

## ANALISIS PERFORMA CNN SEQUENTIAL DAN MOBILENETV2 UNTUK KLASIFIKASI TULISAN TANGAN BERBASIS HAND TRACKING MEDIAPIPE

Salsabil Rinaldi, Dhiyaurrahman Hamizan Putra Haikal,  
Akhamad Zufar Rafif, Dodik Ariyanto, Faldiena Marcelita

Teknologi Rekayasa Perangkat Lunak, IPB University  
Jalan Kumbang No. 14, Babakan, Bogor, Indonesia  
dodikariyanto@apps.ipb.ac.id

### ABSTRAK

Kemampuan menulis merupakan keterampilan dasar yang penting dalam proses pembelajaran, khususnya dalam pengenalan huruf dan angka. Namun, anak dengan disleksia sering mengalami kesulitan dalam mengenali, membedakan, dan menuliskan karakter yang memiliki bentuk visual serupa, sehingga diperlukan media pembelajaran yang lebih interaktif dan adaptif. Penelitian ini bertujuan untuk menganalisis performa model *Convolutional Neural Network* (CNN) Sequential dan MobileNetV2 dalam klasifikasi tulisan tangan berbasis hand tracking menggunakan MediaPipe Hands. Metode yang digunakan meliputi pengumpulan dataset tulisan tangan huruf kapital, huruf kecil, dan angka, pra-pemrosesan data berupa resize, grayscale, normalisasi, dan augmentasi, serta pengujian model pada setiap kategori karakter. Evaluasi dilakukan menggunakan metrik accuracy, precision, recall, F1-score, grafik akurasi dan loss, serta *confusion matrix*. Hasil pengujian menunjukkan bahwa Sequential CNN lebih optimal pada kategori angka dengan F1-score 97,28%, sedangkan MobileNetV2 lebih unggul pada huruf kapital dengan F1-score 96,83% dan huruf kecil dengan F1-score 97,77%. Hasil ini menunjukkan bahwa pemilihan model berdasarkan kategori karakter dapat meningkatkan performa klasifikasi tulisan tangan.

**Kata kunci :** Klasifikasi Tulisan Tangan, CNN, Sequential, MobileNETV2, MediaPipe Hands

### 1. PENDAHULUAN

Pendidikan inklusif bagi anak dengan kebutuhan khusus, khususnya disleksia, merupakan salah satu fokus penting dalam sistem pendidikan modern. Literasi yang mencakup kemampuan membaca dan menulis menjadi pondasi dasar pendidikan karena sangat dibutuhkan dalam pengembangan akademik seseorang [1].

Namun, anak dengan disleksia seringkali menghadapi kesulitan mengeja, mengenali bentuk huruf, serta menulis tangan [2].

Gangguan ini sering kali mencakup hambatan motorik halus, di mana tulisan tangan menjadi tidak jelas, tidak konsisten, dan memiliki kecepatan penulisan yang sangat lambat [3].

Sementara itu, tidak ada perbedaan yang signifikan antara pembelajaran menulis di sekolah umum bagi siswa penderita hambatan tersebut [3].

Kondisi ini menunjukkan bahwa anak dengan disleksia membutuhkan pendekatan pembelajaran yang lebih sesuai dengan karakteristik dan kebutuhannya. Kesulitan ini tidak hanya memengaruhi performa anak di sekolah, tetapi juga memengaruhi rasa percaya diri dan kehidupan sehari-hari anak di masa depan [4].

Perkembangan *computer vision*, *machine learning*, dan teknologi *hand tracking* membuka peluang penerapan *touchless interaction* sebagai media interaksi yang lebih natural dan intuitif, khususnya pada sistem pembelajaran interaktif berbasis gestur tangan [4].

Deteksi tangan dilakukan dengan menggunakan *hand landmark detection* merupakan model yang

dirancang secara akurat dalam menentukan posisi dari 21 titik kunci pada tangan dalam bentuk koordinat 2.5D [5].

Salah satu framework yang potensial adalah MediaPipe Hands dari Google, yang mampu melakukan pelacakan 21 *landmark* tangan secara *real-time* dengan presisi spasial tinggi [6]. Teknologi ini dinilai efisien karena mampu berjalan dengan baik pada perangkat berspesifikasi CPU standar atau perangkat *edge*, tanpa membutuhkan perangkat keras tambahan berbiaya tinggi.

Tantangan utama dalam sistem ini adalah memilih arsitektur model klasifikasi yang tepat untuk mengenali pola tulisan tangan anak disleksia yang sering kali tidak beraturan dan unik. Permasalahan variasi bentuk karakter yang tinggi ini menuntut adanya model klasifikasi yang optimal. Sebagai rencana penyelesaian masalah, penelitian ini menerapkan dan membandingkan dua pendekatan arsitektur Deep Learning yang berbeda. Model Convolutional Neural Network (CNN) telah terbukti sebagai metode paling efektif dalam mengekstraksi fitur spasial secara otomatis dari citra tulisan tangan [7].

Convolutional Neural Network (CNN) merupakan salah satu algoritma Deep Learning yang mengambil masukan dalam bentuk gambar dan menentukan aspek atau objek mana dalam gambar yang dapat digunakan oleh mesin untuk "belajar" mengenali gambar yang digunakan untuk membedakan antara gambar yang satu dengan gambar yang lain [8].

