

رسالة محمد

یادگیری عمیق

مدرس: محمدرضا محمدی

زمستان ۱۴۰۱

شبکه‌های عصبی کانولوشنی

Convolutional Neural Networks

لایه‌های متصل محلی

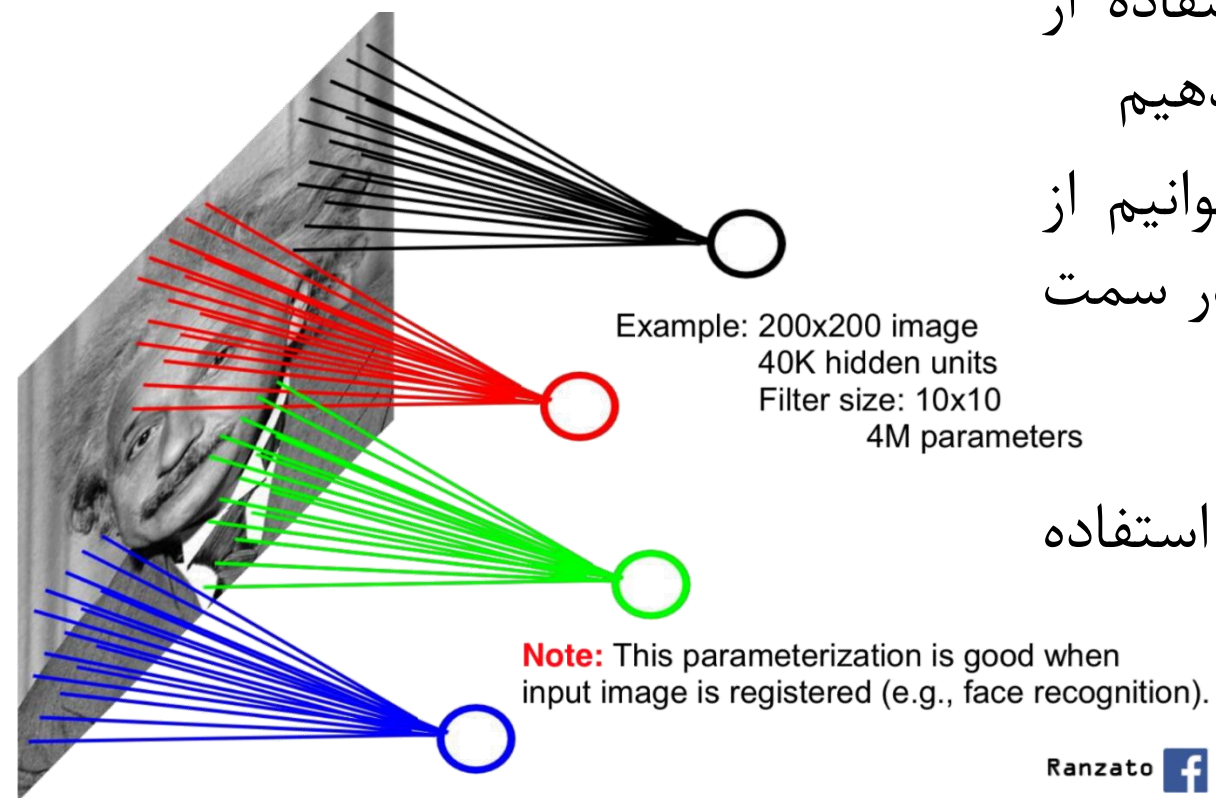
- بسیاری از ویژگی‌هایی که چشم انسان به راحتی می‌تواند تشخیص دهد، ویژگی‌های محلی هستند

- ما می‌توانیم لبه‌ها، بافت‌ها و حتی شکل‌ها را با استفاده از شدت پیکسل‌ها در ناحیه کوچکی از تصویر تشخیص دهیم

- اگر می‌خواهیم یک ویژگی را تشخیص بدهیم، می‌توانیم از همان آشکارساز در گوشه پایین سمت چپ تصویر و در سمت راست بالای تصویر استفاده کنیم

- ما می‌توانیم از وزن‌های یکسان در هر مکان از تصویر استفاده کنیم

- اشتراک وزن‌ها (weight sharing)



کانولوشن



$$G_y$$

+1	0	-1
+2	0	-2
+1	0	-1

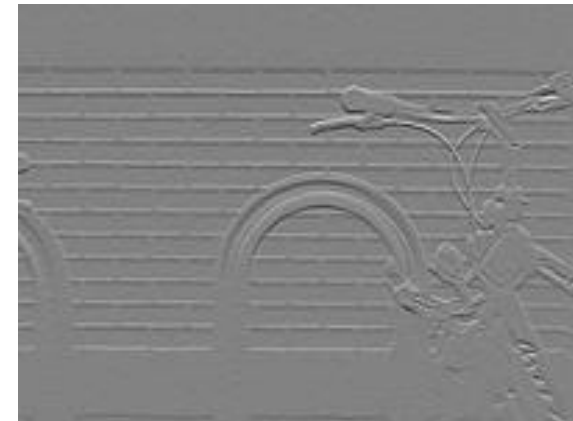


3_0	3_1	2_2	1	0
0_2	0_2	1_0	3	1
3_0	1_1	2_2	2	3
2	0	0	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

