**PENERAPAN MACHINE LEARNING UNTUK KLASIFIKASI PENYAKIT PADA TANAMAN ANGGUR MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN) DAN K-NEAREST NEIGHBOR (KNN)**



Disusun oleh:

Khairuz Zuhdi - 221011400174

Bintang Syaputra - 221011402754

Rizal Fazri - 221011400157

Satrio Panca Nugroho - 221011403120

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS PAMULANG

2024

DAFTAR ISI

[ABSTRAK 4](#_Toc186648558)

[BAB I 5](#_Toc186648559)

[PENDAHULUAN 5](#_Toc186648560)

[1.1 Latar belakang 5](#_Toc186648561)

[1.2 Tujuan 8](#_Toc186648562)

[1.3 Manfaat 9](#_Toc186648563)

[BAB II 12](#_Toc186648564)

[KAJIAN TEORI 12](#_Toc186648565)

[2.1 Teori terkait 12](#_Toc186648566)

[2.2.2 K-Nearest Neighbor 13](#_Toc186648567)

[2.2.3 Klasifikasi 14](#_Toc186648568)

[2.2.4 Morfologi 15](#_Toc186648569)

[2.2.5 EfficientNet 15](#_Toc186648570)

[BAB III 18](#_Toc186648571)

[METODOLOGI PENELITIAN 18](#_Toc186648572)

[3.1 Desain Penelitian 18](#_Toc186648573)

[3.2 Data Set 18](#_Toc186648574)

[3.2.1 Metode KNN 19](#_Toc186648575)

[3.3 Preprocessing Data 20](#_Toc186648576)

[3.3.1 Convolutional Neural Network (CNN) 21](#_Toc186648577)

[3.4 Model Machine Learning 26](#_Toc186648578)

[3.4.1 Convolutional Neural Network (CNN) 26](#_Toc186648579)

[3.4.2 EfficientNet 27](#_Toc186648580)

[3.4.3 Inception 27](#_Toc186648581)

[BAB IV 28](#_Toc186648582)

[PEMBAHASAN 28](#_Toc186648583)

[4.1 IMPORT LIBRARY 28](#_Toc186648584)

[4.2 MEMUAT DATA SET 29](#_Toc186648585)

[4.3 Menampilkan Gambar Dari Dataset 31](#_Toc186648586)

[4.4 Analisis Tingkat Kesalahan 32](#_Toc186648587)

[4.5 Menampilkan loss curves 33](#_Toc186648588)

[4.6 Membuat Prediksi pada Data Test 36](#_Toc186648589)

[4.7 Menampilkan 25 gambar random dari dataset yang diberi label 37](#_Toc186648590)

[4.8 TEST LABEL KLASIFIKASI REPORT DAN CONFUSION MATRIX 38](#_Toc186648591)

[4.9 KLASIFIKASI REPORT 40](#_Toc186648592)

[4.10 Membuat Tabel Confusio matrix 41](#_Toc186648593)

[BAB V 43](#_Toc186648594)

[KESIMPULAN DAN SARAN 43](#_Toc186648595)

[5.1 Kesimpulan 43](#_Toc186648596)

[5.2 Saran 43](#_Toc186648597)

[DAFTAR PUSTAKA 44](#_Toc186648598)

# ABSTRAK

Pemanfaatan teknologi machine learning dalam klasifikasi penyakit pada tanaman anggur merupakan terobosan signifikan di bidang pertanian, terutama untuk mengatasi kendala identifikasi penyakit yang selama ini bergantung pada metode manual. Penelitian ini dirancang untuk mengembangkan sistem berbasis Convolutional Neural Network (CNN) dan K-Nearest Neighbor (KNN) yang dapat secara otomatis mendeteksi serta mengklasifikasikan penyakit pada daun anggur. Data yang digunakan mencakup 4000 citra daun anggur yang dibagi menjadi tiga kelompok: pelatihan, validasi, dan pengujian. Hasil penelitian menunjukkan bahwa metode CNN unggul dalam hal akurasi karena kemampuannya mengenali pola-pola kompleks tanpa memerlukan langkah preprocessing yang rumit, sedangkan KNN menawarkan fleksibilitas dan kemudahan implementasi. Dengan kedua metode ini, sistem dapat mempercepat proses identifikasi, meminimalkan kesalahan manusia, dan memberikan solusi berbasis teknologi untuk meningkatkan kualitas dan produktivitas hasil panen. Penelitian ini memberikan kontribusi penting dalam modernisasi sektor pertanian, khususnya dalam pengelolaan penyakit pada tanaman anggur.

# BAB I

# PENDAHULUAN

## 1.1 Latar belakang

Indonesia merupakan salah satu negara agraris dengan potensi besar dalam sektor pertanian, termasuk dalam budidaya tanaman anggur. Tanaman anggur memiliki nilai ekonomi yang tinggi dan menjadi komoditas penting baik untuk pasar lokal maupun internasional. Namun, salah satu tantangan utama dalam budidaya anggur adalah keberadaan berbagai jenis penyakit yang dapat menyerang tanaman ini. Penyakit pada tanaman anggur, seperti fine mold, fleece mold, dan dark spoil, dapat menyebabkan penurunan kualitas dan kuantitas hasil panen, yang pada akhirnya merugikan para petani. Identifikasi dini terhadap penyakit tanaman anggur menjadi hal yang sangat penting untuk mencegah kerugian yang lebih besar. Namun, proses identifikasi ini sering kali dilakukan secara manual oleh para petani atau ahli pertanian, yang membutuhkan pengalaman, keahlian, dan waktu yang tidak sedikit. Pendekatan manual ini juga rentan terhadap kesalahan karena sifat gejala penyakit yang terkadang mirip satu sama lain. Seiring dengan berkembangnya teknologi, penerapan kecerdasan buatan (Manufactured Insights) dan pembelajaran mesin (Machine Learning) telah menjadi solusi potensial dalam menyelesaikan permasalahan ini. Salah satu metode pembelajaran mesin yang banyak digunakan dalam klasifikasi citra adalah Convolutional Neural Organize (CNN). CNN memiliki kemampuan untuk menganalisis dan mengenali pola-pola pada gambar dengan tingkat akurasi yang tinggi, sehingga sangat cocok untuk diterapkan dalam pengenalan dan klasifikasi penyakit pada tanaman anggur.

Penerapan CNN dalam klasifikasi penyakit pada tanaman anggur tidak hanya mempercepat proses identifikasi, tetapi juga meningkatkan akurasi hasil deteksi. Dengan demikian, petani dapat mengambil tindakan pengendalian yang tepat dan cepat untuk mencegah penyebaran penyakit. Penelitian ini bertujuan untuk mengembangkan demonstrate berbasis CNN yang mampu mengklasifikasikan jenis penyakit pada tanaman anggur secara otomatis, sehingga dapat memberikan kontribusi dalam peningkatan efisiensi dan produktivitas sektor pertanian. Banyak penelitian telah dilakukan untuk mengklasifikasi penyakit pada jenis-jenis daun. Salah satunya adalah klasifikasi penyakit pada daun anggur, menggunakan citra daun yang disegmentasi menggunakan thresholding, Gray Level Co-Occurrence Framework (GLCM) dan Radon Change untuk mengekstraksi fitur citra, dan metode Support Vector Machine (SVM) untuk mendeteksi penyakit. Hasilnya menunjukkan akurasi sebesar 89.90% (Shweta, dkk, 2018). Kelemahan dari metode SVM yaitu memiliki beberapa parameter seperti C (yang mengatur tingkat toleransi terhadap kesalahan) dan part yang perlu diatur dengan benar untuk mencapai klasifikasi terbaik dan tidak cocok untuk dataset yang besar. Klasifikasi lainnya menggunakan metode Convolutional Neural Arrange (CNN), dipadukan dengan teknik Exchange Learning. Metode ini memiliki keunggulan yaitu dilakukan pra-pelatihan dengan menggunakan informasi yang telah dipelajari dari demonstrate CNN yang telah dilatih sebelumnya kemudian mengubah informasi ini untuk tugas klasifikasi baru. Pendekatan ini dapat mengurangi waktu pelatihan dan memerlukan lebih sedikit information untuk mencapai performa yang baik.Kelemahan Exchange Learning yaitu negative exchange, perlu diingat bahwa exchange learning bekerja lebih baik ketika pola-pola information kelas baru sebanding dengan polapola yang ada di pra-pelatihan model. Jika pola-pola yang ada di antara kedua tugas sangat berbeda, exchange learning mungkin tidak bekerja dengan baik, dan demonstrate harus dilatih dari awal menggunakan dataset kelas baru. Keunggulan metode CNN tidak memerlukan preprocessing, ekstraksi fitur dan klasifikasi dapat dilakukan sekaligus. Akurasi CNN mencapai 91% untuk mengenali penyakit (Gulavnai & Patil,2019).

