

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Dalam penyusunan penelitian ini, penulis mendapatkan referensi dari peneliti sebelumnya yang berkaitan dengan latar belakang masalah pada penelitian ini berikut: penelitian yang dilakukan oleh Salma Hayani, Mohamed Benaddy, Othmane El Meslouhi dan Mustapha Kardouchi. Bahasa yang digunakan adalah Bahasa arab dengan 28 alfabet dan juga angka 0 sampai dengan 1. Dimana CNN yang digunakan dalam penelitian ini. Yang digunakan Melati data Adam *Optimizer* dengan *Learning rate 0.03*. Dengan data yang digunakan sebanyak 7869 citra dimana didapatkan akurasi yang terbaik pada data latih yang digunakan sebanyak atau sebesar 80% dari data total[4].

Penelitian yang berjudul "Pengenalan Angka Sistem Isyarat Bahasa Indonesia Menggunakan Metode Convolutional Neural Network" dilaksanakan pada tahun 2019 oleh Mochamad Bagus Setiyo Bakti dan Yuliana Melita Pranoto[5]. Mereka memanfaatkan arsitektur LeNet untuk struktur CNN mereka. Dataset yang digunakan terdiri dari 450 gambar citra angka dalam Sistem Isyarat Bahasa Indonesia (SIBI), yang dibagi menjadi 10 kelas, yaitu angka 0 hingga 9. Dataset ini dibagi menjadi dua bagian: 80% untuk pelatihan dan 20% untuk pengujian. Pelatihan dilakukan dalam tiga tahap percobaan dengan menggunakan 25 epoch, 50 epoch, dan 100 epoch. Dalam percobaan pertama, pelatihan menggunakan 25 epoch menghasilkan akurasi sebesar 67,66% dengan kerugian sebesar 0,95%. Saat dilakukan pengujian pada 20% data, diperoleh akurasi sebesar 79,23%. Dalam percobaan kedua, pelatihan dilakukan

dengan 50 epoch dan menghasilkan akurasi 89,44% serta kerugian 0,37%. Pengujian berikutnya pada data menghasilkan akurasi sebesar 90,44%. Percobaan ketiga menggunakan 100 epoch untuk pelatihan dan berhasil mencapai akurasi 96,44% dengan kerugian sebesar 0,13%. Pengujian akhir pada data menghasilkan akurasi yang sangat tinggi, yaitu 98,89%.

Penelitian yang dilakukan oleh [6] pada penelitian ini menemukan bahwa deteksi tangan paling optimal pada jarak 110 cm dari kamera. Hasil keseluruhan penelitian ini cukup memuaskan, meskipun beberapa huruf menunjukkan akurasi di bawah 90% karena adanya kemiripan antar isyarat. Penelitian ini berhasil mengimplementasikan Face Recognition pada dataset wajah 25 orang, masing-masing dengan 10 data wajah. Dataset ini dibagi menjadi data latih dan data uji. Dengan menggunakan framework Caffe Deep Learning untuk deteksi wajah, diperoleh hasil 100% sesuai kriteria yang ditentukan. Untuk klasifikasi, CNN menghasilkan persentase 98%, precision 98,4%, recall 98%, dan accuracy 99,84%.

Penelitian yang dilakukan oleh [7] menemukan bahwa akurasi sistem dalam mengidentifikasi dan mengenali citra jamur sangat dipengaruhi oleh parameter jaringan internal, seperti laju pembelajaran, jumlah neuron pada layer tersembunyi, dan jumlah iterasi. Persentase akurasi tertinggi mencapai 93%, yang menunjukkan hasil yang baik. Untuk penelitian lebih lanjut, disarankan untuk mencoba metode yang tidak mengubah citra menjadi skala abu-abu terlebih dahulu.

