Why do deep convolutional networks generalize so poorly to small image transformations?

Summary: Convolutional Neural Networks (CNNs) are commonly assumed to be invariant to small image transformations: either because of the convolutional architecture or because they were trained using data augmentation. Recently, several authors have shown that this is not the case: small translations or rescalings of the input image can drastically change the network's prediction. In this paper, we quantify this phenomena and ask why neither the convolutional architecture nor data augmentation are sufficient to achieve the desired invariance. Specifically, we show that the convolutional architecture does not give invariance since architectures ignore the classical sampling theorem, and data augmentation does not give invariance because the CNNs learn to be invariant to transformations only for images that are very similar to typical images from the training set. We discuss two possible solutions to this problem: (1) antialiasing the intermediate representations and (2) increasing data augmentation and show that they provide only a partial solution at best. Taken together, our results indicate that the problem of insuring invariance to small image transformations in neural networks while preserving high accuracy remains unsolved.

MSC:
- 68T07 Artificial neural networks and deep learning
- 62H30 Classification and discrimination; cluster analysis (statistical aspects)
- 62M45 Neural nets and related approaches to inference from stochastic processes
- 68T05 Learning and adaptive systems in artificial intelligence

Keywords:
- machine learning
- deep convolutional neural networks
- generalization

Software:
- ScatNet

References:
[7] Why do deep convolutional networks generalize so poorly to small image transformations?
[38] Bolle Zhou, Aditya Khosla, ‘Agata Lapedriza, Aude Oliva, and Antonio Torralba. This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically...