By Bruce Lee

Applications of Deep Learning to Different Optical Systems

Optical Information System Laboratory

李杰恩

Chieh-En Lee (Bruce)

celee@nctu.edu.tw

Department of Photonics
National Chiao Tung University
30010 Hsinchu, Taiwan

Presentation for The Seminar of Computational Optics Group, 2020

Introduction to Optical Information System Laboratory

Advisor

Chung-Hao Tien (Peter)

chtien@nctu.edu.tw

Research Interests

- Optical system
- Digital image processing
- Computational imaging
- Machine vision
- Machine learning

Current Projects

Most of our research projects are dependent to the companies in Hsinchu Science Park.

Today

- Introduction to Optical Information System Laboratory

- Deep Learning for Different Optical Systems
  - Optical Coherence Tomography System
  - Phase Retrieval System
  - Instance Segmentation
  - Image Pre-correction System

- Summary

The world around you

- As a human, we are able to transform perceptions (vision, olfaction, hearing, …) into information (words, images, concepts, …).

Cat

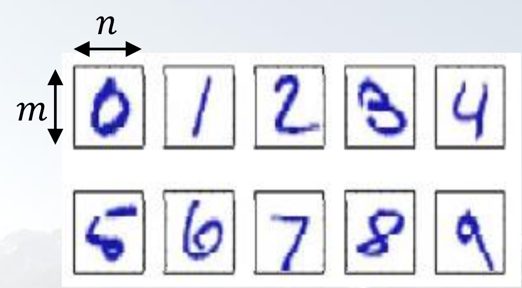
發大財

- **How does your brain translate this information?**
- **How does we introduce this world to our baby?**

The world inside the machine

- To a machine, they are just data with different combinations of “0101…”

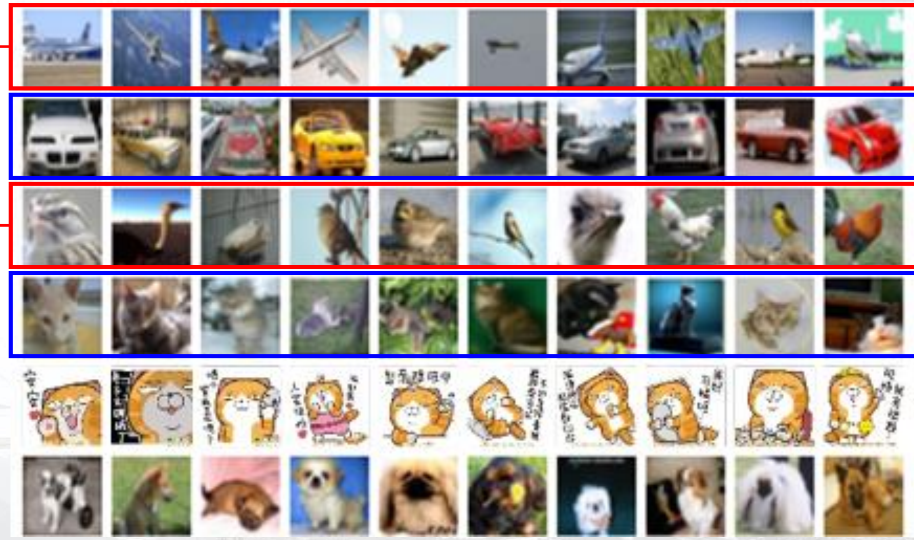
Handwritten digits

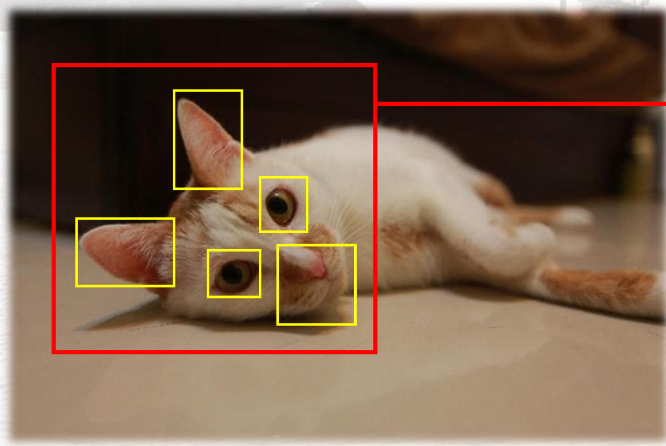


Each image has $m \times n$ pixels.

Machines do not have the ability to tell you what it means.

Vision is the key for human accessing the world





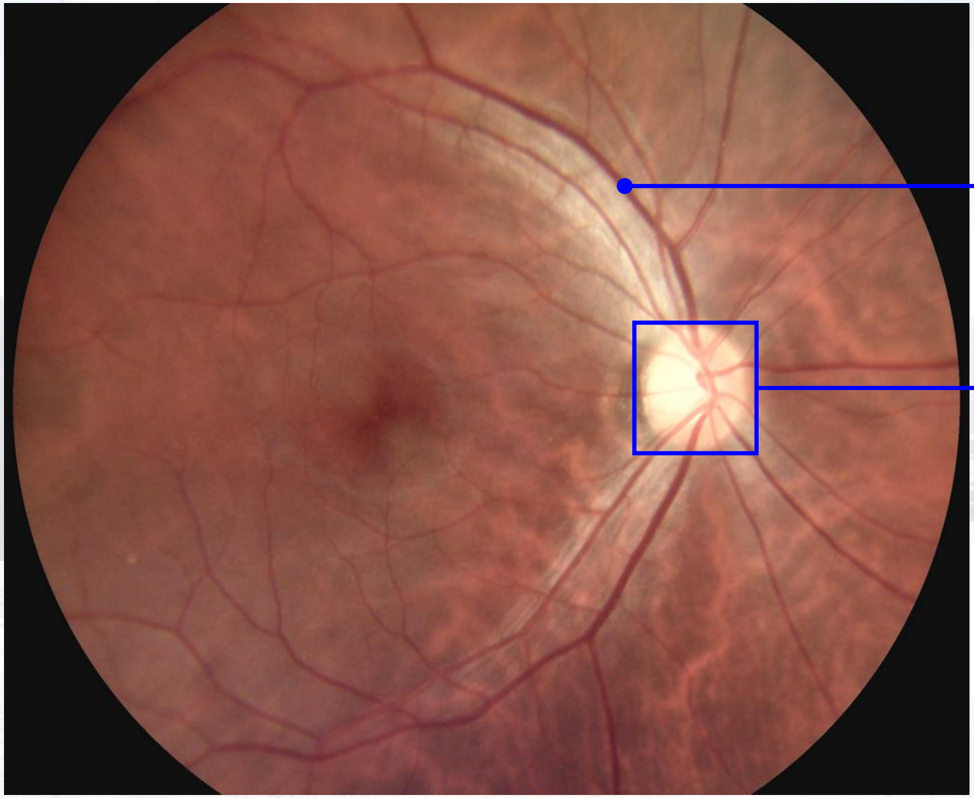
There is a face which consists of

- ✓ 2 cat-like ears
- ✓ 2 cat-like eyes
- ✓ 1 cat-like mouth

⇨ This is a cat's face.

Can we expect machine to do this?

Convolution makes it possible



A line-shape filter can receive a peak response

A circle-shape filter can receive a peak response

These are human-defined filter!

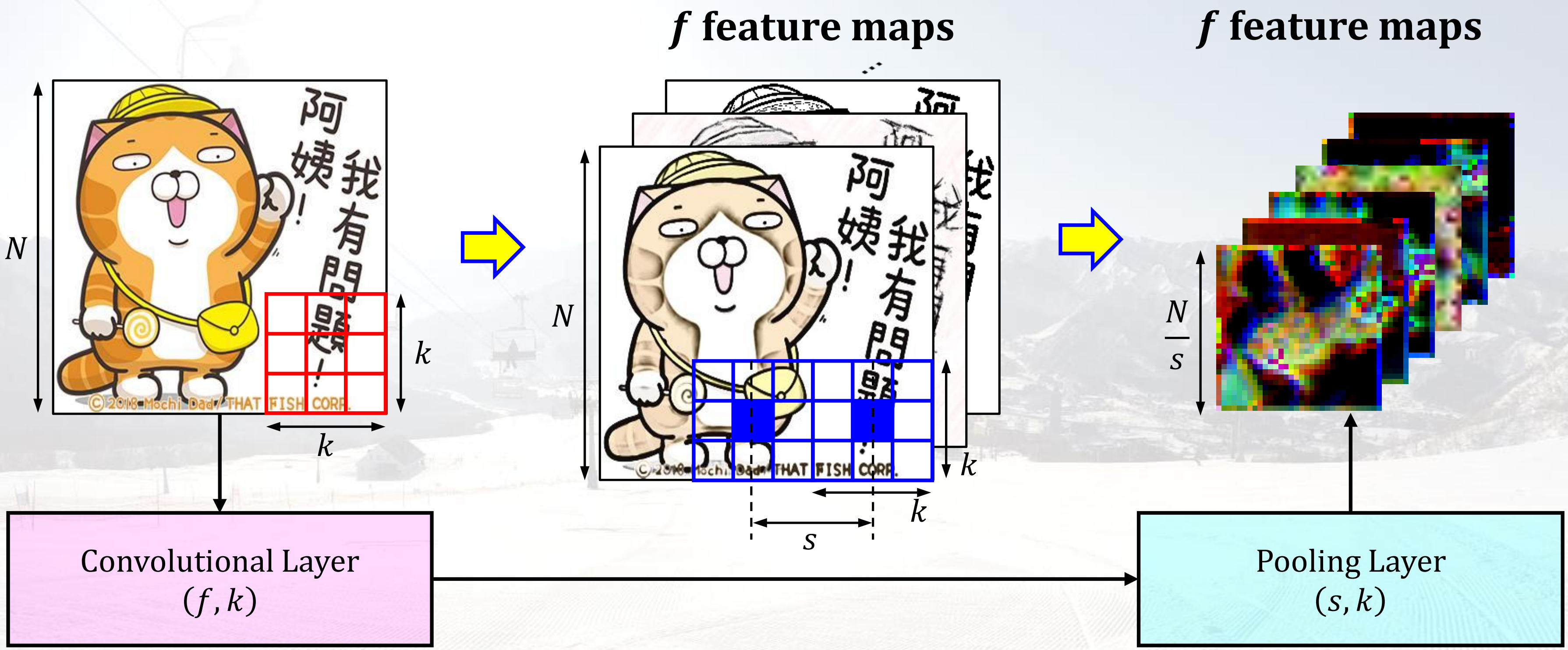
This time, we don't tell machines what kind of filters they should use. We don't even tell machines what they should learn.

