

Teachable Machine: Deteksi Dialek Sumba Timur (Kambera) Menggunakan Layanan *Open Source*

Edwin Ariesto Umbu Malahina

Program Studi Teknik Informatika D3, STIKOM Uyelindo Kupang, Kota Kupang, Nusa Tenggara Timur 85111 INDONESIA (tel. 0380 – 8430672, fax. 0380-8430672, email: edwinariesto@gmail.com)

[Diterima: 29 Mei 2023, Direvisi: 10 Oktober 2023]

Corresponding Author: Edwin Ariesto Umbu Malahina

INTISARI — Studi penelitian ini dilaksanakan untuk mengembangkan sistem deteksi fonetik dialek Kambera dari bahasa lokal Sumba Timur berbasis *framework* TensorFlow yang akan diimplementasikan pada aplikasi *mobile*. Sebagai bagian dari inisiatif ini, penelitian ini telah mengompilasikan sebuah *dataset* representatif dari sampel fonetik dialek Kambera. Tujuan utamanya adalah untuk meningkatkan presisi dalam pengenalan fonetik. Menggunakan dialek Kambera sebagai studi kasus, data diekstraksi dan dilatih menggunakan layanan Teachable Machine berbasis *open-source*. Penelitian ini mengadopsi pendekatan berbasis *convolutional neural network* (CNN) yang dikombinasikan dengan metode *Mel frequency cepstral coefficients* (MFCC) untuk ekstraksi fitur yang lebih akurat. Setelah proses pengumpulan data, pelatihan model, pengujian, dan implementasi, model diintegrasikan ke dalam *platform* Android untuk kepentingan publik yang ingin memahami bahasa dialek Kambera dari Sumba Timur. Pengembangan dan pengujian sistem ini dirancang untuk mendeteksi dan menginterpretasikan fonetik bahasa lokal Sumba Timur dengan dialek Kambera, memberikan kontribusi signifikan dalam optimalisasi pengenalan fonetik, serta menyediakan *dataset* untuk kepentingan penelitian berkelanjutan. Sistem ini juga berfungsi sebagai alat edukasi linguistik yang dapat diakses dan mendukung inklusi serta diversifikasi linguistik dalam teknologi digital. Evaluasi empiris menunjukkan bahwa tingkat presisi deteksi dialek rata-rata secara keseluruhan mencapai 98,3% hingga 99,6%, dengan tingkat kepuasan pengujian kepada pengguna mencapai 99,33%. Hasil ini menegaskan bahwa sistem yang dikembangkan memiliki kapabilitas deteksi yang sangat efisien dan baik.

KATA KUNCI — *Convolutional Neural Network, Mel Frequency Cepstral Coefficients, Teachable Machine, Deteksi Suara, Sumba Timur, Open Source, TensorFlow, Dialek Kambera.*

I. PENDAHULUAN

Layanan berbasis *open source* adalah jenis layanan atau perangkat lunak yang menggunakan sumber terbuka. Tujuan utamanya adalah memungkinkan pengembang dan komunitas untuk berkontribusi dalam memperbaiki *bug* serta menambahkan dan memanfaatkan fitur ataupun fungsi [1], [2]. Layanan tersebut juga dapat berkolaborasi dengan layanan lain yang berbeda [3]-[5]. Teachable Machine adalah sebuah *platform* berbasis *open source* yang memfasilitasi pengembangan model klasifikasi bahasa konteks pembelajaran mesin dengan memanfaatkan pustaka TensorFlow. *Platform* ini mengadopsi teknik *deep learning* dalam jaringan saraf tiruan untuk mengenali pola suara. Dalam kajian ini, upaya dikonsentrasikan pada simulasi pengenalan fonetik dari bahasa daerah Sumba Timur, dengan fokus khusus pada terminologi dasar. Studi terdahulu telah memvalidasi efikasi Teachable Machine, mencatat tingkat akurasi, presisi, dan sensitivitas deteksi yang impresif, yang berfluktuasi dalam kisaran 90% hingga 100% [6]-[8]. Tingkat akurasi yang sangat baik ini menjadikan Teachable Machine sebagai layanan yang direkomendasikan dalam penelitian ini.

Penelitian ini mengidentifikasi isu kunci yang menjadi fokus dalam topik penelitian, yaitu urgensi pelestarian bahasa daerah sebagai bagian integral dari warisan budaya bangsa. Di Indonesia, terdokumentasi adanya 718 bahasa daerah yang berstatus sebagai warisan budaya yang perlu dipertahankan dan dilindungi. Namun, sejumlah bahasa daerah saat ini menghadapi krisis yang mendalam. Bahasa-bahasa tersebut mengalami penurunan drastis dalam hal jumlah penutur asli dan tidak lagi ditransmisikan kepada generasi muda. Hal ini menimbulkan risiko kehilangan warisan linguistik yang berharga, yang dapat mengikis kekayaan dan keanekaragaman budaya bangsa. Oleh karena itu, ada kebutuhan mendesak untuk mengambil langkah-langkah proaktif dalam memastikan

transmisi bahasa daerah kepada generasi berikutnya untuk mempertahankan keberagaman linguistik dan mempromosikan pelestarian bahasa daerah [9], khususnya bahasa Sumba Timur, yang terdiri atas beberapa dialek, yaitu dialek Rindi, Lumbu Manggit, Kambera, Melolo, Uma Ratu Nggai (Umbu Ratu Nggai), Lewa, Kanatang, Mangili-Waijelo (Wai Jilu, Waidjelu, Rindi, Waijelo), dan Sumba Selatan [10]. Saat ini, dialek-dialek tersebut menghadapi krisis kehilangan penutur asli dan transmisi ke generasi berikutnya. Sebagai solusi, penelitian ini mengusulkan pemanfaatan kecerdasan buatan melalui layanan Teachable Machine, yang memungkinkan pengembangan sistem deteksi suara berbasis pembelajaran mesin dengan memanfaatkan pustaka TensorFlow. Penelitian menciptakan potensi untuk mengembangkan *dataset* suara bahasa daerah yang dapat memfasilitasi penelitian lebih lanjut dalam upaya pelestarian bahasa daerah serta melakukan optimalisasi pengenalan suara dengan mengombinasikan metode *Mel-frequency cepstral coefficients* (MFCC) di dalam *framework* TensorFlow.

