

Deteksi Bahasa Isyarat Dalam Pengenalan Huruf Hijaiyah Dengan Metode YOLOV5

Dadang Iskandar Mulyana, Muhammad Faizal Lazuardi, Mesra Betty Yel

Program Studi Teknik Informatika, STIKOM CKI.

Jl.Raden Inten II No.8, RT.5/RW.14, Duren Sawit, Kec. Duren Sawit, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta 13440.

E-mail: faizalgoku19@gmail.com

Naskah Masuk: 15 Juli 2022; Diterima: 06 Agustus 2022; Terbit: 18 Agustus 2022

ABSTRAK

Abstrak - Bahasa isyarat adalah Bahasa yang menggunakan gerakan tangan dan tubuh, serta ekspresi dalam menyampaikan kata dan kalimat. Belajar huruf Hijaiyah adalah langkah awal untuk bisa membaca Al-Qur'an. Anak penyandang tunarungu dan tunawicara memiliki IQ di bawah rata-rata anak normal, yang menyebabkan proses belajar mereka lebih lambat dan memerlukan metode khusus. Ada sekelompok besar tunarungu-bisu di seluruh dunia, dan bahasa isyarat adalah alat komunikasi utama dalam komunitas ini. Penyandang tunarungu dan tunanetra perlu untuk dapat berkomunikasi dengan orang lain yang mampu mendengar, dan orang yang mendengar juga perlu memahami bahasa isyarat, yang menghasilkan permintaan yang besar untuk pelajaran bahasa isyarat. Sudah banyak penelitian yang membahas tentang pendeteksian sebuah objek menggunakan citra digital untuk mengenali macam-macam bahasa isyarat. Penelitian kali ini penulis menerapkan metode YOLOV5 untuk mendeteksi Bahasa isyarat khususnya huruf hijaiyah. Pada penelitian ini penulis menggunakan dataset sebesar 1014 gambar dengan kelas huruf hijiyah dari alif sampai ya. Hasil penelitian yang di peroleh menggunakan metode YOLOV5 terbukti dapat mengenali objek secara konsisten dengan nilai tingkat akurasi yang cukup tinggi yaitu 95%

Kata kunci: Bahasa isyarat, YOLOV5, Objek deteksi, Hijaiyah

ABSTRACT

Abstract - Sign language is a language that uses hand and body movements, as well as expressions in conveying words and sentences. Learning Hijaiyah letters is the first step to being able to read the Qur'an. Children who are deaf and speech impaired have IQs below the average normal child, which causes their learning process to be slower and requires special methods. There is a large group of deaf-mutes around the world, and sign language is the main means of communication in this community. Deaf and blind people need to be able to communicate with others who can hear, and hearing people also need to understand sign language, which results in a great demand for sign language lessons. There have been many studies that discuss the detection of an object using digital images to recognize various sign languages. In this study, the author applies the YOLOV5 method to detect sign language, especially hijaiyah letters. In this study, the author uses a dataset of 1014 images with hijiyah letter classes from alif to ya. The results obtained using the YOLOV5 method are proven to be able to recognize objects consistently with a fairly high accuracy value of 95%

Keywords: Sign Language, YOLOV5, Object Detection, Hijaiyah

Copyright © 2022 Universitas Muhammadiyah Jember.

1. PENDAHULUAN

Didunia tidak semua anak terlahir dengan kondisi yang normal, salah satu contohnya adalah penyandang tunarungu, dimana mereka mengalami kesulitan untuk mendengar. Salah satu metode yang digunakan agar mereka dapat berkomunikasi dengan kita sebagai orang normal adalah salah satunya dengan bantuan bahasa isyarat. Banyak diantara penyandang keterbatasan tersebut yang beragama islam, dan belajar Al-Qur'an adalah hukumnya wajib bagi setiap muslim tanpa terkecuali. Tetapi kita tidak bisa mengajarkan mereka dengan cara orang normal, karena penyandang tuna rungu tidak bisa mendengarkan apa yang kita ucapkan, disitulah peran bahasa isyarat agar mereka mengerti dengan apa yang kita ucapkan dan ajarkan kepada mereka. Teknologi bantu yang digunakan dalam membantu anak berkebutuhan khusus seperti tunarungu merupakan pengembangan perangkat keras dan perangkat lunak yang disesuaikan dengan kemampuan yang dimiliki oleh anak tersebut. Teknologi bantu lainnya berpusat pada pemrosesan dan tampilan suara yang bervariasi, sementara peneliti lain mengungkapkan penggunaan teknologi bantu yang

