Lecture Notes
PART II

CmpE258
Spring 2023

March 21 (Tue).

Midterm is scheduled on the 23rd (Thu). 1 hr Exame. 16:30–17:30

Then 15 minutes for prep & Uploading the file.

Example: Softmax Activation Function. MNIST CNN

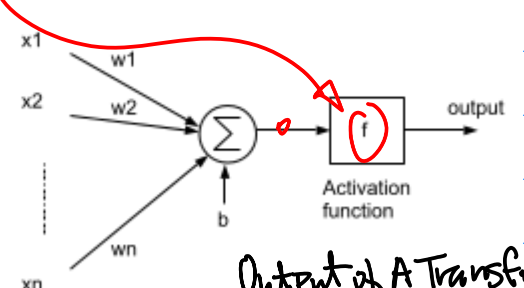
Note:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad \text{for } i = 1, \dots, K \text{ and } \mathbf{z} = (z_1, \dots, z_K) \in \mathbb{R}^K \quad \dots (1)$$

Dimension

Let K=1

1° $f(z)$ Notation



Output of A Transfer function

2° Index $f(z_i) = \dfrac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$ ...(2)

Total No. of Output Neurons = K for Hand Written Digits Recognition K=10;

Where

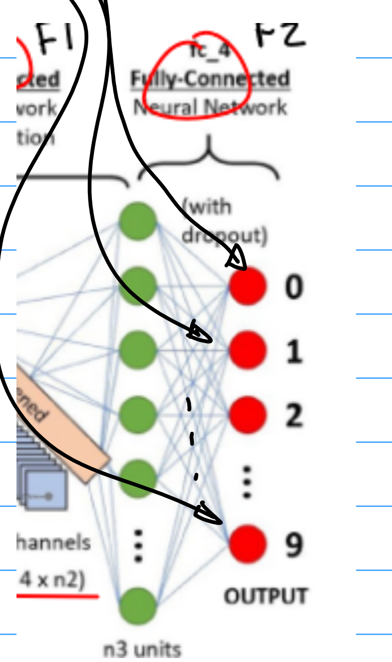
$$\sum_{j=1}^{K} e^{z_j} = e^{z_1} + e^{z_2} + \dots + e^{z_K} \geq e^{z_i} \quad \dots (3)$$

Output from Each Neuron:

$$f(z_1) = \frac{e^{z_1}}{\sum_{j=1}^{10} e^{z_j}} \quad \dots (4)$$

for Digit "0" (1st Output)

And



$$f(z_2) = \frac{e^{z_2}}{\sum_{j=1}^{10} e^{z_j}}$$ for Digit "1" (2nd output) ...(5)

$\vdots$

$$f(z_{10}) = \frac{e^{z_{10}}}{\sum_{j=1}^{10} e^{z_j}}$$ for Digit "9" (10th o/p)

if Add Eqn (1) + Eqn (2) + ...

$$f(z_1) + f(z_2) + \dots + f(z_{10})$$

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

Property 1.

$$= \frac{\sum_{j=1}^{10} e^{z_j}}{\sum_{j=1}^{10} e^{z_j}} = 1. \quad \dots (6)$$

CmpE258
Spring 2023

The softmax function takes as input a vector z of K real numbers, and normalizes it into a probability distribution consisting of K probabilities proportional to the exponentials of

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

See Eqn (6).

Now, move to the 2nd Half of DCNN.

2022F-108a-Yolo-architecture-loss-function-2022-10-10.pdf

Typical Classification/ Recognition Results are in Bounding Boxes

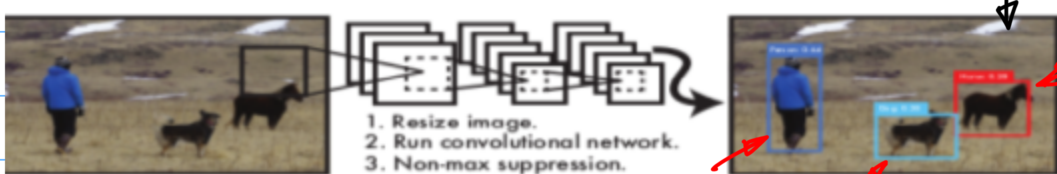
Unified, Real-Time Object Detection

://arxiv.org/pdf/1506.02640v5.pdf

h Divvala, Ross Girshick, Ali Farhadi
, Allen Institute for AI ,Facebook AI Research http://pjreddie.com/yolo/

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/

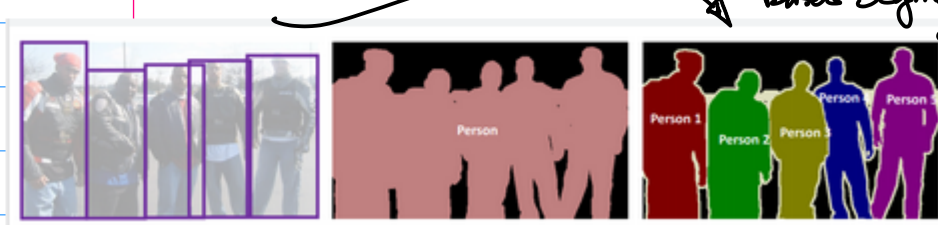


del runs at 45 FPS. A smaller version, Fast YOLO, runs astounding 155 itperforms DPM (deformable parts models) and R-CNN.

Figure 1: The YOLO Detection Sy image to 448 × 448, (2) runs a co network, and (3) thresholds the re

Then, we would like to Achieve Semantic Segmentation.

Pixel by pixel Based Segmentation/ Detection/ Recognition



PART II (After the midterm)

April 4 (Tue)

Road map: Yolo (You-Only-Look-ONCE) → Semantic Segmentation.

