

A Deep Learning Based Approach for Detecting Fetal Stress Using CTG Signal

S. M. Seeni Mohamed Aliar Maraikkayar*^{ORCID}, R. Tamilselvi^{ORCID}, M. Parisa Beham^{ORCID}

Department of Electronics and Communication Engineering, Sethu Institute of Technology, Kariapatti 626115, India

Corresponding Author Email: seenimohamedali@sethu.ac.in

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.430113>

ABSTRACT

Received: 17 February 2025

Revised: 16 October 2025

Accepted: 21 January 2026

Available online: 28 February 2026

Keywords:

fetal stress, deep learning, VGGNet, GoogleNet, AlexNet, accurate

Detecting fetal stress accurately and promptly is critical for preventing adverse perinatal outcomes. Cardiotocography (CTG) is a standard monitoring technique widely used during pregnancy to assess fetal health by recording fetal heart rate (FHR) and uterine contraction (UC) signals. However, traditional methods of interpreting CTG signals are often subject to observer variability and may result in delayed or inaccurate diagnoses. Also, traditional feature-based classification methods require manual reading of the morphological characteristics from the FHR curve to capture the signal's features. This process is time-consuming, costly, and prone to calibration bias. To address this, we propose a deep learning-based approach for detecting fetal stress using CTG signal data. This study introduces a novel Enhanced-VGGNet-based method for classifying FHR signals, which reduces human error. This algorithm learns superficial features directly from the data, using AlexNet and Google enabling the real-time diagnosis of FHR data. The dataset employed in this study is from the Kaggle dataset which includes numerous attributes for feature extraction. The proposed enhanced VGGNet serves as the classification model, categorizing the signal into Normal, Suspicious, and Pathological classes. This automated approach outperforms traditional machine learning classification methods, which struggle to incorporate a greater number of attributes for more accurate feature extraction. The efficiency of the proposed methodology is then assessed using several performance parameters. It provides sensitivity, specificity, and accuracy of 95.8%, 91.2% & 96.6% respectively which shows its outperformance compared with other methods.

1. INTRODUCTION

The modern society is also facing the ever-growing demands of health needs and the ever-growing standards of living, and the same can be said of the future generations. The population of individuals who would desire quality services for their children during pregnancy and as delivery babies would increase. To facilitate healthy growth of the foetus and detection of potential problems at their earlier stages of development, parents will need to have real-time physiological records of the foetus, its heart rate, and the intensity of its contractions during the development [1]. Psychological or physiological stress could cause significant variance in the physiological parameters of some hospitalized patients that are not present in their natural conditions.

Recent research has noted that out of every three individuals who are provided with home blood pressure equipment, their measurements are expected to be in accordance with their clinical diagnosis, particularly in individuals who already have high blood pressure and those who are at risk of developing heart disease conditions [2]. The fact that physiological indicators of blood pressure, blood oxygen, and EKG of pregnant women and fetuses were measured was important due to the fact that Professor Johnson, through his remote fetal monitoring, had established that the determination of the physiological indicators was essential. The techniques would

also be applicable in determining the position of the baby adequately and providing a comfortable space for the mother. Therefore, the application of remote fetal monitoring technology would be applicable at home during the perinatal stage effectively to improve the quality of care during the perinatal period.

The paper will capture real-time fetal health and abnormal fetal development based on a remote fetal heart monitoring system, which involves neural networks to classify the data of fetal heart rate monitoring. The data collection is long and cumbersome, and the qualifications of professional medical specialists are required due to the professional specificity that can be offered in classifying the data on classical fetal heart rate (FHR) as the process is time-consuming. In practice, the investigation of these FHR attributes is time- and labour-intensive, and therefore it cannot be handy to provide immediate medical outcomes. This is retrogressive in medical diagnosis. However, the recently developed neural network models have the ability to identify valuable information in fetal heart rate data in the form of the corresponding algorithms automatically and generate successful classifications with the actual results. In this way, we research the problem of data prediction of fetal cardiac monitoring through neural networks.

However, the problems encountered in the use of fetal heart rate monitoring data are excessive noise, large amounts of

information, and high error rates, which will eventually not result in positive outcomes of fetal heart rate classification. The issue that ought to be addressed in this study is the manner in which to preprocess data and come up with a model of a neural network that would suit the data. In this study, different neural network models of the FHR data categorisation are applied. It is the most applicable contribution to the identification of the most appropriate model of FHR data classification because it compares and contrasts the performance of a large number of models, summarises the findings, and, based on the optimal model, informs medical practice.

2. RELATED WORKS

There are means for the selection of the features of the Cardiotocography (CTG) signals and their classification based on their nature. Algorithms with different natures produce different results in terms of efficiency. Different classification algorithms have been applied in the past to classify cardiotocographic detection. The paragraph is relevant to the present study as it talks about current methods of deep learning-based neural network models to detect fetal stress. An MCNN suggested by Petrozziello et al. [3] was more precise when predicting cord acidemia at birth compared to Clinical Practice and prior computerised methods. Zhao et al. [4] proposed an 8-layer deep convolutional neural network (CNN) model to predict fetal acidemia automatically. The determination of whether the fetal heart rate (FHR) is normal or abnormal should be done with a CNN model that is weighted with a voting scheme suggested by Liang and Li [5]. This was initially accomplished by Gao and Lu [6], who obtained the baseline properties of the fetal heart rate, and later, fetal heart rate segmentation was carried out in the form of a Long Short-Term Memory network. Fasihi et al. [7] designed a 1-D convolutional neural network with a new shallow architecture to improve the quality of fetal state analysis. The idea of predicting the classes of fetal states with the help of artificial intelligence was suggested by Irajil et al. [8]. The deep learning (DL) model adopted by Mohammad et al. [9] is a CNN model to identify infants with a possible low Apgar score.

In an attempt to determine the fitness of the fetus, Mohammed et al. [10] concentrated on the observation of the procedures that are used on fetal heart rate variability (FHRV) signals that are monitored during pregnancy. Ogasawara et al. [11] propose a deep neural network-based system (CTG-Net) to determine whether a fetus is healthy or not. Boudet et al. [12] propose a typical signal processing framework, where their methods are applied to fetal heart rate analysis, including a signal viewer for expert annotation and an evaluation procedure using multiple performance metrics. In addition to determining features of a guideline, the features were also isolated in different domains, and the most important features were chosen and used in the classification, and the creation of an automated model that would analyze the CTG signal took the next step further [13, 14]. It is an automated machine learning method of categorisation. In the study by Cömert et al. [15], a time-frequency representation of the FHR signal was demonstrated through the creation of a specific machine learning prognostic model, which may be used to determine fetal distress. In another study, Cömert [16] applied a traditional machine learning method to assess the diagnosis of

fetal distress based on FHR collected during the first and the second stage of labour. Bursa and Lhotska [17] developed a convolutional neural network to detect CTG signals. To provide the CNN model with a continuous wavelet transform family known as the complex Morlet wavelet, they provided it with a time-frequency image of FHR data. In the FHR data obtained in stage I labour, the classification was determined to be 94.1%. The AlexNet model has enabled Cömert and Kocamaz [18] to make use of transfer learning to come up with a deep convolutional neural network. They were supported by the high classification accuracy of 94.32% that was obtained through the application of the STFT to generate a time-frequency image of the FHR signal and input it into the AlexNet model. Zhao et al. [4] created another deep learning model. They used the CWT families DB and sym to input into the deep learning model to create the input images. The risk in both the mother and the unborn infant during pregnancy was determined by a ResNet-50 convolutional neural network that was optimised [19]. They provided the deep learning model with numerous time-domain characteristics, which were learned following the use of a conventional machine learning method. After the extraction of features, an optimised ResNet was then used to classify them with a 94.63% accuracy. As described in the literature related to it, the input image to a deep learning model was created by using a set of approaches to express the time-frequency of the FHR signal; this way, greater classification performance was obtained. The scanned pictures of cardiotocograms (CTGs) were used even in certain publications [20, 21], providing 2D images and then applied to deep learning models.

Deep learning was applied in the study by Frasch et al. [20] to detect fetal distress that could be prevented with the help of CTG signal images. Their study was on personal information, and the accuracy of classification was 93.65%. A deep learning network to classify scanned CTG image signals was created in the study by Romagnoli et al. [22]. They reported 0.73 under the receiver operating characteristic curve. The other article by Saini et al. [21] is an article on a deep convolutional network for the recognition of fetal distress. They identified three classes of fetal distress based on the image of CTG signal scanning, and fetal distress was divided into normal, mild, and severe. Most of the corresponding literature based on deep learning algorithms showed massive potential in predicting fetal distress, but in the study by Ogasawara et al. [11] and the article by Saini et al. [21], the accuracy was lower. The best results are in the study by Cömert and Kocamaz [18] and the study by Zhao et al. [4], which can be argued to have been implemented, where the time-frequency analysis of FHR at various stages of labour was conducted. In addition, the signal was pre-classified graphically in the study by Frasch et al. [20], and it is therefore very susceptible to false positives. The performance analysis is an issue, and the model presented in the study by Parvathavarthine and Balasubramanian [19] is not fully validated yet, so the absence of significant details in the above-mentioned works was addressed with the help of positive motivational techniques and the importance of key factors to govern the accurate determination of fetal distress in the FHR signal of the CTG. The key contribution of the current study is the following:

In the first stage of preprocessing the raw FHR signal, the missing values and unwanted artefacts were removed.

