

# Integrasi YOLOv9 Dan Fine-Tuned Segment Anything Model Untuk Pengenalan Komponen Alat Pengukur Analog

**Diah Asmawati<sup>1\*</sup>, Chastine Fathichah<sup>2</sup>**

<sup>1,2</sup>Departemen Teknik Informatika, Institut Teknologi Sepuluh Nopember Surabaya  
Jalan Raya ITS, Sukolilo, Surabaya 60111  
6025232014@student.its.ac.id<sup>1</sup>, chastine@its.ac.id<sup>2</sup>

## Info Artikel

### *Riwayat Artikel:*

Diterima: 4 November 2025

Direvisi: 6 Desember 2025

Disetujui: 30 Januari 2026

Diterbitkan: Maret 2026

### *Kata Kunci:*

Alat Pengukur Analog;  
Segment Anything Model;  
Segmentasi;  
YOLOv9.

## ABSTRAK

Otomatisasi pembacaan alat pengukur analog merupakan tantangan yang signifikan di lingkungan industri, terutama karena citra alat pengukur sering diambil dalam kondisi pencahayaan dan kualitas visual yang bervariasi. Penelitian ini mengusulkan pendekatan hibrida berbasis visi komputer yang mengintegrasikan model deteksi YOLOv9 dan model segmentasi Segment Anything Model yang telah disesuaikan menggunakan dataset citra alat pengukur analog. YOLOv9 digunakan untuk mendeteksi area komponen penting seperti jarum dan garis-garis skala, menghasilkan area kotak pembatas yang berfungsi sebagai masukan prompt bagi SAM. Model SAM kemudian melakukan segmentasi presisi tinggi pada area tersebut untuk memisahkan objek utama dari latar belakang citra. Dataset pelatihan terdiri dari 1010 citra dengan anotasi mask biner. Hasil evaluasi menunjukkan bahwa model Fine-Tuned SAM ViT-B mencapai nilai Precision 0.81, Recall 0.946, dan IoU sebesar 0.778 melampaui performa model SAM dasar dan YOLOv9 Segmentation. Segmentasi jarum penunjuk juga menunjukkan peningkatan signifikan dengan nilai IoU mencapai 0.837. Pendekatan menunjukkan efektivitas integrasi deteksi dan segmentasi berbasis model transformer dalam menghasilkan sistem pengenalan komponen alat pengukur analog yang presisi, efisien dan adaptif untuk aplikasi pembacaan nilai otomatis di industri.

## *Penulis Korespondensi:*

Diah Asmawati,

Departemen Teknik Informatika, Institut Teknologi Sepuluh Nopember, Surabaya  
6025232014@student.its.ac.id

## 1. PENDAHULUAN

Perangkat digital merupakan teknologi yang lazim digunakan dalam sistem pengukuran di lingkungan industri. Namun, dalam praktiknya hal ini belum sepenuhnya menghilangkan penggunaan alat pengukur analog. Keandalan mekanis serta biaya perawatan yang relatif rendah menjadikan alat analog tetap menjadi pilihan utama [1]. Meskipun demikian, proses pembacaan nilai secara manual masih membutuhkan kehadiran operator dan rentan terhadap kesalahan manusia, terutama ketika berada di lingkungan yang berisiko tinggi[2][3].

Otomatisasi pembacaan alat pengukur analog menghadapi berbagai tantangan, terutama dalam pengolahan citra yang diperoleh dari perangkat kamera dengan berbagai macam kualitas. Kondisi pencahayaan yang berubah-ubah, variasi sudut pengambilan gambar, serta derau visual dapat menghambat proses identifikasi komponen penting seperti jarum indikator dan angka skala [4]. Oleh karena itu, diperlukan pendekatan berbasis visi komputer yang tangguh terhadap variasi lingkungan agar sistem mampu memberikan hasil yang konsisten.

Dalam konteks pembacaan alat pengukur analog secara otomatis, semakin detail dan presisi komponen alat pengukur yang dikenali oleh sistem, semakin akurat pula nilai yang dapat diekstraksi. Deteksi dan segmentasi yang tepat pada area seperti jarum, garis skala, serta area numerik memungkinkan penghitungan sudut atau posisi jarum secara matematis untuk memperoleh nilai aktual alat pengukur.

Untuk mengidentifikasi komponen yang lebih detail dan spesifik, beberapa penelitian sebelumnya memanfaatkan segmentasi semantik sebagai alternatif. Objek yang biasanya disegmentasi meliputi jarum penunjuk dan garis-garis skala, bahkan terdapat penelitian yang memanfaatkan segmentasi latar belakang untuk mengoreksi sudut penampang alat [5]. Metode berbasis CNN yang dimodifikasi dan dikembangkan untuk hasil segmentasi yang semakin akurat. Algoritma bernama *Value Acquisition Module* (VAM) yang mengaplikasikan gabungan ResNet18 dan *Feature Pyramid Network* membaginya ke dalam dua tugas untuk segmentasi pengambilan komponen jarum dan secara paralel juga untuk pengenalan angka [6].

