Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

Aitor Alvarez-Gila, Joost van de Weijer, Estibaliz Garrote
aitorshuffle.github.io

Venice, 2017/10/23

The task

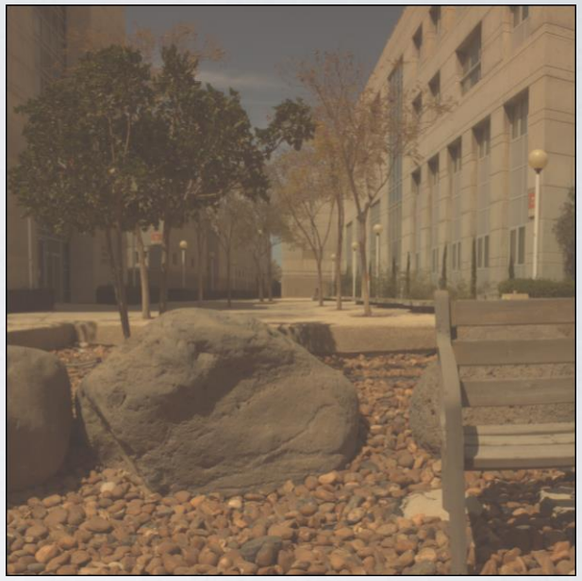
Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

The task

Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

The task

(Hyper)-Spectral Image reconstruction from RGB



RGB

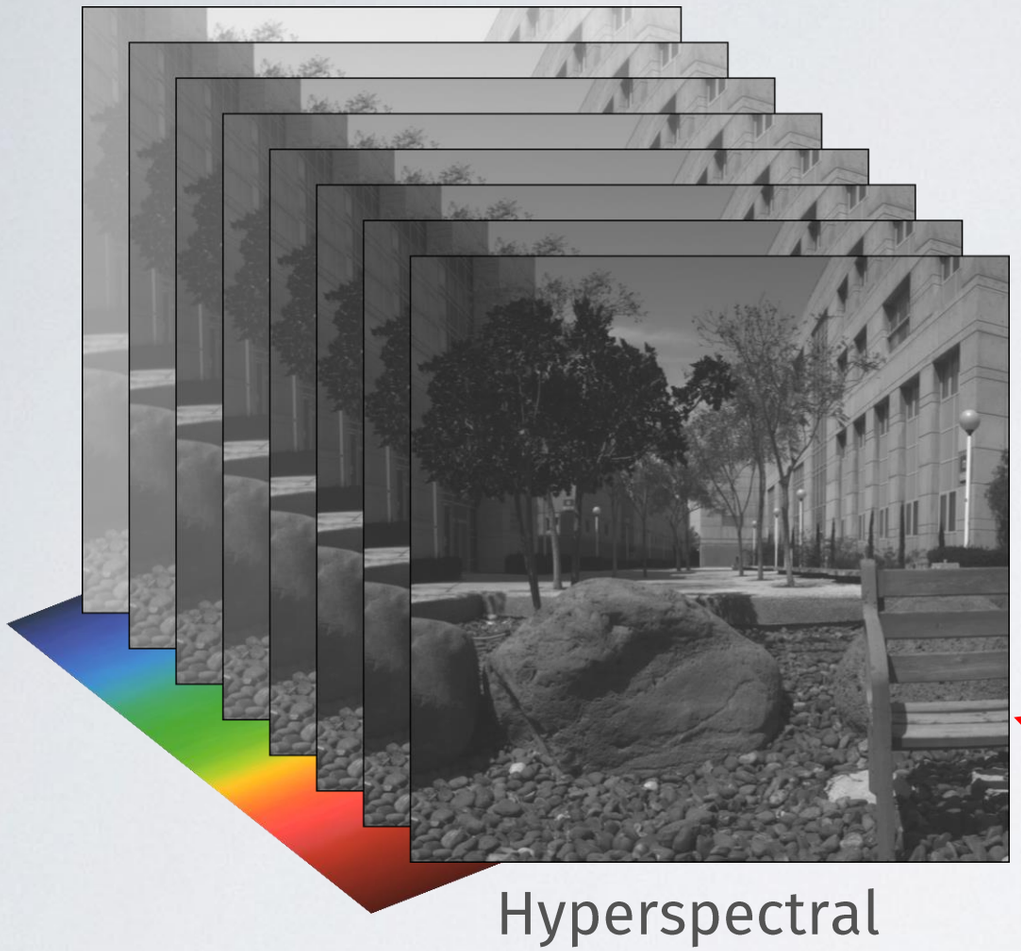
- Learn a image to image mapping s.t., for each pixel:

$$\mathbb{R}^3 \rightarrow \mathbb{R}^c \text{ with } c >> 3$$

- In particular: $c = 31(400 : 10 : 700nm)$
- Heavily underconstrained (e.g. metamers), non-linear problem

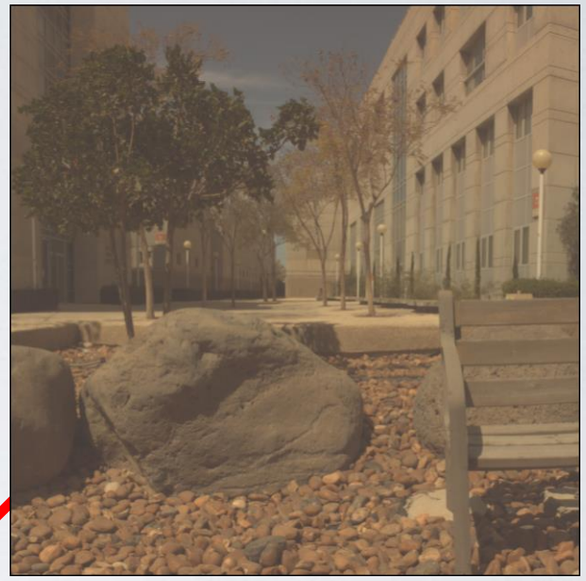
The task

(Hyper)-Spectral Image reconstruction from RGB



$HS \longleftarrow RGB$

Our task



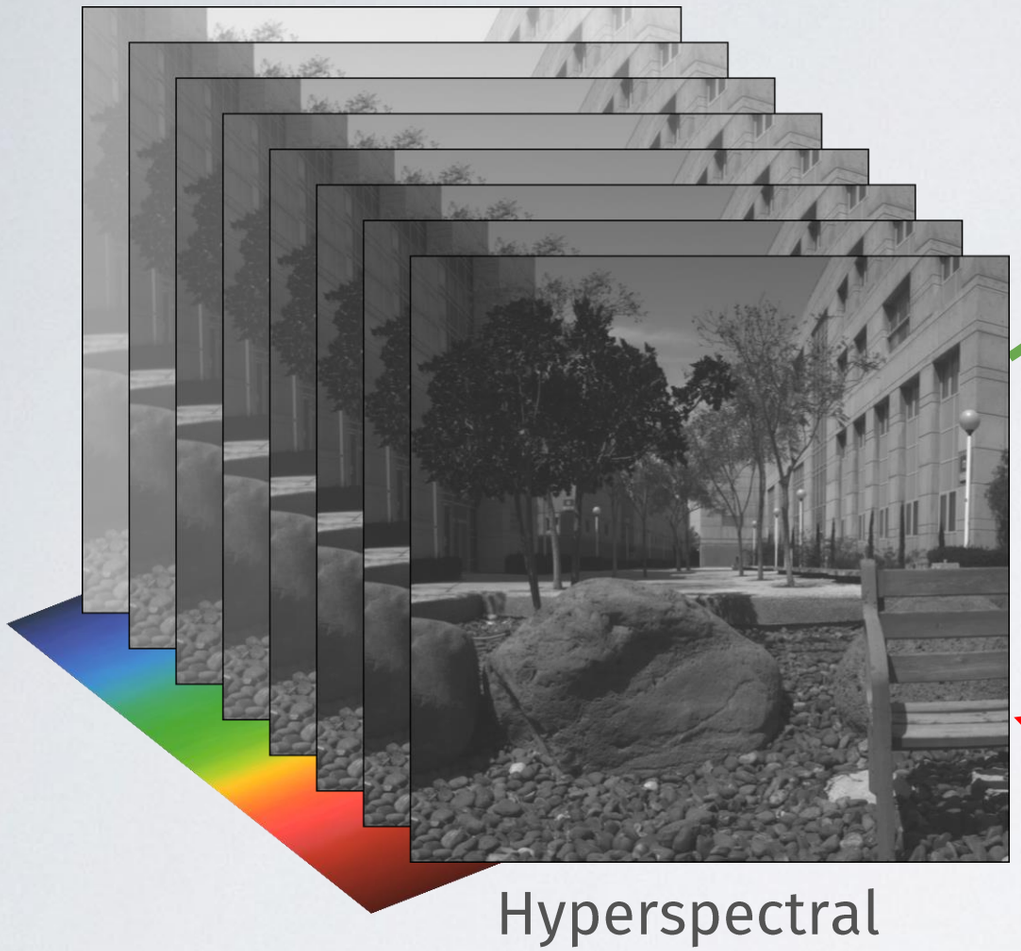
- Learn a image to image mapping s.t., for each pixel:

$$\mathbb{R}^3 \rightarrow \mathbb{R}^c \text{ with } c >> 3$$

- In particular: $c = 31(400 : 10 : 700nm)$
- Heavily underconstrained (e.g. metamers), non-linear problem

The task

(Hyper)-Spectral Image reconstruction from RGB

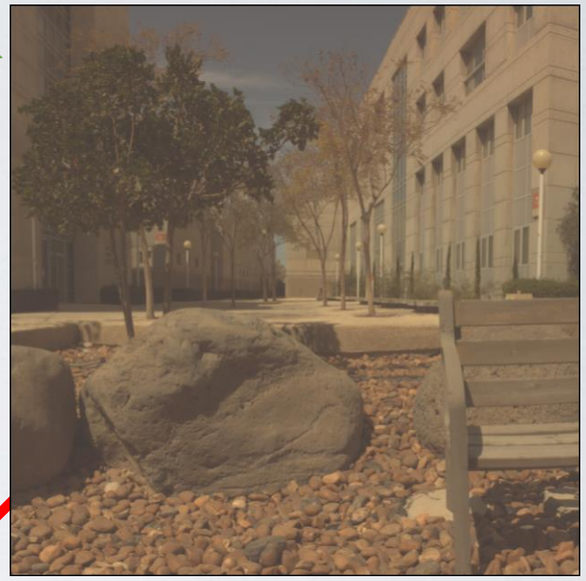


(Deterministic)

$$HS \longrightarrow RGB$$

$$HS \longleftarrow RGB$$

Our task



- Learn a image to image mapping s.t., for each pixel:

$$\mathbb{R}^3 \rightarrow \mathbb{R}^c \text{ with } c >> 3$$

- In particular: $c = 31(400 : 10 : 700nm)$
- Heavily underconstrained (e.g. metamers), non-linear problem

Color
Camera Image formation

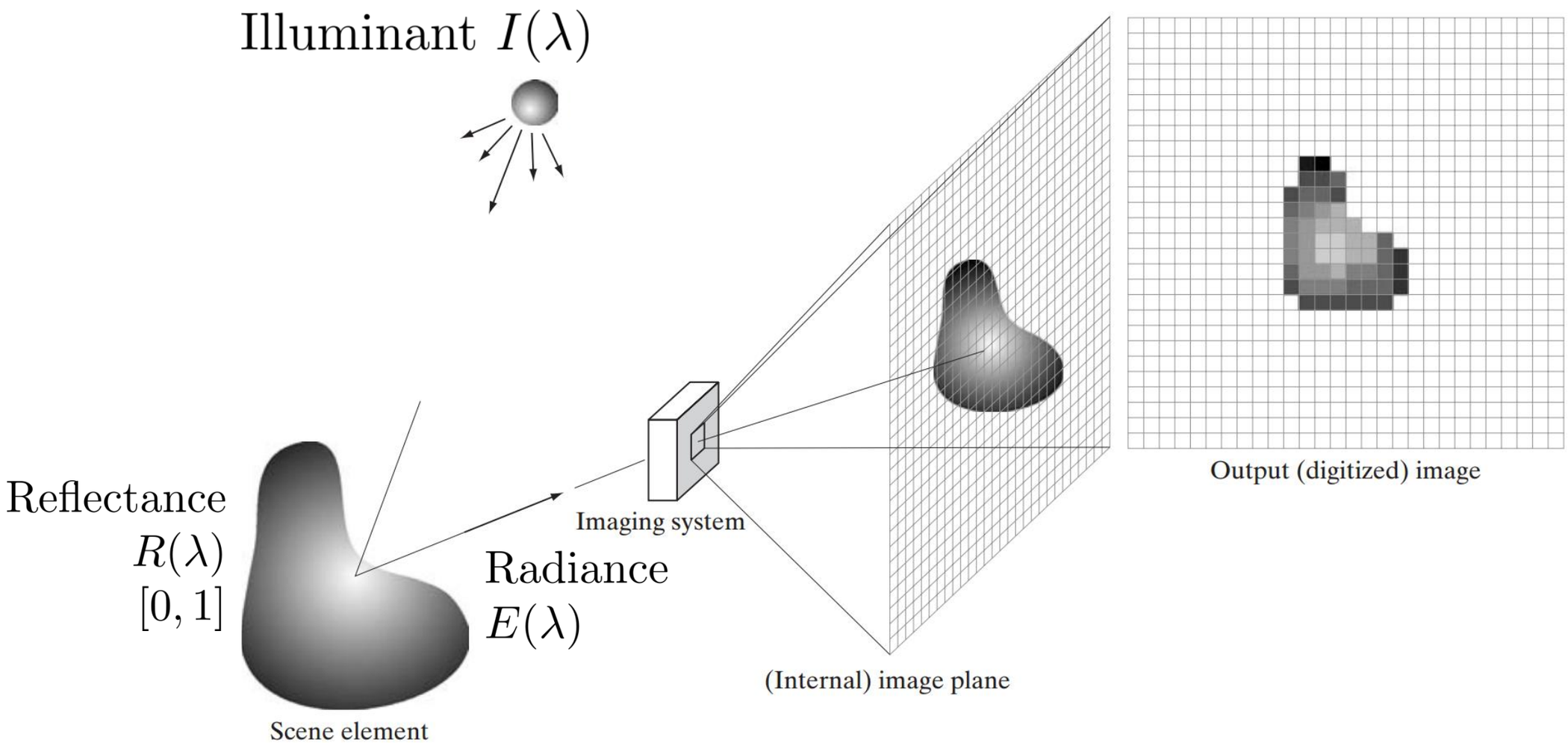
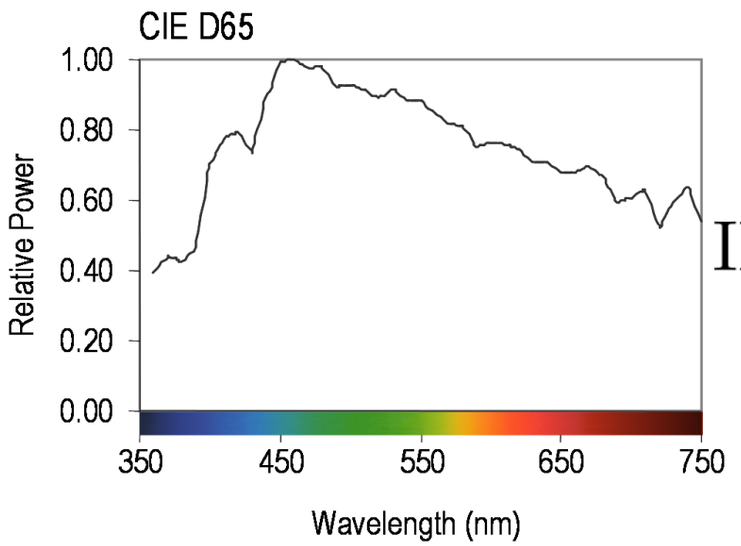


Image:
R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3 edition. Upper Saddle River, N.J: Pearson, 2007.
M. S. Brown, "Understanding the In-Camera Image Processing Pipeline for Computer Vision," IEEE CVPR - Tutorial, 2016.

Color

Camera Image formation



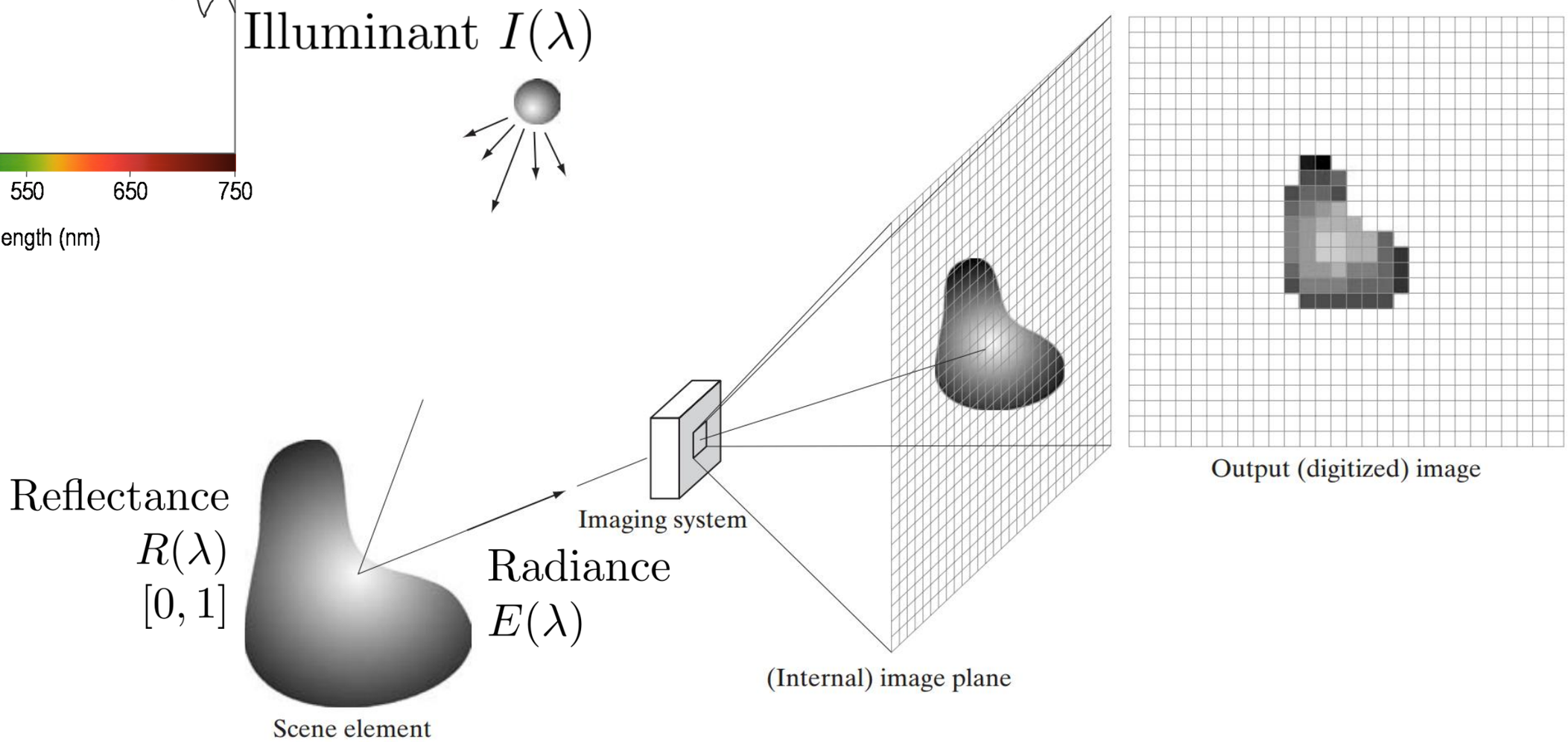
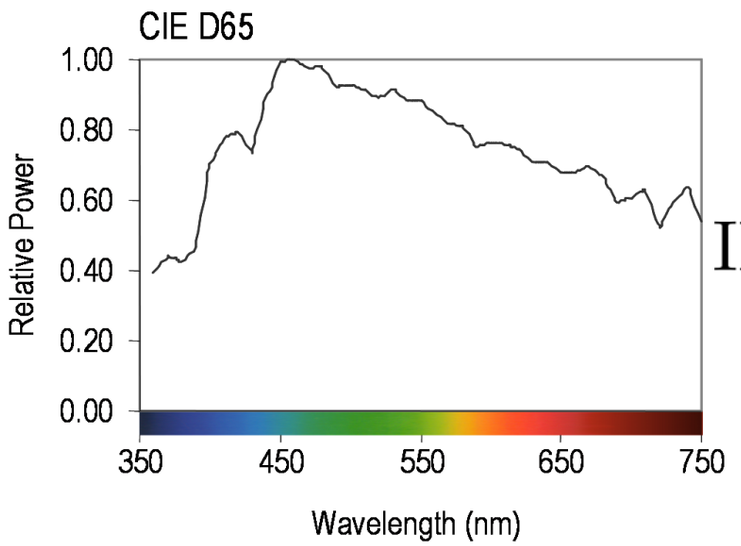


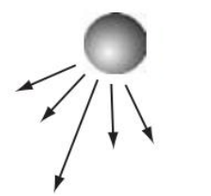
Image:
R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3 edition. Upper Saddle River, N.J: Pearson, 2007.
M. S. Brown, "Understanding the In-Camera Image Processing Pipeline for Computer Vision," IEEE CVPR - Tutorial, 2016.

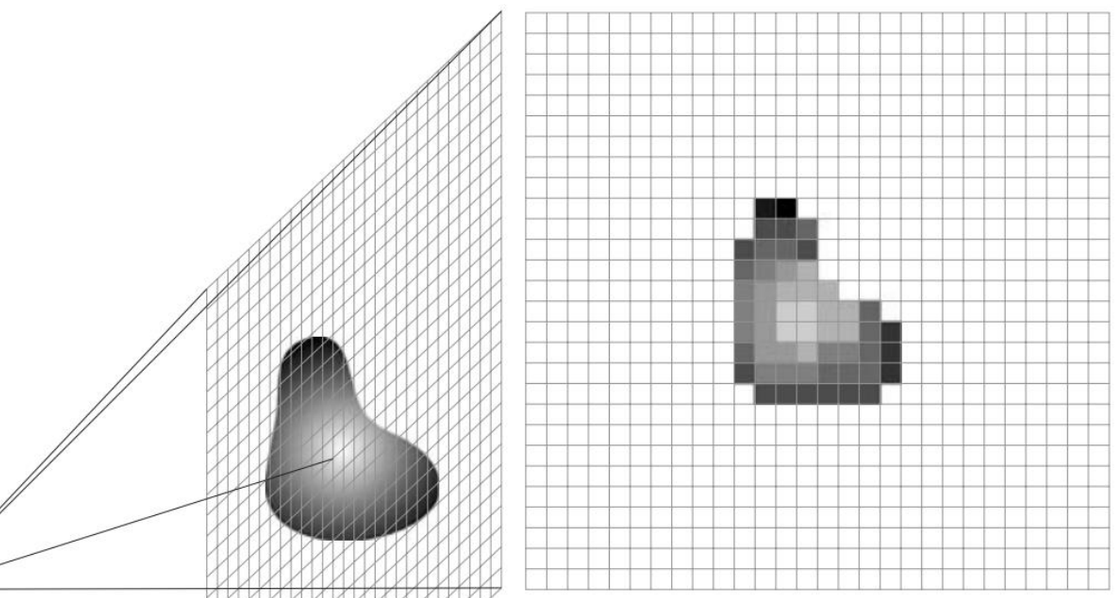
Color

Camera Image formation



Illuminant $I(\lambda)$



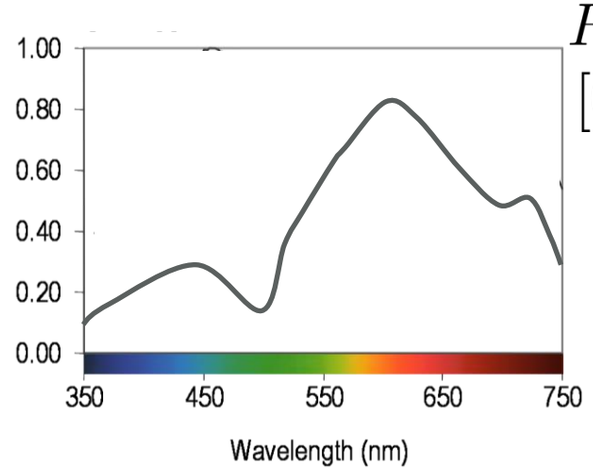


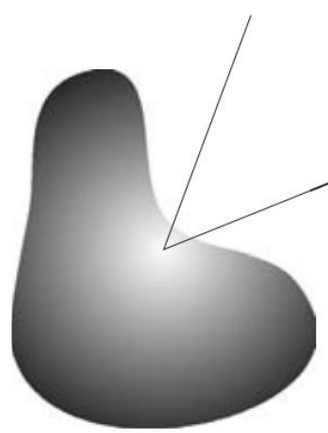
Output (digitized) image

Imaging system

(Internal) image plane

Reflectance
$R(\lambda)$
$[0, 1]$





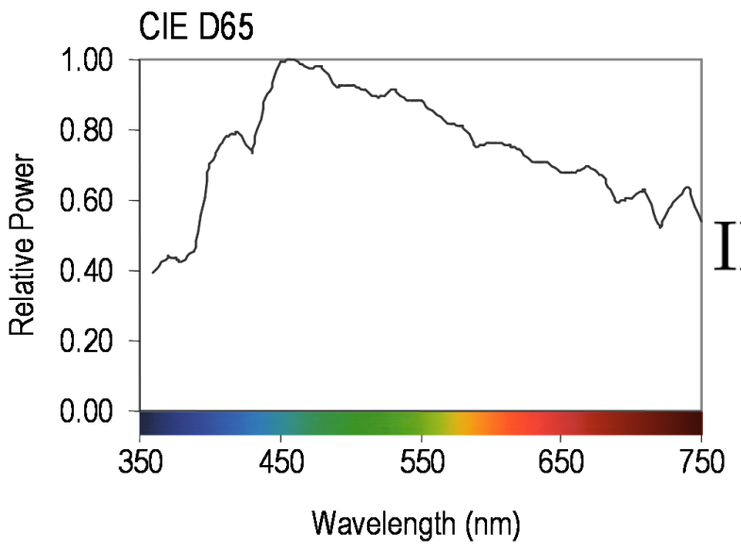
Scene element

Radiance
$E(\lambda)$

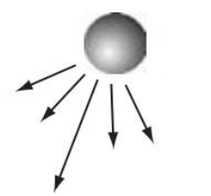
Image:
R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3 edition. Upper Saddle River, N.J: Pearson, 2007.
M. S. Brown, "Understanding the In-Camera Image Processing Pipeline for Computer Vision," IEEE CVPR - Tutorial, 2016.

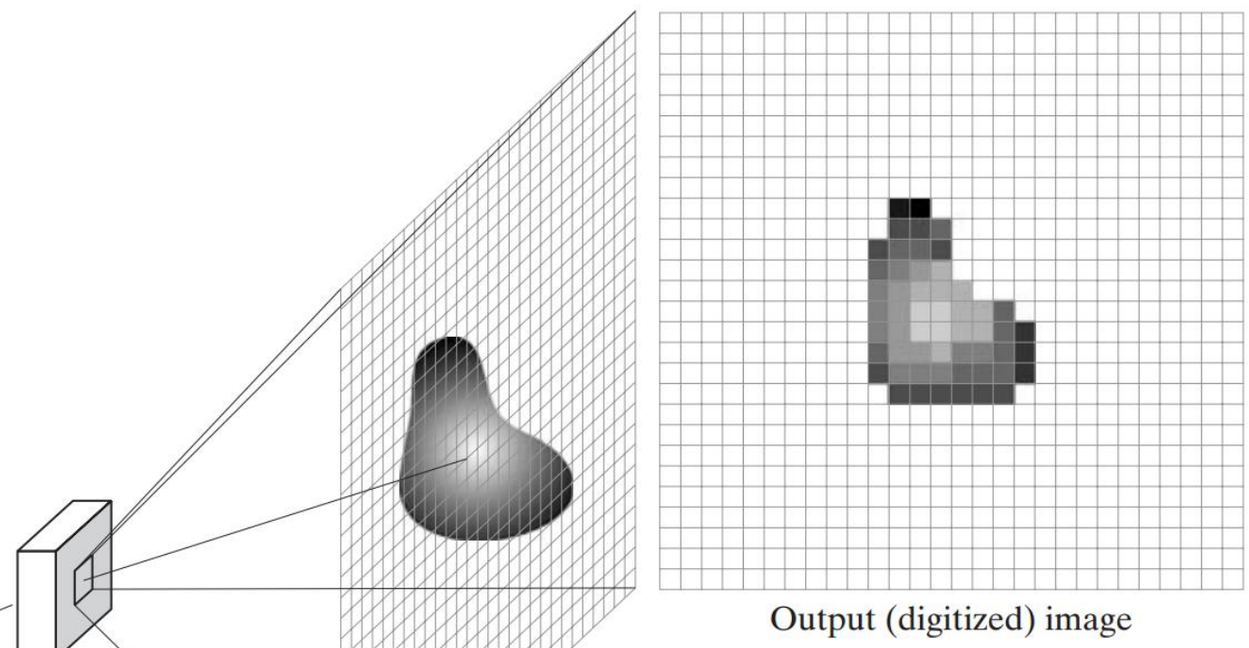
Color

Camera Image formation

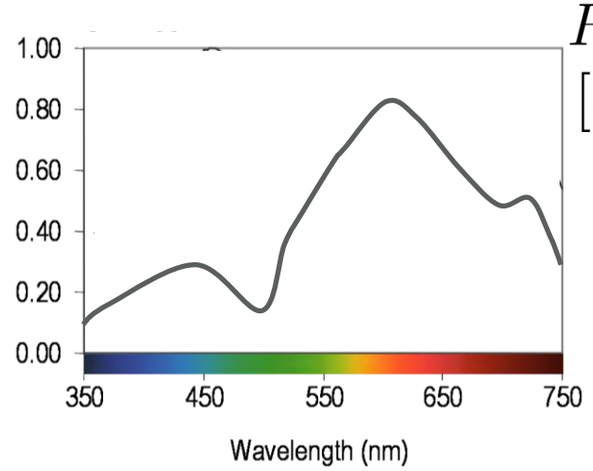


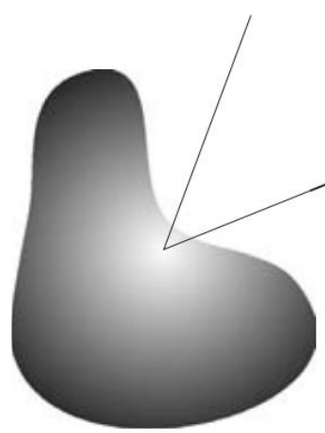
Illuminant $I(\lambda)$





Reflectance
$R(\lambda)$
$[0, 1]$





Scene element

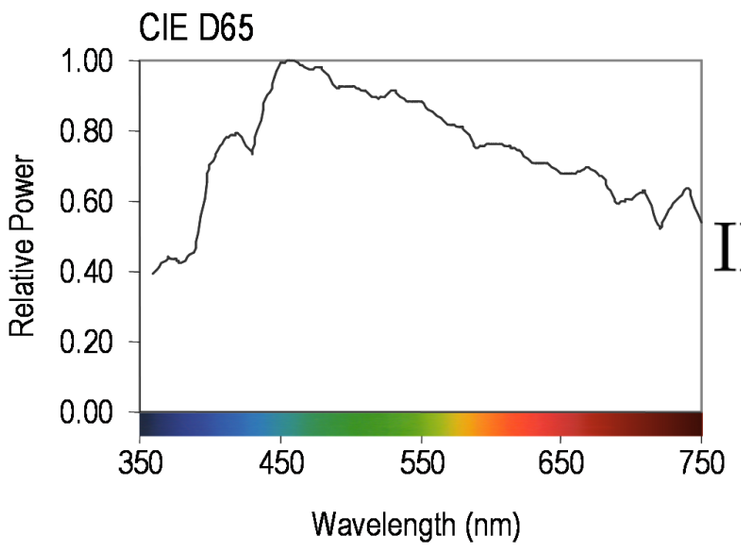
Radiance
$E(\lambda)$

$$E(\lambda) = I(\lambda) \times R(\lambda)$$

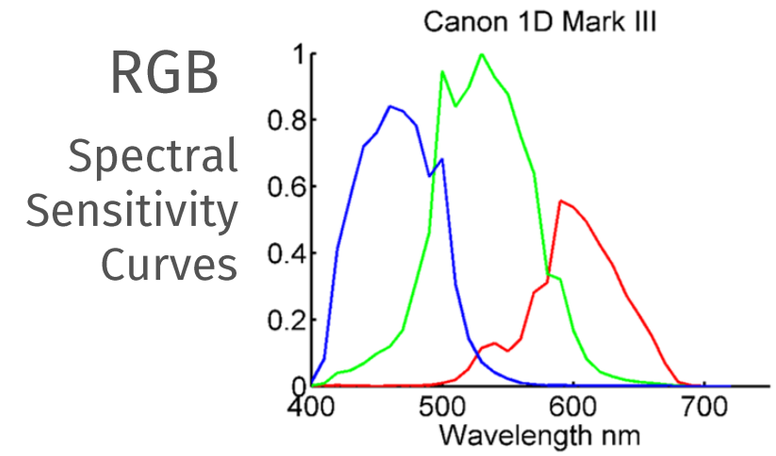
Image:
R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3 edition. Upper Saddle River, N.J: Pearson, 2007.
M. S. Brown, "Understanding the In-Camera Image Processing Pipeline for Computer Vision," IEEE CVPR - Tutorial, 2016.

