

Reporte sobre clasificación de contenido visual.

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

REPORTE SOBRE CLASIFICACIÓN DE CONTENIDO VISUAL.

Para este proyecto resulta fundamental seleccionar los datos (contenido visual) que se usarán tanto para el entrenamiento de los modelos seleccionados como para su prueba. Dado que este proyecto busca mejorar los resultados de un método detector de Deepfake en contenido visual de diferentes calidades, es importante poder establecer cuándo un contenido visual cumple estas características. Para esta tarea se buscó un método o algoritmo que permitiera hacer la clasificación según la calidad.

Se consideraron diversos parámetros del contenido visual como el bitrate, el color, la iluminación entre otros, sin embargo, ninguna de estos parámetros por sí solo puede servir para clasificar el contenido visual. Si bien es cierto que existen muchos parámetros no existe una fórmula o ponderación que nos permita automatizar el proceso.

En la industria de los servicios de distribución de vídeo existen métricas que permiten evaluar cuánta calidad ha perdido un vídeo comprimido respecto al vídeo original. Existen métricas subjetivas en las que se usa la percepción de un grupo de personas para evaluar la calidad y métricas objetivas que mediante software o hardware que permiten evaluar la calidad. Como parte de este proceso de clasificación se planteó la hipótesis de que estas métricas de evaluación de calidad para un video original respecto a sus compresiones podían servir para comparar clips de video diferentes usando un video como referencia. Las métricas evaluadas fueron la VMAF desarrollada por Netflix que busca emular la percepción humana para determinar la calidad y PSNR que mide el nivel de distorsión midiendo el error cuadrático medio entre la señal original y la distorsionada.

Para evaluar esta hipótesis se usó el dataset “The Facebook DeepFake detection challenge (DFDC)” (Dolhansky, Brian, et al., 2019). Inicialmente tomamos seis videos del dataset (1920 x 1080); 2 de calidad baja, 2 de calidad media y 2 de calidad alta esto basados en una métrica subjetiva según la percepción de los tres estudiantes que hacen parte de este proyecto, usamos el software libre “ffmpeg-quality-metrics” (Werner Robitza, 2020) para obtener las métricas VMAF y PSNR; los resultados de ambas métricas resultaron coherentes a lo clasificado sin embargo la muestra de datos era muy pequeña para ser concluyente. Se realizó la misma tarea con 50 videos (usando siempre un video como referencia) se verificaron con percepción humana, obteniendo un 42% de precisión.

Lo anterior permitió concluir que las métricas de evaluación de calidad para un video original respecto a sus compresiones no funcionan para comparar la calidad de clips de videos diferentes y por lo tanto no es útil para clasificar el contenido visual según la calidad. Al no encontrar un método que permitiera clasificar contenido visual evaluando la calidad de forma algorítmica, se decidió clasificar el contenido como datasets de primera generación y de segunda generación. Donde los de primera

generación “incluye caras sintetizadas de baja calidad, límites de empalme visibles, falta de coincidencia de colores, partes visibles de la cara original, y orientaciones faciales sintetizadas inconsistentes, y los de segunda generación mejoran la cantidad y los problemas de calidad que presentan los datasets de primera generación de deepfakes” (Li, Yang, Sun, Qi, & Lyu, 2019).

BIBLIOGRAFÍA

- Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (27 de noviembre de 2019). *Celeb-df: A new dataset for deepfake forensics*. Obtenido de arXiv: <https://arxiv.org/pdf/1909.12962.pdf>
- Dolhansky, B., Howes, R., Pflaum, B., Baram, N., & Ferrer, C. C. (2019). *The Deepfake Detection Challenge (DFDC)*. Obtenido de arXiv: <https://arxiv.org/pdf/1910.08854.pdf>
- Slhck (2020). Calculate quality metrics with Ffmpeg (SSIM, PSNR, VMAF). Obtenido de: <https://github.com/slhck/ffmpeg-quality-metrics>