



Perceptual Invariance in Humans and Machines

Arturo Deza





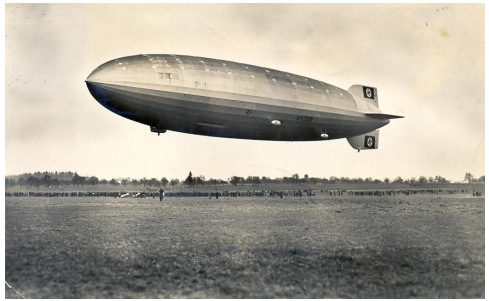
















Marr 1982 (Vision)

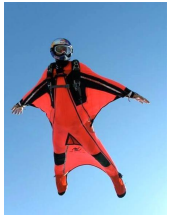


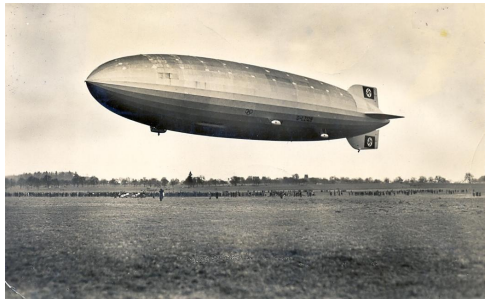




Feathers or Wings?













Courtesy from Bill Freeman's
SVRHM talk @ NeurIPS 2019

Marr 1982 (Vision)

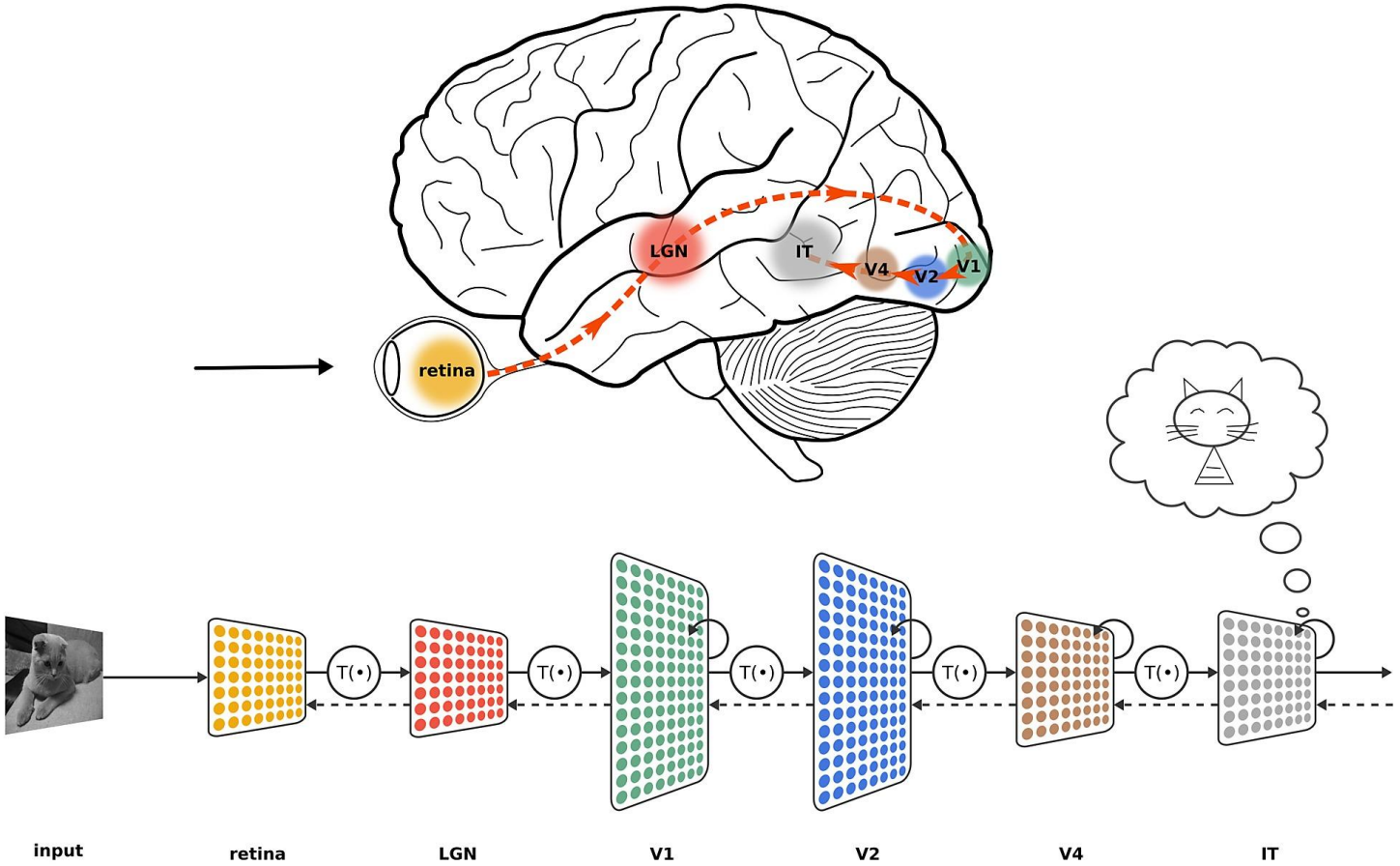
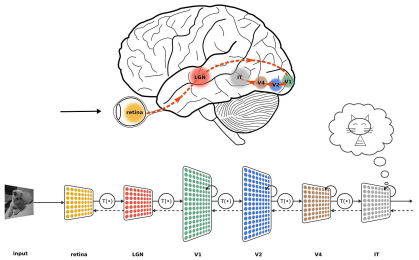
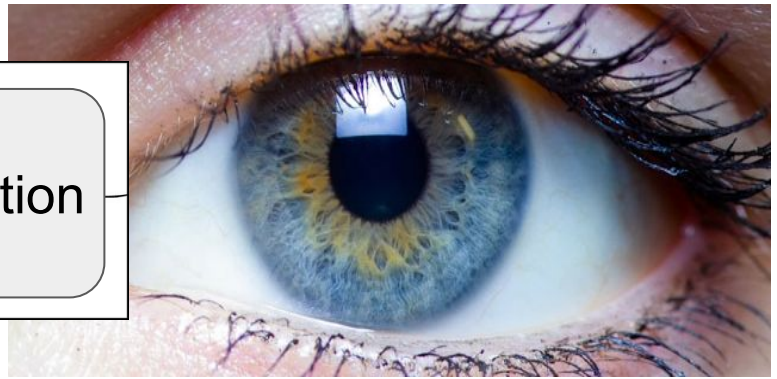


Figure: Jonas Kubilius

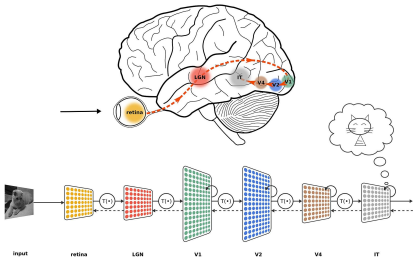




Convolution

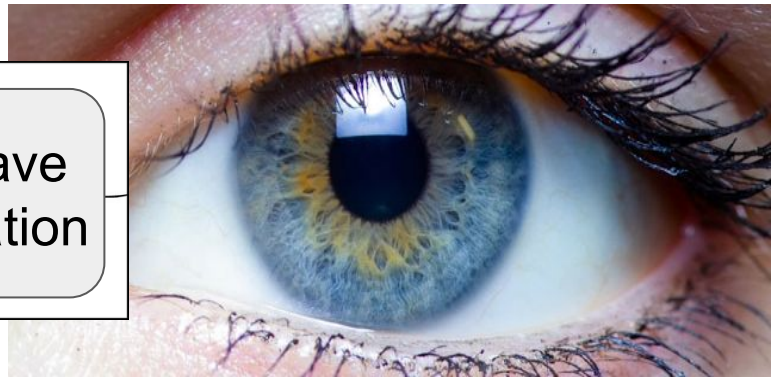






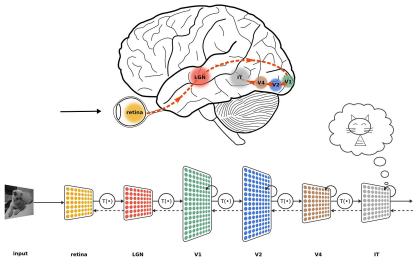


Half-Wave Rectification



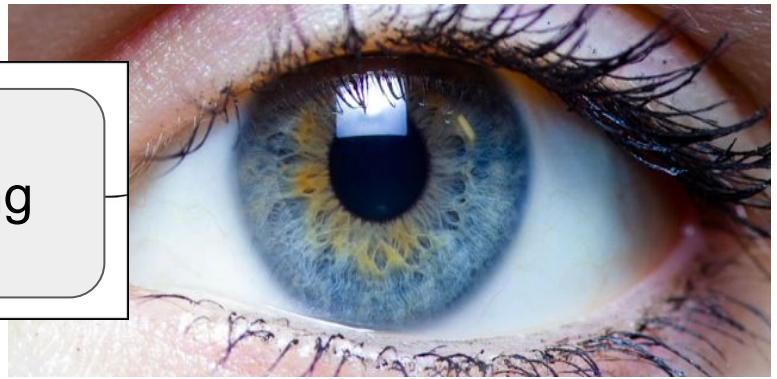






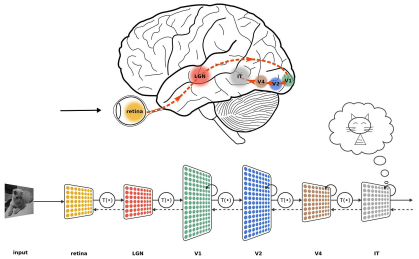


Pooling



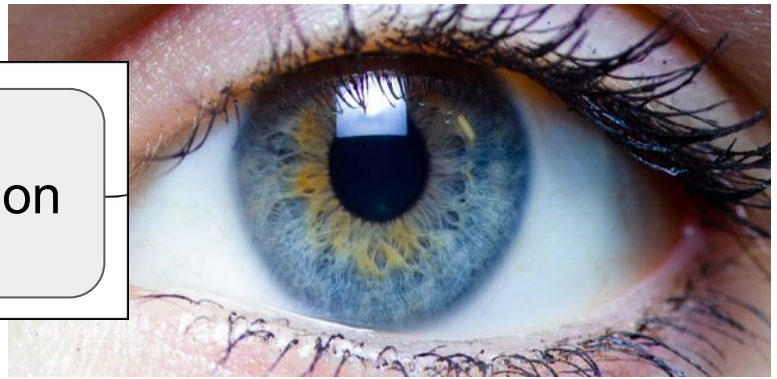








Foveation



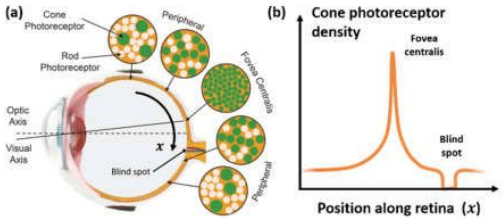
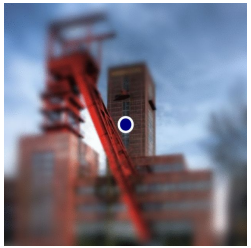
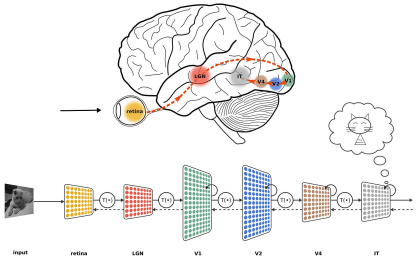


Figure from Chan & Jalaly, 2018.



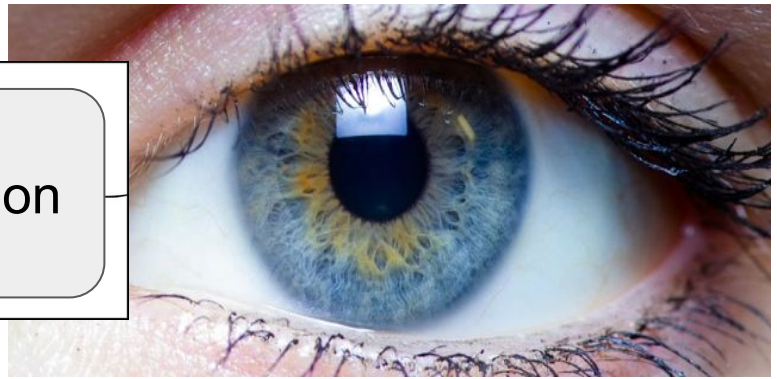








Foveation



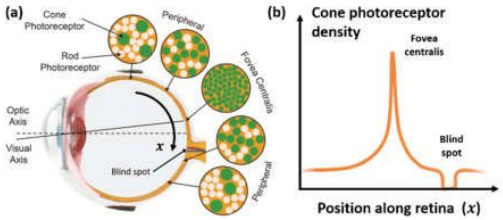
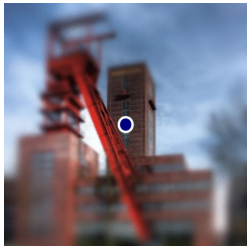


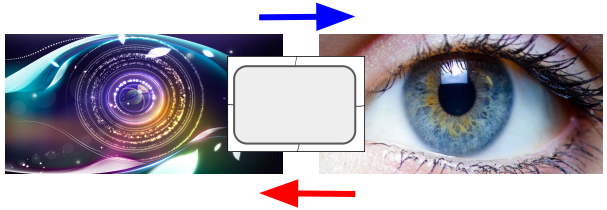
Figure from Chan & Jalaly, 2018.

