SAFE-EYE

ÉVALUATION DE LA SÉCURITÉ VISUELLE D'IMAGES OBSCURES PAR CNN

ÉTAT DE L'ART

L'évolution des technologies actuelles implique une augmentation du **volume des données en ligne** ainsi qu'une augmentation du risque **d'atteinte à la vie privée**.

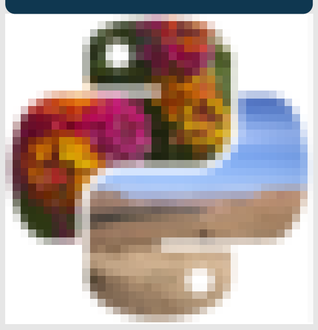
Des techniques d'obscurcissement des images ont ainsi été développées pour protéger les informations sensibles des images.

OBSCURCISSEMENT

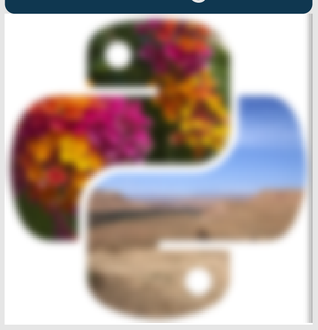
L'obscurcissement pour une image consiste à **modifier voire supprimer des éléments sensibles** tout en conservant certaines caractéristiques visuelles permettant son traitement.

Il en existe plusieurs catégories : *"Face Morphing", "Inpainting", Chiffrement, transfert de style, filtres d'image, etc*.
Dans notre cas, nous avons essentiellement utilisé des techniques basées sur les filtres d'image :

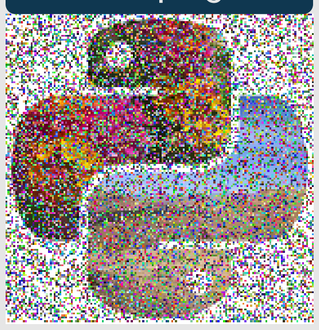
Pixélisation



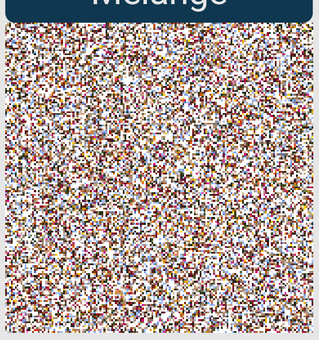
Floutage



Masquage

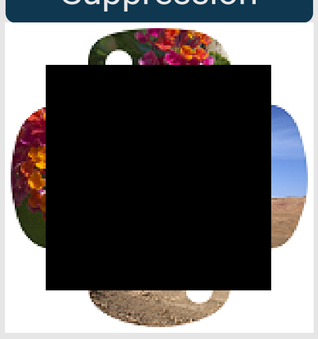


Mélange



Bruitage

Suppression



MÉTRIQUES

Les différences pixel à pixel (**PSNR**) ne sont pas intéressantes pour mesurer la qualité visuelle.
Pour imiter le système visuel humain, nous utilisons des mesures de similarité telles que :

PSNR

- **SSIM** permettant de comparer des images selon trois caractéristiques : la Luminance, le Contraste et la Structure.

SSIM

- **HaarPSI** permettant d'utiliser les amplitudes des coefficients d'ondelettes de Haar à haute fréquence pour définir les similitudes locales.

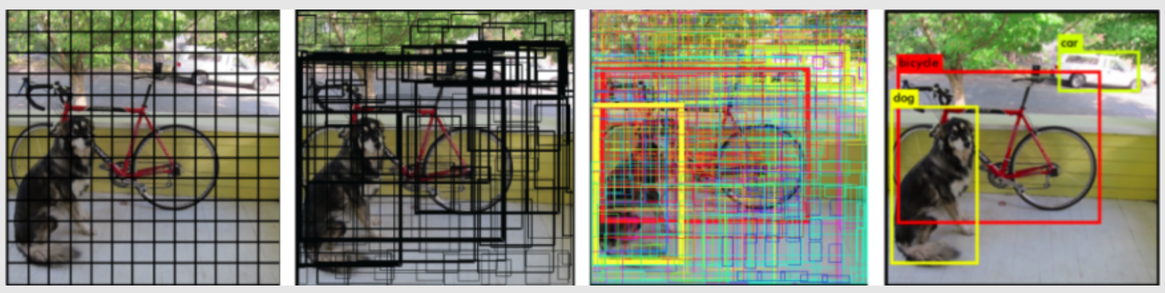
HaarPSI

RECONNAISSANCE

Il existe **plusieurs types de contenu sensible** que l'on peut chercher à obscurcir (biométrique, non biométrique, confidentiel, censuré). Notre application propose deux types de reconnaissance :

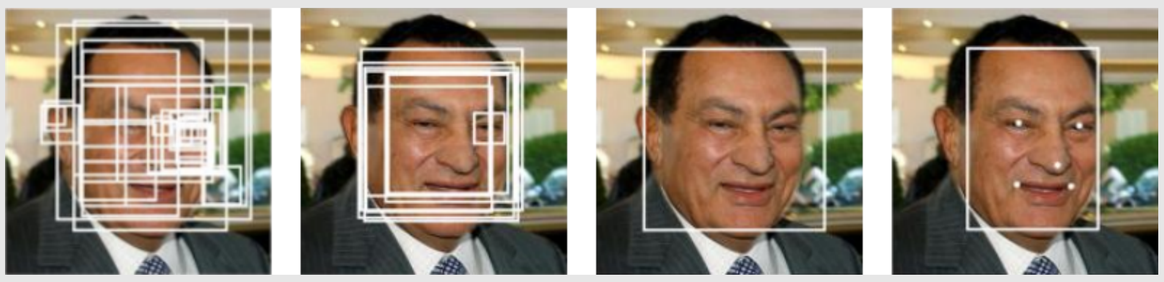
YOLOv3 : *You Only Look Once (Version 3)*

Entraîné sur la base de données "COCO" (80 classes d'objets), ce modèle fonctionne via une grille de cellules permettant de créer des boîtes englobantes. Un algorithme '*Non-maximum Suppression*' est utilisé pour unifier les boîtes englobantes superposées finales.



MTCNN : *Multi-task Cascaded Convolutional Networks*

Ce modèle permet de détecter efficacement les visages humains avec des repères faciaux. Il fonctionne à l'aide de trois autres CNN : (*P-Net, R-Net, O-Net*) permettant la création des boîtes englobantes, leurs simplifications et enfin la détection des repères faciaux.



SÉCURITÉ VISUELLE

Les méthodes d'obscurcissement sont plus ou moins efficaces selon un **compromis "*intimité-utilité*"**. Ce compromis est défini par la capacité à divulguer des informations traitables sans pour autant révéler le contenu.

La naturalité peut également être un objectif de sécurité en soi pour les techniques de remplacement (*Face Morphing*). Cette propriété subjective permet de dissimuler le fait que l'image ait été obscurcie.

Plusieurs mesures (objectives et subjectives) sont nécessaires pour avoir un indice de la sécurité visuelle d'une image. De plus certaines méthodes d'obscurcissement sont vulnérables à des **attaques permettant leur réversibilité** :

- Defloutage (Flou Gaussien)
- Inpainting (Masquage)
- Super Résolution (Pixélisation)
- Decypher (Mélange)

CONCLUSION

D'après nos résultats, la métrique **HaarPSI est la plus intéressante pour corréler les avis subjectifs**. De plus elle semble indiquer un **seuil de sécurité visuelle lorsqu'elle approche de 0.25 ou moins**. Néanmoins, nos attaques permettant de produire des valeurs de HaarPSI plus élevées semblent être décorrélées des avis subjectifs indiquant que certaines attaques ajoutent plus de confusion que de précision.

FLORENTIN DENIS
KHÉLIAN LARVET