

به نام خدا

عنوان:

تکلیف سوم درس شبکه های عصبی

استاد:

دکتر منصوری

دانشجو:

محمد علی مجتهد سلیمانی

4033904504

تاریخ:

۱۴۰۳/۰۹/۳۰

سوال ۱

❖ الف)

سوال ۱
الف)

input = $428 \times 512 \times 3 \rightarrow RGB$ تصاویر

تعداد فیلتر = ۱۲۸

سایز فیلتر = 3×3

strides = 2

padding = 0

Filter

$$\text{output} = \left\lfloor \frac{N + 2P - F}{\text{stride}} \right\rfloor + 1$$

$$\text{output} \rightarrow \left\lfloor \frac{428 - 3}{2} \right\rfloor + 1 = 312 + 1 = 313 \quad \left\lfloor \frac{512 - 3}{2} \right\rfloor + 1 = 254 + 1 = 255$$

تعداد فیلتر = $313 \times 255 \times 128$

bias

تعداد = $(3 \times 3 \times 3 + 1) \times 128 = 3584$

تعداد فیلتر

Receptive field = 3×3 → سایز فیلتر

لایه اول

است و هر دور می تواند 3×3 سایز از دوری را ببیند. (Receptive)

$$RF = (RF_{\text{Previous}} - 1) \times \text{stride} + \text{filtersize}$$

لایه دوم

$$RF = (3 - 1) \times 2 + 3 = 7$$

همانطور که در تصویر مشخص است بر اساس سایز ورودی و تعداد فیلتر و سایز فیلتر ($3 \times 3 \times 3$) و توانستیم سایز خروجی را محاسبه بکنیم.

تعداد پارامتر هم برابر است با سایز فیلتری که داریم بعلاوه bias در هر لایه ضربدر تعداد فیلتری که اعمال شده است که همانطور که بیان شد ۱۲۸ بود.

سایز فیلتر $3*3*3$ هست به سبب اینکه ورودی ما دارای ۳ channel است.

سایز خروجی ما $313*255*128$ است که نحوه محاسبات در عکس آمده است. ۱۲۸ بخاطر ۱۲۸ فیلتر اعمال شده است.

برای بدست آوردن receptive field در ۲ لایه اقدام کردیم:

لایه اول: در این لایه به سبب اینکه سایز فیلتر ما $3*3$ است به این معنا است که هر نورون ما میتواند به اندازه $3*3$ از ناحیه ورودی را ببیند.

لایه دوم: در این لایه به سبب اینکه سایز فیلتر ما در این لایه $3*3$ است همچنین گام ۲ است سایز receptive را $7*7$ بدست آوردیم.

به طور کلی receptive field به ناحیه ای گفته میشود از فضای ورودی که هر نورون میتواند به آن منطقه دید داشته باشد یا حساس باشد. به این دلیل است که ما در CNN ها بر اساس local connectivity کار میکنیم به این معنی که هر نورون در لایه کانولوشن میتواند به یک ناحیه بسیار کوچک از لایه قبلی متصل شود نه مانند شبکه های عصبی قدیمی که نورون در لایه فعلی به تمام نورون ها در لایه های قبلی متصل میشد. هر چه که در شبکه عمیق تر شویم میزان دید ما هم بیشتر میشود و یک نورون در لایه های میانی نواحی بسیار بزرگتری از لایه های قبلی را میتواند ببیند.

بخش ب در ادامه قرار دارد.

(ب)

(ب)

۱۲۸ × ۵۲۵ × ۳۱۳ = خروجی لایه

قبل

* برای تبدیل به یک لایه تمام متصل باید این ابعاد رو flat کنیم و یک بوزا

Flat

bias

به دست بیاریم. بعد با bias جمع کنیم.

$(313 \times 525 \times 128 + 1) \times 79115$

۳۱۳ × ۵۲۵

حاصل ضرب

۳۵۱۴ = تعداد پارامتر

کانونولشن

۱۱۵, ۴۱۵, ۵۱۲, ۶۱۵ = تعداد پارامتر

لایه تمام متصل