



PEMBELAJARAN BAHASA ISYARAT BERBASIS CLOUD DAN MACHINE LEARNING DI PUSBISINDO

**Marisa Premitasari¹, Caecilia Sri Wahyuning², Fifi Herni Mustofa³, Jasman
Pardede⁴, Muhammad Azhari⁵, Ardi Fajar Maulana⁶**

^{1,2,3,4,5,6} Institut Teknologi Nasional Bandung, Bandung, Indonesia

Email : marisa@itenas.ac.id

ABSTRAK

Keterbatasan media pembelajaran interaktif untuk Bahasa Isyarat Indonesia (BISINDO) menyulitkan proses belajar mandiri huruf abjad isyarat (A-Z), khususnya bagi pemula yang tidak memiliki akses langsung ke pengajar atau komunitas Tuli. Kegiatan PKM ini menghasilkan aplikasi web pembelajaran BISINDO yang memungkinkan pengguna mempelajari 26 huruf abjad isyarat secara mandiri dan mempraktikkannya langsung di depan kamera, dengan umpan balik otomatis dari model pengenalan gestur tangan. Sistem dibangun dengan arsitektur client-side: model klasifikasi citra MobileNetV3-Large yang di-fine-tune dari bobot ImageNet1K diekspor ke format ONNX dan dijalankan sepenuhnya di peramban (browser) pengguna melalui ONNX Runtime Web (WebAssembly), sehingga tidak memerlukan server inferensi dan menjaga privasi video pengguna karena tidak ada gambar yang dikirim keluar perangkat. Deteksi keberadaan tangan menggunakan MediaPipe HandLandmarker sebagai gerbang (gate) sebelum citra diteruskan ke model klasifikasi, mengurangi prediksi palsu saat tidak ada tangan di depan kamera. Proses pengembangan model melalui beberapa iterasi signifikan: perbaikan skema pembagian data (split) yang semula bocor akibat duplikasi augmentasi menaikkan akurasi validasi dari 27,6% menjadi 84,0%, dan penambahan pemotongan (crop) citra pada area bounding box tangan menaikkannya lebih lanjut hingga 90,4% (validasi) dan 84,7% (uji) untuk arsitektur Large. Aplikasi diuji secara fungsional pada mode latihan interaktif, di mana pengguna diberi umpan balik visual dan audio ketika berhasil memperagakan huruf target. Hasil kegiatan ini menunjukkan bahwa kombinasi model klasifikasi ringan berbasis mobile-oriented CNN dan inferensi on-device dapat menjadi solusi pembelajaran BISINDO yang ringan, dapat diakses melalui perangkat apa pun tanpa instalasi, dan tidak bergantung pada infrastruktur cloud untuk komputasi model.

KEYWORDS

BISINDO, bahasa isyarat, pengenalan gestur tangan, MobileNetV3, on-device inference, pembelajaran berbasis web

ARTICLE HISTORY

Received 19 April 2026

Revised 21 May 2026

Accepted 20 June 2026

CORRESPONDENCE : Marisa Premitasari @ marisa@itenas.ac.id

PENDAHULUAN

Bahasa Isyarat Indonesia (BISINDO) merupakan bahasa isyarat alami yang digunakan komunitas Tuli di Indonesia sebagai media komunikasi utama. Meskipun kesadaran akan pentingnya BISINDO semakin meningkat, ketersediaan media belajar mandiri yang interaktif khususnya yang memberi umpan balik



langsung atas ketepatan gestur yang diperagakan pengguna masih terbatas. Materi belajar yang beredar umumnya berbentuk gambar atau video statis satu arah, tanpa mekanisme untuk memverifikasi apakah pengguna telah memperagakan isyarat dengan benar. Perkembangan model visi komputer berbasis convolutional neural network (CNN) yang efisien, seperti keluarga MobileNet, membuka peluang untuk menjalankan model pengenalan gestur tangan langsung di perangkat pengguna (on-device), tanpa memerlukan server backend yang mahal untuk inferensi. Pendekatan ini relevan untuk konteks pembelajaran daring di Indonesia yang harus mempertimbangkan keterbatasan infrastruktur dan biaya operasional jasa cloud computing berkelanjutan. Pada kegiatan PKM ini, tim mengembangkan aplikasi web pembelajaran abjad BISINDO (26 huruf, A-Z) bernama "Belajar BISINDO" yang terdiri atas dua komponen utama: (1) modul edukasi berisi galeri huruf dan halaman detail per huruf, dan (2) modul latihan interaktif yang memanfaatkan kamera perangkat untuk mengenali gestur tangan pengguna secara real-time dan memberikan umpan balik otomatis. Berbeda dengan pendekatan berbasis cloud speech-to-text yang bergantung pada layanan vendor eksternal (mis. AWS, GCP) untuk memproses data pada server jarak jauh, sistem pengenalan gestur pada kegiatan ini dirancang untuk berjalan sepenuhnya di sisi klien (client-side inference), sehingga aplikasi tetap ringan diakses dan tidak menimbulkan biaya langganan komputasi model per penggunaan.

