KLASIFIKASI HAMA DAN PENYAKIT TANAMAN PADI DENGAN METODE CONVOLUTIONAL NEURAL NETWORK

PROPOSAL SKRIPSI



Oleh:

Dzakiyyah Azahra Mooduto

NIM : 531420046

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK

UNIVERSITAS NEGERI GORONTALO

2023

# DAFTAR ISI

[DAFTAR ISI ii](#_Toc159376786)

[BAB 1 3](#_Toc159376787)

[PENDAHULUAN 3](#_Toc159376788)

[1.1. Latar Belakang 3](#_Toc159376789)

[1.2. Rumusan Masalah 6](#_Toc159376790)

[1.3. Ruang Lingkup Penelitian 6](#_Toc159376791)

[1.4. Tujuan Penelitian 6](#_Toc159376792)

[1.5. Manfaat Penelitian 6](#_Toc159376793)

[BAB 2 4](#_Toc159376794)

[TINJAUAN PUSTAKA 4](#_Toc159376795)

[2.1. Tanaman Padi 4](#_Toc159376796)

[2.2. Hama dan Penyakit Tanaman Padi 9](#_Toc159376797)

[2.3. Citra Digital 11](#_Toc159376798)

[2.4. Deep Learning 12](#_Toc159376799)

[2.5. Convolutional Neural Network 13](#_Toc159376800)

[2.6. Tensorflow 16](#_Toc159376801)

[2.7. Penelitian Terkait 21](#_Toc159376802)

[BAB 3 11](#_Toc159376803)

[METODOLOGI PENELITIAN 11](#_Toc159376804)

[3.1. Lokasi dan Waktu Penelitian 11](#_Toc159376805)

[3.2. Rancangan Penelitian 27](#_Toc159376806)

[3.3. Data Penelitian 30](#_Toc159376807)

[BAB 4 32](#_Toc159376808)

[HASIL DAN PEMBAHASAN 32](#_Toc159376809)

[4.1. Hasil Penelitian 32](#_Toc159376810)

[4.2. Pembahasan 47](#_Toc159376811)

[4.2.1. Hasil Pengujian Terbaik 47](#_Toc159376812)

[4.2.2. Model *Deep Learning* Pengujian Akhir 49](#_Toc159376813)

[BAB V 53](#_Toc159376814)

[PENUTUP 53](#_Toc159376815)

[5.1. Kesimpulan 53](#_Toc159376816)

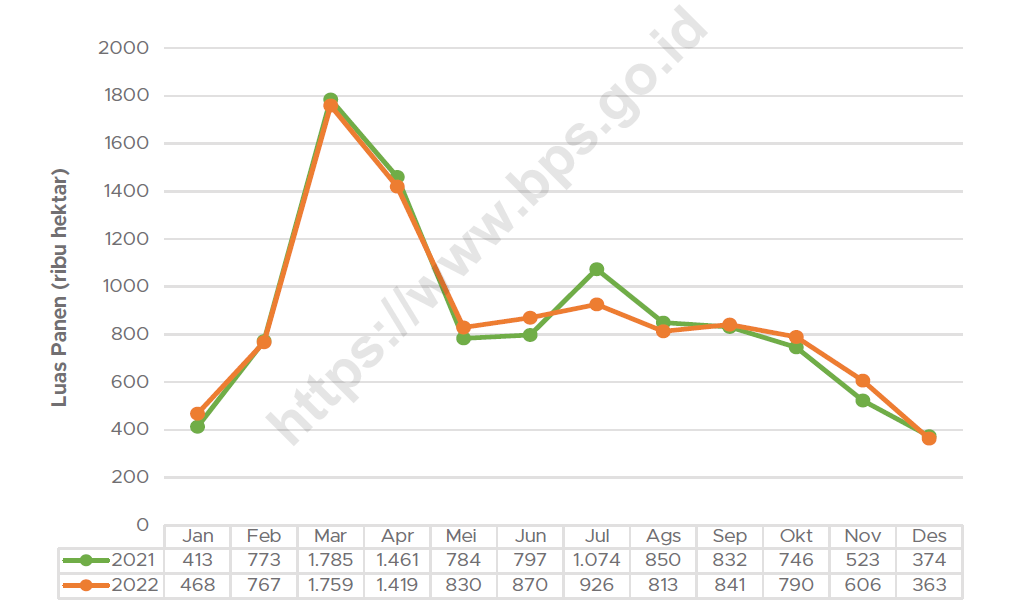
[5.2. Saran 54](#_Toc159376817)

# BAB 1

# PENDAHULUAN

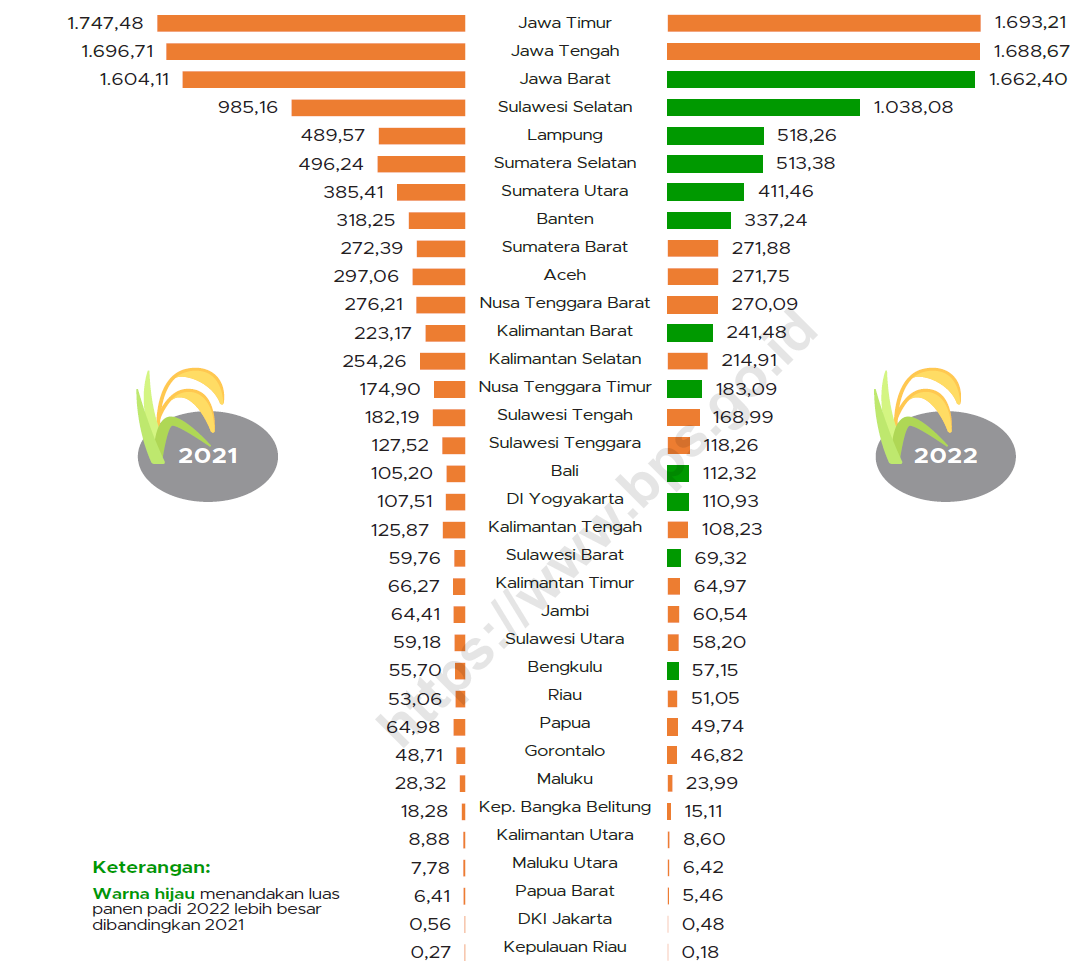
## Latar Belakang

Padi (*Oryza sativa*) yang menghasilkan beras merupakan salah satu makanan pokok di Asia, khususnya Indonesia maupun negara tropis lainnya. Produksi padi tiap tahunnya sangat dibutuhkan demi memenuhi kebutuhan pangan di negara Indonesia. Jumlah penduduk yang semakin meningkat menyebabkan tingkat konsumi dan permintaan pasar dari beras akan semakin meningkat. Oleh karena itu, produksi beras harus dipastikan berada dalam angka yang konstan dan tidak mengalami penurunan.



Berdasarkan hasil survei Kerangka Sampel Area (KSA), pola panen padi di Indonesia pada 2022 secara umum hamper mirip dengan pola panen padi pada 2021. Jika dibandingkan dengan 2021, luas panen padi 2022 mengalami kenaikan sebesar 40,87 ribu hektar (0,39 persen). Namun, jika dilihat secara lebih detail menurut provinsi, terdapat sebanyak 12 provinsi yang mengalami peningkatan lua

panen padi dibandingkan 2021. Sementara itu, 22 provinsi lainnya mengalami penurunan luas panen padi pada 2022 dibandingkan tahun sebelumnya. Berdasarkan gambar di bawah, Provinsi Gorontalo mengalami penurunan dari Angka 48,71 ribu hektar ke 46,82 ribu hektar. Angka tersebut tidak didapatkan secara tiba-tiba tetapi terdapat faktor yang menyebabkan gagal panen di beberapa wilayah (Khasanah & Astuti, t.t.).



Faktor penyebab gagal panen dalam produktivitas tanaman padi sangatlah banyak, salah satunya adalah hama dan penyakit pada daun padi (Saputra dkk., 2021). Serangan dari organisme pengganggu tanaman (OPT) sering menyebabkan rusaknya tanaman padi khususnya di bagian daun (Anam Alidrus dkk., 2021). Selain itu, keterlambatan proses diagnosis dan tindaklanjut secara manual menyebabkan hama dan penyakit pada tanaman padi mencapai pada tahap sangat parah dan tidak tertolong. Hal tersebut terjadi karena minimnya pengetahuan petani serta tidak adanya antisipasi sejak dini, sehingga membutuhkan proses manual dan berakibat pada terjadinya gagal panen.

Hama dan penyakit pada tanaman padi bisa dilihat dari perubahan akar, batang, biji, daun, dan lain-lain. Namun tidak semua manusia dapat mengetahui tentang hama dan penyakit tersebut. Kerugian dari hama dan penyakit tanaman padi lainnya belum banyak dilaporkan. Hama dan penyakit akibat serangan OPT tersebut terus berfluktuasi dari satu tempat ke tempat lainnya maupun dari waktu ke waktu. Direktorat Jenderal Tanaman Pangan mengklasifikasi OPT (Organisme Pengganggu Tanaman) utama ke dalam 7 jenis, di antaranya adalah penggerek batang padi putih (PBPP), ulat grayak, wereng batang coklat (wbc), tikus, hama putih palsu (HPP), hawar daun bakteri/kresek. dan *blast.* Tujuh jenis hama dan penyakit tanaman yang beberapa tahun terakhir berpengaruh besar terhadap terjadinya gagal panen di produksi padi Indonesia.

Dinas Pertanian Provinsi Gorontalo menghadapi masalah kurangnya literasi dan antisipasi dari petani terhadap hama dan penyakit tanaman padi dengan beberapa pendekatan, salah satunya adalah dengan menciptakan aplikasi Balintan Smart Mobile. Aplikasi ini digunakan untuk berbagai aktivitas dan kepentingan dalam pelayanan pertanian seperti pemantauan dan pelaporan aktivitas petugas lapang, informasi seputar pertanian, layanan pelaporan petani, dan sebagainya. Akan tetapi, dari berbagai fungsi dan manfaat dari aplikasi tersebut, Dinas Pertanian Provinsi Gorontalo membutuhkan sebuah sistem pendeteksi hama dan penyakit pada tanaman padi untuk petani.

Dewasa ini, sistem klasifikasi citra sering digunakan untuk menjadi solusi sebuah permasalahan. Klasifikasi citra pada umumnya menggunakan algoritma *Convolutional Neural Network* (CNN). CNN adalah sebuah metode jaringan syaraf yang sangat terkenal dan cukup diminati. CNN digunakan untuk mengolah data dengan struktur grid yaitu salah satunya berupa citra dua dimensi dan juga mampu memproses data dengan dimensi tinggi seperti video. CNN bekerja mirip dengan *neural network standard*, hanya saja perbedaannya yaitu menggunakan kernel dua dimensi atau dimensi tinggi di setiap unit pada lapisan CNN yang akan dikonvolusi. Untuk melakukan penggabungan fitur dari spasial dengan bentuk spasial yang serupa media *input* maka digunakan kernel yang ada di dalam CNN. Kemudian pada CNN untuk mengurangi jumlah variabel agar lebih mudah untuk dilakukan pembelajaran yaitu dengan menggunakan beberapa parameter. Berdasarkan dari Namanya yaitu *convolutional neural network* yang menunjukkan bahwa metode jaringan tersebut menggunakan operasi dari matematika yang disebut sebagai konvolusi. CNN juga dilatih untuk mempelajari fitur dari objek untuk dapat memprediksinya (Jinan & Hayadi, 2022).

