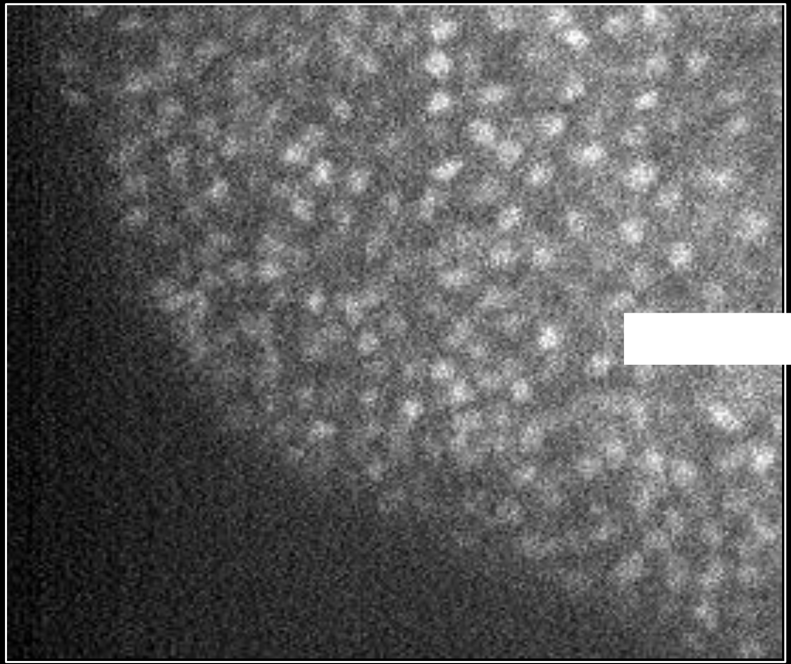
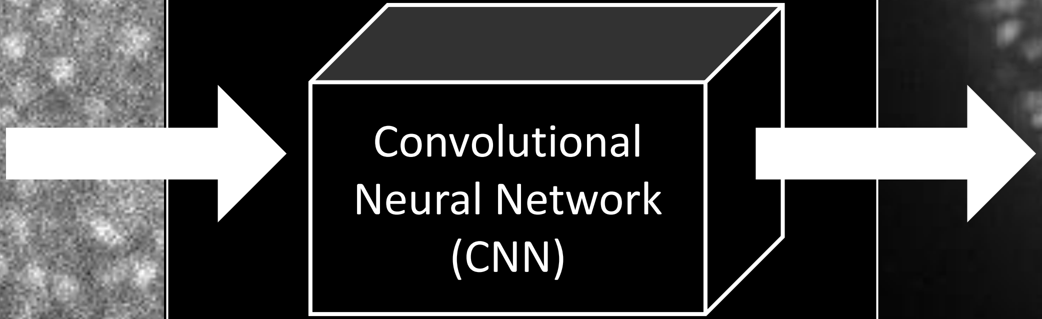
UNIVERSITY OF BIRMINGHAM

Current Topics in Data Science and AI

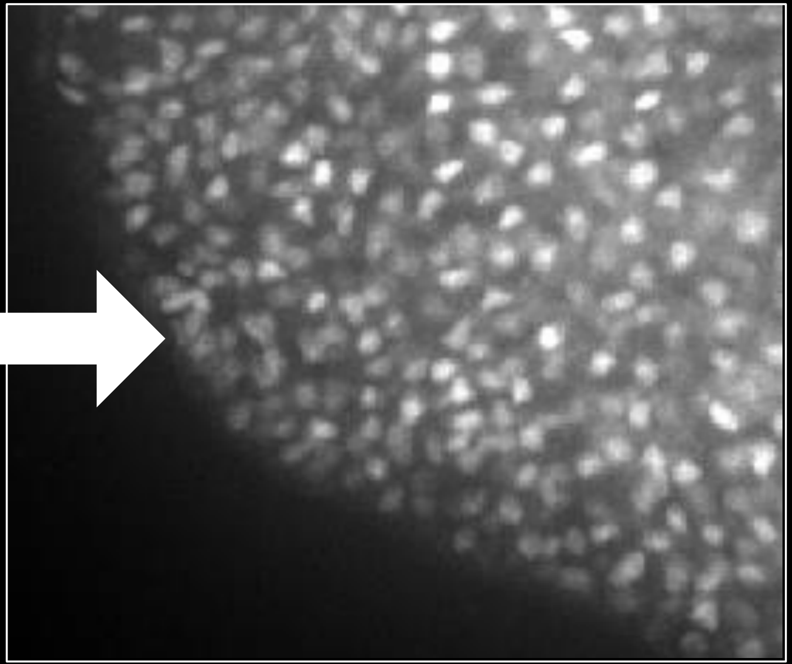
Self-Supervised Denoising in Microscopy

Deep Learning for Denoising





- 100k – 1M parameters
- Have to be trained



Low exposure:

- Low photo toxicity ☺
- Low bleaching ☺
- Noisy ☹

High exposure:

- Strong photo toxicity ☹
- Strong bleaching ☹
- Less noise ☺

Applied to microscopy (CARE): Weigert et al. 2018

Self-Supervised
Noise2Void

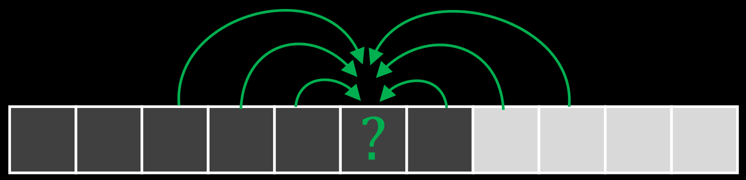
You only need individual noisy images!

Krull, Buchholz, and Jug 2019

Noise2Void – Assumptions

$p(x,y)=p(x)p(y)$

can estimate central signal



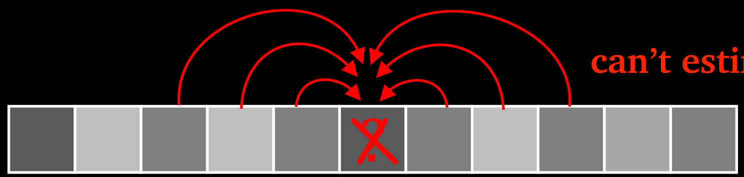
Signal $\boldsymbol{s} = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m)$

Noisy image ($\boldsymbol{x}$) aka *observation*

$=$

$+$

can't estimate central noise



Noise $\boldsymbol{n} = (\boldsymbol{n}_1, \ldots, \boldsymbol{n}_m)$

independent

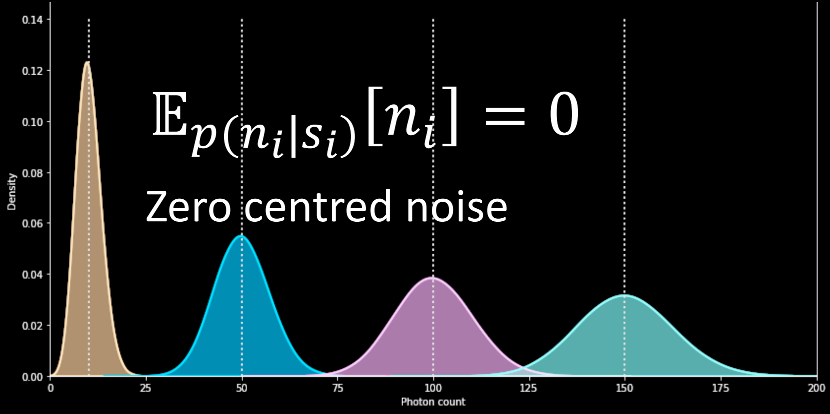
additive noise

$$p(\boldsymbol{n}|\boldsymbol{s}) = p(n_1 \ldots n_m|\boldsymbol{s}) = \prod_{i=1}^{m} p(n_i|s_i)$$

$$n_i \perp\!\!\!\perp n_j \mid s_i$$

$$p(\boldsymbol{x}|\boldsymbol{s}) = p(x_1 \ldots x_m|\boldsymbol{s}) = \prod_{i=1}^{m} p(x_i|s_i)$$

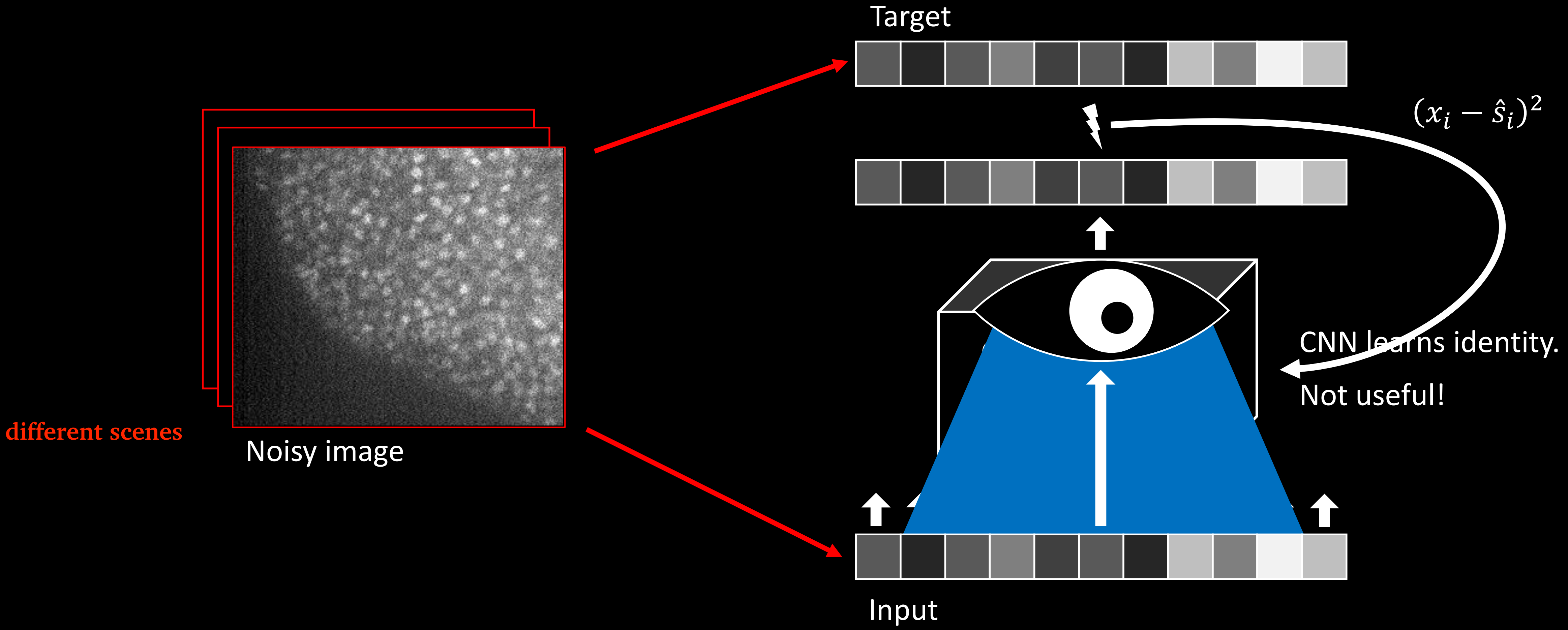
$$x_i \perp\!\!\!\perp x_j \mid s_i$$



$$\mathbb{E}_{p(n_i|s_i)}[n_i] = 0$$

Zero centred noise

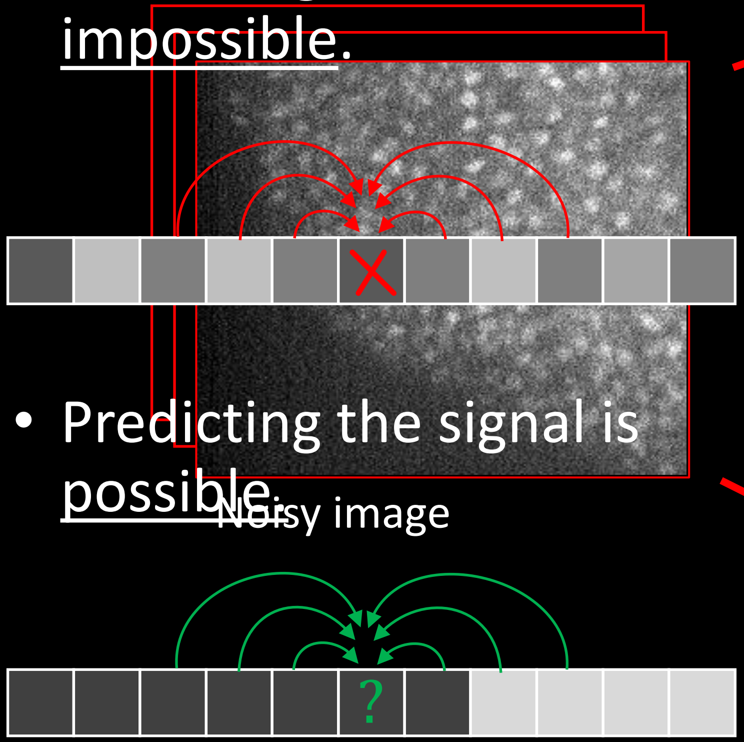
Noise2Void


different scenes
Noisy image
Target
$(x_i - \hat{s}_i)^2$
CNN learns identity.
Not useful!
Input

