Cardiotocogram Data Generation using CT-GANs for Enhancement of Performance metrics used for Classification Models

A Venkata Sriram [1,2*], *P Rajesh Kumar*[3], and *L Alekhya*[4]

[1] Lecturer, Electronics and Communication Department, Government Polytechnic, Parvathi Puram, Andhra Pradesh, India
[2]Electronics and Communication Department, Andhra University College of Engineering, Visakhapatnam, India
[3]Professor, Electronics and Communication Department, Andhra University College of Engineering, Visakhapatnam, India.
[4]Assistant Professor, Electronics and Communication Department, Lendi Institute of Engineering and Technology, Vizianagaram, Andhra Pradesh, India.

Abstract. Autonomous Fetal distress classification is crucial to monitor the Cardiotocogram (CTG) signals for fetal well-being and ensure timely medical intervention during pregnancy. However, the imbalanced nature of the existing CTG datasets impose significant challenges in developing autonomous models. To address this issue, the study provides a novel approach in generating synthetic CTG data using Conditional Tabular Generative Adversarial Networks (CT-GANs). By augmenting the existing CTG dataset with high-quality synthetic samples from GANs, the performance metrics are enhanced with this approach. The obtained results of the synthetic CTG data classification with algorithms namely k-NN and Random Forest with and without Grey Wolf optimisation are compared with the same classification models of Imbalanced CTG data. Hence the approach of Random Forest model with Greywolf optimization for CT-GANs generated balanced data outperforms most widely used techniques with an Overall accuracy of 94.49%, mean sensitivity of 95.23%, mean specificity of 97.7%, mean precision of 95.31%, mean MCC of 91.68% mean Kappa score of 91.63%, Weighted F1-score of 94.50% and averaged AUC-ROC of 99.40% with Random Forest algorithm for multiclass classification of CTG synthetic data.
Keywords: *Cardiotocograms (CTG), Conditional Tabular Generative Adversarial Networks (CT-GANs), Greywolfoptimization, Mathew's Correlation Coefficient (MCC), Kappa score, k-NN, Random Forest.*

1 Introduction

Development of the fetus is very important for the enrichment of the future generation. When a baby is born too early before gestational period called as preterm labour, it may lead to the death of the fetus. To eliminate such scenario or to detect any abnormality in the fetus, monitoring of the Cardiotocogram (CTG) signal play a cruicial role. This helps in monitoring the fetal heart rate and uterine contractions. Based on CTG signal early diagnosis can be handled. To detect the abnormality without any human intervention, automation is done to differentiate normal and abnormal condition of the fetus based on CTG signals. With the advent of technologies in the recent years many machine-based algorithms are used for automating the detection systems for early recognition and diagnosis of the abnormality condition.

CTG signal interpretation follows some specific guidelines given by the FIGO. According to these guidelines, the CTG traces are categorised into three classes namely Normal, Suspicious and Pathological. To categorize into three states the CTG trace has to be analyzed based on baseline fetal heart rate, variability, accelerations, decelerations and the pattern of uterine contractions. Typically, fetal heart rate ranges between 110 to 160 beats per minute (bpm). Any deviation from the normal values of these parameters lead to abnormality condition. Hence these parameters are used to determine the fetal well-being and guide the clinical decisions to be made during the pregnancy and labor.

CTG dataset from UCI repository holds the samples of the three classes with the mentioned parameters taken digitally using Sisporto 2.0 [1] software. This repository is open access and widely used by most machine learning engineers across the globe. This CTG dataset is imbalanced do not contain any null values and can be directly applied to data preprocessing technique. Once data cleaning and normalization is done, the samples of each class are given to CT-GANs for generating the most correlated data with minimum mean square error to make the CTG dataset balanced. GANs are special architectures used for generating the synthetic data which mimics the original data samples. It is constructed with two neural networks Generator and Discriminator both are vividly used to generate the samples. This balanced CTG data is given to Greywolf algorithm which is a Swarm based Metaheuristic optimisation to select the most appropriate features. The balanced data with most relevant features is given to classification algorithms namely k-NN and Random Forest algorithm with bagging method to classify the three categories present in the dataset. The proposed method with RF model outperforms in terms of not only accuracy and F1 score but also in sensitivity, specificity, and precision along with mean MCC, Kappa score and average AUC when compared to other traditional classification methods on CTG data. The block diagram of the proposed methodology is shown in the Fig. 1 below.

