



Klasifikasi Jenis Mangga Menggunakan Algoritma Convolutional Neural Network

Risma Yati*, Tatang Rohana, Adi Rizky Pratama

Fakultas Ilmu Komputer, Teknik Informatika, Universitas Buana Perjuangan Karawang, Karawang, Indonesia

Email: ^{1,*}if19.rismayati@mhs.ubpkarawang.ac.id, ²tatang.rohana@ubpkarawang.ac.id, ³adi.rizky@ubpkarawang.ac.id

Email Penulis Korespondensi: if19.rismayati@mhs.ubpkarawang.ac.id

Abstrak—Mangga yang bernama latin *Mangnifera IndicaL.* Merupakan buah yang berasal dari India dan menyebar sampai ke Indonesia. Terdapat berbagai jenis variasi mangga dengan bentuk dan warna yang berbeda sesuai dengan jenisnya. Untuk membedakan setiap mangga dilihat dari bentuk dan warnanya. Akan tetapi jika dalam proses panen petani mangga harus memilih secara manual memerlukan waktu yang lama dan berpotensi salah dalam menentukan jenisnya. Sehingga diperlukannya teknologi yang dapat mempermudah dalam membedakan jenis mangga berdasarkan bentuknya. Penelitian ini bertujuan untuk membuat model dengan akurasi terbaik pada proses klasifikasi 5 jenis mangga berdasarkan bentuknya. Data yang digunakan pada penelitian kali ini adalah 5 jenis mangga yaitu Mangga Apel, Mangga Arumanis, Mangga Gedong Gincu, Mangga Golek, dan Mangga Manalagi. Digunakan 375 citra mangga sebagai data set. Data set sebelum masuk kedalam proses training sebelumnya dilakukan tahap pre processing yang meliputi proses augmentasi dan resize. Setelah dilakukan tahap augmentasi jumlah citra menjadi 2250. Citra tersebutlah yang dipakai dalam proses training dan pengujian. Data set dibagi ke dalam 3 bagian yaitu 70% data training, 20% data validasi, dan 10% data test. Selanjutnya yaitu proses segmentasi, segmentasi yang digunakan dalam penelitian ini adalah segmentasi otsu. Proses klasifikasi menggunakan arsitektur Convolutional Neural Network (CNN) dengan 3layer convolution 16,32 dan 64, juga menggunakan optimizer Adam. Dilakukan 4 skenario percobaan untuk mencari nilai akurasi terbaik dengan membedakan learning rate dan batch size. Dari hasil pengujian dengan confusion matrix nilai akurasi terbaik dihasilkan dari hyperparameter input size100x100, epoch 100, learning rate 0,001 dan batch size 15 dengan nilai accuracy 99,56%, precision 100%, recall 100%, dan f1-score 100%.

Kata Kunci: Klasifikasi; Jenis Mangga; Bentuk; Convolutional Neural Network; Confusion matrix

Abstract—The name of the mango is *Mangnifera IndicaL.* It originated in India and spread to Indonesia. There are various types of mango variations with different shapes and colors according to the type. To distinguish each mango is seen by its shape and color. However, if in the harvest process mango farmers have to choose manually it takes a long time and potentially mistaken in determining the type. So it needs technology that can make it easier to differentiate the type of mango based on its shape. The study aims to create models with the best accuracy on the process of classifying 5 types of mango based on its shape. The data used in the research this time there are 5 types of mango that will be classified, namely Mangga Apel, Arumanis mango, Mangga Gedong Gincu, Golek mango and Mangga Manalagi. Used 375 images of mango as data sets. The data set before entering the previous training process is undergoing a pre-processing phase that includes the augmentation and resize process. The number of images increased to 2250. The data set is divided into three parts: 70% training data, 20% validation data, and 10% test data. Next is the process of segmentation, the segmentation used in this research is otsu segmentation. The classification process uses the Convolutional Neural Network (CNN) architecture with 3 layers of convolution 16,32 and 64, also using the Adam optimizer. 4 experimental scenarios were performed to find the best accuracy value by distinguishing between learning rate and batch size. From the confusion matrix test results, the best accuracy values were obtained from the input hyperparameter size100x100, epoch 100, learning rate 0,001 and batch size 15 with accurate values of 99.56%, precision 100%, recall 100%, and f1-score 100%.

Keywords: Classification; Type of Mangas; Forms; Convolutional Neural Network; Confusion matrix

1. PENDAHULUAN

Mangga (*Mangifera IndicaL.*) adalah buah yang asli dari India dan menyebar ke Asia Tenggara hingga Indonesia.[1] Dataran rendah dan hawa panas membuat tanaman mangga tumbuh dengan baik. Akan tetapi pada ketinggian yang mencapai 600 mdpl tanaman mangga dapat tumbuh.[2] Kandungan nutrisi dalam buah mangga baik bagi kesehatan yang dapat menaikkan daya tahan tubuh dan mengurangi resiko kerusakan mata dan sariawan yaitu vitamin A dan C.[3][4] Buah mangga sangat digemari masyarakat Indonesia sehingga masyarakat menanam pohon mangga untuk keperluan pribadi atau komersil.[4] Terdapat varietas dan jenis mangga yang populer di Indonesia yaitu jenis Mangga Arumanis, jenis Mangga Golek, jenis Mangga Endog, jenis Mangga Manalagi, jenis Mangga Laliwo, jenis Mangga Madu, jenis Mangga Pakel, jenis Mangga Kemang, dan jenis Mangga Kweni. Dari setiap jenis tersebut mempunyai bentuk yang bervariasi seperti bentuk pipih, bulat memanjang, bundar, dan bulat telur. Ukuran dari Panjang sekitar 2,5-30 cm, dengan warna yang berbeda-beda seperti hijau, kuning, merah, atau campuran dari ketiganya. Dari beragam jenis mangga yang tersebar di Indonesia tidak menjamin kelengkapan serta kemudahan terkait informasi dan akses dalam informasi terkait varietas mangga. Untuk meningkatkan produksi buah dan pemilihan jenis mangga unggulan yang baru diperlukannya ketersediaan informasi mengenai varietas buah mangga. [5]

