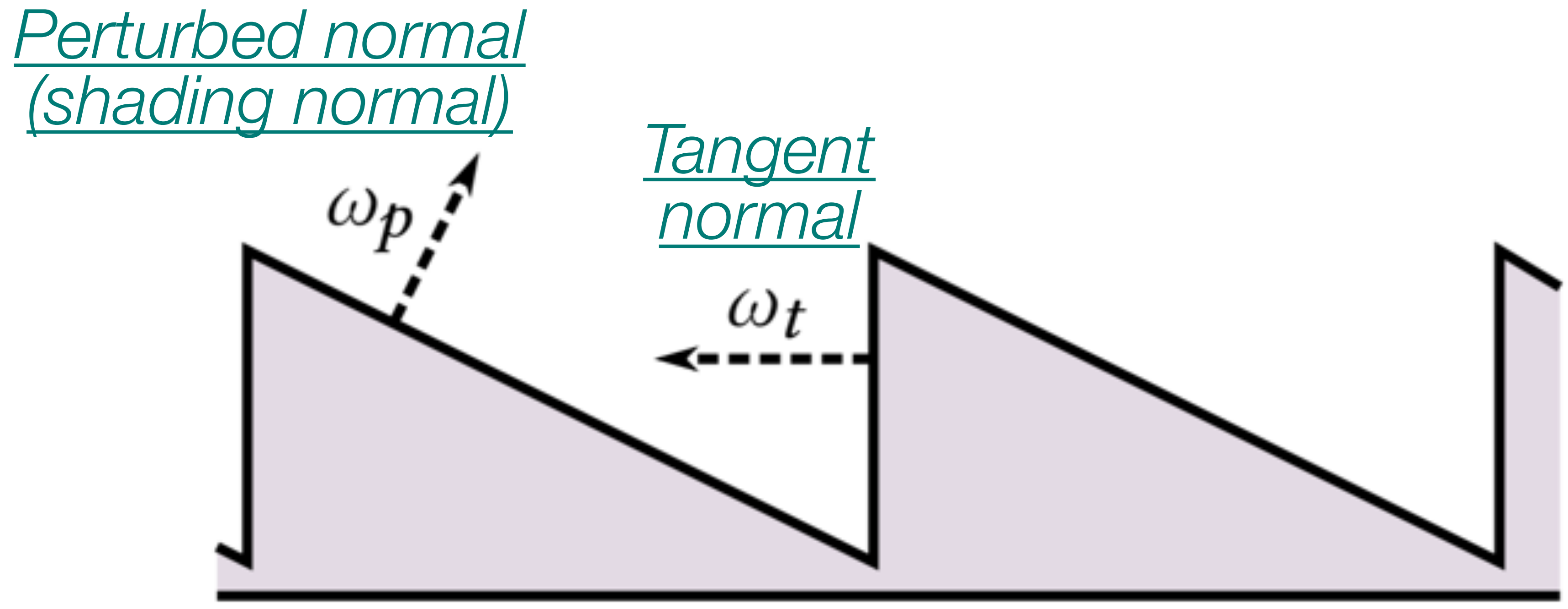
Denoising with Machine Learning

Team 5: Cheolmin Lee, Minki Jo, Nick Heppert

Modeling Microsurface

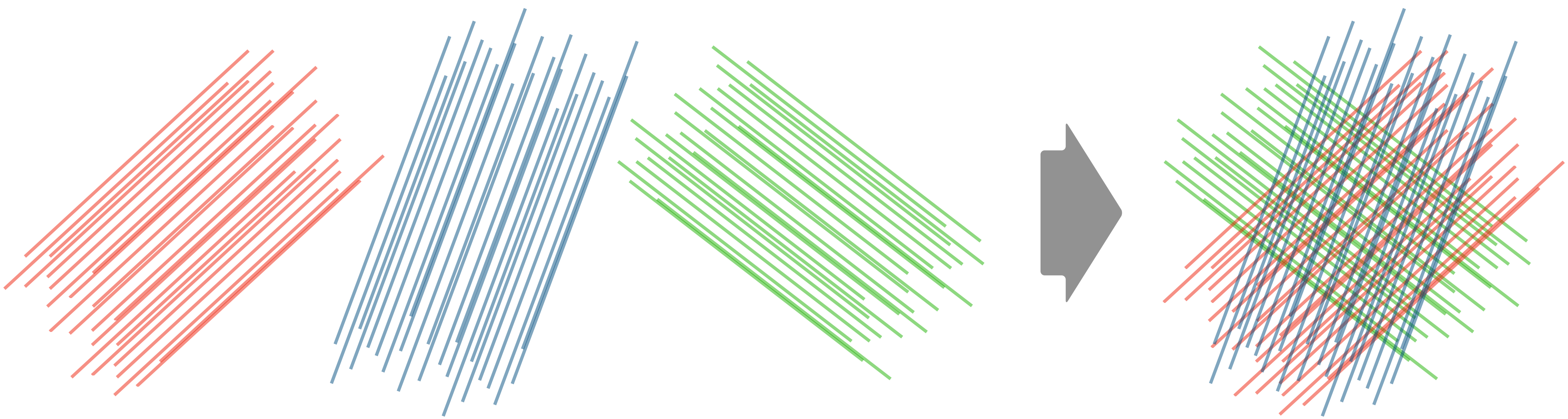
- Add tangent facet that compensates for the perturbed normal such that the average normal of the microsurface remains the geometric normal.



SVBRDF

- Compute a combination of scratch BRDFs weighted by area:

$$\bar{\rho}(\mathbf{x}, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = \sum \alpha_k(\mathbf{x}) \rho_{s,k}(\boldsymbol{\omega}_o, \boldsymbol{\omega}_i)$$

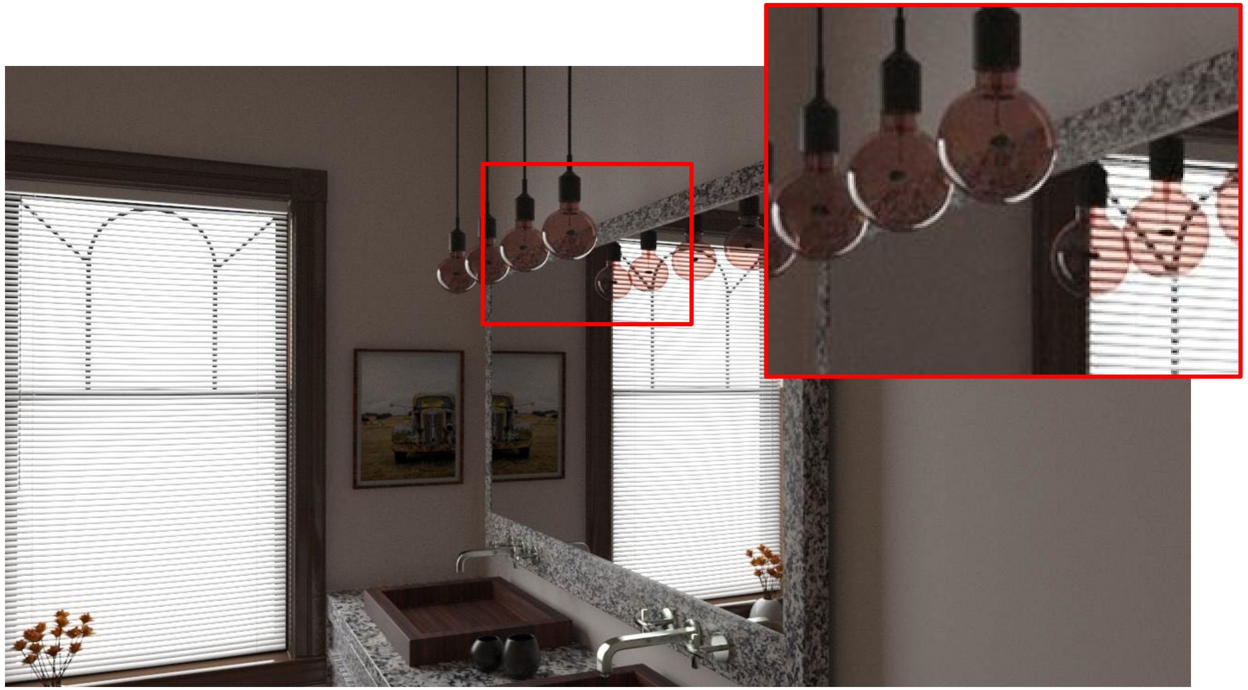


$$\rho(\mathbf{x}, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = \begin{cases} \bar{\rho}/\bar{\alpha}(\mathbf{x}) & \text{if } \bar{\alpha}(\mathbf{x}) > 1 \\ \bar{\rho} + (1 - \bar{\alpha}(\mathbf{x}))\rho_b & \text{otherwise.} \end{cases}$$

Introduction

- Creating images with high samples per pixels (spp) takes a lot of time
- Cut down time by creating low samples images → Noisy
- De-Noising techniques

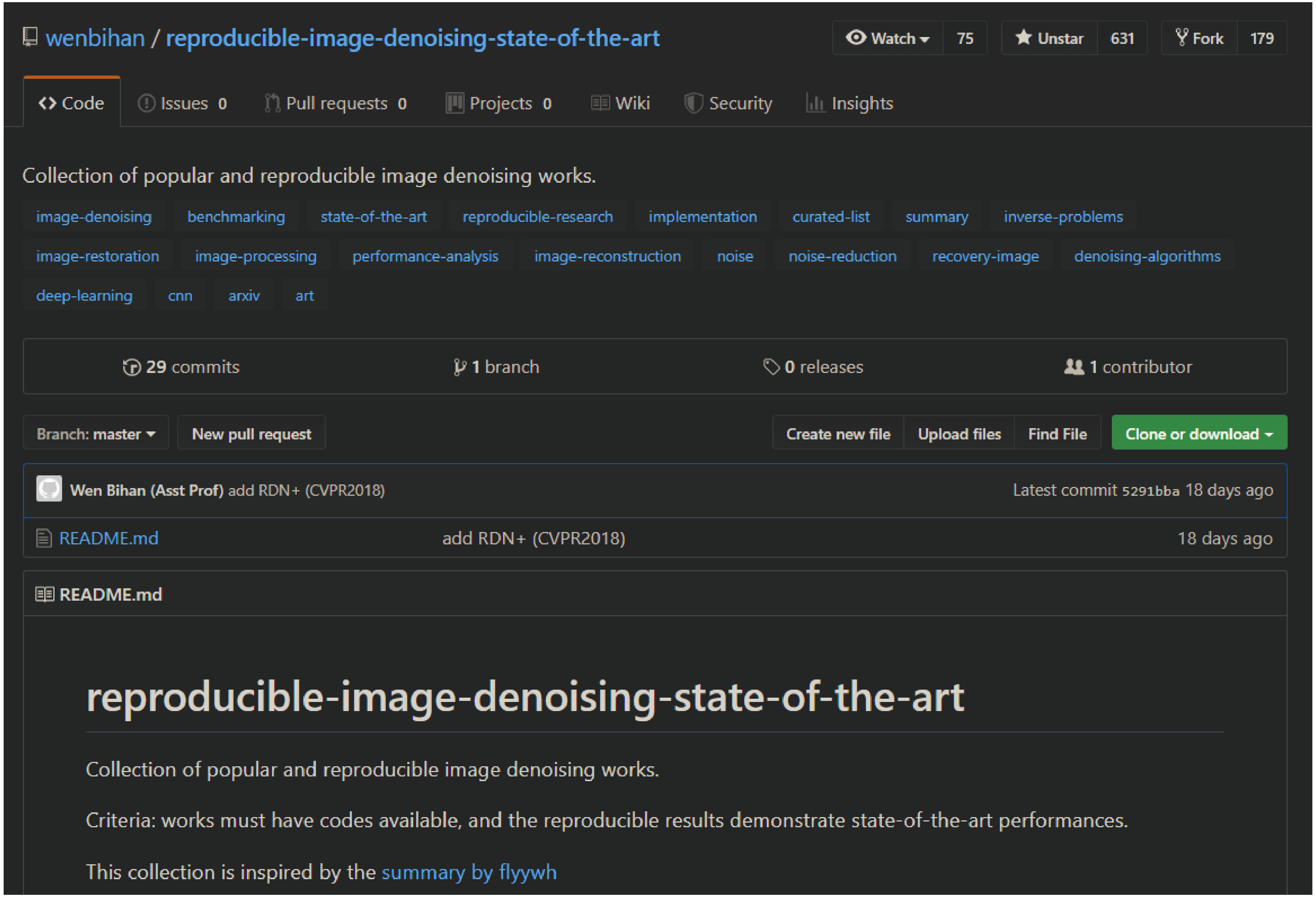
128spp



8192spp

Introduction

Why these papers?



Introduction

Why these papers?

- Current state of the art models

- CBDNet [Web] [Code] [PDF]
  - Toward Convolutional Blind Denoising of Real Photographs (Arxiv), Guo et al.
- Noise2Noise [Web] [TF Code] [Keras Unofficial Code] [PDF]
  - Noise2Noise: Learning Image Restoration without Clean Data (ICML 2018), Lehtinen et al.
- UDN [Web] [Code] [PDF]
  - Universal Denoising Networks- A Novel CNN Architecture for Image Denoising (CVPR 2018), Lefkimmiatis.
- N3 [Web] [Code] [PDF]
  - Neural Nearest Neighbors Networks (NIPS 2018), Plotz et al.
- NLRN [Web] [Code] [PDF]
  - Non-Local Recurrent Network for Image Restoration (NIPS 2018), Liu et al.
- RDN+ [Web] [Code] [PDF]
  - Residual Dense Network for Image Restoration (CVPR 2018), Zhang et al.

Sparsity and Low-rankness Combined

- STROLLR-2D [PDF] [Code]
  - When Sparsity Meets Low-Rankness: Transform Learning With Non-Local Low-Rank Constraint for Image Restoration (ICASSP 2017), Wen et al.

Combined with High-Level Tasks

- Meets High-level Tasks [PDF] [Code]
  - When Image Denoising Meets High-Level Vision Tasks: A Deep Learning Approach (IJCAI 2018), Liu et al.