Penelitian terkait mengenai klasifikasi citra penyakit padi sudah banyak dilakukan sebelumnya, seperti (Swathika dkk., 2021) dalam penelitiannya mengembangkan sebuah sistem identifikasi penyakit pada daun tanaman padi menggunakan CNN *based Deep Learning*. Dataset yang digunakan adalah 3081 gambar padi dengan penyakit dan padi dalam keadaan normal. Dataset tersebut dikumpulkan dari dataset public seperti Kaggle. Dari 3081 gambar, 2772 (90%) gambar dipakai untuk proses *training* dan 10% digunakan untuk *testing* model yang dikembangkan dalam 10 epochs. Setelah tahap *tuning* dan *optimizing*, akurasi yang didapatkan adalah sebesar 70%. Penelitian dari Kawcher Ahmed, Tasmia Rahman Shahidi, Syed Md. Irfanul Alam dan Sifat Momen adalah mendeteksi penyakit pada daun tanaman padi. Penelitian ini menggunakan teknik machine learning untuk mengembangkan model. Metode yang digunakan adalah *supervised classification algorithms* diantaranya *K-Nearest Neighbor* (KNN), *Decision Tree, Naïve Bayes,* dan *Logictic Regression*. Data yang digunakan adalah data yang dibuat secara manual dengan memisahkan daun dengan penyakit menjadi tiga kelas penyakit yang berbeda, yaitu *bacterial leaf blight*, *brown spot*, dan *leaf smut*. Peneliti menggunakan teknik *10-fold cross validation* pada dataset yang ada. Akurasi yang didapatkan pada akhirnya bermacam-macam di antaranya *Logistic Regression* sebesar 75.463%, KKN (K=1) 98.8462%, *Decision Tree* 94.9074%dan *Naïve Bayes* 58.7963%.

Berdasarkan masalah klasifikasi hama dan tanaman padi , adanya penelitian untuk membuat sistem klasifikasi hama dan penyakit pada tanaman padi menggunakan *Convolutional Neural Network.* Sistem klasifikasi menggunakan *public dataset* dari Kaggle serta data primer dari Dinas Pertanian Provinsi Gorontalo. Jumlah klasifikasi hama dan penyakit sebanyak 7 kelas yaitu Penggerek Batang Padi Putih (PBPP), Ulat Grayak, Wereng Batang Coklat (Wbc), Tikus, Hama Putih Palsu (HPP), Kepinding Tanah, Walang Sangit, Hawar Daun Bakteri/Kresek, dan *Blast*.

## 1.2. Rumusan Masalah

Bagaimana klasifikasi hama dan penyakit pada tanaman padi menggunakan pengolahan citra digital dengan metode Convolutional Neural Network ?

## 1.3. Ruang Lingkup Penelitian

1. Metode yang digunakan dalam merancang sistem adalah Convolutional Neural Network (CNN).
2. Klasifikasi digunakan dalam 5 kelas, yaitu normal, penggerek batang, hama putih palsu (HPP), hawar daun bakteri/kresek. dan *blast.*
3. Dataset berasal dari *Kaggle.com* dan Dinas Pertanian Provinsi Gorontalo.
4. Format file citra yang digunakan adalah \*.jpeg

## 1.4. Tujuan Penelitian

Merancang sistem untuk klasifikasi hama dan penyakit pada tanaman padi menggunakan olah citra digital dengan metode CNN.

## 1.5. Manfaat Penelitian

1. Membantu Dinas Pertanian Provinsi Gorontalo untuk mengedukasi petani dan orang awam dalam mendeteksi dini hama dan penyakit tanaman padi secara tepat serta akurat.
2. Melalui pengujian ini, peneliti dapat menentukan seberapa baik akurasi model dalam mengidentifikasi objek yang dikenali.
3. Menggunakan arsitektur terbaik untuk klasifikasi hama dan penyakit pada tanaman padi menggunakan metode *Convolutional Neural Network* (CNN).

# BAB 2

# TINJAUAN PUSTAKA

## 2.1. Tanaman Padi

Padi merupakan salah satu tanaman budidaya terpenting dalam peradaban manusia. Masyarakat Indonesia sangat membutuhkan hasil tanaman padi sebagai pangan utama sehari-hari. Ada beberapa tanaman budidaya yang dimanfaatkan masyarakat Indonesia sebagai bahan pangan seperti sagu dan umbi-umbian(Ayunina, 2023).

Tanaman padi memerlukan waktu 3-6 bulan untuk tumbuh mulai dari pembenihan sampai panen, tergantung varietas dan kondisi lingkungan tempat tumbuh padi tersebut(Warti, 2006). Pertumbuhan padi dapat dikategorikan menjadi 3 tahap yaitu: vegetatif, reproduktif dan pemasakan.

Pertumbuhan padi tidak dapat dipisahkan dari hambatan dan serangan hama maupun penyakit tanaman padi. Berbagai cara dilakukan oleh pemerintah dan lembaga terkait untuk menurunkan faktor sukses panen dari padi di Indonesia. Akan tetapi, faktor hama dan penyakit masih terus mempengaruhi kehilangan hasil. Hama dan penyakit tersebut di antaranya adalah penggerek batang, kepinding tanah, keong sawah, penyakit blas, kresek, bercak coklat dan pucuk putih (Ayunina, 2023). Faktor abiotik juga berperan seperti curah hujan, irigasi, serta perubahan dan anomali iklim (Yulistiani dkk., 2018).

## 2.2. Hama dan Penyakit Tanaman Padi

Hama serta penyakit masih menjadi kendala bagi petani karena di setiap musim tanam terjadi ledakan hama pada budidaya padi. Beberapa hama utama yang berpotensi merusak tanaman padi adalah penggerek batang, ulat grayak, tikus, kepinding tanah dan walang sangit. Hal serupa yang sering terjadi, penyakit merupakan salah satu kendala utama produksi berbagai tanaman pertanian. Banyak kerugian yang diakibatkan karena penyakit tanaman yang terlambat untuk didiagnosis dan sudah mencapai tahap yang parah dan menyebabkan gagal panen. Beberapa penyakit yang sangat berpotensi merusak tanaman padi adalah hawar daun, blas, dan wereng coklat (Dypanuryansyah, 2018).

1. Penggerek Batang.

Penggerek batang termasuk hama paling penting pada tanaman padi yang sering menimbulkan kerusakan berat dan kehilangan hasil yang tinggi. Di lapangan, keberadaan hama ini ditandai oleh kehadiran ngengat (kupu-kupu), kematian tunas-tunas padi (sundep, *dead heart*), kematian malai (beluk, *white head*), dan ulat (*larva*) penggerek batang (Puslitbang Tanaman Pangan dkk., 2007).

Hama ini dapat merusak tanaman pada semua fase tumbuh, baik pada saat pembibitan, fase anakan, maupun fase berbunga. Gejala tanaman yang terserang hama penggerek batang adalah tanaman muda, daun yang tampak merah kekuningan, pucuk tanaman layu mengering, jika dicabut mudah patah. Jika batang padi dibuka, terlihat Larva di dalam batang. Serangan ini disebut sundep. Pada tanaman generatif Larva memotong tangkai buah (pangkal malai) sehingga gabah menjadi hampa (Murad, Mappa, & Rudin, 2021).

1. Hawar daun Bakteri (*bacterial leaf blight* – BLB).

Hawar daun bakteri (HDB) merupakan penyakit bakteri yang tersebar luas dan menurunkan hasil sampai 36%. Penyakit terjadi pada musim hujan atau musim kemarau yang basah, terutama pada lahan sawah yang selalu tergenang, dan dipupuk N tinggi (>= 250 kg urea/ha) (Puslitbang Tanaman Pangan dkk., 2007).

Penyakit HDB menghasilkan dua gejala khas, yaitu kresek dan hawar. Kresek adalah gejala yang terjadi pada tanaman berumur <30 hari (pesemaian atau yang baru dipindah). Daun-daun berwarna hijau kelabu, melipat, dan menggulung. Dalam keadaan parah, seluruh daun menggulung, layu dan mati, mirip tanaman yang terserang penggerek batang atau terkena air panas (*lodoh*). Sementara itu, hawar merupakan gejala yang paling umum dijumpai pada pertanaman yang telah mencapai fase tumbuh anakan sampai fase pemasakan (Wijayanto, Kiswanto, & Manurung, 2013).

1. Blas (Blast)

Semula penyakit blas dikenal sebagai salah satu kendala utama pada padi gogo, tetapi sejak akhir 1980-an, penyakit ini juga sudah terdapat pada padi sawah beririgasi. Penyakit yang mampu menurunkan hasil yang sangat besar ini disebabkan oleh jamur pathogen *Pyricularia grisae* (Puslitbang Tanaman Pangan dkk., 2007).

Penyakit blas menimbulkan dua gejala khas, yaitu blas daun dan blas leher. Blas daun merupakan bercak coklat kehitaman, berbentuk belah ketupat, dengan pusat bercak berwarna putih. Sementara itu, blas leher berupa bercak coklat kehitaman pada pangkal leher yang dapat mengakibatkan leher malai tidak mampu menopang malai dan patah. Kemampuan patogen membentuk strain dengan cepat menyebabkan pengendalian penyakit ini sangat sulit (Murad dkk., 2021).

1. Hama Putih Palsu (*leaffolder)*.

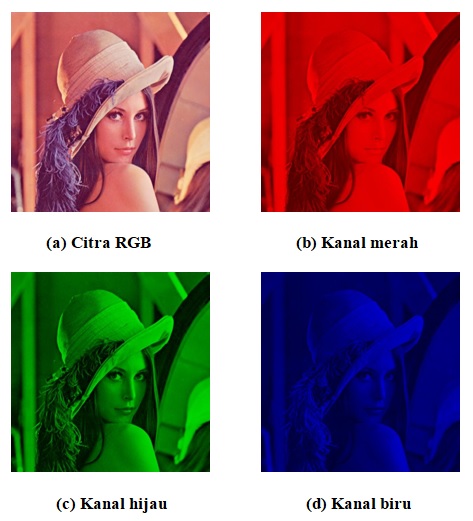
Hama Putih Palsu (HPP) jarang menjadi hama utama padi. Serangannya menjadi berarti bila kerusakan pada daun pada fase anakan maksimum dan fase pematangan mencapai >= 50%. Kerusakan akibat serangan larva HPP terlihat dengan adanya warna putih pada daun di pertanaman (Puslitbang Tanaman Pangan dkk., 2007).

Larva memakan jaringan hijau daun dari dalam lipatan daun meninggalkan permukaan bawah daun yang berwarna putih. Siklus hidup hama ini 30-60 hari. Tanda pertama adanya infestasi HPP adalah kehadiran ngengat berwarna kuning coklat yang memiliki 3 buah pita hitam dengan garis lengkap atau terputus pada bagian sayap depan (Murad dkk., 2021).

## 2.3. Citra Digital

Sebuah citra digital dapat didefinisikan oleh sebuah matriks dua dimensi f(x,y) yang terdiri dari M kolom dan N baris, dimana perpotongan antara kolom dan baris disebut piksel atau elemen terkecil dari sebuah citra. Citra digital merupakan citra yang bisa diproses oleh komputer, dan citra yang dihasilkan oleh perangkat digital bisa secara langsung diproses oleh komputer. Citra analog tak dapat direpresentasikan oleh komputer dan menyebabkan tidak diprosesnya citra analog tersebut. Untuk dapat diproses oleh komputer citra analog harus dikonversikan ke dalam citra digital (Dewi, 2018).

Pada ruang warna RGB ini masing-masing piksel memiliki warna tertentu yaitu merah (*Red*), hijau(*Green*), dan biru(*Blue)*. Ruang warna RGB ini terdiri dari tiga matriks yang mewakili nilai-nilai merah, hijau, dan biru setiap pikselnya.



Pada Gambar di atas, terdapat citra warna 8-bit. Lalu sebagai contoh, sebuah vector r = (x,y,z). Kemudian untuk warna, variabel-variabel tersebut diganti oleh komponen R (*Red*), G(*Green*), B(*Blue*). Setiap warna dasar memiliki nilai minimum nol (0) dan nilai maksimum nol (0) dan nilai maksimum sebesar 255 bit (Rizky, 2023).

