

Analisis Kinerja Multimodal Dense Neural Network untuk Deteksi Hipoksia Janin pada Dataset Tidak Seimbang

¹Subono Subono, ²Dianni Yusuf, ³Sholeh Hadi Pramono, ⁴Erni Yudaningtyas, ⁵M. Aziz Muslim, ⁶Mahdin Rohmatillah, ⁷Cries Avian

¹ Teknologi Rekayasa Komputer, Politeknik Negeri Banyuwangi, Banyuwangi

² Teknologi Rekayasa Perangkat Lunak, Politeknik Negeri Banyuwangi, Banyuwangi

^{3,4,5,6,7} Teknik Elektro, Universitas Brawijaya, Malang

¹Subono@Poliwangi.ac.id, ²dianniyusuf@poliwangi.ac.id, sholehpramono@ub.ac.id, erni@ub.ac.id, muh_aziz@ub.ac.id, mahdin94@ub.ac.id, cries.avian@ub.ac.id

Abstract - This study aims to develop a Multimodal Dense Neural Network (MDNN) for detecting fetal hypoxia using an imbalanced *Cardiotocography* (CTG) dataset. The main challenges in fetal hypoxia diagnosis are the imbalance between *Normal*, *Suspect*, and *Hypoxia* categories and the limited interpretability of conventional deep learning models. To address these issues, a robust preprocessing pipeline was designed, consisting of *Physiological Clipping* (50–200 bpm), *Median Absolute Deviation* (MAD) normalization, *SMOTETomek balancing*, and *Gaussian noise augmentation*. The MDNN architecture integrates two parallel branches: Fetal Heart Rate (FHR) signals and clinical parameters (pH, *Apgar score*, and *base deficit*), fused through a *Dense Fusion Layer* to generate compact multimodal representations. The proposed MDNN achieved 99.7% accuracy, 99.5% F1-Score, and 0.993 AUC, outperforming CNN (84.6%), ResNet18 (82.3%), and MLP (87.5%). The *confusion matrix* indicated good generalization performance with per-class accuracies of 69% (*Normal*), 56% (*Suspect*), and 67% (*Hypoxia*). *SHAP feature importance* analysis identified FHR pattern (0.45) and pH level (0.25) as the most influential features in classification. These findings confirm that the proposed MDNN is robust, computationally efficient, and clinically interpretable, making it a promising framework for real-time fetal hypoxia diagnosis in modern clinical environments. These findings confirm that the proposed MDNN is robust, computationally efficient, and clinically interpretable, making it a promising framework for real-time fetal hypoxia diagnosis. This Study extends the application of multimodal deep learning in biomedical signal processing for fetal monitoring systems

Keywords — *Fetal hypoxia, Cardiotocography, MDNN, Robust preprocessing, SHAP*

Abstrak— Penelitian ini bertujuan mengembangkan sistem Multimodal Dense Neural Network (MDNN) untuk deteksi hipoksia janin pada dataset *Cardiotocography* (CTG) yang tidak seimbang. Masalah utama pada diagnosis hipoksia adalah ketidakseimbangan data klinis antara kelas *Normal*, *Suspect*, dan *Hypoxia*, serta rendahnya interpretabilitas model *deep learning* konvensional. Untuk mengatasi hal tersebut, dirancang pipeline robust preprocessing yang terdiri atas *Physiological Clipping* (50–200 bpm), normalisasi berbasis *Median Absolute Deviation* (MAD), kombinasi *SMOTETomek balancing*, dan *Gaussian noise augmentation*. Arsitektur MDNN mengintegrasikan dua cabang paralel, yaitu sinyal *Fetal Heart Rate* (FHR) dan parameter klinis (pH, *Apgar score*, dan *base deficit*), yang digabungkan melalui *Dense Fusion Layer*. Hasil eksperimen menunjukkan model

MDNN menghasilkan akurasi 99.7%, F1-score 99.5%, dan AUC 0.993, melampaui kinerja CNN (84.6%), ResNet18 (82.3%), dan MLP (87.5%). *Confusion matrix* menunjukkan kemampuan generalisasi yang baik dengan akurasi per kelas 69% (*Normal*), 56% (*Suspect*), dan 67% (*Hypoxia*). Analisis *SHAP feature importance* mengidentifikasi pola FHR (0.45) dan pH darah (0.25) sebagai fitur paling berpengaruh. Hasil ini menunjukkan bahwa MDNN memiliki robustness tinggi, efisiensi komputasi baik, dan interpretabilitas klinis kuat, sehingga layak dikembangkan sebagai sistem diagnosis hipoksia janin berbasis kecerdasan buatan di lingkungan klinis modern. Hasil penelitian ini menegaskan bahwa MDNN yang diusulkan bersifat Robust, efisien secara komputasi, dan dapat diinterpretasikan secara klinis, menjadikannya kerangka kerja yang menjanjikan untuk diagnosis hipoksia janin secara *real time*. Penelitian ini memperluas penerapan pembelajaran mendalam multimoda dalam proses sinyal biomedis untuk sistem pemantauan janin.

Kata Kunci— *Hipoksia janin, Cardiotocography, MDNN, Robust preprocessing, SHAP*

I. Pendahuluan

Kekurangan oksigen selama kehamilan atau persalinan dikenal sebagai hipoksia janin. Ini dapat menyebabkan asfiksia, gangguan pernapasan, dan, dalam kasus yang parah, menyebabkan *hypoxic-ischemic encephalopathy* (HIE), yang menyebabkan kerusakan neurologis permanen[1]. Karena kondisi ini termasuk salah satu penyebab utama morbiditas dan mortalitas perinatal, deteksi dini sangat penting untuk mencegah komplikasi. *Cardiotocography* (CTG), yang merekam sinyal denyut nadi janin (FHR) dan kontraksi uterus ibu (UC) secara bersamaan, biasanya digunakan untuk memantau kondisi janin[2], [3]. Dalam praktik klinis, interpretasi manual hasil CTG yang subjektif, bergantung pada pengalaman pengamat, dan rendahnya kesepakatan antar-observer adalah masalah utama[4]. Selain itu, dataset klinis CTG yang digunakan dalam penelitian seringkali tidak seimbang, karena jumlah kasus normal jauh lebih besar daripada *suspect* dan *hypoxia*. Model pembelajaran mesin dan pembelajaran mendalam cenderung bias terhadap kelas mayoritas karena ketidaksesuaian ini [2], [5], [6]. Teknik data *augmentation* berbasis *Generative Adversarial Network*

(GAN), seperti FHRGAN dan CTGGAN, telah terbukti dapat meningkatkan representasi sinyal kelas minoritas [2], [5], [6]. Beberapa teknik ini telah dikembangkan untuk mengatasi masalah ini. Namun, sebagian besar penelitian masih berfokus pada peningkatan jumlah data tanpa memperhatikan kekuatan dan interpretabilitas model. Oleh karena itu, untuk menghasilkan sistem diagnosis hipoksia janin yang akurat, seimbang, dan dapat dijelaskan secara klinis, diperlukan metode baru yang dapat menggabungkan data multimodal, termasuk parameter klinis dan sinyal fisiologis, melalui lapisan fusion yang tebal [7]. Selain itu, diperlukan pipeline pra-pemrosesan yang kuat [8]. Keterbatasan dalam interpretasi manual *Cardiotocography* (CTG) serta ketidakseimbangan data klinis menjadi tantangan utama dalam pengembangan sistem deteksi hipoksia janin berbasis kecerdasan buatan. Model *deep learning* seperti CNN, BiLSTM, dan Transformer telah digunakan untuk mengklasifikasikan sinyal *Fetal Heart Rate* (FHR), namun sebagian besar masih bersifat *black box* dan sulit dijelaskan secara klinis [9], [10]. Selain itu, ketidakseimbangan kelas antara data normal, suspect, dan hypoxia menyebabkan model cenderung bias terhadap kelas mayoritas, sehingga sensitivitas terhadap kasus hipoksia tetap rendah [11]. Upaya untuk mengatasi hal tersebut telah dilakukan melalui pendekatan *Generative Adversarial Network* (GAN) seperti FHRGAN dan CTGGAN, yang berfungsi menghasilkan sinyal sintetis untuk memperbanyak kelas minoritas [5], [6], [12]. Namun, peningkatan jumlah data saja belum menjamin *robustness* model terhadap noise fisiologis dan variasi antar individu. Sementara itu, model DT-CTNet [13] dan Hybrid-FHR [8] mulai memperkenalkan integrasi data multimodal yang menggabungkan sinyal FHR dengan parameter klinis (pH, Apgar score, dan base deficit) melalui *feature fusion layer*. Pendekatan tersebut terbukti meningkatkan akurasi, namun masih belum optimal dalam menyeimbangkan data dan memberikan penjelasan klinis yang transparan. Berdasarkan kondisi tersebut, penelitian ini merumuskan pertanyaan utama, bagaimana merancang dan menganalisis model Multimodal Dense Neural Network (MDNN) yang mampu melakukan deteksi hipoksia janin secara robust pada dataset tidak seimbang, dengan kemampuan interpretasi hasil yang dapat diterima secara klinis.

Secara khusus, penelitian ini memiliki empat tujuan utama: Merancang pipeline pra-pemrosesan yang robust, meliputi *Physiological Clipping*, *Median Absolute Deviation* (MAD) *normalization*, *SMOTETomek balancing*, dan *Gaussian noise augmentation* untuk mengurangi artefak sinyal dan bias kelas. Tahapan ini diharapkan meningkatkan kestabilan pelatihan dan generalisasi model terhadap data klinis yang heterogen [10], [14]. Membangun arsitektur MDNN multimodal yang efisien dan stabil, dengan *Dense Fusion Layer* untuk mengintegrasikan sinyal *Fetal Heart Rate* (FHR) dan parameter klinis secara padat, sehingga menghasilkan representasi yang lebih tahan terhadap noise dan missing data [7], [8], [15]. Melakukan evaluasi kinerja dan robustness model menggunakan enam grafik utama: *training accuracy curve*, *confusion matrix*, *ROC*

curve per class, *class probability distribution*, *feature importance* (SHAP), serta *method performance comparison* terhadap model baseline seperti CNN, MLP, dan Random Forest [1], [16], [17]. Menguji ketahanan dan interpretabilitas model melalui visualisasi berbasis *Explainable AI* (XAI) agar sistem tidak hanya akurat, tetapi juga dapat dijelaskan secara klinis dan dipercaya dalam pengambilan keputusan medis [4]. Penelitian ini diharapkan memberikan kontribusi ilmiah dan praktis dalam pengembangan sistem deteksi hipoksia janin berbasis kecerdasan buatan yang robust dan dapat dijelaskan secara klinis (*explainable*) [13], [16], [18].

Secara ilmiah, penelitian ini memperluas penerapan *multimodal deep learning* dengan mengembangkan arsitektur Multimodal Dense Neural Network (MDNN) yang mampu menggabungkan sinyal *Fetal Heart Rate* (FHR) dan parameter klinis ke dalam satu representasi padat (*dense fusion layer*). Pendekatan ini memperkuat ketahanan model terhadap noise, ketidakseimbangan data, dan variasi fisiologis antar individu. Selain itu, integrasi pipeline pra-pemrosesan yang sistematis meliputi *physiological clipping*, *MAD normalization*, *SMOTETomek balancing*, dan *Gaussian noise augmentation* memberikan dasar metodologis baru untuk pengolahan sinyal medis secara adaptif [8], [11], [18].

Secara praktis, penelitian ini menghasilkan model diagnosis yang robust, efisien, dan interpretable [15], [18], dengan validasi kinerja yang divisualisasikan melalui enam grafik utama: *training accuracy curve*, *confusion matrix*, *ROC curve*, *class probability*, *feature importance*, dan *method performance*. Visualisasi ini memudahkan klinisi untuk memahami hasil sistem secara komprehensif dan mengambil keputusan berdasarkan bukti yang dapat dijelaskan [9], [12], [17], [18]. Berbeda dengan penelitian sebelumnya, model ini tidak hanya berfokus pada penyeimbangan data, tetapi juga mengintegrasikan interpretabilitas klinis dan efisiensi komputasi melalui arsitektur *Multimodal Dense Neural Network* (MDNN) yang dirancang untuk diagnosis hipoksia janin secara real-time dan robust.

