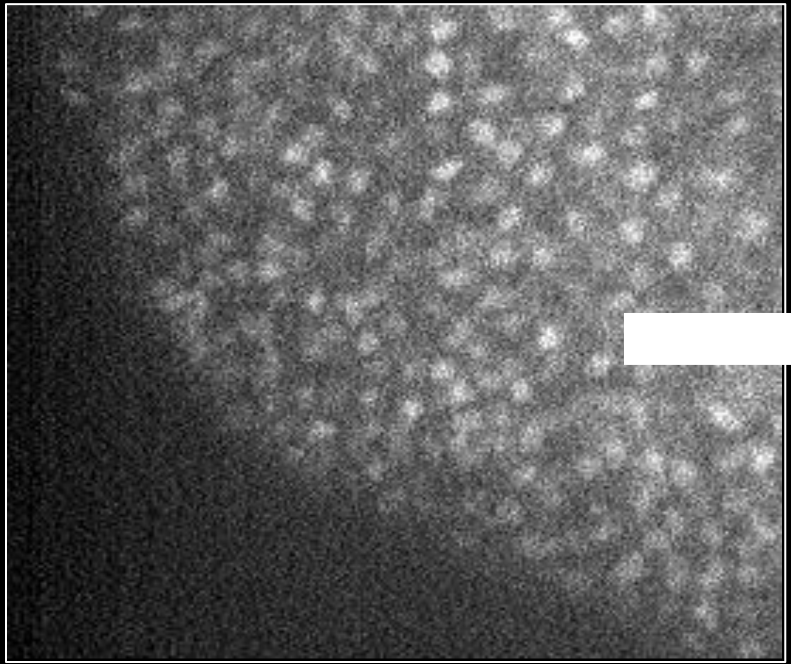
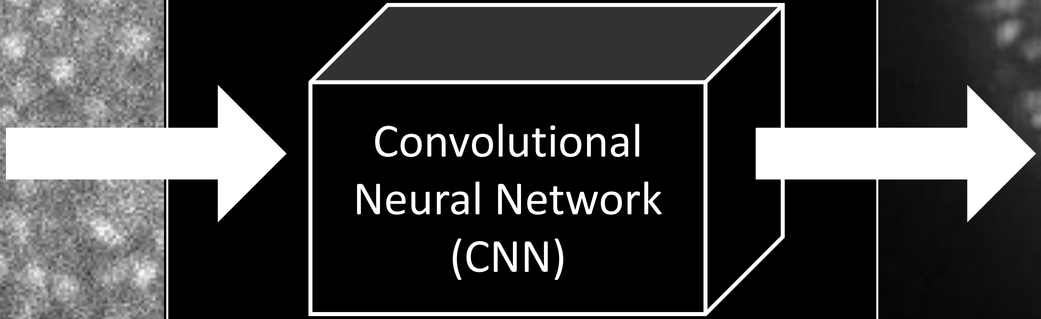
UNIVERSITY OF BIRMINGHAM

