

Implementasi YOLOv11 pada Video Ekokardiografi untuk Deteksi Ventricular Septal Defect pada Anak

Dwi Amalia Putri¹, Amri Muhaimin², Muhammad Nasrudin³

^{1,2,3}Sains Data, Universitas Pembangunan Nasional Veteran Jawa Timur

²amri.muhaimin.stat@upnjatim.ac.id

³nasrudin.fasilkom@upnjatim.ac.id

Corresponding author email: 22083010042@student.upnjatim.ac.id

ABSTRAK

Penyakit Jantung Bawaan (PJB) merupakan kelainan yang umum terjadi pada bayi. Di Indonesia, kasus PJB setiap tahunnya diperkirakan mencapai sekitar 50.000 kasus, namun hanya sekitar 7.500 kasus yang mendapatkan layanan intervensi medis. Salah satu jenis PJB yang paling sering dijumpai pada anak adalah *Ventricular Septal Defect* (VSD), yaitu kelainan dengan lubang pada dinding pemisah antara bilik jantung. Kondisi ini dapat menyebabkan aliran darah tidak normal dan berdampak pada fungsi jantung. Penelitian ini bertujuan untuk memprediksi letak kebocoran pada bilik jantung berdasarkan video ekokardiografi jantung. Hasil deteksi tersebut diharapkan dapat membantu deteksi dini VSD yang menjadi rujukan bagi tenaga medis dalam penanganan pasien. Untuk mendukung proses diagnosis yang efisien dan akurat, penelitian ini menerapkan model *transfer learning* dengan algoritma YOLOv11 yang dilatih ulang menggunakan data *video ultrasound* dengan tampilan *apical four chamber*. YOLO dipilih karena keunggulannya dalam mendeteksi objek secara cepat dan presisi. Proses pembangunan model dilakukan melalui beberapa tahapan, yaitu ekstraksi frame dari *video ultrasound*, pelabelan data, augmentasi, pembuatan model deteksi, hingga pengujian model. Hasil pengujian menunjukkan bahwa model YOLOv11 memiliki nilai mAP yang cukup baik dan mampu mendeteksi area defek secara tepat. Temuan ini menunjukkan bahwa metode YOLOv11 berpotensi mempercepat proses diagnosis serta meningkatkan akurasi deteksi VSD pada anak.

Keywords: Deteksi Objek, Ekokardiografi, *Ventricular Septal Defect*, YOLOv11

I. PENDAHULUAN

Penyakit Jantung Bawaan (PJB) adalah jenis kelainan bawaan yang paling sering terjadi [1]. Kelainan pada struktur dan fungsi jantung muncul akibat gangguan perkembangan jantung dan pembuluh darah besar pada fase pertumbuhan janin dalam kandungan [2]. Di Indonesia layanan pemeriksaan dan penanganan PJB masih sangat rendah. IDAI menyebut, hingga tahun 2024, dari perkiraan kasus PJB setiap tahunnya yang mencapai 50.000 kasus, hanya sekitar 7.500 kasus yang mendapatkan layanan intervensi medis [3]. Salah satu jenis PJB yang paling umum terjadi utamanya pada anak yaitu *Ventricular Septal Defect* (VSD) [4]. VSD ditandai dengan adanya lubang pada sekat ventrikel kiri dan kanan [5]. Kondisi ini dapat menyebabkan aliran darah tidak normal dan berdampak pada fungsi jantung. Pasien dengan VSD dapat mengalami gangguan pertumbuhan, komplikasi, hingga kematian [4]. Untuk meminimalisir resiko, perlu identifikasi lebih dini agar VSD pada anak dapat ditangani lebih cepat.

VSD dapat didiagnosis melalui pemeriksaan non invasif, ekokardiografi. Dengan tampilan apikal, setiap jantung termasuk VSD dapat terlihat. Namun terdapat hambatan dalam interpretasi hasil citra *ultrasound* seperti kualitas hasil yang kurang baik, interpretasi membutuhkan keterampilan profesional yang rentan penafsiran subjektif, dan struktur septum defek yang tampak tidak jelas oleh mata telanjang [6]. Metode konvensional tidak cukup sebagai metode yang cepat dan akurat

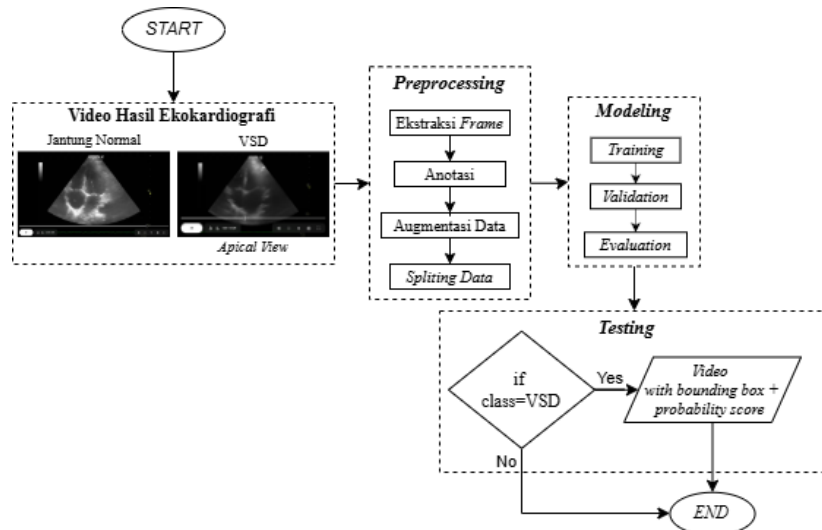
mengidentifikasi posisi defek septum ventrikel pada anak. Berdasarkan hal tersebut, diperlukan sistem deteksi otomatis VSD yang dapat menangkap informasi yang tidak mudah terlihat dalam hasil ekokardiografi.

