Implementasi Convolutional Neural Network dalam Mengenali Image Angka Tulisan Tangan

Andara Najla Jilan*, Drs. M. Yusuf Fajar, Erwin Harahap

Prodi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*andara_10060220025@unisba.ac.id,myusuffajar111@gmail.com,erwin2h@unisba.ac.id

Abstract. Advances in information technology and artificial intelligence, particularly in the field of machine learning, have had a significant impact on various aspects of daily life. Machine learning's ability to learn from data and experience, identify patterns, make decisions, and even perform tasks previously only possible for humans, has revolutionized numerous industries. Within the realm of machine learning, deep learning stands out as a prominent approach. Deep learning employs artificial neural networks with intricate structures to comprehend and process data. One popular deep learning algorithm is the Convolutional Neural Network (CNN). CNNs have found extensive applications, especially in image recognition tasks. By leveraging CNNs, computers can accurately identify objects and patterns within images. The deep learning process underlying CNNs involves complex mathematical computations. It begins with feature extraction to maximize the significance of features from images, which are transformed into matrices. Subsequently, the available data is trained to develop a highly accurate CNN model. This research delves into the mathematical underpinnings of how deep learning, specifically using CNN algorithms, can recognize handwritten digit images. The employed CNN algorithm achieves an impressive accuracy of 99% in recognizing handwritten digit images.

Keywords: *CNN, deep learning, machine learning, image.*

Abstrak. Perkembangan teknologi informasi dan kecerdasan buatan, terutama dalam bidang machine learning, telah memiliki dampak signifikan pada berbagai aspek kehidupan sehari-hari. Kemampuan machine learning memungkinkan komputer untuk belajar dari data dan pengalaman, mengidentifikasi pola, membuat keputusan, dan bahkan melakukan tugas-tugas yang sebelumnya hanya bisa dilakukan oleh manusia.Dalam dunia machine learning, terdapat salah satu pendekatan yaitu deep learning. Deep learning adalah artificial intelligence yang menggunakan neural networks dengan struktur yang lebih kompleks untuk memahami dan memproses data. Salah satu algoritma deep learning yang populer adalah Convolutional Neural Network (CNN). CNN telah digunakan dalam berbagai permasalahan, terutama dalam permasalahan mengenai pengenalan image. Penggunaan CNN memungkinkan komputer untuk mengenali objek dan pola dalam gambar dengan akurasi yang tinggi. Proses deep learning yang terjadi di dalam algoritma CNN merupakan perhitungan matematika. Diawali dari proses feature extraction untuk memaksimalkan fitur-fitur penting dari image yang sudah diubah menjadi matriks dan melatih data yang ada hingga didapatkan model dengan algoritma CNN berakurasi tinggi. Penelitian ini berfokus pada bagaimana deep learning dengan menggunakan algoritma CNN dapat mengenali image angka tulisan tangan. Algoritma CNN yang digunakan dapat menghasilkan model berakurasi 99% dalam mengenali image angka tulisan tangan.

Kata kunci: *CNN, deep learning, machine learning, image*.

Corresponding Author
Email: myusuffajar111@gmail.com

A. Pendahuluan

Perkembangan teknologi informasi dan kecerdasan buatan, khususnya dalam ranah kecerdasan buatan, telah membawa dampak signifikan pada berbagai aspek kehidupan sehari-hari. Kecerdasan buatan dibangun menggunakan algoritma komputer yang bermacam-macam seperti *Natural Language Processing, Computer Vision, dan Machine Learning*. *Machine learning* adalah cabang dari algoritma komputer yang masih terus berevolusi yang dirancang untuk meniru kecerdasan manusia dengan belajar dari lingkungan sekitarnya (El Naqa, I., Murphy, 2015). *Machine learning* memungkinkan komputer untuk belajar dari data dan pengalaman, memungkinkannya untuk mengidentifikasi pola, membuat keputusan, dan bahkan melakukan tugas-tugas yang biasanya memerlukan kecerdasan manusia.

Deep learning adalah sub bidang *dari machine learning* yang menggunakan neural networks dengan struktur yang lebih dalam (*deep neural networks*) untuk memahami dan memproses data. *Deep learning* adalah pendekatan pemodelan berbasis komputer yang disusun oleh banyak lapisan (*layer*) pemrosesan yang digunakan untuk memahami representasi data (Mishra et al., 2021). *Deep learning* menggunakan paling sedikit dua lapis *neural network* dalam prosesnya. Ada beberapa jenis deep learning seperti *Feed Forward Neural Network, Multi Layer Perceptron, dan Convolutional Neural Network*.

