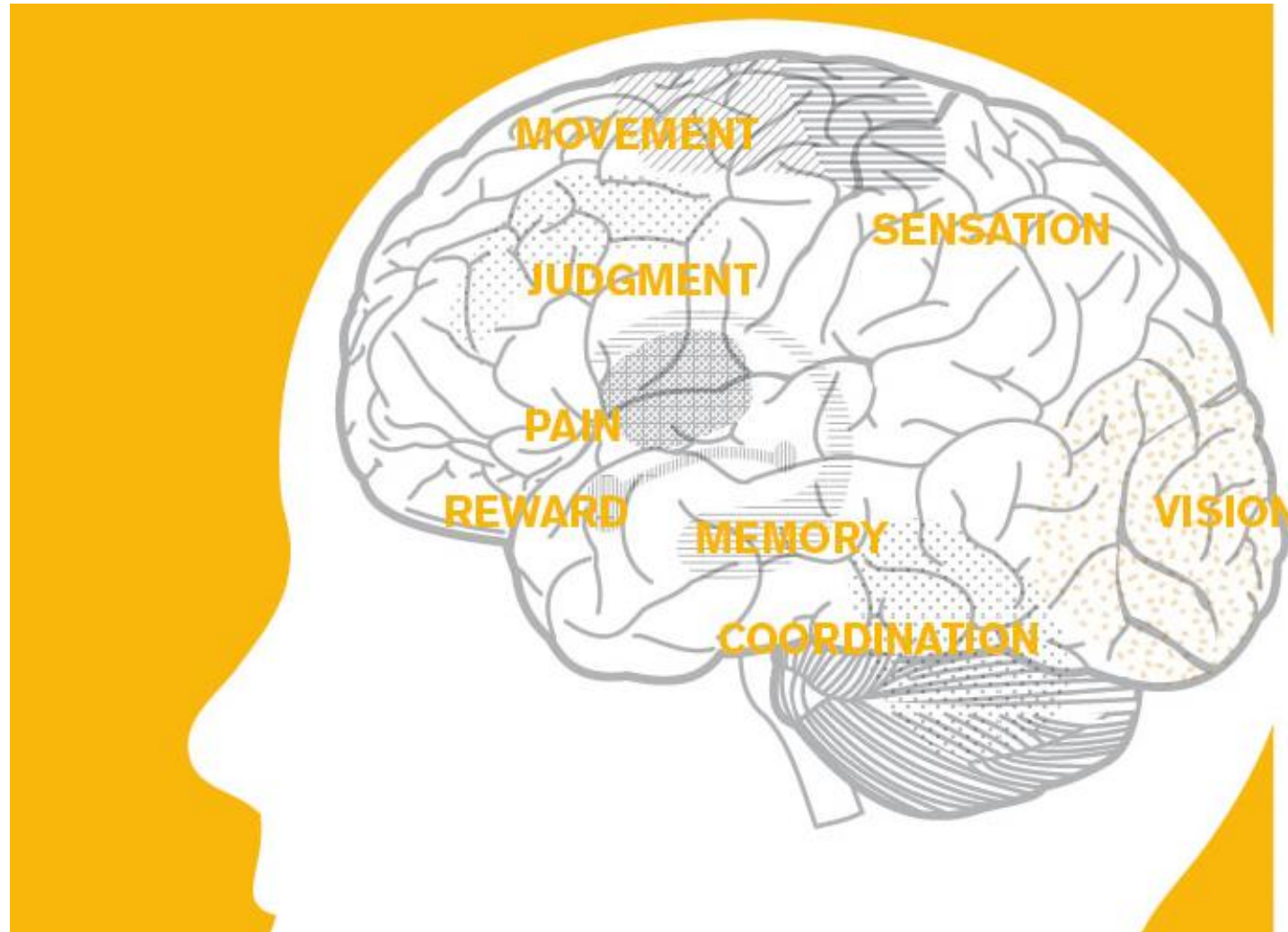


Convolutional Neural Networks (CNN)

گروه دایچه . dayche.com




Humans' brain



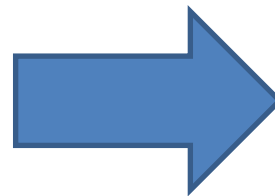
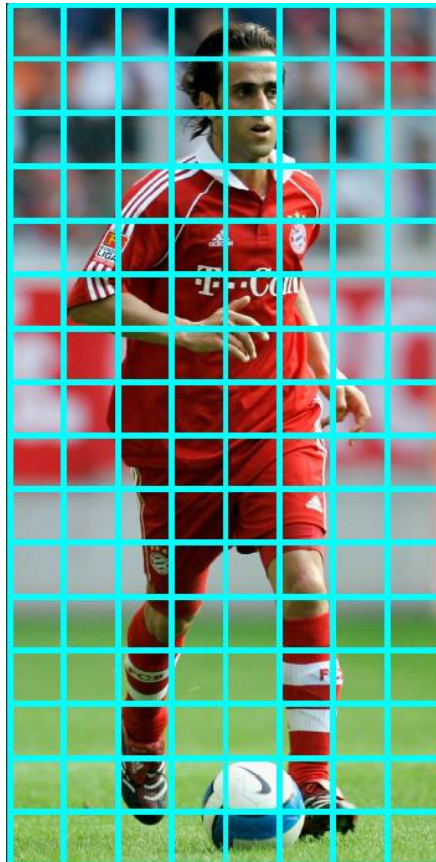
تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

منابع داده‌ای متفاوتی به انسان کمک می‌کنند تا تصمیم‌گیری کند. طبق مطالعات انجام شده، 90 درصد از اطلاعاتی که یک انسان از محیط اطراف خود دریافت می‌کند از طریق چشم است.



شبکه‌های عصبی کانولوشنی، دسته‌ای از شبکه‌های عصبی هستند که با داده‌های مشبک شده کار می‌کنند.

- تصویر
- سری زمانی

هر پیکسل معرف یک متغیر است

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 


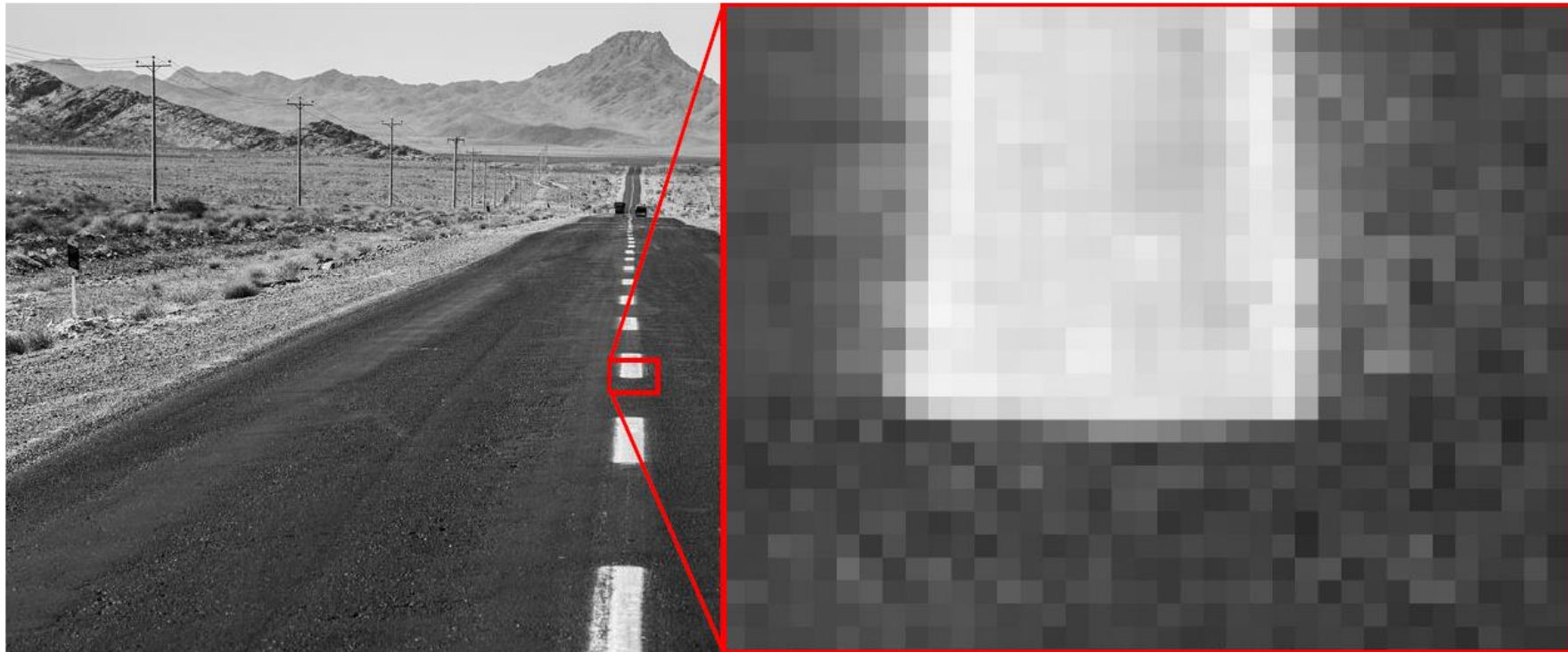
dayche.com | گروه دایچه 

Image data – how a human see an image



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 


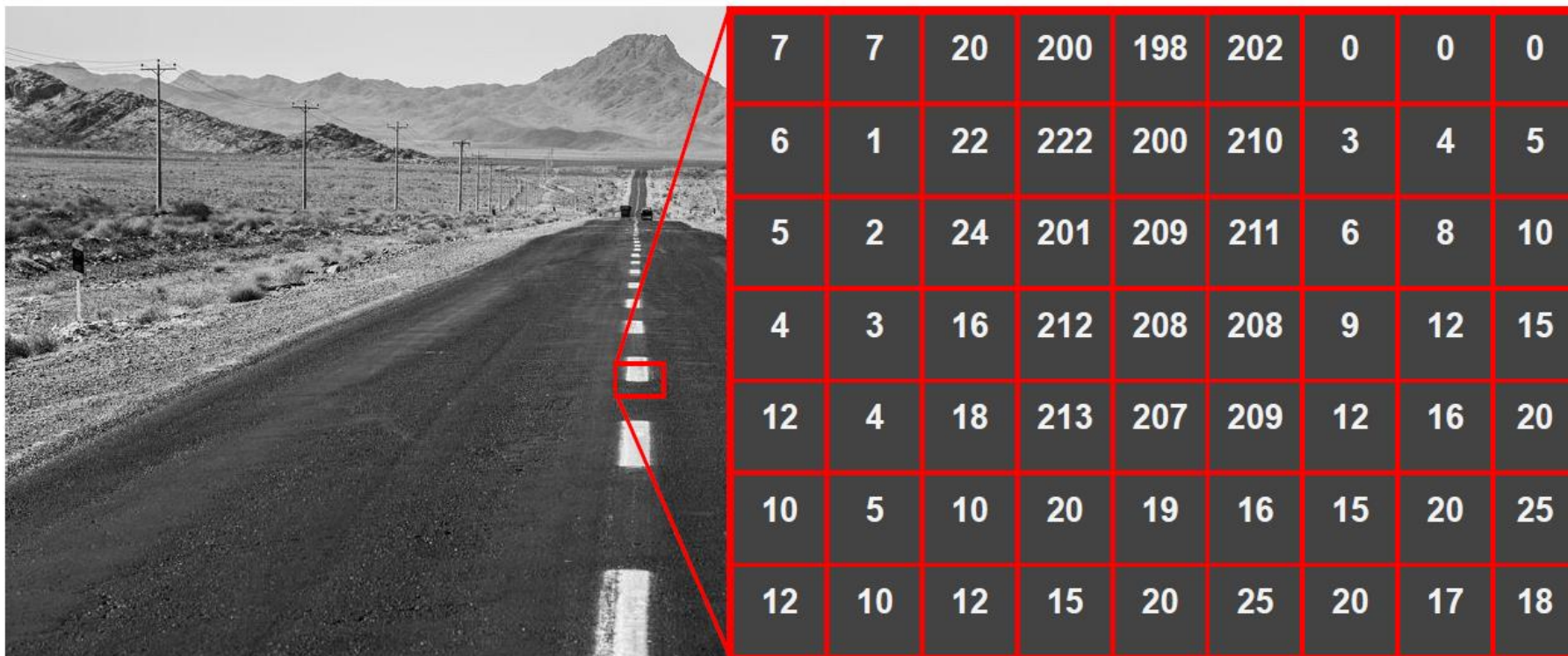
گروه دایکه | dayche.com 

Image data – how computer see an image



Gray image – A 2D array (one dimensional tensor)



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایکه | dayche.com

Image data – how computer see an image



Colored image – Three 2D array (Three dimensional tensor)



تولید محتوا: وحید محمدزاده ایوقی

daychegroup


daychegroup

گروه دایکه | dayche.com


RGB color model



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

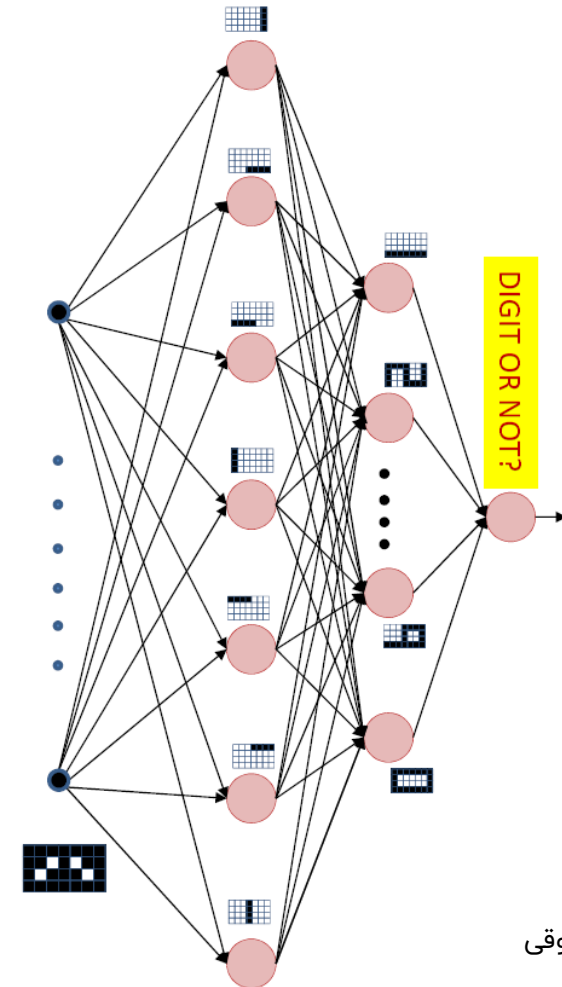
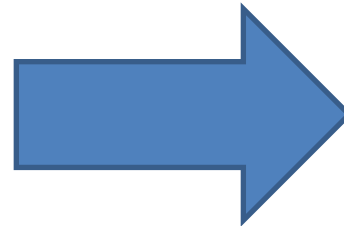
daychegroup 