By Bruce Lee

Fundamentals: convolution and pooling

Optical Information System Laboratory

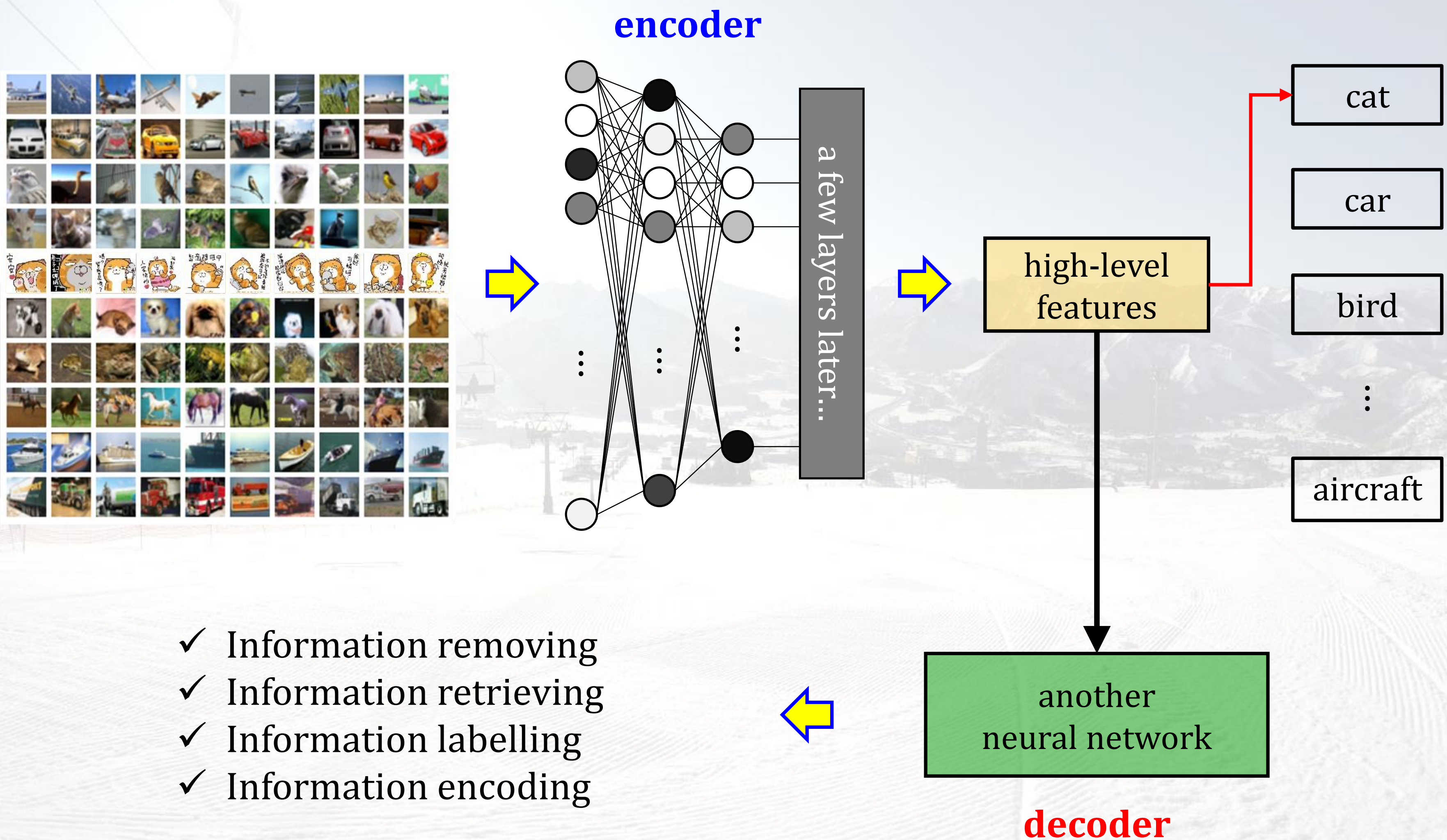
- Illustration of each blocks



f: number of filters
k: kernel size

s: stride
k: kernel size

An encoding/decoding process

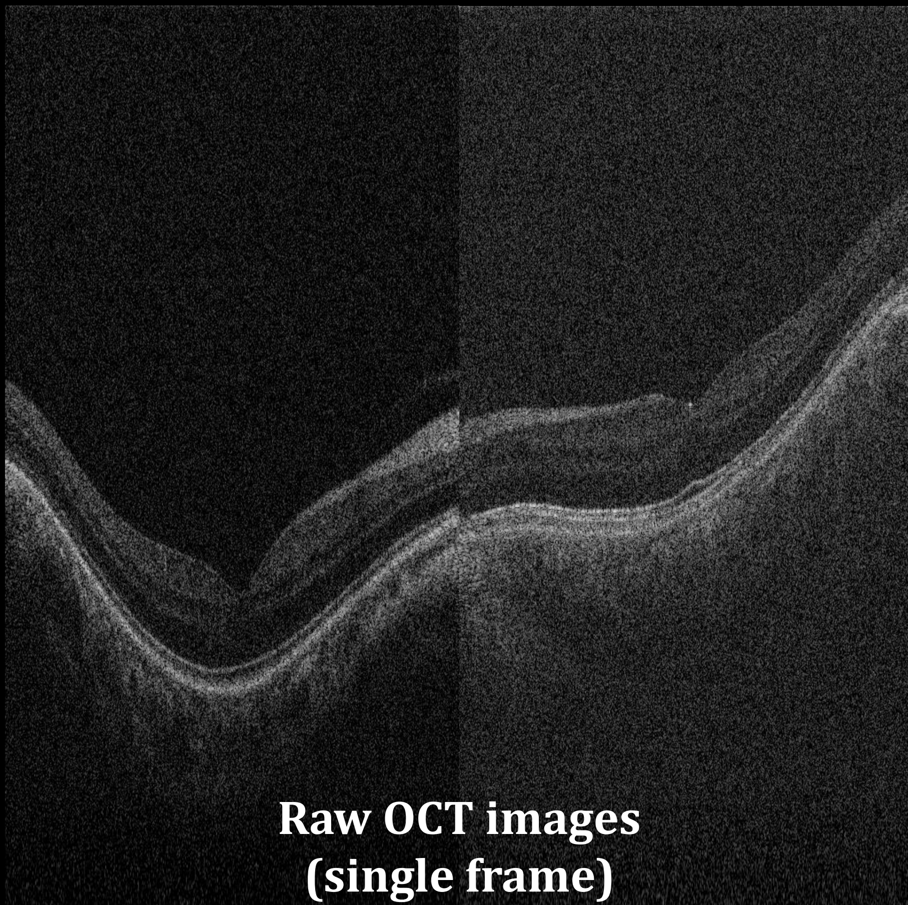


- ✓ Information removing
- ✓ Information retrieving
- ✓ Information labelling
- ✓ Information encoding

