Pengembangan Aplikasi Kontrol Volume Audio Menggunakan Gesture Tangan Berbasis Mediapipe dan Opencv

Mustofa Raul Kurnia[1*], Abdul Aziz Muktadir[1], Ferdiansyah[1], Muhammad Khaerul Umami[1]

[1] Program Studi Informatika, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
Jl. Gatot Subroto No.8, Cimone, Kec. Karawaci, Kota Tangerang, Banten, Indonesia

e-mail korespondensi: mustofaraul78@gmail.com

Abstrak - Perkembangan teknologi computer vision memungkinkan hadirnya metode interaksi baru yang lebih intuitif tanpa memerlukan perangkat input fisik. Penelitian ini bertujuan mengembangkan aplikasi kontrol volume audio berbasis gesture tangan menggunakan MediaPipe dan OpenCV sebagai solusi alternatif terhadap pengaturan volume secara manual. Metode penelitian yang digunakan adalah Research and Development (R&D) melalui tahap analisis kebutuhan, perancangan, implementasi, dan pengujian teknis. Sistem dikembangkan untuk mengenali empat jenis gesture utama, yaitu *jempol* untuk menaikkan volume, *telunjuk* untuk menurunkan volume, *5 jari* untuk melakukan mute toggle, dan *peace* untuk fitur unmute. Berdasarkan pengujian akurasi, sistem mampu mengenali gesture secara real-time dengan tingkat keberhasilan yang stabil pada kondisi pencahayaan normal. Selain itu, aplikasi dapat berjalan tanpa perangkat keras tambahan dan memberikan respons perubahan volume yang konsisten. Hasil penelitian menunjukkan bahwa aplikasi Gesture Audio Controller ini layak digunakan sebagai metode alternatif kontrol audio yang lebih praktis dan modern.

Kata Kunci : Gesture tangan; MediaPipe; OpenCV; Kontrol volume; Aplikasi R&D

Abstract *- The advancement of computer vision technologies has enabled new forms of interaction that are more intuitive and do not require physical input devices. This study aims to develop a hand-gesture-based audio volume control application using MediaPipe and OpenCV as an alternative to conventional manual volume adjustment. The research employs a Research and Development (R&D) methodology consisting of needs analysis, system design, implementation, and technical evaluation. The system is designed to recognize four primary gestures: thumb to increase volume, index finger to decrease volume, five-fingers for mute toggle, and peace gesture for unmute. Based on accuracy testing, the application can recognize gestures in real time with stable performance under normal lighting conditions. Furthermore, the system operates without additional hardware and provides consistent volume adjustment responses. The results indicate that the Gesture Audio Controller is feasible to be used as a practical and modern alternative for audio control.*

Keywords : Hand gesture; MediaPipe; OpenCV; Volume control; R&D application

1. Pendahuluan

Perkembangan teknologi *computer vision* telah mendorong munculnya berbagai bentuk interaksi manusia–komputer (*human–computer interaction*) yang semakin natural dan intuitif. Salah satu pendekatan yang banyak dikembangkan adalah penggunaan gesture tangan sebagai metode input tanpa sentuhan (*touchless interaction*), di mana pengguna dapat mengendalikan sistem komputer tanpa menggunakan perangkat input fisik seperti mouse atau keyboard. Pendekatan ini dinilai lebih fleksibel dan praktis, terutama pada kondisi tertentu seperti saat presentasi, aktivitas multimedia, atau ketika pengguna tidak memungkinkan untuk menyentuh perangkat secara langsung.

Kemajuan signifikan dalam bidang pengenalan gesture tangan terjadi dengan diperkenalkannya MediaPipe, sebuah *framework computer vision* yang dikembangkan oleh Google Research untuk membangun *perception pipeline* secara efisien dan *real-time* (Lugaresi et al., 2019). MediaPipe menyediakan berbagai solusi siap pakai yang dapat dijalankan pada perangkat umum dengan kebutuhan komputasi yang relatif rendah. Salah satu modul utamanya adalah MediaPipe Hands, yang mampu mendeteksi dan melacak hingga 21 titik landmark tangan secara *real-time* hanya dengan menggunakan kamera RGB biasa (Zhang et al., 2020).

