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 Solution **Duration:** 1 hour 50 mins **Full points:** 308

Name: ______________________ **Roll No:** ______________

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

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

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

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

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 $\in \mathbb{R}^d$
$\texttt{net}_{\texttt{E}}:1$	$3\times 32\times 32$	$16\times 28\times 28$	$16\times 3\times 5\times 5$	16×1
$\texttt{net}_{\texttt{E}}:2$	$16\times 28\times 28$	$16\times 14\times 14$	-	-
$\texttt{net}_{\texttt{E}}:3$	$16\times 14\times 14$	$64\times 14\times 14$	$64\times 16\times 3\times 3$	64×1
$\texttt{net}_{\texttt{E}}:4$	$64\times 14\times 14$	$64\times 7\times 7$	-	-

(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 $\in \mathbb{R}^d$
$\texttt{net}_{\texttt{D}}:1$	$64 \times 7 \times 7$	$16 \times 7 \times 7$	$16 \times 64 \times 3 \times 3$	16×1
$\texttt{net}_{\texttt{D}}:2$	$16 \times 7 \times 7$	$16 \times 14 \times 14$	-	-
$\texttt{net}_{\texttt{D}}:3$	$16 \times 14 \times 14$	$3 \times 14 \times 14$	$3 \times 16 \times 5 \times 5$	3×1
$\texttt{net}_{\texttt{D}}:4$	$3 \times 14 \times 14$	$3 \times 28 \times 28$	-	-

(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 $\in \mathbb{R}^d$
$\texttt{net}_{\texttt{C}}:1$	$3,136 \times 1$	160×1	$160 \times 3,136$	160×1
$\texttt{net}_{\texttt{C}}:2$	160×1	84×1	84×160	84×1
$\texttt{net}_{\texttt{C}}:3$	84×1	10×1	10×84	10×1

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

	#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{E}}:1$	940,800	940,800	0	1,881,600
$\texttt{net}_{\texttt{E}}:2$	0	0	9,408	9,408
$\texttt{net}_{\texttt{E}}:3$	1,806,336	1,806,336	0	3,612,672
$\texttt{net}_{\texttt{E}}:4$	0	0	9,408	9,408

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

	#`add`	#`mul`	#`logic`	Total
$\texttt{net}_{\texttt{D}}:1$	451,584	451,584	0	903,186
$\texttt{net}_{\texttt{D}}:2$	0	0	0	0
$\texttt{net}_{\texttt{D}}:3$	235,200	235,200	0	470,400
$\texttt{net}_{\texttt{D}}:4$	0	0	0	0

(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
$\texttt{net}_\texttt{C}:1$	501,760	501,760	0	1,003,520
$\texttt{net}_\texttt{C}:2$	13,440	13,440	0	26,880
$\texttt{net}_\texttt{C}:3$	840	840	0	1,680

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 $\tilde{\mathbf{x}} \in \mathbb{R}^{3\times 28\times 28}$ 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 $\mathbf{Q} = \mathbf{R}\mathbf{S} + \mathbf{T}$, 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 $\frac{\partial J_2(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^d$
$\texttt{net}_\texttt{E}:1$	$16\times 3\times 5\times 5$	16×1	$16\times 3\times 5\times 5$	16×1
$\texttt{net}_\texttt{E}:2$	-	-	-	-
$\texttt{net}_\texttt{E}:3$	$64\times 16\times 3\times 3$	64×1	$64\times 16\times 3\times 3$	64×1
$\texttt{net}_\texttt{E}:4$	-	-	-	-

(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 $\frac{\partial J_2(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^d$
$\texttt{net}_\texttt{D}:1$	$16\times 64\times 3\times 3$	16×1	-	-
$\texttt{net}_\texttt{D}:2$	-	-	-	-
$\texttt{net}_\texttt{D}:3$	$3\times 16\times 5\times 5$	3×1	-	-
$\texttt{net}_\texttt{D}:4$	-	-	-	-

(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 $\frac{\partial J_2(\cdot)}{\partial \mathbf{b}} \in \mathbb{R}^d$
$\texttt{net}_\texttt{C} : 1$	-	-	$160 \times 3,136$	160×1
$\texttt{net}_\texttt{C} : 2$	-	-	84×160	84×1
$\texttt{net}_\texttt{C} : 3$	-	-	10×84	10×1

(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
$\texttt{net}_\texttt{E} : 1$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	$939,600$	$940,800$	0	$1,880,400$
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	$12,528$	16	0	$12,544$
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	$939,600$	$940,800$	0	$1,880,400$
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	$12,528$	16	0	$12,544$
$\texttt{net}_\texttt{E} : 2$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
$\texttt{net}_\texttt{E} : 3$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	$1,797,120$	$1,806,336$	0	$3,603,456$
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	$12,480$	64	0	$12,544$
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	$1,797,120$	$1,806,336$	0	$3,603,456$
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	$12,480$	64	0	$12,544$
$\texttt{net}_\texttt{E} : 4$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	0	0	0	0

(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
$\texttt{net}_{\texttt{D}}:1$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	$435,456$	$451,584$	0	$887,040$
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	768	16	0	784
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
$\texttt{net}_{\texttt{D}}:2$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
$\texttt{net}_{\texttt{D}}:3$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	$23,400$	$235,200$	0	$469,200$
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	585	3	0	588
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
$\texttt{net}_{\texttt{D}}:4$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	0	0	0	0

(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
$\texttt{net}_\texttt{C}:1$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	501, 760	501, 760	0	1, 003, 520
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	160	160	0	320
$\texttt{net}_\texttt{C}:2$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	13, 440	13, 440	0	26, 880
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	13, 440	13, 440	0	26, 880
$\texttt{net}_\texttt{C}:3$	$\frac{\partial J_1(\cdot)}{\partial \mathbf{w}}$	0	0	0	0
	$\frac{\partial J_1(\cdot)}{\partial \mathbf{b}}$	0	0	0	0
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{w}}$	840	840	0	1, 680
	$\frac{\partial J_2(\cdot)}{\partial \mathbf{b}}$	840	840	0	1, 680

End of questions.