

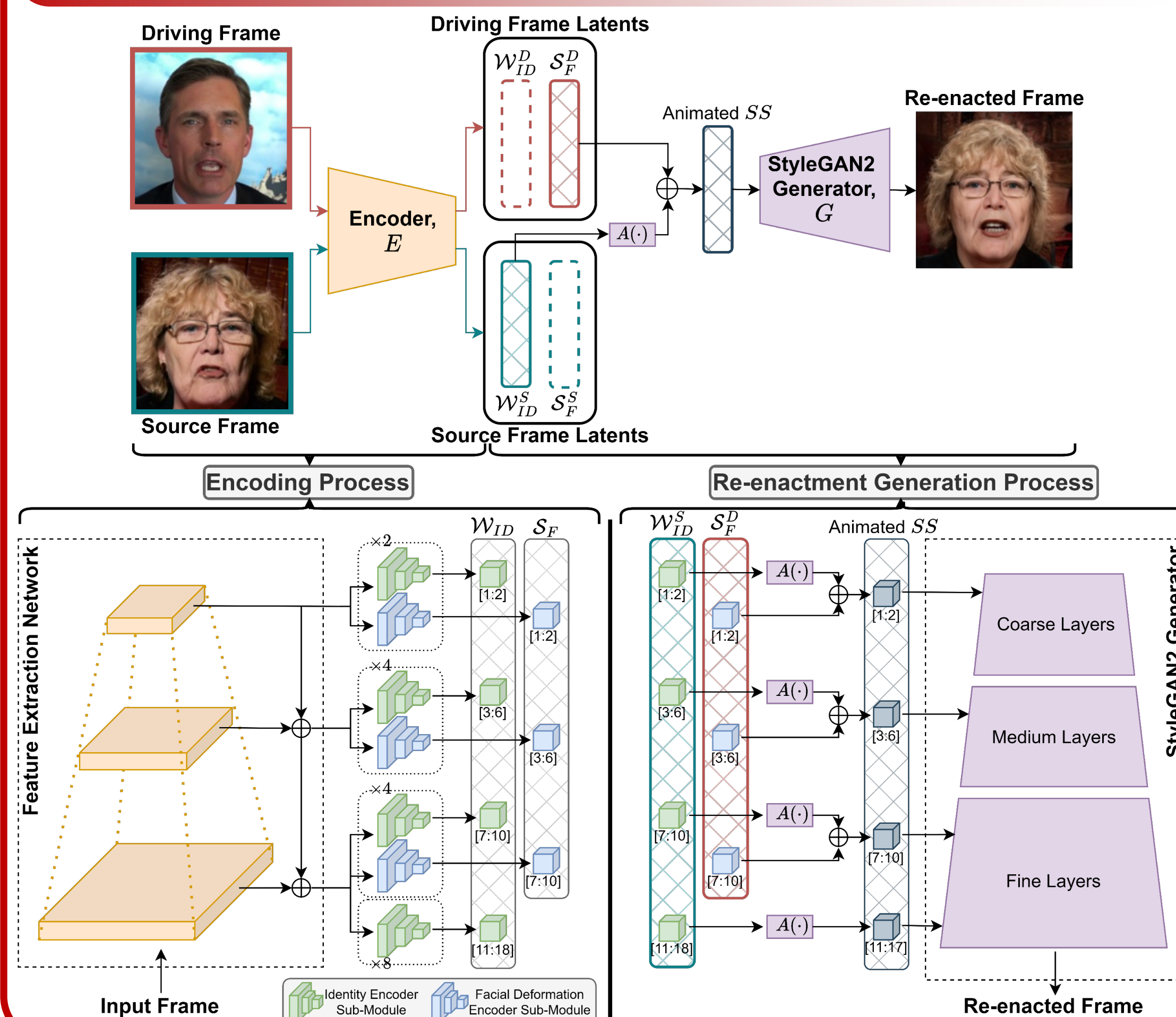


Trevine Oorloff and Yaser Yacoob
University of Maryland, College Park, Maryland, USA

Motivation

- Existing face re-enactment approaches rely on at least one of the following:
 - Explicit 2D/ 3D structural priors
 - Flow based warping
- Existing one-shot re-enactment methods have **poor one-shot robustness** (i.e., robustness to diverse expressions and head poses of the source frame)
 - inconsistent predictions
 - inability to capture fine details and facial accessories
 - poor generalization
 - artifacts, etc.
- Considering **StyleGAN2 latent space manipulations** it is evident that the pre-defined latent space of a pre-trained StyleGAN2 has **implicit 3D facial priors embedded** within it
- We conjecture that **StyleGAN2's latent spaces** are **not yet fully exploited for re-enactment** and the use of explicit structural representations is **redundant** and **limits the performance** of StyleGAN2 to the capacity limits of such structural priors

