Klasifikasi Audio Deepfake Berbasis Transfer Learning YAMNet dan Regresi Logistik

Hakam Dzakwan Diash[1], Dwi Arman Prasetya[2], Alfan Rizaldy Pratama[3]
[1, 2, 3]*Program Studi Sains Data, Universitas Pembangunan Nasional Veteran Jawa Timur*
[1]22083010069@student.upnjatim.ac.id
[3]alfan.fasilkom@upnjatim.ac.id
Corresponding author email: arman.prasetya.sada@upnjatim.ac.id

ABSTRAK

Audio *deepfake* menjadi ancaman serius bagi keamanan publik karena dapat disalahgunakan untuk penipuan, manipulasi informasi, dan penyebaran disinformasi secara masif. Penelitian ini mengusulkan pendekatan klasifikasi yang mengintegrasikan YAMNet sebagai ekstraktor fitur berbasis *transfer learning* dengan model klasifikasi regresi logistik. YAMNet digunakan untuk mengekstraksi representasi fitur audio yang kaya dari sinyal suara, sedangkan regresi logistik berperan dalam memodelkan perbedaan antara sampel audio asli dan palsu. Dataset *In-the-Wild* digunakan dengan pembagian 70% data pelatihan (22.245 sampel) dan 30% data pengujian (9.534 sampel). Berdasarkan hasil pengujian, model regresi logistik mencapai akurasi 98.53%, presisi 0.9890, *recall* 0.9876, dan *F1-score* 0.9883 yang menunjukkan kinerja sangat baik dan seimbang dalam mengklasifikasikan kedua kelas tersebut. Selain itu, waktu inferensi yang sangat cepat, yaitu 0.0276 detik untuk satu kali prediksi terhadap seluruh data uji, menunjukkan efisiensi komputasi yang tinggi. Hasil ini menegaskan bahwa kombinasi YAMNet dan Regresi Logistik mampu menghasilkan sistem klasifikasi yang akurat, ringan, dan efisien, serta memiliki potensi tinggi untuk diterapkan dalam upaya mitigasi penyalahgunaan teknologi audio sintetis pada berbagai aspek keamanan digital.

Keywords: Audio *Deepfake*, YAMNet, Regresi Logistik

I. PENDAHULUAN

Teknologi informasi dan komunikasi (TIK) telah menjadi komponen penting dalam masyarakat modern, terutama saat interaksi tatap muka terbatas. Ketergantungan yang tinggi pada platform komunikasi virtual seperti konferensi video dan aplikasi pesan instan, yang sangat mengandalkan audio, menegaskan peran penting audio dalam penyebaran informasi di era digital [1]. Sejalan dengan meningkatnya peran penting media audio tersebut, teknologi Kecerdasan Buatan (AI) mengalami kemajuan pesat, khususnya dalam inovasi *Generative Artificial Intelligence* (Gen-AI). Teknologi ini menunjukkan kemampuan canggih dalam menciptakan konten multimedia baru, salah satunya sintesis suara [2]. Pemanfaatan kemampuan inilah yang kemudian memfasilitasi kemunculan fenomena audio *deepfake*, yaitu rekaman suara buatan yang dirancang untuk meniru suara individu tertentu dengan tingkat kemiripan yang tinggi.

Kemampuan Gen-AI untuk menghasilkan tiruan suara yang sangat realistis telah membuka berbagai celah ancaman yang serius. Dalam ranah sosial dan politik, *audio deepfake* dapat digunakan sebagai alat penyebar disinformasi dan berita palsu, dengan potensi merusak reputasi individu serta mengganggu stabilitas [3]. Selain itu, di sektor keamanan dan finansial, ancaman ini menjadi lebih nyata. Teknologi ini dapat dimanfaatkan untuk melakukan penipuan berbasis suara yang dikenal sebagai *voice phishing* untuk mengelabui korban, serta berpotensi membobol sistem keamanan yang

menggunakan autentikasi biometrik suara [4]. Meningkatnya insiden penyalahgunaan ini menunjukkan adanya kebutuhan mendesak untuk mengembangkan metode klasifikasi yang efektif dan akurat.

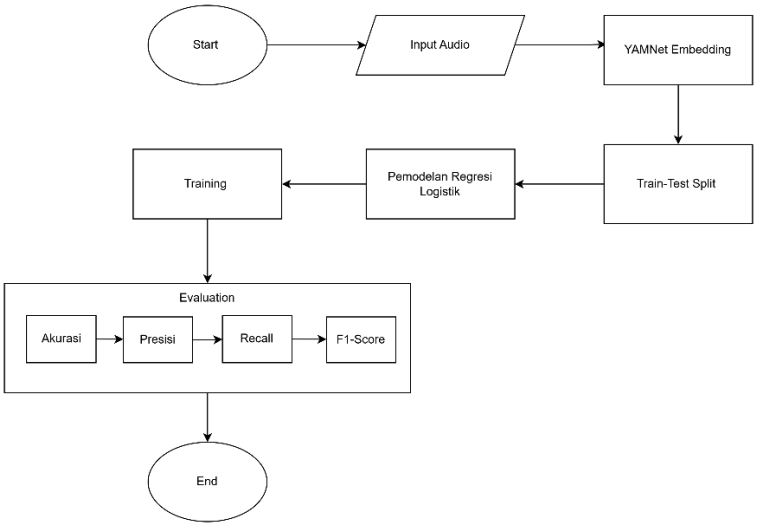
Model *deep learning* banyak digunakan untuk melakukan klasifikasi audio *deepfake*, namun metode tersebut memiliki kompleksitas komputasi yang tinggi. Sebagai alternatif, regresi logistik menawarkan pendekatan klasifikasi yang lebih ringan dan efisien secara komputasi, terutama pada dataset berdimensi tinggi [5]. Selain efisien, performa regresi logistik sangat bergantung pada fitur yang digunakan, di mana pemilihan fitur yang informatif terbukti mampu meningkatkan akurasi tanpa menambah beban komputasi yang signifikan [6]. Oleh karena itu, kombinasi antara *feature extractor* modern seperti YAMNet dengan regresi logistik dapat menghasilkan sistem klasifikasi audio *deepfake* yang efisien dan tetap akurat.

Untuk memenuhi kebutuhan regresi logistik terhadap fitur yang relevan dan berkualitas tinggi, diperlukan metode ekstraksi fitur yang mampu merepresentasikan karakteristik audio secara komprehensif. YAMNet merupakan model berbasis MobileNetV1 yang telah dilatih menggunakan dataset AudioSet untuk mengklasifikasikan ratusan kategori suara, dan terbukti efektif sebagai *feature extractor* untuk berbagai tugas audio [7]. Melalui pendekatan *transfer learning*, representasi fitur yang dihasilkan YAMNet dapat digunakan ulang tanpa memerlukan pelatihan dari awal, sehingga secara signifikan mengurangi kebutuhan komputasi dan data pelatihan [8]. Penelitian terkini menunjukkan bahwa penggunaan YAMNet sebagai *feature extractor* mampu menghasilkan representasi audio yang ringkas namun informatif, sehingga mempermudah proses klasifikasi menggunakan model yang ringan seperti regresi logistik [7, 9]. Representasi yang dihasilkan YAMNet dinilai stabil dan efisien secara komputasi, menjadikannya solusi yang tepat untuk memenuhi kebutuhan regresi logistik terhadap fitur yang relevan tanpa menambah beban pemrosesan secara signifikan.

Berdasarkan uraian sebelumnya, tantangan utama dalam pengembangan sistem klasifikasi audio *deepfake* terletak pada kebutuhan akan model yang mampu mencapai akurasi tinggi dengan efisiensi komputasi yang optimal. Pendekatan berbasis *deep learning* murni sering kali memerlukan sumber daya besar, sedangkan regresi logistik menawarkan solusi yang lebih ringan namun sangat bergantung pada kualitas fitur yang digunakan. Untuk mengatasi hal tersebut, penelitian ini mengusulkan penggabungan YAMNet sebagai *feature extractor* dengan regresi logistik sebagai model klasifikasi. Integrasi kedua pendekatan ini diharapkan dapat menghasilkan sistem klasifikasi audio *deepfake* yang efisien secara komputasi, memiliki performa klasifikasi yang andal, serta berpotensi diterapkan pada skenario pemrosesan waktu nyata.