INDEX

1 Introduction

2 Approach

3 Experiment & Results

Non-Local Neural Networks

NIPS 2018

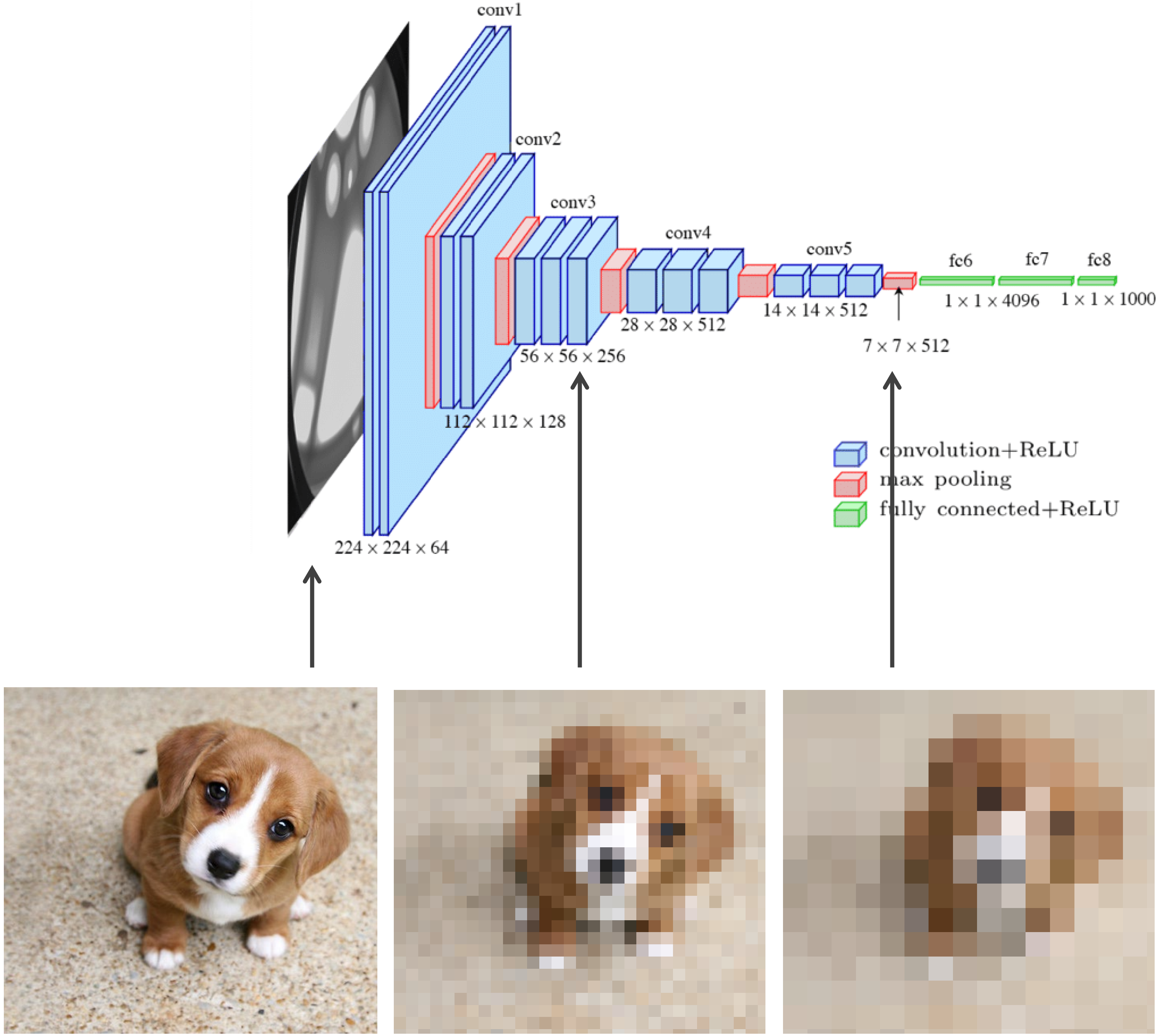
Introduction

1. Problem

CNN for Denoising

Convolutional Neural Network

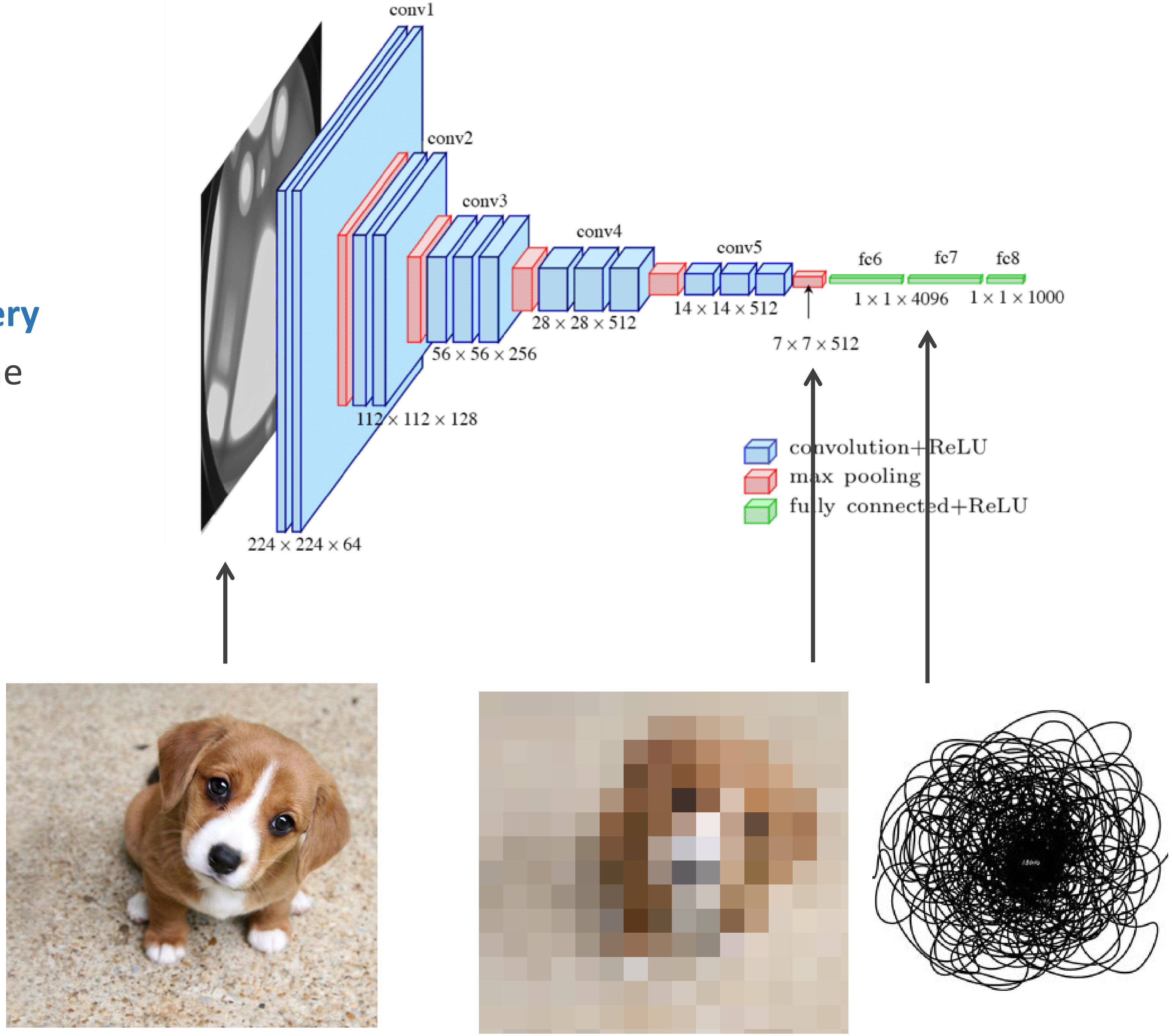
- VGG



CNN for Denoising

Convolutional Neural Network

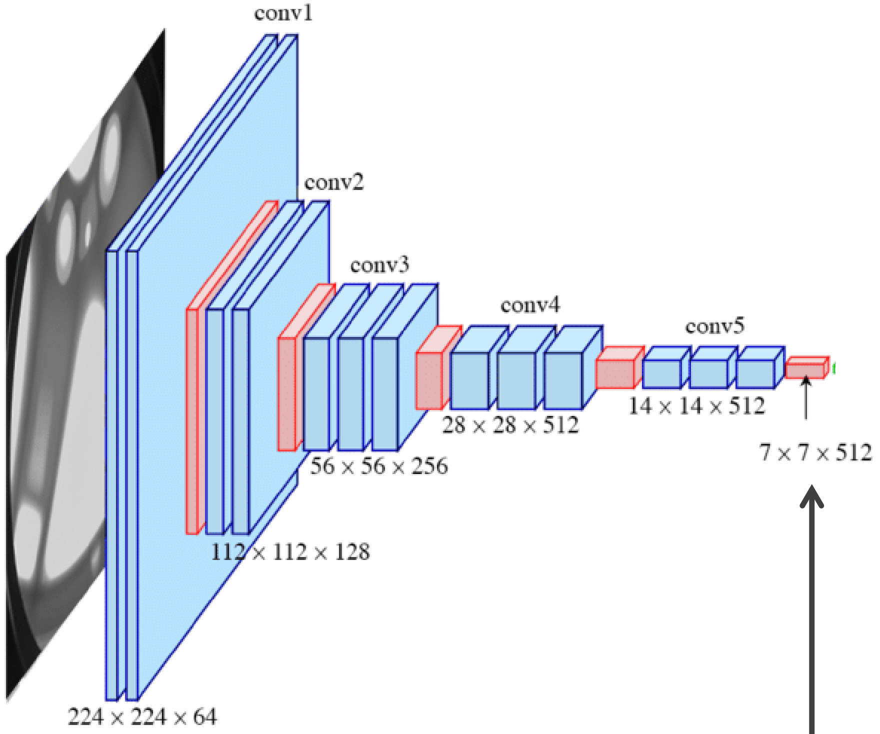
- VGG
  The FC(Fully connected) layer **lose every local feature** which is important for the image data.

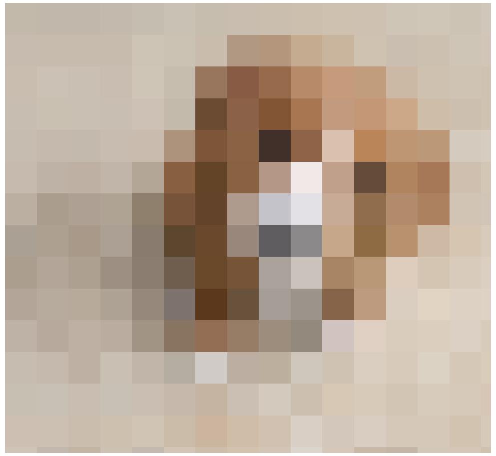


CNN for Denoising

Convolutional Neural Network

- FCN
  Fully Convolutional Network is the network that has the **convolutional layer only**.
- Since the FCN **does not lose the Local Feature**, most of the **Computer Vision tasks** has been used the FCN structure.

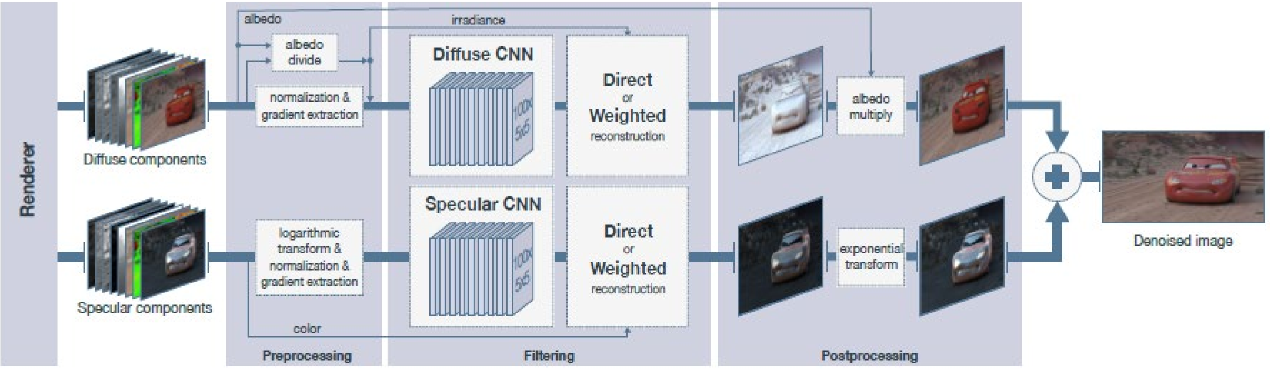




CNN for Denoising

Convolutional Neural Network

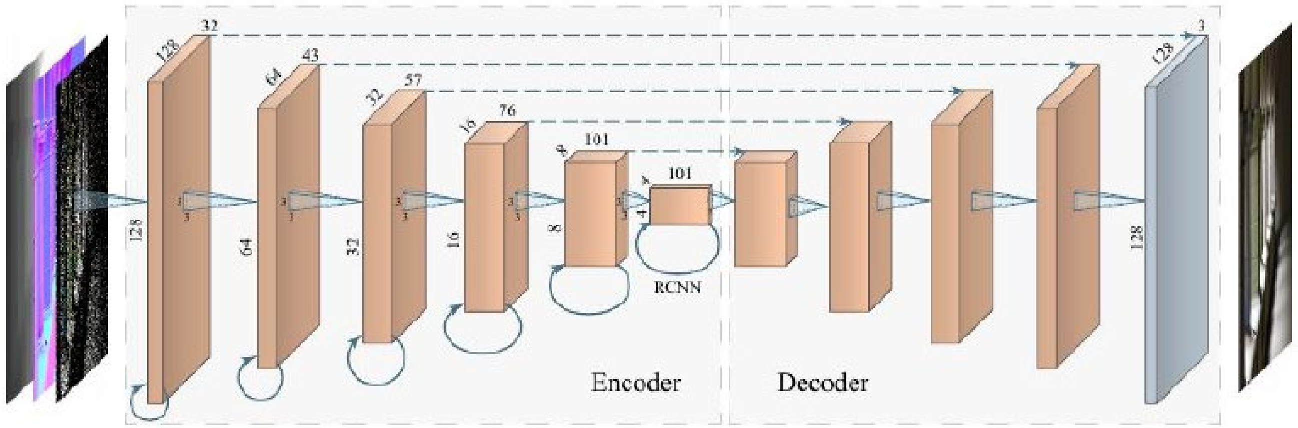
- Many denoising models such as KPCN and RDA use the FCN



focus on the filtering core of the denoiser—the network architecture and the reconstruction filter—and later describe data decomposition and preprocessing that are specific to the problem of MC denoising.

4.1 Network Architecture

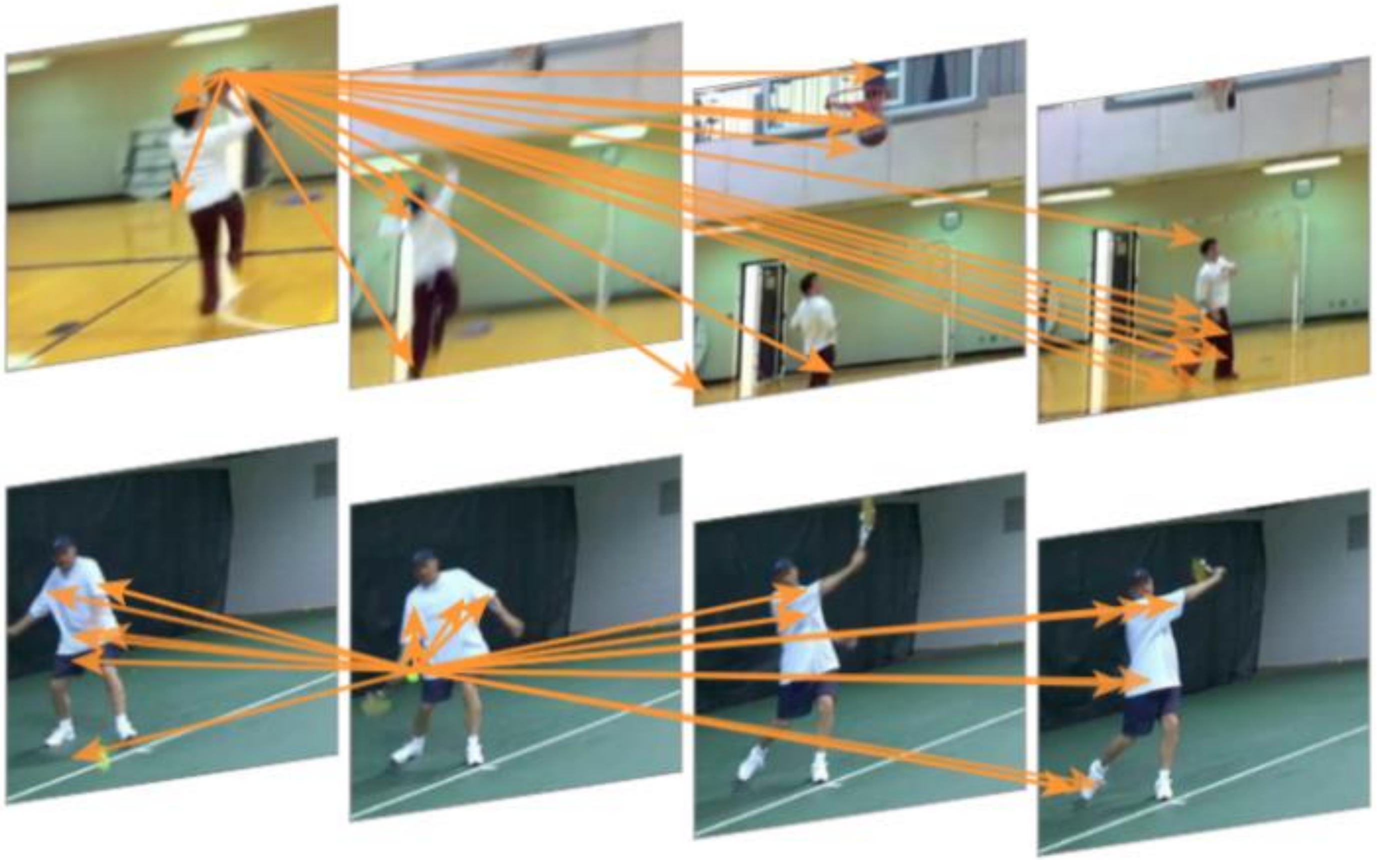
We use deep fully convolutional networks with no fully-connected layers to keep the number of parameters reasonably low. This reduces the danger of overfitting and speeds up both training and inference. Stacking many convolutional layers together effectively



Discussion

Recalibrate Features?

- Global Representation
- Global Context
- Long-range Dependencies
- Shorter Paths

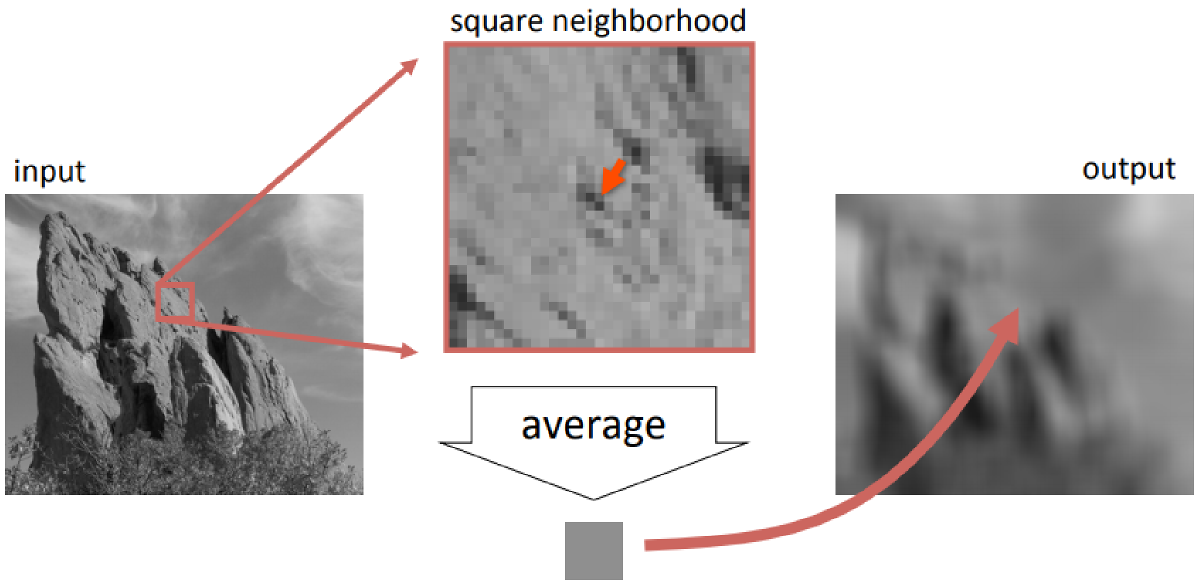


Approach

1. Motivation
2. Approach

Problem: Denoising

Solution: Smoothing
Box Filter



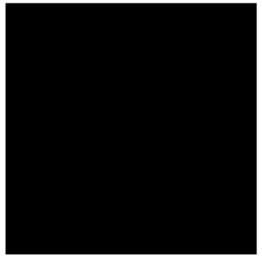
$$\boldsymbol{BA[I]_{\mathbf{p}}} = \sum_{\mathbf{q}\in S} B_\sigma(\mathbf{p}-\mathbf{q})\, I_{\mathbf{q}}$$

result at pixel **p**

sum over all pixels **q**

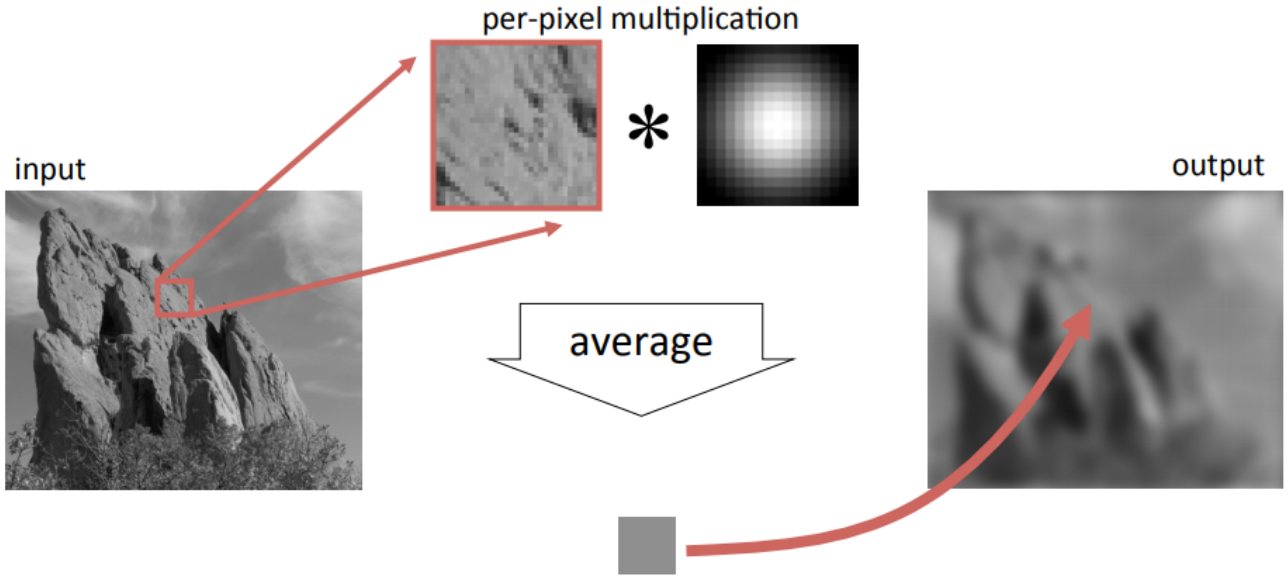
intensity at pixel **q**

normalized box function

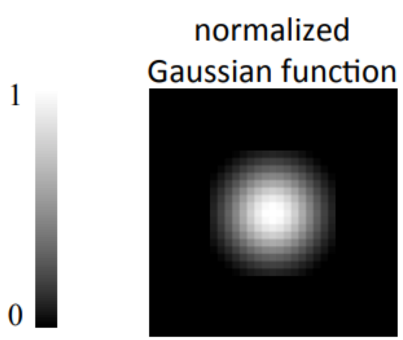


Slides from BIL 717 Image Processing 2012

Gaussian Filter

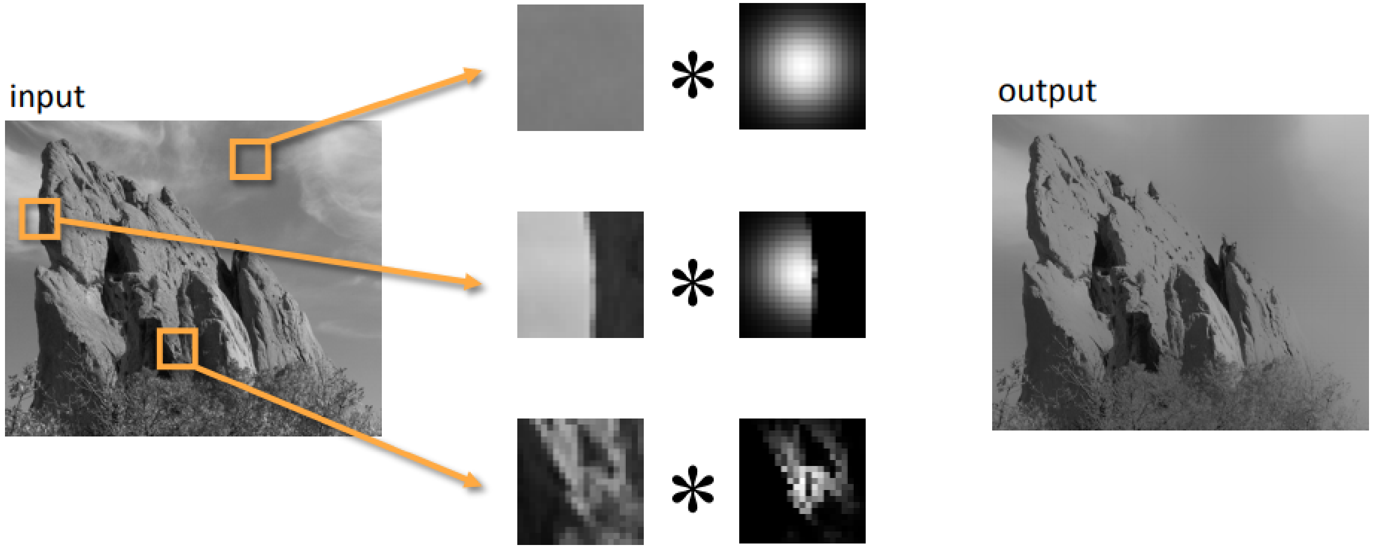


$$GB[I]_{\mathbf{p}} = \sum_{\mathbf{q} \in S} \boxed{G_\sigma(\| \mathbf{p} - \mathbf{q} \|)} I_{\mathbf{q}}$$



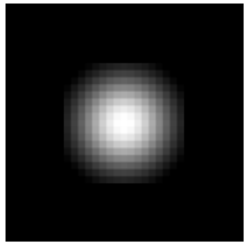
Slides from BIL 717 Image Processing 2012

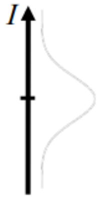
Bilateral Filter



The kernel shape depends on the image content.

$$BF[I]_{\mathbf{p}} = \underbrace{\frac{1}{W_{\mathbf{p}}}}_{\text{normalization factor}} \sum_{\mathbf{q}\in S} \overbrace{\underbrace{G_{\sigma_s}(\|\mathbf{p}-\mathbf{q}\|)}_{\textbf{\textit{space}}\text{ weight}}}^{\text{not new}} \overbrace{\underbrace{G_{\sigma_r}(|I_{\mathbf{p}}-I_{\mathbf{q}}|)}_{\textbf{\textit{range}}\text{ weight}}}^{\text{new}} I_{\mathbf{q}}$$





Slides from BIL 717 Image Processing 2012

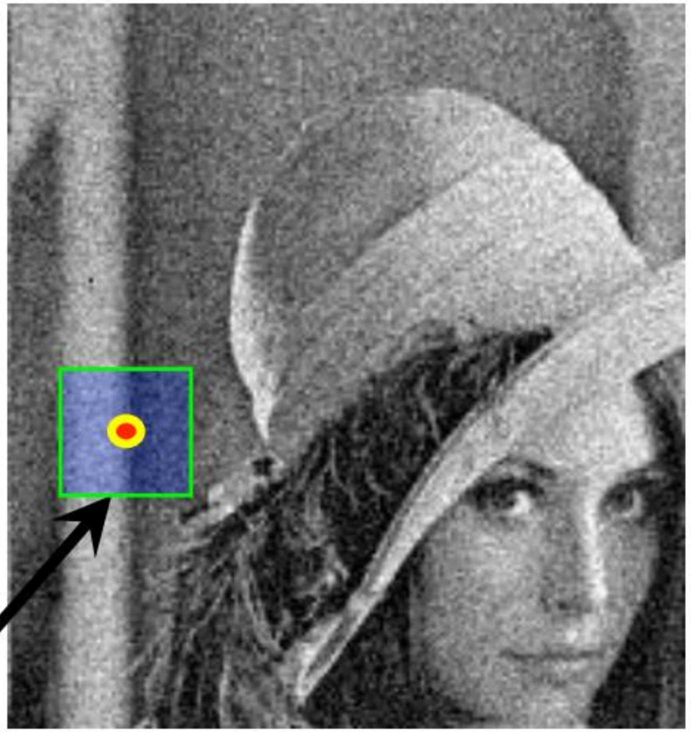
Non-local Filter

- Average Similar Pixels
- Do not Average non-Similar Pixels

Problem)
Not Enough Similar Pixels in LOCAL REGIONS
→ Get More Samples in Non-LOCAL REGIONS

Non-local Filter

NL-Means Method:
Buades (2005)



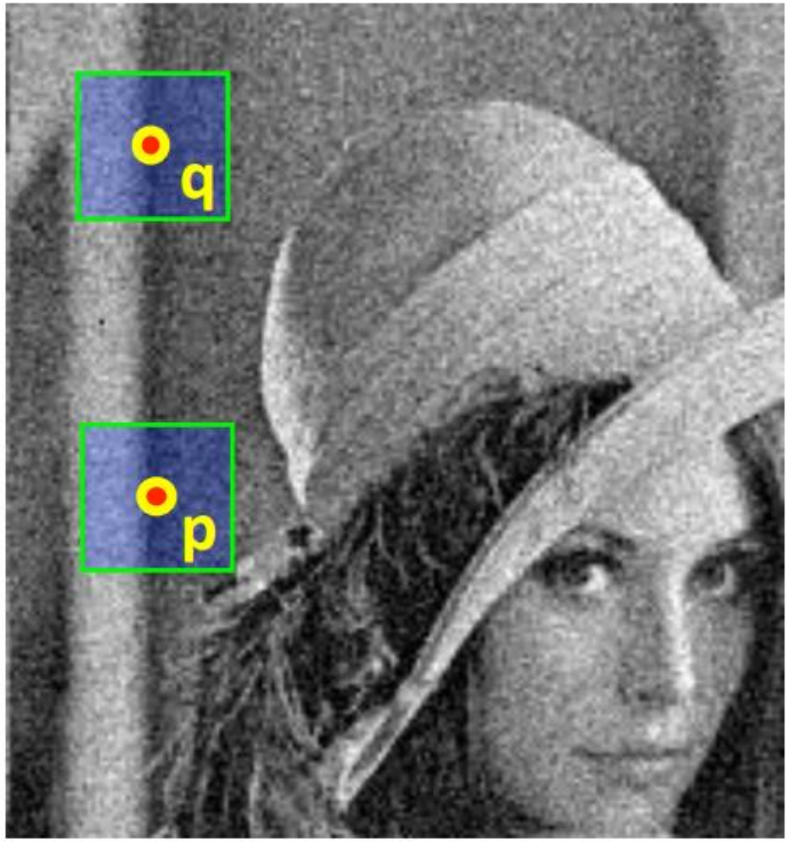
- For each and every pixel **p:**
  - Define a small, simple fixed size neighborhood;

Non-local Filter

NL-Means Method:
Buades (2005)

<u>'Similar'</u> pixels **p, q**
→ **SMALL**
vector distance;

$$|| V_p - V_q ||^2$$

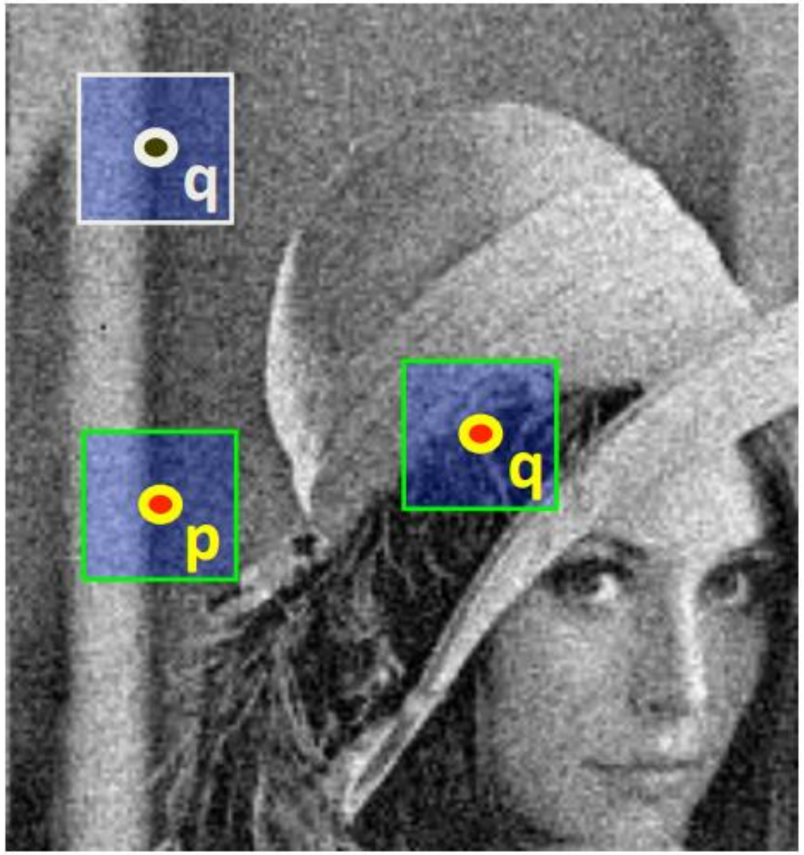


Non-local Filter

NL-Means Method:
Buades (2005)

'Dissimilar' pixels **p, q**
→ **LARGE**
vector distance;

$$|| V_p - V_q ||^2$$



Slides from BIL 717 Image Processing 2012

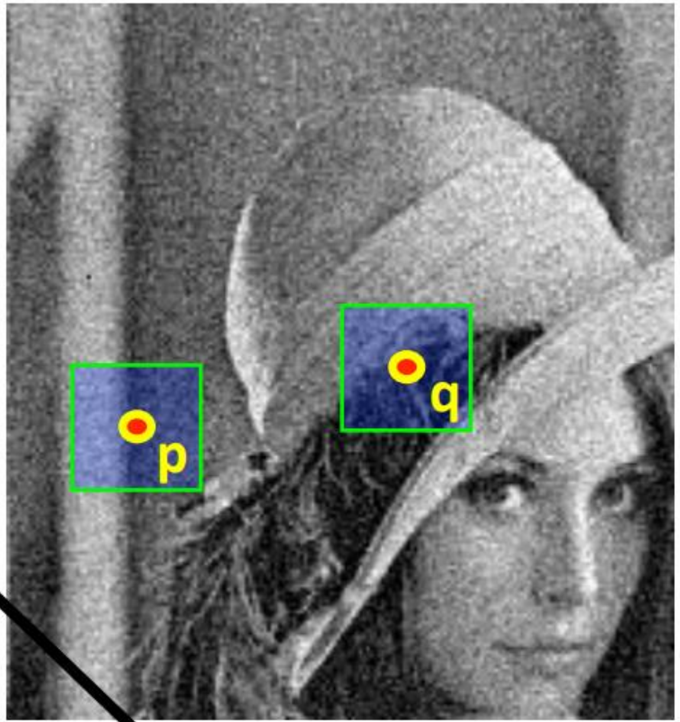
Non-local Filter

NL-Means Method:
Buades (2005)

p, q neighbors define a vector distance;

$\boxed{\| V_p - V_q \|^2}$

Filter with this:
No spatial term!



$$NLMF[I]_{\mathbf{p}} = \frac{1}{W_{\mathbf{p}}} \sum_{\mathbf{q}\in S} \cancel{G_{\sigma_s}(\|\mathbf{p}-\mathbf{q}\|)} G_{\sigma_r}\left(\| \vec{V}_{\mathbf{p}} - \vec{V}_{\mathbf{q}} \|^2\right) I_{\mathbf{q}}$$

$$BA[I]_p = \frac{1}{W} \sum_q I_q$$

$$G[I]_p = \frac{1}{W} \sum_q G_\sigma(\|p-q\|_2) I_q$$

$$G[I]_p = \frac{1}{W} \sum_q G_\sigma(\|p-q\|_2) G_{\sigma_r}\left(\|I_p - I_q\|_1\right) I_q$$

$$NLMF[I]_p = \frac{1}{W} \sum_q G_\sigma\left(\|V_p - V_q\|_2\right) I_q$$

Slides from BIL 717 Image Processing 2012

Non-local Filter

Inputs

$$\boxed{NLMF[I]_p} = \frac{1}{W}\sum_q \boxed{G_\sigma\left(\left\|V_p - V_q\right\|_2\right)}\boxed{I_q}$$

Output Value

Representation
(Probability Distribution)

Target Value (Pixel)
vs All Values (Pixel)

Non-local Operation

Inputs

$$\boxed{y_i} = \frac{1}{C(x)} \sum_j \boxed{f(x_i, x_j)} \boxed{g(x_j)}$$

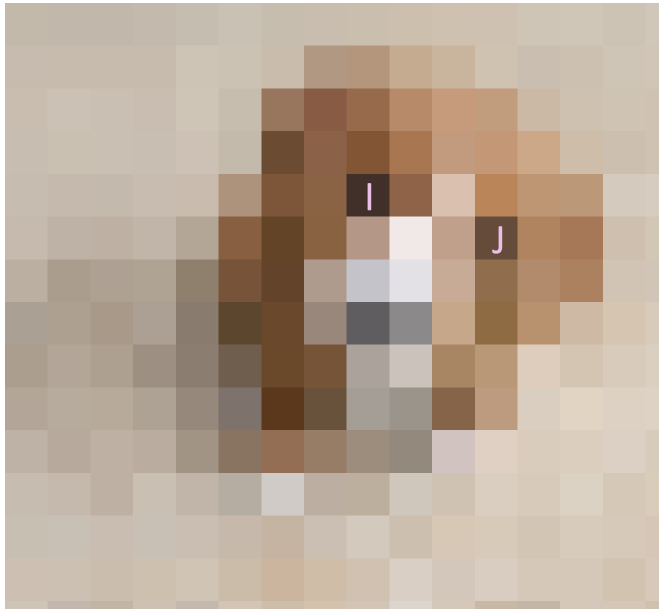
Output Value

Representation
(Probability Distribution)

Target Value (Pixel)
vs All Values (Pixel)

NON-Local Layer

$$y_i = \frac{1}{C(x)} \sum_j \boxed{f(x_i, x_j)} g(x_j)$$



Another Representation of Non-Local Pixels

= Weighted Sum of All Pixels with Similarity

+Learning···

Similarity

$$y_i = \frac{1}{C(x)} \sum_j \boxed{f(x_i, x_j)} g(x_j)$$

- Gaussian $\quad f(x_i, x_j) = \exp(x_i^T \cdot x_j)$
- Embedded Gaussian $\quad f(x_i, x_j) = \exp(\theta(x_i^T) \cdot \phi(x_j))$
- Dot Product $\quad f(x_i, x_j) = \theta(x_i^T) \cdot \phi(x_j)$
- Concatenation $\quad f(x_i, x_j) = ReLU(w_f^T [\theta(x_i) \cdot \phi(x_j)])$

Input Representation

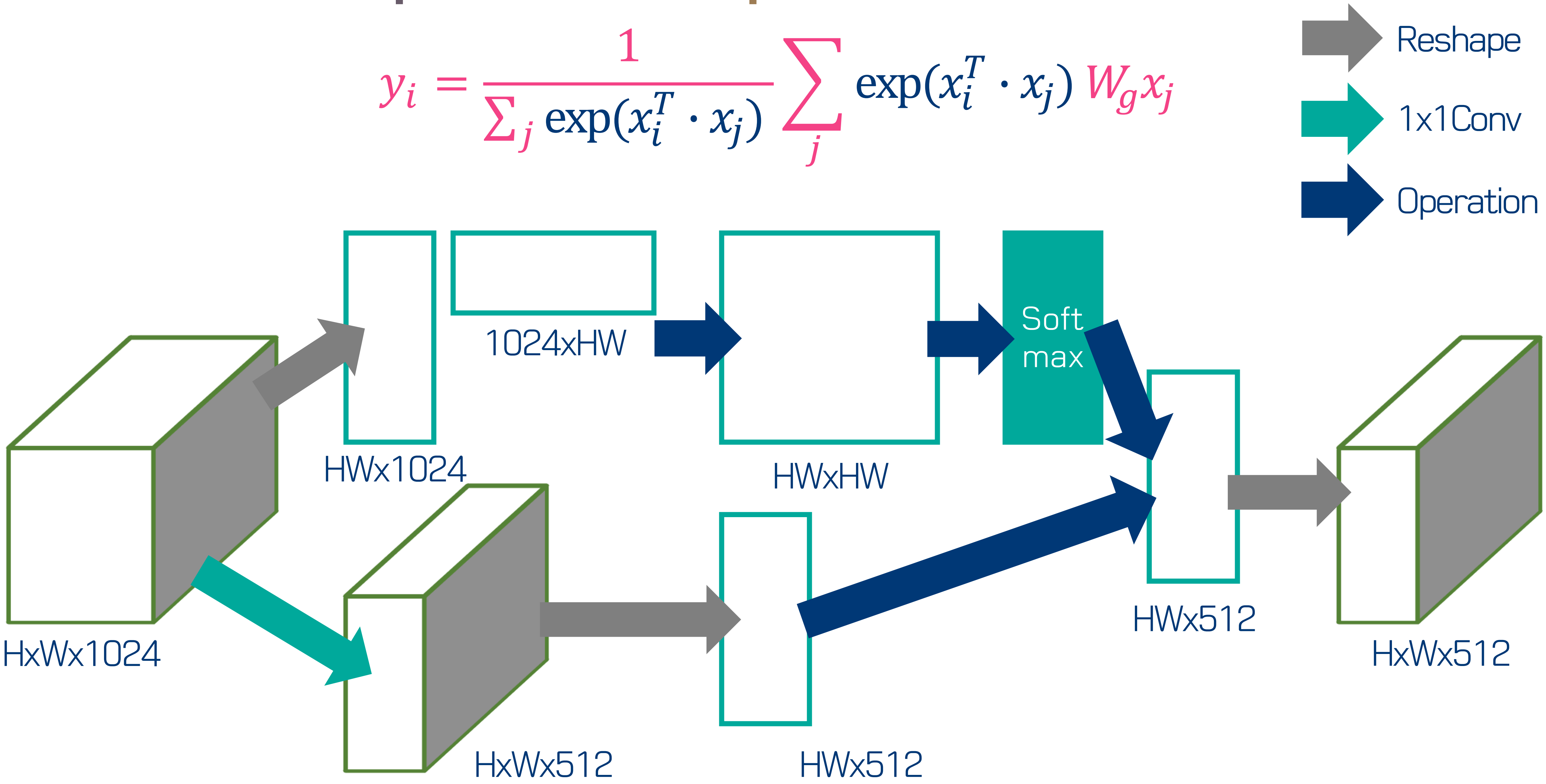
For Feature Extraction

$$y_i = \frac{1}{C(x)} \sum_j f(x_i, x_j) \boxed{g(x_j)}$$

$$g(x_j) = W_g x_j$$

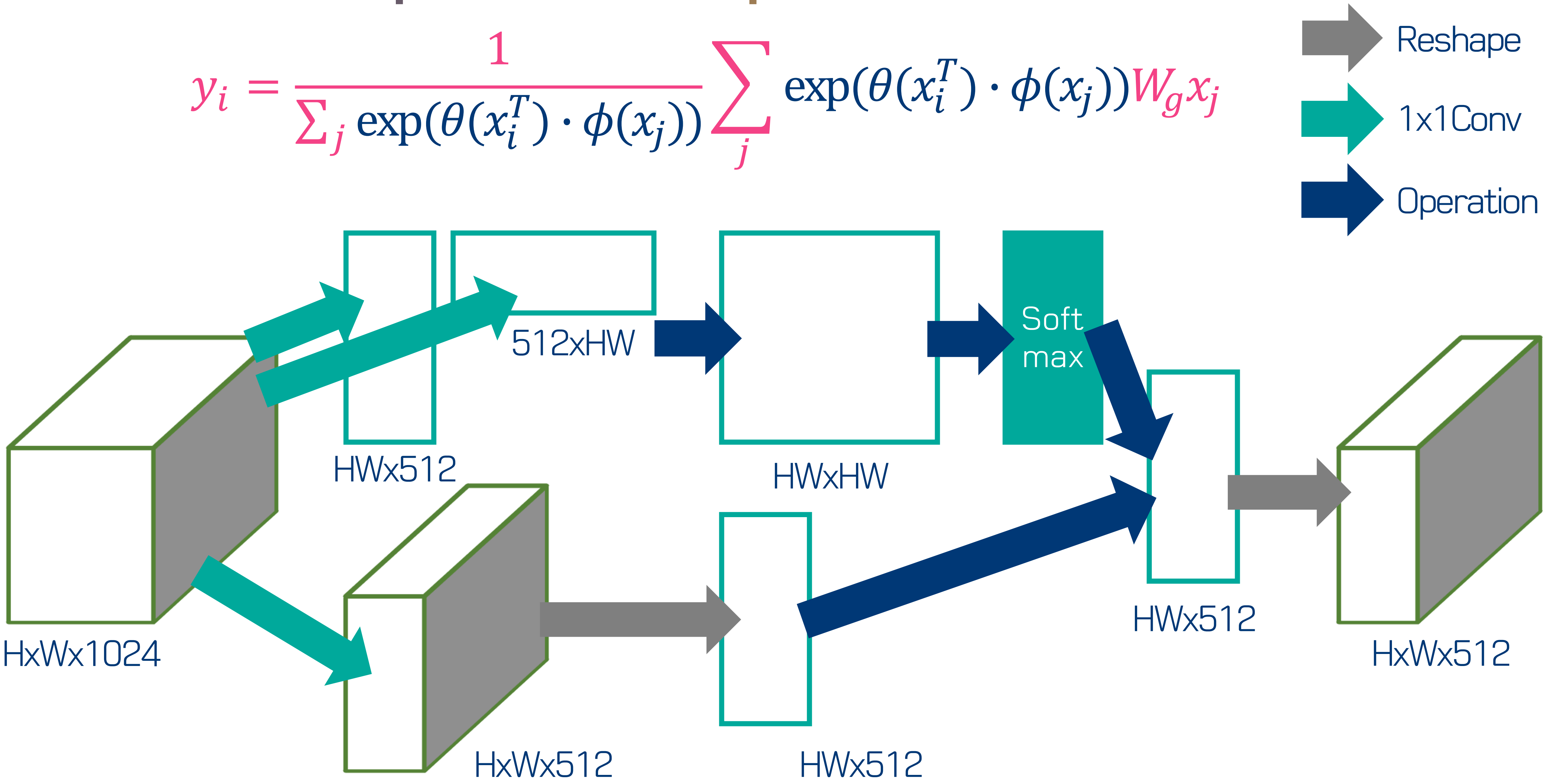
Non-local Operation Implementation

$$y_i = \frac{1}{\sum_j \exp(x_i^T \cdot x_j)} \sum_j \exp(x_i^T \cdot x_j)\, W_g x_j$$



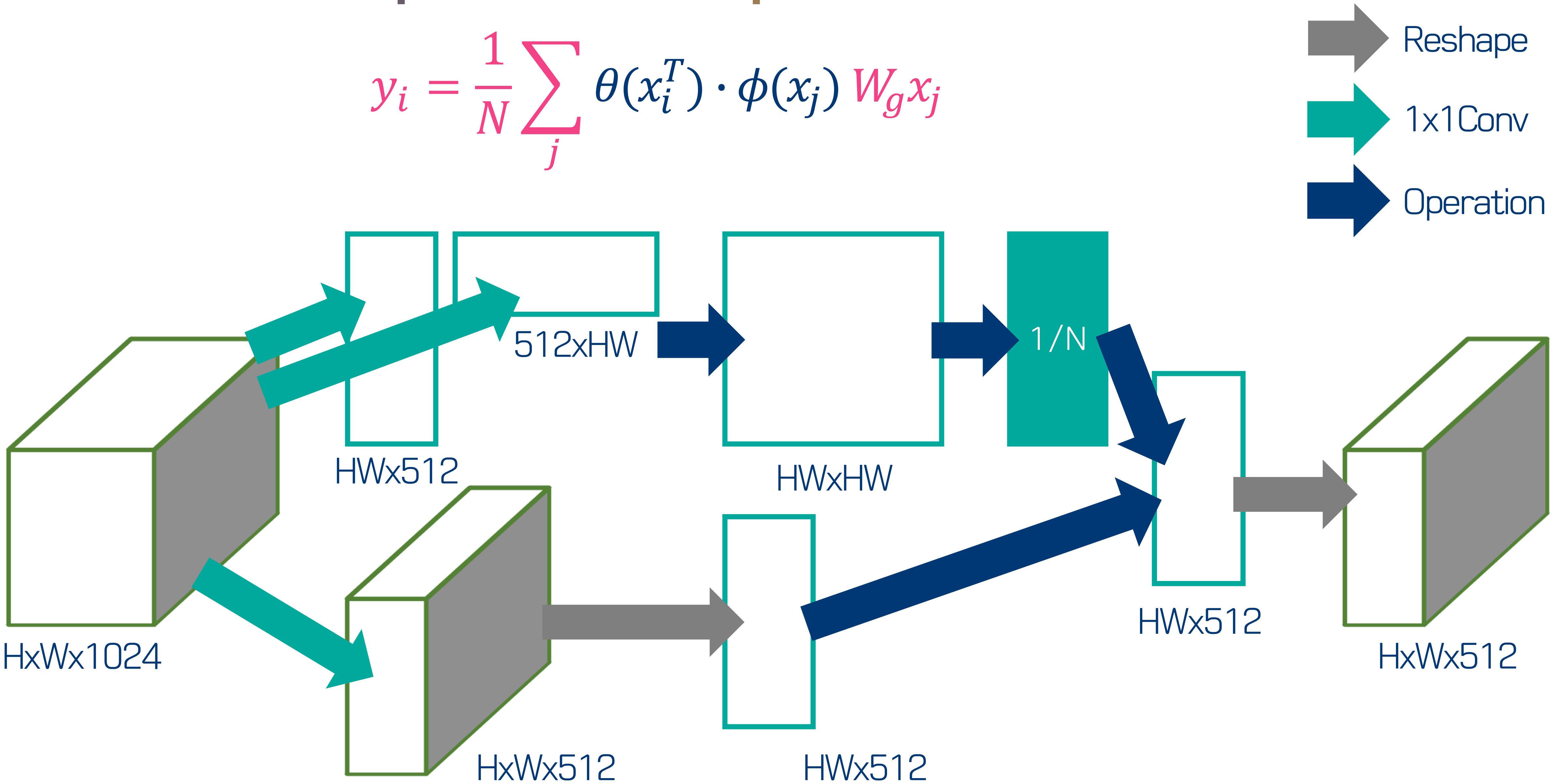
Non-local Operation Implementation

$$y_i = \frac{1}{\sum_j \exp(\theta(x_i^T) \cdot \phi(x_j))} \sum_j \exp(\theta(x_i^T) \cdot \phi(x_j)) W_g x_j$$



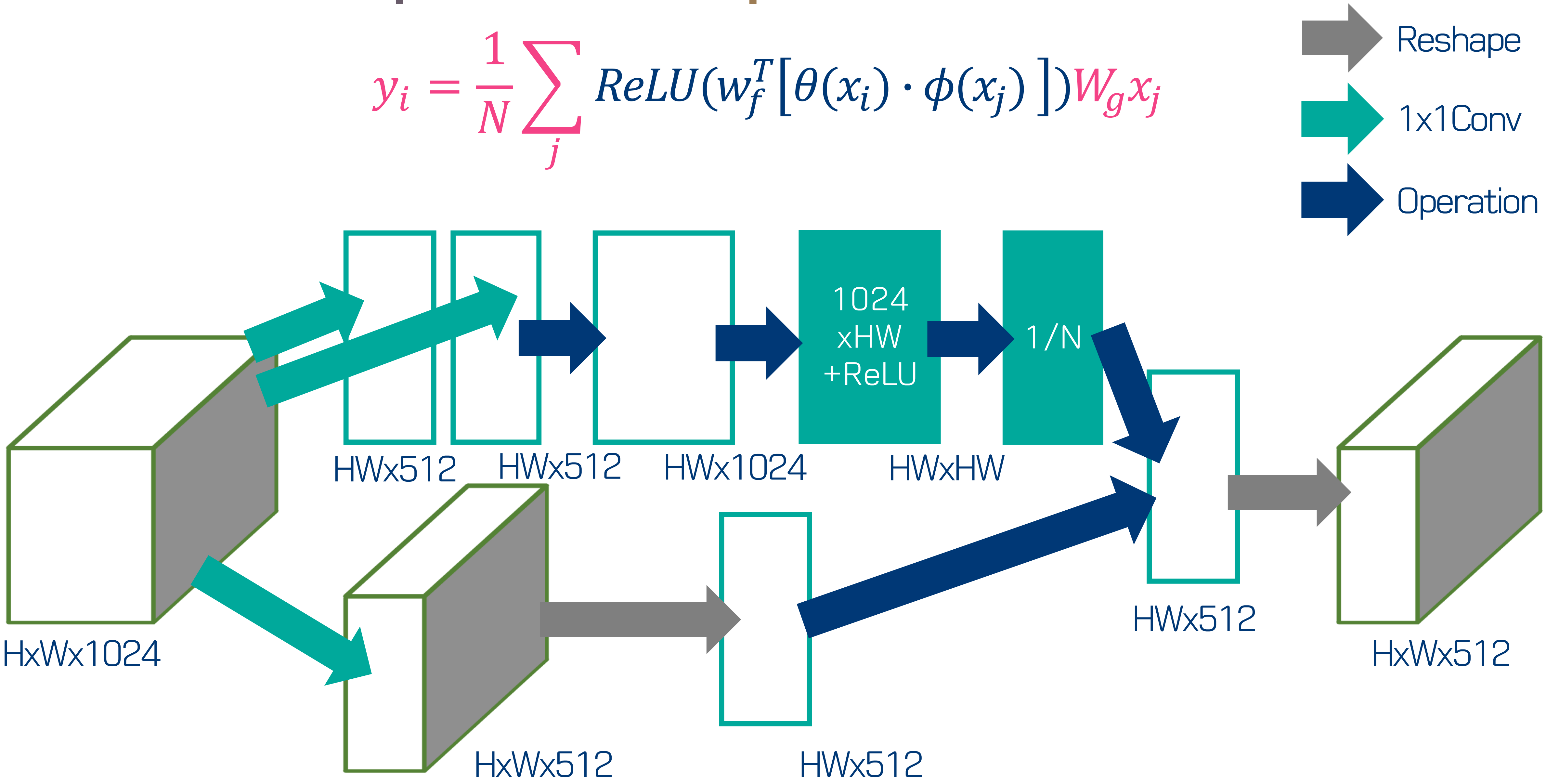
Non-local Operation Implementation

$$y_i = \frac{1}{N}\sum_j \theta(x_i^T)\cdot\phi(x_j)\, W_g x_j$$



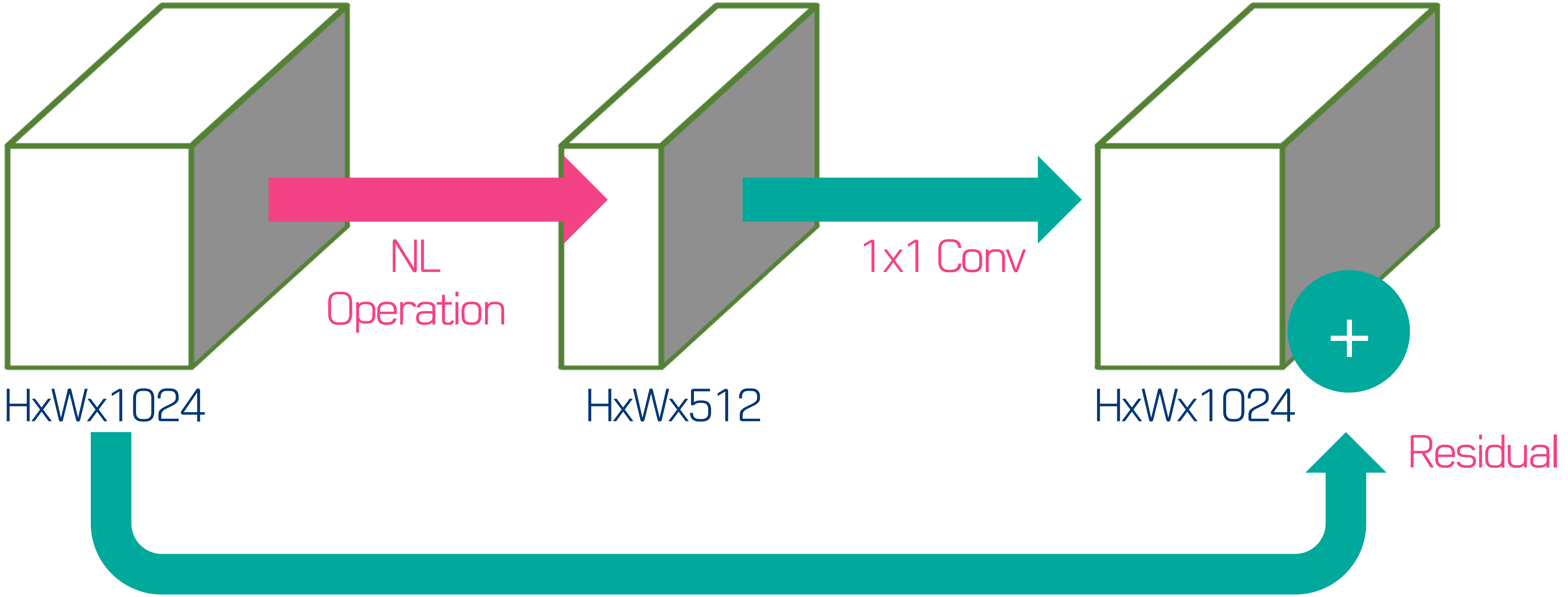
Non-local Operation Implementation

$$y_i = \frac{1}{N}\sum_j ReLU(w_f^T[\theta(x_i)\cdot\phi(x_j)])W_g x_j$$

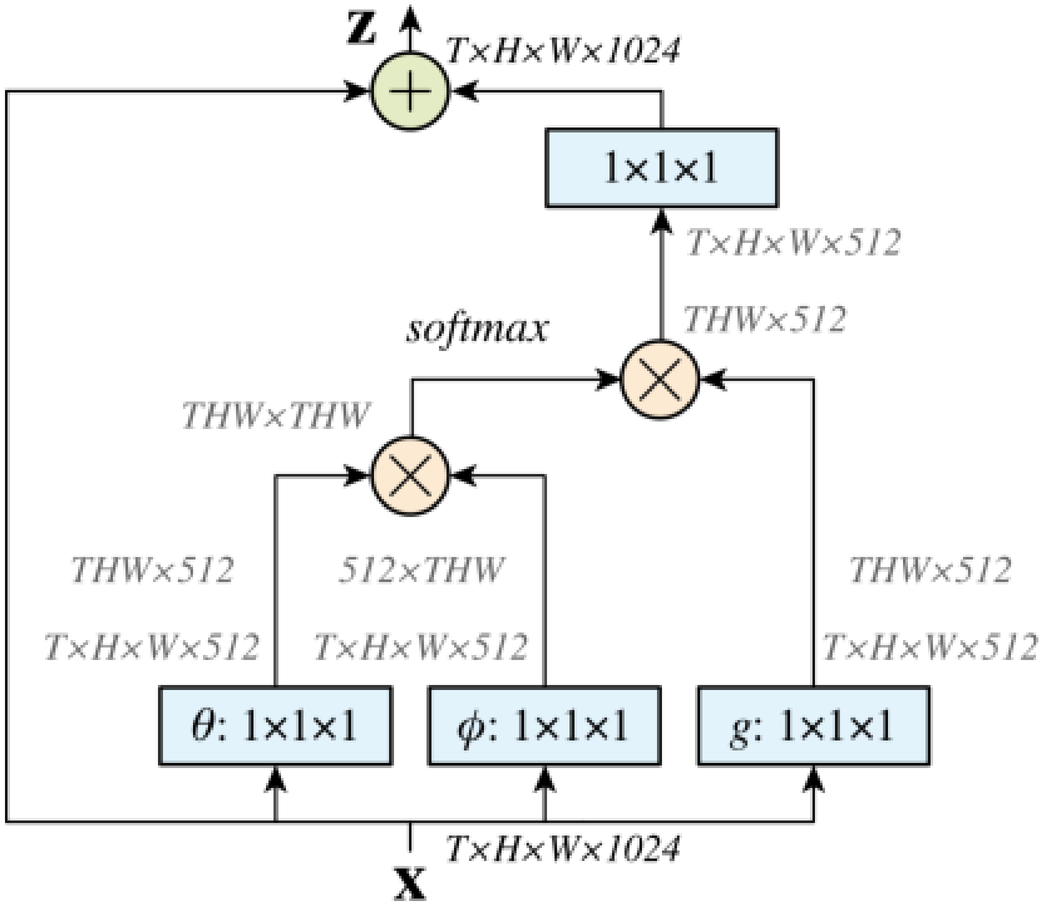


Non-local Block

$$z_i = W_z y_i + x_i$$



in Paper (Video case)



NON-Local Layer

$$y_i = \frac{1}{C(x)} \sum_j \boxed{f(x_i, x_j)}\, g(x_j)$$

Another Representation of Non-Local Pixels
= Weighted Sum of All Pixels with Similarity
+ Learning?

Experiments

1. Experiments

Noise2Noise: Learning Image Restoration without Clean Data

ICML 2018

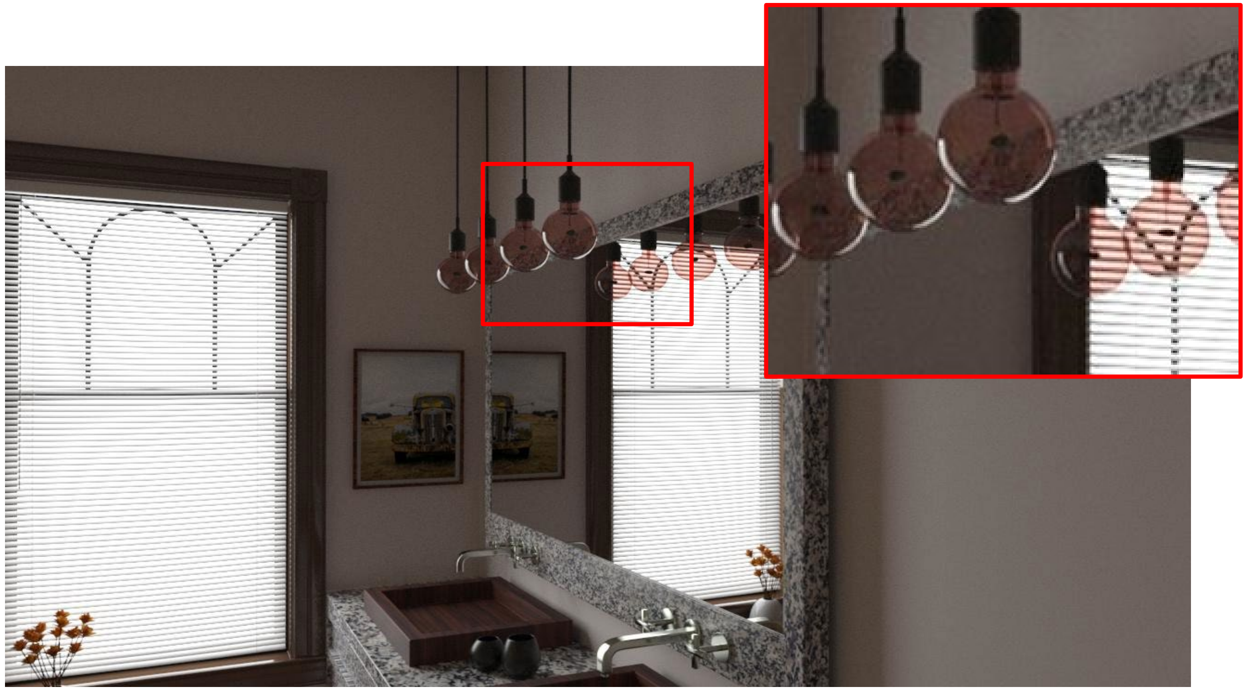
Introduction

1. Problem

Introduction – Problem

- Creating images with high samples per pixels (spp) takes a lot of time
- Cut down time by creating low samples images → Noisy
- De-Noising techniques

128spp



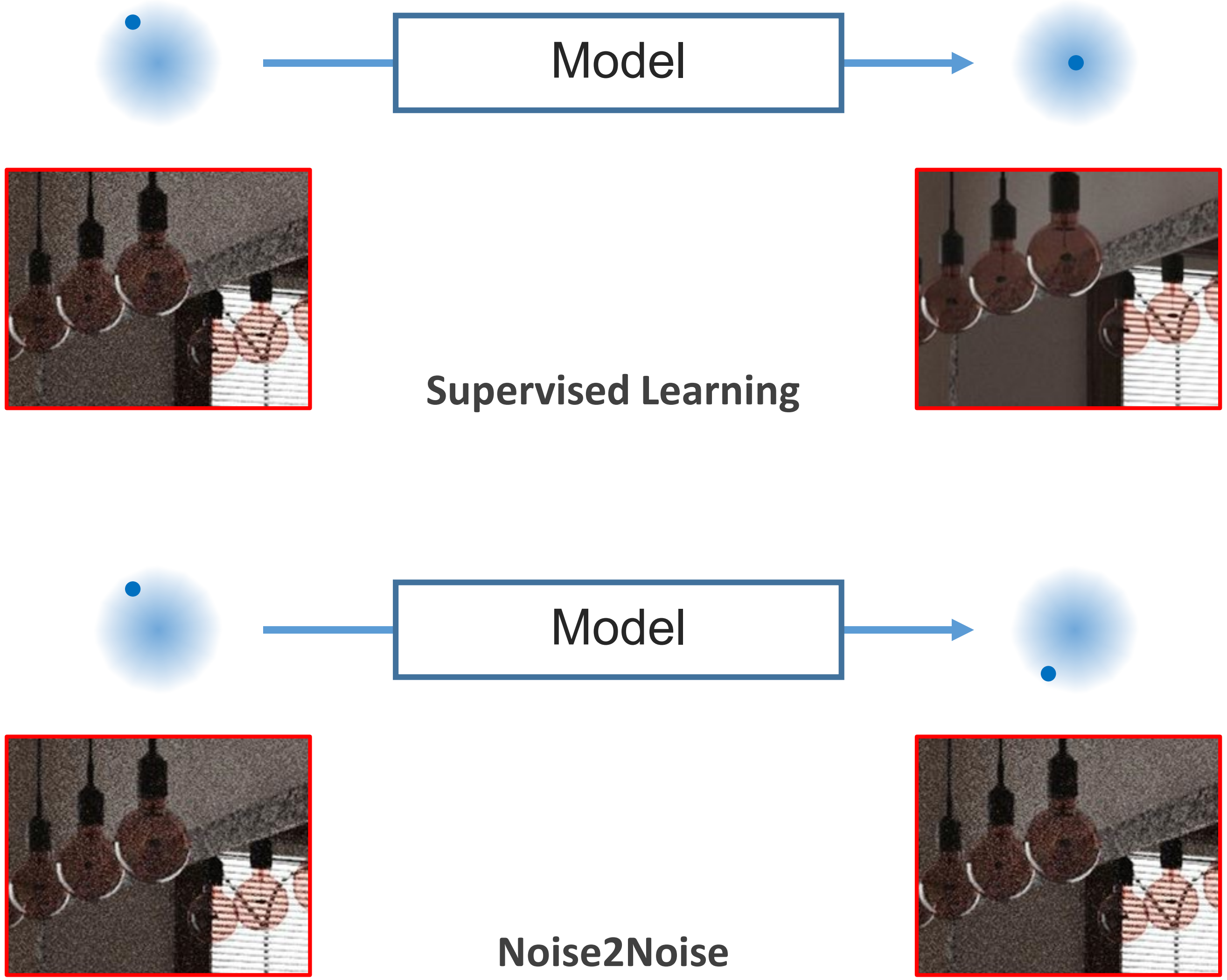
8192spp

Additional Approach

Additional Approach

- Noise2Noise
  N2N is the current state of the art model for the single RGB denoising problem.

