

به نام خدا

تمرین سری ششم - درس مبانی بینایی کامپیوتر

سید محمد علی نخاری - شماره دانشجویی: 99521496

### سوال دوم (الف)

برای به دست آوردن تعداد پارامترهای قابل آموزش در ماژول داده شده باید تعداد پارامترها به ازای هر لایه را به دست آورده و در نهایت تمامی پارامترها را با هم جمع کنیم تا تعداد کل پارامترهای قابل آموزش محاسبه شود. نکته دیگر که باید به آن توجه داشت skip connection موجود در ماژول است که پارامتری برای آموزش نداشته و صرفاً داده را از لایه ای به لایه دیگر منتقل میکند:

برای محاسبه تعداد پارامترهای هر لایه از رابطه زیر استفاده میکنیم:

$$(\text{filter\_size} * \text{\#input\_channels} + 1(\text{bias})) * \text{\#filters}$$

در ابتدا به صورت موازی ورودی داده شده وارد سه لایه کانولوشنی با اندازه  $1*1$  و 32 فیلتر می شود، پس تعداد پارامترها در این مرحله به صورت زیر محاسبه میشود:

$$((1*1*3 + 1) * 32) * 3 = 384$$

پس از این مرحله با توجه به اینکه فیلتر  $1*1$  تصویر را صرفاً از لحاظ عمق دچار تغییر میکند و ابعاد مکانی بدون تغییر می مانند خروجی از این سه لایه به صورت  $32 * n * n$  خواهد بود.

در این مرحله دو تا از خروجی های به دست آمده از لایه قبلی وارد یک لایه کانولوشنی  $3*3$  با 32 فیلتر میشوند که تعداد پارامترها به صورت زیر قابل محاسبه است:

$$((3*3*32 + 1) * 32) * 2 = 18496$$

• با توجه به اینکه خروجی لایه ها در نهایت باید با هم concat شوند فرض میکنیم که از padding هم در زمان کانولوشن استفاده میکنیم.

در مرحله بعد یکی از خروجی های به دست آمده، وارد لایه کانولوشنی  $3*3$  با 32 فیلتر میشوند که تعداد پارامترها در این مرحله به صورت زیر است:

$$((3*3*32 + 1) * 32) = 9248$$

در نهایت خروجی های به دست آمده از مراحل قبل را با یکدیگر concat کرده و وارد یک لایه کانولوشنی  $1*1$  دیگر با 256 فیلتر میکنیم. ورودی این لایه به صورت  $n * n * 96$  میباشد:

$$((1*1*96 + 1) * 256) = 24832$$

تعداد کل پارامترها برای ورودی داده شده به این ماژول برابر است با حاصل جمع تمامی تعداد پارامترهای محاسبه شده در مراحل بالا:

$$384 + 18496 + 9248 + 24832 = 52960 \rightarrow \text{تعداد کل پارامترهای ماژول}$$

برای به دست آوردن میدان تاثیر این ماژول از لایه ورودی شروع به حرکت کرده و در ابتدا مقدار میدان تاثیر را برابر با  $1*1$  در نظر میگیریم و به ازای هر لایه کانولوشنی با اندازه فیلتر  $k*k$  به اندازه  $k-1$  به مقدار میدان تاثیر اضافه میکنیم. در زمانی که نیاز به concat داشته باشیم و یا زمانی که به skip connectio ها میرسیم، از میان میدان تاثیر های به دست آمده، حداکثر مقدار آن را قرار میدهیم. با توجه به توضیحات داده شده، میدان تاثیر این ماژول برابر با  $5 * 5$  خواهد بود.

(ب)

A) *Conv2D(filters=16, kernel\_size=(3, 3), padding='valid')*  
*Conv2D(filters=32, kernel\_size=(3, 3), padding='valid')*

لایه اول:

- تعداد فیلترها = 16
- اندازه فیلتر = (3,3)
- تعداد کانال های تصویر ورودی = 3
- تعداد پارامترهای قابل آموزش  $16 * (3*3*3 + 1) = 448$
- خروجی لایه :  $16 * n-2 * n-2$

لایه دوم:

- تعداد فیلترها = 32
- اندازه فیلتر = (3,3)

- تعداد کانال های تصویر ورودی = 16
- تعداد پارامتر های قابل آموزش  $32 = (3 \times 3 \times 16 + 1) \times 32 = 4640$
- خروجی لایه :  $32 \times n-4 \times n-4$
- تعداد کل پارامتر ها :  $448 + 4640 = 5088$
- میدان تاثیر:  $5 \times 5$  (روش به دست آمدن در قسمت قبل توضیح داده شد)

B) *LocallyConnected2D(filters=16, kernel\_size=(3, 3), padding='valid')*  
*LocallyConnected2D(filters=32, kernel\_size=(3, 3), padding='valid')*

لایه اول:

- تعداد فیلتر ها = 16
- اندازه فیلتر =  $(3,3)$
- تعداد کانال های تصویر ورودی = 3
- تعداد پارامتر های قابل آموزش:  $16 \times (3 \times 3 \times 3 + 1) \times ((n-2) \times (n-2))$
- خروجی لایه :  $16 \times n-2 \times n-2$

لایه دوم:

- تعداد فیلتر ها = 32
- اندازه فیلتر =  $(3,3)$
- تعداد کانال های تصویر ورودی = 16
- تعداد پارامتر های قابل آموزش:  $32 \times (3 \times 3 \times 16 + 1) \times ((n-4) \times (n-4))$
- خروجی لایه :  $32 \times n-4 \times n-4$

تعداد کل پارامتر ها:  $((n^2 - 4n + 4) \times 448) + ((n^2 - 8n + 16) \times 4640)$

- میدان تاثیر:  $5 \times 5$  (در این حالت هم همانند قسمت قبل میدان تاثیر  $5 \times 5$  خواهد بود چرا که در نهایت هر پیکسل با فیلتر های  $3 \times 3$  در هر مرحله کانالو شده است)
- از لحاظ تعداد کل پارامتر های قابل آموزش در حالت locally connected وابسته به اندازه تصویر ورودی است و به طور کلی میتوان گفت به ازای یک اندازه برای دو

- قسمت A, B تعداد پارامترها در حالت دوم بیشتر خواهد بود چرا که برای هر ناحیه فیلتری با وزن ها جدا در نظر گرفته میشود و باید جداگانه آموزش بیابد.
- از لحاظ میدان تاثیر با توجه به اندازه لایه های کانولوشنی هر دو برابر و  $5*5$  خواهند بود.