Pipeline



Main Contributions

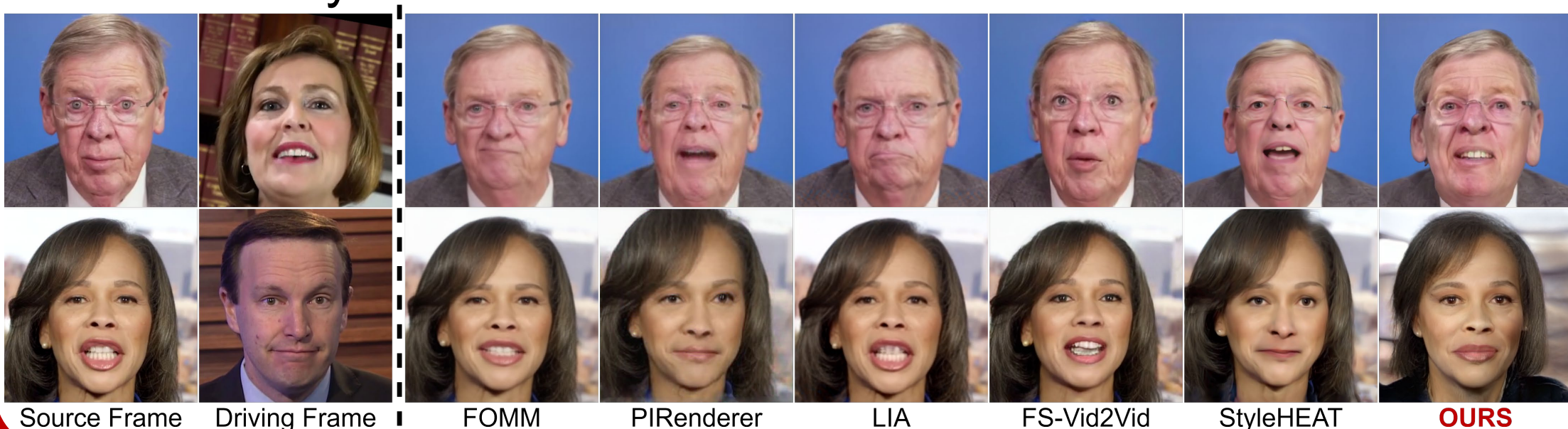
- A novel StyleGAN2-based hybrid latent space framework to facilitate **one-shot face re-enactment at 1024²**
 - without relying on explicit structural priors for guidance
 - accommodating latent-based **attribute edits**
 - robust** to diverse facial expressions and head poses of the source frame
 - achieving **state-of-the-art** results both **quantitatively and qualitatively** (quantitative improvement of up to 12% in cross-identity re-enactment and 50% in one-shot robustness)
- To the best of our knowledge, we are the first,
 - to **decompose identity and facial deformation** within the pre-trained **StyleGAN2's predefined latent spaces** itself
 - to handle **robustness to diverse head pose and expressions** of the source frames
 - to propose a **manifold adjustment technique** handling both source **identity reconstruction** and **non-homogeneity** of the latent space in the task of latent-based re-enactment

Qualitative Results (Re-enactment)