Terdapat berbagai jenis variasi mangga dengan bentuk dan warna yang berbeda sesuai dengan jenisnya. Untuk membedakan setiap mangga dilihat dari bentuk dan warnanya. Dalam proses panen petani mangga harus memilih secara manual satu persatu dan memisahkannya sesuai jenis mangga nya. Proses memilih secara manual memerlukan waktu yang lama dan berpotensi salah dalam menentukan jenisnya. Ketika petani salah dalam



melakukan kesalahan pada proses memilih jenis mangga maka akan berpengaruh terhadap proses distribusi dan perdagangan pada konsumen atau pembeli. Sehingga diperlukannya teknologi yang dapat mempermudah dalam membedakan jenis mangga berdasarkan bentuknya. Untuk itu pada penelitian ini dilakukan pembuatan model yang akan digunakan untuk proses klasifikasi jenis mangga. Telah dilakukan beberapa penelitian mengenai mangga, baik klasifikasi tingkat kematangan, maupun jenis varietas mangga. Salah satunya tingkat kematangan mangga Badami menggunakan metode CNN. Pada penelitian tersebut citra yang digunakan adalah data uji berjumlah 25 citra dan citra latihan berjumlah 179 citra, adapun nilai akurasi dari pengujian model yaitu 97,2%. [6] Penelitian lain Dengan menggunakan metode Hue Saturation Intensity (HSI) yaitu membedakan tingkat kematangan Mangga Harumanis. 20 data digunakan untuk penelitian ini. Data tersebut terdiri dari 10 citra mangga matang sebagai data latihan, dan 10 citra sebagai data uji. Data uji tersebut terdiri atas 5 citra mangga yang sudah matang dan 5 citra mangga yang masih mentah. Adapun Hasil dari pengujian tersebut adalah 87%. [1] Ada Metode lain yang digunakan menentukan tingkat kematangan mangga yaitu metode GLCM yang disatukan bersama metode K-Means Clustering untuk mengekstraksi fitur buah mangga supaya bisa diaplikasikan pada varietas mangga lainnya. Metode yang digunakan untuk mengklasifikasi tingkat kematangan mangga yaitu LS-SVM. Terdapat 240 data set yang digunakan. 240 set merupakan hasil gabungan dari beberapa data jenis mangga, yaitu 80 data Mangga Harumanis, 80 data Mangga Manalagi, dan 80 Mangga Kent. Dari 80 data mangga memiliki 2 jenis data yaitu 40 data Mangga yang sudah matang dan 40 data Mangga yang masih Mentah. Dari proses pengujian menghasilkan akurasi dengan nilai 98,33% dengan data uji. [2] Penelitian berikutnya menggunakan citra jenis daun Mangga Golek, jenis daun Mangga Arumanis, dan jenis daun Mangga Manalagi sebagai data set nya, lalu digunakan metode Convolutional Neural Network (CNN) untuk mengklasifikasikan varietas mangga berdasarkan bentuk dan teksturnya. Hasil tes membutuhkan waktu rata-rata 2 detik untuk diselesaikan, 52 detik merupakan waktu terlama pada saat proses tes. Nilai akurasi yang didapatkan rata-rata nilai 1. [4]

Berdasarkan permasalahan diatas, maka peneliti akan melakukan penelitian untuk membedakan 5 jenis mangga. Citra mangga dari 5 jenis mangga dikumpulkan lalu dilakukan proses pre processing, segmentasi dan metode yang digunakan untuk membuat model klasifikasi jenis mangga berdasarkan bentuknya dengan menggunakan algoritma Convolutional Neural Network. Hasil yang sangat signifikan untuk pengenalan citra digital dapat dikaitkan dengan pendekatan Convolutional Neural Network. Hal tersebut karena CNN didasarkan pada sistem pengenalan gambar di korteks visual otak manusia. [7] Beberapa penelitian terdahulu telah membuktikan hasil akurasi yang bagus menggunakan algoritma CNN diantaranya Performa Metode Convolutional Neural Network dengan berdasarkan Citra Ct-Scan Thoax yang dipakai dalam Klasifikasi Covid-19 Varian Omicron. Dalam penelitian ini dilakukan pengujian dengan membedakan jumlah data set. Pengujian pertama menggunakan data sebanyak 400 data, iterasi (epoch) 30 kali, dan step_per_epoch 30 menghasilkan nilai akurasi 0.88 atau 88% dan nilai Loss 0.3119. Pada pengujian ke-dua menggunakan data sebanyak 800 data, dengan iterasi (epoch) 30 kali, dan step_per_epoch 30 menghasilkan akurasi yang bernilai 0.93 atau 93% dan Loss yang bernilai 0.2711. [8] Selanjutnya penelitian mengenai deteksi otomatis jerawat wajah metode yang digunakan adalah Convolutional Neural Network (CNN) adapun diperoleh hasil dari proses learning metode CNN mendapatkan nilai akurasi 99,8% hingga 100% [7] Selain itu terdapat pula sistem klasifikasi penyakit kulit menggunakan algoritma CNN menggunakan 1125 citra latihan dan 375 citra uji, hasil yang maksimal didapatkan pada saat parameter hidden layer 5, digunakan optimizer Adam, serta ukuran nilai dari learning rate yaitu bernilai 0,001 adapun hasil dari sistem adalah sebagai berikut : akurasi bernilai 96,53%, precision bernilai 95%, recall bernilai 95%, f1-score bernilai 95% dan Loss bernilai 0,2486. [9] Terdapat pula penelitian yang membandingkan nilai 6 learning rate yaitu 0,01, 0,001, 0,0001, 0,00001, 0,000001 dan 0,0000001. Untuk batch size membandingkan 3 angka yaitu 5, 10, 15. Convolutional Neural Network (CNN) merupakan arsitektur yang digunakan pada proses pengujian, hasil yang maksimal dari perbandingan tersebut adalah nilai learning rate yang bernilai 0,0001 dan learning rate yang bernilai 0,00001. Adapun nilai batch size yang paling baik nilainya adalah angka batch size dengan nilai paling kecil yaitu 5. [10]

Algoritma Convolutional Neural Network (CNN) memiliki kinerja yang bagus dalam proses pengklasifikasian. [7][8] [9], [10][10] Terdapat pula pada salah satu penelitian Learning rate dan batch size berpengaruh terhadap hasil akhir dari proses deep learning [10] dan uraian permasalahan mengenai pemilihan jenis mangga secara manual memerlukan waktu yang lama. Maka penelitian ini bertujuan untuk mendapatkan model dengan akurasi terbaik untuk proses klasifikasi 5 jenis mangga berdasarkan bentuknya. Digunakan algoritma Convolutional Neural Network (CNN) sebagai metode penelitiannya dan akan dilakukan empat skenario untuk mendapatkan akurasi terbaik dengan berdasarkan dua hyperparameter yaitu learning rate dan batch size. Selain itu model yang dihasilkan dari penelitian ini dapat dikembangkan lagi yang nantinya akan berguna untuk petani dalam proses pemilihan mangga secara otomatis, dan meminimalisir terjadinya kesalahan dalam proses pemilihan.

2. METODOLOGI PENELITIAN

2.1 Objek Penelitian

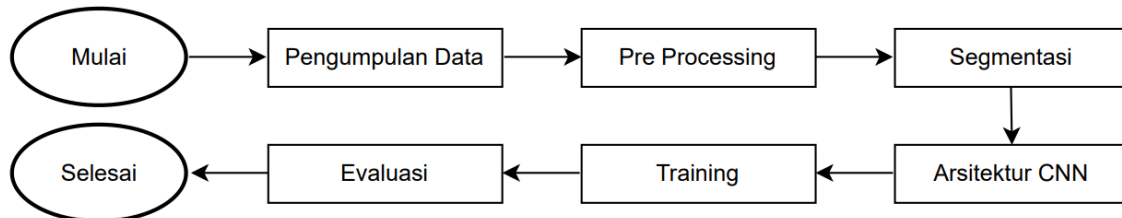
Citra mangga merupakan objek yang digunakan untuk proses penelitian ini. Terdapat lima jenis citra mangga yaitu Mangga Apel, Mangga Arumanis, Mangga Gedong Gincu, Mangga Golek, dan Mangga Manalagi. Citra mangga



digunakan sebagai data set yang akan dibagi untuk proses training, validasi saat pembuatan model menggunakan algoritma Convolutional Neural Network (CNN), dan confusion matrix digunakan saat tahapan evaluasi.

2.2 Tahapan Penelitian

Proses dari setiap tahapan penelitian merupakan hal yang dilakukan untuk mencapai tujuan penelitian. Adapun proses tahapan penelitian diawali dengan pengumpulan data, pre processing, segmentasi, pembuatan arsitektur CNN, training, dan evaluasi. Tahapan dari proses penelitian ditunjukkan pada Gambar 1 dibawah ini.



Gambar 1. Tahapan Penelitian

Gambar 1 merupakan tahap dari proses penelitian. Proses penelitian dimulai dengan pengumpulan data berupa 5 jenis citra mangga yaitu Mangga Apel, Mangga Arumanis, Mangga Gedong Gincu, Mangga Golek, dan Mangga Manalagi. Tahapan selanjutnya adalah proses pre processing, pada tahap ini dilakukan dua proses yaitu resize dan augmentasi. Selanjutnya adalah proses segmentasi dengan menggunakan segmentasi otsu. Tahap selanjutnya membuat arsitektur CNN, proses pembuatan arsitektur terdiri dari Convolution, Pooling, Flattening, dan Full Connection. Tahap selanjutnya adalah proses training, proses training menggunakan optimizer Adam, dan dilakukan 4 skenario untuk mendapatkan nilai akurasi terbaik berdasarkan 2 hyperparameter yaitu learning rate dan batch size. Setelah proses training selesai dan didapatkan sebuah model maka tahap terakhir adalah mengevaluasi model dengan menggunakan confusion matrix.

2.3 Pengumpulan Data

Pada penelitian ini menggunakan data primer dengan lima jenis citra mangga, yaitu Mangga Apel, Mangga Arumanis, Mangga Gedong Gincu, Mangga Golek, dan Mangga Manalagi. Gambar diambil menggunakan kamera smart phone dengan resolusi 12 MP. Dari hasil akuisisi didapatkan 375 citra. Rincian dari masing-masing citra adalah sebagai berikut:

Tabel 1. Jumlah Citra Mangga

Jenis Mangga	Jumlah Citra
Mangga Apel	75
Mangga Arumanis	75
Mangga Gedong Gincu	75
Mangga Golek	75
Mangga Manalagi	75

Tabel 1 merupakan jumlah citra pada setiap jenis mangga. Citra tersebut merupakan citra asli yang dihasilkan dari proses akuisisi.



Gambar 2. Citra Buah Mangga

Gambar 2 merupakan citra 5 jenis mangga yang akan digunakan pada penelitian ini. (a) Mangga Apel, (b) Mangga Arumanis, (c) Mangga Gedong Gincu, (d) Mangga Golek, dan (e) Mangga Manalagi. Mangga dari hasil akuisisi selanjutnya dilakukan proses pre processing. Setelah proses pre processing selesai data dibagi menjadi tiga yaitu sebagai data train, data validasi, dan data test. Data training adalah data yang digunakan algoritma untuk mengklasifikasikan dan melatih model untuk mendeteksi kelas baru yang belum pernah dilihat sebelumnya. Data ini biasanya memiliki rasio paling tinggi di antara ketiga set data. Data validation adalah data yang digunakan saat training yang berfungsi untuk mengoreksi nilai error. Sedangkan data test digunakan ketika model sudah selesai dilatih dan digunakan untuk membandingkan sebgus apa performa model dari hasil training yang



dilakukan. [11] Data dibagi dengan jumlah persentase 70%, 20%, dan 10%. Rincian dari pembagian data terdapat pada tabel 2

Tabel 2. Pembagian Data Set

Jenis Mangga	Training	Validasi	Test
Mangga Apel	70%	20%	10%
Mangga Arumanis	70%	20%	10%
Mangga Gedong Gincu	70%	20%	10%
Mangga Golek	70%	20%	10%
Mangga Manalagi	70%	20%	10%

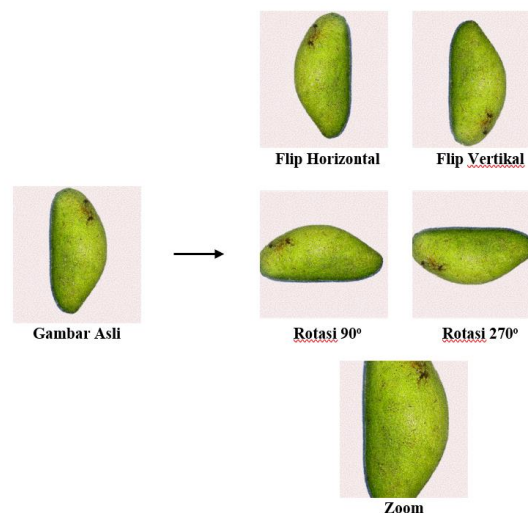
Tabel 2 merupakan tabel pembagian data set menjadi tiga bagian dari masing-masing jenis mangga yaitu data training, validasi dan data test.

2.4 Pre processing

Pre processing merupakan tahap pemrosesan data (pada penelitian ini citra) agar mendapatkan data yang layak untuk digunakan pada tahap berikutnya. [1] Pada tahap pre processing meliputi proses Augmentasi dan Resize.

a. Augmentasi

Augmentasi yaitu gambar atau citra yang ada akan dimanipulasi. [12] Augmentasi memiliki kemampuan meningkatkan variasi gambar yang digunakan untuk memperkecil overfitting. [13] Proses Augmentasi yang digunakan pada penelitian ini adalah flip horizontal, flip vertikal, rotasi 90°, rotasi 270°, dan zoom. Dari proses augmentasi jumlah citra menjadi 2250 citra. Citra hasil Augmentasi ditunjukkan pada Gambar 3.



Gambar 3. Hasil Augmentasi

Gambar 3 merupakan salah satu proses augmentasi yang dilakukan pada data set. Terlihat gambar asli yang kemudian dilakukan augmentasi yaitu flip horizontal, flip vertikal, rotasi 90°, rotasi 270°, dan zoom.

b. Resize

Resize merupakan merubah ukuran piksel suatu citra. [14] Tujuan dari resize ini adalah untuk memperkecil ukuran citra dan mempercepat saat diproses. [15] Pada penelitian ini hasil dari akuisisi citra memiliki ukuran yang beragam. Sehingga ukuran citra dibuat menjadi seragam menjadi ukuran 100x100 piksel dengan proses resize.

2.5 Segmentasi

Proses membagi atau memisahkan sebuah gambar menjadi beberapa area (region) atau objek dengan karakteristik yang identik dikenal dengan segmentasi. Hal ini memungkinkan untuk pengambilan individual dari setiap objek gambar dan penggunaannya sebagai input untuk proses lainnya. [16] Metode Otsu digunakan dalam penelitian ini. Histogram grayscale ditransformasikan menjadi citra biner menggunakan metode Otsu. Analisis diskriminan digunakan dalam metode Otsu untuk mengidentifikasi ciri-ciri yang membedakan dua ataupun lebih kelompok yang muncul secara alami. Analisis diskriminan akan memaksimalkan variabel-variabel ini untuk membedakan antara objek latar depan dan latar belakang. [17] Penjelasan metode otsu adalah sebagai berikut. Pertama-tama, probabilitas nilai intensitas I dalam histogram dihitung melalui persamaan 1:

$$P(i) = \frac{n_i}{N}, p(i) \geq 0 \sum_1^{256} p(i) = 1 \quad (1)$$



Ket :

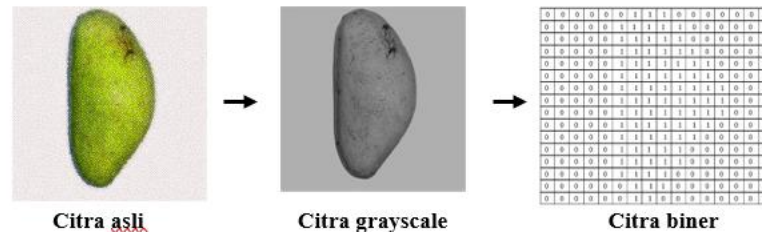
P : Probabilitas

i : nilai intensitas

ni : Jumlah piksel berintensitas

N : Jumlah Semua Piksel

Jumlah piksel berintensitas I dan N dinyatakan dengan ni.[18]



Gambar 4. Citra Segmentasi

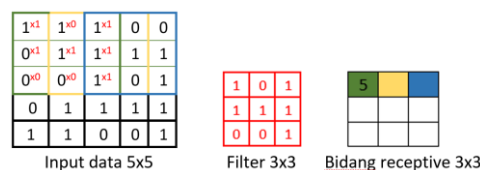
Gambar 4 merupakan gambar dari proses segmentasi hasil citra merupakan citra biner yang digunakan dalam proses training.

2.6 Arsitektur Convolutional Neural Network (CNN)

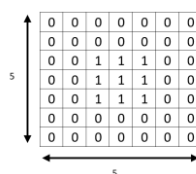
Convolutional Neural Network (CNN) termasuk ke dalam algoritma deep learning. Merupakan penyempurnaan dari Multilayer Perceptron (MPL). CNN merupakan algoritma yang dirancang untuk memproses gambar dan suara dalam dua dimensi. Memiliki kemampuan untuk langsung belajar dari gambar.[19] CNN meniru cara otak manusia memproses informasi visual untuk mengidentifikasi objek. Komputer kini dapat melihat dan membedakan berbagai objek berkat CNN. Kemampuan ini dikenal sebagai image recognition.[20] Susunan neuron disetiap lapisan neuron konvolusional adalah 3D (lebar, tinggi, kedalaman). Dengan lebar, tinggi adalah ukuran layer, dan kedalaman merupakan jumlah layer.[21] Feature Extraction Layer dan Fully Connected Layer merupakan dua bagian besar Arsitektur Convolutional Neural Network dibagi menjadi, yaitu. Proses “encoding” suatu citra ke dalam feature map berupa angka yang merepresentasikan citra tersebut berlangsung pada Feature Extraction Layer. [8] Tahapan dalam membangun arsitektur CNN yaitu :

a. Convolutional Layer

Convolutional Layer yaitu proses konvolusi akan terjadi pada setiap data yang bersentuhan dengan convolutional layer. Activation map atau featur map 2D akan dibuat sebagai hasil dari perubahan lapisan setiap filter agar cocok dengan semua data input. Hal ini dioptimalkan dengan mengatur ukuran filter, stride dan padding.[22] Proses convolution dilakukan sebanyak 3 kali, yaitu menggunakan Layer Filter 16, Layer Filter 32, dan Layer Filter 64. Ketiga proses tersebut menggunakan kernel 3x3 dan stride 1. Aktifasi yang digunakan pada saat proses convolutional adalah ReLU (Rectified Linear Unit). ReLU merupakan sebuah operasi untuk mengenalkan nonlinieritas dalam meningkatkan representasi model. [23]



Gambar 5. Proses Convolution

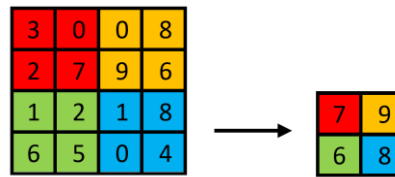


Gambar 6. Proses zero padding 2 pada data 3x3

Gambar 5 menunjukkan Proses convolution dengan stride 1 dengan input data 5x5, filter 3x3 dan bidang receptive 3x3. Adapun pada Gambar 6 menunjukkan proses zero padding 2 pada data 3x3.

b. Pooling

Proses membuat ukuran gambar lebih kecil dikenal sebagai pooling atau subsampling. Max Pooling adalah metode pooling yang digunakan dalam penelitian ini. Setelah membagi output convolution layer menjadi beberapa grid yang lebih kecil, Max Pooling mengkompilasi nilai maksimum dari setiap grid menjadi matriks gambar yang lebih kecil,

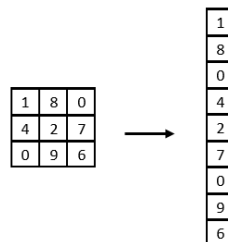


Gambar 7. Perhitungan Max-Pooling

Gambar 7 menunjukkan proses perhitungan dari max pooling. Output lapisan dibagi menjadi beberapa grid kemudian diambil nilai yang paling besar dan terbentuklah matrix baru.

c. Flattening

Merubah data Pooling yang mempunyai array 2 dimensi dirubah menjadi data satu dimensi single vector merupakan proses dari Flattening.

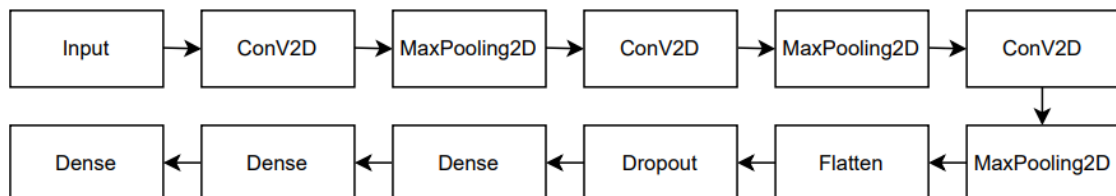


Gambar 8. Proses Flattening

Gambar 8 menunjukan proses Flattening gambar, sebelah kiri merupakan contoh gambar dari hasil pooling dengan ukuran matrix 3x3. selanjutnya dibentuk ke dalam 1 vektor dengan 9 baris.

d. Full Connection

Selanjutnya adalah proses dropout sebanyak satu kali. Parameter dikontrol melalui prosedur ini untuk mencegah overfitting. Lalu dilakukan proses dense sejumlah 3 kali. 1024unit node digunakan saat proses dense kesatu. 512unit node digunakan saat proses dense kedua, dan proses dense yang ketiga digunakan jumlah unit code sama dengan jumlah class yaitu 5. Pada saat proses dense kesatu dan kedua menggunakan aktifasi ReLU serta yang ketiga aktifasi softmax yang digunakan.



Gambar 9. Arsitektur CNN

Diatas telah dijelaskan mengenai pembuatan arsitektur Convolutional Neural Network (CNN). Gambar 9 menunjukkan struktur yang dijelaskan diatas dan digunakan dalam arsitektur CNN.

2.7 Training

Training dilakukan untuk mendapatkan model dengan akurasi terbaik. Proses training menggunakan algoritma CNN, jupyter notebook digunakan sebagai tools dalam proses pengujian, dan Bahasa pemrograman yang digunakan adalah python 3.10. Training menggunakan optimizer Adam yang digunakan agar model yang dibuat lebih akurat.[24] Learning rate yang akan diterapkan pada training dataset ditentukan sebelum optimizer Adam dapat digunakan. Salah satu nilai parameter yang digunakan dalam menentukan nilai koreksi bobot selama training adalah learning rate. Selain itu, nilai untuk batch size, atau jumlah contoh pelatihan yang digunakan dalam satu iterasi, ditetapkan. Proses training akan dilakukan dengan membedakan learning rate dan batch size.

2.8 Evaluasi

Evaluasi pada penelitian ini menggunakan confusion matrix. Confusion matrix adalah alat yang digunakan dalam proses evaluasi model dalam mengklasifikasi untuk memperkirakan objek yang benar ataupun salah.[25] Data uji merupakan data test yaitu 10% dari keseluruhan data. Adapun nilai yang diukur ada empat adalah nilai True Positive (TP), nilai True Negative (TN), nilai False Positive (FP), serta nilai False Negative (FN). Pengujian suatu kelas x TP merupakan jumlah data kelas x yang teridentifikasi benar, FP merupakan jumlah data bukan kelas x



yang teridentifikasi kelas x, FN merupakan jumlah data kelas x yang teridentifikasi sebagai kelas bukan x, dan TN merupakan jumlah data kelas bukan x yang teridentifikasi benar.[26]

Tabel 3. Ilustrasi Confusion matrix

	Predicted : Yes	Predicted : No
Actual : Yes	TP	FN
Actual : No	FP	TN

Tabel 3 merangkum keempat istilah yang telah dijelaskan diatas, dan merupakan penggambaran dari confusion matrix. Macro avg atau nilai rata-rata hasil nilai dari accuracy, nilai dari precision, nilai dari recall, dan nilai dari f1-score digunakan sebagai nilai matrix untuk 5 kelas, yaitu Mangga jenis Apel, Mangga jenis Arumanis, Mangga jenis Gedong Gincu, Mangga jenis Golek, dan Mangga jenis Manalagi. Persamaan yang dipakai dalam proses menghitung nilai accuracy, nilai precision, nilai recall, dan nilai F1-score yaitu rumus 2,3,4 dan 5.

$$Accuracy(\%) = \frac{(TP+TN)}{(TP+FP+TN+FN)} \tag{2}$$

$$Precision(\%) = \frac{TP}{(FP+TP)} \tag{3}$$

$$Recall(\%) = \frac{TP}{(FN+TP)} \tag{4}$$

$$F1 - Score(\%) = \frac{(2 \times recall \times precision)}{(recall+precision)} \tag{5}$$

3. HASIL DAN PEMBAHASAN

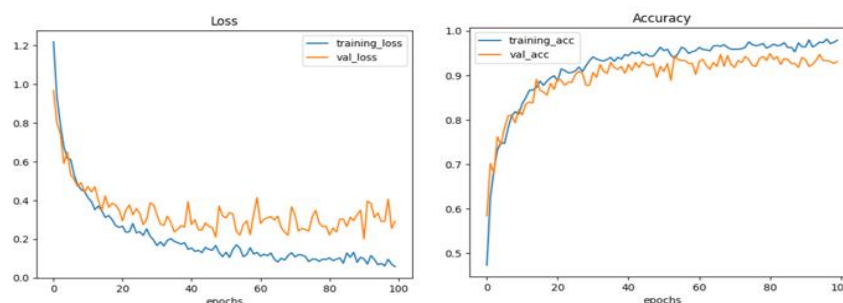
3.1 Hasil Training

Proses training menghasilkan model python. Training dilakukan dengan menggunakan metode Convolutional Neural Network (CNN) dengan tool jupyter notebook, dan Bahasa pemrograman python 3.10. Data yang dipakai untuk proses training dan pengujian adalah data hasil augmentasi yang berjumlah 2250 citra. Data yang digunakan untuk proses training adalah sebanyak 1575 citra, dan proses validasi sebanyak 450 citra dari 5 jenis mangga yang akan diklasifikasi. Training menggunakan optimizer Adam untuk mengoptimalkan nilai akurasi. Untuk mendapatkan akurasi terbaik dilakukan 4 skenario training dengan membedakan nilai learning rate dan batch size. Hyperparameter ke-1 dengan ukuran nilai learning rate adalah 0,001 dan batch size 10, hyperparameter ke-2 dengan ukuran nilai learning rate adalah 0,001 dan batch size 15, hyperparameter ke-3 dengan ukuran nilai learning rate adalah 0,0001 dan batch size 10, dan hyperparameter ke-4 dengan ukuran nilai learning rate adalah 0,0001 dan batch size 15. Semua proses training menggunakan epoch dan input size yang sama yaitu dengan epoch 100 dan input size 100x100.

