

به نام خدا

عنوان:

بررسی سوال سوم تکلیف شبکه عصبی

استاد:

دکتر منصوری

دانشجو:

محمد علی مجتهد سلیمانی

۴۰۳۳۹۰۴۵۰۴

تاریخ:

۱۴۰۳/۰۹/۳۰

✓ بنده در این فایل موارد الف و ب را بررسی خواهم کرد و گزارش پروژه پیاده سازی شده در فایل های مربوط به آن تحت عنوان "Explanation Part" عنوان شده است.

## سوال اول

(الف)

تکنیک separable convolution قصد دارد که تعداد پارامتر و هزینه محاسبات را کاهش بدهد بدون اینکه به طور قابل توجه کارایی را کاهش بدهند. این کار به این صورت انجام میشود که کانولوشن استاندارد را به دو قسمت جدا و کانولوشن های کوچکتر تقسیم میکند:

✓ Depthwise convolution:

در این مرحله ما یک فیلتر کانولوشن به هر channel ورودی به صورت جداگانه اعمال میکنیم.

اگر channel ورودی ما به اندازه  $X$  چنل داشته باشد ما  $X$  فیلتر به صورت جداگانه که هر کدام روی یک channel اعمال میکند.

در این مرحله ما اطلاعات spatial را بدست میاوریم داخل هر channel ولی اطلاعات را میان channel های مختلف ترکیب نمیکند.

## ✓ Pointwise convolution

در این مرحله ما یک کانولوشن  $1*1$  اعمال میکنیم تا خروجی های مرحله قبلی را ترکیب و یکی بکنیم.

این کانولوشن به صورت موثر یک weighted sum بر روی تمام channel ها بر روی هر موقعیت spatial اعمال میکند. تا یک feature map جدید بسازد که اطلاعات channel های مختلف را با هم ترکیب میکند.

اگر channel خروجی ما  $f$  تا داشته باشد، ما  $f$  تا فیلتر  $1*1$  اعمال میکنیم.

مزیت این روش این است که هزینه محاسباتی را کاهش میدهد به این دلیل است که depthwise فقط داخل هر channel محاسبه میکند و pointwise هم kernel بسیار کوچکی دارد. این امر سبب میشود که بسیار نسبت به کانولوشن معمولی سریعتر شود.

همچنین این روش مقادیر بسیار کمتری پارامتر برای یادگیری دارد. و باعث کاهش سایز مدل میشود و بسیار برای memory مناسب است. و کمتر به overfitting متمایل میشود.

## بخش ب

سوال ۳  
بخش ب

کانولوشن عادی

$$\text{output} = \left\lfloor \frac{N + P - F}{\text{stride}} \right\rfloor + 1 = \left\lfloor \frac{14 - 5}{1} \right\rfloor + 1 + 1 = 12$$

input =  $N = 14$   
 $P = 0$   
 $S = 1$

تعداد فیلتر = ۲۵۴

output size =  $12 \times 12 \times 254$

تعداد پارامتر =  $(\text{filter height} \times \text{filter width} \times \text{input channel} + \text{bias}) \times \text{filter}$   
 $= (5 \times 5 \times 3 + 1) \times 254 = 19454$

تعداد فیلتر =  $\text{output height} \times \text{output width} \times \text{output channel} \times \text{filter height} \times \text{filter width} \times \text{input channel}$   
 $12 \times 12 \times 254 \times 5 \times 5 \times 3 = 2745600$

در کانولوشن عادی همانطور که در تمرین اول توضیح دادم تعداد پارامتر و تعداد ضرب و سائز خروجی را بدست آوردیم و فرمول ها را نوشتیم.

## separable Convolution

✓ output sizes  $12 \times 12 \times 256$   
 به ساندکافوشن عا (۵)

✓ Depthwise:  $(\text{filter height} \times \text{filter width} \times 1) \times \text{input channel}$   
 $= (5 \times 5 \times 1) \times 3 = 15$

✓ Pointwise  $(1 \times 1 \times \text{input channel} + 1) \times \text{number of filter}$   
 $(1 \times 1 \times 3 + 1) \times 256 = 1024$

total  $= 15 + 1024 = 1102$

✓ Depthwise =  $\text{output height} \times \text{output width} \times \text{input channel}$   
 $\times \text{filter height} \times \text{filter width} \times 1$   
 $= 12 \times 12 \times 3 \times 5 \times 5 \times 1 = 10800$

Pointwise  $= \text{output height} \times \text{output width} \times \text{no of filter} \times 1 \times 1$   
 $= 12 \times 12 \times 256 \times 1 \times 1 \times 3 = 110592$   
 input channel

total  $= 10800 + 110592 = 121392$

در این بخش separable convolution را محاسبه کردیم و مقادیر مشخص است  
 هم تعداد پارامتر هم تعداد ضرب و هم ساینز خروجی را بدست آوردیم.

