

Pengenalan Bahasa Isyarat Dengan Menggunakan Teknik Augmentasi Gambar Menggunakan Convolutional Neural Network (CNN)

Achareeya Wicaksa Putra Pribadi

E-mail : 20081010194@student.upnjatim.ac.id

Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur

(Naskah masuk: dd mmm yyyy, diterima untuk diterbitkan: dd mmm yyyy)

Abstrak

Penelitian ini bertujuan mengembangkan sistem pengenalan gestur tangan dalam bahasa isyarat menggunakan *Convolutional Neural Network (CNN)* dengan fokus pada peningkatan akurasi melalui augmentasi gambar. Tantangan utama yang diatasi adalah variasi kompleks dalam bahasa isyarat. Metodologi eksperimental digunakan dengan desain yang memastikan kontrol optimal terhadap variabel. Pengumpulan data melibatkan partisipasi komunitas penutur bahasa isyarat dengan representasi gerakan yang beragam. Proses pelatihan model CNN melibatkan pengoptimalan dan preprocessing data. Pengujian model menggunakan dataset independen, dengan analisis menggunakan metrik confusion matrix. Selain itu, penelitian ini mengembangkan metode augmentasi gambar yang canggih. Hasilnya menunjukkan kemampuan model dalam prediksi gestur tangan secara waktu nyata. Kesimpulan menyatakan kontribusi penelitian terhadap efektivitas komunikasi bagi komunitas tuna rungu. Saran meliputi pengembangan lebih lanjut, peningkatan augmentasi gambar, kolaborasi dengan komunitas, evaluasi dalam konteks nyata, dan keterlibatan ahli bahasa isyarat.

Kata Kunci: *Convolutional Neural Network (CNN), preprocessing data, prediksi gestur tangan, bahasa isyarat.*

1. PENDAHULUAN

Penelitian ini mengenai pengembangan sistem pengenalan gestur tangan dalam bahasa isyarat menggunakan Convolutional Neural Network (CNN). Bahasa isyarat memiliki peran krusial dalam memungkinkan komunikasi bagi orang bisu, namun tantangan utama yang dihadapi adalah akurasi pengenalan gestur tangan, terutama akibat variasi kompleks dalam bahasa isyarat. Meskipun telah ada upaya menggunakan CNN, tingkat akurasi masih menjadi isu utama. Oleh karena itu, penelitian ini bertujuan untuk meningkatkan akurasi pengenalan gestur tangan, terutama melalui penggunaan teknik augmentasi gambar. Teknologi ini diharapkan dapat memberikan kontribusi signifikan dalam meningkatkan efektivitas komunikasi dan aksesibilitas bagi orang bisu. Penelitian akan melibatkan serangkaian langkah, mulai dari pengumpulan data hingga integrasi dengan situasi dunia nyata, dengan fokus pada pengembangan metode augmentasi gambar yang canggih. Dengan demikian, penelitian ini diharapkan dapat memberikan solusi konkret terhadap tantangan akurasi dalam pengenalan bahasa isyarat, sekaligus meningkatkan pemberdayaan komunitas orang bisu.

2. METODOLOGI

2.1 Desain Penelitian

Penelitian ini mengadopsi desain penelitian eksperimental dengan pendekatan pra-eksperimen untuk menguji dan menganalisis efektivitas sistem pengenalan gestur tangan dalam bahasa isyarat menggunakan Convolutional Neural Network (CNN). Desain eksperimental dipilih karena memberikan kerangka kerja yang kokoh untuk mengevaluasi dampak intervensi teknologi terhadap komunikasi orang bisu.

Desain eksperimental ini mencakup serangkaian langkah yang terstruktur, mulai dari pengumpulan data hingga analisis hasil. Dengan demikian, penelitian ini dirancang untuk menghasilkan bukti empiris yang kuat terkait kinerja sistem pengenalan bahasa isyarat.