- We will try to merge N2N and KPCN model if we have enough time.

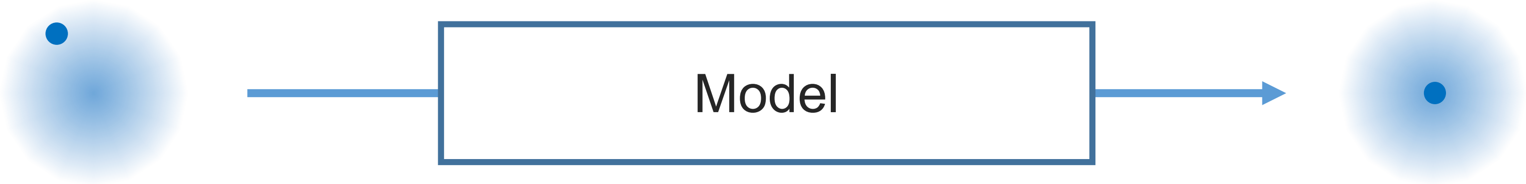


[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Problem

Current Denoising

- Current models take the noisy input and learns to produce the clean target.



Noisy Image

Supervised Learning

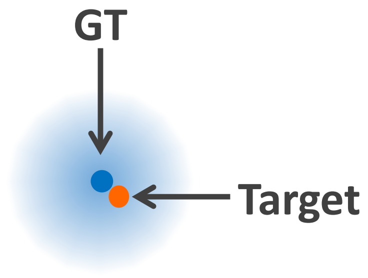
Clean Image

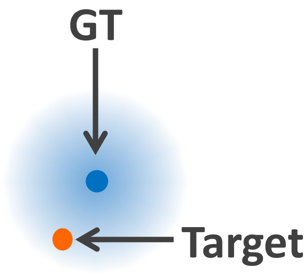
[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Problem

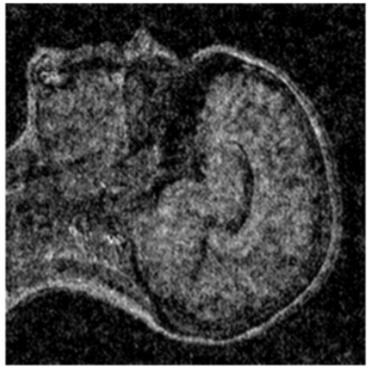
Current Denoising

- However, in some cases, getting a clean image target with zero noise (Ground Truth) is impossible.
- Medical image such as MRI scan, Montecarlo rendering image are one of those cases.





Montecarlo Rendering



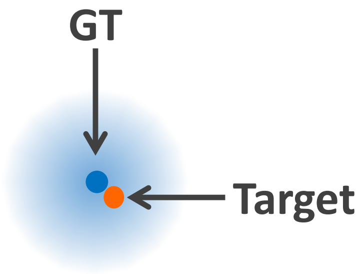
MRI scan

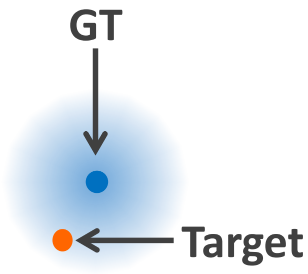
[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Problem

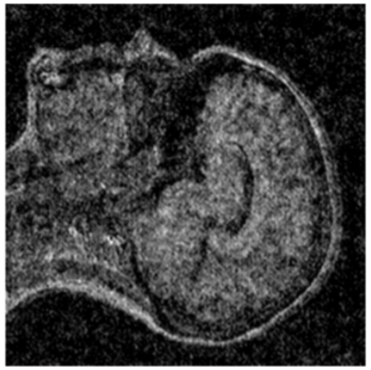
Current Denoising

- In these cases, normal supervised learning method is not the best because the target itself is noisy.





Montecarlo Rendering



MRI scan

[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Approach

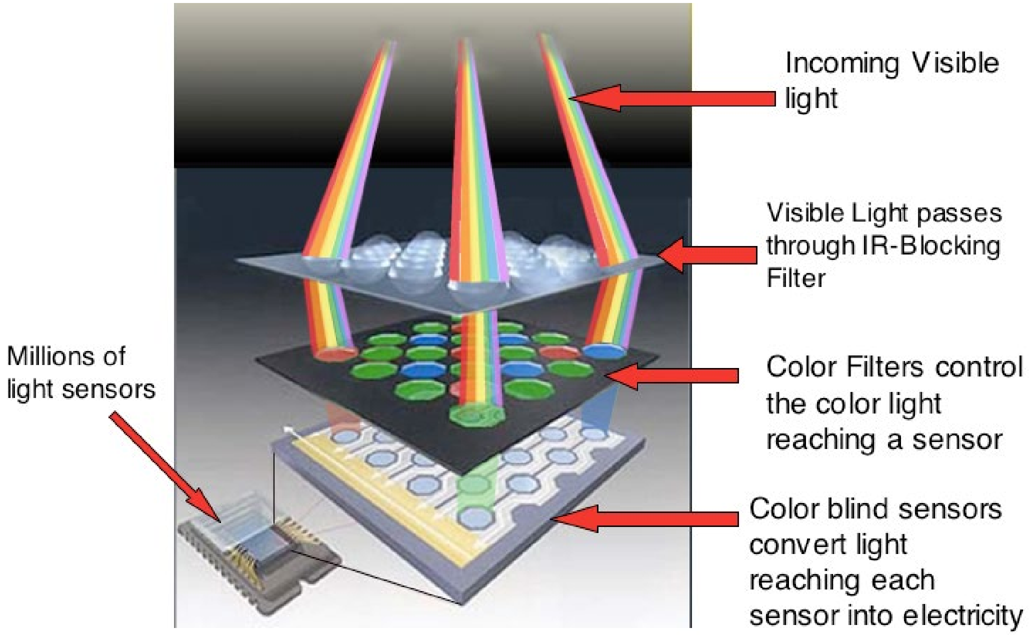
1. Motivation
2. Approach

Motivation

Current Denoising

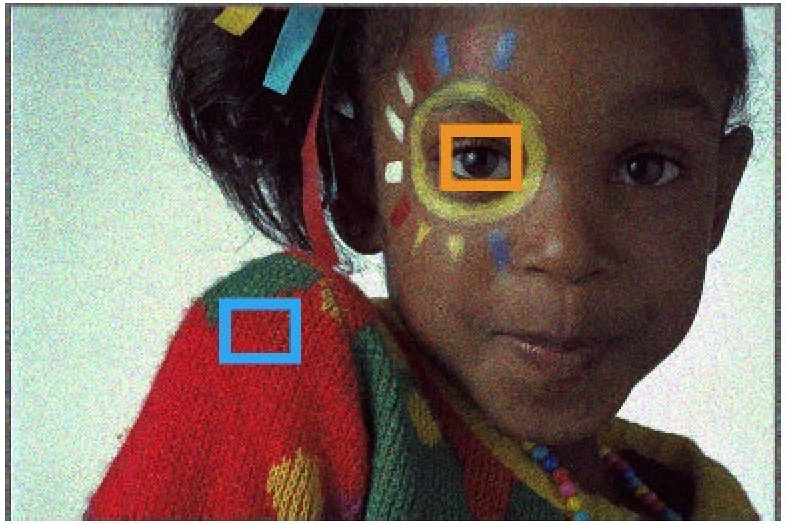
- How the RGB camera get clean image?

Motivation

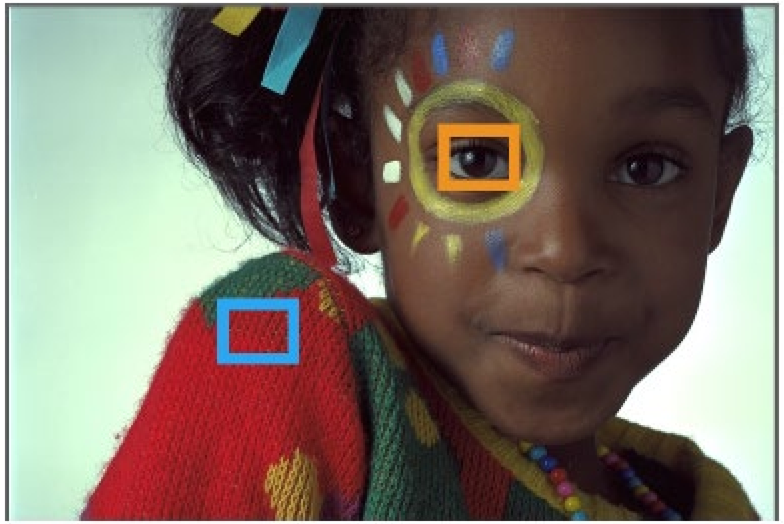


Current Denoising

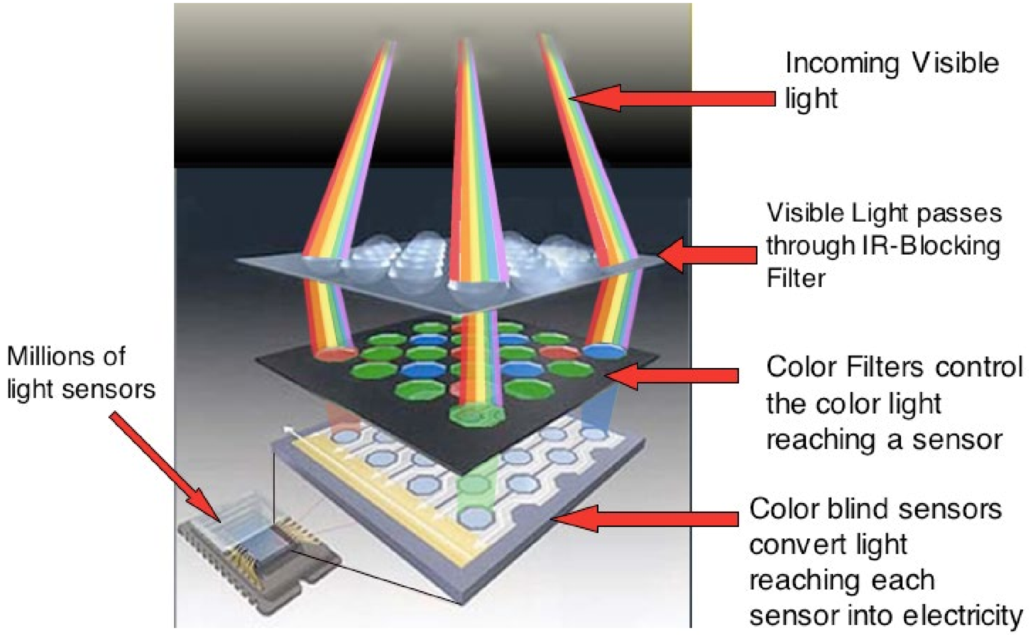
- If the camera sensor shot only one time, the image must be noisy too.
- However, the camera takes many shot during the exposure time, and take average of the color value after the filtering. In this way, we can remove the noise of the image.



x 100 =



Motivation

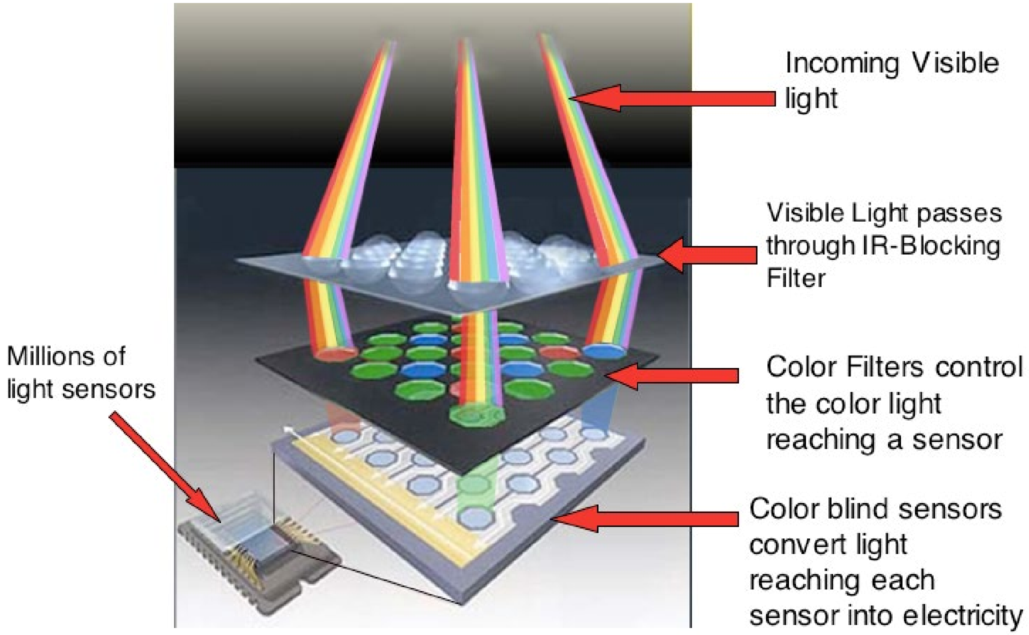


Current Denoising

- Therefore, when the camera get not enough number of light signal (short exposure), the camera will produce the noisy image.

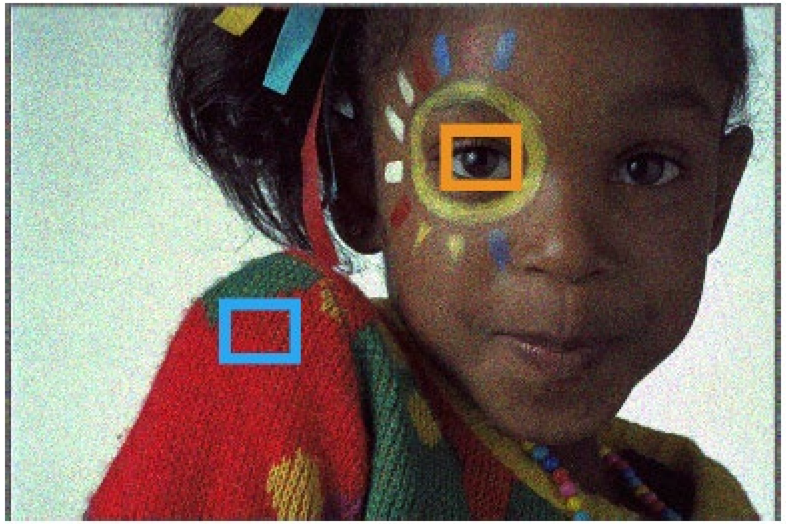
x **10** =

Motivation

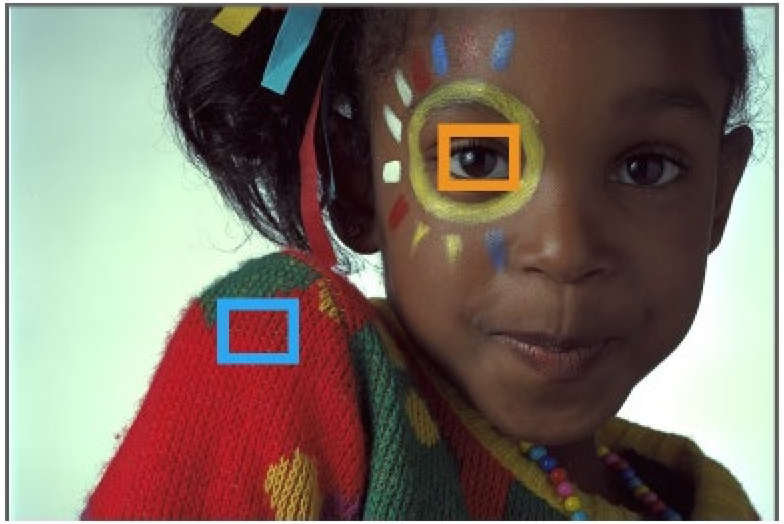


Current Denoising

- This method is possible because the random noise on the camera sensor is **Zero mean**



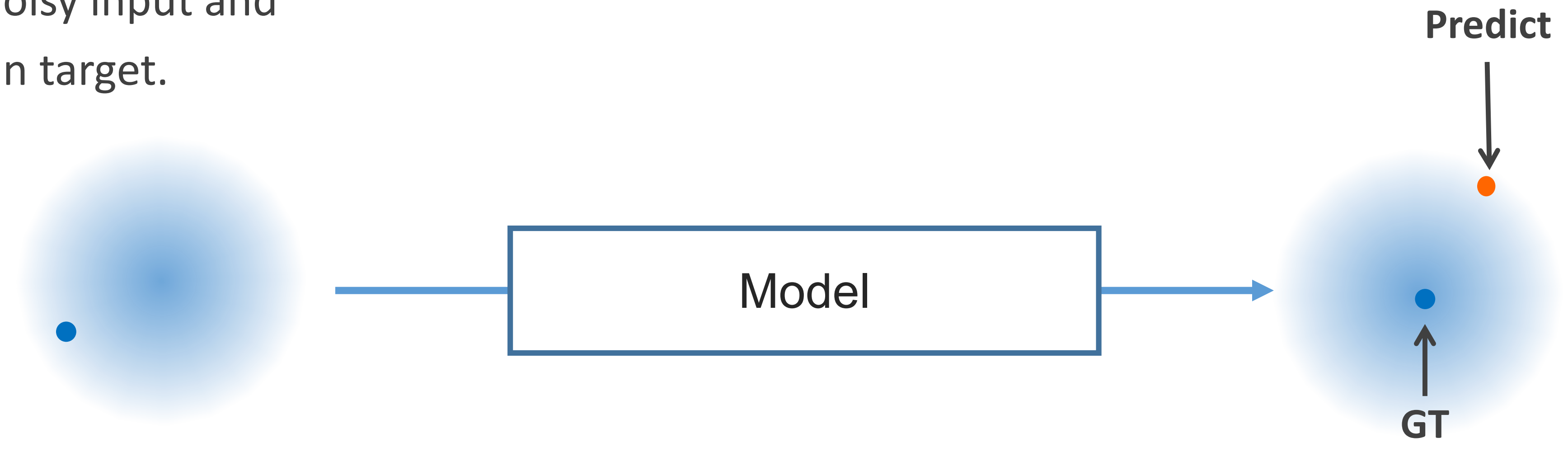
x 100 =



Approach

Supervised Denoising

- Current models take the noisy input and learns to produce the clean target.



