CMPE258
Fall 2023

1/

August 22 (Tue)

Organizational meeting.

1. Class Material on github
github/hualili

aiv100	Add files via upload
deep-learning-2020S	Add files via upload
deep-learning-2022s	Add files via upload
deep-learning-2023s	Add files via upload
facial-detect	Add files via upload
lec1Capture/CMakeFiles	Delete CMakeDirectoryInformation.
lecOpenCV_GL	openGL and openCV sample comm
riscv	Create readme.txt
20-2021S-0-7-1convnets-NumeralD...	Add files via upload

Course and Contact Information

Instructor(s): Harry Li
Office Location: Engineering Building, Room 267A
Telephone: (650) 400-1116 for text messaging only
Email: hua.li@sjsu.edu
Office Hours: M.W. 3:00-4:00 pm
IN-Person.
Class Days/Time: Tuesdays and Thursdays 4:30-5:45 pm.
Classroom: Engineering Building Room 337
Prerequisites: CMPE 255 or CMPE 257 or instructor consent. Computer Engine
Engineering majors only.

Course Description

2. Prerequisites Requirements
Bring your Proof to the next Class.

3. Emphasis on "Deep Neural Networks" &
Semantic Segmentation

Course Description

Deep neural networks and their applications to various problems, e.g., speech recognition, image segmentation, detection and recognition of temporal and spatial patterns, and natural language processing. Covers underlying theory, the range of applications to which it has been applied, and learning from very large data sets.

Note: Definition (HI): (Human Intelligence)
is Symbolic Representation of
Learnt Experience.

CMPE258

Fall 2023

4. Projects. 2

plus 1 team project (Semester Long) } ~~30%~~ 25% ~~30~~ pts

5. In-Person Class; CANVAS is utilized to post Homework/Project Requirements, and to collect the submission of the Homework, as well as for the exams.

This course ~~is an online course~~ NO. The students must have Internet connectivity and Zoo
his/her machine. The students must participate in the class activities and submit all as
exams to SJSU CANVAS. The syllabus, faculty contact information on the syllabus, I
projects, and exam papers are all available on CANVAS. See University Policy F13-2

6.

Grading Information

Quiz, Homework, Projects	30%
Midterm Examination	30%
Final Examination	40%

Note:

1° CANVAS To Be up by the end of the day, Friday;

2° Tools & Software To be installed (will provide "Readme" as Ref)

OpenCV

Python.

T.F. Version 2.0 or higher

7. Textbooks & References

Textbook

- Deep Learning with Python, 1st or 2nd Edition, by François Chollet, ISI
  ISBN-10: 9781617294433, https://github.com/hualili/opencv/blob/mast
  2018F-6-DeepLearningCh02.pdf
- Robot Vision by B.K. P. Horn, the MIT press, ISBN 0-262-08159-8, c
  Hill).
- Reference textbook Learning OpenCV, Computer Vision with the Ope
  Kaebler, O'Reilly Publisher, ISBN 978-0-596-51613-0, 2011.

Today's Topic: Intro. to Deep Convolutional Neural Networks.

8. Software Tools & Dev. Environment

Python. Pycharm.

AnnoConda.

TensorFlow.

Note: Rm268 Available ~~per~~ on Approved Basis.

August 24 (Thursday)

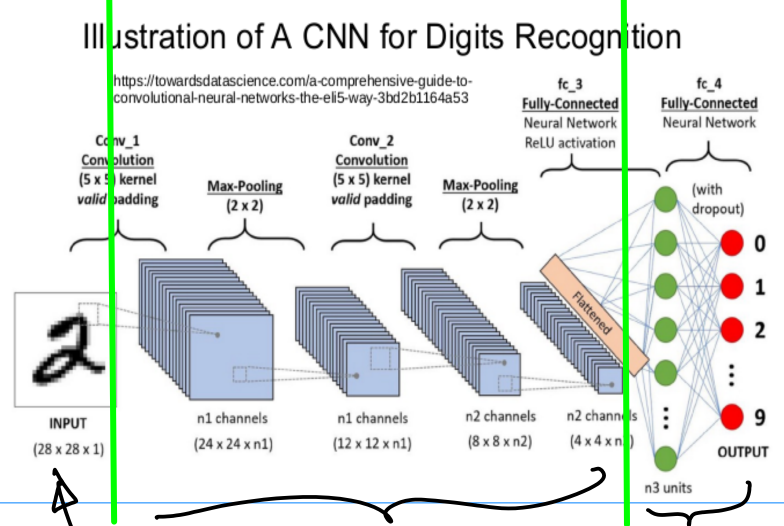
Ref: github.

2022F-103-NN-Intro-Python-v5-2022-8-25

Note: Lab Space for the Class Rm268.

Example: Architecture Overview.

Note 1.



Preprocessing Computer Vision.

Convolution Layers

Feature Extraction

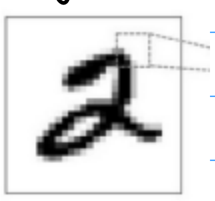
To produce Feature Vectors.

$$\vec{u} = (u_1, u_2, \cdots, u_N)$$

Feed Forward Neural Networks.

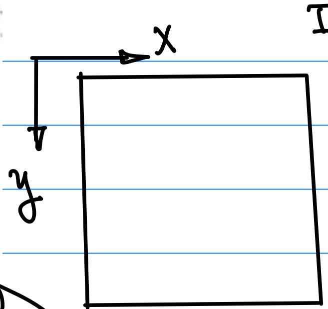
Decision Making

2. Image Definition



M, N, Channel

$I(x, y; t)$

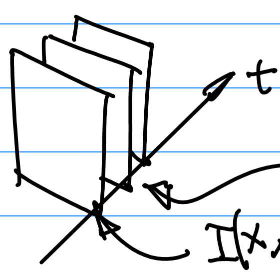


Note: for x & y definition see Ref. Next page.

$M \times N$

No. of col. Per Row.

No. of Rows per frame



$I(x, y; t+\Delta t)$

$I(x, y; t)$

Ref: on the github.

2023S-101-Note-cmpe258-2023-03-16.pdf

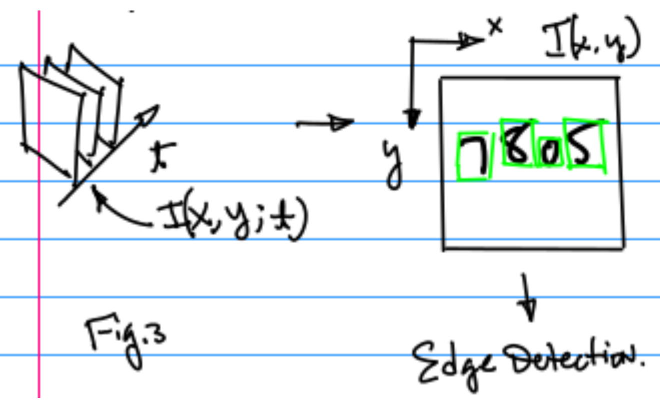
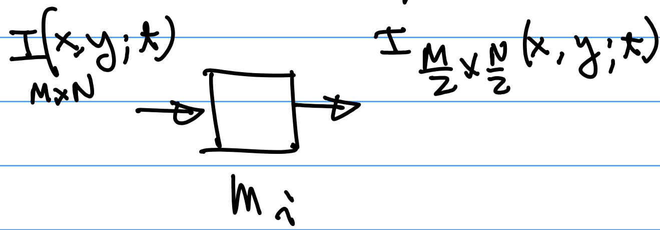


Fig.3

3. For the Convolutional Lay, we denote it as C_i, where $i = 1, 2, \cdots, N$
for Max pooling Layer, we denote it as M_j, where $j = 1, 2, \cdots, K$
so, for the example, we have

$C_1 M_1 \to C_2 M_2$

Maxpooling for Resolution Reduction / Feature Reduction.



Note: BioLogic Inspiration, Human Visual Perception System, Retina, ~ 110 million photo Receptors

Comparison to Web Cam.

1080p Resolution: 1920x1080

$\sim 2^{11} = 2 \cdot 2^{10} = 2K$

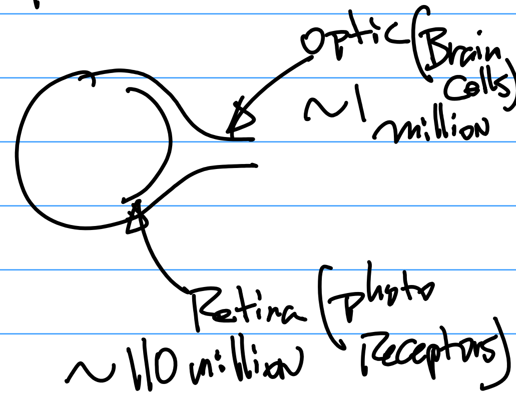
$\sim 2^{11}$

$\underbrace{\qquad}$

$2^{22} = 2^2 \cdot 2^{20}$

$= 4$ Meg

$\simeq 4$ million.



4. At the 3rd Segment (Blocks) of the Architecture, we denote <u>Feed Forward Neural Network</u> as F_M

No. of Neurons / Nodes

for example, F_{10} (10 Nodes) for the output Layer.

CMPE258

F2023.

August 29 (Tue)

Note: 1° CANVAS is up.

2° Homework Assignment

a. Honesty pledge to Be Signed / Signed Copy has to be uploaded to CANVAS.

b. Software Tools Installation. (Opt)

① OpenCV. By Friday

Next Tuesday. Bring Your Laptop w/ OpenCV installed.

Please use smart phone to take a photo, And upload the photo to your laptop to display.

Note: Sample Code (Python) was posted on the github.

② Anaconda Installed on your Laptop.

Note: Readme for Anaconda installation was posted on the github.

③ Create ChatGPT Account.

Python Interface to ChatGPT API (3.5 Version) is to be utilized in your Team Project.

3° Form A Team for the Semester Long Project.

Example: Consider 2D Convolution Technique.

Ref: github / hualili / opencv → 2022

Note 1:

$$c(n_1, n_2) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} a(k_1, k_2)\, b(n_1 - k_1, n_2 - k_2) \quad \ldots (1)$$

Image Kernel

$n_1 = x$, $n_2 = y$

Summation Index: k_1, k_2 are for x and y.

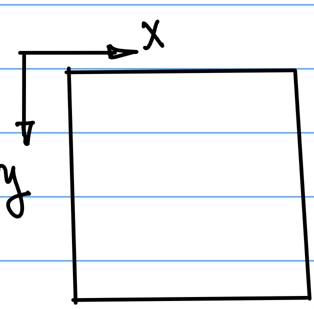


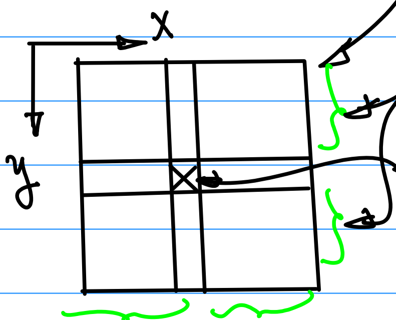
Fig.1

Note 2. $a(x,y)$ or $a(k_1,k_2)$ as an Image. $I(x,y)$

$b(x,y)$: a Kernel for 2D Convolution, $b(k_1,k_2)$

Note 3. Kernel $b(x,y)$ Can be rewritten $K(x,y)$

Size of A Kernel is denoted as $K \times K$



$K(x,y)$

The No. of Rows Above the P.O.I. = No. of Rows Beneath the P.O.I.

Pixel of Interests.

K x K

Fig. 2

$K = 3, 5, 7, \ldots$ (odd Number)

The No. of Col. Right to the P.O.I. = The No. of Col. Left to the P.O.I.

$I(x,y)$

0	0	100	100
0	0	100	100
0	0	100	100
0	0	100	100

4x4

Fig. 3a

8 bit Grayscale: $2^8 - 1 = 255$

-1	0	+1
-1	0	+1
-1	0	+1

Fig. 3b

Step 1. Take the Kernel, and place it at the initial position.

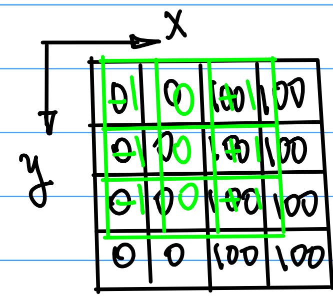


Fig. 3c

$$-1 \times 0 + 0 \times 0 + 1 \times 100 + \\ -1 \times 0 + 0 \times 0 + 1 \times 100 + \\ -1 \times 0 + 0 \times 0 + 1 \times 100 \\ = 300$$

$I(x,y) * K(x,y)$

	300		

Fig. 3d

Step 2. Shift $K(x,y)$ to the Right By 1 pixel (on the Same Row).

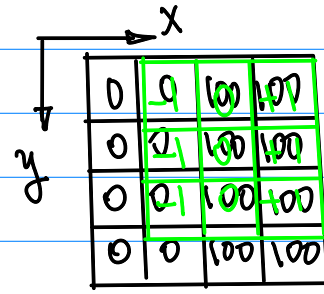


Fig. 3e

$$-1 \times 0 + 0 \times 100 + 1 \times 100 + \\ -1 \times 0 + 0 \times 100 + 1 \times 100 + \\ -1 \times 0 + 0 \times 100 + 1 \times 100 = 300$$

$I(x,y) * K(x,y)$

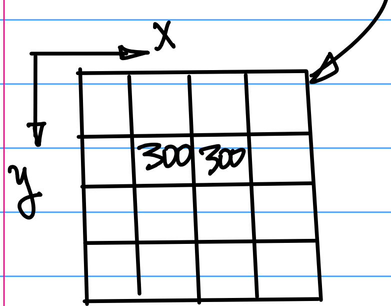


Fig.3F

Step3. $I(x,y) * K(x,y)$

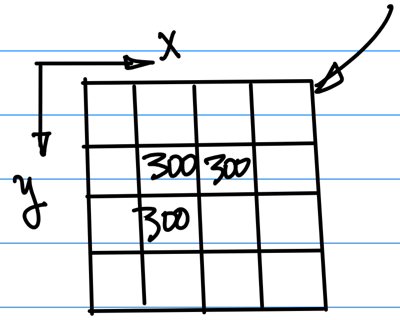


Fig.3G

Step3. $I(x,y) * K(x,y)$

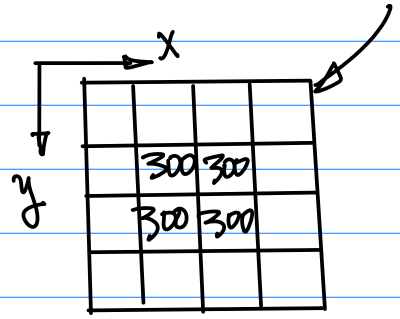


Fig.3H

Note: 2D Convolution (Continuous Case)

given Image $f(x,y)$
a Kernel $h(x,y)$

$$\iint_{\Omega} f(u,v)\, h(x-u, y-v)\, du\, dv \quad \cdots (2)$$

August 31 (Th).

Note: 1° Honesty Pledge on CANVAS, Signed pdf due this Friday. (ON CANVAS). Homework ON CANVAS

2° Will post Anaconda Installation & OpenCV Installation, display an image.

(1) Screen Capture of the Activated CONDA Environment, with personal identifier;

(2) Screen Capture of the OpenCV Display with P.I.D.

Sample Code for the display to Be Re-posted on github, 2023F

Example: Theoretical Aspects for 2D Convolution.

Ref: 1° PPT on the github.

2° 2023S-100-note-...2023-03-21.pdf

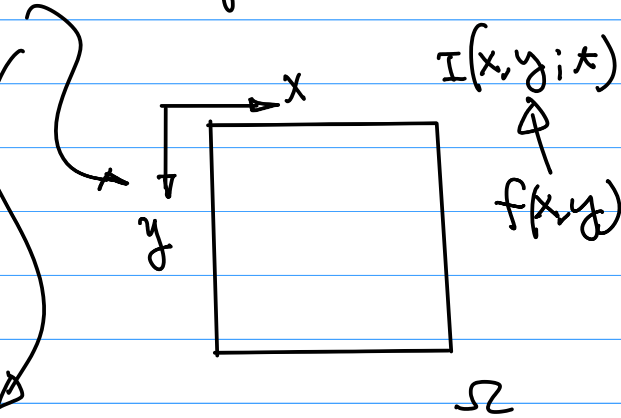
$$\iint_{\Omega} f(u,v)\, h(x-u, y-v)\, du\, dv \quad \cdots (1)$$

Image

Kernel

Note1. Ω: Image Plane(s)