Application I

Deep Learning for
Optical Coherence Tomography System

Noise suppression for OCT images



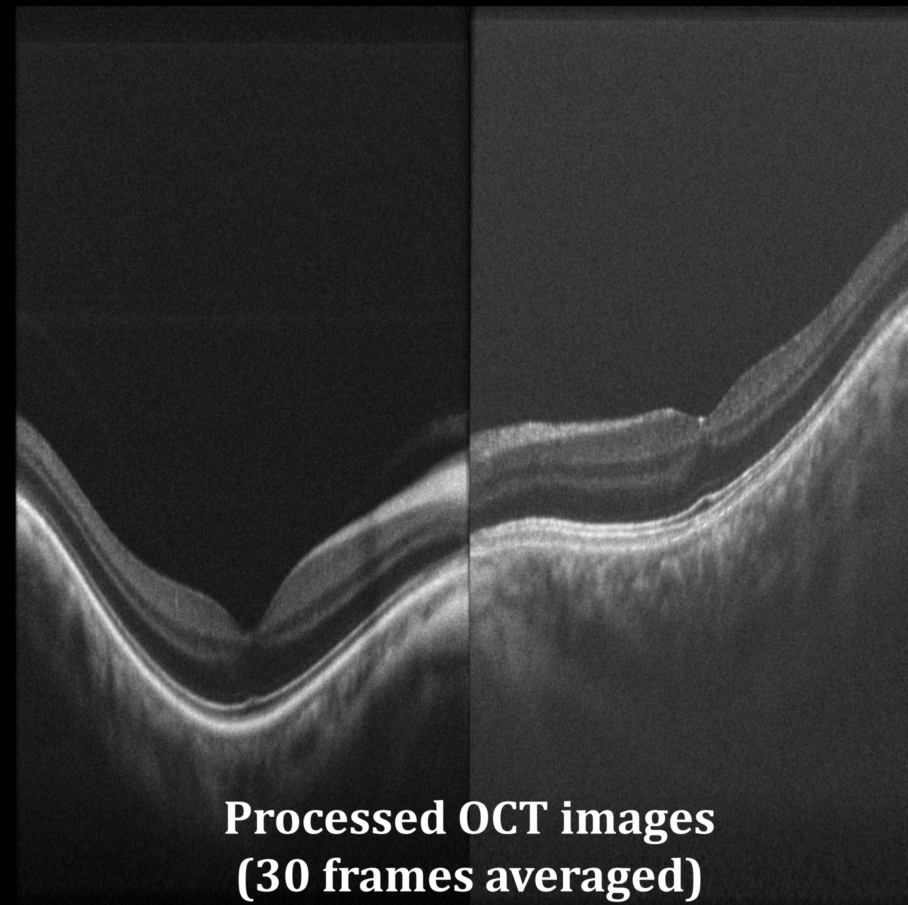
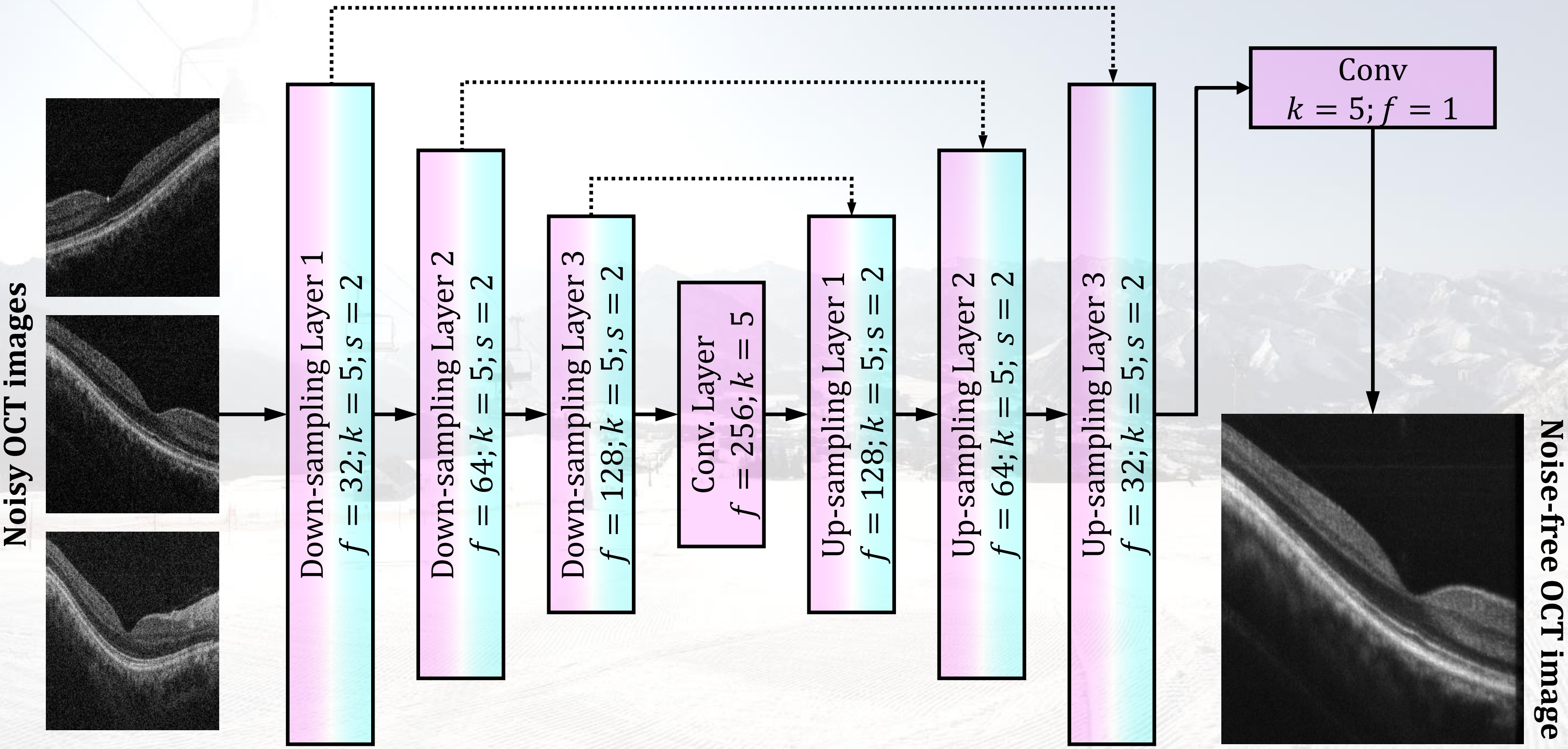


image registration and averaging
(30 frames or more)

execution time: $\sim 10^1$ s

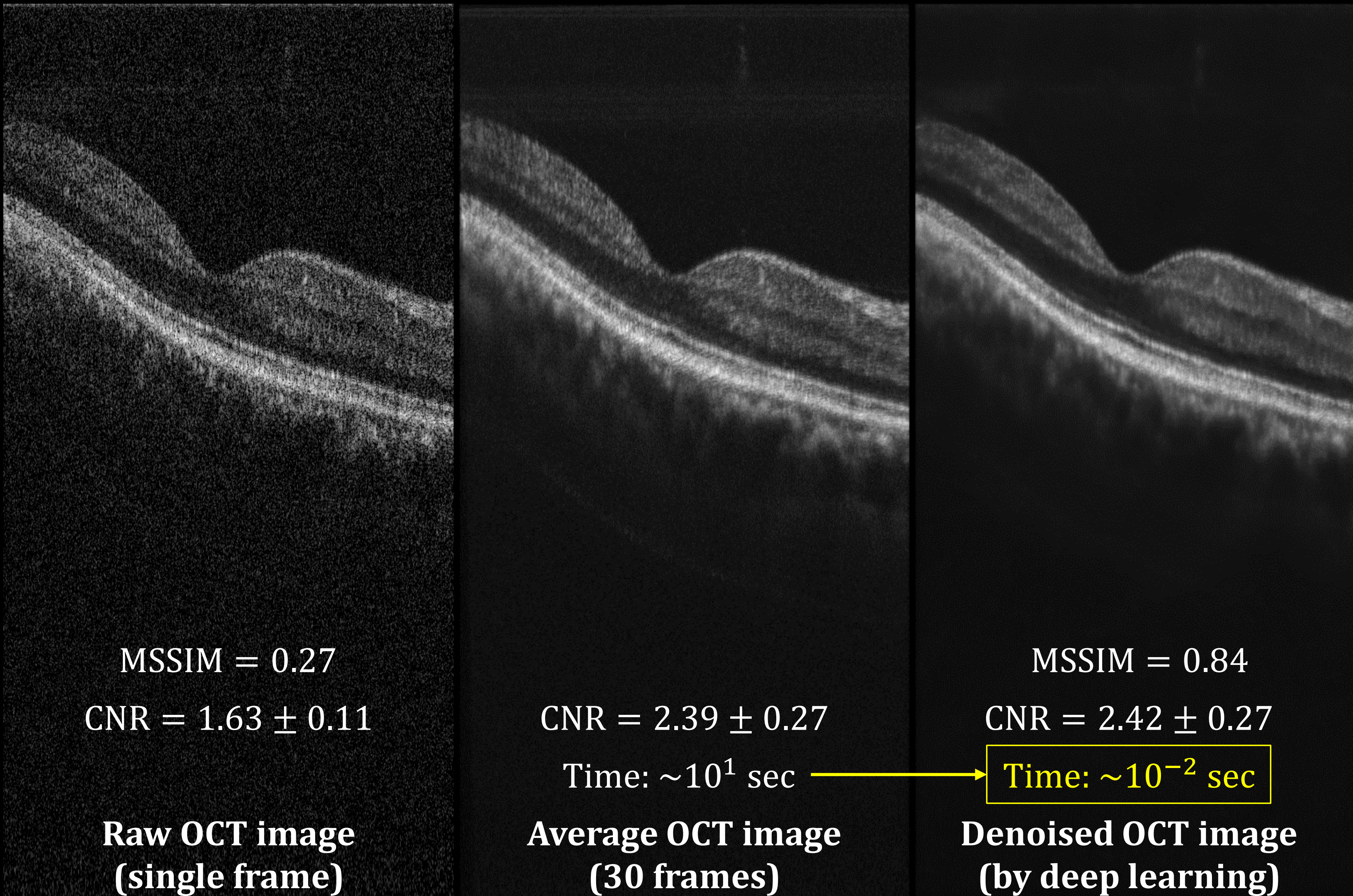
Deep learning for the noise suppression of OCT images



