

Deep into DeepFakes

Draw a line b/w Real & Fake

Pooja Ranjan

Lead Data Scientist, IQM



Contributor, The House of AI



Google
Developer
Groups

Reuters Uses AI To Prototype First Ever Automated Video Reports

Feb 7, 2020



Accessibility

Reporting & Media



Digital De-aging

For the final shot of *Rogue One: A Star Wars Story* (Walt Disney Studios Home Entertainment, 2017), a digitized image of Carrie Fisher was modeled on her appearance from the late 1970s.

David France | Identity protection with deepfakes

Identity Protection

Digital Resurrection

In Conversation with David France

DEEPEFAKERY

satire, human rights, art and journalism in a time of infodemics

Using AI-generated Face Doubles in Documentary: Welcome to Chechnya

Watch later Share

Humanitarian documentaries

Education & Training

A Voice Deepfake Was Used To Scam A CEO Out Of \$243,000

Sep 3, 2019

Deepfake video of Zelenskyy could be 'tip of the iceberg' in info war, experts warn

MARCH 16, 2022 · 8:26 PM ET

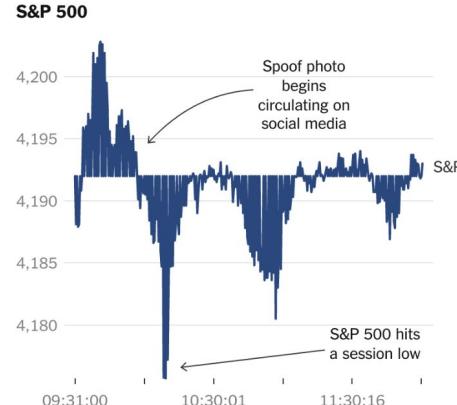


The New York Times

An A.I.-Generated Spoof Rattles the Markets

A stock sell-off driven by a since-debunked picture underscored fears about how artificial intelligence could be used for nefarious purposes with big consequences.

May 23, 2023



Source: Sentieo/AlphaSense • By The New York Times

A Voice Deepfake Was Used To Scam A CEO Out Of \$243,000

Sep 3, 2019

Fraud & Identity Theft

Deepfake video of Zelenskyy could be 'tip of the iceberg' in info war, experts warn

MARCH 16, 2022 · 8:26 PM ET



Political Stability

Privacy & Reputation

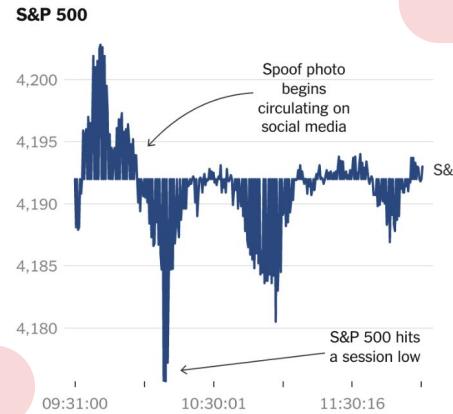
The New York Times

An A.I.-Generated Spoof Rattles the Markets

A stock sell-off driven by a since-debunked picture underscored fears about how artificial intelligence could be used for nefarious purposes with big consequences.

May 23, 2023

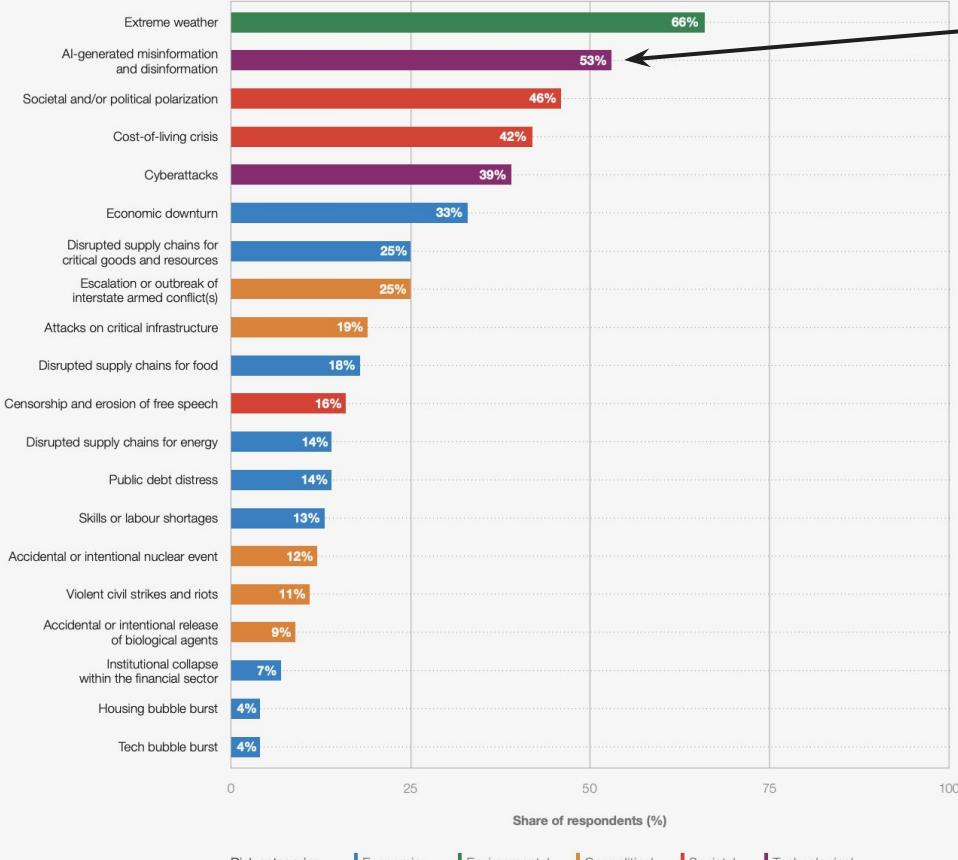
Misinformation



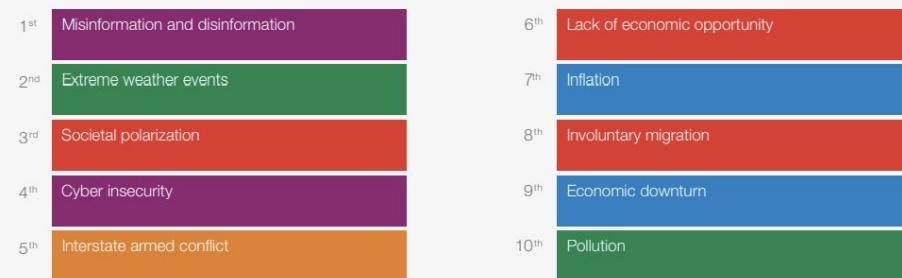
National Security

Source: Sentieo/AlphaSense • By The New York Times

FIGURE 1.2

Current risk landscape*"Please select up to five risks that you believe are most likely to present a material crisis on a global scale in 2024."*

Misinformation & disinformation

Global risks ranked by severity over the short term (2 years)*"Please estimate the likely impact (severity) of the following risks over a 2-year period."*

Risk categories | Economic | Environmental | Geopolitical | Societal | Technological

Source

World Economic Forum Global Risks Perception Survey 2023-2024.

Statement on AI Risk



Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.

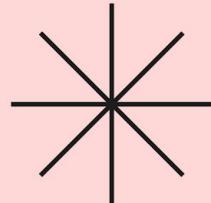
- Geoffrey Hinton**
Emeritus Professor of Computer Science, University of Toronto
- Yoshua Bengio**
Professor of Computer Science, U. Montreal / Mila
- Demis Hassabis**
CEO, Google DeepMind
- Sam Altman**
CEO, OpenAI
- Dario Amodei**
CEO, Anthropic
- Dawn Song**
Professor of Computer Science, UC Berkeley

Jurisdiction	Bill Number	Bill Title	Bill Status	Bill Summary
California	A 2655	Defending Democracy from Deepfake Deception Act of 2024	Enacted	Establishes the Defending Democracy from Deepfake Deception Act for the purpose of preventing the online dissemination of manipulated media and disinformation meant to deceive voters and to prevent voting. Requires a large online platform to block the posting or sending of materially deceptive and digitally modified or created content related to elections, during certain periods before and after an election. Requires a large online platform to develop procedures for state residents to report content.
Florida	H 919	Artificial Intelligence Use in Political Advertising	Enacted	Requires certain political advertisements, electioneering communications, or other miscellaneous advertisements to include a specified disclaimer; specifies requirements for the disclaimer; provides for criminal and civil penalties; authorizes any person to file certain complaints; provides for expedited hearings.
Minnesota	H 1370	Public Safety	Enacted	Relates to public safety; establishes a cause of action for nonconsensual dissemination of deepfake sexual images; establishes the crime of using deepfake technology to influence an election; establishes a crime for nonconsensual dissemination of deepfake sexual images; proposes coding for new law in State Statutes; provides a definition of deepfake.
New Hampshire	H 1432	Deepfakes	Enacted	Establishes the crime of fraudulent use of deepfakes and sets penalties therefor; establishes a cause of action for fraudulent use of deepfakes; prohibits registration of lobbyists who have been found to have fraudulently used deepfakes in certain cases; provides that a person is guilty of a Class B felony if the person knowingly creates, distributes, or presents any likeness in video, audio, or any other media of an identifiable individual that constitutes a deepfake for certain purposes.



About Me!

- Married to Data & AI for 11+ years
- Data Demystifier in Media & Ad Tech
- Love to share about AdTech, Data & AI
- Love to travel



Which face is **Real**?



<https://whichfaceisreal.com/>

Which face is **Real**?



<https://whichfaceisreal.com/>

StyleGAN

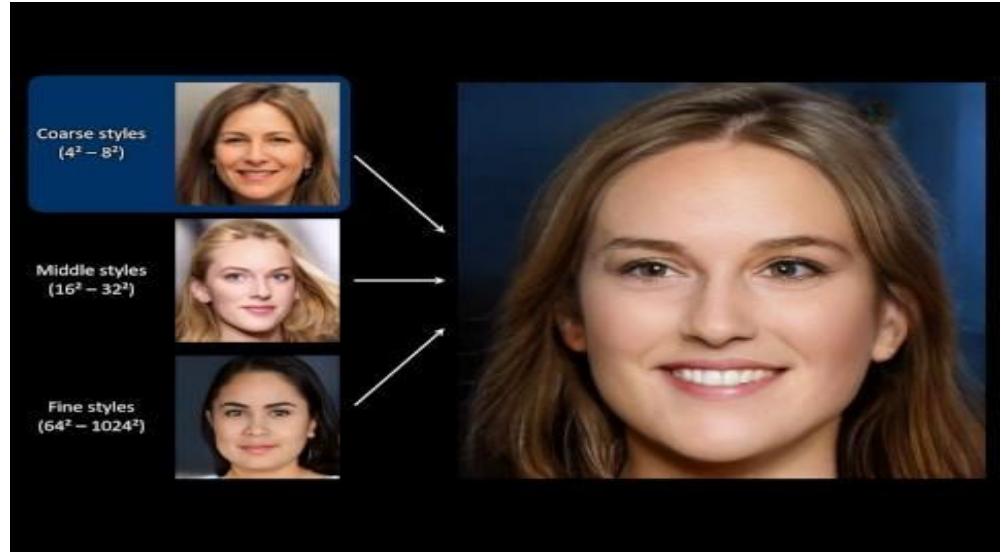
A Style-Based Generator Architecture for Generative Adversarial Networks

StyleGAN — Official TensorFlow Implementation

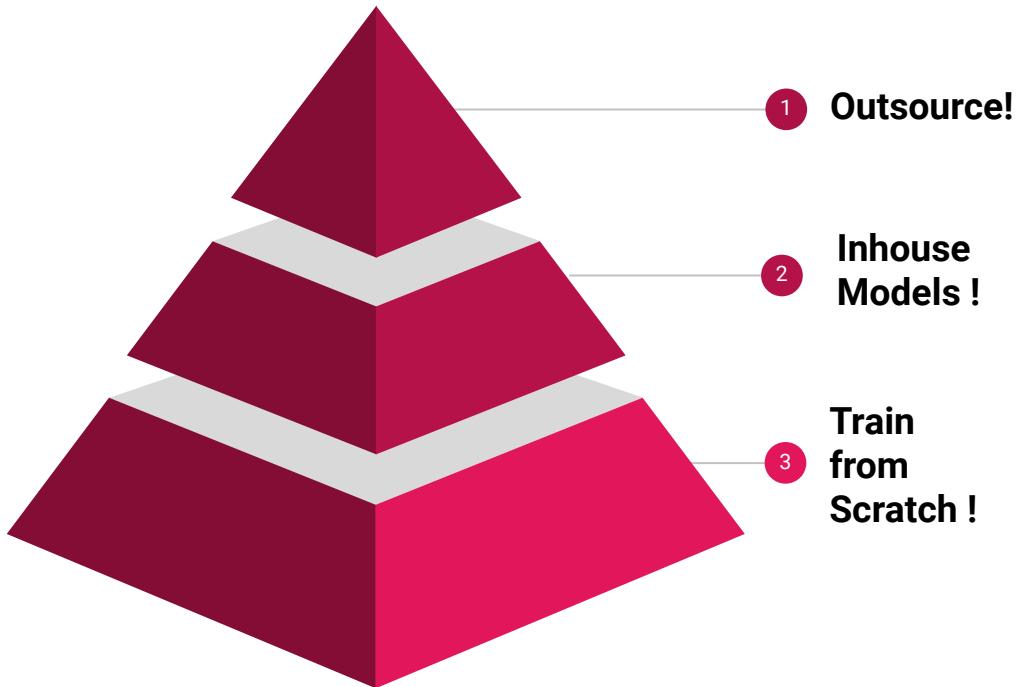
python 3.6 tensorflow 1.10 cudnn 7.3.1 license CC BY-NC



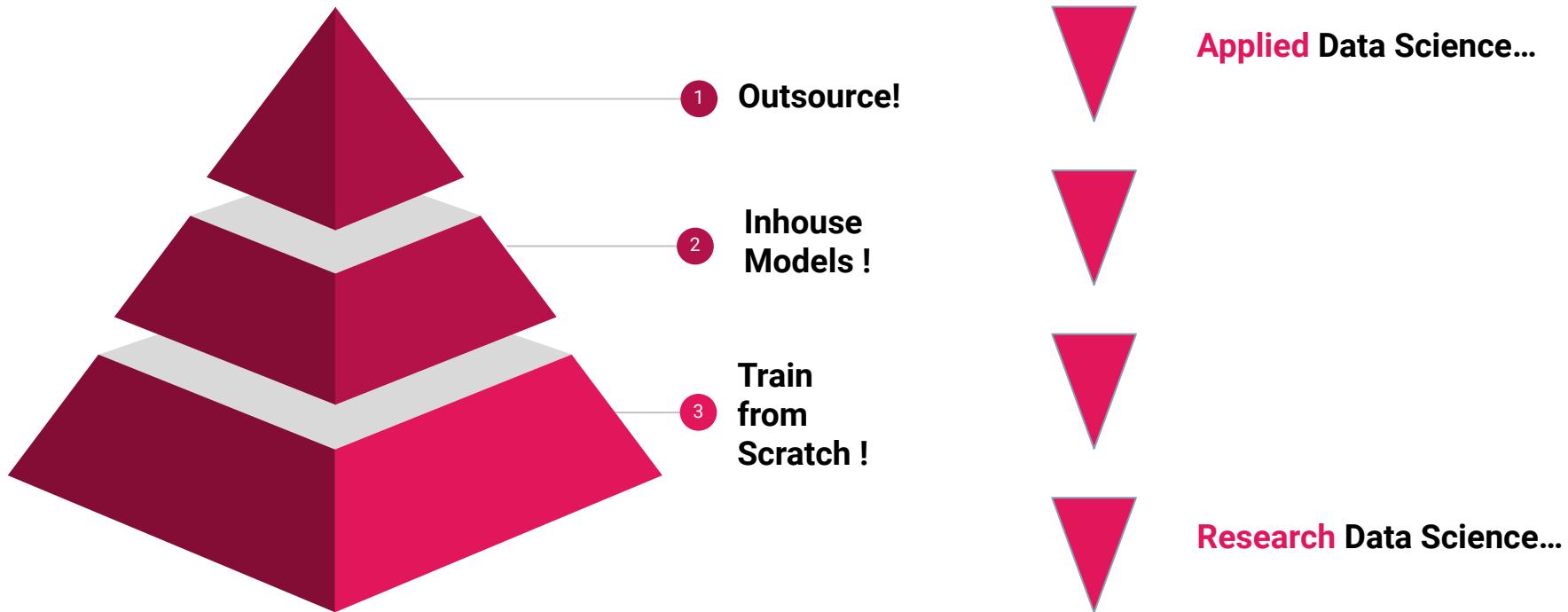
Picture: These people are not real – they were produced by our generator that allows control over different aspects of the image.



DeepFake Detection in Industries



DeepFake Detection in Industries





Outsource!

