ELAD HOFFER

PERSONAL INFORMATION

Born in Israel, 11 October 1986

email	elad.hoffer@gmail.com
website	http://www.DeepLearning.co.il
phone	(M)054-4228891

PROFESSIONAL INTERESTS

Machine Learning, Deep Learning, Computer Vision, Signal Processing

WORK EXPERIENCE

2015–2016 Deep Learning researcher, ICRI-CI

Intel Deep Learning researcher at Intel's Collaborative Research Center for Computational Intelligence.

2013–2015 Visual Algorithms and Neural Networks, CVG

Intel Researching and developing Deep Learning capabilities for computer-vision tasks. Part of the Algorithms team in Intel's Computer-Vision Group.

2011-2013 Emulation student, Network Division

Intel Created FPGA prototypes and automation scripts for emulation of network devices.

2005-2009 Commanding Officer , Artillery Corps

IDF Rank: *Captain (Reserve duty)*
Served as a battery commander (artillery), leading 90 soldiers.
Currently serves on active reserve duty.

EDUCATION

2014-Present Technion, Israel Institute of Technology

PhD (Direct track) Electrical Engineering

Research: *Deep Learning of Representations*
Description: My research explores the machine learning technique known as "Deep Learning" which uses artificial neural networks to learn useful data representations.
Advisor: Prof. Daniel Soudry

2010-2014 Technion, Israel Institute of Technology

BSc Electrical Engineering GPA: 90 · *Cum Laude*
Specialized in Computer Engineering, Signal Processing.
Final Project: Real-Time Movie Subtitles Extraction - using image processing and computer-vision techniques.

2004 High-school "Itzhak Rabin", Gan-Yavne

Graduated with honors
Studied Computer Science and Physics
Final Project: Handwriting Recognition.

May 2015 **Deep metric learning using Triplet network**

ICLR 2015 Workshop

In this paper we propose the deep triplet network model, which aims to learn useful representations by distance comparisons. We demonstrate using various datasets that our model learns a better representation than that of its immediate competitor, the Siamese network, and discuss future possible usage as a framework for unsupervised learning.
Authors: Elad Hoffer, Nir Ailon

May 2016 **Deep unsupervised learning through spatial contrasting**

NIPS 2016 Workshop

In this work we present a novel approach for unsupervised training of Convolutional networks that is based on contrasting between spatial regions within images. This criterion can be employed within conventional neural networks and trained using standard techniques such as SGD and back-propagation, thus complementing supervised methods.
Authors: Elad Hoffer, Itay Hubara, Nir Ailon

November 2016 **Semi-supervised deep learning by metric embedding**

ICLR 2017 Workshop

In this work we explored a new training objective for deep networks that is targeting a semi-supervised regime with only a small subset of labeled data. This criterion is based on a deep metric embedding over distance relations within the set of labeled samples, together with constraints over the embeddings of the unlabeled set.
Authors: Elad Hoffer, Nir Ailon

May 2017 **Train longer, generalize better: closing the generalization gap in large batch training of neural networks**

NIPS 2017 Oral presentation

We examine the initial high learning rate training phase. We find that the weight distance from its initialization grows logarithmically with the number of weight updates. We therefore propose a ”random walk on random landscape” statistical model which is known to exhibit similar ”ultra-slow” diffusion behavior. Following this hypothesis we conducted experiments to show empirically that the ”generalization gap” stems from the relatively small number of updates rather than the batch size, and can be completely eliminated by adapting the training regime used.
Authors: Elad Hoffer, Itay Hubara, Daniel Soudry

April 2018 **Fix your classifier: the marginal value of training the last weight layer**

ICLR 2018

Neural networks are commonly used as models for classification for a wide variety of tasks. Typically, a learned affine transformation is placed at the end of such models, yielding a per-class value used for classification. This classifier can have a vast number of parameters, which grows linearly with the number of possible classes, thus requiring increasingly more resources. In this work we argue that this classifier can be fixed, up to a global scale constant, with little or no loss of accuracy for most tasks, allowing memory and computational benefits. Moreover, we show that by initializing the classifier with a Hadamard matrix we can speed up inference as well. We discuss the implications for current understanding of neural network models.
Authors: Elad Hoffer, Itay Hubara, Daniel Soudry

April 2018 **The Implicit Bias of Gradient Descent on Separable Data**

ICLR 2018

We show that gradient descent on an unregularized logistic regression problem with separable data converges to the max-margin solution. The result generalizes also to other monotone decreasing loss functions with an infimum

at infinity, and we also discuss a multi-class generalizations to the cross entropy loss. Furthermore, we show this convergence is very slow, and only logarithmic in the convergence of the loss itself. This can help explain the benefit of continuing to optimize the logistic or cross-entropy loss even after the training error is zero and the training loss is extremely small, and, as we show, even if the validation loss increases. Our methodology can also aid in understanding implicit regularization in more complex models and with other optimization methods.
Authors: Daniel Soudry, Elad Hoffer, Nathan Srebro

COMPUTER SKILLS

Programming Python, C++, Matlab, Lua, Perl, CUDA, Erlang, Verilog, Julia

Environments Linux, Microsoft Windows, Microsoft Office

Other PyTorch/Torch, TensorFlow, Open-CV, LaTeX

OTHER INFORMATION

Awards 2010-2013 · Dean's honor list - Technion Electrical Engineering Dept.

2018 · The Porat Award for PhD students

Languages Hebrew · Native

English · Fluent

August 10, 2018