

# Klasifikasi Kendaraan Berbasis Suara Di Lalu Lintas: Implementasi TinyML

Hendra Kusumah<sup>1</sup>, Nissa Nabbillah<sup>2</sup>, Syifa Audina<sup>3</sup>

<sup>1</sup>Program Studi Sistem Komputer Universitas Raharja

<sup>2,3</sup>Program Studi Sistem Informasi Universitas Raharja

E-mail: \*<sup>1</sup>hendra.kusumah@raharja.info, <sup>2</sup>nissa.nabbillah@raharja.info,  
<sup>3</sup>syifa.audina@raharja.info

## Abstrak

Mengemudi sudah mendarah daging dalam budaya Indonesia. Jika memenuhi persyaratan, seluruh warga negara Indonesia (WNI) berhak mendapatkan Surat Izin Mengemudi (SIM). Oleh karena itu, jumlah pengendara cukup tinggi di Indonesia sehingga menyebabkan tingginya tingkat kecelakaan lalu lintas. Menurunnya kemampuan pendengaran seseorang menjadi salah satu faktor pendukung tingginya kecelakaan lalu lintas. Penelitian ini berfokus pada pengembangan alat yang mampu mengklasifikasikan kendaraan melalui audio kendaraan saat sedang berjalan sehingga dapat membantu pengendara yang memiliki permasalahan pendengaran. Penelitian ini memanfaatkan *Arduino Nano 33 BLE Sense* yang memiliki mikrofon tertanam dan menggunakan *platform* terbuka *Edge Impulse*. Model yang digunakan berupa model jaringan syaraf buatan sederhana dengan 4 layer konvolusi 1-D dan 1 *hidden* layer. Dataset yang digunakan diambil secara manual melalui *platform* terbuka *Youtube* yang direkam langsung pada *platform Edge Impulse* melalui mikrofon pada *Arduino Nano*. Proses training model dijalankan selama 80 *epochs*. Model dirasa mendapatkan akurasi cukup baik pada saat training dan testing, yaitu dengan nilai akurasi masing-masing 96,3% dan 90,9%. Dengan hasil yang cukup ini, model dirasa memumpuni untuk diimplementasikan walaupun masih memerlukan pengembangan lebih lanjut.

**Kata Kunci**—Kecelakaan Lalu Lintas, Pendengaran, *Arduino*, Audio, *Edge Impulse*

## Abstract

*Driving is deeply embedded in Indonesian culture. All Indonesian citizens (WNI) who meet the qualifications can obtain a driver's license (SIM). As a result, the number of drivers in Indonesia is relatively significant, resulting in a high rate of traffic accidents. One of the reasons leading to the rising number of road accidents is hearing loss. This study focuses on building a tool that can classify vehicles based on the audio of the vehicle while it is running in order to assist drivers with hearing impairments. This study employs the Arduino Nano 33 BLE Sense, which includes an embedded microphone, and the Edge Impulse open platform. A simple artificial neural network model with four 1-D convolution layers and one hidden layer was utilized. The dataset was obtained manually using the YouTube open platform and was captured directly on the Edge Impulse platform using the Arduino Nano's microphone. The model training procedure was repeated 80 epochs. During training and testing, the model is regarded to be accurate, with accuracy scores of 96.3% and 90.9%, respectively. With these adequate results, the model is judged eligible for implementation, despite the fact that it still requires further refinement.*

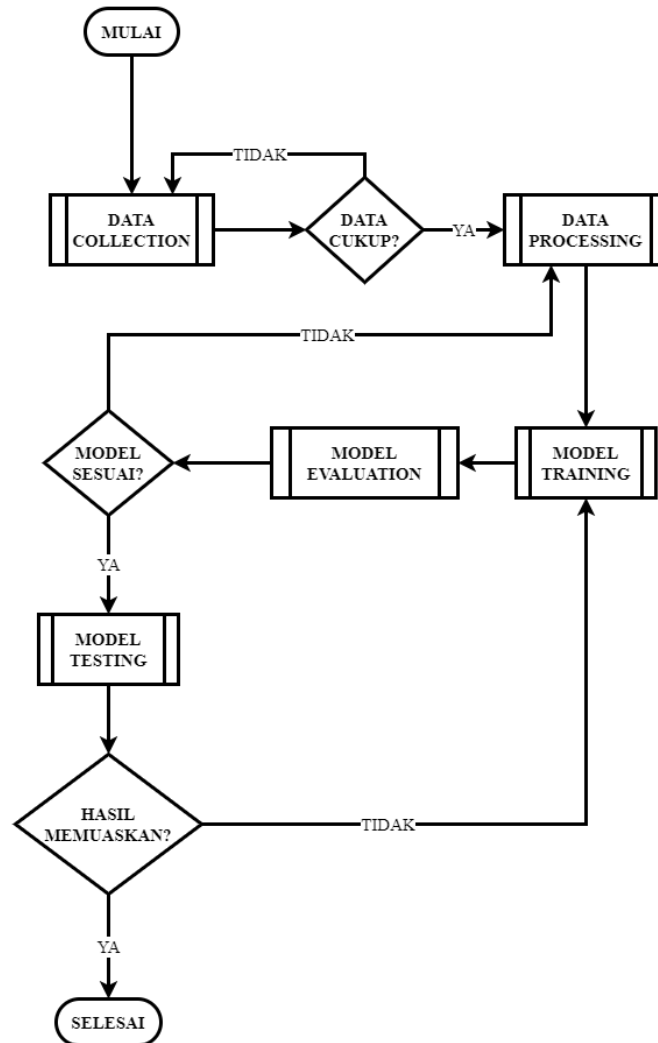
**Keywords**—Traffic accidents, Hearing, *Arduino*, Audio, *Edge Impulse*