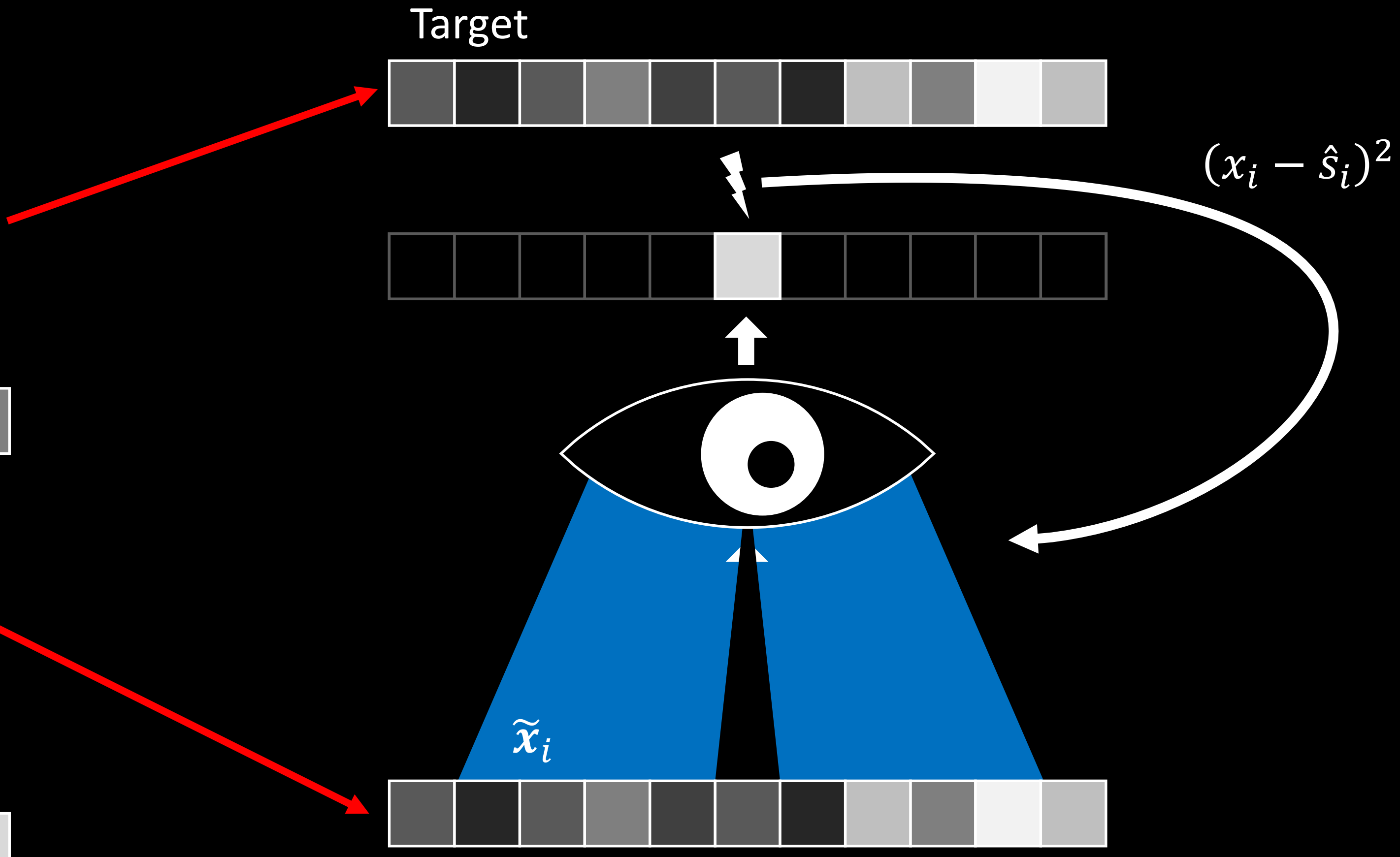
Noise2Void - Blind Spot Network

Why does it work?

- Predicting the noise is impossible.

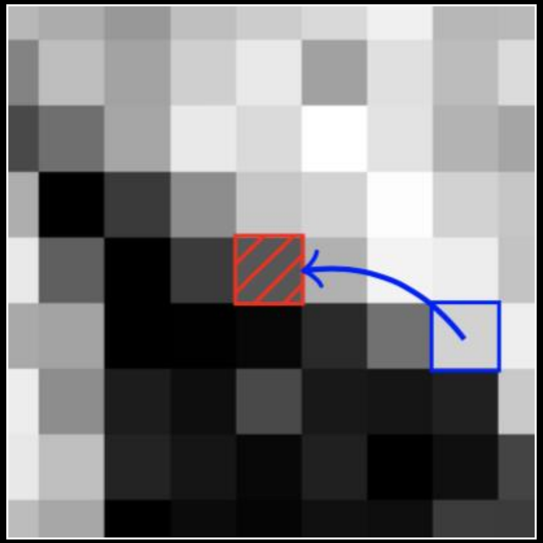


- Predicting the signal is possible.

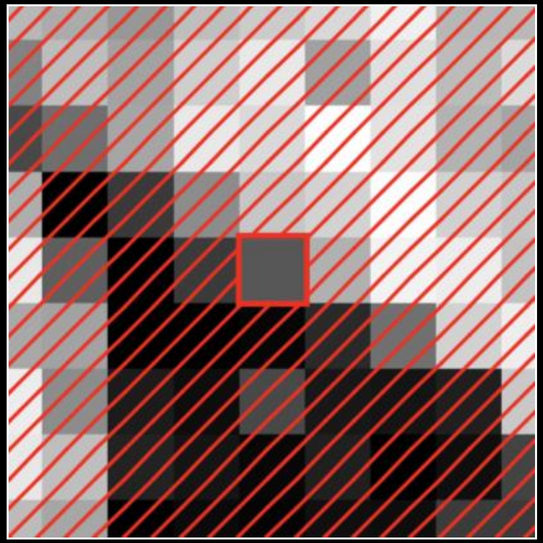


Noise2Void - Blind Spot Implementation

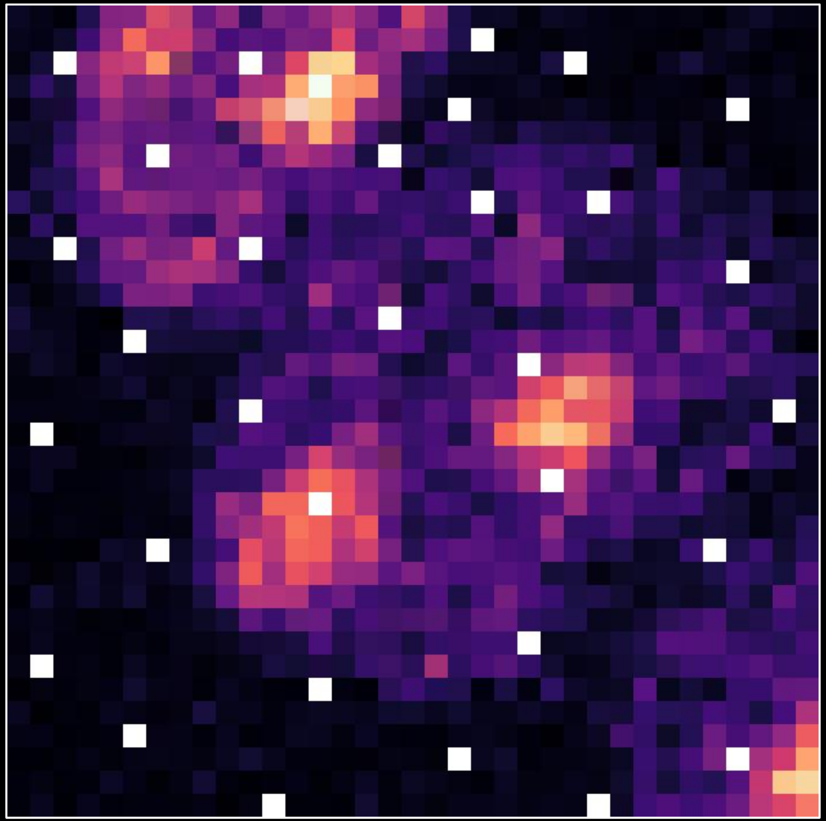
input



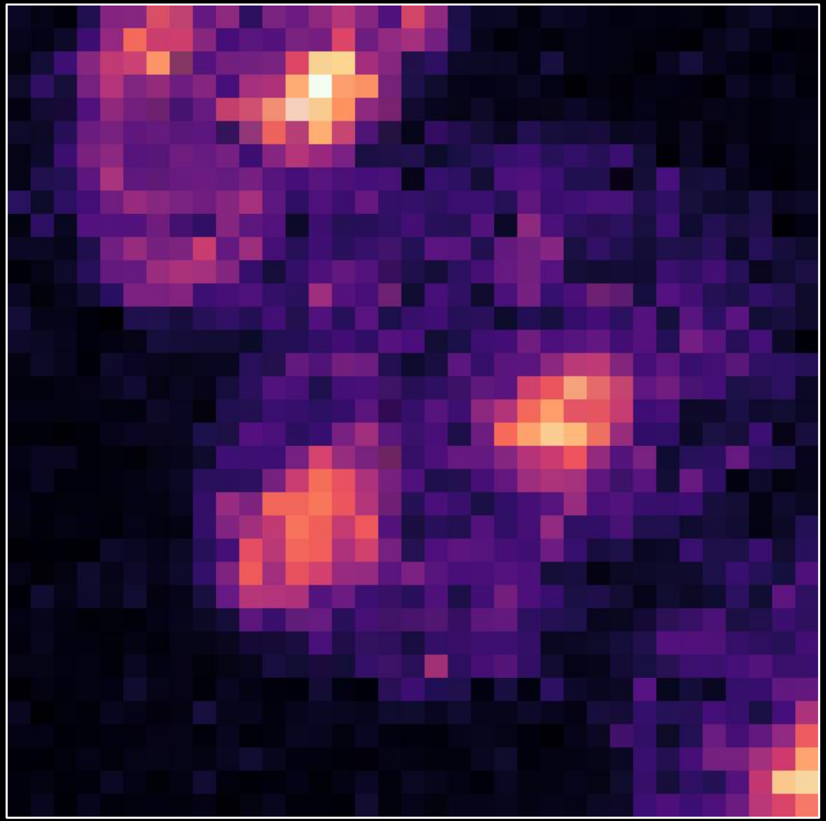
target



Noise2Void - Blind Spot Implementation

