

Implementasi Algoritma *Convolutional Neural Network* (CNN) Untuk Klasifikasi Daun Dengan Metode *Data Mining* SEMMA Menggunakan Keras

Yosia Amado Suwitono, Fergie Joanda Kaunang
Teknik Informatika, Fakultas Teknologi Informasi, Universitas Advent Indonesia
*email:1981011@unai.edu

DOI: <https://doi.org/10.31603/komtika.v6i2.8054>

Received: 25-10-2022, Revised: 09-11-2022, Accepted: 18-11-2022

ABSTRACT

Each plant has its own variations and characteristics. The plant has her three main parts: leaves, roots and stems. Most plants have many leaves, so it's easy to tell them apart. However, due to the limited capacity of the human brain, humans are usually unable to identify plants by their leaves. *classification* is a technique that can be used to distinguish between objects. *classification* requires using the correct method to obtain results with maximum accuracy. The Convolutional Neural Network (CNN) algorithm is a neural network commonly used for image *classification*. Supported by the SEMMA *Data Mining* method, the model building stage is also maximized. Using the right models and mechanisms will give you the best results. Four models built with different eras and lot sizes yielded different accuracy results. From this research, the best results showed that the model built with the epoch mechanism was 100, with a stack size of 30, an overall accuracy of 98%, and a *loss* of 0.0537. From his analysis of the four models built, it was concluded that epochs and stack size affect model accuracy. The SEMMA *Data Mining* method has also proven to ease the process of creating structured models for researchers.

Key Word: CNN, image classification, SEMMA model, neural network

ABSTRAK

Tumbuhan memiliki variasi dan ciri khasnya masing-masing. Ada tiga bagian utama dalam tumbuhan yaitu daun, akar dan batang. Sebagian besar tanaman memiliki daun yang sangat banyak sehingga mudah untuk didapatkan untuk membedakan tanaman satu dengan lainnya. Namun orang pada umumnya tidak dapat mengidentifikasi tanaman menggunakan daun karena terbatasnya kemampuan otak manusia. Klasifikasi adalah salah satu teknik yang dapat digunakan untuk dapat membedakan antar sebuah objek. Klasifikasi harus menggunakan metode yang tepat agar tercapai hasil akurasi yang maksimal. Dengan algoritma *Convolutional Neural Network* (CNN) adalah *neural network* yang banyak digunakan untuk mengklasifikasikan citra. Didukung dengan metode *Data Mining* SEMMA juga memaksimalkan tahapan dalam membuat sebuah model. Dengan model yang tepat dan menggunakan mekanisme yang tepat akan memperoleh hasil yang maksimal. Dari keempat model yang dibangun menggunakan *epoch* dan *batch size* yang berbeda didapatkan hasil *accuracy* yang bervariasi. Dari penelitian ini menunjukkan hasil tertinggi model yang dibangun dengan mekanisme *epoch* sebanyak 100 dan *batch size* sebanyak 30 didapatkan total *accuracy* sebesar 98% dan *loss* sebanyak 0,0537. Dari analisis ke empat model yang dibangun ditarik sebuah kesimpulan bahwa *epoch* dan *batch size* mempengaruhi besar kecilnya *accuracy* sebuah model. Dibuktikan juga bahwa metode *Data Mining* SEMMA dapat mempermudah peneliti dalam proses pembuatan sebuah model yang terstruktur.

Kata-kata kunci: CNN, Klasifikasi Citra, SEMMA Model, Neural Network

PENDAHULUAN

Jumlah jenis tanaman ada sangat banyak sekali. Antara satu tanaman dengan tanaman lainnya bervariasi, mulai dari berbeda semuanya hingga perbedaan yang sangat tipis yang hanya daunnya saja atau daunnya sama tetapi struktur tanamannya berbeda. Dalam ilmu morfologi tumbuhan terdapat tiga organ atau struktur yang membangun suatu tumbuhan yaitu akar, batang, dan daun. Tumbuhan juga memiliki organ-organ sekunder yaitu bunga, buah, duri dan lain-lain [1].

Daun merupakan salah satu bagian tumbuhan yang mudah untuk dibedakan walau dengan mata telanjang. Namun terkadang satu atau beberapa jenis tanaman tertentu memiliki bentuk daun yang sama hanya sedikit perbedaan baik dari warna daun maupun bentuk struktur tanaman itu sendiri. Gen memainkan peran penting dalam pembentukan morfologi tanaman karena digunakan untuk membedakan spesies tanaman[2]. Saat ini, satu spesies tanaman dapat memiliki bentuk yang sangat banyak. Dengan adanya keanekaragaman bentuk tanaman ini maka proses klasifikasi daun dapat menjadi proses yang sangat melelahkan terlebih jika dilakukan hanya dengan cara konvensional.

Kemajuan teknologi saat ini memberikan dampak yang sangat besar termasuk dalam bidang keanekaragaman hayati secara khusus pada keanekaragaman jenis tumbuhan atau tanaman. Salah satu kemajuan teknologi yang ada saat ini yang dapat digunakan untuk membedakan jenis daun berdasarkan bentuk daun adalah *data mining* dan *machine learning*. *Machine learning* adalah salah satu teknologi mutakhir yang berkembang pesat pada era ini dimana sebuah mesin diprogram dengan mengembangkan sebuah algoritma untuk dapat belajar dengan sendirinya[3]. *Data mining* sendiri diperlukan untuk mengumpulkan informasi ataupun pola-pola sebagai bahan pengetahuan dari *machine learning* itu sendiri. Dalam *data mining* terdapat sebuah teknik klasifikasi untuk gambar 2 dimensi yang dikenal sebagai teknik *image classification* [4].

Image classification adalah klasifikasi gambar dengan cara mengelompokkan objek dan menkonversinya menjadi angka yang ditentukan sehingga dapat mudah mengenali objek[5]. Proses ini diawali dengan *image processing* untuk memperjelas data gambar yang akan dibaca oleh mesin. Beberapa pengembang membangun *library* perangkat lunak yang dapat digunakan untuk penelitian maupun untuk mengembangkan sebuah model untuk mengklasifikasikan tanaman. Salah satu *library* yang telah dibuat dan banyak digunakan saat ini adalah tensorflow. Tensorflow sendiri merupakan sebuah *machine learning framework* yang dikembangkan oleh Google pada akhir tahun 2015. Salah satu kelebihan dari tensorflow adalah kemampuan diferensiasi otomatisnya. Sebuah eksperimen dengan jaringan baru dapat dilakukan tanpa harus mendefinisikan banyak perhitungan *key (key calculation)* [6].