Pemanfaatan MediaPipe Hands dan OpenCV telah diterapkan dalam berbagai aplikasi interaksi berbasis gesture. (Chandhan et al., 2023) mengembangkan sistem *Air Canvas* yang memungkinkan pengguna menggambar di udara dengan memanfaatkan deteksi gesture tangan berbasis MediaPipe dan OpenCV. Selain itu, (Khanum & B, 2022) menunjukkan bahwa gesture tangan dapat digunakan secara efektif untuk mengendalikan presentasi sebagai alternatif antarmuka konvensional, sehingga meningkatkan kenyamanan interaksi pengguna. MediaPipe juga banyak digunakan dalam penelitian pengenalan gesture tangan berbasis visi komputer modern. Penelitian yang

Copyright (c) 2026 Mustofa Raul Kurnia, Abdul Aziz Muktadir, Ferdiansyah, Muhammad Khaerul Umami
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

dipublikasikan oleh (Nur Budiman et al., 2022) menunjukkan bahwa MediaPipe mampu mengenali gestur Gerakan jari untuk mengontrol volume. Pendekatan serupa juga diterapkan pada penelitian *multimodal gesture recognition* yang mengombinasikan MediaPipe dengan analisis fitur spasio-temporal untuk meningkatkan keandalan pengenalan gesture tangan (Cao et al., 2024).

Dalam konteks pengendalian audio, gesture tangan juga mulai diterapkan sebagai mekanisme kontrol volume. (Datar & Jain, 2024.) mengembangkan sistem *wireless music control* berbasis gesture tangan menggunakan MediaPipe dan OpenCV, yang menunjukkan bahwa pengaturan volume audio dapat dilakukan secara dinamis berdasarkan pergerakan tangan pengguna. Pendekatan serupa juga ditunjukkan oleh (Vijayan et al., 2025) melalui pengembangan sistem kontrol volume pada aplikasi piano virtual dua dimensi, yang menegaskan potensi gesture tangan sebagai metode kontrol audio yang intuitif dan responsif.

Meskipun demikian, implementasi aplikasi kontrol volume audio berbasis gesture tangan yang bersifat sederhana, ringan, dan mudah diimplementasikan pada sistem desktop masih relatif terbatas. Beberapa penelitian sebelumnya berfokus pada aplikasi tertentu atau sistem dengan kompleksitas yang cukup tinggi, sehingga kurang praktis untuk penggunaan sehari-hari. Oleh karena itu, diperlukan pengembangan aplikasi kontrol audio berbasis gesture tangan yang mudah diakses, berjalan secara *real-time*, dan tidak memerlukan perangkat keras tambahan.

Kebaruan (*novelty*) dari penelitian ini terletak pada pengembangan aplikasi Gesture Audio Controller yang memanfaatkan MediaPipe Hands dan OpenCV untuk mengontrol volume audio komputer menggunakan empat gesture utama, yaitu jempol, telunjuk, lima jari, dan *peace*. Aplikasi ini diintegrasikan langsung dengan sistem kontrol audio pada komputer dan dilengkapi mekanisme stabilisasi gesture untuk meningkatkan konsistensi serta akurasi deteksi.

Berdasarkan latar belakang tersebut, permasalahan penelitian ini dirumuskan sebagai berikut:

1. Bagaimana merancang aplikasi kontrol volume audio berbasis gesture tangan menggunakan MediaPipe dan OpenCV?
2. Bagaimana mendeteksi gesture tangan secara *real-time* dengan akurasi yang baik?
3. Bagaimana mengintegrasikan gesture tangan dengan sistem pengaturan volume audio pada komputer?
4. Bagaimana hasil pengujian teknis sistem, khususnya tingkat akurasi pengenalan gesture?

Penelitian ini bertujuan untuk mengembangkan aplikasi Gesture Audio Controller yang mampu mengenali gesture tangan secara *real-time* dan mengontrol volume audio secara otomatis. Selain itu, penelitian ini bertujuan untuk mengevaluasi performa teknis sistem guna memastikan kelayakannya sebagai metode kontrol audio alternatif berbasis gesture tangan.

