

Klasifikasi Kualitas dan Kematangan Pisang Cavendish Menggunakan Convolutional Neural Network

*Arya Widya Hastungoro¹, Aditya Dwi Putro Wicaksono², Yesy Diah Rosita³

^{1,2,3}Teknik Informatika, Fakultas Informatika, Telkom University

^{1,2,3}Jl. DI Panjaitan No.128, Kabupaten Banyumas, Jawa Tengah

Email: ¹20102273@ittelkom-pwt.ac.id, ²aditya@ittelkom-pwt.ac.id, ³yesy.diah@ittelkom-pwt.ac.id

ABSTRACT

This research aims to develop a classification model using Convolutional Neural Networks (CNN) to determine the ripeness and quality of Cavendish bananas. The model classifies bananas into four categories: good quality unripe (MHBS), poor quality unripe (MHBK), good quality ripe (MGBS), and poor quality ripe (MGBK), using a total of 1,000 images. In this study, the classification process of the ripeness and quality of Cavendish bananas was carried out based on automatic feature extraction using CNN, after which an evaluation was carried out using a confusion matrix to assess model performance. The research developed 36 models with variations in parameters such as the number of epochs, batch size, and dataset split. The analysis results indicate that the number of epochs significantly affects the model's accuracy, with an increase in the number of epochs leading to higher accuracy. However, the dataset split scenario and batch size do not have a significant impact on the model's overall accuracy. Evaluation shows that the highest accuracy of 95% was achieved by the model with a 90:10 dataset split, a batch size of 16, and 20 epochs.

Keywords : deep learning; CNN; classification; cavendish banana

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model klasifikasi menggunakan Convolutional Neural Network (CNN) untuk menentukan kematangan dan kualitas pisang Cavendish. Model ini terdiri dari empat kategori: mentah kualitas bagus (MHBS), mentah kualitas buruk (MHBK), matang kualitas bagus (MGBS), dan matang kualitas buruk (MGBK) dengan total 1000 gambar. Dalam penelitian ini, dilakukan proses klasifikasi kematangan dan kualitas pisang Cavendish berdasarkan ekstraksi fitur otomatis yang dimiliki oleh CNN setelah itu dilakukan evaluasi menggunakan confusion matrix untuk menilai kinerja model. Penelitian ini membangun 36 model dengan variasi parameter seperti jumlah epoch, batch size, dan split dataset. Hasil analisis menunjukkan bahwa jumlah epoch berpengaruh signifikan terhadap akurasi model, dengan peningkatan jumlah epoch menyebabkan peningkatan akurasi. Namun, skenario pembagian dataset dan batch size tidak memiliki pengaruh yang signifikan terhadap akurasi model secara keseluruhan. Evaluasi menunjukkan bahwa akurasi tertinggi mencapai 95% diperoleh oleh model dengan split dataset 90:10, batch size 16, dan jumlah epoch 20.

Kata kunci : *deep learning; CNN; klasifikasi; pisang cavendish*

1. PENDAHULUAN

Pisang merupakan satu dari banyaknya buah yang memiliki potensi besar di Indonesia. Merujuk Badan Pusat Statistik (BPS), jumlah produksi pisang di Indonesia mencapai 9,24 juta ton pada tahun 2022, meningkat sebesar 5,72% dibandingkan tahun 2021 yang mencapai 8,74 juta ton (Annur, 2023). Buah pisang mempunyai nilai ekonomi tinggi dan potensi pasar yang masih terbuka lebar (Putro Wicaksono, 2023). Indonesia memiliki sekitar 200 jenis pisang yang beraneka ragam. Berbagai varietas pisang tumbuh di berbagai wilayah, seperti pisang ambon, pisang Cavendish, pisang kepok, pisang raja, pisang tanduk, pisang barangan, pisang susu, dan pisang nangka. Pisang Cavendish menjadi salah satu jenis pisang yang sangat populer dan diminati oleh masyarakat saat ini (Gampur et al., 2022).