Deteksi objek dapat digunakan sebagai metode efektif untuk mendeteksi VSD pada citra ekokardiografi karena memungkinkan sistem mengidentifikasi dan melokalisasi area abnormal seperti defek septum sebagai objek dengan karakteristik visual khusus [7], [8], [9]. Deteksi objek menggabungkan klasifikasi dan lokalisasi objek sehingga model tidak hanya menentukan bahwa terkategori VSD tetapi juga menunjukkan posisi area defek pada gambar. Beberapa penelitian sebelumnya telah mengkaji metode deteksi otomatis untuk kasus VSD. Chen dkk [8] mendeteksi 3 tipe VSD pada data ekokardiografi color doppler dengan kombinasi algoritma YOLOv4 dan DenseNet, diperoleh evaluasi mAP-50 sebesar 72.61%. Dalam artikel berbeda, Chen dkk [9] mendeteksi 3 tipe VSD dan ASD dengan membandingkan algoritma YOLOv5, YOLOv7, and YOLOv9. Penelitian tersebut juga mengombinasikan teknik augmentasi data dan diperoleh evaluasi yang lebih baik pada YOLO versi 9 dengan evaluasi mAP lebih dari 94% di setiap perlakuan. Yang dkk[10] mendiagnosis PJB dari 5 tampilan citra ultrasound jantung normal dan abnormal dengan beberapa jenis YOLOv5 dan didapatkan model terbaik YOLOv5l dengan akurasi mencapai 92,59%. Penelitian tersebut menyebut jika YOLOv5 lebih unggul daripada FastRCNN-ResNet50 dan FastRCNN-MobileNet50. Terdapat pula penelitian lain seperti deteksi VSD pada anak dengan YOLOv4 melalui citra ultrasound tampilan *apical four chamber* [11] dan deteksi VSD pada anak dengan Faster-RCNN melalui citra ultrasound tampilan *apical four chamber* dan parasternal [12]. Penelitian tersebut menjadi dasar penelitian penulis dalam membangun model deteksi VSD dengan YOLO versi 11. YOLOv11 juga sudah banyak digunakan pada data medis seperti pada penelitian oleh Chourib [13], yang mengembangkan model deteksi dan segmentasi tumor otak melalui transfer learning YOLOv11 dan *postprocessing*, didapatkan hasil mAP yang tinggi. Penelitian oleh Hsu dkk [14], yang mengembangkan model deteksi otomatis untuk diagnosis DDH menggunakan YOLOv11 yang dimodifikasi dan menghasilkan performa yang tinggi dan cepat.

Pendekatan *deep learning* yang mengandalkan *transfer learning* dapat memungkinkan diagnosis penyakit jantung struktural meskipun data pelatihan terbatas. Algoritma *You Only Look Once* (YOLO) terkenal sebagai algoritma detektor objek dengan satu kali proses pemindaian (*single forward pass*) pada gambar atau video [15], sehingga cocok digunakan untuk deteksi kelainan struktural jantung dalam citra medis secara real time. komponen baru pada YOLOv11 membuat algoritma tersebut mampu memberikan akurasi lebih baik pada objek kecil sekaligus dengan latensi yang lebih rendah. Penelitian ini berupaya mengoptimalkan *object detection* berbasis YOLO versi 11 dengan teknik augmentasi data untuk mendeteksi VSD secara lebih spesifik, cepat, dan akurat. Temuan yang dihasilkan memberikan wawasan berharga bagi tenaga medis dalam mendiagnosis VSD.

