Advances in Knowledge-Based Technologies

Proceedings of the
Master and PhD Seminar
Summer term 2019, part 2

Software Competence Center Hagenberg
SCCH Room 0/2
1 July 2019

Software Competence Center Hagenberg
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 800
Fax +43 7236 3343 888
www.scch.at

Fuzzy Logic Laboratorium Linz
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 431
Fax +43 7236 3343 434
www.flll.jku.at

Deep Learning Approaches to OCT-Image Classification for Technical Materials

Laura Peham

Department of Knowledge-Based Mathematical Systems - Johannes Kepler University Linz
and
RECENDT - Research Center for Non Destructive Testing GmbH
Altenberger Straße 69, 4040 Linz, Austria

June 27, 2019

Abstract

Optical coherence tomography (OCT) is a non destructive and non-contacting imaging technology and can visualize the internal structures of various materials. In recent years it became popular for industrial applications for the non destructive testing of materials.
Machine learning models are able to learn from very complex datasets and make predictions for new data based on this knowledge without the need of user-defined mathematical rules describing the task to solve. In my thesis I will use three different machine learning methods to examine the automatic classification of OCT-images in the context of technical materials. **Deep learning** is a machine learning method which deals with deep artificial neural networks and one big advantage is that it can take raw images as input and there is no need to manually extract features out of them. On the other hand there are many machine learning methods with the need to do so, for such methods it is required to think about appropriate features beforehand. **Random Forests** consist of decision trees, which are basically trees which split the data according to specific features at the nodes and ending up at labelled leafs. **Support vector machines** try to find a border to separate the data according to the classes and then maximize the margin of the border to get decision borders which easily separate the different classes.
The goal of my master thesis is to apply these three methods to OCT-images of technical materials and compare the results. The first task is to categorize 3D-printed objects with respect to their material (expressed by different color pigments) with only their (greyscale) OCT-image as input. Given are objects in the colors green, grey, red and transparent with two different objects in each color where in most of the cases the OCT-recordings of these four colors are well distinguishable. Therefore I recorded 12800 images (3200 per color) of different areas on the objects including surface defects, interior defects, plain surfaces and different slopes. After parameter tuning the results on the test set are the following: Deep Learning reaches an accuracy of 99,49%, Random Forests an accuracy of 95,29% and Support Vector Machine and accuracy of 95,19%. For Deep Learning it is possible to visualize what the network learns in creating images for each unit to which the unit is maximally responsive to. For the other methods there is unfortunately no good way to visualize the obtained models.
The second task will be the categorization of OCT-images of different materials with respect to their coating-thickness. This will be split in two sub-problems: The first will be coating on paper like different envelopes or cigarette paper and it will be a binary classifier distinguishing between uncoated and coated. The second will concern the coating on pills simulated with hairspray on Dextro Energy and PEZ. Here the classifier should be able to distinguish between uncoated, coated and thick-coated.

Moment-Based Domain Adaptation: Learning Bounds and Algorithms

Werner Zellinger

2019-07-01

Abstract

Domain adaptation is a statistical problem that generalizes the problem of learning from finitely many examples. The classical assumption of identically distributed examples is relaxed by allowing the future data to follow a distribution different from the one of the training data. In this work we study the problem of domain adaptation under a postulate of weak similarity between training and future distribution. In particular, we study the problem of domain adaptation by only assuming that finitely many moments of the training and future distribution are similar. We underpin the practical relevance of our postulate by constructing new moment-based algorithms and testing its performance on several large scale datasets for domain adaptation. Theoretical bounds on the generalization error of statistical classifiers are provided. Our ideas are applied on industrial problems of cyclical manufacturing, analytical chemistry and generative image modeling. Our main findings are two-fold: (a) Moment-based algorithms often outperform algorithms based on stronger concepts of similarity, and, (b) a high performance of moment-based domain adaptation algorithms can be expected for large samples of smooth distributions.

Explainable AI in the case of Learning Problems with an Interpretable Base Model

By Florian Sobieczky
Software Competence Center Hagenberg, A-4232 Hagenberg

After reviewing AI scenarios involving reference to the physical sciences, we propose a model agnostic scheme to achieve interpretability for an AI learning model which delivers its explanations in the framework of an underlying physical or statistical model. Examples are presented from various problem-sets, including classification and regression.