Buah pisang Cavendish dibudidayakan di Indonesia oleh petani perorangan dan perusahaan besar (Muktianto & Indriyani, 2022). Pengelolaan buah pisang di Kabupaten Banyumas terutama di Desa Kedondong saat ini hanya dikelola dalam skala kecil di pekarangan atau perkebunan milik

penduduk setempat menghasilkan buah pisang berkualitas memerlukan ketelitian dan pengetahuan yang tinggi. Faktor-faktor seperti waktu panen, cuaca, dan tingkat kelelahan petani dapat mempengaruhi kualitas hasil panen. Hasil wawancara dengan petani pisang Cavendish pada tanggal 21 November 2023 di Desa Kedondong Kabupaten Banyumas menunjukkan bahwa penilaian kematangan pisang masih dilakukan secara manual berdasarkan warna dan ukuran, yang dapat menyebabkan ketidak-konsistenan dalam hasil pemilahan dan kerugian bagi petani jika kesalahan terjadi. Oleh karena itu, pentingnya deteksi teknologi untuk menilai kualitas dan kematangan pisang lebih akurat guna mencegah kerugian dalam pemilahan dan pemasaran.

Penelitian tentang klasifikasi kualitas dan kematangan buah telah mendapatkan perhatian yang signifikan dalam beberapa tahun terakhir. Salah satu teknik yang banyak digunakan dalam tugas ini adalah pembelajaran mesin, khususnya *Convolutional Neural Network* (CNN). CNN telah terbukti efektif dalam pemrosesan citra karena kemampuannya dalam mengekstraksi fitur warna dan tekstur secara otomatis melalui filter yang dapat belajar mandiri.

Dalam konteks ini, penelitian oleh (Barburiceanu et al., 2021) menunjukkan bahwa CNN mampu mencapai akurasi hingga 99,74% dalam mengidentifikasi penyakit daun tanaman apel, melalui ekstraksi fitur tekstur.

Selain itu, (Mehenag et al., 2020) menggunakan CNN untuk mengklasifikasikan tujuh jenis buah, yaitu apel, pisang, mangga, jeruk, pir, nanas, dan stroberi, dengan hasil akurasi model mencapai 98,74%.

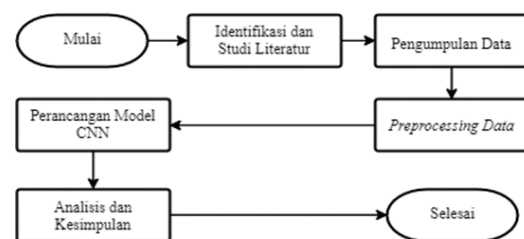
Penelitian ini bertujuan untuk menjelaskan dan menggambarkan penerapan model CNN dalam klasifikasi kematangan dan kualitas pisang Cavendish serta melakukan evaluasi menggunakan *confusion matrix* untuk menilai kinerja model. Perbedaan dari penelitian sebelumnya adalah fokus pada jenis buah tertentu, yaitu pisang Cavendish, yang merupakan salah satu varietas pisang yang populer di pasaran, sementara itu penelitian oleh (Barburiceanu et al., 2021), berfokus pada penyakit daun tanaman apel dan penelitian oleh (Mehenag et al., 2020) mencakup berbagai jenis buah, penelitian ini secara khusus menargetkan varietas pisang Cavendish untuk mengembangkan model yang lebih

spesifik dan optimal. Meskipun ada penelitian sebelumnya yang menggunakan CNN untuk klasifikasi buah secara umum, tidak ada penelitian yang secara khusus mengkaji kualitas dan kematangan pisang Cavendish. Celah ini memberikan peluang untuk mengembangkan model yang lebih spesifik.

Penelitian ini memiliki potensi untuk memberikan kontribusi terhadap pengembangan teknologi pemrosesan citra dalam dunia pertanian, khususnya dalam hal membantu proses pemilahan secara tepat dan efisien.

2. METODE

Pada penelitian ini, tahapan atau metode penelitian digambarkan pada Gambar 1.



Gambar 1. Diagram alir penelitian

2.1. Identifikasi Masalah

Sebelum memulai penelitian, penting untuk memahami masalah yang ingin diselesaikan, terutama pada kasus data yang kompleks. Hal ini meliputi menentukan apakah masalah tersebut

termasuk dalam klasifikasi, regresi, pengelompokan, atau jenis tugas *machine learning* lainnya (Pusporani et al., 2019). Dalam penelitian ini, dataset yang digunakan memiliki karakteristik klasifikasi multi-kelas, yang terdiri dari empat kategori yaitu, mentah kualitas bagus (MHBS), mentah kualitas buruk (MHBK), matang kualitas bagus (MGBS), dan matang kualitas buruk (MGBK).

