

N° d'ordre : D -

THESE

présentée

devant l'Institut National des Sciences Appliquées de Rennes

en vue de l'obtention du

DOCTORAT

spécialité : signal, Image, vision

par M Nicolas BEUVE

Intitulé : Détection de vidéos deepfakes basée sur l'apprentissage profond

Directeurs de Thèse : Wassim Hamidouche, Olivier Déforges

Date, heure et lieu de soutenance : 20 novembre 2023, INSA Rennes, Amphi B
10h00

Membres du jury (nom, prénom, titre et établissement de rattachement, fonction)

DUGELAY Jean-Luc, H, PrU, Eurecom Nice

AOUADA Djamilia, F, Professeur Assistant, Université du Luxembourg

TALEB-AHMED Abdelmalik, PrU, Université de Lille

KACETE Amine, Ingénieur Recherche, B-com, Cesson-Sévigné

HAMIDOUCHE Wassim, MdC INSA Rennes

DEFORGES Olivier, PrU INSA Rennes

RESUME DE LA THESE

Les récentes avancées en apprentissage profond, plus précisément concernant les modèles génératifs, ont conduit au développement d'outils permettant de modifier l'identité ou l'expression faciale d'individus dans une vidéo. Ces vidéos modifiées, appelées deepfakes, représentent un moyen efficace de propager de fausses informations et posent une menace majeure, surtout lorsque la cible du deepfake est une personnalité politique.

Dans ce contexte, de nombreux chercheurs ont entrepris de développer des méthodes de détection automatique de deep-fakes afin de les intercepter avant leur propagation. Les utilisateurs de ces modèles de détection ne se trouvent pas nécessairement au même niveau dans la chaîne de propagation d'un deepfake, ce qui implique des besoins différents en matière de détecteurs. En effet, on distingue deux grandes catégories de détecteurs : les détecteurs passifs, qui supposent que l'origine de la vidéo n'est pas connue, et les détecteurs actifs, qui supposent que l'on a accès à la vidéo originale avant que le deepfake ne soit produit.

Dans cette thèse, des contributions concernant les deux types de détecteurs sont proposées, avec deux contributions améliorant les détecteurs passifs : une nouvelle fonction de coût et de nouvelles méthodes d'inférence pour des modèles d'apprentissage profond. De plus, une troisième contribution utilise des watermarks semi-fragiles pour effectuer la détection active de deepfake.

Mot clés : Deepfake, Apprentissage Profond