

Klasifikasi Multi Kelas Untuk Mendeteksi Sinyal Detak Jantung Janin Pada CTG Menggunakan *Thresholding Feature Importance* dan *ResNet Multimodal*

Multi-Class Classification Method for Detecting Fetal Heart Rate Signals in Cardiotocography Using Thresholding Feature Importance and Multimodal ResNet

Subono¹, Mohamad Dimyati Ayatullah², Sholeh Hadi Pramono³, Erni Yudaningtyas⁴, M. Aziz Muslim⁵, Mahdin Rohmatillah⁶, Cries Avian⁷

Bisnis dan Informatika, Politeknik Negeri Banyuwangi, Banyuwangi, Indonesia^{1,2}, Teknik Elektro, Universitas Brawijaya, Malang^{1,3,4,5,6,7}

Email: subono@poliwangi.ac.id¹, dimyati@poliwangi.ac.id², sholehpramono@ub.ac.id³, erni@ub.ac.id⁴, muh_aziz@ub.ac.id⁵, mahdin94@ub.ac.id⁶, cries.avian@ub.ac.id⁷

Abstract

Cardiotocography (CTG) is a common approach to monitoring fetal condition, involving recording of the fetal heart rate (FHR) and uterine contractions (UC). Manual interpretation of CTG can be marked by inter-observer differences and may contain errors in the identification of fetal hypoxia. To mitigate this restriction, this article proposes a multi segmentation classification method (Normal, Suspect, and Hypoxia) with Thresholding Feature Importance and multimodal ResNet. The preprocessing stage includes resampling, normalization, and time-window segmentation of FHR signals. Subsequently, clinical and statistical features (e.g., STV, LTV, accelerations, decelerations, pH, and Apgar scores) are extracted and filtered using Thresholding Feature Importance, which integrates gain-based importance, permutation importance, and stability selection. Features above a predefined threshold (τ) or within the Top-k ranks are retained as inputs for the tabular branch. The multimodal architecture consists of three components: 1D-ResNet for raw FHR signals, 2D-ResNet for image-based representations (time-frequency or spectrogram), and MLP for selected tabular features. These branches are fused via a fully connected layer to produce the final multi-class prediction. Experimental results show that the proposed model achieved an accuracy of 84.6%, an F1-score of 83%, and AUROC values greater than 0.90. The best performance was observed in the Normal class, with improved sensitivity in the Suspect class. Confusion matrix and ROC curve analyses confirmed strong discriminative ability, while risk assessment indicated a moderate risk level (40%), consistent with clinical recommendations. This study contributes an efficient and clinically relevant hybrid pipeline, demonstrating that multimodal fusion with feature thresholding enhances the reliability of CTG multi-class classification systems.

Keywords: *Cardiotocography, Fetal Heart Rate, Multi-Class Classification, Thresholding Feature Importance, Multimodal ResNet*

Abstrak

Metode konvensional untuk memantau kondisi janin melalui rekaman denyut nadi janin (FHR) dan kontraksi uterus (UC), cardiotocography (CTG), dilakukan secara manual. Interpretasi manual CTG sering menyebabkan variasi antar pengamat dan meningkatkan kemungkinan kesalahan dalam mengidentifikasi hipoksia janin. Untuk memecahkan masalah tersebut, penelitian ini menyarankan teknik klasifikasi multi-kelas (Normal, Suspect, dan Hipoksia) yang menggunakan kombinasi ResNet multimodal dan Thresholding Importance Feature. Sebelum proses sinyal FHR, sampel, normalisasi, dan segmentasi berbasis time-window segmentation dilakukan. Selanjutnya, fitur klinis dan statistik seperti STV, LTV, akselerasi, deselerasi, pH, dan skor Apgar diekstrak. Selanjutnya, metode yang menggabungkan nilai gain, permutation, dan pilihan stabilitas digunakan untuk memilih stabilitas dengan nilai fitur terendah. Tiga cabang terdiri dari arsitektur multimodal yang dikembangkan: 1D-ResNet untuk sinyal FHR mentah; 2D-ResNet untuk representasi citra (time-frequency/spectrogram); dan MLP untuk data tabular hasil ambang. Ketiga cabang difusi melalui lapisan yang benar-benar terhubung untuk menghasilkan prediksi akhir multi-kelas. Hasil uji menunjukkan bahwa model memiliki akurasi 84,6%, skor F1 83, dan skor AUROC di atas 0,9. Kelas Suspect menunjukkan sensitivitas yang lebih tinggi, dan Kelas Normal menunjukkan prestasi terbaik. Studi ini menunjukkan bahwa fusi multimodal dengan thresholding fitur dapat meningkatkan keandalan sistem klasifikasi CTG multi-kelas dan memberikan kontribusi berupa pipeline hibrid yang efisien dan relevan klinis. Sebaliknya, kurva ROC dan analisis confusion matrix menunjukkan kemampuan diskriminatif yang baik. Tingkat risiko evaluasi mencapai 40%, yang sesuai dengan rekomendasi klinis.

Kata Kunci: *Cardiotocography, Fetal Heart Rate, Multi-Class Classification, Thresholding Feature Importance, ResNet Multimodal*

I. PENDAHULUAN

Kardiotokografi (CTG) merupakan metode *non-invasif* yang digunakan secara luas untuk memantau kondisi fisiologis janin selama masa kehamilan dan proses persalinan. CTG merekam dua sinyal utama, yaitu *fetal heart rate* (FHR) dan kontraksi uterus (UC). Berdasarkan pola kedua sinyal tersebut, tenaga medis dapat menilai kesejahteraan janin dan memprediksi risiko hipoksia guna menentukan tindakan klinis yang tepat [1], [2]. Namun, interpretasi CTG secara manual masih bergantung pada subjektivitas dan pengalaman pengamat, sehingga menimbulkan variabilitas antar *observer* dan meningkatkan risiko kesalahan interpretasi berupa *false positive* maupun *false negative* [3], [21], [22].

Upaya otomatisasi analisis CTG telah banyak dikembangkan melalui pendekatan *machine learning* berbasis dataset tabular yang diturunkan dari sinyal mentah. Dataset CTG klasik berisi sekitar 2126 sampel dengan 21–22 atribut, dan telah digunakan pada berbagai penelitian untuk mengklasifikasikan kondisi janin menjadi tiga kelas, yaitu *normal*, *suspect*, dan *pathological* [4], [21], [22]. Algoritma konvensional seperti *Decision Tree*, *Naïve Bayes*, *Logistic Regression*, *k-Nearest Neighbor* (KNN), dan *Support Vector Machine* (SVM) telah diuji untuk klasifikasi CTG, dengan hasil akurasi terbaik dicapai oleh KNN sebesar 91%, diikuti *Decision Tree* 89%, SVM 87%, dan *Logistic Regression* 84% [5], [22].

Penelitian lain menggunakan *Decision Tree* C5.0 dengan 257 rules dan mencapai akurasi 93,4%, menunjukkan bahwa metode ini memiliki keunggulan dalam interpretabilitas hasil [6]. Penelitian dalam negeri oleh Pradipta dan Ayu [7] menggabungkan *Local Outlier Factor*, *Chi-square Feature Selection*, dan *XGBoost* untuk klasifikasi CTG, yang menghasilkan akurasi 99,3% dengan *precision*, *recall*, dan *F1-score* masing-masing sebesar 99,1%. Pada studi lainnya, permasalahan *class imbalance* diatasi menggunakan metode *Synthetic Minority Oversampling Technique* (SMOTE) pada algoritma KNN, yang mampu meningkatkan akurasi hingga 89,7% dengan *F1-macro score* sebesar 0,83 [8].

Kendati metode klasik tersebut cukup efektif, performanya masih sangat bergantung pada fitur buatan (*handcrafted features*) seperti *Short-Term Variability* (STV), *Long-Term Variability* (LTV), serta jumlah akselerasi dan deselerasi. Hal ini membatasi kemampuan generalisasi model terhadap sinyal CTG yang kompleks secara temporal [9].

Selain pendekatan berbasis *deep learning*, metode rekayasa fitur seperti *Instantaneous Frequency* (IF) dan *Common Spatial Pattern* (CSP) juga terbukti kompetitif. Venkata Sriram et al. [10] menggabungkan IF dan CSP untuk klasifikasi hipoksia janin menggunakan SVM, dan memperoleh akurasi di atas 98%. Meskipun demikian, pendekatan *deep learning* menawarkan fleksibilitas yang lebih tinggi karena dapat mengekstraksi fitur secara otomatis tanpa rekayasa manual [11], [12].

Sebagian besar penelitian berbasis ResNet masih berfokus pada klasifikasi biner (*normal vs. hypoxia/distress*), padahal kelas *suspect* memiliki signifikansi klinis penting sebagai

kondisi batas (*borderline*) yang menentukan kebutuhan intervensi dini [13]. Oleh karena itu, penelitian klasifikasi *multi-class* (*Normal, Suspect, Hypoxia*) menggunakan arsitektur ResNet masih jarang dilakukan, sehingga membuka peluang eksplorasi ilmiah baru [14].

