



PENERAPAN PRE-TRAINED NETWORKS UNTUK DETEKSI JENIS TANAMAN HERBAL BERDASARKAN CITRA DAUN

Refido Arjunal Akmal, Sri Hartati, Aliyah Kurniasih*

Program Studi Ilmu Komputer, Universitas Ary Ginanjar, Indonesia

Abstrak: Indonesia memiliki jenis tanaman herbal yang cukup banyak. Tanaman herbal memiliki kebermanfaatan untuk kesehatan yang digunakan masyarakat secara luas sebagai pengobatan tradisional. Akan tetapi keberadaan jenis tanaman herbal tidak cukup banyak dikenali oleh masyarakat. Melihat pentingnya tanaman ini sebagai pengobatan dan masih belum banyaknya jenis tanaman herbal yang belum diketahui masyarakat, maka diperlukan suatu teknologi yang dapat mengenali jenis tanaman herbal dari citra daun. Teknologi *deep learning pre-trained* model seperti ResNet50, VGG16 dan DenseNet201 yang dikombinasikan dengan arsitektur CNN digunakan untuk membuat model *machine learning*, dan dibangun juga *base* model CNN. *Accuracy performance metrix* digunakan untuk melihat kinerja model, dan dihasilkan model terbaik pada *pre-trained networks* DenseNet201 model dengan *accuracy* sebesar 100% yang akan mampu mengenali data baru secara akurat.

Kata kunci: *pre-trained networks*, daun, tanaman herbal

I. PENDAHULUAN

Indonesia memiliki lebih dari 38.000 jenis tanaman, diantaranya lebih dari 20.000 merupakan tanaman kategori herbal dan yang baru terdata jenisnya sekitar 1.000 jenis serta yang sudah dimanfaatkan untuk pengobatan herbal baru sekitar 300 jenis (Yulianto, 2017). Tanaman herbal mengandung senyawa-senyawa baik dan bermanfaat untuk kesehatan, sehingga secara tradisional di gunakan sebagai pilihan alternatif untuk pengobatan (Haryono dkk, 2020). Banyaknya jenis tanaman herbal membuat masyarakat kesulitan mengetahui dan membedakan jenis-jenis tanaman herbal sehingga menyebabkan

keberadaan tanaman herbal masih belum banyak di ketahui masyarakat.

Hal yang paling mudah untuk dapat mengetahui jenis-jenis tanaman herbal adalah dengan mengenali bentuk daun. Daun memiliki karakteristik bentuk serta kandungan senyawa yang berbeda-beda setiap jenisnya (Pujiati & Rochmawati, 2022). Melihat kebermanfaatan tanaman herbal, dan perkembangan teknologi saat ini, maka di perlukan teknologi AI (*Artificial Intelligence*) yang dapat membantu masyarakat dalam mengenali jenis tanaman herbal secara otomatis berdasarkan citra daun.

Teknologi AI memanfaatkan pemodelan dengan algoritma *deep learning*. *Deep learning* erat kaitannya dengan *neural network*, yaitu jaringan saraf tiruan yang mirip dengan saraf pada otak manusia. *Deep learning* dirancang untuk mensimulasikan cara

*) aliyah.kurniasih@uag.ac.id

kerja otak manusia dengan berupaya menciptakan jaringan syaraf tiruan itu sendiri (Zamachsari & Puspitasari, 2021).

