

Pengembangan Sistem Pendeteksi Objek dengan MobileNet-SSD

Boy Erdyansyah Gunawan¹, Enryco Hidayat², Mohammad Habim Hazidan Rifqi³,

Yovi Ibnu Nasikhin⁴, Adi Fajri Firmansyah⁵, Basuki Rahmat⁶

^{1,2,3,4,5} Program Studi Informatika, Fakultas Ilmu Komputer

¹21081010021@student.upnjatim.ac.id

²21081010023@student.upnjatim.ac.id

³21081010031@student.upnjatim.ac.id

⁴21081010128@student.upnjatim.ac.id

⁵21081010175@student.upnjatim.ac.id

⁶basukirahmat.if@upnjatim.ac.id

Corresponding author : 21081010023@student.upnjatim.ac.id

Abstrak

MobileNet SSD adalah salah satu arsitektur convolutional neural network (CNN) yang populer digunakan untuk deteksi objek dalam gambar dan video real-time. Dalam penelitian ini, MobileNet SSD akan diuji dalam deteksi objek pada gambar statis dan real-time melalui kamera. MobileNet SSD V2 dirancang untuk mendeteksi satu objek atau beberapa objek secara real-time dengan menggunakan depthwise separable convolution yang secara signifikan mengurangi jumlah parameter jika dibandingkan dengan regular convolution. Hasil dari penelitian ini menunjukkan deteksi yang baik, namun terdapat kesalahan deteksi pada kondisi tertentu, terutama ketika objek berada dalam posisi kurang tepat. Sebagai solusi dari permasalahan yang ada perlu dilakukan augmentasi data, preprocessing yang lebih baik, optimasi model, penyesuaian resolusi kamera, serta algoritma penyesuaian pencahayaan dan stabilisasi gambar. Langkah-langkah ini bertujuan untuk meningkatkan akurasi deteksi dan mengurangi kesalahan identifikasi objek, sehingga aplikasi MobileNet SSD dapat digunakan secara lebih luas dan dapat diandalkan dalam berbagai situasi. Dengan demikian, MobileNet-SSD dapat menjadi solusi yang efektif dan efisien untuk deteksi objek dalam berbagai kondisi dan lingkungan.

Kata Kunci: MobileNet, SSD, CNN, deteksi objek, gambar, real-time.

Salah satu metode yang populer dalam deteksi objek adalah MobileNet-SSD V2. Sebab dengan metode ini mampu melakukan deteksi objek dengan menghasilkan akurasi tinggi dan area terdeteksi untuk setiap kategori objek dalam gambar. MobileNet-SSD V2 dikenal karena kemampuannya untuk mendeteksi berbagai objek dengan akurasi yang tinggi, memastikan bahwa setiap kategori objek yang ada dalam gambar dapat diidentifikasi dengan tepat. Kombinasi antara SSD dan MobileNet V2 memberikan performa yang optimal dalam deteksi objek, menjadikannya pilihan yang ideal untuk aplikasi yang memerlukan kecepatan dan akurasi tinggi dalam pengenalan objek [3].

Penelitian ini akan fokus pada pengembangan deteksi objek menggunakan MobileNet-SSD. Tujuan dari penelitian ini adalah untuk mengetahui kecepatan dan akurasi deteksi objek menggunakan MobileNet-SSD, serta untuk mengetahui bagaimana metode ini dapat diterapkan dalam berbagai aplikasi pengawasan. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan teknologi pengawasan yang lebih efektif dan efisien, serta membantu meningkatkan keselamatan dan efisiensi operasional robot pengawasan.

I. PENDAHULUAN

Dalam era teknologi modern, pengembangan sistem pengawasan yang efektif dan efisien telah menjadi sangat penting. Salah satu aplikasi yang sangat relevan adalah pengembangan robot pengawasan yang dapat mendeteksi objek secara real-time dan akurat. Dalam konteks ini, deteksi objek menjadi kunci penting dalam memastikan keamanan dan efisiensi operasional robot tersebut [1].

Deteksi objek telah menjadi topik penelitian yang sangat aktif dalam beberapa tahun terakhir, dengan berbagai metode dan algoritma yang dikembangkan untuk meningkatkan kecepatan dan akurasi deteksi. Salah satu metode yang sangat populer dan efektif adalah menggunakan MobileNet-SSD, yang telah menunjukkan hasil yang sangat baik dalam berbagai aplikasi, termasuk pengawasan objek bergerak [2].

II. MATERIAL & METODE

Pendeteksian objek adalah bagian dari ilmu computer vision yang memproses gambar untuk mendeteksi objek yang ada dalam sebuah gambar. Dalam konteks ini, objek yang dituju adalah berbagai objek yang ada dalam sebuah citra. Setiap objek memiliki fitur khusus yang membedakan mereka dalam sebuah citra. Ada banyak teknik yang dapat digunakan untuk mendapatkan fitur pada citra sesuai dengan kebutuhan objek yang akan dideteksi [4]. Dalam penelitian ini, digunakan beberapa bahan atau metode diantaranya :

A. Artificial Intelligence

Artificial Intelligence atau kecerdasan buatan adalah sistem komputer yang mampu melakukan tugas-tugas yang

biasanya memerlukan kecerdasan manusia. Teknologi ini dapat membuat keputusan dengan cara menganalisis dan menggunakan data yang tersedia dalam sistem. Proses yang terjadi dalam *Artificial Intelligence* mencakup learning, reasoning, dan self-correction. Proses ini mirip dengan proses yang dilakukan manusia dalam menganalisis sebelum mengambil keputusan.

Menurut John McCarthy pada tahun 1956, *Artificial Intelligence* bertujuan untuk memahami dan memodelkan proses berpikir manusia serta merancang mesin agar dapat meniru perilaku manusia. Kecerdasan dalam konteks ini mencakup pengetahuan dan pengalaman, kemampuan penalaran (bagaimana membuat keputusan dan mengambil tindakan), serta moral yang baik [5].

Dorongan untuk mencapai kondisi yang memberikan kemampuan kepada perusahaan untuk tetap tumbuh, berkembang, dan menghasilkan keuntungan yang wajar, didorong oleh kekhawatiran yang sering menghantui para pebisnis terkait dengan perubahan dinamis dalam lingkungan industri yang menghadapi Revolusi Industri 4.0. Pada era ini, permintaan terhadap produk yang sesuai dengan keinginan konsumen dan dapat diperoleh dalam waktu relatif singkat menjadi tema utama dalam pengembangan produk. Faktor biaya per unit yang minimal dan mutu produk yang tinggi tidak lagi cukup untuk membangun daya saing. Daya saing ditentukan oleh kinerja dalam pengiriman, kemampuan memenuhi keinginan konsumen, dan kualitas keterlibatan dalam menangani isu-isu lingkungan.

B. Deep Learning

Deep learning adalah suatu metode dalam teknologi kecerdasan buatan (AI) yang memanfaatkan jaringan saraf tiruan (*Artificial Neural Network*) yang berlapis-lapis (multi-layer). Dalam beberapa penelitian, deep learning digunakan untuk mengembangkan teknologi pengenalan citra, seperti pengenalan wajah dan klasifikasi citra. Salah satu contoh implementasi *deep learning* yang paling sering digunakan adalah *Convolutional Neural Network* (CNN), yang dirancang khusus untuk pengenalan dan klasifikasi gambar [6].

Selain itu, *deep learning* memberikan kontribusi besar dalam *computer vision*, terutama dalam *object detection*. Implementasi model *deep learning* untuk deteksi objek telah banyak digunakan di berbagai produk riset dan komersial seperti self-driving car dan CCTV cerdas. Namun, penggunaannya memerlukan perangkat yang mahal seperti GPU atau PC dengan spesifikasi tinggi.

MobileNets, arsitektur convolutional neural network (CNN), hadir sebagai solusi untuk mengatasi kebutuhan akan sumber daya komputasi yang berlebih. Arsitektur MobileNet dikembangkan oleh peneliti Google untuk penggunaan pada perangkat mobile [7].

Perbedaan utama antara MobileNet dan CNN konvensional terletak pada penggunaan *depthwise convolution* dan *pointwise convolution* untuk mengurangi kompleksitas model.

C. Deteksi Objek

Deteksi objek adalah teknologi komputer yang digunakan untuk mengidentifikasi dan menemukan objek-objek tertentu, seperti manusia, gedung, atau mobil, dalam gambar dan video digital. Proses ini melibatkan pencarian posisi dan ukuran setiap objek dalam gambar. Tujuannya adalah untuk menemukan semua objek dari satu atau beberapa kategori, terlepas dari ukuran, posisi, pose, sudut pandang, dan kondisi pencahayaan objek tersebut [8].

