

Systematic Literature Review of Modern Instance segmentation : Dari CNN hingga Transformer dan Open-Vocabulary Models

Putri Maulidia¹, Salma Nurrisa²

Universitas Majalengka, Majalengka, Indonesia^{1,2}

Maulidiap91@gmail.com¹, salmanurrisa94@gmail.com²

Abstrak— Penelitian ini membahas perkembangan metode *instance segmentation* dalam domain *Computer vision* melalui pendekatan Tinjauan Literatur Sistematis. Masalah utama dalam penelitian ini adalah pesatnya pertumbuhan metode *instance segmentation* yang menyebabkan banyak variasi model, teknik, dan domain aplikasi, sehingga dibutuhkan analisis yang sistematis untuk memahami tren dan arah perkembangannya. Penelitian dilaksanakan dengan metode *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* melalui proses identifikasi, pemilihan, dan analisis terhadap 30 artikel yang sepenuhnya terdaftar di SCOPUS dan diambil dari berbagai basis data akademis seperti *ScienceDirect*, *SpringerLink*, *IEEE Xplore*, *arXiv*, *MDPI*, *Wiley Online Library*, *ACM Digital Library*, dan *Nature Scientific Reports*. Penelitian mengungkapkan bahwa teknik berbasis *deep learning* seperti *Convolutional Neural Network*, *YOLO*, *Transformer*, *model difusi*, dan model bahasa-visual telah berkembang pesat dalam meningkatkan akurasi segmentasi, efisiensi komputasi, dan kemampuan generalisasi model. Selain itu, penggunaan data sintetik, pembelajaran multimodal, dan teknik *open-vocabulary* menjadi tren utama dalam perkembangan *instance segmentation modern*. Penelitian ini menyajikan penjelasan mendalam tentang evolusi teknologi, masalah penelitian, serta kesempatan untuk pengembangan metode *instance segmentation* di masa depan.

Abstract— This study examines the development of *instance segmentation* methods in the field of *Computer vision* through a *Systematic Literature Review* approach. The main issue in this study is the rapid growth of *instance segmentation* methods, which has led to a wide variety of models, techniques, and application domains, necessitating a systematic analysis to understand trends and the direction of their development. The research was conducted using the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* guidelines through a process of identifying, selecting, and analyzing 30 articles fully indexed in SCOPUS and sourced from various academic databases such as *ScienceDirect*, *SpringerLink*, *IEEE Xplore*, *arXiv*, *MDPI*, *Wiley Online Library*, *ACM Digital Library*, and *Nature Scientific Reports*. The study reveals that deep learning-based techniques such as *Convolutional Neural Networks*, *YOLO*, *Transformers*, *diffusion models*, and *vision-language models* have advanced significantly in improving segmentation accuracy, computational efficiency, and model generalization capabilities. Additionally, the use of synthetic data, multimodal learning, and open-vocabulary techniques has emerged as a major trend in modern *instance segmentation* development. This research provides an in-depth explanation of technological evolution, research challenges, and opportunities for the future development of *instance segmentation* methods.

Keywords— *Computer vision*, *Deep Learning*, *Instance segmentation*, *Open-Vocabulary Models*, *Transformer*

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc/4.0/) license.



1. Pendahuluan

Deep learning berkembang pesat dalam bidang *computer vision* karena mampu mempelajari representasi fitur secara otomatis dan menyelesaikan permasalahan pengolahan citra yang kompleks dengan akurasi tinggi [1]. Salah satu metode yang menarik perhatian adalah *instance segmentation*, yaitu teknik yang dapat mendeteksi, mengklasifikasikan, serta memisahkan setiap objek secara individual pada level piksel. Tidak seperti *semantic segmentation* yang hanya memberikan label kelas untuk setiap piksel, *instance segmentation* dapat membedakan objek dengan kelas yang sama namun berada di lokasi berbeda dalam sebuah gambar. Teknologi ini sangat krusial karena telah digunakan di berbagai sektor seperti mengemudi otonom, pertanian presisi, citra medis, robotika, pengawasan video, serta penginderaan jauh. Evolusi metode *instance segmentation* dimulai dengan pendekatan berbasis CNN dan usulan wilayah seperti *Mask R-CNN*, kemudian meluas ke metode *real-time* seperti *YOLOACT* [2], metode berbasis *embedding* [3], hingga *Transformer* dan model bahasa-visual modern [4], [5]. Selain itu, sejumlah penelitian terbaru juga menggunakan *synthetic data*, *diffusion model*, *weakly supervised learning*, dan *open-vocabulary*

segmentation untuk memperbaiki kemampuan generalisasi model dalam situasi data yang terbatas [6], [7], [8], [9]. Studi lain mengungkapkan bahwa penerapan ekstraksi fitur multi-skala dan mekanisme perhatian dapat meningkatkan akurasi segmentasi untuk objek yang kecil dan kompleks [10], [11], [12], [13]. Perkembangan ini menunjukkan bahwa *instance segmentation* tidak hanya berkaitan dengan peningkatan akurasi, tetapi juga efisiensi komputasi, kemampuan generalisasi, dan adaptabilitas terhadap objek baru [4], [14], [15], [16].

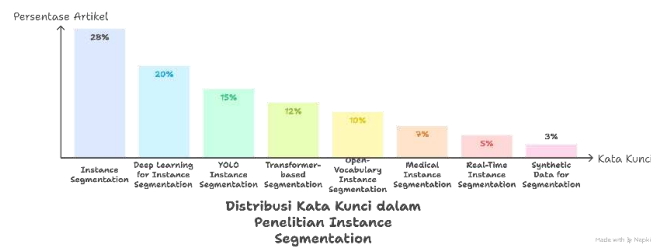
Beragam studi sebelumnya telah menciptakan metode *instance segmentation* di berbagai bidang aplikasi. Dalam sektor pertanian, segmentasi instan digunakan untuk mengidentifikasi daun tanaman, buah, gulma, dan perkembangan tanaman hidroponik secara otomatis [14], [15], [17], [18], [19], [20]. Dalam sektor medis, teknik ini digunakan untuk segmentasi tulang belakang vertebra, gigi, serta citra mikroskopis sel [12], [21], [22], [23], [24]. Di sisi lain, dalam bidang penginderaan jauh dan pengemudian otonom, *instance segmentation* digunakan untuk memisahkan marka jalan, bangunan, kanopi pohon, serta estimasi jarak kendaraan [7], [11], [25], [26], [27], [28]. Studi lain juga mengulas kemajuan teknik modern seperti segmentasi berbasis *Transformer*, *Mask Transformer*, segmentasi interaktif, segmentasi *instance video*, dan segmentasi kosakata terbuka yang memungkinkan model mengenali objek baru tanpa pelatihan khusus [4], [5], [29], [30]. Selain itu, banyak penelitian menggunakan model difusi dan generasi data sintetis untuk memperbaiki mutu data pelatihan agar kinerja model menjadi lebih konsisten [6], [8], [9]. Beragam pendekatan tersebut menunjukkan bahwa penelitian *instance segmentation* terus maju dari metode konvensional yang berbasis CNN ke arah pendekatan modern yang lebih adaptif, efisien, dan kontekstual. Meskipun perkembangan penelitian *instance segmentation* sangat pesat, masih terdapat beberapa kesenjangan penelitian yang ditemukan. Sebagian besar penelitian sebelumnya hanya berfokus pada pengembangan satu metode tertentu atau penerapan pada domain spesifik tanpa memberikan analisis menyeluruh mengenai evolusi metode *instance segmentation* dari generasi awal hingga pendekatan modern berbasis *Transformer* dan *vision-language model* [5], [10], [14], [15], [21]. Selain itu, penelitian review yang tersedia umumnya hanya membahas domain tertentu seperti pertanian atau *interactive segmentation* sehingga belum memberikan gambaran komprehensif mengenai tren perkembangan metode modern secara umum [20], [30]. Tantangan lain yang masih sering ditemukan adalah segmentasi objek kecil, *overlapping object*, kebutuhan komputasi tinggi, keterbatasan data berlabel, serta rendahnya performa model pada kondisi *real-time* [2], [11], [13], [26], [31]. Di sisi lain, perkembangan teknologi terbaru seperti *diffusion model*, *open-vocabulary segmentation*, *synthetic data generation*, dan *weakly supervised learning* mulai banyak digunakan namun belum banyak dirangkum secara sistematis dalam satu kajian literatur [4], [6], [7], [8], [9].

