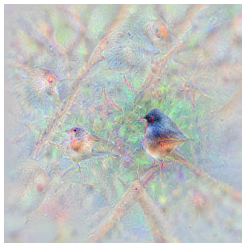


Réseaux de neurones à convolution

Introduction

Laurent Siksous

January 29, 2022



Outline

- 1 Introduction
- 2 Principes
- 3 Représentations visuelles
- 4 Exemple de LeNet-5
- 5 Bibliography

Introduction

- Les réseaux de neurones convolutifs sont très similaires aux réseaux de neurones ordinaires.
- Ils sont constitués de neurones qui ont des poids et des biais apprenables.
- Chaque neurone reçoit des entrées, effectue un produit scalaire et éventuellement applique une non-linéarité.
- Les architectures ConvNet supposent explicitement que les données en entrée sont des images, ce qui permet d'encoder certaines propriétés en dur dans l'architecture.
- Ceux-ci rendent alors la fonction de transfert plus efficace à mettre en œuvre et réduisent considérablement la quantité de paramètres dans le réseau.

Introduction (2/2)

- Pouvez vous identifier le nombre suivant sous forme vectorielle aplatie sans réorganiser les pixels dans un tableau 2D ?

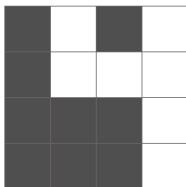


Figure: La disposition spatiale des caractéristiques (pixels) est importante car nous voyons dans une perspective relativiste. C'est là que les réseaux de neurones convolutifs brillent.

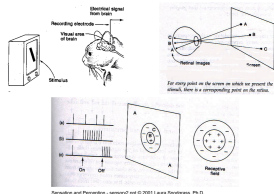
Analogies avec le cortex visuel

HIERARCHY OF VISUAL CELLS

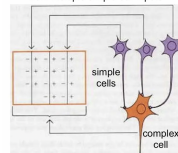
receptive field

[1]

Kuffler's Cat



response of a complex cell, constructed from multiple simple cell inputs



Convolution

- En mathématiques pures, la convolution est un moyen de mesurer comment la forme d'une fonction est modifiée par la forme d'une autre fonction.

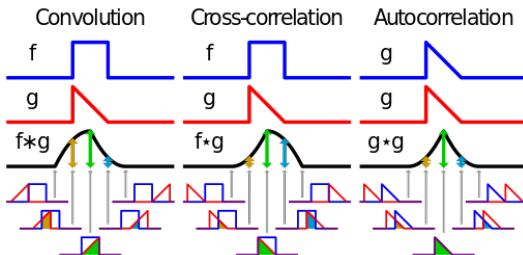







Figure: Produit de 2 fonctions

Histoire

- Le terme est pour la première fois employé par D'Alembert en 1754 dans son ouvrage Recherches sur différens poins importants du système du monde [2].
- Les opérations de convolution apparaissent ensuite dans les travaux de Laplace, Fourier, Poisson, notamment.
- Jusque dans les années 1950, plusieurs autres termes lui sont substitué: Faltung (qui signifie pliage en allemand), produit de composition, intégrale de superposition ou intégrale de Carson.

Filtres à convolution

test

Name	Kernel	Image Result
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Mean Blur	$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$	
Laplacian	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
Gaussian Blur	$\begin{bmatrix} 1/16 & 2/16 & 1/16 \\ 2/16 & 4/16 & 2/16 \\ 1/16 & 2/16 & 1/16 \end{bmatrix}$	

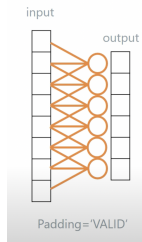
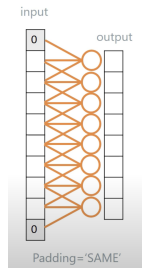
Feature maps

- Dans un réseau de neurones convolutif, chaque filtre h_i est répliqué sur l'ensemble du champ visuel.
- Ces unités (neurones) répliquées partagent les mêmes paramètres (vecteur de poids et biais) et forment une carte de caractéristiques (feature map) appelées aussi cartes d'activation (activation map).

Stride

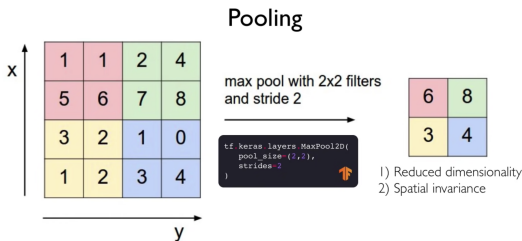
Padding

- Le padding consiste à ajouter des zéros autour du périmètre de l'entrée pour permettre au filtre de la voir en son intégralité.
 - "SAME" padding : la taille de sortie est la même que la taille d'entrée. Cela nécessite que la fenêtre de filtre sorte de la carte d'entrée. Les parties où la fenêtre de filtre est en dehors de la carte d'entrée constituent le padding.
 - "VALID" padding : pas de padding. La fenêtre de filtre reste à l'intérieur de la carte d'entrée tout le temps (dans des positions valides), de sorte que la taille de sortie est plus petite que l'entrée.



