Fake video detection

Popov Nikolay

Plan:

1) Techniques of fake media generation
2) Common steps of face swap
3) Common steps of reenactment
4) Baseline algorithm
5) Proposed improvement
6) Datasets
7) Results

1) Fake media creation techniques

A. Face swap
B. Reenactment
C. Non-existent portrait synthesis
D. Attribute Editing
E. Morphing Attacks

2) Face swap

1) **Pre-processing:**
   - Face detection/segmentation
   - Face landmarks
   - Face crop

2) **Generative network:**
   - Encode-decoder
   - CycleGAN
   - face recognition model + cGAN

3) **Post-processing:**
   - Boundary smoothing or blending

2) Face swap

Vulnerabilities:

- Generated face is merged with the old background
- New face is generated using convolutional generators
- Background for the resulting image is taken from the pristine image

3) Reenactment

Head pose and face expression are stored in:

- feature maps from encoder network
- 3DMM head parameters

Reenactment pipelines:

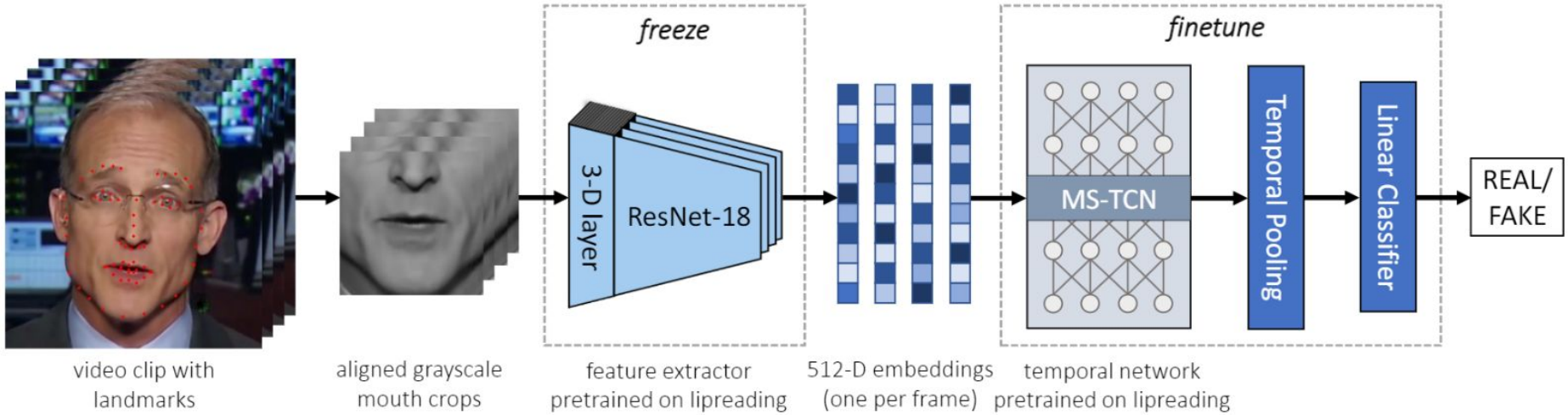
- Convolutional encoder-decoder
- Encoders for source and target images, decoder to create from mixed feature maps
- Encoder to map images to StyleGAN's W+ space, another net for modifying the latent code for inputed 3DMM parameters
- GAN conditioned on facial landmarks and identity embedding

Reenactment either includes the face merging step or performs full synthesis.

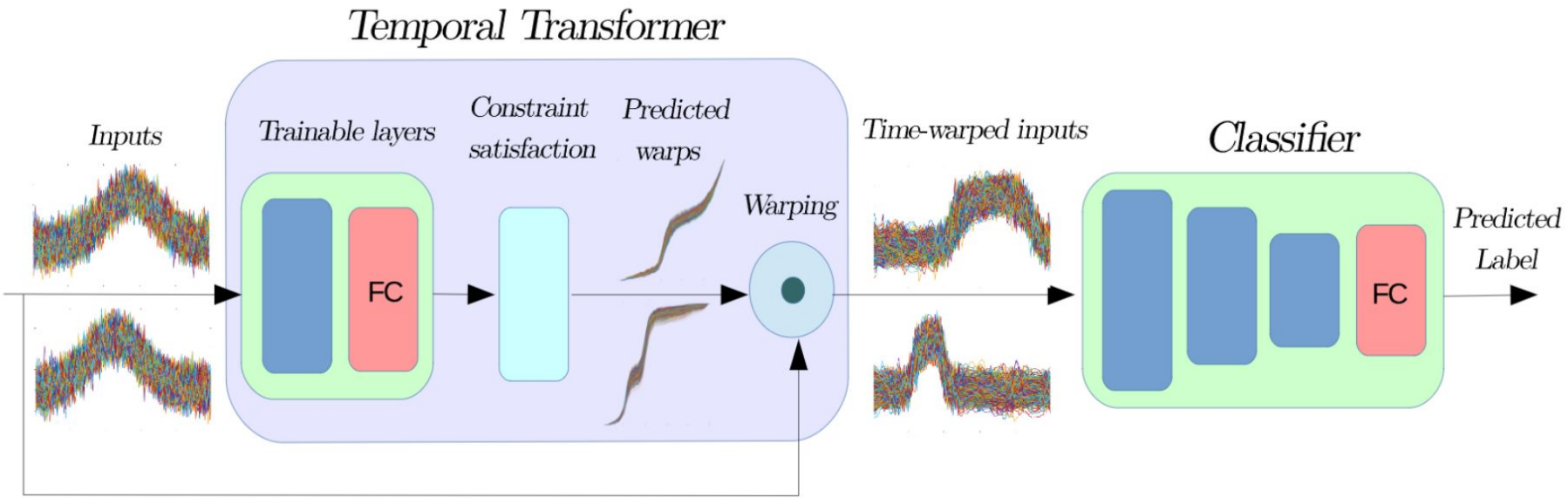
Common vulnerabilities of the fake videos:

