

OPTIMASI AKURASI DETEKSI SIKAP TANGAN MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK

Jidan Dewa Saksana*¹⁾, Ema Utami²⁾, Hanif Al Fatta³⁾

Magister Teknik Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta, Indonesia

Article Info

Kata Kunci: Convolutional Neural Network
Deep Learning; Hands Gesture Recognition;
Image Processing

Article history:

Received 25 September 2024

Revised 23 October 2024

Accepted 16 November 2024

Available online 1 December 2025

DOI :

<https://doi.org/10.29100/jipi.v10i4.6735>

* Corresponding author.

Corresponding Author

E-mail address:

dewa.saksana@students.amikom.ac.id

ABSTRAK

Penelitian ini akan menggunakan pendekatan berbasis deep learning yang telah terbukti berhasil dalam deteksi sikap tangan, yakni Convolutional Neural Networks (CNN). Studi mengusulkan metode optimasi yang mencakup peningkatan preprocessing data, arsitektur CNN yang disesuaikan, dan strategi augmentasi data untuk meningkatkan model terhadap variasi input. Hasil eksperimen menunjukkan bahwa metode optimasi yang diusulkan berhasil meningkatkan akurasi deteksi gerakan tangan dibandingkan dengan pendekatan standar. Pengujian dilakukan pada dataset publik dengan metrik evaluasi yang berfokus pada akurasi.

ABSTRACT

This study will use a deep learning-based approach that has been proven successful in hand gesture recognition, namely Convolutional Neural Networks (CNN). The study proposes an optimization method that includes improved data preprocessing, customized CNN architecture, and data augmentation strategies to improve the model against input variations. Experimental results show that the proposed optimization method successfully improves the accuracy of hand gesture detection compared to standard approaches. Testing is carried out on public datasets with evaluation metrics that focus on accuracy.

I. PENDAHULUAN

DALAM beberapa dekade terakhir deteksi sikap tangan atau Hands Gesture Recognition telah menjadi area penelitian yang penting dalam bidang komputer vision dan pengenalan pola. Deteksi gerakan tangan adalah bidang studi yang mencoba mengidentifikasi gerakan manusia melalui algoritma matematika juga dapat digunakan di beberapa bidang seperti komunikasi bahasa isyarat dan interaksi manusia dengan komputer. Akan tetapi perubahan skala dan sudut, serta latar belakang membuat pengenalan gerakan cukup menantang [1].

Meskipun telah ada kemajuan signifikan dalam pengolahan deteksi sikap tangan, masih terdapat tantangan dalam mencapai akurasi yang tinggi [2]. Beberapa faktor yang mempengaruhi akurasi termasuk variasi posisi tangan, perubahan pencahayaan, penutupan sebagian tangan, dan kerumitan lingkungan. Oleh karena itu kesempatan untuk mengoptimalkan performa deteksi sikap tangan masih ada [3].

Algoritma yang paling sering digunakan pada pengolahan citra antara lain Naïve Bayes [4], Support Vector Machine (SVM) [5], Recurrent Neural Network [6], Long Short-Term Memory (LSTM) [7], dan Convolutional Neural Network (CNN) [8].

Penelitian sebelumnya mencoba melakukan eksperimen pengolahan gambar menggunakan pendekatan Naïve Bayes dan Support Vector Machine kemudian mendapatkan hasil yang buruk pada data training yang banyak [9]. Adapun metode Recurrent Neural Network (RNN) dan Long Short-Term Memory juga dapat digunakan dalam pengolahan gambar, akan tetapi metode ini lebih umum digunakan untuk dataset urutan seperti teks. Setiap teknik memiliki keunggulan masing-masing tergantung pada spesifik dan tujuan penelitian [10].

Convolutional Neural Network menjadi salah satu metode yang populer digunakan dalam penelitian *image processing*. Arsitektur CNN dirancang khusus untuk mengenali pola dalam gambar seperti bentuk dan warna. Selama proses data training CNN memiliki sifat adaptif yang membuat metode ini lebih ideal mendeteksi pola gambar [11]. CNN memiliki *feature extraction* membuatnya dapat mengekstraksi fitur dari gambar secara otomatis melalui lapisan konvolusi dan *pooling*, sedangkan Naive Bayes dan SVM memerlukan ekstraksi fitur secara manual. Kemudian adanya fungsi *deep learning* memungkinkan model untuk belajar secara lebih mendalam, sedangkan Naive Bayes dan SVM lebih sederhana dan tidak memanfaatkan *deep learning* [12].