Current Topics in Data Science and AI

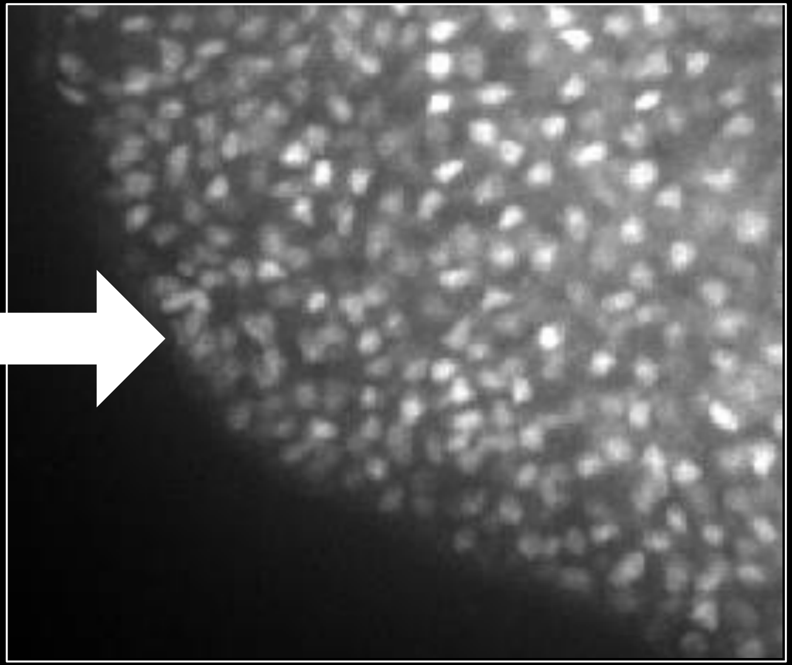
Denoising in Microscopy

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

Traditional Supervised Training

You need clean data.

