Review of PhD Thesis

Neural Network-Based Generative Models For Anomaly Detection

by Vít Škvára

The PhD thesis is submitted in the form of a text document, some of whose chapters are heavily based on previously published articles of the applicant. The text is of the highest formal quality – language and typesetting are virtually error-free. The quality of the illustrations is really outstanding. The text cites a very high number of very relevant literary sources. The structure of the text is adequate.

The review of the state of the art and of the knowledge that the thesis deals with (namely neural networks, various flavours of autoencoders and generative NNs) is very detailed and of high quality. It is very convincing that the applicant has a very good knowledge of his narrower (as well as wider) field and that he is able to handle mathematics and terminology, and to express ideas in the domain.

The scientific contribution of the applicant is presented mostly in chapters 4 and 5 which seem to be based on two previously written articles (the second one currently under review?).

Chapter 4 does not present new scientific methodology, but rather ("empirically") compares existing approaches to anomaly detection. The reviewer finds it difficult to draw solid and noteworthy conclusions there. The datasets are mostly very small (low resolution of the samples, low variability of the data, low count of samples) and somewhat obsolete and in the context of current computer vision trends rather uninteresting. The comparison shows that some of the anomaly detectors perform better than others, but the reviewer does not see important findings that would improve understanding of what is happening in the black-box models or that would show distinct paths for future research.

Chapter 5 presents the main scientific contribution of the author: a new approach to "semantic" anomaly detection based on generative adversarial networks and autoencoders and their variants. The datasets that the proposed design is tested on are mostly synthetic and small (meaning both sample size/resolution and sample count). That makes the evaluation and the overall contribution somewhat questionable – it is difficult to estimate how valuable the proposed solution would be when facing real-world problems and visual data of real-world properties (the image resolution in industrial quality inspection, which is a prominent application domain of visual anomaly detection, is by orders of magnitude somewhere else than the datasets used for benchmarking). It is perfectly legitimate to use small-scale datasets for the development of new theoretical concepts and for shedding light on the properties of studied machine learning models. Such insights are not vivid in the discussed chapter 5 of the thesis. The practical applicability (and worthiness of further study and consideration) of the proposed methodology would be better shown by using data sets that are not synthetic and close in their proportions to the real-world imagery data used today.

The topic of the doctoral thesis is relevant and timely – anomaly detection still is a very valid subject from both theoretical and application perspectives and the development of smart deep-learning models seems to be the way to attack the problem. The applicant – in his thesis – shows good knowledge of the state of the art in (visual) machine learning. It is clear that he carefully studied approaches that are available and sought to transform his acquired knowledge and understanding into a new arrangement of neural networks that would overcome present limitations and outperform existing solutions. However, the scientific contribution presented in chapters 4 and 5 of the thesis is not overwhelming. Apparently, the newly proposed architecture presented in chapter 5 has not yet been published in a peer-reviewed manner.

Questions to be answered during the defense of the thesis:

1. Has the main contribution of the thesis, presented in chapter 5, been published in a peer-reviewed journal or conference proceedings?
2. How to evaluate the applicability of the presented contribution to real-world problems?
3. Figure 1.1c illustrates the meaning of the term “semantic anomaly”. Does the methodology proposed in chapter 5 truly address this class of anomalies? How is it demonstrated / confirmed in the experimental evaluation?
4. Section 5.4.5, namely footnote 4, prefers leave-one-in scenario over the leave-one-out, even in contrast with the literature [52]. The word “anomaly” itself suggests that the sample of interest should be something rare, *many* behaviours or appearances should be perceived as normal, anomalies are typically scarce and out-standing samples. Anomaly detection data sets and problems usually contain large amounts of normal data and a few anomalies. The leave-one-out would thus be a natural way to evaluate (in accordance with [52] and other sources). Choosing the leave-one-in scenario is thus unusual and raises concerns. Can you defend this choice? Can you also provide evaluation of your method in leave-one-out manner? Does training / evaluation in this way constitute a problem for your method?

The applicant seems to be very knowledgeable in machine learning for computer vision and in (visual) anomaly detection. The text of the thesis is both from the formal and from the factual point of view of high quality. However, it appears that the main contribution of the thesis has not been sufficiently published in peer-reviewed scientific media. The evaluation of the core method for “semantic anomaly detection” is only evaluated on synthetic data. These facts constitute certain reservations about recommending the defense of the thesis.

During the defense, the applicant should clarify these concerns and answer the questions raised explicitly above.

In Brno, February 26th, 2024

prof. Ing. Adam Herout, PhD.