

Pengenalan Tanaman Obat Liar dengan Metode Convolutional Neural Network

Diah Puspito Wulandari¹, Muhammad Emirreza Pahlevi¹, Hany Boedinugroho¹

¹Dept. Teknik Komputer Institut Teknologi Sepuluh Nopember Surabaya, Indonesia

Email: diah@ee.its.ac.id, pemirreza@gmail.com, Hanny@its.ac.id

Abstrak

Tanaman adalah dasar dari semua sumber kehidupan di bumi yang memberikan kita semua makan dan juga oksigen. Pemahaman yang baik tentang tanaman obat sangat penting untuk membantu dalam mengidentifikasi spesies atau jenis tanaman obat liar yang masih belum diketahui oleh banyak masyarakat sekitar, jenis tanaman ini juga untuk membantu meningkatkan industri obat menyeimbangkan ekosistem serta produktivitas dan keberlanjutan pertanian. Hasil dari percobaan menggunakan fitur CNN ini menunjukkan konsistensi dan keunggulan dibandingkan lainnya. Klasifikasi citra tanaman obat liar masih dianggap sebagai tantangan dan masalah yang belum terpecahkan, hal ini dikarenakan tumbuhan di alam memiliki macam bentuk dan representasi warna jadi berdasarkan citra model bentuk tanaman berdasarkan kelengkungan dengan memanfaatkan ukuran integral agar bisa mengetahui fungsi kelengkungan tersebut kemudian dilakukan proses klasifikasi data citra tanaman obat liar tersebut sehingga bisa teridentifikasi lalu dikembangkan.

Kata kunci: Tanaman obat, Image Processing, CNN, Klasifikasi

Diterima Redaksi: 05-02-2024 Selesai Revisi: 15-02-2024 Diterbitkan Online: 15-03-2024
DOI: <https://doi.org/10.59378/jcenim.v2i1.43>

I. LATAR BELAKANG

Tanaman obat liar adalah dasar dari semua kehidupan di muka bumi ini yang telah memberikan kita semua oksigen serta pengobatan. Di negara kita ini sangat beragam macam tanaman yang sering kita jumpai dimanapun berada. Indonesia adalah salah satu negara yang kaya akan sumber daya alamnya khususnya nabati. Sering kita mengabaikan tumbuhan yang tumbuh di pinggir jalan, dipekarangan rumah, dipinggir aliran sungai ataupun di hutan, ada yang menganggapnya sebagai tanaman liar dan tumbuhan pengganggu/parasite. Masyarakat hanya mengenal tanaman liar dari dampak negatifnya saja yang tumbuh disekitar pekarangan rumah. Masyarakat belum mendapatkan informasi tanaman liar yang berkhasiat jadi perlu dilakukan penelitian lebih lanjut mengenai pengenalan tanaman obat liar.

Di seluruh sejarah Indonesia bahkan dunia, sumber daya tanaman tetap menjadi bagian dari masyarakat manusia. Setelah memenuhi kebutuhan primer seperti makanan dan tempat untuk berlindung, manusia telah mencari obat yang efektif diantara tanaman untuk menyembuhkan berbagai macam penyakit. Tanaman obat tradisional digambarkan sebagai pengobatan pribumi untuk menjaga kesehatan serta mencegah terjadinya penularan virus atau infeksi menular. Pemahaman yang baik tentang tanaman obat sangat penting untuk membantu dalam mengidentifikasi spesies atau jenis tanaman obat liar yang masih belum diketahui oleh masyarakat sekitar, juga untuk membantu meningkatkan industri obat, menyeimbangkan ekosistem serta produktifitas manusia di dunia medis/kesehatan.

Beberapa penelitian terkait yang sudah dipelajari yaitu pengolahan citra dengan metode CNN yang sederhana, fleksibel dan umum untuk segmentasi pada sebuah objek.

