

~~32x32x3~~

الف ۱

① $224 \times 224 \times 3 \rightarrow 224 \times 224 \times 64 \rightarrow ②$

$$w_2 = \frac{w_1 - F + 2P}{5} + 1 = \frac{224 - 3 + 2}{1} + 1 = 224$$

of Parameters = $(3 \times 3 \times 3 + 1) \times 64 = 28 \times 64 = 1792$

② ~~Relu~~ $224 \times 224 \times 64 \rightarrow 3$

(zero) 0 parameter

③ $112 \times 112 \times 64$

$$0 = \frac{I - K + 2P}{5} + 1 = \frac{224 - 2 + 0}{2} + 1$$

zero parameter

~~FC~~ $\rightarrow FC \rightarrow 10 \text{ output}$

parameters = $(112 \times 112 \times 64 + 1) \times 10 = 8028170$

$(224 \times 224 \times 3 + 1) \times 10 = 1505290$

ب

ج

لایه کانولوشن معمولی محدودیت‌های زیادی دارد که Deformable برای رفع یکی از این محدودیت‌ها معنی بردی بودن دست به کار شده است. درواقع این لایه با جمع کردن یک یا چند خاص به اندازه کانولوشن معمولی حساب می‌شود.

درواقع این لایه می‌تواند ویژگی‌های محدوده دیگری را نسبت به کانولوشن معمولی یاد بگیرد.

$$X = (3, 3)$$

$$w = (2, 2)$$

$$Y = (2, 2)$$

$$Y = WX$$

$$(4, 1) = (4 \times 9)(9 \times 1)$$



الف ٢

كبيراً

الوزن

$$WX = \begin{bmatrix} w_{00}x_{00} + w_{01}x_{01} + w_{10}x_{10} + w_{11}x_{11} \\ w_{00}x_{01} + w_{01}x_{02} + w_{10}x_{11} + w_{11}x_{12} \\ w_{00}x_{10} + w_{01}x_{11} + w_{10}x_{20} + w_{11}x_{21} \\ w_{00}x_{11} + w_{01}x_{12} + w_{10}x_{21} + w_{11}x_{22} \end{bmatrix}$$

$$w = \begin{bmatrix} w_{00} & w_{01} & 0 & w_{10} & w_{11} & 0 & 0 & 0 & 0 \\ 0 & w_{00} & w_{01} & 0 & w_{10} & w_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{00} & w_{01} & 0 & w_{10} & w_{11} & 0 \\ 0 & 0 & 0 & 0 & w_{00} & w_{01} & 0 & w_{10} & w_{11} \end{bmatrix} \quad X = \begin{bmatrix} x_{00} \\ \vdots \\ x_{22} \end{bmatrix}$$

الوزن

4x9

9x1

$$\begin{bmatrix} w_{00} & w_{01} & 0 & 0 & w_{10} & w_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{00} & w_{01} & 0 & 0 & w_{10} & w_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{00} & w_{01} & 0 & 0 & w_{10} & w_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{00} & w_{01} & 0 & 0 & w_{10} & w_{11} \end{bmatrix}$$

$$W = \begin{bmatrix} 1 & 2 & 0 & 0 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 4 & 1 \end{bmatrix}$$

$$Y = W^T X$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} 2 \\ 1 \\ 3 \\ 6 \end{bmatrix} =$$

$$\begin{bmatrix} 2 \\ 1 \\ 1 \\ 2 \\ 8 \\ 2 \\ 4 \\ 1 \\ 3 \\ 6 \\ 6 \\ 12 \\ 12 \\ 3 \\ 24 \\ 6 \end{bmatrix}$$

$$\Rightarrow$$

$$\begin{bmatrix} 2 & 1 & 1 & 2 \\ 8 & 2 & 4 & 1 \\ 3 & 6 & 6 & 12 \\ 12 & 3 & 24 & 6 \end{bmatrix}$$

R-CNN (الف ۳) : Region-based
یک روش است که سعی می کند اشیاء را در تصویر مکان یابی کند این روش بدون صورت کاری می کند که ابتدا یک الگوریتم شیء region تعریف نماید صحنه ای که در آن این شیء ها می تواند با توجه به الگوریتم متفاوت باشد و R-CNN نسبت به الگوریتم agnostic است) سپس ^{این} شیء ها را در شبکه CNN می بیند و با توجه به الگوریتم تعریف نموده انتخاب می شود این به حال است به شبکه ۲۰۴۸ صرفاً یکبار محل تصویر به یک داردی شود و دیگر الگوریتم region proposal نداریم به همین خاطر سرعت ۲۰۴۸ و صحت ۲۰۴۸ پیدا می کند.

region-based Fully convolutional R-FCN networks

این شبکه یک شبکه ای است که به صورت کامل کانولوشن انجام می دهد و حساب می کند که یک برداری که تصویر باشد در آن نداشته باشد بدلی که کاره ابتدا یک روش است و بعداً خاص احوال می شود تا بتوان موقعیت را بهتر مشخص کرد.

به صورت دقیق تر این روش با استفاده از کانولوشن کامل و مستقل به تمام جابجایی ها در کل تصویر با شبکه های کنار هم تفاوت این روش با RCNN همان استفاده از کانولوشن مستقل آن است که وقت کمتری را در انداخته موارد دارد.

۲۰۴۸ : ^{only} ۲۰۴۸ این شبکه با این هدف طراحی شده که نسبت به ^{once} ۱۰۰۰ که سریع تر و بدلی کاربرد در حال realtime بسیار مناسب است.

۳ (ب) ۲۵۷۵

۳ (ج) SPP در واقع یک Pooling است که محدودیت ~~یک~~ سبب است
 که به حذف می کند در واقع این را به بررسی اندیشی را به Conv اعمال می شود

و ویژگی ها را یکجا می کند و یک خروجی با طول ثابت به را به Fully Connected تبدیل

می دهد. به بیان دیگر یک تصحیح در مرحله های آخر شبکه داریم زیرا به

Crop و warp را حذف می کند