merupakan alat manajemen diri yang bermanfaat bagi penyandang disabilitas mental dalam menyelesaikan aktivitas sehari-harinya [1]. Al-Qur'an merupakan kitab suci sekaligus pedoman hidup untuk umat Islam. Karena huruf Hijaiyah dan m engoreksi Makhraj. pentingnya membaca Al-Qur'an dengan benar, setiap Muslim harus memperhatikan Makhraj untuk membaca huruf Hijaiyah (huruf Arab) [2].

Meskipun orang selalu bertanya “apa kesulitan dan tantangan dalam pendeteksian objek?”, sebenarnya pertanyaan ini tidak mudah untuk dijawab dan bahkan mungkin terlalu digeneralisasikan. Karena berbeda objek yang dideteksi, berbeda pula tujuan dan kendala yang dihadapi satu sama lain [3]. Kemajuan teknologi membawa perubahan yang cukup signifikan dalam deteksi objek. Dibandingkan dengan deteksi objek gambar diam, deteksi objek video lebih menantang karena variasi tampilan yang drastis terjadi pada frame video. Variasi tampilan yang umum adalah blur, yang dapat menurunkan akurasi deteksi secara signifikan [4]. Bahasa Isyarat merupakan cara manusia untuk berkomunikasi yang paling awal. Misalkan ada orang bepergian ke negara asing mana pun, mengabaikan bahasa mereka sepenuhnya, dia dapat dengan mudah, dan secara naluriah, menemukan cara untuk membuat tanda-tanda untuk setidaknya minum air, makan, dan mendapatkan tempat untuk tidur. Sepanjang zaman ini, orang tuli dan buta serta keluarga di sekitarnya telah mengembangkan, selama beberapa dekade, serangkaian gerakan tangan, ekspresi wajah, serta gerakan bibir untuk dapat berkomunikasi satu sama lain. Tanda-tanda ini berbeda beda antar negara dan bahasa yang berbeda, walaupun maksudnya memiliki beberapa kesamaan [5]. Belajar Isyarat sangat sulit bagi seorang pemula tanpa bantuan praktisi bahasa isyarat yang terlatih. Pembelajaran melalui buku dinilai kurang efektif karena tidak mudah menjelaskan bahasa isyarat ke dalam buku dengan menggunakan gambar. Kesulitan dalam memahami bahasa lisan dan bentuk tulisannya, keterbatasan kemampuan bahasa isyarat dari para guru dan tingginya biaya orang tua yang dikeluarkan dalam mendidik anak tunarungu mereka adalah faktor-faktor yang berdampak negative terhadap pembelajaran bahasa isyarat. Selain orang tunarungu, tunawicara, guru, orang tua, pekerja sosial, dan peneliti perlu mempelajari tanda. Sulit bagi mereka untuk mengikuti program pelatihan tanda belajar. Bagi orang tua yang mendengar dan berbicara dari anak tunarungu, kurangnya mekanisme belajar ditambah dengan kemampuan berbicara mereka membuat mereka lebih menyukai membaca bibir daripada menggunakan bahasa isyarat. Kendala seperti ini yang membuat anak sulit berkomunikasi dengan baik, seperti berbeda variasi bahasa lisan, bahasa isyarat antar regional. Ini adalah masalah komunikasi dalam komunitas tunarungu itu sendiri [6].

