# Generative Landmarks

David Ferman and Gaurav Bharaj

Introduction

We pose landmark detection as an image translation problem. We capture two sets of unpaired marked (with paint) I^M_{real} and unmarked videos I^U_{real}.

We then use a generative adversarial network and cyclic consistency to predict deformations of landmark templates that simulate markers on unmarked images until these images are indistinguishable from ground-truth marked images.
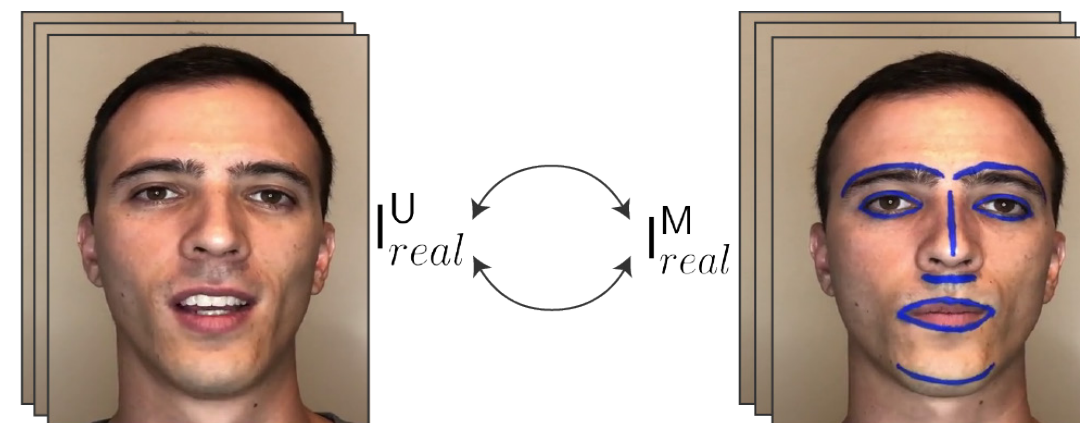
Our novel method does not rely on manually labeled priors, is temporally consistent, and image class agnostic.

Related Works

Off-the-shelf landmark detection methods suffer inconsistencies due to ambiguity in manual landmark annotation as well as temporal imperfections of frame-to-frame labeling, as landmarks are difficult to define precisely [1, 2]. As a result, landmark detection models suffer from temporal jitters, which have been addressed by a boundary-focused approach [1] as well as optical flow [2].

Method: Generative Landmarks

We train a landmark deformation network that takes-in unmarked images and predicts landmark deformations, learning from the marked images in an unsupervised fashion. Rather than predicting these landmarks directly, we use a template with predefined spatial landmarks.

The network intrinsically predicts landmark deformations that are applied as offsets to the template and rendered onto the unmarked images.

Input Image   Inferred Image   Inferred Landmarks
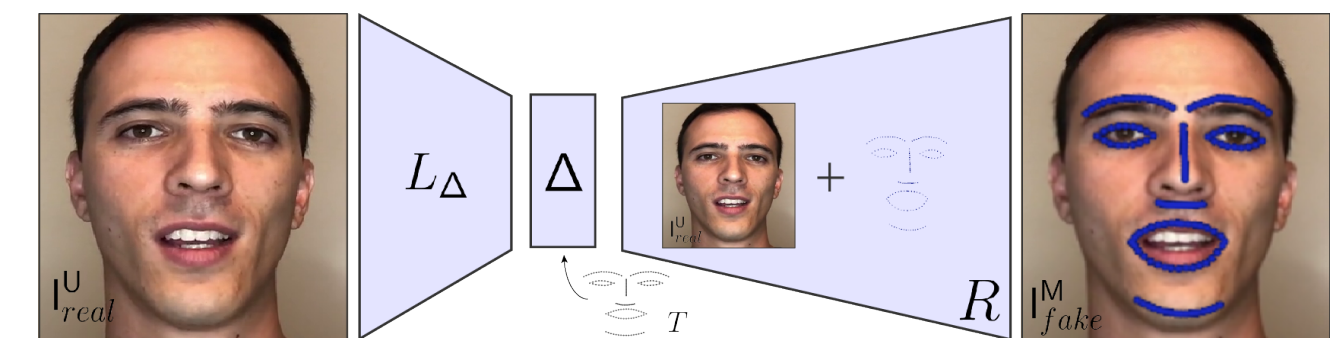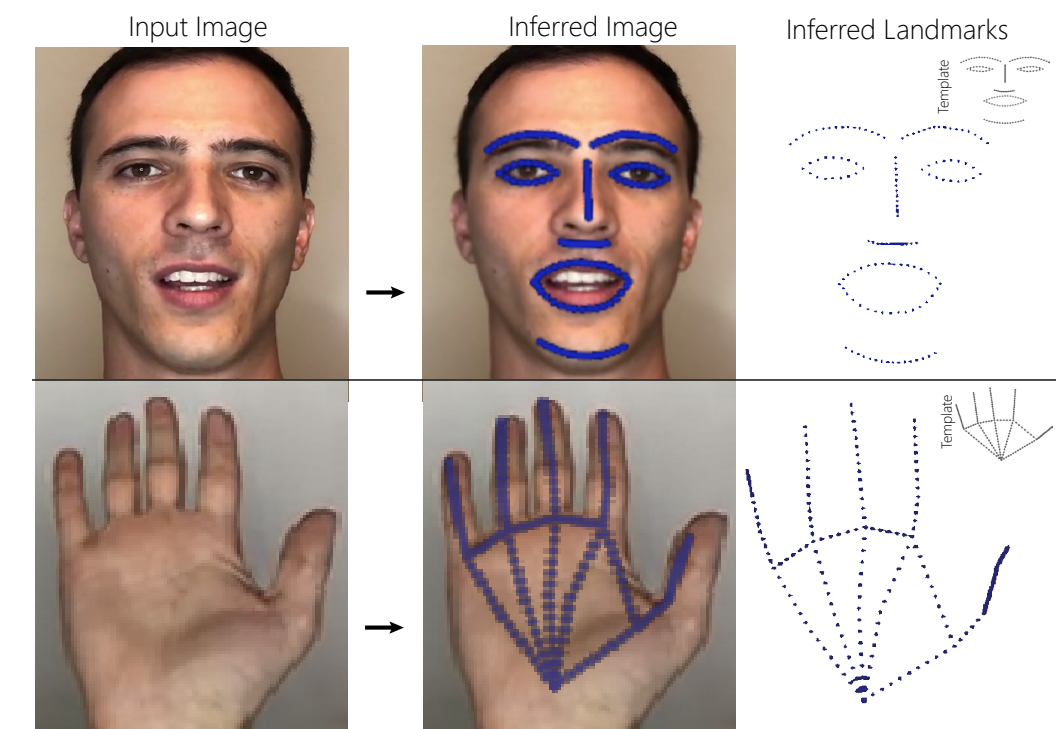
References

[1] Wu W., Qian C., Yang S., Wang Q., Cai Y., Zhou Q.: Look at boundary: A boundary-aware face alignment algorithm, IEEE CVPR 2018

[2] Dong X., Yu S.-I., Weng X., Yang Y., Sheikh Y.: Supervision by registration: An unsupervised approach to improve the precision of facial landmark detectors, IEEE CVPR 2018