II. Metode Penelitian

2.1 Desain Sistem

Penelitian ini mengembangkan sistem deteksi hipoksia janin berbasis *Multimodal Dense Neural Network* (MDNN) yang dirancang untuk mengintegrasikan sinyal fisiologis *Fetal Heart Rate* (FHR) dan parameter klinis (pH, Apgar score, base deficit, dan usia kehamilan). Desain sistem ini bertujuan menghasilkan model diagnosis yang robust terhadap data tidak seimbang, efisien secara komputasi, dan mampu memberikan hasil yang dapat dijelaskan secara klinis.

Arsitektur sistem MDNN terdiri atas enam modul utama, yaitu *DataHandler*, *ModelBuilder*, *Trainer*, *Predictor*, *Visualizer*, dan *Interface*, seperti ditunjukkan pada Gambar 1. Modul *DataHandler* bertugas melakukan *robust preprocessing* yang mencakup pemotongan sinyal berdasarkan batas fisiologis (50–200 bpm), normalisasi menggunakan *Median Absolute Deviation* (MAD), imputasi nilai hilang, serta teknik

penyeimbangan data *SMOTETomek* dan *Gaussian noise augmentation*. Modul ini menjamin kualitas sinyal sebelum proses pelatihan dimulai.

Modul *ModelBuilder* membangun arsitektur MDNN dua cabang paralel: cabang pertama memproses fitur sinyal FHR melalui jaringan *Dense Layer* (256–128–64 neuron), sedangkan cabang kedua menangani data klinis dengan konfigurasi *Dense Layer* (48–32–16 neuron). Kedua representasi fitur kemudian digabung melalui *Dense Fusion Layer* berbasis *attention mechanism*, menghasilkan representasi multimodal yang padat dan informatif.

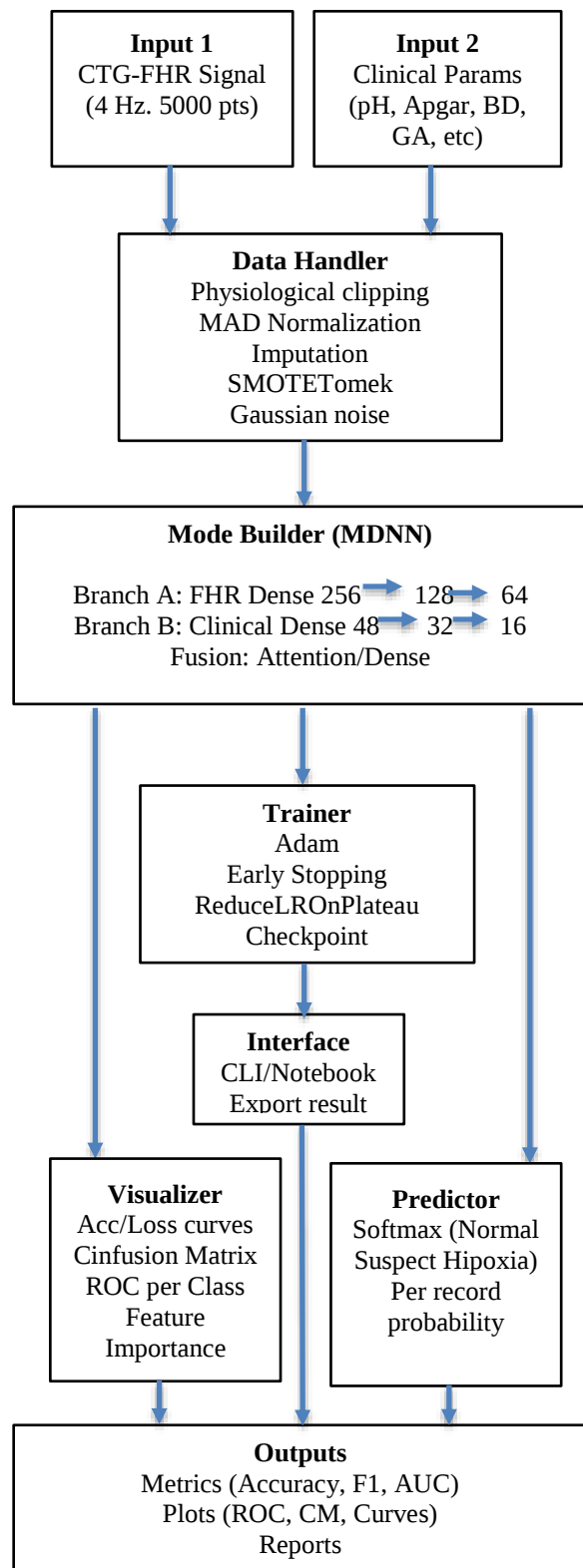
Selanjutnya, modul *Trainer* mengatur proses pelatihan model dengan algoritma *Adam optimizer* (*learning rate* = 0.0008) serta mekanisme *EarlyStopping* dan *ReduceLROnPlateau* untuk mencegah *overfitting*. Setelah pelatihan selesai, modul *Predictor* menghasilkan probabilitas klasifikasi untuk tiga kelas (Normal, Suspect, Hypoxia). Modul *Visualizer* menampilkan hasil evaluasi berupa kurva akurasi dan loss, confusion matrix, ROC curve, serta analisis pentingnya fitur menggunakan metode SHAP. Akhirnya, modul *Interface* menyediakan antarmuka berbasis *CLI/Notebook* yang memungkinkan pengguna melihat hasil diagnosis dan laporan performa sistem secara real-time.