Arsitektur CNN Sequential menawarkan kesederhanaan dalam pengembangan dan efisiensi dalam menangani variasi gaya tulisan. Pada sistem ini, arsitektur tersebut digunakan sebagai model dasar karena memiliki struktur yang sederhana dan efisien dalam menangani variasi gaya tulisan. Di sisi lain, MobileNetV2 digunakan sebagai model pembanding karena memiliki arsitektur *inverted residual* dan *depthwise separable convolutions* yang dirancang khusus untuk efisiensi komputasi maksimal dan latensi rendah [9].

Hingga saat ini, meskipun CNN Sequential dan MobileNetV2 telah banyak diuji untuk berbagai tugas klasifikasi objek, masih sedikit penelitian yang secara spesifik membandingkan performa kedua model ini dalam konteks input real-time dari MediaPipe untuk intervensi disleksia. Oleh karena itu, penelitian ini bertujuan untuk menganalisis dan membandingkan performa metrik klasifikasi dari model CNN Sequential dan MobileNetV2 dalam mengklasifikasikan tulisan tangan anak disleksia. Melalui analisis perbandingan ini, diharapkan dapat ditentukan arsitektur model yang paling optimal berdasarkan setiap kategori karakter untuk mendukung proses pembelajaran mereka secara mandiri.

## 2. TINJAUAN PUSTAKA

### 2.1. Penelitian Terdahulu

Beberapa penelitian terdahulu telah mengkaji penerapan teknologi *Convolutional Neural Network* (CNN) dalam klasifikasi tulisan tangan serta penggunaan *hand tracking* berbasis MediaPipe dalam sistem *air writing*. CNN banyak digunakan karena mampu menghasilkan tingkat akurasi yang tinggi dalam pengenalan huruf dan angka tulisan tangan, baik pada dataset statis maupun *real-time*.

Penelitian yang dilakukan oleh Ley dkk. menggunakan dataset EMNIST yang terdiri dari gabungan huruf kapital, huruf kecil, dan angka dengan distribusi data yang lebih seimbang. Penelitian tersebut juga melakukan eksperimen penambahan lapisan pada arsitektur CNN untuk meningkatkan toleransi terhadap *noise*. Hasil penelitian menunjukkan bahwa model mampu mencapai akurasi sebesar 86,17% pada dataset EMNIST *balanced* dan 97,46% pada dataset MNIST [10].

Penelitian lain oleh Abdiansah dkk. menerapkan CNN untuk pengenalan aksara Jawa yang terdiri dari 20 karakter dasar. Dataset yang digunakan berjumlah 1.000 citra tulisan tangan yang diperoleh dari responden terampil. Dengan menggunakan optimizer Adam dan proses pelatihan sebanyak 150 *epoch*, sistem berhasil memperoleh tingkat akurasi sebesar 99,83%. Hasil tersebut menunjukkan bahwa CNN mampu mengenali pola dan lekukan karakter tradisional yang kompleks dengan tingkat presisi yang tinggi [11].

Selanjutnya, Espinoza-Vicuña mengembangkan aplikasi seluler untuk mengenali pola tulisan tangan penderita disleksia menggunakan kombinasi model

CNN dan U-Net. Penelitian tersebut menemukan bahwa tulisan penderita disleksia memiliki karakteristik tertentu, seperti kecepatan menulis yang rendah dan adanya titik henti pada trajektori huruf. Model yang dikembangkan berhasil mencapai tingkat akurasi di atas 95%, sehingga menunjukkan potensi penggunaan CNN dalam analisis pola tulisan tangan anak dengan disleksia [12].

Selain penelitian terkait klasifikasi tulisan tangan, beberapa penelitian juga telah mengembangkan sistem *air writing* dan *hand tracking* menggunakan MediaPipe. Nagaraju dkk. merancang sistem penulisan digital berbasis pelacakan tangan menggunakan pustaka MediaPipe dan OpenCV. Sistem tersebut memungkinkan pengguna menulis di layar hanya menggunakan gerakan jari, dimana satu jari digunakan untuk menulis dan dua jari digunakan sebagai perintah untuk menghapus tulisan pada layar [13].

Penelitian oleh Steven Suki dkk. mengintegrasikan MediaPipe dengan model CNN kustom untuk membangun sistem *virtual mouse*. Dalam penelitian tersebut, MediaPipe digunakan untuk mendeteksi 21 *landmark* tangan dalam koordinat 2.5D sebagai proses ekstraksi fitur. Hasil pengujian menunjukkan tingkat akurasi validasi sebesar 99,9% pada kondisi latar belakang yang terkendali, namun performa menurun menjadi 66% pada latar belakang yang kompleks. Hal ini menunjukkan bahwa kualitas lingkungan sangat memengaruhi performa sistem berbasis *hand tracking* [5].

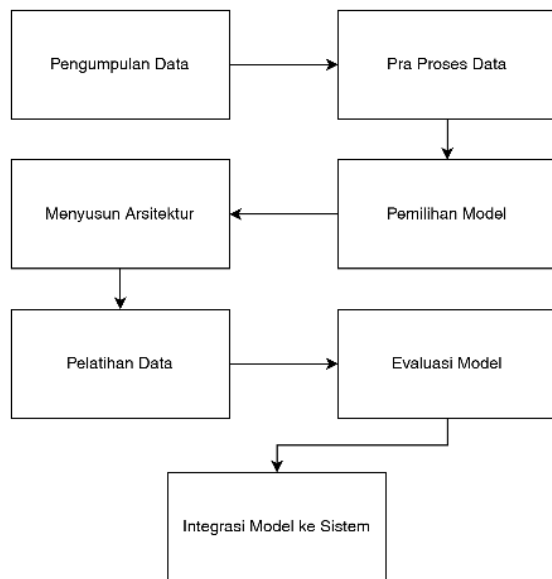
Sementara itu, Putri dkk. meneliti ketahanan sistem MediaPipe terhadap faktor lingkungan seperti pencahayaan dan jarak kamera. Sistem yang dikembangkan menganalisis perpindahan *landmark* tangan untuk menentukan arah gerakan pengguna. Hasil penelitian menunjukkan bahwa MediaPipe mampu mencapai akurasi 100% pada jarak 200 cm dengan pencahayaan terang. Namun, performanya menurun secara signifikan hingga di bawah 30% pada kondisi pencahayaan redup dan jarak 450 cm. Penelitian tersebut menunjukkan bahwa sistem berbasis MediaPipe memiliki sensitivitas tinggi terhadap kondisi lingkungan [6].

Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa kombinasi CNN dan MediaPipe memiliki potensi yang besar dalam pengembangan sistem pengenalan tulisan tangan berbasis *air writing*. Sebagian besar penelitian sebelumnya lebih berfokus pada pengembangan sistem interaksi atau pengenalan karakter secara umum, tanpa melakukan analisis mendalam terhadap kestabilan dan konsistensi performa model klasifikasi. Oleh karena itu, penelitian ini dilakukan untuk menganalisis dan membandingkan performa arsitektur CNN, yaitu Sequential CNN dan MobileNetV2, dalam proses klasifikasi tulisan tangan berbasis *hand tracking*.

### 3. METODE PENELITIAN

Proses klasifikasi data menggunakan Algoritma CNN dimulai dengan mengumpulkan data di lapangan sebagai dataset. Data ini digunakan untuk *training* dan *testing*. Setelah itu, dilakukan *preprocessing* untuk merancang model Algoritma CNN, kemudian diuji untuk mendapatkan akurasi. Tahap terakhir melibatkan pengujian Algoritma CNN menggunakan *Confusion Matrix* [14].

Tahapan penelitian dirancang untuk menganalisis performa model *Convolutional Neural Network* (CNN) dalam proses klasifikasi tulisan tangan berbasis *hand tracking*.



Gambar 1. Diagram tahapan penelitian

#### 3.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle berjudul *Handwritten English Characters and Digits Dataset* oleh Sujay Mann. Dataset terdiri atas data tulisan tangan huruf kapital (A–Z), huruf kecil (a–z), dan angka (0–9) dengan total 62 kelas karakter. Dataset tersebut sebelumnya telah dipisahkan menjadi data *training* dan data *testing*. Pada masing-masing kelas terdapat 44 citra data *training* dan 11 citra data *testing*, sehingga total data yang digunakan dalam penelitian ini berjumlah 3.410 citra, yang diperoleh dari perhitungan  $(44 \times 62) + (11 \times 62)$ .

Selain menggunakan dataset publik, penelitian ini juga memanfaatkan dataset tambahan hasil penyesuaian sendiri untuk mendukung proses pengujian. Dataset tambahan tersebut digunakan untuk menyesuaikan variasi bentuk tulisan tangan pengguna pada kondisi penggunaan secara *real-time* agar model mampu mengenali pola tulisan yang lebih beragam.

#### 3.2. Pra Proses Data

Proses pra-pemrosesan data yang dilakukan meliputi *resize* ukuran data, *grayscale*, normalisasi nilai piksel, augmentasi data, serta pengolahan hasil

gerakan tulisan tangan yang ditangkap melalui *hand tracking* menggunakan MediaPipe Hands.

Citra asli dari dataset yang berukuran  $1200 \times 900$  piksel diubah menjadi  $64 \times 64$  piksel. *Resize* dilakukan agar seluruh citra memiliki ukuran yang seragam.

Setelah itu, dilakukan proses *grayscale*, yaitu mengubah citra dari 3 *channel* warna RGB menjadi 1 *channel grayscale*. Teknik ini digunakan untuk mereduksi dimensi data, mempercepat proses komputasi, serta menghilangkan informasi warna yang tidak terlalu relevan dalam proses pengenalan bentuk huruf maupun angka.

Normalisasi nilai piksel dilakukan dengan mengubah rentang nilai piksel dari 0–255 menjadi 0–1, yaitu dengan membagi setiap nilai piksel dengan 255. Proses ini bertujuan untuk membuat skala data lebih kecil dan stabil sehingga model lebih mudah mempelajari pola citra selama proses pelatihan.

Selanjutnya, augmentasi data diterapkan pada data *training* untuk menambah variasi bentuk tulisan tangan. Pada penelitian ini, augmentasi dilakukan dengan transformasi berupa rotasi dan translasi. Rotasi digunakan untuk mensimulasikan variasi kemiringan tulisan, sedangkan translasi digunakan untuk mensimulasikan perubahan posisi karakter pada citra.

Kemudian dataset dibagi kembali menggunakan metode *randomized split* menjadi tiga subset data, yaitu 80% data *training*, 10% data *testing*, dan 10% data *validation*. Berdasarkan total data pengujian yang digunakan pada sistem, diperoleh 176 data untuk *training*, 22 data untuk *testing*, dan 22 data untuk *validation*.

#### 3.3. Pemilihan Model

Tahap pemilihan model dilakukan untuk menentukan arsitektur *Convolutional Neural Network* (CNN) yang paling optimal dalam proses klasifikasi tulisan tangan pada penelitian ini. Pada tahap ini digunakan dua model, yaitu Sequential CNN sebagai model dasar (*baseline model*) dan MobileNetV2 sebagai model pembanding yang lebih ringan dan efisien. Kedua model diuji untuk mengetahui kemampuan masing-masing dalam mengenali pola tulisan tangan berdasarkan kategori karakter yang digunakan, yaitu angka, huruf kapital, dan huruf kecil.

#### 3.4. Menyusun Arsitektur

Pada tahap ini, akan dilakukan penentuan arsitektur dari CNN, yang berkaitan dengan jenis model, jumlah lapisan konvolusi, lapisan subsampling, dan jumlah lapisannya [15].

Penelitian ini menggunakan dua arsitektur model, yaitu CNN Sequential dan MobileNetV2.

Tabel 1. Arsitektur CNN Model Sequential

| Layer                   | Ukuran             | Dimensi Keluaran   |
|-------------------------|--------------------|--------------------|
| Input                   | (64, 64, 1)        | -                  |
| Convolutional 2D + ReLu | 32 (3 x 3) filters | (None, 64, 64, 32) |

| Layer                   | Ukuran          | Dimensi Keluaran   |
|-------------------------|-----------------|--------------------|
| Max Pooling 2D          | (2 x 2) filters | (None, 32, 32, 32) |
| Convolutional 2D + ReLu | (3 x 3) filters | (None, 32, 32, 64) |
| Max Pooling 2D          | (2 x 2) filters | (None, 8, 8, 128)  |
| Flatten                 | -               | (None, 8192)       |
| Dense + ReLu            | 256 perceptrons | (None, 256)        |
| Dropout (rate=0.5)      | -               | (None, 256)        |
| Dense + Softmax         | 10 perceptrons  | (None, 10)         |

Tabel 2. Arsitektur CNN Model MobileNetV2

| Layer                     | Ukuran                 | Dimensi Keluaran   |
|---------------------------|------------------------|--------------------|
| Input                     | (64, 64, 3)            | -                  |
| Data Augmentation         | Rotation & Translation | (None, 64, 64, 32) |
| Preprocessing             | MobileNetV2 Preprocess | (None, 64, 64, 3)  |
| MobileNetV2 (Base)        | Pre-trained ImageNet   | (None, 2, 2, 1280) |
| Global Average Pooling 2D | -                      | (None, 1280)       |
| Dense + ReLu              | 256 perceptrons        | (None, 256)        |
| Dropout (rate=0.5)        | -                      | (None, 256)        |
| Dense + Softmax           | 26 perceptrons         | (None, 26)         |

Dengan penyusunan dua arsitektur tersebut, penelitian ini dapat membandingkan kemampuan model Sequential CNN dan MobileNetV2 dalam mengenali pola tulisan tangan. Sequential CNN memiliki struktur yang lebih sederhana dan digunakan sebagai pembanding dasar, sedangkan MobileNetV2 memiliki kemampuan ekstraksi fitur yang lebih kompleks dan efisien. Hasil dari proses ini kemudian digunakan untuk menentukan model terbaik pada

#### 4.1. Evaluasi Metrik Klasifikasi

Tabel 3. Perbandingan Metrik Klasifikasi Model CNN Sequential dan MobileNetV2

| Model       | Kelas               | Akurasi | Precision | Recall | F-1 Score |
|-------------|---------------------|---------|-----------|--------|-----------|
| Sequential  | Angka (0 - 9)       | 97,27%  | 97,39%    | 97,27% | 97,28%    |
|             | Huruf Besar (A - Z) | 95,10%  | 95,43%    | 95,10% | 95,06%    |
|             | Huruf Kecil (a - z) | 95,28%  | 95,77%    | 95,28% | 95,27%    |
| MobileNetV2 | Angka (0 - 9)       | 89,09%  | 90,86%    | 89,09% | 89,11%    |
|             | Huruf Besar (A - Z) | 96,85%  | 97,26%    | 96,85% | 96,83%    |
|             | Huruf Kecil (a - z) | 97,73%  | 98,07%    | 97,73% | 97,77%    |

Berdasarkan hasil pelatihan dan pengujian yang dilakukan, diperoleh bahwa model Sequential CNN memberikan performa yang lebih baik pada proses klasifikasi karakter angka (0–9). Oleh karena itu, hasil klasifikasi untuk kategori nomor pada sistem menggunakan model Sequential CNN.

masing-masing kategori karakter, yaitu angka, huruf kapital, dan huruf kecil.

#### 3.5. Pelatihan Data

Tahap pelatihan dilakukan menggunakan dataset yang telah diproses sebelumnya. Proses pelatihan model dilakukan menggunakan TensorFlow dan Keras dengan parameter seperti *epoch*, *batch size*, dan *optimizer* Adam untuk menghasilkan model klasifikasi yang optimal.

