

## PENERAPAN CONVOLUTIONAL NEURAL NETWORK PADA PENGENALAN BAHASA ISYARAT INDONESIA SECARA REAL-TIME

Rifki Ryan Maulana<sup>1</sup>, Abdul Kholid<sup>2</sup>, Edy Widodo<sup>3</sup>

Prodi Sistem Informasi, Universitas Satya Negara Indonesia<sup>1,2</sup>

Prodi Teknik Informatika, Universitas Pelita Bangsa<sup>3</sup>

rifki.ryan@gmail.com, abdulkholiq@usni.ac.id, edy.widodo@gmail.com

Correspondent author : abdulkholiq@usni.ac.id

Tgl. Diterima	Tgl. Revisi	Tgl. Disetujui	Tgl. Terbit
7 Agustus 2025	12 Agustus 2025	20 Agustus 2025	01 September 2025

### Abstract

Sign language is a form of visual communication used by individuals who are deaf or speech-impaired. However, many people in the general public still lack understanding of sign language, which hinders communication between people with disabilities and their surroundings. This research aims to develop a real-time alphabet translator system for Indonesian Sign Language (BISINDO), implemented as an Android application. The system utilizes a Convolutional Neural Network (CNN) model based on the MobileNetV2 architecture, which is trained to recognize 26 alphabet letters from hand gesture images sized 128x128 pixels in RGB format. The dataset was collected and processed through augmentation and divided into training, validation, testing, and evaluation sets. The model was trained using transfer learning and fine-tuning methods and then converted into TensorFlow Lite (.tflite) format for deployment on Android devices. Evaluation results show that the model achieved an average accuracy of 93% on the evaluation dataset. Testing the Android application also demonstrated good real-time performance in recognizing hand gestures. This application is expected to help bridge communication between people with disabilities and the general public through practical and accessible technology.

**Keywords** : Indonesian Sign Language, BISINDO, CNN, MobileNetV2, TensorFlow Lite, Android, image classification, deaf communication

### Abstrak

Bahasa isyarat merupakan salah satu bentuk komunikasi visual yang digunakan oleh penyandang tunarungu dan tunawicara. Namun, masih banyak masyarakat umum yang belum memahami bahasa isyarat, sehingga menghambat komunikasi antara penyandang disabilitas dengan lingkungan sekitarnya. Penelitian ini bertujuan untuk mengembangkan sebuah sistem penerjemah huruf alfabet Bahasa Isyarat Indonesia (BISINDO) secara real-time berbasis aplikasi Android. Sistem ini memanfaatkan model Convolutional Neural Network (CNN) menggunakan arsitektur MobileNetV2 yang telah dilatih untuk mengenali 26 kelas huruf alfabet berdasarkan citra tangan berukuran 128x128 piksel dalam format RGB. Dataset dikumpulkan dan diproses melalui tahap augmentasi dan pembagian menjadi data pelatihan, validasi, pengujian, dan evaluasi. Model dilatih dengan metode transfer learning dan fine-tuning, serta dikonversi ke format TensorFlow Lite (.tflite) untuk implementasi di perangkat Android. Hasil evaluasi menunjukkan bahwa model memiliki akurasi rata-rata sebesar 93% pada data evaluasi. Pengujian aplikasi Android juga menunjukkan kinerja real-time yang baik dalam mengenali isyarat tangan. Dengan adanya aplikasi ini, diharapkan dapat membantu menjembatani komunikasi antara penyandang disabilitas dengan masyarakat umum melalui teknologi yang praktis dan mudah diakses.

**Kata Kunci** : Bahasa Isyarat Indonesia, BISINDO, CNN, MobileNetV2, TensorFlow Lite, Android, klasifikasi gambar, komunikasi tunarungu

## PENDAHULUAN

Bahasa Isyarat Indonesia (BISINDO) merupakan sistem komunikasi visual yang berkembang secara alami di dalam komunitas tunarungu Indonesia dan digunakan secara luas dalam interaksi sehari-hari. Namun, keterbatasan pemahaman bahasa isyarat di kalangan masyarakat umum masih menjadi hambatan signifikan dalam komunikasi dua arah, yang berdampak pada keterbatasan akses sosial, pendidikan, dan layanan publik bagi penyandang tunarungu. Kesenjangan komunikasi ini menegaskan perlunya solusi berbasis teknologi yang bersifat inklusif, mudah diakses, dan dapat digunakan secara luas.

Kemajuan teknologi kecerdasan buatan, khususnya dalam bidang computer vision dan deep learning, telah mendorong pengembangan sistem pengenalan gestur berbasis citra. Convolutional Neural Network (CNN) dikenal memiliki kemampuan unggul dalam mengekstraksi fitur visual kompleks dari citra digital. Namun, implementasi CNN pada perangkat bergerak masih menghadapi tantangan utama berupa keterbatasan daya komputasi dan kebutuhan respons real-time.

Untuk menjawab tantangan tersebut, penelitian ini mengusulkan penerapan arsitektur CNN ringan MobileNetV2 yang diimplementasikan dalam aplikasi Android untuk pengenalan huruf alfabet BISINDO secara real-time. Sistem yang dikembangkan tidak hanya melakukan klasifikasi citra tangan, tetapi juga menyusun huruf menjadi kata dan mengonversinya menjadi suara melalui modul text-to-speech, sehingga berpotensi menjadi solusi komunikasi yang aplikatif dan praktis.

## LANDASAN TEORI

### 2.1 Tinjauan Pustaka

Berbagai penelitian sebelumnya telah mengeksplorasi pengenalan bahasa isyarat menggunakan pendekatan machine learning dan deep learning. Abdurrahman (2019) memanfaatkan sensor Kinect 2.0 untuk memperoleh data kedalaman tangan, yang memberikan akurasi cukup tinggi namun memerlukan perangkat khusus dengan biaya relatif mahal dan keterbatasan portabilitas. Pendekatan ini dinilai kurang praktis untuk penggunaan sehari-hari di masyarakat luas.

Penelitian lain oleh Sholawati et al. (2022) dan Budiman et al. (2023) menggunakan kamera konvensional dan CNN untuk pengenalan bahasa isyarat, namun masih menghadapi kendala pada fleksibilitas penggunaan, terutama terkait jarak kamera, sudut pengambilan citra, dan kondisi pencahayaan. Selain itu, sebagian besar penelitian tersebut berfokus pada pengenalan huruf atau gestur secara terpisah tanpa integrasi lanjutan dalam bentuk sistem komunikasi yang utuh.

Berbeda dengan penelitian terdahulu, penelitian ini menawarkan pendekatan yang lebih komprehensif dengan menggabungkan pengenalan huruf BISINDO berbasis CNN, penyusunan huruf menjadi kata, serta konversi teks ke suara dalam satu aplikasi Android. Selain itu, penggunaan arsitektur MobileNetV2 dan TensorFlow Lite memberikan keunggulan dari sisi efisiensi komputasi dan kesiapan implementasi real-time pada perangkat mobile.

### 2.2 Teori Umum

Deep learning merupakan pendekatan pembelajaran mesin yang memungkinkan model mempelajari representasi data secara hierarkis melalui banyak lapisan tersembunyi. Dalam konteks pengolahan citra, Convolutional Neural Network (CNN) dirancang untuk mengekstraksi fitur spasial seperti tepi, tekstur, dan bentuk melalui operasi konvolusi dan pooling.

Pada sistem pengenalan bahasa isyarat berbasis citra tangan, CNN bertugas memetakan citra input ke dalam kelas tertentu yang merepresentasikan huruf alfabet. Proses ini melibatkan beberapa tahapan utama, yaitu praproses citra, ekstraksi fitur, klasifikasi, dan evaluasi performa. Evaluasi model dilakukan menggunakan metrik kuantitatif seperti akurasi, precision, recall, dan F1-score untuk mengukur keseimbangan antara kemampuan deteksi dan tingkat kesalahan klasifikasi.

Penggunaan transfer learning memungkinkan pemanfaatan model pralatih pada dataset berskala besar seperti ImageNet, sehingga mempercepat proses pelatihan dan meningkatkan performa model pada dataset yang relatif terbatas. Pendekatan ini sangat relevan untuk pengembangan sistem pengenalan bahasa isyarat dengan jumlah data yang tidak terlalu besar.