Sebuah penelitian terdahulu menggunakan algoritma *K-Nearest Neighbor* (KNN) untuk melakukan klasifikasi gambar dalam membedakan varietas jeruk dengan hasil akurasi model sebesar 75% [7]. Penelitian lainnya [8], melakukan penelitian untuk mengklasifikasi buah pisang dengan data sebanyak 5 jenis pisang. Hasil dari penelitian yang menggunakan algoritma *Neural Network* ini menunjukkan bahwa nilai *accuracy* tertinggi sebesar 100% dengan lama pemrosesan 22 detik dan rata-rata nilai *accuracy* mencapai 89,95%. Penelitian terdahulu lainnya melakukan klasifikasi citra daging menggunakan algoritma *Support Vector Machine (SVM)* yang menunjukkan hasil *accuracy* sebesar 87,5% dengan *error* paling kecil 14,2% [9].

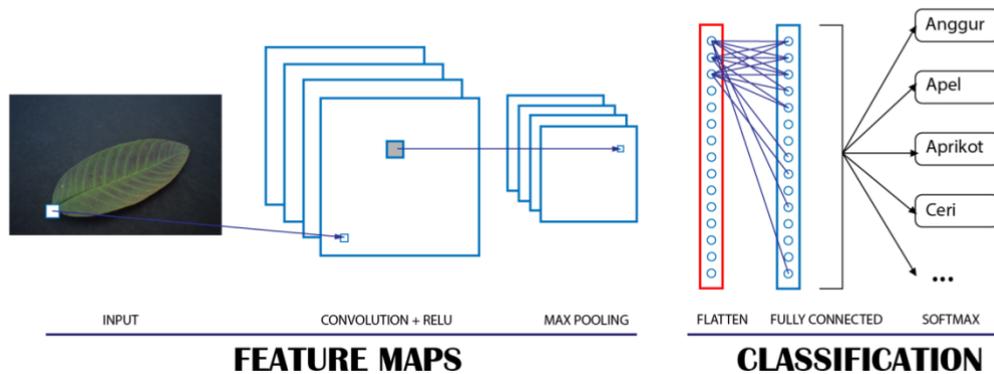
Berdasarkan penelitian-penelitian terdahulu dan dengan adanya kemajuan teknologi yang telah disebutkan sebelumnya, peneliti melihat kesempatan untuk membangun sebuah model dalam mengklasifikasi daun menggunakan teknik klasifikasi pada citra (*image classification*) menggunakan metode *Data Mining* SEMMA. Model akan dibangun menggunakan algoritma *Convolutional Neural Network (CNN)* untuk mencapai *accuracy*, *precision*, dan *recall*

terbaik. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan model-model klasifikasi daun pada suatu tanaman .

METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah CNN yang merupakan salah satu algoritma dari *deep learning* yang banyak digunakan dengan data citra yang dapat mengenali objek-objek pada suatu citra[10]. CNN terbentuk dari *neuron* yang tersusun membentuk sebuah filter dengan memiliki panjang dan tinggi yang didalamnya terdapat *weight*, bias dan aktivasi. Secara menyeluruh, CNN tidak jauh berbeda dari algoritma *deep learning neural network* lainnya.

CNN terdiri atas serangkaian layer tersebut terdiri atas *input layer*, *convolution layer*, *non-linear layer*, *pooling layer*, dan *fully connected layer* [11]. *Layer-layer* tersebut adalah proses jalannya algoritma untuk mengenali sebuah objek dalam sebuah citra. Cara kerja CNN adalah dengan meniru jaringan otak saraf manusia. Arsitektur CNN disajikan pada Gambar 1.

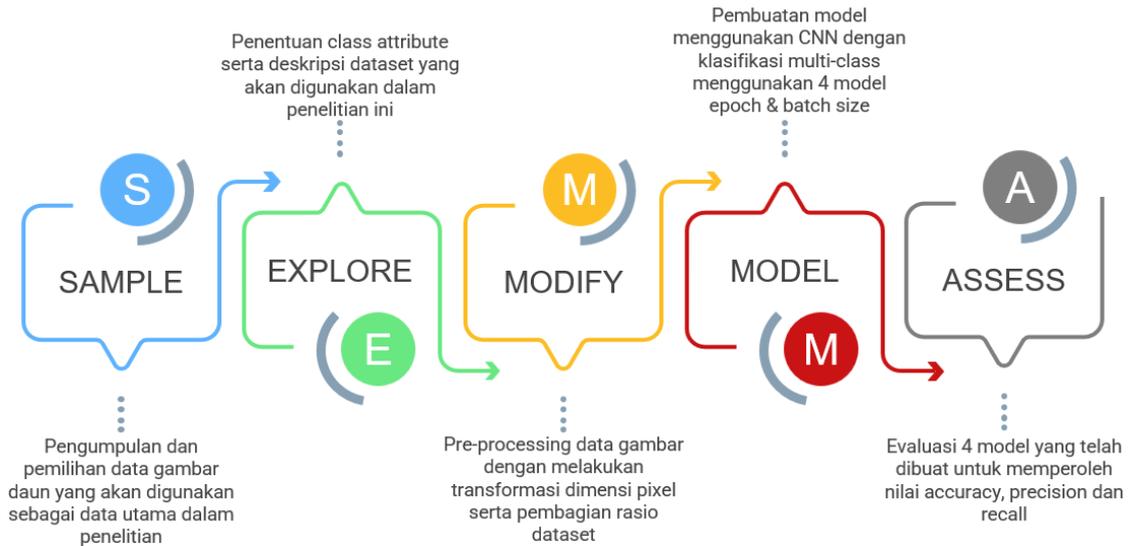


Gambar 1. Arsitektur CNN

Penelitian ini menggunakan algoritma CNN karena algoritma ini memiliki ekstraksi fitur otomatis yang tidak memerlukan bantuan manusia dan CNN dapat mengenali lebih dari satu dimensi objek sehingga dapat mengakses informasi secara maksimal yang terdapat dalam data dan selanjutnya dilakukan ekstraksi fitur [12].

