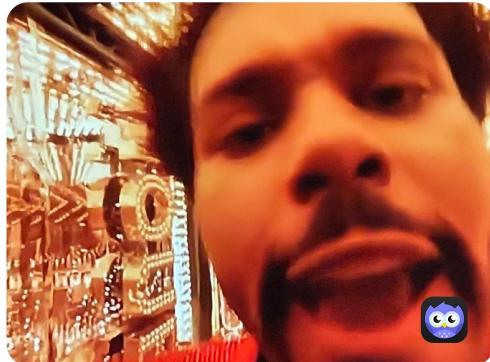


Session - 10

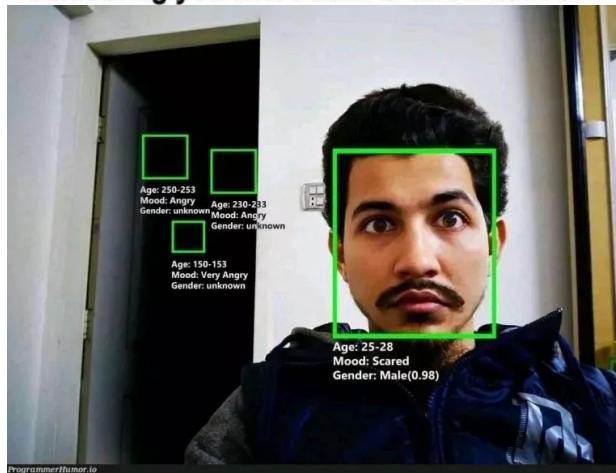
Siamese Network

April 03, 2024

- what the face recognition sees when I go on my phone after my alarm wakes me up

