

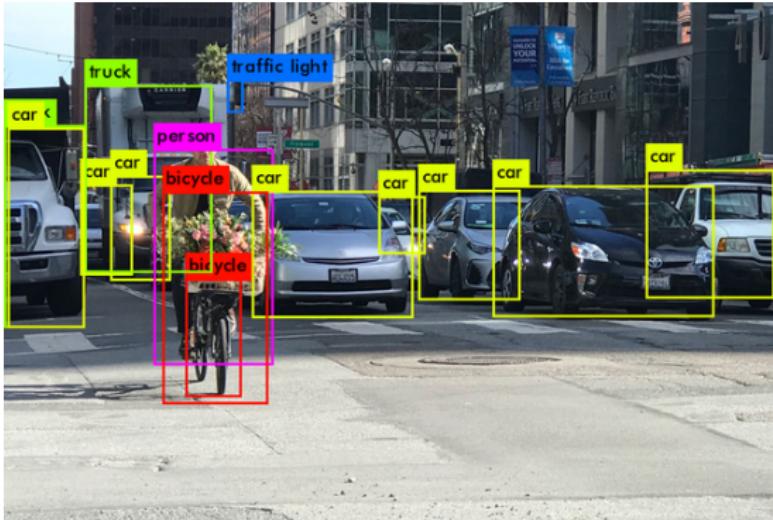
Detail-Preserving Pooling in Deep Networks

Saeedan F, Weber N, Goesele M et al.

Mars 2021



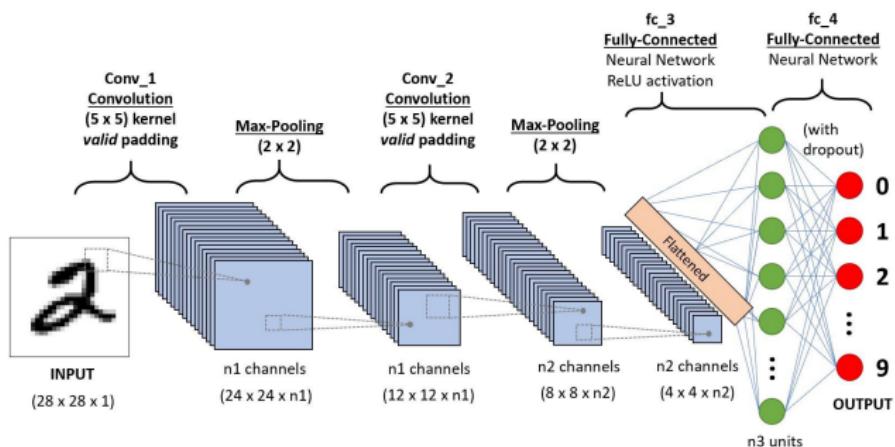
CNN : Applications



Source : A Quora Post

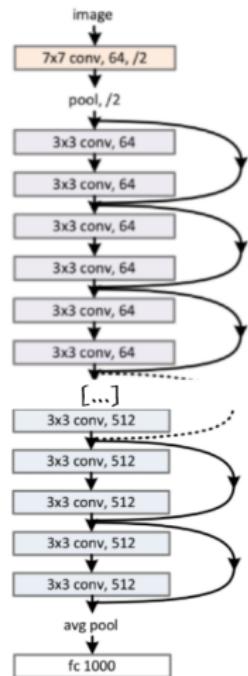
- Classification d'images
- Détection d'objets
- Reconnaissance de caractères
- Détection d'émotions
- Extraction de caractéristiques

CNN : Détails



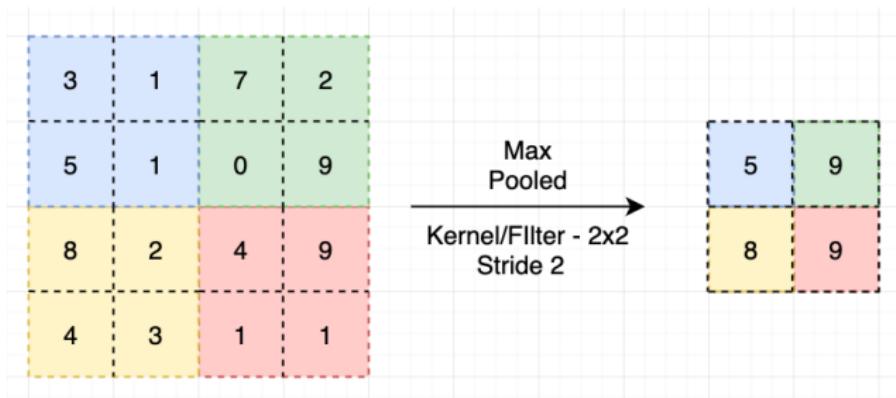
Source : Towards Data Science

34-layer residual



Source : Residual Learning Paper

Le Pooling : Principe



Source : Artificial Intelligence in Plain English

- Max pooling
- Average pooling
- Min pooling
- Pooling mixte
- L_p pooling
- Pooling spectral
- Strided convolution