Noisy Image

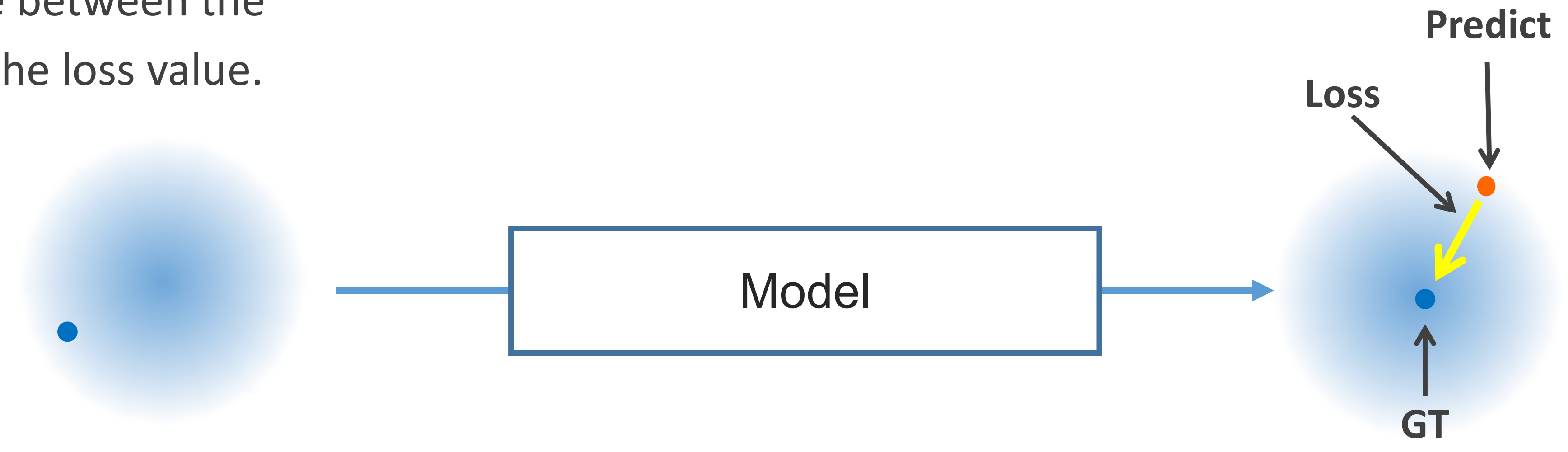
Predicted Image

[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Approach

Supervised Denoising

- The model take difference between the target and prediction for the loss value.



Noisy Image

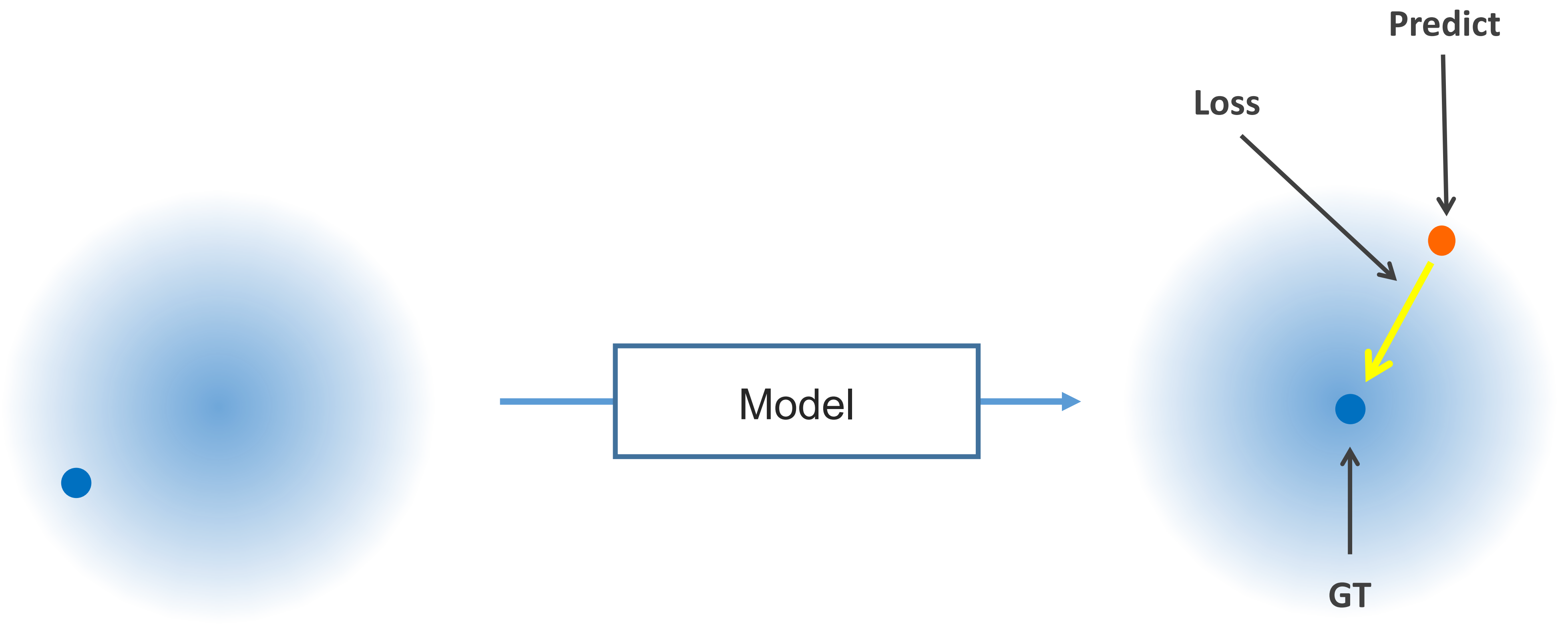
Predicted Image

[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Approach

Supervised Denoising

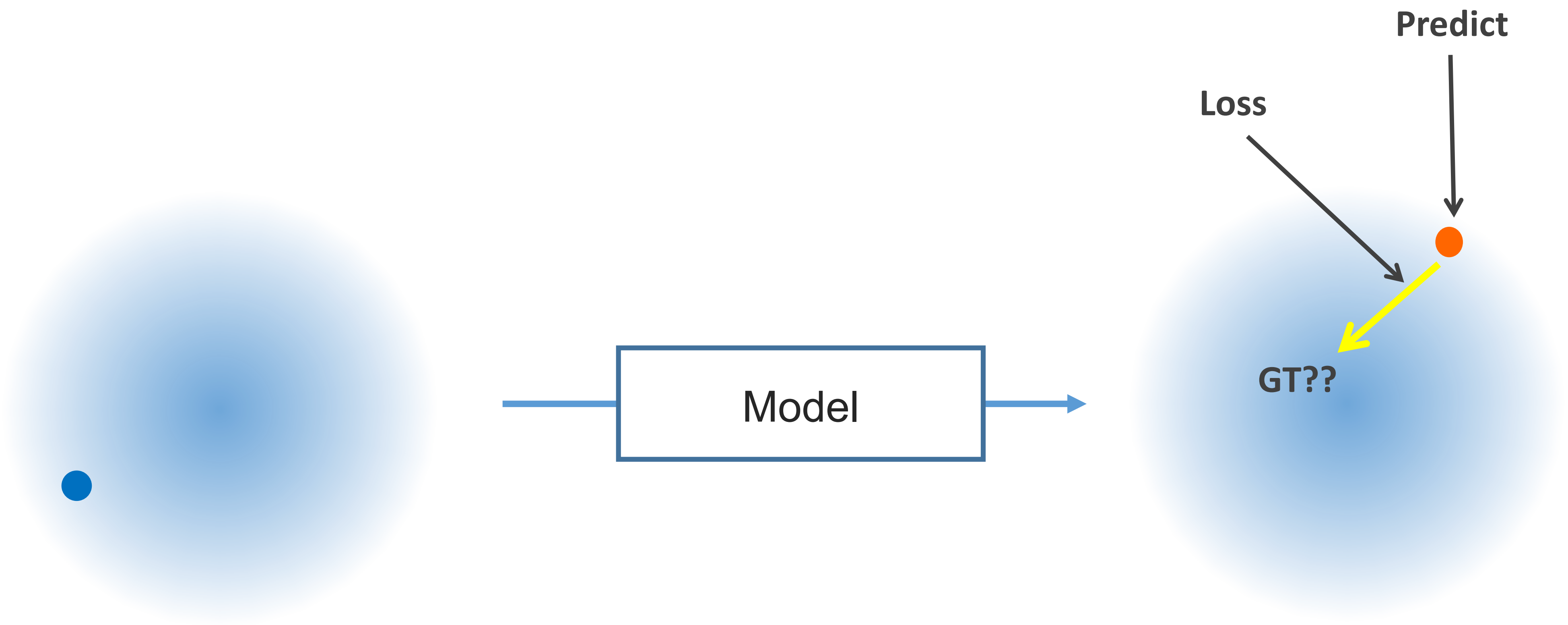
- The model take difference between the target and prediction for the loss value.



Approach

Supervised Denoising

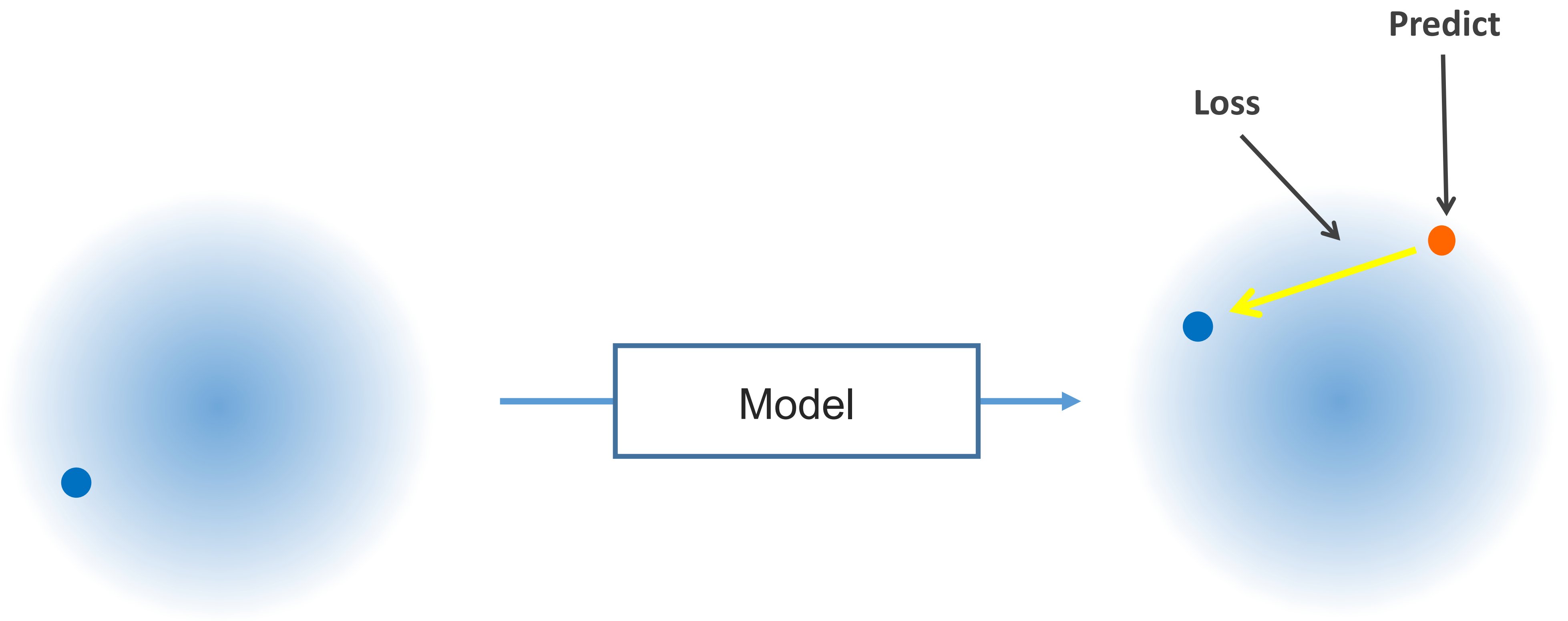
- However, in some cases, there is no GT target.



Approach

Unsupervised Denoising

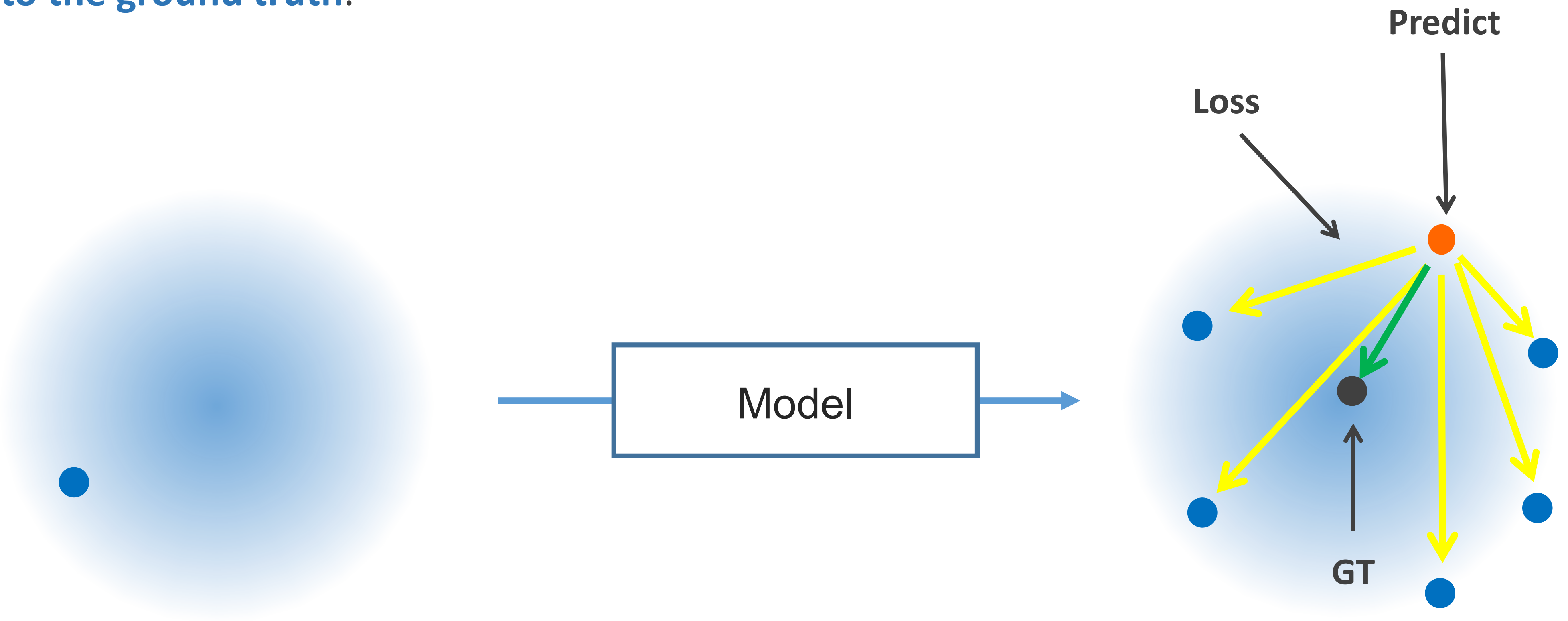
- Therefore, instead of predicting the clean target, N2N infer **another noisy data**.



