Pengembangan Model Segmentasi Lesi Citra Pra-Kanker Serviks Menggunakan YOLOv11-*small*

Syafaat Muhammad Bahril Huda
Jurusan Sistem Komputer,
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia

Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
syafaatbaturaja@gmail.com

Muhammad Naufal Rachmatullah*
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
naufalrachmatullah@gmail.com

Ade Iriani Sapitri
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
adeirianisapitri13@gmail.com

Akhiar Wista Arum
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
wistaarum16@gmail.com

Anggun Islami
Intelligent System Research Group,
Fakultas Ilmu Komputer,
Universitas Sriwijaya
Palembang, Indonesia
anggunislami2@gmail.com

***Abstract*—Segmentasi semantik pada citra pra-kanker serviks merupakan langkah krusial untuk diagnosis dini, namun terhambat oleh tantangan seperti kontras rendah dan batas objek yang tidak jelas. Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model segmentasi spesialis menggunakan arsitektur *You Only Look Once* v11. Pendekatan model spesialis ini diusulkan sebagai alternatif terhadap model multi-kelas tunggal, yang seringkali berpotensi mengalami kesulitan dalam mempelajari dan menyeimbangkan fitur-fitur yang sangat kontras antara kelas anatomis yang besar dengan kelas lesi yang kecil dan subtil. Pendekatan ini membandingkan strategi pelatihan model secara terpisah untuk setiap kelas : *Columnar Area*, *Cervical Area*, dan *Lesion* dengan penelitian sebelumnya yang menggunakan model multi-kelas tunggal. Metodologi penelitian melibatkan pemisahan *dataset* 307 citra pra-kanker serviks menjadi tiga subset terpisah berdasarkan kelas. Setiap subset digunakan untuk melatih model YOLOv11-*small-seg* secara independen. Kinerja model dievaluasi menggunakan metrik *Intersection over Union* dan *Dice Coefficient*. Hasil penelitian menunjukkan kinerja yang sangat tinggi untuk kelas *Cervical Area*, dengan IoU mencapai 0,8493 dan *Dice* 0,8978. Sebaliknya, model spesialis menunjukkan performa yang jauh lebih rendah untuk kelas *Lesion* (IoU 0,3991) dan kelas *Columnar Area* (IoU 0,3359). Kesimpulan dari penelitian ini adalah bahwa strategi model spesialis sangat efektif untuk struktur anatomis yang besar dan jelas seperti *Cervical Area*, namun tidak optimal untuk kelas yang lebih kecil dan kompleks seperti *Lesion* dan *Columnar Area*.**

Keywords* — *Deep Learning*, Kanker Serviks, Model Spesialis, Segmentasi Semantik, Segmentasi Citra Medis, *YOLOv11

I. PENDAHULUAN

Segmentasi semantik merupakan salah satu tantangan fundamental dalam visi komputer, khususnya dalam analisis citra medis untuk deteksi dini kanker serviks [1], [2]. Pendekatan ini berbeda secara signifikan dari segmentasi instans daripada membedakan setiap instans objek secara individual, segmentasi semantik mengklasifikasikan setiap piksel ke dalam kelas tertentu tanpa membedakan identitas objek spesifik[3], [4], [5].

Dalam konteks citra pra-kanker serviks, kemampuan untuk memisahkan wilayah anatomis penting seperti *Columnar Area* (CA), *Cervical Area*, dan *Lesion* dari latar belakang menjadi sangat krusial untuk mendukung diagnosis dini dan perencanaan tindakan medis. Akan tetapi, tantangan praktis cukup signifikan. Citra medis pada domain ini umumnya menampilkan kontras yang rendah, batas objek yang tidak jelas, dan variasi tekstur yang beragam antar pasien, sehingga membuat tugas segmentasi menjadi lebih kompleks dibanding domain citra lainnya.

Perkembangan terkini dalam arsitektur *deep learning* telah membuka peluang baru untuk meningkatkan akurasi segmentasi. *You Only Look Once* (YOLO), meskipun awalnya dikembangkan untuk deteksi objek, telah berkembang dalam versi terbarunya untuk mengintegrasikan kemampuan segmentasi secara *end-to-end* dalam satu *pipeline* yang terpadu. Pendekatan ini menarik karena YOLO menawarkan keseimbangan antara kecepatan inferensi yang tinggi dan kemampuannya beradaptasi dengan berbagai *domain* aplikasi, membuatnya relevan untuk aplikasi citra medis[8], [9].

Riset sebelumnya telah menyelidiki penggunaan satu model YOLO untuk segmentasi multi-kelas CA, *Cervical Area*, dan *Lesion* secara bersamaan. Hasil yang dilaporkan menunjukkan nilai *Intersection over Union* (IoU) sebesar 0,419 untuk kelas *Lesion*, 0,377 untuk CA, dan 0,770 untuk *Cervical Area* [10]. Meskipun hasil tersebut cukup menjanjikan, pendekatan multi-kelas menghadapi limitasi *inherent* ketika menangani perbedaan karakteristik visual yang signifikan antar kelas dimana setiap kelas memiliki tingkat kesulitan segmentasi yang berbeda, dan optimasi simultan tidak selalu menghasilkan performa terbaik untuk semua kelas [11].

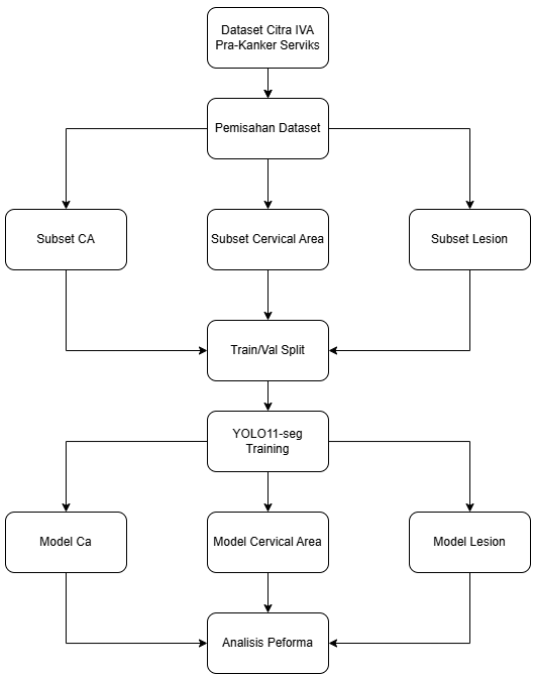
Penelitian ini mempertimbangkan strategi alternatif dengan mengembangkan model segmentasi berbasis YOLO menggunakan skema dua kelas: objek target seperti CA, *Lesion*, atau *Cervical Area* tergantung pada model spesialis yang dikembangkan dan latar belakang. Oleh karena itu, penelitian ini merumuskan hipotesis bahwa model spesialis yang dilatih secara independen akan menunjukkan

peningkatan kinerja yang signifikan dibandingkan dengan model multi-kelas tunggal yang dilaporkan sebelumnya [10]. Secara lebih spesifik, hipotesis utama yang diajukan adalah bahwa model spesialis untuk kelas *Lesion* CA yang secara visual lebih kompleks, lebih kecil, dan memiliki batas tidak jelas akan mampu mencapai nilai IoU yang secara terukur melampaui angka 0,419 untuk *Lesion* dan 0,377 untuk CA. Strategi ini tidak hanya menghasilkan *baseline* kinerja yang lebih kuat untuk segmentasi citra pra-kanker serviks, namun juga memberikan fondasi yang kokoh bagi pengembangan metode lanjutan di masa depan [3], [12].

II. METODOLOGI

A. Alur Kerja Penelitian

Metodologi penelitian ini, sebagaimana diilustrasikan pada Gambar 1, berfokus pada pengembangan model segmentasi *specialist*. Tahap awal meliputi pemisahan *dataset* citra pra-kanker serviks yang telah teranotasi menjadi tiga subset kelas: CA, *Cervical Area*, dan *Lesion*. Setiap *subset*, setelah dibagi menjadi data latih dan validasi, digunakan untuk melatih model segmentasi semantik berbasis arsitektur YOLOv11-*seg*, dengan tujuan utama menganalisis performa dari setiap model yang dilatih secara spesifik pada kelasnya masing-masing.



Gambar 1. Alur Kerja Penelitian

B. Model YOLOv11

YOLOv11 merupakan model segmentasi yang dikembangkan setelah objek deteksi nya yang sangat baik dalam mendeteksi objek. Tidak hanya dapat mendeteksi model bisa memisahkan antara objek dengan background sehingga objek membentuk kontur yang membuat kita mengenali objek apa yang sedang ditargetkan[4], [13], [14], [15]. Cara kerja YOLOv11 untuk segmentasi instan dibagi menjadi tiga bagian utama, yaitu *Backbone*, *Neck*, dan *Head.*

1. *Backbone* (Ekstraksi Fitur)

Backbone berfungsi sebagai tulang punggung untuk ekstraksi fitur dan merupakan tahap pertama dalam proses. Pada penelitian ini, arsitektur yang digunakan adalah YOLOv11-*small-seg*. Pemilihan varian *small* ini didasarkan pada pertimbangan untuk mencapai keseimbangan optimal (*trade-off*) antara akurasi dan kecepatan komputasi untuk aplikasi citra medis. Varian *small* memiliki kompleksitas arsitektur yang lebih rendah dibandingkan varian yang lebih besar (misalnya, *medium* atau *large*), sehingga memungkinkan waktu inferensi yang lebih cepat. Hal ini merupakan faktor krusial untuk potensi implementasi klinis di masa depan yang seringkali menuntut pemrosesan data yang efisien tanpa mengorbankan performa deteksi secara signifikan [11], [16]. Persamaan umum untuk operasi konvolusi adalah sebagai berikut:

$$\sum_{i=1} \omega i x i = \omega i x i + b \qquad (1)$$

2. *Neck* (Penggabungan Fitur)

Setelah ekstraksi fitur dilakukan oleh *Backbone* (jaringan tulang punggung), dihasilkan serangkaian peta fitur (*feature maps*) hierarkis. Peta fitur dari lapisan dangkal (awal) memiliki resolusi spasial yang tinggi namun minim informasi semantik. Sebaliknya, peta fitur dari lapisan dalam (*deep*) memiliki representasi semantik yang kaya namun mengalami degradasi resolusi spasial yang signifikan akibat operasi *downsampling* berulang [17].