### 2.3 Teori Khusus

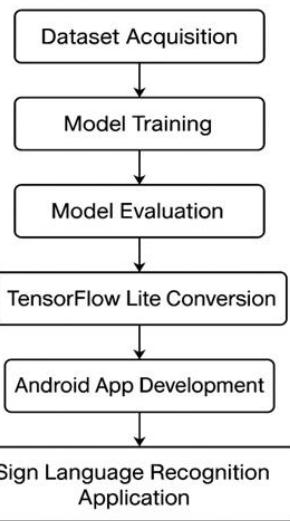
MobileNetV2 merupakan arsitektur CNN yang dikembangkan untuk mendukung pembelajaran mendalam pada perangkat dengan keterbatasan sumber daya, seperti smartphone dan perangkat IoT. Arsitektur ini mengadopsi depthwise separable convolution, yang memisahkan proses konvolusi spasial dan penggabungan kanal, sehingga secara signifikan mengurangi jumlah parameter dan beban komputasi.

Selain itu, MobileNetV2 menerapkan konsep inverted residuals dan linear bottlenecks yang mempertahankan kualitas representasi fitur sambil menekan kompleksitas model. Kombinasi ini menjadikan MobileNetV2 mampu mencapai keseimbangan antara akurasi dan efisiensi, sehingga sangat sesuai untuk aplikasi real-time berbasis Android.

Untuk mendukung implementasi di perangkat mobile, model CNN dikonversi ke format TensorFlow Lite (.tflite). TensorFlow Lite menyediakan optimasi inferensi, latensi rendah, dan konsumsi memori yang lebih kecil, sehingga memungkinkan model dijalankan secara stabil pada perangkat Android dengan spesifikasi menengah.

## METODOLOGI PENELITIAN

Metodologi penelitian ini disusun secara sistematis dan terdiri dari beberapa tahapan utama, sebagaimana ditunjukkan pada gambar berikut :



Gambar 1. Metodologi Penelitian

**Tahap pertama** adalah pengumpulan dataset (dataset acquisition), berupa citra tangan yang merepresentasikan 26 huruf alfabet Bahasa Isyarat Indonesia. Seluruh citra diseragamkan dalam ukuran  $128 \times 128$  piksel dengan format RGB untuk menjaga konsistensi input model.

**Tahap kedua** adalah praproses dan augmentasi data. Augmentasi dilakukan melalui rotasi citra, penyesuaian kecerahan, serta penambahan noise untuk meningkatkan variasi data dan kemampuan generalisasi model terhadap kondisi lingkungan yang berbeda.

**Tahap ketiga** adalah pelatihan model CNN menggunakan arsitektur MobileNetV2 dengan pendekatan transfer learning. Lapisan dasar MobileNetV2 digunakan sebagai feature extractor, kemudian ditambahkan lapisan kustom seperti GlobalAveragePooling2D, Dense, dan Dropout untuk meningkatkan performa klasifikasi dan mengurangi risiko overfitting. Selanjutnya dilakukan fine-tuning pada lapisan tertentu untuk menyesuaikan bobot model dengan karakteristik dataset BISINDO.

**Tahap keempat** adalah evaluasi model menggunakan metrik akurasi, precision, recall, F1-score, serta confusion matrix untuk menganalisis performa klasifikasi antar kelas.

**Tahap kelima** adalah konversi model ke format TensorFlow Lite, yang kemudian diintegrasikan ke dalam aplikasi Android. Aplikasi ini menggabungkan modul kamera untuk akuisisi citra real-time dan modul text-to-speech untuk menghasilkan keluaran suara berdasarkan huruf dan kata yang terdeteksi.

## 1. Perancangan dan Pelatihan Model

Pada tahap ini dilakukan perancangan arsitektur dan pelatihan model Convolutional Neural Network (CNN) untuk mendeteksi huruf-huruf pada bahasa isyarat. Model ini dibangun menggunakan pustaka TensorFlow dan Keras, yang menyediakan antarmuka tingkat tinggi untuk pengembangan model deep learning.

