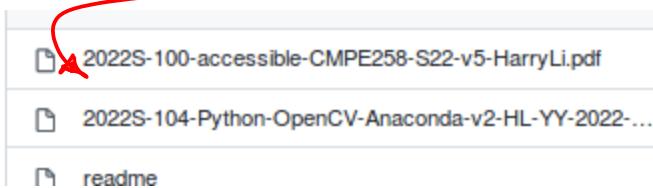


CMPE258 Spring2022

Feb1st. Organizational meeting.

1. Today's Topics "Greensheet"



Naming Convention Yr + Semester + ID
+ Name + Date

Contact Information: Email: hua.li@sjsu.edu

Text message to (650) 400-1116.

Office Hours: M.W. 4:30-5:30pm.

Zoom (link to be shared in the email)

Join from PC, Mac, Linux, iOS or Android: <https://sjsu.zoom.us/j/85616325978>

pwd=MzIRbDJXVHBDQ2gIU0RPM2tYc045Zz09

Password: 451032

On-Line materials on github

<https://github.com/hualili/opencv/tree/master/deep-learning-2020S>

<https://github.com/hualili/opencv/tree/master/deep-learning-2022s>

Also, CANVAS → mostly for Assignments and projects.

All Assignment / Projects are posted on Both github & SJSU CANVAS.

Lecture Material consists of P.P.T. posted on github, and Lecture Notes (White-Board Written Notes)

Core Emphases of the Class: Deep Convolutional

Neural Networks, And Their Application in Image Analysis, Video Analysis.

1. Text Book:

<https://github.com/hualili/opencv/blob/master/IP120-AI-DL/2018F/2018F-6-DeepLearningCh02.pdf>

2. Computer Vision Book By Horn

as a reference for Convolution & Image Segmentation, Contours Analysis (Binary Image).

3. OpenCV Reference Book (and Edition) together with Online Document (OpenCV).

Note: OpenGL (GL: Graphics Library) is just for Reference purpose, no need for this class. (but maybe helpful for the future research).

Unity is game Development Platform, interactive 3D Graphics Design platform.

Programming Languages:

1 Python. 3.6 or 3.7

2 C/C++ Feb. 8th.

Homework (Due A week from today)
No Submission. Submit A Screen

Capture that shows OpenCV installed successfully, with jpg or png file with Naming of the file as follows:

FirstName_LastName-SID-OpenCV.jpg

This Homework will be posted on CANVAS, Submission is on CANVAS.

Homework, Installation of Tensor Flow, Due 2 weeks
Feb 15th.

Submission: Screen Capture that shows the installation is successful. Submission on CANVAS.

Submit jpg, png file with the Naming convention as follows:

FirstName_LastName-SID-TF.jpg

Note: Optional, for Edge AI Computing, Consider using NVidia Jetson NAND (4GB) version.

5% Bonus

Grading:

Homework, projects : 30%
5% 25%

Project 1. Computer Vision for Preprocessing, plus Deep Convolutional Neural Nets
To give Real Time Detection Result of Last 4 Digits of a Student ID.

10%

Project 2. "Semester Long" project, with technical requirements (List)
Team project. 4 person Team.

Each person has clear definition of the tasks (Programming/Coding) And Balanced Contribution.

Final PPT, Demo Presentation
15%.

Midterm Exam: 30%.
Need to use your Laptop Computer, to Run/Execute code, modify the code.

Final 40%

Introduction

Topics: Neural Networks formulation
(Basic Building Blocks)
Digital Images/Videos.

Example: A Single Neuron Formulation
(Some kind Brain Cell)

Step 1. Summation function.



Fig 1.

" Σ " Summation function.

Note: $\sum_{k=1}^N x_k = x_1 + x_2 + \dots + x_N$

Step 2. Inputs

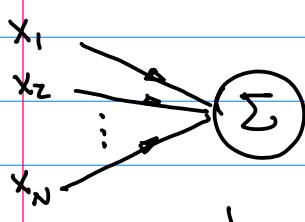


Fig 2.

Step 3. Weights (Knowledge)

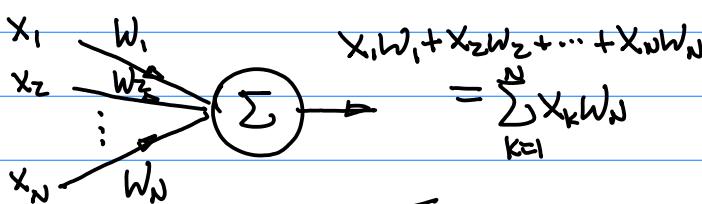


Fig 3.

$$\sum_{k=1}^N x_k w_k = x_1 w_1 + x_2 w_2 + \dots + x_N w_N \dots (1)$$

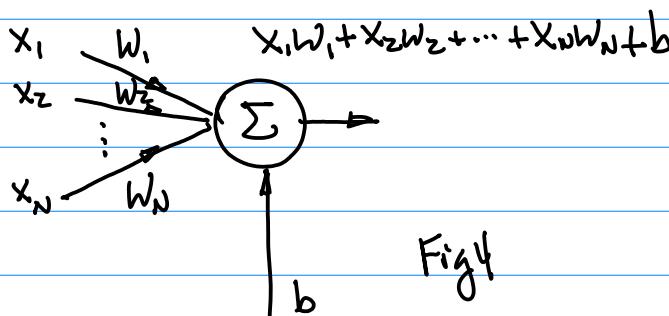


Fig 4

$$\sum_{k=1}^N x_k w_k = x_1 w_1 + x_2 w_2 + \dots + x_N w_N + b \quad \underline{\underline{}} \quad \dots (1b)$$

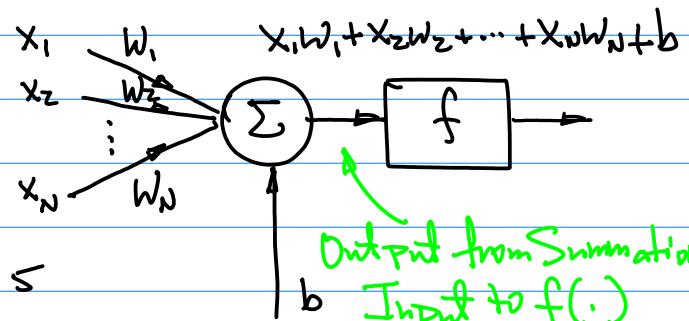


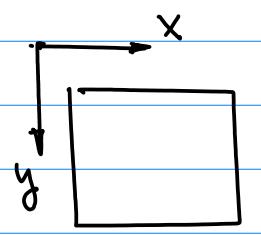
Fig 5

Note: Activation function f , denoted as $f(\cdot)$ (A function of Independent Variable " \cdot ", or A function of Input " \cdot ")

$$\begin{aligned} f(\cdot) &= f(x_1 w_1 + x_2 w_2 + \dots + x_N w_N + b) \\ &= f\left(\sum_{k=1}^N w_k x_k + b\right) \\ &\dots (2) \end{aligned}$$

Summary: The output of a single is given by Eqn (2). Where Activation function $f(\cdot)$ can take different forms, it affects the Learning, Learning Speed.

Example: Digital Image, $I(x, y)$



$M \times N$

$I(x,y)$

Intensity, And/or

Color of An Image

(x, y) Location of A picture element, "pixel"

In Case of a Single pixel, (x, y) is the location of this pixel, I is Color/Intensity of the pixel

Note: For An Image $I(x,y)$

its features include Resolution $M \times N$

Pixel Depth:

bpp (Bit per pixel)

1. Notation for Input $x_i, i=1, 2, \dots, N$

$\{x_i | i=1, 2, \dots, N\} \dots (1)$

$(x_1, x_2, \dots, x_N) \dots (1b)$

$I(x,y)_{mn}$ OR $I(x,y)$

$\frac{1024 \times 768}{\text{pixels/Row}} = \text{Rows}$

Introduce Superscript j for Experiment j

Input $x_i^j, i=1, 2, \dots, N; j=1, 2, \dots, P$

Hence Eqn(1) Becomes

$\{x_i^j | i=1, 2, \dots, N; j=1, 2, \dots, P\}$

$(x_1^j, x_2^j, \dots, x_N^j)$ for Experiment j

2. Notation for Weight

$w_i, \text{ for } i=1, 2, \dots, N$

Hence,

$(w_1, w_2, \dots, w_N) \dots (2)$

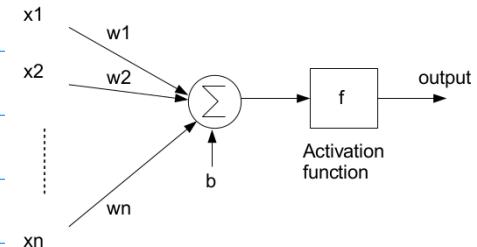


Fig.1

$I(x,y)_{mn}$ OR $I(x,y)$

$\frac{1024 \times 768}{\text{pixels/Row}} = \text{Rows}$

For A color Image, A pixel depth very often is equal to 24 (bpp)

r, g, b Primitive color of red (r), green (g), blue (b) has 8 bits quantization level, e.g.

r: [0, 255], g: [0, 255], b: [0, 255].

Feb 8th (Tue)

Today's Topics : 1° Introduction, Basic Building Blocks, Math Formulation.
2° Sample Python Code for OpenCV.

3. Inputs & weights

$$x_i \quad w_i$$

$$w_i x_i \text{ for } i=1, 2, \dots, N$$

$$(x_1, x_2, \dots, x_N) \cdot (w_1, w_2, \dots, w_N)$$

$$= w_1 x_1 + w_2 x_2 + \dots + w_N x_N = \sum_{i=1}^N w_i x_i$$

$$(w_1, w_2, \dots, w_N) \cdot (x_1, x_2, \dots, x_N)$$

$$= w_1 x_1 + w_2 x_2 + \dots + w_N x_N = \sum_{i=1}^N w_i x_i$$

4. Transfer Function, Denoted as

$$\tilde{h} \text{ OR } h(\cdot)$$

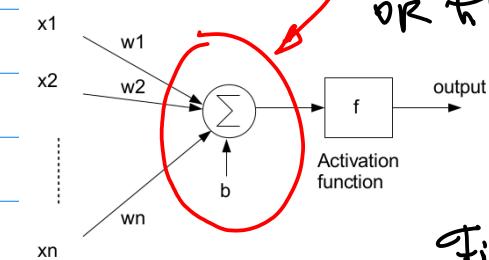


Fig. 2

$$h = \sum_{i=1}^N w_i x_i + b = w \cdot \vec{x} + b$$

... (3)

↑
Offset ("Bias")

\tilde{h} , or $h(\cdot)$, or Eqn (3), or
 $h(w_i; b)$, or $h(w_i)$

5. Activation Function: f

Acts like a switch, ON/OFF &
Attenuate the Output

$$f, f\left(\sum_{i=1}^N w_i x_i + b\right), \text{ or}$$

$$f(h(w_i; b)), \text{ or } f(h(\cdot))$$

The output of A Neuron is denoted as

$$y, \text{ and } y = f\left(\sum_{i=1}^N w_i x_i + b\right) \\ = f(h(w_i; b)) \quad \dots (4)$$

5. Outputs for A Neural Network =

$$y_k, \text{ for } k=1, 2, \dots, Q$$

if $Q=1$, y_1 , also better to just y , for a single neuron as y

Question: What output? whose output?

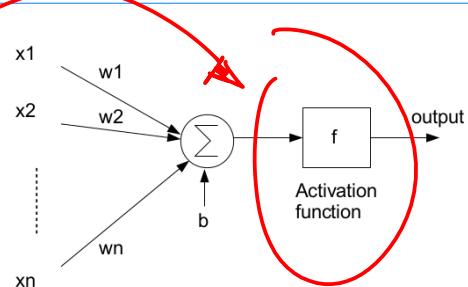
Output from the Network (or Network)

$$\tilde{y} \text{ (Tilde)}$$

Ground Truth is denoted as y

6. Loss function, Objective Function,

OR difference



"Supervised Learning"

For Experiment j (multiple experiments)

$$\begin{aligned} \text{Output from the NN (Neuron or Network)} & \quad \text{Ground Truth for the experiment} \\ \text{for } j=1, \quad \tilde{y}_j - y_j & \\ j=2, \quad \tilde{y}_2 - y_2 \\ \vdots \\ j=P, \quad \tilde{y}_P - y_P \end{aligned}$$

Therefore, put all these differences (Loss) together

$$\begin{aligned} (\tilde{y}_1 - y_1) + (\tilde{y}_2 - y_2) + \dots + (\tilde{y}_P - y_P) \\ = \sum_{k=1}^P (\tilde{y}_k - y_k) \quad \dots (b) \end{aligned}$$

Note:

$$y = |x|$$

To deal with the issue of Absolute Value of a function, Let's square it.

Hence Eqn (b) becomes

$$\sum_{k=1}^P (\tilde{y}_k - y_k)^2 \quad \dots (bb)$$

7. Objective function

$$L \triangleq \sum_{k=1}^P (\tilde{y}_k - y_k)^2 \quad \dots (7)$$

OR,

$$L(w_i) \triangleq \sum_{k=1}^P (\tilde{y}_k - y_k) \quad \dots (7b)$$

$$= \sum_{k=1}^P (\tilde{y}_k(h(w_i)) - y_k) \quad \dots (7c)$$

8. To generalize the result in Eqn

(7c) for multiple output, we have the following formulation.

From Eqn (7)

$$\sum_{k=1}^P (\tilde{y}_{k_1} - y_{k_1})^2$$

k_1 : Output k_1 for multiple output

$$y_j, j=1, 2, \dots, k_1, k_1+1, \dots, M$$

We can count All outputs loss.

$$\sum_{k=1}^P (\tilde{y}_{j_1} - y_{j_1})^2, \sum_{k=1}^P (\tilde{y}_{j_2} - y_{j_2})^2, \dots$$

$$\dots \sum_{k=1}^P (\tilde{y}_{j_M} - y_{j_M})^2$$

Hence

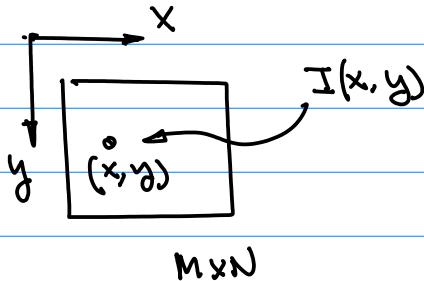
$$\sum_{k_1=1}^m \sum_{k_2=1}^P (\tilde{y}_{j_{k_1}} - y_{j_{k_1}})^2 \quad \dots (8)$$

CMPE258 Feb 8, 22

$$L(w_i) \triangleq \frac{1}{2} \sum_{h_2=1}^m \sum_{k_2=1}^p (\hat{y}_{j_{h_2}}^{k_2} - y_{j_{h_2}}^{k_2})^2$$

... (8b)

Example: Digital Image $I(x, y)$



Installation of OpenCV → Python OR
(C/C++, But
Python is Better)

Different Packages
for ML/DL, may need different version
of Python, And different packages

Anaconda is a well developed,
Adopted tool for Package management.

1. Check github Class Reference

2022S-104a-Python-OpenCV-Anaconda-v2-HL-YY-2022...