Secara umum, materi BISINDO yang diajarkan di PUSBISINDO terbagi menjadi dua kategori utama: isyarat abjad (huruf A–Z, digunakan terutama untuk mengeja nama atau istilah yang belum memiliki isyarat baku) dan isyarat kosakata/kalimat (kata dan frasa sehari-hari yang memiliki bentuk isyarat tersendiri, tidak dieja huruf per huruf). Kegiatan PKM ini berfokus pada kategori pertama, yaitu abjad BISINDO (A–Z), karena isyarat abjad bersifat statis (satu gestur per huruf, tanpa gerakan berkelanjutan) sehingga lebih sesuai untuk didekati dengan model klasifikasi citra tunggal seperti MobileNetV3. Isyarat kosakata umumnya bersifat dinamis (rangkaian gerakan dari waktu ke waktu) dan

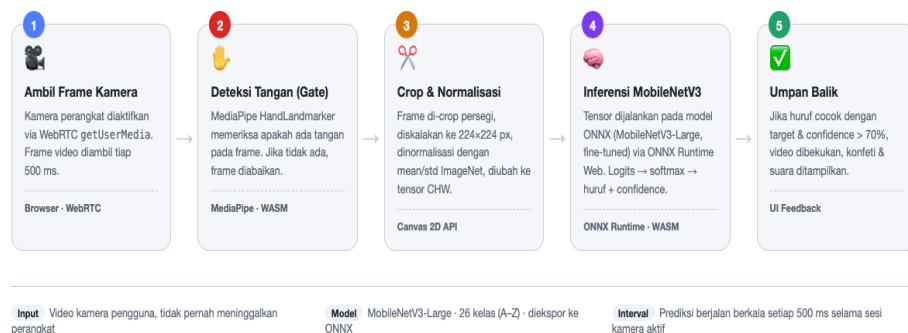
memerlukan pendekatan pengenalan gestur berbasis video/temporal (mis. arsitektur CNN-LSTM), sehingga berada di luar cakupan (scope) kegiatan PKM ini dan diusulkan sebagai pengembangan lanjutan.

METODE

Aplikasi dibangun menggunakan Next.js (React) sebagai framework front-end. Seluruh proses pengenalan gestur BISINDO, mulai dari deteksi tangan, praproses citra, hingga inferensi model, dilakukan langsung di browser pengguna (client-side inference). Gambar 1 menunjukkan alur sistem pengenalan gestur BISINDO yang terdiri atas lima tahap. Karena seluruh proses berjalan di sisi klien, video dari kamera tidak perlu dikirim ke server. Dengan demikian, aplikasi tidak membutuhkan server khusus untuk inferensi maupun biaya komputasi cloud secara berkelanjutan, seperti pada pendekatan yang menggunakan layanan cloud AWS atau GCP.

Gambar. Alur Sistem Pengenalan Gestur BISINDO (Client-Side Inference)

Seluruh proses berjalan di peramban pengguna – tidak ada frame video yang dikirim ke server.



Gambar 1. Alur Sistem Pengenalan Gestur BISINDO (Client-Side Inference)

HASIL dan PEMBAHASAN

Kegiatan PKM ini dilaksanakan melalui beberapa tahapan: pengumpulan dan penyiapan data, pengembangan dan evaluasi model, pengembangan aplikasi web, serta pengujian fungsional.

