

## **Entregable 2.**

### **Caracterización del conjunto de datos**

Estudiantes:

Bayron Daymiro Campaz Hurtado

Juan David Diaz Monsalve

Santiago Gutierrez Bolaños

Tutor:

Christian Camilo Urcuquí López, Msc

Universidad ICESI  
Facultad de Ingeniería  
Ingeniería de Sistemas  
Cali  
2020

Actualmente existen diferentes conjuntos de datos creados con el fin de entrenar y evaluar métodos de detección de Deepfake. Estos conjuntos de datos cuentan con diversas características y propiedades que varía de conjunto a conjunto. El objetivo de este documento es realizar una breve descripción de cada uno de estos conjuntos teniendo en cuenta elementos como la cantidad, la resolución y la variabilidad de las imágenes que los componen, así como otros atributos. Cabe aclarar que se han escogido los conjuntos de datos más relevantes desde la academia en este ámbito.

## **DF-TIMIT**

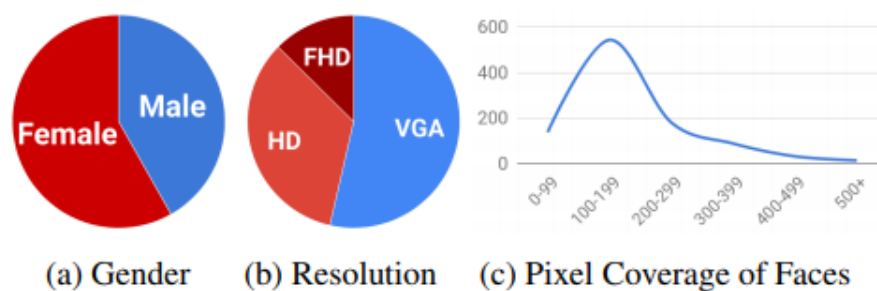
El Deepfake-TIMIT dataset incluye 640 vídeos Deepfake generados con Faceswap-GAN basado en el conjunto de datos Vid-TIMIT. Los vídeos se dividen en dos subconjuntos de igual tamaño: DF-TIMIT-LQ y DF-TIMIT-HQ, con caras sintetizadas de  $64 \times 64$  y  $128 \times 128$  píxeles, respectivamente.

Seleccionaron manualmente 16 pares de personas de aspecto similar de la base de datos DF-TIMIT. Para cada 32 sujetos, entrenaron dos modelos diferentes, en el documento, conocidos como modelos de baja calidad (LQ), con  $64 \times 64$  de entrada / salida, y de alta calidad (HQ), con tamaño de  $128 \times 128$ . Dado que hay 10 vídeos por persona en la base de datos DF-TIMIT, generaron 320 vídeos correspondientes a cada versión, lo que resulta en un total de 640 vídeos con caras intercambiadas. No se realizó ninguna manipulación en el canal de audio. Este conjunto de datos fue lanzado en diciembre de 2018. (Korshunov & Marcel, 2018)

## **FF-DF**

FaceForensics ++ es un conjunto de datos que consta de 1000 secuencias de vídeo originales que han sido manipuladas con cuatro métodos automatizados de manipulación facial: Deepfakes, Face2Face, Faceswap y NeuralTextures. Los datos provienen de 977 vídeos de youtube y todos los vídeos contienen una cara frontal rastreable, lo que permite que los métodos automáticos de manipulación generen falsificaciones realistas. (Rossler et al., 2019)

La distribución de los datos respecto al género, la resolución y el cubrimiento de píxeles en las caras es como sigue:



**Figura 1.** Composición del conjunto de datos FF-DF. En el gráfico (b) VGA denota 480p, HD denota 720p, y FHD denota resolución 1080p. El gráfico (c) muestra el número de secuencias (eje y) con una altura de píxel del cuadro delimitador dado (eje x).

Este conjunto de datos fue lanzado en enero de 2019.

## DFD

El Google/Jigsaw Deepfake Detection Dataset tiene 3068 vídeos Deepfake basados en 363 vídeos originales de 28 individuos de varias edades, géneros y grupos étnicos. No se revela mayor detalle del algoritmo de generación usado, pero si establecen que es una mejora a un algoritmo de generación de Deepfake básico. Este conjunto de datos fue lanzado en septiembre de 2019. (Li, et al., 2019)

## DFDC

Los vídeos de este conjunto de datos incluyen condiciones de iluminación y ángulos del rostro variados. Los creadores de este conjunto de datos manifiestan que los participantes pudieron grabar sus vídeos con el fondo que desearan, lo que produjo fondos visualmente diversos. La aproximación de la distribución general de género y raza en este conjunto de datos es 74% femenino y 26% masculino; y 68% caucásicos, 20% afroamericanos, 9% del este asiático y 3% del sur asiático. Este conjunto de datos consta de 66 individuos. Este conjunto de datos fue creado usando dos diferentes algoritmos de generación, sin embargo, no se especifican cuáles son. DFDC fue lanzado en octubre de 2019. Finalmente, este conjunto de datos tiene un total de 4,464 clips de entrenamiento y 780 clips de prueba. (Dolhansky, et al., 2019)

## CELEB-DF

El conjunto de datos Celeb-DF consta de 590 vídeos reales y 5,639 vídeos Deepfake (correspondientes a más de dos millones de cuadros de vídeo). La duración promedio de todos los vídeos es de aproximadamente 13 segundos con una velocidad de cuadro estándar de 30 cuadros por segundo. Los vídeos reales se eligen de vídeos de YouTube disponibles al público, correspondientes a entrevistas de 59 celebridades con una distribución diversa en sus géneros, edades y grupos étnicos. El 56.8% de los sujetos en los vídeos reales son hombres y el 43.2% son mujeres. El 8,5% tiene 60 años o más, el 30,5% tiene entre 50 y 60 años, el 26,6% tiene 40 años, el 28,0%

tiene 30 años y el 6,4% tiene menos de 30 años. El 5,1% son asiáticos, el 6,8% son afroamericanos y el 88,1% son caucásicos. Además, los vídeos reales exhiben una gran variedad de cambios en aspectos tales como los tamaños de cara de los sujetos (en píxeles), orientaciones, condiciones de iluminación y fondos. Los vídeos de Deepfake se generan intercambiando caras para cada par de los 59 sujetos. Los vídeos finales están en formato MPEG4.0 y no cuentan con sonido. Celeb-DF fue lanzado en noviembre de 2019. (Li, et al., 2019)

Al terminar esta fase de comprensión de los datos se decidió hacer énfasis en la detección de Deepfakes en el conjunto de vídeos de Celeb-DF, dado que es el conjunto de datos que más dificulta la detección de contenido en los métodos detectores actuales. (Li, et al., 2019)

## REFERENCIAS

- Korshunov, P., & Marcel, S. (2018). Deepfakes: a new threat to face recognition? assessment and detection. arXiv preprint arXiv:1812.08685. Recuperado de <https://arxiv.org/pdf/1812.08685.pdf>
- Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). Faceforensics++: Learning to detect manipulated facial images. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1-11). Recuperado de <https://arxiv.org/pdf/1901.08971.pdf>
- Dolhansky, B., Howes, R., Pflaum, B., Baram, N., & Ferrer, C. C. (2019). The Deepfake Detection Challenge (DFDC) Preview Dataset. *arXiv preprint arXiv:1910.08854*. Recuperado de <https://arxiv.org/pdf/1910.08854.pdf>
- Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2019). Celeb-df: A new dataset for deepfake forensics. *arXiv preprint arXiv:1909.12962*. Recuperado de <https://arxiv.org/pdf/1909.12962.pdf>
- Google. (10 de Febrero de 2020). Google. Recuperado de Google: <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=es-419>
- Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). Faceforensics++: Learning to detect manipulated facial images. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1-11).[Figura 1]. Recuperado de <https://arxiv.org/pdf/1901.08971.pdf>