Convolutional Neural Network (CNN) adalah pendekatan deep learning yang banyak digunakan untuk menyelesaikan permasalahan yang kompleks(Achyar & Rohaeni, 2022). Algoritma CNN akan menggunakan *deep learning* untuk mempelajari nilai image berbentuk matriks dan mengklasifikasikan *image-image* tersebut ke dalam kelasnya masing-masing.Image angka tulisan tangan dipakai karena angka merupakan gambar yang memiliki karakteristik dan makna tertentu, sehingga model yang dihasilkan dapat dimanfaatkan untuk *image-image* yang lain (Kumar et al., 2021).

Berdasarkan uraian di atas maka rumusan masalah pada penelitian ini adalah seberapa akurat model CNN yang dihasilkan oleh proses *deep learning*. Tujuan dari penelitian ini adalah mengetahui keakuratan model CNN yang dihasilkan oleh algoritma.

B. Metodologi Penelitian

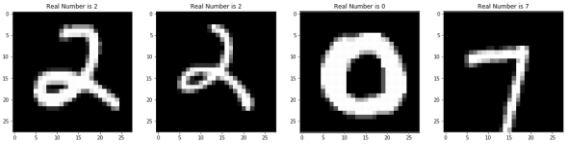
Penelitian ini menggunakan data gambar tulisan tangan dari *database* MNIST (LeCun, n.d.) dan algoritma CNN pada python. Struktur CNN yang digunakan meliputi *preprocessing data, feature extraction, forward propagation, backward propagation*, dan penghitungan akurasi melalui *confusion matrix*.

Pada penelitian, digunakan jumlah *batch* sebesar 64 dan *epoch* sebesar 5, 10, 15, 20 dengan asumsi *stride* adalah 1 dan tanpa *padding* (Lin, 2022).

C. Hasil Penelitian dan Pembahasan

Akurasi Model Convolutional Neural Network berdasarkan Confusion Matrix

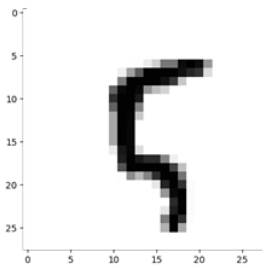
Langkah awal sebelum memasuki CNN adalah mempersiapkan data atau *preprocessing* data(Pratiwi et al., 2020). Pada tahap ini, dilakukan pelabelan untuk memisahkan data yang akan digunakan sebagai data *training* dan data *testing*. Dalam penelitian ini, data secara keseluruhan berjumlah 70000 data dengan 60000 data *training* dan 10000 data *testing* yang disebar ke dalam 9 kelas dari 0 sampai 9. Kedua data kemudian akan dilabeli sebagai (x_train, y_train) untuk data *training* dan (x_test, y_test) untuk data *testing*. Data training hasil pelabelan dapat dilihat pada Gambar 1.



Gambar 1. Pelabelan Data

Setelah pelabelan data selesai, kemudian akan dilakukan pengecekan jumlah data. Jumlah data

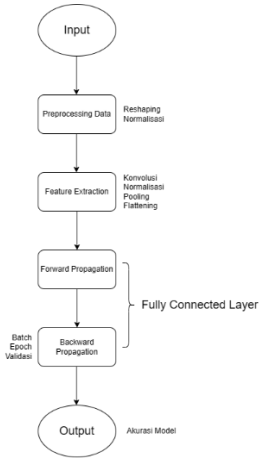
dipisahkan ke dalam 10 kelas sesuai dengan angkanya dari 0 sampai 9. Setelah dipisahkan, dilakukan proses *reshaping* dan normalisasi. Proses *reshaping* dijalankan untuk menyeragamkan ukuran semua *image*. Algoritma CNN menerima ukuran *image* sebesar 28x28 pixel, maka semua *image* harus berukuran 28x28 pixel. Untuk mempermudah dan mempercepat CNN dalam mengolah data, maka skala *input* akan diubah. Hal ini dilakukan dengan membagi *value* warna dengan 255 agar *value* berada pada rentang 0-1. Warna putih akan menghasilkan *value* sebesar 0 dan warna hitam menghasilkan *value* sebesar 1. Hasil *preprocessing* data ditunjukkan pada Gambar 2.



Gambar 2. Hasil *Preprocessing Data*

Dalam CNN, terdapat tiga proses yang terjadi dalam algoritma pengenalan *image* yaitu: *feature extraction, forward propagation*, dan *backward propagation*. *Feature extraction* akan menghasilkan output yang menjadi input pada *forward propagation*. Dalam *feature extraction*, *image* akan mengalami proses konvolusi, normalisasi, dan *pooling* untuk menghasilkan *feature map*. *Feature map* ini kemudian akan melalui proses *flattening* yang mengubahnya dari matriks bujur sangkar menjadi vektor berdimensi satu.

