

پاسخ سوال جلسه پنجم (۱-۵)

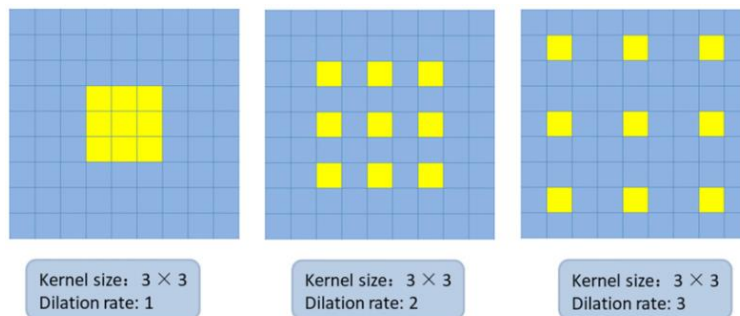
سوال : در Dilation و Stride با اندازه P با چه فرمولی سایز کوچک می شود؟ آیا سایز نصف می شود؟

۱- Dilated Convolution:

استفاده از Dilation در عملکرد Convolution به تنهایی باعث کوچک شدن سایز تصویر نمی شود. بلکه، Dilation بر روی اندازه و اندازه مکانی فیلتر اثر می گذارد و تعداد نقاطی که فیلتر روی آن ها اعمال می شود را تغییر می دهد و باعث کاهش فضای نمونه برداری (downsampling) می شود که در نتیجه می تواند کاهش سایز تصویر را تسهیل کند. با افزایش Dilation rate، فاصله بین نقاطی که فیلتر بر روی آن ها اعمال می شود، بیشتر می شود و در نتیجه ابعاد خروجی کاهش می یابد. اگر فرض کنیم اندازه فیلتر را F و Dilation rate را P بنامیم، تعداد نقاطی که فیلتر بر روی آن ها اعمال می شود به صورت زیر قابل محاسبه است:

$$(F - 1) * P + 1 = \text{تعداد نقاط}$$

اما این نکته را باید توجه داشته باشید که Dilation به تنهایی نمی تواند سایز تصویر را تغییر دهد و همچنین نمی تواند سایز تصویر را نصف کند تأثیر آن به ترکیب با سایر لایه ها و پارامترهای شبکه بستگی دارد. برای کاهش سایز تصویر، معمولاً از عملیات نمونه برداری (downsampling) مانند pooling یا استفاده از stride بزرگتر استفاده می شود



۲- Strided Convolution:

Stride در Convolution باعث کاهش ابعاد تصویر نمی شود. بلکه، باعث کاهش تعداد نقاطی می شود که فیلتر روی آن ها اعمال می شود. سایز تصویر خروجی با استفاده از stride، اگر فرض کنیم اندازه ورودی را W و اندازه فیلتر را F و padding مقدار p باشد به صورت زیر قابل محاسبه است:

$$\text{سایز خروجی} = (W - F + 2p) / \text{stride} + 1$$

بنابراین، stride نهایتاً تعداد نقاطی را کاهش می دهد که فیلتر روی آن ها اعمال می شود، اما سایز تصویر نصف نمی شود. اگر بخواهیم سایز تصویر را نصف کنیم، باید از دو عملکرد متوالی کاهش نمونه برداری (downsampling) با استفاده از stride برای هر کدام از ابعاد عرض و ارتفاع استفاده کنیم. به عنوان مثال، اگر stride را برابر با ۲ در هر ابعاد قرار دهیم، سایز تصویر خروجی نصف سایز تصویر ورودی خواهد بود.