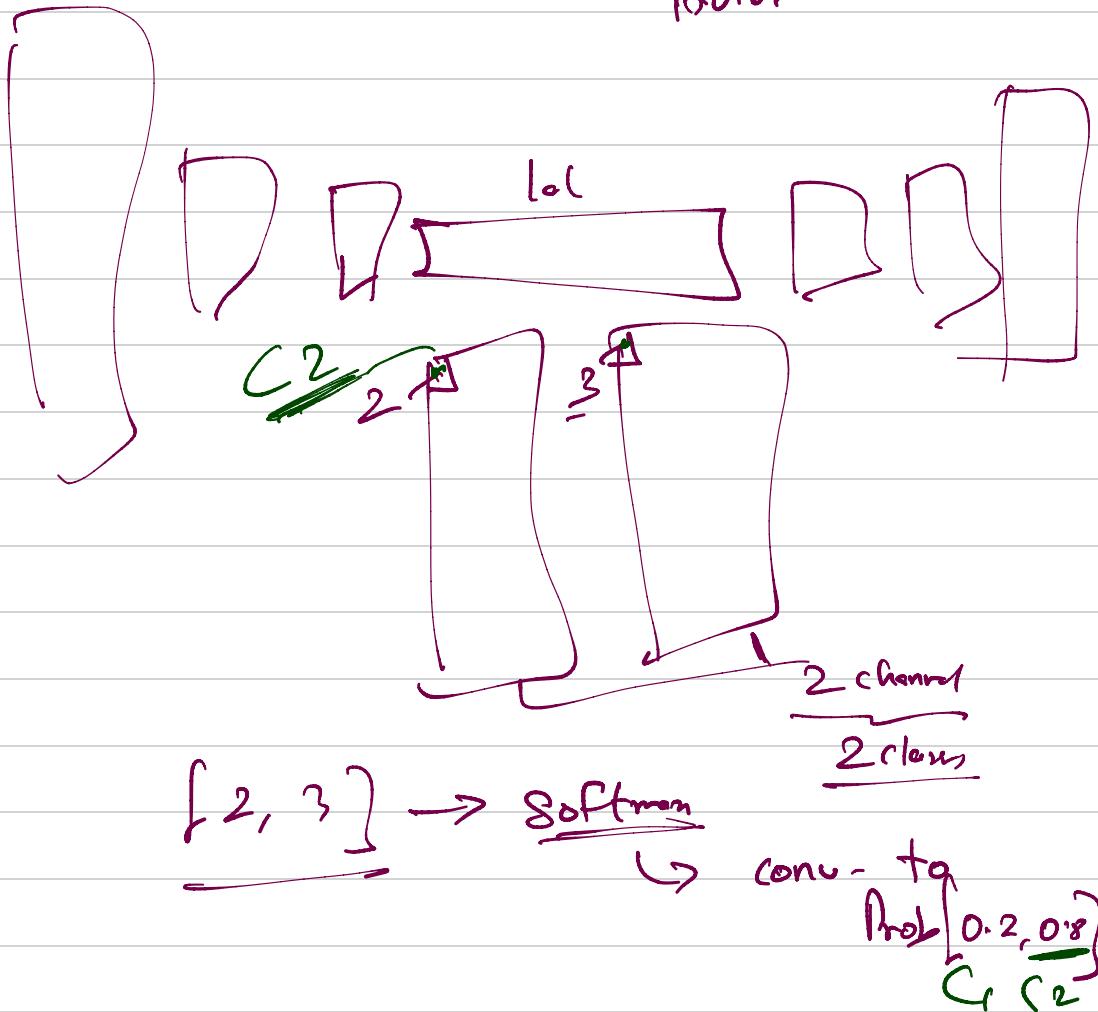
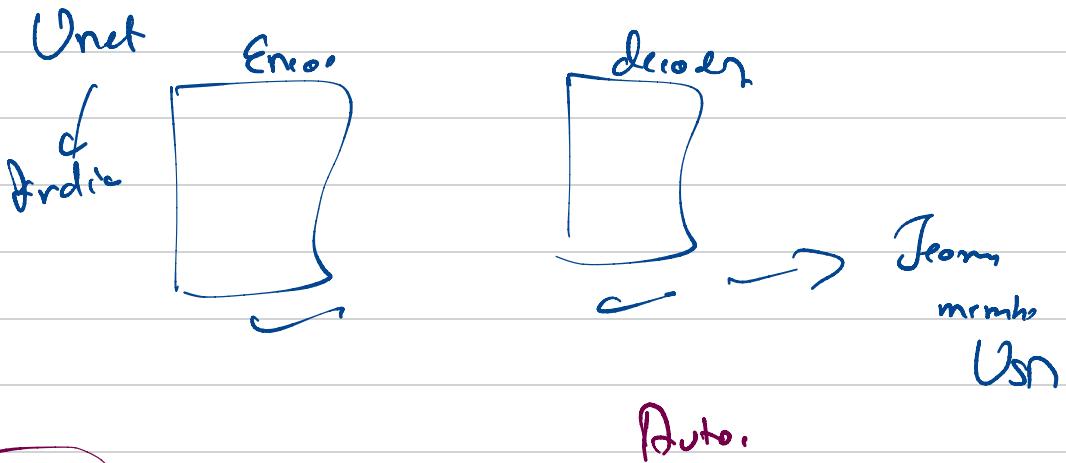