CARE – Traditional Supervised Training

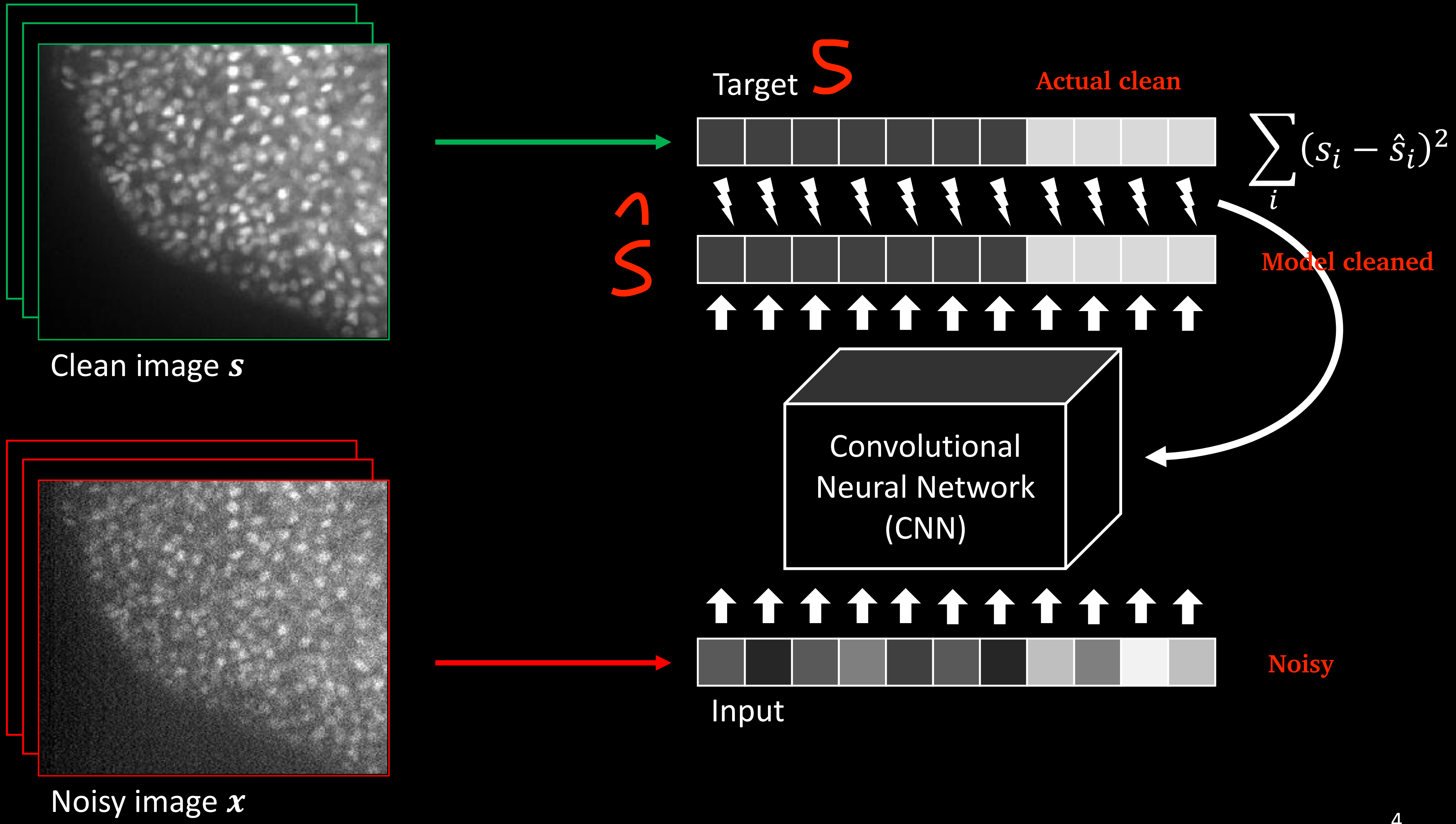


Image Generation Model

$$\boldsymbol{s} \longrightarrow \boldsymbol{x}$$

$p(\boldsymbol{s})$

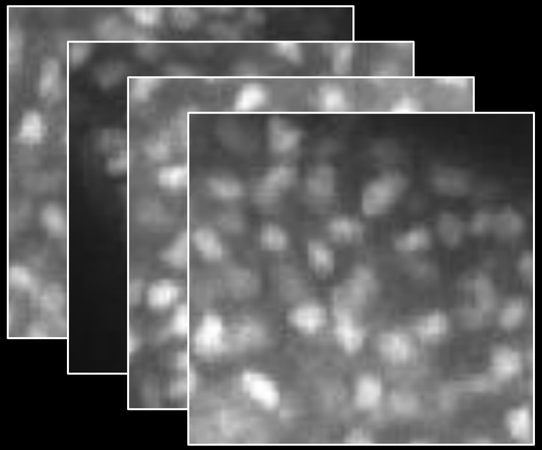
The distr. of clean images

$p(\boldsymbol{x}|\boldsymbol{s})$ Given clean images, what will be noisy images?

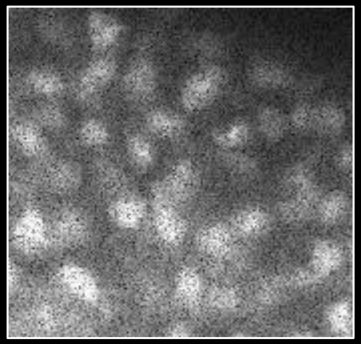
or

$p(\boldsymbol{n}|\boldsymbol{s})$ with $\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{n}$

The noise distribution clean+noise



