UG[2]: a Video Benchmark for Assessing the Impact of Image Restoration and Enhancement on Automatic Visual Recognition

UNIVERSITY OF NOTRE DAME

*Rosaura G. Vidal[*1], Sreya Banerjee[*1], Klemen Grm[2], Vitomir Struc[2], Walter J. Scheirer[1]*

(1) Dept. of Computer Science & Engineering, University of Notre Dame; (2) University of Ljubljana, Slovenia

(= denotes equal contribution)*

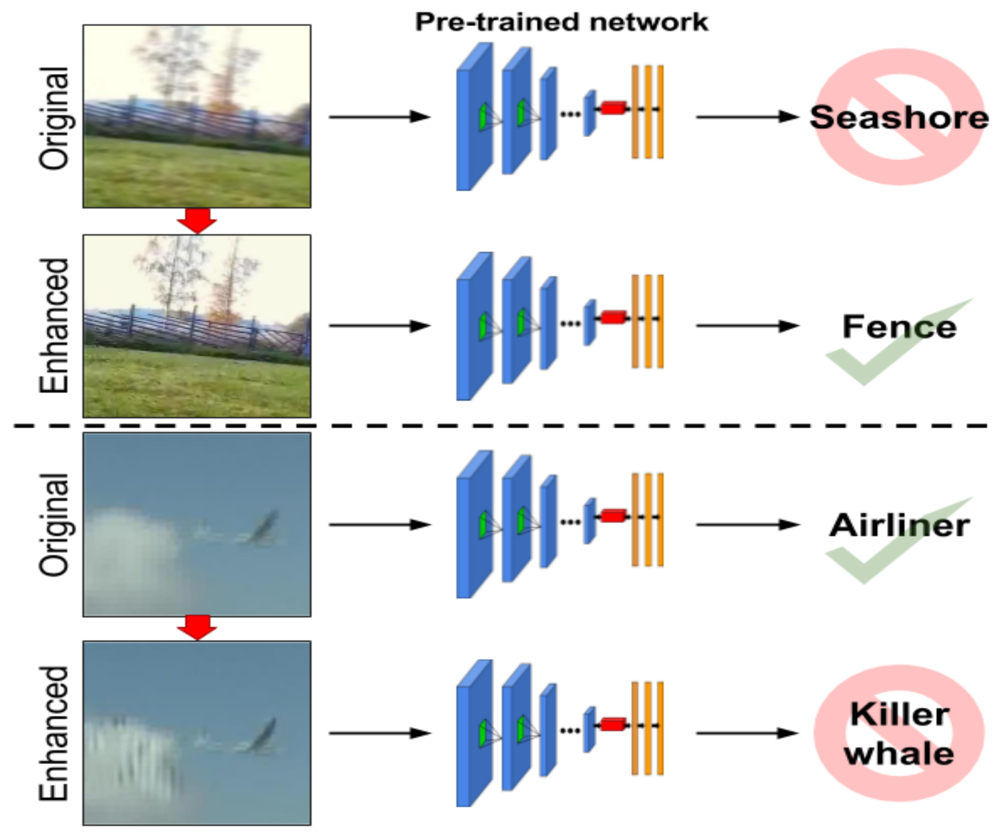
ID: 53

WACV 2018

Motivation

Deep learning-based algorithms are the "go to" method for automatic visual recognition systems.

But how do these methods perform when they are given input that have been processed by algorithms from computational photography which are meant to address artifacts such as blur, noise, and mis-focus?



(Top) In principle, image restoration and enhancement techniques should improve visual recognition performance by creating higher quality inputs for recognition models. This is the case when a Super Resolution Convolutional Neural Network[1] is applied to the image in this panel.

(Bottom) In practice, we often see the opposite effect --- especially when new artifacts are unintentionally introduced, as in this application of Deep Deblurring[3].

In order to answer this question, we:
- Introduce a ***new video benchmark dataset, UG²*** representing both ideal conditions and common aerial image artifacts.
- Evaluate the influence of image aberrations and other problematic conditions on object recognition models - VGG16, VGG19, InceptionV3 and ResNet50.
- Measure impact and suitability of basic and state-of-the-art image and video processing algorithms used in conjunction with common object recognition models.

UG² consists of three collections:
- 50 Creative Commons tagged videos taken by fixed-wing unmanned aerial vehicles (UAV) obtained from YouTube
- 61 videos recorded by pilots of fixed wing gliders.
- 178 controlled videos captured on the ground.

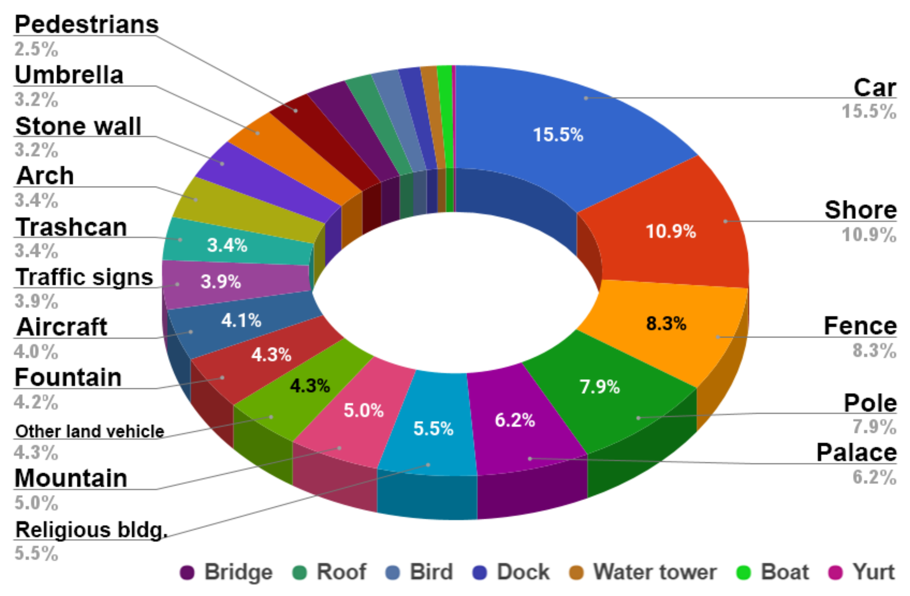
UG² can be found at: https://goo.gl/AjA6En

Acknowledgement: *Funding was provided under IARPA contract **#2016-16070500002**. This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA). The views and conclusions contained herein are those of the organizers and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.*

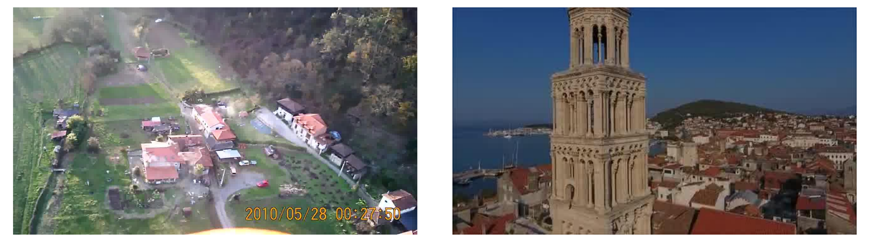
Dataset

289 videos have **1,217,496** frames of which **159, 464** frames have object-level annotations, representing **228** ImageNet classes, combined into **37** super-classes encompassing visually similar ImageNet categories and two additional classes for **Pedestrian** and **Resolution chart** images

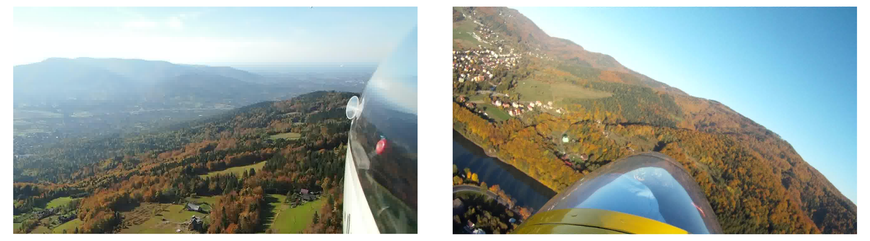
Shared class distribution



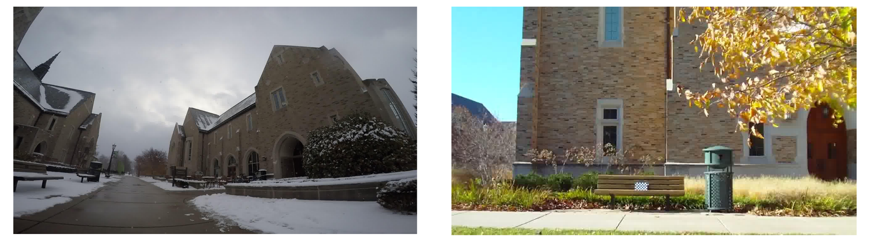
Shared class distribution of at least 2 UG² collections



