

Safe-Eye

Évaluation de la sécurité visuelle d'images obscures par CNN

FLORENTIN DENIS
KHÉLIAN LARVET



Plan

1

Les méthodes d'obscurcissement

2

Mesurer la sécurité visuelle

3

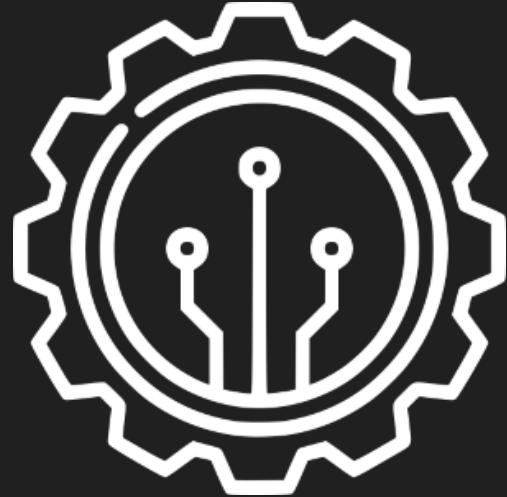
Reconnaisances et Attaques

4

Interfaçage Tkinter

Introduction

Pourquoi évaluer la sécurité visuelle ?



L'évolution des technologies facilite la création de contenu.

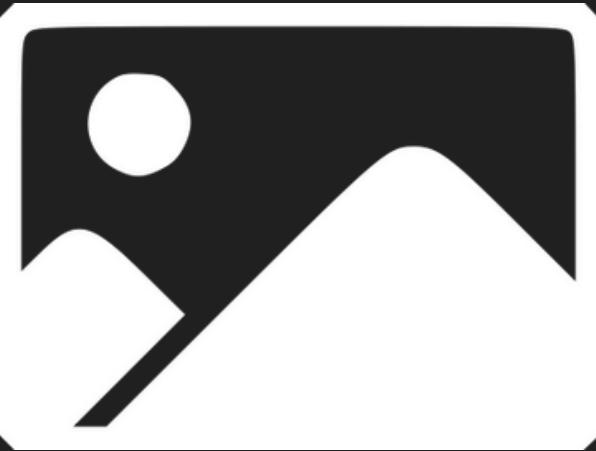
Stockage sur des bases de données tierces non fiables.

Confidentialité compromise et risque d'atteinte à la vie privée.

Introduction

Les processus composant notre application

Image



Informations



Obscurcissement



Reconnaissance



Évaluation



Introduction

Technologies utilisées



Les méthodes d'obscurcissement

Capacité à modifier voire supprimer des éléments sensibles d'une image tout en conservant certaines caractéristiques visuelles permettant son traitement.

Masquage & Bruitage

Filtres & Convolutions
(Pixellisation, Floutage)

Déformations & Distorsions

"Inpainting"

Chiffrement
(P3, Mélange, JPEG)

Transformations esthétiques
(Transfert de style, Cartoons)

Anonymisation d'image
("Face Morphing")

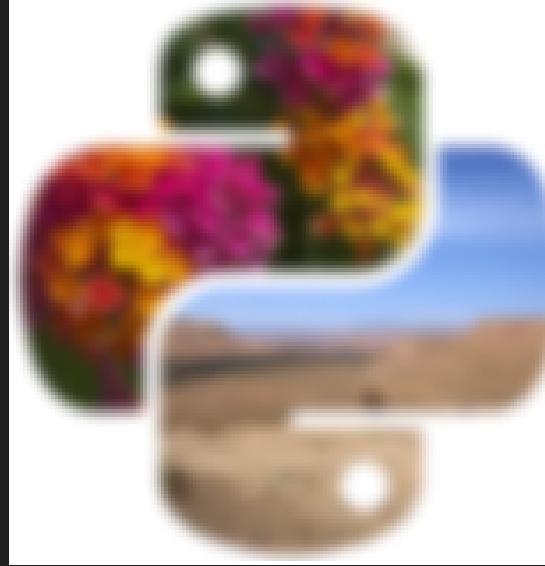
Abstraction
(Silhouette, Avatar)