A number of images required mannually



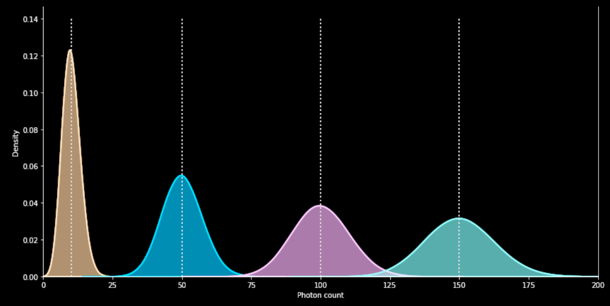


Image Generation Model

$$s \longrightarrow x$$

$p(s)$
The distr. of clean images

$p(x|s)$
or
$p(n|s)$ with $x = s + n$
The noise distribution

$$p(\boldsymbol{x}, \boldsymbol{s}) = p(\boldsymbol{x}|\boldsymbol{s})p(\boldsymbol{s})$$ Product rule

$$p(\boldsymbol{x}) = \int p(\boldsymbol{x}, \boldsymbol{s})\, ds$$ Marginalisation

$$p(\boldsymbol{s}|\boldsymbol{x}) = \frac{p(\boldsymbol{x}, \boldsymbol{s})}{p(\boldsymbol{x})}$$ Conditional probability

