[CS550]-Machine Learning **2023-24M**

Questions

Name: Rohit Ashwani RollNo: 12141390 Email: rohitas@iitbhilai.ac.in

Faculty Name: Dr. Gagan Raj Gupta

Question 1

What is the output size after applying 64 convolution filters of size 7x7 with a stride of 1 and no padding to an input image of size 224x224x3? Also calculate the number of parameters , number of computations , and total amount of memory required ?

Solution

Formula for calculating the output size
Output height = $\lfloor \frac{N+2P-F}{S} \rfloor + 1$
where N = Input height
P = Padding height
F = Filter height
S = Stride
Output height = $\lfloor \frac{224+2*0-7}{1} \rfloor + 1$
Output height = 218
Since there are 64 such filters therefore the final output of this convulution operation will be **218x218x64**.

For simplicity we are assuming that there are no biases
Number of parameters
Number of parameters in one filter = (7x7x3) = 147
Number of parameters in 64 filters = 147x64 = 9408

Number of computations
Number of additions due to one filter = Output size x Number of parameters in one filter
Number of additions due to one filter = 218 x 218 x 147

Number of multiplications due to one filter = Output size x Number of parameters in one filter
Number of multiplications due to one filter = 218 x 218 x 147

Total number of computations due to one filter = 218 x 218 x 147 x 2
Total number of computations due to 64 filters = 218 x 218 x 147 x 2 x 64

Assuming each value is stored in float of 4 bytes
Number of gradients = Number of parameters
Total amount of memory required

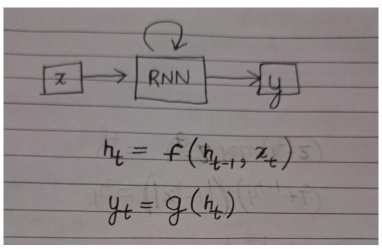
= Memory required for input + Memory required for output + Memory required for parameters + Memory required for gradients
= Input size x 4 + Output size x 4 + Number of parameters x 4 + Number of gradients x 4
= 224 x 224 x 3 x 4 + 218 x 218 x 64 x 4 + 9472 x 4 + 9472 x 4
= 12844032

Question 2

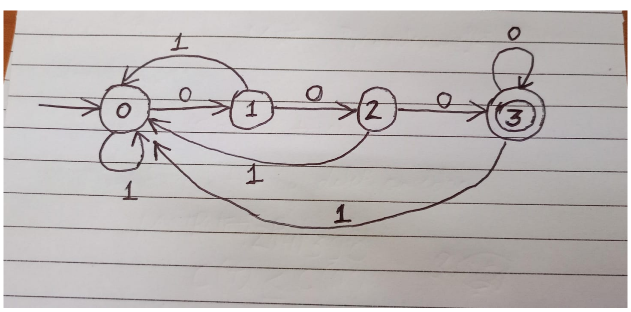
Design an RNN to detect the following pattern in the input
(a) More than two consecutive 0.
(b) More than two consecutive crossing of x_t above a predefined threshold τ within Δ interval.

Solution

A general RNN has the following structure.



(a) Drawing the state machine that will detect more than two consecutive zeros in the input sequence



Converting this state machine in the form of an RNN

Function takes current state and input character and tells the next state

h_{t-1}	x_t	$h_t = f(h_{t-1}, x_t)$
0	0	1
1	0	2
2	0	3
3	0	3
0	1	0
1	1	0
2	1	0
3	1	0