Poll The Audience



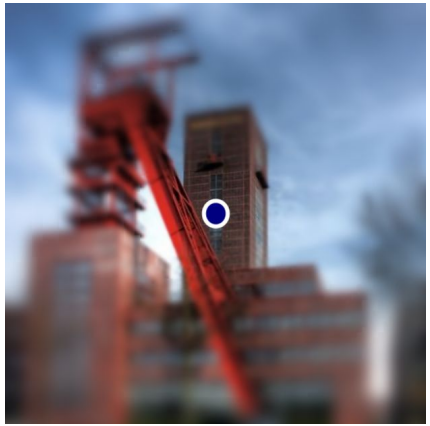


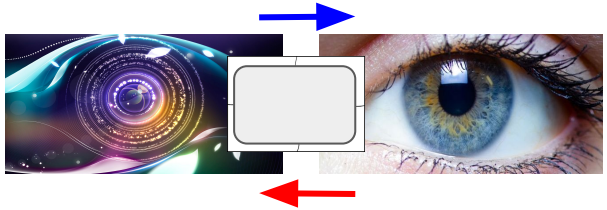




Efficiency vs Functionality

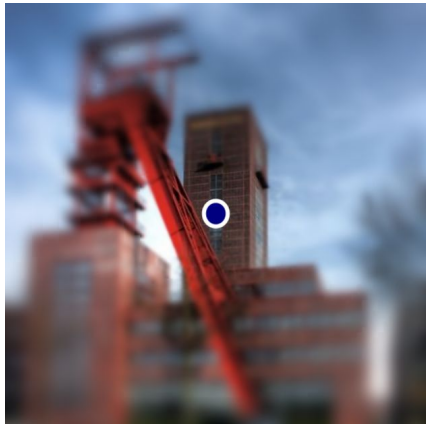


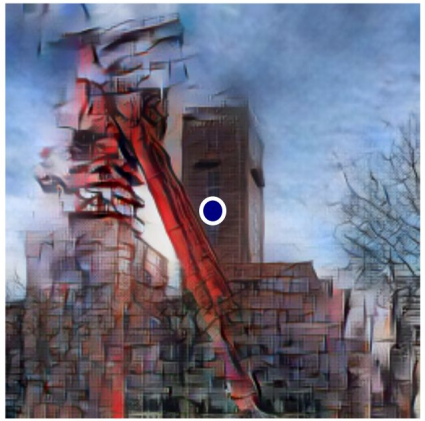




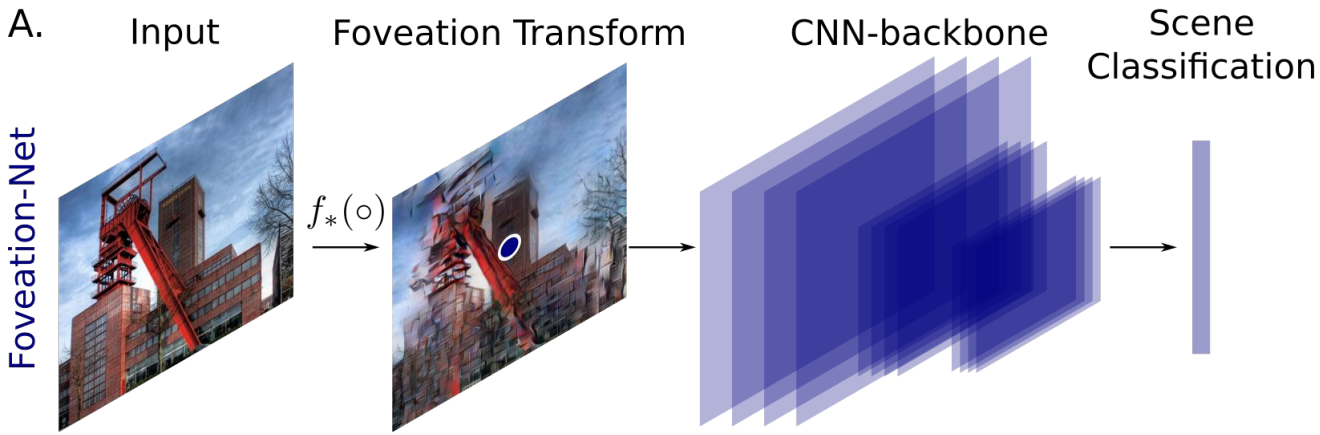
Efficiency vs Functionality



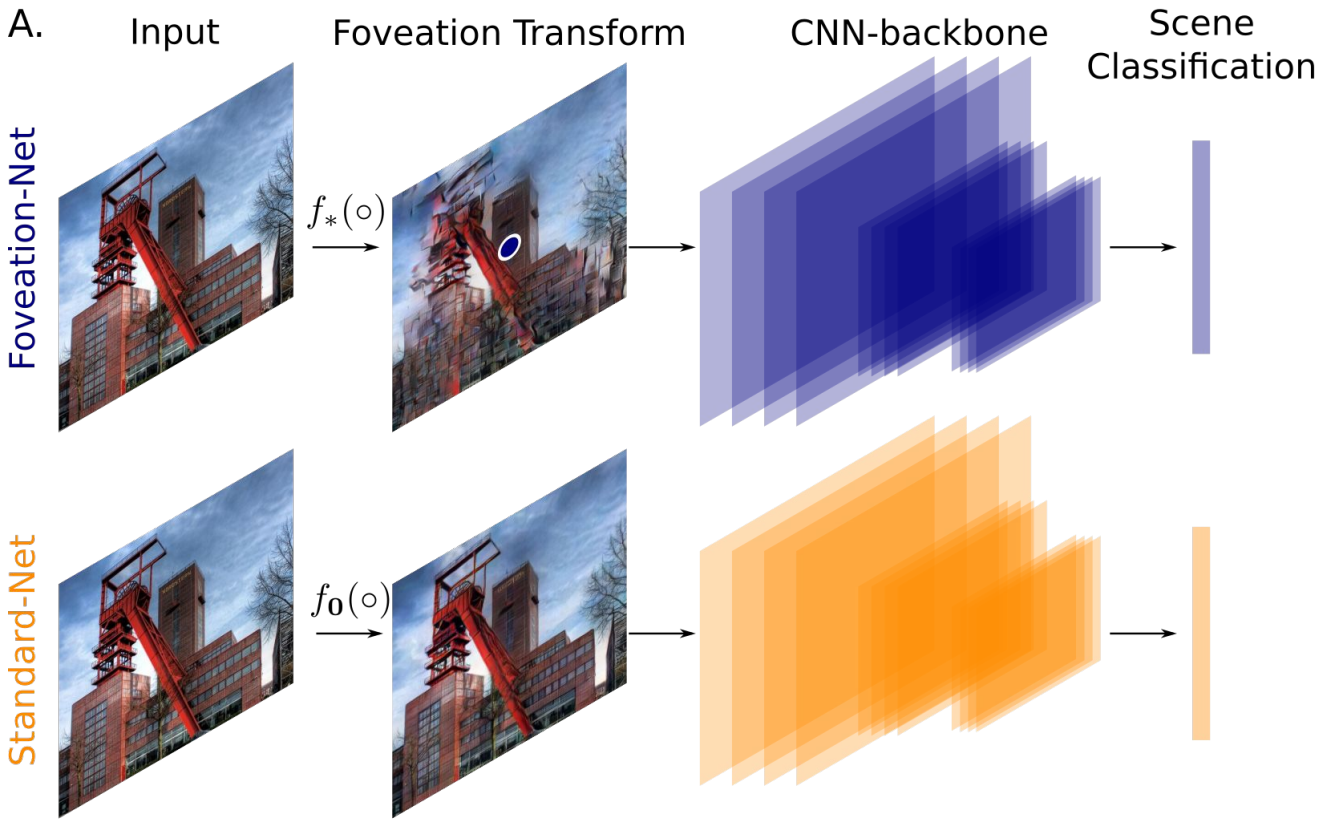




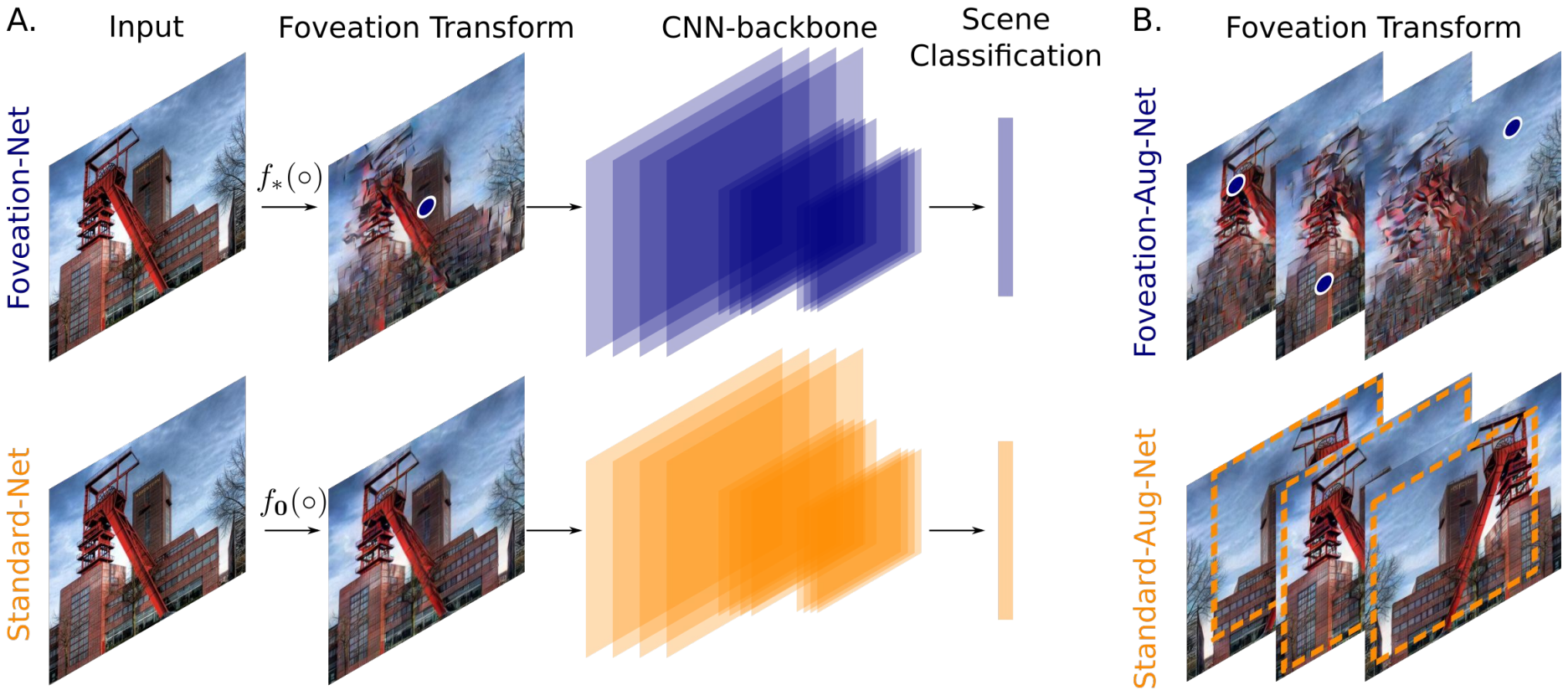
Foveated Perceptual Systems



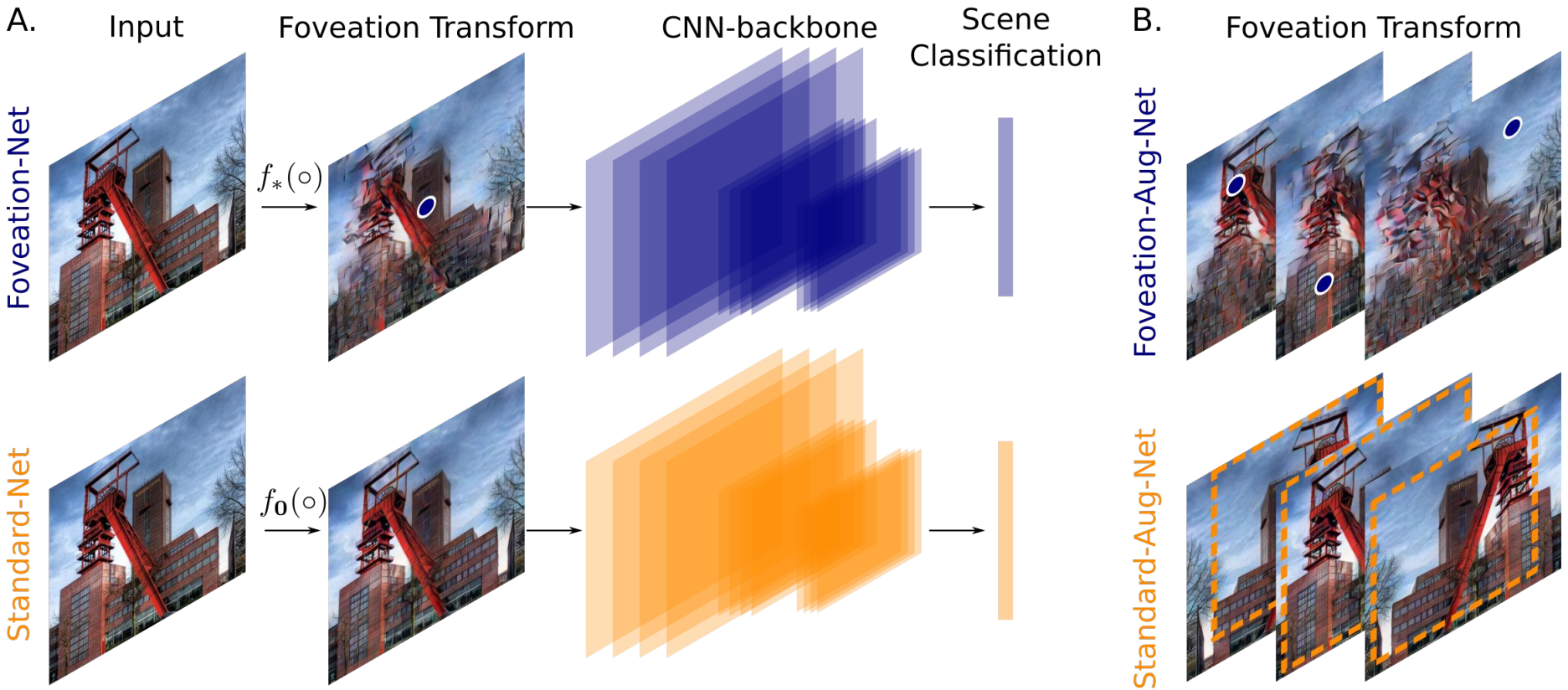
Foveated Perceptual Systems



Foveated Perceptual Systems

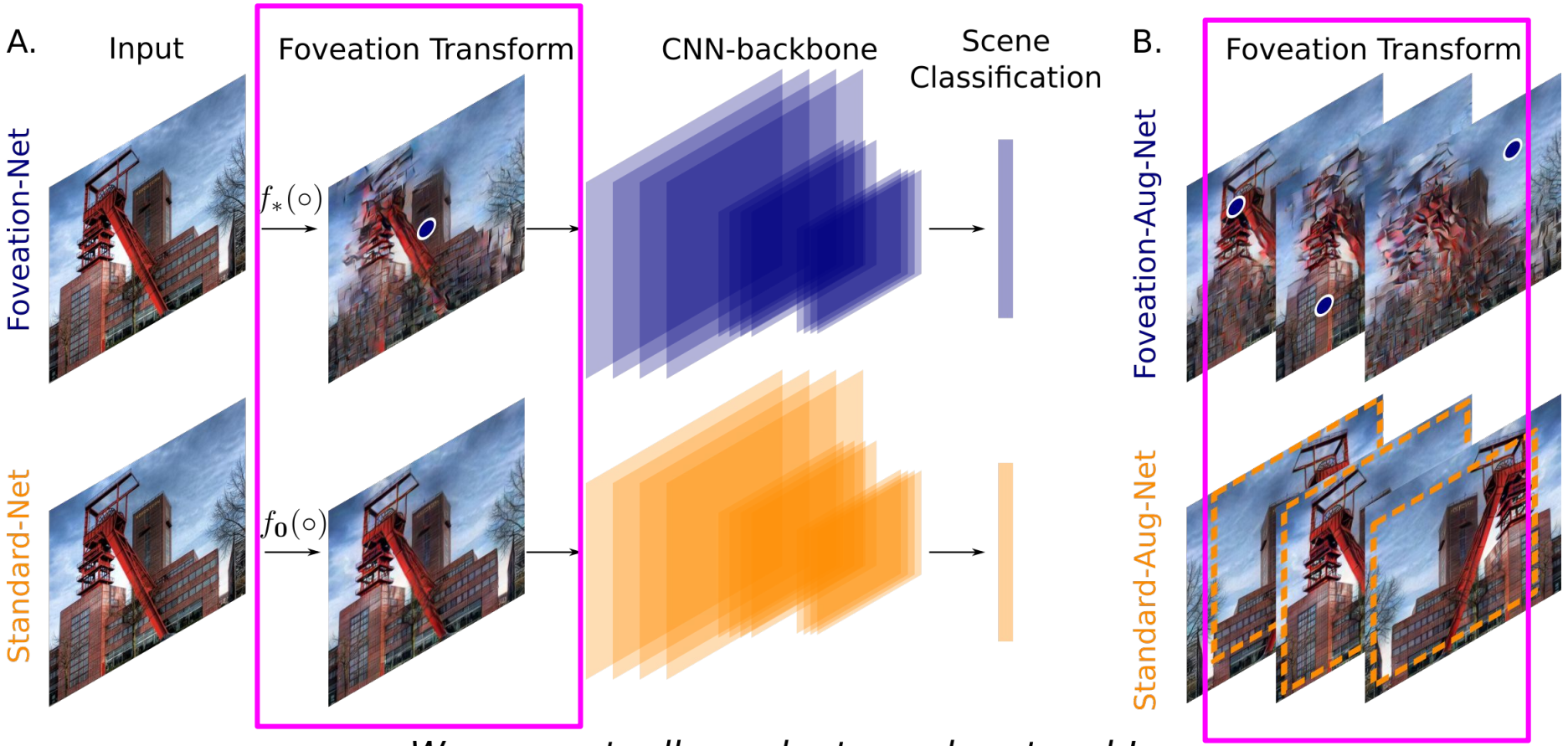


Foveated Perceptual Systems



We perceptually evaluate each network!

Foveated Perceptual Systems

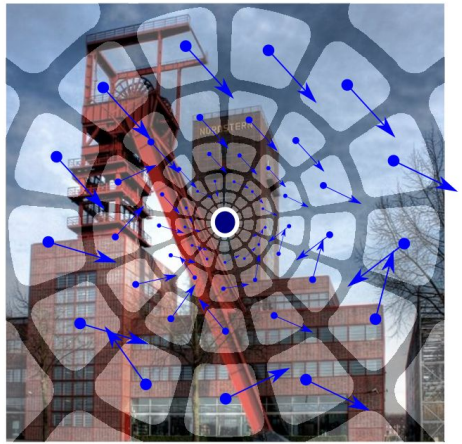


We perceptually evaluate each network!

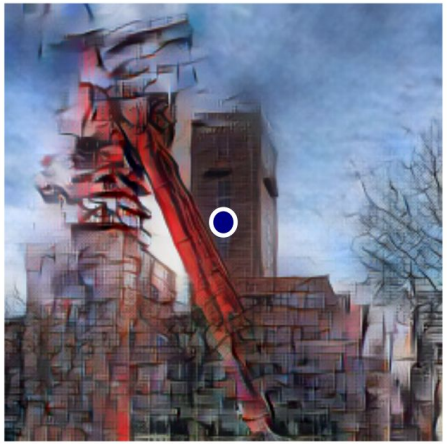
Image



Foveation Transform: $f_*(\circ)$



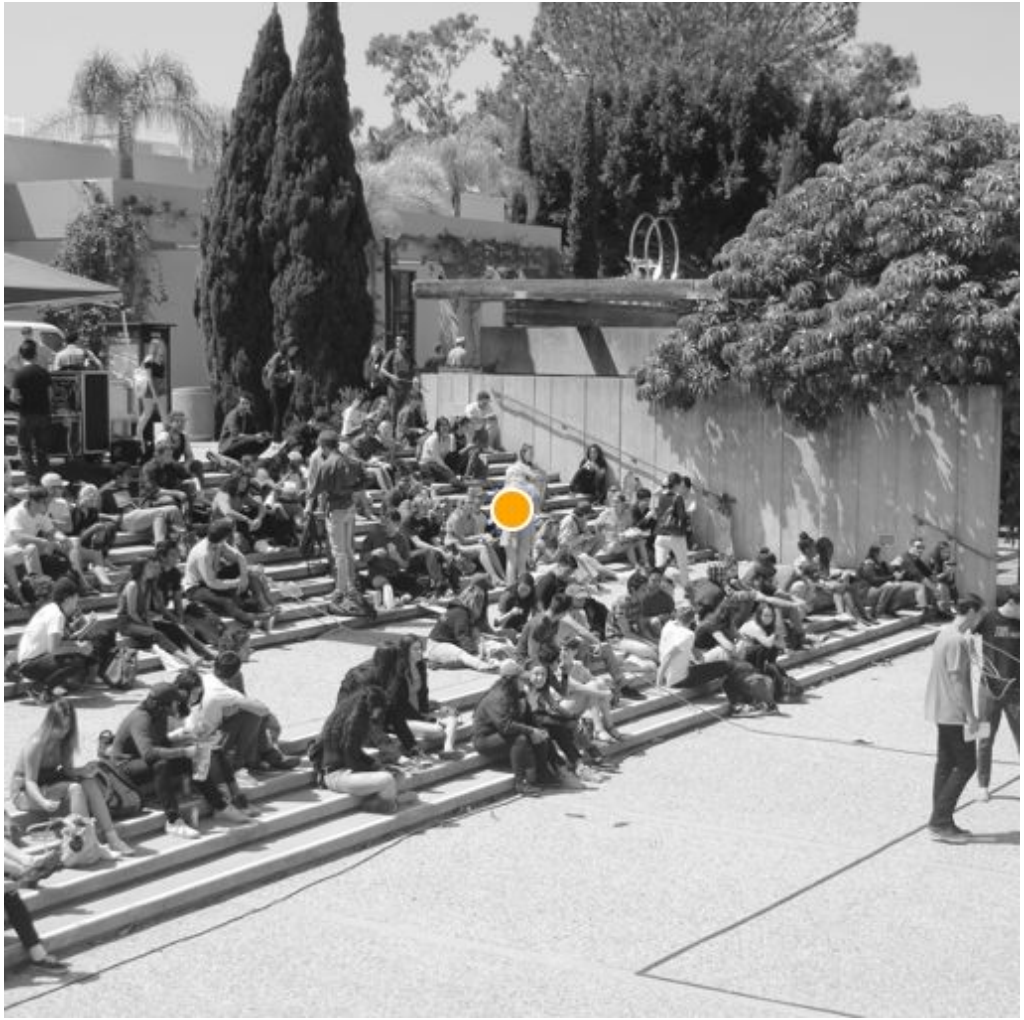
Foveated Image



Deza et al. ICLR 2019

I. Foveation Transform

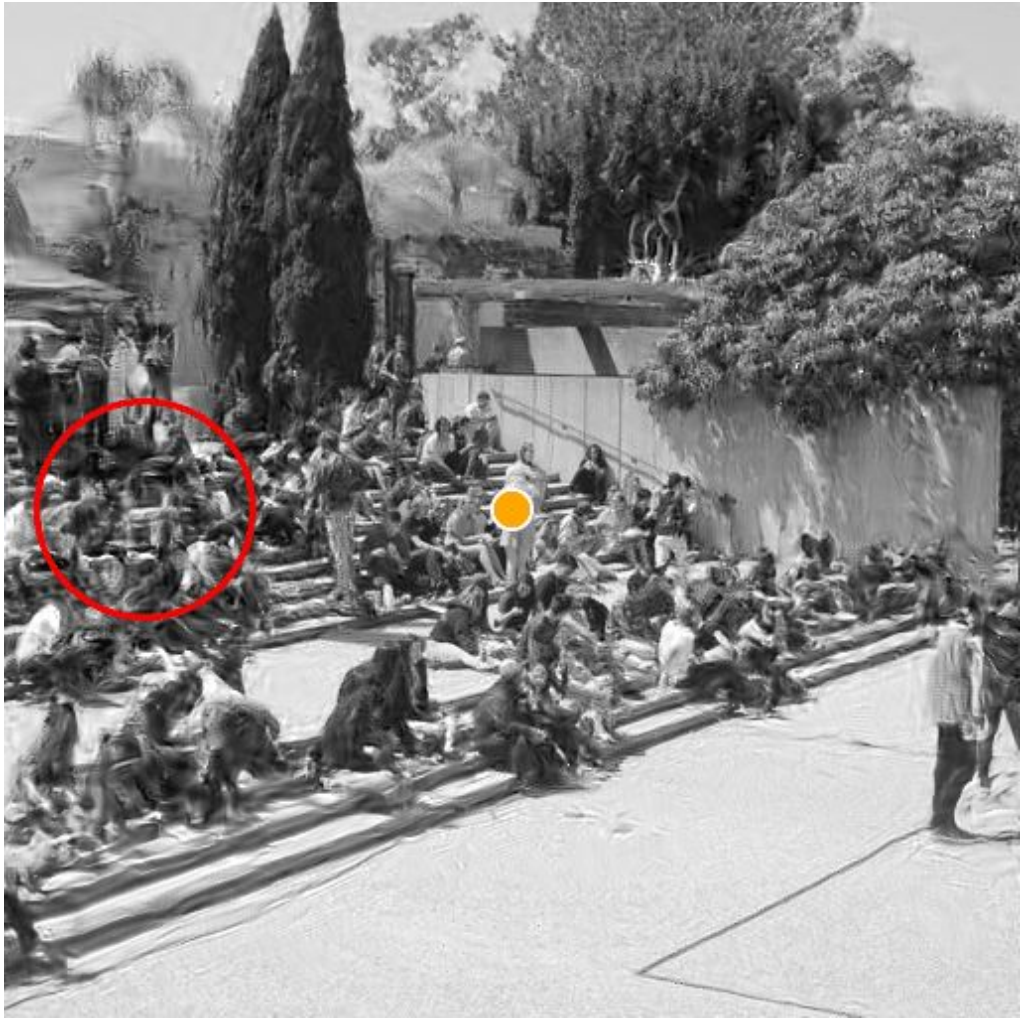
(and some clues on how this is related to Perceptual Invariance)





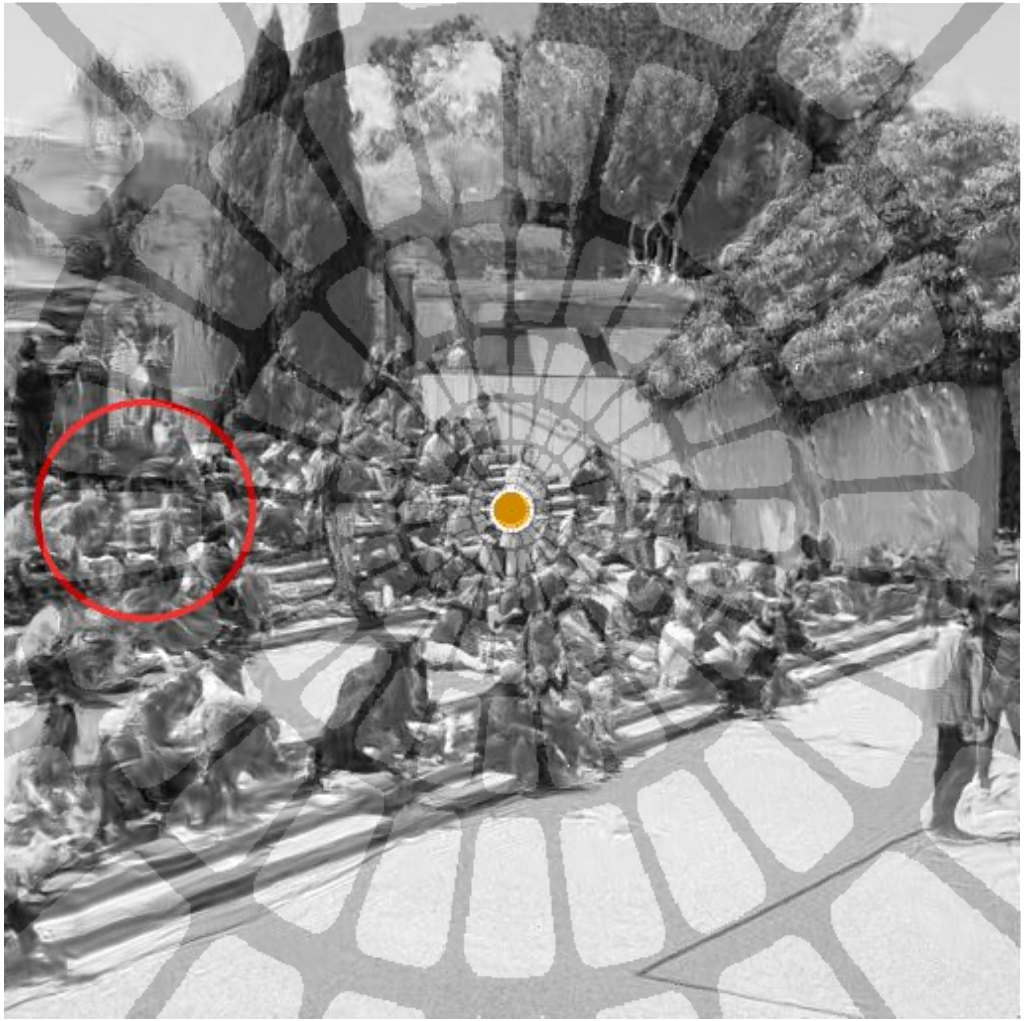
Rosenholtz (2011)

