

BAB I

PENDAHULUAN

1.1 Latar Belakang

Manusia membutuhkan komunikasi verbal untuk terlibat dalam aktivitas sosial. Sayangnya, mereka yang mengalami gangguan bicara dan/atau pendengaran tidak dapat berinteraksi secara verbal dengan orang lain. Mereka berkomunikasi menggunakan bahasa isyarat. Bagi mereka yang mengalami kesulitan mendengar atau berbicara, bahasa isyarat telah diciptakan sebagai sarana komunikasi yang dapat diandalkan (Rahaman et al., 2024).

Komunikasi selalu memegang peranan penting dalam kehidupan manusia. Kemampuan untuk berinteraksi dengan orang lain dan mengekspresikan diri merupakan kebutuhan dasar manusia. Akan tetapi, berdasarkan pola asuh, pendidikan, masyarakat, dan sebagainya, perspektif dan cara kita berkomunikasi dengan orang lain dapat sangat berbeda dengan orang-orang di sekitar kita. Selain itu, memastikan bahwa kita dipahami sesuai dengan maksud kita (Katoch et al., 2022).

Berkomunikasi merupakan kebutuhan semua individu karena setiap individu harus berkomunikasi dengan lingkungan. Berkomunikasi juga membuat seseorang mendapat informasi sehingga dapat dijadikan acuan untuk beradaptasi. Penggunaan bahasa verbal dengan berbicara mengeluarkan suara adalah cara komunikasi individu, namun hal itu tidak dapat dilakukan saat berkomunikasi dengan individu yang memiliki keterbatasan dalam mendengar. Masih banyak yang sulit mengenal bahasa isyarat, maka solusinya adalah membuat sistem untuk klasifikasi bahasa isyarat (Iduar et al., 2021). Kelompok orang ini telah mengembangkan bahasa mereka untuk berkomunikasi di antara mereka sendiri, yang dikenal sebagai Bahasa isyarat. Bahasa isyarat menggunakan gerakan tubuh manusia visual dan gestur untuk mengekspresikan pikiran seseorang. Ada juga bagian dalam bahasa isyarat yang berbeda di mana alfabet bahasa asli diekspresikan dengan tanda tangan ejaan jari (Bora et al., 2022).

Manusia normal tidak memiliki banyak kesulitan dalam berinteraksi satu sama lain dan dapat mengekspresikan diri dengan mudah melalui ucapan, gerakan, bahasa tubuh, membaca, menulis, dan berbicara yang banyak digunakan di antara mereka. Akan tetapi, orang-orang yang mengalami gangguan bicara hanya mengandalkan bahasa isyarat, yang membuat mereka lebih sulit untuk berkomunikasi dengan mayoritas lainnya (Katoch et al., 2022).

Komunikasi yang efektif di antara manusia dilakukan dengan menggunakan gerakan. Orang menggunakan gerakan untuk mengekspresikan pikiran dan perasaan mereka, menarik perhatian, dan menekankan komunikasi. Gerakan berbeda dengan bahasa isyarat. Gerakan bergantung pada pengguna dan dapat digunakan secara acak saat berbicara dan berinteraksi, sedangkan bahasa isyarat perlu dipelajari seperti halnya bahasa lain (Wangchuk et al., 2021).

Penggunaan bahasa isyarat telah ada sejak abad ke-5 SM. Seperti yang dikatakan oleh Socrates, "Jika kita tidak memiliki suara atau lidah, dan ingin menyampaikan sesuatu satu sama lain, bukankah kita akan mencoba membuat isyarat dengan menggerakkan tangan, kepala, dan bagian tubuh kita yang lain, seperti yang dilakukan orang bisu saat ini?" Sejak abad ke-5 SM, telah ada berbagai versi bahasa isyarat. Namun, versi yang menjadi fokus dalam proyek ini adalah *American Sign Language (ASL)* (Ardiansyah et al., 2021).

Pengaruh signifikan yang mungkin dimiliki komunikasi yang jelas terhadap kehidupan mereka yang mengidentifikasi diri sebagai tuna rungu atau sulit mendengar. Bahasa-bahasa yang kaya seperti BSL dan ASL, masing-masing dengan sintaksis dan leksikonnya sendiri, sangat penting untuk membina hubungan antarpribadi dan bertukar ide. Kombinasi pembelajaran mendalam dan teknologi visi komputer telah membuka cara-cara baru untuk menutup kesenjangan komunikasi di dunia digital modern. Dengan menggabungkan kecerdasan buatan dengan kerumitan bahasa isyarat, penelitian ini berharap dapat menyediakan alat bagi orang-orang untuk mengenali, menafsirkan, dan memahami BSL dan ASL dengan mudah dan akurat (Koyineni et al., 2024).