When your face recognition system detects something you don't want to know about



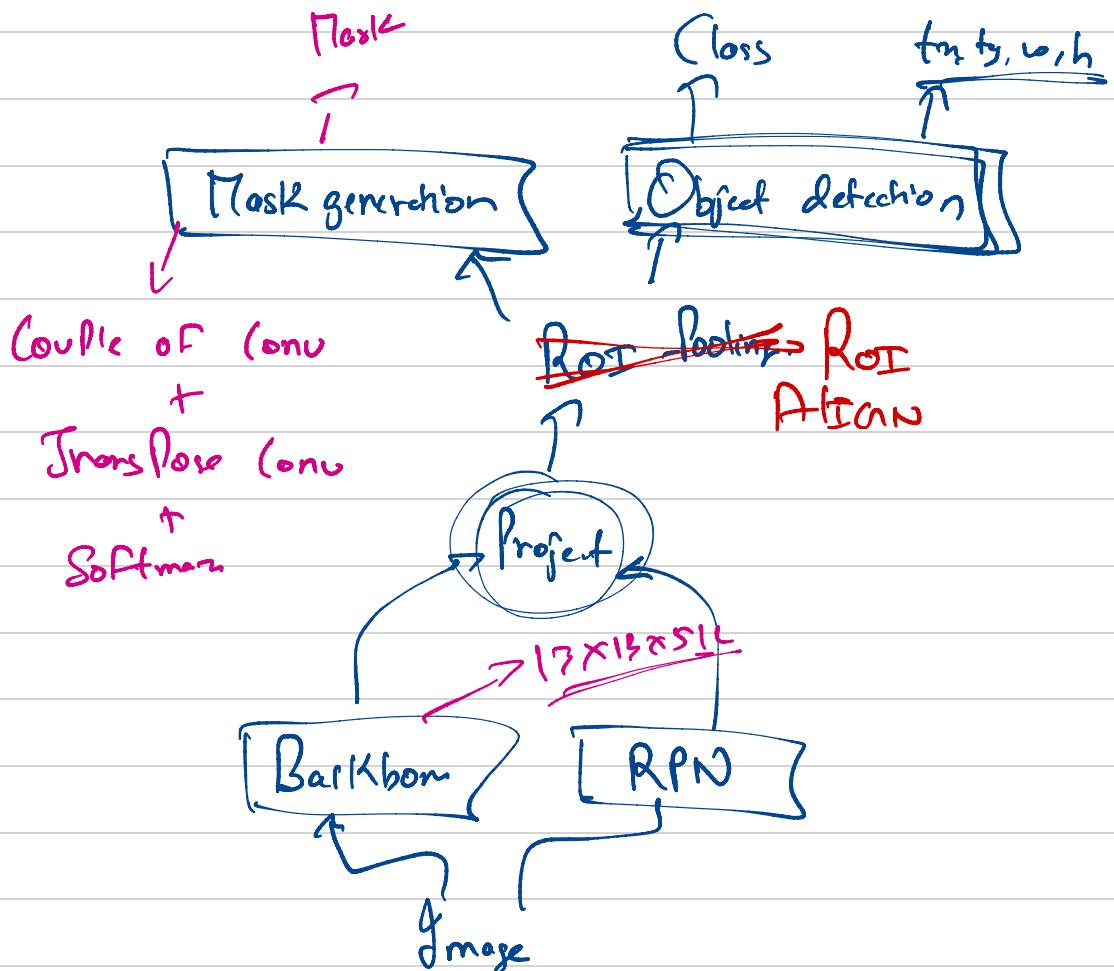
Agenda

- ① UNET
- ② Mask RCNN
- ③ Siamese Networks → Contrastive
→ Triplet loss

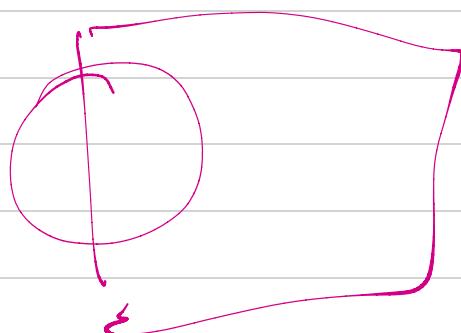
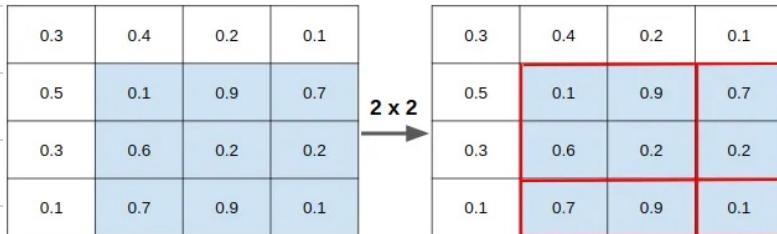


Mask - RCNN

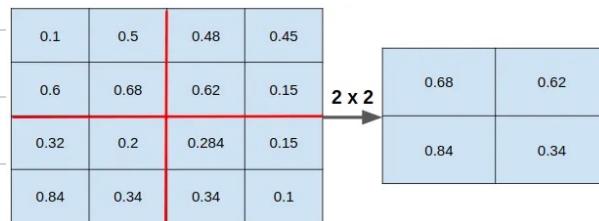
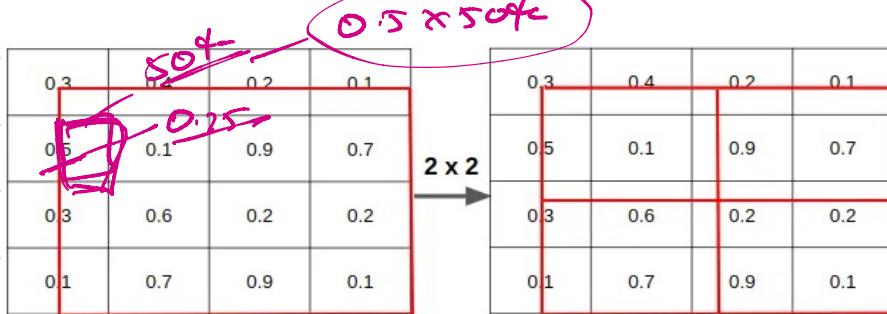
Faster RCNN



