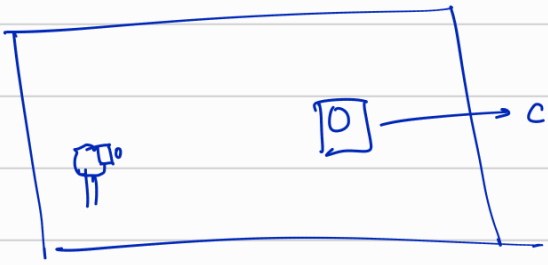
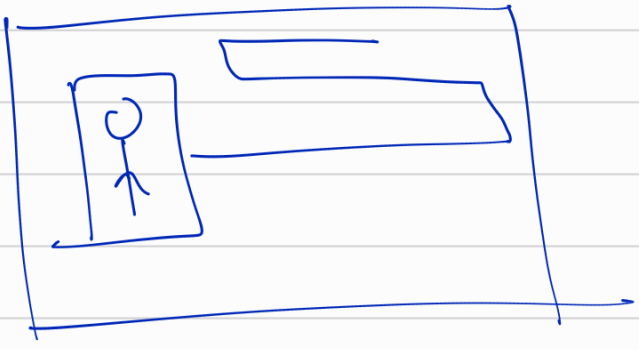
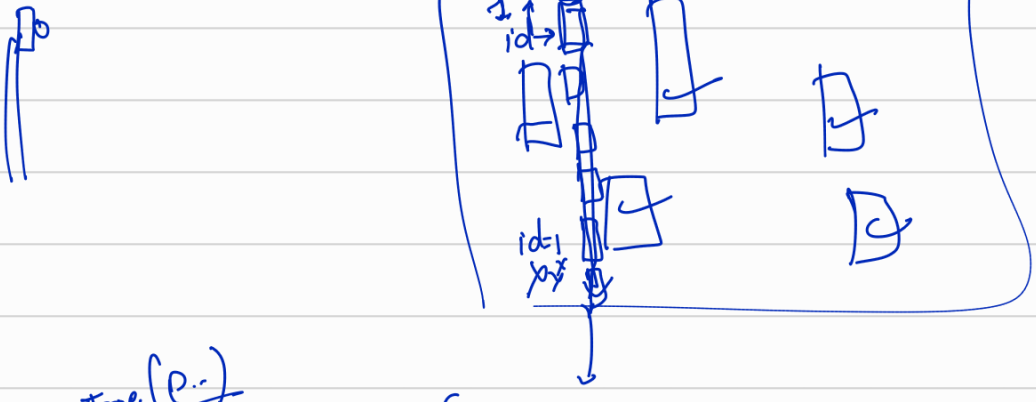
Object detection

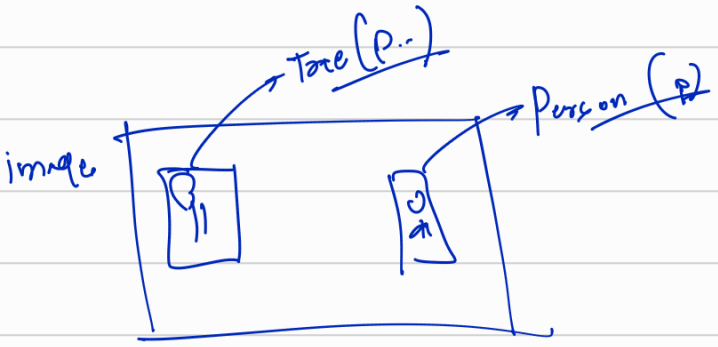
↳ Object Tracking

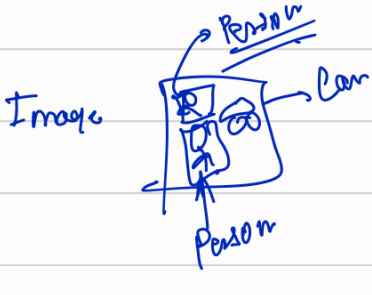


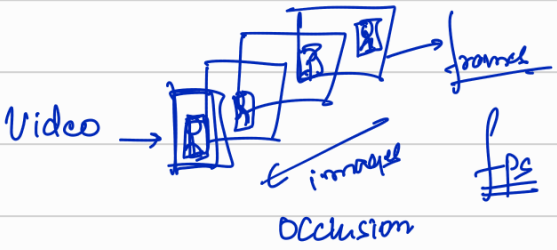


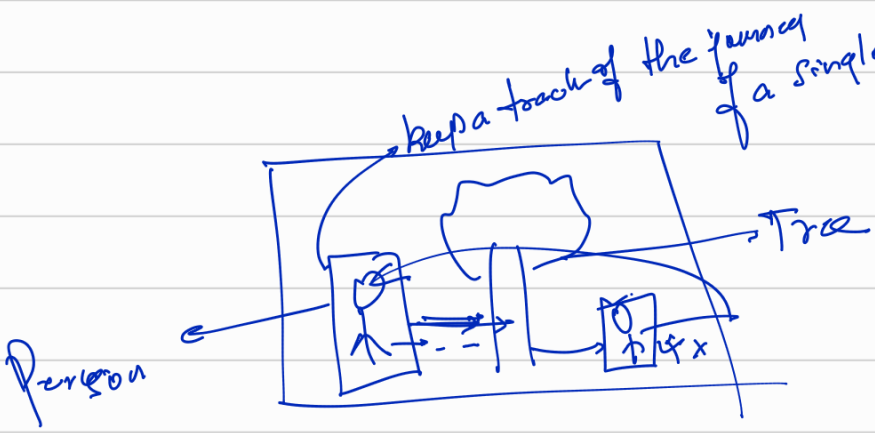
CCTV footage → Traffic cams

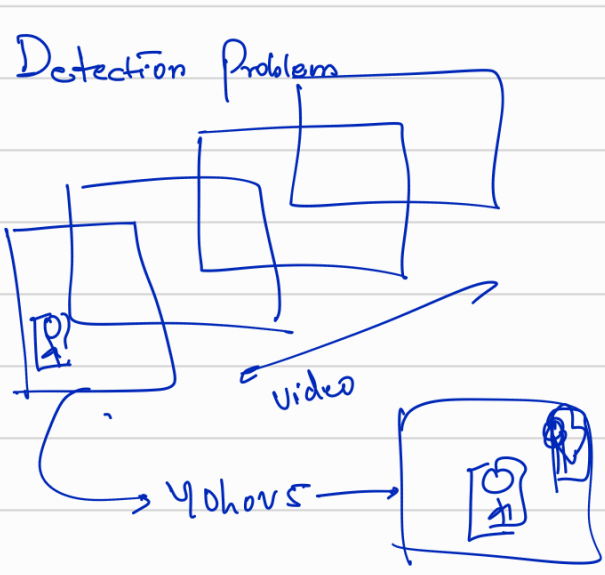


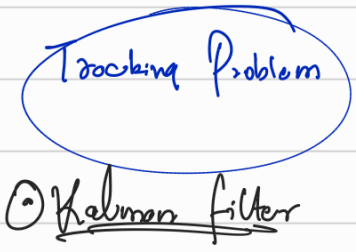


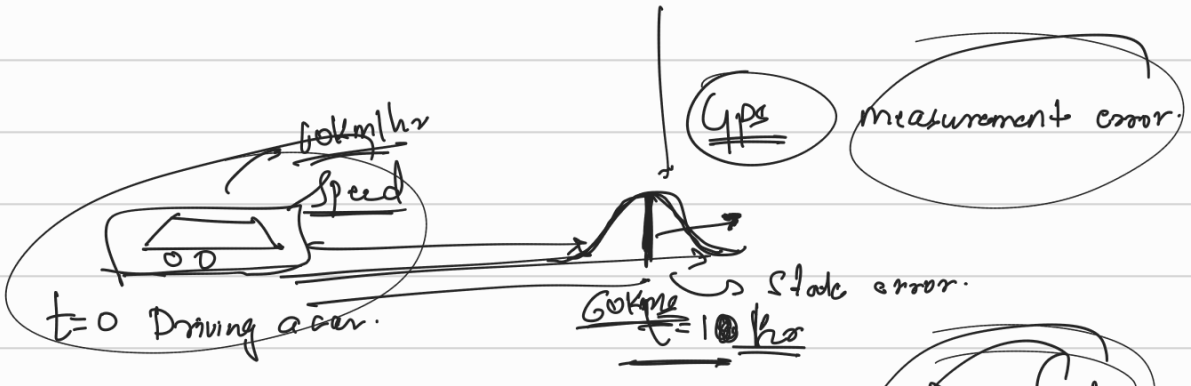


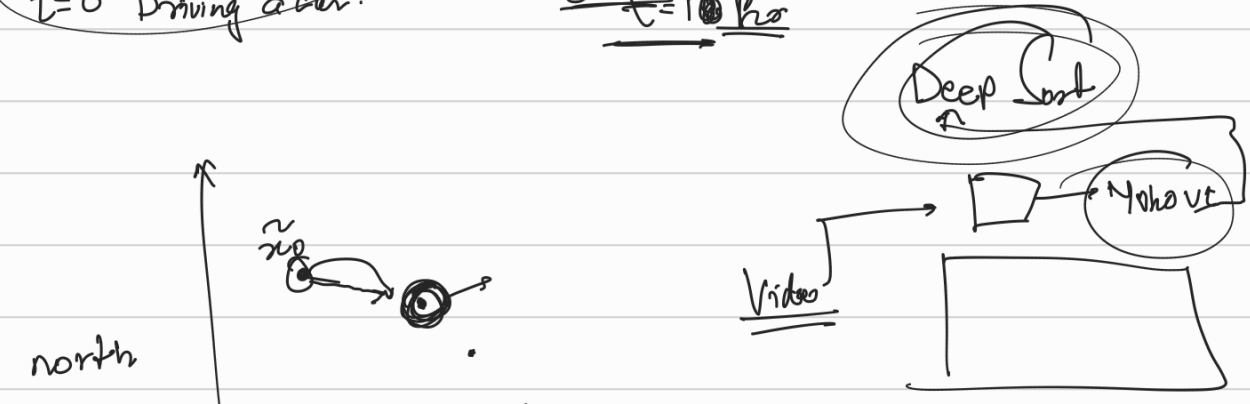


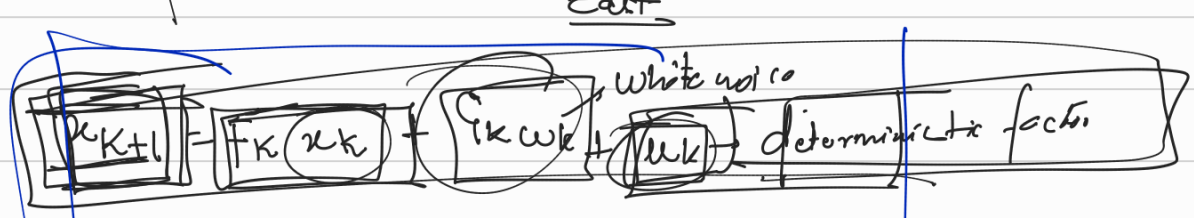

keep a track of the journey of a single object
Person
Tree


Detection Problem
video
Yolov5

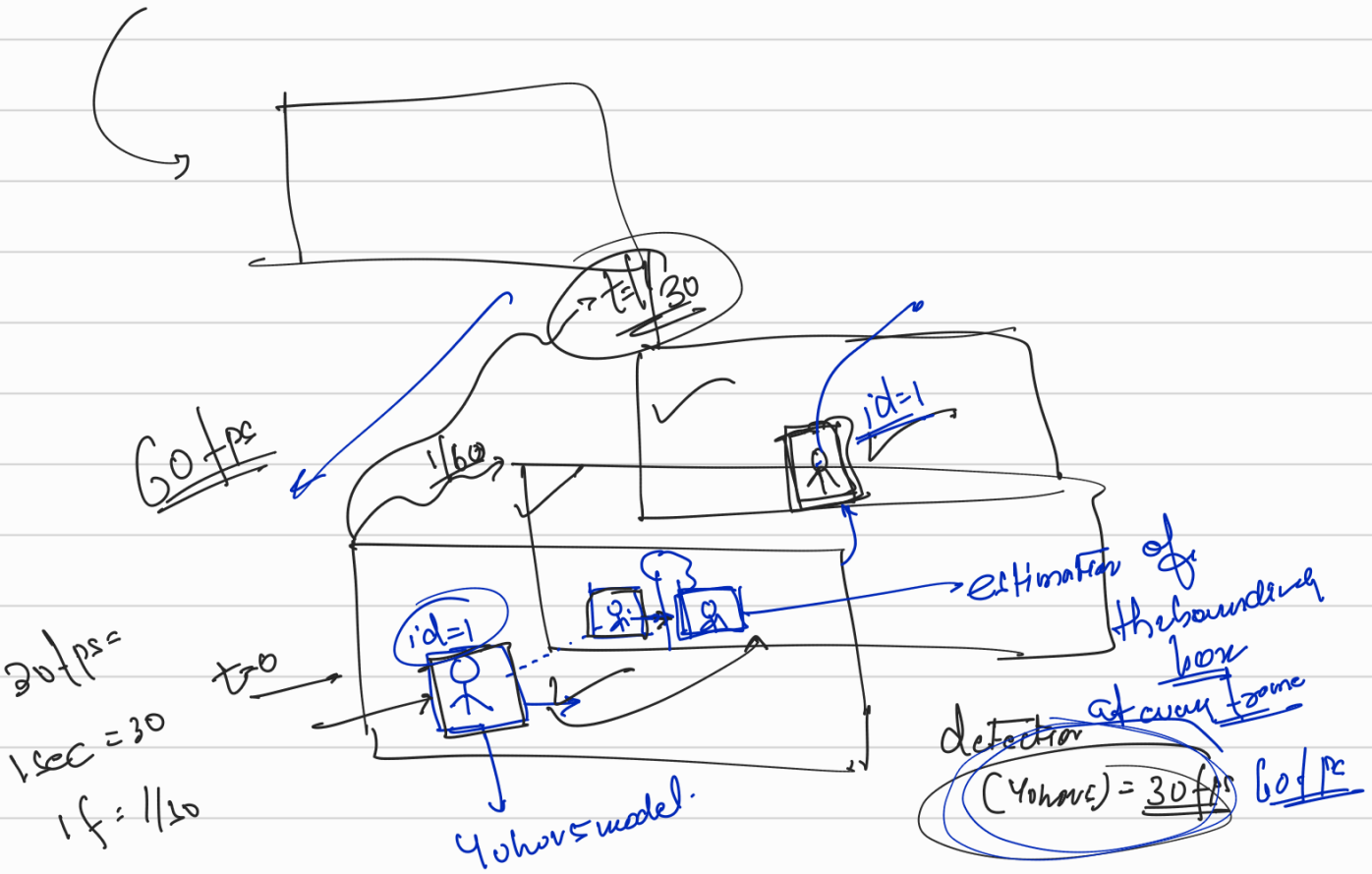

Tracking Problem
Kalman filter

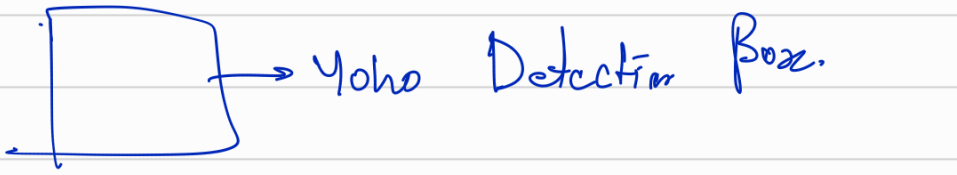

GPS
measurement error.
60km/hr
Speed
t=0 Driving accr.
State error.
60kmph
t=10 hr

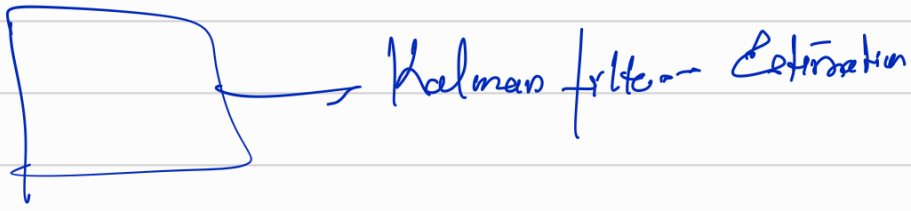

Deep Sort
Yolov5
Video
north

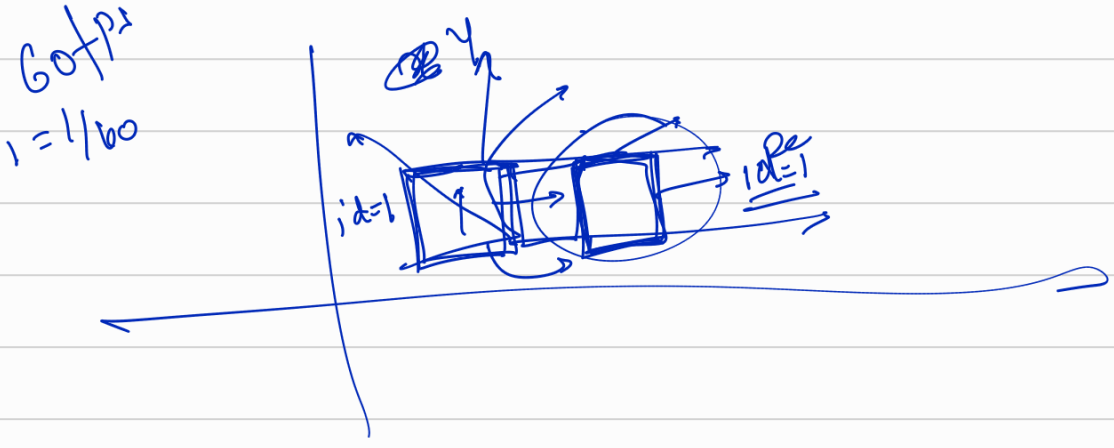

calt
white noise
deterministic facto

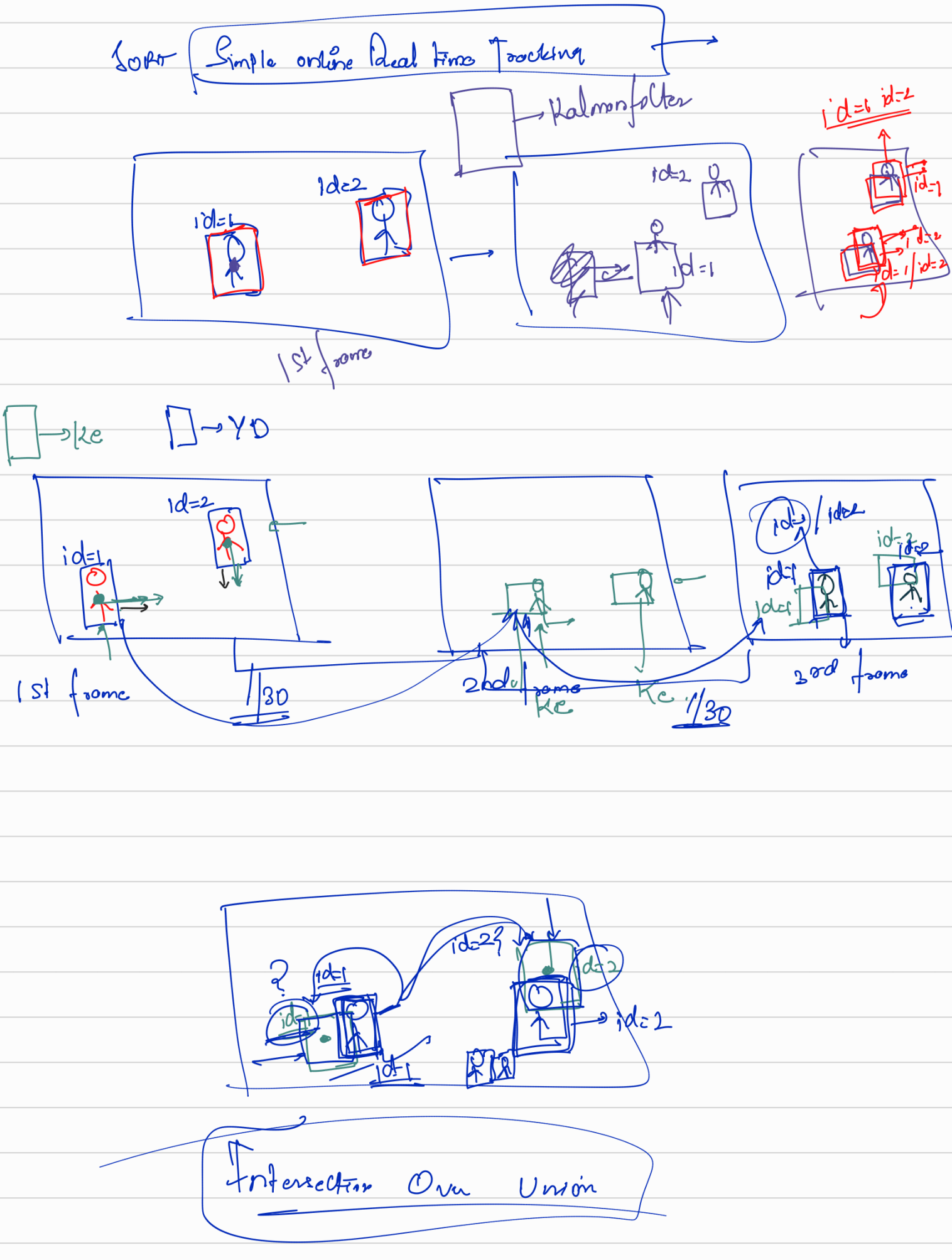
$Z_k = H_k x_k + V_k$ → measurement / observed position; → measurement noise

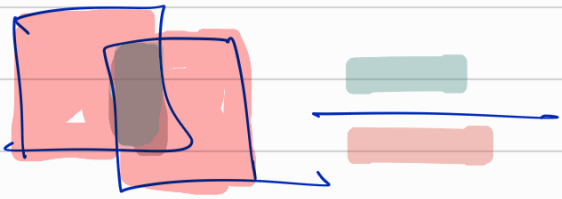


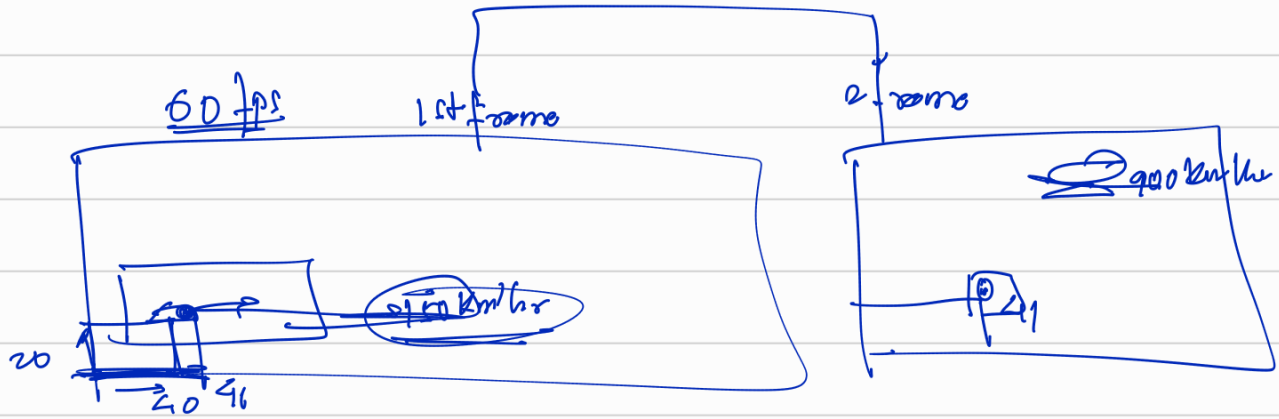







SORT
Simple online Real time Tracking
Kalman filter
id=1
id=2
1st frame
id=1 id=2
2nd frame
3rd frame
KF
YD
1/3D
Intersection Over Union



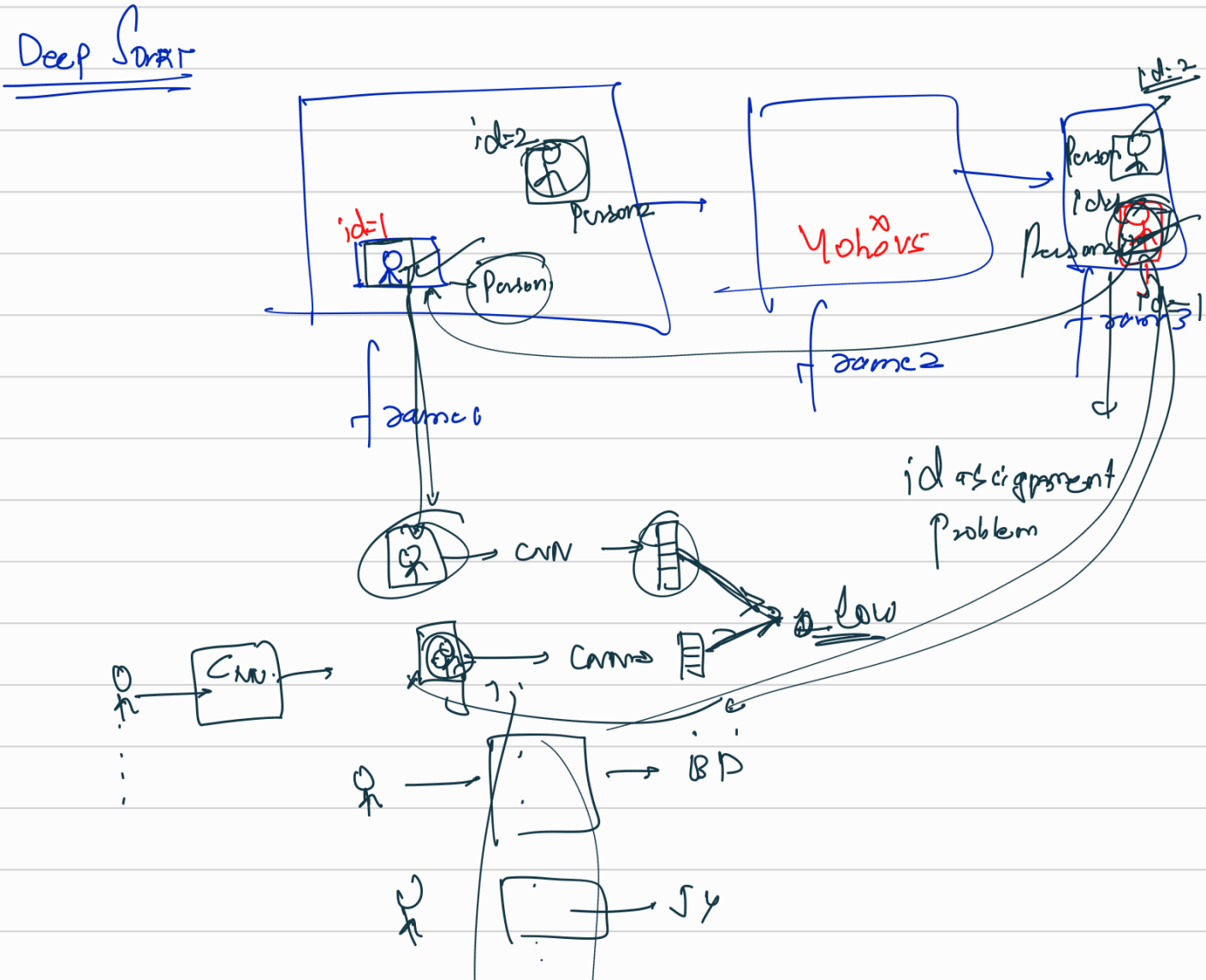


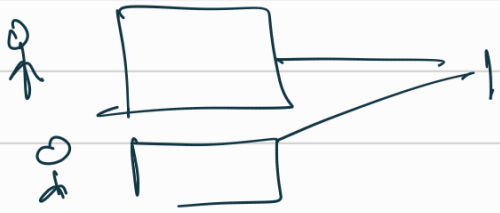
1:1m
150 Km/hr

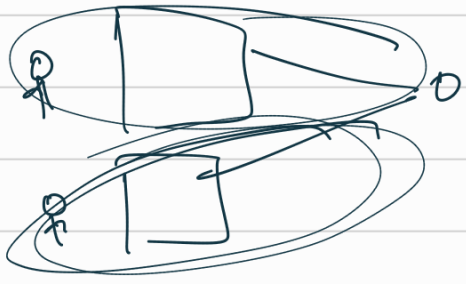
1/60 seconds

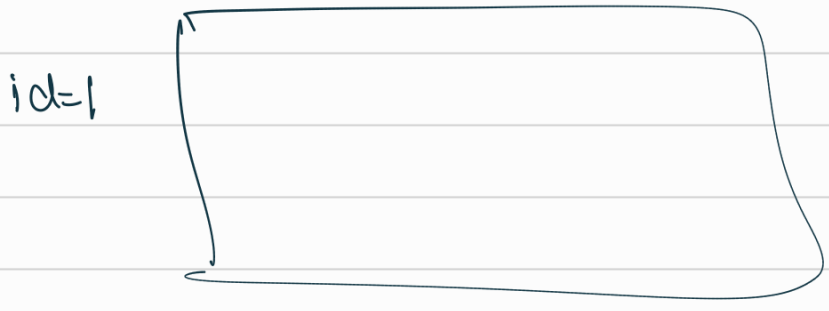
150 Km × 1/3600 × 1/60 = 0.001 km = 1 m

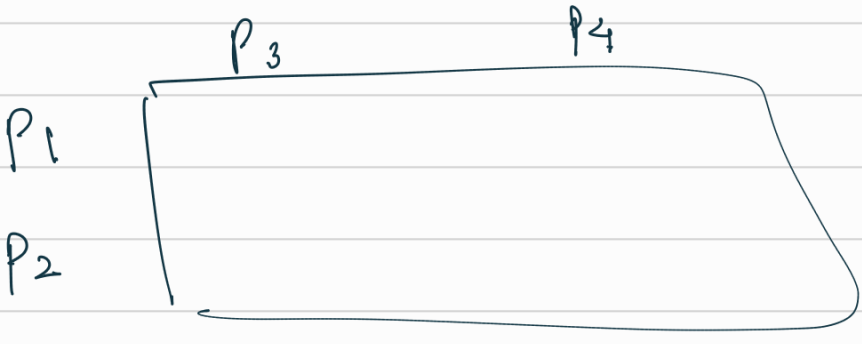
Deep SORT











P2, P3 P2 P4

Deep SORT

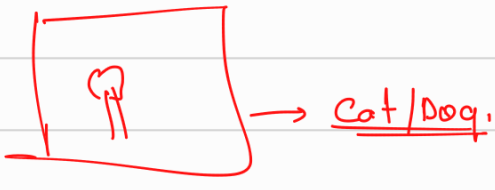
Combined output (IOU + Euclidean dist)

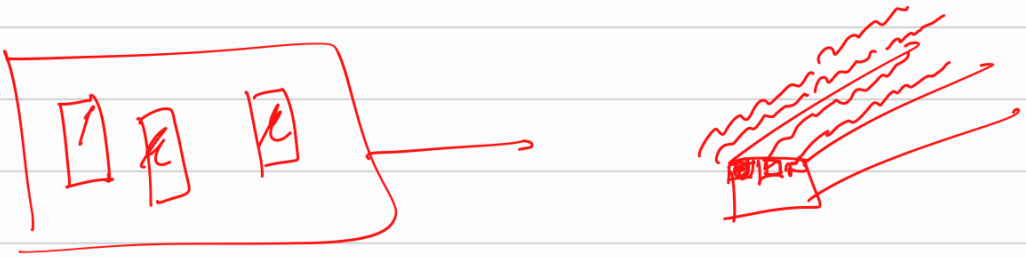
9:20 pm → 9:30 pm

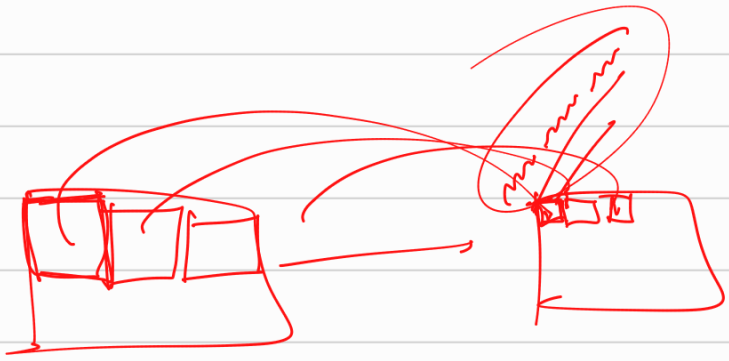
9:30 → 10:15 (Code).

