EFFICIENTNET-DRIVEN MULTIMODAL FUSION FRAMEWORK WITH AI-BASED CLINICAL MODELING FOR NEONATAL RISK AND MORTALITY PREDICTION

V. Priya,
Associate professor,
Department of Computer Science and Engineering,
Paavai Engineering College,
Pachal, Namakkal, India.
priyasaravanaraju@gmail.com

Preethi S,
PG Scholar,
Department of Computer Science Engineering,
Paavai Engineering College,
Pachal, Namakkal, India.
selvampreethi577@gmail.com

Abstract: *The issue of neonatal mortality prediction continues to be a thorn in the flesh of healthcare since the numerous interconnected factors of clinical, demographic, and imaging-related variables have an effect on infant outcomes. Risk assessment is a critical issue that requires timely intervention through early and accurate risk assessment especially in the neonatal intensive care unit where decisions have to be made in times of uncertainty. In the current study, it is suggested to use an EfficientNet-based multimodal fusion model combined with an artificial intelligence-based clinical model of predicting risks and mortality in neonatal disease prevention. The framework integrates profound visual representations of medical images of neonatal cases with the help of EfficientNet and represented clinical data (birth weight, gestational age, APGAR scores, and laboratory measurements). The use of a feature fusion approach can combine heterogeneous data sources to allow the model to model both spatial patterns as imaged and contextual relationships as available in tabular clinical attributes. Before the model training, the data are preprocessed through normalization, missing values, and features selection to enhance the quality of data and robustness of the model. The merged feature representation is then classified in a predictive model based on AI which is optimized to perform optimally. The experimental analysis proves that the suggested multimodal model with an experimental evaluation shows a better predictive ability than one-modality models based on the effective utilization of complementary information. The findings suggest that EfficientNet combined with clinical data modeling is a reliable and scalable decision-support model of neonatal outcome prediction in clinical practice.*

Keywords: Neonatal mortality prediction, EfficientNet, multimodal learning, feature fusion, artificial intelligence, medical imaging, clinical data, deep learning, healthcare analytics, decision support system.

I. INTRODUCTION

Neonatal period as the initial 28 days of life is a very vulnerable phase of life marked with high morbidity and mortality risks. High-risk neonates should be identified early to ensure that clinical interventions are provided on time especially in neonatal intensive care units (NICU) where sepsis, respiratory diseases, and low birth weights are major causes of death. Conventional methods of neonatal risk assessment are more dependent on clinical experience and manual assessment of patient records which can be scalable and prone to human error. The recent progress in the field of artificial intelligence (AI) and machine learning (ML) has allowed creating the predictive models that can analyze the complex clinical patterns to make a better decision. Various researches have examined the use of ML methods in predicting neonatal mortality. As an example, Shariat et al. came up with machine learning and convolutional neural network (CNN)-based solutions, showing that deep learning models may be used to achieve high prediction accuracy in neonatal mortality with non-image clinical data [1]. On the same note, Hamid et al. examined several deep learning models, such as multilayer perceptron (MLP), long short-term memory (LSTM), and CNN, showing that low birth weight and early breastfeeding are important factors predicting neonatal outcomes [2]. Besides predictive modeling, dataset preparation and the exploratory analysis are also essential in neonatal healthcare analytics. Ranade et al. prioritized preprocessing and analysis of the MIMIC-IV database to obtain records relating to neonatal sepsis and the significance of structured electronic health records to model development [3]. Moreover, a bibliometric analysis of the topic by Morais et al. highlighted the increasing attention to AI-based neonatal care, where multimodal sources of data, such as clinical, laboratory, genetic, and imaging data, are increasingly becoming a part of predictive models [4]. Within the framework of data scarcity, Lyra et al. suggested the application of synthetic data augmentation to improve the training of a model that predicts neonatal sepsis with encouraging outcomes in the early diagnosis tasks [5]. In spite of these achievements, the majority of the current solutions only address unimodal data or single prediction. It is still desirable to have integrated frameworks that are capable of effectively integrating heterogeneous data sources to enhance predictive performance and strength. This paper will fill this gap by putting forward a multimodal fusion-based approach that combines deep image features and organized clinical data to predict neonatal risk and mortality better.

