

Evaluasi Kinerja Arsitektur Segmentasi Semantik untuk Identifikasi Area Penyakit Daun Padi

AZARIA SYAHLA FITAN ADIBAH¹, RISKYANA DEWI INTAN PUSPITASARI^{2,*}, ELLY MATUL
IMAH²

¹*Program Studi S1 Sains Data, Universitas Negeri Surabaya, Indonesia*

²*Program Studi S1 Kecerdasan Artifisial, Universitas Negeri Surabaya, Indonesia*

Abstrak

Ketahanan pangan nasional sangat dipengaruhi oleh produktivitas tanaman padi yang rentan terhadap berbagai penyakit daun padi. Deteksi manual yang masih bergantung pada pengamatan visual sering kali bersifat subjektif dan kurang efisien. Penelitian ini mengusulkan pendekatan segmentasi semantik berbasis *deep learning* untuk mendeteksi area penyakit daun padi secara otomatis. Tiga arsitektur segmentasi semantik, yaitu U-Net, PSPNet, dan DeepLabV3+, dibandingkan menggunakan *backbone* ResNet34 dan konfigurasi pelatihan yang seragam untuk memperoleh evaluasi yang objektif. Dataset yang digunakan merupakan Rice Leaf Disease with Segmentation Labels yang terdiri atas citra daun padi beserta anotasi piksel untuk beberapa jenis penyakit daun padi. Tahapan penelitian meliputi pra-pemrosesan data, augmentasi citra, pelatihan model, serta evaluasi menggunakan metrik *Intersection over Union* (IoU) dan akurasi piksel. Hasil eksperimen menunjukkan bahwa DeepLabV3+ memberikan performa terbaik dengan nilai IoU sebesar 0,9804 dan akurasi piksel sebesar 0,9235, melampaui U-Net dan PSPNet. Hasil penelitian menunjukkan bahwa pendekatan segmentasi semantik mampu memberikan deteksi area penyakit daun padi secara lebih detail dan akurat sehingga berpotensi mendukung pengembangan sistem pertanian presisi berbasis kecerdasan buatan.

Kata Kunci *DeepLabV3+; PSPNet; Penyakit daun padi; Segmentasi semantik; U-Net*

Abstract

National food security is strongly influenced by rice productivity, which is vulnerable to various rice leaf diseases. Manual detection based on visual observation is often subjective and inefficient. This research proposes a deep learning-based semantic segmentation approach for automatically detecting rice leaf disease regions. Three semantic segmentation architectures, namely U-Net, PSPNet, and DeepLabV3+, were compared using the same ResNet34 backbone and identical training configurations to ensure objective evaluation. The dataset used was the Rice Leaf Disease with Segmentation Labels dataset, consisting of rice leaf images and pixel-level annotations for several rice leaf diseases. The research stages included data pre-processing, image augmentation, model training, and evaluation using Intersection over Union (IoU) and pixel accuracy metrics. Experimental results showed that DeepLabV3+ achieved the best performance with an IoU score of 0.9804 and pixel accuracy of 0.9235, outperforming U-Net and PSPNet. The findings indicate that semantic segmentation can provide more detailed and accurate detection of rice leaf disease regions and has the potential to support the development of artificial intelligence-based precision

*Corresponding author. Email: riskyanauspitasari@unesa.ac.id or ellymatul@unesa.ac.id.

agriculture systems.

Keywords *DeepLabV3+; PSPNet; Rice leaf disease; Semantic segmentation; U-Net*

1 Pendahuluan

Padi merupakan salah satu komoditas pertanian utama yang memiliki peran penting dalam menjaga ketahanan pangan nasional. Produktivitas padi dapat mengalami penurunan akibat berbagai penyakit daun, seperti *leaf blast*, *bacterial leaf blight*, dan *brown spot* (Mau et al., 2020; Howard et al., 2025; Sopialena et al., 2021). Penyakit tersebut dapat merusak jaringan daun, mengganggu proses fotosintesis, serta menurunkan hasil panen (He et al., 2023; Bag et al., 2023). Oleh karena itu, deteksi dini penyakit daun padi menjadi hal yang penting untuk mendukung pengelolaan penyakit secara tepat dan meminimalkan kerugian di sektor pertanian. Deteksi penyakit daun padi secara manual masih umum dilakukan melalui pengamatan visual. Namun, pendekatan tersebut sering kali bersifat subjektif, tidak konsisten, dan sangat bergantung pada pengalaman pengamat dalam memperoleh hasil identifikasi yang akurat (Bock et al., 2020). Perkembangan teknologi *computer vision* dan *deep learning* membuka peluang untuk mengembangkan sistem deteksi penyakit tanaman secara otomatis dengan tingkat akurasi yang lebih tinggi (Khan et al., 2025; Li et al., 2022). Salah satu pendekatan yang umum digunakan adalah klasifikasi citra, tetapi metode ini hanya mampu mengidentifikasi kategori penyakit tanpa memberikan informasi detail mengenai lokasi area terinfeksi pada permukaan daun (Roy, 2021; Setiawan et al., 2023).

Segmentasi semantik merupakan pendekatan *computer vision* yang mampu melakukan klasifikasi pada tingkat piksel sehingga area penyakit dapat diidentifikasi secara lebih presisi. Pendekatan ini memungkinkan model membedakan area daun sehat dan area terinfeksi dalam suatu citra. Beberapa arsitektur segmentasi semantik yang banyak digunakan dalam analisis citra antara lain U-Net (Ronneberger et al., 2015), PSPNet (Zhao et al., 2017), dan DeepLabV3+ (Chen et al., 2018). U-Net menggunakan struktur *encoder-decoder* dengan *skip connection* untuk mempertahankan informasi spasial secara efektif. PSPNet memanfaatkan *pyramid pooling module* untuk menangkap informasi kontekstual pada berbagai skala, sedangkan DeepLabV3+ menerapkan *atrous convolution* dan *Atrous Spatial Pyramid Pooling* (ASPP) untuk meningkatkan performa segmentasi pada objek dengan batas yang kompleks.

Beberapa penelitian sebelumnya telah menerapkan metode segmentasi semantik untuk deteksi penyakit tanaman menggunakan berbagai arsitektur *deep learning*. Namun, evaluasi komparatif menggunakan konfigurasi pelatihan yang identik dan *backbone* yang sama masih terbatas sehingga perbandingan performa secara objektif menjadi sulit dilakukan. Selain itu, penelitian mengenai segmentasi penyakit daun padi masih terus berkembang untuk memperoleh hasil segmentasi yang lebih detail dan akurat. Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pendekatan segmentasi semantik berbasis *deep learning* untuk mendeteksi area penyakit daun padi secara otomatis. Penelitian ini membandingkan performa tiga arsitektur segmentasi semantik, yaitu U-Net, PSPNet, dan DeepLabV3+, menggunakan *backbone* ResNet34 dan konfigurasi pelatihan yang identik. Evaluasi model dilakukan menggunakan metrik *Intersection over Union* (IoU) dan *pixel accuracy* untuk menganalisis efektivitas masing-masing arsitektur dalam melakukan segmentasi area penyakit daun padi.

2 Metode Penelitian

Penelitian ini mengusulkan pendekatan segmentasi semantik berbasis *deep learning* untuk mendeteksi penyakit daun padi secara otomatis. Metodologi penelitian meliputi persiapan dataset, pra-pemrosesan data, pelatihan model, dan evaluasi performa. Tiga arsitektur segmentasi semantik, yaitu U-Net, PSPNet, dan DeepLabV3+, diimplementasikan dan dievaluasi menggunakan *backbone* serta konfigurasi pelatihan yang sama untuk memastikan perbandingan performa dilakukan secara objektif.

2.1 Dataset

Dataset yang digunakan dalam penelitian ini adalah *Rice Leaf Disease with Segmentation Labels* yang diperoleh dari platform Kaggle (Li et al., 2022). Dataset terdiri atas 450 citra daun padi yang dilengkapi dengan mask segmentasi tingkat piksel. Mask tersebut mencakup beberapa kelas semantik, yaitu *background*, daun sehat, *bacterial blight*, *blast*, dan *brown spot*. Dataset juga memiliki variasi kondisi pencahayaan, ukuran lesi, dan tekstur citra yang merepresentasikan kondisi pertanian di lapangan.

Gambar 1 menunjukkan beberapa contoh citra pada dataset yang merepresentasikan jenis utama penyakit daun padi yang digunakan dalam penelitian ini.



Gambar 1: Contoh citra daun padi yang menunjukkan empat jenis penyakit utama

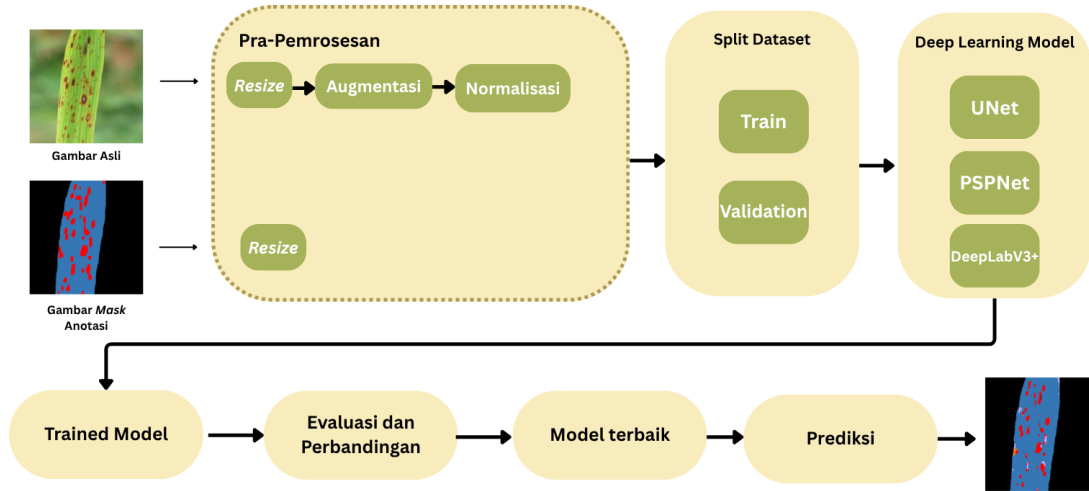
Dataset memiliki variasi yang cukup besar pada pencahayaan, ukuran lesi, orientasi daun, dan kondisi latar belakang yang mencerminkan lingkungan pertanian nyata. Variasi tersebut meningkatkan kompleksitas tugas segmentasi sekaligus memberikan tantangan dalam mengevaluasi ketahanan dan kemampuan generalisasi model. Selain itu, keberagaman karakteristik penyakit pada dataset memungkinkan model mempelajari pola visual yang lebih representatif untuk tugas segmentasi semantik. Oleh karena itu, dataset ini sesuai digunakan untuk mengevaluasi efektivitas arsitektur segmentasi semantik dalam mendeteksi area penyakit daun padi.

2.2 Tahapan Penelitian

Tahapan penelitian pada penelitian ini ditunjukkan pada Gambar 2. Tahapan tersebut menggambarkan alur kerja keseluruhan dari pendekatan segmentasi semantik yang diusulkan, dimulai dari persiapan dataset, pra-pemrosesan data, pelatihan model, proses prediksi, hingga evaluasi performa. Pada tahap awal, citra daun padi beserta mask segmentasinya dikumpulkan dan dipersiapkan untuk proses pelatihan. Data kemudian melalui beberapa tahapan pra-pemrosesan,

meliputi *image resizing*, normalisasi, augmentasi, dan konversi tensor untuk memastikan kompatibilitas dengan arsitektur *deep learning*. Dataset yang telah diproses selanjutnya dibagi menjadi data pelatihan dan data pengujian dengan rasio 80% dan 20%.

Tahap pelatihan melibatkan implementasi tiga arsitektur segmentasi semantik, yaitu U-Net (Ronneberger et al., 2015), PSPNet (Zhao et al., 2017), dan DeepLabV3+ (Chen et al., 2018), menggunakan *backbone* ResNet34 dan konfigurasi pelatihan yang identik. Setiap model menghasilkan prediksi segmentasi tingkat piksel untuk mengidentifikasi area penyakit pada citra daun padi. Pada tahap akhir, performa model dievaluasi menggunakan metrik *Intersection over Union* (IoU) dan *pixel accuracy* untuk menganalisis efektivitas masing-masing arsitektur dalam melakukan segmentasi area penyakit daun padi. Hasil evaluasi kemudian dibandingkan untuk menentukan model segmentasi semantik yang paling efektif dalam mendeteksi penyakit daun padi.



Gambar 2: Tahapan Penelitian

Gambar 2 menunjukkan alur keseluruhan pendekatan segmentasi semantik yang diusulkan. Tahapan penelitian dimulai dari persiapan dataset dan tahap pra-pemrosesan data untuk memastikan seluruh citra dan mask segmentasi telah distandarisasi sebelum proses pelatihan dilakukan. Data yang telah diproses kemudian digunakan untuk melatih dan mengevaluasi arsitektur U-Net, PSPNet, dan DeepLabV3+ menggunakan konfigurasi yang identik. Selanjutnya, hasil segmentasi yang dihasilkan masing-masing model dievaluasi menggunakan metrik IoU dan *pixel accuracy* untuk membandingkan performa dalam mendeteksi area penyakit daun padi.

2.3 Pra-pemrosesan Data

Tahap pra-pemrosesan data dilakukan untuk memastikan konsistensi, kualitas data, dan kompatibilitas sebelum proses pelatihan model. Dataset yang digunakan dalam penelitian ini terdiri atas citra daun padi dan mask segmentasi tingkat piksel yang merepresentasikan beberapa kelas semantik, yaitu *background*, daun sehat, *bacterial blight*, *blast*, *blight*, dan *brown spot* (Li et al., 2022). Setiap citra daun padi dipasangkan dengan mask segmentasi yang sesuai berdasarkan kecocokan nama file untuk menjaga keselarasan anotasi selama proses pelatihan segmentasi semantik. Tahap ini penting dilakukan untuk mencegah ketidaksesuaian label yang dapat memengaruhi proses pembelajaran dan performa evaluasi model.

Mask segmentasi yang awalnya direpresentasikan dalam format warna RGB dikonversi menjadi label indeks kelas untuk klasifikasi tingkat piksel. Proses *color mapping* dilakukan untuk memetakan setiap nilai RGB ke dalam kategori kelas semantik yang sesuai. Konversi ini memungkinkan model melakukan segmentasi semantik secara lebih efektif dengan mengubah mask menjadi representasi label numerik yang sesuai dengan kebutuhan kerangka kerja *deep learning*. Seluruh citra dan mask segmentasi kemudian diubah ukurannya menjadi resolusi tetap sebesar 384×384 piksel untuk memastikan kesesuaian dengan dimensi input yang dibutuhkan oleh arsitektur segmentasi semantik. Normalisasi citra juga diterapkan dengan mengubah rentang intensitas piksel menjadi $[0,1]$ guna meningkatkan stabilitas pelatihan dan mempercepat konvergensi model (Sane and Agrawal, 2017). Selanjutnya, citra yang telah diproses dikonversi ke dalam representasi tensor agar proses komputasi selama pelatihan model menjadi lebih efisien.

Dataset dibagi menjadi data pelatihan dan data validasi dengan rasio masing-masing sebesar 80% dan 20%. Rasio 80:20 dipilih untuk menyediakan jumlah data pelatihan yang memadai dalam mempelajari pola lesi sekaligus mempertahankan data evaluasi yang representatif selama pengujian. Untuk menjaga konsistensi distribusi kelas pada setiap subset, diterapkan metode *stratified random sampling* berdasarkan kategori penyakit. Strategi ini memastikan setiap subset memiliki sampel representatif dari seluruh kelas penyakit dan mengurangi potensi bias selama proses evaluasi model.

Teknik augmentasi data diterapkan hanya pada data pelatihan untuk meningkatkan kemampuan generalisasi model dan mengurangi risiko *overfitting*. Proses augmentasi meliputi rotasi acak hingga $\pm 35^\circ$ dan *horizontal flipping*. Rotasi acak diterapkan karena orientasi daun padi dapat bervariasi bergantung pada sudut pengambilan gambar di lingkungan pertanian nyata, sedangkan *horizontal flipping* digunakan untuk meningkatkan variasi orientasi tanpa mengubah karakteristik lesi. Teknik augmentasi tersebut dipilih karena mampu mempertahankan informasi semantik citra sekaligus meningkatkan keragaman data.

Seluruh proses augmentasi diterapkan secara konsisten pada citra input dan mask segmentasi yang bersesuaian untuk menjaga keselarasan antara label piksel dan area penyakit. Pendekatan ini memastikan posisi lesi tetap sinkron setelah transformasi dilakukan sehingga model segmentasi semantik dapat mempelajari hubungan spasial secara akurat selama pelatihan. Pada arsitektur U-Net dan PSPNet, tahapan pra-pemrosesan diimplementasikan menggunakan utilitas TensorFlow *tf.data.Dataset*. Sementara itu, DeepLabV3+ memanfaatkan pustaka Albumentations untuk proses augmentasi, normalisasi, dan konversi tensor menggunakan *ToTensorV2()*. Data yang telah diproses selanjutnya digunakan sebagai input pada proses pelatihan model segmentasi semantik. mentation, normalization, and tensor conversion using *ToTensorV2()*. The processed data were subsequently used as input for semantic segmentation model training.

2.4 Arsitektur Segmentasi Semantik

Penelitian ini mengevaluasi tiga arsitektur segmentasi semantik, yaitu U-Net, PSPNet, dan DeepLabV3+, untuk deteksi penyakit daun padi secara otomatis. Arsitektur tersebut dipilih karena merepresentasikan pendekatan yang berbeda dalam mengekstraksi informasi spasial dan kontekstual pada tugas segmentasi semantik. Untuk memastikan perbandingan yang objektif, seluruh arsitektur diimplementasikan menggunakan *backbone* ResNet34 dan konfigurasi pelatihan yang identik. Segmentasi semantik bertujuan untuk mengklasifikasikan setiap piksel dalam citra ke dalam kategori semantik yang telah ditentukan (Long et al., 2015). Pada penelitian ini, kelas semantik yang digunakan meliputi *background*, daun sehat, *bacterial blight*, *blast*, *blight*, dan *brown spot*. Dibandingkan dengan teknik pengolahan citra konvensional, metode segmentasi semantik

berbasis *deep learning* mampu memberikan informasi spasial yang lebih detail dan pemahaman kontekstual yang lebih baik untuk mendeteksi area penyakit (Bagga and Goyal, 2025). Proses segmentasi secara umum dapat direpresentasikan sebagai fungsi pemetaan:

$$f : X \rightarrow Y \quad (1)$$

dengan X merepresentasikan ruang citra masukan dan Y merepresentasikan ruang label semantik tingkat piksel. Pada proses segmentasi semantik, setiap citra masukan akan dipetakan ke dalam representasi label yang menunjukkan kategori setiap piksel pada citra. Tujuan utama segmentasi semantik adalah mengidentifikasi dan mengelompokkan setiap piksel berdasarkan kelas semantik yang sesuai, seperti area daun sehat, area penyakit, dan latar belakang. Dengan pendekatan ini, model tidak hanya mampu mengenali keberadaan penyakit, tetapi juga menentukan lokasi dan batas area penyakit secara lebih detail pada tingkat piksel. Setiap piksel x_{ij} pada citra akan diklasifikasikan ke dalam kelas semantik yang sesuai, yaitu y_{ij} , sehingga menghasilkan mask segmentasi yang merepresentasikan distribusi area penyakit pada daun padi.

2.4.1 U-Net

U-Net merupakan arsitektur segmentasi semantik yang diperkenalkan oleh Ronneberger et al. (2015) untuk segmentasi citra biomedis. Arsitektur ini mengadopsi struktur simetris *encoder-decoder* dengan *skip connection* yang berfungsi mempertahankan informasi spasial selama proses rekonstruksi fitur. Karena kemampuannya dalam melokalisasi batas objek secara detail, U-Net menjadi salah satu arsitektur yang paling banyak digunakan pada berbagai aplikasi segmentasi semantik, termasuk analisis penyakit tanaman di bidang pertanian. Selain itu, struktur arsitektur yang relatif sederhana membuat U-Net mampu menghasilkan performa segmentasi yang baik meskipun dilatih menggunakan dataset dengan jumlah terbatas.

Komponen *encoder*, yang juga disebut sebagai *contracting path*, bertugas mengekstraksi fitur visual secara hierarkis melalui operasi konvolusi dan *max-pooling* secara berulang. Setiap operasi konvolusi dapat dinyatakan secara matematis sebagai berikut:

$$y(i, j) = \sum_m \sum_n x(i - m, j - n) \cdot w(m, n) \quad (2)$$

dengan $x(i, j)$ merepresentasikan citra masukan, $w(m, n)$ menunjukkan kernel konvolusi, dan $y(i, j)$ merepresentasikan *feature map* yang dihasilkan. Proses konvolusi memungkinkan model mempelajari pola visual penting, seperti tekstur, bentuk lesi, dan karakteristik area penyakit pada daun padi. Semakin dalam lapisan konvolusi yang digunakan, semakin kompleks representasi fitur yang dapat dipelajari oleh model.

Setelah proses konvolusi, dilakukan aktivasi menggunakan fungsi aktivasi Rectified Linear Unit (ReLU) yang dirumuskan sebagai berikut:

$$f(x) = \max(0, x) \quad (3)$$

Fungsi aktivasi ini berperan dalam menambahkan sifat non-linear pada jaringan sehingga model mampu mempelajari representasi visual yang lebih kompleks (Goodfellow et al., 2016). Selain itu, penggunaan ReLU juga membantu mempercepat proses pelatihan dan mengurangi risiko *vanishing gradient* pada jaringan *deep learning*. Oleh karena itu, fungsi aktivasi ReLU banyak digunakan pada berbagai arsitektur segmentasi dan klasifikasi citra modern.

Untuk mengurangi dimensi spasial sekaligus mempertahankan representasi fitur dominan, diterapkan operasi *max-pooling* pada tahap *encoder* (Goodfellow et al., 2016). Operasi *max-pooling* dirumuskan sebagai berikut:

$$P(i, j) = \max_{(m, n) \in R} x(m, n) \quad (4)$$

dimana R merepresentasikan wilayah *pooling* yang bersesuaian dengan posisi piksel (i, j) . Operasi ini membantu mengurangi kompleksitas komputasi dan meningkatkan efisiensi pelatihan model. Selain itu, *max-pooling* juga memungkinkan model menjadi lebih robust terhadap variasi kecil pada posisi objek dalam citra.

Komponen *decoder*, yang juga disebut sebagai *expanding path*, bertugas merekonstruksi peta segmentasi melalui proses *upsampling*. Salah satu karakteristik utama U-Net adalah mekanisme *skip connection* yang menggabungkan representasi fitur dari *encoder* dan *decoder* sebagai berikut:

$$F_U = [F_{enc}, F_{dec}] \quad (5)$$

dengan F_{enc} merepresentasikan *feature map* dari *encoder* dan F_{dec} merepresentasikan *feature map* dari *decoder*. Mekanisme ini memungkinkan jaringan mempertahankan informasi spasial beresolusi tinggi sehingga akurasi batas segmentasi dapat ditingkatkan. Pada tugas segmentasi penyakit daun padi, *skip connection* membantu model menghasilkan segmentasi area lesi yang lebih detail dan presisi, terutama pada bagian batas penyakit yang memiliki pola kompleks.

2.4.2 PSPNet

PSPNet diperkenalkan oleh Zhao et al. (2017) untuk meningkatkan performa segmentasi semantik melalui pemanfaatan informasi kontekstual global menggunakan *Pyramid Pooling Module* (PPM). Berbeda dengan jaringan segmentasi konvensional yang umumnya berfokus pada fitur spasial lokal, PSPNet mengagregasikan representasi kontekstual dari berbagai *receptive field* untuk meningkatkan pemahaman terhadap keseluruhan citra. Pendekatan ini memungkinkan model memahami hubungan antarobjek dan area pada citra secara lebih menyeluruh. Dalam tugas segmentasi penyakit daun padi, kemampuan memahami konteks global menjadi penting karena area penyakit sering muncul dengan ukuran, bentuk, dan distribusi spasial yang bervariasi.

Diberikan sebuah *feature map* masukan $F \in \mathbb{R}^{H \times W \times C}$, representasi fitur hasil *pooling* pada skala piramida s_i dirumuskan sebagai berikut:

$$P_i = \phi_{s_i}(F) \quad (6)$$

dimana ϕ_{s_i} merepresentasikan operasi *spatial pooling* pada skala s_i . Proses *pooling* dilakukan pada beberapa ukuran skala berbeda untuk memperoleh informasi kontekstual dari berbagai tingkat representasi spasial. Pendekatan ini memungkinkan model menangkap informasi global dan lokal secara bersamaan sehingga proses segmentasi menjadi lebih stabil terhadap variasi ukuran objek. Selain itu, strategi multi-skala membantu model mengenali pola penyakit yang memiliki karakteristik visual berbeda-beda pada setiap citra.

Fitur hasil *pooling* selanjutnya diubah kembali ke resolusi fitur awal melalui proses interpolasi. Representasi kontekstual multi-skala yang dihasilkan dirumuskan sebagai berikut:

$$F_{PPM} = [F, \hat{P}_1, \hat{P}_2, \dots, \hat{P}_n] \quad (7)$$

dengan $\hat{P}_i$ merepresentasikan *feature map* hasil *upsampling* pada skala piramida ke- i . Seluruh representasi fitur tersebut kemudian digabungkan untuk menghasilkan representasi kontekstual yang lebih kaya sebelum dilakukan proses segmentasi akhir. Penggabungan fitur multi-skala ini membantu model mempertahankan detail spasial sekaligus memahami konteks global citra secara lebih efektif. Dengan representasi fitur yang lebih komprehensif, model mampu menghasilkan segmentasi yang lebih konsisten pada area penyakit dengan batas yang kompleks.

Melalui proses agregasi multi-skala tersebut, PSPNet mampu menangkap informasi kontekstual lokal dan global secara bersamaan. Kemampuan ini meningkatkan performa segmentasi pada area penyakit dengan ukuran lesi yang bervariasi dan distribusi spasial yang tidak beraturan. Selain itu, pemanfaatan *Pyramid Pooling Module* membantu model menghasilkan segmentasi yang lebih stabil pada area dengan latar belakang kompleks dan kontras penyakit yang rendah. Oleh karena itu, PSPNet menjadi salah satu arsitektur segmentasi semantik yang efektif untuk analisis citra pertanian dan deteksi penyakit tanaman.

2.4.3 DeepLabV3+

DeepLabV3+ merupakan arsitektur segmentasi semantik lanjutan yang diperkenalkan oleh [Chen et al. \(2018\)](#). Arsitektur ini menggabungkan struktur *encoder-decoder* dengan *Atrous Spatial Pyramid Pooling* (ASPP) untuk meningkatkan ekstraksi fitur kontekstual sekaligus mempertahankan resolusi spasial. Kombinasi tersebut memungkinkan model menghasilkan segmentasi yang lebih detail pada objek dengan batas kompleks. Dalam analisis penyakit daun padi, kemampuan mempertahankan detail spasial menjadi penting untuk mengidentifikasi area lesi secara lebih akurat pada tingkat piksel.

Salah satu komponen utama pada DeepLabV3+ adalah *atrous convolution* atau *dilated convolution*. *Atrous convolution* memperluas *receptive field* tanpa meningkatkan kompleksitas komputasi maupun mengurangi resolusi *feature map*. Operasi *atrous convolution* secara matematis dirumuskan sebagai berikut:

$$y[i] = \sum_k x[i + r \cdot k]w[k] \quad (8)$$

dengan $x[i]$ merepresentasikan sinyal fitur masukan, $w[k]$ menunjukkan filter konvolusi, dan r merepresentasikan *dilation rate* yang mengatur jarak antar elemen kernel konvolusi. Penggunaan *dilation rate* memungkinkan model menangkap informasi kontekstual yang lebih luas tanpa kehilangan detail spasial penting pada citra. Selain itu, pendekatan ini membantu meningkatkan kemampuan segmentasi pada area penyakit dengan pola lesi yang menyebar dan tidak beraturan.

Untuk menangkap informasi kontekstual pada berbagai skala spasial, DeepLabV3+ menerapkan beberapa operasi *atrous convolution* dengan nilai *dilation rate* yang berbeda melalui modul ASPP. Representasi ASPP dirumuskan sebagai berikut:

$$F_{ASPP} = [A_{r_1}(F), A_{r_2}(F), \dots, A_{r_n}(F)] \quad (9)$$

dimana $A_{r_i}(F)$ merepresentasikan keluaran *atrous convolution* yang dihasilkan menggunakan *dilation rate* r_i . Penerapan beberapa skala *dilation rate* memungkinkan model memperoleh representasi fitur multi-skala yang lebih kaya. Strategi ini membantu model mengenali area penyakit dengan ukuran lesi yang bervariasi dan distribusi spasial yang kompleks pada citra daun padi.

Fitur kontekstual multi-skala yang telah diekstraksi kemudian diproses melalui tahap *decoder* untuk merekonstruksi prediksi segmentasi yang lebih padat dengan lokalisasi batas objek

yang lebih akurat. Kemampuan mempertahankan informasi spasial sekaligus menangkap konteks global membuat DeepLabV3+ menunjukkan performa yang unggul pada berbagai tugas segmentasi semantik. Selain itu, arsitektur ini mampu menghasilkan segmentasi yang lebih stabil pada objek dengan batas yang tidak beraturan dan kontras yang rendah. Oleh karena itu, DeepLabV3+ menjadi salah satu arsitektur segmentasi semantik yang efektif untuk deteksi penyakit daun padi dengan karakteristik lesi yang kompleks dan beragam.

2.5 Metrik Evaluasi

Performa model segmentasi semantik dievaluasi menggunakan metrik *pixel accuracy* dan *Intersection over Union* (IoU) (Long et al., 2015; Chen et al., 2018). Kedua metrik tersebut digunakan untuk mengukur kemampuan model dalam mengklasifikasikan piksel serta menentukan kualitas segmentasi area penyakit pada citra daun padi. Penggunaan lebih dari satu metrik evaluasi diperlukan agar performa model dapat dianalisis secara lebih komprehensif. Selain itu, kombinasi metrik evaluasi membantu memberikan gambaran yang lebih jelas mengenai kemampuan model dalam menghasilkan segmentasi yang akurat dan konsisten. Akurasi piksel mengukur proporsi piksel yang berhasil diklasifikasikan dengan benar terhadap jumlah keseluruhan piksel pada citra. Metrik tersebut dirumuskan sebagai berikut:

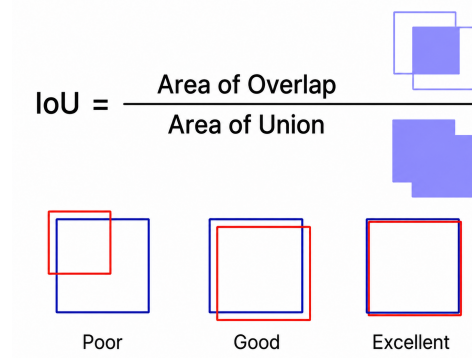
$$\text{Akurasi piksel} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

TP merepresentasikan piksel *true positive*, TN menunjukkan piksel *true negative*, FP merepresentasikan piksel *false positive*, dan FN menunjukkan piksel *false negative*. Nilai *pixel accuracy* yang tinggi menunjukkan bahwa model mampu melakukan klasifikasi piksel dengan baik pada keseluruhan citra. Namun, metrik ini memiliki keterbatasan ketika distribusi kelas tidak seimbang karena nilai akurasi dapat tetap tinggi meskipun segmentasi area penyakit belum optimal. Oleh karena itu, diperlukan metrik tambahan yang mampu mengevaluasi kualitas segmentasi secara lebih detail pada area objek.

Untuk mengatasi keterbatasan akurasi piksel, digunakan juga metrik *Intersection over Union* (IoU) sebagai metrik utama evaluasi segmentasi semantik (Jaccard, 1912; Rahman and Wang, 2016). IoU mengukur tingkat kesamaan area tumpang tindih antara mask segmentasi hasil prediksi dan anotasi *ground truth*. Metrik ini lebih sensitif terhadap kualitas segmentasi batas objek sehingga banyak digunakan dalam berbagai penelitian segmentasi citra. Dengan menggunakan IoU, performa segmentasi area penyakit dapat dievaluasi secara lebih representatif pada tingkat piksel. Metrik IoU dirumuskan sebagai berikut:

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (11)$$

Gambar 3 menunjukkan konsep *Intersection over Union* (IoU) yang digunakan dalam evaluasi segmentasi semantik. Area *intersection* merepresentasikan bagian tumpang tindih antara hasil prediksi dan *ground truth*, sedangkan area *union* merepresentasikan gabungan keseluruhan area dari kedua region tersebut. Nilai IoU yang lebih tinggi menunjukkan performa segmentasi yang lebih baik dan kemampuan lokalisasi area penyakit yang lebih akurat. Semakin besar area tumpang tindih yang dihasilkan, semakin konsisten model segmentasi semantik dalam mengidentifikasi area penyakit pada tingkat piksel. Sebaliknya, nilai IoU yang rendah menunjukkan bahwa mask segmentasi hasil prediksi memiliki perbedaan yang signifikan dibandingkan anotasi *ground truth*, sehingga segmentasi area penyakit menjadi kurang akurat.



Gambar 3: Diagram Intersection over Union (IoU)

3 Hasil dan Pembahasan

Bagian ini menyajikan hasil eksperimen yang diperoleh dari implementasi arsitektur segmentasi semantik untuk deteksi penyakit daun padi. Pembahasan difokuskan pada evaluasi performa kuantitatif, analisis perbandingan antar arsitektur yang diusulkan, serta interpretasi kemampuan segmentasi masing-masing model dalam mengidentifikasi area penyakit pada tingkat piksel. Eksperimen dilakukan menggunakan konfigurasi pelatihan yang identik dan dataset pengujian yang sama untuk memastikan perbandingan performa dilakukan secara objektif. Pendekatan tersebut memungkinkan evaluasi yang lebih adil terhadap kemampuan setiap arsitektur dalam melakukan segmentasi area penyakit daun padi.

Tabel 1 menunjukkan ringkasan performa segmentasi semantik yang dihasilkan oleh U-Net, PSPNet, dan DeepLabV3+ pada berbagai jumlah *epoch* pelatihan.

Tabel 1: Perbandingan Hasil Performa Segmentasi Semantik Model

Model	Epoch	Akurasi	IoU
U-Net	50	0.9645	0.4000
	100	0.9761	0.4006
	200	0.9745	0.4246
PSPNet	50	0.9732	0.5298
	100	0.9721	0.5824
	200	0.9776	0.5975
DeepLabV3+	50	0.9705	0.7879
	100	0.9775	0.8905
	200	0.9235	0.9804

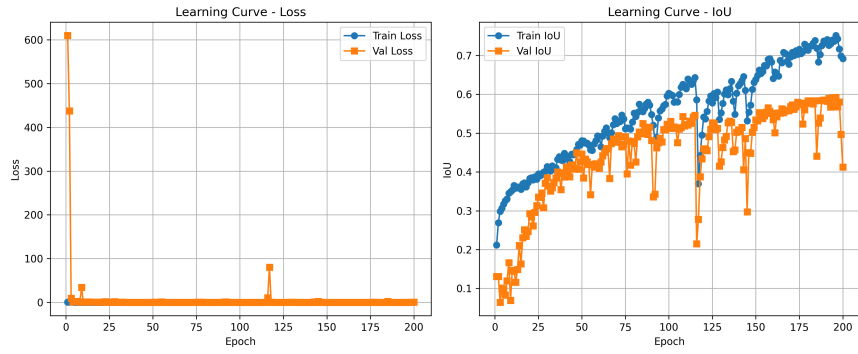
Hasil eksperimen menunjukkan adanya perbedaan kemampuan segmentasi yang signifikan di antara arsitektur yang dievaluasi. DeepLabV3+ secara konsisten menghasilkan performa terbaik pada seluruh tahap pelatihan dengan memperoleh nilai *pixel accuracy* sebesar 0.9804 dan nilai IoU sebesar 0.9235 pada *epoch* ke-200. Nilai tersebut menunjukkan bahwa arsitektur mampu mengidentifikasi area penyakit secara efektif sekaligus mempertahankan lokalisasi spasial batas lesi secara akurat. Performa tinggi yang dicapai DeepLabV3+ terutama dipengaruhi oleh mekanisme *Atrous Spatial Pyramid Pooling* (ASPP) yang memungkinkan ekstraksi infor-

masi kontekstual dari berbagai *receptive field* secara simultan (Chen et al., 2018). Pada tugas segmentasi semantik, khususnya analisis penyakit tanaman, area lesi sering muncul dengan bentuk, tekstur, dan ukuran yang bervariasi. Oleh karena itu, kemampuan mempertahankan detail spasial sekaligus menangkap fitur kontekstual global menjadi sangat penting untuk meningkatkan akurasi segmentasi.

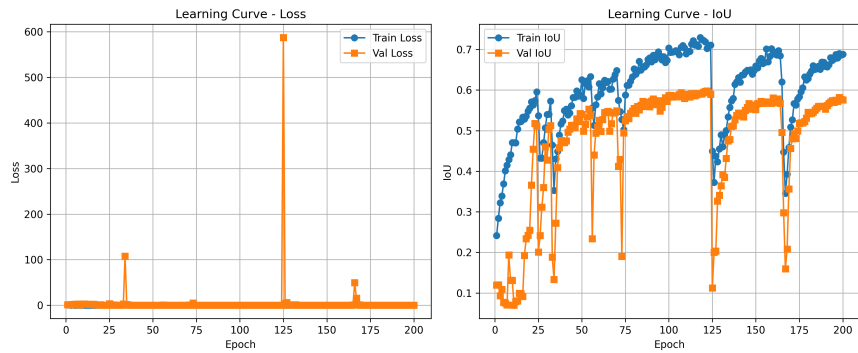
Dibandingkan dengan DeepLabV3+, U-Net menghasilkan nilai IoU yang jauh lebih rendah meskipun memperoleh nilai *pixel accuracy* yang relatif tinggi. Hasil tersebut menunjukkan bahwa U-Net mampu mengklasifikasikan area piksel dominan dengan baik, tetapi masih mengalami keterbatasan dalam melakukan segmentasi batas penyakit secara presisi. Struktur *encoder-decoder* dengan *skip connection* memungkinkan U-Net mempertahankan detail spasial halus, namun kemampuan representasi kontekstualnya masih lebih terbatas dibandingkan arsitektur yang memanfaatkan mekanisme ekstraksi fitur multi-skala. PSPNet menunjukkan performa menengah di antara U-Net dan DeepLabV3+. *Pyramid Pooling Module* (PPM) yang diterapkan pada PSPNet memungkinkan agregasi informasi kontekstual dari berbagai skala spasial sehingga menghasilkan kemampuan segmentasi yang lebih baik dibandingkan U-Net (Zhao et al., 2017). Meskipun demikian, arsitektur tersebut masih menghasilkan nilai IoU yang lebih rendah dibandingkan DeepLabV3+, yang menunjukkan bahwa pembelajaran kontekstual berbasis ASPP memberikan representasi fitur yang lebih efektif untuk tugas segmentasi penyakit daun padi dengan pola lesi yang tidak beraturan dan latar belakang yang kompleks.

Pengamatan penting lainnya dari hasil eksperimen adalah peningkatan performa yang konsisten seiring bertambahnya jumlah *epoch* pelatihan. DeepLabV3+ menunjukkan tren konvergensi yang paling stabil selama proses pelatihan, yang mengindikasikan kemampuan optimasi dan stabilitas pembelajaran fitur yang lebih baik. Sementara itu, U-Net dan PSPNet menunjukkan peningkatan performa yang lebih lambat, khususnya pada nilai IoU, yang dapat mengindikasikan keterbatasan dalam mengekstraksi fitur kontekstual yang lebih diskriminatif. Nilai IoU sebesar 0.9235 yang diperoleh DeepLabV3+ menunjukkan tingkat kesamaan yang tinggi antara mask segmentasi hasil prediksi dan anotasi *ground truth*. Hasil tersebut mencerminkan kemampuan arsitektur dalam melakukan lokalisasi area penyakit secara akurat pada tingkat piksel. Selain itu, performa yang diperoleh juga tetap kompetitif dibandingkan beberapa penelitian segmentasi semantik terbaru pada analisis penyakit tanaman (Li et al., 2022; Bagga and Goyal, 2025). Temuan ini menunjukkan bahwa kombinasi strategi pra-pemrosesan data yang tepat dengan arsitektur segmentasi semantik yang canggih mampu menghasilkan performa segmentasi yang sangat akurat tanpa memerlukan modifikasi model yang terlalu kompleks.

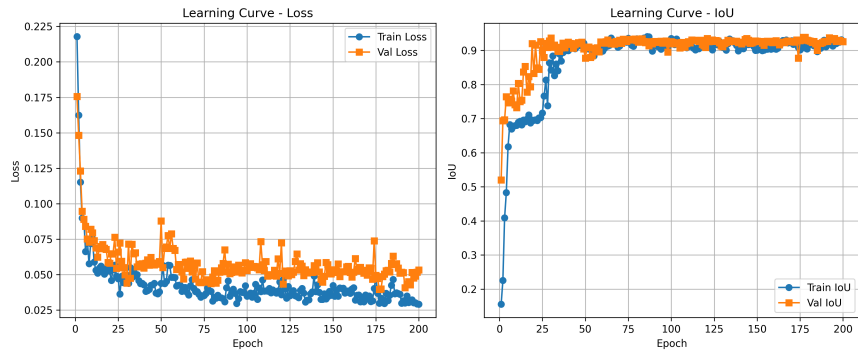
Pada penelitian ini, nilai akurasi yang dilaporkan secara khusus merepresentasikan *pixel accuracy* dan bukan akurasi klasifikasi tingkat citra. *Pixel accuracy* mengevaluasi proporsi piksel yang berhasil diklasifikasikan dengan benar pada seluruh kelas semantik, termasuk area daun sehat, area penyakit, dan latar belakang. Oleh karena itu, proses evaluasi lebih menekankan pada presisi segmentasi pada tingkat piksel yang lebih sesuai untuk tugas segmentasi semantik dengan kebutuhan lokalisasi lesi secara detail. Selain evaluasi kuantitatif, dilakukan juga analisis kualitatif melalui visualisasi kurva performa pelatihan pada masing-masing arsitektur. Gambar 4 menunjukkan kurva pembelajaran U-Net, PSPNet, dan DeepLabV3+ selama 200 *epoch* pelatihan.



(a) U-Net



(b) PSPNet

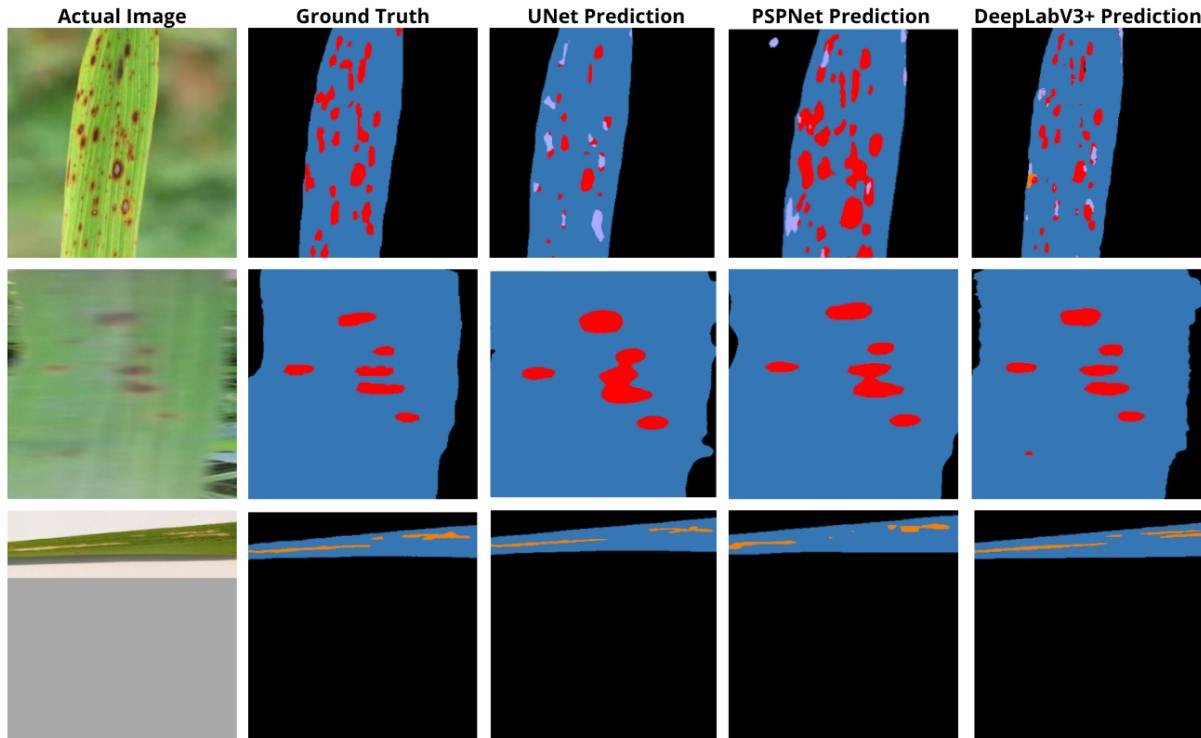


(c) DeepLabV3+

Gambar 4: Kurva Pembelajaran selama 200 *epochs* untuk (a) U-Net, (b) PSPNet, dan (c) DeepLabV3+.

Kurva pembelajaran pada Gambar 4 juga menunjukkan adanya perbedaan performa validasi yang cukup jelas di antara arsitektur yang dievaluasi. PSPNet dan DeepLabV3+ menghasilkan tren validasi IoU yang relatif stabil selama proses pelatihan, sedangkan U-Net menunjukkan fluktuasi performa validasi yang lebih besar sehingga mengindikasikan kemampuan generalisasi yang lebih lemah dan kecenderungan *overfitting* yang lebih tinggi. DeepLabV3+ menunjukkan perilaku konvergensi tercepat dengan kurva pelatihan dan validasi yang mulai stabil setelah sekitar 50 *epoch* sambil tetap mempertahankan nilai IoU yang tinggi hingga tahap akhir pelatihan.

Sebagai perbandingan, PSPNet menunjukkan pola peningkatan yang lebih bertahap, sedangkan U-Net membutuhkan durasi pelatihan yang lebih panjang untuk mencapai performa segmentasi yang relatif stabil. Temuan ini semakin menegaskan efektivitas DeepLabV3+ dalam mempelajari representasi fitur semantik yang kuat untuk tugas segmentasi penyakit daun padi (Chen et al., 2018).



Gambar 5: Hasil deteksi penyakit daun padi pada seluruh jenis penyakit

Gambar 5 menunjukkan perbandingan visual antara anotasi *ground truth* dan hasil segmentasi yang dihasilkan oleh U-Net, PSPNet, dan DeepLabV3+ pada berbagai kategori penyakit daun padi. Visualisasi tersebut menunjukkan kemampuan masing-masing arsitektur dalam mengidentifikasi area penyakit pada berbagai karakteristik lesi, tekstur yang kompleks, dan kondisi pencahayaan yang bervariasi. Secara umum, U-Net dan PSPNet mampu melokalisasi area terinfeksi, namun kedua arsitektur tersebut masih menghasilkan batas segmentasi yang kurang presisi dan prediksi *false positive*, terutama pada area dengan lesi yang saling tumpang tindih atau pola penyakit yang tidak beraturan. Hasil tersebut menunjukkan bahwa kemampuan representasi fitur lokal saja belum cukup untuk menghasilkan segmentasi yang optimal pada kondisi citra yang kompleks.

PSPNet menghasilkan area segmentasi yang lebih konsisten dibandingkan U-Net karena kemampuannya dalam melakukan agregasi fitur kontekstual. Namun, arsitektur tersebut masih menunjukkan kecenderungan *over-segmentation* pada beberapa area daun sehat sehingga mask prediksi meluas melebihi batas penyakit yang sebenarnya. Sebaliknya, DeepLabV3+ menghasilkan segmentasi yang paling akurat dan konsisten secara visual dibandingkan seluruh model yang dievaluasi. Mask prediksi yang dihasilkan DeepLabV3+ terlihat sangat mendekati anotasi

ground truth, termasuk pada sampel dengan area lesi kecil, menyebar, atau memiliki kontras rendah. Hasil tersebut menunjukkan bahwa arsitektur mampu mempertahankan informasi spasial secara efektif sekaligus menangkap fitur kontekstual global secara lebih baik (Chen et al., 2018).

Selain itu, perbandingan visual juga menunjukkan perbedaan kualitas segmentasi yang cukup jelas di antara arsitektur yang dievaluasi. U-Net menunjukkan deviasi terbesar terhadap mask *ground truth*, terutama pada area batas lesi yang sering mengalami *under-segmentation* dan prediksi yang terfragmentasi. PSPNet menunjukkan peningkatan kontinuitas segmentasi, tetapi masih cenderung melebihi estimasi area penyakit pada beberapa sampel. Sementara itu, DeepLabV3+ menghasilkan segmentasi yang lebih seimbang dengan batas yang lebih halus dan lokalisasi area penyakit yang lebih akurat. Mask segmentasi yang dihasilkan terlihat lebih stabil dan lebih sesuai dengan struktur lesi aktual pada anotasi *ground truth*.

Hasil evaluasi kualitatif tersebut konsisten dengan hasil evaluasi kuantitatif yang telah dibahas sebelumnya, di mana DeepLabV3+ memperoleh nilai IoU dan *pixel accuracy* tertinggi dibandingkan seluruh arsitektur segmentasi semantik yang dievaluasi. Performa unggul DeepLabV3+ menunjukkan efektivitas ekstraksi fitur kontekstual multi-skala untuk tugas segmentasi penyakit daun padi dengan distribusi lesi yang tidak beraturan dan latar belakang citra yang kompleks (Chen et al., 2018; Long et al., 2015). Secara keseluruhan, hasil eksperimen menunjukkan bahwa DeepLabV3+ memberikan performa segmentasi terbaik baik pada evaluasi kuantitatif maupun kualitatif. Arsitektur tersebut menunjukkan kemampuan generalisasi yang lebih kuat, perilaku segmentasi yang lebih stabil, dan lokalisasi area penyakit tingkat piksel yang lebih akurat dibandingkan U-Net dan PSPNet. Hasil penelitian ini menunjukkan bahwa DeepLabV3+ sangat sesuai untuk aplikasi deteksi penyakit daun padi berbasis segmentasi semantik pada analisis citra pertanian.

4 Kesimpulan

Penelitian ini berhasil mengimplementasikan dan mengevaluasi arsitektur segmentasi semantik untuk deteksi penyakit daun padi menggunakan U-Net, PSPNet, dan DeepLabV3+ dengan *backbone* ResNet34. Hasil eksperimen menunjukkan bahwa metode segmentasi semantik mampu mengidentifikasi area penyakit pada citra daun padi secara akurat pada berbagai karakteristik lesi dan kondisi lingkungan yang berbeda. Di antara seluruh arsitektur yang dievaluasi, DeepLabV3+ menghasilkan performa segmentasi terbaik dengan memperoleh nilai IoU dan *pixel accuracy* tertinggi dibandingkan U-Net dan PSPNet. Performa unggul DeepLabV3+ menunjukkan bahwa ekstraksi fitur kontekstual multi-skala menggunakan *Atrous Spatial Pyramid Pooling* (ASPP) mampu memberikan representasi semantik yang lebih efektif untuk tugas segmentasi penyakit daun padi. Selain itu, arsitektur tersebut juga menunjukkan perilaku konvergensi yang lebih stabil dan kemampuan generalisasi yang lebih baik selama proses pelatihan.

Hasil evaluasi kualitatif semakin menegaskan bahwa DeepLabV3+ menghasilkan mask segmentasi yang lebih konsisten dan lebih mendekati anotasi *ground truth*, terutama pada citra dengan batas lesi yang tidak beraturan, area penyakit yang menyebar, dan kondisi latar belakang yang kompleks. Sementara itu, U-Net dan PSPNet masih menunjukkan keterbatasan dalam melakukan segmentasi batas lesi secara akurat dan mempertahankan konsistensi spasial pada beberapa sampel. Secara keseluruhan, hasil penelitian menunjukkan bahwa DeepLabV3+ sangat sesuai untuk aplikasi deteksi penyakit daun padi berbasis segmentasi semantik. Pendekatan yang diusulkan berpotensi mendukung sistem pemantauan pertanian cerdas melalui analisis area penyakit daun padi yang lebih akurat dan otomatis. Penelitian selanjutnya dapat difokuskan

pada perluasan dataset, integrasi arsitektur ringan untuk aplikasi *real-time*, serta peningkatan ketahanan segmentasi pada kondisi lingkungan yang lebih beragam.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Program Studi Sains Data Universitas Negeri Surabaya serta seluruh pihak yang telah memberikan dukungan dan masukan selama pelaksanaan penelitian ini.

Referensi

- Bag, M. K., Raghu, S., Banerjee, A., Prabhukarthikeyan, S. R., Baite, M. S., and Yadav, M. (2023). Durable resistance of rice to major and emerging diseases: Present status. *Open Agriculture Journal*, 17(1).
- Bagga, P. and Goyal, D. (2025). Recent advances in deep learning-based semantic segmentation techniques. *Artificial Intelligence Review*.
- Bock, C. H., Barbedo, J. G. A., Ponte, E. M. D., Bohnenkamp, D., and Mahlein, A. K. (2020). From visual estimates to fully automated sensor-based measurements of plant disease severity: status and challenges for improving accuracy. *Phytopathology Research*, 2(1).
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *ECCV*.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
- He, Z., Zhang, Z., Valè, G., Segundo, B. S., Chen, X., and Pasupuleti, J. (2023). Editorial: Disease and pest resistance in rice. *Frontiers in Plant Science*, 14.
- Howard, K., Kachigunda, B., Spencer, R., Hewitt, C. L., and Bayliss, K. L. (2025). Rice blast in the indo-pacific region impacting food security highlights the need for better biosecurity. *NeoBiota*, 100:371–400.
- Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist*, 11(2):37–50.
- Khan, S. U., Alsuhaibani, A., Alabduljabbar, A., Almarshad, F., Altherwy, Y. N., and Akram, T. (2025). A review on automated plant disease detection: motivation, limitations, challenges, and recent advancements for future research. *Journal of King Saud University - Computer and Information Sciences*, 37(3).
- Li, Z. et al. (2022). A copy paste and semantic segmentation-based approach for the classification and assessment of significant rice diseases. *Plants*, 11(22).
- Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440.
- Mau, Y. S., Ndiwa, A., and Oematan, S. (2020). Brown spot disease severity, yield and yield loss relationships in pigmented upland rice cultivars from east nusa tenggara, indonesia. *Biodiversitas Journal of Biological Diversity*, 21(4).
- Rahman, M. A. and Wang, Y. (2016). Optimizing intersection-over-union in deep neural networks for image segmentation. *International Symposium on Visual Computing*, pages 234–244.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *MICCAI*.
- Roy, R. (2021). Plant leaf disease detection using svm. *International Journal for Research in Applied Science and Engineering Technology*, 9(5):1066–1074.

- Sane, P. and Agrawal, R. (2017). Pixel normalization from numeric data as input to neural networks: For machine learning and image processing. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 2221–2225.
- Setiawan, R., Zein, H., Azdy, R. A., and Sulistyowati, S. (2023). Rice leaf disease classification with machine learning: An approach using nu-svm. *Indonesian Journal of Data Science*, 4(3):136–144.
- Sopialena, Suyadi, Jannah, R., and Tantiani, D. (2021). Control of bacterial leaf blight disease in several varieties of rice plants (*oryza sativa* l.) by using bacteria of *paenibacillus polymyxa* mace. *IOP Conference Series: Earth and Environmental Science*, 800(1).
- Zhao, H., Shi, J., Qi, X., Wang, X., and Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6230–6239.