Untuk mengatasi disparitas ini, peta fitur dari berbagai level di *Backbone* diteruskan ke bagian *Neck* (leher arsitektur). *Neck* memainkan peran krusial sebagai jembatan yang mengagregasi fitur-fitur tersebut. Tujuannya adalah untuk menciptakan sebuah representasi fitur tunggal atau multi-skala yang kaya, baik secara informasi spasial (lokasi) maupun semantik (konteks) [18], [19].

Proses agregasi ini umumnya mengadopsi mekanisme *Feature Pyramid Network* (FPN) atau varian-variannya (misalnya PANet). Mekanisme ini melibatkan dua jalur utama:

a) Jalur *Top-Down* (Atas-ke-Bawah): Fitur semantik dari lapisan terdalam terlebih dahulu melalui proses *Upsample* (peningkatan skala). Operasi ini, yang dapat diimplementasikan melalui interpolasi (*nearest-neighbor*) atau konvolusi transpos (*transposed convolution*), bertujuan untuk memperbesar dimensi spasial peta fitur agar selaras dengan dimensi peta fitur dari lapisan yang lebih dangkal.
b) Koneksi Lateral dan Konkatenasi (*Concat*): Fitur yang telah di-*upsample* tersebut kemudian digabungkan dengan peta fitur dari lapisan yang lebih dangkal (yang kaya akan informasi spasial) melalui koneksi lateral. Proses penggabungan ini umumnya menggunakan operasi *Concat* (konkatenasi), yang menumpuk *tensor* fitur pada dimensi *channel*, sehingga mengintegrasikan kedua jenis informasi tersebut secara efektif.

Representasi fitur yang telah tergabung ini selanjutnya diproses melalui blok-blok komputasi tambahan, seperti yang diidentifikasi sebagai C3K2 dan C2PSA dalam konteks arsitektur spesifik. Di dalam blok-blok ini, operasi konvolusi lanjutan digunakan untuk menyempurnakan dan merelasikan fitur-fitur yang baru teragregasi[20].

Fungsi aktivasi digunakan untuk memperkenalkan non-linearitas ke dalam model. Tanpa non-linearitas, tumpukan lapisan konvolusi hanya akan setara dengan satu transformasi linier tunggal, yang secara drastis membatasi kapasitas model untuk mempelajari pola-pola yang kompleks dan non-linier dalam data.

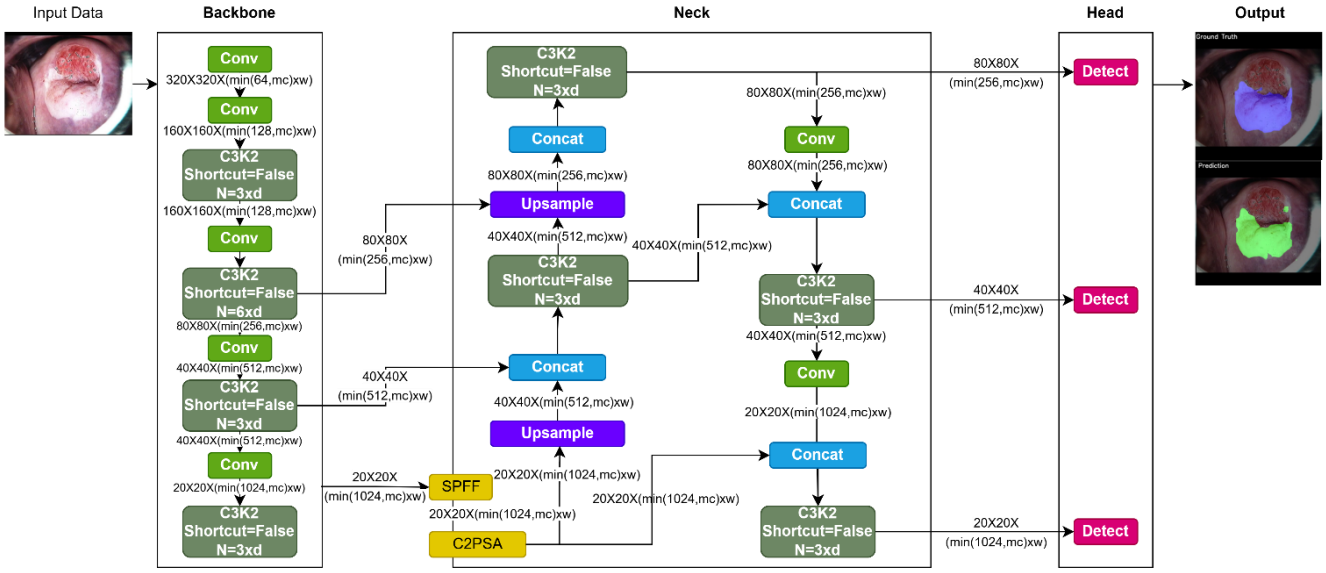
Salah satu fungsi aktivasi yang paling umum digunakan dalam arsitektur *deep learning* modern adalah *Rectified Linear Unit* (ReLU). Fungsi ini memetakan semua input negatif ke nol dan mempertahankan input positif. Secara matematis, fungsi ini didefinisikan sebagai:

$$ReLU(x) = max(0, x) \qquad (2)$$

Keunggulan utama ReLU terletak pada efisiensi komputasinya dan kemampuannya dalam mitigasi masalah *vanishing gradient* (melemahnya gradien) selama proses propagasi balik (*backpropagation*), yang memungkinkan pelatihan jaringan yang lebih dalam dan lebih cepat dibandingkan dengan fungsi aktivasi sigmoidal tradisional. Meskipun demikian, varian modern seperti *Leaky* ReLU atau SiLU (*Swish*) juga sering diimplementasikan untuk meningkatkan performa lebih lanjut.

3. *Head* (Prediksi Akhir)

Tahap terakhir adalah *Head*, yang bertanggung jawab untuk menghasilkan prediksi akhir. Arsitektur ini memiliki dua *head* yang bekerja secara paralel. *Detection Head* menerima peta fitur dari *Neck* untuk mendeteksi setiap objek, menghasilkan kotak pembatas (*bounding box*), label kelas, dan serangkaian koefisien masker (*mask coefficients*). Secara bersamaan, *Segmentation Head* (juga disebut *Proto Branch)* menghasilkan satu set masker prototipe yang bersifat umum. Masker akhir untuk setiap objek kemudian dibentuk dengan menggabungkan koefisien dari *Detection Head* dengan masker prototipe. Hasilnya adalah sebuah topeng presisi yang menguraikan bentuk objek piksel-demi-piksel, sehingga menyelesaikan tugas *segmentation* secara menyeluruh[11], [16]. Arsitektur YOLO11 secara umum ditunjukan pada Gambar 2.



Gambar 2. Arsitektur YOLOv11[21]

III. EKSPERIMEN DAN HASIL

Bagian ini akan menjelaskan proses penelitian yang dilakukan, mulai dari tahap awal hingga akhir. Fokus utama penelitian ini adalah penggunaan data pra-kanker serviks, khususnya yang teridentifikasi sebagai lesi. Proses segmentasi citra medis ini dilakukan menggunakan model YOLOv11-*seg*.

A. *Dataset* dan *Pre-processing*

Dataset yang digunakan dalam penelitian ini diperoleh dari RSUP Mohammad Hoesin dimana berisi citra serviks *pre* dan *post* dari *International Agency for Research on Cancer* (IARC), yang telah tergabung dalam satu *directory*. *Dataset* ini berjumlah 307 yang terbagi menjadi 169 data IARC dan 138 gambar dari RSUP Mohammad Hoesin [22].

Fokus penelitian ini adalah mengidentifikasi ada atau tidak nya lesi pada pasien karena lesi adalah kondisi yang abnormal dan menjadi ciri utama dalam identifikasi pra-kanker serviks.

Dalam tahap *pre-processing*, data sudah dilabel dan dibagi menjadi *train* dan *valid* pada penelitian sebelumnya menjadi 251 *train* dan 56 *valid* dengan aturan *split* 80% *train* dan 20% *valid*. Data ini kami olah kembali dengan melakukan pemecahan berdasarkan kelasnya yakni kelas CA, *Cervical Area*, dan *Lesion*. Saat melakukan pemecahan ini jumlah data mengalami pengurangan karena tidak setiap data memiliki 3 kelas tersebut. Proses memecah *dataset* menjadi sesuai kelas masing-masing ini dilakukan dengan cara konversi nilai pada data label yang kemudian di ambil hanya kelas yang ditentukan kemudian di masukan ke dalam *directory* sesuai kelasnya.

B. Setup Eksperimen

Eksperimen dilakukan menggunakan *dataset* citra pra-kanker serviks untuk melatih model spesialis pada setiap kelas (CA, *Cervical Area*, dan *Lesion*). Model YOLOv11-*small-seg* yang telah dipilih akan dilatih secara terpisah untuk setiap kelas guna mengoptimalkan kemampuannya dalam

mempelajari karakteristik unik dari masing-masing struktur anatomis.