Adapun LeNet merupakan salah satu arsitektur dari CNN, memiliki kinerja yang sangat baik dalam tugas pengenalan tulisan tangan dan dapat diterapkan dalam berbagai data gambar sederhana [13]. LeNet menunjukkan efisiensi yang tinggi dengan waktu training relatif singkat dibandingkan dengan arsitektur lainnya. LeNet dirancang khusus untuk pengenalan karakter tulisan tangan, seperti yang ditemukan dalam dataset MNIST. Arsitekturnya relatif sederhana dibandingkan dengan model CNN modern. Terdiri dari beberapa lapisan konvolusi, lapisan *pooling*, dan lapisan *fully connected*. Menggunakan suatu algoritma dapat menghasilkan pengklasifikasi yang efektif pada dataset tertentu. Namun ketika diterapkan pada dataset lain, pengklasifikasi yang dilatih dengan algoritma yang sama bisa saja menunjukkan kinerja yang lebih rendah [14].

Berdasarkan beberapa penelitian di atas, penelitian ini memanfaatkan tiga dataset yang berbeda yakni Hand Gesture Recognition Dataset, CIFAR-10, dan MNIST. Tujuan penelitian adalah untuk mengevaluasi apakah CNN dapat menghasilkan akurasi lebih tinggi dari LeNet yang pada dasarnya dirancang khusus untuk tugas pengenalan tulisan tangan. Melalui pengujian pada berbagai jenis data, diharapkan dapat diperoleh wawasan yang mendalam mengenai performa dan kemampuan CNN dan LeNet dalam berbagai konteks pengenalan pola serta memberikan panduan bagi peneliti lain terkait topik.

II. METODOLOGI PENELITIAN

Beberapa langkah yang digunakan optimasi akurasi pada penelitian ini antara lain

A. Data Collection

Dataset yang dipakai pada saat penelitian ini didapatkan dari website Kaggle. Citra sikap tangan memiliki jumlah 200 citra disetiap folder dengan format PNG pada setiap kelasnya.

B. Data Preprocessing

Tahap preprocessing yang diterapkan dalam penelitian ini adalah augmentasi dengan menggunakan beberapa teknik seperti membalik, rotasi, menggeser, dan memperbesar gambar. Dataset yang digunakan dalam penelitian ini berjumlah 20000 citra yang dibagi dalam beberapa folder.

1. Data Split

Pemisahan data dilakukan menggunakan `train_test_split` yang ada pada library `pyhton`. Data training, data yang digunakan untuk menjalankan proses pelatihan model. Data yang digunakan dalam untuk proses pelatihan sebesar 80%. Data testing, data yang digunakan untuk menguji kinerja model. Data yang digunakan dalam penelitian ini untuk proses pengujian sebesar 20%.

2. Data Training

Tahap pelatihan model dalam penelitian ini menggunakan model Convolutional Neural Network (CNN) dengan arsitektur LeNet. Pelatihan model yang dilakukan menggunakan augmentasi data untuk menambah variasi pada data training dan menggunakan konfigurasi parameter seperti *input size*, *epoch*, *batch size*, *optimizer*, dan *learning rate* untuk membantu proses saat pelatihan model

C. Deskripsi Metode

1. Convolutional Neural Networks

CNN dipilih karena kemampuannya yang unggul dalam menangani data berbentuk grid, seperti gambar, dengan memanfaatkan lapisan konvolusi untuk mengekstraksi fitur penting secara otomatis. CNN mampu mengenali pola visual yang kompleks tanpa perlu fitur ekstraksi manual, sehingga sangat efektif untuk pengenalan gambar dan deteksi objek. Namun model CNN buruk dalam menangani latar belakang yang kompleks dan tidak konsisten dalam kumpulan dataset [15].

