

# Klasifikasi Fase Pertumbuhan Tanaman Padi Berdasarkan Data Citra Multitemporal Satelit Landsat-8 Dengan Metode *Convolutional Neural Network* (Studi Kasus Sampel Survei KSA Kabupaten Poso, Provinsi Sulawesi Tengah)

Syafniya Zilfah Aniesiy dan Kartika Fithriasari

Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember

Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia

*e-mail:* kartika\_f@statistika.its.ac.id

**Abstrak**— Survei KSA hasil kolaborasi antara BPS dengan BPPT pada tahun 2018 bertujuan untuk mengestimasi luas panen padi dengan mengamati fase pertumbuhan padi yang dilakukan dengan pengambilan sampel berukuran 300 meter  $\times$  300 meter yang diambil secara acak. Namun Survei KSA tersebut cukup banyak mengeluarkan biaya dan sumber daya lainnya. Berbeda apabila menggunakan citra satelit dimana citra satelit sudah populer digunakan untuk memprediksi tanaman oleh komunitas penginderaan jauh karena memiliki banyak informasi vegetasi dengan cepat, global, dan gratis. Citra satelit yang digunakan adalah satelit Landsat-8 dengan akses gratis. Tidak seperti Survei KSA yang melakukan pengambilan sampel secara langsung, penelitian ini menggunakan metode *CNN* yang lebih terukur, murah dan *real time* dengan data penginderaan jauh yang tersedia untuk umum. Penelitian ini fokus klasifikasi fase pertumbuhan padi dilakukan di Kabupaten Poso, Provinsi Sulawesi Tengah, Indonesia sebagai studi kasus. Citra satelit yang digunakan adalah satelit Landsat-8. Untuk tahapan yang dilakukan pada penelitian ini dimulai dari pengambilan data dari citra satelit Landsat-8, melakukan *pre-processing*, kemudian membagi data menjadi *training* dan *testing*, dilanjutkan dengan penerapan metode *CNN* dan dilakukan evaluasi model menggunakan *sensitivity*, *specificity*, *accuracy*, dan *Cohen Kappa Index* dan diperoleh akurasi sebesar 0,2647, *sensitivity* sebesar 0,2488, *specificity* sebesar 0,2485 dan *cohen kappa index* sebesar 0,0648 sedangkan pada *CNN* dengan data hasil *pre-processing brightness* didapatkan akurasi 0,2687, *sensitivity* sebesar 0,25, *specificity* sebesar 0,2499 dan *cohen kappa index* sebesar 0,0673. Padi yang digunakan pada sampel KSA cenderung lebih banyak pada fase panen dan cenderung sedikit pada fase persiapan lahan. Prediksi panen tanaman padi gagal dalam memprediksi panen tanaman padi karena hanya dapat memprediksi secara tepat pada 30 titik amatan dengan rincian 8 titik prediksi fase vegetatif awal, 2 titik prediksi fase vegetatif akhir, 8 titik prediksi fase generatif, 11 titik prediksi fase panen dan 1 titik prediksi fase persiapan lahan dari total 112 titik amatan.

**Kata Kunci**— *CNN*, KSA, Landsat-8, Padi

## I. PENDAHULUAN

Badan Pusat Statistik (BPS) bekerjasama dengan Badan Pengkajian dan Penerapan Teknologi (BPPT) serta instansi terkait mengembangkan metode Survei Kerangka Sampel Area (KSA) pada tahun 2018 bertujuan untuk mengetahui estimasi luas panen padi dengan mengamati fase pertumbuhan padi [26]. Survei KSA adalah survei dengan pengambilan sampel pada area berukuran 300 meter  $\times$  300 meter yang diambil secara acak dimana tahap selanjutnya dilakukan pengamatan fase pertumbuhan padi setiap bulan [34]. Metode KSA menunjukkan hasil bahwa perbedaan MSE (*Mean Square Error*) dan optimasi biaya yang dilakukan pada seluruh desain pengambilan sampel menunjukkan hasil yang optimum dengan memberikan tingkat akurasi dan efisiensi yang lebih baik [41]. Melihat dari hasil Survei KSA pada tahun 2019, luas panen padi yang diperkirakan sebesar 10,68 juta hektar mengalami penurunan sebanyak 700,05 ribu hektar (6,15 persen) dibandingkan luas panen padi pada tahun 2018.

Kelemahan pelaksanaan Survei KSA berada pada biaya yang cukup tinggi untuk melaksanakan Survei KSA karena peningkatan level estimasi memerlukan penambahan sampel yang berakibat pada penambahan sumber daya manusia dan sumber daya lainnya. Selain itu, informasi variasi individu non-sampel yang kurang, mengingat sampel terpilih menjadi dasar dari estimasi KSA [36]. Pemanfaatan data penginderaan jauh dengan menggunakan Landsat-8 sering digunakan untuk penelitian yang berhubungan dengan vegetatif di permukaan bumi yang mudah diakses dan tidak berbayar serta bisa digunakan untuk pemantauan fase pertumbuhan padi [30]. Citra satelit sudah populer digunakan untuk memprediksi tanaman oleh komunitas penginderaan jauh karena memiliki banyak informasi vegetasi dengan cepat, global, dan gratis.

*CNN* merupakan metode yang melakukan klasifikasi dengan objek gambar (*image*). Dalam penelitian ini fokus klasifikasi fase pertumbuhan padi dilakukan di Kabupaten Poso, Provinsi Sulawesi Tengah, Indonesia sebagai studi kasus. Berbeda dengan penelitian sebelumnya, penelitian ini dilakukan dengan metode *CNN* yang telah dilakukan modifikasi pada *pre-processing* data *image* yaitu dengan mengubah

*brightness* dan *contrast* dan metode *CNN* dengan arsitektur lainnya yaitu *CNN* dengan melakukan perubahan arsitektur pada layer-nya.

Perlu suatu cara untuk memaksimalkan pemanfaatan *big data* dari penginderaan jarak jauh (*remote sensing*) data multitemporal satelit Landsat-8 dengan metode *Convolutional Neural Network* dengan harapan hasil akurasi mencapai 90 persen. Peneliti menentukan model terbaik menggunakan *sensitivity*, *accuracy*, dan *Cohen Kappa Index* dalam perolehan keakuratan dan kebaikan klasifikasi pengamatan fase pertumbuhan padi pada Survei KSA dengan data multitemporal menggunakan citra satelit Landsat-8.

## II. TINJAUAN PUSTAKA

### 2.1 Hasil Penelitian Terdahulu

Metode *Convolutional Neural Network* (*CNN*) telah dilakukan untuk analisis. Berikut merupakan beberapa hasil penelitian terdahulu yang menggunakan metode *CNN* dan beberapa penelitian dengan topik pengklasifikasian lahan dengan metode berbeda seperti yang terlihat pada Tabel 1.

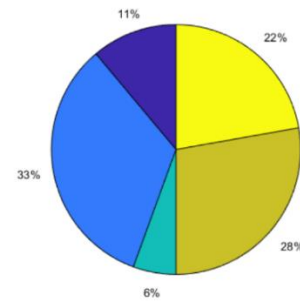
**Tabel 1.**

Hasil Penelitian Terdahulu

| Tahun | Penulis          | Bahasan  |
|-------|------------------|--|
| 2019  | Marsuhandi, dkk  | Pemanfaatan penginderaan jauh untuk pengklasifikasian jenis lahan padi menggunakan <i>Random Forest</i> dan <i>Boosting</i> diperoleh <i>Random Forest</i> dengan nilai akurasi tinggi 76.52%            |
| 2019  | Triscowati, dkk  | Klasifikasi fase pertumbuhan padi data multitemporal citra Landsat-8 dengan menggunakan <i>Random Forest</i> diperoleh akurasi 0.7   |
| 2021  | Magdalena, dkk   | Diperoleh akurasi sebesar 95,45%, loss 0,2457, serta rata-rata dari masing-masing nilai precision, recall dan f1-score sebesar 0,92 dengan metode <i>CNN</i> dalam klasifikasikan 5 jenis tutupan lahan. |
| 2021  | Miranda & Aryuni | Metode <i>CNN</i> diperoleh hasil klasifikasi dengan akurasi sebesar 98.4% dalam klasifikasi tutupan lahan dengan data Citra Satelit Sentinel-2  |

### 2.2 Statistika Deskriptif

Statistika deskriptif merupakan bagian dari statistik yang mempelajari tentang tata cara penyajian data dimana pada penyajian data ini dilakukan dengan analisa data dan interpretasi data yang bertujuan membantu pengambilan keputusan yang efektif sehingga mudah untuk dipahami. [38]. Penyajian data dapat dilakukan dengan visualisasi data menjadi berbagai diagram seperti *box-plot*, *pie chart*, *line chart* dan lain sebagainya. Gambaran mengenai visualisasi data terlihat pada Gambar 2.