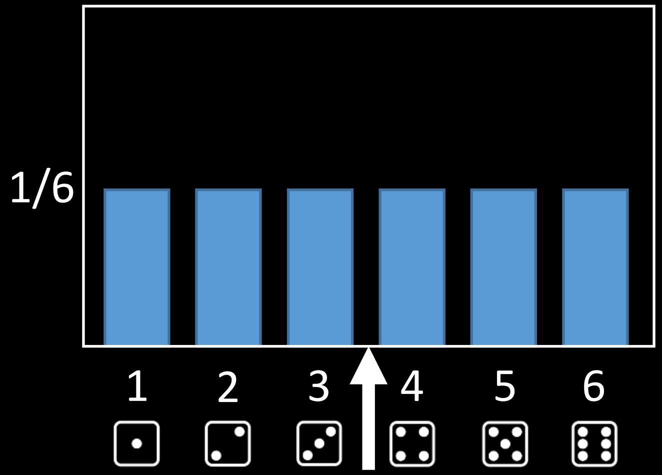
The MMSE Estimate

Network tries to map $\boldsymbol{x}$ to $\boldsymbol{s}$, but
$p(\boldsymbol{s}|\boldsymbol{x})$ is a probability distribution.

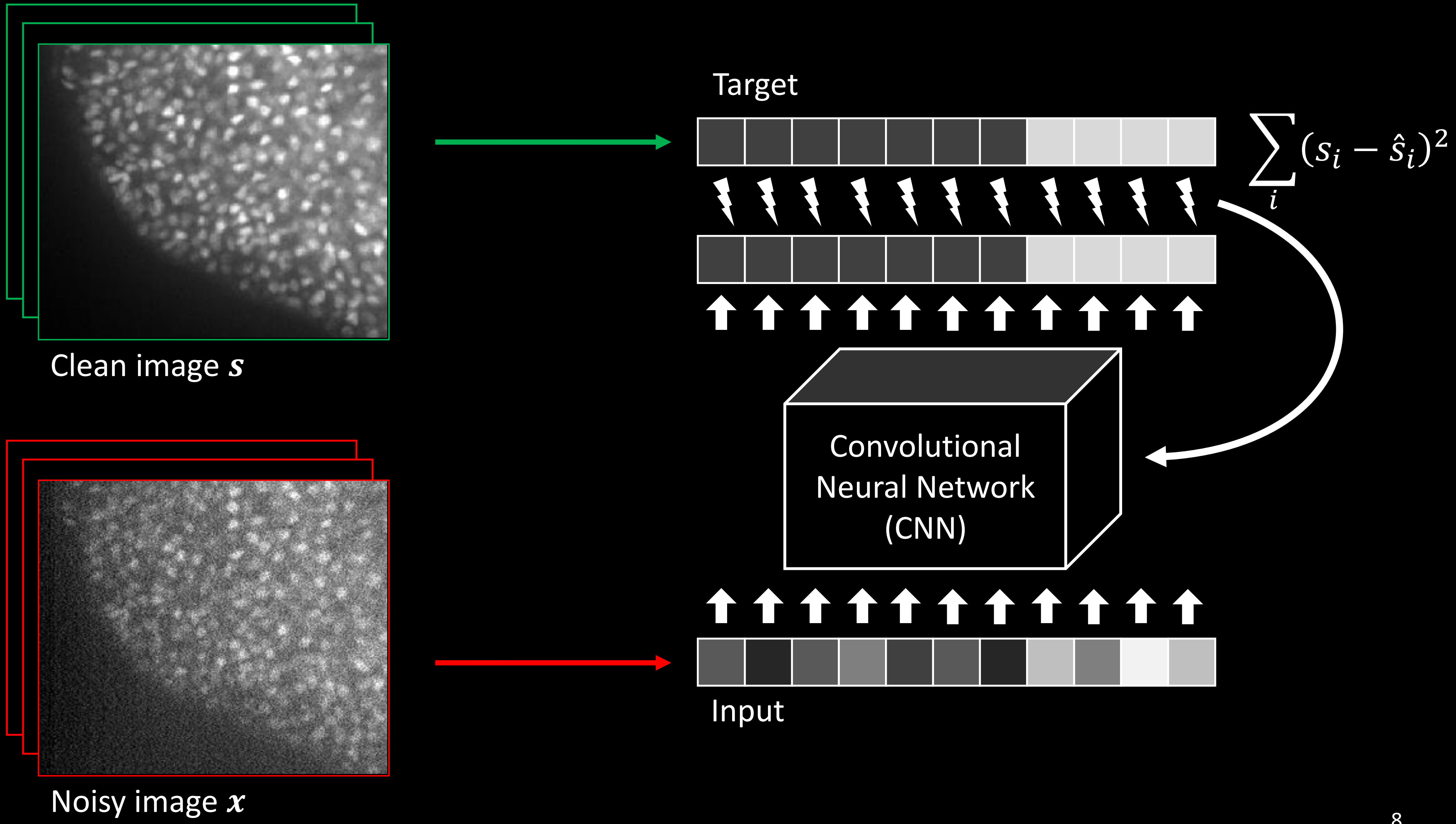
minimising $\sum_i (s_i - \hat{s}_i)^2 \longrightarrow \hat{\boldsymbol{s}} \approx \mathbb{E}_{p(s_i|\boldsymbol{x})}[s_i]$

Expected Valu estimate clean images

Rolling a dice: 



CARE – Traditional Supervised Training



Noise2Noise

You only need noisy data!

Lehtinen et al. 2018

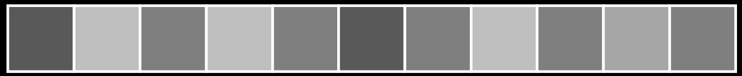
Noise2Noise

Noisy image $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_m)$

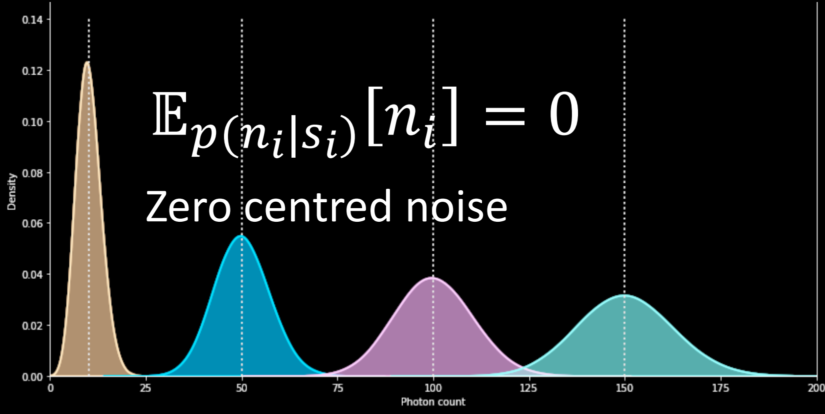
$=$

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

$+$



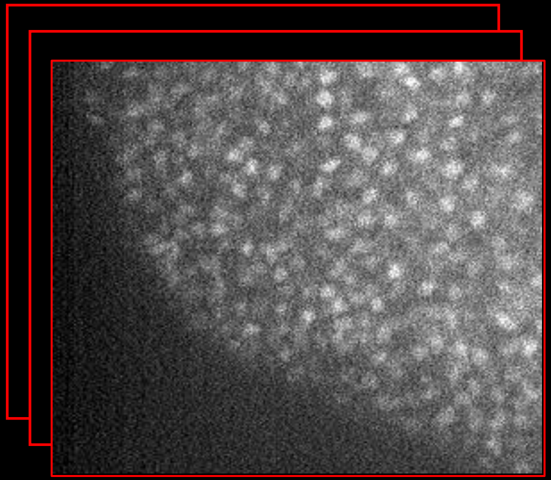
Noise $\boldsymbol{n} = (\boldsymbol{n}_1, \dots, \boldsymbol{n}_m)$



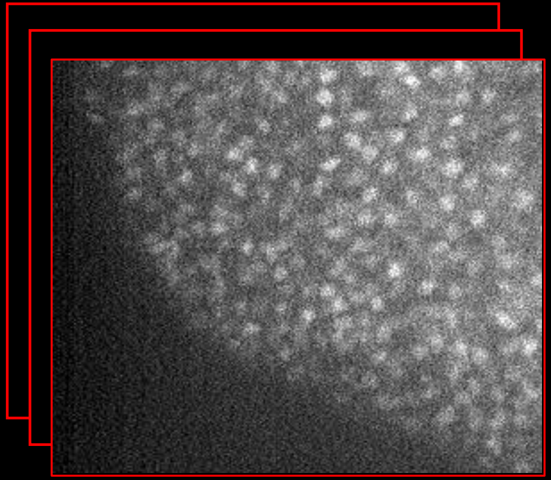
$$\mathbb{E}_{p(n_i|s_i)}[n_i] = 0$$

Zero centred noise

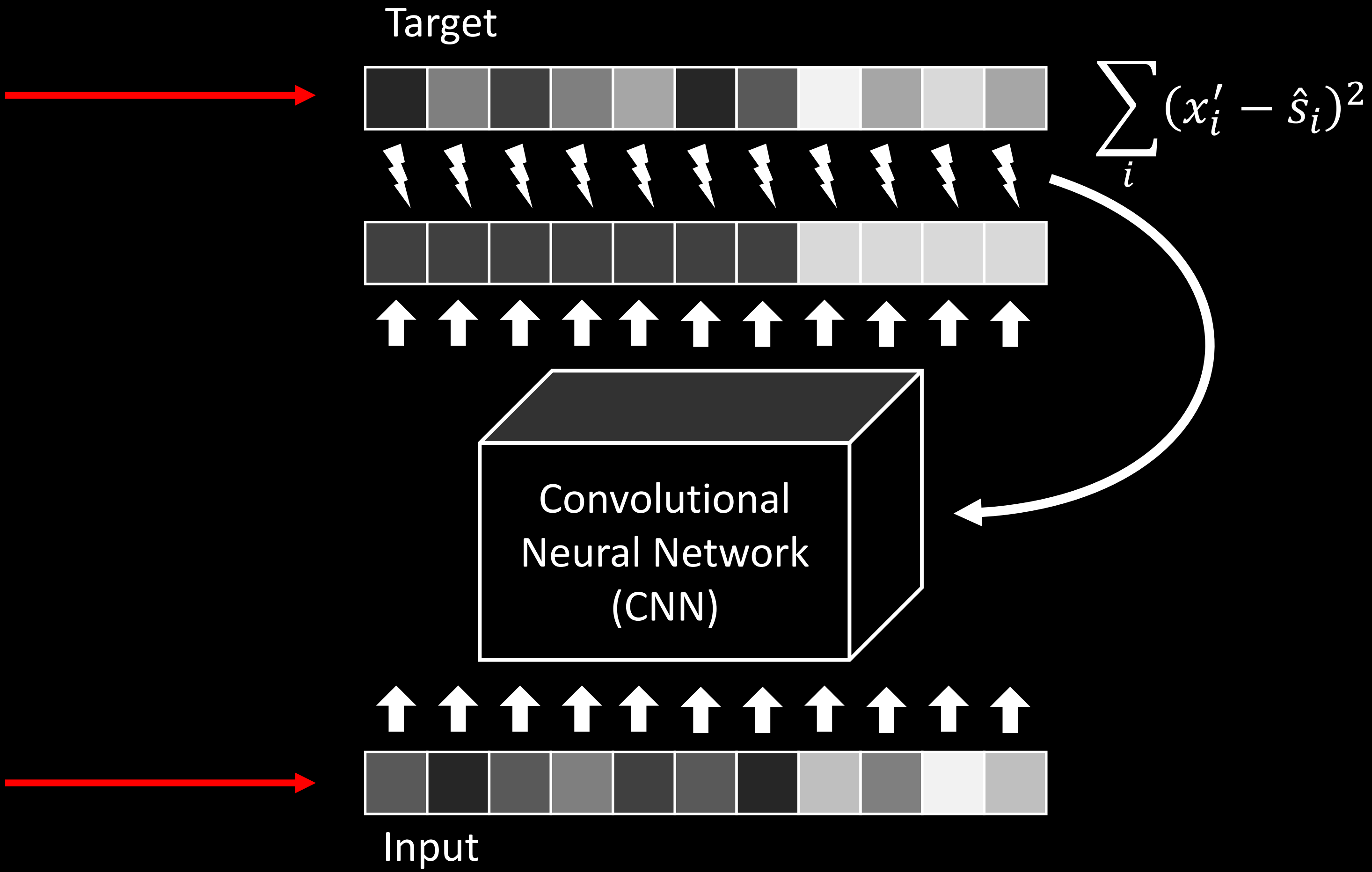
Noise2Noise training



Noisy image: $\boldsymbol{x}' = \boldsymbol{s} + \boldsymbol{n}'$

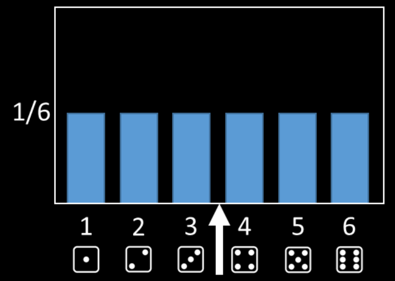


Noisy image: $x = \boldsymbol{s} + n$



Applied to cryo EM: Buchholz et al. 2018

Noise2Noise - Why does it work?



Supervised:

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

actual   model predicted

N2N:

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

another noisy   model predicted

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

Noise2Noise - Why does it work?

