10/11.

CS182 HW5.

1./(a). 1. CNN uses fewer parameters than FCN, CNN is more efficient.
2. CNN notices the spatial connection of image data, so its performance is better than FCN.

(b).

$$\begin{matrix} 1 & x \\ 4 & y \\ 0 & z \\ -2 & \\ 3 & \end{matrix} \qquad \begin{cases} x+4y=-2 \\ 4x-2z=2 \\ 3z-2y=11 \end{cases} \Rightarrow \begin{cases} x=2 \\ y=-1 \\ z=3 \end{cases}$$

∴ filter is: $[2, -1, 3]$

(c).

a	b	c
x		ax
y		ay
z	x	az+bx
	y	by
	z	bz+cx
		cy
		cz

$\Rightarrow [ax, ay, az+bx, by, bz+cx, cy, cz]$.

for 2D transpose conv:

output = input + k − 1 = 2 + 2 − 1 = 3.

$$\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & -2 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 3 & -2 \\ 0 & -1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 3 & -3 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 & -1 \\ 0 & 3 & 1 \end{bmatrix}$$

$+ \rightarrow \begin{bmatrix} -1 & 3 & -2 \\ 3 & -3 & 1 \\ 0 & 3 & 1 \end{bmatrix}$

$\begin{bmatrix} -k & 2k \\ 3k & k \end{bmatrix}$.

2./(a). 1. A. 2. B.

(b). $y_i = \gamma \hat{x}_i + \beta = \gamma x_i - \gamma\mu + \beta$.

$$\frac{\partial L}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial L}{\partial y_j} \cdot \frac{\partial y_j}{\partial x_i}$$

$$= \left[\sum_{j=1}^{n} -\frac{\gamma}{n} \frac{\partial L}{\partial y_j} \right] + \left[\gamma \frac{\partial L}{\partial y_i} \right]$$

$$= \gamma \left[\frac{\partial L}{\partial y_i} - \frac{1}{n} \sum_{j=1}^{n} \frac{\partial L}{\partial y_j} \right]$$

3/(a). $3 \times 3 \times 3 \times 4 = 108$

(b). $3 \times 3 \times 3 + 3 \times 4 = 39$

4./(a).