Hasil *flattening* ini kemudian akan menjadi input dalam proses *forward propagation*. Setelah melewati proses *forward propagation, output* dari proses ini kemudian menjadi *input* untuk proses *backward propagation*. Perjalanan melewati dua kali iterasi ini disebut dengan satu *epoch*. *Image* akan melewati beberapa *epoch* sesuai dengan jumlah *epoch* yang diberikan. Alur implementasi CNN pada penelitian ini dapat dilihat pada Gambar 3.



Gambar 3. Alur Implementasi CNN

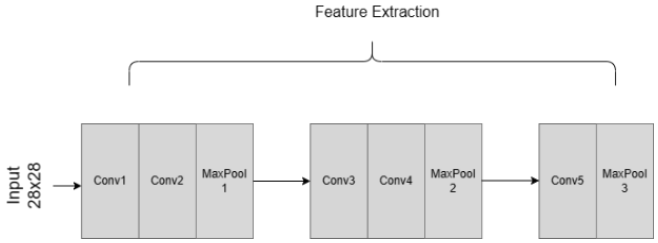
Dalam *feature extraction*, data pertama-tama akan melalui layer konvolusi. Pada *layer* konvolusi, *image* akan dibaca sebagai matriks dan dioperasikan dengan *filter* untuk menghasilkan *output image* berupa matriks yang baru. *Output* hasil konvolusi ini kemudian akan diaktivasi agar tidak ada nilai yang negatif sebelum melalui proses *pooling*. Proses *pooling* berjalan sebanyak 3 kali setiap dua kali konvolusi menggunakan *Max Pooling*. Proses *feature*

extraction bertujuan untuk memperbanyak jumlah data dengan membuang informasi yang tidak relevan dan mengecilkan area kerja algoritma hanya pada area yang berisi informasi-informasi penting saja. Perhitungan yang terjadi di dalam proses feature extraction dapat dilihat pada formula berikut.

Tabel 1. Proses Feature Extraction

Input	Hidden Layer 1	Hidden Layer 2	Softmax	Output
$\begin{bmatrix} a_{11} \\ \vdots \\ a_{2561} \end{bmatrix}$	$[a_{11}] \cdot \frac{\partial softmax(x_i)}{\partial x_i}$	$[a_{11}] \cdot \frac{\partial softmax(x_i)}{\partial x_i}$	$[a_{11}] \cdot softmax_1$	O_1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$[a_{2561}] \cdot \frac{\partial softmax(x_i)}{\partial x_i}$	$[a_{2561}] \cdot \frac{\partial softmax(x_i)}{\partial x_i}$	$[a_{2561}] \cdot softmax_{256}$	O_{10}

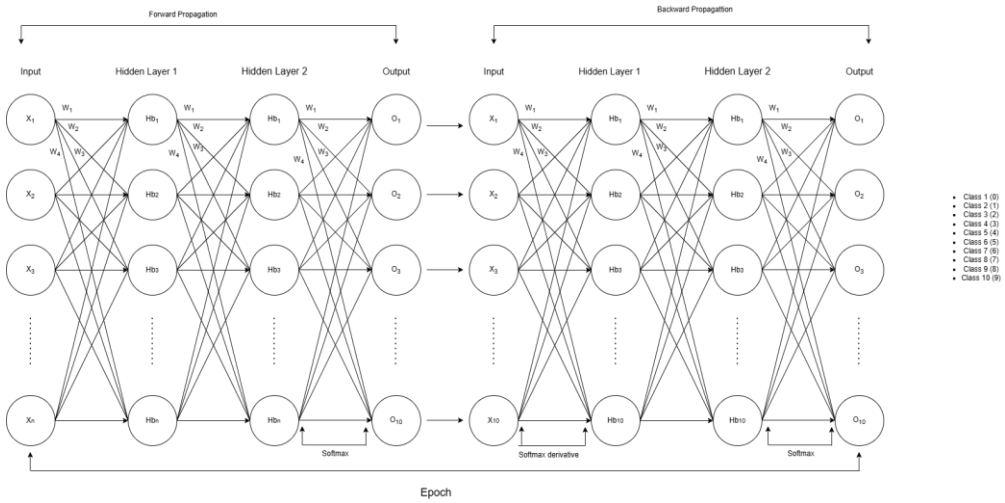
Pada Gambar 4 ditunjukkan fase-fase di dalam *feature extraction*.



