SaTC: CORE: Small: Combating AI Synthesized Fake Media Beyond Detection (SaTC-2153112)

Siwei Lyu (siweilyu@buffalo.edu), University at Buffalo, State University of New York

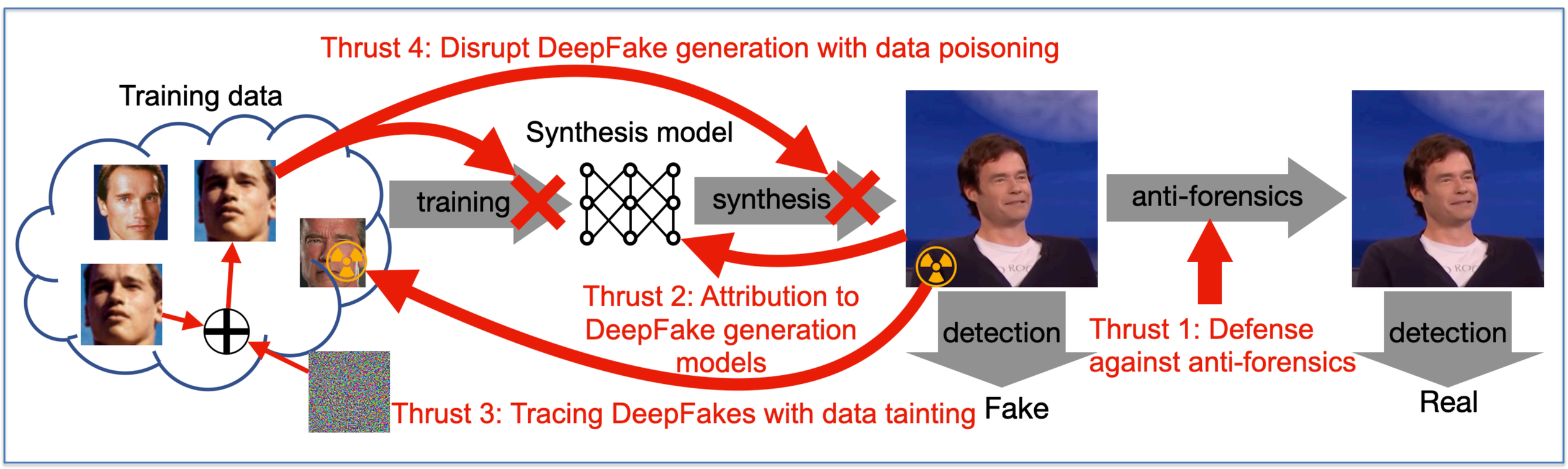
Project URL: https://www.nsf.gov/awardsearch/showAward?AWD_ID=2153112&HistoricalAwards=false

Challenges:

- AI synthesized fake media have negative societal impacts as a means of disinformation.
- Passive detection is not sufficient to provide timely and effective protection for users.

Scientific Impact:

- The results from this project will provide deeper insights about the capacities and vulnerabilities of the state-of-the-art detection methods.
- The proactive and active approaches complements the detection methods for more comprehensive defense.