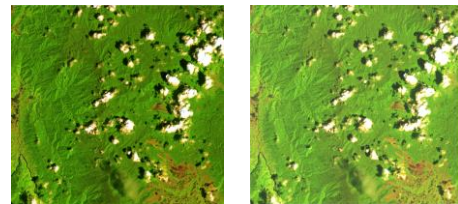
**Gambar 2.** Gambaran Visualisasi Data (Sumber: Mathworks, 2017)

### 2.3 Pre-Processing

*Pre-processing* adalah tahapan untuk mengubah data mentah (*raw data*) menjadi data atau informasi yang lebih bersih dan kedepannya dapat digunakan untuk pengolahan data yang lebih baik lagi. Penelitian ini data yang digunakan adalah data dari citra satelit Landsat-8 yang melakukan pengambilan citra secara terus menerus dan menghasilkan data citra yang kurang akurat. Maka diperlukan *pre-processing* data untuk mendapatkan hasil data yang lebih baik lagi. *Pre-processing* yang dilakukan penulis untuk mendapatkan data yang lebih baik sebagai berikut:

#### 1. Kecerahan (*Brightness*)

Perbaikan pada kecerahan citra dapat dilakukan dengan melakukan penambahan atau pengurangan sebuah konstanta dari setiap *pixel* dalam suatu citra. Nilai *pixel* hasil dari pengubahan dapat lebih kecil atau sama dengan derajat keabuan minimum (0) atau lebih besar sama dengan derajat keabuan maksimum. [29]. Gambar 3 memperlihatkan sebelum dan sesudah operasi penambahan kecerahan. Hasil dari penambahan atau pengurangan kecerahan akan diambil yang terbaik dan selanjutnya digunakan sebagai data *image* dalam pemodelan *CNN*.

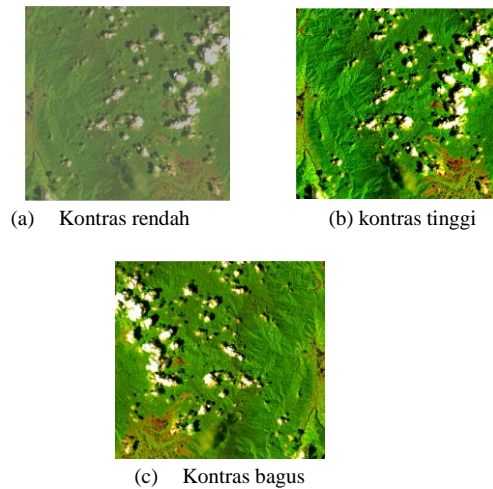


**Gambar 3.** Gambar padi sebelum dan sesudah penambahan kecerahan

#### 2. Peregangan Kontras (*Contrast Stretching*)

Kontras menyatakan sebaran terang dan gelap di dalam sebuah gambar. Citra dapat dikelompokkan dalam tiga kategori kontras. Citra kontras-rendah dapat diperbaiki dengan operasi perenggangan kontras dimana pada operasi ini nilai keabuan *pixel* akan direntangkan dari 0 hingga 255 (pada citra 8-bit) [29].

Selain citra kontras-rendah, terdapat citra kontras-tinggi dimana memiliki nilai jangkauan keabuan yang lebar tetapi terdapat area lebar yang didominasi oleh warna gelap begitupun sebaliknya. Citra kontras-bagus ketika memperlihatkan jangkauan nilai keabuan yang lebar tanpa ada nilai keabuan yang mendominasi [29]. Gambar 2.3 memperlihatkan perbedaan ketiga kategori kontras.



**Gambar 3.** Gambar padi dengan tiga macam kontras

#### 2.4 K-Fold Cross Validation

Sebelum melakukan analisis prediksi dengan algoritma *machine learning*, maka diperlukan metode untuk mengukur performa dari sebuah model. *K-Fold Cross Validation* adalah metode yang digunakan untuk menilai keakuratan dari sebuah model yang dibangun berdasarkan dataset tertentu. Nilai akurasi hasil eksperimen dari seluruh *iterasi* didapatkan dengan mengambil nilai rata-rata [40].

Dengan performa kinerja dapat dihitung dengan menggunakan rumus sebagai berikut.

$$CVA = \frac{1}{K} \sum_{j=1}^K X_j \quad (2.2)$$

dimana,

$K$  = Banyak *fold* yang digunakan

$X$  = Kinerja model (*accuracy, specificity, sensitivity, Cohen Kappa Index*)

KSA adalah metode pengambilan sampel random, maka *k-fold cross validation method* adalah metode yang baik digunakan untuk mengurangi bias yang berhubungan dengan pengambilan sampel yang random [17].

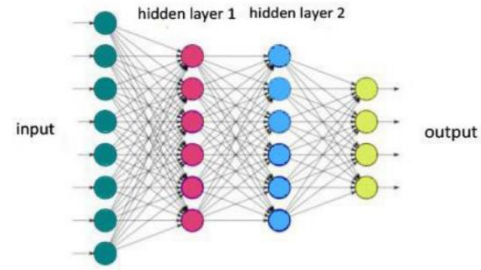
#### 2.5 Machine Learning

*Machine learning* merupakan bagian dari kecerdasan buatan yang bertujuan untuk memahami atau mengenali struktur suatu data dan mengkonversi data tersebut ke dalam suatu model. *Machine learning* dapat didefinisikan sebagai suatu algoritma yang dapat membuat sistem menjadi cerdas dengan mempelajari data-data yang tersedia dimana algoritma atau program tersebut tidak didefinisikan secara eksplisit [35].

#### 2.6 Deep Learning

*Deep Learning* merupakan salah satu kecerdasan buatan dengan cara meniru cara kerja otak manusia dalam memproses data dan menciptakan pola yang kemudian digunakan dalam pengambilan keputusan.. *Convolutional Neural Network (CNN)*, *Multilayer Perceptrons (MLP)*, dan *Recurrent Neural Network (RNN)*. Pada umumnya *deep learning* memiliki lebih

dari 3 layer yaitu *input layer*,  $N$  *hidden layer* dan *output layer* [9]. Gambaran mengenai arsitektur *deep learning* terlihat pada Gambar 4.

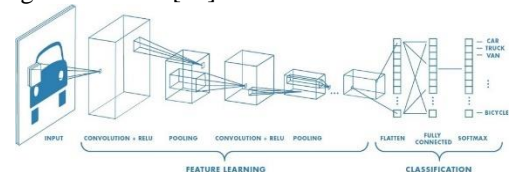


**Gambar 4.** Arsitektur *Deep Learning* (Sumber: Buduma, 2017)

Keunggulan dari *deep learning* ada pada performanya yang baik dalam menyelesaikan masalah kompleks, mengurangi kebutuhan untuk rekayasa ciri, dan memiliki arsitektur yang mampu beradaptasi terhadap permasalahan baru dengan mudah. Sedangkan untuk kekurangan dari *deep learning* adalah memerlukan banyak data, proses *training* yang memakan waktu, dan *overfitting* [2][8].

#### 2.7 Convolutional Neural Network

*Convolutional Neural Network* adalah pengembangan *Multi Layer Perceptron (MLP)* yang digunakan untuk mengolah data dua dimensi. *CNN* merupakan bagian dari *Deep Learning* karena banyak digunakan pada implementasi data citra dan memiliki kedalaman jaringan yang cukup baik. Metode *CNN* dapat digunakan dalam pengenalan wajah, analisis dokumen, klasifikasi gambar, klasifikasi video, dan lain sebagainya. Metode *CNN* digunakan untuk mengklasifikasi gambar dengan memproses gambar yang diinput kemudian diklasifikasikan pada kategori tertentu.. [27].



**Gambar 5.** Arsitektur *Convolutional Neural Network* (Sumber: Google)

Terlihat pada Gambar 5 bahwa arsitektur *CNN* secara umum terdiri dari dua tipe lapisan, yaitu :

##### 1. Feature Learning

Lapisan ini berada pada awal arsitektur tersusunnya *CNN* terdiri dari beberapa lapisan dan setiap lapisannya tersusun atas *neuron* yang terkoneksi dengan *local region* dari lapisan sebelumnya. Pada lapisan ini berfungsi untuk menerima *input* gambar secara langsung dan memprosesnya hingga menghasilkan *output* berupa vektor yang akan diolah pada lapisan berikutnya.

##### 2. Classification Layer

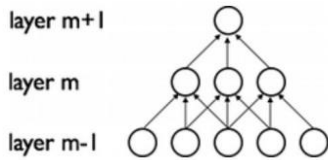
Pada layer ini tersusun dari beberapa lapisan dan setiap lapisan tersusun atas *neuron* yang terkoneksi secara penuh (*fully connected*) dengan lapisan lainnya. Pada lapisan ini menerima *input* dari hasil *output* layer bagian *feature learning* berupa vektor kemudian diproses pada *flatten*.

Layer pada *Convolutional Neural Network* terdapat 3 jenis

layer, yaitu:

### 1. Convolutional Layer

*Convolutional Layer* menggunakan korelasi lokal dengan memaksakan pola konektivitas lokal di antara layer yang berdekatan pada setiap *convolutional layer*. Layer yang lebih atas  $m$  telah didapati dari sebuah subhimpunan satuan dari layer yang lebih rendah. Dibandingkan dengan MLP, keuntungan dari *convolutional layer* adalah jumlah parameter yang berkurang signifikan yang disebabkan oleh parameter bagi [24].



Gambar 6. Jaringan dengan 3 convolution layer (Sumber: Ma,2015)

Pada layer ini dilakukan operasi konvolusi pada *output* dari layer sebelumnya. Layer tersebut adalah proses utama yang menjadi dasar dari sebuah CNN.. Setiap *kernel* dari *convolutional layer* memiliki *weight* sebesar ukuran dari *kernel* tersebut yang kemudian akan dilakukan perubahan dengan menggunakan *backpropagation* [12].

*Convolutional Layer* merupakan perkalian antara data *input* dengan *kernel* dimana *output* dari perkalian tersebut disebut dengan *feature map*. Setiap satu *convolutional layer* dapat memiliki lebih dari satu jenis *kernel* dimana masing-masing *kernel* memiliki *weight* yang berbeda dan mengekstraksi fitur yang berbeda pula. Pada pengolahan citra, *convolutional layer* menggerakkan suatu *kernel W* berukuran  $c \times d$  pada citra  $X$  berukuran  $(m + c - 1) \times (n + d - 1)$  yang dinyatakan pada persamaan sebagai berikut [18].

$$F_{m,n} = b + \sum_{c=1}^C \sum_{d=1}^D W_{c,d} X_{(rgb),m+c-1,n+d-1} \quad (2.3)$$

dimana

$F_{m,n}$  = *feature map*

$b$  = bias pada *feature map*

$W_{c,d}$  = bobot pada *convolutional kernel*

$X$  = Matriks *input*

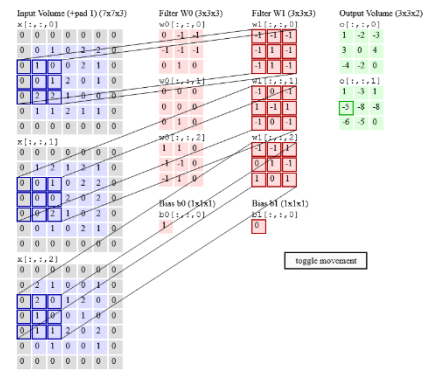
$m = 1, 2, \dots, M$ .  $M$  merupakan panjang *pixel* pada *feature map*

$n = 1, 2, \dots, N$ .  $N$  merupakan lebar *pixel* pada *feature map*

$c = 1, 2, \dots, C$ .  $C$  adalah panjang *convolutional kernel*

$d = 1, 2, \dots, D$ .  $D$  adalah lebar *convolutional kernel*

Seluruh *kernel* ini akan membentuk citra baru (citra hasil perkalian *dot-product*) dengan kedalaman *channel* sesuai dengan jumlah *kernel* [12]. Layer ini menerima input dengan ukuran 2 dimensi spasial (matriks). Gambar 7 merupakan gambaran cara kerja filter pada *convolutional layer*.



Gambar 7. Filter dari convolution layer (Sumber: CS231n, 2021)

Dimensi dari *convolutional layer* dengan input  $n_H^{[l-1]} \times n_w^{[l-1]} \times n_c^{[l-1]}$  diperoleh *output* dimensi  $n_H^{[l]} \times n_w^{[l]} \times n_c^{[l]}$  dengan perhitungan  $n_H^{[l]}$  sebagai berikut.

$$n_H^{[l]} = \left\lceil \frac{n_H^{[l-1]} + 2p - f}{s} + 1 \right\rceil \quad (2.4)$$

dimana,

$f$  = ukuran filter

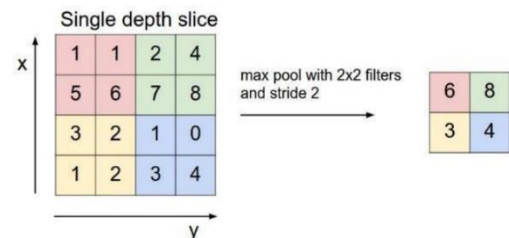
$p$  = *padding*

$s$  = *stride*

### 2. Pooling Layer

Pada layer ini dilakukan pengurangan dimensi pada setiap *activation map* tetapi informasi terpenting tetap dipertahankan. Citra input dibagi menjadi kumpulan persegi yang tidak saling tumpang tindih [39]. Jenis *pooling layer* yang umumnya digunakan adalah *max pooling*.

Cara kerja *max pooling* adalah dengan membagi *output* hasil *convolutional layer* menjadi beberapa *grid* kecil yang kemudian diambil nilai maksimal dari tiap *grid* tersebut untuk disusun matriks citra yang telah direduksi [14].



Gambar 8. Ilustrasi max pooling layer (Ilahiyah & Nilogiri, 2018)

Perhitungan dimensi dari *Pooling Layer* dengan input  $n_H^{[l-1]} \times n_w^{[l-1]} \times n_c^{[l-1]}$  diperoleh *output* dimensi  $n_H^{[l]} \times n_w^{[l]} \times n_c^{[l]}$  dengan perhitungan  $n_H^{[l]}$  seperti *convolutional layer*, namun karena *max pooling* tidak memiliki *padding* atau *padding* bernilai 0 maka perhitungan dimensi dari *max pooling layer* adalah sebagai berikut.

$$n_H^{[l]} = \left\lceil \frac{n_H^{[l-1]} - f}{s} + 1 \right\rceil \quad (2.5)$$

dimana,

$f$  = ukuran filter

$p$  = *padding*

$s$  = *stride*

### 3. Fully-Connected Layer

Pada *layer* ini merupakan tradisional *Multi Layer Perceptron (MLP)* yang menggunakan fungsi aktivasi *softmax*. *Layer* ini berfungsi untuk menghitung nilai setiap kelas. *Neuron* pada *layer* ini memiliki koneksi penuh terhadap seluruh aktivasi *layer* sebelumnya. Aktivasinya dapat dihitung dengan perkalian matriks dengan sebuah *offset* bias. Pada *layer* ini memiliki beberapa *hidden layer*, fungsi aktivasi, *layer output*, dan fungsi *loss*. Untuk menghindari *overfitting*, *CNN* memiliki regularisasi *Dropout* dimana *Dropout* akan memberikan nilai 0 pada sebuah *neuron* secara acak Ketika proses *backpropagation* dan *forward learning* [33].