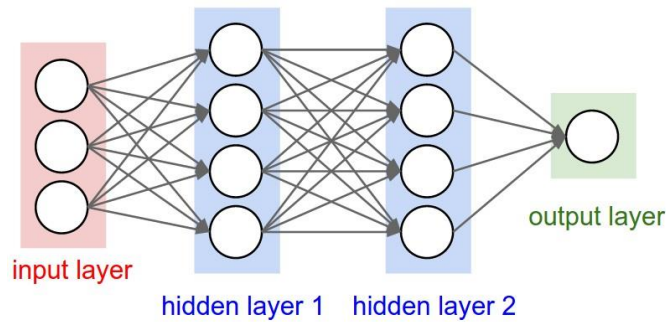
2. LeNet

LeNet salah satu arsitektur CNN paling awal, dipilih karena kesederhanaannya dan efisiensinya dalam mengolah citra gambar serta klasifikasi yang lebih sederhana. Struktur LeNet terdiri dari beberapa lapisan konvolusi dan pooling diikuti oleh lapisan fully connected menjadikannya ideal sebagai model dasar atau sebagai pengenalan awal terhadap penggunaan CNN dalam pengolahan citra.

D. Convolution Neural Network

1. Gambaran Umum

Dikenal sebagai jaringan saraf konvolusional atau CNN merupakan jenis jaringan saraf khusus untuk memproses data yang memiliki topologi grid yang dikenal. Nama "jaringan saraf konvolusional" menunjukkan bahwa jaringan tersebut menggunakan operasi matematika yang disebut konvolusi. Konvolusi adalah jenis operasi linier yang khusus. Jaringan konvolusional merupakan gambaran sederhana jaringan saraf yang menggunakan konvolusi sebagai pengganti perkalian matriks umum di setidaknya satu lapisannya [16].



Gambar. 1. CNN Model Layer

Gambar 1 menggambarkan jaringan syaraf menerima input berupa vektor dan mengubahnya melalui serangkaian hidden layer. Setiap hidden layer terdiri dari satu set neuron, di mana setiap neuron terhubung sepenuhnya ke semua neuron pada layer sebelumnya dan mengirimkan output ke semua neuron pada lapisan berikutnya. Tidak ada koneksi antar neuron dalam lapisan yang sama, sehingga neuron tersebut berfungsi secara independen satu sama lain. Layer terakhir yang terhubung disebut "output layer" dan dalam pengaturan klasifikasi layer ini mewakili nilai kelas.

a. Operasi Konvolusi

Operasi konvolusi dilakukan antara filter (kernel) dan input (citra) untuk menghasilkan output (*feature map*).

$$(I * K)(i, j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i + m, j + n) * K(m, n)$$

Di mana:

- $I(i, j)$ adalah nilai piksel pada posisi (i, j) dari input (citra).
- $K(m, n)$ adalah nilai filter (kernel) pada posisi (m, n) .
- M dan N adalah ukuran filter (kernel), misalnya 3×3 .
- $(I * K)(i, j)$ adalah hasil konvolusi pada posisi (i, j) .

b. Aktivasi :

Setelah konvolusi, fungsi aktivasi seperti ReLU (*Rectified Linear Unit*) sering digunakan untuk menambahkan non-linearitas:

$$f(x) = \max(0, x)$$

Di mana (x) adalah nilai dari *feature map* setelah konvolusi.

c. Pooling

Pooling layer melakukan *downsampling* untuk mengurangi dimensi dari *feature map*, salah satu teknik umum adalah *Max Pooling*:

$$P(i, j) = \max_{m,n} \{F(2i + m, 2j + n)\}$$

Di mana:

- $P(i, j)$ adalah nilai setelah *pooling*.
- F adalah *feature map* input untuk *pooling*.
- *Pooling* biasanya dilakukan dengan ukuran 2×2 dan stride 2.

d. Fully Connected

Setelah beberapa lapisan konvolusi dan *pooling*, data yang telah direduksi dihubungkan ke *fully connected* layer, yang diperlakukan seperti jaringan saraf tradisional:

$$y = W * x + b$$

Di mana:

- x adalah input (flattened feature map).
- W adalah bobot.
- b adalah bias.
- y adalah output, yang biasanya melalui fungsi softmax untuk klasifikasi.

E. Optimasi Parameter

Penelitian ini menggunakan parameter *learning rate* yang disesuaikan dengan *optimizer Adam* untuk meningkatkan kinerja mode. *Adam*, yang merupakan singkatan dari *Adaptive Moment Estimation*, adalah algoritma optimasi yang sangat efisien karena mampu menyesuaikan laju pembelajaran setiap parameter berdasarkan sejarah gradiennya. Dengan kombinasi ini penelitian dapat mencapai konvergensi lebih cepat dan meningkatkan akurasi model secara keseluruhan.

F. Model Evaluasi

Metrik evaluasi adalah model evaluasi yang menghasilkan akurasi dalam CNN. Akurasi klasifikasi merupakan salah satu metrik evaluasi yang paling umum digunakan. Akurasi klasifikasi adalah persentase prediksi yang benar dari total prediksi yang dibuat oleh model. Nilai akurasi berfungsi sebagai perbandingan dari beberapa tes data yang dilakukan dengan CNN maupun arsitektur LeNet. Alur penelitian termuat dalam Tabel I.

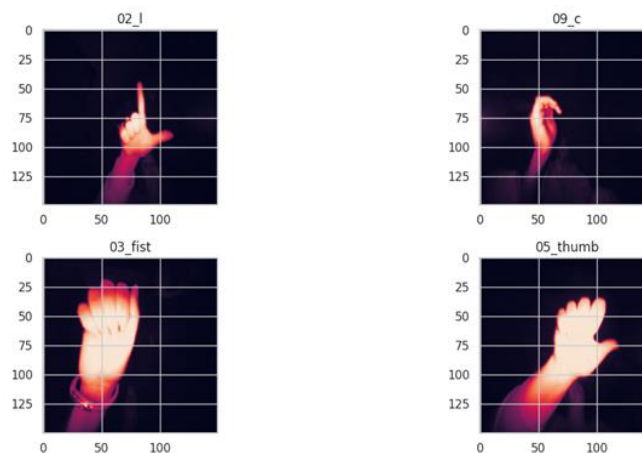
TABEL I
ALUR PENELITIAN

Tahapan	Aktivitas	Deskripsi
<i>Data Collection</i>	Pengumpulan Data	Mengambil dataset dari website Kaggle berjudul Hands Gesture Recognition Dataset
<i>Pre-processing</i>	Augmentation	Melakukan pemrosesan data seperti rotasi, pergeseran, balik, dan perbesaran/perkecil gambar
	Split Data	Membagi data menjadi 80% untuk data training dan 20% untuk data testing
<i>Machine Learning</i>	CNN	Melakukan penerapan metode CNN dalam mengekstraksi dataset
	LeNet	Melakukan penerapan metode LeNet dalam mengekstraksi dataset
<i>Evaluation</i>	Model Evaluasi	Menggunakan metrik standar untuk mengetahui hasil akurasi
<i>Analyze</i>	Analisis Hasil Akurasi	Analisis hasil untuk menilai efektivitas model dalam optimisasi akurasi
<i>Discussion</i>	Kesimpulan	Pembahasan dan kesimpulan berserta saran dari hasil penelitian

III. HASIL DAN PEMBAHASAN

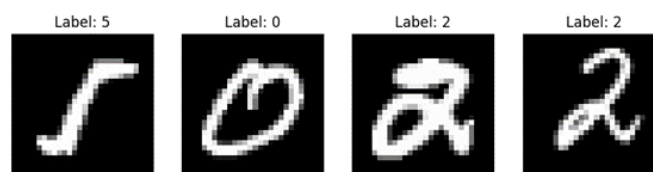
A. Dataset

Hands Gesture Recognition Dataset adalah kumpulan data yang berisi lebih dari 20.000 citra tangan yang berbeda. Dataset ini dikumpulkan dari berbagai sumber dan mencakup berbagai gerakan tangan, seperti mengangkat jari, menggenggam, dan membuka telapak tangan. Setiap citra memiliki anotasi yang menunjukkan jenis gerakan tangan yang terjadi. Dataset ini sangat berguna untuk pelatihan model pengenalan gerakan tangan menggunakan teknik *deep learning*, seperti CNN. Semakin banyak dan bagus dataset yang diolah maka hasil deteksi akurasi akan semakin baik [17]. Adapun penggunaan dataset MNIST dan CIFAR-10 sebagai bahan pembandingan untuk menguji performa model pada varian dataset. Gambar 2 menunjukkan beberapa citra tangan acak dari Hands Gesture Recognition Dataset.