Freeman & Simoncelli (2011)



Rosenholtz (2011)

Freeman & Simoncelli (2011)



Rosenholtz (2011)

Freeman & Simoncelli (2011)



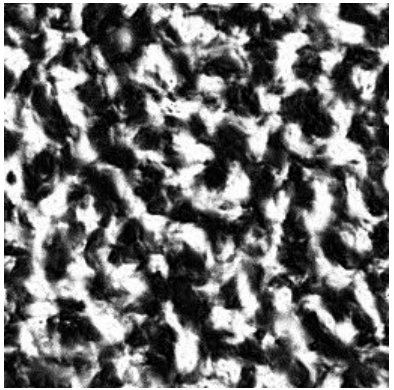




$$f(\cdot) \longrightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$



$$f(\cdot) \longrightarrow \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad \begin{bmatrix} \hat{x_1} \\ \hat{x_2} \\ \vdots \\ \hat{x_N} \end{bmatrix} \longleftarrow f(\cdot)$$



Balas, Nakano & Rosenholtz (2009)

Portilla & Simoncelli (2000)

Texture Synthesis



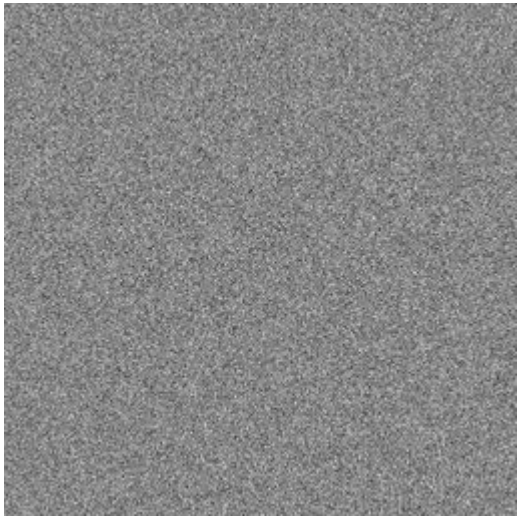
$$\xrightarrow{f(\cdot)} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} \hat{x_1} \\ \hat{x_2} \\ \vdots \\ \hat{x_N} \end{bmatrix} \xleftarrow{f(\cdot)}$$



Balas, Nakano & Rosenholtz (2009)

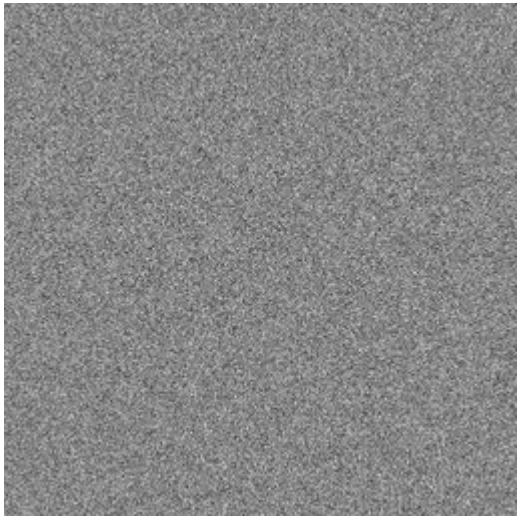
Portilla & Simoncelli (2000)

4-6 hours



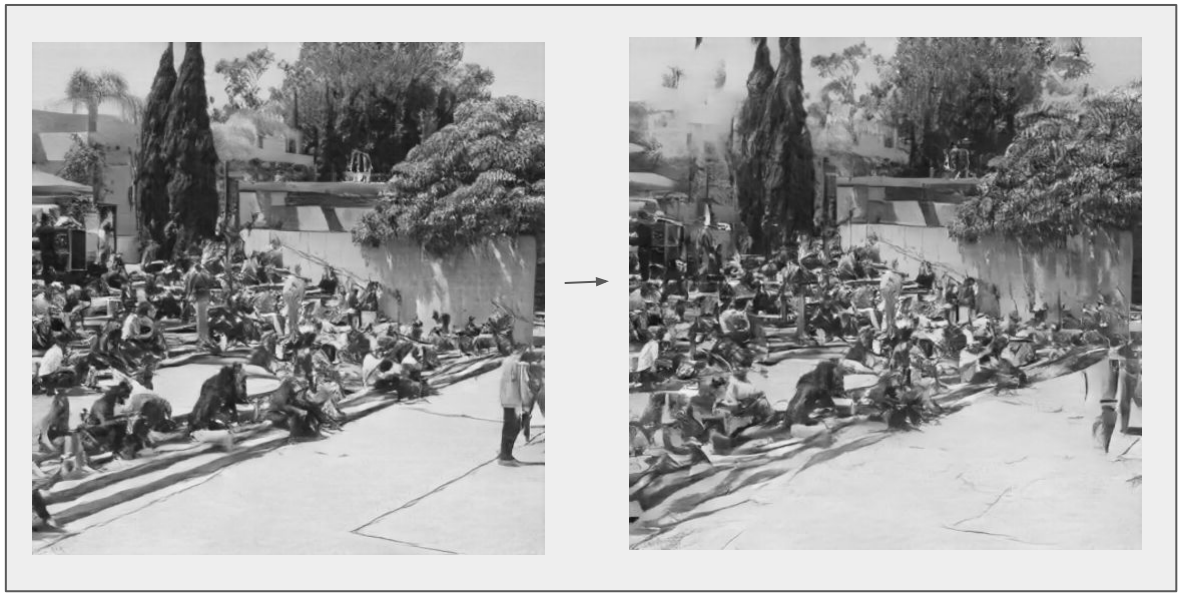
Freeman & Simoncelli (2011)

4-6 hours



Freeman & Simoncelli (2011)

1-2 seconds



Deza, Jonnalagadda & Eckstein. ICLR 2019.

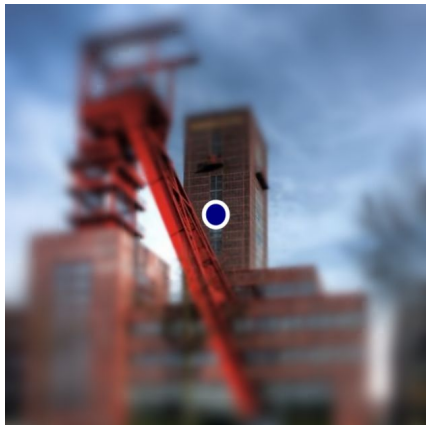
Metamer:

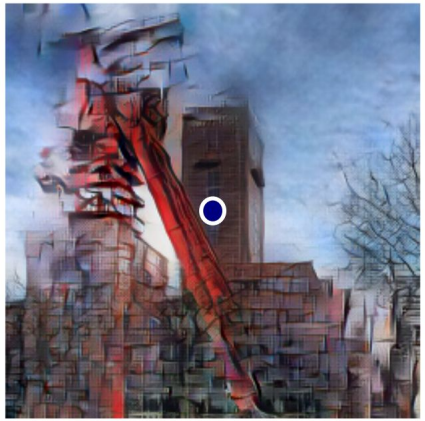
Two physically different stimuli that are perceptually indistinguishable to each other

Metamer:

Two physically different stimuli that are perceptually indistinguishable to each other





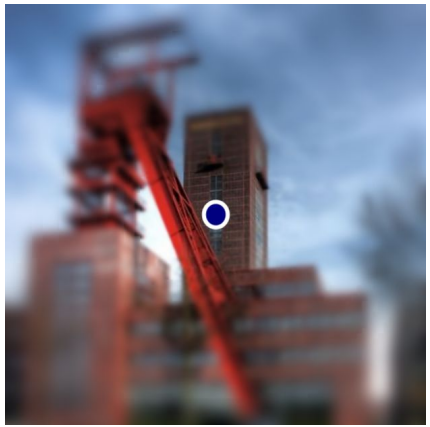


Metamer:

Two physically different stimuli that are perceptually indistinguishable to each other

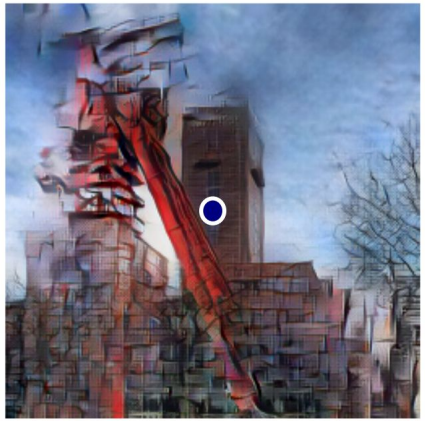


Retina + V1



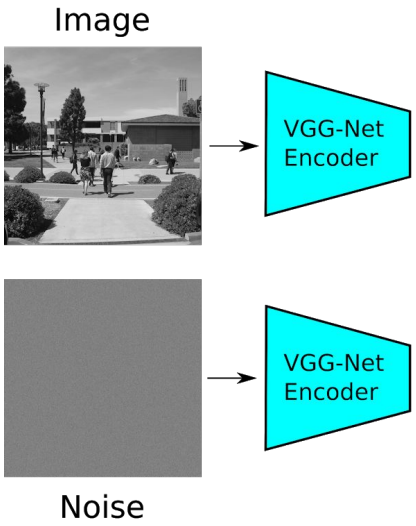
Geisler & Perry, 1998
Wang & Cottrell, 2017

V2



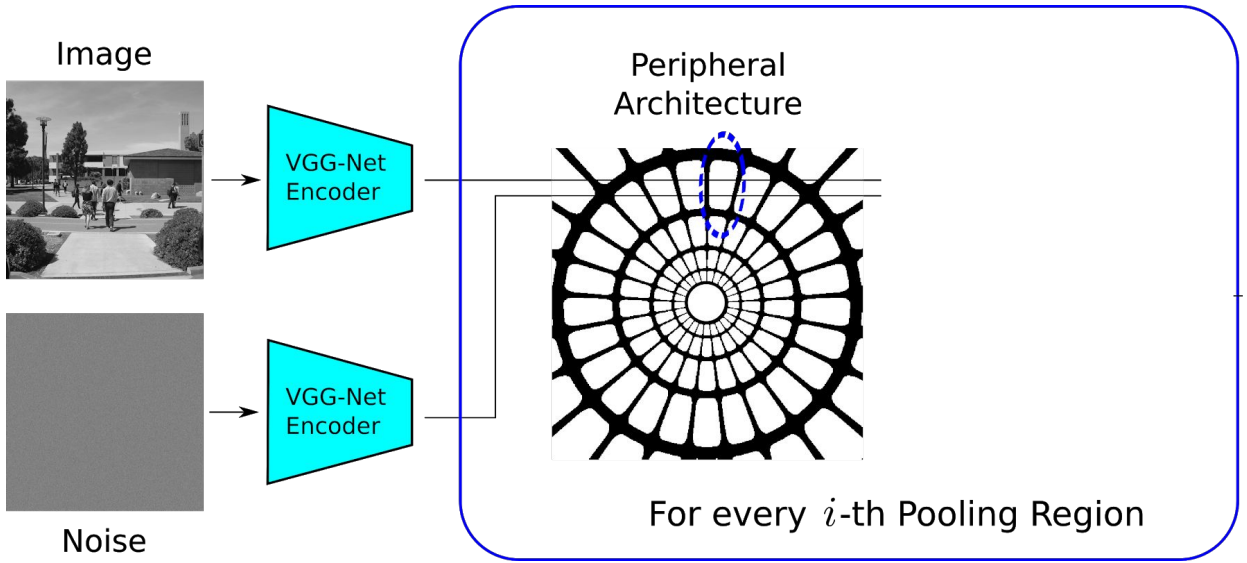
Freeman & Simoncelli, 2011
Deza & Konkle, 2020

A new metamer rendering pipeline



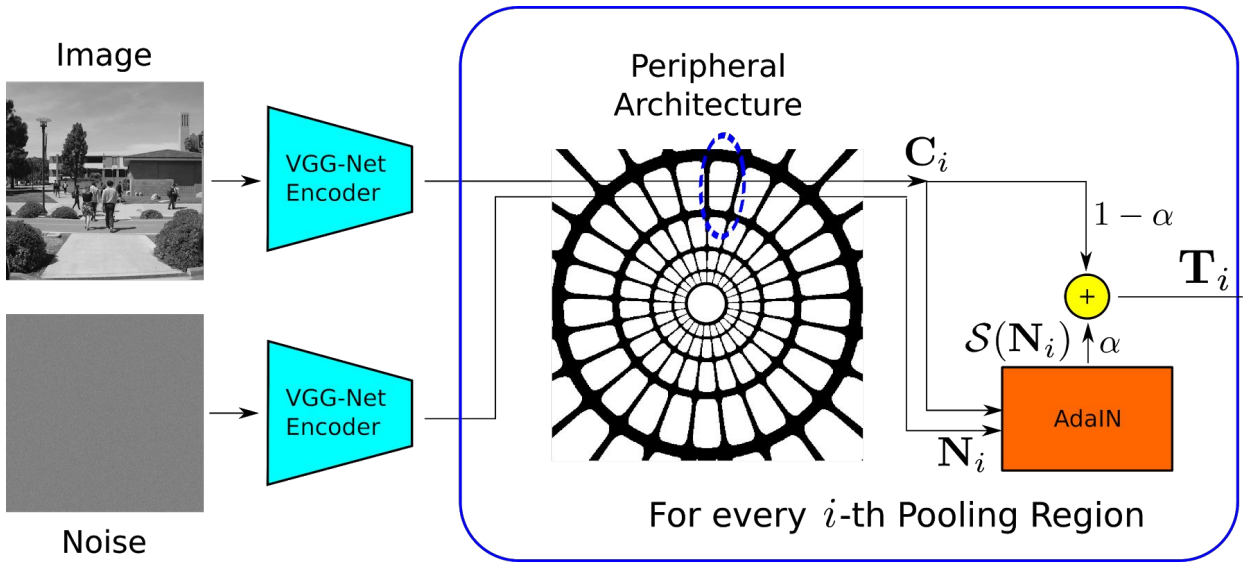
Deza, Jonnalagadda & Eckstein. ICLR 2019.

A new metamer rendering pipeline



Deza, Jonnalagadda & Eckstein. ICLR 2019.

A new metamer rendering pipeline



Deza, Jonnalagadda & Eckstein. ICLR 2019.

Case Study:

What is the maximum amount of texture-driven distortion we can put in this receptive field?

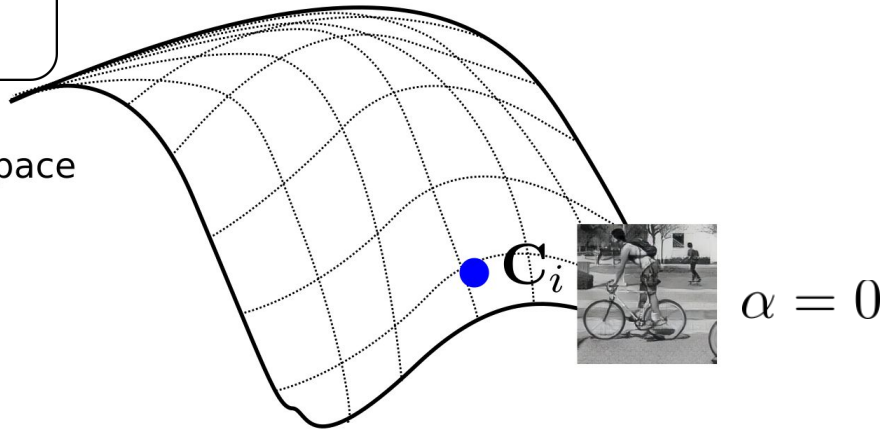


Case Study:

What is the maximum amount of texture-driven distortion we can put in this receptive field?



Encoded Space

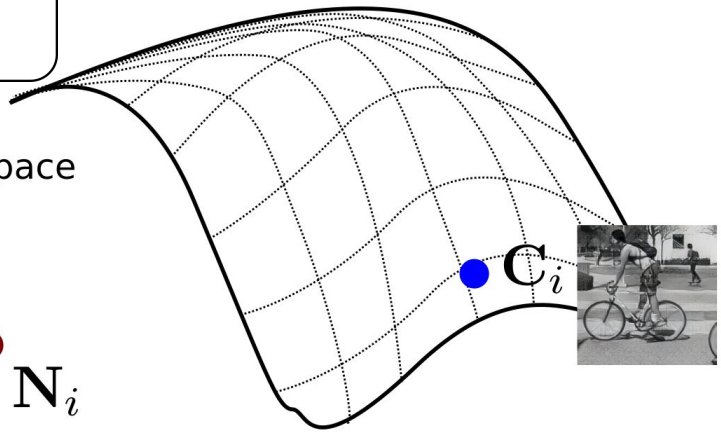


Case Study:

What is the maximum amount of texture-driven distortion we can put in this receptive field?



Encoded Space



Case Study:

What is the maximum amount of texture-driven distortion we can put in this receptive field?



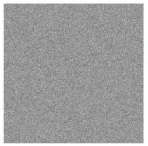


$\alpha = 1$

$\mathcal{S}(\mathbf{N}_i)$

Encoded Space

$\mathcal{S}(\circ)$



$\mathbf{N}_i$

$\mathbf{C}_i$



$\alpha = 0$

Case Study:

What is the maximum amount of texture-driven distortion we can put in this receptive field?





$\alpha = 1$

$\mathcal{S}(\mathbf{N}_i)$



$\alpha = 0.4$

Encoded Space

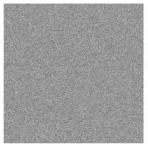
$\mathbf{T}_i$

$\mathcal{S}(\circ)$

$\mathbf{C}_i$



$\alpha = 0$



$\mathbf{N}_i$

Case Study:

What is the maximum amount of texture-driven distortion we can put in this receptive field?