Berdasarkan permasalahan tersebut, penelitian ini menawarkan kebaruan berupa kajian *Systematic Literature Review* yang secara khusus membahas perkembangan modern *instance segmentation* mulai dari metode berbasis CNN, YOLO, *embedding-based segmentation*, *Transformer*, hingga *open-vocabulary* dan *vision-language models*. Berbeda dengan penelitian review sebelumnya yang hanya berfokus pada domain tertentu, penelitian ini mengintegrasikan berbagai pendekatan *instance segmentation* dari banyak bidang aplikasi seperti pertanian, medis, *autonomous driving*, robotika, *surveillance*, dan *remote sensing* [14], [17], [21], [26], [27]. Penelitian ini juga menganalisis tren penggunaan *synthetic data*, *diffusion model*, *weakly supervised learning*, *multi-scale attention*, serta *Transformer* modern yang menjadi arah perkembangan terbaru *instance segmentation* [4], [6], [7], [9], [10], [12]. Dengan demikian, penelitian ini diharapkan mampu memberikan gambaran komprehensif mengenai perkembangan metode, tantangan penelitian, peluang pengembangan, serta arah masa depan *instance segmentation* dalam bidang *computer vision*.

2. Metodologi Penelitian

Studi ini menerapkan metode *Systematic Literature Review* (SLR) untuk menemukan, menilai, dan menganalisis kemajuan teknik *instance segmentation* yang didasarkan pada *deep learning*. Metode SLR dipilih karena dapat memberikan ulasan penelitian yang sistematis, terorganisir, dan objektif mengenai beragam penelitian sebelumnya. Proses penelitian dilakukan dengan mengikuti langkah-langkah *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) yang meliputi tahap identifikasi, penyaringan, kelayakan, dan inklusi. Metode ini diterapkan untuk memastikan bahwa artikel yang dipilih benar-benar berkaitan dengan topik penelitian dan memiliki mutu ilmiah yang memadai.

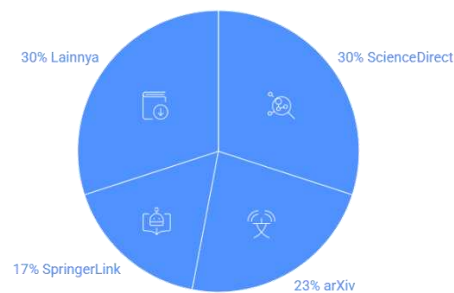
Untuk memastikan Kumpulan data yang komprehensif dan representatif, strategi pencarian disusun secara sistematis dengan menggunakan kombinasi kata kunci bidang segmentasi instance. Kata kunci utama adalah “*Instance segmentation*”, dilengkapi dengan istilah tambahan seperti “*Deep Learning for Instance segmentation*”, “*YOLO Instance segmentation*”, “*Transformer-based Segmentation*”, “*Open-Vocabulary Instance segmentation*”, “*Medical Instance segmentation*”, “*Real-Time Instance segmentation*”, serta “*Synthetic Data for Segmentation*”. Kata kunci dipilih berdasarkan perkembangan tren penelitian saat ini dalam bidang *computer vision*, terutama pada teknik segmentasi yang didasarkan pada deep learning, Transformer, dan model visi-bahasa yang sering diadopsi dalam publikasi ilmiah terbaru. Gambar 1 mempresentasikan distribusi frekuensi dari kata kunci yang digunakan dalam proses pencarian literatur.



Gambar 1. Frekuensi Kata Kunci yang Digunakan dalam Pencarian Literatur di Berbagai Basis Data Utama (2020-2026)

Sumber literatur dalam studi ini dihimpun dari berbagai basis data ilmiah internasional terindeks Scopus, meliputi ScienceDirect, SpringerLink, IEEE Xplore, MDPI, Wiley Online Library, ACM Digital Library, dan Nature Scientific Reports. Di samping itu, penelitian ini juga mengintegrasikan artikel pracetak (*preprint*) dari repositori arXiv. Langkah ini diambil secara selektif guna mengakomodasi dan menangkap tren algoritma serta model *deep learning* paling mutakhir (*state-of-the-art*) di bidang *instance segmentation* yang perkembangannya berjalan sangat akseleratif.

Distribusi Sumber Publikasi Artikel Instance Segmentation



Gambar 2. Distribusi Artikel yang dipilih Berdasarkan Sumber Basis Data

Untuk mendukung proses analisis literatur, penelitian disusun berdasarkan 10 Pertanyaan Penelitian (RQ) yang terkait dengan kemajuan metode, penerapan, serta kendala *instance segmentation*. RQ berfungsi sebagai pedoman dalam mengenali tren penelitian, membandingkan metode yang diterapkan, dan menilai arah kemajuan teknologi dalam bidang ini.



Gambar 3.Peta pikiran penelitian

- RQ1: Apa saja pendekatan utama dalam *instance segmentation* yang berkembang saat ini?
- RQ2: Bagaimana peran data dalam meningkatkan performa *instance segmentation* ?
- RQ3: Teknik apa yang digunakan untuk meningkatkan akurasi *instance segmentation* ?
- RQ4: Bagaimana peran Transformer dalam perkembangan *instance segmentation* ?
- RQ5: Bagaimana perkembangan *instance segmentation* menuju open-vocabulary?
- RQ6: Bagaimana *instance segmentation* dapat dilakukan tanpa data berlabel?
- RQ7: Apa saja domain aplikasi *instance segmentation* ?
- RQ8: Apa tantangan utama dalam *instance segmentation* ?
- RQ9: Bagaimana keseimbangan antara akurasi dan kecepatan dalam *instance segmentation* ?
- RQ10: Apa arah perkembangan masa depan *instance segmentation* ?