Color

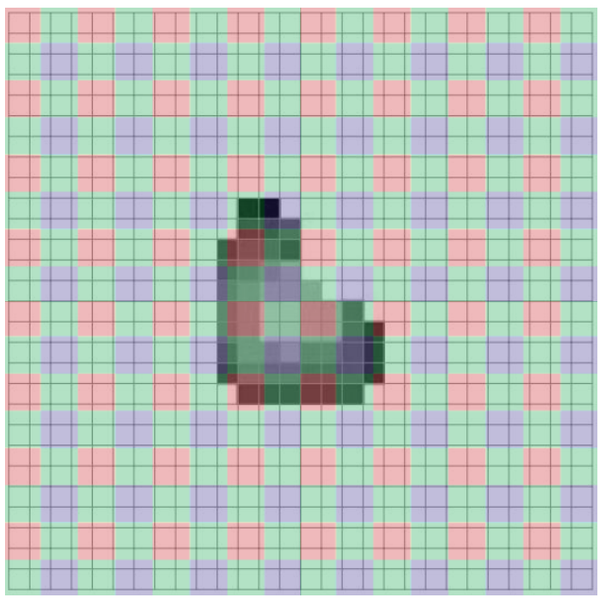
Camera Image formation



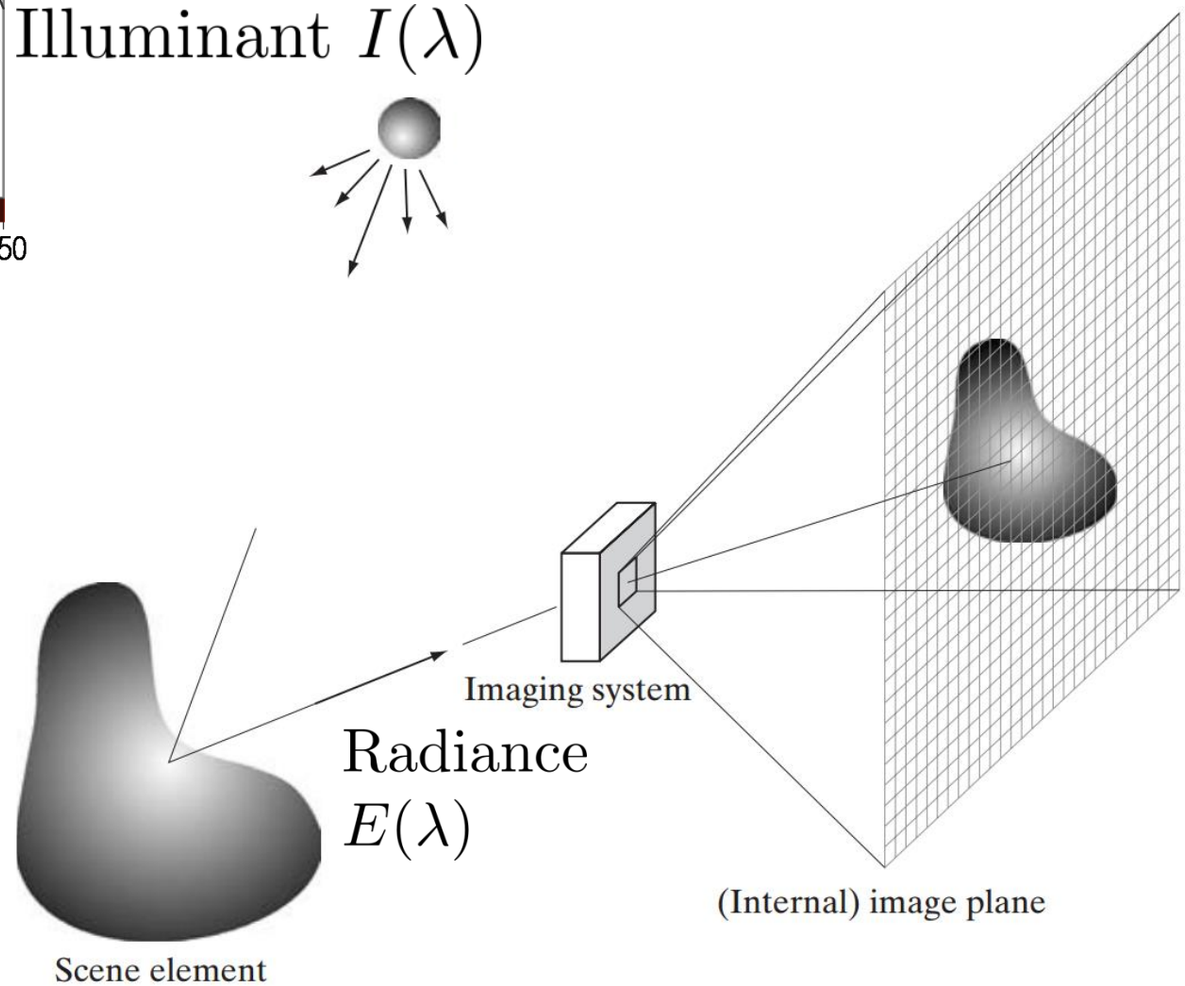
Illuminant $I(\lambda)$



RGB Spectral Sensitivity Curves

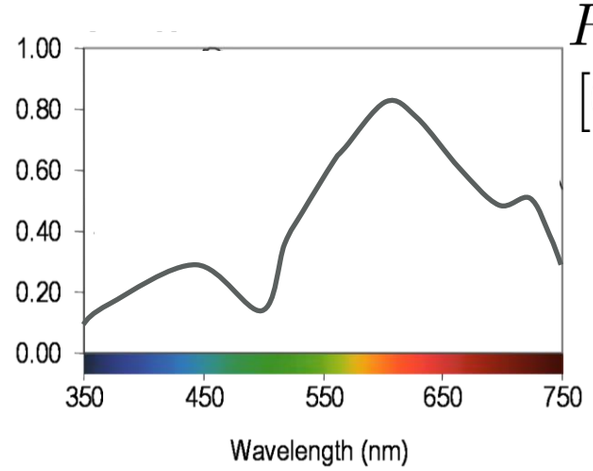


Color Filter Array



Reflectance $R(\lambda)$ $[0, 1]$

Radiance $E(\lambda)$



$$E(\lambda) = I(\lambda) \times R(\lambda)$$

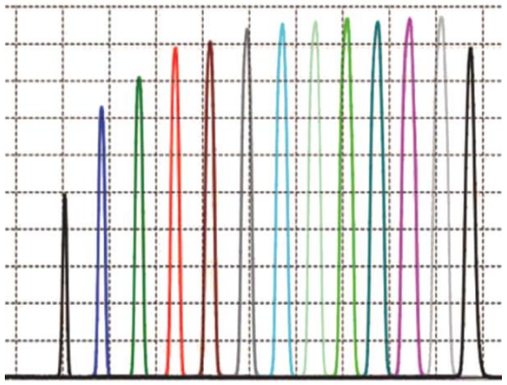
Image:
R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3 edition. Upper Saddle River, N.J: Pearson, 2007.
M. S. Brown, "Understanding the In-Camera Image Processing Pipeline for Computer Vision," IEEE CVPR - Tutorial, 2016.

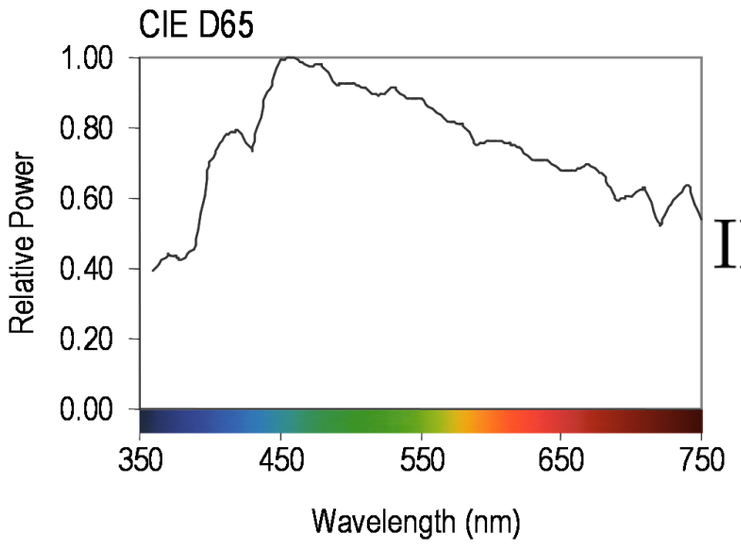
Color

Camera Image formation

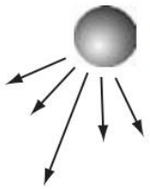
Hyperspectral

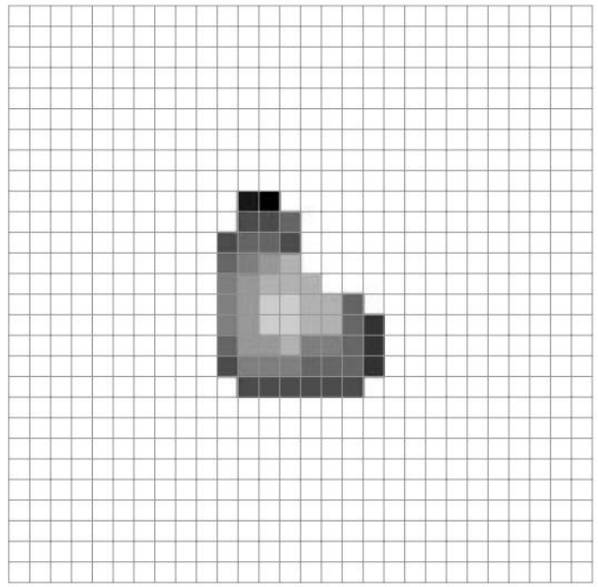
Spectral Sensitivity Curves





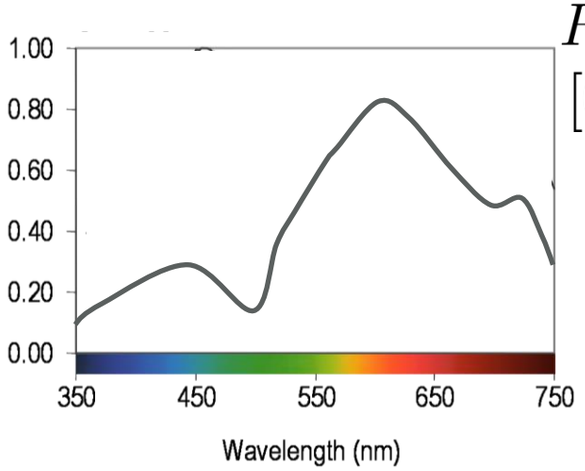
Illuminant $I(\lambda)$

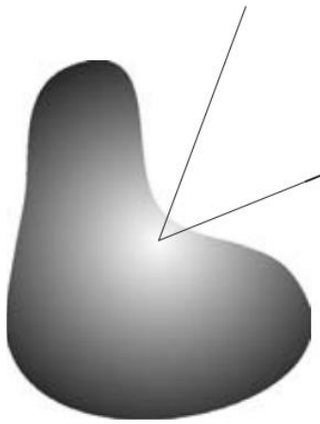




Color Filter Array

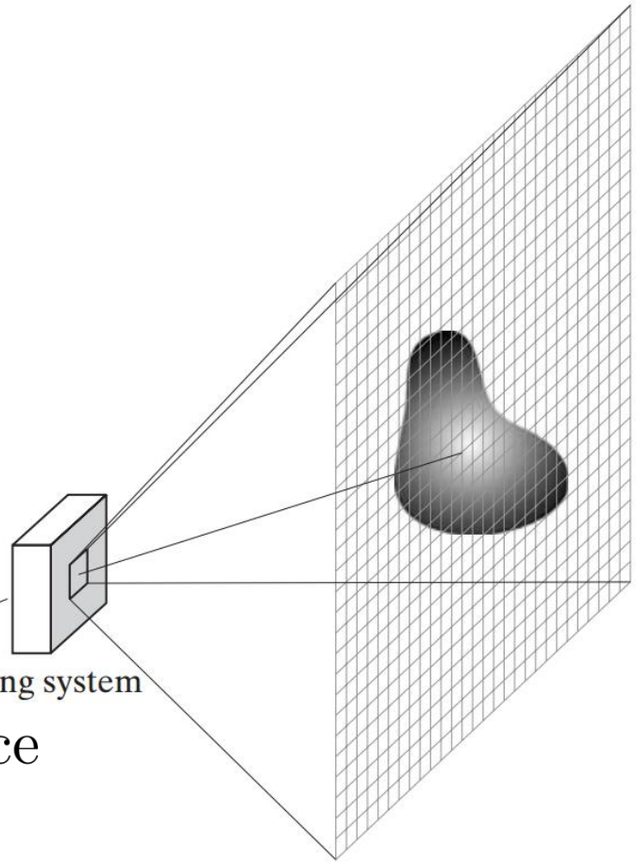
Reflectance $R(\lambda)$ $[0, 1]$





Scene element

Radiance $E(\lambda)$



Imaging system

(Internal) image plane

$$E(\lambda) = I(\lambda) \times R(\lambda)$$

Image:
R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3 edition. Upper Saddle River, N.J: Pearson, 2007.
M. S. Brown, "Understanding the In-Camera Image Processing Pipeline for Computer Vision," IEEE CVPR - Tutorial, 2016.

Related work

Spectral reconstruction

- From RGB + extra help:
  - Low-res HS image

[Cao2011] X. Cao, X. Tong, Q. Dai, and S. Lin, "High resolution multispectral video capture with a hybrid camera system," in CVPR 2011, 2011, pp. 297–304.

  - Multiplexed light

[Park2007] J. I. Park, M. H. Lee, M. D. Grossberg, and S. K. Nayar, "Multispectral Imaging Using Multiplexed Illumination," in 2007 IEEE 11th International Conference on Computer Vision, 2007, pp. 1–8.

[Parmar2008] M. Parmar, S. Lansel, and B. A. Wandell, "Spatio-spectral reconstruction of the multispectral datacube using sparse recovery," in 2008 15th IEEE International Conference on Image Processing, 2008, pp. 473–476.

[Goel2015] M. Goel et al., "HyperCam: Hyperspectral Imaging for Ubiquitous Computing Applications," in Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York, NY, USA, 2015, pp. 145–156.

- No spatial info considered

[Nguyen2014] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, "Training-Based Spectral Reconstruction from a Single RGB Image," in Computer Vision ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Springer International Publishing, 2014, pp. 186–201.

[Arad2016] B. Arad and O. Ben-Shahar, "Sparse Recovery of Hyperspectral Signal from Natural RGB Images," in *ECCV*, 2016, pp. 19–34.

- Leveraging spatial context:

[Robles-Kelly2015] A. Robles-Kelly, "Single Image Spectral Reconstruction for Multimedia Applications," in ACM Multimedia, 2015.

[Galliani2017] S. Galliani, C. Lanaras, D. Marmanis, E. Baltsavias, and K. Schindler, "Learned Spectral Super-Resolution," arXiv:1703.09470 [cs], Mar. 2017.

Related work

Spectral reconstruction

- From RGB + extra help:
  - Low-res HS image

[Cao2011] X. Cao, X. Tong, Q. Dai, and S. Lin, "High resolution multispectral video capture with a hybrid camera system," in CVPR 2011, 2011, pp. 297–304.

  - Multiplexed light

[Park2007] J. I. Park, M. H. Lee, M. D. Grossberg, and S. K. Nayar, "Multispectral Imaging Using Multiplexed Illumination," in 2007 IEEE 11th International Conference on Computer Vision, 2007, pp. 1–8.

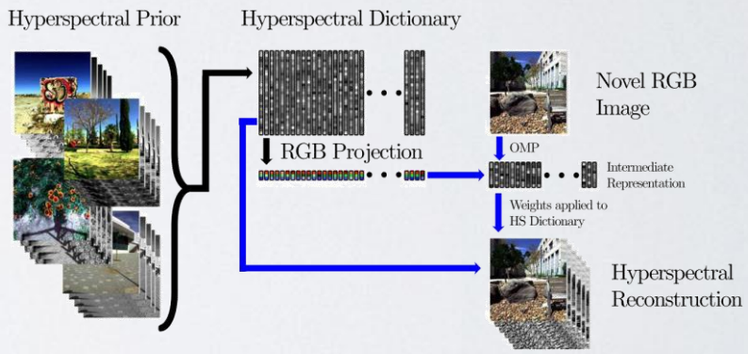
[Parmar2008] M. Parmar, S. Lansel, and B. A. Wandell, "Spatio-spectral reconstruction of the multispectral datacube using sparse recovery," in 2008 15th IEEE International Conference on Image Processing, 2008, pp. 473–476.

[Goel2015] M. Goel et al., "HyperCam: Hyperspectral Imaging for Ubiquitous Computing Applications," in Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York, NY, USA, 2015, pp. 145–156.

- No spatial info considered

[Nguyen2014] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, "Training-Based Spectral Reconstruction from a Single RGB Image," in Computer Vision ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Springer International Publishing, 2014, pp. 186–201.

[Arad2016] B. Arad and O. Ben-Shahar, "Sparse Recovery of Hyperspectral Signal from Natural RGB Images," in *ECCV*, 2016, pp. 19–34.



- Leveraging spatial context:

[Robles-Kelly2015] A. Robles-Kelly, "Single Image Spectral Reconstruction for Multimedia Applications," in ACM Multimedia, 2015.

[Galliani2017] S. Galliani, C. Lanaras, D. Marmanis, E. Baltsavias, and K. Schindler, "Learned Spectral Super-Resolution," arXiv:1703.09470 [cs], Mar. 2017.

Related work

Spectral reconstruction

- From RGB + extra help:
  - Low-res HS image

[Cao2011] X. Cao, X. Tong, Q. Dai, and S. Lin, "High resolution multispectral video capture with a hybrid camera system," in CVPR 2011, 2011, pp. 297–304.

  - Multiplexed light

[Park2007] J. I. Park, M. H. Lee, M. D. Grossberg, and S. K. Nayar, "Multispectral Imaging Using Multiplexed Illumination," in 2007 IEEE 11th International Conference on Computer Vision, 2007, pp. 1–8.

[Parmar2008] M. Parmar, S. Lansel, and B. A. Wandell, "Spatio-spectral reconstruction of the multispectral datacube using sparse recovery," in 2008 15th IEEE International Conference on Image Processing, 2008, pp. 473–476.

[Goel2015] M. Goel et al., "HyperCam: Hyperspectral Imaging for Ubiquitous Computing Applications," in Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York, NY, USA, 2015, pp. 145–156.

- No spatial info considered

[Nguyen2014] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, "Training-Based Spectral Reconstruction from a Single RGB Image," in Computer Vision ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Springer International Publishing, 2014, pp. 186–201.

[Arad2016] B. Arad and O. Ben-Shahar, "Sparse Recovery of Hyperspectral Signal from Natural RGB Images," in *ECCV*, 2016, pp. 19–34.

- Leveraging spatial context:

[Robles-Kelly2015] A. Robles-Kelly, "Single Image Spectral Reconstruction for Multimedia Applications," in ACM Multimedia, 2015.

[Galliani2017] S. Galliani, C. Lanaras, D. Marmanis, E. Baltsavias, and K. Schindler, "Learned Spectral Super-Resolution," arXiv:1703.09470 [cs], Mar. 2017.

Related work

Spectral reconstruction

- From RGB + extra help:
  - Low-res HS image

[Cao2011] X. Cao, X. Tong, Q. Dai, and S. Lin, "High resolution multispectral video capture with a hybrid camera system," in CVPR 2011, 2011, pp. 297–304.

  - Multiplexed light

[Park2007] J. I. Park, M. H. Lee, M. D. Grossberg, and S. K. Nayar, "Multispectral Imaging Using Multiplexed Illumination," in 2007 IEEE 11th International Conference on Computer Vision, 2007, pp. 1–8.

[Parmar2008] M. Parmar, S. Lansel, and B. A. Wandell, "Spatio-spectral reconstruction of the multispectral datacube using sparse recovery," in 2008 15th IEEE International Conference on Image Processing, 2008, pp. 473–476.

[Goel2015] M. Goel et al., "HyperCam: Hyperspectral Imaging for Ubiquitous Computing Applications," in Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York, NY, USA, 2015, pp. 145–156.

- No spatial info considered

[Nguyen2014] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, "Training-Based Spectral Reconstruction from a Single RGB Image," in Computer Vision ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Springer International Publishing, 2014, pp. 186–201.

[Arad2016] B. Arad and O. Ben-Shahar, "Sparse Recovery of Hyperspectral Signal from Natural RGB Images," in *ECCV*, 2016, pp. 19–34.

- Leveraging spatial context:

[Robles-Kelly2015] A. Robles-Kelly, "Single Image Spectral Reconstruction for Multimedia Applications," in ACM Multimedia, 2015.

[Galliani2017] S. Galliani, C. Lanaras, D. Marmanis, E. Baltsavias, and K. Schindler, "Learned Spectral Super-Resolution," arXiv:1703.09470 [cs], Mar. 2017.

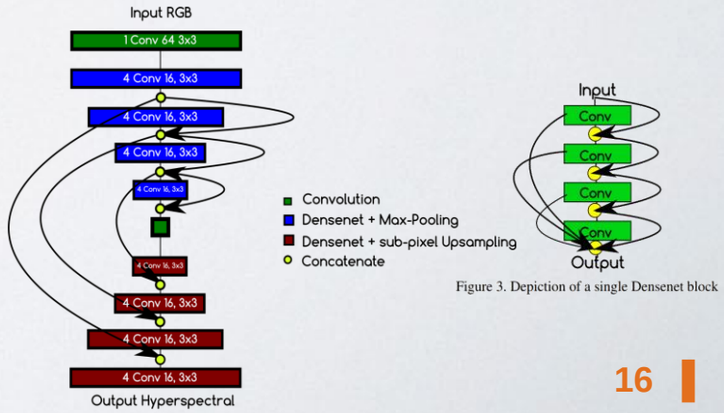


Figure 3. Depiction of a single Densenet block

Dataset

ICVL dataset [Arad2016]

- ~200 images
- Specim PS Kappa DX4 + rotary
- 1392×1300 spatial resolution over 519 spectral bands (400-1,000nm at roughly 1.25nm increments.
- Downsampled to 31 spectral channels from 400nm to 700nm at 10nm increments

[Arad2016] B. Arad and O. Ben-Shahar, "Sparse Recovery of Hyperspectral Signal from Natural RGB Images," in *ECCV*, 2016, pp. 19–34.

Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

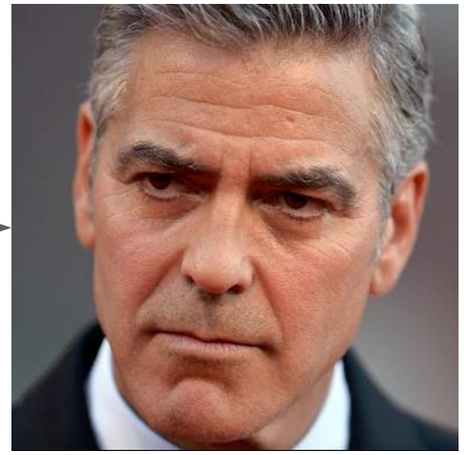
Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

Generative Adversarial Networks (GANs)

$$G^* = \underset{G}{\operatorname{argmin}} \max_{D} \mathcal{L}_{GAN}$$

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{y \sim p_{data}(y)}[\log D(y)] + \\ + \mathbb{E}_{z \sim p_{noise}(z)}[\log(1 - D(G(z)))]$$

Real world images
(train set)

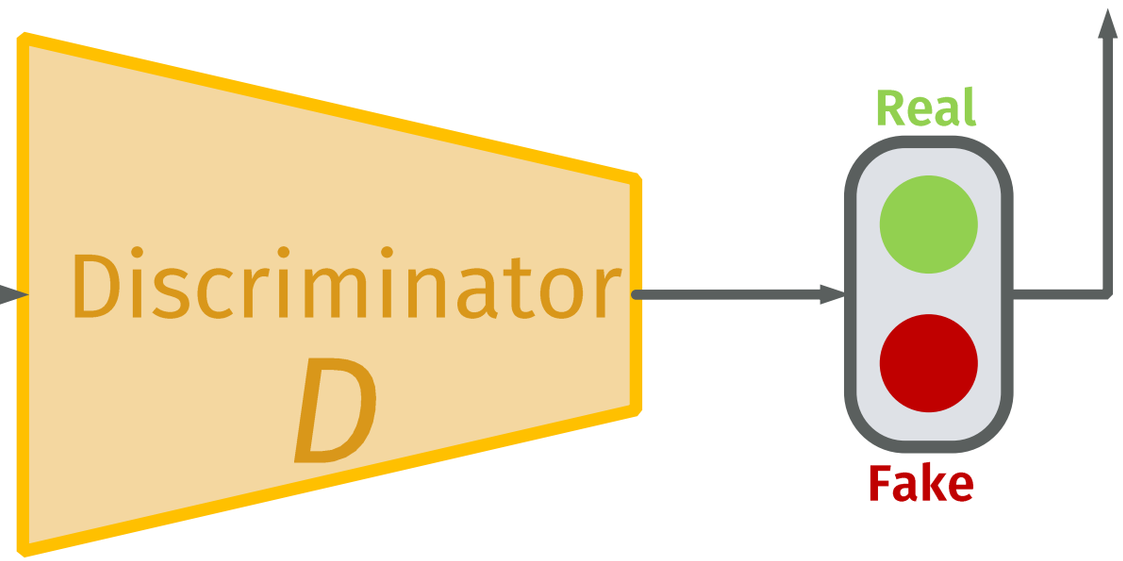


Real sample

Fake sample

$z \sim p_{noise}(z)$

Random
noise vector



https://blog.openai.com/generative-models/
https://www.youtube.com/watch?v=J0o6LhaUSsc&vl=en

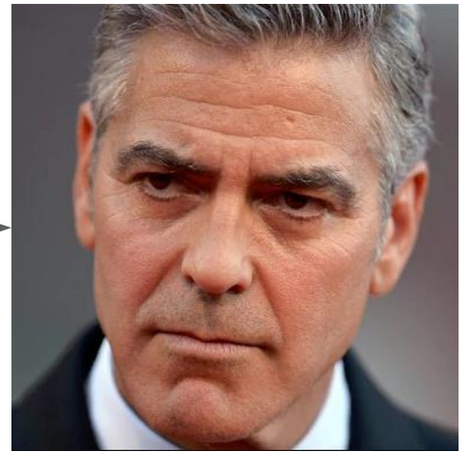
[Goodfellow2014] I. Goodfellow et al., "Generative Adversarial Nets," in Advances in Neural Information Processing Systems 27, 2014, pp. 2672-2680.

Generative Adversarial Networks (GANs)

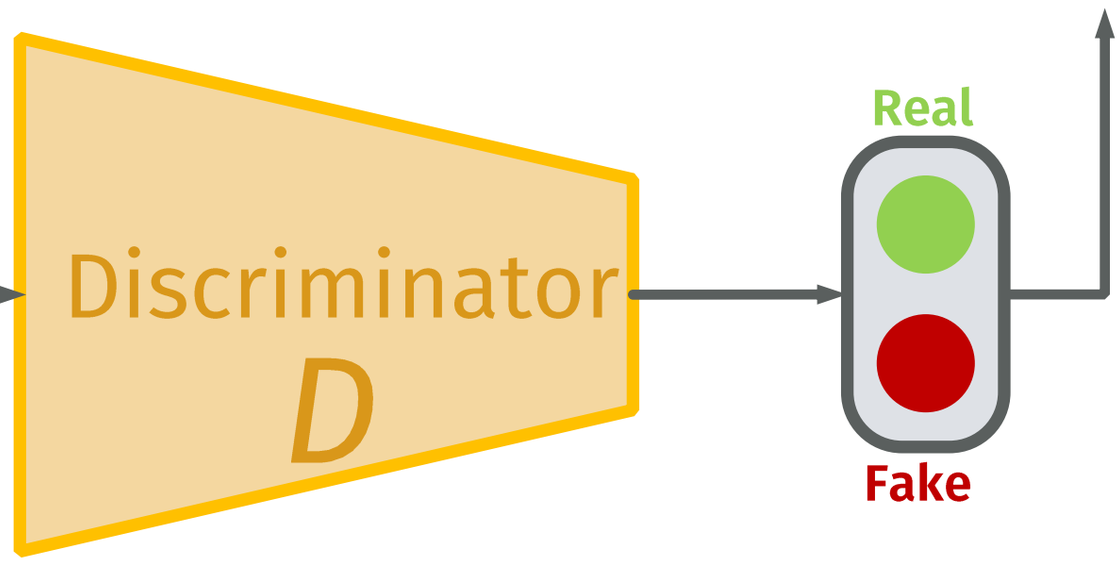
$$G^* = \underset{G}{\operatorname{argmin}} \max_{D} \mathcal{L}_{GAN}$$

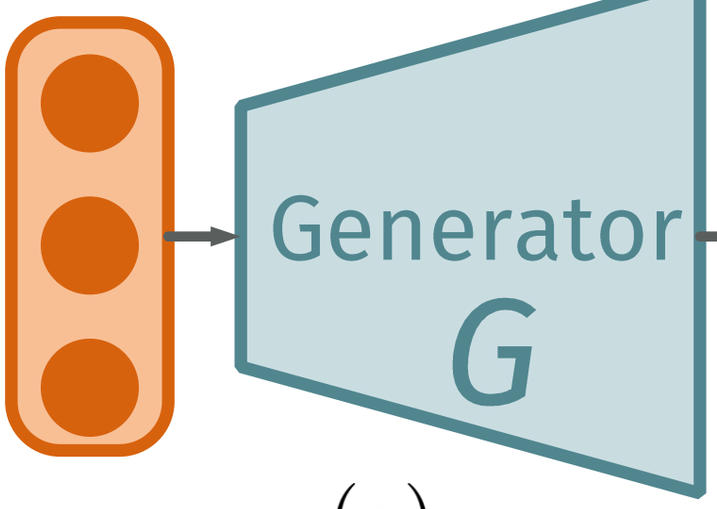
$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{y \sim p_{data}(y)}[\log D(y)] + \\ + \mathbb{E}_{z \sim p_{noise}(z)}[\log(1 - D(G(z)))]$$

Real world images
(train set)



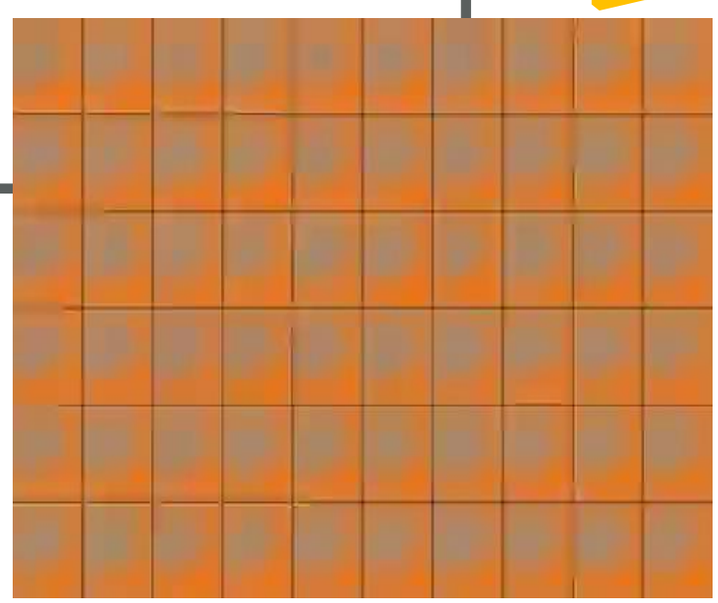
Real sample





$z \sim p_{noise}(z)$

Random
noise vector



https://blog.openai.com/generative-models/
https://www.youtube.com/watch?v=J0o6LhaUSsc&vl=en

[Goodfellow2014] I. Goodfellow et al., "Generative Adversarial Nets," in Advances in Neural Information Processing Systems 27, 2014, pp. 2672-2680.