II. DASAR TEORI

A. Tanaman Obat

Banyak tanaman yang digunakan sebagai bahan obat dalam berbagai keperluan pengobatan, oleh karena itu pekerjaan dilakukan pengenalan untuk mengklasifikasi tanaman obat yang benar dari gambar tanaman obat liar, sehingga gambar dari citra tersebut bisa dianggap sebagai kumpulan data pelatihan agar dapat mengklasifikasi tanaman obat lalu dapat bisa mengidentifikasinya [1].

B. Image Processing

Image processing adalah suatu bentuk pengolahan dengan input berupa gambar (image) dan ditransformasikan menjadi gambar lain sebagai keluarannya dengan teknik tertentu. Dengan memakai metode image processing ini dalam menentukan tingkat akurasi tanaman obat liar berdasarkan warna, bentuk serta ukuran agar proses identifikasi bisa terpenuhi dengan cara melihat hubungan presentase antara daun satu dengan yang lainnya [2].

C. Data Augmentation

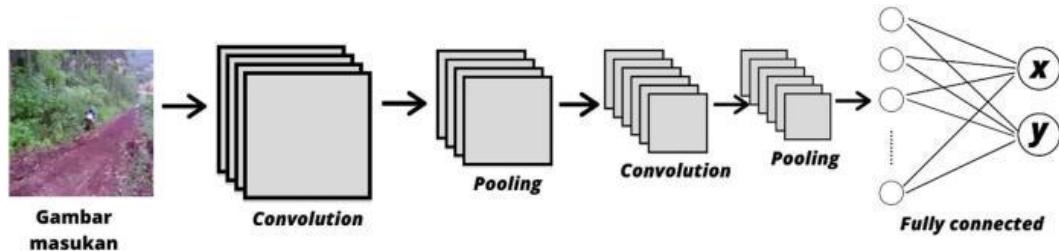
Data augmentasi adalah teknik memanipulasi sebuah data tanpa kehilangan inti atau esensi dari data yang telah didapatkan. Augmentasi data merupakan suatu proses dalam pengolahan gambar untuk mengubah atau memodifikasi citra gambar tanaman obat sedemikian rupa sehingga komputer akan mendeteksi bahwa gambar tanaman obat yang diubah adalah gambar yang berbeda, kemudian hasil dari proses augmentasi ini masih dapat mengetahui bahwa gambar yang diubah tersebut adalah gambar foto tanaman yang sama [3].

D. Convolutional Neural Network

Convolutional Neural Network atau disingkat dengan CNN adalah jenis jaringan saraf juga termasuk bagian dari Deep learning yang mendominasi berbagai tugas visi komputer dan telah menarik perhatian di berbagai bidang. CNN memiliki lapisan yang lebih tersembunyi dan arsitektur yang lebih kompleks serta memiliki kemampuan pembelajaran fitur dan representasi fitur yang lebih canggih. CNN ini bertujuan untuk mempermudah proses klasifikasi gambar (*images*) serta memberikan ketepatan deteksi objek ataupun citra.

E. ResNet-50

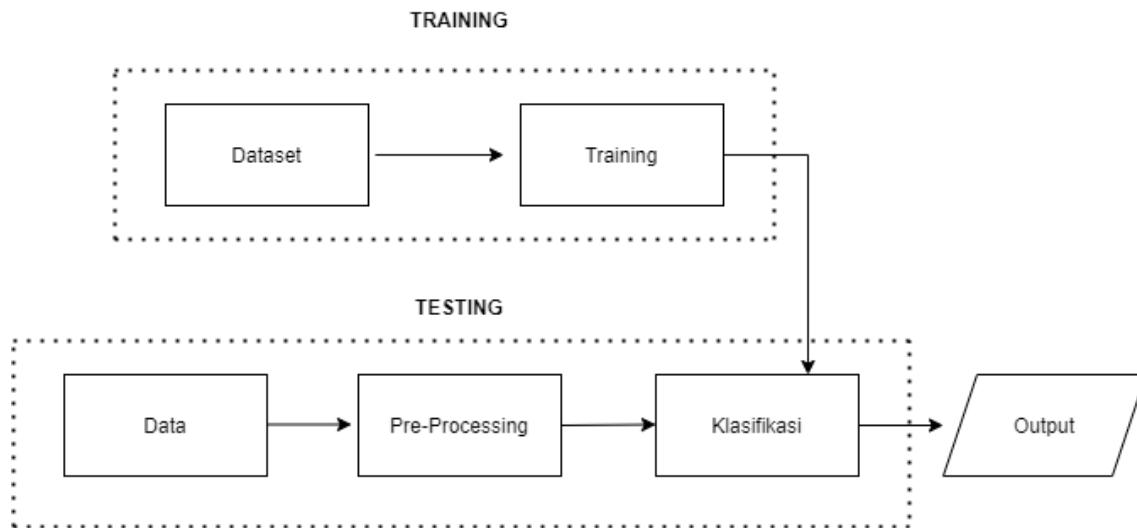
ResNet-50 merupakan salah satu arsitektur dari CNN yang memperkenalkan sebuah konsep baru yaitu *shortcut connections*. Munculnya konsep *shortcut connection* yang ada pada arsitektur ResNet-50 ini memiliki keterkaitan dengan *vanishing gradient problem* yang terjadi ketika usaha memperdalam struktur suatu *network* dilakukan. Bagaimanapun memperdalam suatu *network* dengan tujuan meningkatkan agar performansinya tidak bisa dilakukan hanya dengan cara menumpuk *layer*. Semakin dalam suatu *network* dapat memunculkan *vanishing gradient problem* yang bisa membuat *gradient* menjadi sangat kecil yang berakibat pada menurunnya perfomansi atau akurasi [4].



Gambar 1: Ilustrasi arsitektur pada CNN

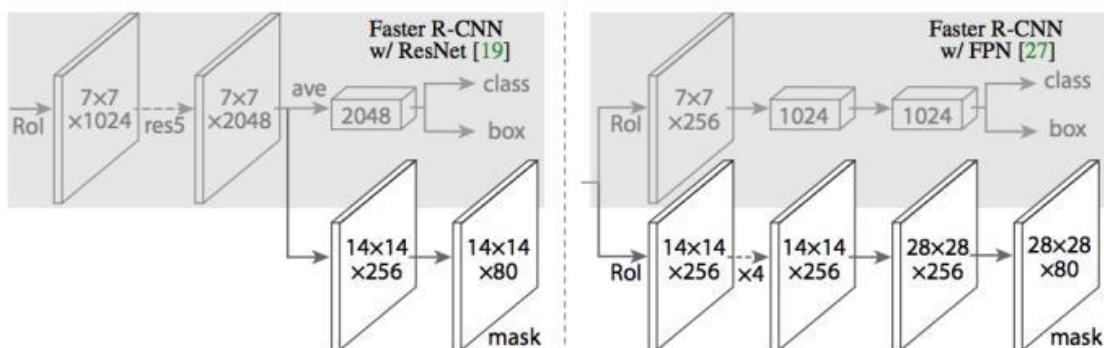
III. DESAIN SISTEM

Penelitian ini bertujuan untuk membuat sebuah model yang bisa mengklasifikasikan berbagai macam tanaman obat liar berupa *dataset* yang sudah dikumpulkan lalu diimplementasikan ke sebuah proses dari data mentah ke yang jadi. *Dataset* berupa citra tanaman obat liar dari berbagai sudut lalu masuk ke tahap *training*, dari *training* tersebut akan menghasilkan berbagai banyak citra model tanaman lalu bisa teridentifikasi.



Gambar 2: Blok diagram penelitian

Langkah selanjutnya adalah men-*training* data yang sudah diekstrasi dari citra tanaman obat tersebut menggunakan CNN agar mendapatkan sebuah citra model. Lalu selanjutnya untuk mengekstrak gambar tersebut menggunakan arsitektur ResNet-50 yang bertujuan untuk mempermudah struktur pada *shortcut connection* agar bisa meminimalisir hilangnya fitur-fitur penting pada saat proses konvolusi. Secara keseluruhan ResNet-50 terdiri dari 5 *stage* proses konvolusi yang kemudian dilanjutkan *average pooling* diakhiri dengan *fully connected layer* sebagai *layer* predksi.



Gambar 3: Arsitektur ResNet-50

A. Pengumpulan Data (Dataset)

Dataset yang dibutuhkan pada penelitian ini berupa tanaman obat liar 5 macam yaitu tanaman putri malu, sirih, binahong, kelor dan ciplukan. Total gambar yang didapatkan ada 5000 file citra dataset.



Gambar 4: Gambar lia jenis tanaman

B. Ekstraksi Fitur

Setelah mendapat gambar sesuai keinginan selanjutnya adalah ekstraksi fitur-fitur yang dibutuhkan dari citra gambar. Fitur yang diambil dari gambar tersebut seperti warna daun, bentuk daun, dan bentuk batangnya.

C. Training & Testing

Setelah mendapatkan fitur yang dibutuhkan seperti warna daun, bentuk daun dan bentuk batangnya. Langkah selanjutnya adalah melakukan *training* menggunakan CNN (*Convolutional Neural Network*) untuk mendapatkan *output* yang diinginkan.

D. Output

Output yang dihasilkan merupakan hasil dari 5 macam tanaman obat yang sudah terklasifikasi serta mengetahui apa jenis tanaman obat tersebut beserta nilai probabilitasnya.

IV. PENGUJIAN DAN ANALISA

Langkah pengujian dalam penelitian ini dilakukan seperti pada tahap berikut ini.

A. Find Dataset

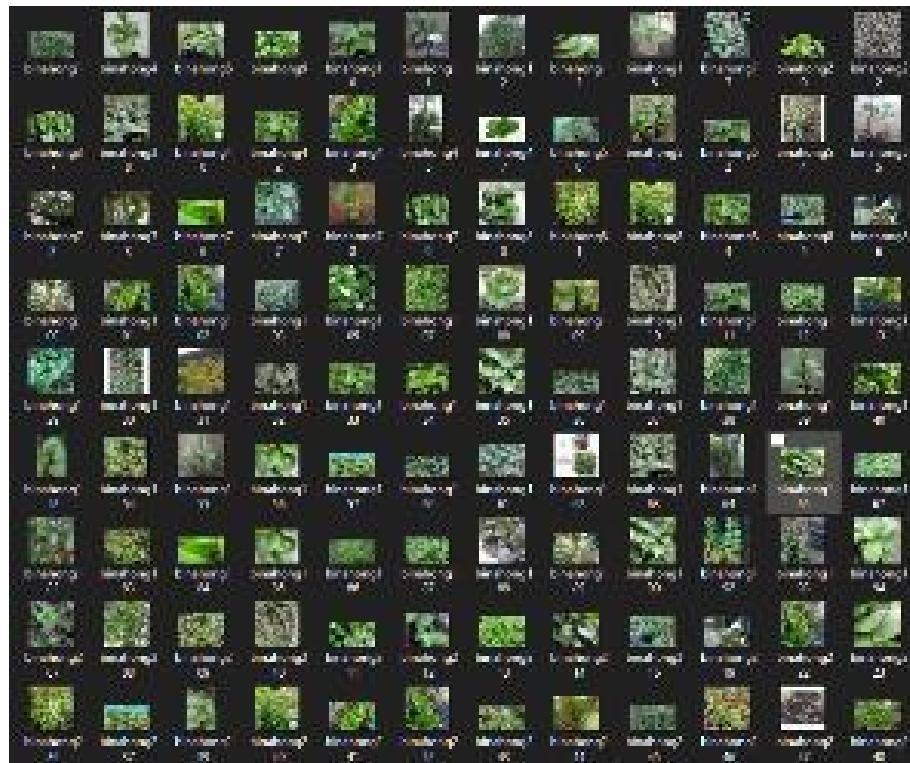
Langkah awal saat pengujian sebelum memulai eksperimen yaitu dengan mengumpulkan gambar *dataset* tanaman obat liar yang didapatkan dari mengunduh gambar pada Google Images lalu dipilih dan disortir gambarnya terlebih dahulu sebelum ditindaklanjuti ke proses pemfilteran. Berikut pada Gambar 5 merupakan contoh *dataset* yang sudah diunduh dari internet.