## 2. Perancangan Arsitektur Model

Arsitektur CNN yang dirancang mengacu pada model MobileNetV2, yaitu arsitektur CNN yang ringan dan dioptimalkan untuk perangkat dengan sumber daya terbatas, seperti smartphone. MobileNetV2 memiliki keunggulan dalam efisiensi komputasi tanpa mengorbankan akurasi secara signifikan. Berikut arsitektur CNN pada penelitian ini terdiri dari beberapa komponen utama sebagai berikut :

**Tabel 1.** Arsitektur Model

No	Lapisan	Fungsi
1	Input Layer	Menerima input citra berukuran 128x128 piksel dengan 3 channel warna.
2	Base Model: MobileNetV2	Sebagai ekstraktor fitur visual. Pada tahap awal, layer ini dibekukan (non-trainable).
3	GlobalAveragePooling2D	Mereduksi dimensi output dari feature extractor menjadi 1D.
4	Dense Layer (128 unit)	Layer fully-connected dengan aktivasi ReLU untuk pembelajaran fitur lanjut.
5	Dropout (rate 0.3)	Untuk mengurangi overfitting dengan cara mengabaikan 30% neuron saat training.
6	Output Layer	Dense layer dengan 26 neuron dan aktivasi Softmax, menghasilkan probabilitas untuk masing-masing huruf A-Z.

## 3. Strategi Pelatihan Model

Pelatihan model dilakukan dalam dua tahap, yaitu dengan pendekatan Transfer Learning dan Fine Tuning menggunakan bahasa pemrograman Python dengan bantuan Visual Studio Code. Metode ini digunakan untuk mengatasi keterbatasan jumlah data pelatihan dan mempercepat proses pelatihan model. Transfer learning memungkinkan pemanfaatan model deep learning yang telah dilatih sebelumnya (pretrained model) pada dataset besar seperti ImageNet, kemudian diadaptasi untuk tugas klasifikasi bahasa isyarat BISINDO.

## 4. Rancangan Pengujian Model

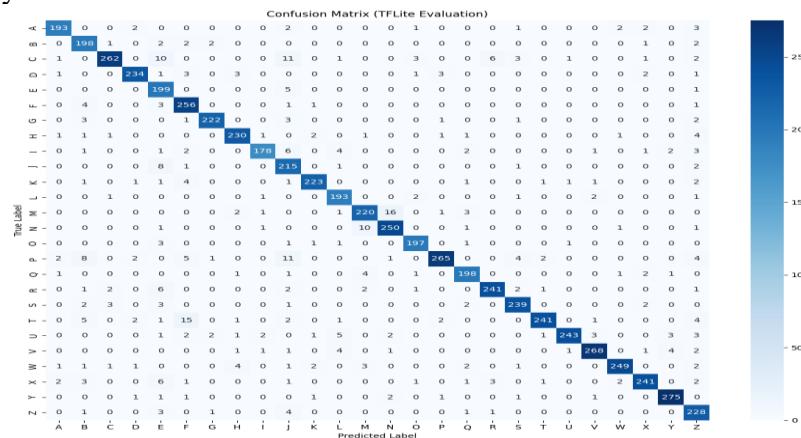
Setelah model CNN berhasil dirancang dan dilatih menggunakan dataset yang telah dipersiapkan, tahap selanjutnya adalah melakukan pengujian model untuk mengukur performa dan akurasinya dalam mengenali huruf-huruf bahasa isyarat.

Adapun rancangan pengujian model ini bertujuan untuk menilai akurasi model dalam mengklasifikasikan citra gesture, mencari misclassification, dan mengevaluasi kinerja dari model. Dimana Dataset dalam pengujian model ini dibagi menjadi Training Set (60%), Validation Set (15%), Testing Set (15%), dan Evaluation Set (10%).

Sementara untuk Metode Pengujian Model ini dilakukan dalam dua tahap; pada tahap pertama Pengujian Model (.h5) dengan menggunakan Bahasa Pemrograman Python untuk mencari evaluasi berupa Akurasi, Confusion Matrix, Classification Report (precision, recall, f1-score); kemudian pada tahap kedua Pengujian Model (.tflite) dilakukan dengan Bahasa Pemrograman Python yang menguji dataset untuk menguji ulang model setelah di konversi, sehingga terpantau apakah terdapat pengurangan akurasi yang signifikan atau tidak.

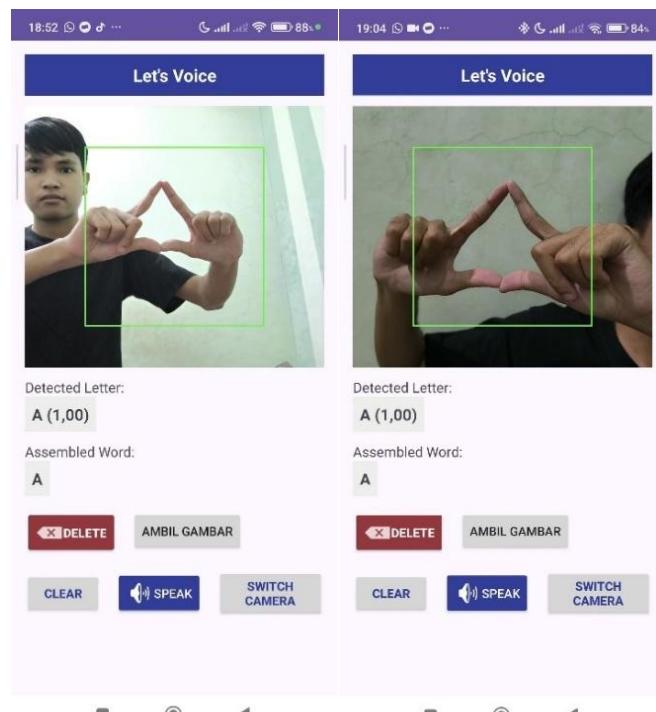
## HASIL DAN PEMBAHASAN

Hasil rancangan dan pengujian menghasilkan dataset augmentasi yang diperoleh dari pengolahan dataset asli dan dataset latar belakang menggunakan teknik seperti rotasi, penambahan noise, dan perubahan kecerahan acak. Dataset tersebut digunakan dalam pelatihan model dengan pendekatan transfer learning menggunakan arsitektur MobileNetV2 melalui dua tahap, yaitu transfer learning dan fine-tuning. Model hasil pelatihan disimpan dalam format .h5 sebagai model utama, kemudian dikonversi ke format TensorFlow Lite (.tflite) agar dapat diimplementasikan pada aplikasi Android untuk mendukung klasifikasi bahasa isyarat secara real-time.



Gambar 1. Hasil Evaluasi Model TFLite

Model MobileNetV2 yang telah dikonversi ke TensorFlow Lite menunjukkan performa klasifikasi yang baik dalam mengenali 26 huruf alfabet BISINDO. Sebagian besar huruf terkласifikasi dengan benar, terlihat dari dominasi nilai diagonal yang tinggi pada setiap kelas, misalnya huruf C (262), F (256), Q (265), dan X (275). Nilai ini menunjukkan konsistensi model dalam mendeteksi huruf dengan tingkat akurasi yang tinggi. Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi (misclassification), terutama pada huruf yang memiliki kemiripan bentuk tangan, seperti I, J, dan Y, yang cenderung terdeteksi silang. Kesalahan tersebut kemungkinan disebabkan oleh variasi posisi tangan, pencahayaan, serta keterbatasan jumlah data pada kelas tertentu. Secara keseluruhan, hasil ini mengindikasikan bahwa model memiliki generalisasi yang baik dengan rata-rata akurasi tinggi, sehingga layak diimplementasikan pada aplikasi Android untuk pengenalan bahasa isyarat secara real-time.



Gambar 2. Hasil Aplikasi Deteksi Bahasa Isyarat

Aplikasi yang dikembangkan berhasil melakukan deteksi huruf bahasa isyarat secara real-time menggunakan kamera perangkat Android. Pada Gambar ditunjukkan bahwa sistem mampu mengenali isyarat tangan dengan tepat, ditandai dengan munculnya huruf hasil klasifikasi pada bagian Detected Letter beserta nilai probabilitas prediksinya. Huruf yang terdeteksi kemudian disusun secara otomatis pada bagian Assembled Word, yang dapat diperbarui sesuai input isyarat berikutnya. Selain itu, aplikasi dilengkapi fitur Delete untuk menghapus huruf, Clear untuk mengosongkan kata, serta Speak yang mengaktifkan modul text-to-speech (TTS) sehingga kata yang terbentuk dapat diucapkan. Hasil pengujian menunjukkan bahwa aplikasi dapat berjalan dengan baik dan responsif, serta mampu menjadi sarana bantu komunikasi antara penyandang tunarungu dengan masyarakat umum secara lebih praktis dan mudah diakses.

## KESIMPULAN

Hasil penelitian menunjukkan bahwa penerapan CNN dengan arsitektur MobileNetV2 efektif untuk pengenalan huruf alfabet Bahasa Isyarat Indonesia secara real-time pada perangkat Android. Model yang dikembangkan mencapai performa akurasi tinggi dengan efisiensi komputasi yang baik setelah dikonversi ke TensorFlow Lite. Integrasi model ke dalam aplikasi Android membuktikan bahwa sistem yang diusulkan berpotensi menjadi solusi praktis dan aplikatif dalam menjembatani komunikasi antara penyandang tunarungu dan masyarakat umum.

