

Convolutional Neural Networks for Medical Image classification+ VLAD aggregation

Réaliser Par :

Houssem FARHAT

Encadré Par :

Ronan SICRE

Plan

Réseau de neurone convolutionnel(CNN)

Fixed feature extractor

Machine à vecteurs de support (SVM)

Vector of locally aggregated descriptors (VLAD)

Évaluation expérimentale

Discussion et Perspectives

Réseau de neurone convolutionnel (CNN)

Classe d'architectures NN
profondes.

reconnaissance d'objet

a récemment connu un
grand succès.

Possède trois types de
couches:

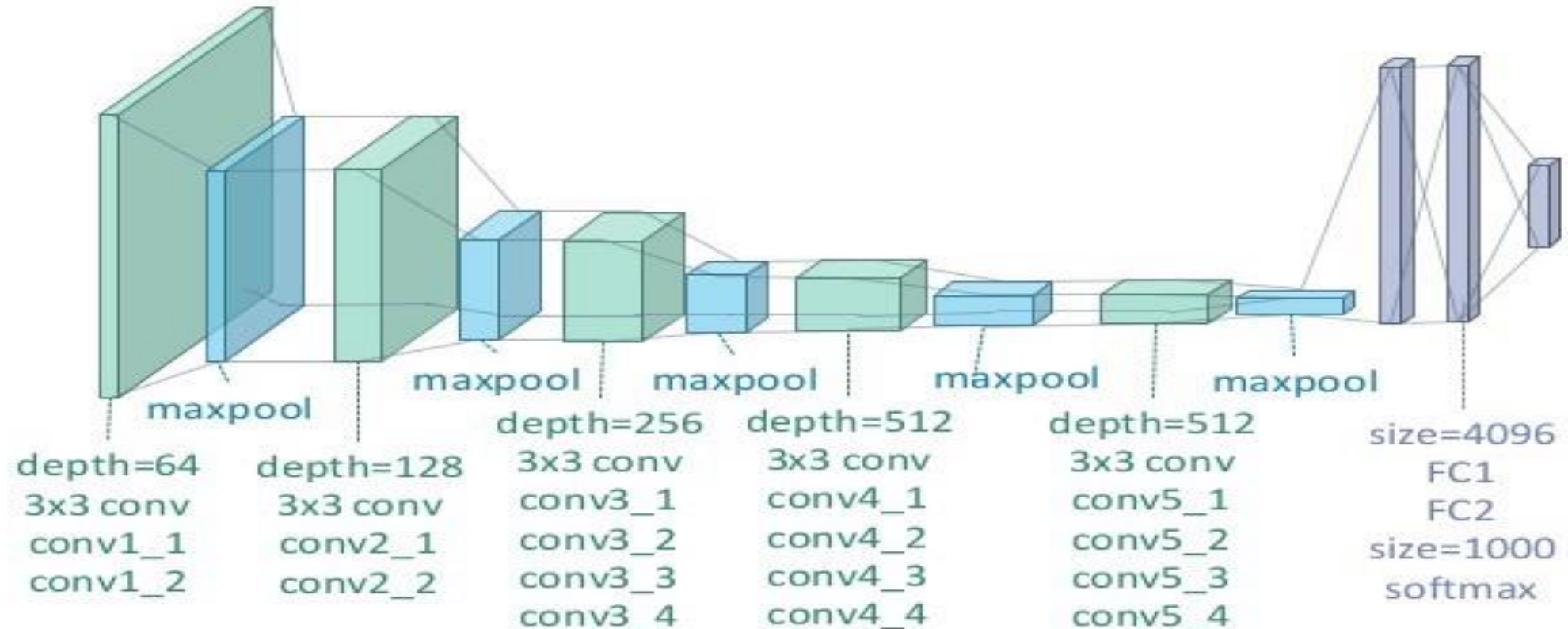
Une couche de convolution
(convolutional layer)

Une couche de
regroupement (pooling
layer)

Une couche entièrement
connectée

Réseau de neurone convolutionnel (CNN)

VGG 19



Réseau de neurone convolutionnel (CNN)

Il faut un jeu de données volumineux pour entraîner le réseau



Beaucoup de temps pour le faire




Beaucoup de ressource matérielle

Fixed feature extractor


On retire la dernière couche de classification du réseau.



On gèle les poids du modèle et on les utilise comme variable d'extraction fixe.



Les codes CNN sont ensuite extraits grâce à la variable fixe pour toutes les images.



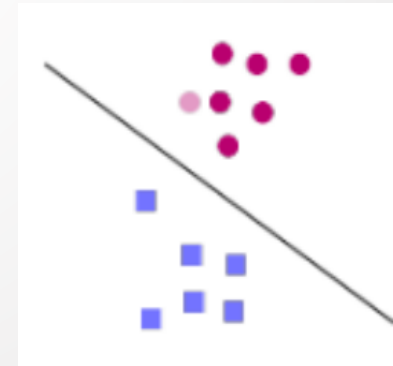
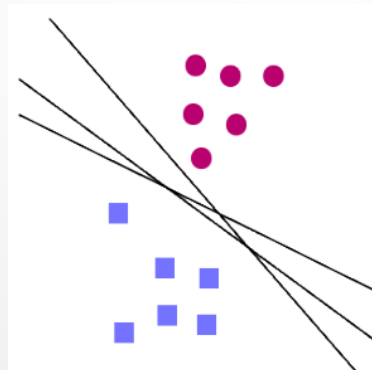
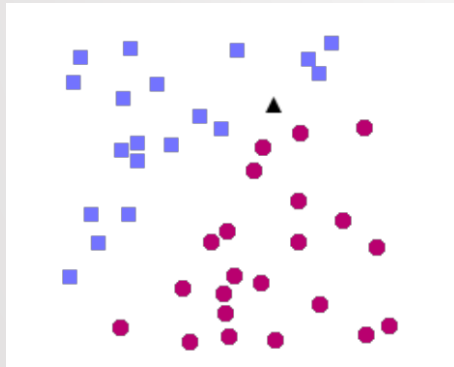
Une classification est effectuée pour le nouvel ensemble de données avec un SVM linéaire par exemple.

Machine à vecteurs de support (SVM)

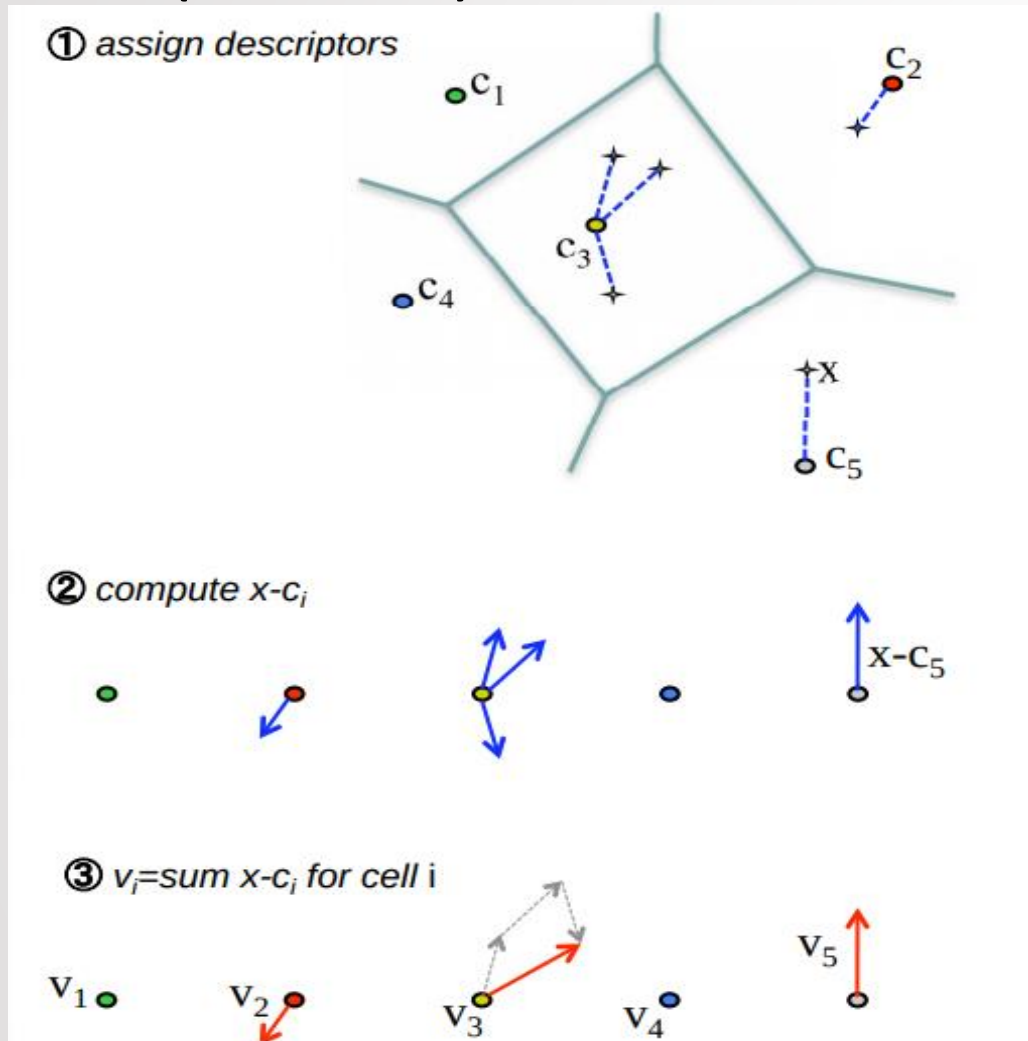
Methode relativement récente (Vapnik et Chervonenkis en 1995)

