BAB I
PENDAHULUAN

1.1 Latar Belakang

Stomata merupakan pori-pori mikroskopis pada permukaan daun dan batang tanaman yang berperan penting dalam pertukaran gas, yaitu penyerapan karbon dioksida (CO_2) untuk fotosintesis serta pelepasan oksigen (O_2) dan uap air melalui transpirasi. Fungsi stomata sangat krusial dalam menjaga keseimbangan air dan suhu tanaman, serta menjadi indikator penting dalam studi adaptasi tanaman terhadap lingkungan. Karakteristik seperti ukuran dan kerapatan stomata mencerminkan efisiensi metabolisme tumbuhan dan digunakan dalam analisis fisiologis, terutama pada kondisi cekaman lingkungan seperti kekeringan dan peningkatan CO_2 atmosferik (Haworth et al., 2023).

Stomata memiliki peran yang sangat penting dalam mengatur keseimbangan antara proses fotosintesis dan transpirasi tanaman. Melalui mekanisme buka-tutup sel penjaga (*guard cells*), stomata mengontrol difusi gas seperti karbon dioksida (CO_2) untuk fotosintesis dan pelepasan uap air ke atmosfer. Efisiensi fungsi stomata berpengaruh langsung terhadap produktivitas dan adaptasi tanaman terhadap stres lingkungan. Menurut Feng et al., (2021), variasi ukuran dan kerapatan stomata merupakan respons fisiologis yang adaptif terhadap perubahan intensitas cahaya dan ketersediaan air. Penelitian oleh Mohanasundaram et al., (2022) juga menunjukkan bahwa spesies tanaman dengan kerapatan stomata tinggi memiliki kemampuan fotosintesis lebih efisien dalam kondisi CO_2 rendah, tetapi lebih rentan terhadap kehilangan air. Sementara itu, Huang et al. (2023) menegaskan bahwa pengaturan morfologi dan distribusi stomata menjadi indikator utama dalam menilai ketahanan tanaman terhadap kekeringan serta perubahan iklim global. Dengan demikian, analisis karakteristik stomata, seperti ukuran, bentuk, dan kerapatan, tidak hanya penting untuk memahami fungsi fisiologis tanaman, tetapi juga menjadi dasar bagi pemuliaan tanaman yang lebih adaptif terhadap lingkungan yang berubah.

Pengamatan stomata secara manual masih sering digunakan dalam banyak penelitian. Namun, proses manual ini memiliki keterbatasan dalam hal efisiensi waktu, subjektivitas penilai, serta potensi kesalahan dalam mengidentifikasi struktur stomata terutama dalam jumlah besar atau citra berkualitas rendah (Nature Research Intelligence., 2023). Seiring dengan berkembangnya teknologi citra digital dan kecerdasan buatan, muncul peluang untuk mengotomatisasi proses ini secara lebih akurat, efisien, dan konsisten.

Urgensi penelitian dalam otomatisasi deteksi stomata berbasis kecerdasan buatan semakin meningkat dalam beberapa tahun terakhir. Hal ini didorong oleh kebutuhan untuk memperoleh data morfologi stomata secara cepat, akurat, dan dalam jumlah besar guna mendukung analisis fisiologis tanaman. Sultana et al., (2021) menunjukkan bahwa penerapan *deep learning* pada pengenalan stomata daun kedelai mampu meningkatkan efisiensi pengolahan data hingga lima kali lipat dibandingkan metode konvensionalLebih lanjut, Murphy et al., (2025) menegaskan bahwa penerapan

kecerdasan buatan dalam *image-based plant phenotyping* menjadi kunci dalam memahami respon adaptif tanaman terhadap stres lingkungan. Oleh karena itu, penelitian mengenai otomatisasi deteksi stomata berbasis AI menjadi semakin mendesak untuk dikembangkan, terutama sebagai langkah menuju sistem analisis morfologi tanaman yang efisien, objektif, dan dapat diandalkan dalam berbagai kondisi lingkungan.

Salah satu pendekatan yang menjanjikan adalah dengan menggunakan metode deteksi objek berbasis *deep learning*. *YOLO* (*You Only Look Once*) merupakan algoritma deteksi objek satu tahap yang terkenal karena kecepatannya dan akurasi tinggi dalam memproses gambar secara *real-time*. Penelitian sebelumnya menunjukkan bahwa *YOLOv5* berhasil mendeteksi stomata dengan cukup baik pada citra mikrograf tanaman seperti jagung, beras, dan kapas (Qiu et al., 2022). Seiring dengan evolusinya, *YOLO* kini telah mencapai versi ke-11 (*YOLOv11*), dengan berbagai peningkatan arsitektur seperti modul *C2f*, *attention mechanism*, dan *Spatial Pyramid Pooling Fast* (SPPF) yang secara signifikan meningkatkan kemampuan deteksi objek kecil seperti stomata.

Namun, penggunaan algoritma deteksi objek tidak terlepas dari kualitas citra yang diolah. Citra mikroskopis sering kali mengalami berbagai kendala visual seperti pencahayaan yang tidak merata, *noise*, dan kontras rendah yang menyulitkan proses deteksi. Oleh karena itu, diperlukan proses prapengolahan (*preprocessing*) citra guna meningkatkan kualitas visual sebelum dilakukan klasifikasi. Dua metode praproses yang umum digunakan adalah *Gamma Correction* dan *CLAHE* (*Contrast Limited Adaptive Histogram Equalization*). *Gamma Correction* bertujuan untuk menyesuaikan kecerahan citra melalui manipulasi distribusi intensitas piksel, sedangkan *CLAHE* meningkatkan kontras lokal tanpa memperkuat *noise* secara berlebihan.

CLAHE dan *Gamma Correction* telah terbukti meningkatkan performa berbagai model deteksi, khususnya dalam domain medis. Misalnya, Benchabane & Charif (2025) menggabungkan *CLAHE* dan Gamma *Correction* dalam analisis *X-ray COVID-19* menggunakan arsitektur DenseNet201,. Penerapan Dalam konteks botani, *CLAHE* juga telah digunakan untuk menajamkan tepi sel epidermis dan membuka struktur stomata dengan lebih jelas, sehingga membantu dalam proses deteksi otomatis. Kombinasi keduanya memiliki potensi besar, namun penerapannya pada klasifikasi stomata menggunakan arsitektur *YOLOv11* masih belum banyak dijelajahi dalam literatur terkini.

Lebih lanjut, penelitian terkini oleh Zhang et al., (2023) memperkenalkan metode *rotating stomata detection* dengan model *DeepRSD*, yang mampu mendeteksi orientasi acak stomata dan mengukur *stomatal conductance* secara otomatis. Metode ini tidak hanya mendeteksi bentuk, tetapi juga menyertakan fungsi kerugian khusus untuk *conductance*, menghasilkan akurasi hingga 94,3 % pada daun jagung menandakan bahwa klasifikasi stomata memiliki peran penting dalam evaluasi fisiologis tanaman yang lebih kompleks

Dengan demikian, penelitian ini bertujuan untuk menggabungkan kekuatan dari arsitektur deteksi terkini (*YOLOv11*) dan dua teknik peningkatan kualitas citra (*Gamma Correction* dan *CLAHE*) untuk membangun sistem klasifikasi stomata yang unggul.

Melalui evaluasi performa dengan metrik seperti *accuracy*, *precision*, *recall*, dan *mAP*, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pengembangan sistem analisis citra mikroskopis berbasis AI untuk bidang fisiologi tanaman dan pertanian presisi.

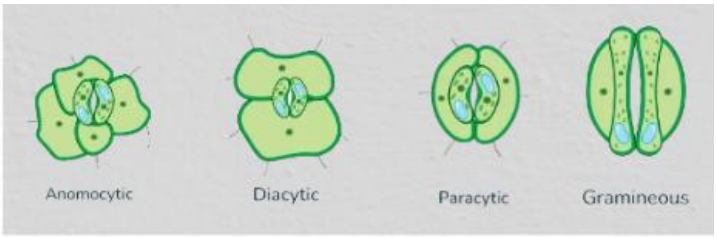
1.2 Landasan Teori

1.2.1 Stomata

Stomata adalah pori-pori mikroskopis yang umumnya terdapat pada permukaan epidermis daun dan batang muda tanaman. Struktur ini dibentuk oleh dua sel penjaga (*guard cells*) yang mengapit celah sempit yang disebut pori stomata. Fungsi utama stomata adalah sebagai jalur pertukaran gas antara jaringan internal tumbuhan dan atmosfer, termasuk dalam proses fotosintesis dan respirasi (Zhang et al., 2022).

Stomata memiliki peran penting dalam pengambilan karbon dioksida (CO_2) dari atmosfer untuk fotosintesis serta dalam pelepasan oksigen (O_2) dan uap air hasil transpirasi. Oleh karena itu, stomata menjadi salah satu komponen kunci dalam regulasi gas dan keseimbangan air pada tumbuhan. Dalam konteks fisiologi tanaman, parameter stomata seperti jumlah, ukuran, dan kerapatan sering digunakan sebagai indikator dalam menilai efisiensi metabolisme, adaptasi terhadap lingkungan, serta ketahanan terhadap stres abiotik (Nature Research Intelligence., 2023).

Stomata dapat diklasifikasikan berdasarkan susunan dan jenis sel tetangga yang mengelilingi sel penjaga. Beberapa tipe morfologi stomata yang umum ditemukan dalam berbagai spesies tanaman antara lain:



Gambar 1. Tipe Stomata

Source: https://www.geeksforgeeks.org/biology/types-of-stomata/

1. **Anomositik (*Anomocytic* (A))**: Stomata ini dikelilingi oleh sel-sel epidermis biasa yang tidak memiliki ciri khusus, ukuran dan bentuknya tidak berbeda nyata dari sel lainnya.
2. **Diastik (*Diacytic* (D))**: Stomata ini dikelilingi oleh dua sel tetangga yang posisinya saling tegak lurus terhadap celah stomata.
3. **Parasitik (*Paracytic* (P))**: Memiliki satu atau dua sel tetangga yang letaknya sejajar dengan arah celah stomata.

4. **Graminoid *(Graminaceous* (G))**: Stomata memiliki sel penjaga berbentuk dumbbell (seperti barbel), khas pada tanaman rumput.

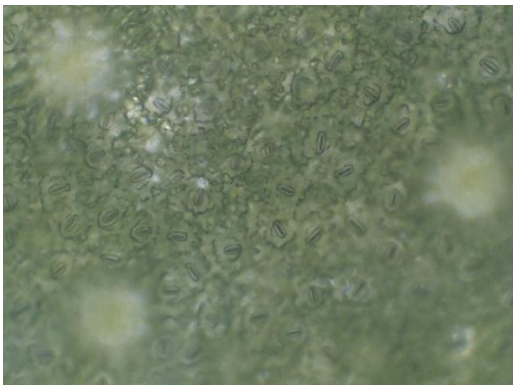
Jenis dan distribusi morfologi stomata ini bervariasi antar spesies tanaman dan sering digunakan dalam kajian taksonomi, ekologi, serta identifikasi adaptasi terhadap lingkungan. Dalam praktik ilmiah, pengamatan stomata juga penting untuk mendukung studi-studi lanjutan dalam bidang fisiologi tanaman, agronomi, dan bioteknologi tanaman.