C. Evaluasi Model

Evaluasi model adalah proses krusial untuk mengukur performa secara objektif dan menguji kemampuan generalisasi pada data *valid* yang belum pernah dilihat sebelumnya untuk mencegah *overfitting*. Kinerja model segmentasi ini dinilai menggunakan metrik utama, yaitu IoU yang mengukur tingkat kesesuaian antara area prediksi dan area sebenarnya, serta *Dice Coefficient* yang juga mengukur keselarasan serupa namun lebih sensitif terhadap kesalahan segmentasi kecil. Hasil dari kedua metrik ini memungkinkan perbandingan untuk menentukan model dengan akurasi terbaik.[23], [24], [25].

Melalui hasil evaluasi ini, perbandingan performa antar model dapat dilakukan untuk menentukan model terbaik yang paling akurat.

D. Hasil Eksperimen

Pada bagian ini berisi hasil dari penelitian yang telah dilakukan, Bagaimana performa model *specialist* dan visualisasi hasil prediksi dari model yang telah dilatih.

1) Performa Model *Specialist*

Fokus sub-bab ini adalah menyajikan hasil kuantitatif secara keseluruhan. Ini adalah ringkasan utama yang menunjukkan seberapa baik model dalam melakukan tugasnya. Hasil lengkap performa model di rangkum pada Tabel 1.

Tabel 1 Performa model *specialist*

MODEL	KELAS					
	CA		Cervical Area		Lesion	
	IoU	Dice Score	IoU	Dice Score	IoU	Dice Score
YOLO11S	0.3359	0.4277	0.8493	0.8978	0.3991	0.5128

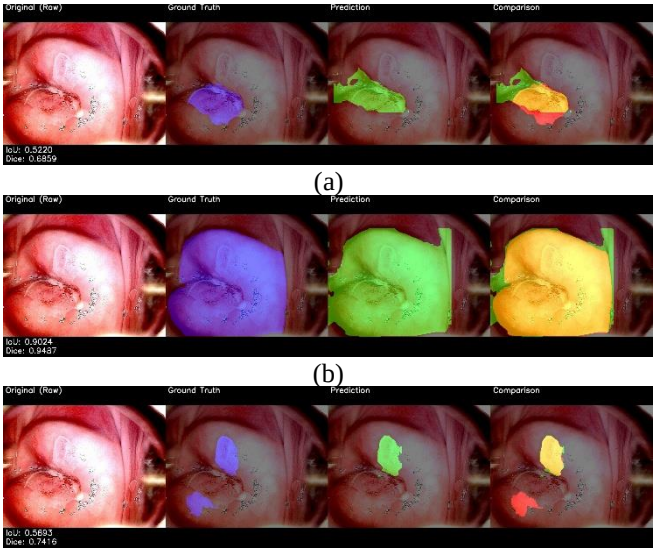
Berdasarkan tabel 1 tersebut, terlihat jelas bahwa model menunjukkan kinerja yang sangat tinggi dan akurat pada kelas *Cervical Area*. Sebaliknya, performa model tercatat secara signifikan lebih rendah untuk kelas CA dan *Lesion*. Temuan ini mengindikasikan bahwa model memiliki kinerja yang sangat terspesialisasi, dengan akurasi prediksi yang kuat untuk kelas *Cervical Area* namun belum mampu melakukan generalisasi yang efektif pada kelas lainnya.

2) Visualisasi Hasil

Evaluasi juga dilakukan melalui visualisasi hasil segmentasi untuk menginspeksi performa model secara visual. Metode ini bertujuan untuk membandingkan secara langsung antara *output* prediksi model dengan anotasi *ground truth* pada citra masukan.

Hasilnya disajikan dalam format 4 panel: (a) Citra Asli, (b) *Ground Truth* (ungu), (c) Prediksi Model (hijau), dan (d) Komparasi yang menampilkan superposisi keduanya untuk analisis tumpang tindih. Panel komparasi menggunakan skema warna untuk mengidentifikasi area irisan (*True Positive*) yang berwarna oranye, area prediksi berlebih (*False Positive*) yang berwarna hijau, dan area *ground truth* yang gagal dideteksi (*False Negative*) yang berwarna merah.

Analisis visual yang lebih komprehensif untuk setiap kelas target disajikan secara terperinci pada Gambar 3.



Gambar 3. Visualisasi Prediksi
(a) CA (b) *Cervical Area* (c) *Lesion*

Gambar 3 menyajikan komparasi visual hasil prediksi untuk setiap kelas. Teramati bahwa model mampu menghasilkan segmentasi dengan tumpang tindih (*overlap*) yang sangat tinggi untuk kelas CA (a) dan *Cervical Area* (b), yang ditandai dengan dominasi area *True Positive* (oranye).

Sebaliknya, pada kelas *Lesion* (c), teridentifikasi adanya area *False Negative* (merah) yang signifikan. Hal ini mengindikasikan bahwa performa model lebih rendah pada kelas *Lesion*, yang kemungkinan disebabkan oleh ukuran target yang lebih kecil dan variabilitas bentuk yang lebih tinggi.

