COMSATS University Islamabad, Attock Campus

Department of Computer Science

Program: BS(AI)

Fall 2023: Assignment 2		**Course: Deep Learning (AIC467)**
Dated: 18/10/2023	**Due Dated: 24/10/2023**	**Marks: 20**
Name:		
Note: - Don't write anything on Question Paper except your Name & Reg. No.		

The student will submit the hard copy to CR before the due timing. The CR is responsible for submitting the class assignments till the due date in the office.

Question#1 [CLO-2(SO:2-4)] [10 Marks]

Find the Activation Shape (Output Shape), Activation Size (Output Size) and the Number of Trainable Parameters for the following CNN model.

Note: Two types of convolution layers are used in the example: 1) depth-wise not separable Conv2D(), *and 2) depth-wise separable* SeparableConv2D().

Hyper Parameters	Activation Shape	Activation Size	Number of Trainable Parameters
Input(16,16,4)	(16,16,4)	1024	0
ConvolutionLayer1_Conv2D(number of filters=10, filter size=5, stride=1 padding=null)	?	?	?
PoolingLayer1(filter size=2, Stride=2, Padding=null)	?	?	?
ConvolutionLayer2_SeparableConv2D(number of filters=16, filter size=3, stride=1 padding=yes)	?	?	?
PoolingLayer2(filter size=2, Stride=2, Padding=null)	?	?	?
FullyConnectedLayer3(number of neurons=100)	?	?	?
FullyConnectedLayer4(number of neurons=10)	?	?	?
OutputLayer(number of neurons=5, activation function=softmax)	?	?	?

Question#2 [CLO-2(SO:2-4)] [10 Marks]

For the following CNN model, apply the Stochastic Gradient Descent Backpropagation optimization technique, and update the weights (convolution layer (filter) weights and fully connected layer weights).
Do the following steps:

1. Forward pass: make a table where the activation shape (dimension + values), activation size and number of trainable parameters be clearly visible.
2. Calculate the cost/error value using the cost function.
3. Backpropagate the error and update the weights.
4. Calculate the output the for the same input using the forward pass again: make a table where the activation shape (dimension + values), activation size and number of trainable parameters be clearly visible.
5. Calculate the cost/error value using the cost function.
6. Compare the error values, calculated in step 2 and step 5.
7. What is the predicted output.

Consider the actual output=1.

<table>
<tr><td>Layer</td><td>Hyperparameters</td></tr>
<tr><td>Input</td><td>R=
<table>
<tr><td>1</td><td>2</td><td>1</td><td>0</td><td>3</td><td>1</td><td>1</td><td>2</td></tr>
<tr><td>3</td><td>1</td><td>2</td><td>1</td><td>1</td><td>1</td><td>2</td><td>2</td></tr>
<tr><td>2</td><td>0</td><td>1</td><td>1</td><td>2</td><td>0</td><td>1</td><td>1</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>2</td><td>1</td><td>1</td></tr>
<tr><td>2</td><td>2</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td><td>2</td></tr>
<tr><td>1</td><td>0</td><td>1</td><td>2</td><td>1</td><td>1</td><td>3</td><td>0</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>3</td><td>1</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>2</td><td>1</td><td>2</td><td>2</td><td>1</td></tr>
</table>
G=
<table>
<tr><td>1</td><td>1</td><td>3</td><td>1</td><td>2</td><td>1</td><td>1</td><td>2</td></tr>
<tr><td>2</td><td>1</td><td>2</td><td>1</td><td>3</td><td>3</td><td>2</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>2</td><td>2</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>2</td><td>2</td><td>2</td><td>1</td><td>1</td></tr>
<tr><td>3</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr>
<tr><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>2</td><td>1</td><td>1</td></tr>
<tr><td>1</td><td>0</td><td>0</td><td>1</td><td>2</td><td>1</td><td>2</td><td>1</td></tr>
<tr><td>1</td><td>2</td><td>1</td><td>1</td><td>0</td><td>1</td><td>2</td><td>1</td></tr>
</table>
B=
<table>
<tr><td>1</td><td>1</td><td>1</td><td>2</td><td>2</td><td>2</td><td>1</td><td>2</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>2</td></tr>
<tr><td>2</td><td>2</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr>
<tr><td>3</td><td>2</td><td>1</td><td>3</td><td>1</td><td>2</td><td>1</td><td>2</td></tr>
<tr><td>1</td><td>2</td><td>2</td><td>1</td><td>1</td><td>1</td><td>1</td><td>2</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>2</td><td>2</td><td>2</td><td>0</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td></tr>
</table></td></tr>
<tr><td>Conv1 (Separable)</td><td>Padding=Null, Stride=1,
Each bias weight=1
Filter1=
<table>
<tr><td>1</td><td>0</td><td>1</td></tr>
<tr><td>0</td><td>0</td><td>0</td></tr>
<tr><td>-1</td><td>0</td><td>-1</td></tr>
</table>
Filter2=
<table>
<tr><td>0</td><td>1</td><td>0</td></tr>
<tr><td>1</td><td>0</td><td>-1</td></tr>
<tr><td>0</td><td>-1</td><td>0</td></tr>
</table></td></tr>
<tr><td>Pool1</td><td>Padding=Null, Stride=2, Filter Size=2, Max Pooling</td></tr>
<tr><td>Conv2 (Separable)</td><td>Padding=Yes (Zero Value),
Each bias weight=0
Filter1=
<table>
<tr><td>1</td><td>0</td><td>1</td></tr>
<tr><td>0</td><td>-4</td><td>0</td></tr>
<tr><td>1</td><td>0</td><td>1</td></tr>
</table>
Stride=1,</td></tr>
<tr><td>Pool2</td><td>Padding=Null, Stride=1, Filter Size=2, Average Pooling</td></tr>
<tr><td>FC3</td><td>Number of units=2, Each hidden unit weight=0.5, Bias weight=0, Activation function=Relu</td></tr>
<tr><td>Output</td><td>Number of units=1, Each output unit weight=1, Bias weight=1, Activation function=Sigmoid</td></tr>
</table>