Importance du Pooling

- Réduit progressivement la taille des couches cachées
- Réduit le nombre de paramètres
- Améliore l'invariance à certaines distorsions
- Augmente l'importance donnée aux pixels qui contribuent à la détection des caractéristiques
- Réduit le risque de surapprentissage

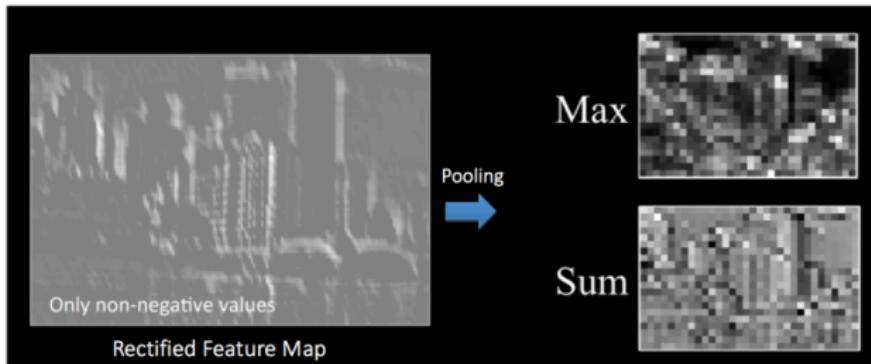


Source : Colossal

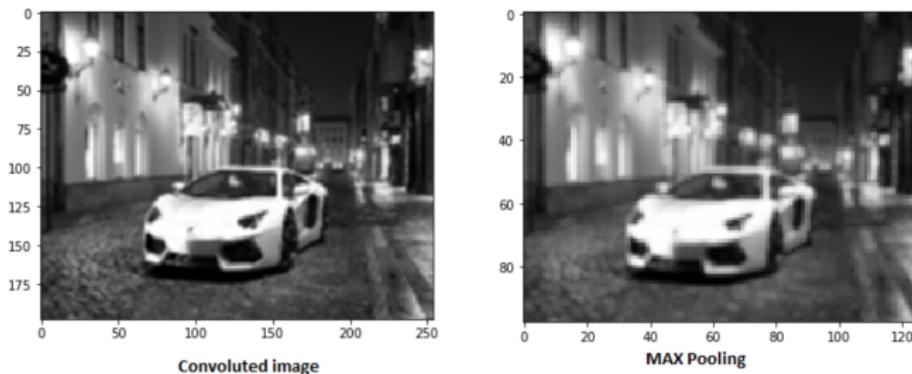
Problèmes du Pooling

- Processus avec perte
- Non dérivable
- Perturbe le flux des gradients
- Il faut souvent déterminer la meilleure stratégie avec des tests empiriques

Problèmes du Pooling



Source : Présentation de Rob Fergus



Source : Analytics Vidhya

Solution Proposée : Detail-Preserving Pooling (DPP)

- Inspiré du système visuel humain qui se concentre sur les changements spatiaux locaux
- Basé sur le filtre bilatéral inverse, utilisé pour accentuer les différences plutôt que de les punir
- Les petits détails transportent plus d'informations que les grandes régions avec des couleurs similaires
- Méthode de pooling adaptative qui amplifie les changements spatiaux et préserve les détails structurels importants
- Complètement dérivable
- Avec des paramètres qui permettent d'ajuster le niveau de détail conservé