$\alpha = 1$

$\mathcal{S}(\mathbf{N}_i)$



$\alpha = 0.4$

Encoded Space

$\mathcal{S}^{\perp}(\mathbf{N}_i)$

$\mathbf{T}_i$

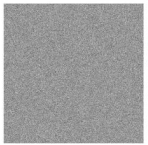
$\mathcal{S}(\circ)$

$\mathcal{S}^{\parallel}(\mathbf{N}_i)$

$\mathbf{C}_i$



$\alpha = 0$

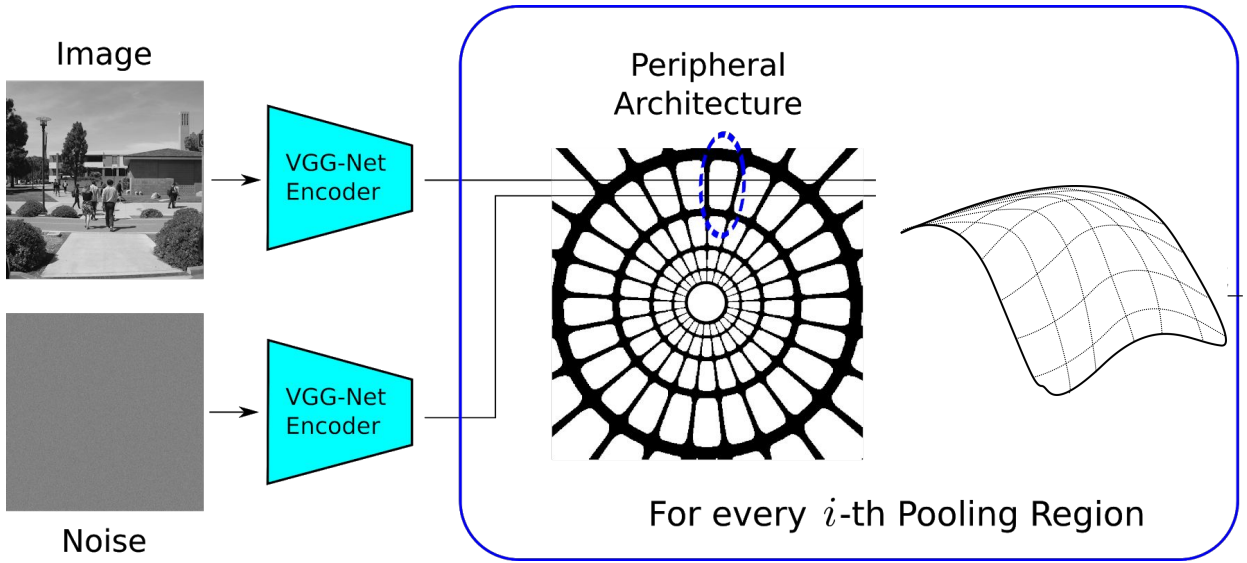


$\mathbf{N}_i$

P

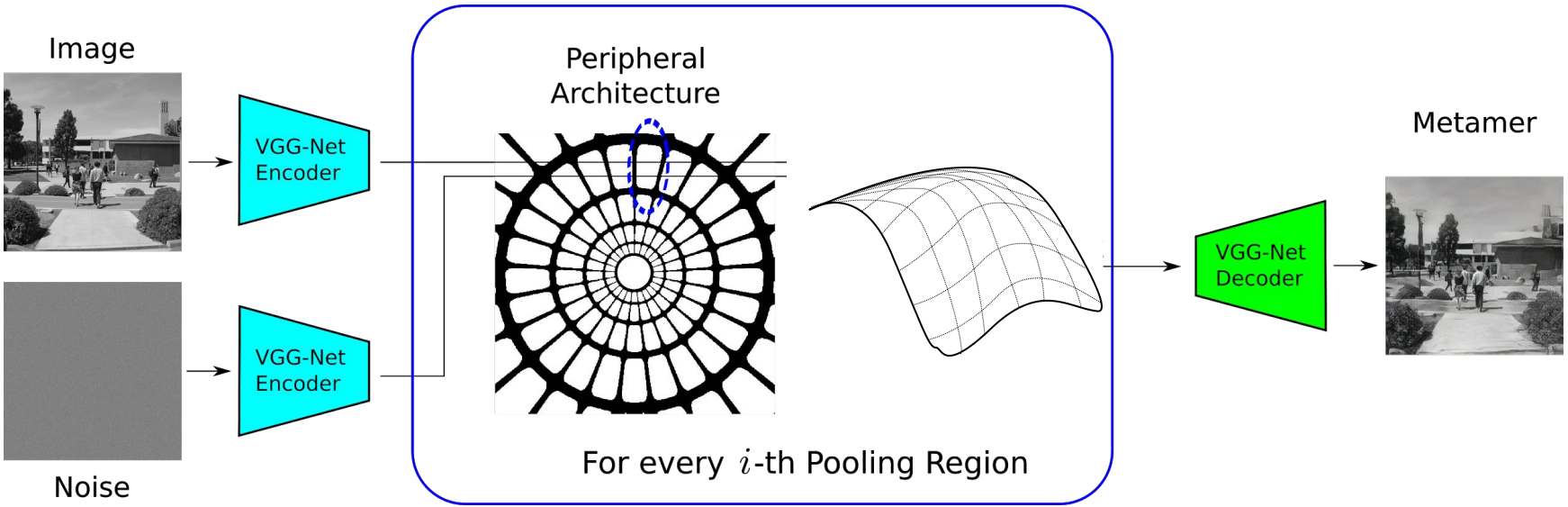
Perceptual Space

A new metamer rendering pipeline



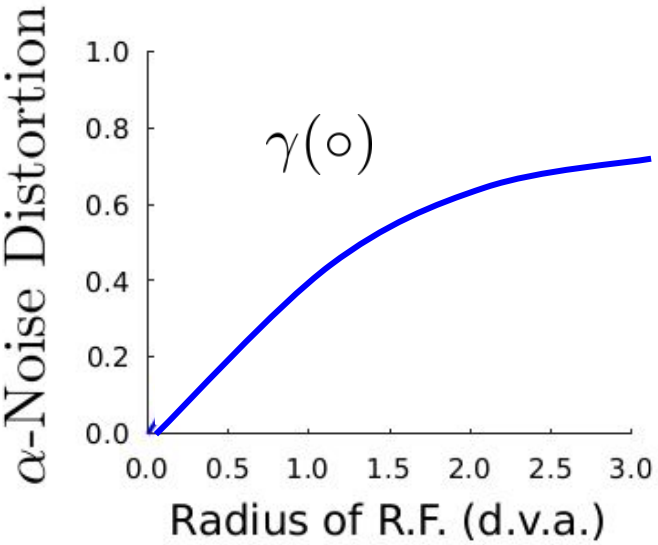
Deza, Jonnalagadda & Eckstein. ICLR 2019.

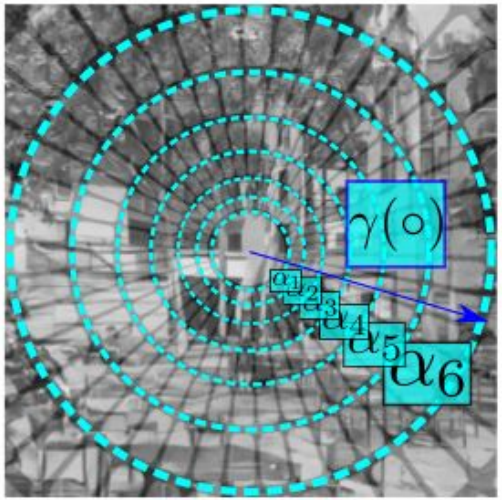
A new metamer rendering pipeline



Deza, Jonnalagadda & Eckstein. ICLR 2019.

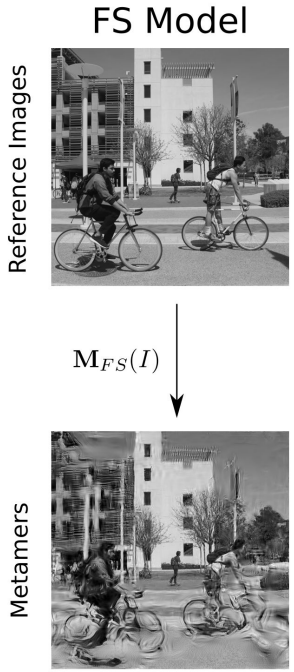
How much can we distort in each pooling region?





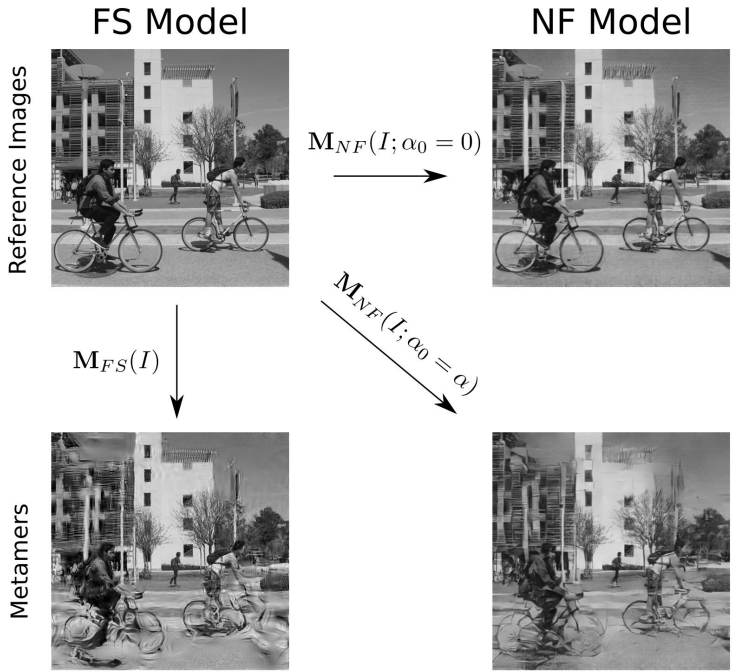
Deza, Jonnalagadda & Eckstein. ICLR 2019.

Perceptual Optimization Procedure



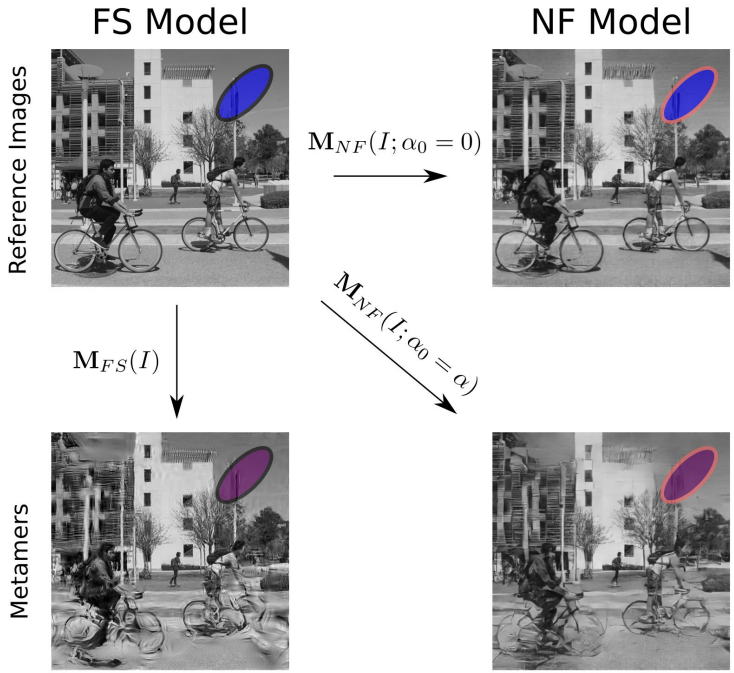
Deza, Jonnalagadda & Eckstein. ICLR 2019.

Perceptual Optimization Procedure



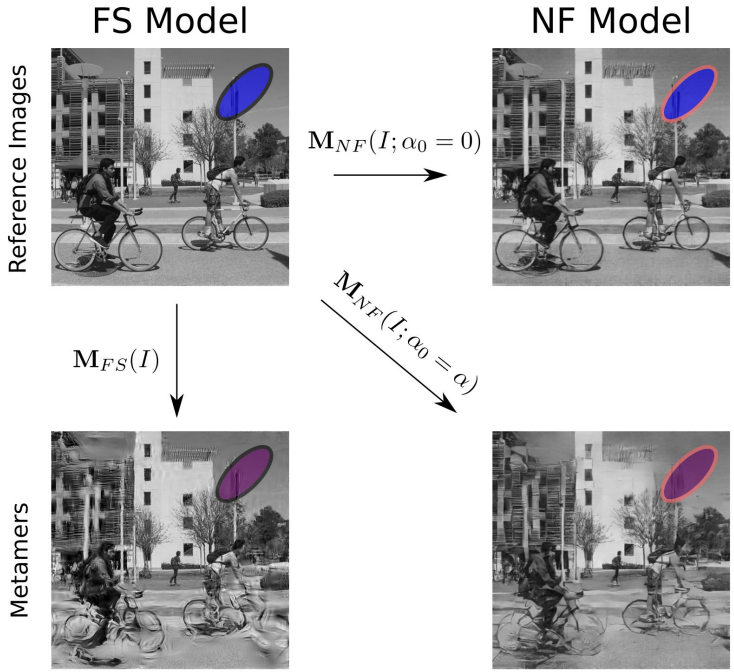
Deza, Jonnalagadda & Eckstein. ICLR 2019.

Perceptual Optimization Procedure



Deza, Jonnalagadda & Eckstein. ICLR 2019.

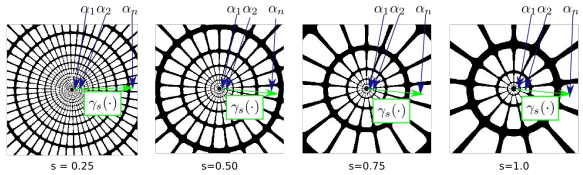
Perceptual Optimization Procedure



Minimize : $(\text{SSIM}(\,,\,) - \text{SSIM}(\,,\,))^2$

for every pooling region in the visual field

Estimation of maximum distortion in the visual field



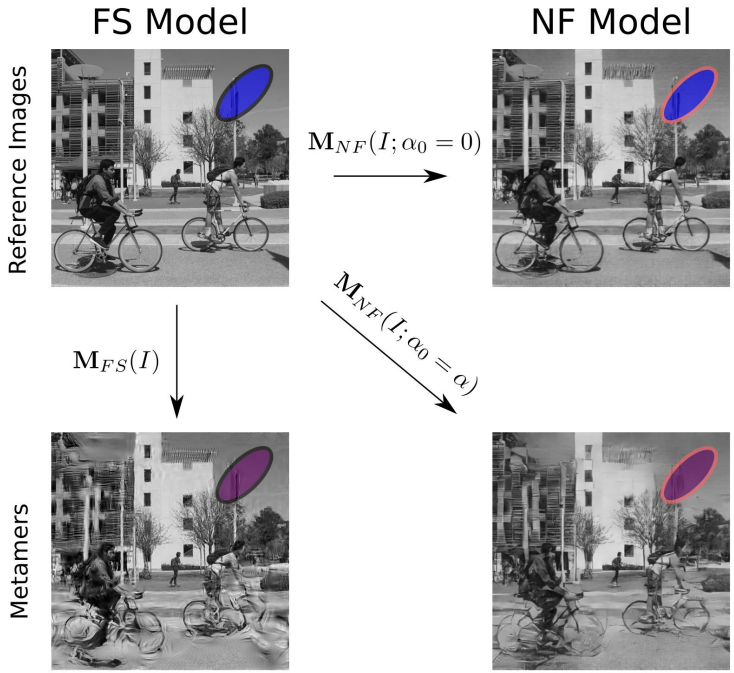
$$s_0, \bar{\alpha_0} = \text{argmax}_{s,\bar{\alpha}} \mathbb{E}[d'(s,\bar{\alpha}|\theta_{obs})]$$
$$\text{s.t. } 0 < d'(s,\bar{\alpha}|\theta_{obs}) < \epsilon$$

$$\downarrow \alpha = \gamma(\circ; s)$$

$$s_0 = \text{argmax}_s\ \mathbb{E}[d'(s,\gamma(\circ;s)|\theta_{obs})]$$
$$\text{s.t. } 0 < d'(s,\gamma(\circ;s)|\theta_{obs}) < \epsilon$$

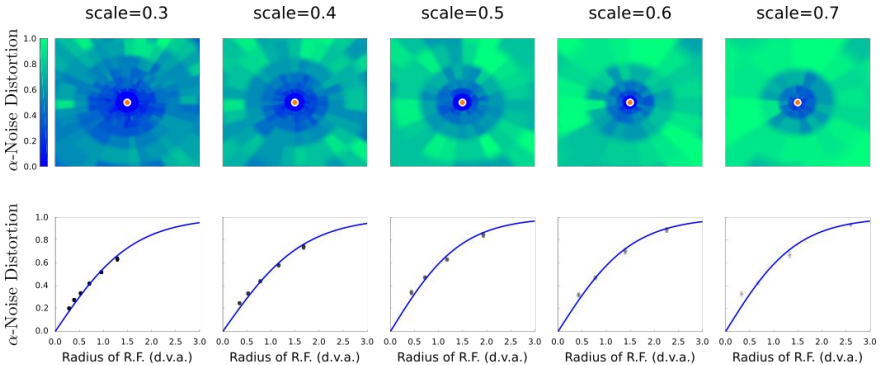
Deza, Jonnalagadda & Eckstein. ICLR 2019.

Perceptual Optimization Procedure



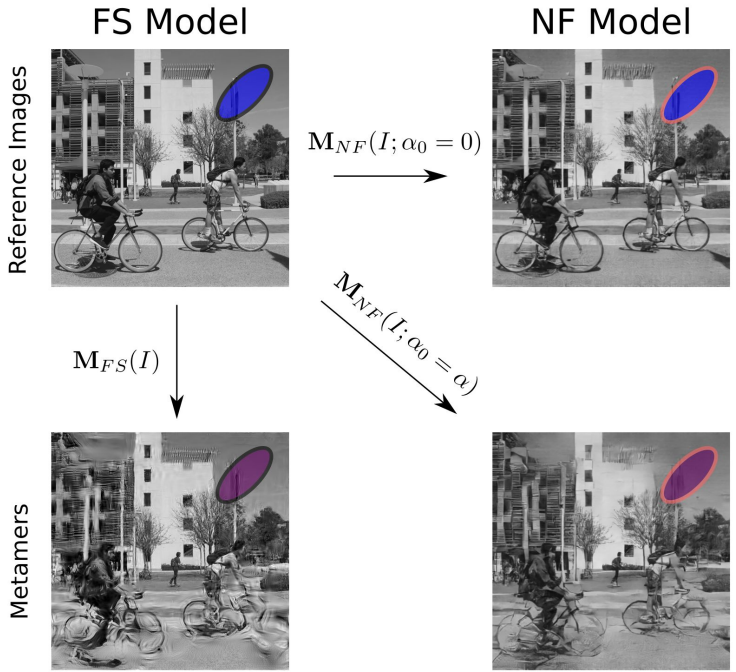
Minimize : $(\text{SSIM}(\,,\,) - \text{SSIM}(\,,\,))^2$

for every pooling region in the visual field



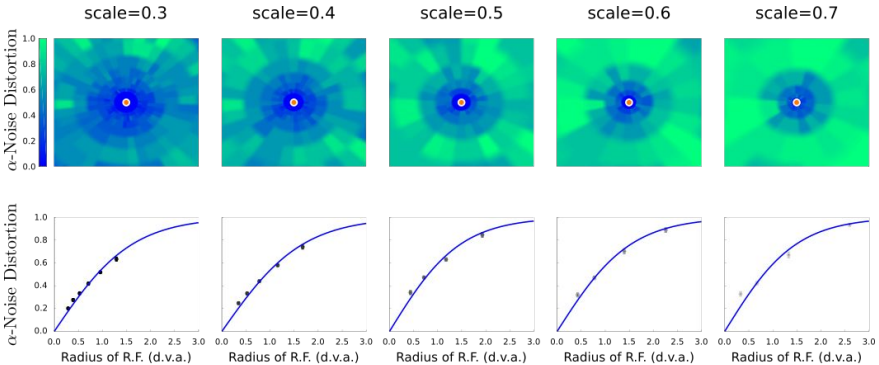
Deza, Jonnalagadda & Eckstein. ICLR 2019.

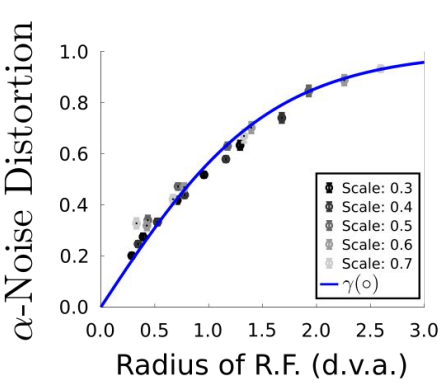
Perceptual Optimization Procedure

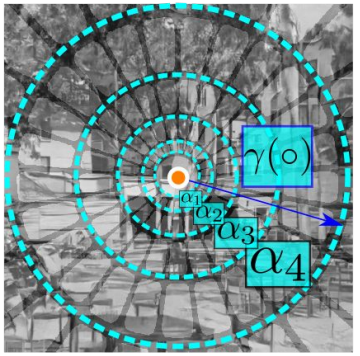


Minimize : $(\text{SSIM}(\bigcirc,\bigcirc) - \text{SSIM}(\bigcirc,\bigcirc))^2$

for every pooling region in the visual field





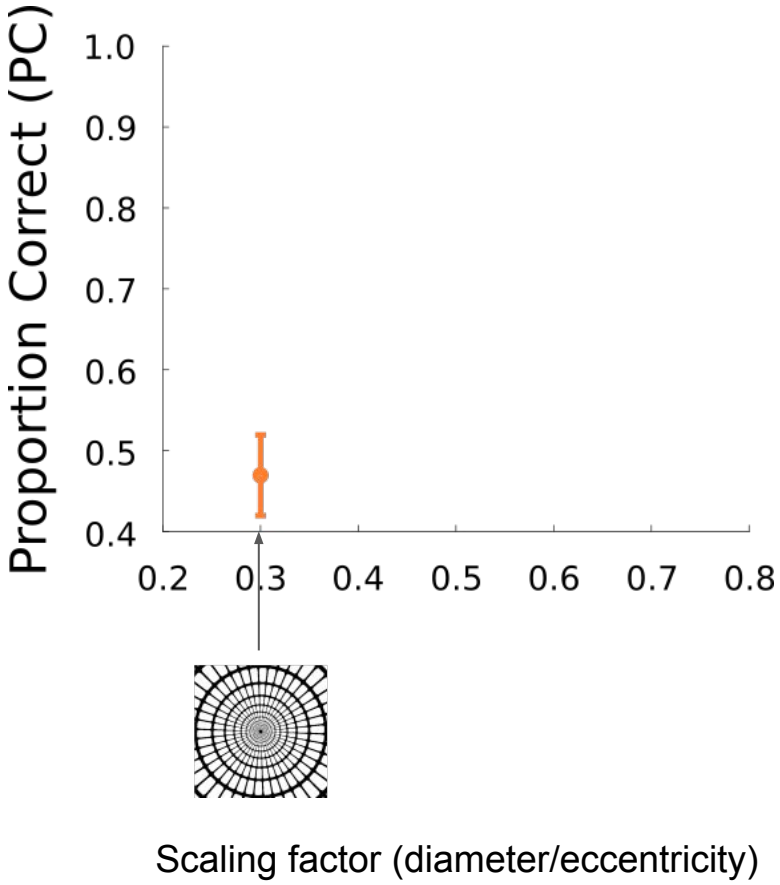


scale = 0.5

Deza, Jonnalagadda & Eckstein. ICLR 2019.