1. Pengumpulan dan Penyiapan Dataset

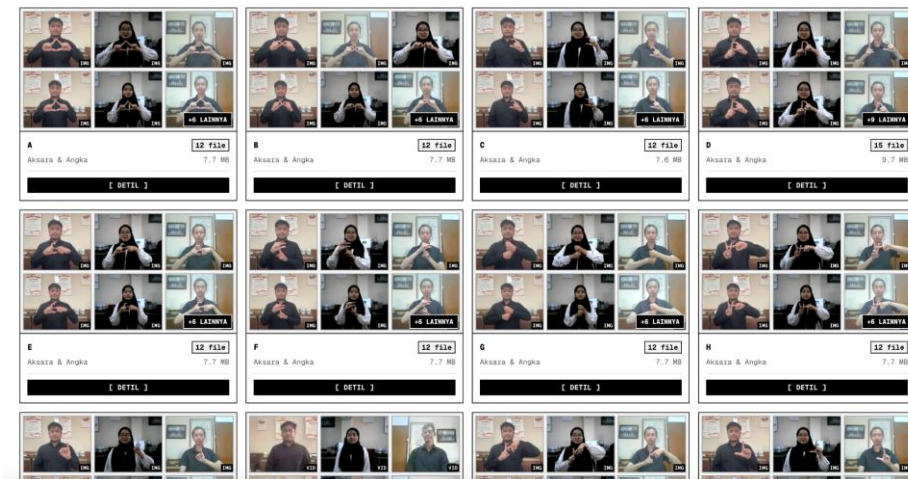
Pengumpulan data gestur tangan dilakukan secara langsung untuk

mendapatkan dokumentasi isyarat huruf abjad BISINDO (A–Z) yang variatif. Proses ini melibatkan partisipasi relawan dengan mengambil sampel peragaan tangan dari berbagai sudut dan kondisi pencahayaan ruangan agar data yang diperoleh menyerupai kondisi penggunaan nyata.



Gambar 2. Dokumentasi Pengambilan Dataset Gestur Tangan BISINDO

Proses pengambilan data seperti pada Gambar 2 menjadi acuan utama dalam membangun bahan latih sistem pengenalan isyarat. Seluruh foto yang terkumpul kemudian dirapikan dan dikelompokkan berdasarkan kelas huruf masing-masing sebelum digunakan dalam proses pelatihan sistem.



Gambar 3. Contoh Isyarat BISINDO

Gambar 3 memperlihatkan struktur penyimpanan data mentah hasil dokumentasi, dikelompokkan per folder untuk masing-masing huruf abjad (A–Z). Setiap folder kelas berisi 12–15 berkas foto dan video pendek dari beberapa

relawan peraga yang berbeda, diambil pada sudut pengambilan dan kondisi pencahayaan yang bervariasi. Pengelompokan pada tahap ini menjadi dasar sebelum data diproses lebih lanjut (pelabelan ulang, cropping area tangan, dan pembagian ke subset latih/validasi/uji) sebagaimana dijelaskan pada bagian Metodologi.

Tabel 1. Tabel Isyarat BISINDO

Subset	Jumlah Citra	Rentang Citra per Kelas
Latih (train)	1.643	54 – 74
Validasi (Valid)	332	9 – 18
Uji (test)	346	9 – 17
Total	2.321	-

2. Pengembangan Aplikasi

Aplikasi web "Belajar BISINDO" dirancang agar sederhana dan ramah pengguna, khususnya bagi pemula yang ingin belajar mandiri. Aplikasi ini terbagi menjadi dua bagian utama:

- Modul Belajar: Menampilkan galeri kartu huruf A sampai Z. Saat salah satu huruf dipilih, pengguna dapat melihat bentuk isyarat beserta petunjuk cara memperagakannya secara tepat.
- Modul Latihan (Interaktif): Mengaktifkan kamera perangkat pengguna untuk mempraktikkan isyarat secara langsung. Sistem akan membaca gerakan tangan dan memberikan tanggapan berupa animasi atau suara saat peragaan isyarat sudah benar.



Gambar 4. Tampilan Antarmuka Aplikasi Web Belajar BISINDO

Tampilan antarmuka aplikasi pada Gambar 4 dibuat bersih dan intuitif agar pengguna dapat dengan mudah berpindah antara materi pembelajaran dan menu latihan langsung di depan kamera.

3. Pengujian Fungsional

Tahap uji coba dilakukan untuk melihat sejauh mana aplikasi dapat digunakan dengan lancar oleh masyarakat umum. Pengguna diminta mencoba modul latihan, memperagakan beberapa huruf di depan kamera laptop maupun ponsel, dan merasakan respons dari sistem.