DPP : Étapes

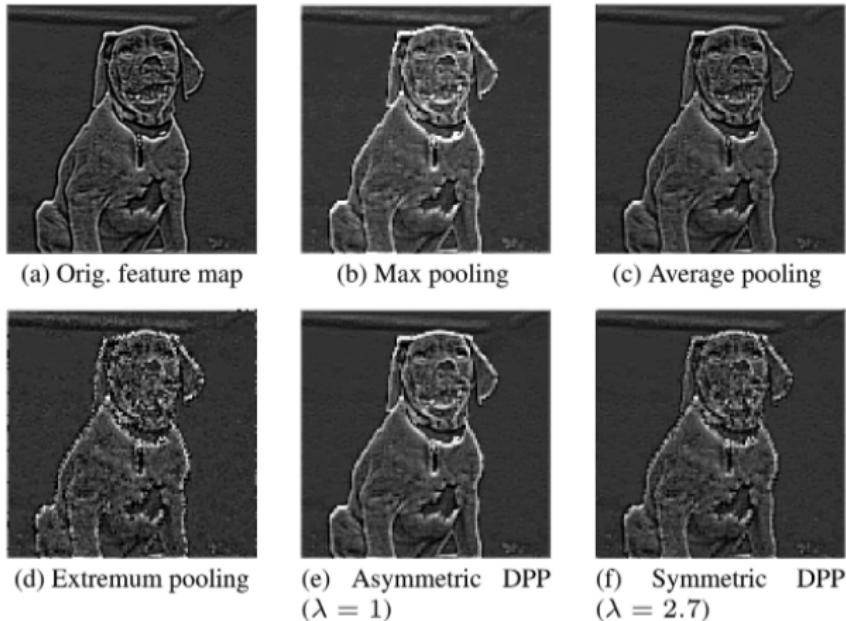
Pour chaque carte de caractéristiques :

- ① Sous-échantillonnage
- ② Prise en compte de la distance entre les valeurs du pixel central et ceux dans la fenêtre
- ③ Normalisation des résultats
- ④ Obtention de la nouvelle valeur

Source : vdumoulin

DPP : Détails

- Deux fonctions pour calculer les distances
- Une version complète et une autre allégée
- Une version stochastique



DPP : Résultats

	Method	VGG	NIN	ResNet
Deterministic methods	Strided conv.	8.43±0.20	10.97±0.10	6.23 ^(*)
	Max	7.43±0.20 ^(*)	9.42±0.07	6.52
	Average	7.12±0.18	8.75±0.15	6.33
	NIN	—	9.01±0.11 ^(*)	—
	Mixed (50/50)	7.27±0.20	8.68±0.23	6.05
	Gated	7.25±0.14	8.67±0.22	7.12
	L_2	7.15±0.18	8.65±0.12	7.29
	Lite-DPP _{Asym}	<u>7.10</u> ±0.15	8.62±0.10	6.17
Stoch.	Full-DPP _{Asym}	<u>7.17</u> ±0.18	8.73±0.05	6.23
	Lite-DPP _{Sym}	<u>7.19</u> ±0.10	8.58±0.11	6.05
	Full-DPP _{Sym}	7.02 ±0.18	8.70±0.14	5.97
	Stochastic	7.67±0.10	8.92±0.09	5.83
S3pool	S3pool	7.21±0.14	<u>7.23</u> ±0.08	<u>5.55</u>
	Lite-S3DPP _{Sym}	—	7.13 ±0.09	5.42

Conclusion

- Nouvelle couche de pooling dont l'objectif est de préserver le plus de détails possible
- Un sous-échantillonage qui se concentre sur la préservation de l'information avec la prise en compte de la distance entre les valeurs des pixels
- Performe des tâches similaires au max et average pooling
- Des résultats meilleurs sur plusieurs architectures

Fin

Des Questions ?

People with no idea about AI
saying it will take over the world:

My Neural Network:



Source : Joke Jet

L'article

Détails sur les CNN

Détection de chiffres interactive

Pourquoi les fonctions d'activation sont nécessaires

La convolution en détail

Les couches de pooling

Revue sur les méthodes de pooling

Cartes de Caractéristiques

Source : Présentation de Rob Fergus

DPP : Détails

$$\mathcal{D}_{\alpha,\lambda}(I)[p] = \frac{1}{\sum_{q' \in \Omega_p} w_{\alpha,\lambda}[p, q']} \sum_{q \in \Omega_p} w_{\alpha,\lambda}[p, q] I[q]$$

$$w_{\alpha,\lambda}[p, q] = \alpha + \rho_\lambda \left(I[q] - \tilde{I}[p] \right)$$