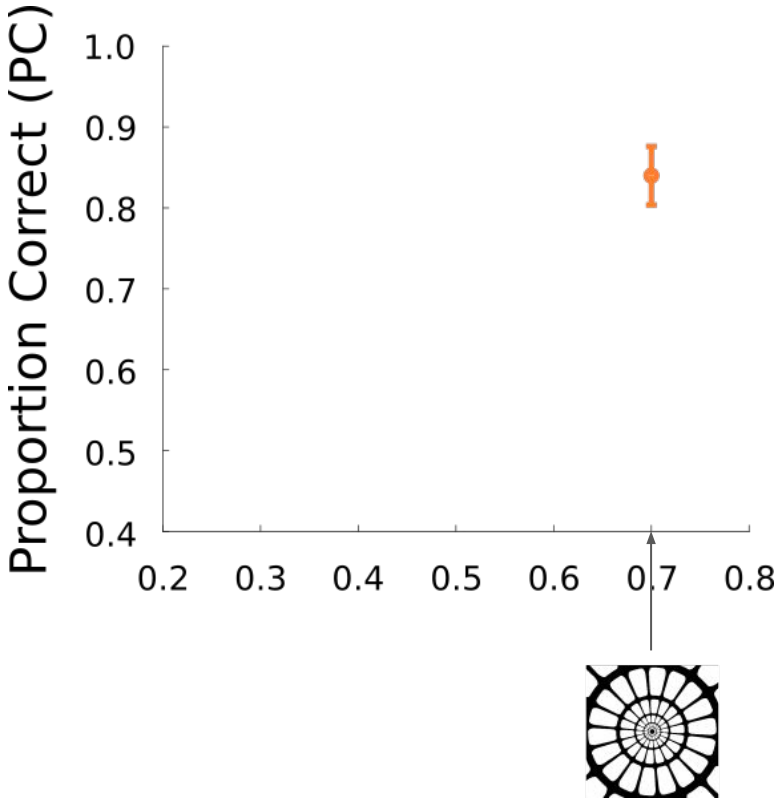
How do we test this?





How do we test this?

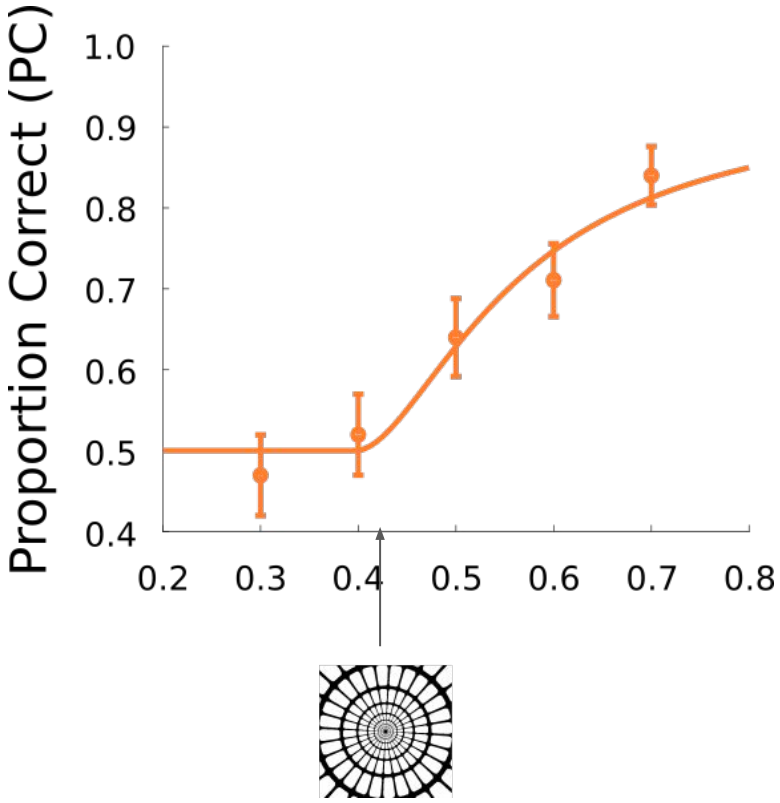




Scaling factor (diameter/eccentricity)

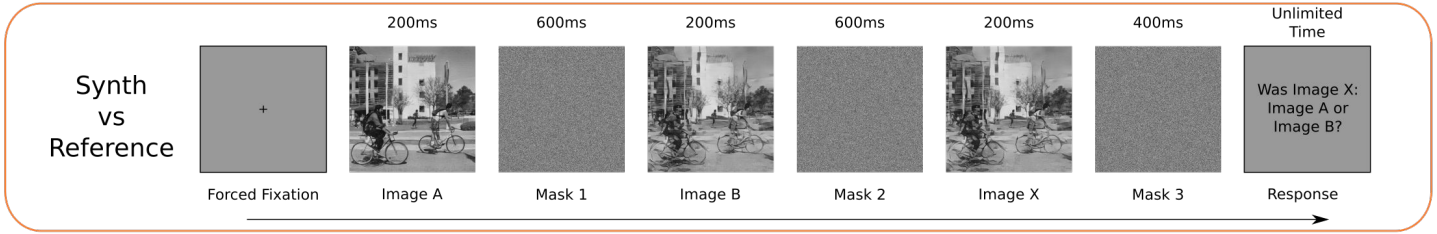
How do we test this?

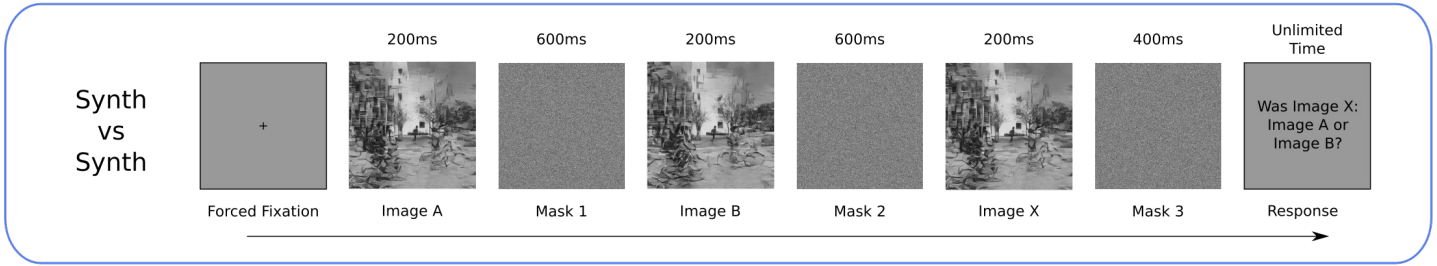




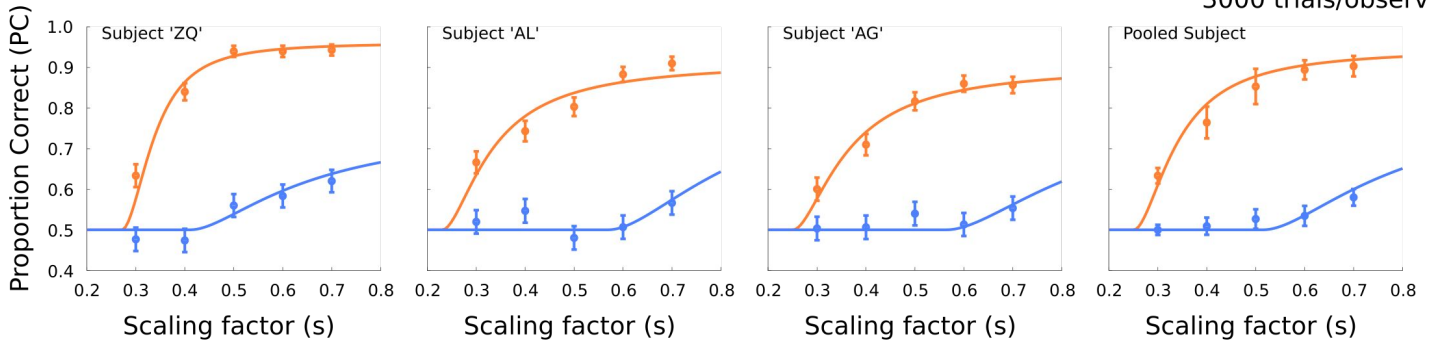
Scaling factor (diameter/eccentricity)

Human Psychophysics: ABX Task





3000 trials/observer



Results

Reference



Metamer (V1)



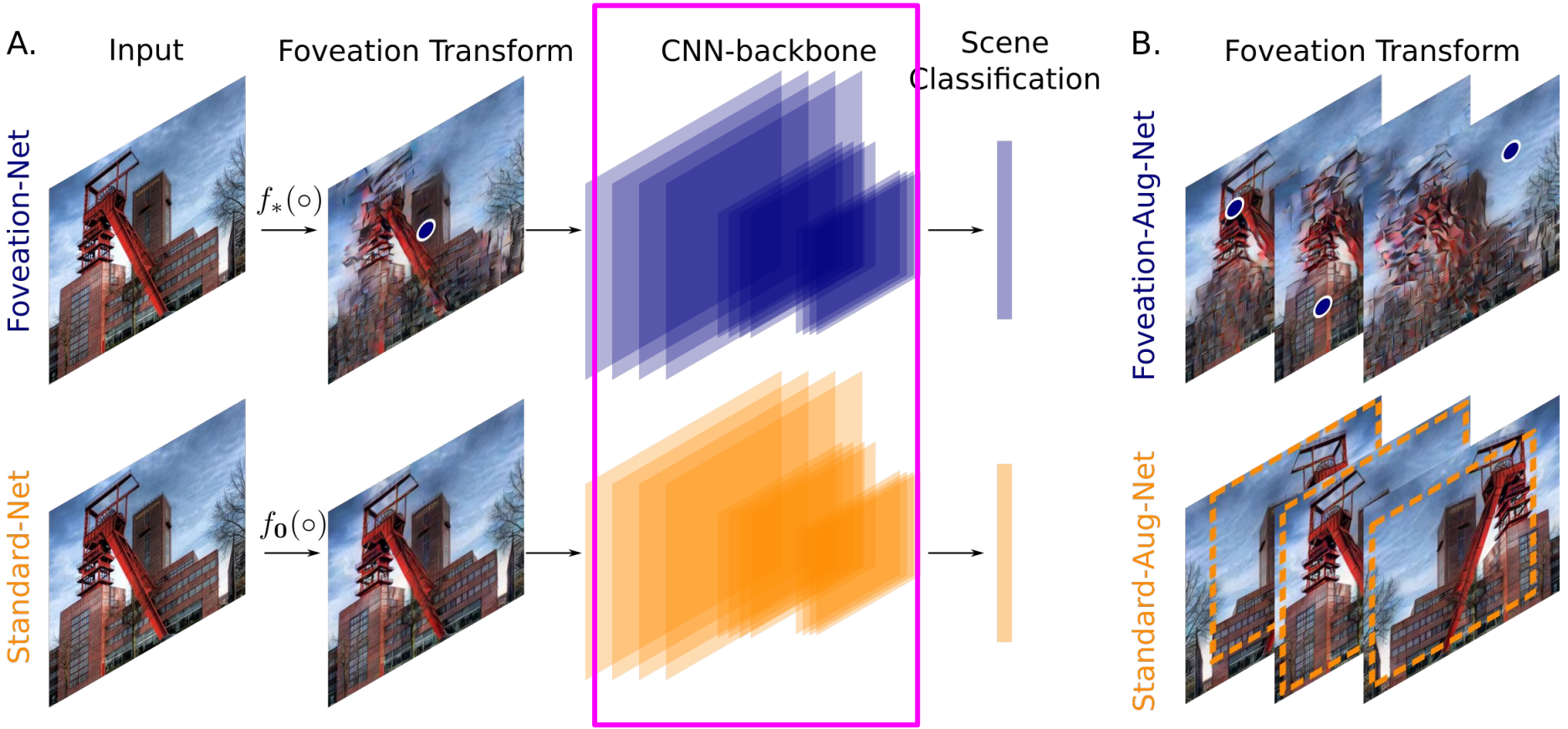
Metamer (V2)



Metamer (V2)



Foveated Perceptual Systems



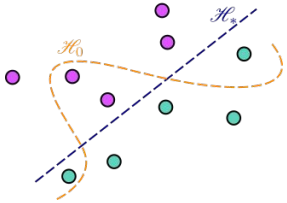
We perceptually evaluate each network!

Deza & Konkle. 2020.

II. Perceptual Implications of the Foveation Transform

a.k.a. what happens if you build perceptual invariance into deep neural networks

Generalization



Standard-Net



Foveation-Net



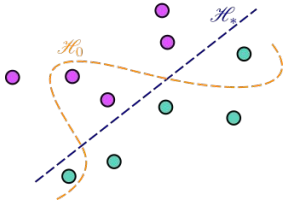
Standard-Aug-Net

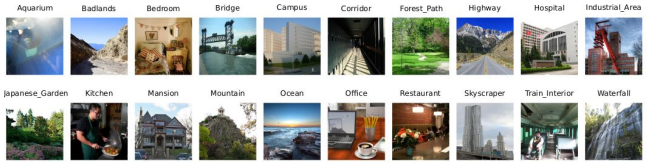


Foveation-Aug-Net



Generalization



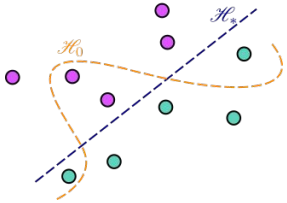


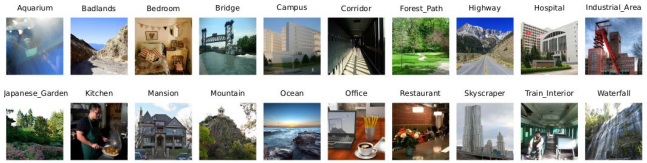
Scene Categorization

Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



Generalization





Scene Categorization

Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net

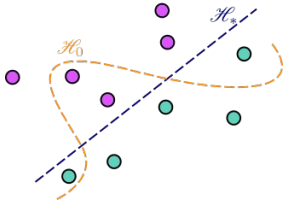


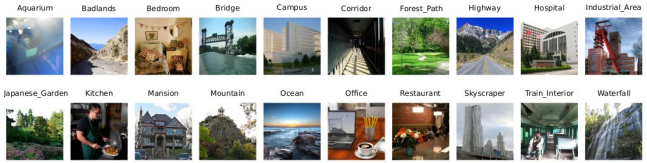
Feathers vs Wings





Generalization





Scene Categorization

Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



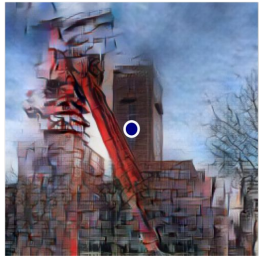
Feathers vs Wings





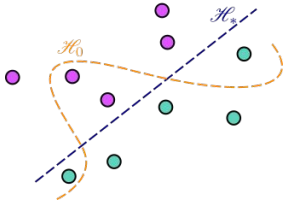
10 Foveation-Nets
10 Foveation-Aug -Nets

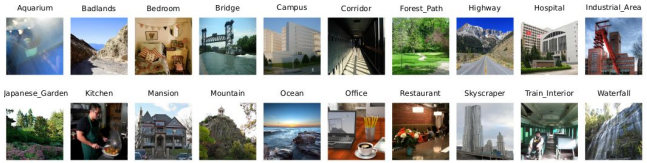
10 Standard-Nets
10 Standard-Aug -Nets





Generalization





Scene Categorization

Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



$(\blacksquare, \ , \ , \ )$ Generalization

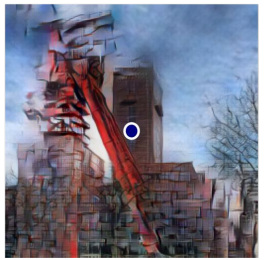
Feathers vs Wings



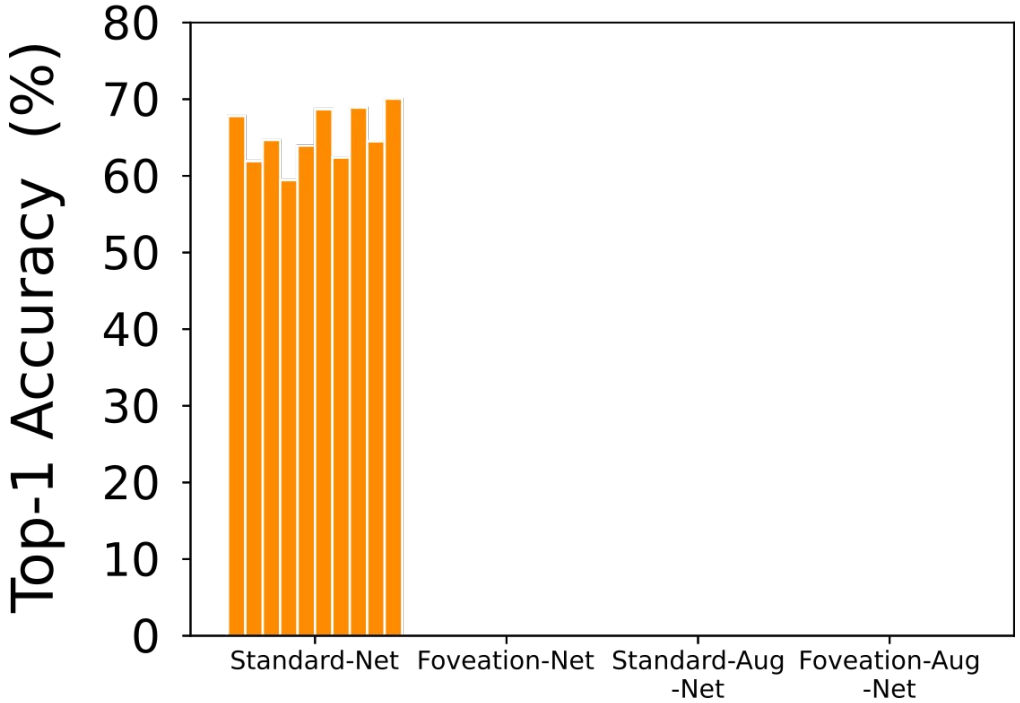


10 Foveation-Nets
10 Foveation-Aug -Nets

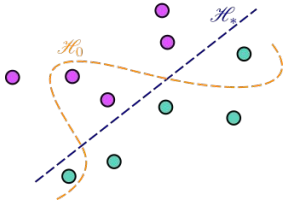
10 Standard-Nets
10 Standard-Aug -Nets

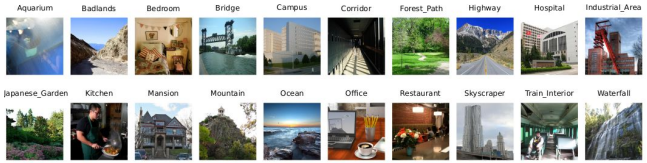






Generalization





Scene Categorization

Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



$(\blacksquare, \blacksquare, \blacksquare, \blacksquare)$ Generalization

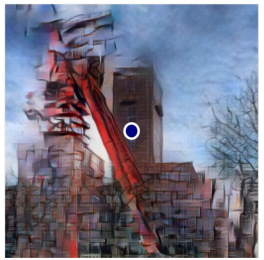
Feathers vs Wings



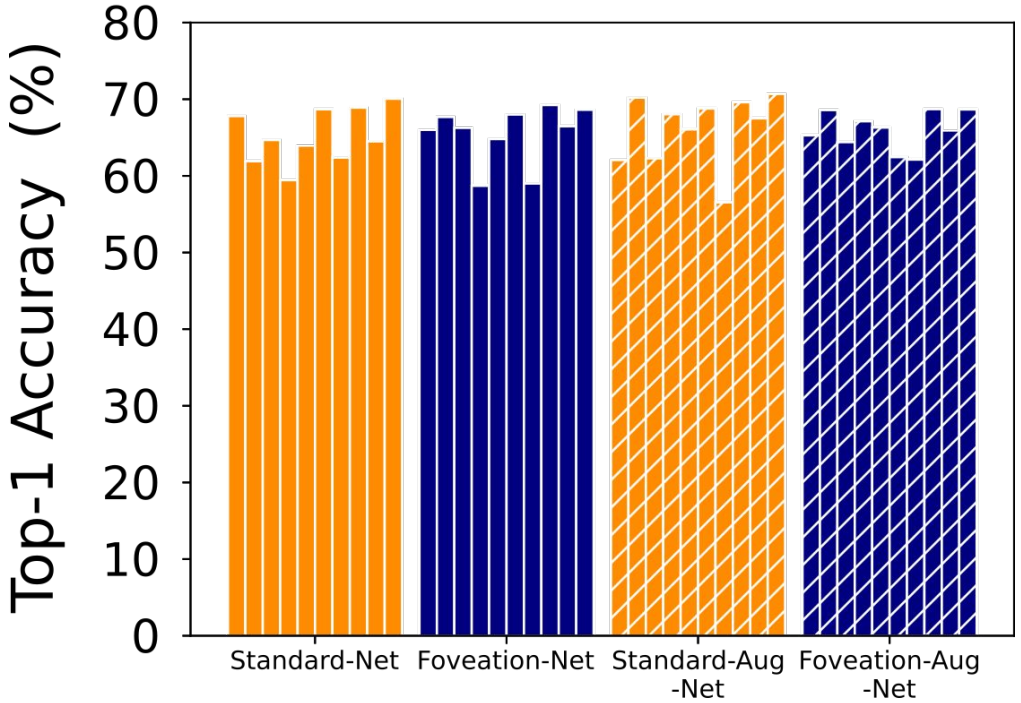


10 Foveation-Nets
10 Foveation-Aug -Nets

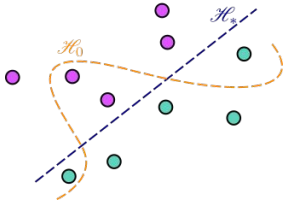
10 Standard-Nets
10 Standard-Aug -Nets

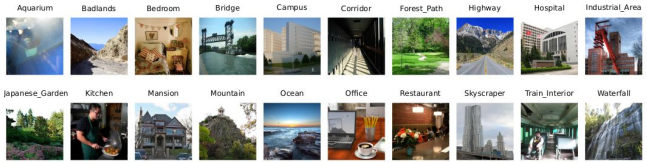






Generalization





Scene Categorization

Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



(■, ■, ◪, ◪) Generalization

(◆) Cross-Generalization

Feathers vs Wings

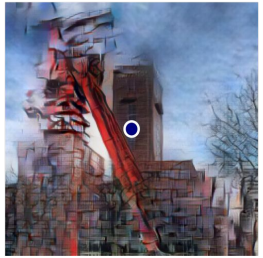


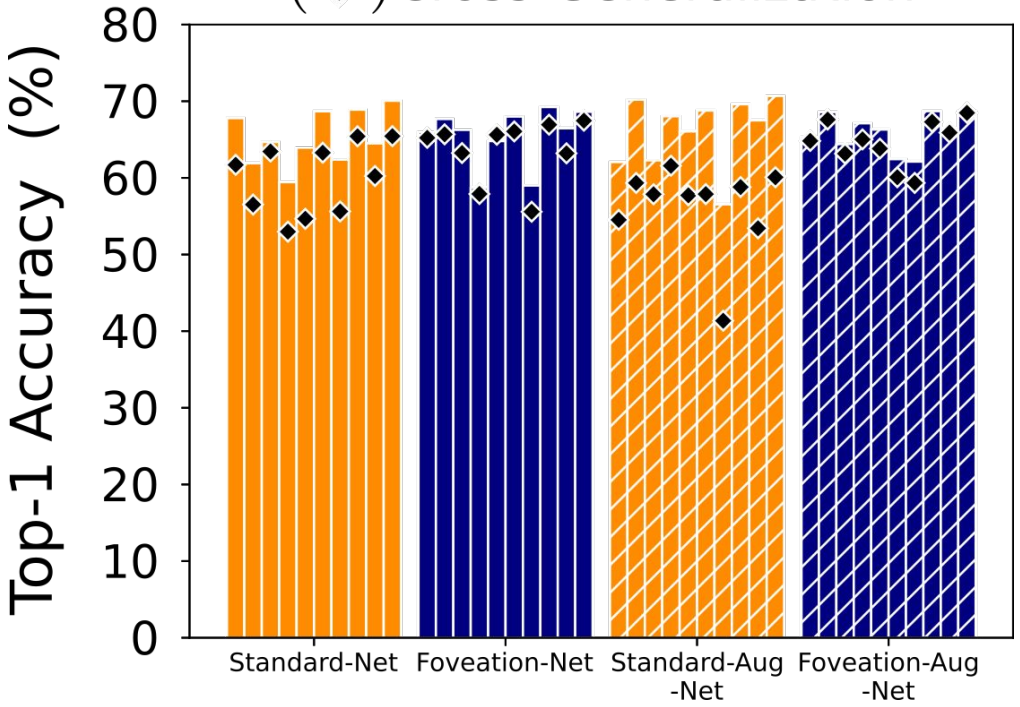


10 Foveation-Nets
10 Foveation-Aug -Nets

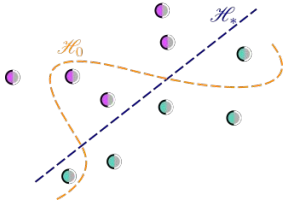
10 Standard-Nets
10 Standard-Aug -Nets







Robustness to Occlusion

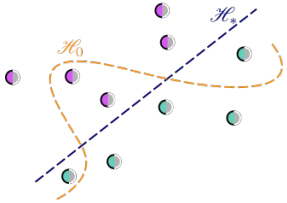




Standard-Net | Foveation-Net | Standard-Aug-Net | Foveation-Aug-Net



Robustness to Occlusion





Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



How does the accuracy of each perceptual system change as we *occlude* the visual stimulus?

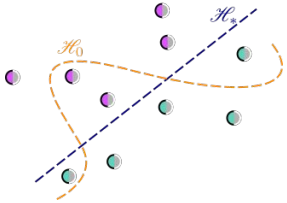
Glaucoma



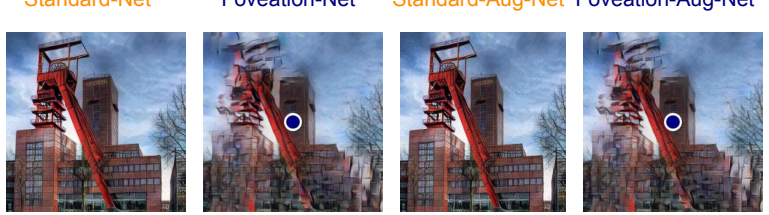
Scotoma



Robustness to Occlusion







How does the accuracy of each perceptual system change as we *occlude* the visual stimulus?

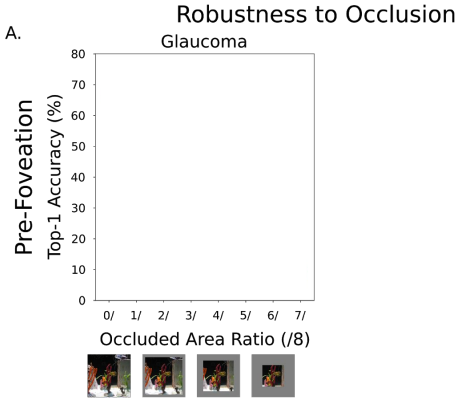
Glaucoma



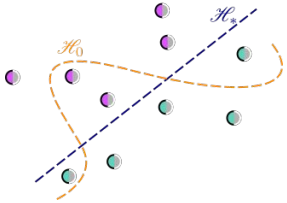
Scotoma



Robustness to Occlusion



Robustness to Occlusion





Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



Robustness to Occlusion

How does the accuracy of each perceptual system change as we *occlude* the visual stimulus?

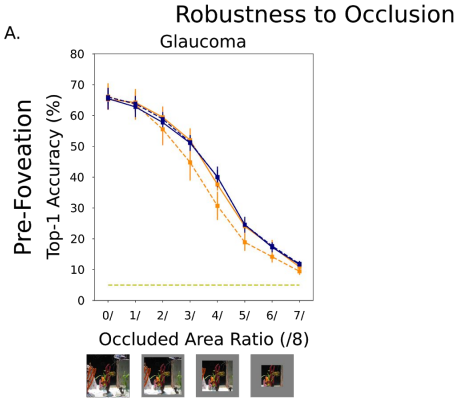
Glaucoma



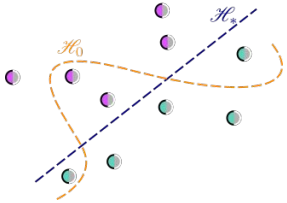
Scotoma



A.



Robustness to Occlusion





Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



How does the accuracy of each perceptual system change as we *occlude* the visual stimulus?

Glaucoma

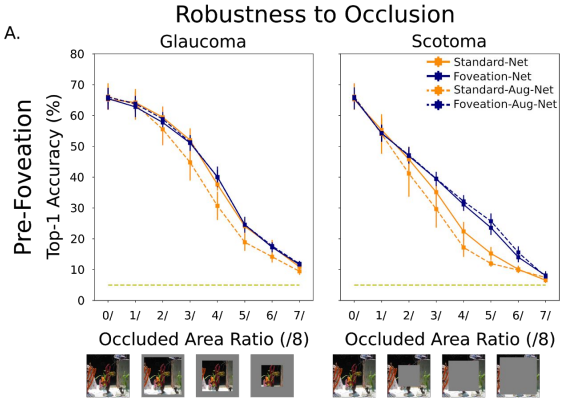


Scotoma

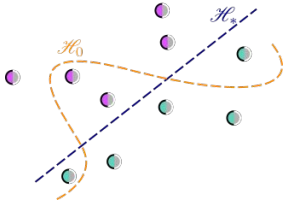


Robustness to Occlusion

A.



Robustness to Occlusion



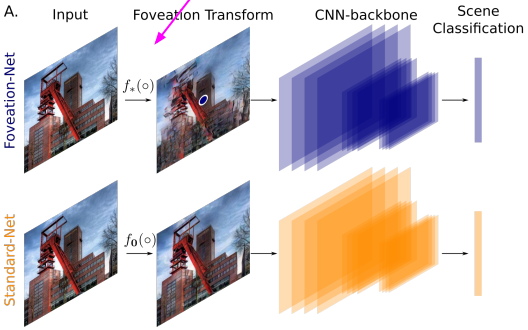


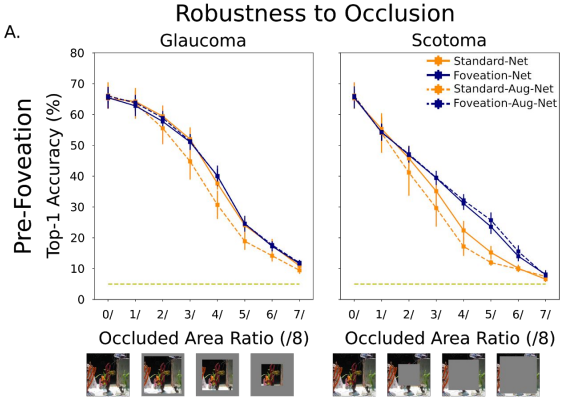
Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net

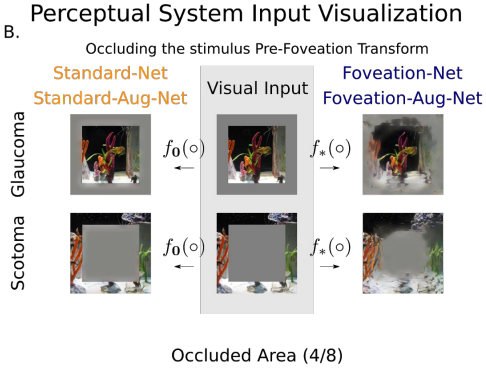


Robustness to Occlusion

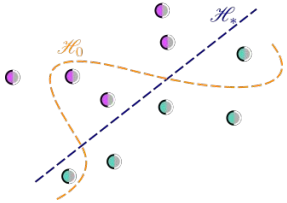
How does the accuracy of each perceptual system change as we *occlude* the visual stimulus?







Robustness to Occlusion

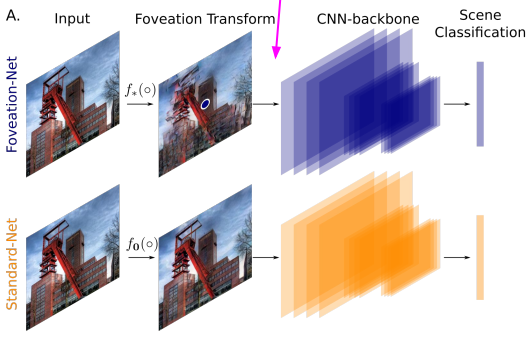


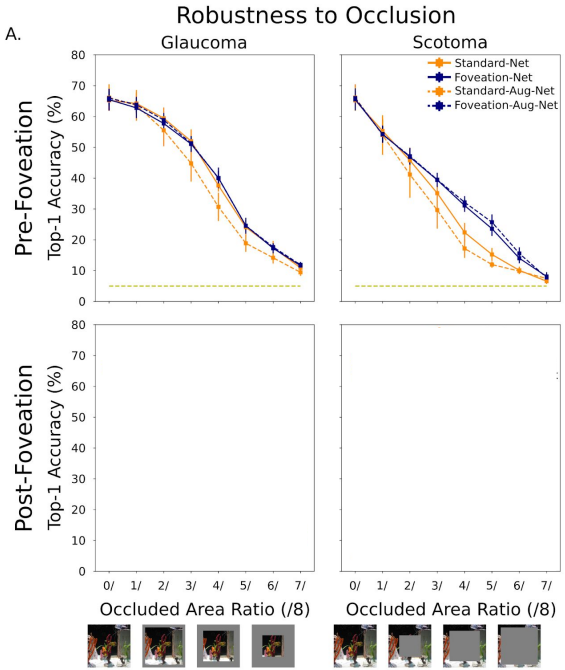


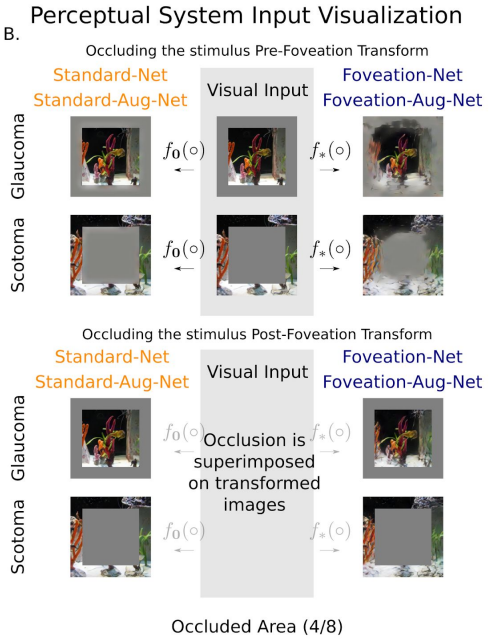


Robustness to Occlusion

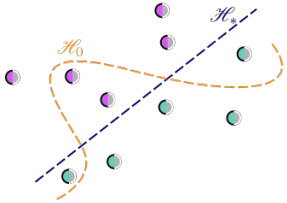
How does the accuracy of each perceptual system change as we *occlude* the visual stimulus?







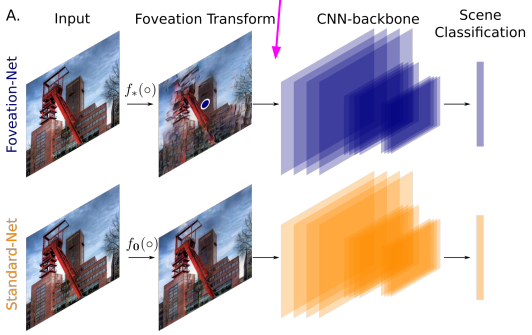
Robustness to Occlusion



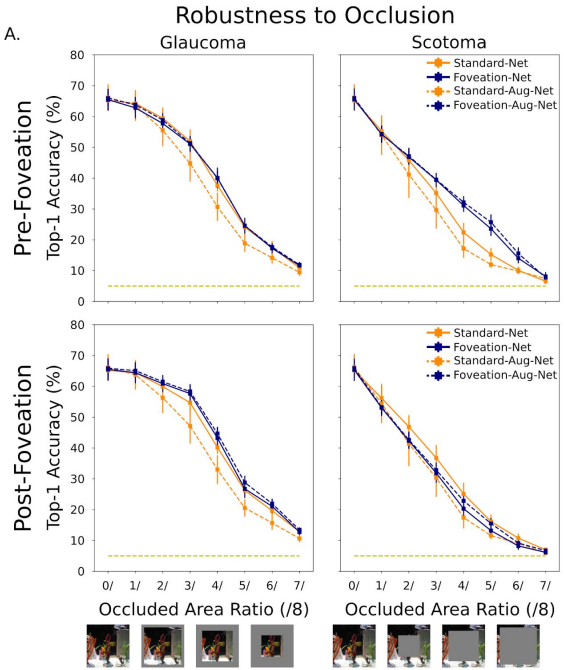


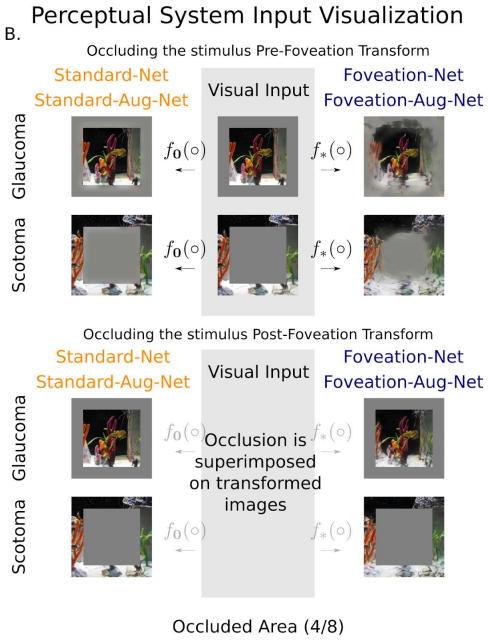


How does the accuracy of each perceptual system change as we *occlude* the visual stimulus?



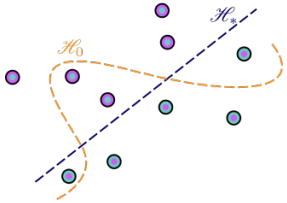
Robustness to Occlusion





Robustness to Scotomal Occlusion stems from the foveation process

Windowing Cue-Conflict





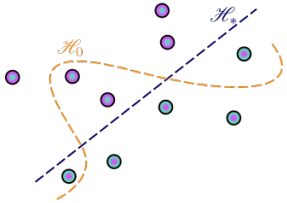
Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



How does the accuracy of each perceptual system change as we *bias* the visual stimulus?



Windowing Cue-Conflict

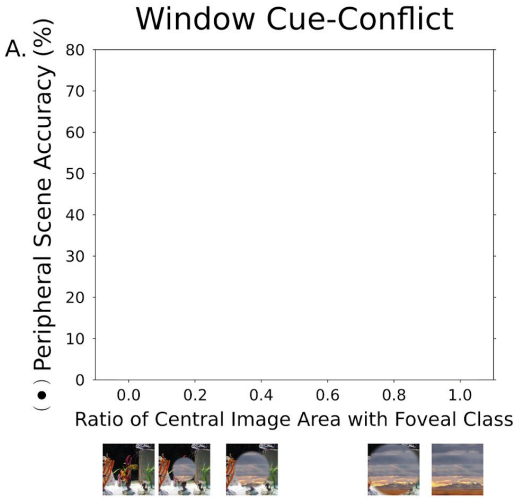


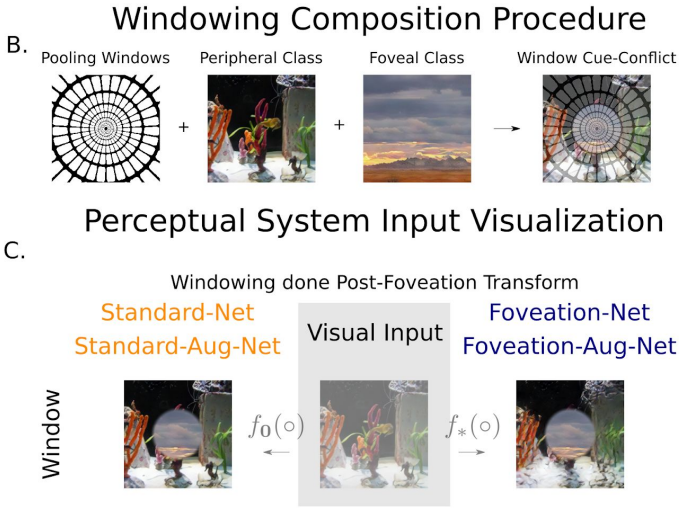




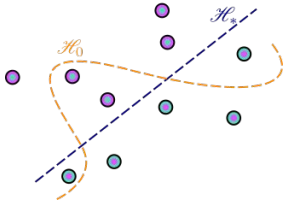
How does the accuracy of each perceptual system change as we *bias* the visual stimulus?