The following is the entire outline of the sections of this work, related documentation and literature are given in the section II. Section III include the data

* Corresponding author: avsriram406@gmail.com

 https://doi.org/10.1051/e3sconf/202459108007

preprocessing, feature selection and classification tasks applied on the CTG data considered. The results and the experimental analysis are given in the Section IV and Conclusion and future work are discussed in the section V.

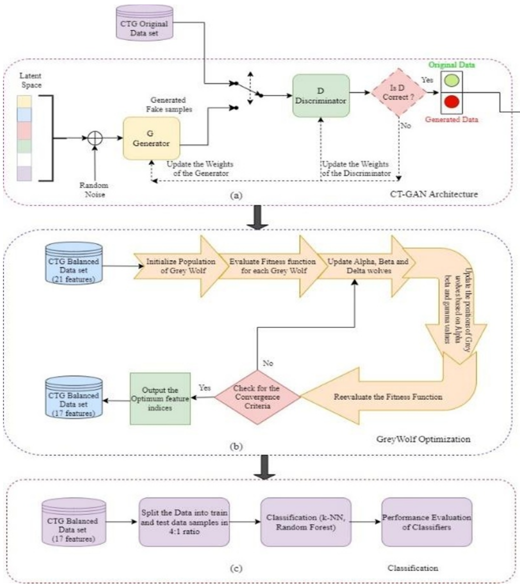

Fig. 1. Proposed Methodology

2 Related Works

Cardiotocogram is the most commonly usesd dataset to analyze the fetal health condition. Various works are carried out to detemine the fetal well-being. Rongdan Zeng et al. in [2] extracted Time-Frequency features using Wavelet Transforms on CTG signals and obtained more effective results with sensitivity: 85.2%, specificity: 66.1% with ECSVM classifier. Especially, the AUC-ROC curve (0.77 vs. 0.64) is suggestively increased with the ECSVM vs. SVM.

N. Chamidah et al. in [3] reconstructed the features of CTG data with k-means clustering and classified the data with an accuracy of 90.64% using SVM. T. H. Dang et al. in [4] used SVM and k-NN classifier models and achieved an accuracy of 92% and 89% respectively in detecting the abnormal cardiotocogram signal.

J. Piri et al. in [5] employed the SMOTE based oversampling technique to deal with discrepancy in the CTG data and distance-based multi-objective Ant lion optimization (MOALO-CD) is used as Feature selection method to boost the performance of the various classifiers. V. Shukla et al. in [6] implemented Random rest, XG Boost classifiers on CTG data and identified 356 and 358 samples respectively by each out of 377 total samples with an accuracy of 94%. Y. Salini et al. in [7] classified cardiotocogram data using various machine learning models and achieved an overall accuracy of 93% for Random Forest model, 90% for k-Nearest Neighbour and Gradient Boosting Classifier and 89%, 85% and 81% for logistic regression, decision tree and Support vector classifier respectively.

R. Hephzibah et al. in [8] used SmoteENN an oversampling and undersampling technique to balance the data samples of CTG and achieved an accuracy of 93% for Random Forest classifier. S.P. Potharaju et al. in [9] has used SMOTE (Symmetric Minority Over Sampling Technique) on imbalanced CTG dataset to make it balanced and quicken the classification performance of various classifiers.

Suman, J. V et al. in [10] has implemented device modelling using FinFET for application of machine learning with optimization techniques. This can be used to implement the real time acquisition system for sensing and recording the CTG data. Alekhya, L. et al. in [11] has converted the ECG signal data to 2-D scalograms images and applied to VGG16net an extension of convolutional neural network and achieved an accuracy of 96.94% with RF algorithm in classifying multiclass labels of Electrocardiogram signals.

3 Materials and Methods

The fetal well-being analysis is major criteria for normal delivery to mother. Any disturbance in the fetal health condition can be risky to both mother and the child. Early recognition and diagnosis accordingly are important to eradicate fetal deaths. The proposed methodology given in Fig. 1. uses CTG dataset available online free access data to detect automatically the abnormality in the fetal data using machine learning algorithms.

3.1 Data Preprocessing

The dataset used is obtained from UCI machine learning repository available in free access online [12]. This dataset is attained from Sisporto 2.0 digital acquisition system and made available for free access. It majorly contains Fetal Cardiotocogram signal extracted feature values with measurements from fetal heart rate (FHR) and Uterine contractions (UC) labelled into three classes namely Normal (N), Pathological (P) and Suspect (S) by obstetricians. The data contains around 2126 records with 1655 Normal, 295 Suspect and 176 pathological. The entire dataset is of the size 2126 rows and 22 columns of which 21 are features and 1 column contains the three labels of the data.

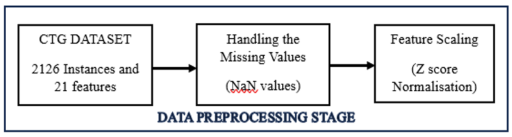

Fig. 2. CTG Dataset Preprocessing

The steps involved in the data preprocessing stage is shown in the Fig. 2. To apply these values to the proposed model, initially the missing values must be handled by either removing them or replacing them with mean of the feature column where that value is present.

Once the missing values are handled, the feature scaling is applied using z-score normalisation to improve convergence speed of gradient descent algorithms and enhances the potential of the machine learning models. After proper preprocessing, the data set can now be applied to the proposed methodology as mentioned in the Fig. 1.

3.2 CTG Dataset Balancing with CT-GANs

Clearly the dataset considered is Imbalanced which means there are no equal number of records for each category of label. This kind of data provide bias towards majority class which inturn gives inadequate classification accuracy of the model. For this kind of data mean MCC and kappa score metrics give appropriate value in the classification of the model. To improve these performance metrics, data must be balanced. Various oversampling methods such as SMOTE [9] has been used to make the data balanced. In this work, a Generative Adversial Network namely CT-GANs for tabular data was proposed to generate the synthetic data which is most correlated with existing CTG data with least mean square error possible. This generated data is is combined with the original CTG data to make the dataset balanced to improve the performance metrics of the classifier models.

Conditional Tabular Generative Adversarial Networks popularly called as CT-GANs, a Generative AI model used to generate the synthetic data with relevance to the applied data. This generation is done using CTGAN Synthesizer from Synthetic Data Vault (SDV) a python-based library made available to public and is powered by DataCebo Inc. The synthesizer class can be imported from SDV package and can be applied to tabular data under single data category. The generated data is combined along with original CTG data to make it balanced. This data is applied to feature reduction stage.

3.3 Feature Reduction by Grey Wolf Optimizer

The Grey Wolf Optimization (GWO) algorithm is encouraged by the social hierarchy and chasing process of grey wolves. It has emerged as a powerful swarm-based metaheuristic algorithm for solving optimization problems, including feature selection tasks. GWO effectively models the dynamics of wolf packs, which allows it to explore and exploit the solution space efficiently, making it suitable for high-dimensional data like CTG. Fig. 3. given below depicts the process of grey wolf optimisation algorithm in feature selection process of CTG dataset.