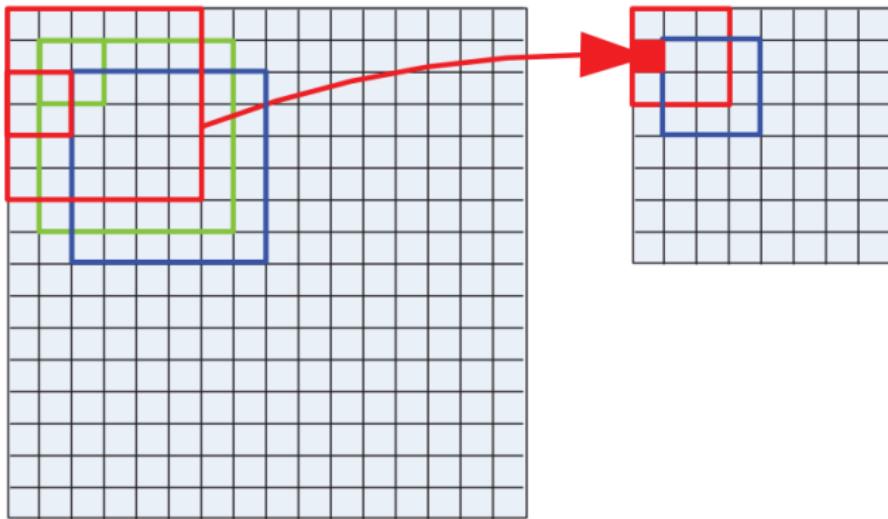
$$\rho_{\text{Sym}}(x) = \left(\sqrt{x^2 + \epsilon^2} \right)^\lambda$$

$$\rho_{\text{Asym}}(x) = \left(\sqrt{\max(0, x)^2 + \epsilon^2} \right)^\lambda$$

Alpaga



Perte Pooling



Source : Region Segmentation Paper

VGG

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Source : VGG Paper

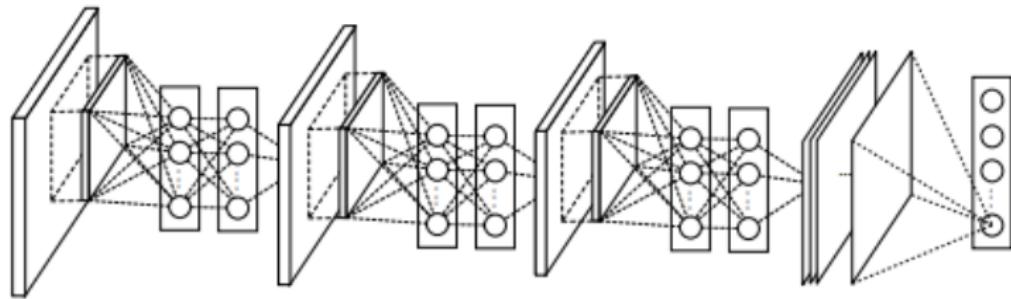
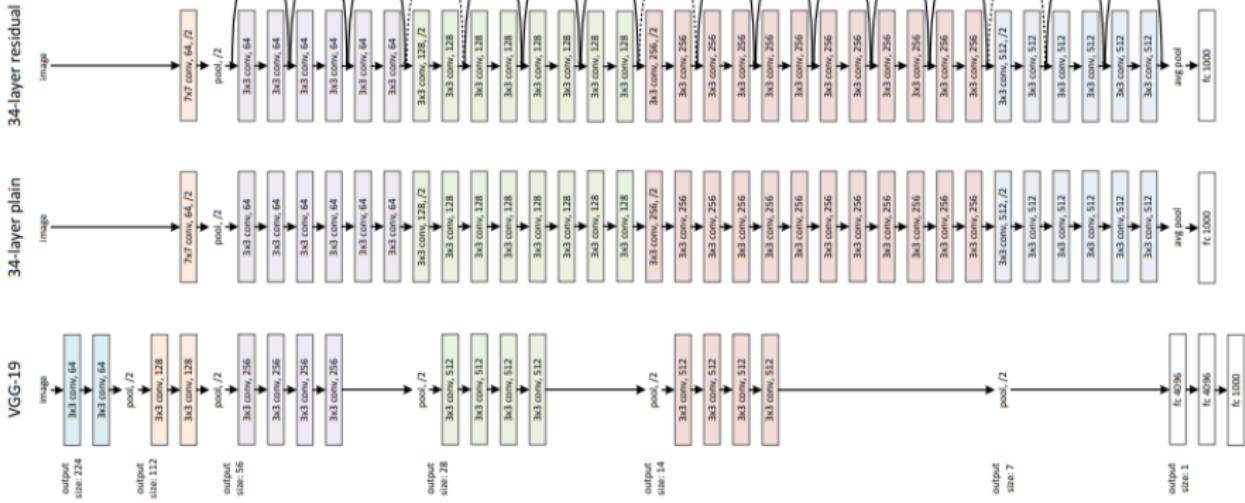


Figure 2: The overall structure of Network In Network. In this paper the NINs include the stacking of three mlpconv layers and one global average pooling layer.

Source : NIN Paper

ResNet



Source : Residual Learning Paper