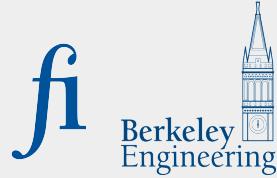


# Robust Detection of Manipulated Videos



**Team:**  
 Pranay Pasula (EECS)  
 Ujjwal Singhania (EECS)  
 Yijiu Zhong (EECS)

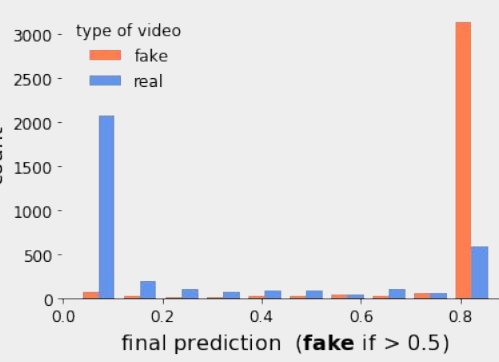
**Advisors:**  
 Dawn Song (EECS)  
 Ruoxi Jia (EECS)

Artificial videos are becoming **indistinguishable** from real content. For example, videos of **world leaders** have been manipulated to make them appear to say **incendiary** things. Previous approaches have failed to generalize over varied deepfake generation methods. We use machine learning and computer vision to create a classifier that **detects deepfake videos** and is **robust to unseen manipulation methods**.



**Facebook DFDC Dataset:**

- 119,146 videos
- Created by various deepfake methods to test generalization
- Skewed (86% fake videos)



**Classification Results:**

- **F1 Score:** 0.90 out of 1.00
- **Accuracy:** 88.8%
- **Fake Detection Rate:** 97.8%
- **Real Detection Rate:** 78.1%

\* We note that the risk of failing to detect a fake video is generally higher than the risk of misclassifying a real video

