

به نام خدا

استاد: دکتر مرضیه داود آبادی
درس مبانی یادگیری عمیق

نام: فاطمه زهرا بخشنده
شماره دانشجویی: 98522157

گزارش تمرین 4:

سوال اول:

عملیات کانولوشن دو بعدی به صورت زیر است:

$$S(i, j) = (I \star K)(i, j) = \sum_m \sum_n I(i + m, j + n) K(m, n)$$

ابتدا کرنل را 180 درجه می چرخانیم (این کرنل چون متقارن است، با خودش یکی می شود):

1	1	1
1	-8	1
1	1	1

ابتدا با توجه به ابعاد کرنل، به تصویر zero padding اضافه می کنیم.

0	0	0	0	0	0	0
0	0	0	10	0	0	0
0	0	0	10	0	0	0
0	0	0	10	0	0	0
0	0	0	10	0	0	0
0	0	0	10	0	0	0
0	0	0	0	0	0	0

سپس از خانه (0,0) شروع کرده و کرنل را روی تصویر حرکت می دهیم. طبق فرمول بالا، اعدادی که روی هم قرار می گیرند در هر ضرب شده و باهم جمع می شوند. و نتیجه convolution تصویر با کرنل بدست می آید.

0	20	-70	20	0
0	30	-60	30	0
0	30	-60	30	0
0	30	-60	30	0
0	20	-70	20	0

تحلیل نتیجه: کرنل استفاده شده، کرنل لاپلاسیان است، که در واقع مشتق دوم تصویر را به صورت تقریبی محاسبه می کند.

$$\Delta^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

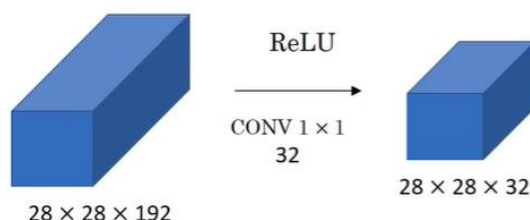
لاپلاسیان تغییرات شدت روشنایی را برجسته میکند و لبه ها و تغییرات تصویر را تشخیص می دهد. همانطور که دیدیم این تصویر دارای یک لبه عمودی بود که این کرنل آن را پیدا کرده است. و جاهایی که تغییرات بیشتر است (با توجه به zero padding)، مقادیر بزرگتر شده اند.

سوال دوم:

استفاده از کانولوشن های 1 در 1 می تواند چند مزیت داشته باشد.

- 1- این نوع کانولوشن بدون تغییر ابعاد (width و height)، عمق را کاهش می دهد. در واقع اگر بخواهیم ابعاد را تغییر دهیم می توانیم از pooling layer استفاده کنیم. اما اگر بخواهیم تعداد channels را تغییر دهیم، آنگاه می توانیم از کانولوشن 1×1 استفاده کنیم. برای مثال در شکل زیر، تعداد channel ها 192 تاست و می خواهیم آن ها را کاهش دهیم و خروجی سمت چپ را داشته باشیم. در این صورت از 32 تا فیلتر $1 \times 1 \times 192$ استفاده می کنیم و ابعاد خروجی $28 \times 28 \times \text{number of filters}$ خواهد شد که در اینجا number of filters برابر 32 است.

Using 1×1 convolutions



- 2- اگر از کانولوشن 1×1 با تعداد فیلتر برابر هم استفاده کنیم، در واقع یک لایه غیر خطی (در شکل بالا با تابع فعالسازی Relu) اضافه کرده ایم که توانایی یادگیری پیچیده تری دارد. در واقع بدون تغییر عمق، nonlinearity اضافه می کند.
- 3- این نوع کانولوشن خاصیت های مکانی را در نظر می گیرد. یک روش نگاه به آن این است که یک نورون را در نظر بگیریم که 192 تا input می گیرد و هرکدام از این 192 عدد را در یک slice مکانی ضرب می کند. و در واقع این 192 کانال، هر مکان را در 192 وزن مختلف ضرب می کنند، و سپس یک Relu Non Linearity به آن اضافه می کنند. و سپس خروجی می دهند. در نتیجه اگر تعداد زیادی فیلتر داشته باشیم، آنگاه تعداد زیادی unit خواهیم داشت، که به عنوان ورودی، همه ویژگی های یک مکان را گرفته، آن ها را ترکیب کرده و یک ویژگی جدید میسازد.
- در واقع می توان گفت کانولوشن 1×1 یک fully connected neural network در خود دارد. یک network که طبق مثال شکل بالا، 192 عدد را ورودی گرفته، و number of filters عدد خروجی می دهد. این ایده، به اصطلاح Network in Network (شبکه درون شبکه) نیز معروف است.
- خواص بالا باعث شده این نوع کانولوشن در شبکه هایی مانند inception که تعداد لایه های بالایی دارند استفاده شود. در واقع در برخی لایه ها از کانولوشن 1×1 استفاده می شود تا قدرت شبکه را افزایش داده و عمق را کاهش دهد، و بتوان لایه های بیشتری نیز به شبکه اضافه کرد.