Function takes current state and gives the output character

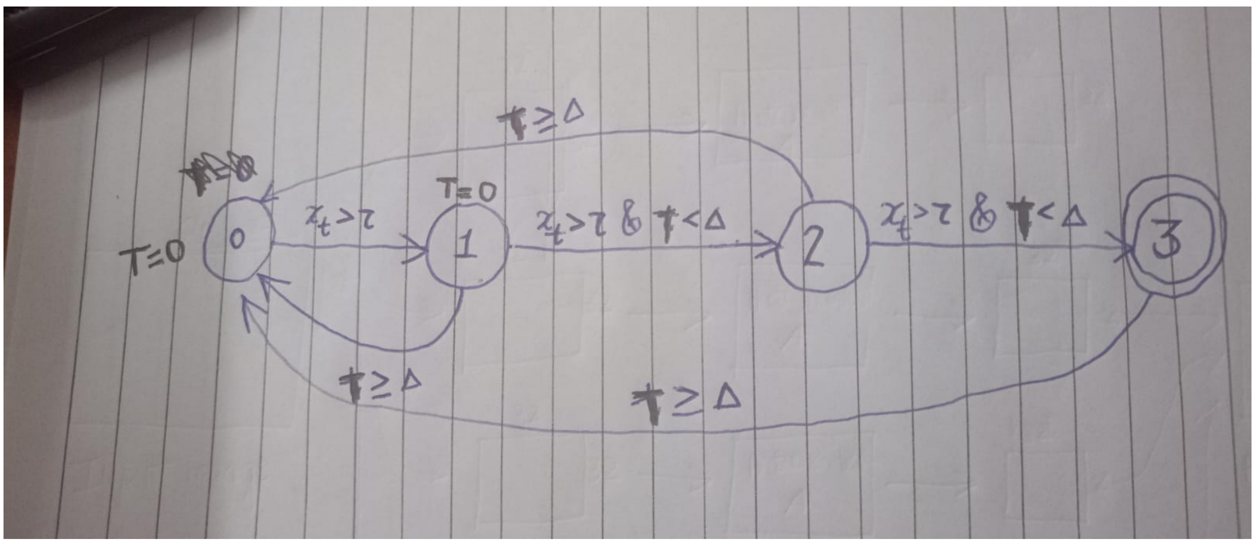
h_t	$y_t = g(h_t)$
0	0
1	0
2	0
3	1

Final RNN
$h_t = f(h_{t-1}, x_t) = min(3, (1 - x_t)(h_{t-1} + 1))$

$y_t = g(h_t) = ReLU(h_t - 2)$

h_{t-1} is the previous hidden state
x_t is the current input character
h_t is the current hidden state
y_t is the output character

(b) Drawing the state machine that detects more than two consecutive crossing of x_t above a predefined threshold τ within Δ interval.



Converting this state machine in the form of an RNN

Function takes current state, input character and current time and tells the next state

h_{t-1}	$a_t = x_t > \tau$	$b = T > \Delta$	$h_t = f(h_{t-1}, x_t)$
0	0	0	0
1	0	0	1
2	0	0	2
3	0	0	3
0	0	1	0
1	0	1	0
2	0	1	0
3	0	1	0
0	1	0	1
1	1	0	2
2	1	0	3
3	1	0	0
0	1	1	0
1	1	1	0
2	1	1	0
3	1	1	0

Function takes current state and gives the output character

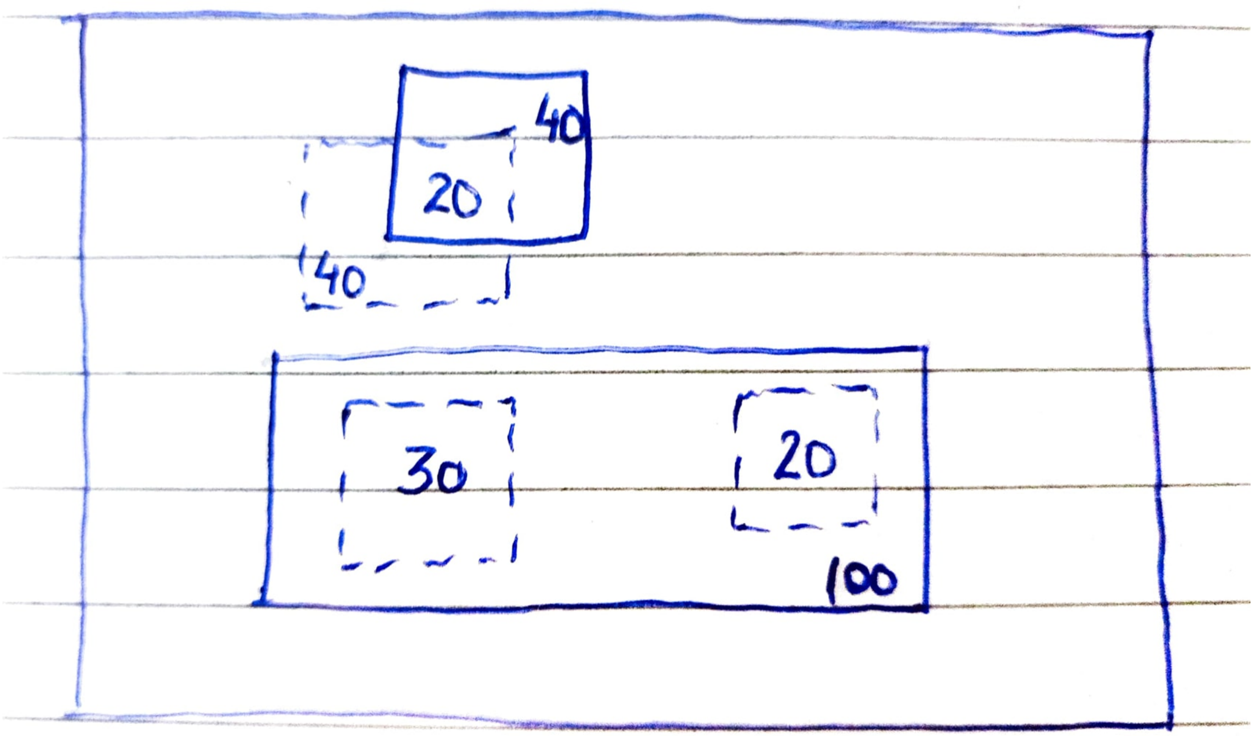
h_t	$y_t = g(h_t)$
0	0
1	0
2	0
3	1

Final RNN
$h_t = f(h_{t-1}, x_t, T) = (1-b)((1-a_t)(h_t) + (a_t(min(h_t+1, 3))))$

$y_t = g(h_t) = ReLU(h_t - 2)$

h_{t-1} is the previous hidden state
x_t is the current input character
h_t is the current hidden state
y_t is the output character
a_t is a boolean telling whether $x_t > \tau or not$
T is the current time (setted to zero after reaching state 0 and state 1 every time.)
b is a boolean telling whether the $T > \Delta$ or not

Question 3



Find the IoU for each predictions and also calculate the mAP. Provided that there is only one class

For the first prediction:
IoU = $\frac{AreaofUnion}{AreaofIntersection} = \frac{20}{40+40-20} = \frac{20}{60} = 0.333$

For the second prediction:
IoU = $\frac{AreaofUnion}{AreaofIntesection} = \frac{30}{100} = 0.3$

For the third prediction:
IoU = $\frac{AreaofUnion}{AreaofIntersection} = \frac{20}{100} = 0.2$

Calculating mAP

Since all the predictions have IoU values less that 0.5 therefore all of them will be marked as negative.
So when we will plot the Precision versus Recall curve it will be on X-axis as the precision will be 0 and recall will also be 0. So the average precision = Area under the Precision-Recall curve = 0.

Since there is only one class, mAP = Average Preision of this class = 0.

Question 4

A is an autoencoder for input image of size 32x32. Design an autoencoder for input image of size of 64x64 using A.

Solution

Task: Transform a 64x64 image using an existing autoencoder A built for 32x32 images.

Approach:
Splitting: Divide the 64x64 image into four 32x32 regions.
Usage of Autoencoder A: Apply the 32x32 autoencoder A to each of these smaller regions separately.
Encoding: Get encodings for each of the four regions using A's encoding part.
Combination: Combine these four smaller region encodings to create a final encoding for the entire 64x64 image.

Image explaining the flow