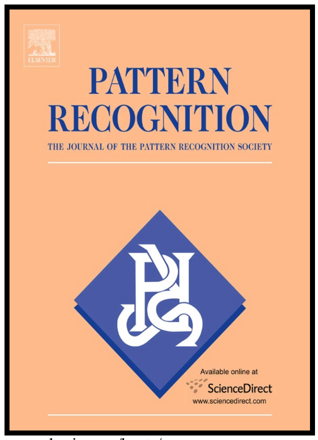
Author's Accepted Manuscript



www.elsevier.com/locate/pr

Comparing Two Classes of End-to-End Machine-Learning Models in Lung Nodule Detection and Classification: MTANNs vs. CNNs

Nima Tajbakhsh, Kenji Suzuki

PII: S0031-3203(16)30279-5
DOI: http://dx.doi.org/10.1016/j.patcog.2016.09.029
Reference: PR5891

To appear in: *Pattern Recognition*

Received date: 30 January 2016
Revised date: 4 September 2016
Accepted date: 21 September 2016

Cite this article as: Nima Tajbakhsh and Kenji Suzuki, Comparing Two Classe of End-to-End Machine-Learning Models in Lung Nodule Detection an Classification: MTANNs vs. CNNs, *Pattern Recognitior* http://dx.doi.org/10.1016/j.patcog.2016.09.029

Comparing Two Classes of End-to-End Machine-Learning Models in Lung Nodule Detection and Classification: MTANNs vs. CNNs

Nima Tajbakhsh and Kenji Suzuki[a],

[a]*Department of Electrical and Computer Engineering, Illinois Institute of Technology,
3300 S Federal St, Chicago, IL 60616, USA (e-mail: {ntajbakh, ksuzuki}@iit.edu).*

Abstract

End-to-end learning machines enable a direct mapping from the raw input data to the desired outputs, eliminating the need for hand-crafted features. Despite less engineering effort than the hand-crafted counterparts, these learning machines achieve extremely good results for many computer vision and medical image analysis tasks. Two dominant classes of end-to-end learning machines are massive-training artificial neural networks (MTANNs) and convolutional neural networks (CNNs). Although MTANNs have been actively used for a number of medical image analysis tasks over the past two decades, CNNs have recently gained popularity in the field of medical imaging. In this study, we have compared these two successful learning machines both experimentally and theoretically. For that purpose, we considered two well-studied topics in the field of medical image analysis: detection of lung nodules and distinction between benign and malignant lung nodules in computed tomography (CT). For a thorough analysis, we used 2 optimized MTANN architectures and 4 distinct CNN architectures that have different depths. Our experiments demonstrated that the performance of MTANNs was substantially higher than that of CNN when using only limited training data. With a larger training dataset, the performance gap became less evident even though the margin was still significant. Specifically, for nodule detection, MTANNs generated 2.7 false positives per patient at 100% sensitivity, which was significantly (p<.05) lower than the best performing CNN model with 22.7 false positives per patient at the same level of sensitivity. For nodule classification, MTANNs yielded an area under the receiver-operating-characteristic curve (AUC) of 0.8806 (95% CI: 0.8389 to 0.9223), which was significantly (p<.05) greater than the best performing CNN model with an AUC of 0.7755 (95% CI: 0.7120 to 0.8270). Thus, with limited training data, MTANNs would be a suitable end-to-end machine-learning model for detection and classification of focal lesions that do not require high-level semantic features.

Keywords: deep learning, patch-based machine learning, image-based machine learning, massive-training artificial neural network, convolution neural network, focal lesions, classification, computer-aided diagnosis, lung nodules

1. Introduction

End-to-end learning machines are particular machine learning models that seek a direct mapping from the raw input image data to the target output, eliminating the need for the design of an intermediate feature space. As a result, such learning machines require less engineering effort and fewer user interventions to produce the desired outputs. Yet, they have achieved extremely good results for many computer vision tasks, breaking state-of-the-art performance records previously held by the heavily-engineered and hand-crafted approaches such as part-based models [1] and bag of visual words [2]. Although originally developed in the computer vision community, end-to-end machine-learning models have now found their ways to a variety of disciplines including natural language processing [3, 4, 5], drug discovery [6, 7, 8], and medical image analysis [9, 10, 11, 12, 13].

In this paper, we consider two classes of end-to-end learning machines, namely, massive-training artificial neural networks (MTANNs) and convolutional neural networks (CNNs). Although MTANNs have been actively used for a number of medical image analysis tasks over the past two decades, CNNs have recently emerged in the field of medical imaging, as a promising technique. It would be interesting to study how these two successful learning machines, which both stem from artificial neural networks (ANNs), compare to each other theoretically and experimentally. To that end, we investigated the performance of MTANNs and CNNs in lung nodule detection and classification, two well-studied topics in the field of medical image analysis. To our knowledge, no prior research has compared the effectiveness of MTANNs and CNNs in neither computer vision nor medical image analysis fields.

2. Material and Methods

2.1. Massive-Training Artificial Neural Networks (MTANNs)

As the extension of neural filters [14, 15], MTANNs can accommodate various pattern-recognition tasks [16, 17, 18] such as detection of focal lesions and classification of lesion types. MTANNs come at 2 major models: 1) 2D MTANNs, which are designed for processing 2D images, and 2) 3D MTANNs, which are the generalized form of 2D MTANNs and are designed for processing volumetric data. The first appearance of