Abstract of the doctoral thesis

Thesis title: Deep learning methods for computer vision-based industrial inspection applications

Author: mgr inż. Przemysław Dolata

Computer vision is finding increasingly more uses in the automation of industrial inspection tasks. Methods based on deep learning have unlocked even more application avenues, previously inaccessible to classical approaches. The concept of transfer learning and the growing availability of large, high-performance vision models have made it easier than ever before to deploy computer vision systems in practice. But despite that progress, there still exist domains where off-the-shelf deep learning models do not adapt with similarly superlative performance, necessitating further research.

The goal of this thesis is to develop specialized deep learning methods for computer vision-based industrial inspection applications in domains where the standard transfer learning approach does not yield satisfying results. Such application domains include agriculture, food processing, or certain process control scenarios. Due to the unique requirements or limitations in those domains, the generic deep models either do not transfer well or do not fully solve the vision problems. In order to achieve the goal of the thesis, three research questions were formulated and answered in a series of six publications.

The first research question concerned the design of convolutional neural network-based image classifiers for the special *multi-view* classification setting, in which the sample to be classified comprises a tuple of images. Such scenarios arise in certain quality inspection problems, where objects need to be imaged using multiple cameras, for example, in the food industry. Standard classification models are not able to reason from multiple images without substantial, non-trivial modifications. This question was addressed in Publications 1 and 2, discussing the design and training of double-stream convolutional neural networks. Publication 1 explores methods of fusing information from multiple inputs, while Publication 2 proposes a special case of a viewpoint-specific classification model.

The second research question was related to the optimization of the U-Net for semantic segmentation in low-latency applications. Certain tasks of industrial process control, for example, the monitoring of laser processes in additive manufacturing, could be automated using computer vision. But those tasks require models with exceptionally high inference speeds, since real-time processing is crucial for control. The U-Net, while offering high segmentation quality, is too computationally inefficient for this purpose. An architectural optimization method was proposed in Publication 3, which was awarded the Best Paper award at FedCSIS 2019.

The final research question involved the applications of deep learning-based instance segmentation methods for the purpose of size estimation of small objects in dense scenes from single images. The small size of objects, their large number and high density, often causing occlusions, and the need to accurately estimate their size, even from non-3D images, are the frequent challenges in industrial quality inspection, for example, in agriculture or food processing. At the same time, generic deep models trained on large datasets struggle significantly in those conditions. Three sub-questions were formulated in order to study those issues, and three Publications (4, 5, and 6) were presented to address them. The first sub-question, addressed in Publication 4, concerned the problem of size estimation of small three-dimensional objects in single images without depth information. The second

sub-question asked for a method of training *amodal* instance segmentation models, that is, models capable of segmenting occluded objects, without amodal data annotations; such an approach was presented in Publication 5. The third sub-question was related to improving the detection of small objects in high-density images, with a novel detection method proposed in Publication 6.

In total, this thesis summarizes findings from six Publications, all of which I am the first author of. To date, those Publications have been cited 37 times, confirming their academic impact. Additionally, four of them were a result of research activity funded by research and development grants, highlighting their practical value in the field of industrial inspection.