

Classification of Harvest-Non Harvest in Rice Plant Image Using Convolutional Neural Network Algorithm

Revina Rahmadani, Yenni Kurniawati*, Dony Permana, Dina Fitria

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: yennikurniawati@fmipa.unp.ac.id

Submitted : 11 Juni 2024

Revised : 06 Agustus 2024

Accepted : 12 Agustus 2024

ABSTRACT

Rice (*Oryza Sativa*) is a staple crop crucial to Indonesia's national food security. To ensure adequate rice production, the government, through the Central Agency of Statistics (BPS), routinely conducts the Area Sampling Frame (ASF) survey to estimate the area and timing of rice harvests. However, manual classification of rice images faces limitations in terms of time efficiency and speed, as it requires significant time for BPS at the regency and city levels to validate the input data. Therefore, this issue can be addressed by using Convolutional Neural Network (CNN) algorithms, which are expected to provide more accurate and efficient results.. CNN is a method to detect and recognize objects in digital images. The hyperparameters in this study are batch size 32, Adam optimizer, and several epoch values. The epochs used were epoch 5, epoch 25, epoch 50, epoch 75, and epoch 100. The results showed that the highest accuracy value of the training data was in epoch 100 with an accuracy value of 93%, while the accuracy of the testing data was 82,3%. Based on the accuracy table, the classification results are included in the good classification category, so it can be concluded that the model with CNN has good performance in predicting harvest-non harvest in rice plant images.

Keywords: Convolutional Neural Network (CNN), image of rice plant, The Area Sample Framework (ASF) survey.



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Tanaman padi (*Oryza sativa* L.) merupakan tanaman penghasil beras yang berperan penting dalam menjaga ketahanan pangan nasional. Oleh karena itu, untuk menjaga ketahanan pangan nasional pemerintah berupaya untuk memastikan produksi padi di Indonesia mencukupi kebutuhan konsumsi masyarakat Indonesia. Salah satu yang dilakukan adalah dengan melakukan perhitungan terhadap produksi panen nasional dengan tepat dan akurat. Informasi mengenai data ketersediaan beras nasional berdasarkan luas panen diperlukan untuk menjamin ketahanan pangan nasional yang akurat. Upaya yang dilakukan pemerintah untuk memperkirakan waktu panen dan mengestimasi luas panen adalah dengan menggunakan metode Kerangka Sampel Area (KSA).

Kegiatan Survei KSA merupakan survei rutin yang dilakukan oleh Badan Pusat Statistik (BPS) setiap bulannya. Survei KSA adalah survei berbasis area dengan cara pengamatan langsung terhadap bagian-bagian sampel yang sudah ditetapkan lokasinya. Kegiatan ini bertujuan untuk meningkatkan metode pengumpulan data konvensional menjadi lebih ilmiah, modern, dan objektif, serta menjadikan data yang dihasilkan lebih akurat dan tepat waktu dengan memanfaatkan perkembangan teknologi (BPS, 2017).



(a)



(b)

Gambar 1. (a) Fase Generatif (b) Fase Panen

Tiap bulannya petugas KSA memotret hasil amatan dengan menggunakan *handphone* berbasis android, dimana hasil potret tersebut akan diklasifikasikan secara manual oleh petugas survei dan dikirimkan ke server pusat untuk diolah. Klasifikasi manual oleh petugas KSA dapat dilihat pada Gambar 1. Pada Gambar 1 (a) merupakan fase generatif, dimana fase tersebut termasuk ke dalam kategori non panen. Fokus pada penelitian ini yaitu mengklasifikasikan fase pertumbuhan padi ke dalam kategori panen-non panen. Fase non panen merupakan periode di mana tanaman padi tidak menghasilkan gabah. Fase ini terdiri dari fase vegetatif awal, vegetatif akhir, dan generatif, sedangkan fase panen adalah periode di mana gabah sudah matang dan siap untuk dipanen.