Training GPU: NVIDIA GeForce RTX 2080 Ti

Memory requirement: less than 1 GB for executing the trained model

Results


MSSIM = 0.27
CNR = 1.63 ± 0.11
Raw OCT image
(single frame)
CNR = 2.39 ± 0.27
Time: ~10¹ sec
Average OCT image
(30 frames)
MSSIM = 0.84
CNR = 2.42 ± 0.27
Time: ~10⁻² sec
Denoised OCT image
(by deep learning)

Application II

Deep Learning for Phase Retrieval System

Phase retrieval

- Two kinds of phase information

Fourier phase

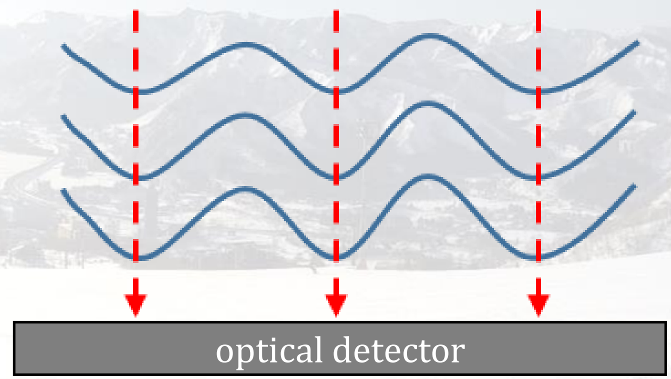
$|F(\mathbf{k}_r)| \longleftarrow I(\mathbf{r})$

$\mathcal{F}\{I(\mathbf{r})\} = F(\mathbf{k}_r)\exp[\boxed{j\phi(\mathbf{k}_r)}]$

crucial information of image

Local phase (wavefront)

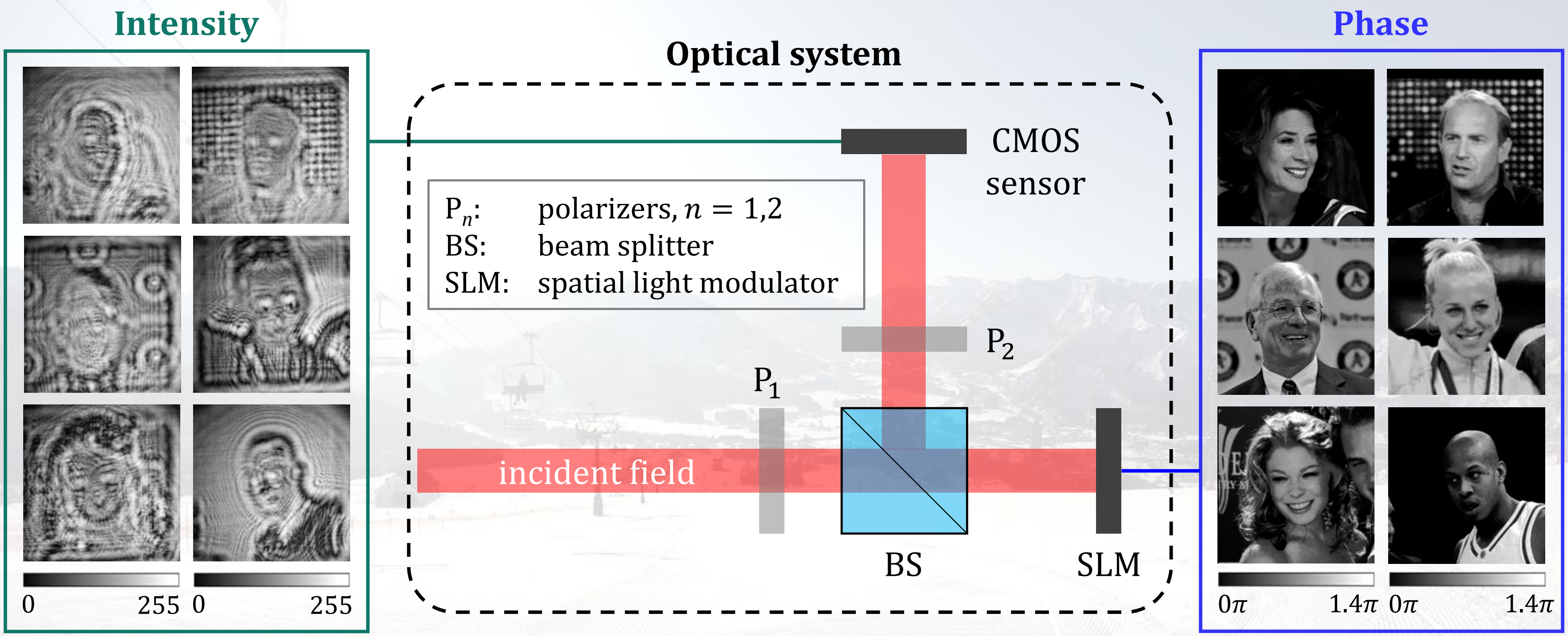
$E(\mathbf{r}) = A(\mathbf{r})\exp[-j\phi(\mathbf{r})]$



$I(\mathbf{r}) = |E(\mathbf{r})|^2 = |A(\mathbf{r})|^2$

the phase information is lost

Deep learning for phase retrieval

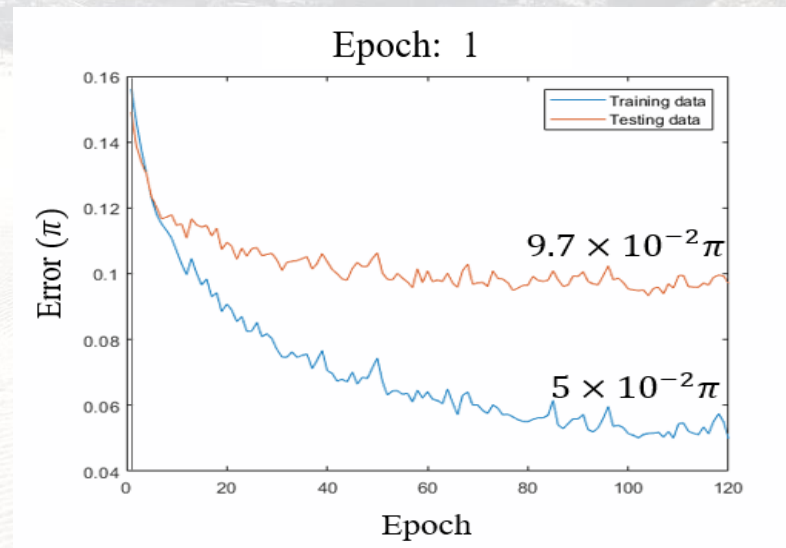


$$I(\mathbf{r}) \xrightarrow{\text{recovery the phase from intensity}} \phi(\mathbf{r})$$

~~Iterative Fourier transform algorithm~~ → **Deep learning**

Ref: A. Sinha, J. Lee, S. Li, and G. Barbastathis, Optica **4**, 1117-1125 (2017).
G. Barbastathis, A. Ozcan, and G. Situ, Optica **6**, 921-943 (2019).

Results



Ref: Y. Yao, *et al.*, Proc. SPIE **11197**, 111970N (2019).