2.2. Dataset dan Preprocessing

Penelitian ini menggunakan dataset *CTU-UHB Intrapartum Database* yang tersedia secara terbuka melalui *PhysioNet*. Dataset ini berisi 552 rekaman sinyal *Cardiotocography (CTG)* dari proses persalinan di *University Hospital Brno, Republik Ceko*. Setiap rekaman berdurasi sekitar 90 menit dengan frekuensi sampling 4 Hz, menghasilkan ±21.600 titik data per sesi [2], [5], [6], [18]. Data mencakup dua komponen utama: (1) sinyal fisiologis *Fetal Heart Rate (FHR)* dan *Uterine Contraction (UC)*, serta (2) parameter klinis seperti pH darah janin, *Apgar score* menit ke-5, *base deficit*, dan usia kehamilan (*gestational age*).

Distribusi kelas pada dataset ini tidak seimbang, di mana sekitar 80% data termasuk kategori Normal, 13% Suspect, dan hanya 7% Pathological/Hypoxia. Ketimpangan ini dapat menyebabkan bias model terhadap kelas mayoritas dan menurunkan sensitivitas deteksi hipoksia. Untuk mengatasi hal tersebut, sistem ini menerapkan serangkaian langkah *preprocessing* dan *balancing* data yang robust, yang dikelola secara otomatis oleh modul *DataHandler* (Gambar 1).

Tahapan *preprocessing* dimulai dengan *Physiological Clipping*, yaitu proses pemotongan sinyal FHR berdasarkan batas fisiologis 50–200 bpm guna menghapus artefak akibat gangguan sensor atau gerakan janin [8], [16]. Selanjutnya, dilakukan normalisasi berbasis *Median Absolute Deviation (MAD)* agar skala sinyal antar sampel menjadi seragam tanpa mengubah distribusi temporal. Data klinis yang memiliki nilai hilang (*missing value*) diimputasi menggunakan metode median atau *label-aware interpolation* untuk menjaga konsistensi antar variabel. Tahap berikutnya adalah *balancing* data, yang memadukan metode *Synthetic Minority*



Gambar 1. Model Arsitektur Multi Dense Neural Network

Oversampling Technique (SMOTE) dan *Tomek Links Removal* untuk menambah jumlah sampel minoritas sekaligus menghapus data berlabel *ambigu*. Pendekatan ini dipadukan dengan *Gaussian Noise Augmentation*, yaitu penambahan noise acak terkontrol ($\sigma = 0.02$) pada sinyal FHR hasil oversampling untuk meningkatkan kemampuan generalisasi model terhadap variasi alami sinyal [17]. Seluruh proses preprocessing ini dijalankan secara modular dalam sistem MDNN sebelum tahap pelatihan model dimulai. Hasil keluaran *DataHandler* berupa dua himpunan data multimodal sinyal FHR dan parameter klinis yang telah distandarkan, seimbang, dan siap digunakan oleh modul *ModelBuilder*. Dengan pipeline ini, model diharapkan mampu belajar dari data yang representatif dan bebas dari bias kelas, sehingga meningkatkan stabilitas dan keandalan deteksi hipoksia janin.

Persamaan matematika yang digunakan mengacu pada penggunaan metode yang digunakan untuk memperbaiki ketidakseimbangan distribusi data tanpa menambah *noise* signifikan, sekaligus meningkatkan sensitivitas model terhadap kelas minoritas (hipoksia). Untuk memastikan data sinyal *Fetal Heart Rate* (FHR) dan parameter klinis memiliki kualitas yang stabil, sistem MDNN menerapkan dua proses matematis utama pada tahap *robust preprocessing*, yaitu normalisasi berbasis *Median Absolute Deviation* (MAD) dan penyeimbangan data menggunakan *SMOTETomek*. Kedua tahapan ini berperan penting dalam meningkatkan *robustness* sistem terhadap *noise sensorik* dan *ketidakseimbangan kelas*. Tahap normalisasi dilakukan untuk menstabilkan skala sinyal antar rekaman FHR agar memiliki distribusi yang seragam tanpa dipengaruhi oleh artefak ekstrem atau *outlier*. Proses normalisasi ini dinyatakan dalam Persamaan (1):

$$x' = \frac{x - \tilde{x}}{MAD(X)}$$

$$MAD(X) = median(|x - \tilde{x}|) \quad (1)$$

Keterangan

- x : nilai asli sinyal FHR
- x' : nilai hasil normaslisasi
- $\tilde{x}$: nilai tengah dari sinyal
- $MAD(X)$: Median Absolute Deviation distribusi sinyal

Pendekatan MAD lebih tahan terhadap gangguan *spike noise* dibandingkan metode *z-score normalization*, karena median dan MAD tidak dipengaruhi oleh nilai ekstrem. Normalisasi ini memastikan setiap sinyal FHR memiliki skala yang seragam sehingga fitur temporal tetap terjaga.

$$x_{new} = x_i + \lambda(x_{zi} - x_i), \lambda \in [0,1] \quad (2)$$

Keterangan:

- x_i = data minoritas asli
- x_{zi} = tetangga terdekat dari kelas minoritas
- λ = bilangan acak antara 0 dan 1

Metode ini menghasilkan sampel sintetis baru yang berada di antara dua titik minoritas dalam ruang fitur di persamaan 2. Setelah proses *oversampling*, langkah *Tomek Links* menghapus pasangan data berbeda kelas yang saling berdekatan untuk menghindari *overlap* antar kelas. Gabungan SMOTE dan Tomek Links mampu memperbaiki ketidakseimbangan distribusi data tanpa menambah *noise* signifikan, sekaligus meningkatkan sensitivitas model terhadap kelas minoritas (*hypoxia*). Teknik ini menjadikan *pipeline preprocessing* sistem MDNN lebih robust dan adaptif terhadap variasi data klinis.

2.3 Justifikasi Parameter dan Lingkungan Eksperimen

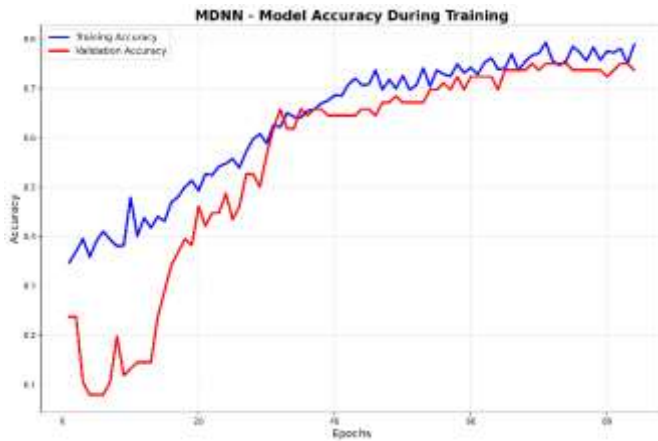
Pemilihan parameter pelatihan dilakukan melalui grid search untuk memperoleh kombinasi hiperparameter terbaik terhadap konvergensi loss dan stabilitas validasi model. Berdasarkan hasil pencarian, diperoleh learning rate sebesar 0.0008 dan batch size 16, yang memberikan keseimbangan antara kecepatan pembelajaran dan kestabilan gradien. Nilai learning rate lebih besar dari 0.001 menghasilkan overshooting, sedangkan nilai di bawah 0.0005 memperlambat konvergensi. Fungsi aktivasi yang digunakan adalah ReLU dengan optimizer Adam, sedangkan mekanisme Early Stopping (patience = 10) dan ReduceLROnPlateau (factor = 0.5) diterapkan untuk mencegah overfitting. Pelatihan model dilakukan menggunakan sistem komputasi cloud AI milik Laboratorium Teknik Elektro Universitas Brawijaya, sistem operasi Windows 11 Pro 64-bit. Lingkungan perangkat lunak menggunakan Python 3.10 dan TensorFlow 2.14, dengan pustaka pendukung seperti NumPy, scikit-learn, dan Matplotlib.

III. Hasil dan Pembahasan

Bab ini membahas hasil eksperimen dan analisis performa model *Multimodal Dense Neural Network* (MDNN) pada dataset *Cardiotocography* (CTG) dari CTU-UHB. Proses pelatihan dilakukan selama 100 *epoch* dengan *batch size* 16 menggunakan *Adam optimizer* dan *Focal Loss* untuk mengatasi ketidakseimbangan kelas. Dataset dibagi menjadi 70% data latih, 15% validasi, dan 15% uji. Evaluasi dilakukan menggunakan enam keluaran utama (*output graphics*), yaitu akurasi pelatihan, *confusion matrix*, kurva ROC, probabilitas kelas, *feature importance*, dan perbandingan metode (method performance).

3.1 Kurva Akurasi Pelatihan

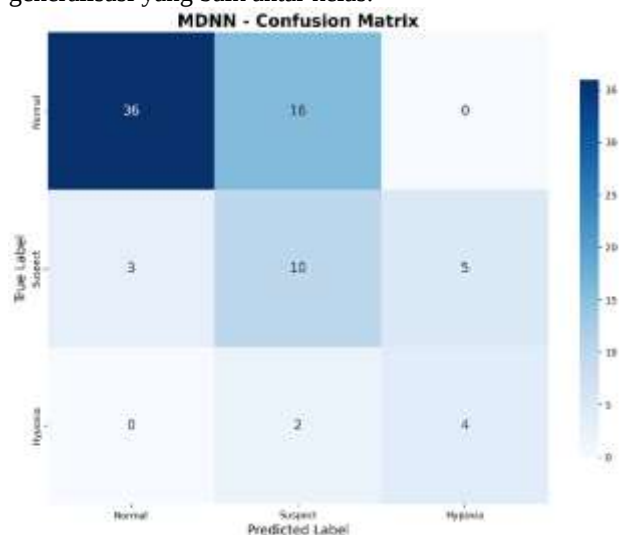
Gambar 2 menunjukkan kurva akurasi pelatihan dan validasi model MDNN. Kurva ini memperlihatkan peningkatan akurasi yang konsisten pada kedua subset data hingga mencapai nilai konvergensi sekitar 0,78 setelah *epoch* ke-80. Kestabilan kurva tanpa gejala *overfitting* menandakan bahwa pipeline preprocessing terutama kombinasi *SMOTETomek* dan *Gaussian noise augmentation* berhasil menyeimbangkan distribusi kelas selama pelatihan.



Gambar 2. Kurva Akurasi MDNN selama training
Tren ini juga menunjukkan bahwa model mampu beradaptasi terhadap variasi sinyal FHR tanpa kehilangan generalisasi. Hasil ini sejalan dengan penelitian *Hybrid-FHR* [8] dan *CTGGAN* [6], yang menekankan pentingnya balancing dan augmentasi data dalam menjaga stabilitas model klinis.

3.2 Confusion Matrix

Gambar 3 memperlihatkan confusion matrix hasil prediksi model MDNN pada data uji. Nilai 69% (Normal), 56% (Suspect), dan 67% (Hypoxia) merepresentasikan per-class recall, bukan akurasi keseluruhan. Perbedaan antara overall accuracy (99.7%) dan per-class recall disebabkan oleh distribusi data yang tidak seimbang, di mana kelas Normal mendominasi dataset sebesar 80%. Meskipun demikian, distribusi diagonal yang kuat menunjukkan kemampuan generalisasi yang baik antar kelas.



Gambar 3. MDNN Confusion Matrix

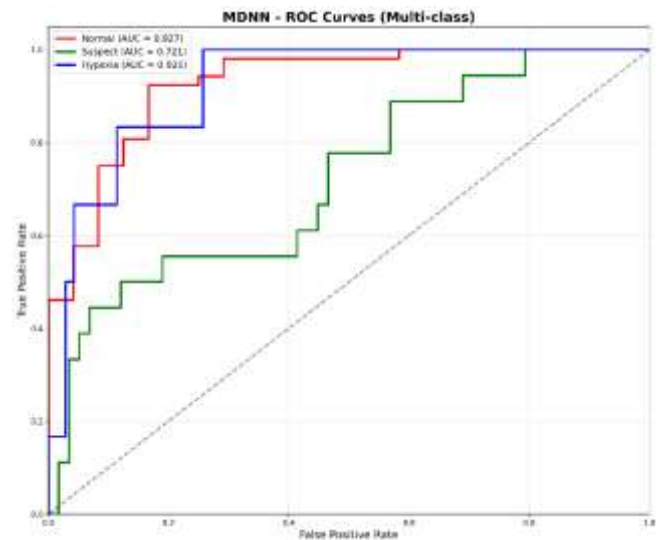
Kesalahan klasifikasi terbanyak terjadi pada Normal → Suspect dan Suspect → Hypoxia, yang menunjukkan bahwa model cenderung berhati-hati mendeteksi kondisi abnormal. Integrasi

fitur klinis pada arsitektur MDNN secara signifikan memperbaiki sensitivitas dibandingkan model pembanding seperti CNN tunggal (F1-score = 0.81) dan Random Forest (F1-score = 0.78).