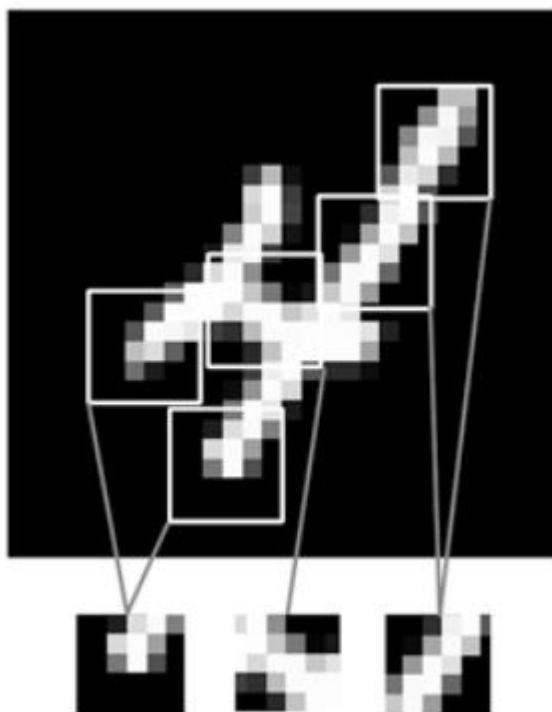
$$G_x$$

+1	+2	+1
0	0	0
-1	-2	-1



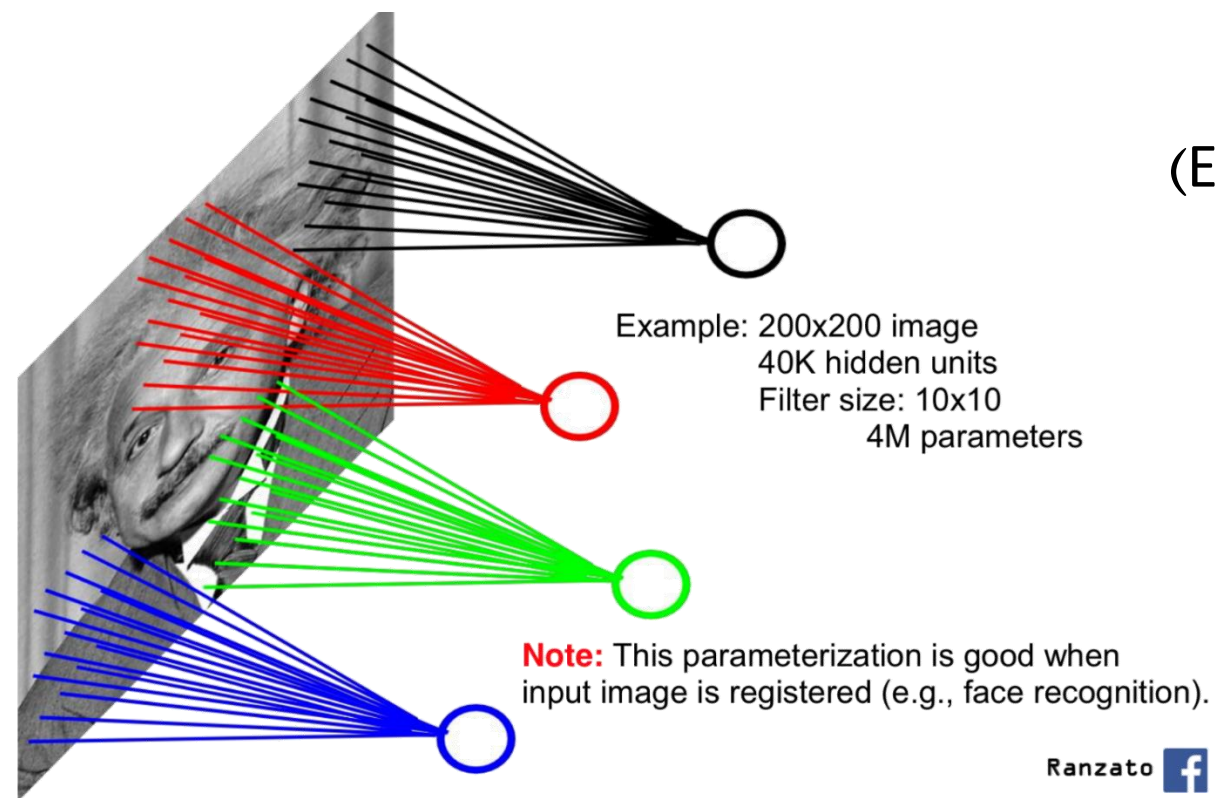
لایه کاملاً متصل در مقایسه با لایه کانولوشنی

- لایه‌های کاملاً متصل الگوهای سراسری را در فضای ویژگی ورودی خود می‌آموزند
 - به عنوان مثال، برای یک رقم MNIST، الگوهایی که شامل همه پیکسل‌ها هستند
- لایه‌های کانولوشنی الگوهای محلی را یاد می‌گیرند



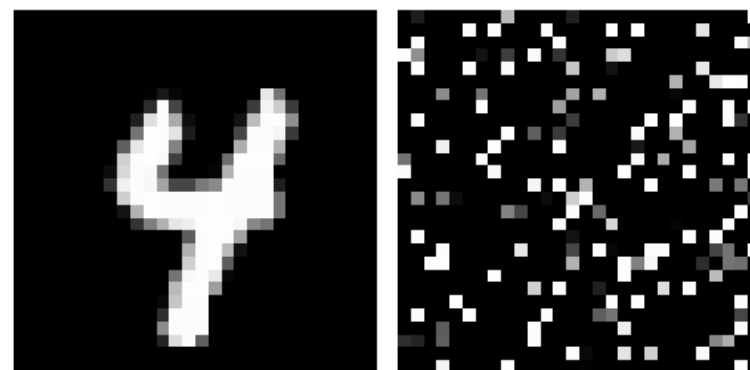
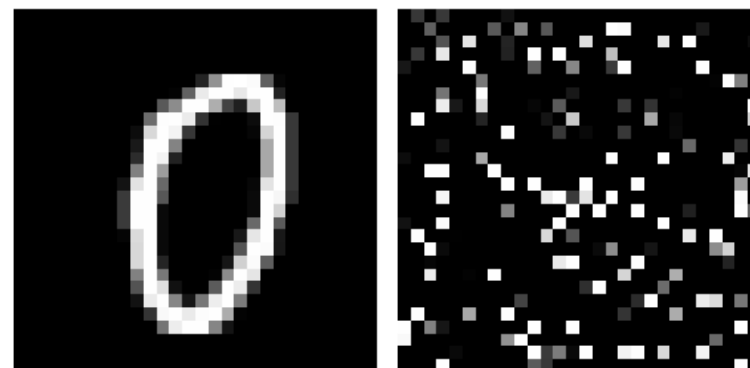
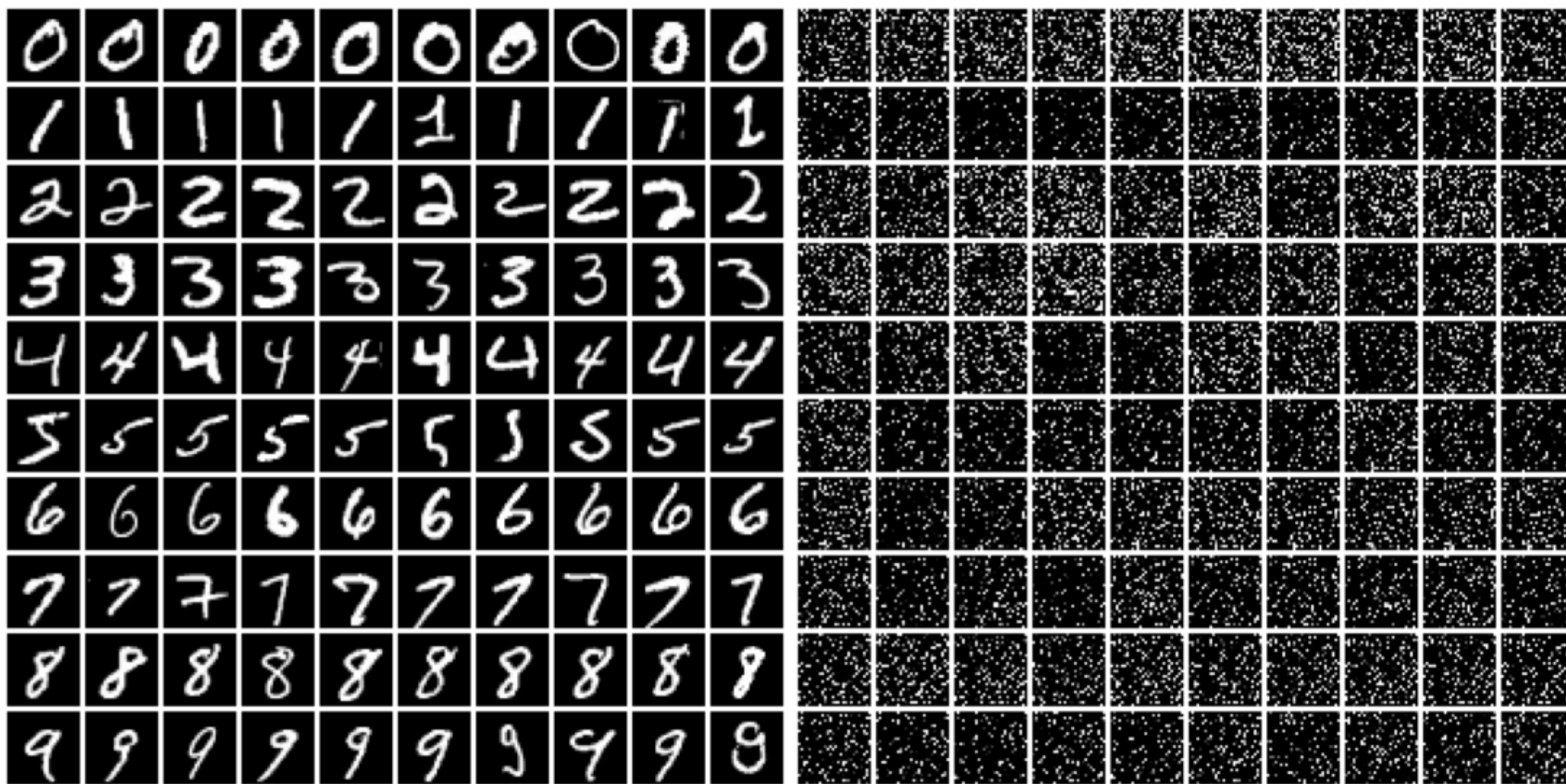
لایه کانولوشنی

- اتصالات تنک (Sparse interactions)
- اشتراک وزن‌ها (Parameter sharing)
- تعداد پارامتر بسیار کمتر از لایه کاملاً متصل
- بازنمایی‌های هم‌تغییر (Equivariant representations)
- توانایی کار با ورودی‌های با ابعاد مختلف
- قابلیت زیاد برای پردازش موازی



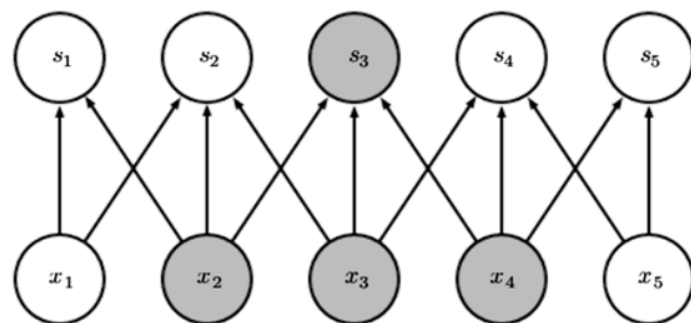
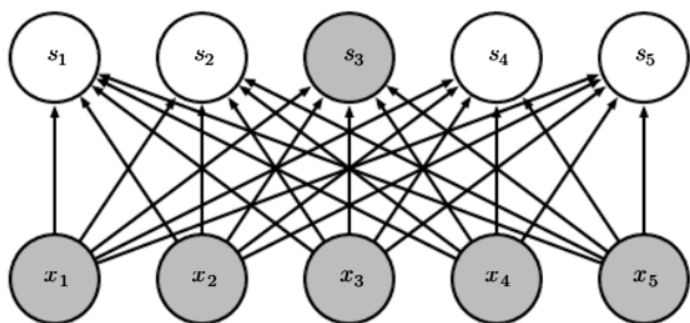
جابجایی مقادیر

- ترتیب مقادیر برای لایه کاملاً متصل هیچ تفاوتی نمی‌کند!



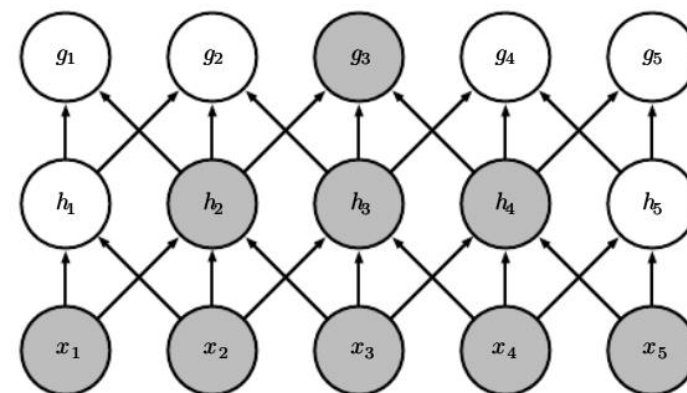
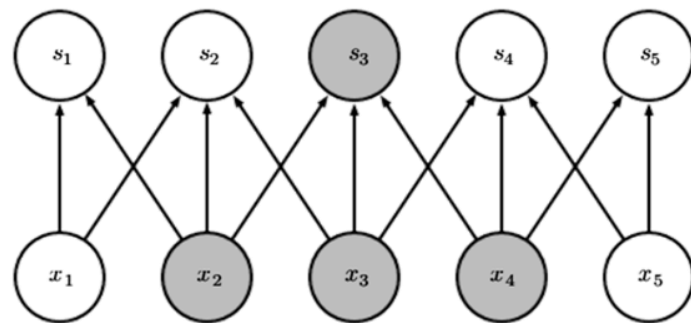
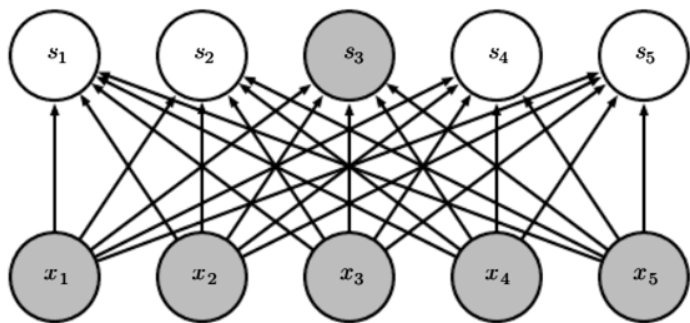
اتصال محلی

- در لایه‌های کاملاً متصل، هر واحد خروجی به تمام واحدهای ورودی متصل است
 - در لایه‌های کانولوشنی، معمولاً هر واحد تنها به برخی از واحدهای ورودی متصل است
 - برای مثال، هنگام پردازش یک تصویر، ورودی ممکن است هزاران یا میلیون‌ها پیکسل داشته باشد، اما می‌توانیم ویژگی‌های کوچک و بامعنایی مانند لبه‌ها را با هسته‌هایی که فقط دهه‌ها یا صدها پیکسل را استفاده می‌کنند، تشخیص بدهیم
- باید پارامترهای بسیار کمتری را ذخیره کنیم



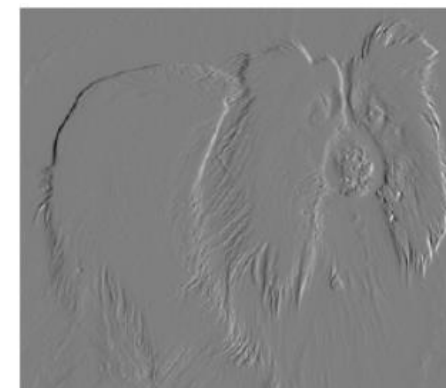
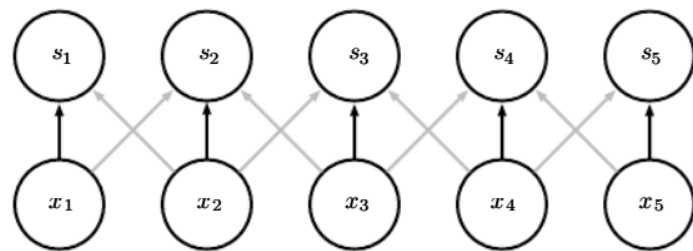
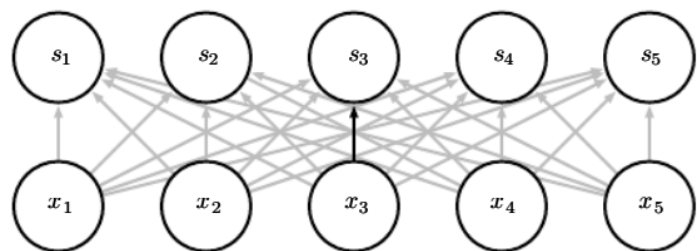
اتصال محلی

- در یک شبکه کانولوشنی عمیق، واحدها در لایه‌های عمیق‌تر ممکن است به طور غیرمستقیم به بخش بزرگ‌تری از ورودی وابستگی داشته باشند
- این کار به شبکه اجازه می‌دهد تا بتواند ویژگی‌های پیچیده را به صورت سلسله‌مراتبی و با استفاده از چنین اتصالات محلی بیاموزد
- میدان تاثیر واحدها در لایه‌های عمیق‌تر می‌تواند بسیار بزرگ باشد



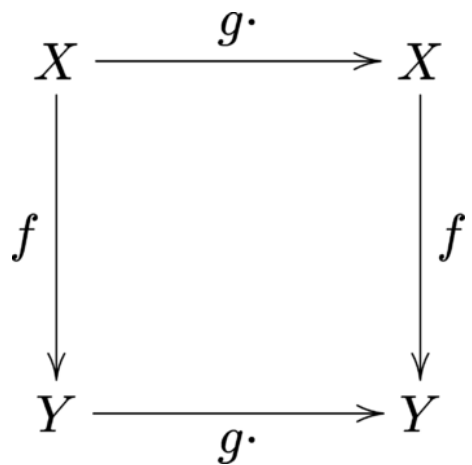
اشتراک وزن‌ها

- اشتراک پارامتر به استفاده از یک پارامتر برای بیش از یک تابع در یک مدل اشاره دارد
- در یک لایه کانولوشنی، هر یک از پارامترهای هسته در هر موقعیت از ورودی استفاده می‌شود
- در برخی موارد، ممکن است نخواهیم پارامترها را به اشتراک بگذاریم!



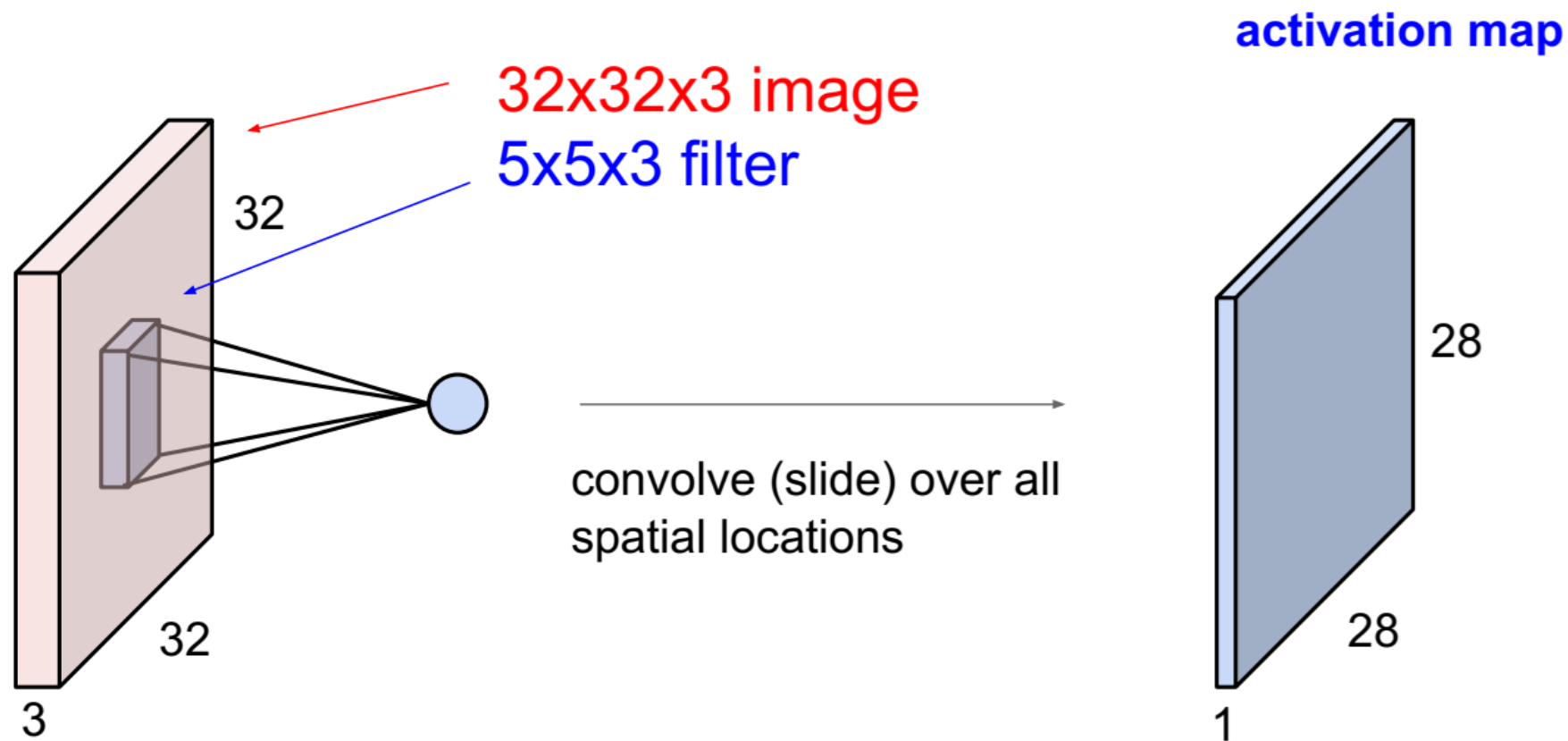
بازنمایی‌های هم‌تغییر

- هم‌تغییر بودن یک تابع به این معنی است که اگر ورودی تغییر کند، خروجی نیز به همان صورت تغییر می‌کند
- در مورد کانولوشن، شکل خاص به اشتراک‌گذاری پارامترها باعث می‌شود که لایه نسبت به جابجایی هم‌تغییر شود
- کانولوشن به طور طبیعی نسبت به برخی از تبدیل‌های دیگر، مانند تغییر در مقیاس یا چرخش یک تصویر، هم‌تغییر نیست



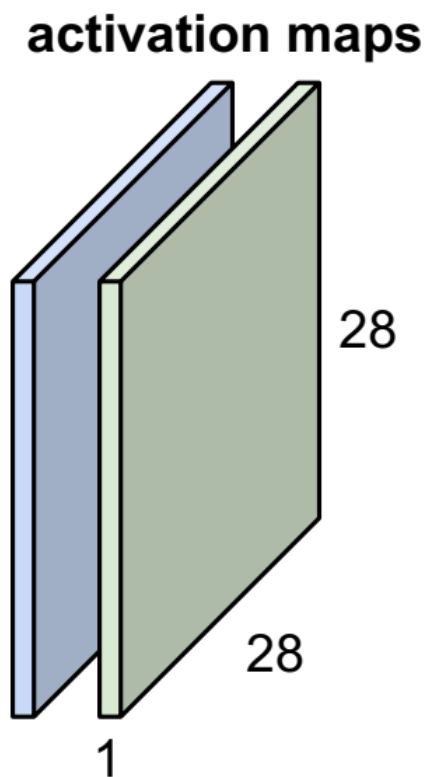
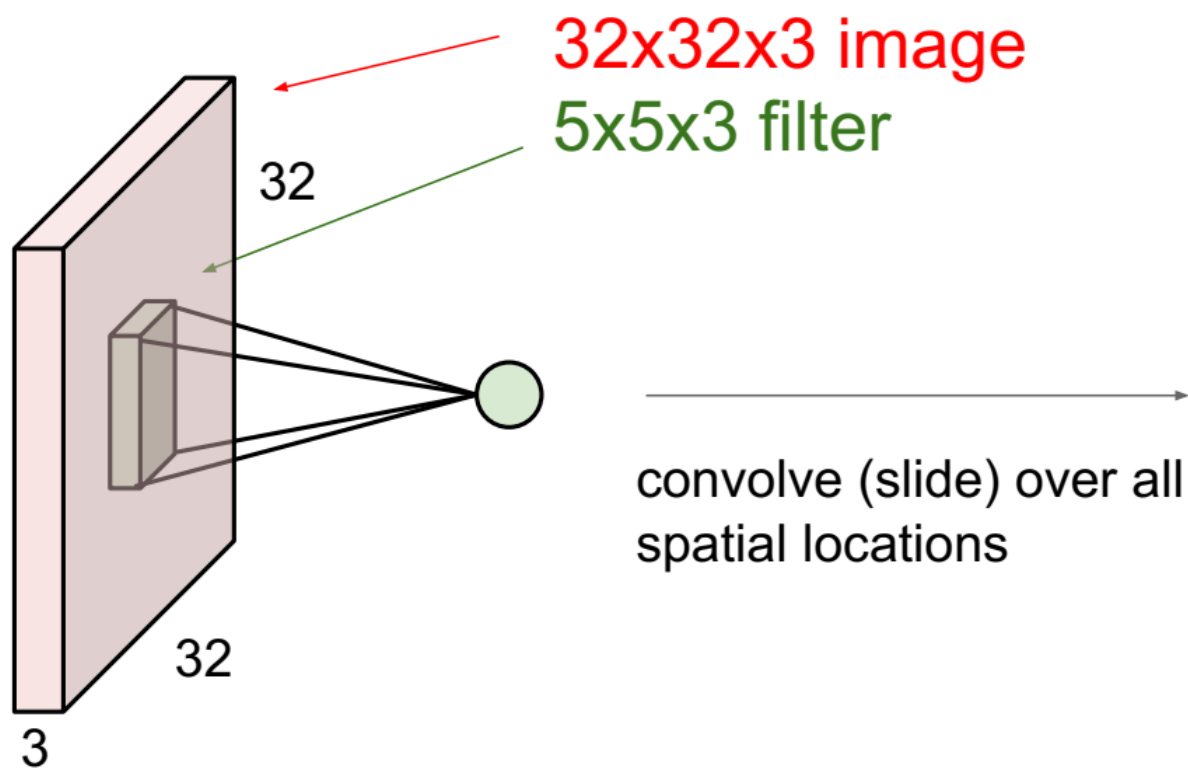
$$f(g(x)) = g(f(x))$$

لایه کانولوشنی



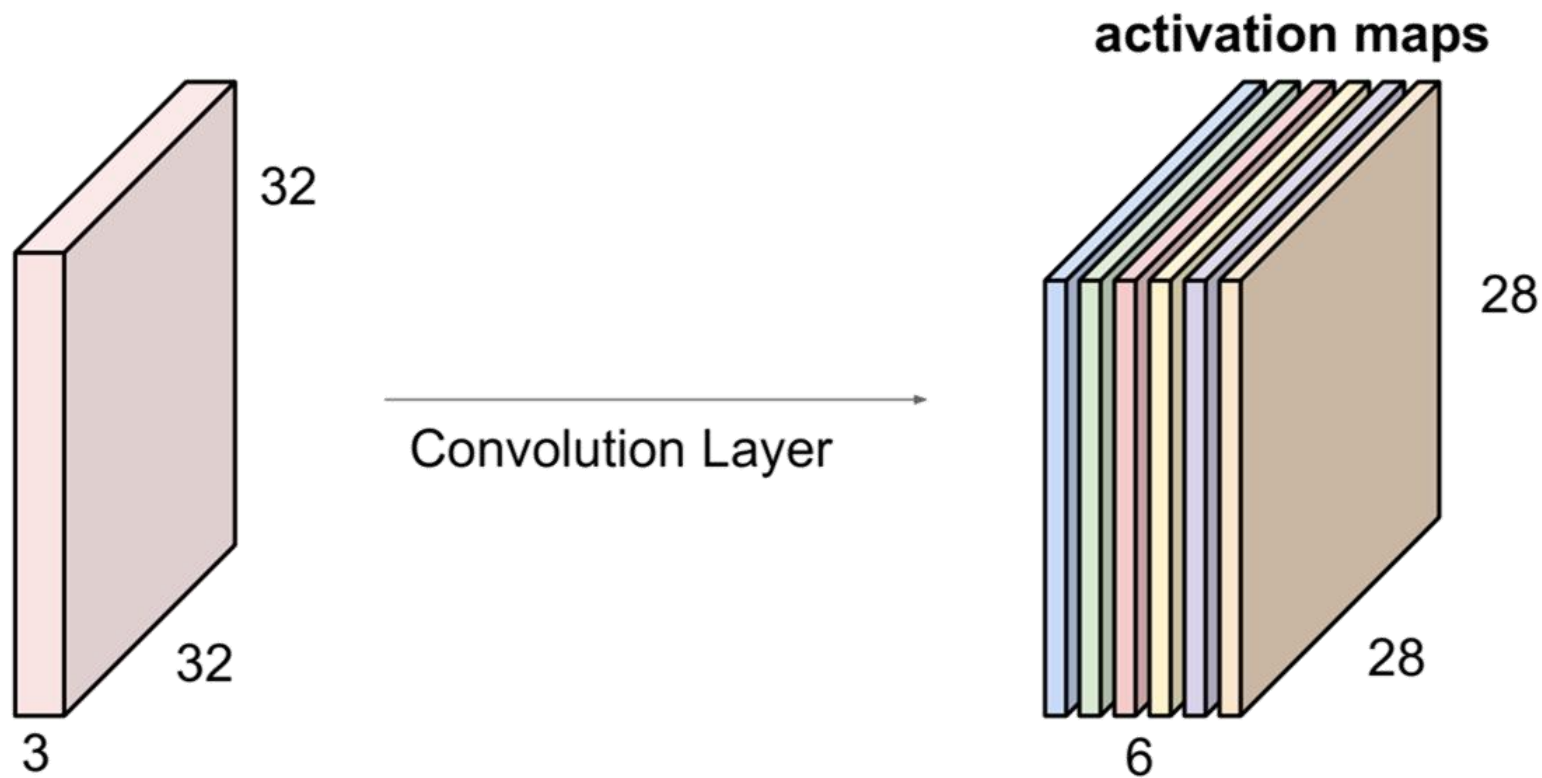
لایه کانولوشنی

- یک فیلتر دوم را در نظر بگیرید (رنگ سبز)



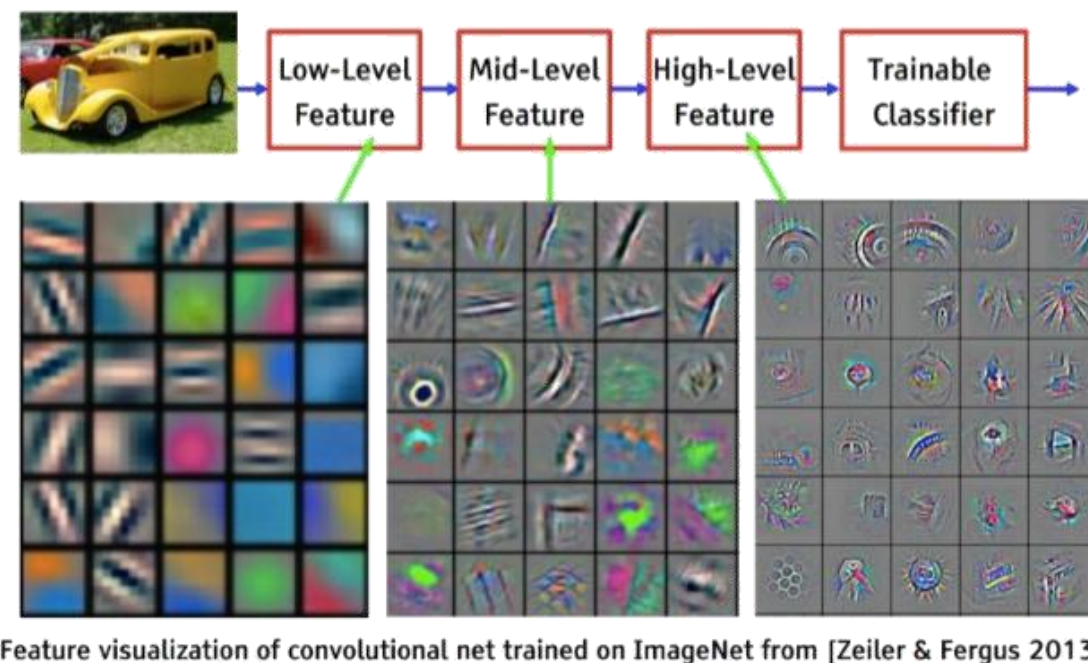
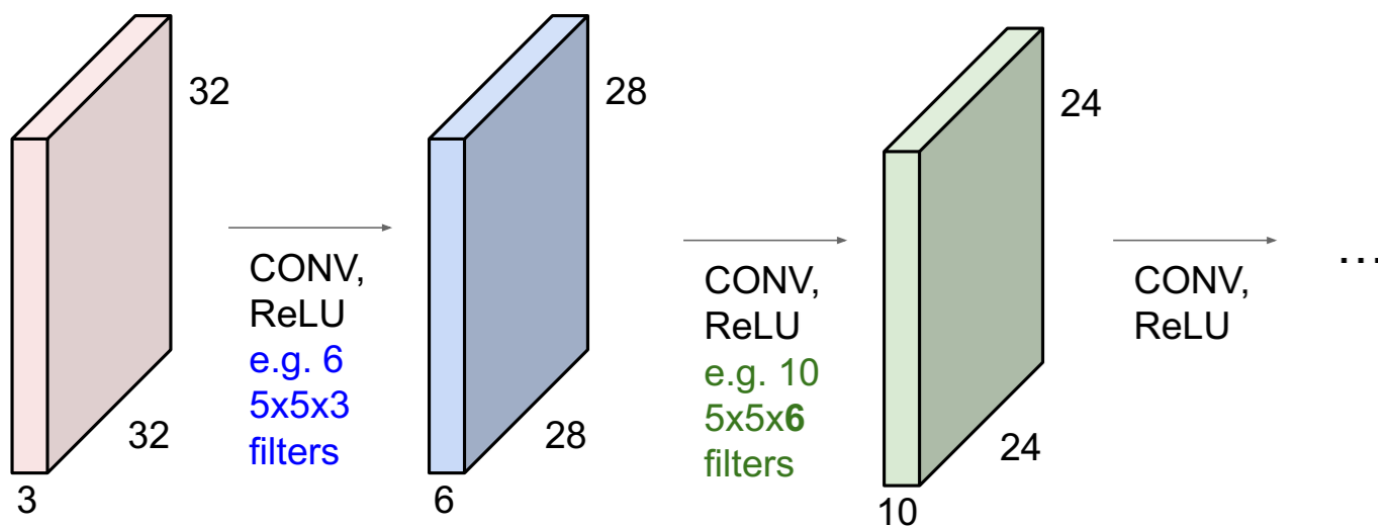
لایه کانولوشنی

- در لایه کانولوشن از چند فیلتر مجزا استفاده می کنیم



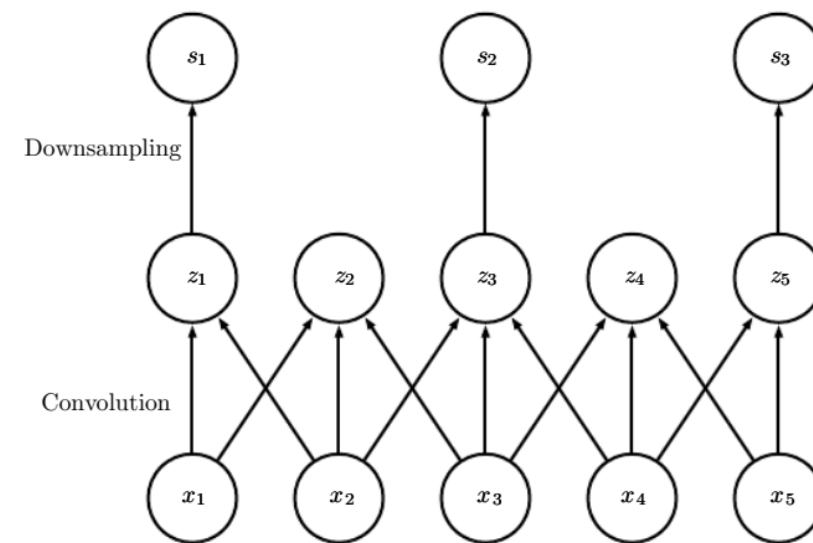
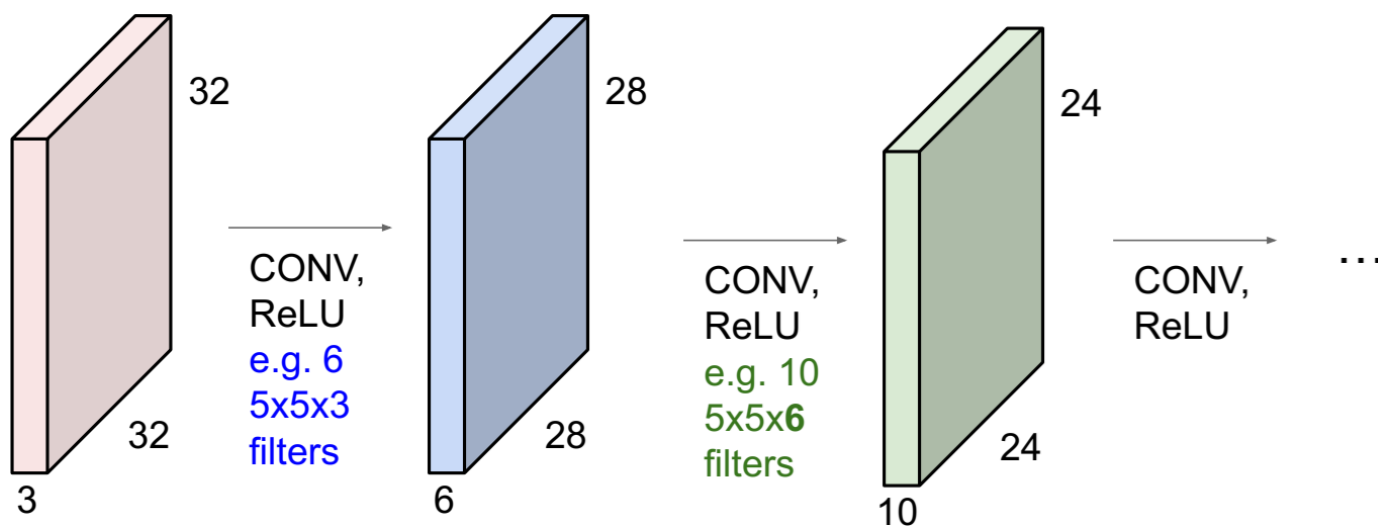
شبکه کانولوشنی (ConvNet)

- شبکه کانولوشنی دنباله‌ای از لایه‌های کانولوشنی با توابع فعال‌سازی است



گام (Stride)

- برای کاهش هزینه محاسباتی می‌توانیم از برخی موقعیت‌ها پرش کنیم
 - به قیمت استخراج نشدن ویژگی‌ها با رزولوشن کامل
 - به نوعی downsampling انجام می‌شود



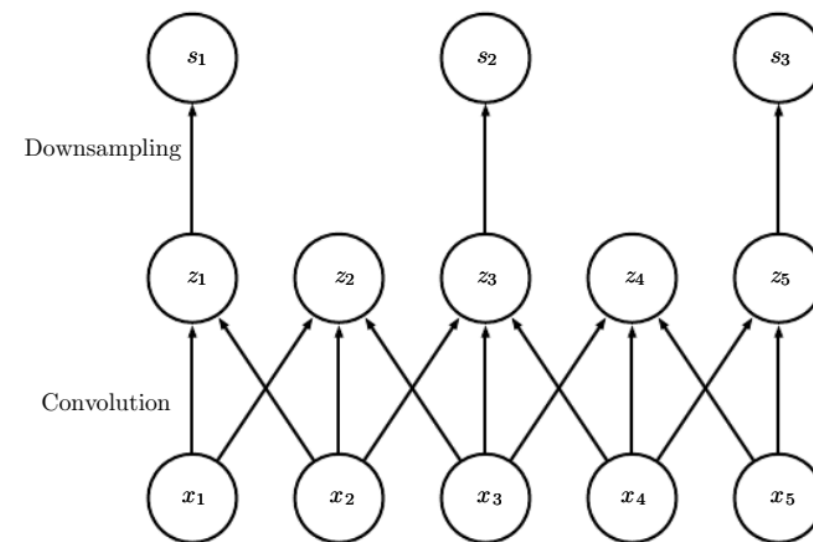
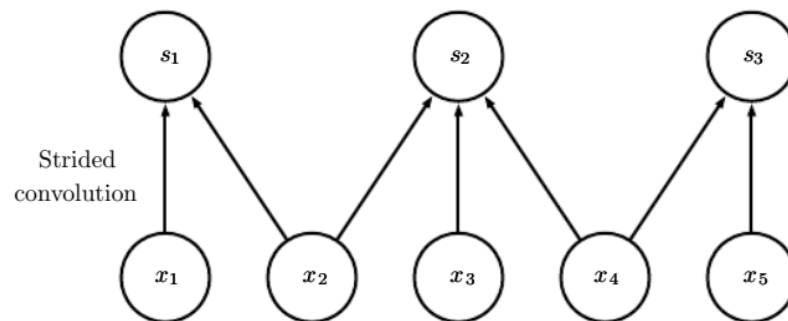
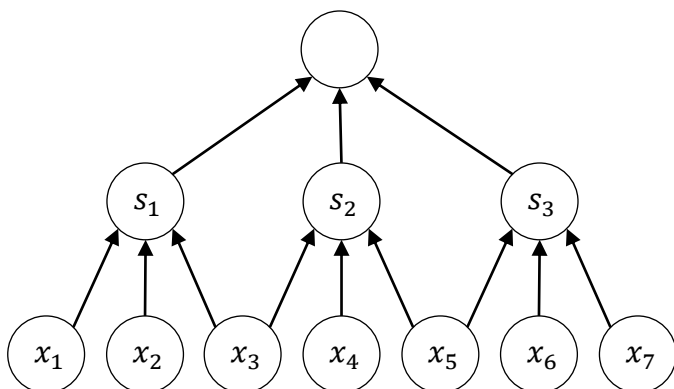
گام (Stride)

• برای کاهش هزینه محاسباتی می‌توانیم از برخی موقعیت‌ها پرش کنیم

- به قیمت استخراج نشدن ویژگی‌ها با رزولوشن کامل

- به نوعی downsampling انجام می‌شود

- باعث افزایش میدان تاثیر می‌شود



ادغام (Pooling)

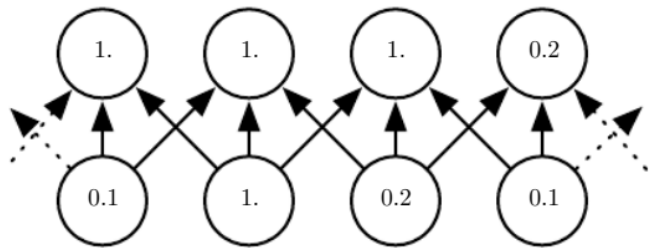
- یک تابع ادغام، خروجی شبکه در یک موقعیت مشخص را با یک مشخصه آماری از مقادیر در همسایگی آن جایگزین می‌کند

- به عنوان مثال، عملیات ادغام حداکثر (Max Pooling)، مقدار بیشینه در یک محدوده مشخص را محاسبه می‌کند

- سایر توابع ادغام پرکاربرد عبارتند از میانگین، نرم L2، انحراف معیار، و میانگین وزن‌دار

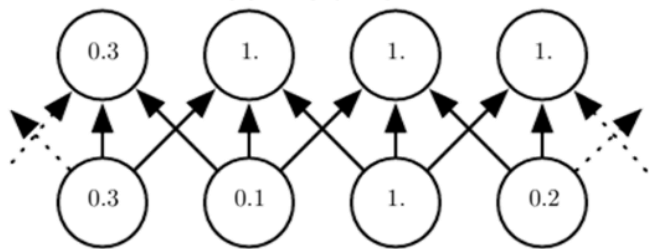
- در همه موارد، ادغام کمک می‌کند تا بازنمایی نسبت به جابجایی‌های کوچک ورودی حساسیت کمتری داشته باشد

POOLING STAGE



DETECTOR STAGE

POOLING STAGE



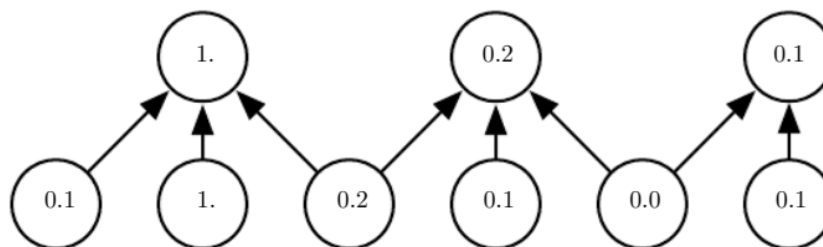
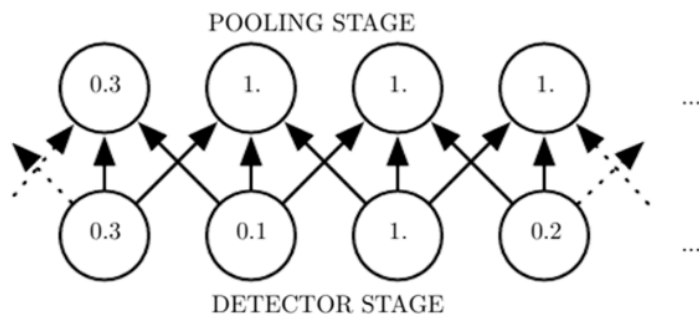
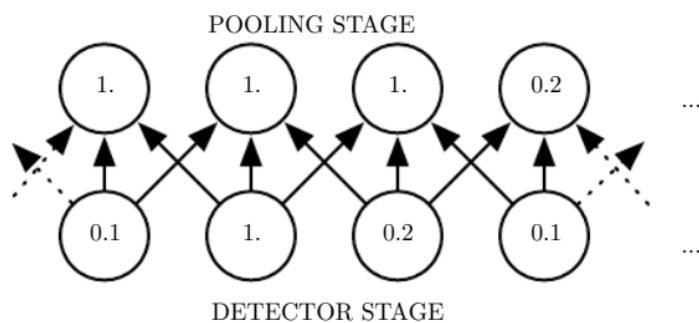
DETECTOR STAGE

ادغام (Pooling)

- یک تابع ادغام، خروجی شبکه در یک موقعیت مشخص را با یک مشخصه آماری از مقادیر در همسایگی آن جایگزین می‌کند

- می‌توانیم پس از ادغام مقادیر، برای کاهش محاسبات بعدی و افزایش کارایی، برخی مقادیر را حذف کنیم

- این کاهش در اندازه ورودی لایه‌های کاملاً متصل می‌تواند منجر به بهبود بازدهی آماری و کاهش نیاز به حافظه برای ذخیره‌سازی پارامترها شود



مقایسه

Model	Parameters	Batch Time (ms)	Train Loss	Train Acc.	Test Loss	Test Acc.
MLP	1,746,510	0.696	1.193	0.572	1.452	0.499
CNN	978,126	17.5	0.014	0.999	2.819	0.631
CNN+Stride	272,846	5.28	0.073	0.987	1.767	0.657
CNN+Pool	272,846	7.65	0.043	0.995	1.317	0.741