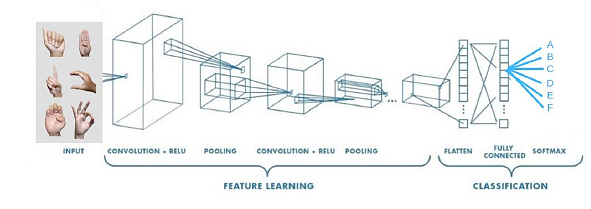
## 2.4. Deep Learning

*Deep learning* adalah bagian dari *Machine Learning* yang berusaha untuk memprediksi beberapa lapisan implementasi menjadi *output* gambar. Kinerja saraf dalam otak manusia menginspirasi sistematika kerja dari *Deep learning*. Teknik tingkat tinggi dari *Deep learning* memungkinkan sistem mengenali dan mengklasifikasikan gambar, animasi, serta suara(Rao Moparthi dkk., 2021). *Deep learning* biasanya digunakan untuk mengatasi masalah mulai dari kumpulan data berdimensi rendah hingga tinggi. Secara teknis *Deep learning* dapat diartikan teknik dalam *neural network* atau jaringan saraf tiruan (JST) dengan percepatan proses pembelajaran JST tersebut menggunakan banyak lapisan jaringan. Dengan menggunakan *Deep learning*, waktu yang dibutuhkan untuk training akan semakin sedikit karena masalah hilangnya gradien pada propagasi balik akan semakin rendah. *Convolutional Neural Network* (CNN) adalah arsitektur pembelajaran mendalam yang paling umum digunakan dalam banyak metode pola.

## 2.5. Convolutional Neural Network

*Convolutional Neural Network* (CNN) merupakan variasi dari *Multilayer Perceptron* (MLP) yang digunakan untuk menganalisis data dua dimensi berupa gambar. Karena kedalaman jaringannya yang besar dan aplikasi yang luas untuk menggambarkan data, CNN diklasifikasikan sebagai *Deep Neural Network*. Konvolusi adalah matriks dengan fungsi memfilter citra. CNN adalah arsitektur yang dapat dilatih dalam beberapa tahap. Input dan output dari setiap tahapan adalah beberapa array yang disebut peta fitur. Keluaran setiap tahapan berupa peta fitur hasil olahan dari semua lokasi masukan.

Proses tersebut adalah proses *training*. Proses pembentukannya terdiri dari 3 lapisan yaitu, *convolutional layer*, *activation layer*, dan *pooling layer*. Lapisan konvolusi (*convolutional layer*) adalah tingkat pertama dalam arsitektur CNN. Tahap ini diakhiri dengan memanfaatkan kernel dalam ukuran tertentu. Jumlah kernel yang digunakan ditentukan oleh jumlah fitur yang dibuat. Fungsi aktivasi (*activation layer*) umumnya merupakan fungsi aktivasi ReLU (*Rectifier Linear Unit*), mengikuti tahap kedua. Lapisan penyatuan (*pooling layer*) digunakan pada langkah ketiga dan prosedur ini diulang sampai peta fitur yang memadai diperoleh untuk melanjutkan ke jaringan saraf yang sepenuhnya terhubung. *Fully connected network* merupakan *output class*(Khoiruddin dkk., 2022).



1. Convolutional Layer.

*Convolutional layer* adalah lapisan pertama yang menerima input gambar langsung ke arsitektur. Operasi dalam lapisan Convolutional Layer identic dengan operasi konvolusi, yang terdiri dari operasi filter kombinasi linier pada wilayah lokal gambar. Filter mewakili bidang reseptif neuron yang terkait dengan wilayah lokal dari gambar input. Sebuah lapisan konvolusi melakukan operasi konvolusi pada output dari lapisan sebelumnya. Lapisan ini merupakan proses utama yang mendasari CNN. Tujuan dari citra yang berbelit-belit adalah untuk mengekstrak fitur dari citra masukan. Secara umum, operasi konvolusi dapat didefinisikan dengan menggunakan rumus:

………………….(1)

Pada fungsi s(t) menghasilkan output tunggal yaitu *feature map*, ketika argumen pertama berupa input yang merupakan x dan argument kedua yang merupakan w sebagai kernel atau filter. Jika melihat input sebagai citra dua dimensi, maka (t) bisa diasumsikan sebagai sebuah piksel(Sheila dkk., 2023).

1. Aktivasi Rectified Linear Units (ReLU).

Aktivasi *Rectified Linear Units* (ReLU) adalah prosedur untuk memperkenalkan nonlinear dan meningkatkan representasi model. ReLU mengubah semua nilai piksel negative menjadi nol, memungkinkan *feature map* digambarkan lebih akurat dalam kaitannya dengan gambar. Persamaan dibawah ini merupakan aktivasi untuk ReLU :

……………………(2)

Dengan input sama dengan output ketika *x* positif and *output* sama dengan nol ketika x negatif (Rizky dkk., 2023).

1. Pooling layer.

*Pooling layer* mencakup filter dengan ukuran dan *stride* tertentu. Tiap pergeseran ditentukan oleh jumlah *stride* yang digeser pada seluruh bagian *feature map* atau *activation map*. *Pooling layer* yang umum digunakan merupakan *max pooling* dan *average pooling*. Fungsi dari *pooling layer* adalah untuk mereduksi *input* secara spasial (mengurangi jumlah parameter) dengan operasi *down-sampling*. Gambar di bawah merupakan ilustrasi dari *pooling layer*(Anam Alidrus dkk., 2021).

1. Fully-connected layer

Semua neuron aktivasi dari lapisan sebelumnya terhubung ke neuron di lapisan berikutnya disebut dengan *fully-connected layer*. Setelah *convolutional layer* dan *pooling layer, fully Connected Layer* akan diterapkan. *Output* dari prosedur *Convolution Layer* dan *Pooling Layer* adalah *feature map*, yang kemudian diratakan dan dimasukkan ke dalam *fully connected layer*.. Lapisan Linear termasuk dalam lapisan ini untuk menentukan nilai setiap kelas. Lapisan ini juga memiliki fungsi aktivasi, yaitu *LogSoftmax*, yang digunakan untuk menentukan probabilitas setiap kelas, dengan kemungkinan kerugian *negative log-likelihood* yang digunakan dalam perhitungan *loss*. Persamaan *softmax* dapat dilihat pada persamaan berikut.

……………………………..(3)

Dengan 𝑓𝑗 adalah probabilitas setiap kelas, 𝑧𝑗 adalah nilai vector ke-𝑗, dan 𝑘 adalah adalah jumlah kelas (Anam Alidrus dkk., 2021).

## 2.6. Tensorflow

Tensorflow adalah *library* perangkat lunak yang dikembangkan oleh Tim Google Brain dalam organisasi penelitian Mesin Cerdas Google, untuk tujuan pembelajaran mesin dan penelitian jaringan syaraf dalam. Tensorflow memiliki perkakas, pustaka, dan sumber daya komunitas yang komprehensif serta fleksibel yang dapat memudahkan jejaring serta peneliti dalam mengembangkan *machine learning* serta tim pengembang dapat dengan mudah membangun dan menerapkan aplikasi yang didukung algoritma *machine learning*. Tensorflow menggabungkan aljabar komputasi dan teknik pengoptimalan kompilasi, mempermudah penghitungan dengan mempersingkat waktu dalam melakukan perhitungan (Dewi, 2018).

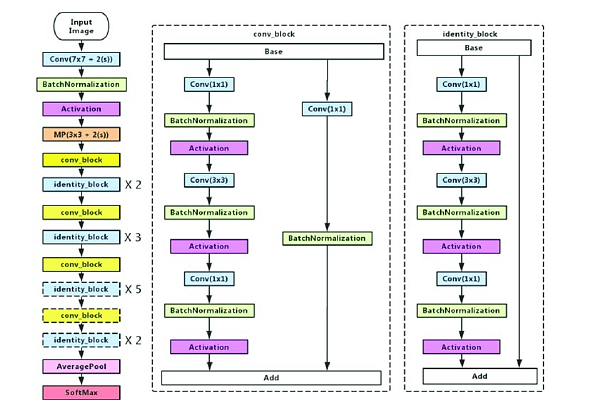
Berikut adalah fitur utama dari Tensorflow.

1. Mendefinisikan, mengoptimalkan, dan menghitung secara efisien ekspresi matematis yang melibatkan *array multidimension* (tensors)
2. Pemrograman pendukung jaringan syaraf dalam dan teknik pembelajaran mesin.
3. Penggunaan GPU yang transparan, mengotomatisasi manajemen dan optimalisasi memori yang sama dan data yang digunakan. Tensorflow bisa menulis kode yang sama dan menjalankannya dengan baik di CPU, GPU, atau pun TPU.
4. Skalabilitas komputasi yang tinggi di seluruh mesin dan kumpulan data yang besar.

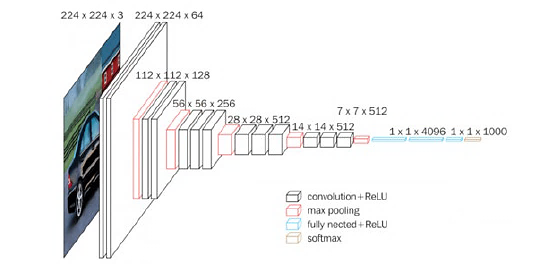
Tensorflow menyediakan *Object Detection API* untuk mempermudah pengembangan aplikasi *deep learning* untuk membuat perangkat lunak pendeteksian objek. TensorFlow *Object Detection API* adalah *open source framework* yang dapat digunakan untuk mengembangkan, melatih, dan menggunakan model deteksi objek. Sistem tersebut sudah banyak diterapkan pada berbagai produk Google seperti pencarian gambar, deteksi wajah dan plat nomor kendaraan pada Google Streetview, Google Assistant, dan lainnya (Sheila dkk., 2023).

Tensorflow *Object Detection* juga menyediakan banyak model yang telah dilatih dengan dataset *Common Object in Context* (COCO) 2017 yang tediri dari 300.000 gambar berisi 90 kategori objek. Model-model tersebut memiliki kecepatan dan tingkat akurasi masing-masing, sesuai dengan arsitektur modelnya. Model tersebut diantaranya adalah model SSD ResNet50 V1 FPN 640x640, VGG-16, dan AlexNet.

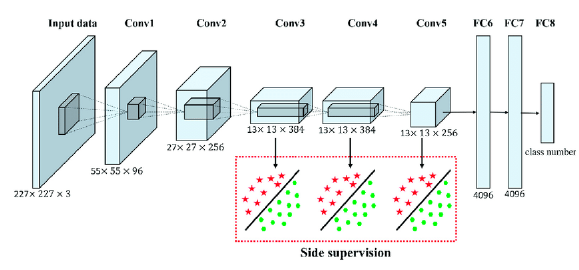
SSD ResNet50 V1 FPN 640x640 merupakan arsitektur deteksi objek FPN SSD berdasarkan ResNet-50. ResNet-50 adalah *convolution neural network* (CNN) yang memiliki kedalaman 50 lapisan, ResNet merupakan kependekan dari *residual network*. Setiap unit *residual* dalam jaringan terdiri dari dua lapisan konvolusi tanpa lapisan penyatuan, normalisasi *batch*, aktivasi ReLU, dan kernel 3x3. Satu tahapan pada model tersebut menggunakan *Feature Pyramid Network* (FPN) *back-bone* pada arsitekturnya. Input dari model SSD ResNet50 V1 FPN 640x640 berupa tensor berukuran 1x640x640x3, ukuran tersebut menunjukan nilai batch size berjumlah 1 dengan gambar RGB berukuran 640x640. Output dari model berupa nama kelas objek, nilai probabilitas, koordinat kotak deteksi, dan jumlah kotak deteksi yang diprediksi(Dewi, 2018).



Arsitektur model selanjutnya, VGG-16 merupakan salah satu arsitektur CNN yang diperkenalkan oleh *Visual Geometry Group* (VGG) dari Universitas Oxford yang berpartisipasi dalam ILSVRC pada tahun 2014. VGG terdiri dari 16 *hidden layer*, konvolusi dengan filter 3x3, *strides =* 2,2, *padding* yang sama, dan *max pooling* 2x2. Setiap Citra yang melewati *max pooling* fitur ukurannya akan direduksi setengah dan filter akan berjumlah 64 bertambah sebesar kelipatan dua dari lapisan konvolusi pertama, hingga blok terakhir VGG-16 (Rizky dkk., 2023).



Berbeda dari arsitektur model sebelumnya, AlexNet terbagi atas dua kelompok layer. Layer pertama adalah *extraction layer* yang terdiri dari *convolutional layer, pooling layer,* dan ReLU yang tersusun secara berselang-seling. AlexNet dikembangkan oleh Alex Krizhevsky, Ilya Sutskever, dan Geoffrey Hinton. Pengembangan AlexNet menghasilkan arsitektur model yang terdiri atas lima *convolution layer*, tiga *pooling layer*, dan lima ReLU. *Extraction layer* terletak pada bagian awal arsitektur sehingga langsung menerima gambar dan *output* dari *extraction layer* yang berupa gambar vektor. *Layer* kedua adalah *classification layer*, yang meliputi *fully connected layer*, ReLU dan *softmax* (Swathika dkk., 2021). Tersusun secara berselang-seling antara tiga *fully connected layer*, dan dua ReLU dan diakhiri oleh *softmax*. Mendapatkan input berupa gambar vektor yang merupakan hasil dari pengolahan pada *extraction layer* dan output dari *classification layer* hasil dari keseluruhan pemrosesan berupa citra yang telah terklasifikasi.



## 2.7. Penelitian Terkait

a. Hasil Penelitian Rizal Amegia Saputra, dkk (2021)

“Penerapan Algoritma Convolutional Neural Network dan Arsitektur MobileNet Pada Aplikasi Deteksi Penyakit Daun Padi” oleh Rizal Amegia Saputra, Sri Wasyianti, Adi Supriyatna, Dede Firmansyah Saefudin dari Universitas Bina Sarana Informatika dalam Jurnal Swabumi. Penelitian ini bertujuan untuk melakukan ekstrak fitur penyakit daun padi dengan algoritma Neural Network, penelitian ini menunjukkan bahwa algoritma Neural Network dapat digunakan dalam klasifikasi penyakit daun padi melalui aplikasi berbasis Android.

Dalam penelitian ini, peneliti mengembangkan sebuah model klasifikasi penyakit daun padi dengan algoritma CNN dan arsitektur MobileNet dengan *feature extraction* dan *feature map*. Peneliti menggunakan dataset yang diambil dari UCI repository dengan objek data terdiri dari tiga jenis penyakit yaitu *bacterial leaf blight, brown spot,* serta *leaf smut*. Proses *training* dan *tuning* dari model yang dikembangkan yaitu sebanyak 100 epochs.

Hasil dari penelitian ini menunjukkan aplikasi yang dikembangkan mampu mengklasifikasi gambar jenis penyakit daun padi menggunakan algoritma CNN dengan arsitektur MobileNetV1 dan *Feature Extraction* memiliki akurasi yang tinggi yaitu sebesar 92% namun dengan jumlah data cukup kecil dan penggunaan *epochs* sebanyak 100, mengakibatkan nilai validasi kesalahan lebih tinggi dibandingkan nilai kesalahan pada proses pelatihan, sehingga akan terjadi *overfitting*.

Adapun perbedaannya terdapat pada objek klasifikasi. Penelitian ini terfokus pada tiga penyakit daun padi yaitu *bacterial leaf blight, brown spot,* serta *leaf smut.* Selain itu, peneliti menggunakan arsitektur model MobileNetV1 dan mengambil data dari *public dataset*.

b. Hasil Penelitian Annida Purnamawati, dkk (2020)

“Deteksi Penyakit Daun Padi pada Tanaman Padi menggunakan Algoritma Decision Tree, Random Forest, Naïve Bayes, SVM dan KNN” oleh Annida Purnamawati, Wawan Nugroho, Destiana Putri, dan Wahyutama Fitri Hidayat dari STMIK Nusa Mandiri dan Universitas Bina Sarana Informatika, Jurusan Ilmu Komputer dalam Jurnal Nasional Informatika dan Teknologi Jaringan. Penelitian ini bertujuan untuk melakukan *Machine learning classifier*. Untuk dapat melihat klasifikasi dari jenis penyakit maka peneliti mencari akurasi dengan menggunakan perbandingan berbagai macam metode algoritma yaitu dengan Decision Tree, Random Forest, Naïve Bayes, SVM dan KNN.

Dalam penelitian ini, peneliti mengusulkan perbandingan metode klasifikasi dalam dataset Rice Leaf Diseases yang terbagi atas tiga kelas yaitu *bacterial leaf blight*, *brown spot*, dan *leaf smut*. Peneliti menggunakan analisis dengan data kualitatif dimana merupakan suatu Analisa yang digunakan apabila angka-angka dan perhitungan menggunakan rumus yang ada hubungannya dengan penelitian.

Hasil dari penelitian ini menunjukkan bahwa pada pengembangan *machine learning classifier* dengan metode algoritma Random Forest, Decision Tree, dan Naïve Bayes terjadi *overfitting*. Sementara itu, untuk pengembangan dengan metode algoritma SVM terjadi *underfitting*. *Good models* didapatkan dengan menggunakan algoritma KNN dengan nilai akurasi sebesar 87%.

Adapun perbedaannya terdapat pada metode penelitian. Penelitian ini menggunakan lima metode sekaligus untuk dilakukan perbandingan yaitu Random Forest, Decision Tree, Naïve Bayes, SVM serta KNN.

c. Hasil Penelitian Kawcher Ahmed, dkk. (2021)

Penelitian dari Kawcher Ahmed, Tasmia Rahman Shahidi, Syed Md. Irfanul Alam dan Sifat Momen adalah mendeteksi penyakit pada daun tanaman padi. Penelitian ini menggunakan teknik machine learning untuk mengembangkan model. Metode yang digunakan adalah *supervised classification algorithms* diantaranya *K-Nearest Neighbor* (KNN), *Decision Tree, Naïve Bayes,* dan *Logictic Regression*. Data yang digunakan adalah data yang dibuat secara manual dengan memisahkan daun dengan penyakit menjadi tiga kelas penyakit yang berbeda, yaitu *bacterial leaf blight*, *brown spot*, dan *leaf smut*. Peneliti menggunakan teknik *10-fold cross validation* pada dataset yang ada. Akurasi yang didapatkan pada akhirnya bermacam-macam di antaranya *Logistic Regression* sebesar 75.463%, KKN (K=1) 98.8462%, *Decision Tree* 94.9074%dan *Naïve Bayes* 58.7963%.

# BAB 3

# METODOLOGI PENELITIAN

## Lokasi dan Waktu Penelitian

1. Lokasi Penelitian

Penelitian ini berlokasi di Provinsi Gorontalo, khususnya di lahan tanaman padi dibawah pembinaan Dinas Provinsi Gorontalo.

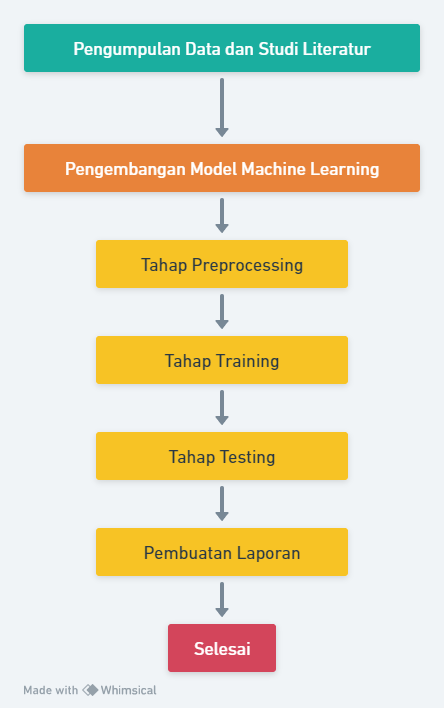
1. Waktu Penelitian

Waktu Penelitian ini dilakukan selama 6 bulan, terhitung dari bulan Agustus 2023 sampai dengan Januari 2024.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No | Tahap Penelitian | Tahun 2023-2024 | | | | | |
| Bulan | | | | | |
| Agust | Sept | Okt | Nov | Des | Jan |
| 1 | Observasi, Pengumpulan Data Tahap Awal dan Studi Literatur |  |  |  |  |  |  |
| 2 | Pengumpulan Dataset Citra |  |  |  |  |  |  |
| 3 | Pengembangan Sistem : Tahap *Preprocessing* dan Tahap *Training* |  |  |  |  |  |  |
| 4 | Tahap *Testing* |  |  |  |  |  |  |
| 5 | Pembuatan Laporan |  |  |  |  |  |  |

## Rancangan Penelitian

Rancangan Penelitian adalah sebagai model pendekatan penelitian sekaligus merupakan rancangan analisis data. Rancangan penelitian menggunakan model pengembangan sistem *machine learning,* dapat dilihat pada bagan di bawah ini:



1. Pengumpulan Dataset dan Studi Literatur

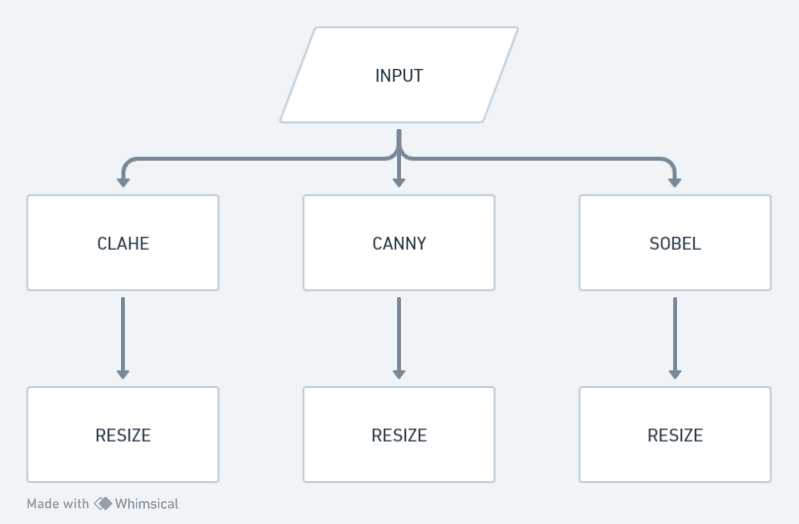
Pada tahap ini, peneliti mengumpulkan data yang diperlukan berdasarkan permasalahan yang diangkat. Data yang dikumpulkan menjadi acuan dalam membangun model klasifikasi hama dan penyakit tanaman padi. Data diperoleh dari pelaksanaan observasi, wawancara, *public dataset*, sumber bacaan dan pengumpulan data gambar secara langsung di lapangan.

1. Pengembangan Model Machine Learning

Adapun alur proses yang akan digunakan peneliti dalam perancangan model klasifikasi hama dan penyakit tanaman padi adalah metode pengembangan model *machine learning*.

1. Tahap *Preprocessing*

Pada *preprocessing* ini merupakan tahap pertama dari pengolahan citra sebelum citra dikirim ke tahap pelatihan. *Preprocessing* digunakan untuk meningkatkan kualitas data yang dikumpulkan. CLAHE, Sobel Edge Detection, Conny Edge Detection, dan Resize adalah teknik *preprocessing* yang digunakan pada penelitian ini. Berikut merupakan alur kerja pada tahapan *preprocessing*.



1. Tahap *Training*

Pada tahapan ini, sistem akan melakukan pengenalan terhadap citra hama dan penyakit tanaman padi untuk diklasifikasikan sesuai dengan kelasnya. Proses klasifikasi menggunakan TensorFlow melalui dua tahap yaitu *training* dan *validation*. Tahap *training* dilakukan untuk mencari parameter terbaik pada citra yang kemudian dijadikan acuan untuk proses klasifikasi.

1. Tahap *Testing*

Pada tahap ini akan dilakukan pengujian oleh model sistem yang telah dibuat dan dilatih sebelumnya. Performasi sistem akan dinilai dari parameter nilai akurasi, *precision*, *recall*, *fl-score*, dan *loss* dari model-model yang telah dihasilkan oleh data pengujian. Peneliti akan melakukan tiga tahapan pengujian performansi sistem terhadap jumlah *hidden layer*, *optimizer*, dan *learning rate*.Is

## Data Penelitian

* 1. Sumber Data

Sumber data dari penelitian ini diambil dari *Public Dataset* dan Dinas Pertanian Provinsi Gorontalo.

* 1. Jenis Data
     + 1. Data Primer

Data primer dalam penelitian ini diperoleh dari hasil koordinasi dan komunikasi dengan Dinas Pertanian Provinsi Gorontalo.

* + - 1. Data Sekunder

Data sekunder dalam penelitian ini didapatkan dari penelitian-peneltian ini yang memiliki topik permasalahan yang sama dengan penelitian ini. Penelitian tersebut menjadi literatur sekaligus bahan bagi peneliti untuk mempelajari proses pembangunan sistem model machine learning untuk klasifikasi hama dan penyakit tanaman padi.

