



Projet Image

Débruitage d'images

Compte-rendu n°6

Adam Bonbon

Alaric Hunot-Martin

Louis Jean

Master 2 IMAGINE
Université de Montpellier

24 novembre 2024

Table des matières

1	Introduction	2
2	Exploration des réseaux de neurones	2
3	Entraînement de CGNet	2
3.1	Progrès et défis	2
3.2	Améliorations de notre modèle	2
4	Analyse préliminaire avec NAFNet	3
4.1	Méthodologie	3
4.2	Résultats	4
4.3	Interprétation	4
5	Création d'un nouveau modèle pour de la vidéo	4
6	Amélioration de l'interface graphique	5
7	Conclusion	6
8	Perspectives et suggestions d'amélioration	6

1 Introduction

2 Exploration des réseaux de neurones

Cette semaine, nous avons continué notre travail sur le débruitage d'images, avec un focus sur deux modèles : **CGNet (Cascaded Gaze Network)** et **NAFNet (Nonlinear Activation Free Network for Image Restoration)**. Le but est d'explorer leurs performances dans des contextes spécifiques et de comparer leurs résultats.

3 Entraînement de CGNet

CGNet, un modèle robuste mais lourd, ne propose pas de versions pré-entraînées sur le dépôt des auteurs, en raison de sa taille conséquente. Nous avons donc opté pour une première approche consistant à réentraîner le modèle sur la base de données **SIDD (Smartphone Image Denoising Dataset)**, comme indiqué dans l'article original.

3.1 Progrès et défis

Nous avons du réduire la taille du batch pour que l'entraînement soit faisable sur les machines de l'université. L'entraînement a débuté le mercredi et se poursuivra sur les machines de la faculté, offrant une meilleure performance que nos ordinateurs personnels. Le programme d'entraînement sauvegarde régulièrement les modèles, permettant une reprise sans perte de progrès. L'objectif sera de reprendre l'entraînement lundi pour qu'on puisse tester ce modèle avec nos images.

3.2 Améliorations de notre modèle

Nous avons poursuivi notre démarche d'amélioration en cherchant à rapprocher davantage notre modèle CNN de l'architecture de CGNet.

L'objectif était de conserver les forces de CGNet, notamment sa légèreté et son efficacité contextuelle, tout en augmentant ses capacités pour mieux répondre à des tâches complexes comme le débruitage vidéo.

Pour cela, nous avons apporté plusieurs modifications.

Dans les blocs contextuels, nous avons ajouté une couche de convolution supplémentaire pour renforcer l'extraction des caractéristiques locales et globales.

Nous avons également introduit des convolutions factorisées qui permettent de mieux capturer les relations spatiales horizontales et verticales, tout en limitant l'augmentation du nombre de paramètres.

Ces changements s'inscrivent dans la continuité des principes de CGNet, qui vise à maximiser l'efficacité sans alourdir l'architecture.

Nous avons aussi décidé d'intégrer des mécanismes d'attention adaptative.

Ces blocs d'attention aident le modèle à se concentrer sur les parties les plus importantes de l'image, tout en atténuant l'impact des zones moins pertinentes.

De plus, nous avons ajouté un bloc résiduel avec attention, combinant des connexions résiduelles pour préserver les informations essentielles et des mécanismes d'attention pour capturer des rela-

tions plus subtiles.

Ces améliorations permettent au modèle d'être plus robuste face aux variations importantes comme celles induites par le bruit.

Malgré toutes ces améliorations, nous constatons que les performances du modèle stagnent.

Cela s'explique par le fait que nous sommes repartis de zéro pour entraîner le modèle avec les nouvelles modifications.



Figure 1: Résultats du npuveau modèle

4 Analyse préliminaire avec NAFNet

En attendant les résultats complets de CGNet, nous avons exploré les performances de **NAFNet**, un modèle introduit par Chen *et al.*. Deux variantes ont été testées :

- NAFNet entraîné sur la base SIDD.
- NAFNet entraîné sur la base GoPro.

4.1 Méthodologie

Nous avons utilisé une petite base d'images en noir et blanc contenant :

- Une image originale d'un avion.
- La même image avec un bruit de type sel et poivre.
- La même image avec un bruit gaussien.

Les images débruitées par chaque modèle ont ensuite été comparées en utilisant le score NIMA.

4.2 Résultats

Image	Score NIMA
Image originale	4.953
Image bruitée (sel et poivre)	4.756
Image bruitée (gaussien)	4.565
Image débruitée (NAFNet GoPro, gaussien)	5.155
Image débruitée (NAFNet SIDD, gaussien)	4.208
Image débruitée (NAFNet GoPro, sel et poivre)	4.695
Image débruitée (NAFNet SIDD, sel et poivre)	4.624

Table 1: Comparaison des scores NIMA pour différentes images.

4.3 Interprétation

Le fait que l'image débruitée gaussienne avec le modèle GoPro ait un meilleur score que l'image originale peut indiquer que le débruitage améliore certains aspects visuels évalués par NIMA, tels que le contraste ou la netteté. En revanche, les résultats relativement faibles pour les images débruitées sel et poivre par rapport à l'image bruitée originale ainsi que l'image débruitée gaussienne avec la base SIDD, pourraient suggérer une spécialisation du modèle pour des types de bruit différents ou pour des images différentes.

5 Crédit d'un nouveau modèle pour de la vidéo

Nous avons cherché à adapter notre modèle de débruitage d'images pour l'appliquer au débruitage de vidéos, en utilisant le dataset DAVIS 2017, connu pour ses vidéos de haute qualité.

Comme pour le dataset BSD500 dans le contexte des images, le bruit a été ajouté manuellement aux vidéos en combinant un bruit "sel et poivre" et un bruit gaussien généré aléatoirement.

L'objectif était de tester si un modèle conçu pour les images pouvait être efficace sur des séquences vidéo bruitées en traitant les frames une à une.

Les résultats montrent une réduction notable du bruit dans les frames, mais également des artefacts qui doivent être améliorés pour obtenir une qualité visuelle satisfaisante.

Ce travail ouvre des perspectives intéressantes pour approfondir les techniques de débruitage vidéo, notamment en intégrant des méthodes capables d'exploiter les relations temporelles entre les frames pour des résultats plus cohérents et réalistes.

Pour faciliter l'évaluation de nos résultats, nous avons mis à disposition un lien Google Drive contenant des exemples de vidéos bruitées et débruitées.

Malheureusement, certaines frames sont un peu désynchronisées dans la vidéo.



Figure 2: Une frame de la vidéo bruitée

6 Amélioration de l'interface graphique

Nous avons également rapidement amélioré l'interface graphique, en ajoutant notamment la méthode par filtre moyenneur, en permettant le contrôle de la taille de la fenêtre glissante et en affichant des métriques.

