CS540 Introduction to Artificial Intelligence
Lecture 7

Young Wu
Based on lecture slides by Jerry Zhu and Yingyu Liang

June 10, 2019

Calculator on Midterm

Admin

- Do you want calculators allowed on midterm?
- Not allowed $\Rightarrow$ questions will contain expressions without numbers or nice numbers.
- Allowed $\Rightarrow$ questions will contain specific numbers, possibly require rounding etc.
- A: indifferent
- B: Yes on regular midterm and alternative midterm
- C: No on regular midterm and alternative midterm
- D: Yes on regular midterm, No on alternative midterm
- E: No on regular midterm, Yes on alternative midterm

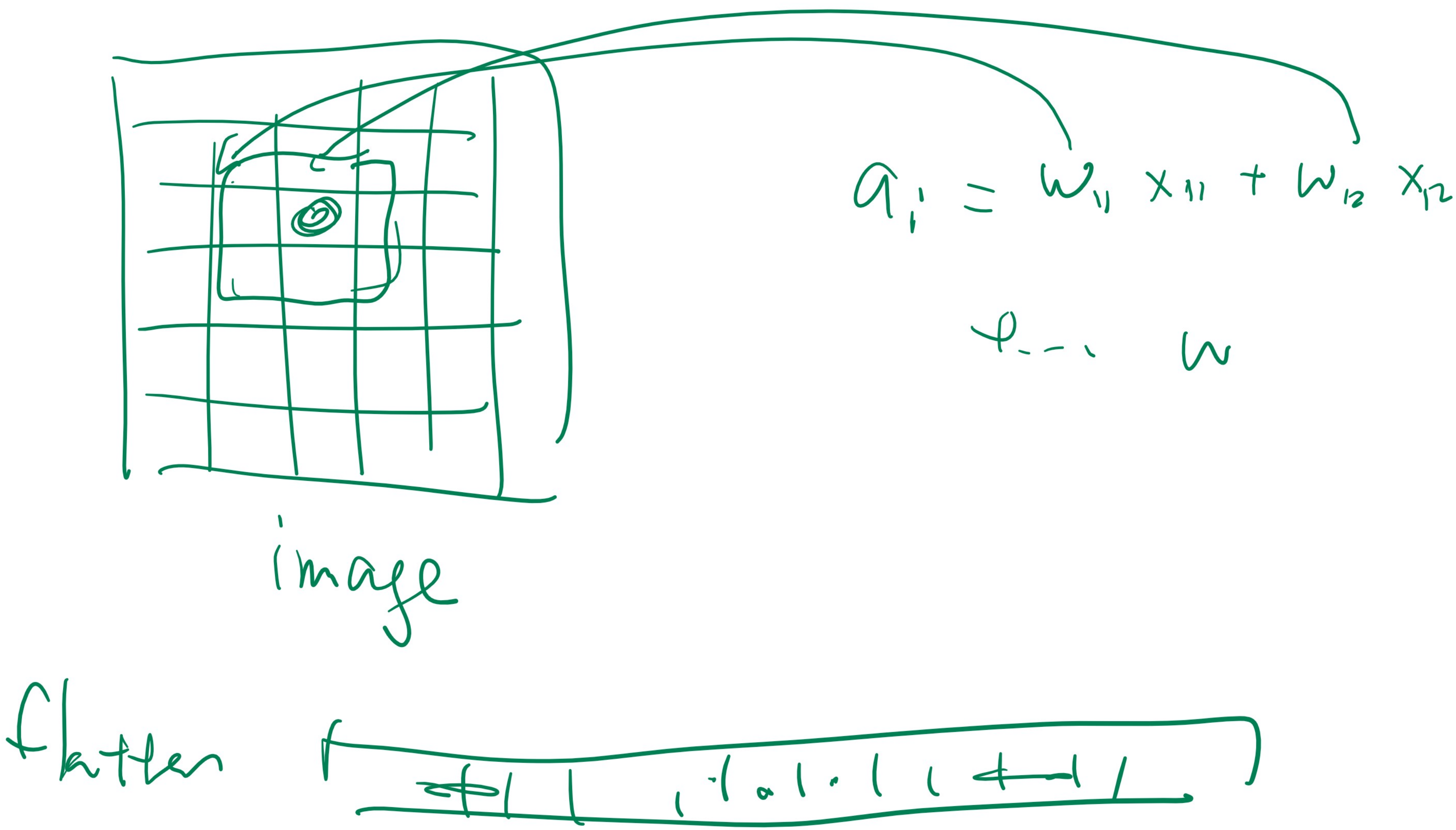
Homework 2

Admin

- What is your classification accuracy on Homework 2?
- A: 50 – 75 percent.
- B: 75 – 90 percent.
- C: 90 – 100 percent.
- D: Not finished yet.
- E: Not started yet.