1. The second phase was the transformation of 1D FHR data into a 2D image.

2. The second step is the AlexNet- and GoogleNet-based feature extraction.
3. The classification step was characterized by the training and testing of a proposed improved version of the VGGNet model.
4. An extensive experiment with the FHR signal was conducted in the last step.
5. The study employs a new classification algorithm to monitor fetal heart rate using fewer features based on the data of other studies and backgrounds that will be verified.

The suggested work shows that it can be classified. Such deep learning and machine learning models are able to enhance the accuracy of the results.

3. SYSTEM ARCHITECTURE

Different deep learning frameworks, including GoogleNet and an AlexNet-based model are proposed in the suggested study so that to extract the data and estimate fetal stress using the fetal heart rate. The above two methods are assessed on performance and then they are classified.

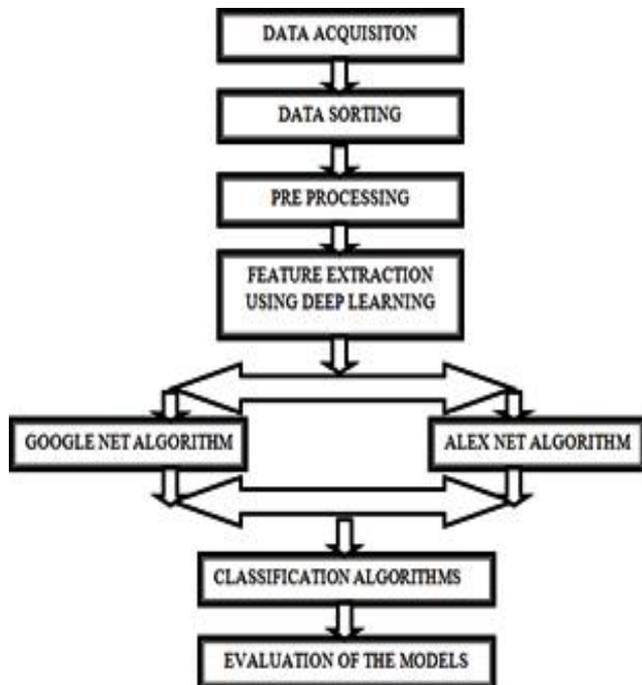


Figure 1. Overall flow diagram

The overall process map of the suggested process is depicted in Figure 1. Heart rate of the fetus is the first signal to be inputted. The provided signal of fetal heart rate is transformed into a matrix. The structured matrices are changed into images to be utilized as a deep learning source of feature

extraction. The obtained images are in the pre-processing stage. The wavelet transform does the denoising of the given image in this pre-processing. The image is subjected to a smoothing filter, which follows after the denoising stage and the image is passed to the feature selection block. There are two deep learning models utilized in the extraction of features, which are GoogleNet-based CNN Models or AlexNet-based CNN Models. After extracting the features, four algorithms are used to perform classification and these are Support Vector Machine, K-Nearest Neighbour, Decision Tree and improved VGGNet. The outcomes of the classification algorithms are subsequently tested.

3.1 Dataset and data sorting

In this work, the Kaggle dataset [18] is used. It has 2126 fetal CTGs and their related diagnostic characteristics. This dataset's occurrences have been categorised into 1655 regular classes, 295 suspicious classes, and 176 pathological classes. Each case has different attributes. The sample dataset is shown in Table 1. Preparing raw data into higher-quality data requires pre-processing. It is not possible to fully process the UCI ML [18] data for classification, as multiple structures need to be modified. Nominal data in the form of classes is necessary for specific classification techniques. In this step, the outliers are eliminated. Also, it is needed to convert part of the class properties for classification into nominal data. The following data sorting process was performed on the dataset to improve detection accuracy. Two independent datasets were used to ensure training and external validity. The model itself was trained and internally tested only on the public Kaggle CTG dataset, which had 2,126 samples (1,655 normal, 295 suspicious, 176 pathological). The external test on a secondary dataset in Pixel Scans Diagnostic Centre (Trichy, India), which contains 350 anonymised CTG images, was postponed to evaluate the generalisation on the actual data. Both datasets were not amalgamated; and an equal amount of preprocessing and normalisation was done in both. The ethical approval of the private dataset was obtained with the help of the Pixel Scans Institutional Ethics Committee (Approval ID: PX-CTG/21/IRB). All identifiers of patients were removed before analysis. The study was conducted as per the national biomedical research guidelines and the declaration of Helsinki. The FHR signal curve is missing a heart value, which is counted and the sample with a ratio of more than 10s is rejected.

Breakpoint detection on the FHR signal curve is done and samples whose consecutive breakpoints are greater than 30s are rejected.

- Repair of missing values in the curve of the FHR signal using linear interpolation.
- A heart rate that is five times less than 10 bpm (beats per minute) is considered to be a steady heart rate in the FHR signal curve that can be used to reduce noise.

Table 1. Dataset sample

Bval	astv	mvs	pal	mv1	HW	Hmi	Hma	hnp	hnz	Hmo	Hmea	Hme	Hvar	Hte	FH
123	73	0.5	43	2.4	64	62	126	2	0	120	137	121	73	1	2
132	17	2.1	0	10.4	130	68	198	6	1	141	136	140	12	0	1
133	16	2.1	0	13.4	130	68	198	5	1	141	135	138	13	0	1
134	16	2.4	0	23	117	53	170	11	0	137	134	137	13	1	1

3.2 Pre-processing

The FHR signals are pre-processed to eliminate noise. Here, pre-processing is performed using a wavelet transform. Then the de-noised signal was smoothed by a smoothing filter. The pre-processed signals are then analysed and converted into matrices. The converted matrices are transformed to form an image. The transformed images are fed into the feature extraction block. Wavelet denoising was performed using the Daubechies (db4) wavelet at level 3 decomposition with adaptive soft-thresholding ($\sigma = 2 \log n$) to remove high-frequency noise and preserve FHR morphology. After the wavelet reconstruction, the baseline drift was further stabilised by using Gaussian smoothing (kernel size = 5, 1.5). The hybrid method saved more than 95 per cent of the signal energy and raised the signal-to-noise ratio (SNR) to 28.3 dB, compared with 19.4 dB. A comparative analysis using Empirical Mode Decomposition (EMD) and the Short-Time Fourier Transform (STFT) revealed that wavelet-Gaussian preprocessing achieved the best morphology preservation, and it was therefore used in all subsequent analyses.

3.3 Feature extraction

A pre-processed data is then formed into an image to be used as a feature. Machine learning feature selection uses a first set of measured data to generate derived values or features that are supposed to be informative and non-redundant. This eases the stages of later learning and generalisation and in certain cases results in superior human interpretations. Dimensionality reduction is connected with feature selection.

The input data to an algorithm may be considered unnecessary or too large to manipulate, which may be narrowed down to a few values, or a feature vector (e.g., the precise value in feet and meters, or the repetition of the image in form of pixels). The stage of choosing the first characteristics is referred to as feature selection. The selected features are expected to contain the information that is relevant in the input data to complete the task one wants to achieve but with this reduced representation instead of using the whole original data. In this case, deep learning is applied to extract features. Two distinct deep learning methods are being used: GoogleNet and AlexNet that are both grounded on CNNs.

The GoogleNet Architecture has a total of 27 layers of pooling that form 22 levels. It has 9 linearly stacked inception modules. The inception modules have their terminus points linked to the global average pooling layer. One of the biggest contributions to neural network research that was made, especially in CNNs, was the Inception Network. Inception Networks has now been made available in three versions Inversion versions 1, 2 and 3. This original edition of the Inception network is known as GoogleNet. Overfitting can occur when there are several deep layers in constructing the network. We proposed GoogleNet design, which is a solution to this problem, provided on the basis of the idea of filters that can be of various sizes. This notion is simply extending the network and not enriching it.

The CNN-based AlexNet model was used in this work. An architecture of convolutional neural network is called AlexNet. AlexNet was made up of eight networks five convolutional layers in the first stage with max pooling layers in certain cases and three fully connected layers in the final stages. It also worked well in training as compared to tanh and sigmoid in the case of the non-saturating ReLU activation function. AlexNet is undoubtedly one of the simplest methods

of learning the principles and methods of deep learning. It is not a complicated architecture, by comparison with some of the state-of-the-art CNN designs of recent times. Features were extracted using GoogleNet and AlexNet and redundant features eliminated. This extracted feature is the input to the classification model then.

3.4 Classification using machine learning and proposed enhanced VGGNet

The proposed enhanced VGGNet architecture was used in this process and the findings were compared to three machine learning algorithms namely Decision Tree, K-Nearest Neighbours (KNN) and Support Vector machine (SVM). In a higher dimensional space, SVM defines hyperplanes as defined by the set of points at which the dot product with a particular vector is a constant; this space of vectors is orthogonal (and, thus, minimal) space defining a hyperplane. The distance that is perpendicular to the closest data points to the hyperplane is called the margin (both sides). The support vectors are the points that have the least margin distance. The process of grouping is used to compute margin value. In this case, margin is applied as an operator.

The Pattern recognition KNN is a method of classification that is based on the nearest neighbours of a new sample whose classification is based on a training data set and the nearest k neighbours. In KNN classification, a data is classified according to the closest data of the data. This proximity is achieved by giving the value of K which is put on the model. In case of $k = 1$ it would be moved over to an nearest class to 1. The categorization performed by K-NN is categorized as a local approximation and all the computations are delayed until a time when the user intends to test the function. The same is that in this technique distance is used to cluster the available data, this is to say that when the scale of the features utilized is radically different or when the features utilized in the procedure are not of the same physical quantities the normalization of the training information can greatly enhance the quality of the method.

The decision trees depend on various methods of establishing whether or not a node should be sub-divided into two or more sub-nodes. Production of the sub-nodes increases homogeneity of the physical sub-node. That is, it can claim that the purity of the node is improving in reference to the target variable. Results of the feature selection algorithms are used in order to create ten classes. This further divides into these ten groups which are further subdivided into three; Pathological, Suspicious and Normal. The chosen feature is then feed forwarded to the improved VGGNet.

There are four layers in the architecture of the proposed skip network of VGGNet as illustrated in Figure 2. It is a layered architecture. Figure 2(a) shows the created broadened skip network of broad semantic biomedical signal training. Figure 2(b) is the description of the properties of the network which were obtained at every output stage. Additional skipping module: The amount of feature map layouts of the final stage of production once all the other multi-scale filter coefficients are done is kept at 12 although the amount of feature maps produced by producing 1×1 convolution will also give the same number of feature maps as the input provided the number of channels on the input channel is C_{in} . The overall output characteristic maps of the skip module are therefore $(C_{in} + 12)$. Figure 2(c) depicts the upsampling and concatenation and convolution with a specific number of the number of output

feature maps and combining them to merge the generated feature maps of the extended skip module and the decoder. The algorithm of a combination of the skip module and feature maps is depicted.

One of the several VGG Net-like encoders is made up of a series of 3×3 convolution, batch normalisation, and a nonlinear activation. The encoder has been configured to apply max pooling to the spatial resolution of the feature maps right after every VGG-block (except the final one). The last two convoluted blocks (big kernels) of the encoder are designed to capture a bigger context, then there is batch normalisation, nonlinear activation and two additional convoluted blocks. In order to compute these huge convolutions, separable kernels are computed to minimize the number of computations. The activation used in the network is nonlinear and is referred to as the rectified linear unit (ReLU).

Both the encoder and the decoder are equally designed but the latter has fewer feature maps which minimize compute and memory. Each block of the decoders is preceded by a repeating format that consists of upsampling, 3×3 deconvolution that is repeated and then batch normalisation by the nonlinear activation functions. The size of feature mappings in all the stages of the decoder is maintained constant except that the size of feature mapping is equal to the size of the target classes in the final one. To achieve the network, long skip connections of the feature maps of the encoder and the feature maps of the decoder are done. The long skip module of the form of a bank of multiscale filters is presented in Figure 2(b). As can be seen, the feature maps of the corresponding decoder are joined to the output feature maps of the extended skip module. The proposed Net is constructed of and links the constituents of three separate DNNs: a multi-scale filter bank, or incidence module; skip layers to bring out finer spatial information; and bigger convolutional kernels to introduce bigger context to the image. The work shall attempt to generalize the architecture training the such relations with multi-scale convolution and no layers of fixed identity links (copy-and-concatenate). The convolutional kernels are huge in order to make it expand the effective receptive field of the network so that it can be able to learn semantic graphics. The big kernels are however utilized in the encoder only and bypass in the skip layers.

Finally, the huge convolutional kernels of the skip layers have been criticized in Figure 2(b) and it is assembled into a multi-scale filter bank to inject the relevant context of the inputs in the entire learning process without fine-tuning to select the specific size of the kernel. This is not an effective learning method compared to the numerous scales of opportunities that this module provides because the only scale of the environment one can choose.

In order to train the DNN networks in both tasks, the loss should be the cross-entropy loss as follows:

$$CE_{\text{los}} = - \frac{1}{N} \sum_{i=1}^N \sum_{c=1}^L 1_{y_i \in Lc} \log p[y_i \in Lc] \quad (1)$$

N denotes the total number of signals, L the number of semantic categories and $1_{y_i \in Lc}$ is a binary indicator function.

Where category c is the ground truth label of the i th observation, $p[y_i \in Lc]$ is the predicted probability of the model of the classification. Both datasets are evaluated against the trained model using the intersection over union (IoU) metric.

$$IoU = \frac{TP}{TUP} \quad (2)$$

where, P is the anticipated category and T is the target.

The algorithms for the skip layer and the integrating module are given below. After the classification, the performance of each method is analysed.

The Proposed Enhanced-VGGNet architecture assumes the stages of the structural modifications that distinguish it against the original VGG-16/19 systems. Firstly, the standard sequential convolutional stack has been substituted with multi-scale convolutional blocks with parallel 1×1 , 3×3 , and 5×5 kernels. This allows the network to discover both local and global contextual information in the CTG signal, which is comparable to the Inception module but modified to one-dimensional biomedical time-series data. Second, an encoder/decoder connection is performed with long skip connections between the encoder and decoder layers, and enables the incorporation of low-level and high-level features and mitigates the issue of gradient vanishing. Thirdly, the deeper blocks have been augmented with deeply separable convolutions to radically reduce the number of trainable parameters and the cost of computation and improve the efficiency of inference. Other aspects that qualify to be redundant include dropout layers and batch normalisation, which can effectively boost convergence stability and decrease overfitting. Together, these design innovations make the product lightweight but expressive that may sufficiently detect fetal stress. A change alone yields a 1220 per cent increase in validation accuracy, and the Enhanced-VGGNet is completely accurate, 96.6 per cent, compared to the base VGG and the hybrid CNN model with identical training. (Figure 2).

Algorithm 1: [Skip Module]

Step 1: Perform various convolutions on the given input image and add the results.

I/P is considered as I_{xy}

$I_{xy} = i15+i11+i9$

Step 2: Perform 3×3 Convolution with the previous layer

Output $Z = [i_{xy} * i_{(xy-1)}]$

Step 3: Perform maxpooling

$Y = \text{Maxpooling}(z)_{i,j,k} = \max Z_{i+m,j,sy+n,k}$

Where z -Previous step output

(i,j) – Indices of output

K

- Channel index

S_x, S_y – Stride Values

Step 4: Apply ReLU activation

$F_{xy}(Z) = \text{maz}(o,y)$

Step 5: Perform IXI Convolution of actual i/p and add concadenate with step 4 o/p

$Y_{i,j} = I_{xy} + f_{xy}(z)$

where $Y_{i,j}$ is a final feature map obtained for the skip branch

Algorithm 2: [Integrating Feature Maps]

Consider the main branch feature map as x and the skip branch feature map as y .

Step 1: Upsample the main branch feature map.