2.2. Pengumpulan Data

Pengumpulan dataset pisang dilakukan secara mandiri dengan sudut pengambilan gambar sekitar 60° hingga 70° dan jarak sekitar 35 cm hingga 40 cm dari objek pisang. Pengambilan gambar dilakukan pada jam 11 sampai 13 siang bertempat di kebun pisang Cavendish, Desa Kedondong, Kecamatan Sokaraja, Kabupaten Banyumas. Total ada 400 gambar yang terkumpul dalam format .jpg, dengan masing-masing kategori memiliki 100 gambar dengan empat kategori. Pada Gambar 2, menunjukkan buah pisang dalam keadaan MHBS sedangkan pada Gambar 3, menunjukkan buah pisang dalam keadaan MHBK dan pada Gambar 4, menunjukkan buah pisang dalam keadaan MGBS sedangkan pada Gambar 5,

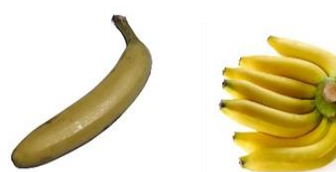
menunjukkan buah pisang dalam keadaan MGBK.



Gambar 2. Mentah Bagus



Gambar 3. Mentah Buruk



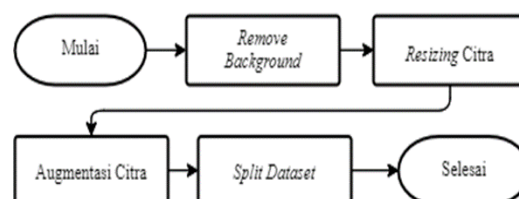
Gambar 4. Matang Bagus



Gambar 5. Matang Buruk

2.3. Preprocessing

Alur *preprocessing* yang diambil oleh peneliti dalam tahap ini mencakup *remove background*, *resizing citra*, *augmentasi citra*, dan *split dataset*. Gambar 6 menggambarkan alur dari tahap-tahap *preprocessing* tersebut.

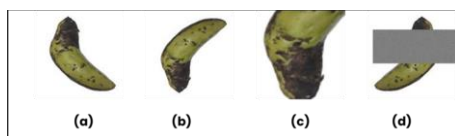


Gambar 6. Tahap *preprocessing* data

Pada bagian *resizing* citra, citra semula berukuran 3000 x 4000 *pixel* diubah menjadi 150x150 *pixel*.

2.4. Augmentasi Citra

Pada proses augmentasi citra, gambar-gambar diperbanyak menjadi 1000 citra dengan 250 citra tiap kategorinya dengan teknik *flip left-right(a)*, *flip top-bottom(b)*, *random zoom 50%(c)*, dan *random erasing(e)*. Berikut ini hasilnya pada Gambar 7.



Gambar 7. Hasil Augmentasi

2.5. Split Dataset

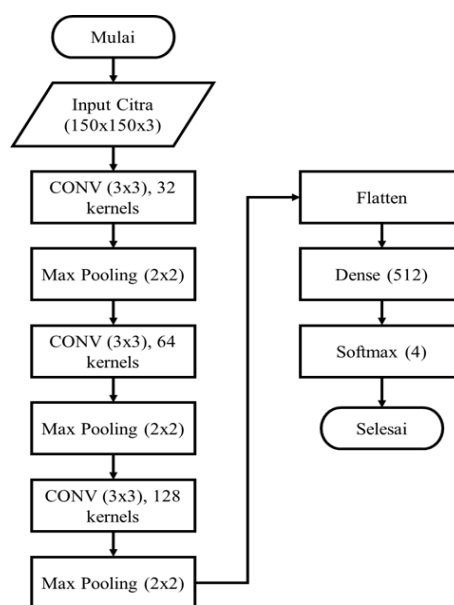
Pada proses *split dataset*, dibagi dua bagian yaitu pelatihan dan pengujian dengan proporsi yang berbeda: 60:40, 70:30, 80:20, dan 90:10.

2.6. Perancangan Model CNN

Proses pembuatan model ini menggunakan *input shape* dengan ukuran (150, 150, 3), yang berarti menggunakan citra dengan resolusi 150 x 150 piksel dan tiga *channel* warna (R, G, B) karena citra berwarna.