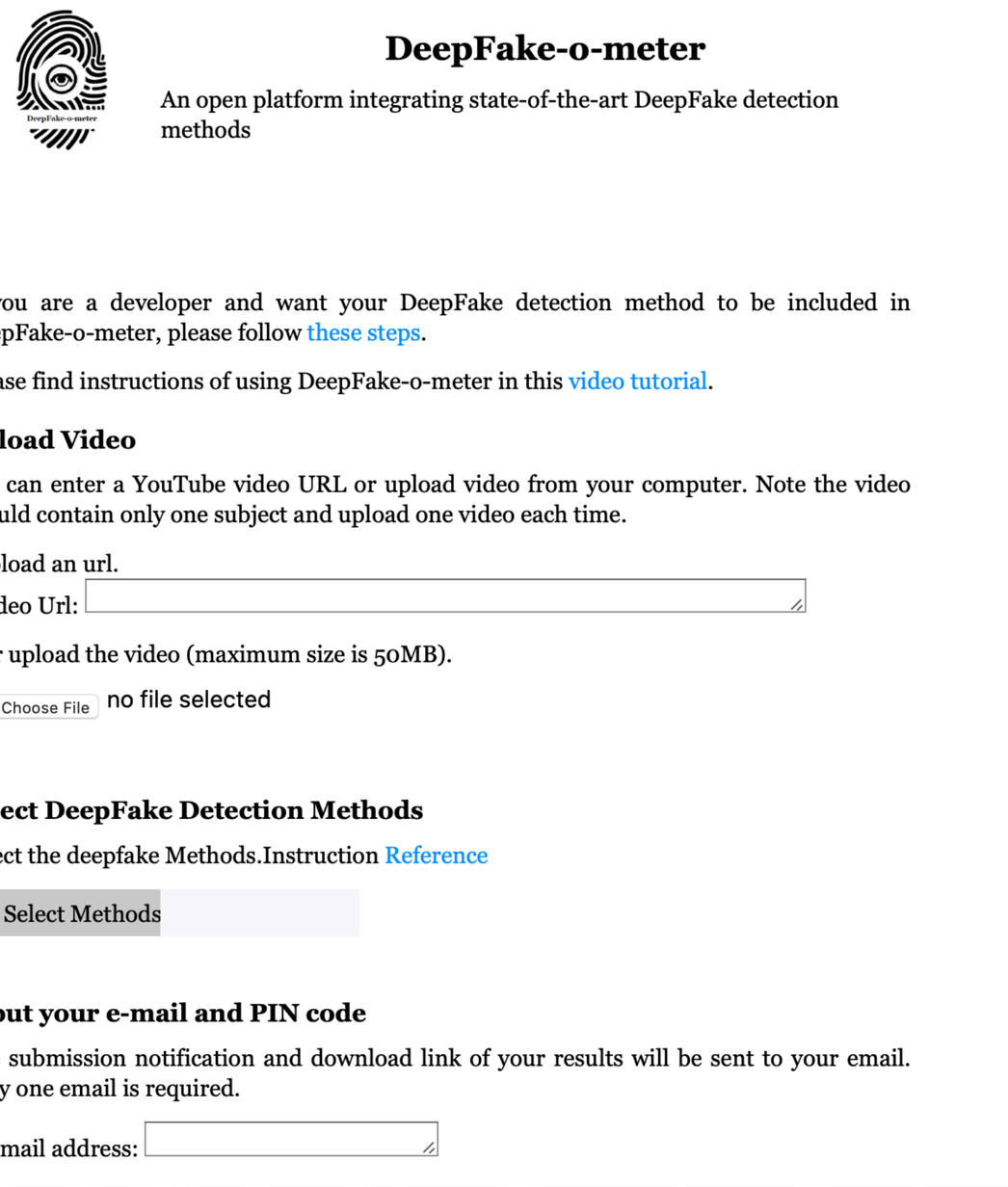


Solution:

- Develop proactive approaches to disrupt training of the AI synthesis models,
- Develop active approaches to add traceable marks to synthesized media,
- Develop approaches to attribute the specific model used to generate a DeepFake,
- Develop approaches to defend detectors from anti-forensics attacks.

Solution:

- Collaborate with NYS Assembly on legislations concerning synthesized media.
- Provide DeepFake-o-meter (http://zinc.cse.buffalo.edu/ubmdfl/deep-o-meter), an open-source free online platform to aggregate state-of-the-art DeepFake detection methods for users.
- Design new course on Multimedia Forensics, with a special focus on DeepFake Forensics.
- Encourage women and URM participation in research in the PI's research lab (UB Media Forensics Lab).



DeepFake-o-meter

An open platform integrating state-of-the-art DeepFake detection methods

If you are a developer and want your DeepFake detection method to be included in DeepFake-o-meter, please follow these steps.

Please find instructions of using DeepFake-o-meter in this video tutorial.

Upload Video

you can enter a YouTube video URL or upload video from your computer. Note the video should contain only one subject and upload one video each time.

•Upload an url.

Video Url:

• Or upload the video (maximum size is 50MB).

Choose File no file selected

Select DeepFake Detection Methods

Select the deepfake Methods.Instruction Reference

Select Methods

Input your e-mail and PIN code

The submission notification and download link of your results will be sent to your email. Only one email is required.

• E-mail address: