

Sign Catcher: Inovasi Penerjemah Bahasa Isyarat

Muhammad Zacky Al Giffari
Department of Data Science
School of Computing Telkom
University Bandung, Indonesia
zackyalgiffari@student.telkomu
niversity.ac.id

Nabila Janatri Iswibowo
Department of Data Science
School of Computing Telkom
University Bandung, Indonesia
nabilajanatri@student.telkomuni
versity.ac.id

Muhammad Zaki Rabbani
Department of Data Science
School of Computing Telkom
University Bandung, Indonesia
zakirab@student.telkomuniversi
ty.ac.id

Abstract— Penelitian ini bertujuan untuk mengembangkan aplikasi penerjemah bahasa isyarat secara real-time menggunakan teknologi visi komputer dan pembelajaran mesin. Metode yang digunakan meliputi pengumpulan data gestur tangan, pra-pemrosesan data, deteksi dan anotasi tangan menggunakan MediaPipe, serta penerjemahan gestur menjadi karakter alfabet menggunakan Support Vector Machine (SVM). Model berhasil mencapai akurasi 98% dalam mengenali 26 gestur tangan yang mewakili alfabet bahasa isyarat Amerika. Penelitian ini menunjukkan potensi penerapan teknologi ini dalam meningkatkan aksesibilitas komunikasi bagi komunitas tuna rungu..

Kata kunci : Bahasa Isyarat, Deteksi Gestur, MediaPipe, Support Vector Machine (SVM), *computer vision*, pembelajaran mesin.

I. INTRODUCTION

Dengan kemajuan dalam pembelajaran mesin dan teknologi *computer vision*, interaksi manusia-komputer telah berevolusi melampaui perangkat input tradisional seperti keyboard dan mouse. Salah satu bidang penelitian HCI yang menjanjikan adalah penerjemahan gestur tangan ke dalam teks, yang memiliki potensi signifikan untuk aplikasi dalam bidang aksesibilitas. Laporan ini memperkenalkan sebuah aplikasi yang dirancang untuk mendeteksi dan menerjemahkan gerakan tangan ke dalam karakter abjad secara real-time[1].

Laporan ini memperkenalkan sebuah aplikasi inovatif bernama *sign catcher* yang memanfaatkan MediaPipe Google untuk deteksi dan anotasi gerakan tangan, yang dikombinasikan dengan algoritma *Support Vector Machine* (SVM) untuk menerjemahkan gerakan-gerakan ini menjadi karakter alfabet. Dengan menggunakan model pre-trained dari MediaPipe, sistem ini dapat mendeteksi dan melacak posisi tangan dengan presisi tinggi, sementara algoritma SVM digunakan untuk mengidentifikasi pola dari vektor fitur yang dihasilkan oleh anotasi gerakan tangan. Kombinasi teknologi ini memungkinkan aplikasi untuk menerjemahkan gerakan tangan kompleks menjadi teks yang dapat dimengerti.

II. METODE

Metode penelitian ini dimulai dengan Pengumpulan Data, di mana gambar atau video dari berbagai gerakan tangan yang mewakili setiap karakter alfabet direkam dalam kondisi yang berbeda seperti sudut pandang, ukuran tangan, dan pencahayaan. Data ini kemudian diberi label sesuai dengan alfabet yang diwakili. Pra-pemrosesan Data melibatkan normalisasi koordinat landmark tangan untuk konsistensi skala dan augmentasi data dengan rotasi, skala, dan

pencahayaan untuk meningkatkan variasi dan generalisasi model.

Selanjutnya, Deteksi dan Anotasi Tangan dilakukan menggunakan MediaPipe Google untuk melacak 21 titik kunci pada tangan dengan presisi tinggi. Koordinat landmark ini diubah menjadi vektor fitur yang menggambarkan posisi relatif dari setiap titik kunci. Vektor fitur ini kemudian dimasukkan ke dalam algoritma Support Vector Machine (SVM). SVM dilatih dengan dataset yang terlabel untuk mengenali pola dari vektor fitur dan mengklasifikasikan gerakan tangan menjadi karakter alfabet yang sesuai. tangan menjadi teks secara akurat dalam waktu nyata.

A. Dataset

Data yang dipakai terdiri dari serangkaian foto yang memperlihatkan gestur tangan untuk setiap huruf dalam alfabet, sesuai dengan standar Bahasa Isyarat Amerika (American Sign Language). Setiap huruf diwakili oleh gambar yang diambil dari sudut pandang, kecerahan, dan perspektif berbeda, hal ini dilakukan untuk memastikan algoritma yang dilatih dapat mengenali gestur tangan dari berbagai kondisi. Setiap huruf dilatih dengan kurang lebih 8000 foto.

B. Mediapipe

Hands

MediaPipe Hands merupakan modul dalam framework MediaPipe yang memanfaatkan teknologi machine learning untuk deteksi dan pelacakan tangan secara real-time. Modul ini beroperasi dengan mengidentifikasi 21 landmark penting pada tangan, termasuk ujung jari, buku jari, dan pergelangan tangan. Proses deteksi dimulai dengan penggunaan model jaringan saraf untuk mendeteksi area palm, yang kemudian diikuti oleh model pelacak untuk menentukan posisi tepat dari setiap landmark tangan. Landmark ini memberikan informasi detail tentang orientasi, bentuk, dan posisi tangan dalam ruang tiga dimensi. MediaPipe Hands melakukan normalisasi koordinat landmark untuk memastikan konsistensi dalam representasi tangan, serta mengaplikasikan teknik pra-pemrosesan untuk meningkatkan kestabilan dan akurasi deteksi. Kecepatan dan efisiensi menjadi fokus utama, sehingga aplikasi yang menggunakan MediaPipe Hands dapat memberikan respons dalam waktu nyata dengan latency yang rendah, memungkinkan interaksi yang lebih natural antara manusia dan komputer dalam berbagai konteks, seperti game, augmented reality, dan aplikasi bantuan aksesibilitas.

C. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah salah satu algoritma klasifikasi yang paling banyak digunakan dalam machine learning. SVM bekerja dengan mencari hyperplane terbaik yang memisahkan data dalam ruang fitur, dengan memaksimalkan margin antara kelas yang berbeda. Margin ini adalah jarak antara hyperplane dan titik-titik data terdekat dari masing-masing kelas, yang disebut sebagai support vectors. SVM mampu menangani data yang tidak linear dengan menerapkan fungsi kernel, yang memetakan data ke ruang fitur yang lebih tinggi untuk meningkatkan pemisahan antar kelas. Parameter regularisasi dalam SVM mengontrol kompleksitas model untuk menghindari overfitting pada data pelatihan. SVM telah diterapkan dalam berbagai bidang seperti pengenalan pola, analisis citra, dan bioinformatika, di mana klasifikasi akurat dan robust dari data merupakan kebutuhan utama. Dalam konteks aplikasi deteksi dan klasifikasi gerakan tangan, SVM digunakan untuk mempelajari pola dari vektor fitur yang dihasilkan oleh MediaPipe Hands, sehingga dapat mengenali dan menerjemahkan gerakan tangan menjadi karakter alfabet dengan tingkat akurasi yang tinggi dalam waktu nyata.

III. HASIL DAN PEMBAHASAN

Model pengenalan gestur tangan yang dikembangkan dalam penelitian ini menunjukkan hasil yang sangat baik dalam mengenali 26 gestur tangan yang mewakili alfabet bahasa isyarat Amerika. Model ini mencapai akurasi sebesar 98%, menunjukkan kemampuannya yang handal dalam mengklasifikasikan gestur tangan dengan benar..

A. Pengumpulan dan Persiapan Data

Data yang digunakan dalam penelitian ini adalah gambar tangan yang membentuk huruf-huruf alfabet dalam bahasa isyarat. Data ini dikumpulkan dan diatur dalam folder terpisah untuk setiap huruf, seperti yang ditunjukkan pada poster, terdapat lebih dari 8000 gambar untuk setiap huruf isyarat (A-Z).

B. Ekstraksi Ciri

Kode menggunakan library MediaPipe untuk mendeteksi dan mengekstrak koordinat landmark tangan dari gambar. Landmark ini merupakan titik-titik kunci pada tangan yang mewakili posisi sendi dan ujung jari. Koordinat landmark ini kemudian diubah menjadi vektor ciri yang akan digunakan sebagai input untuk model pembelajaran mesin.

C. Pelatihan Model

Model SVM dilatih menggunakan data latih yang telah disiapkan. SVM akan mempelajari pola hubungan antara vektor ciri landmark tangan dengan huruf alfabet yang sesuai. Pada tahapan ini, model akan mengoptimalkan parameter internalnya agar mampu mengklasifikasikan data baru dengan akurat.

D. Evaluasi Model

Setelah model selesai dilatih, performanya dievaluasi menggunakan data uji. Data uji ini merupakan sekumpulan data yang belum pernah dilihat oleh model sebelumnya. Dengan mengukur akurasi prediksi model pada data uji, kita

dapat mengetahui seberapa baik model dapat digeneralisasi untuk mengenali huruf alfabet dalam bahasa isyarat dari gambar baru.

E. Penyimpanan Model

Model SVM yang telah dilatih disimpan ke dalam file 'svm_model.pkl' menggunakan library pickle. Model ini dapat digunakan kembali di kemudian hari tanpa perlu melakukan pelatihan ulang.

IV. ANALISIS

Hasil akurasi 94% yang dicapai oleh model ini merupakan indikator kuat dari kemampuannya dalam mengenali gestur tangan secara akurat dan real-time. Namun, akurasi tinggi ini perlu dianalisis lebih lanjut untuk memahami faktor-faktor yang berkontribusi terhadap performa model dan potensi area perbaikan.

A. Optimasi Model

Selain optimasi model, teknik pra-pemrosesan yang baik juga berperan penting dalam mencapai akurasi tinggi. Ini menunjukkan bahwa pemilihan dan penerapan algoritma SVM yang tepat, serta penyetelan parameternya, memiliki dampak signifikan terhadap kinerja model. Analisis lebih lanjut dapat dilakukan untuk mengeksplorasi apakah ada ruang untuk optimasi lebih lanjut dalam model SVM, seperti penggunaan kernel yang berbeda atau penyesuaian parameter regularisasi.

B. Teknik Pra-pemrosesan

Normalisasi koordinat landmark tangan dan augmentasi data dengan rotasi, skala, dan pencahayaan adalah beberapa teknik yang digunakan dalam penelitian ini. Analisis lebih lanjut dapat dilakukan untuk mengevaluasi efektivitas masing-masing teknik pra-pemrosesan dan mengidentifikasi teknik tambahan yang dapat meningkatkan performa model.

C. Evaluasi Lebih Lanjut

Akurasi 94% yang dilaporkan didasarkan pada dataset tertentu. Penting untuk melakukan evaluasi lebih lanjut dengan dataset yang lebih beragam dan menantang untuk mengukur kemampuan generalisasi model. Uji coba dengan pengguna yang berbeda-beda, termasuk mereka dengan variasi ukuran tangan, warna kulit, dan gaya bahasa isyarat,

akan memberikan pemahaman yang lebih komprehensif tentang kinerja model di dunia nyata.

Secara keseluruhan, analisis mendalam terhadap hasil penelitian ini akan memberikan pemahaman yang lebih baik tentang faktor-faktor yang mempengaruhi kinerja model pengenalan gestur tangan. Informasi ini akan sangat berharga dalam mengarahkan pengembangan lebih lanjut, baik dalam hal optimasi model, teknik pra-pemrosesan, maupun pengumpulan data yang lebih representatif.

V. KESIMPULAN

Penelitian ini telah berhasil mengembangkan model pengenalan gestur tangan untuk alfabet bahasa isyarat Amerika menggunakan kombinasi MediaPipe Hands dan Support Vector Machine (SVM). Model ini mencapai akurasi 98%, mendemonstrasikan potensi penerapan teknologi visi komputer dan pembelajaran mesin dalam meningkatkan aksesibilitas komunikasi bagi komunitas tuna rungu.

Meskipun demikian, penelitian ini juga mengidentifikasi beberapa area yang memerlukan penelitian lebih lanjut. Optimasi model, teknik pra-pemrosesan, dan pengumpulan data yang lebih beragam merupakan langkah-langkah penting untuk meningkatkan kinerja dan keandalan model. Selain itu, evaluasi lebih lanjut terhadap kemampuan generalisasi model dan performa real-time akan memberikan informasi berharga untuk pengembangan aplikasi yang lebih luas.

Secara keseluruhan, penelitian ini memberikan kontribusi signifikan dalam bidang pengenalan gestur tangan dan membuka peluang untuk pengembangan teknologi yang lebih inklusif dan bermanfaat bagi masyarakat.

REFERENSI

- [1] M. Oudah, A. Al-Naji, and J. Chahl, "Hand Gesture Recognition Based on Computer Vision: A Review of Techniques," *Journal of Imaging*, vol. 6, no. 8. MDPI, Jul. 01, 2020. doi: 10.3390/JIMAGING6080073.