Penelitian ini menghasilkan model penerjemah bahasa isyarat berbasis Convolutional Neural Network (CNN) yang mampu mengidentifikasi gerakan tangan dalam rekaman video dengan akurasi tinggi. Sistem ini menggunakan standar bahasa

isyarat Indonesia (BISINDO) dan berfungsi secara efisien untuk merajut komunikasi antara pengguna dan sistem. Implementasi sistem ini memungkinkan respons instan terhadap gerakan tangan yang terdeteksi dalam video, mendukung interaksi seketika antara pengguna dan sistem. Evaluasi kinerja menunjukkan bahwa sistem memiliki tingkat akurasi yang tinggi dalam mengenali dan menerjemahkan gerakan tangan dalam bahasa isyarat Indonesia (BISINDO)[8].

2.2 Landasan Teori

2.2.1 Bahasa Isyarat

Bahasa isyarat adalah bahasa komunikasi yang digunakan oleh penyandang tunarungu untuk menyampaikan informasi melalui gerakan tangan, ekspresi wajah, dan gerakan tubuh. Asalnya berkembang secara alami di komunitas tunarungu sebagai bentuk komunikasi visual. Pada abad ke-18, bahasa isyarat mulai diakui dan dipelajari secara formal, dengan tokoh penting seperti Abbe de l'Épée yang mendirikan sekolah pertama untuk anak-anak tunarungu di Prancis. Setiap negara memiliki versi bahasa isyaratnya sendiri, seperti American Sign Language (ASL) di Amerika Serikat dan British Sign Language (BSL) di Inggris. Bahasa isyarat memungkinkan komunikasi yang efektif dan ekspresif tanpa memerlukan suara, menjadikannya penting bagi komunitas tunarungu dalam kehidupan sehari-hari[9].

2.2.2 Sistem Isyarat Bahasa Indonesia (SIBI)

SIBI (Sistem Isyarat Bahasa Indonesia) adalah bahasa isyarat yang mengadaptasi gerakan ASL (*American Sign Language*). Bahasa isyarat abjad SIBI memiliki 26 gerakan yang mewakili 26 alfabet yang terdiri dari 24 peragaan berupa gerakan pasif serta 2 peragaan berupa gerakan aktif (J dan Z) menggunakan satu

tangan. SIBI sering digunakan dalam komunikasi formal, dalam penerapannya digunakan dalam proses belajar mengajar di SLB. Peraturan Menteri Pendidikan Republik Indonesia nomor 0161/U/1994 mengatur tentang penggunaan SIBI dalam SLB[9].

2.2.3 Deep Learning

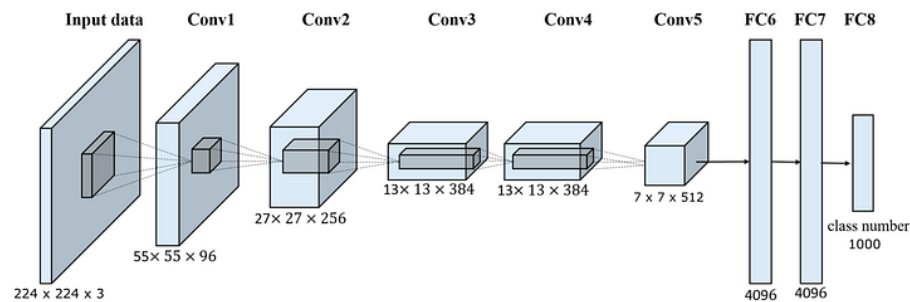
Deep learning adalah cabang *machine learning* yang menggunakan jaringan saraf tiruan untuk memodelkan data kompleks. Dimulai pada 1940-an oleh Warren McCulloch dan Walter Pitts, jaringan saraf tiruan mengalami kemajuan signifikan pada 1980-an dengan algoritma backpropagation oleh Geoffrey Hinton. Pada 2000-an, *deep learning* berkembang pesat berkat peningkatan komputasi, data besar, dan inovasi arsitektur seperti CNN dan RNN. Puncaknya, pada 2012, AlexNet memenangkan kompetisi ImageNet, meningkatkan popularitas *deep learning* yang kini digunakan dalam pengenalan gambar, pemrosesan bahasa alami, dan pengemudi otomatis.