Banyak proses model *data mining* yang dapat digunakan diantaranya *Knowledge Discovery Databases (KDD)*, *Cross Industry Standard Process (CRISP-DM)*, dan *Sample, Explore, Modify, Model, Access (SEMMA)*. Masing-masing proses memiliki fungsi dan kegunaannya sendiri. Penelitian ini menggunakan proses model SEMMA karena keunggulan SEMMA lebih terpusat pada aspek pengembangan model dari *data mining* sehingga pembuatan model lebih terstruktur.

SEMMA adalah sebuah metode yang mudah dipahami dan digunakan sebagai acuan dalam sebuah proyek *data mining*. *Sample* adalah sebuah proses untuk mengumpulkan data dan informasi yang signifikan. *Explore* adalah tahapan untuk mencari kumpulan data yang terkait dengan ide-ide yang akan dibangun. *Modify* adalah sebuah proses untuk memodifikasi data dan mengelompokkan variabel variabel. *Model* merupakan proses memodelkan data untuk memprediksi hasil yang diinginkan. *Access* merupakan penilaian data untuk evaluasi [13]. Tahapan metode SEMMA disajikan pada Gambar 2.

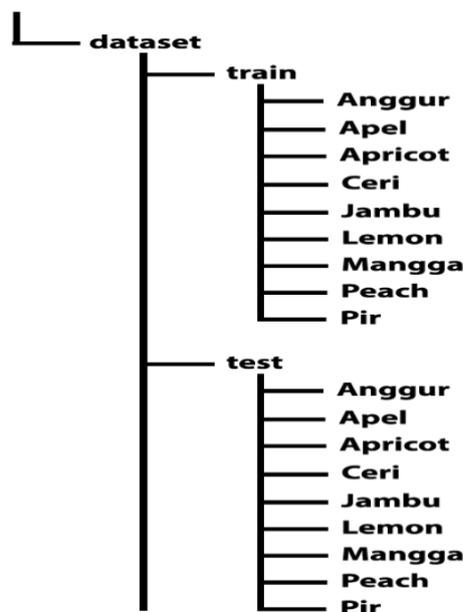


Gambar 2. Metode SEMMA

Dalam penelitian ini menggunakan tensorflow yang merupakan *library* populer yang dikembangkan oleh Google untuk mengembangkan sebuah *machine learning*. Selain itu juga menggunakan Keras yaitu salah satu *library* yang digunakan untuk membangun dan melatih sebuah model menggunakan *neural network*.

SAMPLE

Data yang digunakan dalam penelitian ini adalah dua buah *dataset* yang terdapat di bank dataset Mendeley dan di referensi oleh Tensorflow dataset *plantaek* [14] dan *plant_leaves* [15]. Dari dua dataset tersebut data dipilah secara manual dan digunakan daun dari pohon berbuah yang secara umum dikenali seperti anggur, apel, aprikot, ceri, jambu, lemon, mangga, persik, dan pir. Total gambar yang diperoleh dari dua dataset tersebut sebanyak 1608 gambar dengan direktori dataset seperti pada Gambar 3.



Gambar 3. Direktori Dataset

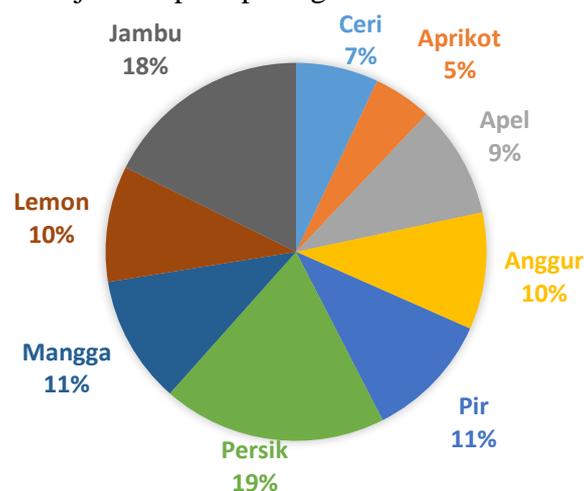
EXPLORE

Dari konten dua dataset tersebut memiliki dimensi, konsep gambar yang sama, sehingga terbentuklah sebuah ide untuk mengklasifikasikan gambar berdasarkan 9 kelas tersebut. Karena banyaknya *class* yang digunakan pada proses ini peneliti menentukan *class mode categorical* untuk *augmentation* menggunakan generator dan *loss function sparse categorical crossentropy* karena model yang akan dibuat adalah *multi-class* dan data label berbentuk *integer*.

Peneliti mengembangkan sebuah ide untuk membangun 4 model untuk dianalisis pengaruh *epoch* dan *batch size* terhadap suatu kepingan sebuah mesin. Dengan batasan *epoch* 100 sebagai yang tertinggi dengan *batch size* tanpa batasan tergantung pada peneliti. Atribut *machine learning* yang digunakan dalam pembelajaran adalah piksel data dan juga warna pada gambar karena satu jenis daun dapat memiliki warna yang berbeda beda tergantung pada masa hidupnya dan jenis tanamannya sehingga semakin banyak warna yang dipelajari akan semakin baik hasil yang akan diprediksi.

MODIFY

Di tahap *modify* akan dilakukan *pre-processing image*. Pada tahap *sample* sudah dilakukan penggabungan data dan kemudian pembersihan data secara manual. Data-data yang tidak relevan dan tidak dibutuhkan untuk pembelajaran mesin akan dibuang beserta suatu kelas yang memiliki data berlebih dan mengganggu keseimbangan akan dikurangi sehingga pembagian data disajikan seperti pada gambar 4.