Seiring perkembangan metode segmentasi, muncul berbagai pendekatan yang melahirkan algoritma baru untuk meningkatkan kinerjanya. Metode U-Net merupakan salah satu yang secara luas digunakan untuk segmentasi objek. Salah satu turunan metode U-Net yaitu U2Net juga diterapkan pada beberapa penelitian terkait alat pengukur untuk mengekstraksi komponen[5][7]. Selain itu, pendekatan *encoder-decoder* seperti DeepLabV3+, yang merupakan pengembangan dari seri DeepLab, turut digunakan dalam segmentasi semantik jarum pengukur untuk memperoleh hasil yang lebih akurat dan stabil[8].

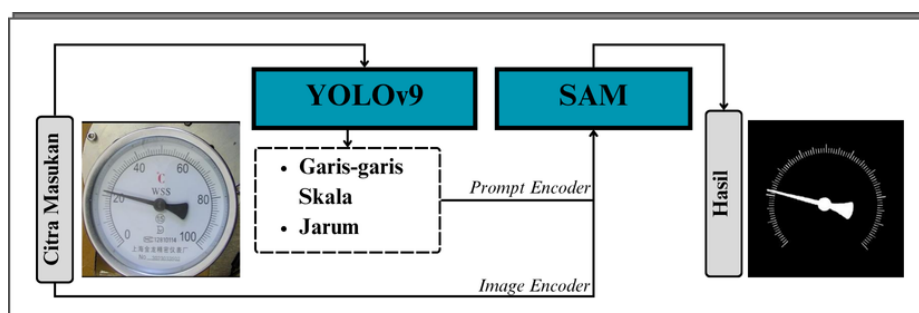
Perkembangan pesat dalam bidang visi komputer telah menghasilkan berbagai model canggih yang mampu memahami konteks visual secara mendalam. Salah satu inovasi di bidang ini adalah Segment Anything Model (SAM) yang dikembangkan oleh Meta AI [9]. Beberapa studi menunjukkan kinerja SAM yang tinggi mampu diimplementasikan di berbagai bidang seperti bidang medis dan infrastruktur bangunan [10][11][12]. Pada kasus segmentasi pada aliran sungai dalam pengindraan jauh jarak dekat, pendekatan *encoder-decoder* yang digunakan SAM mampu mengungguli beberapa teknik lain seperti U-Net, PSPNet, DeepLabV3+, PAN dan LinkNet[13]. Kemampuan SAM dalam menghasilkan mask piksel yang presisi menjadikan kandidat kuat untuk diterapkan pada citra alat pengukur analog yang kompleks, terutama setelah dilakukan penyesuaian (*fine-tuning*) agar dapat menyesuaikan dengan karakteristik alat pengukur di domain industri.

Selain itu, terdapat pula teknologi deteksi yang semakin mutakhir dengan model *You Only Look Once* (YOLO). YOLOv9 sebagai bagian dari evolusi YOLO menawarkan kecepatan dan efisiensi tinggi dalam mendeteksi objek pada citra dengan akurasi yang kompetitif [14][15]. Meskipun tersedia berbagai varian YOLO, seperti YOLOv8 yang telah stabil dan muncul lebih dulu, maupun yang lebih kini seperti YOLOv11, pemilihan YOLOv9 dalam penelitian ini didasarkan pada karakteristik arsitektur yang lebih sesuai untuk deteksi objek berukuran kecil dan medium, serta berstruktur halus, seperti jarum dan garis skala pada alat pengukur [16]. YOLOv9 memperkenalkan *Programmable Gradient Information* (PGI) dan *Generalized Efficient Layer Aggregation Network* (GELAN) yang mampu meningkatkan aliran informasi dan sensitivitas fitur. Selain itu, Dengan kemampuan prediksi waktu nyata, YOLOv9 dapat berfungsi sebagai modul deteksi awal untuk menemukan area komponen penting meliputi jarum dan garis-garis skala sebelum dilakukan segmentasi detail oleh SAM. Integrasi kedua model ini membentuk kerangka kerja bertahap dan menjadi kontribusi dalam penelitian ini. Dengan memanfaatkan kekuatan masing-masing sehingga menghasilkan sistem pengenalan komponen alat pengukur

analog yang presisi dan dapat diadaptasi untuk aplikasi pembacaan nilai alat pengukur secara otomatis.

## 2. METODE PENELITIAN

Pengenalan komponen alat pengukur analog pada penelitian ini melewati beberapa tahapan. Pertama, sistem menyiapkan citra masukan berupa penampang alat ukur analog secara utuh. Selanjutnya, model YOLOv9 digunakan untuk mendeteksi objek jarum dan garis-garis angka pada citra, menghasilkan kotak pembatas (*bounding box*) yang menandai area masing-masing objek. Hasil deteksi tersebut kemudian digunakan sebagai *prompt* masukan bagi model SAM, yang memfokuskan proses segmentasi dari citra yang sama. Proses segmentasi ini menghasilkan masker biner yang membedakan latar belakang citra dengan objek utama jarum dan garis-garis skala. Keseluruhan tahapan proses pengenalan komponen dalam studi ini ditunjukkan secara visual pada Gambar 1.



Gambar 1. Arsitektur sistem pengenalan komponen alat pengukur analog menggunakan YOLOv9 dan SAM.

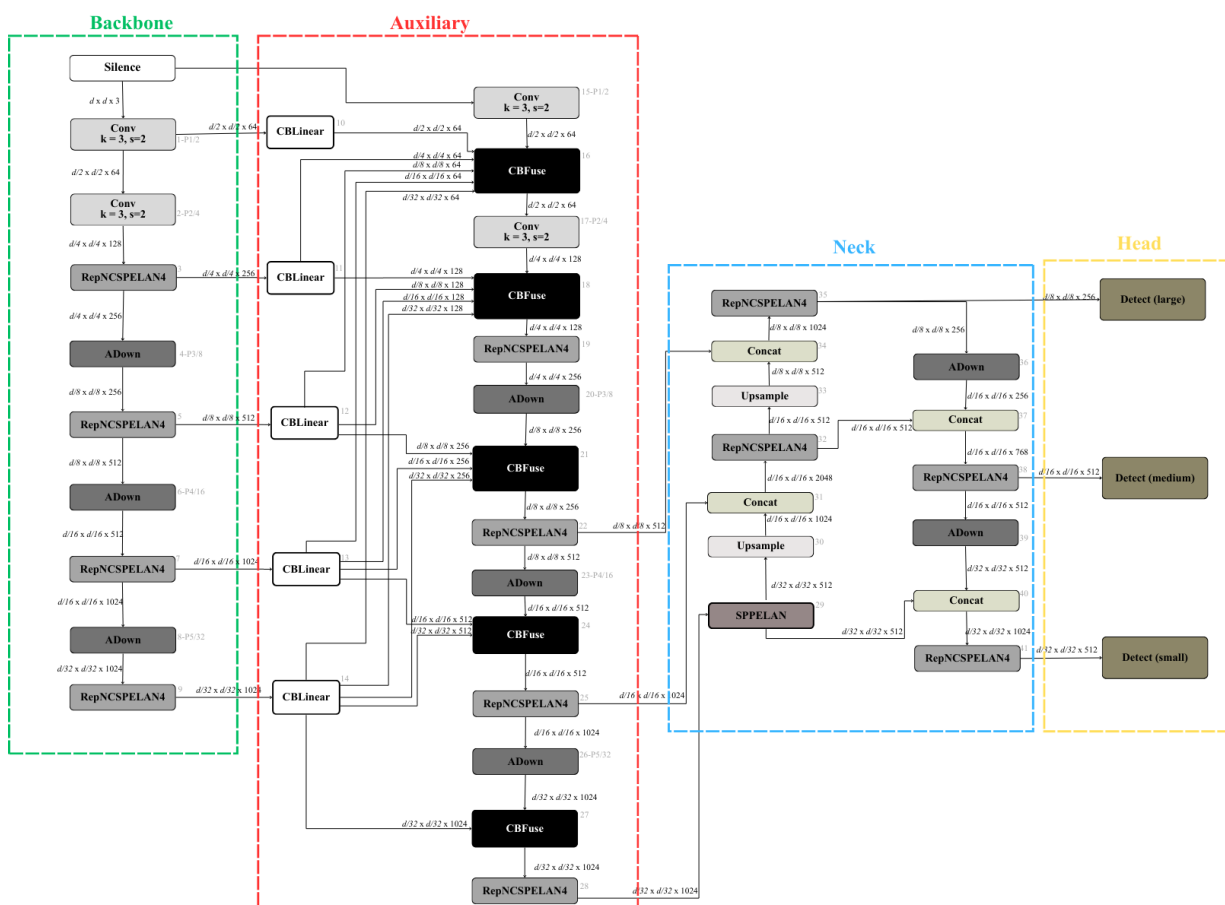
### 2.1. Deteksi Area Komponen YOLOv9

Proses deteksi komponen dalam studi ini melibatkan model YOLOv9. YOLOv9 merupakan model terbaru dari seri YOLO yang dirilis pada tahun 2024 [14]. Model ini adalah pengembangan dari seri YOLOv7 dan YOLOv8, dengan fokus pada peningkatan kecepatan, akurasi, dan efisiensi komputasi. Algoritma ini mampu mendeteksi objek dalam waktu nyata dengan akurasi yang tinggi [15]. YOLOv9 memperkenalkan dua inovasi dalam arsitekturnya yaitu *Programmable Gradient Information (PGI)* dan *Generalized Efficient Layer Aggregation Network (GELAN)*. PGI dirancang untuk meningkatkan stabilitas pelatihan model dan membantu mengatasi masalah *information bottleneck* demi mempertahankan informasi penting saat melewati lapisan pada jaringan. PGI terdiri dari tiga komponen utama : *Main Branch, Auxiliary Reversible Branch, dan Multi-level Auxiliary Information*. Di sisi lain, GELAN merupakan gabungan dua arsitektur jaringan yaitu CSPNet dan ELAN yang sudah dikenalkan pada arsitektur YOLOv7. Arsitektur varian model YOLOv9 yaitu GELAN-E yang digunakan dalam penelitian ini dapat dilihat pada Gambar 2.

Model YOLOv9 dalam kasus penelitian ini memprediksi kotak pembatas komponen jarum, garis-garis skala dan teks angka pada alat pengukur analog yang ditemukan dalam gambar (bisa satu objek atau lebih). Prediksi terdiri dari nilai koordinat titik  $(x,y)$ , lebar dan tinggi area  $(w,h)$ , serta *confidence score* pada kotak pembatas. Koordinat  $x,y$  sesuai dengan titik pusat kotak relatif terhadap batas sel grid. Lebar  $w$  dan tinggi  $h$  diprediksi relatif terhadap keseluruhan gambar. Sedangkan confidence adalah nilai IoU antara kotak pembatas prediksi dengan nilai sebenarnya.

Pada tahapan segmentasi komponen dalam penelitian ini, area yang difokuskan hanya garis skala dan jarum penunjuk sebagai hasil prediksi utama. Sementara itu, teks skala dimanfaatkan sebagai *Region Of Interest (ROI)* untuk proses pengenalan karakter dalam rangka memperoleh nilai akhir pembacaan alat pengukur [17][18]. Detail hasil prediksi area yang dihasilkan pada proses YOLOv9 dapat dilihat pada Gambar 3. Untuk menilai performa dan akurasi deteksi

YOLOv9 secara kuantitatif, *Average Precision (AP)* digunakan sebagai metrik evaluasi. AP mempertimbangkan kurva presisi dan sensitivitas berdasarkan ambang batas skor *confidence*.

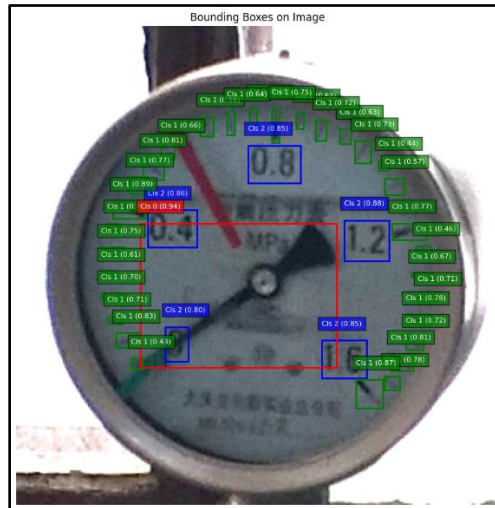


Gambar 2. Arsitektur YOLOv9 (Model Gelan-E).

## 2.2. Segmentasi Komponen SAM

*Segment Anything Model (SAM)* adalah sebuah model yang dapat diberi masukan instruksi (*promptable*) berbasis encoder-decoder yang inovatif dan dikembangkan oleh tim Meta AI untuk tugas segmentasi citra. Model ini melibatkan pelatihan pada set data yang masif SA-1B. Set data ini mengandung lebih dari 1 miliar mask yang didapat dari 11 juta gambar. Dengan demikian model mampu meningkatkan pengelompokan objek gambar yang tidak terlihat [9].

SAM terdiri dari tiga komponen utama yaitu Pengode Citra (*Image Encoder*), Pengode Prompt (*Prompt Encoder*), dan Dekoder Mask (*Mask Decoder*). Pengode *prompt* dapat diklasifikasikan ke dalam dua set isyarat : *sparse* (titik,kotak,teks) dan *dense*(mask). Titik dan kotak direpresentasikan dalam pengodean posisi yang dijumlahkan dengan penyisipan yang dipelajari untuk setiap jenis instruksi. Teks dalam bentuk bebas direpresentasikan dalam penyandian teks siap pakai dari CLIP. Sedangkan mask disisipkan menggunakan konvolusi dan dijumlahkan elemen demi elemen dengan penyematam gambar (*image embedding*). Dekoder mask secara efektif memetakan penyematam gambar, *prompt* dan token keluaran ke sebuah mask. Desain ini menggunakan modifikasi blok dekode Transformer yang diikuti oleh kepala prediksi mask yang dinamis.



Gambar 3. Hasil deteksi komponen alat pengukur analog oleh YOLOv9 berupa kotak pembatas untuk objek jarum, garis-garis skala dan teks skala.

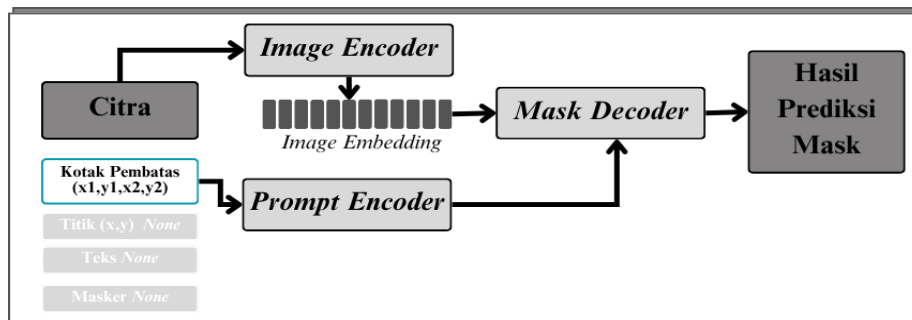
Pada tahapan dalam penelitian ini, metode SAM yang bertanggung jawab melakukan segmentasi semantik untuk mengekstraksi area komponen penting. Segmentasi memisahkan komponen-komponen utama alat pengukur, khususnya jarum dan garis-garis skala, dari latar belakang citra. Berkaitan dengan deteksi oleh YOLOv9 pada tahapan sebelumnya, hasil prediksi kotak pembatas dijadikan masukan dan pedoman bagi SAM, di samping masukan citra yang juga mengalami pengodean. Masukan ini diaplikasikan pada teknik pengodean *prompt*. Tujuannya agar informasi kontekstual dapat dimanfaatkan sehingga fokus pada area yang spesifik dan menghasilkan segmentasi yang lebih akurat, serta mempermudah perhitungan akhir dari alat pengukur. Arsitektur SAM dalam tahapan pembacaan alat pengukur ini dapat dilihat pada Gambar 4. Meskipun SAM dapat menjadikan titik posisi, teks maupun mask sebagai masukan *prompt*, namun dalam penelitian ini cukup dengan kotak pembatas terdeteksi oleh YOLOv9 yang digunakan. Sementara itu, informasi kategori kelas pada YOLOv9 akan tetap disimpan sebagai label karena segmentasi SAM menghasilkan kelas-agnostik yang tidak menyertakan klasifikasi kelas pada hasil prediksi mask [19].