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

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

Marginalisation

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

Product rule

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

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

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

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

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

Cond. independence

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

Expected value

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

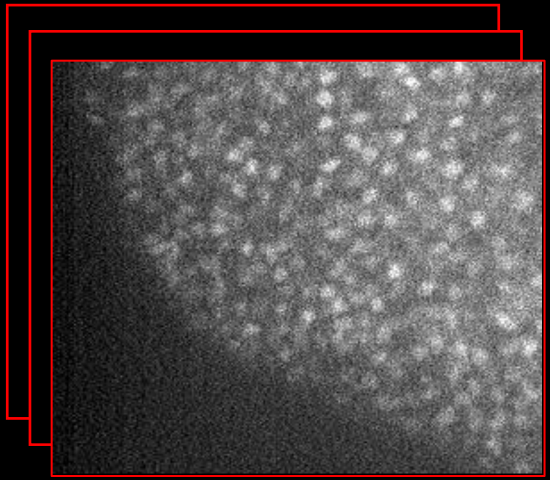
Zero centred noise

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

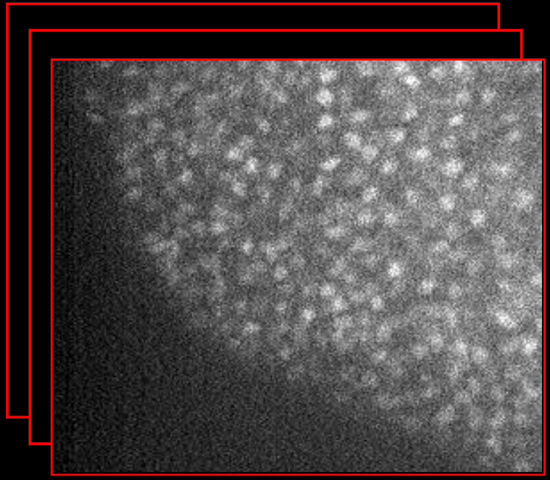
Expected value

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

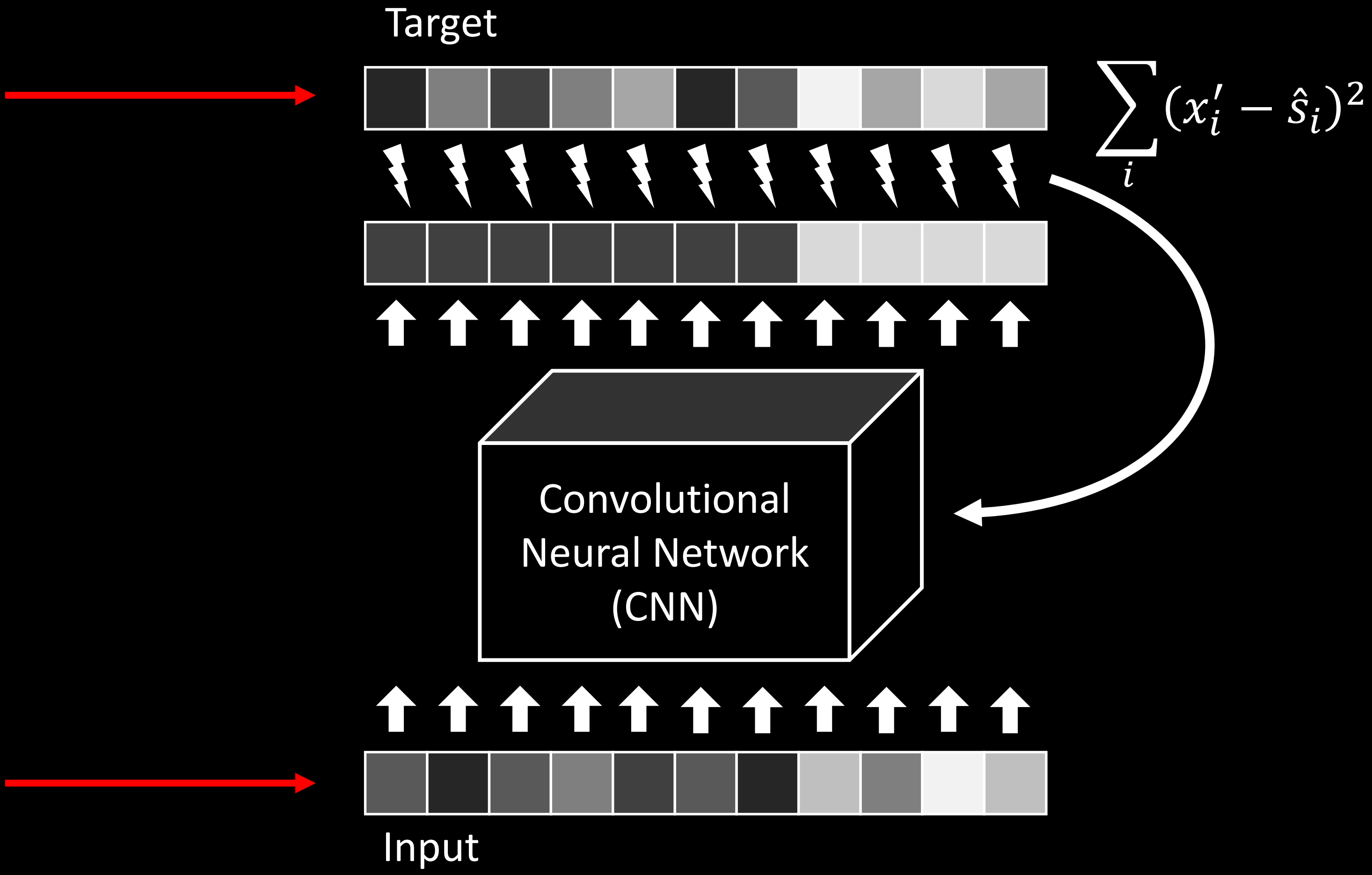
Noise2Noise training



Noisy image: $\boldsymbol{x}' = \boldsymbol{s} + \boldsymbol{n}'$



Noisy image: $x = \boldsymbol{s} + n$



Summary

- Traditional supervised training:
  - Tries to map noisy image to clean image.
  - Impossible: map to expected value of clean image.
  - Downside: requires clean images during training.

- Noise2Noise training:
  - Requires no clean data.
  - Tries to map noisy image to noisy image.
  - Impossible: also map to expected value of clean image.
  - Downside: Still requires image pairs.

Refer to the whole image sets, because of no access