Image Preprocessing Diagram

Motivation



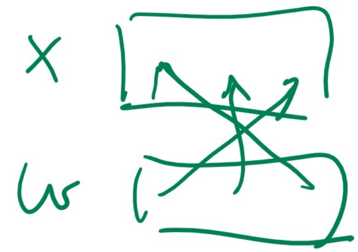
Description of Algorithm

Description

- Convolve the input image with a filter.
- Pool the output of convolution.
- Feed the output of pooling into a neural network.

One Dimensional Convolution

Definition



- The convolution of a vector $x = (x_1, x_2, ..., x_m)$ with a filter $w = (w_{-k}, w_{-k+1}, ...w_{k-1}, w_k)$ is:

$$a = (a_1, a_2, ..., a_m) = x * w$$

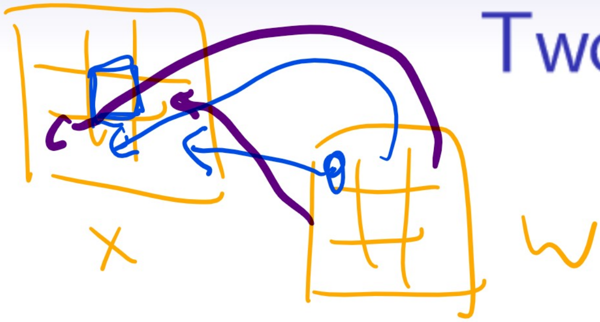
$$a_j = \sum_{t=-k}^{k} w_t x_{j-t}, j = 1, 2, ..., m$$

linear combination.

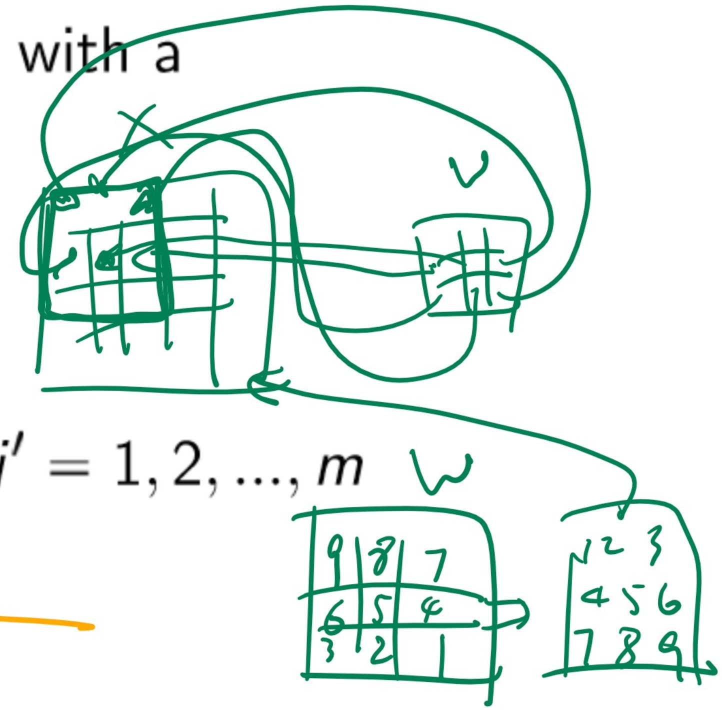
- w is also called a kernel (different from the kernel for SVMs).
- The elements that do not exist are assumed to be 0.

Two Dimensional Convolution

Definition



- The convolution of an $m \times m$ matrix X with a $(2k+1) \times (2k+1)$ filter W is:



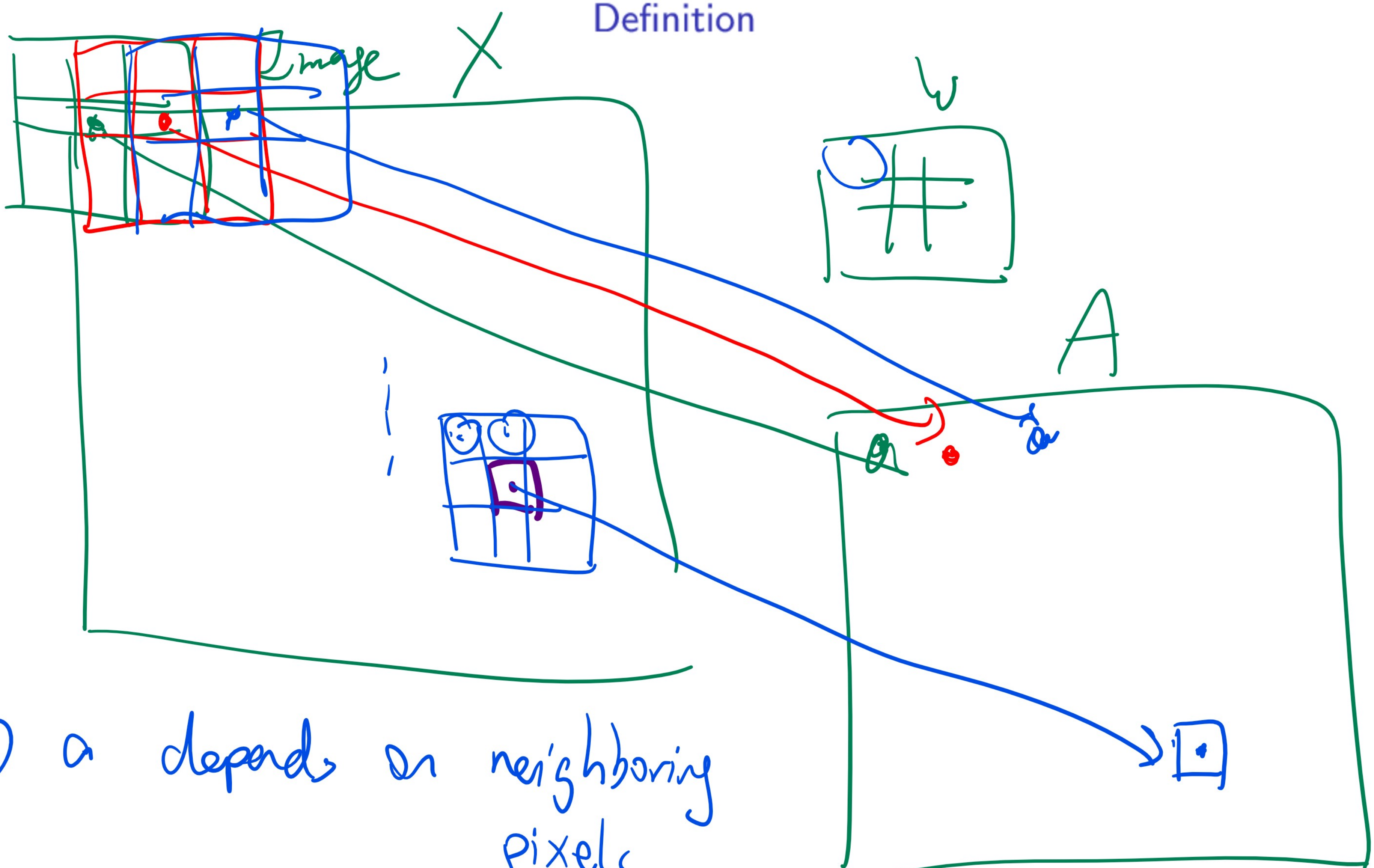
$$A = X * W$$

$$A_{j,j'} = \sum_{t=-k}^{k} \sum_{t'=-k}^{k} W_{t,t'} X_{j-t,j'-t'}, \quad j, j' = 1, 2, ..., m$$

- The matrix W is indexed by (t, t') for $t = -k, -k+1, ..., k-1, k$ and $t' = -k, -k+1, ..., k-1, k$.
- The elements that do not exist are assumed to be 0.

Convolution Diagram

Definition



① a depends on neighboring pixels

② all locations share the same W.

Convolution Example, Part I

Quiz (Graded)

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

- A: $\begin{bmatrix} -1 & -3 & -3 \\ 0 & 0 & 0 \\ 1 & 3 & 3 \end{bmatrix}$, B : $\begin{bmatrix} -3 & -3 & 3 \\ -4 & -4 & 4 \\ -3 & -3 & 3 \end{bmatrix}$
- C: $\begin{bmatrix} -3 & -4 & -3 \\ -3 & -4 & -3 \\ 3 & 4 & 3 \end{bmatrix}$, D : $\begin{bmatrix} -1 & 0 & 1 \\ -3 & 0 & 3 \\ -3 & 0 & 3 \end{bmatrix}$
- E: none of the above

Convolution Example, Part II

Quiz (Graded)

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

- A: $\begin{bmatrix} -1 & -3 & -3 \\ 0 & 0 & 0 \\ 1 & 3 & 3 \end{bmatrix}$, B : $\begin{bmatrix} -3 & -3 & 3 \\ -4 & -4 & 4 \\ -3 & -3 & 3 \end{bmatrix}$
- C: $\begin{bmatrix} -3 & -4 & -3 \\ -3 & -4 & -3 \\ 3 & 4 & 3 \end{bmatrix}$, D : $\begin{bmatrix} -1 & 0 & 1 \\ -3 & 0 & 3 \\ -3 & 0 & 3 \end{bmatrix}$
- E: none of the above

Pooling

Definition

- Combine the output of the convolution by max pooling,

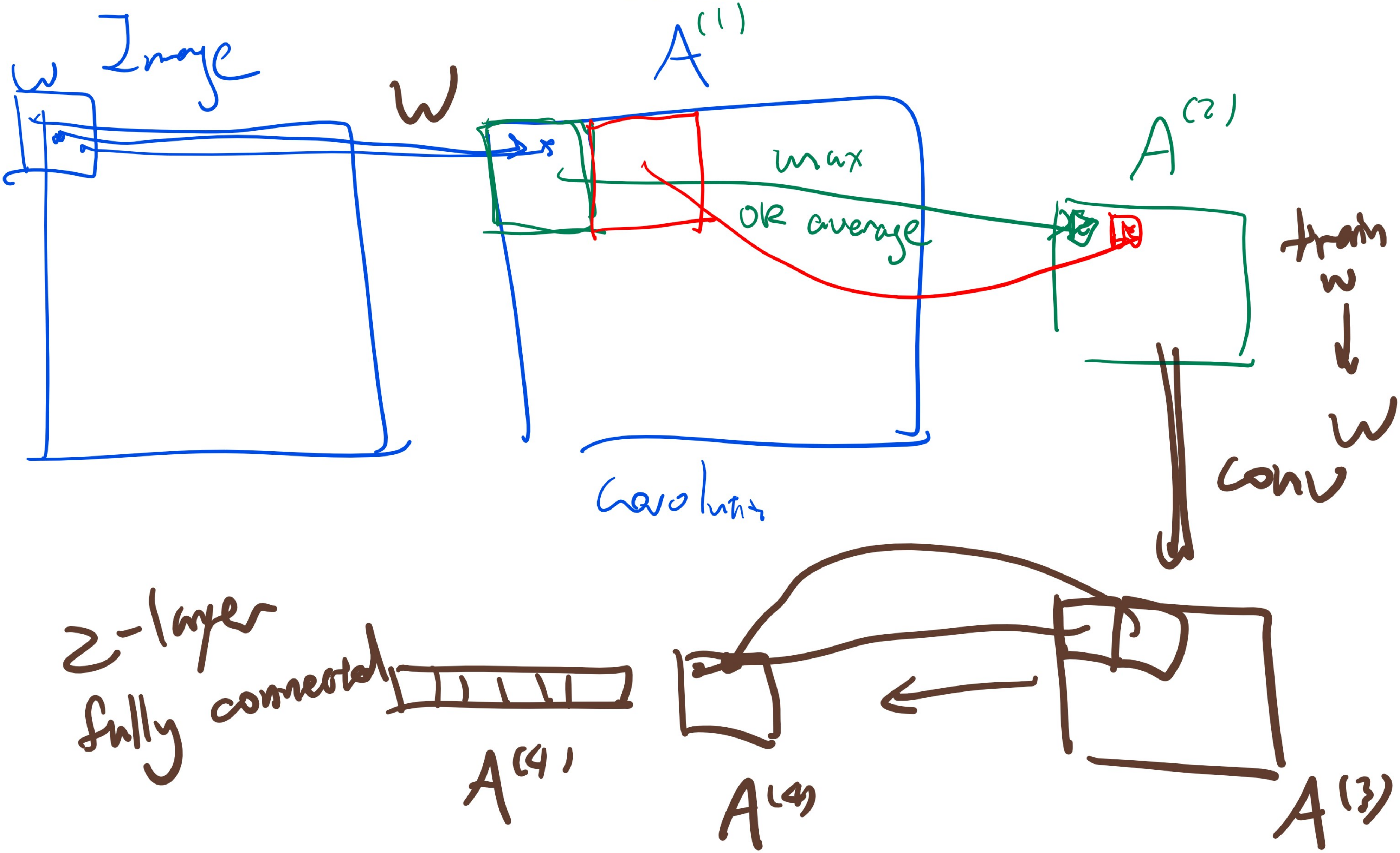
$$a = \max\{x_1 \ldots x_m\}$$

- Combine the output of the convolution by average pooling,

$$a = \frac{1}{m}\sum_{j=1}^{m} x_j$$

Pooling Diagram

Definition



Training Convolutional Neural Networks

Discussion

- The training is done by gradient descent.
- The simplest way to update convolutional layers is to use finite difference approximation.

$$\frac{\partial C}{\partial w} \approx \frac{C\left(w+\varepsilon\right)-C\left(w-\varepsilon\right)}{2\varepsilon}, \varepsilon > 0, \varepsilon \approx 0$$

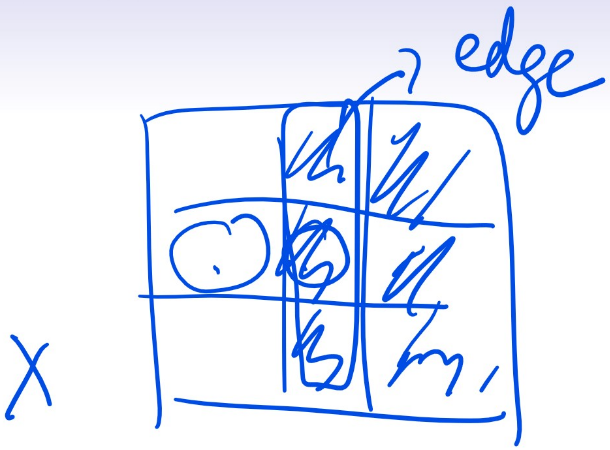
Generating Features for Images

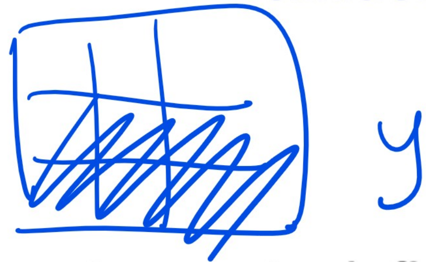
Motivation

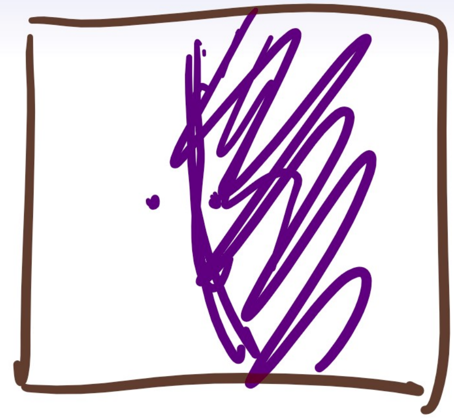
- Instead of training convolutional layers to generate features in images, there are predefined filters and other techniques to find features such as edges, corners, or blobs in images.

Image Gradient

Definition







- The gradient of an image is defined as the change in pixel intensity due to the change in the location of the pixel.

$$\frac{\partial I(x,y)}{\partial x} \approx \frac{I\left(x+\frac{\varepsilon}{2}, y\right) - I\left(x-\frac{\varepsilon}{2}, y\right)}{\varepsilon}, \varepsilon = 1$$

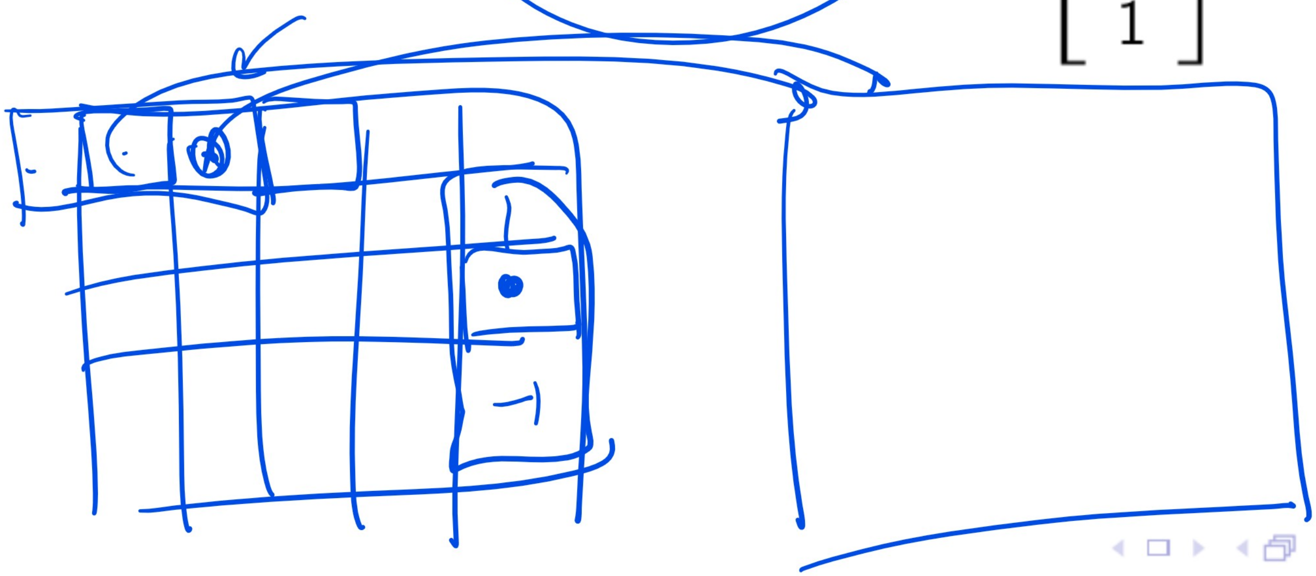
$$\frac{\partial I(x,y)}{\partial y} \approx \frac{I\left(x, y+\frac{\varepsilon}{2}\right) - I\left(x, y-\frac{\varepsilon}{2}\right)}{\varepsilon}, \varepsilon = 1$$

