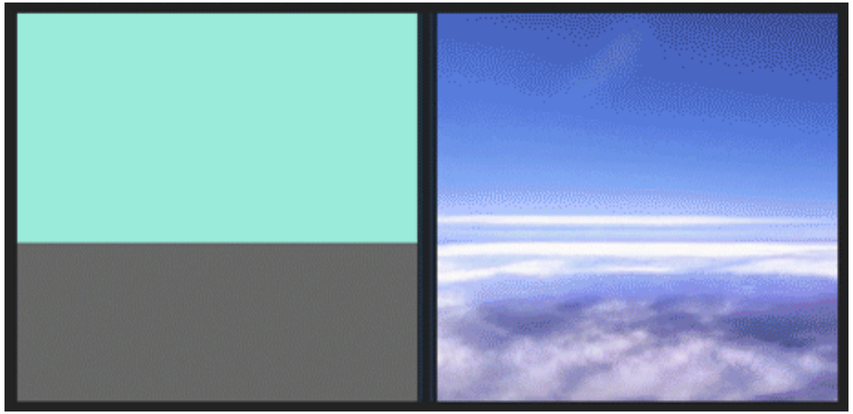
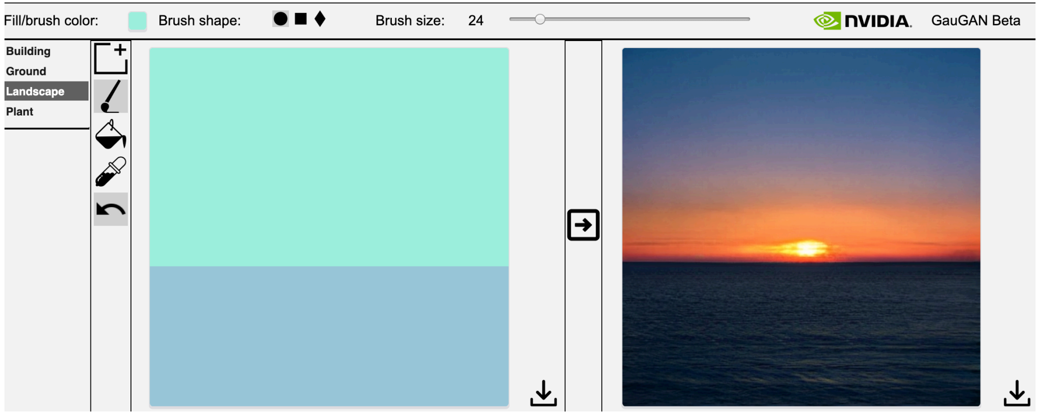
DERIN OGRENMEYLE RESIM SENTEZLEME

Aysegul Dundar

IMAGE SYNTHESIS

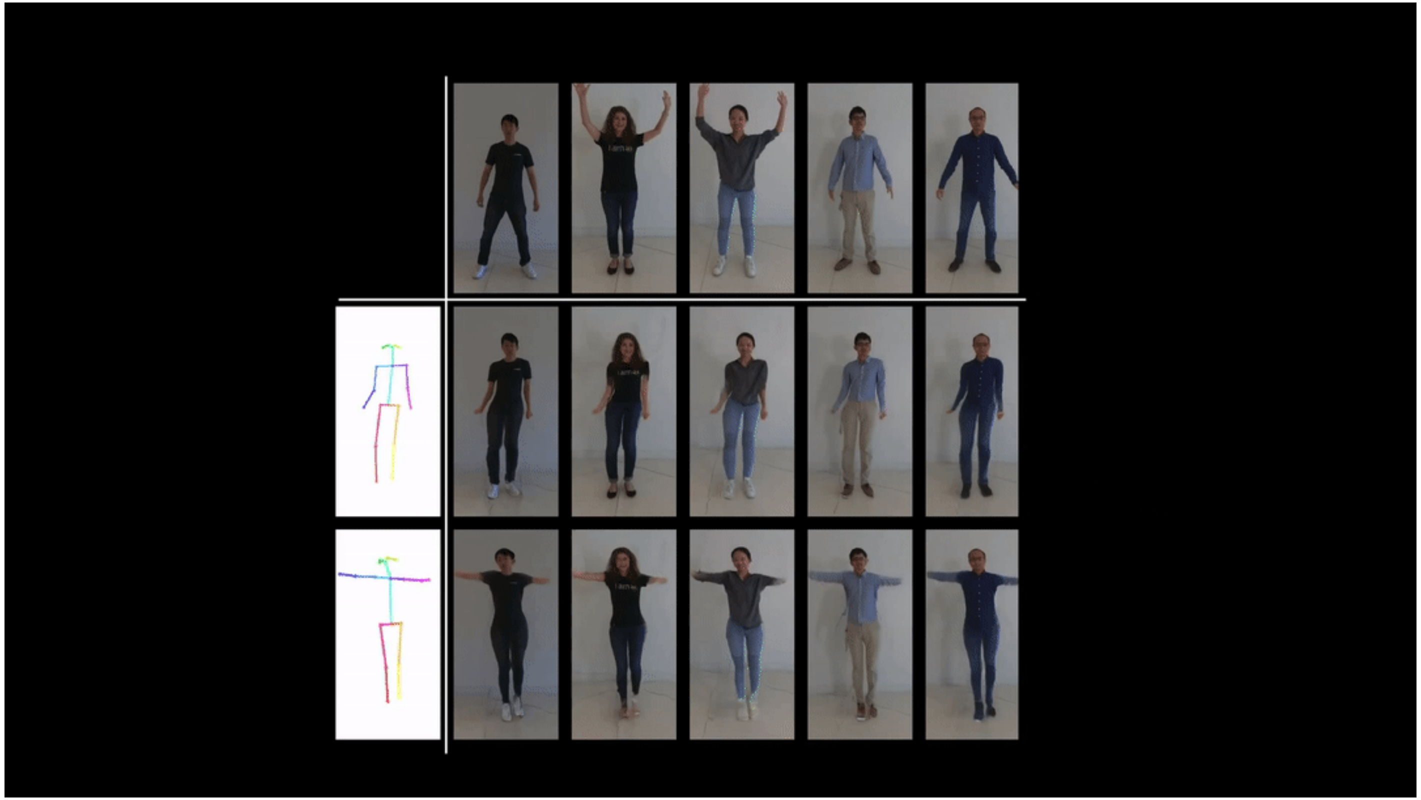




Demo: https://www.nvidia.com/en-us/research/ai-playground/

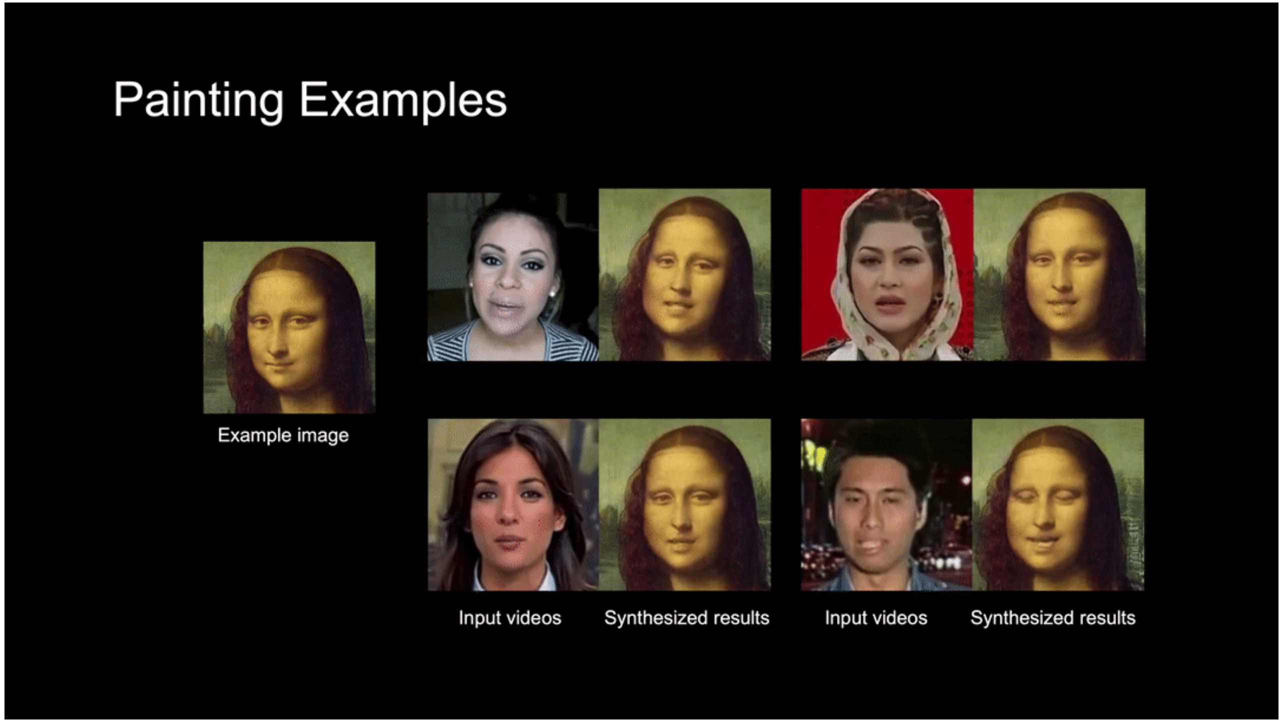
Aysegul Dundar

VIDEO SYNTHESIS



Demo: https://github.com/NVlabs/few-shot-vid2vid/

VIDEO SYNTHESIS

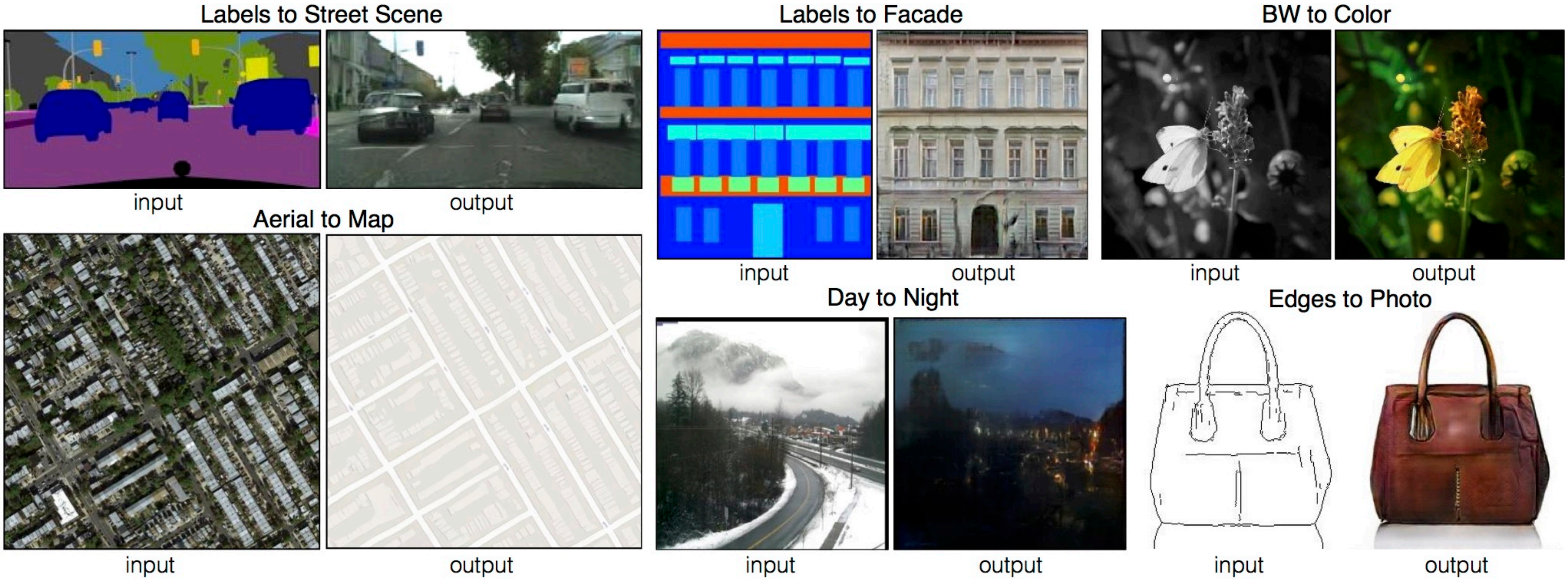


Demo: https://github.com/NVlabs/few-shot-vid2vid/

IMAGE INPAINTING

Demo: https://www.nvidia.com/research/inpainting/

IMAGE SYNTHESIS



Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." CVPR. 2017.

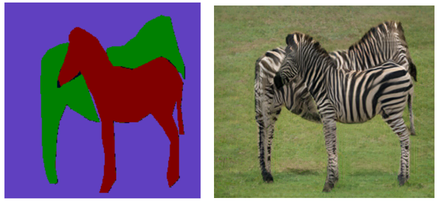
OUTLINE

- GAN and image synthesis

OUTLINE

- GAN and image synthesis

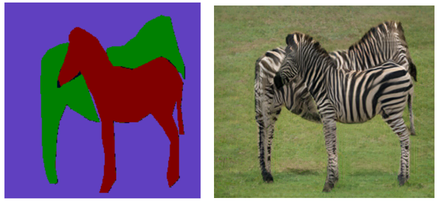
- Conditional image synthesis



OUTLINE

- GAN and image synthesis

- Conditional image synthesis



- Image inpainting

OUTLINE

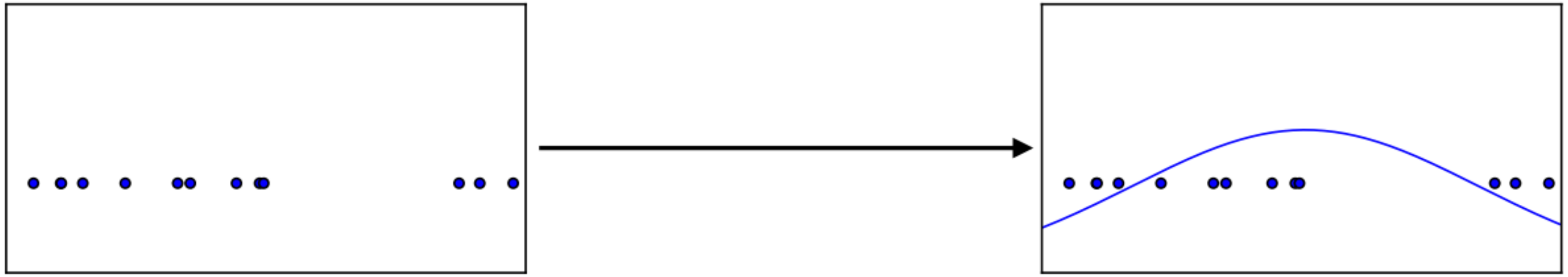
- GAN and image synthesis

- Conditional image synthesis

- Image inpainting

GENERATIVE MODELS

Density estimation



Training data
$p_{data}(x)$

Generated samples
$p_{model}(x)$

Top figure: Goodfellow, NIPS 2016 Tutorial: Generative Adversarial Networks
Dataset: http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html
Generated faces by: https://github.com/NVlabs/stylegan

MANIFOLD ASSUMPTION

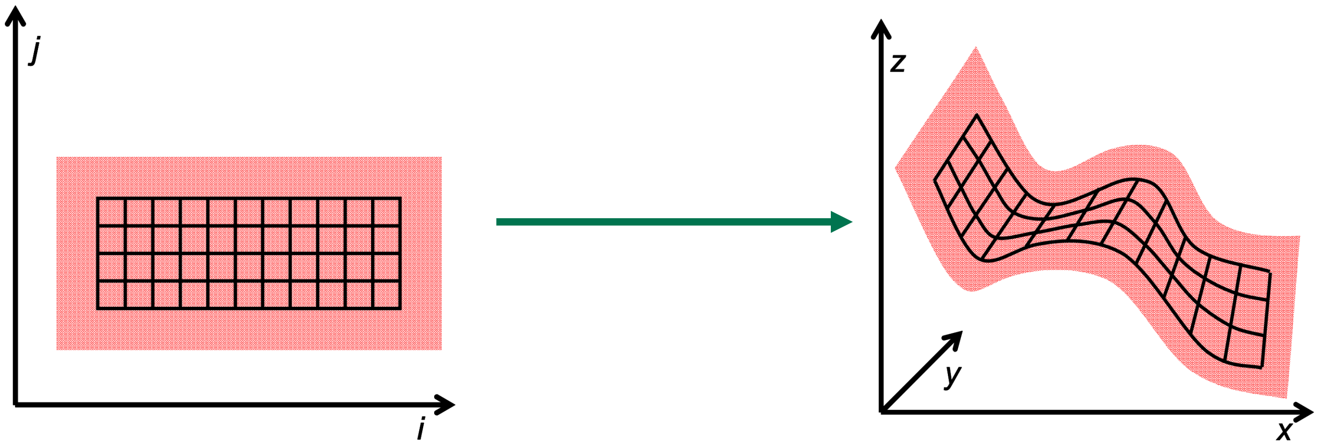
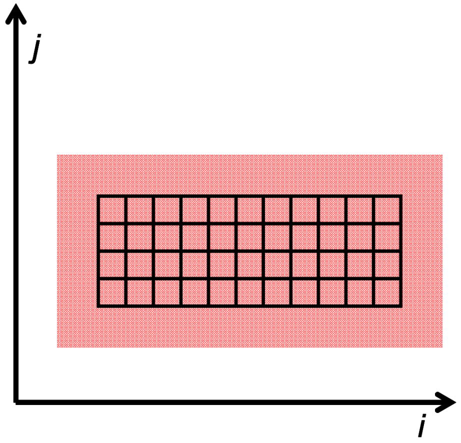
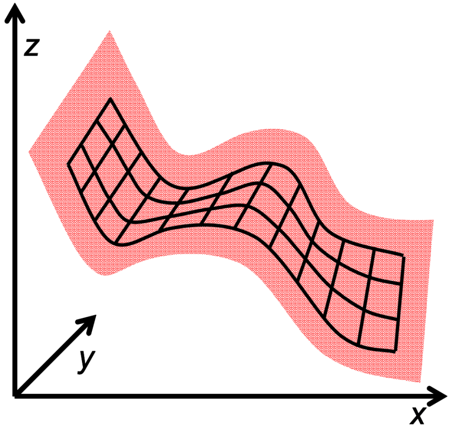


Figure credit: Ward, Aaron D., and Ghassan Hamarneh. "3D surface parameterization using manifold learning for medial shape representation." Medical Imaging 2007: Image Processing. Vol. 6512. International Society for Optics and Photonics, 2007.
Ming-Yu Liu, CVPR 2017 Tutorial: Theory and Applications of Generative Adversarial Networks

MANIFOLD ASSUMPTION





Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." ICLR (2014).

GENERATIVE MODELS

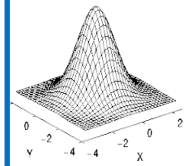
- Generative Adversarial Networks
- Variational Auto Encoders
- Flow-based Generative Model
- Pixel RNN / Pixel CNN
- Hidden Markov Model
- Gaussian Mixture Model
- ….

