

Review Report

Review Cycle Up to date
Submitted on 6/8/16 3:00 PM
Evaluation Indifferent
Confidence Medium
Argumentation Reviewer's Feedback Summary (0=weak .. 5=strong):

Originality	2
Quality	2
Presentation	3
Relevance to ICANN	4
Reviewer's confidence	3

General Contents Summary, Comments and Feedback to Author:

As the title suggests, the paper seeks to compare the performance of the LeNet and GoogLeNet on a set of 23 image classification problems first introduced in F. Fleuret et al., Comparing machines and humans on a visual categorization test. PNAS, 108(43):17621–17625, 2011, to measure human and machine performance on visual categorization tasks with the overall goal of determining whether GoogLeNet performs better performance LeNet.

To check this, the authors use the Caffe implementations of LeNet and GoogLeNet, with some hyperparameter changes and ADAM optimization procedure.

The authors report that, over the entire problem set, GoogLeNet's performance only marginally improves on that of LeNet. However, it seems that a problem subset requires shape comparisons to solve the classification problem, which appears to be outside the reach of both CNN methods. On the other subset, GoogLeNet's performance clearly improves on that of LeNet. Human performances are also considered, although not those reported on the PNAS paper but on an adaptation the authors consider better suited to a Machine Learning setting.

The authors seem to conclude that CNNs generally appear not to be able to solve problems containing shape comparison, and even when they do, this may be done to some hidden problem feature due to side effects on the image generation process.

This seems to be a reasonable suggestion but one may find the paper somewhat inconclusive and not only from a man-machine perspective but perhaps more importantly on the CNN models studied. GoogLeNet

is a clear successor of LeNet but there have been many other recent proposals of CNNs having a better performance at least on recent ImageNet competitions (AlexNet, Inception, VGG, several residual nets; although submitted after the reviewed paper, see for instance Canziani et al, arXiv:1605.07678). The authors seem to somewhat agree on this on their discussion that can be seen as opening several new questions but not closing others.

Review Report

Review Cycle Up to date
Submitted on 6/8/16 3:00 PM
Evaluation Weak Accept
Confidence High
Argumentation Reviewer's Feedback Summary (0=weak .. 5=strong):

Originality : 3
Quality : 3
Presentation : 2
Relevance to ICANN : 4
Reviewer's confidence : 5

The authors compare two well known methods which employ convolutional neural networks (CNNs), for classifying images into abstract classes. They compare the performance of LeNet (an older method) to that of GoogLeNet (a more recent proposal) at classifying randomly generated images, which are differentiated by an abstract property and, in this way, analyse the progress made in the area of CNNs. They compare the results of the CNNs with a framework presented by Fleuret and collaborators, where the performance of humans is also considered.

In the problems which they analysed, they show that problems which require comparison of shapes are not well solved by either of the CNNs, whereas humans and the “boosting method” employed by Fleuret and collaborators perform well for this class of problems. CNNs perform well in problems which do not require comparison of shapes.

In the second paragraph of page 5 (and later in the paper), the authors explain and sustain that the performance measurements of CNNs for some problems was impaired by the possibility that the generation process of the images imparts some unwanted pattern into the images. The fact that the accuracy does not change for a modified problem was attributed to the fact that the “the CNNs are exploiting some unintended pattern in the data and comparing the

shapes does not contribute to the classification”. It is necessary to verify/show that other causes do not lead to the same result, i. e. the authors need to substantiate their statement.

The text needs some improvements in order to allow good understanding and to conform to LNCS formatting requirements.

1. It is necessary to define all acronyms in the paper the first time they are used. For example, SVRT is not well defined in the paper. In the case of SVRT, it is also important to give a brief description.
2. The caption of Table 1 needs to be rewritten for clear understanding.
3. References need to be written in LNCS format and order.
4. The structure of some paragraphs needs to be reviewed. Some paragraphs have only a single small phrase.

Final Report

Review Cycle Up to date

Submitted on 6/10/16 3:44 PM

Argumentation Dear Authors, The paper is Accepted pending revisions indicated by the reviewers. Please insert your revisions without a need for a separate explanation. Please resubmit the revised version as soon as possible ! The final assignment to a specific session will be determined at a later stage. Thank you for your participation to ICANN 2016. -- ICANN 2016 Program Committee

IMPORTANT NOTICE ***** Only contributions from registered presenters will be included in the Springer volume. Register to the conference as soon as possible before June 27th to enjoy the early registration fee and to include your contribution in the Lecture Notes in Computer Science volume.