Gambar 4. Distribusi Data dalam Dataset

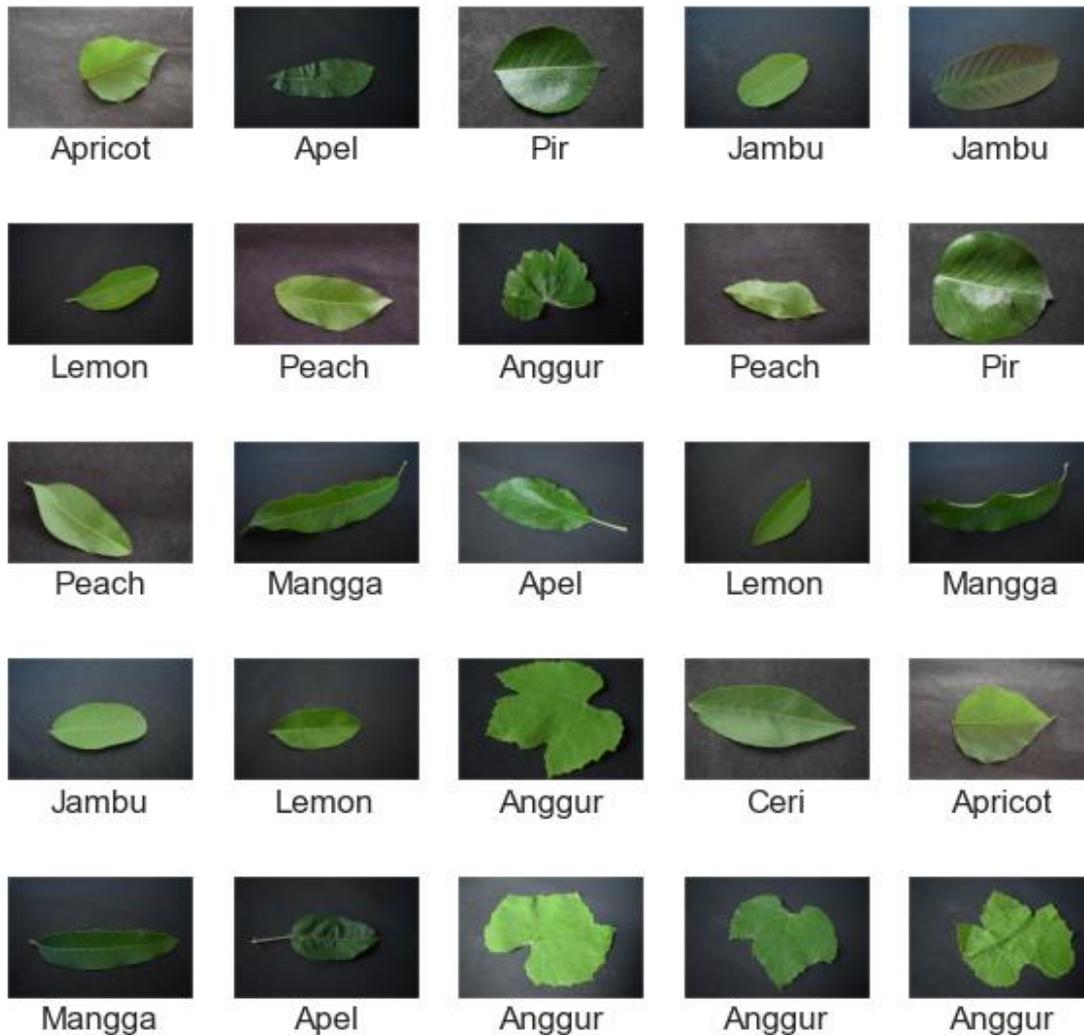
Data citra memiliki dimensi panjang dan lebar 6000 x 4000 piksel. Ukuran dimensi ini cukup besar sehingga kurang efisien saat proses oleh *machine learning* yang mempunyai ukuran ideal 256 x 256 piksel. Semakin besar ukuran data, hasil akan menjadi samar karena terlalu banyak fitur yang digunakan serta akan semakin berat dan lama dalam memproses model. Oleh karena itu perlu dilakukan *pre-processing image* dengan kompresi dimensi data menggunakan tools Caesium Image Compressor untuk memaksimalkan ukuran ideal *machine learning* yaitu 252 x 168 untuk setiap citra dalam dataset. Beberapa citra daun yang digunakan dalam penelitian ini seperti pada gambar 5.



Gambar 4. Ukuran 6000 x 4000 px

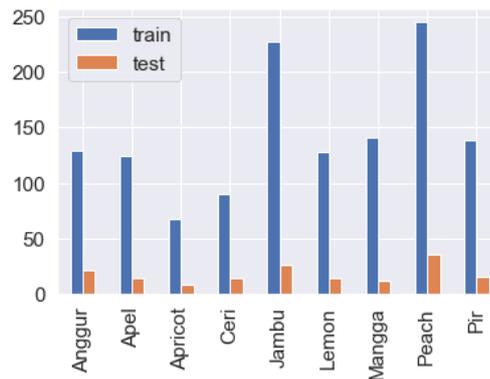


Gambar 5. Ukuran 252 x 168 px



Gambar 5. Sampel Data Secara Random

Dataset citra daun yang terdiri dari sembilan kelas dan sembilan direktori dibagi menjadi dua *set* yaitu *set train* dan *set val* dengan *train ratio* 0.8 yang berarti 80% data akan dialokasikan ke dalam direktori *train* dengan sub folder 9 kelas begitu juga 20% data akan dialokasikan ke dalam direktori *validation*. Komposisi data train dan data test disajikan dalam Gambar 6.

Gambar 6. Rasio *Data Train* dan *Data Val*

MODEL

Peneliti membuat model sesuai dengan arsitektur CNN dengan proses *input data shape* sebesar (252, 168, 3), dengan mengatur *feature maps* tiap *layer* menggunakan aktivasi yang disarankan yaitu ReLu. Pada proses model klasifikasi menggunakan aktivasi softmax yang dapat menggunakan data kategori atau kelas lebih dari dua [16]. Klasifikasi dilakukan *multi-class* menggunakan *batch size* dan epoch menggunakan 4 model tersebut (semakin tinggi *batch size* memori yang dibutuhkan akan semakin banyak pula), *target size* berdimensi (168, 252), dengan langkah dalam setiap *epoch* nya sebanyak 20. *Epoch* adalah proses pada jaringan algoritma pada sekali putaran yang ditentukan oleh step per *epoch* itu sendiri sebagai banyak jaringan yang dilewati. Dalam satu putaran *epoch* jumlah *batch* tergantung dari *batch size* dimana *batch size* adalah jumlah data dalam satu *batch* [17]. Jika menggunakan *batch size* 15 maka dari 1608 data akan dibagi per *batch* dengan total data 15.

