

PENGENALAN KARAKTER SIBI MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK

Fadhil Maulana Suryawan¹⁾ Jeanny Pragantha, Ir., M.Eng²⁾ Teny Handayani, S.Kom., M.Komⁿ⁾

¹⁾²⁾³⁾ Teknik Informatika, FTI, Universitas Tarumanagara

Jl. Letjen S Parman no 1, Jakarta 11440 Indonesia

¹⁾ email : fadhil.535170107@stu.untar.ac.id, ²⁾ email : jeannyp@fti.untar.ac.id, ³⁾ email : tenyh@fti.untar.ac.id

ABSTRACT

Communication between humans is an important thing for the activities of daily life. However, Humans were created with the advantages and disadvantages of each person and one of them is the difficulty of establishing or communicating and interacting for people who are deaf and speech impaired. Meanwhile, what solutions can be given about these shortcomings. So that there is no gap in society.

This has led to the development of sign language recognition systems, which can automatically translate sign language into text and speech with effective pre-processing and accurate sign classification. According to recent developments in the field of deep learning, neural networks may have broad implications and implementations for sign language analysis. In the proposed system, Convolutional Neural Network (CNN) is used to classify sign language images because convolutional networks are faster in feature extraction and image classification than other classifiers.

CNN architecture is carried out in 3 stages, 25 epochs, 50 epochs and 100 epochs. Based on the experiments conducted, the accuracy value obtained continues to increase in each stage, starting from 91.03%, 92.69% to the highest accuracy value in the training process of 94.25%. Likewise, the data prediction process also increases in each stage, starting from 90.62%, 93.75% until the highest accuracy value in data prediction is same 95.83%.

Key words

Convolutional Neural Network, Deep learning, Detection, Sign Language

1. Pendahuluan

Bahasa isyarat adalah cara paling ekspresif untuk komunikasi antara orang-orang tunarungu, di mana informasi sebagian besar disampaikan melalui gerakan tangan/lengan. Bahasa isyarat adalah bahasa alami lengkap dengan aturan bahasa dan vokabularinya.[1]. Bahasa isyarat tidaklah universal dan tidak bisa saling

dimengerti satu sama lain,[2] meskipun ada beberapa kemiripan menonjol di dalam bahasa isyarat.

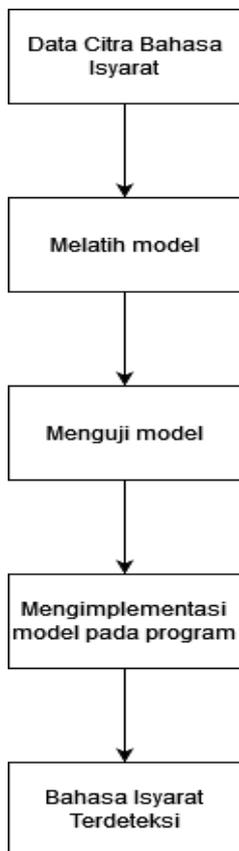
Seperti bahasa mana pun, bahasa isyarat antar satu negara dengan negara lainnya berbeda-beda. Meskipun bahasa isyarat merupakan sebuah cara yang alami untuk bertukar informasi dengan orang tunarungu, bahasa isyarat dapat bervariasi dari satu bagian dunia ke bagian lain. Oleh karena itu, orang tuli dan bisu mengalami kesulitan dalam berkomunikasi dengan orang-orang pada umumnya karena orang normal tidak memahami bahasa isyarat. Karena Bahasa isyarat tidak sering digunakan oleh orang umum tata bahasanya juga berbeda dari satu negara ke negara lain. Di Indonesia terdapat dua sistem isyarat, yakni Bahasa isyarat Indonesia (BISINDO) dan Sistem Isyarat Bahasa Indonesia (SIBI), tetapi yang merupakan bahasa isyarat alamiah adalah BISINDO. Pada kenyataannya, bahasa isyarat bervariasi.[3]

Ada penelitian sebelumnya tentang Angka Sistem Isyarat Bahasa Indonesia Dengan Menggunakan Metode Convolutional Neural Network [4] Dalam paper ini membahas pengenalan isyarat angka SIBI dengan menggunakan metode Convolutional Neural Network (CNN). Arsitektur CNN yang digunakan adalah arsitektur LeNet. Arsitektur CNN diproses dalam 3 tahap, 25 epoch, 50 epoch dan 100 epoch. Berdasarkan percobaan yang dilakukan nilai akurasi yang di dapat terus meningkat dalam tiap tahapnya, mulai 67.66%, 89.44% sampai nilai akurasi tertinggi dalam proses training sebesar 96.44%. Begitupun dalam proses prediksi data juga mengalami kenaikan dalam tiap tahapnya, mulai 79.23%, 90.45% sampai di dapatkan nilai akurasi tertinggi dalam prediksi data 98.89%. Dari total 90 data testingset, hanya sekali keasalahan dalam prediksi data.

