

شبکه‌های عصبی کانولوشن

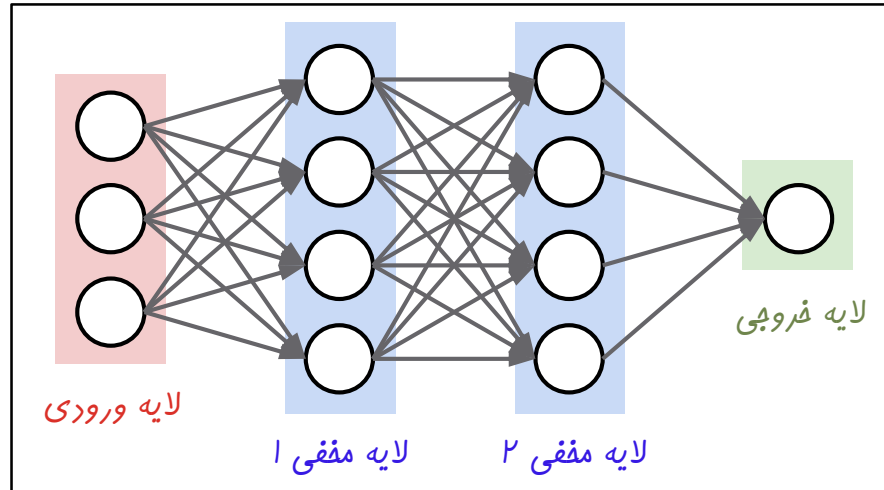
سید ناصر رضوی www.snrazavi.ir

۱۳۹۷

یادآوری: بهینه‌سازی

۲

□ گرادیان کاهشی. [با دسته‌های کوچک]

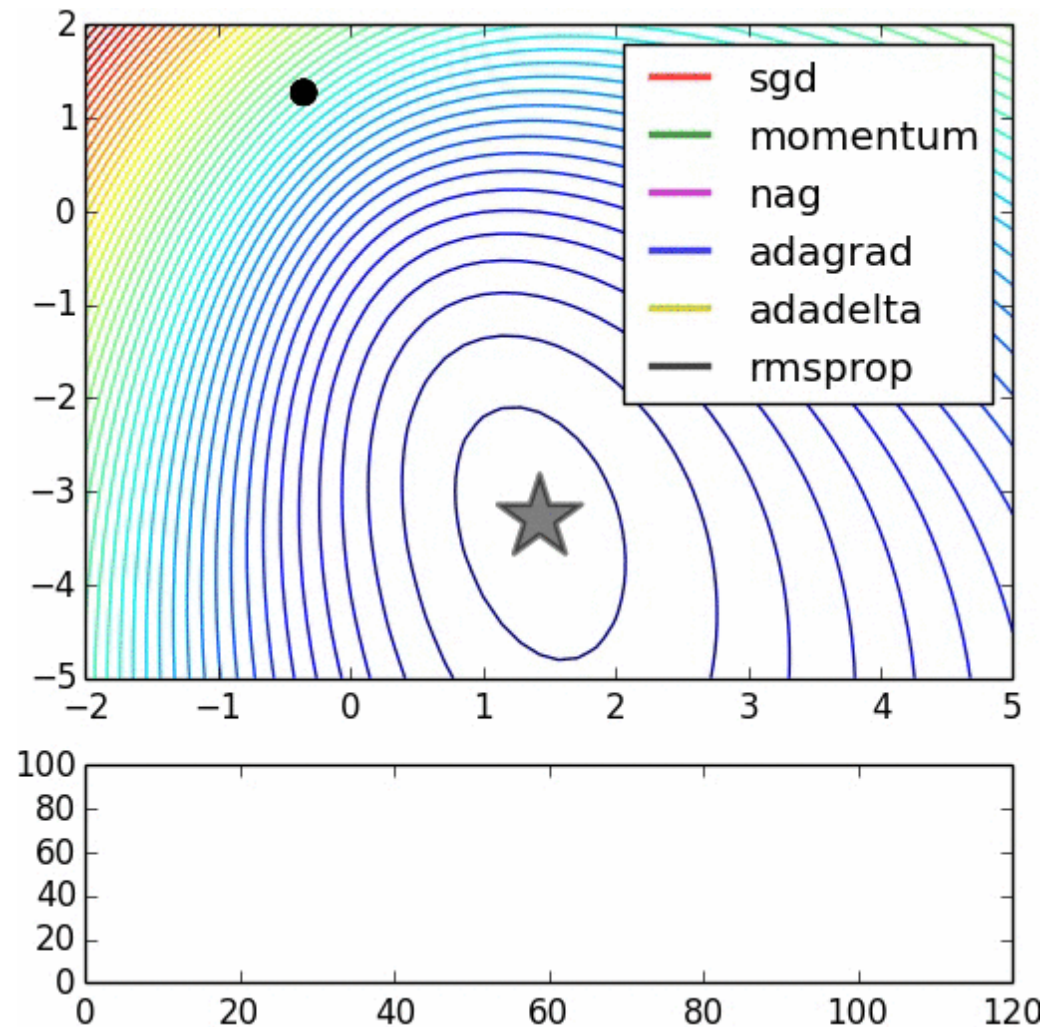


حلقه:

۱. یک دسته از داده‌ها را به طور تصادفی انتخاب کن.
۲. محاسبات رو به جلو را در طول گراف انجام بده و مقدار تابع هزینه را محاسبه کن.
۳. محاسبات رو به عقب را به منظور محاسبه گرادیان‌ها انجام بده.
۴. مقدار پارامترها را با استفاده از گرادیان‌های محاسبه شده به روز رسانی کن.

یادآوری: روش‌های به روز رسانی پارامترها

۳



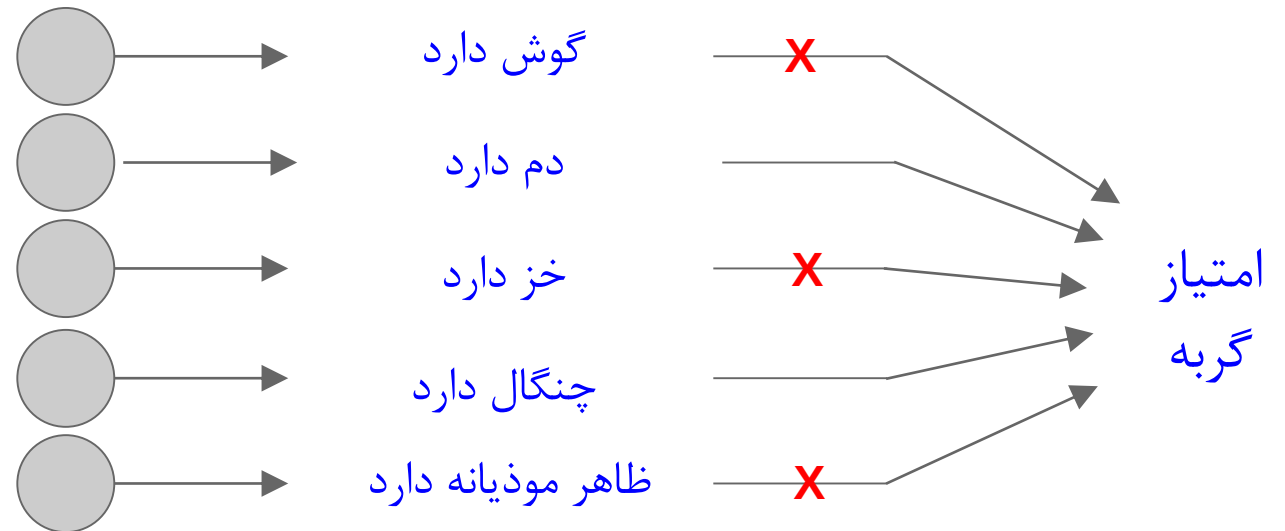
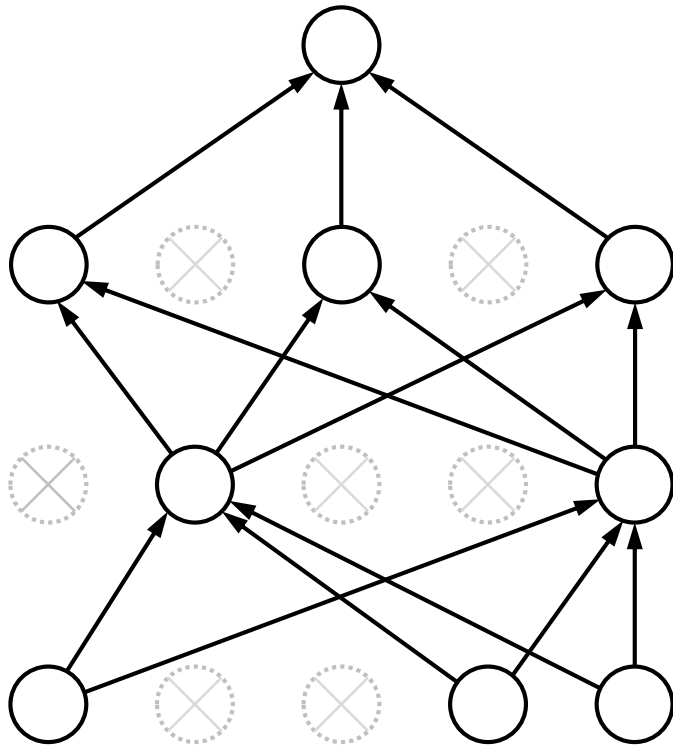
□ به روز رسانی پارامترها.

یادآوری: دورریزی [Dropout]

۴

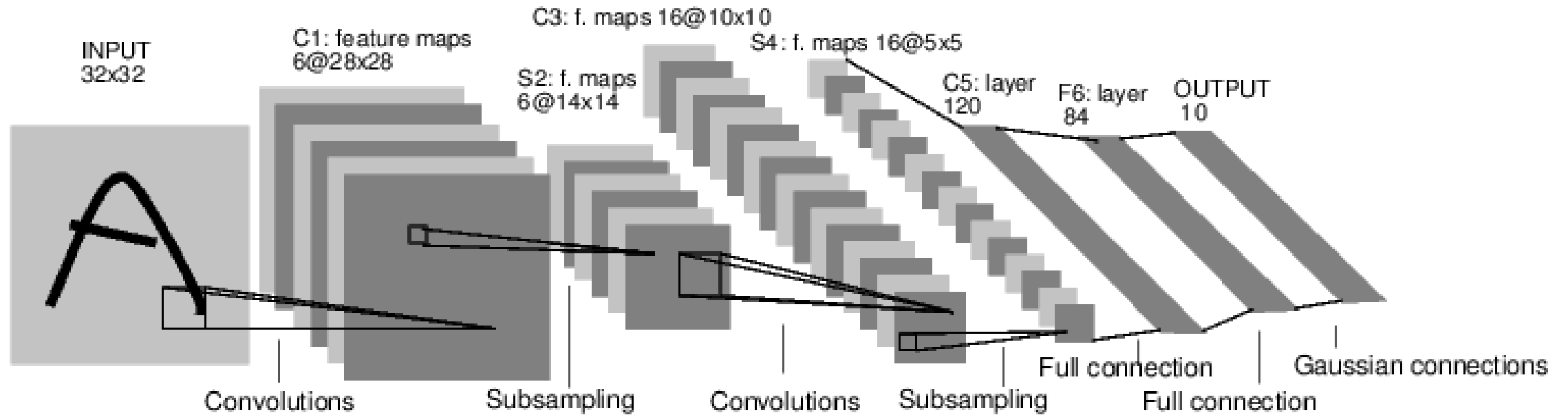
□ چگونه ممکن است حذف تصادفی برخی نورون‌ها ایده خوبی باشد؟

اجبار شبکه به داشتن بازنمایی دارای افزونگی



شبکه‌های عصبی کانولوشن

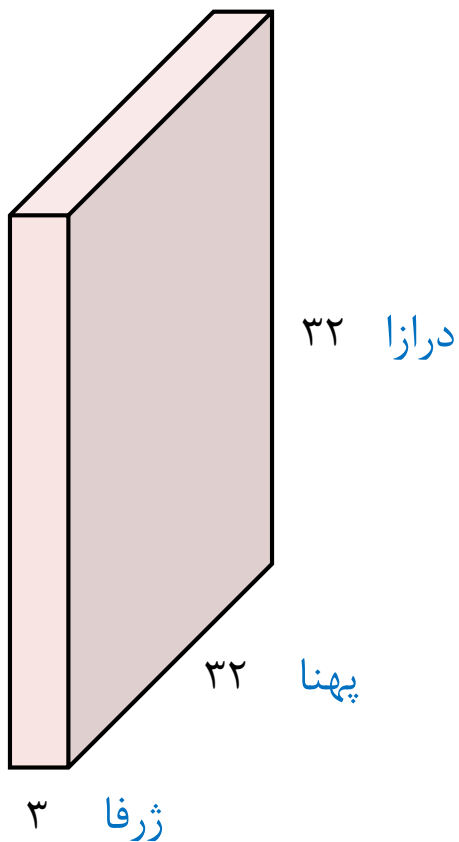
۵



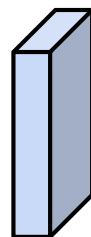
لایه کانولوشن

۶

تصویر $32 \times 32 \times 3$



فیلتر $5 \times 5 \times 3$



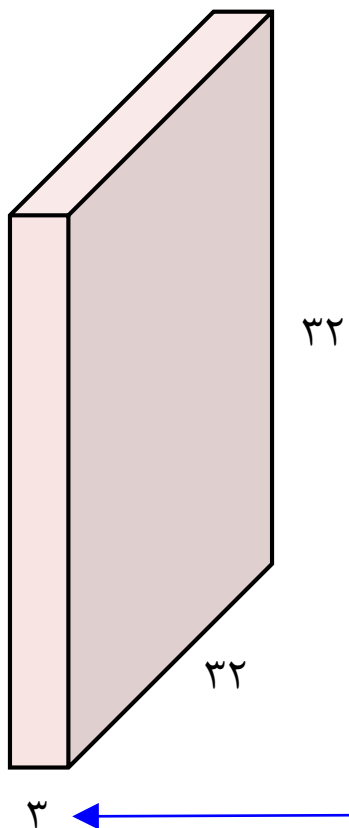
کانولوشن فیلتر با تصویر

یعنی «فیلتر را بر روی تصویر حرکت بده و هر بار ضرب داخلی فیلتر و بخشی از تصویر که فیلتر بر روی آن قرار دارد را محاسبه کن.»