Orang yang mengalami gangguan bicara hanya dapat bergantung pada bahasa isyarat, yang membuat mereka lebih sulit berkomunikasi dengan sebagian besar orang lainnya. Ini menunjukkan adanya kebutuhan untuk pengenalan bahasa isyarat yang dapat mengenali dan mengubah bahasa isyarat menjadi bahasa lisan atau tulisan dan sebaliknya. Pengenal semacam itu, bagaimanapun, masih terbatas, mahal, dan sulit digunakan (Katoch et al., 2022).

Bahasa isyarat adalah media komunikasi bagi penyandang tunarungu. Namun, terkadang mereka perlu berkomunikasi dengan orang normal yang tidak peka terhadap bahasa isyarat. Masalah ini biasanya diselesaikan dengan bantuan penerjemah manusia. Di India, terdapat sekitar 250 penerjemah untuk setiap satu juta orang penyandang tunarungu. Untuk menggantikan penerjemah manusia, teknologi dapat membantu membangun penerjemah mesin (Kumar et al., 2022).

Dengan menggunakan bahasa isyarat, penyandang tuli dan bisu dapat berkomunikasi dengan cara tertentu. Namun, hanya sebagian orang yang memahami bahasa isyarat. Ketika penyandang tuli atau bisu perlu berbicara di depan umum, biasanya mereka memanfaatkan bantuan seorang penerjemah. Penelitian ini bertujuan untuk mencari teknik yang dapat menyediakan penerjemah bahasa isyarat terbaik, yang berarti harus efisien, akurat, dan dapat diakses oleh semua orang (Ardiansyah et al., 2021).

Algoritma deep learning diterapkan pada solusi ini untuk menghasilkan sistem yang cepat, murah, ringan, dan mudah diterapkan, yang dapat digunakan sebagai inti dari sistem pengenalan bahasa isyarat yang lengkap (Koyineni et al., 2024).

Ada banyak pendekatan yang telah diusulkan untuk klasifikasi gerakan tangan statis dan dinamis. Di antaranya, *Convolutional Neural Network* (CNN) telah muncul sebagai salah satu teknik yang menjanjikan di bidang pengenalan pola dan klasifikasi gambar. CNN analog dengan jaringan saraf tiruan tradisional. CNN memungkinkan untuk menggunakan fitur khusus gambar yang tertanam dalam jaringan, membuatnya cocok untuk video dan gambar digital. Ini menyediakan operasi yang terfokus, bersama dengan pengurangan parameter yang dapat dilatih dibandingkan dengan ANN. Data gambar membutuhkan sumber daya komputasi yang besar jika diproses oleh ANN. Operasi seperti penggabungan dan konvolusi memperkenalkan generalisasi dan dimensionalitas yang berkurang dalam jaringan. Kemajuan terbaru dan pengenalan berbagai arsitektur CNN padat menghasilkan peningkatan efisiensi dan akurasi sistem pengenalan gerakan tangan dinamis. Namun, sistem gerakan tangan dinamis di dunia nyata berjuang dengan berbagai tantangan terbuka termasuk, kondisi pencahayaan yang berbeda, waktu pemrosesan, deteksi segmen tangan, dan banyak lagi (Singh, 2021). Baru-baru ini para peneliti menggunakan *Convolutional Neural Network* untuk memecahkan masalah ini. Pigou dkk. menggunakan *Convolutional Neural Network* normal dengan 2 aliran input dan beberapa pra-pemrosesan. Sementara Huang dkk. menggunakan *Convolutional Network 3D* dengan input 5 saluran (*Red, Green, Blue, Alpha, Skeleton*). Hasilnya sangat bagus. Namun ada penelitian baru tentang pengenalan tindakan yang menjadi state of the art yaitu Carreira dkk (Suharjo et al., 2021).

CNN merupakan suatu sistem dengan menerapkan model kerja *Deep Learning*. Penelitian ini menggunakan model CNN karena dinilai sebagai model yang paling sesuai fungsinya. Dalam artikel ilmiahnya Indra mengatakan kemampuan CNN diklaim sebagai model terbaik yang digunakan untuk memecahkan permasalahan deteksi objek dan pengenalan objek karena

tidak memerlukan komputasi yang besar dalam prosesnya. Menurut Yann (LeCun et al., 2015) Ini adalah salah satu masalah penting yang dapat diselesaikan dengan Jaringan Saraf Konvolusional (CNN). Perkembangan terbaru dalam CNN telah memungkinkan untuk memperluas area masalah ini dari pengenalan karakter Bahasa Inggris atau pengenalan angka ke pengenalan karakter bahasa daerah, namun masih belum banyak penelitian yang dilakukan di domain bahasa daerah (Katoch et al., 2022).