OR



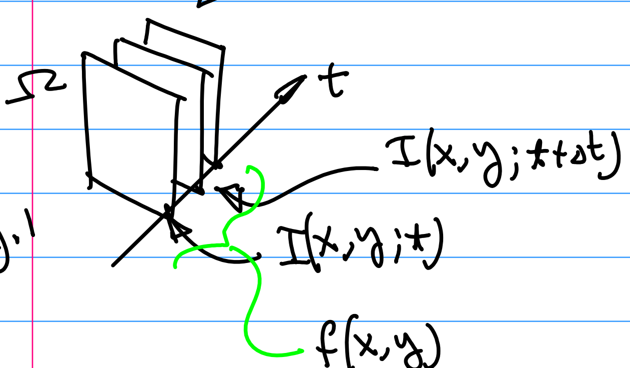


Fig.1

Note2: Kernel

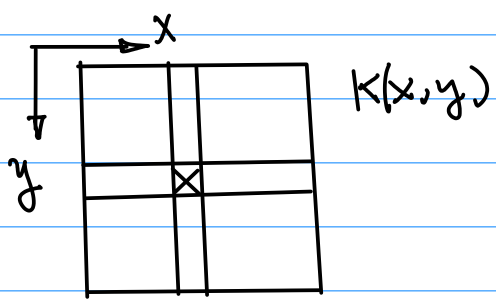


Fig.2a

↓ (1) $K(-x,y)$

Flip the Kernel w.r.t y-axis

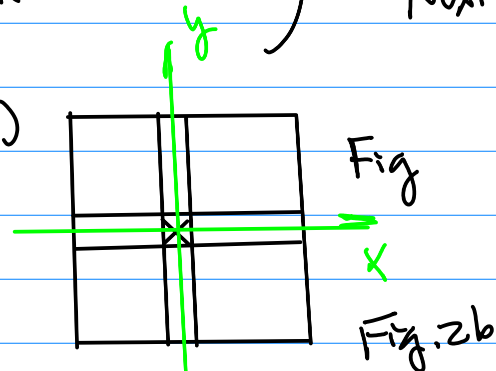


Fig.2b

(2) Flip the Kernel w.r.t. x-axis. (since $K(-x,y)$ changed to $K(-x,-y)$)

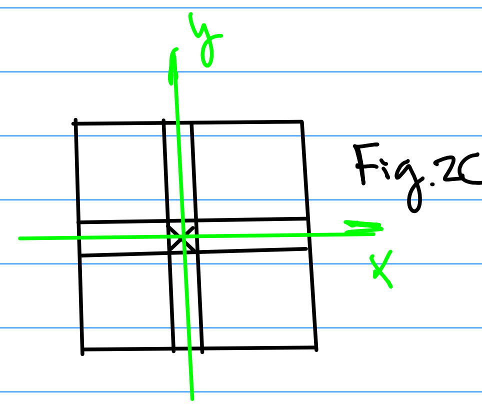
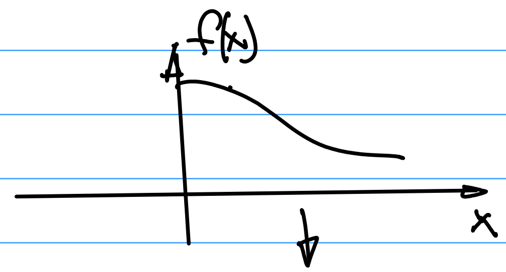
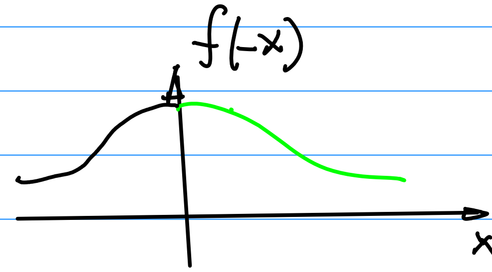


Fig.2c

Note: for 1D Case





Next, make $K(-x,-y)$ to $K(x_0-x, y_0-y)$

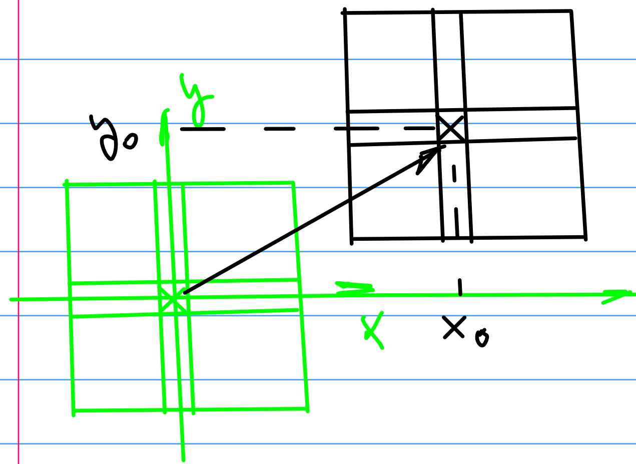


Fig. 2d

Note: $\frac{k-1}{2}$ for the top Rows
$\frac{k-1}{2}$ for the Bottom Rows
Lost

$$2 \times \frac{k-1}{2} = k-1$$

Similarly for the col.s.

Therefore,
Input Dimension $M \times N$ → Output After Convolution $M-(K-1) \times N-(K-1)$ ...(3)

Now, for Discrete/Digital Image (OR Feature plane)

Replace $\iint$ → by $\sum_{k_1} \sum_{k_2}$

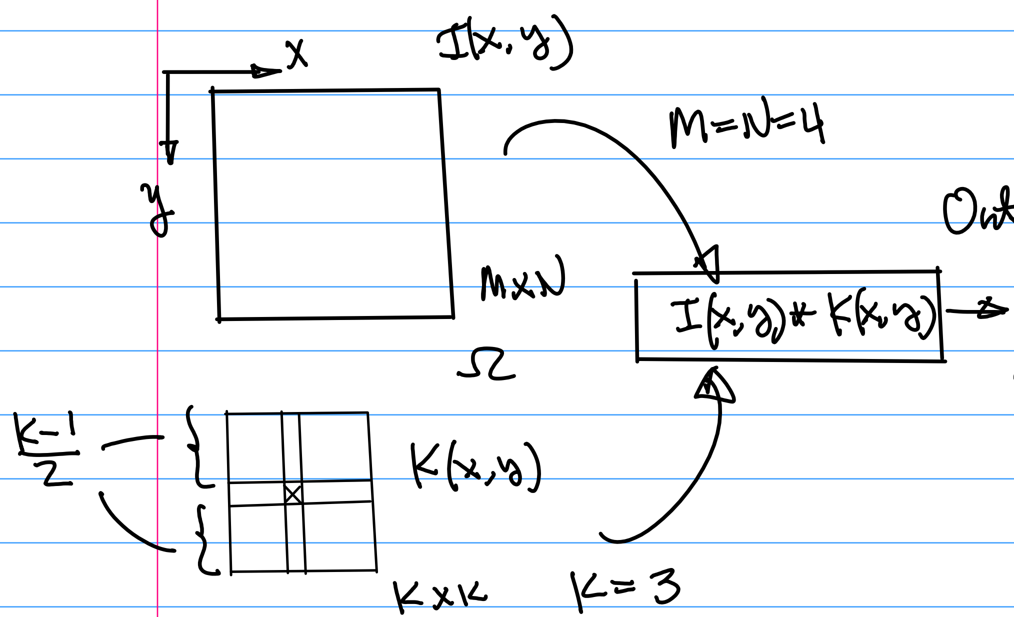
$$\boxed{c(n_1, n_2) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} a(k_1,k_2) b(n_1-k_1, n_2-k_2)} \quad \cdots (2)$$

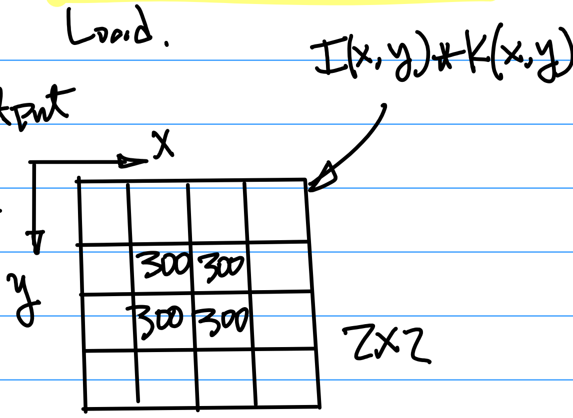
Image Kernel

Sept. 5 (Tue)

Note: 1° CANVAS OpenCV, Anaconda Installation;
2° Homework Coming ON Thursday.
a. Installation of T.F. Version 2.0 or Higher;
b. ChatGPT 3.5 API Down Load.

Dimension Change Due to the Convolution.





Example: Continuation on Convolutional Layers and Architecture for Hand written Digits Recognition.

Based on Eq(3), pp.9.

$M \times N$ Image $\xrightarrow[K \times K \text{ Convolution}]{}$ $M-(K-1) \times N-(K-1)$

From MNIST Architecture, the Input Image $I(x, y)$ with $M \times N$ $(M = N = 28)$. the output feature Layers Resolution 24×24. $\rightarrow$ $K = \underbrace{(28-24)}_{\text{Resolution difference.}} + 1$

Now, the total No. of feature layers $= n|$

Since one Convolution (e.g. using one $K \times K$ Kernel to perform Convolution) produces 1 feature Layer,

Therefore

$n|$ Convolutional feature Layes is the result of $n|$ Convolutions, e.g. $n|$ different Kernels.

Note: Multiple feature Layers are due to the inspiration of Human Visual Perception System, e.g. "Early" Vision System.

LoG Laplace of Gaussian.

Gabor Operators

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad \cdots (2)$$

$$G(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{\cancel{2\sigma^2}\ 2(\sigma_x^2 + \sigma_y^2)}} \quad \cdots (1)$$

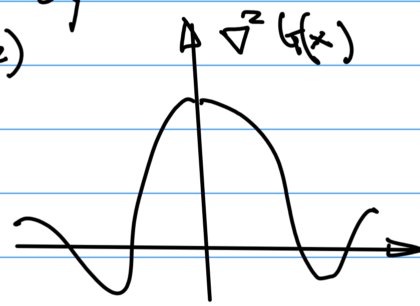


Fig 1.

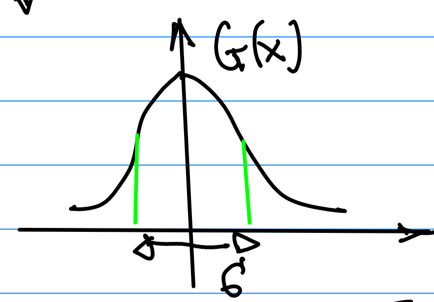


Fig.2

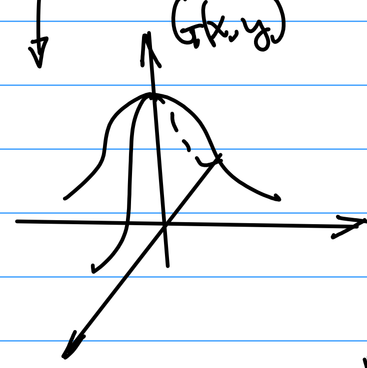
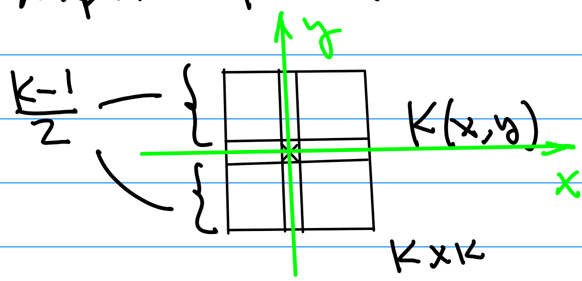