Gambar. 2. Hands Gesture Recognition Dataset

Lecun pada penelitiannya memaparkan basis data digit tulisan tangan MNIST memiliki set pelatihan berisi 60.000 citra dan set pengujian berisi 10.000 citra. Dataset telah dinormalisasi ukurannya dan dipusatkan dalam gambar berukuran tetap. Basis data ini cocok bagi seseorang yang ingin mencoba mempelajari teknik dan metode pengenalan pola dengan sedikit upaya pada praproses dan pemformatan. Dataset ini berisi pengenalan tulisan tangan yang menjadi keahlian LeNet. Gambar 3 menunjukkan beberapa citra penulisan tangan acak dari MNIST Dataset.



Gambar. 3. MNIST Dataset

CIFAR-10 adalah dataset yang terdiri dari 60.000 gambar berwarna berukuran 32x32 piksel yang terbagi dalam 10 kelas berbeda antara lain : pesawat, mobil, burung, kucing, rusa, anjing, katak, kuda, kapal, dan truk. Dari jumlah tersebut, 50.000 gambar digunakan untuk data latih dan 10.000 gambar digunakan untuk data uji. Dataset ini dirancang untuk menjadi tantangan dalam tugas pengenalan gambar karena variasi yang tinggi di dalam setiap kelasnya. Gambar 4 menunjukkan beberapa citra penulisan tangan acak dari CIFAR-10 Dataset.



Gambar. 4. CIFAR-10 Dataset

B. Data Split

Rasio 80:20 pada split data mengacu pada pembagian dataset menjadi dua bagian dengan proporsi tertentu. Dalam hal ini 80% dari dataset digunakan untuk *data training*, sementara 20% sisanya digunakan untuk *data testing*. Rasio ini umumnya digunakan dan dianggap optimal untuk memastikan model memiliki akurasi yang baik tanpa mengorbankan ukuran dataset yang cukup besar untuk evaluasi. Dengan rasio 80:20 diharapkan dapat memastikan model terlatih dengan baik dan memiliki data yang cukup untuk menguji kinerjanya.

Perbandingan yang tepat antara data latih dan data uji membantu memastikan bahwa tidak terjadi *overfit* pada data latih. Perbandingan yang seimbang membantu menjaga integritas model dan memastikan bahwa model tersebut mencapai rasio terbaik dan siap untuk digunakan dalam lingkungan nyata. Demi mengetahui apakah model bekerja dengan baik, pengujian akan dilakukan dengan rasio yang lain karena tidak dapat dilakukan dengan hanya sekali dalam menentukan nilai parameter yang baik [18].

C. Pengujian

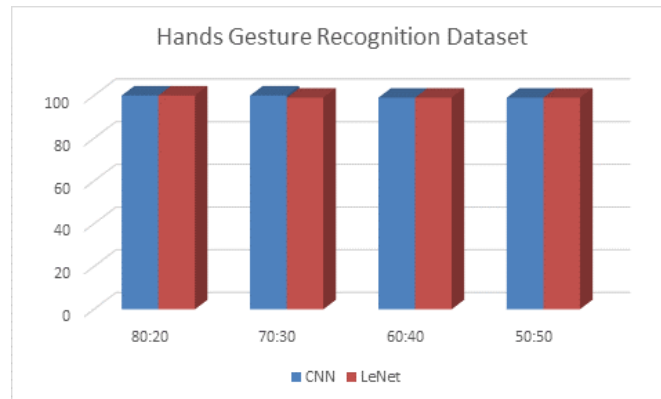
Tahap awal pengujian melibatkan library *TensorFlow* pada platform *Google Colaboratory*. *TensorFlow* dapat menangani berbagai detail di balik layar dengan baik, sehingga memudahkan proses pengujian dan pengembangan model [19]. Dengan menggunakan *TensorFlow*, pengguna dapat memanfaatkan kemampuan komputasi yang lebih efisien dan terstruktur, yang pada akhirnya meningkatkan produktivitas dan hasil dari pengujian tersebut.