ACCESS

Model klasifikasi yang telah dibuat pada proses sebelumnya akan dinilai dan dievaluasi seberapa baik kinerja model tersebut. Evaluasi dapat dilakukan dari meninjau hasil akhir model berupa *accuracy*, *precision* and *recall*.

Nilai *accuracy* adalah tingkat ketepatan model dengan prediksi benar sesuai dengan fakta yang ada. Nilai *loss* adalah jumlah *error* yang terjadi dalam setiap iterasi. Semakin rendah *loss* maka semakin baik model yang dibuat. *Precision* adalah persentase data yang di prediksi positif. *Recall* adalah persentase data yang aktual positif. *F1 score* adalah rata-rata *precision* dan *recall* dengan indikasi 0 sebagai nilai terburuk dan 1 sebagai nilai terbaik [18].

HASIL

Implementasi sistem yang digunakan adalah Virtual Machine yang dikembangkan oleh Google yaitu Google Colab Pro Research dengan menggunakan bahasa pemrograman python dan *library* Tensorflow Keras. Dari pembuatan *machine learning* klasifikasi daun menggunakan algoritma CNN diperoleh *data train* yang digunakan dalam pelatihan model dan *data validation* yang dilakukan berurutan setelah *training*. Pengujian dilakukan dengan dimensi (*batch size*, *epoch*) berdasarkan 4 model yaitu (15,50), (15,100), (30,50), (30,100) sesuai dengan Tabel 1,2 dan 3.

Tabel 1. Hasil *Accuracy* Berdasarkan Jumlah *Epoch* dan *Batch Size*

<i>batch size \ epoch</i>	Data Train <i>accuracy</i>		Data Validation <i>accuracy</i>	
	50	100	50	100
15	91%	89,67%	86,67%	92%
30	92,09%	95,96%	94,67%	98%

Tabel 2. Hasil *Precision*, *Recall* dan *F1 Score* Berdasarkan Jumlah *Epoch* dan *Batch Size*

<i>batch size \ epoch</i>	Data Train			Data Validation		
	<i>precision</i>	<i>recall</i>	F1 <i>score</i>	<i>precision</i>	<i>recall</i>	F1 <i>score</i>
15, 50	92,59%	88%	90,15%	88,76%	85,33%	86,99%
30, 50	92,78%	89,96%	91,31%	95,93%	93,33%	94,61%
15, 100	91,02%	88,33%	89,63%	93,05%	90,67%	91,82%
30, 100	96,42%	95,62%	96,01%	98%	98%	98%

Tabel 3. Hasil *Loss* Berdasarkan Jumlah *Epoch* dan *Batch Size*

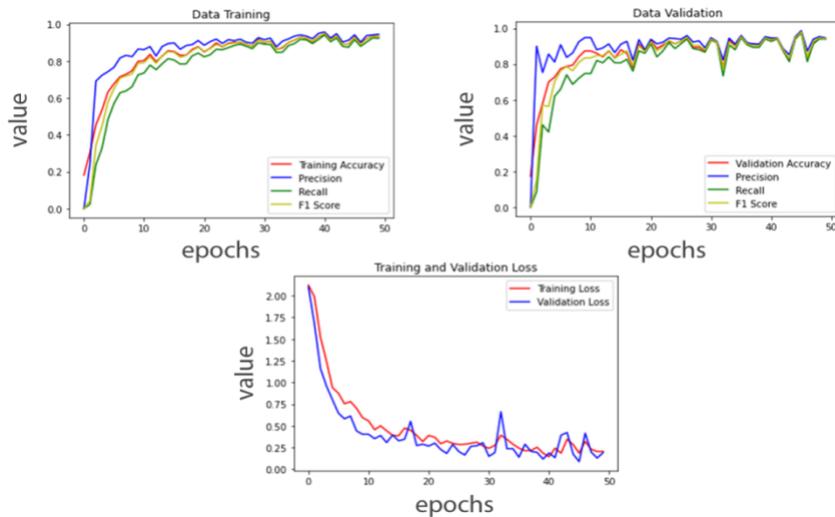
<i>batch size \ epoch</i>	<i>loss Train</i>		<i>loss Validation</i>	
	50	100	50	100
15	0,2940	0,3363	0,3406	0,1863
30	0,2142	0,1279	0,2069	0,0537

Dalam tabel 1,2 dan 3 menginterpretasikan nilai nilai *accuracy*, *precision*, *recall*, *f1 score* dan *loss* dari 4 model yang dibuat berdasarkan *epoch* dan *batch size*. Ditinjau dari Tabel 2, model yang menggunakan *batch size* 30 dan *epoch* 100 memiliki *accuracy validation* tertinggi yaitu 98% dan yang terendah sebesar 86,67% menggunakan *batch size* 15 dan *epoch* 50.