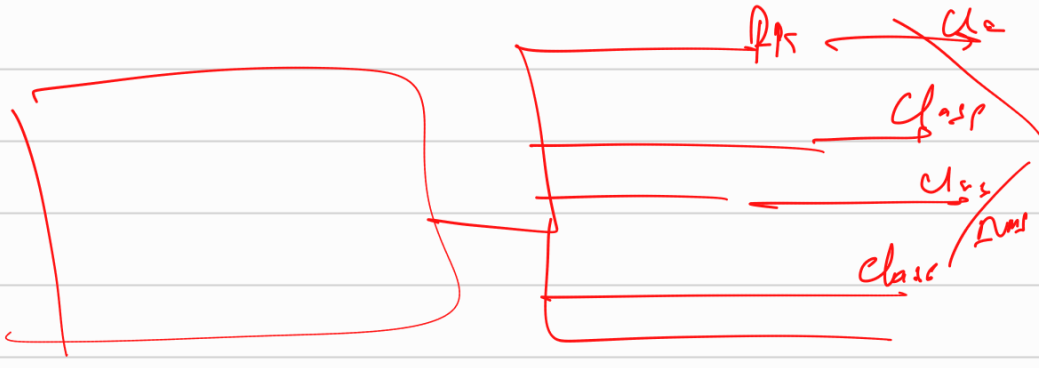
CV → Image classification

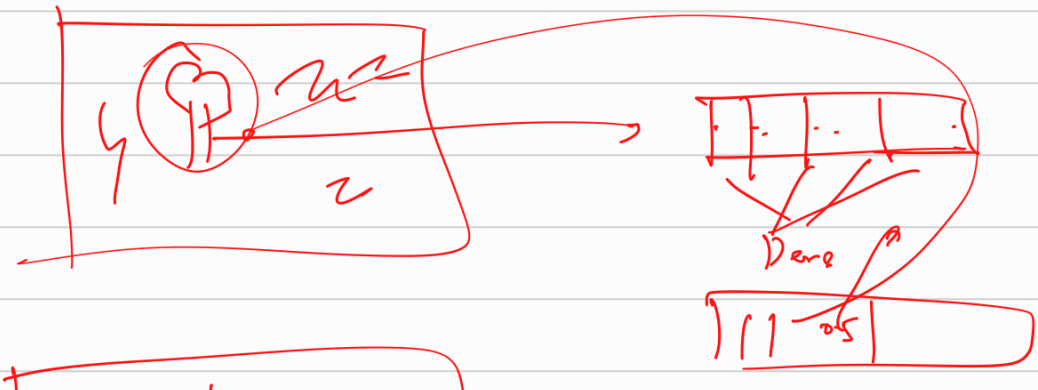
- fast rcnn video
- Resources for Object tracking

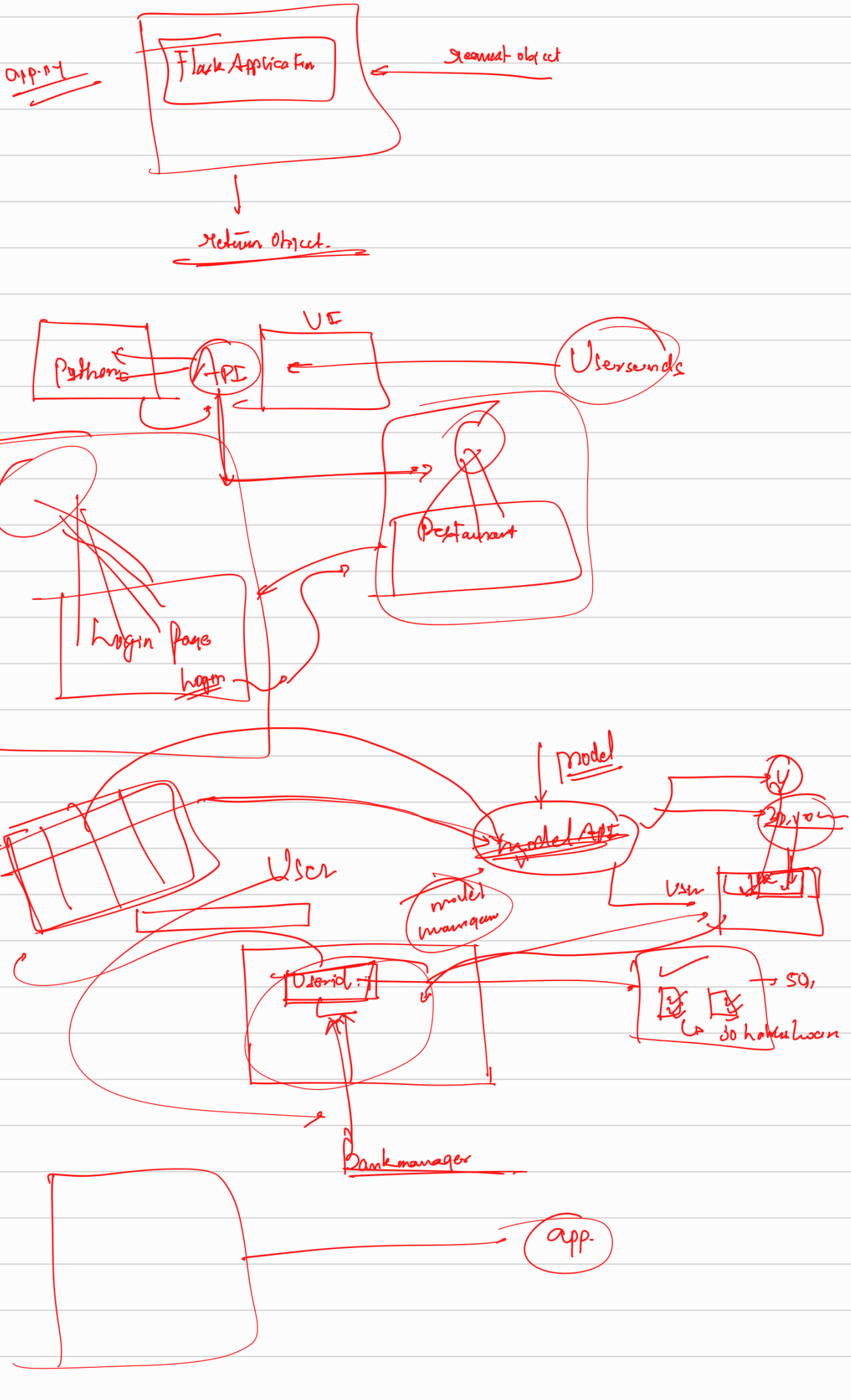


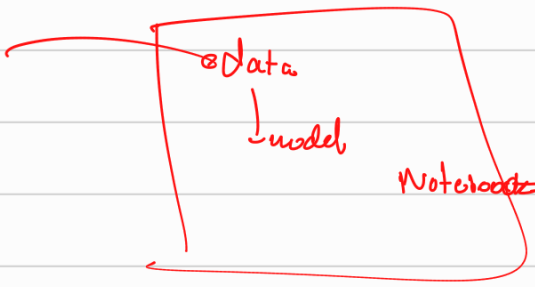








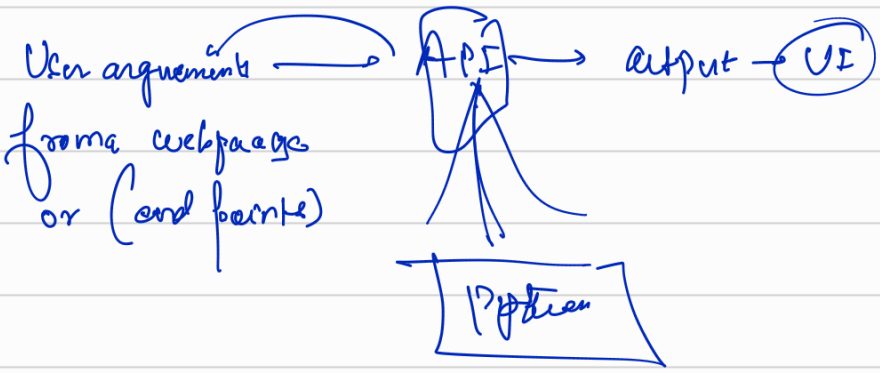

app.py
Flask Application
request object
Return Object.
UI
Pathons
API
Usersends
Restaurant
Login Page
login
model
User
User
model manager
Userid
Bankmanager
app.

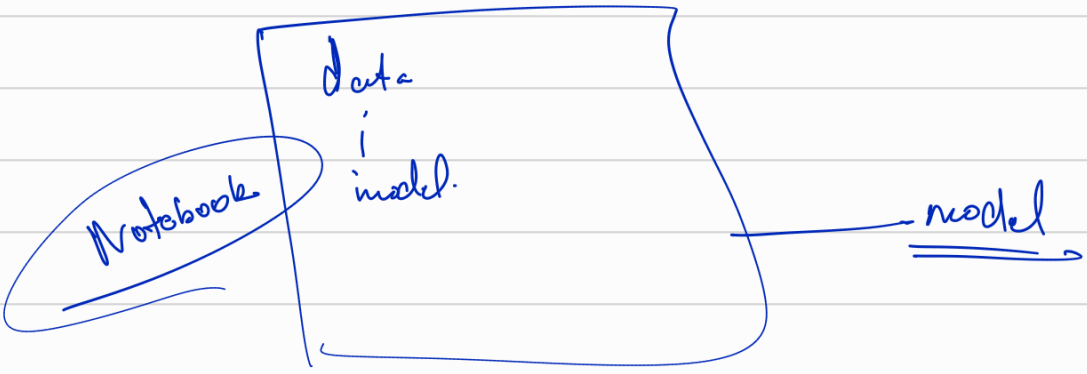


app.

