SNT – PHOTOGRAPHIE NUMÉRIQUE
FAKE NEWS
COMMENT SAVOIR SI UNE IMAGE/VIDÉO EST TRUQUÉE ?

https://youtu.be/-5BoejLNc9A (< 4 min)

https://youtu.be/hhlnFMOEFas (≈ 2 min)

/!\ L'outil Amnesty YouTube Dataviewer mentionné dans cette vidéo est très simple d'utilisation et donne parfois des résultats intéressants, mais il est assez limité : si une vidéo a été très légèrement modifiée, par exemple raccourcie, l'outil n'est pas en mesure de détecter son origine. Par ailleurs, il ne fonctionne qu'avec des vidéos YouTube.

1. LE BON SENS ET L'OBSERVATION

La première chose à faire est de se demander qui a mis l'image en ligne. Le fait que cela provienne d'un compte « certifié » ou détenu par une personnalité n'indique pas nécessairement que l'image soit vraie, mais si elle a été publiée par un compte créé une heure avant, vous pouvez légitimement douter.
La « fiabilité » du site sur lequel elle circule est importante. Si vous avez un doute sur un site, tapez son nom sur Google, regardez s'il a une page Wikipédia. Demandez-vous s'il a déjà été épinglé pour avoir diffusé de fausses informations.

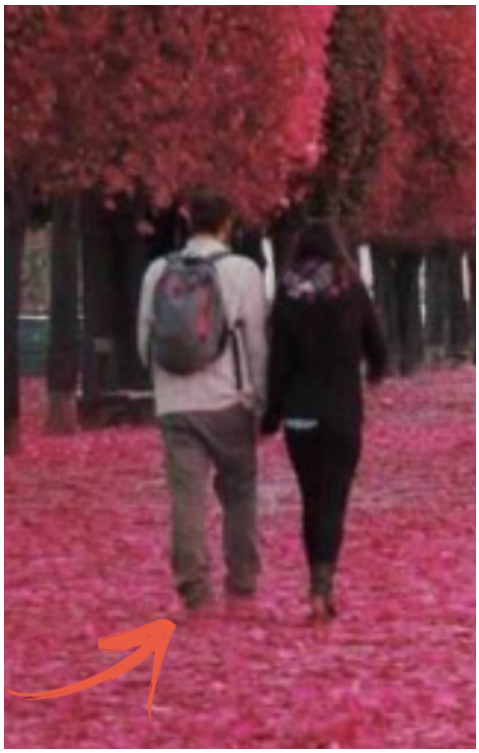
Parfois, il suffit de s'attarder deux minutes sur une image pour voir si elle est fausse.
Prenons l'exemple d'une capture d'écran d'un tweet :

Rien qu'en regardant la police d'écriture ou en la comparant à celle d'autres tweets, on peut remarquer que seul le troisième n'est pas un montage.

S'il s'agit d'un écran de smartphone du même type ou modèle que le vôtre, prenez votre smartphone et regardez si vous avez les mêmes informations au même endroit : est-ce que l'heure est affichée de la même façon ? Est-ce que les bulles de conversation ont la même forme ? Les couleurs sont-elles similaires ?

N'hésitez pas aussi à zoomer sur l'image : si vous remarquez qu'elle est très pixelisée à certains endroits et pas à d'autres, c'est une bonne raison pour être méfiant.

Par exemple, la photo suivante est évidemment un montage.
C'est joli, certes, mais les couleurs sont trop vives pour que ce soit réaliste, non ?
Et en zoomant sur les pieds du passant à gauche, on remarque que ses chaussures ont disparu…



Autre exemple : si on vous présente cette photo en vous disant que ça se passe en France, un peu de bon sens (regardez la tenue des policiers) !

