

Stage weken voor de opleiding Master Informatica

Titel: Computer Vision: GANs for Synthetic Data

Gegevens bedrijf:

Naam: Vintecc

Tel: 0495928460

Externe promotor(en): Glenn

mailadres(sen): glenn@vintecc.com

Andere begeleiders: Frederik, Garben

mailadres(sen): frederik.vaneecke@vintecc.com

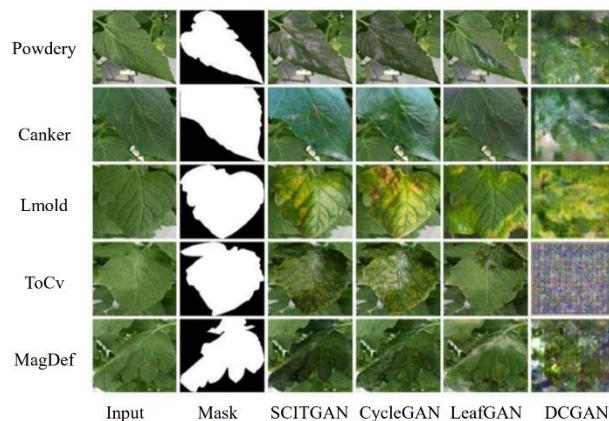
garben.tanghe@vintecc.com

Adres waar de student zal werken: Hof ter Weze 3, 8800 Roeselare

Kort (min 120 uur) of lang (min 200 uur): kort / lang

Korte beschrijving van de opdracht:

At Vintecc, we are often working on computer vision tasks related to agriculture processes. For these projects, we need a rich and diverse dataset to train our models on. Working with agricultural data, involving natural products, brings an additional challenge because of the high amount of variation possible within the data. For example, when building a potato detector, you should account for the fact that no two potatoes are the same. We are already using synthetic data as a solution for this type of problem where we render images with all possible variations. However, the problem partially remains since some effects are hard to simulate, think of dirt or sand on the potato. Therefore, we are looking into generative models that could learn these domain-specific features and include them within synthetic data as a form of data augmentation. A similar approach was taken by researchers working on disease detection on leaves, the image below shows how they could generate realistic data including disease-like features starting from real input data.



Goals

The goal of this internship is to create a data augmentation workflow that seamlessly plugs into an existing Python package. This package would enable us to easily create realistic-looking datasets, starting from synthetic data.

Methodology

In the first part, you would start with the existing literature on this topic. Then, you would quickly transition to the implementation part. A substantial portion of this internship would comprise of training GANs and visually comparing their results.

Technologieën die aan bod zullen komen:

Computer vision, deep learning, generative adversarial networks, synthetic data, style transfer