Les méthodes d'obscurcissement

Pixélisation



Floutage



Masquage



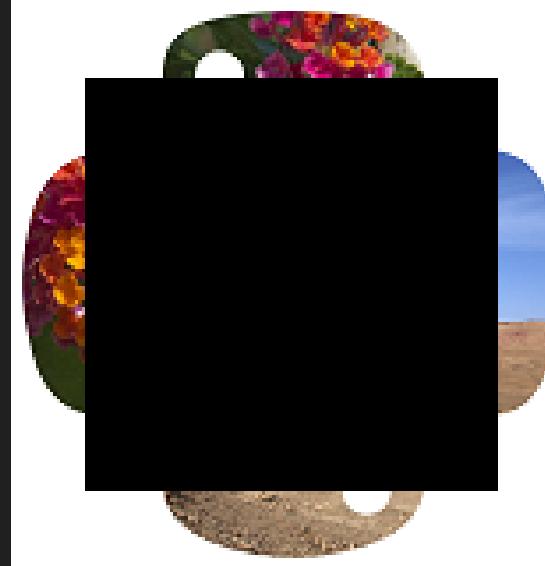
Mélange



Bruitage



Suppression

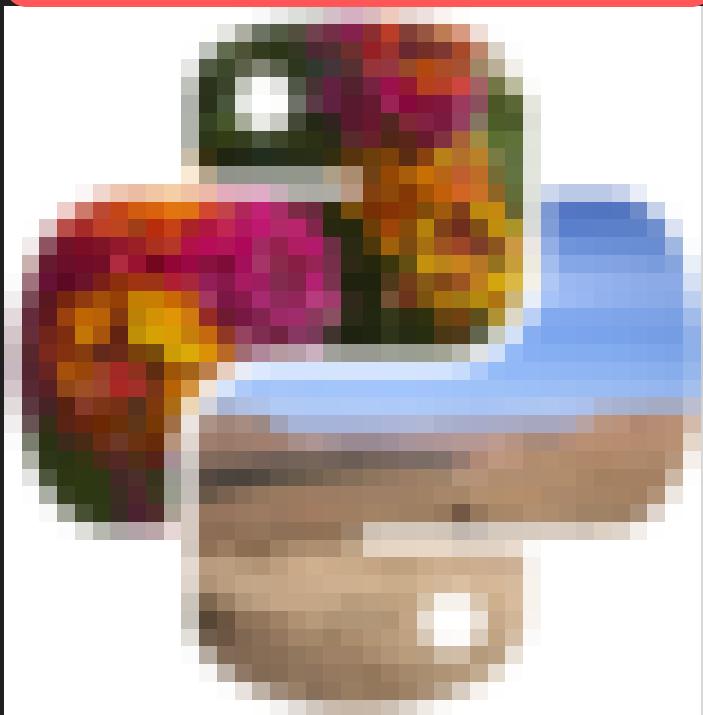


Les méthodes d'obscurcissement

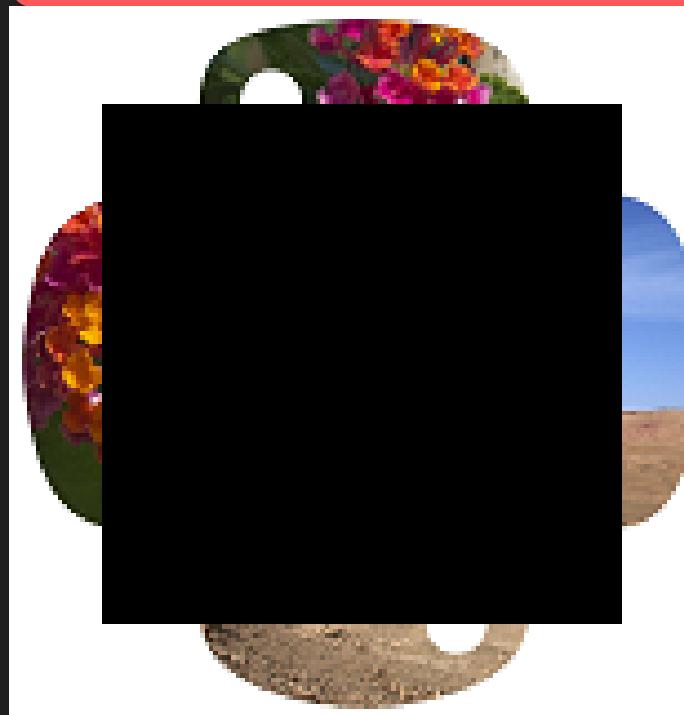
Compromis "intimité-utilité" : Capacité à divulguer des informations traitables sans en révéler le contenu.

La naturalité : Capacité à dissimuler le fait que l'image ait été obscurcie.

Pixélisation



Suppression



[1] - Fig. 9. Face Morphing*

Mesurer la sécurité visuelle

PSNR

(Peak Signal to Noise Ratio)

Rapport entre la puissance maximale d'un signal et la puissance du bruit.