Approach

Unsupervised Denoising

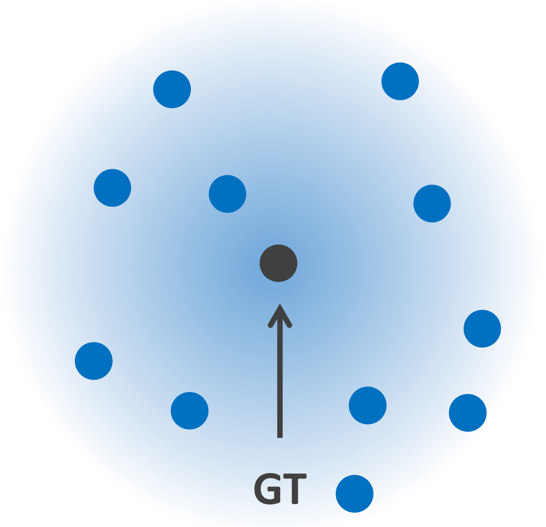
- If the **mean of the noise is zero**, the average of the gradients that model takes is **same with the gradient to the ground truth**.

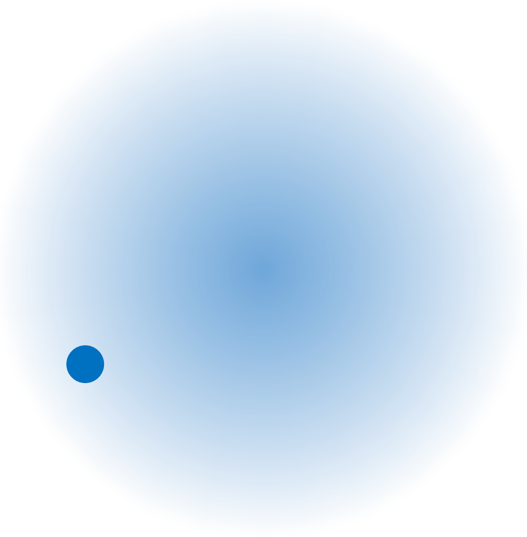


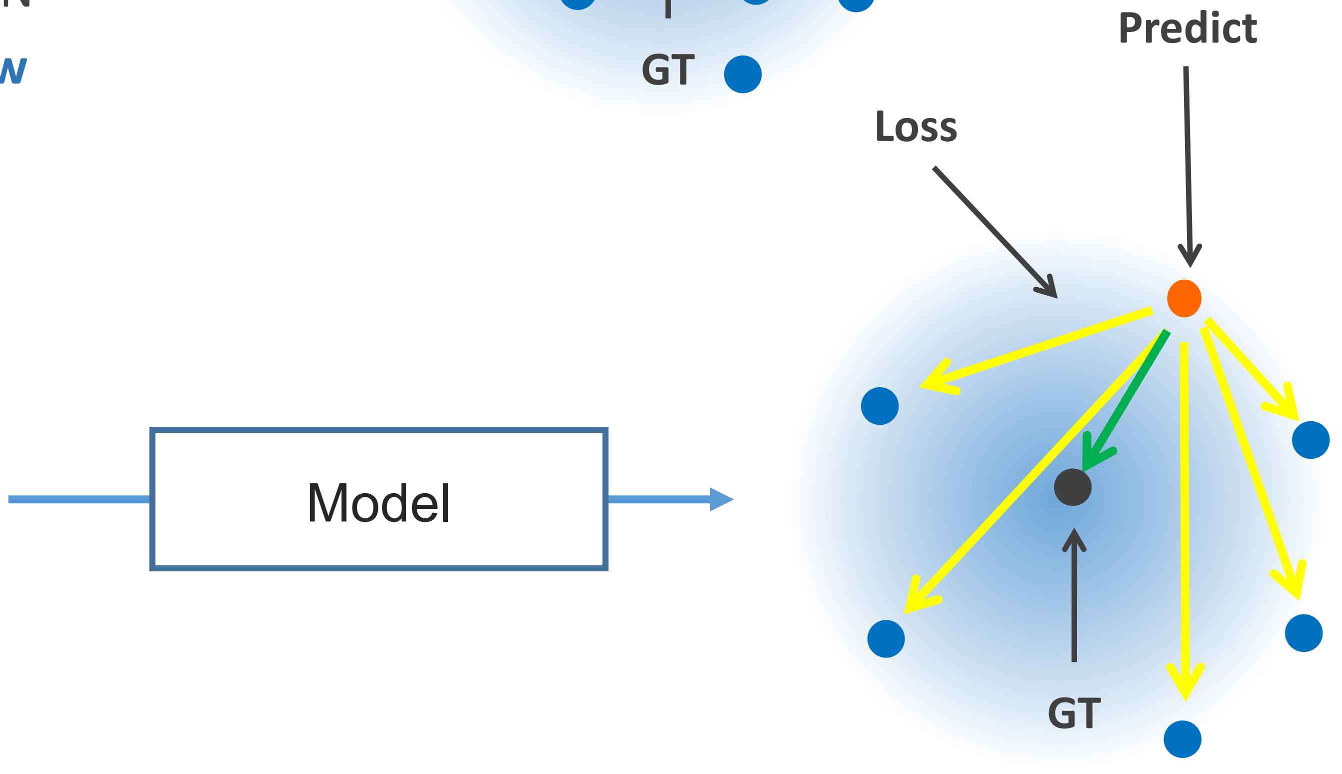
Approach

What's the difference with taking average directly from noisy images?

- In order to get a meaningful ground truth, **large number of images are required**. N2N learn to find the mean value with **only few random samples**.







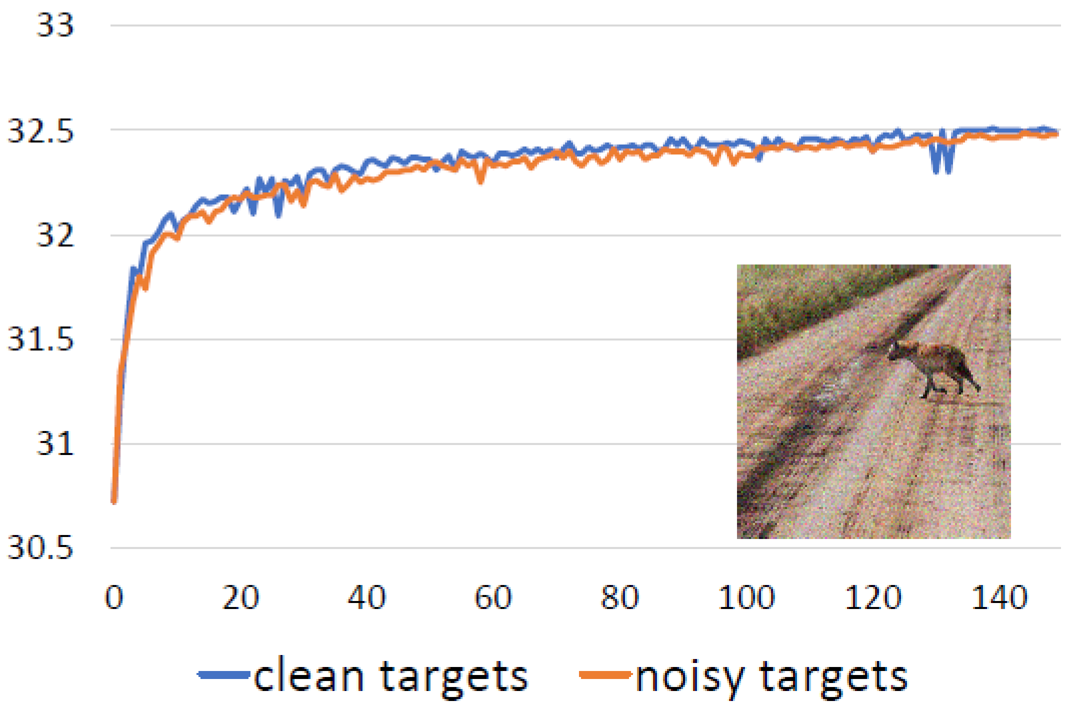
Experiments

1. Experiments

Experiment

Characteristics of N2N

- During the **training**, the N2N model **cannot succeed** in transforming one instance of the noise to another. Therefore, the training loss does not decrease well.
- However, It shows almost similar performance with supervised model at the **test accuracy**.



(a) **White Gaussian,** $\sigma = 25$

[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Experiment

Removing texts

- The ‘clean target’ below means the Supervised learned model with clean data, and rest of the results are produces by N2N.

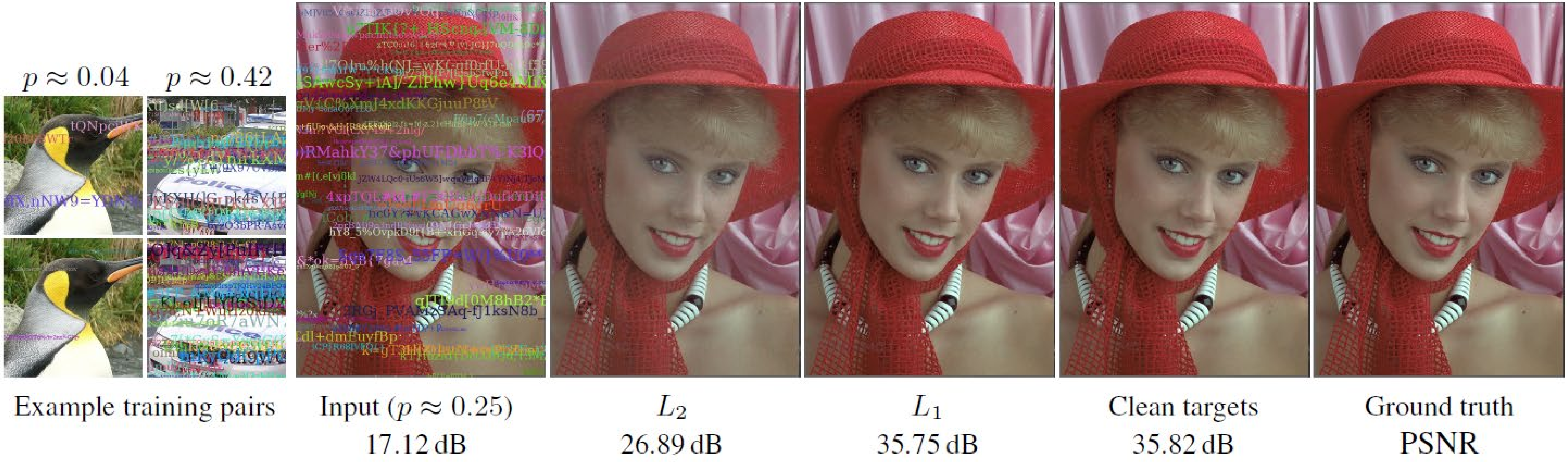


Figure 3. Removing random text overlays corresponds to seeking the median pixel color, accomplished using the L_1 loss. The mean (L_2 loss) is not the correct answer: note shift towards mean text color. Only corrupted images shown during training.

Experiment

Monte Carlo rendering denoising

- The ‘clean target’ below means the Supervised learned model with clean data, and rest of the results are produces by N2N.
- It takes 9 channel (RGB, RGB albedo, 3D normal vector of each pixel)

(a) Input (64 spp), 23.93 dB (b) Noisy targets, 32.42 dB (c) Clean targets, 32.95 dB (d) Reference (131k spp)

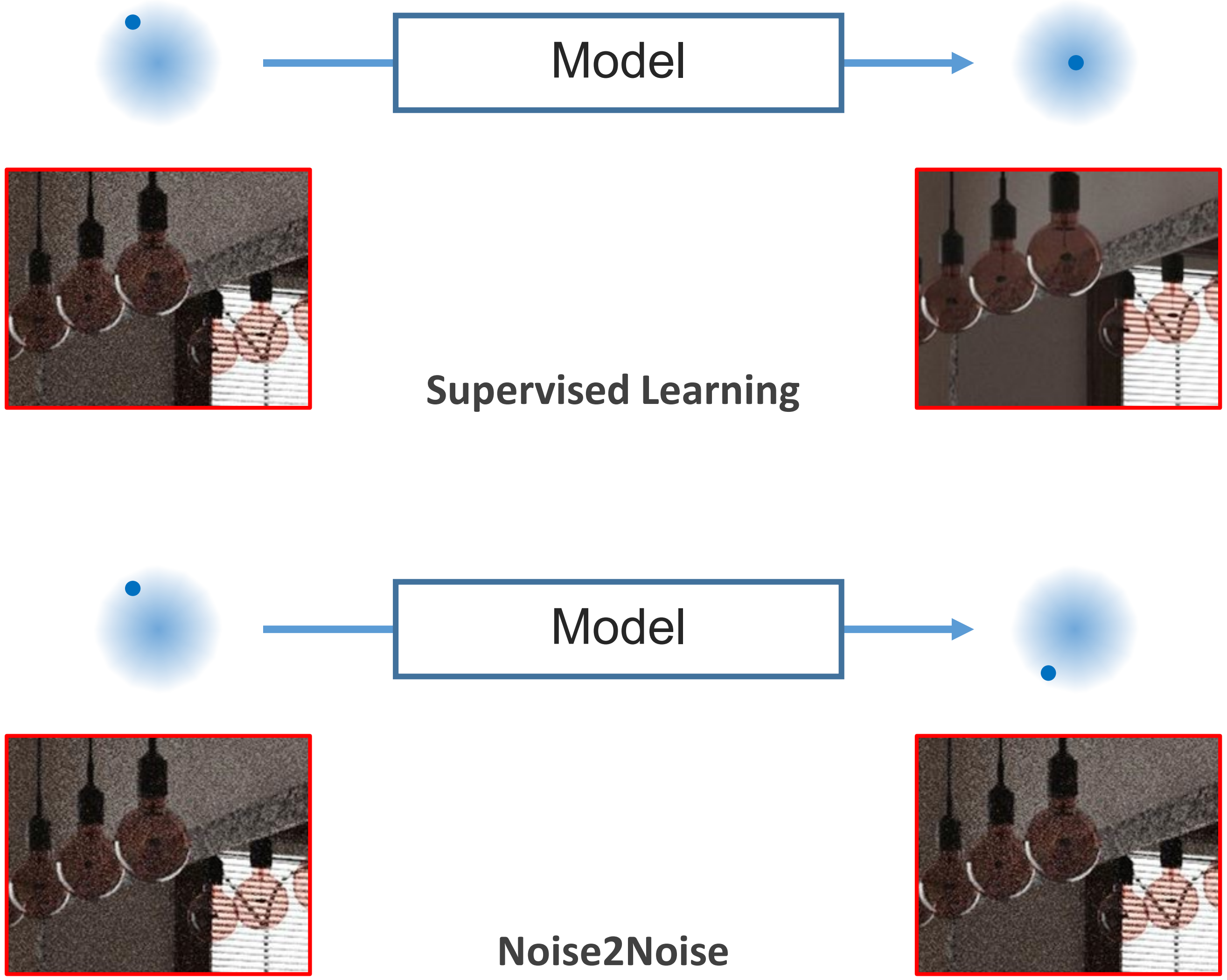
Figure 7. Denoising a Monte Carlo rendered image. (a) Image rendered with 64 samples per pixel. (b) Denoised 64 spp input, trained using 64 spp targets. (c) Same as previous, but trained on clean targets. (d) Reference image rendered with 131 072 samples per pixel. PSNR values refer to the images shown here, see text for averages over the entire validation set.

Additional Approach

Additional Approach

- Noise2Noise
  N2N is the current state of the art model for the single RGB denoising problem.

- We will try to merge N2N and KPCN model if we have enough time.



[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

Thank You!

Reference

[Liu et. al. 17] Learning Efficient Convolutional Networks through Network Slimming, ICCV2017

[Zhang et. al. 18] Image Super-Resolution Using Very Deep Residual Channel Attention Networks, ECCV2018

[Lehtinen et. al. 18] Noise2Noise: Learning Image Restoration without Clean Data, ICML2018

[Bako Et al. 17] "Kernel-Predicting Convolutional Networks for Denoising Monte Carlo Renderings." ACM Transactions on Graphics 36, no. 4 (July 20, 2017)