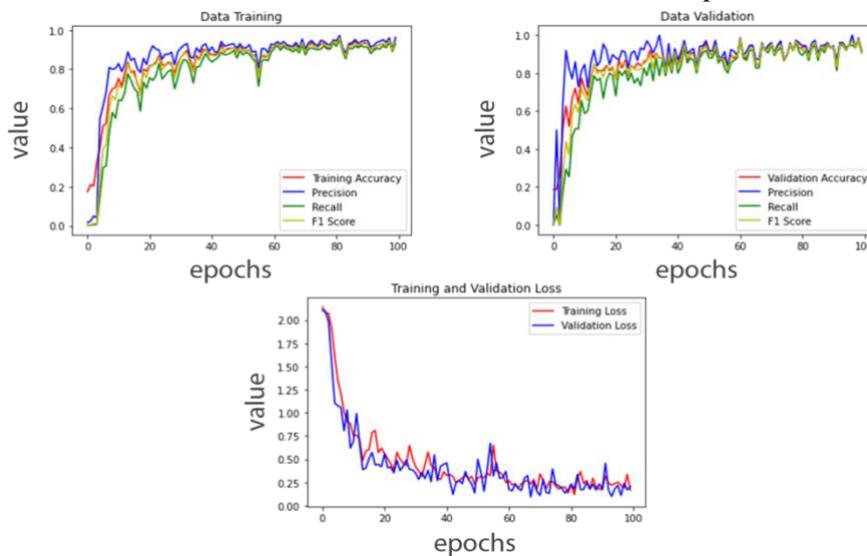
PEMBAHASAN

Berdasarkan hasil penelitian menggunakan CNN dengan metode SEMMA untuk mengklasifikasikan 9 kelas daun membuktikan memudahkan peneliti dalam setiap proses dengan terstruktur. Diawali dengan mencari data yang valid digunakan, mengeksplorasi ide bersangkutan dengan data yang didapatkan, *pre-processing image* daun menjadi 252x168x3 piksel, membuat sebuah model klasifikasi *multi-class* dan evaluasi model untuk proses terakhir.

Untuk membandingkan 3 nilai utama *accuracy*, *precision* dan *recall*, peneliti memilih menggunakan *accuracy* sebagai acuan karena nilai *precision* dan *recall* yang dianggap mendekati. Dari keempat model yang dibangun yaitu (15,50), (15,100), (30,50), (30,100) secara bertahap, nilai *accuracy* meningkat dari 86,67%, 92%, 94,67% hingga 98%. Namun pada *batch size* dan *epoch* (15,100) dan (30,50) membuktikan bahwa *batch size* yang lebih besar walaupun dengan *epoch* yang lebih sedikit memiliki *accuracy* yang lebih besar daripada *batch size* yang lebih kecil dan *epoch* yang lebih banyak. Grafik perbandingan pelatihan (30,50) dan (15,100) disajikan pada Gambar 7 dan 8.



Gambar 7. Grafik Pelatihan $batch\ size = 30$, dan $epoch = 50$

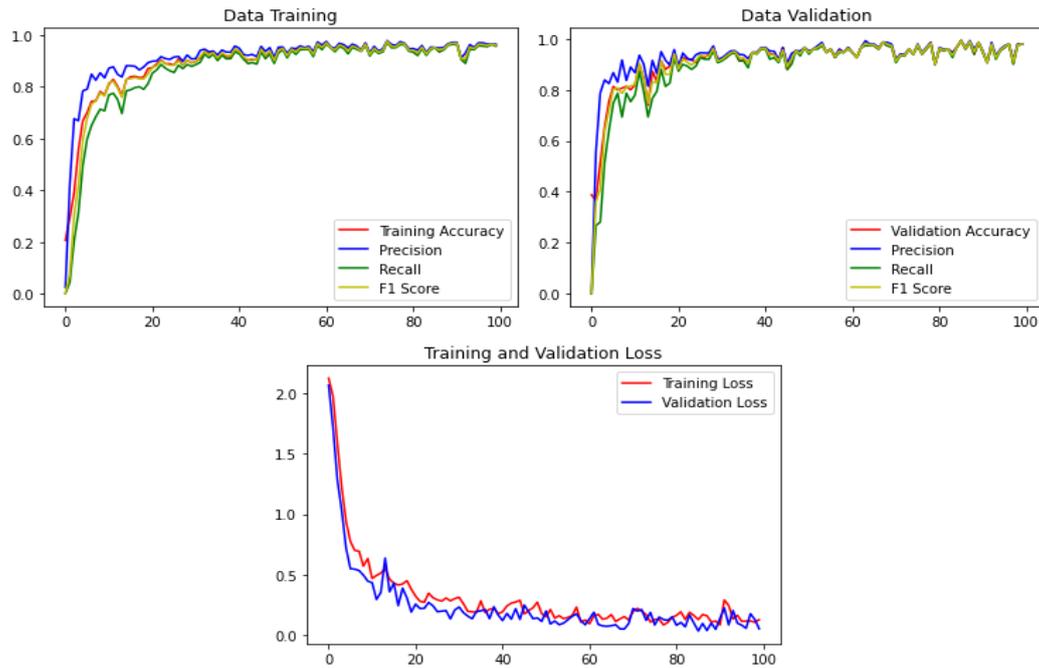


Gambar 8. Grafik Pelatihan $Batch\ Size = 15$, dan $Epoch = 100$

Dari perbandingan kedua model tersebut, model (15, 100) memiliki fluktuasi tinggi dan hampir terjadinya *overfitting* namun dari grafik tersebut pelatihan berjalan secara normal. Sedangkan model (30,50) grafik terlihat cukup stabil dengan fluktuasi yang tidak terlalu tinggi dan tidak terlalu rendah. Untuk hasil model (30,100) disajikan dalam tabel 4 dan gambar 9.

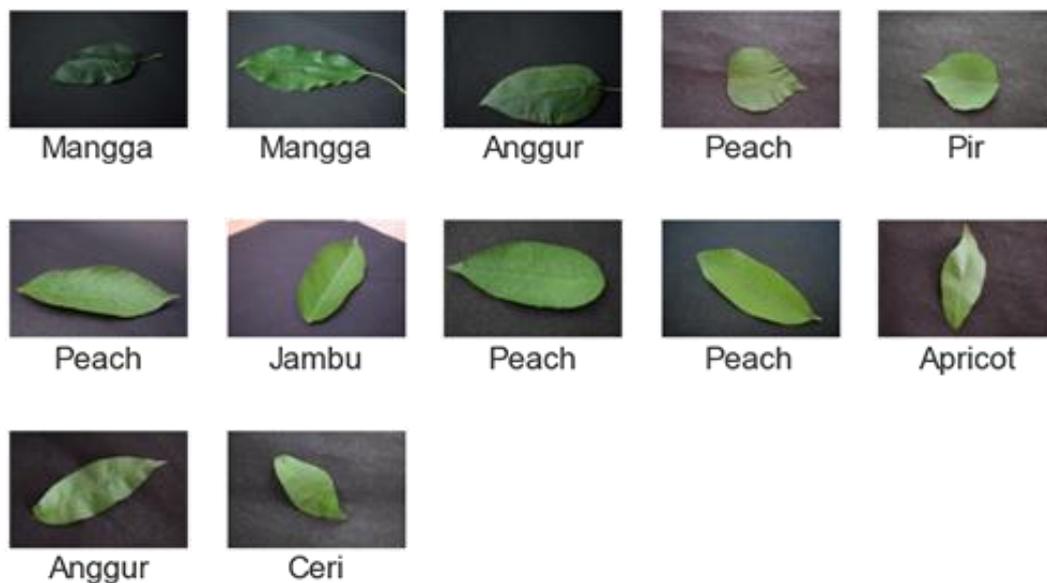
Tabel 4. Hasil Pelatihan Model $Batch\ Size = 30$, dan $Epoch = 100$