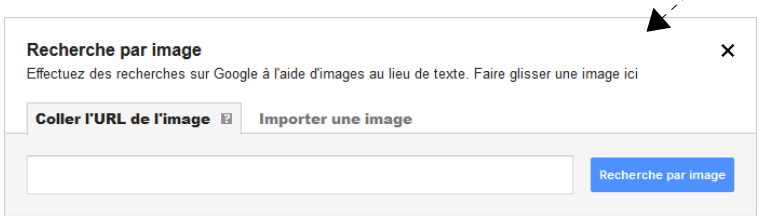
De manière générale, on se méfie de tout ce qui est trop beau, trop spectaculaire, insolite ou incroyable.

2. LA RECHERCHE INVERSÉE

https://youtu.be/rN_Vh9wZEyI
(≈ 2 min)

Dans la vidéo, on vous conseille de faire une capture d'écran (rappel : avec Firefox, faire « clic droit » et « Effectuer une capture d'écran »), mais si c'est possible, préférez (pour ne pas altérer l'image) copier l'adresse URL de l'image (« clic droit » et « Copier l'adresse de l'image ») et aller dans Google Images.

Vous pouvez aussi enregistrer directement l'image (« clic droit » et « Enregistrer l'image sous… ») lorsque c'est possible.



Google Images indexe même les images qui ne datent que de quelques heures. Il est exceptionnel pour identifier le contenu textuel lié aux images. Cependant, il ne parvient pas à identifier les variations importantes (recadrage, édition, etc.). Google n'est pas très fort pour trouver d'autres exemples d'un visage ou de personnes d'apparence similaire, mais assez bon pour trouver la version originale et non recadrée d'une photo à partir de laquelle une capture d'écran a été prise.

Mentionnons donc deux autres moteurs de recherche inversée :

• **Yandex** est très efficace, avec une capacité effrayante à reconnaître les visages, les paysages et les objets. Ce site russe s'appuie fortement sur le contenu généré par les utilisateurs, tels que les sites d'avis touristiques (par exemple FourSquare et TripAdvisor) et les réseaux sociaux (par exemple, les sites de rencontres), pour des résultats remarquablement précis avec des requêtes de reconnaissance faciale et paysagère.
Ses points forts résident dans les photographies prises dans un contexte européen ou ex-soviétique.

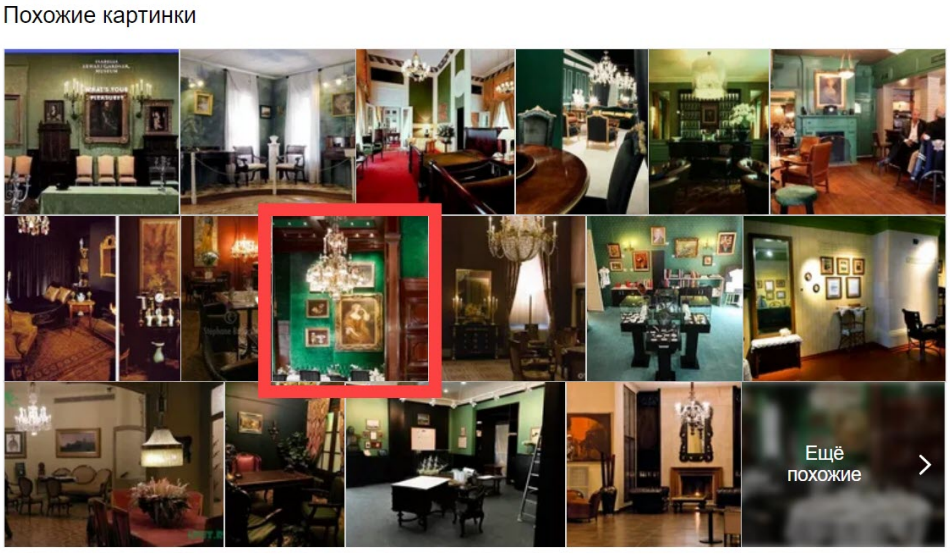
Alors que Google et Bing peuvent simplement rechercher d'autres photographies montrant une personne avec des vêtements similaires et des caractéristiques faciales générales, Yandex recherchera ces correspondances, ainsi que d'autres photographies d'une correspondance faciale.