Image Derivative Filters

Definition

- The gradient can be computed using convolution with the following filters.

$$w_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}, w_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$



Sobel Filter

Definition

- The Sobel filters also are used to approximate the gradient of an image.

$$W_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, W_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Decomposition of Filters

Definition

- The Sobel filters can be decomposed into two one dimensional filters.

$$W_x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}, W_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$$

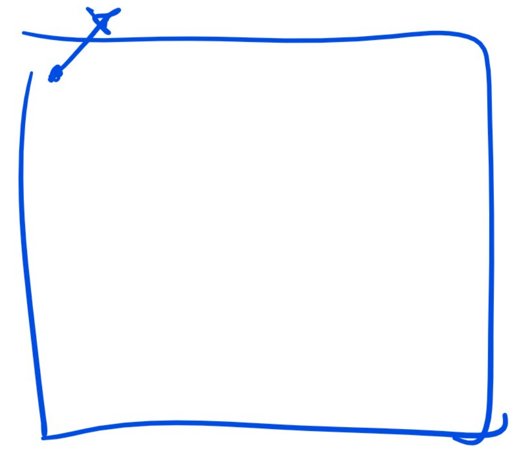
- It is significantly faster to do two one dimensional convolutions than to do one two dimensional convolution.

Gradient of Images

Definition

- The gradient of an image I is $(\nabla_x I, \nabla_y I)$.

$$\nabla_x I = W_x * I, \nabla_y I = W_y * I$$



- The gradient magnitude is G and gradient direction Θ are the following.

$$G = \sqrt{\nabla_x^2 + \nabla_y^2}$$

$$\Theta = \arctan\left(\frac{\nabla_y}{\nabla_x}\right)$$

Laplacian of Image

Definition

- The Laplacian of an image I is defined as the sum of the second derivatives.

$$\nabla^2 I(x,y) = \frac{\partial^2 I(x,y)}{\partial^2 x^2} + \frac{\partial^2 I(x,y)}{\partial^2 y^2}$$

$$\frac{\partial^2 I(x,y)}{\partial^2 x^2} \approx \frac{I(x+\varepsilon,y) - 2I(x,y) + I(x-\varepsilon,y)}{\varepsilon^2}, \varepsilon = 1$$

$$\frac{\partial^2 I(x,y)}{\partial^2 y^2} \approx \frac{I(x,y+\varepsilon) - 2I(x,y) + I(x,y-\varepsilon)}{\varepsilon^2}, \varepsilon = 1$$

Laplacian Filter

Definition

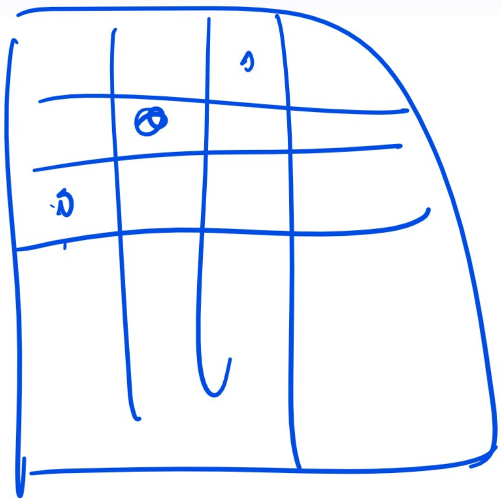
- The Laplacian can be computed using convolution with the following filters.

$$W_L = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\nabla^2 I = W_L * I$$

Edge Detection

Discussion



- Both the gradient and Laplacian of an image can be used to find edge pixels in an image.
- Images usually contain noise. The noises are not edges and are usually removed before computing the gradient.

2 Dimensional Gaussian Filter

Definition

- The Gaussian filter is used to blur images and remove noise in the image. A Gaussian filter with standard deviation σ is the following.