SSIM

(Structural Similarity Index Measure)

Similitudes locales selon trois composantes : Luminance, Contraste et Structure.

HaarPSI

(Haar Perceptual Similarity Index)

Similitudes locales définies par les amplitudes des coefficients d'ondelettes de Haar à haute fréquence.

Mesurer la sécurité visuelle

Plusieurs types de contenu sensible

Biométrique

Biométrique léger

Non biométrique

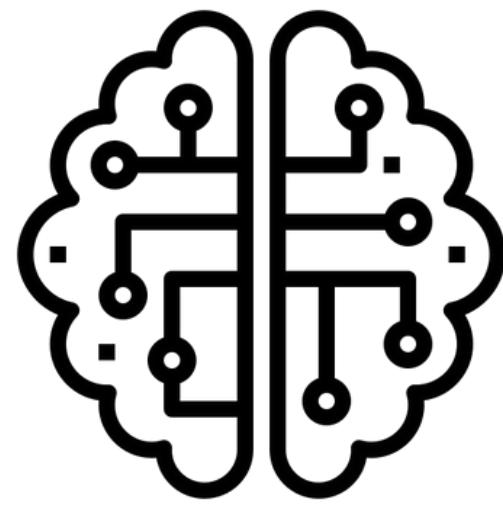
Confidentiel

Censuré

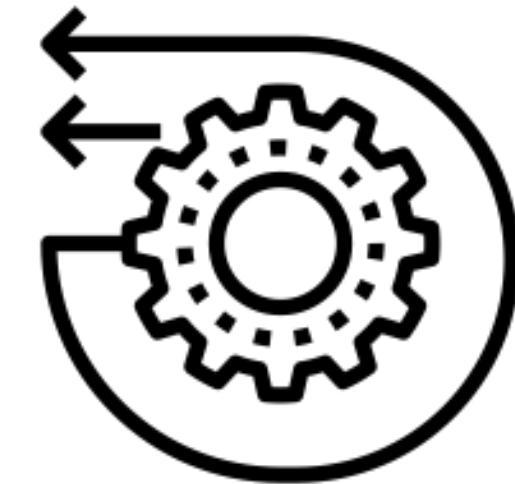
Plusieurs mesures de sécurité



**Reconnaissance
par un utilisateur**



**Reconnaissance
par un classifieur**



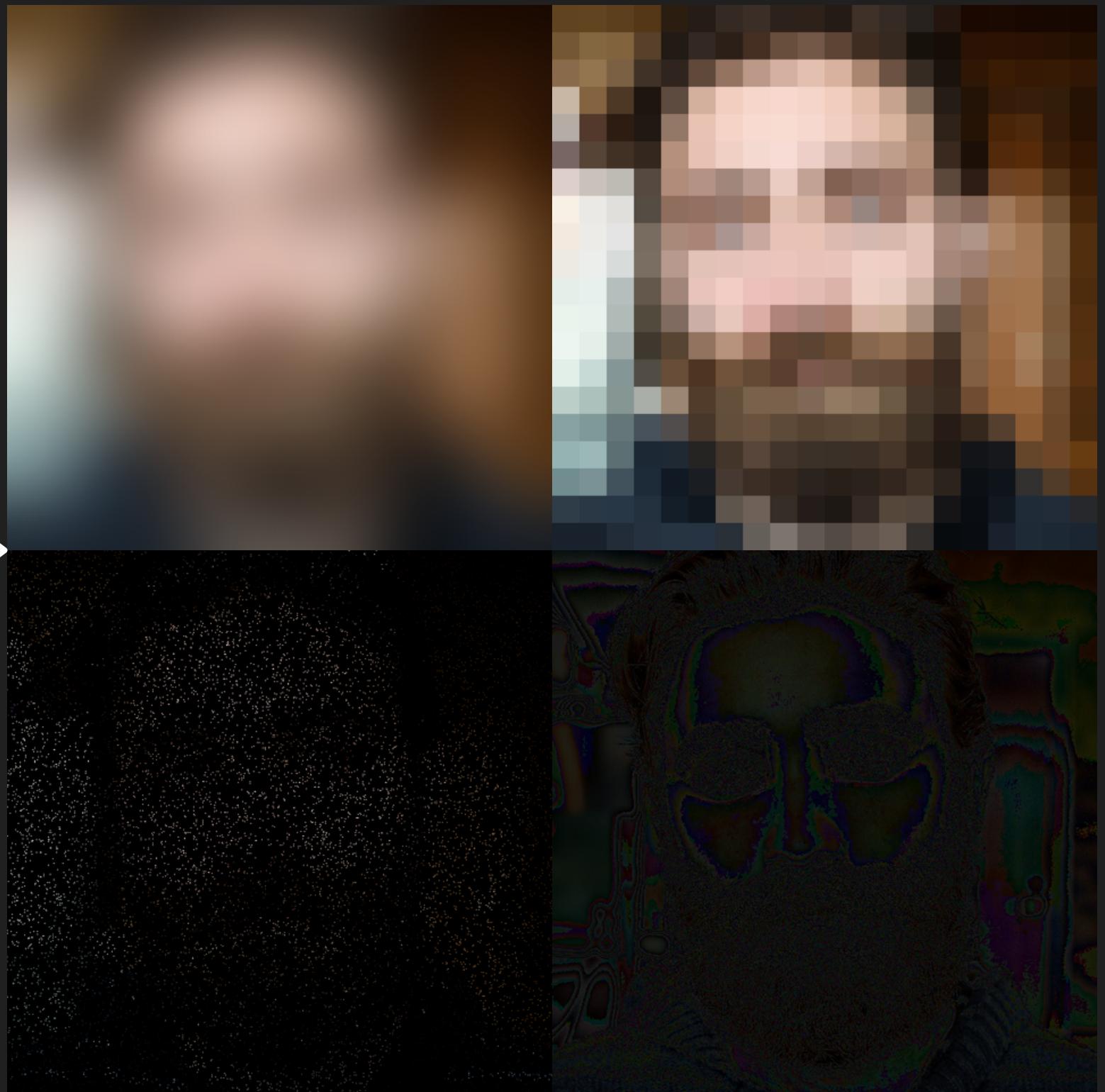
**Reconnaissance
après une attaque**

Mesurer la sécurité visuelle



HaarPSI = 0.25

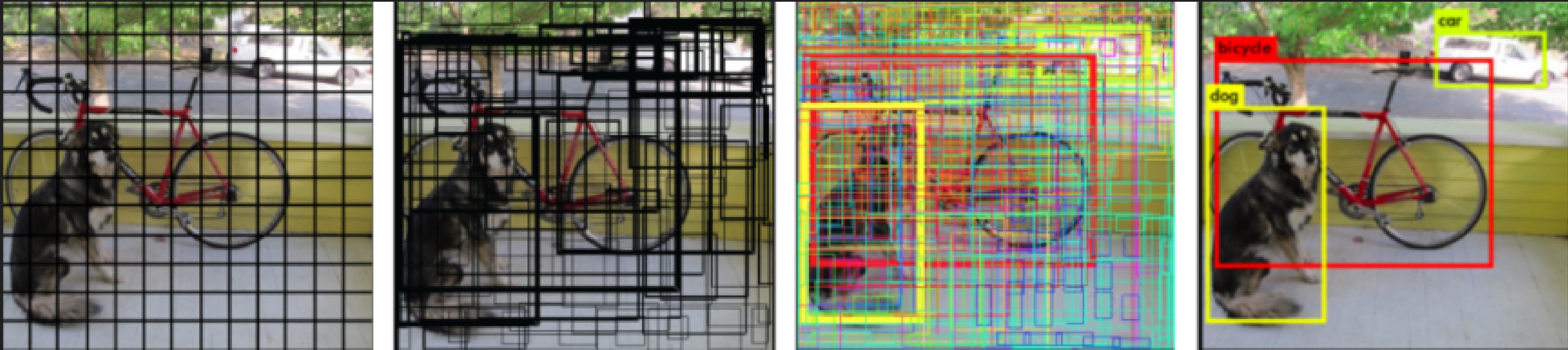
HaarPSI = 1.0



Reconnaisances et Attaques

YOLOv3 : You Only Look Once (Version 3)