Gambar 5: Contoh dataset yang diperoleh dari Google Images downloader

B. Ekstraksi Fitur Gambar

Gambar *dataset* pada Google Images yang berkaitan dengan 5 macam tanaman obat liar tersebut yaitu sirih, binahong, ciplukan, kelor dan putri malu dengan total gambar 3048 file, kemudian gambar tersebut akan di *filter* supaya tidak kecampur dengan data kotor dari internet serta formatnya diseragamkan menjadi format JPG.



Gambar 6: Contoh hasil filterisasi ke format JPG

C. Pengujian Gambar Berdasarkan Image Processing

Sebelum melakukan *image processing* ada proses yang harus dilakukan terlebih dahulu yaitu melakukan *setting* parameter yang akan digunakan dalam proses *training* dan juga *testing*, jadi gambar tersebut diubah menjadi ukuran 224×224 karena model yang akan digunakan menggunakan *input size* 224×224 . Selain itu beberapa teknik augmentasi yang disediakan oleh PyTorch. Setelah dilakukan *resize* gambar dan juga ada beberapa teknik augmentasi data.

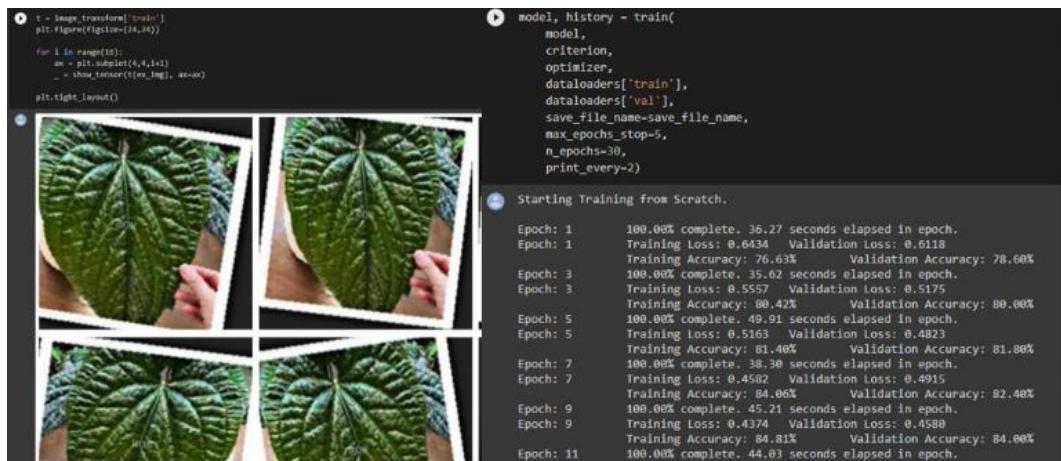
```
▶ image_path = "/content/sirih123.jpg"
display_prediction(image_path, model, topk=5)
```