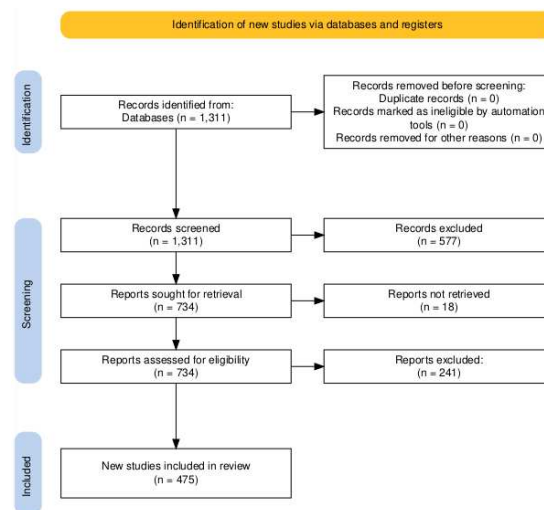
Pertanyaan penelitian ditujukan untuk mengarahkan analisis literatur secara teratur di bidang *Computer vision*, terutama *instance segmentation* yang mencakup berbagai aspek penting, termasuk evolusi metode dari pendekatan berbasis CNN ke Transformer dan model-vocabulary terbuka, peran data dan teknik untuk meningkatkan akurasi, serta penerapan di berbagai sektor seperti kesehatan, pertanian, dan transportasi. Selain itu, RQ juga berorientasi pada pengenalan tantangan utama, seperti keterbatasan data, kompleksitas objek, serta kompromi antara akurasi dan efisiensi komputasi. Dengan merumuskan RQ secara sistematis, penelitian ini bertujuan tidak hanya untuk merangkum hasil-hasil penelitian sebelumnya, tetapi juga memberikan analisis mendalam mengenai tren, keunggulan, serta arah perkembangan *instance segmentation* di masa yang akan datang.

Tabel 1. Kriteria Inklusi dan Eksklusi Seleksi Artikel

No.	Parameter Seleksi	Kriteria Inklusi (Inclusion Criteria)	Kriteria Eksklusi (Exclusion Criteria)
1.	Rentang Waktu Publikasi	Artikel ilmiah yang dipublikasikan dalam rentang waktu tahun 2020 hingga 2026 untuk merepresentasikan tren <i>instance segmentation</i> modern.	Artikel ilmiah yang dipublikasikan sebelum tahun 2020.
2.	Sumber Data & Indeksasi	Artikel ilmiah yang bersumber dari basis data akademik utama (ScienceDirect, SpringerLink, IEEE Xplore, arXiv, MDPI, Wiley, ACM, dan Nature Scientific Reports) serta telah terindeks oleh SCOPUS.	Artikel yang tidak terindeks oleh SCOPUS, literatur abu-abu (<i>grey literature</i>), artikel blog, situs komersial, maupun tesis/disertasi yang tidak dipublikasikan secara formal.
3.	Fokus & Topik Utama	Artikel yang secara spesifik membahas pengembangan, modifikasi, atau analisis komparatif metode <i>instance segmentation</i>	Artikel yang membahas bidang <i>computer vision</i> secara umum tanpa fokus spesifik pada <i>instance segmentation</i> , seperti klasifikasi citra

		(berbasis CNN, YOLO, Transformer, model difusi, model bahasa-visual, atau <i>open-vocabulary</i>).	umum, deteksi objek murni, atau <i>semantic segmentation</i> biasa.
4.	Domain Implementasi	Artikel yang menguji atau mengimplementasikan metode segmentasi pada domain pertanian presisi, citra medis, sistem mengemudi otonom (<i>autonomous driving</i>), robotika, pengawasan video (<i>surveillance</i>), atau penginderaan jauh (<i>remote sensing</i>).	Artikel yang menerapkan metode di luar domain teknis tersebut atau hanya berfokus pada pembahasan teori matematika murni tanpa implementasi visi komputer yang aplikatif.
5.	Tipe Dokumen	Dokumen berupa artikel penelitian utama (<i>research article</i>) atau makalah konferensi (<i>conference paper</i>) yang menyajikan metodologi terstruktur dan hasil eksperimen empiris.	Dokumen yang berupa artikel ulasan pustaka (<i>review paper</i>), editorial, opini, abstrak pendek pra-konferensi, atau bab buku (<i>book chapter</i>).
6.	Aksesibilitas Teks	Artikel yang menyediakan akses dokumen teks lengkap secara utuh (<i>full-text available</i>) untuk kebutuhan analisis mendalam.	Artikel yang tidak dapat diakses naskah lengkapnya (<i>full-text not retrievable</i>) karena kendala hak akses/berbayar atau hanya menyediakan ringkasan abstrak saja.
7.	Kriteria Seleksi Sekunder	Artikel dari kelompok inklusi umum yang memiliki tingkat kebaruan (<i>novelty</i>) tinggi, menyajikan arsitektur model mutakhir secara rinci, memberikan kontribusi signifikan terhadap penyelesaian masalah kritis, serta mencantumkan metrik performa (<i>mAP</i>) secara komprehensif untuk dipilih sebagai 30 artikel inti.	Artikel dari kelompok inklusi umum yang karakteristik pembahasannya dinilai repetitif, metodologinya kurang jelas, atau tidak memberikan kontribusi kebaruan yang signifikan terhadap perkembangan arsitektur <i>instance segmentation</i> modern.

Berdasarkan parameter seleksi pada Tabel 1, prosedur penyeleksian artikel dilakukan bertahap melalui tiga fase inti yaitu identifikasi (*identification*), penyaringan (*screening*), dan inklusi (*included*). Seluruh tahapan tersebut digambarkan secara sistematis pada diagram alur PRISMA di bawah ini.



Gambar 4. Pemilihan Studi Menggunakan Metode Diagram Alur PRISMA, yang Terdiri dari Tahap identifikasi, penyaringan, dan inklusi

Sumber : PRISMA 2020-compliant flow diagrams

Gambar 4 menampilkan diagram alur PRISMA [32], yang menunjukkan prosedur komprehensif dalam penyeleksian artikel, mencakup fase pencarian, inklusi, dan eksklusi. Diagram tersebut terdiri dari tiga fase inti, yaitu identifikasi (*identification*), penyaringan (*screening*), dan inklusi (*included*). Pada fase

identifikasi, tinjauan literatur dilakukan melalui berbagai basis data akademik terkemuka seperti ScienceDirect, arXiv, SpringerLink, IEEE Xplore, MDPI, Wiley Online Library, ACM Digital Library, dan Nature Scientific Reports dalam rentang publikasi tahun 2020–2026. Seluruh artikel yang diintegrasikan dalam tahapan awal penelitian ini merupakan publikasi ilmiah yang telah terindeks SCOPUS, sehingga kualitas dan kredibilitas sumber literatur dapat dipertanggungjawabkan secara akademis. Proses pengumpulan literatur dalam penelitian ini menerapkan metode PRISMA 2020 (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) untuk menjamin proses seleksi yang sistematis, terstruktur, dan transparan. Melalui pendekatan ini, hasil kajian literatur yang diperoleh menjadi lebih kredibel dan memiliki validitas akademik yang tinggi. Pada tahap awal identifikasi, pencarian berbasis kata kunci pada basis data bereputasi terindeks SCOPUS berhasil menghimpun sebanyak 1.311 artikel yang relevan dengan topik penelitian.

