

Klasifikasi Penyakit Daun Tanaman Tomat Menggunakan EfficientFormer

Yaya Wihardi^{#1}, Erlangga^{#2}, Amirah Dzatul Himmah^{#3}, Herbert Siregar^{#4}

Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam
Universitas Pendidikan Indonesia
Bandung, Indonesia

¹yayawihardi@upi.edu

²erlangga@upi.edu

³amirahdzh@upi.edu

⁴herbert@upi.edu

Abstract—Tomatoes are crucial crops for Indonesian farmers, but they often suffer from diseases caused by fungi, bacteria, and viruses, leading to potential 40% yield loss. Current methods of spotting these diseases by eye result in costly and ineffective use of pesticides. This study focuses on a new way to classify tomato plant diseases using EfficientFormer. This method aims for high accuracy and fast inference time. The model reached an impressive 92% accuracy and takes just 0.4 seconds to identify diseases. This new approach could help farmers spot tomato plant diseases more accurately and quickly, potentially reducing economic losses and excessive pesticide use in Indonesia.

Key Words— EfficientFormer, MobileNet, tomatoes leaf diseases, disease identification

Abstrak— Tomat merupakan salah satu tanaman hortikultura yang banyak dibudidayakan oleh petani di Indonesia. Namun, tanaman ini cukup rentan terhadap serangan hama penyakit, baik yang diakibatkan oleh jamur, bakteri, maupun virus. Penyakit ini dapat mengakibatkan penurunan hasil panen yang signifikan hingga 40%. Metode konvensional identifikasi penyakit melalui pengamatan visual telah menyebabkan penggunaan pestisida yang tidak efektif dan mahal bagi para petani. Meskipun telah ada upaya sebelumnya dalam pengembangan model deteksi penyakit pada tanaman tomat, namun terdapat keterbatasan dalam akurasi serta ruang lingkup identifikasi. Oleh karena itu, penelitian ini fokus pada pengembangan model klasifikasi penyakit tomat berbasis EfficientFormer dengan penekanan pada akurasi tinggi dan efisiensi komputasi. Hasil menunjukkan bahwa model yang dikembangkan berhasil mencapai akurasi sebesar 92% dengan waktu inferensi yang cepat, hanya 0.4 detik. Diharapkan bahwa model ini dapat menjadi solusi inovatif dalam membantu petani untuk mengidentifikasi penyakit tanaman tomat secara lebih cepat dan akurat, dengan potensi mengurangi kerugian ekonomi serta penggunaan pestisida yang berlebihan dalam praktik pertanian tomat di Indonesia..

Kata kunci— EfficientFormer, MobileNet, penyakit tanaman tomat, klasifikasi penyakit

I. PENDAHULUAN

Tomat merupakan salah satu tanaman hortikultura yang banyak dibudidayakan oleh petani di Indonesia. Namun, tanaman ini cukup rentan terhadap serangan hama penyakit, baik yang diakibatkan oleh jamur, bakteri, maupun virus. Serangan penyakit ini tidak hanya mengakibatkan kerugian secara ekonomi bagi petani, tetapi juga berpotensi menurunkan hasil panen hingga mencapai 40% [1]. Meskipun demikian, proses pengendalian dan penanggulangan hama penyakit yang umum dilakukan oleh para petani tomat ini masih cukup konvensional [2]. Dibutuhkan pemanfaatan teknologi terintegrasi yang dapat membantu para petani menjaga keberhasilan bahkan meningkatkan hasil panennya.

Penyakit pada tanaman tomat pada umumnya dapat diidentifikasi melalui perubahan penampakan pada daun, baik dari segi bentuk, warna, maupun teksturnya [3]. Sejauh ini para petani cenderung mengandalkan pengamatan visual secara mandiri, padahal dibutuhkan keahlian khusus untuk menentukan jenis penyakit apa yang menyerang dan tindakan apa yang harus dilakukan [4]. Akibatnya seringkali pemakaian pestisida berlebih dan tidak spesifik, sehingga selain membunuh hama penyakit juga membunuh serangga yang justru berguna untuk mengendalikan hama. Selain tidak efektif, praktik ini juga dapat menambah beban biaya produksi. Untuk mengatasi masalah tersebut diperlukan pendekatan yang lebih inovatif dan terintegrasi untuk membantu para petani dalam mengenali, mendiagnosis, dan mengatasi permasalahan hama penyakit tanaman tomat dengan lebih akurat dan efisien.

Sejumlah penelitian terdahulu telah mencoba mengatasi permasalahan ini dengan mengembangkan model-model untuk mendeteksi penyakit pada tanaman tomat. Namun, sebagian besar dari model-model tersebut memiliki keterbatasan, baik dalam hal cakupan identifikasi penyakit maupun dalam tingkat akurasi yang dihasilkan. Ishaq & Waqas [5] mencoba memanfaatkan Histogram Oriented Gradient dan Support Vector Machine untuk mendeteksi penyakit Late Blight yang

disebabkan oleh gangguan jamur. Hasilnya cukup memuaskan, hasil pengujian mendapatkan akurasi sebesar 82%. Meskipun demikian, model yang dikembangkan baru dapat mendeteksi satu jenis penyakit saja.

Thuseethana [6] mencoba melakukan deteksi 9 jenis penyakit sekaligus yang meliputi bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, target spot, mosaic virus, yellow leaf curl virus, dan two-spotted spider mite. Hasilnya cukup baik, mendapatkan nilai akurasi sebesar 91% untuk model mobilenetv2 dengan arsitektur jaringan siamese. Namun, jaringan siamese ini membutuhkan database citra pembandingan yang cukup besar untuk mengukur kemiripan dengan citra kueri dan juga membutuhkan waktu komputasi yang lama.

Beberapa tahun terakhir muncul arsitektur transformer [7], yaitu VisionTransformer [8], Swin-Transformer [9], Data-Efficient Image Transformer [10], dan lainnya, yang mampu mengungguli Convolutional Neural Network sebagai basis pengembangan MobileNetV2 [11]. Transformer memberikan hasil yang lebih baik dan membutuhkan waktu training yang lebih cepat mengingat proses training dapat dilakukan secara paralel. Meskipun demikian, waktu inferensi masih menjadi tantangan yang belum terselesaikan.

Mengingat kekurangan Vision Transformer tersebut, Li dkk [12] mencoba mengembangkan model Transformer yang mampu melakukan inferensi secepat bahkan sedikit lebih cepat dibanding MobileNetV2. Hasilnya, bukan hanya dari sisi kecepatan inferensi, tapi juga mampu memberikan akurasi yang lebih baik. Berdasarkan keunggulan metode tersebut, dalam penelitian ini dikembangkan model klasifikasi penyakit tanaman tomat dengan menggunakan EfficientFormer untuk mengatasi permasalahan keterbatasan akurasi dan waktu inferensi pada klasifikasi penyakit tanaman tomat yang telah dipaparkan sebelumnya. Sebagai pembandingan, dikembangkan pula model klasifikasi penyakit tanaman tomat berbasis MobileNetV2. Penelitian ini diharapkan mampu memberi Solusi alternatif dan inovatif dalam membantu petani untuk mengidentifikasi penyakit tanaman tomat secara lebih cepat dan akurat, dengan potensi mengurangi kerugian ekonomi serta penggunaan pestisida yang berlebihan dalam praktik pertanian tomat di Indonesia.

II. METODE

Dalam penelitian ini digunakan dataset PlantVilage [13] guna membangun model klasifikasi penyakit pada tanaman tomat. Sebelum digunakan pada tahap training, data tersebut dilakukan pre-processing terlebih dahulu dengan melakukan brightness enhancement [14], dan normalisasi. Adapun model klasifikasi yang digunakan meliputi MobileNetV2 [11] dan EfficientFormer [12].

A. Dataset

Dataset yang digunakan adalah PlantVilage [13]. Dataset ini terdiri dari 54,309 daun tanaman baik yang sakit maupun yang sehat dari 14 jenis tanaman dengan resolusi 256x256 piksel. Namun, pada penelitian ini hanya digunakan dataset daun tomat yang terdiri dari 18,160 buah citra meliputi 9 jenis

penyakit yaitu: bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, target spot, mosaic virus, yellow leaf curl virus, dan two-spotted spider mite. Adapun cuplikan citra untuk setiap jenis penyakit tersaji pada gambar 1.



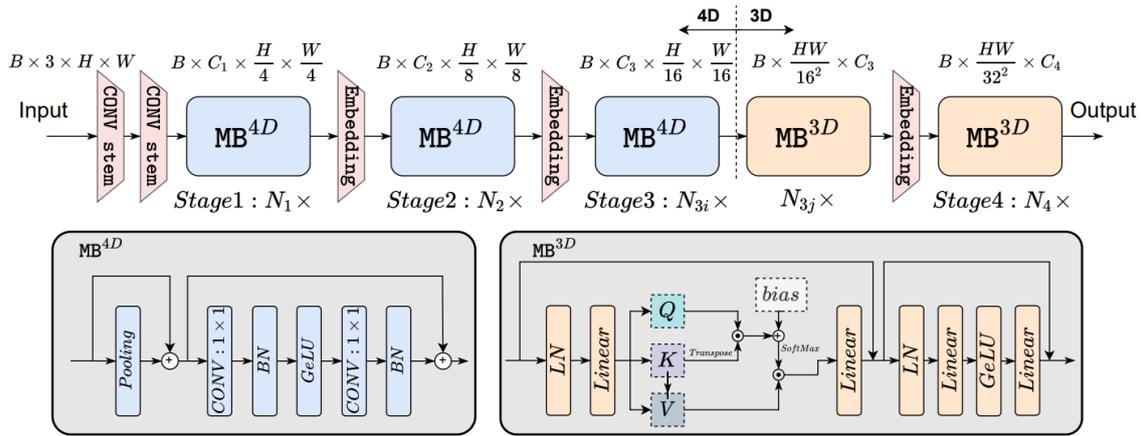
Gambar 1. Dari kiri ke kanan: bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, target spot, mosaic virus, yellow leaf curl virus, two-spotted spider mite.

Pada dataset ini, jumlah data setiap kelasnya tidak merata/imbalance. Adapun distribusinya sebagaimana terlihat pada tabel 1.

TABEL 1. DISTRIBUSI JUMLAH DATA PADA SETIAP KELAS

Kelas	Jumlah Data
bacterial spot	2.127
early blight	1.000
late blight	1.909
leaf mold	1.000
septoria leaf spot	1.771
target spot	1.404
mosaic virus	1.000
yellow leaf curl virus	5.357
two-spotted spider mite	1.676
healty	1.591

Data pada tabel 1. merupakan data asli tanpa proses augmentasi. Untuk mengatasi dataset yang imbalance, dalam penelitian ini dilakukan pula proses augmentasi data, sedemikian hingga jumlah data setiap kelas menjadi seimbang sesuai dengan jumlah data pada kelas dengan jumlah terbanyak. Augmentasi dilakukan dengan melakukan upsampling melalui proses rotasi, translasi, cropping, dan edge padding agar tetap mempertahankan resolusi aslinya. Total data setelah proses augmentasi menjadi sebesar 53.570 buah citra.



Gambar 2. Arsitektur EfficientFormer Pada Klasifikasi Penyakit Daun Tanaman Tomat [12]

TABEL 2. HASIL EKSPERIMEN

Model	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Waktu Inferensi (ms)
MobileNetV2	PlantVillage With Augmentation	84.33	85.26	82.79	84.01	511
MobileNetV2	PlantVillage Without Augmentation	80.46	82.77	79.29	80.99	503
EfficientFormer	PlantVillage With Augmentation	92.06	90.71	91.63	91.16	421
EfficientFormer	PlantVillage Without Augmentation	90.68	88.51	91.33	89.90	429

Data kemudian dibagi menjadi dua bagian yaitu 80% untuk proses training dan 20% untuk proses testing atau evaluasi. Pada tahapan pengembangan model, data training juga dibagi menjadi dua bagian yaitu 80% untuk proses training dan 20% untuk proses validasi.

B. Preprocessing

Sebelum pengembangan model dilaksanakan, pada dataset dilakukan praproses pengolahan data terlebih dahulu supaya memberikan hasil yang optimal dan sesuai dengan kebutuhan input pada setiap model. Praproses pengolahan ini meliputi proses resizing, brightness enhancement [14], dan normalisasi [15]. Resizing dilakukan mengingat kebutuhan input model adalah 224x224x3, sedangkan resolusi citra yang tersedia adalah 256x256x3. Adapun Normalisasi dan brightness enhancement dilakukan guna mengurangi kompleksitas komputasi dan mengoptimalkan proses training pada model. Praproses dilakukan baik pada data training maupun pada data testing dan validation.

C. Arsitektur MobileNet V2

Dalam proses klasifikasi, digunakan model MobileNetV2 [11] sebagai baseline yang dikombinasikan dengan dense layer sejumlah kelas pada dataset tomat plantvillage. Adapun model ditilnya dapat dilihat sebagaimana tersaji pada tabel 3.

Input model ini berupa citra berukuran 224x224 piksel dengan 3 channel warna yaitu Red, Green, dan Blue. Sebagaimana terlihat pada tabel 3., terdapat 12 blok operasi yang digunakan, meliputi blok Convolution 2D (Conv2D), Bottleneck, Average Pooling (Avgpool), dan Dense Layer. Pada setiap blok digunakan parameter faktor ekspansi (t), jumlah

channel (c), jumlah pengulangan (n), dan stride (s) yang masing-masing nilainya tersaji pada tabel 3. Pada setiap blok bottleneck dilakukan proses ekspansi sebanyak faktor t guna meningkatkan dimensi dari channel input. Kemudian pada setiap channel dilakukan operasi konvolusi secara terpisah yang dikenal sebagai depthwise convolution. Proses selanjutnya, channel-channel yang dihasilkan lalu dilakukan proyeksi kembali menjadi c buah channel sesuai parameter yang telah ditetapkan sebagaimana tersaji pada tabel 3. Terakhir, pada blok paling ujung digunakan Dense layer dengan fungsi aktivasi softmax guna melakukan klasifikasi terhadap luaran dari blok-blok sebelumnya. Luaran dari Dense layer ini berupa nilai probabilitas untuk setiap kelas data.

TABEL 3. MODEL KLASIFIKASI CITRA PENYAKIT DAUN TANAMAN TOMAT BERBASIS MOBILENETV2

Input	Operator	t	c	n	s
224 ² x 3	Conv2d	-	32	1	2
112 ² x 32	bottleneck	1	16	1	1
112 ² x 16	bottleneck	6	24	2	2
56 ² x 24	bottleneck	6	32	3	2
28 ² x 32	bottleneck	6	64	4	2
14 ² x 32	bottleneck	6	96	3	1
14 ² x 96	bottleneck	6	160	3	2
7 ² x 160	bottleneck	6	320	1	1
7 ² x 320	Conv2d 1x1	-	1280	1	1
7 ² x 1280	Avgpool 7x7	-	-	1	-
1 x 1 x 1280	Conv2d 1x1	-	k	-	-
1280	Dense 10	-	-	1	-

D. Arsitektur EfficientFormer

Selain menggunakan mobilenetV2, model utama yang digunakan untuk menyelesaikan klasifikasi penyakit pada tanaman tomat ini digunakan pula EfficientFormer [12] yang berbasis pada arsitektur Transformer. Adapun detail arsitekturnya sebagaimana terlihat pada gambar 2. Input data untuk model ini sama dengan input yang digunakan pada MobileNetV2. Melalui pemanfaatan mekanisme self-attention yang lebih efisien, diharapkan model ini mampu memberikan kecepatan inferensi yang lebih baik tanpa mengorbankan kinerja model secara keseluruhan.

Arsitektur ini terdiri dari patch embedding dan serangkaian meta transformer block (MB). Patch embedding memproses input yang terdiri dari sejumlah batch size (B) dengan ukuran HxW piksel, dalam hal ini 224x224 piksel, menjadi patch-patch yang lebih kecil yaitu 16x16 piksel. Luaran dari blok patch embedding ini kemudian diproses oleh MB. Proses pada MB terbagi ke dalam beberapa stage sebagaimana terlihat pada gambar 2., dimana dalam setiap stage-nya memproses fitur dengan ukuran yang sama. MB sendiri terdiri dari 2 bagian, yaitu MB4D yang diimplementasikan dengan memanfaatkan pendekatan Convolutional Neural Network (CNN) [11], [16], dan MB3D menggunakan pendekatan Vision Transformer [8]. Adapun luaran akhir yang didapat, sama seperti pada mobilenetv2 berupa nilai probabilitas untuk setiap kelas datanya.

E. Metode Evaluasi

Evaluasi unjuk kerja model dilakukan dengan menggunakan metric accuracy, precision, recall, dan F1-score. Selain itu, dilakukan pula pengukuran waktu inferensi guna melihat seberapa cepat proses inferensi dilakukan pada perangkat keras eksperimen.

III. HASIL DAN PEMBAHASAN

Proses klasifikasi dilakukan dengan menggunakan dua model, yaitu MobileNetV2 dan EfficientFormer. Kemudian pada setiap modelnya, dilakukan dua kali eksperimen yaitu pertama dengan menggunakan dataset asli dan kedua dengan menggunakan dataset yang telah dilakukan augmentasi. Eksperimen dilakukan dengan menggunakan tensorflow v2 di atas GPU NVIDIA GTX 1080. Adapun hasilnya sebagaimana terlihat pada tabel 2.

Proses training dilakukan dengan menggunakan teknik gridsearch di atas library ScikitLearn. Dalam setiap eksekusi training, membutuhkan waktu antara tiga sampai dengan lima jam, dengan jumlah epoch berkisar antara 300-500. Adapun optimizer yang digunakan adalah Adam dan SGD.

Hasil eksperimen menunjukkan bahwa model yang digunakan mampu melakukan klasifikasi lebih baik pada dataset yang disertai augmentasi. Hal ini terjadi karena dataset plantvillage merupakan dataset yang imbalance, sehingga proses augmentasi cukup berpengaruh pada unjuk kerja model.

Jika dilihat dari sisi model yang digunakan, EfficientFormer mampu memberikan hasil yang lebih baik dibandingkan dengan MobileNetV2. Hal ini terlihat dari kinerja dan waktu inferensinya yang selalu lebih unggul pada setiap skema eksperimen. Meskipun demikian, perbedaan waktu inferensi

antara kedua model tidak terlalu berbeda jauh, sehingga masih memungkinkan proses optimasi pada masing-masing model untuk mendapatkan hasil yang lebih baik lagi.

IV. KESIMPULAN

Berdasarkan hasil eksperimen menunjukkan bahwa EfficientFormer lebih baik dibanding MobileNetV2 dalam melakukan klasifikasi penyakit pada daun tanaman tomat. EfficientFormer berhasil mencapai akurasi sebesar 92% dengan waktu inferensi yang cepat, hanya 0.4 detik. Adapun MobileNetV2 hanya mencapai akurasi sebesar 84.33% dengan waktu inferensi yang tidak begitu jauh berbeda yaitu 0.5 detik. Selain itu, proses augmentasi data juga berpengaruh positif pada peningkatan akurasi dari kedua model.

Pengembangan model-model ini diharapkan dapat menjadi solusi inovatif dalam membantu petani untuk mengidentifikasi penyakit tanaman tomat secara lebih cepat dan akurat, dengan potensi mengurangi kerugian ekonomi serta penggunaan pestisida yang berlebihan dalam praktik pertanian tomat di Indonesia.

V. REFERENSI

- [1] G. W. Sasmito, "Sistem pakar diagnosis hama dan penyakit tanaman hortikultura dengan teknik inferensi forward dan backward chaining," *J. Teknol. dan Sist. Komput.*, vol. 5, no. 2, pp. 69–74, 2017, doi: 10.14710/jtsiskom.5.2.2017.70-75.
- [2] E. Erlangga, H. Siregar, and Y. Wihardi, "Pengembangan Framework Mobile Learning pada Pertanian Sayuran," *J. Comput. & Bisnis*, vol. 13, no. 2, pp. 58–65, 2020.
- [3] H. Rehana, M. Ibrahim, and M. H. Ali, "Plant disease detection using Region-Based Convolutional Neural Network." 2023.
- [4] S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," *Proc. - 1st Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2015*, pp. 768–771, Jul. 2015, doi: 10.1109/ICCUBEA.2015.153.
- [5] M. Ishaq and M. Waqas, "Early detection of Late Blight Tomato Disease using Histogram Oriented Gradient based Support Vector Machine," *arXiv Prepr. arXiv2306.08326*, 2023.
- [6] S. Thuseethan, P. Vigneshwaran, J. Charles, and C. Wimalasooriya, "Siamese Network-based Lightweight Framework for Tomato Leaf Disease Recognition," *arXiv Prepr. arXiv2209.11214*, 2022.
- [7] A. Vaswani *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems*, 2017, vol. 2017-December.
- [8] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for Image Recognition at scale," 2021.
- [9] Z. Liu *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," 2021, doi: 10.1109/ICCV48922.2021.00986.
- [10] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, "Training data-efficient image transformers & distillation through attention," in *Proceedings of Machine Learning Research*, 2021, vol. 139.
- [11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Jun. 2018.
- [12] Y. Li *et al.*, "EfficientFormer: Vision Transformers at MobileNet Speed," in *Advances in Neural Information Processing Systems*, 2022, vol. 35.
- [13] D. P. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics." 2016.
- [14] M. A. Haque *et al.*, "Deep learning-based approach for identification of diseases of maize crop," *Sci. Reports 2022 121*, vol. 12, no. 1, pp. 1–14, Apr. 2022, doi: 10.1038/s41598-022-10140-z.
- [15] Y. Wihardi, W. M. Kristy, Erlangga, A. Turnip, I. N. Yulita, and Endroyono, "A normalized cross-correlation convolutional neural

- network (CNN-NCC) for exemplar-based object detection,” *AIP Conf. Proc.*, vol. 2734, no. 1, Oct. 2023, doi: 10.1063/5.0155742/2917154.
- [16] M. Tan and Q. Le, “EfficientNet: Rethinking model scaling for Convolutional Neural Networks,” in *Proceedings of the 36th International Conference on Machine Learning*, 2019, vol. 97, pp. 6105–6114, [Online]. Available: <https://proceedings.mlr.press/v97/tan19a.html>.