

ABSTRAK

Saputra, I Gede Tiar Eka (2026), Klasifikasi Penyakit Daun Cabai Rawit Menggunakan Komparasi Arsitektur Swin Transformer Dan ConvNeXt. Tesis, Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha.

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I : Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs dan Pembimbing II: Ir. Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.

Kata Kunci : Klasifikasi Citra, Swin Transformer, ConvNeXt, Penyakit Daun Cabai Rawit, Deep Learning

Penelitian ini bertujuan untuk mengevaluasi dan membandingkan performa arsitektur Swin Transformer (Swin-T) dan ConvNeXt (ConvNeXt-T) dalam mengklasifikasikan empat kelas kondisi daun cabai rawit (Bercak Daun, Klorosis, Mosaik, dan Sehat) di lingkungan terbuka (open field). Penelitian ini menggunakan dataset primer berjumlah 1.991 citra yang dikumpulkan langsung dari lahan perkebunan di Desa Sebudi, Karangasem, Bali, dan telah divalidasi secara langsung oleh pakar dari Dinas Pertanian. Dataset tersebut dibagi secara sistematis menjadi 80% (1.592 citra) untuk skenario pelatihan dan 20% (399 citra) diisolasi sebagai unseen test set murni. Pada tahap pelatihan, citra diperkaya melalui berbagai teknik augmentasi data dinamis secara on-the-fly (rotasi, flip, shear, blur, dan random erasing) guna mencegah overfitting. Kedua arsitektur dimodifikasi menggunakan skema transfer learning berbasis bobot ImageNet menggunakan varian Tiny, dan dievaluasi stabilitasnya melalui metode Stratified 5-Fold Cross Validation. Hasil penelitian pada fase validasi silang (5-Fold CV) menunjukkan bahwa kedua arsitektur memiliki tingkat stabilitas yang sangat baik dan sama-sama mencapai rata-rata akurasi yang identik sebesar 98,99%. Namun, pada pengujian generalisasi akhir menggunakan unseen test set, Swin Transformer terbukti lebih presisi dan unggul dengan tingkat akurasi mencapai 99,25%, presisi 99,26%, recall 99,25%, dan f1-score 99,25%, mengungguli ConvNeXt yang memperoleh akurasi sebesar 98,75%. Ditinjau dari segi efisiensi komputasi, Swin Transformer juga mencatatkan performa yang jauh lebih optimal dengan total waktu pelatihan 114,37 menit (39,3% lebih cepat dibandingkan ConvNeXt yang membutuhkan waktu 188,58 menit) dengan ukuran model yang sangat ramping yaitu 110,32 MB. Disimpulkan bahwa arsitektur Swin Transformer merupakan model yang paling tangguh, optimal, dan efisien untuk diimplementasikan dalam sistem cerdas klasifikasi penyakit tanaman cabai rawit dengan ketersediaan dataset primer berskala menengah di lingkungan nyata.

ABSTRACT

Saputra, I Gede Tiar Eka (2026), *Classification of Bird's Eye Chili Leaf Diseases Using a Comparison of Swin Transformer and ConvNeXt Architectures*. Thesis, Computer Science, Postgraduate Program, Ganesha University of Education.

This thesis has been approved and examined by Supervisor I: Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs and Supervisor II: Ir. Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.

Keywords: Image Classification, Swin Transformer, ConvNeXt, Bird's Eye Chili Leaf Disease, Deep Learning

This study aims to evaluate and compare the performance of the Swin Transformer (Swin-T) and ConvNeXt (ConvNeXt-T) architectures in classifying four classes of bird's eye chili leaf conditions (Leaf Spot, Chlorosis, Mosaic, and Healthy) in an open field environment. This study uses a primary dataset of 1,991 images collected directly from plantations in Sebudi Village, Karangasem, Bali, which has been directly validated by experts from the Department of Agriculture. The dataset was systematically divided into 80% (1,592 images) for the training scenario and 20% (399 images) isolated as a pure unseen test set. During the training phase, the images were enriched through various dynamic on-the-fly data augmentation techniques (rotation, flip, shear, blur, and random erasing) to prevent overfitting. Both architectures were modified using an ImageNet weight-based transfer learning scheme employing the Tiny variant, and their stability was evaluated through the Stratified 5-Fold Cross Validation method. The research results in the cross-validation phase (5-Fold CV) showed that both architectures demonstrated excellent stability and achieved an identical average accuracy of 98.99%. However, in the final generalization test using the unseen test set, the Swin Transformer proved to be more precise and superior with an accuracy rate of 99.25%, precision of 99.26%, recall of 99.25%, and f1-score of 99.25%, outperforming ConvNeXt which obtained an accuracy of 98.75%. In terms of computational efficiency, the Swin Transformer also recorded much more optimal performance with a total training time of 114.37 minutes (39.3% faster than ConvNeXt which took 188.58 minutes) alongside a very streamlined model size of 110.32 MB. It is concluded that the Swin Transformer architecture is the most robust, optimal, and efficient model to be implemented in an intelligent system for classifying bird's eye chili plant diseases with the availability of a medium-scale primary dataset in a real environment.