Selanjutnya, tahap penyaringan (*screening*) awal dilakukan terhadap seluruh artikel yang telah dikumpulkan. Pada proses ini, sebanyak 577 artikel dieliminasi karena tidak memenuhi kriteria kesesuaian awal, sehingga menyisakan 734 artikel untuk dilanjutkan ke tahap pengambilan data dan evaluasi kelayakan (*retrieval and eligibility assessment*). Pada fase tersebut, dilakukan analisis yang lebih mendalam terhadap isi artikel, termasuk metode penelitian, dataset, pendekatan model, hasil eksperimen, serta kontribusi penelitian terhadap perkembangan *instance segmentation*. Dalam prosesnya, ditemukan sebanyak 18 artikel yang tidak dapat diakses secara penuh (*full-text not retrievable*) sehingga harus dieksklusi. Evaluasi kelayakan kemudian dilanjutkan terhadap artikel yang tersisa dengan mempertimbangkan kriteria inklusi yang ketat, seperti kejelasan metode, kelengkapan informasi, serta relevansi hasil terhadap perkembangan metode *instance segmentation* modern. Dari evaluasi mendalam ini, sebanyak 241 artikel dieliminasi karena tidak memenuhi standar inklusi yang ditetapkan.

Pada tahapan *included* pertama, diperoleh sebanyak 475 artikel yang memenuhi kriteria kelayakan umum dan lolos dari penyaringan awal berbasis konten. Namun, guna menghasilkan analisis komparatif yang lebih mendalam, fokus, dan tajam, dilakukan tahapan seleksi akhir (penyaringan sekunder). Dari total 475 artikel yang memenuhi kriteria inklusi tersebut, dilakukan telaah kritis (*critical appraisal*) untuk memilih artikel yang memiliki kontribusi paling signifikan, kebaruan (*novelty*) tertinggi, serta metodologi yang paling relevan dengan fokus utama penelitian ini. Berdasarkan proses penyaringan sekunder tersebut, dipilih 30 artikel inti (*core articles*) sebagai representasi utama yang dianalisis secara komprehensif dalam studi ini.

Reduksi dari 475 artikel menjadi 30 artikel inti ini dilakukan secara objektif untuk memastikan bahwa sintesis literatur tidak hanya luas secara cakupan, tetapi juga memiliki kedalaman analisis yang optimal. Ke-30 artikel utama inilah yang kemudian dibedah secara sistematis untuk mengidentifikasi perkembangan metode *instance segmentation*, tren arsitektur model, domain implementasi, performa metode, tantangan penelitian, serta arah pengembangan teknologi di masa mendatang. Dengan penerapan metode PRISMA yang berlapis ini, proses pemilihan literatur menjadi sangat transparan, sistematis, dan terstruktur, sehingga hasil analisis yang diperoleh menjadi lebih komprehensif, objektif, dan memiliki validitas akademik yang kokoh.

3. Hasil dan Pembahasan

Analisis Sintesis Kuantitatif dan Tren Perkembangan Literatur

Untuk memetakan arah perkembangan teknologi *instance segmentation*, dilakukan agregasi dan sintesis data secara menyeluruh terhadap 30 artikel ilmiah terpilih dalam rentang waktu publikasi tahun 2019 hingga 2026. Sintesis ini difokuskan pada tiga parameter utama, yaitu kronologi tren publikasi tahunan, komparasi performa arsitektur, dan peta kecenderungan domain implementasi.

A. Tren Jumlah Publikasi Per Tahun

Analisis kronologis terhadap literatur yang ditinjau menunjukkan peningkatan volume riset yang signifikan, menggambarkan tingginya urgensi topik ini dalam domain *computer vision*. Perkembangan publikasi tahunan dipetakan dalam beberapa fase penting:

- i. Fase Fondasi Real-Time (2019–2021): Publikasi pada periode awal ini didominasi oleh perancangan model pionir satu tahap [2] serta implementasi dasar pada pengawasan video [31] dan pertanian presisi [3], [17]. Klaster ini tercatat sebesar 10% dari total literatur yang dikaji, di mana fokus utama riset tertuju pada penyelesaian masalah *trade-off* antara akurasi spasial dan efisiensi waktu komputasi.

- ii. Fase Transisi Arsitektur (2022–2024): Seiring dengan diperkenalkannya mekanisme *self-attention* global melalui model transformator 3D [5], tinjauan literatur sistematis [20], [30], serta pemanfaatan *Generative AI* untuk augmentasi data melalui model difusi [6], frekuensi publikasi melonjak drastis hingga mencapai porsi 43%. Riset pada periode ini mulai bergeser secara masif dari komputasi berbasis konvolusi (*CNN-based*) menuju arsitektur transformatif yang lebih adaptif.
- iii. Fase Maturitas Multimodal dan Open-World (2025–2026): Puncak densitas publikasi ditemukan pada kluster tahun terbaru yang berkontribusi sebesar 47% dari keseluruhan data sampel. Tren riset pada fase ini didominasi oleh konvergensi antara representasi visual dan bahasa melalui *Vision-Language Models* [4], pemanfaatan model *state space* [10], integrasi kerangka kerja open-source berbasis YOLO [14], [15], serta eskalasi pengolahan citra medis mutakhir [9], [12], [21], [22], [23] dan penginderaan jauh [11], [13], [27].

B. Perbandingan Performa Lintas Metode

Berdasarkan sintesis hasil eksperimen dari artikel yang dianalisis, performa komparatif antar-kategori metode dipengaruhi secara langsung oleh karakteristik arsitektur intinya:

- i. Pendekatan Berbasis CNN (Two-Stage): Arsitektur berbasis konvolusi dua tahap dan modifikasinya [26] mempertahankan keunggulan dalam stabilitas nilai *Average Precision* (AP), khususnya pada kondisi objek statis. Namun, ketergantungan pada *region proposal networks* menyerap latensi komputasi yang tinggi, menjadikannya kurang ideal untuk skenario waktu nyata (*real-time applications*).
- ii. Pendekatan Berbasis YOLO (One-Stage): Model segmentasi satu tahap seperti pada [14], [15] unggul dalam efisiensi kecepatan proses (*Frame per Second / FPS*) yang tinggi. Keunggulan ini dicapai melalui pemrosesan *mask* secara paralel, meskipun memiliki kelemahan bawaan berupa penurunan akurasi (*trade-off*) saat dihadapkan pada segmentasi objek yang berukuran mikro atau saling bertumpang tindih (*overlapping*).
- iii. Pendekatan Berbasis Transformer dan Mamba: Representasi fitur menggunakan mekanisme perhatian global [5], [12] atau pendekatan ruang keadaan (*state space*) [10] memberikan lompatan nilai AP tertinggi pada kondisi *dataset* yang kompleks. Melalui pemahaman konteks global, metode ini efektif mengatasi tantangan oklusi (*occlusion*). Kendati demikian, tuntutan terhadap kapasitas memori GPU dan kebutuhan data pelatihan berskala besar menjadi batasan operasional utamanya.
- iv. Model Open-Vocabulary / Vision-Language: Pendekatan teranyar ini [4], [6] mendefinisikan ulang kemampuan generalisasi model dalam mengenali objek di luar data pelatihan. Walaupun performa AP pada kelas objek yang belum pernah dilatih (*unseen classes*) masih berada di bawah model *fully-supervised*, metode ini secara radikal memangkas alokasi waktu untuk proses pelabelan data manual (*annotation cost*).