Penelitian-penelitian sebelumnya masih berfokus pada konsep pemetaan dialek bahasa [11]-[13], kesantunan dalam penuturan bahasa [14], sejarah bahasa [15], sampai pada antropologi linguistik bentuk lingual [16], sedangkan penelitian saat ini berfokus pada optimalisasi kemampuan deteksi dialek yang mengombinasikan algoritma *convolutional neural network* (CNN) yang ada pada *framework* TensorFlow dengan pendekatan metode MFCC yang akan diuji perbandingan dan keakuratannya. Di sisi lain, penelitian ini juga bertujuan untuk melakukan digitalisasi dalam bentuk aplikasi dan pembuatan sampel *dataset* (basis data) dialek Kambera sebagai upaya pelestarian bahasa daerah Sumba Timur untuk mencegah kepunahan bahasa dengan memanfaatkan sistem cerdas yang mampu memberikan kontribusi dalam pendataan bahasa daerah Sumba Timur.

Penelitian ini juga diharapkan bermanfaat bagi siapa pun yang ingin mempelajari bahasa Sumba Timur. Selain itu, tidak menutup kemungkinan bagi penelitian selanjutnya untuk memanfaatkan *dataset* yang disediakan, baik menggunakan maupun mengembangkannya kembali.

Pembahasan masalah dan studi literatur yang dilakukan menjadi pedoman dalam pengembangan serta berkontribusi dalam melakukan proses pengoptimalan deteksi suara menggunakan MFCC yang akan ditambahkan pada kerangka kerja *framework* TensorFlow. Oleh karena itu, perangkat lunak berbasis *mobile* yang dikembangkan dapat mendeteksi dialek suara Kambera lebih akurat.

II. METODOLOGI

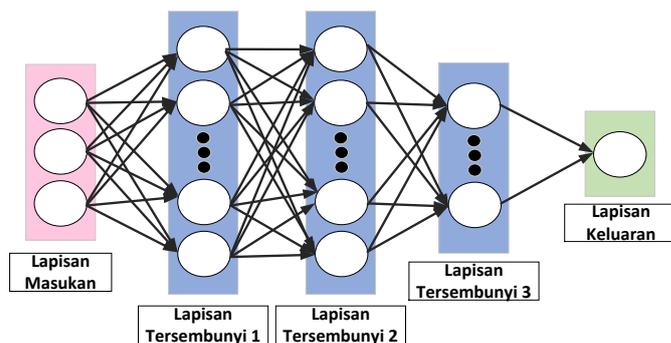
Metodologi yang diadopsi dalam fase pengumpulan data ini mengikuti pendekatan eksploratif terhadap dialek Kambera. Proses ini akan menghasilkan data berupa transkrip percakapan serta rekaman audio yang diperoleh dari tiga narasumber masyarakat lokal yang memiliki keahlian dalam dialek Kambera. Tujuan utama pendekatan ini adalah untuk mengidentifikasi dan mendokumentasikan variasi yang ada dalam kosakata, pengucapan, struktur tata bahasa, intonasi, serta fitur linguistik lain dari dialek Kambera. Analisis ini bertujuan untuk membedakan dan memahami nuansa unik dialek Kambera serta menciptakan fondasi yang kuat untuk penelitian linguistik lebih lanjut.

Dalam pengembangan dan pengujian perangkat lunak, digunakan metode yang memanfaatkan layanan *framework* TensorFlow. *Framework* ini menyediakan infrastruktur yang diperlukan dan memiliki keunggulan dalam efisiensi pengembangan kode dan model pembelajaran mesin [17], [18]. Proses pelatihan *dataset* dalam TensorFlow dikendalikan oleh CNN, yang merupakan metode klasifikasi data [19]. Dalam konteks ini, data sampel suara dialek diekstraksi dan fitur-fiturnya diklasifikasikan secara terpisah, sehingga membentuk sistem deteksi berbasis pembelajaran [20]. Dalam penelitian ini, data linguistik Kambera dilatih menggunakan Teachable Machine. Teachable Machine menyediakan algoritma CNN yang dapat digunakan untuk mengklasifikasikan suara dialek Kambera untuk mendapatkan tingkat akurasi dalam bentuk persentase, serta menambahkan metode MFCC yang akan membantu dalam ekstraksi fitur yang lebih baik dari data suara, sehingga dapat meningkatkan akurasi pengenalan suara.

A. MODEL TRANSFER LEARNING

Framework TensorFlow.js merupakan perpustakaan *open source* untuk pembelajaran mesin yang operasional dalam lingkungan JavaScript. Hal ini memungkinkan pengembangan untuk memasukkan elemen pembelajaran mesin dalam aplikasi web dan *mobile* serta kompatibel dengan berbagai perangkat yang mendukung JavaScript [21]. *Transfer learning* adalah sebuah pendekatan pengetahuan sebelumnya dari model pembelajaran mesin, seperti CNN, yang digunakan ulang untuk memahami tugas-tugas baru yang serupa. Dalam konteks ini, model yang telah terlatih dalam pengenalan suara dapat digunakan kembali untuk klasifikasi suara lainnya, sehingga menghemat waktu dan sumber daya pelatihan.

CNN sendiri telah diaplikasikan pada berbagai jenis penelitian di berbagai bidang dengan studi kasus berbeda, seperti pada deteksi suara untuk melakukan klasifikasi [22], [23], baik mendeteksi suara depresi atau emosional [24], [25], mendeteksi musik [26], merealisasikan keamanan sistem berbasis suara [27], [28], mendeteksi fonem bahasa [29], [30], mendeteksi suatu penyakit yang dikenali dari suara [31]-[34],



Gambar 1. Arsitektur *convolutional neural network*.

sampai pada pengenalan suara hewan [35], [36]. Selain itu, masih banyak studi kasus penelitian lainnya yang dapat ditemukan dan dijadikan riset berkelanjutan ke depan.

Gambar 1 menampilkan arsitektur CNN yang dimulai pada bagian kiri (kotak merah) dengan setiap lapisannya berisi sejumlah neuron (lingkaran putih). *Multi-layer perceptron* (MLP) menerima data masukan satu dimensi dan mengirimkannya melalui jaringan untuk menghasilkan keluaran. Setiap hubungan antara neuron pada dua lapisan yang bersebelahan memiliki parameter bobot satu dimensi yang memengaruhi kualitas model. Pada setiap lapisan, data masukan mengalami operasi linear dengan memperhitungkan nilai bobot yang ada. Kemudian, hasil komputasi tersebut ditransformasikan menggunakan fungsi aktivasi yang bersifat nonlinear [37].

B. ALUR KERJA SISTEM

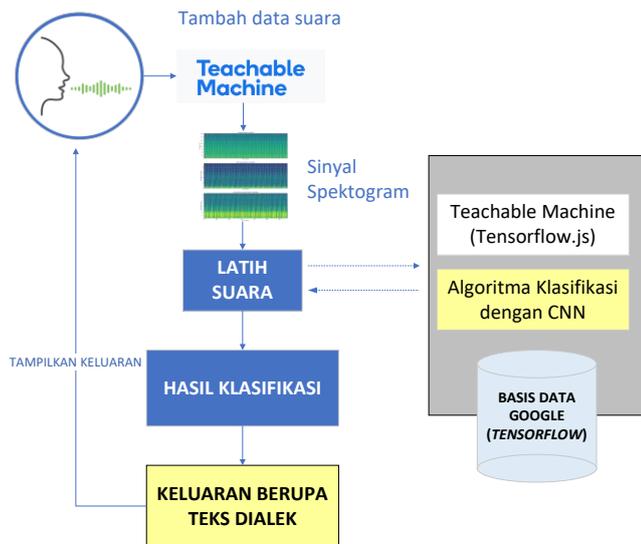
Dalam penelitian ini, proses pengenalan audio dialek Kambera dalam bahasa Sumba Timur dilakukan menggunakan *framework* TensorFlow dari layanan Teachable Machine sebagai alat pengolahan data. Konsep yang digunakan dalam alat ini didasarkan pada prinsip pembelajaran mesin, yaitu sebuah sistem dapat belajar dan menganalisis data tanpa harus diprogram secara eksplisit. Hal ini bertujuan untuk membantu berbagai kalangan dalam memahami konsep alur kerja pembelajaran mesin dengan menciptakan dan menggunakan model eksperimen klasifikasi pengembang secara mandiri [38].

Sistem kerja ini dimulai dengan menyiapkan sampel suara bahasa Sumba Timur dialek Kambera untuk data klasifikasi suara. Kemudian, dilanjutkan dengan menyediakan sampel derau audio untuk menghilangkan suara bising saat pengujian. Sampel suara dan derau audio diimpor ke Teachable Machine dan diubah menjadi spektrogram. Lalu, dilakukan pelatihan data otomatis di Teachable Machine. Selanjutnya, dilakukan pengujian akurasi. Jika sukses, pengembang dapat mengeksplor data ke Javascript untuk dikembangkan menjadi aplikasi.

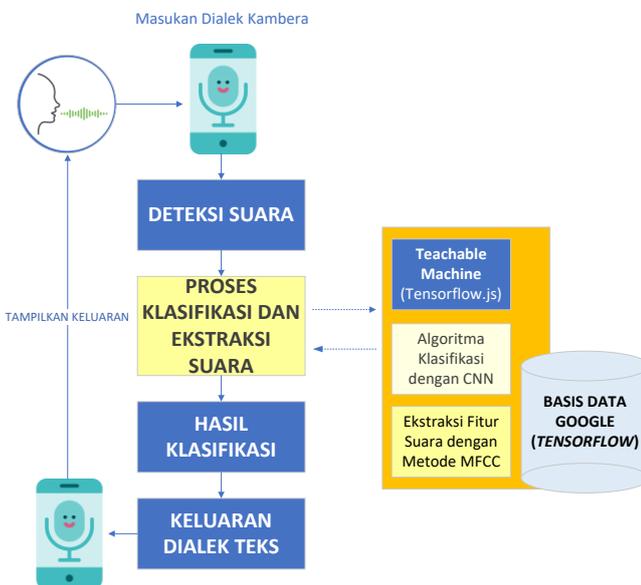
C. PELATIHAN MODEL

Diagram alir program berfungsi sebagai representasi visual dari program komputer melalui simbol-simbol diagram alir. Jenis diagram alir ini menggambarkan langkah-langkah proses suatu program dengan detail serta memperlihatkan langkah-langkah penyelesaian unit masalah yang saling terkait dan berinteraksi satu sama lain guna mencapai tujuan tertentu. Diagram alir ini menunjukkan aliran sistem atau fungsi yang dijalankan oleh sistem.

Diagram alir proses ditunjukkan pada Gambar 2. Proses dimulai dengan pengguna memasukkan suara bahasa Sumba Timur dialek Kambera ke aplikasi Teachable Machine untuk disimpan sebagai sinyal audio dan derau audio. Lalu, sinyal diubah menjadi spektrogram dan sinyal derau untuk



Gambar 2. Diagram alir kelola data dialek suara ke dalam website Teachable Machine.

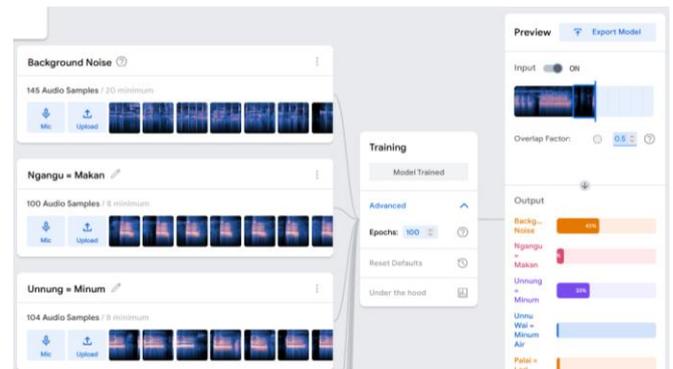


Gambar 3. Diagram alir pengguna melakukan deteksi dialek Kambera dengan aplikasi mobile web yang dikembangkan.

mendeteksi dan menghapus gangguan suara nantinya. Suara tersebut dilatih untuk pengenalan suara. Pelatihan dikelola oleh algoritma CNN untuk klasifikasi tanpa metode MFCC untuk ekstraksi fitur audio. Data hasil kemudian disimpan di basis data Google dan sistem menghasilkan klasifikasi suara. Selanjutnya, hasil klasifikasi ditampilkan sebagai teks sesuai masukan suara pengguna.

Hasil proses pada Gambar 2 kemudian diekspor ke dalam bentuk file TensorFlow.js. Nantinya akan ditambahkan secara manual fungsi metode MFCC di dalamnya untuk ekstraksi fitur audio.

Tahap-tahap pada Gambar 3 menggambarkan proses, yang dimulai ketika pengguna memasukkan suara ke dalam aplikasi mobile. Selanjutnya, suara tersebut mengalami proses pengolahan dan identifikasi dengan memanfaatkan algoritma CNN, yang didukung oleh ekstraksi fitur suara menggunakan metode MFCC sebelumnya pada file TensorFlow.js. Penambahan metode MFCC ini bertujuan untuk meningkatkan kualitas data suara yang ada, yang pada akhirnya akan meningkatkan akurasi dalam pengenalan suara. Jika hasil



Gambar 4. Data dialek suara dilatih terlebih dahulu pada layanan Teachable Machine.

identifikasi berhasil ditemukan, aplikasi mobile akan menampilkan teks yang sesuai dengan akurasi presentasi pengenalan suara ke artikulasi teks dari bahasa Kambera yang dimasukkan.

