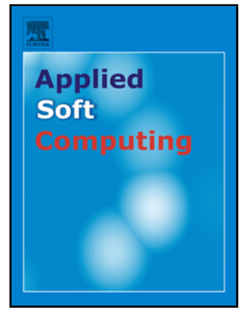
Accepted Manuscript



Title: Semi-supervised Deep Rule-based Approach for Image Classification

Authors: Xiaowei Gu, Plamen P. Angelov

PII: S1568-4946(18)30154-6
DOI: https://doi.org/10.1016/j.asoc.2018.03.032
Reference: ASOC 4781

To appear in: *Applied Soft Computing*

Received date: 10-10-2017
Revised date: 9-1-2018
Accepted date: 21-3-2018

Please cite this article as: Xiaowei Gu, Plamen P.Angelov, Semi-supervised Deep Rule-based Approach for Image Classification, Applied Soft Computing Journal https://doi.org/10.1016/j.asoc.2018.03.032

Semi-supervised Deep Rule-based Approach for Image Classification

Xiaowei Gu[1] and Plamen P. Angelov[1,2*]

[1]School of Computing and Communications, Lancaster University, Lancaster, LA1 4WA, UK

[2]Technical University, Sofia, 1000, Bulgaria (Honorary Professor)

e-mail: {x.gu3, p.angelov}@lancaster.ac.uk

Highlights

- A novel semi-supervised deep rule-based (SSDRB) classifier with a prototype-based nature is introduced;
- The SSDRB classifier is able to generate human interpretable IF…THEN… rules through the semi-supervised learning process in a self-organising and highly transparent manner;
- The SSDRB classifier is able to perform classification on out-of-sample images;
- The SSDRB classifier outperforms the state-of-art approaches in classification accuracy.

Abstract- In this paper, a semi-supervised learning approach based on a deep rule-based (DRB) classifier is introduced. With its unique prototype-based nature, the semi-supervised DRB (SSDRB) classifier is able to generate human interpretable IF…THEN… rules through the semi-supervised learning process in a self-organising and highly transparent manner. It supports online learning on a sample-by-sample basis or on a chunk-by-chunk basis. It is also able to perform classification on out-of-sample images. Moreover, the SSDRB classifier can learn new classes from unlabelled images in an active way becoming dynamically self-evolving. Numerical examples based on large-scale benchmark image sets demonstrate the strong performance of the proposed SSDRB classifier as well as its distinctive features compared with the "state-of-the-art" approaches.

Keywords- semi-supervised learning; deep rule-based (DRB) classifier; prototype-based models; fuzzy rules; self-organising classifier; transparency and interpretability.

1. Introduction

Due to the more accessible electronics and the boom in the information technologies, an astronomic amount of images are being produced and uploaded on the web every day. Studying and classifying these images is of paramount importance [1]. Facing the strong demand, however, traditional fully supervised machine learning approaches are insufficient [2], [3]. In most real applications, labelled images are scarce and it is expensive to obtain them. Despite the fact that unlabelled images are abundant, supervised approaches are not able to effectively use them.

Semi-supervised machine learning approaches [4]–[9] consider both the labelled and unlabelled data. The goal of the semi-supervised learning is to use the unlabelled data to improve the generalisation. Cluster assumption states that the decision boundary should not cross high density regions, but lie in low density regions [6]. Virtually all the existing successful semi-supervised approaches rely on the cluster assumption in a direct or indirect way from estimating or optimizing a smooth classification function over labelled and unlabelled data [10], [11].

There are two major branches of semi-supervised approaches, SVM-based and graph-based approaches [8], [12]. Semi-supervised SVMs [13]–[15] are extensions of the traditional SVMs [2] to a semi-supervised scenario. Traditional SVMs maximise the separation between classes based on the training data via a maximum-margin hyperplane [2], while semi-supervised classifiers balance the estimated maximum-margin hyperplane with a separation of all the data through the low-density regions. Graph-based approaches [8], [12], [16] use the labelled and unlabelled data as vertices in a graph and build pairwise edges between the vertices weighted by similarities. In general, both types of semi-supervised approaches are computationally expensive and they consume a lot of computer memory, thus, are not suitable for large-scale datasets. Moreover, they are not applicable to the out-of-sample data and require full retraining when more training samples are given [12].

*Corresponding Author