C. Kecenderungan Domain Aplikasi

Sintesis distribusi kluster artikel menunjukkan bahwa implementasi *instance segmentation* modern mengelompok kuat pada beberapa domain strategis berikut:

- i. Domain Pertanian Presisi (*Smart Agriculture*): Menjadi kluster aplikasi terbesar yang memuat sekitar 30% artikel riset. Fokus penerapan mencakup lokalisasi buah [15], fenotipe tanaman [19], deteksi tanaman pengganggu [17], estimasi volume [18], serta identifikasi patologi daun [1]. Domain ini didominasi oleh penggunaan varian YOLO dan arsitektur hibrida guna mengakomodasi mobilitas perangkat keras di lapangan yang terbatas.
- ii. Domain Citra Medis (*Medical Imaging*): Berkontribusi sebesar 27% dari literatur yang ditinjau. Aplikasi terfokus pada segmentasi struktur anatomi kompleks seperti vertebra dari CT-Scans [21], deteksi gigi pada radiografi CBCT [22], [24], penghitungan sel mikroskopis [23], pembedahan robotik [9], serta pemanfaatan model perhatian [12]. Domain ini menjadi pendorong utama adopsi arsitektur 3D untuk menjamin presisi diagnostik piksel tingkat tinggi.
- iii. Domain Penginderaan Jauh (*Remote Sensing*) & Transportasi Otonom: Memiliki porsi penyebaran sebesar 23%. Riset terkonsentrasi pada analisis citra satelit [10], ekstraksi marka jalan [26], pemetaan kondisi infrastruktur berbasis pesawat tanpa awak [28], pemetaan kanopi hutan [27], deteksi batas bangunan [25], estimasi jarak kamera fisheye [7], serta segmentasi multi-modal gabungan Kamera dan LiDAR [3]. Tantangan variasi skala objek yang ekstrem memicu tingginya adopsi fitur multi-skala di domain ini.
- iv. Domain Robotika & Pengawasan (*Surveillance*): Mengisi sisa 20% sebaran publikasi, dengan fokus riset pada pelacakan manusia pada kamera pengawas [31], peningkatan segmentasi resolusi tinggi [13], serta optimasi alur kerja diagram alir metodologis seperti PRISMA [32].

Tabel 1 menyajikan ringkasan analisis perbandingan kategori metode segmentasi instansi dengan memperhatikan karakteristik arsitektur, domain penerapan, keunggulan, dan kekurangan berdasarkan tinjauan terhadap 30 publikasi penelitian yang digunakan dalam kajian ini.

Tabel 2 : Ringkasan Perbandingan Kategori Metode *Instance segmentation*

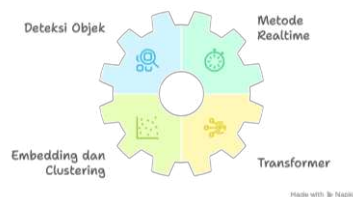
No	Kategori Metode	Contoh Model/Paper	Sorotan Arsitektur	Domain Implementasi	Kelebihan	Keterbatasan
1	CNN-based (Awal)	Mask R-CNN, Embedding-based Segmentation	Convolution dan region proposal	Autonomous Driving	Akurasi segmentasi tinggi	Komputasi besar dan inferensi lambat
2	YOLO-based	YOLACT, GenYOLO-leaf, Improved YOLOv8	Real-time parallel segmentation	Pertanian, Surveillance	Cepat dan efisien	Kurang optimal pada objek kecil
3	Transformer-based	Mask3D, swin Transformer	Self-attention dan global context	Medis, 3D Segmentation	Pemahaman konteks lebih baik	Membutuhkan sumber daya komputasi tinggi
4	Vision-Language Models	Open-Vocabulary Segmentation	Integrasi visual dan bahasa	Open-world Segmentation	Dapat mengenali objek baru	Proses training kompleks
5	Synthetic Data-based	MosaicFusion, Robotic Suturing	Synthetic data generation	Medis, Robotics	Mengurangi keterbatasan dataset	Perbedaan distribusi data nyata
6	Weakly/Unsupervised	Object-Centric Embeddings, SAPNet++	Label minim dan clustering	Microscopy, General Segmentation	Mengurangi kebutuhan anotasi	Akurasi belum stabil
7	Multi-scale Feature Learning	MambaRSIS, ST-multi-ISWNet	Multi-scale feature aggregation	Remote Sensing	Deteksi objek berbagai ukuran	Kompleksitas model meningkat
8	Medical Segmentation	ToothSeg, 3D U-Net	Segmentasi citra medis 3D	CT-Scan, Dental Imaging	Detail segmentasi tinggi	Sensitif terhadap noise citra
9	<i>Interactive segmentation</i>	Deep <i>Interactive segmentation</i>	User-guided segmentation	General Segmentation	Fleksibel dan adaptif	Membutuhkan interaksi pengguna
10	Real-time Segmentation	YOLACT, AGRO-YOLO-V	Lightweight dan fast inference	Robotics, Agriculture	Mendukung aplikasi real-time	Trade-off dengan akurasi

Berdasarkan tabel tersebut, perkembangan *instance segmentation* menunjukkan adanya pergeseran perhatian penelitian dari peningkatan akurasi segmentasi ke pengembangan model yang lebih efisien, fleksibel, dan dapat beroperasi pada berbagai domain aplikasi. Penelitian terkini berfokus pada kemampuan generalisasi model, efisiensi komputasi, penggunaan data sintetik, serta penggabungan antara visual dan bahasa untuk mendukung sistem penglihatan komputer generasi mendatang.

Diskusi Berdasarkan Pertanyaan Penelitian

Pada subbagian ini merujuk pada pertanyaan penelitian (RQ) yang telah ditetapkan untuk menyelidiki secara mendetail kemajuan metode *instance segmentation* dalam area *Computer vision*. Setiap RQ dianalisis berdasarkan sintesis dari 30 artikel yang telah dipilih, meliputi aspek perkembangan metode dari CNN hingga Transformer dan model berbasis bahasa, fungsi data dalam meningkatkan kinerja model, serta teknik-teknik yang diterapkan untuk meningkatkan akurasi dan efisiensi. Selain itu, analisis juga mengidentifikasi tantangan utama seperti kurangnya data, kompleksitas objek, dan kebutuhan komputasi yang tinggi. Melalui pendekatan ini, penelitian tidak hanya menyajikan ringkasan pustaka, tetapi juga menawarkan analisis menyeluruh terkait tren, keunggulan, serta arah perkembangan *instance segmentation* di masa mendatang

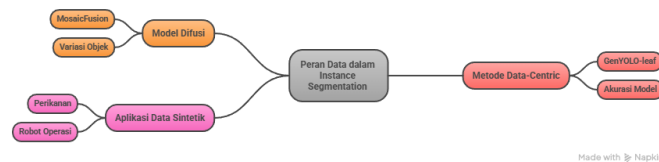
- RQ1: Apa saja pendekatan utama dalam *instance segmentation* yang berkembang saat ini?