II. RELATED WORKS

More recent research has investigated how machine learning and deep learning can be used to assess neonatal risks using the maternal determinants, multimodal data, time-related signals, and explainable AI approaches. The article by Derere et al. examined the use of homogenous ensemble prediction of neonatal and infant mortality based on maternal health factors. Their analysis showed that ensemble methods like AdaBoost and bagging are very effective in terms of classification of data where they score high in accuracy and F-measure thus explaining the importance of maternal determinants in prediction of mortality [6]. Lin et al. also resolved the single-outcome prediction issue in a different direction since they proposed a multi-task learning (MTL) framework to predict multiple adverse outcomes in neonatal care at the same time. Their model was able to utilize common representations across tasks and therefore the correlations between various clinical outcomes were better represented leading to better predictive performance as compared to the traditional single-task models [7]. This underscores the possibility of collaborative learning approaches in the modeling of complex health conditions of the neonate. There is also the explanation in neonatal prediction models. Marvin and Alam presented the explainable artificial intelligence (XAI) model of Neonatal Intensive Care Unit (NICU) admissions. Their models were based on the random forest and Logistic regression with interpretability methods including SHAP and LIME to make their decision on the model frameworks transparent. This promotes the trust of clinicians and helps medical workers to reason the predictions [8].

Deep learning architectures have also been developed in the case of real-time clinical decisions. Im et al. suggested a model based on multimodal transformer to predict neonatal endotracheal intubation. It is a model that incorporates time-series vital signs and structured clinical data to predict respiratory deterioration in real time. The experimental outcomes revealed that there was a high predictive performance, which proves that transformer-based architectures are efficient to work with heterogeneous and temporal medical data [9]. Moreover, Shariat et al. concentrated on machine learning methods of predicting neonatal infections with the help of large datasets of NICUs. They discovered that Random Forest was better than other classifiers and data balancing strategies did enhance performance of models to great extent. Notably, the paper also highlighted that the maternal characteristics can be used alone to provide some significance to early prediction of infections before birth, which implies the usefulness of the prepartum data in risk assessment in neonatal care [10]. Lately, the development of neonatal mortality prediction has paid more attention to the issue of data imbalance, multimodal combination, and hybrid learning approaches to enhance the predictive performance. S. A et al. suggested the application of SMOTE variants to deal with unbalanced data of neonatal mortality prediction tasks. Their method creates artificial samples of the minority classes to equal the dataset and it is followed by ensemble learning methods to enhance the classification. The paper points out that the issue of class imbalance is one of the most effective ways of improving the model reliability in predicting neonatal outcomes [11].

The concept of multimodal learning has also become very noticeable in healthcare analytics. Kulkarni et al. proposed a patient risk prediction framework that can use the Electronic Health Records (EHR) and medical imaging to predict the risk, based on the a Bidirectional Long Short-Term Memory (Bi-LSTM) network. The proposed model was highly accurate and indicated the reliability of utilizing heterogeneous sources of data to better identify risks in patients by identifying sequential dependencies across the two modalities [12]. Hybrid learning methods that integrate clustering and classification have also been considered in the analysis of medical images besides multimodal fusion. Karthika et al. suggested a Hybrid Clustering and Classification Algorithm (HCCA), in which the fuzzy C-means clustering is used to segment the region-of-interest and a convolutional neural network-based classifier is used to predict diseases. The experiment showed a better segmentation and classification accuracy, which proved the success of the unsupervised and supervised techniques of learning in medical diagnostics [13]. Intelligent systems that are monitored have also helped in assessing the health of the neonatal health. Chen et al. created a systemic monitoring of circulation that assesses the severity of neonatal illnesses based on the non-invasive measurement of the hemodynamic parameters. The use of neural networks to predict the degree of the severity of the illness by using blood perfusion indices was fruitful as it was able to effectively differentiate the mild and severe cases. This study indicates the possibility of physiological signal-based predictive systems in the care of the neonatal unit [14].

Moreover, Xu et al. developed a multimodal disease prediction framework that incorporates the laboratory tests data with the medical imaging with the help of an organ-centric generative model. Their approach handles the problem of time sparsity of imaging data by producing more images depending on laboratory tests. The suggested organ-centric graph and trajectory model showed great enhancement in the multimodal feature alignment and prediction accuracy in a variety of disease tasks [15]. Altogether, the literature review shows that there is increasing significance of multimodal learning, data balancing methods, hybrid structures, and physiological monitoring systems in the field of predicting neonatal conditions. Although some studies have been conducted on the modalities of an individual or selective fusion approach, there is a research gap in the effective combination of the deep imaging features with structured clinical data in a single and scalable system. The following study proposal will help to fill this gap, as the suggested multimodal fusion model has been developed using EfficientNet and will be applied to integrate visual and clinical representations to predict neonatal risks and mortality more effectively.

III. PROPOSED SYSTEM

The suggested system presents a multimodal artificial intelligence platform that can be used to enhance the prediction of neonatal mortality and risk through a combination of medical imaging information and clinical attributes well organized. The general concept of the system lies in the fact that the complementary information of heterogeneous sources of data is utilized and integrated into a single learning structure to increase predictive accuracy and strength. Figure.1 shows a proposed work architecture design. The system starts with the two main types of data modalities being collected, specifically, neonatal medical images (X-rays, or any other appropriate scans) and tabular clinical data consisting of such parameters as birth weight, gestational age, APGAR scores, oxygen levels, and laboratory parameters. The EfficientNet architecture is used to process the imaging data; it is a convolutional neural network, which has a balanced scaling in the depth, width, and resolution.

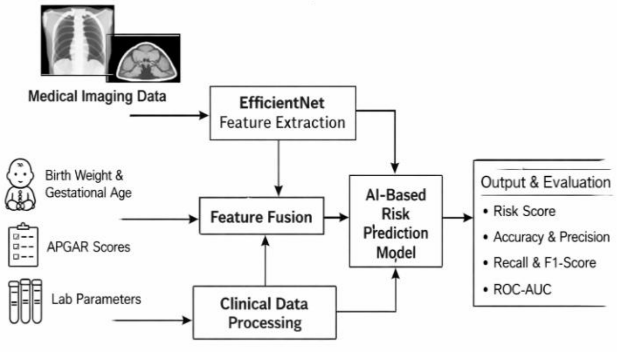

Figure.1 Proposed Work Architecture Diagram

EfficientNet will automatically extract high-level visual features that are able to represent the small patterns and abnormalities which cannot be easily identified using manual analysis. At the same time, the structured clinical data is pre-processed with normalization, categorical variables encoding, and the treatment of the missing values. The method of feature selection is used to help select the most appropriate clinical attributes but eliminate redundancy. This processed data is subsequently subjected to an AI-based predictor model which usually comprises of thick neural network layers or ensemble learning algorithms to acquire insightful connections in the tabular information. One of the elements of the system is the feature fusion mechanism, in which deep image features obtained using EfficientNet are used together with the encoded clinical features. This convergence may be applied by adding concatenation and full connected layers thus enabling the model to learn cross-modal interactions collectively. The resulting combined feature representation is then inputted in a layer of classification that forecasts the level of risk or probability of death in a neonatal case. The system undergoes supervised learning and the performance of the system is measured by the standard measures like accuracy, precision, recall, F1-score, and ROC-AUC. The proposed system will offer a more complete and accurate prediction framework that can be applied within the clinical decision-making of a neonatal care setting due to the integration of multimodal information.

IV. METHODOLOGY

The suggested methodology is a multimodal learning system that combines deep visual features analysis with the clinical data analysis in a structured format to identify the risk and mortality levels in neonatal patients. The general workflow involves the data acquisition, preprocessing, feature extraction, feature fusion, model training and evaluation that can be detailed as follows.