Deep learning menggunakan neural network dengan input metadata yang diolah melalui fungsi transformasi non-linier berlapis. Terdiri dari lapisan tersembunyi (hidden layer), jaringan ini melatih metadata berdasarkan output neuron sebelumnya. Neural network pada *deep learning* membentuk hirarki sederhana dengan banyak lapisan (multi-layer).[10].

2.2.4 Algoritma Convolutional Neural Network (CNN)

Convolution Neural Network (CNN) adalah jenis jaringan saraf memiliki multi lapisan yang sangat khusus. Pada dasarnya CNN ini dirancang untuk mengenali pola visual langsung dari gambar dengan pemrosesan minimal. CNN, atau *Convolutional Neural Networks*, berbeda dari MLP, atau Multi-Layer Perceptrons, karena dua alasan

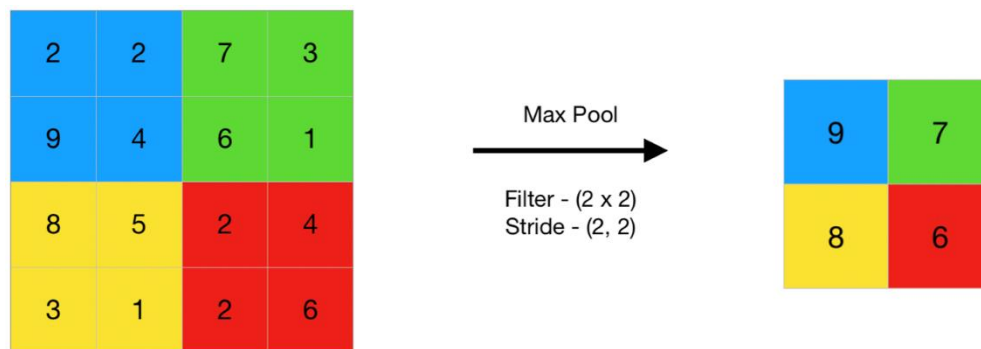
utama. Pertama, CNN mempertahankan struktur spasial dari gambar dengan memproses data melalui lapisan konvolusi yang mempertimbangkan hubungan lokal antara piksel, yang penting untuk memahami fitur visual seperti tepi, bentuk, dan tekstur. Kedua, CNN menggunakan lapisan tersembunyi yang lebih kompleks, termasuk lapisan *konvolusi*, lapisan *pooling*, dan lapisan *fully connected*, yang bekerja bersama untuk efektif mengekstrak dan mempelajari fitur dari data input dengan cara yang lebih efisien dan representatif[11]. Berikut tahapan proses dari CNN dapat dilihat pada Gambar 2.1 di bawah ini.



Gambar 2. 1 Tahapan proses CNN

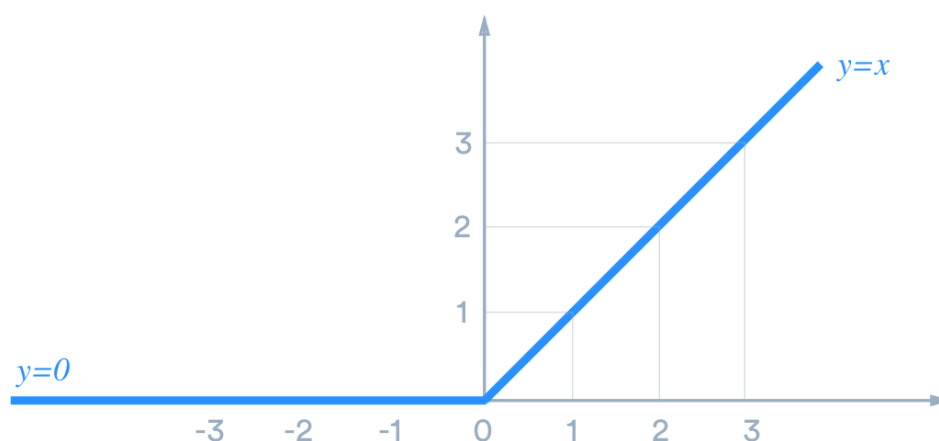
Lapisan *pooling* berfungsi untuk mereduksi dimensi dari peta fitur dengan cara mengurangi jumlah elemen dalam tinggi dan lebar, sambil mempertahankan kedalaman dari peta tersebut. Terdapat dua jenis metode *pooling*[12] yang umum digunakan, yaitu *average pooling* dan *max pooling*. *Max pooling* umumnya lebih efektif dalam praktek karena metode ini memilih nilai maksimum dari setiap segmen pada peta fitur. Hal ini menghasilkan peta fitur yang lebih fokus pada fitur paling dominan dalam segmen tersebut, dibandingkan dengan *average pooling* yang hanya menghitung rata-rata nilai dalam segmen, yang bisa mengaburkan fitur-fitur