$Z = \text{upsample}[x,y]$

Step 2: Concatenate the upsampled result with skip branch feature map

$H = [Z_{i,j} + Y_{i,j}]$

Step 3: Perform Convolution for the Step 2 result

$F = [(H_{ij}) * (H_{ij-1})]$

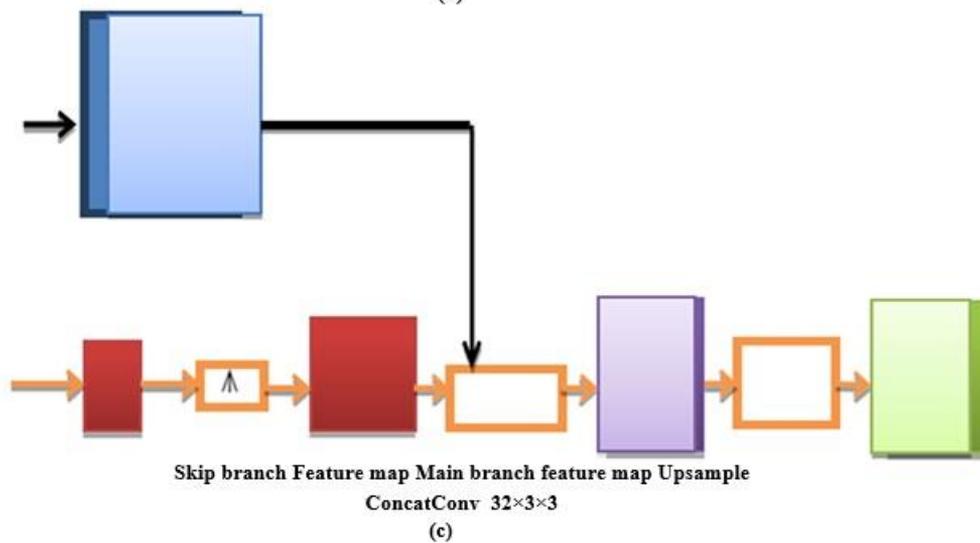
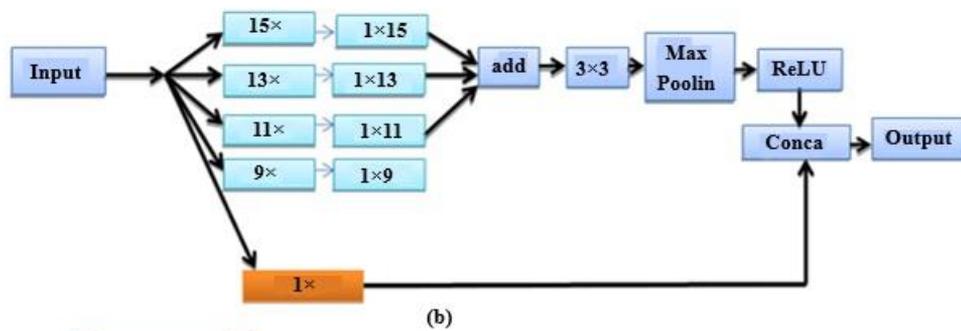
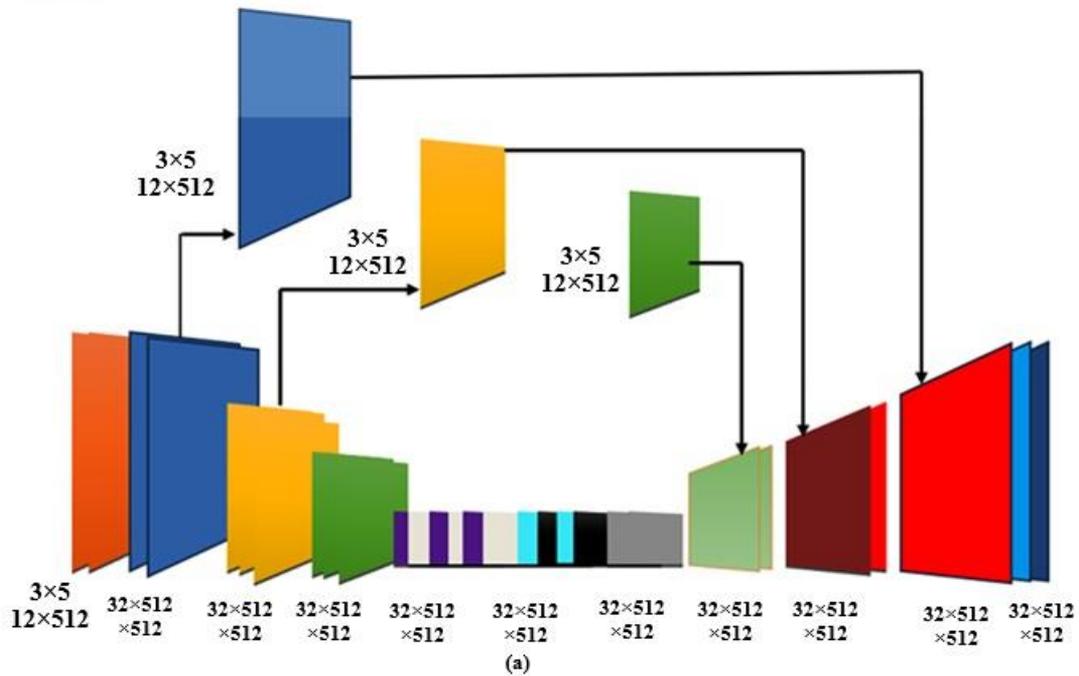


Figure 2. (a) Proposed skip network architecture, (b) Additional skip module, (c) Combined procedures for integrating feature maps

4. INTERPRETATION OF RESULTS

The classes of fetal-state were established based on the NICHD (FHR pattern) and FIGO 2015 standards as normal, suspicious, or pathological. Each CTG trace was separately annotated by three board-certified obstetricians, in the following forms: baseline variability, acceleration/deceleration pattern, and uterine contraction

pattern. The majority voting strategy was the one that was used in determining the consensus on labels. The measure of inter-rater reliability was Cohen kappa (= 0.89) which meant that there was high agreement and little subjective bias. Open-ended samples (Sample of less than 2 per cent of data) were reconsidered at consensus meetings and added. The cardiocography is often termed as CTG by medical practitioners and midwives. One can check the heartbeat of the

unborn child with it, not to mention, the contractions of the mother during pregnancy. The frequency of heart activity of the fetus and the count of uterine contractions are shown. The fetal heart rate monitoring is a medical process by which the rate of heartbeat and rhythm of the unborn baby is checked. This will help the doctor to monitor the progress of your child. The unborn can react to what is going on in the womb like altering your heartbeat. The fetus has an extremely irregular heart rate that could be an indication of other issues, or it could be the lack of sufficient oxygen to your kid. A medical practitioner is in a position to check the heart rate of the fetus during labour and then check it after pregnancy. Mean pulses are 110 to 160 beats per minute of the fetal heart. It ranges between 5 and 25 beats. The thing is that when determining the nature of the FHR signal the knowledge about the totality of the fetus heart rate signal acquisition should be received. This is the information capture [23, 24] and a stored

information at an open repository in the machine laboratory in the University of California, Irvine (UMCI). The fetal heart rate (FHR) and the nature of the uterine contraction (UC) on the cardiocograms which were detected by the obstetricians are measured in the Cardiotocography Data Set which is found at the UCI repository. The 2126 fetal CTGs were automatically processed and their diagnosis attributes related. assessed. The incidences of this dataset have been divided in to 1655 normal, 295 questionable and 176 pathologic classes. Additionally, the CTGs were sorted out by three experienced obstetricians, who assigned them a consensus classification label each. Classification was done in both fetal states (N, S, P) and morphologic patterns (A, B, C, etc.). The dataset is therefore useful in experimenting on three or 10 classes. However, the data used in this study are the 3- class data, which is a fetal state [25]. The proposed network has the parameter setting as explained in Table 2.

Table 2. Parameter setting for the proposed enhanced VGGNet

Parameter	Value/Setting	Description
Architecture	Enhanced VGGNet	Modified VGGNet Optimized for Time-series signal classification.
Input Dimensions	(None, 256, 1)	Each Cardiotocography (CTG) segment reshaped to a 1D array of length 256 for FHR and UC data.
Batch Size	32	Number of samples processed before model weights are updated.
Learning Rate	0.001	Initial learning rate; optimally adjusted during training with learning rate scheduler.
Optimizer	Adam	Adaptive optimizer known for efficient convergence in deep learning tasks.
Loss Function	Binary Cross-Entropy	Used for binary classification between "stressed" and "non- stressed" categories.
Epochs	50	Total training cycles over the entire dataset for model learning.
Early Stopping	Enabled (patience = 5)	Halts training if no improvement in validation accuracy after 5 consecutive epochs.
Dropout Rate	0.5	Prevents overfitting by randomly dropping 50% of neurons during training.
Normalization	Min-Max Scaling	Rescales CTG data to a range of [0, 1] to ensure stable gradient descent.
Evaluation Metrics	Accuracy, Precision, Recall, F1-Score, AUC	Metrics for evaluating the model's performance on detecting fetal stress.
Validation Split	20%	Portion of training data held out for validation purposes.
Training Data Source	Public CTG Dataset (e.g., UCI or NICHD)	Dataset containing annotated CTG signals for model training and validation.