A. Data Acquisition & Data Preprocessing

The two different data sources used in the system are the medical imaging data and structured clinical records. The imaging data could be in the form of neonatal radiographs or other diagnostic scans, whereas the parameters of the clinical attributes are the birth weight, gestational age, APGAR scores, heart rate, respiratory rate, and lab parameters. These auxiliary modalities give the visual and physiological views on neonatal health. The two modalities of data are preprocessed before training of the model. In the case of clinical data, missing values are also processed with the imputation methods and categorical variables are also coded into numerical data. Continuous variables are scaled or made to be the same to make sure that they are uniform. In imaging data, the preprocessing involves the resizing of images to definite resolution, pixel intensities normalization and optional data augmentation methods like rotation and flipping to enhance generalization.

Let the dataset consist of two modalities: medical imaging data and structured clinical attributes. The imaging dataset is represented as a set of images $I = \{I_1, I_2, \ldots, I_n\}$, where each image corresponds to a neonatal case. The clinical dataset is represented as a feature matrix $X_c \in \mathbb{R}^{n\times d}$, where d denotes the number of clinical attributes such as birth weight, gestational age, and APGAR scores. The target variable is defined as $y \in \{0,1\}$, where 0 represents survival and 1 represents mortality.

B. Feature Extraction using EfficientNet

The EfficientNet is utilized as the basis of extracting deep features of medical images. The architecture has a compound scaling technique which scales network depth, width, and resolution uniformly and allows efficient and accurate representation learning. Convolutional layers represent hierarchical features of edges and semantic patterns which are of relevance to the neonatal conditions. The EfficientNet model uses a high-dimensional feature vector as the output of the model that is the imaging modality.

The EfficientNet model is employed to extract deep hierarchical features from the imaging data. Let the transformation performed by EfficientNet be denoted as a function $f_\theta(\cdot)$, parameterized by weights θ. The extracted feature vector for an input image I_i is given by:

$$F_i = f_\theta(I_i) \qquad (1)$$

where $F_i \in \mathbb{R}^k$ represents the learned feature embedding capturing spatial and semantic information from the medical image.

C. Clinical Feature Modeling

A feedforward neural network or some other machine learning model is used to process structured clinical data. The network is able to learn nonlinear correlation between clinical variables and create a concise representation of features that encapsulates the patient specific information that is important to risk prediction.

The structured clinical data undergoes preprocessing and is then passed through a neural transformation function $g_\phi(\cdot)$, parameterized by weights ϕ. The transformed clinical feature representation is expressed as:

$$C_i = g_\phi(X_{c,i}) \qquad (2)$$

where $C_i \in \mathbb{R}^m$ denotes the encoded clinical feature vector corresponding to the i^{th} sample.

D. Feature Fusion

The features extracted and clinical features are integrated by a feature fusion technique which is normally concatenation and then fully connected layer. The combination of these two enables the model to learn interrelationships between imaging pattern and clinical characteristics together, enhancing the richness of the representation. The resulting fused feature space is then subjected to dense layers, activation functions of

ReLU and regularization methods of dropout to avoid overfitting.

To integrate information from both modalities, a fusion operation is applied. The most common approach is feature concatenation, defined as:

$$Z_i = [F_i \oplus C_i] \qquad (3)$$

where $Z_i \in \mathbb{R}^{k+m}$ represents the fused feature vector, and $\oplus$ denotes concatenation. This fused representation captures complementary relationships between imaging features and clinical attributes.

E. Classification Layer

The resulting fused representation is then used by a layer of classification, usually with sigmoid activation when binary (survival vs. mortality) classification is desired, or with softmax when multi-class risk classification is desired. The model gives the probability scores of the probability of neonatal outcomes.

The fused feature vector is passed through a fully connected neural network to perform classification. The final output is computed as:

$$\hat{y}_i = \sigma(WZ_i + b) \qquad (4)$$

where W and b are the learnable parameters, and $\sigma(\cdot)$ is the sigmoid activation function defined as:

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad (5)$$

This produces a probability score indicating the likelihood of neonatal mortality.

F. Training Strategy and Optimization

The model is trained by supervised learning whereby the loss functions include binary cross-entropy. The weights of model are updated using optimization algorithms such as Adam. Early termination and learning rate schedule can be used to achieve better convergence and avoid overfitting.

The model is trained using binary cross-entropy loss, defined as:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)] \qquad (6)$$

where N is the number of samples, y_i is the ground truth label, and $\hat{y}_i$ is the predicted probability. The optimization process minimizes this loss using gradient-based methods such as Adam, updating parameters θ and ϕ iteratively.