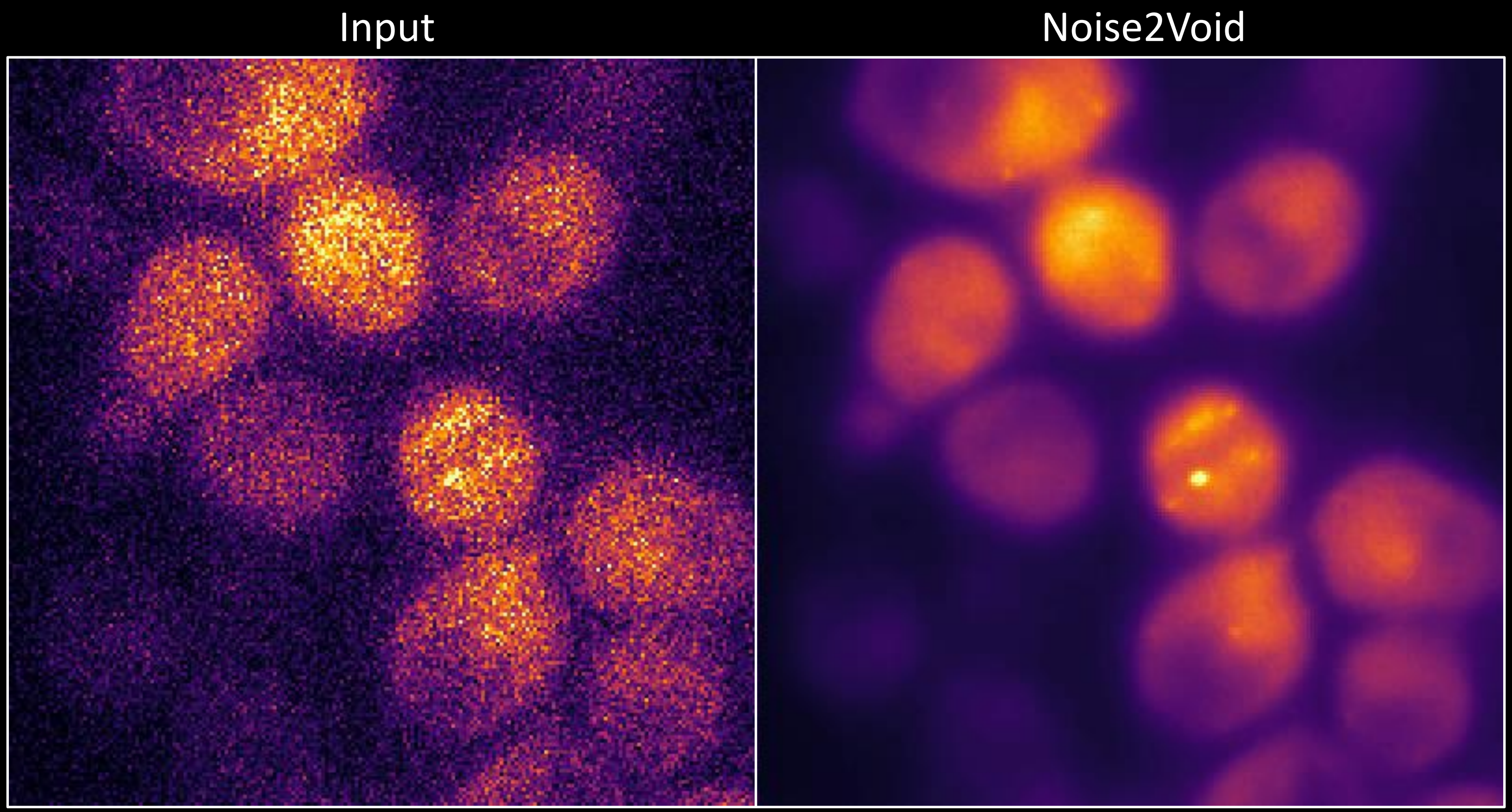


Input

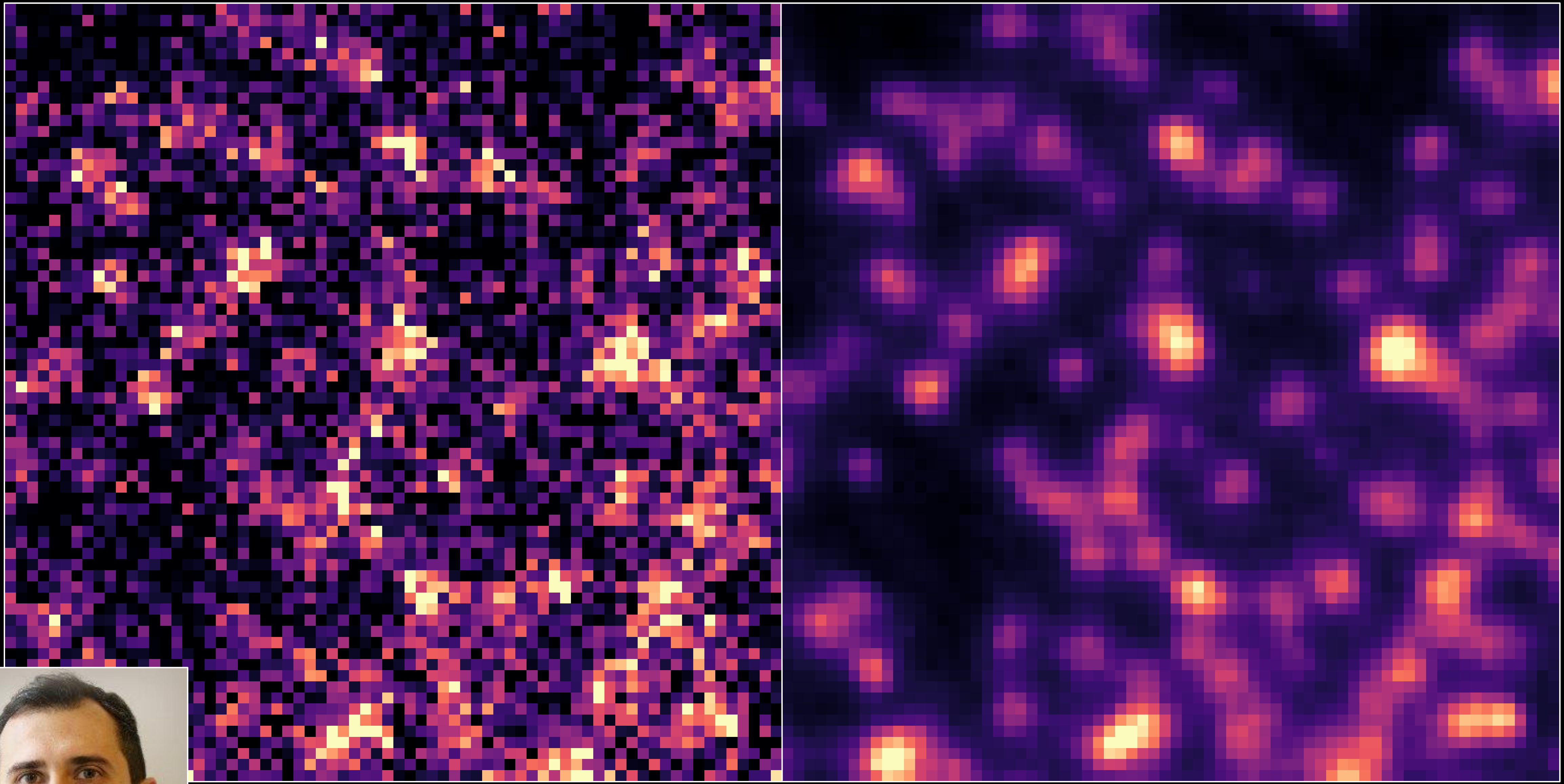


Target

Noise2Void - Results

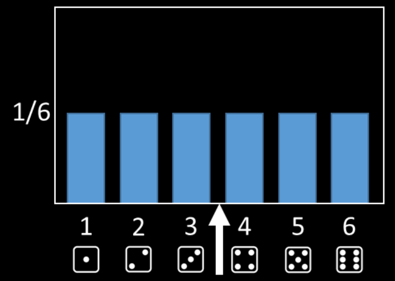


Data by Stephanie Heinrich



Data by Davide Calebiro

Noise2Void - Why does it work?



Supervised:

Minimising $\sum_i (s_i - \hat{s}_i)^2 \longrightarrow \hat{s}_i \approx \mathbb{E}_{p(s_i|\boldsymbol{x})}[s_i]$

N2V: 

Minimising $\sum_i (x_i - \hat{s}_i)^2 \longrightarrow \hat{s}_i \approx \mathbb{E}_{p(x_i|\tilde{\boldsymbol{x}}_i)}[x_i]$

$$= \mathbb{E}_{p(s_i|\tilde{\boldsymbol{x}}_i)}[s_i]$$

Why it works:

$$\mathbb{E}_{p(x_i|\tilde{\boldsymbol{x}}_i)}[x_i]$$

$$= \int x_i \, p(x_i|\tilde{\boldsymbol{x}}_i) \, dx_i$$

Marginalisation

$$= \int x_i \int p(x_i, s_i|\tilde{\boldsymbol{x}}_i) \; ds_i \, dx_i$$

Product rule

$$= \int x_i \int p(x_i|s_i, \tilde{\boldsymbol{x}}_i) p(s_i|\tilde{\boldsymbol{x}}_i) \; ds_i \, dx_i$$

$$= \int \int x_i p(x_i|s_i, \tilde{\boldsymbol{x}}_i) p(s_i|\tilde{\boldsymbol{x}}_i) \; ds_i \, dx_i$$

$$= \int \int x_i p(x_i|s_i, \tilde{\boldsymbol{x}}_i) p(s_i|\tilde{\boldsymbol{x}}_i) \; dx_i \, ds_i$$

$$= \int p(s_i|\tilde{\boldsymbol{x}}_i) \int x_i \, p(x_i|s_i, \tilde{\boldsymbol{x}}_i) \; dx_i \, ds_i$$

$$= \int p(s_i|\tilde{\boldsymbol{x}}_i) \int x_i \, p(x_i|s_i, \tilde{\boldsymbol{x}}_i) \; dx_i \, ds_i$$

Cond. independence

$$= \int p(s_i|\tilde{\boldsymbol{x}}_i) \int x_i \, p(x_i|s_i) \; dx_i \, ds_i$$

Expected value

$$= \int p(s_i|\tilde{\boldsymbol{x}}_i) \, \mathbb{E}_{p(x_i|\boldsymbol{s}_i)}[x_i] \; ds_i$$

