Ques No 1

Input Layer:-

Input Shape : (16, 16, 4)

Activation Size : 16x16x4 = 1024

Trainable Parameters : 0

Convolution layer 1:-

no Filters = 10

Filter size = 5

Stride = 1

valid padding

width, height = $(16-5+2\times0)/1 + 1 = 12$

Activation Shape: (12, 12, 10).

Activation Size = 12x12x10 = 1440

Trainable Parameters = $(5\times5\times4+1)\times10$

= 1010

Pooling Layer 1:-

Filter size : 2

Stride : 2

width, height = $(12-2+2\times0)/2 + 1 = 6$

Activation Shape: 6x6x10

Activation Size = 360

Trainable Parameters = 0

Convolution Layer 2

no Filters = 16

Filter size = 3

Stride = 1

Activation shape = (6, 6, 16)

Activation size = 576

→ Depth wise : $(3 \times 3 \times 1) \times 1$

→ Point wise : $1 \times 1 \times 16 \times 16 + 16$

Trainable Parameters = $(3 \times 3 \times 1 + 1) \times 16 + (1 \times 1 \times 16 \times 16 + 16)$

= 4544

Pooling Layer 2:-

Filter size = 2

Stride = 2

width, height : $(6 - 2 + 2 \times 0)/2 + 1 = 3$

Activation shape = (3, 3, 16)

Activation size : $3 \times 3 \times 16 = 144$

Fully Connected Layer 3:-

no Neuron = 100

Activation shape = (100,)

Activation size = 100

Trainable Parameters = $(3 \times 3 \times 16 \times 100) + 100$

= 43300

Fully Connected Layer 4:

no of neuron : 10

Activation shape = (10,)

Activation size : 10

Trainable Parameters : (10x100)+10

= 1010

Output Layer :-

no of neurons : 5

Activation shape : (5,)

Activation size = 5

Trainable Parameters : (10x5)+5

= 55

QUES No 2

Layer	Activation Shape	Activation Size	Trainable Para
Input	(8,8,3)	192	0
Conv1	(3,3,2)	72	19
Pool1	(3,3,2)	18	0
Conv2	(3,3,1)	9	10
Pool2	(1,1,1)	1	0
FC3	(2,1)	2	4
Output	(1,1)	1	2

Cost / Error

$$J(\theta) = -[1 \times \log(0.5) + (1-1) \times \log(1-0.5)]$$

$$= 0.693$$

Backpropagate:-

$$\Delta J(w\text{-}out) = (a-r) \times a\text{-}in$$

$$a\text{-}in = f(\Sigma(w\text{-}fc \times a\text{-}conv))$$

$$a\text{-}in = max(0, \Sigma(w\text{-}fc \times a\text{-}conv))$$

$$\Delta J(w\text{-}out) = (0.5-1) \times max(0, \Sigma(w\text{-}fc \times a\text{-}conv))$$

Convolution Layer 2

Fully Connected:

$$\Delta J(w\text{-}fc) = (w\text{-}out^T \times \Delta J(w\text{-}out)) \times a\text{-}conv$$

Convolution Layers.

$$\Delta J(w\text{-}conv) = \Sigma(\Delta J(w\text{-}fc) \times w\text{-}fc^T \times f'(a\,conv) \times a\text{-}prev$$

$$\Delta J(w\text{-}conv) = \Sigma(\nabla J(w\text{-}fc) \times w\text{-}fc^T) \times 1(a\text{-}conv > 0) \times a\,prev$$

$$w = w - \alpha \times \Delta J(w)$$

for

w-out = 0.5

w-fc = [[0.1, 0.2], [0.3, 0.4]]

a-conv = 0.6, 0.7

α = 0.1

$\Delta J(w\text{-}out) = -0.25$

$\Delta J(w\text{-}fc) = [[-0.15, -0.175]]$

$\Delta J(w\text{-}conv) = [[0.33, 0.39]]$

w-out: $[[0.5] - 0.1 \times -0.25 = 0.525$

w-fc: $[0.1, 0.2], [0.3, 0.4] - 0.1 \times [[-0.15, -0.175]$

$= [[0.115, 0.225], [0.315, 0.425]]$

w-conv: $[0.8, 0.9] - 0.1 \times [0.33, 0.39]$

$= [0.47, 0.51]$

The table for activation shape, size is the same.

$$a\text{-}out = f\left(\Sigma\left(w\text{-}out \times a\text{-}fc\right) + (v)\right)$$

$$= f\left(\Sigma\left(w\text{-}out \times f\left[\Sigma\left(w_{conv} \times a\text{-}prev\right)\right]\right)\right)$$

$$a_{out} = f\left(0.525 \times f\left(0.115\, a\text{-}prev1 + 0.225\, a\text{-}prev2 + 0.375\, a_{prev3}\right.\right.$$

$$\left.\left. + 0.125\, a_{prev4}\right)\right)$$

The output will be closer to 1 than 0.5.

5) Calculate cost/error value using the cost function

for $a_{out} = 0.9$

$$J(d) = -\log(a_{out})$$

$$= -\log(0.9)$$

$$= 0.105$$

6) Compare Error

Error has reduced from 0.693 to 0.105.

7) What is predicted output.

0.9. Predicted output is 0.9.