Algoritma CNN (*Convolutional Neural Network*) merupakan salah satu dari algoritma *deep learning* yang di rancang secara otomatis dan adaptif untuk mempelajari hirarki spasial melalui *backpropagation* dengan menggunakan *layer* penyusun seperti *convolutional layer*, *pooling layer*, *flatten layer*, dan *fully connected layer* (Yamashita dkk, 2018). *Convolutional Neural Network* (CNN) adalah model di mana setiap *layer* pada CNN disebut *feature map*. *Feature map* dari setiap *deep layer* adalah gambar *multi-channel* yang diinduksi di mana piksel dapat dianggap sebagai *feature* spesifik (TiaraSari & Hermawan, 2023). Kelebihan dari CNN yaitu menghasilkan akurasi model yang bagus untuk data citra, namun akurasi tersebut akan sangat rentan terhadap *overfitting* (Haryono dkk, 2020). Kondisi *overfitting* merupakan dimana model cenderung menghasilkan akurasi yang cukup tinggi yaitu menghafal pola data latih dengan sangat baik, akan tetapi model tidak cukup baik dalam mengenali data baru (generalisasi yang rendah) (Haryono et al, 2020)(Setya Nugraha & Hermawan, 2023). CNN juga memerlukan waktu yang cukup lama dalam memproses data yang besar pada data citra.

Transfer learning merupakan metode menggunakan jaringan saraf yang sudah di latih sebelumnya lalu mengurangi jumlah parameter dengan cara mengambil beberapa bagian dari model yang di latih untuk di gunakan dalam mengenali model baru (Firmansyach dkk, 2023). Dalam *transfer learning* terdapat istilah *pre-trained*, dimana *pre-trained* merupakan *transfer learning* yang nantinya akan melatih model dengan menggunakan data besar.

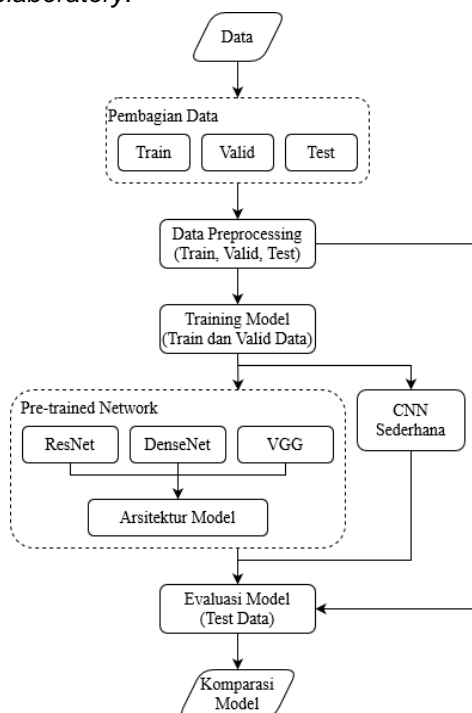
CNN memiliki *pre-trained* model yang cukup kuat terhadap data yang besar dan dapat mencegah terjadinya *overfitting*. *Pre-trained* model memiliki tingkat ke dalam lapisan yang berbeda-beda jika di dibandingkan dengan penggunaan algoritma CNN yang sederhana (Pujiati & Rochmawati, 2022). *Pre-trained* model diantaranya ResNet (*Residual Network*), melakukan pemetaan identitas data citra yang di lakukan dengan prinsip *skip connection* atau melewati beberapa *layer* untuk menghindari kehilangan gradien (Wahid Islahfari dkk, 2022). Selanjutnya ada DenseNet (*Dense Convolutional Network*) bekerja dengan menghubungkan setiap *layer* ke *layer* yang lain dengan cara *feed-forward*, yang ringan pada masalah gradien, penyebaran fitur, mendorong penggunaan kembali fitur, dan fungsionalitasnya mengurangi jumlah parameter (Widyaya & Budi, 2021). Kemudian terdapat VGG (*Visual Geometry Group*) menggunakan *convolutional layer* dengan *filter* 3x3 dan *stride* 1 yang berada di *padding same* dan *maxpool layer* dengan *filter* 2x2 dari *stride* 2 yang secara konsisten di seluruh arsitektur (Ibrahim et al, 2021).