1. Uji Pertama

Penelitian dilakukan menggunakan Hands Gesture Recognition Dataset untuk menguji kinerja CNN dan arsitektur LeNet. Dalam penelitian ini, saya melakukan pengujian secara berkala untuk memantau performa model dalam mengenali gerakan tangan. Pengujian ini mencakup beberapa iterasi untuk mendapatkan perbandingan akurasi yang lebih mendetail. Hasil dari penelitian ini menunjukkan perbedaan tingkat akurasi antara model CNN yang umum dan arsitektur LeNet, yang masing-masing memberikan perspektif yang unik dalam proses deteksi dan pengenalan gerakan tangan.

TABEL II
 HASIL AKURASI TES PERTAMA

Tes	Dataset	Rasio	Akurasi	
			CNN	LeNet
I	Hands Gesture Recognition Dataset	80:20	100%	100%
		70:30	100%	99.9%
		60:40	99.9%	99.9%
		50:50	99.9%	99.9%



Gambar. 5. Perbandingan akurasi CNN dan LeNet pada Hands Gesture Recognition Dataset

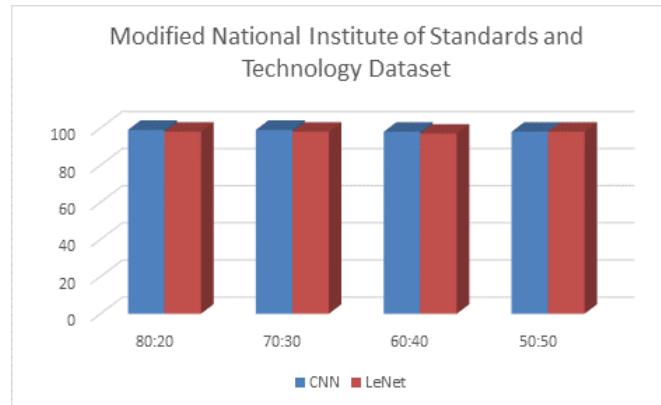
Gambar 5 dan Tabel II merupakan serangkaian pengujian dengan mencoba berbagai rasio pembagian data untuk training, termasuk rasio 80:20, 70:30, 60:40, dan 50:50. Tujuan dari variasi rasio ini adalah untuk mengidentifikasi pengaruh proporsi data pada kinerja model dalam mengenali gerakan tangan. Setiap rasio memberikan wawasan yang berbeda mengenai bagaimana model dapat beradaptasi dengan data yang lebih banyak atau lebih sedikit untuk pelatihan, serta bagaimana hal ini memengaruhi akurasi prediksi pada data uji. Hasil dari pengujian ini membantu dalam menentukan rasio yang paling optimal untuk digunakan dalam penelitian lebih lanjut.

2. Uji Kedua

Setelah melakukan pengujian dengan Hands Gesture Recognition Dataset, saya melanjutkan ujicoba dengan dataset lain pada Tabel III, yaitu Modified National Institute of Standards and Technology dataset (MNIST). Tujuan dari eksperimen ini adalah untuk mengevaluasi kinerja model CNN dan LeNet pada berbagai jenis data, tidak hanya pada pengenalan gerakan tangan.

TABEL III
 HASIL AKURASI TES KEDUA

Tes	Dataset	Rasio	Akurasi	
			CNN	LeNet
II	MNIST Dataset	80:20	99.0%	98.4%
		70:30	99.0%	98.6%
		60:40	98.6%	97.9%
		50:50	98.8%	98.4%



Gambar. 6. Perbandingan akurasi CNN dan LeNet pada *MNIST Dataset*

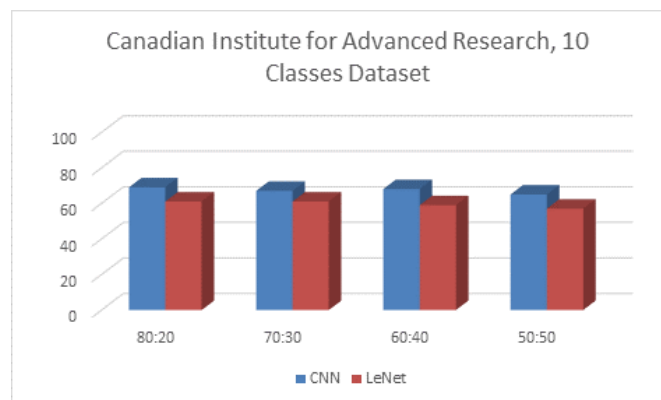
Gambar 6 menunjukkan perbandingan antara CNN dan LeNet menggunakan dataset MNIST. Hasil pengujian menunjukkan bahwa CNN memiliki akurasi yang lebih unggul dibandingkan dengan LeNet. Dataset MNIST, yang terdiri dari gambar grayscale dari angka tulisan tangan, digunakan untuk menguji kemampuan kedua model dalam mengklasifikasikan angka. CNN, dengan arsitektur yang lebih kompleks dan kemampuan untuk menangkap fitur yang lebih dalam, berhasil mencapai tingkat akurasi yang lebih tinggi, membuktikan keunggulannya dalam tugas klasifikasi gambar dibandingkan dengan LeNet.

3. Uji Ketiga

Kemudian pada Tabel IV terjadi pengujian terakhir dengan Canadian Institute for Advanced Research, 10 classes(CIFAR-10).

TABEL IV
 HASIL AKURASI TES KETIGA

Tes	Dataset	Rasio	Akurasi	
			CNN	LeNet
III	CIFAR-10 Dataset	80:20	69.9%	61.8%
		70:30	67.4%	61.4%
		60:40	68.3%	59.3%
		50:50	65.8%	57.5%



Gambar. 7. Perbandingan akurasi CNN dan LeNet pada *CIFAR-10 Dataset*

Gambar 7 menunjukkan bahwa CNN memiliki akurasi yang lebih unggul dibandingkan dengan LeNet. Dataset CIFAR-10, yang terdiri dari gambar berwarna dalam 10 kelas yang berbeda, digunakan untuk menguji kemampuan kedua model dalam mengklasifikasikan gambar. CNN dengan arsitektur yang lebih kompleks dan kemampuan untuk menangkap fitur yang lebih dalam berhasil mencapai tingkat akurasi yang lebih tinggi, membuktikan keunggulannya dalam tugas klasifikasi gambar dibandingkan dengan LeNet.

Dari semua pengujian menunjukkan hasil bahwa model CNN secara konsisten menghasilkan akurasi yang lebih tinggi dibandingkan dengan LeNet pada semua pengujian. Temuan ini memperkuat keunggulan arsitektur CNN dalam menangani berbagai jenis data gambar, baik untuk tugas pengenalan gerakan tangan maupun untuk dataset umum seperti MNIST dan CIFAR-10. CNN adalah metode untuk hasil klasifikasi terbaik dan peningkatan juga diusulkan dalam penelitian yang lain [20].

IV. KESIMPULAN

Hasil menunjukkan bahwa CNN lebih kompleks dibandingkan dengan LeNet. Model CNN menunjukkan keunggulan akurasi dari setiap pengujian dataset yakni Hand Gesture Recognition Dataset (100% vs 99.9%), MNIST (99.0% vs 98.6%), maupun CIFAR-10 (67.4% vs 61.4%). Implementasi deep learning pada dataset sikap tangan memperoleh kesimpulan yakni identifikasi citra menggunakan CNN berhasil dilakukan menghasilkan akurasi yang lebih baik. Sedangkan identifikasi citra menggunakan arsitektur LeNet berhasil dilakukan dengan cara yang sama menghasilkan akurasi dengan nilai di bawah CNN meski dengan dataset yang berbeda. Pada proses pengujian data menggunakan epoch sebanyak 5, menggunakan optimizer adam, learning rate 0,001, batch size 32, perbandingan data training dan data testing dengan berbagai rasio. Semua percobaan training dilakukan menggunakan Google Colaboratory.