Tabel 4. Hyperparameter yang digunakan

Hyperparameter Ke	Nilai Input size	Nilai Epoch	Nilai Learning rate	Nilai Batch size
1	100x100	100	0,001	10
2	100x100	100	0,001	15
3	100x100	100	0,0001	10
4	100x100	100	0,0001	15

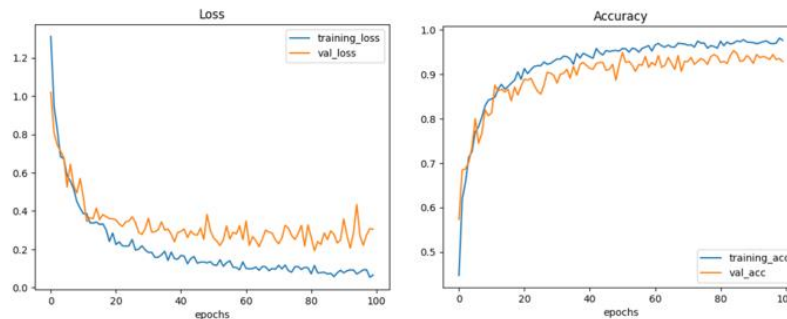
Tabel 4 merupakan rangkuman dari keempat hyperparameter yang digunakan. terdapat 4 hyperparameter yang akan dilakukan pada penelitian kali ini untuk mendapatkan model dengan nilai akurasi terbaik. Setelah proses penentuan hyperparameter telah ditentukan selanjutnya melakukan proses training. Hasil proses training dijelaskan dibawah ini :



Gambar 10. Grafik Loss dan Accuracy Hyperparameter ke-1

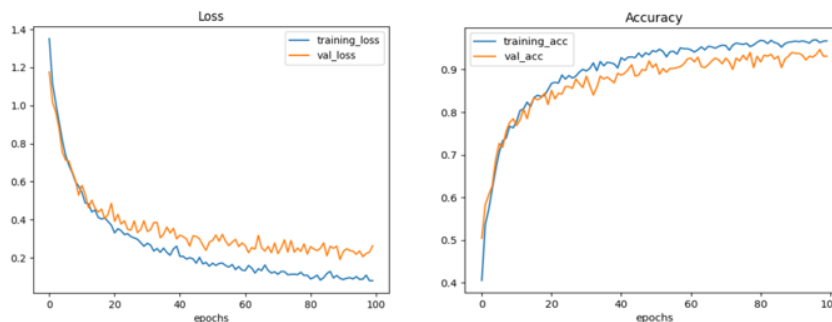


Gambar 10 menunjukkan grafik performance dari proses pengujian hyperparameter ke-1 dengan hyperparameter nilai input size adalah 100x100, nilai epoch adalah 100, nilai learning rate adalah 0,001, dan nilai batch size adalah 10. Nilai akurasi terbaik proses training ada pada epoch 81 yaitu dengan nilai Loss 0,1024, nilai accuracy 0,9651, nilai val_Loss 0,2206, dan nilai val_acc 0,94889.



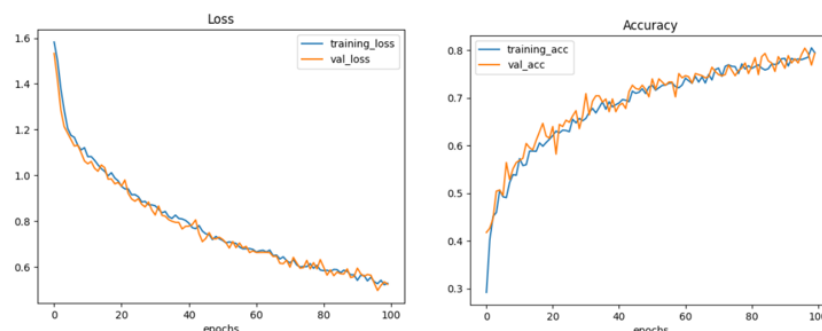
Gambar 11. Grafik Loss dan Accuracy hyperparameter ke-2

Gambar 11 menunjukkan grafik performance dari proses pengujian hyperparameter ke-2 dengan hyperparameter nilai input size adalah 100x100, nilai epoch adalah 100, nilai learning rate adalah 0,001, dan nilai batch size adalah 15. Nilai akurasi terbaik proses training ada pada epoch 85 yaitu dengan nilai Loss 0,0800, nilai accuracy 0,9702, nilai val_Loss 0,2853, dan nilai val_acc 0,9533.



Gambar 12. Grafik Loss dan Accuracy hyperparameter ke-3

Gambar 12 menunjukkan grafik performance dari proses pengujian hyperparameter ke-3 yaitu dengan hyperparameter nilai input size adalah 100x100, nilai epoch adalah 100, nilai learning rate adalah 0,0001, dan nilai batch size adalah 10. Nilai akurasi terbaik proses training ada pada epoch 98 yaitu dengan nilai Loss 0,1092, nilai accuracy 0,9632, nilai val_Loss 0,2227, dan nilai val_acc 0,94667.



Gambar 13. Grafik Loss dan Accuracy hyperparameter ke-4

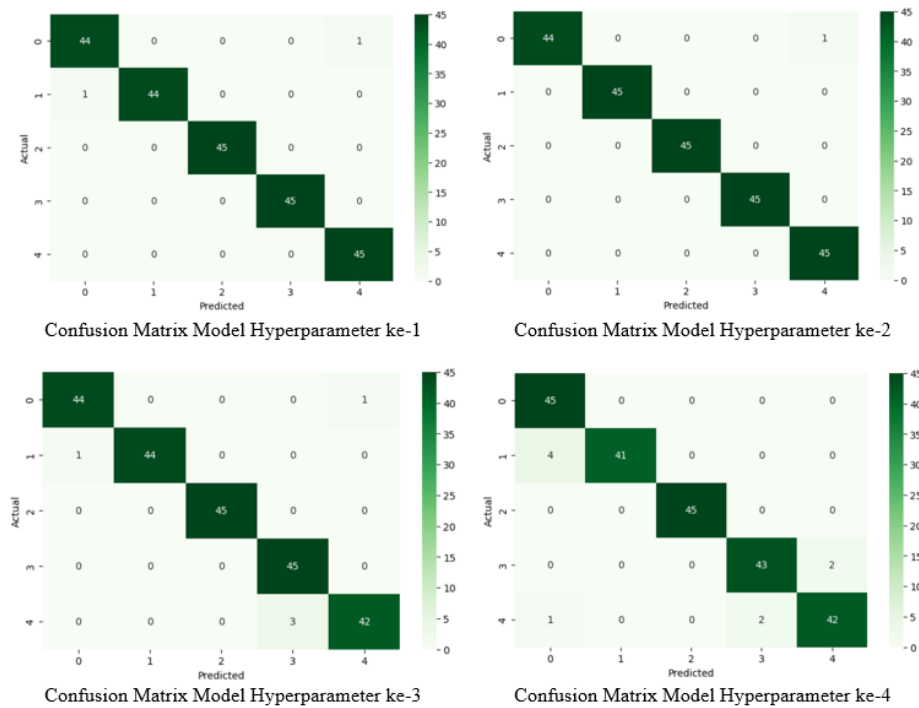
Gambar 13 menunjukkan grafik performance dari proses pengujian hyperparameter ke-4 yaitu dengan hyperparameter nilai input size adalah 100x100, nilai epoch adalah 100, nilai learning rate adalah 0,0001, dan nilai batch size adalah 15. Nilai akurasi terbaik proses training ada pada nilai epoch 100 yaitu dengan nilai Loss 0,0939, nilai accuracy 0,9689, nilai val_Loss 0,2170, dan nilai val_acc 0,9444.

