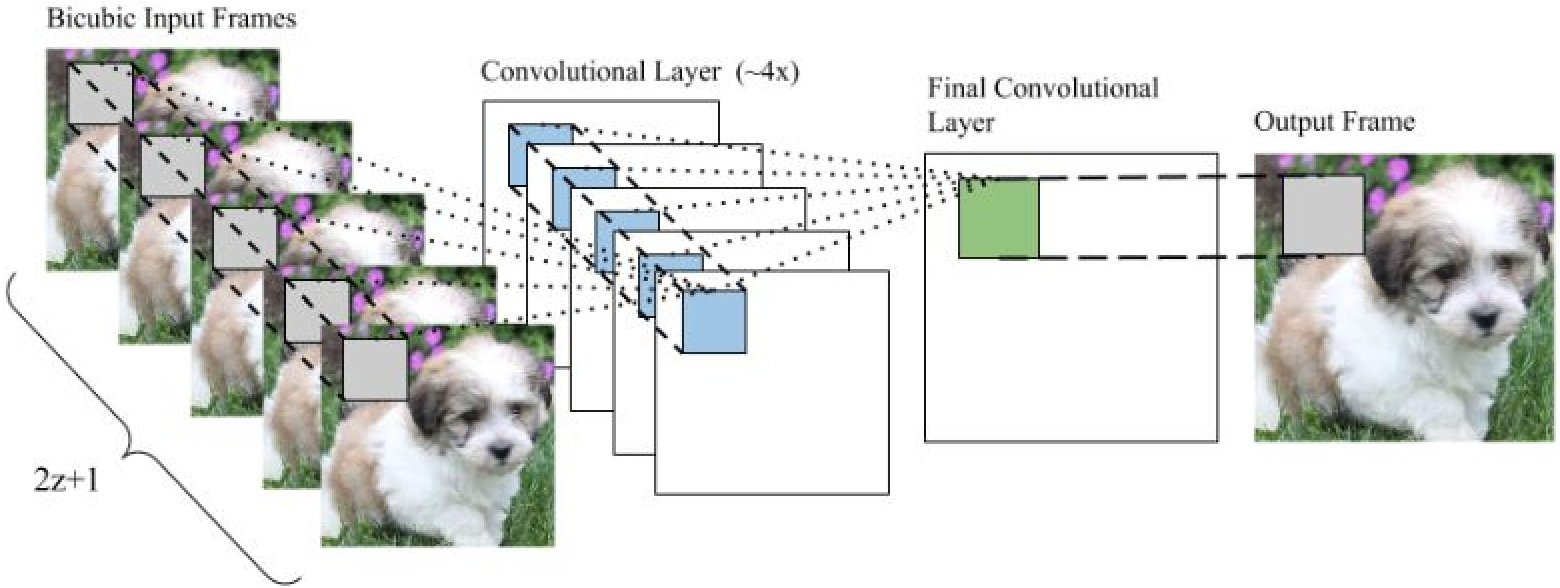
Recurrent Neural Network



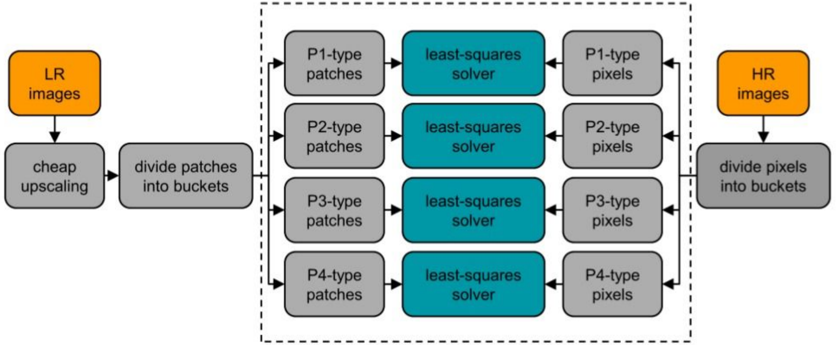
: Feedforward convolution : Recurrent convolution : Conditional convolution



