DTEC

# TruePersona

Product Information

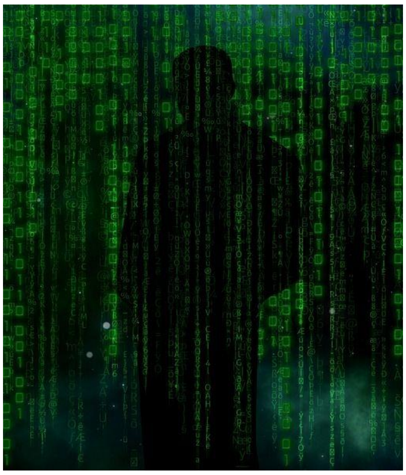
Bank heists and phishing attacks using **voice impersonation** are on the rise due to recent advances in sophisticated **AI deepfake voice cloning technology**. Fake news also poses an increasingly serious problem. Modern state of the art machine learning makes artificial voices virtually indistinguishable from the impersonated human voice just by listening.

Thanks to its state of the art **voice cloning detection** engine, with ***TruePersona*** now you can detect in real time artificially generated voices used by cybercriminals and impostors and thus prevent fraud. Tenths of different technologies of voice synthesis are modeled and detected, being effective even against advanced *vocoders*. New cloning technologies are specifically modeled and added to the recognition engine as they are released for keeping maximum accuracy.

State of the art AI deepfake voice cloning technology can work in two operating modes: **Text-To-Speech** (TTS) and **Voice Modulation** (voice changer). In TTS, human voice recordings are used to train a synthesizer on advance that can then be used to input text and generate a voice mimicking the cloned person. On the other hand, a Voice Modulator changes on the fly a human input voice into the impersonated target voice, adjusting acoustic parameters based on the reference voice to be cloned, and thus there is no need to input text. ***TruePersona*** can analyse both types of attacks and detect subtle audio signal features present in real human voices and not in synthetics voices, and vice versa.

***TruePersona*** is distributed as a **SDK (*Software Development Kit*)** that exports its functionalities through a powerful yet easy to use **API** (*Application Programming Interface*). This highly efficient C++ API allows easy integration into any final application and operating environment.

In today's digital world with humanlike AIs on the rise, ***TruePersona*** is a much needed weapon against fraud:



- **Fake news in social media, radio and TV**: impersonation of celebrities and authorities.
- **Banks heists by phone or video calls**: impersonation of big clients or even colleagues by cybercriminals.
- **Phishing to citizens by phone**: gathering of confidential data and banking accounts, luring family members to make a bank transfer.
- **Malicious AI**: is the doctor I'm talking to remotely by phone or video-call a human being or an AI entity?

In addition, ***TruePersona*** can also work together with ***BioVox*** as state of the art anti-spoofing technology in authentication applications based on voice biometrics.

## PRODUCT

- Solution for detection of artificially generated voice and AI deepfake voice cloning.

## KEY FEATURES

- Effective against both **TTS** and advanced **vocoders**.
- **Language independent**: voice impersonation attacks are detected no matter the language.
- State of the art AI recognition engine: based on advanced **machine learning** algorithms.
- **Advanced audio features extraction**: the audio signal is processed and analysed before feeding the AI recognition engine.
- **Easy to integrate API** for on-premise solutions.
- **Highly optimized C++ recognition engine**: can be integrated into embedded systems.

## TECHNICAL SPECIFICATIONS

- Audio required for synthetic speech detection:
  - Minimum: 5s.
  - Recommended: >12s.
- Verification time: 3X real time (using a single core of the recommended CPU below).
- Disk space for installation: 250 MB.
- RAM usage: <300 MB (typical, depending of the processed audio).
- Supported audio formats: WAV PCM linear 16 bits 8/16 KHZ (recommended), MP3.
- Proprietary C++ API.
- Number of TTS and voice cloning technologies modeled: 23 (new ones will be added as they are released).
- EER[1]: < 1% (dependent on audio quality and channel noise).
- Minimum recommended CPU: Intel Core i5-4460@3.20 GHz or equivalent.

## SUPPORTED PLATFORMS

- Windows® 10, 11.
- Linux, several distributions.

[1] *Equal Error Rate*: the value in which the two opposite error rates (whenever one is reduced, the other one is increased as a consequence) associated to any classifier are made equal. These error rates are: *FRR* (False Rejection Rate - a legitimate human voice is rejected as fake) and *FAR* (False Acceptance Rate – a deepfake artificially generated voice is accepted as human).