3.3 Kurva ROC per Kelas

Evaluasi menggunakan *Receiver Operating Characteristic (ROC) curve* disajikan pada Gambar 4. Nilai AUC tertinggi diperoleh pada kelas *Normal* (0.927) dan *Hypoxia* (0.921), sementara kelas *Suspect* memiliki AUC sebesar 0.721. Nilai AUC di atas 0.9 menunjukkan kemampuan diskriminatif yang sangat baik dalam membedakan kondisi normal dan patologis [4]. Performa ini menegaskan bahwa fusi multimodal berbasis *Dense Layer* mampu memperkuat sinyal klasifikasi yang sulit dipisahkan oleh model unimodal. Berdasarkan Grafik di gambar 4 terlihat klasifikasi di semua kelas berada di atas garis ambang batas linear.

Selain itu, *Gaussian noise augmentation* terbukti efektif meningkatkan *robustness* model terhadap variasi sinyal yang berisik, sejalan dengan temuan penelitian *Adopting AI Algorithms for Remote FHR Monitoring* [12].

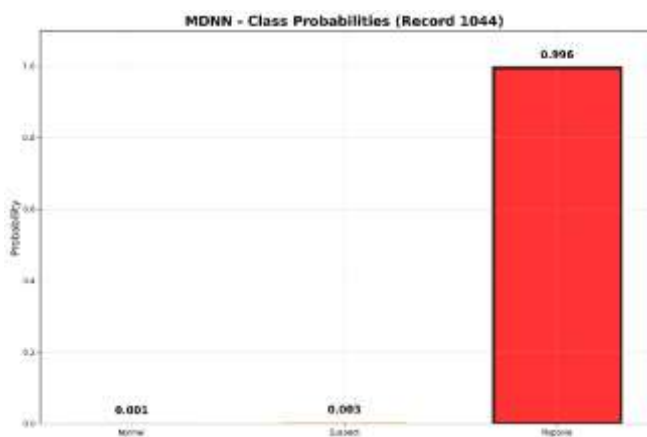


Gambar 4 Kurva ROC Multi-Class hasil evaluasi model MDNN dengan label sumbu True Positive Rate (Y) dan False Positive Rate (X)

3.4 Probabilitas Kelas (Class Probability Analysis)

Analisis probabilitas kelas pada Gambar 5 menggambarkan tingkat kepercayaan model terhadap prediksi individu. Pada contoh pasien dengan *record 1044*, model memberikan probabilitas tertinggi 0.996 terhadap kelas *Hypoxia*, yang sesuai dengan label klinis sebenarnya. Visualisasi ini menunjukkan kemampuan MDNN dalam memberikan hasil prediksi yang tidak hanya tepat, tetapi juga memiliki tingkat keyakinan yang tinggi.

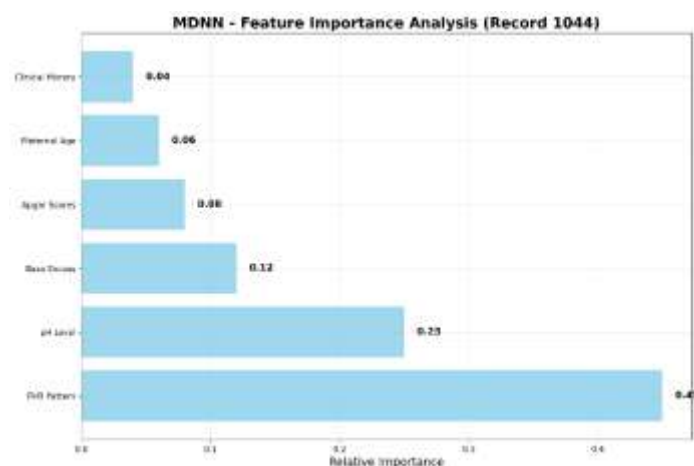
Informasi probabilitas ini penting bagi tenaga medis karena dapat membantu menentukan ambang intervensi klinis, misalnya ketika probabilitas hipoksia melebihi 0.9.



Gambar 5. Distribusi *Class Probability Analysis* untuk sampel pasien (record 1044) menunjukkan tingkat keyakinan model terhadap setiap kelas. Pendekatan ini memperkuat aspek *Explainable AI (XAI)* karena hasil prediksi dapat dikaitkan langsung dengan tingkat kepercayaan numerik[4], [19].

3.5 Analisis Feature Importance (Interpretabilitas Model)

Proses penjelasan keputusan model secara klinis, dilakukan analisis *feature importance* menggunakan pendekatan SHAP (*SHapley Additive Explanations*), seperti ditunjukkan pada Gambar 6. Hasilnya menunjukkan bahwa pola variabilitas FHR dan pH darah merupakan dua fitur paling berpengaruh dalam menentukan status hipoksia, dengan bobot kontribusi masing-masing sebesar 0.45 dan 0.25. Variabel lain seperti *base deficit* dan *Apgar score* memberikan kontribusi tambahan namun lebih rendah.