II. METODOLOGI PENELITIAN

Metodologi penelitian ini dirancang untuk menggambarkan alur proses dalam membangun sistem klasifikasi audio *deepfake* berbasis *transfer learning* menggunakan YAMNet dan regresi logistik. Setiap tahapan disusun secara sistematis agar menghasilkan model klasifikasi yang efisien dan akurat, mulai dari tahap *input* data audio hingga evaluasi performa model. Proses diawali dengan pemberian data audio sebagai masukan, kemudian dilakukan ekstraksi fitur menggunakan YAMNet untuk memperoleh representasi *embedding* audio. Hasil ekstraksi fitur tersebut selanjutnya dibagi menjadi data pelatihan dan pengujian (*train-test split*) untuk proses pemodelan regresi logistik. Setelah model dilatih, kinerjanya dievaluasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-score* untuk mengukur performa klasifikasi. Tahapan lengkap dari proses penelitian ini dapat dilihat pada Gambar 1.



Gambar 1 *Flowchart* Metodologi Penelitan

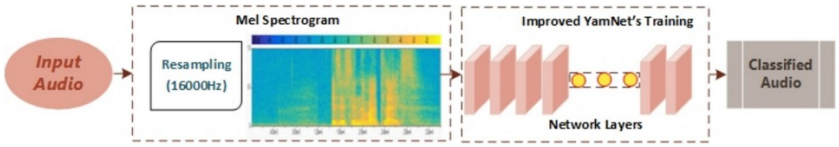
II.1 Input Audio

Penelitian ini menggunakan dataset sekunder "In-the-Wild", yang berisi rekaman audio dalam kondisi dunia nyata dengan variasi lingkungan dan melibatkan 54 pembicara berbeda [10]. Dataset ini mencakup suara asli (*bona-fide*) dan suara tiruan (*spoof*). Data bersumber dari berbagai konteks publik seperti pidato, wawancara, dan unggahan daring, sehingga mencerminkan karakteristik akustik yang kompleks dan representatif terhadap kondisi sebenarnya. Secara keseluruhan, dataset terdiri atas 31.779 file audio berformat WAV, dengan 19.963 sampel berlabel *bona-fide* dan 11.816 sampel berlabel spoof. Seluruh file telah di-*downsample* menjadi 16 kHz agar seragam dan sesuai dengan kebutuhan *input* model YAMNet. Pemilihan dataset ini memungkinkan model yang dikembangkan diuji pada kondisi yang lebih realistis, dengan variasi kebisingan, latar belakang, dan intonasi yang menantang, sehingga memberikan gambaran performa klasifikasi yang lebih akurat terhadap kondisi dunia nyata.

II.2 YAMNet *Embedding*

Pada tahap ini, model YAMNet digunakan untuk melakukan ekstraksi fitur otomatis dari sinyal audio yang telah melalui proses prapengolahan. YAMNet berperan sebagai *feature extractor* berbasis *transfer learning* yang mampu mengubah data audio mentah menjadi representasi fitur atau embedding berdimensi tinggi. Representasi ini memuat informasi penting mengenai pola spektral dan temporal dari sinyal suara yang relevan untuk proses klasifikasi. Dengan menggunakan model yang telah dilatih sebelumnya pada dataset AudioSet [11], YAMNet memungkinkan ekstraksi fitur dilakukan secara efisien tanpa perlu pelatihan ulang dari awal. Hasil *embedding* yang dihasilkan kemudian digunakan sebagai *input* bagi model regresi logistik pada tahap pemodelan selanjutnya.