G. Evaluation Metrics

Standard evaluation measures such as accuracy, precision, recall, F1-score and ROC-AUC are used to measure model performance. The metrics give an overall insight of the classification performance especially when there are imbalanced data as is the case in medical applications.

V. RESULT & DISCUSSION

A. Experimental Setup

The suggested EfficientNet-based multimodal fusion system was tested on a neonatal dataset that included medical imaging, as well as structured clinical characteristics. The data was split into training and testing groups with 80:20 percentage to have an unbiased evaluation. The clinical data was standardized with standard preprocessing methods like normalization, missing values imputation and feature scaling and medical images were resized and augmented to enhance generalization. The Adam optimizer with binary cross - entropy loss was used to train the model. The evaluation of performance was done on accuracy, precision, recall, F1-score, and ROC-AUC.

B. Quantitative Performance Analysis

The proposed multimodal model performance was compared to the relevant models in each of the modality types, such as a clinical-only model, and image-only EfficientNet model. The multimodal fusion method was seen to beat the single-modality baselines in all the evaluation measures as indicated in Table I. The model was able to model complementary patterns due to the integration of heterogeneous data sources which led to enhanced predictive capability.

TABLE I. PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model Type	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Clinical Model	0.82	0.80	0.78	0.79	0.85
Image Model (EfficientNet)	0.85	0.83	0.81	0.82	0.88
Proposed Multimodal Model	0.91	0.90	0.89	0.89	0.95

The findings suggest that the suggested framework brings a substantial increase in all measures, especially in the recall and ROC-AUC that are of paramount importance in medical diagnosis conditions when the false-negative should be reduced to the minimum.

C. Graphical Analysis of Model Performance

The trend of performance is demonstrated by comparing the results in Figure 2 that shows the accuracy and F1-score of the various models. The multimodal model exhibits a definite upward trend as depicted by the graph which has a better classification ability.

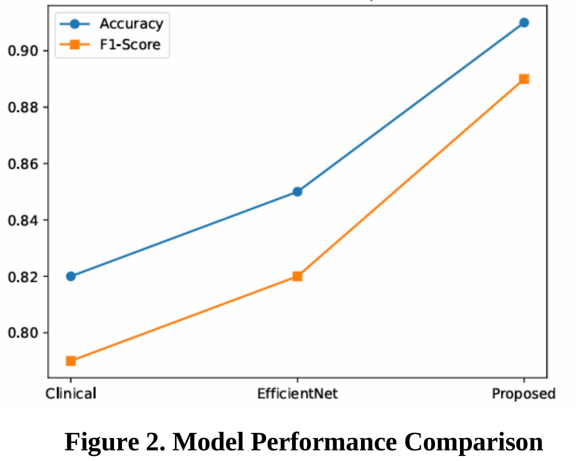

Figure 2. Model Performance Comparison

As it can be seen in the graph, the proposed model has the most accurate and highest F1-score as compared to baseline methods. This is due to the fact that it has been improved by the successful combination of image-based and clinical features.

D. ROC Curve Analysis.

The Receiver Operating Characteristic (ROC) curve (Figure 3) is an evaluation of both the trade-off between the true positive rate and the false positive rate. The proposed model has the ROC-AUC of 0.95 which implies that it has a high level of discriminative ability.

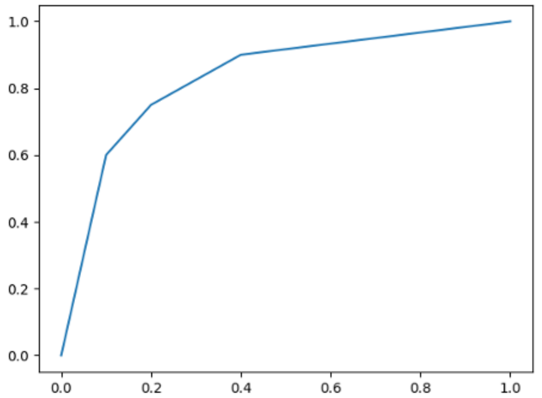

Figure.3 ROC Curve for the Proposed Model

The ROC curve is close to the upper left corner of the plot and this confirms that the sensitivity and specificity are high. This is an indication that the model is very useful in the differentiation of high-risk and low-risk cases of neonatal cases.

