

Présentation de la lecture d'article

Jules Bataller Beltran Kenewy Diallo Kossi Abotsi
Lorenzo Gaggini

2025-02-10

Sommaire

1. **Contexte** * Enjeux de la Visualisation des Réseaux de Neurones Convolutionnels * Le cadre statistiques

2. Principaux outils

- ▶ Description des principales couches : convolution, ReLU, pooling, et couches entièrement connectées.
- ▶ Présentation du DeconvNet et de ses méthodes (unpooling, rectification inversée, convolution transposée) pour visualiser les activations.

3. Résultats expérimentaux

- ▶ Visualisations comparatives des activations, impact de l'occultation sur les prédictions.

4. Conclusion

Contexte

Enjeux de la Visualisation des Réseaux de Neurones
Convolutionnels

Contexte

Le cadre statistiques

Principaux outils

Fonctionnement des réseaux de neurones convolutionnels (CNN)

- ▶ Les réseaux de neurones convolutionnels (CNN) traitent des données structurées en grilles, comme les images et se compose de plusieurs types de couches :
 - ▶ **Couche de convolution** :
 - ▶ Applique des filtres sur l'image pour extraire des caractéristiques locales (bords, textures).
 - ▶ **Fonction d'activation (ReLU)** :
 - ▶ Permet de modéliser des relations complexes.
 - ▶ **Couche de pooling** :
 - ▶ Rend le réseau plus robuste aux variations (translations, redimensionnements).
 - ▶ **Couches entièrement connectées** :
 - ▶ Les cartes d'activation sont aplaties et traités comme dans un réseau de neurones classique.

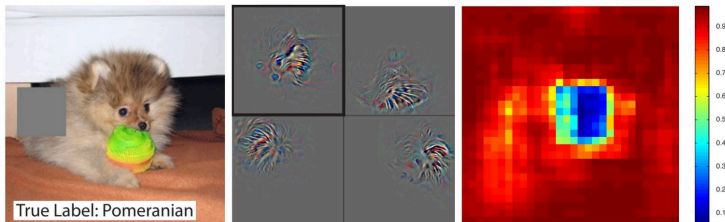
Principaux outils

DeconvNet : Interprétation des CNN

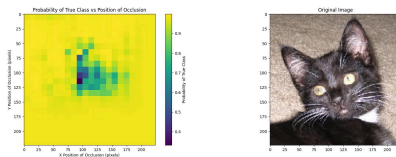
- ▶ Le DeconvNet permet d'interpréter les activations d'un CNN en reconstruisant l'image d'origine à partir des caractéristiques extraites.
 - ▶ **Unpooling (Dé-pooling) :**
 - ▶ Permet de restaurer la structure de l'activation.
 - ▶ **Rectification inversée :**
 - ▶ Applique l'inverse de la fonction d'activation ReLU utilisée dans le CNN.
 - ▶ **Convolution transposée :**
 - ▶ Permet de reconstruire une image en visualisant les caractéristiques identifiées par le CNN.
- ▶ Cette approche offre une meilleure compréhension des motifs détectés à chaque niveau du réseau, en identifiant quelles parties de l'image d'origine ont influencé les activations.

Résultats expérimentaux

Visualisation des activations (Heatmap)



(a) Résultat de l'article original : Image d'entrée, Projection de la carte de caractéristiques la plus forte, Probabilité de la classe correcte.



(b) Résultat de la simulation reproduite

Figure 1: Impact de l'occultation sur les activations des caractéristiques et les prédictions de classe

Résultats expérimentaux

Effet de la rotation sur les performances du modèle

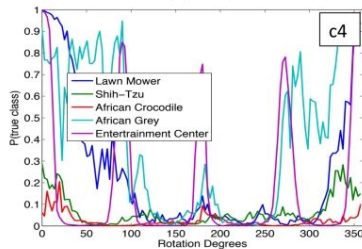
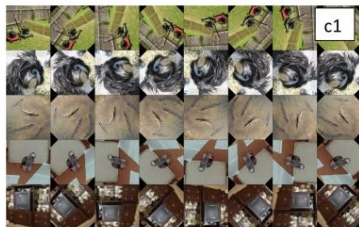


Figure 2: Effet de la rotation sur la classification

Résultats expérimentaux

Exemples de visualisations (1/2)

