Sujet de stage Master 2/Ingénieur

DETECTION DE DEEPFAKE VIDEO

Mots-clés : apprentissage, vidéos, forensic

PREAMBULE

Le laboratoire GREYC (UMR 6072) est un laboratoire basé en Normandie, composé de plus de 230 membres dont une centaine de chercheurs ou enseignants-chercheurs permanents, dans 6 équipes de recherche couvrant de nombreux domaines de l'informatique, l'automatique et l'électronique.

L'équipe de recherche en cybersécurité SAFE du GREYC, basée à Caen, mène des activités de recherche dans le domaine de la sécurité informatique autour de 3 thèmes : 1) Biométrie, 2) Architecture et modèle de sécurité et 3) Sciences de l'investigation (Forensic). L'équipe propose un sujet de stage sur la **détection de deepfake vidéo**.

Ce sujet s'inscrit dans le cadre de travaux en cours dans le thème Forensic de l'équipe. Il fait suite à la réalisation d'un premier prototype fonctionnel de détection. Afin de poursuivre ces travaux prometteurs, l'équipe recherche un étudiant de niveau Master2 pour un stage de cinq mois du 1er mars au 31 juillet 2023.

CONTEXTE

Avec l'amélioration rapide du traitement des images et des vidéos, un grand nombre d'outils logiciels d'édition d'images et de vidéos puissants ont vu le jour, ce qui permet de modifier ou de générer facilement des images et des vidéos numériques sans traces visibles. Ces images/vidéos malveillantes fournissent des informations falsifiées, qui sont généralement utilisées à des fins illégales ou trompeuses, telles que la falsification de preuves, la propagande politique, etc. Parmi tous les schémas de falsification, nous nous attaquerons à la détection de vidéos falsifiées, car cette technique est principalement déployée pour permettre au fraudeur de se faire passer pour une autre personne.
En fait, la capacité d'usurper l'identité d'une personne sur une vidéo en direct, au téléphone ou même pendant un appel vidéo va poser un défi aux systèmes de vérification en ligne, qui reposent sur la reconnaissance du visage ou de la voix, soit par un humain, soit automatiquement par un logiciel.

OBJECTIFS

Le problème des vidéos truquées notamment en remplaçant les visages a fait l'objet d'une grande attention ces deux dernières années, notamment après l'avènement de la technologie deepfake qui manipule des images et des vidéos avec des outils d'apprentissage profond. L'algorithme deepfake peut remplacer les visages de la vidéo

cible par des visages de la vidéo source en utilisant des autoencodeurs ou des réseaux antagonistes génératifs. Avec cette technologie, il est extrêmement simple de générer des vidéos avec des visages, à condition d'avoir accès à de grandes quantités de données. Néanmoins, certains éléments permettent de repérer un hypertrucage. La plupart des hypertrucages donnant l'impression que quelque chose ne va pas, vient du fait que les erreurs induites par le processus de génération laissent des traces résiduelles [*Korshunov and Marcel, 2018*].
Cette sensation provient du fait qu'une partie de la vidéo diffusée provient d'une seconde vidéo.
La détection basée sur des Réseaux de Neurones est couramment utilisée dans la littérature, où la tâche de détection des vidéos truquées est considérée comme une tâche de classification ordinaire [*Bonettini et al., 2021* ; *Li et al., 2020b*]. Les caractéristiques de cohérence temporelle sont également exploitées pour détecter les discontinuités entre les images adjacentes d'une fausse vidéo. Pour trouver des caractéristiques distinctives, les artefacts visuels générés par le processus de fusion sont exploités dans les tâches de détection [*Montserrat et al., 2020*]. Les approches récemment proposées se concentrent sur des caractéristiques plus fondamentales, où l'empreinte de la caméra et les schémas basés sur les signaux biologiques montrent un grand potentiel dans les tâches de détection [*Frank et al., 2020*].

TRAVAIL

Durant ce stage, vous serez amené à :

- Réaliser un état de l'art sur l'utilisation des signaux faibles pour la détection des hypertrucages, notamment, sur l'analyse statistique d'identifier des zones de l'image provenant de sources ;
- Développer un algorithme d'extraction des caractéristiques dites profondes à l'aide d'un réseau de neurones à convolution (CNN) qui pourront être combinées à des informations spatio-temporelles ;
- Un classifieur sera ensuite entrainé afin de réaliser une classification de la vidéo en deux catégories : "vidéo falsifiée" ou "vidéo non falsifiée".

Une attention particulière sera portée sur la mise à l'échelle des méthodes développées, c.-à-d. la capacité à détecter des modifications indépendamment de la longueur de la vidéo, et de leur positionnement dans la vidéo.

LIEU

GREYC, équipe SAFE
Bat F, ENSICAEN
6, bd Maréchal Juin, 14 032 - Caen

Condition d'exercice : Travail en présentiel

PROFIL RECHERCHE

Etudiant.e en Master 2 Recherche ou en dernière année d'école d'ingénieur spécialisé en informatique, image et/ou intelligence artificielle.

Une solide formation en machine learning est indispensable.

Des connaissances et expériences solides en traitement d'images/vidéos, apprentissage profond et programmation (Python, TensorFlow/PyTorch) sont nécessaires.

CANDIDATURE

Pour postuler, envoyer par email aux encadrants un dossier avec CV, lettre de motivation, relevés de notes des deux dernières années de formation, ainsi que toute pièce susceptible de renforcer la candidature (lettre de recommandation, etc.).

CONTACT

Christophe Charrier (christophe.charrier@unicaen.fr)
Emmanuel Giguet (emmanuel.giguet@unicaen.fr)

REFERENCES

[*Korshunov and Marcel, 2018*] Korshunov, P. and Marcel, S. (2018). Deepfakes: a new threat to face recognition? assessment and detection. arXiv preprint arXiv:1812.08685

[*Bonettini et al., 2021*] Bonettini, N., Cannas, E. D., Mandelli, S., Bondi, L., Bestagini, P., and Tubaro, S. (2021). Video face manipulation detection through ensemble of cnns. In2020 25th International Conference on Pattern Recognition (ICPR), pages 5012–5019. IEEE.

[*Li et al., 2020b*] Li, L., Bao, J., Zhang, T., Yang, H., Chen, D., Wen, F., and Guo, B. (2020b). Face x-ray for more general face forgery detection. in 2020 ieee. In CVF Conference on Computer Vision and Pattern Recognition (CVPR), volume 2.

[*Montserrat et al., 2020*] Montserrat, D. M., Hao, H., Yarlagadda, S. K., Baireddy, S., Shao, R., Horváth, J., Bartusiak, E., Yang, J., Guera, D., Zhu, F., et al. (2020). Deepfakes detection with automatic face weighting. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pages 668–669

[*Frank et al., 2020*] Frank, J., Eisenhofer, T., Schönherr, L., Fischer, A., Kolossa, D., and Holz, T. (2020). Leveraging frequency analysis for deep fake image recognition. In International conference on machine learning, pages 3247–3258. PMLR