## 1. PENDAHULUAN

Mengemudi sudah mendarah daging dalam budaya Indonesia. Mulai dari mengendarai kendaraan roda dua hingga kendaraan roda banyak seperti bus dan truk. Jika memenuhi syarat, seluruh warga negara Indonesia (WNI) berhak mendapatkan Surat Izin Mengemudi (SIM) [1]. Meningkatnya jumlah pengemudi juga meningkatkan jumlah kecelakaan lalu lintas. Menurut data Badan Pusat Statistik (BPS), kecelakaan lalu lintas di Indonesia pada tahun 2021 tumbuh sebanyak 116.411 kejadian atau 3,62% dari 100.028 kejadian pada tahun 2020 [2]. Tingginya angka kecelakaan tersebut disebabkan oleh beberapa variabel [3]–[5] antara lain Budaya lalu lintas Indonesia, kesadaran pengemudi terhadap peraturan perundang-undangan yang berlaku, kondisi sarana dan prasarana transportasi, dan cuaca ekstrim. Selain itu, kompetensi pengemudi [6] merupakan elemen penting dalam prevalensi kecelakaan lalu lintas yang tinggi di Indonesia. Kemampuan mendengar [7] merupakan salah satu kompetensi utama pengemudi. Besar kemungkinan gangguan pendengaran berperan penting dalam mengemudi. Hal ini disebabkan kemampuan pendengaran pengemudi yang terbatas, yang dapat mengurangi perhatian terhadap lingkungan sekitar saat berkendara di lalu lintas. Pengemudi dengan gangguan pendengaran, di sisi lain, tetap diizinkan untuk mengemudi jika mereka memenuhi persyaratan yang diperlukan. Perizinan bagi pengendara tuna rungu diatur dalam UU RI No. 22 Tahun 2009 tentang lalu lintas, khususnya pasal 80 (e) [8], dan Peraturan Kapolri No. 5 Tahun 2021 yang mengatur tentang pemberian dan penandaan surat izin mengemudi, khususnya pasal 1 angka 15 [9], [10]. Sebagai akibatnya, beberapa orang percaya bahwa alat yang mampu memberikan pemberitahuan awal tentang kendara terdekat di sekitar pengendara sangat penting. Para peneliti telah berupaya mengembangkan studi di bidang pemrosesan audio [11]–[14] dalam beberapa tahun terakhir. Fitur-fitur dalam audio diekstraksi dan digunakan sesuai kebutuhan. Temuan penelitian ini dianggap dapat diterapkan untuk membantu manusia dengan tugas sehari-hari biasa, seperti asisten rumah dan asisten *Google* dengan aktivasi suara [15], [16]. Dengan terobosan baru ini, para peneliti berusaha membangun sejumlah alat yang dapat membantu orang-orang dengan kebutuhan khusus dalam melakukan tugas-tugas tipikal. Selain itu, Alat ini juga dimaksudkan untuk mengurangi angka kecelakaan lalu lintas yang diakibatkan oleh berkurangnya kemampuan pendengaran seseorang.

## 2. METODE PENELITIAN

Metode penelitian yang digunakan pada penelitian ini terdiri atas 5 (lima) tahapan, yaitu diantaranya:



Gambar 1. Alur Penelitian

### A. Data Collection

Tujuan dari *Data Collection* ini yaitu untuk mengumpulkan data-data yang relevan sesuai dengan tujuan penelitian. Data yang digunakan dalam penelitian ini berupa data audio. Data-data yang dikumpulkan berasal dari beberapa sumber terbuka seperti melalui *platform Youtube* [17]. Setelah dilakukan pencarian terhadap data-data yang relevan, kemudian data tersebut dikumpulkan dan direkam dengan menggunakan alat yang diajukan dengan kriteria durasi data audio sepanjang 1 hingga 5 detik. Data yang digunakan terdiri dari 4 kategori audio, yaitu audio motor, audio truk, audio bus, dan audio *noise* di jalan Indonesia. Perekaman terhadap data-data ini dilakukan secara manual langsung pada *platform Edge Impluse* [18] yang terkoneksi dengan mikrokontroler yang digunakan.

