

LI, Yuezun et al. Celeb-df: A large-scale challenging dataset for deepfake forensics. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**. 2020. p. 3207-3216.

O artigo inicia dizendo que as redes neurais profundas (DDNs) não criaram, mas sim ascenderam ainda mais a criação de vídeos e imagens falsos. O uso de DDNs transformou essa criação muito mais eficiente, convincente, rápida e fácil. Várias aplicações surgiram (FakeApp, FaceSwap, DeepFaceLab) e criar imagens falsas se tornou simples e popular.

Com esse aumento e ascensão das deepfakes surge a necessidade e o interesse na criação de métodos de detecção de deepfakes, como o DeepFake Detection Challenge. Assim, surgiram algoritmos dedicados a esse trabalho que podem ser separados em três categorias. Os métodos da primeira categoria consistem em encontrar irregularidades nas faces geradas, como olhos piscando irregularmente, posição da cabeça errada entre outras irregularidades faciais. Assim são capturados padrões comportamentais dos indivíduos para que se possa comparar e detectar os vídeos falsos. A segunda categoria utiliza a detecção de artefatos de imagem gerados no processo criação de vídeos ou imagens falsas. A terceira categoria de detecção consiste no treinamento de várias DDNs utilizando vídeos reais e falsos para classificar se a mídia é alterada.

O autor prossegue falando sobre as bases de dados existentes sobre deepfakes até o momento da escrita do artigo, sendo elas UADFV, DF-TIMIT, FF-DF, DFD e DFDC. Depois, ele conclui que apesar de essas bases proverem vários vídeos que possuem artefatos visuais que são facilmente detectados, o objetivo desse artigo é criar uma base de dados com informações mais relevantes para dar suporte ao futuro de desenvolvimento de métodos de detecção de deepfakes.

A base de dados proposta pelo texto é a Celeb-DF, contendo 590 vídeos reais e 5.639 vídeos falsos, totalizando mais de dois milhões de frames de imagens falsas nessa base. Esses vídeos contêm aproximadamente 13 segundos à 30 quadros por segundo. A captura desses dados foi feita utilizando um algoritmo aprimorado de criação de deepfakes a partir de vídeos na plataforma do YouTube. Utilizando o Celeb-DF e outros datasets com os métodos de detecção de deepfakes existentes até a criação do artigo, o Celeb-DF teve a avaliação de desempenho mais abrangente entre todas. Para os testes foram considerados 9 métodos de detecção de deepfakes: Two-stream, MesoNet, HeadPose, FWA, VA, Xception, Multi-task, Capsule, DSP-FWA. Comparando cada método utilizando cada dataset o Celeb-DF foi o que mais dificultou para os algoritmos de detecção, pelo motivo de seus dados terem maior qualidade, assim, sendo mais fidedignos a vídeos reais.

Ao final, o autor conclui que o trabalho apresentou um novo desafio de criar uma base de dados em larga escala para avaliar e desenvolver algoritmos de detecção de deepfakes. E para o futuro deve-se aumentar a quantidade e qualidade da solução proposta, o Celeb-DF.

--