MobileNetV2 adalah salah satu arsitektur *convolutional neural network* (CNN) yang dirancang untuk mengatasi kebutuhan akan sumber daya komputasi yang berlebihan atau memerlukan komputasi tinggi. Arsitektur ini dikembangkan khusus untuk aplikasi di perangkat *mobile* dan *embedded* yang memiliki keterbatasan daya komputasi [9].

MobileNetV2 menggunakan teknik *depthwise separable convolution* untuk mengurangi jumlah parameter dan kompleksitas model, sehingga cocok digunakan pada perangkat dengan spesifikasi rendah namun tetap mempertahankan akurasi yang baik dalam tugas-tugas seperti pengenalan objek dan klasifikasi gambar. Metode ini memungkinkan pengurangan beban komputasi tanpa mengorbankan kualitas hasil deteksi atau klasifikasi [10].

Selain itu, *Single Shot Detector* (SSD) juga digunakan sebagai pendeteksi objek dalam studi ini. SSD adalah detektor objek satu tahap yang melakukan klasifikasi dan lokalisasi dalam satu jaringan (*single network*), yang membuatnya lebih sederhana dibandingkan detektor dua tahap. Keuntungan dari detektor satu tahap ini adalah proses *end-to-end learning* dan kesederhanaan model. Model arsitektur SSD membutuhkan jaringan tambahan untuk mengekstrak fitur yang diperlukan. Dalam studi ini, digunakan MobileNet sebagai jaringan ekstraksi fitur, sehingga fitur-fitur yang diekstraksi oleh MobileNet akan digunakan sebagai data pelatihan untuk SSD [11].

D. Webcam Laptop

Pada penelitian ini, webcam laptop diintegrasikan sebagai kamera untuk sistem deteksi objek menggunakan MobileNet-SSD v2. Penggunaan webcam laptop memungkinkan pengambilan gambar secara real-time untuk analisis deteksi objek menggunakan arsitektur deep learning yang efisien ini [12].

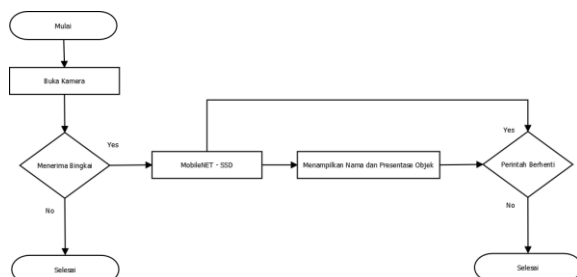
Metode ini memanfaatkan keunggulan MobileNet-SSD v2 dalam menawarkan deteksi objek yang cepat dan akurat, sambil mengurangi kebutuhan akan sumber daya komputasi yang tinggi, yang sering kali menjadi batasan pada aplikasi mobile atau perangkat terbatas lainnya.

Dengan integrasi ini, diharapkan dapat meningkatkan kemampuan deteksi objek dalam konteks aplikasi yang lebih luas, seperti keamanan visual, analisis video, dan aplikasi berbasis interaksi pengguna [13].

Dengan integrasi ini, diharapkan dapat meningkatkan kemampuan deteksi objek dalam konteks aplikasi yang lebih luas, seperti keamanan visual, analisis video, dan aplikasi berbasis interaksi pengguna. Misalnya, dalam bidang keamanan, penggunaan webcam laptop yang dilengkapi dengan MobileNet-SSD v2 dapat mendeteksi aktivitas mencurigakan atau intrusi secara real-time, memberikan respons cepat untuk mengatasi potensi ancaman. Dalam analisis video, teknologi ini dapat digunakan untuk pelacakan objek dan analisis situasional, yang bermanfaat dalam berbagai industri termasuk ritel dan pengawasan [14].

Selain itu, aplikasi berbasis interaksi pengguna seperti asisten virtual atau permainan augmented reality dapat memanfaatkan deteksi objek real-time untuk meningkatkan pengalaman pengguna. Webcam laptop yang dikombinasikan dengan teknologi deteksi objek ini dapat mengenali gerakan dan objek di sekitar pengguna, memungkinkan interaksi yang lebih alami dan responsif. Dengan demikian, integrasi antara webcam laptop dan MobileNet-SSD v2 tidak hanya memberikan solusi yang efisien dan efektif untuk deteksi objek, tetapi juga membuka peluang baru untuk inovasi dalam berbagai aplikasi teknologi [15].

E. Flowchart / Rancangan



Gambar. 1 Flowchart Program

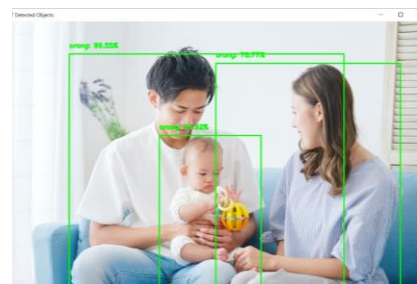
Pada gambar diatas menjelaskan langkah-langkah utama dalam proses deteksi objek menggunakan MobileNet-SSD melalui kamera. Alur dimulai dengan membuka akses ke kamera untuk mengambil gambar secara real-time. Setelah sistem berhasil menerima bingkai dari kamera, langkah berikutnya adalah menggunakan model MobileNet-SSD untuk melakukan deteksi objek pada bingkai tersebut. Hasil dari deteksi objek, termasuk nama objek dan tingkat kepercayaan deteksi, ditampilkan pada layar.

Selanjutnya, sistem memeriksa apakah ada perintah untuk menghentikan proses deteksi. Jika perintah berhenti

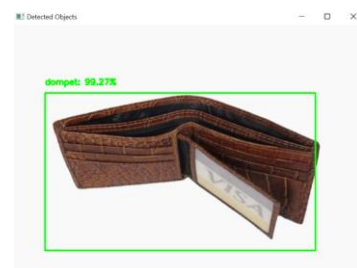
diterima, proses akan berakhir; jika tidak, sistem akan kembali ke langkah penerimaan bingkai untuk melanjutkan deteksi objek. Akhirnya, proses akan selesai setelah tidak ada lagi bingkai yang diterima atau ada perintah untuk mengakhiri deteksi. Dengan demikian, flowchart ini memberikan gambaran tentang bagaimana sistem melakukan deteksi objek secara efisien dan responsif terhadap input dari pengguna. Ini adalah pendekatan yang kuat untuk aplikasi deteksi objek real-time yang memungkinkan interaksi yang lancar antara sistem dan pengguna dalam mendeteksi objek di lingkungan nyata.

III. HASIL DAN PEMBAHASAN

Pada tahap awal evaluasi, kami menguji model MobileNet-SSD dengan melakukan input gambar statis untuk melihat seberapa baik model mendeteksi objek dalam gambar.

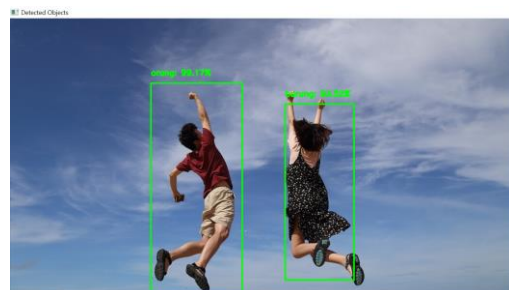


Gambar. 2 Hasil deteksi gambar orang



Gambar. 3 Hasil deteksi gambar dompet

Dari hasil deteksi gambar yang terlihat pada gambar 1 dan gambar, model sudah benar mendeteksi objek dengan akurasi yang tinggi.



Gambar. 4 Hasil contoh salah deteksi gambar

Namun, terdapat beberapa kasus di mana model tidak akurat atau tidak menghasilkan deteksi objek yang salah seperti yang terlihat pada gambar 3. Beberapa masalah yang kami temui termasuk posisi objek yang kurang ideal dan kondisi latar belakang yang kompleks atau *noise*. Ketika objek tidak berada pada posisi yang ideal, seperti terlalu dekat dengan tepi gambar atau sebagian terhalang, model sering gagal mendeteksi objek dengan benar. Selain itu, latar belakang yang kompleks dapat mengganggu kemampuan model untuk mendeteksi objek, seperti pola yang berantakan atau objek-objek lain yang menyerupai target.

Untuk mengatasi masalah ini, terdapat beberapa solusi yang mungkin dapat membantu. Pertama, melakukan augmentasi data pada saat pelatihan dapat membantu mensimulasikan berbagai posisi objek dan kondisi latar belakang. Teknik augmentasi dapat mencakup rotasi, pemotongan, perubahan pencahayaan, dan penambahan noise pada gambar pelatihan. Kedua, penggunaan dataset yang lebih beragam dapat memastikan bahwa dataset pelatihan mencakup berbagai variasi posisi objek dan kondisi latar belakang, yang membantu model belajar mengenali objek dalam berbagai situasi. Ketiga, fokus pada preprocessing data, seperti normalisasi pencahayaan dan pengurangan noise latar belakang, sebelum gambar diproses oleh model dapat meningkatkan akurasi deteksi.

Selanjutnya, kami mengimplementasikan model MobileNet-SSD untuk melakukan deteksi objek secara realtime menggunakan input video dari kamera. Evaluasi kinerja dilakukan dengan mengukur akurasi deteksi serta kecepatan pemrosesan.



Gambar. 5 Hasil deteksi lewat kamera

Dari hasil deteksi gambar yang terlihat pada gambar 4 dan gambar, model sudah benar mendeteksi objek dengan akurasi yang tinggi. Sama seperti yang saat melakukan pengujian dengan input gambar, pengujian lewat kamera juga ditemui permasalahan tidak keakuratan ketika posisi objek yang kurang ideal. Selain itu juga terdapat tantangan lain seperti kecepatan pemrosesan, kondisi pencahayaan yang berubah-ubah, dan gerakan objek serta kamera. Untuk aplikasi realtime, kecepatan pemrosesan menjadi sangat kritis, dan model harus cukup cepat

untuk memproses setiap frame video tanpa menyebabkan lag yang signifikan. Kondisi pencahayaan yang berubah-ubah, misalnya saat kamera bergerak dari area terang ke area gelap, juga dapat mempengaruhi kinerja model. Selain itu, gerakan yang cepat baik dari objek maupun kamera dapat menyebabkan blur, yang mempengaruhi akurasi deteksi.

Selain itu, resolusi kamera juga memainkan peran penting dalam akurasi dan kecepatan deteksi objek. Kami mengamati bahwa resolusi yang lebih tinggi dapat memberikan lebih banyak detail, yang dapat membantu dalam mendeteksi objek yang lebih kecil atau sebagian terhalang. Namun, peningkatan resolusi juga meningkatkan jumlah data yang harus diproses, yang dapat memperlambat kecepatan deteksi. Di sisi lain, resolusi yang lebih rendah dapat meningkatkan kecepatan pemrosesan tetapi mengurangi detail yang tersedia untuk model, yang dapat menurunkan akurasi deteksi.

Untuk mengatasi beberapa tantangan tersebut, terdapat beberapa langkah-langkah yang dapat dilakukan. Pertama, menggunakan algoritma stabilisasi gambar yang dapat membantu mengurangi efek blur yang disebabkan oleh gerakan kamera. Kedua, menerapkan algoritma penyesuaian pencahayaan secara dinamis dapat menjaga kualitas gambar input tetap konsisten di berbagai kondisi pencahayaan. Ketiga, menyesuaikan resolusi kamera untuk mencapai keseimbangan antara detail yang memadai dan kecepatan pemrosesan yang optimal. Sebagai contoh, memilih resolusi sedang yang cukup tinggi untuk memberikan detail yang memadai namun tidak terlalu tinggi sehingga memperlambat pemrosesan.

IV. KESIMPULAN

MobileNet-SSD memiliki potensi besar untuk digunakan dalam berbagai aplikasi deteksi objek, baik dalam gambar statis maupun secara realtime. Dengan implementasi solusi yang ada, seperti augmentasi data, preprocessing yang lebih baik, optimasi model, penyesuaian resolusi kamera, dan algoritma penyesuaian pencahayaan serta stabilisasi gambar, diharapkan kinerja model dapat ditingkatkan lebih lanjut. Penelitian lanjutan yang mencakup pengujian pada dataset yang lebih beragam dan skenario dunia nyata yang lebih kompleks juga diperlukan untuk memastikan bahwa model ini siap digunakan dalam berbagai aplikasi praktis. Dengan demikian, MobileNet-SSD dapat menjadi solusi yang efektif dan efisien untuk deteksi objek dalam berbagai kondisi dan lingkungan.