Perhitungan *fully connected layer* dapat dilakukan pada ukuran 1 dimensi atau vektor sehingga output dari *layer* sebelumnya diubah ke vektor. Banyak *feature map* ditentukan oleh jumlah *output* [32]. Perhitungan *fully connected layer* tertera pada persamaan (2.4).

$$FCL_c = b_c + \sum_{i=1}^I W_{c,d} X_d \quad (2.6)$$

Keterangan:

$FCL_c$  = *fully connected layer output* pada *neuron output* ke- $c$

$b_c$  = nilai simpangan pada tiap *neuron* dari *fully connected layer*

$W_{c,d}$  = bobot *fully connected layer* tiap *neuron input* ke- $c$

$I$  = ukuran vektor *input layer* sebelumnya

$c$  = ukuran vektor *output* dari *fully connected layer*

### 2.8 Evaluasi Model Confusion Matrix dan Cohen Kappa Index

*Confusion matrix* adalah tabel klasifikasi yang diperoleh dari jumlah ketepatan hasil prediksi dengan data aktual pada masing-masing data uji [15]. *Confusion matrix* dapat diartikan sebagai sebuah metode untuk menganalisis seberapa baik kinerja sebuah model yang telah dibuat dalam mengidentifikasi data dari kelas yang berbeda-beda [31]. *Accuracy*, *sensitivity*, dan *spesitifity* merupakan ukuran kebaikan model prediksi. Ketiga nilai tersebut berasal dari perhitungan yang menggunakan *confusion matrix*.

**Tabel 2.**  
*Confusion Matrix*

| Kelas<br>Prediksi | Kelas Aktual |     |           |       |           |     |       |
|-------------------|--------------|-----|-----------|-------|-----------|-----|-------|
|                   | $C_1$        | ... | $C_{g-1}$ | $C_g$ | $C_{g+1}$ | ... | $C_l$ |
| $C_1$             | $TN_g$       |     | $FP_g$    |       | $TN_g$    |     |       |
| $\vdots$          |              |     |           |       |           |     |       |
| $C_{g-1}$         | $FN_g$       |     | $TP_g$    |       | $FN_g$    |     |       |
| $C_g$             |              |     |           |       |           |     |       |
| $C_{g+1}$         | $TN_g$       |     | $FP_g$    |       | $TN_g$    |     |       |
| $\vdots$          |              |     |           |       |           |     |       |
| $C_l$             | $TN_g$       |     | $FP_g$    |       | $TN_g$    |     |       |
|                   |              |     |           |       |           |     |       |

Definisi kelas klasifikasi dituliskan  $C_g$  dengan  $g = 1, 2, \dots, G$ . Dimana  $G$  adalah banyaknya kelas. Diperoleh perhitungan *accuracy*, *sensitivity*, *specificity*, dan *precision* sebagai berikut:

$$accuracy = \frac{\sum_{g=1}^G \frac{TP_g + TN_g}{TP_g + FN_g + FP_g + TN_g}}{G} \quad (2.7)$$

$$sensitivity = \frac{\sum_{g=1}^G \frac{TP_g}{TP_g + FN_g}}{G} \quad (2.8)$$

$$specificity = \frac{\sum_{g=1}^G \frac{TN_g}{FP_g + TN_g}}{G} \quad (2.9)$$

$$precision = \frac{\sum_{g=1}^G \frac{TP_g}{TP_g + FP_g}}{G} \quad (2.10)$$

dimana,

$TP_g$  = *true positive*, model tepat memprediksi kelas ke- $g$ .

$TN_g$  = *true negative*, model tepat memprediksi kelas diluar ke- $g$ .

$FP_g$  = *false positive*, model memprediksi kelas ke- $g$  namun aktualnya di luar kelas ke- $g$ .

$FN_g$  = *false negative*, model memprediksi kelas di luar kelas ke- $g$  namun aktualnya adalah kelas ke- $g$ .

$G$  = jumlah kelas.

Model yang didapatkan berpeluang memiliki akurasi yang tinggi namun sensitivitas dan spesifisitas tidak seimbang ataupun sebaliknya yaitu model yang didapatkan memiliki akurasi klasifikasi yang rendah namun sensitivitas dan spesifisitas seimbang [22].

Pada penelitian ini dilakukan pengukuran keakuratan model dengan menggunakan nilai *Cohen Kappa Index* [11].

**Tabel 3.**

| Proporsi Banyaknya Amatan |           |           |           |
|---------------------------|-----------|-----------|-----------|
| Kelas Aktual              |           |           | Total     |
| Kelas Prediksi            | $Pr_{11}$ | $Pr_{21}$ | $Pr_{1.}$ |
|                           | $Pr_{21}$ | $Pr_{22}$ | $Pr_{2.}$ |
| Total                     | $Pr_{.1}$ | $Pr_{.2}$ | 1         |

$$Kp = \frac{Pr_0 - Pr_e}{1 - Pr_e} \quad (2.11)$$

dimana,

$Kp$  = Nilai *kappa*

$Pr$  = Proporsi nilai amatan

$Pr_0 = Pr_{11} + Pr_{22} + Pr_{33}$

$Pr_e = (Pr_{.1}Pr_{.1}) + (Pr_{.2}Pr_{.2}) + (Pr_{.3}Pr_{.3})$  (Proporsi yang diharapkan secara kebetulan)

Interpretasi dari *Cohen Kappa Index* dengan nilai *kappa*  $< 0$  (sangat jelek), nilai *kappa*  $0 - 0,20$  (jelek), nilai *kappa*  $0,21 - 0,40$  (kurang), nilai *kappa*  $0,41 - 0,60$  (sedang), nilai *kappa*  $0,61 - 0,80$  (baik), dan nilai *kappa*  $0,81 - 1,00$  (sangat baik) [23].



## 2.9 Kerangka Sampel Area

Kerangka Sampel Area (KSA) merupakan metode yang digunakan untuk menghitung luas panen padi tiap bulannya dan dapat menghasilkan estimasi luas panen yang dikembangkan oleh BPS dengan bantuan dari BPPT [6].

Pengambilan sampel KSA dilakukan pada area yang berukuran  $6 \times 6$  km dan  $300 \times 300$  meter untuk seluruh daerah penelitian. Area  $300 \times 300$  meter tersebut dinamakan sebagai segmen dimana per segmen terdiri dari 9 subsegmen yang berukuran  $100 \times 100$  meter dimana setiap segmen dihitung sebagai satu sampel area. Pelaksanaan lapangan Survei KSA dilakukan dengan mengamati fase pertumbuhan padi pada titik tengah subsegmen yang telah ditentukan [4].

## 2.10 Citra Satelit Landsat-8

Satelit Landsat-8 direncanakan memiliki durasi misi selama 5-10 tahun. Satelit ini dilengkapi dua sensor yang merupakan hasil pengembangan sensor dari Landsat sebelumnya. Kedua sensor tersebut yaitu *Onboard Operational Land Imager (OLI)* dan *Sensor Thermal InfraRed Sensors (TIRS)*.

Landsat 8 memiliki beberapa keunggulan dibandingkan satelit sebelumnya. Spesifikasi dari Landsat 8 memiliki kanal-kanal maupun panjang rentang spektrum gelombang elektromagnetik yang ditangkap. Warna objek pada citra tersusun dari 3 warna dasar, yaitu *Red*, *Green* dan *Blue (RGB)*. Semakin banyak kanal penyusun komposit *RGB*, maka warna-warna objek menjadi lebih bervariasi [37].

## III. METODOLOGI PENELITIAN

### 3.1 Sumber Data

Data yang digunakan dalam penelitian ini berupa data gambar (*image*) RGB untuk variabel prediktor serta data label untuk variabel respon. Data label berasal dari sampel Survei KSA di Kabupaten Poso, Provinsi Sulawesi Tengah dari Bulan Januari 2018 hingga Februari 2021. Survei KSA yang terhitung sebanyak 58 segmen dimana segmen tersebut dapat dilakukan analisis periode tanamnya. Periode data Landsat-8 melakukan perekaman setiap 16 hari. Data citra satelit Landsat-8 diperoleh melalui *USGS*.

### 3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian adalah sebagai berikut:

**Tabel 4.**  
Variabel Penelitian

| Variabel | Keterangan  | Skala    |
|----------|---|----------|
| $X_p$    | Nilai <i>Red</i> , <i>Green</i> dan <i>Blue</i> setiap <i>pixel</i> (0-255)   | Interval |
| $Y$      | Kode fase pertumbuhan padi Survei KSA<br>1: fase vegetatif awal,<br>2: fase vegetatif akhir,<br>3: fase generatif,<br>4: fase panen<br>5: persiapan lahan | Nominal  |

Tabel 4 menunjukkan variabel yang akan digunakan dalam metode *CNN* dengan  $0 < i < n_x$ ;  $n_x$  merupakan total *pixel* ( $32 \times 32$ ) dari setiap gambar,  $0 < n < N$ ;  $N=2.523$  citra. Data tersebut

dilakukan *matching* antara data Survei KSA sebagai variabel respon dan data citra satelit Landsat-8.