2. Metode Penelitian

Penelitian ini menggunakan metode Research and Development (R&D) dengan mengacu pada model pengembangan yang dikemukakan oleh Sugiyono (2019). Metode R&D dipilih karena penelitian ini berfokus pada proses pengembangan produk berupa aplikasi *Gesture Audio Controller*, mulai dari perancangan sistem hingga pengujian teknis untuk menghasilkan produk yang fungsional dan dapat digunakan. Model R&D yang digunakan terdiri dari tujuh tahapan, yaitu: (1) potensi dan masalah, (2) pengumpulan data, (3) desain produk, (4) pembuatan produk, (5) uji coba produk, (6) revisi produk, dan (7) produk akhir.

2.1. Potensi dan Masalah

Tahap awal penelitian dimulai dengan mengidentifikasi potensi dan permasalahan pada sistem pengaturan volume audio konvensional. Pengaturan volume pada komputer umumnya masih mengandalkan perangkat input fisik seperti keyboard dan mouse, sehingga kurang mendukung interaksi manusia–komputer yang bersifat natural. Dalam konteks *human–computer interaction* (HCI), diperlukan metode kontrol yang lebih intuitif, fleksibel, dan tidak bergantung pada sentuhan langsung.

Perkembangan teknologi *computer vision* memungkinkan gesture tangan digunakan sebagai metode interaksi tanpa sentuhan dalam pengendalian perangkat komputer, sehingga pengguna dapat berinteraksi dengan sistem secara lebih natural dibandingkan antarmuka konvensional(Gupta et al., 2023). Pemanfaatan gesture tangan juga telah diterapkan untuk mengendalikan fungsi perangkat komputer seperti kursor dan navigasi sistem, yang menunjukkan bahwa interaksi berbasis gesture dapat menggantikan peran perangkat input tradisional (Juzairi Safitli & Huda Aminuddin, 2025)

Selain itu, penggunaan MediaPipe dan OpenCV memungkinkan pendeteksian gesture tangan secara *real-time* dengan kebutuhan komputasi yang relatif rendah, sehingga cocok untuk diimplementasikan pada sistem desktop tanpa perangkat keras tambahan (Dhamodaran et al., 2024)
Berdasarkan kondisi tersebut, pengembangan sistem kontrol volume audio berbasis gesture tangan memiliki potensi untuk menjadi alternatif interaksi yang lebih modern, praktis, dan efisien dibandingkan metode pengaturan volume konvensional.

2.2. Pengumpulan Data

Tahap pengumpulan data dilakukan melalui uji coba gesture tangan secara langsung menggunakan aplikasi Gesture Audio Controller yang dikembangkan. Pengumpulan data bertujuan untuk memperoleh data deteksi gesture tangan yang digunakan dalam pengujian kinerja sistem, khususnya untuk menghitung tingkat akurasi pengenalan gesture.
Pendekatan pengujian langsung (*experimental testing*) umum digunakan dalam penelitian berbasis *computer vision* untuk mengevaluasi performa sistem pengenalan gesture secara objektif (Furqan Rasyid et al., 2023).Data dikumpulkan dengan melakukan pengujian terhadap empat jenis gesture tangan, yaitu gesture jempol, telunjuk, lima jari, dan *peace*. Setiap gesture dilakukan secara berulang di depan kamera dan diproses oleh sistem menggunakan MediaPipe Hands untuk mendeteksi serta mengenali pola gesture tangan. Integrasi MediaPipe, OpenCV, dan algoritma Random Forest mampu menghasilkan sistem pengenalan gesture yang responsif dan andal, serta dapat diterapkan pada lingkungan perbankan untuk mendukung interaksi tanpa hambatan komunikasi(Shaikh, 2024).
Setiap percobaan pengujian dicatat sebagai deteksi benar atau deteksi salah berdasarkan kesesuaian antara gesture yang dilakukan oleh pengguna dan hasil pengenalan gesture oleh sistem. Data hasil pengujian kemudian direkapitulasi dan digunakan sebagai dasar dalam perhitungan tingkat akurasi sistem pada tahap uji coba produk.
Pengumpulan data dilakukan dalam kondisi pencahayaan normal dan jarak kamera yang relatif konstan untuk menjaga konsistensi hasil pengujian.

