

Deep Fake content classification : Is it real or not ?

contact: Sarah A. Carneiro

2023/2024

1 Introduction

Alimentés par des algorithmes d'IA sophistiqués, les deepfakes manipulent l'audiovisuel pour créer des simulations hyper-réalistes de personnes dans des situations fictives. Initialement vus comme divertissants, ils sont devenus des outils de désinformation et de propagande. Leur diffusion aisée sur les réseaux sociaux pose des défis à la communication numérique et à la confiance dans les médias.

Les deepfakes vont au-delà des préoccupations de vie privée pour impacter la politique et la justice. Ils peuvent fausser des vidéos de personnalités publiques et manipuler des preuves judiciaires. Cette prolifération nécessite des stratégies de détection pour distinguer le vrai du faux. Il est donc impératif de développer des mécanismes de détection robustes capables de distinguer entre le contenu authentique et manipulé (Figure 1).

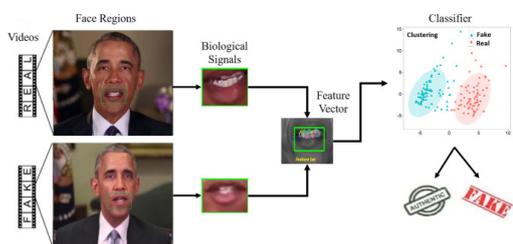


Figure 1: Diagram from: DFT-MF: Enhanced deepfake detection using mouth movement and transfer learning

2 Objectifs

À travers l'utilisation de techniques d'apprentissage automatique et d'algorithmes de traitement d'images, ce projet vise à développer des méthodes de détection de contenu deepfake dans les images et les vidéos. Les objectifs du projet incluent :

- Explorer les ensembles de données deepfake, tels que celui proposé par le Défi de Détection de Deepfake de Kaggle, et les ressources pour comprendre les caractéristiques et les défis associés à la détection de deepfake
- Comparer différents modèles de détection de deepfake et évaluer les performances des modèles de détection de deepfake
- Analyser les forces et les limitations des différentes approches de détection et proposer des stratégies pour améliorer la précision de la détection
- Apprendre à communiquer les résultats en rédigeant un rapport.

3 Jour des Projets

- Démontrer que nous sommes capables, avec différentes méthodes d'apprentissage profond et de traitement d'images, de déterminer si un contenu image/vidéo est faux ou non.

4 Bon à savoir

Les éléments ci-dessous peuvent faciliter le développement du projet :

- Les étudiants devraient être disposés à investir du temps dans l'apprentissage et le développement du projet.
- Il est souhaitable que les étudiants aient une compréhension de l'apprentissage automatique ainsi qu'une certaine compréhension des techniques de traitement d'images.
- Les étudiants devraient être capables de communiquer, de lire et d'écrire en anglais.

5 Initial References

- Agarwal, S., Farid, H., El-Gaaly, T. and Lim, S.N., 2020, December. Detecting deep-fake videos from appearance and behavior. In 2020 IEEE international workshop on information forensics and security (WIFS) (pp. 1-6). IEEE.
- Rahman, Ashifur, et al. "A qualitative survey on deep learning based deep fake video creation and detection method." Aust. J. Eng. Innov. Technol 4.1 (2022): 13-26.