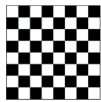


ہے تاکہ خدا

مسعود حمدي

تہذیب ساری سوسائٹی

یا دلیری کیسے



فرض کنید خانه‌های به رنگ سیاه در این جدول را با عدد صفر و خانه‌های به رنگ سفید را با عدد ۲۵۵ نشان دهیم.
میخواهیم عملکرد فیلتر زیر را روی این تصویر بررسی کنیم.

۱	۱	۱
۱	۱	۱
۱	۱	۱

با اعمال این فیلتر روی تصویر نتیجه را بدست آورید. فرض کنید پیکسل‌های موجود در لایه تصویر بدون تغییر باقی می‌مانند.
(ب) توضیح دهید این فیلتر به طور تقریبی چه عملی روی تصویر انجام می‌دهد.

$$\text{سایه‌فروبی} \leadsto k=7 \rightarrow k-1=6 \rightarrow 1=6$$

$$\frac{255 \times 6 + 0 \times 2}{9} = 119,44$$

$$\frac{255 \times 2 + 0 \times 6}{9} = 141,66$$

۱۴۱٫۶۶	۱۱۹٫۴۴	۱۴۱٫۶۶	۱۱۹٫۴۴	۱۴۱٫۶۶	۱۴۱٫۶۶
۱۱۹٫۴۴	۱۴۱٫۶۶	۱۱۹٫۴۴	۱۴۱٫۶۶	۱۱۹٫۴۴	۱۴۱٫۶۶
۱۴۱٫۶۶	//	//	//	//	//
۱۱۹٫۴۴	//	//	//	//	//
۱۴۱٫۶۶	//	//	//	//	//
۱۱۹٫۴۴	//	//	//		//

این مربع ۶ در ۶ که در آنجا به ترتیب ۱۱۹٫۴۴، ۱۴۱٫۶۶ هسته

(ب)

این خایه یک فیلتر میانگین‌گیر (moving average) فقدان است که یک فیلتر پایین‌گذراست و باعث همواری پیکسل‌های تصویری می‌شود و عملاً از تغییرات شدید بین پیکسل‌ها جلوگیری می‌کند.
با فرض ثابت در نظر گرفتن پیکسل‌های لبه (طبق صورت سؤال):



۲. (۵۰ نمره) شبکه عصبی کانولوشنی ای با لایه های ستون چپ جدول در نظر بگیرید. برای هر کدام از لایه های ذکر شده در جدول، ابعاد خروجی لایه و همچنین تعداد پارامترهای هر لایه را وارد نمایید. ابعاد خروجی را به صورت $H \times W \times C$ که به ترتیب نشان دهنده ارتفاع، عرض و عمق خروجی می باشد، نمایش دهید. همچنین نحوه نمایش هر لایه مطابق زیر می باشد:

- $\text{CONV}_{k-N}(S, P)$: یک لایه کانولوشی با N فیلتر، هر کدام به ابعاد $k \times k \times D$ که D عمق لایه قبلی می باشد و همچنین با گام (stride) برابر با S و تعداد پدینگ (padding) P . همچنین، در صورت ذکر نشدن مقادیر S و P هر دو را برابر با ۱ در نظر بگیرید.
- POOL-n : نمایش دهنده یک لایه max-pooling با ابعاد $n \times n$ ، گام n و پدینگ ۰ می باشد.
- FLATTEN : ورودی هموار می کند، معادل `torch.nn.flatten/tf.layers.flatten`
- FC-N : نمایش دهنده یک لایه fully-connected با N نورون می باشد.

Layer	Output Dimension	Number of Parameters
Input	$32 \times 32 \times 3$	0
CONV3-10	$32 \times 32 \times 10$	$(3 \times 3 \times 3 + 1) \times 10 = 100$
ReLU	$32 \times 32 \times 10$	۰
POOL-2	$16 \times 16 \times 10$	۰
CONV3-20(3,2)	$16 \times 16 \times 20$	$(3 \times 3 \times 10 + 1) \times 20 = 182$
ReLU	$16 \times 16 \times 20$	۰
POOL-2	$8 \times 8 \times 20$	۰
FLATTEN	640	۰
FC-10	۱۰	$640 \times 10 + 10 = 6410$

Padding
گام

$32 + 2 = 3 + (k - 1) \rightarrow k = 32 \rightarrow 32 \times 32 \times 10$ (Bios)
 $32 = 2 + 2(k - 1) \rightarrow k = 16$ (maxpool)

خروجی لایه برای $\rightarrow \text{out size} = \left(\frac{\text{in size} + 2 \times \text{pad} - \text{kernel size}}{\text{stride}} \right) + 1$

محاسبه سایز خروجی هر لایه
 خروجی، Padding، stride

۳. (۵۰ نمره) شبکه عصبی کانولوشنی یک بعدی ای مطابق شکل زیر در نظر بگیرید که ورودی‌های آن، ۵ متغیر x_1, x_2, x_3, x_4, x_5 می‌باشند و خروجی آن \hat{y} می‌باشد که با استفاده از آن مقدار تابع هزینه مطابق رابطه ذکر شده در شکل محاسبه می‌شود.

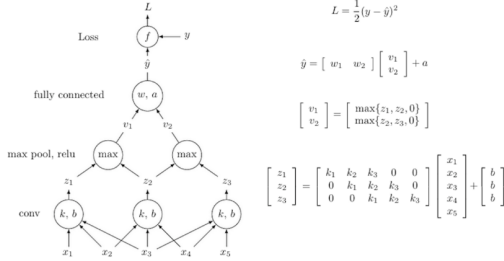
(آ) کدام یک از متغیرهای داده شده، پارامترهای شبکه می‌باشند؟
 (ب) مقادیر $\frac{\partial L}{\partial w_1}$ ، $\frac{\partial L}{\partial w_2}$ و $\frac{\partial L}{\partial w_3}$ را بر حسب y ، \hat{y} و v_1 ها به دست آورید.
 (ج) با فرض اینکه داشته باشیم:

$$\frac{\partial L}{\partial v_1} = \delta_1, \quad \frac{\partial L}{\partial v_2} = \delta_2$$

مطلوب است محاسبه مقادیر $\frac{\partial L}{\partial w_1}$ بر حسب مقادیر δ_1 و δ_2 .
 (د) با فرض اینکه داشته باشیم:

$$\frac{\partial L}{\partial z_1} = \alpha_1, \quad \frac{\partial L}{\partial z_2} = \alpha_2, \quad \frac{\partial L}{\partial z_3} = \alpha_3$$

مطلوب است محاسبه مقادیر $\frac{\partial L}{\partial b}$ ، $\frac{\partial L}{\partial k_1}$ ، $\frac{\partial L}{\partial k_2}$ ، $\frac{\partial L}{\partial k_3}$ ، $\frac{\partial L}{\partial k_4}$ ، $\frac{\partial L}{\partial k_5}$ بر حسب α_i و x_i ها.



شکل ۱: شبکه عصبی کانولوشنی یک بعدی

(ه) در حالت کلی، فرض کنید یک لایه کانولوشن یک بعدی با رابطه زیر داریم:

$$\begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} k_1 & \dots & k_d \\ & k_1 & \dots & k_d \\ & & \ddots & \\ & & & k_1 & \dots & k_d \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b \\ \vdots \\ b \end{bmatrix}$$

و همچنین می‌دانیم که:

$$\frac{\partial L}{\partial z_i} = \alpha_i$$

مطلوب است محاسبه $\frac{\partial L}{\partial b}$ و $\frac{\partial L}{\partial k_i}$ بر حسب α_i و x_i ها.

$$k_1, k_2, k_3, b, w_1, w_2, a \quad (۳)$$

(ب)

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial w_1} = -(y - \hat{y}) \times v_1$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial w_2} = -(y - \hat{y}) \times v_2$$

$$\frac{\partial L}{\partial a} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial a} = -(y - \hat{y}) \times 1$$

(ج)

$$v_1 = \text{Relu}(\underbrace{z_1}_{z'})$$

$$v_2 = \text{Relu}(\underbrace{\max(\text{pool}(z_2, z_3))}_{z''})$$

$$= \max(\text{pool}(z_2, z_3)) \begin{cases} z_2 & z_2 > z_3 \\ z_3 & z_3 > z_2 \end{cases}$$

$$\frac{\partial \max(\text{pool}(z_2, z_3))}{\partial z_1} = \text{Sgn}(z_2 - z_3)$$

$$\frac{\partial v_1}{\partial z'} = \text{Sgn}(v_1) \quad \frac{\partial z'}{\partial z_2} = \text{Sgn}(z_2 - z_3)$$

$$\frac{\partial L}{\partial z_2} = \frac{\partial L}{\partial v_1} \frac{\partial v_1}{\partial z'} \frac{\partial z'}{\partial z_2} = \delta_1 \odot \text{Sgn}(v_1) \odot \text{Sgn}(z_2 - z_3)$$

$$\frac{\partial L}{\partial z_3} = \frac{\partial L}{\partial v_1} \frac{\partial v_1}{\partial z'} \frac{\partial z'}{\partial z_3} + \frac{\partial L}{\partial v_2} \frac{\partial v_2}{\partial z''} \frac{\partial z''}{\partial z_3}$$

$$\frac{\partial v_2}{\partial z''} = \text{Sgn}(v_2) \quad \frac{\partial z''}{\partial z_3} = \text{Sgn}(z_3 - z_2)$$

$$\rightarrow \frac{\partial L}{\partial z_3} = \delta_1 \odot \text{Sgn}(v_1) \odot \text{Sgn}(z_3 - z_2) + \delta_2 \odot \text{Sgn}(v_2) \odot \text{Sgn}(z_3 - z_2)$$

$$\frac{\partial L}{\partial z_v} = \frac{\partial L}{\partial v_r} \frac{\partial v_r}{\partial z''} \frac{\partial z''}{\partial z_v} = \delta_v \odot \text{sgn}(v_r) \odot \text{sgn}(z_v - z_r)$$

$$\frac{\partial L}{\partial k_1} = \frac{\partial L}{\partial z_1} \frac{\partial z_1}{\partial k_1} + \frac{\partial L}{\partial z_r} \frac{\partial z_r}{\partial k_1} + \frac{\partial L}{\partial z_v} \frac{\partial z_v}{\partial k_1} = \alpha_1 x_1 + \alpha_r x_r + \alpha_v x_v \quad (2)$$

$$\frac{\partial L}{\partial k_r} = \frac{\partial L}{\partial z_1} \frac{\partial z_1}{\partial k_r} + \frac{\partial L}{\partial z_r} \frac{\partial z_r}{\partial k_r} + \frac{\partial L}{\partial z_v} \frac{\partial z_v}{\partial k_r} = \alpha_1 x_r + \alpha_r x_v + \alpha_v x_z$$

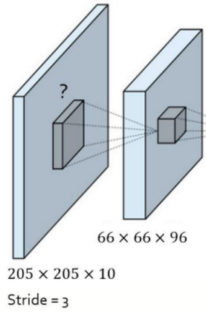
$$\frac{\partial L}{\partial k_v} = \frac{\partial L}{\partial z_1} \frac{\partial z_1}{\partial k_v} + \frac{\partial L}{\partial z_r} \frac{\partial z_r}{\partial k_v} + \frac{\partial L}{\partial z_v} \frac{\partial z_v}{\partial k_v} = \alpha_1 x_v + \alpha_r x_z + \alpha_v x_0$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial z_1} \frac{\partial z_1}{\partial b} + \frac{\partial L}{\partial z_r} \frac{\partial z_r}{\partial b} + \frac{\partial L}{\partial z_v} \frac{\partial z_v}{\partial b} = \alpha_1 + \alpha_r + \alpha_v$$

$$\frac{\partial L}{\partial k_j} = \sum_{i=1}^m \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial k_j} = \sum_{i=1}^m \alpha_i x_{j+i-1} \quad \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} k_1 & \dots & k_d & k_d & \dots \\ & k_1 & \dots & k_d & \dots \\ & & \ddots & & \ddots \\ & & & k_1 & \dots & k_d \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b \\ \vdots \\ b \end{bmatrix} \quad (2)$$

$$\frac{\partial L}{\partial b} = \sum \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial b} = \sum_{i=1}^m \alpha_i$$

(آ) با توجه به ابعاد ورودی و خروجی نشان داده شده در شکل زیر، سائز کرنل مورد استفاده در این عملیات کانولوشنی را بدست آورید. لازم به ذکر است که ابعاد ورودی مشخص شده در شکل، با احتساب $zero-padding$ داده شده اند.



$$\text{past_layer kernel size} + \text{stride}(\text{next layer} - 1)$$

$$205 = \text{kernel} + 3(96 - 1)$$

$$\text{kernel size} = 205 - 96 = 10$$

$$\boxed{\text{kernel size} = 10}$$

(ب) تعداد پارامترهای قابل آموزش یا همان *Learnable* موجود در این لایه کانولوشنی را تعیین نمایید.

(راهنمایی: به پارامتر بایاس موجود در هر کرنل نیز در محاسبات خود توجه داشته باشید.)

(ج) تعداد عملیات ضرب مورد نیاز برای بدست آوردن خروجی را محاسبه کنید. (ضرب های در صفر را نیز در شمارش تعداد ضرب ها لحاظ نمایید.)

تعداد کانالهای لایه کانولوشنی → ۹۶

$$\text{kernel size} \rightarrow 10 \times 10 \times 10 \Rightarrow \text{Learnable parameter} = 96 \times (10 \times 10 \times 10 + 1) = 94096$$

۱ bias → هر کانال

تعداد خروجی

$$10 \times 10 \times 10 \times 96 \times 96 = 907200$$

تعداد kernel

برای هر kernel و هر خروجی