Koordinat kotak pembatas dalam dataset ditentukan dengan memuat kotak pembatas sekecil mungkin (*fitting*) di sekitar mask segmentasi asli untuk tiap kelas pada data acuan. Kotak pembatas ini dilibatkan dalam proses pelatihan sebagai kode *prompt* pada model SAM. Hasil segmentasi dari SAM terdiri atas dua kategori piksel yaitu mask terdeteksi dan latar belakang.

Kumpulan prediksi mask tersebar dan berlabel garis skala dan jarum berdasarkan label kotak pembatas. Perlu diingat bahwa prediksi mask yang didapatkan dari metode SAM tidak menghasilkan identifikasi kelas. Keluaran prediksi komponen oleh YOLOv9 yang dilakukan pada tahapan sebelumnya menjadi acuan untuk dapat mengidentifikasi kelas jarum dan garis skala pada hasil prediksi mask.

Untuk menguji kinerja secara kuantitatif dari model integrasi YOLOv9 dan SAM dalam penelitian ini, maka digunakan perhitungan *Interaction Of Union* (IoU). IoU menunjukkan rasio interseksi dan konkatenasi sampel positif (A) yang diperoleh dari prediksi dan sampel positif berlabel nyata (B). Perhitungan ini dilakukan dengan membandingkan luas area tumpang tindih antara hasil prediksi dan hasil sebenarnya (*Area of Overlap*) dengan luas total dari gabungan hasil prediksi dan hasil sebenarnya (*Area of Union*) seperti dilihat pada Persamaan (1).

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{A \cap B}{A \cup B} \quad (1)$$



Gambar 4. Arsitektur SAM untuk segmentasi komponen alat pengukur.

Lebih lanjut evaluasi dilakukan terhadap nilai presisi (*Precision*) dan sensitivitas (*Recall*), yang diekspresikan pada Persamaan (2) dan (3). Nilai TP (*True Positive*) menunjukkan prediksi benar, yaitu objek atau piksel terdeteksi dengan benar. FP (*False Positive*) merupakan nilai prediksi salah, yaitu objek atau piksel yang tidak ada namun terdeteksi. Sementara FN (*False Negative*) yaitu objek atau piksel ada namun tidak terdeteksi. Presisi mengukur banyaknya prediksi yang benar jika dibandingkan dengan total prediksi positif yang dibuat model. Sedangkan, sensitivitas merupakan perhitungan banyaknya objek yang berhasil dideteksi.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

### 2.3. Dataset

Dataset yang digunakan yaitu citra alat pengukur analog yang diperoleh dari berbagai sumber di internet. Diantaranya dataset publik yang disediakan oleh platform pengembangan model pembelajaran mesin dan kecerdasan buatan seperti *Roboflow* dan *Huggingface*. Dataset ini meliputi berbagai kondisi citra yang mewakili kondisi dunia nyata, seperti distorsi cahaya, posisi alat yang miring, citra alat yang tidak bersih dan kabur. Dataset terdiri dari 1010 gambar alami penampang alat pengukur yang digunakan untuk pemodelan YOLOv9 dan SAM. YOLOv9 memiliki fitur augmentasi yang disediakan pada konfigurasi pelatihannya. Jenis augmentasi yang digunakan dalam rujukan repositori YOLOv9, antara lain *mosaic*, *mix-up*, HSV, serta transformasi geometri seperti *scaling*, *translating*, *flipping* [14].

Anotasi data dilakukan menggunakan perangkat *open-source* Label Studio. Label berupa mask digunakan untuk menandai jarum dan garis-garis skala. Mask juga dimanfaatkan untuk menghasilkan kotak pembatas sebagai label acuan pada proses deteksi komponen dengan YOLOv9.

## 3. HASIL DAN PEMBAHASAN

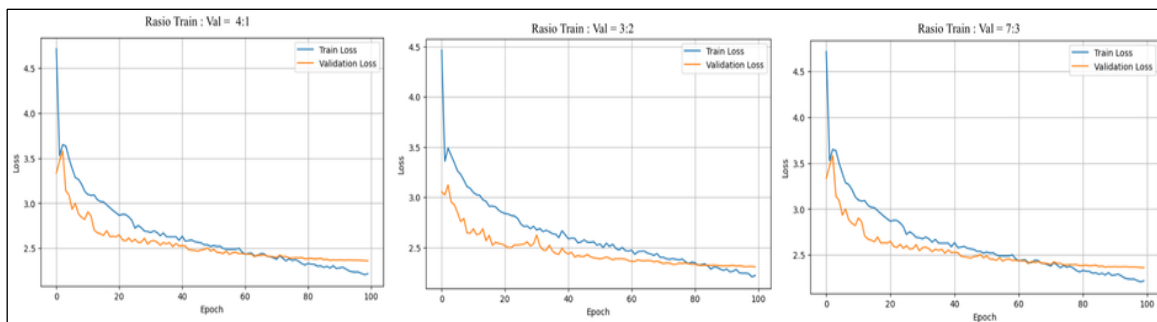
Seluruh eksperimen dilakukan pada lingkungan komputasi berbasis GPU, dengan spesifikasi NVIDIA Tesla T4 GPU dan memori sistem 16GB, Model-model deep learning dilatih menggunakan *framework* Pytorch 2.6.0 dan Python 3.11.

### 3.1 Deteksi Kotak Pembatas

Model ini dilatih dengan dataset citra penampang pengukur analog sebanyak 1010 citra dan diujikan pada beberapa variasi pembagian data latih dan validasi. Hasil evaluasi menunjukkan bahwa model mencapai performa yang sangat baik, dengan nilai mAP75 mencapai 0.897. Jika

dilihat secara spesifik untuk masing-masing komponen diperoleh hasil AP75 yaitu jarum 0.995, garis skala 0.815 dan teks angka skala 0.881. Pada proses pelatihan, Gambar 5 menunjukkan penurunan nilai grafik pada training loss dan validation loss untuk semua variasi rasio data latih dan validasi, antara lain pada rasio train:valid = 4:1, train:valid = 3:2 dan train:valid = 7:3. Hal ini menunjukkan kualitas pelatihan yang sehat dan stabil dengan konvergensi yang tidak overfitted.

Evaluasi lebih lanjut dengan menggunakan confusion matrix yang ditampilkan pada Gambar 6. Berdasarkan data uji yang diberikan, model menunjukkan kemampuan deteksi yang tinggi, dengan tingkat keberhasilan mencapai 100% untuk objek jarum, 94% untuk garis skala dan 98% untuk area angka skala.



Gambar 5. Grafik evaluasi pelatihan YOLOv9 berdasarkan kurva train-validation loss untuk pemodelan deteksi komponen pengukur analog dengan perbandingan beberapa rasio data latih dan data validasi.

		Nilai Prediksi		
		0	1	2
Nilai Benar	0	1.00		
	1		0.94	
	2			0.98
		jarum	garis-skala	angka-skala

Gambar 6. Confusion matrix hasil uji deteksi komponen alat pengukur.s

### 3.2 Evaluasi Segmentasi YOLOv9+Fine-Tuned SAM


























Tahap segmentasi dilakukan dengan memanfaatkan hasil prediksi kotak pembatas untuk komponen jarum penunjuk dan garis-garis skala yang diperoleh dari pada tahapan deteksi sebelumnya. Model yang digunakan adalah model *fine-tuned* SAM yang telah dilatih menggunakan dataset citra penampang alat pengukur. Dataset tersebut dilengkapi dengan mask biner meliputi area jarum dan garis-skala serta informasi kotak pembatas sebagai masukan *prompt* bagi model segmentasi .

Dataset pelatihan terdiri dari 1010 citra, yang dibagi menjadi 808 citra latih dan 202 citra uji. Model pra-latih SAM yang dijadikan titik awal (*checkpoint*) adalah model arsitektur SAM ViT-B dan SAM ViT-L. Selama proses pelatihan, bagian Pengode Citra dan Pengode *Prompt* akan dibekukan untuk mempercepat proses konvergensi model.

Berdasarkan hasil evaluasi, model *fine-tuned* SAM dengan pra latih ViT-B mampu menghasilkan kinerja terbaik dengan nilai IoU, *precision*, dan *recall* masing-masing yaitu 0.81, 0.87, dan 0.92. Pendekatan yang dikembangkan dalam penelitian ini terbukti mampu menghasilkan segmentasi yang lebih terfokus pada area komponen penting, sehingga meningkatkan akurasi dalam ekstraksi fitur.

Model SAM menghasilkan sejumlah mask sesuai jumlah masukan *prompt* kotak pembatas. Ini menunjukkan kemampuan SAM dalam merespon setiap *prompt* dengan segmentasi objek yang berbeda. Jika seluruh masker segmentasi yang dihasilkan tersebut digabungkan, maka dihasilkan hasil akhir yang bersih dan terfokus hanya pada area jarum dan garis skala, serta mengabaikan latar-belakang sepenuhnya.

Tabel 1. Hasil perbandingan model segmentasi jarum dan garis-garis skala menggunakan YOLOv9 Segmentation, SAM, dan Fine-Tuned SAM.

	Citra Asli	Citra Target	YOLOv9+SAM	YOLOv9 Segmentation	YOLOv9 +Fine-Tuned SAM
(a)					
(b)					
(c)					
(d)					
(e)					

Untuk menunjukkan peningkatan performa YOLOv9 berintegrasi dengan *fine-tuned* SAM, dilakukan perbandingan terhadap dua model lain, yaitu YOLOv9 Segmentation dan SAM dasar (tanpa pelatihan ulang) yang mendapatkan masukan *prompt* dari hasil prediksi area oleh model YOLOv9. Secara visual hasil perbandingan model segmentasi citra dalam berbagai kondisi alat pengukur dapat dilihat pada Tabel 1.

Secara kuantitatif, seperti terlihat pada Tabel 2., hasil perbandingan model menunjukkan bahwa MobileNetV2 SAM dasar menunjukkan performa yang lebih rendah dengan *precision* sebesar 0.745, *recall* sebesar 0.802, dan IoU sebesar 0.624. Hasil ini menunjukkan bahwa SAM

pada dasarnya memiliki kemampuan generalisasi yang baik, namun perlu dilakukan fine-tuning agar performa model lebih akurat. YOLOv9 yang memiliki kemampuan dalam tugas spesifik segmentasi juga menunjukkan hasil evaluasi yang baik dengan nilai IoU mencapai 0.738, *precision* sebesar 0.785 dan *recall* sebesar 0.923. Namun, model Fine-Tuned SAM menghasilkan kinerja terbaik dengan *precision* sebesar 0.81, *recall* sebesar 0.946, dan IoU mencapai 0.778. Hal ini membuktikan bahwa proses *fine-tuning* berhasil meningkatkan ketepatan dan konsistensi segmentasi.

Tabel 2. Hasil perbandingan evaluasi model segmentasi berdasarkan Precision, Recall, dan IOU.

Model	Precision	Recall	IOU
YOLOv9 Segmentation	0.785	0.923	0.738
YOLOv9 + SAM	0.745	0.802	0.624
<b>YOLOv9 + Fine-Tuned SAM</b>	<b>0.81</b>	<b>0.946</b>	<b>0.778</b>

Jarum penunjuk merupakan komponen kritis dalam pembacaan nilai alat ukur. Segmentasi yang akurat terhadap jarum berperan penting dalam menjamin keakuratan pembacaan nilai skala. Seperti terlihat pada Tabel 3, hasil eksperimen menunjukkan bahwa pendekatan YOLOv9+ Fine-Tuned SAM mampu mencapai 0.837 segmentasi jarum IoU. Nilai ini lebih tinggi dibandingkan hasil segmentasi jarum pada penelitian yang pernah dilakukan sebelumnya yang mencapai 0.568 [17]. Selain itu, pada segmentasi gabungan antara jarum penunjuk dan garis-garis skala, model YOLOv9 + Fine-Tuned SAM juga menunjukkan kinerja yang lebih baik dengan nilai IoU yang mencapai 0.778 jika dibandingkan metode CBAM + U2NET [2] yang menghasilkan nilai IoU sebesar 0.71. Hasil ini menunjukkan pendekatan pada penelitian ini lebih stabil dan adaptif terhadap variasi bentuk serta kondisi citra alat pengukur analog.

Tabel 3. Hasil perbandingan metrik IoU pada segmentasi jarum dan garis-garis skala.

Model	IoU (Jarum Penunjuk)	IoU (Jarum Penunjuk + Garis-garis skala)
YOLOv8 [17]	0.568	-
CBAM + U2NET [5]	-	0.71
<b>sYOLOv9 + Fine-Tuned SAM</b>	<b>0.837</b>	<b>0.778</b>

#### 4. KESIMPULAN

Penelitian ini berhasil mengembangkan kerangka kerja bertahap pembelajaran mendalam untuk pengenalan komponen alat pengukur analog dengan mengintegrasikan YOLOv9 sebagai modul deteksi dan *fine-tuned* SAM sebagai modul segmentasi. Integrasi keduanya menghasilkan sistem yang mampu mengenali area penting seperti jarum dan garis-garis skala dengan tingkat presisi dan akuisisi tinggi. Berdasarkan hasil pengujian, model YOLOv9+Fine-Tuned SAM ViT B mencapai nilai IoU sebesar 0.778 untuk segmentasi keseluruhan dan 0.837 untuk segmentasi jarum, menunjukkan peningkatan signifikan dibandingkan penelitian sebelumnya berbasis YOLOv8.

Temuan ini menunjukkan bahwa *fine-tuning* SAM pada domain citra alat pengukur dapat meningkatkan kemampuan model dalam memahami struktur visual spesifik, terutama ketika diarahkan oleh kotak pembatas hasil deteksi YOLOv9. Dengan demikian, pendekatan ini tidak hanya efektif dalam mendeteksi dan mengekstraksi komponen penting, tetapi juga potensial untuk diimplementasikan dalam sistem pembacaan alat pengukur analog secara otomatis di lingkungan industri.

Sebagai arah pengembangan selanjutnya, penelitian dapat difokuskan pada integrasi modul pembacaan numerik berbasis OCR. Pendekatan OCR berbasis transformer, seperti TrOCR atau Donut, berpotensi memberikan akurasi tinggi dalam mengenali angka skala di permukaan dial pada berbagai kondisi. Sistem juga dapat dikembangkan untuk meningkatkan *robustness* terhadap gangguan visual dengan teknik pemrosesan citra atau eksplorasi augmentasi. Selanjutnya, optimalisasi model dapat dilakukan untuk memungkinkan diterapkan secara langsung di lapangan pada perangkat *edge computing* dalam komputasi terbatas.