### 3.3 Struktur Data

Struktur data hasil ekstraksi citra yang dihasilkan dari satelit Landsat-8 diubah menjadi bentuk nilai pada sebuah matriks yang dituliskan pada Tabel 3.2. Struktur data pada Tabel 3.2 dilakukan analisis dengan menggunakan metode *CNN*. Bentuk struktur datanya sebagai berikut ini.

**Tabel 5.**

Struktur Data Citra untuk Metode *CNN*

| $p$      | $Y_p$      | $X_{RGB, row, col}$   |
|----------|------------|---|
| 1        | $Y_1$      | $\begin{bmatrix} x_{R,1,1,1} & x_{R,1,1,2} & \cdots & x_{R,1,1,32} \\ x_{R,1,2,1} & x_{R,1,2,2} & \cdots & x_{R,1,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{R,1,32,1} & x_{R,1,32,2} & \cdots & x_{R,1,32,32} \end{bmatrix}$ $\begin{bmatrix} x_{G,1,1,1} & x_{G,1,1,2} & \cdots & x_{G,1,1,32} \\ x_{G,1,2,1} & x_{G,1,2,2} & \cdots & x_{G,1,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{G,1,32,1} & x_{G,1,32,2} & \cdots & x_{G,1,32,32} \end{bmatrix}$ $\begin{bmatrix} x_{B,1,1,1} & x_{B,1,1,2} & \cdots & x_{B,1,1,32} \\ x_{B,1,2,1} & x_{B,1,2,2} & \cdots & x_{B,1,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{B,1,32,1} & x_{B,1,32,2} & \cdots & x_{B,1,32,32} \end{bmatrix}$ |
| 2        | $Y_2$      | $\begin{bmatrix} x_{R,2,1,1} & x_{R,2,1,2} & \cdots & x_{R,2,1,32} \\ x_{R,2,2,1} & x_{R,2,2,2} & \cdots & x_{R,2,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{R,2,32,1} & x_{R,2,32,2} & \cdots & x_{R,2,32,32} \end{bmatrix}$ $\begin{bmatrix} x_{G,2,1,1} & x_{G,2,1,2} & \cdots & x_{G,2,1,32} \\ x_{G,2,2,1} & x_{G,2,2,2} & \cdots & x_{G,2,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{G,2,32,1} & x_{G,2,32,2} & \cdots & x_{G,2,32,32} \end{bmatrix}$ $\begin{bmatrix} x_{B,2,1,1} & x_{B,2,1,2} & \cdots & x_{B,2,1,32} \\ x_{B,2,2,1} & x_{B,2,2,2} & \cdots & x_{B,2,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{B,2,32,1} & x_{B,2,32,2} & \cdots & x_{B,2,32,32} \end{bmatrix}$ |
| $\vdots$ | $\vdots$   | $\vdots$  |
| 2523     | $Y_{2523}$ | $\begin{bmatrix} x_{R,2523,1,1} & x_{R,2523,1,2} & \cdots & x_{R,2523,1,32} \\ x_{R,2523,2,1} & x_{R,2523,2,2} & \cdots & x_{R,2523,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{R,2523,32,1} & x_{R,2523,32,2} & \cdots & x_{R,2523,32,32} \end{bmatrix}$ $\begin{bmatrix} x_{G,2523,1,1} & x_{G,2523,1,2} & \cdots & x_{G,2523,1,32} \\ x_{G,2523,2,1} & x_{G,2523,2,2} & \cdots & x_{G,2523,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{G,2523,32,1} & x_{G,2523,32,2} & \cdots & x_{G,2523,32,32} \end{bmatrix}$   |

Tabel 6.

Struktur Data Citra untuk Metode CNN (lanjutan)

| $p$  | $Y_p$      | $X_{RGB, row, col}$  |
|------|------------|--|
| 2523 | $Y_{2523}$ | $\begin{bmatrix} x_{B,2523,1,1} & x_{B,2523,1,2} & \cdots & x_{B,2523,1,32} \\ x_{B,2523,2,1} & x_{B,2523,2,2} & \cdots & x_{B,2523,2,32} \\ \vdots & \vdots & \ddots & \vdots \\ x_{B,2523,32,1} & x_{B,2523,32,2} & \cdots & x_{B,2523,32,32} \end{bmatrix}$ |

### 3.4 Langkah Analisis

#### a. Tahap Penyusunan Data

1. Mengunduh data KSA BPS pada website <http://ksa.bps.go.id> hingga didapatkan variabel respon berupa label KSA fase pengamatan periode Bulan Februari 2018 hingga Bulan April 2018 dan koordinat sampel KSA di Kabupaten Poso, Sulawesi Tengah.
2. *Fold* data validasi dari Februari 2018 – April 2018 menjadi *5-fold* dan membagi data menjadi *training* dan validasi serta menetapkan data *testing* dan menyimpannya menjadi sebuah *dataset*.
3. Melakukan ekstraksi titik koordinat KSA setiap *dataset* kedalam bentuk file peta *SHP* dalam bentuk titik dengan menggunakan program R.
4. Mengunduh gambar (*image*) sesuai batas koordinat wilayah Poso yaitu 120,2062° BT- 120,9238° BT dan 1,14859° LS- 2,12314° LS melalui *USGS* setiap bulan dari Januari 2018 – Juli 2018 (data *training*). Gambar yang diunduh setiap bulannya terdapat tujuh citra dalam bentuk citra *band 1 – band 7*.
5. Melakukan optimasi pada tiap citra untuk didapatkan hasil gambar (*image*) yang optimum dengan mengubah *brightness* dan *contrast* pada data *image*.
6. Melakukan *rastering* pada setiap citra hingga didapatkan satu citra yang dilakukan analisis.
7. Melakukan *overlay* terhadap titik dalam bentuk *shp* terhadap setiap citra.
8. Melakukan pemotongan terhadap titik yang telah dihasilkan dari langkah *overlay*.
9. Mengubah hasil data citra dari format *cmk* ke format *jpg* serta mengubah ukuran citra dari 120 × 120 *pixel* menjadi ukuran citra 32 × 32 *pixel*.
10. Warna terbentuk dari kombinasi tiga warna R (Red), G (Green), dan B (Blue). Setiap komponen warna utama memiliki tingkat gradasi yang berkisar antara 0-255, semakin kecil nilai maka warna dihasilkan semakin gelap.
11. Data yang didapatkan adalah 3 matriks dengan ukuran panjang dan lebar gambar (32 × 32 *pixel*) yang berisi nilai antara 0-255.
12. Membentuk struktur data untuk dianalisis dengan metodologi *CNN*.

#### b. Tahap *pre-processing* data

1. Melakukan pembersihan data meliputi deteksi kolom yang diunduh dari citra satelit Landsat-8 terdapat *missing data* yang disebabkan oleh tutupan awan yang terjadi pada area penelitian pada data citra.
2. Melakukan *filter* terhadap label KSA sehingga label yang digunakan menjadi *input* untuk label pertumbuhan yaitu label 1 hingga label 7.
- c. Membagi data *training* dan data *testing* dengan menggunakan *K-fold Cross Validation*.
- d. Melakukan klasifikasi dengan menggunakan metode *Convolutional Neural Network (CNN)* dengan langkah-langkah sebagai berikut.
  1. Menentukan arsitektur yang akan digunakan dalam penelitian.
  2. Melakukan *inisialisasi* parameter bias dan bobot.
  3. Melakukan perhitungan operasi konvolusi.
  4. Menghitung fungsi aktivasi *tangen hiperbolik (TanH)*.
  5. Melakukan *max pooling*.
  6. Merubah matriks dari hasil konvolusi menjadi bentuk vektor pada *fully connected layer*.
  7. Melakukan proses klasifikasi dengan menggunakan fungsi *softmax* untuk mendapatkan *predicted value*.
  8. Menghitung nilai *loss* dari *predicted value*.
  9. Mencoba parameter yang telah dihasilkan pada data *testing*.
  10. Menghitung nilai kinerja klasifikasi *sensitifity*, *specificity*, *accuracy*, *cohen kappa index* dan *Matthew Correlation Coefficient* pada hasil klasifikasi data *testing*.
  11. Melakukan langkah ke-1 hingga langkah ke-10 pada *fold* selanjutnya.
  12. Menghitung rata-rata kinerja klasifikasi dari keseluruhan *fold*.
- e. Membandingkan kinerja dari metode *CNN* dengan dan tanpa dilakukan *pre-processing* yang telah didapatkan dengan menggunakan *sensitifity*, *specificity*, *accuracy*, *Cohen Kappa Index*.
- f. Mengambil kesimpulan dan saran.