Note: For installation of Anaconda,
Check github document,
2022S-104-~

Once Anaconda is installed, then
Let's take a look the configuration for

Tensorflow & OpenCV Environment

Step 1.

Running O

Step 1. Create configuration
file .yml for CPU and GPU
version

conda-cpu.yml

```
1 name: yolov4-cpu
2
3 dependencies:
4   - python==3.7
5   - pip
6   - matplotlib
7   - opencv
8   - pip:
9     - opencv-python==4.1.1.26
10    - lxml
11    - tqdm
12    - tensorflow==2.3.0rc0
13    - absl-py
14    - easydict
15    - pillow
```



Name of the environment

17 lines (16 sloc) | 269 Bytes

```
1 name: yolov4-gpu
2
3 dependencies:
4   - python==3.7
5   - pip
6   - matplotlib
7   - opencv
8   - cudnn
9   - cudatoolkit==10.1.243
10  - pip:
11    - tensorflow-gpu==2.3.0rc0
12    - opencv-python==4.1.1.26
13    - lxml
14    - tqdm
15    - absl-py
16    - easydict
17    - pillow
```

Step 2. Create the environment

Step 2. Configure/create the environment in the folder you will run your openCV program by

Sconda env create -f conda-cpu.yml #for CPU
Sconda env create -f conda-gpu.yml #for GPU

file names to be used

Step 3 Activate the environment

Step 4. Run Your Python Code (OpenCV or T.F.)

Running OpenCV with Conda Environment

Step 1. Create configuration file .yml for CPU and GPU version

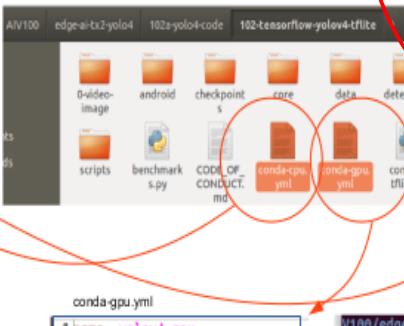
conda-cpu.yml

```
1 name: yolov4-cpu
2
3 dependencies:
4   - python==3.7
5   - pip
6   - matplotlib
7   - opencv
8   - pip:
9     - opencv-python==4.1.1.26
10  - lxml
11  - tqdm
12  - tensorflow==2.3.0rc0
13  - absl-py
14  - easydict
15  - pillow
```

conda-cpu.yml

```
name: yolov4-cpu
dependencies:
- python==3.7
- pip
- matplotlib
- opencv
- pip:
- opencv-python==4.1.1.26
- bml
- lqdm
- tensorflow==2.3.0rc0
- absl-py
- easydict
- pillow
```

Name of the environment to be created



conda-gpu.yml

```
1 name: yolov4-gpu
2
3 dependencies:
4   - python==3.7
5   - pip
6   - matplotlib
7   - opencv
8   - cudnn
9   - cudatoolkit==10.1.243
10  - tensorflow-gpu==2.3.0rc0
11  - opencv-python==4.1.1.26
12  - lxml
13  - tqdm
14  - absl-py
15  - easydict
16  - pillow
```

Step 2. Configure/create the environment in the folder you will run your openCV program by

```
$ conda env create -f conda-cpu.yml      #for CPU
$ conda env create -f conda-gpu.yml      #for GPU
```

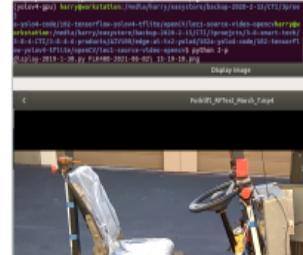
Step 3. Activate the environment with the name that you have created by

```
$ conda activate yolov4-cpu #for CPU
$ conda activate yolov4-gpu #for GPU
```

Name of the environment in the .yml file

Step 4. Now run your Python OpenCV program

\$ python myOpenCV.py image.png



Name, Program Name, Version, Status(Delay, Release), Date, Note

a. Program Header

```
"""
pdisplay.py
Demo read and display image
"""
```

```
import sys
import cv2
import numpy as np
```

```
#
#main(sys.argv[1:])
window_name = 'Display Image'
```

```
imageName = sys.argv[1] #get file name from command line
```

```
src = cv2.imread(imageName, cv2.IMREAD_COLOR)
```

```
if src is None:
    print ('Error opening image!')
    print ('Usage: pdisplay.py image_name\n')
```

```
19
```

b. cv2 OpenCV

Capture of the window

Note: Install Anaconda → Create the environment for T.F. & OpenCV

Execute OpenCV → Activate the environment
Sample Code (from the class github)



```
1 """
2 pdisplay.py
3 Demo read and display image
4 """
5 import sys
6 import cv2
7 import numpy as np
8
9 #main(sys.argv[1:])
10 window_name = 'Display Image'
11
12 imageName = sys.argv[1] #get file name from command line
13
14 src = cv2.imread(imageName, cv2.IMREAD_COLOR)
15
16 if src is None:
17     print ('Error opening image!')
18     print ('Usage: pdisplay.py image_name\n')
19
```

CMPE258 Feb 8, 22

2nd Ref on the github)

9

2022S-103a-notation-neuro-loss-function-2022-2-8-1.pdf

```
22 while True: (3)  
23     cv2.imshow(window_name, src)  
24  
25     c = cv2.waitKey(500)  
26     if c == 27: #ESC  
27         break  
28  
29     ind += 1
```

Keyboard input to exit.

Note: These Python functions are registed (memorize them!)

- (1) import cv2 (2) cv2.imread()
- (3) cv2.imshow().

Homework (Due A week from today)
Feb 15

- 1° Installation of OpenCV.
- 2° Installation of Anaconda.
- 3° Use your Smartphone to take photo, and Save it for OpenCV Program
- 4° Write A Python Program (Ref. Code from the class github is ok)

to display:

- a. Your Name + SID (4 Digits)
- b. Your Smartphone picture.

- 5° Submission: One pdf file And Zipped. On CANVAS .

Feb 15

- Ref: 1° 2022S-103a-Notation
- 2° 2022S-103C

VIII. MINIMIZE THE LOSS FUNCTION

$$\frac{\partial L}{\partial w_{i,k}} = \frac{\partial}{\partial w_{i,k}} \frac{1}{2} \sum_{j=1}^P \sum_{i=1}^M (\hat{y}_i^j - y_i^j)^2 \quad (24)$$

Now, consider to Optimize Neural Network

Performance By minimizing the Loss function.

Mathematical Background

Derivative(s) → Partial Derivatives

a.



c. Gradient

↑

A special kind of Gradient: Descent

To Reduce Loss

"Steepest" ~

↑

Training A Neural Network

↑

Generalize the concept:
"Learning" (Supervised learning)

Derivatives.

Definition $\frac{df(x)}{dx} \triangleq \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$... (1)

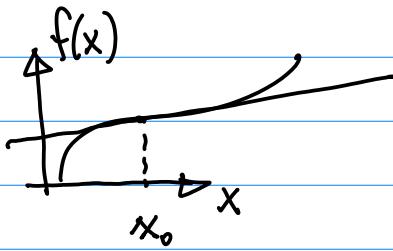
Note Eqn(1) is Based on "Forward Difference".

$f(x + \Delta x)$

forward of x



2022S-103a-notation-neuro-loss-function-2022-2-8.pdf



if $f'(x) = \frac{df}{dx} > 0$, then $f(x)$ at $x=x_0$ will increase

if $f'(x) < 0$, then $f(x)$ will decrease

if $f'(x) = 0$, then $f(x)$ is unchanged.

For a function $f(x_1, x_2)$, the partial

Derivative of $f(x_1, x_2)$

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \dots (z)$$

$$\frac{\partial f(x_1, x_2)}{\partial x_1} \stackrel{def}{=} \lim_{\Delta x_1 \rightarrow 0} \frac{f(x_1 + \Delta x_1, x_2) - f(x_1, x_2)}{\Delta x_1} \dots (za)$$

$$\frac{\partial f(x_1, x_2)}{\partial x_2} \stackrel{def}{=} \lim_{\Delta x_2 \rightarrow 0} \frac{f(x_1, x_2 + \Delta x_2) - f(x_1, x_2)}{\Delta x_2} \dots (zb)$$

$$\frac{\partial f}{\partial x_1} > 0, f \uparrow \text{as } x_1 \uparrow$$

$$\frac{\partial f}{\partial x_2} > 0, f \uparrow \text{as } x_2 \uparrow$$

Define gradient of a function $f(x_1, x_2)$

(Note: Connection to Neural Network,

And a Single Neuron. inputs x_1, x_2, \dots, x_n

And Weights w_1, w_2, \dots, w_N

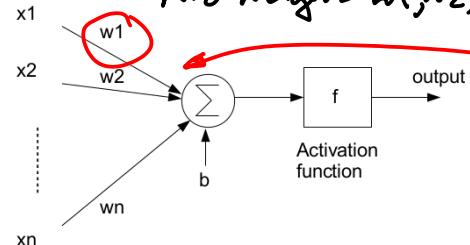


Fig 1.

Defined a vector

$$\begin{pmatrix} \frac{\partial f(x_1, x_2)}{\partial x_1} \\ \frac{\partial f(x_1, x_2)}{\partial x_2} \end{pmatrix} \text{ or } \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix}$$

We can use github Lecture ... (3)

Notes to expand Eqn (3) to N-Dimensional Case

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \dots (3b)$$

For A Single Neuron, we have gradient function as follows.

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial w_1} \\ \frac{\partial f}{\partial w_2} \\ \vdots \\ \frac{\partial f}{\partial w_N} \end{bmatrix} \dots (3c)$$

Guideline for the Development of "Training" Technique :

Negative Gradient gives the steepest Descent of a Loss function f .

Note: Negative gradient gives the Best way to adjust w_i in a Neural Net

To make the Neural Net to Reach to Optimal Solution.
e.g. minimized loss function.

The Approach to Verify the Above Technique:

Generalized loss function, as f , or $f(\cdot)$, or $f(x_1, x_2, \dots, x_n)$, or

$f(\underline{w})$, or $f(w_1, w_2, \dots, w_n)$, or

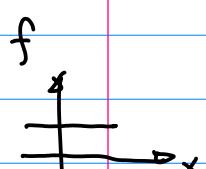
$f(w)$

Simple/Generalized Description to connect independent Variable x_1, x_2, \dots, x_n (or w_1, w_2, \dots, w_n) to the function f (loss function)

Taylor Expansion for $f(x)$

$$f(x) = f(x_0) + \frac{f'(x)}{1!}(x-x_0) + \frac{f''(x)}{2!}(x-x_0)^2 + \dots + \frac{f^{(k)}(x)}{k!}(x-x_0)^k + \dots \quad (4)$$

= Constant + 2nd (Linear function) +



$$\frac{f'(x)}{1!}(x-x_0)$$

Coefficient, Slop. $a(x-x_0)$
 $y = ax+b$ Linear function

$$\frac{f''(x_0)}{2!}(x-x_0)^2 \quad \text{Quadratic Term.}$$

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + T_n(x)$$

$$f(x) \approx f(x_0) + \frac{f'(x_0)}{1!}(x-x_0)$$

Linear function

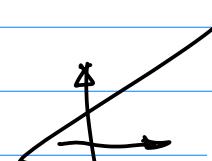
$$f(x) = f(x_0) + \frac{f'(x_0)}{1!} \Delta x \dots (5)$$

We expand function f (Loss) into Multi-dimensional Case, e.g. x_1, x_2, \dots (or w_1, w_2, \dots).

Choose work on 2D Case. e.g.

$$f(x_1, x_2) = f(x_{10}, x_{20}) + \frac{\partial f}{\partial x_1}(x-x_{10}) + \frac{\partial f}{\partial x_2}(x-x_{20}) + \dots \dots (6)$$

$$f(x_1, x_2) \approx f(x_{10}, x_{20}) + \frac{\partial f}{\partial x_1}(x-x_{10}) + \frac{\partial f}{\partial x_2}(x-x_{20}) \dots (7)$$



$$f(x_1, x_2) \approx f(a, b) + \frac{\partial f}{\partial x_1}(x_1 - a) + \frac{\partial f}{\partial x_2}(x_2 - b) \quad (6)$$

Question, if we change independent Variable x_1 and x_2 by the negative gradient of f , then the function f is decreased or increased?

Based on the Theory, we should have loss function f decreased in an optimized way, e.g. "Steepest Descent".

$$(x_1^{k+1}, x_2^{k+1}) = (x_1^k, x_2^k) + [-\eta(\nabla f)^T] \quad (5)$$

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix}$$

$$\begin{cases} x_1^{k+1} = x_1^k + (-\eta) \frac{\partial f}{\partial x_1} & \dots (8a) \\ x_2^{k+1} = x_2^k + (-\eta) \frac{\partial f}{\partial x_2} & \dots (8b) \end{cases}$$

If we use Eqn.(8a), (8b) to update the independent Variables x_1, x_2 (or weights w_1, w_2), then you are sure to be able to find loss function f at this step $k+1$, is smaller than itself, e.g. f function at Step K.

$$f(x_1, x_2) - f(a, b) = (x_1 - a, x_2 - b) \nabla f \quad (10)$$

The Above equation, eqn(10), is Taylor Expansion

Note: from the Analysis below, we have:

$$f(x_1, x_2) - f(a, b) = (\Delta x_1, \Delta x_2) \nabla f = -\left(f_{x_1}^2 + f_{x_2}^2\right) \quad (11)$$

$$f_{x_1} = \frac{\partial f}{\partial x_1}$$

$$f_{x_1}^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 \geq 0$$

$$\text{Similarly } f_{x_2} = \frac{\partial f}{\partial x_2}, \quad f_{x_2}^2 = \left(\frac{\partial f}{\partial x_2}\right)^2 \geq 0$$

$$\text{Therefore } -\left[\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2\right] \leq 0$$

Hence

$$f(x_1, x_2) - f(a, b) \leq 0$$

$$f(x_1, x_2) \leq f(a, b)$$

$$x_1 - a = -\eta \frac{\partial f}{\partial x_1}$$

$$x_1^{k+1} - x_1^k = -\eta \frac{\partial f}{\partial x_1}$$

$$f(x_1, x_2) - f(a, b) = \frac{\partial f}{\partial x_1}(x_1 - a)$$

$$+ \frac{\partial f}{\partial x_2}(x_2 - b)$$

$$x_1^k = a, x_2^k = b \quad f(a, b)$$

$$f(x_1, x_2) \approx f(x_1^k, x_2^k) + \frac{\partial f}{\partial x_1}(x_1 - x_1^k) + \frac{\partial f}{\partial x_2}(x_2 - x_2^k) \quad \text{Eqn(7)}$$

PP.11

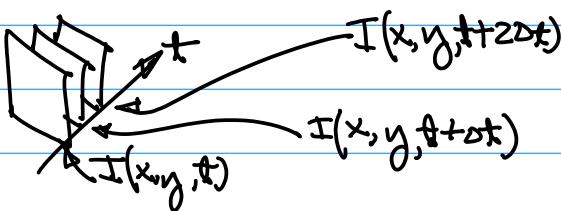
OpenCV

Note: One of the tools for Python

Programming for this class
is PyCharm.

Create A New Project \rightarrow Configuration
of the environment
& packages Needed

Example: OpenCV for Video.



Exercise Use your Smartphone to
Capture a video clip which
has the following feature:

1. White Background, Such as
Printer Paper or White Board,
then, write your last 4 Digits
on this white Background Paper/Board.

2. Be sure to use a marker to give
adequate contrast and width for
each mark. (1080P or 1080P Resolution)

3. Save the video (may have 20 to
Seconds video clip) for the future
use By Next Lecture.

Note: we will discuss Computer
Vision preprocessing techniques
in the next lecture.

Feb 22nd (Tue)

Today's Topics:

1. Theoretical Background on
Deep learning.
2. Preprocessing Techniques
for Handwritten Student ID
Recognition.

Single Neuron \rightarrow Loss function

Build multi-
Layer feed forward
NN \rightarrow Gradient Descent
Technique

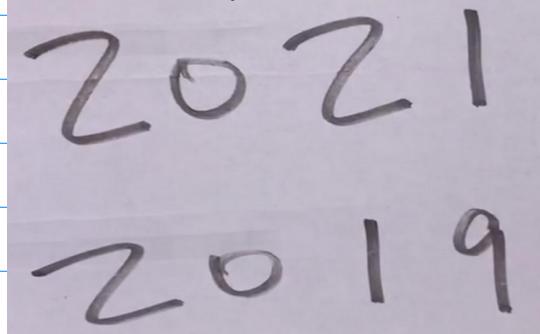
Back Prop. \rightarrow Deep Neural Network
 \rightarrow Convolutional NN
CNN]

Example: Pre-processing
Technique for handwritten Digits
Recognition.

2022S-107a-video-digits1.mp4

2022S-107b-video-digits2-outdoor.mp4

Capture A Single frame from the videos



- a.
- Objectives: To process video, identify ROI (Region of Interest), then
b. Remove Any Random Noise,
c. Enhance the quality of an image,
then feed it into Deep learning Neural Network.

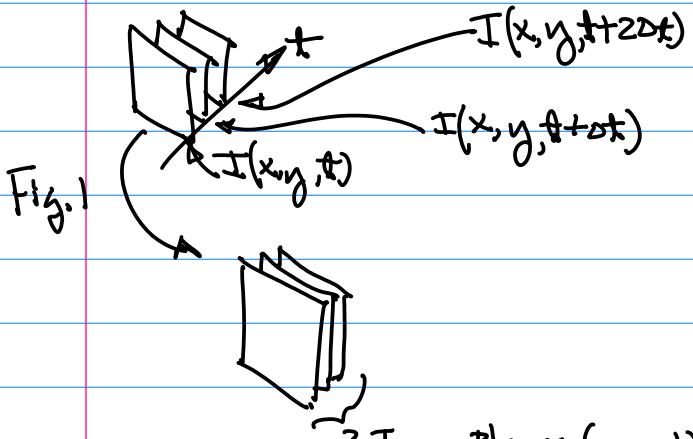
Example: Pre-processing Technique

Binary Image Processing.

Note: 1. A Color image, r, g, b

Red Green Blue

primitive colors, r, g, b : 8 bits per pixel.
24 Bits Color image



Pixel depth, Color Depth: 8 bits for Each Color plane

$$r \in [0, 255], g \in [0, 255], b \in [0, 255]$$

2. Gray Scale Image.

$I(x, y, t)$, drop the time t,
 $I(x, y)$ 8 bit gray scale image

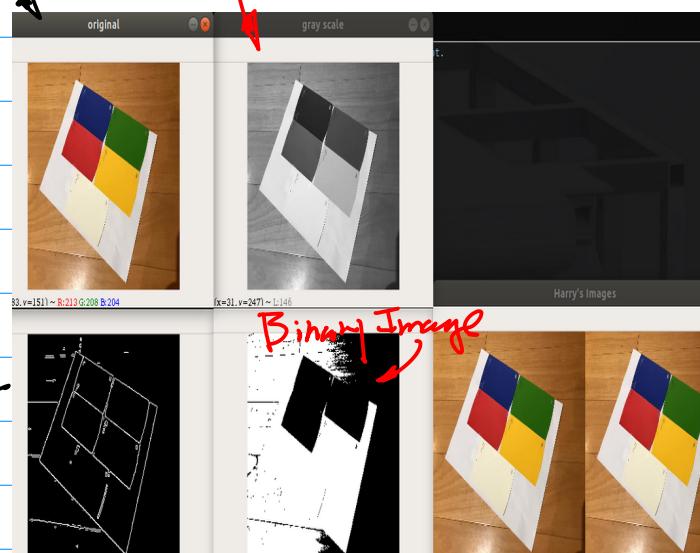
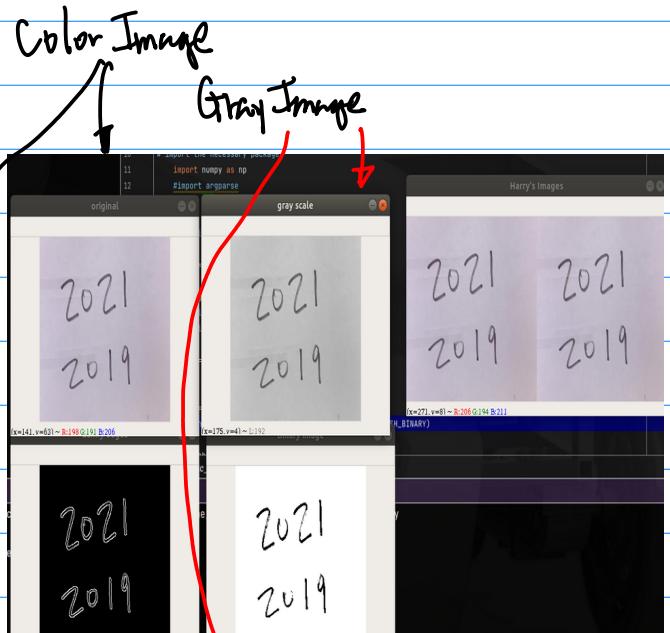
$$I(x, y) \in [0, 255],$$

$$I(x, y) = \frac{1}{3} [r(x, y) + g(x, y) + b(x, y)]$$

Provides a gray conversion.

Sometimes, Weighted Average

$$I(x, y) = \frac{1}{3} [d_1 r(x, y) + d_2 g(x, y) + d_3 b(x, y)] \quad \text{--- (1a)}$$



3. Binarize the gray scale

Image (Color Image Can
be binarized as well)

Binary image is denoted as

$$B(x, y) = \begin{cases} C_1 & \text{if } I(x, y) \geq \text{Threshold} \\ C_2 & \text{else} \end{cases}$$

if $C_1 = 255$, $C_2 = 0$

Threshold is defined by user per
his/her processing need.

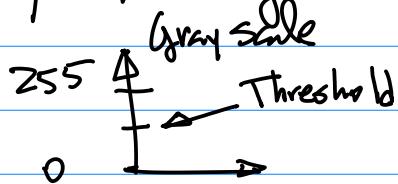
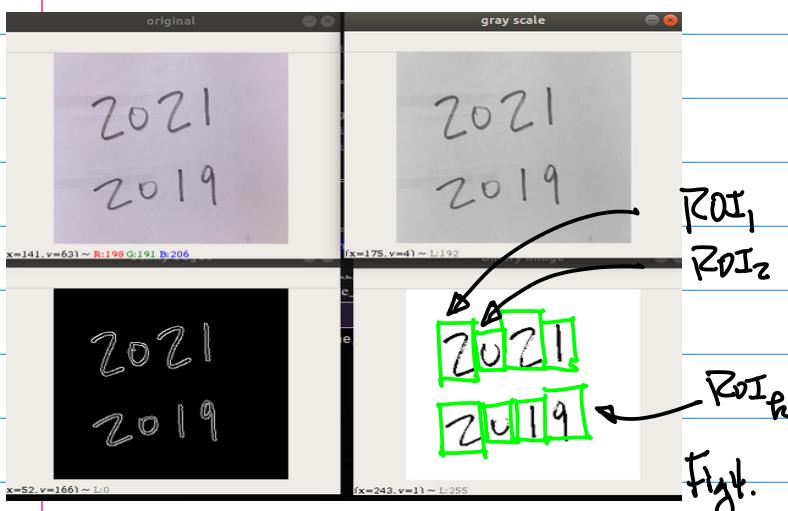


Fig.3.

Ref: For Binary Image Processing

[2022S-108-Intro-Binary.pdf](#)

Task: To Identify the Regions of
Interest (ROI)



Can be different.

Define Aspect Ratio of ROI:

Given a ROI shown in Fig.3

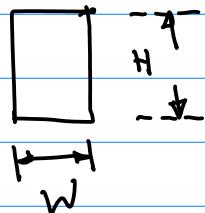


Fig.5

With Height H, Width W, the
Aspect Ratio α is defined as follows

$$\alpha = W/H \quad \dots \quad (3)$$

Suppose $W = 35$ (pixels), $H = 120$ (pixels)

$$\therefore \alpha = \frac{35}{120}$$

Note: Color image, gray scale image,
Binary image

Note: the Aspect Ratio for Each ROI

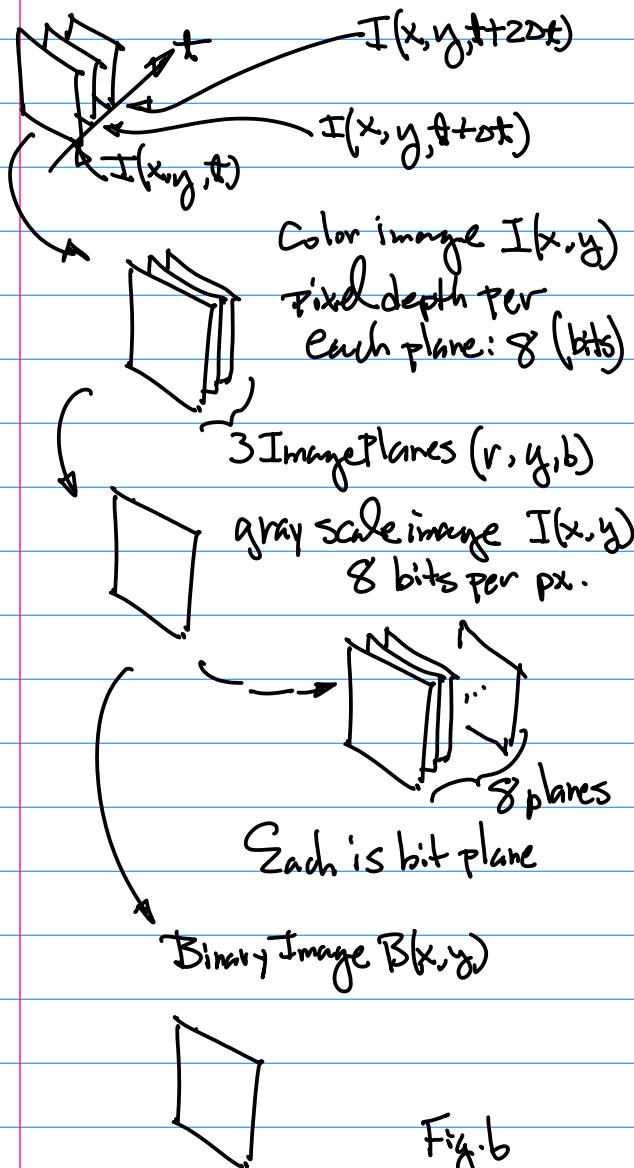


Fig.b

Question: Python code for Tensor Representation

of Each Image Types? e.g.

Color, gray scale, and Binary.

Example: Binarize a gray scale image

2022S-108c-example-binary.pdf

38

```
20     image = cv2.imread(img, cv2.IMREAD_COLOR)
24
26     image = cv2.resize(image, (256, 256))
28
29     gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
```

Note: 1. Binarization function, function's arguments

```
38     ret, thresh1 = cv2.threshold(gray, 127, 255, cv2.THRESH_BINARY)
```

Input Image, threshold, intensity Value, Binarization Keyword

1) Given a digital image $f(x, y)$ find Binary Image Based on threshold $T = 135$.

Sol

Based on $T = 135$ and

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

x	y	$f(x, y)$	
70	70	80	159
75	85	140	121
90	20	101	90
211	209	115	85

(Origin of the Coordinate System)
(0,0)

Fig.1.

K_1	K_2	y
0	0	255
0	0	255
0	255	0
255	255	0

However, for hard calculation, we use (1,1) instead.

K_1 : Index for Each Row,

K_2 : Index for Each Col.

Sample Code in OpenCV Python

2022S-107d-display-binary.py

Example: Algorithms for Binary Image Processing.

2022S-108-Intro-Binary.pdf

Note 1. Binary or Gray Scale image,

$I(x, y)$, $x = 0, 1, 2, \dots, N-1$
 $y = 0, 1, 2, \dots, M-1$

No. of Row \uparrow for Col.

$M \times N$
 \uparrow
Col. (px) per Row Row.

2. Since Digital Image is Collection of Discrete Data, So Rewrite integration-Based formula By Summation.

From Eqn(1) of the handout, we have

$$\iint_{\Omega} B(x, y) dx dy$$

Ω : An Image, like the one below

70	70	80	159
75	85	140	124
90	210	101	90
211	209	115	85

$B(x, y)$: A Binary Image, like

we have

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

$\iint_{\Omega} B(x, y) dx dy$ $\dots (1)$
Summation w.r.t all x's

$\iint_{\Omega} B(x, y) dx dy$ $\dots (2)$
Summation w.r.t all y's

Pattern Recognition For Binary Images

The tool box for pattern recognition for binary images

- 1. Size
- 2. Moments

\bar{x}
 \bar{y}
 \bar{x}^2
 \bar{y}^2 etc.

- 3. Perimeter
- 4. Orientation
- 5. Compositions of the above
Perimeter and moments: vector
- 6. Invariant operators
size invariant
orientation invariant
illumination invariant

Harry Li, PhD, SJSU

Biologically inspired techniques

- Rule 1. Proximity
- Rule 2. Similarity
- Rule 3. Closure
- Rule 4. Good continuation
- Rule 5. Symmetry
- Rule 6. Simplicity

Note: 'Proximity' usage for clean up binary image and remove noise, as well as growing boundary points per 'good continuation' rule to form a better edge map.

Note: Similarity defines an interesting question, how to describe one object is similar, or somewhat similar to others, neural network and fuzzy logic may help.

Shape Descriptor

Invariant "Shape Descriptor"

In this class, we would like to develop

Techniques, Tools Described on the left Side of the PPT.

2022S-108b-BinaryImageFormula.pdf

Example:

$$\bar{x} = \frac{\iint_{\Omega} x \times B(x, y) dx dy}{\iint_{\Omega} B(x, y) dx dy} \dots (1)$$

Where, the Area of $B(x, y)$ is defined as

$$A = \iint_{\Omega} B(x, y) dx dy \dots (2)$$

Therefore, we have

$$\sum_{y=0}^{M-1} \sum_{x=0}^{N-1} B(x, y) \approx \iint B(x, y) dx dy \quad \dots (4)$$

$$\frac{1}{255} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} B(x, y) = \text{Area}$$

Find

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

for the given Binary Image $B(x, y)$

Below,

We have

		K_2	
		0	255
K_1	0	0	0
	255	0	0
255	0	0	0

Sol. From Eqn (4), we have

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

Since image example is very small, we start index from 1, not 0.

Hence,

$$\sum_{y=1}^{M-1} \sum_{x=1}^{N-1} B(x, y) = \sum_{y=1}^4 \sum_{x=1}^4 B(x, y)$$

$$= \sum_{y=1}^4 \left(\sum_{x=1}^4 B(x, y) \right)$$

$$= \sum_{y=1}^4 (B(1, y) + B(2, y) + B(3, y) + B(4, y))$$

$$= B(1, 1) + B(2, 1) + B(3, 1) + B(4, 1) +$$

$$B(1, 2) + B(2, 2) + B(3, 2) + B(4, 2) +$$

$$B(1, 3) + B(2, 3) + B(3, 3) + B(4, 3) +$$

$$B(1, 4) + B(2, 4) + B(3, 4) + B(4, 4)$$

$$= (0+0+0+255)$$

$$0+0+255+0$$

$$0+255+0+0$$

$$255+255+0+0)$$

$$= 5 \times 255$$

After multiplying $\frac{1}{255}$, we have

$$\sum_{y} \sum_{x} B(x, y) = 5$$

Exercise 1: Form 4-Person Team,

elect a Coordinator of the Team.

to prepare for Semester Long Project

Send me email by next week with

the team information.

Exercise 2: Use Existing Code Sample

to write a program to display your Video Clips by next week.

March 1st.

Today's Topics: Preprocessing

for Digits Recognition:

2. Neural Networks (Feed Forward).

Homework Assignment Due A week from Today: March 8.

[2022S-102b-homework2-building-blocks-gradient-2022-3...](#)

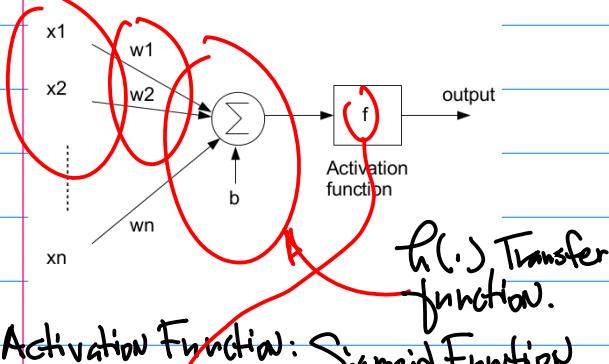
Summation, A Cumulative Loss

$$L_{total} = \frac{1}{2} \sum_{j=1}^P (\tilde{y}^j - y^j)^2$$

Ground Truth

$$y = f\left(\sum_{i=1}^N w_i x_i = W \cdot X + b\right) = f(h(w_i, b))$$

$$N=3, x_1, x_2, x_3, w_1, w_2, w_3$$

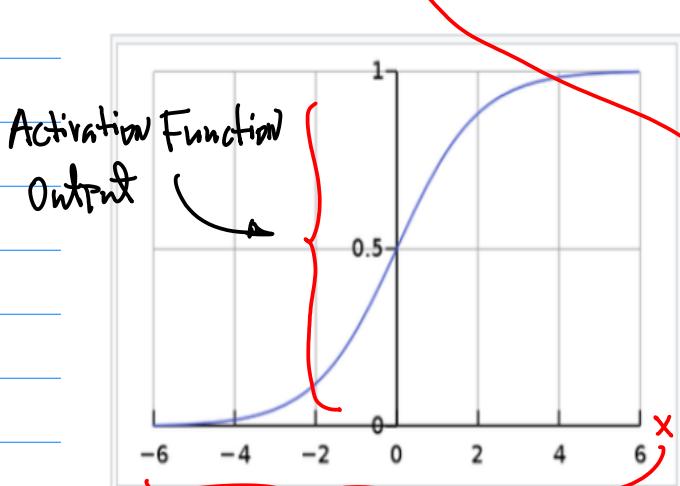


$h(\cdot)$ Transfer function.

Activation Function: Sigmoid Function

Fig 1.

$$S(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x + 1}$$



Input to the Activation function, e.g. $h(w_i, b)$'s output

(4.3) Suppose there are 3 experiments (training) are performed, with the known ground truth as $(11.3, 0.2, 1)$. Use the equation in Figure 3, find the total loss. You can use abstract output symbol such as y^1, y^2 and y^3 in your result provided you have evaluated y^1 output based on the given parameters in this assignment.

$$\tilde{y}^{(1)} = 1.75 \text{ (for Example)}$$

$$\tilde{y}^{(1)} - y^{(1)} = 1.75 - y^{(1)}$$

$$\tilde{y}^{(2)} - y^{(2)}, \dots$$

$$\frac{\partial L}{\partial w_{i,k}} = \frac{\partial}{\partial w_{i,k}} \frac{1}{2} \sum_{j=1}^P \sum_{i=1}^M (\tilde{y}_i^j - y_i^j)^2$$

M=1, Then

ith Output

$$\frac{\partial}{\partial w_{ik}} \frac{1}{2} \sum_{j=1}^P (\tilde{y}_i^j - y_i^j)^2 \Big|_{\tilde{y}_i^j = f(h(w_{ik}, b))}$$

where

$$\tilde{y}_i^j = f(h(w_{ik}, b)) = f\left(\sum_{i=1}^3 w_i x_i + b\right)$$

Therefore

$$\frac{\partial}{\partial w_{ik}} \frac{1}{2} \sum_{j=1}^P (\tilde{y}_i^j - y_i^j)^2 =$$

$$\frac{1}{2} \sum_{j=1}^P (f\left(\sum_{i=1}^3 w_i x_i + b\right) - y_i^j)$$

$$\frac{1}{L} \sum_{i=1}^L (f\left(\sum_{j=1}^3 w_j x_j + b\right) - y_i)$$

w_1, w_2, w_3 ; x_1, x_2, x_3 ;
 b .

then construct feedforward NN.

Ground Truth

Example: Collecting data for training

Name	Weight (lb)	Height (in)	Gender
Alice	133	65	F
Bob	160	72	M
Charlie	152	70	M
Diana	120	60	F

The Loss function

The weights for Each
Neuron w_i to Replace
 x_i

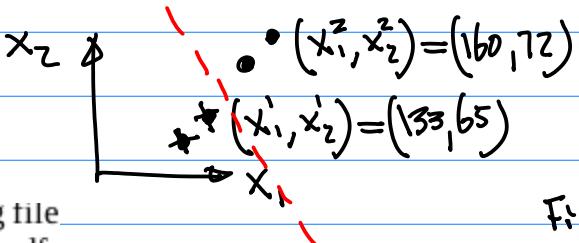
$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \dots \\ \frac{\partial f}{\partial x_i} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Fig.2

Signs	M_{ij}	M_{pq}
V1	133	65
V2	160	72
V3	152	70
V4	120	60

Note: Input vector (Training Data) (x_1, x_2)
2 Dimensional Vector.

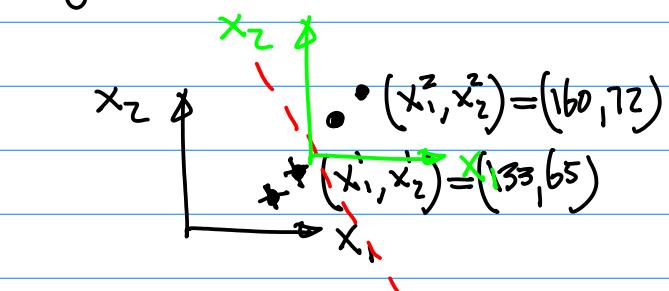
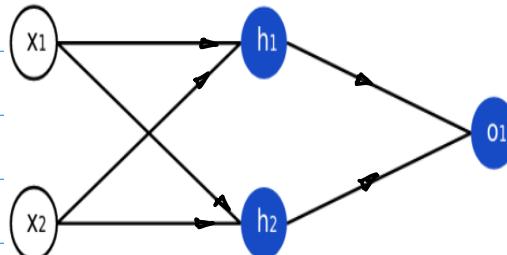
7. Submit your work in one PDF file, then zip it. Use the following file
firstName_lastName_SID(last-4-digits)_cmpe258_CondaOpenCV.pdf.



Basic Building Block
"B3"

Example: Build Feedforward Neural
Network (Multi-Layer NN)

2022S-103c-NN-Intro-Python-v5.pdf



Define "Loss function", Error function

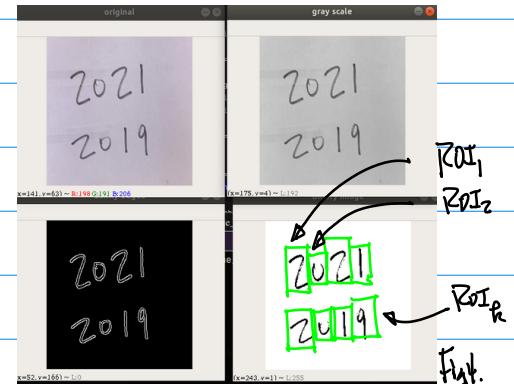
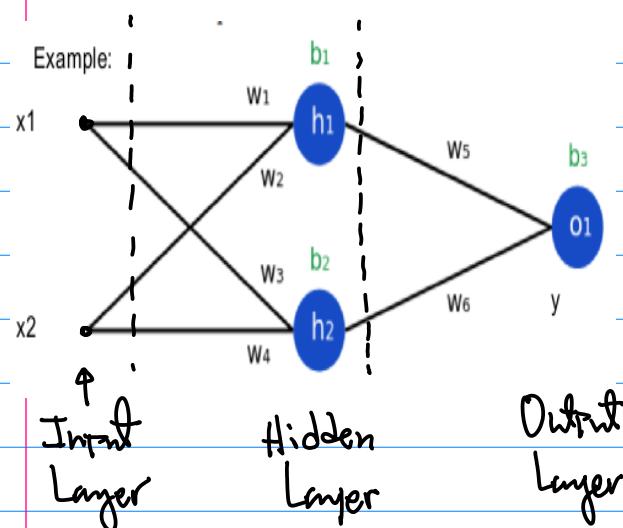
$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{true} - y_{pred})^2$$

mean

Python Code for Basic Building Blocks,

Note: The enumeration for W_i

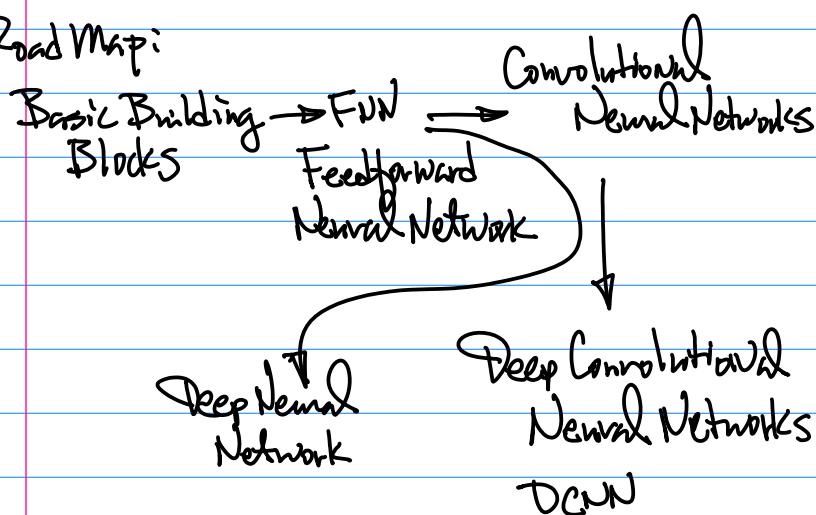
Each Handwritten Digit.



Requirements for this Python Code:

- 1° To be able to make connection Between the code and the Actual Neurons, Neural Network Architecture;
- 2° To modify the code to handle more neurons in the Hidden Layer if needed;
- 3° Run the code,

Road Map:



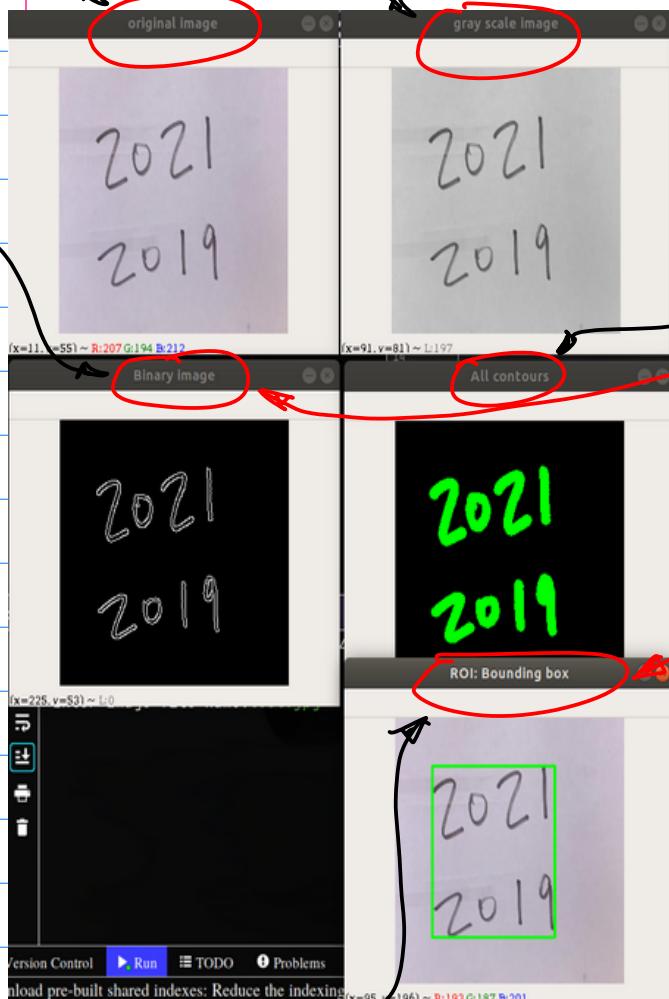
Example: Task to Develop Preprocessing Techniques(s) to define ROI. (to localize

Cmpe258 March 1st, 22

22

Note: a. Color Video \rightarrow Color Image, b.

