Indian Institute of Technology Kharagpur
Department of Electrical Engineering

Subject No.: EE60020 **Subject: Machine Learning for Signal Processing**
Date of Assignment: 15 April 2024 **Semester: Spring 2023-24**
Assignment Number: 4 **Duration: 1 hour 50 mins** **Full points: 308**

Name: ______________________ **Roll No:** ______________

---

1. You are provided with a convolutional neural network defined as follows

$$\begin{aligned}\texttt{net}_{\texttt{E}}(\cdot) \mapsto &(1:\text{Conv2D})\texttt{16c5w1s0p} \rightarrow (2:\text{MaxPool2D})\texttt{2w2s} \\ &\rightarrow (3:\text{Conv2D})\texttt{64c3w1s1p} \rightarrow (4:\text{MaxPool2D})\texttt{2w2s}\end{aligned} \tag{1}$$

$$\begin{aligned}\texttt{net}_{\texttt{D}}(\cdot) \mapsto &(1:\text{Conv2D})\texttt{16c3w1s1p} \rightarrow (2:\text{MaxUnPool2D})\texttt{2w2s} \\ &\rightarrow (3:\text{Conv2D})\texttt{3c5w1s2p} \rightarrow (4:\text{MaxUnPool2D})\texttt{2w2s}\end{aligned} \tag{2}$$

$$\texttt{net}_{\texttt{C}}(\cdot) \mapsto (1:\text{FC})\texttt{160} \rightarrow (2:\text{FC})\texttt{84} \rightarrow (3:\text{FC})\texttt{10} \rightarrow \tag{3}$$

such that it operates as follows when provided with an input $\mathbf{x} \in \mathbb{R}^{3\times 32\times 32}$

$$\mathbf{z} = \texttt{net}_{\texttt{E}}(\mathbf{x}) \tag{4}$$
$$\hat{\mathbf{x}} = \texttt{net}_{\texttt{D}}(\mathbf{z}) \tag{5}$$
$$\hat{\mathbf{y}} = \texttt{net}_{\texttt{C}}(\mathbf{z}) \tag{6}$$

(a) (16 points) Find the following associated with feedforward operation in $\texttt{net}_{\texttt{E}}(\cdot)$

	a in input $\in \mathbb{R}^a$	b in output $\in \mathbb{R}^b$	c in weight $\in \mathbb{R}^c$	d in bias $\in \mathbb{R}^d$
$\texttt{net}_{\texttt{E}}:1$				
$\texttt{net}_{\texttt{E}}:2$				
$\texttt{net}_{\texttt{E}}:3$				
$\texttt{net}_{\texttt{E}}:4$				

(b) (16 points) Find the following associated with feedforward operation in $\texttt{net}_{\texttt{D}}(\cdot)$

	a in input $\in \mathbb{R}^a$	b in output $\in \mathbb{R}^b$	c in weight $\in \mathbb{R}^c$	d in bias $\in \mathbb{R}^d$
$\texttt{net}_{\texttt{D}}$: 1				
$\texttt{net}_{\texttt{D}}$: 2				
$\texttt{net}_{\texttt{D}}$: 3				
$\texttt{net}_{\texttt{D}}$: 4				

(c) (12 points) Find the following associated with feedforward operation in $\texttt{net}_{\texttt{C}}(\cdot)$

	a in input $\in \mathbb{R}^a$	b in output $\in \mathbb{R}^b$	c in weight $\in \mathbb{R}^c$	d in bias $\in \mathbb{R}^d$
$\texttt{net}_{\texttt{C}}$: 1				
$\texttt{net}_{\texttt{C}}$: 2				
$\texttt{net}_{\texttt{C}}$: 3				

(d) (16 points) Find the number of unitary arithmeic operations associated with feedforward operation in $\texttt{net}_{\texttt{E}}(\cdot)$

	#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{E}}$: 1				
$\texttt{net}_{\texttt{E}}$: 2				
$\texttt{net}_{\texttt{E}}$: 3				
$\texttt{net}_{\texttt{E}}$: 4				

(e) (16 points) Find the number of unitary arithmeic operations associated with feedforward operation in $\texttt{net}_{\texttt{D}}(\cdot)$

	#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{D}}$: 1				
$\texttt{net}_{\texttt{D}}$: 2				
$\texttt{net}_{\texttt{D}}$: 3				
$\texttt{net}_{\texttt{D}}$: 4				

(f) (12 points) Find the number of unitary arithmeic operations associated with feed-forward operation in $\texttt{net}_{\texttt{C}}(\cdot)$

	#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{C}}:1$				
$\texttt{net}_{\texttt{C}}:2$				
$\texttt{net}_{\texttt{C}}:3$				

2. Consider that the neural networks in Q. 1 are trained with Mean Squared Error (MSE) as the loss function. Let $J_1(\hat{\mathbf{x}}, \tilde{\mathbf{x}})$ and $J_2(\hat{\mathbf{y}}, \mathbf{y})$ be evaluated with $\tilde{\mathbf{x}} \in \mathbb{R}^{3\times28\times28}$ and $\mathbf{y} \in \mathbb{R}^{10}$ representing the true states of the variables in the loss function. Let $\boldsymbol{\nabla}_1$ denote the derivative of $J_1(\cdot)$ with respect to the output of a given layer, and $\boldsymbol{\delta}_1$ denote the derivative of $J_1(\cdot)$ with respect to the input to that layer, such that for any layer represented mathematically as $\mathbf{Q} = \mathbf{RS} + \mathbf{T}$, where $\mathbf{Q}$ is the output of the neural layer we have $\nabla = \frac{\partial J(\cdot)}{\partial \mathbf{Q}}$, $\mathbf{S}$ is the input to the neural layer we have $\delta = \frac{\partial J(\cdot)}{\partial \mathbf{S}}$, $\mathbf{R}$ denotes the weights and $\mathbf{T}$ denotes the biases in the layer respectively. We have $\frac{\partial J(\cdot)}{\partial \mathbf{R}} = \boldsymbol{\nabla}\mathbf{S}^\top$, $\frac{\partial J(\cdot)}{\partial \mathbf{T}} = \boldsymbol{\nabla}$, $\boldsymbol{\delta} = \frac{\partial J(\cdot)}{\partial \mathbf{S}} = \mathbf{R}^\top \boldsymbol{\nabla}$. Similarly, $\boldsymbol{\nabla}_2$ and $\boldsymbol{\delta}_2$ correspond to these set of operations associated with $J_2(\cdot)$.

(a) (16 points) Find the following associated with error backpropagation operation in $\texttt{net}_{\texttt{E}}(\cdot)$ when $\mathbf{w}$ and $\mathbf{b}$ represent the weights and biases in a layer respectively.

	a in $\frac{\partial J_1(\cdot)}{\partial \mathbf{w}} \in \mathbb{R}^a$	b in $\frac{\partial J_1(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^b$	c in $\frac{\partial J_2(\cdot)}{\partial \mathbf{w}} \in \mathbb{R}^c$	d in $\frac{\partial J_2(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^d$
$\texttt{net}_{\texttt{E}}:1$				
$\texttt{net}_{\texttt{E}}:2$				
$\texttt{net}_{\texttt{E}}:3$				
$\texttt{net}_{\texttt{E}}:4$				

(b) (16 points) Find the following associated with error backpropagation operation in $\texttt{net}_{\texttt{D}}(\cdot)$ when $\mathbf{w}$ and $\mathbf{b}$ represent the weights and biases in a layer respectively.

	a in $\frac{\partial J_1(\cdot)}{\partial \mathbf{w}} \in \mathbb{R}^a$	b in $\frac{\partial J_1(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^b$	c in $\frac{\partial J_2(\cdot)}{\partial \mathbf{w}} \in \mathbb{R}^c$	d in $\frac{\partial J_2(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^d$
$\texttt{net}_{\texttt{D}}:1$				
$\texttt{net}_{\texttt{D}}:2$				
$\texttt{net}_{\texttt{D}}:3$				
$\texttt{net}_{\texttt{D}}:4$				

(c) (12 points) Find the following associated with error backpropagation operation in $\texttt{net}_{\texttt{C}}(\cdot)$ when $\mathbf{w}$ and $\mathbf{b}$ represent the weights and biases in a layer respectively.

	a in $\frac{\partial J_1(\cdot)}{\partial \mathbf{w}} \in \mathbb{R}^a$	b in $\frac{\partial J_1(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^b$	c in $\frac{\partial J_2(\cdot)}{\partial \mathbf{w}} \in \mathbb{R}^c$	d in $\frac{\partial J_2(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^d$
$\texttt{net}_{\texttt{C}} : 1$				
$\texttt{net}_{\texttt{C}} : 2$				
$\texttt{net}_{\texttt{C}} : 3$				

(d) (64 points) Find the number of unitary arithmeic operations associated with error backpropagation operation in $\texttt{net}_{\texttt{E}}(\cdot)$

		#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{E}} : 1$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{E}} : 2$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{E}} : 3$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{E}} : 4$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				

(e) (64 points) Find the number of unitary arithmeic operations associated with error backpropagation operation in $\texttt{net}_{\texttt{D}}(\cdot)$

		#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{D}}:1$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{D}}:2$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{D}}:3$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{D}}:4$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				

(f) (48 points) Find the number of unitary arithmeic operations associated with error backpropagation operation in $\texttt{net}_{\texttt{C}}(\cdot)$

		#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{C}}:1$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{C}}:2$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				
$\texttt{net}_{\texttt{C}}:3$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$				
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$				

End of questions.

This page is intentionally left blank for rough work.

This page is intentionally left blank for rough work.