# R&D SH@WCASE 2023

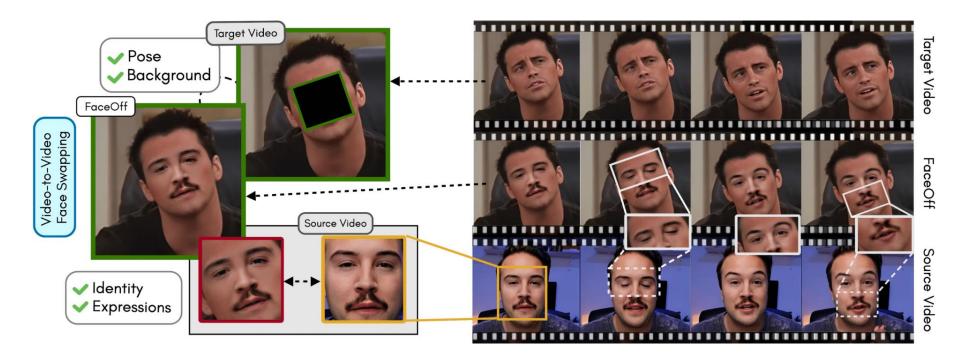
# FaceOff: A Video-to-Video Face-Swapping System (WACV 2023)

## **ABSTRACT**

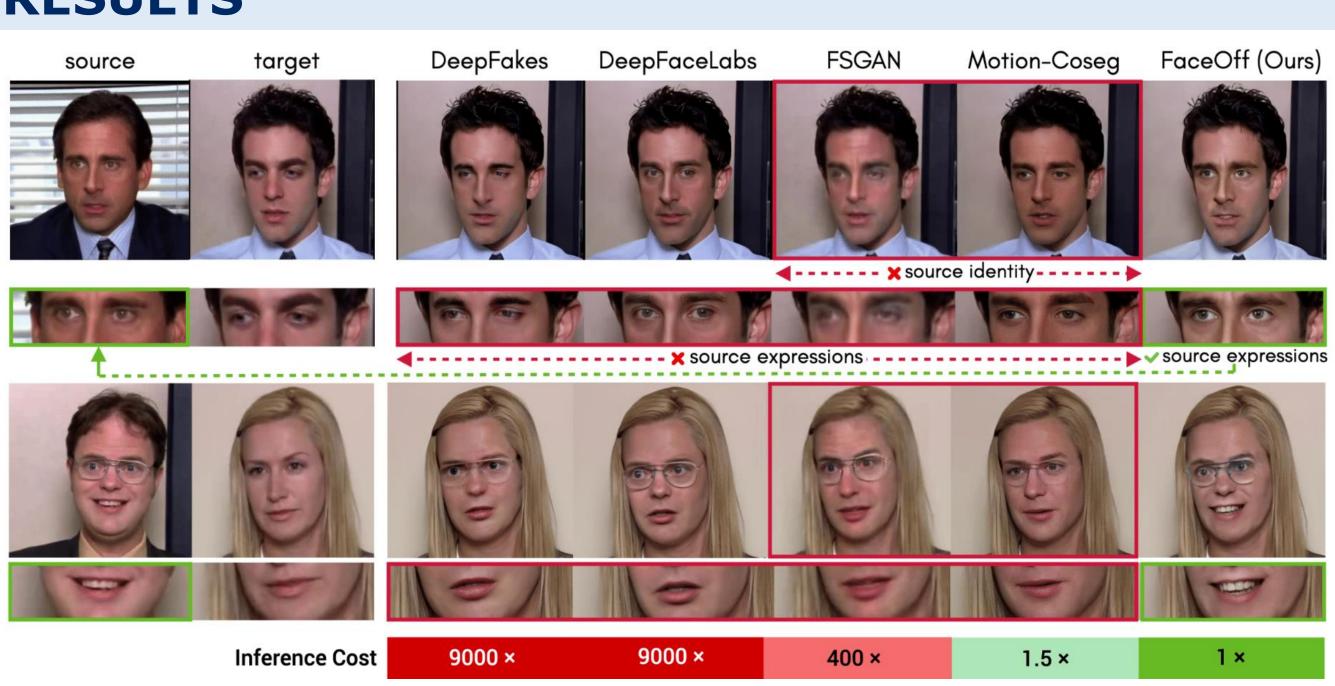
Doubles take the place of the actors in dangerous stunt scenes, and their face is later replaced with actors using expensive CGI techniques. Existing face-swapping techniques preserve the source expressions of the actor important for the scene's context. We introduce video-to-video face-swapping, a novel task of face-swapping that can preserve (1) the identity and expressions of the source actor and (2) the background and pose of the target video. We propose FaceOff, that operates by learning a robust blending operation to merge two face videos following the above constraints.