گروه دایکه | dayche.com 

Multilayer perceptron network




اگر ورودی یک آرایه 28 در 28 باشد، و تعداد نرون‌های لایه هیدن برابر با 100 باشد، تعداد پارامترهای چقدر خواهد بود؟



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

Challenges




0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

تولید محتوا: وحید محمدزاده ایوقی

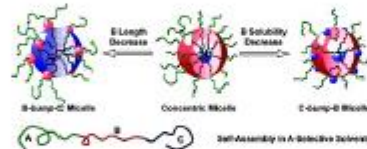
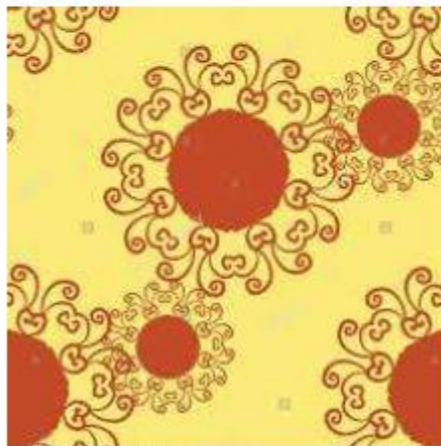
daychegroup 

daychegroup 

گروه دایکه | dayche.com 

Challenges


آیا تصویر داده شده، مستقل از نوع و محل، متعلق به یک گل است؟



تولید محتوا: وحید محمدزاده ایوقی

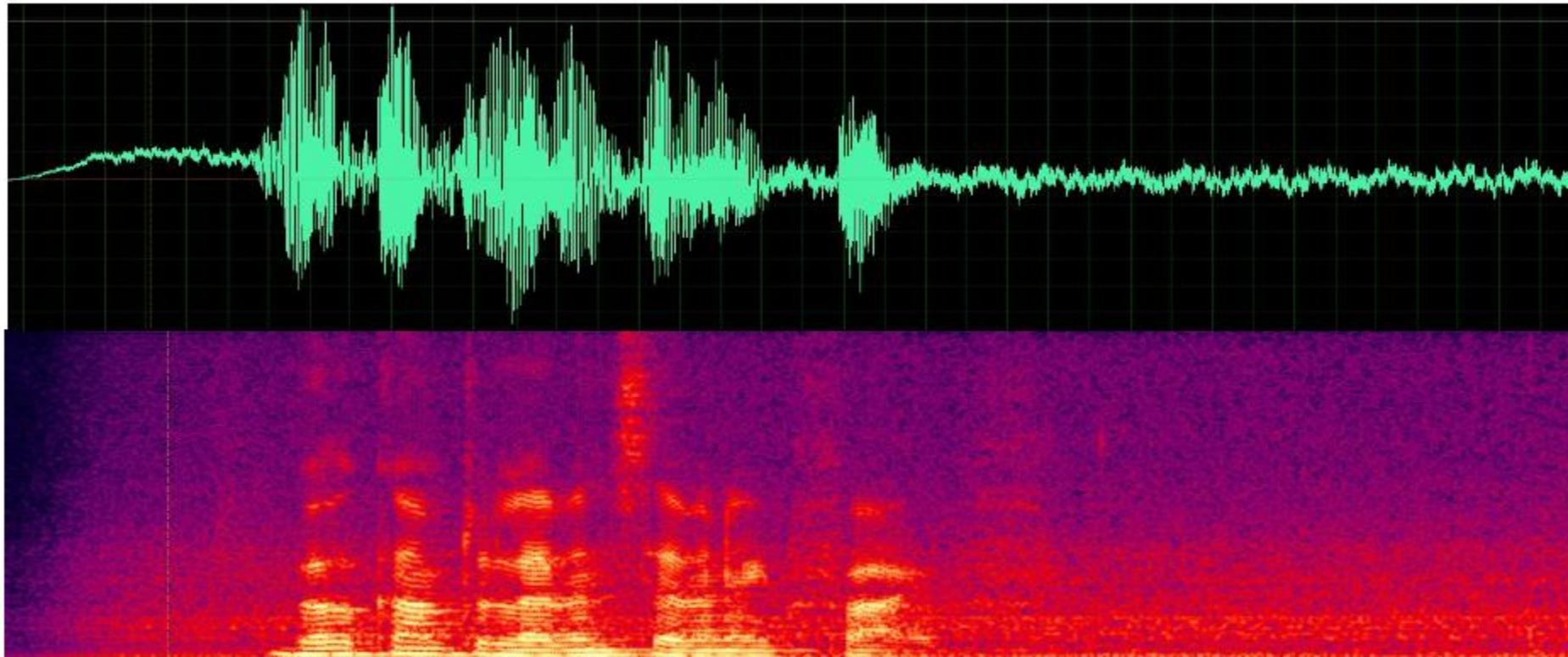
daychegroup 

daychegroup 

گروه دایچه | dayche.com 



آیا کلمه مشخصی، به عنوان مثال یک اسم خاص، در این سیگنال گفتار وجود دارد؟



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

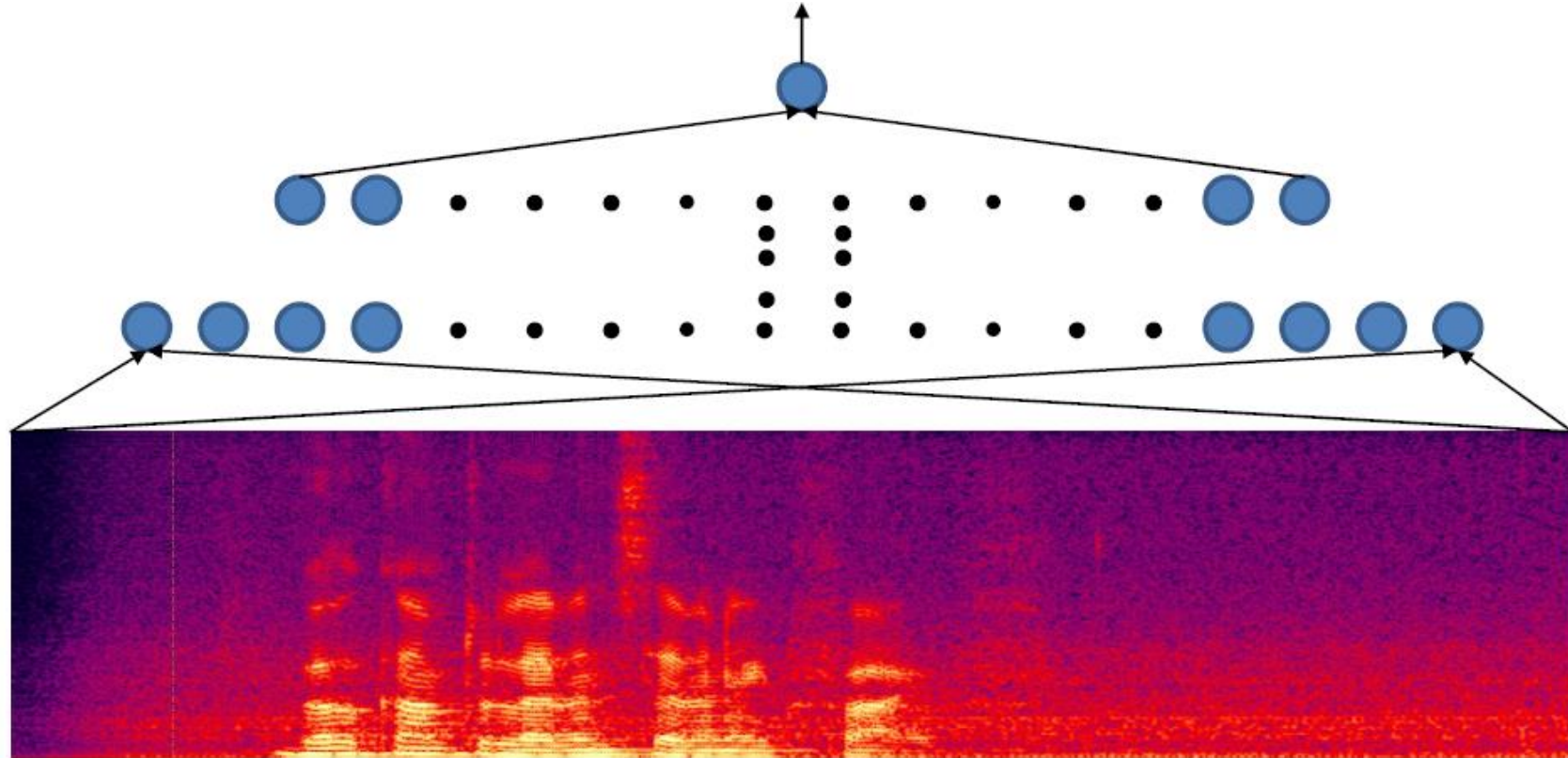
daychegroup 

dayche.com | گروه دایکه 

Can MLP be a solution?



اگر برای تمام سیگنال‌های گفتار، کلمه خواسته شده در محل مشخصی رخ بدهد شبکه عصبی MLP پاسخ مناسبی است.



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

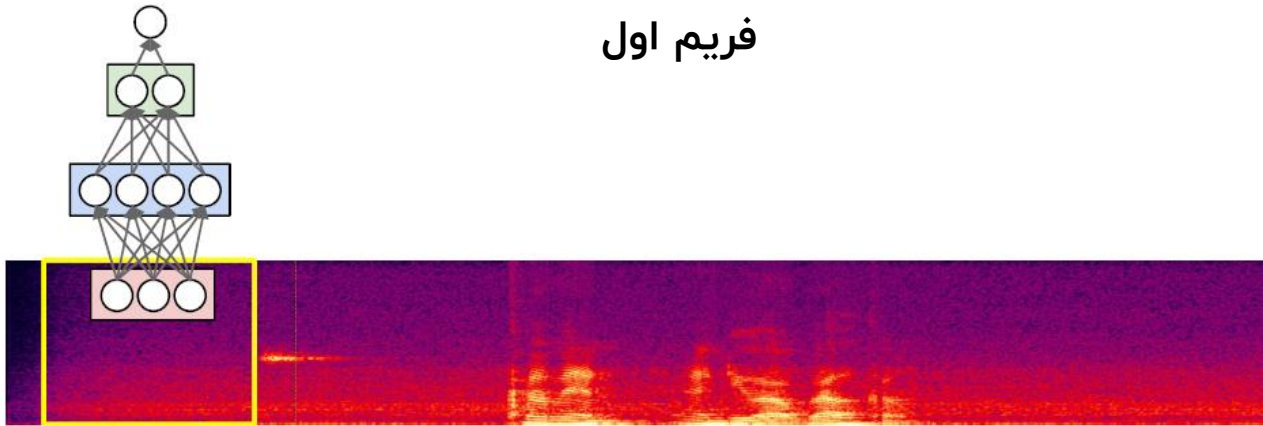
daychegroup

dayche.com | گروه دایکه

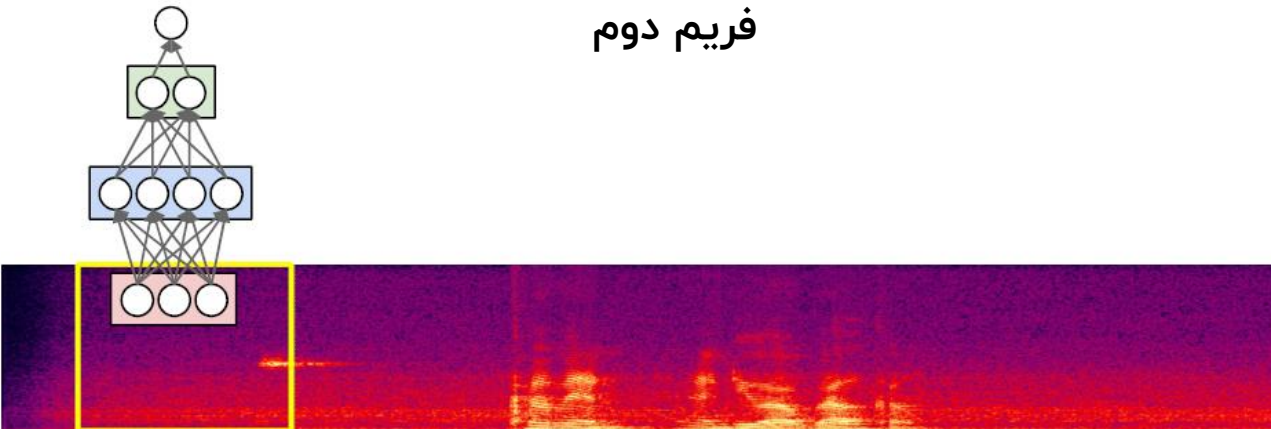
How a human search for an object?



فریم اول



فریم دوم



- فرآیند اسکن کردن توسط مغز در کسری از ثانیه اتفاق می‌افتد.
- در هر فریم از سیگنال ورودی یک شبکه عصبی وجود دارد که این شبکه ورودیهای متناظر با فریم داده شده را پردازش می‌کند.
- هر نرون شبکه عصبی تحت چه شرایطی فعال است؟ به عبارت دیگر یک واحد عصبی پرسپترون چگونه رفتار می‌کند.