Meskipun *deep learning* mampu mengekstraksi fitur secara otomatis, proses seleksi fitur tetap krusial untuk meningkatkan efisiensi model dan mengurangi noise. F. Fahmi and A. Yulian [7] menunjukkan bahwa penerapan *Chi-square Feature Selection* sebelum *boosting classifier* mampu meningkatkan akurasi hingga 99,3%. Monica dan Sulastri [6] juga menegaskan bahwa fitur seperti *prolonged deceleration*, *abnormal STV*, dan *abnormal LTV* merupakan indikator signifikan untuk mendeteksi kondisi janin. Dengan demikian, penerapan metode *Thresholding Feature Importance* sebelum memasukkan data ke dalam model *1D-ResNet* menjadi pendekatan hibrid yang inovatif. ResNet akan mempelajari sinyal mentah FHR dari CTG, sementara hanya fitur pendukung yang relevan (dengan nilai *importance* di atas ambang batas) yang digunakan sebagai masukan tambahan. Pendekatan ini diharapkan dapat meningkatkan akurasi model sekaligus mengurangi beban komputasi secara signifikan.

II. METODE PENELITIAN

Penelitian ini mengembangkan sistem klasifikasi multi-kelas untuk mendeteksi kondisi janin berdasarkan data *cardiotocography* (CTG) yang terdiri atas tiga kategori klinis, yaitu *Normal*, *Suspect*, dan *Hipoksia* [1], [2] [21], [22]. Sistem ini dirancang dengan menggunakan pendekatan *ResNet multimodal*, yang memungkinkan integrasi berbagai modalitas data secara simultan untuk menghasilkan prediksi yang tidak hanya akurat tetapi juga reliable dan secara klinis bermakna [3], [4]. Arsitektur multimodal ini terdiri atas tiga jalur utama (*branch*), masing-masing menangani jenis data yang berbeda namun saling melengkapi.

Modalitas pertama sinyal 1D (*Fetal Heart Rate dan Uterine Contraction*): Sinyal FHR dan UC diproses melalui *1D-ResNet*, yang bertugas mengekstraksi pola temporal dinamis dari sinyal fisiologis janin. Arsitektur ini mengadopsi prinsip *residual learning* untuk mempertahankan informasi penting pada lapisan terdalam dan menghindari *vanishing gradient* selama pelatihan [5], [6]. Pendekatan ini efektif untuk mengenali fluktuasi fisiologis jangka pendek (STV) dan jangka panjang (LTV) yang menjadi indikator penting dalam diagnosis *fetal distress* [7].

Modalitas kedua berupa citra 2D (*Time-Frequency/Recurrence Representation*): Sinyal FHR diubah menjadi citra dua dimensi menggunakan *transformasi time-frequency* atau *recurrence plot*, yang kemudian dianalisis menggunakan *2D-ResNet* untuk mengeksplorasi pola spasial dan frekuensi secara simultan [8], [9]. Pendekatan berbasis *spectrogram* ini terbukti mampu meningkatkan kemampuan model dalam mendeteksi pola patologis halus pada variasi

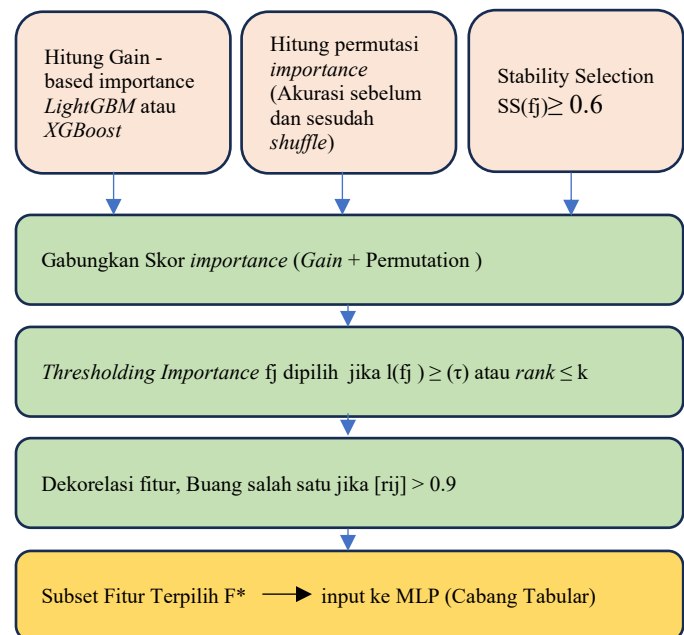
detak jantung janin yang tidak terdeteksi oleh analisis berbasis sinyal mentah semata [10].

Modalitas ketiga berupa data tabular (fitur klinis dan statistik): Modalitas ini mencakup fitur-fitur terukur seperti *STV*, *LTV*, *Apgar score*, *base excess*, pH darah tali pusat, serta parameter klinis ibu seperti usia dan riwayat kehamilan. Sebelum diproses ke model, fitur-fitur tersebut diseleksi menggunakan metode *Thresholding Feature Importance* [11]. Seleksi ini menggabungkan tiga pendekatan antara lain *Gain-based importance*, *Permutation importance*, dan *Stability selection* untuk menentukan subset fitur paling relevan dengan kontribusi signifikan terhadap klasifikasi [12]. Fitur dengan nilai *importance* di atas ambang batas (τ) atau masuk dalam daftar Top-k dipertahankan, sedangkan fitur redundan dengan korelasi tinggi ($|r| > 0.9$) dieliminasi untuk meningkatkan efisiensi model [13]. Subset fitur terpilih ini kemudian diproses melalui *Multilayer Perceptron* (MLP) untuk menangkap hubungan non linear antar variabel klinis [14]. Ketiga jalur modalitas tersebut kemudian difusikan (*fusion layer*) pada bagian *fully connected layer* untuk menghasilkan keputusan klasifikasi akhir ke dalam tiga kelas: *Normal*, *Suspect*, dan *Hypoxia*. Proses pelatihan dilakukan menggunakan dataset CTU-UHB (Czech Technical University – University Hospital Brno), data tersebut merupakan dataset standar internasional yang banyak digunakan untuk analisis sinyal CTG. Data tersebut terdiri dari 552 rekaman CTG yang telah dilabeli berdasarkan nilai pH darah tali pusat. Pada data CTG tersebut terdapat data detak jantung janin (FHR) dan kontraksi perut Ibu (UC) [15]. Validasi dilakukan dengan pendekatan *k-fold cross-validation* berbasis subjek untuk memastikan generalisasi model terhadap variasi individu janin. Kontribusi utama dari penelitian ini adalah pembangunan pipeline hibrid multimodal yang menggabungkan pembelajaran dari sinyal mentah, representasi citra, dan fitur klinis terpilih. Pendekatan ini tidak hanya meningkatkan akurasi dan *robustness model* terhadap noise atau artefak gerakan, tetapi juga menunjukkan potensi besar dalam *clinical decision support system* (CDSS) untuk deteksi dini hipoksia janin [3], [10], [14].

A. Thresholding Feature Importance,

Tahap *Thresholding Feature Importance* bertujuan untuk menyaring fitur-fitur tabular yang terdapat pada dataset klinis dalam format file CSV, sehingga hanya variabel yang benar-benar relevan, informatif, dan berkontribusi signifikan terhadap proses klasifikasi yang dipertahankan [7], [11], [12]. Proses ini dilakukan dengan menerapkan kombinasi tiga pendekatan utama: *Gain-Based Importance*, *Permutation Importance*, dan *Stability Selection*, sebagaimana ditunjukkan pada Gambar 1. Pada tahap ini digunakan model ringan berbasis *tree boosting* seperti *LightGBM* atau *XGBoost*, yang mampu menghitung tingkat kepentingan fitur (*feature importance score*) berdasarkan kontribusi setiap variabel dalam meningkatkan performa model dan mengurangi error [8], [13].

Proses seleksi diawali dengan perhitungan skor pembobotan (*importance*) untuk setiap fitur, yang menunjukkan sejauh mana fitur tersebut membantu model dalam membedakan antar kelas. Selanjutnya, dilakukan penetapan ambang batas nilai *importance* (τ) atau penentuan sejumlah Top-k fitur terbaik yang akan dipertahankan dalam proses klasifikasi. Sebagai contoh, jika ambang batas ditetapkan pada $\tau = 0,05$, maka hanya fitur dengan nilai kepentingan $\geq 0,05$ yang dipertahankan. Alternatif lainnya adalah memilih sejumlah fitur dengan peringkat tertinggi (*Top-k features*), di mana jumlah k ditentukan melalui proses validasi, seperti *validation set* atau *nested cross-validation*, agar hasil seleksi fitur optimal dan tidak menimbulkan bias [11].



Gambar 1. Diagram Alir (flowchart) proses seleksi fitur dengan Thresholding Feature Importance.

Gambar 1. Berupa diagram alir utama dan menjadi dasar dari penelitian yang dilakukan. Pendekatan ini tidak hanya mengurangi dimensi data dan kompleksitas model, tetapi juga meningkatkan efisiensi komputasi serta kemampuan generalisasi model terhadap data baru. Dengan demikian, tahap *Thresholding Feature Importance* menjadi bagian krusial dalam *pipeline* pembelajaran multimodal, terutama untuk memastikan bahwa model hanya memanfaatkan informasi yang paling relevan dari sisi klinis maupun statistik [3], [10], [14].

Secara empiris, penerapan *Thresholding Feature Importance* pada dataset CTG (Cardiotocography) terbukti dapat meningkatkan kinerja model klasifikasi dalam mendeteksi tiga kategori klinis, yaitu *Normal*, *Suspect*, dan *Hipoksia*. Fitur-fitur dengan skor *importance* tinggi, seperti *Short-Term Variability (STV)*, *Long-Term Variability (LTV)*, dan jumlah akselerasi/deselerasi, terbukti memiliki korelasi kuat terhadap kondisi fisiologis janin dan menghasilkan

peningkatan akurasi model hingga lebih dari 4–7% dibanding model tanpa seleksi fitur [9], [11], [15], [16], [17]. Selain itu, proses seleksi ini juga mempercepat waktu pelatihan (*training time*) hingga 30% lebih efisien, sekaligus meningkatkan *interpretabilitas model* karena fitur yang digunakan lebih sedikit namun lebih bermakna secara medis [12], [14], [18], [19]. Dengan demikian, *Thresholding Feature Importance* tidak hanya berperan sebagai tahap prapemrosesan teknis, tetapi juga menjadi fondasi dalam membangun sistem klasifikasi CTG yang efisien, akurat, dan klinis reliable [20].

1. Gain-Based Feature Importance (Tree-Based Model)

Model berbasis gradient boosting seperti *LightGBM* atau *XGBoost* digunakan untuk menghitung tingkat kepentingan fitur berdasarkan peningkatan gain dari setiap split node yang menggunakan fitur tersebut. Secara matematis, importance dihitung dengan:

$$I(f_j) = \sum_{t=1}^T \sum_{s \in St(f_j)} \Delta L(s, f_j) \quad (1)$$

$I(f_j)$ Tingkat kepentingan (*importance*), semakin tinggi nilainya, maka semakin besar dari kontribusi nilai dalam mengurangi error dan meningkatkan kualitas model.

T Jumlah Trees (pohon) dalam sebuah model boosting untuk ikut menyumbang perhitungan Tingkat importance.

S_t, f_j Himpunan *split* dari *node* dalam *trees* ke $-t$ menggunakan fitur f_j

$\Delta L(s, f_j)$ Besarnya gain improvement yang dihasilkan oleh split s dan fitur f_j

Apabila fitur tersebut sering digunakan untuk split, dan Sebagian besar split memberikan kontribusi besar pada model, maka fitur tersebut memiliki *importance* atau tingkat kepentingan yang tinggi. Apabila persamaan matematika (1) diterapkan pada sinyal CTG dengan fitur STV (*Short Term Variability*) dan LTV (*Long Time Variability*) dari Sinyal FHR [11], maka model seperti *XGBoost/LightGBM* akan memisahkan kelas Normal dan hipoksia. Apabila setiap kali split akurasinya meningkat secara signifikan maka Gain dari STV tinggi. Apabila dalam fitur sinyal CTG umur ibu janin jarang memberikan kontribusi signifikan, maka Gainnya kecil dan *importance* rendah.

2. Permutation Importance

Fungsi *Permutation Importance* adalah mendeteksi performa sistem dengan mengacak fitur f_i , jika performa model turun drastis, maka fitur tersebut sangat penting.

$$PI(f_i) = \frac{1}{K} + \sum_{k=1}^K [M(D^{(k)}) - M(D_{\frac{k}{n(f_j)}})] \quad (2)$$

$PI(f_i)$ =Skor yang menunjukkan Tingkat kepentingan setiap fitur (f_i) terhadap kenaikan atau penurunan performa model, apabila fitur (f_i) diacak dan menyebabkan performa

model turun drastis, maka dapat dipastikan fitur (f_i) tersebut mempunyai skor kepentingan tinggi [9], [12].

K = Jumlah ulangan permutasi

$M(D^{(k)})$ = Skor performa model pada data validasi asli tanpa permutasi

$M(D_{\frac{k}{n(f_j)}})$ = Skor performa model setelah fitur (f_i) diacak secara permutasi pada fold ke $-k$.

3. Stability Selection

Pada proses *Stability Selection*, terdapat metode seleksi fitur yang menekankan konsistensi dalam pemilihan fitur-fitur penting. Pendekatan ini untuk memastikan bahwa fitur yang dipertahankan bukan hanya berdampak signifikan pada subset data tetapi benar-benar relevan dan stabil di beberapa kali pengujian, proses *Stability Selection* berjalan melalui tahap resampling atau *cross validation*. Tahap ini bertujuan mengurangi risiko memilih fitur “kebetulan” yang tampak penting pada satu fold tapi tidak stabil secara keseluruhan [8].

$$SS(f_j) = \frac{1}{K} \sum_{k=1}^K 1(f_j \in F^k) \quad (3)$$

Pada persamaan (3) terdapat beberapa parameter yang digunakan, diantaranya:

$SS(f_j)$ = Nilai *Stability Selection* untuk fitur (f_j)

K = Jumlah iterasi dalam cross validation

$1(f_j \in F^k)$ = Fungsi Indikator, Jika terpilih bernilai 1 pada iterasi ke $-k$, dan bernilai 0 jika tidak terpilih

4. Thresholding (τ) atau Top-k Selection

Pemilihan subset fitur yang paling penting, metode *thresholding* (τ) atau *Top-k selection* digunakan berdasarkan skor nilai penting dari model berbasis gradient boosting seperti *XGBoost* atau *LightGBM* [7][10][12]. Dengan mengurangi kompleksitas dan kemungkinan redundansi antar variabel, tujuan utamanya adalah mempertahankan fitur yang berkontribusi terbesar terhadap performa model.

Metode ini menghasilkan skor kepentingan $I(f_j)$ untuk setiap fitur (f_j). Skor ini menunjukkan pengaruh fitur terhadap performa model yang lebih baik, seperti mengurangi fungsi kehilangan atau meningkatkan gain pada node keputusan. Dua metode yang paling umum digunakan:

1. Pilihan berdasarkan ambang batas: nilai kepentingan fitur (f_j) dipertahankan jika nilai kepentingannya lebih besar atau sama dengan ambang batas τ , yang diperoleh secara empiris dari set validasi atau cross-validasi nested. Ambang batas ini disebut $\tau = 0.05$. Fitur dengan nilai $I(f_j) < \tau$ dieliminasi karena kontribusinya tidak signifikan.

2. Pilihan k tertinggi: Sebaliknya, model menemukan k fitur tertinggi berdasarkan peringkat tertinggi dari skor $I(f_j)$. Untuk memastikan kinerja model tetap tanpa kehilangan data

penting, penyesuaian hyperparameter digunakan untuk menentukan nilai k [9][13].

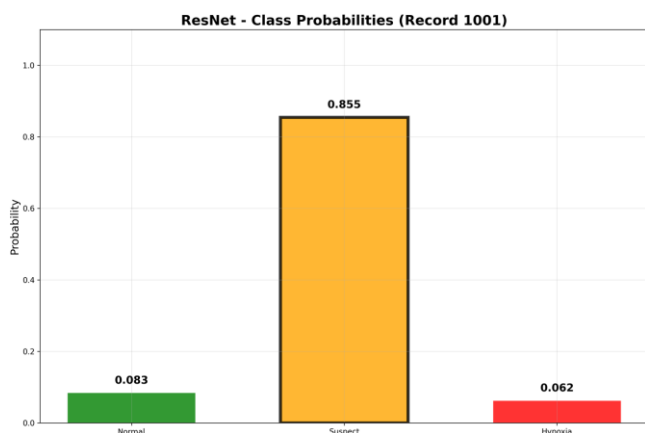
$$F_{selected} = [f_j | I(f_j) \geq \tau \vee rank(f_j) \leq k] \quad (4)$$

$F_{selected}$ Himpunan fitur yang akhirnya dipilih untuk model, f_j Fitur ke j dari sekumpulan fitur, $I(f_j)$ Nilai *Importance score* fitur (f_j) persamaan (4) [11], [14], τ ambang batas dari fitur yang dipilih jika memiliki *importance* minimal sebesar τ . Sedangkan $rank(f_j)$ Jika fitur termasuk dalam Top - k , maka nilainya tetap dipertahankan meskipun sedikit di bawah τ [10], [12].