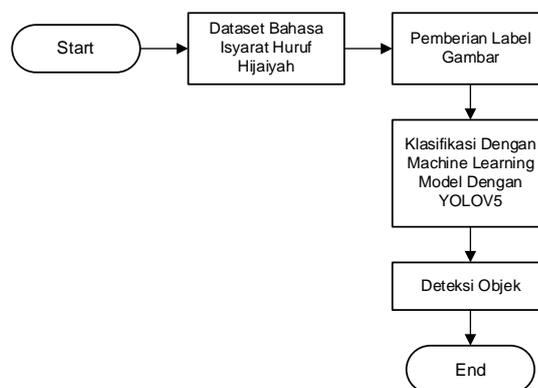
Ada beberapa macam untuk membuat aplikasi pendeteksian objek, salah satunya adalah memakai metode *You Only Look Once* (YOLO). Sistem pendeteksi metode YOLO terbukti lebih cepat dan akurat untuk mendeteksi objek pada gambar atau citra sehingga paling sesuai jika diterapkan untuk *real-time* pendeteksian objek pada video. Dengan menggunakan pendeteksian objek metode YOLO pada suatu sistem dapat membantu mengklasifikasi setiap tanda gerakan atau bahasa isyarat yang diperagakan oleh pengajar kepada muridnya yaitu penyandang tuna rungu secara *real-time* pada rekaman video. Jenis bahasa isyarat yang diperagakan akan terdeteksi otomatis berdasarkan nilai hasil tingkat akurasinya dan klasifikasi

2. METODE PENELITIAN

2.1 Data Penelitian

Pada penelitian ini, dataset yang digunakan diambil dari dataset public website Kaggle.com yang diunggah oleh akun Ammar Sayyed yang berjudul “Arabic Sign Language Dataset 2022 (416 x 416)”. Dataset memiliki label Bahasa isyarat huruf hijaiyah sebanyak 4.247 citra [7].

2.2 Penerapan metodologi

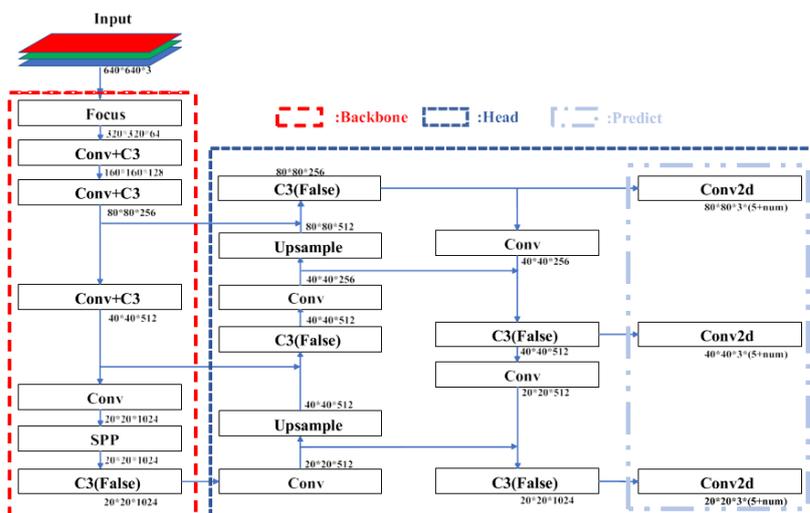


Gambar 1. Tahapan Metodologi Yang Digunakan

2.2.1 Pemberian label gambar

Setelah dataset yang diperlukan sudah terkumpul, tahap selanjutnya adalah pemberian label pada tiap-tiap gambar yang akan ditrain atau diuji. Hal ini berguna agar computer dapat mengenali maksud dari gambar yang akan dideteksi. Untuk memberikan label pada suatu gambar banyak caranya, disini penulis menggunakan bantuan website makesense.ai dari internet untuk memberikan label pada tiap-tiap gambar.

2.2.2 Klasifikasi dengan machine learning model menggunakan YOLOV5



Gambar 2. Model Jaringan YOLOv5 [8].

Seperti yang ditunjukkan pada Gambar 2, model jaringan YOLOv5 dibagi menjadi tulang punggung dan Kepala. Input di-down sampel untuk $\times 2$ kali melalui tulang punggung jaringan, dan kemudian mengambil tiga lapisan fitur terakhir sebagai input kepala. Di tulang punggung jaringan, data pertama melewati lapisan fokus untuk memotong lebar dan ketebalan data menjadi setengah dari aslinya, sangat mempercepat kecepatan transmisi maju jaringan. Kemudian berlalu melalui empat lapisan konvolusi ditambah lapisan C3. Ini mengusulkan komponen residu baru (lapisan C3) yang ditingkatkan sejak Lapisan CSP di jaringan YOLOv4. Dengan bantuan lapisan C3, jaringan tulang punggung dapat mengekstrak fitur yang lebih detail. Selain itu, YOLOv5 menggantikan fungsi aktivasi di Lapisan CSP mengikuti metode fungsi aktivasi yang ditingkatkan di YOLOv4. Akhirnya, perlu dicatat bahwa modul SPP adalah ditambahkan di tengah lapisan ConvCC3 terakhir. Modul ini diusulkan dari YOLOv3, yang dapat memperluas reseptif bidang lapisan fitur dan memainkan peran besar dalam perpaduan fitur Modul kepala. Di bagian kepala YOLOv5, ide utamanya sama seperti YOLOv4, yang ditingkatkan berdasarkan PANet. Perlu dicatat bahwa YOLOv5 menggunakan lapisan C3 yang berbeda pola di kepala daripada tulang punggung. Di Backbone, fiturnya lapisan hanya melakukan pengambilan sampel, sehingga lapisan C3 mengadopsi kombinasi sisa. Namun, di kepala, integrasi jaringan PANet memarut lapisan fitur pengambilan sampel naik-turun, sehingga residu kombinasi tidak lagi digunakan di lapisan C3 [8].

2.2.3 Deteksi objek

Dalam digital processing, deteksi objek adalah suatu proses yang digunakan untuk menentukan keberadaan objek tertentu di dalam suatu citra digital. Proses deteksi tersebut dapat dilakukan dengan berbagai macam metode yang umumnya melakukan pembacaan fitur-fitur dari seluruh objek pada citra input

3. HASIL DAN PEMBAHASAN

3.1 Evaluasi Metode

Penulis menggunakan Google Colab dan menerapkan metode YOLOV5. YOLO merupakan jaringan untuk mendeteksi objek sedangkan YOLOV5 adalah metode versi terbaru yang dikembangkan metode YOLO. Tugas pendeteksian objek untuk menentukan tempat pada sebuah gambar atau citra pada objek yang hadir dan mengklasifikasikan jenis objeknya. Jadi sederhananya ada sebuah gambar

atau citra menjadi inputan, kemudian buat vektor kotak pembatas dan prediksi klas dalam outputnya. Pada penelitian ini, penulis mengumpulkan dataset Gerakan tangan atau Bahasa isyarat huruf hijaiyah sebanyak 30 jenis citra, berikut contoh 5 huruf hijaiyah:

Tabel 3.1 Contoh Gerakan Tangan atau Bahasa Isyarat Huruf Hijaiyah

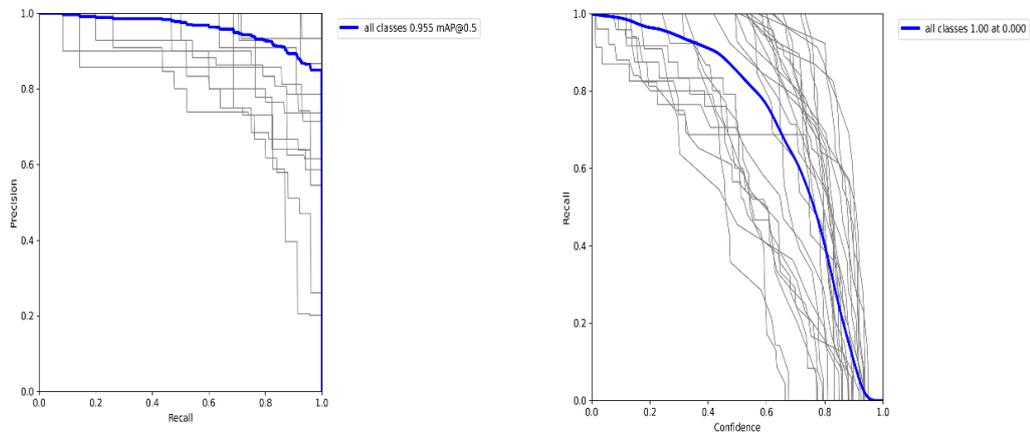
Huruf Hijaiyah	Gerakan Tangan / Bahasa Isyarat
(Alif) ا	
(Ba) ب	
(Ta) ت	
(Tsa) ث	
(Jim) ج	

Setelah dataset dirasa cukup, langkah selanjutnya adalah memberi label pada tiap-tiap gambar supaya computer dapat mengenali nama-nama dari Gerakan tersebut untuk dipelajari. Disini saya menggunakan website makesense.ai untuk memberikan label pada tiap gambar, berikut adalah contoh dari proses pelabelan pada tiap gambar.

Tabel 3.2 Pelabelan Pada Gambar

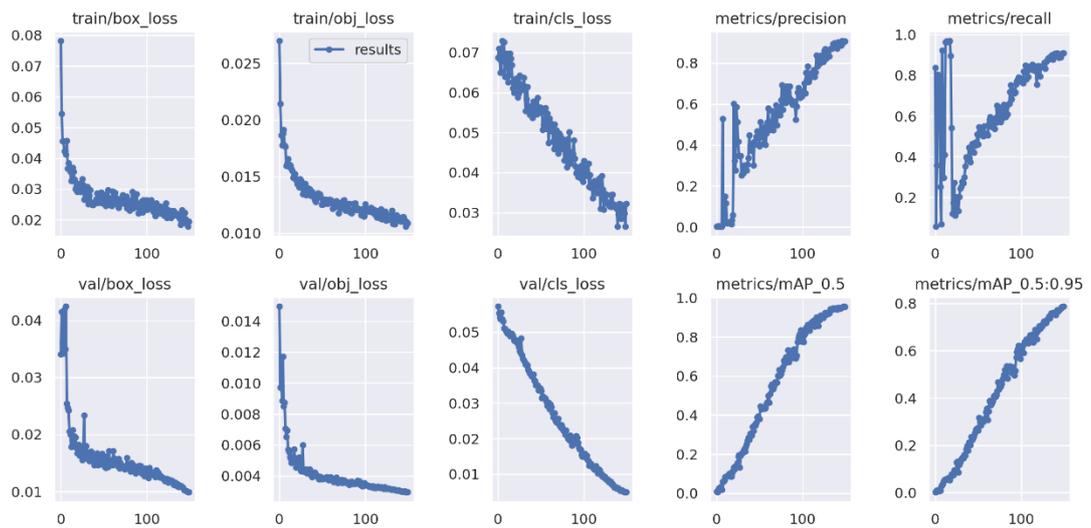
Huruf Hijaiyah	Proses label gambar
(Alif) ا	
(Ta) ت	
(Waw) و	

Pada tahap training penulis menggunakan software berbasis web/cloud yaitu google colab dengan menguji 1104 gambar, 28 class. Gambar-gambar tersebut di uji sebanyak 150 epoch dan batch 16, didapatkan hasil sebagai berikut:

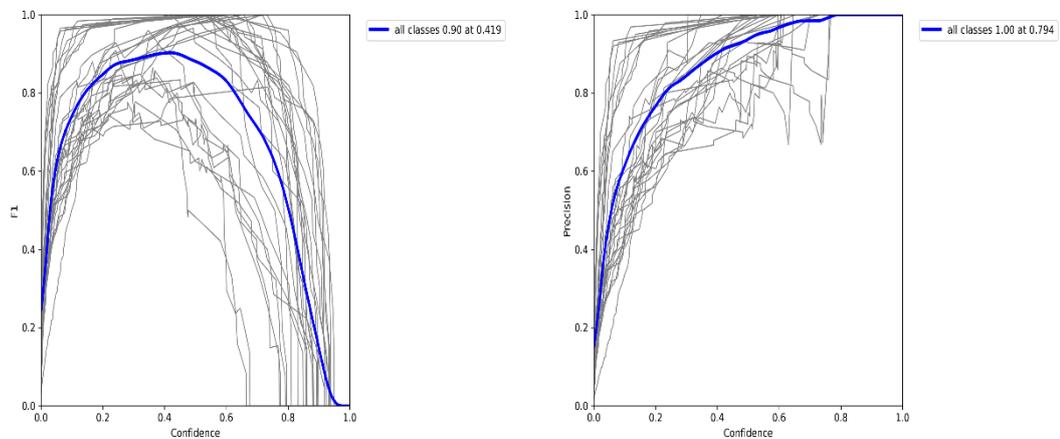


Gambar 3. Nilai Recall

Dari nilai recall yang di peroleh dari tarin diatas dapat dilihat tingkat ke akurasiannya adalah sebesar 0,955 atau sekitar 95 % terhadap jumlah datanya



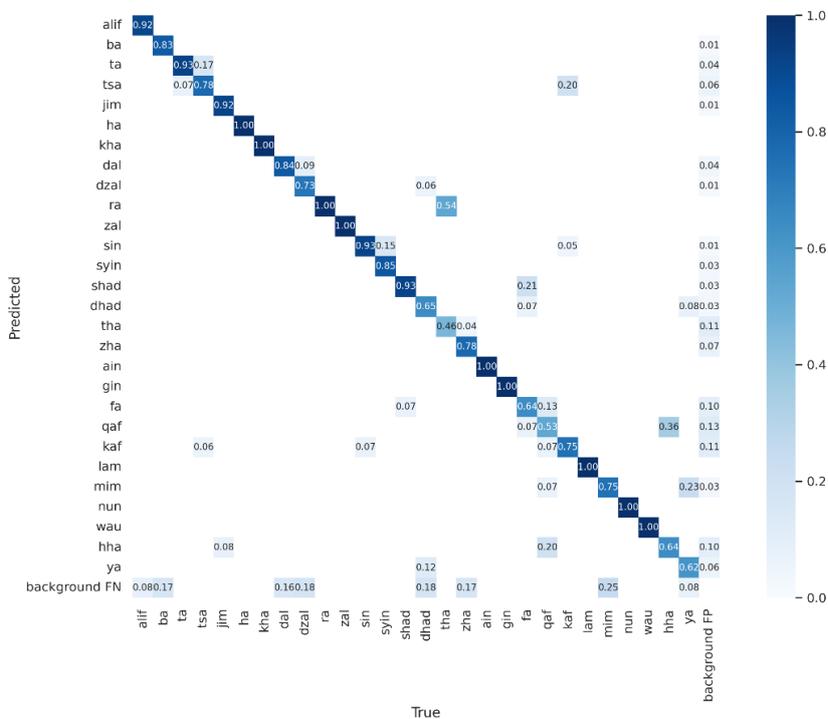
Gambar 4. Hasil Evaluasi Data Training



Gambar 5. Kurva nilai F1 dan *precision* terhadap nilai *confidence*

Pada gambar 5 diatas nilai F1 mendapatkan nilai puncak rata-rata 0,90 terhadap nilai confidence 0,419. Nilai precision mendapatkan nilai rata-rata 1,00 pada nilai confidence 0,794.

3.2 Analisis



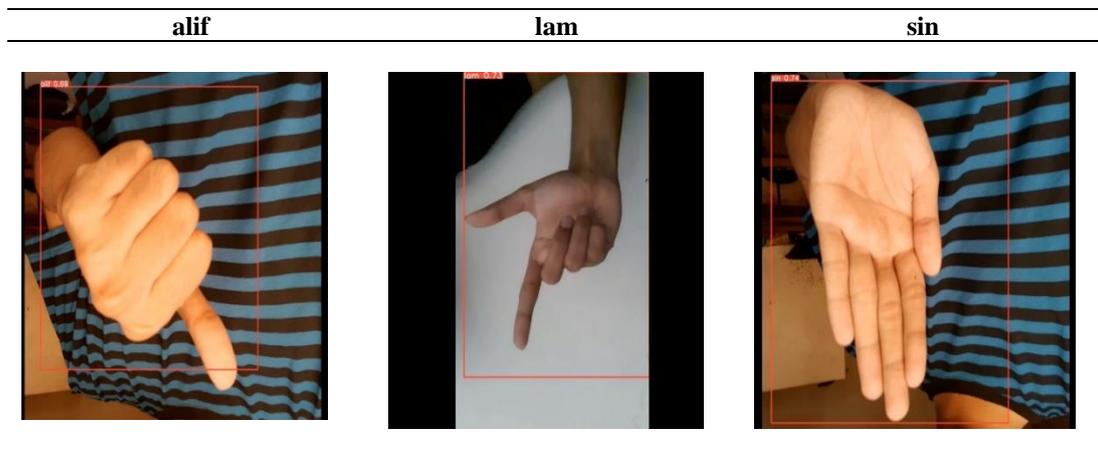
Gambar 6. Confusion_matrix

Dari gambar 6 confusion matrix diatas dapat dilihat tingkat keakurasian yang cukup baik, bahkan sampai ada yang menyentuh angka 1,00 yang artinya sangat akurat, contohnya di huruf ha, kha dan ra dan akurasi nilai terendah ada di huruf tha dengan angka 0,46. Pendeteksian bahasa isyarat huruf hijaiyah menggunakan metode YOLOV5 berjalan dengan lancar dan nilai akurasinya pun cukup tinggi. Dibawah ini adalah gambar pendeteksian bahasa isyarat:



Gambar 7. Angka Prediksi

Pada pengujian pendeteksian video yang diambil juga mendapatkan nilai akurasi yang cukup tinggi juga dan didalam video deteksi bahasa isyarat huruf hijaiyah dapat mendeteksi secara stabil, Berikut ini gambar pengujian deteksi video:



4. KESIMPULAN

Berdasarkan penelitian dan hasil data train, diperoleh hasil yang cukup tinggi. Dimana pada penelitian ini di train sebanyak 150 epoch untuk mendapatkan tingkat akurasi yang tinggi. Hasilnya nilai akurasi menunjukkan angka 0.95 yang artinya mendekati angka 1 yaitu sekitar 95%, dimana di akurasi tersebut objek dapat terdeteksi dengan baik.

REFERENSI

- [1] I. Afrianto, S. L. Tremain, A. F. Faris, and S. Atin, "Hijaiyah Letter Interactive Learning for Mild Mental Retardation Children using Gillingham Method and Augmented ... Cite this paper Related papers Ext ending Tangible Int eract ive Int erfases for Education: A System for Learning Arabic Braille ... Duaa alSaeed Telecommunications Problems and Design Strategies For People With Cognitive Disabilities Hijaiyah Letter Interactive Learning for Mild Mental Retardation Children using Gillingham Method and Augmented Reality," 2019. [Online]. Available: www.ijacsa.thesai.org
- [2] F. W. Wibowo and Institute of Electrical and Electronics Engineers, *2018 International Conference on Information and Communications Technology (ICOIACT) : 6-7 March 2018*.
- [3] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," May 2019, [Online]. Available: <http://arxiv.org/abs/1905.05055>
- [4] Y. Wu, H. Zhang, Y. Li, Y. Yang, and D. Yuan, "Video Object Detection Guided by Object Blur Evaluation," *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.3038913.
- [5] M. A. Bencherif *et al.*, "Arabic Sign Language Recognition System Using 2D Hands and Body Skeleton Data," *IEEE Access*, vol. 9, pp. 59612–59627, 2021, doi: 10.1109/ACCESS.2021.3069714.
- [6] J. Joy, K. Balakrishnan, and M. Sreeraj, "SignQuiz: A quiz based tool for learning fingerspelled signs in indian sign language using ASLR," *IEEE Access*, vol. 7, pp. 28363–28371, 2019, doi: 10.1109/ACCESS.2019.2901863.
- [7] AMMAR SAYED, "Arabic Sign Language Dataset 2022," <https://www.kaggle.com/datasets/ammarsayedtaha/arabic-sign-language-dataset-2022>, 2022.
- [8] H. Zhang, M. Tian, G. Shao, J. Cheng, and J. Liu, "Target Detection of Forward-Looking Sonar Image Based on Improved YOLOv5," *IEEE Access*, vol. 10, pp. 18023–18034, 2022, doi: 10.1109/ACCESS.2022.3150339.