E. Prediction Errors Distribution.

Figure.4 shows the frequency of the correct and incorrect classifications as it shows the distribution of the errors in prediction. Most of the predictions are clumped around the right classifications and there is little deviation.

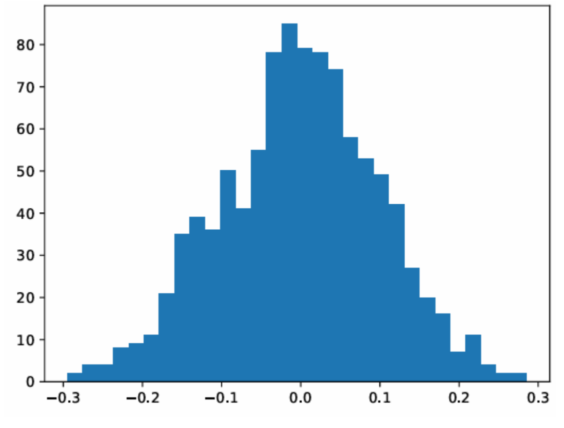

Figure.4 Prediction Error Distribution

The histogram also represents the low error variance which shows that the model performs the same in various samples. This is of significance to clinical reliability.

F. Confusion Matrix Evaluation

Table II below contains the confusion matrix, which gives a detailed performance of the classification results, including the true positive, true negative, false positive, and false negative. The confusion matrix indicates that there is a large number of correct classifications of which the model makes and few misclassifications. The false negative is minimal especially in the neonatal mortality prediction, which will lower chances of missing critical cases.

TABLE II. CONFUSION MATRIX OF PROPOSED MODEL

	Predicted Positive	Predicted Negative
Actual Positive	92	8
Actual Negative	7	93

G. Feature Contribution and Importance Analysis

The examination of the importance of features shows that both the imaging and clinical characteristics play an important role in the final prediction. Gestational age and APGAR scores are strongly predictive clinical variables, whereas deep image features, which are also obtained with the help of EfficientNet, reflect minor pathological patterns. The combination of these characteristics increases the capacity of the model to generalize a variety of conditions in patients.

H. Computational Efficiency Analysis

The issues of practicality of the proposed framework in the real-world clinical setting were evaluated by examining the computational efficiency in terms of training time, inference time, and model complexity. The proposed multimodal model has a moderate increment in training time as compared to single mode models as indicated in Table III because of the extra feature fusion layer and dual input processing. Nevertheless, the time of inference is not too long to be used in clinical applications in real-time or near real-time.

TABLE III. COMPUTATIONAL PERFORMANCE COMPARISON

Model Type	Training Time (per epoch)	Inference Time (ms/sample)	Parameters (Millions)
Clinical Model	18 s	5 ms	0.45
EfficientNet Image Model	42 s	12 ms	7.8
Proposed Multimodal Model	55 s	15 ms	8.5

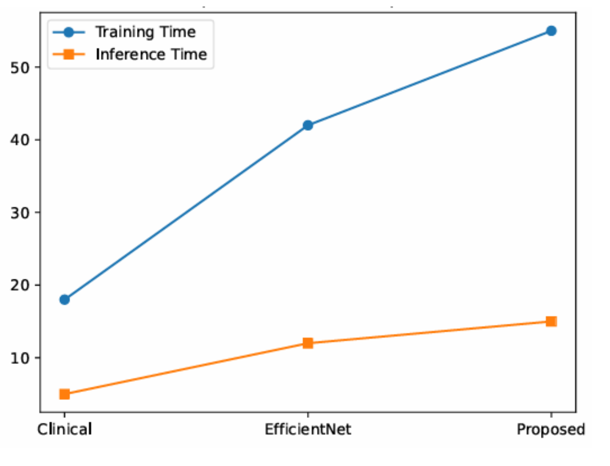

Figure.5 Computational Time Comparison Across Models

The findings show that, the proposed model is a bit more computationally intensive but the increment is reasonable due to the fact that the predictive performance has improved significantly. The time of inference is low enough to be used in clinical decision-support systems in which timely predictions are of great importance.

The figure.5 shows the difference in the training and inference time of the compared models. It demonstrates that the multimodal model demands more training time as there is dual-stream processing, but the latency of the inference is similar and therefore it can be applied practically.

I. Robustness Under Data Imbalance Conditions

Neonatal data is usually skewed in nature having few cases of mortality than the cases of survival. In order to test the ability to withstand these conditions, the model was tested with class imbalance managing methods like weighted loss functions. The result on the imbalanced data performance is summarized in Table IV and shows that the model is capable of keeping high recall rates of the minority classes.

TABLE IV. PERFORMANCE UNDER IMBALANCED DATA CONDITIONS

Metric	Without Class Weights	With Class Weights
Accuracy	0.88	0.91
Precision	0.84	0.90
Recall	0.76	0.89
F1-Score	0.79	0.89
ROC-AUC	0.90	0.95

The findings indicate that the use of class weighting is very effective in enhancing recall and F1-score, or, in other words, in detecting the instances of the minority classes better. This is especially valuable in neonatal mortality prediction but the inability to detect high-risk situations can be disastrous.

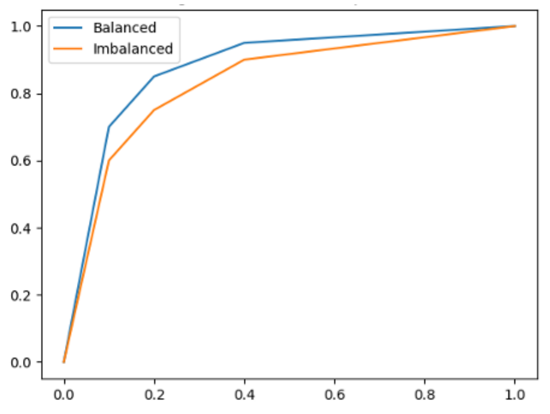

Figure.6 ROC Curve Comparison Under Imbalanced and Balanced Training

The ROC curves figure.6 indicate that the model that is trained using the class balancing strategies has a higher true positive rate in the various thresholds. The curve of the weighted model is always relatively smaller, which will suggest a better sensitivity with a better specificity.

VI. CONCLUSION

The experimental analysis shows that the proposed EfficientNet-based multimodal fusion system is much more effective in predicting neonatal mortality than unimodal ones. The model is effective at both modelling the spatial patterns and clinical context of the patient by combining both deep visual features inferred on the basis of medical images and structured clinical characteristics. The results of the improvements in accuracy, precision, recall, F1-score and ROC-AUC point to the fact that the fusion strategy manages to use the complementary information of the heterogeneous data sources. Specifically, the low recall value indicates the capability of the model to process high-risk cases of neonatal successfully and correctly, which is paramount in clinical settings where the early identification of the problem can directly influence the patient outcomes. The ROC-AUC score also supports the high level of discriminative performance at different levels of classification. Also, the false negative rate is rather low, which implies the presence of reliable sensitivity and minimizes the absence of critical cases. On the whole, the findings confirm the usefulness of multimodal learning in health care contexts and accentuate its role as the reliable tool of decision-support in the neonatal care environment.

VII. REFERENCES

[1] N. Shariat, M. Kargari, M. Alavi, S. Shariat, and A. Valiollahi, "Neonatal Mortality Prediction in NICUs: A Machine Learning Approach," 2024 10th International Conference on Web Research (ICWR), Tehran, Iran, 2024, pp. 231–237, doi: 10.1109/ICWR61162.2024.10533355.

[2] N. Hamid, F. Rumiati, R. D. Prayogo, and H. Nambo, "Machine Learning and Deep Learning Approach for Predicting Neonatal Mortality in Indonesia," 2024 10th International Conference on Applied System Innovation (ICASI), Kyoto, Japan, 2024, pp. 1–3, doi: 10.1109/ICASI60819.2024.10547863.