- Temporal incoherence as consequence of frame-by-frame generation
- Teeth unnaturality - if there is no image with target person smiling
- Eyes inconsistency - coherence of gaze direction and color of human eyes
- Mouth movement incoherence and unnaturality

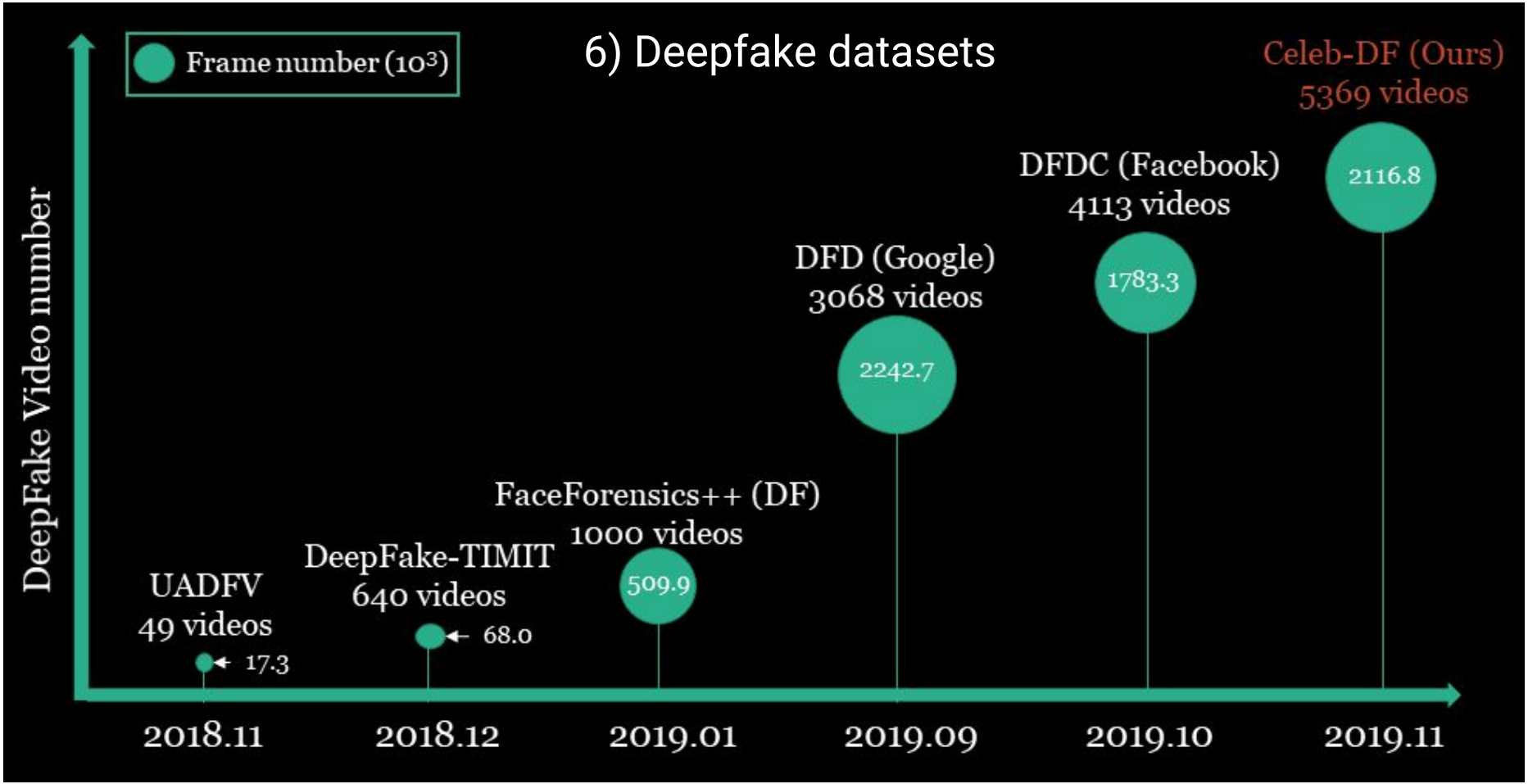
4) Baseline detection method - LipForensics net



5) Improvement - adding TTN



6) Deepfake datasets



Source - https://www.cs.albany.edu/~lsw/celeb-deepfakeforensics.html

7) Results

Model \ Test splits	FF++, c0	FF++, c23	FF++, c40	CelebDF-v2
LipForensics (reported), c0	99.9	-	-	-
LipForensics (reported), c23	-	99.7	-	82.4
LipForensics (public ckpt), c23	*98.9*	*98.7*	*81.3*	*81.5*
LipForensics+TTN, c0.1.c)	99.7	99.5	83.7	78.9
LipForensics+TTN, c23.2.a)	**99.6**	**99.6**	**84.0**	**81.4**

Table 1: Comparison of video-level AUC scores.

Questions