



# Deepfake Detection and Annotation

Yuqing Zheng  
308816

# TL;DR (Motivation, Main Idea, Results)

## Motivation

Combat the ethical challenges of misinformation spread through deepfakes.

## Main Idea

Build a system that integrates state-of-the-art deepfake detection models to identify manipulated media.

## Result

Users can upload images or videos and receive a realness score with visual evidence.

## Impact

Support AI ethics by promoting transparency and trust in digital content.

# Motivation & Problem Statement



Deepfakes are highly realistic fake videos/images generated by AI.

They threaten:

## Politics

fake speeches or public statements

## Social Media

spread of misinformation

## Security

fake IDs or evidence

## Ethical Challenge

Need for detection systems to preserve truth and combat misinformation.



# Literature Review: Related Methods

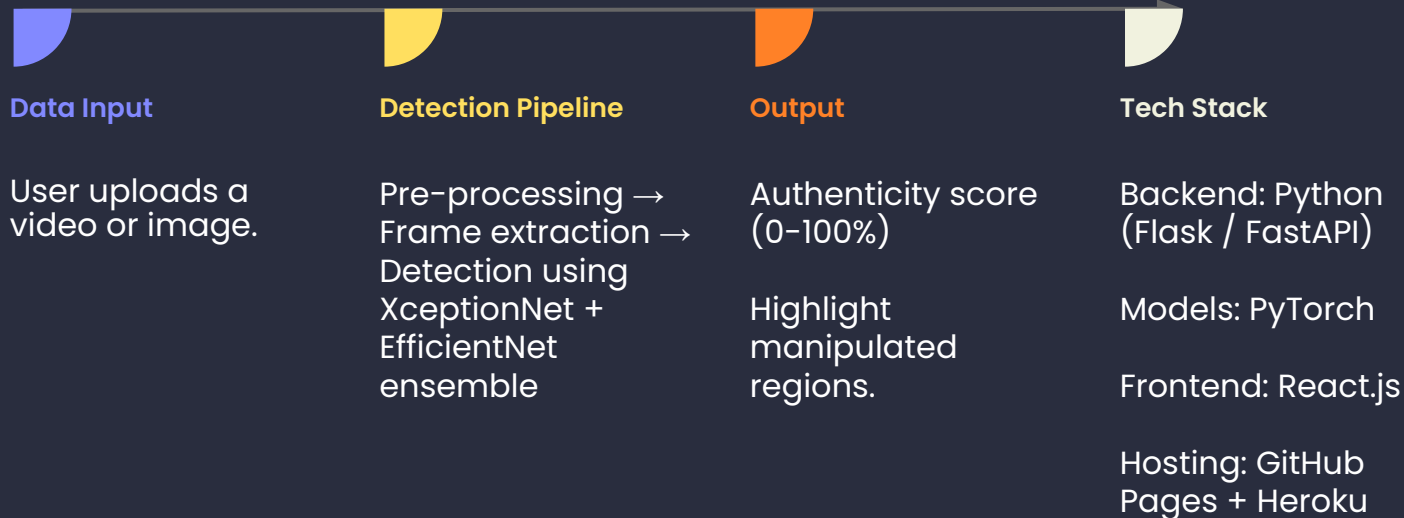
Method	Description	Strengths	Weaknesses
FaceForensics++	Benchmark dataset for deepfake detection	Rich labeled data	Dataset bias
XceptionNet	CNN-based model for frame-level detection	High accuracy on images	Struggles with new fakes
EfficientNet-B7	Lightweight deepfake detection network	Fast inference	Less robust to adversarial attacks
Deepware Scanner API	Cloud-based detection service	Easy integration	Limited transparency

## Key Finding

Ensemble of multiple models improves accuracy and robustness.

# Approach (System Design):

Goal: Build an upload-and-detect system



LINK

- [FaceForensics++ Dataset](#)
- [Deepware API](#)

**Thank you**