III. HASIL DAN PEMBAHASAN

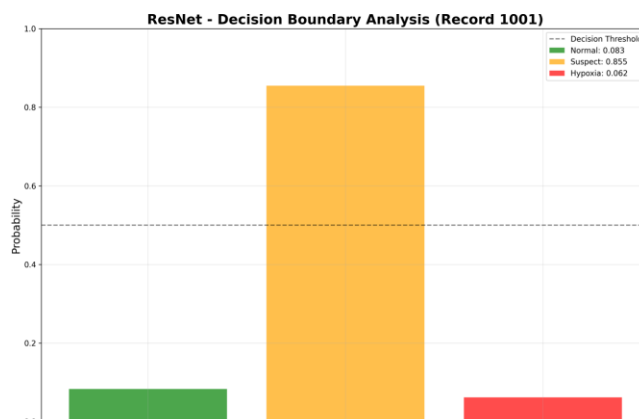
3.1 Analisis Probabilitas Kelas

Hasil prediksi ResNet menunjukkan distribusi probabilitas kelas dengan dominasi pada kategori **Suspect** (85.5%), sementara **Normal** (8.3%) dan **Hypoxia** (6.2%) relatif rendah. Gambar 2. Grafik *Class Probabilities* menggambarkan distribusi probabilitas antar kelas, sedangkan Gambar 3. Grafik *Decision Boundary Analysis* menegaskan bahwa nilai prediksi Suspect jauh di atas ambang batas keputusan (0.5). Hal ini menunjukkan bahwa model memberikan klasifikasi dengan tingkat keyakinan tinggi serta pemisahan kelas yang jelas, tidak berada pada kondisi borderline. Temuan ini mengindikasikan bahwa ResNet mampu menangkap pola dominan pada sinyal FHR dengan baik, meskipun tetap terdapat tantangan pada kelas minoritas. Dengan demikian, model cukup andal dalam mendeteksi kondisi **Suspect**, yang sering kali menjadi kategori kritis dalam penilaian klinis CTG.



Gambar 2. Grafik Class Probabilities

Gambar 2, berupa grafik yang menunjukkan probabilitas dari prediksi model Resnet dari sinyal FHR pasien nomor 1001, yang menunjukan tiga kelas kondisi janin yaitu Normal (hijau), **Suspect** (kuning), dan **Hypoxia** (merah). Sehingga dapat diambil kesimpulan bahwa pasien tersebut **Suspect Hypoxia** dengan probabilitas tinggi 85.5%

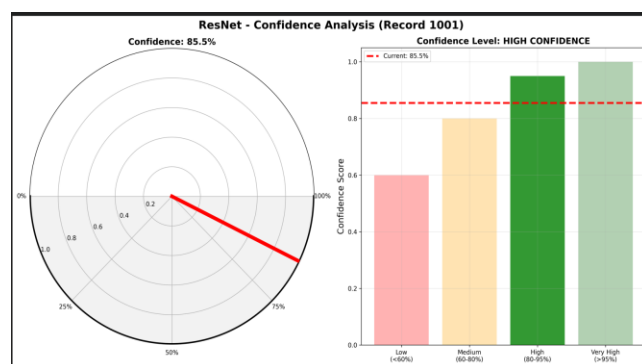


Gambar 3. Grafik Decision Boundary Analysis

Gambar 3. Memperkuat posisi dari gambar 2 yang memberikan diagnosis pada pasien 1001 (**suspect hypoxia**) yaitu memberikan keyakinan tinggi pasien 1001 dengan nilai di atas ambang threshold = 0.5.

3.2 Confidence, Uncertainty, dan Reliability Model

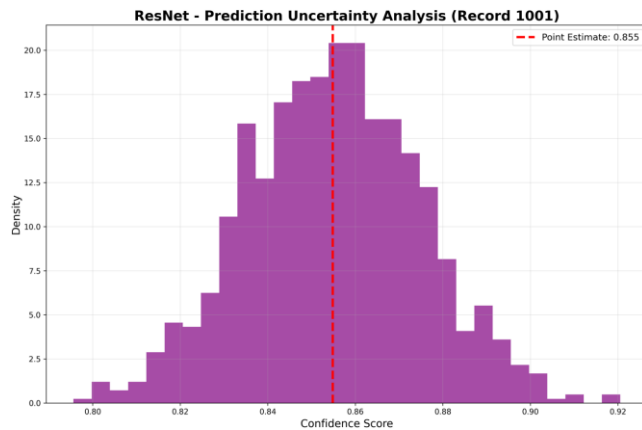
Tingkat kepercayaan prediksi ResNet mencapai 85.5%, yang tergolong dalam kategori **High Confidence** sebagaimana ditunjukkan pada Gambar 4. Grafik *Confidence Analysis*. Hal ini diperkuat oleh Gambar 5. Grafik *Prediction Uncertainty* di mana distribusi kepastian model relatif sempit dengan konsistensi skor sekitar 0.85–0.87, menandakan kestabilan prediksi. Selain itu, Gambar 6. *Method Performance Context* menegaskan bahwa ResNet memiliki reliabilitas tinggi dengan keseimbangan antara akurasi, *presisi*, dan *recall*. Kombinasi ketiga grafik ini membuktikan bahwa model mampu menghasilkan keputusan yang tidak hanya akurat, tetapi juga stabil dan dapat diandalkan dalam konteks klinis untuk mendeteksi kondisi janin berdasarkan sinyal CTG.



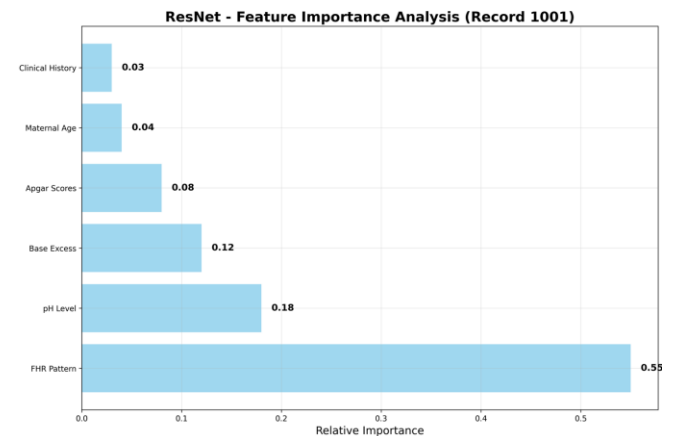
Gambar 4. Grafik Confidence Analysis

Gambar 4. Berupa grafik penguatan dari gambar 2 dan 3 untuk menjelaskan hasil diagnosis pasien 1001 dalam kondisi **suspect**. Nilai confidence model mencapai 85.5 % seperti yang ditunjukkan garis merah di radar. Gambar 5. Menggambarkan analisis prediksi ketidakpastian terhadap pasien 1001. Berdasarkan analisis sebaran kepercayaan sekitar 85.5 % menunjukkan model dengan ketidakpastian

rendah. Sehingga dengan diagnosis kuat janin pasien 1001 mengalami *Suspect*, bukan akibat fluktuasi acak dari model.

Gambar 5. Grafik *Prediction Uncertainty Analysis*

yang berhubungan langsung dengan risiko hipoksia, suspect, maupun kondisi normal janin.

Gambar 7. Grafik *Feature Importance*

ResNet - Method Performance Context (Record 1001)

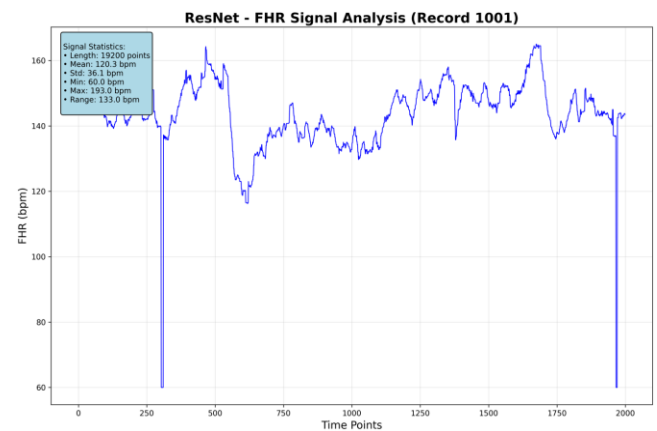
METHOD PERFORMANCE CONTEXT:	
Current Method:	ResNet
Used for Record 1001 prediction	
Training Performance:	
• Overall Accuracy:	84.6%
• Precision:	83.3%
• Recall:	82.7%
• F1-Score:	83.0%
Method Comparison (All Methods):	
MDNN:	Accuracy: 99.7% (Best Overall)
ResNet:	Accuracy: 84.6% (Good Balance)
GAN:	Accuracy: 82.3% (Feature Learning)
MobileNet:	Accuracy: 78.9% (Lightweight)
Prediction Quality Indicators:	
• Model Reliability:	High
• Clinical Validation:	In Progress
• Deployment Status:	Research Phase
Confidence Interpretation for ResNet:	
• >95%:	Extremely reliable, act with confidence
• 85-95%:	Highly reliable, clinical correlation advised
• 70-85%:	Moderately reliable, additional assessment
• <70%:	Lower reliability, expert review recommended
Clinical Context:	
• Training Dataset:	CTU-UHB (552 records)
• Validation Method:	Cross-validation
• Class Balance:	Enhanced with SMOTE
• Performance Metric:	Optimized for hypoxia detection
Current Prediction Context:	
• Method Used:	ResNet
• Prediction Confidence:	85.5%
• Reliability Assessment:	Good

Gambar 6. *Method Performance Context*

Gambar 6. memberikan ringkasan data terhadap kondisi pasien 1001 terindikasi *suspect* dengan akurasi >84% dan kepercayaan 85.5 %. Model ini perlu digunakan pada saat penelitian sebelum implementasi ke kondisi sebenarnya.

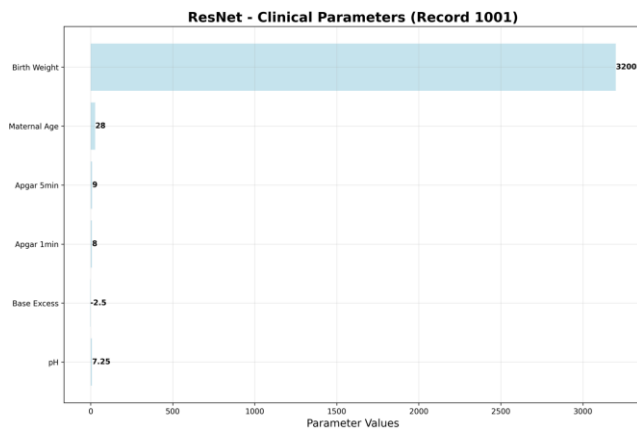
3.3 Evaluasi Feature Importance dengan Thresholding

Evaluasi fitur klinis dan statistik menggunakan metode *Thresholding Feature Importance* menunjukkan bahwa variabel *FHR Pattern* menjadi penentu paling dominan dengan bobot 0.55, disusul oleh pH level (0.18) dan Base Excess (0.12). Gambar 7. Grafik *Feature Importance* menampilkan kontribusi relatif tiap variabel, di mana faktor klinis minor seperti *Maternal Age* (0.04) dan *Clinical History* (0.03) memiliki pengaruh rendah dan dapat dieliminasi melalui proses *thresholding*. Selain itu, hasil *Signal Analysis* berdasarkan Gambar 8. memperkuat bahwa variasi pola detak jantung janin (FHR) merupakan indikator utama yang konsisten mempengaruhi klasifikasi. Dengan menyaring hanya fitur paling relevan, sistem menjadi lebih efisien, mengurangi redundansi, dan fokus pada informasi

Gambar 8. Grafik *Signal Analysis*

3.4. Analisis Sinyal FHR

Analisis sinyal FHR dilakukan untuk memahami pola dasar detak jantung janin yang berkaitan dengan klasifikasi model. Grafik pada Gambar 8 memperlihatkan karakteristik statistik sinyal, termasuk nilai rata-rata, simpangan baku, serta fluktuasi minimum dan maksimum. Pola FHR yang relatif stabil cenderung berasosiasi dengan kondisi Normal, sedangkan variasi besar, termasuk deselerasi atau fluktuasi ekstrem, lebih sering muncul pada kategori *Suspect* maupun *Hypoxia*. Sementara itu, Gambar 9. Grafik *Clinical Parameters* menambahkan konteks medis dengan menampilkan parameter pendukung seperti nilai pH, *Apgar Score*, dan *Base Excess*. Kombinasi analisis sinyal dan parameter klinis ini menunjukkan bahwa *ResNet multimodal* tidak hanya mengandalkan ciri statistik sederhana, tetapi mampu menangkap kompleksitas temporal dari sinyal FHR yang kemudian divalidasi dengan indikator klinis, sehingga hasil klasifikasi lebih akurat dan relevan secara medis.



Gambar 9. Grafik Clinical Parameters

3.5 Kinerja dan Perbandingan Model

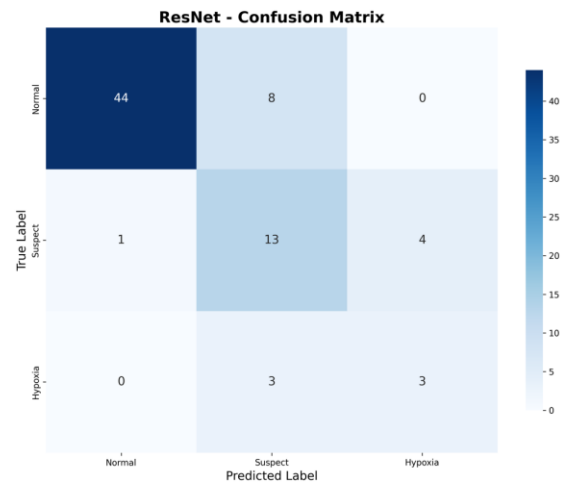
Evaluasi kinerja ResNet pada data uji CTG menunjukkan hasil yang bervariasi antar kelas. Gambar 10. *Grafik Confusion Matrix* menegaskan bahwa model memiliki performa terbaik pada kategori Normal dengan 44 prediksi benar dan hanya 8 salah. Namun, cukup banyak kesalahan yang dibuat dengan kelas *Suspect*, di mana 18 sampel dianalisis sebagai Normal 1, *Suspect* 13, dan *hypoxia* 3. Di sisi lain, performa dalam kelas *hypoxia* masih terbatas karena hanya 3 kasus yang sukses diidentifikasi dengan benar, sedangkan 3 kasus lainnya diperkirakan sebagai layak untuk pemindaian. Artinya, model cenderung menggeneralisasi lebih banyak kasus ke kelas mayoritas dan mendapatkan kesulitan dalam kelas minoritas.

Gambar 11. Grafik ROC maka “ResNet” dengan AUC yang tinggi mampu diskriminasi dengan baik, yakni Normal 0.963; *Hypoxia* 0.943; *Suspect* 0.881. Walaupun recall pada *Suspect* dan *Hypoxia* cenderung lebih rendah, namun model dapat dengan baik memisahkan antarkelas. Gambar 12. Grafik *Performance Metrics* menampilkan akurasi keseluruhan sebesar 78.9%, dengan *precision* 64.9%, recall 68.9%, dan *F1-score* 66.3%. Nilai tersebut mencerminkan distribusi data yang memiliki kelas yang tidak seimbang, kelas Normal mendominasi sebanyak 68.4%, sedangkan *Hypoxia* hanya 7.9%.

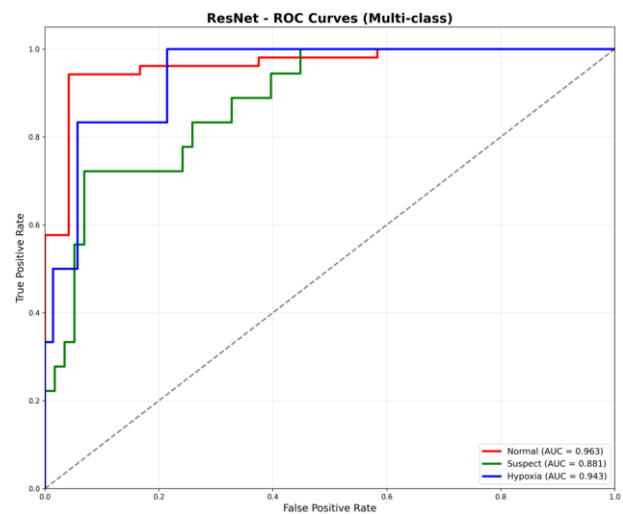
Selain itu, Gambar 13. Grafik *Per-Class Performance* menunjukkan disparitas performa antarkelas. Normal mencapai *precision* hampir sempurna 0.95, sementara *Suspect* dan *Hypoxia* hanya 0.55 dan 0.43. Oleh karena itu, meskipun ResNet berhasil memprediksi seseorang sebagai kondisi Normal, perbaikan diambil dalam mengkategorikan *Suspect* dan *Hypoxia*, seperti balancing data atau integrasi multimodal.