UAV Collection



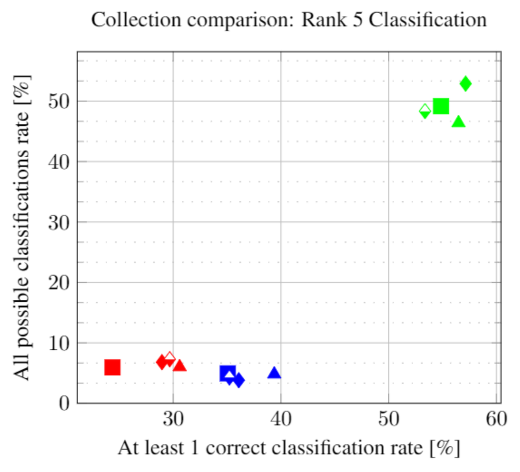
Glider Collection



Ground Collection

Experiments

Baseline Performance



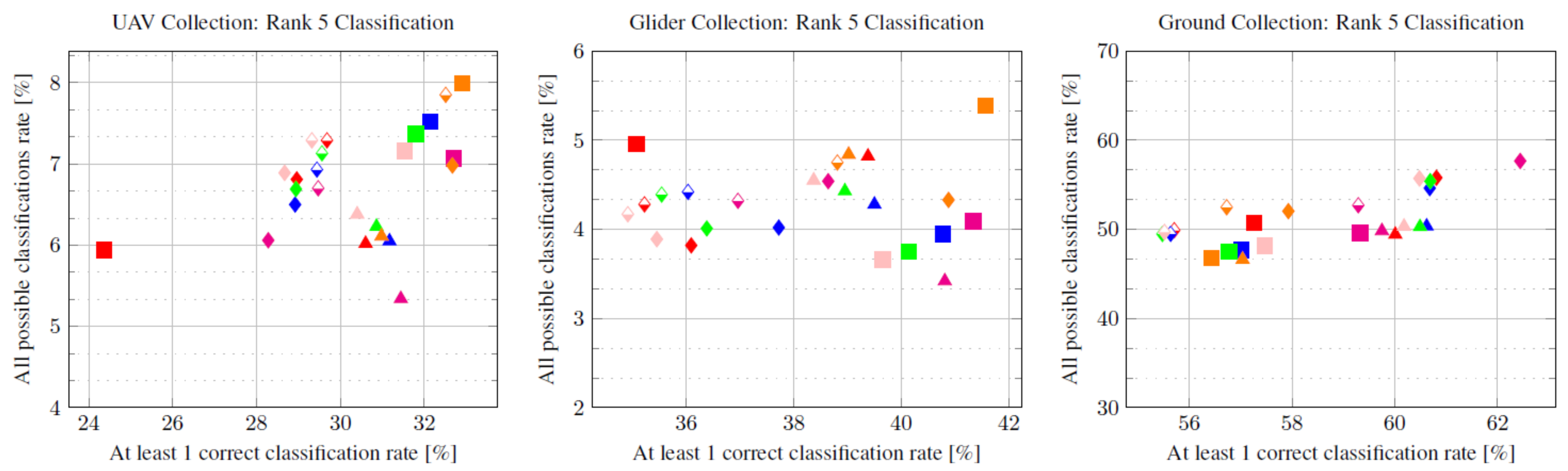
Legends

♦VGG16 ◊VGG19 ■Inception ▲ResNet

● Baseline ● Bilinear ● Bicubic ● NearestNeighbor ● VDSR ● SRCNN

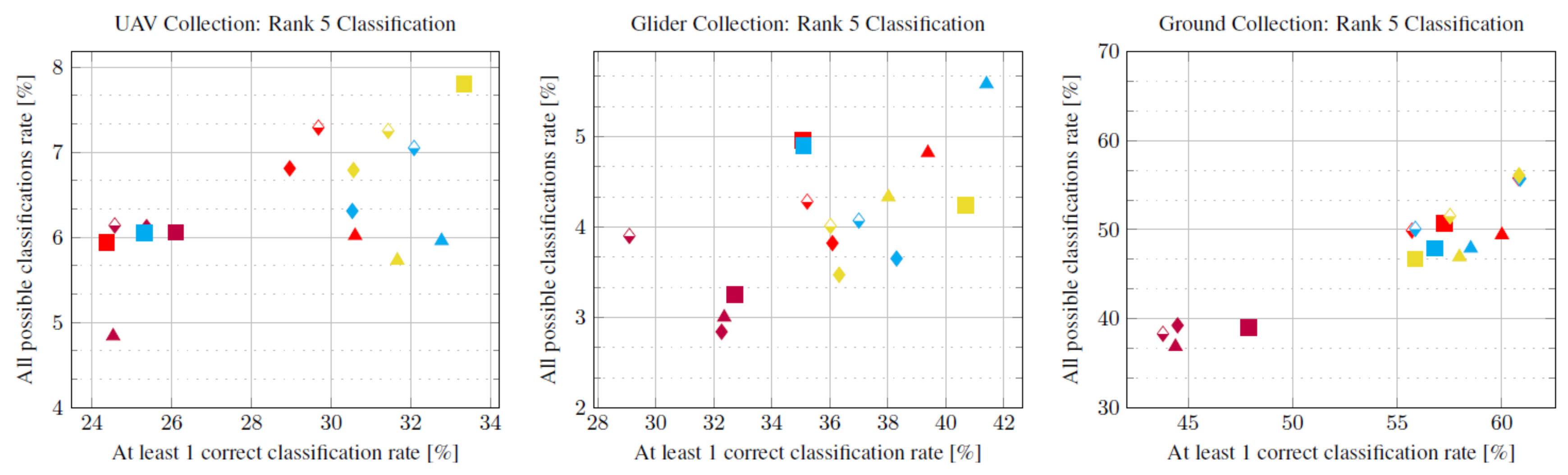
● UAV Collection ● Glider Collection ● Ground Collection

With Image Enhancement Methods[1,2]



Experiments

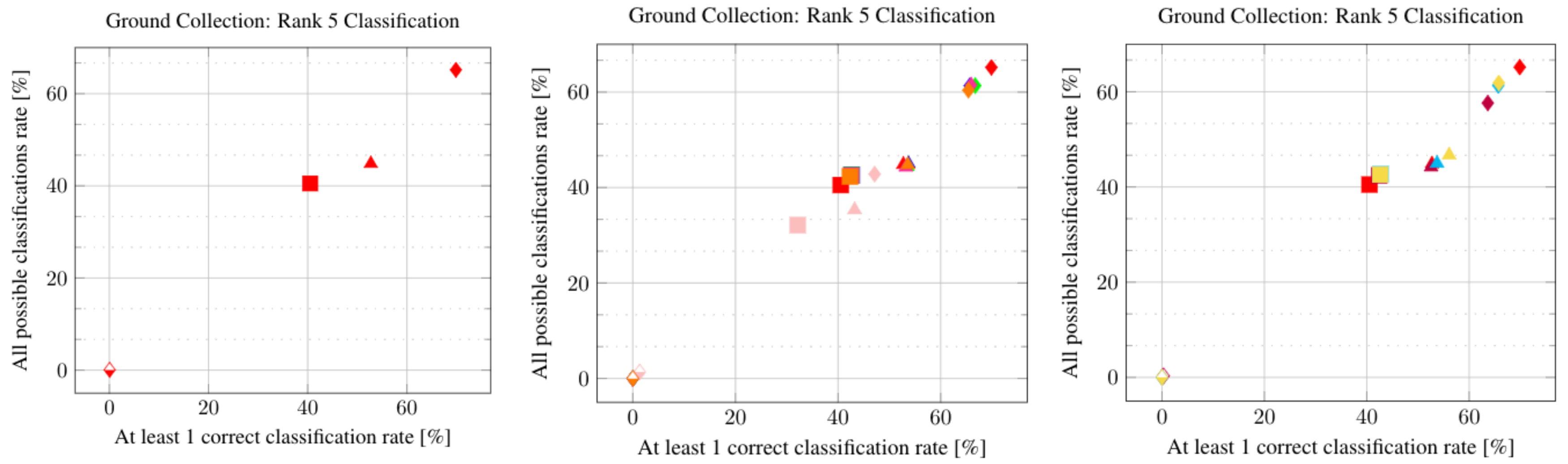
With Image Restoration (Deblurring) Techniques[3, 4]



♦VGG16 ◊VGG19 ■Inception ▲ResNet

● Baseline ● BlindDeconvolution ● DeepDeblurring ● DynamicDeepDeblurring

Effect of Weather on Classification



♦Cloudy ◊Rainy ■Snow ▲Sunny

● Baseline ● Bilinear ● Bicubic ● NearestNeighbor ● VDSR ● SRCNN

● Baseline ● BlindDeconvolution ● DeepDeblurring ● DynamicDeepDeblurring

UG² Prize Challenge

We are organizing a prize challenge at **CVPR 2018,** Salt Lake City, Utah in association with ODNI, IARPA. **$75,000** plus travel money can be won!
Please visit our website: ***http://www.ug2challenge.org/***

IARPA BE THE FUTURE

References

[1] Kim, Jiwon, et al. "Accurate image super-resolution using very deep convolutional networks." IEEE CVPR, 2016.
[2] Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." ECCV, 2014.
[3] S. Su, et al, "Deep Video Deblurring," IEEE CVPR, 2017.
[4] S. Nah, et al. "Deep multi-scale convolutional neural network for dynamic scene deblurring". CoRR, abs/1612.02177, 2016.