Pada lapisan konvolusi (*convolutional layer*), nilai filter yang digunakan adalah 32, 64, dan 128 untuk mendeteksi berbagai fitur gambar, mulai

dari tepi dan tekstur sederhana hingga objek dan pola kompleks (Iswantoro & Handayani, 2022; Molnar, 2022). Lapisan *MaxPooling2D* mereduksi ukuran citra dengan mengambil nilai maksimum dari matriks 2x2, yang membantu mempertahankan informasi penting dalam citra (Alzubaidi et al., 2021). Proses *flattening* mengubah array dua dimensi hasil *pooling* menjadi satu dimensi untuk input ke lapisan *Dense*. Lapisan *Dense* pertama memiliki 512 unit dengan fungsi aktivasi ReLU, sementara lapisan *Dense* terakhir, yang berfungsi sebagai output, memiliki 4 node sesuai jumlah kelas dengan fungsi aktivasi *softmax* untuk klasifikasi *multiclass* (Alzubaidi et al., 2021). Alur perancangan model CNN ini akan dijelaskan dalam Gambar 8.



Gambar 8. Arsitektur CNN

Setelah merancang arsitektur CNN, langkah berikutnya adalah menentukan *hyperparameter*. Perubahan pada *hyperparameter* bertujuan untuk mengamati variasi hasil pada akurasi dan *loss* saat pengujian model, berikut merupakan skenario yang akan diterapkan disajikan pada Tabel 1.

Tabel 1. Skenario *training data*

Parameter	Nilai
<i>Split Dataset</i>	60:40, 70:30, 80:20, dan 90:10
<i>Epoch</i>	10, 20, dan 30
<i>Batch Size</i>	8, 16, dan 32

Berdasarkan Tabel 1, total terdapat 36 model yang dilatih. Setelah pelatihan, dilakukan evaluasi untuk menentukan akurasi terbaik dengan menggunakan *confusion matrix*. Evaluasi ini mencakup penilaian model berdasarkan metrik akurasi, presisi, *recall*, dan *F1-Score*.

3. HASIL DAN PEMBAHASAN

Dalam tahapan augmentasi data, data citra diperbanyak dan berikut ini rincian tentang jumlah data citra yang dihasilkan pada Tabel 2.

Tabel 2. Data sebelum dan sesudah Augmentasi

Pisang Cavendish	Sebelum	Sesudah
MHBS	100	250
MHBK	100	250
MGBS	100	250
MGBK	100	250
Jumlah	400	1000

Dalam implementasi CNN ini, terdapat tiga lapisan konvolusi, dengan lapisan konvolusi terbesar memiliki 128 filter. Citra input yang digunakan memiliki ukuran 150x150 *pixel*.

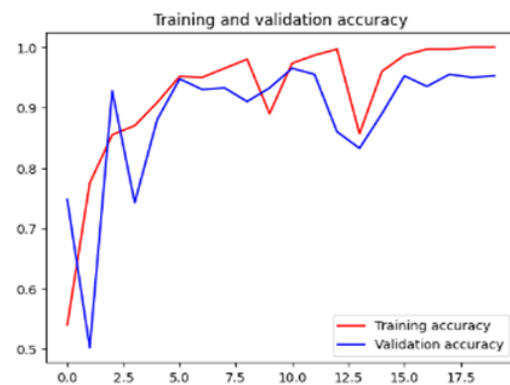
3.1. Hasil Training

Setelah melatih 36 model dengan berbagai skenario, satu model akan dipilih masing-masing satu model dari skema *split dataset* berdasarkan akurasi terbaik untuk dianalisis lebih lanjut. Hasilnya dicatat dalam Tabel 3.

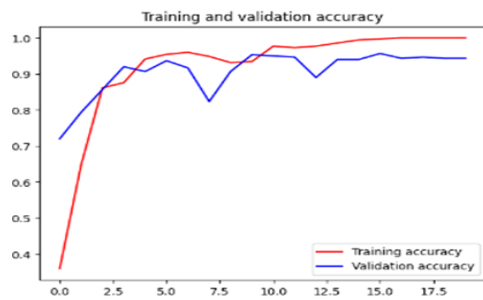
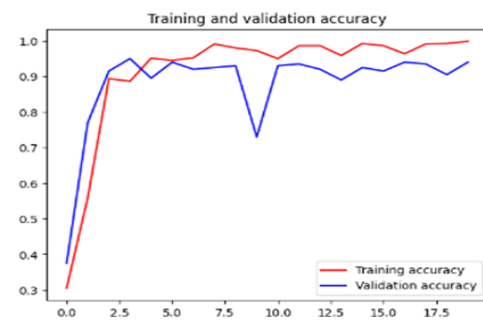
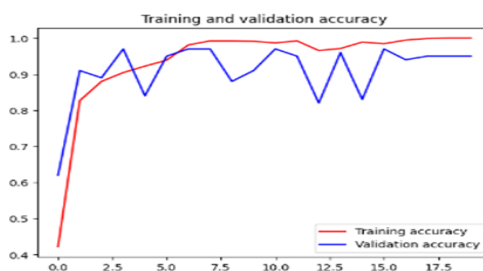
Tabel 3. Model terbaik tiap *Split Dataset*

Split Dataset	Acc	Val Acc	Loss	Val Loss
60:40	100	94	0.003	38
70:30	100	94	0.05	33
80:20	99	94	0.8	40
90:10	100	95	0.085	29

Tabel 3 menunjukkan detail hasil dari empat model yang menggunakan skema *split dataset* yang berbeda Berikut disajikan grafik akurasi dari tiap model pada Gambar 9, 10, 11, dan 12.



Gambar 9. Grafik akurasi *split dataset* 60:40

Gambar 10. Grafik akurasi *split dataset* 70:30Gambar 11. Grafik akurasi *split dataset* 80:20Gambar 12. Grafik akurasi *split dataset* 90:10

3.2. Pembahasan

Setelah *training* selesai dilakukan, berikut analisis hasil dari klasifikasi kualitas dan kematangan pisang Cavendish menggunakan CNN.

a. Pengaruh *hyperparameter*

Pada penelitian ini parameter yang diujikan adalah jumlah *epoch*, skenario *split dataset*, dan jumlah *batch size*. Berdasarkan hasil penelitian, parameter-parameter tersebut memiliki

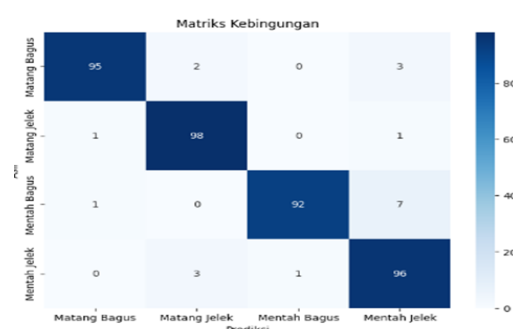
pengaruh terhadap hasil akurasi dan *loss* model. Berdasarkan hasil penerapan ketiga skenario jumlah *epoch*, maka didapatkan hasil bahwa penggunaan *epoch* 30 pada saat pelatihan model akan menghasilkan akurasi yang lebih tinggi dan *loss* yang lebih rendah dibandingkan dengan penggunaan *epoch* 20 atau 10. Hal tersebut menunjukkan bahwa semakin banyak proses pelatihan pada model semakin banyak informasi yang didapatkan sehingga akurasi model lebih tinggi dan *loss* model lebih rendah.

Skenario penggunaan parameter selanjutnya yang digunakan pada penelitian ini adalah *split dataset*. Hasil dari penerapan empat skenario pelatihan data menunjukkan adanya pengaruh pembagian data terhadap akurasi dan *loss* model. Pembagian data dengan rasio 90:10 menghasilkan akurasi tertinggi pada 2 model dan *loss* terendah pada 7 model. Namun parameter ini tidak mempengaruhi akurasi data secara keseluruhan, hal tersebut dikarenakan akurasi tertinggi pada penelitian ini tersebar pada semua skenario *split dataset*.

Parameter terakhir yang diterapkan pada penelitian ini adalah jumlah *batch size*. Hasil dari ketiga

skenario penggunaan *batch size* menunjukkan bahwa perbedaan dalam akurasi model tidak begitu signifikan antara *batch size* yang besar atau kecil, tetapi penggunaan *batch size* 8 cenderung menghasilkan nilai loss yang lebih rendah dibandingkan dengan *batch size* 16 atau 32.

b. *Split Dataset 60:40*



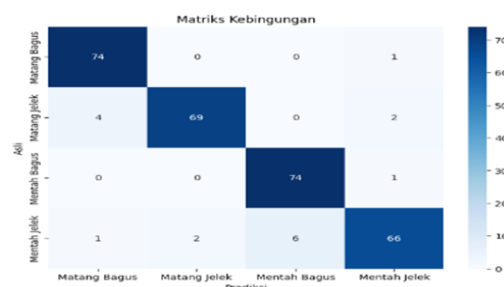
Gambar 13. *Confusion matrix split 60:40*

Pada Gambar 13 menunjukkan hasil klasifikasi dari model CNN dievaluasi menggunakan *confusion matrix* pada *split dataset* 60:40, dengan *batch size* 8 dan *epoch* 20. Hasil rata-rata penghitungan dapat dilihat pada Tabel 4.

Tabel 4. Hasil klasifikasi *split dataset* 60:40

Pisang Cavendish	Presisi	Recall	F1-Score
MGBS	0.98	0.93	0.95
MGBK	0.91	0.93	0.92
MHBS	0.96	0.96	0.96
MHBK	0.89	0.92	0.91
<i>Accuracy</i>			0.94

c. *Split Dataset 70:30*



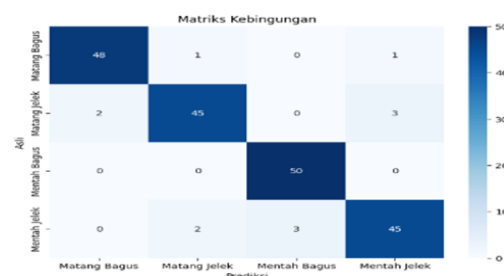
Gambar 14. *Confusion matrix split 70:30*

Pada Gambar 14 menunjukkan hasil klasifikasi dari model CNN dievaluasi menggunakan *confusion matrix* pada *split dataset* 70:30, dengan *batch size* 16 dan *epoch* 20. Hasil rata-rata penghitungan dapat dilihat pada Tabel 5.

Tabel 5. Hasil klasifikasi *split dataset* 70:30

Pisang Cavendish	Presisi	Recall	F1-Score
MGBS	0.94	0.99	0.96
MGBK	0.97	0.92	0.95
MHBS	0.93	0.99	0.95
MHBK	0.94	0.88	0.91
<i>Accuracy</i>			0.94

d. *Split Dataset 80:20*



Gambar 15. *Confusion matrix split 80:20*

Pada Gambar 15 menunjukkan hasil klasifikasi dari model CNN dievaluasi menggunakan *confusion matrix* pada *split dataset* 80:20, dengan *batch size* 16 dan *epoch* 20. Hasil rata-

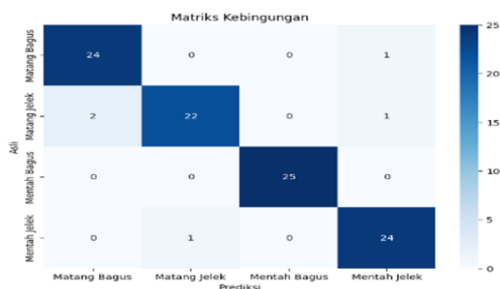
rata penghitungan dapat dilihat pada

Tabel 6.

Tabel 6. Hasil klasifikasi *split dataset* 80:20

Pisang Cavendish	Presisi	Recall	F1-Score
MGBS	0.96	0.96	0.96
MGBK	0.94	0.90	0.92
MHBS	0.94	1.00	0.97
MHBK	0.92	0.90	0.91
<i>Accuracy</i>			0.94

e. *Split Dataset* 90:10



Gambar 16. *Confusion matrix* split 90:10

Pada Gambar 16 menunjukkan hasil klasifikasi dari model CNN dievaluasi menggunakan *confusion matrix* pada *split dataset* 80:20, dengan *batch size* 16 dan *epoch* 20. Hasil rata-rata penghitungan dapat dilihat pada Tabel 7.

Tabel 7. Hasil klasifikasi *split dataset* 90:10

Pisang Cavendish	Presisi	Recall	F1-Score
MGBS	0.92	0.96	0.94
MGBK	0.96	0.88	0.92
MHBS	1.00	1.00	1.00
MHBK	0.92	0.96	0.94
<i>Accuracy</i>			0.95

Berikut perolehan *accuracy* yang didapatkan *split dataset* 60:40, 70:30, 80:20, 90:10 dapat dilihat pada Tabel 8.

Tabel 8. Hasil Testing

Split Datas	Presisi	Recall	F1-Score	Accuracy
60:40	94%	94%	94%	94%
70:30	94%	94%	94%	94%
80:20	94%	94%	94%	94%
90:10	95%	95%	95%	95%

4. KESIMPULAN

Berdasarkan hasil yang diperoleh, dapat diambil kesimpulan bahwa model CNN dalam penelitian ini berhasil dalam melakukan klasifikasi kualitas dan kematangan buah pisang Cavendish. Penelitian ini menggunakan model CNN untuk mengklasifikasikan kematangan dan kualitas pisang Cavendish, dengan citra input berukuran 150 x 150 piksel. Model terdiri dari tiga lapisan konvolusi, lapisan *max pooling* 2x2 dengan filter 32, 64, dan 128, satu lapisan *flatten*, dan lapisan *dense*. Berbagai parameter seperti jumlah *epoch* (10, 20, 30), *batch size* (8, 16, 32), dan skema *split dataset* (60:40, 70:30, 80:20, 90:10) dieksplorasi untuk mencari akurasi terbaik. Evaluasi yang dilakukan menggunakan *confusion matrix*, dan model terbaik dengan akurasi 95% pada data pengujian ditemukan pada skema *split dataset* 90:10, *batch size* 16, dan *epoch* 20.

DAFTAR PUSTAKA

Alzubaidi, L., Zhang, J., J. Humaidi, A., Al-Dujaili, A., Duan, Y., Al-

- Shamma, O., Santamaría, J., A. Fadhel, M., Al-Amidie, M., & Farhan, L. (2021, December). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1). 10.1186/s40537-021-00444-8
- Annur, C. M. (2023, June 22). *Jawa Timur Jadi Produsen Pisang Terbesar Nasional 2022, Berapa Jumlahnya?* Databoks. Retrieved May 13, 2024, from <https://databoks.katadata.co.id/datapublish/2023/06/22/jawa-timur-jadi-produsen-pisang-terbesar-nasional-2022-berapa-jumlahnya>
- Barburiceanu, S., Meza, S., Orza, B., Malutan, R., & Terebes, R. (2021, November). Convolutional Neural Networks for Texture Feature Extraction. Applications to Leaf Disease Classification in Precision Agriculture. *IEEE Access*, 9, 160085-160103. 10.1109/ACCESS.2021.3131002
- Fuadi, A., & Suharso, A. (2022). Perbandingan Arsitektur Mobilenet Dan Nasnetmobile Untuk Klasifikasi Penyakit Pada Citra Daun Kentang. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 7(3). <https://doi.org/10.29100/jupi.v7i3.3026>
- Gampur, G., Ordiyasa, I. W., & Mulyani, S. H. (2022, Agustus). Klasifikasi Jenis Pisang Menggunakan Convolutional Neural Network. *Prosiding Seminar Nasional Teknik Elektro, Informatika & Sistem Informasi (SINTaKS)*, 1(1). <https://doi.org/10.35842/sintaks.v1i1.5>
- Iswantoro, D., & Handayani, D. (2022, July). Klasifikasi Penyakit Tanaman Jagung Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Ilmiah Universitas Batanghari Jambi*, 22(2), 900-905. 10.33087/jiubj.v22i2.2065
- Mehenag, K., Julker, N., Forhad, A. M., Pritom, S., & Nakib, T. (2020, July). Fruits Classification using Convolutional Neural Network. *GRD Journals- Global Research and Development Journal for Engineering*, 5(8). ISSN- 2455-5703.
- Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Christoph Molnar.
- Muktianto, A., & Indriyani, V. (2022). Segmentasi Tingkat Kematangan Buah Pisang Cavendish Sangat Matang Berdasarkan Warna Menggunakan Watershed. *JURIKOM (Jurnal Riset Komputer)*, 9(1). <http://dx.doi.org/10.30865/jurikom.v9i1.3828>
- Putro Wicaksono, A. D. (2023, November). Klasifikasi Tingkat Kematangan, Kualitas dan Jenis Buah Pisang Berdasarkan Ciri Warna dan Bentuk Menggunakan Artificial Neural Networks. *Jurnal Teknologi Informasi Indonesia (JTII)*, 7(2), 91-98. <https://doi.org/10.30869/jtii.v7iNo2.955>
- Pusporani, E., Qomariyah, S., & Irhamah, I. (2019, March). Klasifikasi Pasien Penderita Penyakit Liver dengan Pendekatan Machine Learning. *Inferensi*, 2(1), 25-32. <https://doi.org/10.12962/j27213862.v2i1.6810>