FORMATIO | RESSOURCE INFORMATIQUE BUREAUTIQUE

Hypertrucage (ou Deepfake)

Claude Drouin

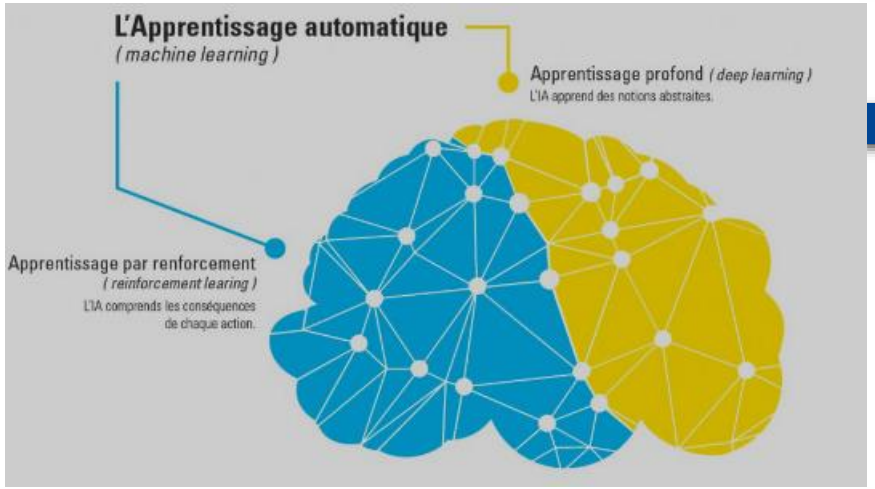
Octobre 2023

D’où vient le nom en anglais Deepfake

- Le nom provient d’un utilisateur de la plateforme Reddit (un réseau social)
- C’est une contraction de « deep learning » + Fakes
- Il fut le premier a partager ses hypertrucages en « remplaçant » des célébrités dans du contenu pour adulte
- Les premières célébrités ciblées étant Emma Watson, Scarlet Johansson et Barak Obama
- L’engouement fut tel qu’il y a maintenant une explosion de contenu hypertruqué

L’Hypertrucage

- Utilise des algorithmes d’apprentissage machine et d’intelligence artificielle
- Manipule et génère du contenu visuel (image, vidéo) et audio
- Résulte en un contenu qui trompe l’auditeur



Un premier exemple

L’humain peut facilement être trompé

- Si on voit and entends avec nos propres yeux et oreilles, nous croyons que ça existe ou c’est vrai, MÊME si c’est peu probable!
- Notre cerveau peut être ciblé et avoir une perception erronée tout comme avec les illusions optiques et figures ambigues (perception bistable)

Exemple

Lapin ou canard?

Vase ou Figures ?

Image Ambigüe (Bistable en anglais)

L’importance d’être éduqué sur l’hypertrucage

- C’est une **technologie accessible** car ça permet à tous, avec un minimum de formation, de données, de matériel et logiciel de produire du contenu hypertruqué
  - Par exemple, Zao, une application mobile asiatique permet à son usager de remplacer des personnages par sa figure dans des émission, films, etc. le tout gratuitement!