Gambar 5. Kegiatan Uji Coba Penggunaan Aplikasi Bersama Pengguna

Berdasarkan kegiatan uji coba yang mendokumentasikan antusiasme pengguna pada Gambar 5, sistem mampu merespons gestur tangan secara cepat tanpa rasa jeda (*lag*) yang mengganggu. Mekanisme umpan balik berupa suara dan animasi konfeti saat latihan berhasil terbukti menambah motivasi pengguna dalam belajar.

4. Kendala dan Temuan

Beberapa pasangan huruf ditemukan rawan tertukar akibat kemiripan bentuk tangan pada resolusi sumber citra yang rendah (96x96 piksel sebelum diskalakan ke 224x224), yaitu D/A, M/N, U/R, P/S, dan Q/S. Sebagian huruf (mis. Q, M) juga bergantung pada konteks kedekatan tangan dengan wajah, yang dapat berkurang ketepatannya bila crop tangan terlalu ketat. Ukuran dataset uji yang



kecil (sekitar 10-17 citra per kelas) turut menyebabkan fluktuasi akurasi per kelas yang cukup besar antar-eksperimen, sehingga penyesuaian parameter padding crop lebih lanjut cenderung mengoptimalkan noise, bukan sinyal yang substantif.

SIMPULAN

Kegiatan PKM ini menghasilkan aplikasi web pembelajaran BISINDO yang memungkinkan pengguna belajar 26 huruf abjad isyarat secara mandiri dan mempraktikkannya langsung di depan kamera dengan umpan balik otomatis. Berbeda dari pendekatan pembelajaran berbasis cloud yang bergantung pada layanan pihak ketiga untuk pemrosesan di server, sistem pengenalan gestur pada aplikasi ini menjalankan model MobileNetV3 sepenuhnya di sisi klien melalui ONNX Runtime Web, sehingga ringan diakses melalui berbagai perangkat tanpa instalasi maupun biaya komputasi server berkelanjutan, sekaligus menjaga privasi video pengguna.

Proses pengembangan model menunjukkan bahwa perbaikan integritas pembagian data (mengatasi kebocoran split akibat duplikasi augmentasi) memberikan lonjakan akurasi terbesar (dari 27,6% menjadi 84,0% validasi), lebih signifikan dibandingkan penggantian arsitektur atau penyesuaian hyperparameter semata. Model akhir (MobileNetV3-Large dengan crop bounding box tangan) mencapai akurasi validasi 90,4% dan akurasi uji 84,7%, dengan kesalahan tersisa terkonsentrasi pada pasangan huruf yang secara visual memang mirip. Pengembangan lanjutan yang dapat dilakukan meliputi penambahan data foto beresolusi lebih tinggi khususnya untuk pasangan huruf yang rawan tertukar, validasi silang (k-fold) untuk estimasi akurasi yang lebih stabil, serta perluasan cakupan dari huruf abjad ke kosakata isyarat BISINDO yang lebih luas.

DAFTAR PUSTAKA

Dwijayanti, S., Hikmarika, H., et al. (2021). Indonesia Sign Language Recognition using Convolutional Neural Network. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 12(10).

Howard, A., et al. (2019). Searching for MobileNetV3. *Proceedings of the*



IEEE/CVF International Conference on Computer Vision (ICCV).

Indonesian Sign Language (BISINDO) Classification Using Xception Transfer Learning Architecture. International Journal of Advances in Data and Information Systems (IJADIS).

Implementation of Hand Gesture Recognition for Indonesian Sign Language (SIBI) Based on MobileNetV2. JKTJ Jurnal Keilmuan Teknologi Informasi (2026).

L2IC and MobileViT-XXS for BISINDO Alphabet Recognition. Journal of Applied Informatics and Computing (JAIC).

Mahesa, T. R., et al. (2020). BISINDO (Bahasa Isyarat Indonesia) Sign Language Recognition Using CNN and LSTM. Advances in Science, Technology and Engineering Systems Journal, 5(5), 35.

MediaPipe Tasks: Hand Landmarker.
https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker
(diakses 2026)

ONNX Runtime Web Documentation. <https://onnxruntime.ai/docs/tutorials/web/>
(diakses 2026)

Prasetyo, D., et al. Integrating the CNN Model with the Web for Indonesian Sign Language (BISINDO) Recognition. Journal of Applied Informatics and Computing (JAIC).

Word Recognition and Automated Epenthesis Removal for Indonesian Sign System Sentence Gestures (MobileNetV2 + LSTM). Indonesian Journal of Electrical Engineering and Computer Science.