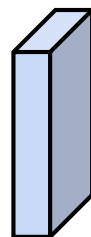
لایه کانولوشن

۷

تصویر $32 \times 32 \times 3$



فیلتر $5 \times 5 \times 3$



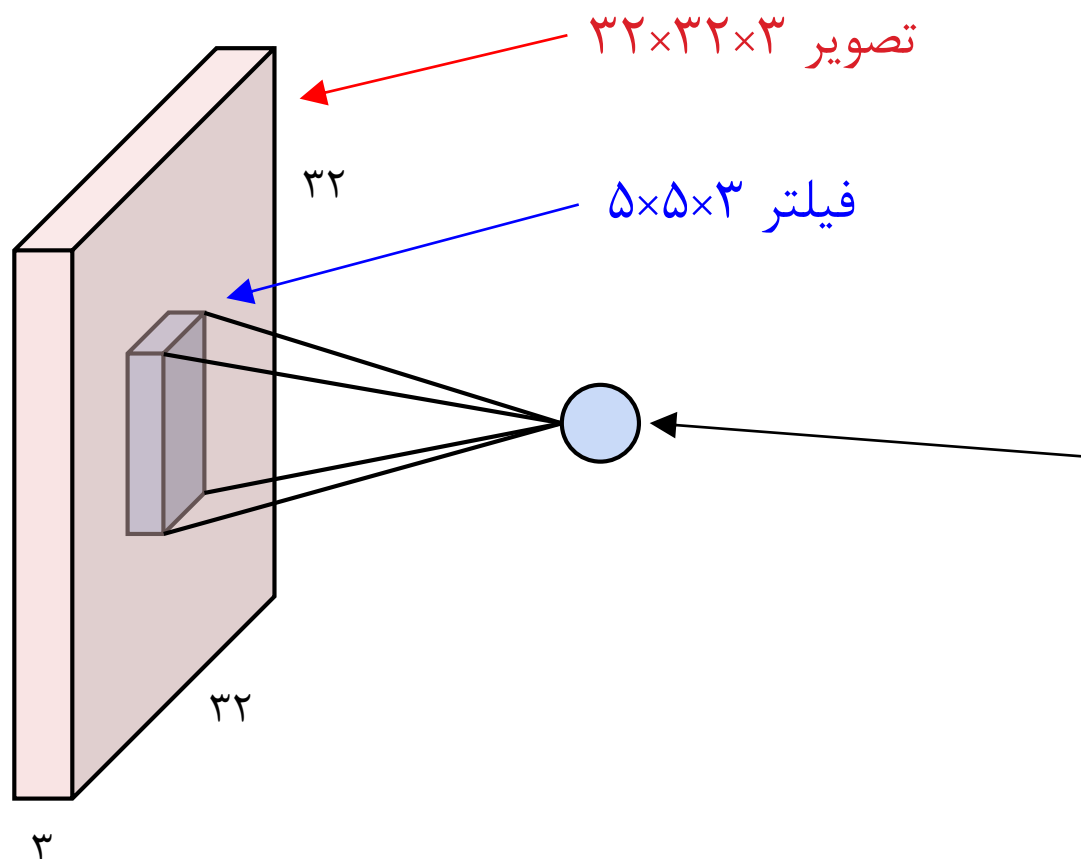
همیشه عمق تصویر و فیلتر
با یکدیگر برابر هستند

کانولوشن فیلتر با تصویر

یعنی «فیلتر را بر روی تصویر حرکت بده
و هر بار ضرب داخلی فیلتر و بخشی از
تصویر که فیلتر بر روی آن قرار دارد را
محاسبه کن.»

لایه کانولوشن

۸



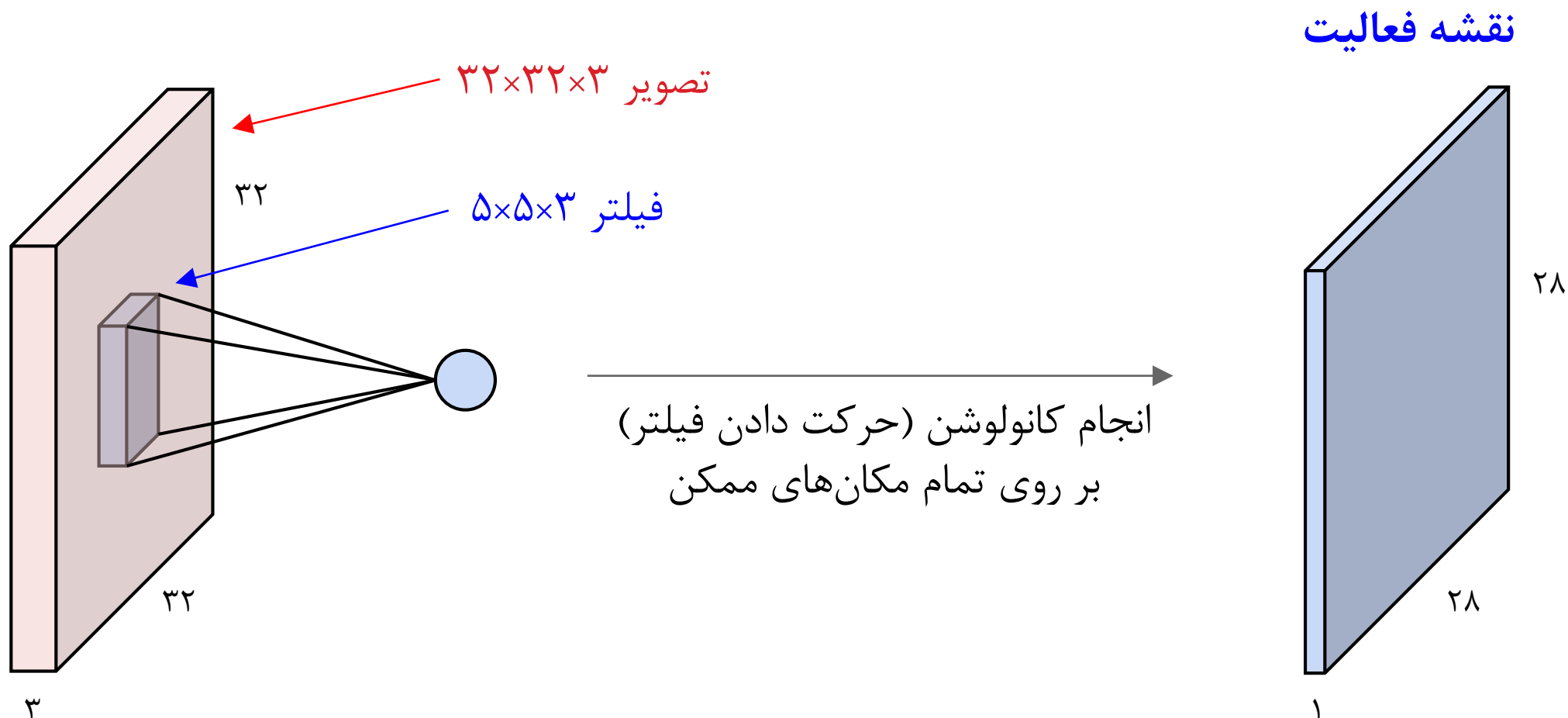
۱ عدد:

به دست آمده از ضرب داخلی فیلتر و بخش کوچکی از تصویر که فیلتر بر روی آن قرار دارد. (یعنی، ضرب داخلی دو بردار ۷۵ بعدی + بایاس)

$$w^T x + b$$

لایه کانولوشن

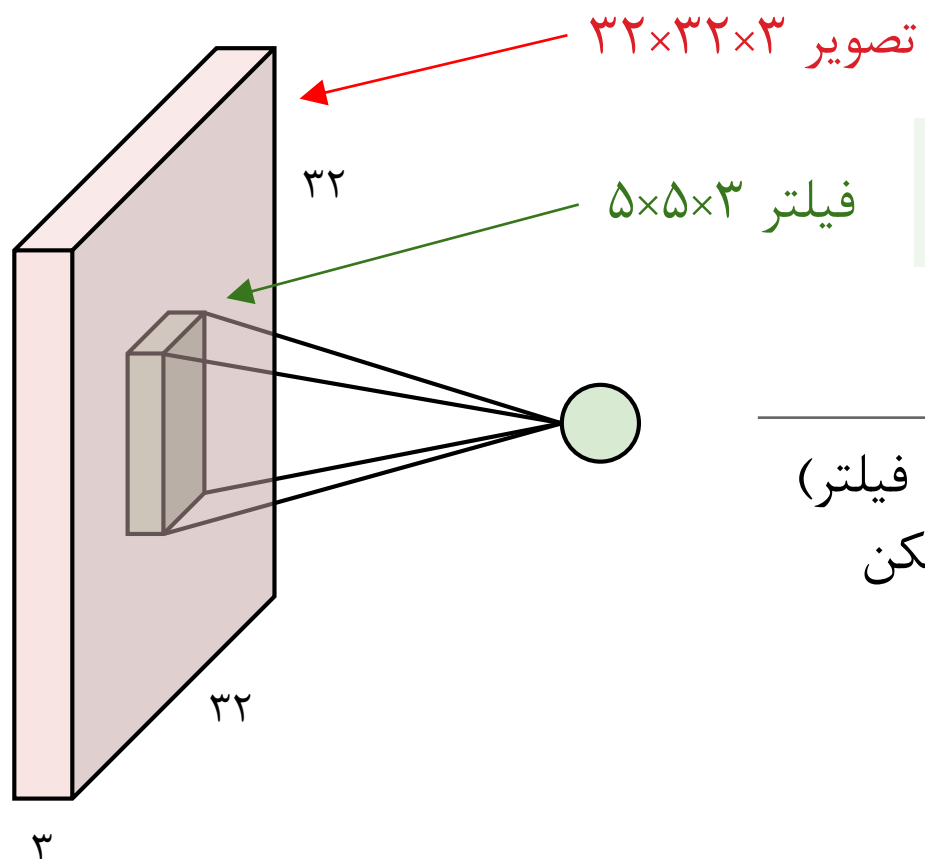
۹



لایه کانولوشن

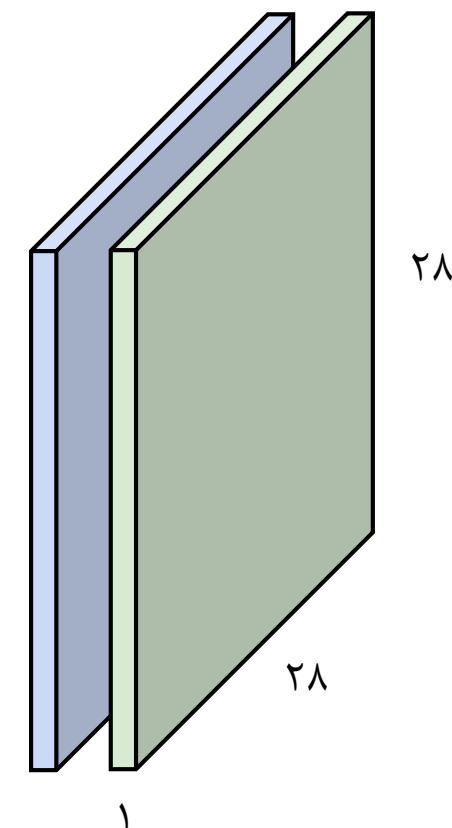
۱۰

نقشه‌های فعالیت



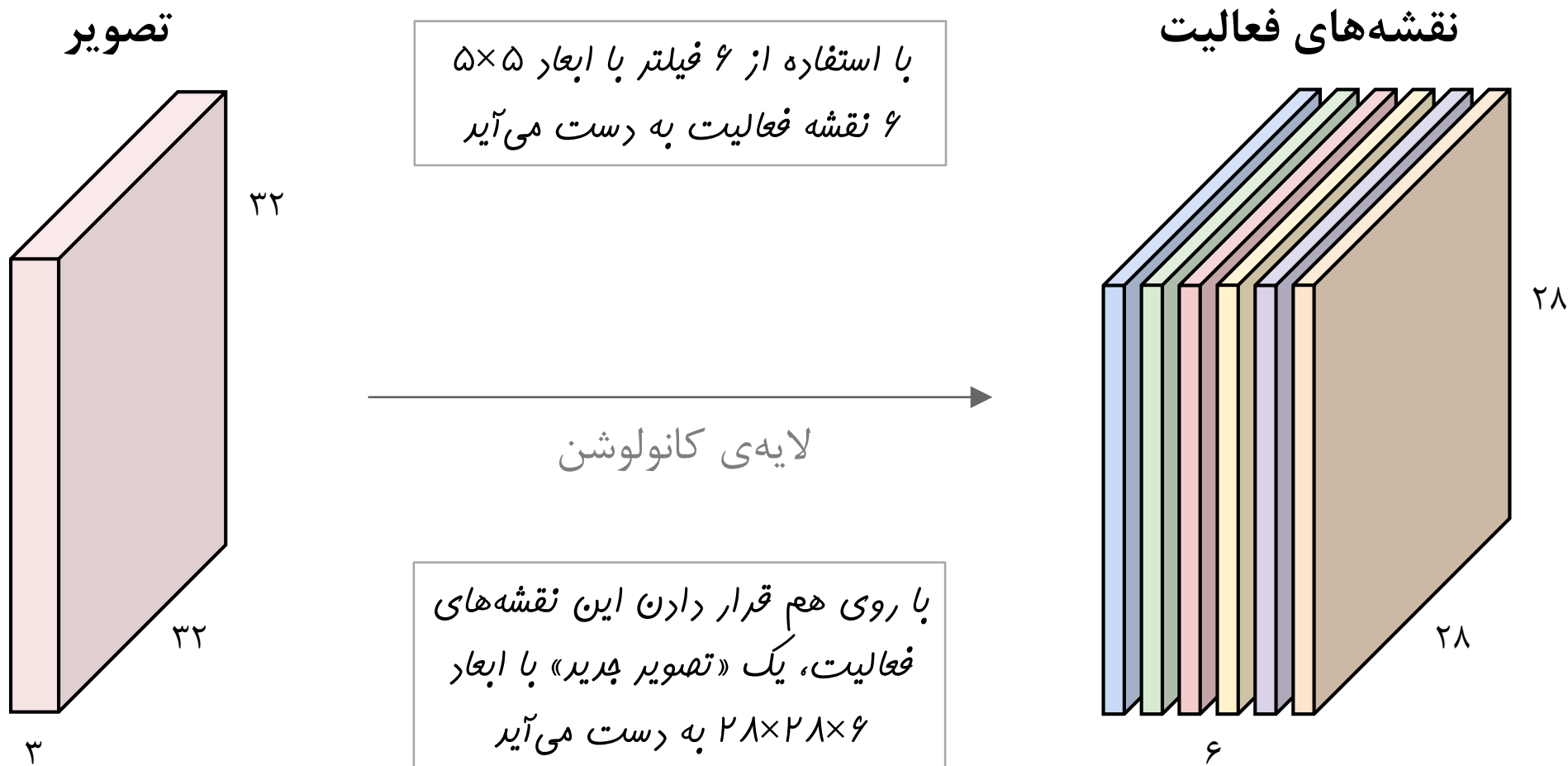
یک فیلتر دیگر با وزن‌های
متفاوت نسبت به فیلتر قبلی

انجام کانولوشن (حرکت دادن فیلتر)
بر روی تمام مکان‌های ممکن



لایه کانولوشن

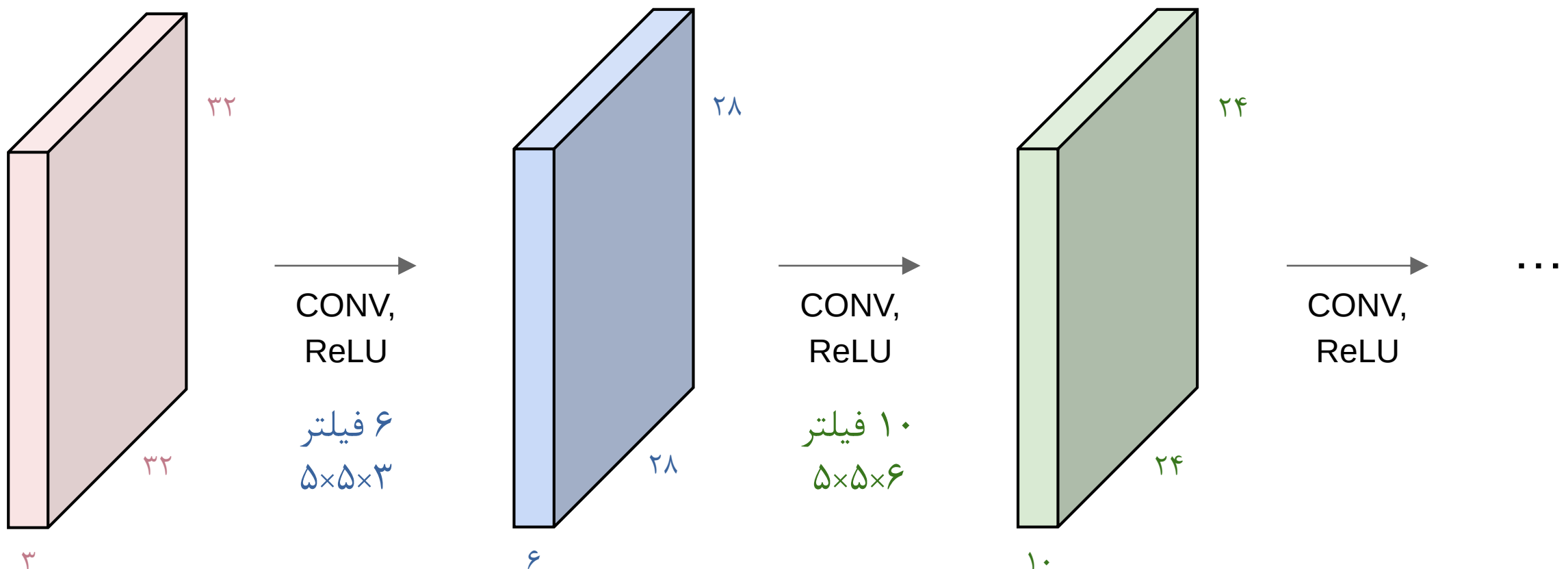
۱۱



شبکه‌های عصبی کانولوشن

۱۲

□ شبکه‌های عصبی کانولوشن. یک دنباله از لایه‌های کانولوشن و توابع فعالیت در میان آنها.



شبکه‌های عصبی کانولوشن

۱۳

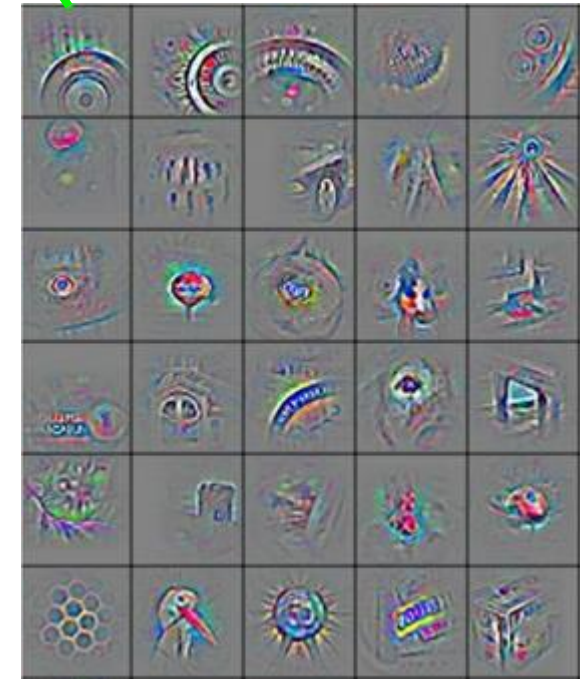
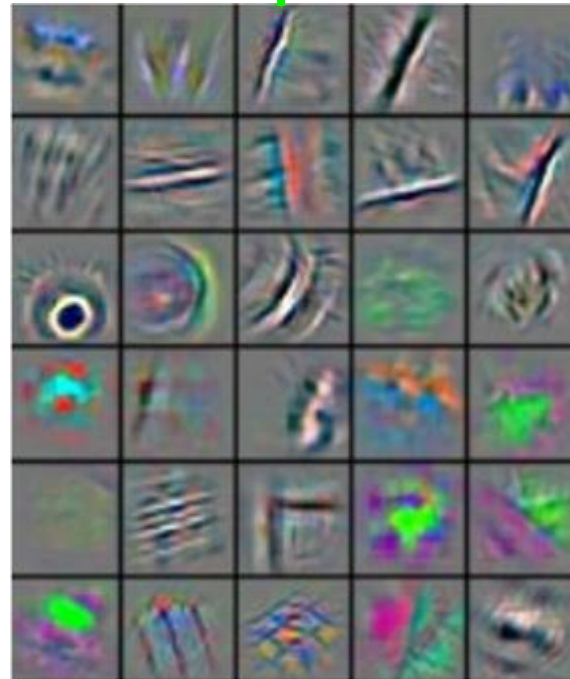


ویژگی‌های
سطح پایین

ویژگی‌های
سطح میانی

ویژگی‌های
سطح بالا

دسته‌بند
قابل آموزش



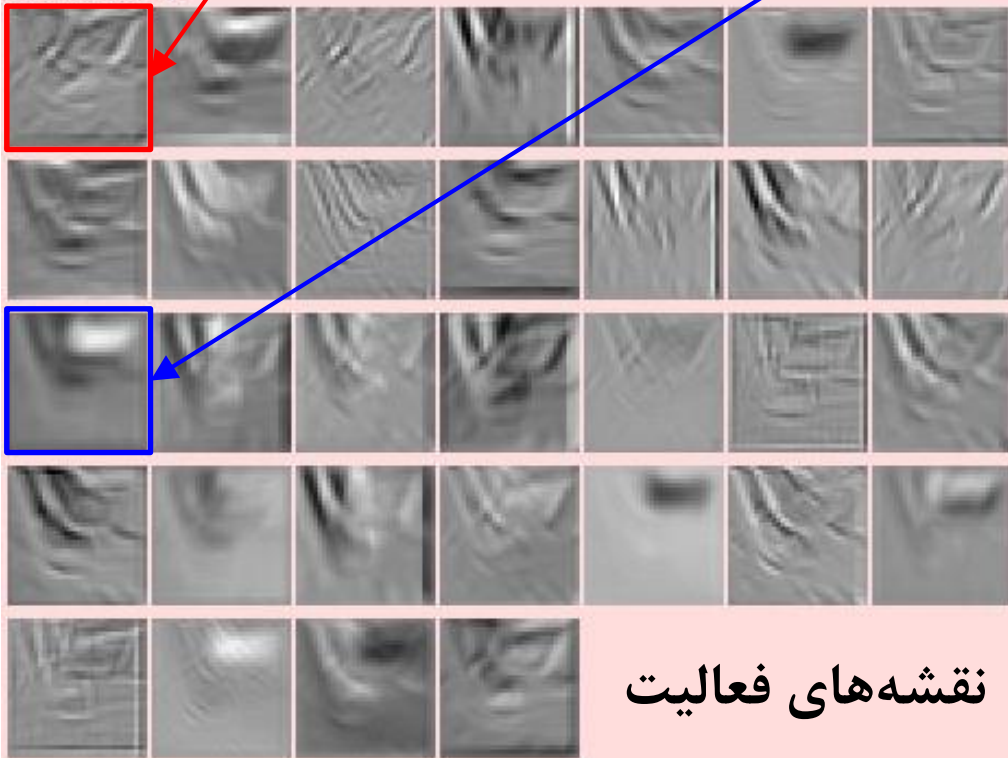
شبکه‌های عصبی کانولوشن

۱۴

Activations:



Activations:



نقشه‌های فعالیت

فیلترهای نمونه با ابعاد 5×5 [۳۲ فیلتر]

[هر فیلتر $\Leftarrow$ یک نقشه‌ی فعالیت]

این لایه‌ها را به این دلیل که با کانولوشن دو سیگنال مرتبط هستند، **لایه‌های کانولوشن** می‌نامیم:

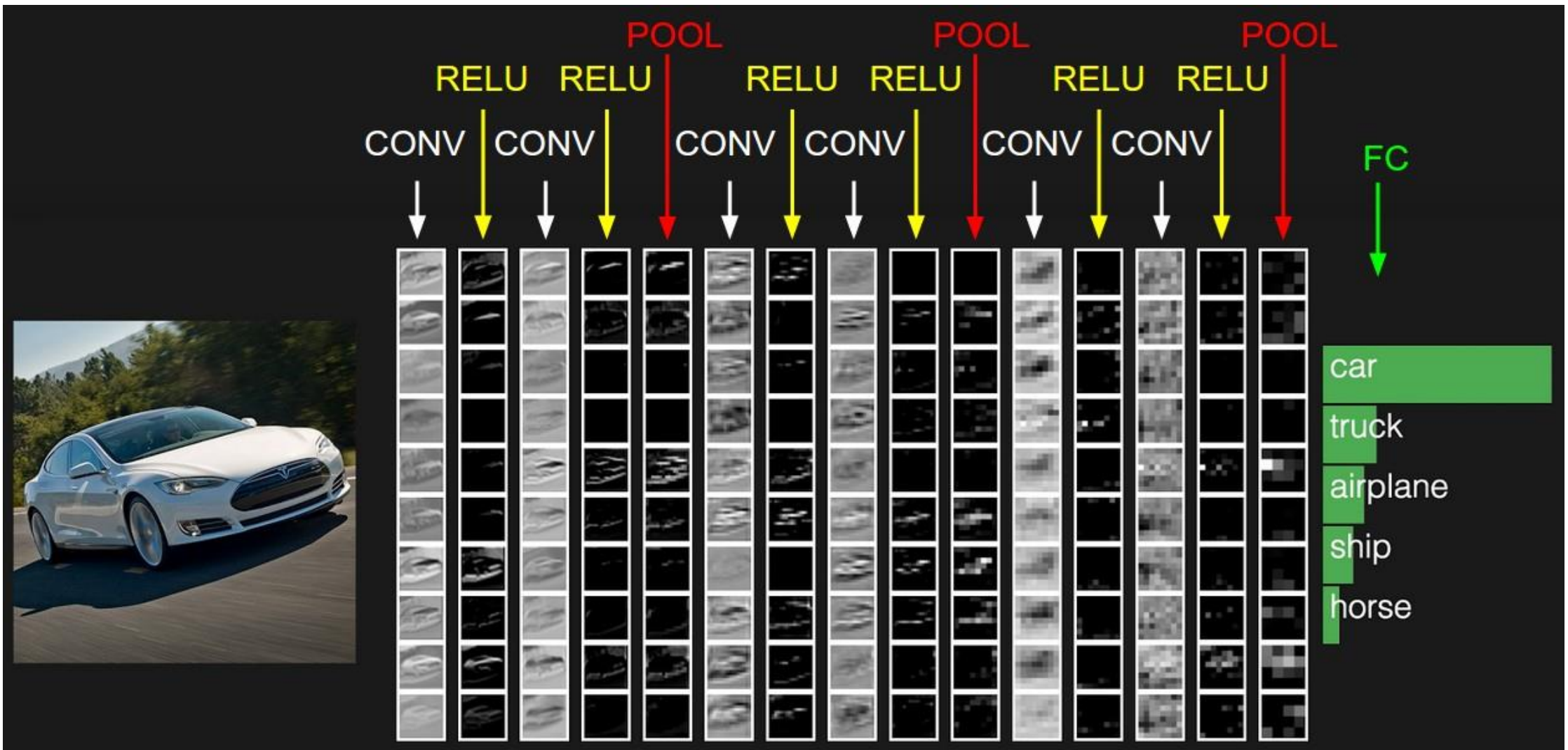
$$f[x, y] * g[x, y] = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1, n_2] \cdot g[x - n_1, y - n_2]$$



ضرب عنصر به عنصر و جمع یک فیلتر و سیگنال

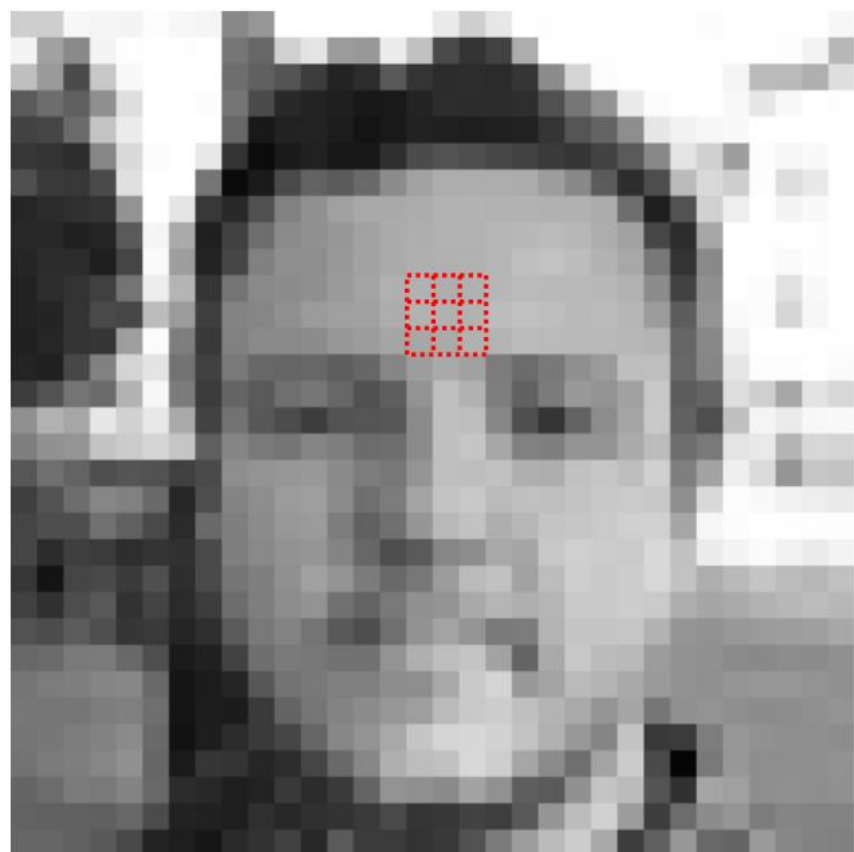
شبکه‌های عصبی کانولوشن

۱۵



لایه کانولوشن: یک نگاه دقیق‌تر

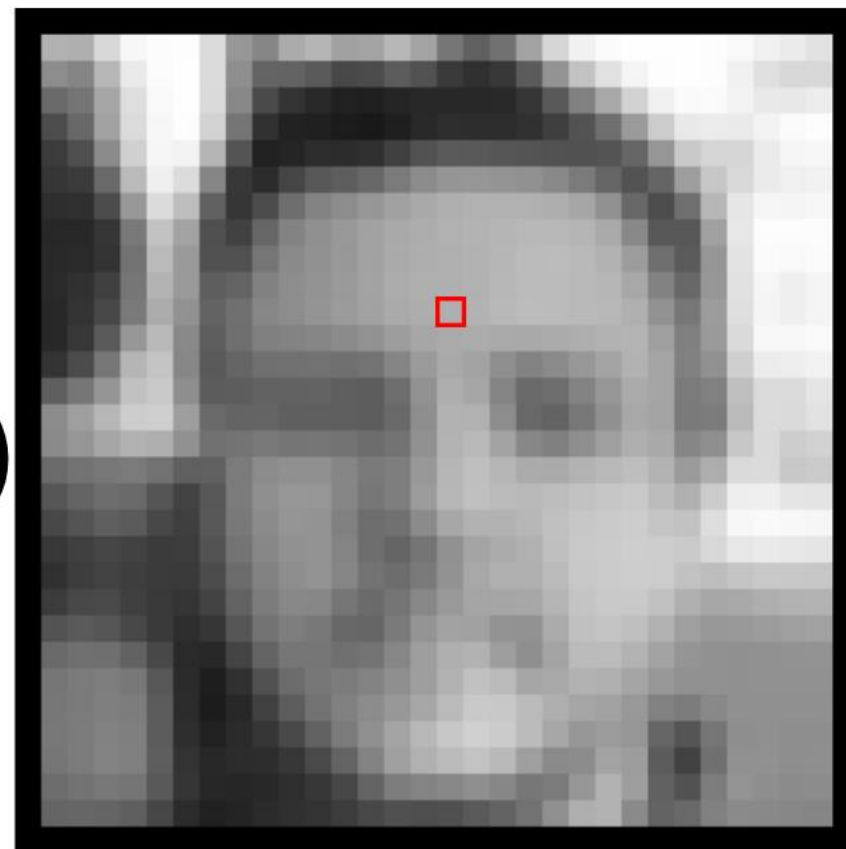
۱۶



input image

$$\begin{pmatrix} 178 + 174 + 177 \\ \times 0.0625 \times 0.125 \times 0.0625 \\ + 179 + 178 + 179 \\ \times 0.125 \times 0.25 \times 0.125 \\ + 168 + 171 + 178 \\ \times 0.0625 \times 0.125 \times 0.0625 \end{pmatrix} = 176$$