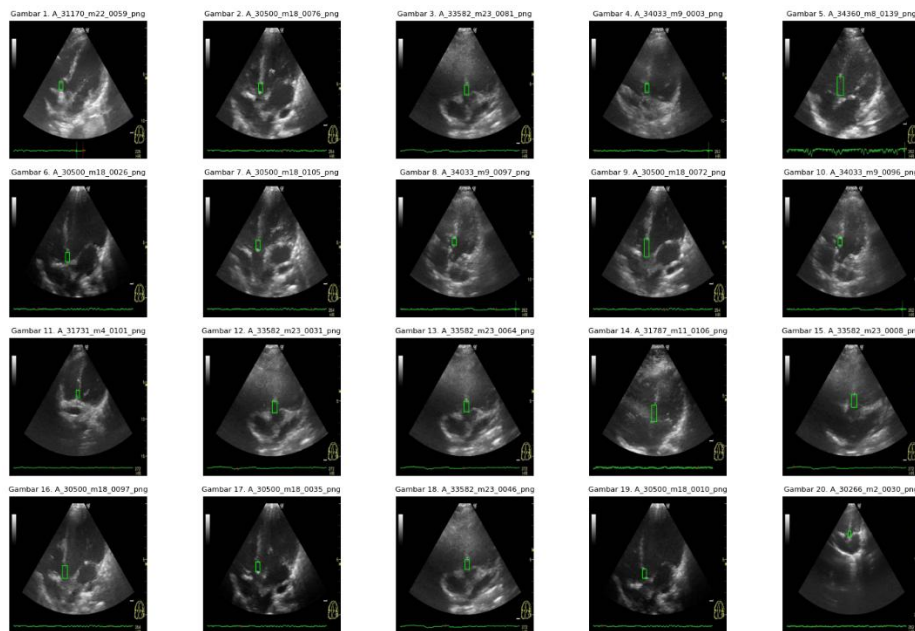
II. METODOLOGI PENELITIAN

Penelitian ini disusun dengan metodologi yang dibagi menjadi 4 tahap utama seperti pada gambar 1. Data yang digunakan merupakan data video jantung pasien normal dan yang terdiagnosis VSD. Data tersebut didapatkan dari hasil pemeriksaan ekokardiografi dengan tampilan *apical four chamber* yang dihimpun dari suatu rumah sakit. Data terdiri atas 24 video yang meliputi 23 video VSD dan 1 video normal. Data diolah menggunakan bahasa pemrograman python.



Gambar 1. Flowchart Penelitian

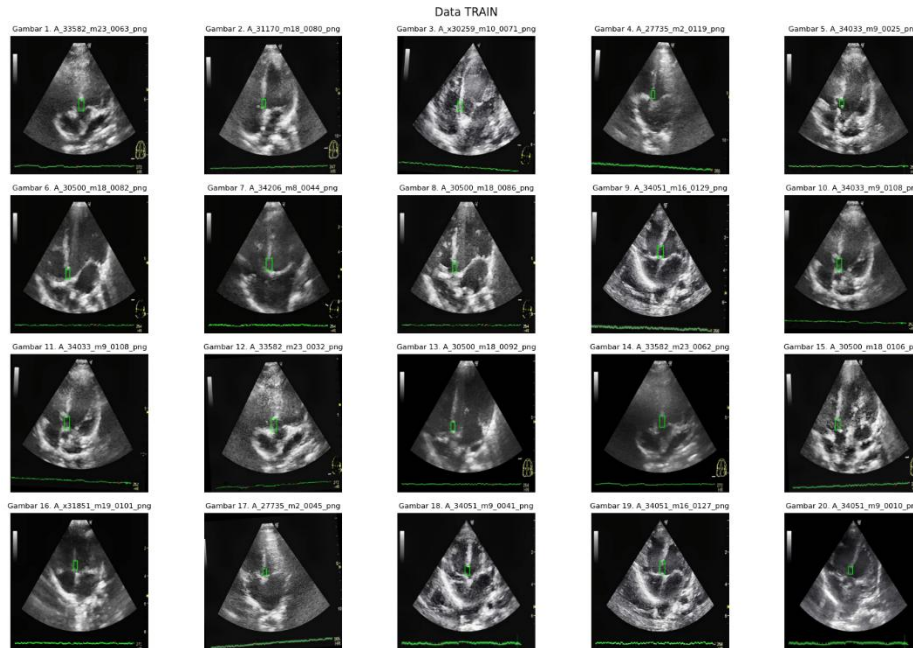
Setelah tahap akuisisi data, tahap selanjutnya yaitu preprocessing data. Untuk memudahkan proses selanjutnya, video *ultrasound* diubah menjadi gambar melalui ekstraksi frame. Banyaknya frame yang dihasilkan bergantung pada durasi video dan kecepatan video (*frame per second*). Keseluruhan gambar dalam video, diekstrak menggunakan pustaka subprocess dari *ffmpeg* dan selanjutnya dipilih hanya gambar yang menunjukkan adanya defek atau celah di sekat ventrikel untuk tahap labeling. Dalam tugas deteksi objek, pelabelan dilakukan dengan memberikan bounding box di area yang menunjukkan adanya objek. Objek yang dimaksud dalam penelitian ini yaitu defek pada sekat ventrikel. Tahap anotasi data dilakukan dengan menggunakan *tools* *labelme*. Hasil anotasi disimpan dalam format *.json* yang kemudian diubah ke format YOLO dataset dengan *python*.