3) Analisis Detail per-Kelas

Melalui hasil evaluasi ini, dilakukan analisis mendalam terhadap performa model YOLOv11s pada setiap kelas target. Analisis ini membandingkan metrik evaluasi kuantitatif yang disajikan dengan bukti visual kualitatif dari Gambar 4.

a) Kelas CA

Kelas CA mencatat performa terendah di antara ketiganya, dengan skor IoU hanya 0.3359 *dan Dice Score* 0.4277. Hasil ini menunjukkan bahwa model mengalami kegagalan signifikan dalam mengenali kelas ini.

Temuan kuantitatif ini didukung oleh bukti visual pada Gambar 4a. Berbeda dengan *Cervical Area*, panel komparasi untuk CA menunjukkan area *False Negative* (merah) yang sangat dominan. Ini mengindikasikan bahwa model hanya berhasil memprediksi sebagian kecil (area hijau/oranye) dari total area *ground truth* (area ungu/merah) yang seharusnya dideteksi. Hal ini kemungkinan disebabkan oleh fitur visual kelas CA yang terlalu subtil, ukuran nya yang kecil, dan kemiripannya dengan latar belakang.

b) Kelas *Cervical Area*

Model YOLOv11-*small-seg* menunjukkan performa yang sangat tinggi dan memuaskan pada kelas *Cervical Area*. Secara kuantitatif, kelas ini mencapai skor IoU sebesar 0.8493 dan *Dice Score* 0.8978, yang merupakan hasil tertinggi di antara ketiga kelas.

Hasil metrik ini divalidasi secara visual pada Gambar 4b. Panel komparasi menunjukkan dominasi area *True Positive* (oranye), yang mengartikan bahwa prediksi model (hijau) hampir sepenuhnya tumpang tindih dengan *ground truth* (merah). Performa superior ini dapat diatribusikan pada karakteristik kelas *Cervical Area* yang memiliki area target relatif besar dengan batas-batas (*boundaries*) yang jelas, sehingga memudahkan model untuk melakukan segmentasi secara akurat.

c) Kelas *Lesion*

Pada kelas *Lesion*, performa model mengalami penurunan signifikan, dengan perolehan IoU 0.3991 dan *Dice Score* 0.5128. Meskipun secara metrik lebih tinggi dari kelas CA, hasil ini masih tergolong rendah dan menunjukkan adanya tantangan segmentasi.

Analisis visual pada Gambar 4c mengonfirmasi temuan ini. Panel komparasi menampilkan adanya area *False Negative* (merah) yang substansial, yang berarti model gagal mendeteksi keseluruhan area lesi yang sebenarnya. Kesulitan ini dapat dijelaskan oleh karakteristik kelas *Lesion* yang secara inheren lebih kompleks: memiliki ukuran objek yang jauh lebih kecil dan bentuk yang tidak teratur (*irregular*).

IV. Diskusi

Penelitian ini mengusulkan strategi model spesialis berbasis YOLOv11-*small-seg* untuk segmentasi citra pra-kanker serviks, dengan hipotesis bahwa fokus pada satu kelas dapat meningkatkan performa dibandingkan model multi-kelas. Hasil eksperimen memberikan gambaran yang beragam.

Pada kelas *Cervical Area*, hipotesis terbukti benar. Model spesialis mencapai IoU 0,8493, sebuah peningkatan signifikan dari nilai 0,770 yang dilaporkan oleh penelitian multi-kelas sebelumnya. Performa superior ini, yang juga didukung oleh visualisasi (Gambar 4b), dapat diatribusikan pada kemampuan model untuk fokus penuh pada karakteristik kelas *Cervical Area*, yang cenderung memiliki area target relatif besar dengan batas-batas yang jelas.

Namun, untuk dua kelas lainnya, strategi ini tidak memberikan hasil yang diharapkan. Performa model spesialis untuk kelas *Lesion* (IoU 0,3991) dan kelas CA (IoU 0,3359) justru menurun dibandingkan dengan laporan penelitian multi-kelas sebelumnya *(Lesion* IoU 0,419 dan CA IoU 0,377). Temuan ini mengindikasikan bahwa penyederhanaan masalah menjadi dua kelas (target vs. latar belakang) tidak serta merta menjamin hasil yang lebih baik, terutama ketika kelas target memiliki karakteristik yang kompleks.

Analisis visual pada Gambar 4a dan 4c menunjukkan adanya area *False Negative* (merah) yang signifikan untuk kelas CA dan *Lesion.* Hal ini mengonfirmasi bahwa model gagal mendeteksi keseluruhan area target. Kegagalan ini kemungkinan besar disebabkan oleh faktor-faktor yang melekat pada kelas tersebut: ukuran objek yang kecil, bentuk yang tidak teratur (*irregular*), fitur visual yang terlalu subtil, dan kemiripan dengan latar belakang.