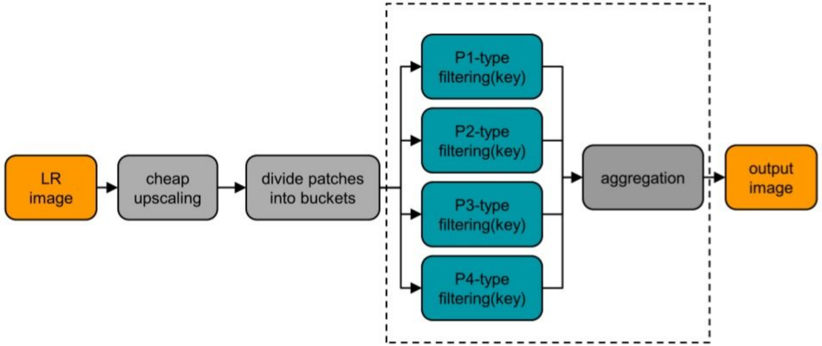
Convolutional Neural Network



Neural Network Alternatives

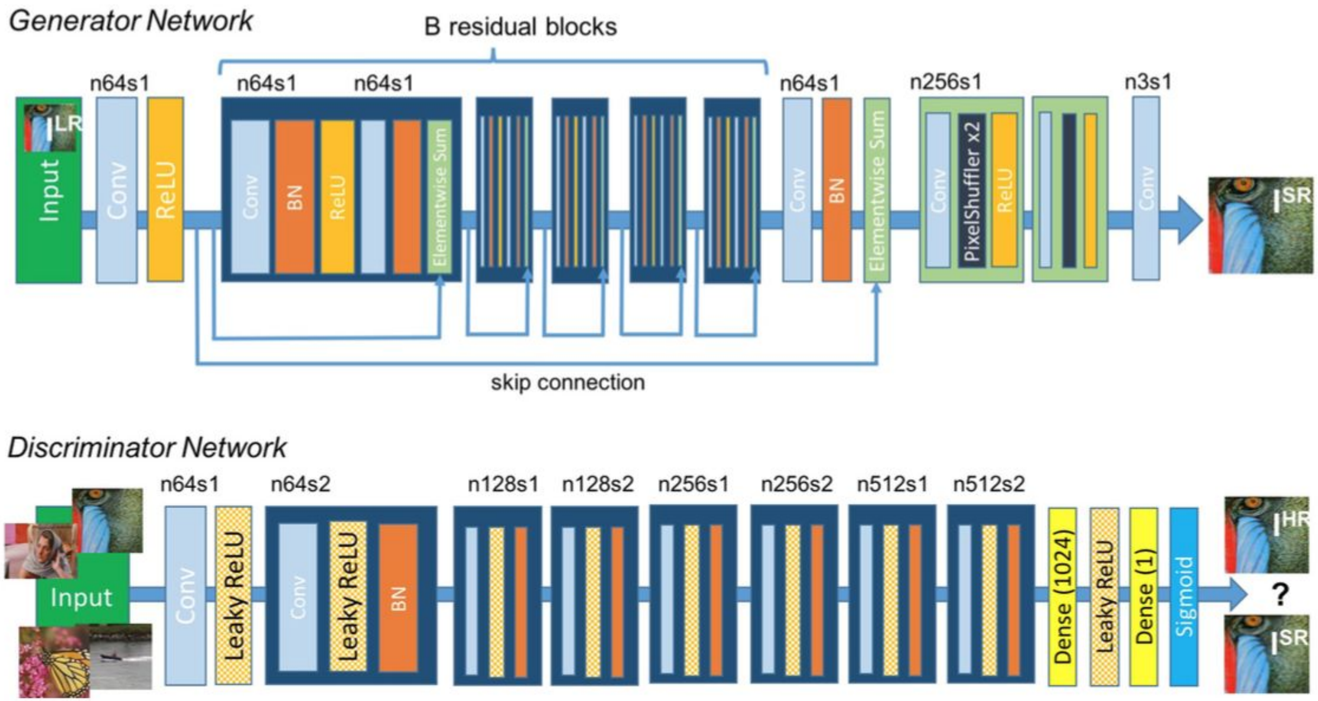


(a) Learning Stage



(b) Upscaling Stage

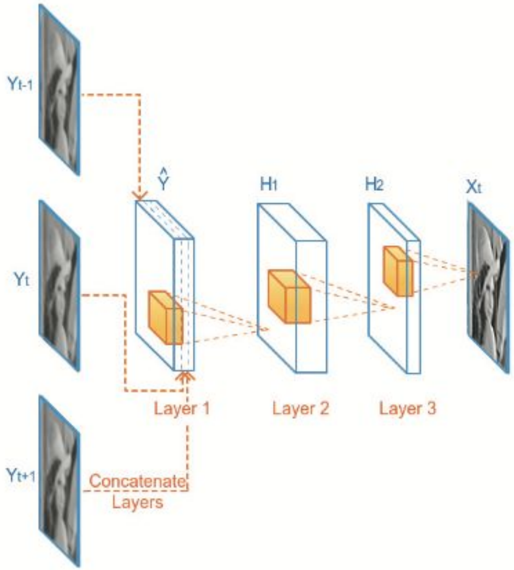
Generative Adversarial Network



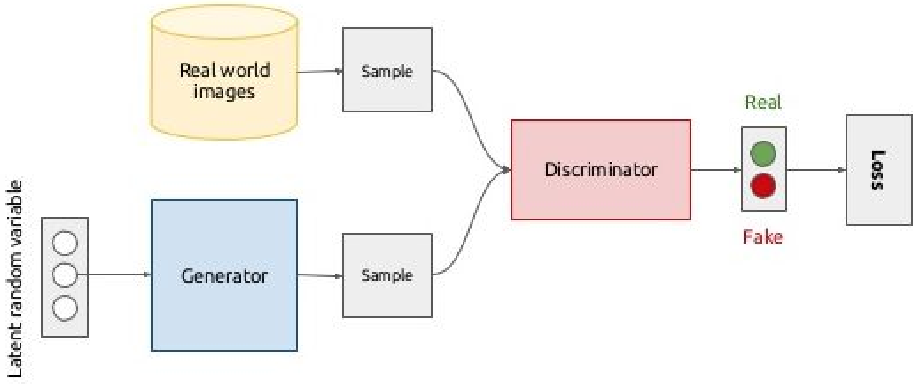
Experiments

Candidate architectures

Convolutional Neural Network



Generative Adversarial Neural Network



Deep learning framework

- Easy to learn
- Supports Tensorflow or Theano
- Efficient implementation of layers
- Enables DRY models