Windowing Cue-Conflict

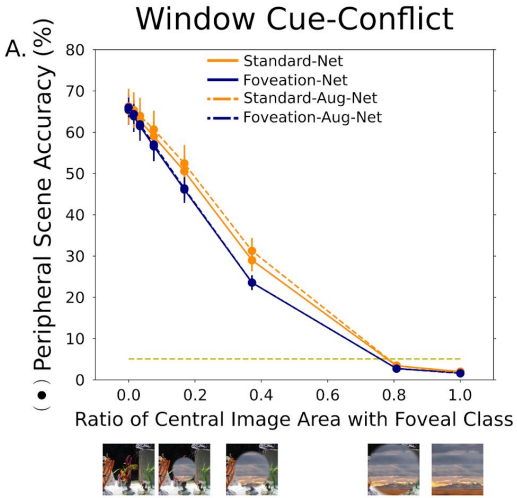


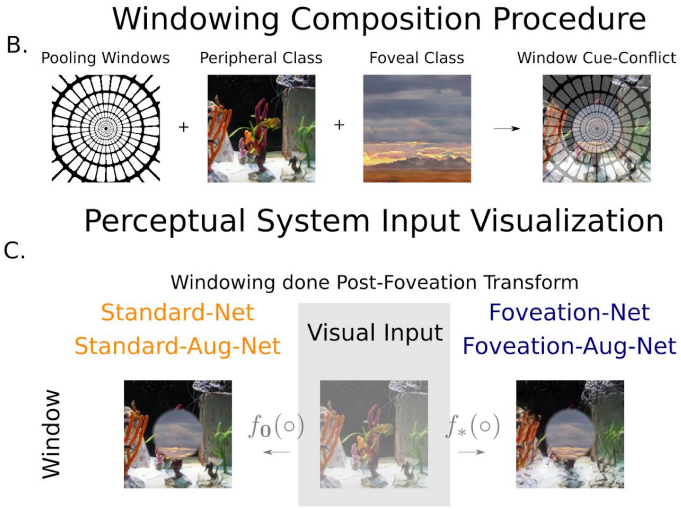




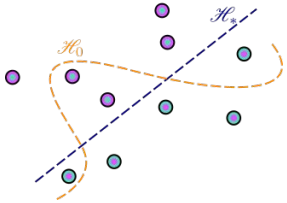
How does the accuracy of each perceptual system change as we *bias* the visual stimulus?







Windowing Cue-Conflict



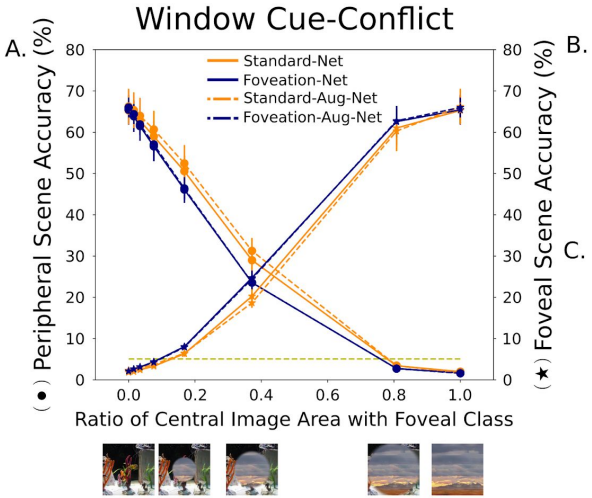


Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net

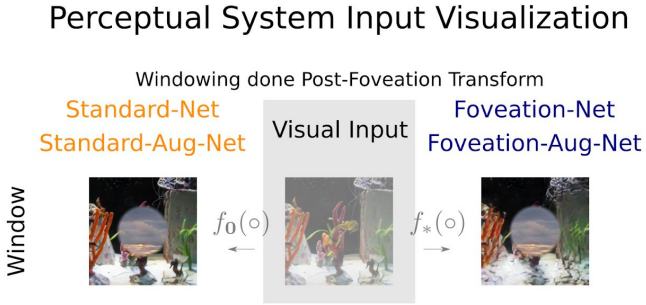


How does the accuracy of each perceptual system change as we *bias* the visual stimulus?

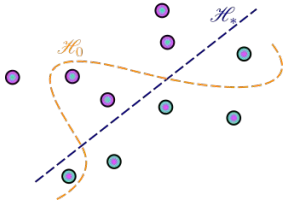








Windowing Cue-Conflict





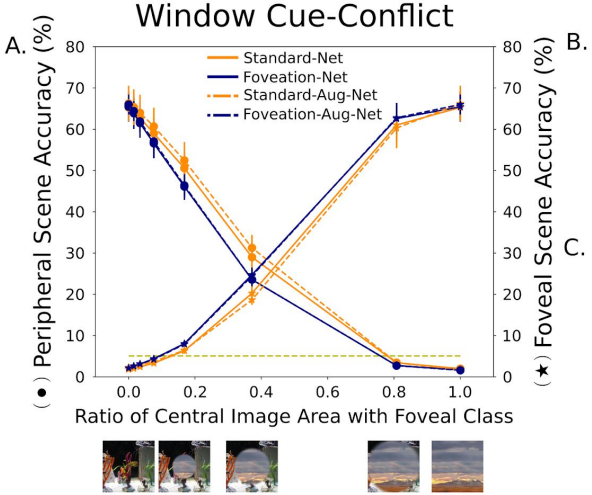
Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



How does the accuracy of each perceptual system change as we *bias* the visual stimulus?



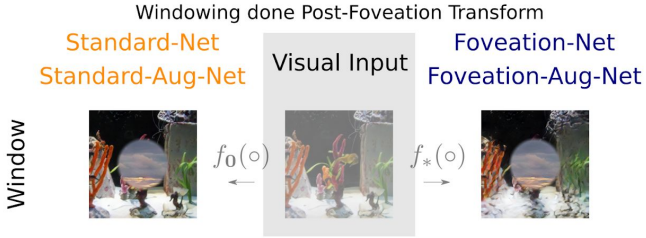
Window Cue-Conflict



Windowing Composition Procedure

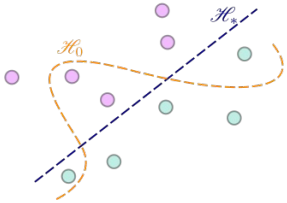


Perceptual System Input Visualization



Foveation-Nets are more biased than Standard-Nets to central image regions

Spatial Frequency Sensitivity

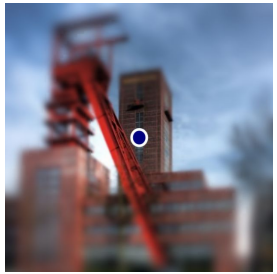




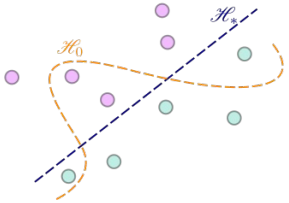
Standard-Net Foveation-Net Standard-Aug-Net Foveation-Aug-Net



How does the accuracy of each perceptual system change as we *filter* the visual stimulus?



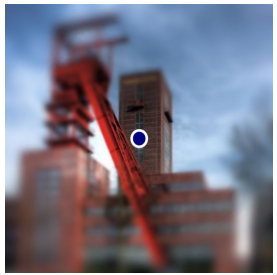
Spatial Frequency Sensitivity



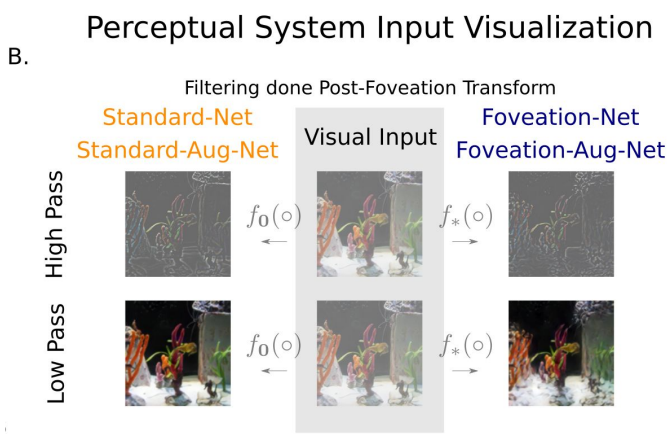




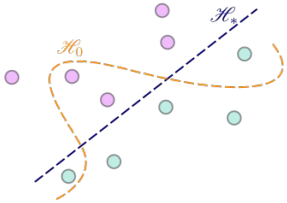
How does the accuracy of each perceptual system change as we *filter* the visual stimulus?



B.



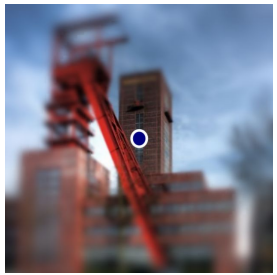
Spatial Frequency Sensitivity

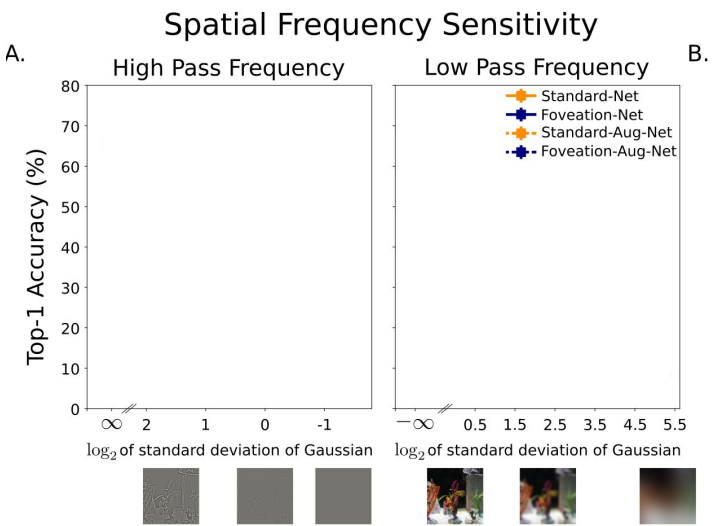


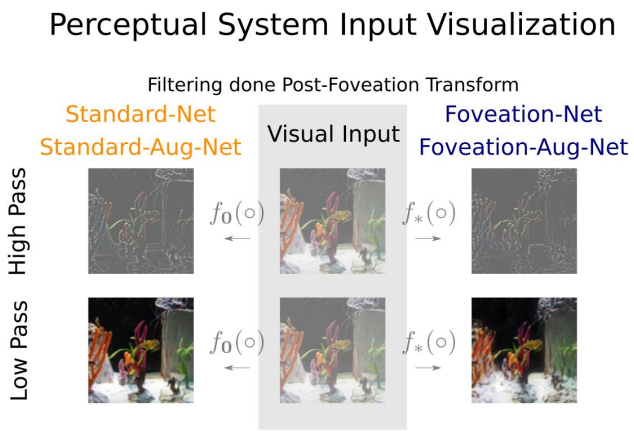




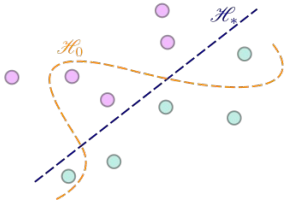
How does the accuracy of each perceptual system change as we *filter* the visual stimulus?







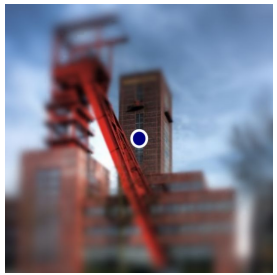
Spatial Frequency Sensitivity

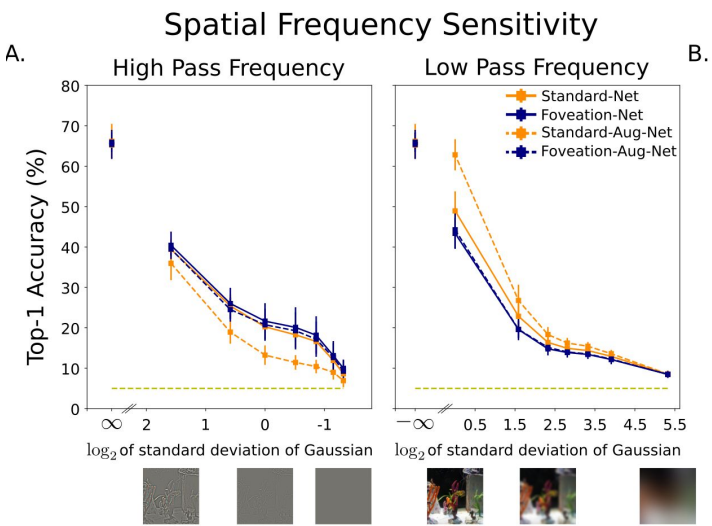


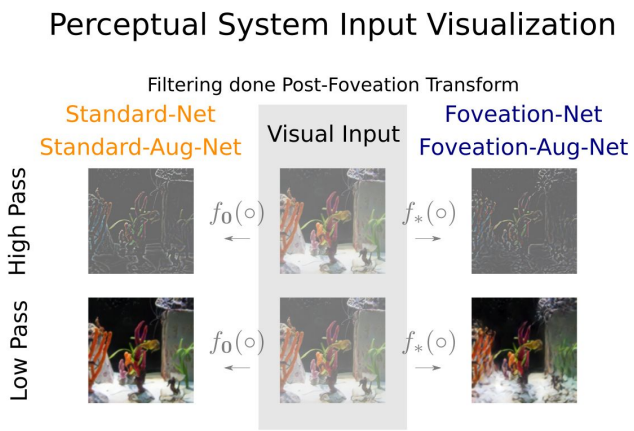




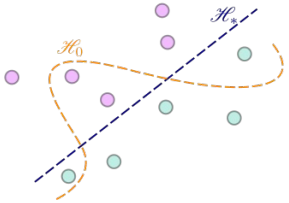
How does the accuracy of each perceptual system change as we *filter* the visual stimulus?







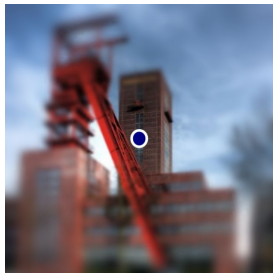
Spatial Frequency Sensitivity

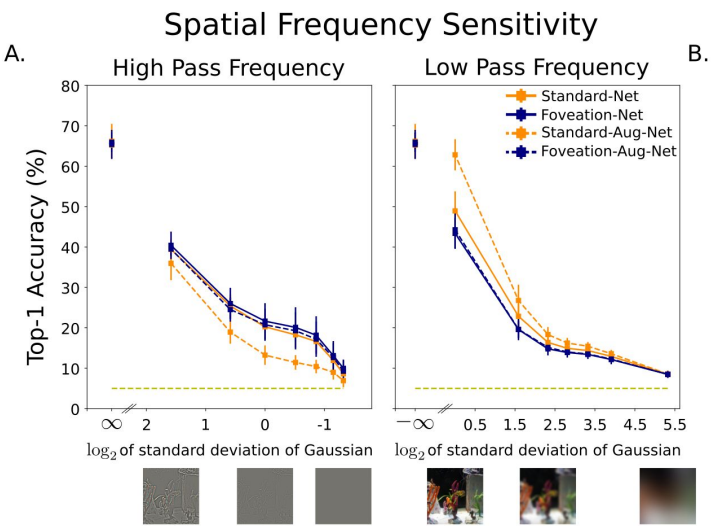


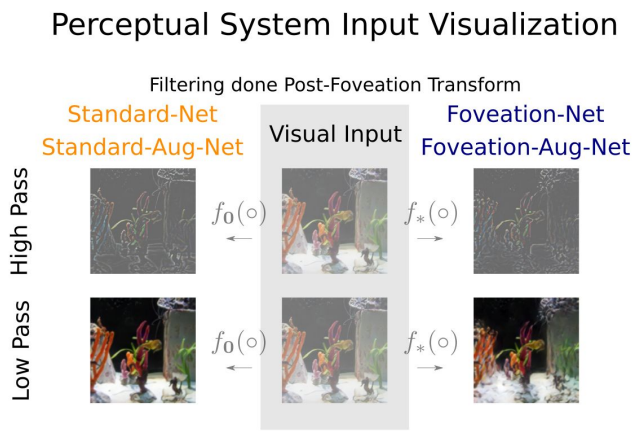




How does the accuracy of each perceptual system change as we *filter* the visual stimulus?

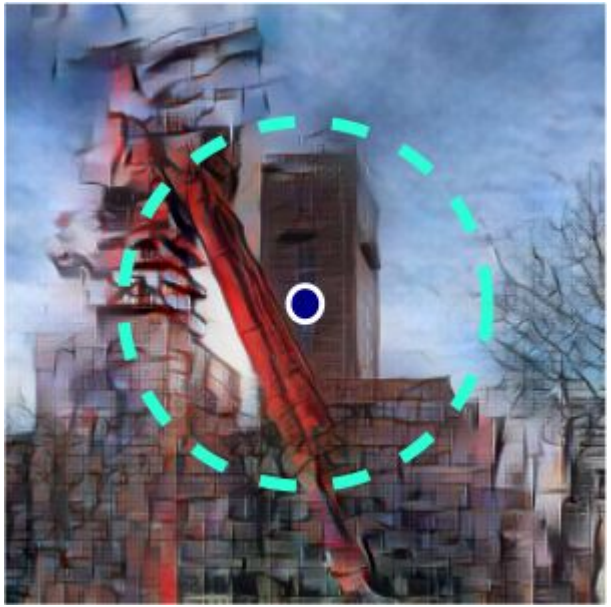




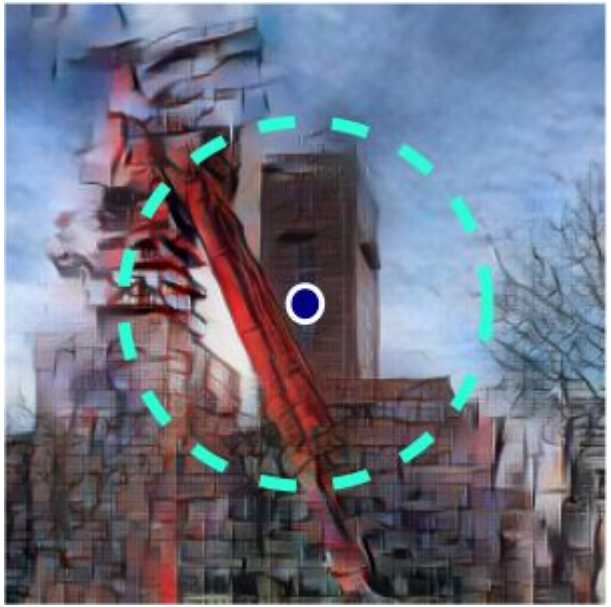


Foveation-Nets are less sensitive than Standard-Nets to low spatial frequency

Emergence of a *Central Focusing Mechanism* for Foveated Systems vs Non-Foveated Systems

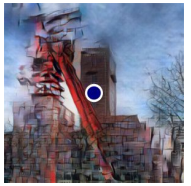


Emergence of a *Central Focusing Mechanism* for Foveated Systems vs Non-Foveated Systems



Nominal Differences:
Eye-Movements for scene perception

Foveation-Net



Foveation-Aug-Net



Greater Differences:
Augmentation increases most effects between non-foveated systems.
Why?

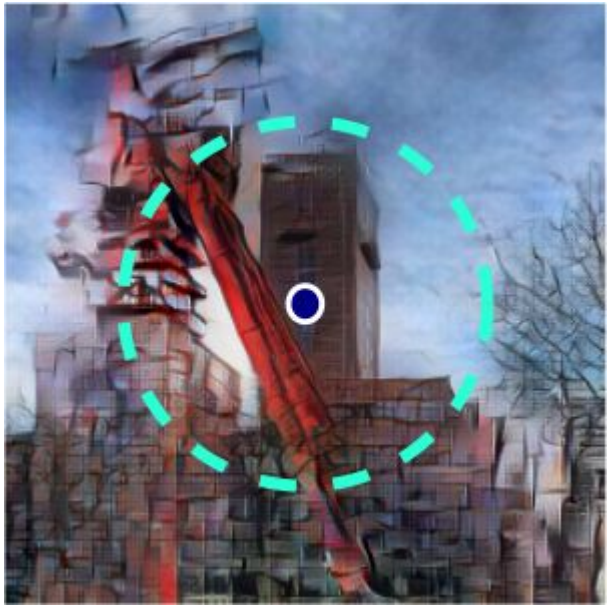
Standard-Net



Standard-Aug-Net



Emergence of a *Central Focusing Mechanism* for Foveated Systems vs Non-Foveated Systems



Foveation-Net | Foveation-Aug-Net



Nominal Differences:
Eye-Movements for scene perception

Standard-Net | Standard-Aug-Net

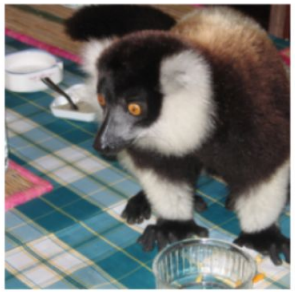


Greater Differences:
Augmentation increases most effects between non-foveated systems. Why?

Scenes vs Objects



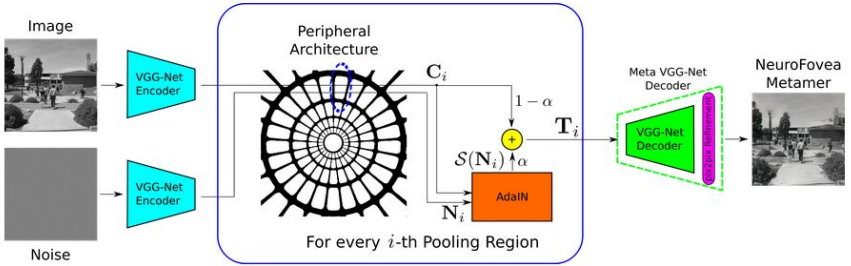
vs



Geirhos et al. 2019

NeuroFovea.lua	Fixed Grayscale conversion	2 months ago
README.md	Update README.md	4 months ago
download_models_and_stimuli.sh	Update download_models_and_stimuli.sh	4 months ago

README.md



Towards Metamerism via Foveated Style Transfer

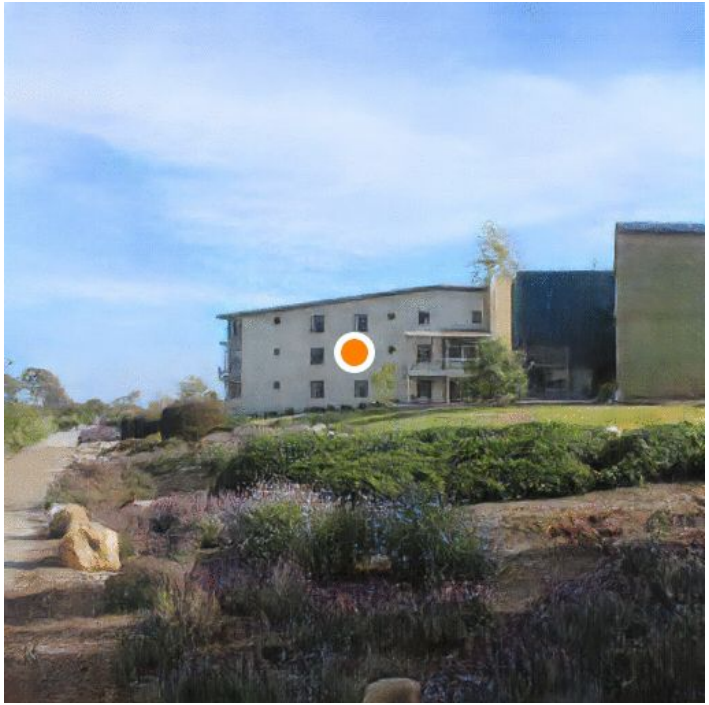
This repository containts the code to reproduce the Metamers used in the paper (Deza, Jonnalagadda, Eckstein. ICLR 2019). Link to the paper and discussion in openreview: https://openreview.net/forum?id=BJzbG20cFQ

This code has been tested successfully on CUDA version 8.0 (Ubuntu 14.04 and 16.04) and CUDA version 10.0 (Ubuntu 18.04).

The code to implement our model is mainly driven by:

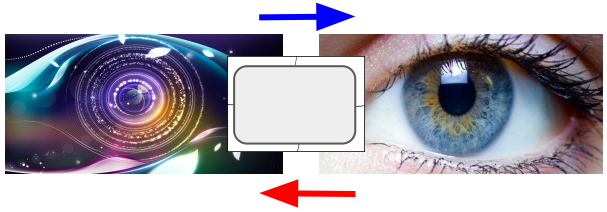
- Adaptive Instance Normalization code: https://github.com/xunhuang1995/AdaIN-style
- pix2pix super-resolution module: https://github.com/phillipi/pix2pix
- The original Metamer code of Freeman & Simoncelli: https://github.com/freeman-lab/metamers
- A set of localized pooling regions stored in the *Receptive_Fields/* folder for the different rate of growth of the receptive fields specified by the scaling factor which should match the human psychophysical testing procedure as specified in the paper.

What is a Metamer?



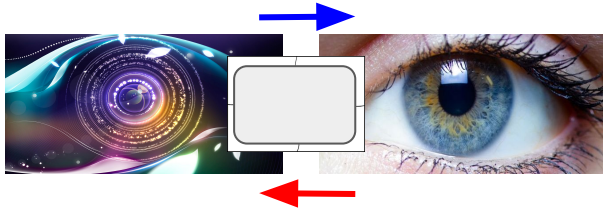
https://github.com/ArturoDeza/NeuroFovea

III. Perspectives & New Directions



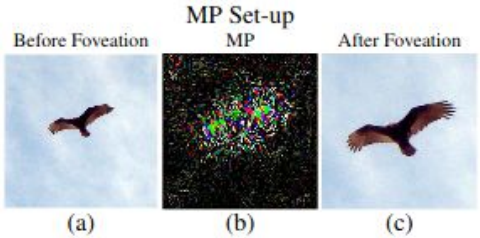


Akbas & Eckstein. 2014/17.

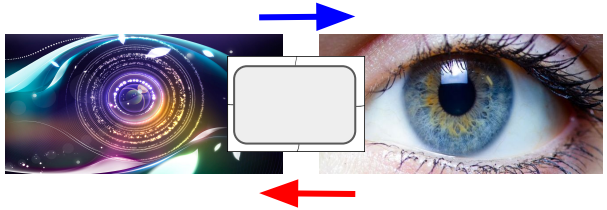




Akbas & Eckstein. 2014/17.

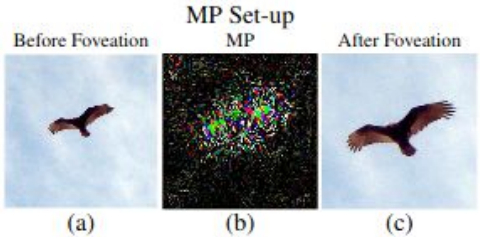


Luo, Boix, Roig,
Poggio & Zhao. 2015.

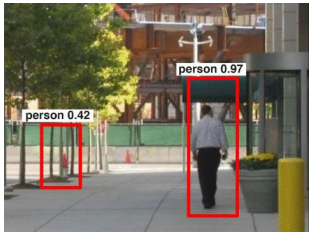




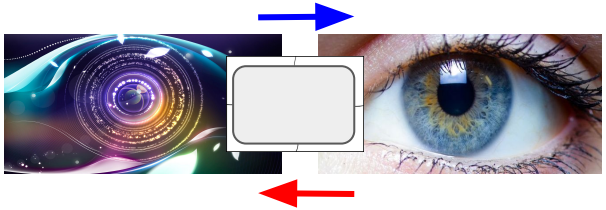
Akbas & Eckstein. 2014/17.



Luo, Boix, Roig,
Poggio & Zhao. 2015.

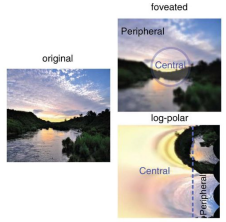


Pramod, Kitti & Arun. 2018.

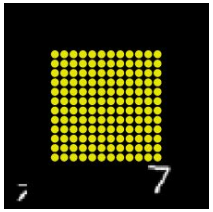




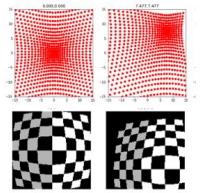
Rosenholtz, Yu & Keshvari. 2019.



Wang & Cottrell. 2017.



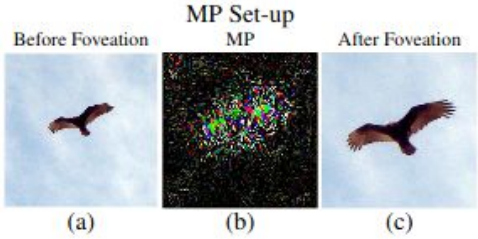
Cheung, Weiss & Olshausen. 2017.



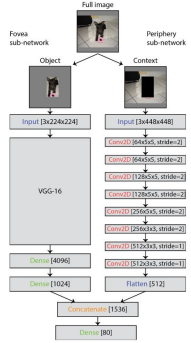
Reddy, Banburski, Pant & Poggio. 2020.



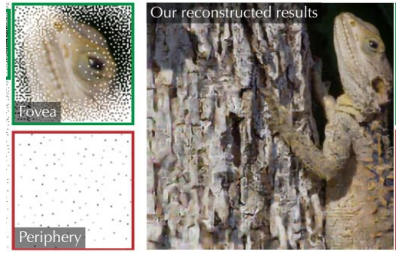
Akbas & Eckstein. 2014/17.



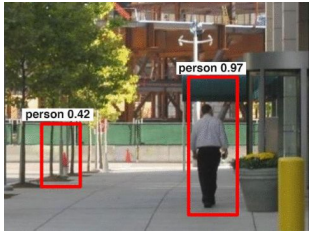
Luo, Boix, Roig, Poggio & Zhao. 2015.



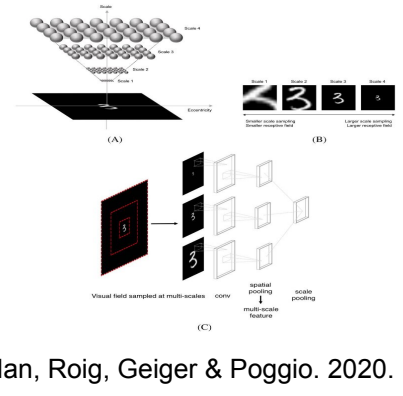
Wu, Wu & Kreiman. 2017.



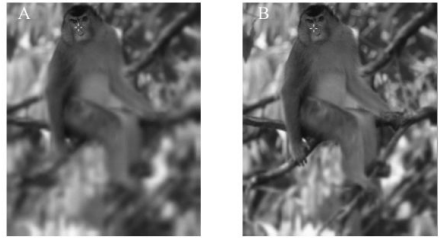
DeepFovea (Facebook Reality Labs). 2019.



Pramod, Kitti & Arun. 2018.



Han, Roig, Geiger & Poggio. 2020.



Geisler & Perry. 1998.

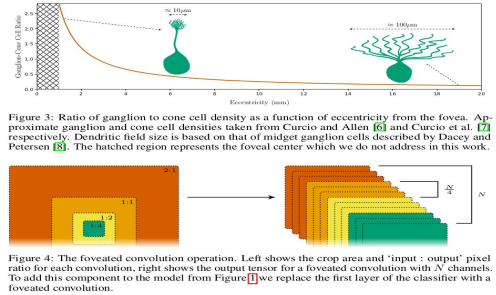


Figure 3: Ratio of ganglion to cone cell density as a function of eccentricity from the fovea. Approximate ganglion and cone cell densities taken from Curcio and Allen [6] and Curcio et al. [7] respectively. Dendritic field size is based on that of midget ganglion cells described by Dacey and Petersen [8]. The hatched region represents the foveal center which we do not address in this work.

Figure 4: The foveated convolution operation. Left shows the crop area and 'input : output' pixel ratio for each convolution, right shows the output tensor for a foveated convolution with N channels. To add this component to the model from Figure 1 we replace the first layer of the classifier with a foveated convolution.

Harris, Niranjan & Harre. 2019.

Metamerizing Adversarial Examples

LOCKHEED MARTIN

I



Bicycles

$\mathbf{A}(\cdot)$

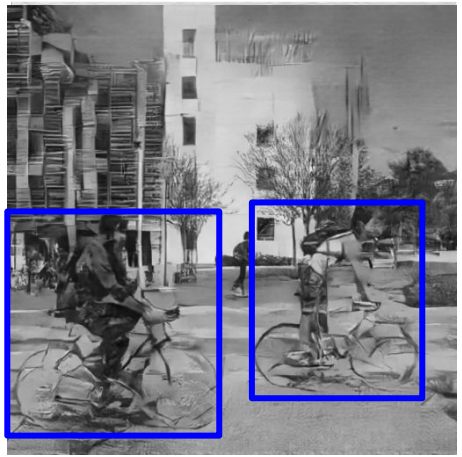
$\mathbf{A}(I)$



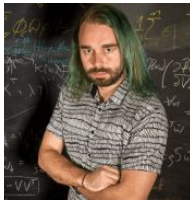
Traffic Light (Green)

$\mathbf{M}(\cdot)$

$\mathbf{M}(\mathbf{A}(I))$

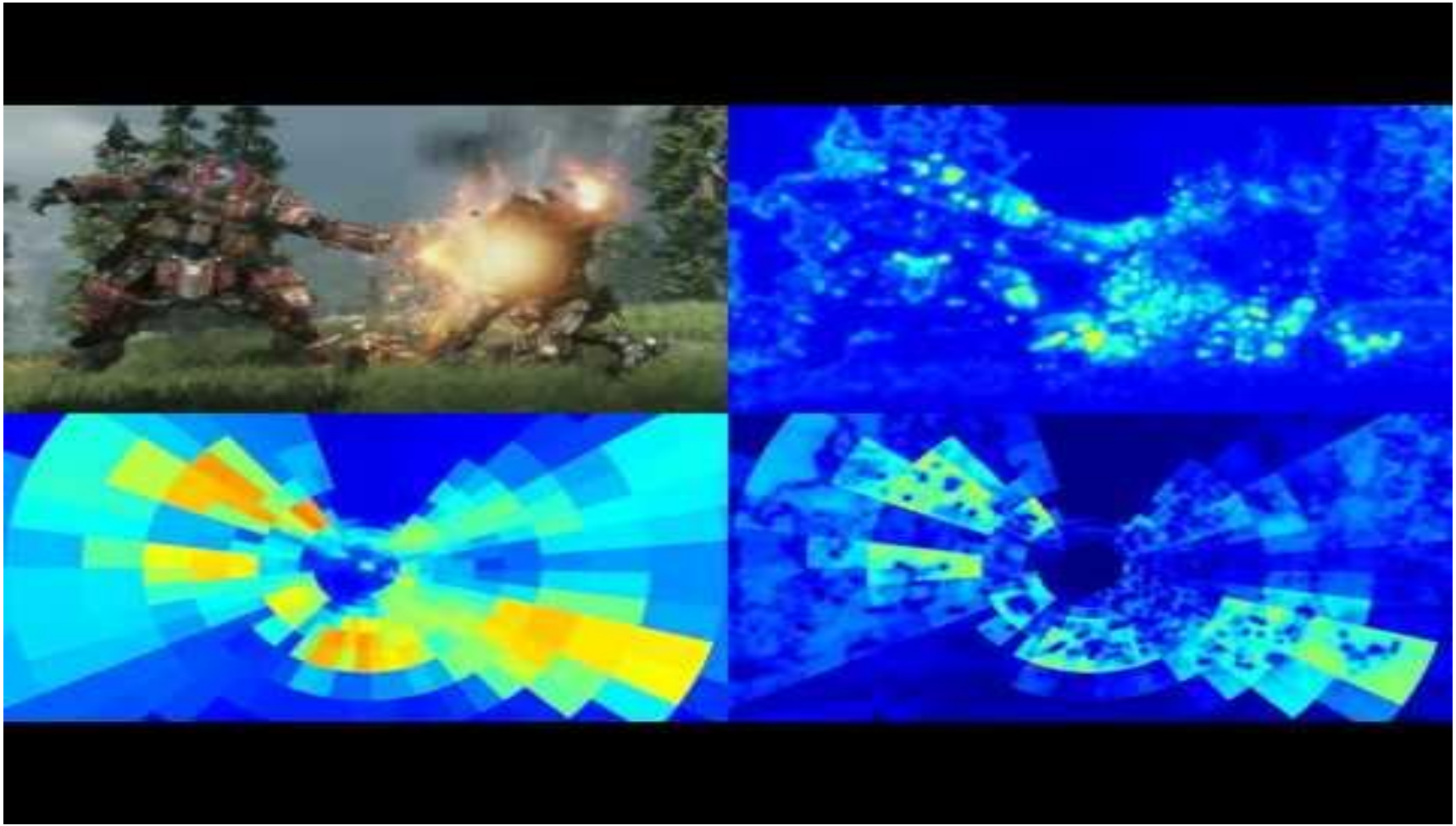


Bicycles



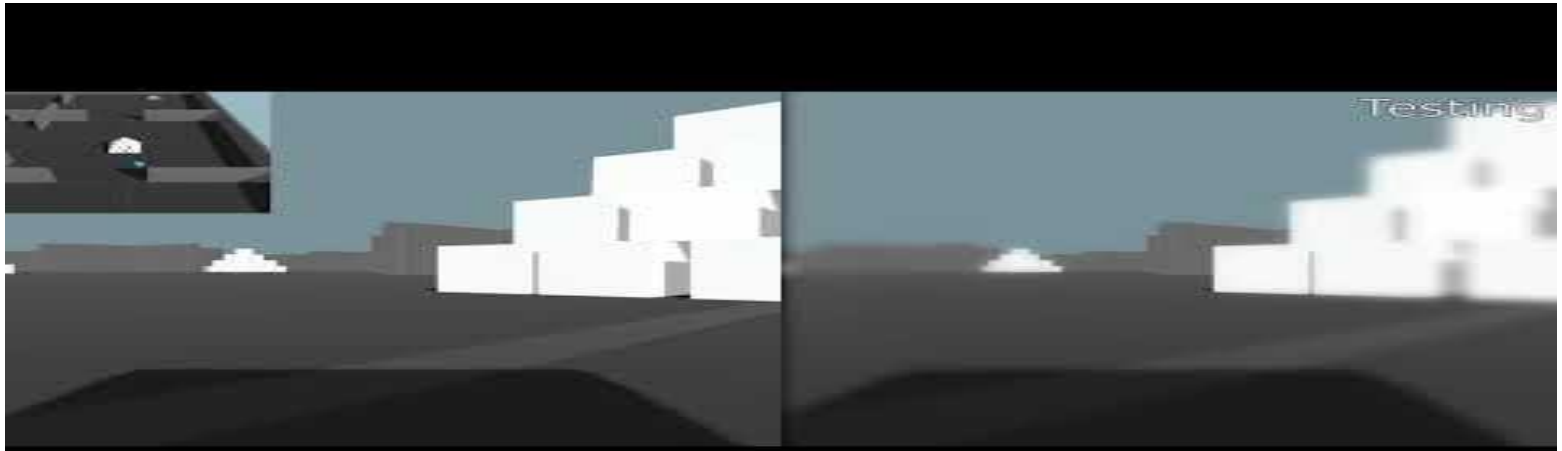
Andy Banburski

Inspired from : Reddy, Banburski, Pant & Poggio (2020)
Luo, Boix, Roig, Poggio & Zhao (2015)



Deza & Eckstein. NeurIPS 2016.

Agents





Ronald Alvarez

Robots





Elian Malkin





SVRHM 2019

Shared Visual Representations in Human & Machine Intelligence

NeurIPS Workshop

The goal of the SVRHM workshop is to disseminate findings from neuroscience, psychology, and cognitive science that might inform modern machine learning methods.

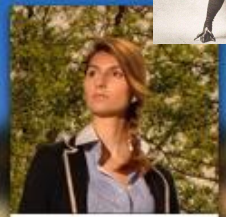
Focus Areas

- Biological Inspiration and Inductive Biases
- Robustness and Generalization
- New Datasets for Comparing Humans and Machines

www.svrhm2019.com



Olivier Henaff
Google DeepMind



Irina Higgins
Google DeepMind



Bill Freeman
MIT



James DiCarlo
MIT



Talia Konkle
Harvard



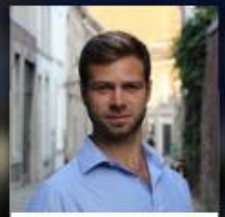
Matthias Bethge
Universität Tübingen



Niko Kriegeskorte
Columbia



Erin Grant
UC Berkeley



Ruairidh Battleday
Princeton



Will Xiao
Harvard

CALL FOR PAPERS

Organized by Arturo Deza, Joshua Peterson, Apurva Ratan Murty, and Thomas Griffiths

SVRHM

2020!

Shared Visual Representations in Human & Machine Intelligence

NeurIPS Workshop

The goal of the SVRHM workshop is to disseminate findings from neuroscience, psychology, and cognitive science that might inform modern machine learning methods.

Focus Areas

- Biological Inspiration and Inductive Biases
- Robustness and Generalization
- New Datasets for Comparing Humans and Machines

www.svrhm.com

— Invited Speakers —



Grace Lindsay

University College London



S.P. Arun

Indian Institute of Science



Miguel Eckstein

UC Santa Barbara



Alyosha Efros

UC Berkeley



Gamaleldin Elsayed

Google Brain



Tim Krietzmann

Donders Institute for Brain, Cognition and Behaviour



Aude Oliva

MIT



Carlos Ponce

Washington University in St. Louis



Bria Long

Stanford



Martin Hebart

Max Planck Institute for Human Cognitive and Brain Sciences

CALL FOR PAPERS

Organized by Arturo Deza, Joshua Peterson, Apurva Ratan Murty, and Thomas Griffiths

Acknowledgements

ICLR 2019

Submitted @ NeurIPS 2020



Aditya Jonnalagadda



Miguel Eckstein



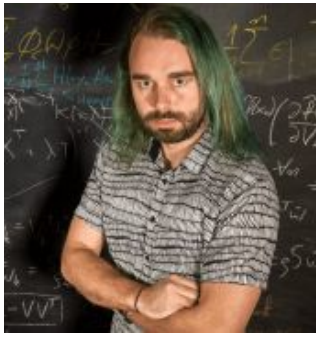
Talia Konkle



Elian Malkin
ICRA 2021?



Ronald Alvarez
CoRL 2021?



Andrzej Banburski
ICLR 2021?



Tomaso Poggio
Course 9-58!