Gambar 5. Evolusi Pendekatan Segmentasi Instance

Pendekatan *instance segmentation* saat ini dapat dibedakan menjadi beberapa kategori utama berdasarkan struktur dan cara kerja model. Pendekatan yang mengandalkan deteksi seperti Mask R-CNN masih sering dipakai karena dapat menghasilkan segmentasi yang tepat melalui proses dua tahap, yaitu deteksi objek dan pembuatan Mask [21], [26]. Akan tetapi, cara ini cenderung rumit dan memiliki durasi komputasi yang besar. Untuk mengatasi masalah tersebut, diciptakan metode realtime seperti YOLACT yang memisahkan proses pembuatan Mask menjadi dua bagian agar bisa di proses secara bersamaan dan lebih efisien [2]. Di samping itu, ada pendekatan yang menggunakan embedding dengan Teknik clustering dalam ruang fitur untuk memisahkan objek tanpa bergantung pada bounding box, seperti yang dilakukan dalam metode discriminative loss [3]. Kemajuan terbaru juga mengarah pada penerapan transformer, seperti Mask3D, yang dapat menangkap keterkaitan global dalam data dan meningkatkan kinerja dalam situasi kompleks seperti citra 3D dan medis [5]. Oleh karena itu, tren menunjukkan peralihan dari pendekatan yang rumit menuju model yang lebih efektif, adaptable, dan menyeluruh.

- RQ2: Bagaimana peran data dalam meningkatkan performa *instance segmentation* ?



Gambar 6. Peran data dalam *Instance segmentation*

Data memainkan peran krusial dalam menentukan kinerja model *instance segmentation*, terutama karena model deep learning sangat tergantung pada jumlah serta variasi data pelatihan. Salah satu metode yang sedang berkembang adalah pemanfaatan Model Difusi untuk menciptakan data sintesis, seperti yang dilakukan dalam penelitian MosaicFusion [6]. Cara ini memungkinkan model untuk dilatih dengan variasi objek yang lebih beragam, sehingga meningkatkan kemampuan generalisasi, terutama dalam situasi dengan keterbatasan data nyata. Selain itu, metode data-centric seperti yang digunakan pada GenYOLO-leaf [14] membuktikan bahwa peningkatan kualitas dan variasi dataset mampu memberikan dampak besar terhadap akurasi model dibandingkan hanya merubah arsitektur. Penggunaan data sintetik juga diterapkan di berbagai bidang seperti perikanan [8] dan robot operasi [9], yang mendukung model belajar dari kondisi yang sulit didapat secara langsung. Dengan demikian, tren terkini menekankan bahwa pengelolaan data memiliki peranan yang sama pentingnya dengan perancangan model dalam meningkatkan kinerja *instance segmentation*.

- RQ3: Teknik apa yang digunakan untuk meningkatkan akurasi *instance segmentation* ?



Gambar 7. Teknik Peningkatan Akurasi Segmentasi Instance

Peningkatan akurasi segmentasi instance dicapai dengan berbagai teknik yang menekankan perbaikan kualitas ekstraksi dan representasi fitur. Salah satu metode utama adalah ekstraksi fitur multi-skala, yang memungkinkan model mendeteksi objek dalam berbagai ukuran, seperti yang diterapkan pada MambaRSIS dan Swin Transformer [10], [12]. Metode ini sangat esensial karena objek dalam gambar sering memiliki skala yang bervariasi dan dapat saling bertumpuk. Di samping itu, penerapan mekanisme perhatian dan penggabungan fitur mendukung model untuk memahami konteks global dan rincian lokal secara bersamaan. Dalam bidang medis, metode seperti self-correction pada ToothSeg [22] memungkinkan model untuk secara otomatis memperbaiki kesalahan prediksi sehingga meningkatkan akurasi hasil

segmentasi. Penggabungan teknik-teknik ini terbukti ampuh dalam menghadapi tantangan seperti occlusion, objek kecil, dan latar belakang yang rumit.

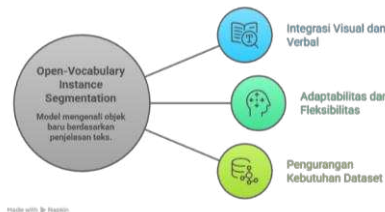
- RQ4: Bagaimana peran Transformer dalam perkembangan *instance segmentation* ?



Gambar 8. Peran Transformer dalam Segmentasi Instance

Transformer berperan penting dalam kemajuan *instance segmentation* karena kemampuannya memahami keterkaitan global antara elemen gambar. Berbeda dengan CNN yang terfokus pada area tertentu, Transformer dapat memahami konteks secara keseluruhan sehingga lebih efisien dalam mengelola objek yang kompleks dan bervariasi. Model seperti Mask3D menunjukkan bahwa pendekatan ini dapat diterapkan secara menyeluruh tanpa memerlukan proses tambahan seperti pengusulan atau pengelompokan [5]. Selain itu, penggunaan Transformer dalam model seperti Swin Transformer [12] menunjukkan peningkatan kinerja pada citra medis melalui mekanisme perhatian berbasis multi-skala. Ini memungkinkan model untuk menjaga rincian minor sambil memahami struktur keseluruhan objek. Oleh karena itu, Transformer menjadi salah satu fokus utama dalam pengembangan segmentasi instance modern berkat fleksibilitas dan kemampuannya untuk mengolah berbagai jenis data.

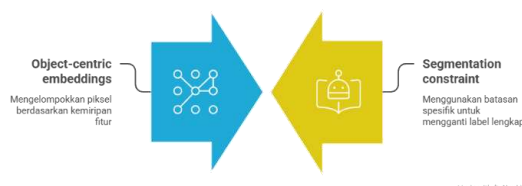
- RQ5: Bagaimana perkembangan *instance segmentation* menuju open-vocabulary?



Gambar 9. Dimensi Open-Vocabulary Instance segmentation

Perkembangan *instance segmentation* kini berfokus pada konsep open-vocabulary, di mana model tidak lagi terikat pada kategori objek tertentu yang telah dilatih sebelumnya. Metode ini mengintegrasikan pemahaman visual dan verbal, sehingga model mampu mengenali objek baru hanya berdasarkan penjelasan teks tanpa perlu pelatihan ulang [4], [29]. Hal ini menjadi krusial dalam aplikasi dunia nyata yang memiliki beragam objek yang sangat luas. Dengan pendekatan ini, model menjadi lebih adaptif dan fleksibel terhadap lingkungan yang baru. Kemampuan ini juga mengurangi kebutuhan akan dataset berlabel besar, karena model dapat menggunakan pengetahuan yang telah dipelajari sebelumnya. Oleh karena itu, *open-vocabulary* menjadi salah satu fokus utama dalam kemajuan *instance segmentation* di masa mendatang.

- RQ6: Bagaimana *instance segmentation* dapat dilakukan tanpa data berlabel?



Gambar 10. Skala Instance segmentation

Segmentasi instansi tanpa data berlabel dilakukan dengan pendekatan pembelajaran unsupervised dan weakly supervised. Dalam pendekatan tanpa pengawasan, model mengidentifikasi pola dan relasi antar piksel tanpa adanya anotasi secara manual, seperti pada teknik object-centric embeddings [23]. Model ini mengelompokkan piksel berdasarkan kemiripan fitur sehingga objek dapat terpisah secara otomatis. Sementara itu, pendekatan weakly supervised memanfaatkan informasi tambahan atau batasan tertentu untuk menggantikan label lengkap, seperti yang dilakukan dalam penelitian berbasis segmentation constraint [7]. Metode ini lebih efektif karena mengurangi kebutuhan akan anotasi data yang mahal dan memakan waktu. Oleh karena itu, metode tanpa label menjadi solusi krusial untuk meningkatkan kapasitas skalabilitas dalam *instance segmentation*.