GENERATIVE MODELS

- **Generative Adversarial Networks**
- Variational Auto Encoders
- Flow-based Generative Model
- Pixel RNN / Pixel CNN
- Hidden Markov Model
- Gaussian Mixture Model
- ….

GENERATIVE ADVERSARIAL NETWORKS

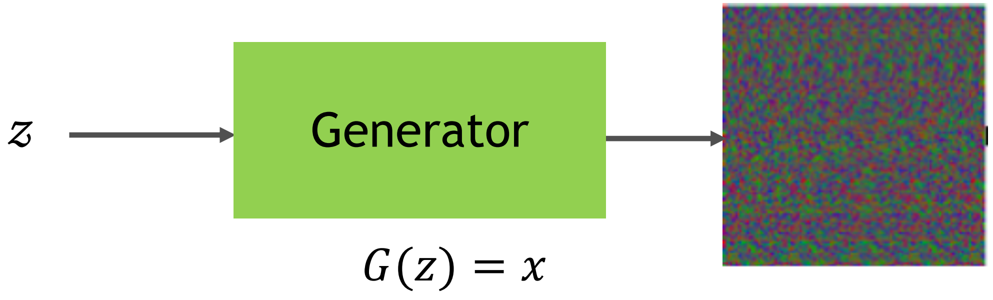
z → Generator →



$G(z) = x$

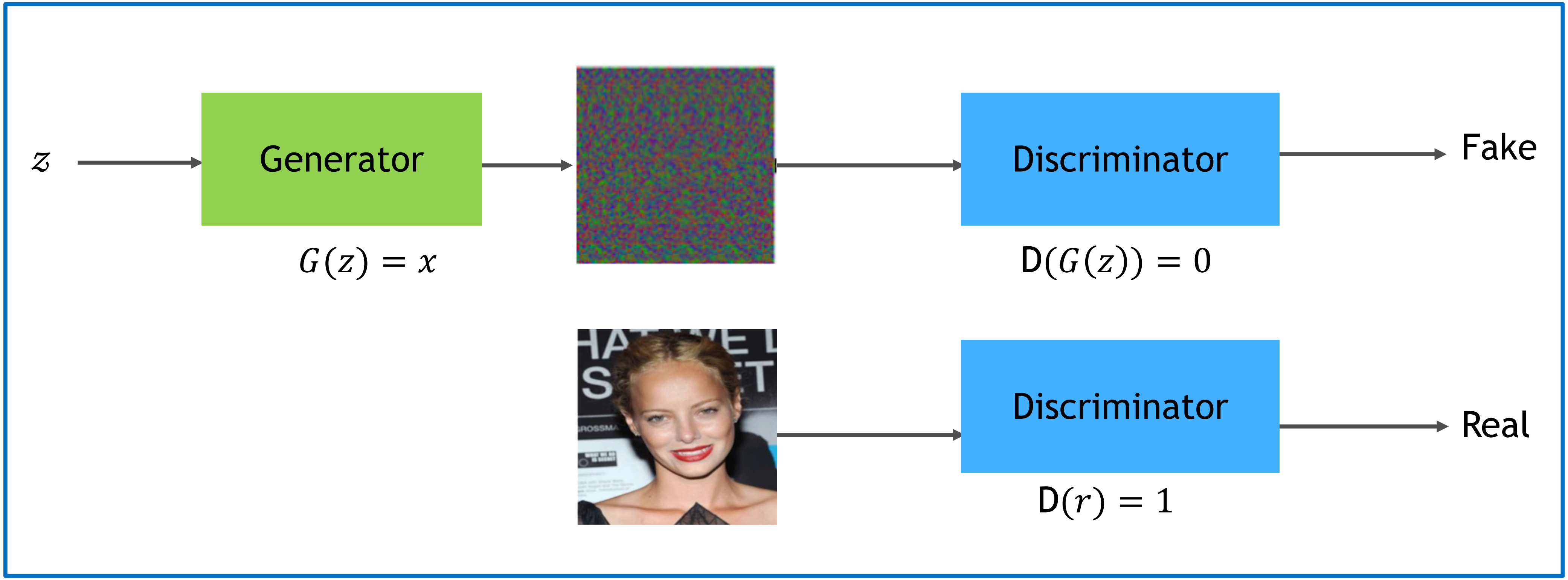
Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

GENERATIVE ADVERSARIAL NETWORKS



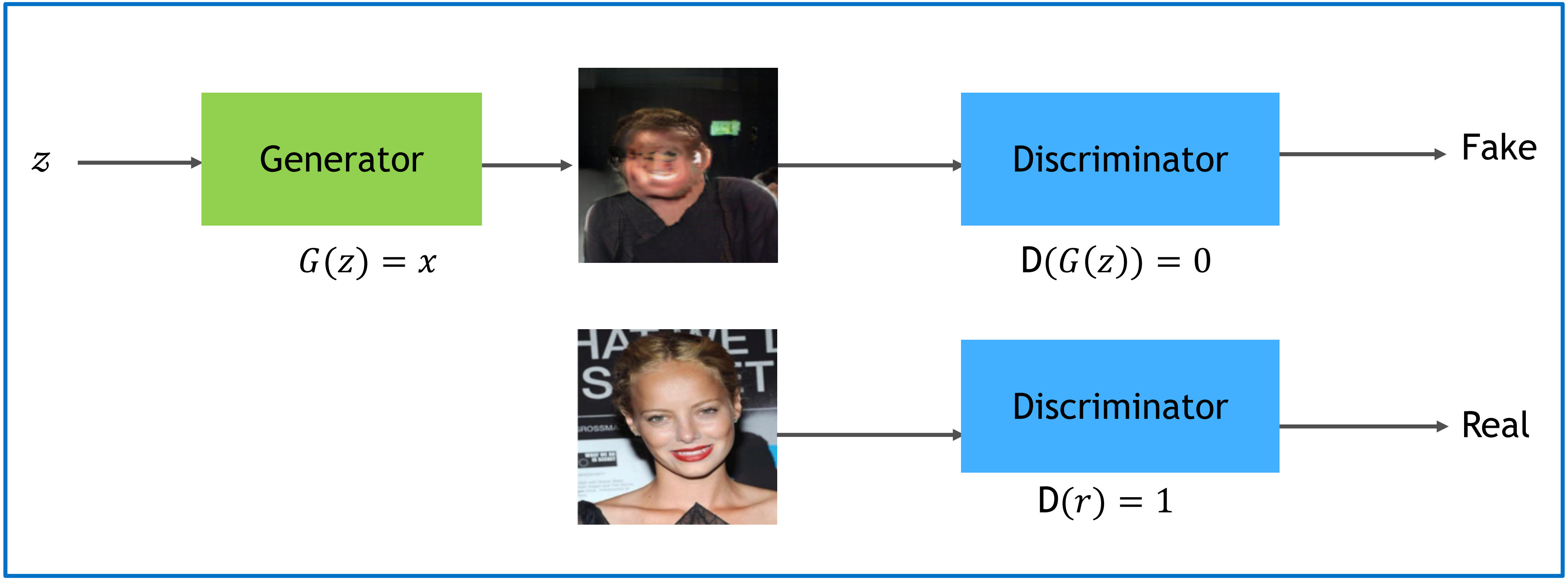
Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

GENERATIVE ADVERSARIAL NETWORKS



Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

GENERATIVE ADVERSARIAL NETWORKS



Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

OBJECTIVE FUNCTION

Train jointly in minimax game

$$\min_G \max_D \quad E_{x \sim p_X}[\log D(x)] + E_{z \sim p_Z}[\log(1 - D(G(z))]$$

Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

OBJECTIVE FUNCTION

Train jointly in minimax game

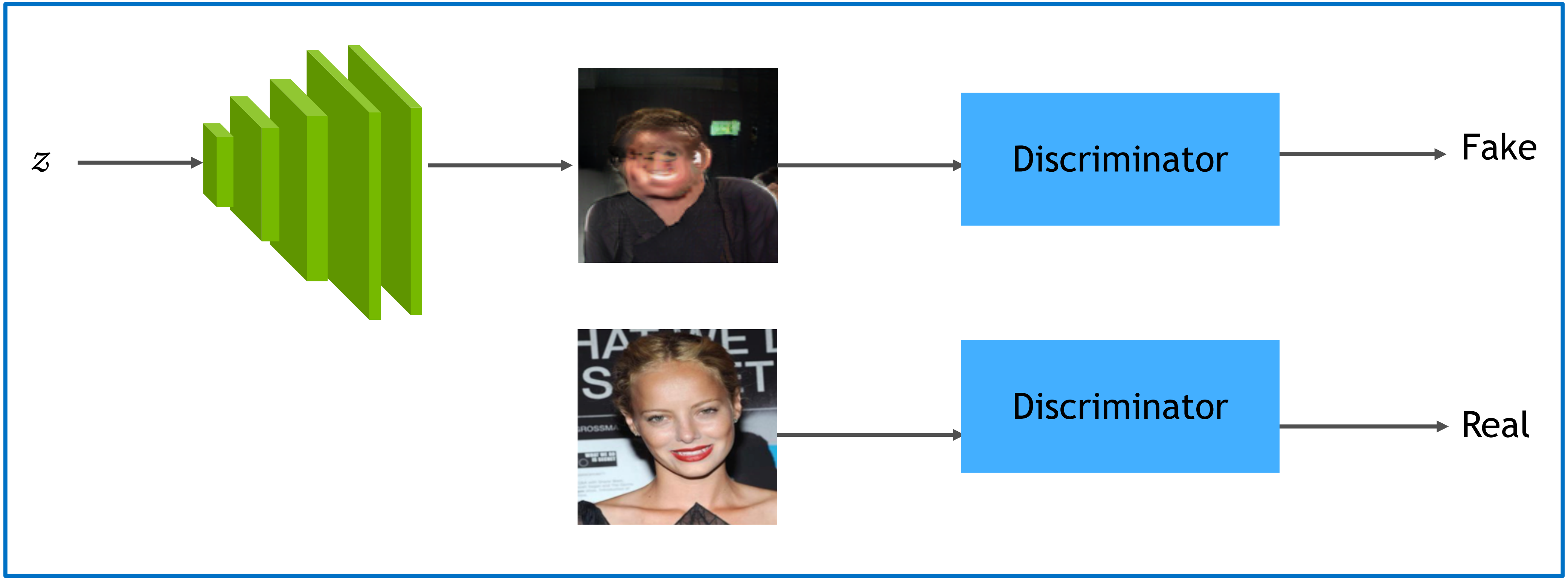
$$\min_G \max_D \quad E_{x\sim p_X}[\log D\boxed{(x)}] + E_{z\sim p_Z}[\log(1 - D(\boxed{G(z)}))]$$

real data   fake data

Discriminator maximize objective -> D(x) ~ 1 and D(G(z)) ~ 0
Generator minimize objective D(G(z)) ~ 1

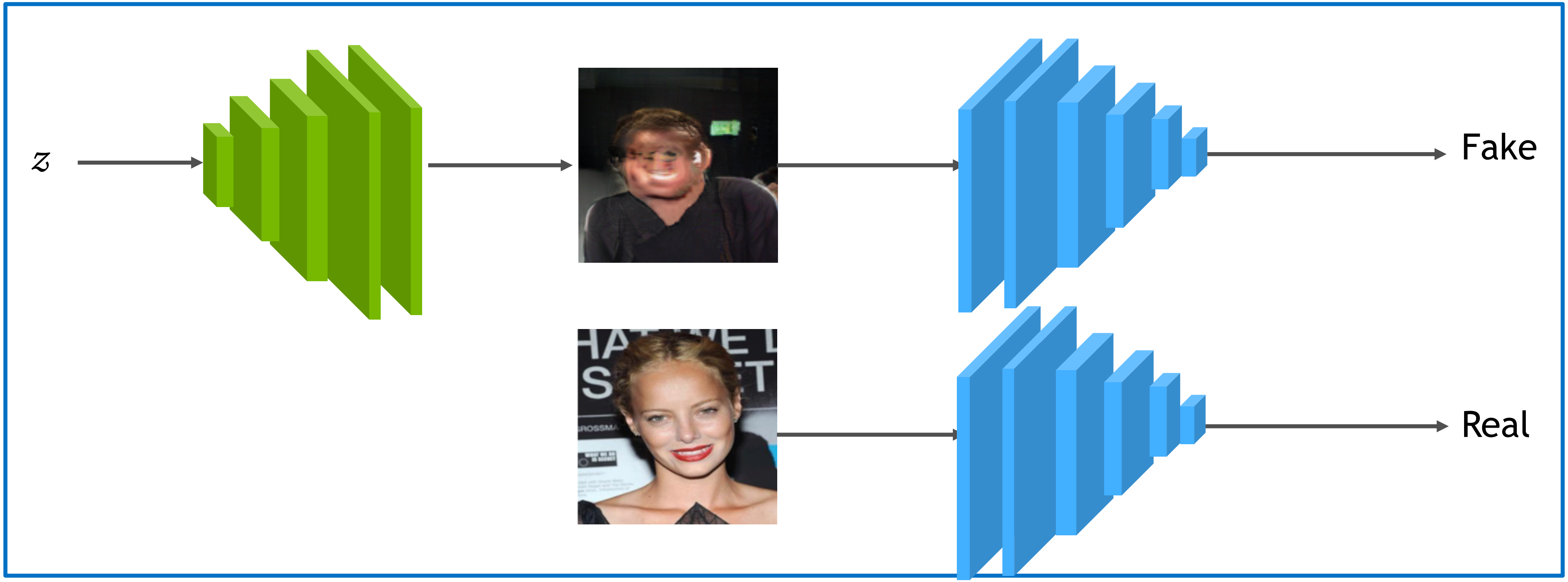
Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

GENERATIVE ADVERSARIAL NETWORKS



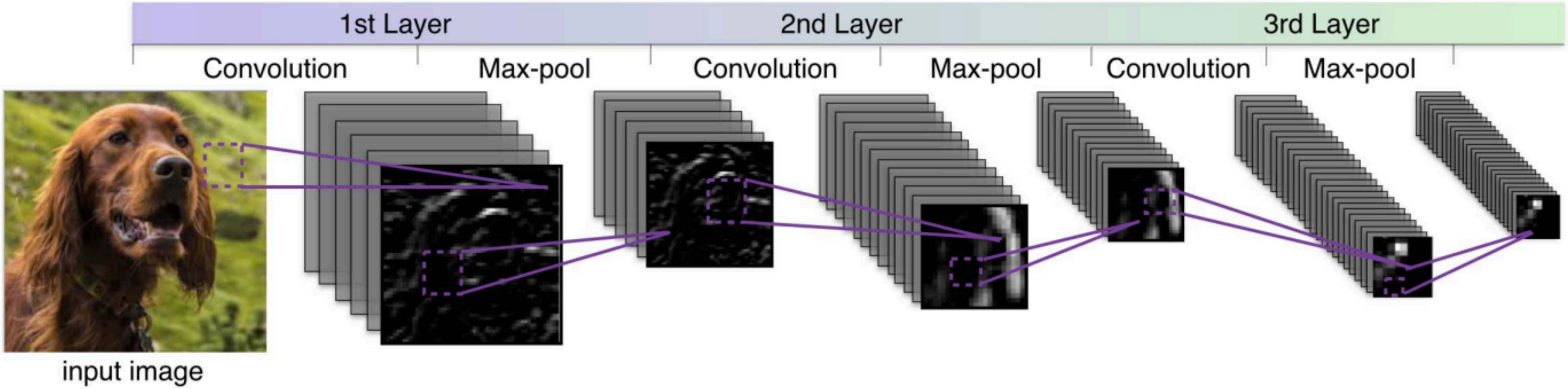
Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

GENERATIVE ADVERSARIAL NETWORKS



Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

DEEP CONVOLUTIONAL NEURAL NETWORKS

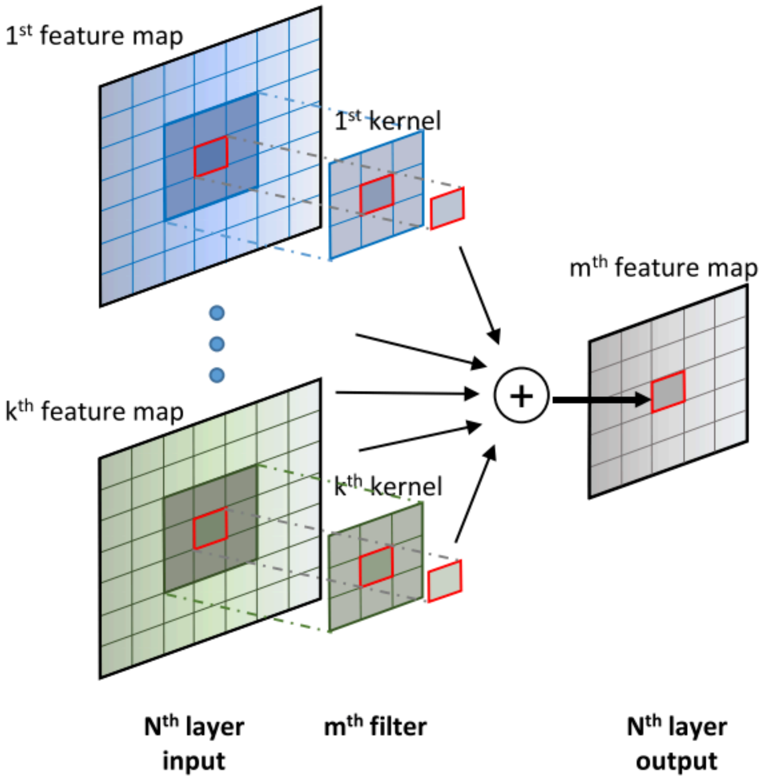


Inspired by Hubel & Wiesel 1962

1) Simple cells (convolution layer) = Detect local features
2) Complex cells (pooling layer) = Pool outputs from neighboring locations

LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." *Neural computation,* 1989.

CONVOLUTION OPERATION



Mark C. F. Sousa

Input Image

252	251	246	207	90
250	242	236	144	41
252	244	228	102	43
250	243	214	59	52
248	243	201	44	54

Feature map

Kernel

1	0	-1
1	0	-1
1	0	-1

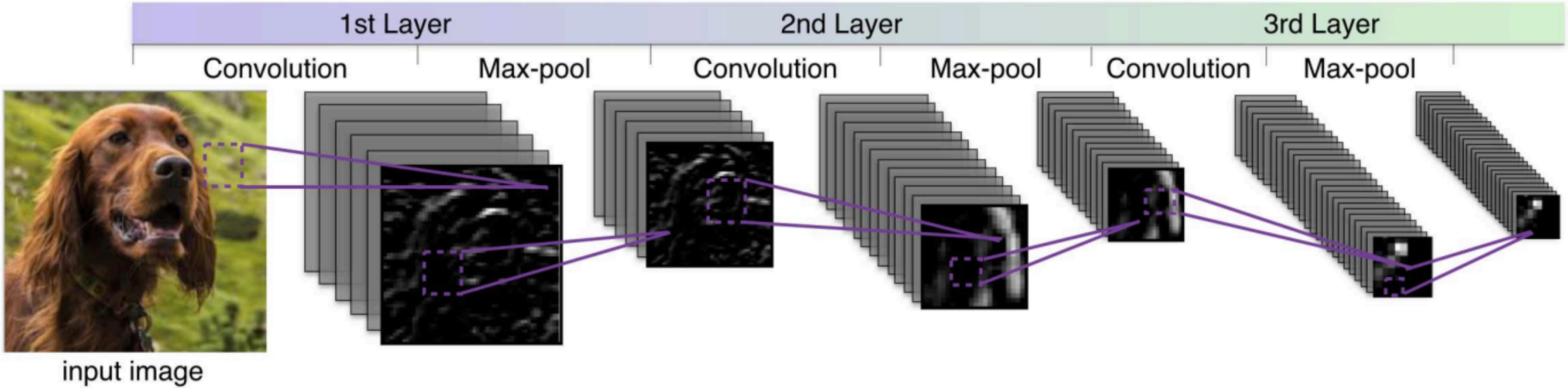
X

Receptive field

=

https://towardsdatascience.com/visualizing-the-fundamentals-of-convolutional-neural-networks-6021e5b07f69

DEEP CONVOLUTIONAL NEURAL NETWORKS

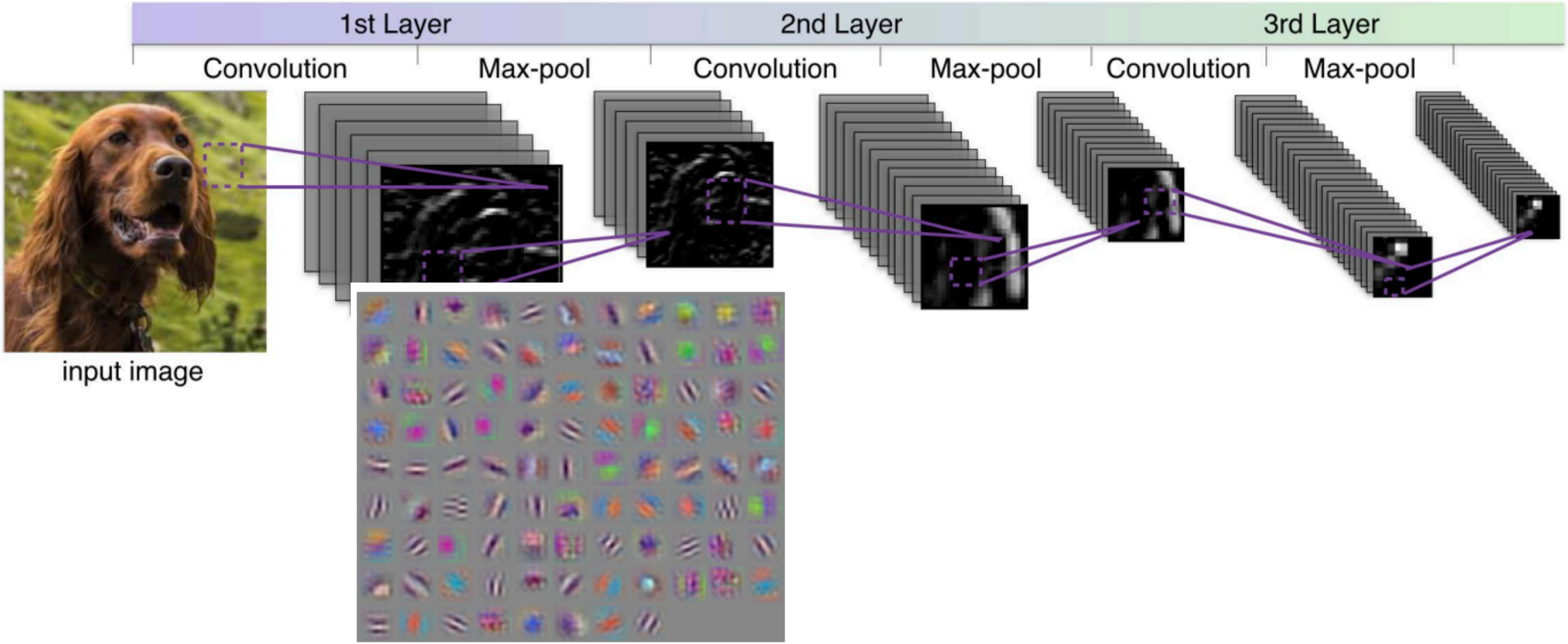


Inspired by Hubel & Wiesel 1962

1) Simple cells (convolution layer) = Detect local features
2) Complex cells (pooling layer) = Pool outputs from neighboring locations

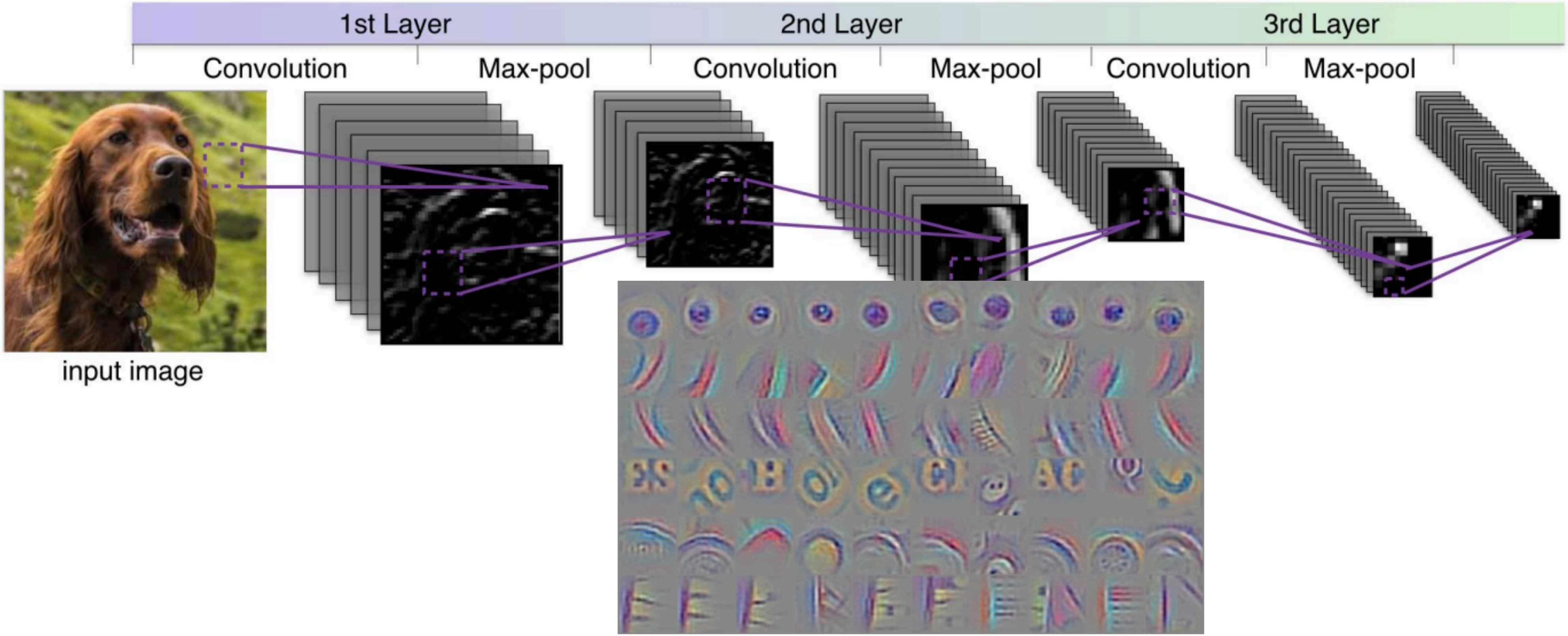
LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." *Neural computation,* 1989.

DEEP CONVOLUTIONAL NEURAL NETWORKS



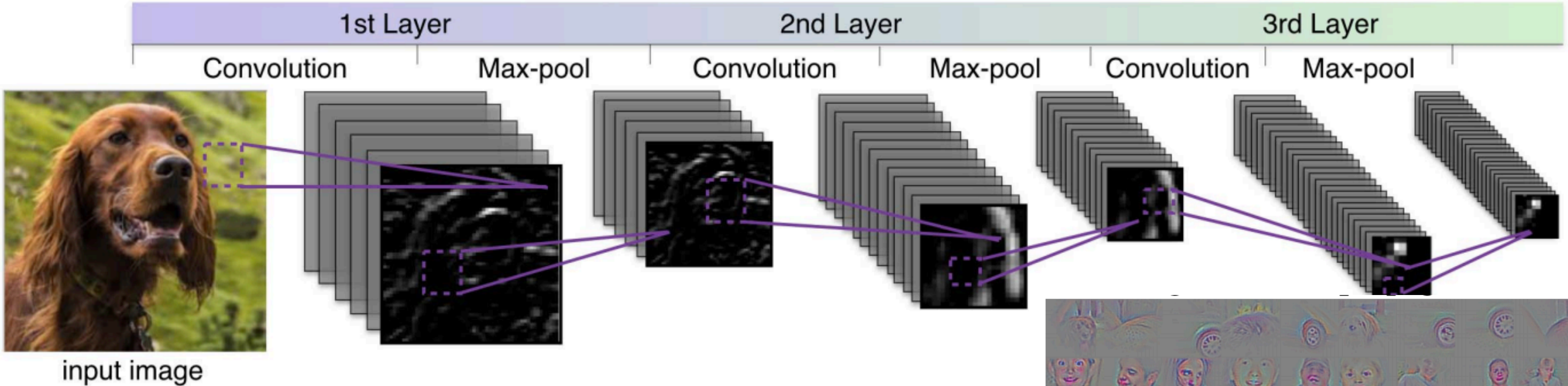
LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." *Neural computation,* 1989.

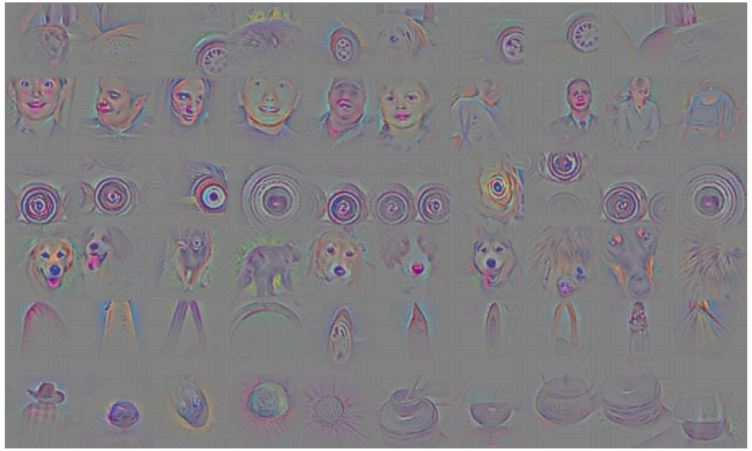
DEEP CONVOLUTIONAL NEURAL NETWORKS



LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." *Neural computation,* 1989.

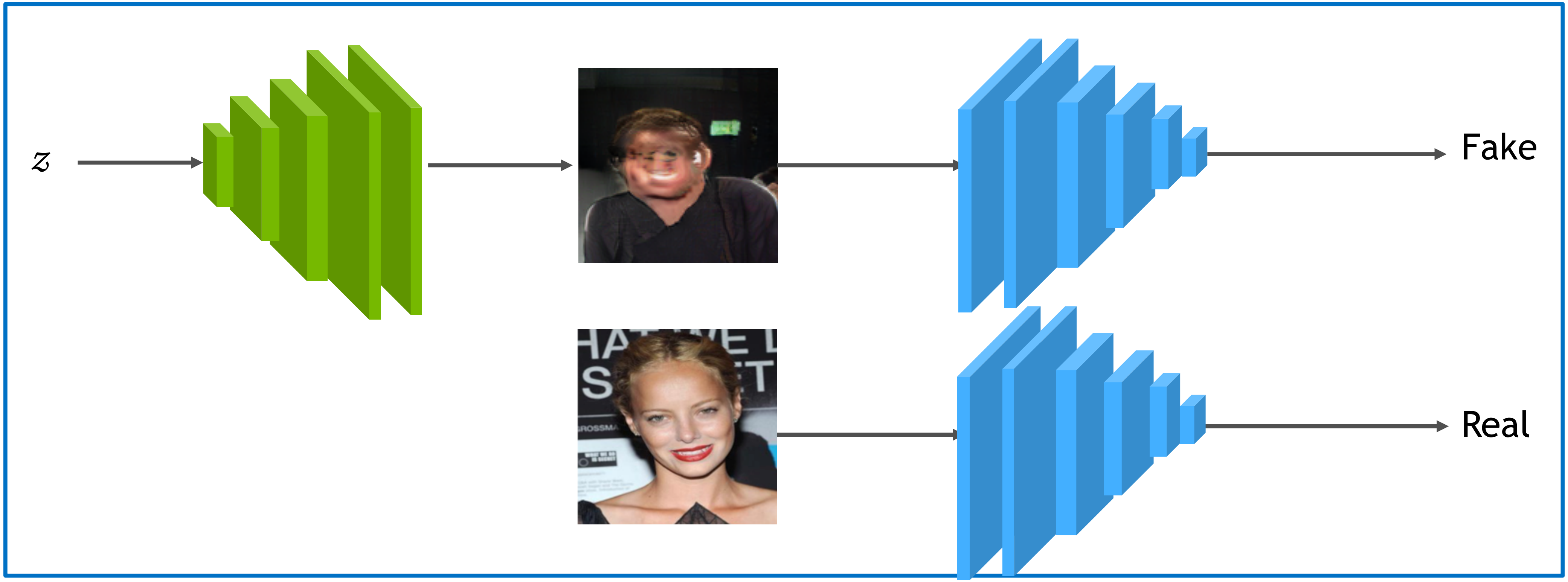
DEEP CONVOLUTIONAL NEURAL NETWORKS





LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." *Neural computation,* 1989.

GENERATIVE ADVERSARIAL NETWORKS



Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

IMAGE SYNTHESIS

GAN 2014

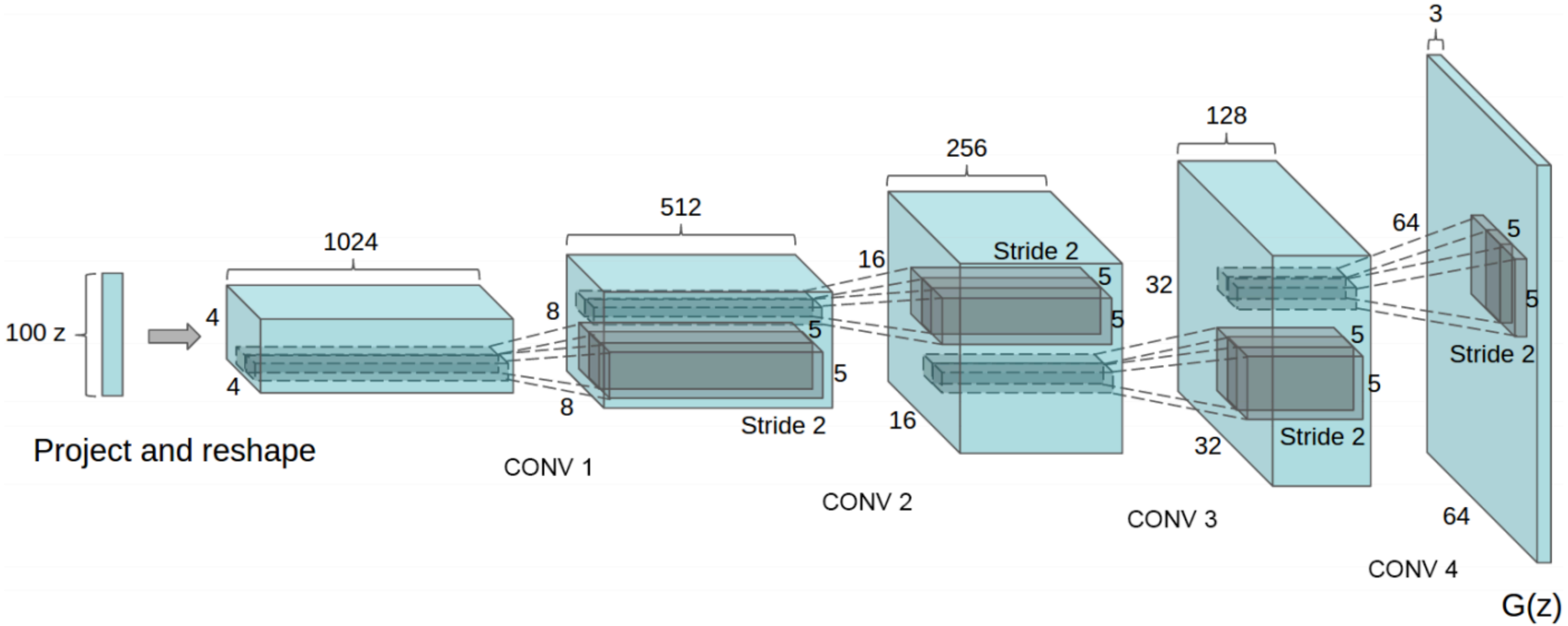
DCGAN 2016

Progressive GAN 2018

StyleGAN 2019

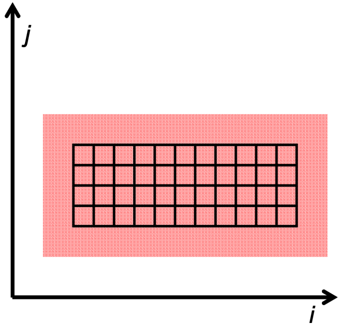
[1] Goodfellow, Ian, et al. "Generative adversarial nets." Neurips. 2014.
[2] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." ICLR. 2016
[3] Karras, Tero, et al. "Progressive growing of gans for improved quality, stability, and variation." ICLR. 2018.
[4] Karras, Tero, Samuli Laine, and Timo a. "A style-based generator architecture for generative adversarial networks." CVPR. 2019.