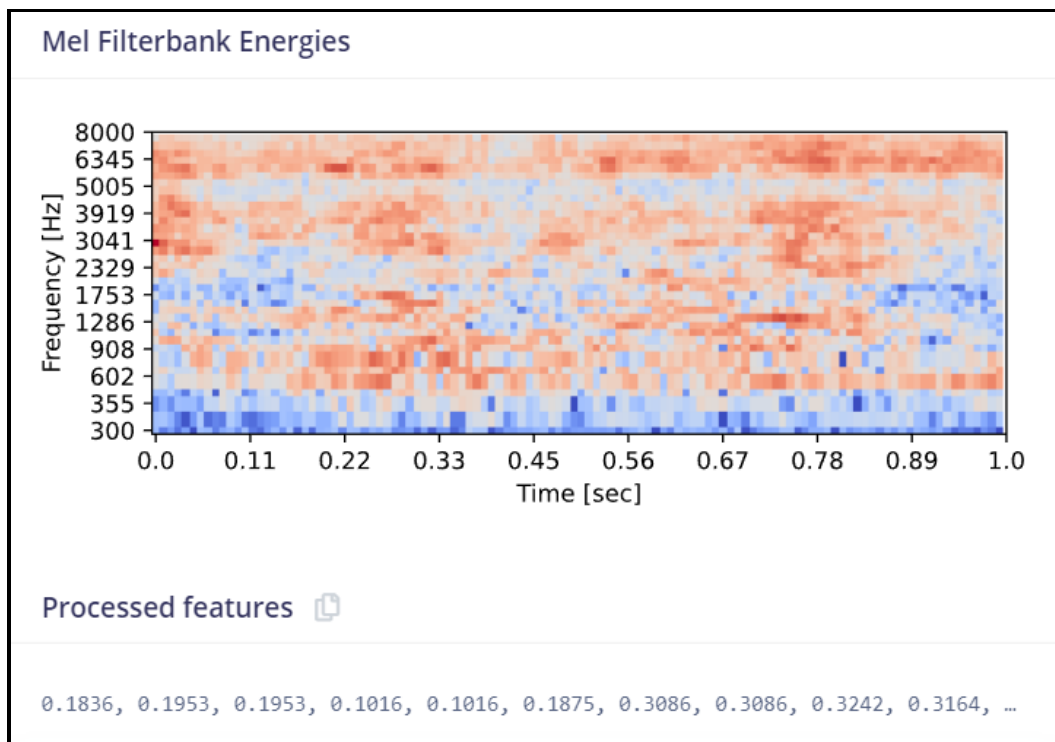
### B. Data Processing

*Data processing* dilakukan untuk mengambil fitur-fitur data pada data audio yang dikumpulkan sehingga dapat diproses untuk klasifikasi. *Data processing* (pemrosesan data)

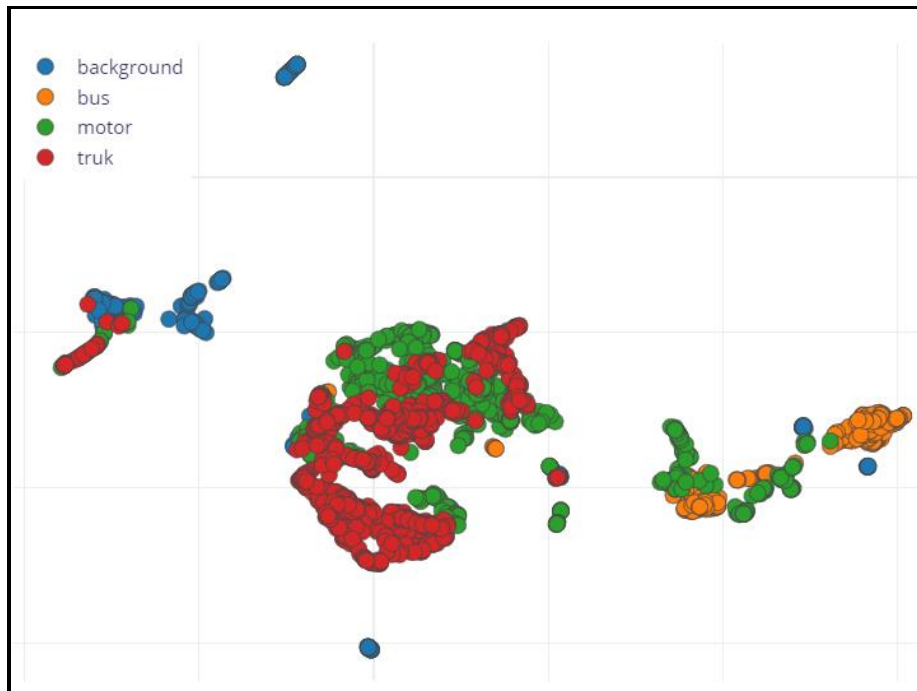
dilakukan melalui 2 tahapan, yaitu MFE (*Mel-Filterbank Energy*) [19], [20] dan *Spectrogram* [21]:

C. *MFE (Mel-Filterbank Energy)*

Sudah diketahui umum bahwa sistem pengenalan ucapan otomatis (*Automatic Speech Recognition*) [21]–[23] sangat bergantung pada *Mel-Filterbank Energy*. Biasanya, *Sub-band Spectral* timbul dari spektrum dengan rentang frekuensi yang lebih sempit [24]. Dalam domain frekuensi, ini menggunakan skala non-linier yang dikenal sebagai skala Mel [25]. Skala Mel unggul dalam memproses data audio, terutama ketika suara yang akan dikategorikan mudah untuk dibedakan oleh telinga manusia. Ini paling umum di luar aplikasi pengenalan suara. Dengan cara yang sama setelah spektrogram dihitung, filter segitiga skala Mel digunakan untuk mengisolasi rentang frekuensi individual. Jumlah filter, Frekuensi rendah, dan Frekuensi tinggi adalah parameter yang dapat dikonfigurasi yang menentukan rentang frekuensi dan kedalaman ekstraksi fitur [26]. Nada skala mel adalah nada yang disetujui secara subyektif oleh pendengar dengan jarak yang sama. Ini bekerja secara efektif dengan suara yang dapat dibedakan oleh telinga manusia karena lebih banyak fitur (lebih banyak bank filter) yang diekstraksi pada frekuensi yang lebih rendah.



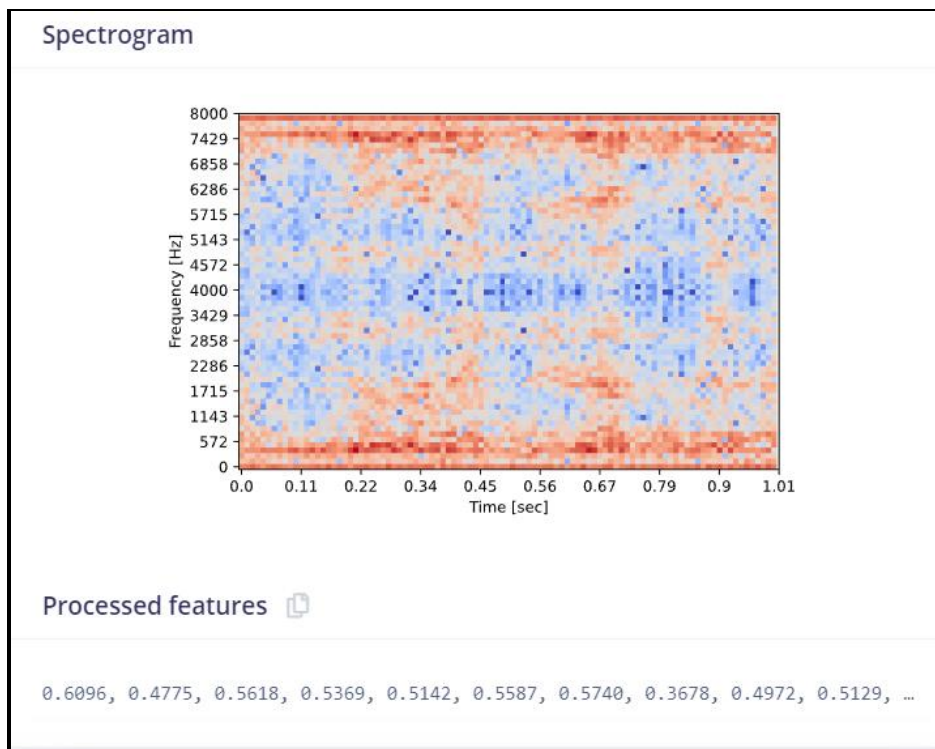
Gambar 2. Parameter fitur MFE yang digunakan