Initialement construits pour la classification binaire

Rechercher une règle de décision basée sur une separation par hyperplan de marge optimale (optimisation sous-contraintes)



Vector of locally aggregated descriptors (VLAD)



- **Idée : représenter une image par un seul vecteur de taille fixe:**
 - 1 vecteur pour un ensemble de n descripteurs locaux
- **Apprentissage: Kmean :**
 - Sortie : k centroids : $C_1, \dots, C_i, \dots, C_k$
- Pour une image donnée
 1. attribuer chaque descripteur au centre le plus proche
 2. accumuler (somme) des descripteurs par cellule
$$v_i := v_i + (x - c_i)$$
- Sortie: descripteur de taille $k \cdot d$
- Normalisation L2
- typiquement k in $[16..256]$ et $d \geq 2048$

Évaluation expérimentale

- **DATA :**

MINI MIT

- 3 catégories
- Entraînement: 120 images
- Test: 120 images

CHEST-XRAY

- 2 catégories
- Entraînement: 5 232 images
- Test: 624 images

KVASIR

- 8catégories
- Entraînement: 6400images
- Test: 1600 images

Évaluation expérimentale

CNN VGG19

Fixed feature extractor:
couche block5_pool

SVM Linéaire

| block5_pool | scales | 0 | 1 | 2 |
|--------------------|------------|------|------|------|
| | Mini MIT | 0,78 | 0,82 | 0,78 |
| | chest_xray | 0,8 | 0,81 | 0,78 |
| | Kavasir | 0,87 | 0,87 | 0,86 |
| | | | | |

Évaluation expérimentale

| CNN VGG19 | Fixed feature extractor: couche block5_pool | PCA 128 | VLAD | SVM Linéaire |
|-------------|--|------------|------|--------------|
| block5_pool | | K | 64 | 256 |
| | | Mini MIT | 0,82 | 0,84 |
| | | chest_xray | 0,77 | 0,78 |
| | | Kavasir | 0,88 | 0,88 |

Discussion et Perspectives

- Résultat $> \sim 0.8$ (baseline)
- Résultats varient selon le jeu de données
 - Amélioration Kvasir et Mini MIT
 - Dégradation Chest-Xray
- Autres Jeux de données
- Différentes couches
- Taille de descripteurs
- Paramètre K de Kmean
- Autres réseaux de neurones