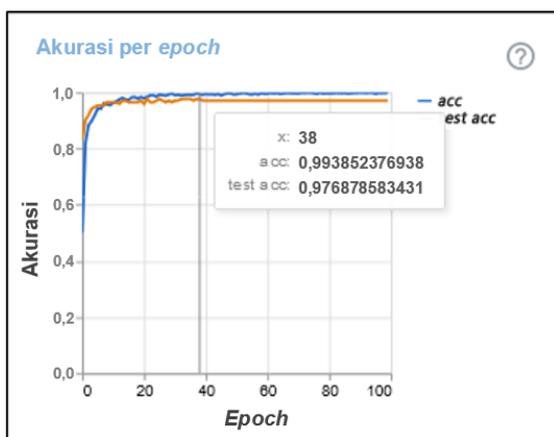
III. HASIL DAN PEMBAHASAN

A. PENGUJIAN TERHADAP MODEL DIALEK SUARA

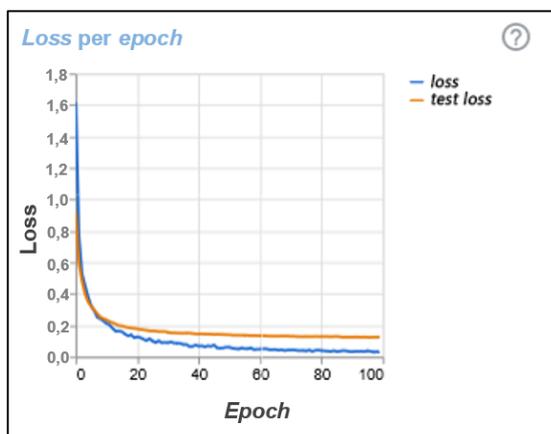
Pengujian terhadap model dialek suara Kambera dalam Teachable Machine melibatkan penggunaan dataset sebanyak seratus sampel per kosakata secara terpisah untuk menguji keakuratan dan kinerja model suara. Kemudian, sampel suara dimasukkan ke Teachable Machine yang telah dilatih dan model akan mengklasifikasikan suara ke dalam label (keterangan nama kelas audio) yang telah ditentukan. Tujuan pengujian ini adalah untuk mengukur kemampuan model dalam mengenali dan mengklasifikasikan suara dengan akurasi tinggi serta memberikan informasi tentang keandalan dan efektivitas model tersebut.

Pada Gambar 4, dalam konteks kerangka kerja yang disediakan oleh Teachable Machine, penelitian ini melibatkan pengujian yang komprehensif terhadap seratus spektrogram yang mewakili setiap label dialek yang relevan. Pengambilan sampel suara dilakukan dengan mengikuti standar durasi rekaman selama 2 s per sampel dialek, dengan opsi untuk merekam ulang jika diperlukan. Selanjutnya, model yang telah melalui proses pelatihan menggunakan platform Teachable Machine diuji menggunakan kelas-kelas dataset yang telah ditetapkan sebelumnya. Evaluasi kinerja model dilakukan dengan menganalisis sejumlah metrik yang meliputi tingkat akurasi, presisi, dan sensitivitas. Hal ini bertujuan untuk mengukur kemampuan model tersebut dalam mengenali dan mengklasifikasikan berbagai varian dialek dengan akurat, untuk memastikan validitas dan efektivitas dari solusi yang dikembangkan dalam konteks pengenalan suara dialek Kambera.

Dalam konteks pengenalan suara menggunakan Teachable Machine, "akurasi per epoch" merujuk pada metrik yang mengukur persentase jumlah data yang diklasifikasikan dengan benar oleh model selama satu siklus pelatihan (epoch). Sementara itu, "loss per epoch" mengukur besarnya kesalahan prediksi model pada data pelatihan selama satu epoch. Tujuan dalam pelatihan model adalah meningkatkan akurasi per epoch sambil mengurangi loss per epoch, yang dapat dicapai melalui penyesuaian parameter model seperti laju pembelajaran atau perubahan pada arsitektur model untuk memastikan bahwa model dapat mempelajari dan mengklasifikasikan suara dengan lebih baik sesuai dengan tujuan aplikasi yang diinginkan. Hasilnya diperlihatkan pada Gambar 5 dan Gambar 6.



Gambar 5. Akurasi per epoch.



Gambar 6. Loss per epoch.

Gambar 5 merupakan tampilan visual grafik akurasi per epoch dalam tangkapan layar pada situs Teachable Machine, yang menampilkan metrik “acc” untuk mengukur akurasi dalam pelatihan model dan menunjukkan kemampuan model memprediksi data pelatihan. Nilai “acc” berkisar antara 0,97 hingga 1, yang berarti model memiliki akurasi antara 97% hingga 100% terhadap data pelatihan. Sementara itu, “test acc” adalah akurasi validasi yang mengukur kemampuan model memprediksi data uji yang tidak digunakan dalam pelatihan. Nilai “test acc” berkisar antara 0,97 hingga 0,99, menunjukkan akurasi antara 97% hingga 99% terhadap data validasi.

Gambar 6 menunjukkan grafik perubahan loss per epoch dalam tangkapan layar pada situs Teachable Machine, yang menggambarkan kemampuan model dalam melakukan prediksi terhadap seluruh dataset pada setiap iterasi pelatihan. Metrik “loss” mengukur selisih prediksi model dari nilai sebenarnya. Makin rendah nilai loss, makin baik model dalam melakukan prediksi, dengan rentang nilai loss antara 0,07 hingga 0,03 (7% hingga 3%).

Selain itu, terdapat metrik test loss, yang mengukur kesalahan model pada data uji atau validasi setiap epoch. Test loss ini penting untuk mengevaluasi kemampuan model dalam menerapkan pembelajaran dari data pelatihan ke data yang belum pernah digunakan sebelumnya. Hasil test loss, dengan rentang nilai antara 0,66 hingga 0,35 (66% hingga 35%) pada data validasi, menunjukkan penurunan yang konsisten dan baik.

B. PENGUJIAN LANGSUNG DIALEK SUARA

Langkah berikutnya dalam penelitian ini adalah evaluasi empiris terhadap aplikasi yang telah dikembangkan, dengan fokus khusus pada pengenalan sepuluh sampel dialek Kambera.