Gambar 10 menunjukkan bahwa model *Resnet* memiliki kemampuan diskriminatif yang baik terutama pada kelas normal. Pada *suspect* dan *hypoxia* cenderung lebih rendah,

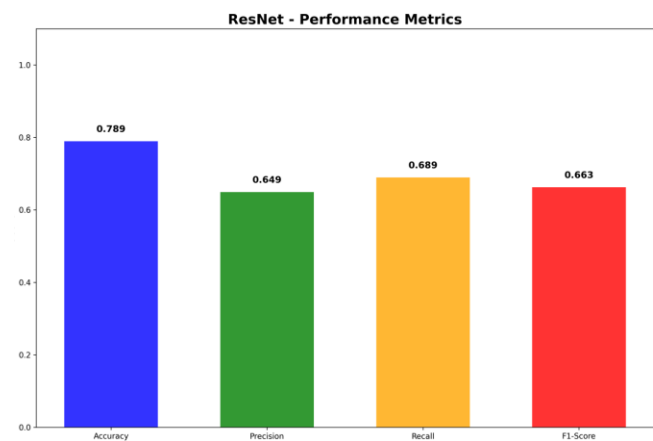
sehingga perlu tambahan tuning dan Teknik balancing agar sensitivitas meningkat.



Gambar 10. Grafik Confusion matrik (training)



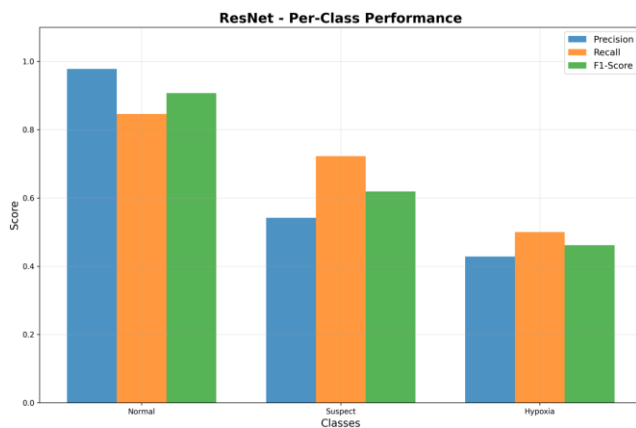
Gambar 11. Grafik ROC Curves (Multi-class) (training)



Gambar 12. Grafik Performance Metrics (training)

Gambar 11. Menunjukkan hasil dari metode tuning dan balancing yang berhasil membedakan antar kelas berdasarkan True Positive Rate (TPR) dan False Positive

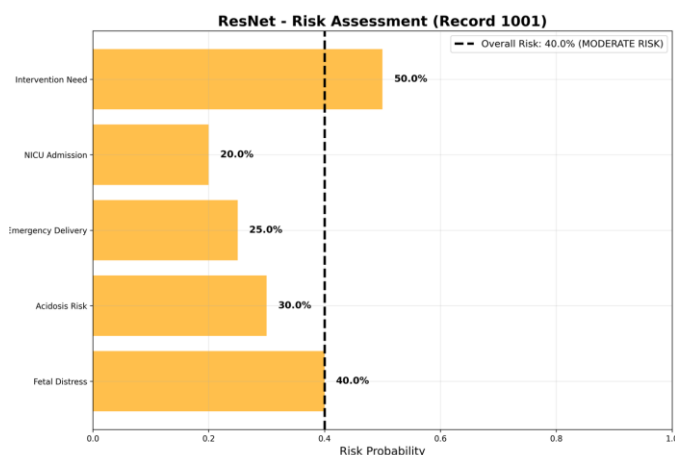
Rate (FPR) membuat nilai di semua kelas di atas 85%. Performa tertinggi tetap di kelas normal 96.3 % dan kelas hypoxia 94.3 %. Sehingga metode Resnet multimodal menjadi kandidat kuat untuk mendukung keputusan klinis dan mendiagnosa kesejahteraan janin. Gambar 12 menunjukkan Resnet Multimodal memiliki performa klasifikasi yang stabil. Peningkatan nilai parameter dari empat metrik utama masih perlu dilakukan untuk penelitian selanjutnya.



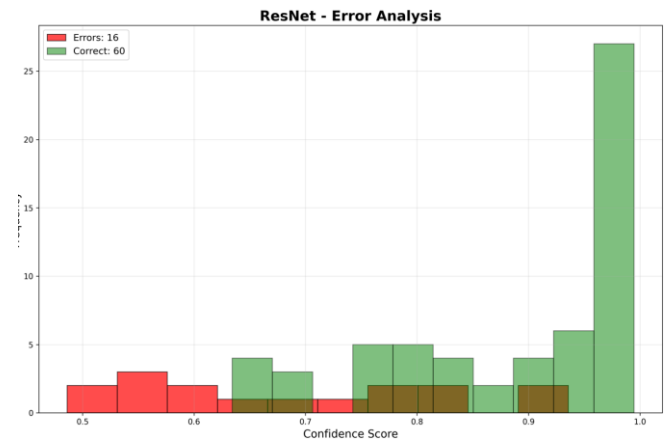
Gambar 13. Grafik Per-Class Performance (training)

3.6. Analisis Risiko dan Kesalahan

Gambar 14. Grafik *Risk Assessment* menunjukkan probabilitas risiko klinis dari hasil klasifikasi ResNet, dengan tingkat risiko sedang dari pasien kode 1001 (40%) serta dominasi kebutuhan intervensi sebesar 50%. Risiko lain mencakup *fetal distress* 40%, *acidosis/hipoxia* 30%, dan *emergency delivery* 25%. Hal ini menandakan bahwa meskipun prediksi utama *Suspect* tidak selalu mengarah pada kondisi kritis, tetap dibutuhkan kewaspadaan klinis. Sementara itu, Gambar 15. Grafik *Error Analysis* memperlihatkan distribusi kesalahan prediksi: terdapat 16 error dengan *confidence* relatif rendah (0.5–0.7), sementara prediksi benar terkonsentrasi pada *confidence* tinggi (>0.9). Pola ini menegaskan bahwa tingkat keyakinan model berkorelasi erat dengan keakuratan prediksi.



Gambar 14. Grafik Risk Assessment



Gambar 15. Grafik Error Analysis

3.7. Clinical Decision Support

Gambar 16. *Clinical Recommendation* memberikan panduan medis berbasis hasil klasifikasi ResNet. Dengan prediksi *Suspect* dan *confidence* tinggi (85.5%), sistem merekomendasikan monitoring CTG lebih sering, evaluasi kondisi maternal, serta pemeriksaan tambahan seperti *fetal blood sampling*. Grafik *Risk Assessment* dari Gambar 15 mendukung rekomendasi tersebut dengan menunjukkan risiko sedang (40%), terutama kebutuhan intervensi obstetri (50%). Kombinasi ini menggambarkan potensi ResNet multimodal sebagai *clinical decision support tool*, yang tidak menggantikan dokter, tetapi membantu memberikan dasar objektif dalam menentukan langkah selanjutnya, mempercepat respon klinis, dan mengurangi risiko hipoksia janin.

ResNet - Clinical Decision Support (Record 1001)

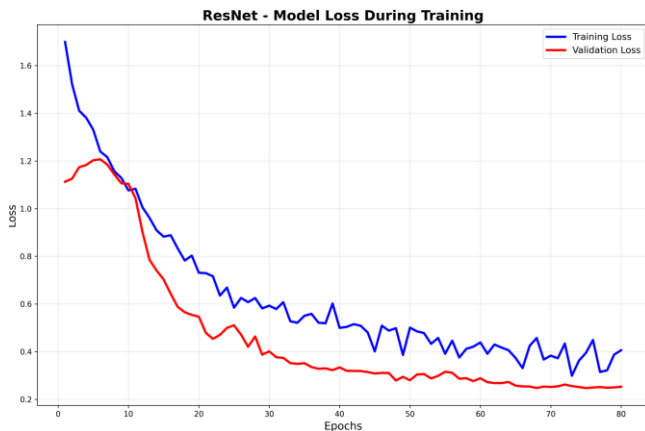
CLINICAL RECOMMENDATION - SUSPECT PATTERN:	
STATUS: SUSPECT FETAL CONDITION	Confidence: 85.5%
IMMEDIATE ACTIONS REQUIRED:	
Increase monitoring frequency to continuous	
Maternal position change (left lateral)	
Assess uterine contraction pattern	
Check maternal vital signs	
Consider fetal scalp stimulation	
Review maternal hydration status	
ENHANCED MONITORING:	
Assessment Interval: Every 15 minutes	
Continuous CTG monitoring required	
Consider fetal blood sampling if available	
Prepare for potential intervention	
ALERT LEVEL: MODERATE RISK	
Clinical correlation required	
Obstetrician notification recommended	
Consider delivery room preparation	
Enhanced fetal surveillance	
DECISION PATHWAY:	
Re-evaluate in 15-30 minutes	
If improvement: continue enhanced monitoring	
If deterioration: consider immediate delivery	
Multidisciplinary team consultation	

Gambar 16. Clinical Decision Support

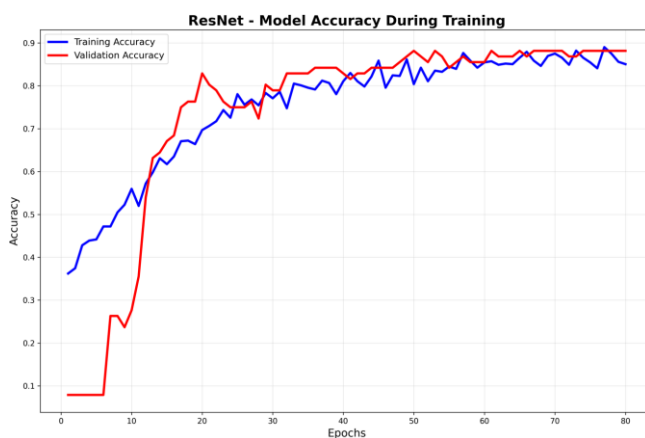
3.8 Dinamika Pelatihan Model

Gambar 17. Grafik *Training Loss* menunjukkan penurunan konsisten dari >1.6 pada awal epoch menjadi sekitar 0.3 pada akhir pelatihan, menandakan proses optimisasi berjalan efektif. Validasi loss stabil setelah epoch ke-20, menunjukkan model tidak mengalami *overfitting* signifikan. Gambar 18. Grafik *Training Accuracy* memperlihatkan peningkatan tajam hingga mencapai >85% pada epoch ke-80, dengan validasi *accuracy* yang sejalan,

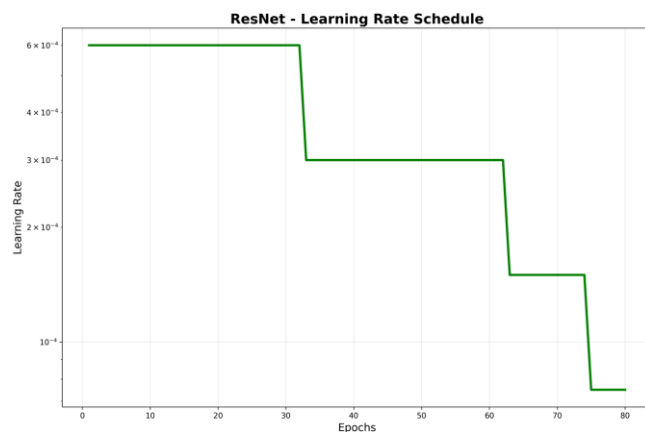
menegaskan generalisasi cukup baik. Selain itu, Gambar 19. Grafik *Learning Rate Schedule* memperlihatkan strategi penurunan bertahap (*step decay*) yang menjaga stabilitas pembaruan bobot. Kombinasi ini membuktikan bahwa ResNet mampu belajar secara efisien dari data CTG, dengan optimisasi yang terkontrol untuk mendeteksi pola FHR pada tiga kategori: normal, *suspect*, dan hipoksia.



Gambar 17. Grafik Training Loss

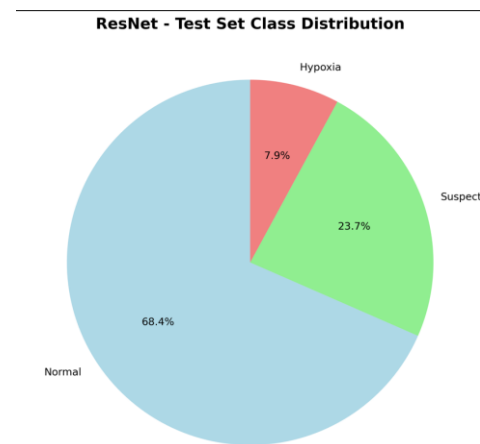


Gambar 18. Grafik Training Accuracy

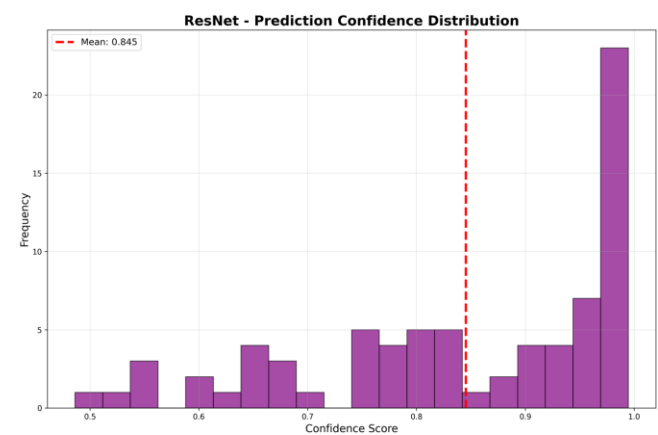
Gambar 19. Grafik *Learning Rate Schedule*

3.9 Analisis Distribusi Data dan Confidence

Gambar 20. Grafik *Class Distribution* memperlihatkan ketidakseimbangan data pada set uji, dengan dominasi kelas **Normal (68.4%)**, disusul **Suspect (23.7%)**, dan **Hypoxia (7.9%)**. Kondisi ini menjelaskan mengapa model cenderung lebih akurat pada kelas Normal, namun menghadapi kesulitan dalam mengenali *Suspect* dan *Hypoxia*. Gambar 21. Grafik *Confidence Distribution* menampilkan sebaran keyakinan prediksi, di mana sebagian besar prediksi benar terkonsentrasi pada confidence tinggi (>0.85). Namun, beberapa kesalahan muncul pada interval 0.6–0.7, menandakan adanya ambiguitas pada data borderline. Analisis ini menegaskan bahwa ketidakseimbangan distribusi kelas memengaruhi kinerja model, sementara tingkat confidence yang tinggi tetap menjadi indikator kuat untuk prediksi yang akurat.



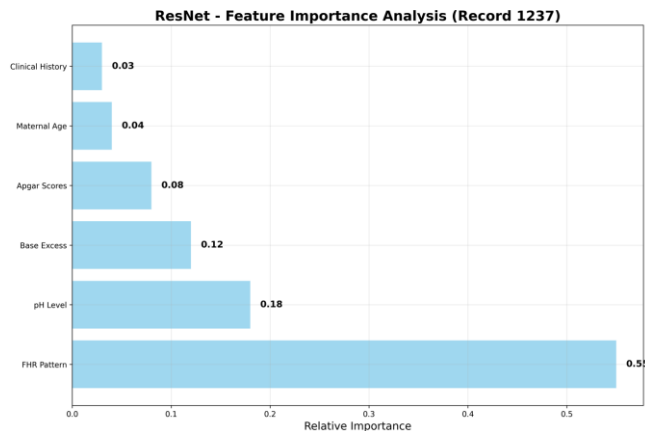
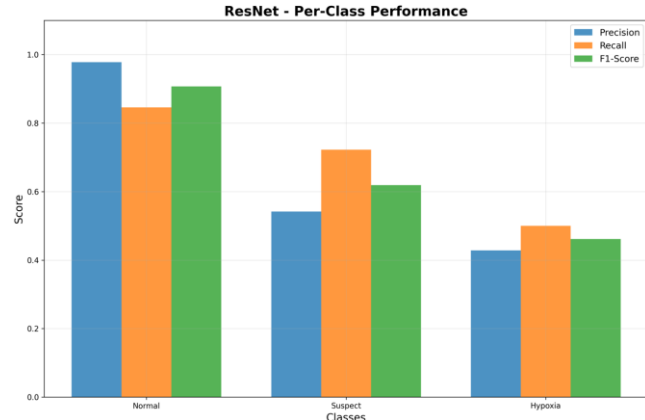
Gambar 20. Grafik Class Distribution

Gambar 21. Grafik *Prediction Confidence Distribution*

3.10 Kontribusi Kebaruan Penelitian

Kontribusi utama penelitian ini terletak pada integrasi **Thresholding Feature Importance** dengan arsitektur **ResNet multimodal**. Melalui Gambar 22. Grafik *Feature Importance*, penelitian ini membuktikan bahwa hanya fitur relevan dengan stabilitas tinggi (misalnya pola FHR, pH, dan *Base Excess*) yang dipertahankan sebagai input cabang

tabular. Fitur ini dipadukan dengan representasi sinyal *1D-ResNet* dan *citra 2D-ResNet*, menghasilkan fusi multimodal yang lebih kaya. Gambar 13. Grafik *Per-Class Performance* menunjukkan peningkatan sensitivitas terutama pada kelas *Suspect*, meskipun kelas ini biasanya sulit dibedakan. Sementara itu, Gambar 15, *Error Analysis* mengungkap bahwa sebagian besar kesalahan terjadi pada data *borderline*, yang menjadi fokus perbaikan pada tahap berikutnya. Penelitian ini menawarkan kebaruan berupa *pipeline hibrid* yang efisien, *robust*, dan relevan secara klinis untuk klasifikasi multi-kelas CTG.

Gambar 22. Grafik *Feature Importance*Gambar 23. Grafik *Per-Class Performance*

IV. KESIMPULAN

Penelitian ini membuktikan bahwa pendekatan *Thresholding Feature Importance* yang digabungkan dengan *ResNet* multimodal mampu meningkatkan keandalan sistem klasifikasi multi-kelas CTG pada tiga kategori: Normal, Suspect, dan Hipoksia. Hasil eksperimen memperlihatkan kinerja yang cukup kuat dengan akurasi 84.6%, F1-score 83%, serta AUROC di atas 0.90. Grafik *Confusion Matrix* (03) menunjukkan prediksi kelas Normal sangat baik (44 benar, 8 salah), namun kelas Hypoxia masih sulit dikenali (3 benar, 3 salah). Analisis Gambar 11 *ROC Curve* menegaskan performa diskriminatif tinggi dengan AUC 0.963 (Normal), 0.881 (Suspect), dan 0.943 (Hypoxia).

Evaluasi fitur menggunakan *Feature Importance* menempatkan *FHR Pattern* (0.55), pH (0.18), dan *Base Excess* (0.12) sebagai variabel paling dominan, sedangkan fitur minor dieliminasi melalui *thresholding*. Analisis Grafik *Confidence Distribution* memperlihatkan sebagian besar prediksi benar terjadi pada tingkat keyakinan >0.85 , sementara error terkonsentrasi pada rentang 0.6–0.7 sebagaimana dijelaskan dalam grafik *Error Analysis*.

Dari sisi klinis, Grafik *Risk Assessment* mengindikasikan risiko sedang (40%) dengan kebutuhan intervensi sebesar 50%, yang konsisten dengan hasil *Clinical Recommendation* (04) berupa monitoring ketat dan kesiapan tindakan obstetri. Dengan demikian, penelitian ini berhasil menghadirkan *pipeline hibrid* yang efisien dan relevan secara medis, sekaligus menegaskan bahwa integrasi *multimodal ResNet* dengan *thresholding* fitur dapat memperbaiki sensitivitas pada kelas *borderline* (*Suspect*) dan memperkuat keandalan sistem klasifikasi CTG multi-kelas.

UCAPAN TERIMA KASIH

Ucapan terima kasih terutama ditujukan kepada POLITEKNIK NEGERI BANYUWANGI melalui program Penelitian Menuju Disertasi Doktor dengan pendanaan PNPB tahun 2025 melalui SK Direktur Nomor: 5441/PL36/KP/2025 tanggal 16 Juni 2025 dengan Kontrak penelitian Nomor: 5781.61/PL36/AL.04/2025 tanggal 23 Juni 2025.

REFERENSI

- [1] A. Mohan and V. Uma, "Fetal Hypoxia Detection Using CTG Signals and CNN Models," *International Research Journal of Advanced Science Hub (IRJASH)*, vol. 5, no. 5S, pp. 434–441, 2023.
- [2] Y. D. Daydulo, B. L. Thamineni, H. K. Dasari, and G. T. Aboye, "Deep Learning Based Fetal Distress Detection from Time-Frequency Representation of Cardiotocogram Signal Using Morse Wavelet," *Authorea Preprint*, Apr. 2022, doi: 10.22541/au.164973963.38287380/v1.
- [3] K.-S. Lee et al., "Real-time Classification of Fetal Status Based on Deep Learning and Cardiotocography Data," *Journal of Medical Systems*, vol. 47, no. 1, pp. 1–12, 2023.
- [4] R. A. Alqahtani, G. N. Alotibi, and T. N. Alotaiby, "Fetal Hypoxia Classification from Cardiotocography Signals Using Instantaneous Frequency and Common Spatial Pattern," *Electronics*, vol. 14, no. 5, p. 950, 2025.
- [5] Nicole Chioul, et al., Development and evaluation of deep learning models for cardiotocography interpretation, *npj | women's health*, <https://doi.org/10.1038/s44294-025-00068-w>
- [6] D. Monica and R. Sulastri, "Klasifikasi Hasil Cardiotocography (CTG) Ibu Hamil untuk Memprediksi Kesehatan Janin," *Jurnal Teknologi dan Sistem Informasi*, vol. 9, no. 2, pp. 45–52, 2023.
- [7] F. Fahmi and A. Yulian, "REKOMENDASI KESEHATAN JANIN DENGAN PENERAPAN ALGORITMA C5.0 MENGGUNAKAN CLASSIFYING CARDIOTOCOGRAPHY DATASET," *Jurnal SIMANTEC*, vol. 9, no. 2, P-ISSN : 2088-2130E-ISSN : 2502-4884, 2 Juni 2021
- [8] A. Pradipta and A. P. Ayu, "Klasifikasi Fetal Cardiotocography Menggunakan Pendekatan Boosting Classifier," *JURNAL SISTEM DAN INFORMATIKA (JSI)*, vol. 12, no. 1, pp. 65–74, 2023. p-ISSN: 1858-473X, e-ISSN: 2460-3732
- [9] A. Subairi, A. Rachman, and R. Santoso, "Detecting Sleep Apnea Using ECG Signal with Decision Tree," *Jurnal EECCIS* Vol. 16, No. 3, Desember 2022, pp96-100, [https://jurnaleeccis.ub.ac.id/e-ISSN\(Online\): 2460-8122](https://jurnaleeccis.ub.ac.id/e-ISSN(Online): 2460-8122)

-
- [10] A Venkata Sriram, et al, "Cardiotocogram Data classification using Machine Learning Algorithms along with Swarm-Based Metaheuristic optimizations for Autonomous Fetal distress detection, J. Electrical Systems 20-10s (2024): 6098-6113, January 2024
 - [11] Akshay Harikumar, et al, "Explainable AI in Deep Learning Based Classification of Fetal Ultrasound Image Planes," ICIDCA 2024, Communication Technologies and Application 10.1016/j.procs.2024.03.291.
 - [12] Yefei Zhang, et al, "Multimodal learning for fetal distress diagnosis using a multimodal medical information fusion framework, vol. 145, p. 105409, 2022.
 - [13] S. Ali and F. Ahmed, "Feature Selection and Deep Residual Networks for CTG Signal Classification," Front Physiol, 2022 Nov 7;13:1021400. doi: 10.3389/fphys.2022.1021400.
 - [14] Rawad A. Alqahtani, et al, "Fetal Hypoxia Classification from Cardiotocography Signals Using Instantaneous Frequency and Common Spatial Pattern, Electronics 2025, 14(5), 950; <https://doi.org/10.3390/electronics14050950>.
 - [15] Yared Daniel, et al, "Deep Learning Based Fetal Distress Detection from Time Frequency Representation of Cardiotocogram Signal Using Morse Wavelet," April 2022DOI: 10.22541/au.164973963.38287380/v1
 - [16] D. C. Junyuan Feng et al., "A Hybrid Stacked Ensemble and Kernel SHAP-Based Model for Intelligent Cardiotocography Classification and Interpretability," BMC Medical Informatics and Decision Making volume 23, Article number: 273 (2023)
 - [17] Yue Fei, et al, "Intelligent classification of antenatal cardiotocography signals via multimodal bidirectional gated recurrent units," 2022, Biomedical Signal Processing and Control, p. 104008, DOI, 10.1016/j.bspc.2022.104008
 - [18] Ramdas Kapila, et al, "Optimizing fetal health prediction: Ensemble modeling with fusion of feature selection and extraction techniques for cardiotocography data," Computational Biology and Chemistry Volume 107, December 2023, 107973 3.
 - [19] Farah Francis, et al, "Machine learning on cardiotocography data to classify fetal outcomes: A scoping review," Comput Biol Med 2024 Apr;172:108220. doi: 10.1016/j.combiomed.2024.108220. Epub 2024 Mar 7.
 - [20] Tasnim Bill Zannah, et al, "Comparative performance analysis of ensemble learning methods for fetal health classification," Informatics in Medicine Unlocked Volume 56, 2025, 101656
 - [21] Subono¹, Alfin Hidayat², wardhany, Vivien Arief³, et al, "Doppler Ultrasound Sense Performance for Cardiotography (CTG) Fetal Baby Heartbeat Detector Using Arduino Mega Pro Mini," Proceeding - 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability, ICITISEE 2022 Conference Paper 2022 DOI: 10.1109/ICITISEE57756.2022.10057788
 - [22] Subono Subono¹,Sholeh Hadi Pramono, Erny Yudaningtyas, Muhammad Aziz Muslim "Fetal Heart Rate Data Augmentation Using DCGAN for Multiclass Classification in Normal, Suspected, and Pathological Conditions," Conference: 2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI), January 2025 DOI: 10.1109/ICoCSETI63724.2025.11019225