Fine-grained Recognition in the Noisy Wild: Sensitivity Analysis of Convolutional Neural Networks Approaches

Erik Rodner[1]
erik.rodner@uni-jena.de

Marcel Simon[1]
marcel.simon@uni-jena.de

Robert B. Fisher[2]
rbf@inf.ed.ac.uk

Joachim Denzler[1]
joachim.denzler@uni-jena.de

[1] Computer Vision Group
Friedrich Schiller University Jena
Germany
www.inf-cv.uni-jena.de

[2] University of Edinburgh
United Kingdom

Overview In this paper, we study the sensitivity of CNN outputs with respect to image transformations and noise in the area of fine-grained recognition. We answer the following questions: (1) how sensitive are CNNs with respect to image transformations encountered during wild image capture?; (2) can we increase the robustness of CNNs with respect to image degradations? and (3) how can we predict CNN sensitivity?

To answer the first question, we provide an extensive empirical sensitivity analysis of common CNN architectures (AlexNet, VGG19, and GoogleNet) across various types of image degradations. We perturb test images of different datasets with noise types including Gaussian and pepper noise, random color shifts, and different geometric image transformations. This allows for predicting CNN performance for new domains comprised by images of lower quality or captured from a different viewpoint. Our experiments show that even small random noise can lead to a dramatic performance decrease.

The question naturally arises if it is possible to increase the robustness either during testing or by adapting the learning. we analyze two intuitive ideas for increasing robustness: data augmentation by applying input dropout to the training data and image pre-processing.

After the empirical analysis, the question remains whether we can quickly detect images with unstable CNN outputs. This question goes beyond a pure sensitivity study but asks for uncertainty estimates often available for Bayesian methods but not for CNNs. We present a novel approach (Figure 1) for estimating the sensitivity given an input using a first-order approximation of the output change.

Take-home message The experiments show that the influence especially of common intensity noise is severe even at low noise levels. The reason is a domain shift between noise-free training and perturbed test data. From our study, we can draw several conclusions:

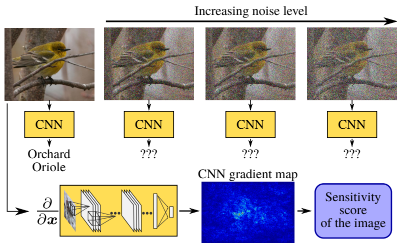

Figure 1: How sensitive are CNNs with respect to image noise and transformations? We study this question and show how to predict CNN sensitivity for a given image.

1. The training images should have the same noise level as the test images and care has to be taken even for small noise applied to intensities.
2. Data augmentation during training or image pre-processing are no solutions as they decrease the accuracy on noise-free images dramatically and are only beneficial for high noise levels at best.
3. Noise sensitivity depends on the CNN architecture and VGG19 has shown to be the most robust one.
4. Sensitivity of CNN outputs can be predicted for small noise levels with our technique allowing for uncertainty estimates of CNN outputs.

These conclusions can be seen as guidelines especially for developers of real-world applications, where, for example, cheap camera sensors deliver low quality images but the training was done on relatively noise-free datasets like ImageNet.