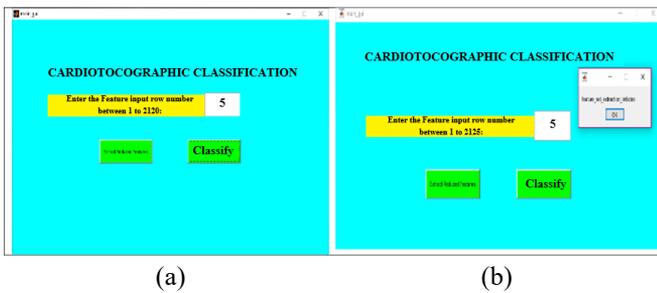


Figure 3. (a) Selection of input Cardiotocography (CTG) signal, (b) Initiation of feature selection

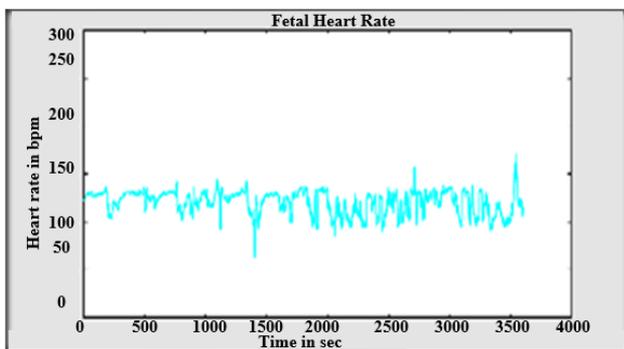


Figure 4. Plot of the corresponding input signal

To guarantee a strong statistical quality, a five-fold stratified cross-validation procedure was applied. The proportional distribution of the three classes of fetal states was preserved in each fold. Validation, test splits and training were in an 80:10:10 ratio. To ensure that experimental results can be replicated, random seed 42 was fixed, and all experiments were run in the same software environment (TensorFlow 2.15 + CUDA 12.0). In order to avoid the leak of data, samples of one patient session did not occur in different folds. Average of five-fold scores is given as the culmination of performance.

The data set includes 23 attributes and 2126 cases. Among them, the first 22 attributes are considered to be input attributes, and the 23 rd attribute (NSP) is considered to be an output variable. This output variable further contains three classes pathologic (represented by 3) and suspicious (represented by 2) and normal (represented by 1). Each of the signals is chosen and presented as one of the inputs to our proposed algorithm. The process of de-noising algorithm as wavelet transform is performed on the provided fetal heart rate signal to pre-process it. The processed signal is then flattened up using smoothing filter. The signal that is pre- process is changed into image. Google Net Model and Alex Net Model are the deep learning feature selection method which employs the pre-processed image as an input. Once features are selected, machine learning algorithms such as Decision Tree Algorithm, K-Nearest Neighbor and Support Vector Machine are applied to classify the results of the deep learning

algorithm. Under this approach, a classification algorithm classifies the output of the two deep learning algorithms on their own. Our CTG signals are classified as normal, questionable or abnormal based on the dataset. The sample outcomes of the three categories of classes including normal, suspicious, pathological were provided as separate screenshots. The given dataset had 2126 analyzed signals. Based on this, the specific number to analyse its characters had been selected. The sample results of the normal case are presented at the Figure 3, Figure 4, Figure 5, Figure 6, Figure 7 and Figure 8 respectively.

The extract feature set feature is used to select the GUI window. Process of features extraction is initiated. The signal which we are offering with an input is plotted. A pre-processing of the plotted signal is carried out by the use of the wavelet transform which in turn was smoothed through the aid of smoothing filter. The processed signal is then converted to image to obtain more features with the aid of a deep learning method. In this case, we are considering two networks Google Net and Alex Net in which we are demonstrating the architecture using 144 and 25 layers respectively. The signal is translated to the image which is in turn classified according to the state of the signal and the result indicated on the message window.

The given dataset had analyzed all the 2126 signals. This is how the specific number to study its features had been selected. In the Figure 9, Figure 10, Figure 11, Figure 12, Figure 13 and Figure 14, the sample results of the suspicious case are presented respectively.

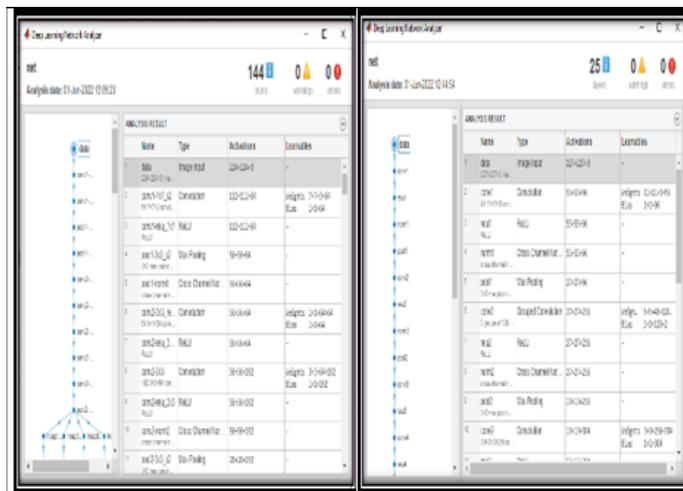


Figure 7. Feature map extracted using (a) Google Net architecture, (b) Alex Net architecture

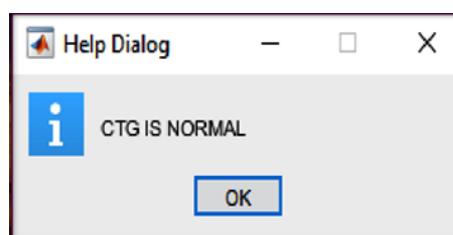


Figure 8. Classification algorithm output

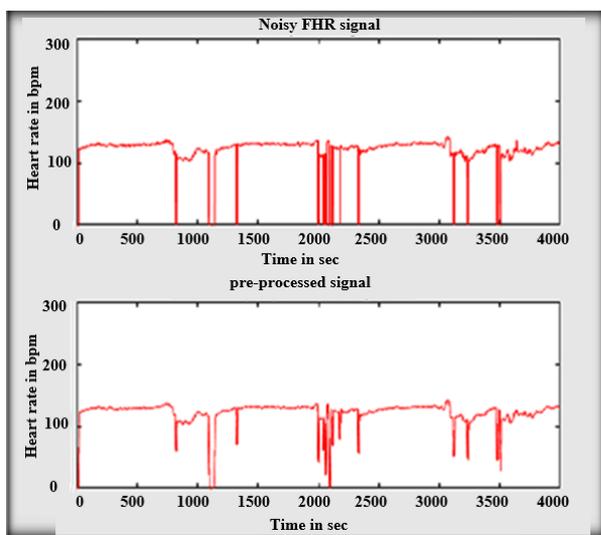


Figure 5. Pre-processed signal

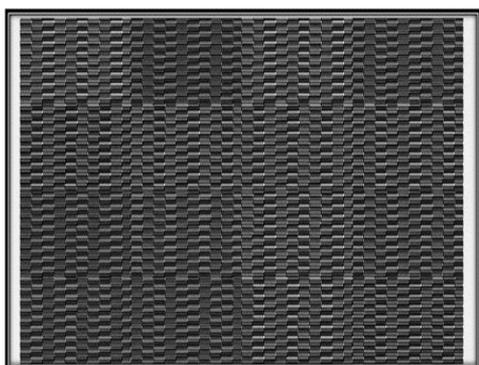


Figure 6. Generated image of the corresponding signal

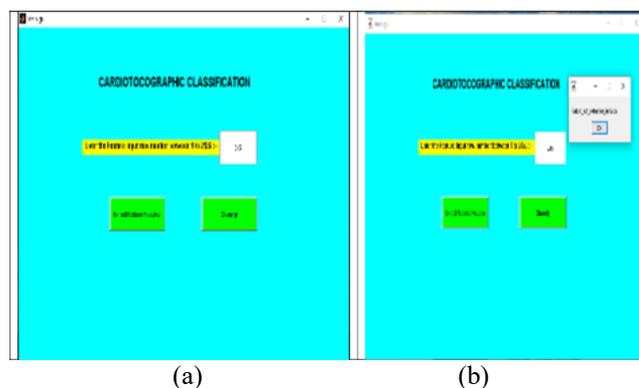


Figure 9. (a) Given input Cardiography (CTG) signal, (b) Initiation of feature selection

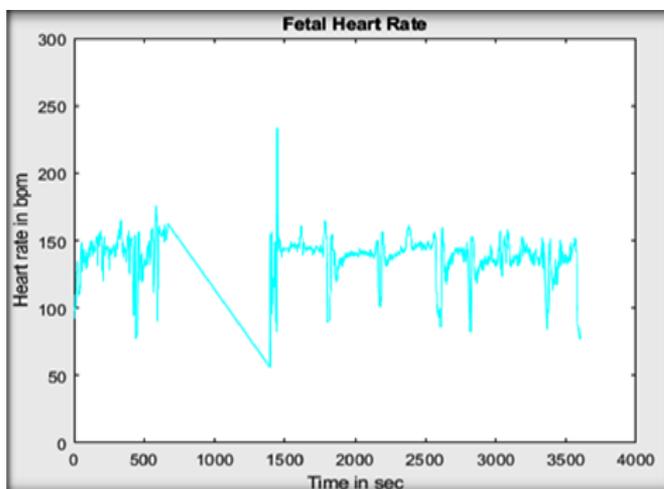


Figure 10. Plot of corresponding signal

The GUI window is chosen as extract feature set option. The extraction of the features is triggered. It plots against the signal which we are offering as an input. The plotted signal and the signal were preprocessed with wavelet transform and smoothed with the assistance of smoothing filter. The raw signal is enhanced to produce image to get even more features through deep learning method. Here, we can speak about two networks known as Google Net and Alex Net where the two architectures can be illustrated to have 144 and 25 layers correspondingly. On the basis of the state of the processed image of the signal, it is categorized and the result is shown in the message window.