GANs

Generated fake samples

D. Berthelot, T. Schumm, and L. Metz, "BEGAN: Boundary Equilibrium Generative Adversarial Networks," Mar. 2017.

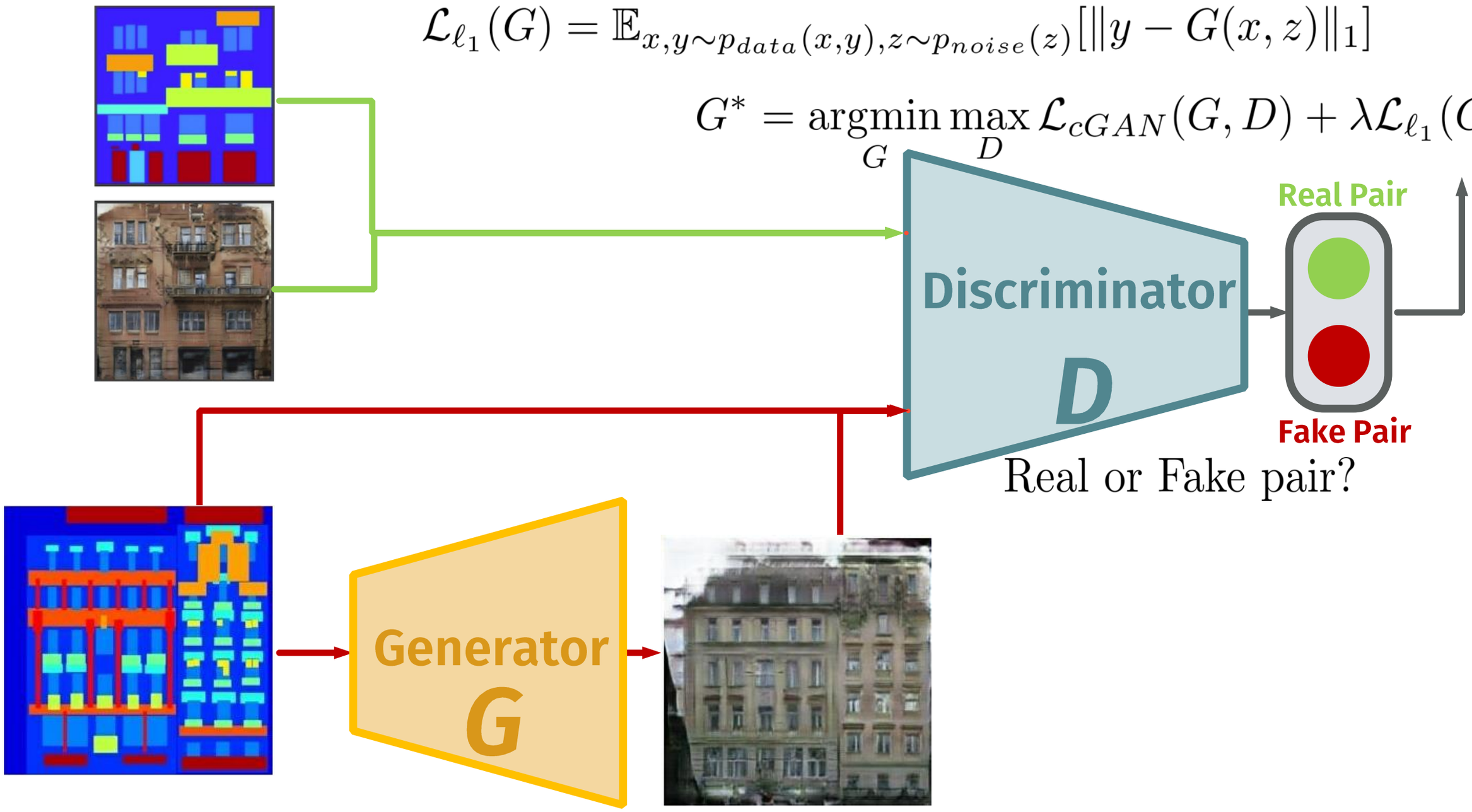
Conditional GANs: pix2pix [Isola2017]

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{x,y\sim p_{data}(x,y)}[\log D(x,y)] + \mathbb{E}_{x\sim p_{data}(x),z\sim p_{noise}(z)}[\log(1 - D(x, G(x,z)))]$$

Sketch-real image aligned pairs

$$\mathcal{L}_{\ell_1}(G) = \mathbb{E}_{x,y\sim p_{data}(x,y),z\sim p_{noise}(z)}[\|y - G(x,z)\|_1]$$

$$G^* = \underset{G}{\operatorname{argmin}} \max_{D} \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{\ell_1}(G)$$



[Isola2017] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-To-Image Translation With Conditional Adversarial Networks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[Hesse2017] Christopher Hesse, "Image-to-Image Demo - Affine Layer," 19-Feb-2017. https://affinelayer.com/pixsrv/.

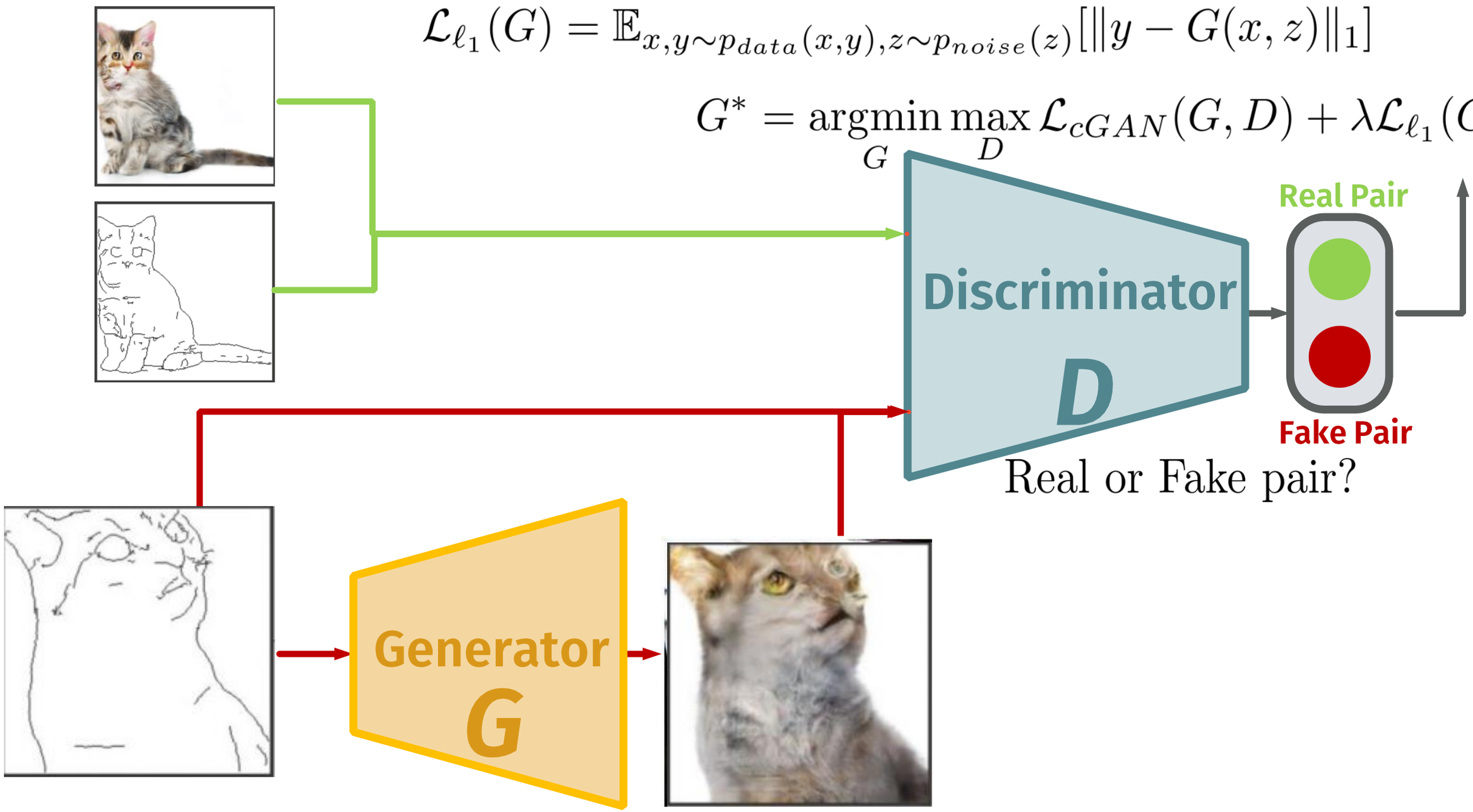
Conditional GANs: pix2pix [Isola2017]

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{x,y\sim p_{data}(x,y)}[\log D(x,y)] + \mathbb{E}_{x\sim p_{data}(x),z\sim p_{noise}(z)}[\log(1 - D(x, G(x,z)))]$$

Sketch-real image aligned pairs

$$\mathcal{L}_{\ell_1}(G) = \mathbb{E}_{x,y\sim p_{data}(x,y),z\sim p_{noise}(z)}[\|y - G(x,z)\|_1]$$

$$G^* = \underset{G}{\operatorname{argmin}} \max_{D} \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{\ell_1}(G)$$



[Isola2017] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-To-Image Translation With Conditional Adversarial Networks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[Hesse2017] Christopher Hesse, "Image-to-Image Demo - Affine Layer," 19-Feb-2017. https://affinelayer.com/pixsrv/.

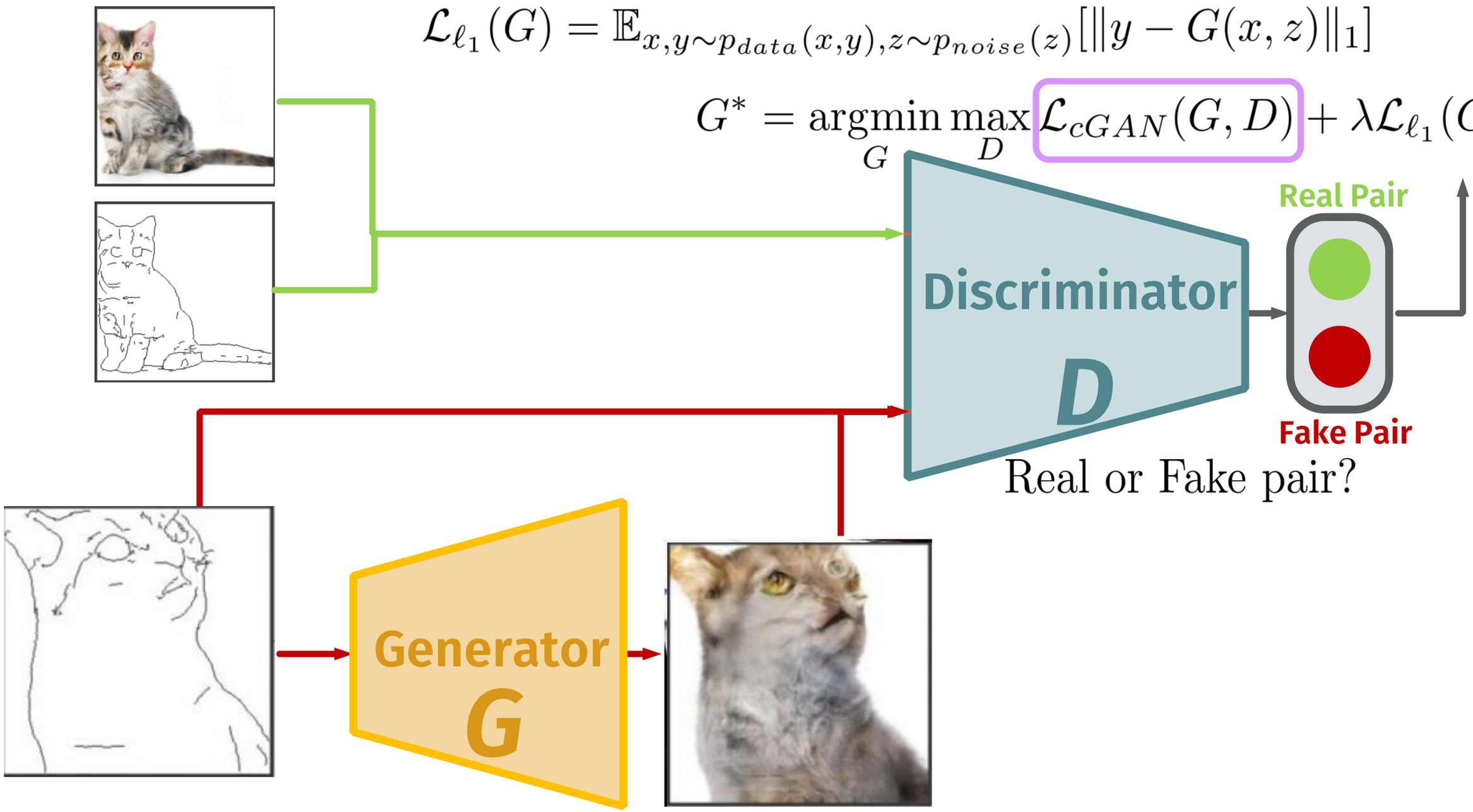
Conditional GANs: pix2pix [Isola2017]

Sketch-real image aligned pairs

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{x,y\sim p_{data}(x,y)}[\log D(x,y)] + \mathbb{E}_{x\sim p_{data}(x),z\sim p_{noise}(z)}[\log(1 - D(x, G(x,z)))]$$

$$\mathcal{L}_{\ell_1}(G) = \mathbb{E}_{x,y\sim p_{data}(x,y),z\sim p_{noise}(z)}[\|y - G(x,z)\|_1]$$

$$G^* = \operatorname*{argmin}_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{\ell_1}(G)$$



[Isola2017] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-To-Image Translation With Conditional Adversarial Networks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[Hesse2017] Christopher Hesse, "Image-to-Image Demo - Affine Layer," 19-Feb-2017. https://affinelayer.com/pixsrv/.

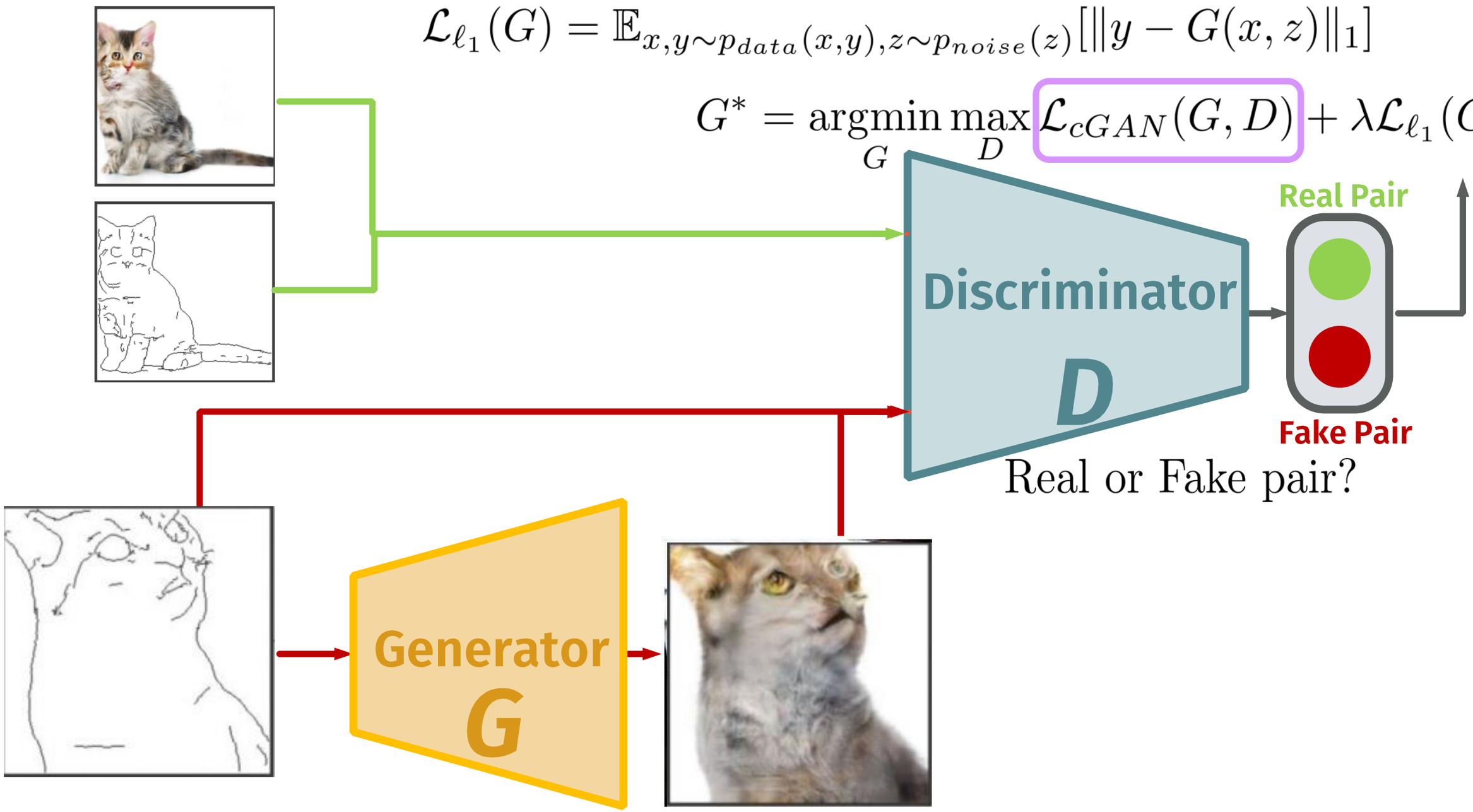
Conditional GANs: pix2pix [Isola2017]

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{x,y\sim p_{data}(x,y)}[\log D(x,y)] + \mathbb{E}_{x\sim p_{data}(x),z\sim p_{noise}(z)}[\log(1 - D(x, G(x,z)))]$$

Sketch-real image aligned pairs

$$\mathcal{L}_{\ell_1}(G) = \mathbb{E}_{x,y\sim p_{data}(x,y),z\sim p_{noise}(z)}[\|y - G(x,z)\|_1]$$

$$G^* = \operatorname*{argmin}_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{\ell_1}(G)$$



[Isola2017] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-To-Image Translation With Conditional Adversarial Networks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

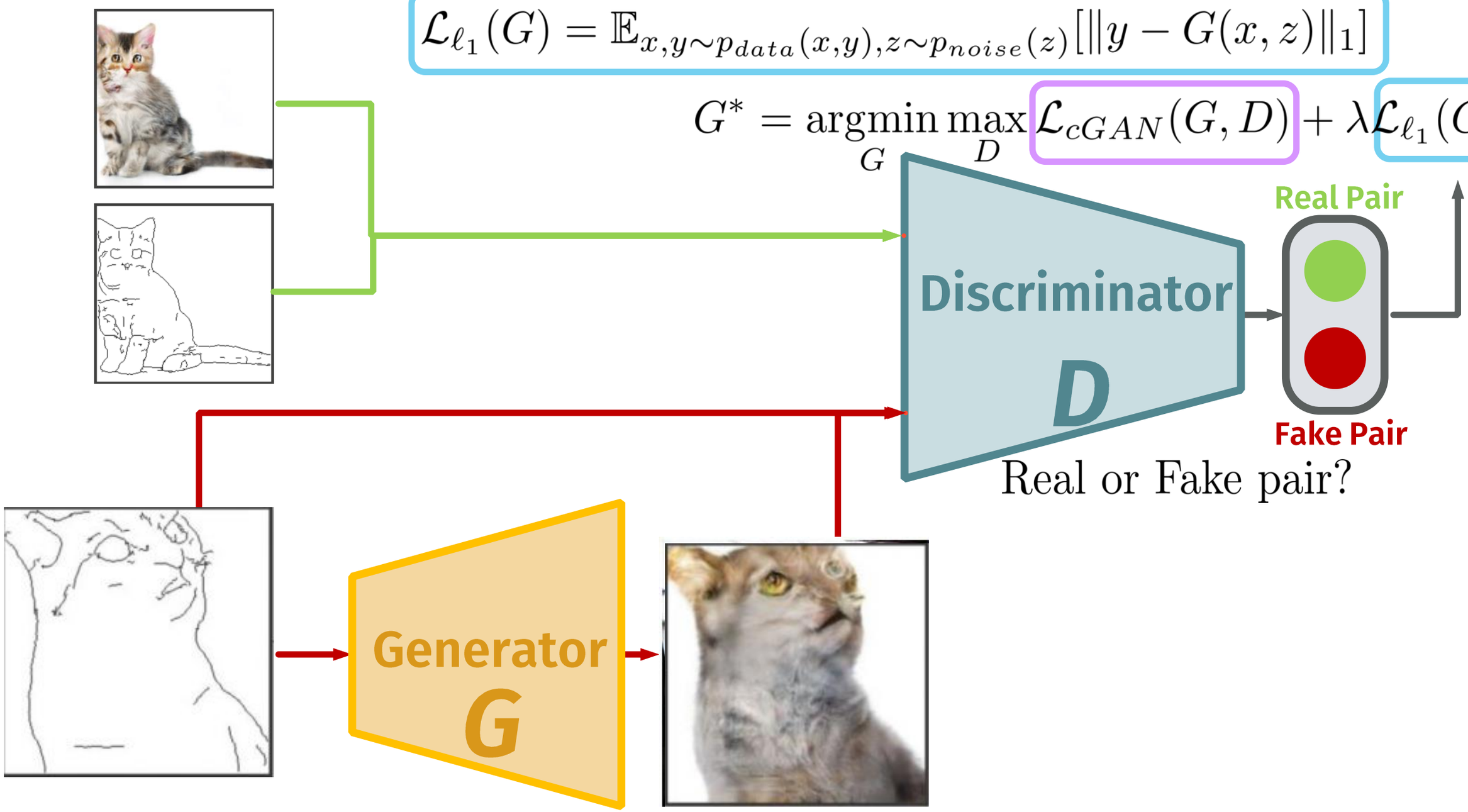
[Hesse2017] Christopher Hesse, "Image-to-Image Demo - Affine Layer," 19-Feb-2017. https://affinelayer.com/pixsrv/.

Conditional GANs: pix2pix [Isola2017]

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{x,y\sim p_{data}(x,y)}[\log D(x,y)] + \mathbb{E}_{x\sim p_{data}(x),z\sim p_{noise}(z)}[\log(1 - D(x, G(x,z)))]$$

$$\mathcal{L}_{\ell_1}(G) = \mathbb{E}_{x,y\sim p_{data}(x,y),z\sim p_{noise}(z)}[\|y - G(x,z)\|_1]$$

$$G^* = \operatorname*{argmin}_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{\ell_1}(G)$$



[Isola2017] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-To-Image Translation With Conditional Adversarial Networks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

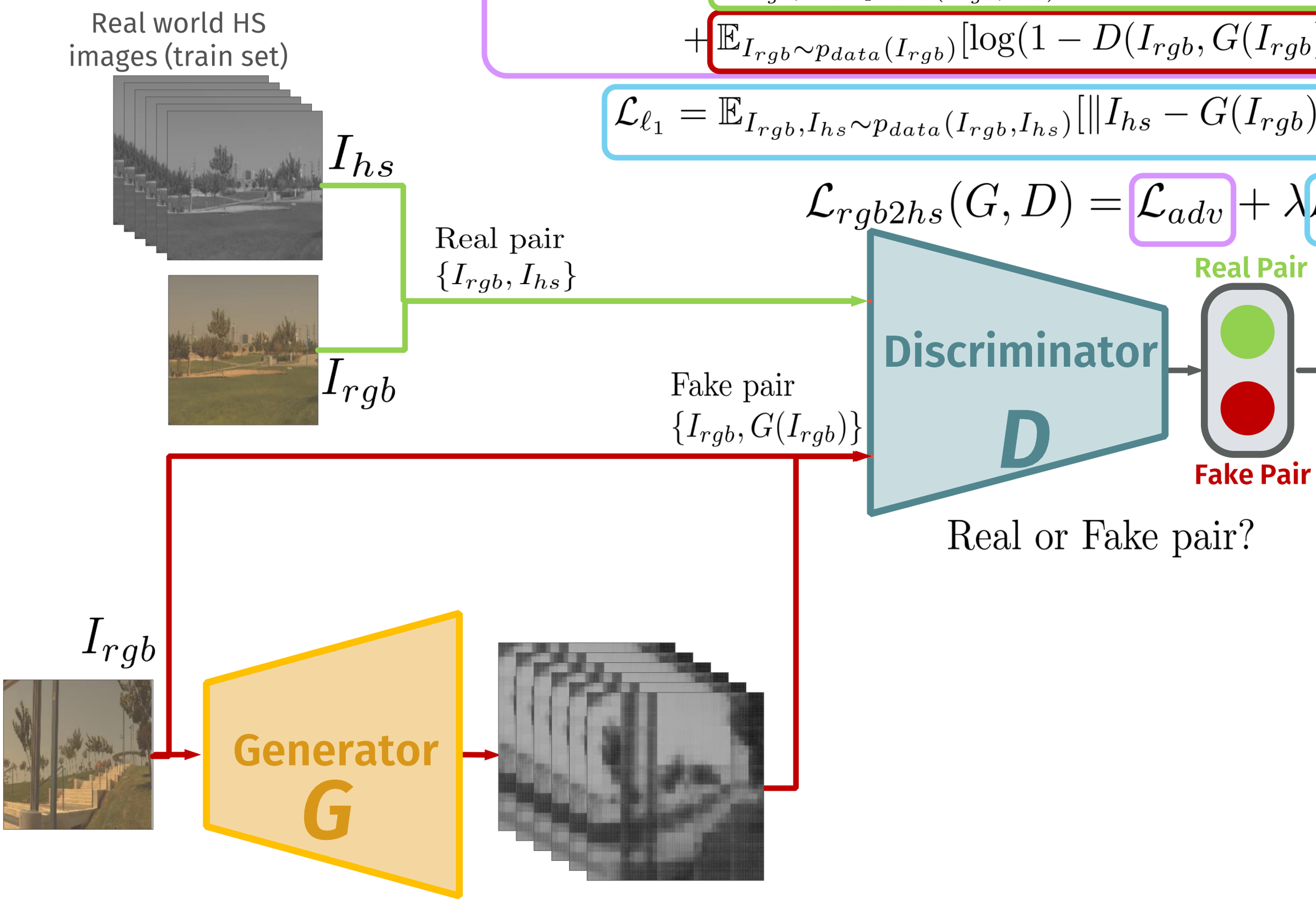
[Hesse2017] Christopher Hesse, "Image-to-Image Demo - Affine Layer," 19-Feb-2017. https://affinelayer.com/pixsrv/.

Adversarial spectral reconstruction networks

$$\mathcal{L}_{adv}(G,D) = \mathbb{E}_{I_{rgb},I_{hs}\sim p_{data}(I_{rgb},I_{hs})}[\log D(I_{rgb},I_{hs})] + \mathbb{E}_{I_{rgb}\sim p_{data}(I_{rgb})}[\log(1 - D(I_{rgb}, G(I_{rgb})))]$$

$$\mathcal{L}_{\ell_1} = \mathbb{E}_{I_{rgb},I_{hs}\sim p_{data}(I_{rgb},I_{hs})}[\|I_{hs} - G(I_{rgb})\|_1]$$

$$\mathcal{L}_{rgb2hs}(G,D) = \mathcal{L}_{adv} + \lambda\mathcal{L}_{\ell_1}$$

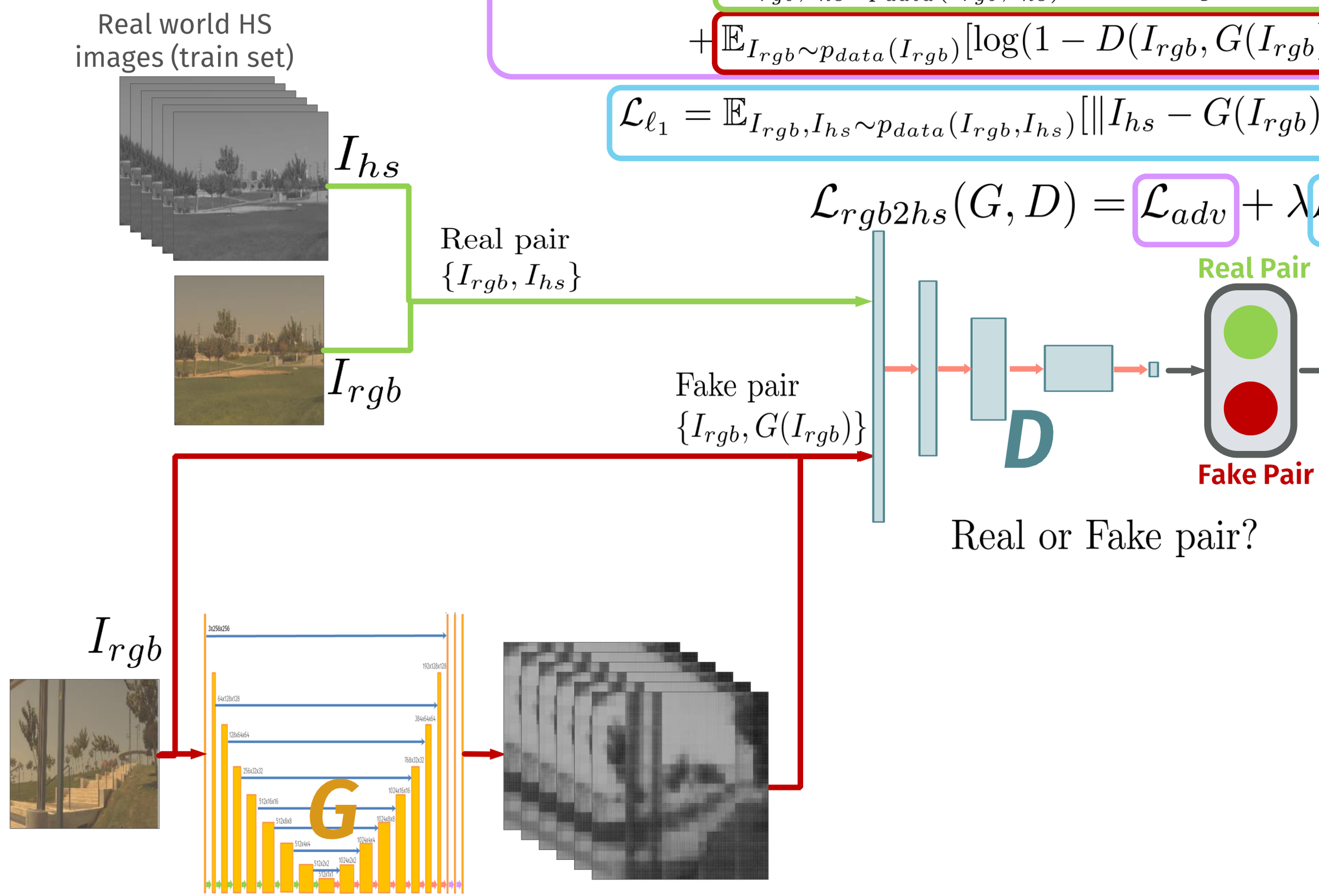


Adversarial spectral reconstruction networks

$$\mathcal{L}_{adv}(G,D) = \mathbb{E}_{I_{rgb},I_{hs}\sim p_{data}(I_{rgb},I_{hs})}[\log D(I_{rgb},I_{hs})] + \mathbb{E}_{I_{rgb}\sim p_{data}(I_{rgb})}[\log(1 - D(I_{rgb},G(I_{rgb})))]$$

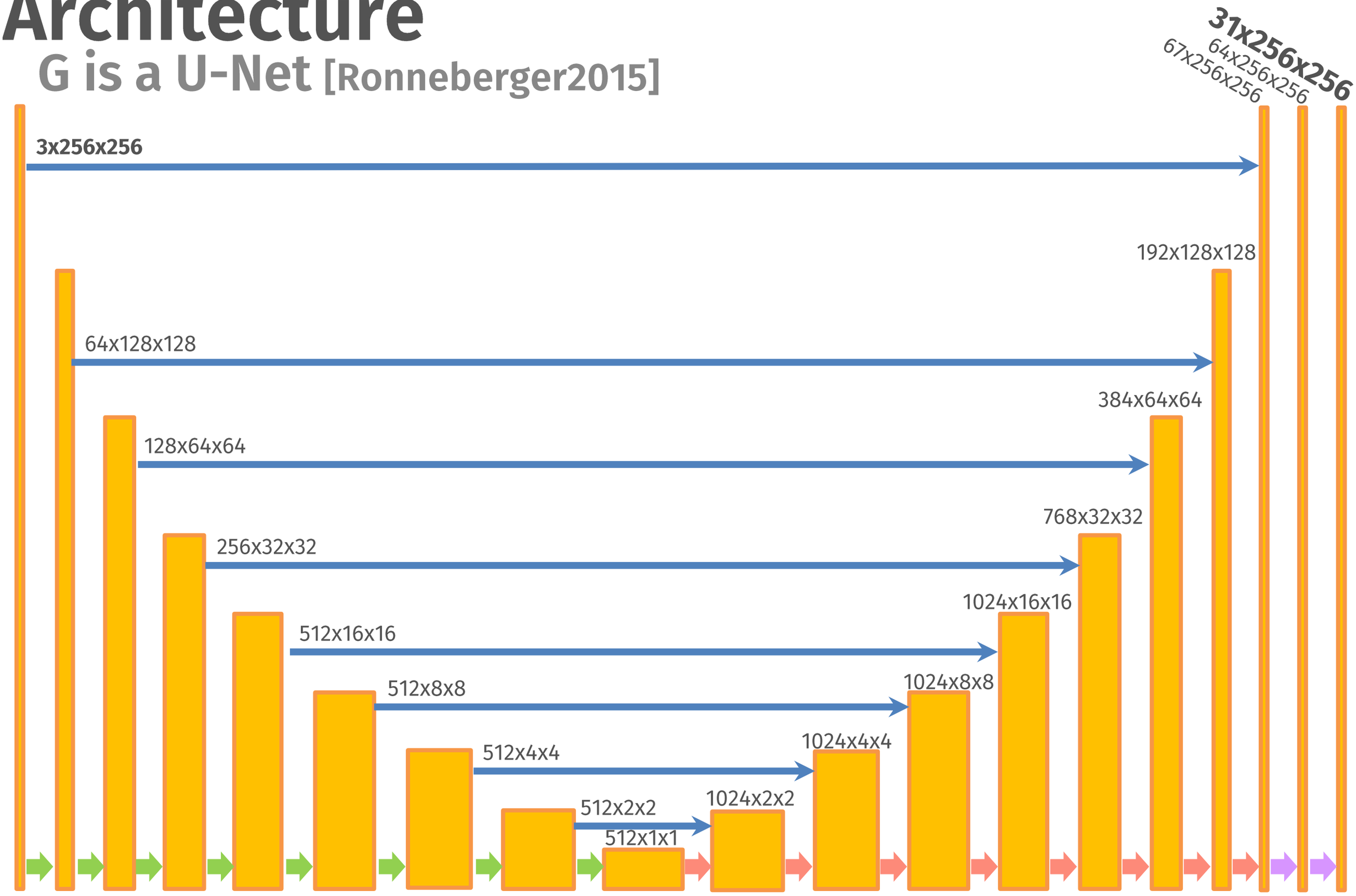
$$\mathcal{L}_{\ell_1} = \mathbb{E}_{I_{rgb},I_{hs}\sim p_{data}(I_{rgb},I_{hs})}[\|I_{hs} - G(I_{rgb})\|_1]$$

$$\mathcal{L}_{rgb2hs}(G,D) = \mathcal{L}_{adv} + \lambda\mathcal{L}_{\ell_1}$$



Architecture

G is a U-Net [Ronneberger2015]



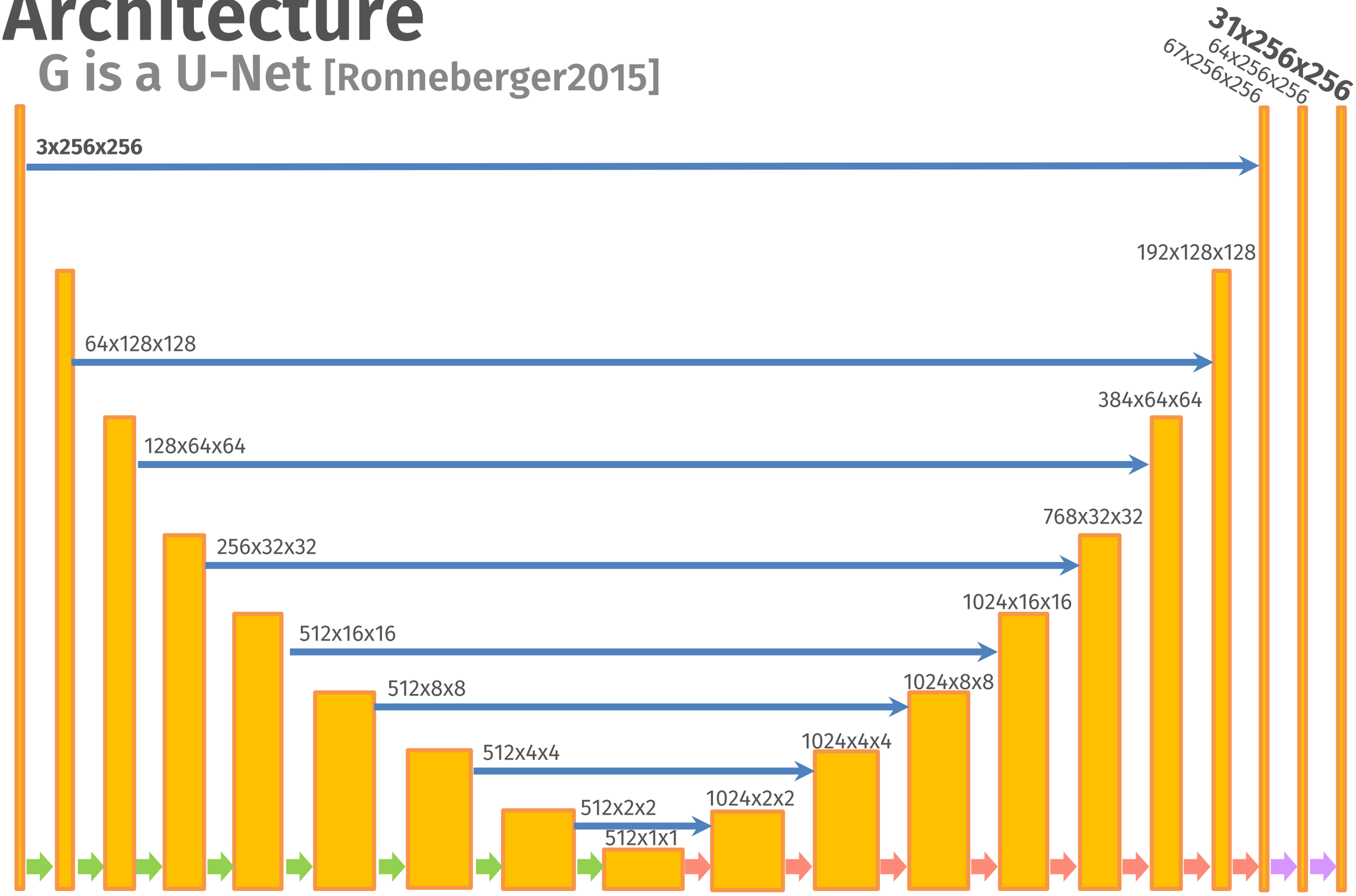
Conv2D(k=3,s=2) + BatchNorm + LeakyReLU

Conv2DTranspose(k=2,s=1) + Dropout(r=0.1) + Merge(→) + LeakyReLU

Conv2D(k=1,s=1) + LeakyReLU

Architecture

G is a U-Net [Ronneberger2015]



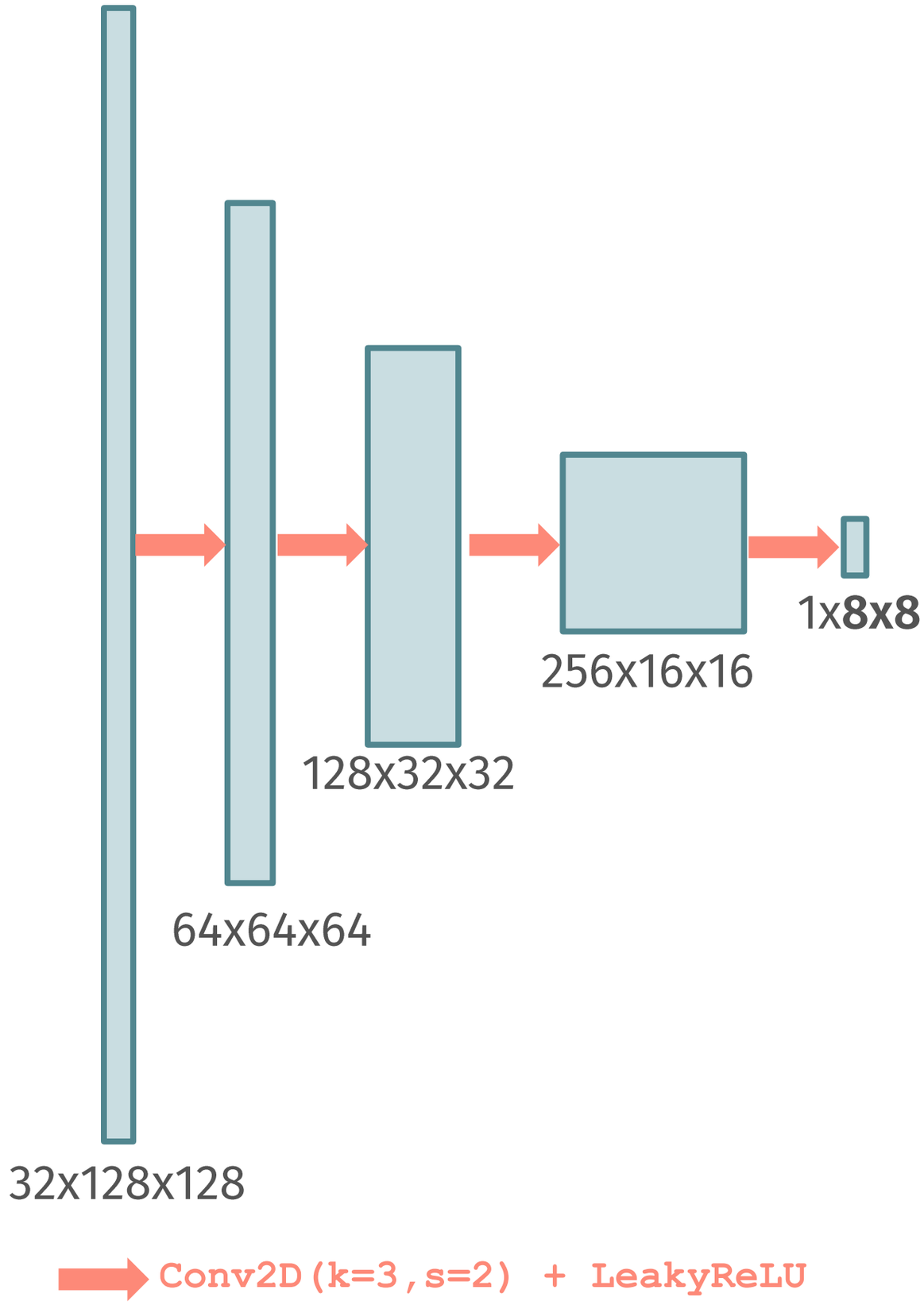
Conv2D(k=3,s=2) + ~~BatchNorm~~ + LeakyReLU

Conv2DTranspose(k=2,s=1) + Dropout(r=0.1) + Merge(→) + LeakyReLU

Conv2D(k=1,s=1) + LeakyReLU

Architecture

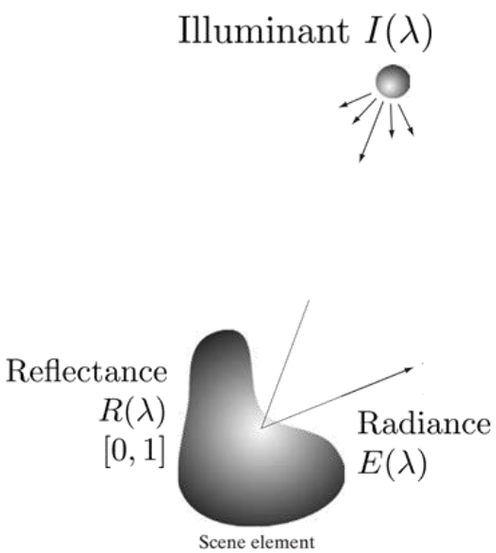
D is a Patch-CNN



Experiments

Dataset preparation

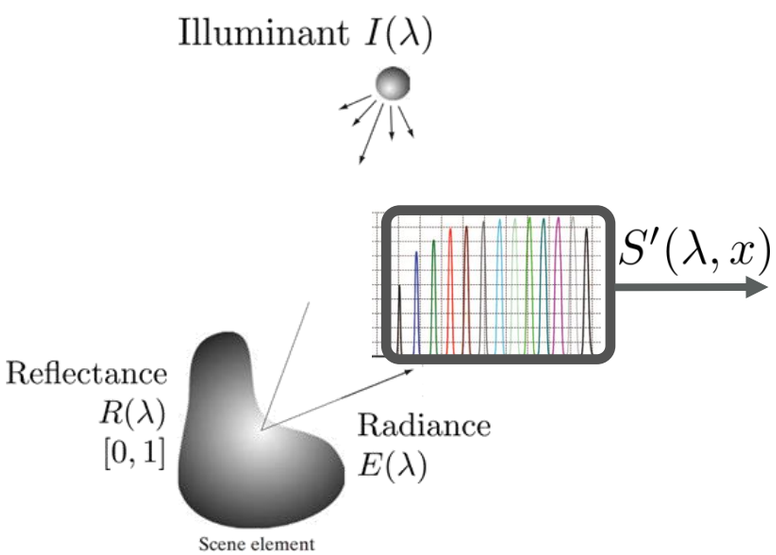
- 2 fold cross-validation (2x100)
- From hyperspectral to sRGB:



Experiments

Dataset preparation

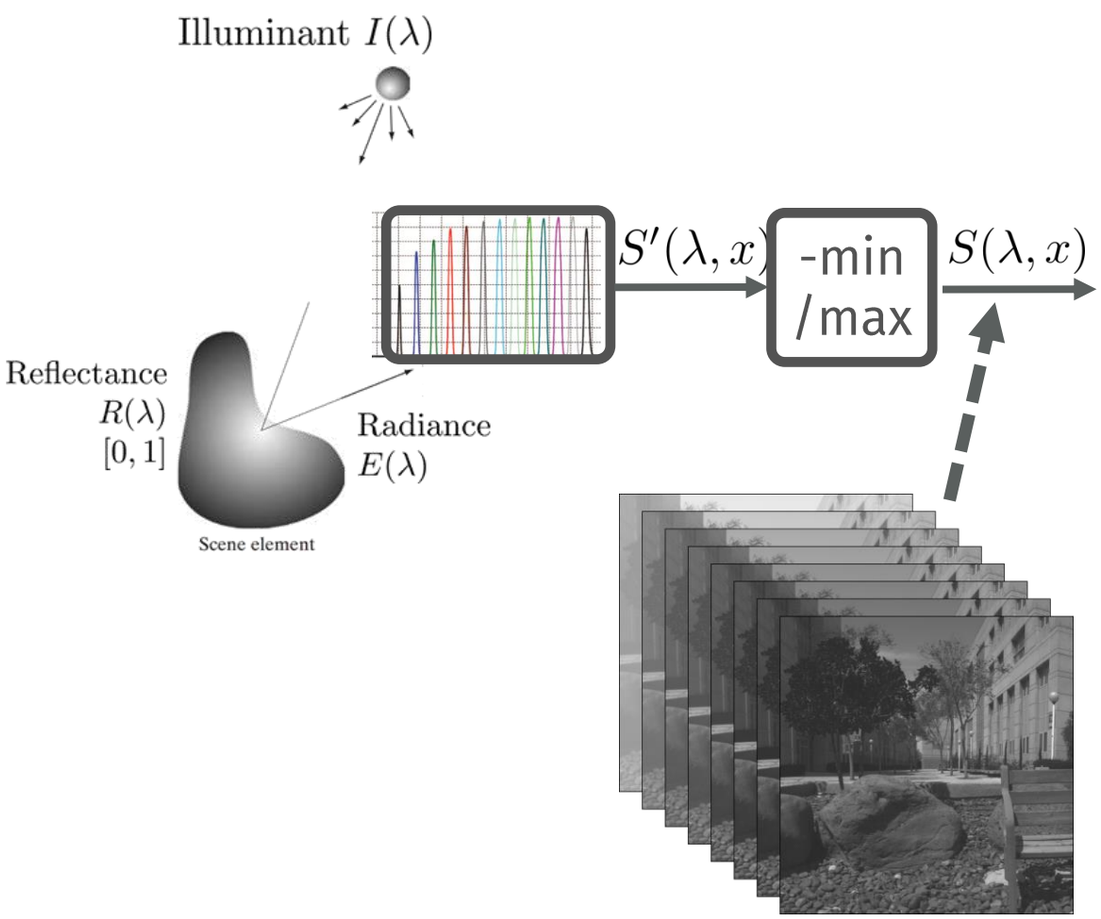
- 2 fold cross-validation (2x100)
- From hyperspectral to sRGB:



Experiments

Dataset preparation

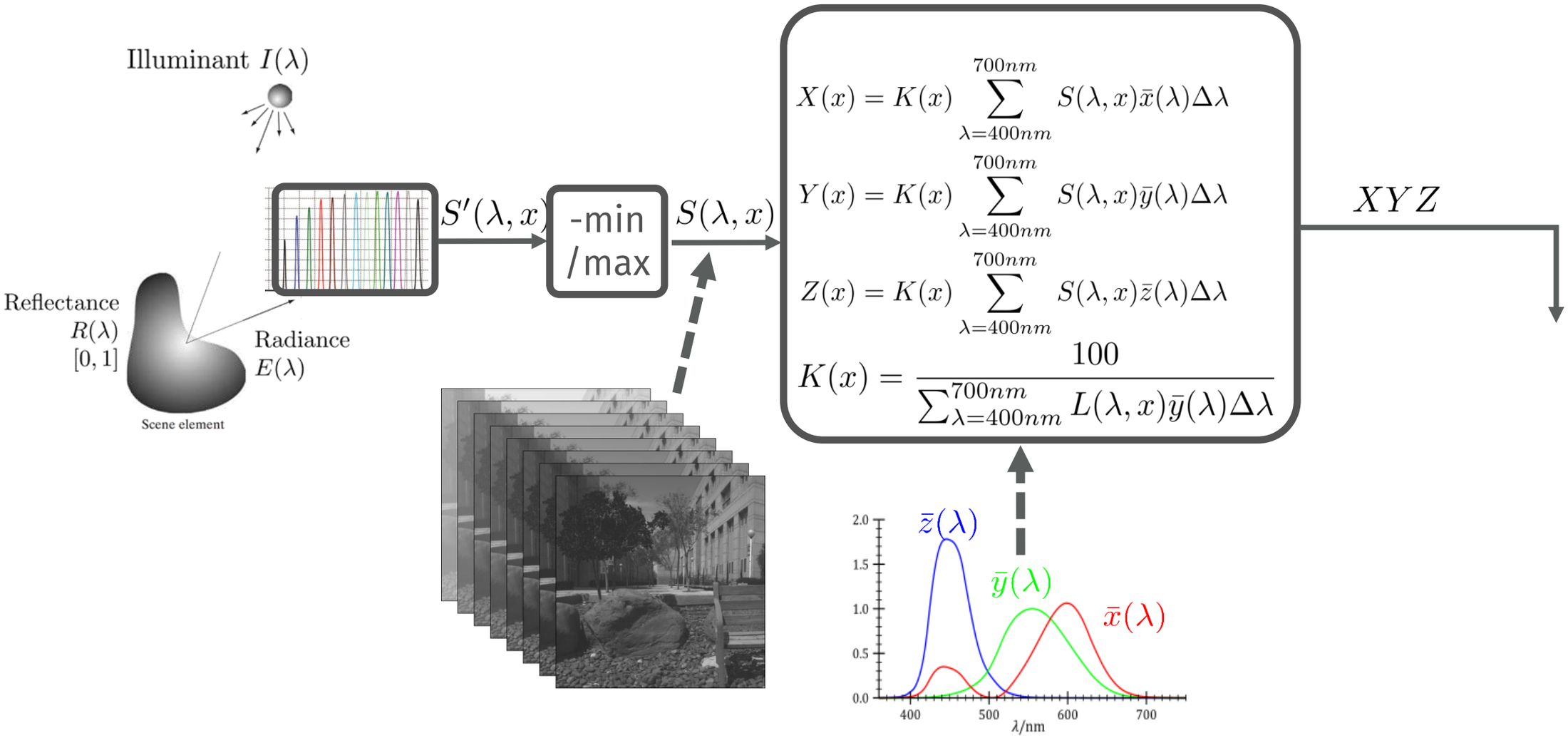
- 2 fold cross-validation (2x100)
- From hyperspectral to sRGB:



Experiments

Dataset preparation

- 2 fold cross-validation (2x100)
- From hyperspectral to sRGB:

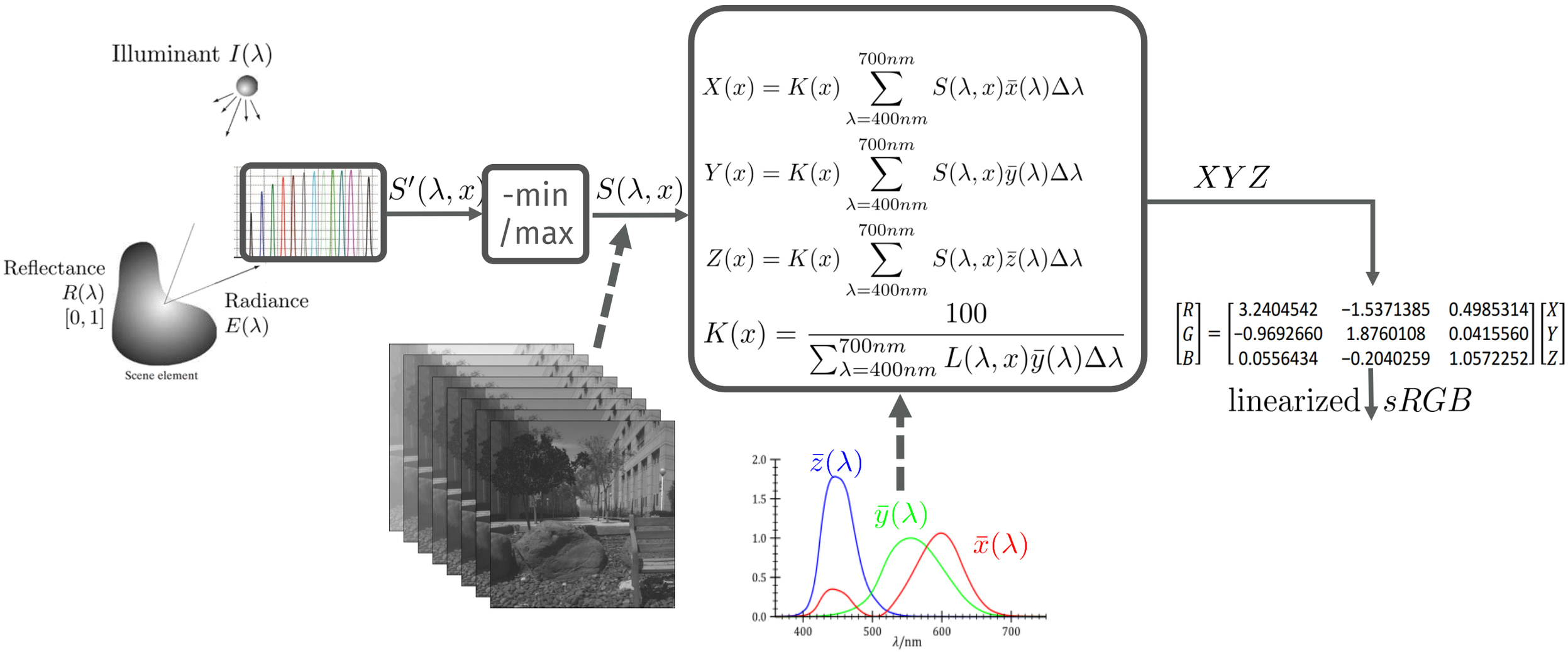


CIE 1964 10°
Color Matching Functions

Experiments

Dataset preparation

- 2 fold cross-validation (2x100)
- From hyperspectral to sRGB:

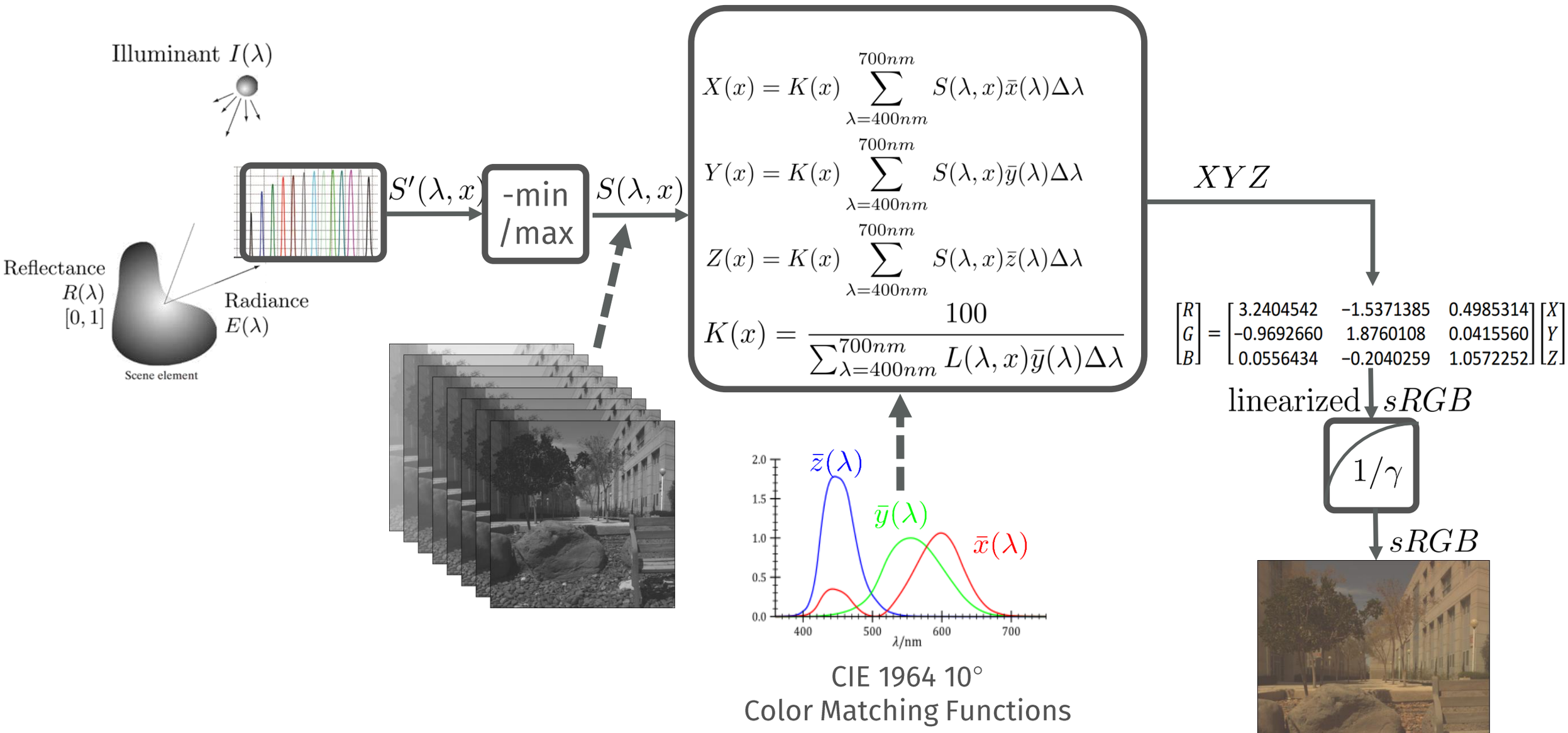


CIE 1964 10°
Color Matching Functions

Experiments

Dataset preparation

- 2 fold cross-validation (2x100)
- From hyperspectral to sRGB:

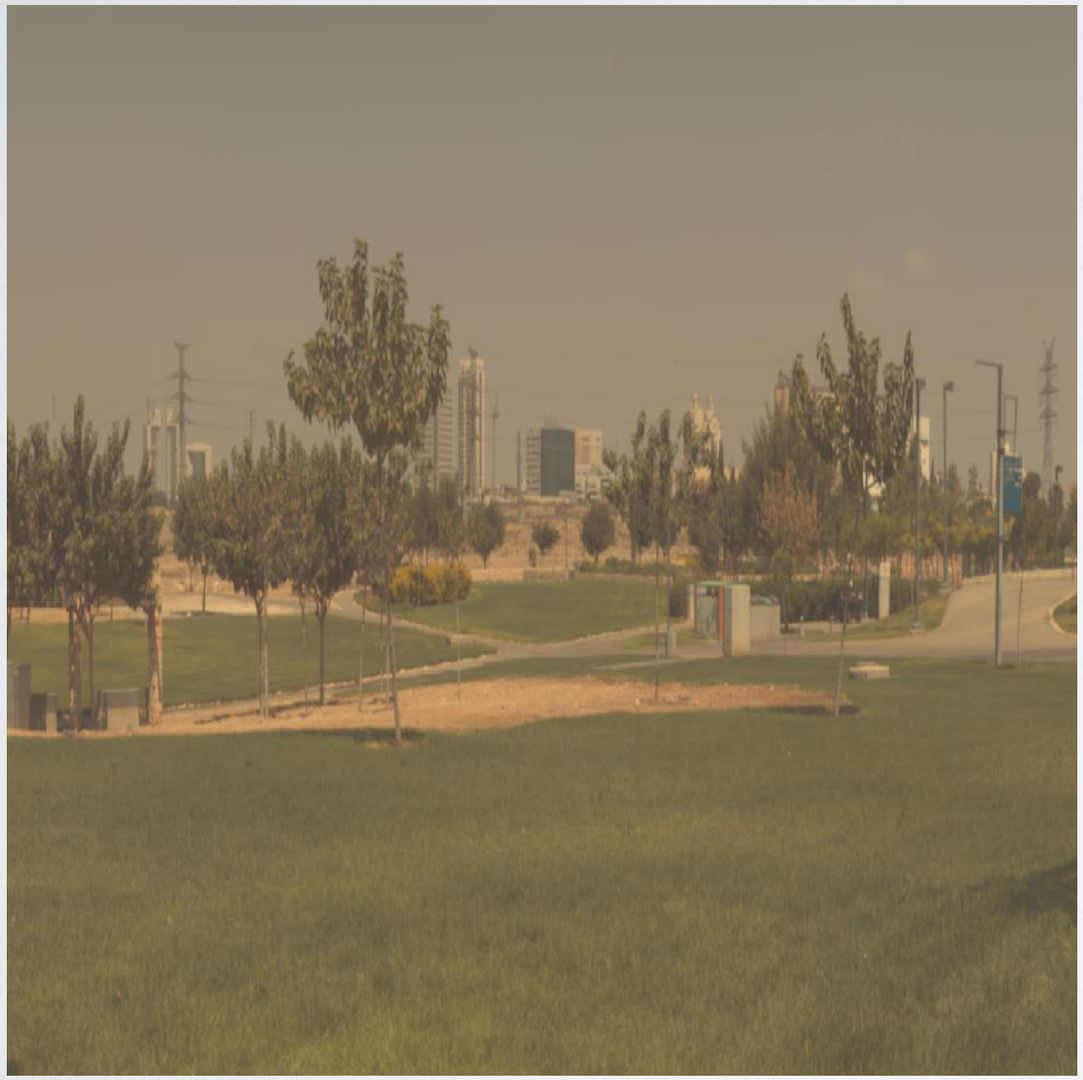


CIE 1964 10°
Color Matching Functions

Experiments

Training and testing with large images

- Training: random crops



1280

1280

- Testing: tile decompositon + reconstruction

1280

1280

Experiments

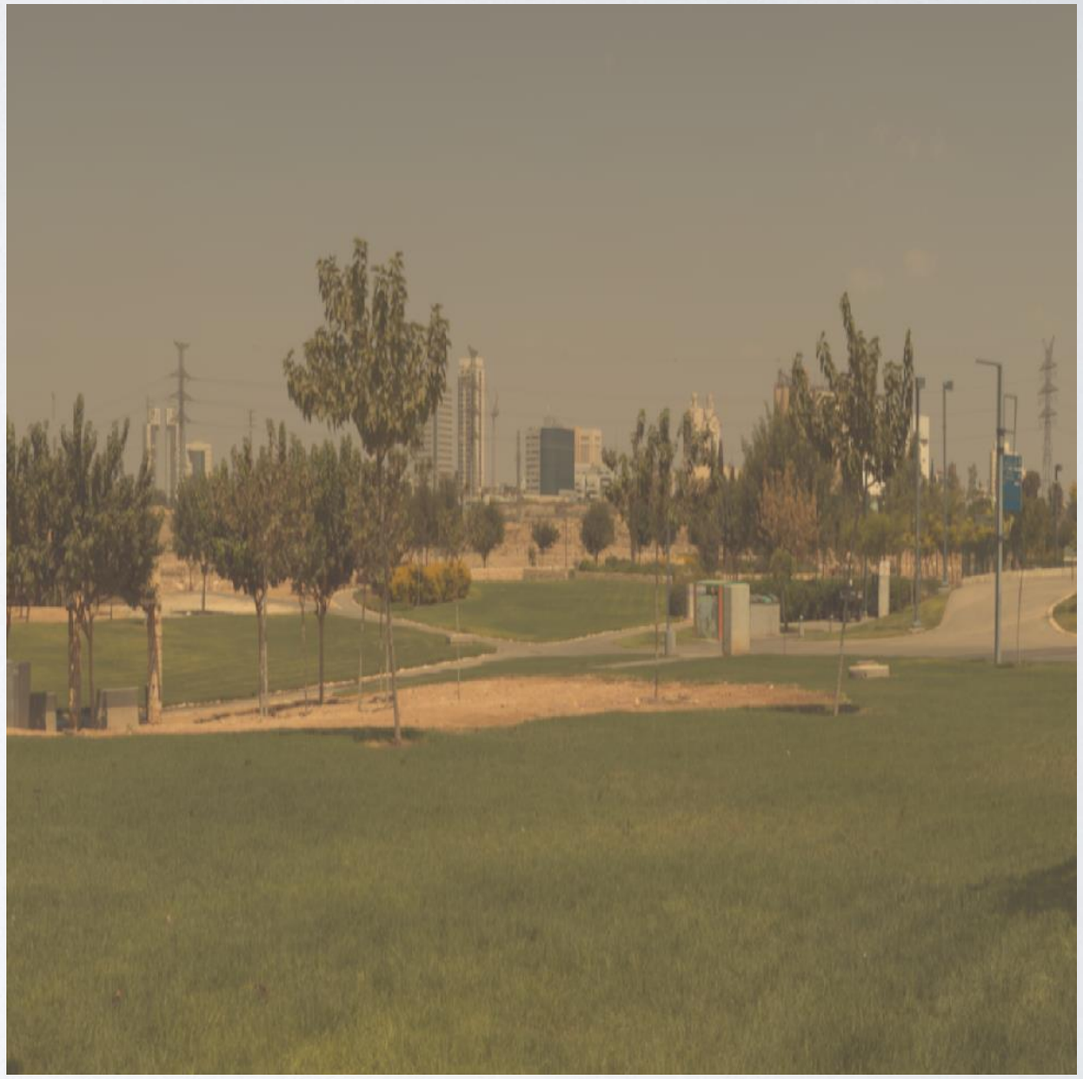
Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction



1280

1280

Experiments

Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction

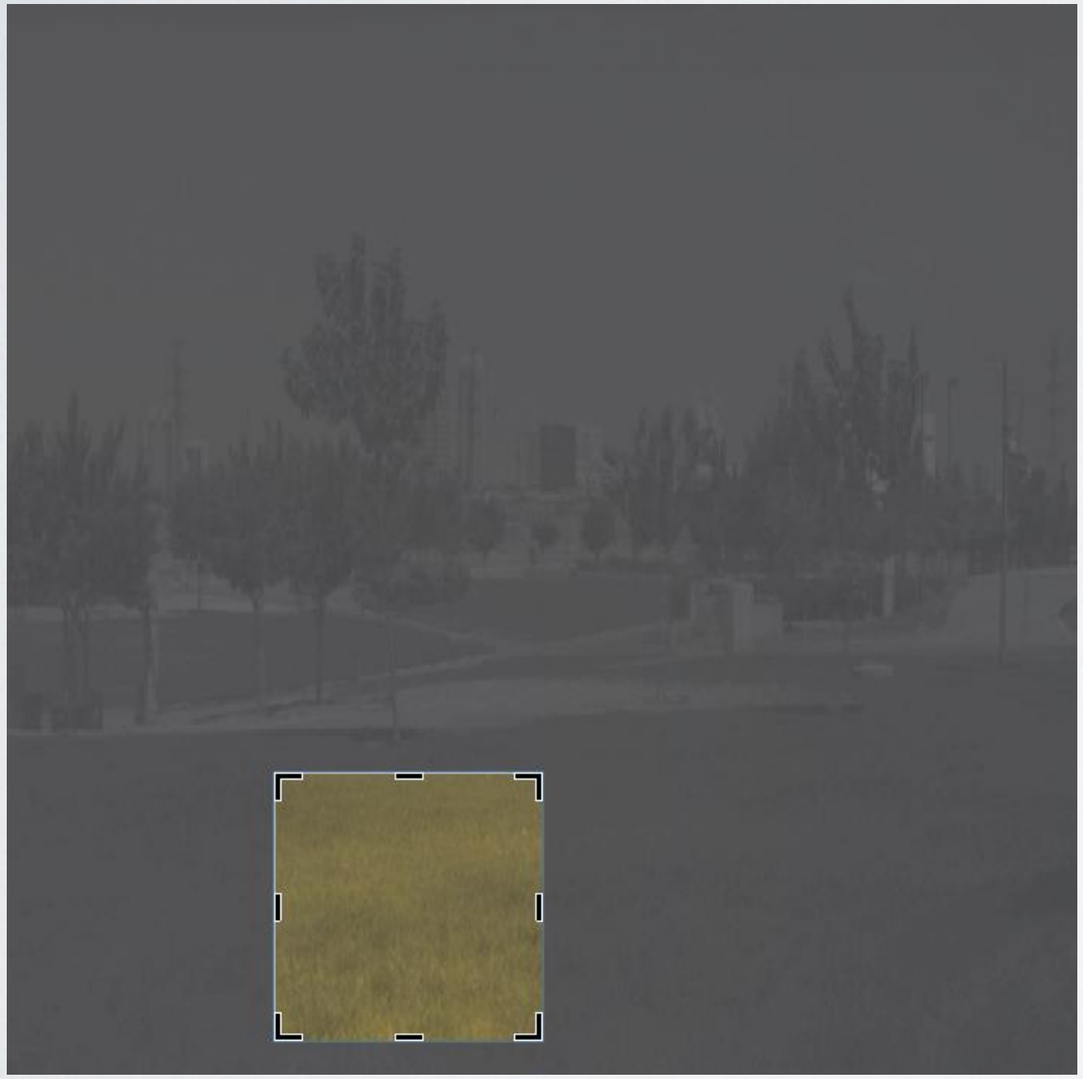
1280

1280

Experiments

Training and testing with large images

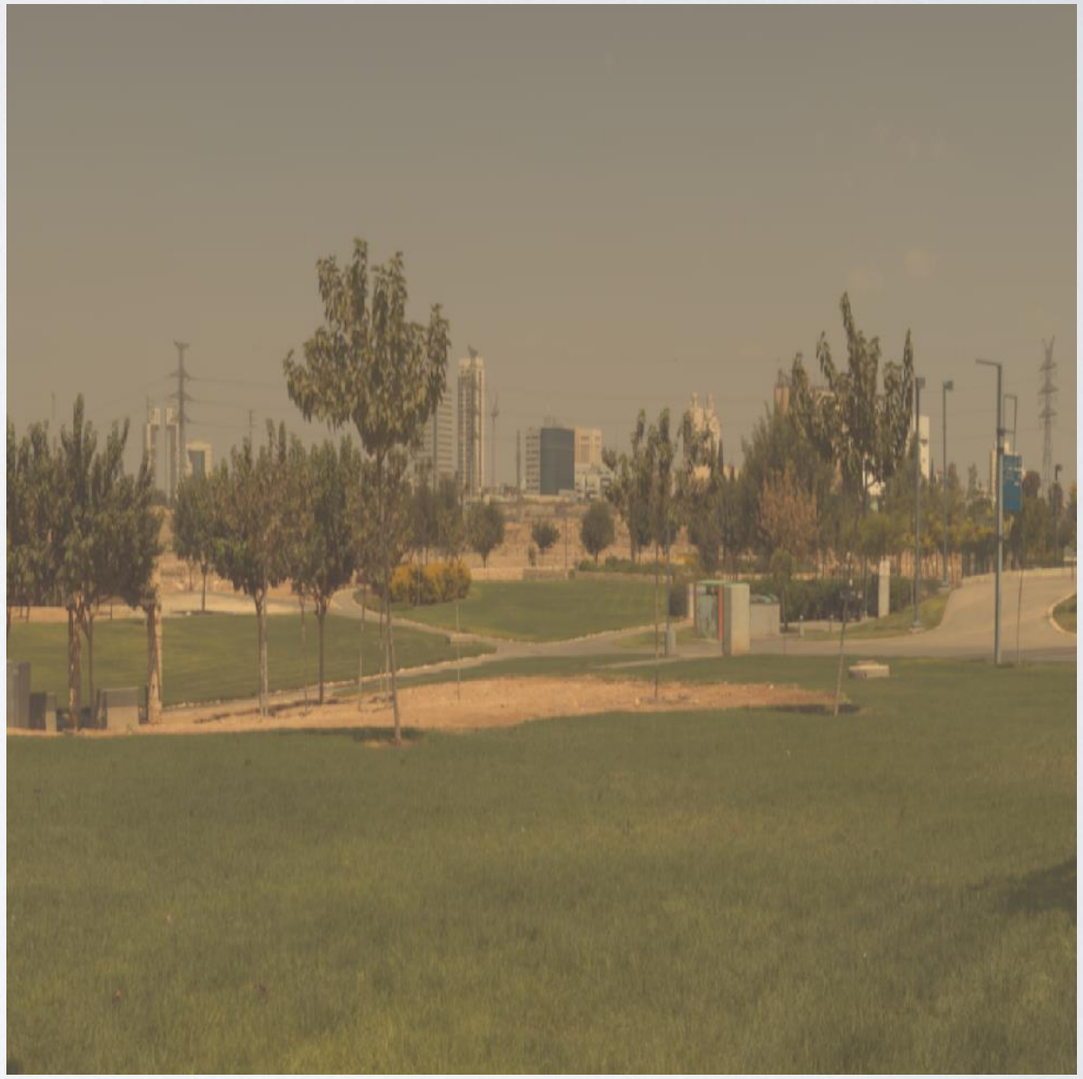
- Training: random crops



1280

1280

- Testing: tile decompositon + reconstruction



1280

1280

Experiments

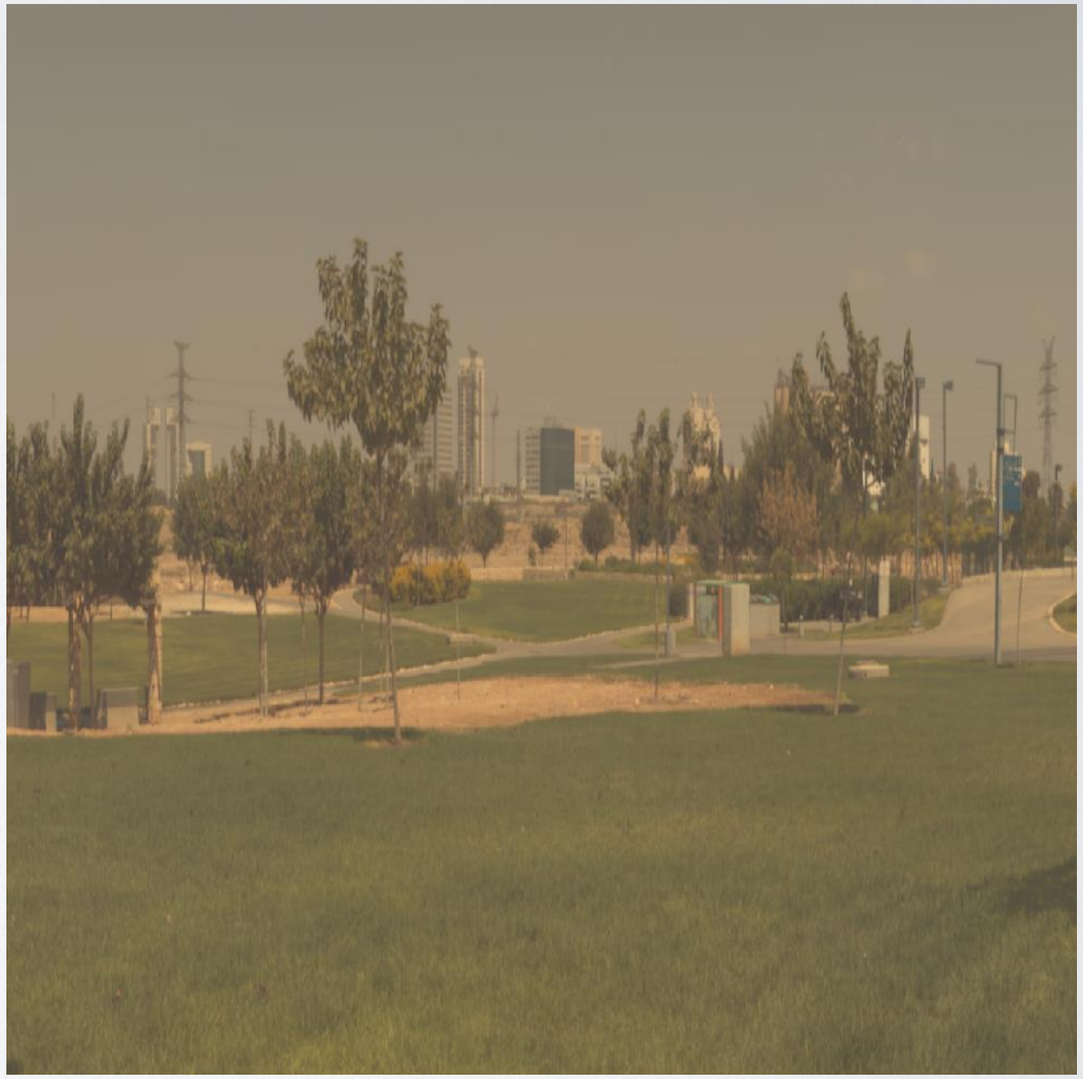
Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction



1280

1280

Experiments

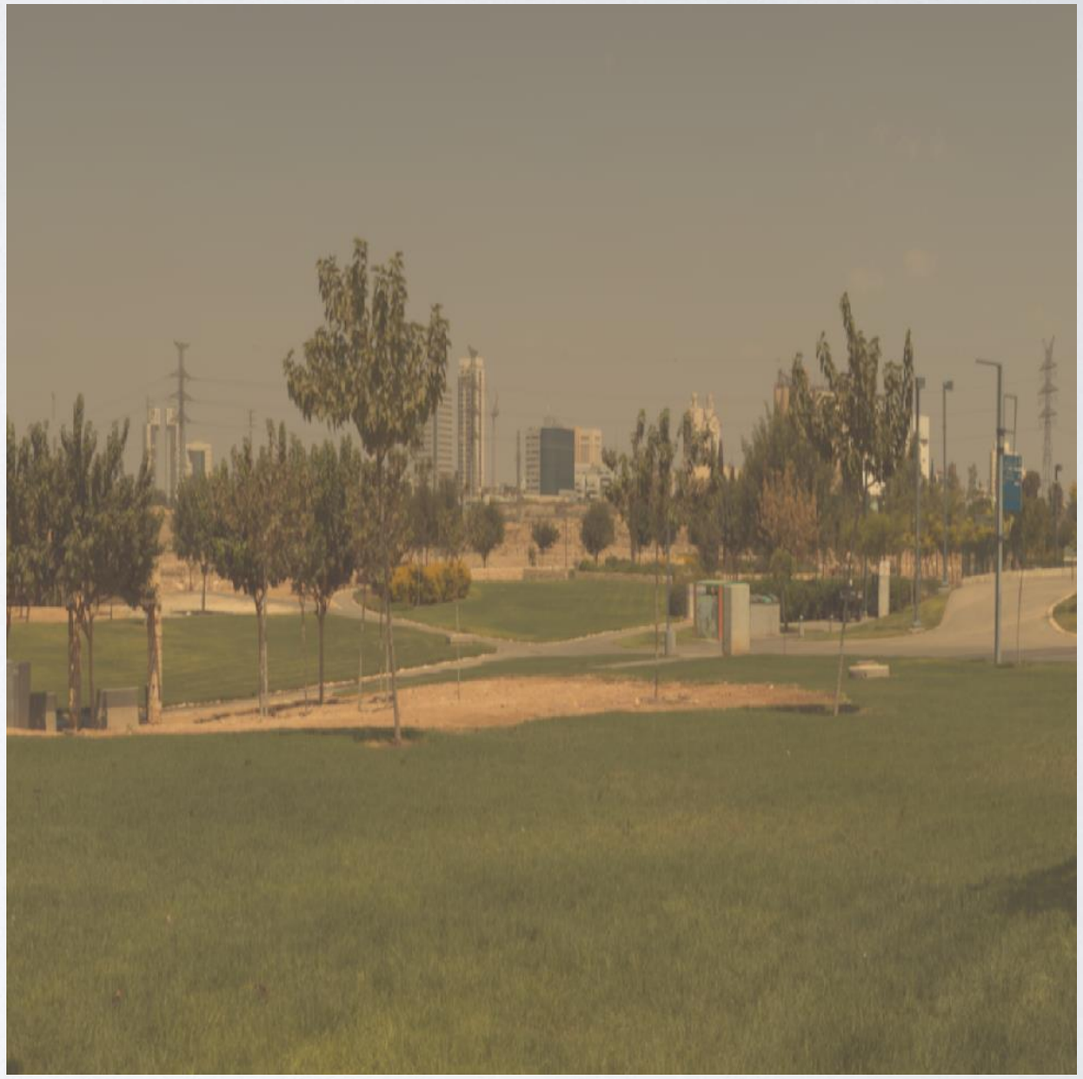
Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction



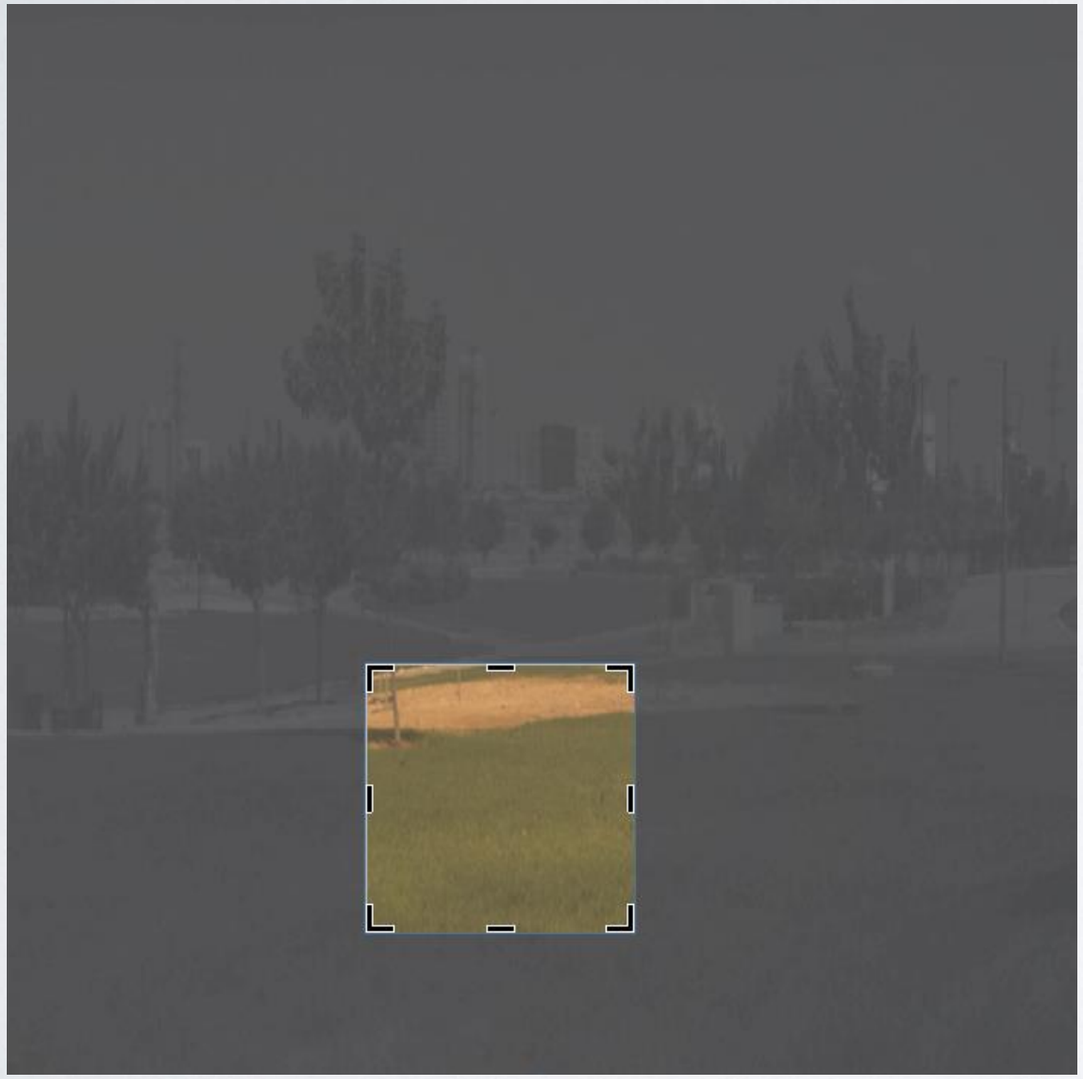
1280

1280

Experiments

Training and testing with large images

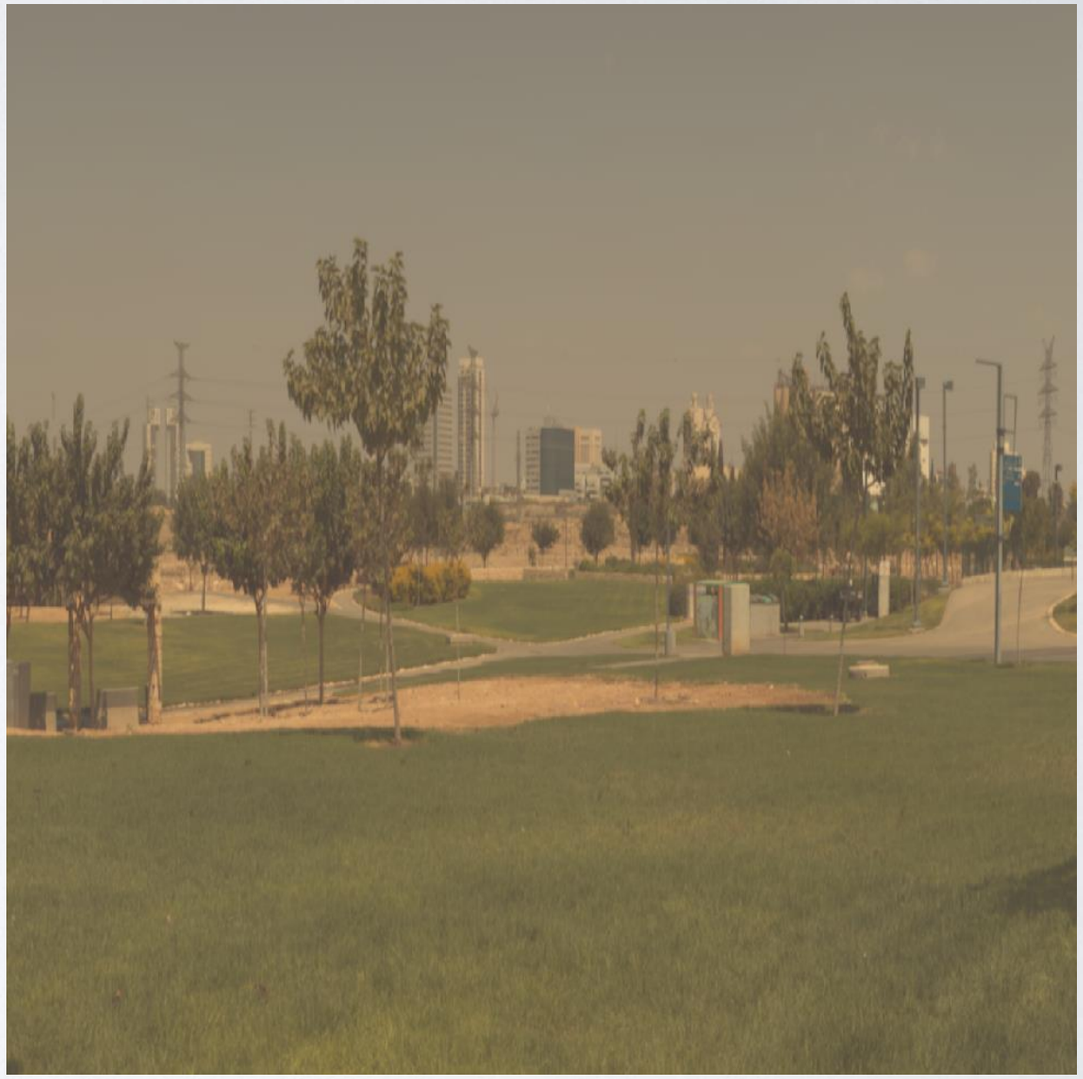
- Training: random crops



1280

1280

- Testing: tile decompositon + reconstruction



1280

1280

Experiments

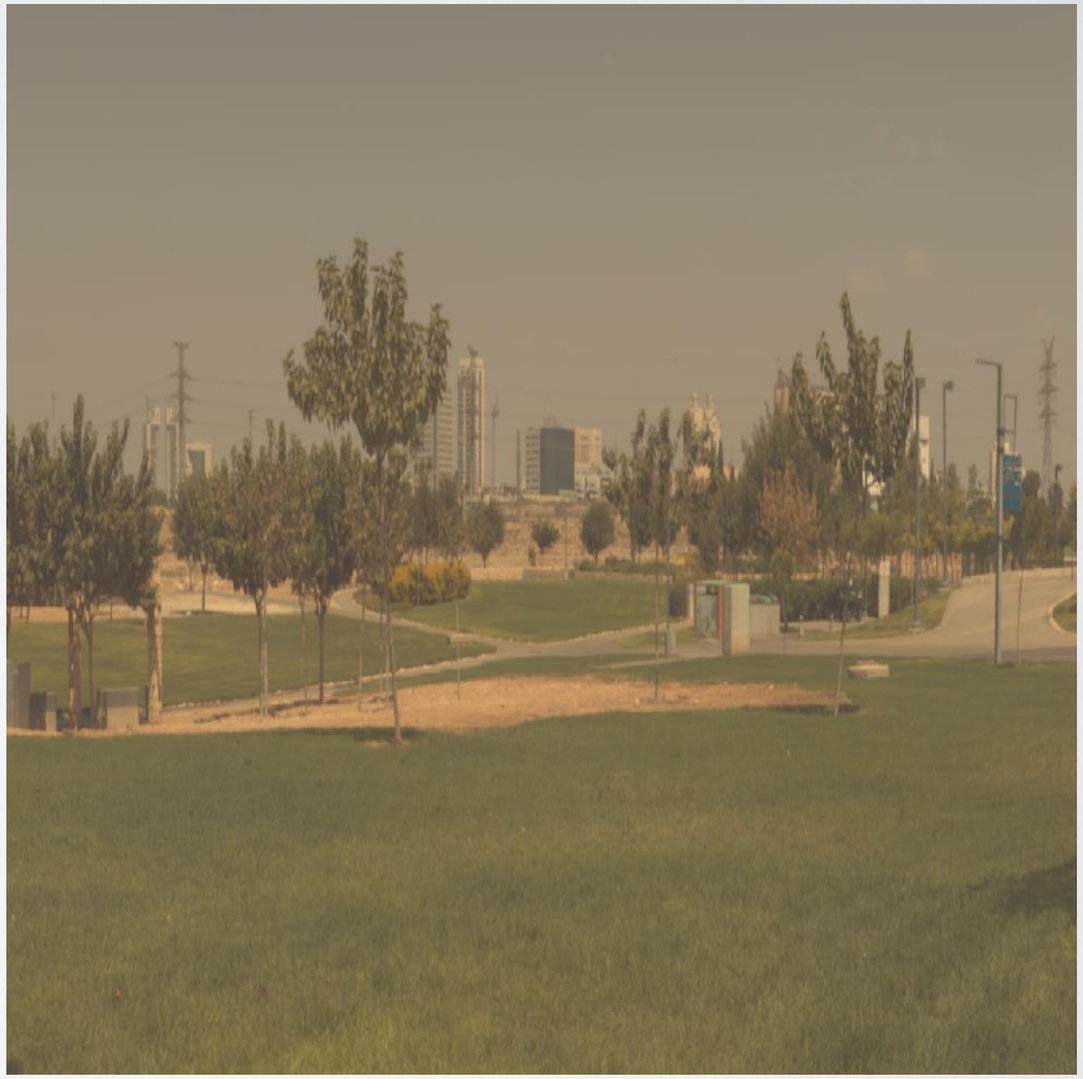
Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction



1280

1280

Experiments

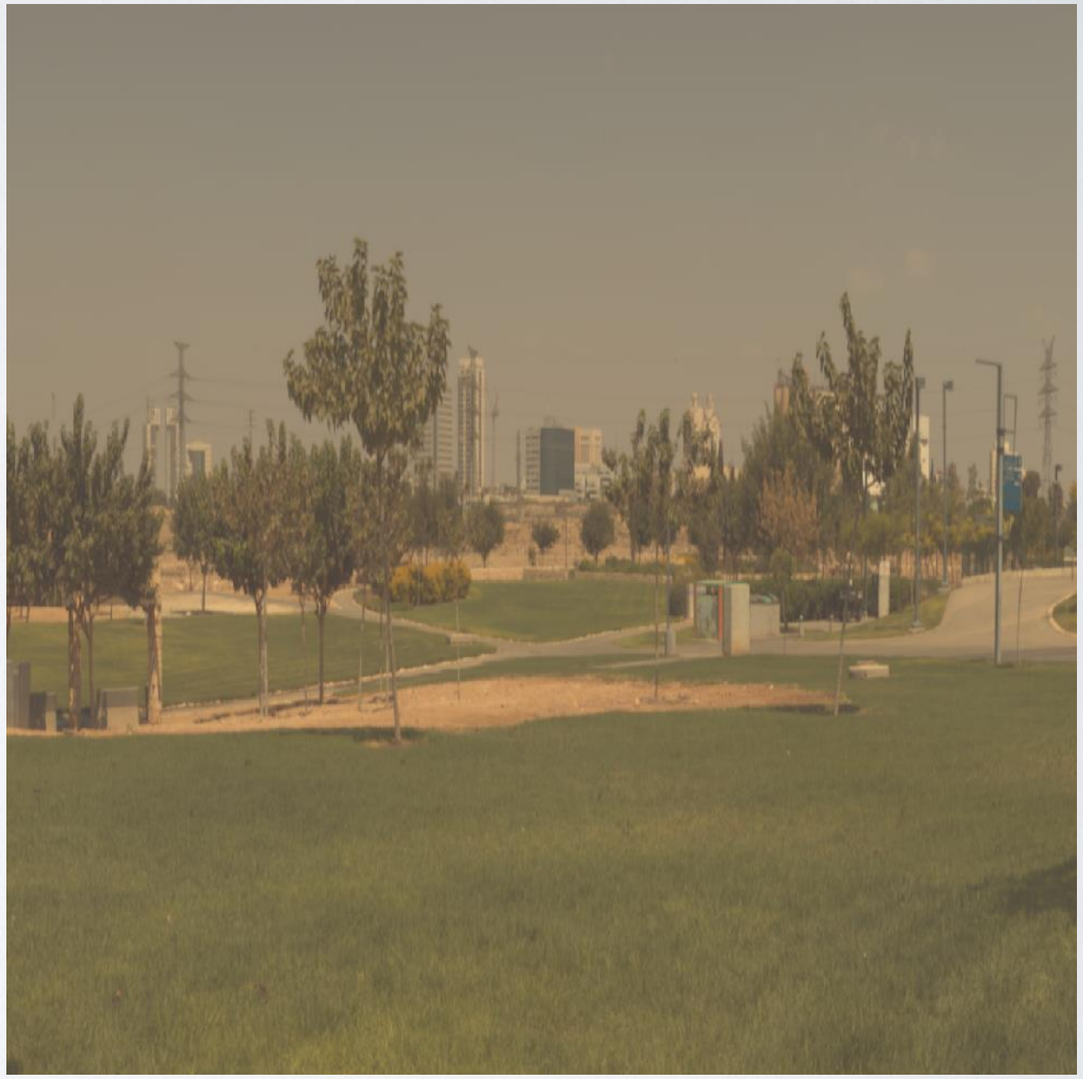
Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction



1280

1280

Experiments

Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction

1280

1280

Experiments

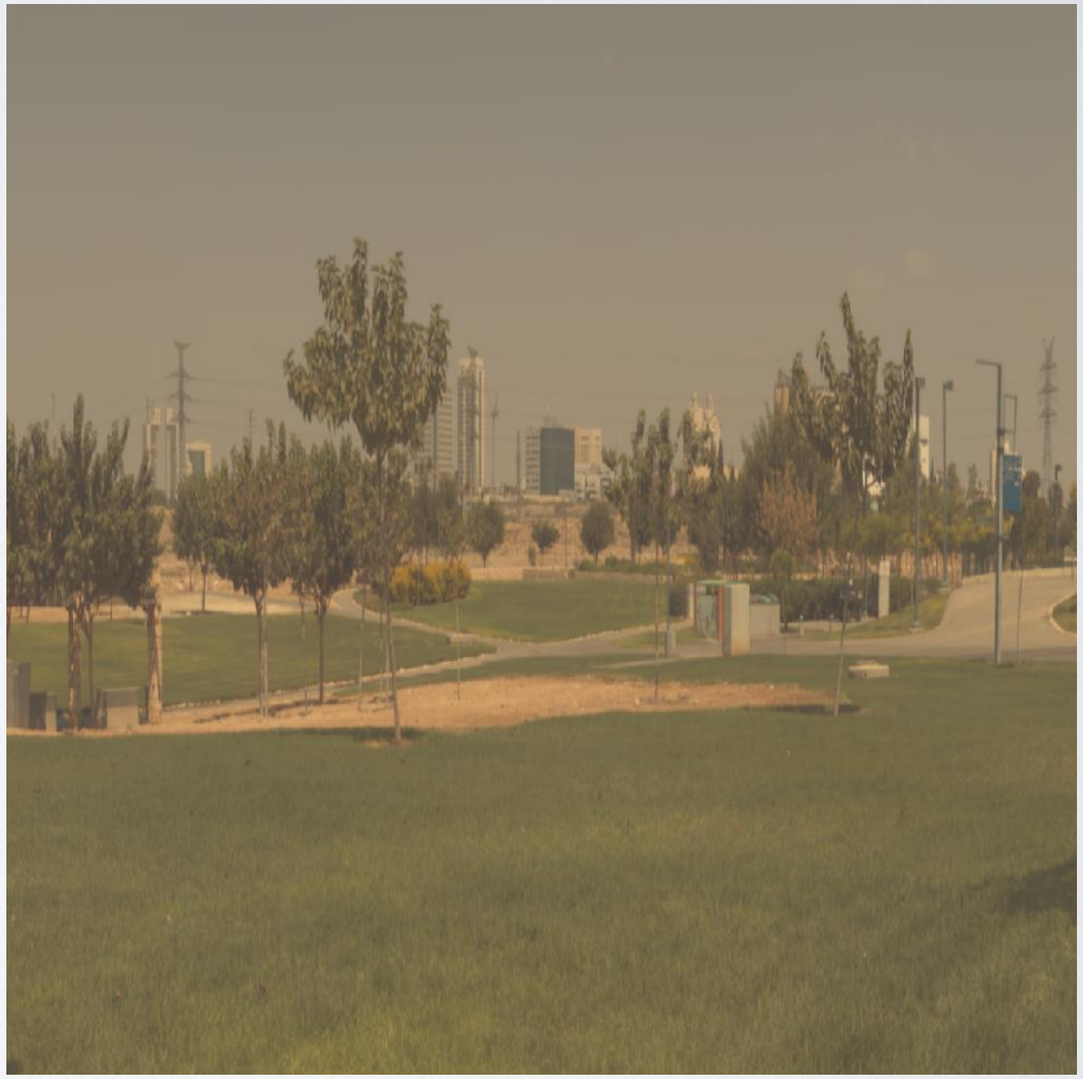
Training and testing with large images

- Training: random crops

1280

1280

- Testing: tile decompositon + reconstruction



1280

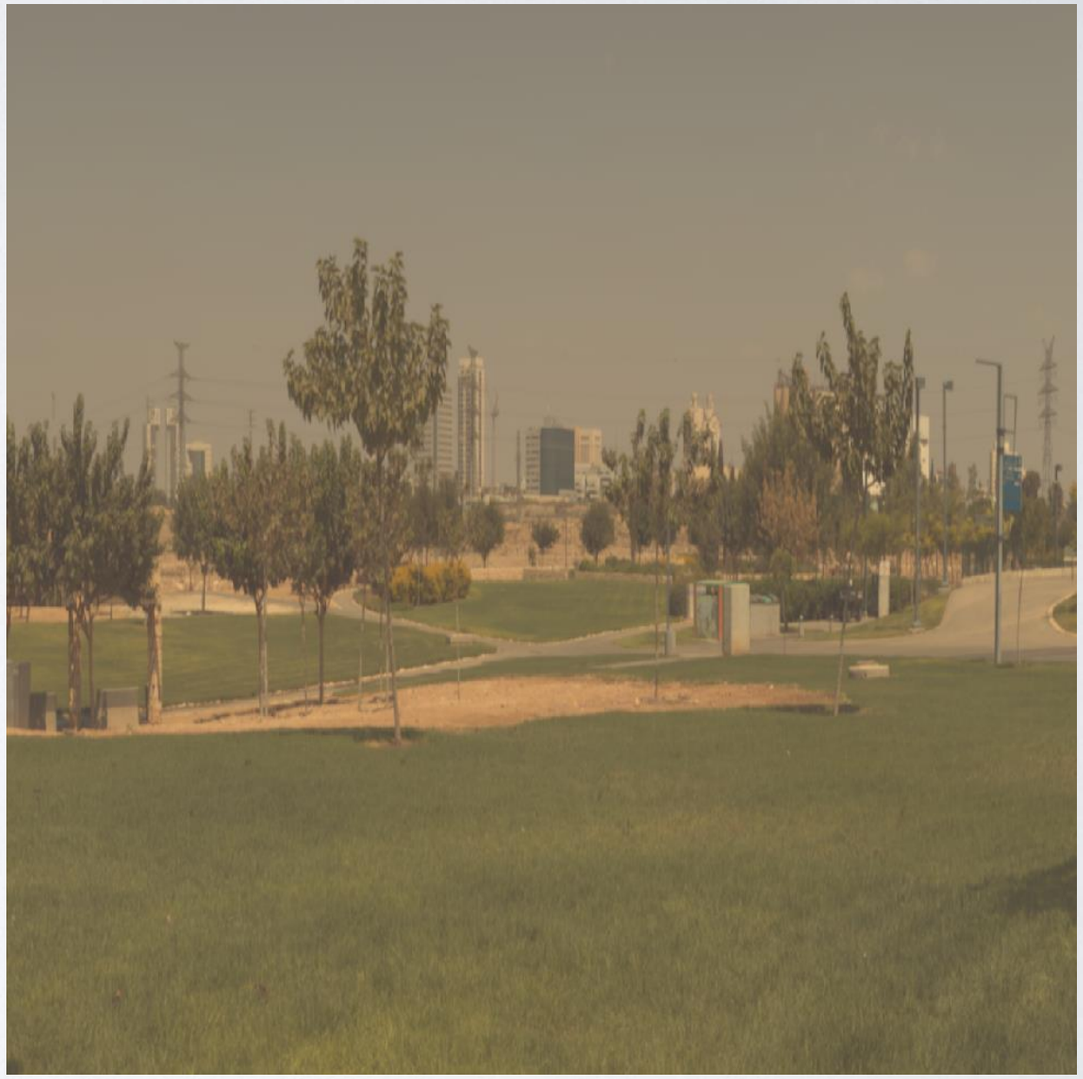
1280

Experiments

Training and testing with large images

- Training: random crops

- Testing: tile decompositon + reconstruction

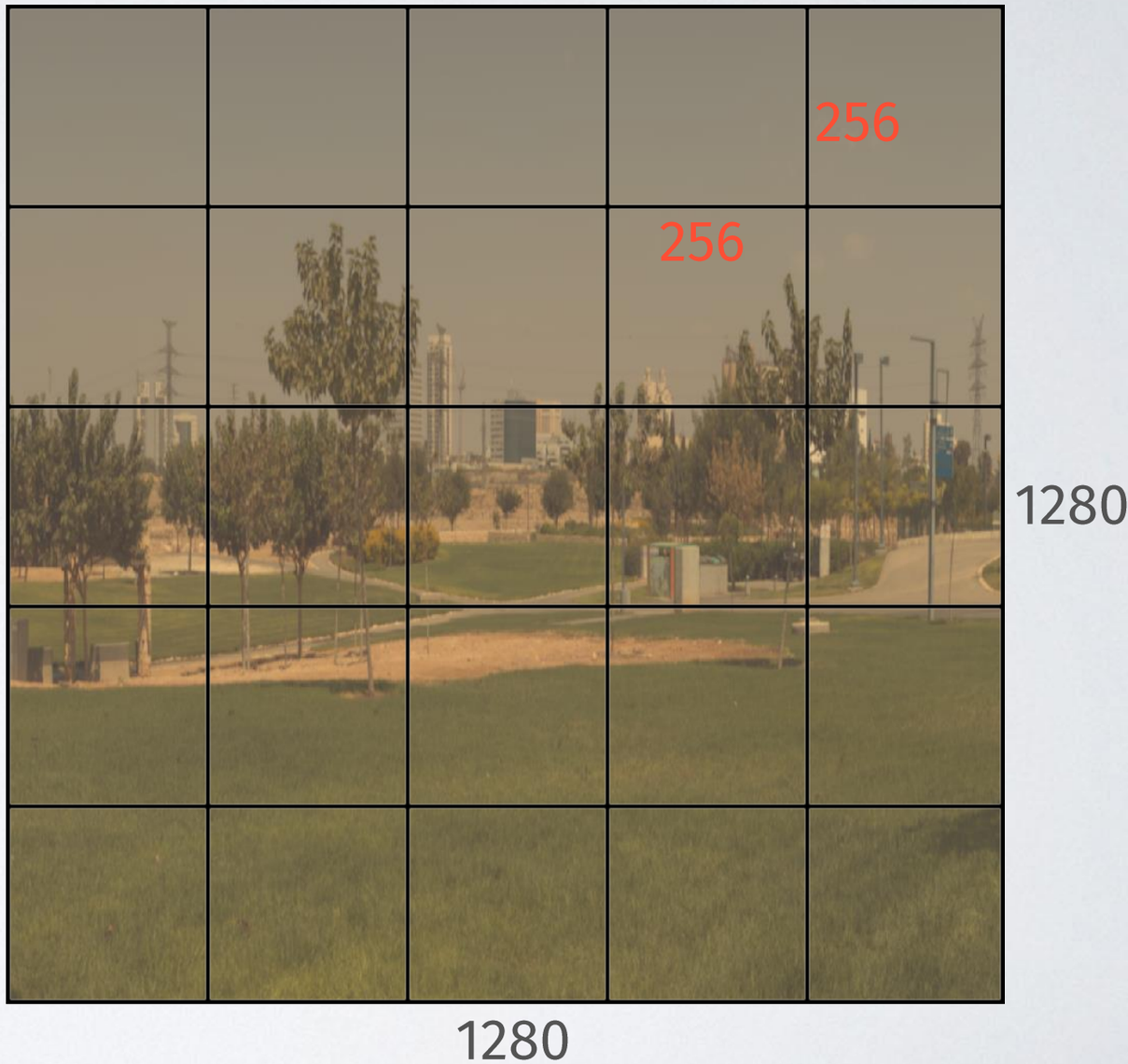


Experiments

Training and testing with large images

- Training: random crops

- Testing: tile decompositon + reconstruction



Results
Global

- Metrics (per pixel, pixel-count weighted):
  - **RMSE**: per-pixel Root Mean SE accross the spectrum dimension
  - **Relative RMSE**: Normalized for radiance level to account for samples with low luminance
  - **GFC**: Goodness of fit coefficient
  - **CIE ΔE_{00}**: Perceptual color difference

Results
Global

- Metrics (per pixel, pixel-count weighted):
  - **RMSE**: per-pixel Root Mean SE accross the spectrum dimension
  - **Relative RMSE**: Normalized for radiance level to account for samples with low luminance
  - **GFC**: Goodness of fit coefficient
  - **CIE ΔE_{00}**: Perceptual color difference

Method	RMSE	RMSERel	GFC	ΔE_{00}
[Arad2016] reported	2.633	0.0756	-	-
[Arad2016] default params	2.440 ± 0.066	0.1003 ± 0.022	–	–
[Arad2016] optimized	2.184 ± 0.064	0.0872 ± 0.004	–	–
[Galliani2017] reported	1.980	0.0587	–	–
Ours	$\mathbf{1.457 \pm 0.040}$	$\mathbf{0.0401 \pm 0.0024}$	$\mathbf{0.99921 \pm 0.00012}$	$\mathbf{2.044 \pm 0.341}$
fold 0	1.452 ± 0.101	0.0383 ± 0.0024	0.99906 ± 0.00001	1.861 ± 0.324
fold 1	1.463 ± 0.022	0.0420 ± 0.0024	0.99936 ± 0.00023	2.228 ± 0.358

Results
Global

- Metrics (per pixel, pixel-count weighted):
  - **RMSE**: per-pixel Root Mean SE accross the spectrum dimension
  - **Relative RMSE**: Normalized for radiance level to account for samples with low luminance
  - **GFC**: Goodness of fit coefficient
  - **CIE ΔE_{00}**: Perceptual color difference

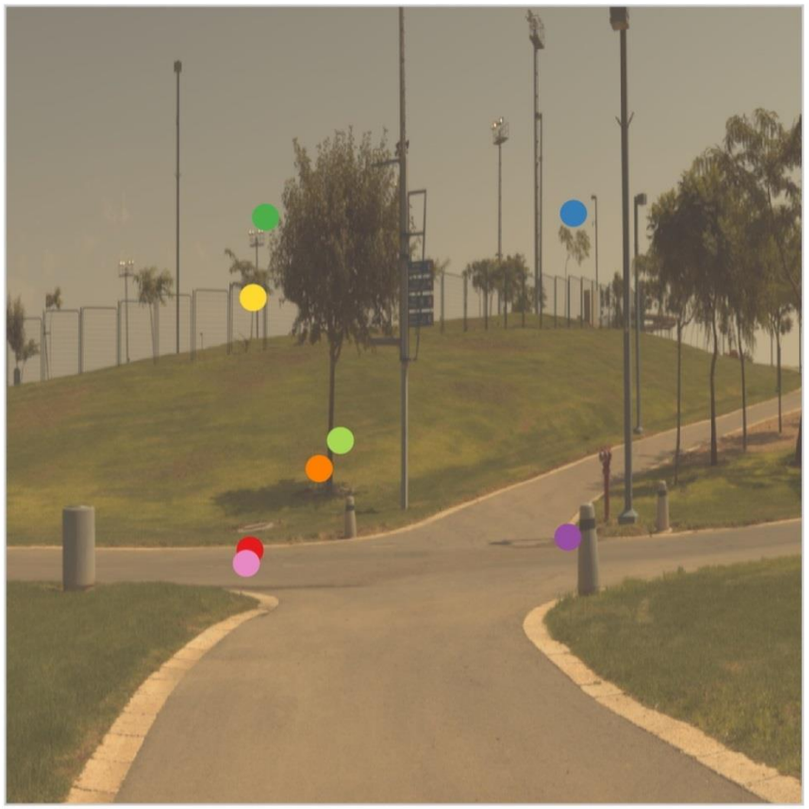
Method	RMSE	RMSERel	GFC	ΔE_{00}
[Arad2016] reported	2.633	0.0756	-	-
[Arad2016] default params	2.440 ± 0.066 -40.3%	0.1003 ± 0.022 -60.0%	–	–
[Arad2016] optimized	2.184 ± 0.064	0.0872 ± 0.004	–	–
[Galliani2017] reported	1.980	0.0587	–	–
Ours	$\mathbf{1.457 \pm 0.040}$	$\mathbf{0.0401 \pm 0.0024}$	$\mathbf{0.99921 \pm 0.00012}$	$\mathbf{2.044 \pm 0.341}$
fold 0	1.452 ± 0.101	0.0383 ± 0.0024	0.99906 ± 0.00001	1.861 ± 0.324
fold 1	1.463 ± 0.022	0.0420 ± 0.0024	0.99936 ± 0.00023	2.228 ± 0.358

Results
Global

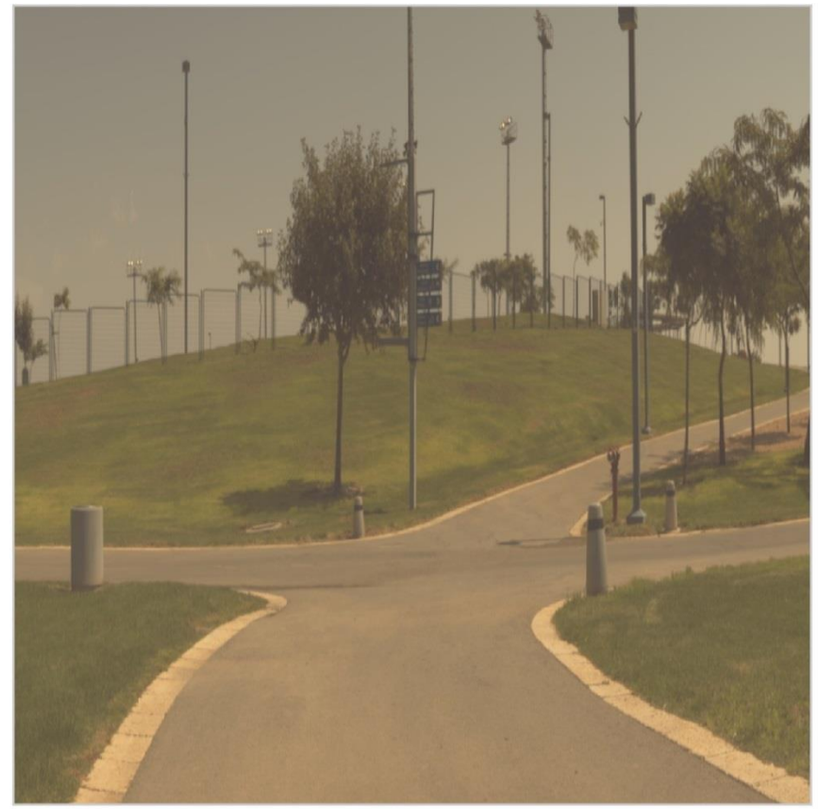
- Metrics (per pixel, pixel-count weighted):
  - **RMSE**: per-pixel Root Mean SE accross the spectrum dimension
  - **Relative RMSE**: Normalized for radiance level to account for samples with low luminance
  - **GFC**: Goodness of fit coefficient
  - **CIE ΔE_{00}**: Perceptual color difference

Method	RMSE	RMSERel	GFC	ΔE_{00}
[Arad2016] reported	2.633	0.0756	-	-
[Arad2016] default params	2.440 ± 0.066 -40.3%	0.1003 ± 0.022 -60.0%	–	–
[Arad2016] optimized	2.184 ± 0.064 -33.2%	0.0872 ± 0.004 -54,0%	–	–
[Galliani2017] reported	1.980	0.0587	–	–
Ours	$\mathbf{1.457 \pm 0.040}$	$\mathbf{0.0401 \pm 0.0024}$	$\mathbf{0.99921 \pm 0.00012}$	$\mathbf{2.044 \pm 0.341}$
fold 0	1.452 ± 0.101	0.0383 ± 0.0024	0.99906 ± 0.00001	1.861 ± 0.324
fold 1	1.463 ± 0.022	0.0420 ± 0.0024	0.99936 ± 0.00023	2.228 ± 0.358

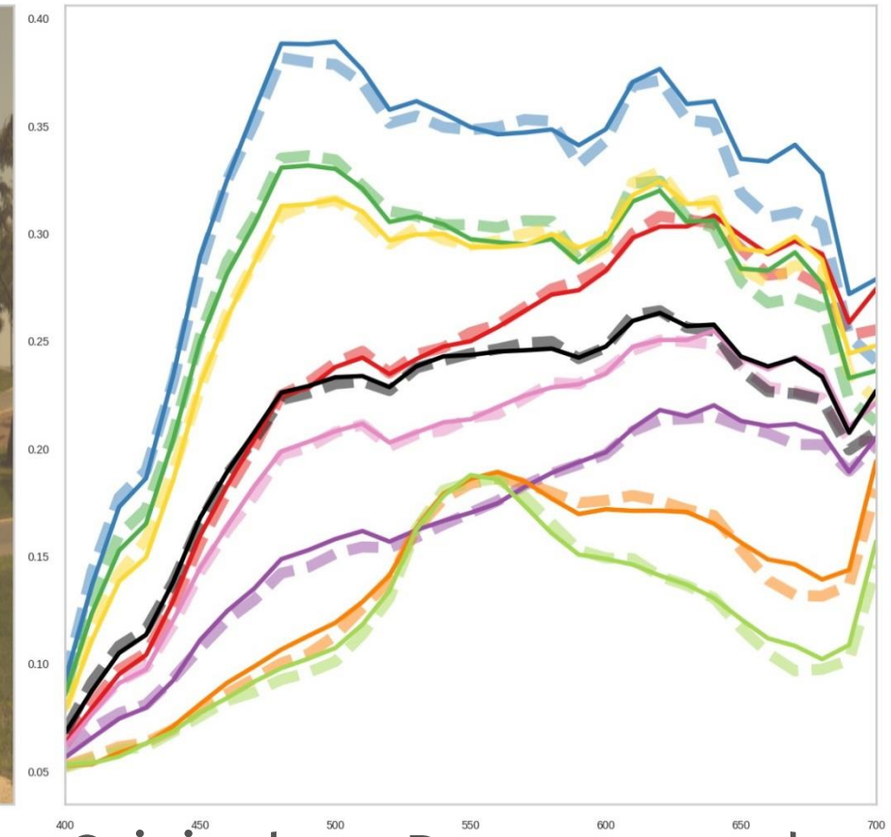
Results



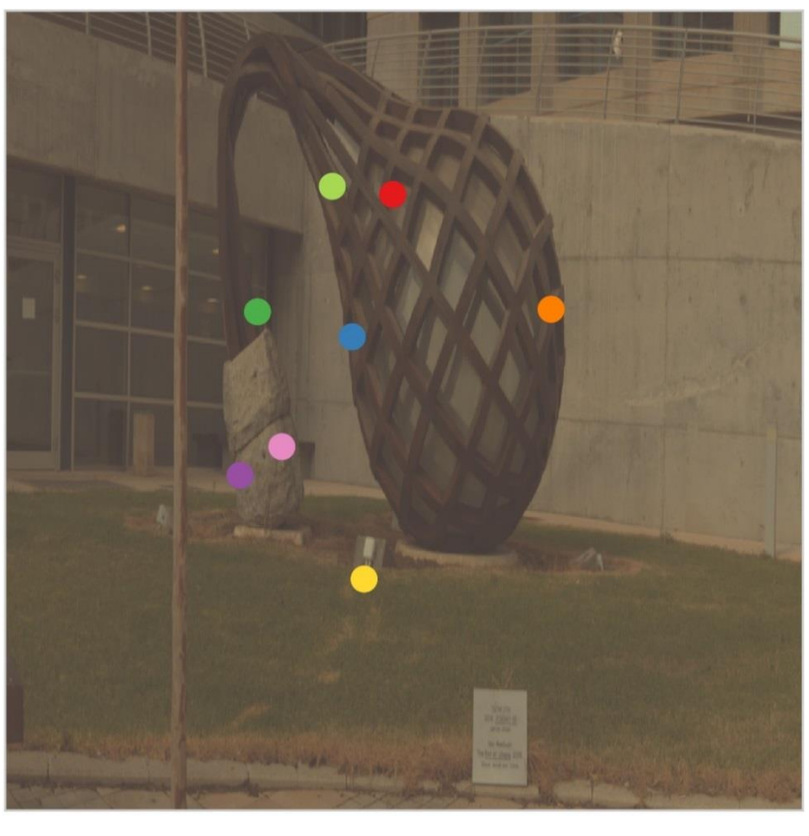
Original (sRGB render)

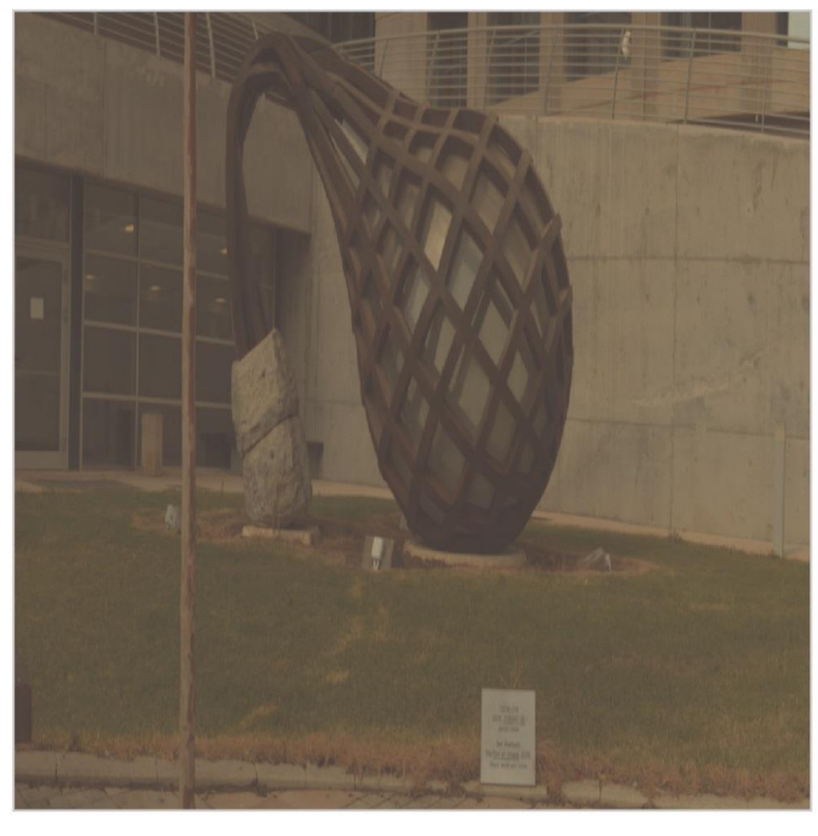


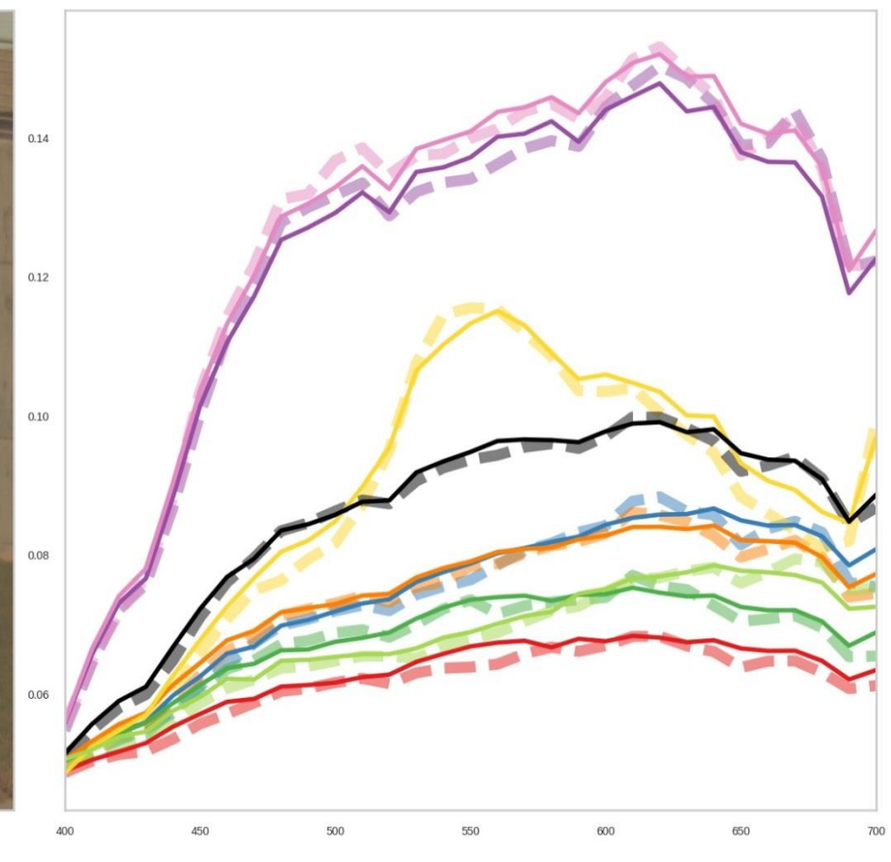
Reconstructed (sRGB render)



Original --- Reconstructed —



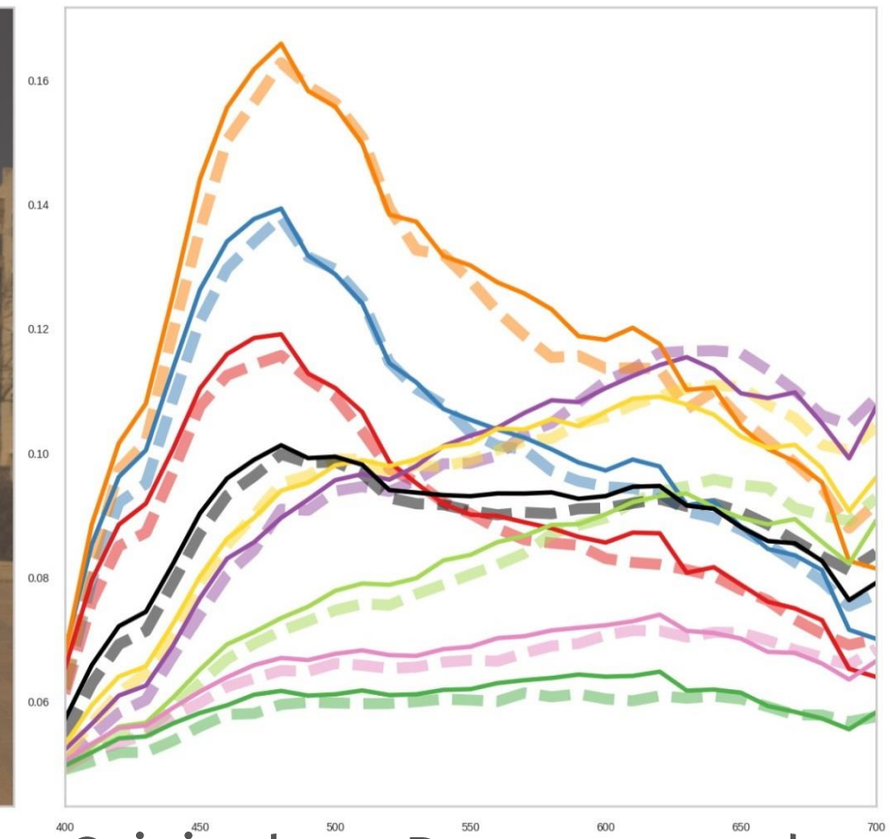




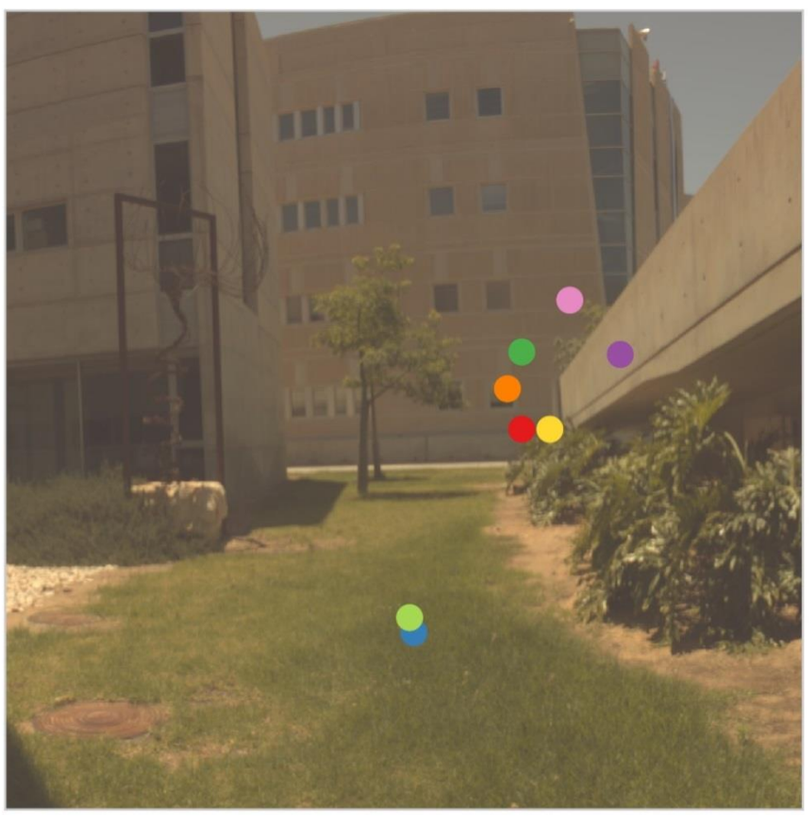
Results

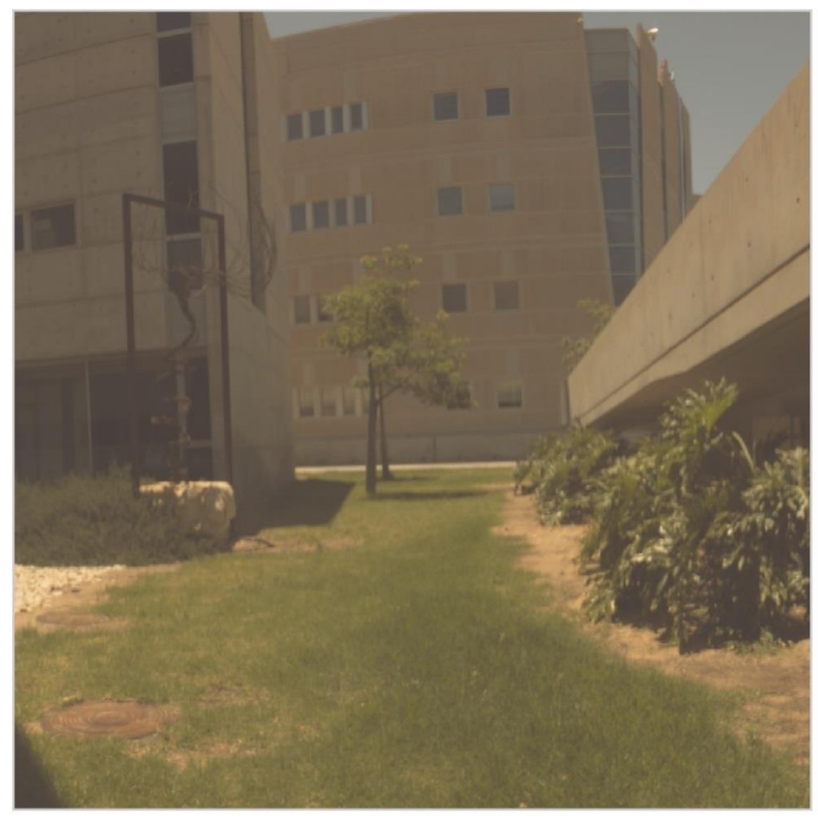
Original (sRGB render)

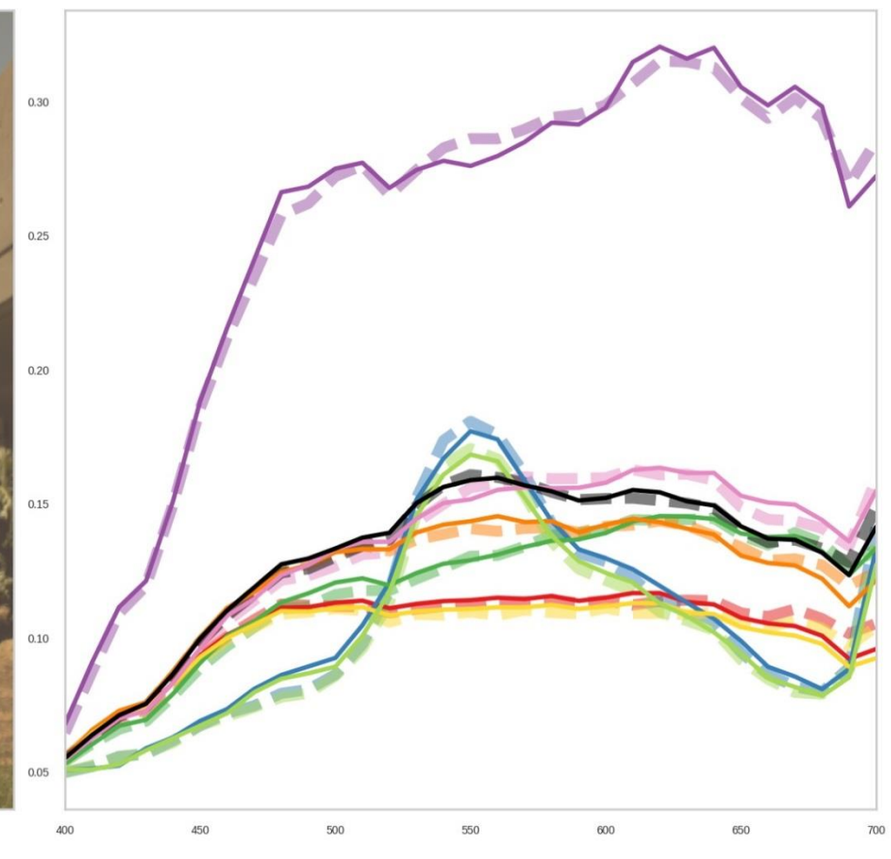
Reconstructed (sRGB render)