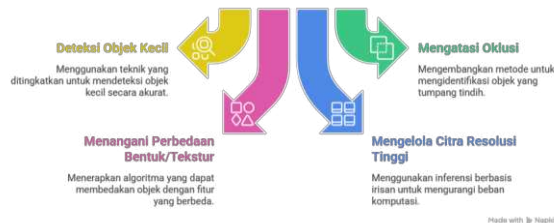
- RQ7: Apa saja domain aplikasi *instance segmentation* ?



Gambar 11. Aplikasi *Instance segmentation*

Segmentasi instansi telah digunakan di berbagai bidang karena kemampuannya dalam memisahkan objek secara rinci pada tingkat piksel. Dalam sektor pertanian, teknik ini diterapkan untuk secara otomatis mengidentifikasi tanaman dan gulma [17]. Dalam bidang kesehatan, diterapkan untuk segmentasi organ atau struktur seperti gigi dan tulang [21], [24]. Sementara itu, dalam transportasi digunakan untuk mengenali marka jalan dan mendukung sistem mengemudi otomatis [26]. Di samping itu, *instance segmentation* juga dimanfaatkan dalam penginderaan jauh untuk analisis gambar satelit [11], dan dalam sektor perikanan serta robotika untuk automasi proses [8], [9]. Ini menunjukkan bahwa teknologi ini sangat fleksibel dan bisa digunakan di berbagai sektor dengan kebutuhan yang beragam.

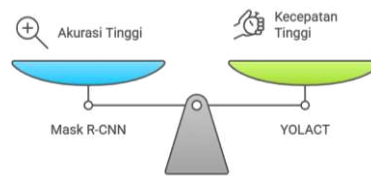
- RQ8: Apa tantangan utama dalam *instance segmentation* ?



Gambar 12. Tantangan *Instance segmentation*

Tantangan utama dalam segmentasi instance mencakup kesulitan dalam mendeteksi objek kecil, objek yang saling tumpang tindih (occlusion), serta perbedaan bentuk dan tekstur objek. Citra dengan resolusi tinggi juga menimbulkan tantangan karena memerlukan komputasi besar dan dapat menurunkan kinerja model jika tidak dikelola dengan tepat. Dalam rangka mengatasi isu ini, sejumlah penelitian merekomendasikan metode seperti inferensi berbasis irisan yang membagi gambar menjadi segmen kecil untuk diproses secara terpisah [13]. Strategi ini membuat model dapat mempertahankan akurasi tanpa secara signifikan menambah beban komputasi. Karena itu, tantangan utama dalam segmentasi instance sangat berkaitan dengan kompleksitas data dan efisiensi pemrosesan.

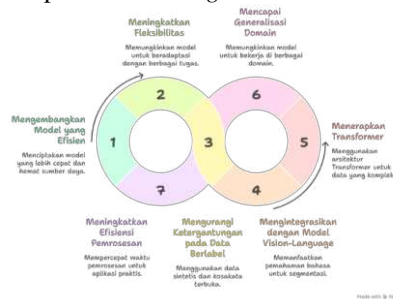
- RQ9: Bagaimana keseimbangan antara akurasi dan kecepatan dalam *instance segmentation* ?



Gambar 13. Akurasi dan Kecepatan dalam Instance segmentation

Ada kompromi antara ketepatan dan kecepatan dalam segmentasi instance. Model dengan tingkat akurasi tinggi seperti Mask R-CNN umumnya memerlukan waktu komputasi lebih lama karena mengandalkan pipeline yang rumit. Sebaliknya, model *real-time* seperti YOLACT dapat membuat prediksi dengan cepat tetapi dengan sedikit pengurangan akurasi [2]. Studi terbaru menekankan pada pengoptimalan arsitektur guna menyeimbangkan kedua elemen tersebut, sehingga model dapat diterapkan dalam aplikasi waktu nyata tanpa mengurangi akurasi secara berarti. Hal ini krusial untuk aplikasi seperti mengemudi otonom dan pengawasan video yang memerlukan reaksi segera.

- RQ10: Apa arah perkembangan masa depan *instance segmentation* ?



Gambar 14. Siklus Peningkatan Instance segmentation

Arah kemajuan segmentasi instance memperlihatkan pergeseran ke arah model yang lebih efisien, fleksibel, dan tidak bergantung pada data berlabel dalam jumlah besar. Pemanfaatan data sintesis, kosakata terbuka, dan integrasi dengan model vision-language telah menjadi tren utama dalam penelitian terkini [4], [6]. Di samping itu, penerapan Transformer juga akan terus maju berkat kemampuan mereka dalam menangani data yang rumit. Di masa mendatang, model segmentasi instance diharapkan dapat berfungsi secara umum di berbagai domain tanpa memerlukan pelatihan ulang yang besar. Di samping itu, efisiensi pemrosesan dan kemampuan secara langsung juga akan menjadi perhatian utama agar teknologi ini bisa digunakan secara luas dalam berbagai aplikasi praktis.

4. Kesimpulan

Berdasarkan hasil *Systematic Literature Review* (SLR), dapat disimpulkan bahwa metode *instance segmentation* berbasis *deep learning* telah berkembang pesat dan menjadi salah satu teknologi penting dalam *computer vision*, dengan pendekatan berbasis CNN dan Mask R-CNN masih mendominasi karena performanya yang stabil, sementara arsitektur berbasis Transformer menunjukkan potensi yang semakin besar dalam memahami konteks global citra. Tren penelitian saat ini bergeser dari sekadar meningkatkan akurasi menuju pengembangan model yang lebih efisien, ringan, mampu bekerja secara real-time, serta memiliki kemampuan generalisasi yang baik untuk implementasi pada perangkat dengan keterbatasan komputasi dan berbagai lingkungan nyata. Penerapan *instance segmentation* juga telah memberikan kontribusi signifikan pada berbagai bidang, seperti medis, pertanian presisi, *remote sensing*, *autonomous driving*, dan pemantauan industri. Oleh karena itu, penelitian mendatang disarankan untuk berfokus pada pengembangan arsitektur hibrida CNN-Transformer yang lebih efisien, penerapan pendekatan *weakly-supervised*, *semi-supervised*, atau *self-supervised learning* guna mengurangi ketergantungan pada anotasi manual, serta integrasi *vision-language learning* dan data sintetik untuk meningkatkan ketahanan model dalam mengenali objek baru yang belum pernah dilatih sebelumnya.

5. Daftar Pustaka

- [1] M. I. Rosadi and M. Lutfi, "Identifikasi Jenis Penyakit Daun Jagung Menggunakan Deep Learning Pre-Trained Model," *Explore IT: Jurnal Keilmuan dan Aplikasi Teknik Informatika*, vol. 13, no. 2, pp. 35–42, 2021, doi: 10.35891/explorit.
- [2] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-Time *Instance segmentation* ," in *Proceedings of the IEEE/CVF International Conference on Computer vision (ICCV)*, 2019, pp. 9157–9166. doi: 10.1109/ICCV.2019.00925.
- [3] K. Geng, "Deep Dual-Modal Traffic Objects *Instance segmentation* Method Using Camera and LiDAR Data for Autonomous Driving," *Remote Sens. (Basel)*, vol. 12, no. 20, p. 3274, 2020, doi: 10.3390/rs12203274.
- [4] H. Liu, "Leveraging Vision-Language Models for Open-Vocabulary *Instance segmentation* and Tracking," *IEEE Robot. Autom. Lett.*, 2025, doi: 10.1109/LRA.2025.3606363.
- [5] J. Schult, "Mask3D: Mask Transformer for 3D Semantic *Instance segmentation* ," *ArXiv*, 2022, doi: 10.1109/ICRA48891.2023.10160590.
- [6] A. Gupta, "MosaicFusion: Diffusion Models as Data Augmenters for Large Vocabulary *Instance segmentation* ," *ArXiv*, 2023, doi: 10.48550/arXiv.2309.13042.
- [7] Z. Zhang and X. Yang, "Weakly Supervised Monocular Fisheye Camera Distance Estimation with Segmentation Constraints," *Electronics (Basel)*, vol. 14, no. 17, p. 3429, 2025, doi: 10.3390/electronics14173429.
- [8] J. S. Dyrstad and E. R. Øye, "Creating Synthetic Data Sets for Training of Neural Networks for Automatic Catch Analysis in Fisheries," *Comput. Electron. Agric.*, vol. 233, p. 110160, 2025, doi: 10.1016/j.compag.2025.110160.
- [9] P. Leoncini, "A Reproducible Framework for Synthetic Data Generation and *Instance segmentation* in Robotic Suturing," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 20, no. 8, pp. 1567–1576, 2025, doi: 10.1007/s11548-025-03460-8.
- [10] L. Pan, "MambaRSIS: Context-aware multi-scale feature aggregation with selective state space model for remote sensing *instance segmentation* ," *Eng. Appl. Artif. Intell.*, 2025, doi: 10.1016/j.engappai.2025.111993.
- [11] Z. Hu, "ST-multi-ISWNet: an internal solitary wave *instance segmentation* algorithm based on spatial-temporal multi-modal under high sea states," *Int. J. Digit. Earth*, 2026, doi: 10.1080/17538947.2026.2640693.
- [12] B. M. Golkhatmi, "Multi-Scale Attention Swin Transformer for Medical Segmentation," *Sci. Rep.*, vol. 15, no. 1, p. 22649, 2025, doi: 10.1038/s41598-025-22649-0.
- [13] M. Mihajlovic and M. Marjanovic, "Enhancing *Instance segmentation* in High-Resolution Images," *Mathematics*, vol. 13, no. 19, p. 3079, 2025, doi: 10.3390/math13193079.
- [14] A. Yildirim and R. Terzi, "GenYOLO-leaf: a data-centric and open source framework for generalizable leaf *instance segmentation* across diverse datasets," *Vis. Comput.*, 2025, doi: 10.1007/s00371-025-04351-4.
- [15] J. Zhou, "Improved YOLOv8 for multi-colored apple fruit *instance segmentation* and 3D localization," *Smart Agricultural Technology*, 2025, doi: 10.1016/j.aiaa.2025.10.013.
- [16] Z. Wei, "SAPNet++: Evolving Point-Prompted *Instance segmentation* with Semantic and Spatial Awareness," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2026, doi: 10.1109/TPAMI.2026.3667694.
- [17] M. Werner, "*Instance segmentation* for the fine detection of crop and weed plants by precision agricultural robots," *Appl. Plant Sci.*, 2021, doi: 10.1145/3589132.3625658.
- [18] A. Sukkuea, "AGRO-YOLO-V: Hybrid *Instance segmentation* and Geometric Modeling for Single-View Orange Volume Estimation," *IEEE Access*, 2026, doi: 10.1109/ACCESS.2026.3667161.
- [19] M. H. Rahman, "Forecasting Growth Dynamics of Hydroponic Kale in Controlled Environment Agriculture Through Vision-Based Phenotyping and Time-Series Modeling," *Smart Agricultural Technology*, vol. 9, p. 101039, 2025, doi: 10.1016/j.atech.2025.101039.
- [20] C. Charisis and D. Argyropoulos, "Deep Learning-Based *Instance segmentation* Architectures in Agriculture: A Review of the Scopes and Challenges," *Smart Agricultural Technology*, vol. 8, p. 100448, 2024, doi: 10.1016/j.atech.2024.100448.
- [21] J. Hill, "Automated *instance segmentation* and registration of spinal vertebrae from CT-Scans with an improved 3D U-Net neural network," *Comput. Biol. Med.*, 2025, doi: 10.1016/j.compbimed.2025.110663.

- [22] N. V Nistelrooij, "ToothSeg: Robust Tooth *Instance segmentation* and Numbering in CBCT using Deep Learning and Self-Correction," *IEEE J. Biomed. Health Inform.*, 2025, doi: 10.1109/jbhi.2025.3650444.
- [23] C.-W. Wang, "A Survey of Deep Learning Methods on Cell *Instance segmentation* ," *Neural Comput. Appl.*, vol. 37, no. 17, pp. 11195–11264, 2025, doi: 10.1007/s00521-025-11119-3.
- [24] G. Jader, "Deep *instance segmentation* of teeth in panoramic X-ray images," in *IEEE SIBGRAP Conference*, 2023. doi: 10.1109/SIBGRAP.2018.00058.
- [25] M. Werner, "Bavaria Buildings - A Novel Dataset for Building Footprint Extraction, *Instance segmentation* , and Data Quality Estimation," in *Proceedings of the ACM*, 2023. doi: 10.1145/3589132.3625658.
- [26] J. Tian, "Road Marking Detection Based on Mask R-CNN *Instance segmentation* Model," in *IEEE CVIDL*, 2020. doi: 10.1109/CVIDL51233.2020.00-92.
- [27] J. L. Vaschetti, E. Arnaudo, and C. Rossi, "TreePseCo: Scaling Individual Tree Crown Segmentation using Large Vision Models," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLVIII-M-7-2025, pp. 275–282, 2025, doi: 10.5194/isprs-archives-XLVIII-M-7-2025-275-2025.
- [28] M. D. Bui, "Drone-Based Road Marking Condition Mapping," *Drones*, vol. 10, no. 2, p. 77, 2026, doi: 10.3390/drones10020077.
- [29] C. Xu, "Deep Learning Techniques for Video *Instance segmentation* : A Survey," *ArXiv*, 2023, doi: 10.48550/arXiv.2310.12393.
- [30] Z. Marinov, "Deep *Interactive segmentation*: A Systematic Review," *arXiv preprint*, pp. 1–27, 2023, doi: 10.48550/arXiv.2311.13964.
- [31] C. Tseng, "Person Retrieval in Video Surveillance Using Deep Learning–Based *Instance segmentation* ," *J. Healthc. Eng.*, vol. 2021, p. 9566628, 2021, doi: 10.1155/2021/9566628.
- [32] N. R. Haddaway, M. J. Page, C. C. Pritchard, and L. A. McGuinness, "PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis," *Campbell Systematic Reviews*, vol. 18, p. e1230, 2022, doi: 10.1002/cl2.1230.

6. Penulis



Putri Maulidia
Fakultas Teknik, Universitas Majalengka, Majalengka, Indonesia
Penulis merupakan Mahasiswa Informatika di Universitas Majalengka



Salma Nurrisa
Fakultas Teknik, Universitas Majalengka, Majalengka, Indonesia
Penulis merupakan Mahasiswa Informatika di Universitas Majalengka