Fig.3

CMPE258

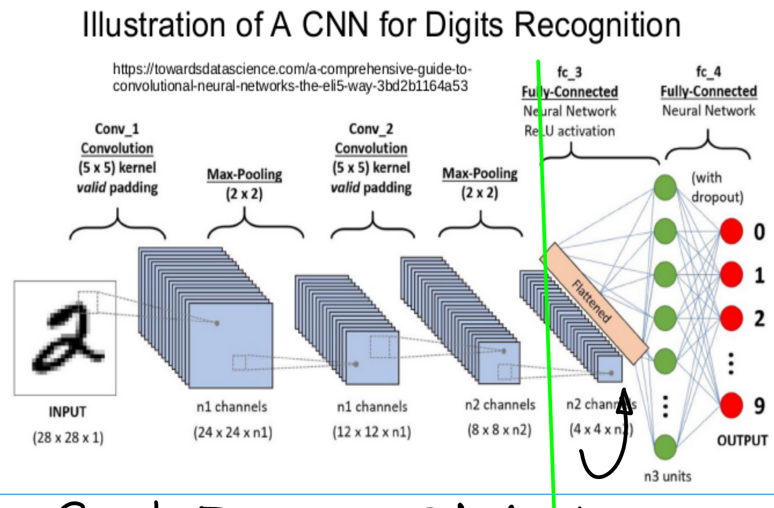
F2023

Take G(x,y) in Fig.3, Map it to K x K.



n1 Kernels Produce n1 Layers.

Now, consider the Architecture of MNIST



Consider Re-arrange 2D feature Layers / Each into 1D Neurons.

Now, consider a technique for feature Reduction,

Given an image I(x,y)

0	0	100	100
0	0	100	100
0	0	100	100
0	0	100	100

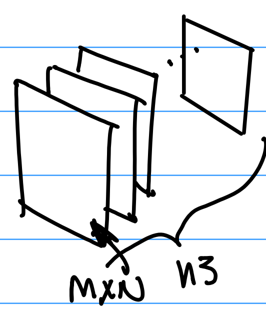
Feature Layer

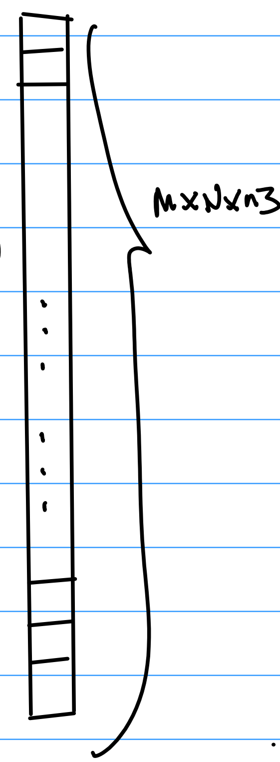
Max pooling $\longrightarrow$

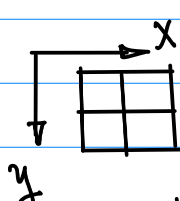
0	100
0	100

Max{0,0,0,0}

Mapping $\longrightarrow$







K x K
where K=2

Example: Sample code OpenCV for Displaying an image

```
cv2.imread(imageName, cv2.IMREAD_COLOR)
cv2.imshow(window_name, src)
```

Sept. 7 (Th).

Note: 1° About OpenCV Reference/Sample Code.

a) On github. for Video Capture

opencv / deep-learning-2022s / 2022S-109c-v2-localization-contours.py

hualili Add files via upload

to get A Single Frame.

b) Sample for Single frame, e.g. .jpg, .png etc Image Input.

2022S-104d ~ Python Code

c) Script for Conda Environment On github:

2022S-104b ~

" -104c ~

d) Readme for Creating Conda Environment.

Note 2. Homework Due A week from today (Sept. 17, Sun)

Requirements:

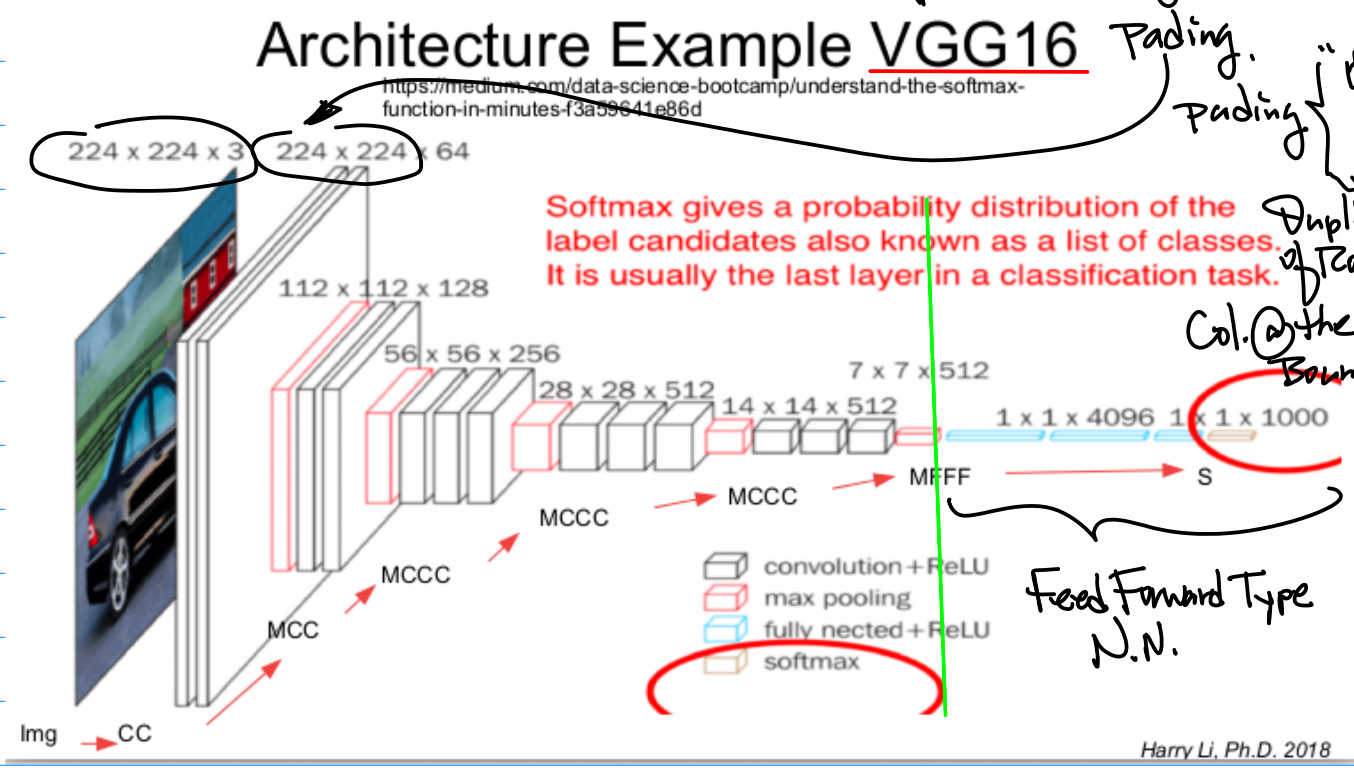
a) Installation of T.F. Version 2.0 or higher

b) Screen Capture that shows T.F. installed Successfully. (with Personal Identifier)

Submission on CANVAS.

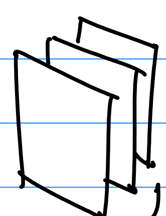
Example: Convolutional Layer, Architecture Analysis.

Note 1. Architecture ; Note 2: Pading v.s. No Pading. "0"s pading; Duplication of Const. Col. @ the Boundary



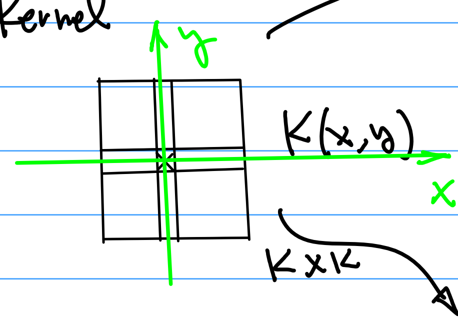
Feed Forward Type N.N.

Note 3: Convolution Discussion.



Colour Image

Kernel



One Example

$$\begin{array}{c} \xrightarrow{x} \\ y\downarrow \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} \end{array}$$

LoG $\nabla^2 G(x,y) \rightarrow$ PP.10.

$\nabla^2 G(x,y;\mu_x,\mu_y,\delta_x,\delta_y)$

$\rightarrow \left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right) G(x,y;\mu_x,\mu_y,\delta_x,\delta_y)$

$I(x,y) = (r(x,y), g(x,y), b(x,y))$

To perform Convolution. $\rightarrow K(x,y)$

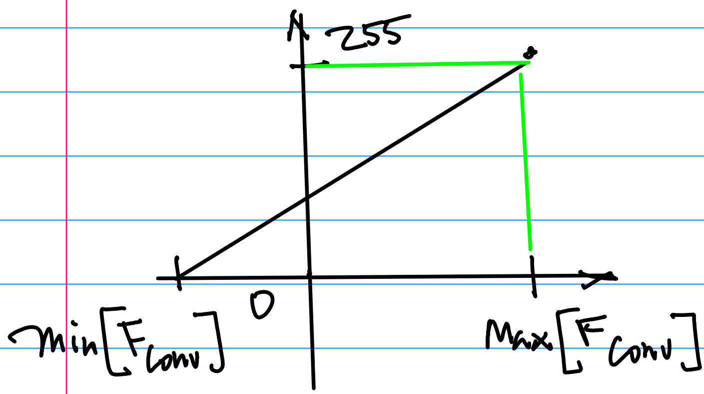
First, kernel is one channel in this case;

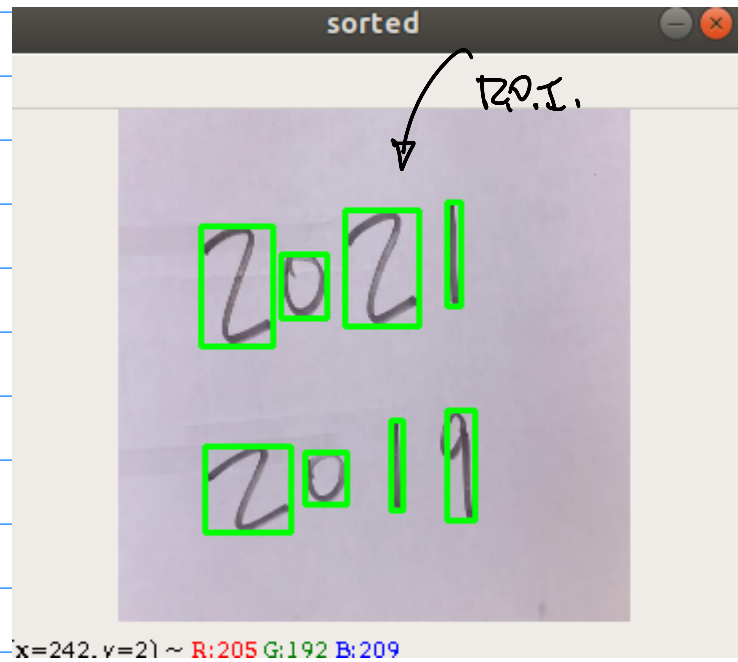
Then, Convolutions. $r(x,y) * K(x,y)$, $g(x,y) * K(x,y)$, $b(x,y) * K(x,y)$

CmpE258

F2023.

$$F_{conv}(x,y) = \frac{1}{3}[r(x,y) * K(x,y) + g(x,y) * K(x,y) + b(x,y) * K(x,y)]$$





Region of Interests.

Consider A Design of S.I.D Recognition System Using T.F & MNIST CNN as a Processing engine.

Requirements:

1° Live CAM Input,
720P or 1080P
1280x720 (M x N)
Col. Row.

1920x1080
(M x N)
Col. Row.

2° Printer Paper (Blank/White)
Black mark to Writ 4 Digits

3° Localize R.O.I. ON Each Digit

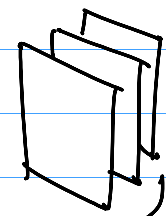
Sept.12 (Tue).

Note 1. For Homework 1 Extended to the Saturday, Requires the Submission of Python Code.

Note 2. Team Project.
Ref: 2023F-103-project-team-2023-9-12.pdf
ON CANVAS as well.
Due on Nov.27 (Monday) 11:59 P.M.

Example: Suppose a color image I(x,y) is given below, and a Convolution Kernel K(x,y) is given as follows

A color Image I(x,y)



Colour Image

x, y

$\frac{1}{6}$

-1	-1	-1
0	0	0
+1	+1	+1

3x3.

Find Convolution Result.

$I_r(x,y)$

0	0	0	255
0	0	0	0
0	0	0	0
0	0	0	1

$I_g(x,y)$

0	0	100	0
0	0	0	100
0	0	5	100
0	0	2	0

$I_b(x,y)$

0	0	255	0
0	0	0	0
0	0	0	0
0	0	255	0

Each channel $*$ $\frac{1}{6}$

-1	-1	-1
0	0	0
+1	+1	+1

Results:

0	-255
0	1

-95	5
2	-98

-255	-255
255	255

Note: In the future, we have Need to Reduce the Number of feature Layers

By introducing Newer Kernel(s) for Convolution.

Homework: Due 1 week from Today. (Sept. 19th).

1° Installation of T.F., Screen Capture the installation Result. (w/ Personal ID).

2° OpenCV Code to Handle

a. Live Web CAM Input Video, and Display it;

b. file input, MPEG4 format Video and display it.

Note: Please use A Printer Paper, Write with a Black marker, 4 Digits of your SID.

Submission: the Code & Video Clips.

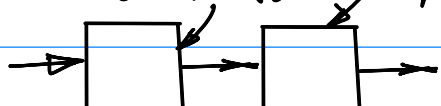
Preprocessing for Hand Written SID Recognition System Design.

Live Web CAM Input

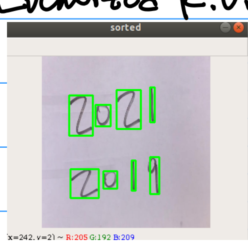
Output (I)

Output stage II

Make I(x,y) Processed to match the Resolution & Channel Requirements of the DCNN Engine Input (e.g. 28x28x1) without Distortion!



Localized R.O.I.



Region of Interests

Ref:

Homework: Due 1 week from today. Sept. 21st.

Use your OpenCV for the following preprocessing functions.

1° Convert the color Image $I(x,y)$ to a gray scale Image. $I_G(x,y)$

2° Binarize the gray scale image to Obtain a Binary Image $B(x,y)$.

3° Perform CANNY Edge Detection on $I_G(x,y)$, and display the edge map.

4° Perform Gaussian Blur on the Gray Scale Image.

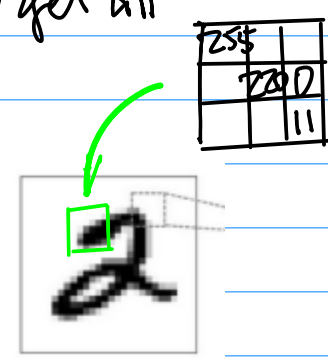
Submission:

5° Screen Capture of 1° to 4°. with Personal Identifier.

Example: Continuation of project Discussion. Preprocessing.

The First Objective: To get all the Bounding Boxes

Color Video → Gray Scale Image.



INPUT
(28 x 28 x 1)

C. Filtering Operation, Conducted by Convolution w/ Predefined Kernel Coefficient.

B. To Be Continued.
Edge Detection CANNY, LoG, Gabor

Note: Input should be a grayscale (8 bits), Output should be 8 bit (1 channel)

A.
Binarization. To Obtain A Representation of A Digit.
Note: Binarization may lead to the Loss of useful information.

Example: Image Binarization.

Given $I_G(x,y)$ or Simply $I(x,y)$
Binarization is defined as

$$B(x,y) = \begin{cases} 255 & \text{if } I(x,y) \geq T \\ 0 & \text{o/w} \end{cases} \quad \ldots (1)$$

Note: T is set for the entire image for Now.

Example from PP. 18, Lecture Notes on Github, 2023S.

Note: Treat the top Left Corner as (1,1) for the Binarized Image Descriptors Calculation (to Void Skewed Result).

Most Important Descriptors are those Built Based on Moments.

$$\frac{\iint_{\Omega} (x-\bar{x})^{m} (y-\bar{y})^{n} B(x,y)\,dx\,dy}{\iint_{\Omega} B(x,y)\,dx\,dy} \quad \ldots (2)$$

order

Image Plane. $M \times N$

Binary Image

Area, Size

$$A = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} B(x,y) \quad \ldots (3)$$

$$\sum_{y=0}^{N-1} \sum_{x=0}^{M-1} (x-\bar{x})^{m} (y-\bar{y})^{n} B(x,y) \quad \ldots (4)$$

Where

$$\bar{x} = \frac{\sum_{y=0}^{N-1} \sum_{x=0}^{M-1} x\,B(x,y)}{A} \quad \ldots (4\text{-}a)$$

and

$$\bar{y} =$$