deep imaging

Lecture 1: Machine Learning and Imaging in a Nutshell

Machine Learning and Imaging

BME 590L

Roarke Horstmeyer

What is an image?

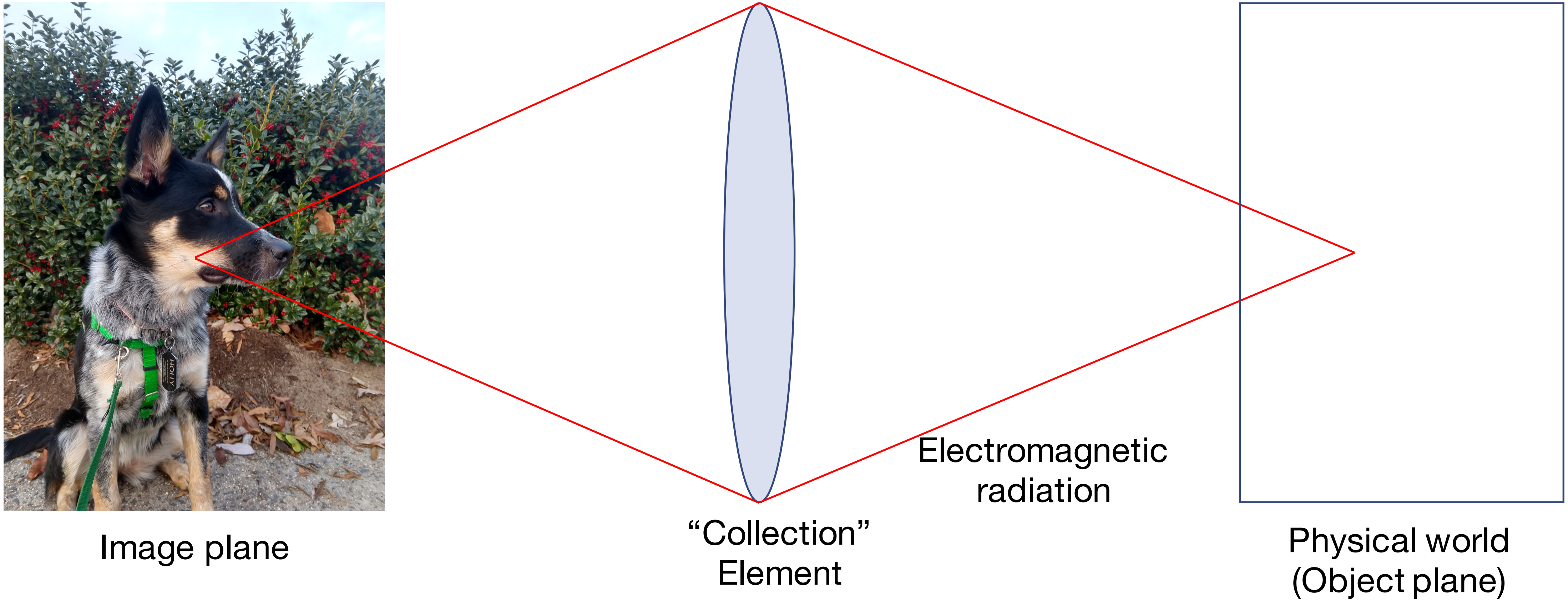
What is an image?

1. “Qualitative” Interpretation

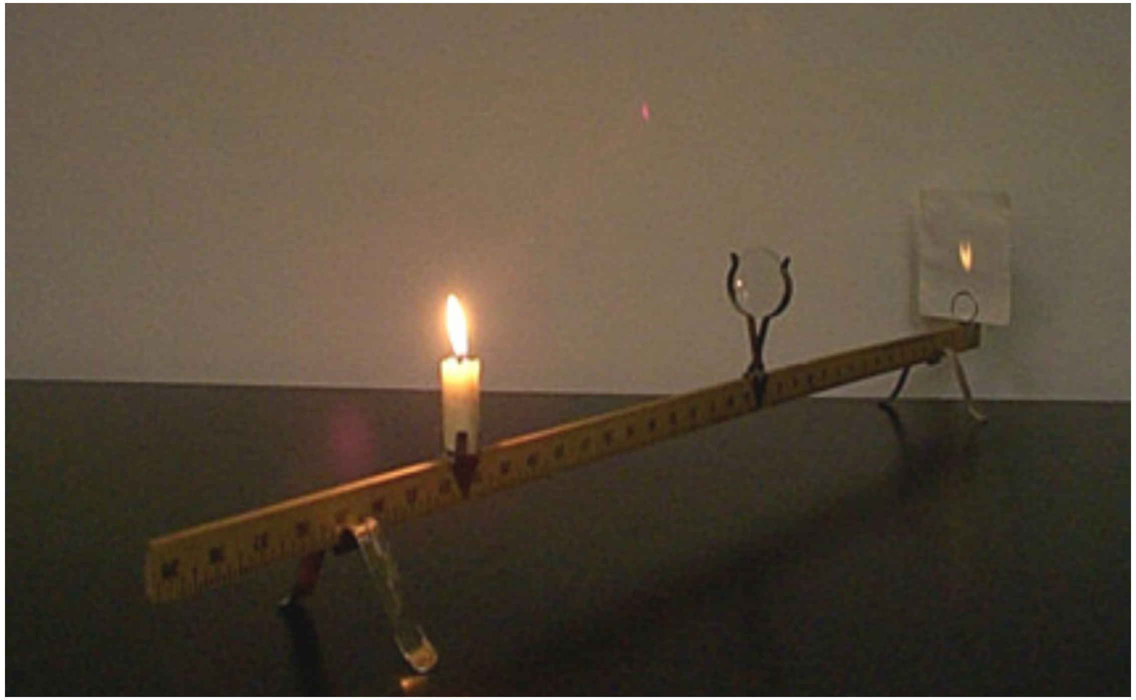
- A re-creation of a visual scene
- A visible impression
- A mental representation or idea

What is an image?

2. “Physical” Interpretation



What is an image?



What is an image?

2. "Physical" Interpretation

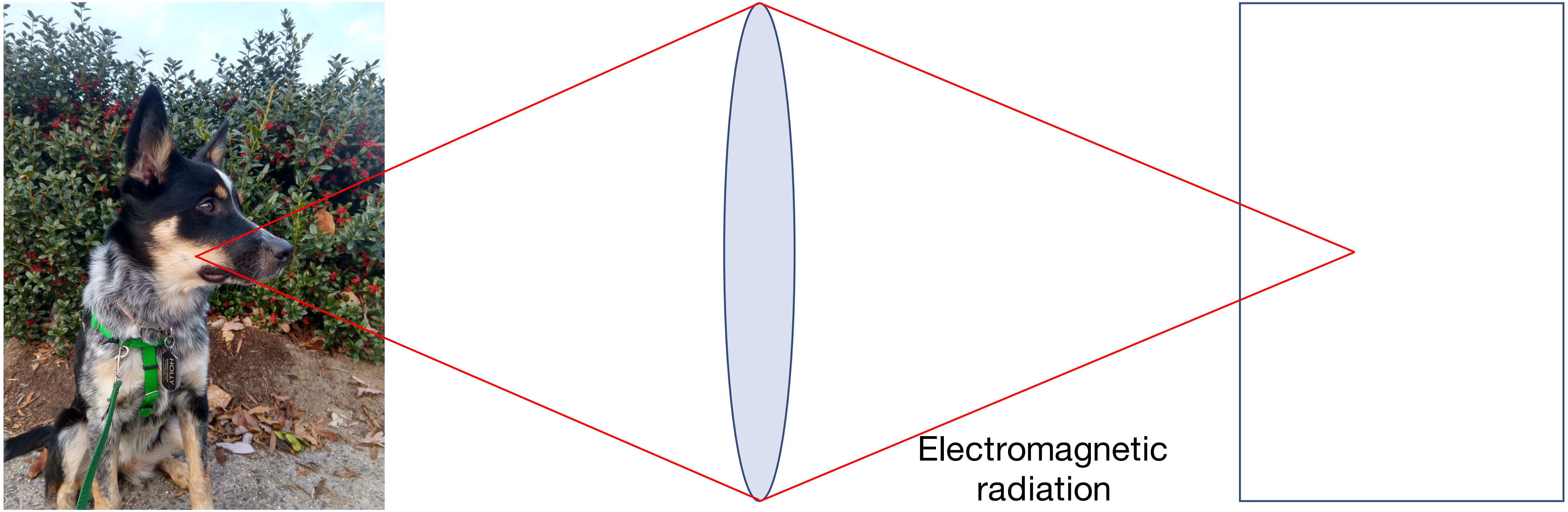


Image plane

"Collection"
Element

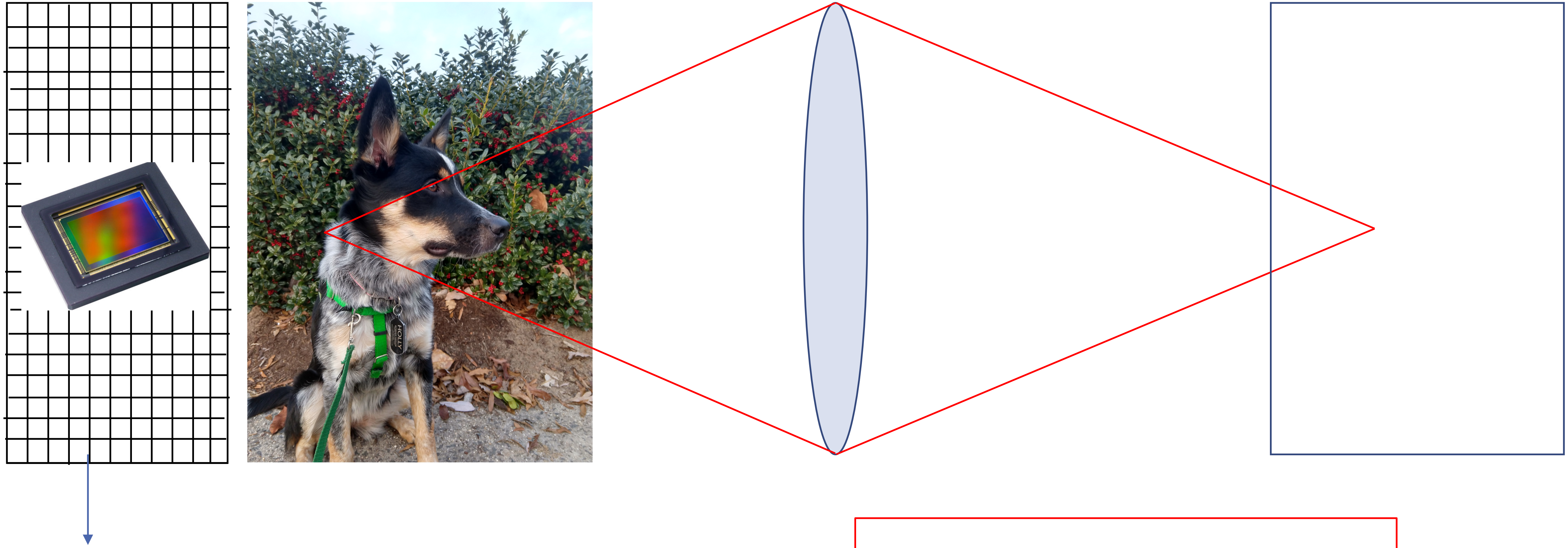
Physical world
(Object plane)

Continuous signal: $\boxed{I(x,y), (x,y) \in (-\infty, \infty)}$ (Physical wave)

What is an image?

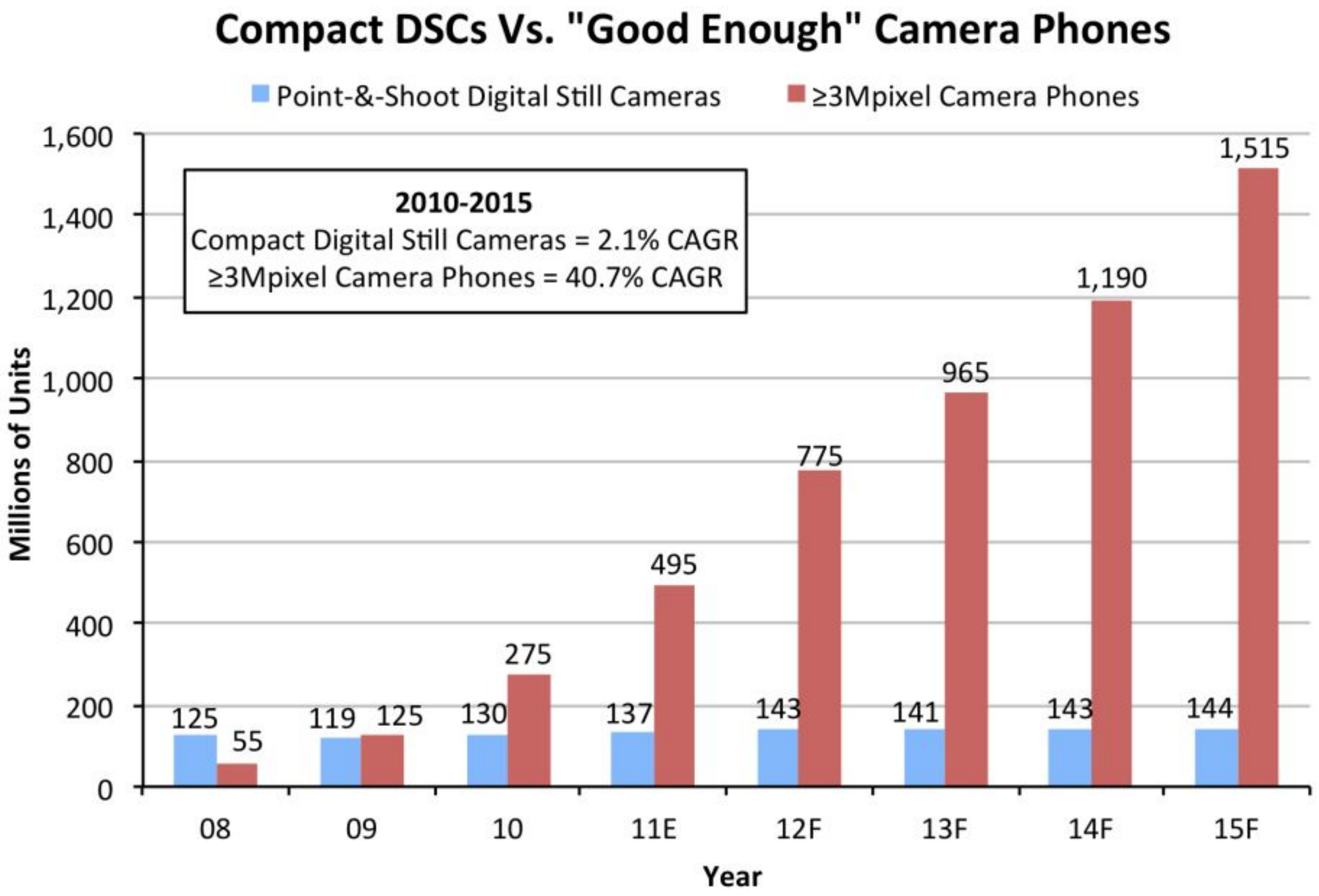
3. “Digital” Interpretation

n x *m* array



Photons to electrons → Digitazation → *Discrete signal*

$$\boxed{I_s(x,y), (x,y) \in Z^{n\times m}}$$



A guess: there are now more discretized images than continuous images in the world

Upcoming digital image sensor arrays can now also detect depth...