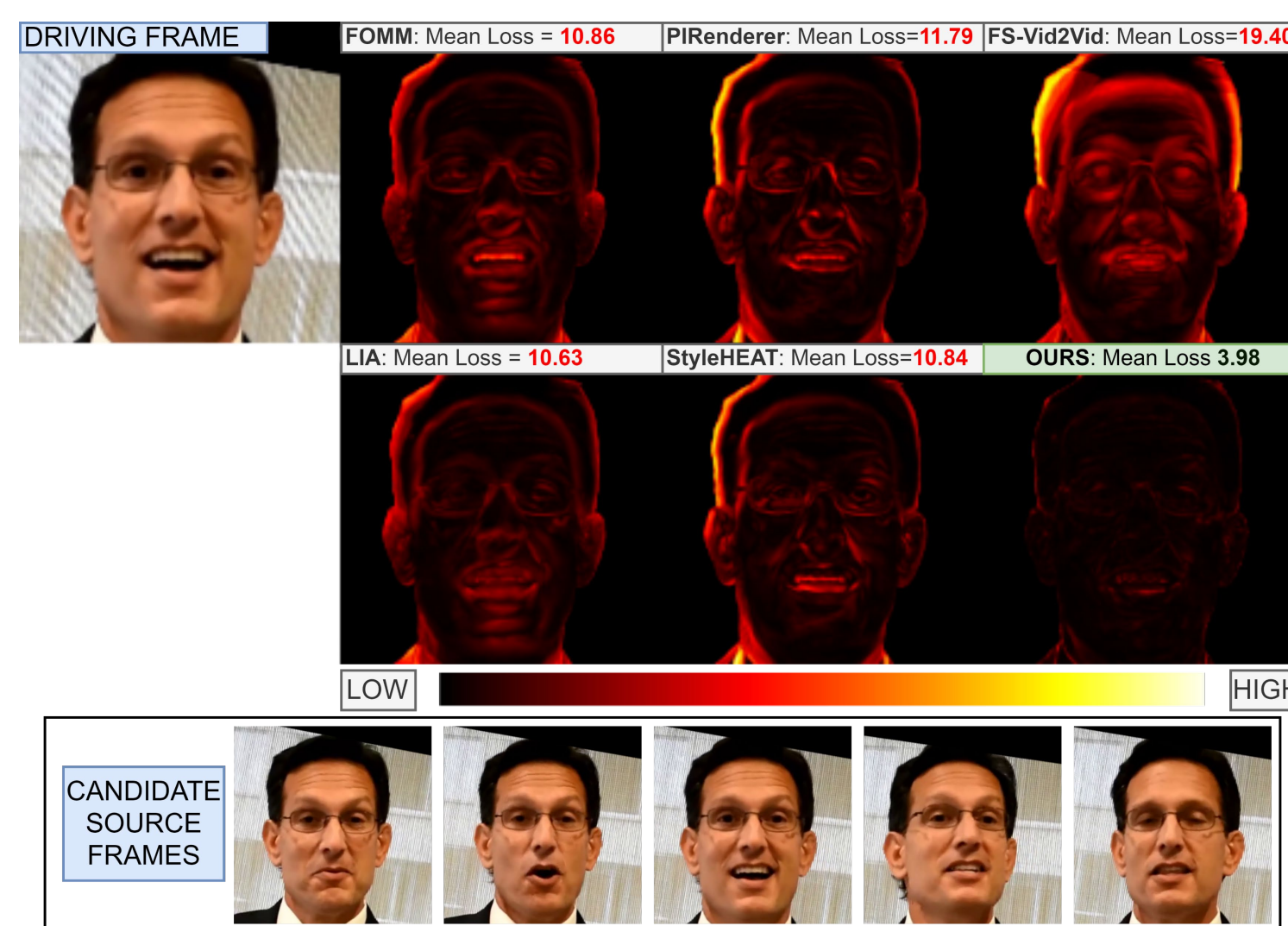
Same-Identity Re-enactment



Cross-Identity Re-enactment



Qualitative Results (One-Shot Robustness)



Quantitative Results

Cross-Identity Re-enactment

Method	FID↓	FVD↓	FVD _w ↓	ρ_{AU+GZ} ↑	ρ_{pose} ↑
FOMM	94.0	529.4	78.4	0.450	0.782
PIRenderer	<u>84.8</u>	417.3	<u>54.2</u>	<u>0.668</u>	<u>0.880</u>
LIA	94.8	536.2	76.7	0.404	0.788
FS-Vid2Vid	90.6	532.7	86.7	0.493	0.745
StyleHEAT	97.2	<u>408.8</u>	58.9	0.645	0.875
Ours	74.3	375.4	50.2	0.718	0.915

One-Shot Robustness

Method	LPIPS↓ $\times 10^{-2}$	L_{ID} ↓ $\times 10^{-1}$	FID↓ $\times 10^1$	FVD↓ $\times 10^2$
FOMM	7.5 ± 2.1	2.3 ± 1.1	3.2 ± 1.0	2.0 ± 0.7
PIRenderer	<u>5.7 ± 0.4</u>	<u>1.2 ± 0.2</u>	<u>2.2 ± 0.2</u>	1.5 ± 0.3
LIA	7.8 ± 2.0	2.3 ± 1.0	3.4 ± 1.0	2.0 ± 0.8
FS-Vid2Vid	7.5 ± 1.1	1.8 ± 0.5	3.3 ± 0.6	2.2 ± 0.5
StyleHEAT	6.0 ± 0.4	1.4 ± 0.2	3.3 ± 0.5	<u>1.5 ± 0.2</u>
Ours	2.9 ± 0.2	1.1 ± 0.1	1.5 ± 0.2	0.8 ± 0.1