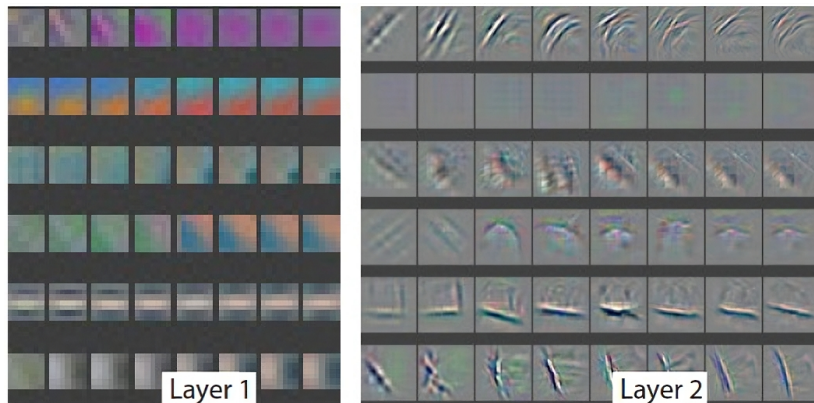


Figure 3: Figure de DeconvNet

Résultats expérimentaux

Exemples de visualisations (2/2)

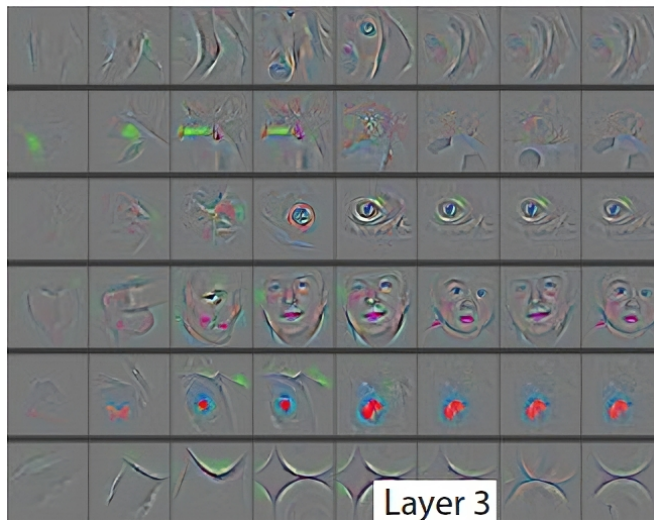
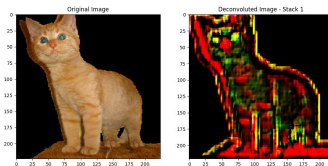
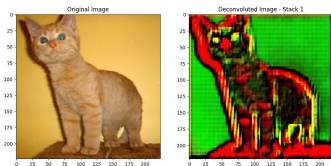
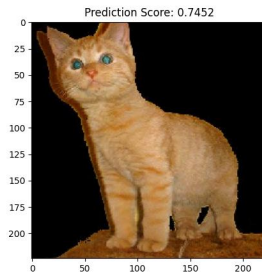
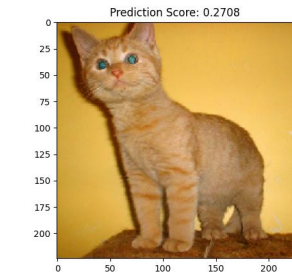


Figure 4: Figure de DeconvNet

Résultats expérimentaux

Simulation deconvolution (1/2)



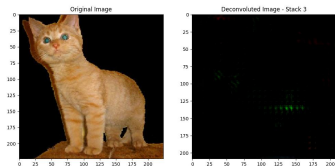
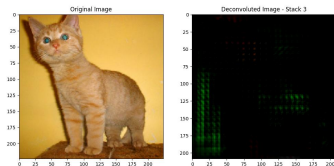
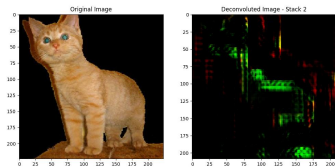
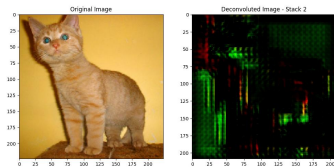
(a) Avec fond de couleur

(b) Sans fond de couleur

Figure 5: Predictions initiales et couche 1

Résultats expérimentaux

Simulation deconvolution (2/2)



(a) Avec fond de couleur

(b) Sans fond de couleur

Figure 6: Couches 2 et 3

Conclusion