• **TinEye** est exceptionnel pour trouver des correspondances partielles. Bien qu'il ne connaisse pas immédiatement les nouvelles images, il identifiera généralement les images largement diffusées, ainsi que les images des médias.

Autres astuces

Si vous savez faire, vous pouvez pixelliser/flouter/brouiller les éléments d'une photographie afin d'inciter le moteur de recherche à se concentrer sur l'arrière-plan.

Exemple : sur cette photo, le téléchargement de l'image exacte ne ramènera pas les résultats montrant où elle a été prise. Après pixellisation de la personne, Yandex nous permet de savoir où l'image a été prise : un hôtel populaire à Vienne.



Source : bellingcat

Imaginons que la France connaisse des inondations importantes. Un de vos amis vous envoie cette photo, avec en commentaire « juste à côté du port de Marseille… OMG ! ».

Fake ou pas fake ?

« Au large des côtes sud-africaines, un énorme requin blanc suit un kayak » :
Fake ou pas fake ?

« L'armée américaine proche du pont de San Francisco, en entraînement. Un requin déboule !!! »
Fake ou pas fake ?

3. L'ANALYSE DES MÉTADONNÉES

Nous l'avons déjà vu, l'analyse des données EXIT (IPTC et XMP) peut donner des informations.
Cela est malheureusement assez rare sur une image partagée sur un réseau social, puisque ce dernier supprime et/ou modifie souvent les métadonnées. Mais ça vaut la peine d'ouvrir ces données et de les analyser. Une idée pourrait être de trouver l'image originale (par exemple sur un blog), via une recherche inversée, puis d'analyser les métadonnées de cette image.

Si les métadonnées indiquent que Adobe Photoshop a été utilisé, on peut bien sûr penser qu'une retouche a été faite… mais retouche ne veut pas dire truquage. On peut vouloir retoucher les couleurs, la luminosité, etc. afin de rendre la photo plus jolie.

Soyons très clairs : l'analyse des images est une tâche complexe. Il n'y a pas de solution à un bouton qui vous dira si une image est réelle ou modifiée numériquement. Les situations, le contenu, les forums et les questions sont aussi nombreux que les options de modification des images : il n'y a pas de résultat d'analyse automatisé. Une vidéo peut être « vraie » mais le commentaire qui l'accompagne totalement « faux ».

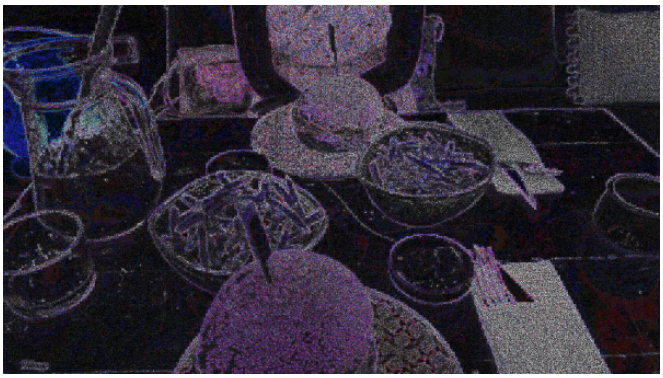
4. L'ANALYSE ELA : COMPLEXE MAIS PARFOIS UTILE

Le site FotoForensics propose un outil qui peut aider à savoir si une image a subi une retouche. Il permet d'afficher les métadonnées mais aussi d'analyser cette image avec des algorithmes complexes, qu'il faut savoir utiliser, dont l'algorithme ELA.
/!\ Le serveur est public, donc évitez d'envoyer une photo de vous sur le site, et surtout n'envoyez pas de contenu pornographique, de nudité ou sexuellement explicite.

L'algorithme ELA (*Error Level Analysis* = analyse de niveau d'erreur) est explicitement conçu pour être utilisé avec JPEG : il permet d'identifier les zones d'une image qui sont à différents niveaux de compression. En effet, l'algorithme JPEG compresse les images avec perte : il enlève une certaine quantité des hautes fréquences et réduit les différences entre les bords, les textures et les surfaces à contraste élevé.
Chaque réencodage (réenregistrement) de l'image ajoute plus de perte de qualité à l'image.
Plus précisément, l'algorithme JPEG fonctionne sur une grille de 8 × 8 pixels. Chaque carré 8 × 8 est compressé indépendamment. Si l'image n'est pas modifiée, tous les carrés 8 × 8 doivent avoir des potentiels d'erreur similaires : chaque carré doit se dégrader approximativement au même rythme. Si une image est modifiée, les zones modifiées apparaîtront avec un niveau d'erreur plus élevé (plus clair).
ELA enregistre l'image à un niveau de qualité JPEG spécifié (souvent 75 %). Ce réenregistrement introduit une quantité d'erreur sur toute l'image. L'image réenregistrée est ensuite comparée à l'image d'origine.

Original

ELA



/!\ Si une image est réenregistrée plusieurs fois, elle peut être entièrement à un niveau d'erreur minimum : la compresser encore (= réduire les hautes fréquences) n'y fera plus rien : l'image est compressée au maximum. Dans ce cas, ELA retournera une image noire et aucune modification ne peut être identifiée à l'aide de cet algorithme. Conclusion : avant de faire l'ELA, toujours chercher la meilleure version de l'image possible (recherche inversée) ! Une photo diffusée sur Facebook ou Twitter ne contient probablement pas beaucoup de détails, puisqu'ils enregistrent les images à un niveau de basse qualité.

« Original »

ELA

Avec les images JPEG, l'image entière doit être à peu près au même niveau. Si une section de l'image est à un niveau d'erreur significativement différent, cela indique probablement une modification numérique :

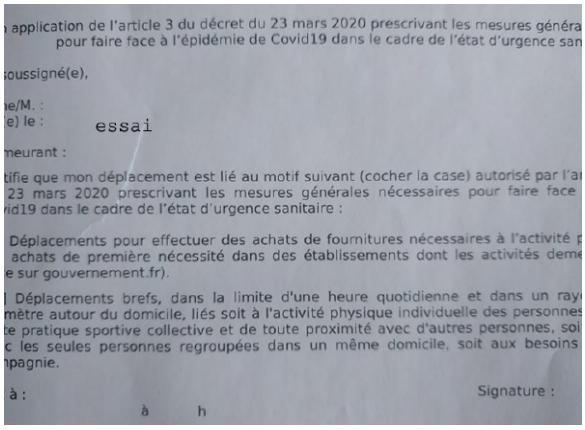
Original

le masque a sans doute été rajouté ou modifié →

ELA

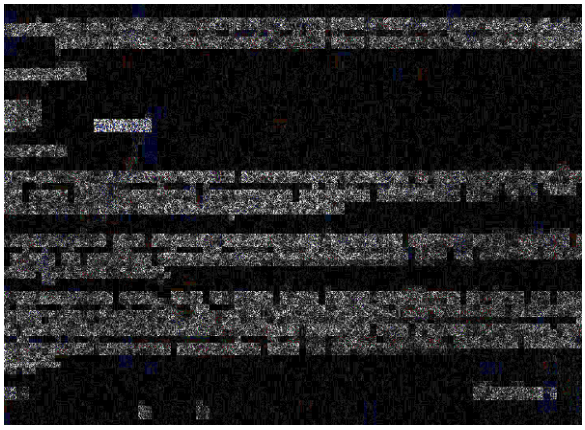
/!\ Les régions à coloration uniforme, comme un ciel bleu uni ou un mur blanc, auront probablement un résultat ELA inférieur (couleur plus foncée) que les bords à contraste élevé, comme du texte :

Original



seul le « essai » a été rajouté avant le dernier enregistrement →

ELA



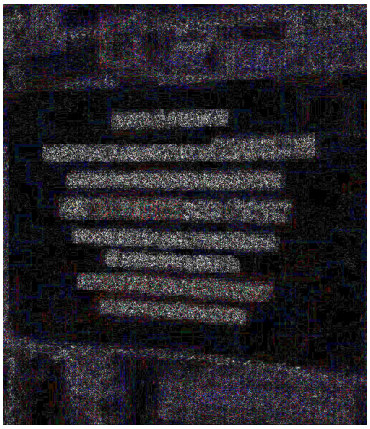
Concrètement, les choses à rechercher :