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

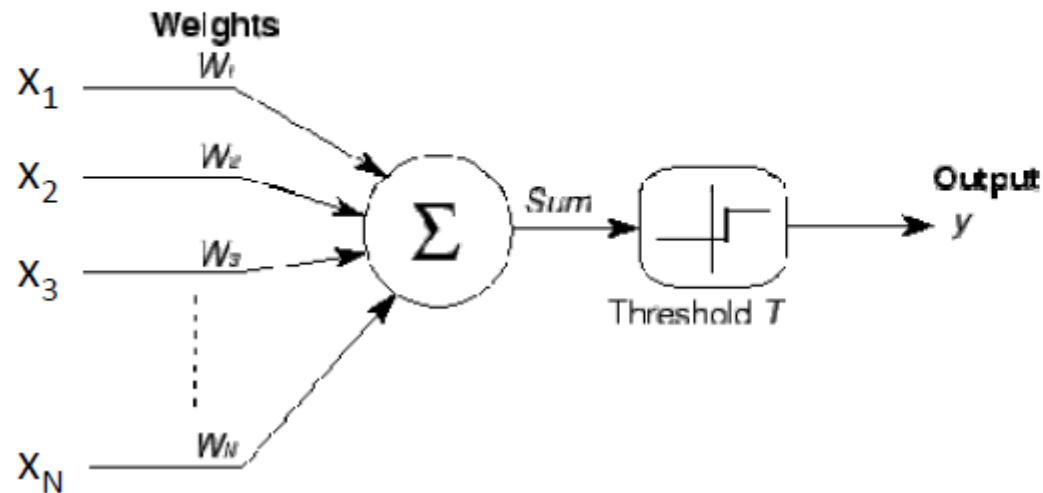
daychegroup

dayche.com | گروه دایکه

What do neurons capture?



اگر همبستگی وزن‌های شبکه با ورودی‌های یک نرون بالا باشد، نرون مد نظر فعال خواهد بود و اطلاعات را منتقل می‌کند.



$$y = \begin{cases} 1 & \text{if } \sum_i w_i x_i \geq T \\ 0 & \text{else} \end{cases}$$

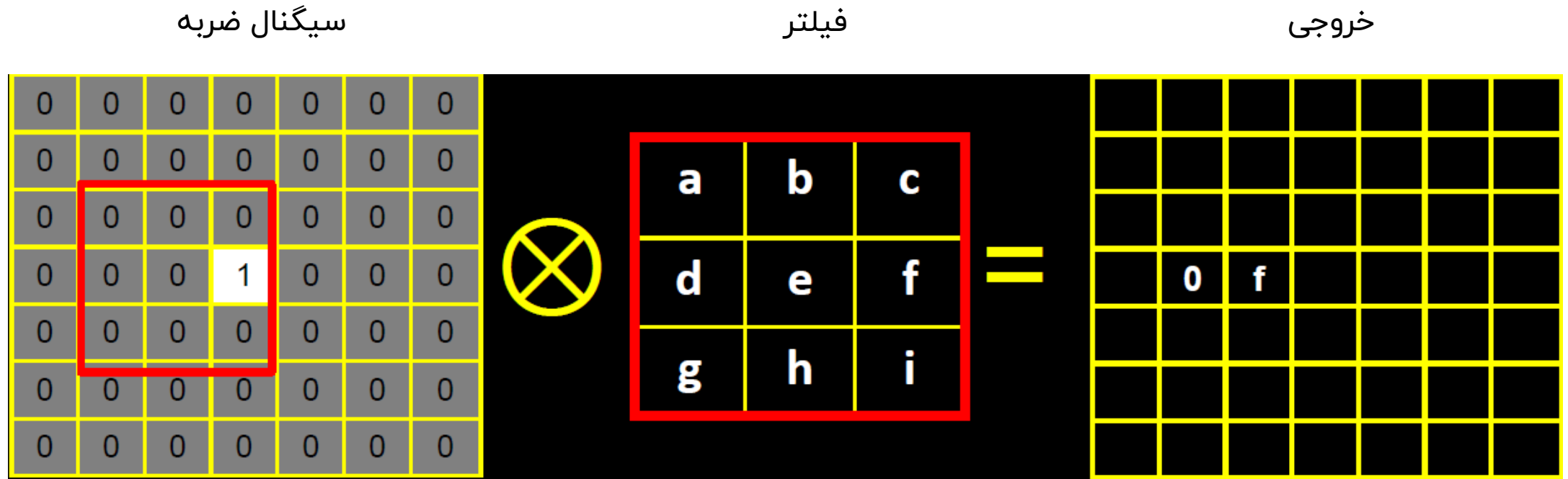
$$y = \begin{cases} 1 & \text{if } \mathbf{x}^T \mathbf{w} \geq T \\ 0 & \text{else} \end{cases}$$

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایکه



$$S(i, j) = \sum_m \sum_n I(m, n) K(i + m, j + n)$$

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

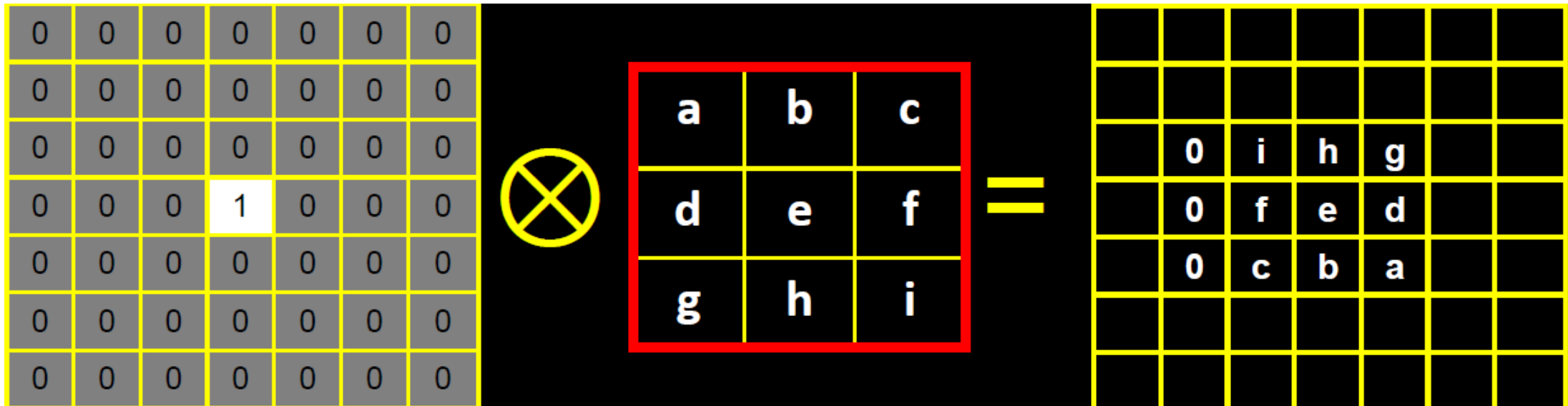
dayche.com | گروه دایکه



سیگنال ضربه

فیلتر

خروجی



چه تفاوتی بین خروجی و فیلتر به کار گرفته شده وجود دارد؟

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

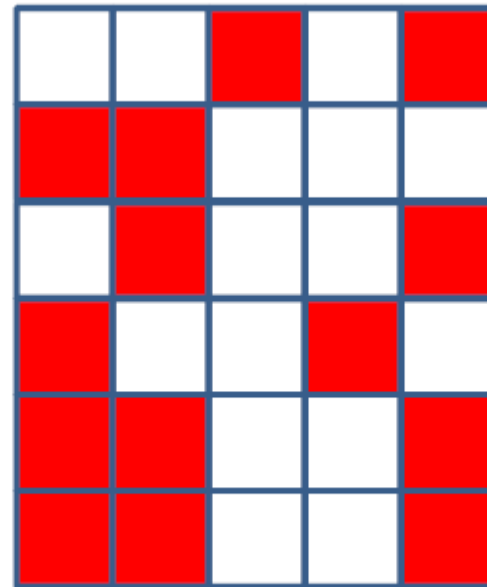
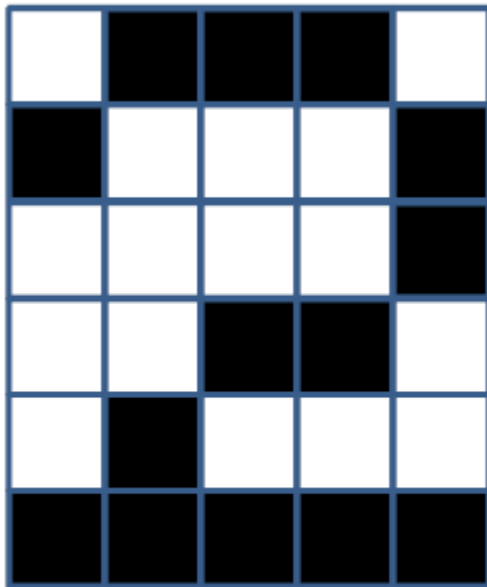
dayche.com | گروه دایچه

Convolution

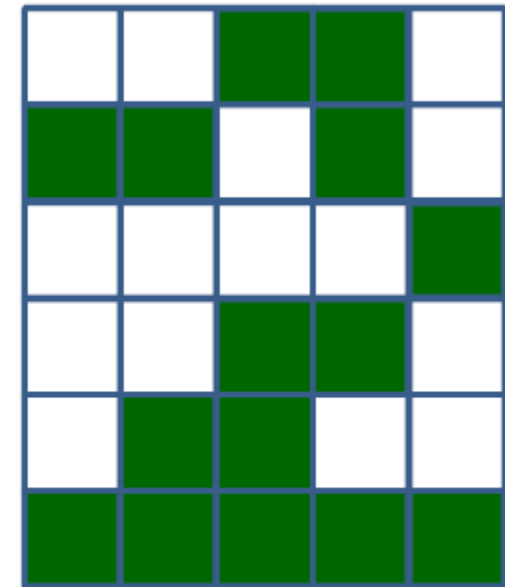


$$S(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n)$$

کانولوشن همان همبستگی را محاسبه می‌کند با این تفاوت که قبل از محاسبه فیلتر به لحاظ افقی و عمودی flip شده باشد.



Correlation = 0.57



Correlation = 0.82

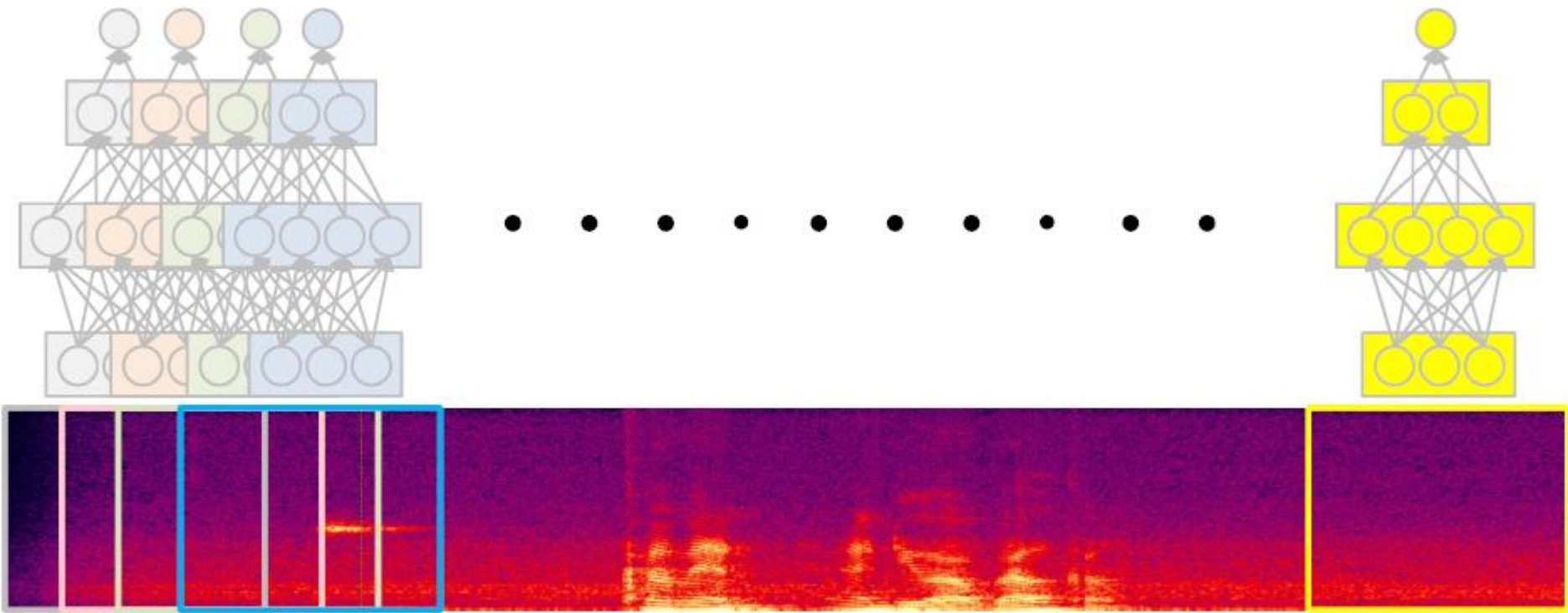
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایکه | dayche.com