REFERENSI

- [1] H. Tanujaya and Lina, "Pengenalan Objek Menggunakan Metode Single Shot Multibox Detector Pada Bahan Sembako," *J. Ilmu Komput. dan Sist. Inf.*, vol. 11, no. 1, pp. 1–6, 2023, doi: 10.24912/jiksi.v11i1.24067.
- [2] A. A. N. MULANA, "Perbandingan Ssd-Mobilenetv2

- Dengan Ssd Lite-MobileNetV2 Menggunakan Raspberry Pi Untuk Keamanan Rumah Secara Real-Time,” 2023, [Online]. Available: https://eprints.untirta.ac.id/id/eprint/30293%0Ahttps://eprints.untirta.ac.id/30293/1/Aan_Mulana_3332160067_Fulltext.pdf
- [3] M. R. Daffa Ulhaq, M. A. Zaidan, and D. Firdaus, “Pengenalan Ekspresi Wajah Secara Real-Time Menggunakan Metode SSD Mobilenet Berbasis Android,” *J. Technol. Informatics*, vol. 5, no. 1, pp. 48–52, 2023, doi: 10.37802/joti.v5i1.387.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *IEEE Xplore*, 2016, doi: <https://doi.org/10.1109/CVPR.2016.90>.
- [5] A. Raup, W. Ridwan, Y. Khoeriyah, S. Supiana, and Q. Y. Zaqiah, “Deep Learning dan Penerapannya dalam Pembelajaran,” *JiIP - J. Ilm. Ilmu Pendidik.*, vol. 5, no. 9, pp. 3258–3267, 2022, doi: 10.54371/jiip.v5i9.805.
- [6] P. A. Nugroho, I. Fenriana, and R. Arijanto, “Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) Pada Ekspresi Manusia,” *Algor*, vol. 2, no. 1, pp. 12–21, 2020.
- [7] M. Sobron and Lubis, “Implementasi Artificial Intelligence Pada System Manufaktur Terpadu,” *Semin. Nas. Tek. UISU*, vol. 4, no. 1, pp. 1–7, 2021, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/semnastek/article/view/4134>
- [8] I. Yuni Wulandari, N. Indroasyoko, R. Mudia Alti, Y. N. Asri, and R. Hidayat, “Pengenalan Sistem Deteksi Objek untuk Anak Usia Dini Menggunakan Pemrograman Python,” *Remik*, vol. 6, no. 4, pp. 664–673, 2022, doi: 10.33395/remik.v6i4.11772.
- [9] K. Falah, M. Gustiana H, and U. Ungkawa, “Karakteristik Metode Mobilenet-SSD Dengan Pre-Trained Model Mobilenet Untuk Objek Bergerak,” *Pros. Disem. FTI Ganjil*, vol. X, no. X, pp. 1–13, 2022.
- [10] I. B. Pakpahan and I. C. Dewi, “Pendeteksian Lubang Pada Jalanan Menggunakan Metode SSD-MobileNet,” *IJEIS (Indonesian J. Electron. Instrum. Syst.)*, vol. 11, no. 2, p. 213, 2021, doi: 10.22146/ijeis.60157.
- [11] K. Aggarwal *et al.*, “Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning,” *Iraqi J. Comput. Sci. Math.*, vol. 3, no. 1, pp. 115–123, 2022, doi: 10.52866/ijcsm.2022.01.01.013.
- [12] N. Nufus *et al.*, “Sistem Pendeteksi Pejalan Kaki Di Lingkungan Terbatas Berbasis SSD MobileNet V2 Dengan Menggunakan Gambar 360° Ternormalisasi,” *Pros. Semin. Nas. Sains Teknol. dan Inov. Indones.*, vol. 3, no. November, pp. 123–134, 2021, doi: 10.54706/senastindo.v3.2021.123.
- [13] J. I. Komputasi, V. No, M. Ssd, V. Mobilenet, and S. Model, “Pembuatan Aplikasi Deteksi Objek Menggunakan TensorFlow Object Detection API dengan Memanfaatkan SSD MobileNet V2 Sebagai Model Pra - Terlatih,” *J. Ilm. Komputasi*, vol. 19, no. 3, pp. 421–430, 2020, doi: 10.32409/jikstik.19.3.68.
- [14] R. A. H. ANI, “Real Time Object Detection and Recognition Basedon Deep Learning,” vol. 11, no. 01, pp. 141–142, 2022.
- [15] Van-Sonvu and N. D. Nam, “Application of MobileNet-SSD Deep Neural Network for Real-Time Object Detection and Lane Tracking on an Autonomous Vehicle,” *ResearchGate*, 2022, doi: http://dx.doi.org/10.1007/978-3-030-91892-7_53.