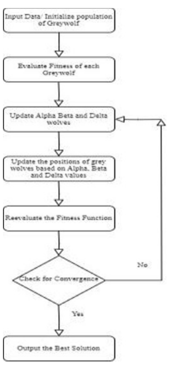

Fig. 3. Grey Wolf Optimiser process

The final position would be updated by alpha, beta and delta values. The entire process of the optimisation process is to select the most prominent features. It outputs the most optimum features out of 21 feature columns of the applied data.

3.4 Classification Models

The original CTG data and balanced data both with reduced and non-reduced form are applied to Machine learning classification models namely k-Nearest Neighbour and Random Forest with Bagging method to evaluate the performance of the models and compare the results of four datasets. Since the data holds multi class labels, all the performance metrics are evaluated based on one vs rest mode of the classification.

4 Experimental Analysis

Experiments were conducted on a computer having an Intel i5 microprocessor and 16 GB Ram and Nvidia GPU of 2 GB, while coding were carried out on MATLAB 2023b version software. The study was conducted on the CTG dataset of size 2126 x 22 which include 21 attributes of which 11 are discrete and 8 are continuous and 1 column indicates labels. Out of 2126 instances 1655 belongs to Normal, 295 to Suspect and 176 to pathologic classes respectively. Hence the data size becomes 2126 x 21 with 3 class classification columns of the data taken as Dataset D1 which indicates the original data.

The Imbalanced data D1 is applied to CT-GANs using Python SDV library to generate most relevant data and make the dataset balanced with an elapsed time of 240 seconds and is denoted as D3. This dataset is of size 5808 x 22, 1936 records for each category. Both the D1 and D3 datasets are applied to the Grey Wolf optimiser that outputs 17 most optimum features with elapsed time of 22.8 seconds for D1 and 135.63 seconds for D3 that which reduces to D2 and D4 data respectively. These datasets along with labels column are divided into train and test dataset in the ratio of 4:1 respectively.

Table. 1. Overall Performance Metrics

CTG Dataset	ML models	Overall Performance Metrics in Percentage			
		Overall Accuracy	Overall Sensitivity	Overall Specificity	Overall Precision
Imbalanced Original – D1	k-NN	88.94	72.74	87.57	84.67
	RF	90.82	79.76	90.16	85.19
Imbalanced Reduced (Greywolf)D2	k-NN	91.29	80.95	90.86	86.25
	RF	93.65	84.57	91.5	94.39
Balanced Original – D3	k-NN	92.17	92.15	96.15	92.15
	RF	93.98	93.81	97.01	94.00
Balanced Reduced (Greywolf)D4	k-NN	93.12	93.06	96.61	93.04
	RF	94.49	94.32	97.27	94.50

Table. 2. Overall Performance Metrics

CTG Dataset	ML models	Overall Performance Metrics in Percentage			
		Weighted F1 score	Mean MCC	Mean kappa	Average AUC
Imbalanced Original – D1	k-NN	89.53	68.01	67.04	92.78
	RF	91.24	74	72.82	87.66
Imbalanced Reduced (Greywolf) –D2	k-NN	91.60	75.70	75.47	93.05
	RF	94.12	82.86	81.70	92.78
Balanced Original – D3	k-NN	92.11	88.27	88.18	98.88
	RF	93.98	90.92	90.86	99.08
Balanced Reduced (Greywolf)-D4	k-NN	93.07	89.63	89.60	98.83
	RF	94.50	91.68	91.63	99.40

All the four datasets are applied to train the classification models k-NN and Random Forest and the test dataset of each are used to evaluate various performance metrics and the respective results of each dataset are compared and tabulated in the Table. 1. and Table. 2.

The experimental results show a best performance with RF model for dataset D4 which is balanced data using CT-GAN with Grey wolf optimiser for feature reduction. It gave an overall accuracy of 94.49% for Random Forest and 93.12% for k-NN algorithm which outperforms the models given by R. Hephzibah et. al in [8] that used SmoteENN model for balancing and achieved an accuracy of 93% for the same RF model. Y. Salini et. al in [7] achieved overall accuracy of 93% for Random Forest model, 90% for k-Nearest Neighbour. The method used achieved better accuracy for both the ML models.

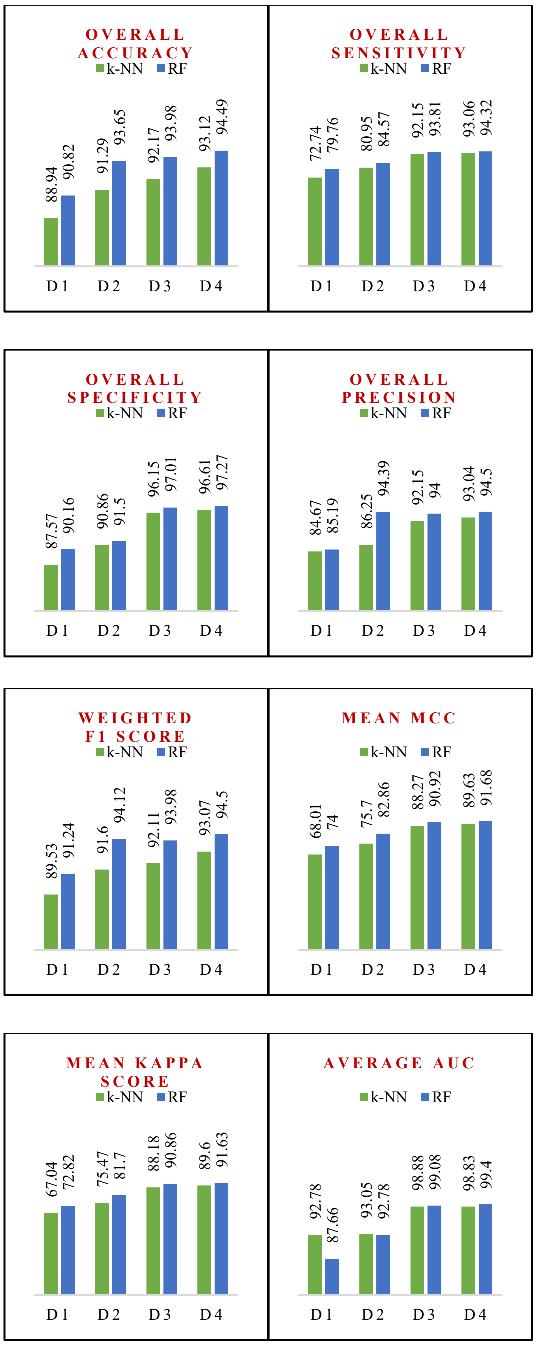

Fig 4. Experimental Results

Also, the other metrics such as MCC and Kappa score is 91.68% and 91.63% which is proportionally high for balanced data D4 when compared to Imbalanced data D1 as given in Fig. 4 which denotes the experimental results. These metrics show the importance of balancing the data. The confusion matrix of the k-NN model and RF model with Reduced Balanced data D4 is given in the Fig. 5 and Fig. 6 respectively below.