## IV. HASIL DAN PEMBAHASAN

### 4.1 Deskriptif Data

Citra satelit digunakan untuk membantu melakukan Survei KSA dengan gratis dan cepat. Batas koordinat Poso yang terletak pada 120,11° BT – 120,8681° BT dan 1,1125° LS – 2,23167° LS yang digunakan untuk pengambilan data gambar citra satelit melalui *USGS*. Sebaran titik KSA antara 120,2062° BT – 120,9238° BT dan 1,14859° LS – 2,42314° LS. Data titik KSA dari setiap titik amatan yang diperoleh dirubah menjadi *file* citra dengan ekstensi *shp* yang akan digunakan sebagai titik acuan *clipping rastering* data gambar yang sebelumnya telah diperoleh melalui *USGS*.

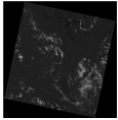

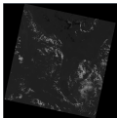
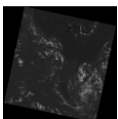
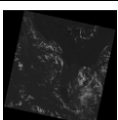
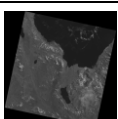
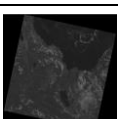
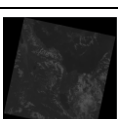
### 4.2 Data Gambar

Pengunduhan data gambar pada *USGS* berdasarkan titik

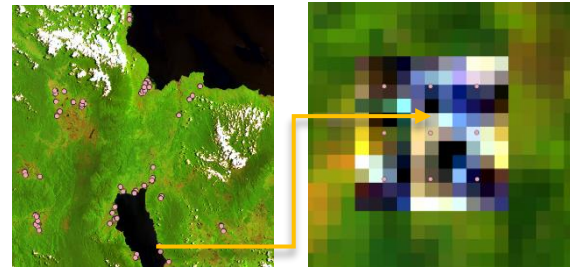
koordinat tersebut dianggap memiliki kriteria tutupan awan kurang dari 50 persen dan bayangan awan kurang dari 30 persen. Peta yang diunduh dari Februari 2018 hingga Mei 2018 sehingga terdapat 4 peta yang diambil pada lokasi yang sama namun dalam waktu yang berbeda sebagai data *training*. Setiap peta memiliki 7 informasi *layer* yaitu citra *band 1* hingga *band 7* sehingga total peta yang diunduh 28 citra yang memiliki resolusi tinggi kemudian dilakukan *raster* pada setiap peta tiap bulannya seperti pada Tabel 7.

**Tabel 7.**

*Raster Citra Landsat-8 Bulan April 2018*

| Sensor | Bentuk Citra Landsat-8  | Hasil Raster Citra   |
|--------|---|--|
| Band 1 |    |  |
| Band 2 |    |  |
| Band 3 |    |  |
| Band 4 |   |  |
| Band 5 |  |  |
| Band 6 |  |  |
| Band 7 |  |  |

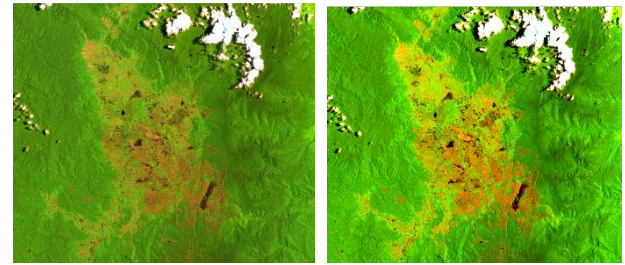
Citra hasil *raster* dilakukan pemotongan atau *clipping* berdasarkan titik shp yang sebelumnya telah didapatkan sehingga diperoleh gambar dengan format *tiff*. Format *tiff* merupakan format citra yang memiliki 4 matriks dengan format *CMYK* (*Cyan, Magenta, Yellow, dan Black*). Hasil dari *clipping* merupakan *input* yang digunakan pada metode *CNN*. Proses *clipping* terlihat pada Gambar 9.



**Gambar 9.** *Clipping Citra Hasil Raster untuk Input Data CNN*

#### 4.3 Pre-processing Brightness dan Contrast

Langkah-langkah yang dilakukan sama seperti perlakuan pada data gambar tanpa *pre-processing brightness* dan *contrast*. Hal yang membedakan terdapat pada tahap *raster*, dimana untuk *preprocessing brightness* dan *contrast* dilakukan setelah citra *raster* terbentuk dengan mengubah nilai *alpha*, *beta* dan *gamma* dengan tujuan memperoleh citra yang lebih jelas seperti Gambar 10.



**Gambar 10.** *Citra Sebelum dan Sesudah Dilakukan Pre-processing Brightness dan Contrast*

Gambar 10 menunjukkan perubahan sebelum dan sesudah dilakukan *pre-processing brightness* dan *contrast* dengan perubahan peningkatan *alpha* (*contrast*) sebesar 30, *beta* (*brightness*) sebesar 20, dan *corrector gamma* sebesar 0,1 dengan software *QGIS* versi 3.16.16.

#### 4.4 Convolutional Neural Network dengan Data Aktual

Keefisienan arsitek yang dibentuk ditandai dari hasil prediksi yang dihasilkan. Kerumitan pada sebuah arsitek *CNN* ini dipengaruhi oleh kerumitan data yang dijadikan input. Ukuran kernel yang digunakan berdasarkan penelitian terdahulu dalam melakukan klasifikasi pada vegetasi di permukaan bumi. Hasil pengolahan data menghasilkan parameter pada setiap layer. Berikut jumlah layer yang dihasilkan pada setiap layer bisa terlihat pada Tabel 8.

**Tabel 8**

*Bentuk Arsitektur dan Jumlah Parameter Terbentuk*

| <i>Layer (type)</i>                   | <i>Output Shape</i> | <i>Parameter</i> |
|---------------------------------------|---------------------|------------------|
| <i>conv2d(Conv2D)</i>                 | (None, 30, 30, 32)  | 896              |
| <i>max_pooling2d (MaxPooling2D)</i>   | (None, 15, 15, 32)  | 0                |
| <i>conv2d_1(Conv2D)</i>               | (None, 13, 13, 64)  | 18.496           |
| <i>max_pooling2d_1 (MaxPooling2D)</i> | (None, 6, 6, 64)    | 0                |
| <i>conv2d_2(Conv2D)</i>               | (None, 4, 4, 128)   | 73.856           |



**Tabel 9.**  
Bentuk Arsitektur dan Jumlah Parameter Terbentuk (lanjutan)

| Layer (type)                   | Output Shape      | Parameter |
|--------------------------------|-------------------|-----------|
| max_pooling2d_2 (MaxPooling2D) | (None, 2, 2, 128) | 0         |
| flatten (Flatten)              | (None, 512)       | 0         |
| dense (Dense)                  | (None, 128)       | 65.664    |
| dense_1 (Dense)                | (None, 128)       | 16.512    |
| dense_2 (Dense)                | (None, 5)         | 645       |

Menggunakan 100 *epoch* sebagai pengukur kinerja dengan menggunakan *k-fold cross validation* dengan 5 *fold* dengan data Februari 2018 hingga April 2018 terlihat pada Tabel 10.

