

الف) قسمت تئوری

۱. فرض کنید در یک شبکه عصبی کانولوشنی ابعاد تصویر اولیه $N \times N$ و سائز فیلتر اعمال شده در تصویر ورودی نیز برابر $F \times F$ باشد. حال، اگر اندازه گام برابر S باشد، نشان دهید سائز خروجی این لایه برابر $1 + \frac{N-F}{S}$ می شود. (فرض کنید $\frac{N-F}{S}$ صحیح است.)
۲. فرض کنید ورودی یک لایه کانولوشنی تصاویری با ابعاد 65×65 و دارای ۳ کانال باشد و لایه کانولوشنی دارای ۱۲۸ فیلتر کانولوشن 5×5 با گام ۲ باشد. ابعاد خروجی و همچنین تعداد پارامترهای این لایه چقدر است؟
۳. در قسمت قبل، فرض کنید با یک لایه تمام متصل از تصویر ورودی به خروجی با ابعاد مشابه برسیم. تعداد پارامترها را در این حالت نیز به دست آورده و با قسمت قبل مقایسه کنید.
۴. توضیح دهید چگونه Max-Pooling باعث مقاومت شبکه نسبت به چرخش در تصویر ورودی می شود.
۵. یک لایه کانولوشن را به تنهایی در نظر بگیرید، فرض کنید سائز تصویر ورودی $M \times N$ باشد، سائز فیلتر را نیز $k_1 \times k_2$ در نظر بگیرید. ثابت کنید مشتق تابع هزینه نسبت به وزن لایه l از رابطه ۱ بدست می آید که در آن $o_{m,n}^l$ نشان دهنده خروجی تابع فعال ساز لایه قبل است که از اعمال f بر روی x بدست می آید و $\delta_{i,j}^l$ نشان دهنده مشتق تابع خروجی به هر عضو فضای ویژگی است. یعنی همان گونه که در معادله ۲ در نظر بگیرید که $x_{i,j}^l$ نشان دهنده المان با اندیس i, j از لایه l است و مشتق تابع هزینه نسبت به آن $\delta_{i,j}^l$ است. این مفهوم نشان دهنده آن است که با عمل کانولوشن با سائز فیلتر بین ورودی مرحل مستقیم لایه ی فعلی و ماتریس گرادیان محلی می توان پس انتشار خطا را انجام داد.

$$\frac{\partial E}{\partial w_{m',n'}^l} = \sum_{i=0}^{M-k_1} \sum_{j=0}^{N-k_2} \delta_{i,j}^l o_{i+m',j+n'}^{(l-1)} \quad (1)$$

$$x_{i,j} = \sum_m \sum_n w_{m,n}^l o_{i+m,j+n}^{(l-1)} + b^l \quad (2)$$

$$\frac{\partial E}{\partial x_{i,j}^l} = \delta_{i,j}^l, o_{i,j}^l = f(x_{i,j}^l)$$

۶. در قسمت قبل نشان دهید که گرادیان منتشر شده در $x_{i,j}^l$ از رابطه زیر به دست می آید:

$$\frac{\partial E}{\partial x_{i,j}^l} = \sum_{m=0}^{k_1-1} \sum_{n=0}^{k_2-1} \delta_{i-m,j-n}^{l+1} w_{m,n}^{l+1} f'(x_{i,j}^l) \quad (3)$$

ب) قسمت عملی

در قسمت عملی این تمرین، با استفاده از پایگاه داده **MNIST** به طراحی و پیاده سازی شبکه عصبی کانولوشنی پرداخته می شود:

۱. یک شبکه عصبی کانولوشنی به صورت زیر در نظر بگیرید:

۶۴ فیلتر کانولوشن با ابعاد 5×5 و گام ۱ با تابع فعال سازی ReLU، لایه Max-Pooling با ابعاد 2×2 و گام ۲، ۶۴ فیلتر کانولوشن با ابعاد 5×5 و گام ۱ با تابع فعال سازی ReLU، لایه Max-Pooling با ابعاد 2×2 و گام ۲، یک لایه تمام متصل با ۲۵۶ نورون و تابع فعال سازی ReLU، لایه آخر نیز ۱۰ نورون (به تعداد کلاس های پایگاه داده) با تابع فعال سازی softmax

این شبکه را با استفاده از بهینه‌ساز SGD و با نرخ یادگیری ۰,۰۱ با تابع هزینه Cross Entropy آموزش دهید.^۱ نمودار تابع هزینه و دقت دسته‌بندی برای داده‌های یادگیری و اعتبارسنجی را در حین آموزش در گزارش خود ارائه کنید. (برای بهبود یادگیری می‌توانید از منظم‌ساز استفاده کنید).

۲. برای شبکه آموزش دیده شده در قسمت ۱، وزن‌های فیلترهای مختلف در لایه کانولوشن اول شبکه را به صورت تصویر ذخیره و در گزارش خود ارائه کنید. توضیح دهید وزن‌های به‌دست آمده چه فیلترهایی روی تصویر ورودی اعمال می‌کند.

۳. برای شبکه آموزش دیده شده در قسمت ۱، با یک تصویر ورودی ثابت مربوط به کلاس (عدد) ۵، خروجی فیلترهای مختلف در لایه‌های کانولوشن شبکه را به صورت تصویر ذخیره و در گزارش خود ارائه کنید. خروجی فیلترهای مختلف شبکه را برای تصویر ورودی مقایسه و درمورد خروجی لایه‌های مختلف شبکه توضیح دهید.

۴. در این قسمت هدف آموزش یک دسته‌بند برای دو دسته متفاوت با استفاده از مفهوم یادگیری انتقالی و استفاده از شبکه آموزش دیده شده در قسمت ۱ است. به این منظور، باید لایه‌های کانولوشن شبکه آموزش دیده شده در قسمت ۱ بدون به‌روزرسانی و ثابت بگیرید، اما لایه‌های تمام متصل به‌روزرسانی شود. در لایه آخر شبکه نیز، از دو نورون برای تعیین دو دسته عدد «۱» و عدد «۴» استفاده کنید. با بهینه‌سازی وزن‌های لایه‌های تمام متصل، دسته‌بند مناسب را آموزش دهید. نمودار تابع هزینه و دقت دسته‌بندی برای داده‌های یادگیری و اعتبارسنجی را در حین آموزش در گزارش خود ارائه کنید.

^۱ توجه شود که پیاده‌سازی های این بخش باید بدون استفاده از tf.layers و با استفاده از عملگرهایی مثل tf.nn.conv2d , tf.nn.max_pool, ... انجام شود.