

Propagating Targets through Noninvertible Layers of Deep Networks

Submission ID 3000313
Submission Type Poster
Topic Artificial Intelligence
Status Submitted
Submitter Jeffrey Seely
Affiliation Columbia University

SUBMISSION DETAILS

Presentation Type Either Poster or Oral Presentation

Presentation Abstract Summary Training deep neural networks relies on effective solutions for credit assignment. While backpropagation is the predominant approach, a method known as target propagation has been shown to be a viable alternative. Whereas backpropagation propagates gradients using the chain rule, target propagation propagates `targets' using inverses. Once targets are propagated to intermediate layers, layer-local cost functions are minimized to update network weights. However, since target propagation relies on inverses, it is not clear how to best define feedback functions of noninvertible layers. This is a fundamental challenge, since essentially all neural networks are noninjective thus noninvertible. Furthermore, noninjectivity is a desirable feature and not something that can or should be circumvented in general. Here, we explore a framework for propagating targets through noninvertible layers. We show that this approach can successfully train networks where noninvertibility is an obvious issue, such as autoencoders with a low-dimensional hidden layer.

Co-author Information

* Presenting Author

First Name	Last Name	Affiliation	E-mail
Jeffrey *	Seely *	Columbia University	jsseely@gmail.com
Raoul-Martin	Memmesheimer	University of Bonn	memmesheimer@fiis.uni-frankfurt.de
Larry	Abbott	Columbia University	labbott@columbia.edu

Keywords

Keywords

target propagation

deep learning

credit assignment