

# Note Méthodologique : Preuve de Concept

## Dataset Retenu

Le dataset utilisé est le **Stanford Dogs Dataset**, qui est une collection de 20 580 images représentant 120 races de chiens. Ce dataset a été conçu pour des tâches de classification d'images et de détection d'objets, mais il est également adapté pour les tâches de super-résolution. Les images sont fournies en haute résolution et sont utilisées pour créer des versions à basse résolution pour l'entraînement et la validation du modèle SRGAN. La préparation du dataset inclut le redimensionnement des images à basse résolution (LR) à l'aide d'une interpolation bicubique, ce qui est essentiel pour l'entraînement de modèles de super-résolution.

## Concepts de l'Algorithme Récent

Le modèle employé est le **SRGAN** (Super-Resolution Generative Adversarial Network), qui utilise un réseau générateur et un réseau discriminateur pour améliorer la qualité des images à basse résolution (LR). Voici les concepts clés :

- **Générateur** : Crée des images haute résolution à partir d'images basse résolution. Il est pré-entraîné avec une fonction de perte MSE (Mean Squared Error) pour établir une base initiale de qualité des images.

- **Discriminateur** : Évalue les images générées pour déterminer si elles ressemblent à des images réelles. Il aide à affiner le générateur en fournissant un retour sur la qualité perceptuelle des images générées.
- **Perceptual Loss** : Mesure la différence perceptuelle entre les images générées et les images réelles en utilisant des caractéristiques extraites par des réseaux de neurones pré-entraînés (par exemple, VGG). Cette perte capture des détails visuels importants que la MSE ne peut pas détecter.
- **Discriminator Loss** : Évalue la performance du discriminateur pour différencier les images générées des images réelles. Une faible valeur de perte indique que le générateur est efficace pour produire des images de haute qualité qui trompent le discriminateur.

## La Modélisation

La méthodologie de modélisation est structurée comme suit :

### 1. **Pré-entraînement du Générateur** :

- Le générateur est initialement formé en utilisant la MSE Loss. Ce pré-entraînement aide à fournir une base de génération d'images avec une certaine qualité.

- Le modèle est entraîné pendant 20 000 étapes avec les images d'entraînement et de validation, et les poids sont sauvegardés.

## 2. **Entraînement Complet avec SRGAN :**

- Le générateur pré-entraîné est affiné en utilisant les métriques **Perceptual Loss** et **Discriminator Loss**. Cela permet d'améliorer la qualité visuelle des images générées.
- Le modèle est entraîné pendant 5 000 étapes pour ajuster les poids du générateur et du discriminateur.
- Les poids du générateur et du discriminateur sont sauvegardés après l'entraînement.

## **Méthodologie d'Évaluation :**

- **Perceptual Loss** : Utilisée pour évaluer la qualité visuelle des images générées en capturant des caractéristiques perceptuelles importantes.
- **Discriminator Loss** : Mesure la capacité du discriminateur à distinguer les images générées des images réelles, fournissant un indicateur de la qualité des images générées.

## Synthèse des Résultats

Les résultats montrent que l'entraînement complet avec SRGAN améliore la qualité des images générées par rapport au pré-entraînement avec MSE Loss. Les valeurs de **Perceptual Loss** et de **Discriminator Loss** fournissent des indicateurs clés de la performance du modèle :

- Les valeurs de **Perceptual Loss** plus faibles indiquent une meilleure qualité visuelle des images générées.
- Les valeurs de **Discriminator Loss** plus faibles indiquent que les images générées sont plus difficiles à distinguer des images réelles par le discriminateur.

Une comparaison des résultats avec les techniques précédentes (comme l'utilisation de MSE Loss uniquement) montrera les avantages en termes de qualité perceptuelle et de réalisme des images générées par le SRGAN.

## Analyse de la Feature Importance Globale et Locale du Nouveau Modèle

- **Feature Importance Globale** : Les couches du générateur qui capturent des détails fins, telles que les couches convolutives profondes, jouent un rôle crucial dans la génération de haute qualité. Les couches du discriminateur

qui apprennent à reconnaître des textures et des structures subtiles sont également importantes.

- **Feature Importance Locale** : L'analyse locale peut inclure la visualisation des cartes de caractéristiques pour comprendre quelles parties des images sont améliorées par le générateur. Les cartes de caractéristiques du discriminateur montrent comment les différentes régions de l'image influencent la décision du modèle.

## Limites et Améliorations Possibles

- **Limites :**

- Les modèles SRGAN peuvent présenter des artefacts de génération, tels que des effets de bord ou des artefacts de sur-échantillonnage.
- L'entraînement des GAN est souvent instable et peut nécessiter des ajustements fins des hyperparamètres.

- **Améliorations :**

- **Régularisation** : Intégrer des techniques de régularisation supplémentaires pour stabiliser l'entraînement, comme la normalisation de lot ou la perte de régularisation.

- **Architectures Améliorées** : Explorer de nouvelles architectures de réseaux, comme les réseaux résiduels ou les mécanismes d'attention, pour réduire les artefacts et améliorer la qualité.
- **Évaluation Complète** : Ajouter des métriques supplémentaires telles que PSNR (Peak Signal-to-Noise Ratio) et SSIM (Structural Similarity Index) pour obtenir une évaluation plus complète de la qualité des images générées.