Bords	Des bords <u>similaires</u> devraient avoir une luminosité <u>similaire</u> dans le résultat ELA. **Tous les bords à contraste élevé doivent se ressembler et tous les bords à faible contraste doivent se ressembler**. Avec une photo originale, les bords à faible contraste doivent être presque aussi lumineux que les bords à contraste élevé.
Textures	**Des textures <u>similaires</u> devraient avoir une coloration <u>similaire</u> sous ELA**. Les zones avec plus de détails sur la surface, comme un gros plan d'un ballon de basket, auront probablement un résultat ELA plus élevé qu'une surface lisse.
Surfaces	Quelle que soit la couleur réelle de la surface, **toutes les surfaces planes doivent avoir à peu près la même coloration sous ELA**.

Original

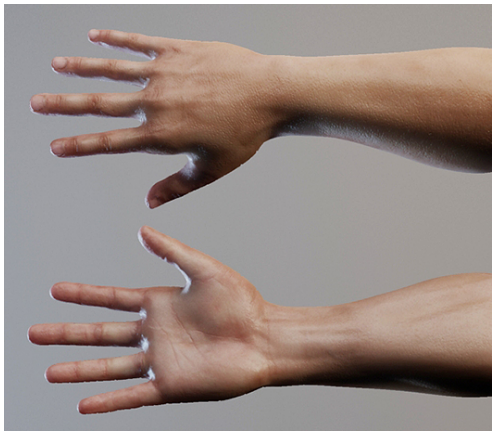
/!\
à priori, aucune modification →
ce qui ne veut pas dire que le message diffusé est vrai !
(ici, après enquête, une erreur de date seulement, mais c'est vrai lorsque l'image est partagée en avril 2020)

ELA

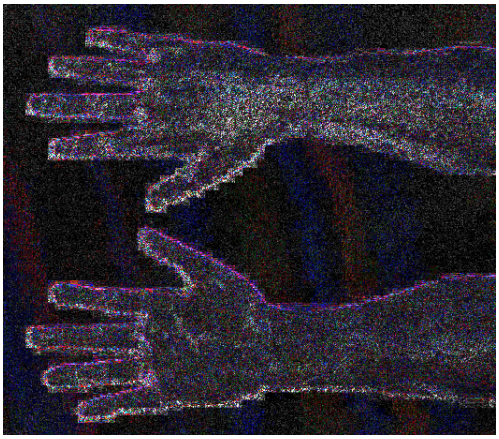


Remarque : la forte présence d'arc-en-ciel suggère qu'un produit Adobe, comme Photoshop ou Lightroom, a été utilisé pour enregistrer l'image (même si certains appareils photo numériques peuvent produire des arcs-en-ciel et que Gimp en produit un peu). Il n'identifie pas les modifications intentionnelles.

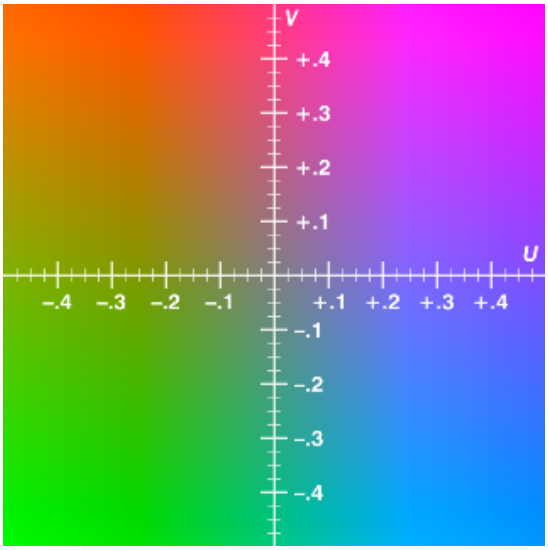
Original



ELA



JPEG convertit les couleurs RGB dans l'espace colorimétrique YUV. Les couleurs à contraste élevé dans la même grille, comme le noir et blanc, l'orange et le bleu ou le vert et le violet (extrémités opposées de l'espace colorimétrique YUV), génèrent généralement des valeurs ELA plus élevées que les couleurs similaires de la même grille.



U = chrominance de bleu

V = chrominance de rouge

quantité de coloration

À retenir sur ELA

/!\ ELA n'est qu'un algorithme qui montre la quantité de différence qui se produit lors d'un réenregistrement JPEG. Plus de blanc signifie plus de changement, et le noir indique aucun changement.

/!\ Avec ELA, "blanc" ne signifie pas modifié ; le blanc signifie un potentiel de niveau d'erreur plus élevé qui doit être comparé à des bords et surfaces similaires sur l'image.

/!\ Quand il s'agit d'analyser des images à partir d'une capture d'écran, ne le faites pas. Chercher une meilleure version de l'image par une recherche inversée.

/!\ Si une image est plusieurs fois compressée en JPEG, ELA donnera une image noire.[1]

/!\ L'interprétation des résultats peut ne pas être concluante : il est important de valider les résultats avec d'autres techniques et algorithmes d'analyse.

1 Le crime organisé l'a bien compris. Un exemple ici : https://www.hackerfactor.com.

5. DERNIERS CONSEILS ET PRATIQUE

Je vous conseille par ailleurs d'utiliser l'excellent outil InVID, qu'on peut ajouter en extension sur les navigateurs Chrome et Firefox.

Une des nombreuses possibilités est d'extraire d'une vidéo (par exemple diffusée sur Facebook) plusieurs images, qui pourront être directement vérifiées par une recherche inversée. Exemple : ----------▶

https://youtu.be/TgOpesGF-9 g (< 2 min)

Si cet outil vous intéresse, le mieux est de l'installer et d'essayer les fonctionnalités (parfois complexes) mais vous pouvez aussi regarder cette vidéo (< 10 min) :

https://youtu.be/om8ZbgDOl_s

Enfin, si vous avez un doute sur une « fake news », il est fort possible que quelqu'un avant vous ait déjà fait une enquête : faire une recherche dans Google, lire les commentaires d'une vidéo… mais aussi (la première chose à faire) aller sur des sites de « fact checking » : AFP, HoaxBuster, Snopes (américain), Hoax-Net (belge)…
Et pensez à faire des recherches, en identifiant les sites sur lesquels l'information est étudiée.

Prêt à pratiquer ?

1. Télécharge cette photo sur ton disque dur :

a) Faire une analyse ELA. Peux-tu en déduire quelque chose ?

b) Fais une recherche inversée sur Google Images.
Obtiens-tu quelque chose ?

c) Fais une recherche inversée sur TinEye.
Obtiens-tu quelque chose ?

d) Sauras-tu retrouver l'original de cette photo, et sa provenance ?

2. Voici l'image d'un énorme insecte ! Fake ou pas fake ?

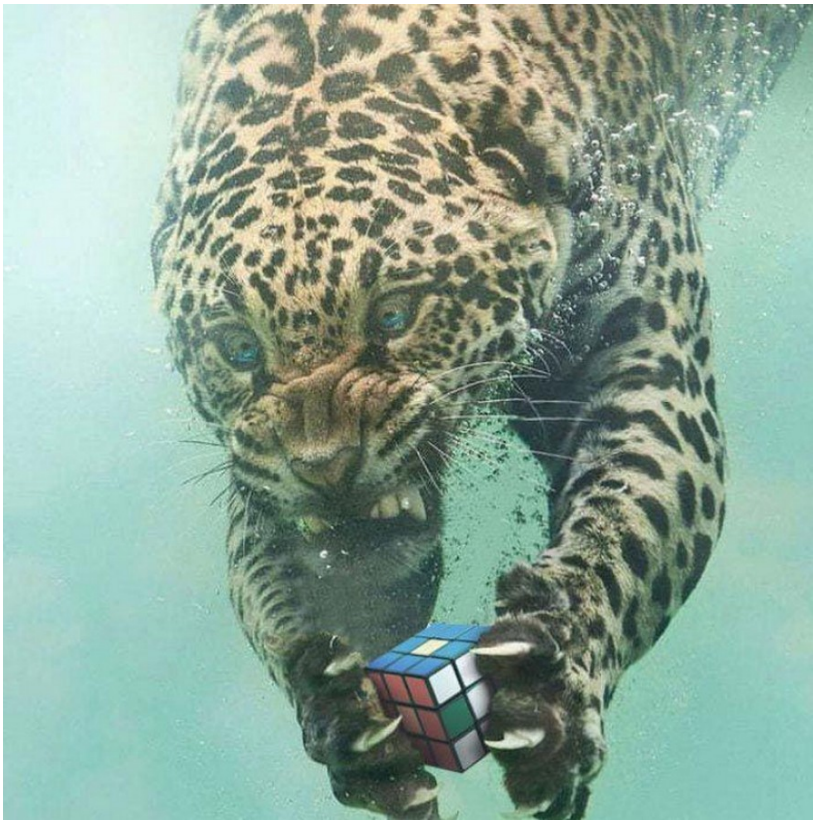
3. Cette photo de tatouage est impressionnante !

On dirait que le tatouage est creusé.
Franchement, bravo à l'artiste !
Fais une analyse ELA de cette image.
Que peux-tu en déduire ?

4. a) Fais une analyse ELA de cette image.
Que peux-tu en déduire ?

b) Peux-tu dire vers quelle année cette « photo » a été prise ?

5. Selon toi, cette photo a-t-elle été modifiée ? Est-ce un « fake » ?



6. Un ami vous dit qu'il est tombé fou amoureux d'une fille en discutant en ligne avec elle depuis des mois. Il s'apprête à la rejoindre : elle habite dans l'Ohio, aux USA. Cet ami vous dit qu'il a confiance, et que cette fille lui a même envoyé une photo de son passeport.

a) Analyse cette image*. Que peux-tu en dire ?

* *pas de lien vers l'image, à toi de te débrouiller*

b) Pourrais-tu retrouver la fille dont la photo figure sur le passeport ?

7. Pourrais-tu retrouver l'endroit exact pris ici en photo ?

Tu trouveras la réponse ici, mais cherche par toi-même d'abord, c'est faisable ! ^_^

6. OMG ! TRÈS DÉSTABILISANT

Avec les progrès en « intelligence artificielle » (IA), on peut désormais créer des visages de personnes qui n'existent pas… Plus précisément, Nvidia a présenté en décembre 2018 son IA capable de générer des visages artificiels à partir de vrais visages. Derrière ce site, se cache Philip Wang, un ingénieur logiciel chez Uber, qui souhaite éduquer aux potentiels dangers de l'IA.
Sur le site, il suffit d'actualiser la page pour voir apparaître des visages tous différents les uns des autres et qui semblent pourtant réalistes. Bien évidemment, certains visages présentent quelques défauts, mais c'est déjà une prouesse… Adresse du site : https://www.thispersondoesnotexist.com/

Si l'intelligence artificielle et le machine learning t'intéressent, je te conseille de lire le livre *Comprendre le Deep Learning : Une introduction aux réseaux de neurones* de Jean-Claude Heudin, mais aussi de regarder les excellentes vidéos sur le sujet (parfois un peu compliquées) de Lê Nguyên Hoang (Science4All sur YouTube). Par exemple, sa vidéo n°49 sur le sujet traite des GAN (Generative Adversarial Network), et les photos créées par le site thispersondoesnotexist.com sont tirées d'un GAN : deux réseaux sont utilisés. Le premier va tenter de créer l'image à partir de données brutes. Le second est considéré comme "un juge" c'est lui qui décidera si l'image correspond aux attentes ou non. Si ce n'est pas le cas, le premier réseau recommencera.