Application III

Deep Learning for Instance Segmentation

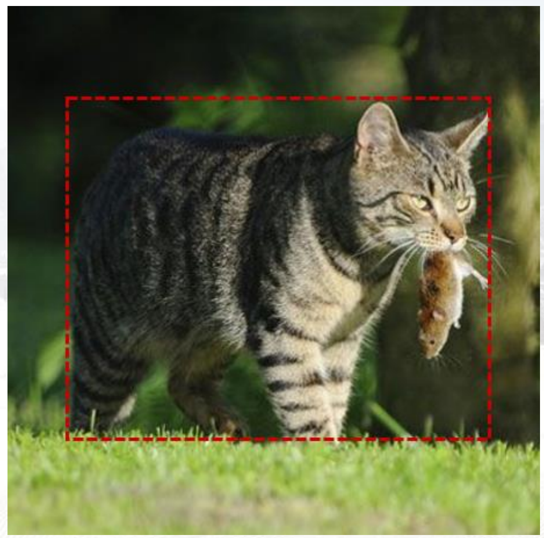
Deep learning for instance segmentation

- Region-based convolutional neural network (R-CNN) and its families

CNN

Classification

R-CNN & Faster R-CNN



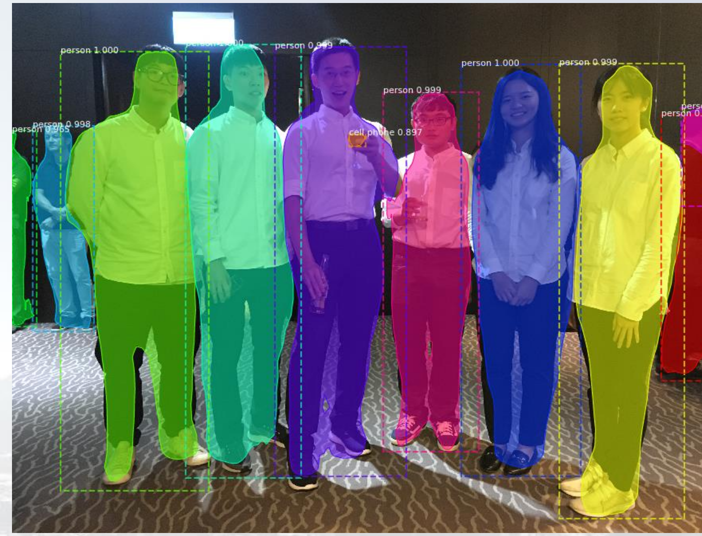
Classification + Object detection (bounding box)

Mask R-CNN

Classification + Segmentation (mask)

Ref: S. Ren, *et al.*, IEEE Trans. Pattern Anal. Mach. Intell. **39**, 1137–1149 (2017).
K. He, *et al.*, in Proc. IEEE Int. Conf. Comput. Vis. (2017), pp. 2980–2988.

Preliminary result for pictures

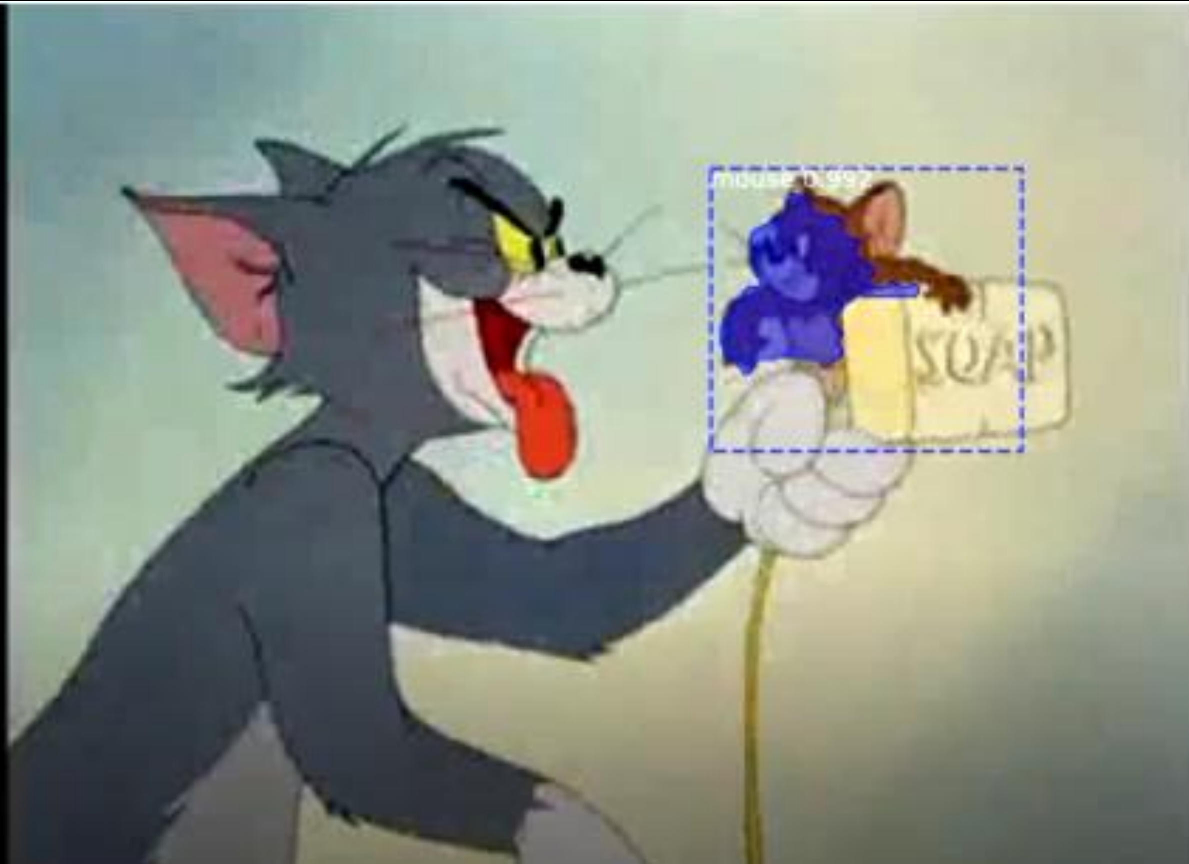


Original Image | **Results of Mask R-CNN**

Courtesy of Cable News Network

Preliminary result for video

- Training on custom dataset (training:testing = 35:5)
- Video specification: 640x360, 20 fps