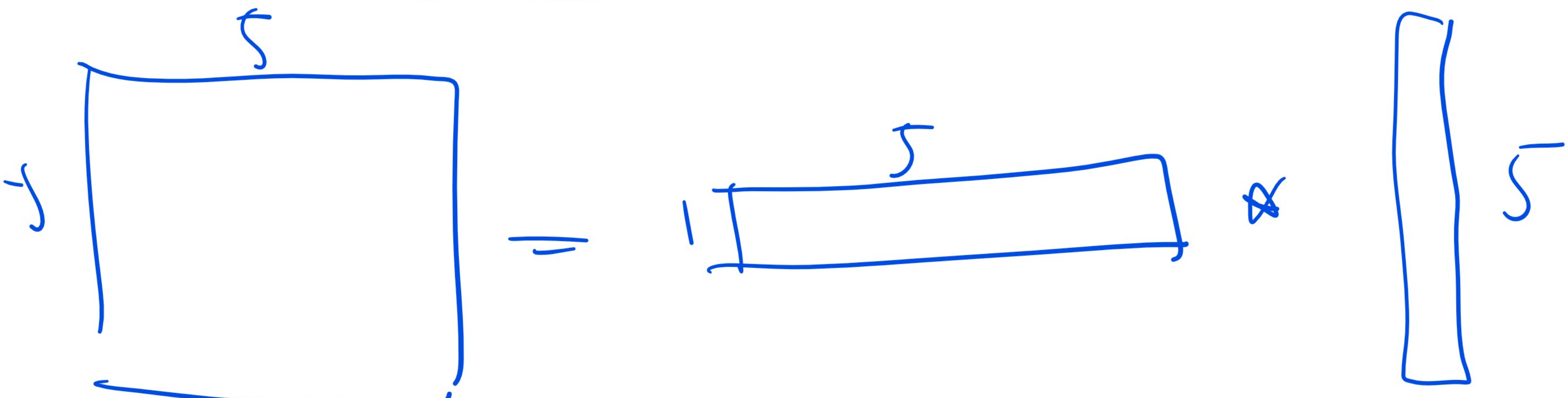
$$W_\sigma : (W_\sigma)_{t,t'} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{t^2 + t'^2}{2\sigma^2}\right)$$

1 Dimensional Gaussian Filter

Definition

- The Gaussian filter can be decomposed into two one dimensional filters as well.

$$W_\sigma = w_\sigma * w_\sigma, (w_\sigma)_t = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{t^2}{2\sigma^2}\right)$$



Gaussian Filter Example 3

Definition

- When filter size $k = 3$, and standard deviation $\sigma = 0.8$:

$$W_\sigma = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

- Sobel filter is approximately the combination of the gradient filter and the Gaussian filter.

Laplacian of Gaussian

Definition

- The Laplacian filter and the Gaussian filter are usually also combined into one filter called Laplacian of Gaussian filter (LoG filter).

$$W_{L,\sigma} : (W_{L,\sigma})_{t,t'} = -\frac{1}{\pi\sigma^4}\left(1 - \frac{t^2 + t'^2}{2\sigma^2}\right)\exp\left(-\frac{t^2 + t'^2}{2\sigma^2}\right)$$

Difference of Gaussian

Definition

- The Laplacian of Gaussian filter is difficult to compute because it cannot be decomposed into two one dimensional filters. Therefore an approximation is used called the Difference of Gaussian filter (DoG filter).