#### 3.6. Evaluasi Model

Tahap evaluasi dilakukan untuk mengukur performa model dalam mengenali tulisan tangan. Performa model dievaluasi dengan menggunakan dataset pengujian, menghitung metrik akurasi, presisi, recall, dan F1-score untuk mengukur kinerja klasifikasi [14].

Selain itu, dilakukan analisis perbandingan performa antara Sequential CNN dan MobileNetV2 untuk mengetahui model dengan performa terbaik.

#### 3.7. Integrasi Model ke Sistem

Dilakukan proses validasi dan pengujian terhadap input gambar yang dihasilkan dari metode *air writing* pada sistem. Hasil deteksi gerakan tangan dari MediaPipe Hands dihubungkan dengan model CNN sehingga input tulisan yang dibuat pengguna melalui gerakan jari dapat diproses secara otomatis dan *real-time*. Setelah proses klasifikasi selesai, hasil prediksi akan ditampilkan pada sistem sebagai output pengenalan tulisan tangan pengguna.

### 4. HASIL DAN PEMBAHASAN

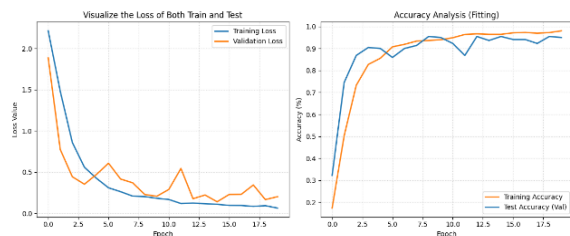
Penelitian ini menggunakan dua model CNN, yaitu Sequential CNN sebagai model dasar (*baseline model*) dan MobileNetV2 sebagai model pembanding yang lebih ringan dan efisien. Kedua model diuji untuk mengetahui kemampuan masing-masing dalam mengenali pola tulisan tangan berdasarkan kategori karakter yang digunakan, yaitu angka, huruf kapital, dan huruf kecil.

Sementara itu, pada proses klasifikasi huruf kapital (A–Z) dan huruf kecil (a–z), model MobileNetV2 menunjukkan performa yang lebih optimal dibandingkan Sequential CNN.

#### 4.2. Grafik Pelatihan Model

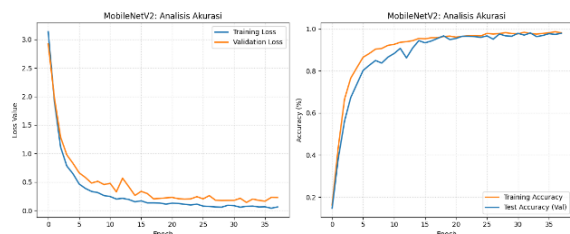
Grafik ini memperlihatkan perbandingan antara nilai *training accuracy* dan *validation accuracy*, serta nilai *training loss* dan *validation loss* pada setiap epoch untuk tiap kategori dengan model pilihan.

Pada kategori nomor, model yang digunakan adalah Sequential CNN. Nilai *training loss* mengalami penurunan yang cukup tajam dari awal *epoch* hingga mendekati nilai yang rendah pada *epoch* akhir. Nilai *validation loss* juga menunjukkan pola penurunan yang relatif serupa, meskipun terdapat sedikit fluktuasi pada beberapa *epoch*. Hal ini menunjukkan bahwa model mampu mempelajari pola angka dengan baik. Pada grafik akurasi, *training accuracy* meningkat secara bertahap hingga mendekati nilai sekitar 98%, sedangkan *validation accuracy* juga berada pada nilai yang tinggi dan relatif stabil di atas 90% setelah beberapa *epoch* awal.



Gambar 2. Grafik *accuracy* dan *loss* pada kategori data nomor dengan model Sequential

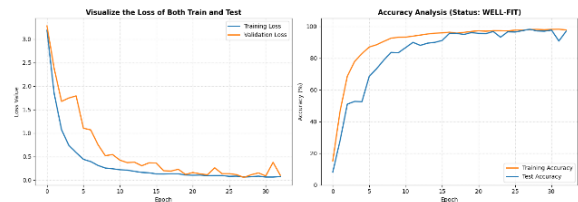
Pada kategori huruf besar, model yang digunakan adalah MobileNetV2. Nilai *training loss* dan *validation loss* mengalami penurunan yang konsisten dari awal hingga akhir proses pelatihan. Pada awal *epoch*, *loss* masih berada pada nilai yang cukup tinggi, namun menurun secara signifikan setelah beberapa *epoch* dan mulai stabil pada *epoch* berikutnya. Grafik akurasi juga menunjukkan peningkatan yang stabil, di mana *training accuracy* dan *validation accuracy* naik secara bertahap hingga mencapai nilai mendekati 98%. Selisih antara akurasi *training* dan validasi yang relatif kecil menunjukkan bahwa model tidak mengalami *overfitting* secara signifikan.



Gambar 3. Grafik *accuracy* dan *loss* pada kategori data huruf besar dengan model MobileNetV2

Sedangkan pada kategori huruf kecil, *training loss* mengalami penurunan tajam sejak awal pelatihan dan terus menurun hingga berada pada nilai yang rendah. Nilai *validation loss* juga menunjukkan penurunan yang cukup baik, meskipun terdapat beberapa fluktuasi kecil pada *epoch* tertentu. *Training accuracy*

meningkat dengan cepat dan mencapai nilai tinggi setelah beberapa *epoch*, sementara *validation accuracy* mengikuti pola peningkatan yang serupa hingga berada pada kisaran 95–98%.

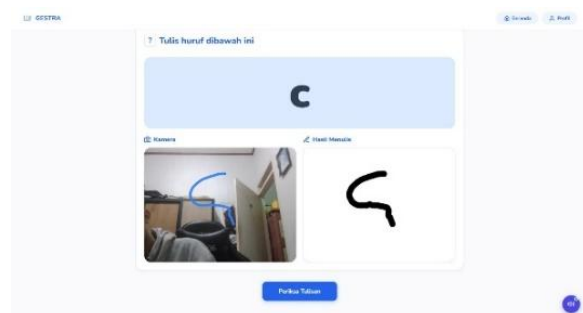


Gambar 4. Grafik *accuracy* dan *loss* pada kategori data huruf kecil dengan model MobileNetV2

Hasil pengujian menunjukkan bahwa kombinasi penggunaan kedua model memberikan performa klasifikasi yang lebih baik dibandingkan hanya menggunakan satu model secara keseluruhan. Sequential CNN digunakan khusus untuk klasifikasi angka karena lebih stabil pada karakter sederhana, sedangkan MobileNetV2 digunakan untuk klasifikasi huruf besar dan huruf kecil karena memiliki kemampuan ekstraksi fitur yang lebih baik pada karakter dengan tingkat kompleksitas lebih tinggi. Dengan demikian, pemilihan model dilakukan berdasarkan hasil performa terbaik dari masing-masing kategori karakter agar sistem mampu menghasilkan klasifikasi tulisan tangan yang lebih akurat dan efisien.

#### 4.3. Hasil Uji Coba Sistem

Model CNN yang telah dilatih kemudian diterapkan ke dalam sistem yang bekerja dengan menangkap gerakan tangan pengguna melalui kamera, kemudian titik ujung jari yang terdeteksi digunakan untuk membentuk lintasan tulisan pada area canvas digital. Hasil tulisan tersebut selanjutnya diproses oleh sistem yang mengirimkan input tersebut ke model klasifikasi yang telah dipilih berdasarkan hasil evaluasi sebelumnya.



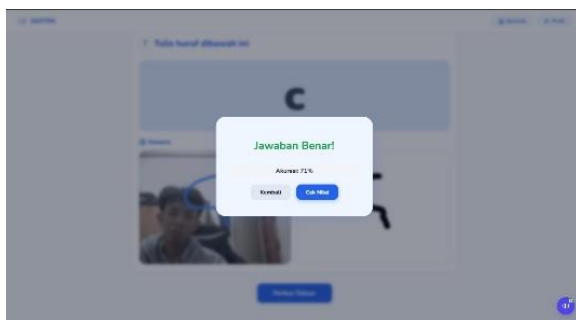
Gambar 5. Uji coba 1 penulisan pada sistem

Pada gambar pertama, sistem menampilkan instruksi “Tulis huruf di bawah ini” dengan target huruf C. Pengguna kemudian menulis huruf C melalui gerakan tangan di depan kamera. Bentuk huruf C yang dihasilkan masih kurang sempurna karena terdapat tambahan garis atau cabang pada bagian bawah huruf.

Cabang tersebut dapat muncul akibat gerakan tangan yang berlebih, posisi jari yang tidak stabil, atau sistem masih mendeteksi pergerakan setelah bentuk huruf selesai dibuat.

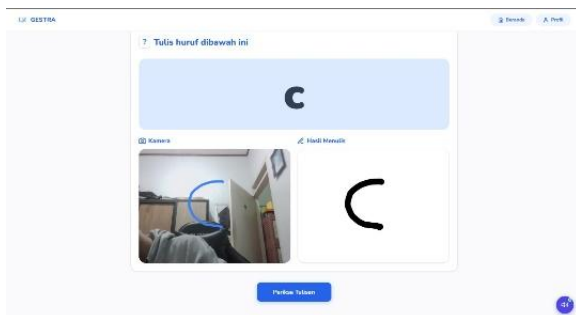
Gambar 6 menunjukkan hasil klasifikasi dari uji coba pertama. Sistem menampilkan pesan “Jawaban Benar!” dengan nilai akurasi sebesar 71%. Hasil ini menunjukkan bahwa model berhasil mengenali tulisan pengguna sebagai huruf C, tetapi tingkat akurasinya belum maksimal.

Nilai akurasi yang lebih rendah disebabkan oleh bentuk huruf C yang memiliki cabang tambahan pada bagian bawah. Cabang tersebut membuat pola tulisan sedikit berbeda dari bentuk huruf C yang ideal, sehingga mempengaruhi proses ekstraksi fitur oleh model CNN.