Proses untuk menemukan sekumpulan demonstrate atau organisms kedalam kelas-kelas tertentu disebut dengan klaifikasi, proses klasifikasi memiliki tujuan untuk menggunakan demonstrate sebagai penentu kelas terhadap suatu objek yang belum diketahui kelasnya . KNN merupakan salah satu metode unutk melakukan klasifikasi berdasarkan information pembelajaran yang memiliki jarak yang terdekat atau mempunyai persamaan ciri yang withering banyak terhadap objek yang diprediksi. Untuk menentukan kedekatan jarak antara tetangga biasanya dilakukan perhitungan menggunakan Euclidean Separate. Ekstraksi fitur atau Include Extraction merupakan suatu proses nutuk mengambl ciri/feature dari sebuahbentuk ataupun tekstur yang kemudan diari nilai tersebut dapat digunakan untuk analisis pada proses klasifikasi .Gray Level Co-occurrence Framework atau GLCM merupakan sebuah metode yang berguna untuk menganalisis ekstraksi tekstur/fitur, di dalam GLCM terdapat sebuah network yang menggambarkan frekuensi kemunculan dua pasangan piksel dalam intensitas tertentu pada gambar berdasarkan jarak dan arah yang telah ditentukan . tekstur dan bentuk gambar yang berbeda pada bidang pengenalan pola dan komputer visi dapat diukur menggunakan GLCM, selain itu GLCM juga telah diganakan di berbagai bidang dan terbukti berhasil dalam melakukan banyak kegiatan klasifikasi . Pada beberapa penelitian menggunakan algoritma K-NN dengan ekstraksi fitur GLCM sebelumnya menjunjukkan bahwa kedua metode tersebut dapat bekerja dengan baik dalam melakukan proses klasifikasi. Pada penelitian “implementasi gray level cooccurrence lattice (glcm) untuk klasifikasi penyakit daun padi” diperoleh akurasi tertinggi sebesar 93,3% waktu melakukan klasifikasi 3 kelas daun dengan jumlah information latih sebanyak 210 dan information uji sebanyak 30 . Sedangkan pada penelitian “Klasifikasi Citra Daun dengan GLCM (Gray Level Co-Occurence) dan K-NN (K-Nearest Neighbor)“ dilakukan klasifikasi terhadap 8 jenis citra daun menggunakan 400 information latih dan 80 information uji, pada proses klasifikasi diperoleh nilai akurasi tertinggi sebesar 98% waktu melakukan klasifikasi menggunakan nilai k = 1, sedangkan nilai akurasi withering kecil diperoleh dengan nilai sebesar 89% waktu menggunakan nilai k = 7 . Sebagai upaya dalam penanganan penyebaran penyakit pada tanaman anggur dilkaukan penelitian untuk menklasifikasikan jenis penyakit daun anggur menggunakan algoritma K-Nearest Neighbor berdasarkan ekstraksi fitur Gray Level Co-occurence Network. Pada penelitian ini dilakukan pengambilan information dari dataset plantVillage yang kemudian diolah sehingga diperoleh nilai ekstraksi fitur dari metode GLCM sehingga dapat dilakukan proses klasifikasi menggunakan KNN.

## 1.2 Tujuan

Penelitian ini bertujuan untuk mengembangkan sistem berbasis Convolutional Neural Network (CNN) dan K-NEAREST NEAIGHBOR yang dapat secara otomatis mengklasifikasikan jenis penyakit pada tanaman anggur. Dengan melakukan perbandingan dari 2 metode ini, diharapkan dapat meningkatkan efisiensi dan akurasi dalam proses identifikasi penyakit, yang selama ini dilakukan secara manual oleh petani atau ahli pertanian.

1. Mengidentifikasi Penyakit Secara Otomatis:

Mengembangkan Metode CNN dan K-NNN untuk mendeteksi dan mengklasifikasikan penyakit pada tanaman anggur dengan akurasi yang lebih tinggi dibandingkan metode manual.

2. Meningkatkan Efisiensi Proses Identifikasi:

Mempercepat proses identifikasi penyakit sehingga petani dapat segera mengambil tindakan pengendalian yang tepat untuk mencegah penyebaran penyakit.

3. Mengurangi Kesalahan Manusia:

Mengurangi kemungkinan kesalahan dalam identifikasi penyakit yang sering terjadi akibat kesamaan gejala antara berbagai jenis penyakit.

4. Menyediakan Solusi Berbasis Teknologi:

Memanfaatkan kemajuan teknologi dalam pembelajaran mesin untuk memberikan solusi yang lebih efektif dalam pengelolaan kesehatan tanaman anggur, sehingga dapat meningkatkan produktivitas dan kualitas hasil panen.

5. Memberikan Kontribusi pada Sektor Pertanian:

Meningkatkan pemahaman dan pengetahuan tentang penggunaan teknologi cutting edge dalam pertanian, khususnya dalam pengelolaan tanaman anggur sebagai komoditas penting di Indonesia.

6. **Penggunaan Teknik Pembelajaran Mesin**:

Menggunakan teknik pembelajaran mesin dan deep learning untuk meningkatkan proses klasifikasi, termasuk pengoptimalan hyperparameter dan preprocessing citra agar model lebih efektif dalam mengenali pola-pola penyakit.

Dengan mencapai tujuan-tujuan tersebut, penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap sektor pertanian, khususnya dalam budidaya tanaman anggur di Indonesia.

## 1.3 Manfaat

Penelitian mengenai penerapan metode Convolutional Neural Organize (CNN) dan K-Nearest Neighbor (K-NN) dalam klasifikasi penyakit pada tanaman anggur memiliki beberapa manfaat yang signifikan. Berikut adalah penjelasan mengenai manfaat dari kedua metode tersebut:

1. Proses Identifikasi yang Cepat:

Dengan kemampuan analisis gambar yang canggih, CNN memungkinkan proses identifikasi penyakit dilakukan dengan lebih cepat dibandingkan metode manual. Hal ini sangat penting bagi petani untuk segera mengambil tindakan pengendalian sebelum penyebaran penyakit semakin meluas.

2. Pengurangan Kebutuhan Pra-Pemrosesan:

Salah satu keunggulan CNN adalah kemampuannya untuk melakukan ekstraksi fitur dan klasifikasi secara bersamaan, sehingga mengurangi kebutuhan akan langkah-langkah pra-pemrosesan yang kompleks Ini membuat implementasi teknologi lebih sederhana dan efisien

3. Adaptasi terhadap Dataset Besar:

CNN dirancang untuk bekerja dengan baik pada dataset besar, sehingga dapat meningkatkan performa deteksi ketika information pelatihan yang digunakan lebih banyak dan beragam.

4. Responsif terhadap Perubahan Information:

K-NN tidak memerlukan pelatihan ulang setelah setiap penambahan information baru, sehingga sangat responsif terhadap perubahan kondisi di lapangan atau penambahan information baru tentang penyakit tanaman anggur.

5. Straight forwardness and Interpretability:

K-NN adalah metode yang mudah dipahami dan diimplementasikan. Dengan menggunakan jarak Euclidean untuk menentukan kedekatan antara information, petani atau ahli pertanian dapat dengan cepat memahami bagaimana klasifikasi dilakukan.

6. Fleksibilitas dalam Ekstraksi Fitur:

K-NN dapat dipadukan dengan teknik ekstraksi fitur seperti Gray Level Co-occurrence Framework (GLCM), yang terbukti efektif dalam meningkatkan akurasi klasifikasi hingga 98lam beberapa penelitian. Kombinasi ini memungkinkan analisis tekstur yang lebih mendalam, sehingga meningkatkan kemampuan deteksi penyakit.  
  
penerapan kedua metode ini tidak hanya meningkatkan efisiensi dan akurasi dalam klasifikasi penyakit pada tanaman anggur, tetapi juga memberikan solusi praktis bagi para petani untuk mengatasi tantangan yang dihadapi dalam budidaya anggur di Indonesia. 

# BAB II

# KAJIAN TEORI

## 2.1 Teori terkait

Menggunakan metode CNN dengan arsitektur EfficientNet untuk pengujian terhadap dataset yang disediakan oleh ImageNet.

Penulis : Tan & Le ( 2019 ), Pendekatan saat ini untuk mendeteksi penyakit daun anggur terutama didasari pada pengenalan visual. Pengenalan visual secara manual dapat memakan waktu yang lama serta menjadi tugas yang sulit sehingga rentan terjadinya kesalahan. Hasil diagnosis yang salah akan mengarah kepada penyalahgunaan pestisida yang akan merusak lingkungan pertumbuhan anggur dan merusak kualitas buah. Oleh karena itu, berbagai teknik spektroskopi telah banyak diterapkan dalam diagnosis dan pemantauan penyakit tanaman. Namun, persyaratan sensor besar dan perangkat yang lengkap menjadi tidak efisien dan memiliki biaya yang tinggi.

Berhubung gejala tanaman yang terinfeksi muncul secara visual pada daun tanaman, Computer Vision bisa menjadi langkah yang efektif dan cepat dalam mengidentifikasi adanya penyakit terhadap daun anggur. Kemajuan terbaru dalam pengembangan computer vision berbasis deep learning dengan arsitektur kompleks yang menggunakan jaringan syaraf tiruan sebagai dasar klasifikasi (Aravind). Salah satu metode dari deep learning yang cukup populer terhadap pengenalan citra yaitu Convolutional Neural Network (CNN).

CNN memiliki pendekatan yang berbeda dengan traditional machine learning yang mencari dan memilih dari sekian banyaknya feature extraction yang akan ditetapkan. CNN mempelajari feature extraction melalui pola gambar secara otomatis dari proses pelatihan, Selain itu, CNN sangat efektif digunakan dengan data yang besar serta berdimensi tinggi.