Penelitian ini bertujuan untuk membangun model *machine learning* yang akan dapat mengenali jenis tanaman herbal dari citra daun tanaman herbal yang memiliki nilai *accuracy* cukup baik. Model *machine learning* dibuat dengan memanfaatkan algoritma *pre-trained networks* seperti *Residual Network* (ResNet) model, *Visual Geomentry Group* (VGG) Model, dan *Densely Connected Convolutional Networks* (DenseNet) model, kemudian menambahkan arsitektur *Convolutional Neural Network* (CNN) pada *pre-trained* model sebelum *output* model. Untuk melihat apakah *pre-trained* model dapat menghasilkan nilai *accuracy* yang lebih baik, CNN *base* model juga dibangun untuk

mempbandingkan antara kinerja model tersebut. Model terbaik ditentukan berdasarkan hasil evaluasi model yaitu nilai *accuracy* model yang paling tinggi dimana nilai tersebut menggambarkan kinerja model secara keseluruhan yang akan dapat memprediksi jenis tanaman herbal.

II. METODOLOGI

Gambar 1 merupakan tahapan-tahapan pada penelitian ini yaitu dari data yang digunakan dilakukan pembagian data, kemudian data *preprocessing*, selanjutnya *training* model menggunakan *pre-trained networks* dan algoritma CNN sederhana, evaluasi model, dan yang terakhir dihasilkan model terbaik. Semua proses tersebut dilakukan dengan bahasa pemrograman *Python* pada *Google Colaboratory*.

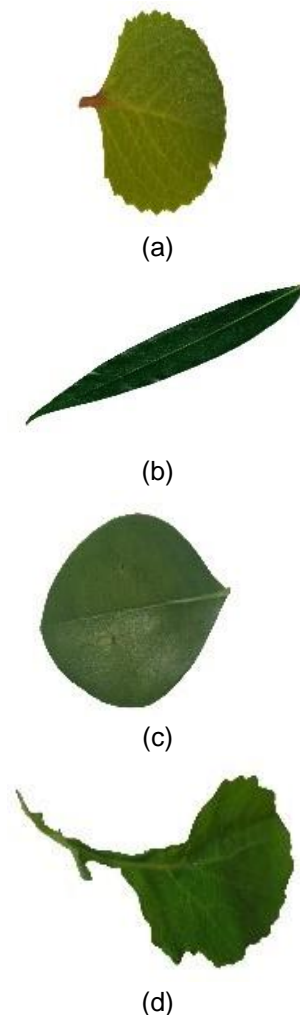


Gambar 1. Alur Penelitian

2.1 Data

Data pada penelitian ini bersumber dari data publik yaitu *website Mendelay Data* (Dharma Adhinata dkk, 2021). Data yang digunakan

hanya terdiri dari lima jenis citra daun tanaman herbal sebanyak 240 citra. Jenis citra daun tanaman herbal terdiri dari *Carissa Carandas* (Buah Samarinda) 74 citra, *Syzygium Jambos* (Jambu Mawar) 56 citra, *Plectranthus Amboinicus* (Torbangan) sebanyak 48 citra, *Brassica Juncea* (Sesawi) 34 citra, dan *Nyctanthes Arbor-tristis* (Srigading) 28 citra. Setiap citra memiliki resolusi yang sama yaitu 1600x1200 piksel dan termasuk ke dalam citra berwarna dengan berbagai kombinasi nilai intensitas dari tiga saluran warna RGB (*Red Green Blue*). Gambar 2 adalah bentuk daun dari citra tanaman herbal yang di gunakan.





(e)

Gambar 2. (a) *Plectranthus Amboinicus*, (b) *Syzygium Jambos*, (c) *Carissa Carandas*, (d) *Brassica Juncea*, (e) *Nyctanthes Arbor-tristis*.

2.2. Pembagian Data

Seluruh data citra daun tanaman herbal dilakukan pembagian data dengan proposisi 70% data *train* yang di gunakan untuk *training* model, 15% data *valid* untuk validasi model pada saat proses *training* model, dan 15% data *test* untuk proses evaluasi model dalam mengukur kinerja model secara keseluruhan.