Epoch	Data Train		Data Validation	
	accuracy	loss	val accuracy	val loss
1	20,67%	2,1231	38,67%	2,0656
2	29,33%	1,9735	36,67%	1,7172
3	39%	1,5883	51,33%	1,2846
98	95,62%	0,1173	90,67%	0,1777
99	96,33%	0,1077	98%	0,1301
100	95,96%	0,1279	98%	0,0537



Gambar 9. Grafik Pelatihan *Batch Size* = 30, dan *Epoch* = 100

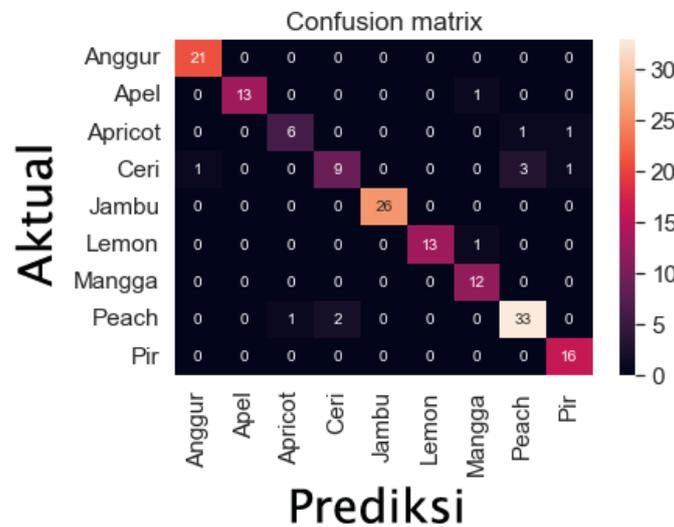
Citra daun yang telah diproses dengan sedemikian rupa diprediksi dengan menggunakan mesin yang sudah dilatih. Dengan dimensi piksel dan warna sebagai karakteristik utama dan mesin hasil terbaik dengan *batch size* 30 dan *epoch* 100 didapatkan kurang lebih 98% akurasi dari total citra daun sebanyak 1608. Untuk kesalahan prediksi citra daun disajikan pada gambar 10.



Gambar 10. Kesalahan Prediksi Citra

Confusion Matrix yang didapat dari 9 kelas daun tersebut pada *epoch* 100 dan *batch size* 30 adalah *True Positive* (TP) sebanyak 149 dari 161 data yang diambil. Sisa 12 data yang tidak didapati TP adalah *True Negative*, *False Positive*, dan *False Negative*. *True Positive* adalah data yang diprediksi benar dan hasil yang ditunjukkan pula benar. *True Negative*

adalah data yang diprediksi tidak benar tetapi sebenarnya memang tidak benar. False Positive adalah data yang diprediksi benar tetapi sebenarnya tidak benar. False Negative adalah data yang diprediksi salah tetapi hasil sebenarnya benar. Untuk confusion matrix yang dihasilkan disajikan pada gambar 11.



Gambar 11. Confusion Matrix Model Batch Size 30, dan Epoch 100

True Positif atau daun yang diklasifikasi benar di representasikan oleh nilai-nilai berwarna yang membentuk diagonal yaitu citra yang diprediksi sesuai dengan fakta pada kelasnya masing-masing. Sedangkan nilai-nilai di luar garis diagonal tersebut menunjukkan hasil yang *error*, tidak sesuai untuk penempatan kelasnya. Jadi meskipun dalam pelatihan mesin mencapai 98% dari pengujian menggunakan citra lain menunjukkan penurunan akurasi dalam mengklasifikasi yang dibuktikan dengan klasifikasi *error* yang lebih tinggi mencapai 8%.

KESIMPULAN DAN SARAN

Selain metode CRISP-DM dan KDD, metode SEMMA terbukti dapat mempermudah peneliti untuk membangun sebuah model dalam kronologi proses yang runtut. Ke empat model yang dibangun menggunakan CNN juga dapat mengklasifikasikan daun ke kelasnya masing-masing dengan hasil yang cukup baik dan bervariasi.

Berdasarkan variasi hasil implementasi terhadap keempat model yang dibangun dapat diketahui bahwa hubungan nilai *accuracy* dengan nilai *loss* pada pelatihan bergantung pada jumlah *epoch* dan banyaknya data yang digunakan. Dibuktikan dengan relevansi antar jumlah *epoch* dan besar *batch size* dimana ketika *epoch* semakin besar dan *batch size* yang semakin banyak maka *accuracy* akan semakin tinggi dan *loss* akan semakin kecil. Pada penelitian ini dengan menggunakan 1608 citra dengan 9 kelas, total *accuracy* tertinggi yang didapatkan sebesar 98% dengan *loss* 0,0537 menggunakan *batch size* 30 dan *epoch* 100.