Karena karakteristik stomata dapat dijadikan parameter fisiologis yang penting, maka proses identifikasi dan klasifikasi stomata dari data citra menjadi fokus penting dalam banyak penelitian. Upaya untuk mempercepat dan mempermudah proses ini kini banyak dikembangkan melalui teknologi pengolahan citra digital dan kecerdasan buatan.

1.2.2 Mikroskop dan Mikrograf

Mikroskop adalah alat optik yang digunakan untuk memperbesar objek kecil agar dapat diamati secara visual. Mikroskop cahaya merupakan jenis yang paling umum digunakan dalam studi biologis, termasuk pengamatan stomata, karena dapat menampilkan struktur jaringan seperti sel epidermis secara jelas tanpa memerlukan teknik preparasi yang kompleks. Selain itu, mikroskop cahaya dapat digunakan bersama kamera digital untuk menghasilkan citra mikroskopis atau mikrograf (Gibbs et al., 2021).

Mikrograf adalah citra hasil tangkapan mikroskop yang digunakan sebagai data visual dalam analisis. Dalam penelitian ini, mikrograf stomata menjadi *input* utama bagi sistem klasifikasi otomatis berbasis *deep learning*. Namun, kualitas mikrograf sangat memengaruhi hasil deteksi. Masalah seperti kontras rendah atau pencahayaan tidak merata sering muncul, sehingga diperlukan prapengolahan citra untuk meningkatkan kualitas visual, adapaun contoh mikrograf stomata pada gambar 2.



Gambar 2. Mikrograf Stomata

Teknik prapengolahan yang umum digunakan adalah *CLAHE* untuk meningkatkan kontras lokal, dan *Gamma Correction* untuk menyesuaikan pencahayaan. Kombinasi kedua teknik ini terbukti efektif meningkatkan kejelasan objek mikroskopis dan mendukung akurasi klasifikasi (Chang et al., 2018). Penelitian oleh Gibbs et al., (2021)

menunjukkan bahwa mikrograf berkualitas tinggi yang telah diproses mampu meningkatkan akurasi deteksi stomata secara signifikan.

1.2.3 Pengolahan Citra

Pengolahan citra digital (*digital image processing*) adalah serangkaian proses matematis dan algoritmik yang diterapkan pada citra digital untuk meningkatkan kualitas visual, mengekstrak informasi, atau mempersiapkan citra untuk analisis lanjutan oleh sistem komputer. Dalam konteks penelitian ini, pengolahan citra digital digunakan untuk memperjelas struktur stomata pada citra mikroskopik sebelum dilakukan klasifikasi menggunakan metode *deep learning* (Gonzalez & Woods, 2022).

Pengolahan citra sangat penting dalam sistem berbasis pembelajaran mesin atau deep learning karena kualitas input sangat menentukan hasil klasifikasi atau deteksi. Pada penelitian ini, tahap praproses berperan memastikan fitur visual stomata tampil jelas dan konsisten sehingga lebih mudah dikenali oleh YOLOv11. Dengan mengurangi noise, menyeimbangkan pencahayaan, dan menajamkan struktur penting, model dapat mengekstraksi fitur secara lebih akurat dan menghasilkan deteksi yang lebih stabil.

1.2.4 Visi Komputer

Visi komputer adalah cabang dari kecerdasan buatan yang memungkinkan komputer untuk menafsirkan informasi visual dari citra atau video secara otomatis. Tujuannya adalah agar sistem komputer dapat “melihat” dan memahami objek seperti manusia, dengan cara mengekstrak dan menganalisis fitur-fitur penting dari gambar digital (Szeliski, 2022).

Dalam penelitian ini, visi komputer digunakan untuk mendeteksi dan mengklasifikasikan stomata dari citra mikroskopis secara otomatis. Pendekatan ini menggantikan metode manual yang cenderung lambat dan subjektif. Dengan bantuan algoritma *deep learning* seperti *YOLOv11*, sistem dapat mengenali objek kecil seperti stomata secara cepat dan akurat. Komponen utama dalam sistem visi komputer mencakup akuisisi citra, praproses, ekstraksi fitur, dan klasifikasi. Semua tahapan ini saling terhubung dan mendukung proses analisis citra secara menyeluruh, mulai dari pengambilan gambar hingga interpretasi hasil.

1.2.5 Deteksi Objek

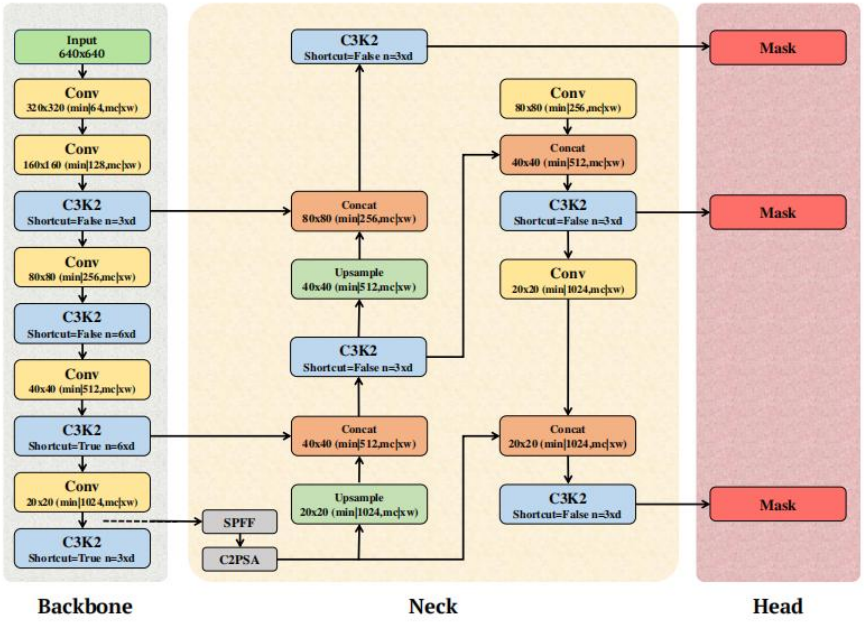
Deteksi objek (*object detection*) adalah cabang dari visi komputer yang berfokus pada identifikasi dan lokalisasi objek-objek tertentu dalam gambar atau video. Proses ini mencakup pengenalan jenis objek serta penentuan posisi objek tersebut dalam gambar. Deteksi objek merupakan langkah penting dalam berbagai aplikasi seperti pengawasan keamanan, kendaraan otonom, analisis video, dan realitas tertambah (*Augmented Reality*) (Zou et al., 2019).

Adapun komponen utama dalam deteksi objek dimulai dengan proses untuk menyiapkan gambar agar lebih mudah dianalisis. Teknik ini termasuk penyesuaian

kontras, pengurangan *noise*, dan transformasi gambar untuk meningkatkan kualitas gambar dan membuat fitur penting lebih menonjol. Mengidentifikasi dan mengekstrak fitur penting dari gambar, seperti tepi, sudut, tekstur, dan pola warna. Fitur ini digunakan sebagai *input* untuk algoritma deteksi objek. Menggunakan model pembelajaran mesin atau pembelajaran mendalam untuk mengenali dan mengklasifikasikan objek dalam gambar berdasarkan fitur yang diekstraksi. Menentukan lokasi objek dalam gambar, biasanya dengan menggambar kotak pembatas (*bounding box*) di sekitar objek yang terdeteksi.

1.2.6 *YOLOv11* (*You Only Look Once*)

YOLO (*You Only Look Once*) merupakan salah satu algoritma deteksi objek paling populer dalam bidang visi komputer. Pendekatan *YOLO* mengklasifikasikan dan melakukan prediksi lokasi objek dalam satu tahap secara langsung pada keseluruhan citra, sehingga menghasilkan sistem deteksi yang cepat dan efisien (Redmon et al., 2016). Sejak pertama kali dikembangkan, *YOLO* terus mengalami penyempurnaan, dan versi terbarunya yaitu *YOLOv11* menawarkan peningkatan baik dari sisi akurasi maupun efisiensi, khususnya dalam mendeteksi objek kecil.



Gambar 3. Arsitektur *YOLO*

Source: https://medium.com/thedeephub/the-architecture-of-yolov11-a-leap-in-object-detection-2ce1e946c74a

Gambar 3 menampilkan arsitektur *YOLOv11,* dimana *YOLOv11* merupakan pengembangan dari arsitektur *YOLOv8* yang dilengkapi dengan beberapa inovasi, seperti modul *C2f* yang menggantikan *C3* untuk ekstraksi fitur yang lebih ringan, *SPPF++*

untuk menangkap informasi *spasial multi-skala*, serta *anchor-free detection head* yang memungkinkan fleksibilitas deteksi terhadap objek dengan ukuran dan bentuk yang bervariasi. Kombinasi modul-modul tersebut menjadikan *YOLOv11* mampu mendeteksi objek kecil seperti stomata dalam citra mikroskopis secara lebih akurat.

YOLOv11 juga meningkatkan proses *Non-Maximum Suppression* (NMS) untuk menghindari deteksi objek yang tumpang tindih, yang sebelumnya menjadi salah satu keterbatasan dalam pendeteksian multi-objek kecil dalam satu bidang citra. Model ini dilatih menggunakan teknik *augmentasi* data dan *loss function* yang telah disesuaikan untuk menyeimbangkan antara klasifikasi dan prediksi lokasi objek.

Beberapa penelitian terkini menunjukkan potensi besar *YOLOv11* dalam bidang botani dan fisiologi tanaman. Misalnya, penelitian oleh (Yang et al., 2025) mengembangkan *StomaYOLO*, model ringan berbasis *YOLOv11* untuk mendeteksi sel penjaga stomata pada daun jagung dengan *mean Average Precision* (*mAP*) mencapai 91,8%, menunjukkan efisiensi tinggi dalam mendeteksi objek mikroskopis dengan variasi bentuk yang kompleks. Selain itu, (Tang et al., 2025) melalui model *YOLOv11-AIU* berhasil menerapkan arsitektur *YOLOv11* untuk deteksi penyakit *early blight* pada tomat, yang juga menyoroti kemampuan model ini dalam mengidentifikasi pola morfologi kecil pada jaringan tanaman. Hasil-hasil tersebut menegaskan bahwa *YOLOv11* memiliki potensi tinggi untuk diterapkan pada analisis citra mikroskopis tanaman, termasuk dalam deteksi dan klasifikasi stomata.

Tabel 1. Perbandingan *YOLOv8-Yolov11*

Versi	***Precision***	***Recall***	***mAP***	***mAP@50–95***	***Val Box Loss***	***Val Cls Loss***	***Val DFL Loss***	***Computation Time (s)***
YOLOv8	0.91931	0.88889	0.94618	0.50840	1.59822	0.71588	1.28544	8376.94
YOLOv9	0.91602	0.89876	0.95256	0.51298	1.57651	0.65708	1.27104	10337.31
YOLOv10	0.85307	0.84475	0.91822	0.51182	3.10694	1.46670	2.50162	7524.26
YOLOv11	0.93348	0.88631	0.95152	0.51488	1.57081	0.71905	1.28165	4542.97

Source: Utiarahman et al., 2025.
https://doi.org/10.1109/NMITCON65824.2025.11188393

Penelitian terkini oleh Utiarahman et al., (2025) melakukan evaluasi komprehensif terhadap empat arsitektur *YOLO* generasi terbaru, yaitu *YOLOv8*, *YOLOv9*, *YOLOv10*, dan *YOLOv11*, dalam mendeteksi serta mengklasifikasikan tipe morfologi stomata pada citra mikroskopis daun tanaman herbal. Berdasarkan hasil pengujian kuantitatif, *YOLOv11* menunjukkan performa terbaik dengan nilai *precision* tertinggi sebesar 0.93348, *mAP* sebesar 0.95152, dan *mAP@50–95* sebesar 0.51488, disertai waktu komputasi tercepat yakni 4.542,97 detik. Model ini juga memiliki nilai *validation loss* yang stabil (*Val Box Loss*: 1.57081; *Val Cls Loss*: 0.71905; *Val DFL Loss*: 1.28165), menandakan efisiensi tinggi dalam proses prediksi.

Sebagai perbandingan, *YOLOv9* memperoleh recall tertinggi sebesar 0.89876 serta *validation loss* terendah, menandakan kemampuannya yang baik dalam mendeteksi seluruh objek stomata secara menyeluruh. Namun demikian, *YOLOv11* dinilai lebih unggul secara keseluruhan karena mampu mencapai keseimbangan optimal antara akurasi deteksi, efisiensi komputasi, dan kestabilan hasil. Sementara itu, *YOLOv10* menunjukkan performa paling rendah akibat nilai *loss* yang tinggi, sedangkan *YOLOv8* masih menunjukkan hasil kompetitif namun belum melampaui dua versi teratas.

Berdasarkan hasil tersebut, dapat disimpulkan bahwa *YOLOv11* menjadi model paling tepat untuk digunakan dalam penelitian ini karena mampu memberikan hasil deteksi yang akurat dengan waktu pemrosesan yang jauh lebih cepat. Keunggulan ini menjadikan *YOLOv11* sebagai model terkini berperforma unggul dalam otomasi analisis citra mikroskopis, khususnya untuk klasifikasi morfologi stomata yang menuntut presisi tinggi dan efisiensi waktu.

Dalam penelitian ini, *YOLOv11* dipilih karena kemampuannya dalam mendeteksi objek kecil, dengan efisiensi tinggi dan arsitektur yang lebih ringan dibandingkan model dua tahap seperti *Faster R-CNN*. Model ini digunakan untuk mendeteksi dan mengklasifikasikan stomata pada citra mikroskopis, khususnya setelah citra ditingkatkan kualitasnya menggunakan teknik *Gamma Correction* dan *CLAHE*.

1.2.7 *Gamma Correction*

Gamma Correction adalah teknik pemrosesan citra digital yang digunakan untuk menyesuaikan kecerahan (*brightness*) dan kontras gambar dengan mengoreksi hubungan *nonlinier* antara intensitas cahaya yang ditangkap oleh sensor dan yang ditampilkan oleh perangkat (W. Wang et al., 2021). Secara umum, gambar digital mentah yang diperoleh dari sensor cenderung terlihat terlalu gelap atau terlalu terang jika tidak dilakukan koreksi gamma. Proses *Gamma Correction* dilakukan dengan menerapkan fungsi:

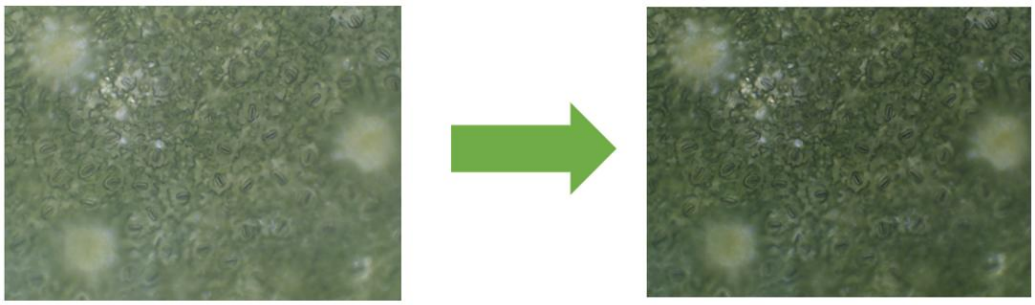
$$\boldsymbol{I_{out} = \frac{I_{in}}{255}} \quad (1)$$

di mana:
I_{in} adalah intensitas asli dari piksel,
I_{out} adalah intensitas keluaran setelah koreksi,
γ adalah nilai gamma (jika γ<1, gambar menjadi lebih terang; jika γ>1, gambar menjadi lebih gelap).

Seiring dengan perkembangan teknologi pengolahan citra dan pembelajaran mendalam, penerapan *Gamma Correction* telah berevolusi dari teknik linier sederhana menjadi pendekatan adaptif berbasis jaringan saraf. Penelitian terbaru oleh (Y. Wang et al., 2023) melalui *Illumination-Aware Gamma Correction and Complete Image Modelling Network* menunjukkan bahwa koreksi gamma adaptif yang mempertimbangkan distribusi pencahayaan lokal mampu meningkatkan kualitas citra secara signifikan pada kondisi cahaya rendah. Pendekatan ini tidak hanya memperbaiki kecerahan, tetapi juga mempertahankan detail tekstur penting yang sebelumnya hilang akibat pencahayaan

tidak merata. Sementara itu, Weng et al., (2024) mengembangkan *Lightweight Low-Light Image Enhancement Network* yang menggabungkan *Gamma Correction* dengan *channel prior*, menghasilkan peningkatan visibilitas citra dengan efisiensi komputasi tinggi, sehingga cocok untuk aplikasi *real-time*. Penelitian-penelitian tersebut menunjukkan bahwa *Gamma Correction* kini bukan sekadar tahap praproses konvensional, melainkan telah menjadi bagian integral dalam pipeline pembelajaran mendalam untuk peningkatan kualitas citra. Dalam konteks penelitian ini, penerapan *Gamma Correction* pada citra mikrograf stomata diharapkan dapat meningkatkan kontras dan visibilitas struktur mikroskopis sehingga mendukung performa model deteksi berbasis YOLOv11 secara lebih optimal.

Penerapan *Gamma Correction* sangat penting dalam pengolahan citra mikroskopis, termasuk mikrograf stomata, karena membantu meningkatkan visibilitas detail struktur halus yang tidak terlihat jelas pada pencahayaan asli (Jeon et al., 2024). Dalam konteks deteksi objek menggunakan *deep learning* seperti *YOLO*, kualitas pencahayaan sangat memengaruhi akurasi deteksi. Oleh karena itu, *Gamma Correction* digunakan sebagai salah satu tahap praproses untuk memastikan kontras dan distribusi cahaya optimal pada gambar *input* sebelum dilakukan pelatihan model (Yu et al., 2021). Gambar 4 menampilkan penerapan *gamma correction.*



(a) Gambar Asli (b) Gambar Hasil Gamma

Gambar 4. Penerapan *Gamma Correction*

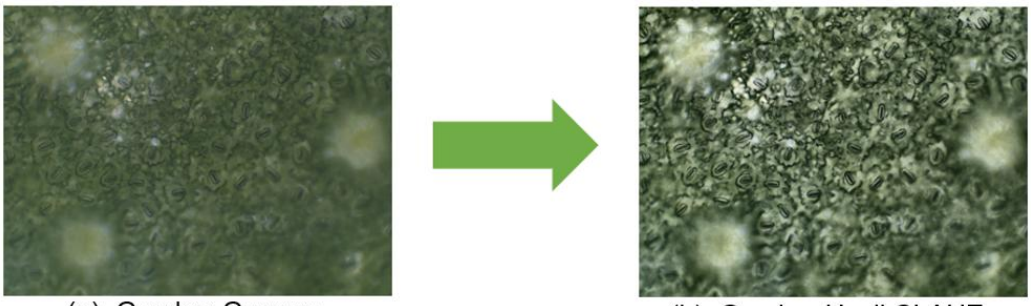
1.2.8 *CLAHE* (*Contrast Limited Adaptive Histogram Equalization*)

CLAHE adalah metode peningkatan kontras citra yang bekerja secara lokal dengan membagi gambar ke dalam beberapa bagian kecil (*tiles*), kemudian melakukan *equalization* pada masing-masing bagian. Berbeda dengan metode histogram *equalization* biasa, *CLAHE* menggunakan parameter *clip limit* untuk mencegah peningkatan kontras yang berlebihan, sehingga mampu mengurangi efek *noise*.

Metode *Contrast Limited Adaptive Histogram Equalization* (CLAHE) telah banyak digunakan dalam peningkatan kualitas citra karena kemampuannya memperbaiki kontras lokal tanpa memperkuat noise secara berlebihan. Dalam penelitian terbaru ŞENGÖZ et al., (2022) memperlihatkan bahwa penggunaan CLAHE sebagai bagian dari pipeline meningkatkan skor F1 dari 93 % menjadi 98 % dalam klasifikasi gambar histopatologi ketika dibandingkan model *CNN* tanpa preprocessing. Selanjutnya, Huang Guoand Xu, (2021)mengembangkan model enhancement citra yang

menggabungkan operator diferensial fraksional orde variabel dengan CLAHE, untuk memperkuat tepi dan tekstur citra.

Dalam beberapa penelitian terkini, penggunaan *CLAHE* terbukti meningkatkan performa deteksi pada citra medis dan mikroskopis, termasuk citra stomata, dengan cara memperjelas detail morfologi yang sulit terlihat (Liu et al., 2021).Oleh karena itu, dalam penelitian ini, *CLAHE* digunakan sebagai tahap praproses untuk meningkatkan kejelasan struktur stomata sebelum dilakukan klasifikasi oleh model *YOLOv11*. Gambar 5 menampilkan penerapan *CLAHE.*



(a) Gambar Gamma (b) Gambar Hasil CLAHE

Gambar 5. Penerapan *CLAHE*

1.3 Perumusan Masalah

Berdasarkan latar belakang masalah di atas, maka rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Sejauh mana model *YOLOv11* mampu mengklasifikasikan stomata dengan tingkat akurasi yang tinggi?
2. Bagaimana pengaruh penerapan *Gamma Correction* dan *CLAHE* terhadap peningkatan kinerja klasifikasi citra stomata pada model *YOLOv11*?

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Tujuan yang ingin dicapai pada penelitian ini yaitu:

1. Meningkatkan akurasi kalsifikasi stomata yang memiliki warna hampir sama dengan daun menggunakan algoritma *YOLOv11*, *Gamma Correction* dan *CLAHE*.
2. Mengevaluasi pengaruh teknik peningkatan *Gamma Correction* dan *CLAHE* terhadap performa model *YOLOv11* dalam mengklasifikasi stomata berbagai kondisi pencahayaan dan latar belakang yang kompleks.

1.4.2 Manfaat

Adapun manfaat yang dapat diperoleh dari penelitian ini yaitu:

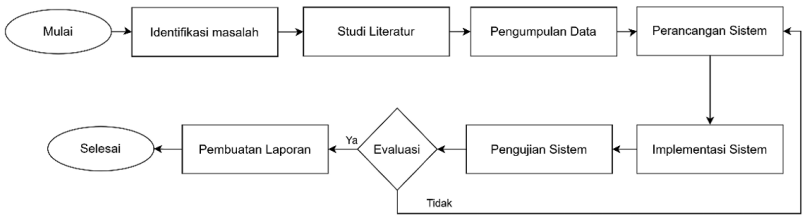
1. Menyediakan sistem klasifikasi stomata yang otomatis, cepat, dan akurat untuk mendukung analisis tanaman.
2. Menjadi referensi dalam penerapan kombinasi *deep learning* dan peningkatan citra untuk objek mikroskopis di bidang informatika dan pertanian.