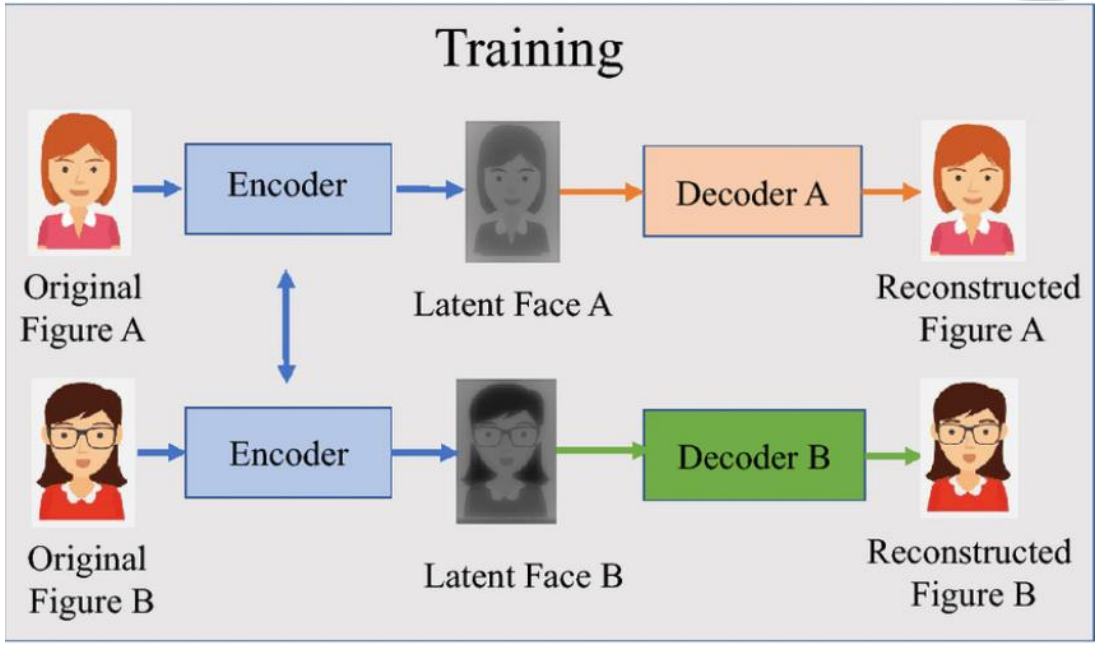
Comment ça fonctionne

- La technologie principale utilisée est l'apprentissage profond (Deep Learning)
- C'est une méthode permettant d'entraîner les réseaux neuronaux profonds (minimum de 3 niveaux de neurones)
- Plus spécifiquement, l'hypertrucage va utiliser l'architecture d'autoencodage permettant d'apprendre dà partir de caractéristiques discriminantes
- On l'utilise ainsi pour remplacer l'image (et aussi l'audio!!!) d'une personne par une autre

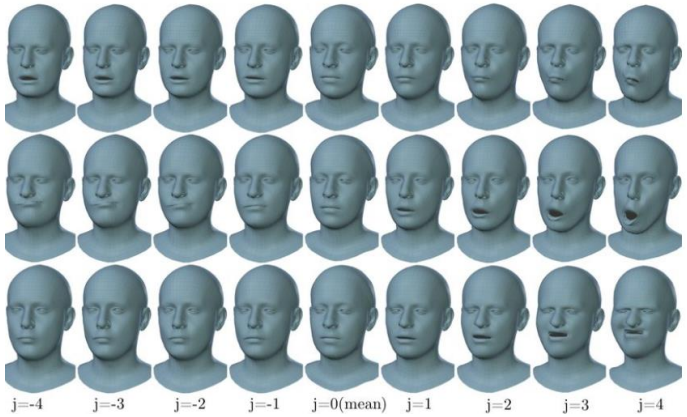
Pour en savoir plus sur l'intelligence artificielle, visitez SLACK pour consulter la présentation sur l'intelligence artificielle à https://clubinformati-0j77622.slack.com/archives/C031H8V6MDE/p1675372096593439

Comment ça fonctionne

- ✓ Étape 1 - Entraînement
  - ✓ entraîner le RNN avec plusieurs vidéos montrant la figure des 2 sujets
  - ✓ Autoencodage des aspects faciaux (mimiques, complexion, etc.)
  - ✓ Le même encodeur est utilisé pour entraîner le modèle à partir des 2 vidéos.
  - ✓ Le processus compresse (encodage) et décompresse (decodage) les images
  - ✓ Suite à l'encodage, le processus crée un ensemble de vecteurs de base (latent Face)