Original --- Reconstructed —



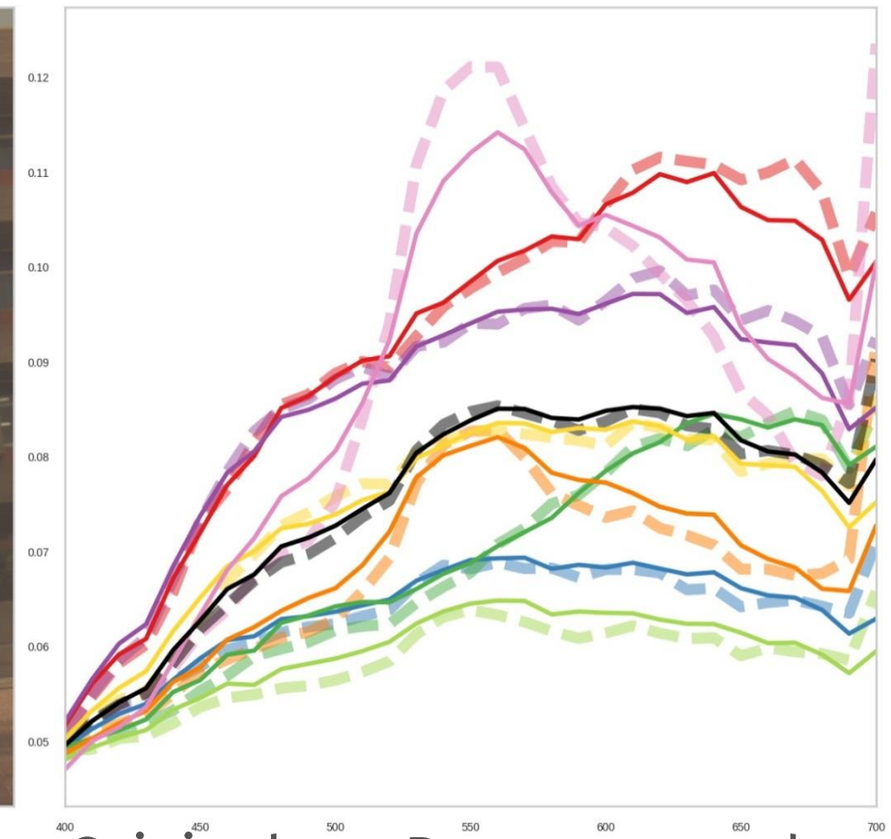




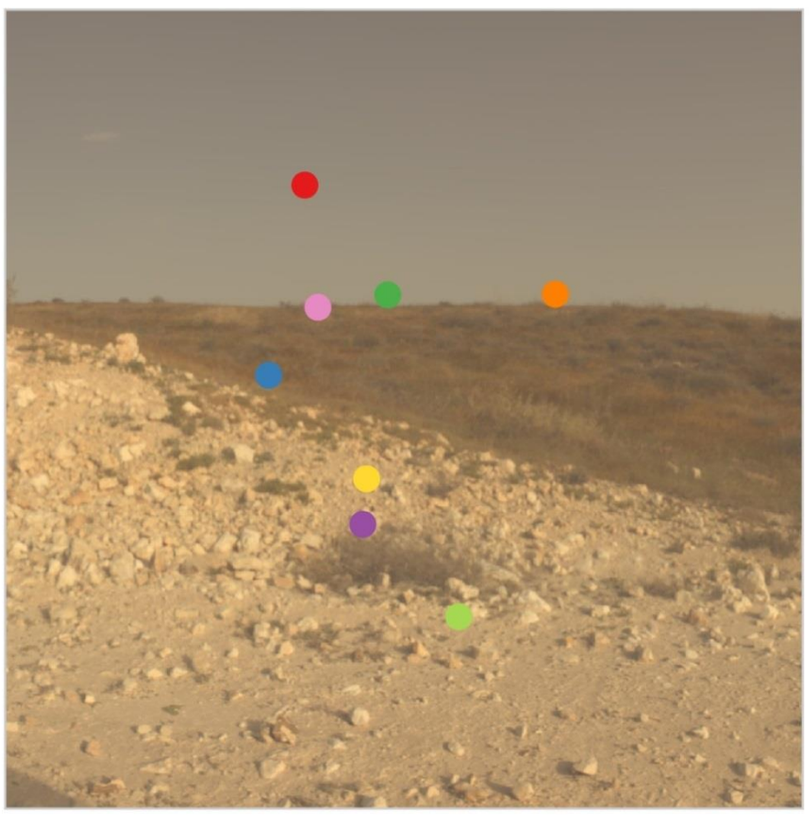
Results

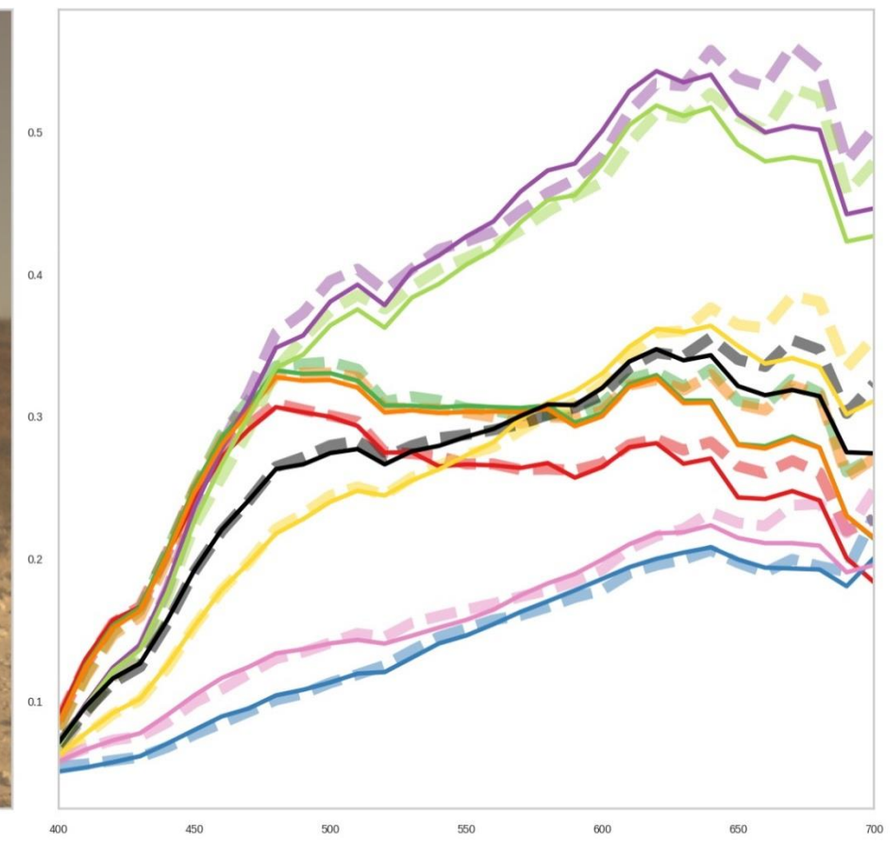
Original (sRGB render)

Reconstructed (sRGB render)



Original --- Reconstructed —





Results

Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

Results

Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

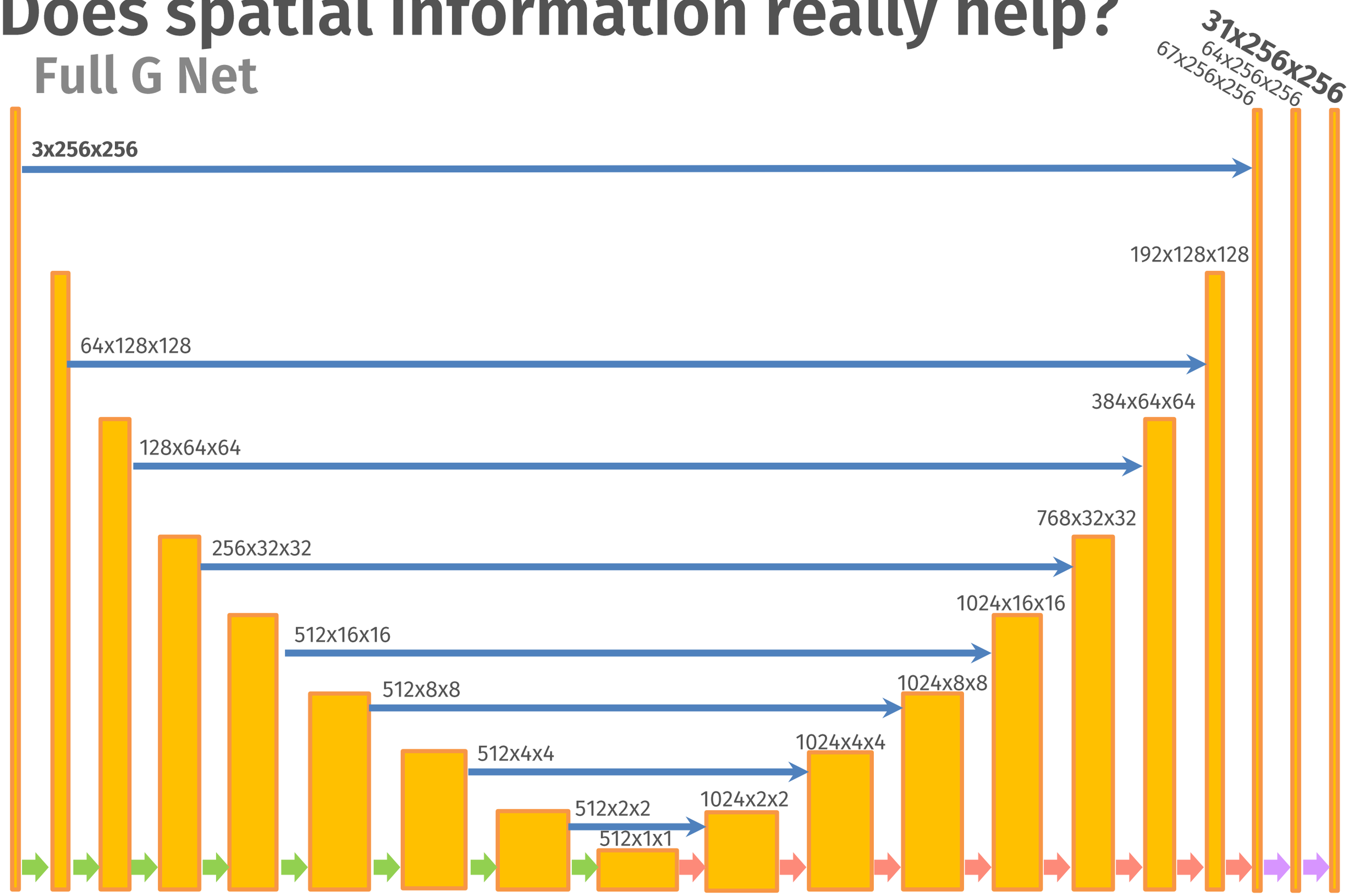
Results

Does spatial information really help?

Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

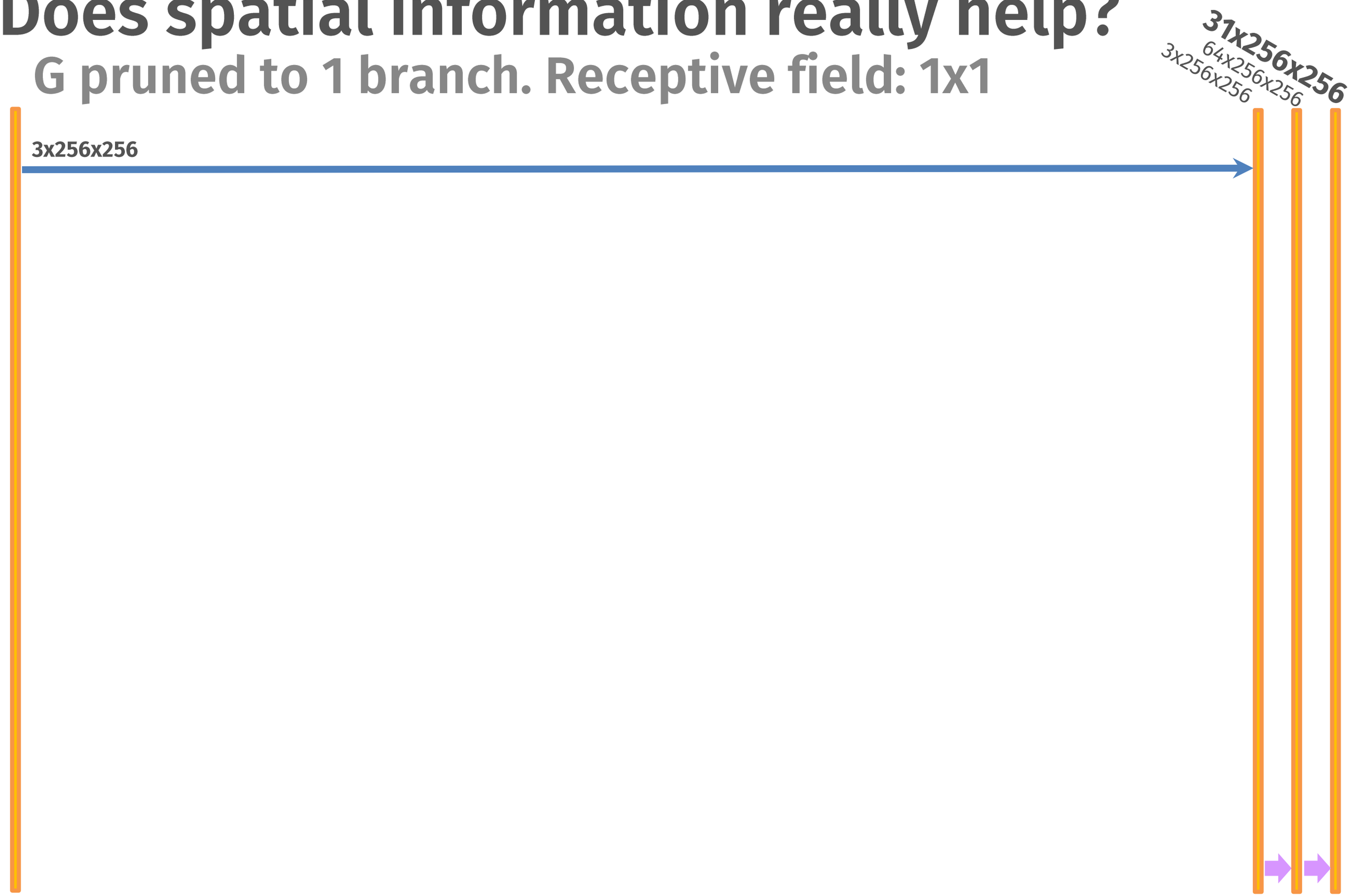
Does spatial information really help?

Full G Net



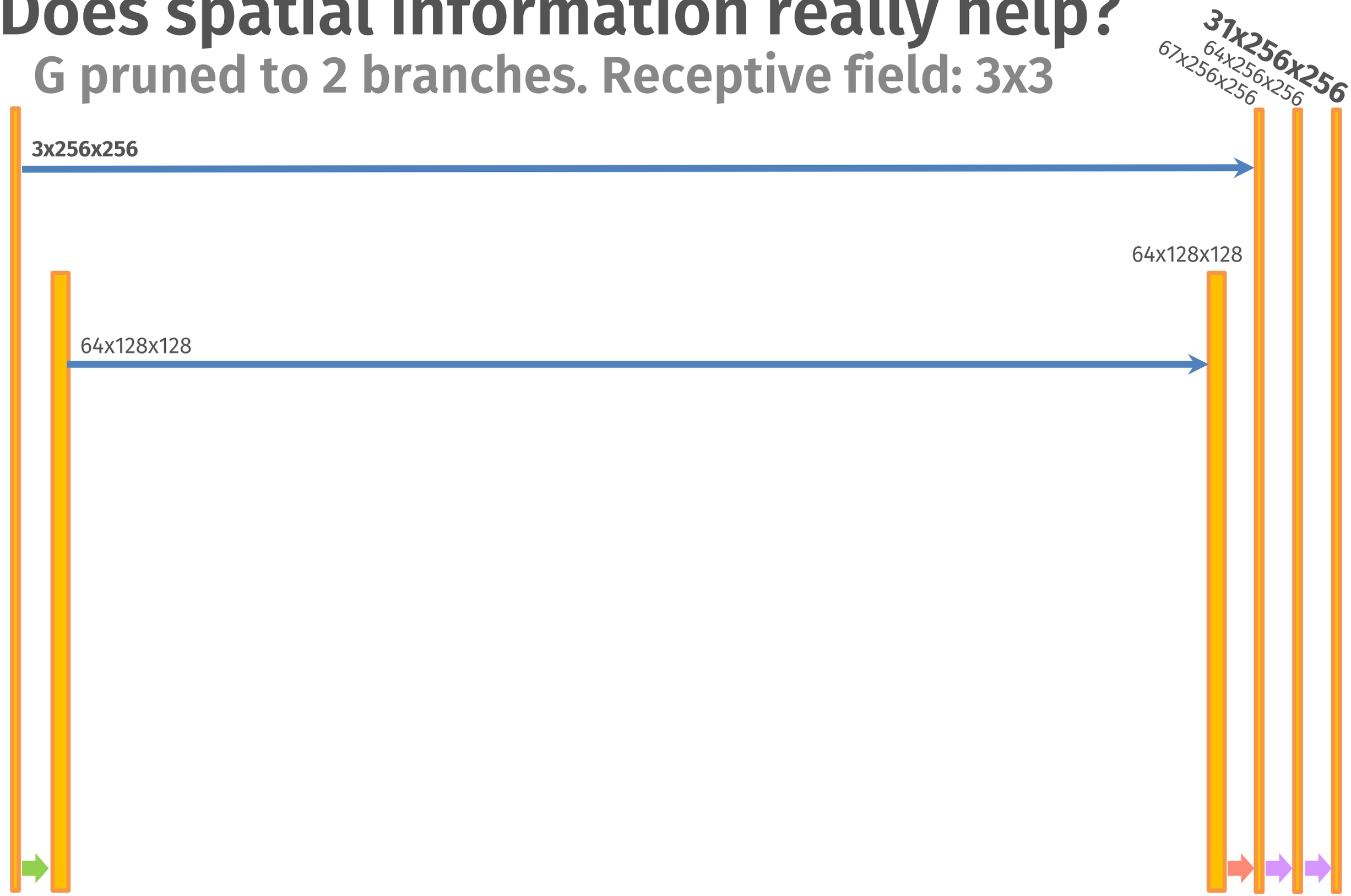
Does spatial information really help?

G pruned to 1 branch. Receptive field: 1x1



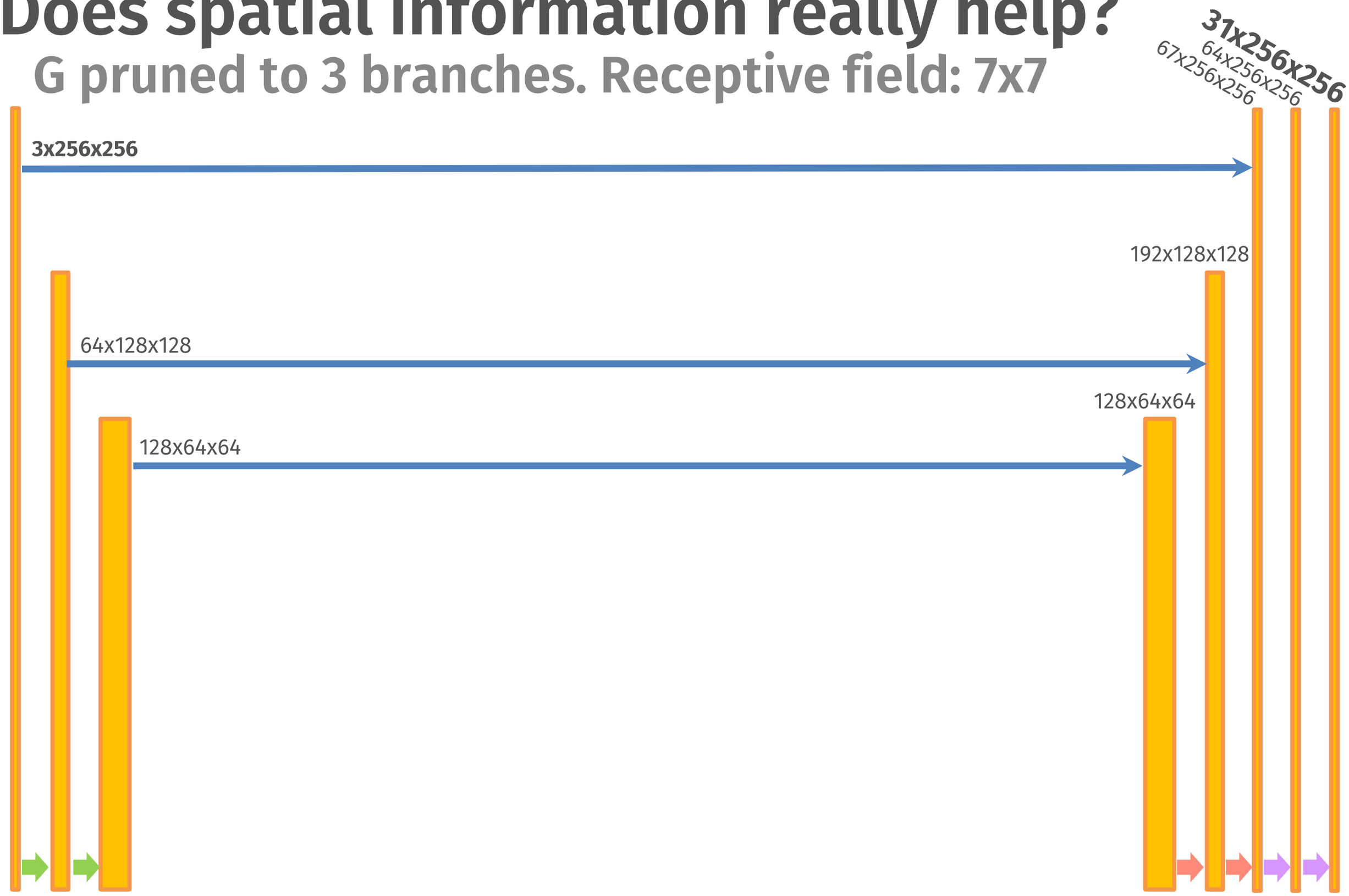
Does spatial information really help?

G pruned to 2 branches. Receptive field: 3x3



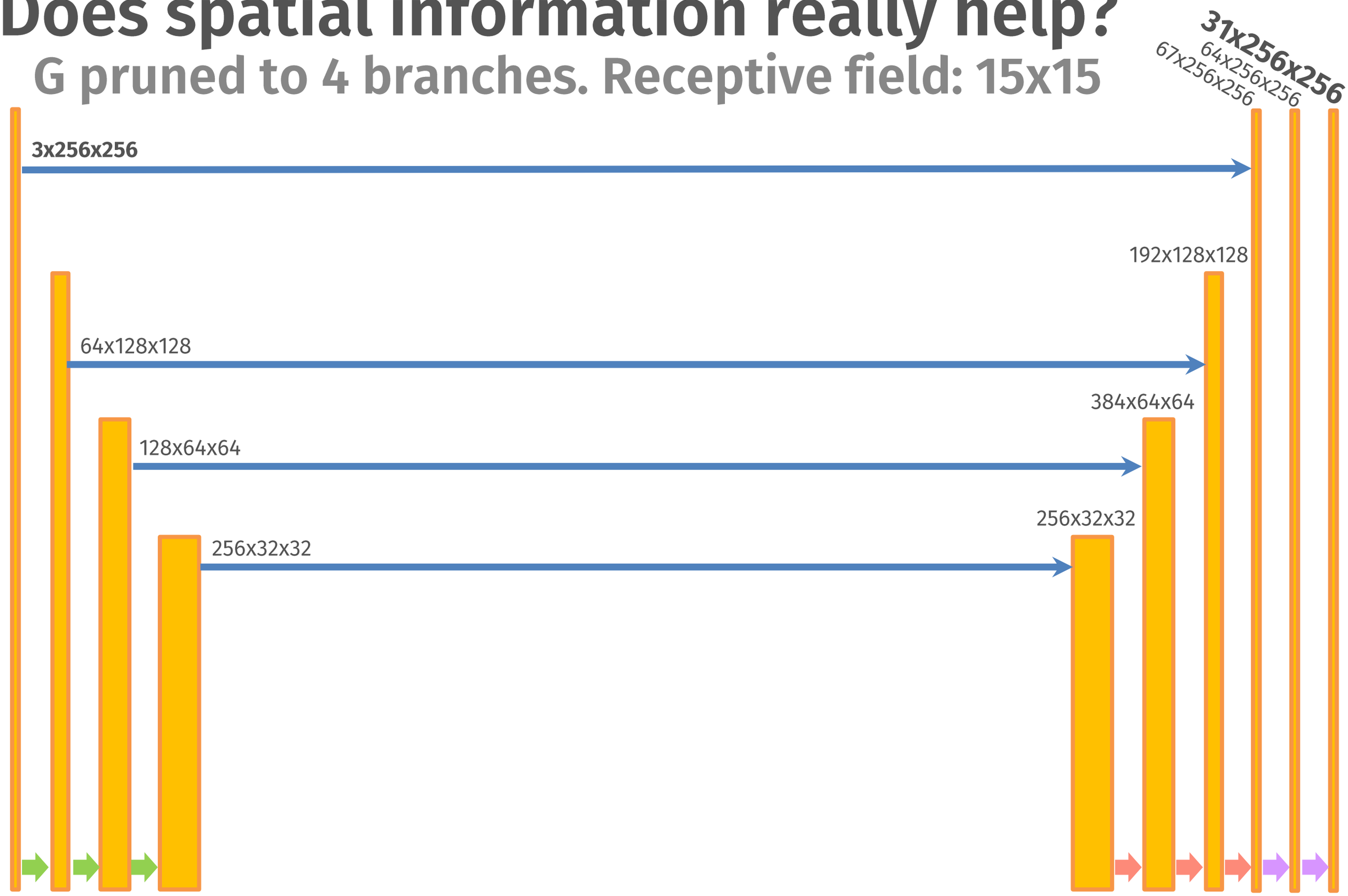
Does spatial information really help?

G pruned to 3 branches. Receptive field: 7x7



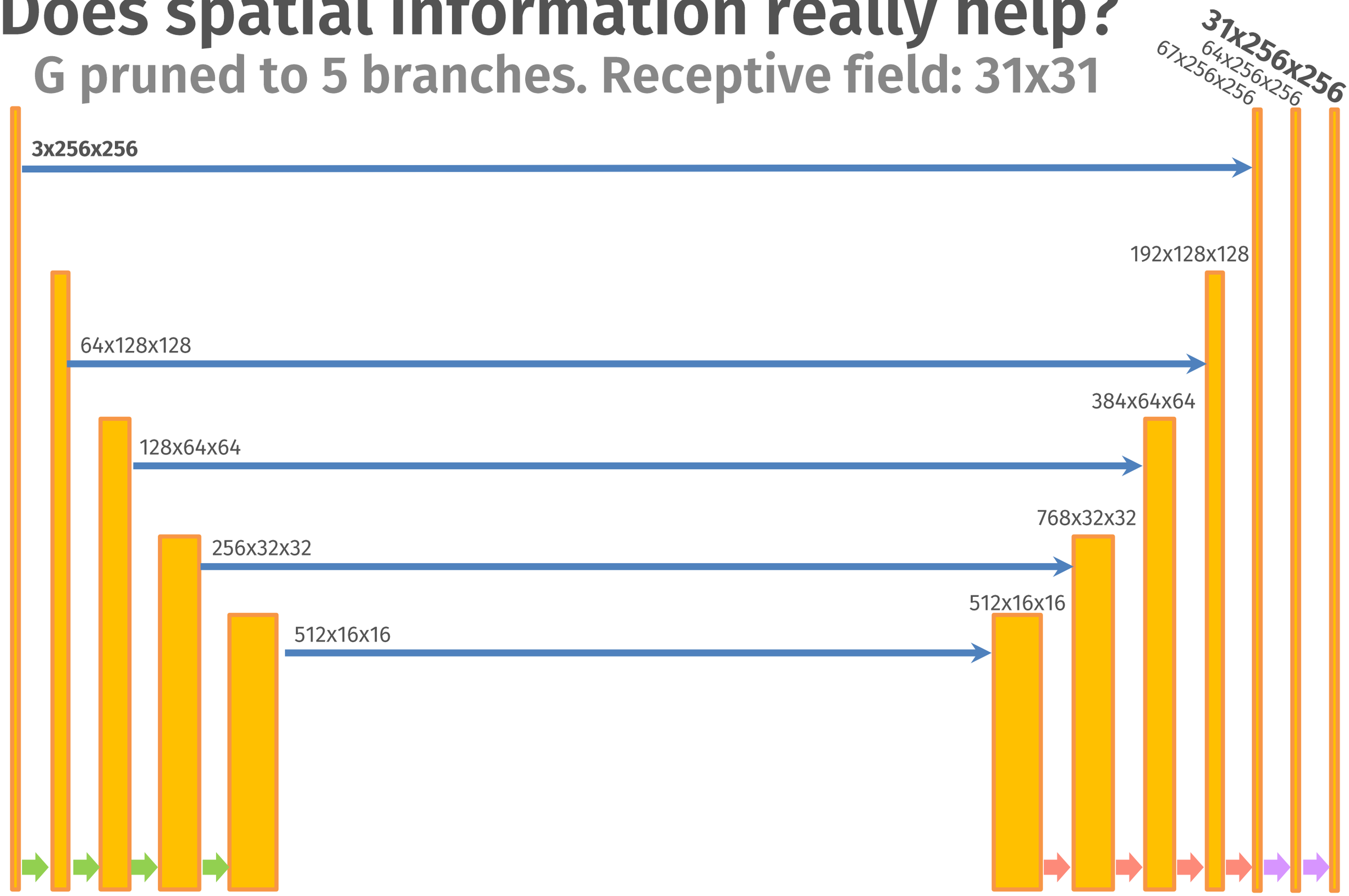
Does spatial information really help?

G pruned to 4 branches. Receptive field: 15x15



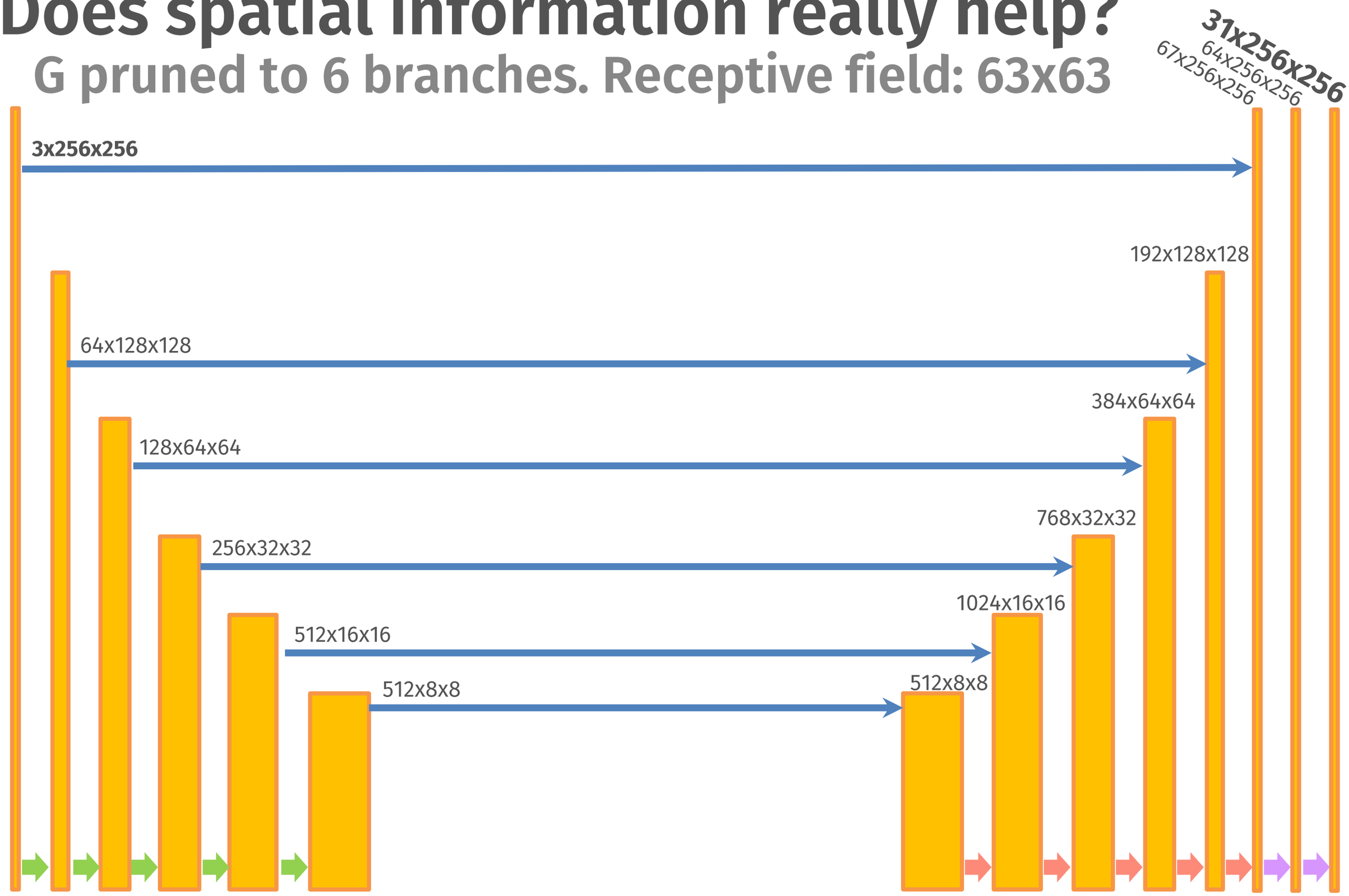
Does spatial information really help?