Hal ini menyebabkan Sistem Pengenalan Bahasa isyarat memainkan peran penting dalam mengembangkan sistem interaksi komputer manusia berbasis gerakan. Dimana sistem dapat menerjemahkan bahasa isyarat secara otomatis ke dalam teks dan karakter dengan *preprocessing* yang efektif dan klasifikasi tanda yang akurat. Dalam sistem yang diusulkan, *Convolutional Neural Network* (CNN) digunakan untuk mengklasifikasikan gambar bahasa isyarat karena jaringannya yang lebih cepat dalam ekstraksi fitur dan klasifikasi gambar dibanding sistem pengklasifikasian lainnya.

2. Metode Penelitian

Pada perancangan sistem yang dibangun menggunakan Convolutional Neural Network (CNN). Sistem dibangun untuk mendeteksi citra bahasa isyarat pada citra dengan tanpa pewarnaan. Model yang dihasilkan akan diimplementasikan pada program. Tahapan sistem dapat dilihat pada gambar 1.



Gambar 1. Alur sistem

2.2 Dataset

Dataset berupa foto simbol alphabet SIBI yang dibuat oleh model. Dataest diambil dari web kaggle denga banyak dataset yang didapat sebanyak 55.120 gambar. Banyak kelas yang dihasilkan sebanyak 26 kelas angka mulai dari A – Z. Contoh alfabet SIBI dapat dilihat pada Tabel 1.

Tabel 1. Daftar kelas huruf SIBI (Lanjutan)

No.	Nama	Isyarat
1.	A	
2.	B	

3.	C	
4.	D	
5.	E	
6.	F	
7.	G	
8.	H	
9.	I	
10.	J	
11.	K	
12.	L	
13.	M	
14.	N	

Tabel 1. Daftar kelas huruf SIBI (Lanjutan)

No.	Nama	Isyarat
15.	O	
16.	P	
17.	Q	
18.	R	
19.	S	

20.	T	
21.	U	
22.	V	
23.	W	
24.	X	
25.	Y	
26.	Z	

Contoh data citra sibi pada dataset yang diambil dapat dilihat pada Gambar 2.



Gambar 2. Bentuk dataset

2.3 Convolutional neural network

Convolutional Neural Network dibuat untuk memproses data masukan yang terbagi menjadi beberapa array, misalnya gambar berwarna yang terdiri dari tiga array 2D yang memiliki intensitas piksel dalam tiga warna. Banyak data dalam bentuk array ganda 1D untuk sinyal dan urutan, termasuk bahasa 2D untuk gambar atau audio, dan 3D untuk video atau gambar volumetrik. Ada empat ide utama di balik ConvNets yang memanfaatkan properti sinyal alami koneksi lokal, bobot bersama, penyatuan, dan penggunaan banyak lapisan.[5]

Model ditraining dengan menggunakan epoch sebanyak 30 kali dan batch size-nya 32. Pada percobaan ini, terdapat dua jenis percobaan yaitu menggunakan layer 2 dan layer 3 tanpa pewarnaan. Pelatihan model untuk citra simbol bahasa isyarat menggunakan 2 convolution

layer, 2 pooling layer dengan ukuran 2x2, 1 flatten layer dan 2 dense layer. Fungsi aktivasi yang digunakan yaitu ReLu dan Softmax, dengan filter sebanyak 16, 32 dan 64 dengan ukuran kernel 3x3. Konfigurasi lapisan untuk citra preperat tanpa pewarnaan dapat dilihat pada Tabel 2