## DAFTAR PUSTAKA

- Abdurrahman, F. M. (2019). Pengenalan Bahasa Isyarat Indonesia dengan Algoritma Convolutional Neural Network (CNN) menggunakan Kinect 2.0. <https://repository.its.ac.id/61176/>
- Adam Nurmansyah, Nanda Rizqia Rhamadhani, Sabrina Alfarissy Nur Hakim, Sri Azhari Agustin, & Siti Hamidah. (2023). Permasalahan Komunikasi Yang Kerap Terjadi Pada Penyandang Disabilitas. *Jurnal Pendidikan, Bahasa Dan Budaya*, 2(2), 200–210. <https://doi.org/10.55606/jpbb.v2i2.1515>
- Ahmad Hania, A. (2017). Mengenal Artificial Intelligence, Machine Learning, & Deep Learning. *Jurnal Teknologi Indonesia*, 1(June), 1–6. <https://amt-it.com/mengenal-perbedaan-artificial-intelligence-machine-learning-deep-learning/>
- Althanio, J. G., & Saragih, R. B. R. (2024). Komunikasi Interpersonal Pada Penyandang Tunarungu Dalam Interaksi Sosial di Kota Bengkulu Interpersonal Communication of Deaf People in Social Interactions in Bengkulu City. 8(2).
- Aufar, Y., & Kaloka, T. P. (2022). Robusta coffee leaf diseases detection based on MobileNetV2 model. *International Journal of Electrical and Computer Engineering*, 12(6), 6675–6683. <https://doi.org/10.11591/ijece.v12i6.pp6675-6683>
- Budiman, S. N., Lestanti, S., Yuana, H., & Awwalin, B. N. (2023). SIBI (Sistem Bahasa Isyarat Indonesia) berbasis Machine Learning dan Computer Vision untuk Membantu Komunikasi Tuna Rungu dan Tuna Wicara. *Jurnal Teknologi Dan Manajemen Informatika*, 9(2), 119–128. <https://doi.org/10.26905/jtmi.v9i2.10993>
- Chollet, F. (2017). Deep Learning with Python. Manning Publications.
- Developers, A. (2025). Mengenal Android Studio. <https://developer.android.com/studio/intro?hl=id>
- Developers, G. (2024). CameraX overview. *Android Developers*. <https://developer.android.com/training/camerax>
- Hartati, S. (2021). Kecerdasan Buatan Berbasis Pengetahuan. UGM PRESS.

Huroniyah, F., & Martita, L. (2023). Implementasi Metode Bahasa Isyarat Indonesia (BISINDO) untuk Meningkatkan Kepercayaan Diri Anak Tunarungu di SLBN Tompokersan Lumajang. *Indonesian Journal of Disability* ....  
<https://disabilitas.uinkhas.ac.id/index.php/IJDR/article/view/17%0Ahttps://disabilitas.uinkhas.ac.id/index.php/IJDR/article/download/17/13>

Pressman, R. S. (2015). *Software Engineering: A Practitioner's Approach* (8th ed.). McGraw-Hill Education.

Ridhovan, A., & Suharso, A. (2022). Penerapan Metode Residual Network (Resnet) Dalam Klasifikasi Penyakit Pada Daun Gandum. *JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 7(1), 58–65. <https://doi.org/10.29100/jipi.v7i1.2410%0D>

Sandler, M., et al. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. <https://arxiv.org/abs/1801.04381>

Sholawati, M., Auliasari, K., & Ariwibisono, F. (2022). Pengembangan Aplikasi Pengenalan Bahasa Isyarat Abjad Sibi Menggunakan Metode Convolutional Neural Network (Cnn). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(1), 134–144. <https://doi.org/10.36040/jati.v6i1.4507>

TensorFlow. (2023). Keras Overview. TensorFlow Official Documentation. <https://www.tensorflow.org/guide/keras>

Wulan Anggraini. (2020). DEEP LEARNING UNTUK DETEKSI WAJAH YANG BERHIJAB MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN) DENGAN TENSORFLOW. 2507(February), 1–9.