Gambar 4. *Feature Extraction*

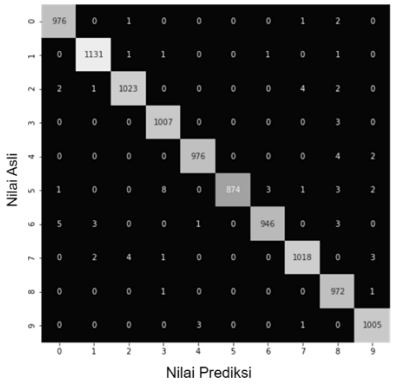
Pada proses *pooling* terakhir, fitur menjadi matriks berdimensi 1x1 dengan kedalaman fitur (*depth*) sebesar 256. *Output* hasil *pooling* kemudian akan dijadikan vektor satu dimensi (*flatten*). Karena matriks *output* hasil konvolusi sudah berdimensi 1x1x256, maka proses ini hanya mengubah bentuk matriks ke dalam bentuk vektor saja menjadi 1x256. Setelah selesai melalui proses *flatten*, maka *output* dapat diolah lebih lanjut ke dalam *Fully Connected Layer* (FCL).

Proses *deep learning* terjadi di dalam FCL. Image akan mengalami dua kali proses iterasi, yaitu *forward propagation* dan *backward propagation*. Satu kali siklus ini dinamakan satu *epoch*. Arsitektur FCL ditunjukkan pada Gambar 5.



Gambar 5. *Fully Connected Layer*

Setelah proses deep learning selesai, akurasi dapat dilihat melalui confusion matrix. Akurasi dihitung per kelas angka kemudian dihitung secara keseluruhan. Gambar 6 menunjukkan *confusion matrix* yang dihasilkan oleh model.



Gambar 6. *Confusion Matrix*

Pada confusion matrix, terdapat empat parameter yang digunakan untuk menghitung akurasi yaitu TP (*True Positive*), TN (*True Negative*), FP (*False Positive*), dan FN (*False Negative*). Parameter ini memiliki kriteria sebagai berikut:

1. Nilai TP adalah hasil model mengenali nilai secara tepat.
2. Nilai TN adalah banyaknya nilai yang salah diprediksi.
3. Nilai FP adalah banyaknya prediksi yang meleset.
4. Nilai FN adalah banyaknya label dan prediksi lain.

Akurasi adalah nilai sampel yang benar di antara semua sampel. Akurasi dianggap akurat atau baik ketika persentase *loss* nya kurang dari 5% (Science Ready, 2024). Nilai akurasi dapat dicari dengan formula berikut:

$$Akurasi: \frac{TP + TN}{TP + FP + TN + FN}$$

Tabel 2. Parameter Akurasi

Angka	TP	FP	FN	TN	Akurasi
0	976	8	4	9012	99%
1	1131	6	4	8859	99%
⋮	⋮	⋮	⋮	⋮	⋮
9	1005	8	4	8983	99%

Sumber: Parameter Akurasi, 2024.

Tabel 1 menunjukkan parameter penghitungan akurasi per kelas berdasarkan *confusion matrix*. Pada angka 0, ada 976 image angka 0 yang dikenali secara tepat, 8 *image* lain yang dikenali sebagai 0, 4 *image* angka 0 yang dikenali sebagai angka lain, dan 9012 angka lainnya. Akurasi yang dihasilkan adalah sebesar 99%. Hal ini menunjukkan bahwa model dapat mengenali angka 0 dengan baik dengan *loss* kurang dari 5%.

Tabel 2. Akurasi Model CNN Berdasarkan Kelas

Karakter	Berhasil	Gagal	Total *Image*	Akurasi
0	976	8	984	99%
1	1131	6	1137	99%
2	1023	6	1029	99%
3	1007	3	1010	98%
4	976	4	980	99%
5	874	0	874	100%
6	946	4	950	99%
7	1018	7	1025	99%
8	972	18	990	98%
9	1005	8	1013	99%

Sumber: Akurasi per Karakter, 2024.

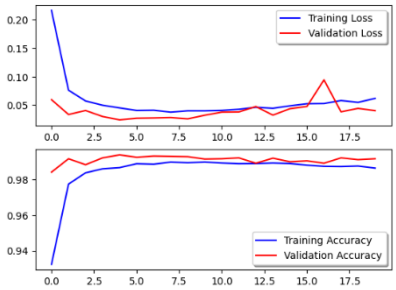
Berdasarkan Tabel 2 di atas, dapat dilihat bahwa model memiliki performa yang baik dengan akurasi berada pada 98%-100%. Akurasi terbesar ada pada angka 5 dengan akurasi sebesar 100%. Hal ini menyatakan bahwa model dapat mengenali angka 5 secara sempurna dan tidak dibaca sebagai angka lain.