Application IV

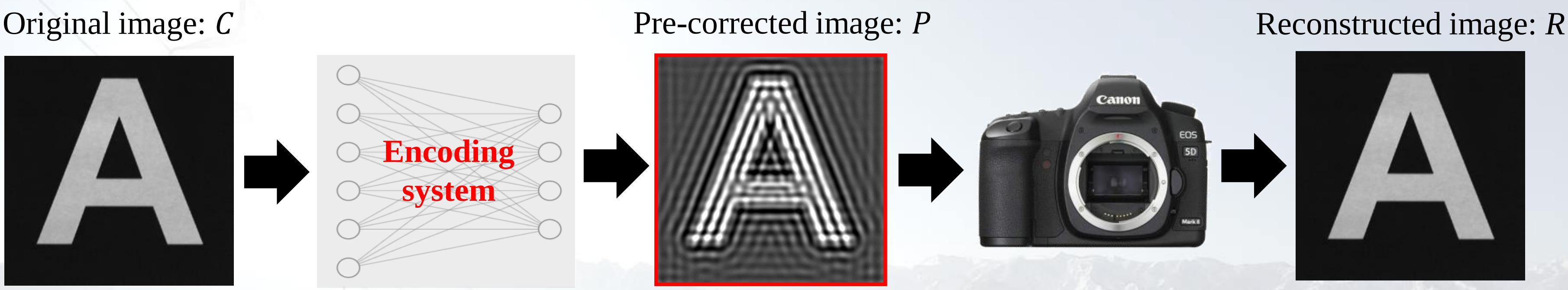
Deep Learning for Image Pre-correction

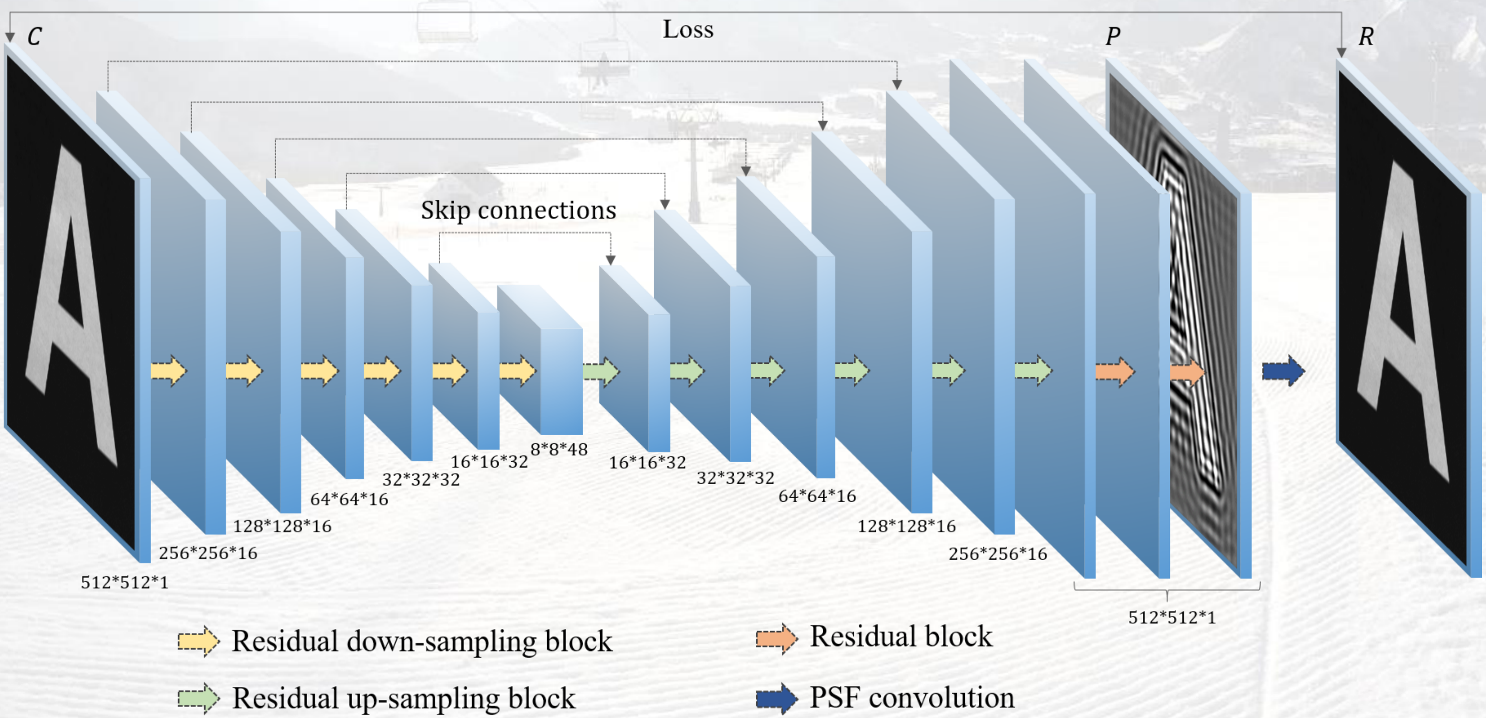
The idea of image pre-correction

Our goal**: *watch a display without wearing glasses***

Flowchart of image pre-correction

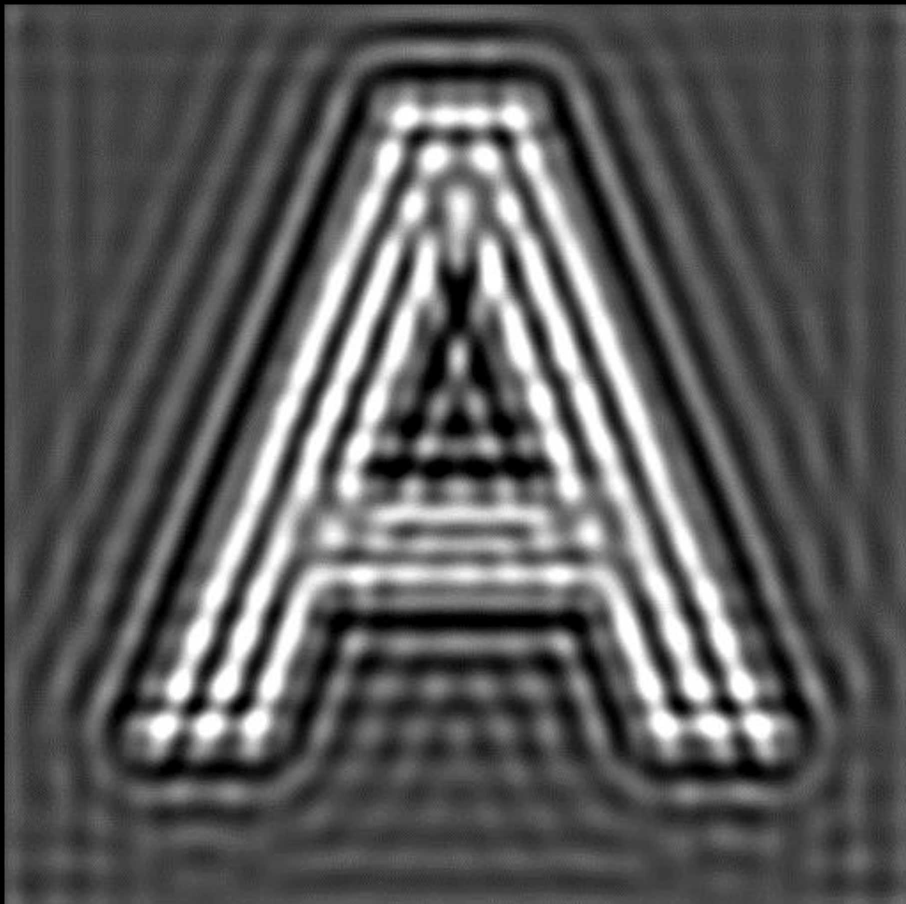
Pre-corrected image generation





- Loss function: MSE
- Optimizer: ADAM
- Learning rate: 10^{-5}

Result of image pre-correction

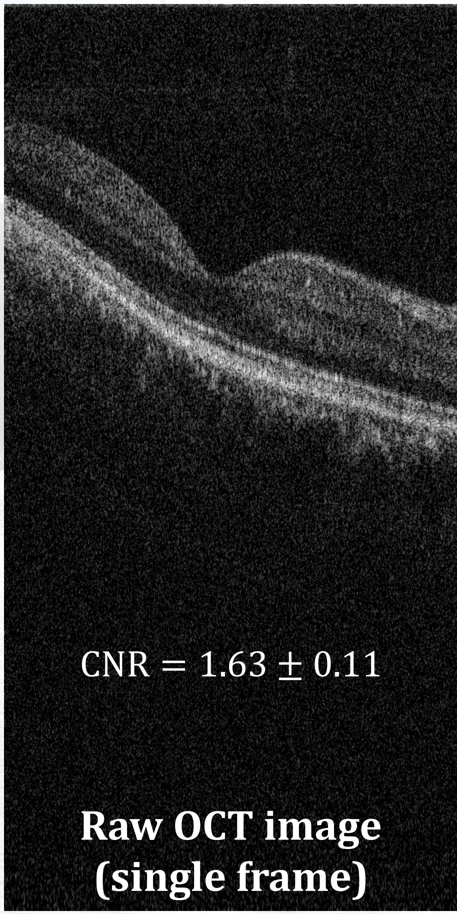


Ref: Chu, Chih-Yu. Myopia pre-corrected image algorithm based on deep learning. 2020. NCTU, MA Thesis.
Yao, Yun-Zhen. Estimation and deep learning-based pre-correction for aberrated imaging system. 2020. NCTU, MA Thesis.

Summary

Summary

- Perception is an encoding & decoding process. This time, we just let machine learn how to do it by itself, without any designate filter by human.



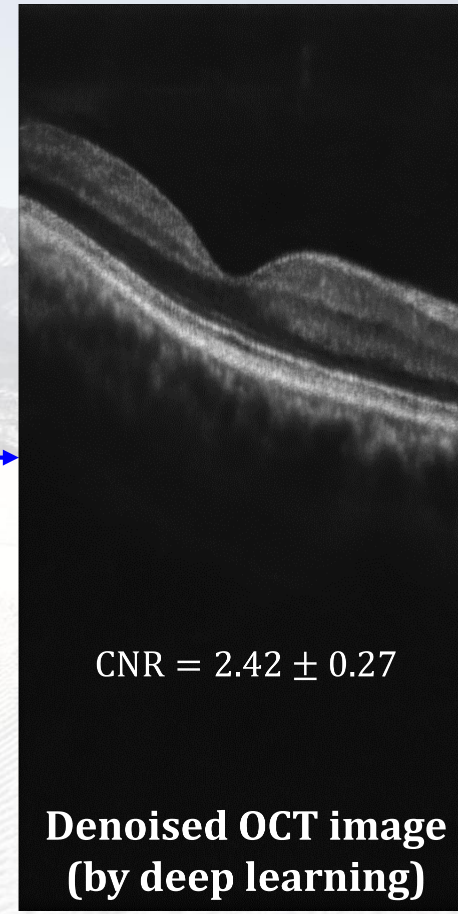
Raw OCT image (single frame)

Once it has been well-trained, we can use it to…

Information removing

encoding → high-level features → decoding

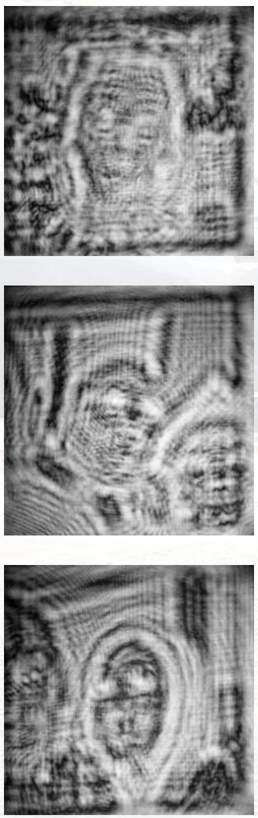
3 orders improvement on execution time ($\sim10^{1} \rightarrow \sim10^{-2}$ s)



Denoised OCT image (by deep learning)

Summary

- Perception is an encoding & decoding process. This time, we just let machine learn how to do it by itself, without any designate filter by human.



