

ABSTRAK

Penyakit pada daun teh dapat menurunkan produktivitas dan kualitas hasil panen tanaman teh. Oleh karena itu, diperlukan sistem yang mampu mengklasifikasikan jenis penyakit pada daun teh secara otomatis dan akurat. Penelitian ini mengangkat permasalahan mengenai efektivitas metode segmentasi dua tahap untuk mendeteksi area terinfeksi pada daun teh serta dampaknya terhadap kinerja model klasifikasi *Convolutional Neural Network* (CNN). Dua rumusan masalah utama yang dijawab dalam penelitian ini adalah: (1) bagaimana kinerja segmentasi dua tahap dalam mengekstraksi area penyakit, dan (2) bagaimana pengaruh segmentasi terhadap akurasi klasifikasi penyakit daun teh menggunakan CNN.

Penelitian ini menggunakan pendekatan segmentasi dua tahap, yaitu segmentasi daun menggunakan *K-Means Clustering* dan segmentasi penyakit menggunakan konversi ruang warna CIELAB serta metode *Otsu thresholding*. Dataset utama diperoleh dari *Kaggle* dan ditambah dengan anotasi manual untuk evaluasi segmentasi, serta lima dataset eksternal dari *Kaggle* dan *Mendeley Data* untuk evaluasi generalisasi model. Proses klasifikasi dilakukan menggunakan arsitektur CNN VGG19 dengan metode *transfer learning*. Tahapan penelitian meliputi pengumpulan data, segmentasi objek, praproses citra, pelatihan dan evaluasi model, serta implementasi sistem berbasis web menggunakan *Streamlit*.

Hasil evaluasi menunjukkan bahwa segmentasi daun memiliki performa spasial yang tinggi (IoU: 0,9465; *Dice*: 0,9724; *Pixel Accuracy*: 0,9871), sedangkan segmentasi area penyakit memiliki akurasi spasial yang lebih rendah (IoU: 0,4341; *Dice*: 0,5592). Pada klasifikasi, model yang dilatih dengan citra asli menunjukkan akurasi validasi sebesar 98,6% dan akurasi uji sebesar 99%, lebih tinggi dibandingkan model dengan citra hasil segmentasi (94,6% dan 96%). Evaluasi terhadap data eksternal juga menunjukkan performa model citra asli lebih baik (68,32%) dibanding segmentasi (64,32%). Penelitian ini menyimpulkan bahwa meskipun segmentasi dapat membantu menyederhanakan data, citra asli tetap memberikan informasi visual yang lebih lengkap dan mendukung performa klasifikasi yang lebih akurat dan stabil. Sistem yang dikembangkan berhasil mengintegrasikan seluruh proses segmentasi dan klasifikasi ke dalam antarmuka web interaktif. Penelitian ini memberikan kontribusi dalam mengkaji pengaruh *preprocessing* spasial terhadap akurasi klasifikasi penyakit tanaman berbasis citra.

Kata Kunci: Daun Teh, Klasifikasi Penyakit, Segmentasi Objek, CNN, VGG19

ABSTRACT

Diseases in tea leaves can significantly reduce both the productivity and quality of tea crop yields. Therefore, an automated and accurate classification system for detecting tea leaf diseases is crucial. This research addresses the effectiveness of a two-stage segmentation method in detecting infected regions on tea leaves and its impact on the performance of a Convolutional Neural Network (CNN) classification model. The study focuses on two main research questions: (1) how effective is the two-stage segmentation in extracting diseased areas, and (2) how does segmentation affect the accuracy of tea leaf disease classification using CNN.

The proposed method consists of two segmentation stages: leaf segmentation using K-Means Clustering and disease region segmentation using CIELAB color space conversion followed by Otsu thresholding. The primary dataset was sourced from Kaggle and manually annotated for segmentation evaluation, with five additional external datasets from Kaggle and Mendeley Data used to assess model generalization. Classification was performed using the VGG19 CNN architecture with transfer learning. The research workflow includes data collection, object segmentation, image preprocessing, model training and evaluation, and system deployment using a Streamlit-based web interface.

Evaluation results show that leaf segmentation achieved high spatial accuracy (IoU: 0.9465; Dice: 0.9724; Pixel Accuracy: 0.9871), while disease segmentation yielded lower accuracy (IoU: 0.4341; Dice: 0.5592). In classification, the model trained on original images achieved higher validation accuracy (98.6%) and test accuracy (99%) compared to segmented images (94.6% validation, 96% test). External dataset evaluation also favored original images (68.32%) over segmented ones (64.32%). This study concludes that although segmentation simplifies visual data, original images provide richer visual information, leading to more accurate and stable classification performance. The developed system successfully integrates segmentation and classification into a single interactive web interface. This research contributes to understanding the impact of spatial preprocessing on image-based plant disease classification accuracy.

Keywords: Tea Leaf, Disease Classification, Object Segmentation, CNN, VGG19