Feature map



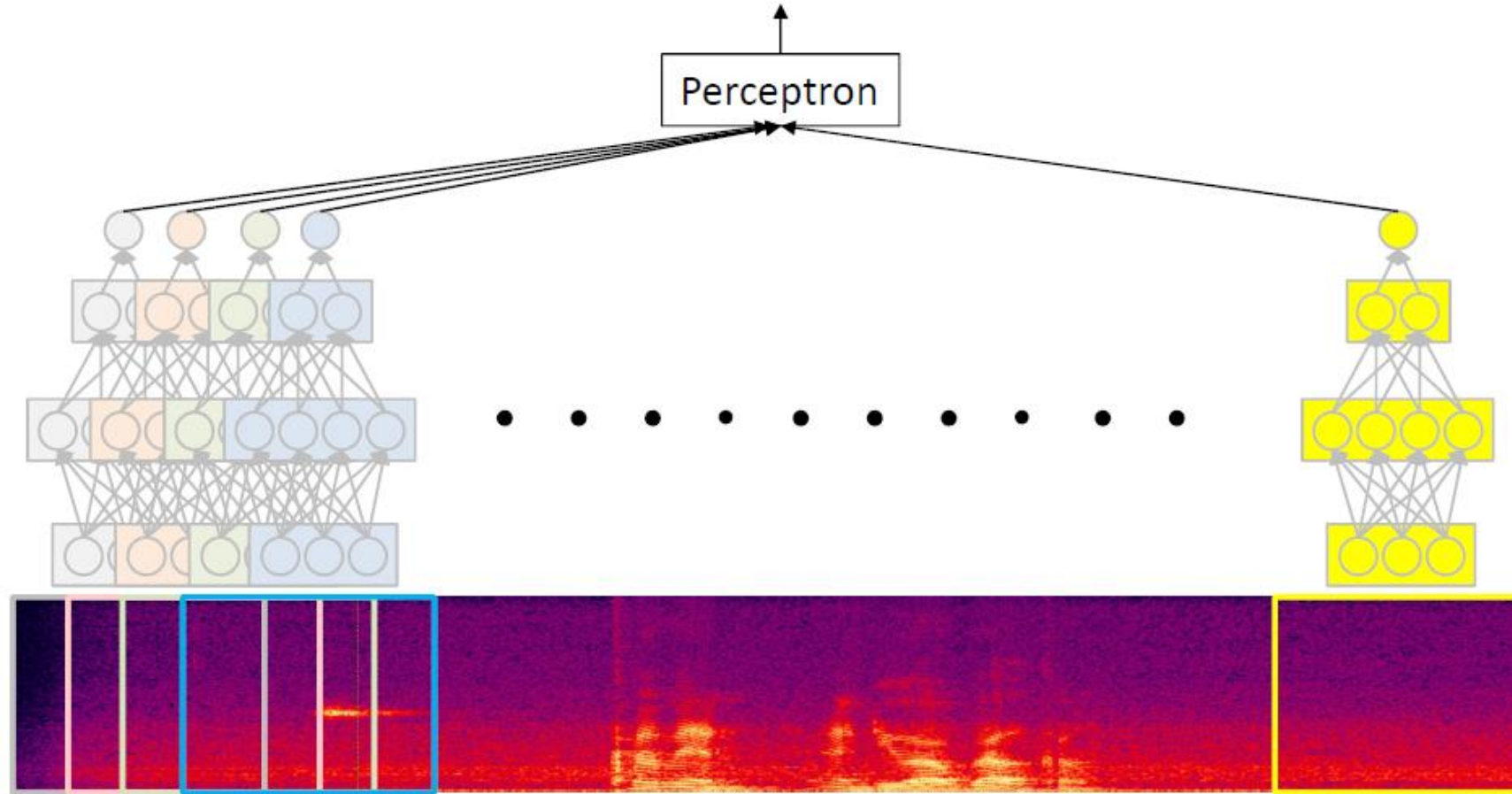
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایکه | dayche.com

Decision making



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

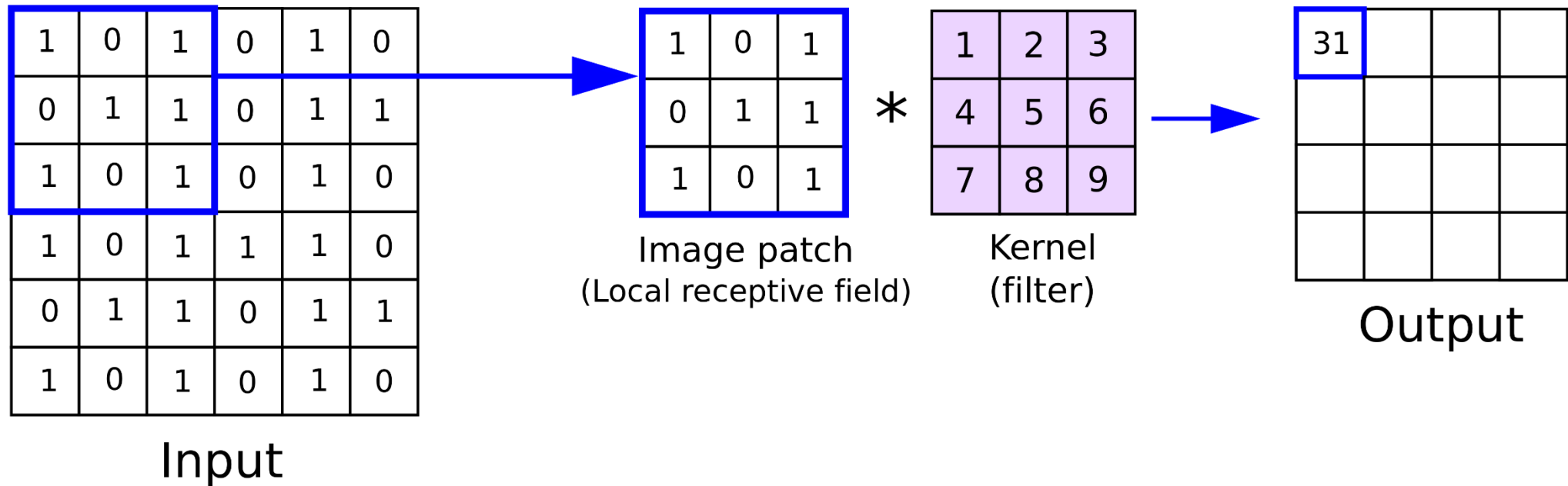
daychegroup

گروه دایچه | dayche.com

What is going on?



ساده ترین مشاهده‌ای که می‌توان داشت کوچک شدن سایز تصویر خروجی است



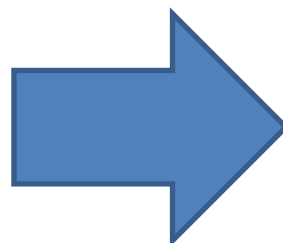
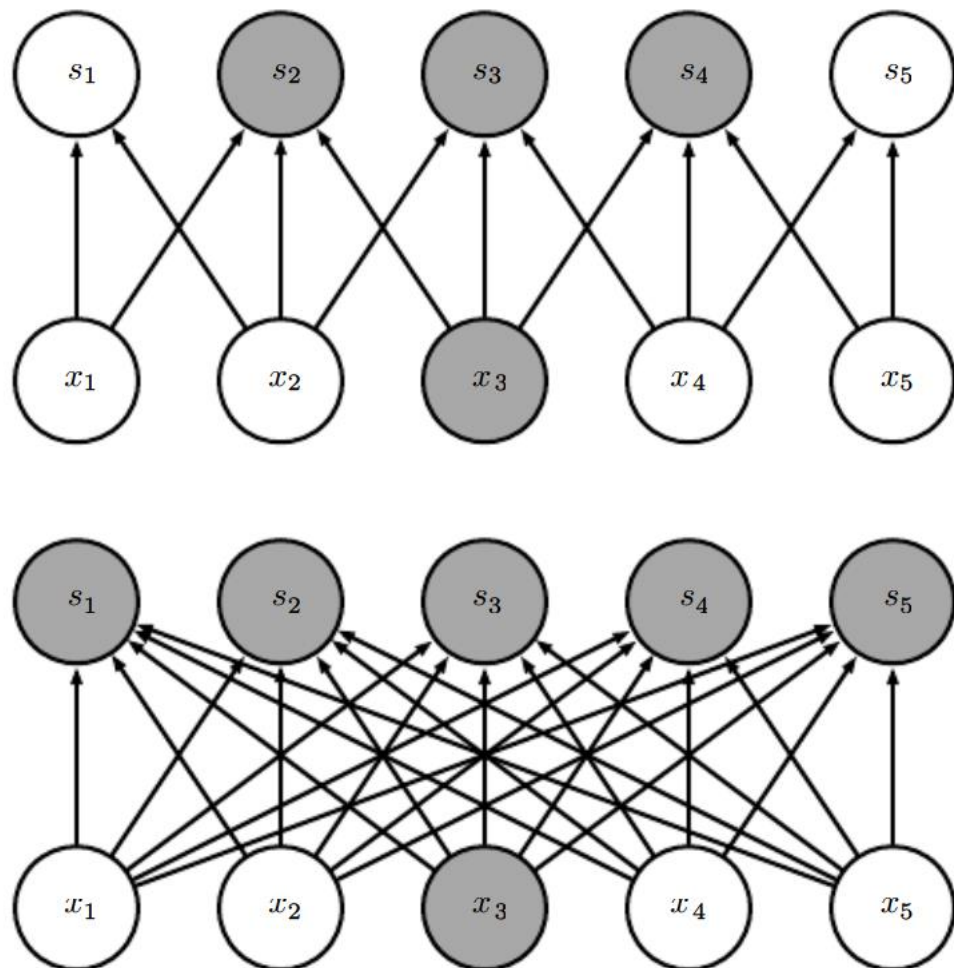
تولید محتوا برای شبکه‌های اجتماعی

daychegroup

daychegroup

dayche.com | گروه دایکه

Sparse connectivity



- اگر تعداد ورودی‌ها برابر با m و تعداد خروجی‌ها برابر با n باشد، در یک شبکه عصبی MLP نیازمند یک ضرب ماتریسی با m سطر و n ستون هستیم.
- در صورت استفاده از کانولوشن این تعداد به محاسبه ضرب ماتریس k سطر و n ستون تقلیل می‌یابد.

دید پایین به بالا

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

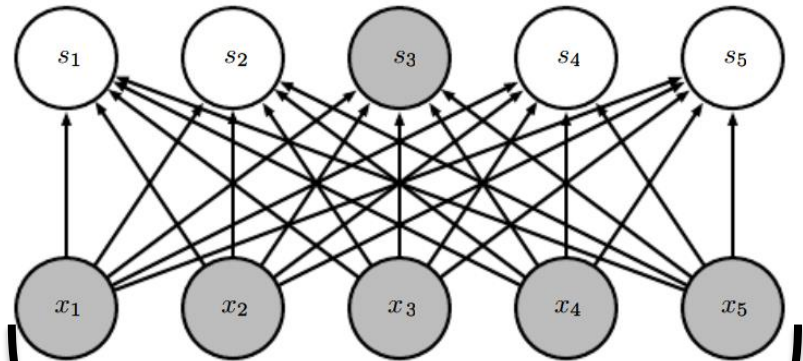
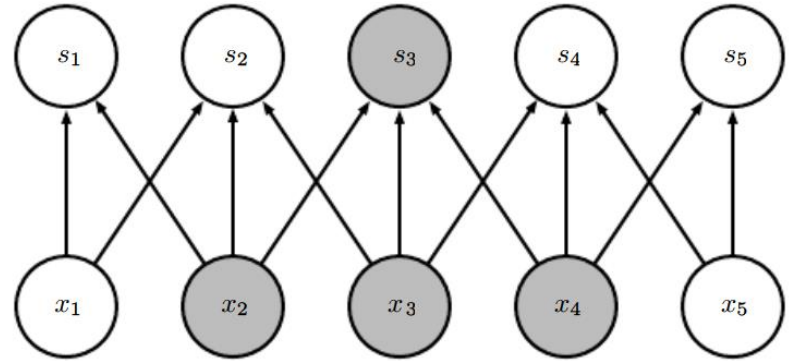
daychegroup

dayche.com | گروه دایچه

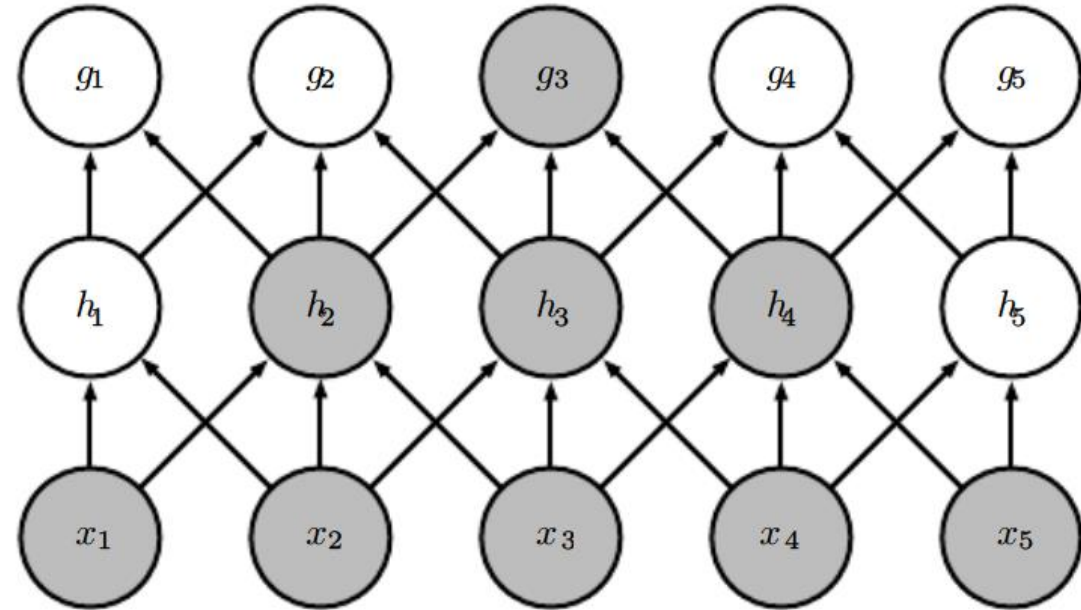
Receptive field



مشکلی که در حذف اتصال کامل و جایگزین کردن کانولوشن وجود دارد کوچک شدن میدان دید هر نرون است. برای حل مشکل شبکه را عمیق می‌کنیم.



Receptive field



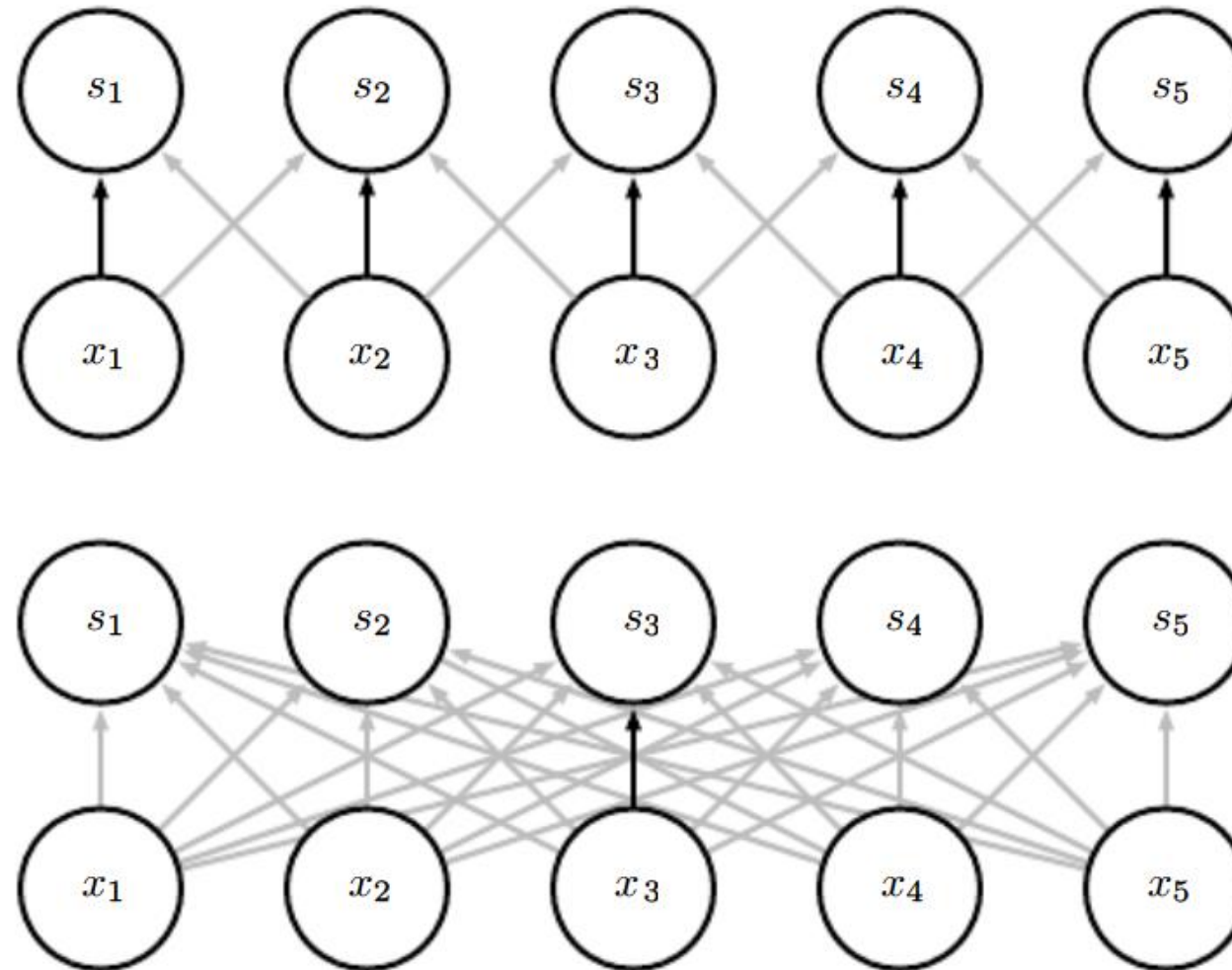
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایچه | dayche.com

Parameter sharing



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

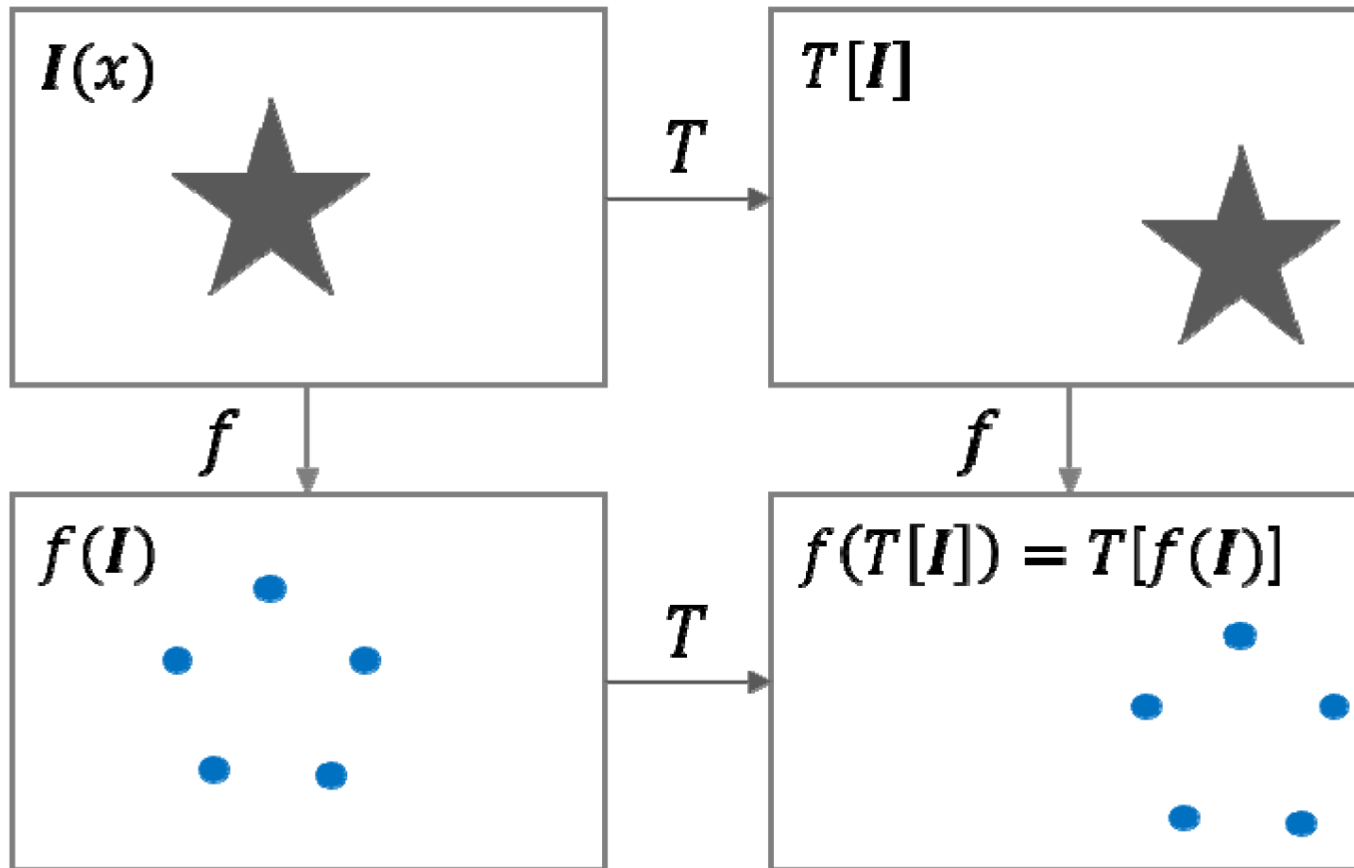
daychegroup

گروه دایچه | dayche.com

Translation invariance



اشتراک‌گذاری پارامترها موجب می‌شود تا ویژگی محاسبه شده برای یک ورودی نسبت به تغییر ورودی مقاوم باشد. وابستگی به تغییر افقی و عمودی را از بین می‌برد. تغییرات نسبت به دوران چطور؟



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

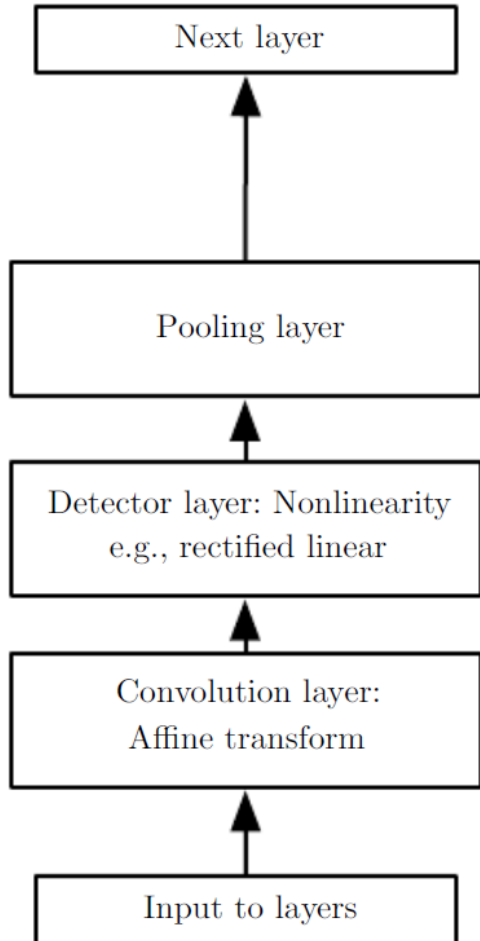
daychegroup

dayche.com | گروه دایکه

Simple convolution layer



ورودی لایه کانولوشن بعدی و یا یک لایه با اتصال کامل



- هر لایه کانولوشن دارای سه مرحله است:
- مرحله اول - کانولوشن (اسکن کردن تمام فضای ورودی به منظور یافتن ویژگی مناسب)
- مرحله تشخیص - عبور ویژگی‌های اسکن شده از یک تابع غیرخطی
- مرحله ادغام

- لایه کانولوشن باعث مقاوم شدن ویژگی خروجی نسبت به جابجایی عمومی می‌شود اما نسبت به دوران مقاوم نیست.
- لایه ادغام مقاومت ویژگی یاد گرفته شده را نسبت به دوران و جابجایی محلی تضمین می‌کند.
- از طریق خلاصه کردن آماره‌هایی از نقشه ویژگی خروجی

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

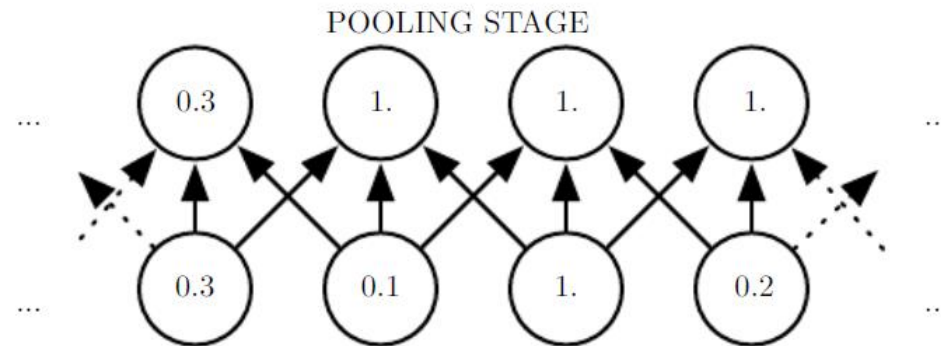
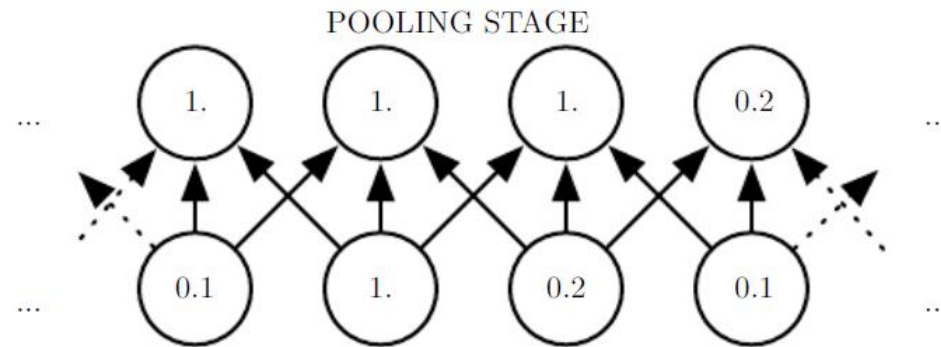
daychegroup

dayche.com | گروه دایچه

Pooling stage – invariance to local translation



- مرحله ادغام، تغییرات نگاشت ویژگی را نسبت به تغییرات کوچک انتقال محلی در ورودی خنثی می‌کند.



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

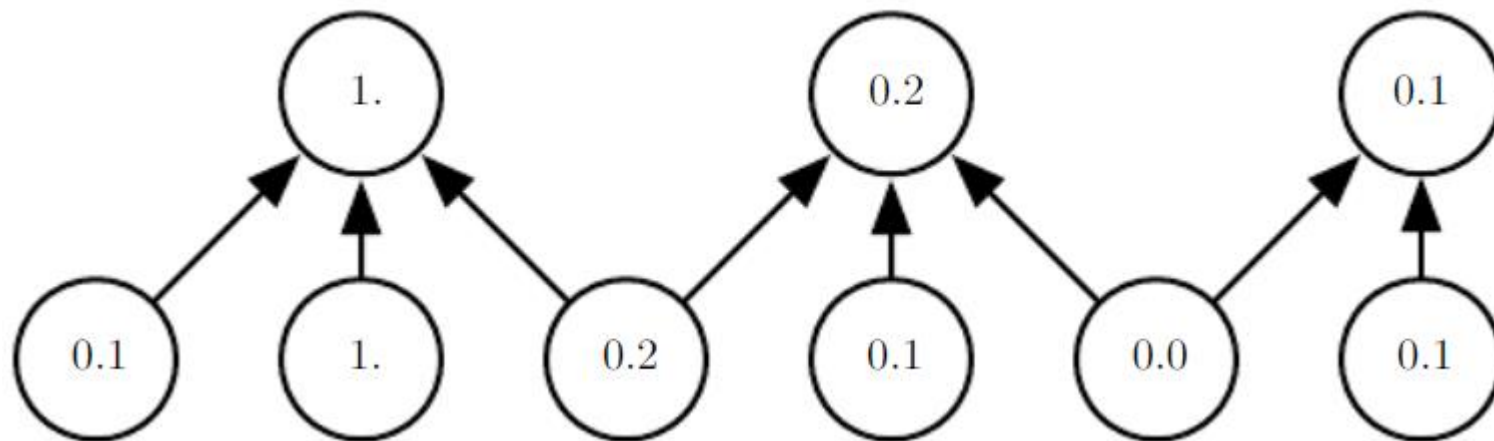
daychegroup

dayche.com | گروه دایچه

Pooling stage – down sampling property



- خلاصه‌ای از ویژگی‌های محاسبه شده در مرحله تشخیص را به دست می‌دهد.
- کاهش نمونه‌ها موجب مقاوم شدن ویژگی آموزش دیده خواهد شد و تنها با تغییر گام ادغام قابل پیاده‌سازی است.



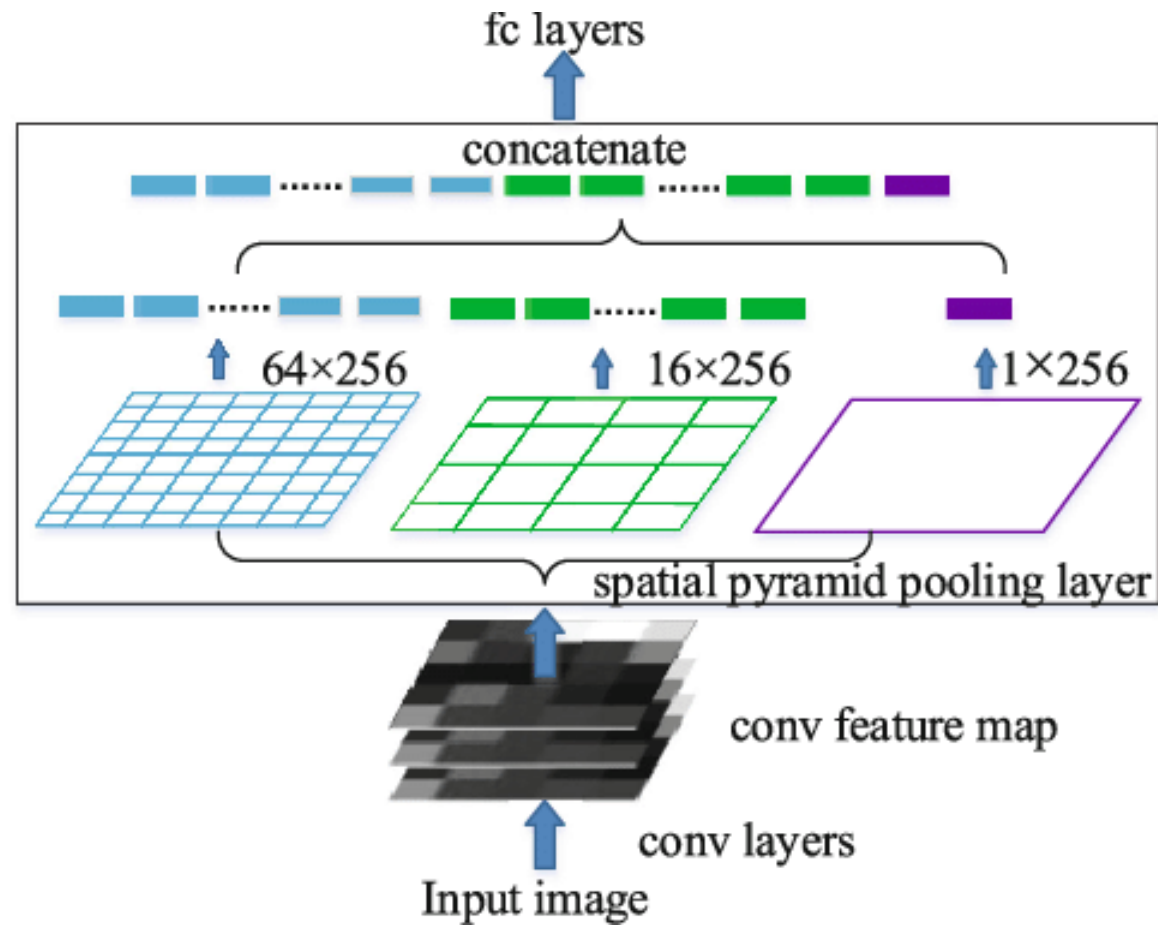
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایچه

Pooling stage – Spatial pyramid pooling



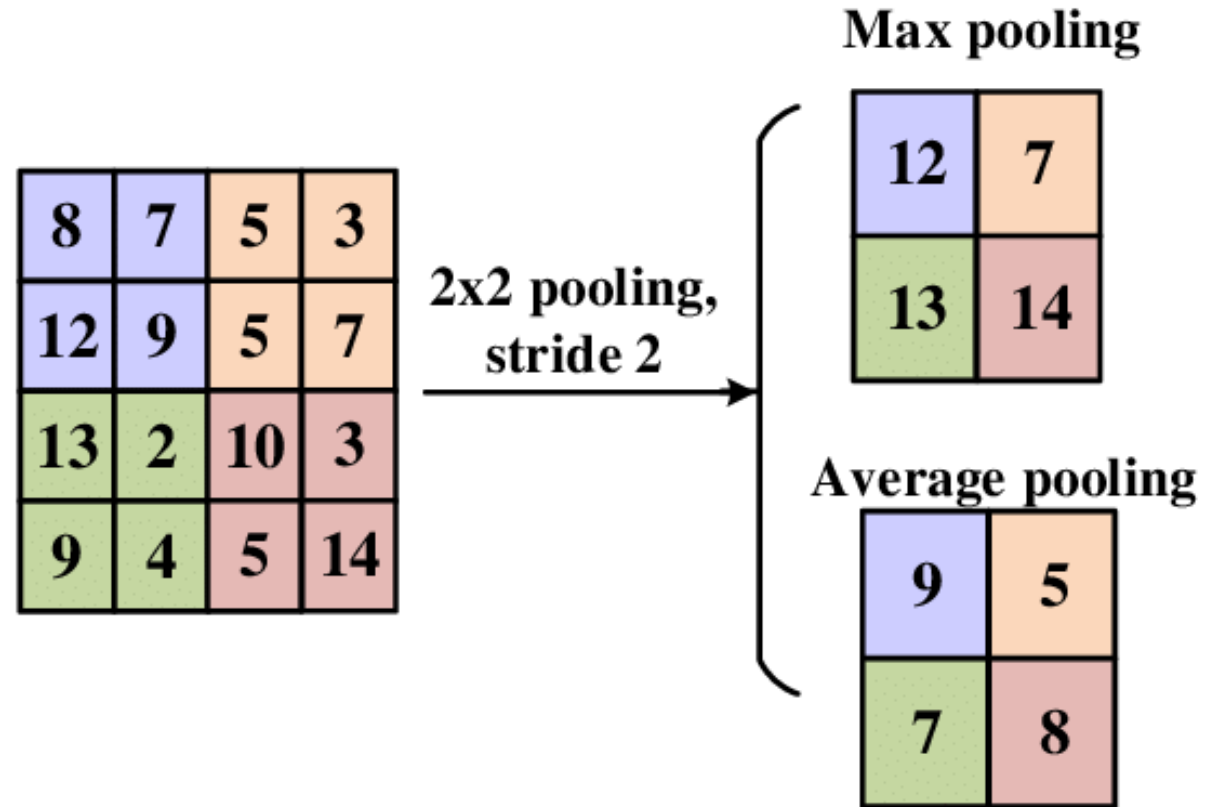
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایچه | dayche.com

Custom pooling operation



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

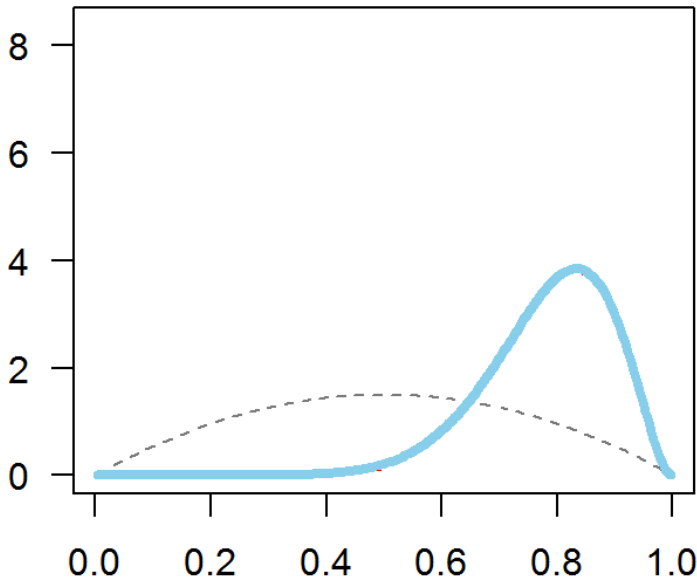
daychegroup

dayche.com | گروه دایچه

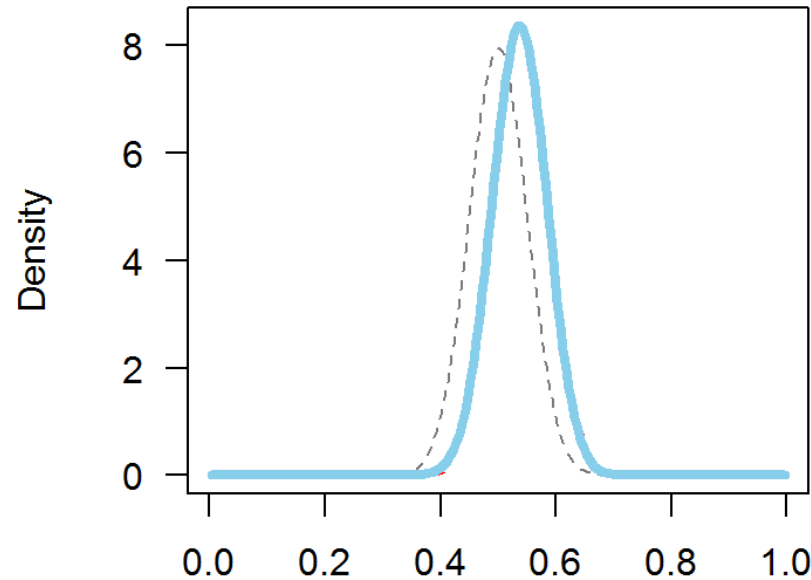
Bayesian interpretation of convolution layer



Weak Prior



Strong Prior



توزیع پیشین استفاده شده در شبکه‌های کانولوشنالی infinitely strong هستند.
• این مطالب تحلیل مطلق هستند و ارتباطی به پیاده‌سازی ندارد.

- توزیع پیشین برای لایه کانولوشن مجبور می‌کند تا وزن‌های پیکسل‌هایی که در یک همسایه قرار می‌گیرند یکسان باشد.
- توزیع پیشین برای عملگر ادغام هر واحد را موظف می‌کند نسبت به تغییرات ورودی مقاوم باشد.
- توزیع پیشین فرض شده برای لایه کانولوشن و ادغام تحت چه شرایطی مناسب نیست؟ (ممکن است موجب underfitting شود)

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

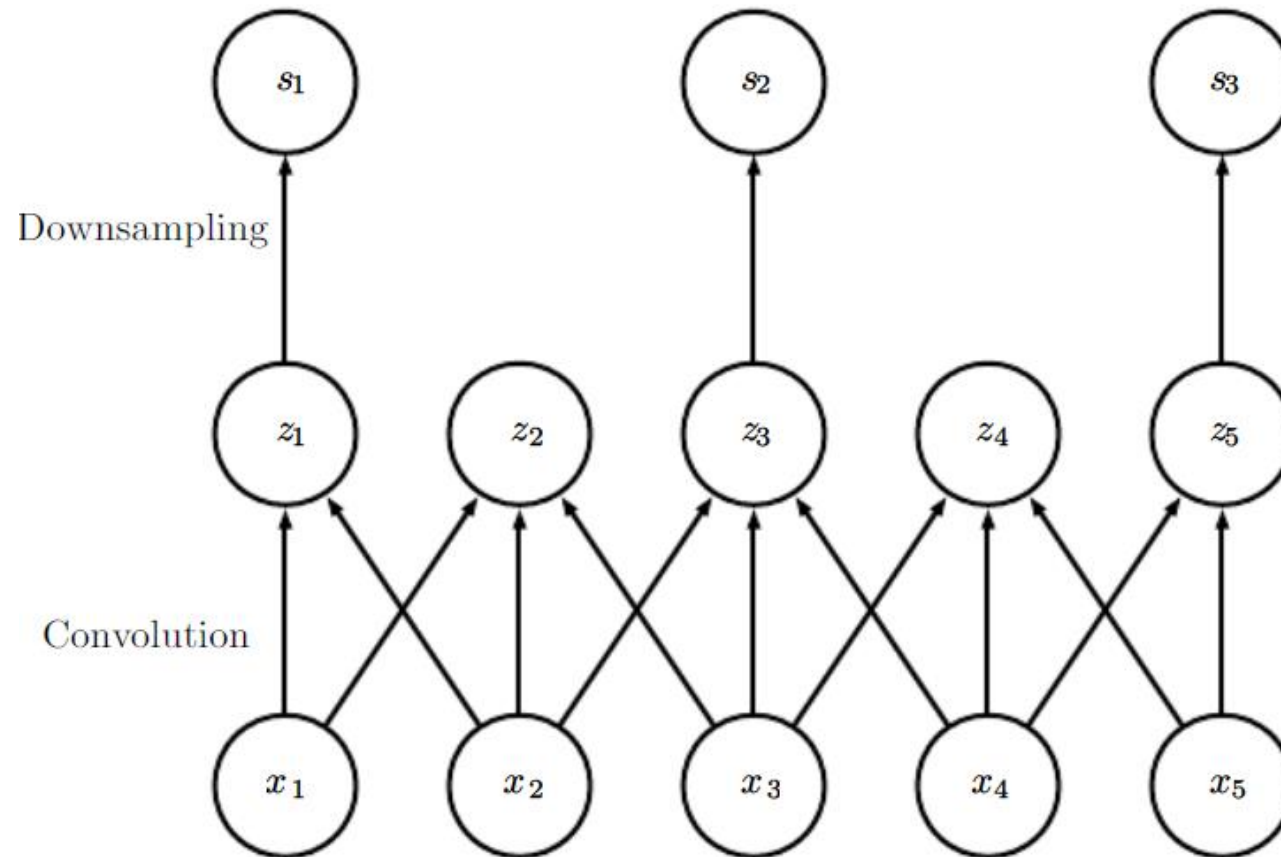
daychegroup

dayche.com | گروه دایچه

A simple convolutional layer



- یک لایه کانولوشن
- ابعاد خروجی لایه کانولوشن بسیار کوچکتر از ابعاد ورودی است. به منظور افزایش قدرت ویژگی نگاشت شده نیازمند عمیق کردن شبکه هستیم، چاره چیست؟



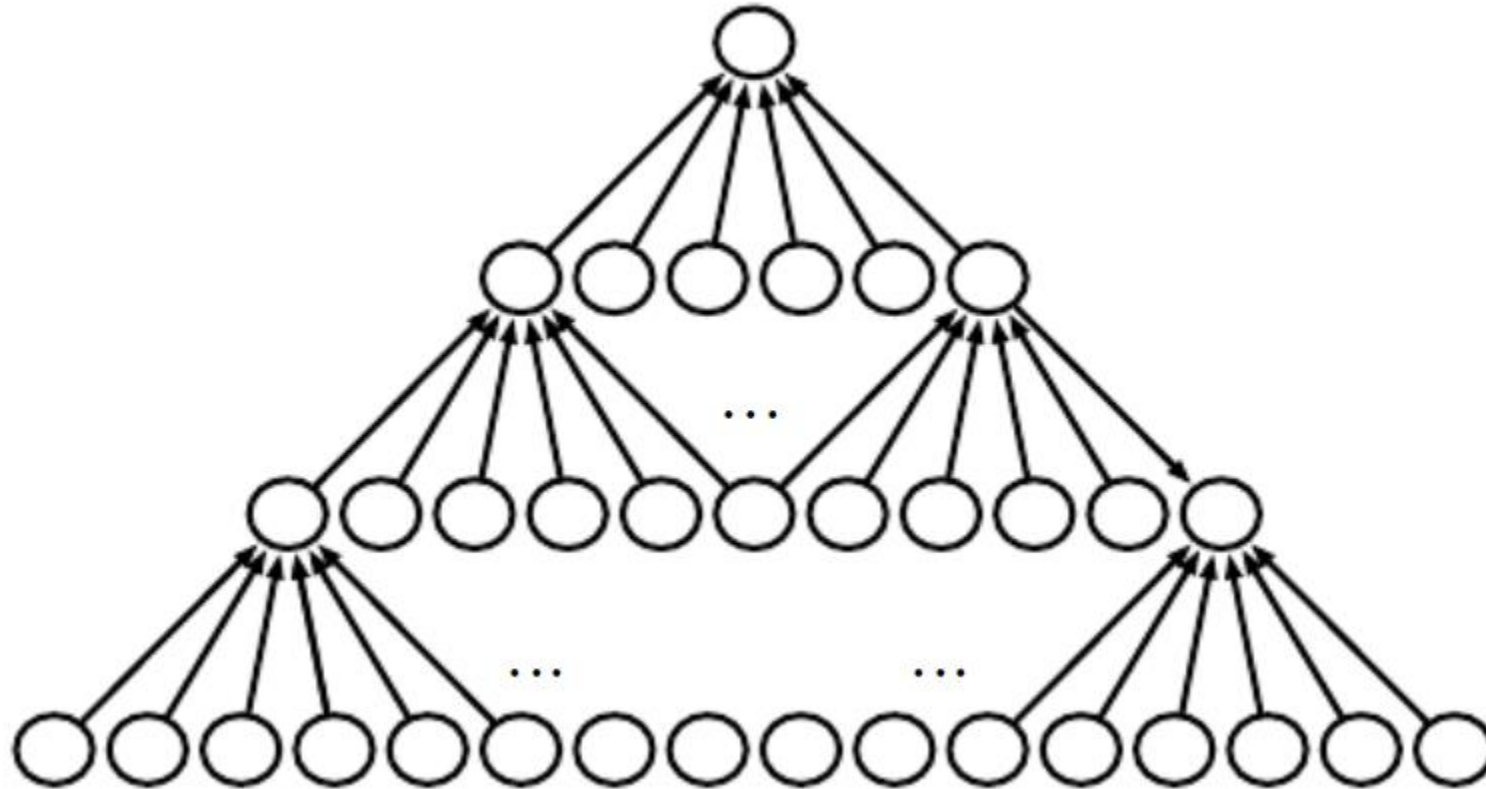
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایکه


Small size of output



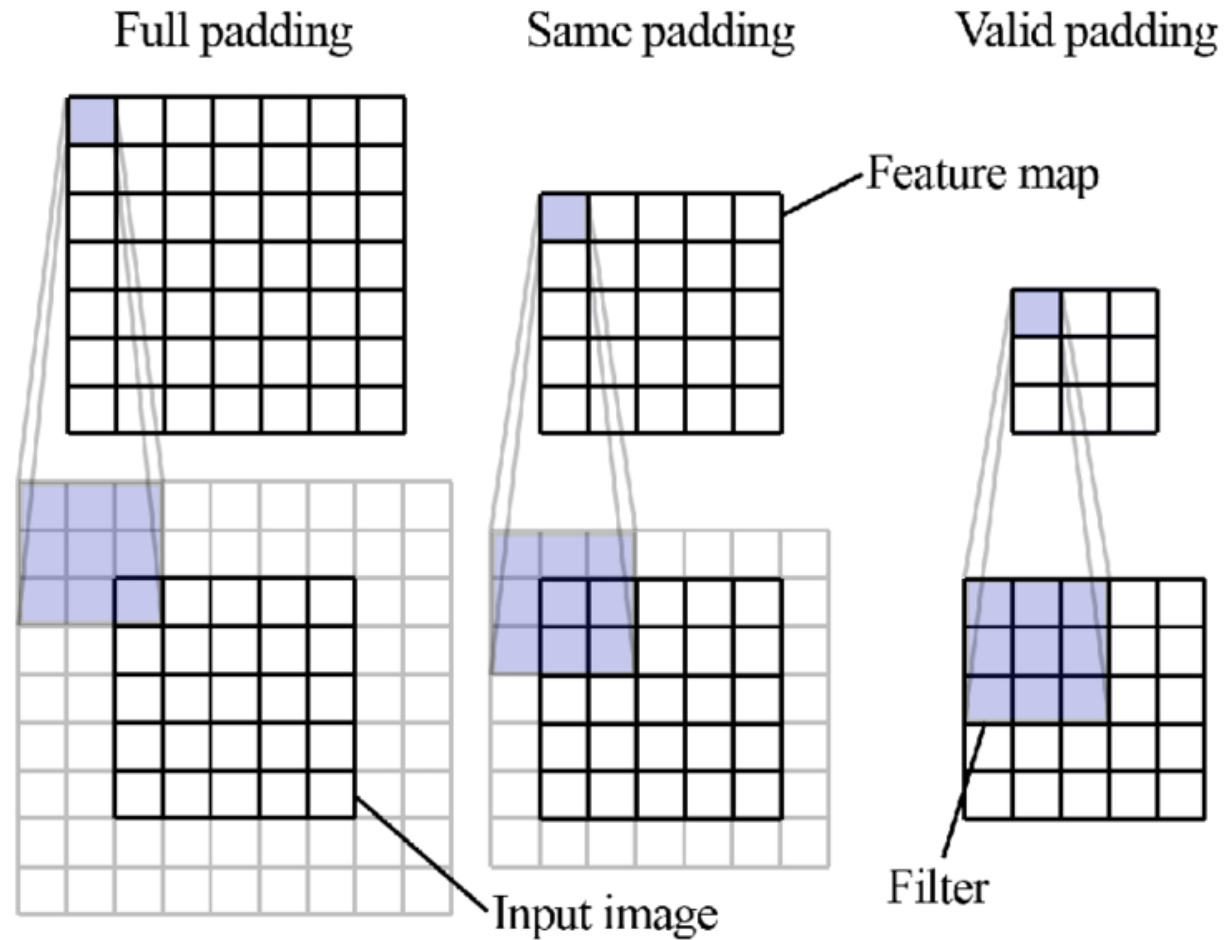
تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

گروه دایکه | dayche.com 

Padding



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایچه


Convolution layer with same padding



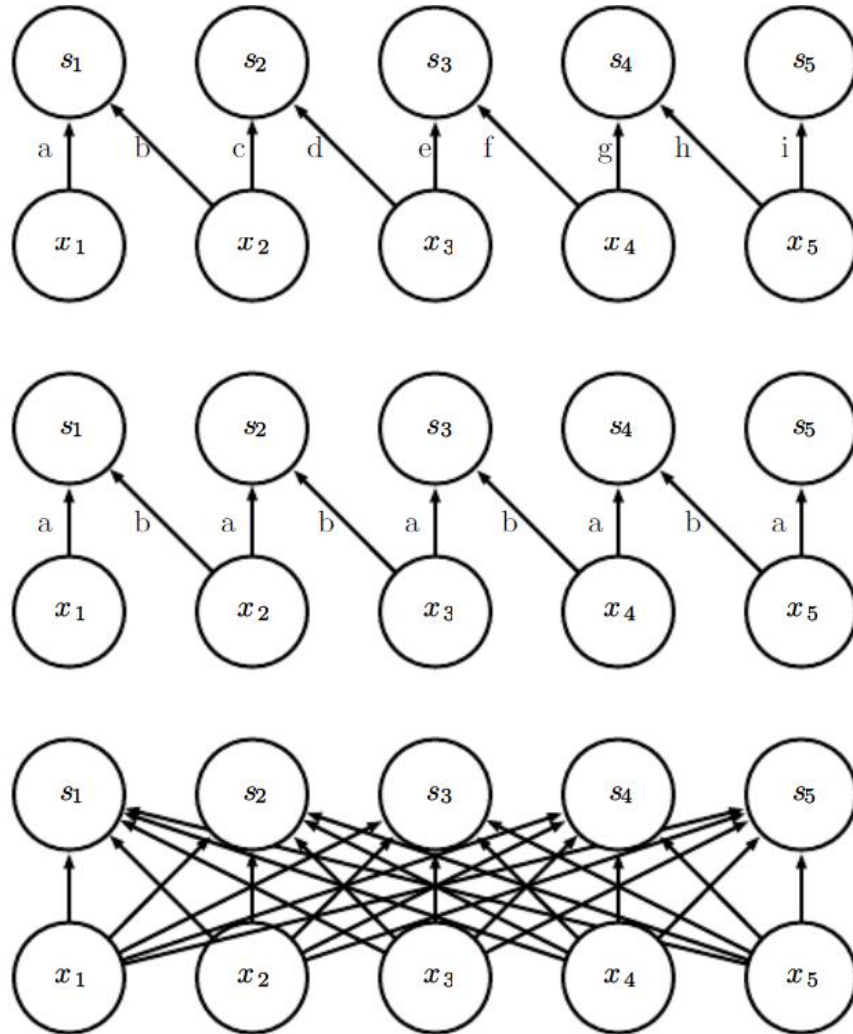
تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

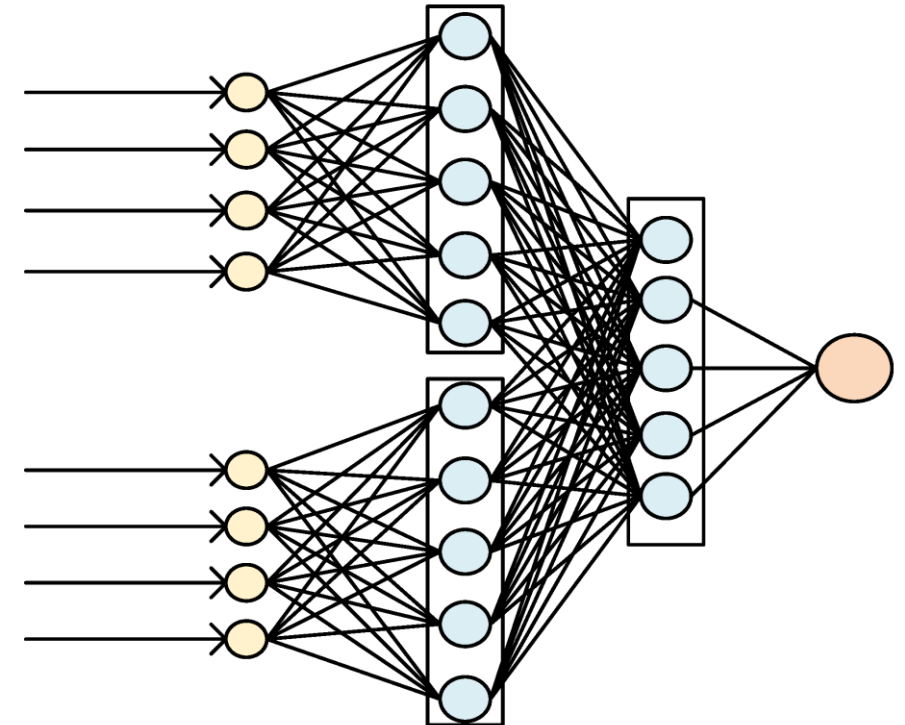
daychegroup 

گروه دایکه | dayche.com 

Local connection – unshared convolution



- شبکه با اتصال محلی (Face recognition)
- نیازمند اطلاعات پیشین درباره مسئله



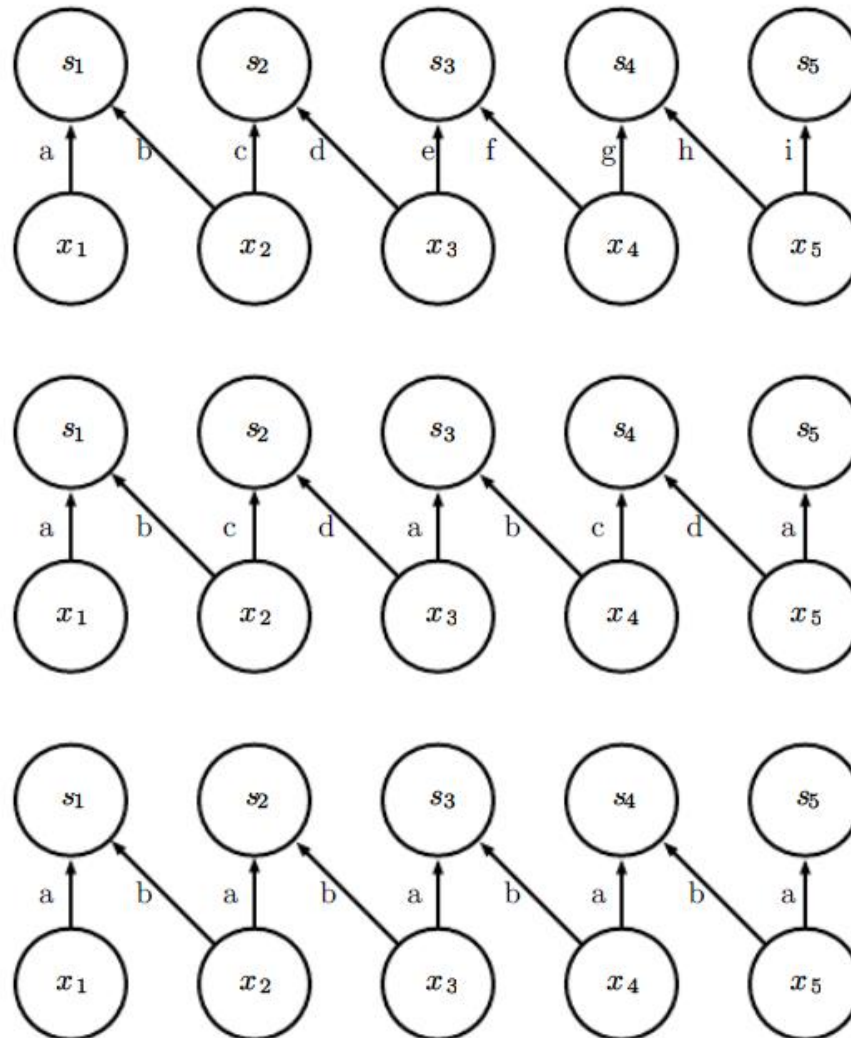
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایچه | dayche.com

Tiled convolution



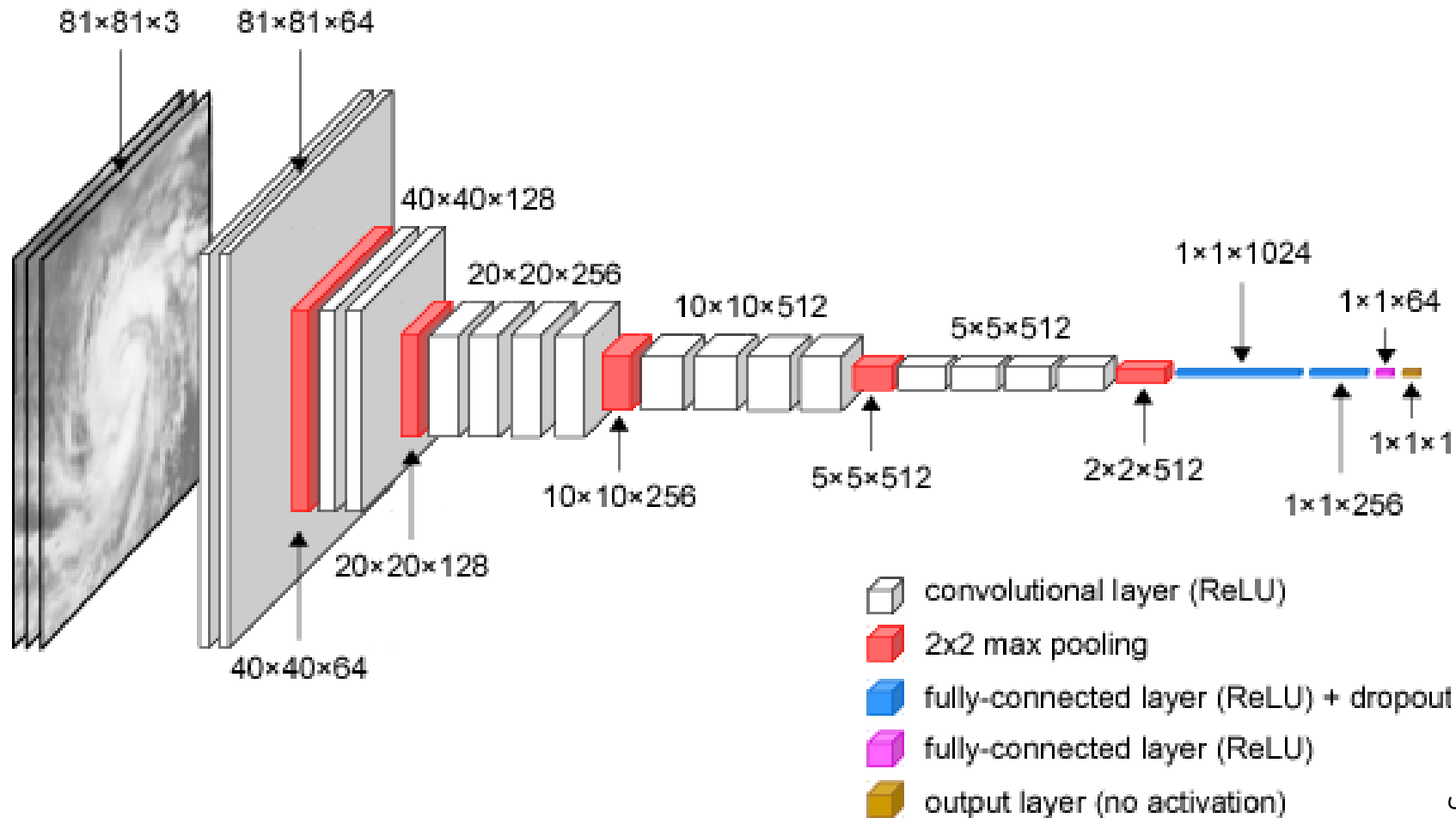
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایچه | dayche.com

Convolutional neural networks



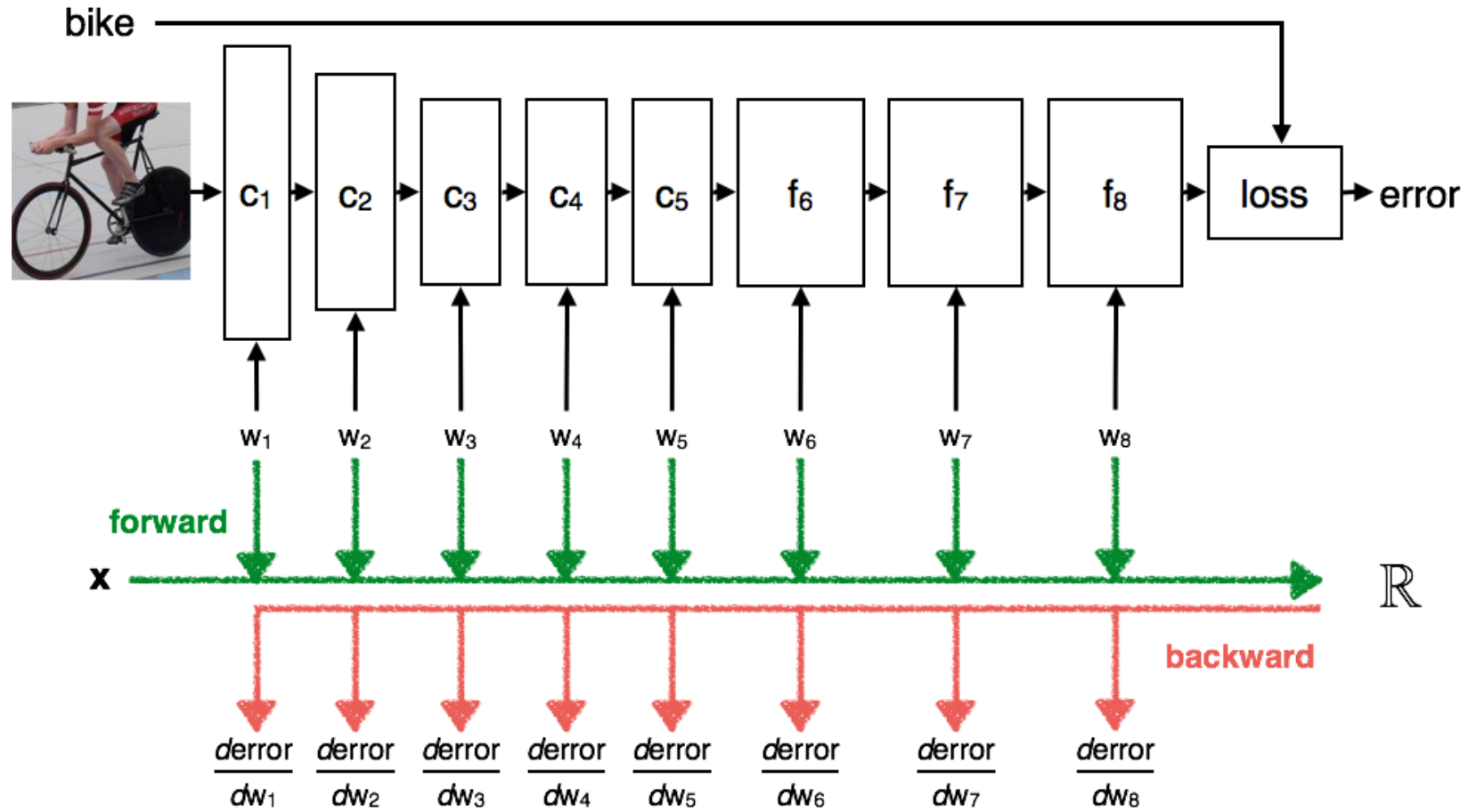
تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایچه | dayche.com

Back propagation



تولید محتوا: وحید محمدزاده ایوقی

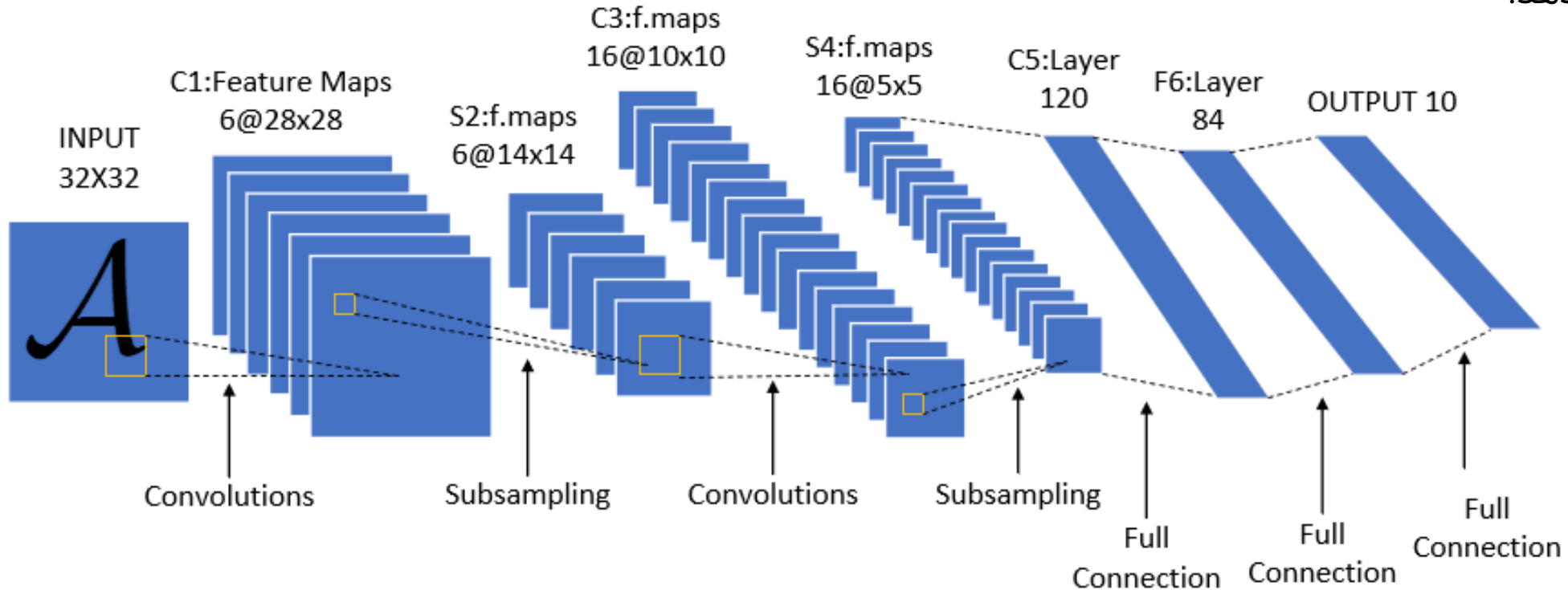
daychegroup

daychegroup

dayche.com | گروه دایچه



استفاده از مدل‌ها و ساختارهای از پیش آموزش دیده شده که به صورت تجربی ارائه شده‌اند، نیازحیاتی به مجموع داده‌های بالا را کاهش می‌دهد.



این شبکه در سال 1998 ارائه شده است و بزرگترین مشکل آن نبود امکانات سخت‌افزاری برای آموزش مدل بود.

تولید محتوا: وحید محمدزاده ایوقی

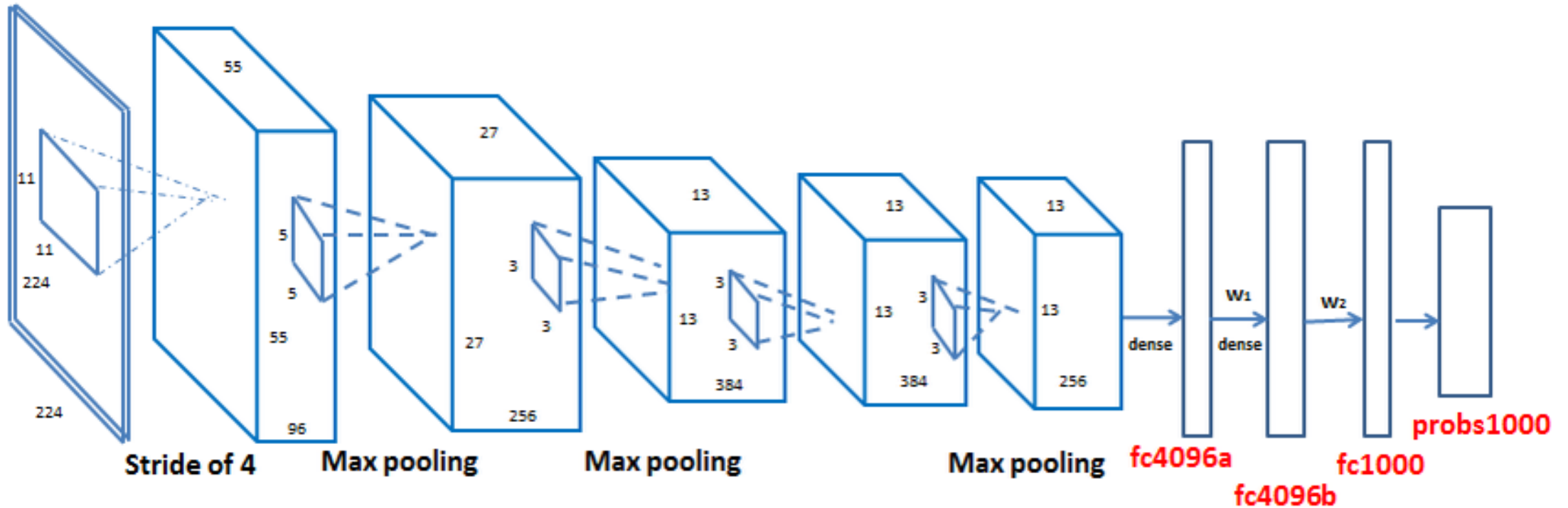
daychegroup

daychegroup

dayche.com | گروه دایکه



در سال 2012 این معماری توسط Alex krizhevsky ارائه شد که برای مشکلات LeNet راه حل داشت.



تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

گروه دایچه | dayche.com