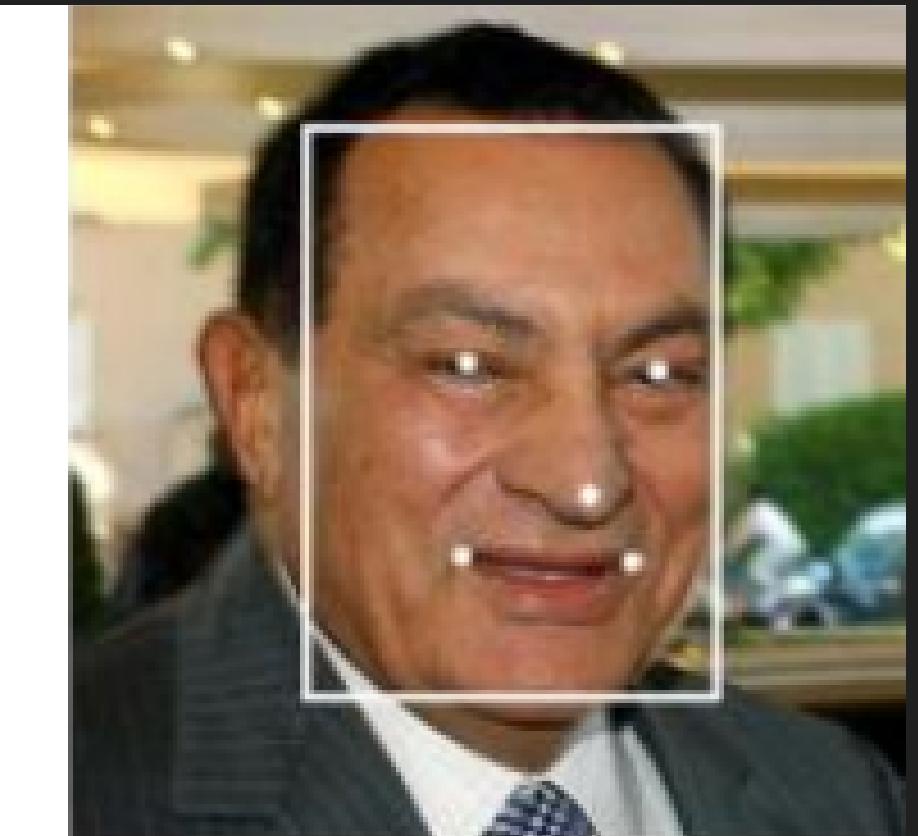
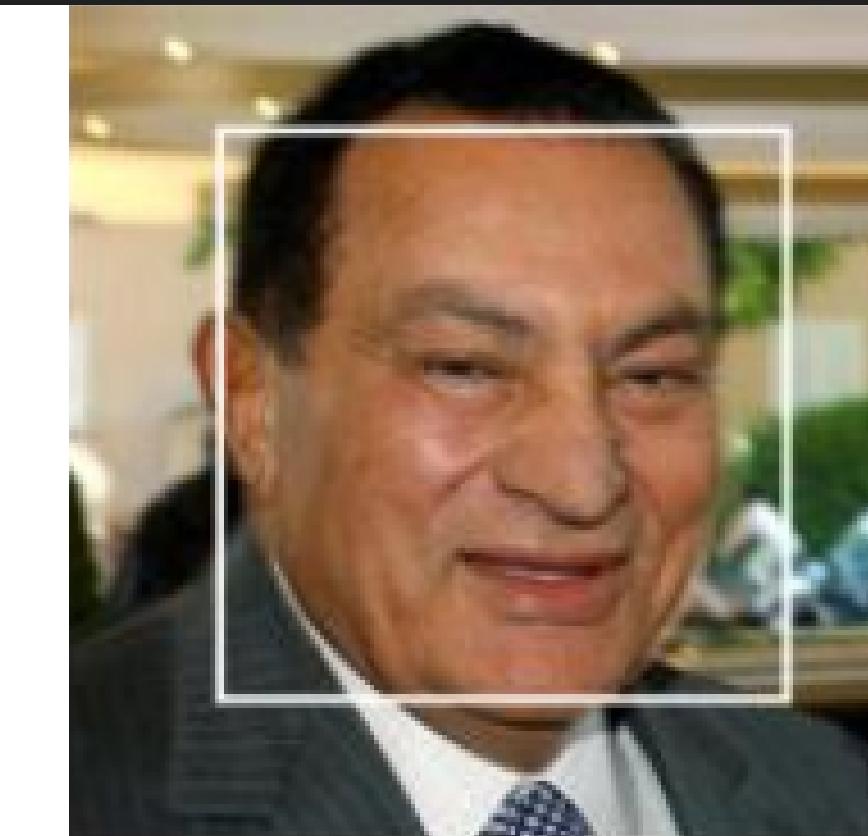
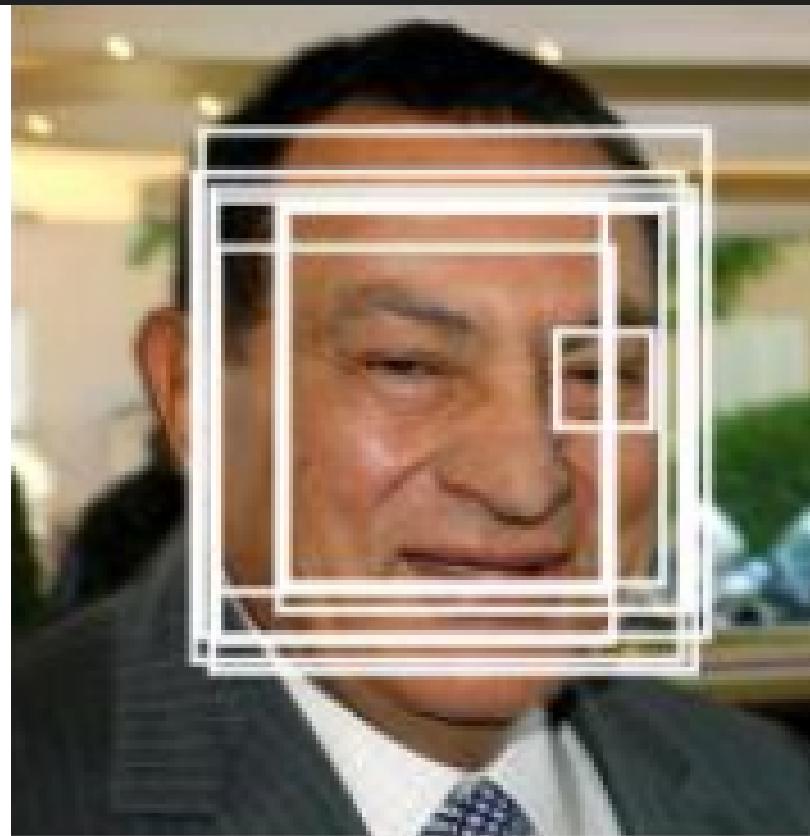