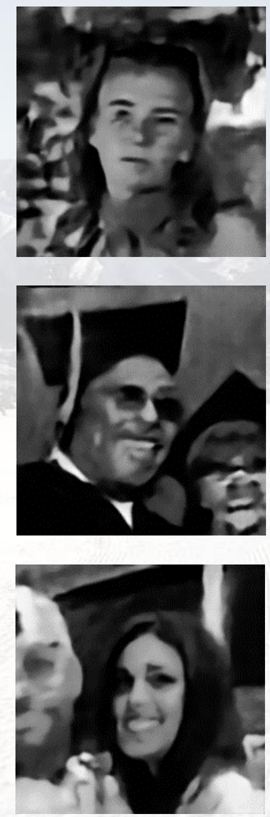
Intensity

Once it has been well-trained, we can use it to…

Information retrieving

encoding → high-level features → decoding

Testing error: $\sim 10^{-1}\pi$



Phase

Summary

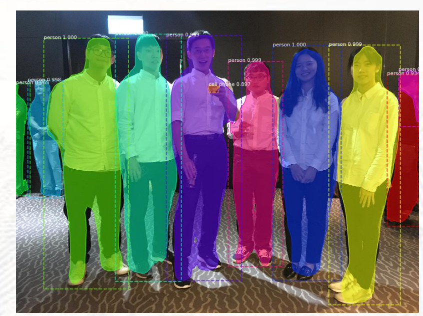
- Perception is an encoding & decoding process. This time, we just let machine learn how to do it by itself, without any designate filter by human.

Once it has been well-trained,
we can use it to…

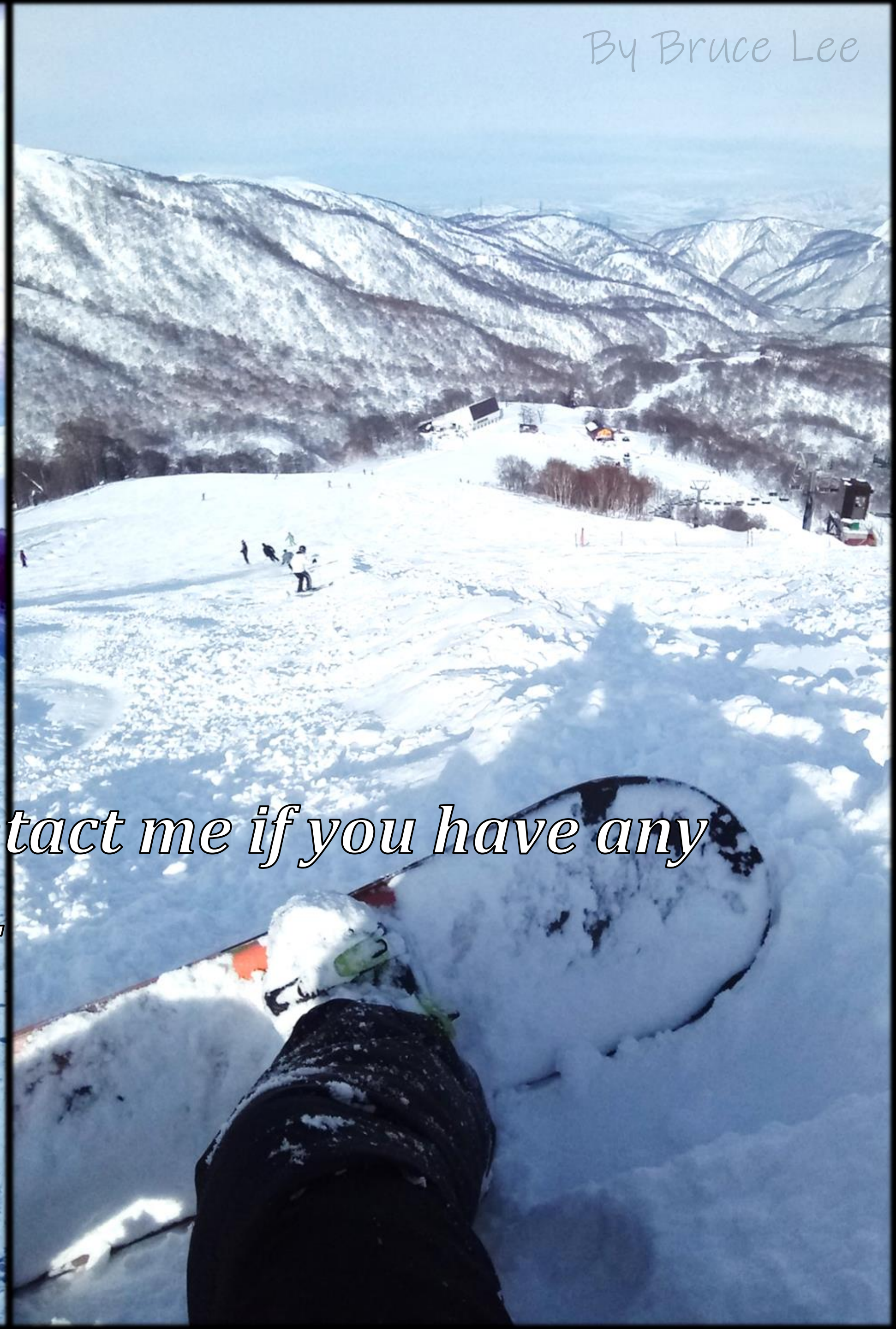
Information labelling

Original image

encoding → high-level features → decoding



Instance segmentation



Thank you!

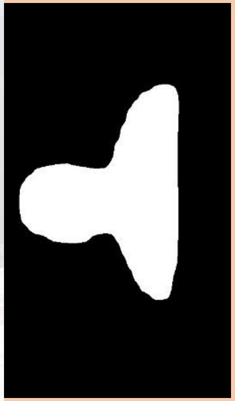
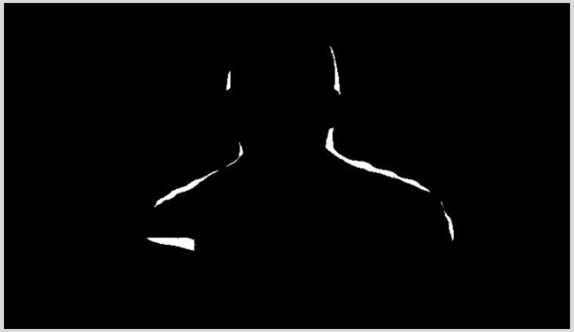
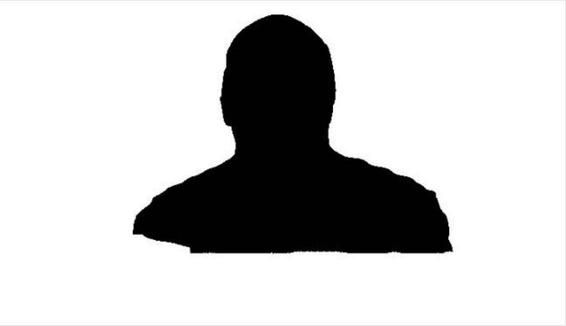
Please feel free to contact me if you have any question or comment.

Chieh-En Lee, Bruce:

celee@nctu.edu.tw


THE DARK KNIGHT
LEGENDARY

Appendix: confusion matrix for MRCNN

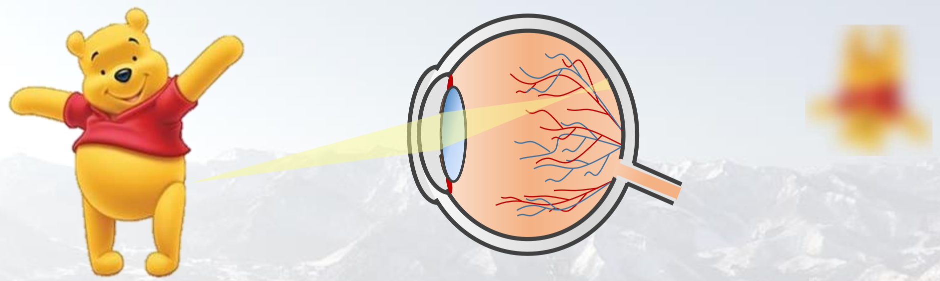
		**Ground Truth** 	
		Positive	Negative
Predicted Result 	Positive	 TP: 57629 pixels	 FN: 2482 pixels
	Negative	 FP: 2731 pixels	 TN: 235238 pixels

Appendix: Mathematical model of image pre-correction

The image formation model can be depicted as following:

$$o(x,y) = i(x,y) \otimes k(x,y)$$

$i(x,y)$ clear image | $k(x,y)$ point spread function | $o(x,y)$ blurry image

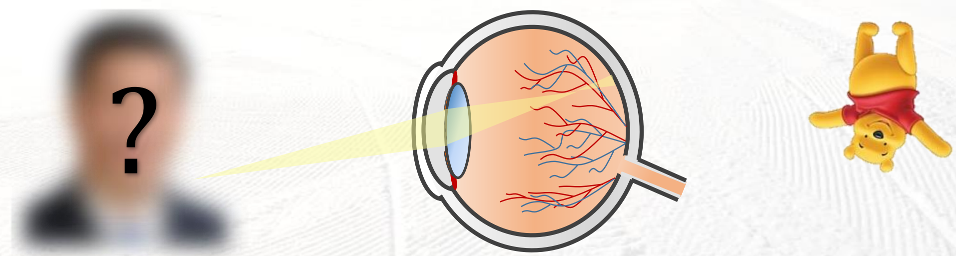


Assume that there exists a ***pre-corrected image*** p which satisfies:

$$i(x,y) = p(x,y) \otimes k(x,y)$$

Then, we should get p from the following equation:

$p(x,y)$ pre-corrected image | $k(x,y)$ point spread function | $i(x,y)$ clear image



$$p(x,y) = i(x,y) \otimes^{-1} k(x,y)$$

where $\otimes^{-1}$ is the ***deconvolution operator***.

Appendix: Flowchart of the experiment of pre-correction

Data acquisition

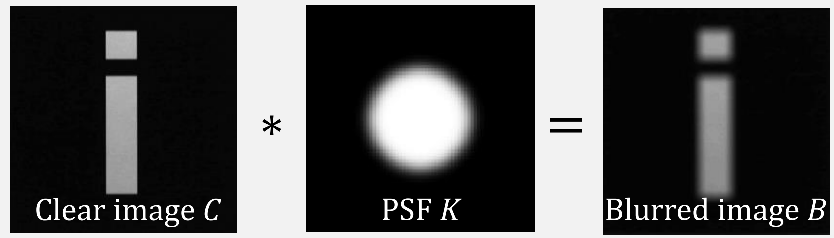
01 Image collection

02 Image registration

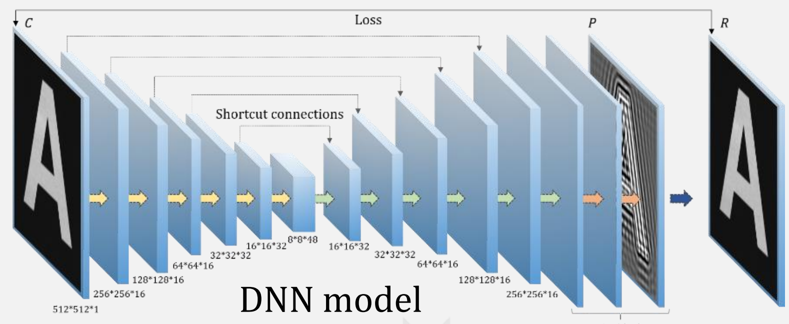
03 Response curve

04 Flat field correction

PSF Estimation



Pre-corrected image generation



Pre-corrected system validation and comparison with traditional methods

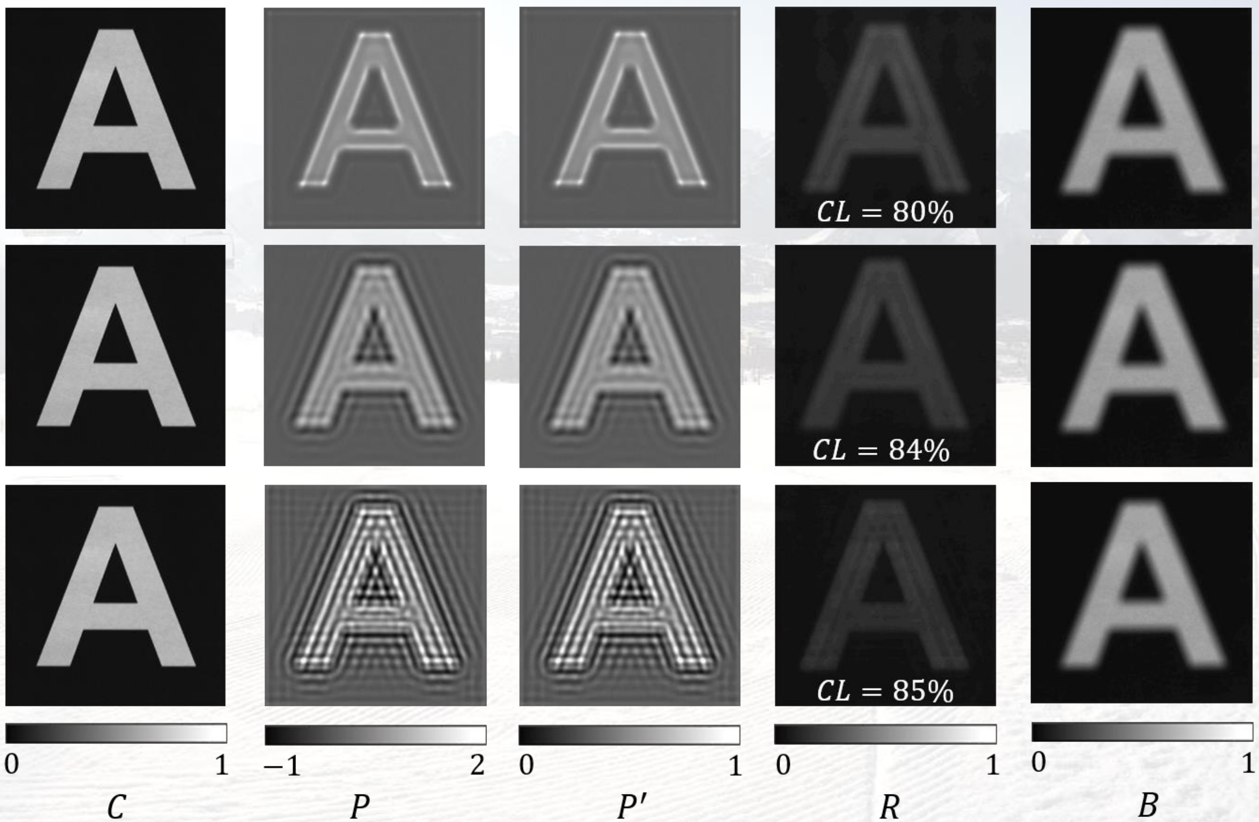
Appendix: Comparisons between different methods

- Lucy-Richardson deconvolution with Iteration $= 10$
- Wiener deconvolution with $\lambda = 10$
- DNN model with $\alpha = 1$ and epoch $= 400$
- In simulation, they have the same contrast loss of 66%.

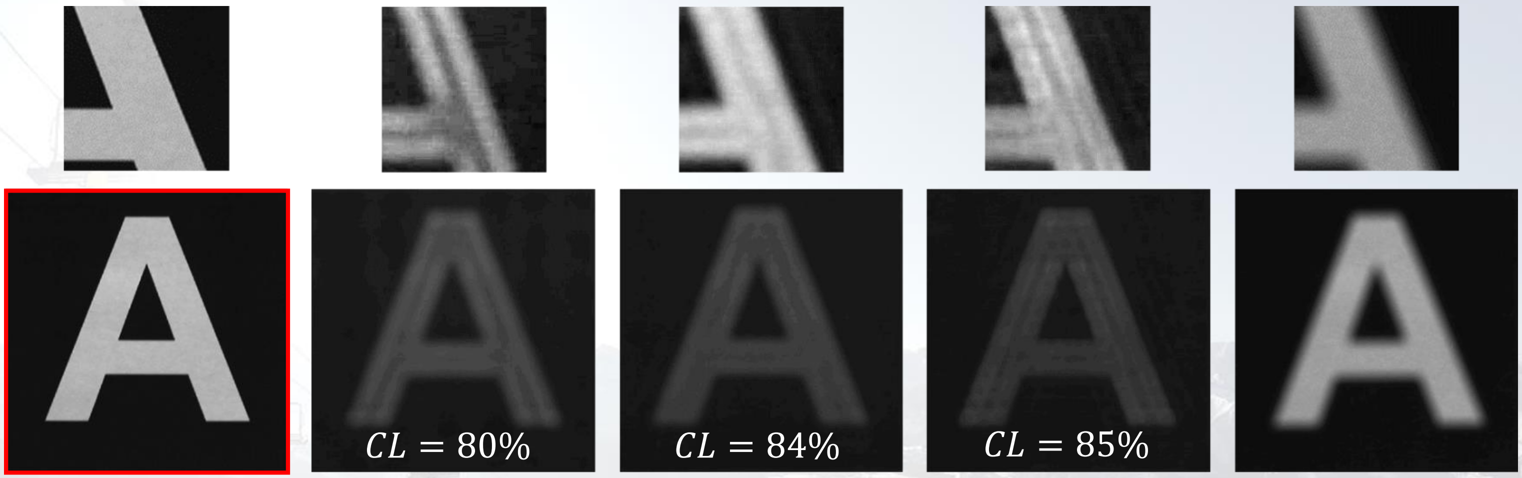
Lucy-Richardson deconvolution

Wiener deconvolution

Deep learning deconvolution



Appendix: Image quality assessment



Clear image | Lucy-Richardson deconvolution | Wiener deconvolution | DNN model | Blurred image

	Lucy-Richardson	Wiener	DNN model	Blurred image
RMSE	0.2210	0.2313	0.2405	0.0805
PSNR	13.1109	12.7132	12.3745	21.8733
SSIM	0.6753	0.6979	0.7097	0.8315
Information Entropy	2.5102	2.3228	2.2563	2.7288