Pada tahap awal, pengujian dilakukan dengan menggunakan framework TensorFlow yang telah dilengkapi dengan algoritma CNN yang telah ada guna mengukur kinerja pengenalan suara dialek. Meskipun hasil awal menunjukkan tingkat kinerja yang dapat diterima dengan masih terdapat beberapa sampel yang mencatat tingkat akurasi relatif rendah (yaitu “Unnu Wai” dengan akurasi minimal 79% dan “Laku” dengan akurasi minimal 85%), terdapat ruang untuk perbaikan yang signifikan. Sebagai respons terhadap temuan tersebut, metode MFCC diterapkan secara terintegrasi untuk mengubah sinyal suara menjadi serangkaian koefisien yang merepresentasikan karakteristik suara dengan lebih baik.

Implementasi strategi ini dilakukan dengan menambahkan metode MFCC dan telah menghasilkan peningkatan yang nyata dalam akurasi pengenalan suara, yaitu mencapai tingkat minimal 97%. Melalui eksplorasi mendalam terhadap kombinasi antara arsitektur CNN dan ekstraksi fitur MFCC, penelitian ini berhasil menciptakan sebuah model pengenalan suara yang andal dan akurat. Temuan ini mengonfirmasi bahwa kolaborasi antara CNN dan MFCC dapat secara efektif memperbaiki pengenalan dialek suara Kambera dan meningkatkan stabilitas serta ketepatan dalam proses identifikasi fonetik. Analisis lebih lanjut menunjukkan bahwa integrasi MFCC membantu dalam mengurangi ambiguitas fonetik dan variabilitas intraspecies dalam data suara, yang pada gilirannya memfasilitasi identifikasi dialek yang lebih tepat.

Keunggulan utama dari pendekatan ini terletak pada cara CNN memanfaatkan hierarki fitur spasial dari data suara, sementara MFCC, dengan kemampuannya dalam karakterisasi sifat frekuensi suara, menyediakan representasi yang baik dan informatif dari karakteristik fonetik yang inheren. Kombinasi sinergis antara CNN dan MFCC menciptakan suatu paradigma baru dalam bidang pengenalan suara. Kedua metode ini saling melengkapi untuk menciptakan sebuah framework yang kuat dalam identifikasi varietas dialek dengan tingkat akurasi yang tinggi. Dalam konteks ini, metode MFCC bertindak sebagai alat yang efisien untuk mereduksi dimensi data suara, sedangkan arsitektur CNN mengambil keuntungan dari struktur lokal dan global dari data tersebut untuk melaksanakan pengenalan pola yang efektif, sehingga memperlihatkan bahwa integrasi strategis antara CNN dan MFCC dapat berfungsi sebagai fondasi yang kuat untuk pengembangan futuristik dalam teknologi pengenalan suara, khususnya dalam konteks multibahasa dan multidialek.

Penelitian ini juga mempertimbangkan fakta bahwa varietas dalam data pelatihan dapat memengaruhi kinerja model. Penelitian ini menemukan bahwa dengan memiliki dataset pelatihan yang beraneka ragam, model yang dihasilkan lebih mampu memahami dan menggeneralisasi berbagai dialek suara Kambera, sehingga meningkatkan kinerja pengenalan suara secara signifikan. Penelitian ini memiliki implikasi penting terhadap penelitian futuristik dalam bidang pengenalan suara, terutama dalam upaya mengurangi kesenjangan teknologi di antara berbagai dialek dan bahasa.

Pada Tabel I terlihat sepuluh dataset sampel dialek kosa kata Kambera yang diuji dengan menggunakan rata-rata dataset sebanyak seratus sampel spektrogram. Terdapat akurasi minimal dari setiap dialek yang ada dan rata-rata minimal koreksi akurasi dialek, yaitu sebesar 92,3% (kurang akurat). Kemudian, rata-rata akurasi maksimal dari dialek yang diuji menggunakan TensorFlow Standar atau bawaan adalah 98,2% (akurat). Jika ditambahkan metode MFCC, yang telah

TABEL I
PENGUJIAN DETEKSI DIALEK MENGGUNAKAN ALGORITMA CNN DARI TENSORFLOW DAN METODE MFCC

No	Arti Dialek (Bahasa Indonesia)	Dialek Kampera (Sumba Timur)	Dataset	CNN - TensorFlow		MFCC	
				Akurasi Minimal (%)	Akurasi Maksimal (%)	Akurasi Minimal (%)	Akurasi Maksimal (%)
1	Makan	Ngangu	100	99	100	99	100
2	Minum	Unnung	100	90	100	99	100
3	Minum Air	Unnu Wai	100	79	89	98	98
4	Lari	Palai	100	91	98	97	100
5	Seperti	Tuna Kadi	100	96	100	98	100
6	Pergi	Laku	100	85	99	97	99
7	Kesana	Laku Nolua	100	98	100	98	99
8	Kemari	Mai Yohu	100	94	99	99	100
9	Nama Saya	Na Tamu Ngu	100	98	98	99	100
10	Kamu	Nyummy	100	93	99	99	100
Rerata akurasi			1.000	92,3 %	98,2 %	98,3%	99,6%

TABEL II
PENGUJIAN PENGGUNA

No	Pertanyaan	Sangat Setuju	Setuju	Netral	Tidak Setuju	Sangat Tidak Setuju
		5	4	3	2	1
1	Apakah antarmuka pengguna di dalam aplikasi mudah digunakan dan dipahami menu-menunya?	10	5	5	0	0
2	Apakah deteksi suara yang di input mudah dikenal oleh sistem aplikasi?	6	12	1	1	0
3	Apakah aplikasi sangat cepat dalam proses deteksi suara dialek yang dimasukan?	16	4	0	0	0
4	Apakah aplikasi yang dikembangkan bermanfaat untuk kalangan umum?	2	6	10	2	0
5	Apakah tingkat akurasi berpengaruh dalam hasil deteksi suara pengguna?	16	4	0	0	0
6	Apakah pemanfaatan repository sistem cerdas dapat membantu dalam pelestarian bahasa lokal daerah khususnya dialek Kampera dari Sumba Timur?	2	18	0	0	0
7	Apakah anda terbantu dengan adanya sistem yang dikembangkan ini?	7	12	1	0	0
Jumlah		59	61	17	3	0
Rumus Likert = $T \times Pn$ (Jumlah responden \times deret skor Likert)		295	244	51	6	0

berkontribusi dalam mengekstraksi fitur yang lebih baik dari data suara, akurasi suara meningkat menjadi 98,3% - 99,6%. Nilai ini lebih tinggi dari pengujian menggunakan TensorFlow standar yang menggunakan algoritma CNN, sehingga kolaborasi kedua metode ini sangat optimal dalam meningkatkan akurasi deteksi suara.

Dalam proses penelitian ini, diperoleh juga beberapa catatan penting, yaitu sebagai berikut.

1. Suara dialek yang sama hanya direkam selama 2 s saja untuk pembagian sampel spektrogram (tidak bisa lebih).
2. Suara dialek harus jelas dan keras untuk menghasilkan gelombang audio yang akurat.
3. Suara bising sangat berpengaruh terhadap hasil persentase akurasi.
4. Kosakata yang mirip dengan *dataset* yang sudah ada berpengaruh pada akurasi.
5. Jumlah sampel kelas yang banyak akan memengaruhi waktu proses pelatihan data.

Pada Tabel II, data mengenai pengujian pengguna disajikan. Pengujian ini dilakukan dengan menggunakan kuesioner yang didistribusikan kepada dua kelompok responden, yaitu lima belas orang dari pengguna umum yang dipilih secara acak dan lima orang yang memiliki keahlian dalam dialek Kampera. Tujuan dari pengujian ini adalah untuk menilai tingkat keakuratan pengenalan dialek Kampera oleh sistem. Analisis data dari kuesioner ini dilakukan dengan menggunakan metode skala Likert, yang melibatkan langkah-langkah perhitungan tertentu sesuai dengan rumus yang akan dijelaskan berikutnya:

$$T \times Pn \quad (1)$$

dengan T merupakan jumlah reponden dan Pn merupakan pilihan angka skala Likert.

$$\begin{aligned} \text{Sangat setuju} &= 59 \times 5 = 295 \\ \text{Setuju} &= 61 \times 4 = 244 \\ \text{Netral} &= 17 \times 3 = 51 \\ \text{Tidak setuju} &= 3 \times 2 = 6 \\ \text{Sangat tidak setuju} &= 0 \times 0 = 0 \end{aligned}$$

Jika semua nilai dijumlahkan, diperoleh nilai 596.

Pengujian melalui data kuesioner yang menggunakan perhitungan dengan skala Likert menghasilkan nilai 596. Selanjutnya, nilai ini dibagi dengan jumlah pertanyaan, yaitu enam pertanyaan (sebagaimana tercantum dalam Tabel II). Perhitungan ini menghasilkan nilai sebesar 99,33%, yang masuk dalam kategori "Sangat Setuju" untuk tingkat persetujuan dari responden terhadap pertanyaan-pertanyaan yang diajukan dalam kuesioner tersebut.

C. PENGUJIAN DAN IMPLEMENTASI SISTEM

Berikut ini adalah hasil dari pengembangan sistem pengenalan dialek suara yang telah diintegrasikan ke dalam sebuah aplikasi web *mobile*. Aplikasi ini dirancang untuk memfasilitasi masyarakat dalam mempelajari dan memahami dialek Kampera, yang merupakan salah satu dialek dari bahasa Sumba Timur.

Pada Gambar 7 ditampilkan antarmuka awal aplikasi pengenalan dialek Kampera saat pertama kali diakses oleh pengguna. Pada halaman ini, tersedia grafis yang



Gambar 7. Antarmuka aplikasi saat pertama kali dibuka.

merepresentasikan pulau Sumba Timur dan terdapat tombol “MULAI” yang jika ditekan oleh pengguna, aplikasi akan merespons dan mengarahkan secara langsung ke halaman proses deteksi pengenalan dialek Kambera dari bahasa Sumba Timur. Halaman tersebut merupakan proses utama aplikasi. Pada halaman tersebut, pengguna aplikasi dapat mengucapkan kalimat berbahasa Kambera secara langsung dan sistem akan melakukan respons serta mengenali ucapan pengguna tersebut untuk dicocokkan pada basis data TensorFlow Google. Aplikasi ini dirancang dengan antarmuka pengguna yang intuitif dan mudah digunakan, yang memungkinkan pengguna dengan berbagai tingkat keahlian teknologi untuk berinteraksi dengan sistem secara efektif. Dengan mengintegrasikan teknologi pengenalan suara berbasis pembelajaran mesin, aplikasi ini menawarkan solusi inovatif untuk tantangan pengenalan dialek di wilayah multibahasa, seperti Sumba Timur. Penggunaan basis data TensorFlow Google dalam infrastruktur sistem ini menunjukkan pemanfaatan teknologi awan dan pembelajaran mesin untuk mengatasi masalah lokal dengan sumber daya global.

Pada Gambar 8, halaman pengenalan dialek bahasa Kambera dipresentasikan. Di sini, pengguna dapat memulai proses pengenalan suara dengan menekan tombol “UCAP” dan kemudian mengucapkan kosakata dari dialek Kambera ke dalam aplikasi. Aplikasi ini, yang dilengkapi dengan teknologi pengenalan suara otomatis, akan menganalisis masukan suara dan, jika mengenali, menampilkan terjemahan ke dalam bahasa Indonesia beserta persentase akurasi pengenalan tersebut. Aplikasi ini memfasilitasi pemahaman pengguna terhadap kosakata dialek bahasa Kambera dari Sumba Timur dengan umpan balik akurasi dalam bentuk persentase.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan terkait pengenalan suara dialek Kambera dari bahasa Sumba Timur, beberapa kesimpulan dan rekomendasi kritis telah diidentifikasi. Analisis dilakukan menggunakan *framework* TensorFlow dengan algoritma CNN menunjukkan bahwa akurasi minimal yang dapat dicapai adalah 92,3%, sedangkan akurasi maksimal mencapai 98,2%. Ketika metode ini dikombinasikan dan dioptimalkan dengan pendekatan MFCC, terjadi peningkatan signifikan dalam akurasi rata-rata, yaitu dalam kisaran 98,3% hingga 99,6%. Peningkatan ini menunjukkan bahwa integrasi MFCC dapat memfasilitasi



Gambar 8. Konversi dan akurasi suara ke teks berdasarkan dialek Kambera.

ekstraksi fitur yang lebih efektif, sehingga meningkatkan akurasi pengenalan suara dari model awal. Selanjutnya, pengujian terhadap pengguna menghasilkan penilaian positif dengan tingkat akurasi sebesar 99,33%, yang menegaskan bahwa sistem ini memiliki tingkat deteksi suara yang sangat baik dan siap untuk diimplementasikan.

