

ABSTRAK

Penelitian ini mengkaji implementasi arsitektur *Vision Transformer* (ViT) untuk klasifikasi penyakit pada daun singkong dengan lima kelas: *Cassava Bacterial Blight*, *Cassava Brown Streak Disease*, *Cassava Green Mottle*, *Cassava Mosaic Disease*, dan daun sehat (*Healthy*). *Dataset* sekunder berisi sebanyak 21.397 citra dari platform Kaggle dikumpulkan dan diolah melalui tahap penyeimbangan (*oversampling* dan *undersampling*) sehingga setiap kelas memiliki 3.000 sampel. Model ViT dengan token *embedding* dan *self-attention* digunakan sebagai pendekatan utama dalam deteksi visual. Eksperimen dilakukan dalam 18 konfigurasi skenario yang memvariasikan jumlah *epoch* (50 dan 70), optimizer (SGD, Adam, AdamW), ukuran *batch size* (8, 16, dan 32), dan rasio pembagian data (60:20:20, 70:15:15, dan 80:10:10). Evaluasi kinerja ditentukan menggunakan metrik *precision*, *recall*, dan *F1-score*. Hasil eksperimen menunjukkan bahwa konfigurasi terbaik dicapai pada Skenario 18 (70 *epoch*, optimizer AdamW, *batch size* 32, rasio 80:10:10), dengan *F1-score* tertinggi sebesar 86% yang seimbang antara *precision* dan *recall*. Optimizer AdamW dan Adam unggul dibandingkan SGD, dengan performa *F1-score* berkisar antara 79% hingga 86%, sedangkan SGD berada di rentang 55% hingga 73%. *Batch size* besar (32) dan rasio data pelatihan yang lebih tinggi (80%) dapat mendongkrak performa model, terutama pada *epoch* lebih besar (70). Temuan ini mengindikasikan bahwa kombinasi *hyperparameter* yang tepat sangat penting dalam optimasi model *Vision Transformer* pada klasifikasi penyakit daun singkong. Khususnya penggunaan optimizer seperti AdamW, dukungan data *training* yang lebih banyak, dan durasi pelatihan yang memadai, terbukti berkontribusi signifikan terhadap kemampuan generalisasi dan akurasi. Kontribusi penelitian ini diharapkan dapat menjadi referensi bagi pengembangan teknologi *computer vision* dalam bidang pertanian, khususnya deteksi penyakit tanaman berbasis ViT.

Kata Kunci: Klasifikasi, Tanaman Daun Singkong, *Vision Transformer*