Tabel 2. Konfigurasi Lapisan Citra Simbol Isyarat layer 2

Layer	Paramater	Output Shape
conv2d (Conv2D)	64, (3,3)	(None, 64, 64, 64)
conv2d_1 (Conv2D)	64, (3,3)	(None, 64, 64, 64)
Activation (Activation)	Relu	(None, 64, 64, 64)
max_pooling2d (MaxPooling2D)	64, (3,3)	(None, 32, 32, 64)
dropout (Dropout)	0.25	(None, 32, 32, 64)
conv2d_2 (Conv2D)	64, (3,3)	(None, 32, 32, 64)
conv2d_3 (Conv2D)	64, (3,3)	(None, 32, 32, 64)
Activation_1 (Activation)	Relu	(None, 32, 32, 64)
max_pooling2d_1 (MaxPooling2D)	64, (3,3)	(None, 16, 16, 64)
Dropout_1 (Dropout)	0.25	(None, 16, 16, 64)
Flatten (Flatten)	-	(None, 4096)
Dense (Dense)	128	(None, 128)
Dropout_3 (Dropout)	0.5	(None, 128)
Dense_1 (Dense)	128	(None, 128)
Dropout_4 (Dropout)	0.5	(None, 128)
Dense_2 (Dense)	26	(None, 26)
activation_3 (Activation)	Softmax	(None, 26)

Pelatihan model untuk citra simbol bahasa isyarat menggunakan 3 convolution layer, 3 pooling layer dengan ukuran 2x2, 1 flatten layer dan 2 dense layer. Fungsi aktivasi yang digunakan yaitu ReLu dan Softmax, dengan filter sebanyak 16, 32 dan 64 dengan ukuran kernel 3x3.

Konfigurasi lapisan untuk citra preparat tanpa pewarnaan dapat dilihat pada Tabel 3

Tabel 3. Konfigurasi Lapisan Citra Simbol Isyarat layer 3

Layer	Paramater	Output Shape
conv2d (Conv2D)	64, (3,3)	(None, 64, 64)
conv2d_1 (Conv2D)	64, (3,3)	(None, 64, 64)
Activation (Activation)	Relu	(None, 64, 64)
max_pooling2d (MaxPooling2D)	64, (3,3)	(None, 32, 64)
dropout (Dropout)	0.25	(None, 32, 64)
conv2d_2 (Conv2D)	64, (3,3)	(None, 32, 64)
conv2d_3 (Conv2D)	64, (3,3)	(None, 32, 64)
Activation_1 (Activation)	Relu	(None, 32, 64)
max_pooling2d_1 (MaxPooling2D)	64, (3,3)	(None, 16, 64)
Dropout_1 (Dropout)	0.25	(None, 16, 64)
conv2d_4 (Conv2D)	64, (3,3)	(None, 16, 64)
conv2d_5 (Conv2D)	64, (3,3)	(None, 16, 64)
Activation_2 (Activation)	Relu	(None, 16, 64)
max_pooling2d_2 (MaxPooling2D)	64, (3,3)	(None, 8, 8, 64)
Dropout_2 (Dropout)	0.25	(None, 8, 8, 64)
Flatten (Flatten)	-	(None, 4096)
Dense (Dense)	128	(None, 128)
Dropout_3 (Dropout)	0.5	(None, 128)
Dense_1 (Dense)	128	(None, 128)
Dropout_4 (Dropout)	0.5	(None, 128)
Dense_2 (Dense)	26	(None, 26)
activation_3 (Activation)	Softmax	(None, 26)

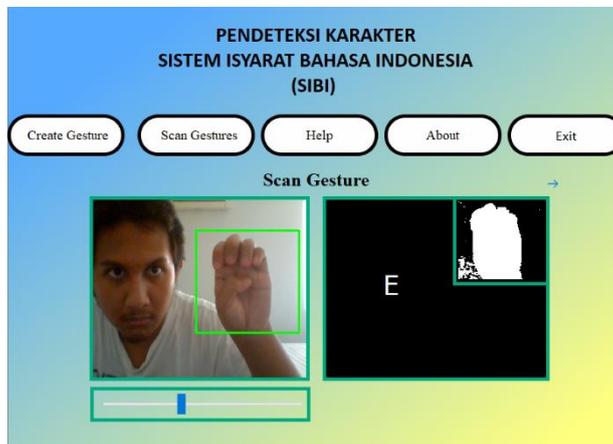
Setelah training selesai, didapatkanlah hasil training. Hasil pelatihan dari model yang dilatih diatas dapat dilihat pada tabel 4.