## **METHODOLOGY**

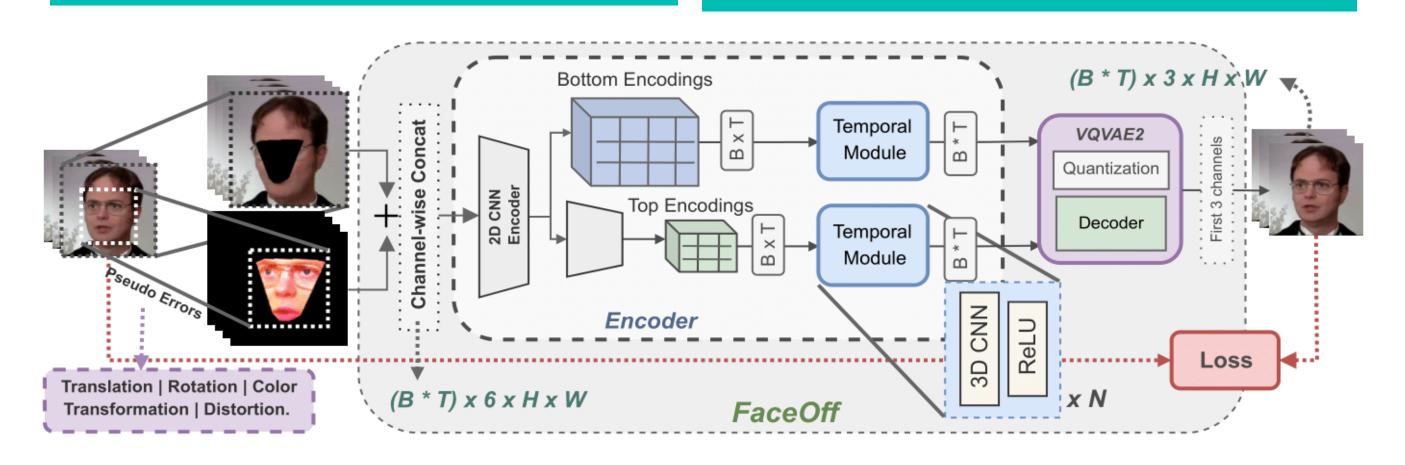
We reduce the face videos to a quantized latent space and blend them in the reduced space using a self-supervised training strategy. We use a single video as the source and target and introduce pseudo motion errors on the source video. Finally, we train a network to 'fix' these pseudo errors to regenerate the source video.



### RESULTS



Self-supervised Training face-swapping in We pose videos as a blending problem given two videos as input, blend the videos in the reduced space to produce a coherent output.

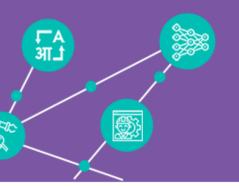


Author names: Aditya Agarwal\*, Bipasha Sen\*, Rudrabha Mukhopadhyay, Vinay Namboodiri, C.V. Jawahar

### **Inference Pipeline**

TECHSSALCHEMY

FaceOff can be directly inferred on any unseen identity without any finetuning by blending the foreground of the source with the background of target.









# **CONCLUSION FUTURE** WORKS

propose FaceOff, a self-We supervised temporal autoencoding network that takes two face videos as input and produces a single coherent blended output. Our work adds a new dimension to movie editing that can potentially save months of tedious manual effort and millions of dollars.

Future Work: We would like to improve the quality of the blended output, reduce temporal jitters, and handle extreme pose variations between source and target videos.

### REFERENCES

- Yuval Nirkin et.al. Fsgan: Subject swapping agnostic face and reenactment.
- Aliaksandr Siarohin, Motion supervised co-part segmentation.
- Ali Razavi et. Al., Generating diverse high-fidelity images with vq-vae-2.

email ID: {aditya.ag, bipasha.sen, radrabha.m}@research.iiit.ac.in