{ Project Assignment Yolo Implementation = Due $2\frac{1}{2}$ weeks. April 20th. (Thursday)

CmpE258

Spring 2023

Homework (In-Class Presentation)

Requirements Due April 6 (Thu).

1° One Paragraph Description (Abstract) of the proposed Semester-Long Project.

2° Title

Team members: First Name, Last Name, Major

Team Coordinator.

Contact E-mail.

3° Abstract Part.

Objective(s): a) What is the proposed work;

b) What is the Coding / Training / Testing Task involved in the project?

c) Anticipated Result? And deliverable?

d) Tools, platform, Programming Language Version, T.F., Pytorch, ChatGPT etc.

Also, Define Python Packages, OpenCV.

Example: On Yolo.

Ref:

2022F-106-README-Tiny-Yolo4-GP...

Note 1: Readme for Yolo github 34

Installation & Testing for the project.

Title: README Tiny Yolo v4 GPU Ubuntu

Document Number: 105-1b

CTI One Corporation

Table 1a. Document History

2022-10-6	Establish this document, document archive: (base) harry@workstation:~/yolo-2022-10-19$	YY

1. Setup YOLO v4 environment

1.1. Clone the GitHub folder;

$ git clone https://github.com/pythonlessons/TensorFlow-2.x-YOLOv3.git

1.2. Create YAML file for building the YOLO v4 Anaconda environment;

Create TensorFlow-2.x-YOLOv3/conda-gpu.yml as the following;

==

name: yolov4-gpu

Ref: Introduction

2022F-108a-Yolo-architecture-loss-function-2022-10-10.pdf

Base-Line Ref for Yolo Technique

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/

Lecture Notes: Base Line Ref/Requirements.

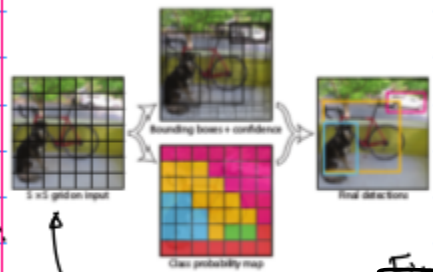
2022F-101-cmpe258-note-2022-11-1.pdf

Example: Notations for YoLo.

Ref, pp36.

CmpE258
Oct.13,22

Fig 1.



~~Fig.2a~~

Note:

1. Image I(x,y) is divided into SxS Grids. Denote it as $G_{p,q}(x,y)$, ...(1) where p,q = 0,1,2,... indicate the

1) Image I(x,y). with Resolution

M x N

No. of Col. No. of Rows.

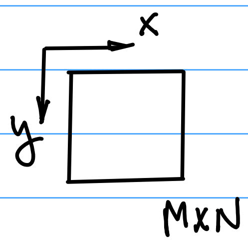


Fig.1b.

Divide I(x,y) into SxS Grids.

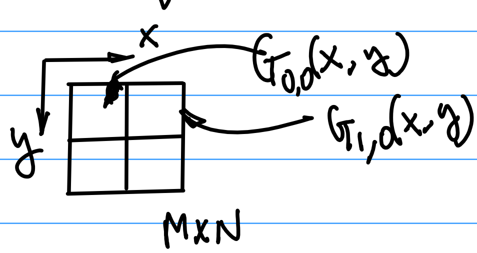
Each Grid is Denoted as

$$G_{pq}(x,y) \quad \dots (1)$$

where p,q = 0,1,2,...,S-1

matches to x

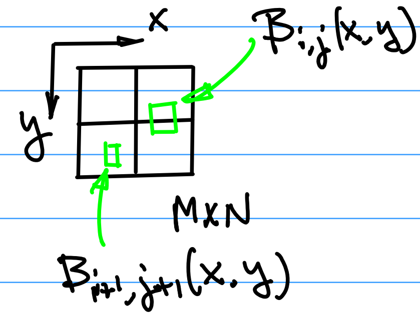
Fig1c



2. Bounding Boxes $B_{ij}(x,y)$ ... (2)

i for x-direction; j for y-direction

Fig1a



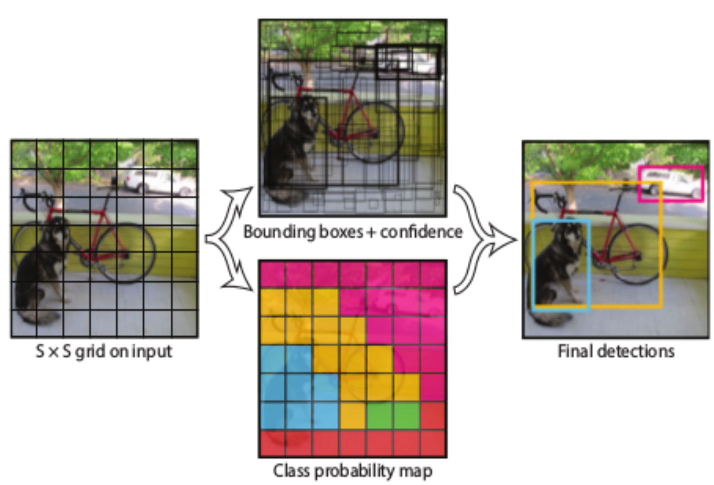
3. Five Parameters to define each Bounding Box.

(x,y) : Location of the top Left Corner of $B_{ij}(x,y)$

w, h : Width and Height of $B_{ij}(x,y)$

f : Confidency Level, Probability distribution to Describe the likelihood of the B.B (B^2) belongs to a certain Class of objects.

$$(x, y, w, h, f) \quad \dots (3)$$



CmpE258

Spring2023

Note1. a. ($G_{pq}(x,y)$ Grid.

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.

Class probability.

b. $B_{ij}(x,y)$

c. (x, y, w, h, f)

Probability

Confidency.