2.2 Rationale Desain Eksperimental

Desain eksperimental dipilih untuk memastikan kontrol yang optimal terhadap variabel-variabel yang dapat mempengaruhi hasil. Dengan melakukan

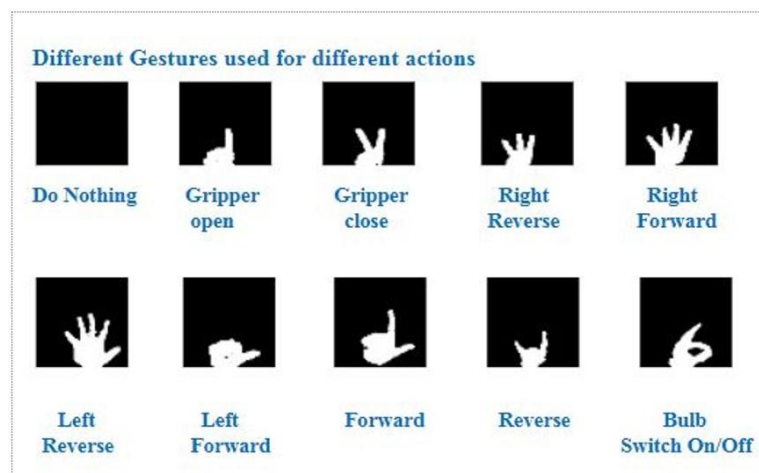
eksperimen, penelitian ini dapat memahami secara mendalam bagaimana perubahan dalam model CNN dan metode augmentasi gambar dapat meningkatkan akurasi pengenalan gestur tangan. Pendekatan ini juga memungkinkan identifikasi hubungan sebab-akibat antara intervensi dan perubahan dalam kinerja sistem.

2.3 Populasi dan Sampel

Populasi penelitian melibatkan komunitas penutur bahasa isyarat. Sampel penelitian akan dipilih secara acak dari komunitas ini dengan memperhatikan representativitas dan keberagaman gerakan, posisi tangan, serta dialek bahasa isyarat. Jumlah sampel yang memadai akan diambil untuk mencapai tingkat generalisasi yang optimal.

2.4 Pengumpulan Data

Proses pengumpulan data dilakukan melalui dua tahap utama. Pertama, melibatkan pengumpulan data gestur tangan dari partisipan dengan mewakili variasi gerakan dan bahasa isyarat. Kedua, mengumpulkan data teknis terkait hasil pengenalan gestur tangan oleh model CNN. Data teknis ini akan mencakup hasil dari setiap iterasi model selama proses pelatihan dan pengujian.



Gambar 2.1 Dataset Tangan

2.5 Preprocessing Data

Data yang terkumpul akan melalui tahap preprocessing yang komprehensif. Proses ini mencakup normalisasi data, pembersihan noise, dan ekstraksi fitur yang relevan dari gestur tangan. Langkah-langkah ini akan memastikan bahwa data yang digunakan untuk melatih dan menguji model CNN memiliki kualitas yang optimal.

2.6 Pembuatan Model CNN

Pembuatan model CNN akan melibatkan identifikasi arsitektur yang sesuai dan pengoptimalan inisialisasi bobot serta bias. Pembuatan model ini akan didasarkan pada data gestur tangan yang telah dipreprocess sebelumnya. Model akan dilatih menggunakan teknik pembelajaran yang sesuai dengan desain eksperimental.

2.7 Pengujian dan Evaluasi Model

Pengujian model CNN dilakukan dengan menggunakan dataset independen yang tidak digunakan selama pelatihan. Metrik evaluasi, seperti akurasi akan dihitung untuk mengevaluasi performa model. Analisis confusion matrix juga akan digunakan untuk mendapatkan wawasan lebih mendalam tentang kekuatan dan kelemahan model.

2.8 Pengembangan Metode Augmentasi Gambar

Selama tahap ini, metode augmentasi gambar akan diperluas dan dioptimalkan. Pendekatan ini bertujuan untuk menciptakan variasi yang lebih banyak dalam dataset pelatihan, memungkinkan model untuk lebih adaptif terhadap kompleksitas bahasa isyarat.

2.9 Integrasi dengan Situasi Dunia Nyata

Pengujian sistem dalam situasi dunia nyata melibatkan interaksi antara pengguna dan sistem dalam konteks sehari-hari. Performa sistem akan dievaluasi dalam situasi nyata, dan perbaikan atau penyesuaian model akan dilakukan berdasarkan temuan dari pengujian tersebut.

2.10 Analisis Statistik

Analisis statistik akan dilakukan untuk menginterpretasikan data hasil eksperimen. Uji hipotesis dan analisis regresi dapat digunakan untuk mengidentifikasi hubungan signifikan antara variabel-variabel yang diamati.

2.11 Etika Penelitian

Aspek etika penelitian, termasuk persetujuan partisipan, privasi data, dan perlindungan hak-hak individu, akan menjadi perhatian utama dalam seluruh tahap penelitian ini. Semua prosedur akan mematuhi pedoman etika penelitian yang berlaku.

2.12 Kerangka Pemahaman

Kerangka pemahaman penelitian ini mencakup penggabungan teori dan praktik, dengan mengacu pada landasan teori yang mendukung pemilihan metode dan langkah-langkah yang diambil dalam penelitian ini. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam mengembangkan sistem pengenalan bahasa isyarat menggunakan Convolutional Neural Network.

3. TINJAUAN PUSTAKA

3.1 Tuna Rungu

Menurut (Soemantri,1996) kata tuna rungu terdiri dari 2 kata, yaitu tuna dan rungu, yang artinya tuna berarti kurang, dan rungu berarti kurang pendengaran. Jadi tuna rungu dapat diartikan sebagai kurangnya pendengaran. Menurut mengemukakan tuna rungu adalah mereka yang kehilangan pendengaran baik sebagai maupun seluruhnya yang menyebabkan pendengarannya tidak memiliki nilai fungsional dalam kehidupan sehari-hari.

3.2 Bahasa Isyarat

3.2.1 Definisi Bahasa Isyarat

Bahasa isyarat adalah sarana berkomunikasi bagi penderita tuna rungu. Senada dengan Reynold and Mann, A. Van Uden (Lani Bunawan, 1997: 11) mengatakan bahasa isyarat adalah bahasa dengan menggunakan tangan, walaupun

dalam kenyataan, ekspresi muka dan lengan juga digunakan untuk berperan. Bahasa isyarat dikembangkan dan memiliki karakteristik sendiri di berbagai negara. Di Indonesia, bahasa isyarat yang digunakan berdasarkan pada SIBI. Ada 4 jenis bahasa isyarat dalam SIBI, yaitu:

1. Isyarat Pokok : melambangkan sebuah kata atau konsep.
2. Isyarat Tambahan : melambangkan awalan, akhiran, dan partikel (imbuhan).
3. Isyarat Bentuk : dibentuk dengan menggabungkan isyarat pokok dan isyarat tambahan.
4. Abjad Jari : dibentuk dengan jari-jari untuk mengeja huruf.

3.2.2 Sejarah Pengembangan Bahasa Isyarat

Sejarah bahasa isyarat dapat ditelusuri kembali ke berbagai budaya dan masyarakat di seluruh dunia. Perkembangannya mencerminkan kebutuhan komunikasi dalam komunitas yang menghadapi tantangan ketika menggunakan bahasa lisan. Dari perkembangan ini, muncullah berbagai bentuk bahasa isyarat yang unik di setiap wilayah.

3.2.3 Peran Teknologi dalam Pengenalan Bahasa Isyarat

Dengan kemajuan teknologi, khususnya dalam bidang Convolutional Neural Network (CNN) dan pengolahan citra, pengenalan bahasa isyarat telah menjadi fokus penelitian yang penting. Teknologi ini memungkinkan pengembangan sistem otomatis yang dapat mengartikan dan merespons isyarat bahasa isyarat dengan akurasi tinggi.

3.3 Convolutional Neural Network (CNN)

3.3.1 Konsep Dasar CNN

Convolutional Neural Networks (CNN) merupakan salah satu algoritma deep learning yang banyak dipakai dalam machine learning khususnya pada kasus citra. Di antara banyaknya model deep learning yang ada, Convolutional Neural Networks terbukti memiliki performa tinggi dalam klasifikasi citra. CNN mirip dengan neural network yang terbentuk dari neurons yang memiliki weight (bobot) dan biases (bias) yang bisa dipelajari. Sebuah CNN terdiri dari satu atau lebih lapisan yang terkoneksi penuh seperti dalam jaringan saraf multilayer standar.

Berbeda dengan neural network biasanya, lapisan pada CNN memiliki neuron yang diatur dalam 3 dimensi: lebar, tinggi dan kedalaman dimana kedalaman merujuk pada dimensi ketiga dari sebuah volume aktivasi, bukan kedalaman dari neural network penuh yang mengacu pada jumlah total dari lapisan dalam sebuah jaringan. Untuk membangun arsitektur CNN terdapat tiga tipe layer utama yaitu convolutional layer, pooling layer, dan fullyconnected layer. Dalam kasus citra, layer konvolusi dan layer-layer lain yang mengikutinya berukuran dua atau tiga dimensi.

3.3.2 Penerapan CNN pada Pengenalan Gestur Tangan

Penerapan CNN dalam pengenalan gestur tangan membawa dampak signifikan dalam meningkatkan akurasi dan kecepatan pengenalan. Model CNN dapat dilatih untuk mengenali pola-pola kompleks yang mewakili gestur tangan dalam berbagai situasi, memberikan kemampuan adaptasi yang baik.

3.4 Teknik Augmentasi Gambar dalam Pengenalan Bahasa Isyarat

Variasi kompleks dalam bahasa isyarat menimbulkan tantangan dalam pengenalan gestur tangan. Teknik augmentasi gambar diperlukan untuk memperkaya dataset pelatihan, sehingga model dapat memahami dan mengenali variasi dengan lebih baik.

3.4.1 Metode Augmentasi Gambar yang Umum Digunakan

Beberapa metode augmentasi gambar yang umum digunakan melibatkan pergeseran, rotasi, zoom, dan flipping. Pemahaman mendalam terhadap metode-metode ini penting untuk memilih strategi augmentasi yang sesuai dengan karakteristik bahasa isyarat.

3.5 Metrik Pengujian

Metrik pengujian adalah kriteria atau parameter yang digunakan untuk mengukur kinerja atau efektivitas suatu sistem. Dalam konteks pengenalan gestur tangan dalam bahasa isyarat menggunakan Convolutional Neural Network (CNN) ini menggunakan metrik pengujian confusion matrik untuk menghitung akurasi.

Confusion matrix (matriks kebingungan) adalah alat evaluasi kinerja yang umum digunakan dalam klasifikasi untuk mengukur sejauh mana model dapat memprediksi kelas dengan benar. Matriks ini memberikan gambaran tentang seberapa baik model dapat membedakan antara kelas positif dan negatif. Matriks kebingungan terdiri dari empat elemen utama :

1. True Positive (TP) : Jumlah contoh positif yang benar-benar diprediksi dengan benar oleh model.
2. True Negative (TN): Jumlah contoh negatif yang benar-benar diprediksi dengan benar oleh model.
3. False Positive (FP) : Jumlah contoh negatif yang salah diprediksi sebagai positif oleh model.
4. False Negative (FN) : Jumlah contoh positif yang salah diprediksi sebagai negatif oleh model.

Rumus menghitung akurasi dengan Confusion Matrix :

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3.6 Penelitian Terdahulu

Penelitian (S. Khetavath et al, 2023) mengusulkan sistem efisien untuk pengenalan gambar isyarat tangan menggunakan teknik pemrosesan gambar canggih. Sistem ini mencakup deteksi warna kulit, operasi morfologi, teknik Heuristic Manta-ray Foraging Optimization (HMFO) untuk pemilihan fitur, dan Adaptive Extreme Learning Machine (AELM) untuk klasifikasi. Tujuannya adalah meningkatkan akurasi dan mengurangi hasil positif palsu serta tingkat kesalahan dibandingkan dengan pendekatan konvensional. Komponen utama sistem ini melibatkan segmentasi gambar isyarat tangan, ekstraksi fitur dengan teknik seperti SIFT dan SURF, klasifikasi menggunakan berbagai algoritma termasuk CNN, GMM, dan HMM, dan pengenalan untuk mengaitkan isyarat dengan tindakan tertentu dalam aplikasi. Sistem ini berfokus pada isolasi isyarat tangan dari latar belakang untuk mengurangi gangguan potensial dan menggunakan fitur yang relevan untuk meningkatkan akurasi klasifikasi. Kelebihan dan kelemahan dari beberapa

karya terkait juga dibahas, termasuk pendekatan berbasis pembelajaran mesin, teknik HMFO dan ELM, optimasi CNN, klasifikasi berbasis deret waktu, dan pengenalan gambar isyarat tangan 3D menggunakan CNN.