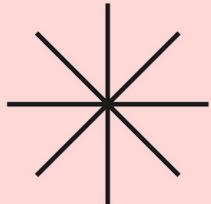
Evaluation

Low Compute

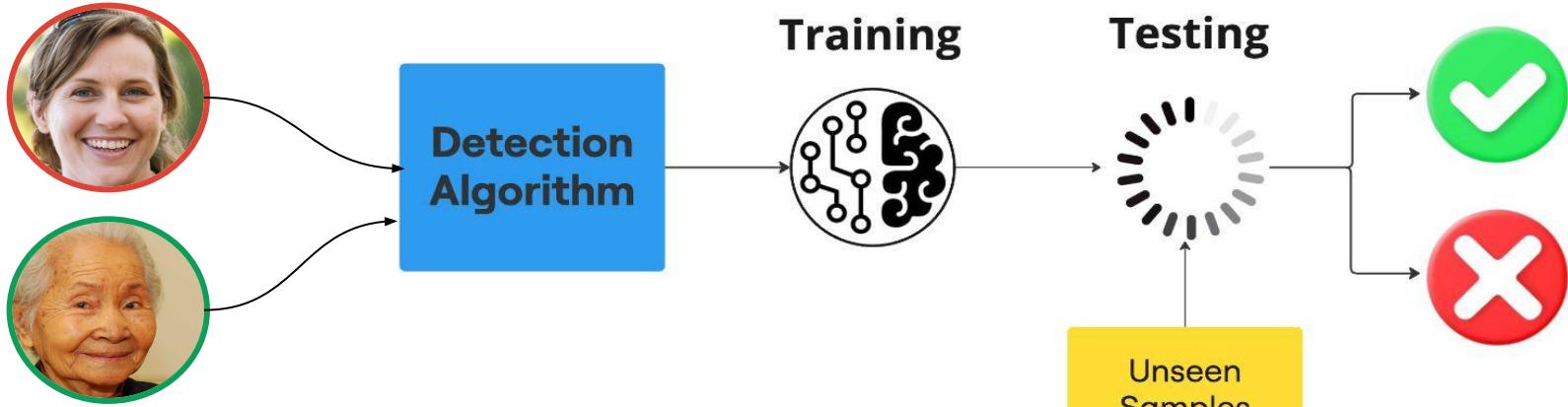
Low Storage



Google
Developer
Groups



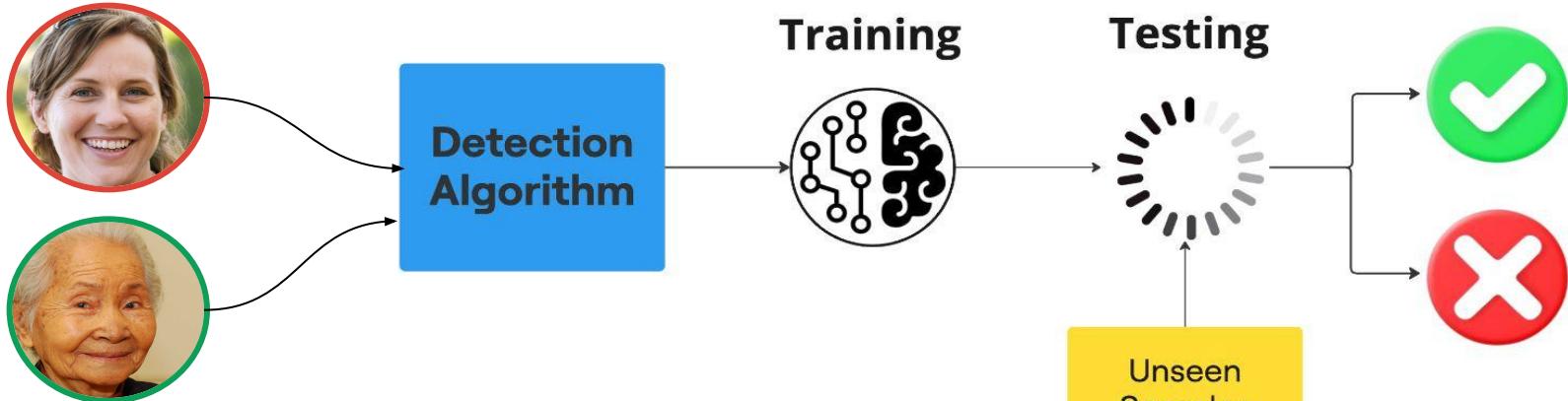
Binary classification



Supervised learning

Classification

Binary classification



Supervised learning

Classification

Data privacy

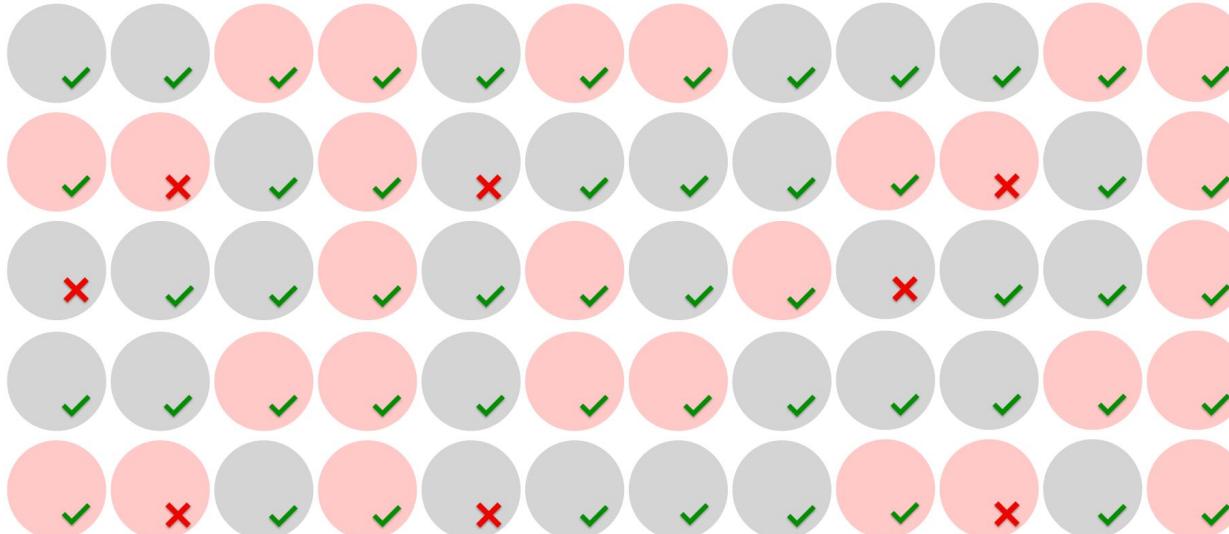
Evaluation

Evaluation - Unskewed

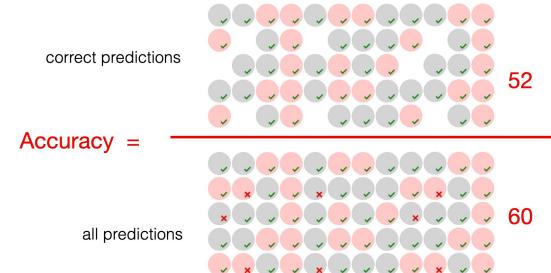


predicted Fake

predicted Real



Total	60
Fake	28
Real	32



Accuracy = 86%

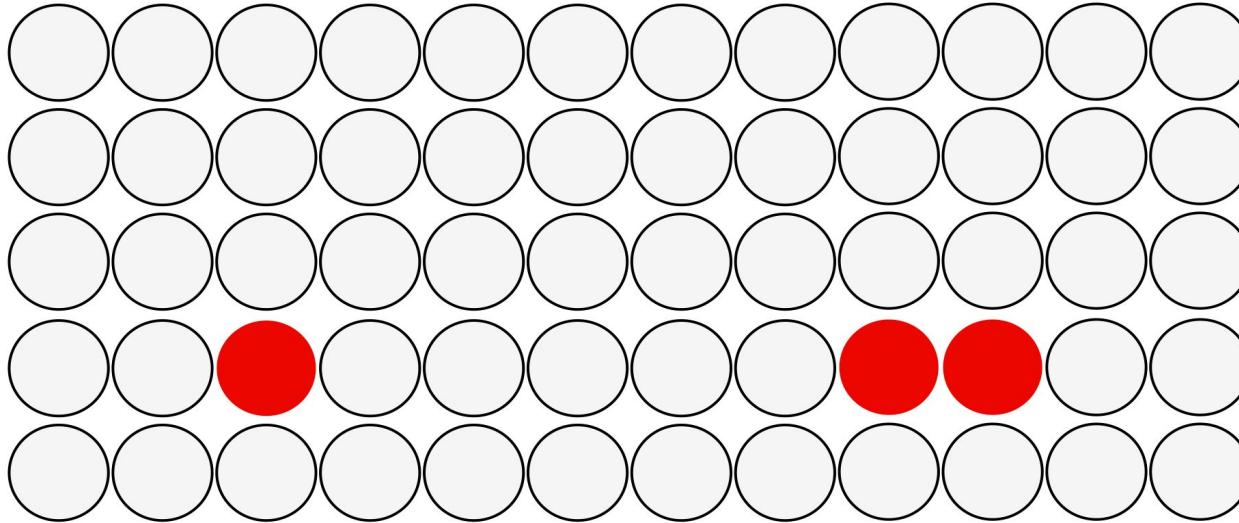
Evaluation- Skewed



actual Fake

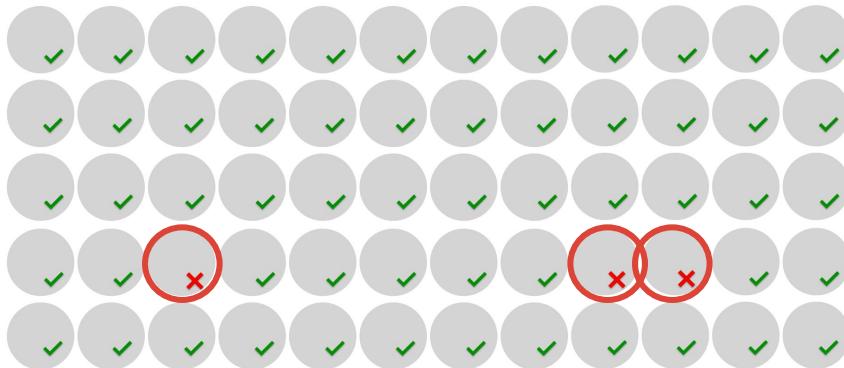


actual Real



Total	60
Fake	3
Real	57

 predicted Fake  predicted Real



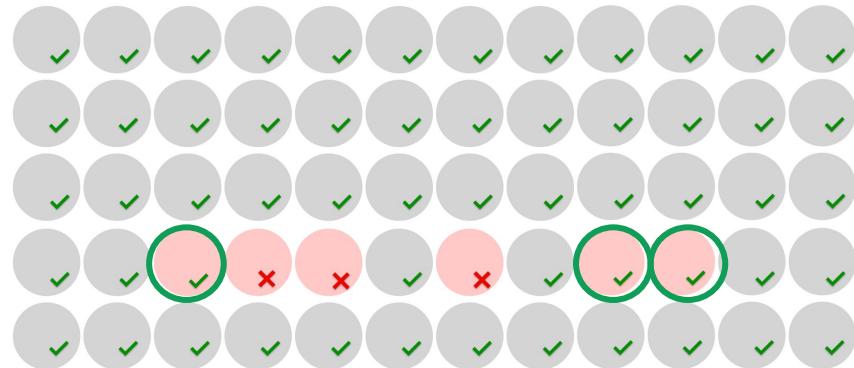
Everything Predicted as
Real

Predictions

		Fake	Real
Actual	Fake	0	3
	Real	0	57

Accuracy = 95%

 predicted Fake  predicted Real



All Fake Samples
Predicted as Fake

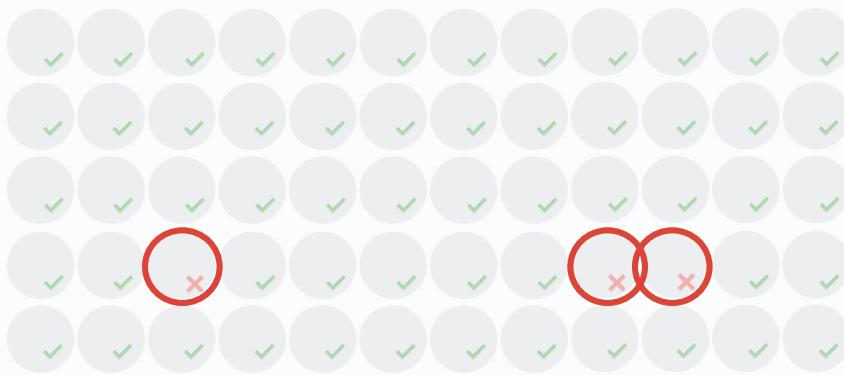
Predictions

		Fake	Real
Actual	Fake	3	0
	Real	3	54

Accuracy = 95%

predicted Fake

predicted Real



Everything Predicted as
Real

Predictions

Fake

Real

Actual

Fake

0

3

Real

0

57

predicted Fake

predicted Real



All Fake Samples
Predicted as Fake

Predictions

Fake

Real

Actual

Fake

3

0

Real

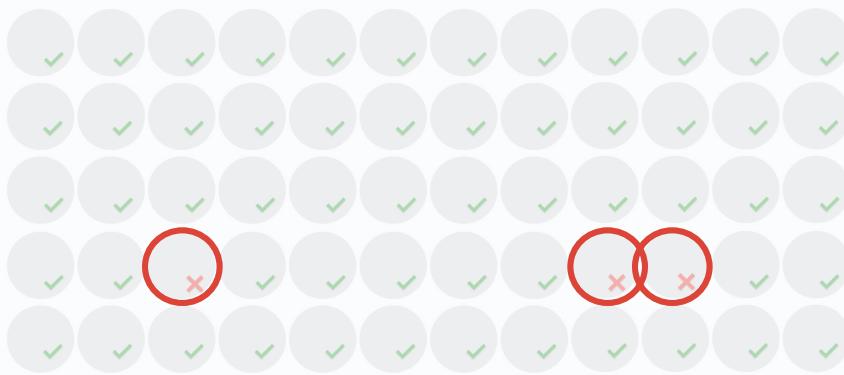
3

54

How many Fakes are
predicted as Fake?

predicted Fake

predicted Real



Everything Predicted as
Real

Predictions

Fake

Real

Actual

Fake

0

3

Real

0

57

Recall = 0%



How many Fakes are
predicted as Fake?

Recall = 100%



predicted Fake

predicted Real



All Fake Samples
Predicted as Fake

Predictions

Fake

Real

Actual

Fake

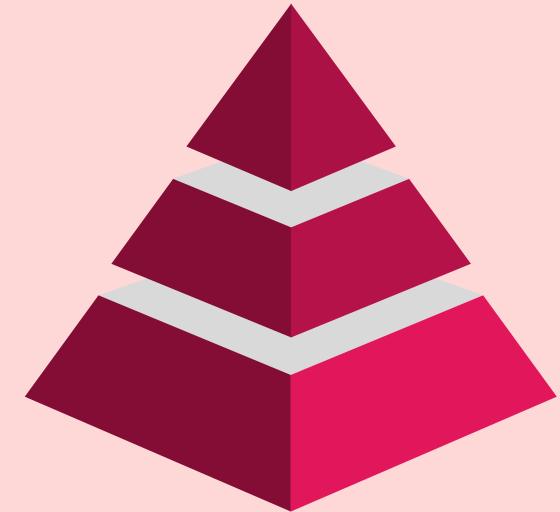
3

0

Real

3

54



Inhouse Models!