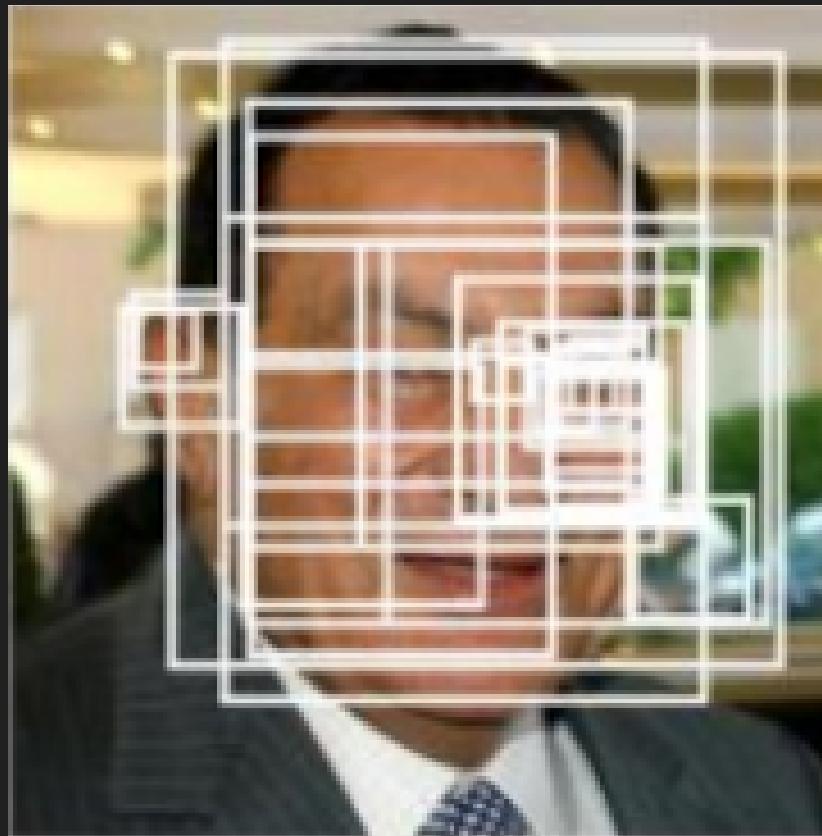
Permet de créer des boîtes englobantes rapidement.



Reconnaisances et Attaques

MTCNN : Multi-task Cascaded Convolutional Networks

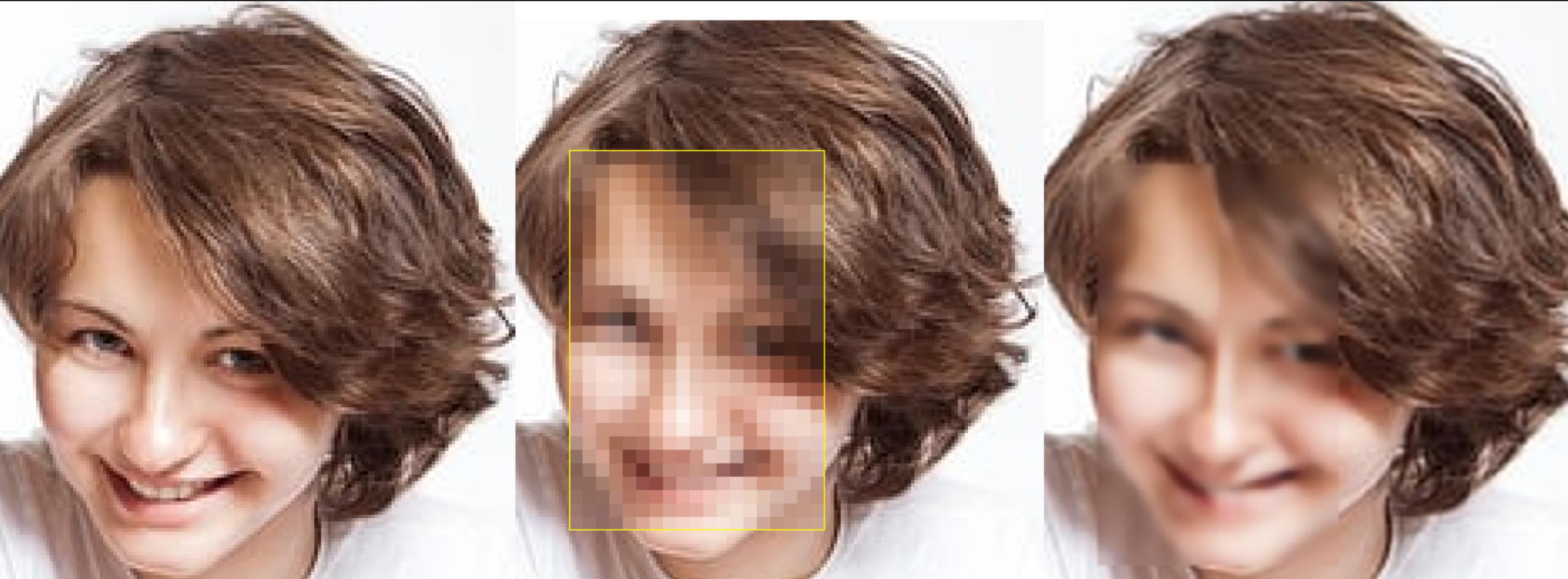
Permet une détection plus avancée sur les visages humains.



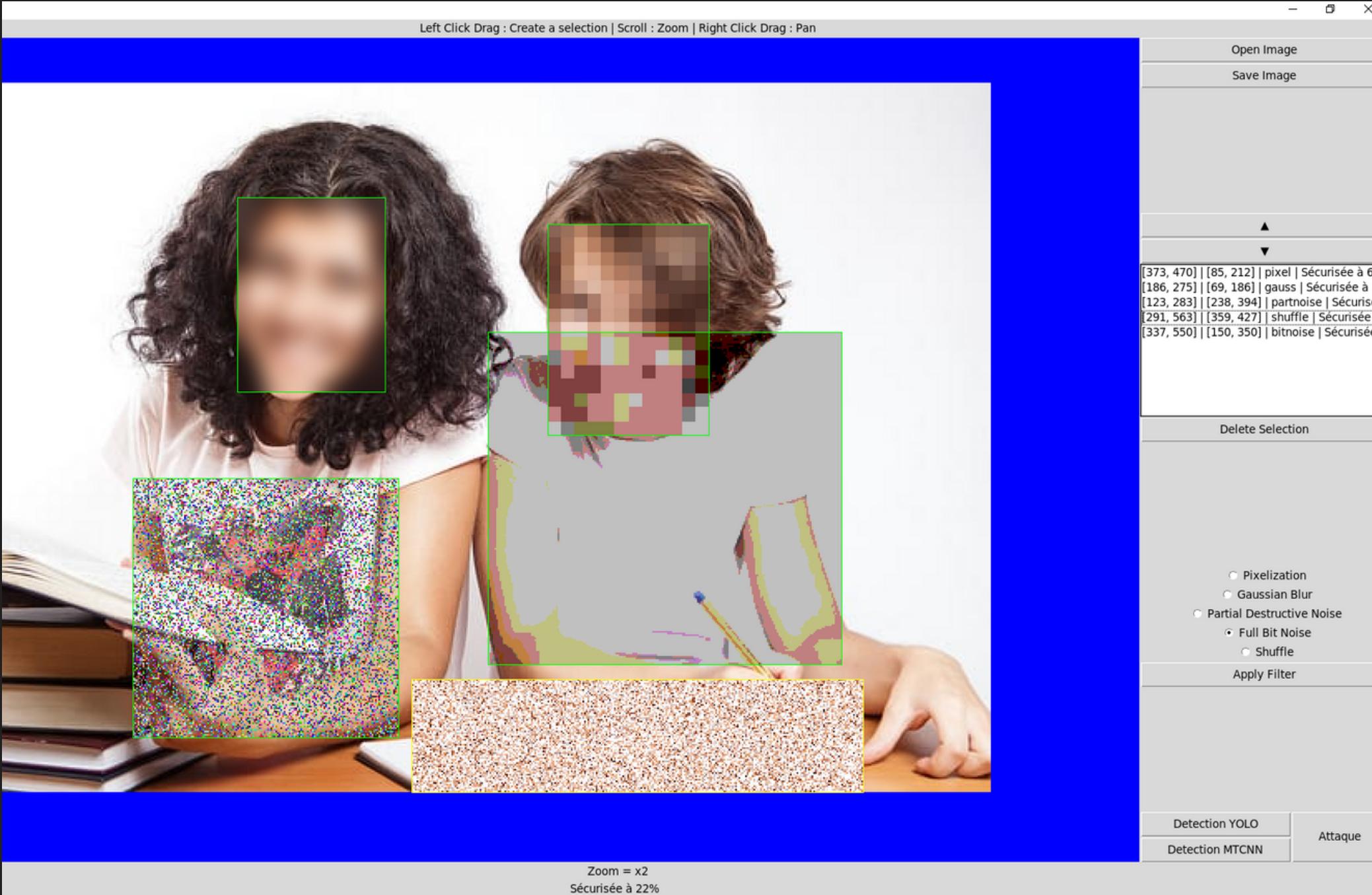
Reconnaisances et Attaques

La réversibilité : Capacité à retrouver le contenu original à partir d'un contenu obscurci.

EDSR : Enhanced Deep Residual Networks for Single Image Super-Resolution



Interfaçage Tkinter

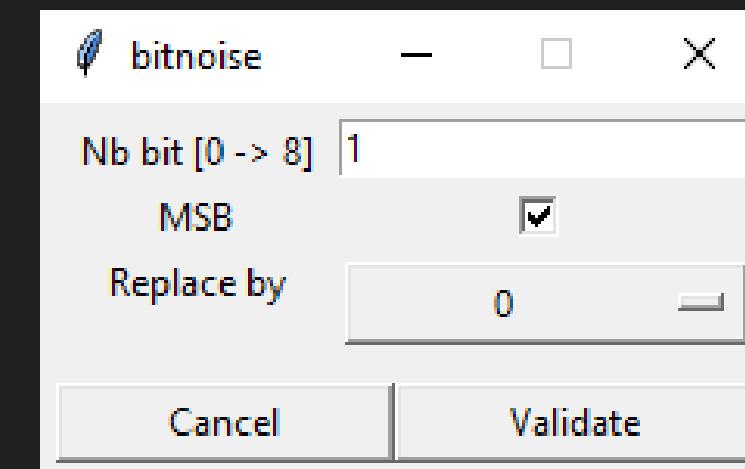


Tkinter

Pillow ImageTK

Selection ROI par "*drag and drop*"

Application et gestion de filtre



Démonstration

Conclusion et Perspectives

La métrique HaarPSI est très intéressante pour corréler les avis subjectifs.

Seuil de sécurité visuel à 0.25
(reconnaissance d'une personne)

Seuil de sécurité visuel à 0.20
(reconnaissance d'un objet)

Nos attaques ne sont pas vraiment optimales et pourraient être améliorées.

Ajouter de nouvelles méthodes d'obscurcissement (naturalité).

Améliorer l'interface utilisateur pour faciliter son utilisation.

Merci de votre attention!

References

[1] - Hanaa Abbas, Roberto Di Pietro (2022) : **Sanitization of Visual Multimedia Content: A Survey of Techniques, Attacks, and Future Directions.**

