Single-frame Regularization for Temporally Stable CNNs

Gabriel Eilertsen[1], Rafał K. Mantiuk[2], Jonas Unger[1]
[1] Linköping University, Sweden [2] University of Cambridge, UK

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

I. Motivation

Main goal

Our objective is to produce temporally stable video results when processing frames with image-to-image neural networks.

Deep neural networks are sensitive beings

Convolutional neural networks (CNNs) are highly sensitive to small changes in input. This is evident from previous work on adversarial examples, where even perceptually indistinguishable changes can result in widely different predictions.

For CNNs used to perform image-to-image mappings of video material, the sensitivity is manifested in temporal artifacts: flickering, unnatural changes of local features, etc.

Existing methods are heavy-weight

Existing neural network-based methods for temporal image-to-image processing often rely on dense motion estimation between frames (optical flow) for overcoming the above problems. The motion information is used in training, and/or to filter the output frames. This means that the CNN architecture needs to be modified, or that extensive post-processing is required. Also, reliable optical flow is a key element, which is difficult to achieve in some situations.

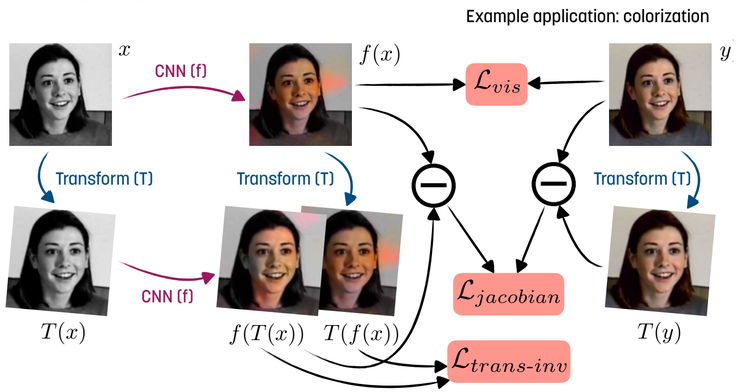
Our solution

We propose a regularization technique, in two different formulations, for stabilizing CNNs in the time domain. It provides a simple strategy for enforcing stability, with the following advantages:

- It is light-weight.
- It can be trained without video or motion information.
- It can be applied to any CNN without architectural modifications.
- It can be used for fine-tuning already trained CNNs.

II. Method

Example application: colorization



Setting

We consider two different loss regularization formulations, $\mathcal{L}_{reg}$, for complementing the visual difference (reconstruction) loss, $\mathcal{L}_{vis}$,

$$\mathcal{L} = (1-\alpha)\mathcal{L}_{vis} + \alpha\mathcal{L}_{reg}.$$

We rely on the transformation function $T(\cdot)$ in both of the regularization formulations, for specifying changes between two consecutive video frames. The changes are simulated by random geometric transformations.

Regularization formulations

(i) **Transform invariance:** We want to minimize temporal incoherence, which can be measured from the differences between frames that cannot be explained by the motion $T(x)$ between the frames,

$$\mathcal{L}_{trans\text{-}inv} = ||f(T(x)) - T(f(x))||_2 .$$

(ii) **Sparse Jacobian:** We consider not only function values, but also partial derivatives over time, in form of the Jacobian. We use a sparse formulation, sampling the Jacobian in one well-selected direction $T(\cdot)$,

$$\mathcal{L}_{jacobian} = || \underbrace{(f(T(x)) - f(x))}_{\text{Gradient of reconstructed frame}} - \underbrace{(T(y) - y)}_{\text{Gradient of ground truth frame}} ||_2 .$$

III. Results

Tested applications

We use colorization and high dynamic range (HDR) video reconstruction as example applications.

(i) We test colorization of faces and HDR reconstruction of artificially generated frames, sampling at different regularization strengths. We find that our regularization not only provides a significant improvement in temporal stability, but also in PSNR.

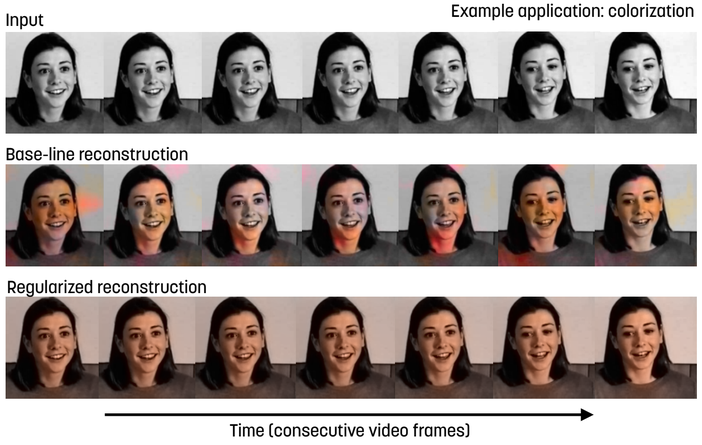
(ii) Applied to state-of-the-art CNNs for colorization and HDR reconstruction, the regularization gives a significant improvement in temporal stability while preserving the PSNR.

Conclusion

Using the proposed light-weight regularization strategy, we are able to demonstrate substantial improvements in temporal stability while preserving the PSNR. For smaller datasets, the PSNR is also boosted.

Transform invariance formulation can sometimes give better performance, while the sparse Jacobian is less sensitive to the regularization strength.

Example application: colorization



http://hdrv.org/hdrcnn/cvpr2019/

LINKÖPING UNIVERSITY
UNIVERSITY OF CAMBRIDGE
European Research Council
Established by the European Commission