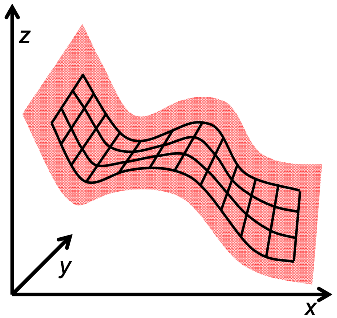
DCGAN



Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." ICLR. 2016

DCGAN - INTERPOLATION

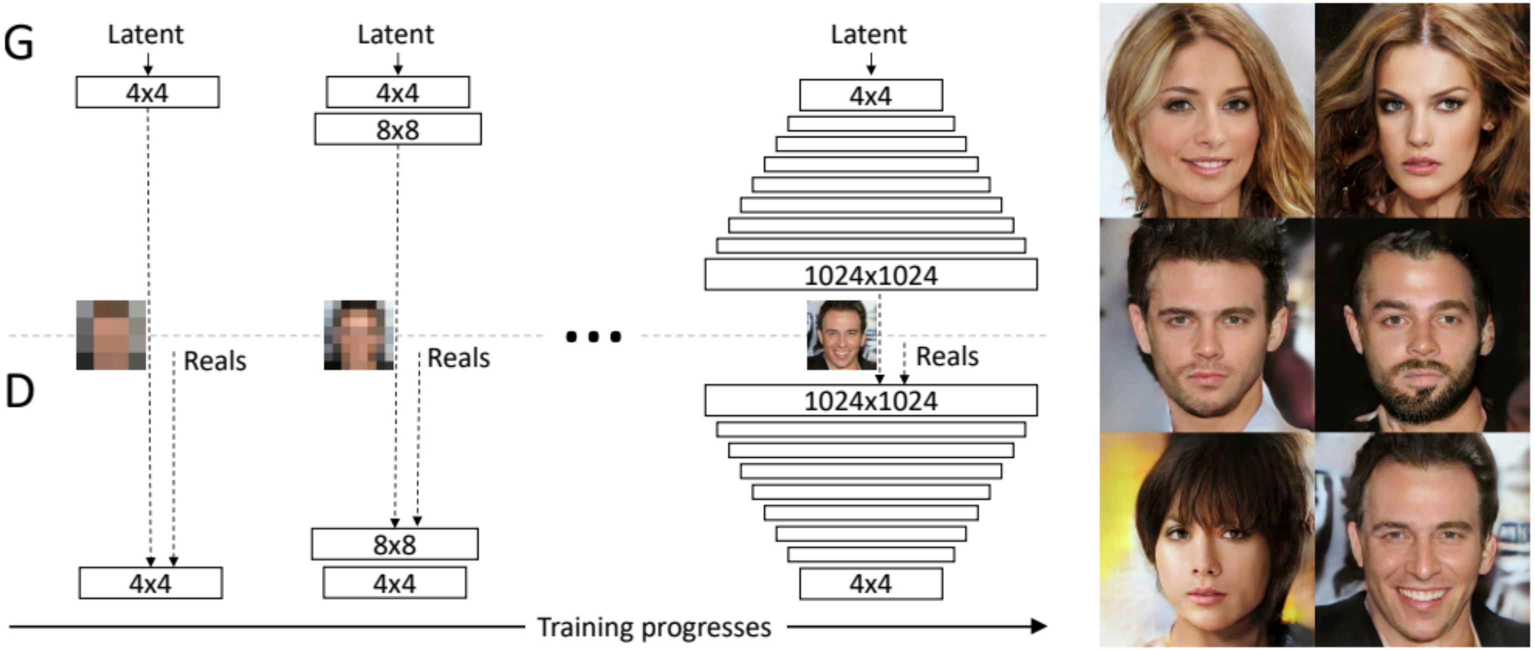




$G(z_1)$ $G(1/2 * z_1 + 1/2 * z_2)$ $G(z_2)$

Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." ICLR. 2016

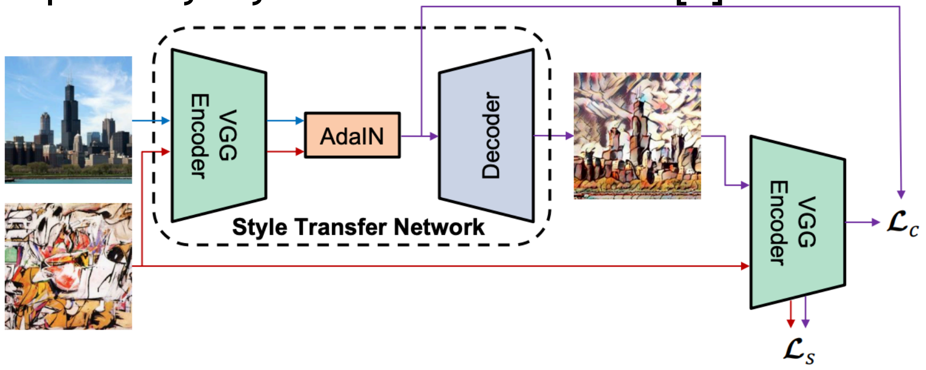
PROGRESSIVE GAN



Karras, Tero, et al. "Progressive growing of gans for improved quality, stability, and variation." ICLR. 2018.

STYLEGAN

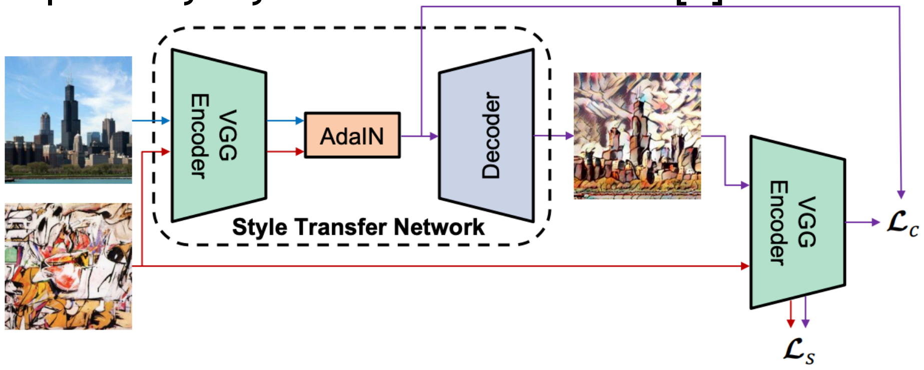
Inspired by style transfer networks [1]



[1] Huang, Xun, and Serge Belongie. "Arbitrary style transfer in real-time with adaptive instance normalization." *ICCV*. 2017.
StyleGAN: Karras, Tero, Samuli Laine, and Timo a. "A style-based generator architecture for generative adversarial networks." CVPR. 2019.

STYLEGAN

Inspired by style transfer networks [1]



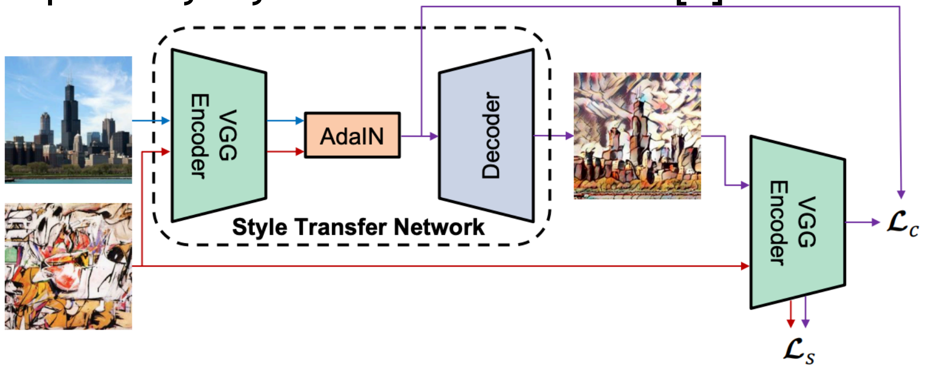
$$\mathrm{AdaIN}(x, y) = \sigma(y)\left(\frac{x - \mu(x)}{\sigma(x)}\right) + \mu(y) \tag{8}$$

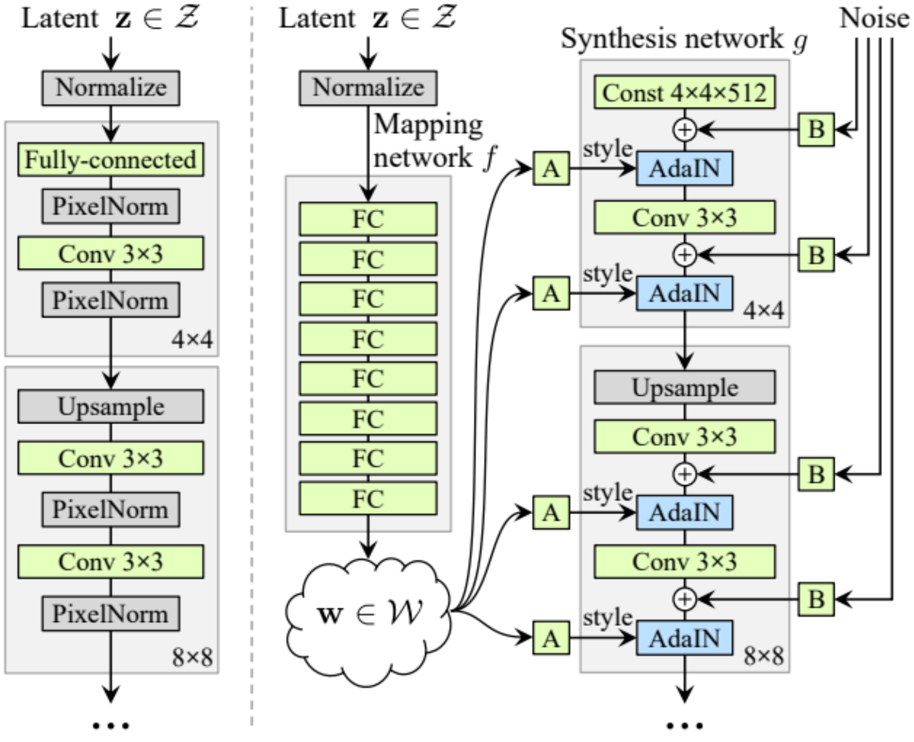
in which we simply scale the normalized content input with $\sigma(y)$, and shift it with $\mu(y)$. Similar to IN, these statistics are computed across spatial locations.

[1] Huang, Xun, and Serge Belongie. "Arbitrary style transfer in real-time with adaptive instance normalization." *ICCV*. 2017.
StyleGAN: Karras, Tero, Samuli Laine, and Timo a. "A style-based generator architecture for generative adversarial networks." CVPR. 2019.

STYLEGAN

Inspired by style transfer networks [1]





(a) Traditional (b) Style-based generator

[1] Huang, Xun, and Serge Belongie. "Arbitrary style transfer in real-time with adaptive instance normalization." *ICCV*. 2017.
StyleGAN: Karras, Tero, Samuli Laine, and Timo a. "A style-based generator architecture for generative adversarial networks." CVPR. 2019.

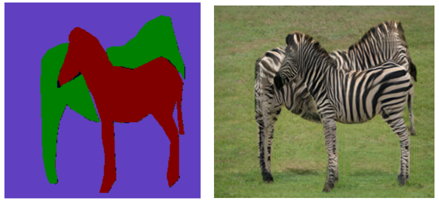
STYLEGAN

Karras, Tero, Samuli Laine, and Timo Aila. "A style-based generator architecture for generative adversarial networks." CVPR. 2019.

OUTLINE

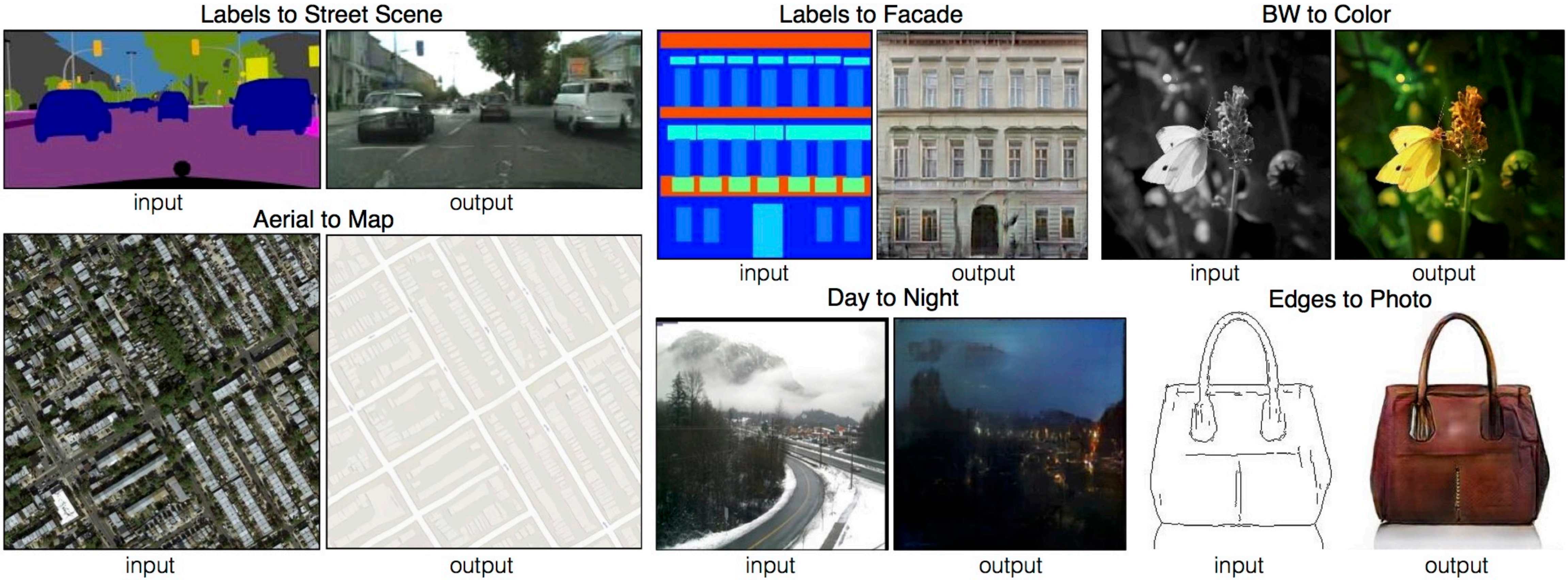
- Image synthesis

- Conditional image synthesis



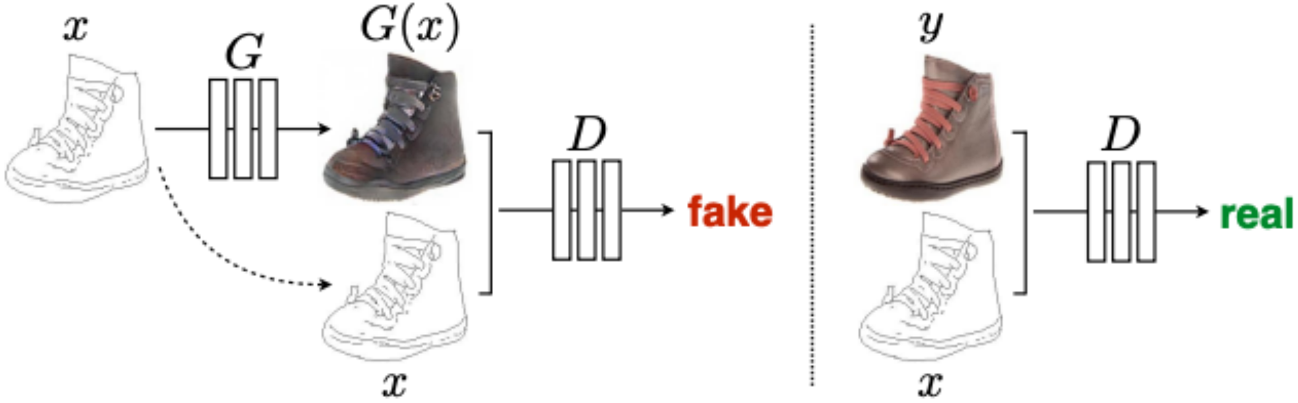
- Image inpainting

CONDITIONAL IMAGE SYNTHESIS



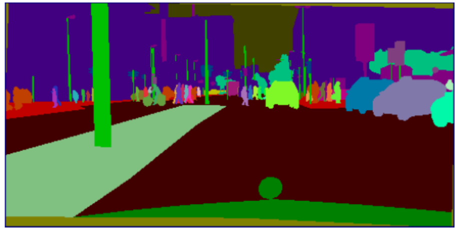
Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." CVPR. 2017.

PIX2PIX

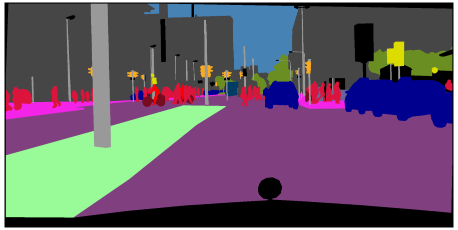


Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." CVPR. 2017.

PIX2PIX-HD



Panoptic map

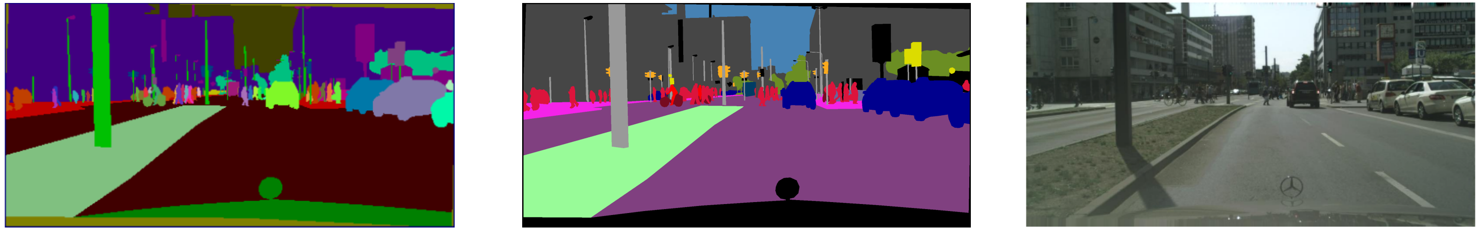


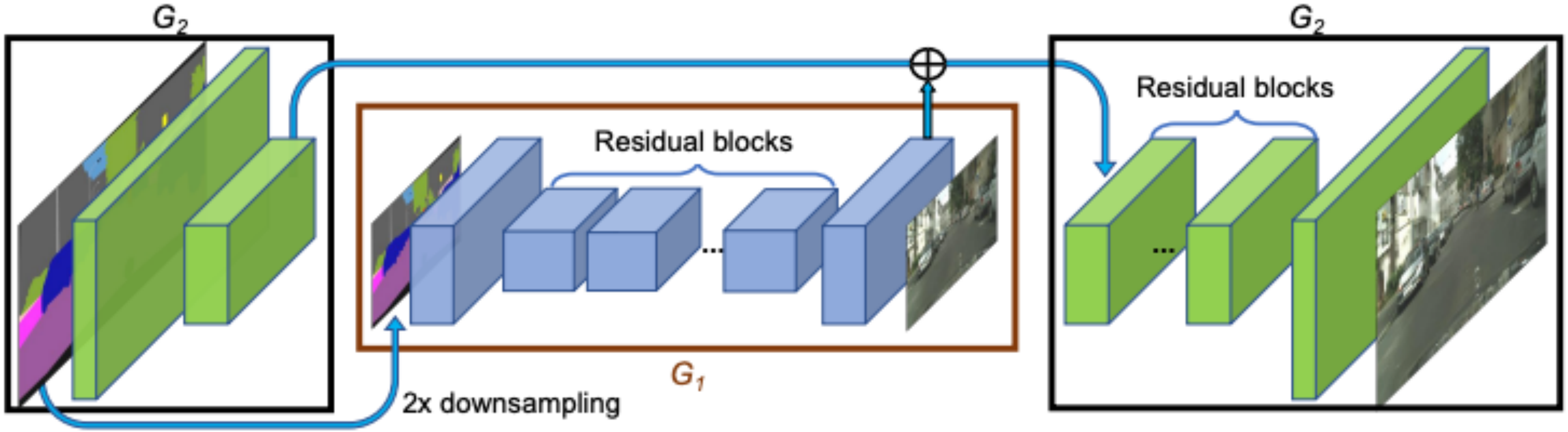
Semantic map

Real image

Wang, Ting-Chun, et al. "High-resolution image synthesis and semantic manipulation with conditional gans." CVPR. 2018.