Gambar 2. Hasil Anotasi

Jumlah gambar hasil anotasi ditingkatkan melalui tahap augmentasi. Selain itu augmentasi digunakan untuk meningkatkan keragaman data dalam melatih model sehingga model tidak hanya baik di data latih namun juga baik di data baru [9]. Dalam augmentasi data, penelitian menggunakan beberapa metode yaitu *image enhancement* dengan SHARPEN, kontras adaptif dengan CLAHE, *noise reduction* dengan Median Blur, dan transformasi geometri meliputi perubahan posisi, ukuran, rotasi dengan ShiftScaleRotate. Berbagai teknik augmentasi tersebut, penulis implementasikan dengan

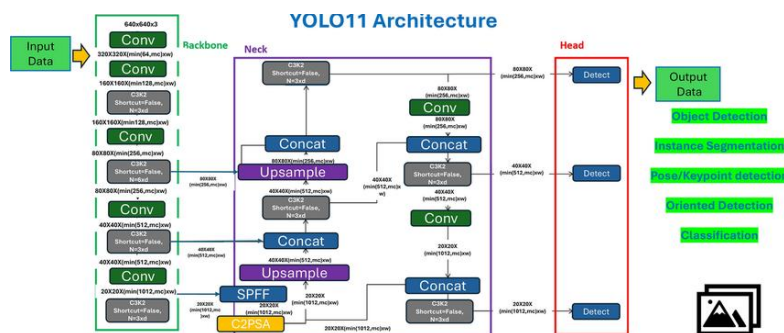
library Albumentations. Dari tahap augmentasi data tersebut, didapatkan hasil seperti pada gambar 2.



Gambar 3. Hasil Augmentasi Data

Pembagian data untuk train, validation, dan test, dilakukan berdasarkan video hasil ekokardiografi. Data train menggunakan 23 video VSD. Data validation menggunakan xx video VSD. Data test untuk mengukur seberapa baik model memprediksi data baru, menggunakan 1 video VSD dan 1 video normal.

Setelah beberapa tahapan persiapan data, data gambar yang sudah disusun dalam format YOLO dataset digunakan untuk membangun model prediktif deteksi defek ventrikel. Model dibangun dengan metode YOLO versi 11. *You Only Look Once* (YOLO) adalah algoritma deteksi objek berbasis jaringan konvolusional (CNN) yang bekerja secara *single-stage* dengan memformulasikan tugas deteksi objek sebagai masalah regresi langsung dari citra piksel ke koordinat *bounding box* dan probabilitas kelas melalui satu jaringan saraf konvolusional yang bersifat *end-to-end*. Arsitektur YOLOv11 merupakan pengembangan terbaru dari keluarga YOLO yang dirilis oleh Ultralytics sebagai penerus YOLOv8, dengan peningkatan signifikan pada efisiensi komputasi, kemampuan generalisasi, dan akurasi multi-task.



Gambar 4. Arsitektur YOLOv11 [16]

Berdasarkan gambar 4, secara umum YOLOv11 terdiri dari tiga komponen utama, yaitu Backbone, Neck, dan Head, yang bekerja secara *end to end* untuk menghasilkan prediksi deteksi objek maupun tugas lainnya. Pada bagian Backbone, YOLOv11 menggunakan struktur bertingkat berbasis *convolutional blocks* dengan mekanisme *shortcut* untuk mempertahankan aliran gradien dan mencegah *vanishing gradient*. Setiap blok konvolusi dieksekusi dengan kernel 3×3 atau 1×1, kemudian diikuti *batch*

normalization dan aktivasi non-linear. *Backbone* tersebut menghasilkan peta fitur multi-skala, mulai dari resolusi tinggi hingga rendah, sehingga dapat menangkap baik detail objek kecil maupun struktur objek besar. Dibandingkan YOLOv8, YOLOv11 memperkenalkan modul C2f dan C2PSA, yaitu varian dari *Cross Stage Partial* (CSP) dengan mekanisme partial self-attention untuk meningkatkan representasi fitur tanpa menambah beban komputasi secara signifikan [17]. C2PS terdiri dari *Partial Self-Attention* (PSA) yang memungkinkan model menangkap hubungan kontekstual antarpiksel pada area penting secara lebih selektif dan efisien dibanding *self attention* penuh.

Pada bagian *Neck*, YOLOv11 mengadopsi struktur *Feature Pyramid Network* (FPN) yang digabungkan dengan *Path Aggregation Network* (PAN). Kombinasi tersebut bertujuan meningkatkan aliran informasi *top-down* dan *bottom-up* sehingga model dapat menggabungkan fitur resolusi rendah (semantik kuat) dengan fitur resolusi tinggi (detail visual). YOLOv11 juga menyertakan modul SPPF (*Spatial Pyramid Pooling-Fast*), yang berfungsi mengompresi konteks global melalui *pooling* bertingkat ukuran besar tanpa menurunkan performa kecepatan. Fitur-fitur dari berbagai level tersebut kemudian disatukan melalui operasi *concat* dan *upsampling* untuk memungkinkan deteksi objek pada berbagai skala. Bagian *Head* YOLOv11 menggunakan mekanisme decoupled head, di mana prediksi kelas dan prediksi bounding box dilakukan melalui jalur yang terpisah. Pendekatan tersebut telah terbukti mempercepat konvergensi dan meningkatkan akurasi deteksi, karena model dapat mengoptimalkan kedua proses secara independen. *Head* menghasilkan output pada tiga tingkat resolusi, sehingga mendukung deteksi objek kecil, sedang, dan besar. Selain itu, desain YOLOv11 telah diperluas menjadi multi-task head yang memungkinkan output untuk *object detection*, *instance segmentation*, *pose estimation*, *oriented object detection*, dan *image classification*, tanpa perlu mengubah struktur backbone secara signifikan [17].

