Master Thesis Proposal

Start: **January 2023**
Contact: **Xiaomeng Zhu (xiazhu@kth.se)**

Topic 1:
Object detection for industrial quality inspection.

Description:
A possible solution to solve the data limitation problem for deep learning in industrial quality inspection is using synthetic data generated from CAD models for training.

In previous research, we have found some classification models that can achieve good performance while training on synthetic data and testing on real data, see Fig. 1. The models are the self-supervised learning model DINO [1], and the state-of-the-art CNN network ConvNet [2].

Fig. 1. Samples of synthetic data and real data.

However, in order to use those classification models for automatic quality inspection in production, we need to convert them into object detection models. Therefore, this thesis work aims to convert those classification models, e.g., DINO [1] and ConvNet [2], into object detection models. The built object detection models will be evaluated on real industry use cases.

Moreover, this thesis work will try to understand how to convert different kinds of classification networks (CNN/transformer) into object detection networks.

Topic 2:
Generating random textures and backgrounds for synthetic industry CAD data.

Description:
A possible solution to solve the data limitation problem for deep learning in industrial quality inspection is using synthetic data generated from CAD models for training.

Currently, there are open-source synthetic training images available, as shown in Fig. 2, Visda2017 [3]; There are also synthetic training images generated from industry CAD models available, as shown in Fig. 3. However, these data are, in general, generated without backgrounds and textures.

a. synthetic image b. real image

Fig. 2. A sample data from Visda2017 [3].

a. synthetic image b. real image

Fig. 3. A sample data generated from an industry CAD model.

In order to achieve domain randomization, this thesis work aims to generate random textures and backgrounds for synthetic data. Some possible model suggestions can be: StyleGAN [4], Pix2Pixel [5], or other state-of-the-art diffusion and GAN models.

Moreover, the thesis work will try to understand how the random textures and backgrounds can help to narrow the domain gap between the synthetic data and real data.

Reference

[1] Caron M, Touvron H, Misra I, et al. Emerging properties in self-supervised vision transformers[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 9650-9660.

[2] Liu Z, Mao H, Wu C Y, et al. A convnet for the 2020s[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 11976-11986.

[3] Peng X, Usman B, Kaushik N, et al. Visda: The visual domain adaptation challenge[J]. arXiv preprint arXiv:1710.06924, 2017.

[4] Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 4401-4410.

[5] Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1125-1134.