Convolutional Neural Network

3 layers

64 9x9 filters

32 5x5 filters

1 5x5 filter

ReLU activations

MSE loss

Adam SGD

Generative Adversarial Network

Wrapped model

Sub-pixel convolution

2 layers in B residual blocks

64 3x3 kernels

Batch normalization

Skip connections

ReLU activations

Adam SGD

Generative Adversarial Network

Wrapped model

8 layers

64->512 3x3 kernels

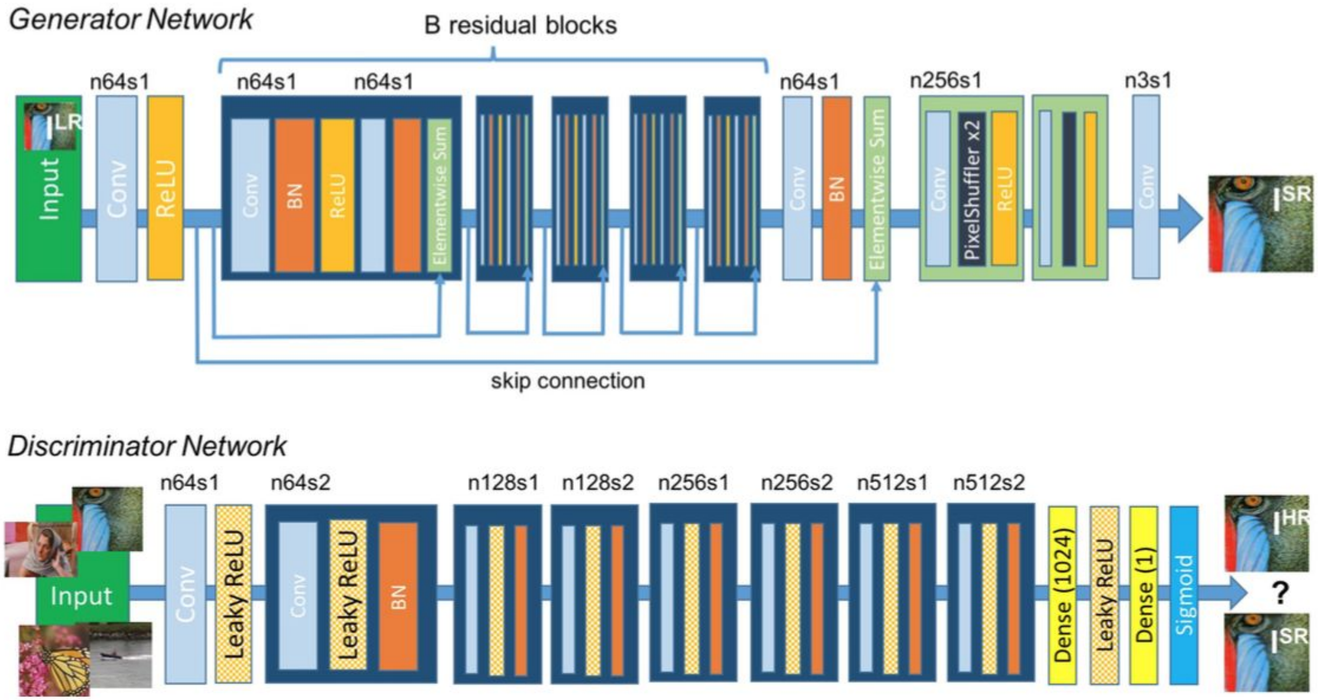
Stride of 1 or 2

Batch normalization

Sigmoid activation

ReLU activations

Adam SGD


Generator Network
B residual blocks
n64s1
n64s1
n64s1
n64s1
n256s1
n3s1
Input
Conv
ReLU
Conv
BN
ReLU
Elementwise Sum
Conv
BN
Elementwise Sum
Conv
PixelShuffler x2
ReLU
Conv
skip connection
Discriminator Network
n64s1
n64s2
n128s1
n128s2
n256s1
n256s2
n512s1
n512s2
Input
Conv
Leaky ReLU
Conv
Leaky ReLU
BN
Dense (1024)
Leaky ReLU
Dense (1)
Sigmoid
?

Results

MYANMAR	Bicubic	VSRNet	CNN	GAN
PSNR (2x)	34.59	38.48	38.37	**39.54**
PSNR (3x)	31.59	34.42	34.48	**35.21**
PSNR (4x)	29.53	**31.85**	31.14	30.98

Evaluation measurement

1. PSNR
2. SSIM
3. MSE / Content Loss
4. Perceptual Loss
5. Generative Loss
6. Mean Opinion Score (MOS)

Evaluation dataset

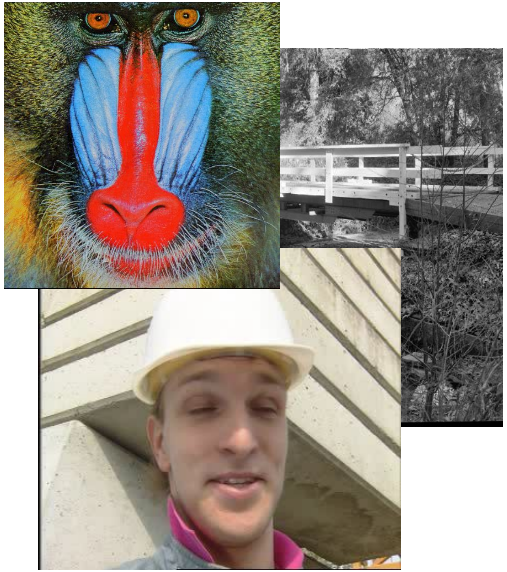
- Current results may not generalize
- Papers all compare against few models
- Papers all use different evaluation sets

Color Palette

Proportional Palette	Hex Color	Area	*Closest Color Name* (Base Color)
	#cccfd7	23.1 %	*Link Water* (Blue)
	#9a754b	18.2 %	*Pale Brown* (Brown)
	#54412d	17.9 %	*Deep Bronze* (Brown)
	#2e2119	13.4 %	*Wood Bark* (Brown)
	#73593b	9.3 %	*Shingle Fawn* (Brown)
	#1a0609	6.5 %	*Black* (Black)
	#4e554e	4.3 %	*Battleship Grey* (Grey)
	#8b7e41	3.7 %	*Highball* (Green)
	#8c9192	2.0 %	*Submarine* (Blue)
	#4e171e	1.6 %	*Bordeaux* (Red)

Set 5

Set 14



BSD100

Challenges

1. Large project scope
2. Extremely large training times
3. Three primary model approaches
4. Lack of universal evaluation dataset

Future work

1. Explore potential universal evaluation datasets
2. Further explore the potential of GAN networks
3. Develop and measure universal speed metrics
4. Consider RNN approaches

Thanks!