

Projet deep learning: CR 3

Hugo Vaillant, Nathan Souvignet

Novembre 2025

Implementation MAE

On a implémenté le code du MAE trouvé sur ce github:

<https://github.com/IcarusWizard/MAE>

On avait voulu utiliser un autre générateur de masques (qui génère une image bruitée plutôt que des gros carrés) mais on a oublié de le changer au moment de l'entraînement.

L'un des objectifs était d'expérimenter en partant d'une version déjà entraînée pour continuer l'entraînement avec d'autres masques.

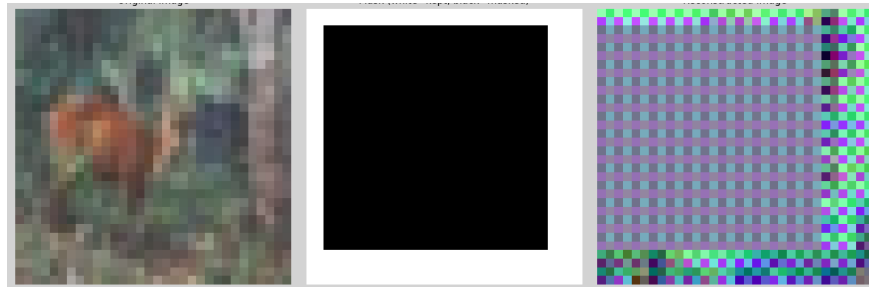


Figure 1: Premier test avec 100 epochs.

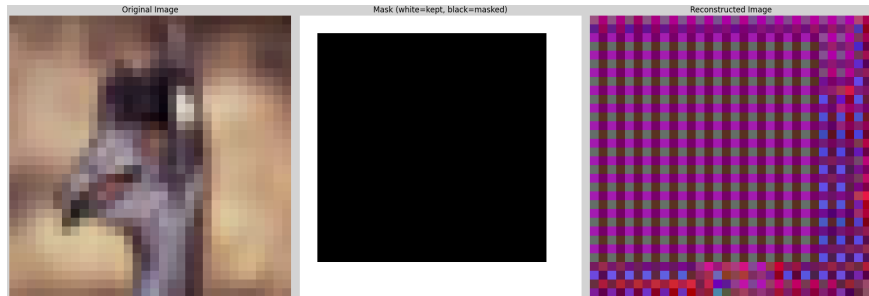


Figure 2: Après 200 epochs, les résultats semblent de moins bonne qualité.

Le MAE représente un certain intérêt car il s'entraîne explicitement sur les zones non masquées et ne se soucie pas du masque.

Nouvelle Méthode

Mais après avoir travaillé avec Pytorch et analysé la méthode MAE, qui ne devait être qu'une partie de la méthode ASUKA. On a remarqué que la méthode est trop avancée pour qu'on l'utilise. De plus, la méthode n'est pas créer de façon à utiliser un GAN.

C'est pour ça que nous avons décidé de partir sur une nouvelle méthode PEN-Net (Pyramid-context ENcoder Network décrits dans l'article [1]) qui encode l'image avec un U-Net, remplit les trous dans les features de façon pyramidale grâce aux ATN (Attention Transfer Network), puis reconstruit l'image via un décodeur et un discriminant adversarial.

References

- [1] Yanhong Zeng, Jianlong Fu, Hongyang Chao, Baining Guo (2019). *Learning Pyramid-Context Encoder Network for High-Quality Image Inpainting* depuis <https://arxiv.org/pdf/1904.07475>