**REFERENSI**

- [1] H. Ninama *et al.*, “Computer vision and deep transfer learning for automatic gauge reading detection,” *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-71270-0.
- [2] G. Salomon, R. Laroca, and D. Menotti, “Image-based Automatic Dial Meter Reading in Unconstrained Scenarios,” *Measurement (Lond)*, vol. 204, Nov. 2022, doi: 10.1016/j.measurement.2022.112025.
- [3] Vue.ai, “The Hidden Costs of Manual Meter Reading: Why Your Utility Company Can’t Afford to Wait for Automation,” 2025. Accessed: Nov. 25, 2025. [Online]. Available: <https://www.vue.ai/blog/ai-transformation/hidden-costs-manual-meter-reading/>.
- [4] J. Peixoto *et al.*, “Development of an Analog Gauge Reading Solution Based on Computer Vision and Deep Learning for an IoT Application,” *Telecom*, vol. 3, no. 4, pp. 564–580, Dec. 2022, doi: 10.3390/telecom3040032.
- [5] H. Fan and Y. Li, “Image Recognition and Reading of Single Pointer Meter Based on Deep Learning,” *IEEE Sens J*, vol. 24, no. 15, pp. 25163–25174, 2024, doi: 10.1109/JSEN.2024.3416436.
- [6] Y. Shu, S. Liu, H. Xu, and F. Jiang, “Read Pointer Meters in complex environments based on a Human-like Alignment and Recognition Algorithm,” Feb. 2023, doi: 10.1007/978-981-99-8761-0\_13.
- [7] P. Ni and P. Mao, “Reading recognition method of pointer instrument based on YOLOv8+U2-Net,” in *2024 6th International Conference on Internet of Things, Automation and Artificial Intelligence, IoTAAI 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 664–668. doi: 10.1109/IoTAAI62601.2024.10692585.
- [8] J. Leon-Alcazar, Y. Alnumay, C. Zheng, H. Trigui, S. Patel, and B. Ghanem, “Learning to Read Analog Gauges from Synthetic Data,” in *Proceedings - 2024 IEEE Winter Conference on Applications of Computer Vision, WACV 2024*, Institute of Electrical and Electronics Engineers Inc., Jan. 2024, pp. 8601–8610. doi: 10.1109/WACV57701.2024.00842.
- [9] A. Kirillov *et al.*, “Segment Anything,” 2023. [Online]. Available: <https://segment-anything.com>.
- [10] L. Ali, F. Alnajjar, M. Swavaf, O. Elharrouss, A. Abd-alrazaq, and R. Damseh, “Evaluating segment anything model (SAM) on MRI scans of brain tumors,” *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-72342-x.
- [11] E. Khalili, B. Priego-Torres, A. Leon-Jimenez, and D. Sanchez-Morillo, “Automatic Lung Segmentation in Chest X-Ray Images Using SAM with Prompts from YOLO,” *IEEE Access*, vol. 12, pp. 122805–122819, 2024, doi: 10.1109/ACCESS.2024.3454188.
- [12] M. Azimi and T. Y. Yang, “Transformer-based framework for accurate segmentation of high-resolution images in structural health monitoring,” *Computer-Aided Civil and Infrastructure Engineering*, Dec. 2024, doi: 10.1111/mice.13211.
- [13] A. Moghimi, M. Welzel, T. Celik, and T. Schlurmann, “A Comparative Performance Analysis of Popular Deep Learning Models and Segment Anything Model (SAM) for River Water Segmentation in Close-Range Remote Sensing Imagery,” *IEEE Access*, vol. 12, pp. 52067–52085, 2024, doi: 10.1109/ACCESS.2024.3385425.
- [14] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, “YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information,” *Computer Vision -- ECCV 2024*, vol. 15089, pp. 1–21, Feb. 2024, doi: [https://doi.org/10.1007/978-3-031-72751-1\\_1](https://doi.org/10.1007/978-3-031-72751-1_1).
- [15] A. Imran, M. Shashishekhara, H. Hamza, and A. A. Gardi, “Real Time American Sign Language Detection Using Yolo-v9,” Jul. 2024. doi: 10.48550/arXiv.2407.17950.
- [16] M. F. Tariq and M. A. Javed, “Small Object Detection with YOLO: A Performance Analysis Across Model Versions and Hardware,” Apr. 2025, [Online]. Available: <http://arxiv.org/abs/2504.09900>

- [17] M. Reitsma, J. Keller, K. Blomqvist, and R. Siegwart, “Under pressure: Learning-based analog gauge reading in the wild,” in *Proceedings - IEEE International Conference on Robotics and Automation*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 14–20. doi: 10.1109/ICRA57147.2024.10610793.
- [18] C. H. Wang, K. K. Huang, R. I. Chang, and C. K. Huang, “Scale-Mark-Based Gauge Reading for Gauge Sensors in Real Environments with Light and Perspective Distortions,” *Sensors*, vol. 22, no. 19, Oct. 2022, doi: 10.3390/s22197490.
- [19] X. Jin, J. Hu, J. Lin, S. Zhang, and L. Cao, “U-SAM: Upgrade Segment Anything Model With Semantic-Aware and Memory-Efficient,” *Institute of Electrical and Electronics Engineers (IEEE)*, Mar. 2025, pp. 1–5. doi: 10.1109/icassp49660.2025.10889270.