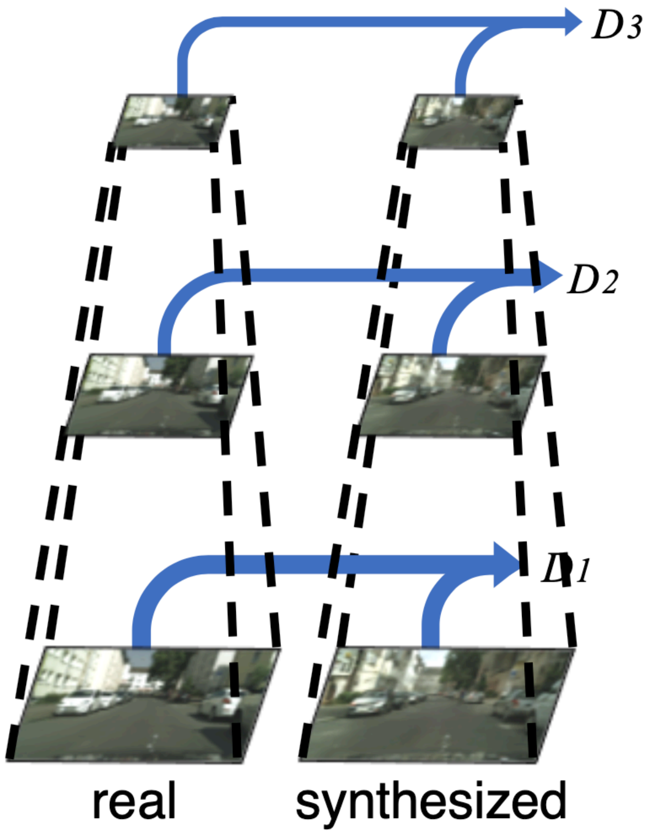
PIX2PIX-HD





Wang, Ting-Chun, et al. "High-resolution image synthesis and semantic manipulation with conditional gans." CVPR. 2018.

PIX2PIX-HD - DISCRIMINATOR

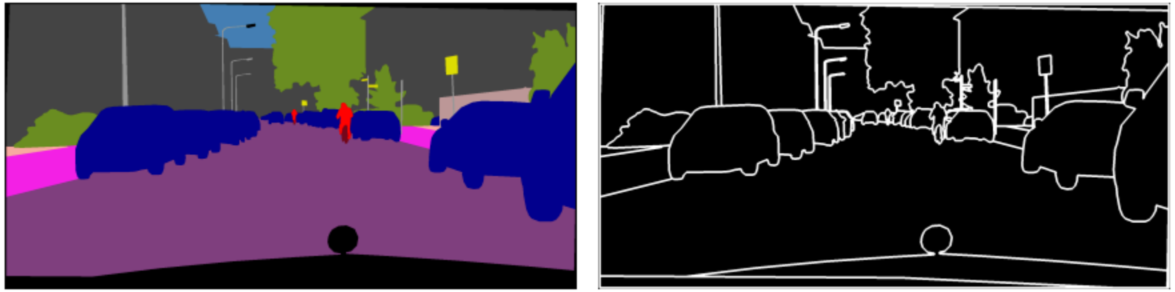


Discriminator outputs multiple scales

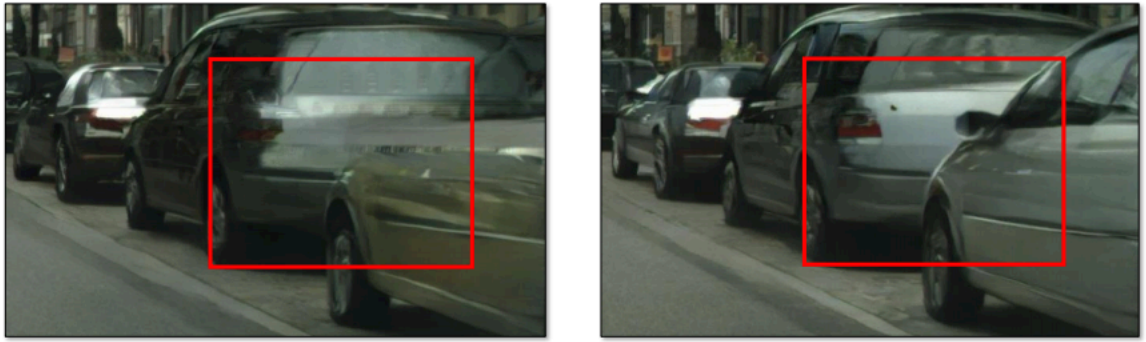
Handles global and local features

Wang, Ting-Chun, et al. "High-resolution image synthesis and semantic manipulation with conditional gans." CVPR. 2018.

PIX2PIX-HD



(a) Semantic labels (b) Boundary map



(a) Using labels only (b) Using label + instance map

Wang, Ting-Chun, et al. "High-resolution image synthesis and semantic manipulation with conditional gans." CVPR. 2018.

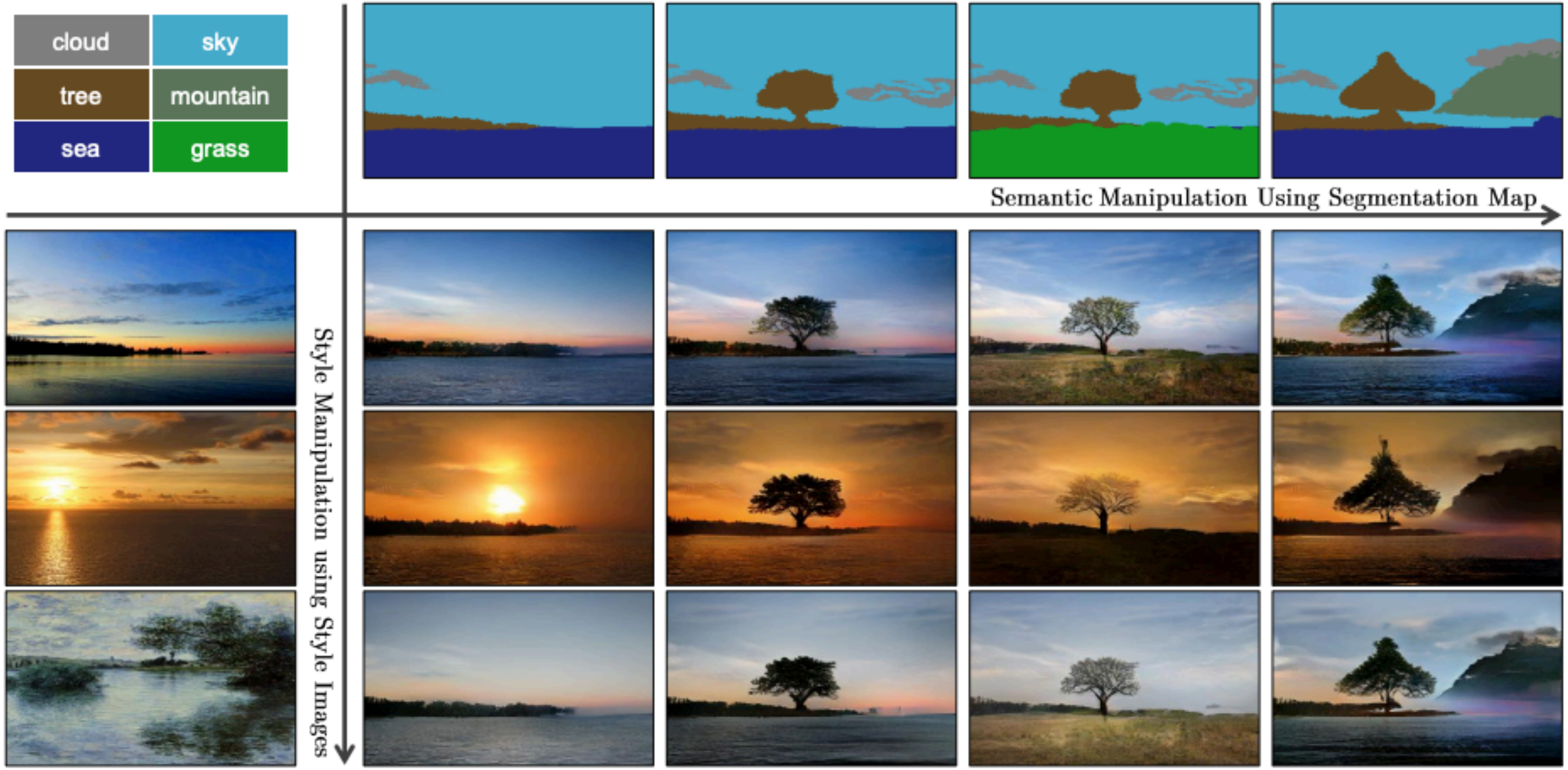
PIX2PIX-HD

original synthesized

original synthesized

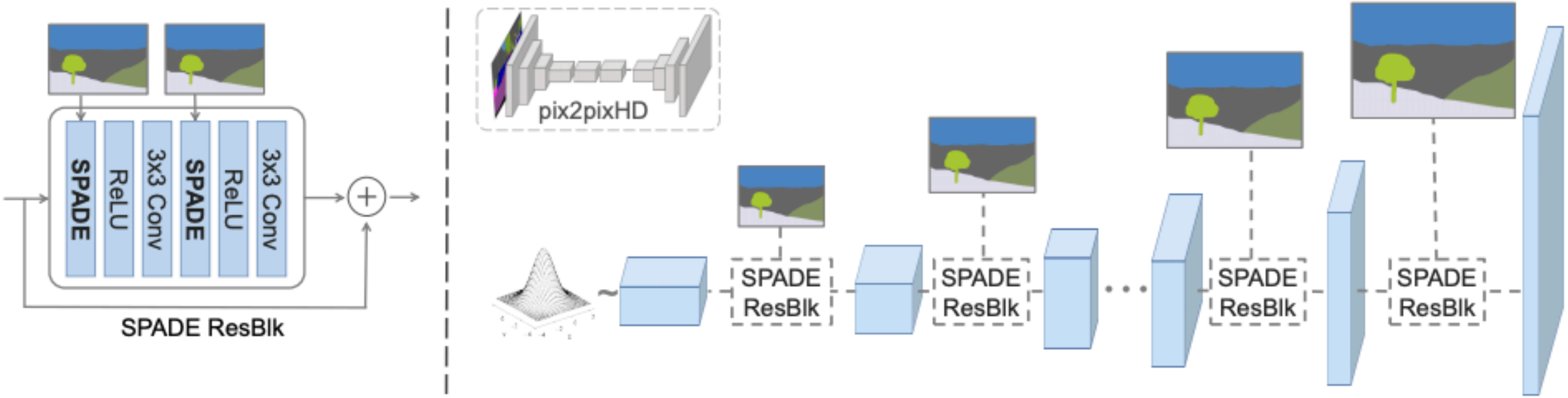
Wang, Ting-Chun, et al. "High-resolution image synthesis and semantic manipulation with conditional gans." CVPR. 2018.

SPADE



Park, Taesung, et al. "Semantic image synthesis with spatially-adaptive normalization."CVPR. 2019.

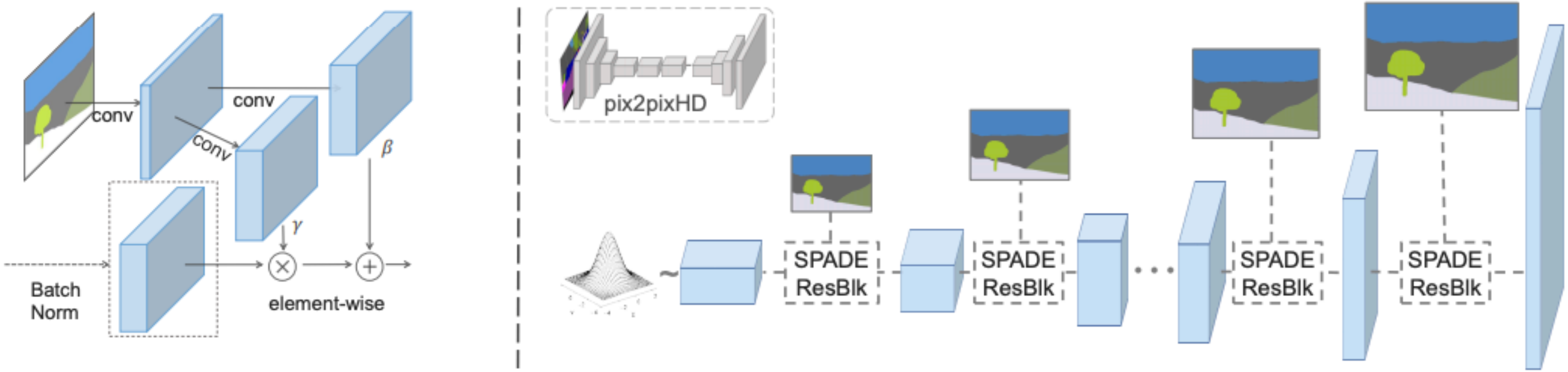
SPADE – SPATIALLY ADAPTIVE NORMALIZATION



Park, Taesung, et al. "Semantic image synthesis with spatially-adaptive normalization."CVPR. 2019.

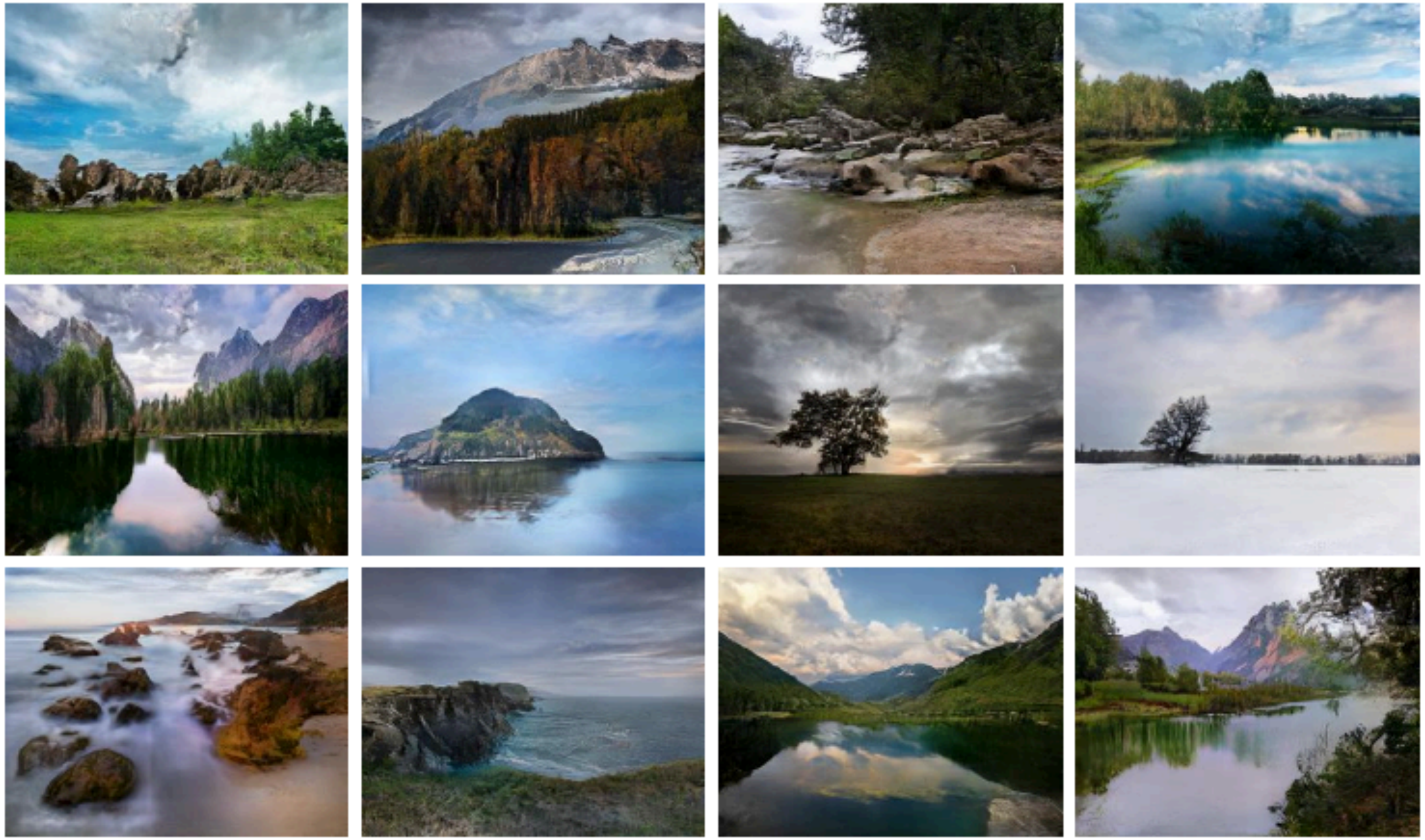
Aysegul Dundar

ARCHITECTURE



Park, Taesung, et al. "Semantic image synthesis with spatially-adaptive normalization."CVPR. 2019.

RESULTS

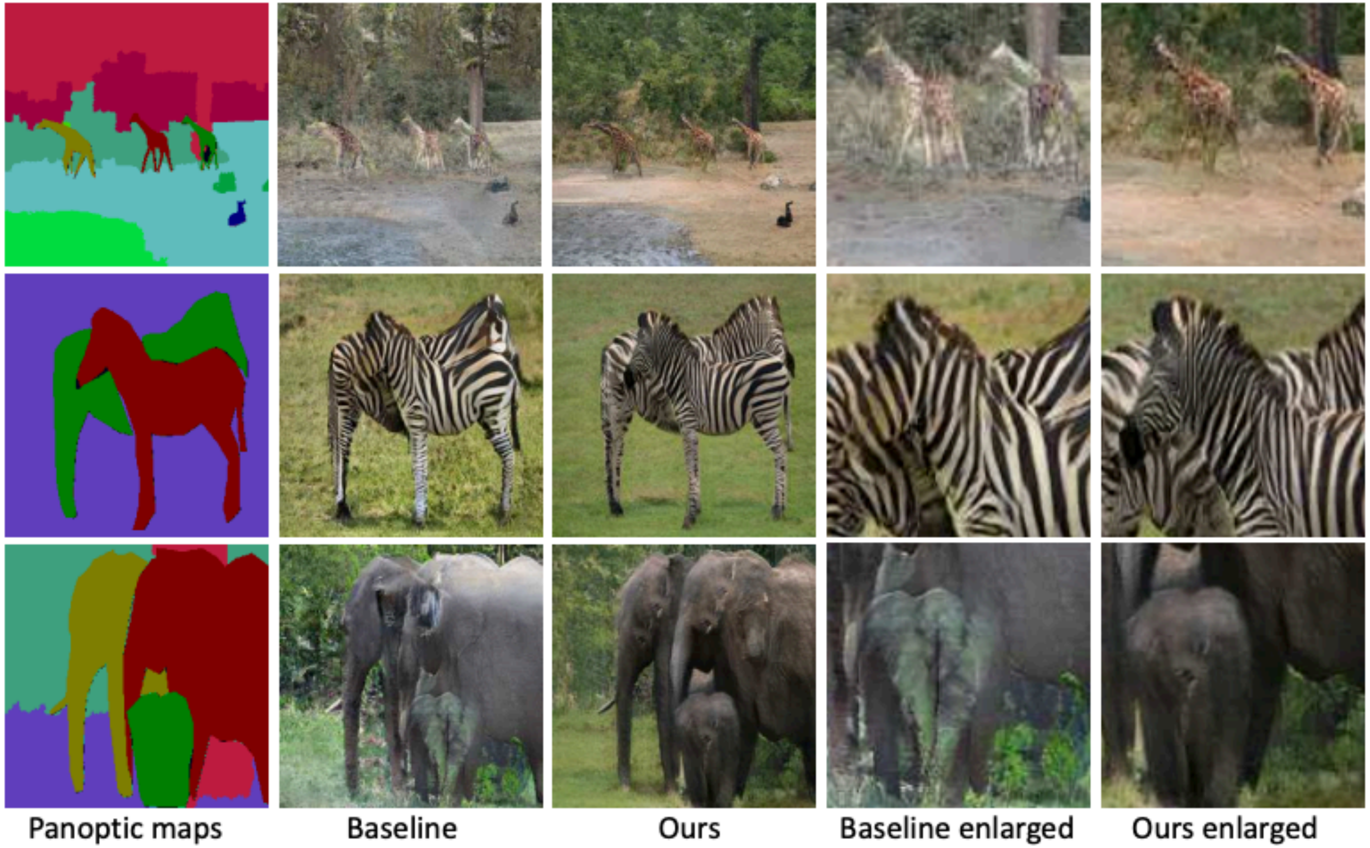


Park, Taesung, et al. "Semantic image synthesis with spatially-adaptive normalization.”CVPR. 2019.

RESULTS

Park, Taesung, et al. "Semantic image synthesis with spatially-adaptive normalization.”CVPR. 2019.

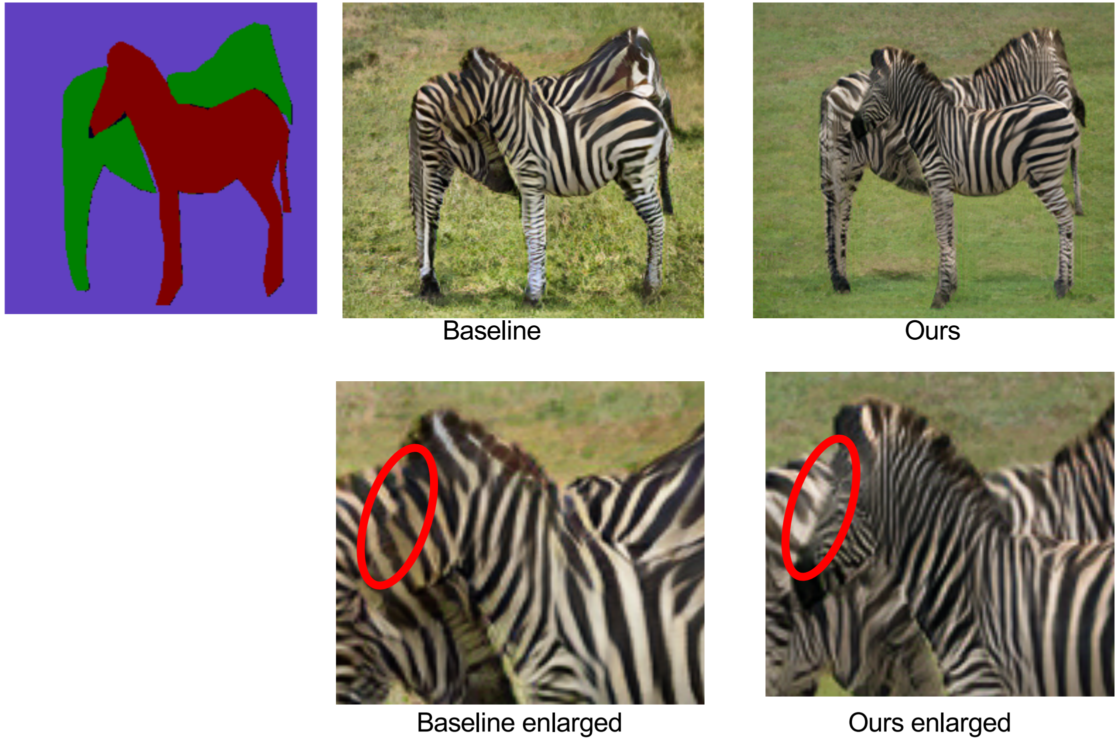
PANOPTIC BASED IMAGE SYNTHESIS



Dundar, Aysegul, et al. "Panoptic-based Image Synthesis." CVPR. 2020.

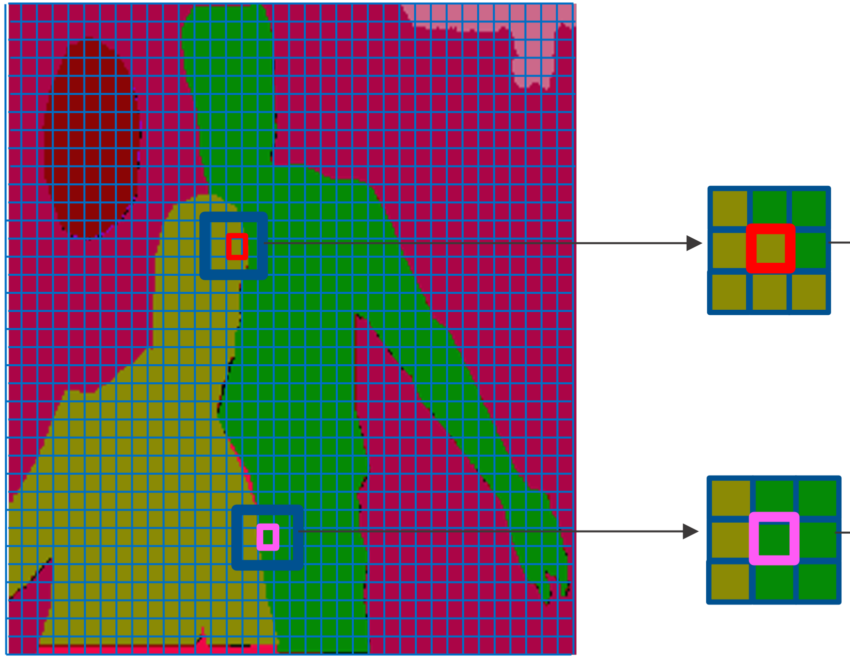
PANOPTIC BASED IMAGE SYNTHESIS

- **Goal:** Synthesizing images given panoptic maps.
- **Limitation of Prior Work:** Conventional convolution layer operate independent of panoptic maps.
- **Proposal:** Use panoptic maps efficiently in convolution layer.



Dundar, Aysegul, et al. "Panoptic-based Image Synthesis." CVPR. 2020.

PANOPTIC BASED IMAGE SYNTHESIS



1	0	0
1	1	0
1	1	1

0	1	1
0	1	1
0	1	1

panoptic map

zoomed patch

mask (M)

masks (M) is reconstructed based on panoptic masks (P)

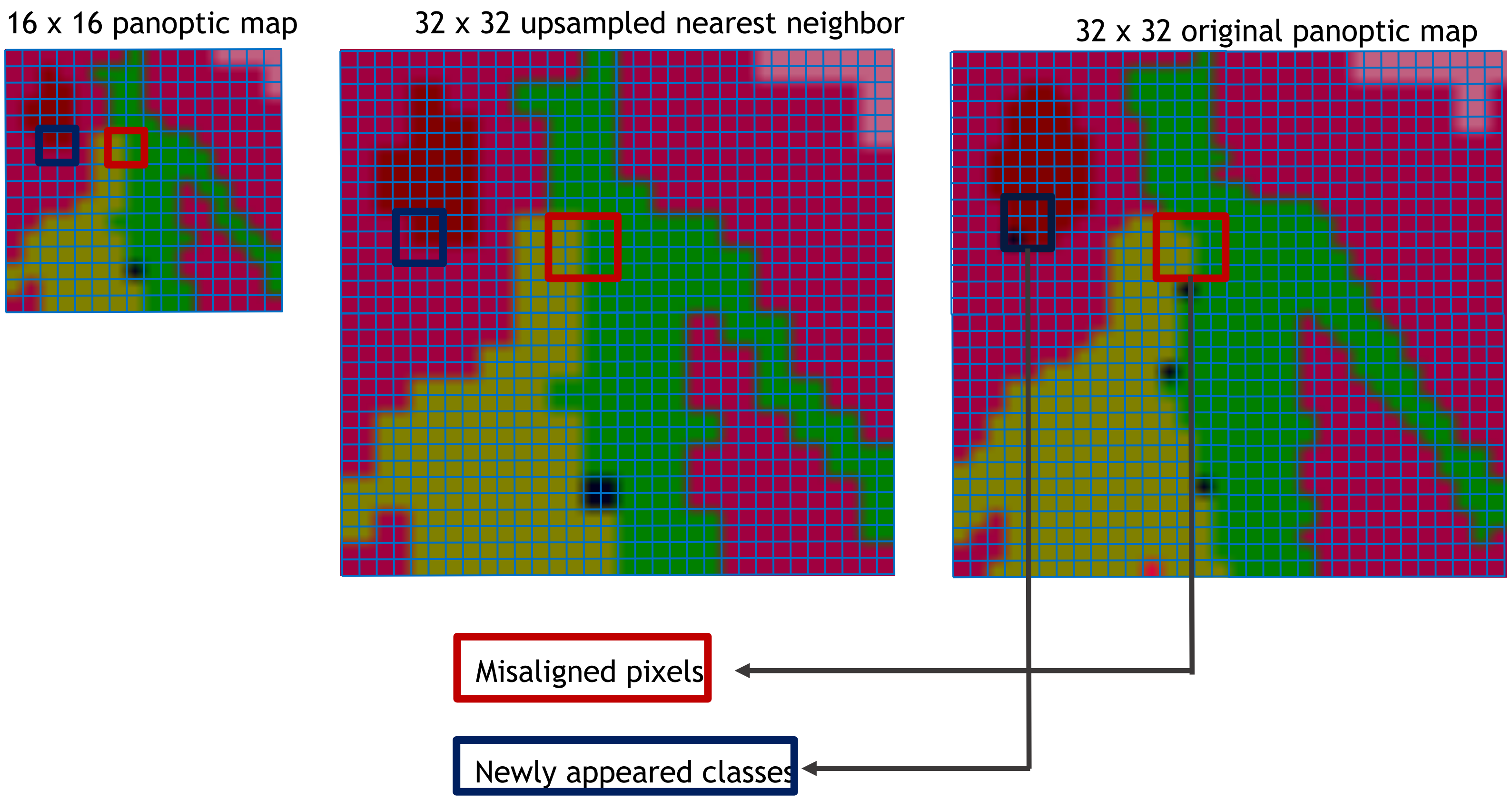
$$m_{(i,j)} = \begin{cases} 1, & \text{if } \mathbf{P}_{(\mathbf{i},\mathbf{j})} == \mathbf{P}_{(\mathbf{center},\mathbf{center})} \\ 0, & \text{otherwise} \end{cases}$$

masks (M) are used to weight convolution results.

$$x' = \begin{cases} \mathbf{W}^T(\mathbf{X} \odot \mathbf{M})\frac{\text{sum}(\mathbf{1})}{\text{sum}(\mathbf{M})} + b, & \text{if sum}(\mathbf{M}) > 0 \\ 0, & \text{otherwise} \end{cases}$$

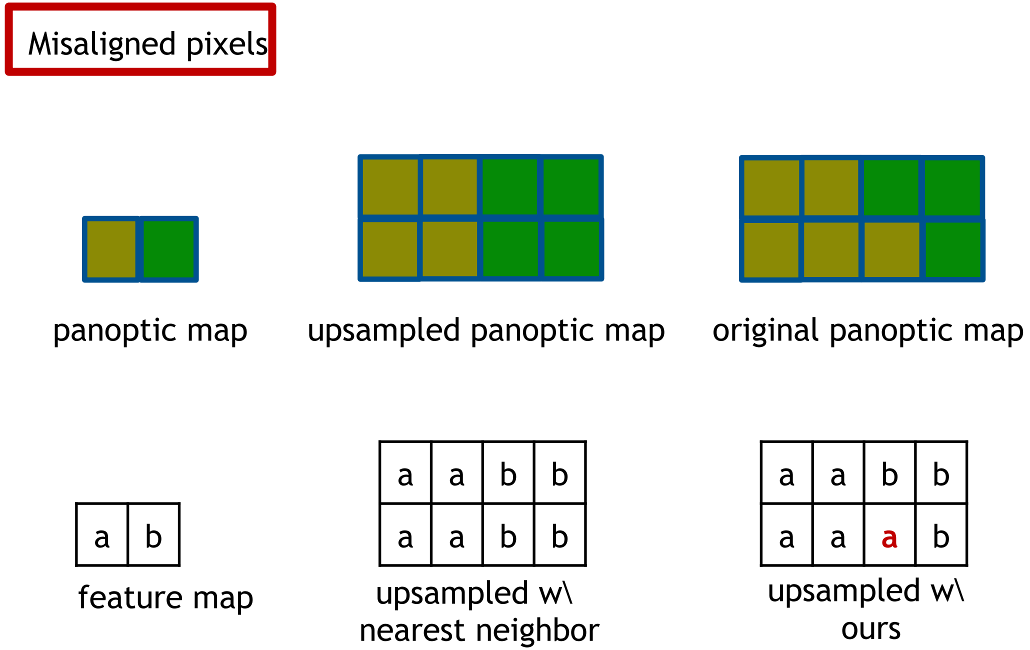
Dundar, Aysegul, et al. "Panoptic-based Image Synthesis." CVPR. 2020.

PANOPTIC BASED IMAGE SYNTHESIS



Dundar, Aysegul, et al. "Panoptic-based Image Synthesis." CVPR. 2020.

PANOPTIC AWARE UPSAMPLING LAYER

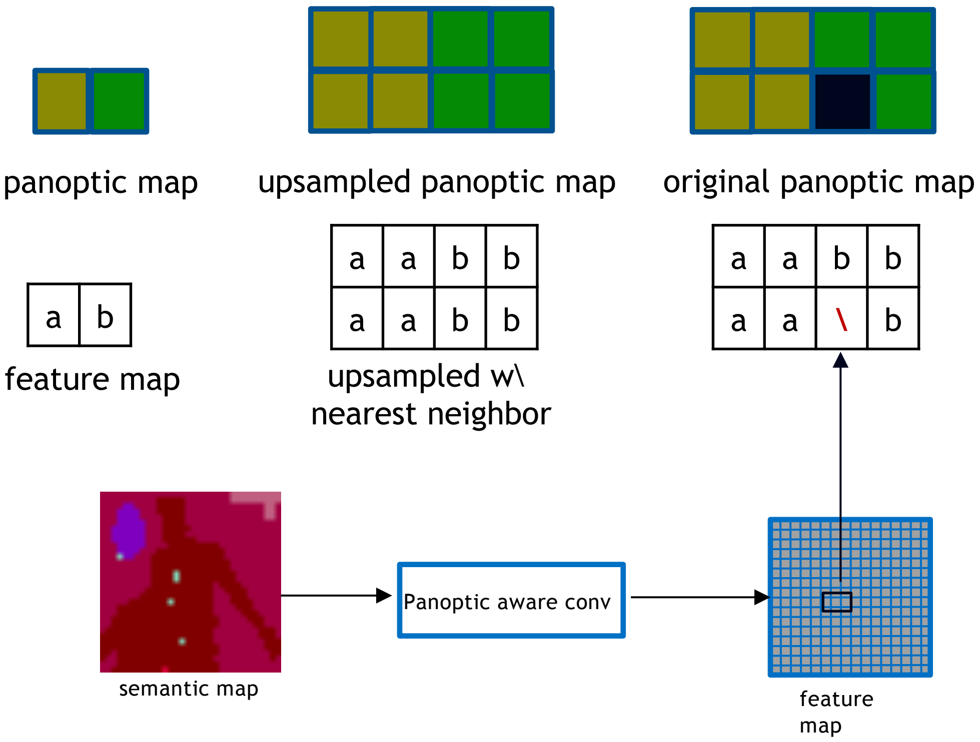


Algorithm 1 Upsampling Alignment Correction.