content

Gambar 7: Contoh hasil proses image processing

D. Pengujian Berdasarkan Training Data

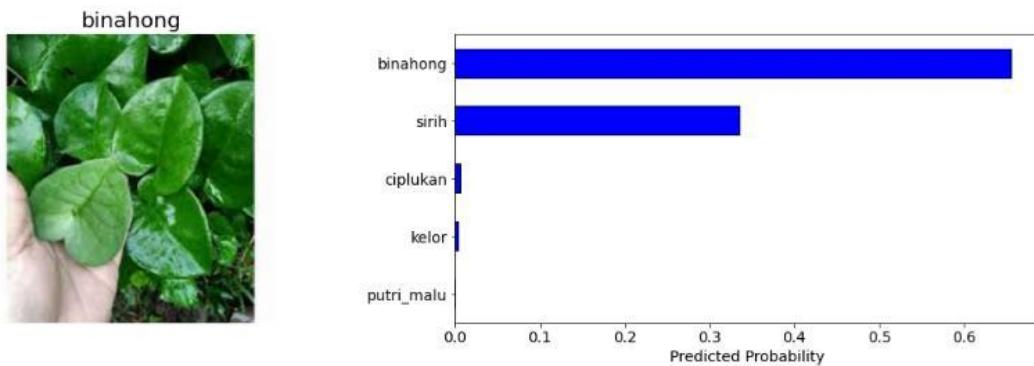
Training dilakukan untuk membentuk model dari arsitektur ResNet-50 yang kemudian digunakan sebagai sistem klasifikasi. Sebelum *training* dijalankan, terdapat beberapa konfigurasi yang harus diatur terlebih dahulu seperti *learning rate*, *batch size* serta parameter lainnya. Untuk menguji hasil model sistem yang sudah dilatih pada tahap *training* selanjutnya dilakukan eksperimen pengetesan sistem terhadap 3048 citra *dataset* yang diuji lalu dipilih secara acak agar mampu mempresentasikan seluruh karakteristik citra yang terdapat dalam *dataset*.



Gambar 8: Contoh hasil proses training

E. Pengujian Berdasarkan Testing Data

Setelah melalui tahap *training* proses berikutnya adalah melakukan uji *testing* hasil klasifikasi agar keluar nilai probabilitasnya yang berfungsi sebagai pembanding antara foto citra tanaman 1 dengan foto citra yang lainnya.

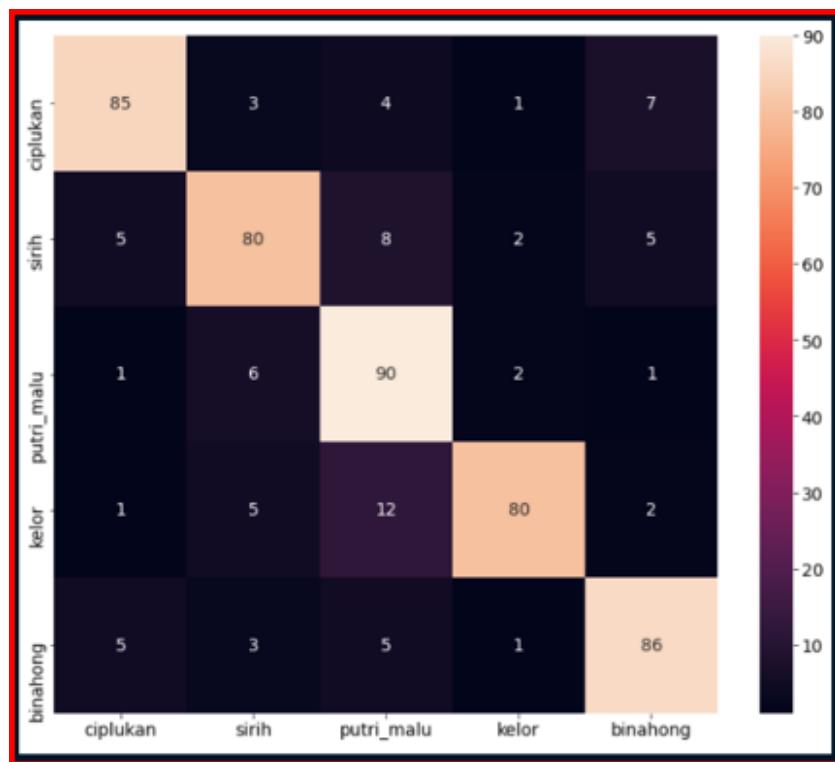


Gambar 9: Contoh hasil nilai probabilitas testing data

F. Confusion Matrix

Confusion matrix adalah pengukuran performa untuk masalah klasifikasi pada tumbuhan obat liar ini dimana hasil keluaran dapat berupa dua kelas atau lebih.

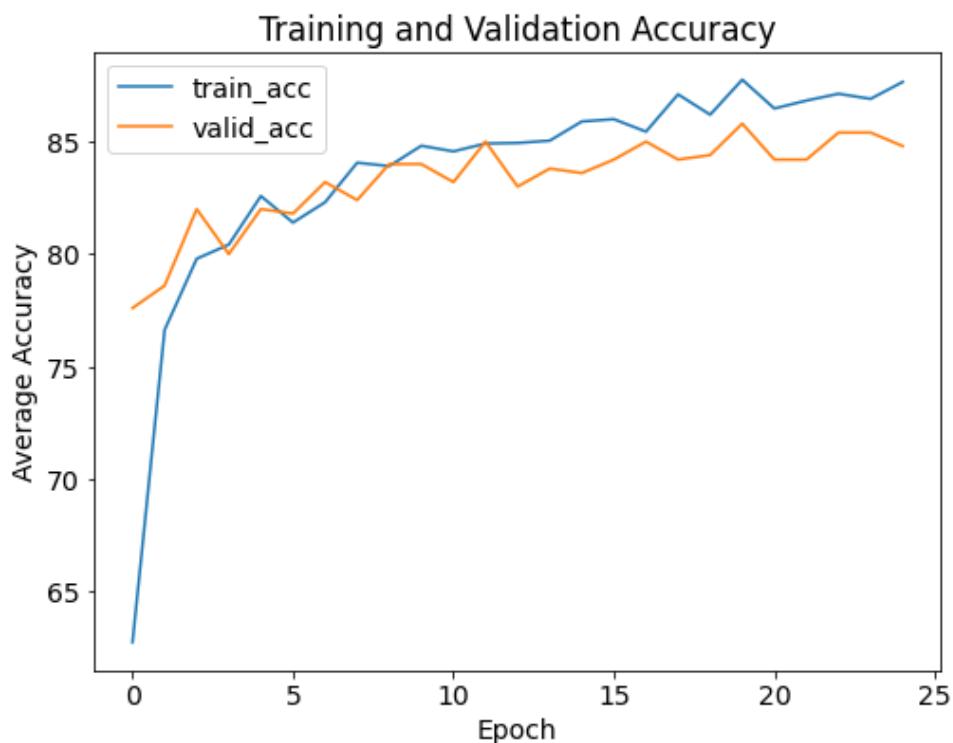
Hasil dari pengujian analisa ciplukan memiliki *True Positive* (TP) sebesar 85 dari 0 sampai 90, sedangkan sirih *True Negative* (TN) sebesar 80. Berikutnya pada tabel berwarna gelap tanaman putri malu, kelor dan binahong memiliki *False Positive* (FP) dan *False Negative* (FN) rata-rata dibawah nilai 10 dikarenakan hasil dari analisa yang telah dilakukan menunjukkan bahwa ketiga tanaman tersebut telah diprediksi bahwa hasilnya bukan seperti tanaman itu sendiri.



Gambar 10: Confusion Matrix

G. Grafik Kurva Loss

Pada kurva *loss* ini dijelaskan hasil dari pengujian data *training* dan juga data *testing* menggunakan iterasi (epoch) sebanyak 0 sampai 25. Dari hasil uji saat melakukan *training* rata-rata nilai akurasinya cukup stabil dari 75% hingga 90% jadi total *average*-nya 85%.



Gambar 11: Grafik Kurva loss

V. KESIMPULAN

Berdasarkan hasil yang telah didapatkan dari eksperimen dapat disimpulkan bahwa peran klasifikasi *images* sangat penting dalam mempengaruhi performa dari sistem klasifikasi CNN dengan ResNet-50 yang telah dirancang. Berikut pada penelitian ini dapat ditarik beberapa kesimpulan antara lain:

1. Nilai Akurasi sebesar 84% terhadap 500 data *test* yang diambil secara acak dari 5000 gambar.
2. Nilai *Precision* yang didapatkan sebesar 85% terhadap 500 data *test* yang diambil secara acak dari 5000 gambar.
3. Nilai F1-Score yang didapatkan sebesar 84% terhadap 500 data *test* yang diambil secara acak dari 5000 gambar.
4. Sistem yang digunakan saat proses klasifikasi hanya 5 macam *class* tanaman obat liar saja yang bisa terdeteksi.
5. Sistem sudah bisa membedakan antara tanaman 1 dengan tanaman lainnya.
6. Sistem sudah bisa memberikan nilai probalitasnya.

Daftar Pustaka

- [1] I. A. M. Zin *et al.*, “Herbal plant recognition using deep convolutional neural network,” *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 5, pp. 2198–2205, 2020.
- [2] I. V. P. D. Reyes, A. M. Sison, and R. P. Medina, “Fused random pooling in convolutional neural network for herbal plants image classification,” *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 3208–3214, 2019.
- [3] S. Patidar, U. Singh, and S. K. Sharma, “Weed seedling detection using mask regional convolutional neural network,” in *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. IEEE, 2020.
- [4] K. H. Mahmud, Adiwijaya, and S. A. Faraby, “Klasifikasi citra multi-kelas menggunakan convolutional neural network,” *e-Proceeding of Engineering*, 2019, telkom University, Bandung.
- [5] S. H. Lee *et al.*, “Deep-plant: Plant identification with convolutional neural network,” in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015.
- [6] J. Hu, Z. Chen, M. Yang, R. Zhang, and Y. Cui, “A multiscale fusion convolutional neural network for plant leaf recognition,” *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 853–857, 2018.
- [7] P. Sharma, Y. P. S. Berwal, and W. Ghai, “Performance analysis of deep learning cnn models for disease detection in plants using image segmentation,” *Information Processing in Agriculture*, 2019.
- [8] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [9] Y. Wu *et al.*, “Convolution neural network based transfer learning for classification of flowers,” in *2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP)*. IEEE, 2018.
- [10] I. Z. Mukti and D. Biswas, “Transfer learning based plant diseases detection using resnet50,” in *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*. IEEE, 2019.
- [11] S. U. Habiba, M. K. Islam, and S. M. M. Ahsan, “Bangladeshi plant recognition using deep learning based leaf classification,” in *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*. IEEE, 2019.
- [12] K. A. M. Han and U. Watchareeruetai, “Black gram plant nutrient deficiency classification in combined images using convolutional neural network,” in *2020 8th International Electrical Engineering Congress (iEECON)*. IEEE, 2020.

- [13] J. A. Villaruz *et al.*, “Philippine indigenous plant seedlings classification using deep learning,” in *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. IEEE, 2018.
- [14] T. Agarwal and H. Mittal, “Performance comparison of deep neural networks on image datasets,” in *2019 Twelfth International Conference on Contemporary Computing (IC3)*. IEEE, 2019.
- [15] I. W. S. E. P., A. Y. Wijaya, and R. Soelaiman, “Klasifikasi citra menggunakan convolutional neural network (cnn) pada caltech 10,” *Teknik Informatika, Fakultas Teknologi Informasi, ITS*, 2015.