Pengamatan dan klasifikasi fase pertumbuhan padi secara manual oleh petugas KSA memiliki keterbatasan dalam efisiensi waktu dan kecepatan, terutama dalam mengolah data citra yang banyak setiap bulannya. Oleh karena itu, proses pengklasifikasian manual oleh petugas KSA dapat diperbaiki dengan menggunakan teknologi modern seperti *Artificial Intelligence* (AI) dengan teknik *deep learning* menggunakan algoritma *Convolutional Neural Network* (CNN), yang diharapkan dapat memberikan hasil yang lebih akurat dan efisien. CNN merupakan metode *Deep Learning* yang digunakan untuk mendeteksi dan mengenali objek pada citra digital. Metode CNN dapat digunakan untuk pengenalan wajah, analisis dokumen, klasifikasi gambar, klasifikasi video, dsb. Kinerja model CNN sangat dipengaruhi oleh pemilihan *hyperparameter* yang tepat. *Hyperparameter* seperti *epoch*, *batch size*, dan *optimizer* berperan penting dalam mengatur proses pelatihan model dan mempengaruhi akurasi serta kecepatan konvergensi.

Terdapat beberapa penelitian terdahulu yang menunjukkan bahwa penyetelan *hyperparameter* yang cermat dapat meningkatkan kinerja model CNN dalam berbagai aplikasi, termasuk klasifikasi citra. Penelitian yang dilakukan oleh Nashrullah dkk., (2020) dengan mengklasifikasikan gambar pornografi menggunakan algoritma CNN dengan beberapa *hyperparameter* seperti *epoch*, *batch size*, dan *optimizer*. Hasil penelitian menunjukkan bahwa dengan menetapkan beberapa *hyperparameter* dapat meningkatkan nilai akurasi, sehingga dapat disimpulkan algoritma CNN dapat bekerja dengan sangat baik dalam mengklasifikasikan gambar pornografi. Selanjutnya penelitian yang oleh Kholilurrahman dkk., (2023). Penelitian tersebut mengklasifikasikan *image processing* pada citra warna daun padi. Hasil penelitian menunjukkan penghitungan penyakit tanaman padi dengan metode CNN memberikan akurasi yang tinggi, sehingga memberikan hasil yang memuaskan untuk dapat dijadikan dasar dalam meningkatkan kualitas bibit. Penelitian ini bertujuan untuk mengklasifikasi citra tanaman padi menjadi kategori panen dan non-panen menggunakan algoritma *Convolutional Neural Network* (CNN) sebagai pengembangan dari penelitian sebelumnya. Penelitian ini diharapkan dapat membantu pemerintah dalam peningkatan efisiensi dalam proses pengolahan data citra fase pertumbuhan padi, sehingga pengolahan data dapat dilakukan dengan lebih cepat dan efisien.

II. METODE PENELITIAN

Jenis penelitian ini menerapkan algoritma CNN dalam mengklasifikasikan fase panen-non panen pada citra tanaman padi. Klasifikasi dilakukan menggunakan data *images* yang diperoleh dari survei KSA Badan Pusat Statistik Kabupaten Padang Pariaman dari bulan November 2023 sampai dengan bulan Februari 2024. Variabel pada data yang digunakan dalam penelitian ini adalah variabel bebas dan variabel terikat. Variabel bebas berupa *image* yang diambil oleh petugas KSA, sedangkan variabel terikat yaitu berupa label KSA dengan 2 kategori utama, yaitu panen dan non panen. Penelitian ini menggunakan dataset yang terdiri dari 354 *image* fase panen dan 1063 *image* non panen, dengan total keseluruhan 1417 *image*. Berikut dibawah ini beberapa contoh *image* dalam dataset penelitian.



Panen



Panen



Panen



A. Tahapan Analisis

Penelitian ini menggunakan algoritma *Convolutional Neural Network* (CNN) yang termasuk dalam metode *deep learning*. Kelebihan metode CNN adalah dapat belajar dan mengekstraksi fitur yang kompleks dalam keadaan jaringan yang tinggi pada data citra. Analisis data dilakukan dengan menggunakan *keras package* dan *tensorflow* pada *Google Colab*. Tahapan analisis dalam penelitian ini dilakukan dengan langkah-langkah sebagai berikut.

1. Tahapan pertama yaitu *input image* fase panen-non panen.
2. Melakukan *preprocessing* data.

Preprocessing dilakukan untuk meningkatkan kualitas gambar, dengan kualitas yang baik akan mempermudah dan mempercepat kinerja algoritma dalam pengenalan gambar (Miranda dkk., 2020) Tahapan *preprocessing* ada 2, yaitu augmentasi data dan pembagian data. Tujuan augmentasi data adalah untuk meningkatkan performa model dengan memberikan lebih banyak ilustrasi tentang bagaimana data dapat digunakan dalam kehidupan nyata (Martias, 2023). Pada penelitian ini, digunakan pembagian dataset dengan perbandingan 75:25, yang artinya 75% untuk data *training* dan 25% untuk data *testing*.

3. Pembuatan Model CNN.

Proses pembuatan model CNN ini dilakukan untuk mengklasifikasikan gambar, dimana model yang dibentuk adalah model *sequential* yang terdiri dari beberapa lapisan. Dataset yang digunakan pada tahap ini adalah *image* fase panen-non panen yang telah diinputkan. Pada CNN, terdapat dua lapisan yaitu *feature learning* dan *classification*. *Feature learning* digunakan untuk mengesktraksi informasi penting dari gambar. Pada *feature learning* menampilkan tiga operasi pada sebuah data *input* yaitu *convolution layer*, *pooling layer*, dan *Rectified Linier Unit* (ReLU). Ketiga operasi tersebut dilakukan secara berulang dengan setiap lapisan mempelajari fitur yang berbeda. Pada tahap *classification* terdiri dari *flatten*, *fully connected layer*, dan fungsi aktivasi *softmax*. Berikut adalah penjelasan pada tahap *classification*.

a. Flatten

Pada tahap ini, *output* dari *feature learning* diubah menjadi vektor sehingga dapat diinput ke dalam *fully connected layer* untuk dilakukan proses klasifikasi.

b. Fully Connected Layer

Fully connected layer memiliki peranan penting yaitu untuk menganalisis dan mengklasifikasikan fitur yang diekstrak dari *input*. Lapisan ini menghasilkan *output* berupa klasifikasi citra yang diinginkan. Persamaan proses *fully connected layer* adalah sebagai berikut:

$$\hat{y}_k = w_0 + \sum_{i=0}^n x_i w_{i,k} ; k = 0,1,2, \dots, t \quad (1)$$

Keterangan :

- | | |
|---|--|
| \hat{y} : Output <i>fully connected layer</i> | x : nilai <i>input</i> dalam <i>flatten</i> |
| w_0 : bias | w : bobot |
| n : banyaknya data pada tiap unit | t : banyaknya target pada layer <i>fully connected</i> |

c. Softmax

Fungsi aktivasi *softmax* merupakan fungsi aktivasi yang digunakan untuk klasifikasi kelas biner. Pada penelitian ini terdapat 2 kategori yaitu 0 untuk fase panen dan 1 untuk non panen.

4. Pelatihan Model CNN.

Data yang digunakan pada tahap ini adalah data *training* yang digunakan untuk melatih model dan data *testing* yang digunakan untuk mengevaluasi performa model selama pelatihan, tetapi tidak digunakan untuk melatih model itu sendiri. *Hyperparameter* yang digunakan pada penelitian ini adalah beberapa nilai *epoch*, *batch size* berukuran 32, dan *optimizer* Adam (*Adaptive Moment Estimation*). Adam adalah algoritma optimasi yang

umum digunakan dan populer dalam *deep learning*. Metode ini menggunakan *moment estimation* untuk memperkirakan *mean* dan *standar deviasi* pada gradien. Hal ini menjadikan algoritma Adam efektif untuk optimasi model yang cepat dan efisien (Enriko dkk., 2023).

5. Melakukan Evaluasi Model.

Confusion matrix digunakan untuk mengevaluasi performa model agar mendapatkan gambaran tentang seberapa baik model dalam mengklasifikasikan data dengan benar. *Confusion matrix* akan membentuk tabel yang berisikan nilai aktual dan nilai prediksi yang menghasilkan nilai *accuracy*, nilai *precision*, dan *recall*. Data yang digunakan pada tahap evaluasi model adalah data *testing* yang digunakan untuk menguji kinerja model setelah pelatihan. Tabel 1 merupakan *confusion matrix* berukuran 2×2 yang dihasilkan oleh proses klasifikasi pada data dua kelas.

Tabel 1. *Confusion Matrix*

Aktual	Prediksi	
	Positif	Negatif
Positif	True Positive (TP)	False Negative (FN)
Negatif	False Positive (FP)	True Negative (TN)

(Sumber: Bramer, 2007)

Keterangan :

- TP adalah kasus dimana ketika banyaknya data aktual positif dan diprediksi sebagai data positif.
- FN adalah kasus dimana ketika banyaknya data aktual positif dan diprediksi sebagai data negatif.
- FP adalah kasus dimana ketika banyaknya data aktual negatif dan diprediksi sebagai data positif.
- TN adalah kasus dimana ketika banyaknya data aktual negatif dan diprediksi sebagai data negatif.

Informasi yang diperoleh dari Tabel 1 sebagai berikut.

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

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

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

Nilai *accuracy* yang diperoleh dapat dikategorikan ke dalam kategori yang terdapat pada Tabel 2.

Tabel 2. Tingkatan Nilai Akurasi

Nilai Akurasi	Tingkatan Klasifikasi
0,90 – 1,00	Excellent Classification
0,80 – 0,90	Good Classification
0,70 – 0,80	Fair Classification
0,60 – 0,70	Poor Classification
0,50 – 0,60	Failure

(Sumber : Gorunescu, 2011)

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini merupakan data survei KSA Badan Pusat Statistik Kabupaten Padang Pariaman dari bulan November 2023 sampai dengan bulan Februari. Data citra yang digunakan adalah *images* tanaman padi fase panen dan non panen dengan jumlah total 1417 *image*, dimana 354 data adalah kategori panen dan 1063 data

adalah kategori non panen. Selanjutnya data citra tersebut di *upload* ke dalam *google drive* dengan nama folder panen dan non panen.

B. Preprocessing

Pada tahap *preprocessing*, ada 2 tahapan yang dilakukan, yaitu augmentasi data dan pembagian dataset. Augmentasi data digunakan untuk memodifikasi data imags. Selanjutnya dilakukan pembagian dataset dengan perbandingan 75% data *training* dan 25% data *testing*. Pembagian dataset ke dalam data *training* dan *testing* terdapat pada Tabel 3.

Tabel 3. Persentase Dataset

Kategori	Jumlah Data	Training		Testing	
		Persen (%)	Pembagian Data	Persen (%)	Pembagian Data
Panen	354	75	265	25	89
Non Panen	1063	75	797	25	226

Jumlah data yang digunakan adalah 1417 *image* yang terdiri dari 354 panen dan 1063 non panen. Proporsi pembagian data *training* dan *testing* adalah 75:25 yang terdapat pada Tabel 3. Data *training* sebesar 75% pada kategori panen adalah sebanyak 265 data dan kategori non panen adalah sebanyak 797 data. Data *testing* sebesar 25% pada kategori panen adalah sebanyak 89 data dan kategori non panen adalah sebanyak 226 data.

C. Model CNN

Model CNN digunakan untuk mengklasifikasikan gambar, dimana model yang dibentuk adalah model sequential yang terdiri dari beberapa lapisan (*layers*). Pada tahapan ini menggunakan *Keras Package* pada bahasa pemograman *Python*. Model yang digunakan pada penelitian ini dapat dilihat pada Tabel 4.

Tabel 4. Model CNN

Layers	Output	Parameter
conv2d (Conv2D)	(None, 222, 222, 32)	896
max_pooling2d (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_1 (Conv2D)	(None, 109, 109, 64)	18.496
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 64)	0
conv2d_2 (Conv2D)	(None, 52, 52, 128)	73.856
max_pooling2d_2 (MaxPooling2D)	(None, 26, 26, 128)	0
flatten (Flatten)	(None, 86528)	0
dense (Dense)	(None, 200)	17.305.800
dropout (Dropout)	(None, 200)	0
dense_1 (Dense)	(None, 500)	100.500
dropout_1 (Dropout)	(None, 500)	0
dense_2 (Dense)	(None, 2)	1.002
Total params : 17.500.550		
Trainable params : 17.500.550		
Non-trainable params : 0		

Berdasarkan Tabel 4 diatas dapat dilihat jumlah parameter yang dilatih dalam model sebanyak 17.500.550 parameter. Ukuran citra atau images setiap layer konvolusi semakin berkurang ukurannya. *Output* dari tahapan konvolusi pertama menghasilkan *output* dengan ukuran $222 \times 222 \times 32$ dan memiliki 896 parameter yang dapat dilatih. Selanjutnya, dilakukan *max pooling* dengan mengurangi ukuran *output* menjadi $111 \times 111 \times 32$. Proses ini diulangi dengan lapisan konvolusi kedua dengan menggunakan 64 filter, diikuti dengan *max pooling* yang mengurangi ukuran *output* menjadi $54 \times 54 \times 64$. Lapisan konvolusi ketiga menggunakan 128 filter dan dilanjutkan dengan *max pooling* yang mengurangi ukuran *output* menjadi $26 \times 26 \times 128$. Lapisan *flatten* digunakan untuk mengubah data menjadi bentuk vektor satu dimensi dengan panjang 86528, yang dihubungkan ke lapisan *dense* pertama dengan 200 neuron dan 17.305.800 parameter yang dapat dilatih. *Dropout* diterapkan untuk mengurangi *overfitting*, diikuti oleh lapisan *dense* kedua dengan 500 neuron dan lapisan dropout kedua. Lapisan *dense* terakhir berfungsi sebagai lapisan output dengan 2 neuron. Hal ini menunjukkan bahwa model dirancang untuk klasifikasi biner, di mana output diklasifikasikan menjadi dua kategori.

D. Pelatihan Model CNN

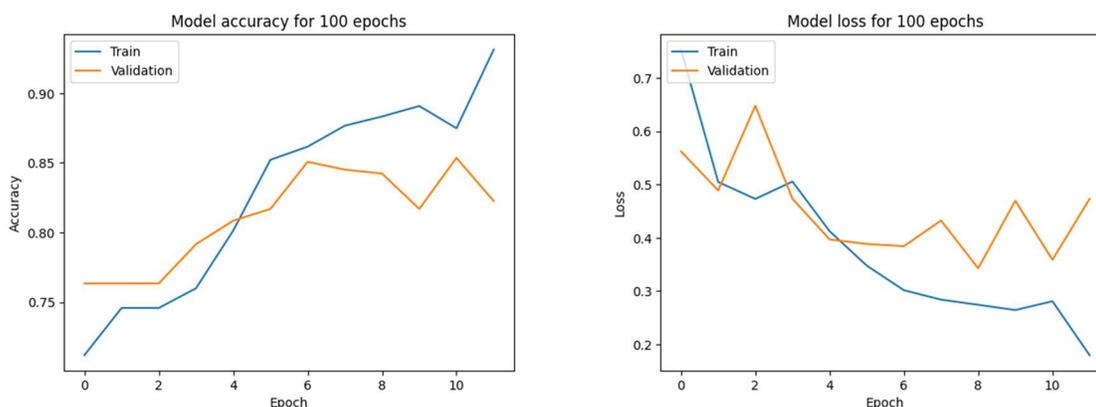
Pada pelatihan model CNN dilakukan dengan membandingkan nilai *epoch*, besar kecilnya nilai *epoch* yang ditentukan mempengaruhi proses pembelajaran dan akan berhenti tepat pada nilai *epoch* yang ditentukan yaitu dengan akurasi diatas 90%. Berdasarkan pada hasil *training* yang telah dilakukan maka diperoleh hasil perbandingan nilai *accuracy* dan *loss accuracy* sebagai berikut:

Tabel 5. Perbandingan Nilai *Epoch* Menggunakan *Optimizer Adam*

<i>Epoch</i>	Data Training		Data Testing	
	<i>Accuracy</i>	<i>Loss</i>	<i>Val Acc</i>	<i>Val Loss</i>
5	0,864	0,325	0,800	0,469
25	0,924	0,187	0,870	0,325
50	0,906	0,216	0,845	0,372
75	0,908	0,214	0,873	0,347
100	0,930	0,166	0,822	0,473

Berdasarkan Tabel 5 dapat dilihat bahwa semakin besar nilai *epoch* yang digunakan, maka nilai akurasi data *training* juga semakin baik. Namun, nilai *epoch* yang terlalu besar dapat menyebabkan perlambatan pada proses *training*. Berdasarkan tabel tersebut dapat dilihat akurasi tertinggi yaitu terdapat pada *epoch* 100 dengan nilai akurasi data *training* sebesar 93% dengan nilai *loss* sebesar 16,6%.

Setelah mendapatkan nilai *epoch* terbaik, selanjutnya melihat grafik akurasi dan *loss* dari nilai *training* dan *testing*. Gambar dibawah ini merupakan grafik yang dihasilkan oleh *optimizer Adam* dengan nilai *epoch* 100.



Gambar 2. (a) Grafik *Accuracy*, (b) Grafik *Loss*

Berdasarkan Gambar 2 (a), terdapat peningkatan yang signifikan pada grafik *training accuracy* maupun grafik *validation accuracy*. Hal ini menunjukkan bahwa model berhasil belajar dari data *training* dan mampu menggeneralisasi pada data yang belum pernah dilihat sebelumnya. Namun, terdapat indikasi potensi *overfitting* pada akhir pelatihan dikarenakan grafik *validation accuracy* mulai melambat dan bahkan sedikit menurun.

Pada Gambar 2 (b), grafik *training loss* menurun secara bertahap, mengindikasikan bahwa model semakin baik dalam memprediksi data yang digunakan untuk pelatihan. Namun, pada grafik *validation loss* terlihat adanya fluktuasi dan ketidakstabilan. Meskipun terjadi penurunan pada awal pelatihan, *validation loss* tidak menurun secara konsisten dan bahkan cenderung meningkat pada beberapa titik. Hal ini menunjukkan terjadinya *overfitting*, di mana model terlalu mempelajari data pelatihan sehingga kurang mampu untuk menggeneralisasi data baru yang belum pernah dilihat sebelumnya.

E. Evaluasi Model

Evaluasi model dilakukan pada *epoch* 100 dengan menggunakan hasil model dalam proses klasifikasi yang disajikan dalam bentuk *confussion matrix*. Tabel 6 menyajikan *confussion matrix* dari model CNN.

Tabel 6. Nilai *Confusion Matrix*

Aktual	Prediksi	
	Panen	Non Panen
Panen	62	22
Non Panen	41	230

Tabel *confusion matrix* diatas menunjukkan hasil evaluasi kinerja model klasifikasi untuk memprediksi panen padi. Kinerja model klasifikasi yang dibangun dapat dievaluasi dengan menghitung nilai *accuracy*, *precision*, dan *recall* menggunakan persamaan (2), (3), dan (4). Hasil evaluasi model yang terbentuk disajikan pada Tabel 7 di bawah ini.

Tabel 7. Hasil Evaluasi Model

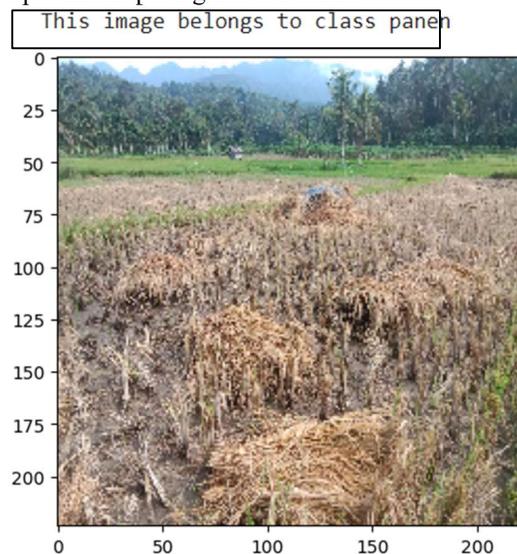
Evaluasi Model	Nilai
<i>Accuracy</i>	82,3%
<i>Precision</i>	60,2%
<i>Recall</i>	73,8%

Tabel 7 memberikan informasi mengenai nilai *accuracy*, *precision*, dan *recall* dari data yang diuji. Didapatkan nilai *accuracy* yaitu sebesar 82,3%, yang artinya model mampu mengklasifikasikan panen-non panen pada citra tanaman padi sebesar 82,3%. *Precision* sebesar 60,2%, artinya dari semua data yang diprediksi sebagai panen, 60,2% adalah prediksi yang benar. *Recall* sebesar 73,8%, artinya algoritma CNN mampu memprediksi 73,8% dari total data aktual panen dengan benar.

Klasifikasi dengan algoritma CNN menghasilkan nilai *accuracy* sebesar 82,3%. Berdasarkan tabel *accuracy* yang terdapat pada Tabel 2, hasil klasifikasi termasuk ke dalam kategori *good classification*, sehingga dapat disimpulkan model dengan CNN memiliki performa yang baik dalam memprediksi panen-non panen pada citra tanaman padi.

F. Klasifikasi Citra

Pada penelitian ini dilakukan klasifikasi citra untuk mengetahui citra padi tersebut termasuk ke dalam fase panen atau fase non panen. Hal tersebut dapat dilihat pada gambar berikut.



Gambar 3. Klasifikasi Citra

Dari Gambar 3 diatas, dapat dilihat gambar dengan kategori panen yaitu gambar yang diambil dari data panen padi. Citra padi tersebut di prediksi masuk ke dalam fase panen. Hal tersebut terbukti benar.

IV. KESIMPULAN

Berdasarkan analisis yang telah dilakukan, banyaknya layer konvolusi yang digunakan yaitu sebanyak 3 layer dan 3 *pooling*, fungsi aktivasi yang digunakan yaitu ReLU dan *softmax* dengan beberapa *hyperparameter* lainnya seperti *batch size*, *optimizer* Adam, dan nilai *epoch*. Pada data *training* didapatkan nilai *accuracy* sebesar 93% pada nilai *epoch* 100, sedangkan nilai *accuracy* pada data *testing* yang didapatkan dari model yang terbentuk yaitu sebesar 82,3%, sehingga dapat disimpulkan model dengan CNN memiliki performa yang baik dalam memprediksi panen-non panen pada citra tanaman padi. Selain itu, hasil klasifikasi yang terbentuk menyatakan bahwa citra panen yang di *inputkan* dapat diklasifikasikan dengan benar. Saran untuk penelitian selanjutnya adalah diharapkan menambahkan lebih banyak *images* fase panen agar dapat memperoleh tingkat akurasi yang lebih akurat. Selain itu penelitian selanjutnya diharapkan melakukan metode penyeimbangan data agar memperoleh nilai akurasi yang lebih baik.

UCAPAN TERIMA KASIH

Terima kasih kepada Badan Pusat Statistik (BPS) Kabupaten Padang Pariaman yang telah bersedia memberikan data *images* KSA yang digunakan pada penelitian ini.

DAFTAR PUSTAKA

- Abu, M. A., Indra, N. H., Rahman, A. H. A., Sapiee, N. A., & Ahmad, I. (2019). A study on image classification based on deep learning and tensorflow. *International Journal of Engineering Research and Technology*, 12(4), 563–569.
- Azmi, Y., Yulistiyono, A., Karyasa, T. B., Putra, R. P., Salama, S. H., Thamrin, N. T., Septiadi, D., Dinata, G. F., Jumiyati, S., & Rizki, F. H. (2022). *Pertanian Terpadu* (S. F. apt. Wafi Nisrin Ramadhani & A. K. Salsabila Syafni Aulia (ed.)). PT. GLOBAL EKSEKUTIF TEKNOLOGI Anggota IKAPI No. 033/SBA/2022.
- BPS. (2017). Pedoman Teknis Pendataan Statistik Pertanian Tanaman Pangan Terintegrasi di Pulau Jawa dengan Metode Kerangka Sampel Area (KSA): KSA2017-Teknis. Badan Pusat Statistik. Jakarta.
- Bramer, M. (2007). *Principle of Data Mining*. London: Springer.
- Enriko, I. K. A., Nizar Gustiyana, F., & Putra, R. H. (2023). Komparasi Hasil Optimasi Pada Prediksi Harga Saham PT. Telkom Indonesia Menggunakan Algoritma Long Short Term Memory. *Jurnal Media Informatika Budidarma*, 7(2), 659–667. <https://doi.org/10.30865/mib.v7i2.5822>
- Gorunescu, F. (2011). *Data Mining: Konsep, model dan teknik* (Vol.12). Sains & Media Bisnis Springer.
- Hawari, F. H., Fadillah, F., Alviandi, M. R., & Arifin, T. (2022). Klasifikasi Penyakit Tanaman Padi Menggunakan Algoritma CNN (Convolutional Neural Network). *Jurnal Responsif: Riset Sains Dan Informatika*, 4(2), 184-189.
- Kholilurrahman, M., Syafei, W. A., & Nurhayati, O. D. (2023). Klasifikasi Image Processing Pada Citra Warna Daun Padi Menggunakan Metode Convolutional Neural Network. *Jurnal Ilmiah Sains*, 23(2), 175–186. <https://doi.org/10.35799/jis.v23i2.50415>
- Martias, M. F. (2023). KLASIFIKASI CITRA DAGING SAPI DAN DAGING BABI MENGGUNAKAN CNN ARSITEKTUR EFFICIENTNET-B6 DAN AUGMENTASI DATA. 2(1), 41–49.
- Miranda, N. D., Novamizanti, L., Rizal, S., Elektro, F. T., & Telkom, U. (2020). Convolutional Neural Network Pada Klasifikasi Sidik Jari Menggunakan Resnet-50 Classification of Fingerprint Pattern Using Convolutional Neural Network in Clahe Image. *Jurnal Teknik Informatika (JUTIF)*, 1(2), 61–68.
- Nurolan, A. (2019). Deteksi Dan Klasifikasi Jenis Kendaraan Berbasis Pengolahan Citra Dengan Metode Convolutional Neural Network (Cnn).
- Rahmadhani, U. S., & Marpaung, N. L. (2023). Klasifikasi Jamur Berdasarkan Genus Dengan Menggunakan Metode CNN. *Jurnal Informatika: Jurnal Pengembangan IT*, 8(2), 169–173. <https://doi.org/10.30591/jpit.v8i2.5229>

- Ruslan, K. (2019). *Memperbaiki Data Pangan Indonesia Lewat Metode Kerangka Sampel Area*. In Center for Indonesian Policy Studies.
- Sirait, T., Ekaria, Prasetyo, A., Mariyah, S., Oktora, S. I., & Yuniarto, B. (2019). Pemodelan Berat Hasil Ubinan, Pola Luas Tanam, dan Ketepatan Klasifikasi Fase Tumbuh Padi dengan Machine Learning Timbang.