**Tabel 10.**  
Kinerja Klasifikasi CNN 100 *epoch* dengan Data Aktual

| K-fold        | Sensitivity   | Specificity   | Accuracy      | Cohen Kappa Index |
|---------------|---------------|---------------|---------------|-------------------|
| Fold 1        | 0,2411        | 0,2390        | 0,2565        | 0,0548            |
| Fold 2        | 0,2616        | 0,2639        | 0,2752        | 0,0750            |
| <b>Fold 3</b> | <b>0,2705</b> | <b>0,2711</b> | <b>0,2897</b> | <b>0,0952</b>     |
| Fold 4        | 0,2284        | 0,2280        | 0,2440        | 0,0416            |
| Fold 5        | 0,2423        | 0,2406        | 0,2579        | 0,0575            |
| Rata-rata     | 0,2488        | 0,2485        | 0,2647        | 0,0648            |

Kinerja terbaik pada arsitektur CNN yang digunakan dalam penelitian ini yaitu dengan menggunakan 100 *epoch* dengan perbandingan 5-Fold dari data aktual dapat terlihat pada kinerja masing-masing *fold*. Terlihat pada Tabel 10 didapatkan *fold* terbaik yaitu pada *fold* 3.

#### 4.5 Convolutional Neural Network dengan Data Brightness

Model yang telah didapatkan dari arsitektur CNN sebelumnya dengan menggunakan 100 *epoch* dengan data hasil *pre-processing brightness* dan *contrast* Februari 2018 hingga April 2018 diperoleh hasil sebagai berikut ini.

**Tabel 11.**  
Kinerja Klasifikasi CNN 100 *epoch* dengan Data Brightness

| K-fold        | Sensitivity   | Specificity   | Accuracy      | Cohen Kappa Index |
|---------------|---------------|---------------|---------------|-------------------|
| Fold 1        | 0,2780        | 0,2791        | 0,2903        | 0,0958            |
| <b>Fold 2</b> | <b>0,2802</b> | <b>0,2754</b> | <b>0,3010</b> | <b>0,1076</b>     |
| Fold 3        | 0,2318        | 0,2319        | 0,2540        | 0,0496            |
| Fold 4        | 0,2283        | 0,2308        | 0,2460        | 0,0343            |
| Fold 5        | 0,2317        | 0,2324        | 0,2520        | 0,0494            |
| Rata-rata     | 0,25          | 0,2499        | 0,2687        | 0,0673            |

Kinerja terbaik pada arsitektur CNN yang digunakan dalam penelitian ini yaitu dengan menggunakan 100 *epoch* dengan perbandingan 5-Fold dari data *pre-processing brightness* dapat terlihat pada kinerja masing-masing *fold*. Terlihat pada Tabel 11 didapatkan *fold* terbaik yaitu pada *fold* 2.

#### 4.6 Perbandingan Kinerja Antar Metode

Perbandingan kinerja dengan CNN dengan data asli dan data hasil *pre-processing brightness* terlihat bahwa tidak terlalu banyak perubahan yang signifikan baik dengan data asli maupun data yang sudah dilakukan *pre-processing brightness* terlihat pada Tabel 12.

**Tabel 12.**  
Perbandingan Kinerja Antar Model

| Data            | Sensitivity | Specificity | Accuracy | Cohen Kappa Index |
|-----------------|-------------|-------------|----------|-------------------|
| Data Asli       | 0,2488      | 0,2485      | 0,2647   | 0,0648            |
| Data Brightness | 0,25        | 0,2499      | 0,2687   | 0,0673            |

Tabel 12 menunjukkan bahwa data hasil *pre-processing brightness* lebih baik namun hanya terlihat kenaikan akurasi yang sedikit dan tidak signifikan, oleh karena itu dilakukan prediksi dengan data *brightness* dengan menggunakan model *fold* 2 dan diperoleh hasil prediksi sebagai berikut ini.

**Tabel 13.**  
Probabilitas Prediksi Label Hasil CNN dengan data *brightness*

|     | Label 1   | Label 2   | Label 3   | Label 4   | Label 5   |
|-----|-----------|-----------|-----------|-----------|-----------|
| 1   | 9,999e-01 | 5,869e-08 | 3,183e-09 | 5,588e-07 | 5,532e-07 |
| 2   | 5,936e-03 | 5,318e-08 | 1,239e-08 | 9,939e-01 | 1,983e-04 |
| 3   | 9,981e-01 | 8,569e-04 | 7,969e-07 | 2,564e-09 | 1,073e-03 |
| ... | ...       | ...       | ...       | ...       | ...       |
| 110 | 8,664e-01 | 1,135e-02 | 3,027e-06 | 9,455e-04 | 1,213e-01 |
| 111 | 7,042e-07 | 3,860e-08 | 9,999e-01 | 9,268e-09 | 4,979e-08 |
| 112 | 1,009e-04 | 1,747e-03 | 3,695e-03 | 9,944e-01 | 1,716e-05 |

**Tabel 14.**  
Hasil Prediksi Model CNN dengan data *brightness*

| Metode                     | Aktual | Prediksi |   |   |    |   |
|----------------------------|--------|----------|---|---|----|---|
|                            |        | 1        | 2 | 3 | 4  | 5 |
| CNN Data Brightness Fold 2 | 1      | 8        | 0 | 1 | 3  | 0 |
|                            | 2      | 7        | 2 | 4 | 10 | 5 |
|                            | 3      | 2        | 5 | 8 | 10 | 7 |
|                            | 4      | 6        | 2 | 5 | 11 | 3 |
|                            | 5      | 4        | 0 | 4 | 4  | 1 |

Pada Tabel 13 dan Tabel 14 masing-masing menunjukkan hasil probabilitas dan *confusion matrix* terhadap data *testing* prediksi pada Bulan Mei 2018 menunjukkan bahwa fase pertumbuhan padi yang terlihat memiliki fase vegetatif awal, fase vegetatif akhir, fase generatif, fase panen, dan fase persiapan lahan dengan total ada 112 titik amatan. Jika dilihat pada Tabel 14 model CNN dengan menggunakan model *fold* 2 dengan menggunakan data hasil *pre-processing brightness* dapat memprediksi secara tepat pada 30 titik amatan dengan rincian 8 titik prediksi fase vegetatif awal, 2 titik prediksi fase vegetatif akhir, 8 titik prediksi fase generatif, 11 titik prediksi

fase panen dan 1 titik prediksi fase persiapan lahan. Berdasarkan *confusion matrix*, dapat dihitung hasil kinerja prediksi seperti nilai *accuracy*, *sensitivity*, *specificity*, *cohen kappa index*. Hasil kinerja prediksi *CNN* terlihat pada Tabel 15.

**Tabel 15.**  
Kinerja Ukuran Kebaikan Prediksi

| Model               | Sensitivity | Specificity | Accuracy | Cohen Kappa Index |
|---------------------|-------------|-------------|----------|-------------------|
| CNN Data Brightness | 0,2945      | 0,2468      | 0,2678   | 0,0843            |

Terlihat dari Tabel 15 untuk prediksi data fase pertumbuhan tanaman padi dengan model *CNN fold 2* dengan data *brightness* menunjukkan hasil yang kurang baik dengan terlihatnya nilai *accuracy* sebesar 0,2678, *specificity* sebesar 0,2468, *sensitivity* sebesar 0,2945 dan *cohen kappa index* sebesar 0,0843. Berdasarkan hasil tersebut, model *CNN* pada penelitian ini belum bisa memprediksi fase pertumbuhan padi Survei KSA dengan baik.

## V. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

1. Dalam penelitian ini menunjukkan bahwa data padi yang digunakan pada sampel KSA ini cenderung lebih banyak pada fase panen dan cenderung sedikit pada fase persiapan lahan.
2. Kinerja yang dihasilkan model *CNN* dengan data aktual untuk memprediksi fase pertumbuhan padi menghasilkan nilai akurasi yang rendah, yaitu sebesar 0,2647, *sensitivity* sebesar 0,2488, *specificity* sebesar 0,2485 dan *cohen kappa index* sebesar 0,0648 sedangkan pada *CNN* dengan data hasil *pre-processing brightness* didapatkan akurasi 0,2687, *sensitivity* sebesar 0,25, *specificity* sebesar 0,2499 dan *cohen kappa index* sebesar 0,0673.
3. Menggunakan model terbaik yaitu dengan menggunakan model *CNN* dengan data *brightness*, prediksi panen tanaman padi diperoleh hasil yang tergolong buruk karena gagal dalam memprediksi panen tanaman padi dimana hanya dapat memprediksi secara tepat pada 30 titik amatan dengan rincian 8 titik prediksi fase vegetatif awal, 2 titik prediksi fase vegetatif akhir, 8 titik prediksi fase generatif, 11 titik prediksi fase panen dan 1 titik prediksi fase persiapan lahan dari total 112 titik amatan.

### 5.2 Saran

Melihat kesimpulan yang diperoleh, maka saran sebagai pertimbangan penelitian selanjutnya adalah sebagai berikut.

1. Menggunakan citra satelit dengan resolusi yang lebih baik lagi agar tidak banyak citra yang tertutup awan, seperti Sentinel-1, Sentinel-2 atau citra berbayar lainnya.
2. Melakukan *cleaning data* yang lebih baik lagi supaya diperoleh data yang lebih optimum.
3. Menggunakan arsitektur *CNN* yang lebih baik lagi supaya hasil yang didapatkan lebih optimum.

4. Memilih lumbung padi nasional sebagai lokasi penelitian

## DAFTAR PUSTAKA

- [1] Agresti, A. (2018). *An Introduction to Categorical Data Analysis*. John Wiley & Sons, Inc.
- [2] Anonim. (2019, August 21). *Mengenal Deep Learning*. Retrieved from <https://inixindojogja.co.id/mengenal-deep-learning/>
- [3] Baig, M. H., Zhang, L., Shuai, T., & Tong, Q. (2014). Derivation of a Tasselled Cap Transformation Based on Landsat 8 at-Satellite Reflectance. *Remote Sensing Letters*, 5(5), 423-431.
- [4] BPS. (2015). *Pedoman Pelaksanaan Uji Coba Sistem Kerangka Sampel Area (KSA)*. Jakarta: BPS.
- [5] BPS. (2021). *Ekonomi Indonesia Triwulan IV 2021 Tumbuh 5,02 Persen (y-on-y)*. Jakarta: BPS.
- [6] BPS, & BPPT. (2018). *Luas Panen dan Produksi Beras 2018*. Jakarta: BPS.
- [7] BPS, & BPPT. (2019). *Luas Panen dan Produksi Padi di Indonesia 2019*. Jakarta: BPS.
- [8] Brownlee, J. (2020, August 14). *What is Deep Learning?* Retrieved from <https://machinelearningmastery.com/what-is-deep-learning/>
- [9] Buduma, N. (2017). *Fundamental of Deep Learning : Designing Next-Generation Machine Intelligence Algorithms*. O'reilly Media, Inc.
- [10] Chawla, N. v., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357. <https://doi.org/10.1613/jair.953>
- [11] Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37-46.
- [12] CS231n. (2021). *CS231n Convolutional Neural Networks for Visual Recognition*. Retrieved from CS231n Course: <http://cs231n.github.io/convolutional-networks/>
- [13] DQLAB. (2021, April 12). *DQLAB*. Retrieved from Pentingnya Preprocessing dalam Pengolahan Data Statistik: <https://www.dqlab.id/pentingnya-preprocessing-dalam-pengolahan-data-statistik>
- [14] Eka Putra, W. S. (2016). Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101. *Jurnal Teknik ITS*, 5(1).
- [15] Fawcett, J. (2005). Criteria for Evaluation of Theory. *Nursing Science Quarterly*, 131-135.
- [16] Furht, B., & Villanustre, F. (2016). *Big Data Technologies and Applications*. USA: Springer.
- [17] Gokgoz, E., & Subasi, A. (2015). Comparison of Decision Tree Algorithms for EMG Signal Classification Using DWT. *Biomedical Signal Processing and Control*, Vol 18, 138-144.
- [18] Gonzalez, R. C., & Woods, R. E. (2008). *Digital Image Processing 3rd*. Upper Saddle River: Pearson.

- [19] Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. Springer: Stanford.
- [20] Ilahiyah, S., & Nilogiri, A. (2018). Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network. *JUSTINDO*, 3(2).
- [21] Irzal Ahmad Sabilla. (2020). *Arsitektur Convolutional Neural Network (CNN) untuk Klasifikasi Jenis dan Kesegaran Buah pada Neraca Buah*.
- [22] Jr., D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression (Vol. 398)*. John Wiley & Sons, Inc.
- [23] Kundel, H. L., & Polansky, M. (2003). Measurement of Observer Agreement. *Radiology*, 228(2), 303-308.
- [24] Ma, M. (2015). *Convolutional Neural Network for Computer Vision and Natural Language Processing*. Graduate Center: The City University of New York.
- [25] Magdalena, R., Saidah, S., Caecar, K. P., & Trisnamulya, A. P. (2021). JEPIN (Jurnal Edukasi dan Penelitian Informatika) Klasifikasi Tutupan Lahan Melalui Citra Satelit SPOT-6 dengan Metode Convolutional Neural Network (CNN). *Jurnal Edukasi Dan Penelitian Informatika*, 7.
- [26] Marsuhandi, A. H., Soleh, A. M., Wijayanto, H., & Domiri, D. D. (2019). Pemanfaatan Ensemble Learning dan Penginderaan Jauh untuk Pengklasifikasian Jenis Lahan Padi. *Seminar Nasional Official Statistics, 2019(1)*, 188-195.
- [27] Mathworks. (2017). *Convolutional Neural Network*. Retrieved from <https://www.mathworks.com/discovery/convolutional-neural-network.html>
- [28] Miranda, E., & Aryuni, M. (2021). SISTEMASI: Jurnal Sistem Informasi Klasifikasi Tutupan Lahan Menggunakan Convolutional Neural Network pada Citra Satelit Sentinel-2. *Jurnal Sistem Informasi*, 10. <http://sistemasi.ftik.unisi.ac.id>
- [29] Munir, R. (2004). *Pengolahan Citra Digital dengan Pendekatan Algoritmik*. Bandung: Informatika.
- [30] Parsa, M., Dirgahayu, D. D., Manalu, J., Carolita, I., & Harsanugraha, W. (2017). Uji Model Fase Pertumbuhan Padi Berbasis Citra Modis Multiwaktu di Pulau Lombok (The Testing of Phase Growth Rice Model Based on Multitemporal Modis in Lombok Island). *Penginderaan Jauh dan Pengolahan Data Citra Digital*, 14(1).
- [31] Santra, A. K., & Christy, C. J. (2012). Genetic Algorithm and Confusion Matrix for Document Clustering. *International Journal of Computer Science Issues (IJCSI)*, 9(1), 322.
- [32] Singh, A. (2020, February 11). *Demystifying the Mathematics Behind Convolutional Neural Network (CNNs)*. Retrieved from Analytics Vidhya: <https://www.analyticsvidhya.com/blog/2020/02/mathematics-behind-convolutional-neural-network/>
- [33] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting, no.15. *Machine Learning Research*, 1929-1958.
- [34] Susanti, D. D. (2019). Perbandingan Penyebaran Sampel dan Varian Ubinan Berdasarkan Ubinan Rumah Tangga dan Ubinan Kerangka Sampel Area (KSA) di Kabupaten Wonosobo Provinsi Jawa Tengah. *Seminar Nasional Official Statistics, 2019(1)*, 202-209.
- [35] Tagliaferri, L. (2017). *An Introduction to Machine Learning*. DigitalOcean.
- [36] Triscowati, D. W., Sartono, B., Kurnia, A., Dirgahayu, D., & Wijayanto, A. W. (2020). Classification of Rice-plant Growth Phase Using Supervised Random Forest Method Based on Landsat-8 Multitemporal Data. *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, 16(2), 187-196.
- [37] USGS. (2018). *Landsat 8 (L8) Data User Handbook*. CRC Press.
- [38] Walpole, R. E. (2007). *Pengantar Metode Statistika*. Jakarta: PT Gramedia Pustaka Utama.
- [39] Xu, H., Zhongwei, L., Xin, L., & Tao, Z. (2019). Log facies recognition based on Convolutional Neural Network. *Shiyou Diqu Wuli Kantan/Oil Geophys. Prospect.*, 1159-1165.
- [40] Yadav, S., & Shukla, S. (2016). Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. *IEEE 6th International Conference on Advanced Computing (IACC)*, 78-83.
- [41] Zikra, H., & Buana, W. P. (2020). Analisis Perbandingan Desain Sampling Survei Kerangka Sampel Area (KSA). *Seminar Nasional Official Statistics, 2020(1)*, 1326-1336