

ABSTRAK

Penyakit dan kekurangan hara pada tanaman kopi merupakan permasalahan penting yang dapat menurunkan kualitas serta produktivitas hasil panen. Identifikasi kondisi kesehatan daun kopi secara konvensional masih banyak dilakukan secara manual dengan mengandalkan pengamatan visual, sehingga memerlukan keahlian khusus dan berpotensi menimbulkan kesalahan, terutama ketika gejala penyakit dan kekurangan hara memiliki kemiripan ciri visual. Seiring dengan perkembangan teknologi pengolahan citra dan kecerdasan buatan, pendekatan berbasis pembelajaran mesin menjadi solusi potensial untuk mengotomatisasi proses identifikasi penyakit dan kekurangan hara tanaman. Oleh karena itu, penelitian ini mengangkat permasalahan bagaimana menerapkan metode berbasis Vision Transformer untuk melakukan klasifikasi penyakit dan kekurangan hara pada daun tanaman kopi citra digital.

Penelitian ini memanfaatkan model Vision Transformer sebagai metode utama klasifikasi citra. Data yang digunakan berupa terdapat delapan kelas, yaitu *healthy*, *cercospora*, *miner*, *phoma*, dan *rust*, serta kekurangan *nitrogen-N*, *phosphorus-P*, dan *potassium-K*. Tahapan penelitian meliputi proses akuisisi data, prapemrosesan data citra, serta pembagian data. Tahap prapemrosesan data dengan melakukan perubahan ukuran citra menjadi 224 x 224 piksel dan normalisasi piksel serta menerapkan proses augmentasi berupa *horizontal flip*, *vertical flip*, *rotate*, dan *color jitter*. Dataset dibagi menjadi tiga dengan rasio 80:10:10 untuk data latih, validasi, dan uji. Selanjutnya, model Vision Transformer dilatih beberapa kali menggunakan nilai *hyperparameter* berbeda untuk menemukan performa terbaik. Proses evaluasi model dilakukan menggunakan data uji dengan memanfaatkan *confusion matrix* dan *classification report* untuk mengukur performa model berdasarkan metrik akurasi, *precision*, *recall*, dan *F1-score*.

Hasil penelitian menunjukkan bahwa model mencapai nilai akurasi keseluruhan sebesar 94,3% dengan nilai rata-rata *F1-score* sebesar 0,924 dengan ukuran *hyperparameter batch size* 16, *learning rate* 1e-4, dan *weight decay* 1e-6. Sebagian besar kelas menunjukkan nilai *precision* dan *recall* yang tinggi, meskipun masih terdapat beberapa kesalahan klasifikasi pada kelas-kelas yang memiliki kemiripan ciri visual. Pengujian tambahan menggunakan 10 citra daun kopi tambahan menunjukkan bahwa model tetap mempertahankan performa yang stabil. Dengan demikian, penelitian ini berhasil menghasilkan model klasifikasi untuk identifikasi kondisi kesehatan daun kopi menggunakan arsitektur Vision Transformer yang diharapkan dapat menjadi dasar pengembangan penelitian di masa mendatang.

Kata Kunci: Vision Transformer, daun kopi, klasifikasi citra, penyakit tanaman, kekurangan hara

ABSTRACT

Diseases and nutrient deficiencies in coffee plants are critical issues that can reduce crop quality and productivity. Conventionally, the identification of coffee leaf health conditions is still largely performed manually through visual observation, which requires specialized expertise and is prone to errors, especially when disease symptoms and nutrient deficiencies exhibit similar visual characteristics. With the advancement of image processing and artificial intelligence technologies, machine learning-based approaches have emerged as a potential solution to automate the identification of plant diseases and nutrient deficiencies. Therefore, this study addresses the problem of applying a Vision Transformer-based method to classify diseases and nutrient deficiencies in digital images of coffee leaves.

This research employs the Vision Transformer model as the primary image classification method. The dataset consists of eight classes: healthy, cercospora, miner, phoma, rust, nitrogen (N) deficiency, phosphorus (P) deficiency, and potassium (K) deficiency. The research stages include data acquisition, image preprocessing, and dataset partitioning. Image preprocessing involves resizing images to 224×224 pixels, pixel normalization, and data augmentation techniques such as horizontal flip, vertical flip, rotation, and color jitter. The dataset is divided into three splits with 80:10:10 ratio for training, validation, and testing data. The Vision Transformer model is trained multiple times using different hyperparameter values to obtain the best performance. Model evaluation is conducted using testing data through a confusion matrix and classification report, measuring performance in terms of accuracy, precision, recall, and F1-score.

According to the experimental data, the suggested model uses the Adam optimizer with a batch size of 16, a learning rate of $1e-4$, and a weight decay of $1e-6$ to attain an overall accuracy of 94,3% and a weighted average F1-score of 0.924. Most classes demonstrate high precision and recall values, although some misclassifications still occur among classes with similar visual characteristics. Additional testing using 10 unseen coffee leaf images indicates that the model maintains stable performance. Therefore, this study successfully develops a coffee leaf health classification model using the Vision Transformer architecture, which is expected to serve as a foundation for future research development.

Keywords: *Vision Transformer, coffee leaves, image classification, plant diseases, nutrient deficiency*