Secara keseluruhan, YOLOv11 mengintegrasikan peningkatan efisiensi struktural, mekanisme *attention* parsial, optimasi arsitektur feature pyramid, dan decoupled head untuk mencapai performa yang lebih stabil dan akurat dibanding generasi sebelumnya. Penulis menggunakan pustaka ultralytics di python untuk membangun model YOLOv11. Dari proses pelatihan dan validasi, model diukur performanya menggunakan metrik evaluasi precision, recall, dan mAP. Metrik presisi mengukur seberapa akurat model memprediksi positif dengan benar. Metrik *recall* mengukur seberapa baik model dapat menemukan semua positif aktual. Dalam diagnosis penyakit, sangat penting meminimalkan kesalahan prediksi terdiagnosis suatu penyakit atau *False Negative* (FN) sehingga metrik recall harus maksimum.

$$P = \frac{\sum TP}{\sum TP + \sum FP} \quad (1)$$

$$R = \frac{\sum TP}{\sum TP + \sum FN} \quad (2)$$

Average Precision (AP) adalah rata-rata precision pada berbagai tingkat *recall* (dari 0 ke 1) yang digunakan untuk mengukur seberapa baik model mempertahankan precision-nya ketika recall meningkat. AP merupakan luas area di bawah kurva *Precision-Recall*. Sehingga mAP dapat ditentukan dengan rumus

$$mAP = \frac{1}{C} \sum_{c=1}^C AP_c \quad (2.12)$$

dengan P : *Precision*, R : *Recall*, C : Total kelas, AP_c : *Average precision* kelas ke- c

III. HASIL DAN ANALISIS

Berdasarkan uraian tahapan penelitian pada subbab sebelumnya, setelah membangun model deteksi VSD dengan YOLOv11, penulis menguraikan hasil pemodelan dalam dua poin besar, sebagai berikut

III.1 Implementasi Model YOLOv11

Model YOLOv11 disusun dengan konfigurasi parameter antara lain, skala ukuran model diatur menjadi nano (n) karena meninjau ukuran data pelatihan, *image size* sebesar 640 yaitu gambar diseragamkan ukurannya menjadi 640x640, z diatur sebanyak

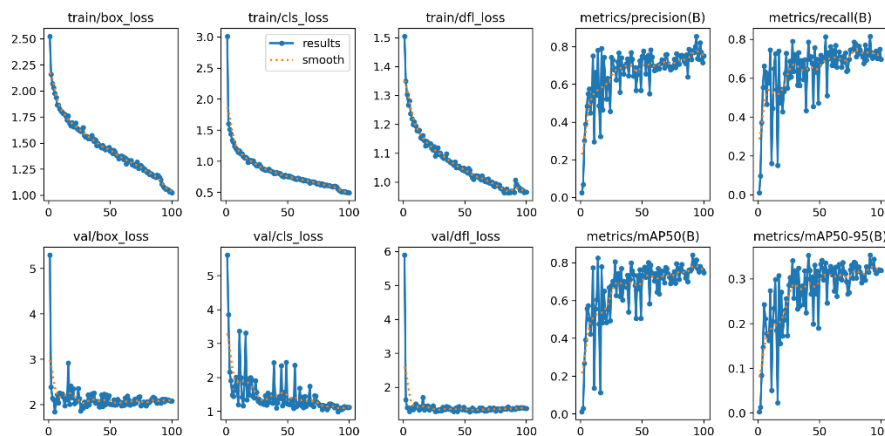
100 yang berarti model belajar dari data sebanyak 100 kali dengan dipantau oleh *patience*. Parameter *patience* bertugas memonitor *validation loss* atau *validation mAP*, disetting sebesar 60 yaitu dalam 60 epoch berturut-turut validasi tidak membaik, training akan dihentikan lebih awal. Parameter selanjutnya yaitu *batchsize* sebesar 16 yang berarti model akan diupdate bobotnya setelah mempelajari 16 data, *batchsize* yang kecil mempengaruhi gradien yang lebih fluktuatif. Parameter *optimizer* seperti AdamW bekerja mengatur bagaimana bobot diperbarui di setiap *batch* dan *epoch*, sehingga kualitas pembaruan bobot sangat dipengaruhi oleh *batch size* dan kompleksitas dataset. Dengan parameter tersebut model deteksi disusun untuk setiap kondisi. Penelitian ini membandingkan performa model YOLOv11 pada dua kondisi preprocessing data. Pada skenario pertama, model dilatih menggunakan data asli tanpa modifikasi (tanpa augmentasi). Pada skenario kedua, model dilatih menggunakan data yang dengan tambahan hasil augmentasi.