Dalam pengembangan model pembelajaran ini, terdapat beberapa kendala dan batasan yang dihadapi, yaitu: durasi rekaman suara dialek dibatasi hanya 2 s untuk pembagian sampel spektrogram; masukan suara dialek yang jelas dan lantang diperlukan saat pelatihan data untuk menghasilkan gelombang audio yang akurat; suara bising (derau) memengaruhi hasil akurasi secara signifikan; kosakata yang mirip memengaruhi *dataset* yang ada terhadap tingkat akurasi; waktu proses pelatihan data meningkat dengan bertambahnya jumlah sampel kelas (> 100); dan spesifikasi komputer yang memadai diperlukan untuk mendukung proses pembelajaran. Spesifikasi komputer yang direkomendasikan mencakup prosesor *server multi-core* atau *multi-threading*, seperti Intel Xeon atau AMD EPYC dengan kecepatan *clock* tinggi, memori minimal 16 GB, ruang penyimpanan minimal 20 GB, dan koneksi internet yang stabil. Rekomendasi ini bertujuan untuk mengoptimalkan kinerja sistem, terutama dalam menghadapi peningkatan beban akibat penambahan sampel uji.

KONFLIK KEPENTINGAN

Selama masa pelaksanaan penelitian riset “*Teachable Machine: Deteksi Dialek Sumba Timur (Kambera) Menggunakan Layanan Open Source*” ini, penulis secara pribadi tidak memiliki konflik kepentingan dengan pihak mana pun.

KONTRIBUSI PENULIS

Konseptualisasi, Edwin Ariesto Umbu Malahina; metodologi, Edwin Ariesto Umbu Malahina; analisis formal, Edwin Ariesto Umbu Malahina; pengujian dan pengembangan sistem, Edwin Ariesto Umbu Malahina.

REFERENSI

- [1] E.A.U. Malahina, R.P. Hadjon, dan F.Y. Bisilisin, “Teachable Machine: Real-Time Attendance of Students Based on Open Source System,” *The IJICS (Int. J. Inform., Comput. Sci.)*, Vol. 6, No. 3, hal. 140–146, Nov. 2022, doi: 10.30865/ijics.v6i3.4928.
- [2] A.T. Murray dan J. Baik, “Opensource Spatial Optimization in GIScience for Strategic Positioning,” *Trans. GIS*, Vol. 27, No. 3, hal. 646–662, Mei 2023, doi: 10.1111/tgis.13033.