Open Source

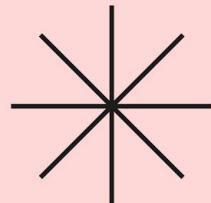
Evaluation

High Compute

Low Storage



Google
Developer
Groups



Open source Detection models

<https://github.com/YZY-stack/DF40>

DF40 - Clip Large

DF40 - Xception

DF40 - i3d

DF40: Toward Next-Generation Deepfake Detection

Zhiyuan Yan¹, Taiping Yao^{2†}, Shen Chen², Yandan Zhao², Xinghe Fu², Junwei Zhu², Donghao Luo², Chengjie Wang², Shouhong Ding², Yunsheng Wu², Li Yuan^{1†}

Oct 2024

<https://github.com/YZY-stack/StA4Deepfake>

Sta4Deepfake

Generalizing Deepfake Video Detection with Plug-and-Play: Video-Level Blending and Spatiotemporal Adapter Tuning

Zhiyuan Yan^{1,2}, Yandan Zhao², Shen Chen², Xinghe Fu², Taiping Yao², Shouhong Ding², Li Yuan¹

Aug 2024

<https://github.com/SCLBD/DeepfakeBench/tree/main>

SRM

Generalizing Face Forgery Detection with High-frequency Features

Yuchen Luo*^{†1,2} Yong Zhang*³ Junchi Yan^{†1,2} Wei Liu^{‡4}

2021 CVPR

Open source Detection models

<https://github.com/YZY-stack/DF40>

DF40 - Clip Large

1.13 GB

DF40 - Xception

83.7 MB

DF40 - i3d

104.2 MB

<https://github.com/YZY-stack/StA4Deepfake>

Sta4Deepfake

856 MB

<https://github.com/SCLBD/DeepfakeBench/tree/main>

SRM

212 MB

 YZY-stack / DF40 Public

DF40: Toward Next-Generation Deepfake Detection

1. Installation

Please run the following script to install the required libraries:

```
sh install.sh
```

2. Download ckpts for inference

All checkpoints/weights of ten models training on our DF40 are released at [Google Drive](#) and [Baidu Disk](#).

Note that:

- If you want to use the CLIP model that is trained on all FS methods of DF40, you can find it at `df40_weights/train_on_fs/clip.pth`. You can use all ckpts under `df40_weights/train_on_xxx_matrix` to reproduce the results of **Protocol-1,2,3** of our paper.
- Similarly, if you want to use the Xception model that is trained **specifically** on the SimSwap method, you can find it in the folder `df40_weights/train_on_fs_matrix/simswap_ff.pth`. You can use all ckpts under `df40_weights/train_on_xxx_matrix` to reproduce the results of **Protocol-4** of our paper.

3. Download DF40 data (after pre-processing)

For quick use and convenience, we provide all DF40 data after pre-processing using in our research. You do **NOT** need to do the pre-processing again but directly use our processed data.

Open source Detection models

<https://github.com/YZY-stack/DF40>

DF40 - Clip Large

1.13 GB

DF40 - Xception

83.7 MB

DF40 - i3d

104.2 MB

<https://github.com/YZY-stack/StA4Deepfake>

Sta4Deepfake

856 MB

<https://github.com/SCLBD/DeepfakeBench/tree/main>

SRM

212 MB

 **YZY-stack / DF40** Public

DF40: Toward Next-Generation Deepfake Detection

2. Download ckpts for inference

All checkpoints/weights of ten models training on our DF40 are released at [Google Drive](#) and [Baidu Disk](#).

Shared with me > df40_weights > train_on_df40 ▾

Name	Owner	Last modified	File size
clip_large.pth	yanzhiyuan1114	Jul 29, 2024	1.13 GB
clip.pth	yanzhiyuan1114	Aug 25, 2024	327.4 MB
i3d.pth	yanzhiyuan1114	Jul 29, 2024	104.2 MB
xception.pth	yanzhiyuan1114	Aug 25, 2024	83.7 MB

3. Run inference

You can then run inference using the trained weights used in our research.

Example-1: If you want to use the Xception model trained on SimSwap (FF) and test it on BlendFace (FF), run the following line.

```
cd DeepfakeBench_DF40
python training/test.py \
--detector_path training/config/detector/xception.yaml \
--weights_path training/df40_weights/train_on_fs_matrix/simswap_ff.pth \
--test_dataset blendface_ff
```

Open source Detection models

<https://github.com/YZY-stack/DF40>

DF40 - Clip Large

1.13 GB

DF40 - Xception

83.7 MB

DF40 - i3d

104.2 MB

<https://github.com/YZY-stack/StA4Deepfake>

Sta4Deepfake

856 MB

<https://github.com/SCLBD/DeepfakeBench/tree/main>

SRM

212 MB

YZY-stack / **StA4Deepfake** Public

Generalizing Deepfake Video Detection with Plug-and-Play: Video-Level Blending and Spatiotemporal Adapter Tuning

1. Installation

Please run the following script to install the required libraries:

```
sh install.sh
```

2. Download ckpts for inference

The pretrained weights of "CLIP-I14 + our StA" training on FaceForensics++ are released at [Google Drive](#).

3. Run demo

You can then infer **one video** (we provide an example video "simswap_000_003.mp4" for illustration) using the pretrained weights.

Specifically, run the following line:

```
python training/demo.py
```

After running the above line, you can obtain all prediction results (fake probabilities) for each clip (clip size is 8).

By default, all frames within the video are used for inference. The frames are organized into clips of 8 frames each, with no overlap between clips. You can try other computation methods for this.

Open source Detection models

<https://github.com/YZY-stack/DF40>

DF40 - Clip Large

1.13 GB

DF40 - Xception

83.7 MB

DF40 - i3d

104.2 MB

<https://github.com/YZY-stack/StA4Deepfake>

Sta4Deepfake

856 MB

<https://github.com/SCLBD/DeepfakeBench/tree/main>

SRM

212 MB

 [SCLBD / DeepfakeBench](#) Public

DeepfakeBench: A Comprehensive Benchmark of Deepfake Detection (NeurIPS 2023 D&B)

★ Detectors (34 detectors):

- 5 Naive Detectors: [Xception](#), [MesoNet](#), [MesoInception](#), [CNN-Aug](#), [EfficientNet-B4](#)
- 19 Spatial Detectors: [Capsule](#), [DSP-FWA](#), [Face X-ray](#), [FFD](#), [CORE](#), [RECCE](#), [UCF](#), [Local-relation](#), [IID](#), [RFM](#), [SIA](#), [SLADD](#), [UIA-VIT](#), [CLIP](#), [SBI](#), [PCL-I2G](#), [Multi-Attention](#), [LSDA](#)
- 3 Frequency Detectors: [F3Net](#), [SPSL](#), [SRM](#)
- 7 Video Detectors: [TALL](#), [I3D](#), [STIL](#), [FTCN](#), [X-CLIP](#), [TimeTransformer](#), [VideoMAE](#)

Release the pre-trained weights for all detectors

For each detector, we release their best-performing checkpoints on the training set, which can be used for testing and evaluating on other datasets.

▼ Assets 15

capsule_best.pth	15 MB	Nov 25, 2023
cnnaug_best.pth	81.3 MB	Nov 25, 2023
core_best.pth	83.7 MB	Nov 25, 2023
fffb4d_best.pth	67.7 MB	Nov 25, 2023
f3net_best.pth	86.2 MB	Nov 25, 2023
ffd_best.pth	83.7 MB	Nov 25, 2023
mesoincep_best.pth	123 KB	Nov 25, 2023
meso4_best.pth	115 KB	Nov 25, 2023
recce_best.pth	183 MB	Nov 25, 2023
spsl_best.pth	83.7 MB	Nov 25, 2023
srm_best.pth	212 MB	Nov 25, 2023
ucf_best.pth	179 MB	Nov 25, 2023
xception_best.pth	83.7 MB	Nov 25, 2023
Source code (zip)		Nov 24, 2023
Source code (tar.gz)		Nov 24, 2023

Open source Detection models

<https://github.com/YZY-stack/StA4Deepfake>

DF40 - Clip Large

1.13 GB

DF40 - Xception

83.7 MB

DF40 - i3d

104.2 MB

<https://github.com/YZY-stack/StA4Deepfake>

Sta4Deepfake

856 MB

<https://github.com/SCLBD/DeepfakeBench/tree/main>

SRM

212 MB

 [SCLBD / DeepfakeBench](#) Public

DeepfakeBench: A Comprehensive Benchmark of Deepfake Detection (NeurIPS 2023 D&B)

★ Detectors (34 detectors):

- 5 Naive Detectors: [Xception](#), [MesoNet](#), [MesoInception](#), [CNN-Aug](#), [EfficientNet-B4](#)
- 19 Spatial Detectors: [Capsule](#), [DSP-FWA](#), [Face X-ray](#), [FFD](#), [CORE](#), [RECCE](#), [UCF](#), [Local-relation](#), [IID](#), [RFM](#), [SIA](#), [SLADD](#), [UIA-VIT](#), [CLIP](#), [SBI](#), [PCL-I2G](#), [Multi-Attention](#), [LSDA](#)
- 3 Frequency Detectors: [F3Net](#), [SPSL](#), [SRM](#)
- 7 Video Detectors: [TALL](#), [I3D](#), [STIL](#), [FTCN](#), [X-CLIP](#), [TimeTransformer](#), [VideoMAE](#)

[DeepfakeBench / training / detectors / srm_detector.py](#) 

If you only want to evaluate the detectors to produce the results of the cross-dataset evaluation, you can use the [test.py](#) code for evaluation. Here is an example:

```
python3 training/test.py \
--detector_path ./training/config/detector/xception.yaml \
--test_dataset "Celeb-DF-v1" "Celeb-DF-v2" "DFDCP" \
--weights_path ./training/weights/xception_best.pth
```



Note that we have provided the pre-trained weights for each detector (you can download them from the [link](#)). Make sure to put these weights in the `./training/weights` folder.

Sta4DeepFake runs

856 MB model

The screenshot shows a code editor interface with the following details:

- File Explorer:** On the left, it shows the project structure for "STA4DEEPCODE". The "test_videos" folder contains several video files named "vid_010....", "vid_011....", "vid_012....", "vid_022....", and "vid_023....". Other folders like "output", "training", and "weights" are also listed.
- Code Editor:** The main editor window displays the "demo.py" script. The code implements a function "main()" that processes input videos from a folder and outputs results to a JSON file. It uses the "os.listdir" function to get a list of files in the input folder, then iterates through them to run inference on each video using the "run_inference" function. The results are stored in a dictionary "results" where the key is the file name and the value is a dictionary containing "probability" and "time_taken". If a file format is unsupported, it prints a warning and sets the probability to None.
- Bottom Bar:** The bottom bar includes tabs for "PROBLEMS", "OUTPUT", "DEBUG CONSOLE", "TERMINAL", "PORTS", and "COMMENTS". The "TERMINAL" tab is active, showing a terminal session with the command "(deepfakebench) pruthvirajdodiya@Pruthvirajs-MacBook-Pro Sta4Deepfake %".
- Status Bar:** The status bar at the bottom right shows icons for "zsh", "outline", and other system status indicators.