Tabel 1. Perbandingan kinerja model YOLOv11 pada Data *Validation*

<i>Preprocessing</i>	Metrik Evaluasi		
	<i>Precision</i>	<i>Recall</i>	mAP-50
tanpa Augmentasi	0.662	0.736	0.703
dengan Augmentasi	0.781	0.816	0.841

Berdasarkan tabel tersebut, terlihat bahwa penerapan augmentasi data menghasilkan peningkatan performa pada seluruh metrik evaluasi. Peningkatan tersebut menunjukkan bahwa augmentasi data sangat membantu dalam meningkatkan kemampuan model untuk menangkap fitur-fitur halus yang penting pada citra dalam mendiagnosis kelainan VSD.

Merujuk pada tabel 1, secara umum performa model YOLOv11 cukup baik dalam mendeteksi celah di septum ventrikel. Nilai *precision* sebesar 0.781 menunjukkan bahwa dari seluruh prediksi positif yang dihasilkan model, sekitar 78.1% merupakan deteksi yang benar, sedangkan sisanya masih merupakan kesalahan prediksi atau *false positive*. Hal tersebut menggambarkan bahwa model sudah cukup baik dalam menghindari deteksi berlebih pada objek yang sebenarnya tidak termasuk kategori VSD. Sementara itu, nilai *recall* sebesar 0.816 mengindikasikan bahwa model mampu mendeteksi sekitar 81.6% dari seluruh objek VSD yang benar-benar ada dalam data uji, namun masih terdapat sekitar 18.4% objek yang tidak terprediksi (*false negative*). Secara keseluruhan, nilai mAP sebesar 0.841 menunjukkan performa model yang baik dalam menggabungkan aspek ketepatan dan kelengkapan deteksi pada berbagai ambang batas IoU. Nilai mAP yang cukup tinggi menegaskan bahwa model mampu memberikan prediksi yang stabil dan konsisten, sehingga layak digunakan sebagai sistem deteksi VSD, meskipun masih memungkinkan dilakukan peningkatan untuk mengurangi kesalahan prediksi dan meningkatkan akurasi deteksi.



Gambar 5. Subplot Evaluasi Hasil Model YOLOv11

Diagram pada gambar di atas menunjukkan perkembangan berbagai komponen loss serta metrik evaluasi selama proses pelatihan model YOLO hingga 100 epoch. Pada bagian atas terlihat kurva *training loss*, yang terdiri dari *box_loss*, *cls_loss*, dan *dfl_loss*. Ketiganya menunjukkan penurunan yang konsisten seiring bertambahnya epoch, menandakan bahwa proses pembelajaran berlangsung stabil dan efektif. *Box_loss* secara bertahap menurun dari nilai awal sekitar 2.3 menjadi mendekati 1.2, menunjukkan peningkatan dalam akurasi regresi bounding box. Sementara itu, *cls_loss* juga mengalami penurunan signifikan dari sekitar 2.6 menjadi kurang dari 1.0, yang berarti kemampuan model dalam mengklasifikasikan objek semakin baik. *dfl_loss* sebagai komponen penting untuk estimasi distribusi batas objek juga menurun halus, mencerminkan peningkatan presisi pada pemodelan lokasi objek.

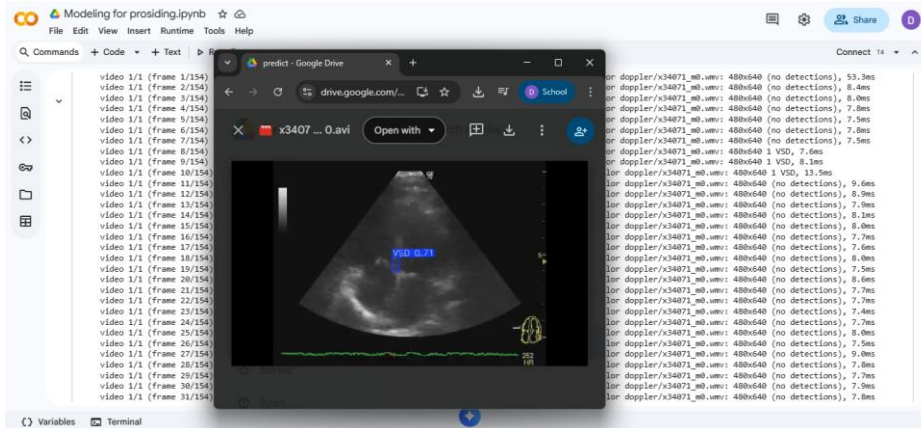
Bagian bawah gambar memperlihatkan *validation loss*, yang mengikuti pola serupa dengan *training loss* namun dengan sedikit fluktuasi. *val/box_loss* dan *val/cls_loss* mengalami variabilitas di beberapa titik, tetapi tetap menunjukkan tren menurun yang stabil, menandakan bahwa model dapat melakukan generalisasi dengan baik. *val/dfl_loss* juga tampak mengalami penurunan meski terdapat beberapa lonjakan, namun pada akhirnya stabil di nilai yang lebih rendah.

Pada bagian kanan, grafik metrik evaluasi precision, recall, mAP50, dan mAP50-95, yang menunjukkan tren peningkatan seiring waktu. Precision dan recall meningkat tajam pada awal pelatihan dan kemudian berfluktuasi di rentang nilai tinggi, menandakan bahwa model semakin akurat dalam mendeteksi objek dan mengurangi kesalahan deteksi. mAP50 meningkat secara konsisten dan mencapai performa yang stabil mendekati nilai optimal setelah sekitar 40 epoch. Hal serupa terlihat pada mAP50-95, yang menggambarkan peningkatan kemampuan model dalam berbagai tingkat ketelitian IoU. Secara keseluruhan, kurva-kurva tersebut menunjukkan bahwa model mengalami proses pembelajaran yang efektif, dengan penurunan loss yang konsisten pada fase training dan validation serta peningkatan metrik evaluasi tanpa indikasi kuat adanya overfitting. Model berhasil meningkatkan akurasi deteksi objek secara stabil sepanjang proses pelatihan.

Jika ditinjau dari waktunya, pelatihan model memerlukan waktu 1,341 jam, selain itu proses validasi model menunjukkan bahwa waktu komputasi untuk setiap gambar sangat efisien, dengan rata-rata waktu 0.5 ms untuk *preprocessing*, 3.2 ms untuk *inference*, dan 3 ms untuk *postprocessing*. Total waktu pemrosesan satu gambar hanya sekitar 6.7 ms, yang berarti model mampu melakukan deteksi secara real-time dengan kecepatan lebih dari 149 gambar per detik (FPS). Dengan demikian, baik proses *training* maupun inferensi berjalan sangat cepat, menandakan efisiensi tinggi dan kecocokan model YOLO11n untuk aplikasi deteksi cepat seperti deteksi VSD pada alur video atau *streaming*.

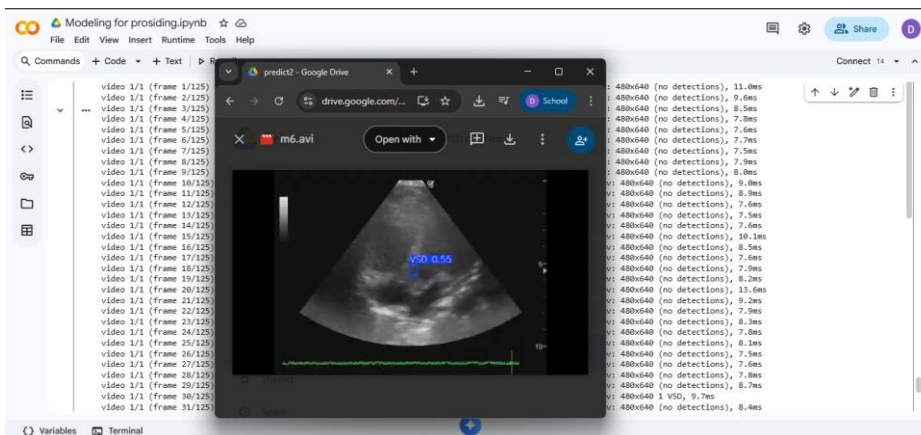
III.2 Testing Deteksi VSD

Model hasil training disimpan dalam format pt, yang selanjutnya dapat diakses untuk memprediksi lokasi VSD dari data gambar atau video hasil ekokardiografi. Pengujian dilakukan dengan 2 video, normal dan VSD dan diperoleh hasil sebagai berikut.



Gambar 6. Testing Video Pasien VSD

Model dapat mendeteksi lokasi VSD yang ditunjukkan dengan adanya bounding box biru pada video. Hasil prediksi juga dilengkapi dengan confidence score yang menggambarkan keyakinan model bahwa ada objek (*objectness*) di dalam bounding box tersebut dan ketepatan ukuran serta posisi defek. Jika ditinjau secara keseluruhan, sebagian besar VSD dalam video berhasil terprediksi.



Gambar 7. Testing dengan Video Pasien Normal

Dengan inputan video pasien normal yang seharusnya tidak ada area defek, model sensitif untuk memprediksi area yang bukan defek sebagai VSD. Kesalahan prediksi berupa deteksi VSD pada citra jantung normal menunjukkan adanya *false positive* yang konsisten. Bias model terhadap kelas positif (VSD) kemungkinan dapat terjadi karena beberapa faktor seperti ketidakseimbangan data, pelabelan yang tidak konsisten, dan sensitifitas model terhadap noise pada citra ekokardiografi. Tingginya angka false positive akan menurunkan nilai precision dan menunjukkan bahwa model belum mampu mengenali karakteristik jantung normal secara optimal. Kesalahan prediksi yang masih ditemukan tidak mengurangi kontribusi utama penelitian ini, yaitu keberhasilan dalam merancang model deteksi VSD berbasis YOLOv11 yang mampu memproses citra ekokardiografi secara cepat dan mendeteksi area yang berpotensi mengandung defek septum. Hasil tersebut menunjukkan bahwa konfigurasi model yang digunakan telah memberikan performa yang stabil, meskipun peningkatan lanjutan tetap diperlukan untuk menyempurnakan performa generalisasi.

IV. KESIMPULAN

Penelitian ini mengembangkan model deteksi VSD yang mampu memprediksi lokasi defek pada sekat ventrikel secara cepat dan akurat pada pasien anak. Model tersebut bekerja dengan data inputan berupa video ekokardiografi dan menggunakan algoritma YOLOv11 untuk melakukan proses pendeteksian. Penelitian ini melatih ulang YOLOv11 menggunakan data VSD yang dilakukan *preprocessing*. Hasil penelitian menunjukkan bahwa penerapan tahapan augmentasi data, yang berfungsi untuk

meningkatkan variasi pada dataset, berhasil meningkatkan performa YOLOv11 dalam mendeteksi VSD secara signifikan, yakni lebih dari 10%. Dengan begitu, augmentasi data direkomendasikan dalam melatih model dengan jumlah yang terbatas seperti data medis. Jika ditinjau dari metrik evaluasi *precision*, *recall*, dan mAP juga waktu komputasi, model YOLOv11 terbukti mampu mendeteksi defek secara baik dari sisi akurasi dan kecepatan. Namun model hasil penelitian ini masih sedikit sensitif ketika memprediksi VSD pada data normal, yang kemungkinan dipengaruhi oleh adanya *noise* dalam data. Lebih lanjut, performa model dapat ditingkatkan dengan menerapkan teknik reduksi *noise* dalam data.

REFERENSI

- [1] *Cardiovascular disability: updating the Social Security listings*. National Academies Press, 2010.
- [2] Fajri Ismayanti and Nurul Munarti, "Gambaran X-Ray Penyakit Jantung Bawaan," *Jurnal Kesehatan Amanah*, vol. 7, no. 1, pp. 19–39, Apr. 2023, doi: 10.57214/jka.v7i1.261.
- [3] Mecca Yumna Ning Prisie, "IDAI soroti rendahnya intervensi penyakit jantung bawaan pada anak," <https://www.antaranews.com/berita/4847921/idai-soroti-rendahnya-intervensi-penyakit-jantung-bawaan-pada-anak>.
- [4] Mohamed H. Alahmadi and Tony I. Oliver, "Ventricular Septal Defect," in *StatPearls*, StatPearls Publishing, 2025. Accessed: Sep. 21, 2025. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK470330/>
- [5] K. Barut, İ. Peñçe, Ö. Çetinkaya Bozkurt, and M. Şişeci Çeşmeli, "Classification of Ventricular Septal Defect Disease Using Deep Learning," *Acta Infologica*, vol. 0, no. 0, pp. 0–0, Feb. 2025, doi: 10.26650/acin.1474115.
- [6] 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.
- [7] Jacob Murel and Eda Kavlakoglu, "What is object detection?," <https://www.ibm.com/id-id/think/topics/object-detection>.
- [8] S. Hsin Chen, C. Wei Wang, I. H. Tai, K. Pen Weng, Y. Hui Chen, and K. Sheng Hsieh, "Modified YOLOv4-DenseNet Algorithm for Detection of Ventricular Septal Defects in Ultrasound Images.," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 7, pp. 101–108, Sep. 2021, doi: 10.9781/ijimai.2021.06.001.
- [9] S.-H. Chen *et al.*, "Optimizing Object Detection Algorithms for Congenital Heart Diseases in Echocardiography: Exploring Bounding Box Sizes and Data Augmentation Techniques," *Rev Cardiovasc Med*, vol. 25, no. 9, Sep. 2024, doi: 10.31083/j.rcm2509335.
- [10] Y. Yang *et al.*, "Classification of normal and abnormal fetal heart ultrasound images and identification of ventricular septal defects based on deep learning," *J Perinat Med*, vol. 51, no. 8, pp. 1052–1058, Oct. 2023, doi: 10.1515/jpm-2023-0041.
- [11] M. Nasrudin, "Deteksi Penyakit Jantung Bawaan pada Anak Jenis VSD Menggunakan You Only Look Once (YOLO) Melalui Video Ultrasound," Institut Teknologi Sepuluh November, 2021.
- [12] N. I. Nirmalasari, "Deteksi Ventricular Septal Defect Berdasarkan Video Ultrasound Jantung Anak pada Parasternal dan Apical View dengan Faster Region-Based Convolutional Neural Network," Institut Teknologi Sepuluh Nopember, Surabaya, 2021.
- [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," *J Imaging*, vol. 11, no. 8, p. 282, Aug. 2025, doi: 10.3390/jimaging11080282.
- [14] W.-S. Hsu, G.-T. Lin, and W.-H. Wang, "Real-Time Ultrasound Diagnosis of Developmental Dysplasia of the Hip Using an Attention-Enhanced YOLOv11

- Model,” *Int J Med Sci*, vol. 22, no. 16, pp. 4236–4249, Oct. 2025, doi: 10.7150/ijms.120138.
- [15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” May 2016, [Online]. Available: <http://arxiv.org/abs/1506.02640>
- [16] R. Sapkota *et al.*, “YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series,” *Artif Intell Rev*, vol. 58, no. 9, p. 274, Jun. 2025, doi: 10.1007/s10462-025-11253-3.
- [17] “Ultralytics YOLO11,” <https://docs.ultralytics.com/models/yolo11/>.