OUSTER

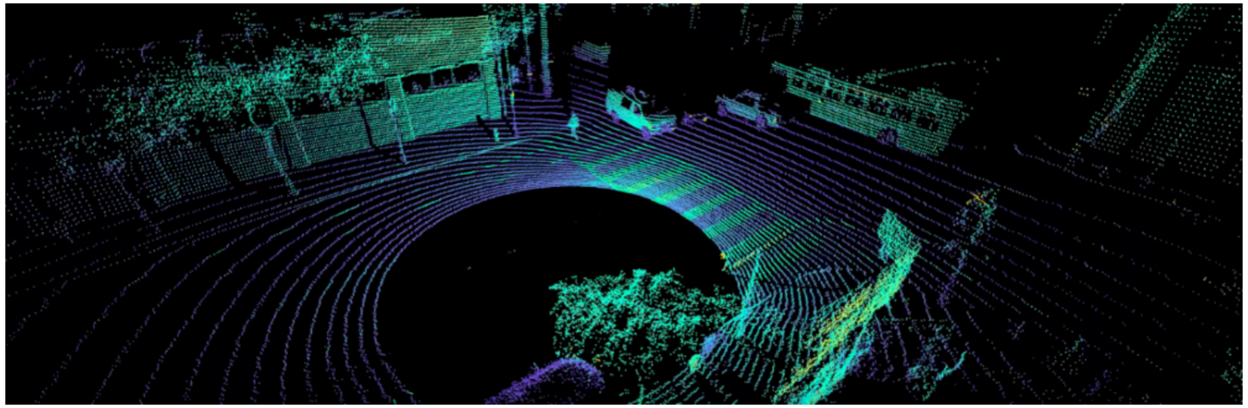
HOME PRODUCTS DOWNLOADS COMPANY CAREERS MEDIA CONTACT TALK TO SALES

A Novel CMOS ASIC with SPAD Detectors

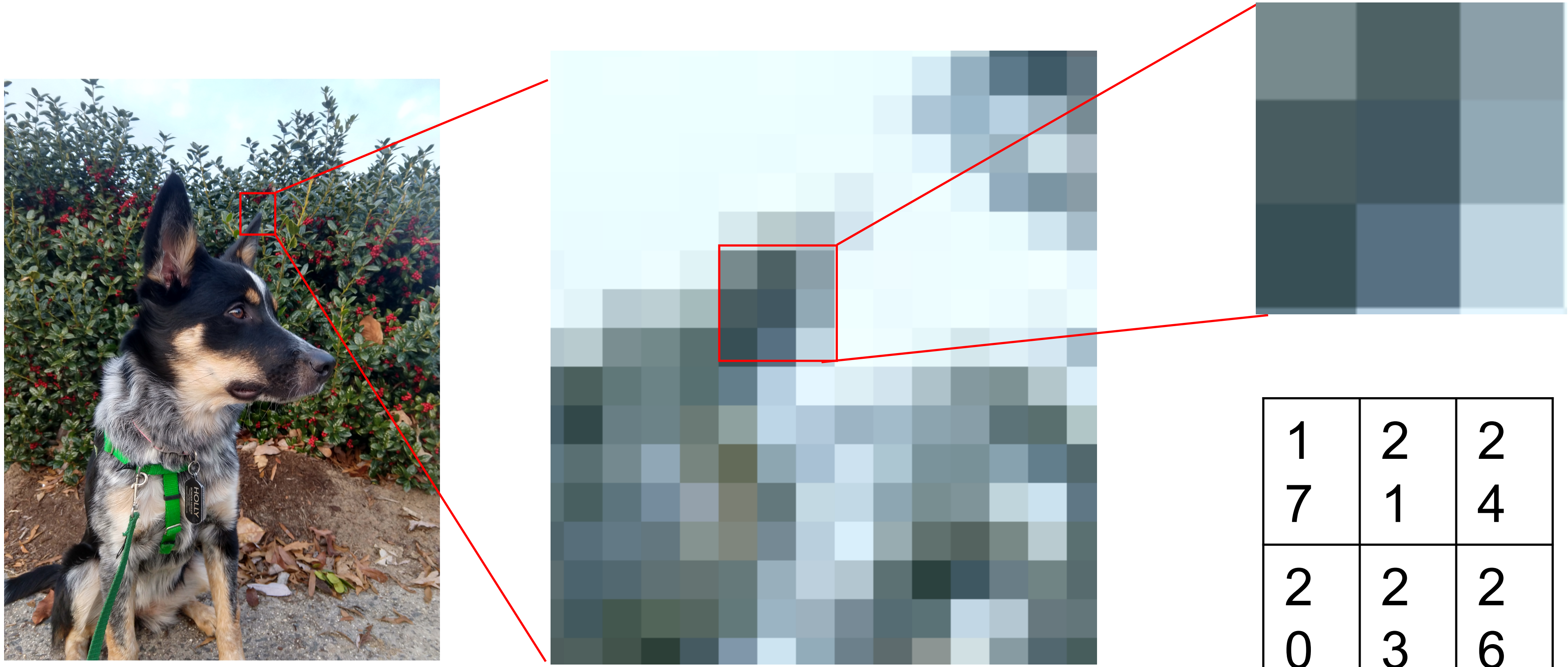
The second chip in our flash lidar is our custom designed CMOS detector ASIC that incorporates an advanced single photon avalanche diode (SPAD) array. SPADs are a relatively new type of photo sensor that creates a binary pulse when a photon is detected as opposed to a traditional camera pixel which generates an analog signal that varies continuously with the amount of light on the detector. SPADs have single photon sensitivity, low noise, and extremely good timing resolution (jitter between 10 ps and 100 ps is typical. That's picoseconds - mere trillionths of second!), all of which make them perfect for detecting and timing the ultra short laser pulses in our lidar.

Like VCSELs, CMOS SPADs have many practical advantages over the traditional approaches from other lidar manufacturers. Most important is that they can be directly integrated in a CMOS wafer, which makes it possible to incorporate massive amounts of signal processing on the silicon die right next to the detectors.

As lidar resolutions and data rates continue to increase, on-chip signal processing is essential - the current OS1-64 detector is capable of counting and storing over **one trillion photons per second into on-chip memory**. This is a titanic amount of data, and we've included over **100 GMACs per second** (1 GMAC = 1 billion multiply accumulate operations) of signal processing logic in over 10 million transistors to ultimately produce the millions of 3D points per second that our customers use to drive cars, map environments, and identify obstacles.



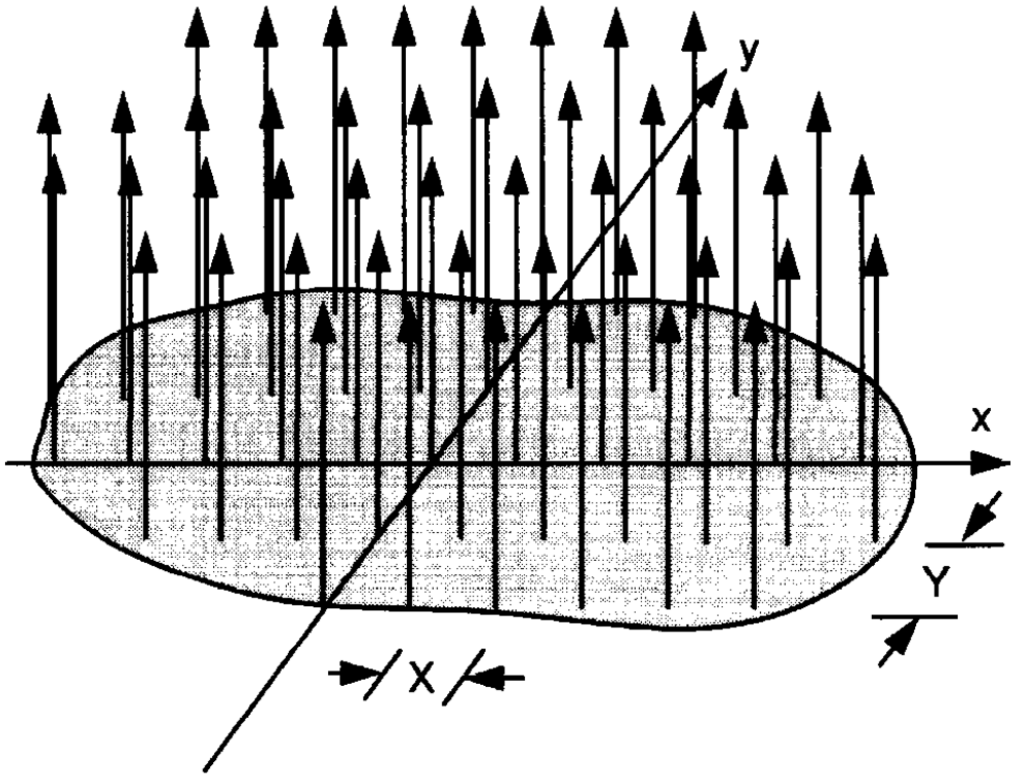
Images as matrices and vectors



17	21	24
20	23	26
22	25	79

Continuous versus discrete representation

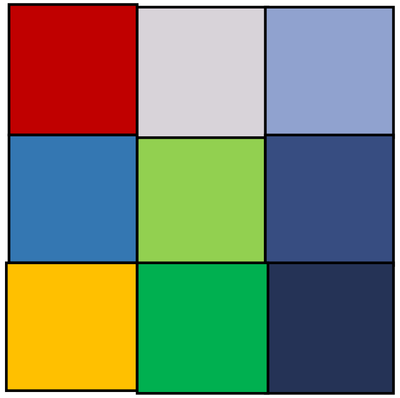
$$I_s(x, y) = \mathrm{comb}(x/X)\mathrm{comb}(y/Y)I(x, y)$$



From J. Goodman, *Introduction to Fourier Optics*

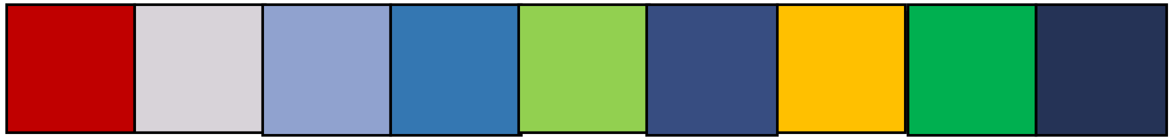
Images unrolled into vectors

$I_s(x,y)$



$\mathrm{vec}_r[\bullet]$ $\longrightarrow$

I_s

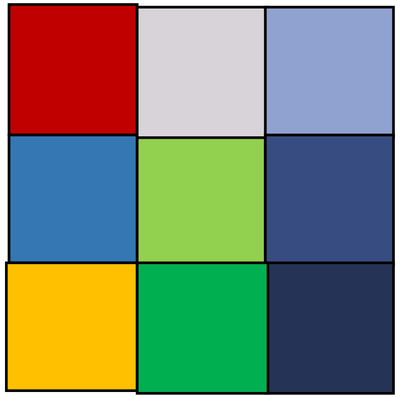


17	21	24
20	23	26
22	25	29

17	21	24	20	23	26	22	25	29

Images unrolled into vectors

$I_s(x,y)$

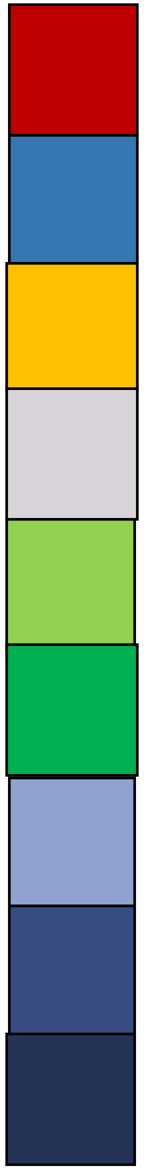


17	21	24
20	23	26
22	25	29

$\mathrm{vec}_c[\bullet]$

$\longrightarrow$

I_s



17
20
22
21
23
25
24
26
29

Example manipulations of images

1. Image addition/subtraction

$$I_o = I_1 + I_2$$

2. Image multiplication

$$I_o = I_1 \odot I_2$$

3. Image transformation: matrix-vector multiplication

$$I_o = \mathbf{W} I_1$$

4. Non-linear image operations

$$I_o = |I_1|^2$$

5. Convolution

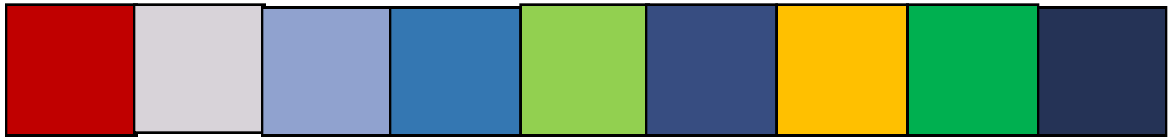
$$I_o = I_1 * I_2$$

Important image manipulation: convolution

$$I_o = I_1 * I_2$$

$$I_o[n] = \sum_{m=-M}^{M} I_1[n-m] I_2[m]$$

I_1



17	21	24	20	23	26	22	25	29

$*$

I_2

1	1	1

=

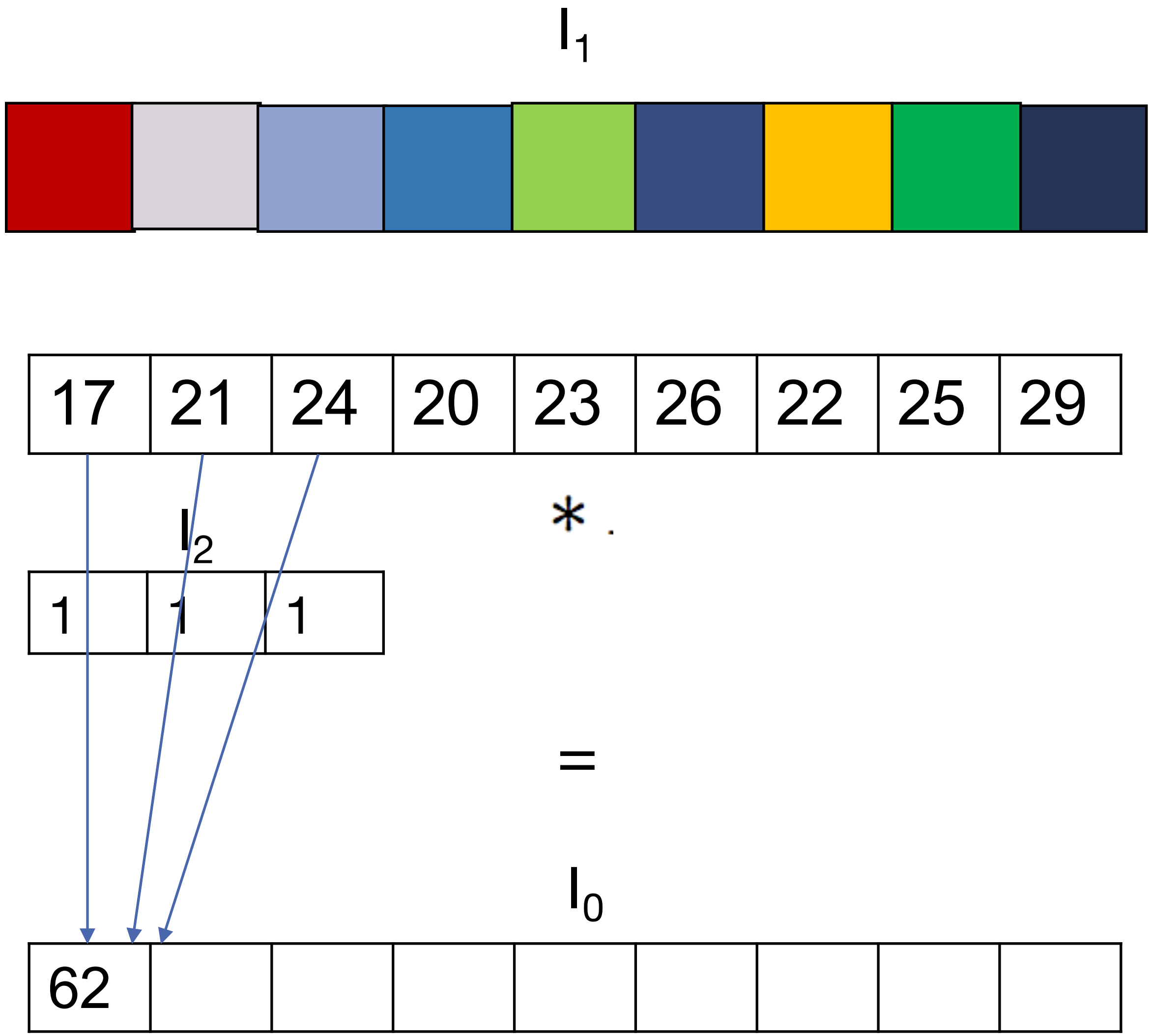
I_0

deep imaging

Important image manipulation: convolution

$$I_o = I_1 * I_2$$

$$I_o[n] = \sum_{m=-M}^{M} I_1[n-m] I_2[m]$$

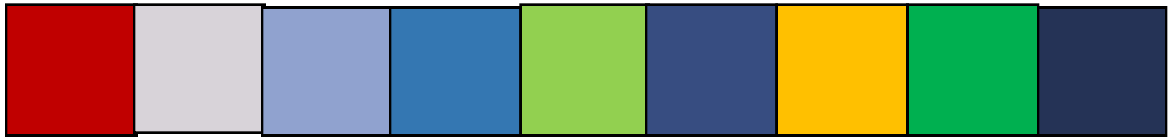


Important image manipulation: convolution

$$I_o = I_1 * I_2$$

$$I_o[n] = \sum_{m=-M}^{M} I_1[n-m] I_2[m]$$

I_1



17	21	24	20	23	26	22	25	29

$*$

I_2

1	1	1

=

I_0

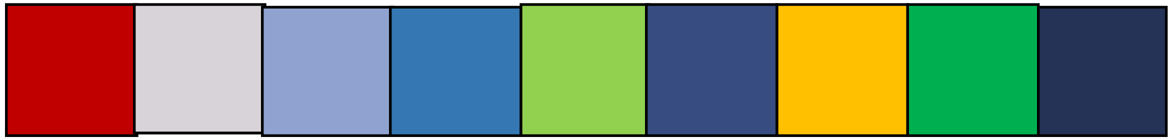
62								

Important image manipulation: convolution

$$I_o = I_1 * I_2$$

$$I_o[n] = \sum_{m=-M}^{M} I_1[n-m]\, I_2[m]$$

I_1



17	21	24	20	23	26	22	25	29

$*$

I_2

1	1	1

=

I_0

62	65							

Important image manipulation: convolution

$$I_o[s,t] = \sum_{l=-L}^{L} \sum_{m=-M}^{M} I_1[s-l, t-m] I_2[l,m]$$

I_1

$*$

I_2

1	1	1
1	1	1
1	1	1

=

I_0

Important image manipulation: convolution

$$I_o[s,t] = \sum_{l=-L}^{L} \sum_{m=-M}^{M} I_1[s-l, t-m] I_2[l,m]$$

I_1

$*$

I_2

-1	0	-1
-2	0	-2
-1	0	-1

=

I_0

Machine learning: “dynamic” image manipulations

deep imaging

Current goal in machine learning : determine image manipulations to highlight features of interest

I_1

$$* \quad W = \begin{bmatrix} w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 \\ w_7 & w_8 & w_9 \end{bmatrix} =$$

Most useful information possible for computer to use

Machine learning: “dynamic” image manipulations

Current goal in machine learning : determine image manipulations to highlight features of interest

I_1

$*$.

W

w 1	w 2	w 3
w 4	w 5	w 6
w 7	w 8	w 9

=

Most useful information possible for computer to use

Determine weights w for particular task: image segmentation, object detection, bw-to-color, etc.

Example tasks for machine learning

Common ML transformations for detected image:

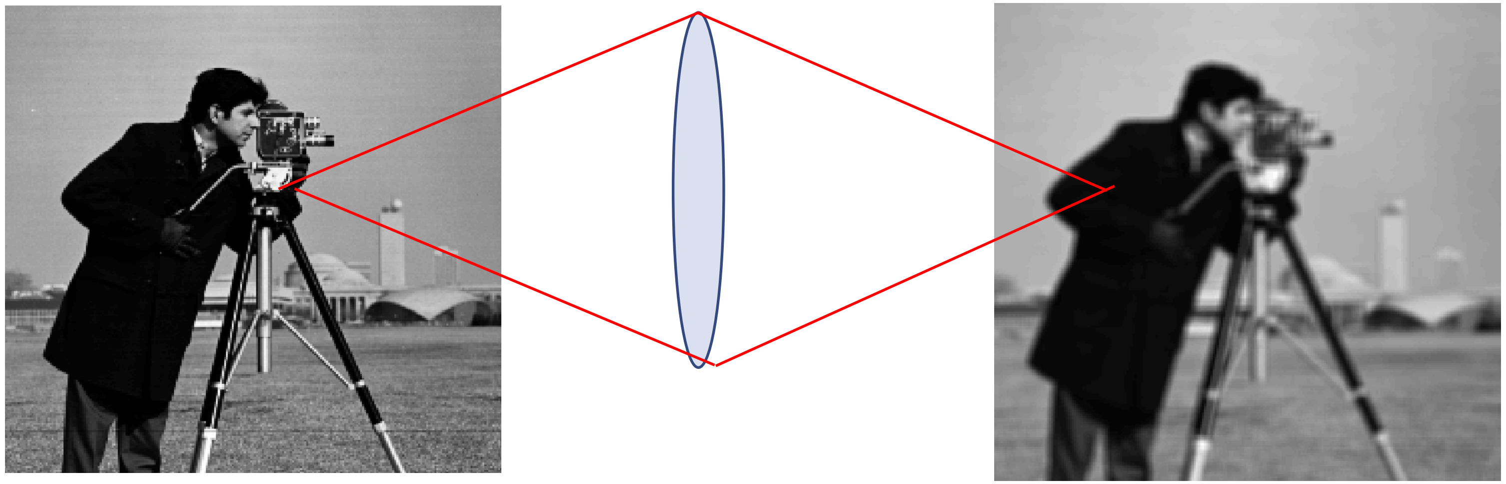
- A vector of different categories (image is of a man, not a dog)
- A vector of coordinates highlighting features of interest (the man's head is contained in the box of pixels from (x,y,x+a,y+b)
- A segmentation map (the line denoting the boundary of the man is 1, rest is 0)

Image formation as a set of discrete equations

- Can also model the behavior of the imaging system before the radiation hits the image detector

Physical world

n x m image I_1



$$I_1 = \mathbf{W_0}\, I_0$$

Image formation as a set of discrete equations

- Can also model the behavior of the imaging system before the radiation hits the image detector

Physical world

n x *m* image I_1

Processed image I_2

$I_1 = \mathbf{W_0}\, I_0$

$I_2 = \mathbf{W_1}\, I_1$

Image formation as a set of discrete equations

- Can also model the behavior of the imaging system before the radiation hits the image detector

Physical world

n x *m* image I_1

Processed image I_2

$I_2 = \mathbf{W_1 W_0}\, I_0$

Linear mapping

deep imaging

Bringing together physical and digital image representations

Physical world

n x m image I_1

$$\text{Task} = \mathbf{W_n} \ldots \mathbf{T_1}[\mathbf{W_1}\,\mathbf{T_0}[\mathbf{W_0}\ I_0]\ldots]$$

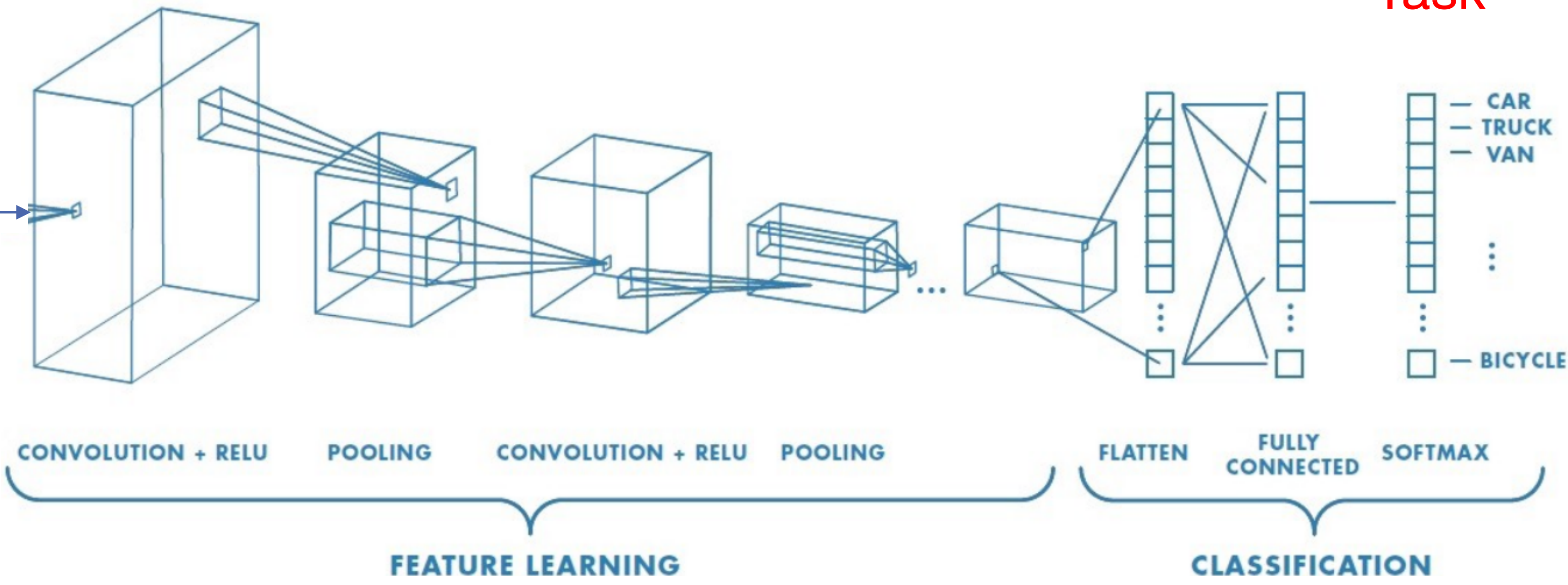
Nonlinear mapping

Bringing together physical and digital image representations

Physical world I_0

Hardware

Image I_1



Task

$$\text{Task} = \mathbf{W_n} \ldots \mathbf{T_1}[\mathbf{W_1}\,\mathbf{T_0}[\mathbf{W_0}\ I_0]\ldots]$$

From https://towardsdatascience.com/

Bringing together physical and digital image representations

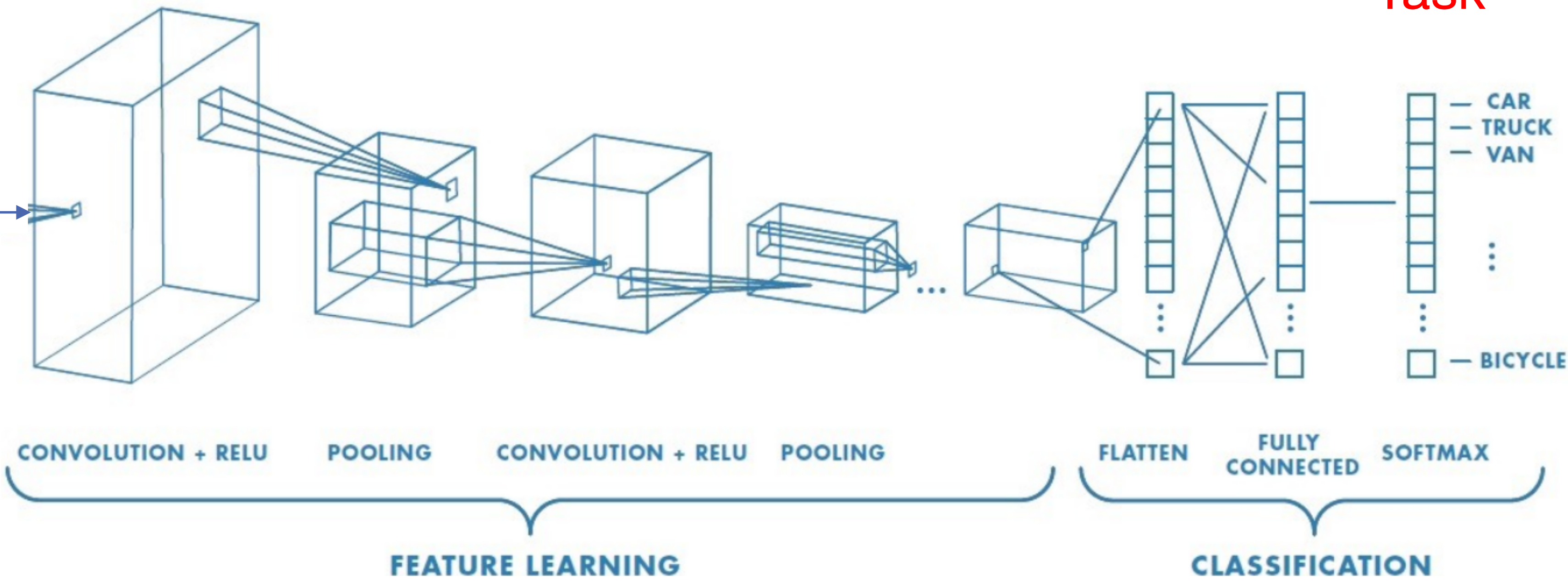
deep imaging

Physical world I_0

Hardware

Image I_1

Task



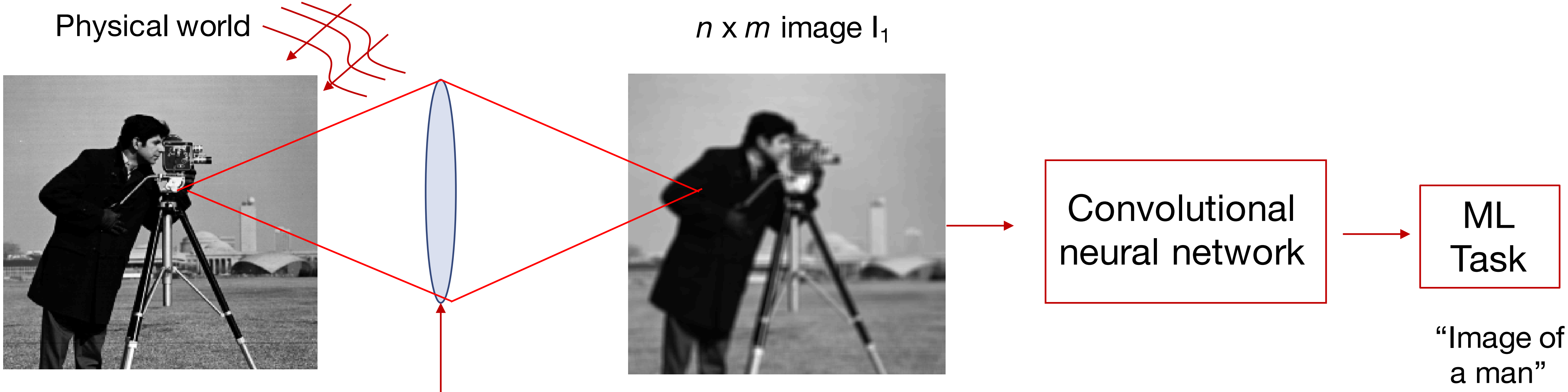
Account for or modify imaging hardware

$$\text{Task} = \mathbf{W_n} \ldots \mathbf{T_1}[\mathbf{W_1}\,\mathbf{T_0}[\mathbf{W_0}\; I_0]\ldots]$$

From https://towardsdatascience.com/

deep imaging

Bringing together physical and digital image representations



Illumination
Imaging device (lens, coils, transducers)
Detector design
Properties of sample of interest

$$\text{Task} = \mathbf{W_n} \ldots \mathbf{T_1}[\mathbf{W_1}\,\mathbf{T_0}[\mathbf{W_0}\; I_0]\ldots]$$

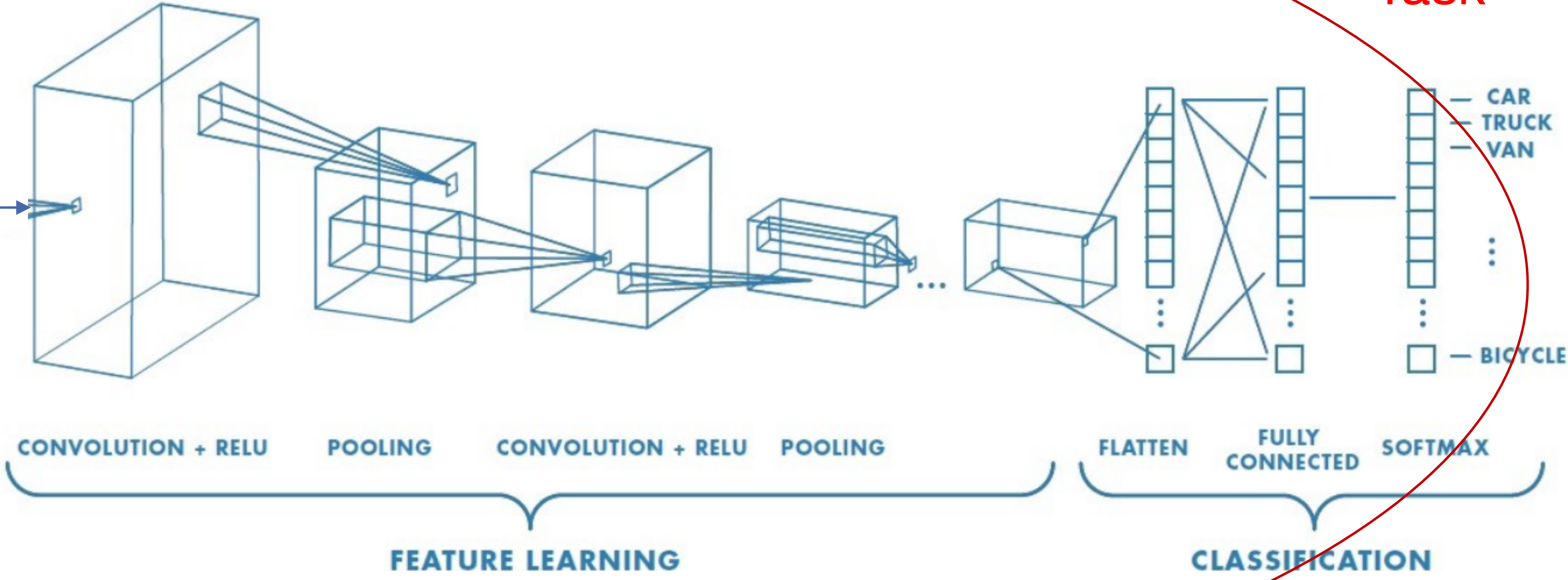
Bringing together physical and digital image representations

Physical world I_0

Hardware

Image I_1

Task



Final project: try to optimize all of this together!

From https://towardsdatascience.com/