- [3] J.M.D.S. Dos Santos, C.A.A.P. Abar, dan M.V. de Almeida, "Automatic Feedback GeoGebra Tasks – Searching and Opensource and Collaborative Intelligent Interactive Tutor," *Proc. 26th World Multi-Conf. Systemics Cybern., Inform.: WMSCI 2022*, 2022, hal. 77–82, doi: 10.54808/WMSCI2022.03.77.
- [4] S. Majumdar dan N.D. Pan, "Combining Opensource GIS and Meta-Analysis to Link Rainfall Trend and Human Activity: Case Study on Gumti and Khawai Drainage Systems, Tripura, India," *Spat. Inf. Res.*, Vol. 28, No. 3, hal. 287–298, Jun. 2020, doi: 10.1007/s41324-019-00288-8.
- [5] H.D. Trung, N.T. Hung, dan N.H. Trung, "Opensource Based IoT Platform and LoRa Communications with Edge Device Calibration for Real-Time Monitoring Systems," dalam *Advanced Computational Methods for Knowledge Engineering. ICCSAMA 2019. Advances in Intelligent Systems and Computing*, H.A.L. Thi, H.M. Le, T.P. Dinh, and N. Nguyen, Eds., Cham, Swis: Springer, 2019, hal. 412–423, doi: 10.1007/978-3-030-38364-0_37.
- [6] M. Jebbar, A. Maizate, dan R.A. Abdelouahid, "Moroccan's Arabic Speech Training and Deploying Machine Learning Models with Teachable Machine," *Procedia Comput. Sci.*, Vol. 203, hal. 801–806, 2022, doi: 10.1016/j.procs.2022.07.120.
- [7] P.Y. Prasad dkk., "Implementation of Machine Learning Based Google Teachable Machine in Early Childhood Education," *Int. J. Early Child. Special Educ.*, Vol. 14, No. 3, hal. 4132–4138, Mei 2022, doi: 10.9756/INT-JECSE/V14I3.527.
- [8] D. Agustian, P.P.G.P. Pertama, P.N. Crisnapati, dan P.D. Novayanti, "Implementation of Machine Learning Using Google's Teachable Machine Based on Android," *2021 3rd Int. Conf. Cybern., Intell. Syst. (ICORIS)*, 2021, hal. 1–7, doi: 10.1109/ICORIS52787.2021.9649528.
- [9] Pengelola Web Kemdikbud (2022) "Pelestarian Bahasa Daerah Menjaga Warisan Bangsa," [Online], <https://www.kemdikbud.go.id/main/blog/2022/02/pelestarian-bahasa-daerah-menjaga-warisan-bangsa>, tanggal akses: 13 Mei 2023.
- [10] (2023) "Bahasa Sumba Timur," [Online], <https://labbineka.kemdikbud.go.id/bahasa/databahasa/3644a684f98ea8fe223c713b77189a77>, tanggal akses: 13-Mei-2023.
- [11] R.M.I. Malo, "Preliminary Study of the Dialects of Kambera," *KULTURISTIK: J. Bhs., Budaya*, Vol. 5, No. 2, hal. 1–6, Jul. 2021, doi: 10.22225/kulturistik.5.2.3654.
- [12] N.W. Kasni, "Strategy to Combine Clauses in Wajjewa Dialect a Sumbanese Language," *e-J. Linguist.*, Vol. 6, No. 2, hal. 93–107, Jul. 2012.
- [13] I.G. Budasi, "Bukti-Bukti Leksikal Pembeda Bahasa Wanokaka dan Anakalang di Sumba NTT," *Mabasan*, Vol. 4, No. 1, hal. 24–42, Jan.–Jun. 2010, doi: 10.26499/mab.v4i1.184.
- [14] I.W. Simpen, A.M. Mbete, I.M. Suastra, dan I.W. Pastika, "Kesantunan Berbahasa pada Penutur Bahasa Kambera di Sumba Timur," *e-J. Linguist.*, Vol. 2, No. 1, hal. 1–15, Mei 2008.
- [15] J.S. Lansing dkk., "Coevolution of languages and Genes on the Island of Sumba, Eastern Indonesia," *Proc. Nat. Acad. Sci. U. S. A.*, Vol. 104, No. 41, hal. 16022–16026, Okt. 2007, doi: 10.1073/pnas.0704451104.
- [16] I. Iswanto, V.J. Arnold, J. Kabnani, dan T. Salau, "Kajian Antropolingistik Bentuk Lingual Umbu dalam Nyanyian Tidur 'Ille Le' pada Masyarakat Melolo, Kabupaten Sumba Timur, Nusa Tenggara Timur (Antropolingistik Study Word 'Umbu' in Sleeping Song 'Ille Le' at the Melolo Community, East Sumba, East Nusa Tenggara)," *Jalabahasa*, Vol. 17, No. 2, hal. 179–191, Nov. 2021, doi: 10.36567/jalabahasa.v17i2.768.
- [17] M. Hoijsink dan A. Planqué-van Hardevel, "Machine Learning and the Platformization of the Military: A Study of Google's Machine Learning Platform TensorFlow," *Int. Political Sociol.*, Vol. 16, No. 2, hal. 1–19, Jun. 2022, doi: 10.1093/ips/olab036.
- [18] Y. Xie, M. He, T. Ma, dan W. Tian, "Optimal Distributed Parallel Algorithms for Deep Learning Framework TensorFlow," *Appl. Intell.*, Vol. 52, No. 4, hal. 3880–3900, Mar. 2022, doi: 10.1007/s10489-021-02588-9.
- [19] T. He dkk., "Bag of Tricks for Image Classification with Convolutional Neural Networks," *2019 IEEE/CVF Conf. Comput. Vis., Pattern Recognit. (CVPR)*, 2019, hal. 558–567, doi: 10.1109/CVPR.2019.00065.
- [20] G. Jiang, H. He, J. Yan, dan P. Xie, "Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox," *IEEE Trans. Ind. Electron.*, Vol. 66, No. 4, hal. 3196–3207, Apr. 2019, doi: 10.1109/TIE.2018.2844805.
- [21] (2022) "What is TensorFlow.js?," [Online], <https://codelabs.developers.google.com/TensorFlowjs-transfer-learning-teachable-machine#1>, tanggal akses: 3-Okt-2022.
- [22] Z. Tariq, S.K. Shah, dan Y. Lee, "Feature-Based Fusion Using CNN for Lung and Heart Sound Classification," *Sens.*, Vol. 22, No. 4, hal. 1–28, Feb. 2022, doi: 10.3390/s22041521.
- [23] U.A. Vishniakou dan B.H. Shaya, "Voice Detection Using Convolutional Neural Network," *Doklady BGUIR*, Vol. 21, No. 2, hal. 114–120, Apr. 2023, doi: 10.35596/1729-7648-2023-21-2-114-120.
- [24] A.Y. Kim dkk., "Automatic Depression Detection Using Smartphone-Based Text-Dependent Speech Signals: Deep Convolutional Neural Network Approach," *J. Med. Internet Res.*, Vol. 25, hal. 1–17, Jan. 2023, doi: 10.2196/34474.
- [25] P.R. Prakash dkk., "A Novel Convolutional Neural Network with Gated Recurrent Unit for Automated Speech Emotion Recognition and Classification," *J. Control, Decis.*, Vol. 10, No. 1, hal. 54–63, Jan. 2023, doi: 10.1080/23307706.2022.2085198.
- [26] Y. Yu, C. Peng, Q. Tang, dan X. Wang, "Monaural Music Source Separation Using Deep Convolutional Neural Network Embedded with Feature Extraction Module," *2022 Asia Conf. Algorithms, Comput., Mach. Learn. (CACML)*, 2022, hal. 546–551, doi: 10.1109/CACML55074.2022.00098.
- [27] P.H. Chandankhede, A.S. Titarmare, dan S. Chauhan, "Voice Recognition Based Security System Using Convolutional Neural Network," *2021 Int. Conf. Comput. Commun., Intell. Syst. (ICCCIS)*, 2021, hal. 738–743, doi: 10.1109/ICCCIS51004.2021.9397151.
- [28] N. Li dkk., "Robust Voice Activity Detection Using a Masked Auditory Encoder Based Convolutional Neural Network," *ICASSP 2021 - 2021 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2021, hal. 6828–6832, doi: 10.1109/ICASSP39728.2021.9415045.
- [29] M. Soundarya, P.R. Karthikeyan, dan G. Thangarasu, "Automatic Speech Recognition Trained with Convolutional Neural Network and Predicted with Recurrent Neural Network," *2023 9th Int. Conf. Elect. Energy Syst. (ICEES)*, hal. 41–45, doi: 10.1109/ICEES57979.2023.10110224.
- [30] M.M. Kamruzzaman, "Arabic Sign Language Recognition and Generating Arabic Speech Using Convolutional Neural Network," *Wirel. Commun., Mob. Comput.*, Vol. 2020, hal. 1–9, Mei 2020, doi: 10.1155/2020/3685614.
- [31] M. Hireš dkk., "Convolutional Neural Network Ensemble for Parkinson's Disease Detection from Voice Recordings," *Comput. Biol., Med.*, Vol. 141, hal. 1–9, Feb. 2022, doi: 10.1016/j.combiomed.2021.105021.
- [32] K.K. Lella dan A. Pja, "Automatic Diagnosis of COVID-19 Disease Using Deep Convolutional Neural Network with Multi-Feature Channel from Respiratory Sound Data: Cough, Voice, and Breath," *Alex. Eng. J.*, Vol. 61, No. 2, hal. 1319–1334, Feb. 2022, doi: 10.1016/j.aej.2021.06.024.
- [33] I. Kwon dkk., "Diagnosis of Early Glottic Cancer Using Laryngeal Image and Voice Based on Ensemble Learning of Convolutional Neural Network Classifiers," *J. Voice*, akan dipublikasikan.
- [34] M. Hireš, M. Gazda, L. Vavrek, dan P. Drotár, "Voice-Specific Augmentations for Parkinson's Disease Detection Using Deep Convolutional Neural Network," *2022 IEEE 20th Jubil. World Symp. Appl. Mach. Intell., Inform. (SAMI)*, 2022, hal. 000213–000218, doi: 10.1109/SAMI54271.2022.9780856.
- [35] T.T. Leonid dan R. Jayaparvathy, "Classification of Elephant Sounds Using Parallel Convolutional Neural Network," *Intell. Automat., Soft Comput.*, Vol. 32, No. 3, hal. 1415–1426, Jun. 2022, doi: 10.32604/iasc.2022.021939.
- [36] S.D.H. Permana dkk., "Classification of Bird Sounds as an Early Warning Method of Forest Fires Using Convolutional Neural Network (CNN) Algorithm," *J. King Saud Univ. – Comput., Inf. Sci.*, Vol. 34, No. 7, hal. 4345–4357, Jul. 2022, doi: 10.1016/j.jksuci.2021.04.013.
- [37] W.S.E. Putra, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101," *J. Tek. ITS*, Vol. 5, No. 1, hal. A65–A69, Mar. 2016, doi: 10.12962/j23373539.v5i1.15696.
- [38] M. Carney dkk., "Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification," *Ext. Abstr. 2020 CHI Conf. Hum. Factors Comput. Syst.*, 2020, hal. 1–8, doi: 10.1145/3334480.3382839.