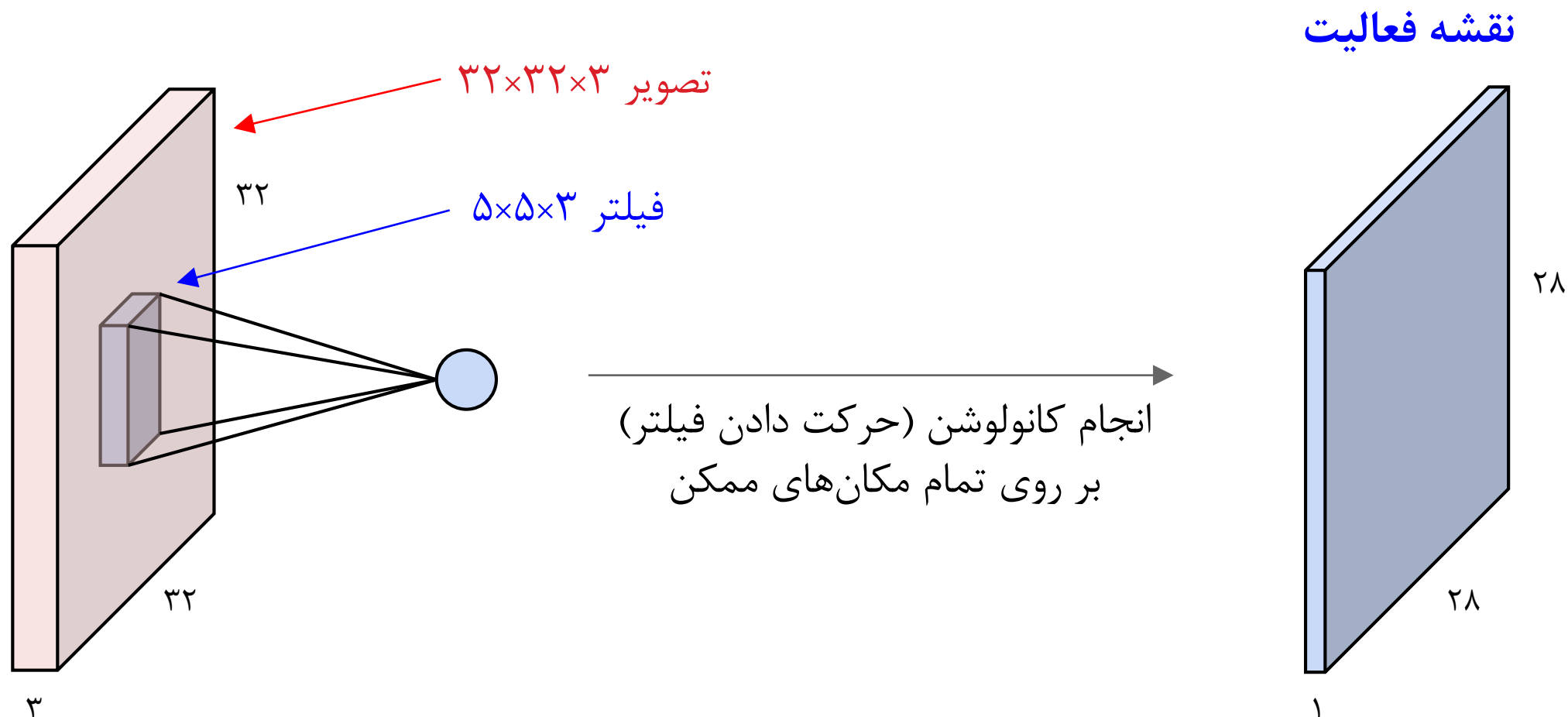
kernel:



output image

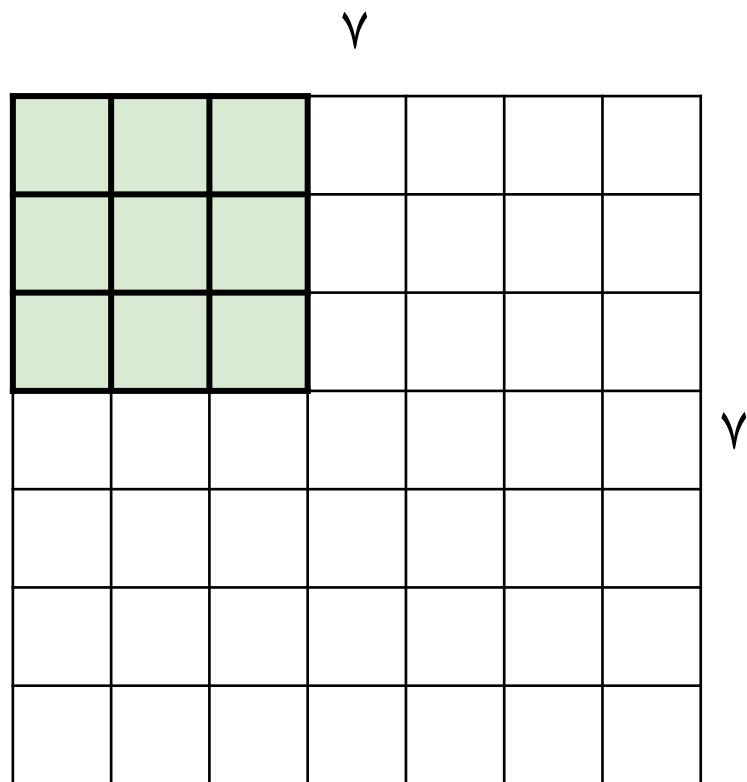
لایه کانولوشن: یک نگاه دقیق‌تر

۱۷



لایه کانولوشن: یک نگاه دقیق‌تر

۱۸

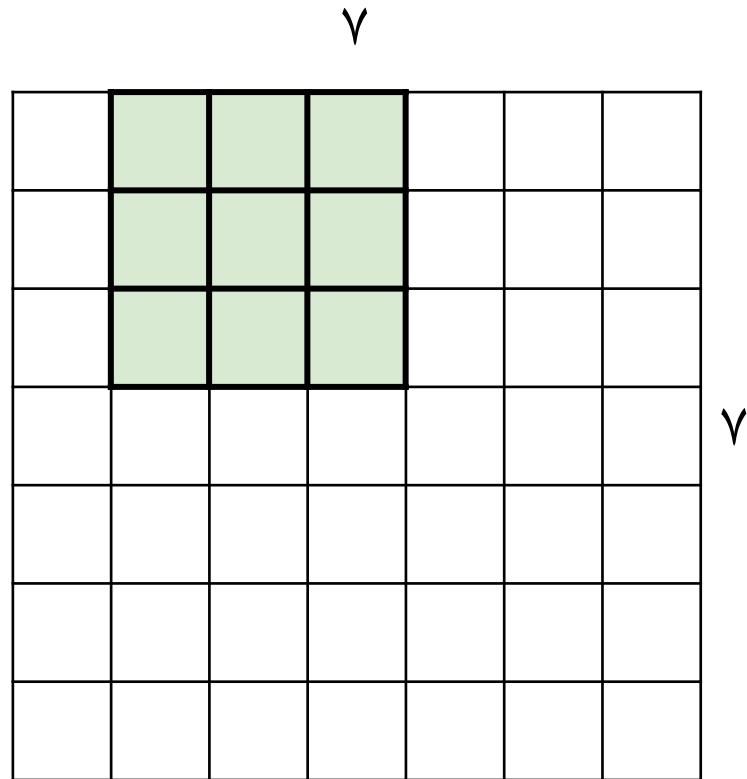


ورودی 7×7

فیلتر 3×3

لایه کانولوشن: یک نگاه دقیق‌تر

۱۹

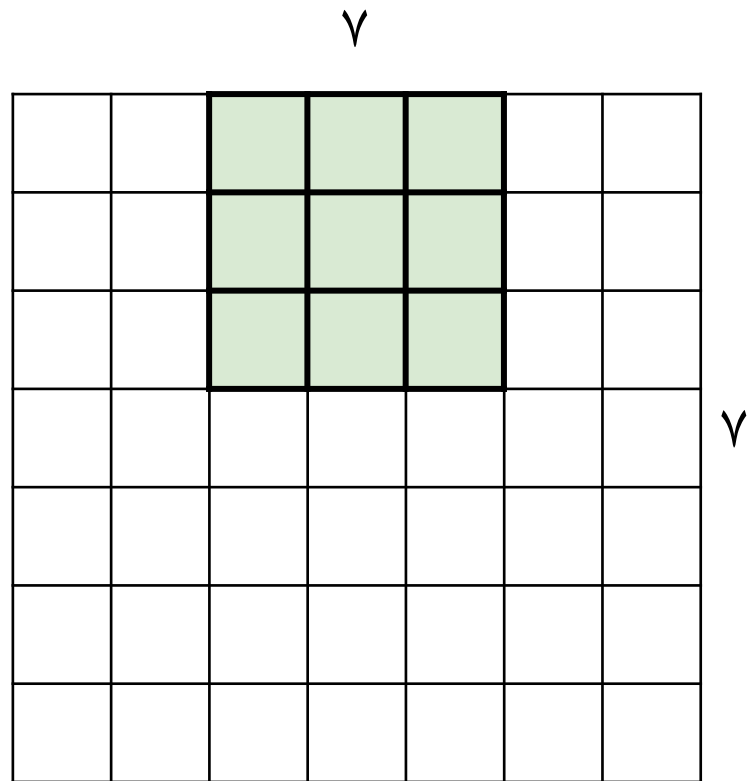


ورودی 7×7

فیلتر 3×3

لایه کانولوشن: یک نگاه دقیق‌تر

۲۰

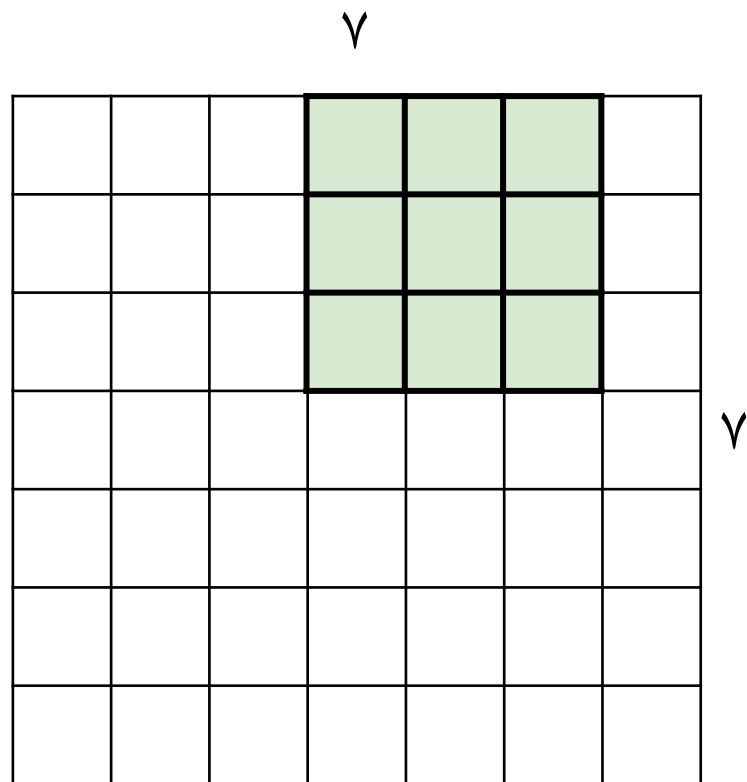


ورودی 7×7

فیلتر 3×3

لایه کانولوشن: یک نگاه دقیق‌تر

۲۱

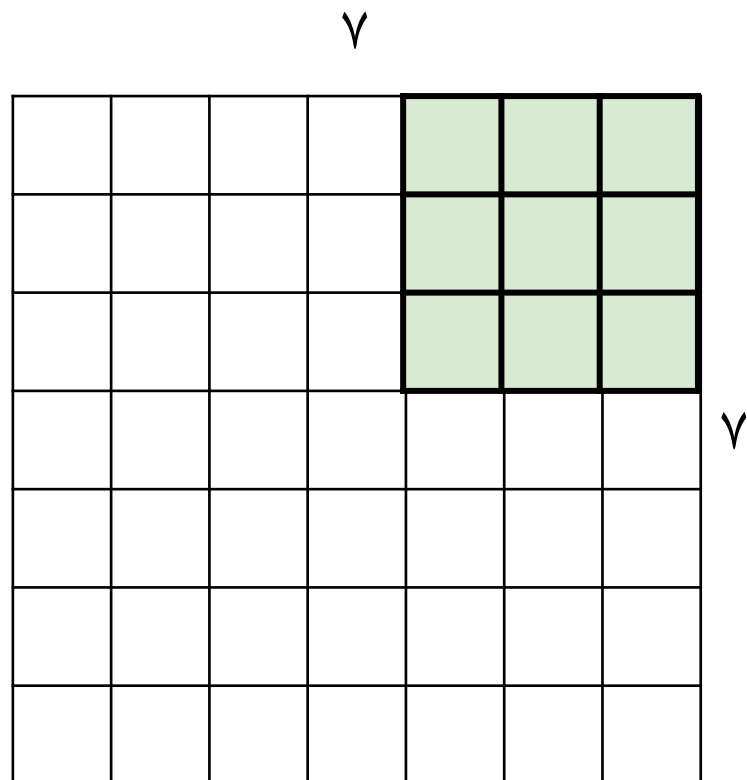


ورودی 7×7

فیلتر 3×3

لایه کانولوشن: یک نگاه دقیق‌تر

۲۲

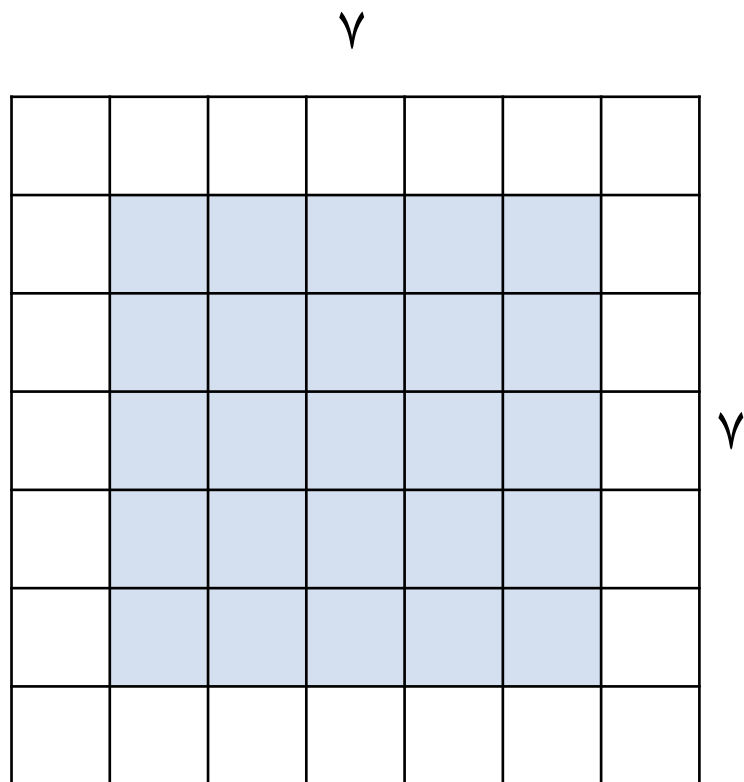


ورودی 7×7

فیلتر 3×3

لایه کانولوشن: یک نگاه دقیق‌تر

۲۳



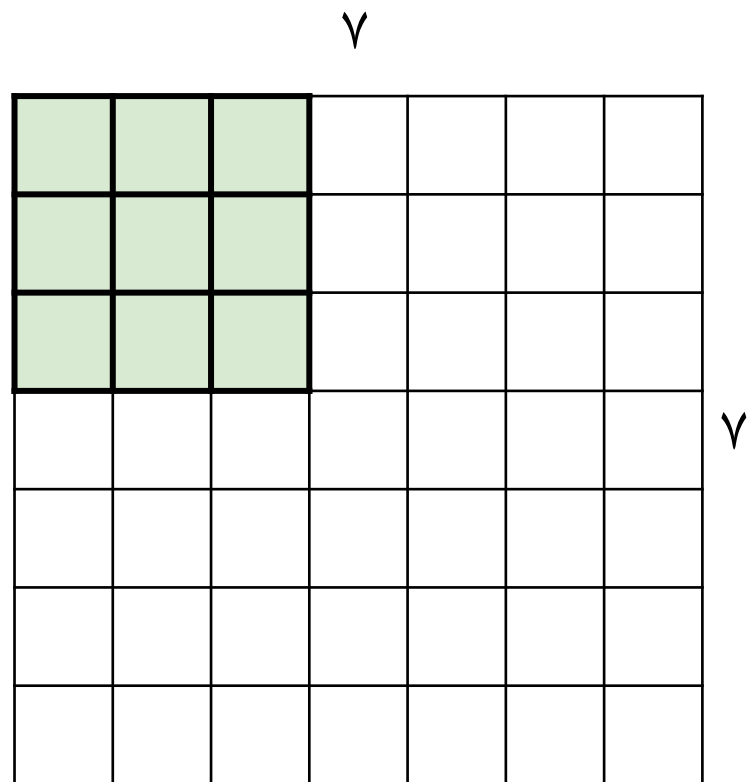
ورودی 7×7

فیلتر 3×3

⇐ خروجی 5×5

لایه کانولوشن: یک نگاه دقیق‌تر

۲۴



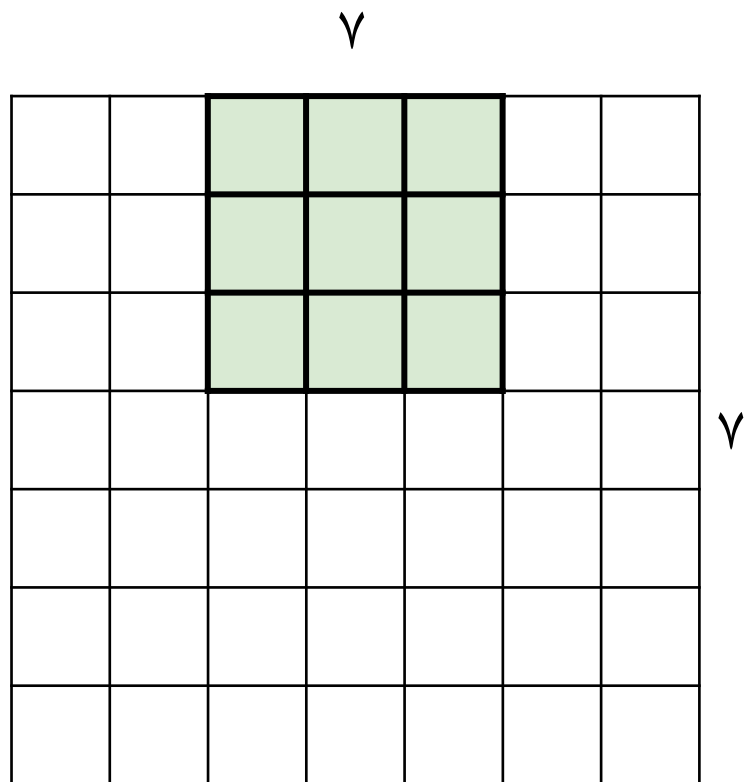
ورودی 7×7

فیلتر 3×3

اندازه گام: ۲

لایه کانولوشن: یک نگاه دقیق‌تر

۲۵



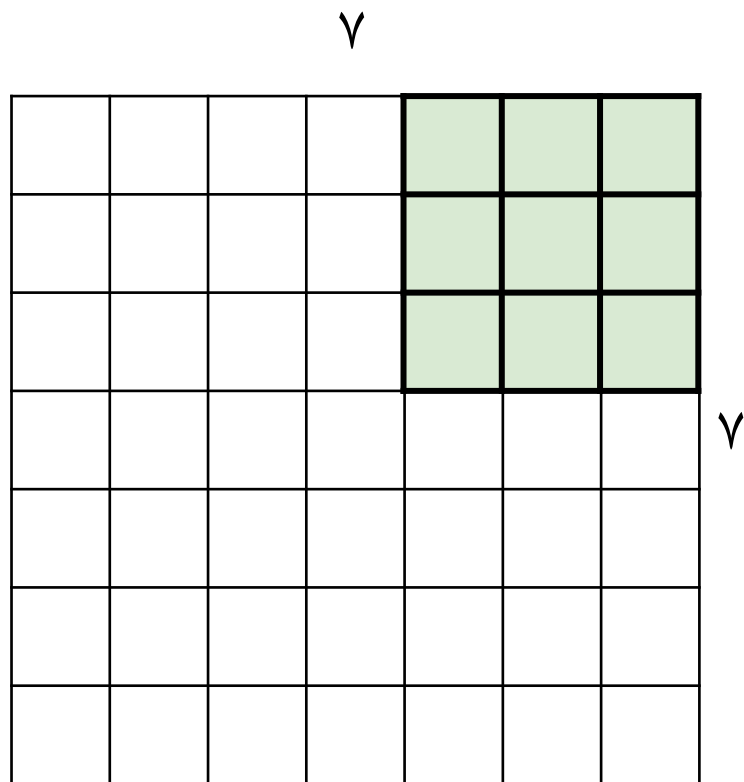
ورودی 7×7

فیلتر 3×3

اندازه گام: ۲

لایه کانولوشن: یک نگاه دقیق‌تر

۲۶



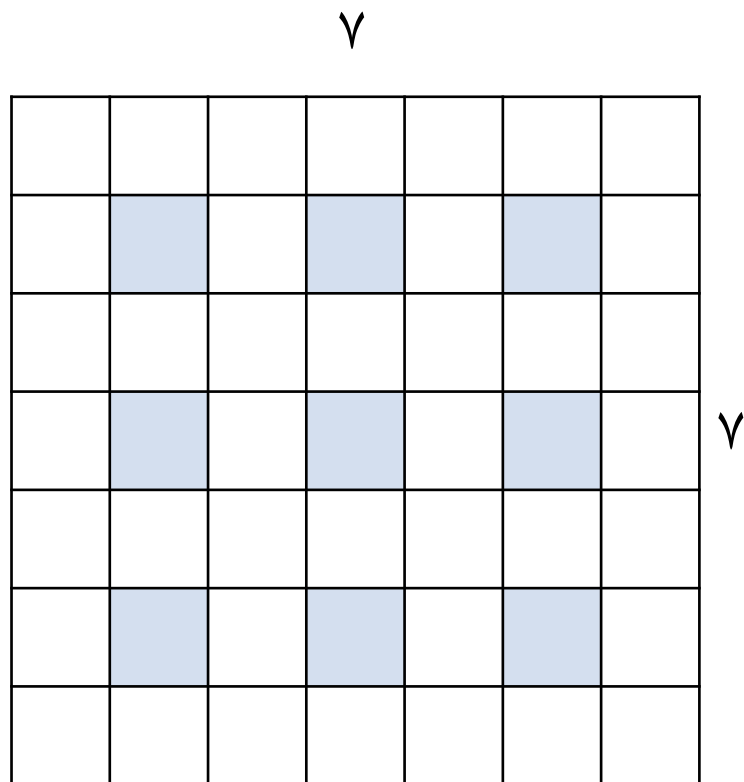
ورودی 7×7

فیلتر 3×3

اندازه گام: ۲

لایه کانولوشن: یک نگاه دقیق‌تر

۲۷



ورودی 7×7

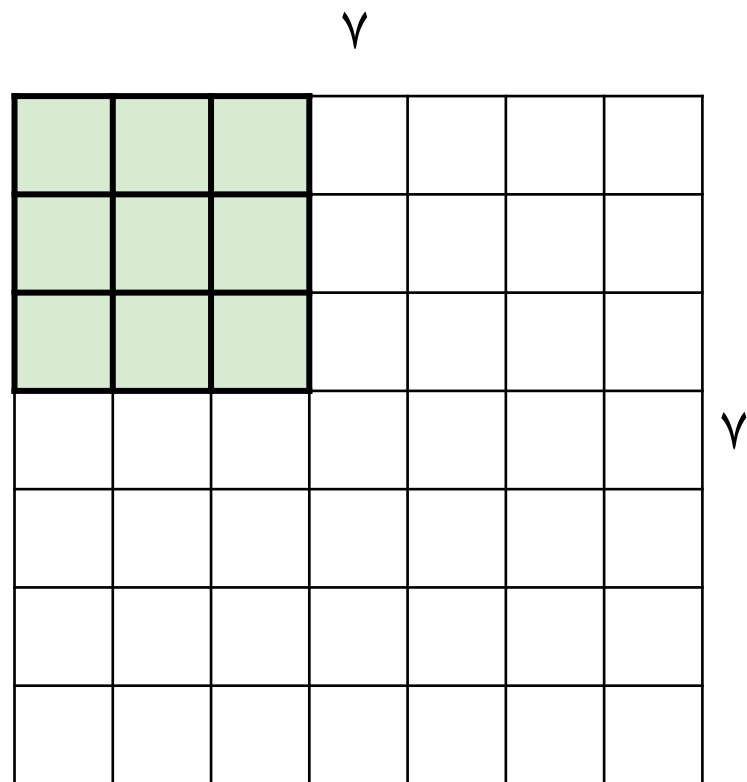
فیلتر 3×3

اندازه گام: ۲

⇐ خروجی 3×3

لایه کانولوشن: یک نگاه دقیق‌تر

۲۸



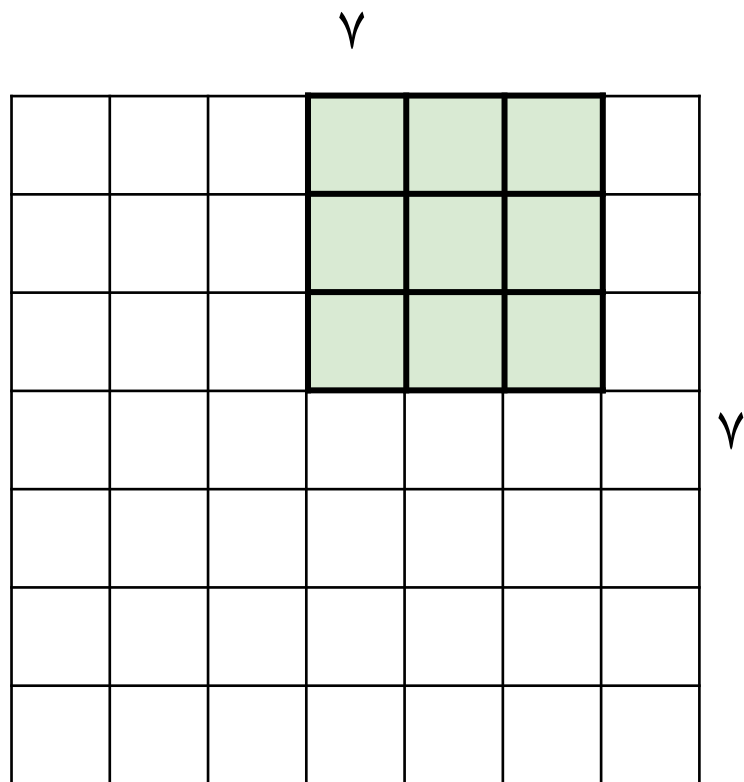
ورودی 7×7

فیلتر 3×3

اندازه گام: ۳

لایه کانولوشن: یک نگاه دقیق‌تر

۲۹



ورودی 7×7

فیلتر 3×3

اندازه گام: ۳

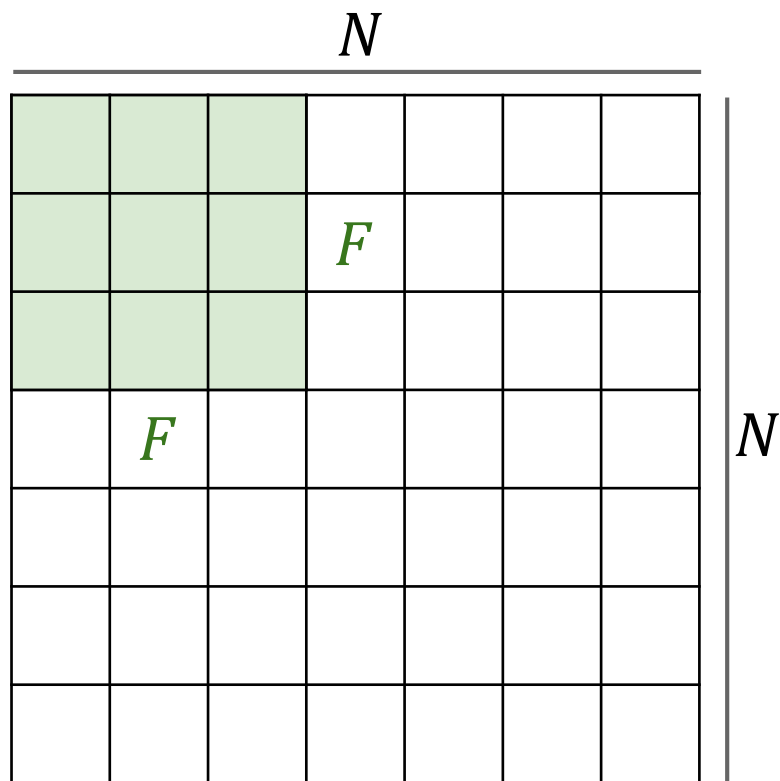
امکان پذیر نیست!

لایه کانولوشن: یک نگاه دقیق‌تر

۳۰

اندازه خروجی:

$$(N - F) / \text{stride} + 1$$



$$N = 7, F = 3$$

$$\text{stride } 1 \Rightarrow (7 - 3) / 1 + 1 = 5$$

$$\text{stride } 2 \Rightarrow (7 - 3) / 2 + 1 = 3$$

$$\text{stride } 3 \Rightarrow (7 - 3) / 3 + 1 = 2.33$$

لایه کانولوشن: گسترش دادن مرزها با صفر

۳۱

0	0	0	0	0	0	0	0	0
0								0
0								0
0								0
0								0
0								0
0								0
0								0
0	0	0	0	0	0	0	0	0

□ مثال. تصویر 7×7

□ فیلتر 3×3 ، با اندازه گام برابر با ۱

□ گسترش مرزها با ۱ پیکسل، اندازه خروجی؟

خروجی 7×7 !

لایه کانولوشن: گسترش دادن مرزها با صفر

۳۲

0	0	0	0	0	0	0	0	0
0								0
0								0
0								0
0								0
0								0
0								0
0	0	0	0	0	0	0	0	0

□ به طور کلی، لایه‌های کانولوشن با اندازه گام برابر ۱، اندازه فیلتر $F \times F$ و صفرگذاری مرزها به اندازه $(F - 1) / 2$ متداول هستند. [اندازه ورودی بدون تغییر]

□ مثال.

□ فیلتر $3 \times 3 \Leftarrow$ صفرگذاری به اندازه ۱

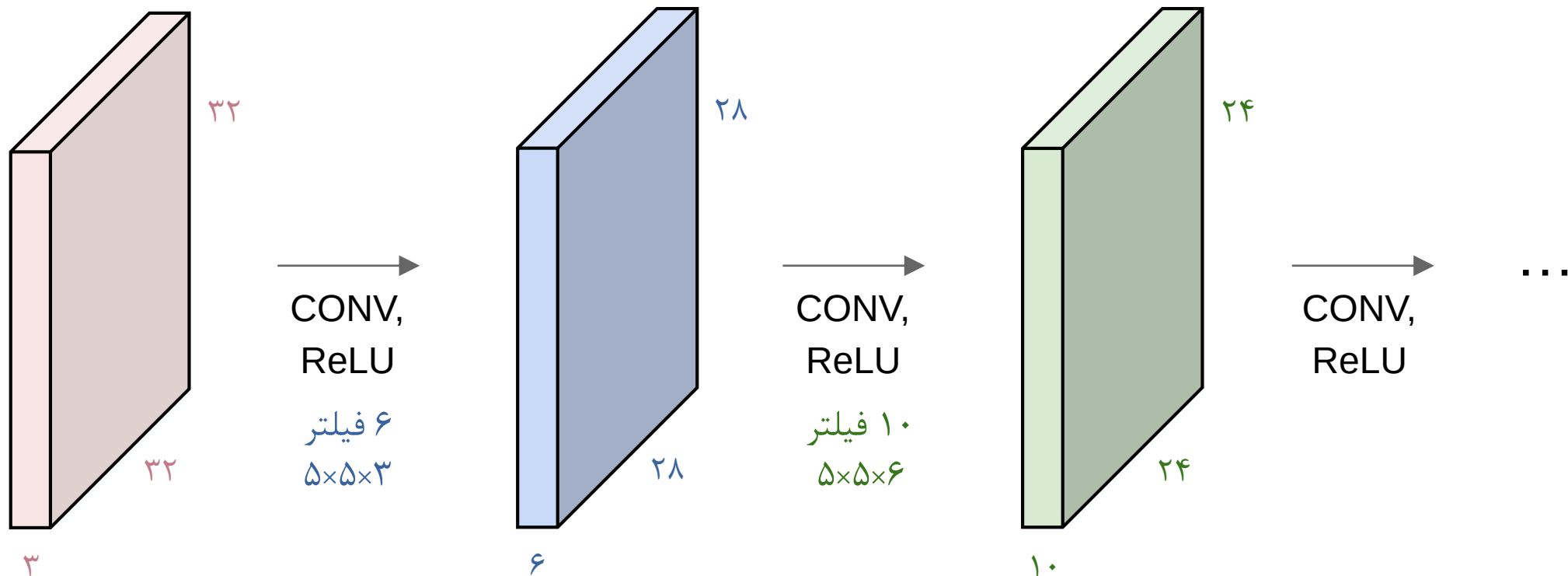
□ فیلتر $5 \times 5 \Leftarrow$ صفرگذاری به اندازه ۲

□ فیلتر $7 \times 7 \Leftarrow$ صفرگذاری به اندازه ۳

یادآوری: شبکه‌های عصبی کانولوشن

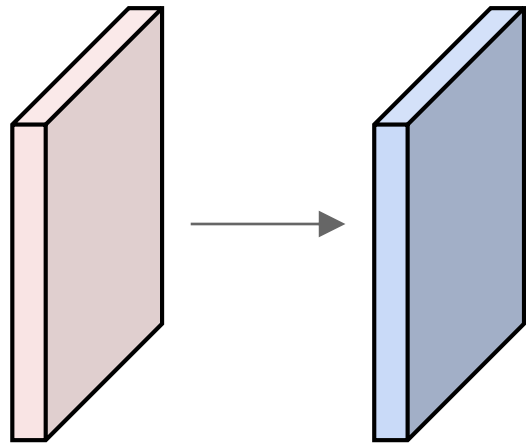
۳۳

- انجام کانولوشن به صورت پی در پی بر روی ورودی، باعث کوچک و کوچک‌تر شدن اندازه ورودی می‌شود.
- کاهش سریع اندازه، ایده خوبی نیست! $[32 \rightarrow 28 \rightarrow 24 \rightarrow \dots]$



لایه کانولوشن: مثال

۳۴



□ اندازه ورودی: $32 \times 32 \times 3$

□ 10 فیلتر 5×5 ، با گام 1 و صفرگذاری 2 پیکسلی

□ اندازه خروجی؟

$$(32 - 5 + 2 \times 2) / 1 + 1 = 32$$

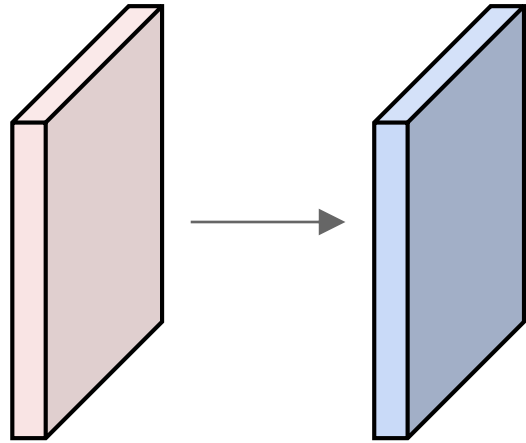
اندازه هر یک از
نقشه‌های فعالیت

$$32 \times 32 \times 10$$

تعداد
نقشه‌های فعالیت

لایه کانولوشن: مثال

۳۵



□ اندازه ورودی: $32 \times 32 \times 3$

□ 10 فیلتر 5×5 ، با گام 1 و صفرگذاری 2 پیکسلی

□ تعداد کل پارامترها در این لایه؟

برای بایاس

$$5 \times 5 \times 3 + 1 = 76$$

تعداد پارامترها
برای هر فیلتر

$$76 \times 10 = 760$$

تعداد کل پارامترها

لایه کانولوشن: جمع‌بندی

۳۶

مقایسه متداول:

$K =$ توانی از ۲ (مانند ۳۲، ۶۴، ۱۲۸ و ۵۱۲)

- $P = 1, S = 1, F = 3$

- $P = 2, S = 1, F = 5$

- $P = ?, S = 2, F = 5$ (هر مقدار مناسب)

- $P = 0, S = 1, F = 1$

□ دریافت یک ورودی با ابعاد $W_1 \times H_1 \times D_1$

□ نیاز به چهار ابرپارامتر:

□ تعداد فیلترها K

□ ابعاد هر فیلتر $F \times F$

□ اندازه گام S

□ میزان صفرگذاری P

□ تولید یک خروجی با ابعاد $W_2 \times H_2 \times D_2$ ، به گونه‌ای که:

$$W_2 = (W_1 - F + 2P) / S + 1 \quad \square$$

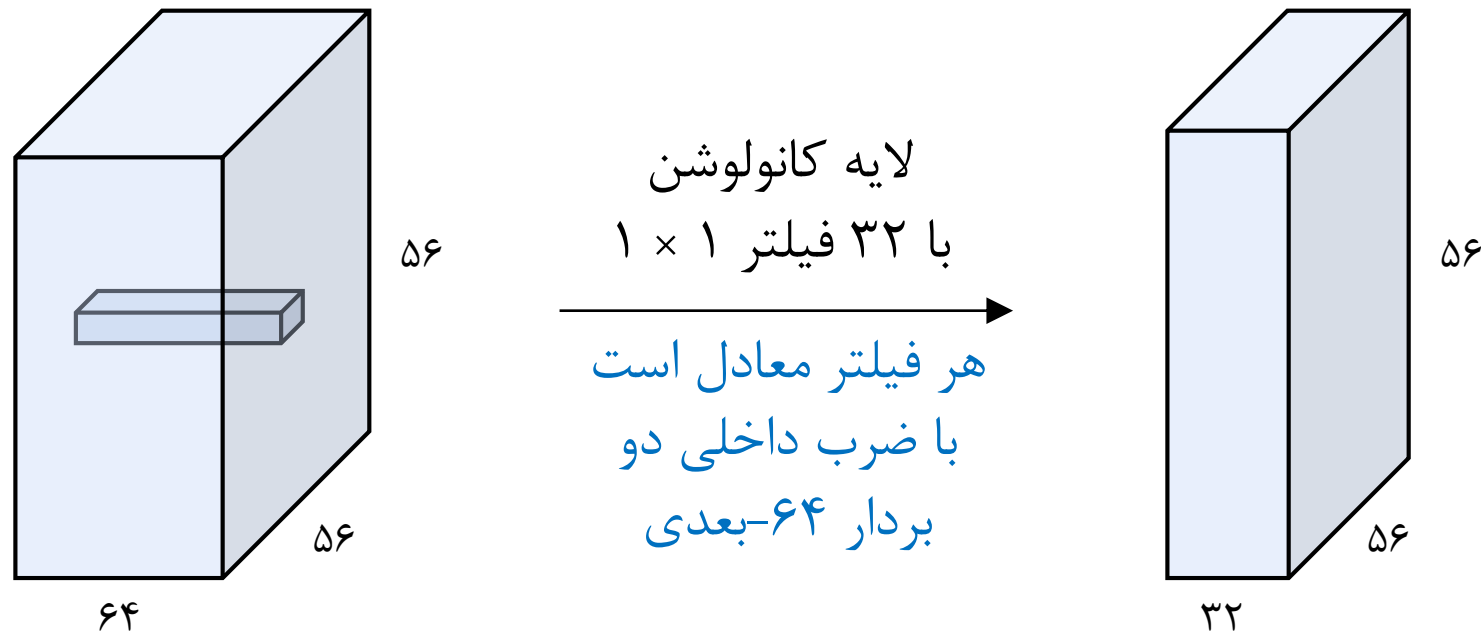
$$H_2 = (H_1 - F + 2P) / S + 1 \quad \square$$

$$D_2 = K \quad \square$$

□ دارای $F \times F \times D_1$ وزن به ازای هر فیلتر و تعداد کل $(F \times F \times D_1) \times K$ وزن و K بایاس.

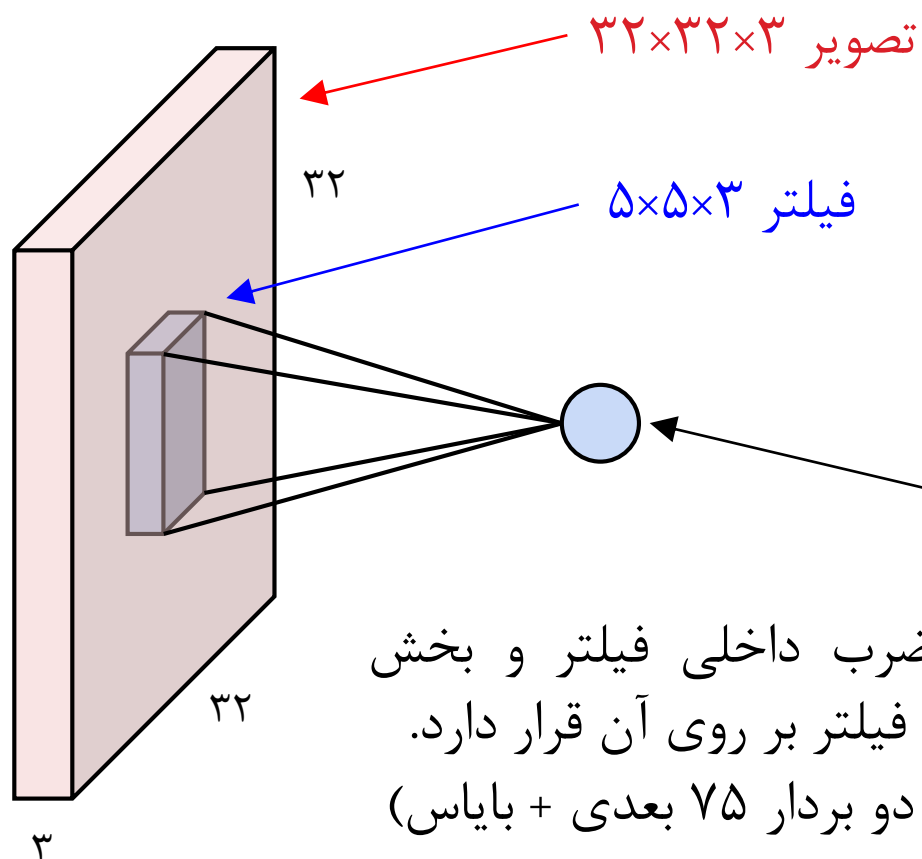
لایه کانولوشن: با فیلترهای 1×1

۳۷

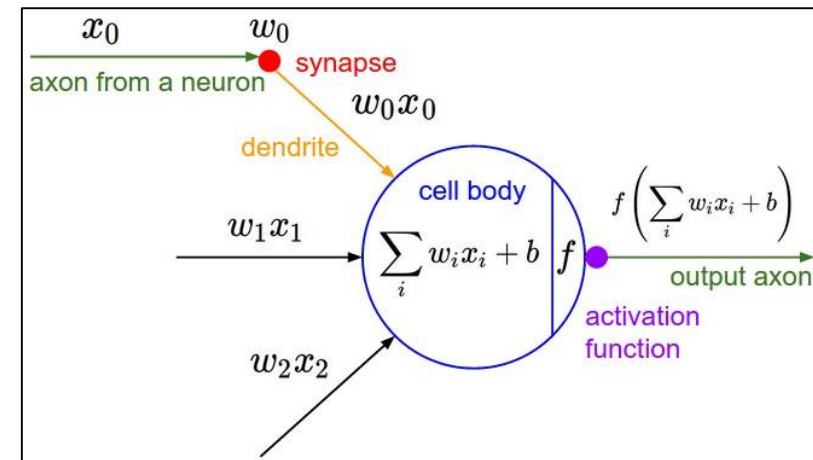


دیدگاه نورونی از لایه کانولوشن

۳۸



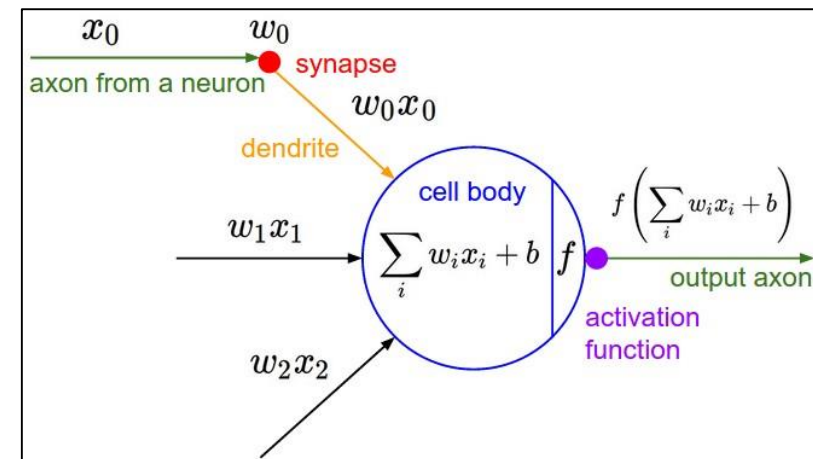
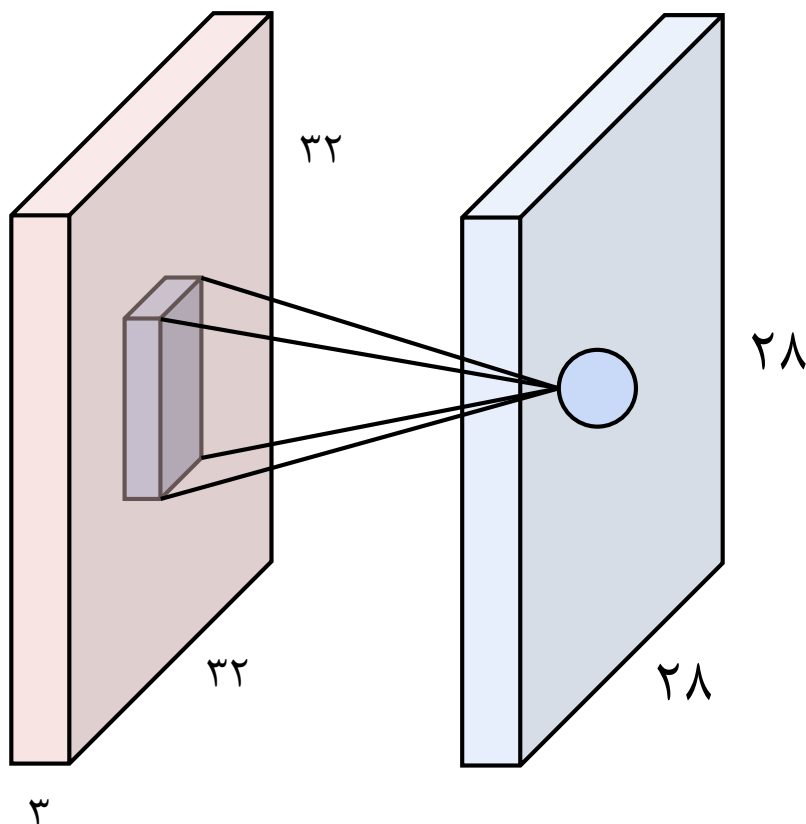
به دست آمده از ضرب داخلی فیلتر و بخش کوچکی از تصویر که فیلتر بر روی آن قرار دارد. (یعنی، ضرب داخلی دو بردار ۷۵ بعدی + بایاس)



یک نورون با اتصالات محلی...

دیدگاه نورونی از لایه کانولوشن

۳۹

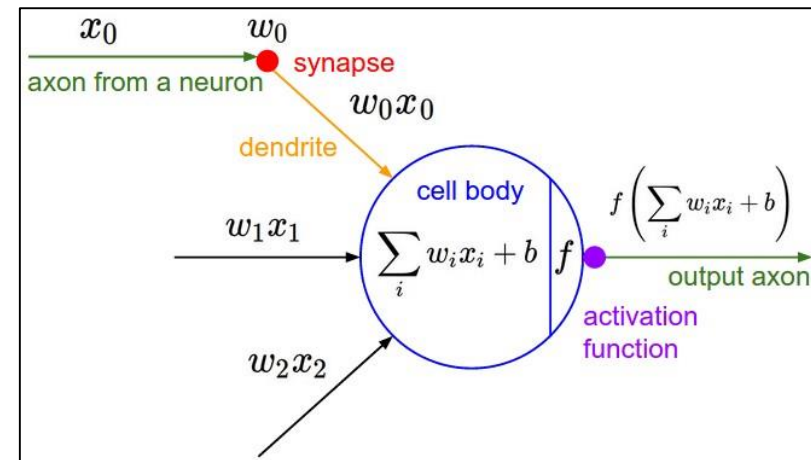
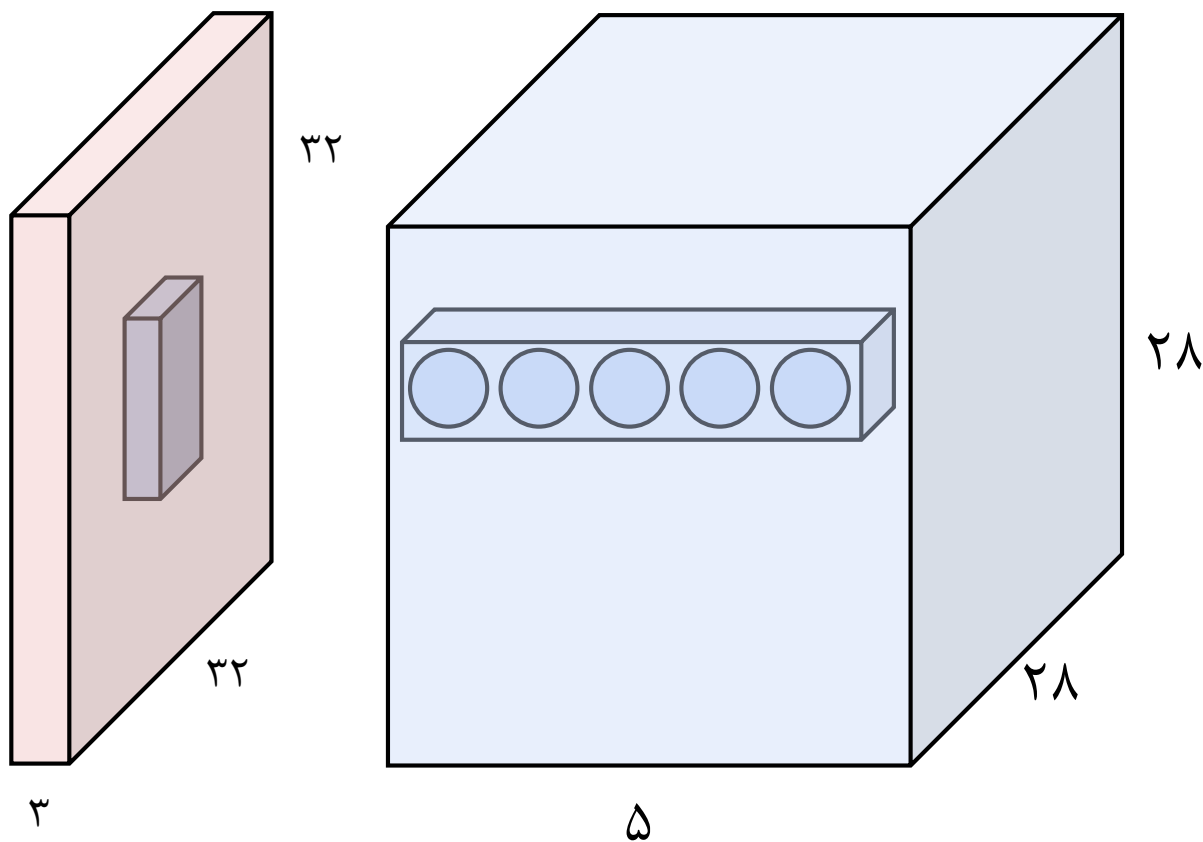


- هر نقشه فعالیت یک لایه 26×26 از خروجی نورون‌ها است:
۱. هر نورون به یک ناحیه کوچک از ورودی متصل است.
 ۲. همه نورون‌ها دارای پارامترهای مشترک هستند.

«فیلتر 5×5 »: یک میدان دید 5×5 برای هر نورون!

دیدگاه نورونی از لایه کانولوشن

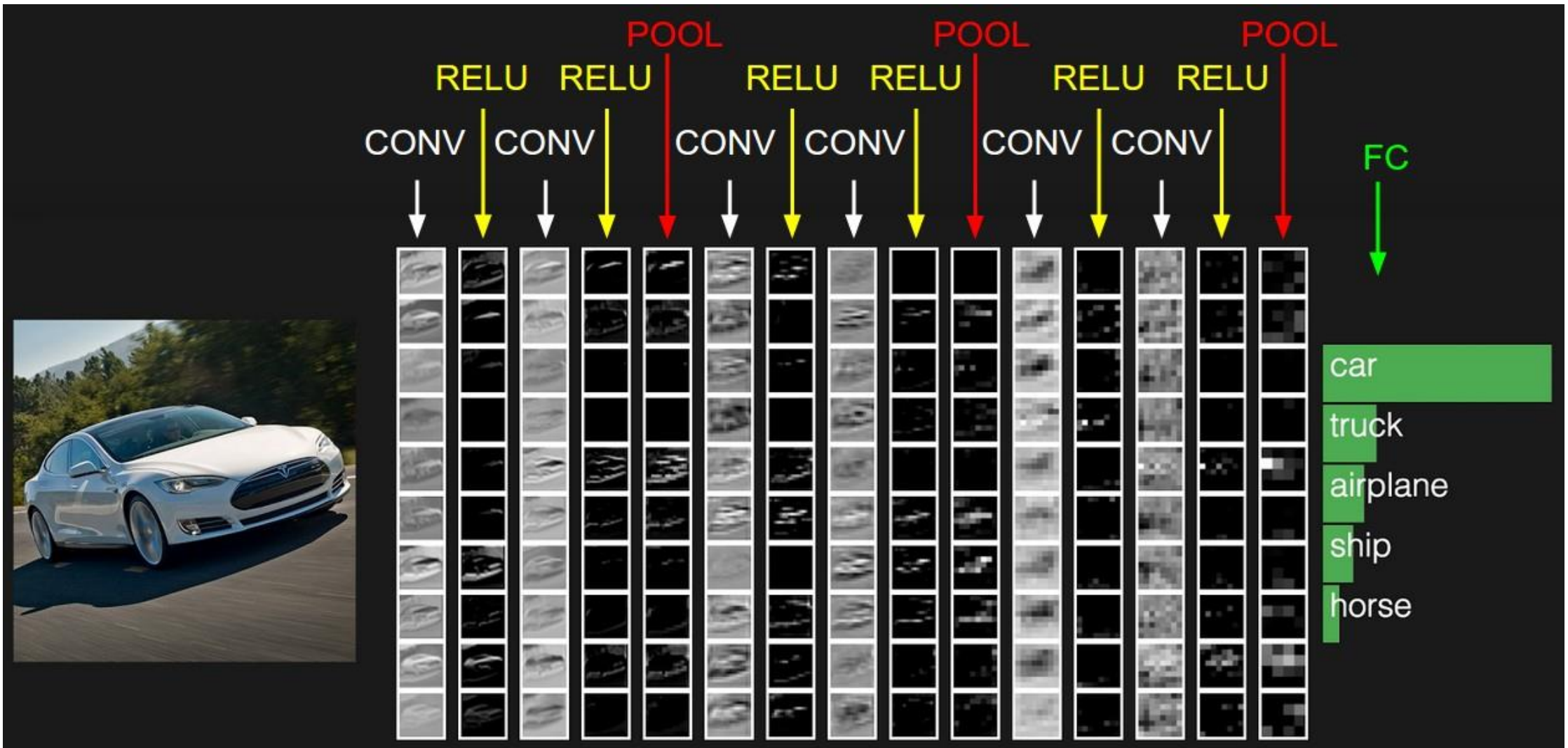
۴۰



به عنوان مثال با ۵ فیلتر، لایه کانولوشن از تعدادی نورون تشکیل شده است که به صورت یک گرید ۳-بعدی قرار گرفته‌اند.
($28 \times 28 \times 5$)

شبکه‌های عصبی کانولوشن: لایه‌های ادغام

۴۱

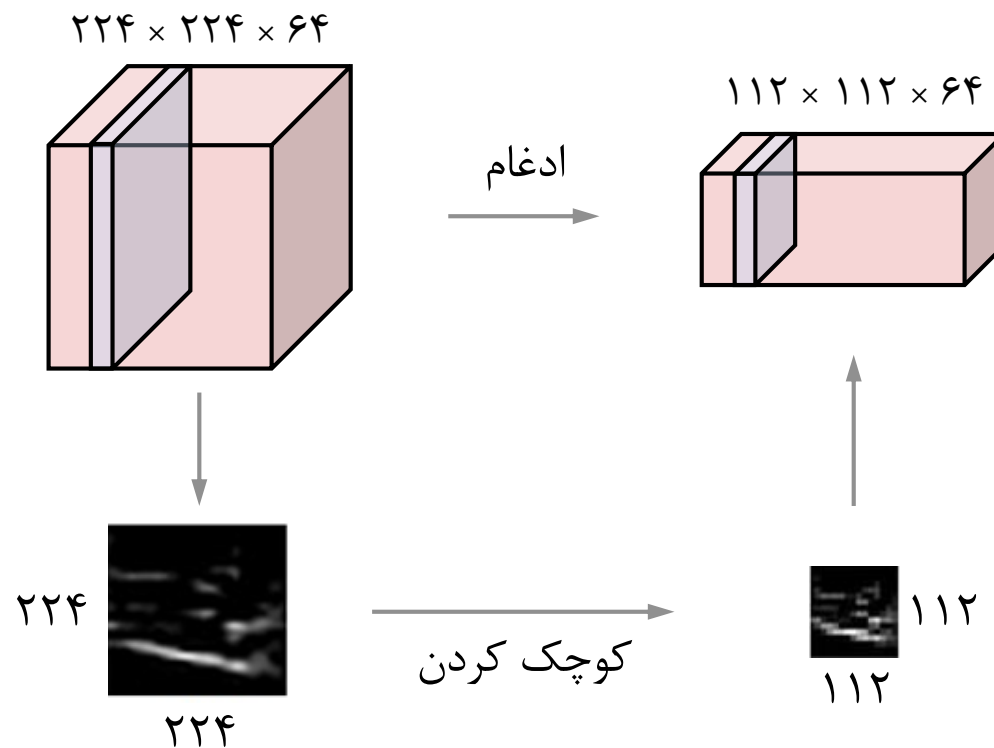


شبکه‌های عمیق کانولوشن: لایه‌های ادغام

۴۲

□ لایه ادغام. کوچک‌تر کردن بازنمایی!

□ بر روی هر نقشه فعالیت به صورت جداگانه عمل می‌کند.

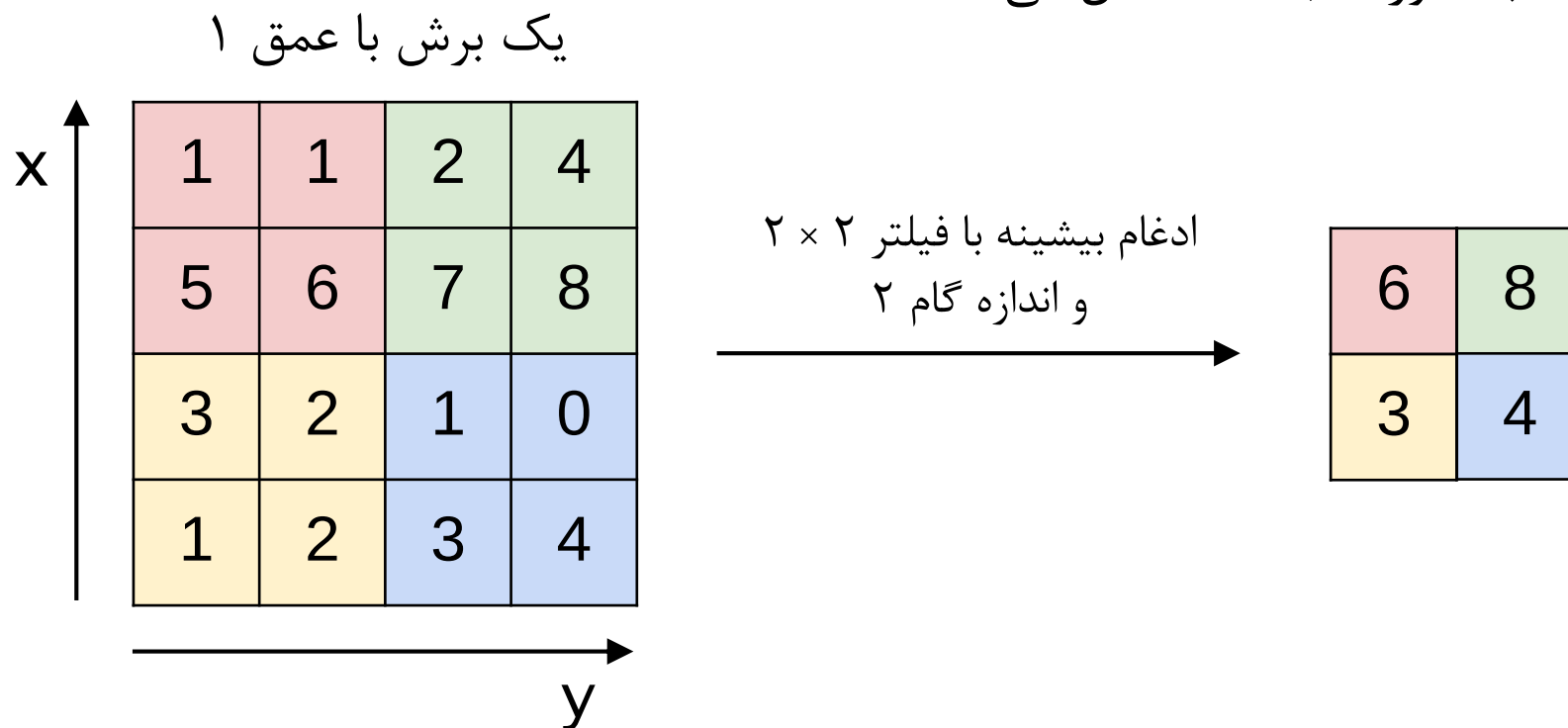


شبکه‌های عصبی کانولوشن: ادغام بیشینه

۴۳

□ لایه ادغام. کوچک‌تر کردن بازنمایی!

□ بر روی هر نقشه فعالیت به صورت جداگانه عمل می‌کند.



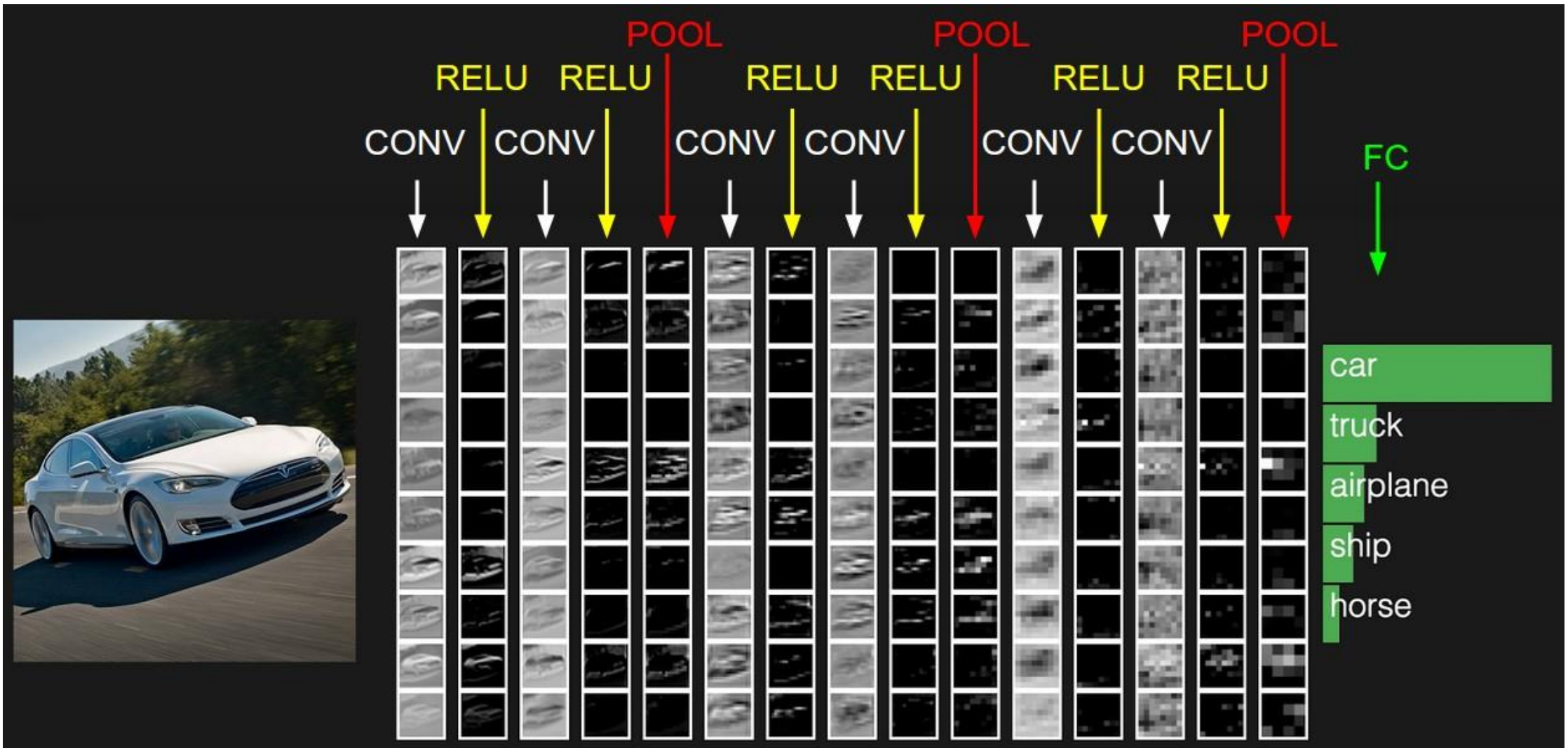
لایه ادغام: جمع‌بندی

۴۴

- دریافت یک ورودی با ابعاد $W_1 \times H_1 \times D_1$
- نیاز به سه ابرپارامتر:
 - ابعاد هر فیلتر $F \times F$ [مانند ۲ یا ۳]
 - اندازه گام S
- تولید یک خروجی با ابعاد $W_2 \times H_2 \times D_2$ ، به گونه‌ای که:
 - $W_2 = (W_1 - F)/S + 1$
 - $H_2 = (H_1 - F)/S + 1$
 - $D_2 = D_1$
- تعداد پارامترها برابر است با **صفر**!
- **توجه.** استفاده از صفرگذاری برای لایه‌های ادغام متداول نیست.

شبکه‌های عصبی کانولوشن: لایه‌های تمام متصل

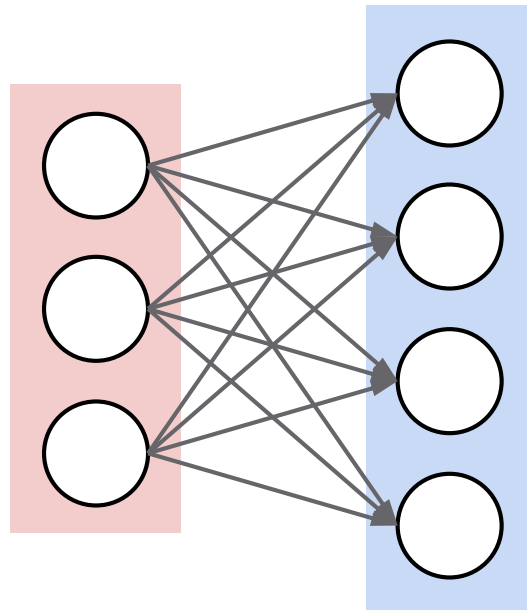
۴۵



شبکه‌های عصبی کانولوشن: لایه‌های تمام متصل

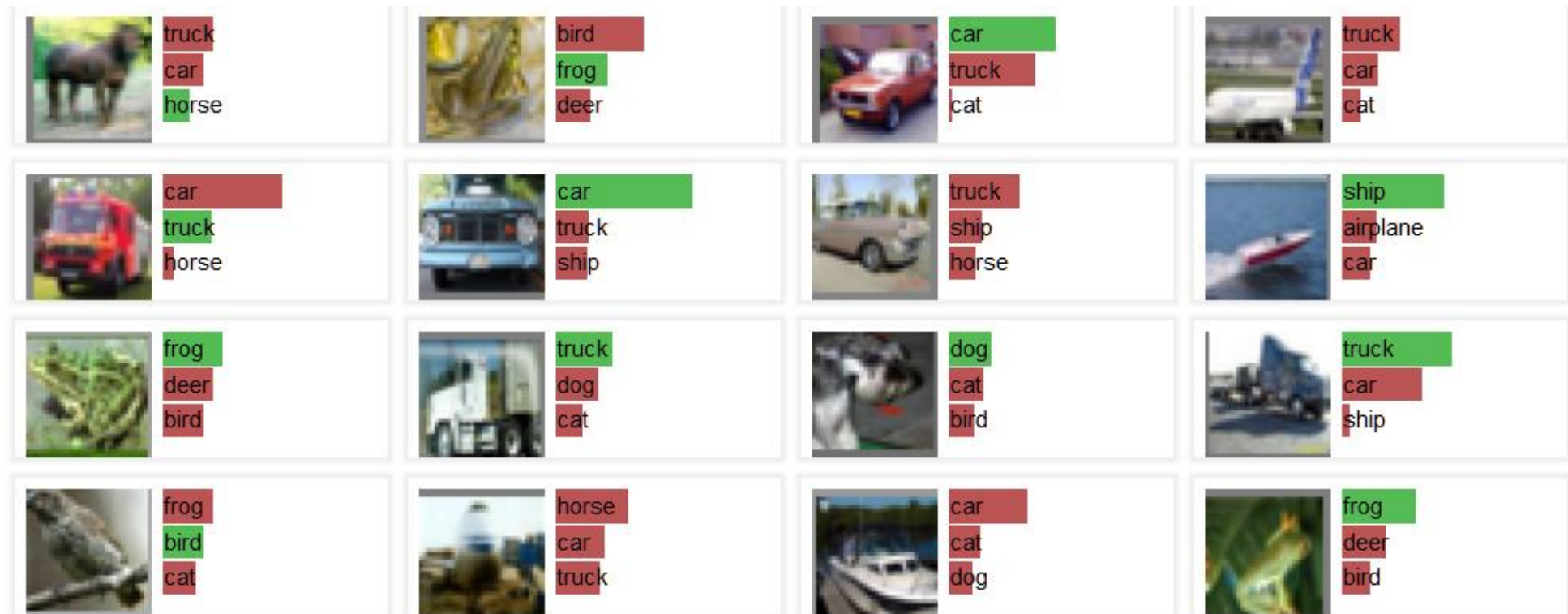
۴۶

- لایه‌های تمام متصل. شامل نورون‌هایی که به تمام ورودی‌ها وصل هستند!
- همانند لایه‌های به کار رفته در یک شبکه عصبی معمولی.



شبکه‌های عمیق کانولوشن: اجرای یک نمونه نمایشی

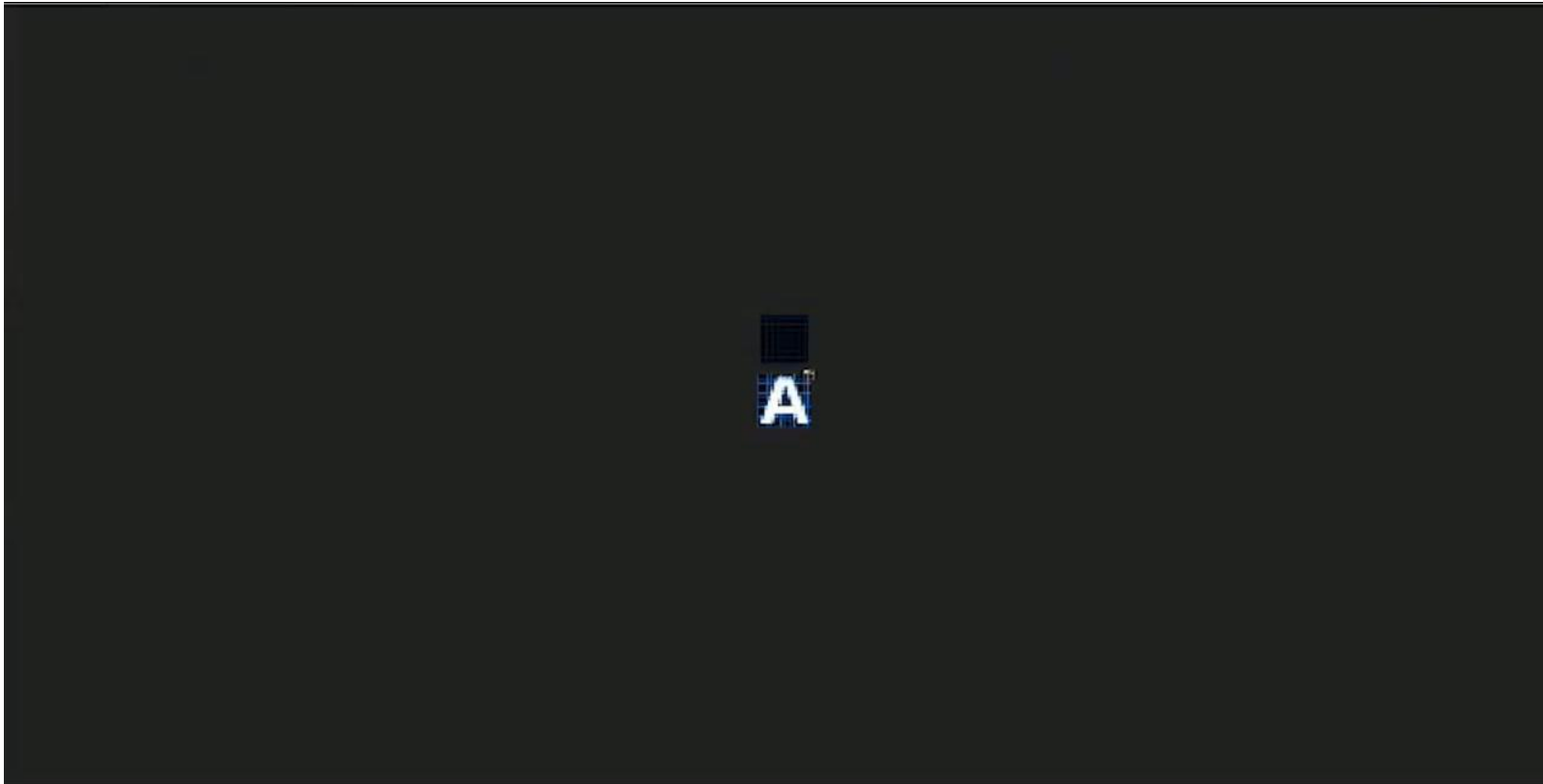
۴۷



اجرای یک نمونه نمایشی بر روی مجموعه داده CIFAR-10

شبکه‌های عصبی کانولوشن: اجرای یک نمونه نمایشی

۴۸



[Otavio Good] A visual and intuitive understanding of deep learning

شبکه‌های عصبی کانولوشن: جمع‌بندی

۴۹

- شبکه‌های عصبی کانولوشن. دنباله‌ای از لایه‌های کانولوشن، ادغام و تمام متصل.
- گرایش به سمت فیلترهای کوچک‌تر و ساختارهای عمیق‌تر.
- گرایش به سمت حذف کامل لایه‌های ادغام و لایه‌های تمام متصل.
- یک الگوی رایج برای ساختار شبکه‌های کانولوشن:

$[(\text{CONV-RELU}) * N - \text{POOL?}] * M - (\text{FC-RELU}) * K, \text{SOFTMAX}$

- به طوری که N حدکثر برابر با ۵، M یک مقدار بزرگ و K بین ۰ و ۲ است.
- اما پیشرفت‌های اخیر مانند ResNet و GoogLeNet این الگو را به چالش کشیده‌اند!