2.3. Data Preprocessing

Data preprocessing yang dilakukan pada data *train*, *valid* dan *test* dari citra daun tanaman herbal yaitu proses normalisasi data menggunakan *Image Data Generator Tensorflow*. Parameter panjang dan lebar citra daun dirubah menjadi *image_size* 200x200 piksel, perubahan ini dilakukan karena pada penelitian ini ukuran citra asli daun yang beresolusi besar (1600x1200 piksel) akan membutuhkan sumber daya RAM (*Random Access Memory*) dengan ukuran lebih dari 12GB (melebihi batas maksimum RAM di dapatkan pada *Google Colab*). Kemudian menggunakan nilai *batch_size* 8 yang akan menentukan jumlah *batch* data di setiap *epoch* pada saat training model yaitu dihasilkan batch 22, dan mengacak data dengan pengaturan *seed* 9.

Data *preprocessing* selanjutnya yaitu *augmentasi* data menggunakan *sequential layer tensorflow* dengan parameter *RandomFlip horizontal*, *input_shape* yaitu *image_size* dengan 3 karena data bertipe

RGB, *RandomRotation* 0.1, *RandomZoom* 0.1, dan *Rescaling* 1./255.

2.4. Training Model

Training model merupakan proses pembuatan model dengan data *train* dan data *valid* menggunakan *pre-trained networks* yaitu model ResNet50, DenseNet201 dan VGG16 yang dikombinasikan dengan arsitektur tambahan, serta *training* model menggunakan algoritma CNN sederhana.

Arsitektur model CNN sederhana menggunakan sequential model *tensorflow* yang terdiri dari tiga *Convolutional2D layer* dengan jumlah *filter (kernel)* 32, dimensi 3 yaitu ukuran filter pada operasi *konvolusi* 3x3 piksel, *padding same*, *activation ReLu*, kemudian setiap *Convolutional2D layer* di ikuti dengan *MaxPooling2D layer*, akan tetapi *Conv2D layer* selanjutnya menggunakan *kernel* 64. Setelah itu menggunakan *Dropout layer* sebesar 0.2. Selanjutnya menggunakan *Flatten layer*, dimana *Flatten layer* digunakan untuk mengubah data *array* dengan matriks dua dimensi menjadi data vektor. Setelah *Flatten layer* menggunakan dua *Dense layer* dengan *neuron* unit 64 dan *activation ReLu*. *Output layer* menggunakan *Dense Layer* dengan *neuron* unit 5 dan *activation softmax*.

Pre-trained model VGG16, ResNet50, dan DenseNet201 menggunakan *weights imagenet*, *include_top False*, *input_shape (image_size, 3)* dan *trainable False*. Kemudian di tambahkan *Dropout layer* sebesar 0.2, selanjutnya *Flatten layer*, dan dua *Dense layer* dengan *neuron* unit 64 dan *activation ReLu*. *Output layer* menggunakan *Dense layer* dengan *neuron* unit 5 dan *activation Softmax*. Semua model dilakukan proses compiler menggunakan *loss function categorical crossentropy* karena jenis data citra daun lebih dari dua class, optimizer Adam dan metrics *accuracy*. Training model menggunakan nilai *epoch* 20.

2.5. Evaluasi Model

Evaluasi Model yaitu proses menguji model pada data test dengan melihat hasil performa model secara keseluruhan menggunakan *performance matrix* seperti *accuracy*, *precision*, *recall* dan *f1-score* (Sulistyowati dkk, 2023).

III. HASIL DAN PEMBAHASAN

3.1. Hasil Pembagian Data

Hasil pembagian data *train* berjumlah 174 citra, data *valid* 41 citra, dan data *test* 37 citra, detail jumlah citra setiap kelas di sajikan pada Tabel 1.

Tabel 1. Hasil Pembagian Data

Data	Kelas	Jumlah
Train	<i>Plectranthus Amboinicus</i>	33
	<i>Syzygium Jambos</i>	39
	<i>Carissa Carandas</i>	51
	<i>Brassica Juncea</i>	23
	<i>Nyctanthes Arbor-tristis</i>	28
Valid	<i>Plectranthus Amboinicus</i>	8
	<i>Syzygium Jambos</i>	9
	<i>Carissa Carandas</i>	12
	<i>Brassica Juncea</i>	6
	<i>Nyctanthes Arbor-tristis</i>	6
Test	<i>Plectranthus Amboinicus</i>	7
	<i>Syzygium Jambos</i>	8
	<i>Carissa Carandas</i>	11
	<i>Brassica Juncea</i>	5
	<i>Nyctanthes Arbor-tristis</i>	6

3.2. Hasil Training Model

Tabel 2 merupakan arsitektur model yang dihasilkan oleh model algoritma CNN sederhana atau CNN *base* model.

Tabel 2. Hasil Arsitektur Model CNN Sederhana

Layer (type)	Output Shape	Param#
sequential (Sequential)	(None, 200, 200, 3)	0
conv2d (Conv2D)	(None, 200, 200, 32)	896
max_pooling2d (MaxPooling2D)	(None, 100, 100, 32)	0
conv2d_1 (Conv2D)	(None, 100, 100, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 50, 50, 64)	0

Layer (type)	Output Shape	Param#
conv2d_2 (Conv2D)	(None, 50, 50, 64)	36928
max_pooling2d_2 (MaxPooling2D)	(None, 25, 25, 64)	0
dropout (Dropout)	(None, 25, 25, 64)	0
flatten (Flatten)	(None, 40000)	0
dense (Dense)	(None, 64)	2560064
dense_1 (Dense)	(None, 64)	4160
dense_2 (Dense)	(None, 5)	325
Total params: 2620869		
Trainable params: 2620869		
Non-trainable params: 0		

Tabel 3 merupakan arsitektur *pre-trained networks* model VGG16 yang dihasilkan.

Tabel 3. Hasil Arsitektur Pre-Trained Networks VGG-16 Model

Layer (type)	Output Shape	Param#
sequential (Sequential)	(None, 200, 200, 3)	0
vgg16 (Functional)	(None, 6, 6, 512)	14714688
dropout (Dropout)	(None, 6, 6, 512)	0
flatten (Flatten)	(None, 18432)	0
dense (Dense)	(None, 64)	1179712
dense_1 (Dense)	(None, 64)	4160
dense_2 (Dense)	(None, 5)	325
Total params: 15898885		
Trainable params: 1184197		
Non-trainable params: 14714688		

Tabel 4 merupakan arsitektur *pre-trained networks* model ResNet50 yang dihasilkan.

Tabel 4. Hasil Arsitektur Pre-Trained Networks ResNet50 Model

Layer (type)	Output Shape	Param#
sequential (Sequential)	(None, 200, 200, 3)	0
resnet50 (Functional)	(None, 7, 7, 2048)	23587712
dropout (Dropout)	(None, 7, 7, 2048)	0
flatten (Flatten)	(None, 100352)	0
dense (Dense)	(None, 64)	6422592
dense_1 (Dense)	(None, 64)	4160
dense_2 (Dense)	(None, 5)	325
Total params: 30014789		
Trainable params: 6427077		
Non-trainable params: 2357712		

Tabel 5 merupakan arsitektur *pre-trained networks* model DenseNet201 yang dihasilkan.

Tabel 5. Hasil Arsitektur Pre-Trained Networks DenseNet201 Model

Layer (type)	Output Shape	Param#
sequential (Sequential)	(None, 200, 200, 3)	0
densenet201 (Functional)	(None, 6, 6, 1920)	18321984
dropout (Dropout)	(None, 6, 6, 1920)	0
flatten (Flatten)	(None, 69120)	0
dense (Dense)	(None, 64)	4423744
dense_1 (Dense)	(None, 64)	4160
dense_2 (Dense)	(None, 5)	325
Total params: 22750213		
Trainable params: 4428229		
Non-trainable params: 18321984		

Tabel 6 merupakan hasil training *pre-trained networks* model dan CNN sederhana yang dihasilkan pada *epoch* ke-20 yang terdiri dari nilai *accuracy*, *validation accuracy*, *loss*, dan *validation loss*. Berdasarkan nilai *accuracy* dan *validation accuracy* yang dihasilkan, model terbaik yaitu pada *pre-trained networks* model VGG16 dan DenseNet201 sebesar 100%. Akan tetapi jika melihat nilai *loss* dan *validation loss* terendah yaitu pada *pre-trained networks* model DenseNet201 dengan *loss* sebesar $3,26 \times 10^{-6}$ dan *validation loss* $1,16 \times 10^{-7}$. Sehingga *training* model terbaik yaitu pada *pre-trained networks* model DenseNet201.

Tabel 6. Hasil Training Model

Pre-Trained Networks	Acc	Val Acc	Loss	Val Loss
CNN	98,3%	97,6%	0,0689	0,2425
VGG16	100%	100%	$4,18 \times 10^{-4}$	0,0093
ResNet50	85,6%	80,5%	0,3983	0,9400
DenseNet201	100%	100%	$3,26 \times 10^{-6}$	$1,16 \times 10^{-7}$

3.2. Hasil Evaluasi Model

Tabel 7 merupakan hasil evaluasi model dengan melihat nilai keakuratan model secara keseluruhan dalam memprediksi data baru. Model terbaik dihasilkan pada *pre-trained*

networks model DenseNet201 dengan nilai akurasi 100% dan nilai *loss* 0,000. Kemudian di ikuti dengan *pre-trained networks* model VGG16 akurasi sebesar 100% dan *loss* 0,001, setelah itu model CNN sederhana dengan akurasi sebesar 97,30% dan *loss* 0,083, dan model dengan akurasi terendah yaitu *pre-trained networks* ResNet50 model dengan akurasi 81,08% dan *loss* 0,617.

Tabel 7. Hasil Evaluasi Model

Pre-Trained Networks	Accuracy	Loss
CNN	97,30%	0,083
VGG16	100%	0,001
ResNet50	81,08%	0,617
DenseNet201	100%	0,000

IV. KESIMPULAN

Evaluasi model menghasilkan nilai *accuracy* tertinggi sebesar 100% pada kedua *pre-trained networks* model yaitu DenseNet201 model dan VGG16 model. Akan tetapi jika melihat nilai *loss* yang dihasilkan pada VGG16 model dengan *loss* sebesar 0,001 dan DenseNet16 model sebesar 0,000, maka dapat disimpulkan model *pre-trained networks* DenseNet201 model yang akan mampu mengenali data baru dalam memprediksi jenis tanaman herbal berdasarkan citra daun secara akurat dibandingkan dengan model yang lainnya pada penelitian ini. Hasil nilai akurasi yang cukup tinggi ini dapat dipengaruhi oleh data asli citra daun yang digunakan dengan *background* putih atau dengan nilai intensitas RGB mengandung kombinasi *red* 255 dan *blue* 255 secara merata pada seluruh citra daun.

DAFTAR PUSTAKA

Adhinata. F. D., Rakhmadani. D. P., Wibowo. M., & Jayadi. A. (2021). A Deep Learning Using DenseNet201 to Detect Masked or Non-Masked Face. *JUITA: Jurnal Informatika*, 9(1), 115-121.

- Firmansyach. W. A., Hayati. U., & Wijaya. Y. A. (2023). Analisa Terjadinya Overfitting Dan Underfitting Pada Algoritma Naïve Bayes Dan Decision Tree Dengan Teknik Cross Validation. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 262-269.
- Haryono, Anam. K., & Saleh. A. (2020). Autentikasi Daun Herbal Menggunakan Convolutional Neural Network dan Raspberry Pi. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 9(3), 278–286.
<https://doi.org/10.22146/v9i3.302>
- Ibrahim. N., Rizal. S., Saidah. S., Syahrin. H., Fardiansyah. S. A., Afghani. A. Z. A., & Hayat. M. H. (2022). Klasifikasi Citra Klon Teh Seri GMB Menggunakan Convolutional Neural Network (CNN) dengan Arsitektur Resnet, Vgg-net, dan Alexnet. *Jurnal Sains Teh dan Kina*, 1(2), 27–39.
- Kurniasih. A., & Manik. L. P. (2022). On the Role of Text Preprocessing in BERT Embedding-based DNNs for Classifying Informal Texts. *International Journal of Advanced Computer Science and Applications*, 13(6), 927–934.
<https://doi.org/10.14569/IJACSA.2022.01306109>
- Pujiati. R., & Rochmawati. N. (2022). Identifikasi Citra Daun Tanaman Herbal Menggunakan Metode Convolutional Neural Network (CNN). *Journal of Informatics and Computer Science (JINACS)*, 3(03), 351–357.
<https://doi.org/10.26740/jinacs.v3n03.p351-357>
- Roopashree. S., & Anitha. J. (2020). *Medicinal Leaf Dataset*.
<https://doi.org/10.17632/nnytj2v3n5.1>
- Nugraha. R. S., & Hermawan. A. (2023). Optimasi Akurasi Metode Convolutional Neural Network Untuk Klasifikasi Kualitas Buah Apel Hijau. *Jurnal Mnemonic*, 6(2), 149–156.
<https://doi.org/10.36040/mnemonic.v6i2.6730>
- Sulistiyowati. T., Purwanto, Zami. F. A., & Pramunendar. R. A. (2023). VGG16 Deep Learning Architecture Using Imbalance Data Methods For The Detection Of Apple Leaf Diseases. *Moneter: Jurnal Keuangan Dan Perbankan*, 11(1), 41–53.
<https://doi.org/10.32832/moneter.v11i1.57>
- Sari. A. T., & Haryatmi. E. (2021). Penerapan Convolutional Neural Network Deep Learning dalam Pendeteksian Citra Biji Jagung Kering. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(3), 265–271.
<https://doi.org/10.29207/resti.v5i2.3040>
- Wahid. M. I., Lawi. A., & Siddik. A. M. A. (2022). Perbandingan Kinerja Model Ensembled Transfer Learning Pada Klasifikasi Penyakit Daun Tomat. *Seminar Nasional Teknik Elektro Dan Informatika (SNTEI)*, 8(1), 286–291. Retrieved from <http://118.98.121.208/index.php/sntei/article/view/3630>
- Widyaya. J. E., & Budi. S. (2021). Pengaruh Preprocessing Terhadap Klasifikasi Diabetic Retinopathy dengan Pendekatan Transfer Learning Convolutional Neural Network. *Jurnal Teknik Informatika Dan Sistem Informasi*, 7(1), 110–124.
<https://doi.org/10.28932/jutisi.v7i1.3327>

- Yamashita. R., Nishio. M., Do. R. K. G., & Togashi. K. (2018). *Convolutional Neural Networks: An Overview and Its Applications in Radiology. Smart Innovation, Systems and Technologies*, 21–30. https://doi.org/10.1007/978-981-15-7078-0_3
- Yulianto. S. (2017). Penggunaan Tanaman Herbal Untuk Kesehatan. *Jurnal Kebidanan Dan Kesehatan Tradisional*, 2(1), 1–7. <https://doi.org/10.37341/jkkt.v2i1.37>
- Zamachsari. F., & Puspitasari. N. (2021). Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(3), 203–212. <https://doi.org/10.29207/resti.v5i2.2952>