Convolutional Neural Network (CNN) adalah salah satu metode *Deep Learning* yang paling umum digunakan untuk menganalisis citra visual. CNN membutuhkan prapemrosesan yang lebih sedikit dibandingkan dengan algoritma klasifikasi gambar lainnya. Jaringan ini mempelajari filter-filter yang biasanya dirancang secara manual dalam sistem lain. Penggunaan CNN mengubah gambar menjadi format yang lebih mudah diproses sambil tetap mempertahankan fitur-fitur penting yang diperlukan untuk membuat prediksi yang akurat (KASAPBAŞI et al., 2022). Dengan transfer learning, model tuning terbaik dapat mencapai akurasi 97,5% yang hanya memiliki 1 kesalahan di antara semua dataset pengujian. tanpa transfer learning model ini tidak akan bekerja dengan dataset yang sangat kecil. 200 data dengan 10 kelas adalah data yang sangat kecil. Namun kombinasi metode ini dapat meningkatkan akurasi hingga 97,5%. Selain itu, dengan metode CNN dapat memangkas beberapa pra-pemrosesan dan ekstraksi fitur dibandingkan dengan metode pengklasifikasi lama seperti SVM, HMM, dll.

Karena model CNN dapat mengekstraksi fiturnya sendiri dengan sangat baik terutama dengan model CNN (Suharjito et al., 2021). Akurasi pelatihan mencapai 99,94%. Akurasi pengujian mencapai 97,62%. CNN outperform dibandingkan model lainnya seperti SVM, KNN, Logistic Regression, dan LeNet-5. Model CNN menghasilkan skor presisi, recall, dan F1 rata-rata tertimbang sebesar 98%. Metode berbasis CNN yang diusulkan memberikan performa terbaik untuk pengenalan angka BSL, menjadikannya solusi yang potensial untuk mengurangi hambatan komunikasi bagi komunitas tuna rungu di Bhutan(Wangchuk et al., 2021).

1.2 Rumusan Masalah

Dalam berkomunikasi antara penyandang tunarungu dan masyarakat umum, terdapat kesenjangan karena tidak semua orang memahami bahasa isyarat. Hal ini menjadi tantangan besar, terutama dalam berinteraksi dan berkomunikasi. Dengan perkembangan teknologi, kita

bisa membuat sistem yang mampu mengenali dan mengklasifikasikan huruf dalam bahasa isyarat secara otomatis.

Berdasarkan hal tersebut, beberapa pertanyaan yang muncul adalah:

1. Bagaimana cara membangun model berbasis CNN yang dapat mengenali huruf-huruf bahasa isyarat?
2. Seberapa akurat model ini dalam mengklasifikasikan huruf bahasa isyarat berdasarkan data yang digunakan?

1.3 Batasan Masalah

Adapun batasan pada penelitian ini, yaitu:

1. Pada penelitian ini menggunakan *American Sign Language* (ASL).
2. Data set untuk penelitian ini menggunakan foto tidak menggunakan video.
3. Metode yang digunakan untuk klasifikasi adalah Convolutional Neural Network (CNN) tanpa membandingkannya dengan metode lainnya.
4. Sistem hanya akan mengenali huruf dalam bahasa isyarat, bukan kata atau kalimat.
5. Pengambilan gambar dilakukan pada jarak 20 sampai 25centimeter dengan background tanpa corak.

1.4 Tujuan Penelitian

Berdasarkan uraian dari rumusan masalah diatas, maka tujuan dari penelitian ini adalah:

1. Membuat model kecerdasan buatan berbasis CNN yang dapat mengenali huruf bahasa isyarat dari gambar tangan.
2. Mengukur sejauh mana model ini dapat mengenali huruf dengan tingkat akurasi yang baik.

1.5 Manfaat Penelitian

Diharapkan hasil penelitian ini dapat menjadi referensi dalam penelitian di bidang pengolahan citra dan kecerdasan buatan, khususnya dalam penerapan CNN untuk bahasa isyarat.

1.6 Sistematika Penulisan

Pada penulisan ini memiliki beberapa bagian sebagai berikut :

BAB I. PENDAHULUAN

Pada bagian pendahuluan berisi tentang latar belakang penulisan, rumusan masalah, batasan masalah, tujuan penelitian, dan manfaat penelitian.

BAB II. LANDASAN TEORI

Pada bagian ini menjelaskan landasan teori yang berhubungan tentang klasifikasi dan prediksi huruf bahasa isyarat yang menggunakan Convolutional Neural Network (CNN), Jupyter Notebook, Python, dan American Sign Language (ASL).

BAB III. METODE PENELITIAN

Bab ini menjelaskan tentang prosedur yang dilakukan dalam penelitian ini meliputi tahapan penelitian, pengumpulan data, klasifikasi gambar.

BAB IV. HASIL DAN PEMBAHASAN

BAB V. KESIMPULAN