Keterbatasan utama dari pendekatan ini mungkin terletak pada proses pemisahan *dataset*. Seperti yang disebutkan dalam metodologi, jumlah data mengalami pengurangan setelah dipecah berdasarkan kelas. Model multi-kelas sebelumnya mungkin mendapat manfaat dari pelatihan pada *dataset* yang lebih besar dan beragam, meskipun harus menangani tiga kelas sekaligus. Sebaliknya, model spesialis untuk CA dan *Lesion* kemungkinan dilatih pada jumlah data yang tidak mencukupi, sehingga menghambat kemampuan generalisasi model. Selain itu, pemilihan varian YOLOv11-*small-seg* , meskipun efisien secara komputasi, mungkin memiliki kapasitas yang kurang untuk mempelajari fitur-fitur yang sangat subtil dari kelas CA dan *Lesion*.

Untuk penelitian di masa depan, beberapa perbaikan dapat dieksplorasi. Penambahan data yang agresif dan spesifik untuk kelas CA dan *Lesion* dapat membantu meningkatkan variasi data. Selain itu, eksplorasi varian model yang lebih besar (misalnya, YOLOv11-*medium* atau *large*) dapat dipertimbangkan untuk melihat apakah arsitektur yang lebih kompleks mampu menangkap fitur-fitur yang sulit tersebut, meskipun dengan mengorbankan kecepatan inferensi.

Kesimpulan

Penelitian ini telah mengembangkan dan mengevaluasi model segmentasi spesialis menggunakan arsitektur YOLOv11-*small-seg* untuk tiga kelas citra pra-kanker serviks: CA, *Cervical Area*, dan *Lesion*. Tujuan penelitian adalah untuk menentukan apakah model yang dilatih secara spesifik per kelas dapat mengungguli pendekatan model multi-kelas yang telah ada sebelumnya.

Hasil penelitian menunjukkan bahwa strategi model spesialis sangat berhasil untuk kelas *Cervical Area*, yang mencapai performa kuantitatif tertinggi dengan IoU 0,8493 dan *Dice Score* 0,8978. Performa ini melampaui hasil dari penelitian multi-kelas sebelumnya, menunjukkan bahwa untuk struktur anatomis yang besar dan jelas, pendekatan terfokus ini sangat efektif.

Namun, strategi ini terbukti tidak berhasil untuk kelas yang lebih menantang. Model spesialis untuk kelas *Lesion* (IoU 0,3991) dan CA (IoU 0,3359) menunjukkan performa yang rendah secara signifikan , bahkan lebih rendah dari yang dilaporkan oleh model multi-kelas. Kegagalan ini diatribusikan pada karakteristik visual kelas tersebut yang kompleks, seperti ukuran target yang kecil, bentuk yang tidak teratur, dan fitur yang subtil, serta potensi berkurangnya jumlah data pelatihan setelah pemisahan *dataset.*

Kesimpulannya, strategi model spesialis YOLOv11 efektif untuk segmentasi struktur yang jelas dan besar, namun tidak menjadi solusi optimal untuk struktur yang kecil dan subtil dalam citra pra-kanker serviks. Penelitian di masa depan harus mempertimbangkan arsitektur model yang berbeda atau teknik penanganan data yang lebih canggih untuk mengatasi tantangan pada kelas *Lesion* dan CA.

Daftar Pustaka

[1] S. Nurmaini *et al.*, "Real time mobile AI-assisted cervicography interpretation system," *Informatics in Medicine Unlocked*, vol. 42, no. July, p. 101360, 2023, doi: 10.1016/j.imu.203.101360.

[2] H. Wang *et al.*, "Cervical vertebral body segmentation in X-ray and magnetic resonance imaging based on YOLO-UNet: Automatic segmentation approach and available tool," *Digital Health*, vol. 11, pp. 1–10, 2025, doi: 10.1177/20552076251347695.

[3] M. Mihajlovic and M. Marjanovic, "Medical Segmentation of Kidney Whole Slide Images Using Slicing Aided Hyper Inference and Enhanced Syncretic Mask Merging Optimized by Particle Swarm Metaheuristics," *BioMedInformatics*, vol. 5, no. 3, pp. 1–24, 2025, doi: 10.3390/biomedinformatics5030044.

[4] R. Sapkota and M. Karkee, "Comparing YOLOv11 and YOLOv8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment," no. September 2024, pp. 1–17, 2025.

[5] R. P. S. Ribeiro and A. von Wangenheim, "Instance Segmentation in Medical Imaging: A Comparative Study of CNN and Transformer-Based Models in a Teledermatology Study-Case," pp. 819–827, 2025, doi: 10.5753/sbcas.2025.7814.

[6] R. Khanam and M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements," vol. 2024, pp. 1–9, 2024.

[7] H. Alanazi, "Optimizing Medical Image Analysis: A Performance Evaluation of YOLO-Based Segmentation Models," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 4, pp. 1167–1174, 2025, doi: 10.14569/IJACSA.2025.01604111.