Klasifikasi penyakit pada daun anggur pernah dilakukan oleh beberapa peneliti sebelumnya. penelitian tersebut menggunakan metode CNN dengan arsitektur VGG-16. Model yang diusulkan mendapatkan akurasi sebesar 97.8%. Penelitian lain pernah dilakukan oleh Thet dkk (2020), penelitian tersebut menggunakan arsitektur VGG-16 dan SVM sebagai klasifikasinya. Berdasarkan penelitian tersebut, VGG-16 fine-tuning menggunakan classifier SVM tidak mendapatkan hasil akurasi yang baik dalam pengujian data testing. SVM tidak efektif digunakan untuk dataset yang besar dan tidak dapat bekerja dengan baik ketika dataset memiliki multi-classes.

### 2.2.2 K-Nearest Neighbor

Salah satu metode machine learning yang dapat digunakan dalam pengklasifikasian adalah metode KNN . KNN merupakan salah satu metode klasifikasi berbasis supervised learning, dalam melakukan klasifikasi metode ini menggunakan pendekatan dengan menghitung jumlah tetangga paling dekat. Terdapat beberapa pendekatan yang dapat digunakan untuk mencari jarak pada saat klasifikasi menggunakan metode KNN, salah diantaranya adalah pendekatan euclidean distance yang dapat dilihat pada rumusan dibawah ;



Variabel [ D ] menyimpan nilai jarak, variabel X2i adalah data latih, variabel X1i adalah data uji, variabel p adalah jumlah data yang digunakan. Proses perhitungan jarak antara data latih dan data uji dilakukan berdasarkan nilai ekstraksi fitur pada masing masing data citra, dimana pada penelitian ini proses pengambilan nilai fitur dilakukan menggunakan metode GLCM. Setelah diperoleh jarak antara data uji dengan setiap data latih maka dilakukan proses penentuan kelas dengan cara menghitung jumlah kelas terbanyak dalam rentang nilai K terdekat. Jumlah kelas yang terbanyak akan digunakan sebagai kelas hasil prediksi untuk data yang diprediksi.

Pada beberapa penelitian menggunakan algoritma K-NN dengan ekstraksi fitur GLCM sebelumnya menjunjukkan bahwa kedua metode tersebut dapat bekerja dengan baik dalam melakukan proses klasifikasi. Pada penelitian “implementasi gray level cooccurrence matrix (glcm) untuk klasifikasi penyakit daun padi” diperoleh akurasi tertinggi sebesar 93,3% waktu melakukan klasifikasi 3 kelas daun dengan jumlah data latih sebanyak 210 dan data uji sebanyak 30 [12]. Sedangkan pada penelitian “Klasifikasi Citra Daun dengan GLCM (Gray Level Co-Occurence) dan K-NN (K-Nearest Neighbor) dilakukan klasifikasi terhadap 8 jenis citra daun menggunakan 400 data latih dan 80 data uji, pada proses klasifikasi diperoleh nilai akurasi tertinggi sebesar 98% waktu melakukan klasifikasi menggunakan nilai k = 1, sedangkan nilai akurasi paling kecil diperoleh dengan nilai sebesar 89% waktu menggunakan nilai k = 7 , Sebagai upaya dalam penanganan penyebaran penyakit pada tanaman anggur dilkaukan penelitian untuk menklasifikasikan jenis penyakit daun anggur menggunakan algoritma K-Nearest Neighbor berdasarkan ekstraksi fitur Gray Level Co-occurence Matrix. Pada penelitian ini dilakukan pengambilan data dari dataset plantVillage yang kemudian diolah sehingga diperoleh nilai ekstraksi fitur dari metode GLCM sehingga dapat dilakukan proses klasifikasi menggunakan KNN.

### 2.2.3 Klasifikasi

Menurut Han 2006, Klasifikasi adalah proses membangun model yang berfungsi untuk membedakan kelas data, bertujuan memprediksi kelas dari suatu objek yang belum diketahui kelasnya. Proses klasifikasi terdiri dari dua tahap utama yaitu fase training dan pembentukan aturan klasifikasi. Pada fase training, data pelatihan dianalisis dan direpresentasikan dalam bentuk aturan klasifikasi (Han, 2006). Klasifikasi melibatkan empat komponen utama: kelas, prediktor, data pelatihan(training dataset) dan data pengujian(testing dataset). Kelas adalah label variabel dependen yang mewakili kategori pada suatu objek, seperti jenis gempa, loyalitas pelanggan atau risiko penyakit jantung. Prediktor adalah variabel independen yang menggambarkan karakteristik data, contohnya tabungan, aset, pendapatan, konsumsi alkohol, atau kebiasaan merokok. Data pelatihan berisi nilai dan prediktor untuk menentukan kelas yang sesuai, sementara data pengujian digunakan untuk mengevaluasi akurasi model klasifikasi pada data baru.

### 2.2.4 Morfologi

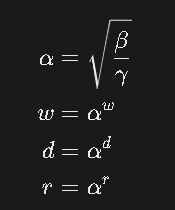
Ekstraksi morfologi adalah metode penting untuk mengidentifikasi penyakit pada daun termasuk daun anggur, dengan menganalisis karakteristik visual seperti bentuk dan tekstur. Teknik ini mencakup fitur seperti eccentricity, solidity, rectangularity, metric, extent, dan elongation untuk mengenali perbedaan morfologi yang spesifik pada daun yang mungkin mengindikasikan infeksi. Pada tahap awal, segmentasi citra dilakukan untuk memisahkan daun dari latarbelakang menggunakan metode thresholding, yang sederhana namun efektif dalam mengisolasi fitur utama daun. Setelah itu, fitur-fitur morfologi diekstraksi dan dianalisis menggunakan algoritma klasifikasi seperti K-Nearest Neighbor (KNN) atau Support Vector Machine (SVM). Pendekatan tersebut terbukti akurat dalam membedakan antara daun sehat dan terinfeksi, menjadikan ekstraksi morfologi sebagai teknik yang efektif dalam deteksi penyakit tanaman. dekat dengan objek tersebut.

Algoritma ini merupakan metode supervised learning dan bekerja dengan cara mengukur jarak terpendek antara data uji dan data latih untuk menentukan kelas data tersebut. Setelah data dikumpulkan dalam kelompok K yang spesifik, kelas mayoritas dalam kelompok tersebut digunakan sebagai kelas prediksi untuk data uji. Algoritma ini memiliki keunggulan dalam ketahanannya terhadap data yang mengandung noise dan efektif pada data latih berjumlah besar. Namun, komputasi algoritma ini cenderung memakan waktu lama jika data latihnya sangat banyak, dan kinerjanya sangat sensitif terhadap ciri-ciri data yang redundan atau kurang. Diagram alur algoritma K-Nearest Neighbors (KNN) dalam klasifikasi citra daun anggur yang mengalami penyakit.

### 2.2.5 EfficientNet

EfficientNet adalah suatu arsitektur jaringan saraf tiruan (neural arrange) yang dikembangkan oleh tim Google pada tahun 2019. Arsitektur ini dirancang untuk meningkatkan efisiensi komputasi dalam demonstrate Convolutional Neural Network (CNN) sementara menjaga atau bahkan meningkatkan akurasi.

Teknik Scaling Multiplier

EfficientNet menggunakan teknik scaling multiplier untuk memperluas ukuran show CNN secara efisien. Teknik ini memungkinkan jaringan untuk diasumsikan secara seragam di seluruh lebar, kedalaman, dan resolusi. Equation yang digunakan untuk skalabilitas multipler adalah:  
  
  
Dimana α,β,γ>1α,β,γ>1.

EfficientNet merupakan solusi yang sangat efektif dalam space pembelajaran mesin, terutama dalam penggunaan CNN. Dengan desain yang efisien dan skalabel, EfficientNet dapat digunakan dalam berbagai aplikasi, mulai dari pengenalan objek hingga diagnosa medis. Meskipun tidak selalu memiliki akurasi tertinggi dibandingkan demonstrate lain, namun performa overallnya yang stabil dan efisien membuatnya populer dalam industri AI saat ini.

**2.2.6 Inception**

Inception adalah arsitektur jaringan saraf konvolusi (CNN) yang dikembangkan oleh tim peneliti di Google, yang pertama kali diperkenalkan dalam makalah "Going Deeper with Convolutions" pada tahun 2014 dan mengalami beberapa pengembangan, termasuk Inception V3 yang dirilis pada tahun 2015. Arsitektur ini dirancang untuk meningkatkan akurasi klasifikasi gambar sambil tetap menjaga efisiensi komputasi.

Fitur Utama Inception  
1. Modul Inception:  
• Modul ini memungkinkan penggunaan beberapa channel konvolusi dengan ukuran yang berbeda secara paralel. Dengan cara ini, jaringan dapat mengekstrak fitur dari gambar pada berbagai skala dan ukuran, menghasilkan representasi yang lebih kaya dari information visualz.

2. Reduksi Dimensi:  
• Inception V3 menggunakan teknik seperti convolutions 1x1 untuk mengurangi jumlah saluran (channels) di lapisan-lapisan yang lebih dalam. Ini membantu mengurangi kompleksitas demonstrate dan meningkatkan kecepatan komputas

3. Group Normalization:  
• Teknik ini diterapkan untuk mempercepat proses pelatihan dan mengurangi ketergantungan pada inisialisasi parameter, sehingga meningkatkan stabilitas dan kecepatan konvergensi pelatihan

4. Aksesori Baru  
• Inception V3 memperkenalkan normal pooling sebelum lapisan softmax untuk membantu mengurangi overfitting dan meningkatkan akurasi pada dataset yang beragam

5. Information Expansion:  
• Selama pelatihan, teknik augmentasi information digunakan untuk meningkatkan kemampuan generalisasi demonstrate dan mencegah overfitting

6. Pre-trained Show:  
• Inception V3 sering kali sudah dilatih sebelumnya pada dataset besar seperti ImageNet, yang memungkinkan demonstrate untuk memperoleh representasi fitur yang baik sebelum dilatih pada tugas spesifik, meningkatkan kinerja dan kecepatan konvergensi

# BAB III

# METODOLOGI PENELITIAN

## 3.1 Desain Penelitian

Tujuan dari penelitian ini adalaha untuk mengembangkan model prediksi risiko penyakit pada tanaman anggur menggunakan algoritma pembelajaran mesin yang dapat mengidentifikasi pada daun tanaman anggur yang memiliki resiko terkena penyakit berdasarkan data yang tersedia. Dataset yang digunakan dalam penelitian ini berisi gambar dari daun tanaman anggur seperti Blackrot,ESCA,Healthy, dan leaf bright yang menyangkut dengan penyakit tanaman anggur. Studi ini membandingkan 2 metode yaitu : convolutional neural network(CNN) dan K-Nearest neighbor (KNN) dengan model EfficientNET dan Inception, hasil yang diinginkan dari studi ini ialah memberikan tingkat akurasi dari 2 metode ini yang memiliki tingkat akurasi yang paling tinggi yang dapat membantu petani budidaya tanaman anggur dalam membudidayakan nya, mengembangakan system prediksi risiko yang lebih cepat dan akurat serta mendukung upaya di rangkaian layanan pertanian.

## 3.2 Data Set

Data yang digunakan pada penelitian ini diambil dari Kaggle dan terdiri dari 4000 citra daun anggur dengan1000 citra untuk masing-masing jenis citra daun1. Datacitra dibagi dengan komposisi 70 % citra pelatihan atausebanyak 2800 citra, 20 % citra validasi atau sebanyak800 citra, dan 10 % citra untuk pengujian atau sebanyak400 citra. Selain itu, pengumpulan data juga dilakukandari citra Googlesebagai data pengujian di luar datasetsebanyak 100 citra. Contoh citra keempat jenis daunanggur ini dapat dilihat pada Gambar 2. Prapemrosesan citra dilakukan setelah data citraRGB dimasukkan. Proses yang dilakukan adalahmelakukan konversi ruang warna citra dari yang semulaRGB menjadi ruang warna citra LAB. Prosesselanjutnya adalah mengekstrak komponen warna B dataset pada citra LAB untuk digunakan pada prosessegmentasi.

A diagram of a plant

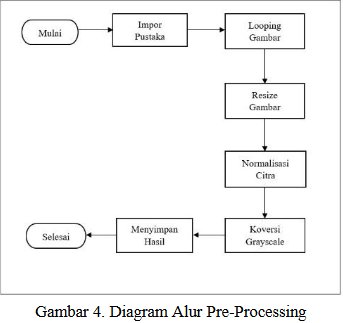
Description automatically generated

### 3.2.1 Metode KNN

Merupakan salah satu metode klasifikasi berbasis supervised learning, dalam melakukan klasifikasi metode ini menggunakan pendekatan dengan menghitung jumlah tetangga paling dekat [17]. Terdapat beberapa pendekatan yang dapat digunakan untuk mencari jarak pada saat klasifikasi menggunakan metode KNN. Proses perhitungan jarak antara data latih dan data uji dilakukan berdasarkan nilai ekstraksi fitur pada masing masing data citra, dimana pada penelitian ini proses pengambilan nilai fitur dilakukan menggunakan metode GLCM. Setelah diperoleh jarak antara data uji dengan setiap data latih maka dilakukan proses penentuan kelas dengan cara menghitung jumlah kelas terbanyak dalam rentang nilai K terdekat. Jumlah kelas yang terbanyak akan digunakan sebagai kelas hasil prediksi untuk data yang diprediksi.

## 3.3 Preprocessing Data

Preprocessing information adalah langkah penting yang dilakukan sebelum information citra daun anggur dapat digunakan dalam demonstrate klasifikasi k-Nearest Neighbors (k-NN). Pada tahap ini, information yang telah dikumpulkan melalui observasi langsung akan diproses dan diolah sehingga lebih siap digunakan untuk pelatihan demonstrate. Preprocessing bertujuan untuk meningkatkan kualitas information dan mengekstraksi fitur-fitur penting dari citra yang dapat digunakan sebagai input bagi algoritma k-NN.

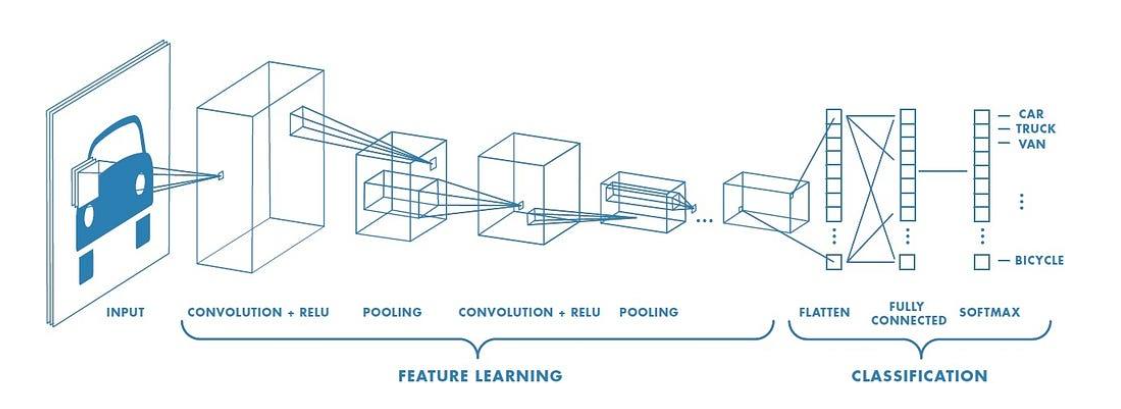


1. Pengolahan Ukuran Citra (Resizing) Citra daun anggur yang diambil melalui observasi langsung diubah ukurannya menjadi 64x64 piksel. Langkah ini bertujuan untuk menyamakan dimensi seluruh citra agar dapat diproses oleh algoritma dengan konsistensi yang baik. Pengubahan ukuran ini juga membantu mempercepat komputasi tanpa menghilangkan terlalu banyak detail penting dari citra.
2. Normalisasi Citra Setelah citra diproses, setiap nilai piksel di-normalisasi dengan membaginya dengan 255, sehingga berada dalam rentang 0 hingga 1. Normalisasi ini penting untuk memastikan bahwa skala nilai tidak terlalu besar, sehingga model dapat berfungsi dengan lebih baik dan tidak terpengaruh oleh rentang nilai yang terlalu lebar.
3. Pengubahan Citra ke GrayscaleUntuk fitur-fitur tertentu, seperti ekstraksi tekstur dan fitur morfologi, citra diubah menjadi grayscale. Pengubahan ini dilakukan karena informasi tekstur dan bentuk lebih mudah dianalisis dalam ruang warna grayscale daripada ruang warna penuh (RGB). Grayscale mengurangi dimensi data dan fokus pada intensitas cahaya, sehingga mempermudah deteksi pola.

### 3.3.1 Convolutional Neural Network (CNN)

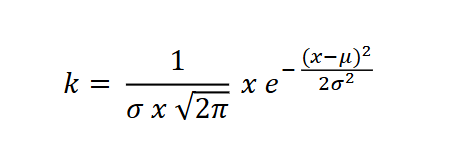
Convolutional Neural Network (CNN) adalah jenis arsitektur jaringan saraf tiruan yang paling sering diterapkan untuk menganalisis citra visual, Karena kemampuan CNN secara otomatis mengekstraksi elemen penting dari data citra maka CNN sangat efektif dalam pengolahan citra dan computer vision(Venkatesan, dkk, 2017).Arsitektur CNN menggunakan lapisan khusus untuk ekstraksi fitur secara hierarkis dari data citra. Konsep kerja CNN memiliki kesamaan dengan Multilayer Perceptron(MLP), namun setiap neuron direpresentasikan dalam bentuk dua dimensi.

CNN memiliki dua proses utama yaitu lapisan deteksi fitur danlapisan klasifikasi. Lapisan deteksi fitur terdapatoperasi konvolusi(+ReLU) dan pooling. Proses ini dilakukan berulang kali pada puluhan atau ratusan lapisan, dan setiap lapisan belajar untuk mengidentifikasi satu atau banyak fitur. Lapisan klasifikasiFully connected layer(FC),menghasilkan vektor dimensi K, di mana K adalah jumlah kelas yang dapat diprediksi jaringan. Probabilitas untuk setiap kelas dari setiap citra yang diklasifikasikan disimpan kedalam vektor ini. Untuk memberikan keluaran klasifikasi, lapisan terakhir arsitektur CNN menggunakan fungsi Softmax(Salman,dkk, 2018).



Gambar 1 ARSITEKTUR CNN

Lapisan konvolusi adalah lapisan yang melakukan operasi konvolusi yaitu kalkulasi perkalian antara citra masukan dan kernel/filter yang menghasilkan matriks baru berupa peta fitur. Filter berupa kisi angka diskrit, memiliki ukuran misalnya panjang 3 piksel dan tinggi 3 piksel, masing-masing elemen merupakan bobot yang diinisialisasi dengan aturan seperti Standard Normal Distibution, Gaussian Blur, Box Blur dan lainnya (Rafael,dkk,2018). Di dalam penelitian ini penentuan kernel menggunakan Standard Normal Distibution, Persamaan (1).



Keterangan :

k : nilai kernel

**x**: variabel random normal

u : mean

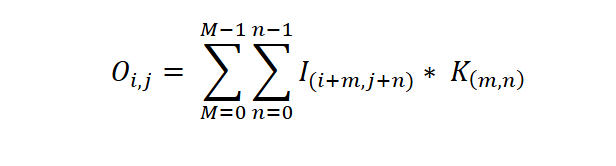
Klasifikasi Penyakit pada Tanaman Anggur Menggunakan Metode Convolutional Neural Network

σ : standard deviasi

π : nilai pi (3.14159265359)

e : bilangan euler (2.718281828459)

Output lapisan konvolusi menghasilkan peta fitur, lihat Persamaan(2), dan terdapat beberapa parameter yang diinisialisasi untuk menghasilkan fitur seperti ukuran kernel, Stride, padding, dan Output Depth(Aniruda, dkk, 2019).



Keterangan

Oi, j : nilai matriks keluaran posisi (i, j)

I(i+m, j+n) : nilai matriks masukan

Symbol ∗ : proses convolution

K(m,n) : nilai kernel

i : koordinat x pada citra

j : koordinat y pada citra

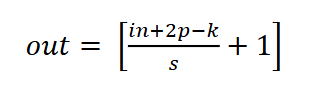
M : lebar citra

N : tinggi citra

m : variabel iterasi untuk lebar citra

n : variabel iterasi untuk tinggi citra

Ukuran citra hasil keluaran dari lapisan konvolusi dihitung menggunakan Persamaan 3



Keterangan :

out : ukuran matriks keluaran

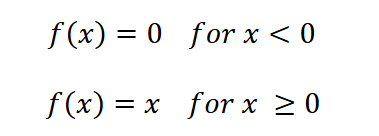
in : ukuran matriks masukan

p : nilai padding

k : nilai kernel

s : nilai stride

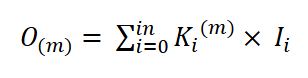
Rectified Linear Unit Layer(ReLU) adalah sebuah fungsi aktivasi untuk mengubah nilai linear menjadi non-linear dengan mengaktifkan dan menonaktifkan neuron (Le Lu, Zheng, Gustavo, & Lin Yang, 2017). Pada persamaan(4),ReLU akan memetakan nilai negatif menjadi nol dan mempertahankan nilainya jika positif. Hal ini memungkinkan CNN untuk memodelkan pola - pola yang lebih kompleks dan meningkatkan kemampuan representasi.



Lapisan pooling digunakan untuk mengurangi dimensi dari peta fitur dengan mereduksi informasi yang tidak penting. Max pooling adalah jenis pooling yang umum digunakan, di mana nilai maksimum dari area tertentu (2x2) diambil sebagai representasi yang lebih ringkas, Persamaan (5).



Lapisan Fully Connected mengubah peta fitur menjadi vektor biasanya ditempatkan di bagian akhir arsitektur. Lapisan ini bertanggung jawab untuk melakukan klasifikasi berdasarkan fitur-fitur yang telah dipelajari sebelumnya. Vektor ini berisi probabilitas untuk setiap kelas dari setiap citra yang diklasifikasikan. Fully Connected Layer tujuan untuk melakukan transformasi pada dimensi data yang dapat diklasifikasi secara linear [8]. Fully Connected Layer memiliki 1 parameter yaitu output depth untuk menentukan panjang keluaran, mewakili setiap nilai kategori hasil klasifikasi. Berikut adalah Persamaan (6)untuk melakukan perhitungan pada Fully Connected Layer.



Keterangan

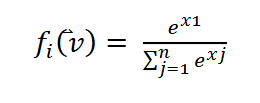
O(m) : nilai matriks keluaran pada posisi ke-m

K : kernel

I : matriks masukan

in : panjang matriks masukan

Lapisan Softmax adalah fungsi klasifikasi linier yang digunakan untuk mendapatkan nilai luaran klasifikasi. Berikut Persamaan (7) untuk mendapatkan nilai luaran probabilitas satu objek.



Keterangan

fi(v) : probabilitas objek ke-i

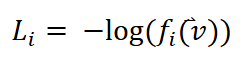
e : bilangan euler bernilai 2.71828183

v : vektor nilai seluruh objek

n : panjang v

i : posisi nilai objek

Softmax merupakan lapisan terakhir dalam arsitektur CNN, berfungsi menghitung nilai loss. Nilai loss ini menunjukkan seberapa besar hasil perhitungan menyimpang dari hasil yang diinginkan dan hasil akhir. Berikut persamaan (8) untuk perhitungan nilai loss pada softmax.



# 3.4 Model Machine Learning

Proyek ini menggunakan beberapa model pembelajaran mesin untuk memprediksi risiko penyakit pada tanaman anggur berdasarkan daun dari data yang tersedia. Tujuan menggunakan model ini adalah untuk menguji pendekatan algoritma yang berbeda dan membandingkan kinerja untuk menemukan metode terbaik untuk memberikan hasil prediksi yang memiliki tingkat akurasi tinggi.

## 3.4.1 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah jenis jaringan saraf tiruan yang dirancang khusus untuk memproses information visual, seperti gambar dan video. CNN menggunakan teknik konvolusi untuk mengekstrak fitur dari input gambar, yang membedakannya dari jaringan saraf biasa yang lebih umum. CNN telah menjadi komponen penting dalam banyak aplikasi teknologi cutting edge, terutama dalam pengenalan objek dan klasifikasi gambar.

Arsitektur CNN terdiri dari beberapa lapisan yang masing-masing memiliki fungsi spesifik:

1. Convolution Layer:

Ini adalah lapisan utama dalam CNN yang melakukan operasi konvolusi pada input. Proses ini melibatkan penggunaan channel atau part yang bergerak di atas gambar untuk mendeteksi fitur-fitur penting seperti tepi dan tekstur. Setiap channel menghasilkan peta fitur (highlight outline) yang merepresentasikan informasi penting dari gambar tersebut.

1. Pooling Layer:

Setelah konvolusi, pooling layer digunakan untuk mengurangi dimensi peta fitur dan mengurangi jumlah parameter yang harus diproses. Teknik pooling yang umum digunakan adalah max pooling, di mana nilai maksimum dari range tertentu diambil untuk mengurangi ukuran peta fitur tanpa kehilangan informasi penting.

1. Completely Associated Layer:

Di bagian akhir arsitektur CNN, terdapat lapisan completely associated (FC) yang menghubungkan semua neuron dari layer sebelumnya ke neuron di layer yield. Lapisan ini bertugas untuk mengklasifikasikan hasil akhir berdasarkan fitur-fitur yang telah diekstraksi sebelumnya.

## 3.4.2 EfficientNet

EfficientNet adalah arsitektur jaringan saraf konvolusional (CNN) yang dibuat untuk meningkatkan akurasi dan penggunaan sumber daya model deep learning. EfficientNet—diumumkan oleh Tan dan Le pada 2019—menggunakan pendekatan penskalaan yang inovatif untuk mengoptimalkan resolusi, kedalaman, dan lebar model secara bersamaan. Ini berbeda dengan mengoptimalkan satu dimensi seperti yang dilakukan oleh banyak arsitektur sebelumnya.  
  
EfficientNet adalah kemajuan besar dalam pengembangan arsitektur CNN. Dengan pendekatan penskalaan yang cerdas dan penggunaan blok MBConv, model ini tidak hanya memiliki akurasi yang tinggi tetapi juga sangat efisien dengan sumber daya. Hal ini membuatnya populer di kalangan peneliti dan praktisi pengajaran mesin untuk digunakan untuk klasifikasi gambar dan analisis data visual lainnya.

## 3.4.3 Inception

Inception adalah arsitektur jaringan saraf konvolusi (CNN) yang dikembangkan oleh tim peneliti Google. Ini pertama kali muncul dalam makalah "Going Deeper with Convolutions" pada tahun 2014 dan kemudian mengalami beberapa versi, salah satunya adalah Inception V3, yang dirilis pada tahun 2015. Arsitektur ini dimaksudkan untuk meningkatkan akurasi klasifikasi gambar sambil meningkatkan efisiensi komputasi.

Inception adalah salah satu inovasi besar dalam bidang deep learning, terutama dalam pengolahan gambar. Arsitektur ini, dengan pendekatan modularnya, dapat menghasilkan representasi fitur yang lebih kaya sambil mempertahankan efisiensi komputasi, yang membuatnya menjadi pilihan yang populer dalam berbagai aplikasi machine learning modern.