Tabel 4. Hasil perbandingan penggunaan layer

Dataset	Loss	Accuracy	Val_Loss	Val_Acc
2 Layer	1.1577	63.47%	0.6347	66.31%
3 Layer	2.725	91.17%	2.868	92.70%

3. Hasil Percobaan

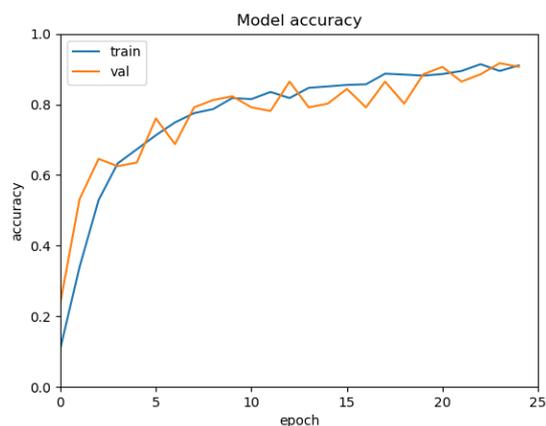
Hasil percobaan dilakukan dengan cara melihat hasil deteksi model bahasa isyarat. Model yang sudah di training akan dilakukan pengujian untuk dievaluasi hasil proses training. Contoh pendeteksian dengan hasil yang benar oleh model dapat dilihat pada Gambar 3.



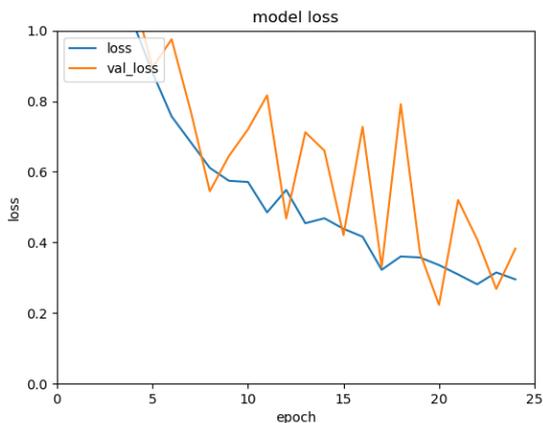
Gambar 3. Hasil pendeteksian yang benar

3.1 Hasil percobaan dalam mencari model yang terbaik

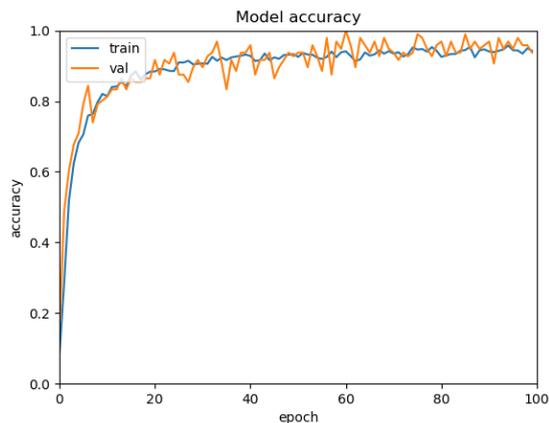
Tahap pertama dilakukan training dengan 25 epoch. Nilai akurasi yang didapat 91.03%. Nilai akurasi tersebut dirasa masih kurang maksimal. Hasil nilai validasi akurasi yang didapatkan saat proses prediksi sebesar 90.62%.