G pruned to 5 branches. Receptive field: 31x31



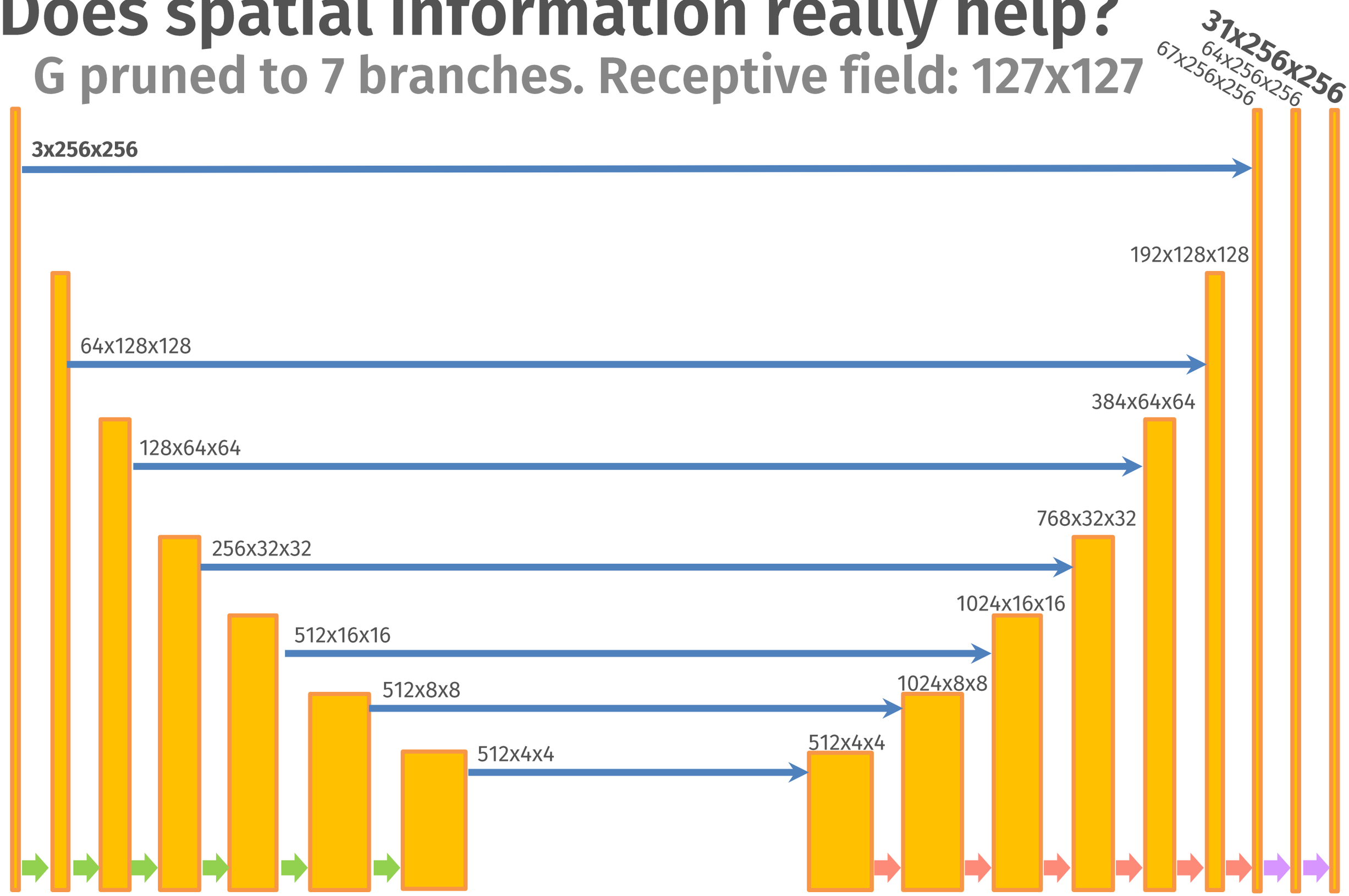
Does spatial information really help?

G pruned to 6 branches. Receptive field: 63x63



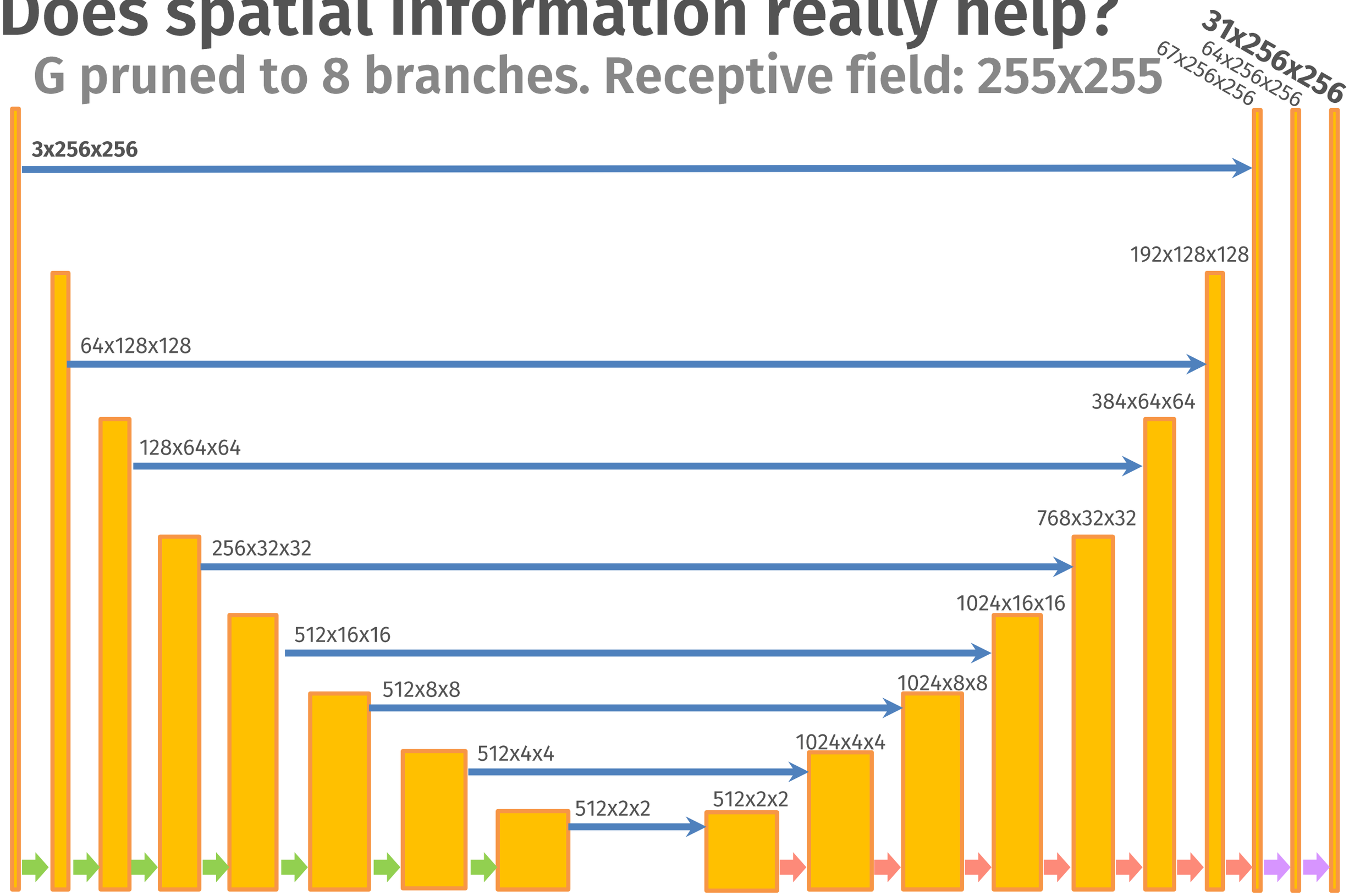
Does spatial information really help?

G pruned to 7 branches. Receptive field: 127x127



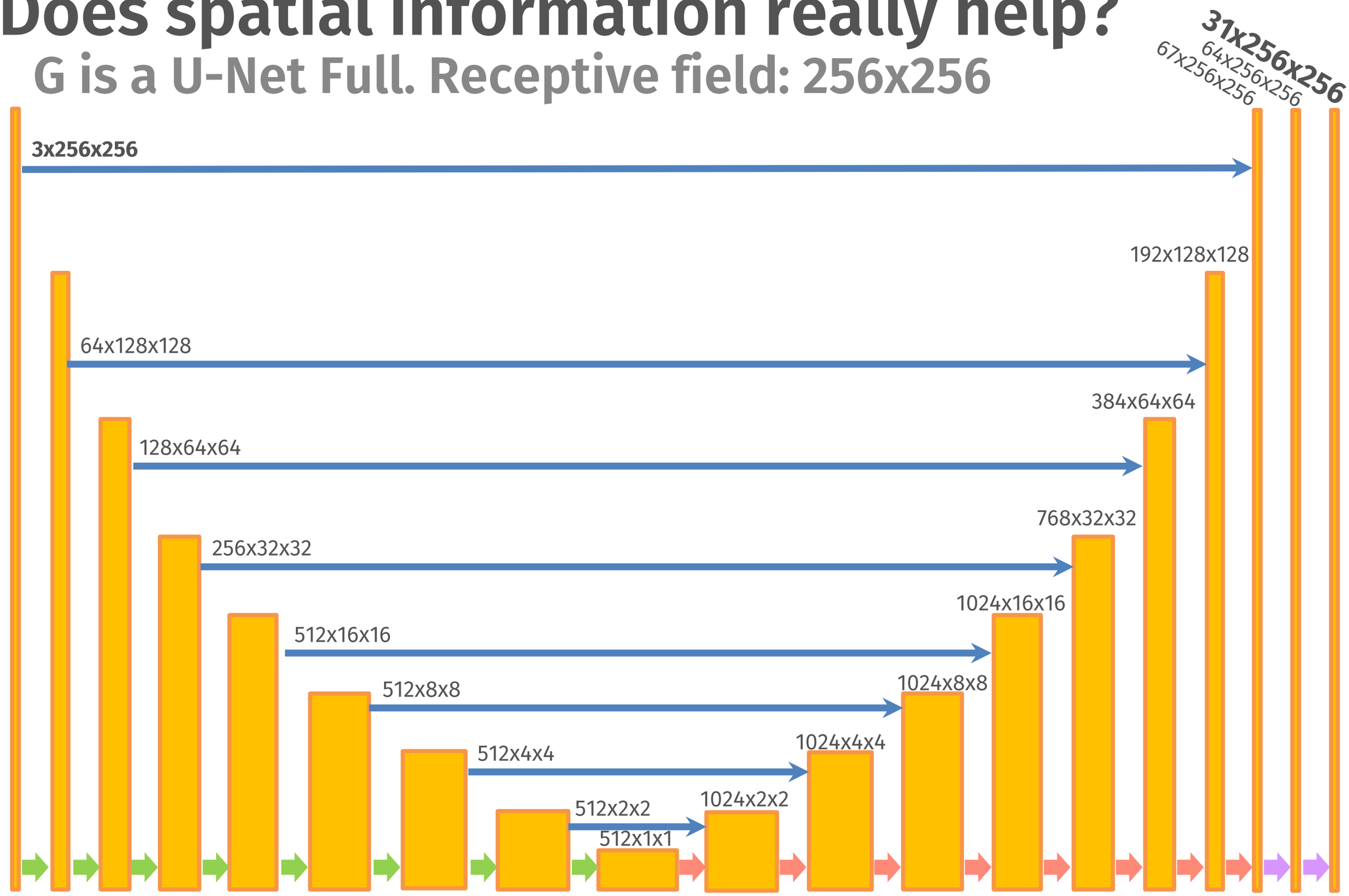
Does spatial information really help?

G pruned to 8 branches. Receptive field: 255x255



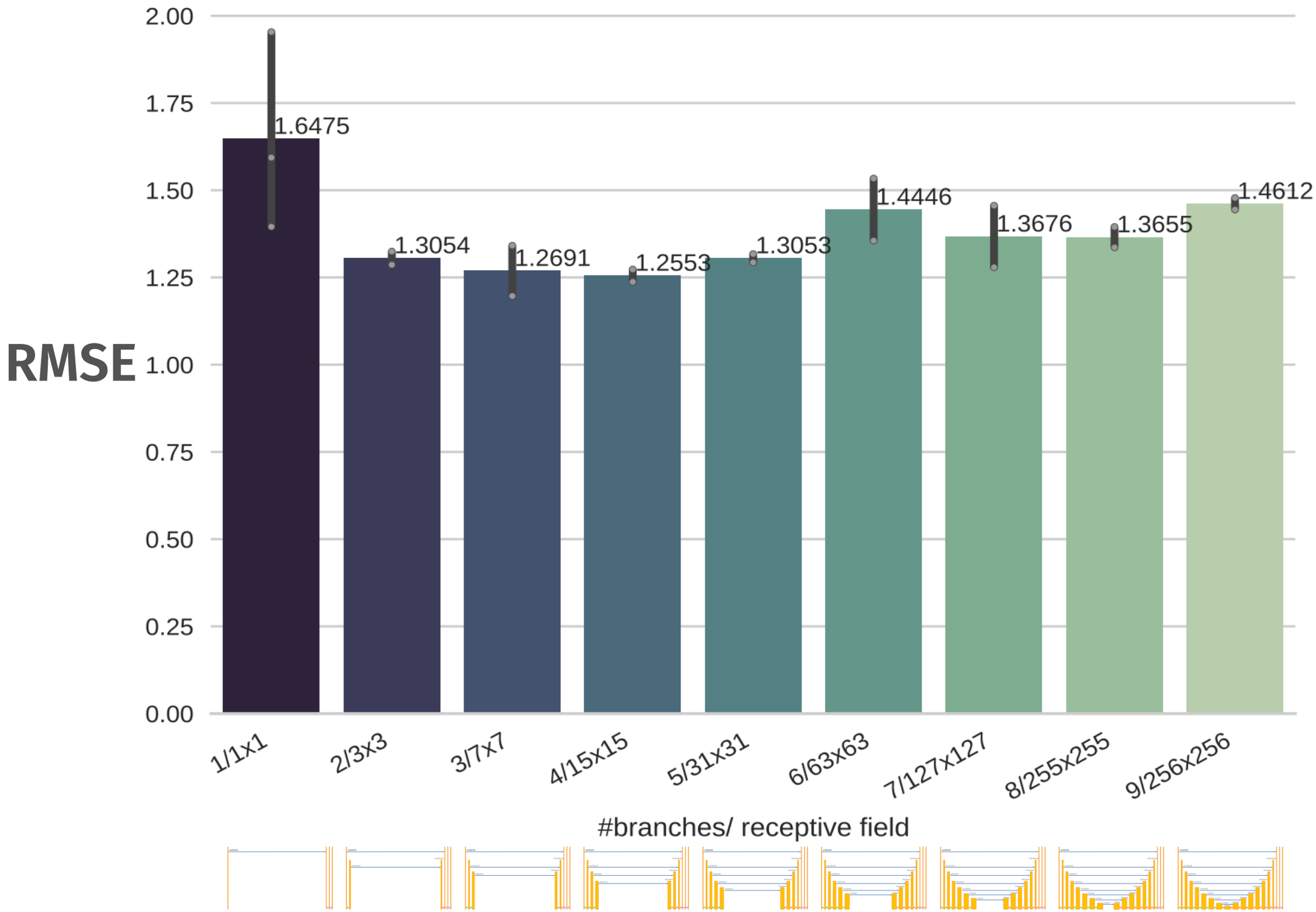
Does spatial information really help?

G is a U-Net Full. Receptive field: 256x256



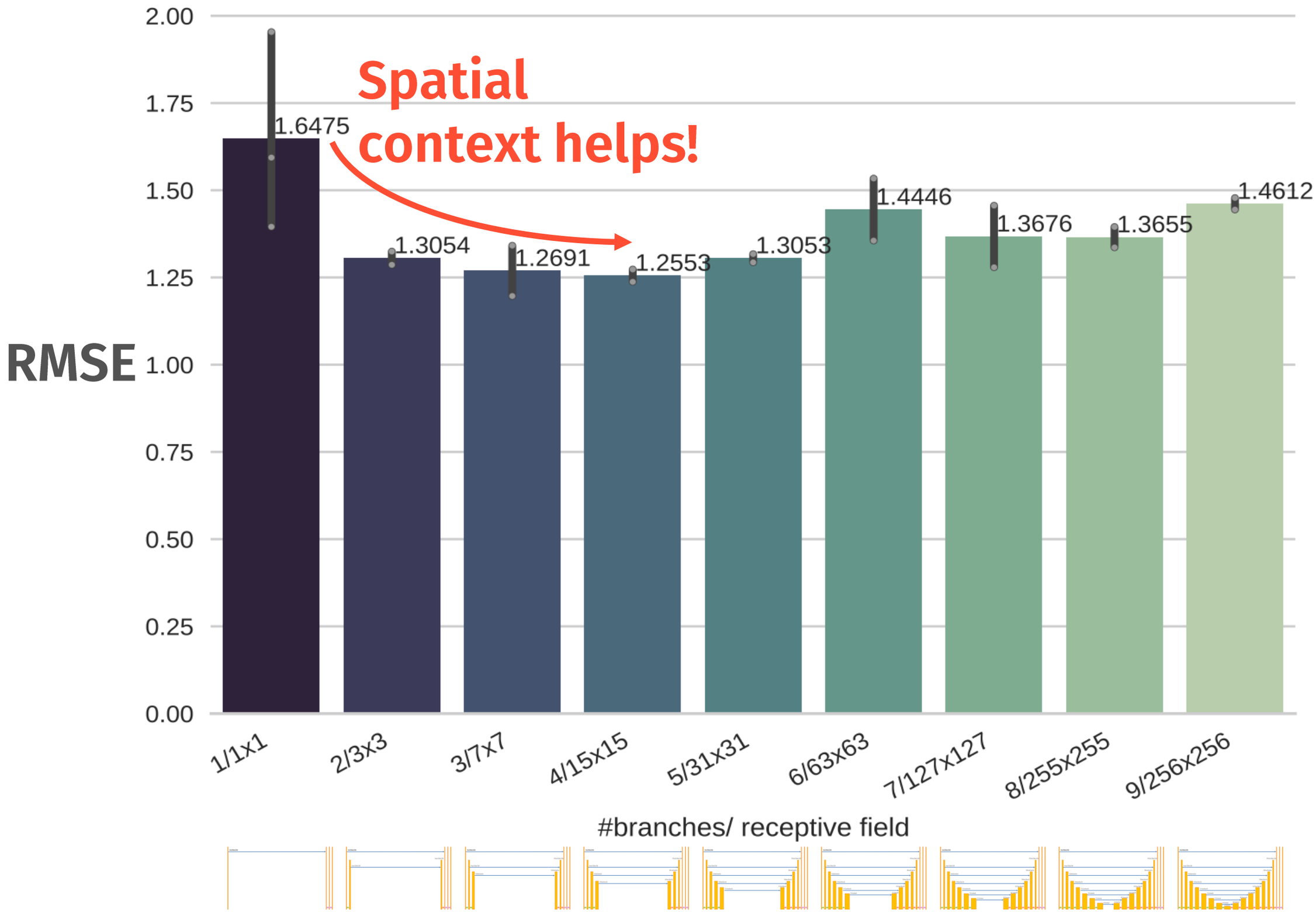
Results

Does spatial information really help?



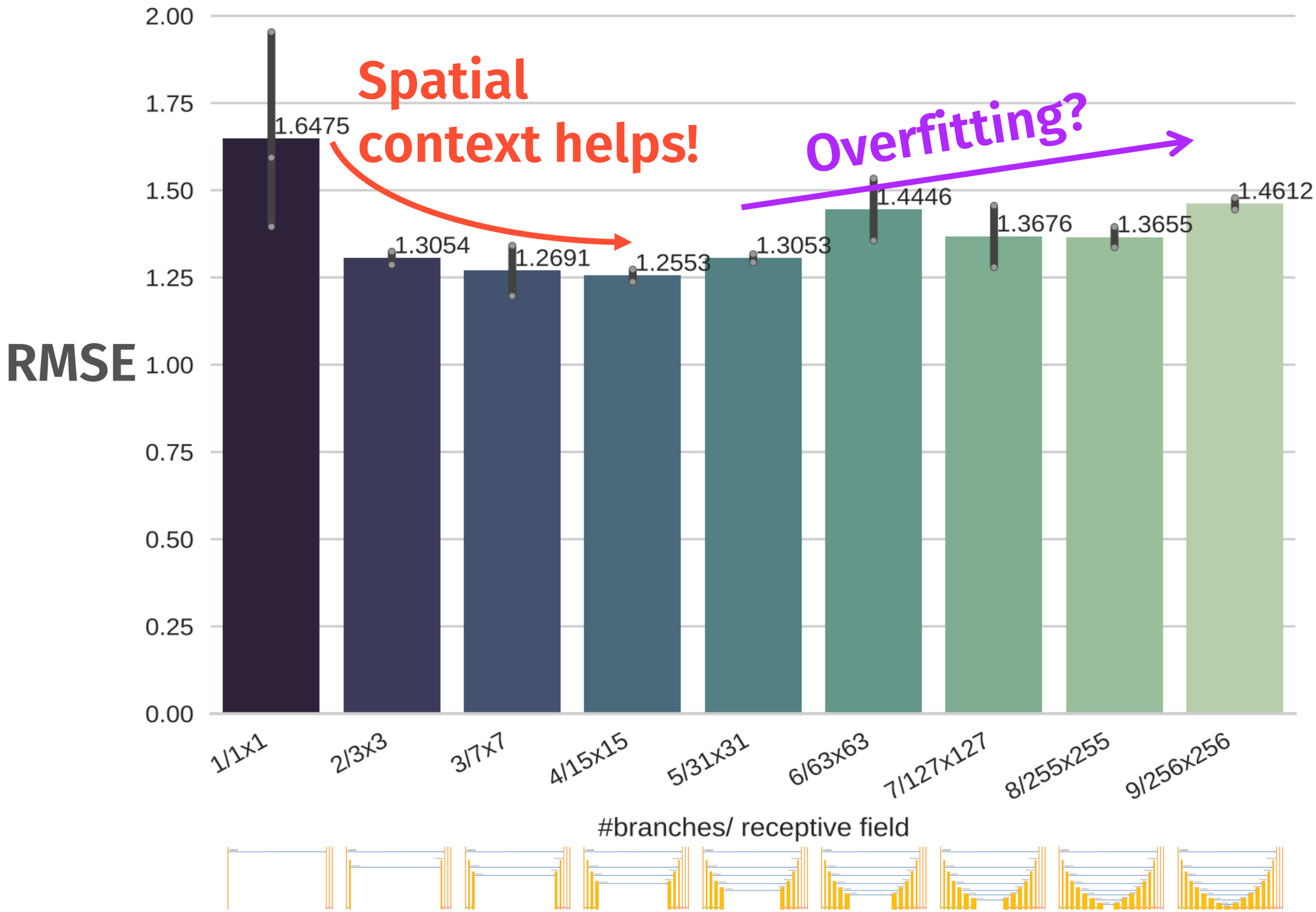
Results

Does spatial information really help?



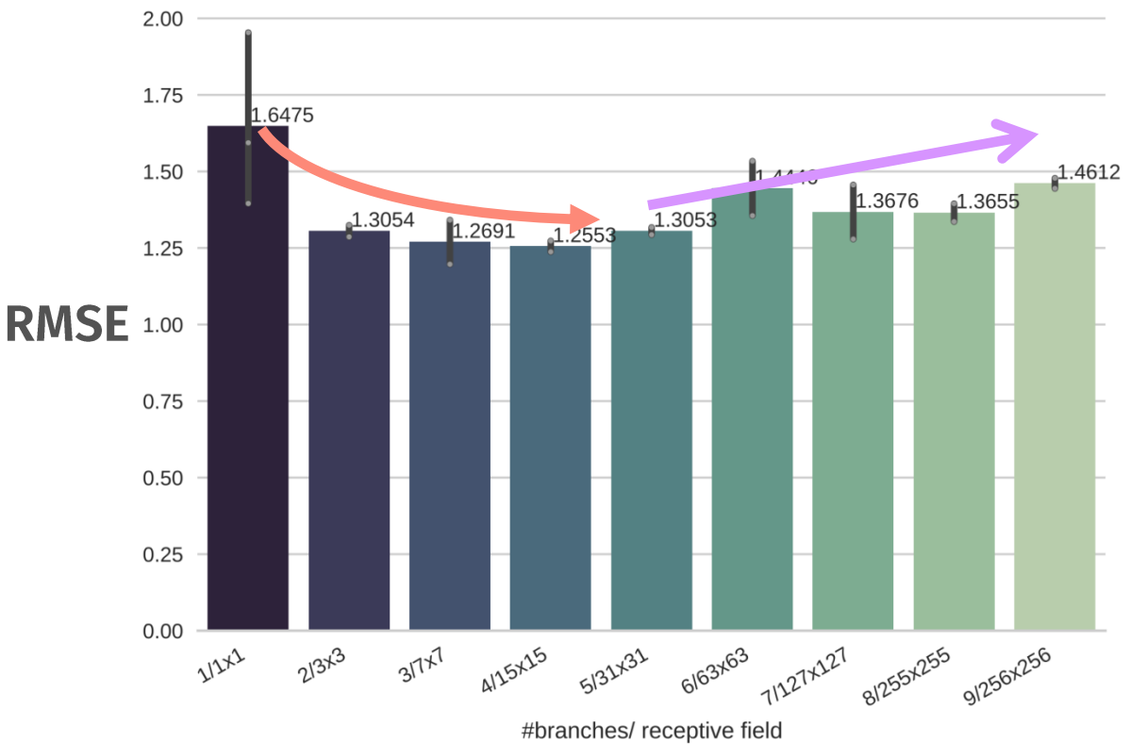
Results

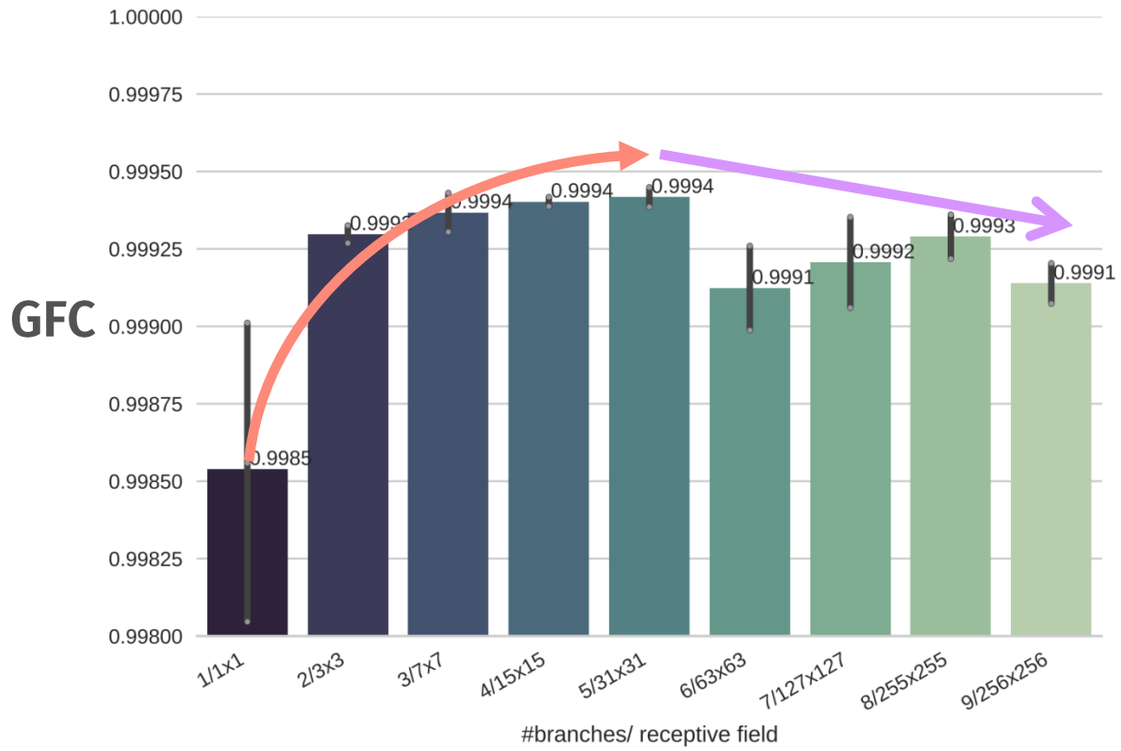
Does spatial information really help?

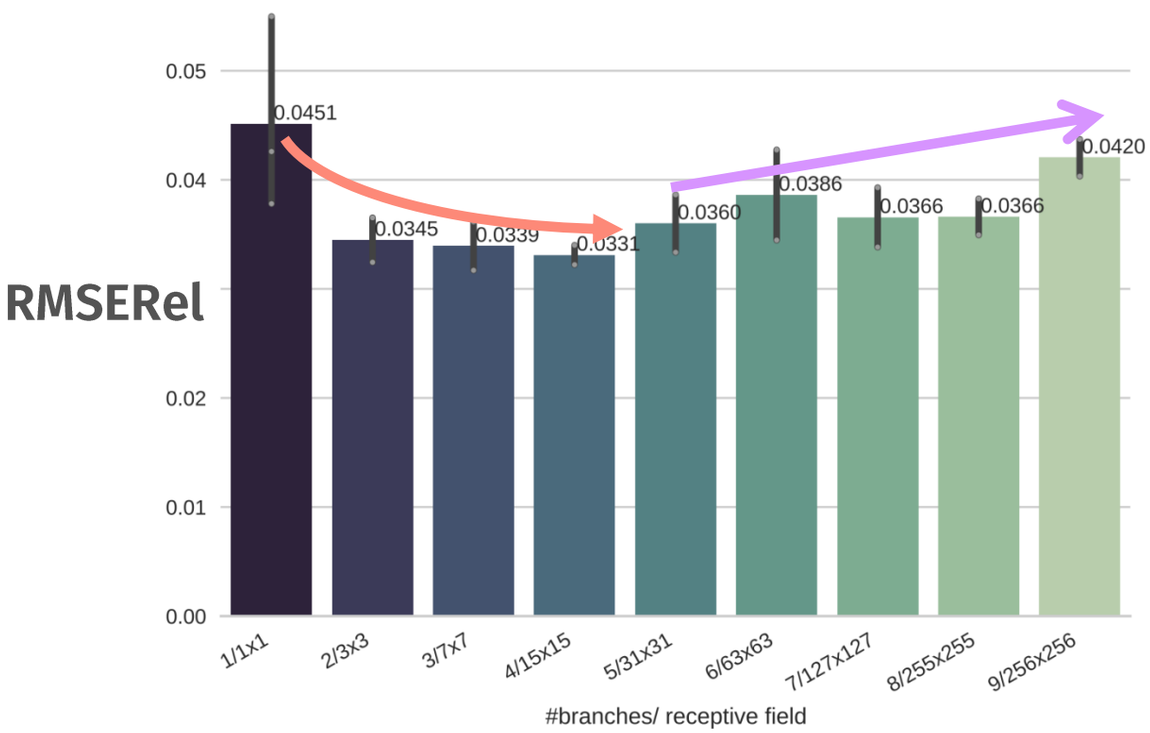


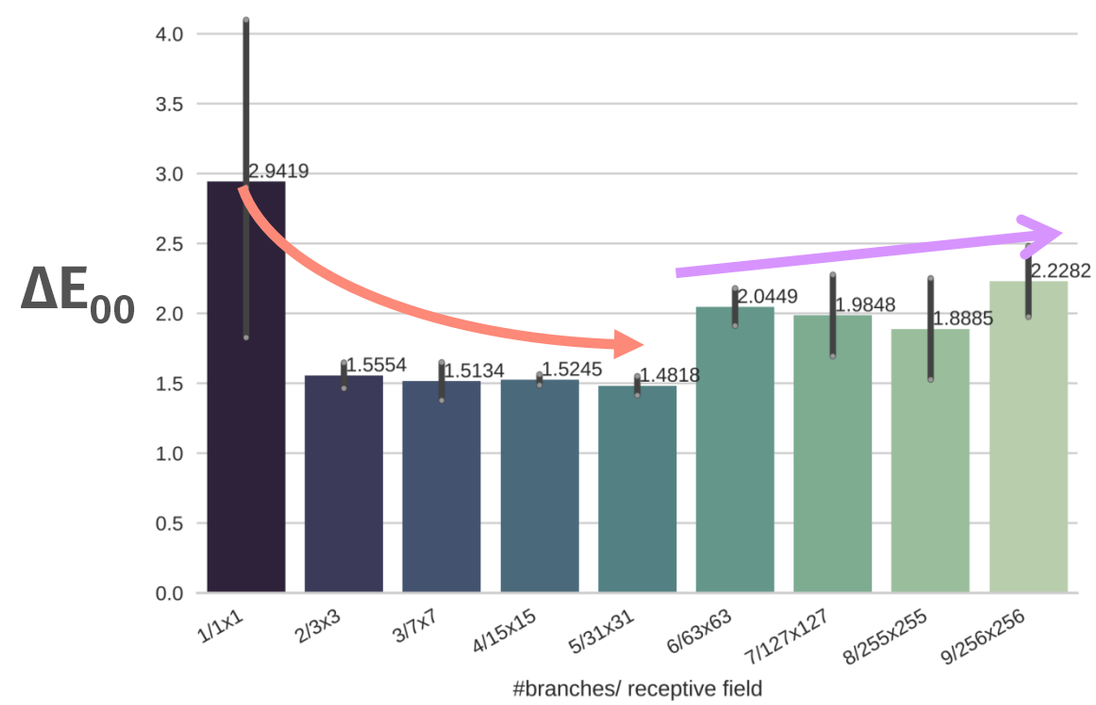
Results

Does spatial information really help?









Conclusions...

- GANs applied to spectral image reconstruction for the 1st time
- State of the art over ICVL dataset
- Spatial context information helps

...and future work

- Evaluate on metamers
- Extend to a variety of spectral sensitivity curves
- Simultaneous {illuminant, spectral reflectance} estimation

Adversarial Networks for Spatial Context-Aware Spectral Image Reconstruction from RGB

Aitor Alvarez-Gila, Joost van de Weijer, Estibaliz Garrote

Personal: **https://aitorshuffle.github.io**

Tecnalia Computer Vision group: **www.computervisionbytecnalia.com**

CVC Learning And Machine Perception team: **http://www.cvc.uab.es/LAMP/**

Venice, 2017/10/23