Pooling

- Une convolution est le processus d'application d'un filtre ("noyau") à une image.
- Le pooling est le processus de réduction de la taille de l'entrée/image par sous-échantillonnage.
- Mais les deux partagent le même principe d'application d'une fenêtre et d'une foulée sur le champ visuel



How else can we downsample and preserve spatial invariance?

Figure: Max pooling

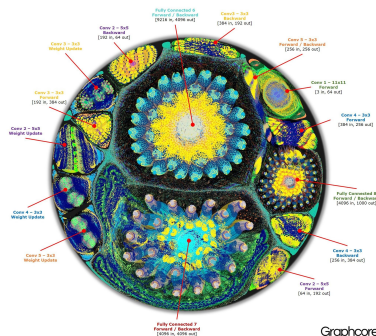
To POOL or CONV ?

Dans leur article de 2014, Striving for Simplicity: The All Convolutional Net [3], Springenberg et al. recommande de supprimer entièrement la couche POOL et de s'appuyer uniquement sur les couches CONV avec une plus grande stride pour gérer le sous-échantillonnage des dimensions spatiales du volume. Leurs travaux ont démontré que cette approche fonctionne très bien sur une variété d'ensembles de données, y compris CIFAR-10 (petites images, faible nombre de classes) et ImageNet (grandes images d'entrée, 1 000 classes).

Cette tendance se poursuit avec l'architecture ResNet, qui utilise également des couches CONV pour le sous-échantillonnage. Il devient de plus en plus courant de ne pas utiliser les couches POOL au milieu de l'architecture réseau et d'utiliser uniquement le Average Pooling à la fin du réseau si les couches FC/Dense doivent être évitées.

Représentation visuelle

- Tenter de comprendre quels calculs sont effectués à chaque couche dans les réseaux de neurones est une direction de recherche très populaire [4].
- Une approche consiste à étudier:
 - chaque couche en tant que groupe
 - le type de calcul effectué par l'ensemble des neurones d'une couche



2D convolutional layer

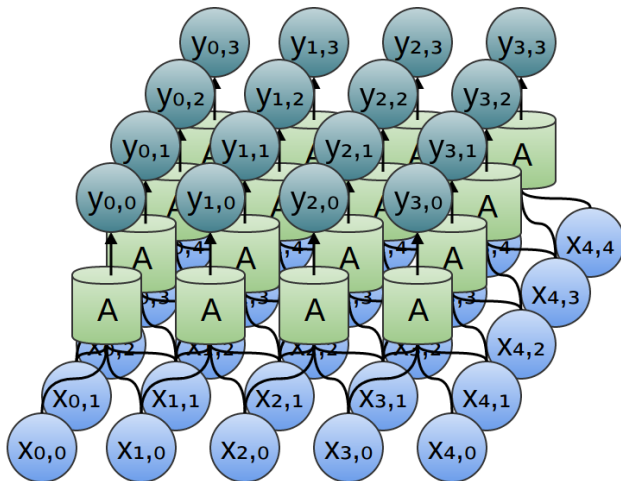


Figure: Conv layer 2D with 1 filter

LeNet-5

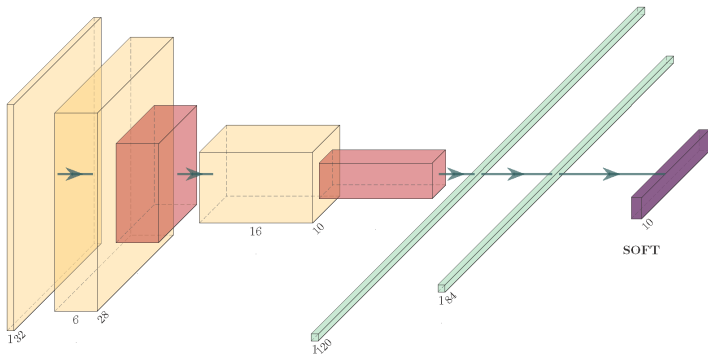


Figure: LeNet architecture

References



David H. Hubel.

Evolution of ideas on the primary visual cortex, 1955–1978: A biased historical account.

Bioscience Reports, 2(7):435–469, July 1982.



Jean Le Rond d'Alembert and undefined.

Recherches Sur Différens Points Importans Du Système Du Monde, Par M. d'Alembert...

1754.



Jost Tobias Springenberg, Alexey Dosovitskiy, Thomas Brox, and Martin Riedmiller.

Striving for Simplicity: The All Convolutional Net.

arXiv:1412.6806 [cs], April 2015.

Comment: accepted to ICLR-2015 workshop track; no changes other than style.



Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs, and Hod Lipson.

Understanding Neural Networks Through Deep Visualization.

arXiv:1506.06579 [cs], June 2015.

Comment: 12 pages. To appear at ICML Deep Learning Workshop 2015.