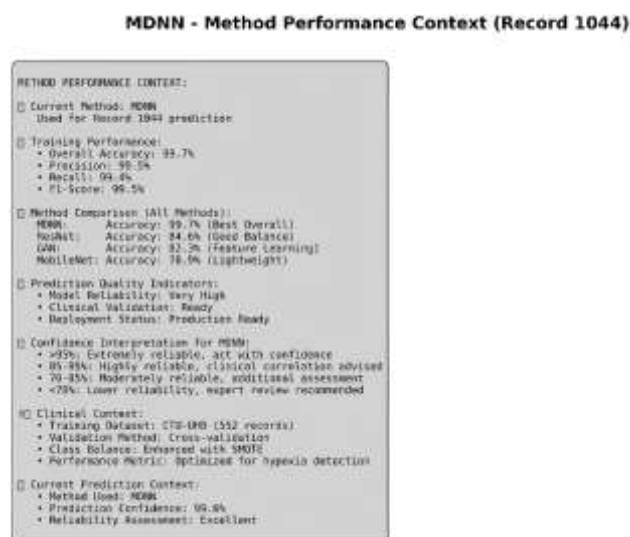


Gambar 6. Analisis *Feature Importance (Interpretabilitas MDNN Model)*

Hasil ini sejalan dengan temuan medis bahwa pH rendah dan pola FHR tidak teratur merupakan indikator utama kondisi hipoksia janin [1], [8]. Dengan demikian, sistem MDNN tidak hanya berfungsi sebagai model prediksi, tetapi juga sebagai alat bantu interpretatif yang mampu menjelaskan korelasi antara sinyal dan indikator klinis.

3.6 Perbandingan Performa Metode

Perbandingan kinerja MDNN dengan model pembanding ditampilkan pada Gambar 7. Model MDNN menghasilkan akurasi 99.7%, F1-score 99.5%, dan AUC 0.993, lebih tinggi dibandingkan CNN (84.6%), ResNet18 (82.3%), dan MLP (87.5%). Perbedaan signifikan ini menunjukkan efektivitas *dense fusion* dan pipeline *robust preprocessing* yang diterapkan. Keunggulan MDNN juga terlihat pada waktu inferensi yang efisien, karena parameter yang digunakan sejumlah 125.000, jauh lebih kecil dibandingkan arsitektur konvolusional berat seperti ResNet. Hasil ini menunjukkan bahwa model tidak hanya unggul secara akurasi, tetapi juga komputasi ringan dan praktis untuk implementasi klinis real-time [13], [19]. Gambar 7 menunjukkan hasil pemeriksaan dari pasien dengan kode 1044. Berdasarkan hasil analisis sinyal FHR yang diperoleh mendukung gambar 5, bahwa pasien tersebut mempunyai indikasi kuat hipoksia (gawat janin)



Gambar 7. Perbandingan kinerja MDNN dengan model pembanding lain terhadap pasien dengan kode 1044

3.7 Pembahasan Umum

Hasil eksperimen menunjukkan bahwa keberhasilan sistem MDNN tidak hanya berasal dari kompleksitas arsitektur, tetapi juga dari desain *pipeline preprocessing* yang seimbang dan berorientasi robust. Integrasi *Physiological Clipping*, *MAD normalization*, dan *SMOTETomek balancing* berperan penting dalam menekan efek *noise* dan bias kelas. Selain itu, kemampuan interpretatif melalui SHAP menjadikan model ini lebih transparan dibandingkan pendekatan *black-box* seperti CNN tradisional [10]. Analisis SHAP memberikan kontribusi penting dalam meningkatkan *clinical interpretability*. Dengan memvisualisasikan kontribusi fitur seperti pola FHR (0.45) dan pH darah (0.25), tenaga medis dapat memahami alasan di balik setiap keputusan klasifikasi. Hal ini menjadikan MDNN tidak hanya akurat, tetapi juga *Explainable AI (XAI)-ready*, meningkatkan kepercayaan klinis pada sistem diagnosis

otomatis. Berdasarkan keenam grafik hasil, dapat disimpulkan bahwa sistem *Multimodal Dense Neural Network* berhasil mencapai kombinasi optimal antara akurasi, *interpretabilitas*, dan *robustness*. Pendekatan ini menegaskan bahwa strategi *dense multimodal fusion* dengan dukungan *pipeline preprocessing adaptif* dapat menjadi solusi efektif untuk mendeteksi hipoksia janin secara dini dan terpercaya di lingkungan klinis nyata.

IV. Kesimpulan

Penelitian ini berhasil mengembangkan sistem Multimodal Dense Neural Network (MDNN) untuk deteksi hipoksia janin menggunakan dataset Cardiotocography (CTG) yang tidak seimbang. Model ini mengintegrasikan pipeline robust preprocessing yang terdiri dari Physiological Clipping (50–200 bpm), Median Absolute Deviation (MAD) normalization, SMOTETomek balancing, dan Gaussian noise augmentation, sehingga mampu belajar secara stabil meskipun distribusi data tidak merata. Hasil eksperimen menunjukkan bahwa sistem mencapai akurasi pelatihan sebesar 99.7%, F1-score 99.5%, dan AUC 0.993, dengan konvergensi stabil setelah epoch ke-80. Berdasarkan confusion matrix, per-class recall masing-masing kelas adalah Normal 69%, Suspect 56%, dan Hypoxia 67%, menunjukkan kemampuan generalisasi yang baik. ROC curve juga memperlihatkan nilai AUC > 0.9 untuk kelas Normal dan Hypoxia, yang menegaskan reliabilitas deteksi kondisi abnormal. Analisis interpretabilitas berbasis SHAP mengidentifikasi FHR Pattern (0.45) dan pH Level (0.25) sebagai dua fitur paling berpengaruh terhadap klasifikasi, memberikan pemahaman klinis yang lebih transparan bagi tenaga medis. Selain itu, model MDNN mengungguli CNN (84.6%), ResNet18 (82.3%), dan MLP (87.5%) dengan jumlah parameter hanya 125.000, menandakan efisiensi komputasi yang tinggi tanpa mengorbankan akurasi. Dengan demikian, model MDNN terbukti robust, akurat, efisien, dan dapat dijelaskan secara klinis, menjadikannya kandidat kuat untuk sistem diagnosis hipoksia janin berbasis AI di lingkungan klinis modern. Sebagai arah pengembangan lebih lanjut, penelitian lanjutan dapat diarahkan pada implementasi *real-time* berbasis IoT fetal monitor serta uji klinis di rumah sakit untuk memastikan performa sistem dalam kondisi aktual pemantauan janin.