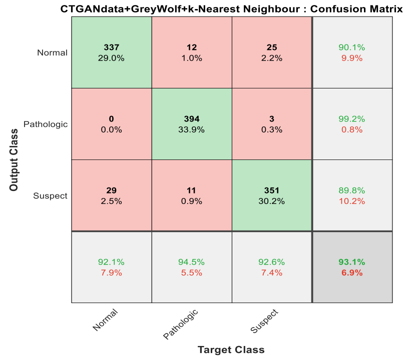

Fig. 5. Confusion Matrix of k-NN model

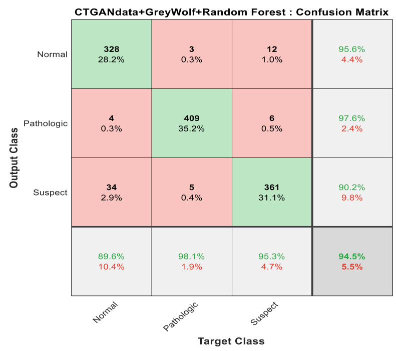

Fig. 6. Confusion Matrix of RF model

5 Conclusion and Future work

Cardiotocogram data analysis is majorly used to detect the fetal health if any abnormalities persist. The early detection and diagnosis of the abnormal condition of the fetus helps in reducing the fetal death rate. The CTG signal which is Imbalanced dataset is applied to CT-GANs and make the data balanced. This balanced dataset is applied to Greywolf optimiser to reduce the feature set. The reduced balanced dataset is applied to k-NN and RF classifiers to detect the N, S, P classes among the data and evaluate the performance metrics.

Experimental findings of the discussed methodology show that balanced reduced data when given to classifiers outperforms other state of the art techniques. In terms of implementing real time system IoT based methods can be used to build such detection system. A quantum-based machine learning algorithms can be used in future to reduce the elapsed time of the model.

The Data that is used in the work is available online in UCI machine learning repository with free access.

Author's contribution includes Data preprocessing, Feature reduction, Classification models, programming in Python and Matlab along with writing and editing the manuscript.

References

1. Ayres-de Campos et al. SisPorto 2.0: a program for automated analysis of cardiotocograms. J Matern Fetal Med. 2000 Sep-Oct;9(5):311-8.
2. Rongdan Zeng et al. Cardiotocography signal abnormality classification using time-frequency features and Ensemble Cost-sensitive SVM classifier, Computers in Biology and Medicine, Volume 130, 2021, 104218, ISSN 0010-4825.
3. N. Chamidah et al. Fetal state classification from cardiotocography based on feature extraction using hybrid K-Means and support vector machine, 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2015, pp. 37-41, doi: 10.1109/ICACSIS .2015.7415166.
4. T. H. Dang et al. Classification of Fetal Status from Cardiotocogram Data by Using Machine Learning, 2023 12th International Conference on Control, Automation and Information Sciences (ICCAIS), Hanoi, Vietnam, 2023, pp. 352-357, doi: 10.1109/ICCAIS59597. 2023.10382381.
5. J. Piri et al. Multi-objective Ant Lion Optimization Based Feature Retrieval Methodology for Investigation of Fetal Wellbeing, 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2021, pp. 1732-1737, doi: 10.1109/ICIRCA51532.2021.9544860.
6. V. Shukla et al. Machine Learning-Based Fetal Assessment: A Non-Intrusive Approach, 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT), Vellore, India, 2024, pp. 1-6, doi: 10.1109/AIIoT58432. 2024.10574555.
7. Y. Salini et al. Cardiotocography Data Analysis for Fetal Health Classification Using Machine Learning Models, in IEEE Access, vol. 12, pp. 26005-26022, 2024, doi: 10.1109/ACCESS .2024.3364755.
8. R. Hephzibah et al. An ensemble technique of SmoteENN and random forest for classification of the cardiotocography data, International Conference on Computer Vision and Internet of Things 2023 (ICCVIoT'23), Coimbatore, India, 2023, pp. 152-157, doi: 10.1049/icp.2023.2869.
9. Sai Prasad Potharaju et al. Data mining approach for accelerating the classification accuracy of cardiotocography, Clinical Epidemiology and Global Health, 7(2):160–164, 2019.
10. Suman, J. V et al. A Review of Machine Learning-Based Device Modeling and Performance Optimization for FinFETs, *International Journal of Scientific Research in Engineering and Management (IJSREM)* 7 (2023): 1-10.
11. Alekhya, L. et al. A new approach to detect cardiovascular diseases using ECG scalograms and ML-based CNN algorithm, 2023. International Journal of Computational Vision and Robotics/ Inderscience publishers. DOI: 10.1504/IJCVR. 2022.10051429
12. Campos,D. et al. Cardiotocography. UCI Machine Learning Repository. (2010). https://doi.org/10.24432/C51S4N.