[8] L. Kurnia Ramadhani and B. Nurul Widyaningrum, "Implementation of YOLO v11 for Image-Based Litter Detection and Classification in Environmental Management Efforts," 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[9] A. Ferdi, "Lightweight G-YOLOv11: Advancing Efficient Fracture Detection in Pediatric Wrist X-rays," Dec. 2024, [Online]. Available: http://arxiv.org/abs/2501.00647

[10] L. He, Y. Zhou, L. Liu, and J. Ma, "Research and Application of YOLOv11-Based Object Segmentation in Intelligent Recognition at Construction Sites," *Buildings*, vol. 14, no. 12, 2024, doi: 10.3390/buildings14123777.

[11] Y. Wu, T. Jiang, Z. Xi, F. Yin, and X. Wang, "YOLO-AR: An Improved Artificial Reef Segmentation Algorithm Based on YOLOv11," *Sensors*, vol. 25, no. 17, pp. 1–20, 2025, doi: 10.3390/s25175426.

[12] M. Kayaoğlu, A. Şengür, S. Çınarsoy Ciğerim, and S. Bor, "Detection of Cervical Vertebrae Using Object Detection and Semantic Segmentation Methods in Lateral Cephalometric Radiographs," *Türk Doğa ve Fen Dergisi*, vol. 14, no. 2, pp. 26–36, 2025, doi: 10.46810/tdfd.1594252.

[13] I. Chourib, "From Detection to Diagnosis: An Advanced Transfer Learning Pipeline Using YOLO11 with Morphological Post-Processing for Brain Tumor Analysis for MRI Images," *Journal of Imaging*, vol. 11, no. 8, pp. 1–24, 2025, doi: 10.3390/jimaging11080282.

[14] R. A. Diptho and S. Basak, "Enhancing Dermatological Diagnosis Through Medical Image Analysis: How Effective Is YOLO11 Compared to Leading CNN Models?," *Ndt*, vol. 3, no. 2, p. 11, 2025, doi: 10.3390/ndt3020011.

[15] P. Hidayatullah, N. Syakrani, M. R. Sholahuddin, T. Gelar, and R. Tubagus, "YOLOv8 to YOLO11: A Comprehensive Architecture In-depth Comparative Review," 2025.

[16] W. Zhang and S. Ji, "Rehabilitation Driven Optimized YOLOv11 Model for Medical X-Ray Fracture Detection," *Sensors*, vol. 25, no. 18, pp. 1–20, 2025, doi: 10.3390/s25185793.

[17] S. Chen, Y. Peng, Y. Liu, P. Wang, and T. Liu, "Enhancing YOLOv11 with Large Kernel Attention and Multi-Scale Fusion for Accurate Small and Multi-Lesion Bone Tumor Detection in Radiographs," *Diagnostics*, vol. 15, no. 16, Aug. 2025, doi: 10.3390/diagnostics15161988.

[18] R. Sapkota and M. Karkee, "Comparing YOLOv11 and YOLOv8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment," Jan. 2025, [Online]. Available: http://arxiv.org/abs/2410.19869

[19] R. A. Diptho and S. Basak, "Enhancing Dermatological Diagnosis Through Medical Image Analysis: How Effective Is YOLO11 Compared to Leading CNN Models?," *NDT*, vol. 3, no. 2, p. 11, May 2025, doi: 10.3390/ndt3020011.

[20] I. Chourib, "From Detection to Diagnosis: An Advanced Transfer Learning Pipeline Using YOLO11 with Morphological Post-Processing for Brain Tumor Analysis for MRI Images," *J Imaging*, vol. 11, no. 8, Aug. 2025, doi: 10.3390/jimaging11080282.

[21] R. Khanam and M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements," *ArXiv*, Oct. 2024, doi: https://doi.org/10.48550/arXiv.2410.17725.

[22] I. A. for R. on Cancer, "IARC Cervical Cancer Image Bank," International Agency for Research on Cancer (IARC). Accessed: Aug. 10, 2024. [Online]. Available: https://screening.iarc.fr/cervicalimagebank.php

[23] S. Chen, Y. Peng, Y. Liu, P. Wang, and T. Liu, "Enhancing YOLOv11 with Large Kernel Attention and Multi-Scale Fusion for Accurate Small and Multi-Lesion Bone Tumor Detection in Radiographs," *Diagnostics*, vol. 15, no. 16, pp. 1–19, 2025, doi: 10.3390/diagnostics15161988.

[24] S. Nurmaini *et al.*, "Accurate detection of septal defects with fetal ultrasonography images using deep learning-based multiclass instance segmentation," *IEEE Access*, vol. 8, pp. 196160–196174, 2020, doi: 10.1109/ACCESS.2020.3034367.

[25] S. Nurmaini *et al.*, "Automatic echocardiographic anomalies interpretation using a stacked residual-dense network model," *BMC Bioinformatics*, vol. 24, n