BAB II
METODE PENELITIAN

2.1 Jenis Penelitian

Jenis penelitian ini merupakan penelitian yang bersifat eksperimental dan analisis sehingga dapat dilakukan dengan metode studi pustaka (*library research*), melalui *beberapa* literatur seperti buku, jurnal, pro siding atau artikel ilmiah lainnya, metode pengumpulan data lapangan (*field research*), perancangan sistem, dan analisis menggunakan metode pengolahan citra (*Image Processing*).

2.2 Tahapan Penelitian



Gambar 6. Tahapan Penelitian

Pada gambar 6 diatas akan di paparkan tahapan penelitian, tahapan yang diusulkan adalah sebagai berikut:

Berikut merupakan penjelasan tentang tahapan-tahapan yang dilakukan selama proses penelitian pada Gambar:

1. Identifikasi masalah adalah tahap awal untuk mengidentifikasi permasalahan yang melatarbelakangi penelitian, yaitu rendahnya akurasi klasifikasi stomata pada citra mikroskopik akibat kualitas citra yang buruk (kontras rendah pencahayaan tidak merata). Permasalahan ini berdampak pada kinerja algoritma deteksi, khususnya dalam konteks citra mikroskopik biologis.
2. Setelah permasalahan diidentifikasi, tahap selanjutnya adalah melakukan studi literatur untuk mendalami teori dan pendekatan yang relevan dengan penelitian ini. Studi ini mencakup penelusuran terhadap teknik-teknik pengolahan citra digital seperti *Gamma Correction* dan *CLAHE*, serta studi mendalam terhadap algoritma *YOLO*, khususnya versi terbaru yaitu *YOLOv11* yang dikenal memiliki keunggulan dalam kecepatan dan akurasi deteksi. Studi literatur ini juga membantu dalam merancang kerangka metodologi penelitian dan mendukung argumen pemilihan metode yang digunakan.
3. Pada tahap ini dilakukan pengumpulan data set berupa citra mikrograf stomata dari berbagai jenis daun. Citra diambil menggunakan mikroskop digital dengan resolusi tinggi untuk memastikan detail stomata terekam dengan baik. Proses pengambilan data dilakukan secara sistematis, dengan memperhatikan pencahayaan dan posisi fokus agar diperoleh variasi data yang merepresentasikan kondisi riil di lapangan.

Data set kemudian dilabeli secara manual untuk keperluan pelatihan dan pengujian model deteksi.

4. Setelah data tersedia, dilakukan perancangan sistem secara menyeluruh, mulai dari alur praproses citra, penerapan algoritma *YOLOv11*, hingga evaluasi performa sistem. Praproses mencakup penerapan *Gamma Correction* untuk meningkatkan pencahayaan dan *CLAHE* untuk meningkatkan kontras lokal pada citra. Desain sistem juga mencakup struktur direktori data set, pemisahan data *train-test*, serta parameter *training* model. Perancangan ini bertujuan agar sistem yang dibangun dapat menangani tantangan deteksi stomata pada citra yang telah ditingkatkan kualitasnya secara efektif.
5. Hasil rancangan sistem kemudian diimplementasikan menggunakan bahasa pemrograman *Python* dengan bantuan pustaka pendukung seperti *OpenCV* untuk pengolahan citra dan *PyTorch* atau *Ultralytics* untuk pemodelan *YOLOv11*. Modul-modul yang dibangun meliputi tahapan *preprocessing* (*Gamma Correction* dan *CLAHE*), pelatihan model *YOLOv11* menggunakan data set hasil peningkatan citra, serta sistem evaluasi untuk mengukur performa deteksi. Proses implementasi dilakukan secara bertahap dan modular untuk memudahkan *debugging* serta pengembangan lebih lanjut.
6. Sistem yang telah dikembangkan selanjutnya diuji menggunakan data uji yang telah disiapkan sebelumnya. Pengujian dilakukan dengan tujuan untuk mengevaluasi performa model dalam mendeteksi dan mengklasifikasikan stomata dari citra hasil praproses. Evaluasi dilakukan menggunakan metrik seperti *accuracy*, *precision*, *recall*, *F1-score*, dan *mean Average Precision* (*mAP*). Hasil dari pengujian ini digunakan untuk menilai efektivitas penerapan *Gamma Correction* dan *CLAHE* terhadap peningkatan performa deteksi menggunakan *YOLOv11*.
7. Tahap evaluasi dilakukan untuk menganalisis apakah sistem yang dibangun telah mencapai target performa yang diinginkan. Jika hasil pengujian menunjukkan performa yang belum optimal, maka dilakukan revisi pada tahap perancangan sistem atau proses pelatihan ulang dengan parameter yang disesuaikan. Evaluasi ini penting untuk memastikan bahwa sistem deteksi stomata yang dibangun benar-benar efektif dalam kondisi nyata dan tidak hanya unggul dalam lingkungan pengujian yang terbatas.
8. Tahap akhir dari proses penelitian ini adalah penyusunan laporan secara sistematis dan menyeluruh. Laporan mencakup latar belakang, rumusan masalah, studi literatur, metodologi penelitian, implementasi sistem, hasil pengujian, evaluasi, dan kesimpulan. Laporan disusun sesuai dengan format penulisan ilmiah yang berlaku sebagai dokumen pertanggungjawaban dari proses dan hasil penelitian yang telah dilakukan.

2.3 Sumber Data

Sumber data yang digunakan dalam penelitian ini adalah citra mikrograf stomata yang diperoleh menggunakan kamera digital yang terpasang pada mikroskop cahaya di Laboratorium AIMP, Universitas Hasanuddin. Data ini digunakan untuk mengklasifikasikan stomata daun dengan bantuan metode *deep learning*. Citra dari

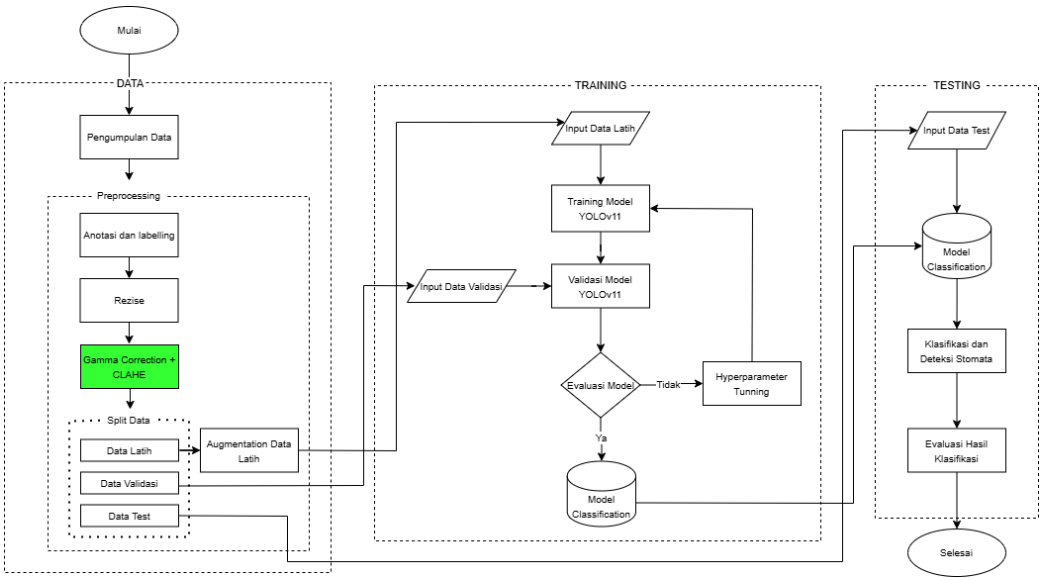
kamera ini kemudian diolah menggunakan teknik *Gamma Correction* dan *CLAHE* untuk meningkatkan kualitas visual sebelum diproses oleh model *YOLOv11*. Data dikumpulkan secara langsung di laboratorium serta didukung oleh data set sekunder dari sumber terbuka sebagai bagian dari proses pelatihan dan pengujian sistem.

2.4 Perangkat Penelitian

1. Spesifikasi Perangkat Keras (*Lenovo Legion 5 Pro 16*)
   a. *NVIDIA RTX 4070 8 GB*
   b. *Intel Core i9 13900HX*
   c. *RAM 64 GB*
   d. *SSD 2 TB*
2. Perangkat Lunak
   a. Sistem Operasi Windows 11-64 bit
   b. Bahasa Pemrograman Python
   c. Jupyter Notebook
   d. Roboflow
   e. Microsoft Office

2.5 Rancangan Sistem

Alur perancangan sistem pada penelitian ini ditunjukkan pada gambar berikut ini:

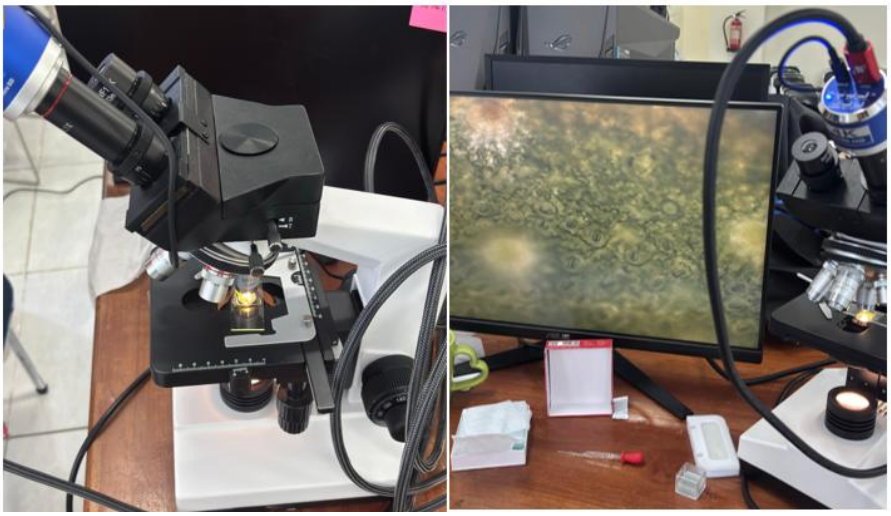


Gambar 7. Rancangan Sistem

Berdasarkan gambar 7 rancangan sistem, proses dilakukan sesuai alur yang telah dirancang untuk mencapai tujuan klasifikasi stomata yang diuraikan sebagai berikut:

2.5.1 Pengumpulan Data

Data dikumpulkan dalam bentuk citra mikrograf stomata dari daun tanaman yang diperoleh menggunakan kamera digital yang terpasang pada mikroskop cahaya. Proses pengambilan data dapat dilihat pada gambar 8.