Gambar 6. Hasil uji coba 1

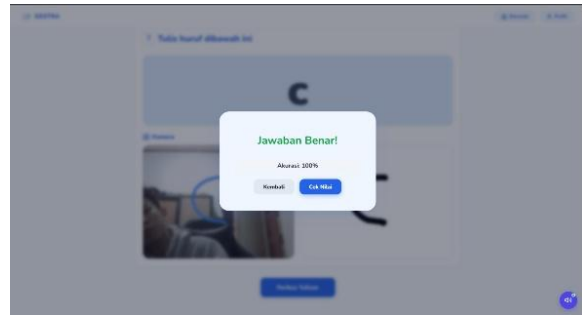
Pada Gambar 7, dilakukan uji coba kembali dengan target huruf yang sama, yaitu huruf C. Pada percobaan ini, pengguna menulis huruf C dengan bentuk yang lebih rapi dan mendekati pola huruf yang seharusnya.



Gambar 7. Uji coba 2 penulisan pada sistem

Gambar 7 menunjukkan hasil klasifikasi dari uji coba kedua. Sistem kembali menampilkan pesan “Jawaban Benar!”, tetapi dengan nilai akurasi yang lebih tinggi, yaitu 100%. Hasil ini menunjukkan bahwa model memiliki tingkat keyakinan penuh terhadap karakter yang ditulis pengguna sebagai huruf C.

Tampilan hasil klasifikasi ini menunjukkan bahwa model yang telah dilatih berhasil dihubungkan dengan sistem secara *real-time*, mulai dari proses deteksi gerakan tangan, pembentukan citra tulisan, hingga prediksi karakter oleh model CNN.



Gambar 7. Hasil uji coba 2

Berdasarkan hasil implementasi tersebut, integrasi model ke sistem dapat berjalan dengan baik karena sistem mampu menerima input tulisan tangan dari pengguna, memprosesnya menjadi citra yang sesuai, dan menampilkan hasil klasifikasi secara langsung. Hasil pengujian juga menunjukkan bahwa kualitas bentuk tulisan sangat memengaruhi nilai klasifikasi. Pada uji coba pertama, huruf C yang memiliki cabang tambahan tetap dikenali sebagai benar, tetapi hanya memperoleh akurasi 71%. Sementara itu, pada uji coba kedua, huruf C yang ditulis dengan bentuk lebih sempurna memperoleh akurasi 100%. Hal ini menunjukkan bahwa semakin jelas dan sesuai bentuk tulisan dengan target karakter, maka semakin tinggi tingkat keyakinan model dalam melakukan klasifikasi.

#### 4.4. Pengujian Sistem Secara *Real-Time*

Untuk mengukur performa dan keandalan model saat diintegrasikan ke dalam sistem secara langsung, dilakukan pengujian operasional secara *real-time*. Pengujian ini dilakukan dengan melibatkan pengguna untuk menuliskan karakter secara bebas di udara menggunakan gerakan jari di depan kamera. Karakter yang diuji terdiri dari 3 kategori utama, yaitu angka (0-9), huruf kapital (A-Z), dan huruf kecil (a-z). Masing-masing kategori karakter diuji sebanyak 30 kali percobaan dengan variasi bentuk tulisan tangan. Hasil pengujian sistem secara keseluruhan dirangkum dalam Tabel 4.

Tabel 4. Hasil Pengujian Operasional Sistem Secara *Real-Time*

| Kategori Karakter           | Angka (0-9)    | Huruf Kapital (A-Z) | Huruf Kecil (a-z) | Total   |
|-----------------------------|----------------|---------------------|-------------------|---------|
| Model Klasifikasi           | Sequential CNN | MobileNet V2        | MobileNet V2      |         |
| Jumlah Percobaan            | 30             | 30                  | 30                | 90      |
| Prediksi Benar              | 28             | 27                  | 26                | 81      |
| Prediksi Salah              | 2              | 3                   | 4                 | 9       |
| Tingkat Keberhasilan Sistem | 93,33%         | 90,00%              | 86,67%            | 90,00 % |

Berdasarkan Tabel 4, hasil pengujian operasional menunjukkan bahwa sistem mampu mengenali input gerakan tangan dengan rata-rata tingkat keberhasilan mencapai 90,00%. Model Sequential CNN yang diterapkan pada kategori angka memberikan performa tertinggi dengan tingkat keberhasilan 93,33%, disusul oleh model MobileNetV2 pada kategori huruf kapital sebesar 90,00%, dan kategori huruf kecil sebesar 86,67%.

## 5. KESIMPULAN DAN SARAN

Sistem menunjukkan bahwa teknologi *hand tracking* dan model *Convolutional Neural Network* (CNN) mampu menangkap gerakan tangan pengguna melalui kamera dan mengolahnya menjadi citra yang dapat diklasifikasikan oleh model. Pengujian terhadap CNN Sequential dan MobileNetV2 menunjukkan bahwa setiap model memiliki karakteristik performa yang berbeda pada masing-masing kategori karakter, sehingga pemilihan model berdasarkan jenis karakter dapat membantu meningkatkan hasil klasifikasi. Selain itu, kestabilan gerakan tangan dan kemiripan visual antar karakter menjadi faktor yang memengaruhi hasil prediksi sistem. Secara umum, integrasi MediaPipe Hands dan CNN dapat diterapkan secara efektif pada sistem klasifikasi tulisan tangan berbasis *air writing* dan berpotensi menjadi media pembelajaran interaktif bagi anak-anak disleksia. Pengembangan selanjutnya dapat dilakukan dengan menambah jumlah serta variasi dataset dan melakukan pengujian langsung kepada pengguna sasaran agar sistem lebih sesuai dengan kebutuhan belajar anak.