penting[13]. Berikut untuk metode umum dari pooling dapat di lihat pada Gambar 2.2 di bawah ini.



Gambar 2. 2 Metode Umum Pooling

Dalam CNN modern, fungsi aktivasi seperti ReLU dan hard tanh telah menggantikan sigmoid dan soft tanh karena lebih efisien dalam melatih model. ReLU, khususnya, mengubah semua nilai negatif menjadi nol dengan fungsi $f(x) = \max(0, x)$, yang mempercepat *konvergensi* selama pelatihan. Berikut fungsi dari RuLe dapat dilihat pada Gambar 2.3 dibawah ini.



Gambar 2. 3 Fungsi RuLe

2.2.5 Python

Python merupakan bahasa pemrograman tingkat tinggi yang dibuat oleh Guido Van Rossum dan dirilis pada tahun 1991. Belakangan ini, Python menjadi sangat populer. Selain itu, Python juga merupakan bahasa pemrograman serbaguna yang dapat digunakan untuk berbagai keperluan. Sebagai contoh, Python bisa digunakan untuk Machine Learning dan Deep Learning.

Dengan menggunakan Python dan berbagai library-nya, produktivitas para pengembang Machine Learning dan Deep Learning dapat meningkat secara signifikan. Bagi pemula yang ingin terjun ke dunia Machine Learning atau Deep Learning, tidak perlu khawatir karena banyak sekali sumber pembelajaran yang tersedia, terutama di internet[15].

2.2.6 Tensorflow

TensorFlow, dikembangkan oleh tim Google Brain, adalah sebuah framework yang dirancang untuk memfasilitasi pengembangan dan eksekusi algoritma pembelajaran mesin. *Framework* ini menggunakan C++ untuk kinerja yang optimal dan mendukung berbagai jenis algoritma, baik untuk pelatihan maupun inferensi dalam model jaringan saraf dalam. TensorFlow dirancang untuk memiliki fleksibilitas tinggi, memungkinkan penggunaannya di berbagai platform seperti CPU dan GPU, dan kompatibel dengan sistem operasi seperti Windows, macOS, Linux, serta perangkat mobile Android dan iOS.

2.2.7 Google Colaboratory

Colaboratory, atau Colab, adalah sebuah platform berbasis cloud yang dikembangkan oleh Google untuk memfasilitasi pembelajaran dan penelitian dalam bidang machine learning. Sebagai sebuah layanan gratis, Colab[16] menyediakan akses ke sumber daya komputasi termasuk GPU tanpa biaya. Platform ini berbasis pada *Jupyter Notebook*, sebuah proyek sumber terbuka, tetapi menawarkan keunggulan karena pengguna dapat menggunakan dan berbagi *notebook Jupyter* secara online tanpa perlu proses unduhan atau instalasi apa pun.

2.3 Kerangka Pikir

Kerangka pikir dalam penelitian ini dapat dilihat pada Gambar 2.4 dibawah ini:



Gambar 2. 4 Kerangkah Pikir