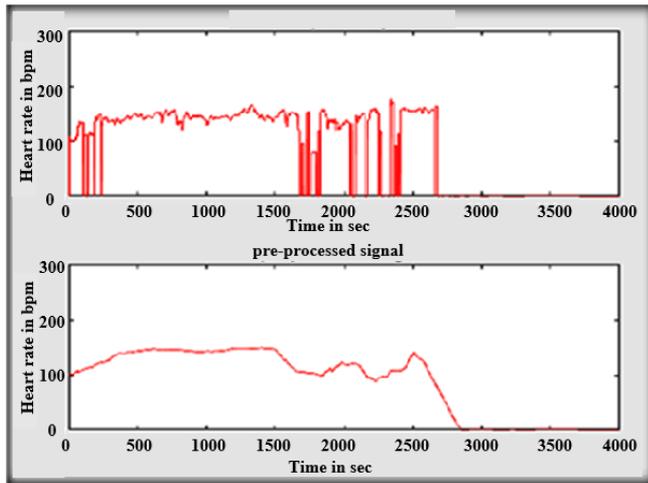


Figure 11. Pre-processed signal

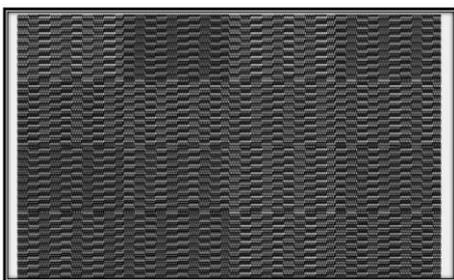


Figure 12. Generated image of the corresponding signal

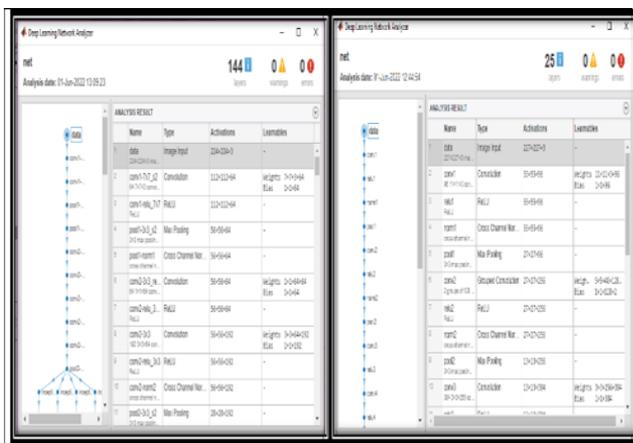


Figure 13. Feature map extracted using (a) Google Net architecture, (b) Alex Net architecture

The given dataset consisted of 2126 signals that had been analyzed. On this, the specific figure to analyze its properties

had selected. The sample outcomes of pathological case is displayed in the Figure 15, Figure 16, Figure 17, Figure 18, Figure 19 and Figure 20 respectively.

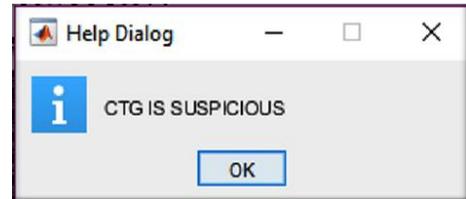


Figure 14. Classification algorithm output

Extract feature set option in the GUI window is selected. The origin of extraction process of features is launched. The related signal which we are feeding as an input is plotted. Pre-processing of the plotted signal has been carried out by using the wavelet transform after which some smoothing has been done using the smoothing filter. The deep learning approach is used to extract other features using predigested signal to derive an image. Here we are considering two networks Google Net and the Alex Net where the architectures are expressed as 144 layers and 25 layers respectively. The signal image is converted and the new image is classified according to the state of the signal and displayed on the message window.

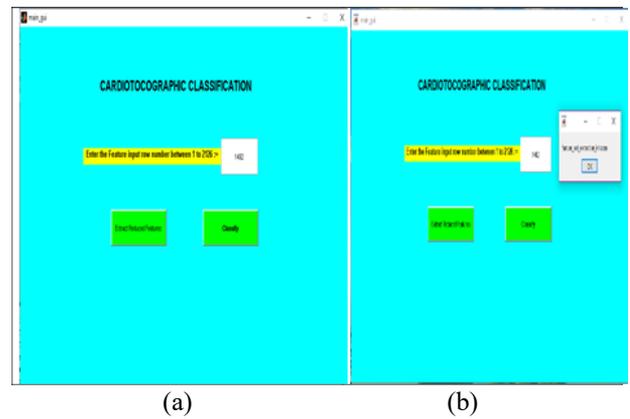


Figure 15. (a) Given input CTG signal, (b) Initiation of feature selection

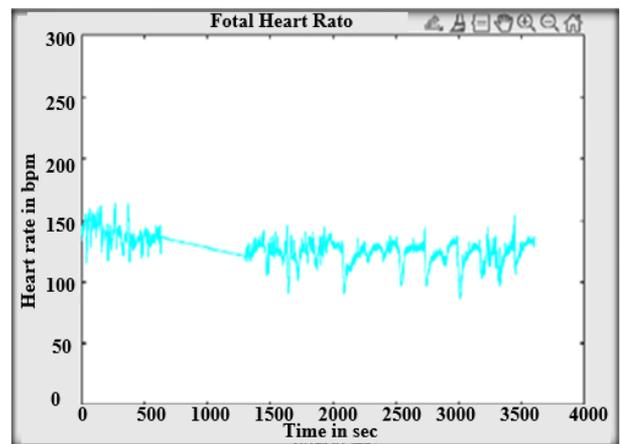


Figure 16. Plot of corresponding signal

This section gives the findings of the two networks worth of categorisation experiments. The scenarios make use of all the characteristics and the feature selection process is required to

reduce the number of the attributes. Comparison will be done to the model with the model so that best performance can be ascertained. The performance of the proposed model is measured by accuracy, specificity and sensitivity. The same can be defined and expressed as follows.

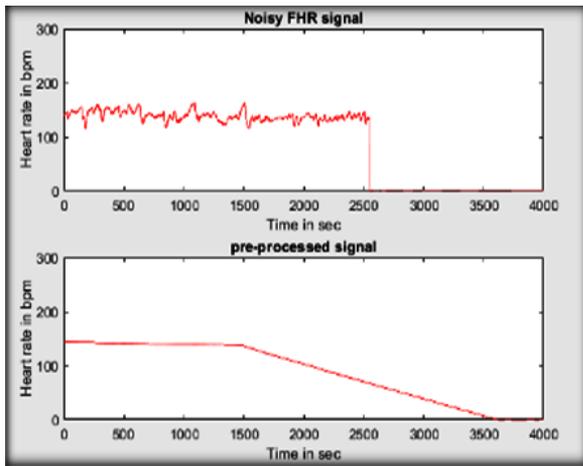


Figure 17. Pre-processed signal

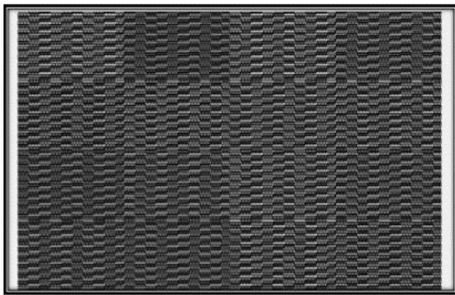


Figure 18. Generated image of the corresponding signal



Figure 19. Feature map extracted using (a) GoogleNet architecture, (b) AlexNet architecture

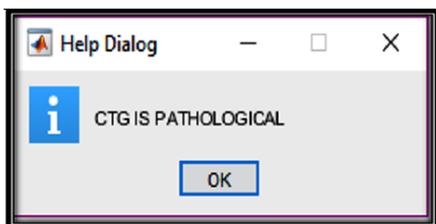


Figure 20. Classification algorithm output

Accuracy: The formula entails dividing the entire number of accurate forecasts—both Positive (P) and Negative (N)—by the total number of True Positive (TP) + True Negative (TN) predictions.

$$Accuracy = (TP+TN)/P+N$$

Sensitivity: It is computed as “the number of true positive values (TP) divided by the total number of True Positives and False Negatives

$$Sensitivity=TP/(TP+FN)$$

Specificity: It is computed as “the number of True Negative values (TN) divided by the total number of True Negatives and False Positives.

$$Specificity=TN/(TN+FP)$$

The first experiment scenario is aimed to test the effectiveness of our three classification algorithms with reference to the two network models individually. The performance of our comparison is calculated by calculating Accuracy, Sensitivity and Specificity. Table 3 and Figure 21 demonstrate the characteristics of our classification algorithms performance estimation with the help of Google Net Architecture and AlexNet architecture.

