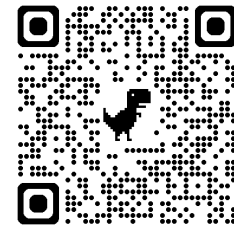




ICLR



A Large-Scale 3D Face Mesh Video Dataset via Neural Re-parameterized Optimization

Kim Youwang Lee Hyun* Kim Sung-Bin* Suekyeong Nam Janghoon Ju Tae-Hyun Oh

POSTECH

KRAFTON

NeuFace-dataset

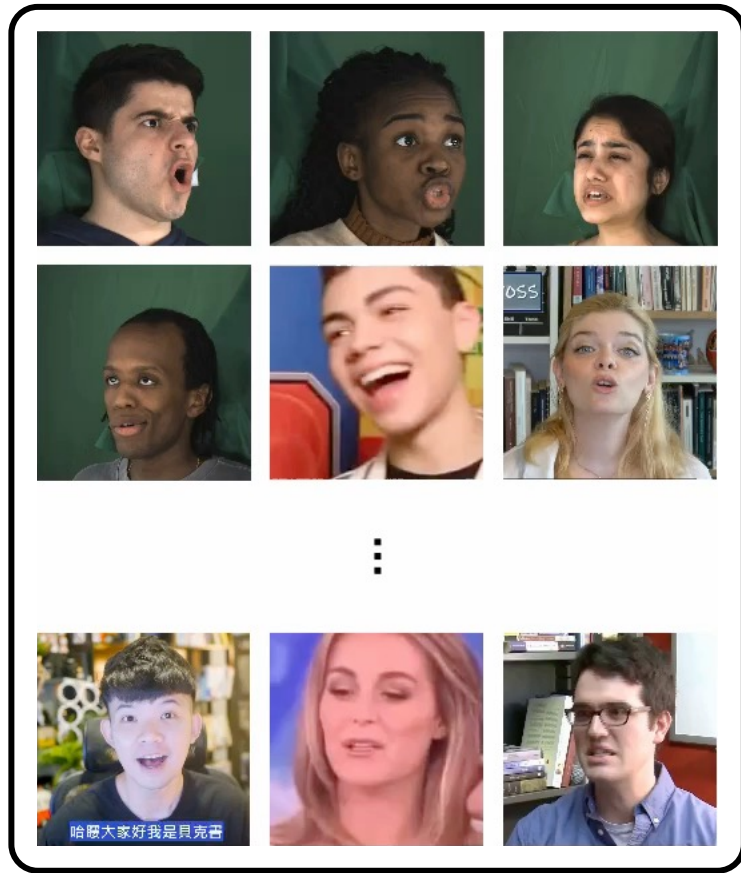
- We present **NeuFace-dataset**, the **largest-scale in-the-wild** 3D face video dataset.
- We provide **accurate** and **spatio-temporally consistent** 3DMM labels for 2D video datasets.