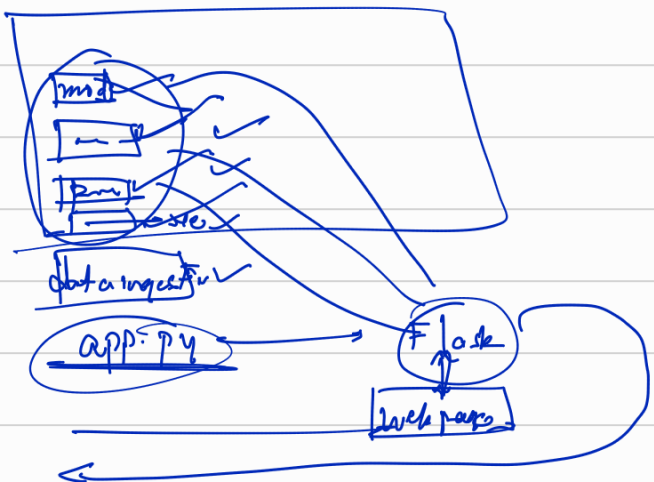
Flask API

arguments → Functions → output





VS code



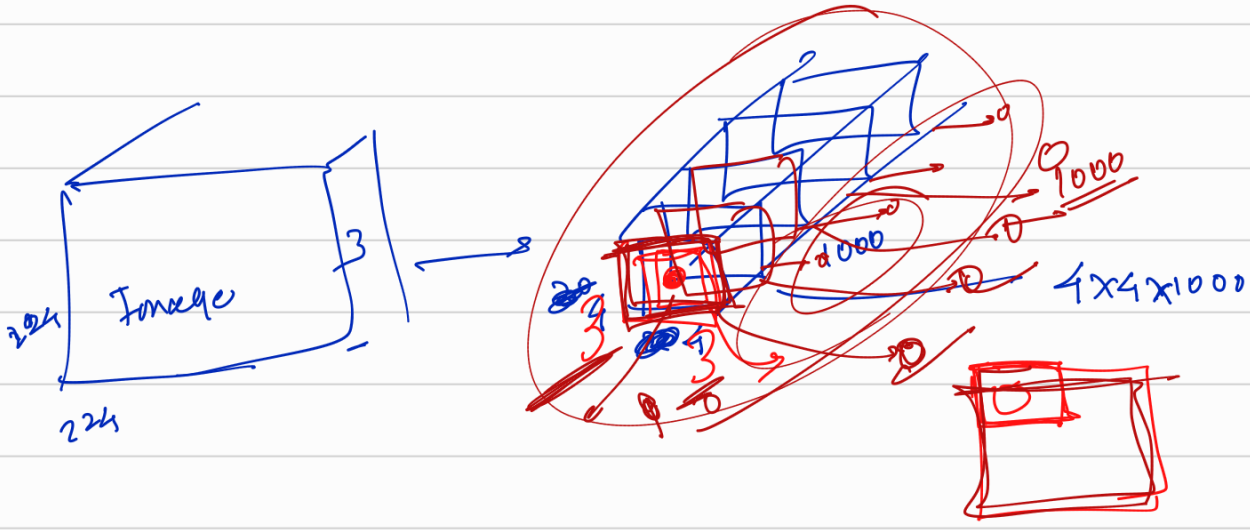
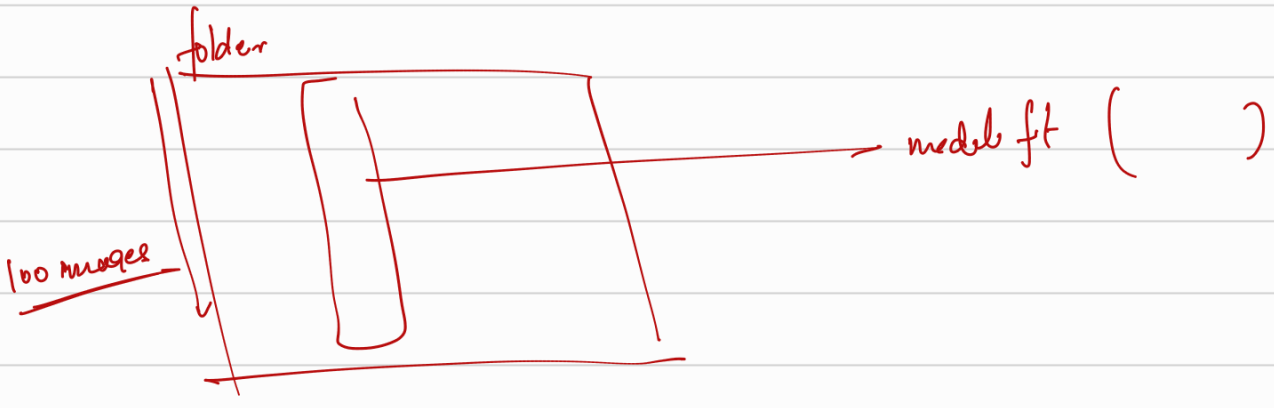
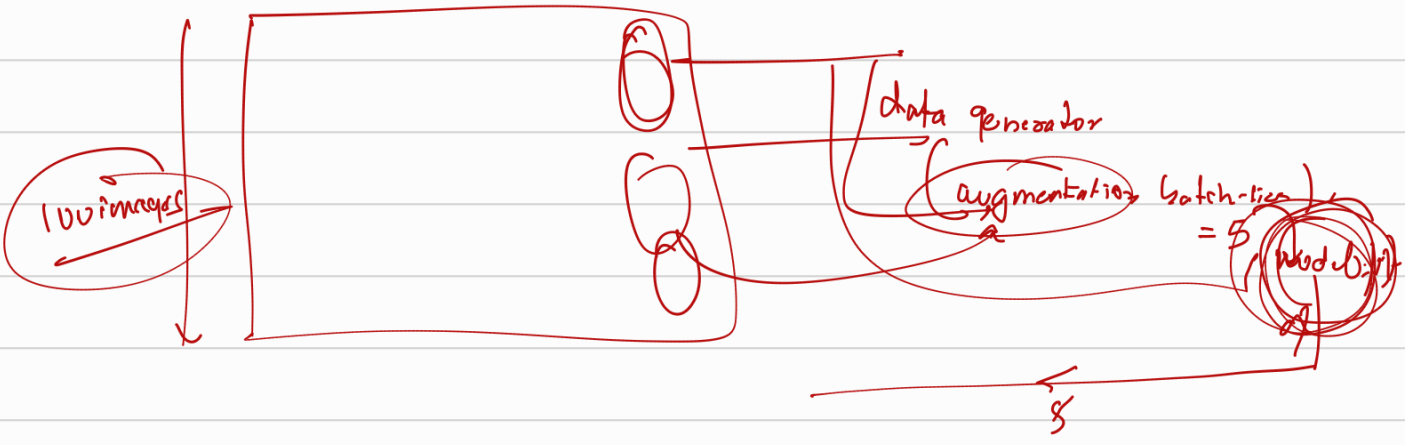
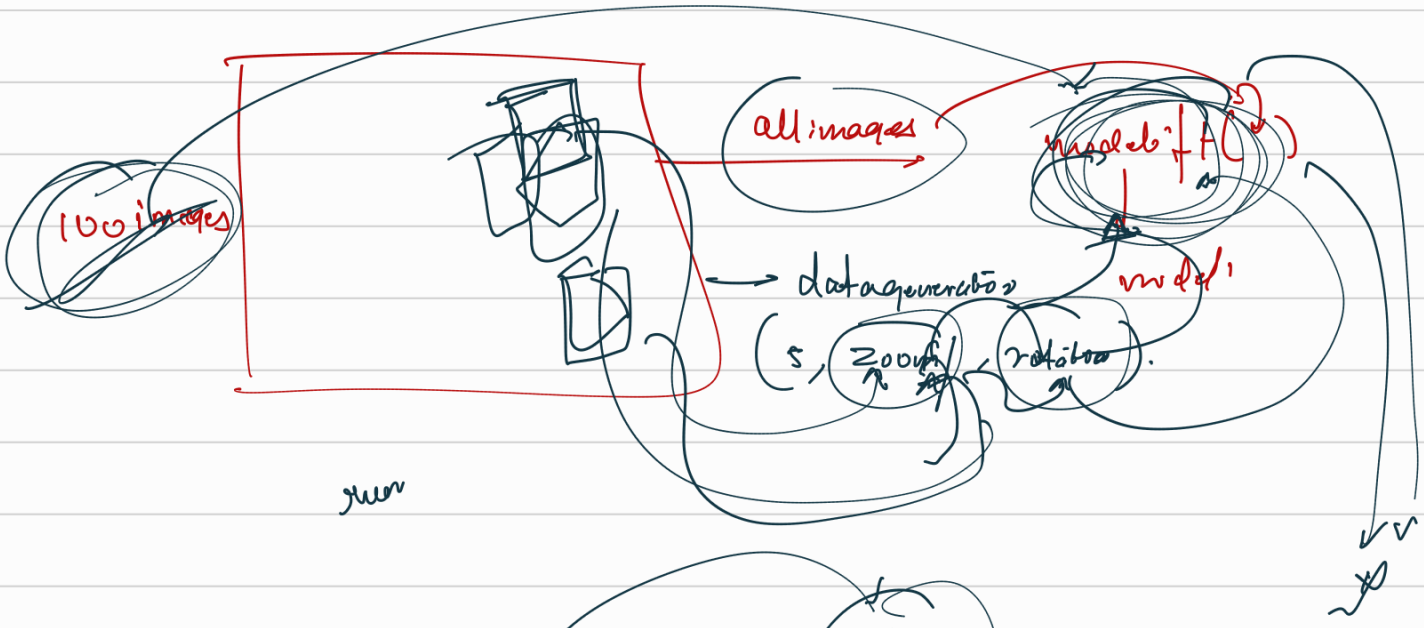
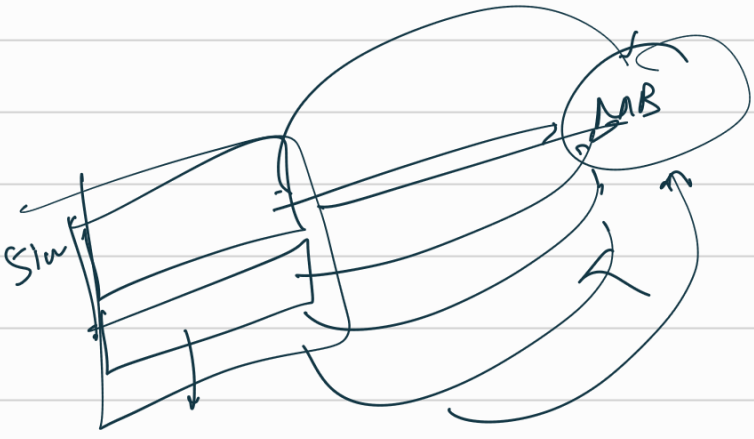


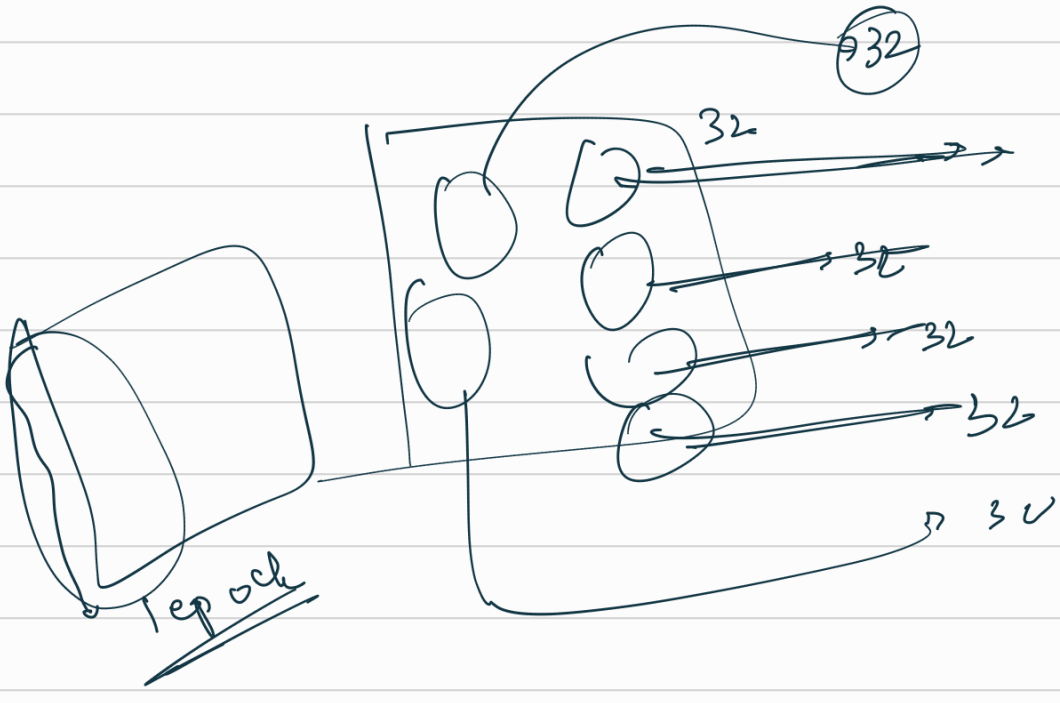
Image
224
224
3
1000
4x4x1000


folder
model.fit ( )
100 images


100 images
data generator
augmentation
batch-size
= 5
model.fit
5


100 images
all images
model.fit()
data augmentation
(5, zoom, rotation)
model
run


MB
Slow


32
32
32
32
32
epoch

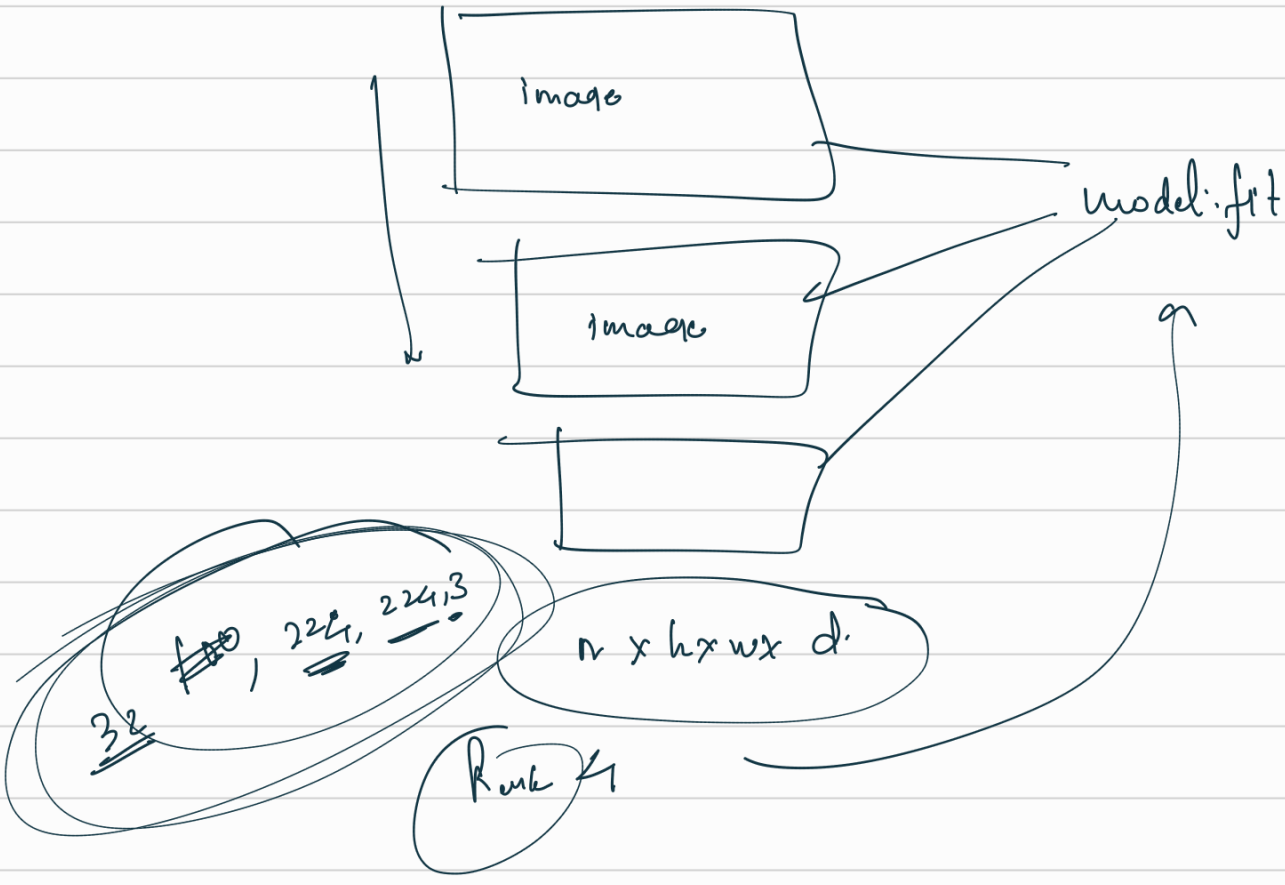
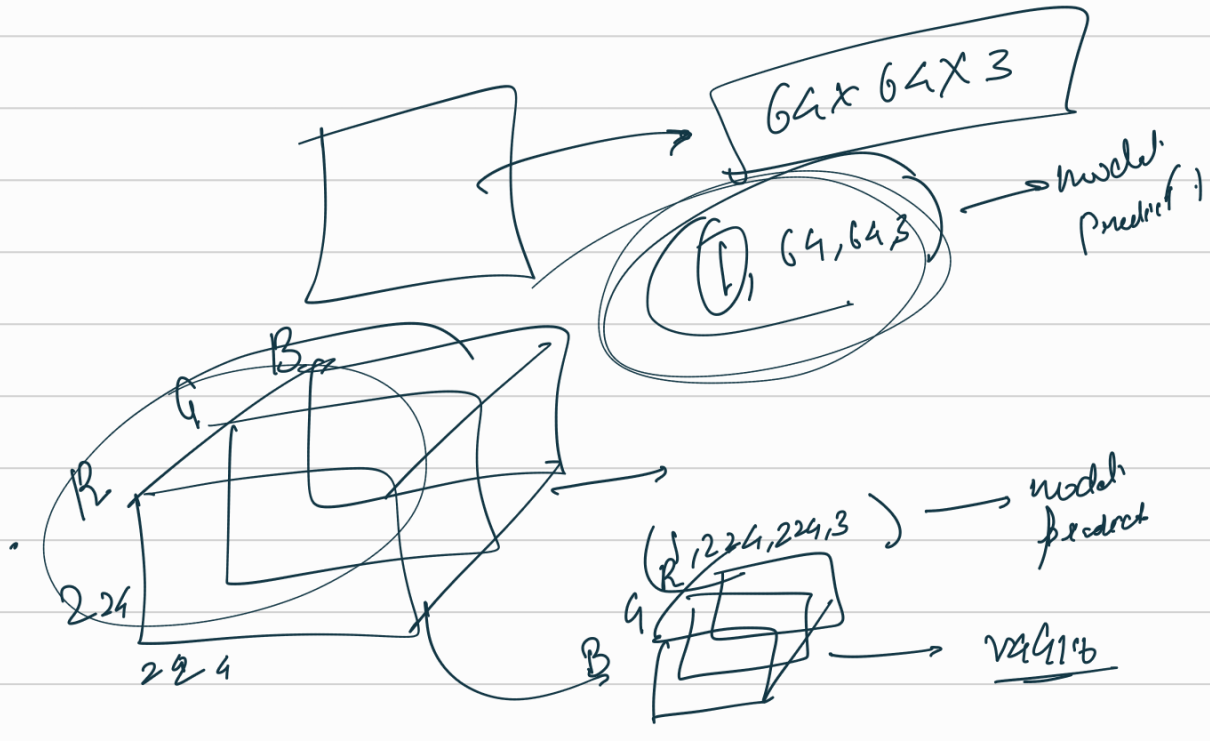
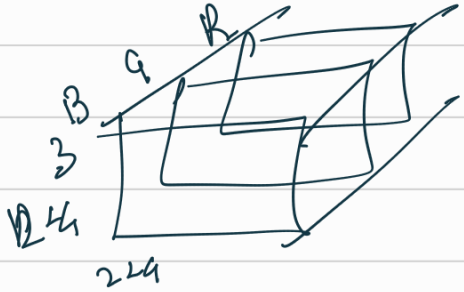
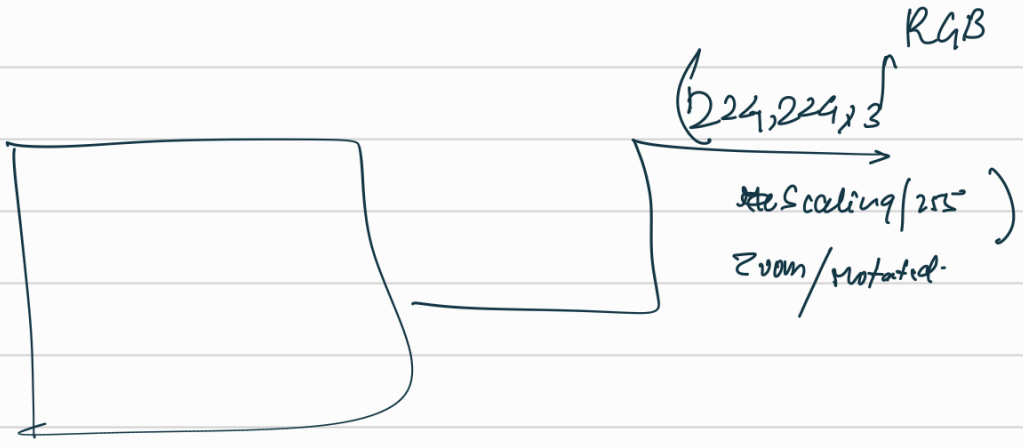


image
model.fit
image
n x h x w x d
Rank 4

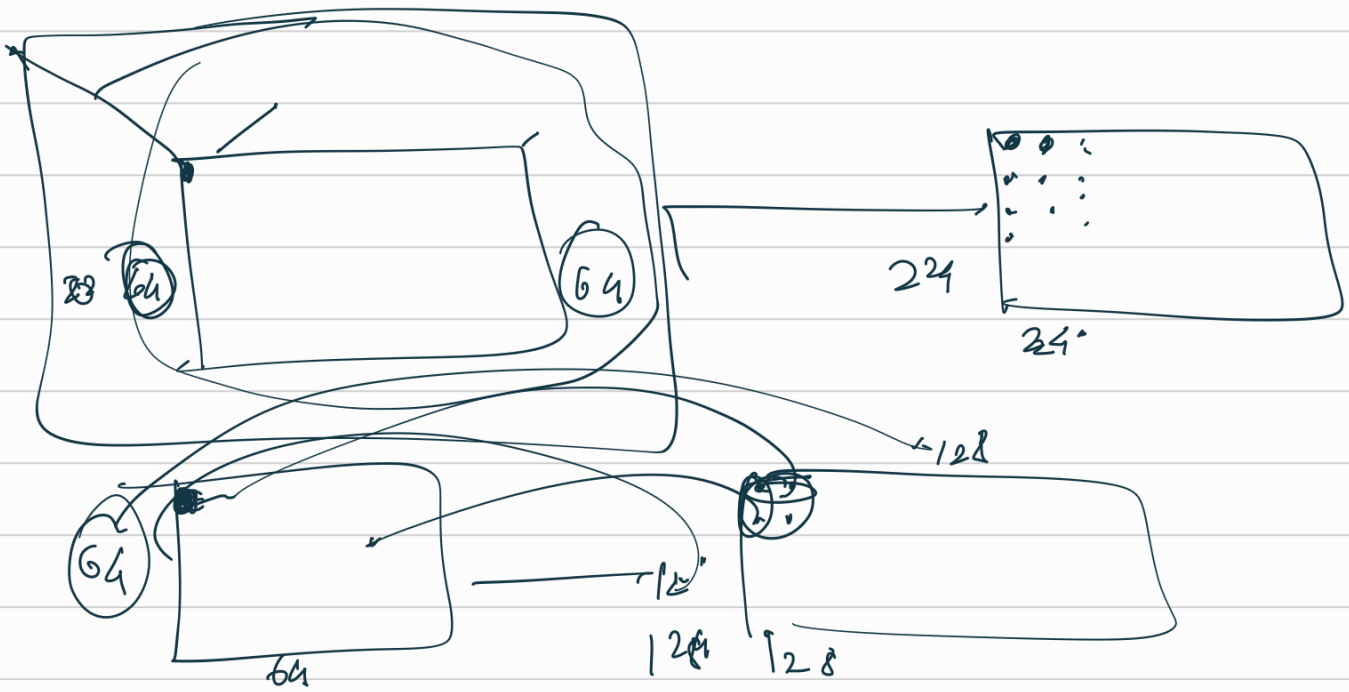
Keras

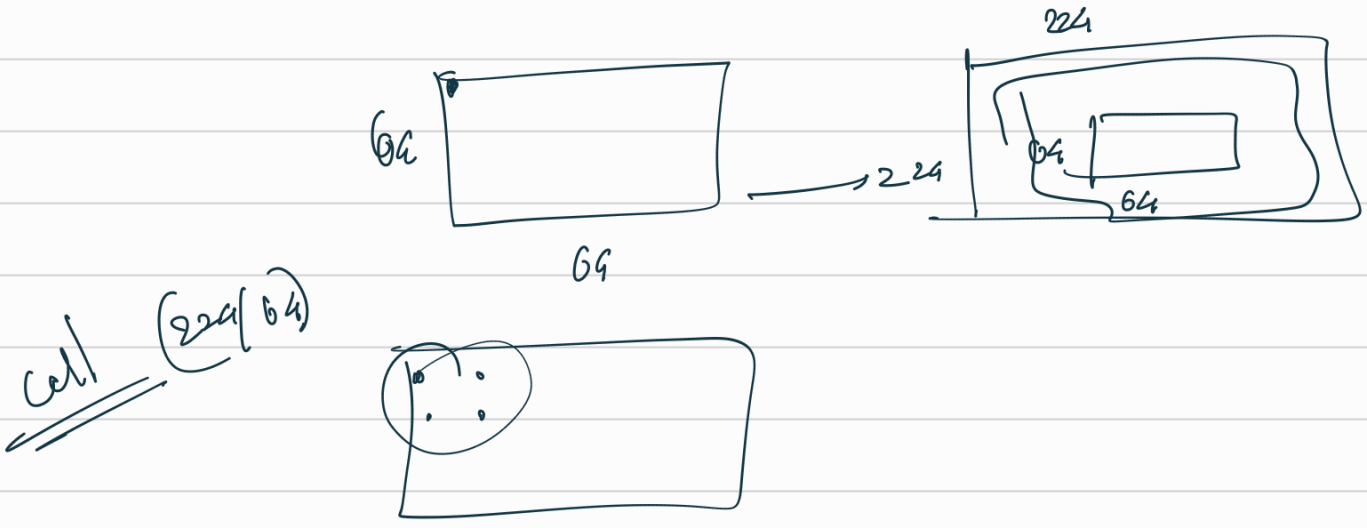

64 x 64 x 3
model.predict()
64, 64, 3
224
224
(R, 224, 224, 3)
model.predict
VGG16

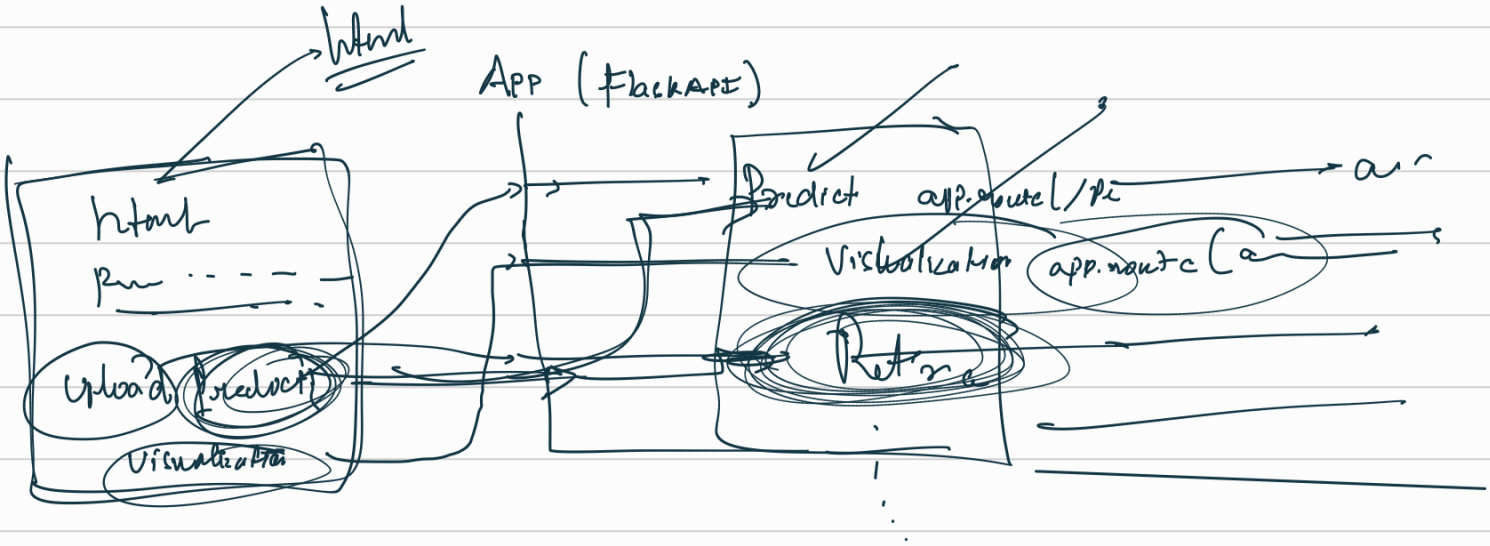

R
G
B
3
224
224


RGB
(224,224,3)
Scaling (/255)
Zoom/rotated.


Resize (1,224,224,3)
Scale (1/255)
any image
RGB → BGR

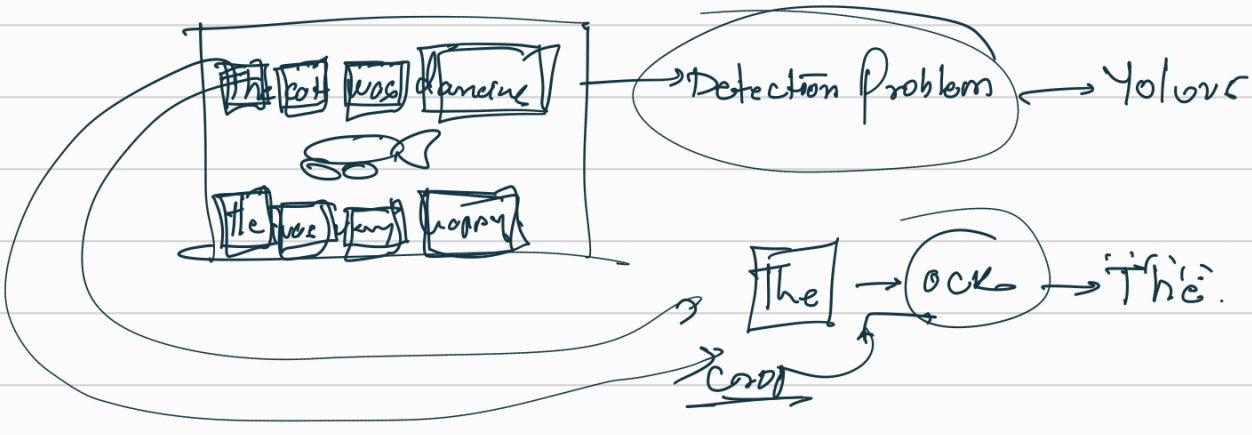

3
64
64
224
224
128
64
64
128
128

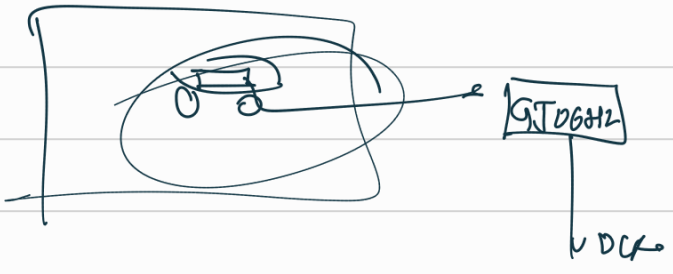

224
64
224
64
64
64
(224, 64)


Html
App (FlaskApi)
Html
Upload
Predict
Visualization
Predict
app.route('/p
Visualization
app.route(
Return

OCR → optical character Recognition

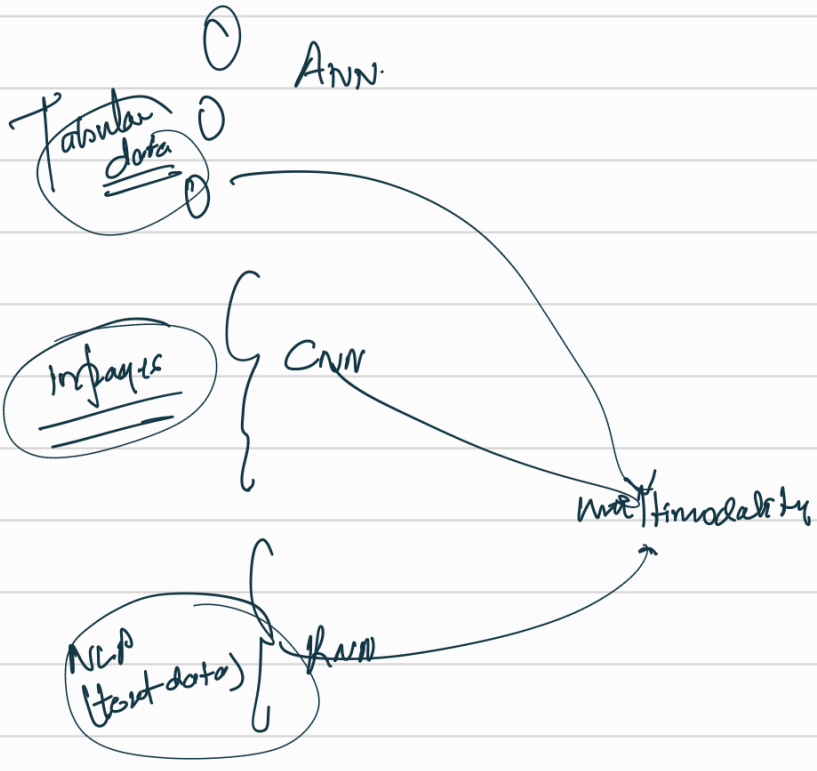
The class is going to be very interesting The class is going to be very interesting we will be great

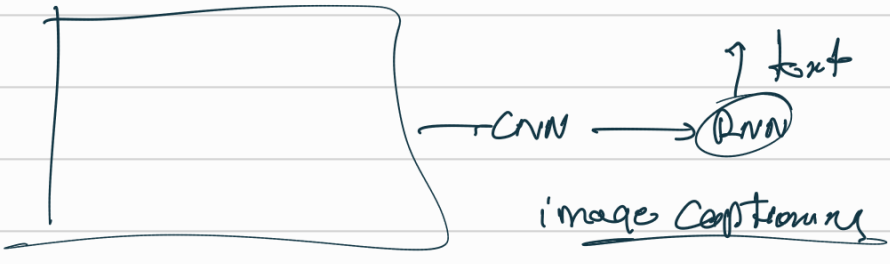


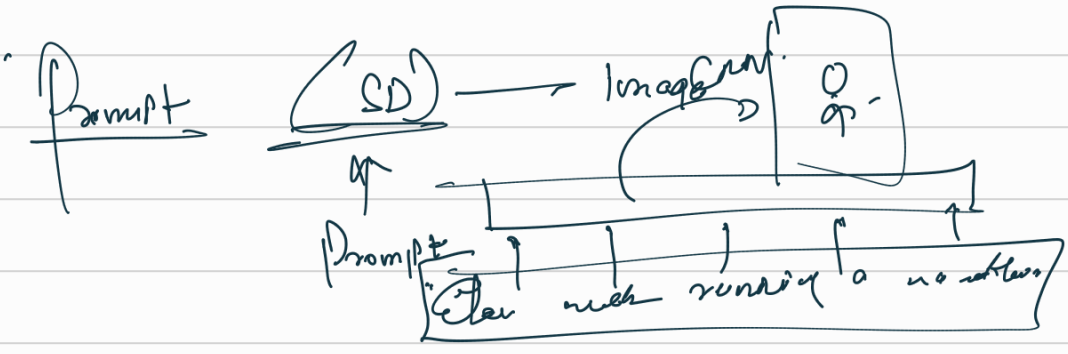


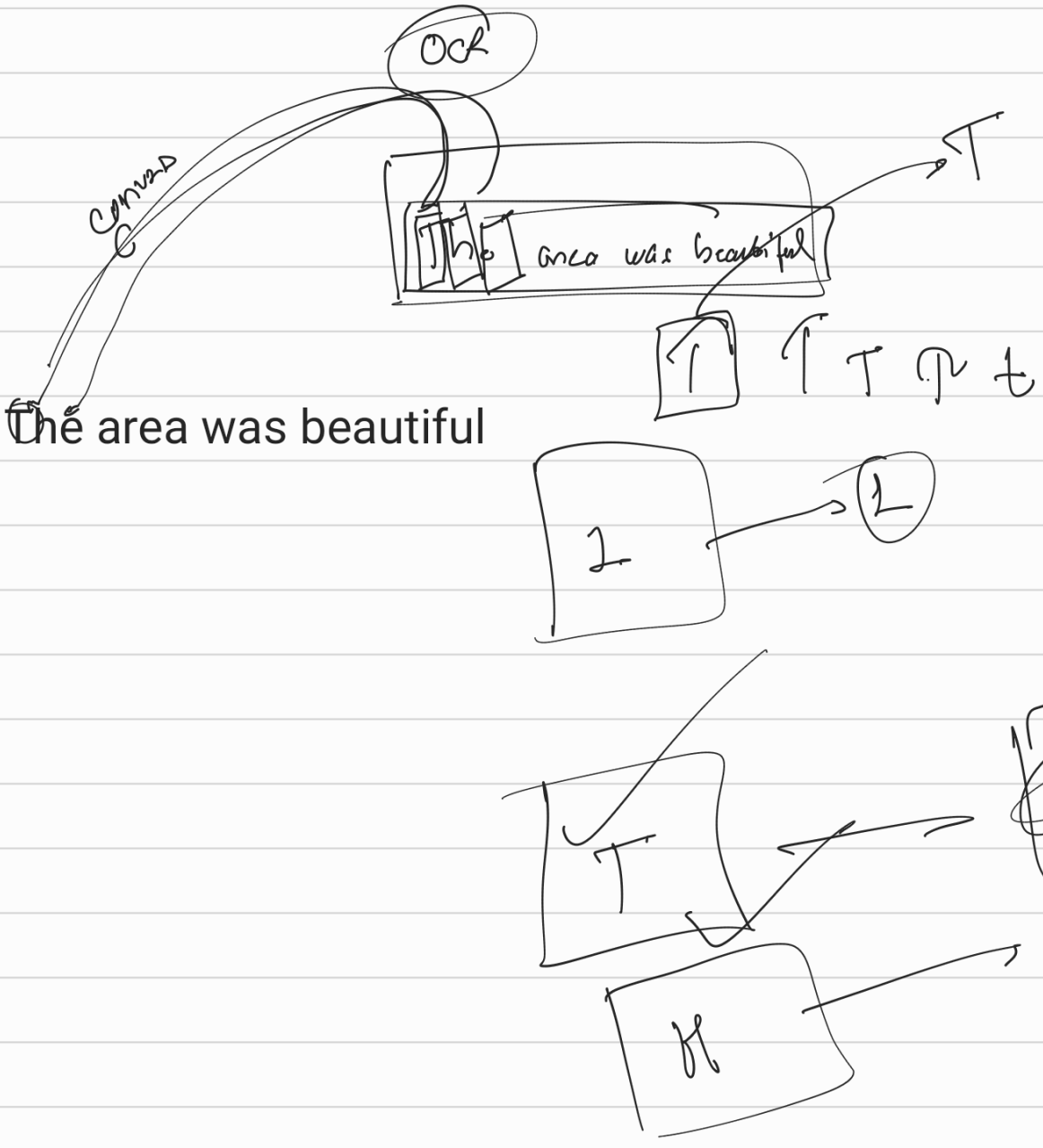
OCR Problem

Architecture to solve an OCR problems.

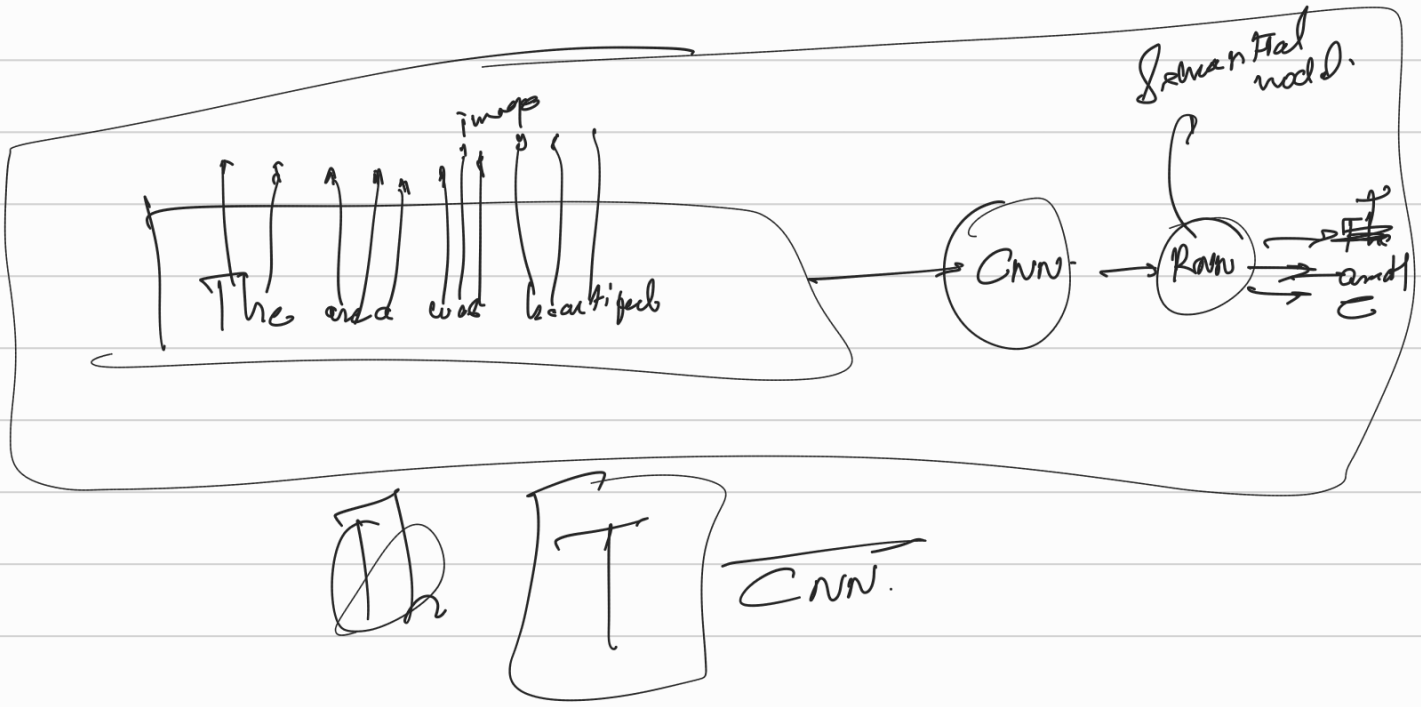


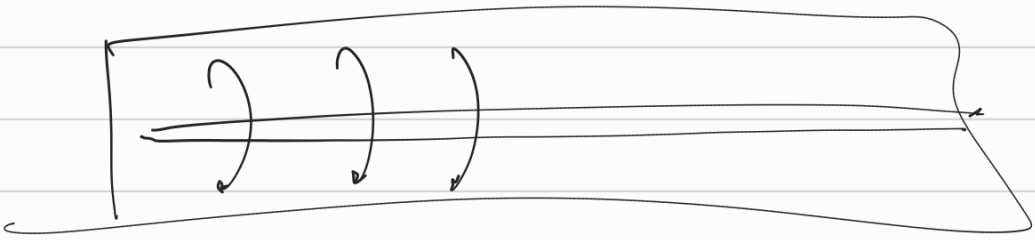


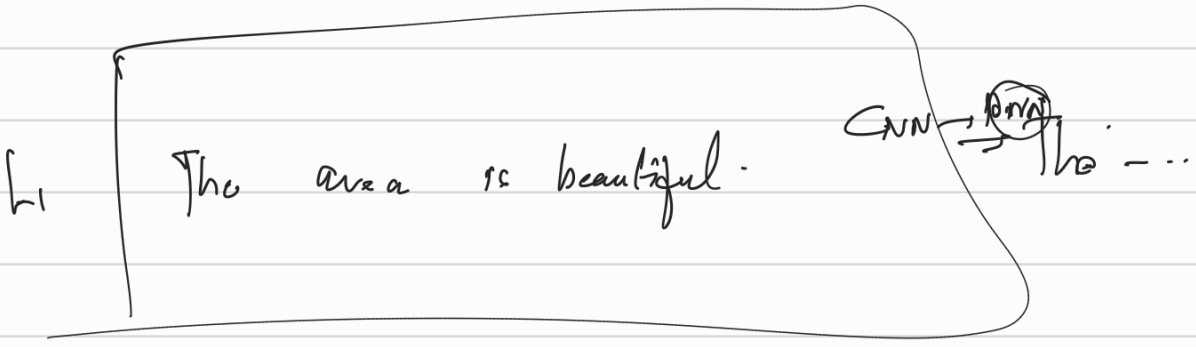




The area was beautiful

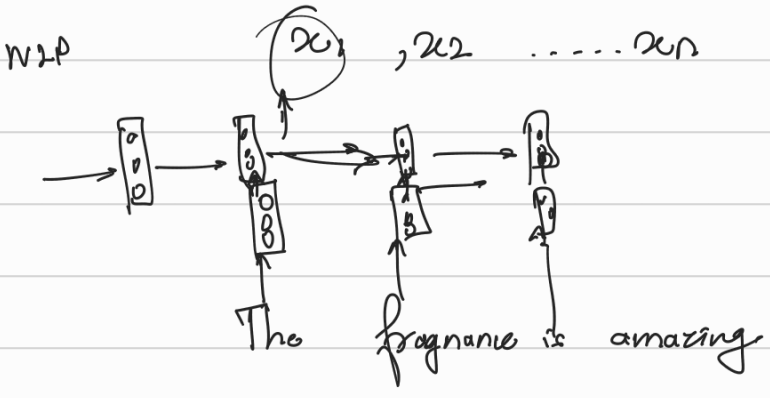


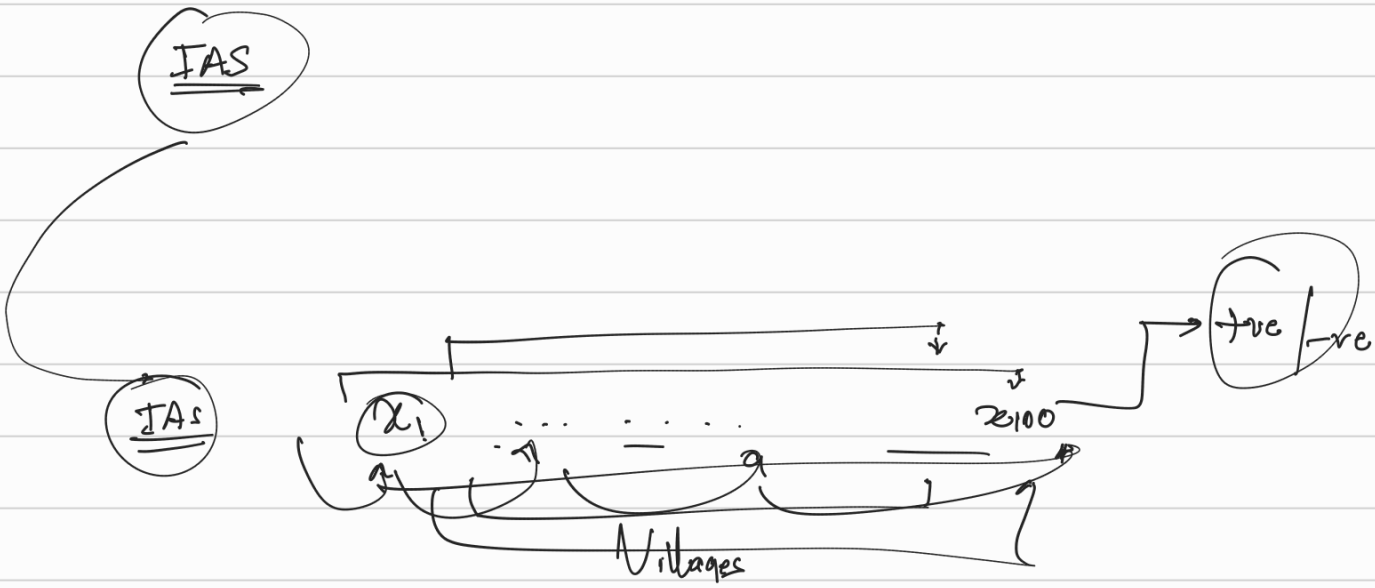


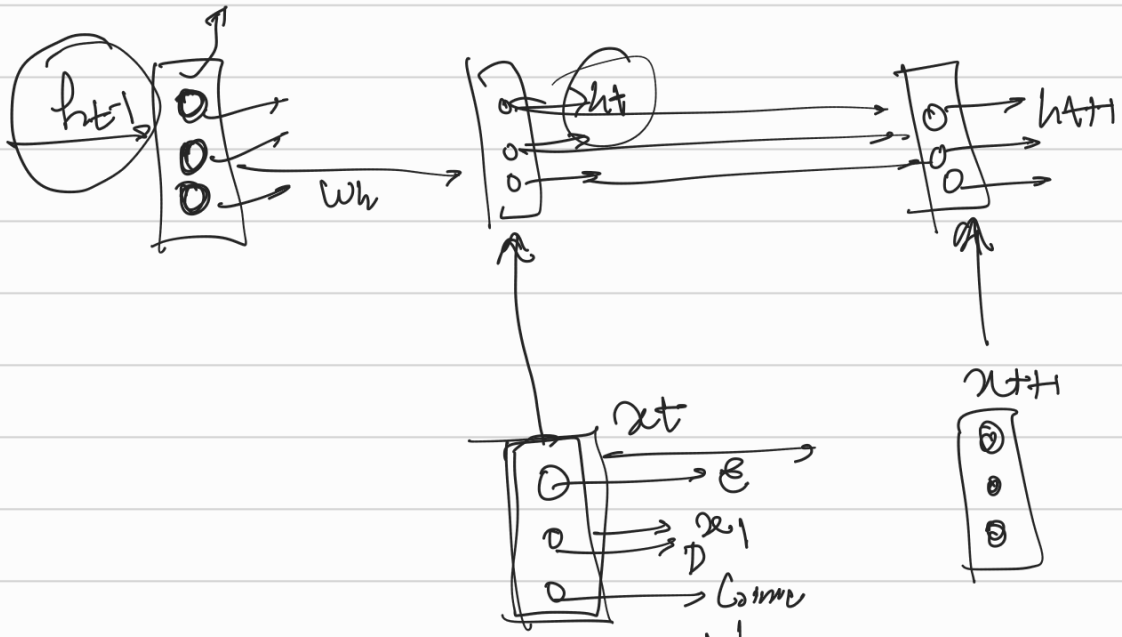


Recurrent neural network

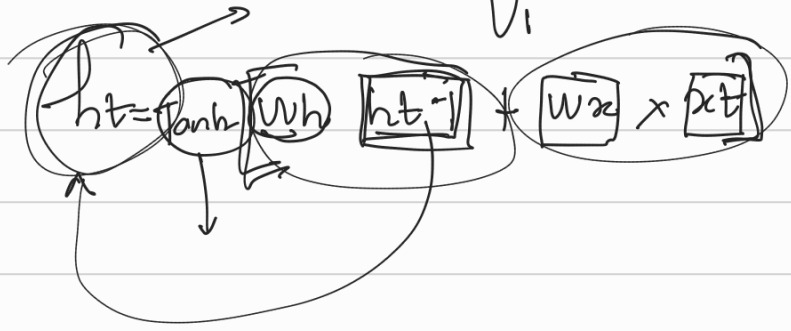
→ Sequential model



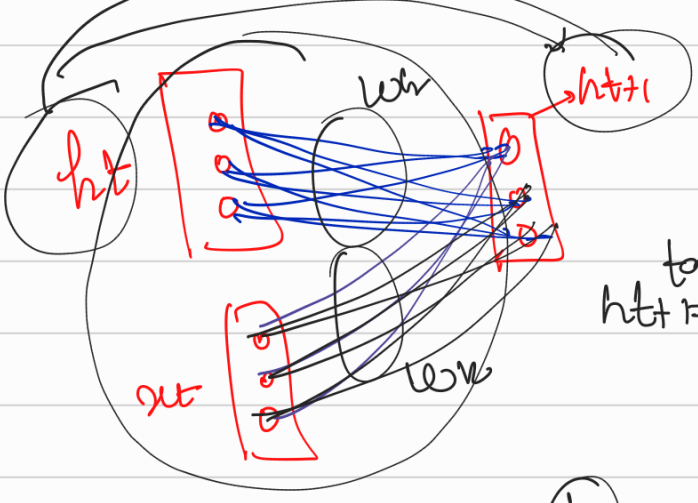




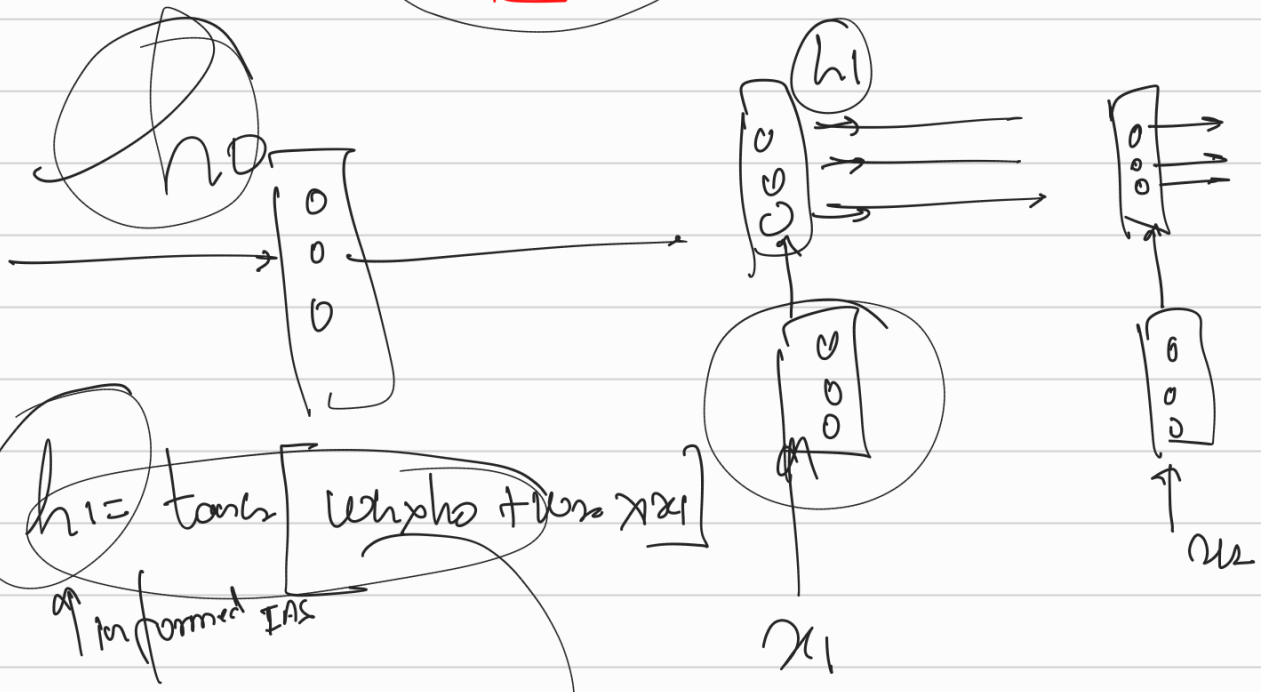
V_1



$$h_{t+1} = \tanh\left[w_h \times h_t + w_x \times x_{t+1} \right]$$



$$h_{t+1} = \tanh\left(W_h(h_t) + W_x x_t\right)$$



$$h_1 = \tanh\left[W_h \times h_0 + W_x \times x_1\right]$$

↑ informed IAS

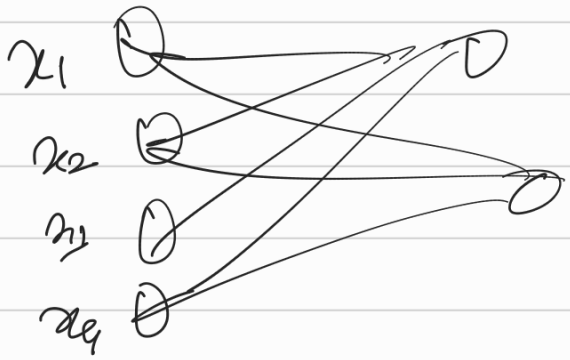
$$h_2 = \tanh\left[W_h \times h_1 + W_x \times x_2\right]$$

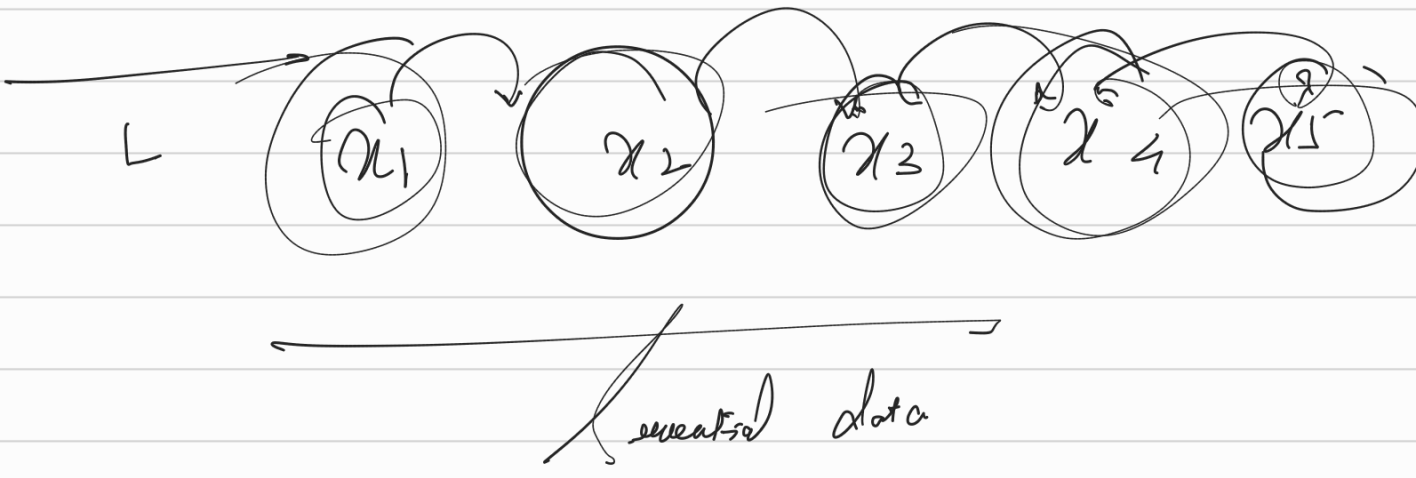
$$h_2 = W_h \times h_1 + W_x x_2$$

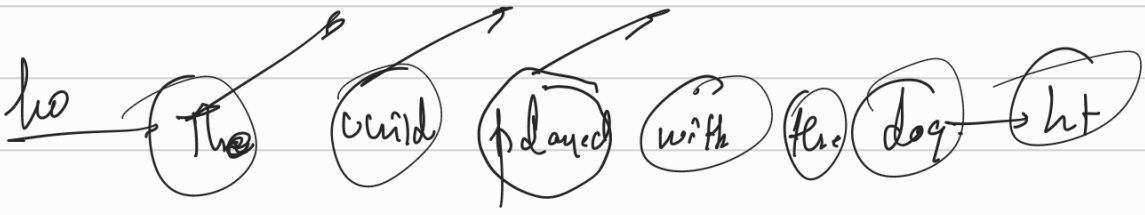
$$h_2 = W_h\left[W_h x_0 + W_x x_1\right] + W_x x_2$$

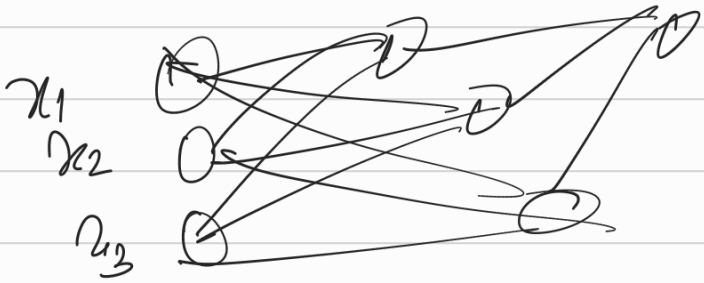
$$h_2 = W_h^2 h_0 + W_h W_x x_1 + W_x x_2$$

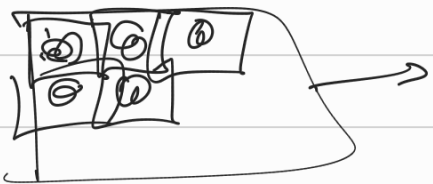
h_t = how information for all the villages from t to 0

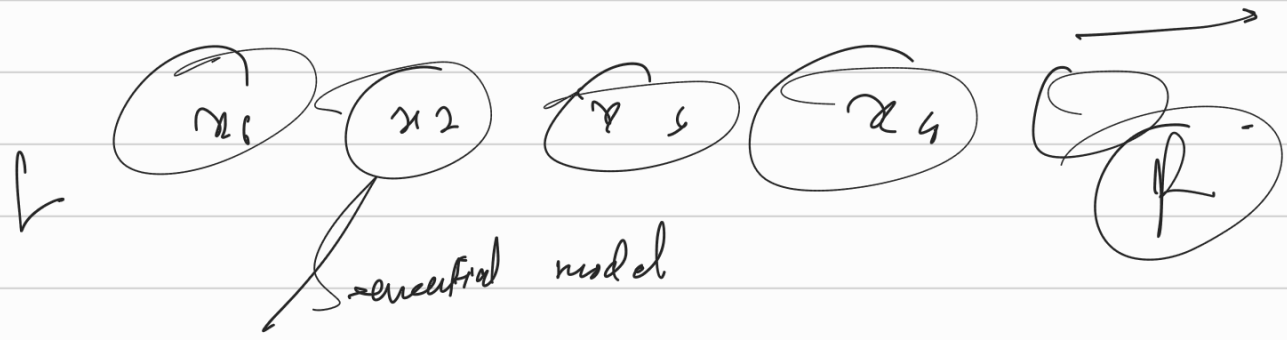

x_1
x_2
x_3
x_4

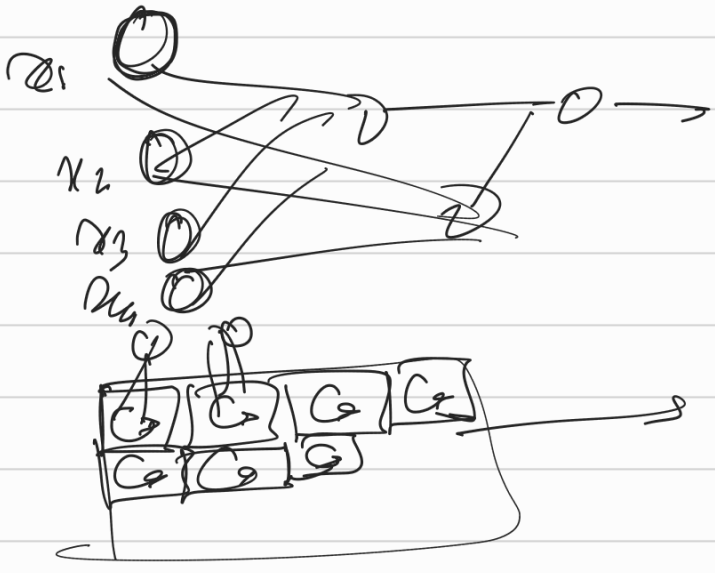

x_1
x_2
x_3
x_4
x_5
created data

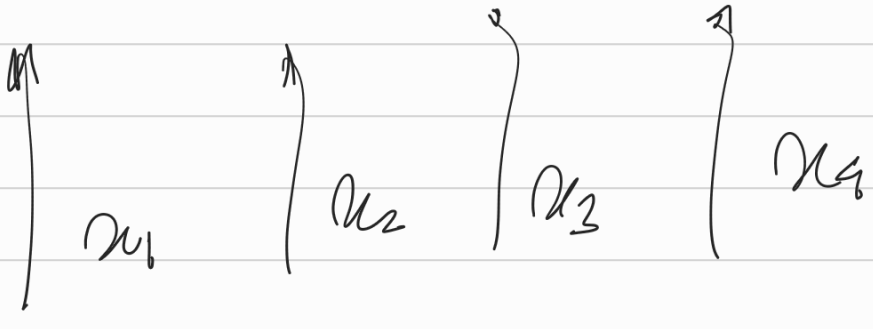

the
child
played
with
the
dog

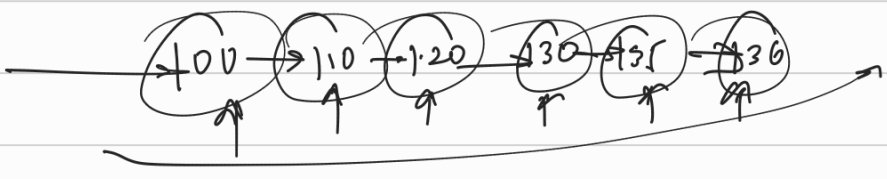

x_1
x_2
x_3

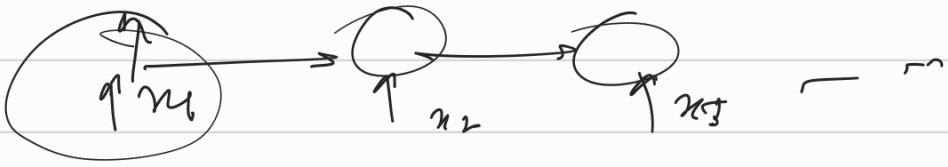


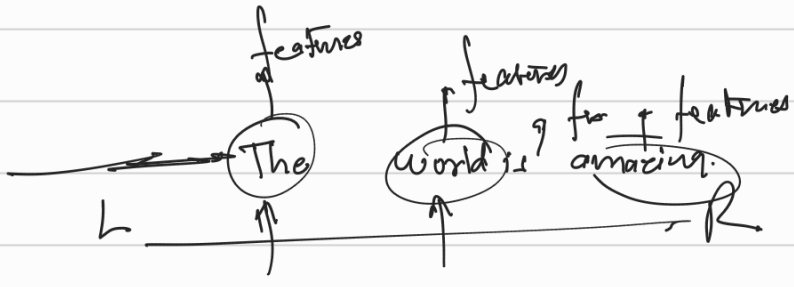

x1
x2
x4
P
sequential model

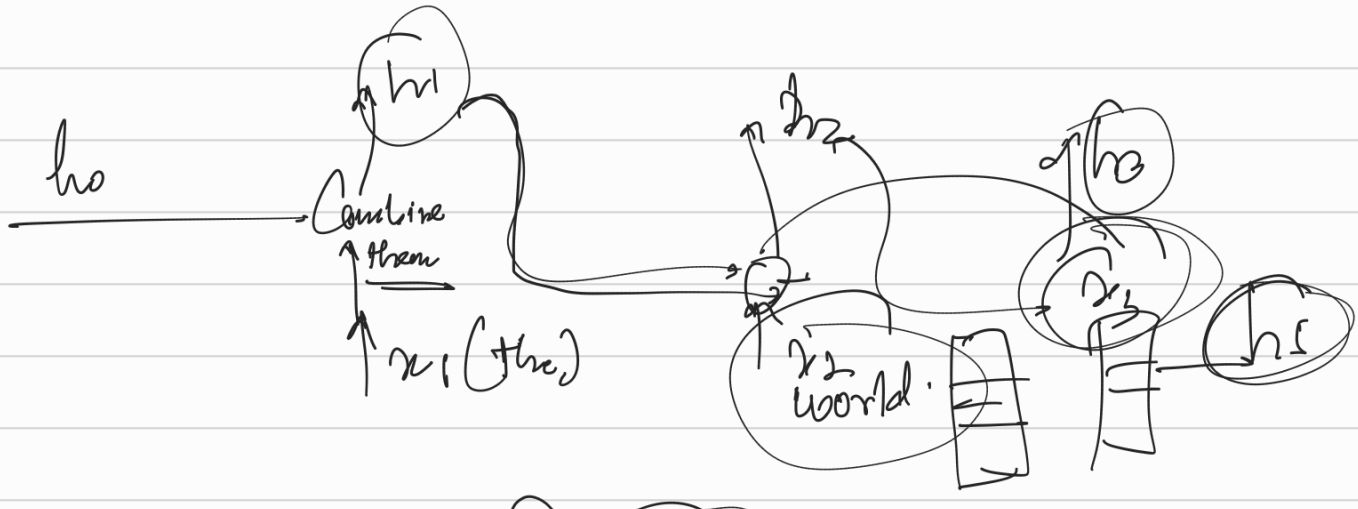

x1
x2
x3
x4


x1
x2
x3
x4

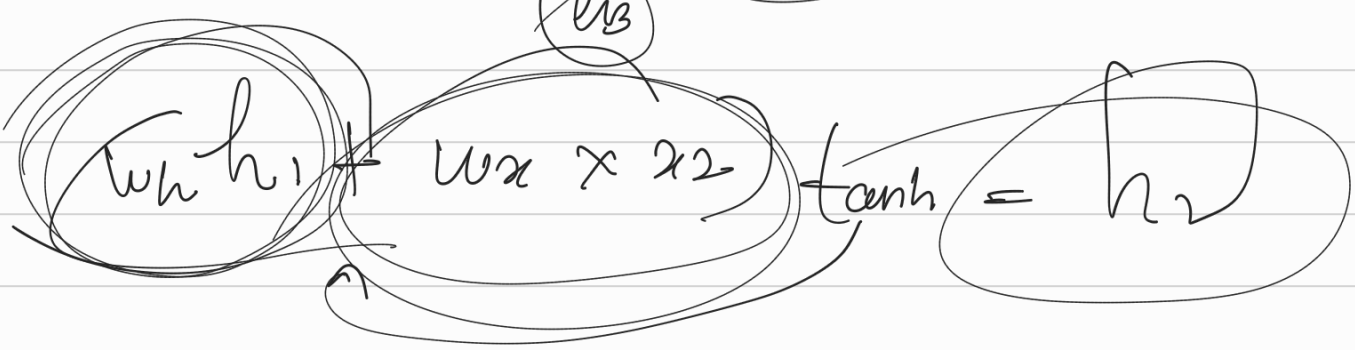

100
110
120
130
135
136

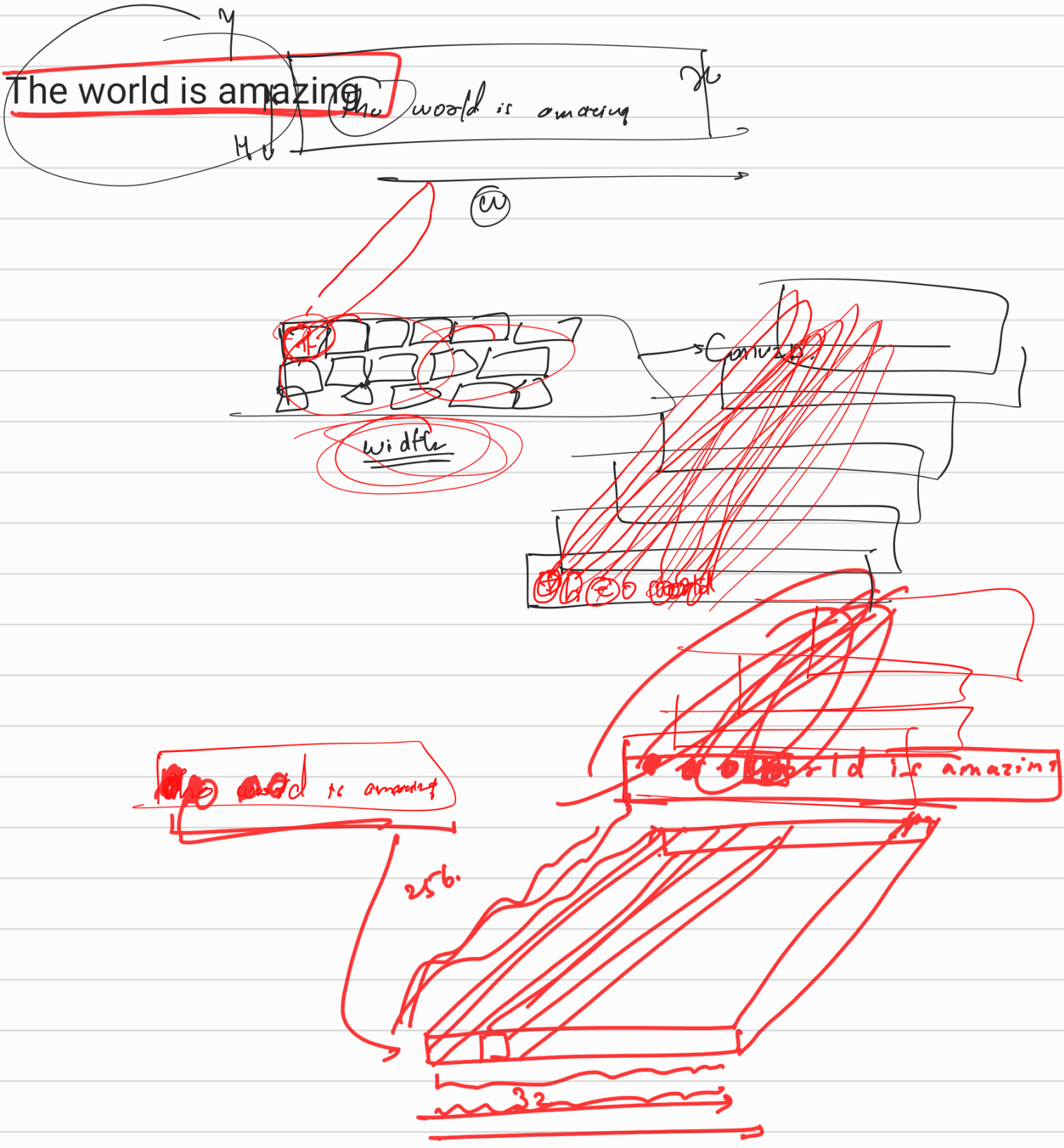




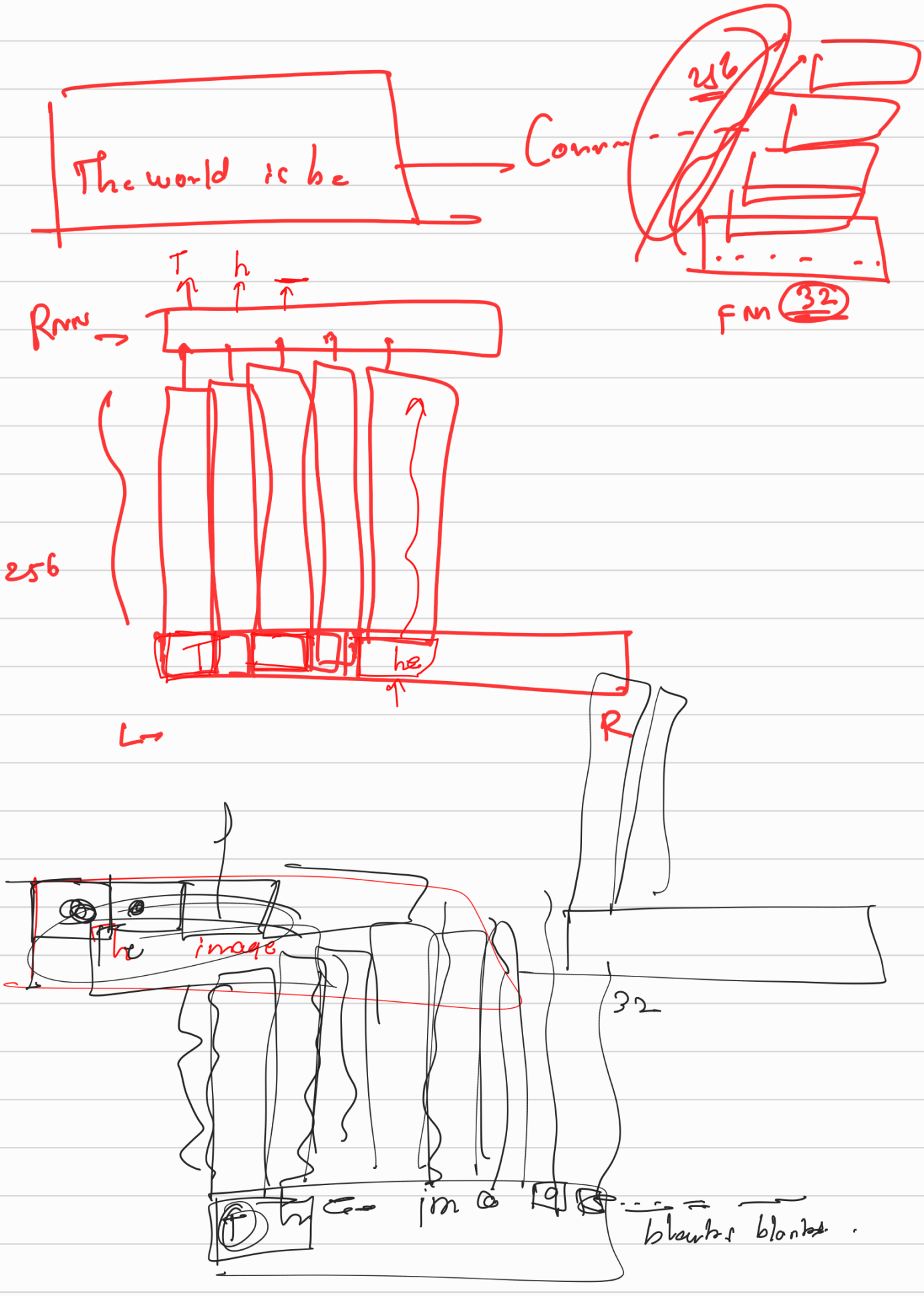


$\tanh(W_{hh} h_0 + W_{hx} x_1)$ $\tanh(W_{hh} h_2 + W_{hx} x_2)$





g have 256 features of 32 characters. (including blank)


The world is be
Conv
256
FM
32
T
h
Rnn
256
he
L
R
the image
32
im
blanks blanks

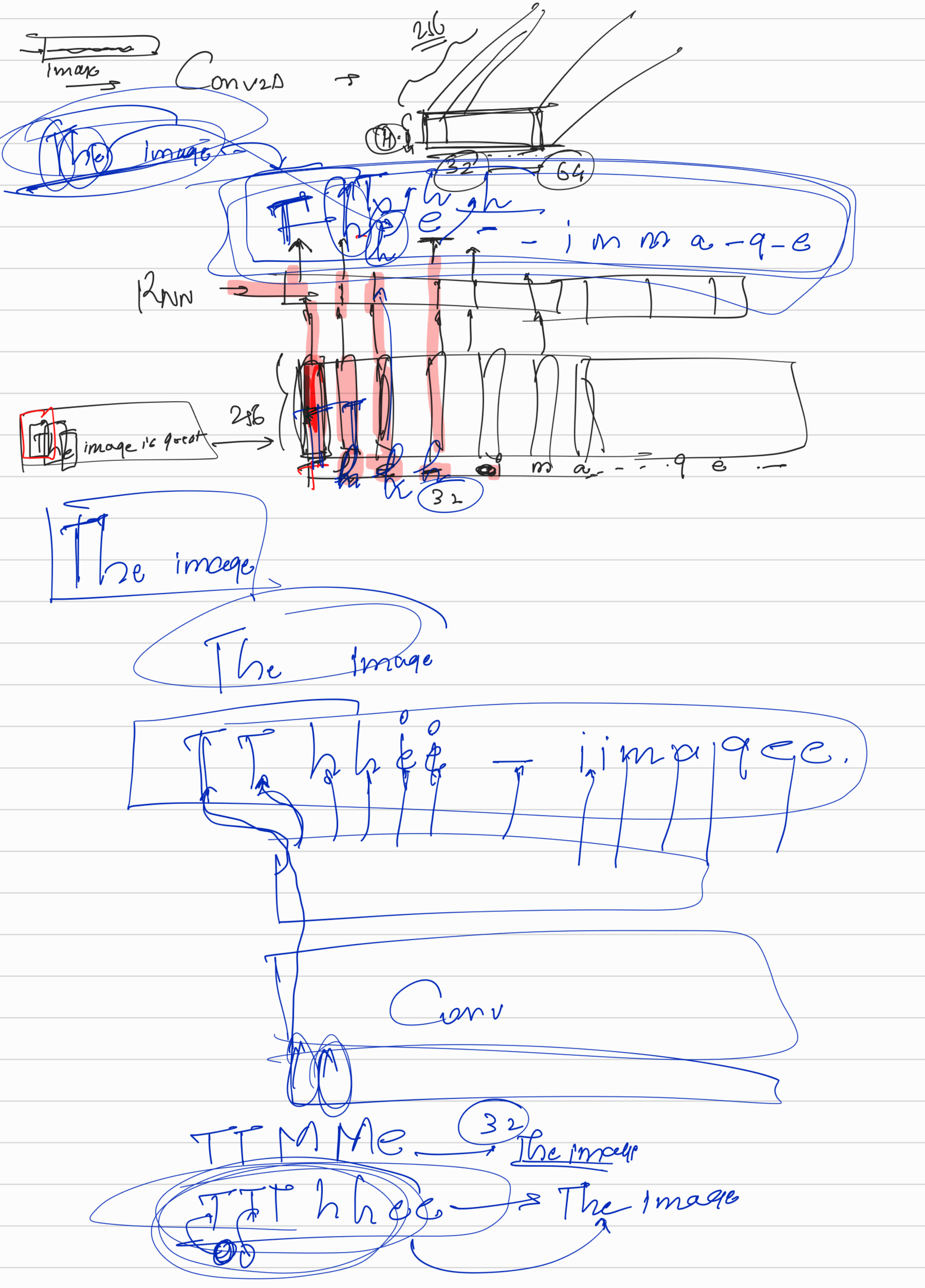
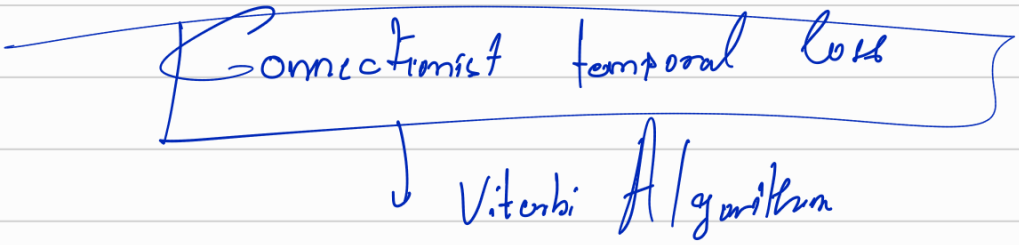
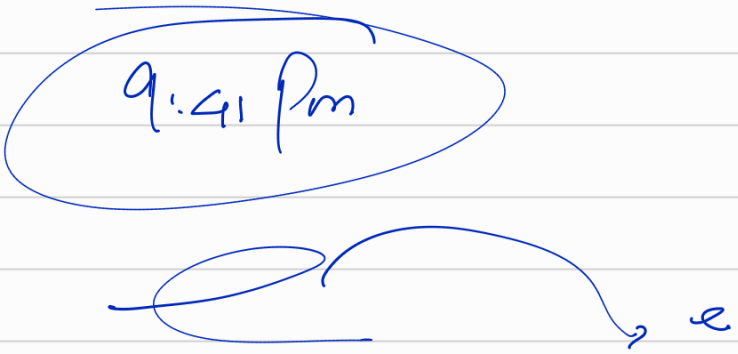


Image
Conv2D
216
The image
32
64
RNN
256
The image is great
32
The image
The image
Conv
32
The image
The image





I am going to eat Burger

I am going to tfalangan o

Jam I am going to Schoo 1 School x... re