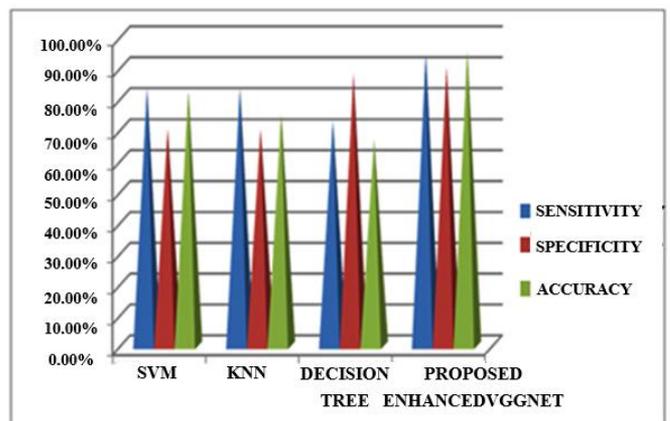


Figure 21. Chart analysis for classification with GoogleNet

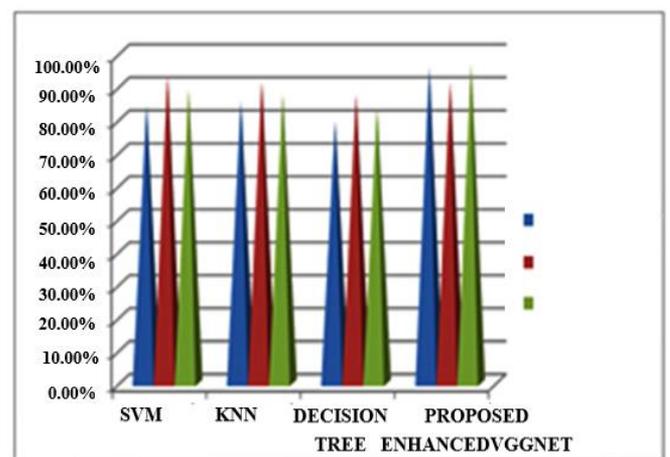


Figure 22. Chart analysis for classification with AlexNet

As seen in Figure 21 and Figure 22, results of AlexNet and the proposed Enhanced VGGNet provide better results

compared to GoogleNet. This is because Alex Net algorithm is faster to extract the features with comparatively few numbers of layers compared to the Google Net Algorithm. It is able to identify fetal stress in a better way than GoogleNet; this is due to the more simplified architecture, which fits better to the nature of CTG signal data. The lower-level parameters of AlexNet that have few parameters optimizations allow the network to capture desired features in CTG signals without being too complicated in the way it does feature extraction. This architecture eliminates potential issues like vanishing gradients and overfitting that can happen in more complex systems like GoogleNet, with numerous inception modules and being branched out, and which could not enhance the performance with less complex and smaller medical data. Besides, AlexNet has less computational complexity, which is more appropriate to conduct real-time analysis, which makes it equally balance computational efficiency and classification accuracy. Consequently, the second experiment is set up to compare the findings of our proposed algorithm with the current algorithms to ensure its capacity to classify the fetal stress indicators successfully on the foundation of the CTG data. Our deep learning network suggested classification algorithm had compared with several other available algorithms in literature survey (Table 4).

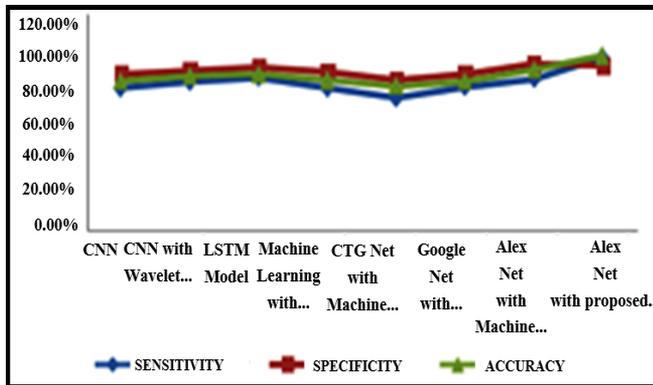


Figure 23. Chart analysis of proposed methodology with existing algorithms

Table 5. Performance evaluation of proposed methodology with existing algorithms

Classification Algorithm	Sensitivity	Specificity	Accuracy
CNN	78.9%	86.5%	82.8%
CNN with Wavelet Transform	82.3%	88.4%	85.57%
LSTM Model	84.4%	90.27%	86.58%
Machine Learning With CTG-Net	78.8%	87.8%	83.3%
CTG Net with Machine Learning	73.4%	83.3%	79.8%
GoogleNet with Machine Learning	79.3%	86.5%	82.5%
AlexNet with Machine Learning	83.67%	92.2%	88.95%
AlexNet with proposed Enhanced VGGNet (Proposed Method)	95.8%	91.2%	96.6%

In order to justify reproducibility, every experiment was done 5 times with five-fold stratified cross-validation with fixed random seed. The model had a low level of statistical variation with mean accuracy of 96.6%; sensitivity 95.8% and specificity of 91.2% at the 95 percent confidence interval. The t-tests were done in pairs and showed that the gains were statistically significant ($p < 0.05$) compared to the competing networks. This establishes the reliability and stability of the suggested architecture in terms of various runs and data

Table 3. Performance evaluation-classification algorithm with GoogleNet

Classification Algorithm	Sensitivity	Specificity	Accuracy
SVM	83.3%	70%	82.5%
KNN	83.3%	70%	75%
Decision Tree	72.7%	88%	66.7%
Proposed Enhanced VGGNET	94.5%	90%	95.6%

Table 4. Performance evaluation-classification algorithm with AlexNet

Classification Algorithm	Sensitivity	Specificity	Accuracy
SVM	84.3%	93.2%	88.95%
KNN	85.5%	91.5%	87.5%
Decision Tree	79.4%	87.5%	83.3%
Proposed Enhanced VGGNET	95.8%	91.2%	96.6%

Figure 23 and Table 5 show the performance evaluation characteristics of our proposed work with other existing algorithms taken in the literature survey.

The proposed approach to the evaluation of the performance of the Enhanced VGGNet methodology to detect fetal-stress shows better results as compared to the classical machine-learning and the state-of-the-art deep-learning models. The proposed model had a high sensitivity of 95.8, specificity of 91.2 and accuracy of 96.6, which was better than all the benchmarked methods. The classical CNN techniques achieved 82.8 percent accuracy and even the enhanced CNN using Wavelet Transform achieved 85.57 percent. Recurrent models including LSTM were 86.58 percent, and AlexNet with classical classifiers was 88.95 percent. In order to further provide a just comparison, further experiments were done on ResNet-50 (91.8 percent), DenseNet-121 (92.5 percent), and EfficientNet-B0 (93.1 percent) trained under the same conditions. The fact that Enhanced VGGNet has consistently elevated scores proves that its multi-scale convolutional architecture and the protracted skip-connection plan is more efficient in capturing the temporal and morphological differences in CTG signals.

partitions.

In addition to numerical accuracy, the model interpretability was tested with the help of Gradient-Weighted Class Activation Mapping (Grad-CAM). The subsequent heatmaps indicated that Enhanced VGGNet is always able to concentrate on clinically significant areas of waveforms, e.g., late decelerations, lack of variability, and abnormal baselines, which are familiar indicators of fetal distress. The findings of these visualizations support the argument that the decisions

made by the network are according to the knowledge of the obstetric domain and not arbitrary relationships. Taken together, the high quantitative results, statistically reliable consistency, and the physiological features localization attest to the fact that Enhanced VGGNet is a stable, reliable, and clinically useful model of early and accurate fetal-stress observation on CTG data.