Gambar 8. Proses Pengambilan Data

2.5.2 *Preprocessing*

a. Anotasi dan *Labeling*

Setelah data citra dikumpulkan, dilakukan proses anotasi dengan menandai letak stomata menggunakan *bounding box* dan memberikan label berdasarkan jenis atau morfologi stomata seperti yang ada pada gambar 9. Tahap ini penting agar model *YOLOv11* dapat belajar mengenali dan membedakan jenis stomata berdasarkan posisi dan klasifikasinya.

Gambar 9. Proses Anotasi Labelling

b. *Rezise*

Citra yang telah dianotasi kemudian diubah ukurannya (*resize*) ke dimensi standar *YOLOv11*, seperti 640x640 piksel. Selanjutnya, citra diproses menggunakan *Gamma Correction* dan *CLAHE* untuk meningkatkan kontras dan kecerahan agar struktur stomata lebih terlihat jelas.

2.5.3 Gamma Correction

Sebelum proses deteksi dilakukan, citra uji terlebih dahulu diproses menggunakan *gamma correction* untuk meningkatkan kecerahan (*brightness*) akibat optik mikroskop. Tujuan dari tahap ini adalah untuk meningkatkan cahaya struktur stomata agar dapat dikenali lebih akurat oleh model.

e. Ambil nilai intensitas piksel asli

Tahap ini menunjukkan matriks nilai piksel asli dari kanal *green* citra berukuran 3×3. Setiap elemen merepresentasikan tingkat kecerahan warna hijau pada skala 0–255. Nilai tinggi (misalnya 255) menunjukkan piksel terang, sedangkan nilai rendah (misalnya 30) menunjukkan piksel gelap.

$$\boldsymbol{G} = \begin{bmatrix} \mathbf{200} & \mathbf{150} & \mathbf{100} \\ \mathbf{50} & \mathbf{128} & \mathbf{220} \\ \mathbf{30} & \mathbf{180} & \mathbf{255} \end{bmatrix} \tag{2}$$

f. Normalisasi nilai piksel

Setiap nilai piksel dibagi dengan 255 agar berada pada rentang 0–1. Langkah ini penting karena banyak operasi pemrosesan citra—termasuk *Gamma Correction* bekerja lebih stabil pada domain ter-normalisasi, dengan rumus:

$$\boldsymbol{I_{out}} = \frac{\boldsymbol{I_{in}}}{\mathbf{255}} \tag{3}$$

Maka hasil normalisasi:

$$\boldsymbol{I_{out}} = \begin{bmatrix} \mathbf{0.784} & \mathbf{0.588} & \mathbf{0.392} \\ \mathbf{0.196} & \mathbf{0.502} & \mathbf{0.863} \\ \mathbf{0.118} & \mathbf{0.706} & \mathbf{1.000} \end{bmatrix} \tag{4}$$

g. Terapkan rumus *Gamma Correction*

Rumus $I_{out} = {I_{in}}^{\gamma}$ digunakan untuk memperbaiki kecerahan citra. Karena nilai $\gamma = 1.5 > 1$, hasilnya membuat piksel gelap menjadi lebih gelap dan piksel terang sedikit lebih cerah. Proses ini menyesuaikan distribusi intensitas agar detail pada area terang lebih menonjol, maka:

$$\boldsymbol{I_{out}} = {\boldsymbol{I_{in}}}^{\mathbf{1.5}} \tag{5}$$

Hasilnya:

$$\boldsymbol{I_{out}} = \begin{bmatrix} (\mathbf{0.784})^{\mathbf{1.5}} & (\mathbf{0.588})^{\mathbf{1.5}} & (\mathbf{0.392})^{\mathbf{1.5}} \\ (\mathbf{0.196})^{\mathbf{1.5}} & (\mathbf{0.502})^{\mathbf{1.5}} & (\mathbf{0.863})^{\mathbf{1.5}} \\ (\mathbf{0.118})^{\mathbf{1.5}} & (\mathbf{0.706})^{\mathbf{1.5}} & (\mathbf{1.000})^{\mathbf{1.5}} \end{bmatrix} \quad (6)$$

$$= \begin{bmatrix} \mathbf{0.693} & \mathbf{0.451} & \mathbf{0.245} \\ \mathbf{0.087} & \mathbf{0.356} & \mathbf{0.802} \\ \mathbf{0.041} & \mathbf{0.593} & \mathbf{1.000} \end{bmatrix} \quad (7)$$

h. Kembalikan ke skala 0–255

Setelah proses *Gamma Correction* selesai, hasil yang semula dalam rentang 0–1 dikembalikan ke skala intensitas citra asli dengan mengalikan 255:

$$\boldsymbol{I_{out}} = \boldsymbol{I_{in}} \times \mathbf{255} \quad (8)$$

Sehingga hasil akhirnya:

$$\boldsymbol{I_{out}} = \begin{bmatrix} \mathbf{177} & \mathbf{115} & \mathbf{62} \\ \mathbf{22} & \mathbf{91} & \mathbf{204} \\ \mathbf{10} & \mathbf{151} & \mathbf{255} \end{bmatrix} \quad (9)$$

2.5.4 CLAHE (*Contrast Limited Adaptive Histogram Equalization)*

Setelah melalui proses perbaikan kecerahan, citra uji kemudian diproses menggunakan *CLAHE* untuk meningkatkan kontras warna (*color contrast*). Tujuan dari tahap ini adalah untuk meningkatkan warna struktur stomata agar dapat dikenali lebih akurat oleh model.

a. Menghitung Histogram

Histogram dibuat dengan menghitung frekuensi kemunculan setiap nilai intensitas pada matriks keluaran I. Karena matriks 3×3 berisi sembilan nilai unik (10, 22, 62, …, 255) dan masing-masing muncul satu kali, tabel histogram menunjukkan frekuensi 1 untuk setiap nilai. Histogram ini memberi gambaran distribusi intensitas di kasus ini distribusi sangat jarang karena semua nilai unik.

$$\boldsymbol{I} = \begin{bmatrix} \mathbf{177} & \mathbf{115} & \mathbf{62} \\ \mathbf{22} & \mathbf{91} & \mathbf{204} \\ \mathbf{10} & \mathbf{151} & \mathbf{255} \end{bmatrix} \quad (10)$$

Tabel 2. Nilai Histogram

Nilai Intensitas (I)	Frekuensi (P)
10	1
22	1
62	1
91	1
115	1
151	1
177	1
204	1
255	1

b. Hitung *CDF* (*Cumulative Distribution Function*)

CDF diperoleh dengan melakukan akumulasi frekuensi histogram dari nilai intensitas terendah ke tertinggi. Untuk tiap nilai I, *CDF*(I) adalah jumlah frekuensi semua intensitas $\leq I$. Karena tiap nilai frekuensinya 1, *CDF* bertambah satu tiap baris: 1, 2, 3, …, 9. *CDF* menunjukkan proporsi kumulatif piksel sampai ambang tertentu dan digunakan untuk meratakan atau meregangkan distribusi intensitas.

Tabel 3. Nilai *CDF*

I	P	*CDF*
10	1	1
22	1	2
62	1	3
91	1	4
115	1	5
151	1	6
177	1	7
204	1	8
255	1	9

c. Normalisasi *CDF* (Manual *CLAHE*)

Normalisasi CDF (Manual *CLAHE*) rumus dan maknanya (Persamaan 11 dan Tabel 3), rumus *CLAHE* manual yang digunakan adalah:

$$\boldsymbol{I}_{\text{baru}} = \frac{(\text{CDF}(\boldsymbol{I}) - \text{CDF}_{\boldsymbol{min}})}{(\boldsymbol{N} - \text{CDF}_{\boldsymbol{min}})} \ (\boldsymbol{L} - \mathbf{1}) \quad (11)$$

Dimana:

$\mathbf{I}_{\text{baru}}$: Nilai intensitas baru setelah peningkatan kontras.

CDF(I): Nilai kumulatif histogram pada intensitas Idi dalam satu blok.

$\text{CDF}_{\min}$: Nilai CDF terkecil pada blok tersebut sebagai acuan normalisasi.

N: Jumlah total piksel dalam blok (*tile*).

L: Jumlah level intensitas.

Dengan $L = 256,\ N = 9,\ \text{CDF}_{min} = 1$

Rumus ini memetakan nilai *CDF* ke rentang 0–255 secara linier; mengurangi CDF_{min}*m*enghindari pemetaan semua nilai ke 0 jika CDF terendah bukan nol. Hasilnya adalah nilai baru untuk setiap intensitas input yang bertujuan menyebarkan kembali intensitas sehingga kontras lokal meningkat. Tabel 3 menunjukkan hasil pemetaan *CDF* ke nilai baru (misalnya CDF=2 → 32, CDF=5 → 128, dst).

Tabel 4. Nilai Hasil Perhitungan *CLAHE*

I	P	*CDF*	Nilai Baru
10	1	1	0
22	1	2	32
62	1	3	64
91	1	4	96
115	1	5	128
151	1	6	160
177	1	7	192
204	1	8	224
255	1	9	255

d. Terapkan Nilai Baru ke Matriks Awal

Setiap elemen pada matriks intensitas awal digantikan dengan nilai baru sesuai tabel pemetaan *CLAHE* sehingga terbentuk matriks hasil:

$$\boldsymbol{G} = \begin{bmatrix} \mathbf{192} & \mathbf{128} & \mathbf{64} \\ \mathbf{32} & \mathbf{96} & \mathbf{224} \\ \mathbf{0} & \mathbf{160} & \mathbf{255} \end{bmatrix} \quad (12)$$

Langkah ini menghasilkan citra dengan kontras yang diregangkan: area gelap menjadi lebih gelap (misalnya 10 → 0) dan area terang dipertegas (misalnya 255 tetap

255), sehingga perbedaan antara tingkat intensitas menjadi lebih jelas dan detail lokal lebih terlihat.

2.5.5 *Split Data*

Data set kemudian dibagi menjadi tiga bagian yaitu data latih untuk proses pelatihan model, data validasi untuk mengevaluasi performa model selama pelatihan, dan data uji untuk mengukur performa akhir model. Pembagian ini dilakukan secara proporsional untuk mencegah *overfitting* dan meningkatkan generalisasi model terhadap data baru. Pada tabel 5 dipaparkan pembagian dataset, dan tabel 6 menampilkan hasil jumlah label dataset.

Tabel 5. *Split* Dataset

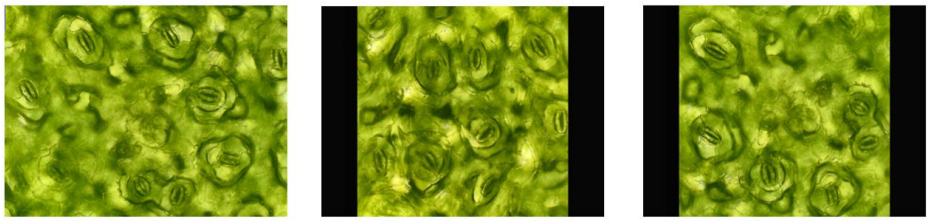
Class	***Train***
Train	1164
Valid	325
Test	174
Total	1663

Tabel 6. Distribusi Label Dataset

Class	***Train***	***Valid***	***Test***	**Total**
Anomocytic	6773	1303	513	8589
Diacytic	14158	3111	896	18165
Graminoid	5992	1276	664	8232
Paracytic	13571	2291	1688	17550
Total	40494	7981	3761	52236

2.5.6 Augmentasi Data Latih

Untuk menambah jumlah dan variasi data latih, dilakukan augmentasi citra. Teknik augmentasi yang digunakan meliputi rotasi dan *flipping* (pencerminan). Tahapan ini bertujuan untuk membuat model lebih tangguh terhadap perubahan posisi dan pencahayaan saat mendeteksi stomata pada kondisi yang berbeda. Setelah di augmentasi dataset *train* menjadi 3374 dari 1164. Dengan distribusi yang dipaparkan pada tabel 7.



(a) Gambar Asli (b) Hasil Augmentasi *Rotate* (c) Hasil Augmentasi *Flipping*

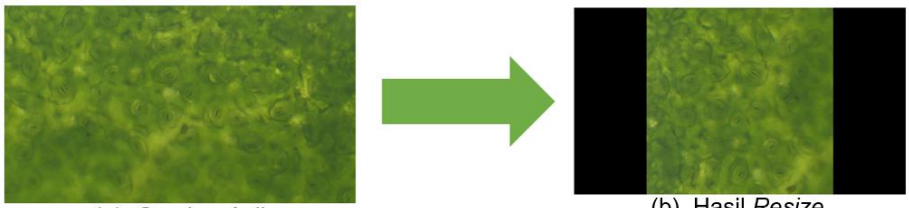
Gambar 10. Penerapan *Augmentasi*

Tabel 7. Distribusi *Augmentasi* Data *Training*

Class	***Train***
Anomocytic	12823
Diacytic	28780
Graminoid	11683
Paracytic	26004
Total	**79290**

2.5.7 *Input* Data Training

Data latih yang telah melalui proses *augmentasi* dan *preprocessing* dimasukkan ke dalam proses pelatihan model *YOLOv11*. Model akan mempelajari fitur visual dari stomata berdasarkan *bounding box* dan label yang telah ditentukan, sehingga mampu melakukan klasifikasi stomata secara otomatis dengan tingkat akurasi yang tinggi.



(a) Gambar Asli (b) Hasil *Resize*

Gambar 11. Inputan YOLOv11

Citra masukan berukuran 640 × 640 × 3 (RGB). Setiap piksel dinormalisasi ke rentang [0,1] menggunakan persamaan:

$$\boldsymbol{I_{norm} = \frac{I_{rgb}}{255}} \tag{13}$$

Misal pada satu piksel di koordinat (0,0) memiliki nilai

$R = 120, \quad G = 200, \quad B = 150$

Maka hasil normalisasinya:

$\mathrm{R} = 0.47, \mathrm{G} = 0.78, \mathrm{B} = 0.59$

2.5.8 *Training* Model *YOLOv11*

Model *YOLOv11* mulai dilatih menggunakan data latih. Pada tahap ini, model mempelajari fitur visual dari stomata, termasuk bentuk, struktur, dan pola distribusinya, agar mampu mendeteksi dan mengklasifikasikan stomata secara otomatis.

a. Lapisan Konvolusi

Proses konvolusi, yang berfungsi untuk mengekstraksi fitur-fitur penting dari citra masukan. Lapisan pertama YOLOv11 melakukan konvolusi 3 kanal menjadi 32

fitur dengan *kernel* 3×3, *stride* 2, dan *padding* 1.
Rumus konvolusi:

$$\mathbf{Y(i,j,k) = \Sigma\Sigma\Sigma W_{m,n,c,k} * X_{i+m,j+n,c} + b_k} \quad (14)$$

Dimana:

$Y(i,j,k)$: Output fitur pada posisi (i,j), kanal k

X: Nilai piksel input

W: Bobot kernel

(m,n,c): Indeks posisi kernel dan kanal input

$\sum$: Penjumlahan seluruh elemen kernel

b_k: Bias kanal keluaran k

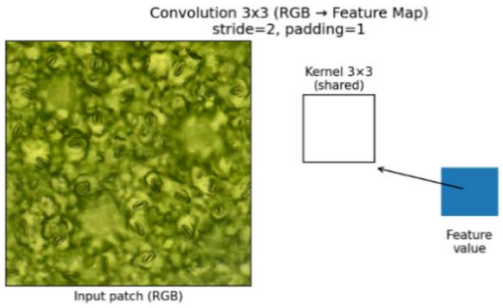
Contoh perhitungan pada fitur pertama (k=0): misal bobot *kernel* rata-rata W=0.1 dan bias b=0.05,

$$\mathbf{Y(0,0,0) = 0.1 \times (0.47 + 0.78 + 0.59) + 0.05 = 0.234} \quad (15)$$

Hasil aktivasi (*SiLU*):

$$\mathbf{f(Y) = Y \times \sigma(Y) = 0.234 \times \frac{1}{1 + e^{-0.234}} = 0.234 \times 0.558 = 0.130} \quad (16)$$

Ukuran keluaran menjadi 320 × 320 × 32.



Gambar 12. Proses Konvolusi

b. Blok C3k2F

Setelah konvolusi dasar, *YOLOv11* menggunakan blok residual untuk mempertahankan informasi dari lapisan sebelumnya dan mempercepat proses konvergensi. Blok *C3k2F* terdiri dari dua konvolusi berurutan dan *shortcut*:

$$\mathbf{O = f\big(W2 * f(W1 * X)\big) + X} \quad (17)$$

Dimana:
X: Input blok

$W1, W2$: Bobot konvolusi pertama dan kedua

f: Fungsi aktivasi (SiLU)

$*$: Operasi konvolusi

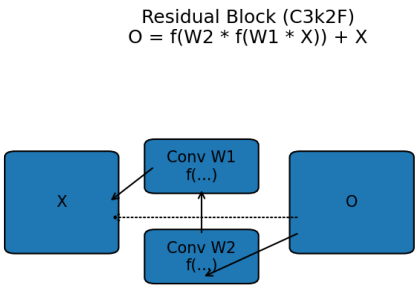
O: Output blok residual

$+X$: Shortcut / skip-connection untuk menjaga informasi awal

Misal hasil f(W_1 * X) = 0.13 dan f(W_2 * X) = 0.40, maka:

$$\mathbf{O} = \mathbf{0.40} + \mathbf{0.13} = \mathbf{0.53} \quad (18)$$

Ukuran tetap 320 × 320 × 32.



Gambar 13. Proses Blok C3k2F

c. Blok C3k2T

Tahapan ini bertujuan memperkaya representasi fitur dengan menambah jumlah kanal. Blok ini meningkatkan kedalaman fitur menjadi 64 kanal:

$$\mathbf{O_T} = \mathbf{f}(\mathbf{W_T} * \mathbf{O}) \quad (19)$$

Dimana:

X: Input dari blok sebelumnya

W_T: Bobot konvolusi untuk penambahan kanal

$*$: Operasi konvolusi

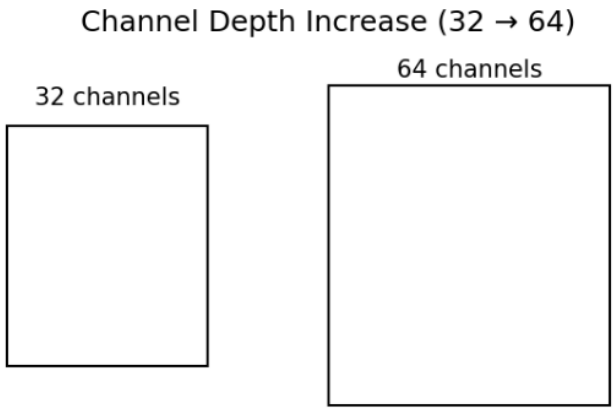
f: Fungsi aktivasi (SiLU)

O_T: Output fitur dengan kanal yang ditingkatkan

Misal bobot rata-rata$W_T = 0.5$, maka:

$$\mathbf{O_T = 0.5 \times 0.53 = 0.265} \quad (20)$$

Ukuran *output* 160 × 160 × 64.



Gambar 14. Proses Blok C3k2T

d. Blok *SPPF* (*Spatial Pyramid Pooling - Fast*)

Lapisan *SPPF* digunakan untuk menangkap informasi konteks dari berbagai skala tanpa mengubah ukuran input. *SPPF* menggabungkan informasi multi-skala dengan *pooling* 5×5, 9×9, dan 13×13:

$$\mathbf{O_s ppf = Concat\left(P_5(O_T), P_9(O_T), P_13(O_T)\right)} \quad (21)$$

Misal hasil pooling rata-rata:

$$\mathbf{P_5 = 0.26, P_9 = 0.27, P_{13} = 0.29} \quad (22)$$

Dimana:

O_T: Input dari blok sebelumnya

P_5, P_9, P_{13}: Pooling dengan ukuran kernel 5×5, 9×9, dan 13×13

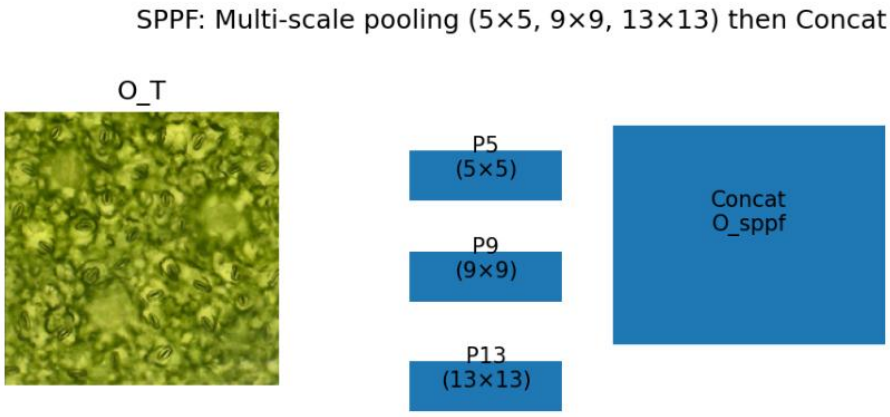
Concat : Menggabungkan hasil pooling multi-skala

O_{sppf}: Output fitur multi-skala dari blok SPPF

Maka:

$$\mathbf{O_{sppf} = [0.26, 0.27, 0.29] = 0.27} \quad (23)$$

Ukuran output menjadi 80 × 80 × 128.



Gambar 15. Proses Blok SPPF

e. Blok C2PSA (*Cross Stage Partial Spatial Attention*)

Mekanisme perhatian bekerja dengan menekankan area penting pada peta fitur dan menekan area yang tidak relevan. *C2PSA* menerapkan mekanisme perhatian spasial dan kanal dengan persamaan:

$$\mathbf{A} = \boldsymbol{\sigma}\left(\mathbf{W_a} * \mathbf{O_{ssppf}}\right) \tag{24}$$

$$\mathbf{O_{c2psa}} = \mathbf{A} \odot \mathbf{O_{sppf}} \tag{25}$$

Dimana:
O_{sppf}: Input dari blok SPPF

W_a: Bobot konvolusi untuk menghasilkan peta perhatian

$*$: Operasi konvolusi

σ: Fungsi sigmoid untuk normalisasi perhatian

A: Peta perhatian (attention map)

$\odot$: Perkalian elemen-per-elemen

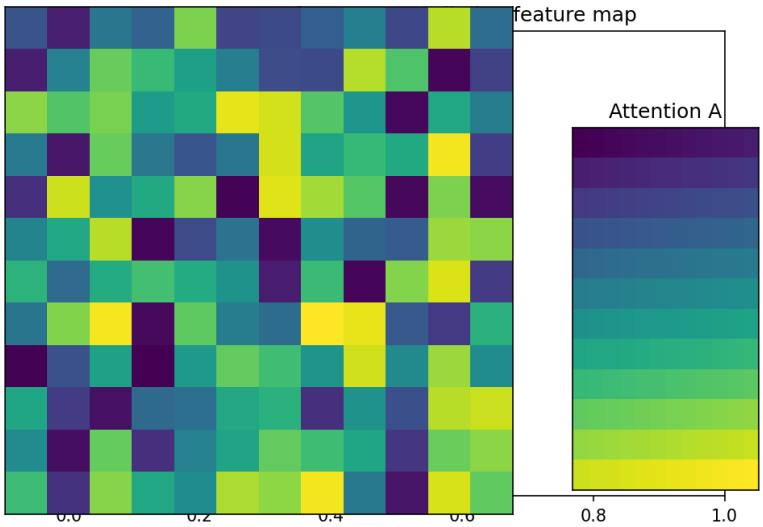
O_{c2psa}: Output setelah perhatian diterapkan

$\text{W}_\text{a} = 0.7$ Maka

$$\mathbf{A} = \boldsymbol{\sigma}(\mathbf{0.7} \times \mathbf{0.27}) = \boldsymbol{\sigma}(\mathbf{0.189}) = \mathbf{0.547} \tag{26}$$

$$\mathbf{O_{c2psa}} = \mathbf{0.547} \times \mathbf{0.27} = \mathbf{0.148} \tag{27}$$

Ukuran tetap 80 × 80 × 128.



Gambar 16. Proses Blok C2PSA

f. *Upsample* dan *Concat*

Setelah fitur diperoleh dari berbagai lapisan, *YOLOv11* melakukan fusi menggunakan mekanisme upsample dan concatenate. Tujuannya agar informasi spasial (resolusi tinggi) dari lapisan bawah dan informasi semantik (abstraksi tinggi) dari lapisan atas dapat digabung.

Upsample menggandakan resolusi:

$$\mathbf{O_{up}} = \mathbf{Interp}\left(\mathbf{O_{c2psa}}, \mathbf{2}\times\right) \rightarrow \mathbf{160 \times 160 \times 128} \tag{28}$$

Kemudian di-*concat* dengan fitur dari tahap sebelumnya:

$$\mathbf{O_{concat}} = \left[\mathbf{O_{up}}, \mathbf{O_T}\right] \tag{29}$$

Ukuran hasil 160 × 160 × 192.

Pada YOLOv11, proses *Concat* dan *Upsample* dilakukan secara multi-skala untuk menggabungkan fitur dari tiga tingkat resolusi berbeda, yaitu:

- 80×80 (skala besar, mendeteksi objek kecil),
- 40×40 (skala menengah),
- 20×20 (skala kecil, mendeteksi objek besar).

Dengan cara ini, model dapat mendeteksi stomata pada berbagai ukuran dan posisi di permukaan daun.

g. Lapisan Deteksi

Tahap ini adalah inti dari proses prediksi YOLOv11. Setiap grid sel pada peta keluaran memprediksi koordinat kotak pembatas (bounding box), tingkat kepercayaan (objectness), dan probabilitas kelas.

Layer deteksi menggunakan ***kernel*** 1×1 untuk menghasilkan prediksi kotak:

$$\boldsymbol{Z = W_d * O_{concat} + b_d} \tag{30}$$

Misal $W_d = 0.3, b_d = 0.05$:

$$\boldsymbol{Z = 0.3 \times 0.148 + 0.05 = 0.094} \tag{31}$$

Setiap skala (80×80, 40×40, 20×20) menghasilkan prediksi dengan format:

$$\boldsymbol{P = (\sigma(x), \sigma(y), \sigma(w), \sigma(h), \sigma(obj), p1, p2, \dots, pn)} \tag{32}$$

dimana:

(x, y): koordinat pusat *bounding box*,

(w, h): ukuran kotak prediksi,

obj: *confidence* bahwa terdapat objek,

p_i : probabilitas tiap kelas stomata.

Tiga *output head* ini kemudian digabungkan untuk menghasilkan hasil deteksi akhir pada berbagai skala.

h. Dekode *Bounding Box*

Setelah prediksi dihasilkan, YOLOv11 melakukan tahap dekode untuk mengonversi koordinat prediksi ke posisi sebenarnya dalam citra. Bounding box divisualisasikan dalam bentuk persegi panjang yang menunjukkan lokasi dan ukuran objek. *YOLOv11* mendekode keluaran jaringan $(\mathrm{t_x}, \mathrm{t_y}, \mathrm{t_w}, \mathrm{t_h})$ menjadi koordinat objek dalam citra asli.

Rumus dekode:

$$\boldsymbol{b_x = \frac{2\sigma(t_x) - 0.5 + c_x}{S}},\ \boldsymbol{b_y = \frac{2\sigma(t_y) - 0.5 + c_y}{S}} \tag{33}$$

$$\boldsymbol{b_w = \left(2\sigma(t_w)\right)^2 \times a_w},\ \boldsymbol{b_h = \left(2\sigma(t_h)\right)^2 \times a_h} \tag{34}$$

Dengan $t_x = t_y = t_w = t_h = 0.094,\ a_w = a_h = 20,\ S = 80,\ c_x = c_y = 5$

$$\boldsymbol{\sigma(0.094) = 0.523} \tag{35}$$

Sehingga

$$\boldsymbol{b_x = \frac{2(0.523) - 0.5 + 5}{80} = 0.069},\ \boldsymbol{b_y = 0.069} \tag{36}$$

$$\boldsymbol{b_w = (1.046)^2 \times 20 = 21.9},\ \boldsymbol{b_h = 21.9} \tag{37}$$

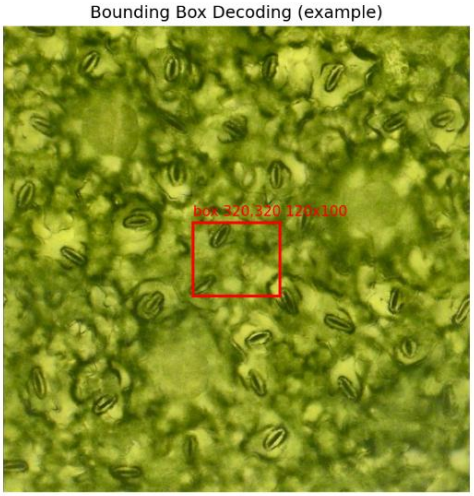
Konversi ke ukuran asli (640x640):

$$\boldsymbol{X_{\text{pusat}} = 0.069 \times 640 = 44.2,} \tag{38}$$

$$\boldsymbol{Y_{\text{pusat}} = 0.069 \times 640 = 44.2} \tag{39}$$

Artinya:

YOLOv11 memprediksi terdapat objek stomata dengan pusat di koordinat (44, 44) dan ukuran sekitar 22×22 piksel pada citra beresolusi 640×640.



Gambar 17. *Bounding Box Decoding*

i. *Output* Klasifikasi (Jenis Stomata)

Setelah lokasi objek diperoleh, YOLO menghitung skor kelas untuk menentukan tipe stomata:

$$\mathbf{p_i = \sigma(W_i * Z + b_i)} \tag{40}$$

Misalkan model memiliki empat kelas: Anomositik, Parasitik, Diasitik, dan Graminoid.
Hasil prediksi (logit) dari layer deteksi: $\mathrm{Z} = [0.094, 0.35, 0.72, 0.56]$

Nilai $\mathrm{Z}_1 = 0.094$ diperoleh dari perhitungan konvolusi terakhir:

$$\boldsymbol{Z_1 = 0.3(0.148) + 0.05 = 0.094} \tag{41}$$

Sementara nilai $\mathrm{Z}_2 = 0.35, \mathrm{Z}_3 = 0.72, \mathrm{dan} \mathrm{Z}_4 = 0.56$ diperoleh dengan rumus yang sama menggunakan bobot dan bias berbeda untuk tiap kelas.

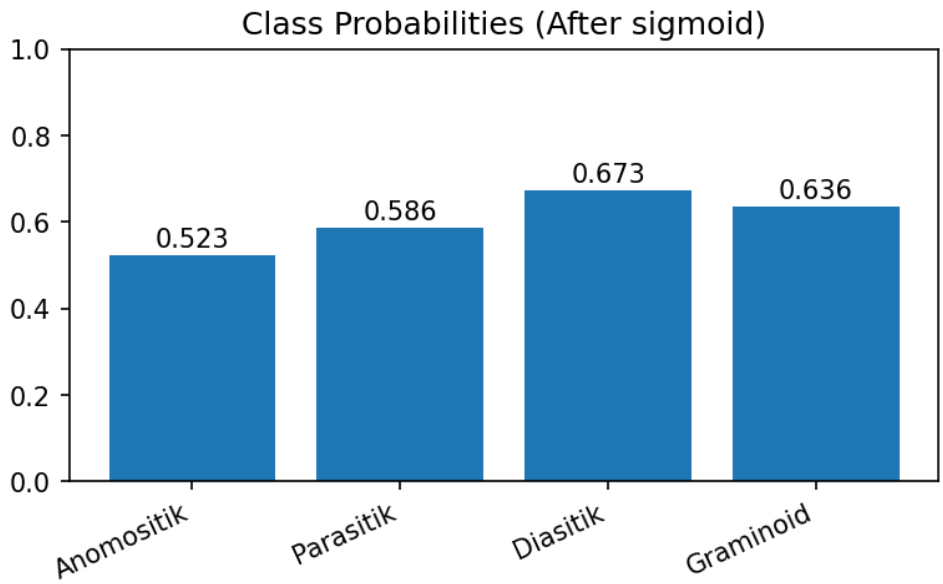
Setelah fungsi sigmoid:

$$\boldsymbol{p} = [\boldsymbol{\sigma}(\mathbf{0.094}), \boldsymbol{\sigma}(\mathbf{0.35}), \boldsymbol{\sigma}(\mathbf{0.72}), \boldsymbol{\sigma}(\mathbf{0.56})] = [\mathbf{0.523}, \mathbf{0.586}, \mathbf{0.673}, \mathbf{0.636}] \quad (42)$$

Kelas dengan nilai probabilitas tertinggi adalah kelas ke 3 (diasitik) dengan 67.3%. Jika pada kondisi lain nilai tertinggi muncul pada kelas ke 4 (graminoid, maka hasil akhir akan ditanyakan sebagai stomata tipe Geaminoid.

Artinya:

YOLOv11 memprediksi terdapat objek stomata berjenis Diasitik dengan pusat koordinat (44,44), ukuran 22 x 22 piksel, dan tingkat keyakinan 67.3%.



Gambar 18. Probabilitas Kelas

Semua nilai probabilitas kelas diperoleh dari perhitungan fungsi sigmoid terhadap logit hasil konvolusi akhir, sedangkan nilai nilai Z_2, Z_3, Z_4 dihitung dengan rumus yang sama seperti contoh Z_1 di atas.

2.5.9 *Input* Data Validasi

Data validasi digunakan sebagai referensi untuk menilai kinerja model secara objektif selama proses pelatihan. Untuk memastikan bahwa model tidak belajar dari data pelatihan, data validasi dalam penelitian ini terdiri dari sekumpulan gambar stomata yang terpisah dari data pelatihan. sehingga hasil prediksi yang dihasilkan benar-benar menunjukkan kemampuan model untuk menangani data baru. Data validasi dimasukkan secara berkala pada setiap interval waktu selama penelitian untuk mengamati tren performa model, termasuk peningkatan akurasi dan penurunan nilai kehilangan. Dengan menganalisis hasil validasi ini, para peneliti dapat mengidentifikasi kemungkinan overfitting, underfitting, atau indikasi bahwa parameter perlu disesuaikan pada tahap pelatihan berikutnya. Dengan demikian, data validasi memainkan peran penting dalam

menjamin bahwa model yang dilatih mampu beradaptasi dengan pola visual stomata secara lebih umum dan tidak hanya menghafal data pelatihan.

2.5.10 Validasi Model *YOLOv11*

Setelah pelatihan selesai, model YOLOv11 diuji dengan data validasi yang telah disiapkan sebelumnya. Proses validasi dilakukan untuk mendapatkan gambaran awal tentang stabilitas dan kemampuan generalisasi model untuk mendeteksi dan mengklasifikasikan berbagai jenis stomata. Hasil prediksi model dinilai berdasarkan tingkat ketepatan deteksi, kesesuaian bounding box, dan konsistensi pengenalan pola morfologi stomata. Selain itu, proses validasi memastikan apakah model telah mempelajari fitur visual yang relevan, seperti bentuk sel tetangga. Dengan menggunakan *YOLOv11*, peneliti dapat memanfaatkan peningkatan arsitektur pada versi terbaru dari *YOLO* untuk memperoleh performa validasi yang lebih baik, terutama dalam hal akurasi deteksi objek berukuran kecil seperti stomata.

2.5.11 Evaluasi Model

Dalam proses evaluasi model, sejumlah metrik standar digunakan, termasuk precision, recall, dan mean average precision (mAP). Precision mengukur tingkat ketepatan model dalam mengidentifikasi objek stomata tanpa menghasilkan deteksi palsu, sedangkan recall mengukur kemampuan model dalam menemukan seluruh objek stomata yang benar-benar ada. Karena mAP memperhitungkan nilai presisi pada berbagai ambang batas Intersection over Union (IoU), itu digunakan sebagai indikator kinerja lengkap. Hasil evaluasi selama penelitian menunjukkan perbedaan dalam kinerja setiap kelas stomata. Ini memungkinkan peneliti untuk menentukan kelas mana yang memerlukan perhatian lebih lanjut. Analisis yang lebih mendalam terhadap hasil evaluasi juga membantu menemukan kemungkinan kelemahan model, seperti kesalahan klasifikasi pada bentuk stomata yang memiliki morfologi yang sangat mirip atau kerusakan kualitas gambar pada objek. Hasil ini kemudian digunakan oleh peneliti untuk melakukan optimasi lanjutan pada tahap tunning hyperparameter.

2.5.12 *Hyperparameter Tuning*

Berdasarkan hasil evaluasi, peneliti kemudian melakukan penyesuaian hyperparameter untuk meningkatkan kinerja model. Beberapa parameter penting dalam arsitektur YOLOv11 telah disesuaikan, termasuk ukuran batch (ukuran batch = 32), dan jumlah epoch (100 epoch). Tujuan dari proses tuning ini adalah untuk mencapai konfigurasi terbaik yang memungkinkan model mempelajari fitur visual stomata dengan lebih baik tanpa mengalami overfitting. Selama penelitian, sejumlah kombinasi parameter diuji secara bertahap untuk mengevaluasi pengaruh mereka terhadap stabilitas hasil prediksi, akurasi deteksi, dan nilai kehilangan. Model dapat meningkatkan nilai precision, recall, dan mAP dengan penyesuaian hyperparameter yang cermat. Selain itu, proses tunning membantu generalisasi model ketika diuji menggunakan data stomata yang lebih bervariasi.

2.5.13 Model *Classification*

Model YOLOv11 ditetapkan sebagai model terakhir dan hasil pelatihan YOLOv11 disimpan dalam format.pt, yang akan digunakan pada tahap analisis berikutnya. Proses pelatihan selesai ketika hasil memenuhi kriteria. Pada saat ini, model dianggap stabil dan

mampu membuat prediksi yang akurat tentang cara menemukan stomata dan mengklasifikasikannya. Model terakhir inilah yang digunakan sebagai sistem prediksi utama penelitian ini. Selain itu, analisis kembali dilakukan pada model terakhir untuk memastikan bahwa bobot dan parameter yang tersimpan mencerminkan kinerja terbaik yang dicapai selama proses pelatihan. Oleh karena itu, model YOLOv11 terakhir telah siap digunakan untuk proses pengujian dengan menggunakan data yang benar-benar baru untuk menilai performa sebenarnya pada kondisi riil.

2.5.14 *Input* Data Testing

Data pengujian adalah sekumpulan gambar stomata yang tidak pernah digunakan dalam proses pelatihan atau validasi. Data tes digunakan dalam penelitian ini untuk menggambarkan performa akhir model dalam skenario nyata, yang menunjukkan perbedaan kondisi gambar mikroskopis. Variasi kualitas gambar, intensitas cahaya, orientasi objek, dan perbedaan jenis jaringan daun disertakan dalam data ini untuk memberikan pengujian yang lebih mendalam.

Pada titik ini, model akan melakukan sendiri untuk mendeteksi dan mengklasifikasikan stomata. Untuk menilai akurasi model di lapangan, hasil prediksi dibandingkan dengan label kebenaran (ground truth). Tahap ini sangat penting karena hasil tes menentukan keberhasilan model dalam menangani data yang kompleks dan bervariasi.

2.5.15 Klasifikasi dan Deteksi Stomata

Proses klasifikasi dan deteksi stomata menggunakan model YOLOv11 yang telah dipersiapkan adalah tahap terakhir dari penelitian ini. Pada titik ini, model dapat menghasilkan output seperti bounding box yang menunjukkan posisi objek, label kelas yang menunjukkan jenis stomata, dan nilai keyakinan yang menunjukkan tingkat keyakinan model terhadap prediksi yang dibuat. Model mengklasifikasikan stomata otomatis menjadi empat kategori: anomocytic, diacytic, paracytic, dan graminoid. Dalam penelitian ini, ditunjukkan bahwa penerapan model dapat mempercepat proses analisis dibandingkan dengan metode manual. Itu juga dapat meningkatkan konsistensi dan objektivitas identifikasi. Selanjutnya, hasil deteksi dan klasifikasi dipelajari untuk membantu memahami secara ilmiah perbedaan morfologi stomata dan tujuan penelitian.

2.5.16 Evaluasi Hasil Klasifikasi

Setelah deteksi dan klasifikasi stomata selesai, hasilnya dibandingkan dengan kebenaran aslinya—atau label sebenarnya—dan dinilai dengan beberapa metrik performa. Accuracy, precision, recall, F1-score, dan mean average precision (mAP) adalah metrik utama yang digunakan dalam penilaian ini. Tujuan dari evaluasi ini adalah untuk mengetahui seberapa akurat model dapat mendeteksi dan mengklasifikasikan stomata dari gambar mikrograf.

Digunakan teknik evaluasi matrix confusion untuk mengetahui seberapa baik model YOLOv11 mendeteksi berbagai jenis stomata pada masing-masing skenario pengolahan gambar—tanpa preprocessing, dengan koreksi Gamma, dan dengan CLAHE. Matrix confusion ini menunjukkan empat kategori utama biner atau multi-kelas, yaitu:

1. *True Positive* (TP): Objek stomata yang berhasil terdeteksi dan diklasifikasikan dengan benar oleh model.
2. *False Negative* (FN): Objek stomata yang tidak berhasil terdeteksi oleh model (*missed detection*).
3. *False Positive* (FP): Bagian citra non-stomata yang secara keliru terdeteksi sebagai stomata oleh model.
4. *True Negative* (TN): Bagian citra yang bukan stomata dan dengan benar tidak diklasifikasikan sebagai stomata oleh model.

Gambar 19. *Confusion Matrix*

Source: https://medium.com/@sserafino5/panduan-tentang-matriks-konfusi-mengukur-akurasi-model-klasifikasi-ad76994f7b91

Gambar 19 merupakan evaluasi dengan *confusion matrix* digunakan karena merupakan pendekatan umum dalam pembelajaran *supervised learning*, yang berfungsi untuk memberikan gambaran kinerja dari model klasifikasi. Beberapa metrik turunan yang digunakan antara lain:

1. *Accuracy*, Menunjukkan tingkat keakuratan model secara keseluruhan dalam mengklasifikasikan citra. Rumusnya adalah.

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP} + \boldsymbol{TN}}{\boldsymbol{TP} + \boldsymbol{TN} + \boldsymbol{FP} + \boldsymbol{FN}} \tag{43}$$

2. *Precision*, Mengukur seberapa banyak prediksi positif yang benar-benar merupakan stomata. *Precision* sangat penting ketika salah klasifikasi sebagai stomata perlu dihindari.

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FP}} \tag{44}$$

3. *Recall*, Menunjukkan kemampuan model dalam mendeteksi semua objek stomata yang ada. *Recall* penting ketika semua objek harus ditemukan, meskipun ada risiko deteksi palsu.

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}} \tag{45}$$

4. F1-Score, Merupakan rata-rata harmonis antara *Precision* dan *Recall*, yang berguna ketika diperlukan keseimbangan antara keduanya.

$$\boldsymbol{F1 - Score = 2 \,.\, \frac{TP}{TP + FN}} \tag{46}$$

5. *Mean Average Precision (mAP),* Rata-rata dari nilai AP untuk semua kelas. Ini digunakan dalam multi-class *classification* dan deteksi objek seperti *YOLO*.

$$\boldsymbol{mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i} \tag{47}$$

Keterangan:

N = Jumlah total kelas yang diuji

AP_i = Nilai *Average Precision* untuk kelas ke-i.

Dengan melakukan evaluasi menggunakan metrik-metrik di atas, maka dapat diketahui skenario *preprocessing* citra mana yang memberikan hasil klasifikasi terbaik untuk mikrograf stomata, serta menentukan kesiapan model untuk digunakan secara *real-world* di bidang anatomi tumbuhan atau botani mikroskopik.

6. *PSNR (Peak Signal-to-Noise Ratio)* adalah metrik evaluasi kuantitatif yang digunakan untuk mengukur kualitas gambar atau video hasil pemrosesan (seperti kompresi, peningkatan kontras, *denoising*, segmentasi, dsb.) dengan cara membandingkannya terhadap gambar referensi (*ground truth*).

$$\boldsymbol{PNSR = 10 \cdot log_{10}(\frac{MAX_I^2}{MSE})} \tag{48}$$

Keterangan:

$MAX_I =$ nilai maksimum piksel dalam citra (255).

MSE = *Mean Squared Error*, yaitu rata-rata kuadrat perbedaan antara citra asli dan citra hasil.

PSNR menghitung berapa besar "*noise*" atau kesalahan yang muncul dalam gambar hasil dibanding gambar aslinya. Makin kecil kesalahan (*noise*), makin tinggi nilai *PSNR*-nya.