YAMNet dibangun berdasarkan arsitektur MobileNetV1, yang mengadopsi konsep *depthwise separable convolution* untuk menghasilkan model dengan bobot ringan namun tetap mampu menangkap karakteristik audio secara mendalam. Arsitektur ini terdiri dari beberapa lapisan konvolusional dan *pooling* yang mengubah spektrum audio menjadi vektor fitur berukuran 1.024. Proses kerja dan struktur jaringan YAMNet secara umum ditunjukkan pada Gambar 2, yang menggambarkan alur transformasi dari sinyal audio masukan hingga terbentuknya *embedding* sebagai hasil ekstraksi fitur.



Gambar 2 Arsitektur dan Proses Ekstraksi Fitur Menggunakan YAMNet [12]

Setelah melalui proses ekstraksi fitur menggunakan YAMNet sebagaimana ditunjukkan pada Gambar 2, setiap file audio direpresentasikan sebagai vektor *embedding* berdimensi 1.024 yang memuat informasi spektral dan temporal utama dari sinyal suara. Representasi ini bersifat informatif, sehingga mampu menggambarkan karakteristik audio secara efisien. Hasil *embedding* dari seluruh sampel kemudian dikumpulkan dan disusun menjadi dataset fitur yang siap digunakan sebagai masukan bagi model regresi logistik pada tahap berikutnya.

II.3 *Splitting Data*

Tahap pembagian data dilakukan untuk memastikan proses pelatihan dan pengujian model berjalan secara seimbang dan representatif terhadap distribusi kelas dalam dataset. Dalam penelitian ini, data dibagi menggunakan metode *stratified splitting* dengan rasio 70% untuk pelatihan dan 30% untuk pengujian. Metode ini mempertahankan proporsi kelas *bona-fide* (asli) dan *spoof* (tiruan AI) pada kedua subset, sehingga mencegah ketidakseimbangan kelas yang dapat menyebabkan bias dalam proses pembelajaran model.

Pendekatan *stratified sampling* terbukti efektif dalam meningkatkan akurasi dan stabilitas klasifikasi karena memastikan bahwa distribusi setiap kelas tetap konsisten antara data pelatihan dan pengujian. Dengan demikian, model dapat mempelajari pola yang merepresentasikan seluruh populasi data secara lebih baik dan menghasilkan performa yang lebih andal [13].

II.4 Implementasi Model Regresi Logistik

Model regresi logistik digunakan sebagai algoritma klasifikasi utama dalam penelitian ini karena kemampuannya yang sederhana namun efektif dalam menangani permasalahan dengan data berdimensi tinggi dan hasil biner. Secara matematis, regresi logistik memanfaatkan fungsi sigmoid untuk mengubah kombinasi linear dari fitur *input* menjadi probabilitas keluaran antara 0 dan 1. Fungsi ini memungkinkan model untuk memprediksi kemungkinan suatu sampel termasuk dalam kelas positif atau negatif berdasarkan bobot fitur yang dipelajari selama proses pelatihan.

Menurut penelitian terdahulu, regresi logistik memiliki peran penting dalam analisis data dan penerapan *machine learning* karena kemampuannya memodelkan hubungan antara variabel prediktor dan hasil kategorikal secara efisien. Proses pembelajarannya merupakan jenis *supervised learning* yang mengevaluasi akurasi menggunakan fungsi *cross-entropy loss* yang menyertakan komponen regularisasi, serta melakukan pembaruan parameter bobot mengikuti aturan *gradient descent* dan *backpropagation* [14]. Model regresi logistik digunakan untuk mengklasifikasikan hasil ekstraksi fitur dari YAMNet ke dalam dua kategori, yaitu *spoof* dan *bona-fide*. Dengan input berupa *embedding* berdimensi 1.024, model ini dilatih untuk menemukan kombinasi bobot optimal yang memisahkan kedua kelas tersebut.

II.5 Evaluasi Model

Evaluasi model dilakukan untuk menilai kinerja regresi logistik dalam mengklasifikasikan audio *bona-fide* dan *spoof*. Pengukuran performa dilakukan menggunakan empat metrik utama, yaitu Akurasi, Presisi, *Recall*, dan *F1-Score*, yang masing-masing dihitung berdasarkan hasil *confusion matrix*.

Akurasi digunakan untuk mengukur proporsi prediksi yang benar terhadap seluruh data uji, baik pada kelas positif maupun negatif. Rumusnya dituliskan sebagai:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Nilai akurasi yang tinggi menandakan bahwa model mampu mengklasifikasikan sebagian besar sampel dengan benar.

Presisi digunakan untuk mengukur ketepatan model dalam memprediksi kelas positif dari semua prediksi yang dikategorikan positif. Rumusnya adalah:

$$Presisi = \frac{TP}{TP + FP} \quad (2)$$

Nilai resisi yang tinggi menunjukkan bahwa model jarang menghasilkan kesalahan prediksi positif.

Recall menggambarkan kemampuan model dalam mendeteksi seluruh data positif yang sebenarnya ada dalam dataset. Dihitung dengan rumus:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Nilai *recall* yang tinggi berarti model mampu mengenali sebagian besar sampel positif dengan benar.

F1-Score merupakan rata-rata harmonis antara presisi dan *recall*, yang memberikan keseimbangan antara kedua metrik tersebut terutama pada data dengan distribusi kelas tidak seimbang. Rumusnya sebagai berikut:

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (4)$$

Keempat metrik tersebut digunakan bersama untuk memberikan gambaran menyeluruh mengenai performa model dari segi ketepatan, sensitivitas, dan keseimbangan antara kesalahan prediksi positif dan negatif [15].

Selain menghitung keseimbangan prediksi, penelitian ini juga mengevaluasi kecepatan inferensi model pada data uji untuk menilai efisiensi komputasi. Pengukuran dilakukan hanya pada tahap prediksi menggunakan model regresi logistik, setelah proses pelatihan selesai. Waktu inferensi dihitung berdasarkan durasi yang dibutuhkan model untuk melakukan satu kali prediksi terhadap seluruh data uji dalam satu proses pengujian penuh.

III. HASIL DAN ANALISIS

III.1 Input Audio

Penelitian ini menggunakan dataset In-the-Wild yang berisi rekaman audio dalam kondisi nyata. Proses pemuatan data dilakukan dengan memindai seluruh file audio berekstensi .wav yang terdapat dalam direktori dataset. Pembacaan setiap file dilakukan menggunakan *library* Librosa, yang berfungsi untuk memuat sinyal audio menjadi bentuk array numerik serta mengekstraksi informasi penting seperti *sampling rate* dan durasi. Penggunaan Librosa memungkinkan proses pembacaan berjalan efisien dan konsisten pada seluruh data.

Berdasarkan proses tersebut, diperoleh 31.779 sampel audio yang siap digunakan untuk pelatihan dan pengujian. Dataset ini terdiri atas dua kelas utama, yaitu *bona-fide* yang merupakan suara asli manusia, dan *spoof* yang merupakan suara tiruan atau *deepfake*.

III.2 YAMNet *Embedding*

Proses ekstraksi fitur dilakukan menggunakan model *pre-trained* YAMNet yang bersumber dari TensorFlow Hub. Setiap file audio yang telah dimuat menggunakan *library* Librosa dikonversi menjadi representasi numerik berupa vektor *embedding* berdimensi 1.024. YAMNet memproses masukan berupa *mel-spectrogram* untuk menghasilkan vektor fitur yang merepresentasikan karakteristik esensial sinyal suara. Contoh hasil *embedding* beserta label kelas yang bersesuaian disajikan pada Tabel 1.

Tabel 1. Contoh hasil *embedding* YAMNet

No	Nilai *Embedding*	Label
1	0.11113991 0.38690013 0.15825415 ... 0.02736065 0.02325923 0.00066478	0
2	0.12385013 0.31513914 0.12581564 ... 0.08742192 0.00805856 0.00057245	1

III.3 *Splitting Data*

Proses pembagian data (*splitting*) dilakukan untuk memisahkan dataset menjadi dua bagian, yaitu data pelatihan dan data pengujian, dengan tujuan untuk mengevaluasi kemampuan generalisasi model. Metode *stratified splitting* digunakan agar distribusi kelas *bona-fide* dan *spoof* tetap seimbang pada kedua subset. Pembagian ini dilakukan dengan proporsi 70% untuk data pelatihan dan 30% untuk data pengujian, serta menetapkan parameter "random_state=5" agar proses pembagian dapat direproduksi dengan konsisten.

Hasil pembagian data menghasilkan 22.245 sampel pada subset pelatihan dan 9.534 sampel pada subset pengujian. Metode *stratified splitting* memastikan bahwa proporsi antara kelas *bona-fide* dan *spoof* tetap seimbang pada kedua subset. Dengan demikian, model dapat mempelajari pola dari setiap kelas secara proporsional dan menghasilkan evaluasi yang lebih representatif terhadap kinerja sebenarnya.

III.4 Implementasi Model Regresi Logistik

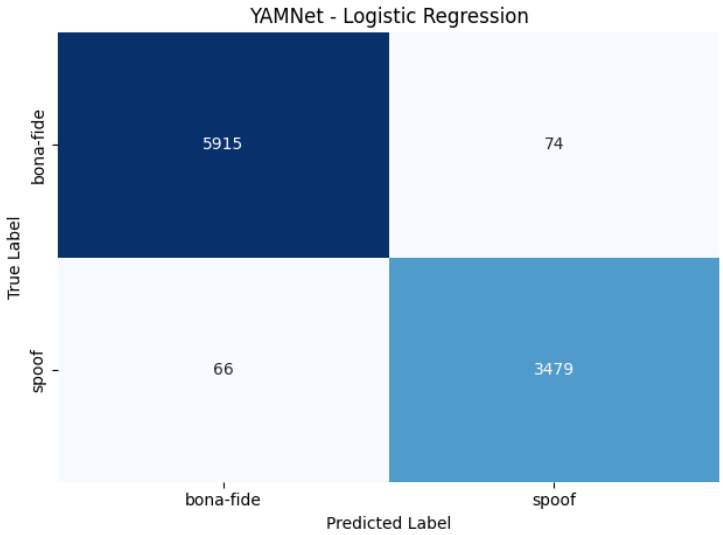
Pelatihan model dilakukan menggunakan *library* Scikit-learn dengan algoritma regresi logistik. Sebanyak 70% dari total data digunakan sebagai data pelatihan, sedangkan 30% sisanya dialokasikan untuk data pengujian, yang diperoleh dari hasil proses *stratified splitting* pada tahap sebelumnya agar proporsi kelas tetap seimbang. Parameter yang digunakan dalam proses pelatihan hanyalah "random_state=5", yang berfungsi menjaga konsistensi hasil pelatihan agar dapat direproduksi pada setiap percobaan.

Selama proses pelatihan, model mempelajari hubungan antara fitur hasil *embedding* dari YAMNet dengan label kelas untuk menentukan batas keputusan optimal antara kategori *bona-fide* dan *spoof*. Model yang telah dilatih kemudian disimpan untuk

digunakan pada tahap evaluasi, di mana performanya akan diuji menggunakan data pengujian untuk menilai tingkat akurasi dan kemampuan generalisasi model.

III.5 Evaluasi Model

Evaluasi kinerja dilakukan untuk mengukur kemampuan model regresi logistik dalam membedakan antara audio asli dan *deepfake* berdasarkan fitur *embedding* hasil ekstraksi dari YAMNet. Pengujian dilakukan menggunakan data uji yang belum pernah dilihat model sebelumnya untuk menilai tingkat generalisasi dan akurasi klasifikasi yang dihasilkan. Visualisasi hasil klasifikasi ditunjukkan pada Gambar 3, yang menampilkan *confusion matrix* dari hasil prediksi model terhadap data uji.



Gambar 3 *Confusion Matrix* Model Regresi Logistik Berbasis YAMNet

Berdasarkan Gambar 3, terlihat bahwa model memiliki kemampuan klasifikasi yang sangat baik dengan dominasi nilai diagonal utama yang menunjukkan jumlah prediksi benar pada kedua kelas. Dapat dilihat jika 5.915 sampel *bona-fide* diklasifikasikan dengan benar dan hanya 74 sampel *bona-fide* yang salah diklasifikasikan sebagai *spoof*. Sementara itu, 3.479 sampel *spoof* berhasil dikenali dengan benar, dengan 66 sampel *spoof* yang salah terdeteksi sebagai *bona-fide*. Proporsi kesalahan yang kecil ini menunjukkan bahwa model memiliki ketepatan tinggi dan mampu membedakan kedua kelas secara konsisten.

Untuk memberikan gambaran kuantitatif terhadap performa model, hasil pengujian ditampilkan dalam Tabel 2, yang memuat nilai empat metrik utama, yaitu akurasi, presisi, *recall*, dan *F1-score*.

Tabel 2. Hasil Evaluasi Performa Model Regresi Logistik

Metrik	Nilai
Akurasi	98.53%
Presisi	0.9890
Recall	0.9876
F1-Score	0.9883

Nilai akurasi sebesar 98.53% menunjukkan bahwa sebagian besar sampel pada data uji berhasil diklasifikasikan dengan benar. Nilai presisi 0.9890 menandakan bahwa model sangat jarang melakukan kesalahan dalam mengidentifikasi *spoof* sebagai *bona-fide*. Sementara itu, nilai *recall* 0.9876 menggambarkan kemampuan model dalam mendeteksi hampir seluruh audio *bona-fide* secara benar. Kombinasi dari presisi dan

recall menghasilkan nilai F1-score 0.9883 yang menunjukkan keseimbangan performa antara kedua metrik tersebut.

Selain hasil metrik evaluasi, dilakukan pengukuran waktu inferensi untuk menilai efisiensi model dalam melakukan prediksi. Berdasarkan hasil pengujian, model regresi logistik membutuhkan waktu sekitar 0.0276 detik untuk melakukan satu kali prediksi terhadap seluruh data uji. Nilai ini menunjukkan bahwa proses inferensi berjalan sangat cepat yang menegaskan karakteristik regresi logistik yang ringan secara komputasi.

IV. KESIMPULAN

Penelitian ini berhasil mengimplementasikan sistem klasifikasi audio *deepfake* dengan menggabungkan YAMNet sebagai ekstraktor fitur dan Regresi Logistik sebagai model klasifikasi yang efektif. Melalui proses ekstraksi fitur berdimensi 1.024 dari YAMNet dan pelatihan model menggunakan data hasil *stratified splitting* (70% pelatihan, 30% pengujian), sistem ini mampu mencapai kinerja optimal dengan akurasi 98.53%, presisi 0.9890, *recall* 0.9876, dan *F1-score* 0.9883. Selain menunjukkan akurasi tinggi, model juga memiliki efisiensi komputasi yang sangat baik dengan waktu inferensi hanya 0.0276 detik untuk satu kali prediksi terhadap seluruh data uji. Hasil ini menegaskan bahwa fitur *embedding* dari YAMNet dapat meningkatkan performa klasifikasi tanpa menambah beban komputasi, sementara Regresi Logistik memberikan klasifikasi yang cepat, stabil, dan efisien. Dengan demikian, pendekatan ini terbukti efektif untuk mengklasifikasikan audio *deepfake* secara cepat dan akurat serta memiliki potensi kuat untuk diterapkan pada sistem keamanan berbasis audio di lingkungan dengan keterbatasan sumber daya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya atas segala bentuk dukungan moral, semangat, dan doa yang telah diberikan oleh berbagai pihak selama proses penyusunan penelitian ini. Dukungan tersebut menjadi sumber motivasi penting dalam menyelesaikan karya ilmiah ini dengan baik.

REFERENSI

[1] R. Komalasari, “Manfaat Teknologi Informasi dan Komunikasi di Masa Pandemi Covid 19,” *TEMATIK*, vol. 7, no. 1, pp. 38–50, June 2020, doi: 10.38204/tematik.v7i1.369.

[2] D. A. Satria, A. Sidauruk, and I. A. Pirmansah, “Pemanfaatan Teknologi Generatif Artificial Intelligence sebagai Media Promosi bagi Guru dan Staf Publikasi di Sekolah Dasar”.

[3] S. W. Nurdin and I. F. Nugraha, “ANCAMAN DEEPFAKE DAN DISINFORMASI BERBASIS AI: IMPLIKASI TERHADAP KEAMANAN SIBER DAN STABILITAS NASIONAL INDONESIA,” vol. 4, no. 01, 2025.

[4] A. Hery, O. Joseph, O. Femi, and H. Luz, “Audio deepfakes: threats to voice assistants and voice-activated systems”.

[5] S. Singh, J. Kubica, S. Larsen, and D. Sorokina, “Parallel Large Scale Feature Selection for Logistic Regression,” in *Proceedings of the 2009 SIAM International Conference on Data Mining*, Society for Industrial and Applied Mathematics, Apr. 2009, pp. 1172–1183. doi: 10.1137/1.9781611972795.100.

[6] R. Islam, S. Mazumdar, and R. Islam, "An Experiment on Feature Selection using Logistic Regression," Jan. 31, 2024, *arXiv*: arXiv:2402.00201. doi: 10.48550/arXiv.2402.00201.

[7] N. H. Valliappan, S. D. Pande, and S. Reddy Vinta, "Enhancing Gun Detection With Transfer Learning and YAMNet Audio Classification," *IEEE Access*, vol. 12, pp. 58940–58949, 2024, doi: 10.1109/ACCESS.2024.3392649.

[8] E. Brusa, C. Delprete, and L. G. Di Maggio, "Deep Transfer Learning for Machine Diagnosis: From Sound and Music Recognition to Bearing Fault Detection," *Appl. Sci.*, vol. 11, no. 24, p. 11663, Dec. 2021, doi: 10.3390/app112411663.

[9] S. Lachenani, H. Kheddar, and M. Ouldzmirli, "Improving Pretrained YAMNet for Enhanced Speech Command Detection via Transfer Learning," in *2024 International Conference on Telecommunications and Intelligent Systems (ICTIS)*, Djelfa, Algeria: IEEE, Dec. 2024, pp. 1–6. doi: 10.1109/ICTIS62692.2024.10894266.

[10] N. M. Müller, P. Czempin, F. Dieckmann, A. Froghyar, and K. Böttinger, "Does Audio Deepfake Detection Generalize?," Aug. 27, 2024, *arXiv*: arXiv:2203.16263. doi: 10.48550/arXiv.2203.16263.

[11] J. F. Gemmeke *et al.*, "Audio Set: An ontology and human-labeled dataset for audio events," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA: IEEE, Mar. 2017, pp. 776–780. doi: 10.1109/ICASSP.2017.7952261.

[12] R. Mahum, A. Irtaza, A. Javed, H. A. Mahmoud, and H. Hassan, "DeepDet: YAMNet with BottleNeck Attention Module (BAM) for TTS synthesis detection," *EURASIP J. Audio Speech Music Process.*, vol. 2024, no. 1, p. 18, Apr. 2024, doi: 10.1186/s13636-024-00335-9.

[13] J. Sadaiyandi, P. Arumugam, A. K. Sangaiah, and C. Zhang, "Stratified Sampling-Based Deep Learning Approach to Increase Prediction Accuracy of Unbalanced Dataset," *Electronics*, vol. 12, no. 21, p. 4423, Oct. 2023, doi: 10.3390/electronics12214423.

[14] A. M. Rakhimovich, K. K. Kadirbergenovich, Z. M. Ishkobilovich, and K. J. Kadirbergenovich, "Logistic Regression with Multi-Connected Weights," *J. Comput. Sci.*, vol. 20, no. 9, pp. 1051–1058, Sept. 2024, doi: 10.3844/jcssp.2024.1051.1058.

[15] Jude Chukwura Obi, "A comparative study of several classification metrics and their performances on data," *World J. Adv. Eng. Technol. Sci.*, vol. 8, no. 1, pp. 308–314, Feb. 2023, doi: 10.30574/wjaets.2023.8.1.0054.