Single Image Reflection Removal Exploiting Misaligned Training Data and Network Enhancements

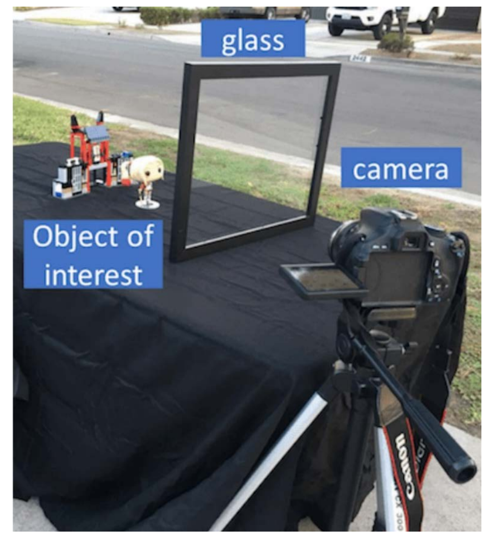
Kaixuan Wei[1] Jiaolong Yang[2] Ying Fu[1] David Wipf[2] Hua Huang[1]
[1] Beijing Institute of Technology, [2] Microsoft Research

Introduction

Challenges of Single Image Reflection Removal

State-of-the-art methods trained on synthetic data generalize poorly to real-world cases. These failures stem from:

- the fundamental ill-posedness of the problem
- the domain gap between the synthetic world and real-life photographs
- the insufficiency of densely-labeled real-world training data



Labor-intensive

Contributions

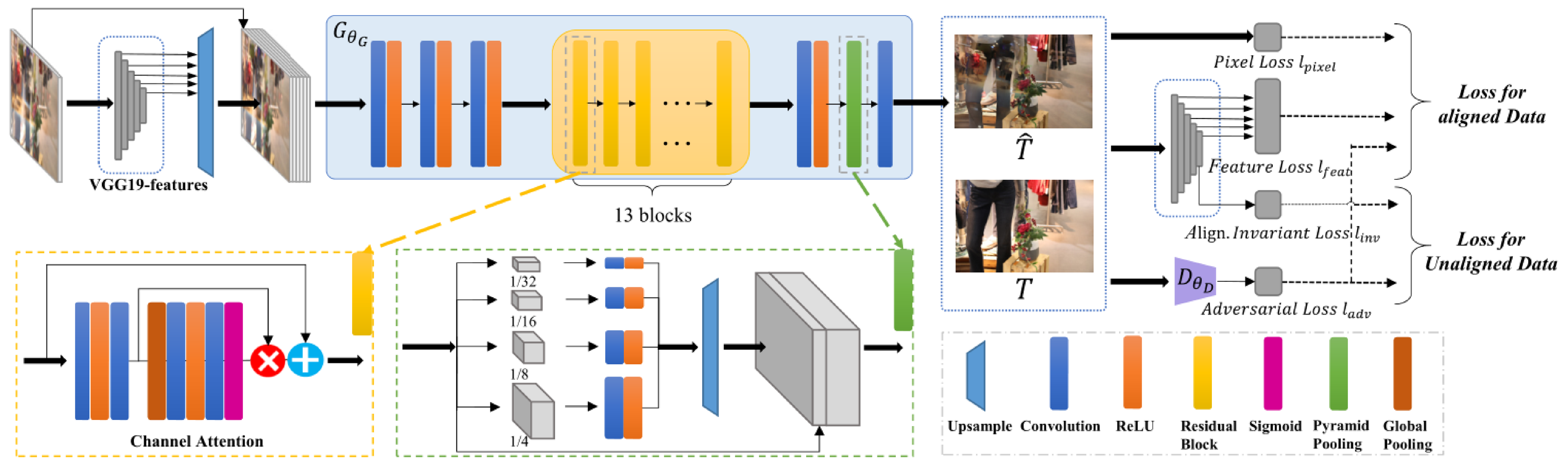
- Context encoding modules that are capable of leveraging high-level contextual clues to reduce indeterminacy within areas containing strong reflections
- Alignment-invariant loss function that facilitates exploiting misaligned real-world training data that is much easier to collect
- A new dataset including 450 unaligned image pairs that is publicly available to the community

No need for a tripod, table, or other special hardware

Our collected unaligned dataset

Our Approach

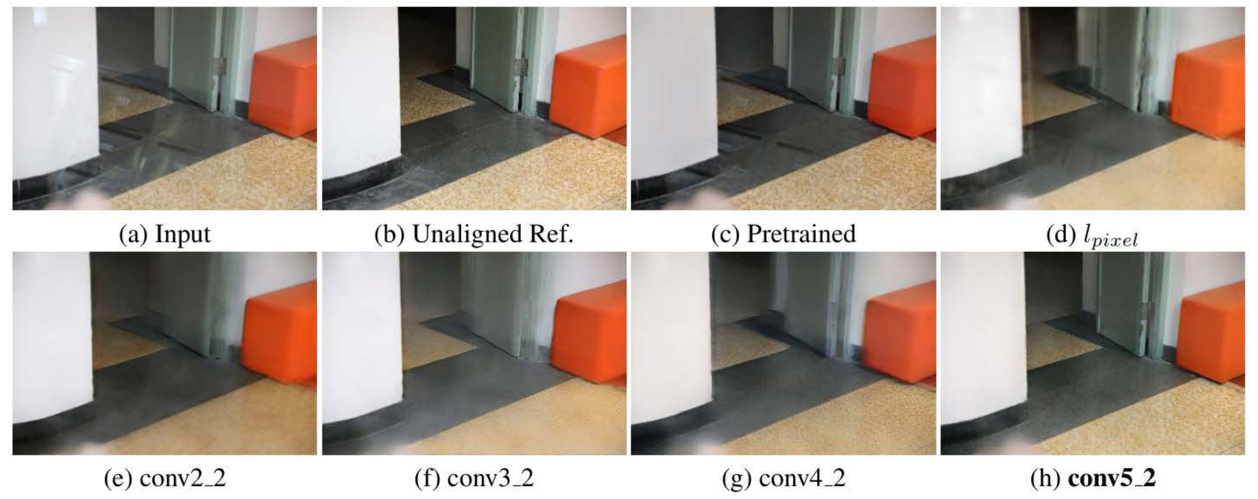
Overview of Our Approach



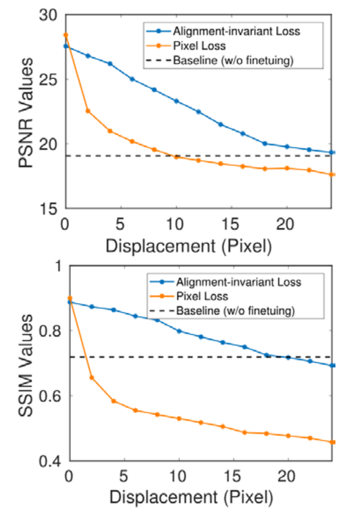
- Context encoding modules
  - Channel-wise context (yellow) : channel attention module
  - Multi-scale spatial context (green) : pyramid pooling module
- Loss function for unaligned data
  - Highest-level VGG feature loss : $l_{inv} = \|\phi_h(T) - \phi_h(\hat{T})\|_1$
  - Adversarial Loss : $l^G_{adv} = -\log(D_{\theta_D}(T,\hat{T})) - \log(1 - D_{\theta_D}(\hat{T},T))$

Alignment-invariant Test

Finetuning using misaligned image pair by different loss functions



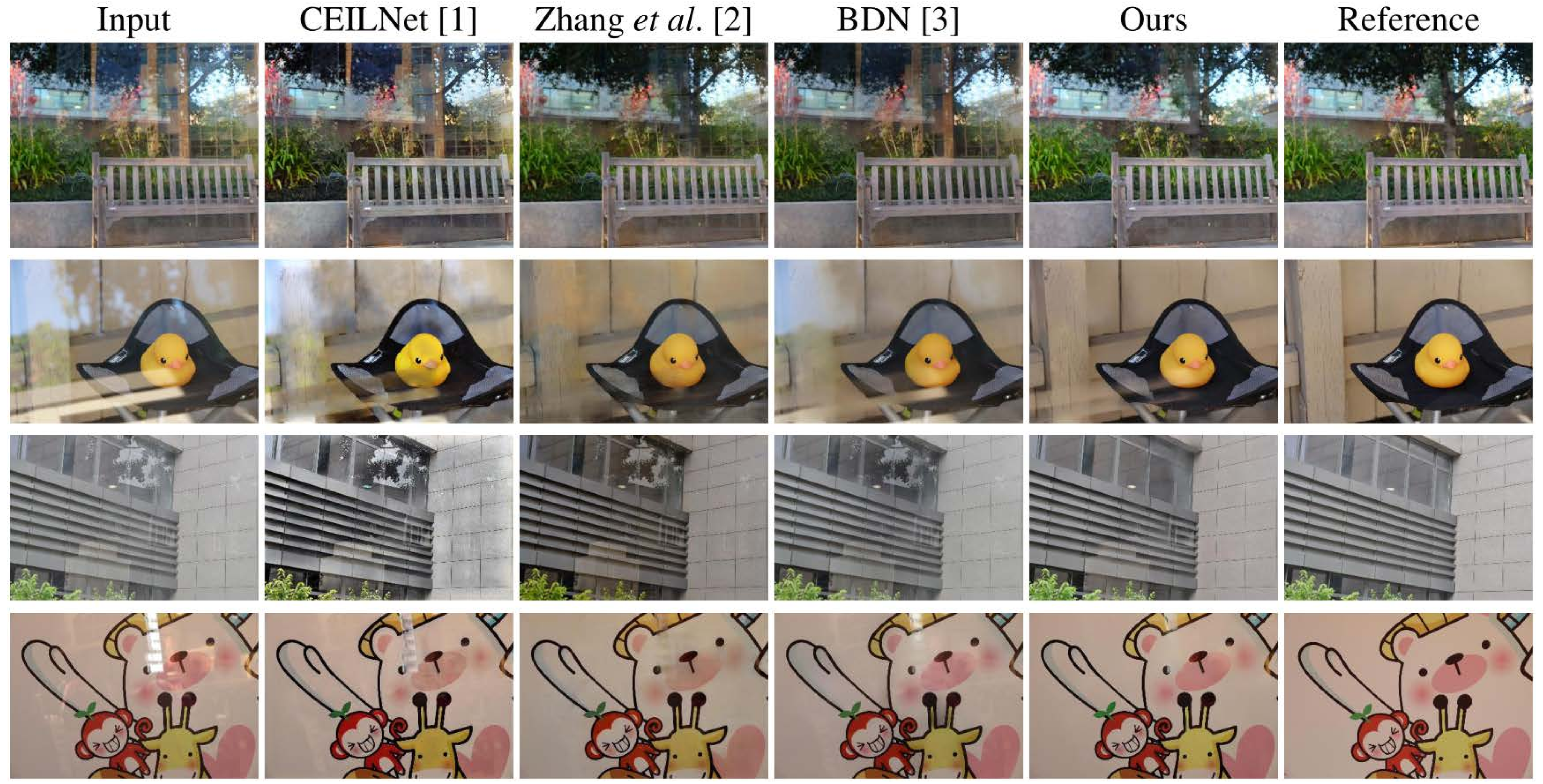
Results finetuned with feature loss based upon different layers of VGG-19. Only the highest-level feature ('conv5_2') yields satisfactory result.



Misalignment toleration analysis

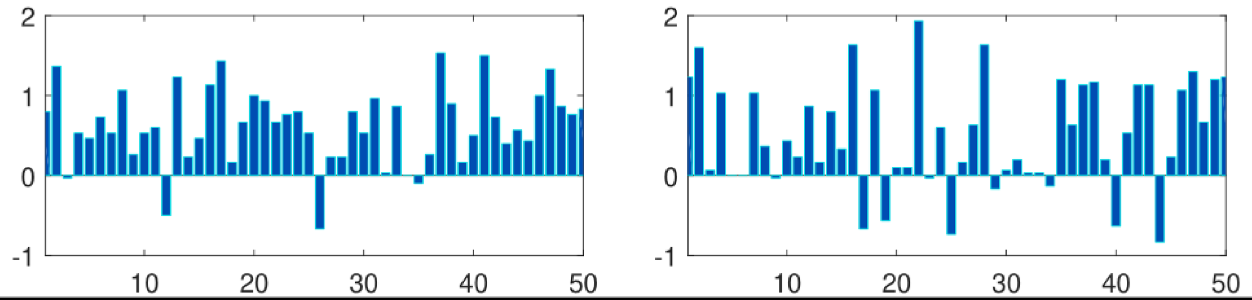
Experimental Results

Method Comparison on Real-world Images

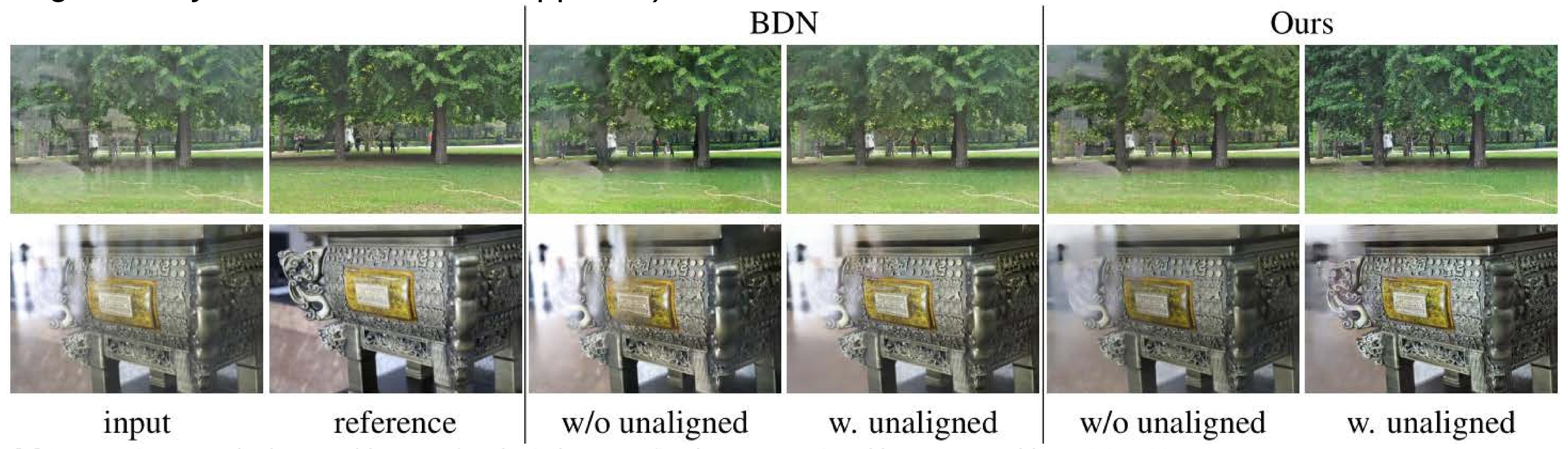


Finetuning with Unaligned Data

Score Range Ratio	BDN	Ours
$(0.25, 2]$	78%	54%
$[-0.25, 0.25]$	18%	36%
$[-2, -0.25)$	4%	10%
Average Score	0.62	0.51



Human preference scores of 50 tested image pairs. (2 indicating the finetuned result is significantly better while -2 the opposite)



[1] Fan et al. A generic deep architecture for single image reflection removal and image smoothing. ICCV, 2017.
[2] Zhang et al. Single image reflection separation with perceptual losses. CVPR, 2018.
[3] Yang et al. Seeing deeply and bidirectionally: A deep learning approach for single image reflection removal. ECCV, 2018.