Evaluation

File	Actual	File size	DF40 -Clip Large	DF40 - Xception	DF40 - i3d	Sta4Deepfake	SRM	Prediction
Video 1	Real	2 MB	97%	27%	20%	15%	32%	Real
Video 2	Real	1 MB	27%	81%	99%	3%	39%	Real
Video 3	Fake	1 MB	78%	99%	82%	99%	72%	Fake
Video 4	Fake	1 MB	94%	99%	100%	100%	72%	Fake
Video 5	Fake	15 MB	71%	75%	60%	73%	60%	Fake

Fakeness probability

Threshold > 70%

Evaluation

Fakeness probability

File	Actual	File size	DF40 -Clip Large	DF40 - Xception	DF40 - i3d	Sta4Deepfake	SRM
Video 1	Real	2 MB	97%	27%	20%	15%	32%
Video 2	Real	1 MB	27%	81%	99%	3%	39%
Video 3	Fake	1 MB	78%	99%	82%	99%	72%
Video 4	Fake	1 MB	94%	99%	100%	100%	72%
Video 5	Fake	15 MB	71%	75%	60%	73%	60%

Evaluation

Fakeness probability

If $P > 70\%$, then Fake

File	Actual	File size	DF40 -Clip Large	DF40 - Xception	DF40 - i3d	Sta4Deepfake	SRM
Video 1	Real	2 MB	97%	27%	20%	15%	32%
Video 2	Real	1 MB	27%	81%	99%	3%	39%
Video 3	Fake	1 MB	78%	99%	82%	99%	72%
Video 4	Fake	1 MB	94%	99%	100%	100%	72%
Video 5	Fake	15 MB	71%	75%	60%	73%	60%

Evaluation

Fakeness probability

If $P > 70\%$, then Fake

Majority Wins!

File	Actual	File size	DF40 -Clip Large	DF40 - Xception	DF40 - i3d	Sta4Deepfake	SRM	Prediction
Video 1	Real	2 MB	97%	27%	20%	15%	32%	Real
Video 2	Real	1 MB	27%	81%	99%	3%	39%	Real
Video 3	Fake	1 MB	78%	99%	82%	99%	72%	Fake
Video 4	Fake	1 MB	94%	99%	100%	100%	72%	Fake
Video 5	Fake	15 MB	71%	75%	60%	73%	60%	Fake

Evaluation

Fakeness probability

If P > 70%, then Fake

Majority Wins!

File	Actual	File size	DF40 -Clip Large	DF40 - Xception	DF40 - i3d	Sta4Deepfake	SRM	Prediction
Video 1	Real	2 MB	97%	27%	20%	15%	32%	Real
Video 2	Real	1 MB	27%	81%	99%	3%	39%	Real
Video 3	Fake	1 MB	78%	99%	82%	99%	72%	Fake
Video 4	Fake	1 MB	94%	99%	100%	100%	72%	Fake
Video 5	Fake	15 MB	71%	75%	60%	73%	60%	Fake
Accuracy			4/5	4/5	3/5	5/5	4/5	5/5
Recall			3/3	3/3	2/3	3/3	2/3	3/3

Evaluation

Fakeness probability

If $P > 70\%$, then Fake

Majority Wins!

File	Actual	File size	DF40 -Clip Large	DF40 - Xception	DF40 - i3d	Sta4Deepfake	SRM	Prediction
Video 1	Real	2 MB	97%	27%	20%	15%	32%	Real
Video 2	Real	1 MB	27%	81%	99%	3%	39%	Real
Video 3	Fake	1 MB	78%	99%	82%	99%	72%	Fake
Video 4	Fake	1 MB	94%	99%	100%	100%	72%	Fake
Video 5	Fake	15 MB	71%	75%	60%	73%	60%	Fake
Accuracy			4/5	4/5	3/5	5/5	4/5	5/5
Recall			3/3	3/3	2/3	3/3	2/3	3/3
			30 sec	15 sec	17 sec	4 min	5 sec	



Train from Scratch!