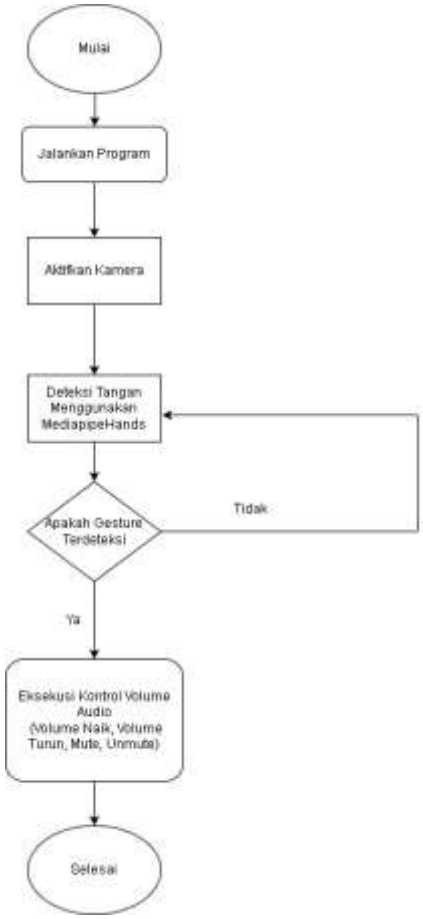
2.3. Desain Produk

Pada tahap desain produk, dilakukan perancangan alur kerja sistem dan penentuan gesture yang digunakan sebagai kontrol volume audio. Alur sistem dirancang secara sederhana dan digambarkan dalam bentuk flowchart, yang meliputi proses pengambilan input kamera, deteksi tangan, klasifikasi gesture, dan eksekusi perintah kontrol volume. Alur kerja sistem tersebut ditunjukkan pada Gambar 1.

Empat gesture tangan dirancang sebagai kontrol utama, yaitu gesture jempol untuk menaikkan volume, telunjuk untuk menurunkan volume, lima jari untuk perintah *mute*, dan gesture *peace* untuk *unmute*. Sistem juga dilengkapi mekanisme stabilisasi gesture dengan memastikan gesture terdeteksi secara konsisten dalam beberapa frame sebelum dieksekusi, guna meningkatkan keandalan dan akurasi deteksi

2.4. Pembuatan Produk

Tahap pembuatan produk dilakukan dengan mengimplementasikan desain sistem ke dalam bentuk aplikasi. Aplikasi dikembangkan menggunakan bahasa pemrograman Python dengan memanfaatkan beberapa pustaka utama, yaitu MediaPipe untuk deteksi dan pelacakan landmark tangan (Lugaresi et al., 2019), MediaPipe Hands untuk pengenalan gesture berbasis 21 titik landmark tangan (Zhang et al., 2020), serta OpenCV untuk pemrosesan citra dan tampilan antarmuka kamera (Bradski, 2000).
Proses klasifikasi gesture dilakukan berdasarkan pola jari yang terangkat dari hasil landmark tangan yang terdeteksi. Setiap gesture yang dikenali kemudian dipetakan ke fungsi kontrol volume sesuai dengan desain yang telah ditentukan. Selain itu, ditambahkan tampilan antarmuka berupa *overlay* informasi volume sebagai umpan balik visual secara *real-time* kepada pengguna.



Gambar 1. Flowchart Gesture Audio Controller

2.5. Uji Coba Produk

Uji coba produk dilakukan untuk mengetahui tingkat akurasi sistem dalam mengenali gesture tangan. Pengujian dilakukan dengan menguji empat jenis gesture, yaitu jempol, telunjuk, lima jari, dan *peace*. Setiap gesture diuji sebanyak 10 kali, sehingga total pengujian berjumlah 40 percobaan.

Setiap percobaan dicatat sebagai deteksi benar atau salah berdasarkan kesesuaian antara gesture yang dilakukan pengguna dan gesture yang dikenali oleh sistem. Tingkat akurasi sistem dihitung menggunakan persamaan sebagai berikut:

$$Akurasi = \frac{Jumlah\ deteksi\ benar}{40} \times 100\% \quad (1)$$

Pengujian dilakukan dalam kondisi pencahayaan normal dan jarak kamera yang relatif konstan untuk menjaga konsistensi hasil.

2.6. Revisi Produk

Berdasarkan hasil uji coba, dilakukan revisi terhadap sistem untuk meningkatkan stabilitas dan konsistensi deteksi gesture. Revisi meliputi penyesuaian parameter deteksi gesture, penyempurnaan mekanisme stabilisasi frame, serta perbaikan tampilan antarmuka agar lebih informatif dan mudah dipahami. Tahap revisi dilakukan secara iteratif hingga aplikasi dapat berjalan sesuai dengan tujuan pengembangan.

2.7. Produk Akhir

Tahap akhir penelitian menghasilkan aplikasi Gesture Audio Controller yang mampu mengenali gesture tangan secara real-time dan mengendalikan volume audio komputer tanpa memerlukan perangkat keras tambahan. Aplikasi ini mendukung empat fungsi utama kontrol volume melalui gesture tangan dan dapat digunakan sebagai bentuk implementasi interaksi manusia–komputer yang lebih intuitif dan modern.

3. Hasil dan Pembahasan

Pengujian dilakukan untuk mengetahui kemampuan sistem dalam mengenali gesture tangan yang digunakan sebagai kontrol volume audio. Empat jenis gesture diuji, yaitu gesture jempol, telunjuk, lima jari, dan peace. Setiap gesture diuji sebanyak 10 kali, sehingga total pengujian berjumlah 40 percobaan.

Hasil pengujian menunjukkan bahwa sistem mampu mengenali gesture tangan dengan tingkat keberhasilan yang baik. Ringkasan hasil pengujian akurasi gesture ditunjukkan pada Tabel 1

Tabel 1. Uji Gesture

No	Gesture	Jumlah Uji	Deteksi Benar	Keterangan
1	Jempol	10	7	Gestur Jempol yang terdeteksi benar berbeda 3 dari Jumlah Uji
2	Telunjuk	10	10	Tidak ada selisih perbedaan dari Jumlah Uji dan Deteksi Benar untuk Gestur Telunjuk
3	5 Jari	10	9	Gestur 5 Jari yang terdeteksi benar berbeda 1 dari Jumlah Uji
4	Peace	10	8	Gestur Peace yang terdeteksi benar berbeda 2 dari Jumlah Uji
Total		40	34	Terdapat Perbedaan 6 nilai dari gestur yang terdeteksi benar dan jumlah uji

Hasil pengujian pada Tabel 1, sistem memperoleh tingkat akurasi sebesar **85%**, yang menunjukkan bahwa gesture tangan dapat dikenali dengan baik oleh aplikasi.

$$Akurasi = \frac{34}{40} x\ 100\% = 85\%$$

Hasil pengujian akurasi gesture, aplikasi Gesture Audio Controller mampu mengenali gesture tangan secara real-time dan mengeksekusi perintah kontrol volume sesuai dengan desain sistem. Nilai akurasi sebesar 85% menunjukkan bahwa sistem dapat digunakan secara fungsional sebagai media interaksi manusia–komputer berbasis gesture. Pemanfaatan MediaPipe Hands dan OpenCV terbukti efektif dalam mendeteksi gesture tanpa memerlukan perangkat keras tambahan. Mekanisme stabilisasi gesture yang diterapkan juga berperan dalam meningkatkan konsistensi deteksi sebelum perintah kontrol volume dijalankan.

Meskipun demikian, hasil pengujian menunjukkan bahwa proses deteksi gesture masih dipengaruhi oleh kondisi pencahayaan dan posisi tangan terhadap kamera. Sistem bekerja lebih optimal pada kondisi pencahayaan normal dengan posisi tangan yang jelas menghadap kamera. Secara keseluruhan, aplikasi yang dikembangkan telah memenuhi tujuan penelitian sebagai sistem kontrol volume audio berbasis gesture tangan yang praktis dan modern.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa aplikasi Gesture Audio Controller berhasil dikembangkan menggunakan metode Research and Development (R&D). Aplikasi ini mampu mengenali gesture tangan secara real-time dan mengontrol volume audio komputer tanpa memerlukan perangkat keras tambahan. Hasil pengujian akurasi menunjukkan bahwa sistem memiliki tingkat akurasi sebesar 85%, yang menandakan bahwa aplikasi dapat berfungsi dengan baik sebagai media interaksi manusia–komputer berbasis gesture tangan. Secara keseluruhan, aplikasi yang dikembangkan telah memenuhi tujuan penelitian dan layak digunakan sebagai alternatif metode kontrol volume audio yang praktis dan modern..

Referensi

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's. In *J Softw Tools*. fttps://cir.nii.ac.jp/crid/1371415748543789317.bib?lang=en

Cao, W., Lu, P., & Cao, W. (2024). Multimodal Gesture Recognition with Spatio-Temporal Features Fusion Based on YOLOv5 and MediaPipe. *International Journal of Pattern Recognition and Artificial Intelligence*, *38*(8). https://doi.org/10.1142/S0218001424550073

Chandhan, T. H., Kumar, K., Raj, N., Nanda, N., Reddy, K., & Zabeeulla, M. (20203). Air Canvas: Hand Tracking Using OpenCV and MediaPipe. In *Science & Technology*. Retrieved https://ssrn.com/abstract=4527784

Datar, Y., & Jain, R. (2024). Wireless Music Control Using Gesture Implementation. In *International Journal of Medical Science and Dental Research*. Retrieved www.ijmsdr.org

Dhamodaran, S., Phukan, P. P., Singh, M., & Nandakum, S. (2024). Implementation of Hand Gesture Recognition using OpenCV. *International Journal of Research Publication and Reviews*, *5*(5), 8039–8050. https://doi.org/10.55248/gengpi.5.0524.1324

Furqan Rasyid, M., Mustafa, M., Mahersatillah, A., Rizal, M., Mushaf, M., & Arifin, A. (2023). Deteksi Mata di Video Smartphone Menggunakan Mediapipe Python. *JOINTECS (Journal of Information Technology and Computer Science)*, *8*, 49–56. https://doi.org/10.31328/jointecs.v8i2.4562

Gupta, A., Chawla, N., Jain, R., Thakur, N., & Devi, A. (2023). *Gesture-Based Touchless Operations: Leveraging MediaPipe and OpenCV*.

Juzairi Safitli, M., & Huda Aminuddin, F. (2025). Controlling Computer Device Cursors Using Hand Gestures by Utilizing OpenCV. In *Intellecta : Journal of Artificial Intelligence Citation* (Vol. 1, Number 1).

Khanum, H., & B, P. H. (2022). Smart Presentation Control by Hand Gestures Using Computer Vision and Google's Mediapipe. *International Research Journal of Engineering and Technology*. www.irjet.net

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). *MediaPipe: A Framework for Building Perception Pipelines*. http://arxiv.org/abs/1906.08172

Nur Budiman, S., Lestanti, S., Marselius Evvandri, S., & Kartika Putri, R. (2022). Pengenalan Gestur Gerakan Jari Untuk Mengontrol Volume Di Komputer Menggunakan Library Opencv Dan Mediapipe. *Antivirus : Jurnal Ilmiah Teknik Informatika*, *16*(2), 223–232. https://doi.org/10.35457/antivirus.v16i2.2508

Shaikh, M. T. (2024). *Implementation Of: Real-Time Hand Gesture Recognition System For Deaf Mute Friendly Banking Using Mediapipe And Opencv*.

Sugiyono. (2019). Metode penelitian dan pengembangan (Research and Development). Bandung: Alfabeta.

Vijayan, R., Mareeswari, V., Sarathi, G., & Sathya Nikethan, R. V. (2025). Hand Gesture-controlled 2D Virtual Piano with Volume Control. *International Journal of Information Technology and Computer Science*, *17*(5), 12–24. https://doi.org/10.5815/ijitcs.2025.05.02

Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C.-L., & Grundmann, M. (2020). *MediaPipe Hands: On-device Real-time Hand Tracking*. http://arxiv.org/abs/2006.10214