3.5 Evaluasi Model

3.6 Implemestasi Sistem

Algoritme yang diterapkan dalam pengolahan citra dengan menggunakan segmentasi k-means clustering,ekstraksi fitur VGG16, dan klasifikasi CNN dapat diterapkan untuk mengklasifikasi citra daun anggur dengan baik. Akurasi yang diperoleh adalah sebesar 99,50 % pada proses pelatihan, 97,25 % pada pengujian 400 citra uji dalam dataset, dan 95 % pada pengujian 100 citra uji di luar dataset.

|  |  |
| --- | --- |
| Kelas | Metrik klasifikasi  Presisi Sensitivitas Skor F1 Support |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Campak hitam | 0,96 | 0,92 | 0,94 | 25 |
| Bercak daun | 0,92 | 0,96 | 0,94 | 25 |
| Daun sehat | 0,93 | 1,00 | 0,96 | 25 |
| Hawar daun | 1,00 | 0,92 | 0,96 5 | 25 |

*Tabel Metrik klasifikasi pada data uji di luar dataset*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kelas | Jumlah citra | | Prediksi | | | |
| Benar | | | Salah |
| Campak hitam | | | 25 | | 23 | 2 | |
| Bercak daun | | | 25 | | 24 | 1 | |
| Daun sehat | | | 25 | | 25 | 0 | |
| Hawar daun | | | 25 | | 23 | 2 | |
| Total | | | 100 | | 95 | 5 | |

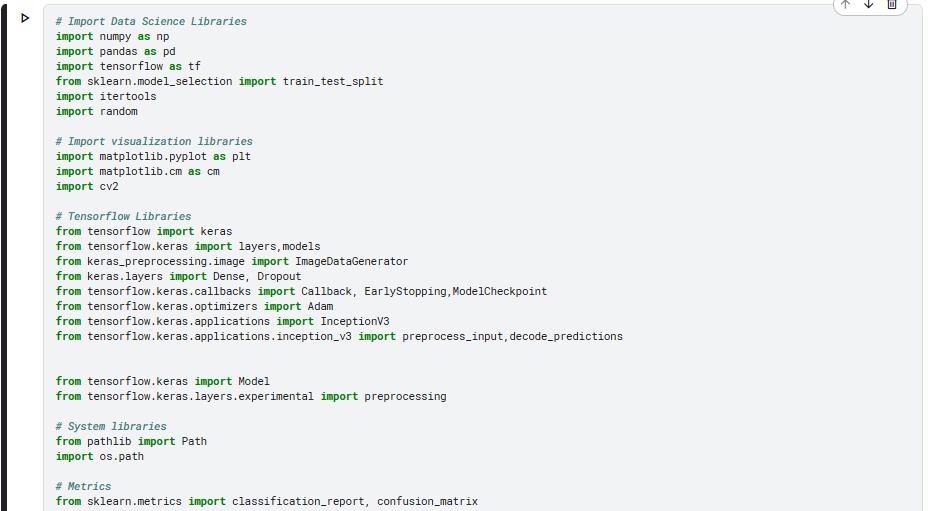
*Tabel Hasil prediksi citra di luar dataset*

# BAB IV

# PEMBAHASAN

## 4.1 IMPORT LIBRARY

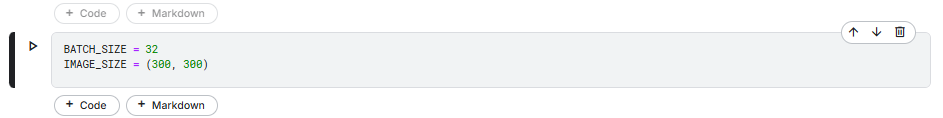
Input :

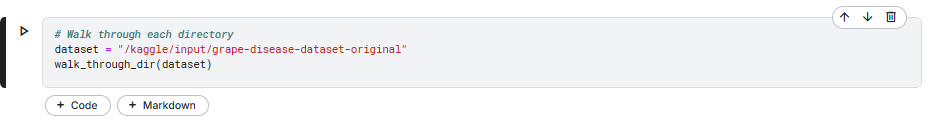


Disini kita import beberapa library yang dibutuhkan untuk melakukan processing data. Ada beberapa library yang digunakan untuk membantu dalam memproses data yang ada.

## 4.2 MEMUAT DATA SET

Input :



Input: :

Output:

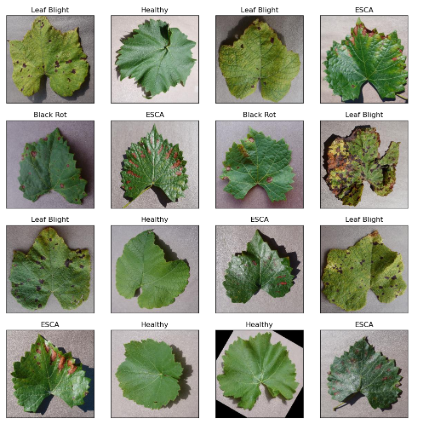
Lalu setelah menambahkan library yang dibutuhkan, kita input dataset yang akan digunakan dalam proses data ini. Data yang digunakan didapat dari KAGGLE (<https://www.kaggle.com/datasets/pushpalama/grape-disease>)

## 4.3 Menampilkan Gambar Dari Dataset

Input:



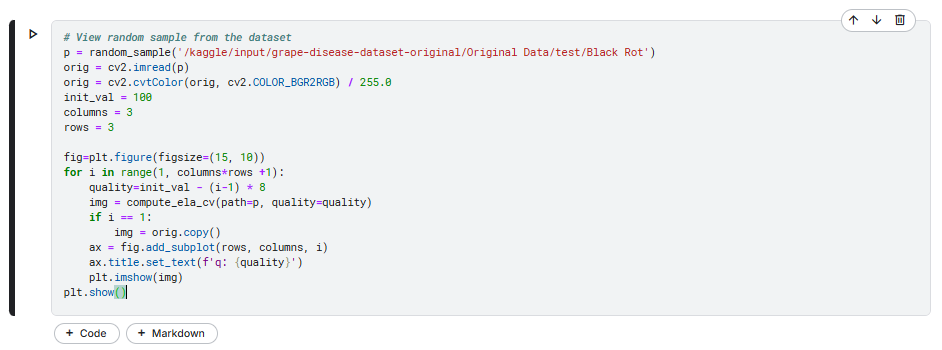
Output:

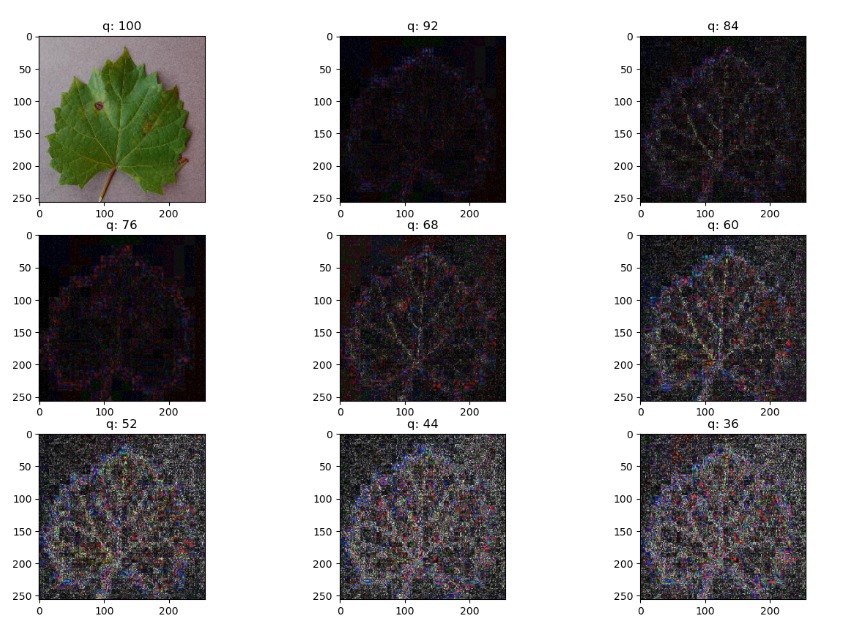


Setelah itu dataset dari Kaggle sudah terinput, kita menampilkan gambar dataset ini , yaitu gambar daun tanaman anggur.

## 4.4 Analisis Tingkat Kesalahan

Input:



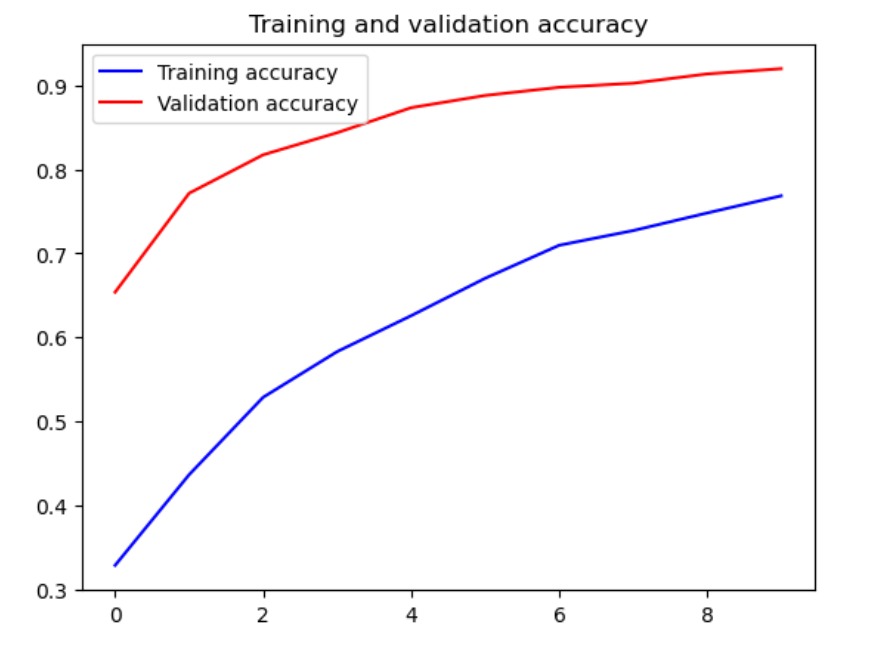
Output:

Pada proses ini menunjukkan serangkaian hasil kompresi pada sebuah gambar daun anggur. Konsep dasarnya adalah mengurangi ukuran gambar dengan cara mengurangi jumlah informasi detail yang terkandung di dalamnya.

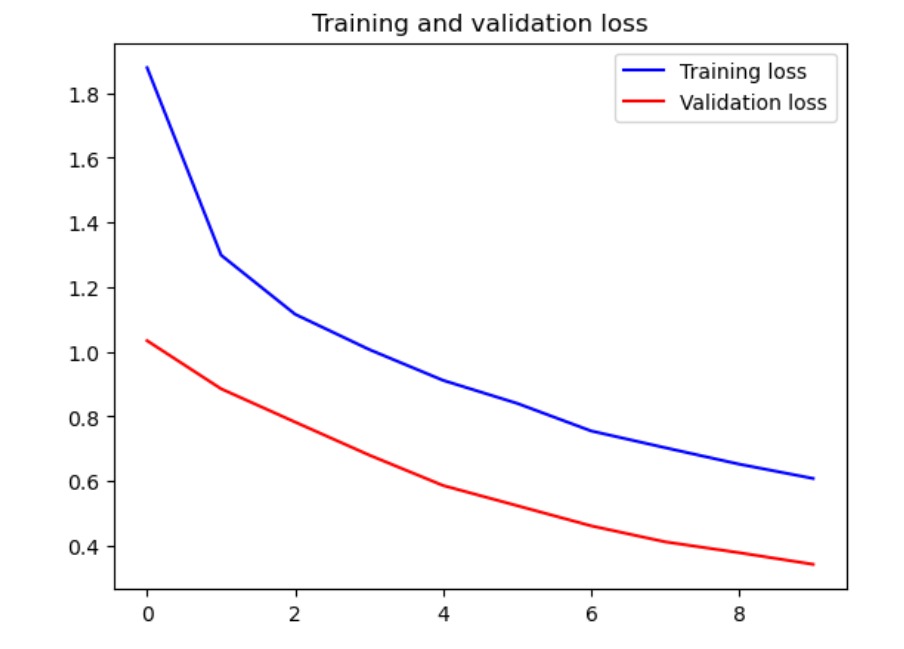
## 4.5 Menampilkan loss curves

Input:



Output1: 1

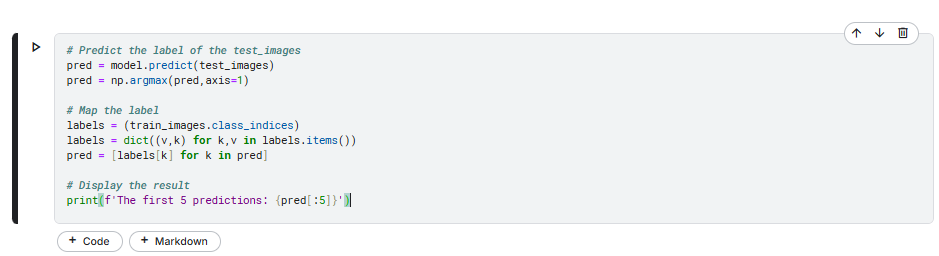
Output 2 :



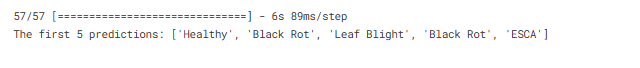
Disini kita menampilkan grafik curves dari dataset yang ada, disitu terlihat ada 2 output. Pada output pertama menampilkan seberapa baik model yang dapat mengklasifikasi data latih dan menunjukan seberapa baik model dapat menggeneralisasikan ke data yang belum pernah dilihat sebelumnya. Sedangakan pada output ke-2 menampilkan seberapa besar kesalahan model dalam memprediksi label kelas pada daa latih dan menunjukan seberapa besar kesalahan model dalam mempredikasi label kelas pada data validasi.

## 4.6 Membuat Prediksi pada Data Test

Input :



Output :



Pada Bagian ini menguji model yang telah dilatih dengan menggunakan data pengujian, yaitu data yang tidak digunakan selama pelatihan. Langkah pertama adalah memuat data pengujian, lalu menjalankan model untuk memprediksi label kelas. Prediksi ini kemudian dibandingkan dengan label asli untuk menghitung akurasi model.

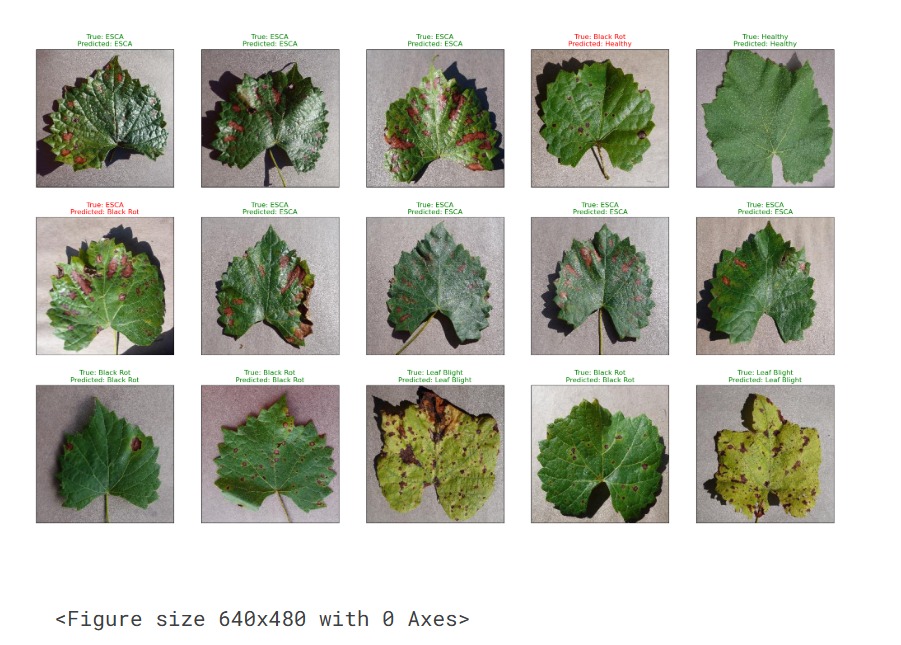
Tujuannya adalah untuk memastikan bahwa model dapat mengenali pola pada data baru dengan baik. Yang Menghasilkan Model label yang diprediksi untuk setiap data dalam set pengujian, serta tingkat akurasi keseluruhan sebagai ukuran kinerja.

## 4.7 Menampilkan 25 gambar random dari dataset yang diberi label

Input:



Output :

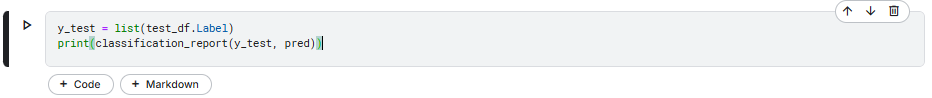


Pada tahap ini menampilkan sampel acak dari dataset untuk menunjukkan bagaimana model memprediksi labelnya. Sebanyak 25 gambar dipilih secara acak dan ditampilkan dalam format grid 5x5. Setiap gambar diberi label prediksi dan label sebenarnya di bawahnya.

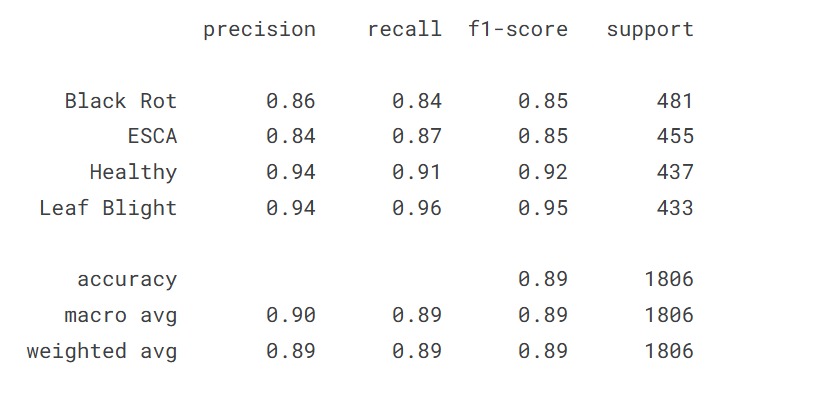
Visualisasi ini berguna untuk memastikan bahwa model bekerja dengan baik dan membantu mendeteksi kesalahan jika label prediksi tidak sesuai dengan label aslinya. Guna Menghasilkan Grid yang terdiri dari 25 gambar beserta label prediksi dan label asli, sehingga mempermudah evaluasi visual terhadap kinerja model

## 4.8 TEST LABEL KLASIFIKASI REPORT DAN CONFUSION MATRIX

Input:



Output :



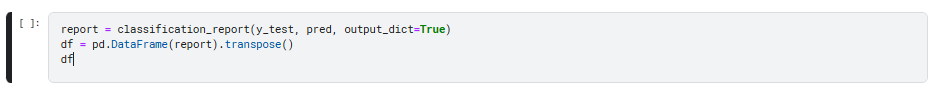
Di bagian ini, model dievaluasi menggunakan dua alat penting: classification report dan confusion matrix. Classification report memberikan ringkasan berupa metrik seperti precision (tingkat ketepatan prediksi), recall (tingkat keberhasilan mengenali data sebenarnya), dan F1-score (kombinasi precision dan recall).

Confusion matrix adalah tabel dua dimensi yang menunjukkan jumlah prediksi benar dan salah untuk setiap kelas. Baris mewakili label asli, sedangkan kolom menunjukkan prediksi model. Matriks ini membantu mengidentifikasi kelas mana yang sering salah diklasifikasikan.

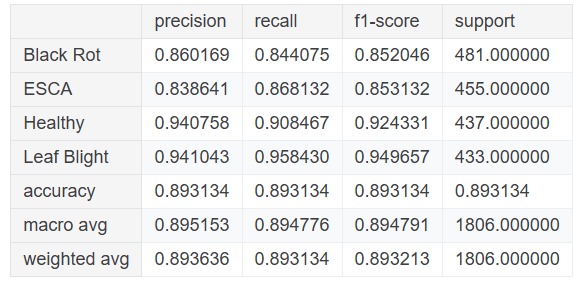
Outputnya Menunjukan Classification report berupa tabel yang merangkum performa model dan confusion matrix yang menggambarkan distribusi prediksi benar dan salah untuk setiap kelas

## 4.9 KLASIFIKASI REPORT

Input:



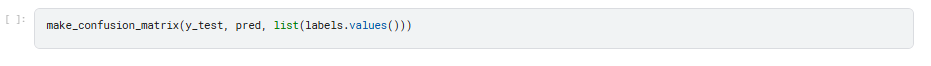
Output



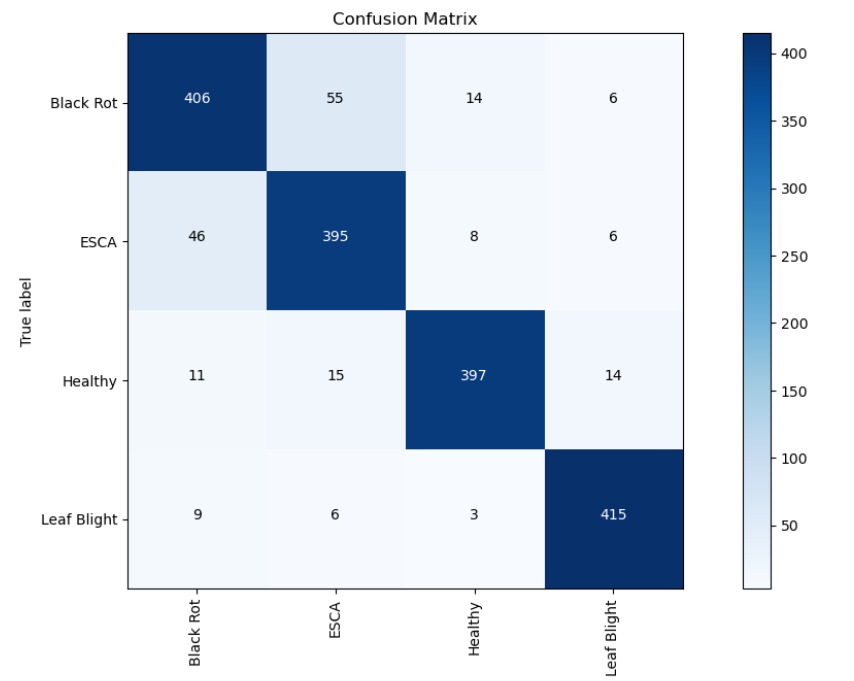
Pada tahap ini hampir sama dengan yang sebelumnya yang membedakan hanya di beberapa bagian dan tampilan outputnya saja. Outputnya Menghasilkan Tabel yang berisi metrik evaluasi untuk setiap kelas, yang memberikan gambaran rinci tentang performa model

## 4.10 Membuat Tabel Confusio matrix

Input:



Output:



Pada Tahap Ini Dilakukan pembuatan sebuah table atau heat yang menampilkan distribusi prediksi model, baik yang benar maupun yang salah. Matriks ini berguna untuk mengidentifikasi pola kesalahan yang konsisten, seperti kelas tertentu yang sering diklasifikasikan ke kelas lain. Informasi ini dapat digunakan untuk memperbaiki model dan meningkatkan performanya.

# BAB V

# KESIMPULAN DAN SARAN

## 5.1 Kesimpulan

Penelitian ini berhasil menunjukkan bahwa penerapan teknologi machine learning, khususnya metode CNN dan KNN, memiliki peran penting dalam mempermudah identifikasi penyakit pada tanaman anggur. CNN terbukti sangat efektif untuk menganalisis dataset besar karena kemampuannya dalam mengenali pola kompleks secara otomatis, menghasilkan tingkat akurasi yang tinggi. Di sisi lain, KNN juga menunjukkan hasil yang baik, terutama ketika dikombinasikan dengan metode ekstraksi fitur seperti Gray Level Co-Occurrence Matrix (GLCM). Meski akurasinya tidak sebaik CNN, metode KNN mudah diterapkan dan lebih fleksibel untuk dataset kecil hingga menengah. Penelitian ini juga menemukan bahwa CNN membutuhkan jumlah data yang besar untuk performa optimal, sementara KNN sensitif terhadap parameter tertentu seperti jumlah tetangga yang digunakan. Dengan demikian, penerapan kedua metode ini tidak hanya mempercepat proses identifikasi penyakit, tetapi juga memberikan solusi praktis dan efisien bagi petani untuk meningkatkan kualitas hasil panen anggur mereka.

## 5.2 Saran

Untuk meningkatkan hasil penelitian ke depan, beberapa rekomendasi perlu dipertimbangkan. Pertama, pengembangan dataset yang lebih besar dan beragam sangat penting untuk memastikan model dapat mengenali pola-pola penyakit secara lebih luas. Kedua, kombinasi metode CNN dan KNN layak untuk dieksplorasi, misalnya dengan memanfaatkan CNN untuk ekstraksi fitur otomatis dan KNN untuk klasifikasi, guna menggabungkan kelebihan kedua metode. Ketiga, sistem yang dikembangkan perlu diuji pada data riil di lapangan untuk mengukur efektivitasnya di luar lingkungan laboratorium. Keempat, optimalisasi parameter seperti pengaturan hyperparameter pada CNN dan pemilihan nilai k yang optimal pada KNN dapat meningkatkan performa model secara signifikan. Selain itu, integrasi teknologi tambahan seperti drone atau perangkat Internet of Things (IoT) dapat membantu dalam proses pengumpulan data gambar tanaman anggur secara otomatis. Dengan menerapkan langkah-langkah ini, solusi yang dihasilkan akan menjadi lebih efektif dan relevan bagi kebutuhan petani dalam mengelola tanaman anggur mereka.

# DAFTAR PUSTAKA

[1] H. Gohzali and S. Irviantina, “Implementation of Fuzzy K-Nearest Neighbor (K-NN) Algorithm to Identify Grape Plant Diseases,” *Instal J. Komput.*, vol. 15, no. June, 2023, [Online]. Available: https://journalinstal.cattleyadf.org/index.php/Instal/article/view/97%0Ahttps://journalinstal.cattleyadf.org/index.php/Instal/article/download/97/40

[2] M. A. Hasan, Y. Riyanto, and D. Riana, “Grape leaf image disease classification using CNN-VGG16 model,” *J. Teknol. dan Sist. Komput.*, vol. 9, no. 4, pp. 218–223, 2021, doi: 10.14710/jtsiskom.2021.14013.

[3] B. Liu, Z. Ding, L. Tian, D. He, S. Li, and H. Wang, “Grape Leaf Disease Identification Using Improved Deep Convolutional Neural Networks,” *Front. Plant Sci.*, vol. 11, no. July, pp. 1–14, 2020, doi: 10.3389/fpls.2020.01082.

[4] M. Ji, L. Zhang, and Q. Wu, “Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks,” *Inf. Process. Agric.*, vol. 7, no. 3, pp. 418–426, 2020, doi: 10.1016/j.inpa.2019.10.003.

[5] K. R. Wardani and L. Leonardi, “Klasifikasi Penyakit pada Daun Anggur menggunakan Metode Convolutional Neural Network,” *J. Tekno Insentif*, vol. 17, no. 2, pp. 112–126, 2023, doi: 10.36787/jti.v17i2.1130.

[6] D. A. N. K. Neighbors, “PENERAPAN METODE MACHINE LEARNING UNTUK PREDIKSI RISIKO PENYAKIT JANTUNG : PERBANDINGAN AKURASI ANTARA LOGISTIC REGRESSION , SUPPORT VECTOR CLASSIFICATION , DECISION TREE , ADABOOST , PROGRAM STUDI TEKNIK INFORMATIKA,” 2024.

[7] F. M. Fathoni, C. A. Putra, and A. L. Nurlaili, “Klasifikasi Penyakit Daun Anggur Menggunakan Metode K-Nearest Neighbor Berdasarkan Gray Level Co-Occurrence Matrix,” *Biner J. Ilm. Inform. dan Komput.*, vol. 3, no. 1, pp. 8–15, 2024, doi: 10.32699/biner.v3i1.6332.

[8] A. Naik, “Comparative Analysis of VGG16, RESNET50, AND CNN Models for Lung Disease Prediction: A Deep Learning Approach,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 12, no. 5, pp. 875–896, 2024, doi: 10.22214/ijraset.2024.61703.

[9] Anissa Ollivia Cahya Pratiwi, “Klasifikasi Jenis Anggur Berdasarkan Bentuk Daun Menggunakan Convolutional Neural Network Dan K-Nearest Neighbor,” *J. Ilm. Tek. Inform. dan Komun.*, vol. 3, no. 2, pp. 201–224, 2023, doi: 10.55606/juitik.v3i2.535.