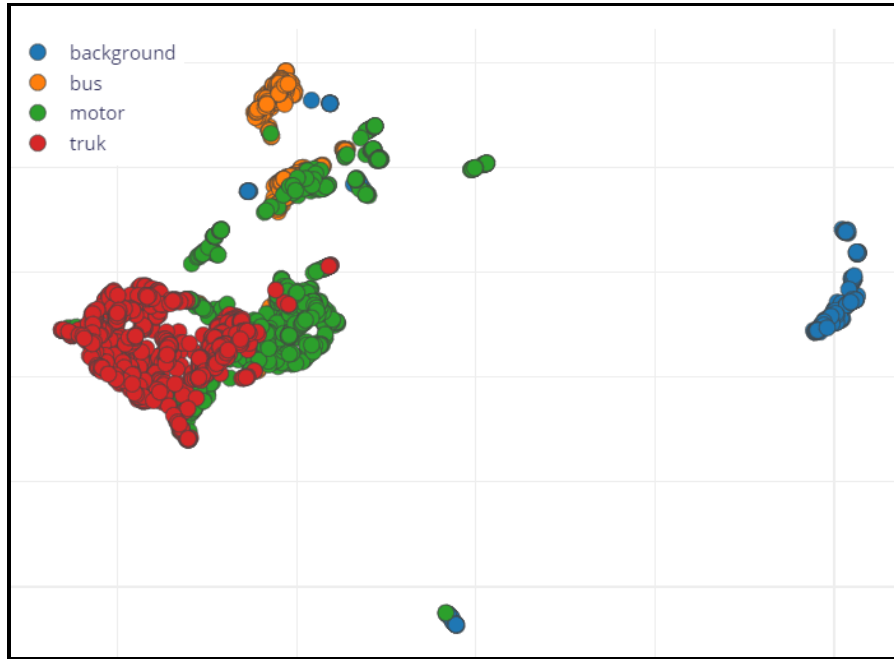
Gambar . Hasil Mel-Filterbank Energy

#### D. Spectograms

Setelah dilakukan pemrosesan MFE, selanjutnya data diesktrak untuk diambil data Spektrogramnya. *Spektrogram* [27] memplot kekuatan sinyal dari waktu ke waktu pada setiap frekuensi bentuk gelombang. *Spektrogram* dapat berupa grafik dua atau tiga dimensi, dengan dimensi ekstra yang diwakili oleh variabel warna keempat.



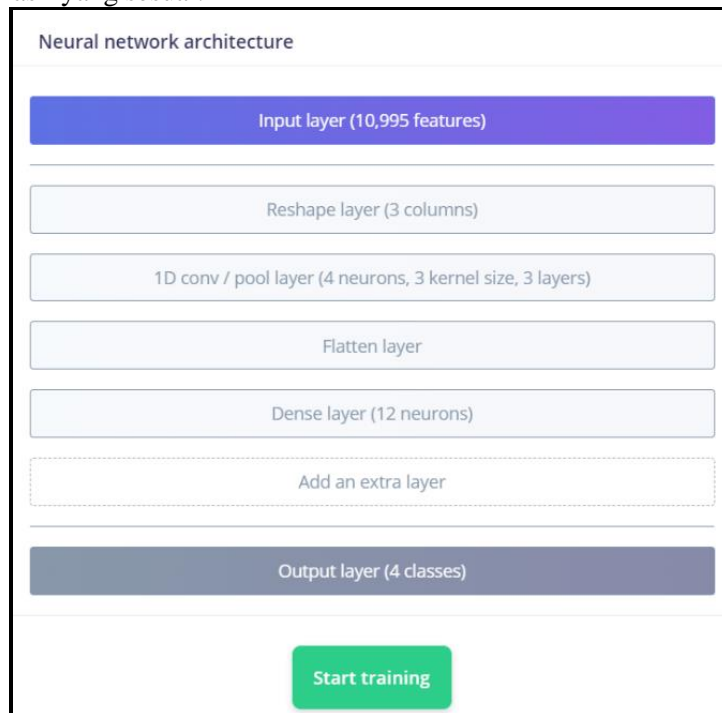
Gambar 4. Parameter fitur spektrogram yang digunakan



Gambar 5. Hasil pemrosesan spektrogram

#### E. Training Process

Penelitian ini memanfaatkan *platform* terbuka *Edge Impulse* yang bertujuan untuk mempermudah proses pelatihan dan pengevaluasian model serta pengetesan model yang telah dibuat. Pada *Edge Impulse*, model dirancang sedemikian rupa sehingga mampu didapatkan hasil yang sesuai.



Gambar 6. Arsitektur model yang diajukan

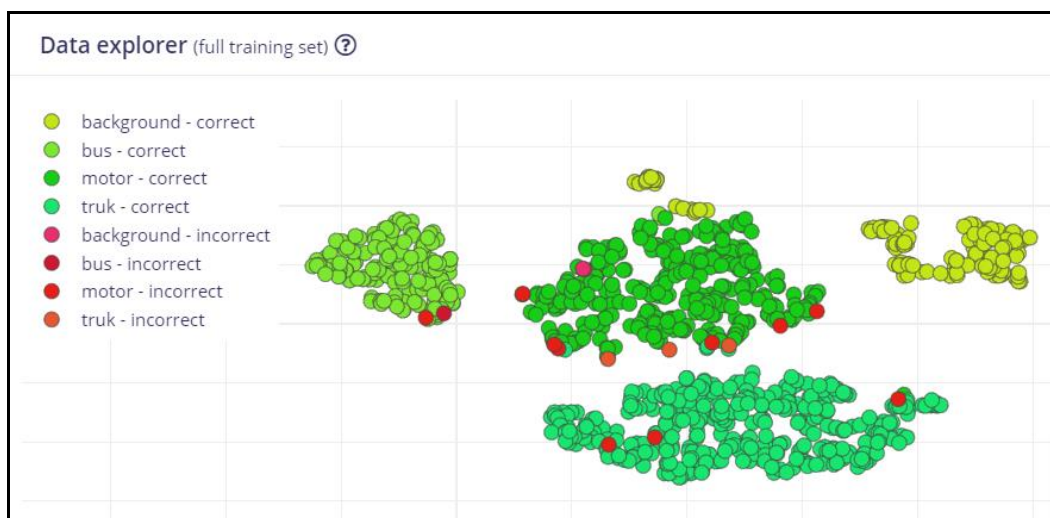
Arsitektur model yang digunakan terdiri atas 4 neuron layer konvolusi 1-D dengan 3 kernel dan fungsi aktivasi ReLU, serta 1 *hidden layer*. Selain itu, layer *Reshape* juga digunakan untuk menyesuaikan dimensi data dengan dimensi layer konvolusi yang

digunakan, serta penggunaan layer *MaxPooling-1D* juga digunakan untuk membantu model dalam memilah nilai yang relevan untuk digunakan dalam pelatihan. Penggunaan layer yang minim ini bertujuan untuk menyesuaikan kebutuhan model dengan kapasitas alat yang digunakan.

### 3. HASIL DAN PEMBAHASAN

Pada penelitian ini, peneliti menggunakan mikrokontroler Arduino Nano 33 BLE Sense dengan 1 MB memory *flash* CPU. Arduino Nano 33 BLE *Sense* ini memiliki beberapa sensor tertanam, terutama *mikrofon*. Proses pengumpulan data dilakukan dengan perekaman audio menggunakan *mikrofon* tertanam pada Arduino ini dilakukan dengan mengkoneksikan Arduino dengan *platform Edge Impulse*. Hasil pengumpulan data berupa data training dengan durasi 24 menit dan data testing dengan durasi 6 menit dengan masing-masing data audio berdurasi 1-5 detik. Selanjutnya data diproses sesuai dengan metode pemrosesan yang diajukan, yaitu melalui pemrosesan MFE (*Mel-Filterbank Energy*) dan *Spectograms*.

Setelah pemrosesan data dilakukan, tahap selanjutnya yaitu persiapan model dengan menentukan arsitektur yang diajukan. Arsitektur model yang dibuat menggunakan arsitektur seminimum mungkin untuk dapat menyesuaikan dengan kapasitas memori dari Arduino Nano yang digunakan. Model dilatih dengan konfigurasi learning rate sebesar 0,0005 dengan *optimizer* Adam dan total *epoch* sebesar 80 epochs. Dari hasil pelatihan ini didapatkan akurasi sebesar 96,3% dengan *loss* sebesar 11%.

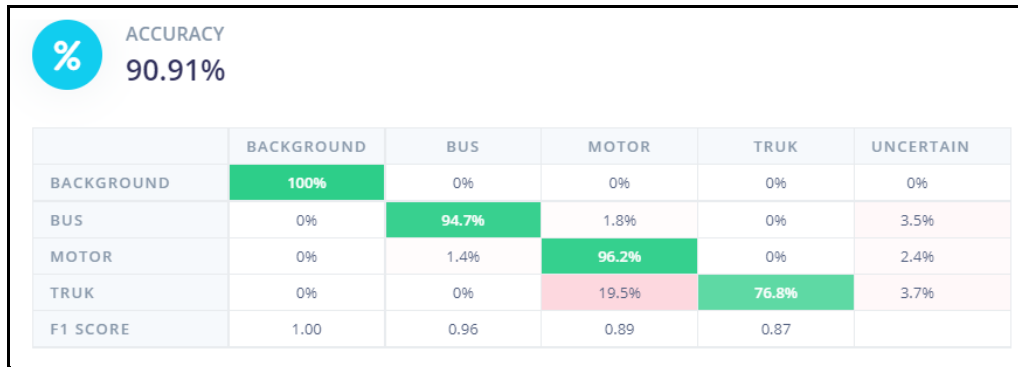


Gambar 7. Hasil training model

	BACKGROUND	BUS	MOTOR	TRUK
BACKGROUND	97.0%	0%	3.0%	0%
BUS	0%	98.8%	1.3%	0%
MOTOR	0%	1.7%	93.9%	4.4%
TRUK	0%	0%	2.7%	97.3%
F1 SCORE	0.98	0.98	0.95	0.97

Gambar 8. Confusion matrix model saat training

Selain akurasi yang tinggi, nilai *F1-Score* dari masing-masing kategori juga cukup tinggi dengan nilai minimum 95% dilihat dari tabel *confusion matrix* di atas. Di sisi lain, pada *platform Edge Impulse* juga terdapat perkiraan total memori *flash* yang mungkin akan digunakan serta estimasi penggunaan RAM yang mungkin digunakan saat di implementasikan pada Arduino Nano. Berdasarkan perhitungan perkiraan tersebut, model yang dibuat diestimasi menggunakan *flash memory* sebesar 50KB dengan penggunaan RAM sebesar 33,6KB dari total 256KB RAM atau hanya sekitar 13% dari RAM yang dimiliki.



Gambar 9. *Confusion mamatrix* model saat testing

Selanjutnya dilakukan testing terhadap model untuk mengetahui seberapa bagus kemampuan model jika diterapkan ke dataset baru. Setelah dilakukan proses testing, dilihat model mendapatkan nilai sebesar 90,91%. Nilai ini dirasa sangat baik, walaupun dari hasil *confusion matrix* pada tabel di atas, dilihat tingkat *confusing* motor dengan truk mencapai 19,5%, namun model masih tetap mampu membedakan antar keduanya.

#### 4. KESIMPULAN

Pengembangan pemrosesan audio memang sudah sangat memumpuni untuk diterapkan pada mikrokontroler. Namun perlu kajian lebih lanjut terhadap implementasinya agar benar-benar mampu membantu manusia. Penelitian ini mampu membuktikan kemampuan klasifikasi model sederhana untuk membedakan suara kendaraan dengan hasil akurasi pelatihan dan pengetasan yang sangat baik, yaitu 96,3% dan 90,9%. Selain itu, nilai *F1-Score* yang didapatkan model yang diusulkan juga cukup baik dengan nilai di atas 95%. Di sisi lain, karena masih dalam tahap pengembangan, dataset yang digunakan masih sangat sedikit dan masih dikumpulkan secara manual. Diharapkan ke depannya dataset yang digunakan dapat lebih memumpuni sehingga model yang diusulkan dapat lebih kredibel dan reliabel.

#### DAFTAR PUSTAKA

- [1] I. D. Akbar, "PROSEDUR MEMPEROLEH SURAT IJIN MENGEMUDI (SIM) (STUDI POLRESTABES SURABAYA)," *DEKRIT*, vol. 12, no. 1, pp. 100–116, Nov. 2022, Accessed: Dec. 15, 2022. [Online]. Available: <https://jurnal.fhubhara.com/index.php/dekrit/article/view/111>
- [2] "Badan Pusat Statistik." <https://www.bps.go.id/indicator/17/513/1/jumlah-kecelakaan-korban-mati-luka-berat-luka-ringan-dan-kerugian-materi.html> (accessed Dec. 15, 2022).
- [3] Y. WA, "ANALISIS PENGARUH HUMAN ERROR, KONDISI LINGKUNGAN DAN KONDISI KENDARAAN TERHADAP KECELAKAAN LALU LINTAS PADA PENGENDARA SEPEDA MOTOR DI JALAN GATOT SUBROTO KABUPATEN CILACAP," 2021.