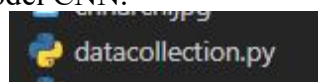
Penelitian selanjutnya yang dilakukan oleh (A. S. M. Miah, M. A. M. Hasan and J. Shin, 2023) memperkenalkan Model Deep Learning Umum yang Berbasis Grafik dengan Perhatian Multi-Cabang untuk pengenalan gerakan tangan dinamis menggunakan dataset rangka. Model ini menggabungkan fitur spasial-temporal, temporal-spasial, dan deep learning umum untuk meningkatkan akurasi. Hasil eksperimen menunjukkan bahwa model yang diusulkan memiliki kinerja yang lebih baik dibandingkan metode yang ada dalam hal akurasi dan biaya komputasi. Model tersebut mengubah koordinat sendi tangan menjadi tensor yang memberi makan jaringan saraf, memungkinkan ekstraksi fitur dan pengenalan yang lebih akurat. Penggunaan mekanisme perhatian membantu fokus pada informasi spasial dan temporal yang relevan, yang lebih meningkatkan kinerjanya. Model ini mengatasi keterbatasan metode sebelumnya yang lebih fokus pada deskriptor fitur dan pembelajaran mesin tradisional. Dengan memanfaatkan deep learning, mekanisme perhatian, dan representasi berbasis grafik dari rangkaian tangan, model ini mencapai akurasi yang lebih tinggi dalam pengenalan gerakan tangan dinamis. Cabang-cabang jaringan saraf, spasial-temporal, temporal-spatial, dan deep learning umum bekerja bersama untuk mengekstrak dan menggunakan informasi spasial, temporal, dan umum dari rangkaian tangan, memungkinkan pengenalan yang akurat dari gerakan tangan dinamis hanya dengan menggunakan informasi rangkaian tangan 3D.

4. HASIL DAN PEMBAHASAN

Hand Gesture Recognition using Convolutional Neural Networks (CNN)" ini memiliki beberapa tahapan utama yang dijelaskan sebagai berikut :

1. Pengumpulan Data

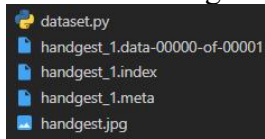
File 'Datacollection.py' digunakan untuk mengumpulkan data pelatihan dan uji. Program ini mengambil 1000 gambar dari setiap gestur dan menyimpannya sebagai set pelatihan. Gambar-gambar ini kemudian digunakan untuk melatih model CNN.



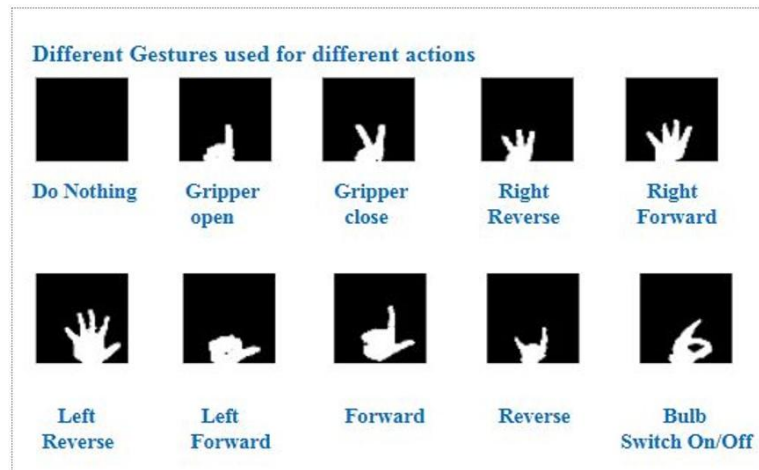
Gambar 3.1 File Datacollection.py

2. Dataset Gestur Tangan

Dataset yang digunakan untuk pelatihan mengandung berbagai gestur tangan yang merepresentasikan aksi rover yang berbeda. Setiap gestur memiliki 1000 gambar, dan mereka dapat dilihat dalam gambar yang disertakan.