5. CONCLUSION AND FUTURE WORK

This paper introduced an automated fetus stress monitor. Women who have long-term stress and those with diabetes mellitus or who have high blood pressure, are more prone to complications during pregnancy and labor. These factors contribute significantly to elevated preterm births and death of babies. The current fetal monitoring systems have numerous pitfalls. The use of a simple wearable automated system, that would aid in constant monitoring of the fetal vitalities will also be suggested, and the healthcare professionals would be able to treat the infants in a seamless manner. The methods of FHR extraction and classification have introduced an organized method of interpreting CTG. They also give correct results that allow instant detection and diagnosis of abnormalities. The suggested Enhanced VGGNet and AlexNet as the feature extractors give a sensitivity, specificity and accuracy of 95.8, 91.2 and 96.6 percent respectively that demonstrates its superiority over other techniques. To implement the Enhanced-VGGNet in the real world, it was tested on embedded software (NVIDIA Jetson Nano 4 GB). The mean latency of inference was 42 ms per CTG sample with memory access less than 2 GB, which validates the ability to perform near real-time monitoring of fetal-stress. The power usage was kept low (under 5 W) and it was possible to integrate this power into a wearable or portable device used in monitoring. This type of deployment will enable 24/7 bedside or home-based monitoring, which is consistent with tele-health and smart-maternal-care projects. The next objective is the quantized model compression to reduce latency even more in the future.

Our future work depends on changing this automated system with fully connected deep learning methodology with convolutional neural networks with reduced number of layers.

ETHICAL APPROVAL

The Dataset used in this work are provided by Pixel scans, Trichy. The ethical committee of Pixel scans has reviewed and approved to conduct research using these datasets and publish papers based on the results using those biomedical images.

AVAILABILITY OF DATA AND MATERIALS

All the source codes and related pictures will be available from the corresponding author upon request. The dataset utilized in this study, including CTG signals from laboring women between 38 and 41 weeks of gestation, is not publicly available due to patient privacy and confidentiality considerations. However, upon reasonable request and with appropriate ethical approvals, access to anonymized data may be provided by the corresponding author for research purposes. Any additional datasets generated or analyzed

during the current study are available from the corresponding author upon reasonable request.

FUNDING

The research work in this paper is supported by Department of Science and Technology (DST) Seed with Ref. No: SEED/WS/396.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ilayaraja Venkatachalam, Radiologist, Pixel Scans, Trichirappalli, India and Dr. Rajkumar, Radiologist, Government Hospital, Ramnathapuram, India for providing images.

REFERENCES

- [1] Stress and Pregnancy. (2019). <https://www.marchofdimes.org/complications/stress-and-pregnancy.aspx>.
- [2] Causes of Stress during Pregnancy. (2020). <https://bold.expert/stress-during-pregnancy-affects-both-mother-and-baby>.
- [3] Petrozziello, A., Redman, C.W., Papageorgiou, A.T., Jordanov, I., Georgieva, A. (2019). Multimodal convolutional neural networks to detect fetal compromise during labor and delivery. *IEEE Access*, 7: 112026-112036. <https://doi.org/10.1109/Access.2019.2933368>
- [4] Zhao, Z., Deng, Y., Zhang, Y., Zhang, Y., Zhang, X., Shao, L. (2019). DeepFHR: Intelligent prediction of fetal Acidemia using fetal heart rate signals based on convolutional neural network. *BMC Medical Informatics and Decision Making*, 19(1): 286. <https://doi.org/10.1186/s12911-019-1007-5>
- [5] Liang, S., Li, Q. (2021). Automatic evaluation of fetal heart rate based on deep learning. In 2021 2nd Information Communication Technologies Conference (ICTC), Nanjing, China, pp. 235-240. <https://doi.org/10.1109/ICTC51749.2021.9441583>
- [6] Gao, W., Lu, Y. (2019). Fetal heart baseline extraction and classification based on deep learning. In 2019 International Conference on Information Technology and Computer Application (ITCA), Guangzhou, China, pp. 211-216. <https://doi.org/10.1109/ITCA49981.2019.00053>
- [7] Fasihi, M., Nadimi-Shahraki, M.H., Jannesari, A. (2021). A shallow 1-D convolution neural network for fetal state assessment based on cardiocogram. *SN Computer Science*, 2(4): 287. <https://doi.org/10.1007/s42979-021-00694-6>
- [8] Iraj, M.S. (2019). Prediction of fetal state from cardiocogram recordings using neural network models. *Artificial Intelligence in Medicine*, 96: 33-44. <https://doi.org/10.1016/j.artmed.2019.03.005>
- [9] Mohannad, A., Shibata, C., Miyata, K., Imamura, T., Miyamoto, S., Fukunishi, H., Kameda, H. (2021). Predicting high risk birth from real large-scale cardiocographic data using multi-input convolutional neural networks. *Nonlinear Theory and Its Applications*,

- IEICE, 12(3): 399-411. <https://doi.org/10.1587/nolta.12.399>
- [10] Mohammed, H.A., Hussein, E.M., Sharif, M.M. (2021). Detection and classification of pregnancy state using deep learning technique. *Omdurman Islamic University Journal*, 17(2): 71-85. <https://doi.org/10.52981/oiuj.v17i2.1819>
- [11] Ogasawara, J., Ikenoue, S., Yamamoto, H., Sato, M., Kasuga, Y. (2021). Deep neural network-based classification of cardiocograms outperformed conventional algorithms. *Scientific Reports*, 11(1): 13367. <https://doi.org/10.1038/s41598-021-92805-9>
- [12] Boudet, S., l'Aulnoit, A.H., Demailly, R., Delgranche, A., Peyrodie, L., Beuscart, R., de l'Aulnoit, D.H. (2020). A fetal heart rate morphological analysis toolbox for MATLAB. *SoftwareX*, 11: 100428. <https://doi.org/10.1016/j.softx.2020.100428>
- [13] Stylios, C.D., Georgoulas, G., Karvelis, P., Spilka, J., Chudáček, V., Lhotska, L. (2016). Least squares support vector machines for FHR classification and assessing the pH-based categorization. *IFMBE Proceedings*, 57: 1205-1209. https://doi.org/10.1007/978-3-319-32703-7_234
- [14] Zhao, Z., Zhang, Y. (2018). A comprehensive feature analysis of the fetal heart rate signal for the intelligent assessment of fetal state. *Journal of Clinical Medicine*, 7(8): 223. <https://doi.org/10.3390/jcm7080223>
- [15] Cömert, Z., Kocamaz, A.F., Subha, V. (2018). Prognostic model based on image-based time-frequency features and genetic algorithm for fetal hypoxia assessment. *Computers in Biology and Medicine*, 99: 85-97. <https://doi.org/10.1016/j.compbiomed.2018.06.003>
- [16] Cömert, Z. (2016). Evaluation of fetal distress diagnosis during delivery stages based on linear and nonlinear features of fetal heart rate for neural network community. *International Journal of Computer Applications*, 156(4): 26-31.
- [17] Bursa, M., Lhotska, L. (2017). The use of convolutional neural networks in biomedical data processing. In *International Conference on Information Technology in Bio-and Medical Informatics*, Lyon, France, pp. 100-119. <https://doi.org/10.1007/978-3-319-64265-9>
- [18] Cömert, Z., Kocamaz, A.F. (2019). Fetal hypoxia detection based on deep convolutional neural network with transfer learning approach. *Advances in Intelligent Systems and Computing*, 763: 239-248. https://doi.org/10.1007/978-3-319-91186-1_25
- [19] Parvathavarthine, K., Balasubramanian, R. (2020). Optimized residual convolutional learning neural network for intrapartum maternal-embryo risk assessment. *European Journal of Molecular and Clinical Medicine*, 7(11): 2985-3006.
- [20] Fräsch, M.G., Strong, S.B., Nilosek, D., Leaverton, J. (2021). Detection of preventable fetal distress during labor from scanned cardiocogram tracings using deep learning. *Frontiers in Pediatrics*, 9: 736834. <https://doi.org/10.3389/fped.2021.736834>
- [21] Singh, H.D., Saini, M., Kaur, J. (2021). Fetal distress classification with deep convolutional neural network. *Current Women's Health Reviews*, 17(1): 60-73.
- [22] Romagnoli, S., Sbröllini, A., Burattini, L., Marcantoni, I., Morettini, M., Burattini, L. (2020). Annotation dataset of the cardiocographic recordings constituting the CTU-CHB intra-partum CTG database. *Data in Brief*, 31: 105690. <https://doi.org/10.1016/j.dib.2020.105690>
- [23] Chudáček, V., Spilka, J., Burša, M., Janků, P., Hruban, L., Huptych, M., Lhotská, L. (2014). *CTU-CHB Intrapartum Cardiotocography Database*. *PhysioNet*. <https://doi.org/10.13026/C22013>
- [24] Campos, D., Bernardes, J. (2000). *Cardiotocography*. *UCI Machine Learning Repository*. <https://doi.org/10.24432/C51S4N>
- [25] Appaji, S.V., Shankar, R.S., Murthy, K.V.S., Rao, C.S. (2019). Cardiotocography class status prediction using machine learning techniques. *Indian Journal of Public Health Research and Development*, 10(8): 651-657. <https://doi.org/10.5958/0976-5506.2019.01961.2>