Bahan pertimbangan pada penelitian selanjutnya dengan menambahkan *epoch* dan *batch size* pada jumlah data yang besar. *Target size* juga dapat dipertimbangkan untuk memberikan *input* yang optimal kepada model. Keakuratan data dapat disebabkan oleh *epoch* dan *batch size* yang terlalu sedikit dengan data yang banyak [19]. Memodifikasi CNN dengan konvolusi dan *pooling* yang lebih banyak akan mempengaruhi hasil yang lebih baik. Selain itu, menambahkan atribut klasifikasi juga dapat memperluas wawasan mesin dalam klasifikasi

DAFTAR PUSTAKA

- [1] K. P.-A. Saputra S. Mochammad Iswan, “Ekstraksi Fitur Morfologi Daun Sebagai Penciri Pada Tanaman Obat,” *Semin. Nas. Apl. Teknol. Inf.*, no. 2018, pp. 13–17, 2018, [Online]. Available: <https://journal.uui.ac.id/Snati/article/view/11150>
- [2] R. I. M. D. Fachrina Wibowo, Rosmayati, “PENDUGAAN PEWARISAN GENETIK KARAKTER MORFOLOGI HASIL PERSILANGAN F2 TANAMAN KEDELAI (*Glycine max* (L.) Merr. PADA CEKAMAN SALINITAS,” *J. Pertan. Trop.*, vol. 6, no. 2, pp. 180–189, 2019, doi: 10.32734/jpt.v8i1.6252.
- [3] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021. [Online]. Available: <https://doi.org/10.1007/s12525-021-00475-2>
- [4] M. Alhag Alobed, A. Ahmed, and A. Osman Ibrahim, “Classification of breast cancer image using data mining techniques,” *American Journal of Data Mining and Knowledge Discovery*, vol. 6, no. 2, pp. 31–35, Nov. 2021.
- [5] K. Fricklas, *Nishant Shukla*. 2018. [Online]. Available: http://manning-content.s3.amazonaws.com/download/3/e0d6f80-038c-49b5-9c3d-1fd3c3bc9e4c/sample_ch10_Shukla_Maching-Learning_January23.pdf
- [6] P. C. Richardson, “Image Classification,” US 8.478,052 B1, 2013 [Online]. Available: <https://patents.google.com/patent/US8478052B1/en>
- [7] Y. Yunitasari, “KLASIFIKASI GAMBARMENGGUNAKAN METODE K-NEAREST NEIGHBOR IMAGE CLASSIFICATION USING K-NEAREST NEIGHBOR METHOD,” in “*Inovasi Disruptif Teknologi Informasi di Era Normal Baru*,” 2020, pp. 451–457. [Online]. Available: <http://prosiding.unipma.ac.id/index.php/SENATIK/article/view/1633/1339>
- [8] H. H. Setiawan, “KLASIFIKASI JENIS BUAH PISANG DENGAN IMAGE PROCESSING MENGGUNAKAN METODE BACKPROPAGATION,” Yogyakarta, 2018. [Online]. Available: https://repository.usd.ac.id/31630/2/135314126_full.pdf
- [9] N. Neneng, K. Adi, and R. Isnanto, “Support Vector Machine Untuk Klasifikasi Citra Jenis Daging Berdasarkan Tekstur Menggunakan Ekstraksi Ciri Gray Level Co-Occurrence Matrices (GLCM),” *J. Sist. Inf. Bisnis*, vol. 6, no. 1, pp. 1–10, 2016, doi: 10.21456/vol6iss1pp1-10.
- [10] M. C. Wujaya and L. W. Santoso, “Klasifikasi Pakaian Berdasarkan Gambar Menggunakan Metode YOLOv3 dan CNN,” *J. INFIA*, vol. 9, no. 1, pp. 2–7, 2021, [Online]. Available: <https://publication.petra.ac.id/index.php/teknik-informatika/article/view/10930/9720>
- [11] R. Jose, “A CONVOLUTIONAL NEURAL NETWORK (CNN) APPROACH TO DETECT FACE USING TENSORFLOW AND KERAS,” *JETIR*, vol. 6, no. 5, pp. 97–103, 2019, [Online]. Available: www.jetir.org
- [12] Y. Vita Via, I. Yuniar Purbasari, and A. Putra Pratama, “ANALISA ALGORITMA CONVOLUTION NEURAL NETWORK (CNN) PADA KLASIFIKASI GENRE MUSIK BERDASAR DURASI WAKTU,” *SCAN J. Teknol. dan Inf.*, vol. 17, no. 1, pp. 35–41, 2022, [Online]. Available: <http://ejournal.upnjatim.ac.id/index.php/scan/article/view/3251/2003>
- [13] U. Shafique and H. Qaiser, “A Comparative Study of Data Mining Process Models (KDD , CRISP-DM and SEMMA),” *Int. J. Innov. Sci. Res.*, vol. 12, no. 1, pp. 217–222, 2014, [Online]. Available: <http://www.ijisr.issr-journals.org/>
- [14] Dataset PlantaeK. [Online]. CC By 4.0. Tersedia di : https://www.tensorflow.org/datasets/catalog/plantae_k
- [15] Dataset Plant_leaves. [Online] . CC By 4.0 Tersedia di : https://www.tensorflow.org/datasets/catalog/plant_leaves

- [16] S. Ilahiyah and A. Nilogiri, “Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network,” *JUSTINDO (Jurnal Sist. dan Teknol. Inf. Indones.*, vol. 3, no. 2, pp. 49–56, 2018, [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JUSTINDO/article/view/2254/1831>
- [17] D. Imam, *Memahami epoch batch size dan Iteration*, Jan 25, 2018. Diakses pada: September 6, 2022. [Online]. Tersedia di: <https://imam.digmi.id/post/memahami-epoch-batch-size-dan-iteration>
- [18] F. Paraijun, R. N. Aziza, and D. Kuswardani, “Implementasi Algoritma Convolutional Neural Network Dalam Mengklasifikasi Kesegaran Buah Berdasarkan Citra Buah,” *KILAT*, vol. 11, no. 1, pp. 1–9, 2022, doi: <https://doi.org/10.33322/kilat.v11i1>.
- [19] M. Wasil, H. Harianto, and F. Fathurrahman, “Pengaruh Epoch pada Akurasi menggunakan Convolutional Neural Network untuk Klasifikasi fashion dan Furniture,” *Infotek J. Inform. dan Teknol.*, vol. 5, no. 1, pp. 53–61, 2022, doi: 10.29408/jit.v5i1.4393.



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)