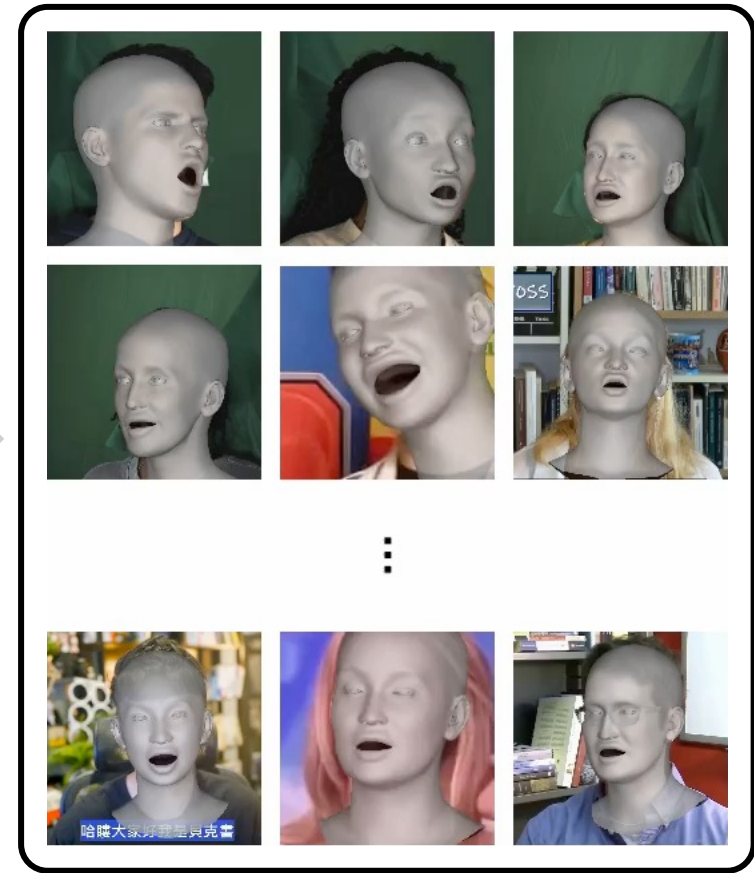
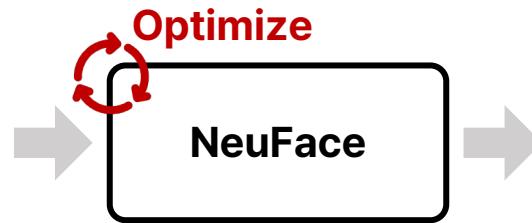
Disclaimer: Videos are not included in our dataset. We augment existing 2D face video datasets with obtained 3DMM labels.

NeuFace-dataset

- We **annotate 3D face meshes** (3DMM) for incomparably large amounts of **2D face video datasets**.
- Obtain 3DMMs for {**MEAD** / **VoxCeleb2** / **CelebV-HQ**} with a novel 3D face optimization, called **NeuFace**.



Multi-view / Monocular / Lab / In-the-wild
"Large-scale" 2D face video datasets



Multi-view / Monocular / Lab / In-the-wild
"Large-scale" face video + **3DMM** datasets

Existing 3D face datasets

	BIWI 3D	COMA	VOCA	(Ours) NeuFace-dataset
Number of sequences	1,100	150	500	1,245,000
Number of identities	14	12	12	21,048
Number of maximum views	1	4	4	7
Total duration	1.4 hrs	0.1 hrs	0.5 hrs	2,090 hrs
Captured environment	In-the-lab	In-the-lab	In-the-lab	In-the-lab + In-the-wild

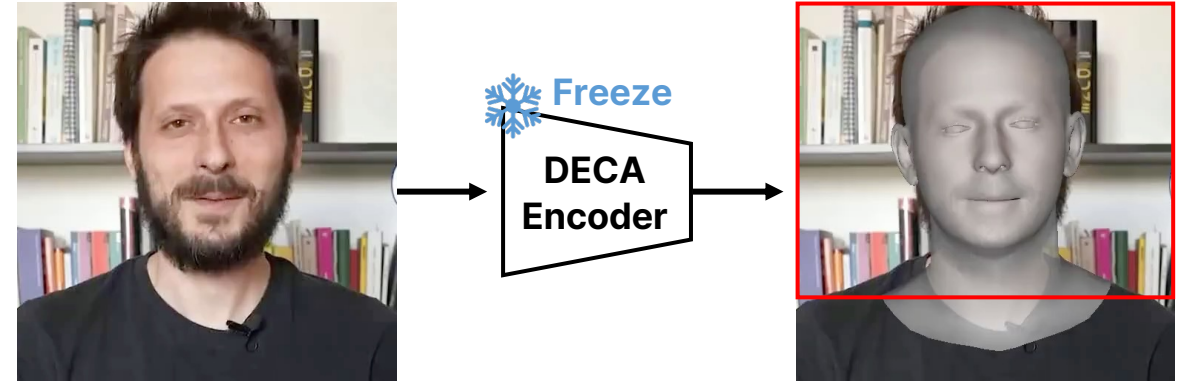
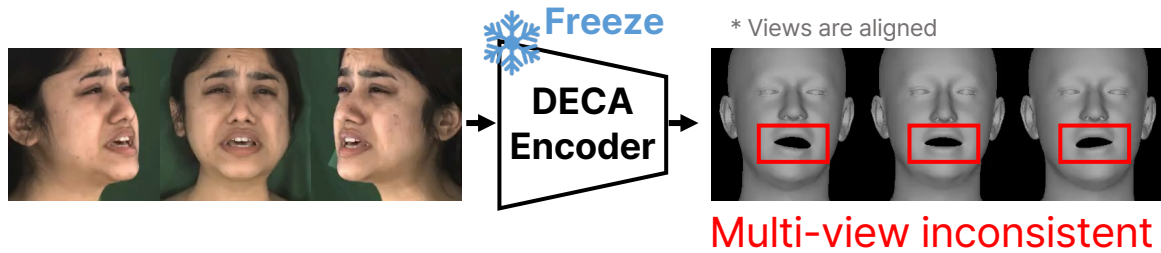
Existing 3D face datasets

	BIWI 3D	COMA	VOCA	(Ours) NeuFace-dataset
Number of sequences	<div>Issues of existing 3D face datasets<ul style="list-style-type: none">- Difficulties in acquisition- Limited sequences & identities- Short video duration- Limited capturing environments</div>			1,245,000
Number of identities				21,048
Number of maximum views				7
Total duration				2,090 hrs
Captured environment				In-the-lab + In-the-wild

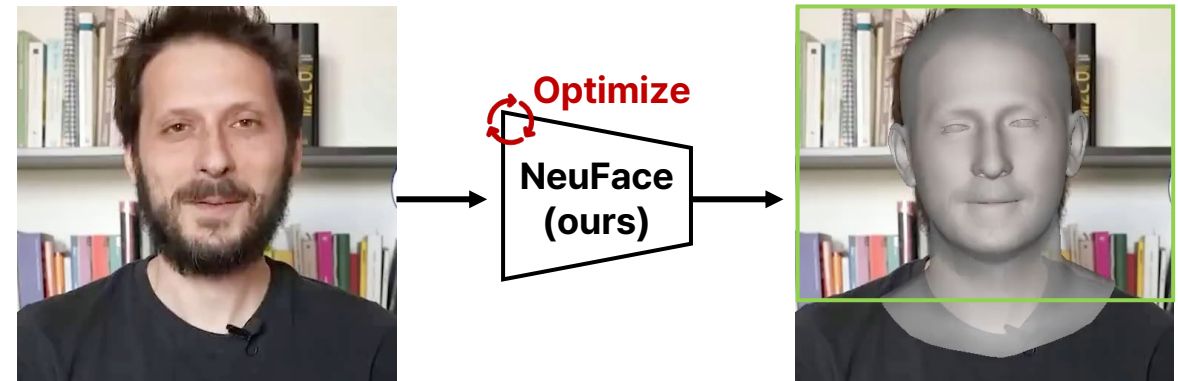
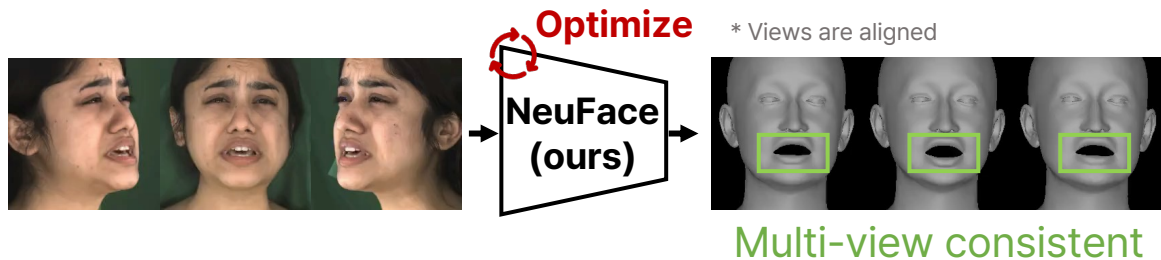
Problems of using off-the-shelf model as annotator

- Off-the-shelf neural models (e.g., DECA) reconstruct **spatio-temporally inconsistent meshes** on videos.

DECA

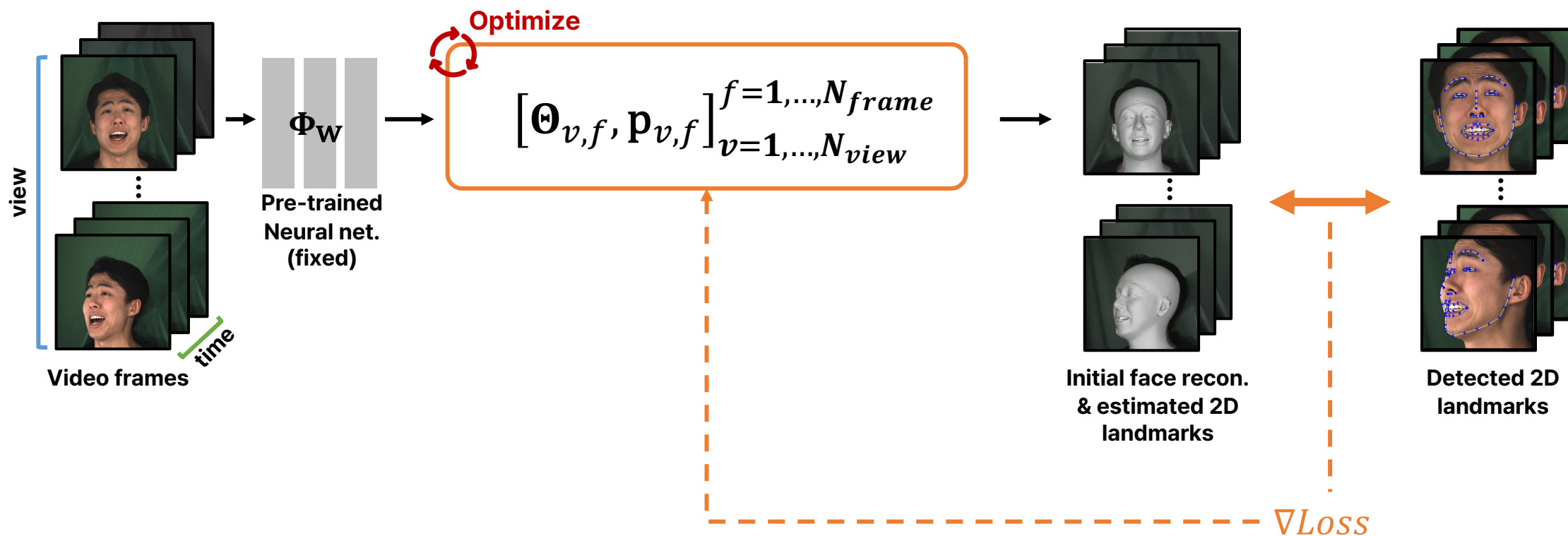


NeuFace (Ours)



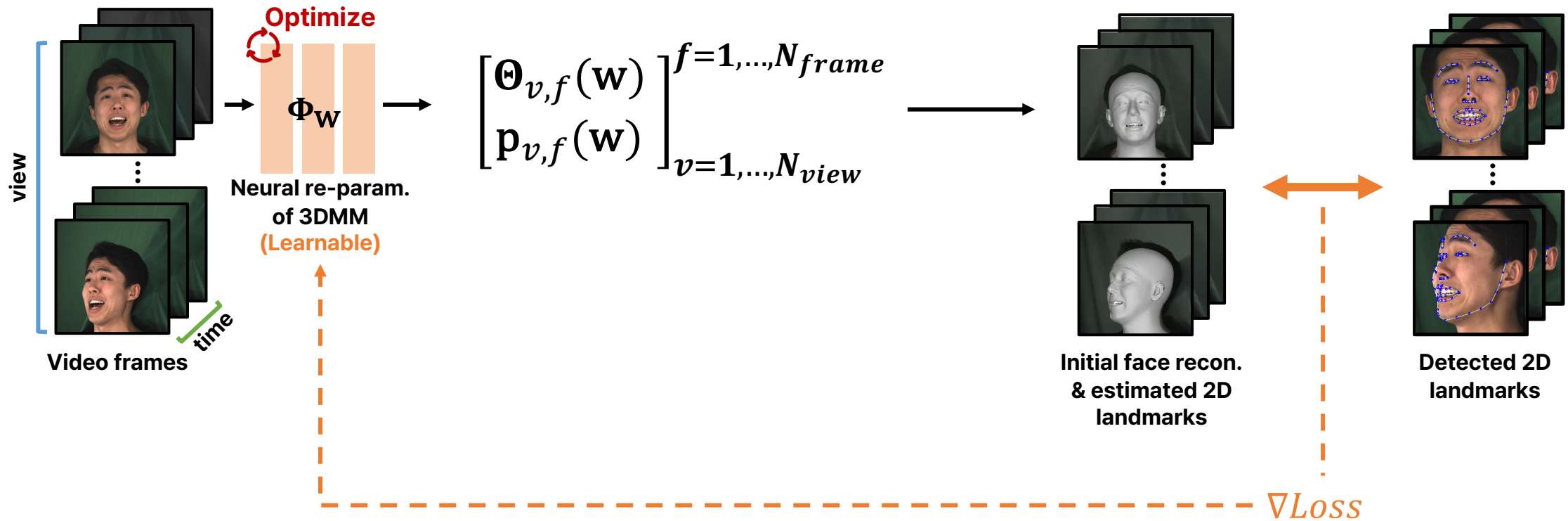
Baseline: Direct 3DMM fitting on videos

- Input: facial videos (multi-view or monocular).
- Supervision: Detected 2D landmarks.
- **Optimization variable:** 3DMM parameters (e.g., FLAME parameters) & camera parameters.

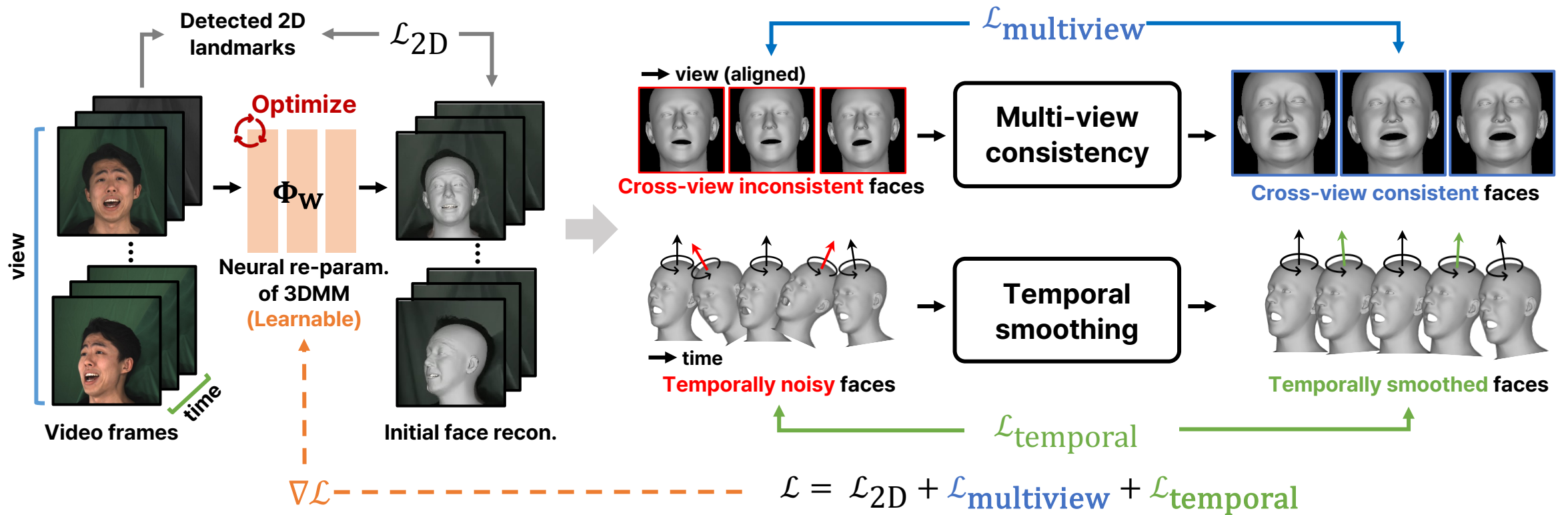


Key idea of NeuFace optimization

- Input: facial videos (multi-view or monocular).
- Supervision: Detected 2D landmarks.
- **Optimization variable: Neural parameters (\mathbf{w})** representing 3DMM & camera parameters.

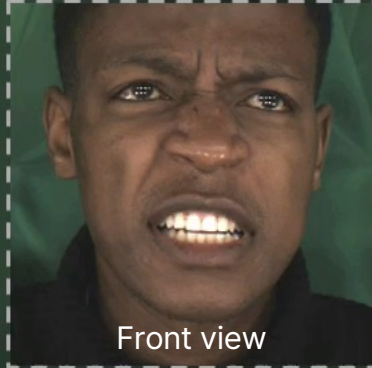


NeuFace optimization: generating 3DMM annotations



Comparison with the baseline (Direct 3DMM optimization)

Motion jitter



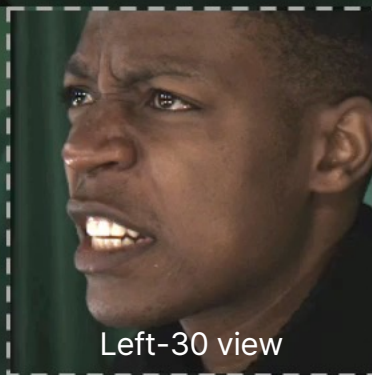
Front view

Mean shape bias



Direct 3DMM optimization

NeuFace (ours)

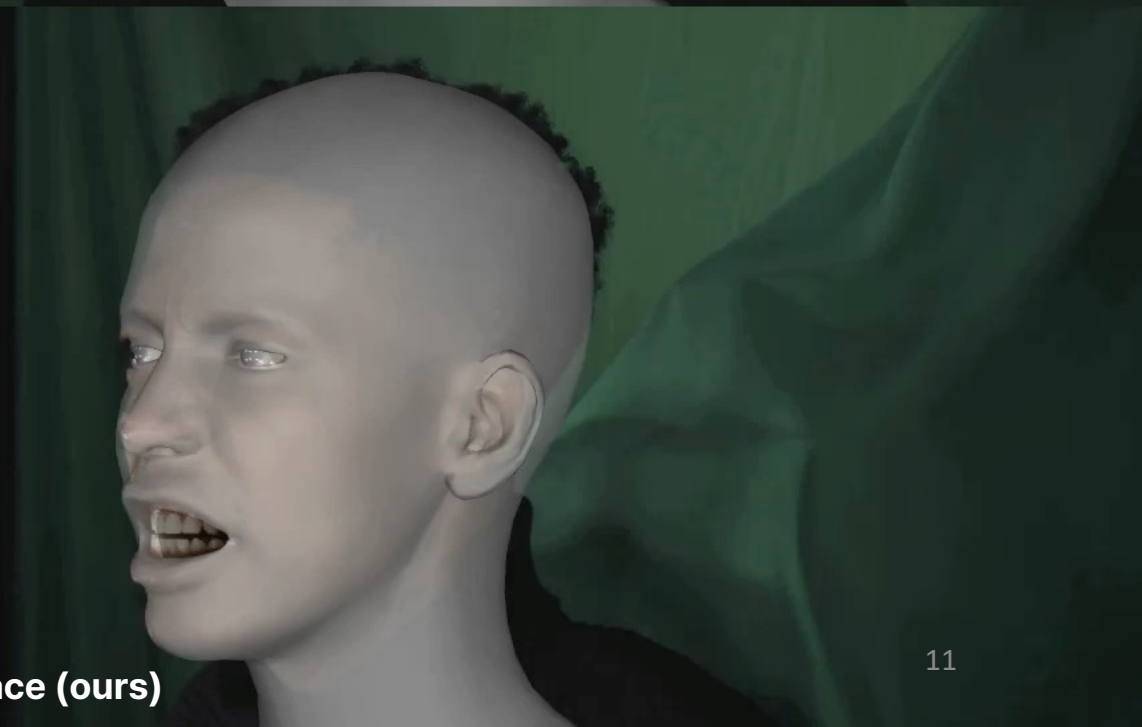


Left-30 view

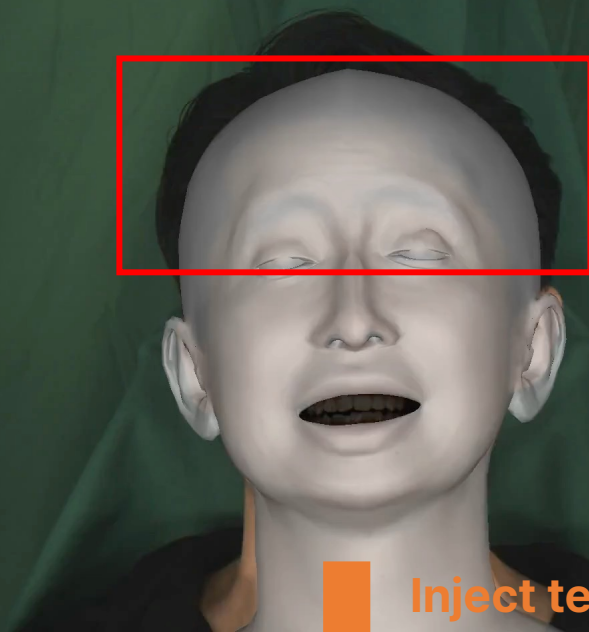


Direct 3DMM optimization

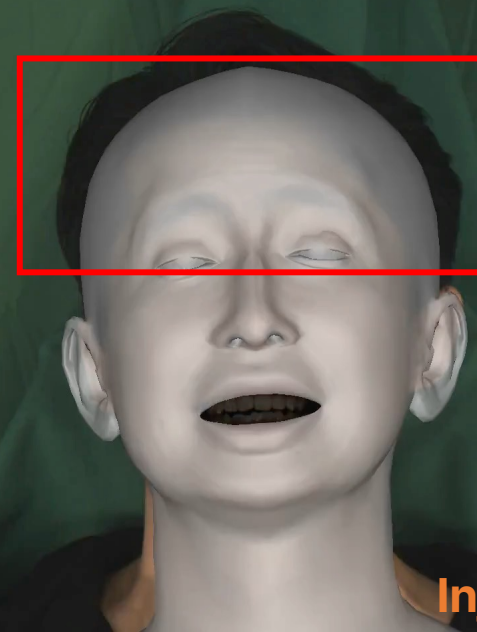
NeuFace (ours)



NeuFace: Loss ablation

\mathcal{L}_{2D} 

Inject multi-view
consistency

 $\mathcal{L}_{2D} + \mathcal{L}_{\text{multiview}}$  $\mathcal{L}_{2D} + \mathcal{L}_{\text{temporal}}$ 