3.2 Evaluasi Model

Setelah melakukan proses training dan mendapatkan model dengan akurasi terbaik, maka model tersebut akan diuji dengan data yang sudah dibagi sejak awal yaitu 10% dari keseluruhan data atau sebanyak 225 citra dengan



masing-masing 45 citra dari setiap jenis mangga. Dalam proses evaluasi menggunakan confusion matrix. Berikut hasil dari pengujian confusion matrix.



Gambar 14. Confusion matrix hasil pengujian model

Gambar 14 merupakan hasil pengujian model dari keempat hyperparameter. Dari hasil pengujian menggunakan model hyperparameter ke-1 mendapatkan nilai accuracy 99,11%, precision bernilai 99%, recall bernilai 99%, dan f1-score bernilai 99%. Pengujian dengan model hyperparameter ke-2 menghasilkan nilai accuracy 99,56%, precision bernilai 100%, recall bernilai 100%, dan f1-score bernilai 100%. Pengujian dengan model hyperparameter ke-3 menghasilkan nilai accuracy 97,78%, precision bernilai 98%, recall bernilai 98%, dan f1-score bernilai 98%. Pengujian dengan model hyperparameter ke-4 menghasilkan accuracy bernilai 96,00%, precision bernilai 96%, recall bernilai 96%, dan f1-score bernilai 96%.

Tabel 5. Hasil Proses Pengujian

Hyperparameter ke-	Input size	epoch	Learning rate	Batch size	Accuracy	Precision	Recall	F1-score
1	100x100	100	0,001	10	99,11%	99%	99%	99%
2	100x100	100	0,001	15	99,56%	100%	100%	100%
3	100x100	100	0,0001	10	97,78%	98%	98%	98%
4	100x100	100	0,0001	15	96,00%	96%	96%	96%

Tabel 5 merupakan rangkuman dari hasil pengujian confusion matrix berdasarkan 4 model yang telah ditraining. Nilai dari proses pengujian Hyperparameter yang ke-2 dengan nilai input size adalah 100x100, nilai epoch adalah 100, nilai learning rate adalah 0,001, dan nilai batch size adalah 15 memiliki akurasi terbaik dengan menghasilkan nilai accuracy 99,56%, precision bernilai 100%, recall bernilai 100%, dan f1-score bernilai 100%. Disusul dengan hyperparameter ke-1 dengan menghasilkan nilai accuracy 99,11%, precision bernilai 99%, recall bernilai 99%, dan f1-score bernilai 99%. Nilai akurasi terbaik selanjutnya adalah hyperparameter ke-3 dengan menghasilkan nilai accuracy 97,78%, precision bernilai 98%, recall bernilai 98%, dan f1-score bernilai 98%. Dan yang terakhir adalah hyperparameter ke-4 dengan menghasilkan nilai accuracy 96,00%, precision bernilai 96%, recall bernilai 96%, dan f1-score bernilai 96%.

4. KESIMPULAN

Berdasarkan proses pengujian pada penelitian yang telah dilaksanakan tentang klasifikasi jenis mangga menggunakan algoritma Convolutional Neural Network (CNN) hasilnya sama dengan tujuan dari proses penelitian ini. Tujuan dari penelitian ini yaitu mendapatkan nilai akurasi terbaik dengan membedakan hyperparameter learning rate dan batch size. Algoritma Convolutional Neural Network (CNN) dapat mengklasifikasi 5 jenis



mangga dengan baik dan menghasilkan nilai akurasi baik. Nilai learning rate adalah 0,001 dan nilai batch size adalah 15 pada hyperparameter ke-2 memiliki nilai akurasi terbaik yaitu nilai accuracy 99,56%, precision bernilai 100%, recall bernilai 100%, dan f1-score bernilai 100%. Disusul dengan hyperparameter ke-1 yang memiliki nilai accuracy 99,11%, precision bernilai 99%, recall bernilai 99%, dan f1-score bernilai 99%. Nilai akurasi terbaik selanjutnya adalah hyperparameter ke-3 dengan nilai accuracy 97,78%, precision bernilai 98%, recall bernilai 98%, dan f1-score bernilai 98%. Dan yang terakhir adalah hyperparameter ke-4 dengan nilai accuracy 96,00%, precision bernilai 96%, recall bernilai 96%, dan f1-score bernilai 96%. Seluruh percobaan memiliki nilai akurasi tinggi. Dari hasil pengujian bisa disimpulkan algoritma Convolutional Neural Network (CNN) layak menjadi salah satu model untuk klasifikasi, dan membuktikan bahwa learning rate dan batch size memang berpengaruh terhadap proses training. Akan tetapi dengan pengolahan citra yang berbeda, data set yang berbeda akan berpengaruh juga terhadap proses training. Karena pada penelitian sebelumnya mendapatkan akurasi terbaik dengan nilai learning rate dan batch size yang berbeda.

REFERENCES

- [1] H. Edha et al., "PENERAPAN METODE TRANSFORMASI RUANG WARNA HUE SATURATION INTENSITY (HSI) UNTUK MENDETEKSI KEMATANGAN BUAH MANGGA HARUM MANIS," 2020.
- [2] C. Bagus Sanjaya and dan Muhammad Imron Rosadi, "KLASIFIKASI BUAH MANGGA BERDASARKAN TINGKAT KEMATANGAN MENGGUNAKAN LEAST-SQUARES SUPPORT VECTOR MACHINE," 2018.
- [3] A. Arip Munawar and D. Suhandy, "Prediction of Vitamin C, Titratable Acidity, and Soluble Solids Content of Mango Fruits Using Near-Infrared Reflectance Spectroscopy," 2020.
- [4] D. Hidayat, "KLASIFIKASI JENIS MANGGA BERDASARKAN BENTUK DAN TEKSTUR DAUN MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN) CLASSIFICATION OF TYPES OF MANGO BASED ON LEAVE SHAPE AND TEXTURE USING CONVOLUTIONAL NEURAL NETWORK (CNN) METHOD," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 5, no. 1, 2022.
- [5] Z. E. Fitri, R. Aprilia, A. Madjid, and A. M. N. Imron, "Ensiklopedia Digital Berdasarkan Klasifikasi Varietas Buah Mangga (*Mangifera* spp.) Menggunakan Algoritma Backpropagation," *Komputika : Jurnal Sistem Komputer*, vol. 11, no. 2, pp. 113–120, Feb. 2022, doi: 10.34010/komputika.v11i2.5513.
- [6] A. Arkadia et al., *Klasifikasi Buah Mangga Badami Untuk Menentukan Tingkat Kematangan dengan Metode CNN*. 2021.
- [7] F. Sudana Putra, D. Otomatis Jerawat Wajah, and M. P. Kurniawan, "Deteksi Otomatis Jerawat Wajah Menggunakan Metode Convolutional Neural Network (CNN)," *JIFOTECH (JOURNAL OF INFORMATION TECHNOLOGY)*, vol. 1, no. 2, 2021.
- [8] O. Nurdiawan, R. Herdiana, I. Ali, and M. Fijriani, "Kinerja Algoritma Convolutional Neural Network dalam Klasifikasi Covid-19 Varian Omicron Berdasarkan Citra Ct-Scan Thoax," *Jurnal Riset Komputer*, vol. 9, no. 5, pp. 2407–389, 2022, doi: 10.30865/jurikom.v9i5.4884.
- [9] M. A. Hanin, R. Patmasari, R. Yunendah, and N. Fu'adah, "SISTEM KLASIFIKASI PENYAKIT KULIT MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) SKIN DISEASE CLASSIFICATION SYSTEM USING CONVOLUTIONAL NEURAL NETWORK (CNN)," Feb. 2021.
- [10] N. Rochmawati et al., "Analisa Learning rate dan Batch size Pada Klasifikasi Covid Menggunakan Deep learning dengan Optimizer Adam," 2021.
- [11] D. Luthfy, C. Setianingsih, and M. W. Paryasto, "Indonesian Sign Language Classification Using You Only Look Once," Feb. 2023.
- [12] A. Aziz, R. Reyhan Zhafari, and M. M. Santoni, *Klasifikasi 10 Spesies Monyet Berdasarkan Citra Menggunakan Convolutional Neural Network*. 2021.
- [13] C. Uswatun Khasanah, A. Kusuma Pertiwi, F. Witamajaya, P. Akbara Surakarta, and J. Sumbing Raya, "Implementasi Data Augmentation Random Erasing dan GridMask pada CNN untuk Klasifikasi Batik Implementation of Random Erasing and GridMask Data Augmentations on CNN for Batik Classification," vol. 13, no. 1, 2023, doi: 10.30700/jst.v13i1.1274.
- [14] I. Wulandari, H. Yasin, and T. Widiharih, "KLASIFIKASI CITRA DIGITAL BUMBU DAN REMPAH DENGAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN)," *Jurnal Gaussian*, vol. 9, no. 3, pp. 273–282, 2020, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [15] S. Sunardi, A. Yudhana, and S. A. Wijaya, "Penerapan Metode Median Filtering untuk Optimasi Deteksi Wajah pada Foto Digital," *Journal of Innovation Information Technology and Application (JINITA)*, vol. 4, no. 1, pp. 51–60, Jun. 2022, doi: 10.35970/jinita.v4i1.1214.
- [16] M. Mellyadi and P. Harliana, "Segmentasi Citra Satelit dalam Observasi dan Konservasi Hutan Lindung Taman Nasional Gunung Lauser Menggunakan Algoritma Fuzzy C-Means," *Hello World Jurnal Ilmu Komputer*, vol. 1, no. 2, pp. 90–96, Jul. 2022, doi: 10.56211/helloworld.v1i2.44.
- [17] A. Susanto and A. Kesehatan, "KOMBINASI SOBEL, CANNY DAN OTSU UNTUK SEGMENTASI CITRA PENGGUNA HELEM SAFETY DAN TANPA HELEM SAFETY," 2022.
- [18] A. Rizky Pratama, A. Ratna Juwita, and T. Al Mudzakir, "Klasifikasi Daging Sapi Berdasarkan Ciri Warna Dengan Metode Otsu dan K-Nearest Neighbor," Apr. 2021.
- [19] S. Ilahiyah and A. Nilogiri, "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network," vol. 3, no. 2, pp. 49–56, 2018.
- [20] Rifkie Primartha and Romi Satria Wahono, *Algoritma Machine Learning*. Bandung: INFORMATIKA, 2021.
- [21] Ilyas and Risnawati, *PENGENALAN TANAMAN CABAI DENGAN TEKNIK KLASIFIKASI MENGGUNAKAN METODE CNN*. 2020.



- [22] P. Adi Nugroho, I. Fenriana, and R. Arijanto, "IMPLEMENTASI DEEP LEARNING MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) PADA EKSPRESI MANUSIA," JURNAL ALGOR, vol. 2, no. 1, 2020, [Online]. Available: <https://jurnal.buddhidharma.ac.id/index.php/alogor/index>
- [23] A. Ratna Juwita et al., "IDENTIFIKASI CITRA BATIK DENGAN METODE CONVOLUTIONAL NEURAL NETWORK," 2021, vol. 6, no. 1, pp. 192–208.
- [24] E. Oktafanda, "Klasifikasi Citra Kualitas Bibit dalam Meningkatkan Produksi Kelapa Sawit Menggunakan Metode Convolutional Neural Network (CNN)," Jurnal Informatika Ekonomi Bisnis, pp. 72–77, Aug. 2022, doi: 10.37034/infec.v4i3.143.
- [25] I. Nawangsih, I. Melani, S. Fauziah, and A. I. Artikel, "PELITA TEKNOLOGI PREDIKSI PENGANGKATAN KARYAWAN DENGAN METODE ALGORITMA C5.0 (STUDI KASUS PT. MATARAM CAKRA BUANA AGUNG)," Jurnal Pelita Teknologi, vol. 16, no. 2, pp. 24–33, 2021.
- [26] U. Muhammadiyah Jember, M. Ainur Rohman, and D. Arifianto, "Penerapan Metode Euclidean Probability and Confusion Matrix dalam Diagnosa Penyakit Koi Application of the Euclidean Probability and Confusion Matrix Methods in the Diagnosis of Koi Disease," 2021. [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JST>