Tabel 3. Akurasi Model CNN Berdasarkan Epoch

Batch	*Epoch*	*Accuracy*
64	5	99.14%
	10	99.34%
	15	99.21%
	20	99.17%

Sumber: Akurasi Model CNN Berdasarkan *Epoch*, 2024.

Dari Tabel 3, akurasi model berdasarkan *epoch* menghasilkan jumlah akurasi yang berbeda-beda. Akurasi model tetap berada pada angka 99% dengan perbedaan kecil pada angka di belakang koma. Ini menunjukkan bahwa model CNN yang dihasilkan cukup baik dan akurat dalam mengenali *image* angka tulisan tangan. Akurasi terbaik berada pada 10 *epoch*. Dapat dilihat berdasarkan tabel bahwa semakin banyak *epoch*, akurasi cenderung menurun. Grafik *loss* dan akurasi yang terjadi pada CNN dapat dilihat pada Gambar 7.



Gambar 7. Grafik *Loss* dan *Accuracy*

Pada Gambar 7, dapat dilihat bahwa *loss* menurun, tetapi mengalami kenaikan saat jumlah *epoch* semakin besar. Bahkan *validation loss* mengalami kenaikan yang signifikan (spike) setelah 15 *epoch* kemudian turun kembali sebelum epoch ke 17.5. *Training accuracy* dan *validation accuracy* terlihat naik pada *epoch* kecil, tetapi kemudian mulai stagnan seiring jumlah *epoch*nya membesar.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Model memiliki akurasi yang cukup baik sebesar 99.34% pada 10 epoch.
2. Kinerja dari model cukup baik dengan akurasi pengenalan per karakter sebesar 98%-100%
3. Penelitian ini belum bisa menunjukkan pengaruh parameter *epoch* terhadap akurasi model.
4. Penelitian selanjutnya disarankan untuk mengambil data lain selain MNIST dengan rentang *epoch* yang lebih besar agar lebih terlihat pengaruh *epoch* terhadap model.

Acknowledge

Ucapan terima kasih disampaikan kepada Bapak M. Yusuf Fajar dan Bapak Erwin Harahap sebagai dosen pembimbing yang telah membimbing peneliti dari awal hingga terselesaikannya penelitian ini. Ucapan terima kasih juga disampaikan kepada seluruh pihak yang turut membantu dalam penelitian ini.

Daftar Pustaka

[1] Achyar, G., & Rohaeni, O. (2022). Penggunaan Hybrid K-Means dan General Regression Neural Network untuk Prediksi Harga Saham Indeks LQ45. *Jurnal Riset Matematika*, 111–120. https://doi.org/10.29313/jrm.v2i2.1193

[2] El Naqa, I., Murphy, M. J. (2015). What Is Machine Learning?. In: El Naqa, I., Li, R., Murphy, M. (eds) Machine Learning in Radiation Oncology. In *Machine Learning in Radiation Oncology*. Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-319-18305-3_1

[3] Kumar, S. N., Ahilan, A., Fred, A. L., & Kumar, H. A. (2021). *Chapter 6 - ROI extraction in CT lung images of COVID-19 using Fast Fuzzy C means clustering* (V. E. Balas, O. Geman, G. Wang, M. Arif, & O. B. T.-B. E. T. for M. for P. with C.-19 Postolache, Eds.; pp. 103–119). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-824473-9.00001-X

[4] LeCun, Y. (n.d.). *MNIST database*.

[5] Lin, R. (2022). Analysis on the Selection of the Appropriate Batch Size in CNN Neural Network. *2022 International Conference on Machine Learning and Knowledge Engineering (MLKE)*, 106–109. https://doi.org/10.1109/MLKE55170.2022.00026

[6] Mishra, R. K., Reddy, G. Y. S., & Pathak, H. (2021). The Understanding of Deep Learning: A Comprehensive Review. *Mathematical Problems in Engineering*, *2021*, 5548884. https://doi.org/10.1155/2021/5548884

[7] Pratiwi, B., Handayani, A., & Sarjana. (2020). Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi WSN Menggunakan Confusion Matrix. *Jurnal Informatika UPGRIS*, *6*(2), 66–75.

[8] Science Ready. (2024). *Validity, Accuracy and Reliability Explained with Examples*. https://scienceready.com.au/pages/validity-accuracy-and-reliability