





21.05.2024 Computer Vision Seminar 23/24





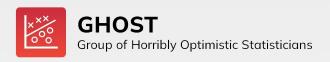
#### 2024: The Super Election Year Countries where a national election is/was held in 2024 Expected number of voters: **2B** (~25% of world pop.) Share of free elections\*: **38%** European Parliament Elections Election moved (martial law) Biggest election India In danger due to coups Burkina Faso Smallest new parliament (16 seats) \* Free and partially free: 75%







General, parliamentary, presidential and economic union elections Sources: Anchor Change Election Cycle Tracker, Statista research



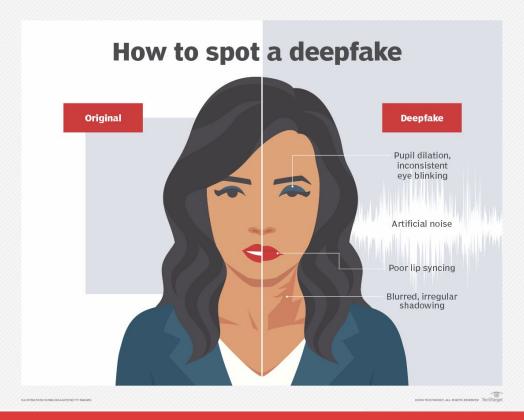


Termin deepfake oznacza zaawansowaną technikę manipulacji treściami multimedialnymi, głównie wideo lub audio, przy użyciu sztucznej inteligencji i głębokiego uczenia.













<u>TikTok Tom Cruise deepfake creator: public shouldn't worry about 'one-click fakes' - The Verge</u>







#### Jak rozpoznać deepfake?

- Zwróć uwagę na twarz.
- Zwróć uwagę na policzki i czoło. Czy skóra wydaje się zbyt gładka lub zbyt pomarszczona?
- Zwróć uwagę na oczy i brwi. Czy cienie pojawiają się w miejscach, których można się spodziewać?
- Zwróć uwagę na okulary. Czy występują jakieś refleksy? Czy jest ich za dużo? Czy kąt padania światła zmienia się, gdy osoba się porusza?





#### Jak rozpoznać deepfake?

- Zwróć uwagę na owłosienie twarzy lub jego brak. Czy włosy na twarzy wyglądają naturalnie?
- Zwróć uwagę na znamiona na twarzy (np. pieprzyki)
- Zwróć uwagę na mruganie. Czy osoba mruga wystarczająco lub zbyt często?
- Zwróć uwagę na ruchy warg. Czy wyglądają naturalnie?





# Rozpoznaj deepfake

**Detect Fakes** 





#### Praca w grupach

- 1. Grupa 1 Types of Deepfakes

  DeepFake: A Deep Learning Approach in Artificial Content Generation | by Brij Gupta | Medium
- 2. Grupa 2 Types of Deepfakes: Techniques of Detecting Deepfakes

  Detecting Deepfakes: Exploring Advances in Deep Learning-Based Media Authentication | by AnalytixLabs | Medium
- 3. Grupa 3 Deepfake a prawo Deepfake laws: is Al outpacing legislation? | Onfido
- 4. Grupa 4 C2PA

  C2PA in DALL · E 3 | OpenAI Help Center



### Wykrywanie deepfake'ów z DL

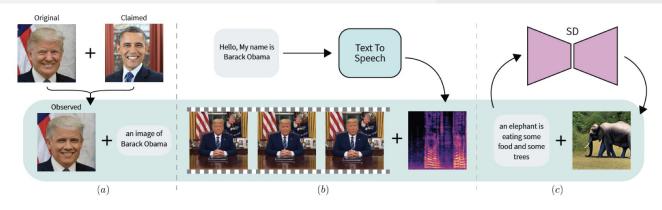


Figure 1: False facts in deepfake attacks. (a) Face forgery: the claimed identity is seamlessly blended into the original image. The observed image is accompanied by a false fact i.e., "an image of Barack Obama". (b) Audio-Visual (AV): fake audio is generated to align with the original video or fake video is generated to align with the original audio. Fake media are accompanied by a false fact, that the video and audio describe the same event. (c) Text-to-Image (TTI): the textual prompt is used by a generative model e.g. Stable Diffusion, to generate a corresponding image. The fake image is accompanied by a false fact, that the caption and the image describe the same content.





#### **FACTOR**

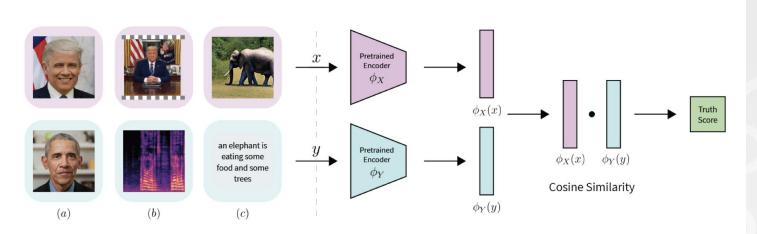


Figure 2: Illustration of FACTOR, our proposed deepfake detection method. FACTOR leverages the discrepancy between false facts and their imperfect synthesis within deepfakes. By quantifying the similarity using the truth score, computed via cosine similarity, FACTOR effectively distinguishes between real and fake media, enabling robust detection of zero-day deepfake attacks.





## Materialy

- The Incredible Creativity of Deepfakes and the Worrying Future of Al | Tom Graham | TED
- Certified Human: How new Intel tech detects deepfakes in real time | Intel
- How Easy Is It to Make and Detect a Deepfake?
- Overview Of How To Create Deepfakes It's Scarily Simple