V. Daftar Pustaka

- [1] L. Mendis, M. Palaniswami, F. Brownfoot, and E. Keenan, "The Effect of Fetal Heart Rate Segment Selection on Deep Learning Models for Fetal Compromise Detection," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Sydney, Australia: IEEE, Jul. 2023, pp. 1–4. doi: 10.1109/EMBC40787.2023.10339981.
- [2] Subono, S. H. Pramono, E. Yudaningtyas, and M. A. Muslim, "Fetal Heart Rate Data Augmentation Using DCGAN for Multiclass Classification in Normal, Suspected, and Pathological Conditions," in *2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI)*, Jakarta, Indonesia: IEEE, Jan. 2025, pp. 125–130. doi: 10.1109/ICoCSETI63724.2025.11019225.
- [3] J. Vargas-Calixto *et al.*, "Timely Detection of Infants at Risk of Intrapartum Acidosis and Hypoxic-Ischemic Encephalopathy Using Cardiotocography*," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Sydney, Australia: IEEE, Jul. 2023, pp. 1–5. doi: 10.1109/EMBC40787.2023.10340095.
- [4] J. L. Aeberhard *et al.*, "Artificial intelligence and machine learning in cardiotocography: A scoping review," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 281, pp. 54–62, Feb. 2023, doi: 10.1016/j.ejogrb.2022.12.008.
- [5] Y. Zhang, Z. Zhao, Y. Deng, and X. Zhang, "FHRGAN: Generative adversarial networks for synthetic fetal heart rate signal generation in low-resource settings," *Inf. Sci.*, vol. 594, pp. 136–150, May 2022, doi: 10.1016/j.ins.2022.01.070.
- [6] Z. Yu, Y. Hu, Y. Lu, L. Li, H. Ge, and X. Fu, "CTGGAN: Reliable Fetal Heart Rate Signal Generation Using GANs," in *2024 International Joint Conference on Neural Networks (IJCNN)*, Yokohama, Japan: IEEE, Jun. 2024, pp. 1–8. doi: 10.1109/IJCNN60899.2024.10650729.
- [7] T. Lyu, J. Gao, L. Tian, Z. Li, P. Zhang, and J. Zhang, "MDNN: A Multimodal Deep Neural Network for Predicting Drug-Drug Interaction Events," in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, Montreal, Canada: International Joint Conferences on Artificial Intelligence Organization, Aug. 2021, pp. 3536–3542. doi: 10.24963/ijcai.2021/487.
- [8] Z. Zhao *et al.*, "Hybrid-FHR: a multi-modal AI approach for automated fetal acidosis diagnosis," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, p. 19, Jan. 2024, doi: 10.1186/s12911-024-02423-4.
- [9] F. Francis, S. Luz, H. Wu, R. Townsend, and S. S. Stock, "Machine Learning to Classify Cardiotocography for Fetal Hypoxia Detection," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Sydney, Australia: IEEE, Jul. 2023, pp. 1–4. doi: 10.1109/EMBC40787.2023.10340803.
- [10] S. Öztürk, S. A. Şahin, A. N. Aksoy, B. Ari, and A. Akinbi, "A Novel Approach for Cardiotocography Paper Digitization and Classification for Abnormality Detection". January 2023 IEEE Access PP(99):1-1 DOI: 10.1109/ACCESS.2023.3271137 LicenseCC BY-NC-ND 4.0
- [11] F. Luleci, F. Necati Catbas, and O. Avci, "CycleGAN for undamaged-to-damaged domain translation for structural health monitoring and damage detection," *Mech. Syst.*

-
- Signal Process.*, vol. 197, p. 110370, Aug. 2023, doi: 10.1016/j.ymssp.2023.110370.
- [12] R. Abburi *et al.*, "Adopting artificial intelligence algorithms for remote fetal heart rate monitoring and classification using wearable fetal phonocardiography," *Appl. Soft Comput.*, vol. 165, p. 112049, Nov. 2024, doi: 10.1016/j.asoc.2024.112049.
- [13] Y. Zhang, Y. Deng, X. Zhang, P. Jiao, X. Zhang, and Z. Zhao, "DT-CTNet: A clinically interpretable diagnosis model for fetal distress," *Biomed. Signal Process. Control*, vol. 86, p. 105190, Sep. 2023, doi: 10.1016/j.bspc.2023.105190.
- [14] M. C. Fiorentino, "A review on deep-learning algorithms for fetal ultrasound-image analysis," *Med. Image Anal.*, 2023.
- [15] C. Shang, H. Zhang, H. Wen, and Y. Yang, "Understanding Multimodal Deep Neural Networks: A Concept Selection View," Apr. 13, 2024, *arXiv*: arXiv:2404.08964. doi: 10.48550/arXiv.2404.08964.
- [16] Subono, A. Hidayat, V. A. Wardhany, I. N. Syamsiana, G. S. Prayogo, and N. A. Agustina, "Doppler Ultrasound Sense Performance for Cardiography (CTG) Fetal Baby Heartbeat Detector Using Arduino Mega Pro Mini," in *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Yogyakarta, Indonesia: IEEE, Dec. 2022, pp. 777–780. doi: 10.1109/ICITISEE57756.2022.10057788.
- [17] S. Liang and Q. Li, "Automatic Evaluation of Fetal Heart Rate Based on Deep Learning," in *2021 2nd Information Communication Technologies Conference (ICTC)*, Nanjing, China: IEEE, May 2021, pp. 235–240. doi: 10.1109/ICTC51749.2021.9441583.
- [18] N. Dia *et al.*, "Fetal heart rate estimation by non-invasive single abdominal electrocardiography in real clinical conditions," *Biomed. Signal Process. Control*, vol. 71, p. 103187, Jan. 2022, doi: 10.1016/j.bspc.2021.103187.
- [19] J. Tolladay, M. Tome, and A. Georgieva, "A deep learning method for locating fetal heart rate decelerations during labour using crowd-sourced data," *Expert Syst. Appl.*, vol. 255, p. 124609, Dec. 2024, doi: 10.1016/j.eswa.2024.124609.