Dataset

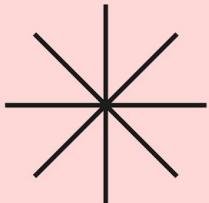
Algorithm

High Compute

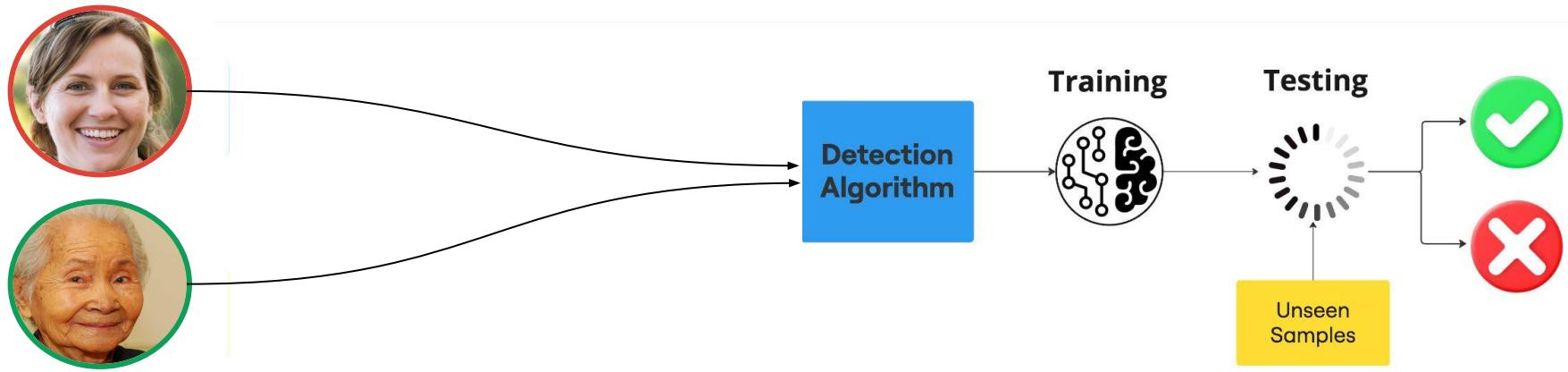
High Storage



Google
Developer
Groups



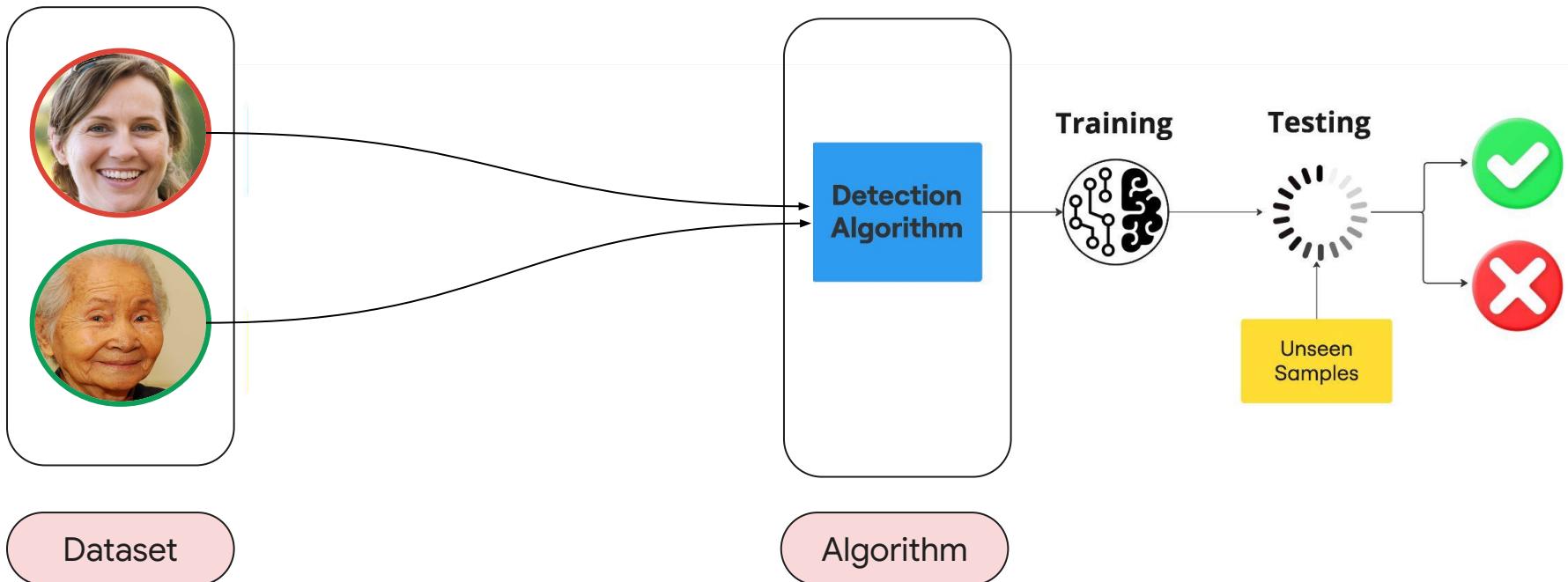
Binary classification



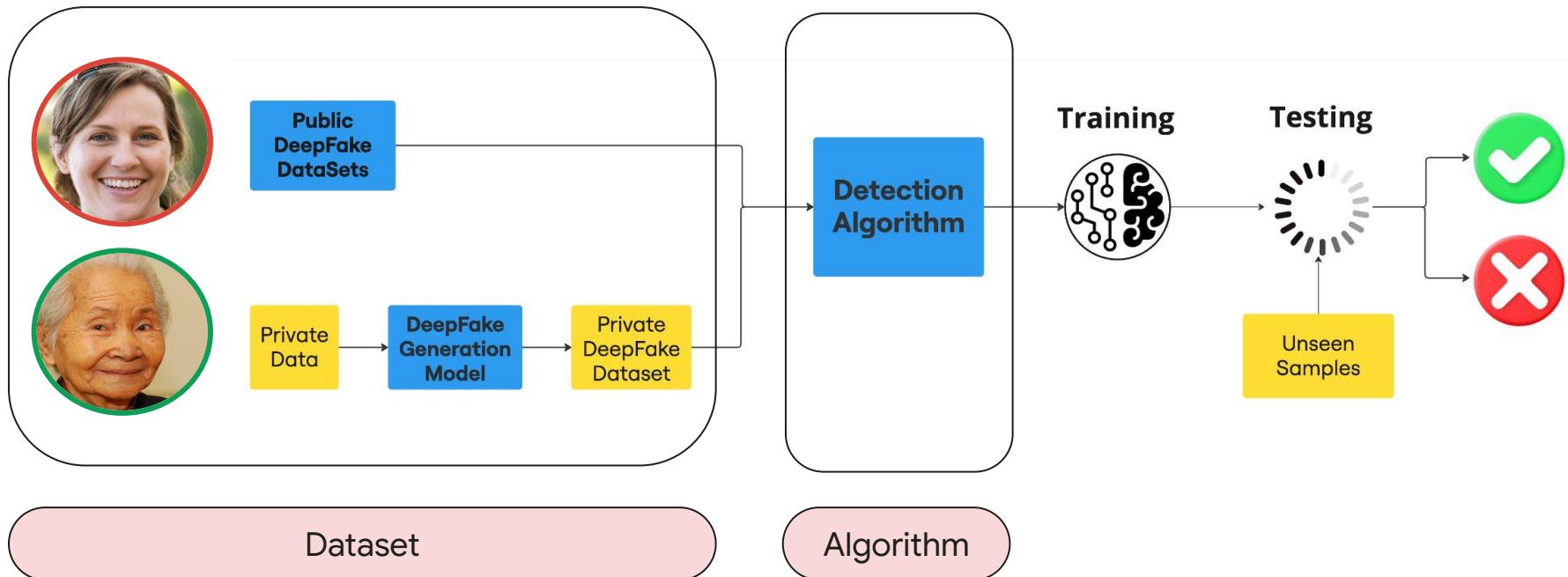
Supervised learning

Classification

Binary classification



Binary classification



Dataset : Ready Public assets

Videos			
Dataset	Real	Fake	Real:Fake
FaceForensics++	1000	4000	1:4
FaceShifter	1000	1000	1:1
DeepfakeDetection	363	3000	1:8
Deepfake Detection Challenge (Preview)	1131	4119	1:4
Deepfake Detection Challenge	23654	104500	1:4
CelebDF-v1	408	795	1:2
CelebDF-v2	590	5639	1:10
DeepForensics-1.0	50000	10000	1:0.2
UADFV	49	49	1:1
DFFD	1000	3000	1:3

Images			
Dataset	Real	Fake	Real:Fake
DFFD	58703	240336	1:4
FFHQ	-	70000	0:1
iFakeFaceDB	-	87000	0:1
ForgeryNet	1438201	1457861	1:1

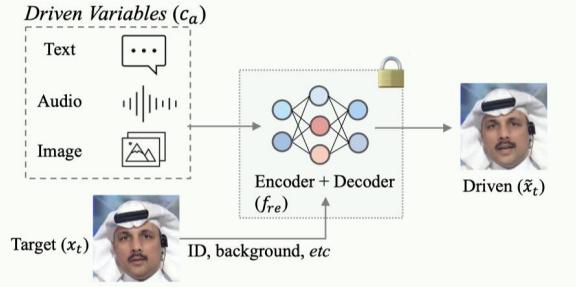
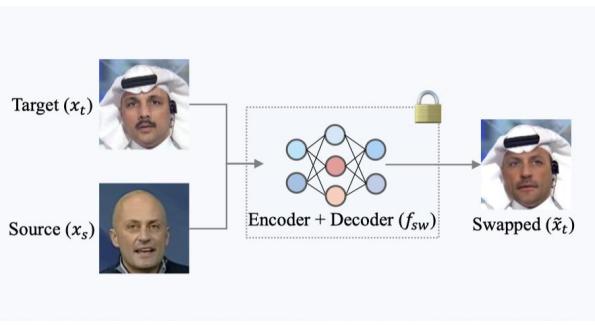
Dataset : Ready Public assets

Videos			
Dataset	Real	Fake	Real:Fake
FaceForensics++	1000	4000	1:4
FaceShifter	1000	1000	1:1
DeepfakeDetection	363	3000	1:8
Deepfake Detection Challenge (Preview)	1131	4119	1:4
Deepfake Detection Challenge	23654	104500	1:4
CelebDF-v1	408	795	1:2
CelebDF-v2	590	5639	1:10
DeepForensics-1.0	50000	10000	1:0.2
UADFV	49	49	1:1
DFFD	1000	3000	1:3

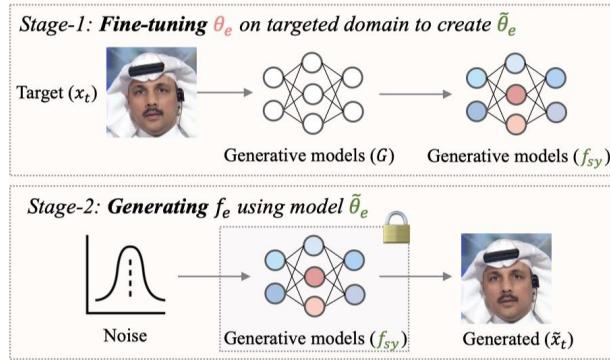
Images			
Dataset	Real	Fake	Real:Fake
DFFD	58703	240336	1:4
FFHQ	-	70000	0:1
iFakeFaceDB	-	87000	0:1
ForgeryNet	1438201	1457861	1:1

Dataset : Generation assets

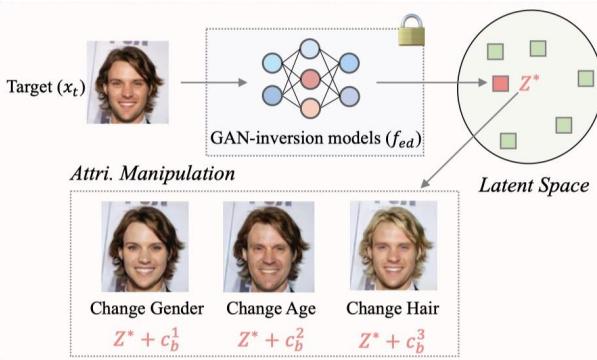
Face Swapping (FS)



Entire Face Synthesis (EFS)



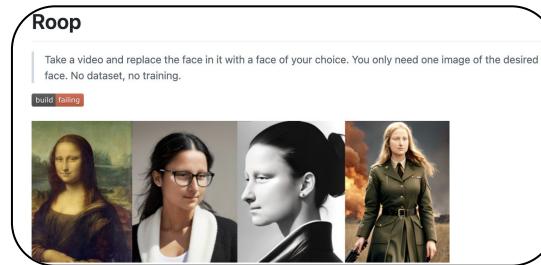
Face Reenactment (FR)



Face Editing (FE)

<https://github.com/s0md3v/roop>

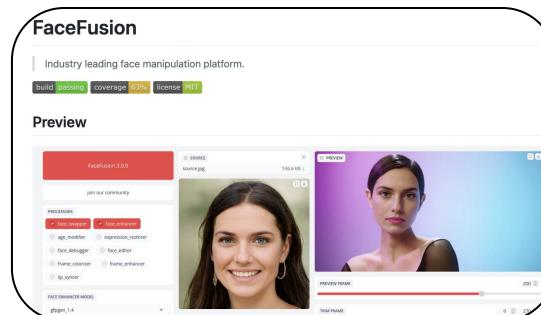
Roop



Face Swapping

<https://github.com/facefusion/facefusion>

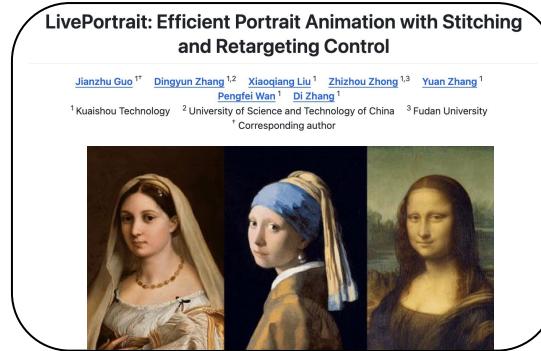
Face Fusion



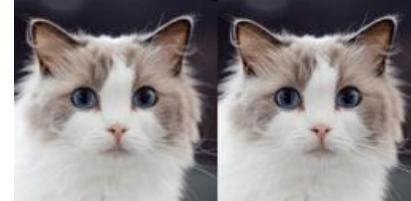
Face Swapping

<https://github.com/KwaiVGI/LivePortrait>

Live Portrait

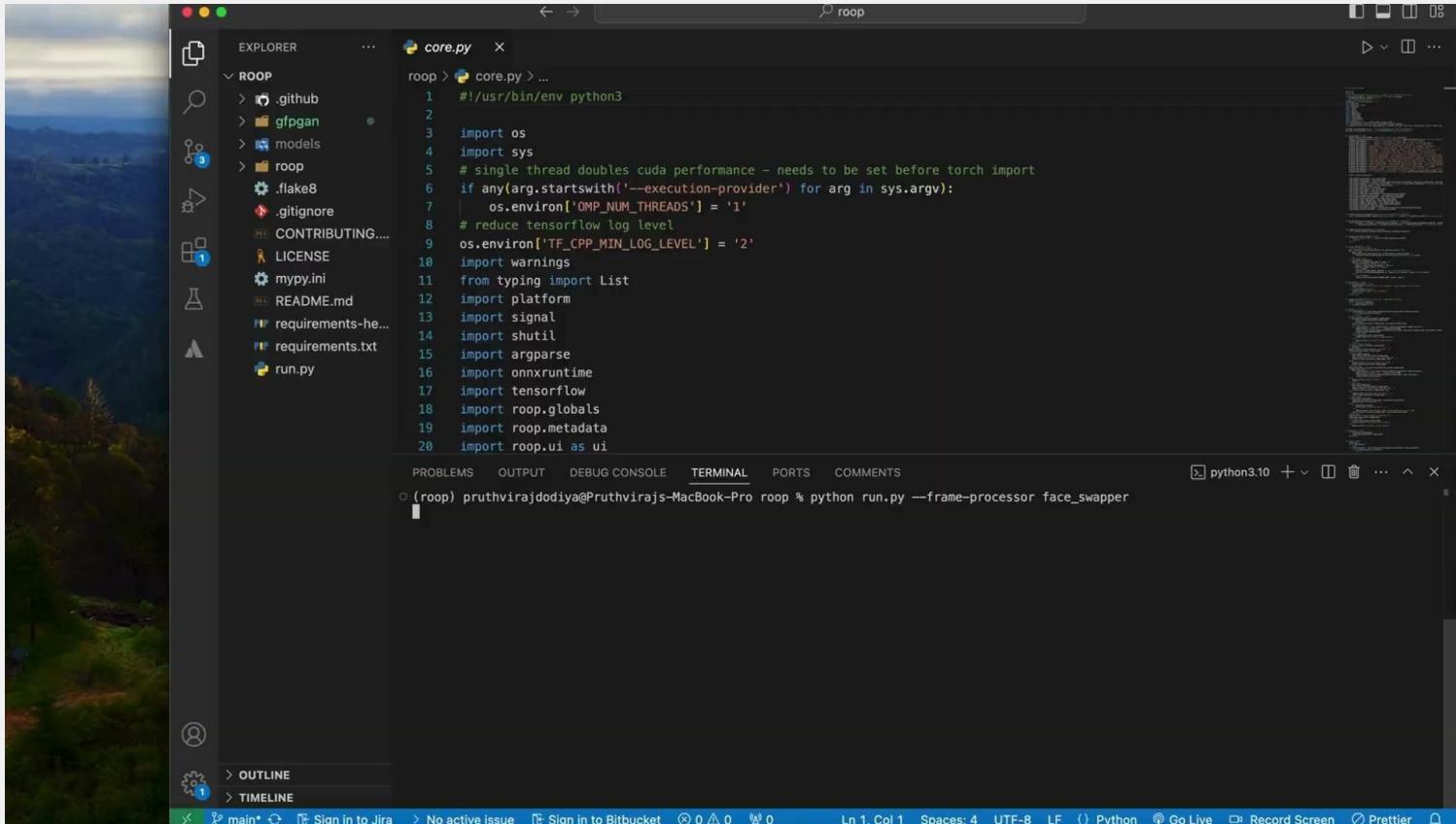


Face Reenactment



Dataset : Generation assets

Roop Demo



The screenshot shows a dark-themed code editor interface, likely Visual Studio Code, displaying a Python project named "roop". The "EXPLORER" view on the left shows the project structure, including files like .github, gfgan, models, roop, .flake8, .gitignore, CONTRIBUTING..., LICENSE, mypy.ini, README.md, requirements-he..., requirements.txt, and run.py. The "core.py" file is open in the main editor area, showing its contents:

```
#!/usr/bin/env python3
import os
import sys
# single thread doubles cuda performance - needs to be set before torch import
if any(arg.startswith('--execution-provider') for arg in sys.argv):
    os.environ['OMP_NUM_THREADS'] = '1'
# reduce tensorflow log level
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '2'
import warnings
from typing import List
import platform
import signal
import shutil
import argparse
import onnxruntime
import tensorflow
import roop.globals
import roop.metadata
import roop.ui as ui
```

The "TERMINAL" tab at the bottom shows the command being run: `(roop) pruthvirajdodiya@Pruthvirajs-MacBook-Pro roop % python run.py --frame-processor face_swapper`. The status bar at the bottom indicates the current file is "main" and the encoding is "UTF-8".

Dataset : Generation assets

Roop Demo



Demo!

Dataset : Generation assets

Roop Demo



Select a face



Select a target

Keep target fps

Keep temporary frames

Skip target audio

Many faces

Start

Destroy

Preview



Dataset : Generation assets

FaceFusion
Demo

The image shows a screenshot of the FaceFusion 3.0.1 software interface. On the left, the 'SOURCE' section displays a portrait of Elon Musk. In the center, the 'PREVIEW' section shows a composite image where Elon Musk's face has been swapped onto the body of Shah Rukh Khan. On the right, the 'TARGET' section shows a portrait of Shah Rukh Khan. The interface includes various configuration options such as processors (face_swapper, age_modifier, expression_restorer, face_debugger, face_editor, face_enhancer, frame_colorizer, frame_enhancer, lip_syncer), execution providers (cpu, coreml), execution thread count, and output image quality.

FaceFusion 3.0.1

SOURCE: elon.jpg (1.3 MB)

PREVIEW: Composite image showing Elon Musk's face swapped onto Shah Rukh Khan's body.

FACE SWAPPER MODEL: inswapper_128_f16

FACE SWAPPER PIXEL BOOST: 128x128

EXECUTION PROVIDERS: CPU (checked), CoreML

EXECUTION THREAD COUNT: 4

EXECUTION QUEUE COUNT: 1

VIDEO MEMORY STRATEGY: Strict

SYSTEM MEMORY LIMIT: 0

OUTPUT IMAGE QUALITY: 80

OUTPUT IMAGE RESOLUTION: 200x252

TARGET: SRK.jpg (7.6 KB)

FACE SELECTOR MODE: reference

REFERENCE FACE: Image of Shah Rukh Khan

FACE SELECTOR ORDER: large-small

FACE SELECTOR GENDER: none

FACE SELECTOR RACE: none

FACE SELECTOR AGE: 0 - 100

FaceFusion 3.0.1

SOURCE: elon.jpg (1.3 MB)

PREVIEW: Composite image showing Shah Rukh Khan's face swapped onto an older man's body.

FACE SWAPPER MODEL: inswapper_128_f16

FACE SWAPPER PIXEL BOOST: 128x128

EXECUTION PROVIDERS: CPU (checked), CoreML

EXECUTION THREAD COUNT: 4

EXECUTION QUEUE COUNT: 1

VIDEO MEMORY STRATEGY: Strict

SYSTEM MEMORY LIMIT: 0

OUTPUT IMAGE QUALITY: 80

OUTPUT IMAGE RESOLUTION: 480x640

TARGET: ai-generated-8527293_640.webp (273.4 KB)

FACE SELECTOR MODE: reference

REFERENCE FACE: Image of an older man

FACE SELECTOR ORDER: large-small

FACE SELECTOR GENDER: none

FACE SELECTOR RACE: none

Algorithm : Detection

FFAA : Face Forgery Analysis Assistant



USER:



USER: A human face is visible in this image. Is it genuine or a forgery? On what grounds do you decide?

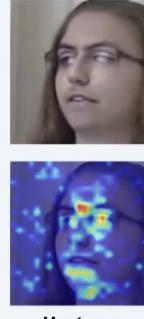
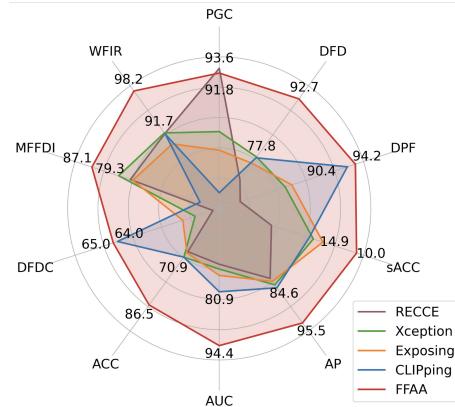


Image description: The image shows a person with long hair, wearing glasses. The face is oriented slightly to the side, with the eyes looking in different directions. The lighting appears to be even, and the resolution is relatively low, making some details less clear. The person appears to be in their twenties or thirties, with a light skin tone. The expression is neutral, with no significant facial decorations or accessories.

Forgery reasoning: Upon close examination, there are several abnormalities in the facial features. **The eyes appear to be misaligned, with one eye looking in a different direction than the other**, which is unusual for a natural face. The skin texture around the eyes and mouth seems inconsistent, with some areas appearing smoother and less detailed than others. **The glasses also seem to be poorly integrated with the face, with some unnatural blending around the edges**. These inconsistencies suggest that the face has been manipulated. **The overall integration of the face with the hair and surrounding environment appears unnatural**, further indicating that the face has been altered.

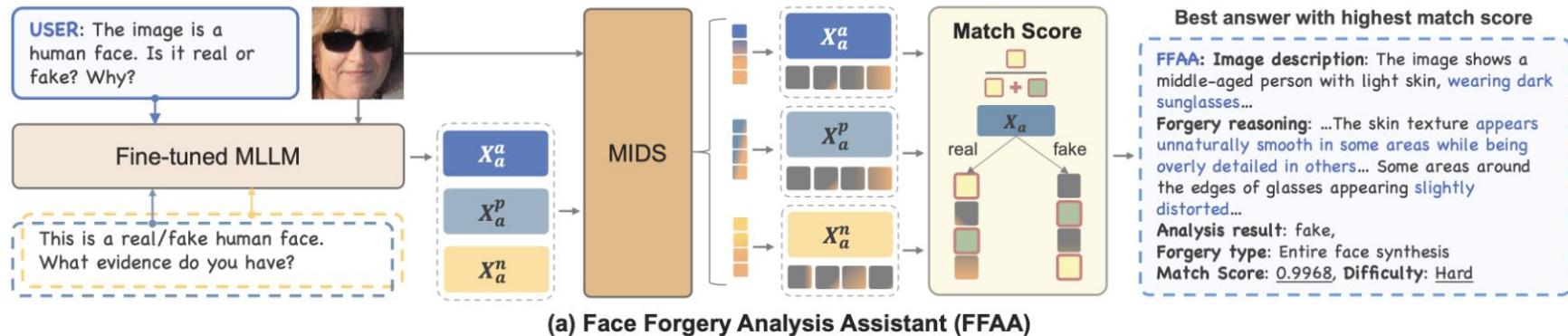
Analysis result: fake, **Forgery type:** Identity exchange

Match score: 0.9877; **Difficulty:** hard



Algorithm : Detection

FFAA : Face Forgery Analysis Assistant



MLLM : Multimodal Large Language model

MIDS : Multi-answer Intelligent Decision System

You are an AI master specialized in face forgery analysis. You'll be provided with a human face image. You need to think step by step to determine whether the face is real or fake.

Below are the three types of face forgery:

1. Entire face synthesis: Generate an entire non-existent face through some technologies.
2. Identity exchange: Use existing videos or images to replace the source character's face with the target character's face, thereby changing the source character's identity.
3. Facial attribute manipulation: This involves altering facial expressions or features without changing the identity, such as skin color, hair color, gender, age, decorations, etc.

Below are the format and requirements of answers:

Image description: Describe this image in terms of lighting, resolution, face orientation, gender, age, skin color, expression, face decoration, etc. Answer in one paragraph of 50 to 100 words;

Forgery reasoning: Analyze whether there are any abnormalities compared with real face images and where the abnormalities are. Let's think step by step. First, check whether there are any abnormalities in the facial organs, such as being too smooth, blurred, pixelation, or uncoordinated. Then check whether there are any fusion artifacts, inconsistent skin color and other issues that often occur at forged boundaries. Next check whether the various features of the face have been tampered with or added in terms of attributes, such as color, texture, naturalness, consistency, expressions, etc. Finally, check whether the entire face is normally integrated with the hair, decorations, and surrounding environment. In addition, you can provide other useful clues. Answer in one paragraph of 100 to 150 words;

Analysis result: real | fake

Probability: a floating-point number in (0.0, 1.0)

Forgery type: Entire face synthesis | Identity exchange | Facial attribute manipulation | None

Below are requirements for answers in the conversation:

1. Strictly follow the format and requirements of answers, and do not say anything else irrelevant.
2. Avoid giving general explanations. Focus on and analyze the specific image. Then make a reasonable and specific explanation.
3. Do not bold the output content.

- **Real:** [Image] The image is a real face. Please analyze the reason and answer the question. [Question prompt].

- **Identity exchange:** [Image] The image is a fake face whose forgery type is identity exchange. Please analyze the reason and answer the question. [Question prompt]. [Reference prompt].

- **Facial attribute manipulation:** [Image] The image is a fake face whose forgery type is identity exchange. Please analyze the reason and answer the question. [Question prompt]. [Reference prompt].

- **Entire face synthesis:** The image is a fake face whose forgery type is entire face synthesis. Please analyze the reason and answer the question. [Question prompt]

- "The image is a human face image. Is it real or fake? Why?"
- "The photo depicts a human face. Can you tell if it is authentic or manipulated? Please explain your reasoning."
- "This picture shows a human face. Do you think it is genuine or counterfeit? What are your reasons?"
- "We see a human face in this image. Is it an actual photo or a fabrication? What leads you to that conclusion?"
- "There is a human face in this image. Is it original or has it been altered? Provide your justification."
- "The picture contains a human face. Is it factual or doctored? What makes you say that?"
- "The visual shows a human face. Is this a real image or a fake one? Could you explain your assessment?"
- "In this photograph, a human face is shown. Is it natural or manipulated? What leads you to this judgement?"
- "A human face is visible in this image. Is it genuine or a forgery? On what grounds do you decide?"
- "The image features a human face. Is it authentic or fabricated? What evidence do you have?"

- **Mask:** [Mask Image] The mask shows the actually modified areas on the face which you should pay more attention to. The mask can help you analyze, but you cannot mention it in your answer.

- **Original Image:** [Original Image] The image shows the original, unaltered version before manipulation. It can assist in your analysis, but you cannot mention it in your answer.

System Prompt

Data source specific prompt

Question prompt

Reference prompt

Algorithm : Detection

FFAA : Face Forgery Analysis Assistant



USER: The image is a human face image. Is it real or fake? Why?

Ground truth: **Fake**

LLaVA-v1.6-34b: The image you've provided **appears to be a genuine photograph** of a person wearing sunglasses. The person's facial features, such as the wrinkles, the shape of the nose, and the expression, all seem consistent with natural human features...
(Real)

GPT4o: The image appears to be a close-up of a human face wearing sunglasses... Upon closer examination, I notice the following:
The skin texture appears consistent, but there might be minor inconsistencies that are hard to detect at a glance ...
Given these observations, it is challenging to definitively conclude without more advanced tools. However, from a basic visual inspection, it leans towards appearing real but could still potentially be AI-generated with sophisticated techniques.
(Real)

FFAA: **Image description:** The image shows a middle-aged person with light skin, wearing dark sunglasses. The person has short, wavy hair that appears to be light brown. The face is oriented towards the camera with a neutral expression...
Forgery reasoning: Upon close examination, there are several indicators that suggest the face is synthesized. The skin texture appears unnaturally smooth in some areas while being overly detailed in others, creating an inconsistent appearance. The transition between the skin and hairline is somewhat blurred and lacks the natural integration seen in real images. Additionally, the sunglasses seem to be unnaturally fused with the face, with some areas around the edges appearing slightly distorted...These inconsistencies are common signs of entire face synthesis.
Analysis result: fake, **Probability:** 0.95, **Forgery type:** Entire face synthesis
Match Score: 0.9968, **Difficulty:** Hard
(Fake)



USER: This picture shows a human face. Do you think it's genuine or counterfeit? What are your reasons?

Ground truth: **Fake**

LLaVA-v1.6-34b: The image you've provided **appears to be a close-up of a human face, and it looks like a genuine photograph.** There are no obvious signs of being counterfeit, such as unnatural lighting, inconsistent textures, or other visual cues that might suggest it's a digitally created or manipulated image...
(Real)

GPT4o: The image appears to depict a human face with a neutral or slightly displeased expression. To determine whether the image is genuine or counterfeit, I would consider several factors: Facial Features and Proportions:... Image Quality and Details:... Contextual Clues:... Expression and Emotion:... If you suspect the image might be AI-generated or manipulated, tools and software designed for detecting deepfakes or image manipulation could provide a more definitive answer.
(Refuse to judge)

FFAA: **Image description:** The image shows a close-up of a young person's face with medium to dark skin tone. The individual appears to be wearing glasses and has a neutral or slightly concerned expression. The lighting is natural, and the background is slightly blurred, suggesting an outdoor setting. The resolution is relatively low, and the face is oriented slightly to the left.

Forgery reasoning: Upon close examination, there are several abnormalities in the facial features. The area around the glasses, particularly where they meet the skin, appears unnaturally smooth and lacks the expected texture and detail. The skin tone around the glasses is inconsistent with the rest of the face, indicating possible manipulation. Additionally, the edges of the glasses and the surrounding skin show signs of blending artifacts, which are common in identity exchange forgeries. The overall integration of the face with the hair and background also appears slightly off, with minor inconsistencies in lighting and shading.

Analysis result: fake, **Probability:** 0.95, **Forgery type:** Identity exchange
Match Score: 0.8366, **Difficulty:** Hard
(Fake)

Algorithm : Detection

FFAA : Face Forgery Analysis Assistant

FFAA: Face Forgery Analysis Assistant

FFAA: Multimodal Large Language Model based Explainable Open-World Face Forgery Analysis Assistant

[[Paper](#)] [[Project Page](#)]

[Zhengchao Huang](#), [Bin Xia](#), [Zicheng Lin](#), [Zhun Mou](#), [Wenming Yang](#), [Jiaya Jia](#)

Release

- [2024/09/26] The code and the model weights of FFAA have been made public!
- [2024/08/19] FFAA has been published on Arxiv!

Contents

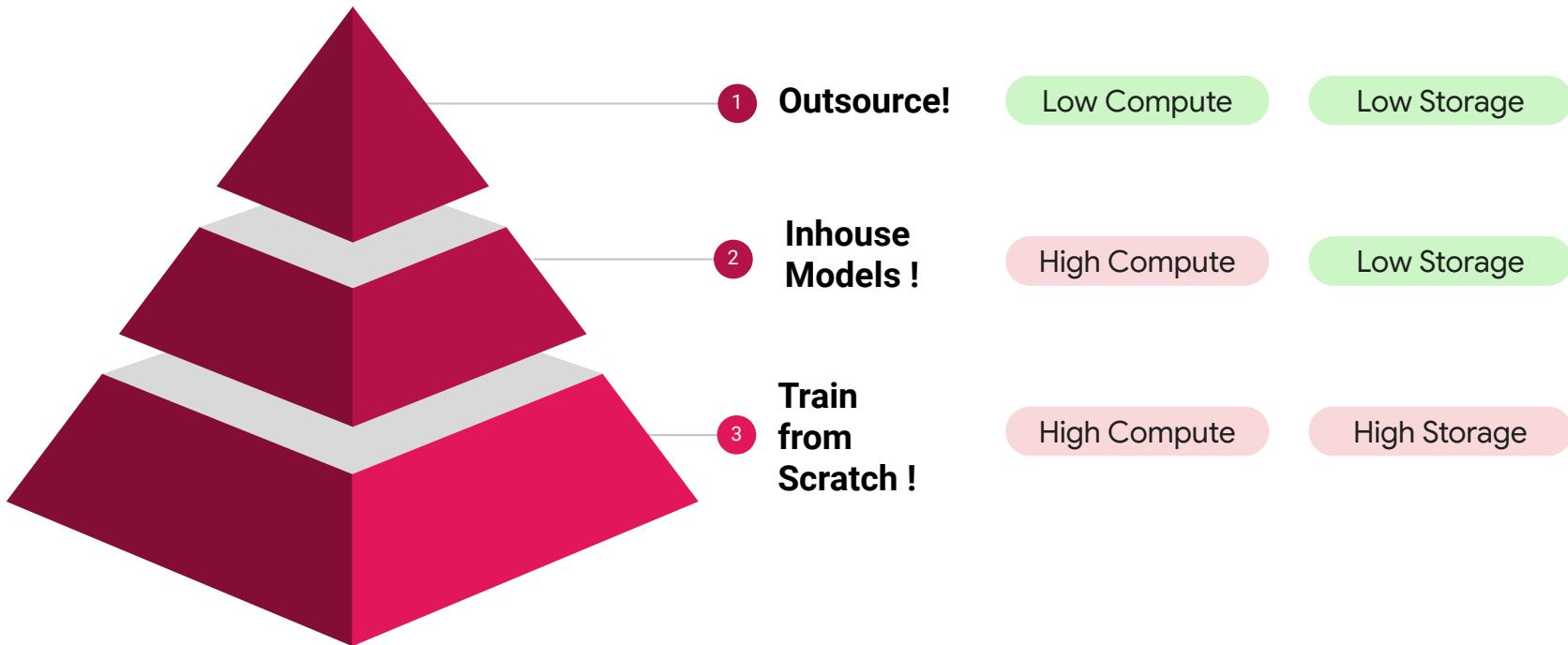
- [Install](#)
- [Model Zoo](#)
- [Dataset](#)
- [Train](#)
- [Evaluation](#)
- [Inference](#)

<https://ffaa-vl.github.io/>

High Compute

High Storage

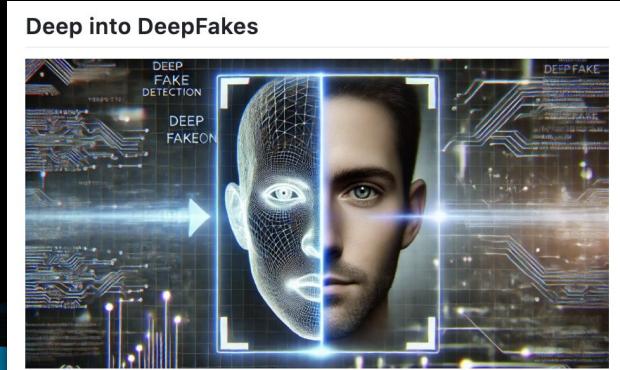
DeepFake Detection in Industries



References



https://github.com/PoojaRanjan01/Deep_into_DeepFakes



Aggregated references for datasets, code and papers for Deepfake detection research.

Talk : Deep into DeepFakes : Drawing a line between Real and Fake

Open source Detection models

- DF40: "DF40: Toward Next-Generation Deepfake Detection", 2024: [Paper](#), [Code](#)
 - StaDeepfake: "Generalizing Deepfake Video Detection with Plug-and-Play: Video-Level Blending and Spatiotemporal Adapter Tuning", 2024: [Paper](#), [Code](#)
 - SRF: "Generalizing Face Forgery Detection with High-frequency Features", 2021: [Paper](#), [Code](#)
 - DeepfakeBench: "DeepfakeBench: A Comprehensive Benchmark of Deepfake Detection", 2023: [Paper](#), [Code](#)

Dataset

Ready Public Assets

Video and Image Datasets : Download Links

Generation Assets

- Roop: [Code](#)
 - Face Fusion: [Code](#)
 - Live Portrait: [Code](#)

Algorithm

- "FFAA: Multimodal Large Language Model based Explainable Open-World Face Forgery Analysis Assistant": [Paper](#) [Code](#)

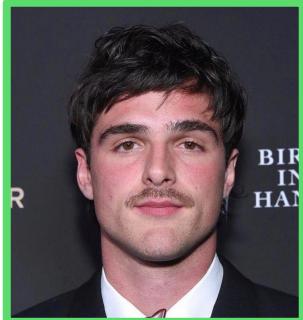
Which face is Real?



Which face is **Real**?



Which face is Real?



Which face is Real?



Pruthviraj Dodiya



Kathan Soni



The House
of
AI!



Talk soon Again...



Let's Start!





Thank You!



<https://www.linkedin.com/in/poojaranjan22>