* 1. Teknik Pengumpulan Data

Adapun teknik pengumpulan data yaitu sebagai berikut:

* + - 1. Wawancara

Pada tahap ini peneliti melakukan wawancara dengan pihak terkait yaitu Pengendali Organisme Pengganggu Tanaman (POPT), tenaga pertanian dari Dinas Pertanian Provinsi Gorontalo, dan petani padi.

* + - 1. Studi Pustaka

Peneliti mempelajari dan mengaitkan literatur yang berhubungan dengan permasalahan yang dihadapi yaitu literatur tentang mendeteksi dan klasifikasi hama dan penyakit tanaman padi. Langkah ini dilakukan sebagai landasan teori serta pedoman dalam menganalisa masalah.

* + - 1. Observasi

Teknik ini dilakukan dengan mengumpulkan data dan informasi mengenai hama dan penyakit tanaman padi. Observasi dilakukan dengan melakukan peninjauan langsung pada lahan yang terdampak OPT padi.

* + - 1. Dokumentasi

Peneliti mengambil sejumlah besar fakta dan data tersimpan dalam bentuk dokumentasi, yaitu gambar. Tujuan dari pengambilan gambar adalah untuk mendapatkan dataset yang kemudian akan dilatih dalam pengembangan model.

# BAB 4

# HASIL DAN PEMBAHASAN

## Hasil Penelitian

Bab ini berisi tentang hasil data dari proses penelitian klasifikasi hama dan penyakit tanaman padi dengan menggunakan metode *convolutional neural network* dengan arsitektur

* + 1. Pengumpulan Data dan Studi Literatur

Data yang digunakan dalam membangun model CNN merupakan data gambar tanaman padi mulai dari batang dan daun. Data ini diperoleh dari 4 dataset yaitu *Paddy Doctor Dataset* (KUTIPAN), *Public Dataset* dari RoboFlow, Sistem Informasi Data OPT & DPT Direktorat Perlindungan Tanaman Pangan, dan dataset hasil observasi di lapangan. Mekanisme pengumpulan Dataset pertama yaitu mengunduh dataset dari situs Kaggle. Dataset yang kedua diunduh dari situs Roboflow. Sementara itu dataset yang ketiga di-*scraping* dari situs atau sistem informasi resmi Direktorat Perlindungan Tanaman Pangan melalui akun Dinas Pertanian Provinsi Gorontalo. Data piksel dari kumpulan gambar tersebut akan digunakan sebagai *input* pada model CNN. Ketiga dataset yang telah terkumpul tersebut digabungkan menjadi sebuah dataset baru. Terdapat beberapa masalah yang ditemukan saat proses penggabungan ini dilakukan seperti format data gambar yang berbeda, gambar yang terduplikasi, data gambar yang rusak dan tidak bisa dibaca, serta data gambar yang belum dilabeli.

Pada dataset *Paddy Doctor Dataset*, terdapat sekitar 13.900 gambar, yang tentunya tidak hanya terdiri dari gambar beberapa kelas hama dan penyakit tanaman padi. Data pada dataset ini memiliki format JPG. Dataset yang diperlukan adalah 5 kelas yaitu normal*,* penggerek batang, hama putih palsu, hawar daun dan blas. Oleh karena itu, data yang diambil dari dataset ini adalah data gambar tanaman padi normal, hama putih palsu, dan blas. Sementara itu, untuk dataset yang diunduh dari situs Roboflow, terdapat sekitar 26.661 gambar yang terdiri dari beberapa label/kelas. Data pada dataset ini memiliki format JPG. Kemudian data yang diambil dari dataset ini adalah data dengan kelas hama putih palsu. Dataset yang diambil dari Sistem Informasi Data OPT & DPT ada sekitar 1.388 data gambar dengan kelas yang sudah sesuai dengan kriteria. Format data gambar dalam dataset ini adalah JPG dan PNG. Akan tetapi, setelah difilter data PNG tergolong sangat sedikit dan diputuskan untuk dikeluarkan dari dataset. Sumber data terakhir yaitu dengan pengumpulan secara langsung di daerah terserang hama/penyakit. Data yang terkumpul yaitu sekitar 500 data gambar.

Seluruh gambar dari keempat dataset tersebut terdiri dari 3 (tiga) kanal warna (*color channel*) yaitu RGB atau *Red*, *Green*, *Blue* (Merah, Hijau, Biru). Di samping itu, gambar-gambar di seluruh dataset tersebut memiliki orientasi gambar yang beragam (vertikal, horizontal, diagonal) dengan berbagai ukuran gambar. Keseluruhan dataset tersebut difilter menjadi 5 kelas yang menjadi fokus utama yaitu normal, penggerek batang (*stem borer*), hama putih palsu (*leaffolder*), hawar daun (*bacterial leaf blight*), dan blas (*blast*). Berikut merupakan contoh gambar dari masing-masing kelas tersebut.



Gambar 1. Tanaman Normal



Gambar 2 Penggerek Batang (Stem Borer)



Gambar 3 Hama Putih Palsu (Leaffolder)



Gambar 4 Hawar Daun (Bacterial Leaf Blight)



Gambar 5 Blas (Blast)

* + 1. Tahap *Preprocessing*

Setelah dataset dikumpulkan dan terdiri dari beberapa kelas yang tidak diperlukan, maka data difilter dan diambil yang sesuai dengan 5 kelas utama. Gambar yang diambil adalah data gambar normal, penggerek batang, hama putih palsu, hawar daun dan blas. Keempat dataset ini kemudian digabungkan berdasarkan kelasnya yang kemudian menghasilkan dataset gabungan baru. Dataset yang duplikat difilter dan dihapus, gambar yang *error* dan tidak sesuai dengan ekstensi dihapus. Adapun dataset gabungan baru memiliki proporsi sebagai berikut.

|  |  |
| --- | --- |
| Nama Kelas | Jumlah Gambar |
| Normal | 1.764 |
| Stem\_borer | 678 |
| Leaffolder | 500 |
| Bacterial\_leaf\_blight | 732 |
| Blas | 1.775 |

Pada tahap *preprocessing* ini dilakukan standarisasi dan normalisasi sebelum data digunakan sebagai masukan (*input*) pada model CNN. *Library* utama yang digunakan pada tahapan ini yaitu TensorFlow dan Keras. Pada penelitian ini, *preprocessing* yang dilakukan yaitu melakukan *resize* dan *rescale* pada setiap gambar agar memiliki ukuran yang sama. *Resize* atau pengubahan ukuran dilakukan saat menyiapkan *input pipeline* sebelum pembangunan model sehingga gambar menjadi berukuran sesuai nilai yang ditetapkan. Sedangkan *rescale* atau penskalaan ulang dilakukan saat mneyusun arsitektur lapisan model (tergabung dalam arsitektur lapisan model).

Untuk meningkatkan diversitas pada data gambar yang digunakan untuk *training*, terdapa sebuah teknik yang digunakan yaitu augmentasi data (*data augmentation*). Teknik augmentasi data bekerja dengan cara menerapkan berbagai jenis transformasi pada data gambar. Teknik ini dapat meningkatkan kinerja dan hasil dari model *machine learning* yang dibangun. Augmentasi yang dilakukan pada dataset yaitu berupa *zoom, rotation*, *shear*, dan *horizontal flip*. Nilai dari setiap parameter tersebut adalah sebagai berikut

|  |  |
| --- | --- |
| Parameter | Nilai |
| Zoom\_range | 0.2 |
| Rotation\_range | 30 |
| Shear\_range | 0.2 |
| Horizontal\_flip | True |

* + 1. Tahap *Training* dan *Testing*

Pada tahapan *training* serta *testing* atau pelatihan model mencakup dua proses yaitu pembentukan model dan menentukan parameter yang terbaik dalam melatih model. Namun, sebelum dilakukan pembangunan model, dataset yang sudah digabungkan dibagi menjadi 2 (dua) bagian yaitu *training* dan *validation*. Untuk memudahkan pembagian ini, dataset gabungan diunggah ke dalam Google Drive dengan folder kelas yang berbeda. Hal ini dilakukan karena Google Drive dapat dipasangkan/dihubungkan dengan Google Colab yang akan memudahkan proses akses data untuk pengembangan model. Adapun versi Python yang digunakan pada Google Colab merupakan Python 3 dan jenis *runtime* yang digunakan bukan merupakan GPU (*Graphics Processing Unit*) atau TPU (*Tensor Processing Unit*). Google Colab yang digunakan adalah versi gratis dan bukan versi yang berbayar. Google Colab dijalankan pada aplikasi Google Chrome.

Berikut spesifikasi Google Colab yang digunakan.

|  |  |
| --- | --- |
| Spesifikasi Google Colab | |
| Arsitektur Sistem | CPU |
| Sistem Operasi | Windows 11 |
| RAM | 12.7 GB |
| Kapasitas Penyimpanan | 107.7 GB |

Setelah *file* baru di Google Colab dibuat, dilakukan pemeriksaan apakah dataset yang diunggah ke Google Drive dapat dibaca atau tidak. Kemudian data tersebut dibagi menjadi dua bagian yaitu *train* dan *validation*. Pembagian ini menggunakan parameter dalam package TensorFlow Keras yaitu *validation split*. Dataset dibagi dengan proporsi 20% *validation* dan 80% *train*.

Model CNN dibangun menggunakan bantuan beberapa *library* dalam bahasa pemograman Python. Adapun *library* yang digunakan pada penelitian kali ini yaitu:

|  |  |
| --- | --- |
| Nama *Library* | Versi |
| Python | 3.10.12 |
| Matplotlib | 3.7.1 |
| Numpy | 1.25.2 |
| TensorFlow | 2.15.0 |
| Keras | 2.15.0 |
| Pathlib | 1.0.1 |

Setelah menyiapkan *library* dan dataset gambar, selanjutnya adalah pembagunan model dengan melakukan beberapa skenario di dalamnya untuk mendapatkan hasil yang paling optimal. Analisis untuk membangun model tersebut ditinjau dari beberapa parameter yaitu *resize*, jenis *optimizer*, nilai *learning rate*, dan *batch size*. Hasil pengujian dari masing-masing parameter divisualisasikan dalam bentuk tabel dan grafik *accuracy* dan *loss*. Kemudian analisis pada keberhasilan sistem diukur menggunakan parameter akurasi, presisi, *recall*, dan *F1-Score*. Dalam pengujian sistem ini memiliki 4 skenario. Berikut adalah hasil dari masing-masing scenario tersebut.

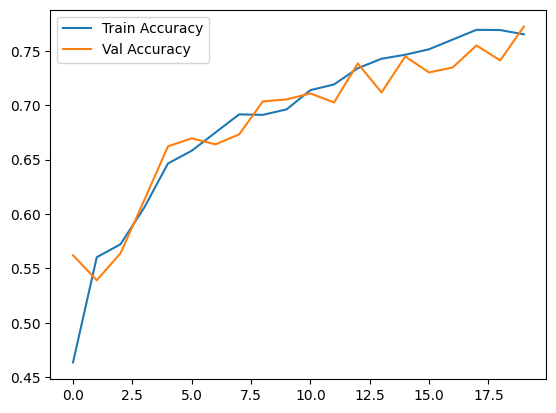
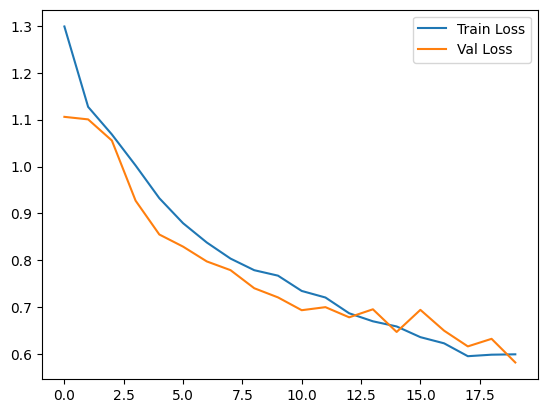
* 1. Pengujian Skenario 1 : *Resize* Citra

Pada skenario pertama dilakukan proses *resize* citra yang akan digunakan pada tahapan ini. Ada 3 ukuran citra yang akan digunakan yaitu 64x64 piksel, 128x128 piksel, dan 150x150 piksel. Dalam skenario ini parameter tetap yang digunakan adalah *optimizer* Adam dengan *epoch* 20 kemudian *learning rate* dengan nilai 0.0001 dan *batch size* dengan nilai 64. Hasil dari pengujian ini ditujukan untuk mengetahui kualitas terbaik pada sistem berdasarkan *image size*.

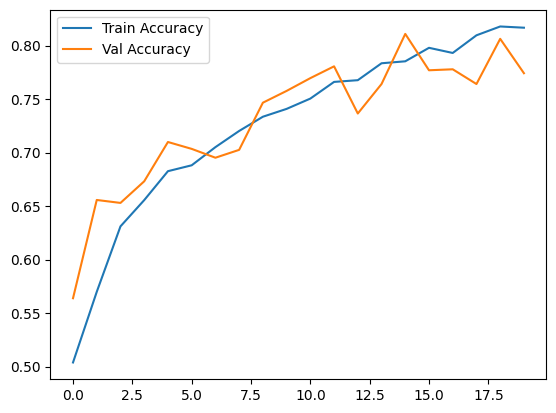
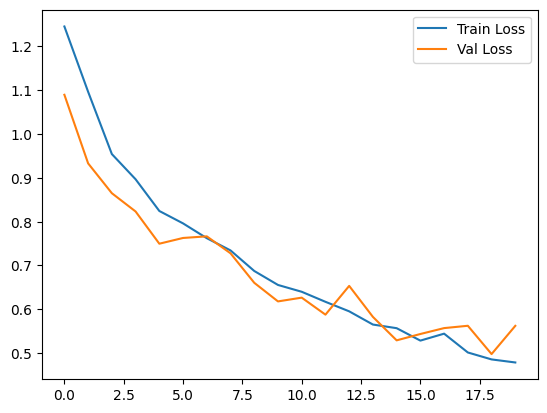
|  |  |  |
| --- | --- | --- |
| *Resize* | Akurasi | Loss |
| 64x64 | 76.51% | 0.5998 |
| 128x128 | 81.66% | 0.4787 |
| 150x150 | 81.69% | 0.4722 |

Dari **TABEL** didapatkan hasil *resize* citra dengan akurasi terbaik pada sistem tersebut yaitu citra dengan ukuran 150x150piksel. Hal itu dinyatakan dengan nilai akurasi yang diperoleh sebesar 81.69% dengan nilai *loss* sebesar0.4722.

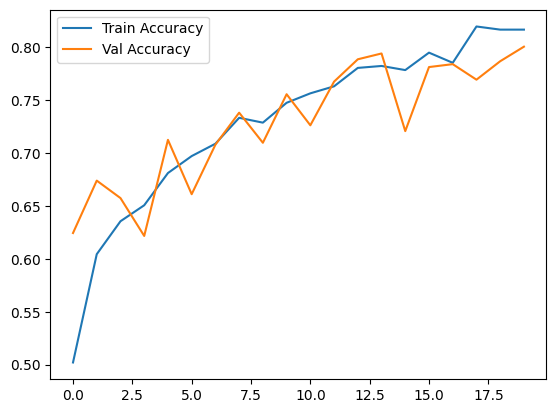
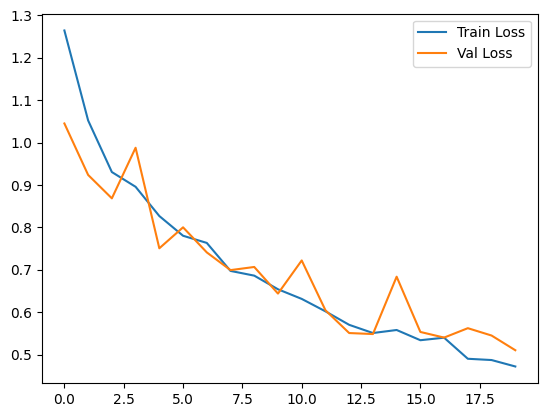
*Input Size 64x64*

*Input Size 128x128*

*Input Size 150x150*

Selanjutnya pada **GAMBAR** menunjukkan grafik yang telah diperoleh dari hasil *training* dan *validation* model dengan *input size* yang berbeda. Berdasarkan hasil pengujian tersebut, menunjukkan hasil stabil dengan akurasi yang paling baik adalah ukuran 150x150 *pixel* terutama di grafik *accuracy* yang mendapatkan angka yang lebih baik dari yang lain. Pada *input size* 128x128 cukup stabil, akan tetapi angka akurasi dengan parameter yang sama tidak lebih baik dari ukuran 150x150. Sementara itu, ukuran 64x64 memiliki angka akurasi paling rendah dengan parameter yang sama.

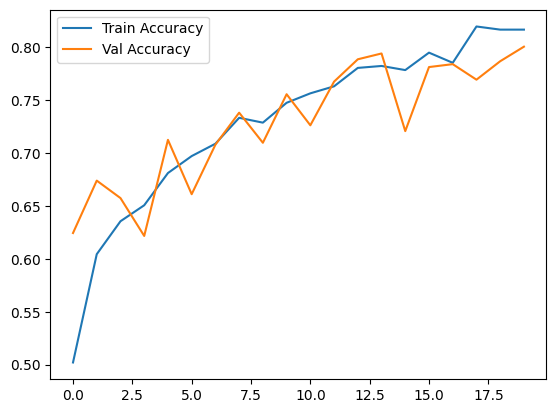
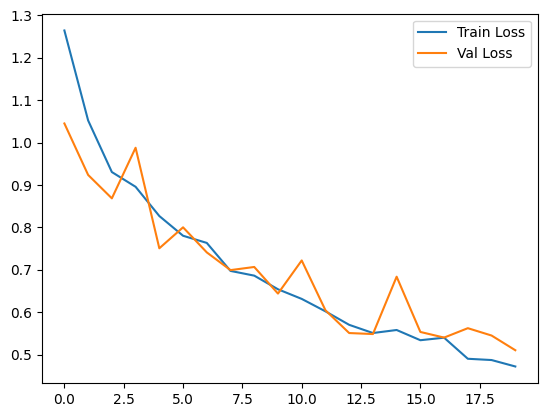
* 1. Pengujian Skenario 2 : Perubahan Jenis *Optimizer*

Pada skenario pertama telah dilakukan pengujian dengan melakukan perubahan pada input citra. Dalam skenario ini dilakukan tahapan perubahan jenis *optimizer* yaitu Adam, Nadam, dan RMSprop. Skenario ini parameter yang tetap adalah *image size*, *learning rate*, *epoch*, dan *batch size*. Seperti pengujian sebelumnya hasil pengujian ini juga dilihat melalui nilai akurasi dan *loss* untuk mengetahui kualitas dari masing-masing *optimizer* guna menentukan parameter tetap selanjutnya.

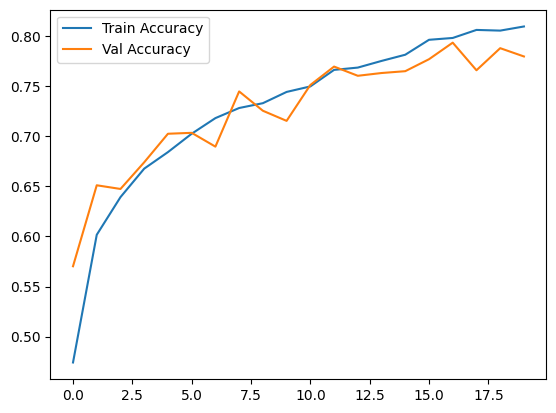
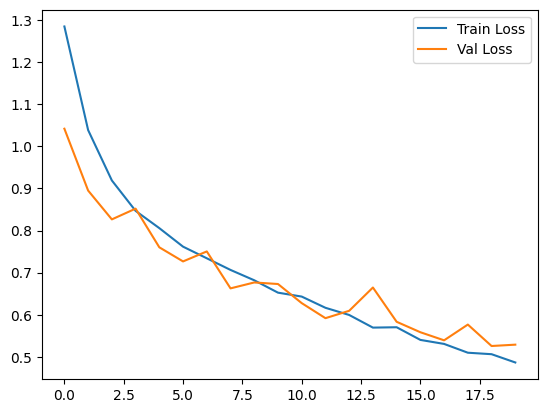
|  |  |  |
| --- | --- | --- |
| *Optimizer* | Akurasi | *Loss* |
| Adam | 81.69% | 0.4722 |
| Nadam | 80.95% | 0.4864 |
| RMSprop | 75.27% | 0.6513 |

Pada skenario 1, parameter tetapnya adalah *optimizer* Adam. Setelah diuji, *optimizer* Adam masih lebih unggul dari kedua parameter *optimizer* lainnya. Hal ini dapat dilihat dari hasil *training* model berdasarkan **TABEL** di atas. Angka yang ditunjukkan dari *optimizer* Adam unggul dari Nadam dan RMSprop. *Optimizer* RMSprop memiliki akurasi paling rendah dengan nilai *loss* paling tinggi.

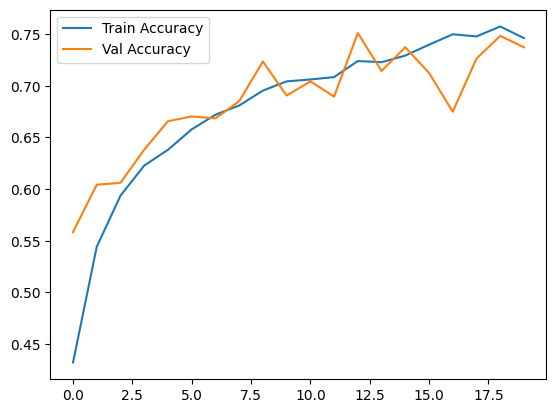
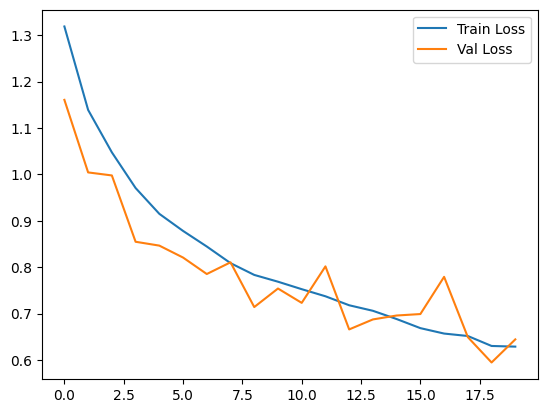
*Optimizer Adam*

*Optimizer Nadam*

*Optimizer RMSprop*

Berdasarkan **GAMBAR**, ditunjukkan bahwa grafik dari keseluruhan pengujian *optimizer* tidak terindikasi *overfitting* atau pun *underfitting*. Ketiga *optimizer* cukup stabil dengan beragam hasil akurasi akhir. Berdasarkan hasil analisis, *fixed* parameter selanjutnya akan menggunakan *optimizer* Adam.

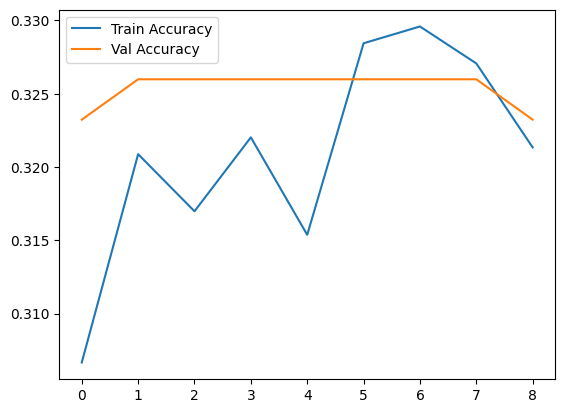
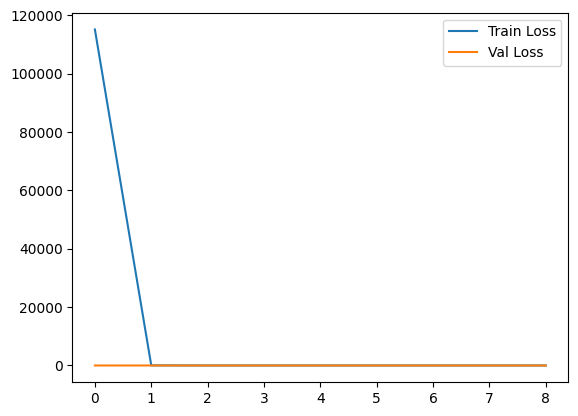
* 1. Pengujian Skenario 3 : Perubahan Nilai *Learning Rate*

Pada skenario ketiga melakukan tahapan perubahan nilai *learning rate*. Adapun *learning rate* yang digunakan adalah 0,01;0,001;0,0001. Dalam skenario ini parameter tetap yang digunakan yaitu *image size*, *optimizer*, *epoch*, dan *batch size*. Hasil pengujian skenario ini akan digunakan untuk mengetahui kualitas terbaik sistem klasifikasi dari nilai *learning rate* untuk menentukan parameter selanjutnya.

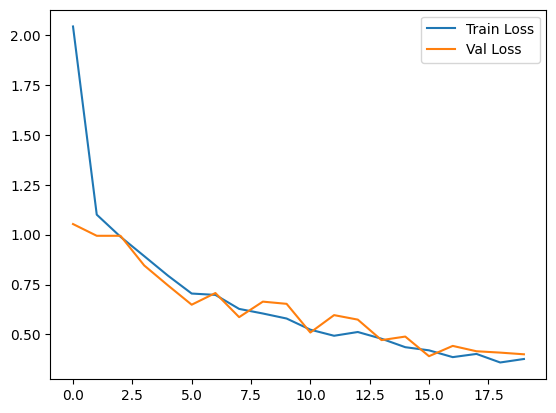
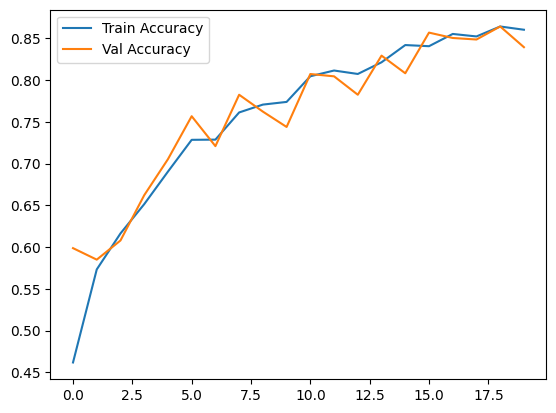
|  |  |  |
| --- | --- | --- |
| *Learning Rate* | Akurasi | *Loss* |
| 0.1 | 32.13% | 1.501 |
| 0.001 | 86.02% | 0.3767 |
| 0.0001 | 81.69% | 0.4722 |

Berdasarkan **TABEL**, diperoleh akurasi terbaik pada pengujian *learning rate* 0,001. Akurasi dan *loss* yang dihasilkan adalah sebesar 86,02% dan 0,3767. Pengujian model dengan parameter 0,1 mengalami *early stop* karena proses training yang melewati syarat maksimal waktu dalam 5 *epochs*, oleh karena itu grafik yang ditunjukkan adalah garis datar setelah iterasi ke-1 dan e*poch* terhenti di angka 8. Sementara itu, *learning rate* 0,0001 cukup stabil namun memiliki waktu yang sangat lama dan angka akurasi yang masih kecil dan *loss* cukup besar.

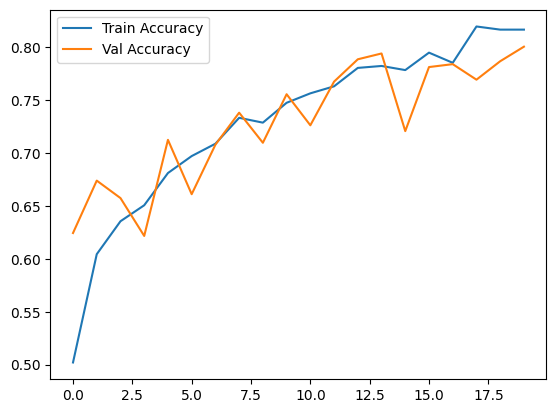
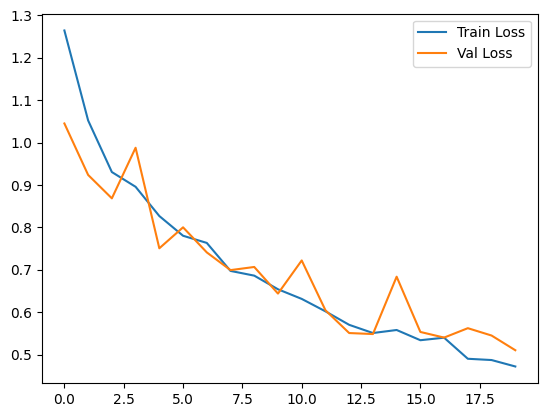
*Learning Rate 0.1*

*Learning Rate 0.001*



*Learning Rate 0.0001*

Pada skenario 2, *learning rate* yang dipakai adalah 0,0001. Namun, setelah dilakukan pengujian dan hasil analisis, penulis menentukan *learning rate* paling baik adalah 0,001. Hal tersebut didukung dengan **GAMBAR** yang menunjukkan akurasi terbaik ada di *learning rate* 0,001. Skenario selanjutnya akan menggunakan *fixed* parameter dengan *learning rate* 0,001.

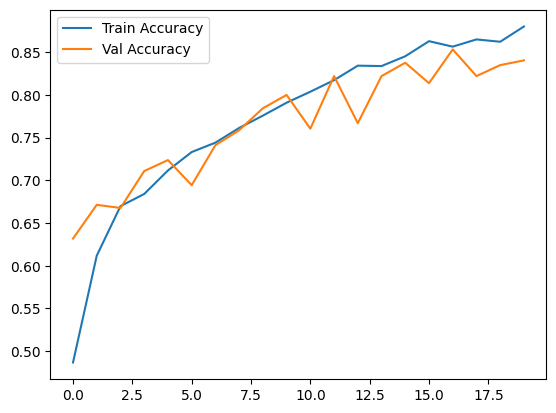
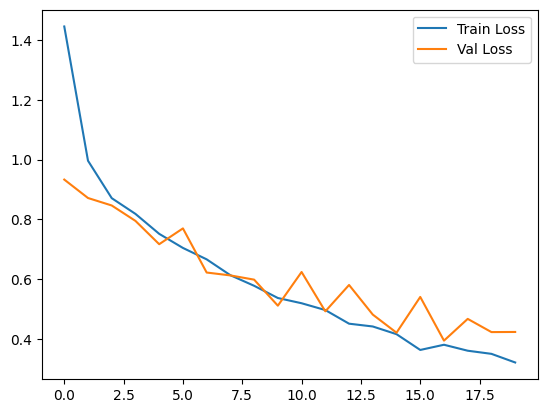
* 1. Pengujian Skenario 4 : Perubahan Nilai *Batch*

Pada skenario keempat melakukan tahapan perubahan ukuran *batch*. Adapun *batch* yang digunakan adalah 32,64,128. Dalam skenario ini parameter tetap yang digunakan yaitu *image size*, *optimizer*, *epoch*, dan *learning rate*. Hasil pengujian skenario ini akan digunakan untuk mengetahui kualitas terbaik sistem klasifikasi dari ukuran *batch.*

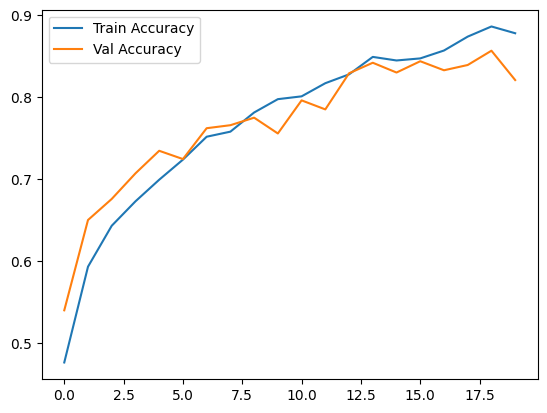
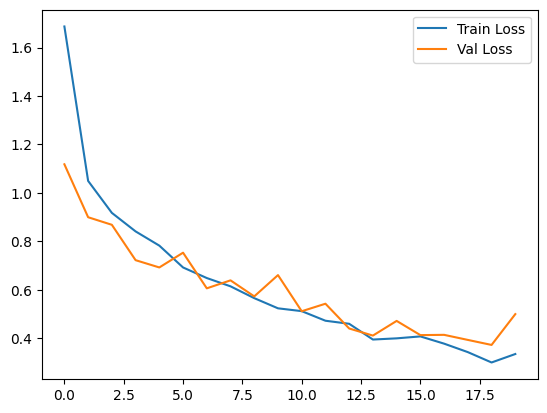
|  |  |  |
| --- | --- | --- |
| *Batch* | Akurasi | *Loss* |
| 32 | 87.99% | 0.3216 |
| 64 | 87.81% | 0.3357 |
| 128 | 86.11% | 0.3750 |

Berdasarkan hasil pada **TABEL** diperoleh akurasi terbaik pada pengujian *batch size* 32 dengan nilai akurasi 87,99% dan *loss* 0,7963. Selain itu, dari tabel di atas, dapat disimpulkan bahwa nilai *batch size* yang semakin besar belum tentu menghasilkan hasil pengujian sistem yang semakin baik.

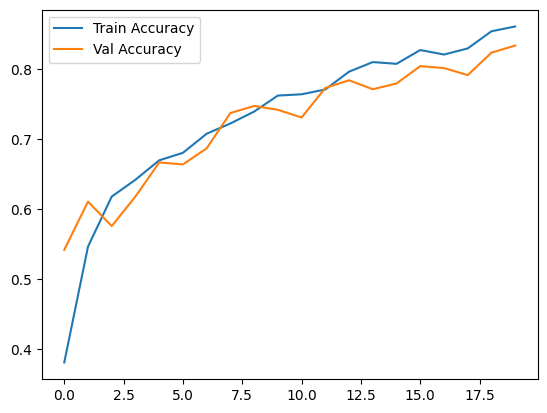
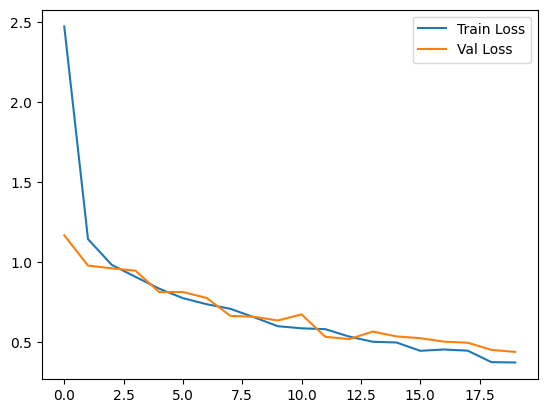
*Ukuran Batch 32*

*Ukuran Batch 64*

*Ukuran Batch 128*

Kemudian pada **GAMBAR** memperlihatkan perbandingan grafik hasil pengujian *batch size*. Pada masing-masing grafik memiliki hasil grafik yang baik dengan tidak adanya *overfitting* ataupun *underfitting*. Pada pengujian ini hasil grafik terbaik dengan akurasi paling baik adalah hasil pengujian nilai *batch size* 32.

## Pembahasan

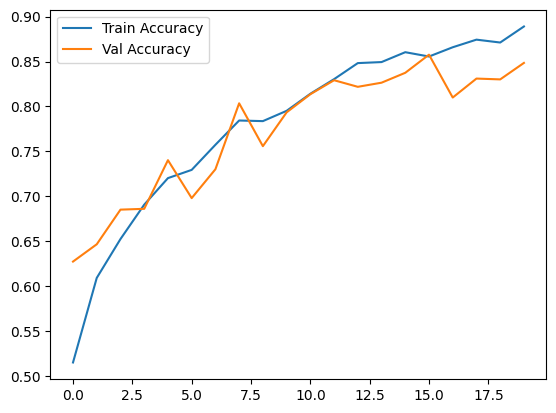
Setelah melakukan tahapan *training* dan *testing*, hasil pengujian 4 paramater berbeda menghasilkan *hyperparameter* yang terbaik untuk klasifikasi hama dan penyakit tanaman padi dengan data yang sudah difilterisasi. Berikut adalah pembahasan dari hyperparameter terbaik dari hasil uji.

### Hasil Pengujian Terbaik

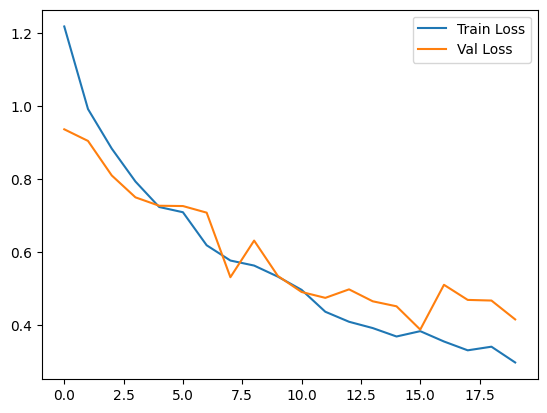
Pengujian sistem klasifikasi hama dan penyakit tanaman padi dengan metode convolutional neural network menggunakan 4 skenario yang berbeda dan dipilih parameter dengan hasil terbaik di setiap skenarionya. Parameter yang paling sesuai dengan sistem adalah ukuran citra 150x150 *pixel*, *optimizer* Adam, *learning rate* 0,001, dan *batch size* 32. Pada **TABEL** ditunjukkan hasil pengujian menggunakan parameter yang sudah ditentukan.

|  |  |  |
| --- | --- | --- |
| Hyperparameter | Akurasi | Loss |
| *Input Size :* 150x150 *pixel* | 88,91% | 0.4154 |
| *Optimizer :* Adam |
| *Learning Rate :* 0,001 |
| *Batch Size* : 32 |

*Model Accuracy*



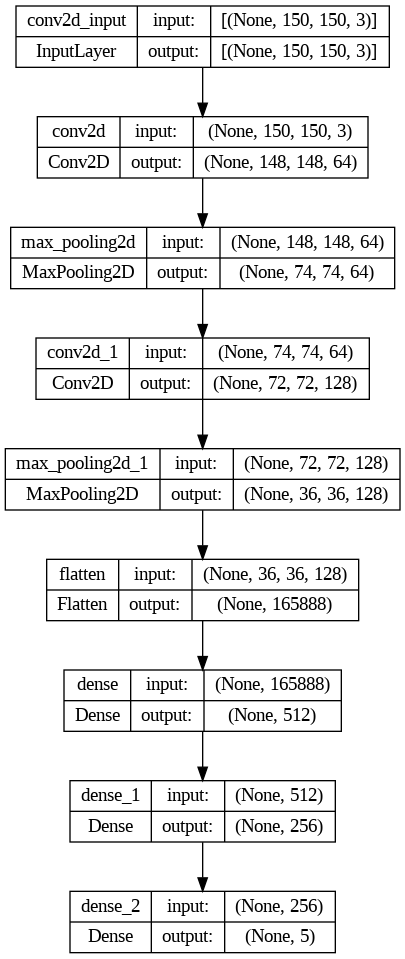
*Model Loss*



Berdasarkan **GAMBAR** grafik di atas, dapat dilihat akurasi dari model dan *loss* stabil dan tidak terdeteksi *overfitting* ataupun *underfitting*. Hasil dari grafik tersebut didapatkan dari hasil penyusunan model dengan mempertimbangkan banyaknya dataset gambar serta kemampuan perangkat dalam melakukan training. *Training Accuracy* dan *Validation* stabil dengan angka yang terus meningkat hingga mencapai angka 88,91%. Sementara itu, nilai *loss* tidak melonjak tinggi dan tetap stabil dengan mencapai angka 0,4154.

### Model *Deep Learning* Pengujian Akhir

Proses pengujian hingga akhirnya menemukan parameter terbaik dalam memaksimalkan nilai akurasi dari *training* data untuk model, maka memerlukan arsitektur model CNN yang optimal dan tidak berlebihan. Dalam penelitian ini, tahap pengujian akhir dari data menggunakan arsitektur yang disusun tanpa melakukan proses *transfer learning*. Adapun arsitektur model CNN dari penelitian dapat divisualisasikan sebagai berikut.



Model CNN di atas memiliki 2 (dua) pasang lapisan konvolusi Conv2D (*Convolutional Layer*) dan *pooling* MaxPool2D yang disusun secara berurutan untuk mengekstrak fitur dari gambar dan kemudian mengurangi dimensi *feature map* yang dihasilkan lapisan konvolusi Conv2D. Dimensi yang berkurang ini dapat dilihat dari ukuran panjang dan lebar gambar sepanjang prosesnya melalui lapisan konvolusi dan *pooling* dari 150x150 (ukuran gambar input) sehingga menjadi 36x36 pada lapisan MaxPool2D sebelum diproses oleh lapisan Flatten. Fungsi aktivasi yang digunakan pada seluruh lapisan konvolusi Conv2D adalah RelU, yang melibatkan nonlinier pada model, mendorong model untuk berfokus pada fitur yang paling penting dan relevan sehingga proses pelatihan menjadi lebih efektif dan efisien (komputasi program lebih cepat).

Setelah diproses oleh lapisan konvolusi dan *pooling*, *feature map* diproses lapisan Flatten sehingga berubah menjadi bentuk vektor 1 (satu) dimensi. Vektor ini kemudian dimasukkan kepada lapisan Dense yang bertindak sebagai *fully-connected layer* untuk melakukan klasifikasi multi kelas. Algoritma optimasi Adam digunakan untuk meminimalisasi *loss function* saat proses pelatihan, dan fungsi kerugian *categorical cross-entropy* digunakan untuk menghitung selisih data yang diprediksi dengan data sebenarnya. Algoritma optimasi Adam digunakan dengan tingkat pembelajaran 0,001 yang memperbaharui bobot selama proses pelatihan model yang dilakukan dalam 20 *epoch*. Metrik yang digunakan untuk mengevaluasi hasil kinerja model merupakan *accuracy*.

Arsitektur model CNN yang dibangun adalah sebagai berikut.

|  |  |  |
| --- | --- | --- |
| Lapisan Model | Konfigurasi Kompilasi Model | Jumlah Epoch |
| * Conv2D dengan 64 filter, kernel 3x3, fungsi aktivasi ‘RelU’, serta menerima bentuk masukan sesuai dengan *image size*, dan memiliki 3 kanal warna (RGB) * MaxPool2D dengan kernel berukuran 2x2 * Conv2D dengan 128 filter, kernel berukuran 3x3, fungsi aktivasi ‘RelU’ * MaxPool2D dengan kernel berukuran 2x2 * Flatten * Dense dengan 512 unit, fungsi aktivasi ‘RelU’ * Dense dengan 256 unit, fungsi aktivasi ‘RelU’ * Dense dengan 5 unit, fungsi aktivasi ‘Softmax’ | * Algoritma optimasi ‘Adam’ dengan tingkat pembelajaran 0,001. * Fungsi kerugian ‘Categorical Cross-Entropy’ * Metrik yang dievaluasi yaitu ‘Accuracy’ | 20 |

# BAB 5

# PENUTUP

## 5.1. Kesimpulan

Berdasarkan penelitian yang dilakukan, dapat dicetuskan sejumlah kesimpulan sebagai berikut.

1. Klasifikasi hama dan penyakit tanaman padi adalah hal yang penting untuk memastikan hama/penyakit padi secara dini sebelum ditangani oleh ahlinya. Oleh karena itu, dibuatlah sebuah model CNN untuk mempermudah petani dalam mendeteksi/klasifikasi hama dan penyakit tanaman padi. Dalam penelitian ini, model CNN dibangun dengan karakteristik yang sudah ditentukan dan melalui tahap *preprocessing* sebelumnya.
2. Untuk menghasilkan model dengan akurasi yang optimal, penelitian ini dilakukan 4 skenario yaitu *image size*, jenis *optimizer*, *learning rate*, dan ukuran *batch size*. Akurasi tertinggi yang diperoleh saat sistem menggunakan kombinasi parameter *resize* citra 150x150 *pixel*, *optimizer* Adam, *learning rate* 0,001, dan *batch size* 32.
3. Pada akhir pengujian model, penelitian ini memperoleh hasil akurasi sistem sebesar 88,91%. Sementara itu, nilai *loss* tidak melonjak tinggi dan tetap stabil dengan mencapai angka 0,4154.

## 5.2. Saran

Berdasarkan penelitian yang sudah dilakukan, sebagai upaya dalam pengembangan penelitian ini di masa yang akan datang nanti, maka terdapat beberapa saran yaitu sebagai berikut.

1. Menambahkan jumlah kelas hama dan penyakit tanaman padi agar lebih maksimal penggunaan model CNN dalam kehidupan.
2. Melakukan tahapan *deploy* model yaitu menintegrasikannya dengan aplikasi dalam bentuk *website* ataupun *mobile*.
3. Melakukan uji dengan *object detection* agar pendeteksian hama dan penyakit oleh mesin dapat lebih akurat dan valid.

DAFTAR PUSTAKA

Anam Alidrus, S., Aziz, M., & Virgantara Putra, O. (2021). Deteksi Penyakit Pada Daun Tanaman Padi Menggunakan Metode Convolutional Neural Network. *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia*.

Ayunina, R. (2023). *TINGKAT SERANGAN DAN KEHILANGAN HASIL TANAMAN PADI OLEH HAMA WALANG SANGIT DAN PENYAKIT BLAS DI KECAMATAN X KOTO SINGKARAK, KABUPATEN SOLOK, SUMATERA BARAT*. Institut Pertanian Bogor, Bogor.

Dewi, S. R. (2018). *DEEP LEARNING OBJECT DETECTION PADA VIDEO MENGGUNAKAN TENSORFLOW DAN CONVOLUTIONAL NEURAL NETWORK (Studi Kasus: Klasifikasi Gambar Meja dan Kursi Motif Ukiran Jepara)*. Universitas Islam Indonesia, Yogyakarta.

Dypanuryansyah, M. R. (2018). *IDENTIFIKASI SAWAH TERSERANG HAMA/PENYAKIT MENGGUNAKAN CITRA LANDSAT 8 (OLI) DI KABUPATEN INDRAMAYU*. Institut Pertanian Bogor, Bogor.

Jinan, A., & Hayadi, B. H. (2022). Klasifikasi Penyakit Tanaman Padi Mengunakan Metode Convolutional Neural Network Melalui Citra Daun (Multilayer Perceptron). *Journal of Computer and Engineering Science*, *1*(2), 37–44.

Khasanah, I. N., & Astuti, K. (t.t.). *LUAS PANEN DAN PRODUKSI PADI DI INDONESIA 2022*. Badan Pusat Statistik.

Khoiruddin, M., Junaidi, A., & Saputra, W. A. (2022). Klasifikasi Penyakit Daun Padi Menggunakan Convolutional Neural Network. *Journal of Dinda*, *2*(1), 37–45. Diambil dari https://www.kaggle.com/tedisetiady/leaf-rice-disease-

Murad, N., Mappa, H., & Rudin, S. (2021). *PENGENALAN HAMA DAN PENYAKIT PADA TANAMAN PADI* (1 ed.; Dr. Ir. M. D. Mario & S. P. , Y. Ismail, Ed.). Gorontalo: Dinas Pertanian Provinsi Gorontalo.

Puslitbang Tanaman Pangan, BPTP SUMUT, BPTP Riau, BPTP Lampung, BPTP DKI, BPTP DIY, … IRRI. (2007). *Hama dan Penyakit Tanaman Padi* (3 ed.).

Rao Moparthi, N., Bhattacharyya, D., Balakrishna, G., & Prashanth, J. S. (2021). Paddy Leaf Disease Detection using CNN. *Volatiles & Essent. Oils*, *8*(4), 800–806.

Rizky, A. T., Fua’adah, R. Y. N., & Rizal, S. (2023). *DETEKSI PENYAKIT PADA TANAMAN JAGUNG BERBASIS PENGOLAHAN CITRA DIGITAL DENGAN METODE CONVOLUTIONAL NEURAL NETWORK*. Bandung.

Saputra, R. A., Wasyianti, S., Supriyatna, A., & Saefudin, D. F. (2021). Penerapan Algoritma Convolutional Neural Network Dan Arsitektur MobileNet Pada Aplikasi Deteksi Penyakit Daun Padi. *JURNAL SWABUMI*, *9*(2). Diambil dari https://archive.ics.uci.edu/ml/datasets/Rice

Sheila, S., Kharil Anwar, M., Saputra, A. B., Pujianto, R., & Sari, I. P. (2023). Deteksi Penyakit pada Daun Padi Berbasis Pengolahan Citra Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Multinetics*, *9*(1), 27–34. Diambil dari https://www.kaggle.com/datasets/tedisetiady/leaf-

Swathika, R., Srinidhi, S., Radha, N., & Sowmya, K. (2021). Disease identification in paddy leaves using CNN based deep learning. *Proceedings of the 3rd International Conference on Intelligent Communication Technologies and Virtual Mobile Networks, ICICV 2021*, 1004–1008. https://doi.org/10.1109/ICICV50876.2021.9388557

Warti. (2006). *PERKEMBANGAN HAMA TANAMAN PADI PADA TIGA SISTEM BUDIDAYA PERTANIAN DI DESA SITU GEDE, KECAMATAN BOGOR BARAT, KOTA BOGOR*. Institut Pertanian Bogor, Bogor.

Wijayanto, B., Kiswanto, & Manurung, G. O. (2013). *Hama dan Penyakit Utama Tanaman Padi* (J. Hadipurwanta, Ed.). Bandar Lampung: Balai Pengkajian Teknologi Pertanian (BPTP) Lampung.

Yulistiani, R., Ramadhan, I., Said, Q., & Santoni, M. M. (2018). Metode Regresi Linier Berganda dan SVR dalam Menentukan Tingkat Pengaruh Cuaca Terhadap Produktivitas Padi di Indonesia. *Sistem Informasi Dan Keamanan Siber (SEINASI-KESI) Jakarta-Indonesia*. Diambil dari http://www.bps.go.id