Initialize: $M^{correction} = 0, F'^{u} = 0,$
for $i \in [0, 2W); j \in [0, 2H)$ **do**
  if $P^u_{i,j} == P^d_{i//2,j//2}$ **then**
    $F'^{u}_{i,j} = F^d_{i//2,j//2}$
    $M^{correction}_{i,j} = 1$
  end if
end for
for $i \in [0, 2W); j \in [0, 2H)$ **do**
  if $P^u_{i,j} == P^d_{i//2+1,j//2}$ **and** $M^{correction}_{i,j} != 1$ **then**
    $F'^{u}_{i,j} = F^d_{i//2+1,j//2}$
    $M^{correction}_{i,j} = 1$
  end if
end for
for $i \in [0, 2W); j \in [0, 2H)$ **do**
  if $P^u_{i,j} == P^d_{i//2,j//2+1}$ **and** $M^{correction}_{i,j} != 1$ **then**
    $F'^{u}_{i,j} = F^d_{i//2,j//2+1}$
    $M^{correction}_{i,j} = 1$
  end if
end for
for $i \in [0, 2W); j \in [0, 2H)$ **do**
  if $P^u_{i,j} == P^d_{i//2+1,j//2+1}$ **and** $M^{correction}_{i,j} != 1$ **then**
    $F'^{u}_{i,j} = F^d_{i//2+1,j//2+1}$
    $M^{correction}_{i,j} = 1$
  end if
end for

PANOPTIC AWARE UPSAMPLING LAYER

Newly appeared classes

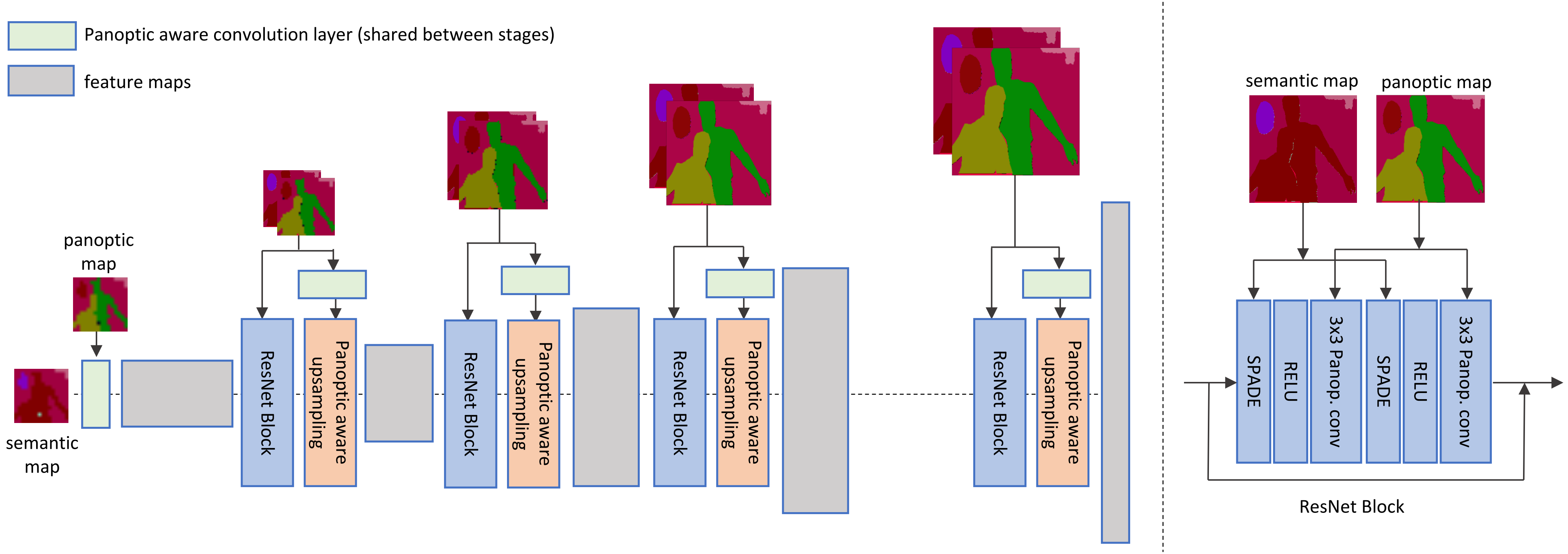


$$F'^{u}_{(i,j)} = F'^{u}_{(i,j)} + \underbrace{(1 - M^{correction}_{i,j}) * f_{holefilling}(S^{u}_{(i,j)})}_{\text{Hole Filling}}$$

$$f_{holefilling} = PanopticAwareConvolution(S^{u})$$

Dundar, Aysegul, et al. "Panoptic-based Image Synthesis." CVPR. 2020.

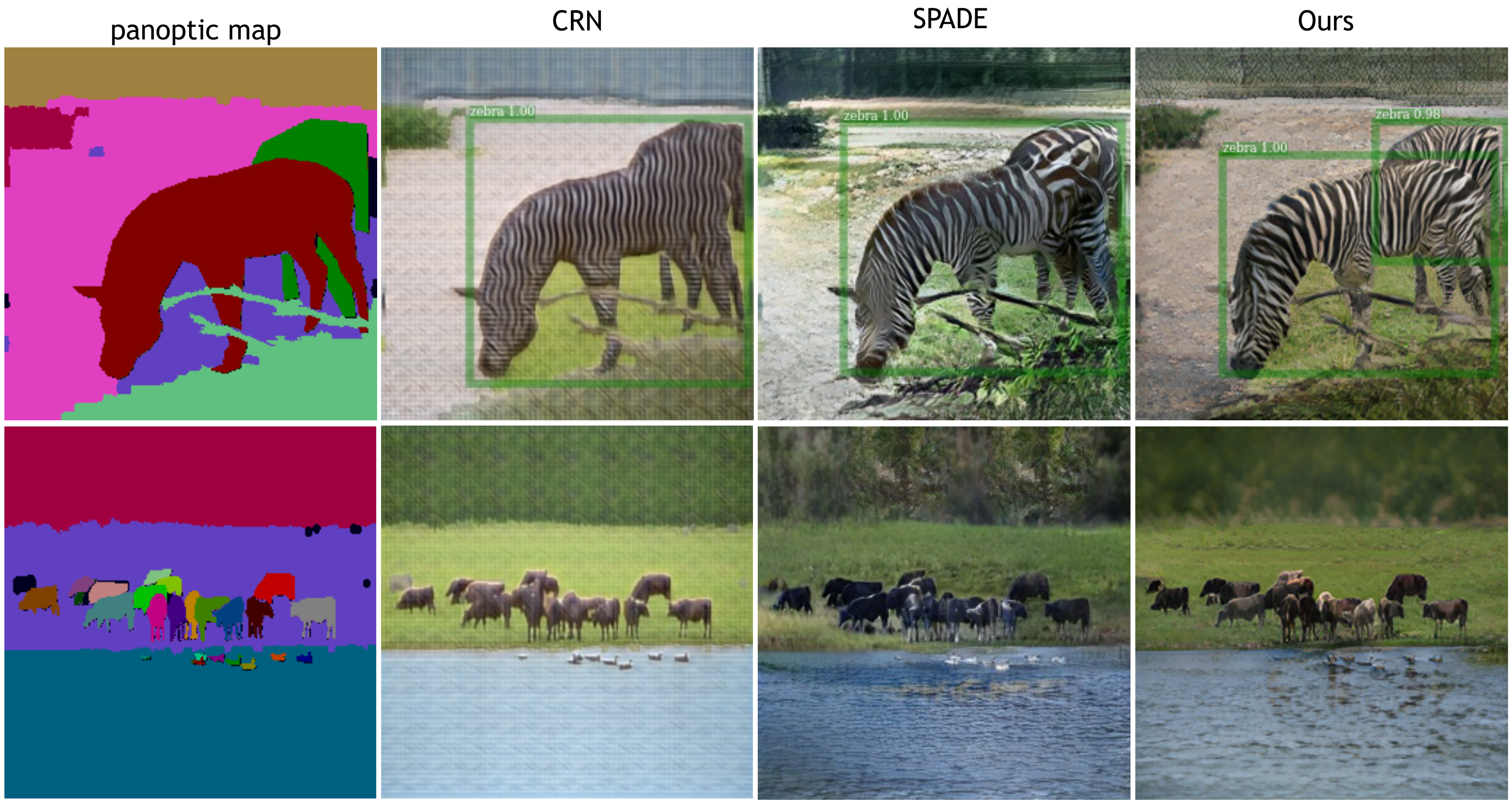
OVERALL ARCHITECTURE



Architecture based on SPADE: Semantic Image Synthesis with Spatially-Adaptive Normalization CVPR2019

Aysegul Dundar

PANOPTIC BASED IMAGE SYNTHESIS



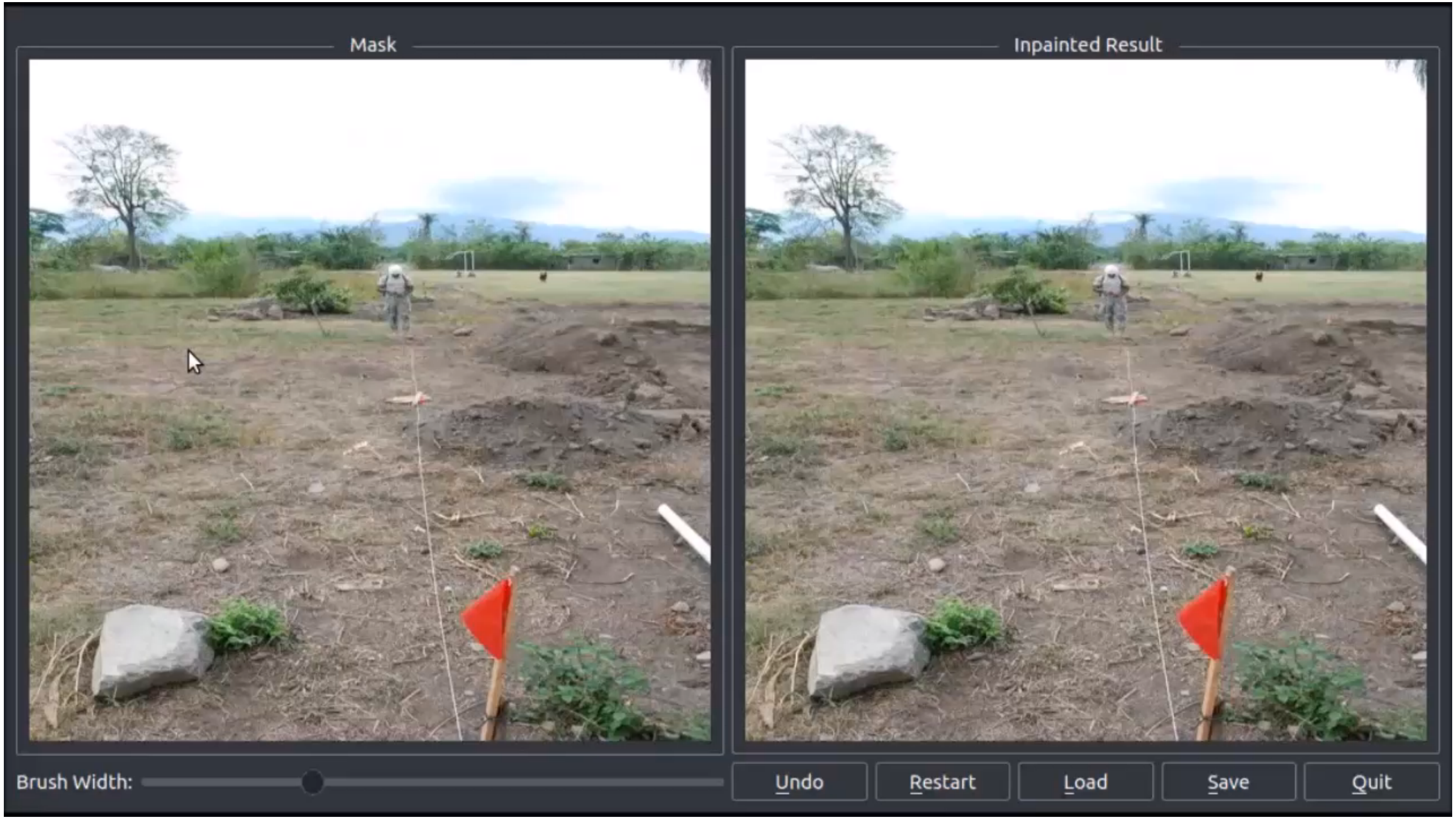
Dundar, Aysegul, et al. "Panoptic-based Image Synthesis." CVPR. 2020.

OUTLINE

- Image synthesis

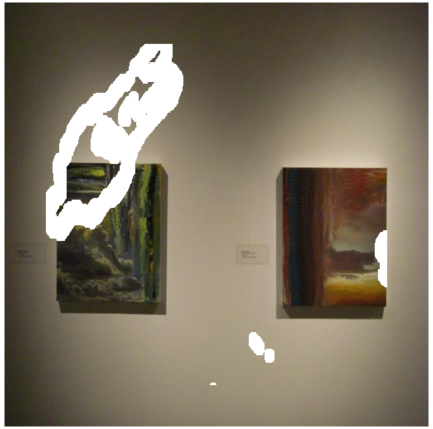
- Conditional image synthesis

- Image inpainting

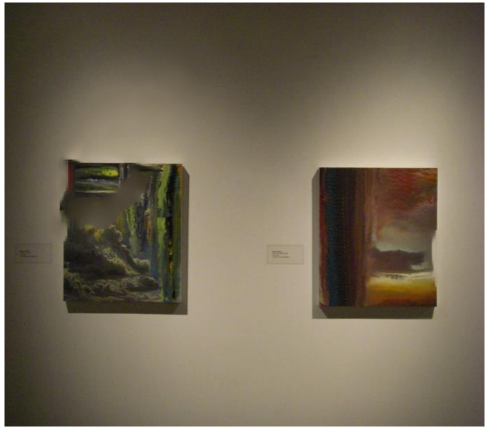


online demo: www.nvidia.com/research/inpainting/

EXISTING WORK



Input

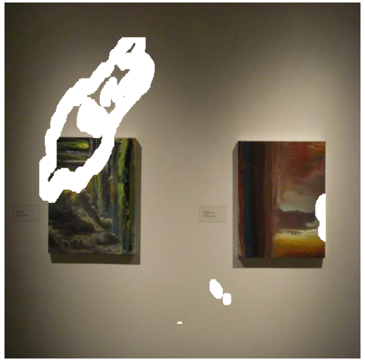


PatchMatch Result

- Adobe Photoshop
- Key idea: fill holes by iteratively searching similar patches
- Can’t create novel thing
- Slow

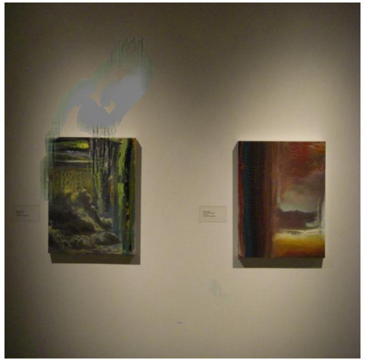
Barnes, C. et al: A randomized correspondence algorithm for structural image editing. TOG 2009. (PatchMatch)

DEEP LEARNING-BASED APPROACH

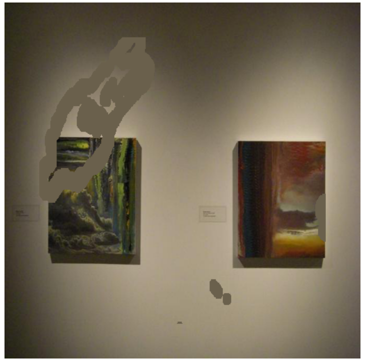


- initialization for hole pixels (pixels with missing values)
- set initial values for holes, e.g. 0 or median values (127.5)
- treat original non-hole pixels and initial hole pixels equally -> confuse the network

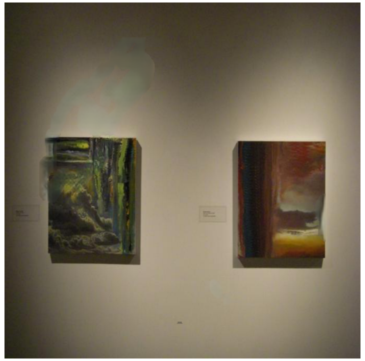
initialize using median value



corresponding output



initialize using mean values



corresponding output

Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." ECCV. 2018.

PARTIAL CONVOLUTION FOR INPAINTING

X: input

1	1	1
1	0	0
1	0	0

M: mask
(1 means nonhole、
0 means hole)

Principles

- Ignore pixels in the hole
- Only use non-hole pixels

Mask/hole-aware convolution

- Re-normalize result using mask size
- Update mask as receptive field becomes larger

Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." ECCV. 2018.

PARTIAL CONVOLUTION FOR INPAINTING

X: input

1	1	1
1	0	0
1	0	0

M: mask
(1 means nonhole、
0 means hole)

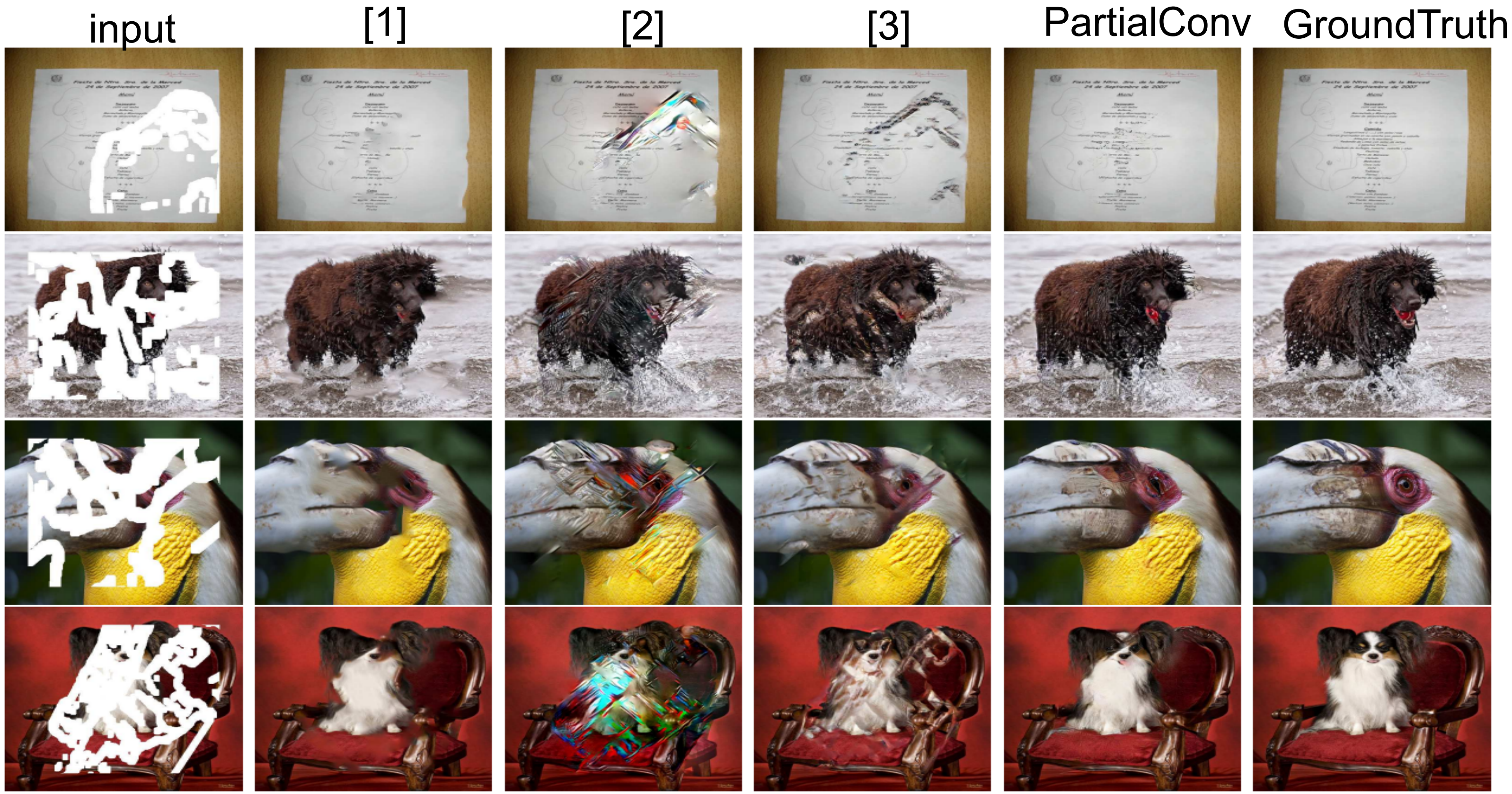
pixel update : $X_i' = W^T(X \circ M) \cdot \frac{K^2}{sum(M)} + b$

mask update : $M_i' = \begin{cases} 1 & if\ sum(M) > 0 \\ 0 & if\ sum(M) = 0 \end{cases}$

mask updating after several partial conv layers
As the receptive field becomes larger, M will all become 1

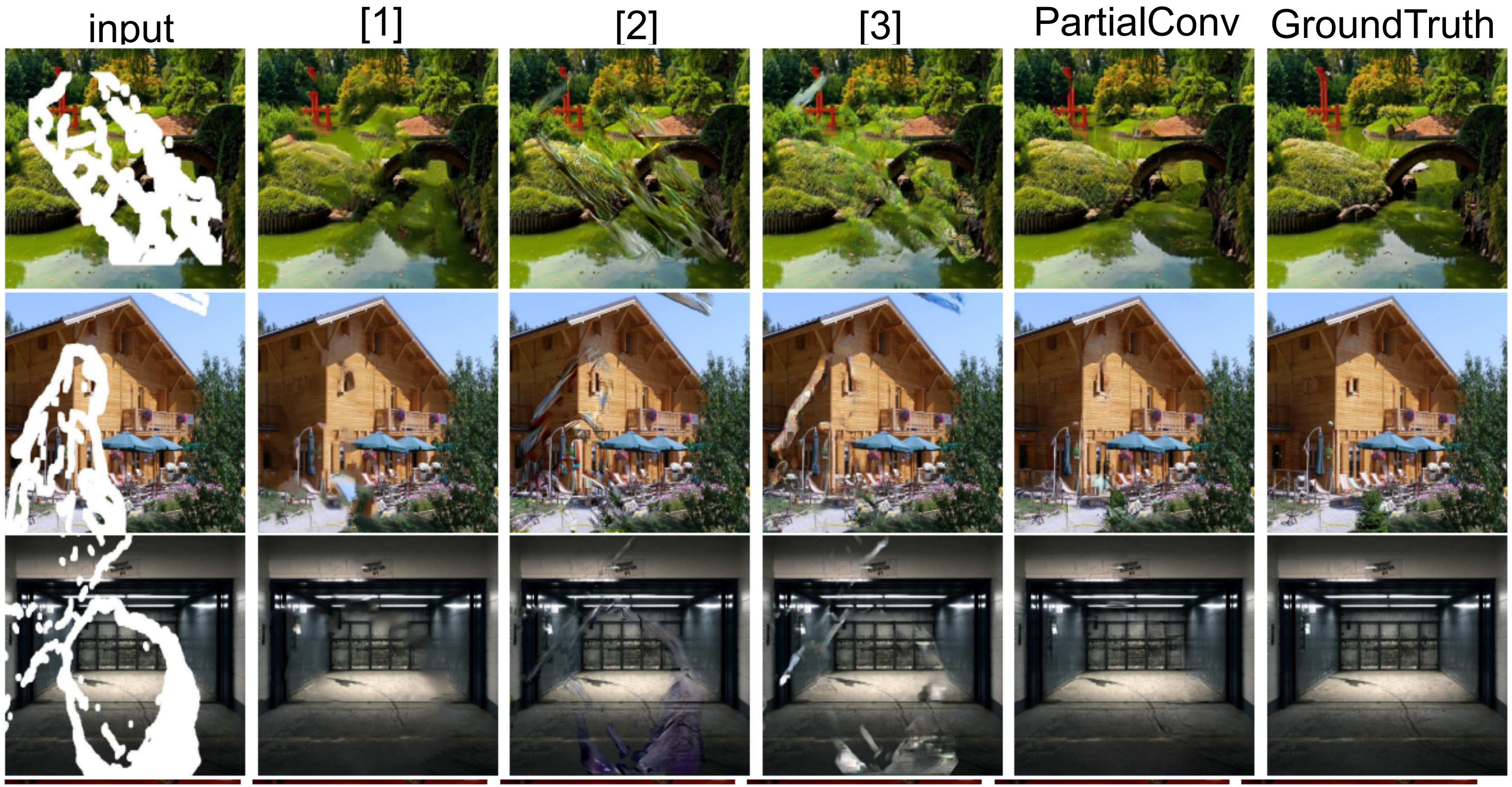
Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." ECCV. 2018.

RESULTS



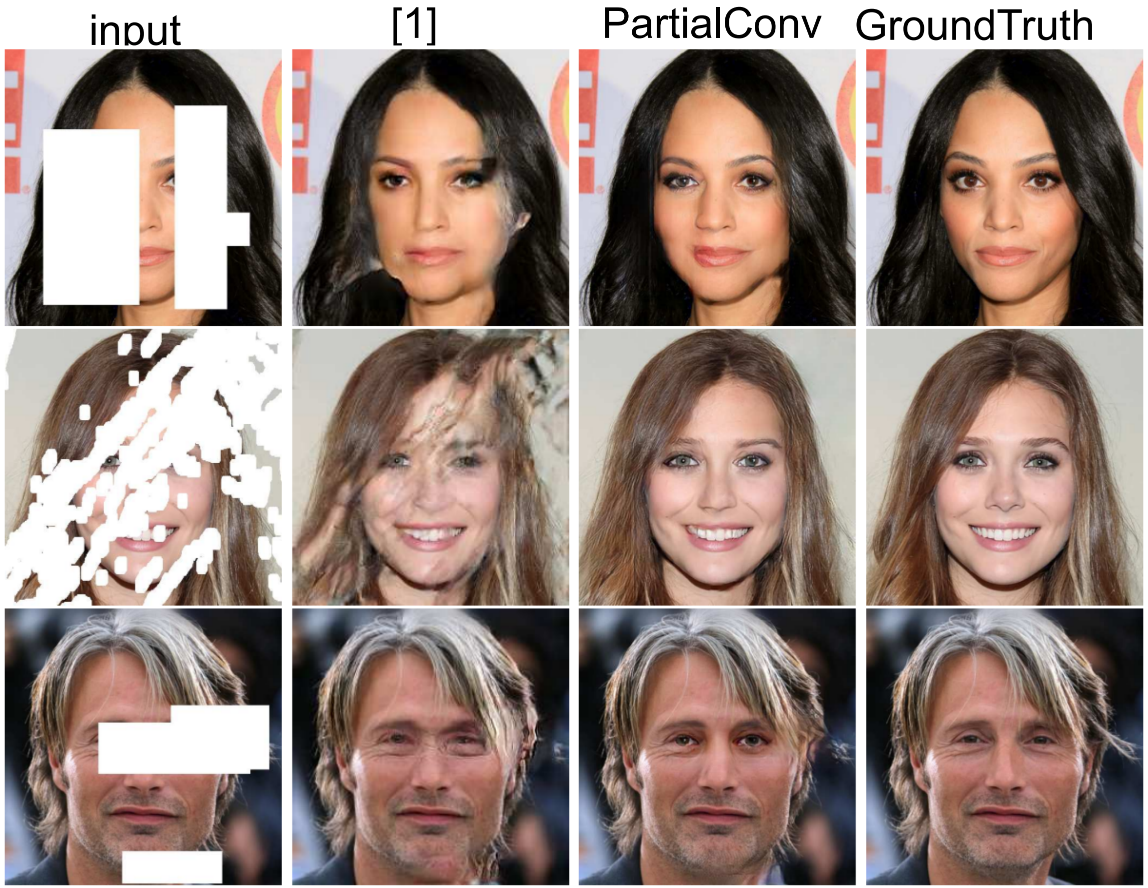
1. Barnes, C. et al: A randomized correspondence algorithm for structural image editing. TOG 2009. (PatchMatch)
2. Iizuka, S., Simo-Serra, E., Ishikawa, H.: Globally and locally consistent image completion. TOG 2017.
3. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., Huang, T.S.: Generative image inpainting with contextual attention. CVPR 2018.
4. Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." ECCV. 2018.

RESULTS



1. Barnes, C. et al: A randomized correspondence algorithm for structural image editing. TOG 2009. (PatchMatch)
2. Iizuka, S., Simo-Serra, E., Ishikawa, H.: Globally and locally consistent image completion. TOG 2017.
3. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., Huang, T.S.: Generative image inpainting with contextual attention. CVPR 2018.
4. Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." ECCV. 2018.

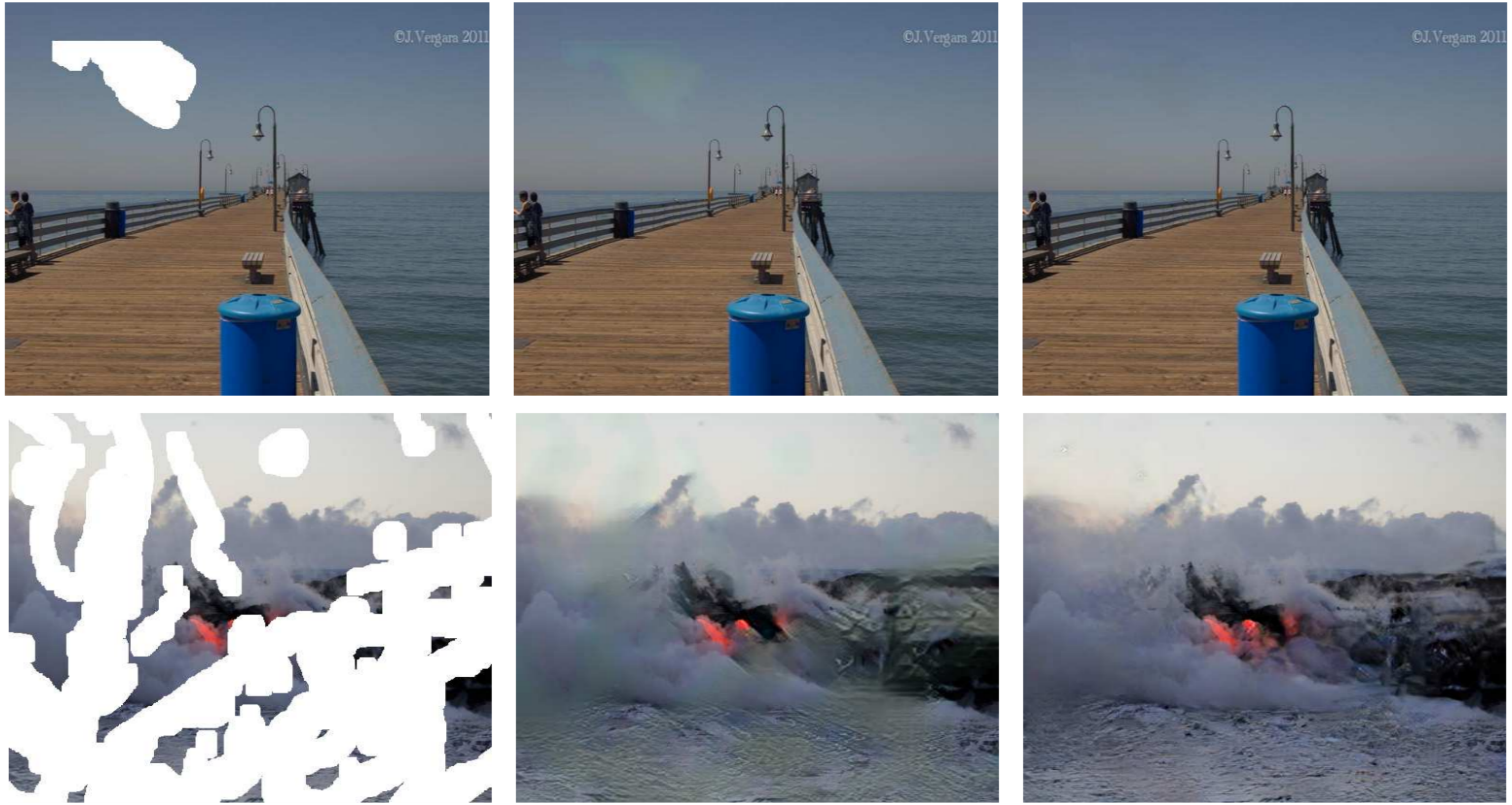
RESULTS



1. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., Huang, T.S.: Generative image inpainting with contextual attention. CVPR 2018.
2. Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." ECCV. 2018.

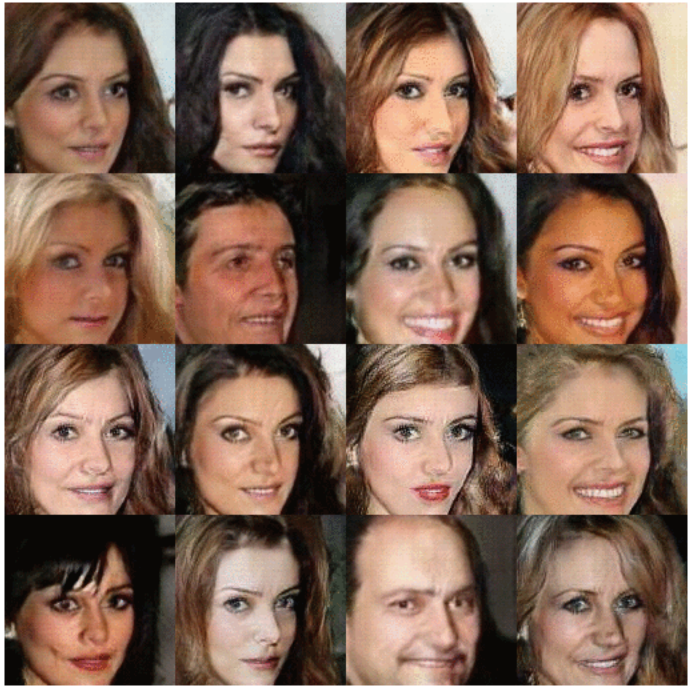
COMPARE WITH TYPICAL CONVOLUTION

input | Conv | PartialConv



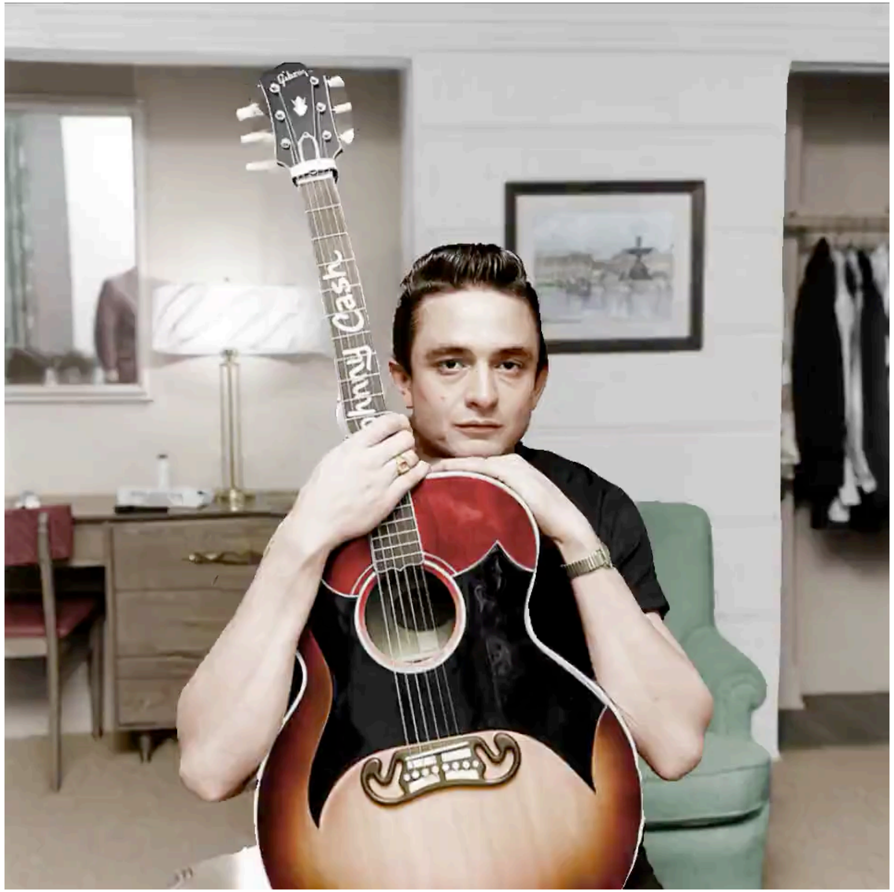
Liu, Guilin, et al. "Image inpainting for irregular holes using partial convolutions." ECCV. 2018.

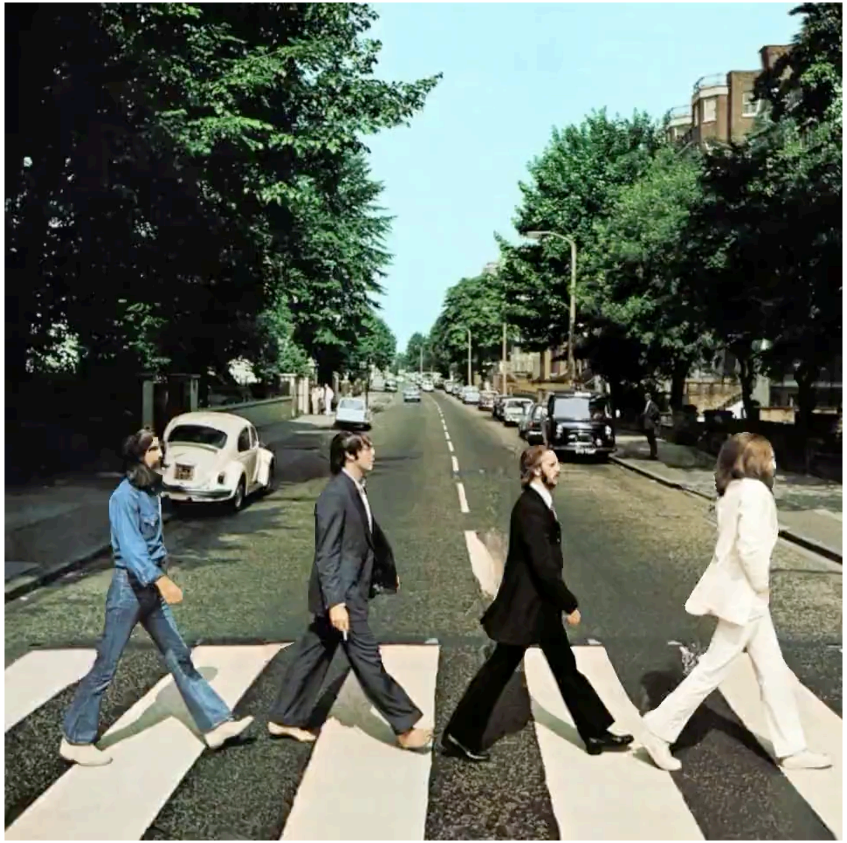
FUTURE DIRECTIONS



Nguyen-Phuoc, Thu, et al. "Hologan: Unsupervised learning of 3d representations from natural images." ICCV. 2019.

FUTURE DIRECTIONS





Shih, Meng-Li, et al. "3D Photography using Context-aware Layered Depth Inpainting." CVPR. 2020.

THANK YOU