- [4] “Beberapa Faktor Penyebab Kecelakaan di Indonesia | Jurnal Ilmu Teknik.” <http://ilmuteknik.org/index.php/ilmuteknik/article/view/75> (accessed Dec. 15, 2022).
- [5] I. Zainafree *et al.*, “EPIDEMIOLOGI KECELAKAAN LALU LINTAS: TANTANGAN DAN SOLUSI,” *Kesehatan Masyarakat*, no. 1, pp. 92–127, May 2022, doi: 10.15294/KM.V11I1.70.
- [6] W. D. Prastiwi, “MEMBEDAH PASAL 77 AYAT (1) UNDANG-UNDANG REPUBLIK INDONESIA NOMOR 22 TAHUN 2009 TENTANG LALU LINTAS DAN ANGKUTAN JALAN DI MASA PANDEMI,” *JISOS: JURNAL ILMU SOSIAL*, vol. 1, no. 7, pp. 639–646, Aug. 25, 2022. Accessed: Dec. 15, 2022. [Online]. Available: <https://www.bajangjournal.com/index.php/JISOS/article/view/3288>
- [7] “Beberapa Faktor Penyebab Kecelakaan di Indonesia | Jurnal Ilmu Teknik.” <http://ilmuteknik.org/index.php/ilmuteknik/article/view/75> (accessed Dec. 15, 2022).
- [8] S. T. Hizbillah, “Tinjauan Yuridis Pasal 35 Ayat 1 Huruf C Peraturan Kapolri Nomor 9 Tahun 2012 tentang Surat Izin Mengemudi Bagi Disabilitas Daksa (Juridical Review Article 35 Paragraph 1 Letter C Regulation of the National Police Chief Number 9 of 2012 concerning Driving Permits for Persons with Disabilities),” *Jurnal HAM*, vol. 11, no. 1, p. 131, Jul. 2022, doi: 10.30641/HAM.2020.11.131-150.
- [9] “Komnas Disabilitas Berhasil Mengadvokasi Pembuatan SIM bagi Tuli - KamiBijak.” <https://www.kamibijak.com/v/komnas-disabilitas-berhasil-mengadvokasi-pembuatan-sim-bagi-tuli> (accessed Dec. 15, 2022).
- [10] “Asyik, Tuna Rungu di Boyolali Kini Bisa Ajukan SIM A dan C Khusus Tuli | MURIANEWS.” <https://www.murianews.com/2022/09/24/319710/asyik-tuna-rungu-di-boyolali-kini-bisa-ajukan-sim-a-dan-c-khusus-tuli> (accessed Dec. 15, 2022).
- [11] L. Schoneveld, A. Othmani, and H. Abdelkawy, “Leveraging recent advances in deep learning for audio-Visual emotion recognition,” *Pattern Recognit Lett*, vol. 146, pp. 1–7, Jun. 2021, doi: 10.1016/J.PATREC.2021.03.007.
- [12] P. Zinemanas, M. Rocamora, M. Miron, F. Font, and X. Serra, “An Interpretable Deep Learning Model for Automatic Sound Classification,” *Electronics 2021, Vol. 10, Page 850*, vol. 10, no. 7, p. 850, Apr. 2021, doi: 10.3390/ELECTRONICS10070850.
- [13] S. Amiriparian, T. Hübner, V. Karas, M. Gerczuk, S. Ottl, and B. W. Schuller, “DeepSpectrumLite: A Power-Efficient Transfer Learning Framework for Embedded Speech and Audio Processing From Decentralized Data,” *Front Artif Intell*, vol. 5, p. 856232, Mar. 2022, doi: 10.3389/FRAI.2022.856232.
- [14] S. Latif, H. Cuayáhuitl, F. Pervez, F. Shamshad, H. S. Ali, and E. Cambria, “A survey on deep reinforcement learning for audio-based applications,” *Artif Intell Rev*, pp. 1–48, Jul. 2022, doi: 10.1007/S10462-022-10224-2/TABLES/5.
- [15] Z. Chen, D. Byrne, and D. El-Zanfaly, “Google Home, Listen: Building Helper Intelligences for Non-Verbal Sound,” *ACM International Conference Proceeding Series*, pp. 619–622, Jun. 2022, doi: 10.1145/3527927.3535202.
- [16] N. AlOtaibi and F. Lombardi, “Privacy and security evaluation of Amazon Echo voice assistant,” *2021 International Conference of Women in Data Science at Taif University, WiDSTaif 2021*, Mar. 2021, doi: 10.1109/WIDSTAIIF52235.2021.9430217.
- [17] “YouTube.” <https://www.youtube.com/> (accessed Dec. 15, 2022).
- [18] “Edge Impulse.” <https://www.edgeimpulse.com/> (accessed Dec. 15, 2022).
- [19] J. G. W. T. Römer, “Real-time pitch detection using resource constrained IoT device,” 2022.
- [20] M. Altayeb, M. Zennaro, and M. Rovai, “Classifying mosquito wingbeat sound using TinyML,” *ACM International Conference Proceeding Series*, pp. 132–137, Sep. 2022, doi: 10.1145/3524458.3547258.
- [21] Y. Gong, Y. A. Chung, and J. Glass, “AST: Audio Spectrogram Transformer,” *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, vol. 1, pp. 56–60, Apr. 2021, doi: 10.48550/arxiv.2104.01778.

- [22] V. Bhardwaj *et al.*, “Automatic Speech Recognition (ASR) Systems for Children: A Systematic Literature Review,” *Applied Sciences* 2022, Vol. 12, Page 4419, vol. 12, no. 9, p. 4419, Apr. 2022, doi: 10.3390/APP12094419.
- [23] S. Alharbi *et al.*, “Automatic Speech Recognition: Systematic Literature Review,” *IEEE Access*, vol. 9, pp. 131858–131876, 2021, doi: 10.1109/ACCESS.2021.3112535.
- [24] X. Hao, X. Su, R. Horaud, and X. Li, “FullsubNet: A full-band and sub-band fusion model for real-time single-channel speech enhancement,” *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2021-June, pp. 6633–6637, 2021, doi: 10.1109/ICASSP39728.2021.9414177.
- [25] K. Rangra and M. Kapoor, “Exploring the mel scale features using supervised learning classifiers for emotion classification,” *International Journal of Applied Pattern Recognition*, vol. 6, no. 3, p. 232, 2021, doi: 10.1504/IJAPR.2021.117204.
- [26] “Audio MFE - Edge Impulse Documentation.” <https://docs.edgeimpulse.com/docs/edge-impulse-studio/processing-blocks/audio-mfe> (accessed Dec. 15, 2022).
- [27] “Spectrogram - Edge Impulse Documentation.” <https://docs.edgeimpulse.com/docs/edge-impulse-studio/processing-blocks/spectrogram> (accessed Dec. 15, 2022).