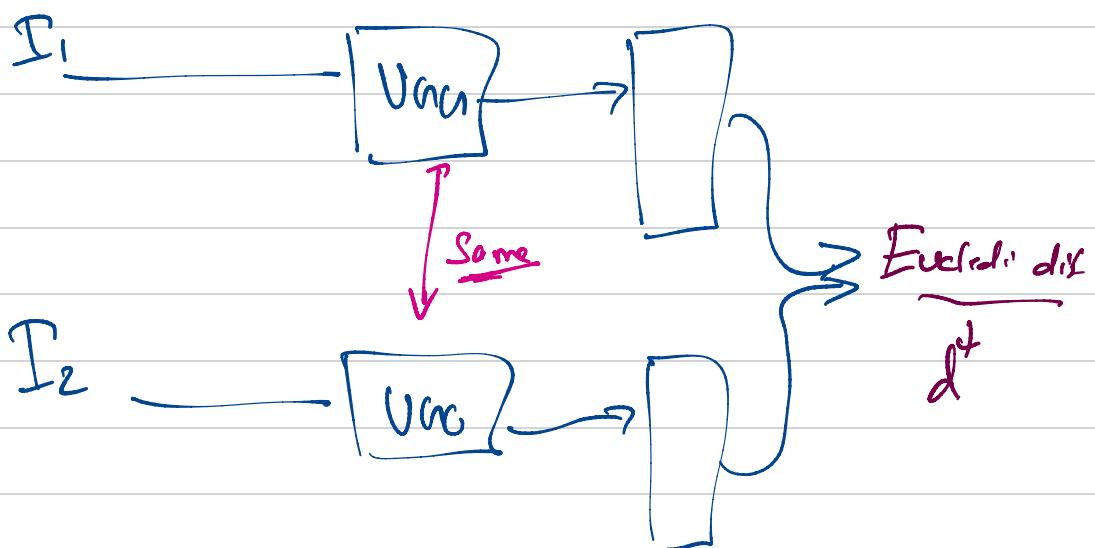
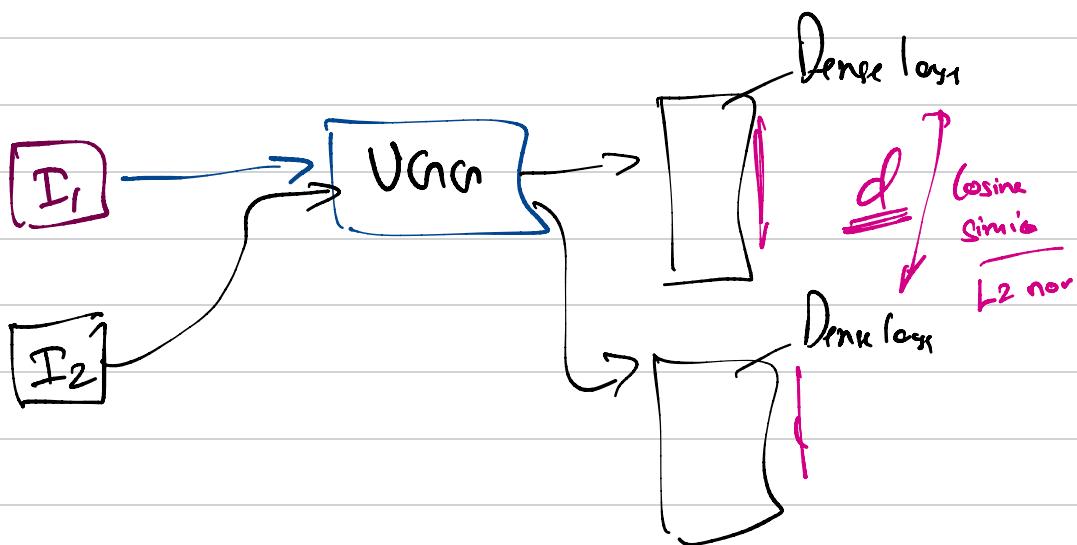
ROI - Region Pooling