Gambar 4. Hasil akurasi dengan epoch 25

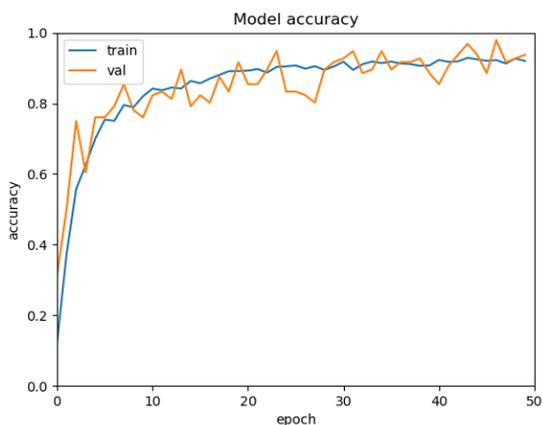


Gambar 5. Hasil Loss dengan epoch 25

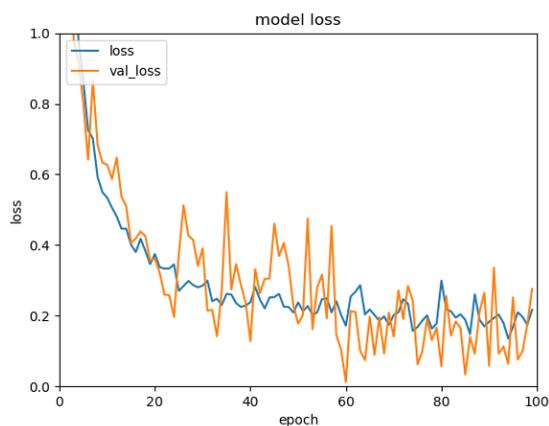


Gambar 8. Hasil akurasi dengan epoch 100

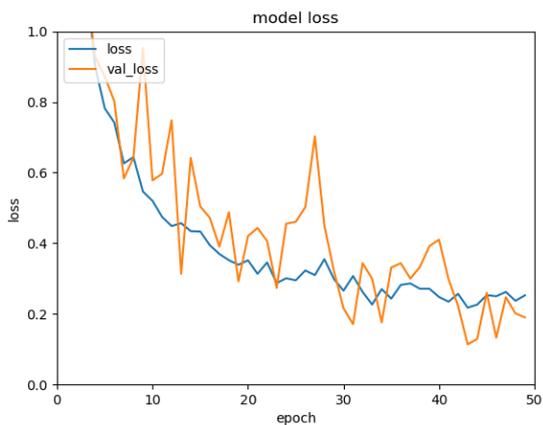
Tahap kedua dilakukan training dengan 50 epoch. Nilai akurasi yang didapat meningkat 92.69% dan nilai akurasi validasi yang didapat meningkat dari sebelumnya 93.75%.



Gambar 6. Hasil akurasi dengan epoch 50



Gambar 9. Hasil Loss dengan epoch 100



Gambar 7. Hasil Loss dengan epoch 50

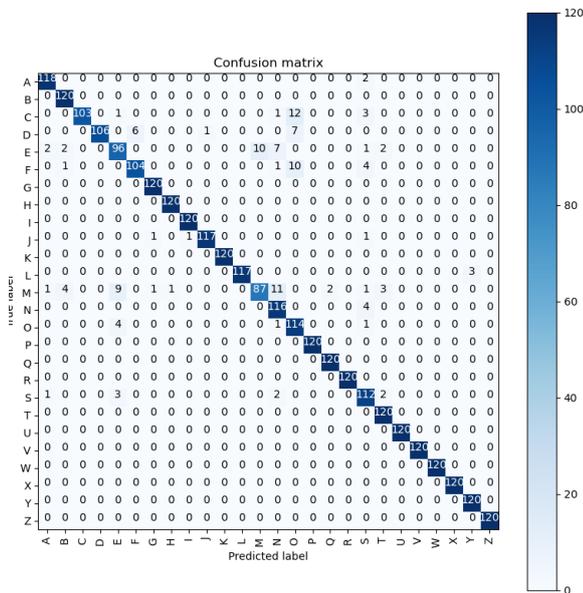
Tahap ketiga dilakukan training kembali sebanyak 100 epoch. Hasilnya nilai akurasi mencapai 94.25% dan nilai akursai validasi meningkat menjadi 95.83%.

Hasil dari preprocessing dataset training dan testing telah mengalami peningkatan pada setiap epochnya. Dalam preprocessing data testing nilai akurasi yang diperoleh dalam pada setiap tahapan mengalami kenaikan dari 91.03%, 92.69%, dan terakhir 94.25%. Begitu juga dengan lossnya yang turun pada setiap epoch. Hal ini dapat dikatakan bahwa banyak penggunaan epoch dalam training data berpengaruh pada tingkat akurasi, bisa jadi nilai akurasi akan semakin bagus. walaupun bila telalu banyak epoch yang dipakai akan mengakibatkan overfitting. Untuk mengetahui perbandingan akurasi dan loss saat proses dapat dilihat tabel hasil konfigurasi epoch pada tabel 5.