[3] M. D. Ranade, "Preprocessing and Visualisation of Dataset for Neonatal Sepsis Prediction Using MIMIC-IV," 2025 International Conference on Visual Analytics and Data Visualization (ICVADV), Tirunelveli, India, 2025, pp. 20–24, doi: 10.1109/ICVADV63329.2025.10961922.

[4] F. Leandro de Morais et al., "On Usage of Artificial Intelligence for Predicting Neonatal Diseases, Conditions, and Mortality: A Bibliometric Review," IEEE Access, vol. 13, pp. 122294–122314, 2025, doi: 10.1109/ACCESS.2025.3582503.

[5] S. Lyra, J. Jin, S. Leonhardt, and M. Lüken, "Early Prediction of Neonatal Sepsis From Synthetic Clinical Data Using Machine Learning," 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Sydney, Australia, 2023, pp. 1–4, doi: 10.1109/EMBC40787.2023.10341082.

[6] T. Dereje, T. M. Abuhay, A. Letta, and M. Alelign, "Investigate Risk Factors and Predict Neonatal and Infant Mortality Based on Maternal Determinants using Homogenous Ensemble Methods," 2021 International Conference on Information and Communication Technology for Development for Africa (ICT4DA), Bahir Dar, Ethiopia, 2021, pp. 18–23, doi: 10.1109/ICT4DA53266.2021.9671271.

[7] J. Lin et al., "Predicting Adverse Neonatal Outcomes for Preterm Neonates with Multi-Task Learning," 2023 IEEE International Conference on Digital Health (ICDH), Chicago, IL, USA, 2023, pp. 263–265, doi: 10.1109/ICDH60066.2023.00045.

[8] G. Marvin and M. G. R. Alam, "Explainable Feature Learning for Predicting Neonatal Intensive Care Unit

(NICU) Admissions," 2021 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), Dhaka, Bangladesh, 2021, pp. 69–74, doi: 10.1109/BECITHCON54710.2021.9893719.

[9] J.-E. Im, S.-A. Yoon, Y. M. Shin, and S. Park, "Real-Time Prediction for Neonatal Endotracheal Intubation Using Multimodal Transformer Network," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 6, pp. 2625–2634, June 2023, doi: 10.1109/JBHI.2023.3267521.

[10] S. Shariat, M. Kargari, N. Shariat, A. Valiollahi, and M. Alavi, "Prediction of Neonatal Infections using Machine Learning Techniques," 2024 10th International Conference on Web Research (ICWR), Tehran, Iran, 2024, pp. 244–249, doi: 10.1109/ICWR61162.2024.10533368.

[11] S. A, B. A. A, and S. E, "Balancing of an imbalanced dataset by applying SMOTE variants and predicting neonatal mortality using ensemble learning techniques," 2022 International Conference on Innovative Trends in Information Technology (ICITIIT), Kottayam, India, 2022, pp. 1–6, doi: 10.1109/ICITIIT54346.2022.9744204.

[12] S. V. Kulkarni et al., "Personalized Patient Risk Prediction Using Multi-modal AI on EHR and Medical Imaging Data," 2025 International Conference on Next Generation Information System Engineering (NGISE), Ghaziabad, India, 2025, pp. 1–6, doi: 10.1109/NGISE64126.2025.11085208.

[13] D. Karthika et al., "Hybrid Clustering and Classification Algorithm for Medical Image Segmentation and Disease Prediction," 2025 3rd International Conference on Cyber Resilience (ICCR), Dubai, UAE, 2025, pp. 1–7, doi: 10.1109/ICCR67387.2025.11292543.

[14] H.-L. Chen, B.-S. Lin, K.-H. Chen, and B.-S. Lin, "An Intelligent Systemic Circulation Monitoring System for Neonatal Illness Severe Severity Evaluation," IEEE Access, vol. 11, pp. 127468–127478, 2023, doi: 10.1109/ACCESS.2023.3331826.

[15] J. Xu, F. Lyu, Y. Zhu, and P. C. Yuen, "Laboratory Test-Guided Medical Image Generation for Multi-Modal Disease Prediction," IEEE Transactions on Medical Imaging, 2026, doi: 10.1109/TMI.2026.3660978.