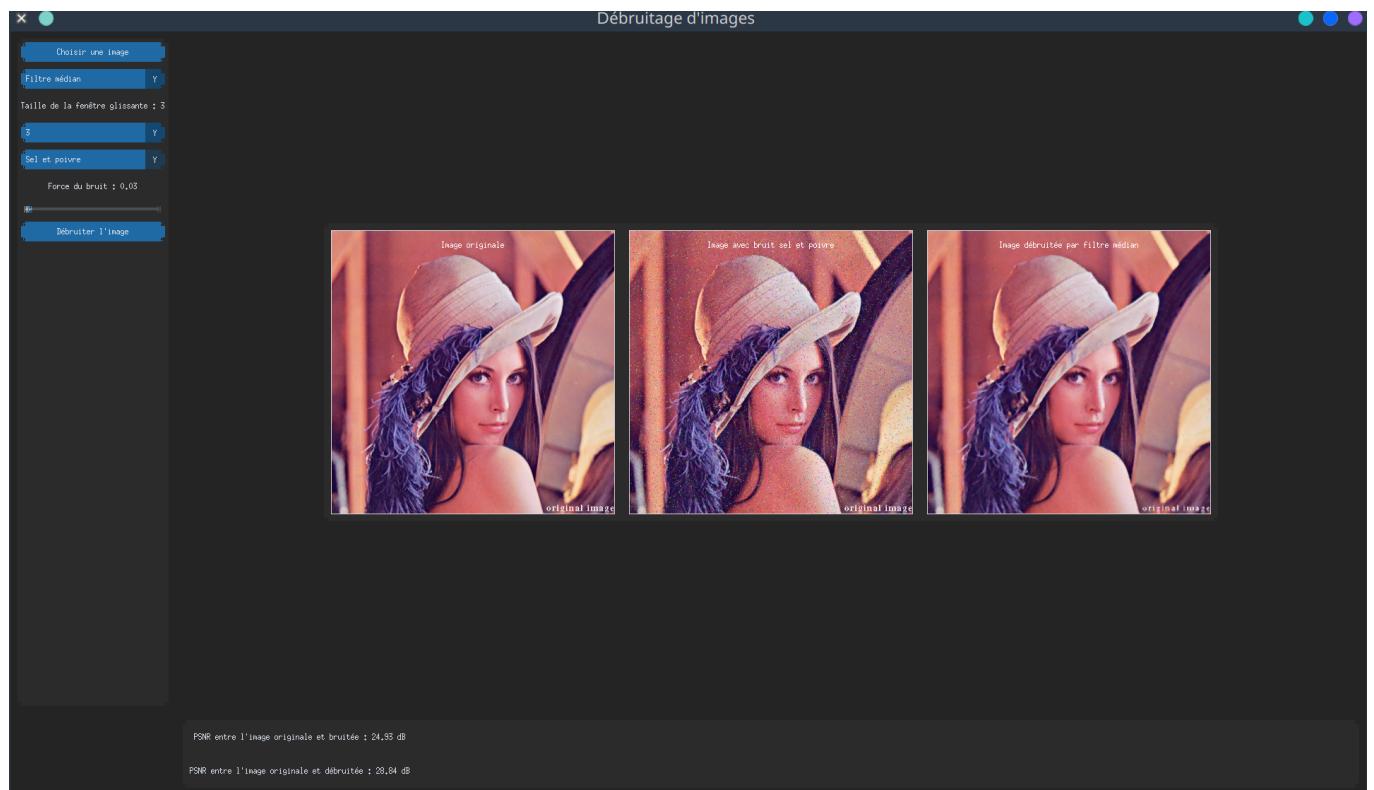


Figure 3: Interface graphique

7 Conclusion

Cette semaine, nous avons réalisé des progrès dans notre exploration des modèles de débruitage d'images. Le réentraînement de CGNet sur la base SIDD, bien que contraint par les limitations matérielles, a permis de poser les bases pour une évaluation approfondie des performances du modèle. L'utilisation de NAFNet nous a fourni des résultats prometteurs.

Les premières tentatives d'adaptation de notre modèle pour le débruitage vidéo ont également mis en lumière des pistes d'amélioration, notamment en ce qui concerne la gestion des artefacts et des relations temporelles entre les frames. Enfin, l'amélioration de l'interface graphique offre une meilleure accessibilité et visualisation des résultats, rendant notre travail plus fonctionnel et interactif.

8 Perspectives et suggestions d'amélioration

Voici quelques pistes et suggestions que nous pourrons faire les semaines à venir :

- **Développement d'un programme d'analyse automatique** : Il serait plus rentable de créer un programme permettant de lire les modèles générés et d'évaluer automatiquement leurs performances en utilisant le score NIMA. Ce programme permettra d'organiser et de visualiser les résultats de manière plus propre et systématique, facilitant ainsi l'interprétation tant quantitative que qualitative.
- **Intégration d'autres métriques** : En complément du score NIMA, nous inclurons les métriques PSNR (Peak Signal-to-Noise Ratio) et SSIM (Structural Similarity Index Measure) pour obtenir une évaluation plus complète des performances de débruitage. Ces métriques permettront d'évaluer plus précisément la fidélité des images débruitées par rapport aux originales et pourront aussi aider pour l'application que nous sommes en train de mettre en place.
- **Amélioration du pipeline d'entraînement** : Nous poursuivrons l'entraînement du modèle CGNet en explorant des hyperparamètres optimaux pour réduire les temps de calcul et améliorer les résultats sur nos jeux de données spécifiques.

Ces efforts devraient non seulement renforcer la robustesse de notre évaluation mais également guider le choix des modèles les plus adaptés à nos objectifs de débruitage.

Merci pour le temps et l'attention que vous avez consacrés à la lecture de ce compte-rendu.