Contribution ID: **16**   Type: **Poster**

Evaluating detection methods of synthetic images

Thursday 24 April 2025 10:45 (30 minutes)

The paper aims to evaluate the effectiveness of available detectors in distinguishing fake from real photos. Deepfakes are artificially generated images or videos created using artificial intelligence or manual tools like Photoshop. New techniques, such as Generative Adversarial Networks (GANs) and Diffusion Models (DMs) enable the rapid generation of highly realistic images. The research utilizes the StyleGAN3 (GAN) and Stable Diffusion XL (DM) models, which were previously fine-tuned on photos of beaten people obtained from the Internet to generate new images. Tests of the detectors were carried out to assess the detector's precision, sensitivity, and resistance to manipulations, such as graphic filters or compression, as well as resistance to fingerprint removal tools. Older detectors that were not trained on the tested generative models struggle to detect fake photos. In contrast, newer detectors trained on new latest models achieve surprisingly good results. These results highlight the need for continuous updates to detection systems to counteract evolving deepfake generation techniques.

Authors: Mr WÓJTOWICZ, Michal (Warsaw University of Technology); BILIŃSKI, Michał (Warsaw University of Technology)

Presenters: Mr WÓJTOWICZ, Michal (Warsaw University of Technology); BILIŃSKI, Michał (Warsaw University of Technology)

Session Classification: Session B (Poster)