

ABSTRACT

DATA BALANCING IN TOMATO LEAF DISEASE DATASET USING COMBINATION OF GAN AND INPUT SPACE AUGMENTATION METHOD

Agriculture plays an important role in our daily life and is part of some countries main industries. These days, the farming industry demands computerized imaging innovation to solve some difficulties, especially plant disease as a major threat. It can aid in the early identification and classification of Tomato leaf diseases by agriculturalists. However, the datasets that are available and used as the input mostly are not sufficient or uneven in the class distribution, this can eventually lead to not maximum performance of the classification models.

This research aims to balance the dataset and at the same time find the best method on how to augment it by experimenting with Generative Adversarial Network and the traditional approaches. The final goal and metrics will be the increased performance of the model, so it can identify Tomato leaf disease effectively.

The results of this study reveal significant insights into the impact of various augmentation methods on model performance. By comparing input space augmentation and CycleGAN-generated images, the research highlights the effectiveness of CycleGAN only in enhancing classification accuracy through realistic data synthesis by 3.69%. However, it also uncovers the complexities and potential drawbacks of combining different augmentation techniques, such as increased background transformations leading to misclassification.

Keywords: Tomato Disease Detection, Learning for Tomato Disease, Tomato Disease

INTISARI

PENYEIMBANGAN DATA PADA DATASET DAUN TOMAT BERPENYAKIT MENGGUNAKAN GAN DAN INPUT SPACE SEBAGAI METODE AUGMENTASI

Pertanian memainkan peran penting dalam kehidupan sehari-hari kita dan merupakan bagian dari industri utama di beberapa negara. Saat ini, industri pertanian membutuhkan inovasi pencitraan terkomputerisasi untuk mengatasi beberapa kesulitan, terutama penyakit tanaman yang menjadi ancaman utama. Inovasi ini dapat membantu dalam identifikasi dini dan klasifikasi penyakit daun tomat oleh para ahli pertanian. Namun, dataset yang tersedia dan digunakan sebagai input sebagian besar tidak memadai atau tidak merata dalam distribusi kelas, yang pada akhirnya dapat menyebabkan kinerja model klasifikasi yang tidak maksimal.

Penelitian ini bertujuan untuk menyeimbangkan dataset dan sekaligus menemukan metode terbaik untuk augmentasi dengan menguji Generative Adversarial Network dan pendekatan tradisional. Tujuan akhir dan metrik yang digunakan adalah peningkatan kinerja model, sehingga dapat mengidentifikasi penyakit daun tomat secara efektif.

Hasil penelitian ini mengungkapkan wawasan signifikan tentang dampak berbagai metode augmentasi terhadap kinerja model. Dengan membandingkan augmentasi *input space* dan gambar yang dihasilkan oleh CycleGAN, penelitian ini menyoroti efektivitas CycleGAN sendiri dalam meningkatkan akurasi klasifikasi melalui sintesis data yang realistis sebesar 3,69%. Namun, penelitian ini juga mengungkapkan kompleksitas dan potensi kelemahan dalam menggabungkan teknik augmentasi yang berbeda, seperti perubahan *background* yang menyebabkan kesalahan klasifikasi.

Kata Kunci: Deteksi Penyakit Tanaman Tomat, Pembelajaran Mesin Penyakit Tomat, Penyakit Tomat