<https://firiuzamedium.com/roi-pooling-vs-roi-align-65293ab741db>



SIAMESE NETWORK



$$\text{Sim. fm} \rightarrow d \approx 0 / v \cdot \log$$

$$\text{Non-Sim fm} \rightarrow d \approx \text{Inf}$$

→ Backpropagated \rightarrow Intrinsic model predictions
so that $d \rightarrow \text{Non-Sim} \uparrow \uparrow \text{vfo}$
 $\rightarrow \text{Sim} \approx 0$

Stable network \rightarrow fm_{opt}

For ongoing

Video \rightarrow Embedding



FAISS
SCANN

10 \approx 20 \approx 1/2

Pinterent

$$f_{m1} \geq f_{m2} \rightarrow \delta = 1 \quad (\text{Contrastive loss})$$

$$\text{"!>"} : \underline{\delta = 0}$$

$$\underline{Y = 1 \rightarrow \underline{\delta}} \quad \underline{\delta = 0} \quad (\text{similar error})$$

$$\underline{Y = 0 \rightarrow \cancel{\delta \rightarrow 0}} \quad \cancel{\text{Margin}}$$

δ b/w f_1 & f_2

is equal to margin
 $f_m \text{ half}$

$$\cancel{1000} \rightarrow \delta \text{ b/w } I_r \text{ & } I_s \quad (Y=0)$$

Contrastive loss =

$$Y \times D^2 + (1-Y) \max(\text{Margin} - D, 0)^2$$

$$\underline{Y=1} \quad (I_1 = I_2)$$

CASE - 1

$$Y=0 \quad \max\left(\text{Margin} - D, 0\right)^2$$

$$D = \underline{\underline{200}}$$

$$\text{bco } I_1 \& I_2 \rightarrow \underline{\underline{200}}$$

$$\text{Margin} = \underline{\underline{1000}}$$

$$\max\left(\frac{1000 - 200}{800}, 0\right)^2$$

CASE 2

$$D = \underline{\underline{2000}}$$

$$\underline{\underline{800}}^2$$

$$\max\left(\frac{1000 - 2000}{800}, 0\right)^2$$

$$= \max\left(\frac{-1000}{800}, 0\right)^2$$

$$= \underline{\underline{0}}^2 = 0$$

$f_{m, f_{1,2}} \rightarrow \underline{\underline{d}} < \underline{\underline{50}} \rightarrow \text{Similar}$

$$\begin{aligned} M &\rightarrow 1000 \\ D &\rightarrow 1800 - \underline{\underline{5000}} \rightarrow 5000 \end{aligned}$$

$$Y = 0$$

D b/w $\underline{I_1}$ & $\underline{I_2} \rightarrow \underline{\underline{1200}}$
 \downarrow
5000

$$\text{C.L.} \quad \text{min} (Margin - D, 0)^2$$

$$Margin = \underline{\underline{1000}}$$

$$= \text{min} (1000 - 1200, 0)^2$$

$$= \text{min} (-200, 0)^2$$

$$= 0^2 = \underline{\underline{0}}$$

TRIPLET Loss

Shivam

Shivam — F

PsiF — F

Anchor Image

True Image

Fake Image

Posit

Negativ

Best Image for a class

$$D(A, P)$$

$$D(A, N)$$

$$D(A, N) - D(A, P) > \underline{\alpha}$$

Shiven \rightarrow 80

Prothamed \rightarrow 30 / 20

$$\begin{array}{r} \text{Find} \rightarrow 160 \\ = \\ = \\ = \\ | \\ | \\ \hline 160 \end{array}$$

20 } Real
30 } Fake

CV

① Code a UNET

② Resnet \rightarrow Residual block
 $\underbrace{\text{diff}}$

free phantom block

Why deeper with Resnet

③ Diff blur dotted line
normal line

④

Kernel / Feature maps / Channel

Feature
entity/
filter

⑤

How does gradient
propagate?Back \rightarrow NN \rightarrow Cover
backpropagation } 2 leftWe have done
in lot no
det.

⑥

Do we need more RF
than Image size

⑦

GAP \rightarrow why do we
use?

⑧

How to calculate Gradient
in max-pooling?

7

Do we need non-linearity
for CNN