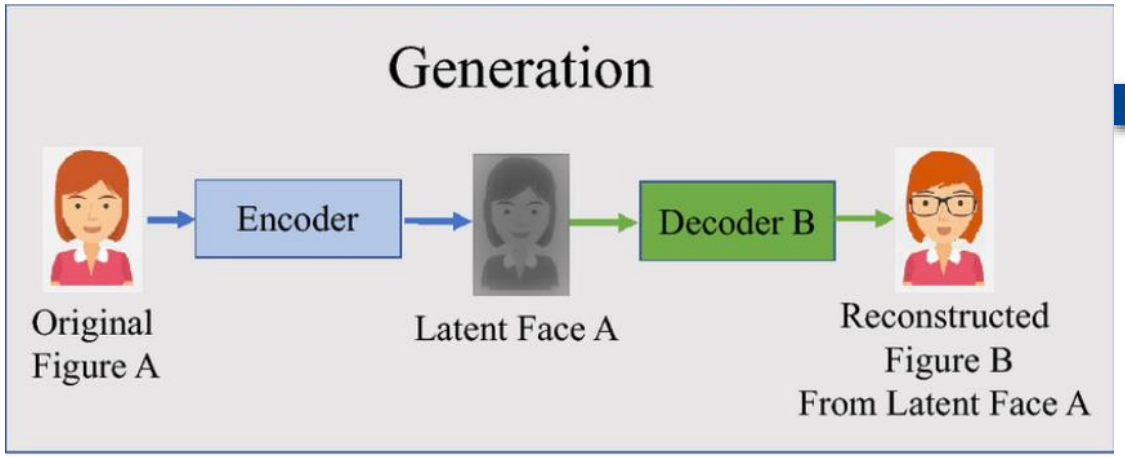


(a) Training Phase



Comment ça fonctionne(bis.)

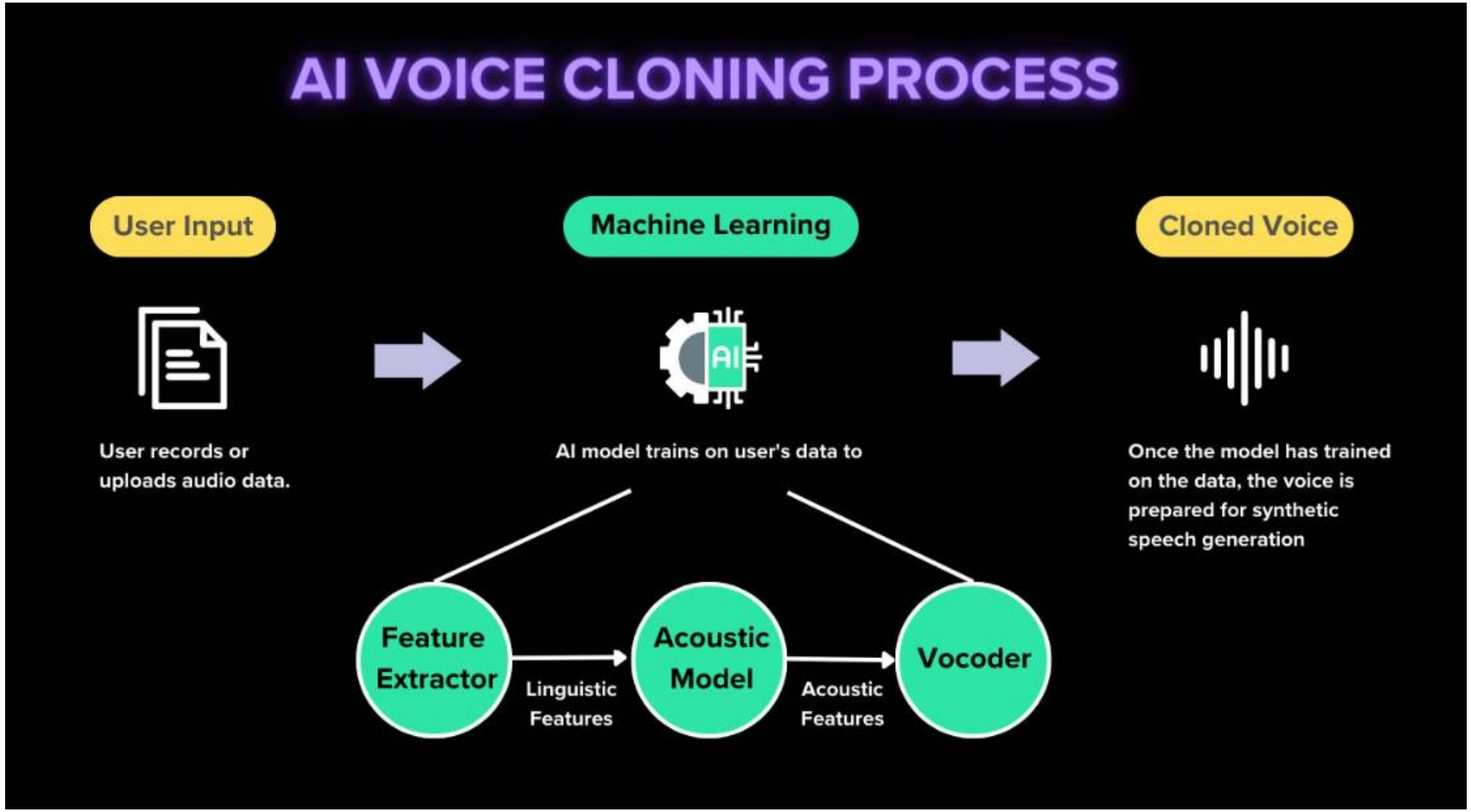
- Étape 2 – Étape de génération
  - Ce qui rend la technologie d'échange de visage possible, c'est de trouver un moyen de forcer les deux visages latents à être codés sur les mêmes caractéristiques.
  - L'hypertrucage a résolu ce problème en faisant en sorte que les deux réseaux partagent le même encodeur, tout en utilisant deux décodeurs différents.



(b) Generation Phase

Un autre exemple

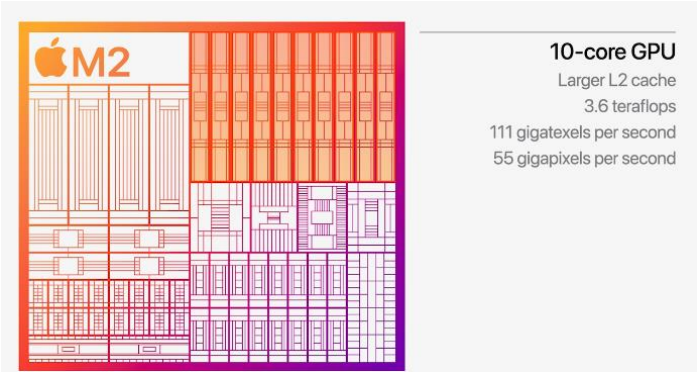
Clonage de la voix



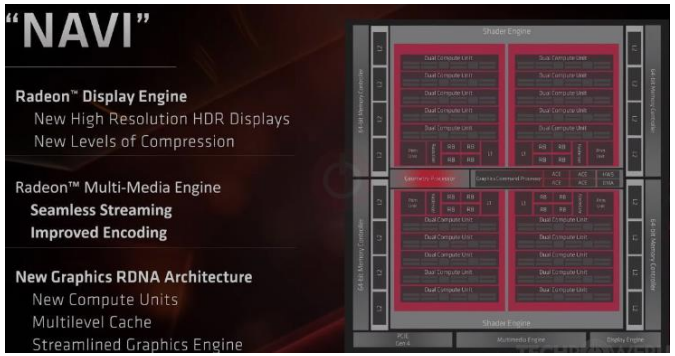
Clonage de la voix

Matériel requis pour le traitement de l'hypertrucage

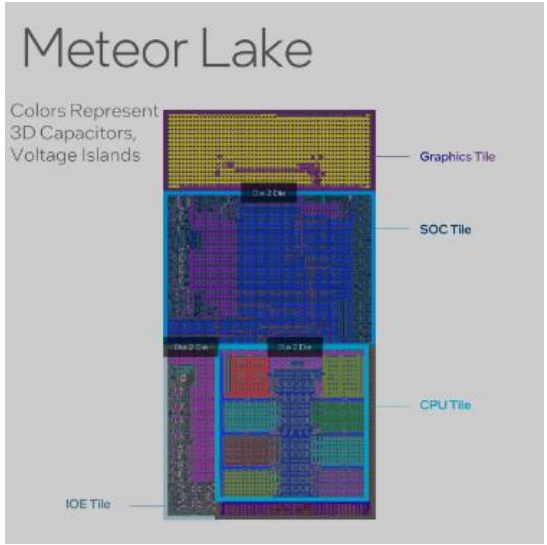
Unité de traitement graphique (GPU en anglais)



Apple



AMD



Intel

Nvidia

Logiciels

Hypetrucage Figure

- Faceswap
- Swapface
- FaceApp
- Zao
- Reface
- SpeakPic
- DeepFaceLab
- FakeApp
- Wombo
- Deepfake webs
- Instagram Deepfake bot
- Deepfake studio

Hypertrucage Voix (Clonage)

- Ressemble.ai
- Descript
- Cereproc
- Respeecher
- Realtime V Cloning
- Ispeech
- VoiceWear
- Readspeak.ai
- ReplicaStudios

Un outil pour permettre de reconnaître l'hypertrucage

https://detectfakes.media.mit.edu/

Questions/Commentaires?