## DAFTAR PUSTAKA

- [1] E. Goyen, C. Grindle, V. Totsika, and S. Sapiets, "Implementation of an Adapted Handwriting Without Tears® Programme for Children with Developmental Disabilities in a Special Education Setting Journal: Tizard Learning Disability Review," *Tizard Learning Disability Review*, vol. 29, no. 1, Apr. 2022, doi: <https://doi.org/10.1108/TLDR-07-2023-0015>.
- [2] A. Ananda, B. Bahrin, D. Ramadhani Anas, H. Nur, F. Psikologi, and U. Negeri Makassar, "SHARING SESSION: PARENTAL SUPPORT KEPADA ORANG TUA YANG MEMILIKI ANAK DISLEKSIA," *Jurnal Kebajikan : Jurnal Pengabdian Masyarakat*, vol. 1, no. 1, Nov. 2022.
- [3] R. Wiryasaputra, "Pengklasifikasian Citra Tulisan Anak Melalui Metode CNN sebagai Pendukung Pendeteksian Dini Disgrafia," *Jurnal Telekomunikasi dan Komputer*, vol. 11, no. 3, p. 233, Dec. 2021, doi: 10.22441/incomtech.v11i3.13769.
- [4] M. Z. Iqbal and A. G. Campbell, "From luxury to necessity: Progress of touchless interaction technology," *Technol. Soc.*, vol. 67, Nov. 2021, doi: 10.1016/j.techsoc.2021.101796.
- [5] S. Suki, C. Lubis, and J. Pragantha, "Implementasi Virtual Mouse Berbasis Hand Gesture Recognition dengan MediaPipe dan Convolutional Neural Network under a Creative Commons Attribution-NonCommercial ShareAlike 4.0 International (CC BY-NC-SA 4.0)," vol. 10, no. 1, 2025, [Online]. Available: <https://journal.uin-alauddin.ac.id/index.php/instek>
- [6] R. D. T. Putri, L. Lasmadi, A. Kusumaningrum, R. Nurdin, and Y. Astuti, "Hand Gesture-Based Human-Computer Interaction using MediaPipe and OpenCV," *Aviation Electronics, Information Technology, Telecommunications, Electricals, and Controls (AVITEC)*, vol. 7, no. 2, p. 223, Jul. 2025, doi: 10.28989/avitec.v7i2.3132.
- [7] I. W. Suartika E. P., A. Y. Wijaya, and R. Soelaiman, "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101," vol. 5, 2016.
- [8] I. Karo, J. Karo, Yunianto, Hariyanto, and M. Falah, "KLASIFIKASI MUTU FISIK TEMPE MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK(CNN)," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 8, no. 2, Dec. 2023.
- [9] R. R. Santosa, A. Sasikirana, Z. L. Ramadhan, Y. Putra, B. Satria, and A. Nugroho, "Perbandingan Kinerja MobileNetV2 dan ResNet50V2 Dalam Klasifikasi Tingkat Kematangan Tomat," *Indonesian Journal of Science*, vol. 3, no. 1, 2025.
- [10] C. S. Ley, A. Syammi, and B. Ab Ghafar, "Handwritten Character Recognition Using Convolutional Neural Network MALAYSIA \*Corresponding Author Designation," *Progress in Engineering Application and Technology*, vol. 2, no. 1, pp. 593–611, 2021, doi: 10.30880/peat.2021.02.01.058.
- [11] L. Abdiansah, A. Eviyanti, N. Lutvi Azizah, P. Studi Informatika, and F. Sains dan Teknologi, "Implementation of Convolutional Neural Networks Algorithm for Javanese Handwriting Recognition Penerapan Algoritma Convolutional Neural Networks untuk Pengenalan Tulisan Tangan Aksara Jawa," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 2, Apr. 2025, doi: <https://doi.org/10.57152/malcom.v5i2.1814>.
- [12] C. E. Espinoza-Vicuña and Y. Nuñez-Medrano, "Dyslexia recognition by handwriting using Convolutional Neural Networks," in *Proceedings of the LACCEI international Multi-conference for Engineering, Education and Technology*, Latin American and Caribbean Consortium of Engineering Institutions, 2023, doi: 10.18687/laccei2023.1.1.1346.
- [13] Mr. M. Nagaraju, "Digital Handwriting Recognition Using Hand Tracking by Using

- Media Pipe and OPENCV Libraries,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 7, pp. 659–666, Jul. 2022, doi: 10.22214/ijraset.2022.44647.
- [14] S. Dewi, F. Ramadhani, and S. Djasmayena, “Klasifikasi Jenis Jerawat Berdasarkan Gambar Menggunakan Algoritma CNN (Convolutional Neural Network),” *Hello World Jurnal Ilmu Komputer*, vol. 3, no. 2, pp. 68–73, Jul. 2024, doi: 10.56211/helloworld.v3i2.518.
- [15] J. Brilliantio, N. Santosa, G. Ardian, and L. Hakim, “Penerapan Convolutional Neural Network untuk Handwriting Recognition pada Aplikasi Belajar Aritmatika Dasar Berbasis Web,” *Jurnal Teknik Informatika Unika St. Thomas (JTIUST)*, vol. 5, no. 2, pp. 2657–1501, Dec. 2020