$$W_{L,\sigma} \approx W_{\sigma} - W_{1.6\sigma}$$

LoG and DoG Diagram

Definition

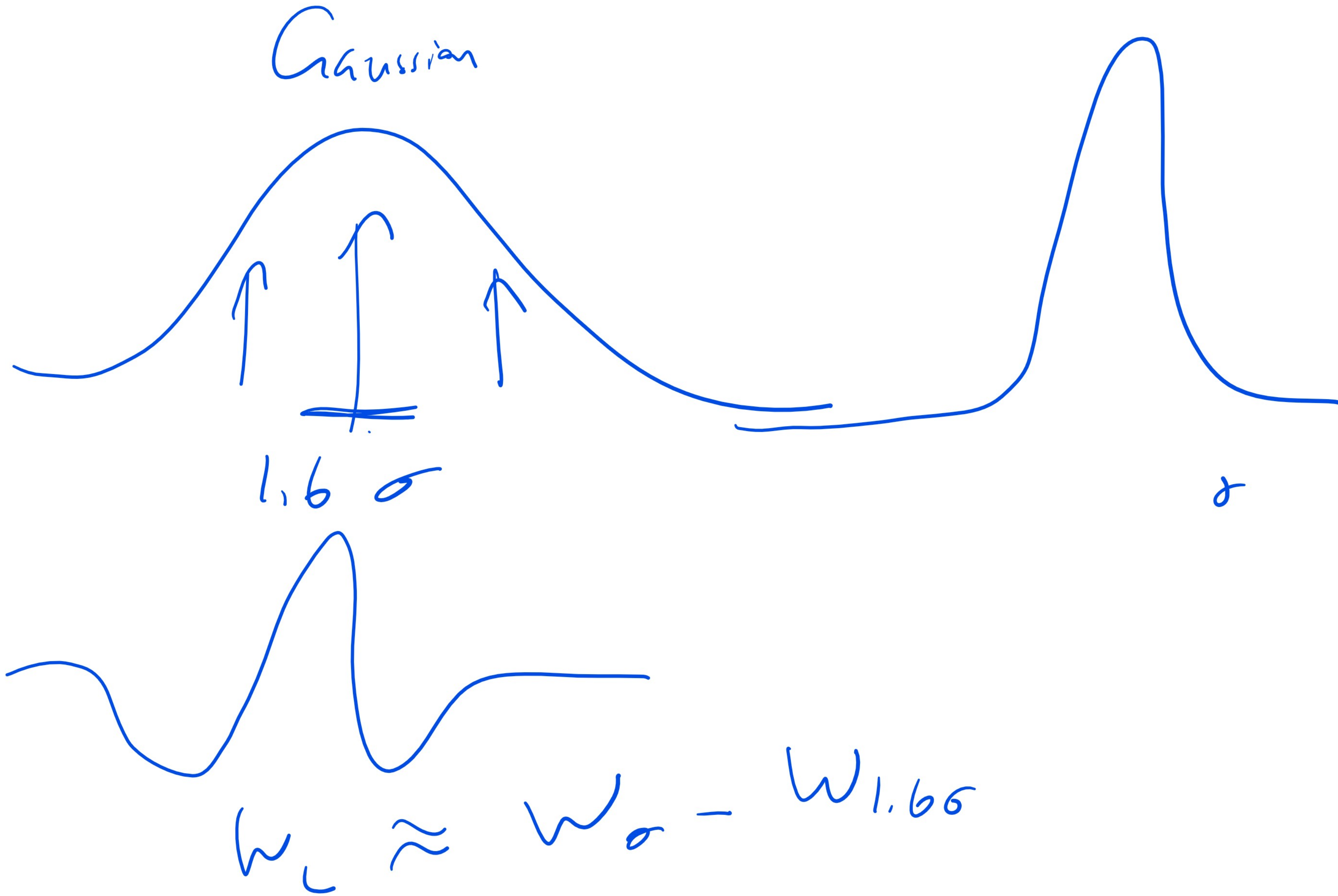
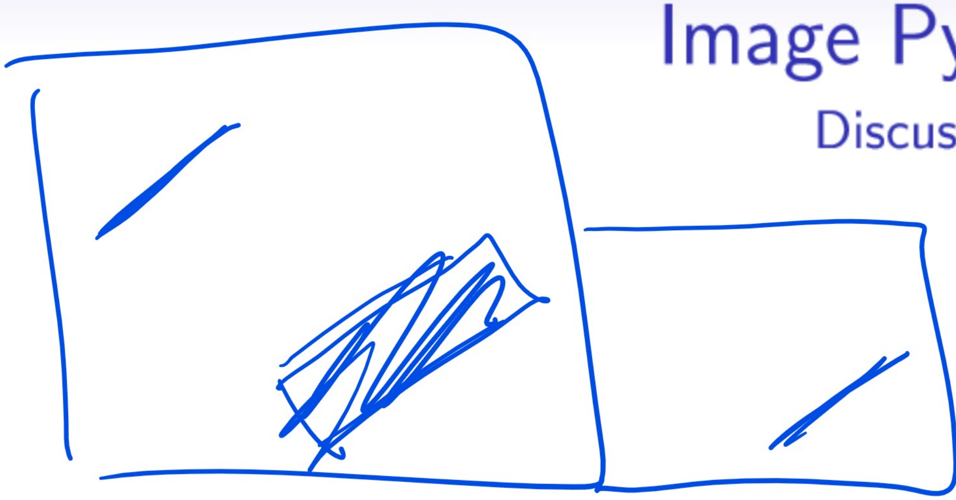
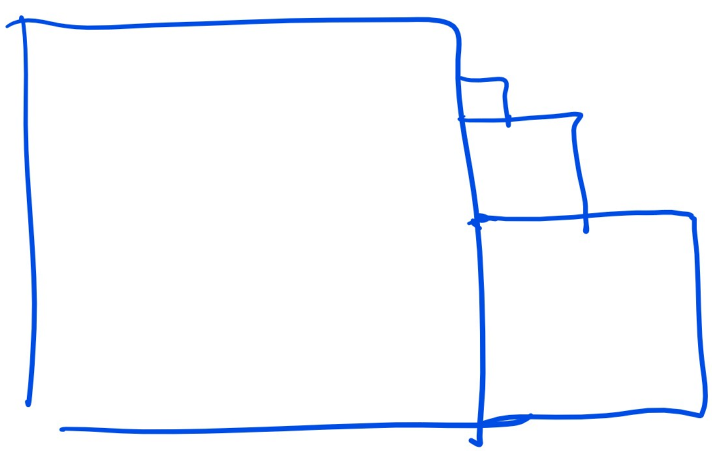


Image Pyramids

Discussion



- There are edges at different scales of the image. Images are blurred and downsampled to get images with different scales.
- An image pyramid contains images at scales $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$



Next Lecture

Admin

- Dandi Chen will lecture on Wednesday on Computer Vision.