Inject temporal
consistency



Inject multi-view
consistency

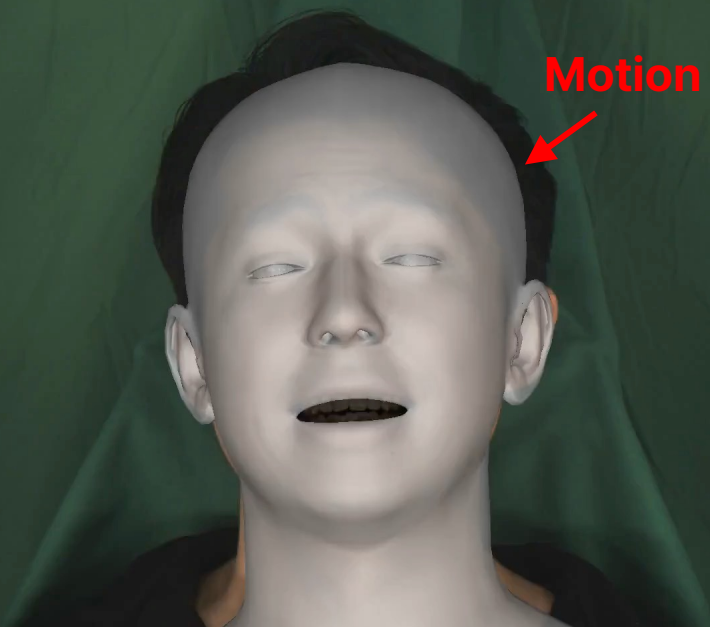
 $\mathcal{L}_{2D} + \mathcal{L}_{\text{multiview}} + \mathcal{L}_{\text{temporal}}$ 

Inject temporal
consistency

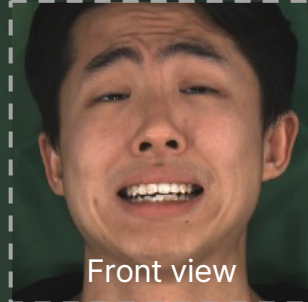


DECA vs. NeuFace

DECA



Motion jitter

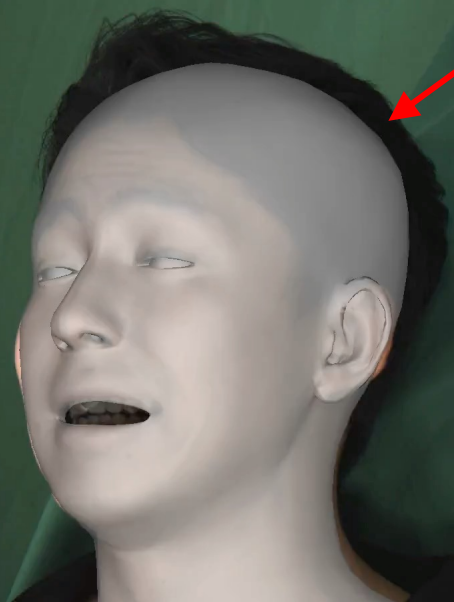


Front view

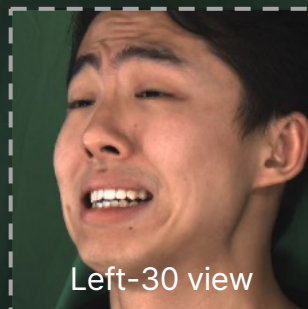
NeuFace (ours)



DECA



Motion jitter



Left-30 view

NeuFace (ours)



Application of the NeuFace-dataset

Learning 3D human facial motion prior

- We train different generative 3D facial motion priors using:
 - (1) existing, small-scale motion capture dataset, **VOCASET**
 - (2) 3D face mesh sequences in **NeuFace-dataset**
- With the learned prior, we **randomly generate long-term motions** (10s), given only a starting frame.

Prior trained with
VOCASET
(existing)



Prior trained with
NeuFace-dataset
(ours)



Learning 3D human facial motion prior

- We train different generative 3D facial motion priors using:
 - (1) existing, small-scale motion capture dataset, **VOCASET : Fails to generate natural facial motion**
 - (2) 3D face mesh sequences in **NeuFace-dataset : Generates natural, smooth facial motion**
- **{Large-scale, diverse, natural}** motions in our dataset help understanding human face dynamics!

Prior trained with
VOCASET
(existing)



Prior trained with
NeuFace-dataset
(ours)



**We expect NeuFace-dataset can open up
new opportunities for the 3D face community.**

Thank you for your attention!