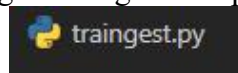
Gambar 3.2 File dataset Gestur Tangan



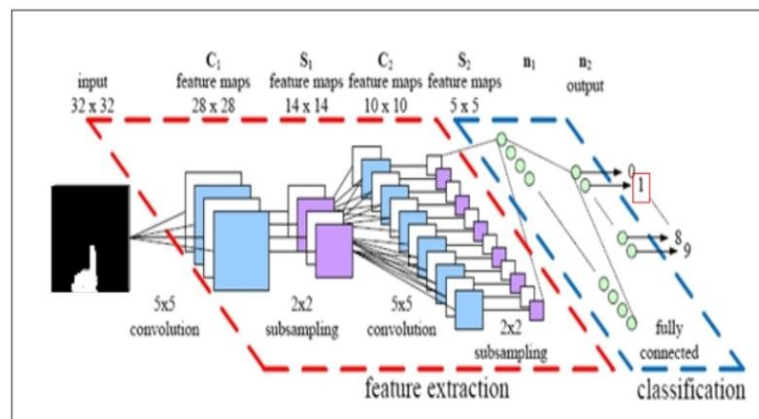
Gambar 3.3 Dataset Tangan

3. Pelatihan Model CNN

Pemodelan CNN dilakukan menggunakan file 'Traingest.py'. Model CNN dilatih dengan menggunakan gambar-gambar yang telah dikumpulkan sebelumnya. Struktur arsitektur CNN dapat dilihat dalam gambar yang disertakan ('cnnarch.jpg'). Selama pelatihan, model menyimpan informasi pembelajaran dalam file meta dan indeks yang akan digunakan pada tahap prediksi.



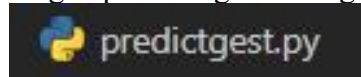
Gambar 3.4 File Traingest.py



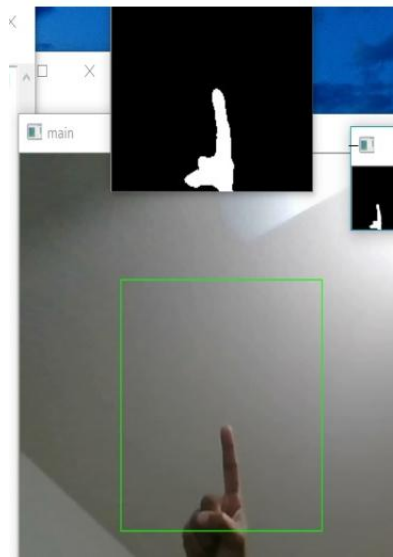
Gambar 3.5 Arsitektur CNN

4. Pengujian Model

File `Predictgest.py` digunakan untuk melakukan prediksi secara waktu nyata terhadap gestur tangan. Ada sepuluh gestur yang berbeda yang telah dilatih dan dapat dikenali. Gambar hasil prediksi waktu nyata dapat dilihat dalam gambar yang disertakan (`results.jpg`). Prediksi memberikan nilai probabilitas untuk setiap gestur. Proyek ini dapat digunakan untuk mengenali gestur tangan secara waktu nyata dan mengaitkannya dengan aksi tertentu pada rover. Prosesnya dimulai dari pengumpulan data, dilanjutkan dengan pelatihan model CNN, dan akhirnya pengujian model dengan prediksi gestur tangan secara waktu nyata.



Gambar 3.6 File Predictgest.py



Gambar 3.7 Proses Training

```
[[1.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000]
 0.000000 0.000000]]
Gesture:0
[[1.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000]
 0.000000 0.000000]]
Gesture:0
[[1.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000]
 0.000000 0.000000]]
Gesture:0
[[0.000000 0.993057 0.000011 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000]
 0.006932 0.000000]]
Gesture:1
```

Gambar 3.8 Hasil

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini fokus pada pengembangan sistem pengenalan gestur tangan menggunakan Convolutional Neural Network (CNN) dalam konteks bahasa isyarat. Dengan memanfaatkan teknologi CNN dan augmentasi gambar, penelitian

ini bertujuan meningkatkan akurasi pengenalan gestur tangan, khususnya dalam bahasa isyarat yang kompleks.

Pengumpulan data dilakukan melalui program `'Datacollection.py'`, menghasilkan dataset gestur tangan yang mencakup variasi gerakan dan bahasa isyarat. Metodologi penelitian mengadopsi desain eksperimental dengan pendekatan pra-eksperimen, memastikan kontrol yang optimal terhadap variabel-variabel yang mempengaruhi hasil.

Arsitektur CNN yang digunakan dalam penelitian ini diimplementasikan melalui `'Traingest.py'`. Proses pelatihan melibatkan pengoptimalan bobot dan bias, dengan data yang telah dipreprocess sebelumnya. Pengujian dan evaluasi model dilakukan dengan menggunakan dataset independen untuk mengukur akurasi dan kinerja model.

Penerapan augmentasi gambar menjadi elemen kunci dalam penelitian ini, dimana teknik augmentasi gambar diperluas dan dioptimalkan untuk menciptakan variasi lebih banyak dalam dataset pelatihan. Hal ini diharapkan dapat membuat model lebih adaptif terhadap kompleksitas bahasa isyarat.

Hasil pengujian menunjukkan kemampuan model untuk melakukan prediksi gestur tangan secara waktu nyata, dengan memberikan nilai probabilitas untuk setiap gestur yang diprediksi. Meskipun kesalahan dan batasan mungkin muncul, penelitian ini memberikan kontribusi dalam meningkatkan efektivitas komunikasi bagi komunitas tuna rungu.

5.2 Saran

1. Pengembangan Lebih Lanjut

Melanjutkan penelitian ini dengan fokus pada pengembangan lebih lanjut dari model, termasuk eksplorasi arsitektur CNN yang lebih canggih dan pengoptimalan parameter untuk meningkatkan akurasi.

2. Peningkatan Augmentasi Gambar

Menginvestigasi metode augmentasi gambar yang lebih canggih dan kompleks untuk menciptakan variasi yang lebih kaya dalam dataset pelatihan.

3. Kolaborasi dengan Komunitas

Melibatkan komunitas tuna rungu dalam pengembangan dan pengujian sistem untuk memastikan bahwa aplikasi yang dikembangkan benar-benar bermanfaat dan sesuai dengan kebutuhan pengguna.

4. Evaluasi dalam Konteks Nyata

Melakukan evaluasi lebih lanjut dalam situasi dunia nyata untuk mengukur efektivitas dan kehandalan sistem dalam kondisi penggunaan sehari-hari.

5. Keterlibatan Ahli Bahasa Isyarat

Menggandeng ahli bahasa isyarat dalam proses pengembangan untuk memastikan bahwa model dapat mengenali variasi bahasa isyarat dengan akurat.

Penelitian ini diharapkan menjadi landasan untuk pengembangan sistem pengenalan gestur tangan yang lebih baik, meningkatkan aksesibilitas dan komunikasi bagi komunitas tuna rungu.

6. DAFTAR RUJUKAN

- Sumantri, Sutjihati. (1996). *Psikologi Anak Luar Biasa*, Jakarta: Depdikbud
- Lani Bunawan. (1997). *Kominikasi total*. Jakarta: Departemen Pendidikan dan Kebudayaan Direktorat Jenderal Pendidikan Tinggi.
- S. Khetavath et al., "An Intelligent Heuristic Manta-Ray Foraging Optimization and Adaptive Extreme Learning Machine for Hand Gesture Image Recognition," in *Big Data Mining and Analytics*, vol. 6, no. 3, pp. 321-335, September 2023, doi: 10.26599/BDMA.2022.9020036.
- A. S. M. Miah, M. A. M. Hasan and J. Shin, "Dynamic Hand Gesture Recognition Using Multi-Branch Attention Based Graph and General Deep Learning Model, " in *IEEE Access*, vol. 11, pp. 4703-4716, 2023, doi: 10.1109/ACCESS.2023.3235368.