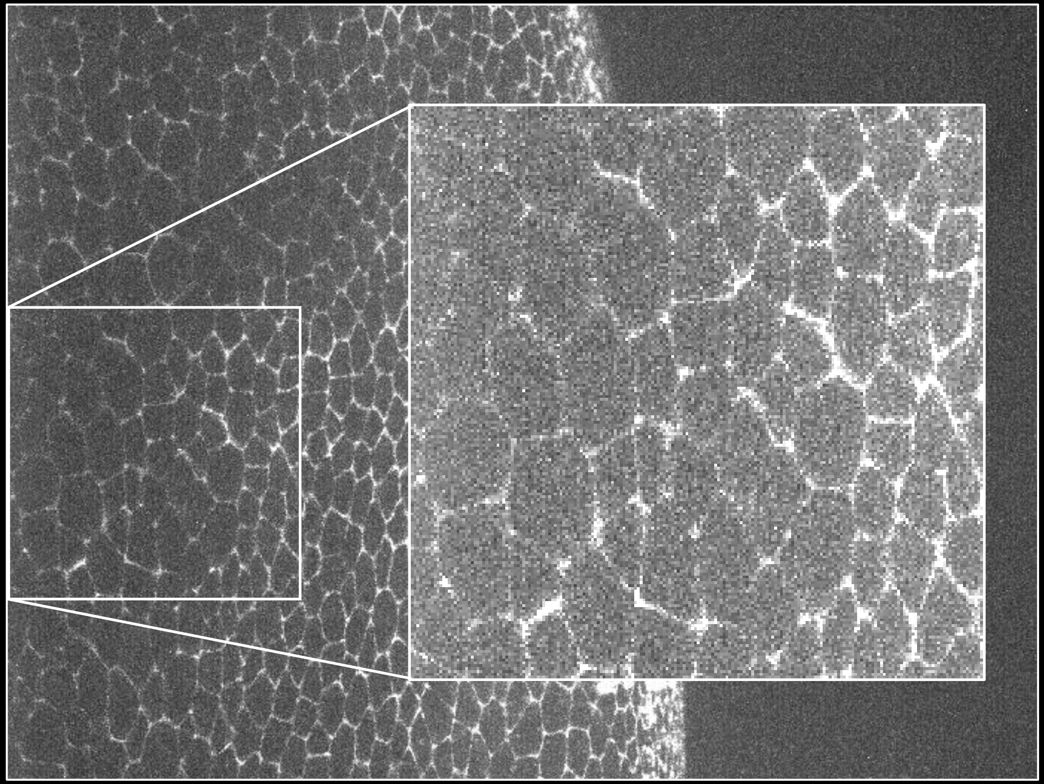
Zero centred noise

$$= \int p(s_i|\tilde{\boldsymbol{x}}_i) s_i \; ds_i$$

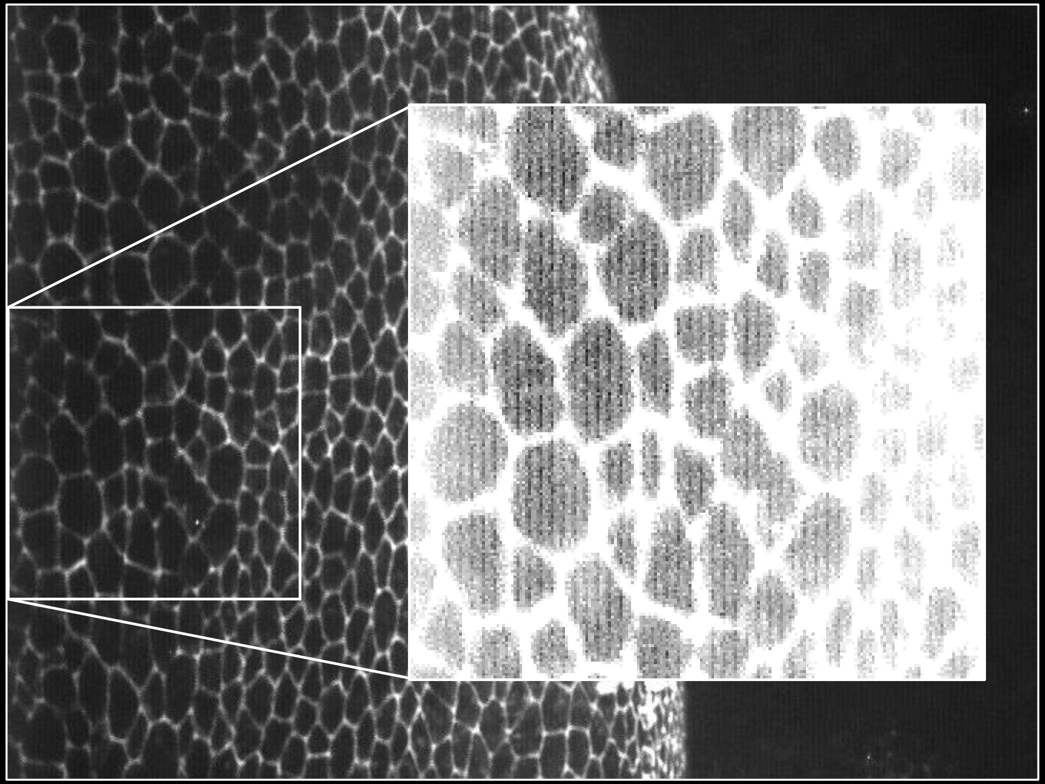
Expected value

$$= \mathbb{E}_{p(s_i|\tilde{\boldsymbol{x}}_i)}[s_i]$$

Noise2Void - results



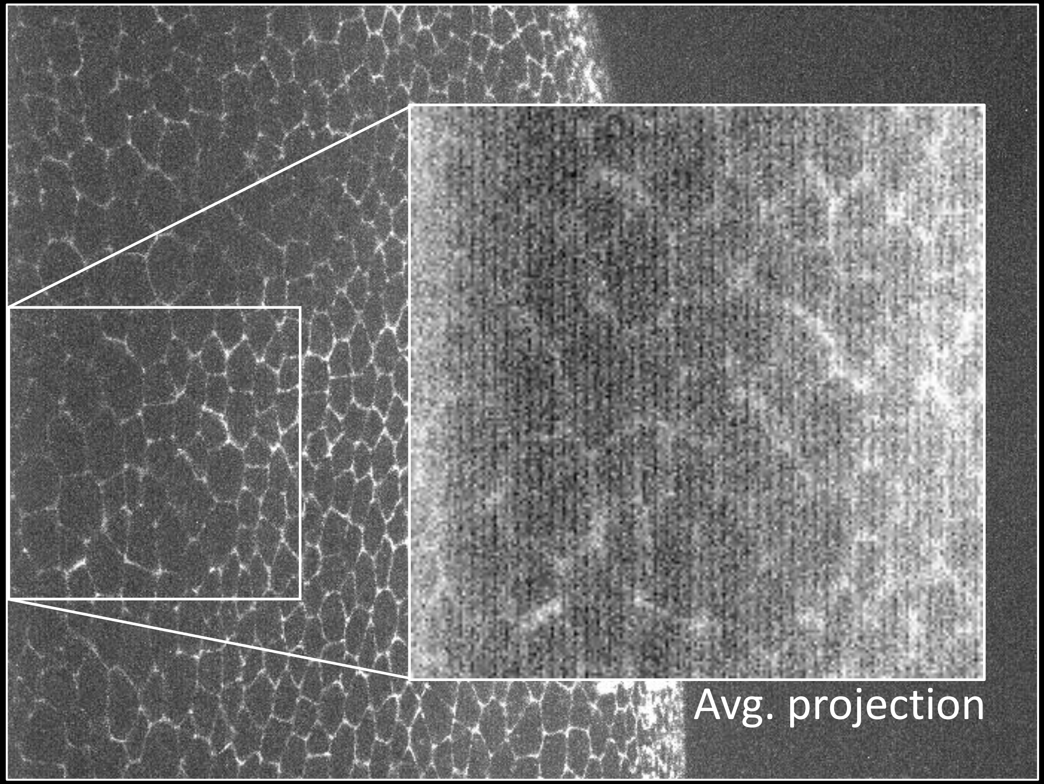
Input (max projection)



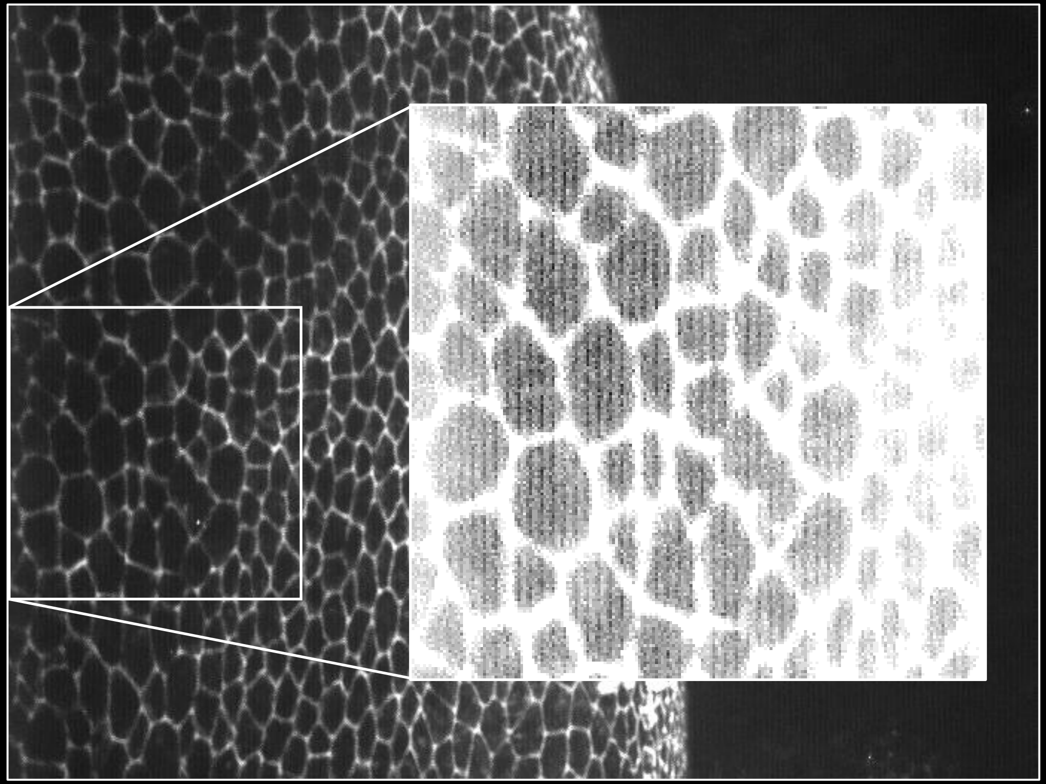
Noise2Void (max projection)

Data by Romina Piscitel, Eaton lab at MPI-CBG

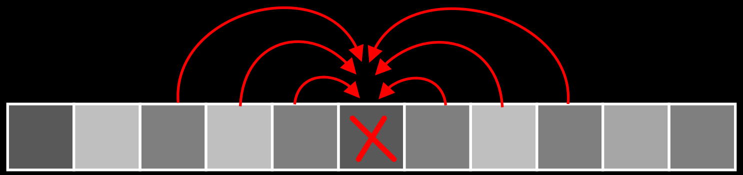
Noise2Void - limitations



Input (max projection)



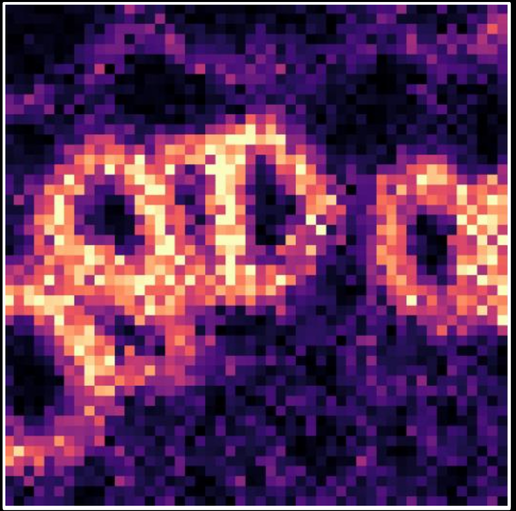
Noise2Void (max projection)



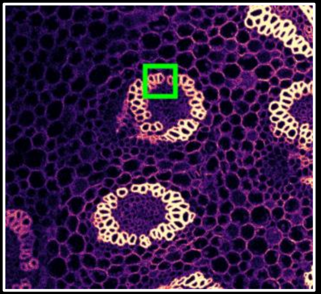
Data by Romina Piscitel, Eaton lab at MPI-CBG

Room for Improvement

Traditional (supervised)

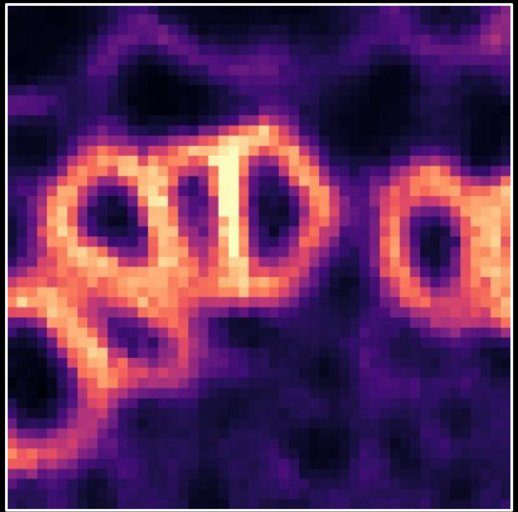


PSNR:36.71

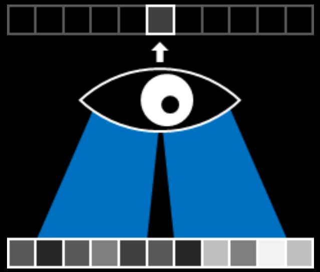


$$\mathbb{E}_{p(s_i|\boldsymbol{x})}[s_i]$$

Noise2Void

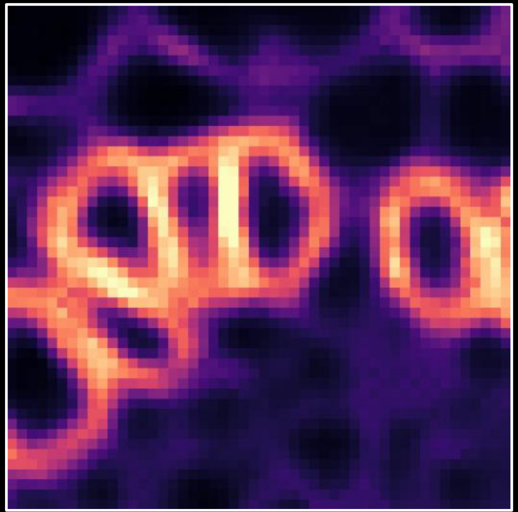


PSNR:35.73



$$\mathbb{E}_{p(s_i|\tilde{\boldsymbol{x}}_i)}[s_i]$$

Probabilistic N2V

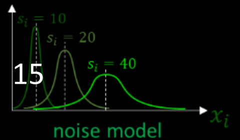


PSNR:36.51

Idea:

- Predict prob. distributions 
- Include probabilistic noise model

$$\mathbb{E}_{p(s_i|\boldsymbol{x})}[s_i]$$



Summary

- Self-Supervised denoising (Noise2Void):
  - Works with unpaired data.
  - Assumptions about the noise.
  - Predict each pixel value from its surroundings.
  - Will remove unstructured noise.
  - Will reveal structured noise.
  - PSNR below supervised training.
- Probabilistic Noise2Void:
  - Improved PSNR
  - Reintroduces information from blind spot.
  - Requires noise model.