Convert Color Image to gray scale



c. Bounding Box

Ref: my Python Code

2022S-109b-countours-roi.py

Input Image, Resize it

```

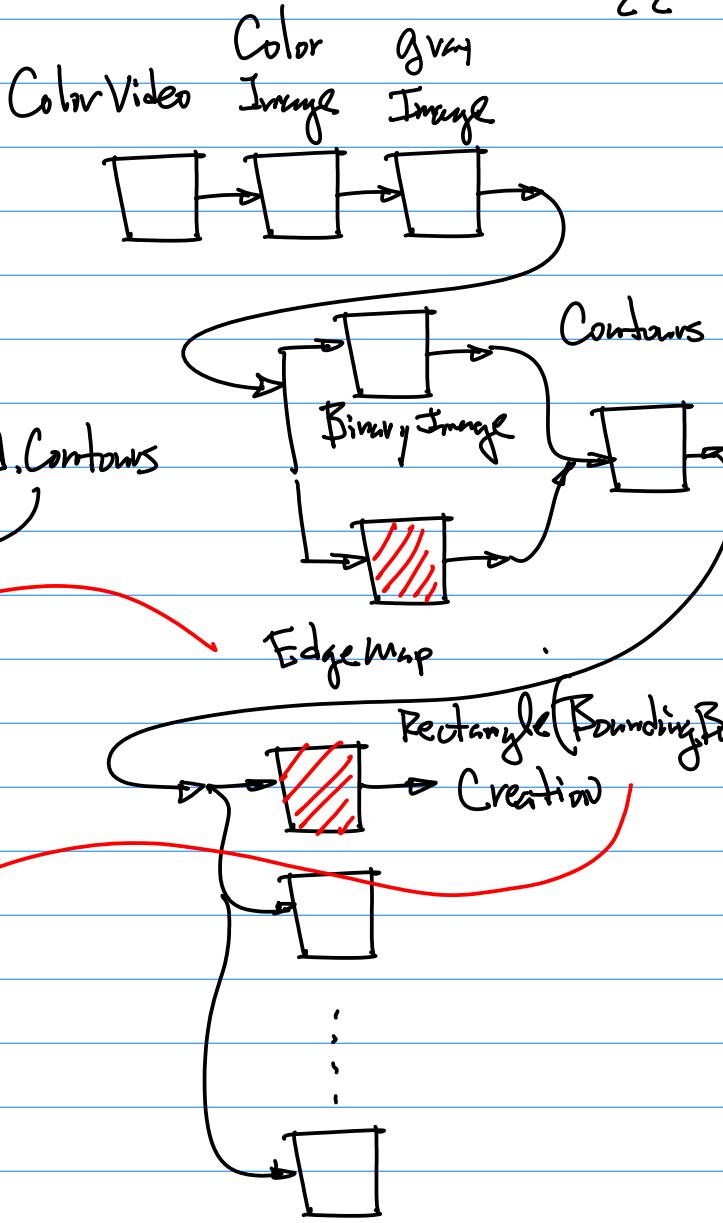
18     img = cv2.imread(image, cv2.IMREAD_COLOR)
19     img = cv2.resize(img, (256, 256))
20     cv2.imshow('original image', img)

22     imggray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
23     cv2.imshow('gray scale image', imggray)

```

Convert the image to gray scale image.

Note: Memorize the Code !



Continued, Edge Detection, Canny Detector

26

thresh = cv2.Canny(imggray, 100, 200)

Produced Edge map, which is a binary image.
 $B(x, y)$.

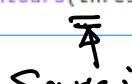
Threshold1 Threshold2

Find Contours

```
29 contours, hierarchy = cv2.findContours(thresh, cv2.RETR_TREE, cv2.CHAIN_APPROX_SIMPLE)
```



Result



Source image

Defines Contour Type

Draw color contour on color image

```
32 thresh1 = cv2.cvtColor(thresh, cv2.COLOR_GRAY2BGR)
```

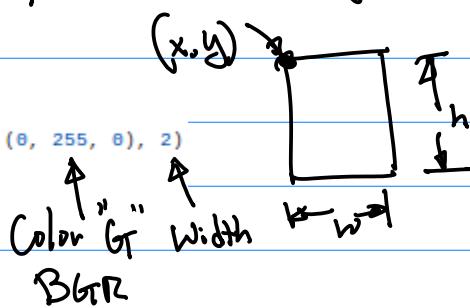
```
33 cv2.drawContours(thresh1, contours, -1, (0, 255, 0), 5) # all contours
```

Compute (x, y) Top left corner of the Bounding Box, w : width, h : height

```
37 x, y, w, h = cv2.boundingRect(thresh)
```

```
38 cv2.rectangle(img, (x, y), (x + w, y + h), (0, 255, 0), 2)
```

```
39 cv2.imshow('ROI: Bounding box', img)
```



Homework March 15, Convolution & Binary Image
midterm Exam Practice & Exam, March 15 & 22nd
Project April 4th (Monday)

March 8th (Tue)

1. Midterm Exam: March 22nd (Tue)

Format of the midterm:

- ① Zoom Based on-Line Exam;
- ② Close-Book, Close-Notes, But One-page Formula is allowed;
- ③ When entering the Exam Session, Video has to Be ON for the entire Session.

- ④ No messaging, No Smartphone use.

⑤ Your Laptop, Desktop machine is ready, Part of the Exam Questions require the execution of your code.

⑥ Have Screen Capture Software tool Ready, Capture Screen, Save as ".jpg" format or ".png" format for example.

⑦ Bring Printer paper to exam to write your Calculation, then use your phone to take photos.

⑧ Convert All photos into PDF Document using

online tool, then merge all the files into one pdf Document.

Note: Please control the resolution of your photos, keep it in 720P (1024×768 pixels) $\approx 1-2$ mByte.

(a) Mark Each page with

- Your First, Last Name,
- Cmpe258
- 4 Digits of SID .

(b) Zip your file,

Name your file:

First-Name+Last-Name+4Digits+Cmpe258-mid.zip

Submission to SJSL Canvas.

Ref: from the Class github



Topics: 1° Convolution Techniques

2° MNIST Neural Network
Convolutional

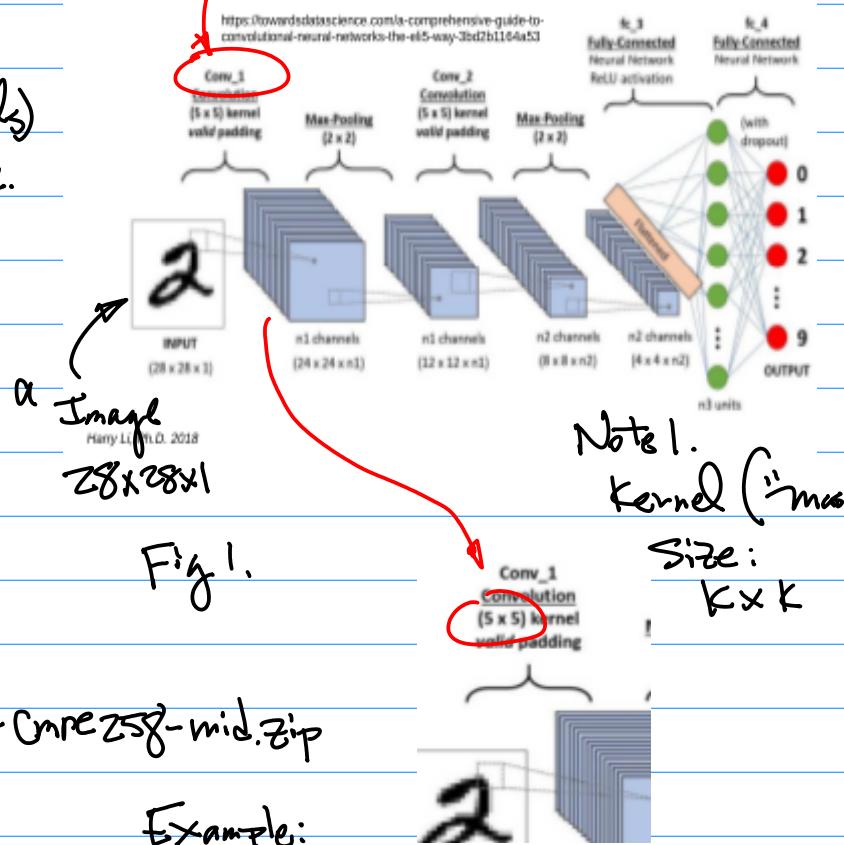
Project Assignment: For Student

ID Recognition (Handwritten ID)
April 4th (Monday).

Example: Convolution

b Convolution Layers

Illustration of A CNN for Digits Recognition



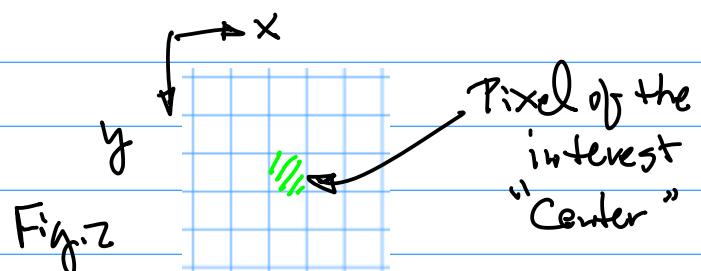
Note 1.
Kernel ("mask")

Size:
 $K \times K$

Example:

K has to be an "odd" Number.

1, 3, 5, ...



1° Kernel size: $K \times K = 5 \times 5$

Center of the kernel, Equal No. of Rows Above & Below the center, Equal No. of Col. Left & Right from the center $\rightarrow K = \text{"Odd Number"}$

"Un-Biased" Kernel

2' Kernel v.s. given image for Convolution.

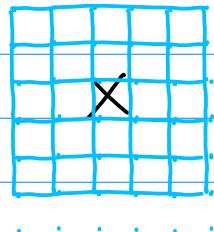
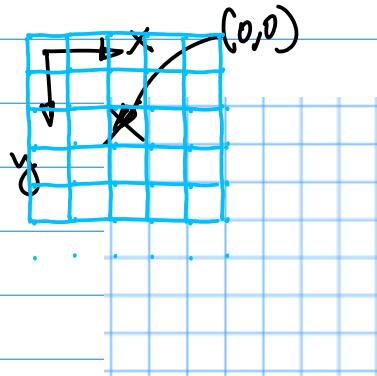
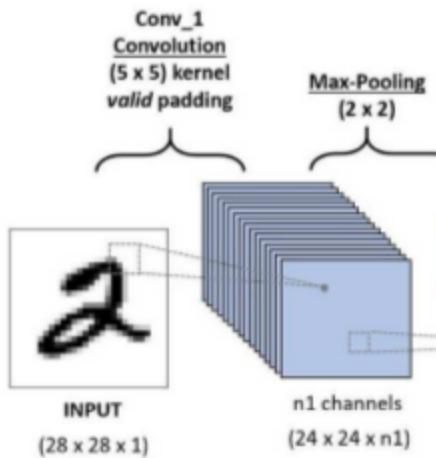


Fig. 3.c



Stage i Stage i+1
→ Convolution
with 5x5 Kernel

Fig. 3.a

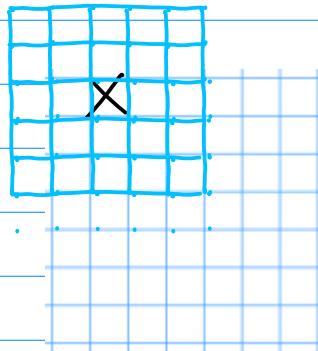
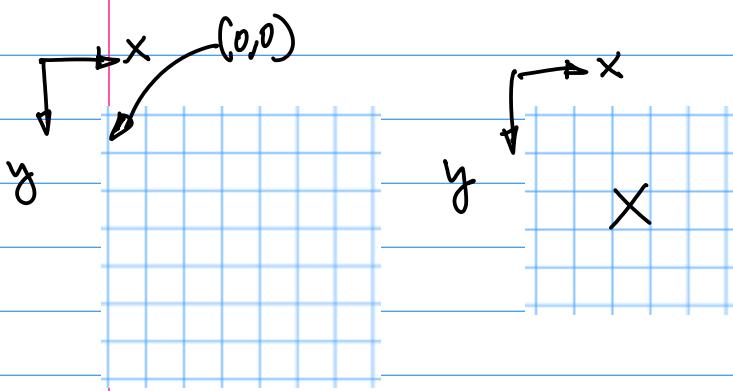


Fig. 3.d



$I(x, y)$
Image

7x7

Fig. 3.b

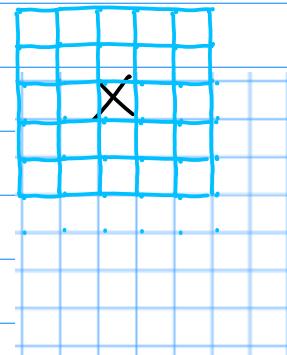


Fig. 3.e

Kernel movement : Starting at the top left hand corner, then $L \rightarrow R$ (Left to Right), $T \rightarrow B$ (Top to Bottom)
One pixel at a time.

Note: The Basic Operations for Convolution, e.g., the movement of a kernel

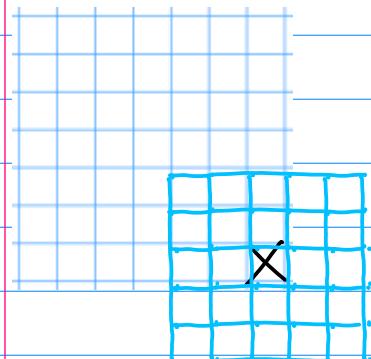


Fig. 3f

$$\iint_{\Omega} I(x-u, y-v) k(u, v) du dv \quad \dots (1)$$

Image Domain:

m x n

Resolution

Kernel, Size
K x K

3. Mathematical Operations

Note: Each kernel ($K \times K$) has to have kernel/pixel coefficients

Example: for $K \times K$ Kernel,

$K=5$, we have $5 \times 5 = 25$

Kernel elements, pixels, each Kernel element, pixel, has its value for processing, feature extraction purpose.

$$\sum_{v=0}^{N-1} \sum_{u=0}^{M-1} I(x-u, y-v) k(u, v) \quad \dots (2)$$

where $m \times N$: Resolution of a given image.

$$\sum_{v=0}^{N-1} \sum_{u=0}^{M-1} I(x-u, y-v) k(u, v) =$$

$$\sum_{v=0}^{N-1} \sum_{u=0}^{M-1} I(u, v) k(x-u, y-v)$$

... (3)

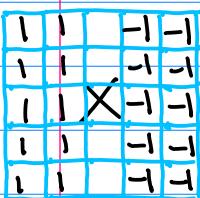


Fig. 4a.

1	1	-1	-1					
1	1	-1	-1					
X	-1	-1						
1	1	-1	-1					
1	1	-1	-1					

Eqn(3):

 $k(x-u, y-v)$ v.s. $k(u, v)$ a. $k(u, v)$, b. $k(-u, -v)$, c. $k(1-u, 1-v)$

(Edge Component Detection, for a Vertical Boundaries of Objects)

Mathematical Definition of 2D Convolution

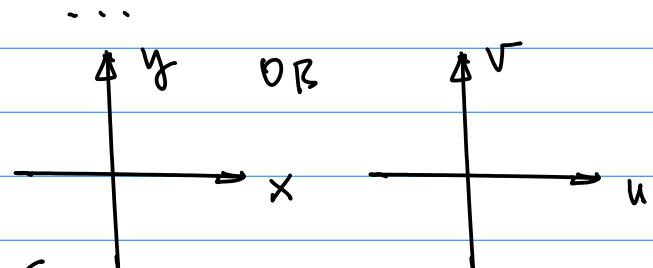


Fig. 5a

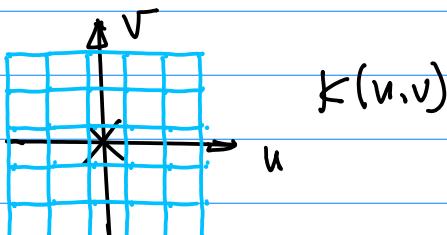


Fig 5b From a.
place the kernel on this
 $u-v$ Coordinate System

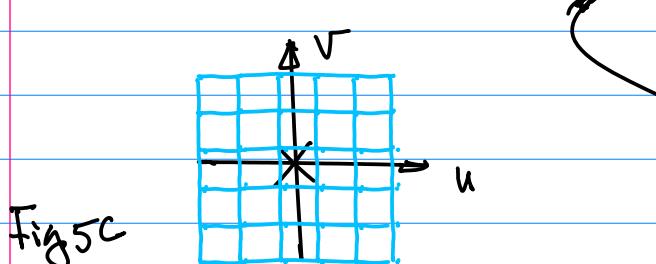
Note:

$$\underbrace{I(u,v) K(x-u, y-v)}$$

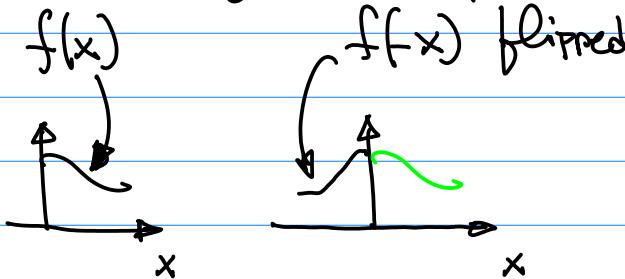
is Product

Note: From Eqn (3),

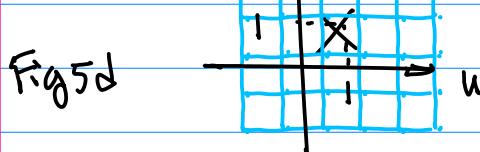
$$\sum_{j=0}^{N-1} \sum_{u=0}^{m-1} \text{Summation of} \\ \text{Each Prodnt.}$$



Flip $K(u,v)$ Twice, once along u , once along v . Example:



Note: Kernel is designed with coefficient to make the "flip". $K(-u, -v)$



We can Shift Kernel $K(x-u, y-v)$ by (x, y) , follow L \rightarrow R, T \rightarrow B one pixel at Time starting at top left corner of the image.

Summary: To carry out 2D convolution in Eqn (3), we perform Operations Characterized as "3-step" operation, e.g.: Shift-multiplication-Summation

Example: Given 4x4 Image Below

10	10	00	00
10	10	00	00
10	10	00	00
10	10	00	00

Fig 6a

and a kernel with
 $K=3$

1	-1	
1	X	-1
1	-1	

Fig 6b

Perform 2D convolution

Sol:

Step 1. Take the kernel, place it at the starting location on the given image

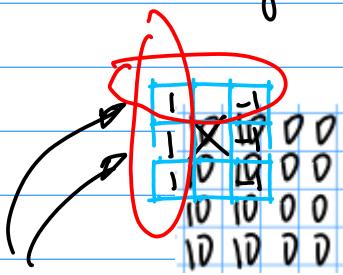


Fig. 7a

have no corresponding image pixels

Note: Always Align the Kernel such that none of its row or col have no corresponding image pixel.

Convolution Result

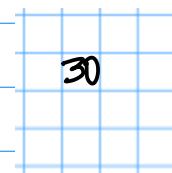


Fig. 7b

Carry out multiplications in "L → R, T → B" fashion

$$\text{1st Row, 1st col. } I(x, y) = \sum(0, 0) = 10$$

Kernel coefficient = 1

$$\text{Product } 1 \times 10 = 10$$

For the rest of the kernel coefficient

$$1 \times 10 + 0 \times 10 + (-1) \times 0 = 10$$

→ 2nd Row of the Kernel
 $1 \times 10 + 0 \times 10 + (-1) \times 0 = 10$
 3rd Row of the Kernel
 $1 \times 10 + 0 \times 10 + (-1) \times 0 = 10$

Add 1st, 2nd, 3rd All together

$$10 + 10 + 10 = 30$$

Continue the Calculation.

Step 2. place Kernel to 1 pixel to its Right

1st Row:

10	10	0	0
10	10	X	0
10	10	0	0
10	10	0	0

10	10	0	0
10	10	0	0
10	10	0	0
10	10	0	0

1	-1
1	X
1	-1

1st Row, kernel; $1 \times 10 + 0 \times 0 + (-1) \times 0 = 10$

2nd : $1 \times 10 + 0 \times 0 + (-1) \times 0 = 10$

3rd : $1 \times 10 + 0 \times 0 + (-1) \times 0 = 10$

Convolution @ this fixed location

$$= 10 + 10 + 10 = 30$$

Convolution Result

30	30
30	30

Finally, to finish the Convolution:

Convolution Result

30	30
30	30

Dimension (Size) of the convolution

Result is smaller due to "Boundary Effect". following the formula below:

$\frac{k-1}{2}$ Size of the Symmetric portion (upper or lower; left or right)

Then, $2 \times \frac{k-1}{2} = k-1$ is the dimension for the Rows and col. to Be Reduced from its original size of the image.

Example: $k=3$, $k-1=2$

Original Image 4×4 is to Be Reduced to $(4-2) \times (4-2) = 2 \times 2$,

Homework Assignment Due Next week

March 14. Check github / Canvas to Be posted on Line today.

March 15 (Tue)

Topics : 1^o Convolutional Neural Network — MNIST ; 2^o

Project for Handwritten SID Recognition
(Due April 11th — Monday, 11:59pm.)

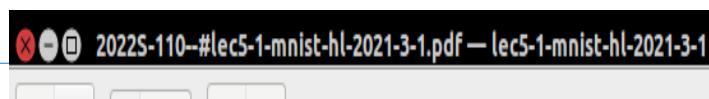
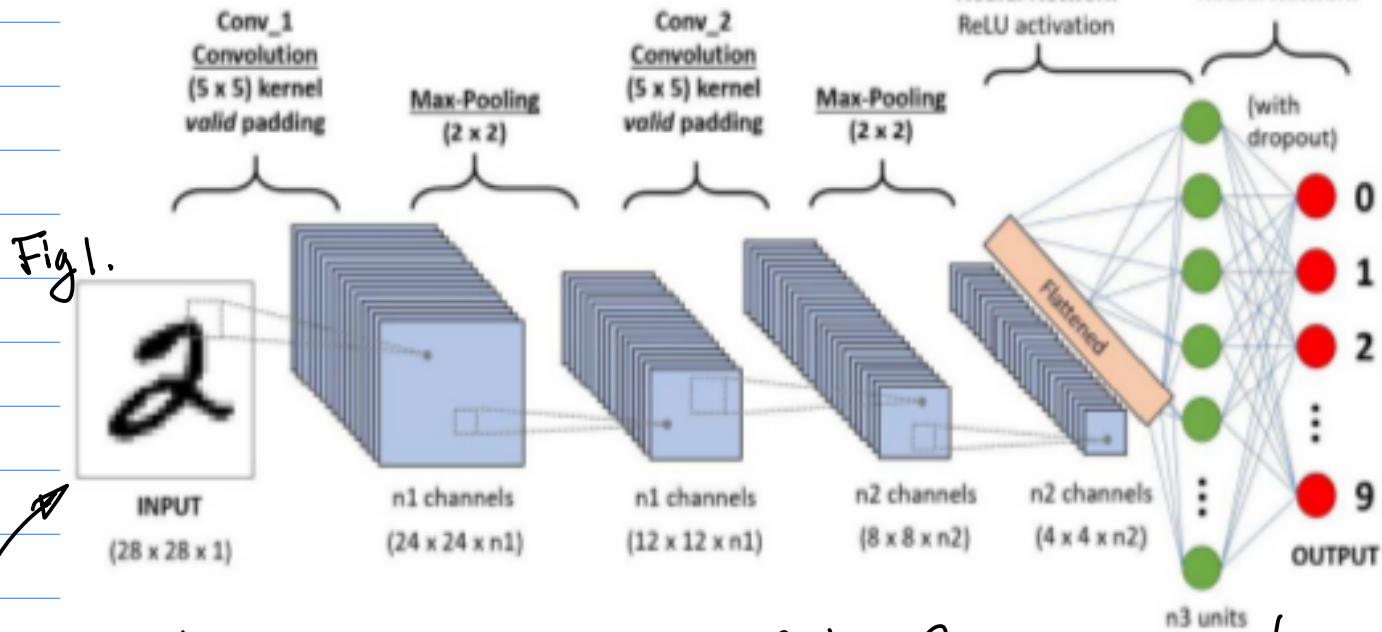


Illustration of A CNN for Digits Recognition

Reference Only
See Sample Code

<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>



a. Preprocessing
Result has to be resize w/o Altering the Aspect

b. 1st 2D Conv. Denoted it as C_1 , 32 kernels/convolutions

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 26, 26, 32)	320

c. MaxPooling

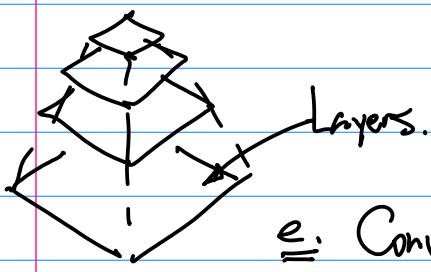
1. 2×2 kernel size
Denoted M_1

2. Find max
Use it as an Output

Input Image Size 28×28 → $K=3$
Output Convolution Result 26×26 → Kernel
 $\max\{15, 34.1, 2, 0.1\} = 34.1$

3. Verify the Dimension of Output Layer $\left(\frac{26}{2} = 13 \text{ for } x, y \text{ dimension}\right)$

max_pooling2d_1 (MaxPooling2 (None, 13, 13, 32))	0
conv2d_2 (Conv2D) (None, 11, 11, 64)	18496



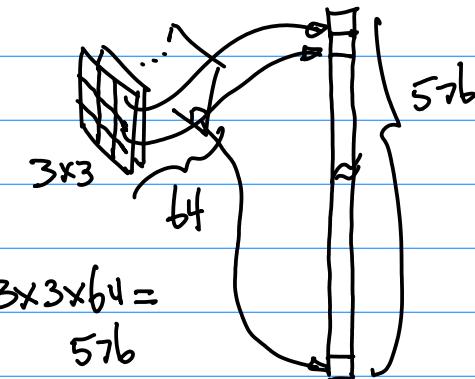
d. Convolution Layer + MaxPooling : $C_2 M_2$ (1st: $C_1 M_1$)

conv2d_2 (Conv2D)	(None, 11, 11, 64)	18496
max_pooling2d_2 (MaxPooling2 (None, 5, 5, 64))	0	

Note: C₃ (3rd Convolution Layer) No. of Kernels / Convolutions

conv2d_3 (Conv2D)	(None, 3, 3, 64)	36928
-------------------	------------------	-------

conv2d_3 (Conv2D)	(None, 3, 3, 64)
flatten_1 (Flatten)	(None, 576)



Note: Dense Layer (For Feed-forward NN)
D₁ and D₂

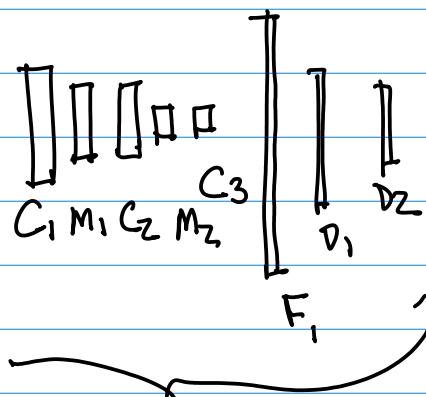
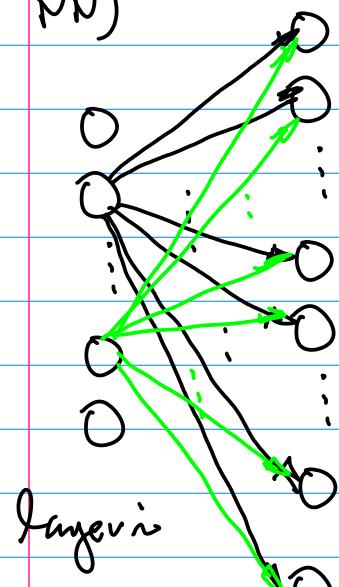


Fig. 4

Layer i+1

Output, Define Activation function, to allow
(1) Output (id of them) to match

Digits 0, 1, ..., 9;

(2) "Probability" sort of, for each

Output, the highest Value will

be selected for the corresponding

Digit. Hence, we have,

$$f(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

... (1)

Where

$$\sum_{j=1}^K e^{z_j} = e^{z_1} + e^{z_2} + \dots + e^{z_{10}}$$

10 Digits

Therefore

$$0 < \frac{e^{z_i}}{e^{z_1} + e^{z_2} + \dots + e^{z_{10}}} < 1$$

in addition,

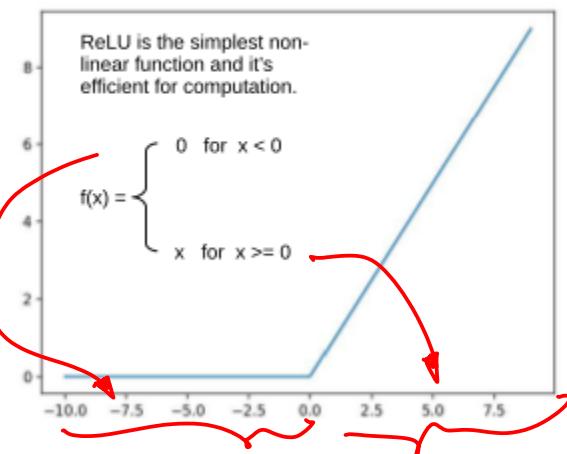
$$\frac{e^{z_1}}{e^{z_1} + e^{z_2} + \dots + e^{z_{10}}} + \frac{e^{z_2}}{e^{z_1} + e^{z_2} + \dots + e^{z_{10}}} + \dots + \frac{e^{z_{10}}}{e^{z_1} + e^{z_2} + \dots + e^{z_{10}}}$$

$$= \frac{e^{z_1} + e^{z_2} + \dots + e^{z_{10}}}{e^{z_1} + e^{z_2} + \dots + e^{z_{10}}} = 1. \quad \underline{\text{Softmax function}}$$

Activation Function: ReLU (Non-Linear)

Rectified Linear Activation Function

<https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/>



Code Walk-Through

```
#-----save trained model
import h5py
model.save('harryTest.h5')
#-end
```

Midterm Review:

1. Scope of the Exam: Up to Today's Lecture Material;

2. 3rd Questions

(1) Basic Concepts, Basic Building Blocks, Input (Vector X), Weight (Vector w)

Transfer function, $f(\cdot)$, $f(\sum x_i w_i + b)$, Activation Function f , or $f(f(\cdot))$ or $f(\sum x_i w_i + b)$

Softmax,

ReLU,

Sigmoid ~

Building/Design Feedforward NN
with Basic Building Blocks.

Block Diagram "Toy Example"

NIST

Loss function Definition.

$$\text{MSE} \quad \frac{1}{N} \sum_{j=1}^N (g_j - f_j)^2$$

$$\frac{1}{M} \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^M (g_{ij} - f_{ij})^2$$

gradient $\nabla L(w_{ij})$

(2) Practical Implementation
(Coding)

Laptop/Desktop Ready to
Execute your program.

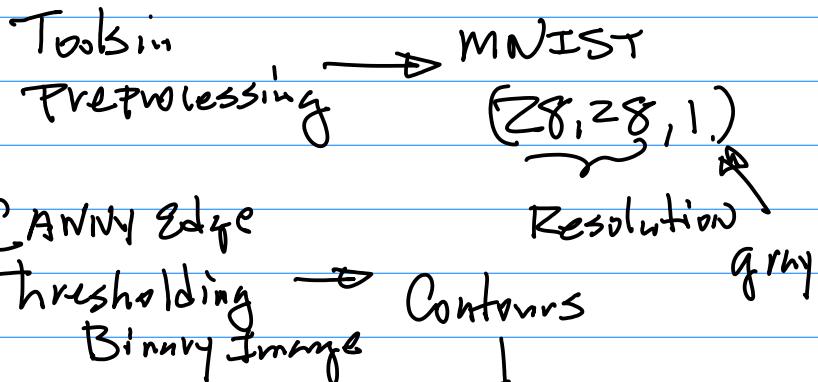
(3) Calculation, Convolution,

Binary Image Processing.

Area, \bar{x} , \bar{y} , higher order terms

$$\bar{x} = \frac{\iint x B(x,y) dx dy}{\iint B(x,y) dx dy}$$

$$\frac{\iint (x - \bar{x})^2 B(x,y) dx dy}{\iint B(x,y) dx dy}, \text{ etc.}$$

Topics: Visit Preprocessing
Technique.

3. One-Hour Exam plus 15 min.
to prepare file uploading.

4. Close Book/Close Notes,
One page Formula is
allowed, just formula, No
Writing or Verb Description
on Formula sheet.

5. Video has to be on during
the entire session.

b. After Exam, 4:45 ~ 4:50
Back to class.

March 22nd (Tue.)

Contour : TREE

| v_prev).

same outer contours. It sets $y[i][3]=-1$ for all the contours.

CV_RETR_EXTERNAL

first = c0

CV_RETR_TREE

first = c0

h00 ✓ h01 ✓

s000 ✓ c010 ✓

✓ h0000 ✓ h0100 ✓

✓ c01000 ✓ c01001 ✓

1st 2nd

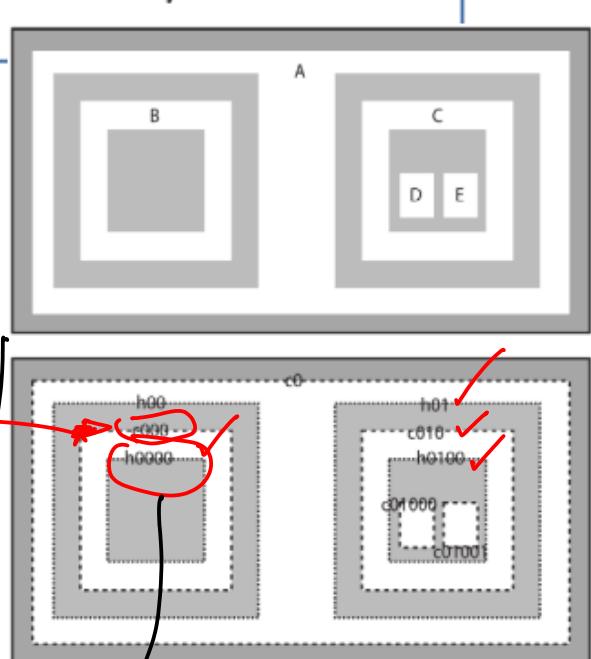
CV_RETR_CCOMP

first = c01000-->c01001-->c010-->c000-->c0
h0100 h0000 h01-->h00

CV_RETR_LIST

first = c01000-->c01001-->h0100-->c010-->c000-->h01-->h00-->c0

retrieves all contours without any hierarchical relationships,



<http://opencvexamples.blogspot.com/2013/09/find-contour.html>

Fig.1

Step1.

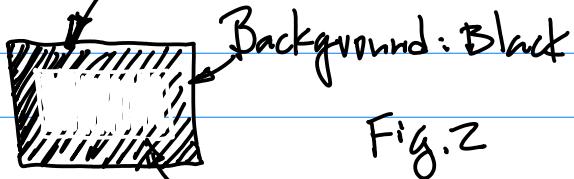


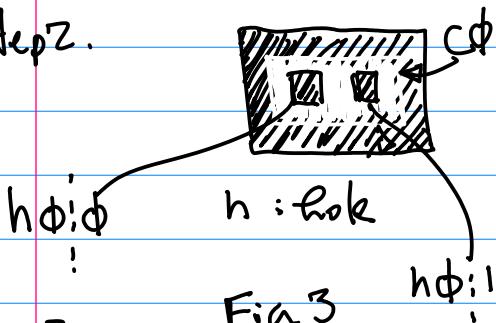
Fig.2

c ϕ

Object: White

Contour No. Enumeration of the Contour

Step2.



h ϕ : ϕ

h : hole

h ϕ :!

Step3.

Fig.3

See Fig.1. Place One Object inside

H ϕ : ϕ

its Parent Contour Number

1st Object

Note: To understand

the tree structure, to be able to generate Contour Tree from a given Pattern, or from a contour tree to illustrate Contour Objects.

h + No. (1st part - Reference to the No. of the Contour it belongs)

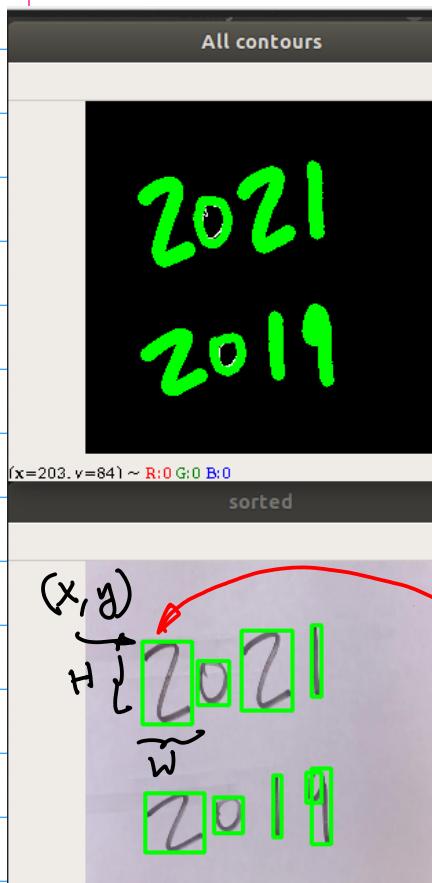
2nd Part: Enumeration of the Contour)

Step4. Continue this process,

we have placed a hole, hence

H ϕ : ϕ

Localization of ROI



```

main.py ^ _pyaev_executive.py ^
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

#img = cv2.imread('test.jpg')
image = input('Enter image file name:')

img = cv2.imread(image, cv2.IMREAD_COLOR)
img = cv2.resize(img, (256, 256))
cv2.imshow('original image', img)

imggray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
cv2.imshow('gray scale image', imggray)

thresh = cv2.Canny(imggray, 100, 200)
cv2.imshow('Canny', thresh)

contours, hierarchy = cv2.findContours(thresh, cv2.RETR_EXTERNAL,
# Draw contour
thresh1 = cv2.cvtColor(thresh, cv2.COLOR_GRAY2BGR)
cv2.drawContours(thresh1, contours, -1, (0, 255, 0), 5) # all cor
cv2.imshow('All contours', thresh1)
#Bounding box

for i in range(len(contours)):
    [x, y, w, h] = cv2.boundingRect(contours[i])
    print("i:", i)
    cv2.rectangle(img, (x, y), (x + w, y + h), (0, 255, 0), 2)
    cv2.imshow('sorted', img)

print(img.shape)

```

April 5th (Tue)

Topics:

1^o Project (10 pts) on Handwritten Digits April 12 Recognition. Due A week from Today

2^o Written Requirements+ Rubrics will be posted on CANVAS.

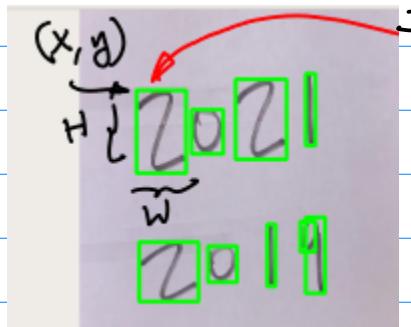
3^o Requirements are given as follows:

a. Input Video (Either Live Video or Video File), Prompt the

User for Input Video Source Selection.

b. Localized ROI has been achieved

Convert to Sgnal Image



Note:

Features/ $x_i, i=1, 2, 3, \dots, K$

Example for Bond Box Size, $i=1$

Fig. 1

X_i : A Class of Features for the Size of Bounding Boxes.

Too Small \rightarrow Noise, filter out
Too Big \rightarrow Unlikely the Right ROI.

$$\begin{matrix} X_1 & X_2 & X_3 & X_4 & f \\ \checkmark & \checkmark & & & \end{matrix}$$

$$f = X_1 X_3$$

Note: Use OpenCV Function to Construct A Square Image.

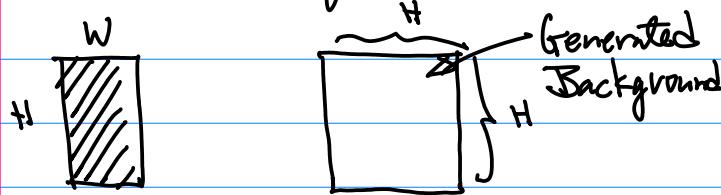
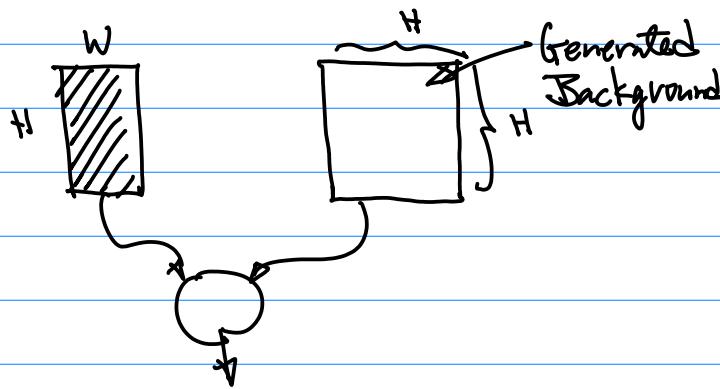


Fig 2.

as dimension for Square Image.



Then, Reduce / Resize the Image to the right size, 28x28

C. Ref: Reshape the image Format

20-2021S-3-load-deployment.py

```

1 from keras.models import load_model
2 import cv2
3 import numpy as np
4 from PIL import Image
5
6 model = load_model('mnist.h5')
7 model.compile(optimizer='rmsprop',
8                 loss='categorical_crossentropy',
9                 metrics=['accuracy'])
10
11 img = Image.open('1.jpg').convert('L')
12 img = np.resize(img, (28, 28, 1))
13 im2arr = np.array(img)
14 im2arr = im2arr.reshape(1, 28*28)
15 y_pred = model.predict_classes(im2arr)
16 print(y_pred)

```

d. Save Detection Result (Video Clips, 5~10 SEC), Be Sure to use your personal SID (Just Last 4 Digits)

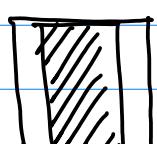
Topic on the 2nd part of this Class Yolo (You Only Look Once)

Source Code distribution for Yolo Reference Paper on Yolo

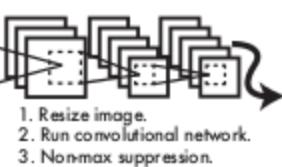
2022S-112-yolo-paper.pdf

Introduction to Yolo Algorithm.

Example:



Note: a. Input Image: Square Image, Resolution of the image: 448×448



64×64
($\frac{1}{4}$)

$I(x,y)$

128×128 ($\frac{1}{4}$)

512×512
 $\approx 448 \times 448$

32×32
 $\approx (28 \times 28)$

For Handwritten Digits.

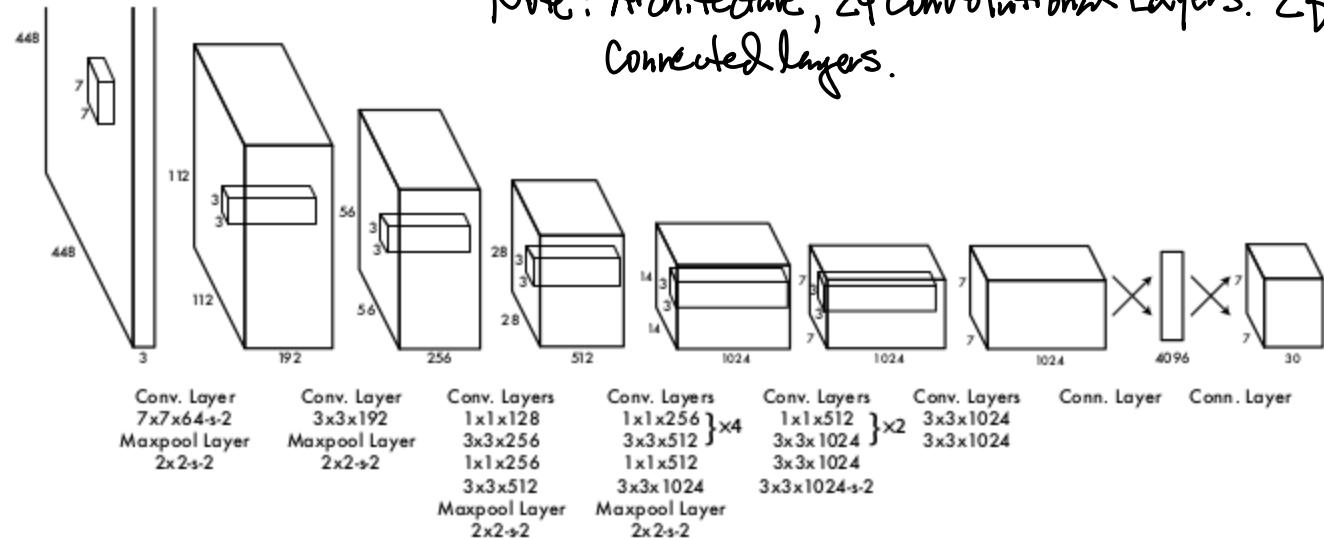


Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

Source Distribution, see Handout to be posted online (github).

b. anaconda to configure the environment

1. README Document.

a. URL for Yolo Source distribution Download

c. Activate the Conda environment

d. 3 Samples to test

First Example, to perform Training

Save the training Result;

2nd works on input image
(Single Frame).

3rd program works on input
video file.

Homework: Due 2 weeks from today.
(April 19th).

1^o Download, Install yolo program.
(yolov4).

2^o Use cellphone to record 5-10
Seconds of video clip for testing
Purpose.

3^o Run Yolo Code to perform default
Detection Task

4^o Submission on CANVAS.

- a. Screen Capture of Yolo program
- b. Processed Video Clip (Be Sure
to use your own Video Clip).

Example: Loss Function(Brief)

Note:

loss function:

$$\begin{aligned}
 & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\
 & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} \left[(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right] \\
 & + \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \\
 & + \lambda_{\text{nobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{nobj}} (C_i - \hat{C}_i)^2 \\
 & + \sum_{i=0}^{S^2} \mathbf{1}_i^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \quad (3)
 \end{aligned}$$

Example: On Probability Classification.

$$\Pr(\text{Class}_i | \text{Object}) * \Pr(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}} = \Pr(\text{Class}_i) * \text{IOU}_{\text{pred}}^{\text{truth}} \quad (1)$$

Definition of Conditional probability.

Example: On Partition of a Test Image

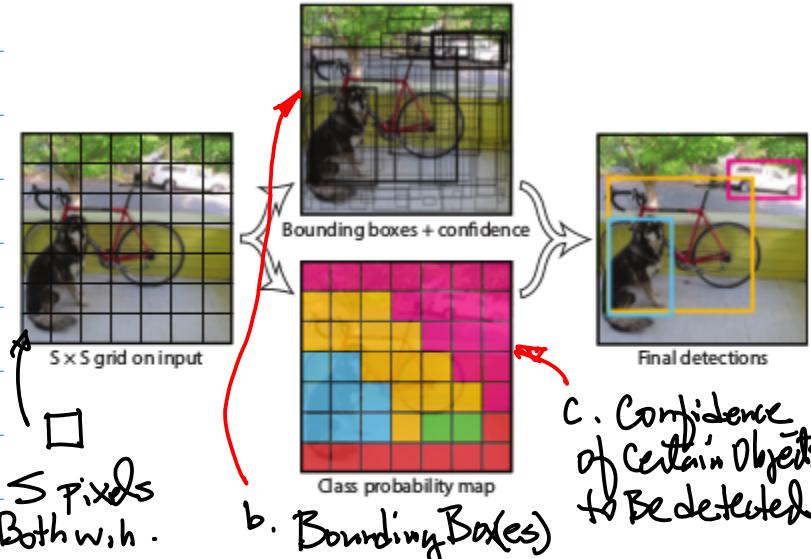


Figure 2: The Model. Our system models detection as a regression problem. It divides the image into an $S \times S$ grid and for each grid cell predicts B bounding boxes, confidence for those boxes, and C class probabilities. These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor.

d. The Classification of Object
Detected is Based on its Confidence Level.

April 12, 22

Topics: 1° YOLO Formulation
 2° Back-propagation (BackProp) Algorithm.

2022S-112-yolo-paper.pdf

You Only Look Once:
 Unified, Real-Time Object Detection

Joseph Redmon*, Santosh Divvala†, Ross Girshick†, Ali Farhadi†

University of Washington*, Allen Institute for AI†, Facebook AI Research†

<http://pjreddie.com/yolo/>

$$\Pr(\text{Class}_i \mid \text{Object}) * \Pr(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}} = \Pr(\text{Class}_i) * \text{IOU}_{\text{pred}}^{\text{truth}}$$

Theoretical Analysis of Yob.

Simple Example As A Starting Point of Our Discussion, from Eqn 1)

Note: 1. $\Pr(\text{Class}_i)$

Probability of a given Object(s) which belongs to a class i

 $i=1, 2, \dots, K$ Let $i=1$ for Simplicity purpose

$$\Pr(\text{Class}_i) = \Pr(\text{Class})$$

2. IOU (Intersection of Union)

 η_{IOU} Coefficient

Simplify

$$\Pr(\text{Class}) \eta_{\text{IOU}} \rightarrow \Pr(\text{Class})$$

Prob(C), Denote class as "C".

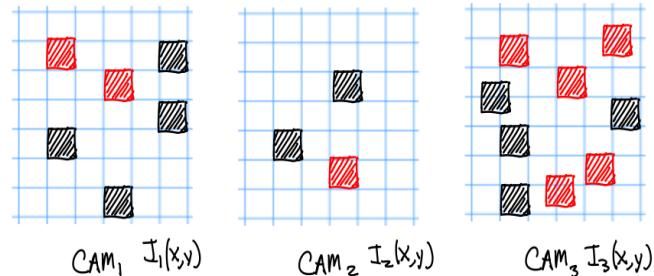
Example: Consider Object Identification,
 In particular, Detect Pedestrian.

Ref:



2022S-114-yolo-designGuide-2022-4-9.pdf

Red, \mathcal{I} : Pedestrians;
 Black, \mathcal{B} : Vehicle;

1. Define An Event R

Detection of Pedestrian.

Fig 1.

Set Relations.

$$R = R\mathcal{I}_1 + R\mathcal{I}_2 + R\mathcal{I}_3 \dots \quad (1)$$

$$R \cap R_1, \mathcal{I}_1 \cap \mathcal{I}_2 \cap \mathcal{I}_3 = \emptyset \quad \text{Empty}$$

2. Formulate the Likelihood of the event R

$$\Pr(R) = \Pr(R\mathcal{I}_1) + \Pr(R\mathcal{I}_2) + \Pr(R\mathcal{I}_3) \dots \quad (2)$$

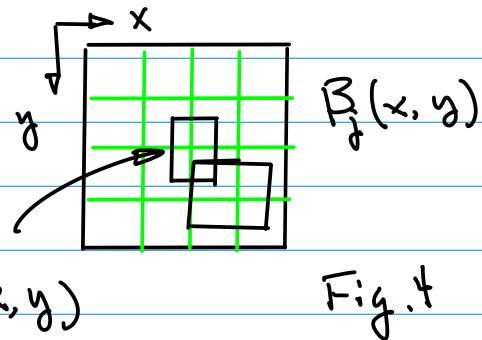
3. Conditional Probability, Bayesian Theory

$$\Pr(R\mathcal{I}_1) = \Pr(R|\mathcal{I}_1) \Pr(\mathcal{I}_1) \dots \quad (3)$$

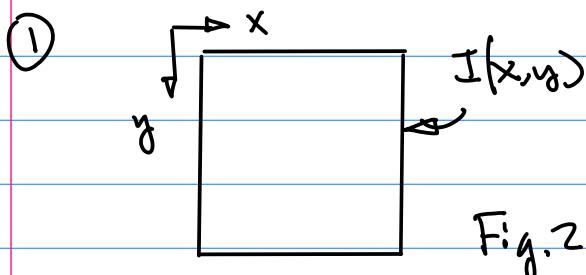
$$\Pr(R\mathcal{I}_2) = \Pr(R|\mathcal{I}_2) \Pr(\mathcal{I}_2)$$

$$\Pr(R\mathcal{I}_3) = \Pr(R|\mathcal{I}_3) \Pr(\mathcal{I}_3)$$

$$\begin{aligned}
 4. \text{ Prob}(R) &= \text{Prob}(R|I_1)\text{Prob}(I_1) \\
 &\quad + \text{Prob}(R|I_2)\text{Prob}(I_2) + \text{Prob}(R|I_3)\text{Prob}(I_3) \\
 &= \sum_{i=1}^3 \text{Prob}(R|I_{i,i})\text{Prob}(I_{i,i}) \\
 &\quad \cdots (4)
 \end{aligned}$$

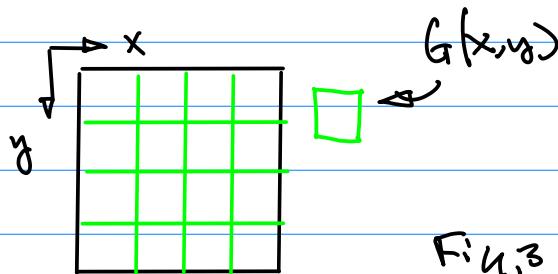


5. Notations for Yolo



(2) Define Grids G(x,y)

$S \times S$
S pixel width, S pixel height



We have multiple grids, denoted as

$G_p(x,y), p=1, 2, \dots \dots (5)$

(3) Define Bounding Box for ROI.
as $B(x,y)$

Note:

$B_i(x,y; w, h; f)$, for $i=1, 2, \dots, M$
 $\overline{\overline{w}}$ $\overline{\overline{h}}$ \overline{f}
 Location Shape Confidence No. of B. Boxes

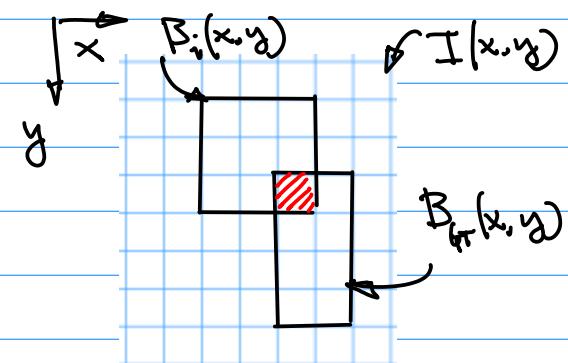
$B_j(x,y; w_j, h_j; f_j) \dots (6)$

Note: Among them, we have one ground Truth Bounding Box

$B_{GT}(x,y; w, h, f)$, $f=1$ for the ground truth.

(4) IoU Definition.

Example: Illustration of IoU



Intersection of Union

$$\text{IoU} = \frac{\text{Intersection}}{\text{Union}} \dots (7)$$

$$\text{IOU} = \frac{\text{Intersection}}{\text{Union}}$$

$$= \frac{B_i(x,y) B_{gr}(x,y)}{B_i(x,y) + B_{gr}(x,y) - B_i(x,y) B_{gr}(x,y)}$$

$$= \frac{1}{9+8-1} = \frac{1}{16} \quad "$$

Physical meaning of IOU.

If $B_i(x,y)$ & $B_{gr}(x,y)$ Are Identical
then $\text{IOU} = \frac{B_i(x,y) B_{gr}(x,y)}{B_i(x,y) + B_{gr}(x,y) - B_i(x,y) B_{gr}(x,y)}$

$$= 1 \quad "$$

Denote IOU as n_{IOU}

From Eqn(1)

$$\text{Prob}(C) n_{\text{IOU}} \dots (8)$$

Note if more than one class,

$$\text{Prob}(C_i) = \text{Prob}(C_i) n_{\text{IOU}_i}$$

b. Consider Multiple Classes.

One class, from one Camera

$$\text{Prob}(C) n_{\text{IOU}}$$

One class, from multiple Cameras, from Eqn

$$\text{Prob}(C) = \sum_i \text{Prob}(C/I_i) \text{Prob}(I_i)$$

Hence

$$\begin{aligned} \text{Prob}(C) n_{\text{IOU}, i} \\ = \sum_{i=1}^k \text{Prob}(C/I_i) \text{Prob}(I_i) \cdot n_{\text{IOU}, i} \end{aligned} \dots (a)$$

Note: Now Expand the Above Analysis Beyond Multiple Cameras, treat Each Camera as input grid $G_i(x,y)$ from a given image, the Above Analysis holds good.

Homework: Due 1 week from today.
(April 19th).

1° Download, Install yolo program.
(yolov4).

2° Use cellphone to record 5-10 seconds of video clip for testing purpose.

3° Run Yolo Code to Perform default Detection Task

Team Project:

1° Presentation PPT. 7 Slides

10 Minutes for each team entire presentation.
PPT.; Demo; Q&A Session

2^o Topics, Deep Learning CNN Applications.

Consider Back Propagation Algorithm.

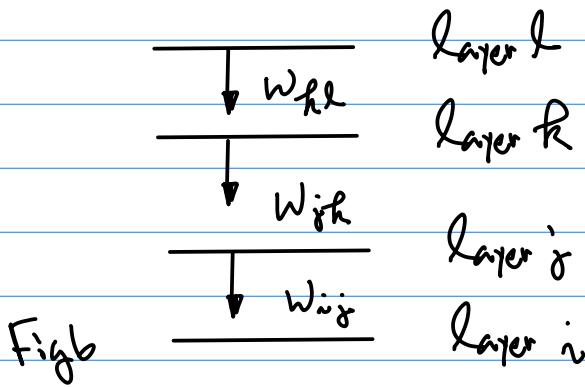
Background

- Multiple Layer Feedforward Neural Networks. MNIST

$$G^{M_1} \dots F \cdot D \cdot D$$



Layers



- For Each Layer.

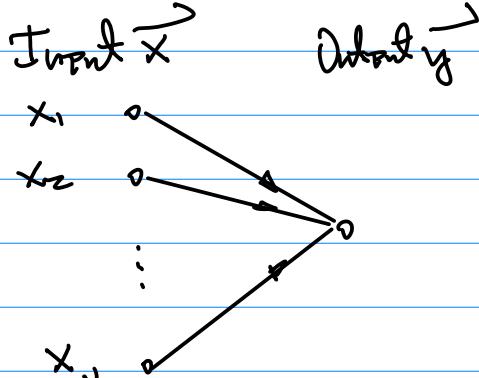
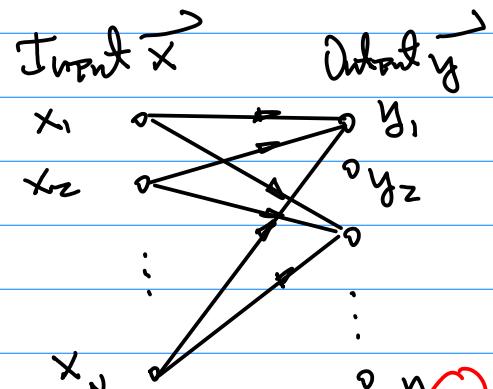


Fig. 7

$$y_j = f\left(\sum_{i=1}^N w_{ij} x_i + b\right) = f(h(w; x_i; b)) \quad \dots (1)$$



For multiple Output Neurons. Fig. 8

- Notation for weights at Each Layer, From the Input Layer i towards the output j, h, and l

$$w_{i,j}$$

The Index of the Output @ the Current Output Layer i
So, we have

The Index of the Input @ Layer j

$$w_{i,j}, w_{j,k}, w_{k,l}$$

- Output $y_j = f\left(\sum_{i=1}^N w_{ij} x_i + b\right)$ for one Output Neuron. ... (2)

Output for Multiple Neurons per Each Layer, and Multiple Layers as well.

In Addition, Let's Denote Experiment (CNN Output y) as \tilde{y} ("tilde y "),

Let ground Truth denoted as y .
Error for one experiment then is defined as

$$\tilde{y} - y \text{ or } y - \tilde{y} \quad \dots (4)$$

Now, Loss function :

4. Loss function.

$$L = \frac{1}{2} \sum_{M=1}^M \sum_{i=1}^N (\tilde{y}_i^M - y_i^M)^2 \quad \dots (5)$$

No. of Experiments No. of Output

Based on the Notation for the Experiments and Ground Truth,

Rewrite Eqn(1) as

$$\begin{aligned} \tilde{y} &= f\left(\sum_{i=1}^N w_i x_i + b\right) = f(h(w; x_i; b)) \\ \text{OR} \\ \tilde{y} &= f(h) \end{aligned}$$

5. The weights update equation

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \delta w_{ij}^{(t)} \dots (6)$$

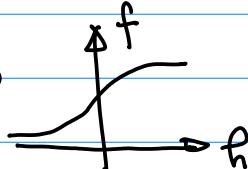
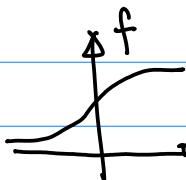
$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}} \left[\sum_{M=1}^M \sum_{i=1}^N (\tilde{y}_i^M - y_i^M)^2 \right]$$

From Here, Apply 'Chain-Rule' for multiple layers. i, j, k, l .

5. Now, Consider Eqn(1) & (2)

Eqn(1) : Has Activation function $f(\cdot)$;
Has Transfer function $h(\cdot)$;

$$y = f\left(\sum_{i=1}^N w_i x_i + b\right) = f(h(w; x_i; b)) \quad \dots (1)$$



$$\sum_{i=1}^N w_i x_i + b \stackrel{?}{=} h(w; x_i; b) = h(i)$$