DAFTAR PUSTAKA

- [1] Li, J., Li, C., Han, J., Shi, Y., Bian, G., and Zhou, S. "Robust Hand Gesture Recognition Using HOG-9ULBP Features and SVM Model". Zhengzhou, China. 2022.
- [2] Sharma, A., Mittal, A., Singh, S., Awatramani, V. "Hand Gesture Recognition using Image Processing and Feature Extraction Techniques". Procedia Computer Science, 2020, 173, hal. 181–190.
- [3] Asif, A., Waris A., Gilani, S., Jamil, M., Ashraf, H., Shafique, M., dan Niazi, I. "Performance Evaluation of Convolutional Neural Network for Hand Gesture Recognition Using EMG". Sensors, 2020, 20, hal. 1-11.
- [4] Hubert, P., Sudaryono, R., dan Suhartono, D. "Classifying Promotion Images Using Optical Character Recognition and Naïve Bayes Classifier". Universitas Bina Nusantara, Jakarta, Procedia Computer Science, 2021, 179, hal. 498–506.
- [5] Abdullah, D. dan Abdulazeez, A. "Machine Learning Applications based on SVM Classification: A Review". Qubahan Academic Journal, 2021, hal. 81-90.
- [6] Gill, H., Khalaf, O., Alotaibi, Y., dan Alghamdi, S. "Multi-Model CNN-RNN-LSTM Based Fruit Recognition and Classification". Intelligent Automation & Soft Computing, 2022, bab 33, no 1, hal. 638-650.
- [7] Rajan, S., Chenniappan, P., Devaraj, S., Madian, N. "Novel deep learning model for facial expression recognition based on maximum boosted CNN and LSTM". 2020.
- [8] Dhiman, G., Juneja, S., Viriyasitavat, W., Mohafez, H., Hadizadeh, M., Islam, M., Bayoumy, I., dan Gulati, K. "A Novel Machine-Learning-Based Hybrid CNN Model for Tumor Identification in Medical Image Processing". Sustainability, 2022, 14, 1447, hal. 1-13.
- [9] Neethu, P., Suguna, R., dan Sathish, D. "An Efficient Method for Human Hand Gesture Detection and Recognition Using Deep Learning Convolutional Neural Networks". 2020.
- [10] Sherstinsky, A. "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network". Elsevier journal "Physica D Nonlinear Phenomena", 2023, bagian 404 hal. 1-43.
- [11] Sahoo, J., Prakash, A., Plawiak, P., dan Samantray, S. "Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network". Sensors, 2020, 22, 706, hal. 1-14.
- [12] Zhao, H. H., dan Liu, H. "Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition". Granular Computing, 2020, 5(3), hal. 411-418.
- [13] Lecun, Y., Bottou, L., Bengio, Y., dan Haffner, P. "Gradient-based learning applied to document recognition". Proceedings of the IEEE, 1998, 86(11), hal. 2278-2324.
- [14] Ewe, E., Lee, C., Kwek, L., dan Lim, K. "Hand Gesture Recognition via Lightweight VGG16 and Ensemble Classifier". Melaka, Malaysia, 2022.
- [15] Ding W, Wang X, Liu H, dan Hu B. "An empirical study of shape recognition in ensemble learning context. In: International conference on wavelet analysis and pattern recognition". Chengdu, 2018, hal. 256–261.
- [16] Goodfellow, I. dkk. "Deep Learning". MIT Press, 2016.
- [17] Sani, A. dan Rahmadinni, S. "Deteksi sikap tangan Berbasis Pengolahan Citra". Politeknik Negeri Batam, Jurnal Rekayasa ElektriKa, 2022, Vol. 18, no. 2.
- [18] Arifah, I., Fajri, F., dan Pratamasunu, G. "Deteksi Tangan Otomatis Pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode YOLO Dan CNN". Journal of Applied Informatics and Computing (JAIC), 2022, bab 6, no. 2.
- [19] Pratamasunu, G., Fajri, F., dan Sari, P. "Deteksi Tangan Otomatis Pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode Deep

Gated Recurrent Unit (GRU)". Jurnal Politeknik Caltex Riau, 2022.

[20] Bhushan, S., Alshehri, M., Keshta, I., Chakraverti, A., Rajpurohit, J., dan Abugabah, A. "An Experimental Analysis of Various Machine Learning Algorithms for Hand Gesture Recognition". Saudi Arabia, 2022.