Tabel 1. Hasil Konfigurasi Epoch pada layer 3

Epoch	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
25	0.2956	91.03%	0.3827	90.62%
50	0.2526	92.69%	0.1902	93.75%
100	0.2167	94.25%	0.1748	95.83%

Untuk pengujian pengklasifikasian model digunakan metode evaluasi adalah Confusion Matrix. Model yang sudah di training akan diuji pada semua dataset, Kriteria untuk model yang dipakai merupakan model yang mempunyai keakuratan paling bagus dan validasi loss serta validasi akurasi yang baik juga. Jadi model yang digunakan adalah model dari hasil epoch 100. Untuk melihat hasil confusion matrix dapat dilihat pada 7.



Gambar 10. Hasil confusion matrix epoch 100

3.2 Hasil percobaan dengan pengguna

Dalam pengujian terhadap 5 pengguna dengan 3 kali percobaan didapatkan total hasil akurasi yang benar 89.7%. Untuk hasil percobaan benar dan salah untuk setiap percobaan dapat dilihat pada tabel 6.

Tabel 2. Hasil Percobaan dengan dengan orang lain

Percobaan	Terdeteksi	Tidak Terdeksi	Akurasi Terdeteksi
1	110	20	84.61%
2	118	12	90.76%
3	122	8	93.84%
Total	350	40	89.7%

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan terhadap pendeteksian karakter sistem isyarat bahasa Indonesia (SIBI), dapat diperoleh beberapa kesimpulan sebagai berikut:

- Dari hasil dari perbandingan percobaan 2 model CNN didapat nilai terbaik dari Training Dataset yang menggunakan CNN 3 layer dengan hasil nilai akurasi Training sebesar 98.1% sedangkan untuk validasi menghasilkan nilai akurasi sebesar 98.3%.
- Dataset untuk Training dan Testing perlu dilakukan proses Image Pre-processing menggunakan metode HSV, hal ini dilakukan agar proses Training dan Testing focus pada bentuk isyarat tangan bukan pada orangnya dan latar belakangnya.
- Akurasi benar yang didapatkan dari pengujian dengan 5 pengguna dan 3 kali percobaan didapatkan total hasil akurasi yang benar 89.7%. Dari total hasil akurasi dapat disimpulkan bahwa akurasi simbol program dapat dibilang sudah bagus.
- Jika citra karakter yang dibuat oleh pengguna saat pendeteksian berbeda dengan data yang sudah dilatih maka output yang dikeluarkan adalah huruf yang paling mirip dari data latih.

5. Saran

Adapun beberapa saran yang bisa dijadikan referensi untuk penelitian kedepannya sebagai berikut:

- Coba gunakan proses Deep Learning lainnya seperti YOLO (*You Only Look Once*), SSD (*Single Shot Detector*), dan MobileNet karena sudah mendukung algoritme deteksi objek.
- Variasi Dataset dapat ditambah kuantitas bahasa isyaratnya, seperti dapat ditambahkan simbol angka sibi ataupun bahasa isyarat yang lain.
- Mencari pendekatan yang lain untuk melatih model. Menggunakan dataset video sebagai masukan baru karena dengan menggunakan dataset video otomatis jangkauannya akan semakin luas. Apabila dataset citra SIBI hanya mendukung pengenalan huruf, dengan menggunakan dataset video mungkin saja bisa mendukung pengenalan kata atau bahkan kalimat pada SIBI.

REFERENSI

- [1] Sandler, Wendy, and Lillo-Martin, Diane. Sign Language and Linguistic Universals. Cambridge:Cambridge University Press, 2006. <https://www.cambridge.org/id/academic/subjects/languages-linguistics/sign-language/sign-language-and-linguistic-universals?format=PB>.
- [2] Perlmutter, David M. What is Sign Language?, <https://www.linguisticsociety.org/content/what-sign-language>, 24 Agustus 2021
- [3] Palfreyman, Nicholas (2015). Sign language varieties of Indonesia: A linguistic and sociolinguistic investigation (Disertasi PhD). University of Central Lancashire. <http://clock.uclan.ac.uk/12132/2/Palfreyman%20Nicholas%20Final%20e-Thesis%20%28Master%20Copy%29.pdf>. 22 Januari 2022
- [4] Bakti, Setiyo.; Bagus, Mochamad.; and Pranoto Yuliana Melita. 2021. "Pengenalan Angka Sistem Isyarat Bahasa Indonesia Dengan Menggunakan Metode Convolutional Neural Network". Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi) 3 (1):011-016.

- [5] LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* Volume 521, 436–444 (2015). <https://doi.org/10.1038/nature14539>.

Fadhil Maulana Suryawan, saat ini sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara.